Computers & Education 162 (2021) 104080

Contents lists available at ScienceDirect

Computers & Education

journal homepage: http://www.elsevier.com/locate/compedu

Capturing schools' digital capacity: Psychometric analyses of the SELFIE self-reflection tool

Patrícia Costa [a,b], Jonatan Castaño-Muñoz [*,a], Panagiotis Kampylis [c,d]

[a] European Commission - Joint Research Centre, Italy
[b] CEMAPRE, Portugal
[c] European Commission - Joint Research Centre, Spain
[d] Greek National Documentation Centre, Greece

ARTICLE INFO

Keywords:
Data science applications in education
Information literacy
Teaching/learning strategies
Elementary education
Secondary education

ABSTRACT

Results from self-reflection tools for schools' digital capacity can lead to evidence-based decisions within the school community and/or the development of an action plan for a better integration of digital technologies. Thus, it is important that the information derived from self-reflection tools is complete, accurate, and relevant. However, usually self-reflection tools do not show evidence of the quality of the information provided. In this paper, we focus on SELFIE, a new, comprehensive, and customisable self-reflection tool for schools' digital capacity, and we analyse the quality of the information that it provides. In particular, we look at discrimination and difficulty item parameters (using item response theory), we analyse the reliability (using Cronbach's alpha and Omega) and the construct validity (using confirmatory factor analysis) of its core items. We find support for the tool quality and conclude that schools using SELFIE are provided with accurate information on their digital capacity. Additionally, we discuss ideas for further improving the tool and future research work. The innovative design of the SELFIE tool and the psychometric analyses of its core items are a novelty in the field of schools' digital capacity and can provide insights for the development of self-reflection tools for school communities.

1. Introduction

1.1. *Schools and digital technologies use*

Digital technologies are rapidly transforming almost every sphere of life. They are also transforming the way we teach and learn in formal, non-formal and informal education (e.g. Rogers, 2014; Skryabin, Zhang, Liu, & Zhang, 2015). The use of digital technologies in formal education comprises two complementary objectives: i) the pedagogical use of digital technologies to support and enhance teaching and learning; ii) the development of digital competence of students and teachers (e.g. Ottestad & Guðmundsdóttir, 2018). In this context, governments, schools and families all over the world considerably invest to integrate digital technologies as a central part of the teaching and learning process (e.g., Conrads, Rasmussen, Winters, Geniet, & Langer, 2017; Escueta, Quan, Nickow, & Oreopoulos, 2017).

Despite the pedagogical use of digital technologies to enhance teaching and learning, the benefits of this considerable investment

* Corresponding author.
E-mail address: jonatan.castano-munoz@ec.europa.eu (J. Castaño-Muñoz).

https://doi.org/10.1016/j.compedu.2020.104080
Received 10 January 2020; Received in revised form 30 June 2020; Accepted 15 November 2020
Available online 28 November 2020

(devices for educational use, internet connections, software, teacher training etc.) are not clear. In general, research has found *outcome gap* (Lim, Zhao, Tondeur, Chai, & Tsai, 2013) as the investment in education technology has not returned the same gains in terms of reduced costs and increased productivity as in other sectors. Studies in this area have struggled to find positive associations between the use of educational technology in schools and students' achievement (e.g. Mora, Escardíbul, & Di Pietro, 2018; Petko, Cantieni, & Prasse, 2017). The meta-analysis by Tamin and colleagues shows a small or moderate significant positive effect size favouring the use of digital technologies over more traditional instruction (Tamim, Bernard, Borokhovski, Abrami, & Schmid, 2011). Other studies suggest that the effectiveness of digital technologies at school depends on the pedagogical practice of teachers (Comi, Argentin, Gui, Origo, & Pagani, 2017; OECD, 2015). The results from an OECD study (2015), based on data from PISA 2009 and 2012, show that the impact of educational technology on student performance is mixed at best. On average, in the past 10 years there has been "no appreciable improvement in student achievement in reading, mathematics or science in the countries that have invested heavily in ICT for education" (OECD, 2015, p. 3). But it largely depends on how the technology is used.

On the other hand, research has also signalled a technology *usage gap* (Lim et al., 2013). School-related usage is much less intensive than outside school usage for other purposes and it can have several consequences. First, it can affect traditional learning outcomes. There is PISA based evidence suggesting that the use digital technologies (in the school and at home) can improve achievement of students with low socioeconomic background and/or low use of technology. However, students who are medium- or high-intensity users of technology typically would not gain from additional use (Rodrigues & Biagi, 2017). Furthermore, research evidence also emphasises the need to better use digital technologies in education and training settings to improve the development of digital competences for students and adults alike (European Commission, 2019c).

Digital competence has been listed as one of the eight key competences for lifelong learning since 2006 (European Parliament & the Council, 2006) and there is a clear need for its early development in educational systems. While, 90% of all jobs in Europe require digital competences (European Commission, 2019b), only 43% of the EU population between 16 and 74 years have acquired basic digital competences (European Commission, 2019b). For young people entering in the labour market the data improves a bit but it is still insufficient. Only around 59% of 16–24 years-old individuals have above basic level digital skills, and this figure declines to 50% at the age group of 25–34 years-old (European Commission, 2020). The results from the International Computer and Information Literacy Study (ICILS) also confirms the need for digital competence development. According to ICILS, in most countries the majority of grade 8 students scored in level 2 of the computer and information literacy scale, which can be defined as the threshold for underachievement in digital competence (Fraillon, Ainley, Schulz, Friedman, & Duckworth, 2019).

In this sense, it can be argued that education systems in general, and schools in particular, are on average not fully leveraging the potential of digital technologies to enhance teaching and learning and the development of digital competence of students and teachers. Many schools lack a school-based digital policy plan and/or a shared vision on the use and integration of digital technologies, which have been identified by the literature as key conditions for improving the digital capacity of schools (Vanderlinde & van Braak, 2010; Vanderlinde, van Braak, & Dexter, 2012). Research also shows that innovative and effective use of digital technologies is mainly bottom-up (e.g., Jeladze & Pata, 2018) and that the shared vision can also be established in this way (Vanderlinde et al., 2012) if the dialogue and conditions are adequate.

1.2. Schools digital capacity

In this context, the concept of digital capacity of educational institutions has become progressively more important (Ilomäki & Lakkal, 2018). This concept is interrelated to similar ones defined by literature. First, to the more general one, *organisational digital capability*, which is defined as "the extent to which the culture, policies and infrastructure of an organisation enable and support digital practices" (Killen, Bentham, & Knight, 2017). Frameworks for the digital capacity of educational organisations have been also developed (e.g. Jisc, n. d.) and recognise that although it goes beyond the capabilities of individuals, still "digitally capable staff and students are essential to organisational success". Secondly, to the more focused one, *schools e-capacity* defined by Vanderlinde and van Braak (2010, p. 543) as "the schools' ability to create and optimise sustainable school level and teacher level conditions to bring about effective ICT change". Finally, the term *digital maturity* has also been defined as the systematic approach towards using technology in school management and teaching and learning practices (Balaban, Redjep, & Calopa, 2018).

Building on these definitions, we define *school's digital capacity* as the extent to which culture, policies, infrastructure as well as digital competence of students and staff support the effective integration of technology in teaching and learning practices.

Research shows that for digital technologies to have a positive pedagogical effect several conditions must be met, not only at policy level but also at school level (Voogt, Knezek, Cox, Knezek, & Brummelhuis, 2011; Vanterlinde & van Braak, 2010). However, many schools do not yet have a clear strategy on how to effectively use digital technologies for teaching and learning (Jeladze & Pata, 2017). Even when schools have a digital strategy (European Commission, 2019c), they still need to regularly review their teaching, learning, and organisational practices to improve their digital capacity and their approach towards digital learning. To do so, schools need (a) a clear understanding of what it means to be digitally capable as an organisation[1] and (b) tools for self-assessing their current state of development and in formulating plans for improvement ((Kampylis, Devine, Punie, & Newman, 2016). The development of school

[1] See below, in sections 2.1 and 2.2.1, what does it mean for an educational organisation to be digitally capable.

self-reflection or self-evaluation tools[2] on this specific topic goes in this direction and has become a common and important practice to improve the digital capacity of schools and optimise the use of digital technologies in education.

1.3. Self-reflection tools for digital capacity development

According to Chapman & Sammons (2013), school self-evaluation can have various purposes (not mutually exclusive): preparation for inspection; raising standards; professional development; building school capacity to respond to and manage change. Antoniou, Myburgh-Louw, and Gronn (2016) found that the school self-evaluation exercises are more effective when they i) are based on the use of validated and theory-based tools; ii) consider all school stakeholders' perspectives; iii) facilitate evidence-based policies for school improvement on multiple levels.

