The youth Digital Skills Indicator

Technical Report · January 2021

DOI: 10.5281/zenodo.4476540

4 authors, including:

Ellen Johanna Helsper
The London School of Economics and Political Science
72 PUBLICATIONS **6,446** CITATIONS
SEE PROFILE

Luc S Schneider
The London School of Economics and Political Science
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Ester van Laar
University of Twente
11 PUBLICATIONS **585** CITATIONS
SEE PROFILE

youth SKILLS

The youth Digital Skills Indicator

Ellen J. Helsper
Luc S. Schneider
Alexander J.A.M. van Deursen
Ester van Laar

Please cite this report as:

Helsper, E.J., Schneider, L.S., van Deursen, A.J.A.M., & van Laar, E. (2020). *The youth Digital Skills Indicator: Report on the conceptualisation and development of the ySKILLS digital skills measure*. KU Leuven, Leuven: ySKILLS.

DISCLAIMER

This project has received funding from the European Union's Horizon 2020 Research & Innovation programme under Grant Agreement no. 870612. The information in this deliverable reflects only the authors' views and the European Union is not liable for any use that may be made of the information contained therein.

DISSEMINATION LEVEL

Public

Project: ySKILLS – Youth Skills
GA: 870612
Call: H2020-SC6-TRANSFORMATIONS-07-2019
Type of action: RIA

The youth Digital Skills Indicator

Report on the conceptualisation and development of the ySKILLS digital skills measure

Work Package 3 – Deliverable 3.3

Due date: **31 December 2020**

Submission date: **31 December 2020**

Lead beneficiary: **London School of Economics and Political Science (LSE)**

Authors:

Ellen J. Helsper, Luc S. Schneider, Alexander J.A.M. van Deursen, Ester van Laar

Table of contents

Executive summary 5

Glossary of key terms 6

1 Introduction 7

1.1 The ySKILLS project 7

1.2 This report 8

2 Final short version of the youth Digital Skills Indicator (yDSI) 10

2.1 Question and answer scale formulation of the digital skills items 11

2.2 Youth Digital Skills Indicator (yDSI) – digital skills items 12

2.3 Question and answer scale formulation of the digital knowledge items 13

2.4 Youth Digital Skills Indicator (yDSI) – digital knowledge items 13

3 Review of the digital skills literature 14

3.1 Academic literature: Conceptual frameworks for digital skills 14

3.2 Conceptual framework for digital skills based on the academic literature 15

4 Literature on (digital) skills measurement 17

4.1 Types of methodologies used to measure digital skills 17

4.2 Best practice guidelines for large-scale population research on digital skills 18

4.2.1 Common errors in digital skills item design 19

4.2.2 Best practice guidelines for skill item survey design 19

4.2.3 Best practice guidelines for question and answer scale design 21

4.3 Selection of academic and grey literature measuring digital skills 22

4.3.1 Academic literature item selection 22

4.3.2 Grey literature item selection 22

4.3.3 Selection of the best items from the academic and grey literature 23

4.4 Findings: Measurement instruments used in digital skills testing 24

4.4.1 Representation of conceptualised skills dimensions in measurement instruments 24

4.4.2 Presence of best and problematic practices in the skills measurement literature 25

4.4.3 Validity and reliability testing in measurement 26

5 Initial digital skills question and answer formulation and item selection 28

6 Validation of the yDSI – Step 1: Cognitive interview and pilot survey testing 30

6.1 Methodology for cognitive interviews validation 30

6.1.1 The sampling cognitive interviews 30

6.1.2 Fieldwork procedure for the cognitive interviews 31

6.2 Methodology for survey pilot validation 32

6.2.1 Pilot survey sampling 32

6.2.2 Pilot survey analytical procedures 34

6.3 Results: Validation through cognitive interviews 35

6.3.1 General issues 35

6.3.2 Country-specific item issues 36

6.3.3 Selection of items based on content validity 37

6.4 Results: Validation through pilot surveys 37

6.4.1 Statistical properties of the digital skills items in the pilot survey 38

6.4.2 Statistical properties of the digital knowledge items in the pilot surveys 43
7 Validation of the yDSI – Step 2: Performance tests 46
7.1 Review of existing performance tests 46
7.2 ySKILLS performance task design 46
7.2.1 Part 1. Information navigation and processing tasks: Navigating 47
7.2.2 Part 2. Critical information navigation and processing: Evaluating 47
7.2.3 Part 3. Communication and interaction skills tasks: Protecting 47
7.2.4 Part 4. Critical communication and interaction tasks: Netiquette 48
7.2.5 Part 5. Content creation and production tasks: Producing, attracting and understanding 48
7.3 Procedure performance tests 48
7.3.1 Classroom setting 49
7.3.2 Individual online sessions 49
7.4 Sample performance test 49
7.5 Results: Validation through performance tests 50
8 Conclusions: Use and construction of yDSI scales 55
8.1 The properties of the short version of the yDSI scale 55
8.2 Long version of the yDSI scales 58
8.2.1 Long version of the yDSI digital skills instrument 58
8.2.2 Long version of the yDSI digital knowledge instrument 61
8.3 Guidelines for the creation of composite scales for analysis in the survey 62
Acknowledgements 63
References 64
Appendices (English versions only) 70
A. Details of adjustments made to the skills survey instrument after partner discussion 70
B. Question and answer formulation and items tested in cognitive interviews and pilot surveys (including source and adaptation notes) 75
C. Instructions for probing for cognitive interviews on skills questions 82
D. Descriptives: Digital skills items for full sample based on pilot survey 92
E. Factor analyses: Digital skills items based on pilot survey 96
F. Reasoning for deletion or modification of items after piloting and cognitive interviews . 100
G. Performance tasks 103

Executive summary

This report presents the youth Digital Skills Indicator (yDSI), a unique, extensively cross-nationally validated measurement tool with 31 items, distributed over digital skills and digital knowledge questions, that can be used for large-scale population research.

The yDSI is the only measurement tool for youth digital skills that has been tested using the full range of validation practices. Over a period of six months, consultation with experts (face validity), cognitive interviews (content validity), pilot surveys (construct validity) and performance tests (criterion validity) with young people were conducted in a wide range of European countries.

A review of the literature led to a framework identifying four dimensions that constitute digital skills: (1) technical and operational skills; (2) information navigation and processing skills; (3) communication and interaction skills; and (4) content creation and production skills.

Across all four dimensions a distinction should be made between being able to use the functionalities of information and communication technologies (ICTs) (functional aspects) and understanding why ICTs are designed and content is produced in certain ways and being able to use that knowledge in managing interactions in and with digital spaces (critical aspects).

Existing publications that report on survey instruments to measure digital skills, tend to cover technical and operational and information navigation and processing skills more than they do communication and interaction and content creation and production skills. Furthermore, functional aspects are more commonly measured than critical aspects of skills.

Many studies that present survey items for the measurement of digital skills fall foul of seven "sins". These studies (1) have basic survey item design flaws; (2) are solely PC-based; (3) are too vague or general; (4) measure outcomes instead of skills; (5) measure use instead of skills; (6) measure attitudes instead of skills; and (7) measure confidence instead of skills.

Seven best practices for digital skills survey design are proposed to prevent the seven problematic practices from occurring: (1) ask participants "Can you do?" or "Do you know how to do?" (skill) rather than "Have you done?" or "Do you do?" (use); (2) avoid device-, app- or activity-specific items; (3) include (functional) digital skills and (critical) digital knowledge items; (4) at least half of the digital knowledge items should involve statements that are untrue; (5) items should ask "Do you know how to do?" (skill) rather than "How good are you at?" or "How do you rate yourself on?" (confidence); (6) items should use truth claims and emphasise the here and now to make the person evaluate their actual personal skills; and (7) answer options should be scale-based and include an option encouraging people to admit to a lack of understanding to avoid social desirability bias.

Cognitive interviews and performance tests showed that many young people did not master a range of skills, including critical information navigation and processing skills. Moreover, these were the hardest to measure cross-nationally. Knowledge around how content was created and produced was also lacking.

Analyses of skewness and kurtosis, confirmatory factor analysis, difficulty estimation and equivalence testing established that the final short version of the yDSI has overall high construct, convergent and discriminant validity. This means that the hypothesised four skills dimensions are clearly present in the yDSI, and that items measure variety within each dimension.

Glossary of key terms

ICTs: Information and communication technologies.

Skills **item** (in a questionnaire): A specific skill a person is asked to evaluate in a questionnaire.

Skills **question** (in a questionnaire): The way in which the question to evaluate a skill is asked.

Skills **answer scale** (in a questionnaire): The possible answer categories on which the person is asked to evaluate the level of their skill.

Cognitive interviews: Qualitative interviews used to validate survey instruments, testing whether respondents have understood the question, the items and the answers in the way intended by the survey research designers.

Longitudinal panel surveys: Surveys that ask the same respondents to fill out the same questionnaire with a certain period of time in between each survey. This is in contrast to longitudinal cohort surveys that ask different respondents to fill out the same questionnaire at different time points. Panel surveys are more appropriate for testing causality in changes over time.

Pilot surveys: Surveys used to test the statistical properties (reliability and validity) of a questionnaire on a subsample of the population.

Performance test: A direct assessment of digital skills in performing authentic tasks (on a digital device).

Validity of a skill survey instrument: The extent to which the questions, answer scales and items in a questionnaire measure the full conceptualisation of the skills researchers intend them to measure.

Reliability of a skill-related survey item: The extent to which an item has the same statistical properties when it is measured at different times, as part of a varying series of items measuring skills, and as part of a different survey.

(Latent variable model) **Equivalence** testing: Testing, through latent variable models, whether the statistical properties of measurements (e.g. factor structures) in different groups are comparable; groups can be countries, gender, education or other groups of individuals.

1 Introduction

1.1 The ySKILLS project

The ySKILLS (Youth Skills) project is funded by the European Union (EU's) Horizon 2020 programme. It involves 15 partners from 13 countries to enhance and maximise the long-term positive impact of the information and communications technology (ICT) environment on multiple aspects of wellbeing for children and young people by stimulating resilience through the enhancement of digital skills. Starting from the view that children are **active agents in their own development**, ySKILLS examines how digital skills mediate the risks and opportunities related to ICT use by 12- to 17-year-olds in Europe (see https://yskills.eu).

The overarching aim of ySKILLS

To enhance and maximise the long-term positive impact of the ICT environment on multiple aspects of wellbeing for all children by stimulating resilience through the enhancement of digital skills.

ySKILLS will **identify the actors and factors** that undermine or can promote **children's wellbeing** in a digital age. The relations between ICT use and wellbeing will be critically and empirically examined over time.

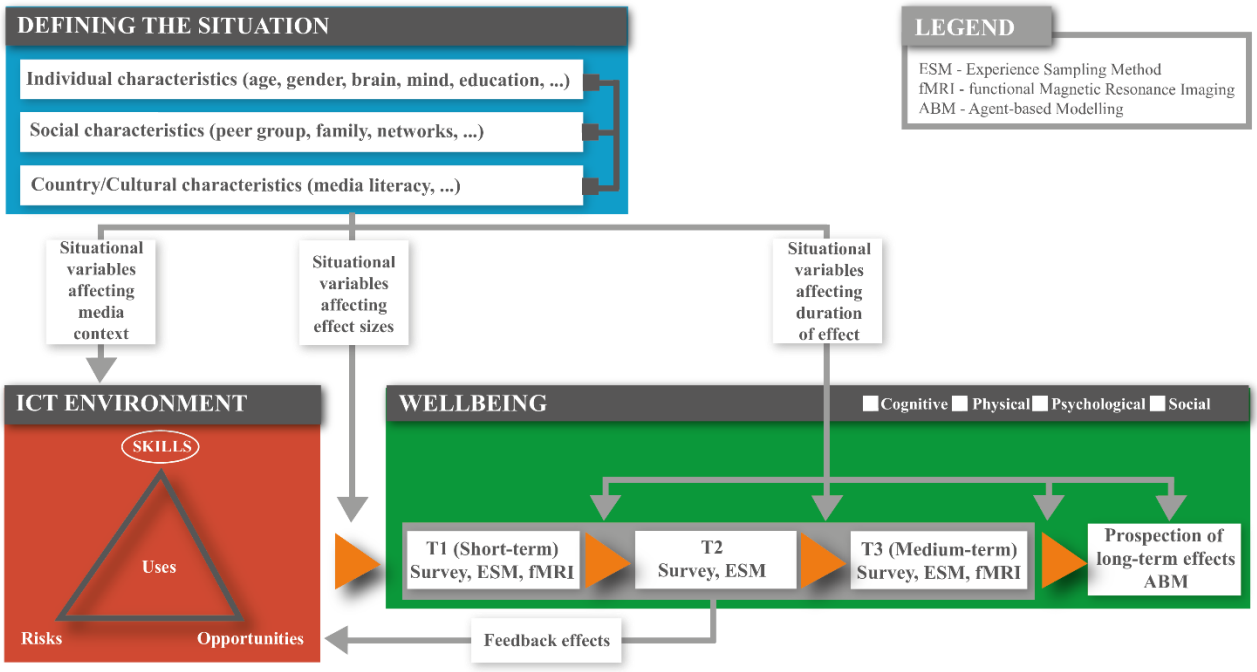
ySKILLS' research objectives

1. To acquire extensive knowledge and better measurement of digital skills.
2. To develop and test an innovative, evidence-based explanatory and foresight model predicting the complex impacts of ICT use and digital skills on children's cognitive, physical, psychological and social wellbeing.
3. To explain how at-risk children (as regards their mental health, ethnic or cultural origin, socioeconomic status and gender) can benefit from online opportunities despite their risk factors (material, social, psychological).
4. To generate insightful evidence-based recommendations and strategies for key stakeholder groups in order to promote European children's digital skills and wellbeing.

This report contributes to achieving objective 1 by reporting on the creation of the youth Digital Skills Indicator (yDSI), a validated survey instrument for the measurement of digital skills amongst youth.

ySKILLS has proposed, and will continue to develop, its **conceptual model** (see Figure 1):

Figure 1. ySKILLS CONCEPTUAL MODEL



This report focuses on the bottom left element of the ySKILLS project – the **conceptualisation and measurement of digital skills** as part of young people's ICT environment. In this model individual, social and country characteristics are antecedents of digital skills and the hypothesised consequences of digital skills are differences in cognitive, physical, psychological and social wellbeing. Haddon et al. (2020) reviewed studies in terms of how existing research hypothesises and measures these relationships with digital skills. They concluded that some elements of digital skills were not measured or that it was unclear whether the measures were of sufficient quality to measure actual skills.

This report focuses purely on the conceptualisation and measurement of skills. For the purposes of this report, other aspects related to the broader concept of digital literacy, such as ICT-related attitudes and confidence, are classified as separate antecedents of ICT use and not conceptualised as digital skills (Haddon et al., 2020). The yDSI presented in this report will be used in the ySKILLS three-wave longitudinal panel survey with 12- to 17-year olds (see https://yskills.eu), but can be adopted by other projects with young people and adults of different generations.