Several tools, online and paper-based, aim to support schools to systematically self-reflect upon or self-evaluate their digital capacity or related concepts (e.g., Balaban et al., 2018; Tanhua-Piiroinen & Viteli, 2017). The analysis of nine tools used in Europe[3](-Kampylis, Devine, Punie, & Newman, 2016) reveals a number of key factors for improving schools' digital capacity that include leadership and governance practices; digital infrastructure and resources; teachers' role and the need for developing their digital competence; the need for integrating digital technologies across the curriculum. Also, in most cases these tools focus mainly on digital infrastructure and frequency of technology use. Moreover, the quality of their indicators is not validated using robust psychometric procedures. Consequently, it cannot be assessed whether they provide high quality information to the schools that use these tools. Also, in these nine tools, only school leaders or school leaders and teachers provide the input about the digital capacity of their schools without the involvement of students. Finally, the analysed tools do not allow customisation but follow a one-size-fits-all approach, providing the same set of questions to all schools regardless of their different needs and settings, which can hinder their relevance and usefulness.

1.4. Research questions

To address these limitations, the European Commission has developed, in collaboration with a panel of experts from across Europe, an online and free self-reflection tool for schools' digital capacity called SELFIE. The development of SELFIE was one of the 11 actions of the Digital Education Action Plan that the Commission adopted in January 2018 (European Commission, 2018). The tool was launched in October 2018 and is available at the time of writing in 32 languages (24 EU official languages as well as Albanian, Georgian, Icelandic, Macedonian, Montenegrin, Russian, Serbian, and Turkish). The tool comprises three questionnaires, which are filled out by school leaders, teachers, and students reflecting on how digital technologies are used in their school. The integrated results aim to provide the school community with a 360° view of where they stand in the use of digital technologies in relevant areas and indicators that measure the digital capacity of the school. Moreover, SELFIE is not a prescriptive tool and rather than focusing on benchmarking, it aims to promote an internal evidence-based debate on the use of digital technologies for teaching and learning in the school at hand. Ultimately, it is expected that this debate would lead to concrete actions and/or the development of an action plan oriented towards a more efficient integration of the technologies for teaching and learning taking into account the school's needs and context. Thus, it is of utmost importance to guarantee the scientific quality of the questionnaires so that the data provided by the SELFIE tool to schools is complete and accurate, capturing well their digital capacity.

To further investigate the quality of the information provided by SELFIE, this paper analyses the anonymised data collected during the first six months since its release, including the responses[4] from school leaders, teachers and students from Europe and beyond. More specifically, following the success factors identified by previous literature (Antoniou et al., 2016) and aiming to overcome the limitations of existing tools described above, we use psychometrics to answer the following research question and three sub-questions:

RQ does the SELFIE tool provide accurate information to school communities on the digital capacity of their schools?

- RQ1) Does SELFIE capture well the digital capacity of general schools allowing a distinction to be made between schools that are in different stages on the use of technology for teaching and learning?
- RQ2) Are the 36 core items[5] of SELFIE allocated to the right thematic area, derived from the theoretical framework that the tool is based upon, providing the school communities with a well-structured and comprehensive overview of the digital capacity of their schools?
- RQ3) Do SELFIE core items work well for different respondent groups and educational levels?

[2] Self-evaluation has been defined as a process of reflection on practice, made systematic and transparent, with the aim of improving pupil, professional and organisational learning (MacBeath, 2005). Therefore, in this paper, we consider that self-reflection using structured tools as a synonym of self-evaluation.

[3] The nine tools analysed were the following (the last two are paper-based, the others are online tools): Assessing the e-Maturity of your School (international), Digital schools of Distinction (Ireland), eLEMER (Hungary), Future Classroom Maturity Model (international), Microsoft Innovative Schools Toolkit & Self-Reflection Tool (international), Opeka (Finland), eLearning roadmap (Ireland), School mentor (Norway), NAACE self-review framework (UK).

[4] By registering in SELFIE, schools consent to aggregated data being used for public reports and presentations. SELFIE privacy policy and data protection is available at https://europa.eu/!MX94Ww.

[5] Items that are compulsory for all schools using SELFIE.

To address these questions we use psychometric methods that are in line with principles and practices to analyse the data from international large-scale student assessments, such as PISA, TIMSS or PIRLS (Martin, Mullis & Hooper, 2017; OECD, 2017).

The rest of the paper is structured as follows: first, we present the conceptual model on which SELFIE is based and its early developments until it became a fully-fledged self-reflection tool for schools' digital capacity. We then describe the tool and the type of data it collects and indicate the details of the methodology that we use for checking items' quality, reliability and construct validity of the instrument. Subsequently, we present the results of these analyses. Finally, we discuss the results and propose ideas for improving the tool and for future research.

2. Background

2.1. Theoretical framework - European Framework of digitally-competent educational organisations (DigCompOrg)

The SELFIE self-reflection tool for schools is the result of several years of work by the European Commission and a team of education experts from across Europe. The theoretical underpinning of SELFIE is the *European Framework for Digitally Competent Educational Organisations* (Kampylis, Punie, & Devine, 2015)(Kampylis, Punie, & Devine, 2015). The methodological approach for the development of DigCompOrg was for the most part qualitative: a review of academic and grey literature; an inventory of existing frameworks and self-assessment tools; an in-depth analysis of selected frameworks and tools; and a number of expert and stakeholder consultations. The framework consists of seven key areas: *Teaching and Learning Practices; Assessment Practices; Content and Curricula; Collaboration and Networking; Professional Development; Leadership and Governance Practices; Infrastructure and Equipment*. The seven key areas of the framework were further expanded into 15 sub-elements that were broken down into 74 descriptors, which are statements that encompass the key aspects of the digital capacity of an educational organisation.

2.2. Tool development

2.2.1. Inception phase

The design of SELFIE started from the areas, sub-elements and 74 descriptors of DigCompOrg framework that include both 'organisational responsibilities' (e.g., Infrastructure) and 'individual responsibilities' (e.g., Teaching and Learning Practices). Therefore, from the outset, the aim was to involve the whole school community (school leaders, teachers and students) in the self-reflection process on the digital capacity of their schools. This would allow for a holistic view of the school's digital capacity in terms of actors involved as well as by capturing both top-down leadership and strategies and bottom-up efforts initiated by staff.

For the initial development of the tool, firstly a meta-analysis of nine existing tools oriented to the use of digital technologies in schools was performed (Kampylis, Devine, Punie, & Newman, 2016). The analysis of the tools resulted in an extensive list of 117 potential items that were mapped against the areas, sub-elements and descriptors of the DigCompOrg framework.

In order to make the tool user-friendly, we faced the challenge of reducing the number of items by identifying those that better capture the digital capacity of schools . For this, we launched two different actions. Firstly, in January 2017, a comprehensive user consultation survey was launched to obtain information about the perceived relevance and quality of the items by each of the three respondent groups.[6] This information was used to reduce the number of items, reformulate those that were not clear, allocate items to specific respondent groups, and decide which ones should be core items (i.e., mandatory for all schools) and which should be optional (i.e. schools can decide whether or not include these in their self-reflection exercise). Secondly, a series of stakeholder consultations (including policymakers, researchers, school leaders, teachers and students) was performed from March to July 2017 to build consensus on the content and design of the items (core and optional) to be included in the SELFIE tool prototype as well as on the possibility to allow the addition of school-created items. These consultations derived information from the results of the user consultation survey.

2.2.2. Pilot phase

Based on the results of the survey and the stakeholder consultation, an early prototype of the SELFIE tool was developed comprising 58 items for school leaders, 45 for teachers and 17 for students (Fig. 1). This prototype was translated into 12 European languages and it was tested in a pilot phase in September–October 2017 (Castaño-Muñoz, Costa, Hippe, & Kampylis, 2018) by 67,714 school leaders, teachers and students from 650 schools in 14 countries.

The data collected during the pilot implementation of the tool allowed us to make an initial analysis of the quality of the items and the factorial structure of the core items from a psychometric perspective. These analyses were complemented by the thematic analysis (Braun & Clarke, 2006) of 12,599 comments collected through the open questions of the tool that were coded into 13,737 excerpts. Moreover, 14 country reports with insights from the implementation of the tool in each of the pilot countries accompanied by twenty-eight case studies in schools that tested SELFIE were conducted in November–December 2017. The results from the analyses of the qualitative and quantitative data were further discussed in a validation workshop with 50 education experts and practitioners who took part in the pilot implementation of the tool (see Fig. 2).

Based on the information from the activities described above, we made some improvements in the tool. First, we selected the

[6] Respondents comprised 224 school leaders, 1,154 teachers and 3,674 students from five European countries (Denmark, Estonia, Ireland, Italy and Spain).