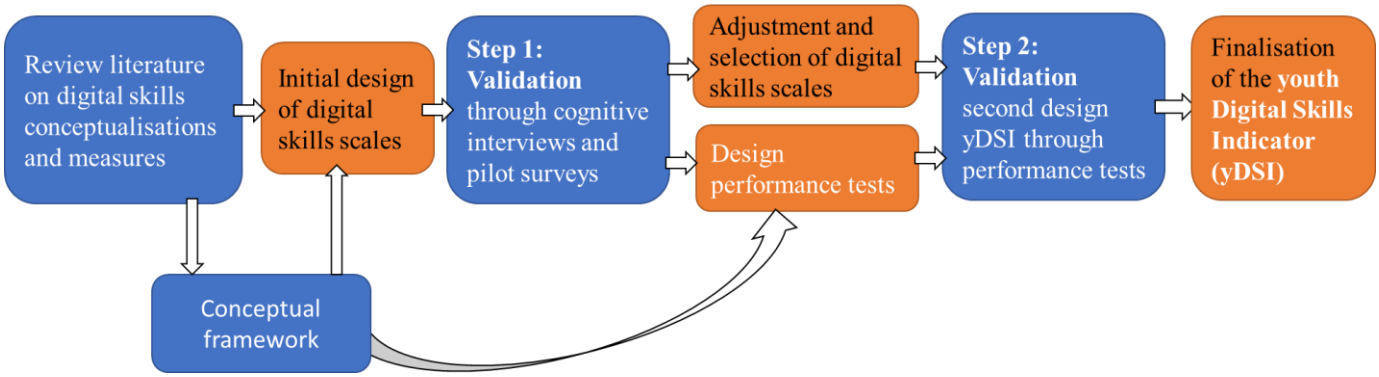
1.2 This report

It has become increasingly clear that specific groups of young people, for example, the psychologically vulnerable and traditionally marginalised, may be less likely to be able to take advantage of online opportunities. Even more concerning is that they may also be less able to avoid more negative outcomes (Haddon et al., 2020). Fundamental in this unequal distribution of outcomes is the unequal distribution of digital skills. Many young people lack advanced and sometimes even basic digital skills, which impedes their participation in increasingly digital societies.

The following definition for digital skills is used in this report: the ability to use ICTs in ways that help individuals to achieve beneficial, high-quality outcomes in everyday life for themselves and for others, and to reduce potential harm associated with more negative aspects of digital engagement (Helsper, 2021; Helsper & van Deursen, 2018). This definition was the starting point for the review of the literature and the design of the ySKILLS measurement tool.

This report describes the development of a high-quality, cross-culturally validated measurement instrument for these skills amongst young people, because so far, such an instrument has been as good as non-existent (see https://yskills.eu). Figure 2 summarises the procedures followed leading to the yDSI.

Figure 2. VISUALISATION OF THE STEPS TAKEN IN THE DEVELOPMENT OF THE yDSI



The report starts with a brief review of the academic literature from which a conceptual framework for digital skills was derived and which led to a review of the existing skills measures. Incorporated in this review is a discussion of best practice principles and common errors (or "sins") in good survey measurement design in relation to skills and competencies in general. This review of the literature, and the conceptual framework derived from it, has been the base of the survey measures for digital skills. The review of digital skills and their measurement also informed the methodology used to validate the developed measures. This report discusses the results of the cognitive interviews and pilot surveys used to validate and revise the initially proposed scales, describes the validation of this revised version through performance tests, and proposes the final version of the skills scales.

The current report starts with the presentation of the validated measurement tool for youth digital skills, the yDSI, before providing a detailed overview of how it was designed. The yDSI instrument is available in the six languages of the ySKILLS survey partners who will be using it for the panel survey (i.e., Estonian, Finnish, German, Italian, Polish and Portuguese). It is also available in Dutch and English where the skills scales were piloted. This report includes the English version; the other versions are available on the ySKILLS website (https://yskills.eu).

2 Final short version of the youth Digital Skills Indicator (yDSI)

This section presents the final short version of the yDSI indicator that was the outcome of the process presented in Figure 2. The yDSI is presented at the beginning so that researchers and others who are mostly interested in using the instrument itself can access it easily. The rest of the report describes in detail the justification for the conceptual framework that underpins the skills dimensions of the instrument (Sections 3 and 4), and the questions, answer scales and item formulation (Sections 4–7).

No such instrument, measuring a comprehensive set of skills for young people, has been designed and validated before. Future research should further validate the instrument based on research with populations of young people from different countries and different sociodemographic groups. The steps for validation and principles for design of indicators set out in this report serve as a guide for best practice in these matters.

Two different types of questions and associated answer scales were formulated as part of the yDSI: one for digital skills and one for digital knowledge items. These two types of measures were developed to be able to capture all elements of functional (the ability to use ICTs) and critical (understanding the ways in which ICTs are designed and content is produced) digital skills. Some elements can be captured in valid ways through self-reports and others through items that test knowledge. In general, functional skills are reasonably well measured through self-reports, while critical skills are better measured through knowledge items (see Section 4).

2.1 Question and answer scale formulation of the digital skills items

The phrasing of the question asking about digital skills is as important as the items that measure young people's skills. To guarantee the highest possible validity and reliability of the measurement instrument, the wording of the question should be followed exactly or, in translations, the closest equivalent in meaning should be used.

> Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones or computers. Reply thinking about **how true this would be of you** if you had to do it now, on your own. **If you do not understand** what the question is asking, tick the box "I do not understand what you mean by this".

The following is added to the instructions for the communication and interaction and the content creation and production skills item blocks only (see Table 1).

> Sometimes there are various examples given; only select "Very true of me" if all of the examples apply to what you do or know.

The formulation and scoring of the answer categories are equally important. The answer scale was carefully piloted and should be used in its entirety to avoid social desirability and skewness issues.

The answer scale used for the digital skills items is:

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	I do not want to answer (99)

Note I: The answer categories should be presented in this order and the scores on the Likert scale (1 through 5, 66 and 99) should not be presented to the respondents; these are only included for coding and analyses.
Note II: For analyses, the "I do not understand what you mean by this" answer category is part of the skills scale and should be converted to zero because it indicates a lack of knowledge as well as a lack of skill and thus ranks below not having a skill (see Section 8.3 for guidelines on how to create and use composite scales in analyses).

2.2 Youth Digital Skills Indicator (yDSI) – digital skills items

There are 25 items on the yDSI (see Table 1) representing the four dimensions of digital skills as established in the conceptual model for digital skills (see Section 3.2).

Table 1.	FINAL yDSI DIGITAL SKILLS SURVEY ITEMS
Dimension	Item
Technical and operational kills	I know how to adjust privacy settings
	I know how to turn off the location settings on mobile devices
	I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition)
	I know how to store photos, documents or other files in the cloud (e.g. Google Drive, iCloud)
	I know how to use private browsing
	I know how to block unwanted pop-up messages or ads
Programming	I know how to use programming language (e.g. XML, Python)[a]
Information navigation and processing[b]	I know how to choose the best keywords for online searches
	I know how to find a website I have visited before
	I know how to find information on a website no matter how it is designed
	I know how to use advanced search functions in search engines
	I know how to check if the information I find online is true
	I know how to figure out if a website can be trusted
Communication and interaction	Depending on the situation, I know which medium or tool to use to communicate with someone (e.g. make a call, send a WhatsApp message, send an email)
	I know when I should mute myself or disable video in online interactions
	I know which images and information of me it is OK to share online
	I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters
	I know how to report negative content relating to me or a group to which I belong
	I know how to recognise when someone is being bullied online
Content creation and production	I know how to create something that combines different digital media (e.g. photos, music, videos, GIFs)
	I know how to edit existing digital images, music and videos
	I know how to ensure that many people will see what I put online
	I know how to change the things I put online depending on how other people react to it
	I know how to distinguish sponsored and non-sponsored content online (e.g. in a video, in a social media post)
	I know how to reference and use content covered by copyright

Notes: [a] Programming is included as a single item; it does not load onto the skills dimensions as the other items do, but is considered important in the literature and interventions and is thus included.

[b] See Section 8.1 for a comment on the statistical properties of the items on this dimension and how they should be used in future research.

2.3 Question and answer scale formulation of the digital knowledge items

Besides the questions that measure digital skills across four dimensions, questions were formulated that measure knowledge about and critical understanding of ICTs. To guarantee the highest possible validity and reliability of the measurement instrument, the wording of these questions and the answer scale should be followed exactly or, in translations, approximated as closely as possible.

> To what extent are the following statements about technologies such as the internet and mobile phones **true or not true**? If you are not sure, please let us know.

This question is accompanied by the following answer scale:

Definitely not true (1)	Definitely true (2)	I'm not sure (3)	I do not want to answer (99)

Note: The answer categories should be presented in this order and the scores on the scale (1 through 3 and 99) should not be presented to the participants; these are only included for coding and analyses.

2.4 Youth Digital Skills Indicator (yDSI) – digital knowledge items

There are six items on the yDSI (see Table 2), distributed along three dimensions, as set out in the conceptual framework.

Table 2.	**FINAL yDSI DIGITAL SKILLS SURVEY ITEMS**
Dimension	Item
Information navigation and processing	The first search result is always the best information source
	Everyone gets the same information when they search for things online
Communication and interaction	The first post I see on social media is the last thing that was posted by one of my contacts
	Whether I like or share a post can have a negative impact on others
Content creation and production	Using hashtags increases the visibility of a post
	Companies pay ordinary people to use their products in videos and content they create

Note: See Section 8.3 for a comment on how these items should be used in future research.

3 Review of the digital skills literature

To get to the final design of the skills measures as presented above, several steps were taken. The first was a review of the academic and grey (i.e., non-academic) literature on digital skills. This led to the conceptual framework that underpinned the design of the yDSI. This review of the literature also had the purpose of examining best practice in the design of measures for large-scale population studies. Therefore, it had a broad scope including studies that measured skills for adults as well as young people. After briefly reviewing this literature, the report looks at measurement instruments used in academic studies. The review of the grey literature consists of mapping the measures used onto the pre-established conceptual framework.

3.1 Academic literature: Conceptual frameworks for digital skills

After an initial focus on a lack of (quality) access to ICTs, researchers and policy-makers shifted their attention to a lack of digital skills as an important obstacle to obtaining the benefits and avoiding the potential harm associated with use of ICTs (van Dijk & van Deursen, 2014). A considerable body of work has sprung up in the last since 2010 that defines and measures digital skills. Early on, digital self-efficacy (the confidence people have in themselves as users) and skills (what people are able to do) were distinguished as two separate aspects of digital literacy. This distinction is important because they are related to different aspects of engagement with ICTs and potential benefits derived from this engagement. Research has tied self-efficacy to a higher propensity to take risks online as well as offline, trying things out, and learning through trial and error (Durndell & Haag, 2002; Gecas, 1989; Huang, Cotten, & Rikard, 2017; Scherer, Rohatgi, & Hatlevik, 2017; Shank & Cotten, 2014).

Higher skill levels, on the other hand, have been shown to be related to greater achievement of positive outcomes and avoidance of negative outcomes of internet use (van Deursen, 2020; van Deursen & Helsper, 2017; van Deursen et al., 2017). What is interesting is that while improvements in skills are often related to higher self-efficacy, the opposite is not necessarily true. In fact, it appears that when high self-efficacy is coupled with low level of skill, this might actually get in the way of people improving their skill levels, as it prevents them from understanding what they do wrong and causes them to attribute failures to factors outside of their control (Broos & Roe, 2006). Digital self-efficacy and confidence are unequally distributed along the lines of those who have more dominant positions in society even when this is not reflected in a difference in skill levels (Haddon et al., 2020). While digital self-efficacy is important for informal learning about, and broad engagement with, ICTs, digital skills are more important in terms of positive participation and wellbeing in society. This is one reason why the ySKILLS project takes skills and not confidence in one's digital abilities as its pivot in explaining participation in digital societies (Haddon et al., 2020).

Early conceptualisations of digital skills often saw them as one-dimensional and focused mostly on technical skills such as installing software, operating a device or programming. This was then labelled "computer literacy" (Richter, Naumann, & Groeben, 2001; Robinson & Thoms, 2001; Selber, 2004). Increasingly, with the broader diffusion of the internet, information navigation became part of the definition of digital skills (Bawden, 2001; Kolle, 2017; Saranto & Hovenga, 2004). These two elements are still part of almost all academic research. With the rise of interactive web2.0, these definitions have been expanded further.

There is currently relative consensus in the academic literature that digital skills can be broadly categorised along a range of dimensions that include the broad categories of technical and operational, information, social and content creation skills (Helsper, 2021; Helsper & van Deursen, 2018; van Deursen, Helsper, & Eynon, 2016). Aspects such as problem-solving and

safety are frequently conceptualised as separate skill domains (van Laar et al., 2017). For the purposes of this report and the construction of the yDSI, these were denominated outcomes of high skill levels. This decision was made because solving a problem can be seen as a positive outcome of skilled use rather than a skill in and off itself, and because safety can be denominated as the prevention of negative outcomes achieved through skilled use.

The definition of digital skills has, furthermore, expanded to include not only practical, functional skills common in the policy and computer science discourse, but also critical, evaluative skills that are more commonly referred to in the pedagogic and media literacy literature. Both functional and critical components of digital skills are essential to take into account when incorporating digital skills into the broader study of wellbeing in increasingly digital societies (Cortesi et al., 2020). Even though these two skills aspects are often hard to disentangle in practice (Haddon et al., 2020; Polizzi, 2020b), they should be disentangled in conceptualisations. This is because having just functional skills (understanding the functionalities of ICTs and being able to use them) is associated with more passive, consumptive participation in digital societies while critical skills (understanding how and why technologies are designed and certain content is produced in particular ways) are essential for more active, constructive participation in society.

3.2 Conceptual framework for digital skills based on the academic literature

The described review of the academic literature, previous reviews for the ITU (Helsper & van Deursen, 2018) and the From Digital Skills to Tangible Outcomes (DiSTO) projects (van Deursen, Helsper, & Eynon, 2016) led to the establishment of a conceptual model for the development of measures that correspond to the four digital skill dimensions, reflecting functional and critical aspects in each dimension: (1) technical and operational skills; (2) information navigation and processing skills; (3) communication and interaction skills; and (4) content creation and production skills (see Figure 3).

These four different skills dimensions can be defined as follows:

Technical and operational: the ability to manage and operate ICTs and the technical affordances of devices, platforms and apps, from 'button' knowledge to settings management to programming.

Information navigation and processing: the ability to find, select and critically evaluate digital sources of information.