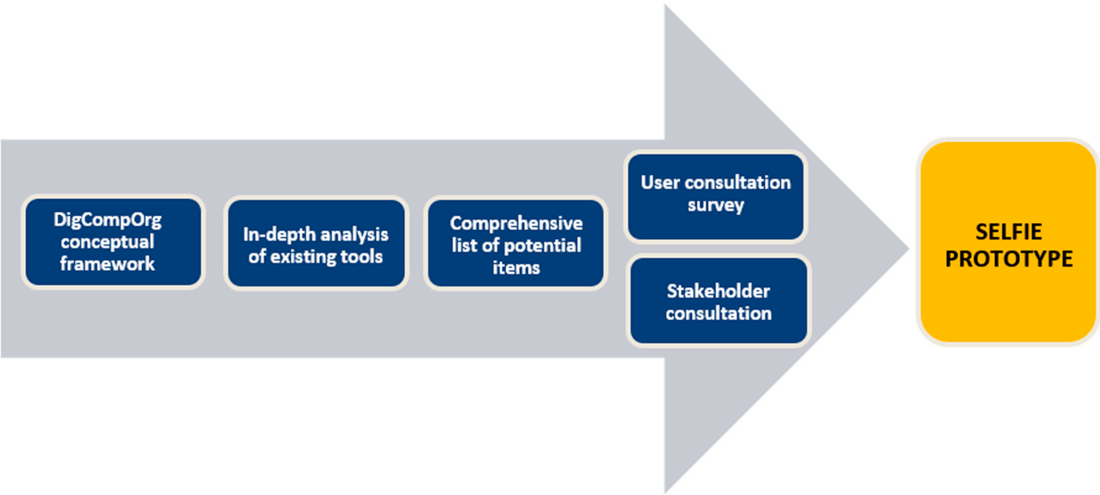

Fig. 1. Overview of the development of the SELFIE prototype.

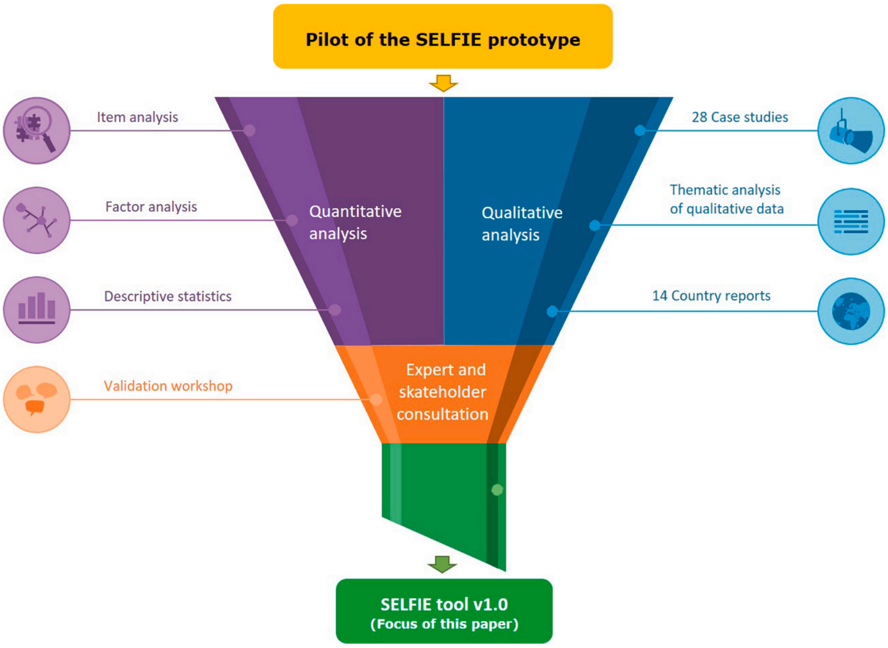

Fig. 2. Overview of the pilot phase.

existing items and added extra items identified as important by the users. Next, we refined the items in two ways: a) homogenising school leaders' (SL), teachers' (T) and students' (S) items formulation to guarantee comparability and b) simplifying items formulation to avoid comprehension difficulties, especially by primary students. Finally, we adjusted the structure of the questionnaire based on the results of the factor analysis. We created a new area referred to as *Student digital competence*, eliminated the *Content and curricula* area[7] and reformulated the items referred to as *Collaboration and networking* area due to validity problems. New items were included covering concepts more related to the original meaning of the *Collaboration and networking* area in the DigCompOrg framework. The items covering this area were provisionally in the area of *Leadership*. Only subsequent factor analysis could confirm or reject the existence of this area as an independent one. In addition, the *Teaching and learning area* was reformulated and divided into two areas: *Teaching and learning* - support and *Teaching and learning – pedagogy*. Fig. 3 summarises the changes made to the structure from the theoretical model to the SELFIE tool (Castaño-Muñoz, Costa, Hippe, & Kampylis, 2018).

[7] Items were allocated into *Teaching and learning* or into the new *Student digital competence* areas.

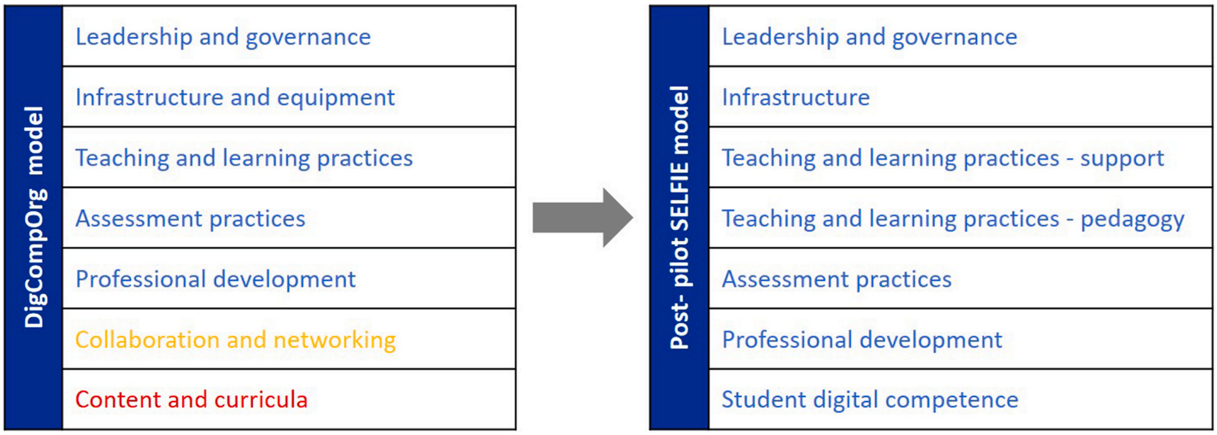

Fig. 3. From DigCompOrg model to SELFIE tool – Areas.

The final build and release of SELFIE v1.0, which is the focus of this analysis, took place on October 25, 2018 (European Commission, 2018). The following points describe the instrument and the analysis performed to check the quality of the tool.

3. Methodology

3.1. Data and instrument

SELFIE is an online[8] tool that can be used by any school free of charge. Slightly different versions of the SELFIE tool are available for *primary (ISCED 1)*, *lower-secondary (ISCED 2)*, *upper-secondary general (ISCED 3)*, *upper-secondary vocational (ISCED 3 – VET) and post-secondary non-tertiary education levels (ISCED 4 – PSNTE)*. Aggregated data from the respondents from all of these education levels is included in the dataset, which includes the replies by 301,935 participants (8,105 school leaders, 40,846 teachers and 252,984 students) that come from 3,110 schools in 33 countries.[9] The distribution of users by educational level is presented in Table 1. The highest number of users participating in SELFIE is in ISCED 2. All data comes from the school year 2018–2019 that was divided into three subsequent, independent sessions: from October 25, 2018 to December 31, 2018 (first wave); January 5, 2019 to April 14, 2019 (second wave); and April 18, 2019 to July 31, 2019 (third wave). In this paper, we focus on the data from the first and second wave[10].

There are the following types of items[11] (questions or statements) in SELFIE:

- A set of *core items* that are the same for every school and school level and are grouped into seven areas[12];
- A set of *optional items* that schools can choose to include or not in the questionnaires for their school leaders, teachers and students – they are grouped into the same seven areas with the core items;
- Up to ten items that schools can add to better suit their needs and special context – these items appear in a new area entitled *Questions added by your school*;
- Additional items about the use of digital technologies for teaching and learning inside and outside school;
- A few demographic questions.

The analysis presented in this paper only considers the core items answered by SELFIE participants (see Table 2).[13] The core items are composed of questions with five answer options: 1: Strongly disagree – In my experience, this is not true at all; 2: Disagree; 3: Slightly agree; 4: Agree; 5: Strongly agree – In my experience, this is very true. The response to these items is obligatory for all respondents. However, in all items respondents have the choice to opt out by selecting the "Not applicable" or "Prefer not to say" answer options.

SELFIE questionnaires can be filled out by three respondent groups (SL, T, S). In this paper, we analyse the items quality and

[8] https://ec.europa.eu/education/schools-go-digital_en.