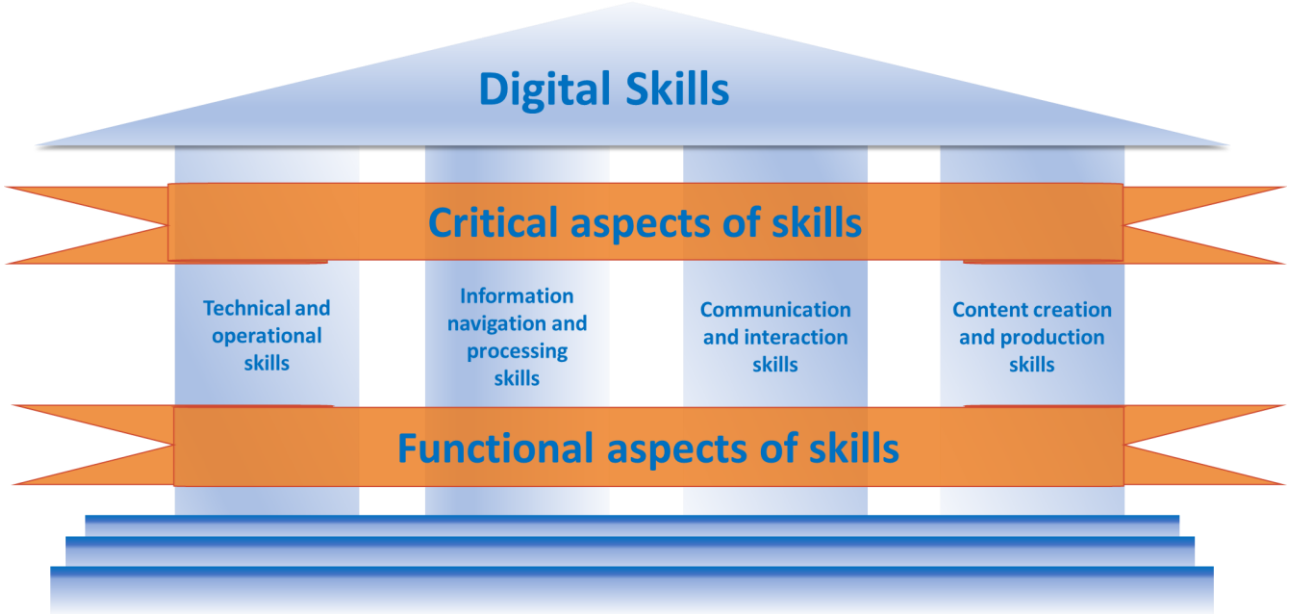
Communication and interaction: the ability to use different digital media and technological features to interact with others and build networks as well as to critically evaluate the impact of interpersonal mediated communication and interactions on others.

Content creation and production: the ability to create (quality) digital content and understand how it is produced and published and how it generates impact.

There are two important further assumptions that underpin this conceptualisation of skills along four dimensions:

- All skills dimensions have functional (understanding technical functionalities and being able to use these) and critical (understanding how and why devices and content are produced in certain ways) aspects.
- In order to participate fully in digital societies, being skilled in all four dimensions is indispensable.

Figure 3. CONCEPTUAL MODEL OF DIGITAL SKILLS AND ITS UNDERLYING FOUR DIMENSIONS AND FUNCTIONAL AND CRITICAL ASPECTS



4 Literature on (digital) skills measurement

One of the reasons that the measurement of digital skills is under increased scrutiny is that knowledge of the level of digital skills is fundamental in the development, implementation and evaluation of policy and interventions. However, to empirically establish what the level of digital skills is requires costly and resource-intensive research methods. This section describes different ways of measuring digital skills and best practice in terms of item and answer scale design.

4.1 Types of methodologies used to measure digital skills

There is a range of methods used to measure digital skills, from indirect measures to performance tests. As this section shows, while there are clear limitations to self-reporting, if they are designed and validated properly, survey instruments are a good alternative for large-scale skills testing, and they are the most cost-effective, least resource-intensive way of doing this.

Indirect measurement: Various large benchmark studies use questionnaires in which respondents are asked which activities they have ever undertaken online. The level of digital skills is subsequently derived from the number of ways in which the person has been digitally engaged. However, the undertaking of an activity does not mean that it is undertaken at a highly competent level, and not undertaking an activity does not mean that the respondent does not have the skills to undertake this activity (Haddon et al., 2020). Other abstract, even more indirect measurements look at the years of schooling, assuming a link between traditional and digital literacy (see, for example, ITU, 2017). There are, indeed, correlations between use and skills and between traditional literacy and digital literacy, although these do not properly reflect actual skill levels. It is possible to conduct large-scale measurements using these proxy indicators, but since there is no observation of actual skills, these indirect measurements do not give a valid representation of digital literacy levels.

Self-assessment: This is the most used method to measure digital skills in the general population. Respondents are asked to evaluate how good they are in relation to a range of skills. This type of research uses answer scales ranging from “very bad” or “novice” to “very good” or “expert”. The advantages of this method are that many questions can be asked in a relatively short space of time, scoring of skill levels is simple, and processing of data is rapid and cost-effective. The disadvantage is that self-assessments are limited in their ability to indicate real competence levels for two main reasons. The first is that it is difficult for people to assess their own skill level, partly because this depends on the comparison group that people use to estimate these skill levels and their own norms about what satisfactory levels are (Herde et al., 2019; Spenner, 1990). The second reason is that self-assessments are subject to social desirability bias, as people often do not want to admit that they are not good at something (Grimm, 2010; King & Bruner, 2000). These kinds of validity issues make self-assessment a limited predictor of actual skills. Some of these issues can be mitigated through good survey design and validation with cognitive interviews, but the problem remains that it is unclear how self-assessment measures relate to actual skills. There is some evidence that skills measured through these non-externally validated self-reports do not predict differences in performance on ICT-related tasks (van Deursen, van Dijk, & Peters, 2012).

Performance tests and observations: This is the most valid way to get a realistic image of a person’s digital competency level. This method is very labour intensive in its development as well as in its implementation and scoring (van Deursen & van Dijk, 2010). The costs of these

types of measurements make them prohibitive for data collection on a large scale. One type of performance test is the interactive standardised test, which is conducted in closed environments where a participant completes tasks in a virtual simulation that replicates real-life digital environments (e.g. software, browser or app). These types of tests are most appropriate for formal educational or training settings. Tasks used for performance tests are, by definition, very specific to the context in which they are applied and cannot measure a broad range of skills. They are often designed for specific professional fields or for specific areas of pedagogy or education. Recent research shows that these are easier to design for functional technical and operational and information navigation skills than for context and relationship-dependent communication and interaction and content creation and production skills. Because of their dependence on context, they are less suitable for large-scale population research.

Self-assessments validated through performance tests: This method involves externally validating the answers of a self-assessment instrument by comparing them to how the respondents perform on a series of tasks. The propositions with the best correspondence to performance on these tests are then used to do large-scale population research. This improves the quality of the self-assessment instruments considerably. The gold standard for large-scale population research would be to use validated longitudinal panel survey design to measure the links between digital skills and participation and wellbeing in digital societies. This is the route followed for the ySKILLS project, as detailed in the rest of this report.

In choosing a measurement instrument a variety of factors have to be taken into account, such as achievability, implementation, reach, reporting and costs as well as the outcomes that are desired for a particular context. Considering the costs and benefits, externally validated self-assessment in a longitudinal panel design is the preferred option for measuring digital skills in larger populations, especially if one needs to know how these levels vary between different groups and what the consequences of differences in skills are.

4.2 Best practice guidelines for large-scale population research on digital skills

An important part of the ySKILLS project consists of a longitudinal study that aims to measure the antecedents and wellbeing consequences of various levels and types of digital engagement and skills (see Section 1.1 and also Haddon et al., 2020). Since this is a large-scale population study where a multitude of factors needs to be measured at three points in time, the most appropriate data collection tool for this is a survey. Therefore, this section goes deeper into the design of measures that appropriately capture the various components of digital skills.

This section lists seven common errors (or “sins”) that should be avoided in digital skills questionnaire item design and outlines the seven most important best practice guidelines for the formulation of items and question and answer scales in measuring digital skills. In doing so, several challenges with existing measures are addressed that hinder linking the results of survey measurements to theoretical conceptualisations of digital skills.

4.2.1 Common errors in digital skills item design

In the context of the ySKILLS project, items and answer scales were designed to be about skills in relation to connected technologies (mobile phones and computers that can access the internet). Previous reviews of the literature (Helsper & van Deursen, 2018; van Deursen et al., 2016) suggest the following seven common "sins" in survey items designed to measure skills related specifically to these connected technologies:

1. They constitute **general bad survey item design** (e.g. double-barrelled two-in-one questions, leading questions, the use of jargon).
2. They are **PC**-based (e.g. keyboard shortcuts, spreadsheets) and are not related to connected technologies.
3. They are too **vague or general** (e.g. "Can you prevent and respond to risk using digital skills and media literacy?", "Are you able to manage your use of technology, taking full advantage of technology while avoiding excessive time online and addiction?").
4. They are about **outcomes** (e.g. "Have you seen fake news?", "Are you able to prevent getting bullied?") instead of skills.
5. They are about **use** (e.g. designing a website or installed virus scans) and not about skills (e.g. "Can you?", "Do you know how to?")instead of skills.
6. They are about **attitudes** (e.g. "Are you open to or excited about trying out new technologies?") instead of skills.
7. They are about **confidence** (e.g. "How good are you at using social media?") instead of skills.

All seven problematic practices listed are present in digital skills item design and should be avoided. Many studies manage to do so, especially when it comes to the first point – that is, most studies use items that are phrased as precisely as possible using the everyday language of the target population (including current examples), and they ask about a single skill at a time (i.e., avoid double-barrelled items). However, there are a few common errors, especially as regards "sins" 4–7, that are persistent, and make many digital skills survey instruments inadequate.

4.2.2 Best practice guidelines for skill item survey design

This section describes seven key best practice guidelines for valid and reliable instrument design of digital skills indicators in survey research and the reasoning behind them. These are, to some extent, related to the common seven "sins" mentioned. However, they go beyond this in that they make suggestions for best practice that are mostly absent in the existing research.

Existing items often measure use of ICTs rather than actual digital skills. However, use of digital tools does not necessarily denote high levels of digital skills, and vice versa (Haddon et al., 2020; Looker & Naylor, 2010). Despite the obvious correlation between skills and use, there are important differences between asking how often people do something and whether they know how to do it (e.g. most people know how to turn off their phone, but many also report rarely doing so; see van Deursen et al., 2014).

Best practice guideline 1: Items should ask about whether participants possess a certain digital skill (e.g. "can do" or "know how to do") rather than about usage (e.g. "have you?" or "do you?").

Second, not only should PC-focused items be avoided when studying connected technologies (see "sin" 2), app- or activity-specific items should be avoided, since such items are too dependent on current social and technological trends that lead participants to adopt or divert from using certain devices or online platforms (van Deursen et al., 2016). These skills related to specific apps or activities may not always be transferrable to the next popular digital tools. For instance, skills that are specific to certain social media platforms (e.g. managing the privacy settings of Instagram) do not transfer to other platforms (e.g. managing privacy settings on Facebook, WhatsApp and LinkedIn). Another example would be knowing what the norms of interaction are in gaming settings but not understanding how these work in online learning environments.

Best practice guideline 2: Avoid device-, app- or activity-specific items to make sure that skills items measure transferrable skills and can continue to be used over time.

In some instances, items designed according to these best practice guidelines are not suitable. Functional and practical skills (e.g. "What can you do?" and "How can you apply this skill?") need different formulations than critical and knowledge-based items (e.g. What do you understand?" and "What are the implications of actions?"). For example, it is impossible to ask whether a person knows or understands something (e.g. "Do you know that a lock item means a website is safe?") since they acquire this knowledge when they are asked the question. This means these items measure the outcome of a skills process (e.g. acquiring knowledge, rather than knowledge itself). Similarly, asking someone about their ability to critically evaluate online interactions or content is difficult. For example, asking people whether they can recognise a certain type of bad behaviour (e.g. adjusting your behaviour to the circumstances or knowing when people have bad intentions or are trying to sell you something) may induce social desirability bias, as the question makes it clear which answer is "good" and which is "bad" (Larson & Bradshaw, 2017; Silvera, Martinussen, & Dahl, 2001). Such items may also be more indicative of confidence (see "sin" 7) rather than skills. These knowledge-based items are better formulated through statements to be evaluated as either true or false.

Best practice guideline 3: Different items should be designed to capture (functional) digital skills and (critical) digital knowledge.

Knowledge-based items based on true–false statements are widely used to test the (factual and critical) knowledge of participants in certain areas (Frisbie, 1973; Schmittlein & Morrison, 1983). In designing these items, a few additional guidelines should be considered. For one, people appear to suffer from "acquiescence" bias, meaning that they tend to guess items being true more often than false (Burton, 2004; Cronbach, 1941), suggesting that false statements should be presented to participants at least in equal proportion to true statements. In addition, evidence shows that negative marking improves the reliability of true–false statements (Burton, 2004, 2005), which also suggests that adding an "I am not sure" option allows for a reliable estimate of participants' level of knowledge and their critical digital skills on the basis of how they score on these knowledge-based items.

Best practice guideline 4: At least half of the skills items in a module that tests people's digital knowledge and critical skills should involve statements that are untrue or unlikely to be true.

4.2.3 Best practice guidelines for question and answer scale design

The formulation of the question is just as important as the phrasing of the items and answer scales to avoid context-related biases. This section contains best practice suggestions for question formulation and answer scale design. Since most surveys rely on self-reports, questions may be more reflective of confidence rather than skill, especially when they ask "How good are you at?" This can lead to individuals from advantaged backgrounds overrating and individuals from more disadvantaged backgrounds underrating their skill levels (Hinostroza et al., 2015; Huang et al., 2017; Shank & Cotten, 2014; Sorgo et al., 2017; Vekiri, 2010; Wamuyu, 2017).

Best practice guideline 5: Items should ask about whether participants possess a certain digital skill (e.g. "Do you know how to do X?") rather than how expert they are in relation to a certain skill (e.g. "How good are you at/rate yourself on the following?").

When asking people to rate their own skill levels, they often use a comparative framework (Herde et al., 2019; Smith & Pettigrew, 2015; van Deursen et al., 2016). That is, they think about how good they are in comparison to the people around them. This means that these items are dependent on the context the participant is in, who they are surrounded by and who they choose to compare themselves to (Helsper, 2017). To avoid this, truth claims that force people to look at themselves honestly in isolation are better (van Deursen et al., 2016). This is done by, for example, asking them to say how true a certain statement is about the way they act ("I know how to do XYZ" with answer options "Not true of me", "Somewhat true of me", "Definitely true of me"), or how much a statement applies to them if they would have to do something now without help (van Deursen et al., 2016).

Best practice guideline 6: Question formulation should be phrased in a way that makes the person evaluate their current personal skills and not their skills in comparison to others (i.e., using truth claims and emphasising the here and now).

The way in which answer options are phrased is crucial to avoid social desirability or "confidence" bias, as well as to allow participants who do not or have not yet use(d) the relevant digital skill to answer the question. For this purpose, including a "I do not understand what this means" option tackles issues around bias, while also making the distinction between participants who know how to do something without having done it and participants who do not know at all (van Deursen et al., 2016). Similarly, using scale rather than dichotomous answer options (e.g. "I know how to do this"/"I do not know how to do this") mitigates the social desirability bias. This is because giving people a large range of options normalises a lack of skill and facilitates respondents admitting that they do not really know how to do something.