[9] Austria, Belgium, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Estonia, France, Finland, Georgia, Greece, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Malta, Mexico, Moldova, Montenegro, Poland, Portugal, Romania, Russia, Serbia, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom.

[10] The analysis of the questionnaires for upper-secondary vocational schools are not included in this paper since these questionnaires include some additional core items that do not appear with the ones for the other ISCED levels

[11] All core and optional items have a help text that provides more information about the highest level 5 of the answer options.

[12] The areas are the following: *Leadership*; *Infrastructure and equipment*; *Continuing professional development* (CPD); T*eaching and learning - support; Teaching and learning – pedagogy; Assessment practices*; and *Student digital competence*.

[13] The optional items and the ones that schools add in their questionnaires are excluded from this analysis as they do not apply to all schools.

Table 1
Number of users by educational level.

		USER TYPE			Total
		School Leaders	Teachers	Students	
EDUCATIONAL LEVEL	**ISCED 1**	2621	13,839	64,314	80,774
	ISCED 2	2873	14,570	99,771	117,214
	ISCED 3	1372	7098	55,208	63,678
	ISCED 3 – VET	934	4307	27,248	32,489
	ISCED 4 – PSNTE	305	1032	6443	7780
TOTAL		8105	40,846	252,984	301,935

Table 2
Number of items per user group, area and ISCED level.

ISCED level	Area	CORE ITEMS		
		User group		
		School Leaders	Teachers	Students
Primary	*Leadership*	5	5	1
	Infrastructure and equipment	6	6	3
	Continuing professional development	3	3	0
	Teaching and learning – support	5	5	0
	Teaching and learning – pedagogy	5	5	3
	Assessment practices	5	5	0
	Student digital competence	7	7	4
	TOTAL	36	36	11
Lower-secondary and upper-secondary general	*Leadership*	5	5	1
	Infrastructure and equipment	6	6	3
	Continuing professional development	3	3	0
	Teaching and learning – support	5	5	1
	Teaching and learning – pedagogy	5	5	5
	Assessment practices	5	5	3
	Student digital competence	7	7	7
	TOTAL	36	36	20

reliability of the three respondent groups. The analysis of the construct validity is restricted to replies provided only by school leaders and teachers. The reason for this is that students' questionnaires have a much lower number of questions, in some areas they have very few or no questions at all (e.g., on *Leadership* or on *Continuing professional development,* respectively) and, in this sense, the validation of areas cannot be done consistently with the other two respondent groups. Furthermore, for students in primary schools, who participate in SELFIE when they are 9 years old or above, the formulation of the items has been further simplified and is quite different from the ones for teachers and school leaders.

The SELFIE core items were designed to allow a distinction to be made among schools with different levels of digital capacity, providing useful information for all types of schools. For the most advanced schools, in terms of digital capacity, the SELFIE tool allows to select optional items and to customise the tool by incorporating schools' own questions, created according to their own needs.

3.2. Statistical methods

The statistical analysis for this paper with reference to the pooled data of wave 1 and 2 was carried out in terms of items quality, reliability of the tool and construct validity of the areas from the DigCompOrg framework. Analyses by ISCED[14] level (1, 2 and 3) were performed for SL and T. For S, the item parameters and reliability are presented for ISCED levels 1,[15] 2 and 3 levels, as the number of items varies across educational levels.

3.2.1. Items quality

The psychometric analyses of the properties and quality of SELFIE core items were performed using statistical procedures and refer to SL, T and S questionnaires. In order to have robust evidence of the capacity of SELFIE items to distinguish among schools with

[14] In SELFIE, schools are categorised according to the 2011 International Standard Classification of Education – ISCED (http://uis.unesco.org/en/topic/international-standard-classification-education-isced).

[15] Only students that are 9 or older participate in the self-reflection process on their schools´ digital capacity through SELFIE tool.

different digitalisation levels we combined the use of two approaches: classical analysis of the items and Item Response Theory (IRT).[16] The findings from both approaches[17] were in line and, in this sense, we opted to present only the items properties based on IRT. We look at the items parameters in terms of discrimination (α) and general location or difficulty (β) using the Generalised Partial Credit Item Response Model – GPCIRM (Muraki, 1993, 1997). This model is theoretically adequate for polytomous items and the item parameters are considered factors that influence the choice of answers and characterise respondents.[18] The item parameters estimates were obtained through the application of the Marginal Maximum Likelihood (MML) estimation procedure using the EM algorithm (Baker & Kim, 2004). Estimation using this procedure was performed using the Parscale software (Muraki & Bock, 2002). This procedure has the advantage of producing estimates that are independent of the tested population. An instrument should be composed of discriminative items and have items with different levels of difficulty.

3.2.2. *Reliability*

In order to empirically verify if SELFIE is a reliable tool, we used Cronbach's alpha, which is a measure of internal consistency. Cronbach's alpha is a function of the average intercorrelations of items and the number of items in the scale. Cronbach's alpha coefficients range from 0.00 to 1.00, with higher coefficients indicating higher levels of reliability and therefore that the items are measuring the digital capacity of schools. Values of Cronbach's alpha higher than 0.9 indicate an excellent internal consistency and values varying between 0.8 and 0.9 suggest good internal consistency. SPSS software, version 24 was used to compute this statistic. In addition, based on the construct validity results (see section 3.2.3) we used the omega coefficient (McDonald, 1999), due to its adequateness for multidimensional constructs with unequal factor loadings. Values of this coefficient larger than 0.70 indicate good scale reliability.

3.2.3. *Construct validity*

We used Confirmatory Factor Analysis (CFA) to investigate the construct validity of the SELFIE instrument and to explore the SELFIE areas/factors. The factors were based on the DigCompOrg framework and its further adaptations from the SELFIE pilot phase, specifying the factors and how they relate to items. The goodness of fit was assessed with the comparative-fit index (CFI); the Tucker–Lewis index (TLI): the root-mean square error of approximation (RMSEA); and standardised root-mean-square residual (SRMR). With respect to the first two, values greater than 0.95 indicate satisfactory fit (Hu & Bentler, 1999). With respect to the last two, values less than 0.08 indicate a satisfactorily low level of noise in the model (Browne & Cudeck, 1992). We use the Mplus 8 software and maximum likelihood (ML) estimation for all analyses. For these analyses, we also consider random samples (25%, 50%, and 75%) of the data in order to perform robustness checks of the results obtained.

4. Results

4.1. *Items quality*

We look at the items characteristics using IRT[19] for SL and T considering the SELFIE pooled data in order to check the psychometric characteristics of the items. These analyses allow us to check if the SELFIE instrument comprises appropriate items (questions or statements).

Table 3 presents the item parameters, estimated separately by area, for SL and T considering the SELFIE pooled data[20]. Results of the item parameters estimates show that the items for SL and T are discriminative (α > 0.4), meaning that can discriminate between schools (and individuals within schools) with a high level of digitalisation and schools with a low level of digitalisation. In general, the SL and T items present acceptable values for the difficulty parameter and the items are mainly easy and of medium difficulty.[21] This is true for the pooled data (Table 3) and also across all ISCED levels (please see Tables A1 and A2 of Appendix A). Table 3 also shows that the less discriminative for SL are items 4.5, 2.5, 2.6, and 2.4. As for T, the less discriminative items are 5.1, 4.5, 2.6, and 2.5. However, considering the pooled data, all SELFIE items for SL and T present acceptable values for this parameter. The results from the psychometric analysis show that all the items from the different respondent groups are acceptable and can be used for further analysis (e.g. CFA).

[16] The respondents selecting the "*prefer not to say*" or "*not applicable*" option were not considered in the analyses.

[17] We looked at the items quality using the classical statistical analysis, considering the following measures: Item-Score correlation and Cronbach's Alpha if Item Deleted. Results are available upon request.

[18] This model also provides information about the four threshold parameters (d) per item that represent deviations from the general location/difficulty. The results are available upon request.

[19] Prior to the SELFIE tool questionnaire, we performed psychometric analyses based on the user consultation survey and on the SELFIE pilot data that included 14 countries. Based on these analyses, some information on the items was collected and decisions were taken in terms of the selection of constructs to be included in the tool and on the allocation of the items per areas.

[20] The pooled data with complete information for the analysis included the responses from 7,444 school leaders and 30,417 teachers. The item parameter estimates by ISCED level can be found in Tables A1 and A2 of appendix A. In the school report, the results are aggregated by area calculated as averages

[21] The classical statistical analysis of the items considering the Item-Score correlation also showed that these values are acceptable or good with all the Item-Score correlations higher than 0.4. The results from additional statistical checks, Cronbach's Alpha if Item Deleted, also indicated that all the items from the different respondent groups are acceptable. This is true for the SELFIE pooled data and also by ISCED level (1, 2 & 3).

Table 3
Item parameter estimates for SL and T – pooled data.