Best practice guideline 7: Answer options need to be scale-based and include an option indicating that a lack of skill or understanding is perfectly okay and normal to avoid social desirability bias.

The next section looks at the extent to which the conceptual framework is reflected in the academic and grey literature on measuring digital skills, and the extent to which these measures presented in this literature demonstrate the digital skills measurement best practice and problematic practices described above.

4.3 Selection of academic and grey literature measuring digital skills

The review of the literature on measures was a targeted review of published research that reports on survey and performance test measures. Its main aim was to collect high-quality measures adhering to the best practice guidelines. What follows is a description of the procedures for the review. The items selected and designed for the initial round of validation followed the seven best practice guidelines for item, question and answer scale design and avoided the common errors as specified in the previous section.

4.3.1 Academic literature item selection

The studies identified by Haddon et al. (2020) served as the basis for the review of measures in the academic literature. Out of the 322 *sources* they identified, all 66 publications were reviewed that scored 3 out of 3 on the "review-specific appropriateness of method and form of evidence". This included an evaluation of the appropriateness of the measures and analysis of digital skills. Each of the 66 papers was classified according to whether it used previously existing data or frameworks, which demographic it targeted and what aspects of digital skills it tested. Further searches were conducted to determine whether items were available online or in accessible publications, and whether questionnaires had been checked for reliability and validity. In addition to these papers, a selection of papers that scored 2 out of 3 in this same category was assessed. Publications with a lower rating were not explored for lack of quality.

The initial total of *items* collected in the academic literature was 428. The selection of items was then narrowed down to account for reliability and validity, as well as by deleting items whose formats (e.g. multiple choice questions) did not fit the requirements for the type of survey items that needed to be designed for the ySKILLS project. A total of 269 items were left to consider following this sorting. These items were subsequently categorised according to the skills dimension conceptualisation for the project (see Figure 3); (1) technical and operational skills; (2) information navigation and processing skills; (3) communication and interaction skills; and (4) content creation and production skills.

4.3.2 Grey literature item selection

A similar selection procedure was followed to get to the most relevant skills measurement tools in the grey, non-academic literature. Here the selection process focused on clearly reported measurement tools being developed or used to evaluate digital skill levels and for which data was available or published. For those studies, the underpinning conceptual frameworks of skills were examined in terms of how these mapped onto the academic conceptualisations in the four dimensions. Cortesi et al.'s (2020) comprehensive report was taken as a starting point, and

ySKILLS partners added a few other international studies that were left out of that review but had a significant impact in Europe and were highlighted by ySKILLS partners and experts in the field as relevant in the context of this project on youth digital skills. This meant that all 35 studies and initiatives discussed in the Cortesi et al. (2020) report were included in the review of measures, complemented by 14 additional reports.

Some of the selected digital skills measures, such as those from DigComp (European Commission, 2020a) and Global Kids Online (Global Kids Online, 2020), were based on academic studies, involved collaborations with academics, or had academic publications based on the data they gathered. However, since their application was primarily in non-academic contexts, they were included as part of the grey literature. The review of the grey literature examined mostly large-scale, international, comparative instruments. There are many local government initiatives and smaller non-governmental organisations (NGOs) active in this field that have set up interventions (see, for examples, MEDICI - European Commission, 2020b). However, since the measurement instruments they use are often not reported or, when they are available, do not include data on actual skill measurement, these are not considered. In addition, many of the local government initiatives and smaller interventions rely on the frameworks set up by larger organisations. Therefore, the general conclusions around measurement and the conceptual frameworks that underpin these smaller initiatives were captured by the review of the larger-scale studies.

All the available digital skills questionnaire items used in the reviewed studies were compiled, resulting in an initial list of 309 items. There was significant overlap between these items and the ones collected in the academic literature. Therefore, 86 items were dropped from the grey literature selection because they had been taken from academic studies that were already included in this review, resulting in 223 remaining items from the grey literature.

4.3.3 Selection of the best items from the academic and grey literature

Merging the academic and grey literature reviews on skills resulted in a total of 492 items. From this list, irrelevant (e.g. “When surfing the internet, I often catch myself saying: Just another few minutes. And then, however, I cannot stop”; see Walther, Hanewinkel, & Morgenstern, 2014), ambiguous, double-barrelled (e.g. “Chooses the data format that best supports the communication, distribution, and sharing of data and knowledge, taking into account the data size and type of users”; see Yoshida, 2018), or outdated items (e.g. “Use a computer to listen to music or watch DVDs”; see ACARA, 2020) were removed. This resulted in a total of 251 remaining “relevant” items. From this list, duplicate items were removed by keeping only one of the items if they were exactly the same in several studies, and choosing items from validated, cross-national studies over those in non-validated or national studies when there were small differences in phrasing or when they measured the same skill with larger differences in phrasing. A total of 136 “unique” items remained that complied with most of the criteria set out in the best practice review.

4.4 Findings: Measurement instruments used in digital skills testing

This section reviews how often the skill dimensions are measured, the extent to which best practice and common errors are represented in the skills items, and the ways in which survey measures have been validated and tested for reliability.

4.4.1 Representation of conceptualised skills dimensions in measurement instruments

The targeted review showed that many tests, instruments and policy documents focus on technical or operational skills, including some form of information navigation and processing skills. Skills related to recent web 2.0 activities were not always included, and even more rarely incorporated are the more critical literacy skills.

Table 3. NUMBER OF QUESTIONNAIRE ITEMS RETRIEVED AND SELECTED FROM THE LITERATURE ACCORDING TO DIMENSION

Skills dimension	*Total* items retrieved	N *relevant* items	N *unique* items
Technical and operational	217	101	55
Information navigation and processing	108	61	23
Communication and interaction	97	49	35
Content creation and production	70	40	23
Total	492	251	136

The literature offered a broad range of skills items for the technical and operational, information navigation and processing, and communication and interaction dimensions of digital skills (see Table 3). There was more diversity in the technical and operational dimension than in the information and interaction skills dimensions, more unique items measuring technical and operational skills and more consensus across studies in how to measure information navigation and processing and communication and interaction skills. However, items related to content creation and production skills were much less frequently included. In addition, when it was measured, items focused predominantly on functional and technical aspects of digital content creation (e.g. uploading or downloading photos). That is, the creation of quality content, understanding of dissemination and a critical view of production and consumption were largely absent. Programming was measured across most studies but seemed to refer more to a technical and operational skill than to a content creation and production skill. While three of the four dimensions used in the current conceptualisation of digital skills were common in the literature, studies in the grey literature in particular tended to focus on only one or two of these (Eurostat, 2019; PISA OECD, 2020a; PIAAC OECD, 2020b). The same can be said for a large part of the academic literature (Gui & Argentin, 2011; Lazonder et al., 2020; Li & Ranieri, 2010; Mason et al., 2018; Nygren & Guath, 2019), although it was much more common for academic literature to discuss at least three of the four digital skills dimensions (Kim, Kil, & Shin, 2014; Le et al., 2019; Rodríguez-de-Dios, van Oosten, & Igartua, 2018).

Several studies did not explicitly refer to frameworks used to underpin the measurement of skills. Their items tended to be bundled together rather than separated according to dimensions. These items were placed into the relevant dimensions following internal discussions and considered as part of the selection process (ACARA 2020; ICILS - National Center for Education Statistics, 2018).

In the academic and grey literature, one dimension of digital skills that was part of a dozen studies, but was not included in the ySKILLS conceptualisation, was problem-solving

(Digcomp - European Commission, 2020a; Kim & Lee, 2013; Siddiq et al., 2016; van Laar et al., 2020). Whilst present in many frameworks, the review showed that measures of this particular dimension of digital skills drew on elements of the other dimensions, or measured an outcome of use rather than a skill. Therefore, problem-solving items complying with best practice guidelines were incorporated into the different dimensions.

Some of the grey literature did provide a clear framework and considered all four dimensions of digital skills but did not measure these skills, as these were intended as self-assessment tools meant to guide training and self-improvement among individual users (DigComp - European Commission, 2020a). A number of studies, including these, phrased their questions in terms of self-efficacy or usage rather than as direct measures of skills (Aesaert & van Braak, 2014; Aesaert et al., 2015, 2017; Areepattamannil & Khine, 2017; ICLS - European Commission, 2019; Lau & Yuen, 2015). These items were considered in terms of their conceptual relevance, especially to inspire the design of new items in areas where there was a lack of items that complied with the best practice criteria, such as content creation and production. When used, these items were rephrased to comply with the best practice criteria and to avoid the common "sins" in digital skills item design.

In conclusion, while many of these tools and questionnaires provided valid options for consideration in the present study, few employed both the established conceptual framework and measures that could be included in the digital skills measures used for this project without adjustments. Items from the studies that did comply with the different projects (DiSTO, 2020; Global Kids Online, 2020; Net Children Go Mobile, 2020) were prioritised in the selection process for the yDSI, particularly when it came to technical and operational, information navigation and processing, and communication and interaction skills.

4.4.2 Presence of best and problematic practices in the skills measurement literature

Few studies manage to avoid all the outlined seven "sins" in the design of their digital skills items (Balea, 2016; DiSTO, 2020; Global Kids Online, 2020; Ponte, 2019). Table 4 shows how often each "sin" was committed across the studies that were considered in the selection and design of the current digital skills measures.

Table 4. NUMBER OF STUDIES THAT SHOW PROBLEMATIC PRACTICES IN DIGITAL SKILLS SURVEY ITEM DESIGN

"Sins"	Number of studies
Bad survey design (e.g. double-barrelled, leading, jargon, etc.)	13
Purely PC-based (e.g. keyboard shortcuts, spreadsheets)	8
Too vague or general	17
About outcomes instead of skills	9
About use and not about skills	5
About attitudes	7
About confidence	13
Studies that did not commit any "sins"	5
Total number of studies	53

The most common shortcoming was to include items that were too vague or general. Examples of this are items such as: "accessing information with a computer" (Areepattamannil & Khine, 2017), "maintaining social relationships [online]" (Kaarakainen, 2019) or "interpreting and

representing information, such as using ICT to synthesize, summarize, compare and contrast information from multiple sources" (Lau & Yuen, 2015). Sometimes more than one "sin" was committed. For example, the first mentioned in the list of examples above is purely PC-based in addition to being too general, the second also relates to use, and the third is also double-barrelled. Another commonly committed "sin" is the phrasing of items in terms of confidence rather than skills (e.g. "How good are you at sending a polite email?"; see Aesaert & van Braak, 2014).

Similar to the occurrence of "sins", the extent to which best practice guidelines have been incorporated into existing empirical research on digital skills varies greatly. Most studies adhere to the first best practice guideline ("Items should ask about whether participants possess a certain digital skill rather than about usage"); only five ask about use as an indicator of skills (see, for example, Alkan & Meinck, 2016). However, as many as 13 did not comply with the fifth best practice guideline and ask about whether participants possess a certain digital skill rather than how expert they are in relation to a certain skill (see, for example, Gastelu, 2013; Moto et al., 2018; OECD, 2020a; Williams-Diehm et al., 2018). As a result, while many existing questionnaires get elements of question formulation right, very few actually phrase their questions and items in the desired way (see, for example, Balea, 2016; Lau & Yuen, 2015; Ponte, 2019).

The second best practice guideline ("Avoid device-, platform- or activity-specific items to make sure that skills items measure transferrable skills and can continue to be used over time") is applied on a more widespread basis (Alkan & Meinck, 2016; Kaarakainen, 2019; Porat, Blau, & Barak, 2018), although some studies still include device-specific items, for example, mobile phone-related items, here and there (Areepattamannil & Khine, 2017; Lee, 2018; Rodríguez-de-Dios, Igartua, & González-Vázquez, 2016). Similarly, the third best practice guideline ("Different items should be designed to capture (functional) digital skills and (critical) digital knowledge") is rarely violated, in the sense that very few studies combine functional and critical elements in the same item. That being said, many studies focus exclusively on functional skills (ACARA, 2020; Aesaert et al., 2017; Balea, 2016; ICILS - European Commission, 2019; Zhong, 2011), thus overlooking the important critical aspect of digital skills. On the other hand, a few studies include a mix of both functional and critical items (Lau & Yuen, 2015; Le et al., 2019; Yoshida, 2018).

Finally, guidelines around question and scale design are applied sporadically. While most studies use scale-based answering options (Gastelu, 2013; Moto et al., 2018; Rodríguez-de-Dios et al., 2016; Tondeur et al., 2011), these options are rarely phrased as truth claims (Christoph et al., 2015; Lee, 2018), and many studies fail to include an option for participants who do not know what the question asks. Even so, studies that incorporate this option tend to explicitly ask participants how well they can perform an action relative to others, thus violating the sixth best practice guideline in the process (Hohlfeld, Ritzhaupt, & Barron, 2013; PISA - OECD, 2020a; Williams-Diehm et al., 2018).

4.4.3 Validity and reliability testing in measurement

There are various indicators to make sure that survey questions and items measure what they are supposed to measure (validity) in consistent ways (reliability).

For validity, distinctions can be made between face validity (does it seem to be a good representation of the constructs measured?), content validity (are all dimensions of the construct reflected in the measures?), construct validity (convergent – does the measure relate as expected to measures of related but different constructs? and discriminant – does it not relate

to factors that are unrelated?), and criterion validity (does it relate to a different (non-survey) measure of the same thing?). For survey research, reliability testing includes test–retest reliability (a construct that is consistent across time and has similar scores when measured at different time points, e.g. someone with high levels of digital skills now should also have high skills in a month's time) and internal consistency testing (responses are consistent across the items that are supposed to measure a similar construct on a multiple-item measure).

Haddon et al. (2020) found that existing academic research uses a range of different validity and reliability checks (e.g. face validity through expert or cognitive interviews, internal consistency checks using Cronbach's alpha). Since no studies measured skills amongst young people over time, no test–retest reliability was established in any of these studies, and construct validity measures were also absent.

The review conducted for this report found that while most studies employ at least one of the validity and reliability testing methods, few use more than one. This means that studies often did either validity or reliability testing, but rarely both. Of the 53 academic sources whose measures and scales were considered and which did not use scales from previous research, 12 did not report validity or reliability tests in the main body of the paper, and only 9 reported using more than one method.

Table 5. NUMBER OF STUDIES USING THE MOST COMMON TYPES OF VALIDITY AND RELIABILITY TESTS

Validity and reliability tests	Number of studies
Face validity through expert consultation	6
Face validity through cognitive interviews	3
Tests for construct validity (mostly exploratory factor analysis)	10
Tests for internal consistency (mostly Cronbach's α)	27
Other (e.g. criterion validity through performance tests)	6

Note: Total number of studies is 53; some studies used more than one measure of validity or reliability testing.

The most prominent checks were for face validity, using cognitive and expert interviews and internal consistency testing, using Cronbach's alpha, to account for reliability of items on pre-established scales (see Table 5). A handful of studies also include exploratory factor analysis (EFA) to test for construct validity, and confirmatory factor analysis (CFA) was far less common.