Item short title	SL - Item code	SL – Pooled		T - Item code	T – Pooled	
		α	β		A	β
Digital strategy	SL_1.1	1.683	−0.537	T_1_1	1.434	−0.506
Strategy development with teachers	SL_1.2	2.011	−0.567	T_1_2	1.823	−0.661
New ways of teaching	SL_1.3	0.972	−1.279	T_1_3	1.153	−1.004
Progress review	SL_1.4	1.102	−0.596	T_1_4	1.395	−0.499
Discussion on the use of technology	SL_1.5	0.953	−0.903	T_1_5	1.180	−0.641
Infrastructure	SL_2.1	1.586	−0.801	T_2_1	1.492	−0.678
Digital devices for teaching	SL_2.2	1.696	−1.002	T_2_2	1.651	−0.887
Internet access	SL_2.3	0.942	−1.276	T_2_3	1.042	−1.075
Technical support	SL_2.4	0.823	−0.884	T_2_4	0.924	−0.809
Data protection	SL_2.5	0.587	−0.950	T_2_5	0.701	−0.944
Digital devices for learning	SL_2.6	0.673	−0.768	T_2_6	0.667	−0.591
CPD needs	SL_3.1	0.850	−0.441	T_3_1	2.018	−0.623
Participation in CPD	SL_3.2	1.340	−0.781	T_3_2	1.497	−0.705
Sharing experiences	SL_3.3	2.732	−0.422	T_3_3	1.824	−0.760
Searching online educational resources	SL_4.1	2.242	−0.302	T_4_1	1.030	−1.509
Creating digital resources	SL_4.2	3.811	−0.144	T_4_2	1.076	−0.778
Using virtual learning environments	SL_4.3	2.586	−0.012	T_4_3	0.977	−0.547
Communicating with the school community	SL_4.4	0.874	−0.691	T_4_4	1.203	−0.985
Keeping data secure	SL_4.5	0.542	−1.286	T_4_5	0.666	−1.205
Tailoring to students' needs	SL_4.8	1.750	−0.977	T_4_8	1.359	−0.894
Fostering creativity	SL_4.9	2.585	−0.724	T_4_9	1.959	−0.840
Engaging students	SL_4.10	1.700	−0.547	T_4_10	1.666	−0.814
Student collaboration	SL_4.11	1.066	−0.619	T_4_11	1.932	−0.605
Cross-curricular projects	SL_4.12	1.107	−0.904	T_4_12	1.150	−0.483
Digital assessment	SL_5.1	0.917	−0.871	T_5_1	0.657	−0.861
Assessing skills	SL_5.2	1.663	−0.466	T_5_2	1.382	−0.555
Timely feedback	SL_5.3	2.641	−0.309	T_5_3	2.175	−0.432
Self-reflection on learning	SL_5.4	3.243	−0.158	T_5_4	2.850	−0.313
Feedback to other students	SL_5.5	2.196	−0.024	T_5_5	1.713	−0.150
Digital skills across subjects	SL_6.1	0.873	−0.639	T_6_1	0.864	−0.699
Safe behaviour	SL_6.2	2.422	−0.835	T_6_2	2.345	−0.848
Responsible behaviour	SL_6.3	2.661	−0.882	T_6_3	2.670	−0.881
Checking quality of information	SL_6.4	2.264	−0.690	T_6_4	2.476	−0.771
Giving credit to others' work	SL_6.5	1.394	−0.523	T_6_5	1.573	−0.666
Creating digital content	SL_6.6	0.986	−0.586	T_6_6	1.010	−0.711
Learning to communicate	SL_6.7	1.052	−0.871	T_6_7	1.107	−0.908

Regarding the parameter estimates of the core items of the S questionnaire (Table A3 of Appendix A), the results show that for all ISCED levels the items are discriminative, except item S_4_10 on engaging students through technology. We decided to reformulate this item. For ISCED 1 the items are of easy difficulty, for ISCED 2 the items are of easy and medium difficulty, and for ISCED 3 the items are of medium difficulty.[22]

4.2. Reliability

Table 4 presents the Cronbach's alpha and Omega analyses for the SL, T and S questionnaires considering the whole tool and by area.[23] The alpha coefficient for the 36 items (tool) answered by the SL varied from 0.970 (for ISCED 2) to 0.976 (in ISCED 1), suggesting that the items have high internal consistency. For T, the alpha coefficient for the whole tool presents its lower value in ISCED 3 (0.968) and the highest in ISCED 1 (0.973). Slightly lower values are found for the three ISCED levels of the S questionnaire. These results indicate that the SL and T full questionnaires have excellent internal consistency for the pooled data and across ISCED levels.

The results of the internal consistency per area show that highest values of the Cronbach's alpha are found for the *Assessment practices*, *Students digital competence* and *Teaching and learning - pedagogy* areas for SL and T. The results by area also revealed that the Cronbach's alpha for all the SELFIE users are acceptable, good or excellent with exception of the area *Teaching and learning – pedagogy*

[22] The test information and standard error curves of the tool considering all the core items for SL, T and S are presented in Figure B1 – Appendix B.
[23] Calculations of the Cronbach's alpha by area were based in the seven factor model (see point 4.3). For SL in the area "Students digital competence" the coefficient was calculated using all the items of the original area. For S the results were based on the original areas. "-" corresponds to the areas "Leadership", "Continuing professional development", "Teaching and learning – support", "Assessment practices" for S where there was only one item belonging to each area.

Table 4
Cronbach's alpha and Omega coefficient.

		Tool (all core items)	Areas						
			Leadership	Infrastructure and Equipment	Continuing professional development	Teaching and learning – support	Teaching and learning - pedagogy	Assessment practices	Students digital competence
Cronbach's alpha									
SL	**Pooled**	**0.973**	0.846	0.866	0.855	0.836	0.926	0.935	0.925
	ISCED 1	**0.976**	0.852	0.866	0.846	0.845	0.933	0.939	0.931
	ISCED 2	**0.970**	0.839	0.864	0.845	0.819	0.919	0.933	0.918
	ISCED 3	**0.974**	0.854	0.874	0.880	0.845	0.930	0.936	0.930
T	**Pooled**	**0.970**	0.862	0.879	0.877	0.811	0.916	0.925	0.935
	ISCED 1	**0.973**	0.859	0.883	0.877	0.811	0.919	0.928	0.942
	ISCED 2	**0.969**	0.862	0.875	0.873	0.812	0.916	0.923	0.931
	ISCED 3	**0.968**	0.868	0.877	0.879	0.805	0.912	0.923	0.932
S	**ISCED 1**	**0.885**	–	0.732	–	–	0.668	–	0.779
	ISCED 2	**0.935**	–	0.701	–	–	0.805	0.791	0.864
	ISCED 3	**0.942**	–	0.714	–	–	0.815	0.804	0.88
Omega coefficient									
SL	**Pooled**		0.915	0.907	0.897	0.879	0.951	0.960	–
	ISCED 1		0.916	0.840	0.868	0.886	0.955	0.961	–
	ISCED 2		0.918	0.833	0.864	0.867	0.946	0.959	–
	ISCED 3		0.910	0.853	0.892	0.890	0.957	0.963	–
T	**Pooled**		0.910	0.932	0.917	0.873	0.943	0.949	0.961
	ISCED 1		0.910	0.935	0.920	0.873	0.946	0.952	0.965
	ISCED 2		0.909	0.931	0.914	0.874	0.944	0.950	0.959
	ISCED 3		0.916	0.931	0.918	0.871	0.941	0.947	0.959

for S that has a questionable value (0.668).

In terms of the omega[24] coefficient per area it varies from 0.833 (for SL – ISCED 2, Infrastructure and equipment area) to 0.965 (for T – ISCED 1, Students digital competence area). All the values are higher than 0.7 revealing a good scale reliability per area.

4.3. *Construct validity: CFA*

For the factor analysis, we consider the T as the main model. Teachers are in the best position to know the level of digital capacity of their schools, as they are both in the classroom and have information about the school context. Moreover, they are at the centre of the learning process in schools, being at the same time in constant contact with school leaders and students. Therefore, they know what is happening both in the classroom and at the school's administration levels.

The results of the CFA confirm a seven-factor structure indicating that the T questionnaire has this number of factors/areas. The area *Teaching and learning* is composed of items measuring support and items measuring *pedagogy*. In this sense, the areas for the T questionnaire are: *Leadership*, *infrastructure and equipment*, *Continuing professional development (CPD)*, *Teaching and learning* – support, *Teaching and learning – pedagogy*, *Assessment practices* and *Student digital competence*. As for the SL, there is a clear correspondence and the same areas are validated by the CFA except the last one.[25]

Fit of indexes for SL and T statistics for the seven-factor CFA model and one-factor CFA model are presented in Table 5. The comparison of the fit indexes for both models indicate that the seven-factor model has a better fit than the one-dimensional model and supports the proposed multidimensionality. In addition, the results show that, for the best model, the values for RMSEA are acceptable and the values for CFI and TLI are above the cut-off value (0.95) indicating that the CFA goodness of fit presents good and acceptable values for the T and SL questionnaires. In general, the results indicate that the SL and T seven-factor models are well fitted.