5 Initial digital skills question and answer formulation and item selection

In order to narrow down the selection of 136 items to the initial selection for the pilot surveys, three sub-components for each dimension of digital skills were conceptualised based on the review of the literature. This improvised sub-classification served to ensure that the selected items captured the breadth of digital skills associated with each dimension, rather than to provide a definitive theorisation of sub-components in each dimension. The components within each dimension were:

- **Technical and operational:** Operating, connecting, customising.
- **Information navigation and processing**: Navigating, interpreting, evaluating.
- **Communication and interaction**: Managing, protecting, netiquette.
- **Content creation and production**: Producing, attracting, understanding.

In the initial design, at least two items were selected for each of these sub-components (five was the maximum number of items representing a sub-component). The items from previous studies were distributed amongst these sub-components and a selection was subsequently made on the basis of: (1) the number of studies the item appeared in, (2) whether it was used in cross-culturally validated studies and (3) how relevant it was in the specific context of its assigned sub-component and the overall project.

When present in multiple studies, items from DiSTO (2020) and Global Kids Online (2020) studies were given preference because these studies cross-culturally validated their questionnaires. Then, in second order of priority, items from other large-scale projects on digital skills amongst children (Net Children Go Mobile, 2020; DKAP - UNESCO, 2020) were used to fill the gaps in the sub-components of digital skills identified. In six instances, items were taken from the rest of the selected literature because these studies were particularly insightful or provided a better fit for the categorisation of the relevant skill than these larger projects (Lau & Yuen, 2015; Porat et al., 2018; Rodríguez-de-Dios et al., 2018).

When possible, the original phrasing was preserved, but several items were rephrased to adhere to the best practice guidelines presented earlier. In some cases, items were substantially amended so that they more accurately captured the desired skill and were appropriate for 12- to 17-year-olds. In this process, the main reasons for the exclusion of items were:

- The item was *not appropriate/too difficult for children*. It referred to online practices that were not relevant to children or it used jargon that was outdated or field-specific (e.g. it refers to a digital skill that is predominantly applied in a work context). All remaining items used language and addressed skills and topics that were deemed appropriate for children.
- The item remained *vague or confusing*. It could be interpreted in two different ways, or it was unclear exactly which digital skill it captured. Every item remaining in the selection measured a unique, clear, unambiguous skill.
- The item *was too general*. It asked about the general skill dimension and could not be placed into a clear sub-component. All remaining items were assigned to one of the sub-components.
- The item was *too specific*. It was device or software-specific and it was potentially not cross-culturally valid (e.g. it asked about the use of a specific program popular in few countries). All remaining items were selected on the basis that they should be equally applicable and understood after translation in all countries included in the ySKILLS project.

A number of items were created specifically for the yDSI to account for remaining gaps of measurement in sub-dimensions; two items were inspired by the DigComp (Vuorikari et al., 2016) and the Essential Digital Skills Framework (EDSF) (Department for Education, 2018),

and six items were created after discussions with ySkills colleagues. Two of these newly designed items were assigned to the dimensions of technical and operational, and communication and interaction skills, respectively. All other items in these dimensions and in the dimension of information processing and navigation were taken from the existing pool of items. The six remaining newly designed items were assigned to the content creation and production dimension. This is because, during the selection process of the content creation items, it became clear that the majority of the existing items focused on the functional aspect of content creation, which did not align with the conceptualisation and sub-components that were established for the purpose of ySKILLS.

The item selection process starting with the 136 unique items from the academic and grey literature resulted in an initial list of 12 technical and operational skills items, 9 information processing and navigation skills items, 12 communication and interaction skills items and 9 content creation and production skills items, covering all sub-components of these skills, while addressing the functional aspect of each of these. In addition to this list of items, 16 statements were created to capture the digital knowledge-based aspects of the information processing, interaction and content creation and production skill dimensions. This means that for the initial design of the skills scales, there were 58 items (42 skills items and 16 digital knowledge items).

Before finalising the items for validation, decisions on rephrasing were made jointly between experts from different strands of the ySKILLS project (see Appendix A for comments on why changes were or were not made). The 58 items were validated through cognitive interviews and pilot surveys (see Appendix B for a full list of these items).

In addition to the digital skills measures, digital critical understanding and knowledge items were developed (see Appendix B for a full list of these items). These were new, although they were inspired by the digital skills items used in previous survey research. No such items were encountered in the existing literature on digital skills in ways that complied with best practice guidelines. In particular they did not comply with best practice guideline 4, that is, with a majority of statements being false, or with best practice guidelines 5 and 6, with items not being about confidence or evaluating a skill in comparison to others (see, for example, Hargittai, 2005). Digital knowledge items were designed to reflect knowledge and understanding within the information navigation and processing, communication and interaction and content creation and production dimensions. Items were created being mindful of capturing various levels of difficulty, in the hopes of distinguishing between respondents with high and low digital skills-related knowledge.

The initial question and answer formulation for the long version of the yDSI (see Appendix B) was very similar to the final version (see Section 2). The instruction to partners was that replicating the design of question and answer scale formulation is essential in the application of the instruments: small changes can mean the properties of the scales change and that the surveys would no longer be comparable. For example, the answer scale for the functional skills items should not be changed to an agree or disagree Likert scale, and under no circumstances should the “I don’t know what you mean by this” option be left out. Nevertheless, translations of items in other languages did not have to be literal but should have the same intrinsic meaning and cognitive properties following the design best practices as set out in this report.

6 Validation of the yDSI – Step 1: Cognitive interview and pilot survey testing

The 58 selected survey items were first validated through cognitive interviews and pilot surveys. This section describes the methodology for validation for each in detail as well as the results obtained through validation. This is a relatively detailed, technical part of the report and those interested in just the final scales can jump straight to Section 8.

Each of these methodologies has its own contribution to make for validation: the cognitive interviews were used mostly to validate the meaning of the items (content validity) and the pilot surveys to test the statistical properties of the items and their consistency across countries and different sociodemographic groups (discriminant and convergent validity). Ethical approval for the ySKILLS project was granted to the coordinating partner KU Leuven, and complied with their internal and the European Commission's ethics guidelines. The procedures and reasoning behind each of these are described below.

6.1 Methodology for cognitive interviews validation

Cognitive interviews are akin to elicitation interviews, which ask people to reflect on an object or example presented in text or image form (Chepp & Gray, 2014). They are specifically intended to make explicit (i.e., elicit) the cognitive (i.e., thinking) processes that allow people to reflect on their thoughts, opinions and behaviours (Willson & Miller, 2014). Without the stimulus of the object or text, this is often very hard to do because most of our decisions and behaviours are made subconsciously.

Cognitive interviewing is, therefore, also a technique that is used in questionnaire and survey design to validate newly developed questionnaires. Interviewees are asked to answer a questionnaire while an interviewer asks them: (1) whether they had any problems understanding any of the items; (2) what they think the question means; and (3) what they were referring to when they gave the answer they gave (probing them to give further examples). This is done to understand whether what the researchers intended a question to mean is aligned with the respondent's interpretation (content validity). It helps to adjust wording for items that are not well understood and to remove items or reword items that measure something completely different (Desimone & Le Floch, 2004; Jobe & Mingay, 1989; Ridolfo & Schoua-Glusberg, 2011). The cognitive interviews test the item wording as well as the answer scale design. What is important here is that the interviewees come from diverse backgrounds (iin terms of country, gender, age and socioeconomic status, SES), so it is clear that the items are consistently interpreted in the same way across different sub-populations. For the ySKILLS project these cognitive interviews included the skills module but also other modules on sociodemographic and wellbeing variables, online opportunities and risks, and positive and negative outcomes of ICT (see Section 1.1). Below, the sampling and fieldwork procedure for the cognitive interviews in general, and the skills scales in particular, is described.

6.1.1 The sampling cognitive interviews

The cognitive interviews were conducted in August/September 2020 in six countries: Estonia, Finland, Germany, Italy, Poland and Portugal. Respondents were recruited through convenience sampling with a focus on equal gender, age and SES distributions. The sample in each country is described in Table 6.

Table 6.	AGE, GENDER AND SES/EDUCATION DISTRIBUTION OF THE COGNITIVE INTERVIEW SAMPLE							
	Age			Gender		SES/education		
Country	12–13	14–15	16–17	Boys	Girls	Low	Middle	High
Estonia	4	2	4	5	5	2	3	5
Finland		12–16		5	5	All relatively high SES		
Germany	4	3	3	5	5	3		7
Italy	4	2	4	5	5	3	3	4
Poland	4	2	4	5	5	4		6
Portugal	4	2	4	5	5	5	2	3

Note: Estonia reported level of education in the family.

6.1.2 Fieldwork procedure for the cognitive interviews

The items were translated into the national languages and trained interviewers from each national team conducted the testing. The interviewers asked the children about their understanding of the questionnaire, that is, whether the meaning described by the children in their own words corresponded with the intended one, whether the children could give examples, whether they missed some response option(s), and if they would be able to answer the questions.

All interviews were recorded and interviewers took field notes about the children's reactions to the questionnaire. This data was used to document the children's responses, comments and questions regarding the tested items. Informed consent was obtained from all the children and from their caretakers or teachers where necessary (depending on their age and national regulations). Blocks of questions were rotated for the testing to prevent the cognitive interviews from becoming too long and overly burdensome for the interviewer and the interviewee. The aim was to examine each block of questions in each age group across all countries.

A spreadsheet was completed on which interviewers indicated for each item for each interviewee which items and answer scales had caused problems. For the skills module, the interviewers asked about one module (e.g. technical and operational skills) at a time, and then asked questions about understanding and meaning for that dimension before moving on to the next set of items.

Since the digital skills measures were a combination of items that were new, had been considerably revised, had never been tested on the particular population of young people, or were never translated into the ySKILLS country languages, all were tested in the cognitive interviews. For this same reason, a set of specific questions for interviewers was provided accompanying the module of digital skills items (see Appendix C). This was provided in addition to the standard instructions for item validation through cognitive interviews.

Not all the children saw all the items. The selection of items was randomly presented from the items in the proposed list. The coordinating teams designed the randomisation to guarantee that all children saw at least one item in each (sub)dimension. They also provided the feedback sheets for the cognitive interviews in general. The comments were collated and fed back to the team designing the skills scales, highlighting the most important comments per question. The examples coming out of the cognitive interviews were also used to inspire the design of the performance tests.

6.2 Methodology for survey pilot validation

The pilot surveys were not part of the original ySKILLS proposal but were added when funds became available due to these not being spent on travel because of the COVID-19 pandemic. The decision to pilot the questionnaire on an older group (18- to 25-year olds) rather than the target age group for the ySKILLS project was based on the practicalities of getting ethical approval and costs, and the speed with which the fieldwork needed to be done. Pilots would not have been possible within the short timescale of the project if they were to have surveyed 12- to 17-year olds. Nevertheless, to validate the statistical properties of the scales, focusing on 18- to 25-year olds was deemed to be a reasonable alternative. The inclusion of this part in the project proved invaluable for validation and item selection.

These pilot surveys contained all the digital skills items and a selection of key sociodemographic variables, such as age, gender and level of education. As the purpose of these pilot surveys was to test the statistical properties of the digital skills items only, they did not include the other variables included in the cognitive interviews or the broader ySKILLS survey (e.g. physical wellbeing, online risks and opportunities). Since the survey participants were adults, only individual informed consent was required, and no caretakers or intermediaries were contacted.

6.2.1 Pilot survey sampling

Data were collected throughout the month of September 2020. These pilot studies ran in each of the countries that are part of the ySKILLS project (Estonia, Finland, Germany, Italy, Poland and Portugal), in addition to the Netherlands and the UK, since the scale was developed in English before being translated into the main language used in each of the ySKILLS countries, and the Netherlands was part of the team that designed the performance tests.

Participants were recruited using the survey respondent platform of Toluna in all countries. Toluna is a market research company with a representative panel sample in countries across Europe.[1] Their participants get rewarded for a number of surveys completed during a year. Survey respondents were reached through three methods: (1) direct or targeted automatic invitations, where participants are sent invitations depending on the sampling and criteria of the study; (2) personal survey centre invitations, where participants are able to log in to their Toluna profile and see if they meet the predefined quotas and criteria; and (3) through real-time recruitment on third party websites, where participants can follow a link taking them to the survey after filling in their demographic information and ensuring that they comply with the predetermined criteria. Furthermore, each potential respondent was sent at most one additional email invite for the same study if they did not participate the first time. Responses were collected through Qualtrics on LSE servers, and answers could not be linked to individuating information of the respondents. LSE complies in its data storage facilities with the EU's General Data Protection Regulations (GDPR).

The target sample was 300 respondents in each country, a minimum of one-third of respondents without higher education (university or equivalent) and a cut-off of 50% female respondents per country was set, to make sure that the sample included a minimum of those with sociodemographic characteristics that are likely to make a difference in how individuals interpret and engage with skills measures. Some flexibility was permitted towards the end of the data collection, depending on the country, especially around the gender quota, to ensure that the benchmark of 300 respondents was reached in every country. The fieldwork process

[1] https://tolunacorporate.com/industry/market-research/

was closely monitored and Toluna staff were updated daily so that they could send invitations to the appropriate demographics to make sure these quotas were approximated as closely as possible.

In addition, only responses from participants who took longer than a pre-set length of time were counted. This time was determined through a trial run of 30 responses in each country, and was set at 7 minutes for all countries, except for Germany and the Netherlands, where it was set at 5 minutes, as participants in these countries seemed to require less time to complete the survey.

Table 7 shows a breakdown of invitations sent out in each country, the number of participants who started the survey, and how many were filtered out by each of the above-mentioned restrictions.

Table 7. INVITATIONS, SURVEY RESPONDENTS AND FILTERED-OUT CASES FOR THE PILOT SURVEY

	Invitations		Surveys	Filtered out based on:			
Country	Estimated reach survey centre	Direct invites sent	Recorded responses	Gender and education quotas	Age	Time	Non-completion
Estonia	0	7,500	607	86	22	110	89
Finland	500	1,800	674	78	64	122	102
Germany	700	200	722	94	78	151	98
Italy	1,400	7,900	1,736	929	144	118	230
Netherlands	1,400	6,100	1,535	610	194	168	263
Poland	2,200	3,900	2,506	1,690	186	68	249
Portugal	800	3,500	1,012	410	85	82	135
UK	1,100	1,100	1,190	371	178	190	150

The final sample was composed of 2,438 individuals, evenly spread out across all eight countries involved in the pilot study. Table 8 shows the breakdown of respondents per country, including gender and education splits.

Table 8. PILOT SURVEY SAMPLE

Country	Number of participants	% (number) of women	% (number) of higher educated
Estonia	300	51 (153)	58 (175)
Finland	308	56 (172)	59 (183)
Germany	301	56 (169)	64 (192)
Italy	315	47 (149)	68 (214)
Netherlands	300	52 (157)	60 (180)
Poland	313	50 (157)	59 (183)
Portugal	300	51 (153)	68 (204)
UK	301	50 (149)	62 (187)
Total sample	2,438	52 (1259)	62 (1518)

6.2.2 Pilot survey analytical procedures

While respondents in most countries were shown all 58 digital skills items, around two-thirds of respondents in Estonia and the UK filled in an earlier version of the survey, which contained the same items, but respondents were only shown a random subset of items for each dimension of digital skills. This issue was rectified as soon as it was spotted, to ensure that the remaining respondent in these countries answered all items, as was done in the other countries. Controls showed that this discrepancy in the data collected in these two countries did not lead to any statistical anomalies or inconsistencies in our findings.

The first step in the analysis was to create descriptive tables that showed the characteristics of each item. This was done to check whether certain items were very problematic in terms of skew (i.e., non-normal distribution with platform or ceiling effects) and kurtosis (i.e., non-normal distributions with more than one peak). They were also used to get a first impression about whether individual items had the same characteristics across countries and across gender, age and education groups.

Both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were subsequently used to examine how the skills items loaded onto their hypothesised dimensions (1) across all countries and (2) within each country. The factor analyses were run on the complete pool of items. Items that did not load onto the digital skills dimension that they were associated with were removed from the pool (see Section 6.4).

Selection from the remaining items incorporated insights gained from the cognitive interviews, as well as considering the skewedness of the items (e.g. if a choice had to be made between two similar items, the least skewed one was selected).

The final step in the selection of items took into consideration whether factor loadings appeared consistent across countries, and whether any item seemed particularly problematic in a certain language. This was done through factorial invariance tests. These tests for measurement invariance are concerned with the equivalence of factors across different groups, in terms of both their measurement and of their structural relationships (Byrne, Shavezlson, & Muthen 1989). Invariance is tested by constraining certain parameters (e.g. factor structures or configural equivalence and factor loadings or metric equivalence of items) to be equal across groups (Byrne 2001; Vandenberg 2002). These statistical tests can point out whether items and groups of items have different meanings across countries and sociodemographic groups (i.e., if there is variance in factor loadings and structures). These analyses examined whether the items of the initial yDSI items grouped in the same way across countries and whether the factor loadings were more or less equivalent.

In addition to the digital skills items, all digital knowledge items were included in the pilot surveys. Selection from these items was based on the relative difficulty of each item, such that each dimension should include one item that was relatively easy and one that was relatively difficult, to make the distinction between high- and low-skilled respondents (discriminative validity). At the end of this process a selection of three items corresponding to the information navigation and processing, communication and interaction, and content creation and production dimensions of digital skills was made.

6.3 Results: Validation through cognitive interviews

The six countries in which the skills scales were validated through cognitive interviews gave detailed feedback (detailed notes on the cognitive interview comments based on initial skills scale proposal are available on request). Some problems were observed across the board in most countries and for most respondents (of a certain age category); others were solely applicable to specific countries.

6.3.1 General issues

A few general comments came up in the translation process and the cognitive interviews that led to adjustments in the question for the skills scales and the answer options for the digital knowledge items. These consisted of:

- Reformulating the question for the skills items to (changes in red): Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones or computers. ~~If you have never done this, then~~ Reply thinking about how true this would be of you if you had to do it now, on your own. If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

These changes were made because the phrasing of the initial sentence ("If you have never done this … if you had to do it now and by yourself") was confusing for several children across the different countries. This was mostly a translation issue: the formulation "and by yourself" in English was difficult to convert into the other languages. The change to "on your own" made the sentence more straightforward to translate and understand. "Or computers" were added at the end of the first sentence to ensure that children did not solely focus on mobile phones.

- Adding the following instructions when there were lists of examples: Sometimes there are various examples given; only select "Very true of me" if all of the examples apply to what you do or know.

This change was made after the cognitive interviews captured several instances in which children answered "Very true of me" on items that included multiple examples but were unable to explain or describe most or all of the examples. This additional line in the instructions clarified to children that they should only report the highest level of skills if they knew how to apply every element of said skill item.

- In the translation process it became apparent that the knowledge-based items answer scale needed to be adjusted. The "Mostly not true" option was deleted because it was hard to make a distinction between that option and the "I'm not sure" option in its meaning. Thus, the final scale for these items consisted of: "Definitely not true", "Definitely true", "I'm not sure", "Don't want to answer".

The next step was to look at specific items that needed to be adjusted based on feedback from most children in most countries. These were seen to indicate more structural problems in the conveyance of meaning and understanding of the questionnaire. No general adjustments were made to items that only one participant found problematic or difficult to understand or that only occurred in one country. Some adjustments were suggested for items that did not make the cut after the pilot survey was analysed (see Section 6.4). Unexpectedly there were more problems with the technical and operational skills items than with the communication and interaction or content creation and production items. The following items were adjusted (original in red in brackets):

- I know how to recognise whether a WI-FI network is safe and secure (I know how to connect to a safe and secure WI-FI network).
- I know how to connect devices to each other using Bluetooth or wireless connections (I know how to connect and install new devices (e.g. a modem, camera, printer)).
- I know how to store photos, documents, contacts or other files in the cloud (I know how to store photos, documents or other files in the cloud).
- I know when I am allowed to use content covered by copyright (I know how to reference and use content covered by copyright).

6.3.2 Country-specific item issues

When issues occurred with items in specific countries only, partners in these countries were asked to check for possible translation issues and to think about how to possibly rephrase the items. Below is a summary of these issues for particular countries and the changes that were made.

Portugal and Italy

- I know how to choose the best keywords for online searches

There were comprehension problems with the word "keywords" in Italy and Portugal. In Italy, the respondents interpreted this as relating to traditional literacy issues (being fluent in your language) rather than critical skills. In both Italy and Portugal, respondents confused it with "password" (this is a language issue, as the translation for both words is similar).

On discussion with the ySKILLS partners from these countries, some suggestions for improvements were made, but it was decided to keep the item as it was, since it had previously been used and translated into Italian and Portuguese in the context of EU Kids Online (EU Kids Online, 2019), where cognitive interviews did not reveal any of these issues.

- I know how to block unwanted pop-up messages or ads

There were some problems with translations of the term "pop-up" in Italy and Portugal. In Italy, the English word "pop-up" was used; however, some children did not understand it. In Portugal, the term was translated; however, some respondents understood it as messages from strangers (perhaps due to this translation).
Discussion with the Italian and Portuguese partners revealed that there is no real translation for 'pop-up' in either language. For lack of a better alternative, the item was kept using the English word.

Germany

- I know how to create something that combines different digital media (photo, music, video, GIF)

Some respondents in Germany did not understand the question, and one was not sure and would have checked "I do not know" as an answer even though she gave good examples of what it could be. This problem did not occur in any other country. Since this was a problem in the German context for very few respondents, they revised their translation. In addition, the change in specifying how to answer questions with a number of examples (see Section 6.3.1) should address some of the underlying confusion.

- Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones. If you do not understand what the question is asking, tick the box "I do not understand what you mean by this"

The phrase "use the internet and technologies such as mobile phones" was problematic for respondents in Germany. Several children did not understand "technologies". It was suggested rephrasing this as: "use of internet, mobile phone and laptop". On this comment in Germany, the decision was made to change the item in all countries "... and technologies such as mobile phones and computers" (see Section 6.3.1).

6.3.3 Selection of items based on content validity

Based on the cognitive interviews, items were excluded because they showed too much confusion and misinterpretation amongst respondents when there was an alternative item measuring a similar subconstruct that had better properties.

- I know how to connect devices to each other using Bluetooth or wireless connections

The cognitive interviews showed that this item was difficult to interpret for many and had to be considerably changed. It was also conceptually similar to another item, so the one that did not require changes was kept.

- I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people

This was conceptually similar to other items in content creation and production sub-dimension of reach, but it was less clearly understood, so this item was deleted.

- I know how to figure out if a website can be trusted

This item was similar to other items in the information navigation and processing dimension, but was deemed to be less distinctive. Indeed, multiple (younger) respondents in the cognitive interviews understood the reliability of a website as the reliability of the information on it, which was captured by another item ("I know how to check if the information I find online is true").[2]

- I know how to use filters and other tools to make a photo or video look more attractive

This item was conceptually similar to other items in the content creation and production dimension, but was deemed to be too specific in terms of the technological application it focused on. While respondents generally understood this item, it seemed to only capture a sub-division of a digital skill that was already captured by another item ("I know how to edit existing digital images, music and videos").

- I know how to recognise whether a WI-FI network is safe and secure

This item was difficult to interpret for many children, and ambiguous. It was rephrased but, on further consideration, it was removed since it was unclear how this item would add value to the broader discussion around digital skills at later stages in the ySKILLS project. Items that did not require changes in phrasing or clarifications as a result of the cognitive interviews were preferred.

6.4 Results: Validation through pilot surveys

This section describes the results of the analyses of the pilot survey that led to the removal of items from the initial yDSI instrument. The decisions for the digital skills items were made

[2] On revision of the factor analyses for the short version of the yDSI and the performance tests, this item was reintroduced.

based on the skewness and kurtosis of the items in different countries, on the factor loadings in different countries and based on the results of equivalence testing. The decisions about selection of the digital knowledge items were made based on their degree of difficulty and the distinctiveness (i.e., discriminatory validity) of the items.

6.4.1 Statistical properties of the digital skills items in the pilot survey

The analysis of the pilot survey that led to the selection of the skills items for the yDSI to be validated through performance test were threefold: descriptives, factor analyses and equivalence testing. These are described below.

6.4.1.1 Descriptives: skewness and kurtosis

Descriptive statistics reporting on means, standard deviations, skewness and kurtosis were run for the full sample, including all countries, as well as for each individual country. At the first stage of item selection, the skewness and kurtosis of each item were considered at the level of the overall sample. As a benchmark, a skewness between –2 and 2 was deemed good, and a skewness between –3 and 3 was deemed acceptable in terms of the distribution of answers for a given item. Of the 41 digital skills items, 17 had a skew below –2. Specifically, 6 items in the technical and operational dimension, 1 item in the information navigation and processing dimension and 10 of the 12 items in the communication and interaction dimension had a skew below –2, indicating that these items may have been "relatively easy" for all participants. Only two had a skew below –3, which means the item was "too easy"; there was little variance in the answers since most participants indicated the highest skill level. These two items were excluded from the selection as a result. All items that showed a kurtosis above 5 were among the deleted 17 items.

While a skewness between –3 and –2 did not constitute a criterion for exclusion in and of itself, extreme skewness and kurtosis values were taken into account when having to choose between two items at subsequent stages of the selection procedure. Items with a greater skew and kurtosis were prioritised for exclusion after looking at the information gathered through the cognitive interviews. In the communication and interaction dimension, the cognitive interviews did not suggest any item should be deleted. Decisions to keep or exclude a specific item over another in the similar sub-dimension were made by looking at these descriptive statistics. In every case, the item with the lower skewness was preserved.

The following items were excluded from the final selection using skewness as the deciding factor:

- I know how to remove apps from a mobile device > skew –3.29
- I know how to remove people from my contact list > skew –3.04
- I know when I should not post pictures or videos of others online > other similar, less skewed items
- I know how to change who I share content with (e.g. just friends, friends of friends or make it public) > other similar, less skewed items
- I know how to make my comments and behaviour appropriate to the online situation > other similar, less skewed items.

In addition, two items were marked for potential deletions, as they were highly skewed (< –2.75), but the decision to delete was left pending the factor analysis and performance tests:

- I know how to block messages from someone I don't want to hear from

- I know how to delete the record of sites I have visited before.

6.4.1.2 Factor analyses

Parallel to the descriptive analysis, exploratory factor analyses (EFA) and confirmatory factor analyses (CFA) were conducted for the full sample, including all countries. The EFA suggests that digital skills items are best split into six distinct factors, but a closer look at the factor loadings of each item reveals that our items are spread out predominantly across four factors, and that these factors closely correspond to the dimensions each item was assigned to before the start of the pilot survey.

The CFA revealed that all but three items did indeed load well onto the factor corresponding with their intended dimension. The items that did not load onto their intended factor were "I know how to use programming language (e.g. XML, Python)", which is intended to be a technical skills item, but did not load onto any of the other factors in the EFA either; "Sometimes I end up on websites without knowing how I got there", which was intended to be an information navigation skill, but did not load onto any other factor either; and "I know who to turn to when someone I know is being bullied or harassed online", which was intended to be a social skill, but loaded more strongly onto the factor associated with creative skills. As these findings were replicated in the CFA of every country separately as well, the latter two items were removed from the list of items to be considered for the final report. The item concerning programming languages was retained, because this item is universal across all existing studies and is seen by many to measure the highest level of technical or content creation skill. It also often appears as a separate dimension instead of as part of another dimension. In addition, it is a core component of many computer and digital literacy training and curricula (Polizzi, 2020a).

Since all skills dimensions are part of the overarching concept of overall digital skills, they should be correlated (see Section 8.3 for suggestions on how to create the scores for the different dimensions). However, since they also measure different dimensions, these correlations cannot be so high as to indicate that they measure the exact same construct. In the pilot survey, the correlation between the different dimensions of digital skills was highest between the technical and operational dimension and the communication and interaction dimension (r=0.82). This would cause a potential multi-collinearity problem in statistical analyses. The other correlations ranged from r=0.56 between the information navigation and processing and communication and interaction dimensions to r=0.67 between the communication and interaction and content creation and production dimensions. While this level of association is acceptable, on the final short scale further selection of items improving discriminatory validity was conducted to improve the distinctiveness of items and dimensions (see Section 8.1).

Table 9.	STATISTICAL PROPERTIES OF PILOT SURVEY SKILLS SCALES (INCLUDING DELETED ITEMS)		
Digital skills dimension	Cronbach's α	Means	SDV
Technical and operational	0.86	4.18	1.15
Information navigation and processing	0.77	3.94	1.13
Communication and interaction	0.88	4.38	1.05
Content creation and production	0.85	3.78	1.26

Base: N=2,438

Note: "Means" represent the average on the scale across all items in a dimension. Scales run from 0 = "I do not understand what you mean by this", to 1 = "Not at all true of me" to 5 = "Very true of me". See Section 8.3 for a description on how scales should be constructed for analysis.

Considering the high alphas for the scales when the data was combined across countries (see Table 9), it would be perfectly acceptable to use all these items for a longer version of the yDSI. Therefore, the recommendation is that:

> Those **studies that** have more space and **focus mostly on digital skills** should use all the items that were used for the pilot test (see Section 8.2). Incorporating the wording changes suggested based on the cognitive interviews and not including the items that had low loadings or a skew below –3.

EFAs and CFAs for individual countries were subsequently conducted to confirm that the factors were consistently present and item loadings were broadly similar across countries. While the data for individual countries were split into eight to thirteen factors in the EFA in different countries using Eigenvalues of 1 as cut-offs (not presented in this report), there were only six or seven factors that had unique and high-item loadings; the other factors had either low-item loadings across the board or consisted of items that already loaded highly onto other factors. In other words, a six or seven factor solution emerged naturally from the data. In the EFA, the following patterns can be observed:

- Estonia did not have clear factors for either *technical and operational* or *information navigation and processing* skills dimensions.
- The *communication and interaction* dimension was split into two factors in Finland.
- Only about half the *technical and operational* items loaded onto one factor in Germany.
- The *technical and operational* skills dimension was split into two factors in Italy.
- In the Netherlands, there was no clear factor for *information navigation and processing* skills, and the *content creation and production* dimension was split into two factors.
- The *information navigation and processing* dimension was split into two factors in both Poland and Portugal.