In addition to model fit indices, factor loadings and squared multiple correlations are examined for the seven-factor model (see Table 6). Squared multiple correlations (R-square) for individual items show that the model is adequately represented by the observed

[24] The coefficient was calculated based on the areas obtained from the CFA results (see point 4.3) and on unidimensional models for every single area.

[25] The area Student digital competence was not considered in the CFA for SL because the items measuring it referred to students' acquisition of digital competence and this information is more reliable when asked to the T.

Table 5
Fit indices for the models: seven factor and one factor.

		Seven-factor CFA model				One-factor CFA model			
		RMSEA	CFI	TLI	SRMR	RMSEA	CFI	TLI	SRMR
T	Pooled	0.073	0.968	0.963	0.029	0.162	0.831	0.819	0.101
	ISCED 1	0.075	0.969	0.965	0.029	0.154	0.917	0.909	0.066
	ISCED 2	0.073	0.967	0.963	0.030	0.169	0.868	0.856	0.084
	ISCED3	0.070	0.968	0.964	0.031	0.164	0.884	0.873	0.082
SL	Pooled	0.078	0.975	0.971	0.030	0.165	0.883	0.872	0.076
	ISCED 1	0.077	0.980	0.977	0.029	0.154	0.917	0.909	0.066
	ISCED 2	0.078	0.973	0.969	0.032	0.169	0.868	0.856	0.084
	ISCED3	0.074	0.978	0.974	0.030	0.162	0.884	0.873	0.082

Table 6
R-square and standardised factor loadings – seven factor for pooled data.

Areas	Item short title	T			SL		
		Items	R-square	Standardised factor loadings	Items	R-square	Standardised factor loadings
Leadership	Digital strategy	T_1.1	0.797	0.893	SL_1_1	0.811	0.9
	Strategy development with teachers	T_1.2	0.784	0.886	SL_1_2	0.823	0.907
	New ways of teaching	T_1.3	0.734	0.857	SL_1_3	0.71	0.843
Infrastructure and equipment	Infrastructure	T_2.1	0.76	0.872	SL_2_1	0.729	0.854
	Digital devices for teaching	T_2.2	0.719	0.848	SL_2_2	0.695	0.834
	Internet access	T_2.3	0.583	0.764	SL_2_3	0.543	0.737
	Technical support	T_2.4	0.624	0.79	SL_2_4	0.584	0.764
	Data protection	T_2.5	0.62	0.787	SL_2_5	0.588	0.767
	Digital devices for learning	T_2.6	0.59	0.768	SL_2_6	0.579	0.761
Continuing professional development	CPD needs	T_3.1	0.786	0.886	SL_3_1	0.748	0.865
	Participation in CPD	T_3.2	0.759	0.871	SL_3_2	0.703	0.839
	Sharing experiences	T_3.3	0.817	0.904	SL_3_3	0.779	0.883
Teaching and learning – support	Searching online educational resources	T_4.1	0.519	0.721	SL_4_1	0.566	0.752
	Creating digital resources	T_4.2	0.598	0.774	SL_4_2	0.667	0.817
	Using virtual learning environments	T_4.3	0.748	0.865	SL_4_3	0.736	0.858
	Communicating with the school community	T_4.4	0.666	0.816	SL_4_4	0.609	0.78
Teaching and learning –pedagogy	Tailoring to students' needs	T_4.8	0.753	0.868	SL_4_8	0.776	0.881
	Fostering creativity	T_4.9	0.787	0.887	SL_4_9	0.825	0.908
	Engaging students	T_4.10	0.741	0.861	SL_4_10	0.782	0.884
	Student collaboration	T_4.11	0.815	0.903	SL_4_11	0.824	0.908
	Cross-curricular projects	T_4.12	0.75	0.866	SL_4_12	0.762	0.873
Assessment practices	Assessing skills	T_5.2	0.78	0.883	SL_5_2	0.786	0.887
	Timely feedback	T_5.3	0.837	0.915	SL_5_3	0.868	0.932
	Self-reflection on learning	T_5.4	0.879	0.938	SL_5_4	0.917	0.957
	Feedback to other students	T_5.5	0.801	0.895	SL_5_5	0.853	0.924
Student digital competence	Safe behaviour	T_6.2	0.811	0.9			
	Responsible behaviour	T_6.3	0.82	0.906			
	Checking quality of information	T_6.4	0.858	0.926			
	Giving credit to others' work	T_6.5	0.785	0.886			
	Creating digital content	T_6.6	0.752	0.867			
	Learning to communicate	T_6.7	0.789	0.888			

measures and between 52% and 88% of the variance on individual items is accounted for by their assigned factors. The combined fit indices, factor loadings and squared multiple correlations supported the factor structure. Further, all observed variables were found to be significant, with t-values > 1.96 ($p < 0.05$). All factor loadings are statistically significant ($p < 0.05$). The same is also true for SL.

However, it has to be noted that in the solution above, some items (4.5 - Keeping data secure, 5.1 - Digital assessment, 6.1 - Digital skills across subjects, 1.4 - Progress review and 1.5 - Discussion on the use of technology) did not fit well in the factors and we excluded them from the final solution. The reasons are as follows.

Firstly, item 4.5 may not fit because keeping the data secure does not always depend on the digital competence of the teachers. The capacity to keep data secure is often related to infrastructure and available software in the school. This is confirmed by the fact that item 2.5, referred to as the existence of data protection systems, works well in the *Infrastructure and equipment* area and is highly correlated with 4.5 ($r_{SL} = 0.64$, $r_T = 0.48$). So, we decided to eliminate item 4.5 from the tool. Secondly, the core items 5.1 and 6.1 may not fit in their area because they refer to a different concept when compared with the other items in the area. Item 5.1 does not refer to an assessment practice using technology but to institutional support for it and item 6.1 does not refer to a specific digital competence as a learning outcome, but to a specific way to deliver digital skills across subjects.[26] Considering the input from the quantitative analysis and qualitative evidence (usability test, qualitative data from open questions and consultation with experts and users performed after the quantitative results were available) items 5.1 and 6.1 became optional.

Finally, for items on "progress review (in teaching and learning with digital strategies)" and "discussion on the use of technology" (items 1.4 and 1.5, respectively) a different explanation can be provided. They are strongly correlated among them ($r_{SL} = 0.6$, $r_T = 0.7$) and conceptually point towards *collaboration and networking*. This was an original area of the DigCompOrg model that was eliminated in post-pilot refinement of the items and that seems to emerge again due to the better quality and validity of the items now included. As two items are very few to test factorial structure, additional CFA analyses were performed to explore the separated existence of this area considering the optional item 1.6.[27] This item refers to the use of digital technologies in schools partnerships with other organisations and together with items 1.4 and 1.5 confirms an independent area entitled *Collaboration and networking* as in the original theoretical foundation: the DigCompOrg framework.

The CFA was also done by ISCED level (ISCED 1, ISCED 2, ISCED 3)[28] and for robustness check considering random samples of the data (50%, 25%, 75%).[29] In general, the results are also good and/or acceptable.

5. Discussion and conclusion

Improving the use of digital technologies in teaching and learning in schools is a policy priority for countries in Europe and beyond (European Commission, 2019c). There are studies that try to capture the development of the use of technologies in schools at regional, national or international level using representative surveys (e.g. European Commission, 2019a). These studies provide useful insights about the overall adoption of digital technologies for teaching and learning in compulsory education. However, they do not provide information about the digital capacity of each school. Yet, school leaders, teachers and students need a practical way to reflect on what goes well and not so well in their own school and to develop actions for a better integration of digital technologies based on accurate information.

While there are some tools that go in this direction, they tend to focus only on some of the dimensions identified as needed for reflecting on and planning for the digital capacity of the schools (Kampylis, Devine, Punie, & Newman, 2016; Kampylis, Punie, & Devine, 2015). Furthermore, they generally have a one-size-fits all approach and do not allow school communities to customise them for their needs. Focusing on content quality, the items included in these tools do not usually use psychometric methods, making it difficult to guarantee the quality of the information provided. Moreover, the vast majority of these tools do not include the students' voice, but only incorporate the viewpoints of school leaders and teachers, although the ultimate goal of embedding digital technologies in schools is to enhance students' learning and digital skills (Kampylis, Devine, Punie, & Newman, 2016).