These results were interesting for potential further exploration but did not inform selection of items. Since the main aim was to test the suitability of the conceptual model, the results that really matter were those of the CFA (see Appendix E). In correspondence with what was found for the CFA on the full sample (for all countries), all but one of the items loaded well onto the factors they were associated with based on the conceptual model. The only exception was "When I have a question, I am able to find information online that is relevant to answering it", which loaded onto the information navigation factor in most countries, but not in Germany. While this might have been a translation issue, consistency across countries is of major

importance in the construction of the final digital skills scale for the main project. Hence, this item was removed from the reduced list of items to be tested in the performance test.

The two items that were pending deletion after the discussion of skewness were: "I know how to block messages from someone I don't want to hear from" and "I know how to delete the record of sites I have visited before". These items did not present any other statistical properties that suggested they should be deleted after the factor analysis. However, since neither of them presented any statistical advantages over the other items, and sufficient items remained in each dimension to create the final selection, they were flagged as potential deletions ahead of the performance tests.

6.4.1.3 Equivalence testing

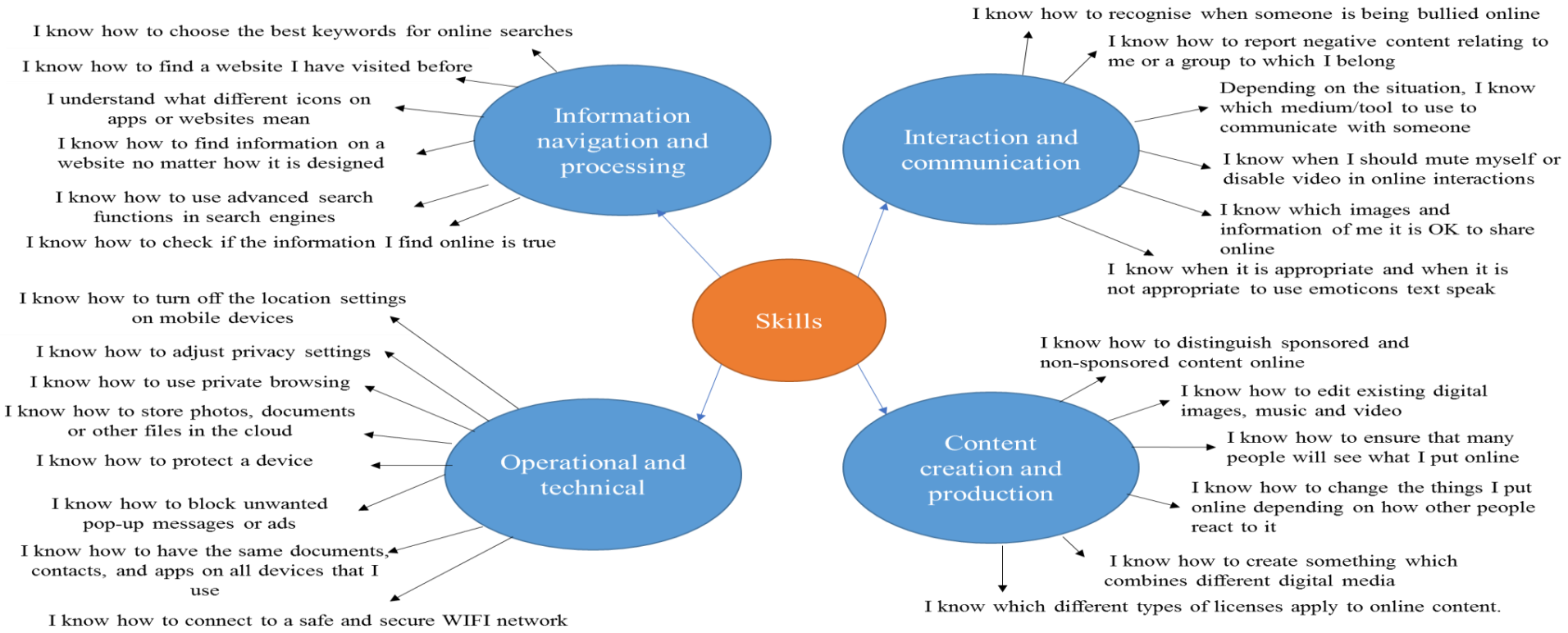
The 26 items that remained after the selection based on the cognitive interviews and the overall and country-specific factor analyses were analysed using equivalence testing to understand if the proposed dimensions and the items loading onto them could be considered similar or equivalent statistically.

The yDSI equivalence testing looked at whether the model with the four skill dimensions fits each country, the different gender and education level groups, and whether the loadings of these items can be considered similar across these groups (a country is considered a group). Equivalence testing is done by comparing a model in which the factor structure is allowed to vary between groups (i.e., the baseline model) with one for which different elements of the model (e.g. the number of factors and items that load onto them, the loadings of the items onto the factors) are fixed to be similar between groups (i.e., a restricted model). If the chi-squared difference between both models is statistically significant, it suggests that the restrictions added to the baseline model do not hold across groups (Chen, 2007; Byrne & Stewart 2006). In other words, a significant Chi-square means that the model that proposes that the factor structure is the same is a worse fit than one in which the factor structure is different between groups; there is no equivalence of factor structures between countries. CFI and RMSEA[3] can also be used to compare models: when sample size is adequate and equal across groups, a change of ≤ -0.01 in CFI supplemented by a change of ≥ 0.015 in RMSEA indicates non-invariance (Chen, 2007).

The model in Figure 4 was the one tested for configural (i.e., the factor structure is the same across countries) and metric equivalence (i.e., the factor loadings are similar across countries). Knowing a programming language was left out because it did not load highly onto the technical and operational skills factor, even though it will be taken up in the final instrument (see Section 6.4.1.2).

[3] Root Mean Square Error of Approximation

Figure 4. LATENT VARIABLE MODEL TESTED FOR EQUIVALENCE ACROSS THE DIFFERENT COUNTRIES



Overall model fit full sample: χ^2_{247}=1,106.345, p=0.00; CFI = 0.96; RMSEA = 0.04 (CI = 0.036 to 0.04). Unconstrained model fit (baseline for six equivalent countries) χ^2_{2114}=7,337.92, p=0.00; CFI = 0.69; RMSEA = 0.04 (CI = 0.036 to 0.04).

Running the model on the full sample, including all the countries, confirms that the factor structure is a good fit to the data with a CFI >0.90 and an RMSEA <0.05 (including the confidence interval completely under 0.10). The analysis showed that a model that compares the factor structures in Estonia, Germany, Italy, Poland, the Netherlands and the UK has reasonable configural invariance. This means that in these countries the items all load onto the four dimensions (factors) as theorised in the conceptual model (model comparison χ^2_{23}=33.19, p=0.08; model fit χ^2_{2137}=7,371.11, p=0.00; CFI = 0.69; RMSEA = 0.37 (CI = 0.36 to 0.38)). However, when Finland and Portugal were added, the configural invariance became significant (model comparison between all countries χ^2_{161}=346.74, p=0.00). This means the model is variant across groups in terms of its factor structure. Thus, in Finland and Portugal some items load onto different skills dimensions.

There was no metric equivalence across the countries (model comparison between all countries χ^2_{217}=1,429.48, p=0.00), not even between the countries that showed configural invariance (model comparison all countries χ^2_{31}=106.28, p=0.00). Establishing metric equivalence with a complex model across this many groups (i.e., countries) tends to be difficult. This means that, even though the factor structure is the same, the weight of individual items in determining the underlying skill differs.

Further equivalence testing was conducted to test for equivalence across gender and education level groups. Configural equivalence (χ^2_{23}=33.57, p=0.07) and even metric equivalence (χ^2_{31}=50.77, p=0.01) was established between educational groups but not between gender groups (χ^2_{23}=48.36, p=0.00 and χ^2_{31}=236.47, p=0.00). The latter is slightly surprising. Nevertheless, it is really encouraging that there is equivalence across different education levels, because this indicates that there are no differences in interpretation between groups, and that there is no issue in terms of the understanding of the items even when there might be different answers (i.e., skill levels) between these groups. The ySKILLS project will examine the gender differences in factor structures and loading in more depth during the analyses of the three-wave longitudinal panel survey.

The items in the latent variable model (see Figure 4) were the ones further validated in the performance test (see Section 7.5), after which the final selection of six items for each dimension was made.

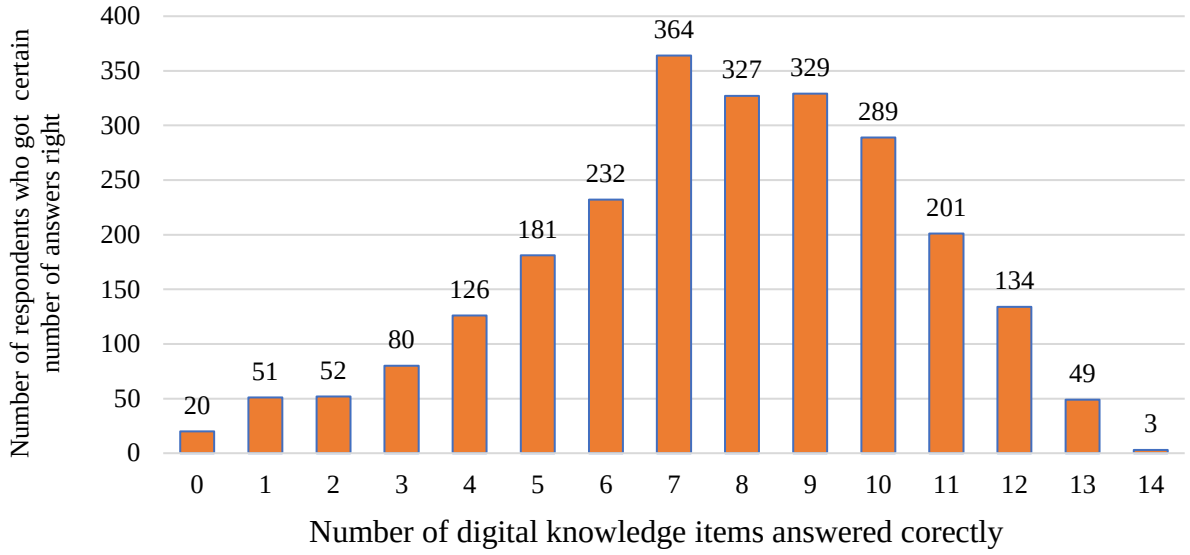
6.4.2 Statistical properties of the digital knowledge items in the pilot surveys

For the digital knowledge items a different procedure for statistical validation had to be followed. Each item was first recoded so that the correct answer was coded as 1. Mostly the correct answer was "Definitely true" or "Definitely false", but in some cases an "I am not sure" answer was coded as correct in addition to the true or the false answer. All remaining answers were coded as 0. The item "The lock icon means a website can be trusted" was removed because the ambiguity of what about the website could be "trusted" (e.g. not easily hacked or the content on the website) meant that all answers could be classified as correct.

In addition, the item "I can freely use an image published with a creative commons license for commercial purposes" was removed because the answer to this item was too dependent on the context (different types of creative commons licenses have different rules about commercial use).

After this, the number of correct answers for each item was sorted by the total percentage of answers that was correct, such that items could be ranked relative to how difficult they were. Out of the whole sample of 2,438 respondents, only 3 got all 14 answers right, while 20 got none of them right. The number of correct answers per participant is nicely distributed across the sample (see Figure 5).

Figure 5. DISTRIBUTION OF CORRECT ANSWERS ON DIGITAL KNOWLEDGE INDICATORS



After this, two more items were removed, because they did not display the expected pattern of incorporation into the list of items that were correct since participants got more answers right. These items were "Online cookies protect my information and activities online from being shared with other companies or organisations" and "It is easy to distinguish content produced by bots from that produced by humans". Both items appeared to be answered better by people who got few answers correct (as well as by those who got most answers correct), which should not be the case if answers to these items are indicative of the participants' knowledge level. The latter item was later reintroduced, as thinking that it is not easy to recognise bots could be an indication of a lack of knowledge (i.e., not understanding what bots are) or of high levels of knowledge (i.e., knowing that it is indeed very difficult to recognise content produced by bots). In that case, the distribution that was found of people at the extreme ends of the scales getting the answer right could be expected. Inclusion of this item into the final questionnaire was left in, pending the findings from the performance tests.

After the initial deletion of the first item, the remaining 13 items were ranked from easiest to most difficult, in order to determine which ones to keep. The dimension that each item corresponds to was factored into this decision, such that each dimension (i.e., information navigation and processing, communication and interaction and content creation and production) would be left with three items. In order to determine discriminant validity, that is, to choose items that clearly distinguish between different types and levels of digital skills, the objective was to take one easy, one intermediary and one difficult item for each dimension of skills (see Table 10).

Table 10.	SELECTED DIGITAL KNOWLEDGE ITEMS WITH THEIR DIFFICULTY RANKING AFTER THE PILOT SURVEY (USED IN PERFORMANCE TESTS)	
Digital skills dimension	Item	Difficulty ranking (1–13)
Information navigation and processing	The first search result is always the best information source	1
	Everyone gets the same information when they search for things online	7
	It is easy to distinguish content produced by bots from that produced by real people	13
Communication and interaction	Before sharing a picture that clearly shows a friend, I should always ask them for permission first.	5
	The first post I see on social media is the last thing that was posted by one of my contacts	10
	Whether I like or share a post can have a negative impact on others	11
Content creation and production	Using hashtags increases the visibility of a post	2
	Companies pay ordinary people to use their products in videos and content they create	4
	When information is backed up on the cloud, it is always encrypted	12

Note I: Difficulty was determined on the overall list of 13 digital knowledge items. A higher score indicates a higher level of difficulty.
Note II: There were no technical and operational skills measured in the digital knowledge items since these were covered with validated measures in the digital skills items.

The selection of knowledge items resulting from the pilot survey were then included in the survey leading up to the performance tests, after which decisions on selection for the short version of the yDSI to be used in the ySKILLS questionnaire was made.

7 Validation of the yDSI – Step 2: Performance tests

Performance tests are a direct assessment of children's digital skills in performing authentic tasks. They are a set of tasks that children have to complete while being monitored by a test leader. Prior research revealed that such practical performance tests can reliably check actual levels of digital skills (van Deursen et al., 2012). The performance tests were the second step in the validation process of the yDSI. They were developed based on the results from cognitive interviews and pilot surveys (see Sections 6.3 and 6.4) and served two goals. First, since the results of performance tests provide a realistic view of people's digital skills, they were compared to the answers of survey questions to test for criterion validity. This testing took place by correlating the digital skills survey question scores with the scores participants obtained on the performance test. Second, observation of behaviours during the performance tests gave in-depth insight into how children complete certain online tasks, which provided information about why some of the skills items developed for the pilot did or did not work as expected.