SELFIE is the European Commission's response to the need to provide schools with a scientifically sound, comprehensive and easy-to-use tool for reflecting on their digital capacity. Conversely to other tools, SELFIE encompasses all key groups in the school community, asking school leaders, teachers and students their opinion on how technologies are used at their school. It provides a 360-degree view of the use of digital technologies in a school following the holistic approach of the DigCompOrg model. Evidently, it was challenging to design a tool for all schools across Europe where there are different settings as well as cultural and contextual differences. However, by allowing schools to customise the tool to their own needs SELFIE can be relevant and useful to any school community that wants to reflect on the digital capacity of their school.

In this paper, we have analysed the psychometric properties of the SELFIE questionnaires to see whether they fulfil the necessary standards to provide accurate information to the schools regarding their digital capacity, which was our guiding research question (RQ). For this, we looked at the quality of the items, as well as the reliability and the construct validity of SELFIE questionnaires from the perspective of different users and in three ISCED levels.

We were able to respond to all the sub-research questions, and found that items are reliable and are discriminative for the schools with more room to improve, confirming its capacity to capture digital capacity of schools and differentiate between schools in different stages on the use of technology for teaching and learning (RQ1). By design, the tool can be used by all types of schools and includes several kinds of items. Our analyses confirm that core items are of low or medium difficulty and are relevant for the vast majority of schools that participate in SELFIE. However, the tool offers, at the same time, more "advanced" schools in terms of digital capacity the possibility to select optional questions and create their own ones so as to better capture the most innovative uses of educational technology. All the other schools can also benefit from this functionality to customise the tool to their specific needs. Items have also good construct validity (RQ2) and work well for different users and educational levels (RQ3). Only one item of the S questionnaire on engaging students on digital learning activities does not work well. We reformulated it as follows: "In our school, teachers give us interesting activities using technology". Analysing in depth the evidence raised for RQ2, the results also showed that the confirmatory

[26] Item 5.1 was tested in the *Leadership* area and item 6.1 was tested in the *Teaching and Learning* area, but in both cases, we got unacceptable values for the factor analysis. Due to this reason, we eliminated the items 5.1 and 6.1 from the analyses.

[27] The results indicate that the model for T (pooled data) is well fitted (RMSEA = 0.068, CFI = 0.967, TLI = 0.962, SRMR = 0.028). The parameters were obtained considering 14,561 teachers. The same is true for the SL considering 7,444 school leaders. More detailed information is available upon request.

[28] More information by ISCED level can be found in appendix A (tables A4 and A5).

[29] Available upon request.

factor analysis for the T questionnaire works perfectly. This is of great importance because teachers are at the centre of the learning process and they know what is happening both in the classroom and in terms of school's digital strategy. The results are slightly less geared towards school leaders. In particular, our results indicate that there are some problems with the last area of the school leader questionnaire (i.e. Student digital competence). An explanation is that SL do not have enough information about students' acquisition of digital competence in the classroom. These findings could imply that we could drop this area from the school leaders' questionnaire. However, they may prefer to keep it as it improves the comparability of the teachers' and school leaders' questionnaires, which is very useful for the schools as indicated from the analysis of the pilot phase data.

As a consequence, the final proposal of the area and items structure[30] is as follows. It is a proposal that merges the results from the underlying theory of the questionnaires and the statistical outcomes presented in this paper. As for the areas, the suggestion is to add the area on *Collaboration and networking* and the other areas stay the same; this means that the areas that appear from our analyses are *Leadership*, *Collaboration and networking, Infrastructure and equipment*, *Continuing professional development*, *Teaching and learning* – support, *Teaching and learning – pedagogy*, *Assessment practices*, and *Student digital competence*. For the items that did not fit well with the existent factors of the CFA, we deleted the item 4.5 on "keeping the data secure" and made the items on "digital assessment" (5.1) and on "digital skills across subjects (6.1) optional due to the importance for the schools. These decisions come from the combination of quantitative analysis and discussion based on a usability test, qualitative data from open questions, and consultation of the results with experts.

Contrary to other self-reflection or self-assessment tools for schools' digital capacity, we checked the quality of the SELFIE questionnaires using some of the standard procedures, similar to those from large scale assessments (e.g., OECD, 2017) and the results are quite satisfactory. However, one has to be aware of the limitations of using self-reflection data. For example, it could be an expression of a perception but not necessarily reflect the true state of the reality in schools. Consequently, in some cases the responses may face bias associated with self-measurement such as the Dunning-Krueger effect (Kruger & Dunning, 1999), self-experiences or social desirability bias (Faddar, Vanhoof, & De Maeyer, 2018) However, these biases do not invalidate the usefulness of SELFIE (and other self-reflection tools approaches). Firstly, because by responding to the tool, individuals can reflect on their knowledge and use of digital technologies in teaching or learning and can identify practices that should be improved. Secondly, there is no clear reason to believe that these effects are stronger in one area than another and, consequently, the relative information about school's strengths and weaknesses at an area or item level remains unbiased. In addition, the fact that SELFIE asks the same questions to school leaders and teachers allows detecting inconsistencies and working as a measure against possible bias. Also, by design, school leaders, teachers, and students do not provide any personal information when they fill in SELFIE questionnaires, so their participation is on an anonymous basis allowing them to be able to provide their input without the fear of being identified. Research shows that sharing experiences and discussing technology usage is a good way to learn how to use it in a meaningful way (Ilomäki & Lakkal, 2018; Jeladze & Pata, 2017). The results from the tool can be used to guide an internal debate or deliberative process to develop concrete actions or strategies for improving the use of digital technologies in a school.

Moreover, the improvement of tools must be periodically revised and refined (Huggins, Ritzhaupt, & Dawson, 2014). As digital technologies are evolving rapidly, developing valid and reliable instruments and assessment tools to measure a construct is a difficult and ongoing process (Hohlfeld, Ritzhaupt, & Barron, 2010). There are several ways to continue to improve SELFIE questionnaires and to analyse their properties in the future. On the one hand, data collected through SELFIE comes from its users: school communities that are usually self-selected to use the tool. Since schools using SELFIE are not representative of the schools in a country or region, we cannot infer that the psychometric analyses presented in this paper would be the same as in a representative sample. For instance, the difficulty level of the core items could vary according to a higher or lower schools' use of digital technologies. For this reason, it would be suitable to run the same analyses with data that is collected from a representative sample. However, it could only be an issue when scaling-up the use of SELFIE to a whole region or country, since schools using SELFIE do not necessarily reflect the population characteristics and in this paper, we have shown that the tool works well for those schools that decide to use it. In addition, an assessment of the measurement invariance of the tool would allow us examining whether the construct has the same meaning under different circumstances. This would be important for a tool that was tested among different languages and cultures. Such analyses could further check the robustness of the results presented in this paper. On the other hand, counterfactual impact evaluations could assist us in verifying whether the use of SELFIE has had a positive causal effect on some outcomes of interest, such as development of strategy on the use of digital technologies, participation in CPD on digital technologies, change in teaching practices and (ultimately) improving teaching and learning in schools.

To sum up, our results indicate that the tool is reasonably reliable and valid for being used in self-reflection process that involves the whole school community. It means that SELFIE complies with the three criteria proposed by Antoniou et al. (2016) for effective self-evaluation exercises: a) it is based in a validated tool that stems from a theoretical framework, b) it is inclusive regarding stakeholders views and c) it facilitates data for decision making on multiple levels. We therefore can conclude that SELFIE can provide a good basis for school communities to reflect on and debate the digital capacity of their schools and to develop action plans to integrate digital technologies in one or more of the different areas covered by the tool. Both the inclusive and flexible design of SELFIE and the robust psychometric examination of its items presented in this paper are a novelty in the field and can offer information on the development of self-reflection or self-assessment tools for schools in the future.

[30] See our proposal for SELFIE full questionnaires in Appendix C.

Author contributions

Patrícia Costa: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing - original draft. Jonatan Castaño Muñoz: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing - original draft, Writing - review & editing . Panagiotis Kampylis: Conceptualisation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Visualization, Validation, Writing - original draft, Writing - review & editing.

Disclaimer

The views expressed in this article are purely those of the authors and should not be regarded as the official position of the European Commission.

Acknowledgments

This research has been funded by the European Commission. The authors would like to thank the reviewers for their thoughtful comments, which have improved the quality of the manuscript. We would also like to thank the colleagues who provided feedback on early versions of the paper: Artur Pokropek, Ralph Hippe, Lilian Weikert, Yves Punie, Georgios Kapsalis, Giorgio Di Pietro, Deirdre Hodson and Tabetha Newman. Thanks also go to the members of the SELFIE community who contributed to the development, testing and implementation of the tool. Finally, special thanks go to the thousands of school leaders, teachers and students who have used the tool so far to reflect on the digital capacity of their schools.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2020.104080.

References

Antoniou, P., Myburgh-Louw, J., & Gronn, P. (2016). School self-evaluation for school improvement: Examining the measuring properties of the LEAD surveys. *Australian Journal of Education*, 1–20, 0(0).

Baker, F., & Kim, S. (2004). *Item response theory: Parameter estimation techniques*. New York: Marcel Dekker.

Balaban, I., Redjep, N. B., & Calopa, M. K. (2018). The analysis of digital maturity of schools in Croatia. *International Journal of Emerging Technologies in Learning, 13*(6), 4–15.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Browne, M., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230–258.

Castaño-Muñoz, J., Costa, P., Hippe, R., & Kampylis, P. (2018). Within-school differences in the views on the use of digital technologies in Europe: Evidence from the SELFIE tool. *Proceedings of the 10th International Conference on Education and New Learning Technologies* (pp. 10417–10426). Palma de Mallorca: IATED.

Chapman, C., & Sammons, P. (2013). *School self-evaluation for school improvement: What works and why? Berkshire*. CfBT Education Trust. https://www.educationdevelopmenttrust.com/EducationDevelopmentTrust/files/d8/d87c1064-51bf-4b7c-b984-85ccb46a54d1.pdf.

Comi, S. L., Argentin, G., Gui, M., Origo, F., & Pagani, L. (2017). Is it the way they use it? Teachers, ICT and student achievement. *Economics of Education Review, 56*, 24–39.

Conrads, J., Rasmussen, M., Winters, N., Geniet, A., & Langer, L. (2017). Overview and analysis of policies for the integration and effective Use of digital technologies in education. Retrieved 12 September 2019, from https://europa.eu/!Tc84QQ.

European Parliament, & the Council. (2006). *Recommendation of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning.* Official Journal of the European Union. L394/310 https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:394:0010:0018:en.pdf.

Escueta, M., Quan, V., Nickow, A. J., & Oreopoulos, P. (2017). Education technology: An evidence-based review. from http://www.nber.org/papers/w23744. (Accessed 12 September 2019).

European Commission. (2018). Digital education action plan. Retrieved 05 September 2019, from https://europa.eu/!Jf36XU.

European Commission. (2019a). 2nd survey of schools: ICT in education. Objective 2: Model for a 'highly equipped and connected classroom. Final Report. Retrieved 25 August 2019, from https://europa.eu/!Bm97bR.

European Commission. (2019b). Digital economy and society index report 2019 - human capital. Retrieved 05 September 2019, from https://europa.eu/!wT97Bt.

European Commission. (2019c). *Digital education at school in Europe - Eurydice report*. Luxembourg: Publications Office of the European Union.

European Commission. (2020). Digital economy and society index. Retrieved from https://bit.ly/31fhC26. (Accessed 12 June 2020) https://bit.ly/31fhC26.

Faddar, J., Vanhoof, J., & De Maeyer, S. (2018). School self-evaluation: Self-perception or self-deception? The impact of motivation and socially desirable responding on self-evaluation results. *School Effectiveness and School Improvement, 29*(4), 660–678.

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Duckworth, D. (2019). Preparing for life in a digital world - IEA international computer and information literacy study 2018 international report. Retrieved January 2020.Retrieved 29 June 2020, from https://www.iea.nl/publications/study-reports/preparing-life-digital-world.

Hohlfeld, T. N., Ritzhaupt, A. D., & Barron, A. E. (2010). Connecting schools, community, and family with ICT: Four-year trends related to school level and SES of public schools in Florida. *Computers & Education, 55*, 391–405.

Huggins, A. C., Ritzhaupt, A. D., & Dawson, K. (2014). Measuring information and communication technology literacy using a performance assessment: Validation of the student tool for technology literacy (ST2L). *Computers & Education, 77*(1), 1–12.

Ilomäki, L., & Lakkal, M. (2018). Digital technology and practices for school improvement: Innovative digital school model research and practice in technology. *Enhanced Learning, 13*(1), 25.

Jeladze, E., & Pata, K. (2017). Change management: Digital transition in the public schools of Georgia. In L. Uden, D. Liberona, & Y. Liu (Eds.), *Learning technology for education challenges: 6th international workshop, LTEC 2017* (pp. 171–185). New York: Springer.

Jeladze, E., & Pata, K. (2018). Smart, digitally enhanced learning ecosystems: Bottlenecks to sustainability in Georgia. *Sustainability, 10*, 8, 2672.

Jisc. (n.d.). Organisational digital capability. Retrieved from https://bit.ly/2Z466nB.

Kampylis, P., Punie, Y., & Devine, J. (2015). *Promoting Effective Digital-Age Learning: A European Framework for Digitally-Competent Educational Organisations*. Luxembourg: Publications Office of the European Union. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC98209/jrc98209_r_digcomporg_final.pdf.
Kampylis, P., Devine, J., Punie, Y., & Newman, T. (2016). Supporting schools to go digital: From a conceptual model towards the design of a self-assessment tool for digital-age learning. *9th International Conference of Education, Research and Innovation* (pp. 816–825). Sevilla (Spain): IATED Academy.
Killen, C., Beetham, H., & Knight, S. (2017). *Developing organisational approaches to digital capability. Jisc guide*. Retrieved 05 September 2019, from https://www.jisc.ac.uk/guides/developing-organisational-approaches-to-digital-capability.
Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134.
Lim, C.-P., Zhao, Y., Tondeur, J., Chai, C.-S., & Tsai, C.-C. (2013). Bridging the gap: Technology trends and use of technology in schools. *Educational Technology & Society, 16*(2), 59–68.
MacBeath, J. (2005). *Background, principles and key learning in self-evaluation: A guide for school leaders*. Nottingham: National College for School Leadership.
Methods and procedures in PIRLS 2016. Retrieved 02 september 2019 from Boston college, TIMSS & PIRLS international study center website:. (2017) In Martin. M. O., Mullis, I. V. S., & Hooper, M. (Eds.), (2017). https://timssandpirls.bc.edu/publications/pirls/2016-methods.html.
McDonald, R. P. (1999). *Test theory: A unified approach*. Mahwah, NJ: Erlbaum.
Mora, T., Escardíbul, J.-O., & Di Pietro, G. (2018). Computers and students' achievement: An analysis of the one laptop per child program in catalonia international. *Journal of Educational Research, 92*, 145–157.
Muraki, E., & Bock, R. D. (2002). *PARSCALE: IRT-based test scoring and item analysis forgraded open-ended exercises and performance tasks (Version 3)*. Chicago: Scientific Software.
Muraki, E. (1993). Information functions of the generalized partial credit model. *Applied Psychological Measurement, 17*, 351–363.
Muraki, E. (1997). A generalized partial credit modelW. J. van der Linden, & R. K. Hambleton (Eds.). . In *Handbook of Modern Item Response Theory* (pp. 153–164). New York: Springer.
OECD. (2015). *Students, computers and learning - making the connection*. https://doi.org/10.1787/9789264239555-en. from . (Accessed 5 September 2019).
OECD. (2017). *PISA 2015 technical report*. from https://www.oecd.org/pisa/sitedocument/PISA-2015-technical-report-final.pdf. (Accessed 2 September 2019).
Ottestad, G., & Guðmundsdóttir, G. B. (2018). Information and communication technology policy in primary and secondary education in Europe. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Handbook of information technology in primary and secondary education* (pp. 1–21). Cham: Springer.
Petko, D., Cantieni, A., & Prasse, D. (2017). Perceived quality of educational technology matters: A secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 test scores. *Journal of Educational Computing Research, 545*(8), 1070–1091.
Rodrigues, M., & Biagi, F. (2017). *Digital technologies and learning outcomes of students from low socio-economic background: An Analysis of PISA 2015*. from https://europa.eu/!Ky63VB. (Accessed 10 September 2019).
Rogers, A. (2014). *The base of the iceberg: Informal learning and its impact on formal and non-formal learning*. Opladen/Berlin/Toronto: Barbara Budrich Publishers.
Skryabin, M., Zhang, J., Liu, L., & Zhang, D. (2015). How the ICT development level and usage influence student achievement in reading, mathematics, and science. *Computers & Education, 85*, 49–58.
Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81*(1), 4–28.
Tanhua-Piiroinen, E., & Viteli, J. (2017). Opeka and Ropeka, the self-assessing services for teachers and principals. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Proceedings of EC-TEL 2017 - data driven approaches in digital education* (pp. 602–605). Cham: Springer.
Vanderlinde, R., & van Braak, J. (2010). The e-capacity of primary schools: Development of a conceptual model and scale construction from a school improvement perspective. *Computers & Education, 55*, 541–553.
Vanderlinde, R., van Braak, J., & Dexter, S. (2012). ICT policy planning in a context of curriculum reform: Disentanglement of ICT policy domains and artifacts. *Computers & Education, 58*(4), 1339–1350.
Voogt, J., Knezek, G., Cox, M., Knezek, D., & Brummelhuis, A.t. (2011). Under which conditions does ICT have a positive effect on teaching and learning? A call to action. *Journal of Computer Assisted Learning, 29*(1), 4–14.