This section briefly describes the kinds of performance tasks that have been designed for previous research and their associated conceptual frameworks before moving on to describe the sampling, procedure and analyses conducted to validate the 39-item version of the yDSI that remained after the validation through cognitive interviews and the pilot survey.

7.1 Review of existing performance tests

Haddon et al. (2020) showed that about one-third of the academic literature on digital skills includes performance tests as tools to measure participants' actual skills. However, these studies predominantly focus on only two of the four dimensions of digital skills that are considered in the present study, namely, technical and operational, and information navigation and processing skills. The absence of studies focusing on measuring performance on interactive and communicative and content creation and production skills presumably arises from the difficulty in scoring or quantifying these. Indeed, while it is easy to assign a score to a technical or operational skill, both communication and interaction and content creation tasks are inherently more subjective when it comes to the evaluation of a successful, high-quality outcome.

Among studies that used performance tests for more easily quantifiable dimensions, they were often combined or intertwined (e.g. information problem-solving). Similarly, information navigation or content creation tasks may implicitly include operational components. This is not necessarily problematic if the scoring of such tasks appropriately captures the various skills that it measures.

Another issue with performance tasks is that participation and scoring for a single participant is time and labour-intensive. This concern is particularly relevant as the population of interest of this project is children, whose attention span may decrease more rapidly, especially if the tasks are cognitively demanding. Because of the time and labour-intensive character of performance testing, most studies were conducted on a small scale, meaning that they targeted small age ranges and did not test for full cross-cultural comparability. Since the ySKILLS survey focuses on children from a broad age range (12- to 17-years old) and runs in six European countries, these limitations had to be taken into account when using the performance tests to validate the yDSI.

7.2 ySKILLS performance task design

This section reports on the design of the performance test used to validate the survey measures. A cycle of task development followed by evaluation was used, both through individual development followed by discussion amongst WP3 partners and through pilot performance tests (five children in the Netherlands and five children in the UK including 12-, 14- and 16-year-olds) followed by

adjustments based on the feedback the children in these pilots gave. This development cycle provided information about task clarity, comprehensibility, applicability, difficulty and whether the tasks were appropriate for children of different ages in different countries. This ensured that the test was perceived to be appropriate and relevant. After several cycles the test was considered sufficiently robust and stable not only in terms of validity, but also in that it was a reliable instrument. What follows is a description of the tasks developed (see Appendix F). Each section roughly corresponds to a particular skills (sub)dimension. Nevertheless, performance tests measure whether someone can successfully complete a task and, therefore, as explained in Section 3.1, all tasks require a broad spectrum of skills and not just those from the skills domain most obviously related to the specific activity. The technical and operational skills are not separately represented because these were integrated in all the tasks developed.

7.2.1 Part 1. Information navigation and processing tasks: Navigating

Remarkably, the design of the information navigation and processing tasks – which are normally considered relatively easy to construct – proved to be quite difficult, as the solutions to the tasks needed to be available in the native language in all participating countries (Dutch, Estonian and Portuguese). A number of search tasks were tried with different topics, and eventually universally solvable tasks were constructed related to Netflix and dinosaurs. All three search tasks were fact-based and had a correct answer. The children were asked to start their search by using a search engine (e.g. Google, Bing). In all tasks, (1) the number of searches, (2) the keywords used, (3) whether an evaluation of the answer took place, and (4) finding of the correct answer were scored. In one of the tasks, the children were asked to specify their search by date and type of information (e.g. news) and they were scored based on whether they narrowed down their search in this way.

7.2.2 Part 2. Critical information navigation and processing: Evaluating

The children were asked to interpret textual and visual information in four social media posts. The four posts represented an ad, phishing, news and fake news. After each post, the child was asked what they thought the purpose of the post was. Although these were open-ended tasks, they were designed so that there was no ambiguity in the interpretation of the answers. The coding scheme scored whether the child correctly identified the intention of the makers of the post.

7.2.3 Part 3. Communication and interaction skills tasks: Protecting

This section contained two tasks in which interactions on social medium platforms were replicated. In the first task, the child was presented with a Snapchat message in which they were invited to a school party by an unknown person who asked them to send a photo. The child was asked how he or she would respond. Initially, WhatsApp was used for this task, but since the pilots showed that that young people were unlikely to receive this kind of message on WhatsApp, the platform was changed to Snapchat. The following things were scored: (1) whether the picture would be shared or not, (2) the reasons for sharing or not sharing a picture, and (3) other steps the child might take (e.g. blocking the sender, telling an adult or a friend).

The second task replicated Facebook posts, even though some participants in the pilot indicated that they were less likely to be active on this platform. The decision was made to stick to Facebook since the aim was to test for transferrable, platform-independent skills, and it would be beneficial to see whether young people could apply their skills in a lesser-known environment. The child was shown two posts on Facebook. The first post revealed a telephone number (publicly shared), and the second a bikini photo (shared only with friends). The children were asked whether the posts were appropriate. These open-ended tasks were again designed to make sure there was no ambiguity of interpretation

of the potential answers. The coding scheme scored: (1) the different responses (appropriate/not appropriate) and (2) the accompanying explanations (e.g. mentioning privacy settings). For the bikini picture post participants could argue that it was either appropriate (because it was only shared with friends) or inappropriate (since it was too revealing, even for friends).

7.2.4 Part 4. Critical communication and interaction tasks: Netiquette

This task presented participants with two discussions on WhatsApp between two children discussing climate change. In each discussion, one person denies climate change and the other advocates that it exists. In the second discussion, the person who is arguing that climate change is real and problematic becomes insulting and leaves little room for the other to give their opinion. The participant was asked whether there was something problematic about the two discussions. This open-ended task was scored on: (1) the different responses (problematic/not problematic) and (2) the accompanying explanations. Only the second discussion with the aggressive element should have been considered problematic.

7.2.5 Part 5. Content creation and production tasks: Producing, attracting and understanding

These tasks tested the ability of the participant to reflect on how content they consume is produced and to create quality content themselves. The first task examined whether children were familiar with ways to share messages or other content with a larger public (e.g. make it go viral). This open-ended task was scored on the number of different ways mentioned. The next task concerned a presentation presumably provided by a classmate. First, the participants were asked whether they were familiar with other ways of sharing a presentation than by email. The different ways of sharing mentioned by the participants were scored. Next, the children were asked to reflect on the design of the presentation shared (the presentation consisted of one slide with a lot of text in small Comic Sans font). The coding consisted of scoring the number of different suggestions for slide improvement. The participants were then asked to create a new slide containing an animal video and to upload this slide. The response was scored based on: (1) whether a new slide was created, (2) whether it was uploaded, and (3) whether it contained an animal video. Finally, the participants were asked to find a copyright-free image containing a polar bear and melting ice. The response was scored based on: (1) whether they searched for copyright-free images, (2) whether they found a copyright-free image, and (3) whether they uploaded the image.

During task completion, the children themselves decided when they were finished or wanted to give up on a task. The recordings of the child's online behaviour resulted in several measures of digital skills: (1) general task completion and (2) the specific skills elements that emerged from the coding (e.g. whether information was evaluated).

7.3 Procedure performance tests

The test started by asking the participating children for their gender and age. Then they were presented with the yDSI skill items selected after the cognitive interviews and the pilot surveys (see Section 7.5). The completion of this part took around 5 minutes in all countries. The test subsequently set out to assess digital skills through the tasks discussed in the prior section. The different parts of the test roughly corresponded with the different skill dimensions. The tasks took around 50 to 60 minutes to complete.

Depending on the possibilities for data collection – as COVID-19 made it impossible for some countries to conduct the performance tests at schools – the tests were held individually or in a classroom setting at school (in the latter case, several children were tested at once). The procedures followed for the two settings were as follows:

7.3.1 Classroom setting

In this set-up the tests were performed in a classroom supervised by a teacher and conducted by trained researchers. Consent forms were completed by both the child and parents before the tests took place. The researcher prepared a classroom in which 15 to 20 children could perform the test at the same time. The computers in the classroom were set up for the tests (internet access, software to create a slide and screen recording options). Before the children entered the classroom, the Qualtrics test page on which the tasks were programmed was opened on the screen.

During the classroom session, the children were welcomed and given verbal instructions about the procedure. Then, the researcher started the screen recording and test. During the tests, no encouragement was given and the researcher refrained from influencing the child's strategies. When a child finished, screen recording was stopped and the recording was saved for later analysis. The performance on the tasks was coded by a researcher watching the video using one form that combined notes on the actions and scoring of the tasks in terms of successful completion and strategies used.

7.3.2 Individual online sessions

In this set-up the child was able to complete the test at home. They used a desktop or laptop computer that was present at home and they were, thus, in most cases familiar with. Therefore, the children were doing the tasks in a setting that mirrored their normal internet use. Before a session took place, consent forms for both the child and parents were completed (and returned via email). Furthermore, it was explained that the assignments could only be performed on a computer or laptop with internet access, and that a program for creating slides should be installed (e.g. PowerPoint). It was also stressed that the assignments could not be completed on a mobile phone. The test leader used a computer for video conferencing (e.g. Teams) that allowed screen sharing and screen recording.

During the online session, the child was first welcomed by the researcher and given a verbal instruction about the procedure. After this explanation, the child was asked to enable screen sharing and open a Qualtrics page to access the assignments. The researcher enabled screen recording and the child started the test. During test completion, no encouragements were given because the pressure to succeed is already higher in an artificial, exam-style setting where the child is monitored and recorded. The researcher refrained from influencing the child's strategies to find a solution. The researcher used a form to take notes of the child's actions. Successful completion of the tasks and other indicators of performance were directly scored during the session.

7.4 Sample performance test

The performance tests were conducted in November 2020 in Estonia, Portugal, Belgium and the Netherlands.

For the purpose of validation, it is important that there is diversity amongst the participating children. Therefore, in all participating countries the children represented different genders and ages (12–17) (see Table 11). In Estonia, three classroom sessions were conducted in one school: one session with 6th grade students (mostly 12-year-olds), one session with 8th grade students (mostly 14-year-olds) and one session with 10th grade students (mostly 16-year-olds, a few aged 17 and one 18). Similarly, in Portugal three classroom sessions were conducted in one school: one session with 8th grade students (aged 12–13 and one child of 14), one session with 9th grades students (aged 14–15), and one session with 12th grades students (aged 16–17). In Belgium and the Netherlands – countries similar in terms of their education and cultural backgrounds, and so their studies were combined – 34

individual sessions were conducted. The total sample across all countries consisted of 143 children of different gender and age groups who participated in the performance test.

Table 11.	PERFORMANCE TEST SAMPLE							
		Estonia		Portugal		Belgium/ Netherlands	Total	
		N	%	N	%	N %	N	%
Gender	Boy	31	53	22	43	13 38	66	46
	Girl	25	43	29	57	21 62	75	52
	Other	2	3	0	0	0 0	2	1
Age	12–13	17	29	16	31	1 3	34	24
	14–15	23	40	17	33	10 29	50	35
	16–17	18	31	18	35	23 68	59	41
	N Total	58		51		34	143	

7.5 Results: Validation through performance tests

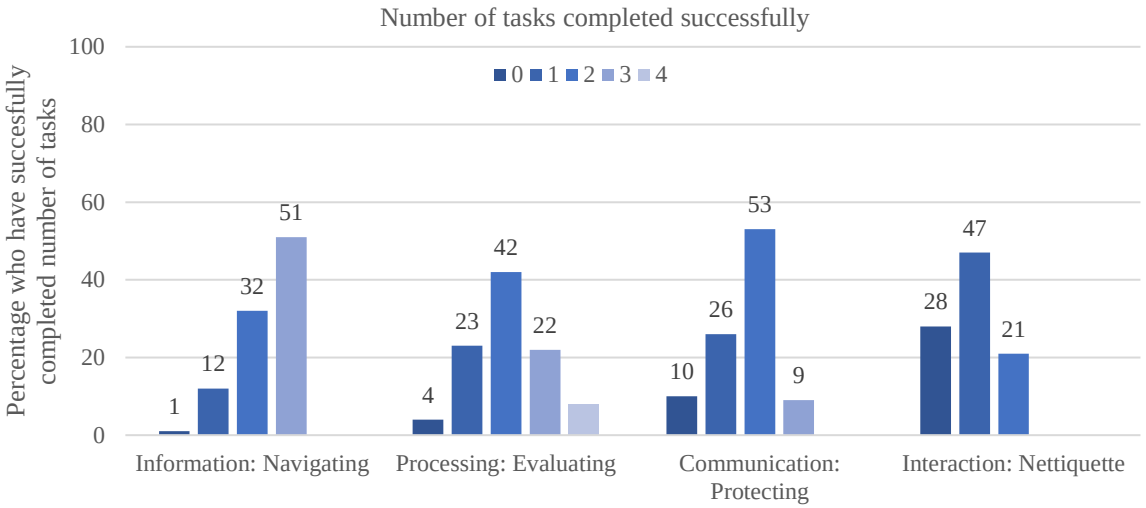
This section describes the results of the analyses of the performance tests which led to the removal or adjustment of items from the yDSI instrument that were selected based on the cognitive interviews and pilot surveys (see Tables 11 and 12).

Table 12.	yDSI SKILLS ITEMS USED IN THE PERFORMANCE TESTS
Skills dimension	Item
Technical and operational	I know how to adjust privacy settings
	I know how to turn off the location settings on mobile devices
	I know how to recognise whether a WI-FI network is safe and secure
	I know how to have the same documents, contacts and apps on all devices that I use
	I know how to delete the record of sites I have visited before
	I know how to protect a device (e.g. with a PIN, screen pattern, finger print, facial recognition)
	I know how to store photos, documents or other files in the cloud
	I know how to use private browsing
	I know how to block unwanted pop-up messages or ads
	I know how to use programming language (e.g. XML, Python)
Information navigation and processing	I know how to choose the best keywords for online searches
	I know how to find a website I have visited before
	I understand what different icons (e.g. , , ,) on apps or websites mean
	I know how to find information on a website no matter how it is designed
	I know how to use advanced search functions in search engines
	I know how to check if the information I find online is true
Communication and interaction	Depending on the situation, I know which medium or tool to use to communicate with someone (make a call, send a WhatsApp message, send an email, etc.)
	I know when I should mute myself or disable video in online interactions
	I know how to block messages from someone I don't want to hear from
	I know which images and information of me it is OK to share online
	I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters
	I know how to report negative content relating to me or a group to which I belong
	I know how to recognise when someone is being bullied online
Content creation and production	I know how to create something which combines different digital media (photo, music, video, GIF)
	I know how to edit existing digital images, music and videos
	I know how to ensure that many people will see what I put online
	I know how to change the things I put online depending on how other people react to it
	I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post)
	I know how to reference and use content covered by copyright
	I know which different types of licenses apply to online content.

Decisions around the digital skills items were made based on observation of actual performances during the taking of the tests, and on how these performances corresponded with the scores on the

skills items. The first important piece of information that the performance tasks (see Section 7.2) gave was how difficult the different tasks were for the young participants (see Figures 6 and 7).

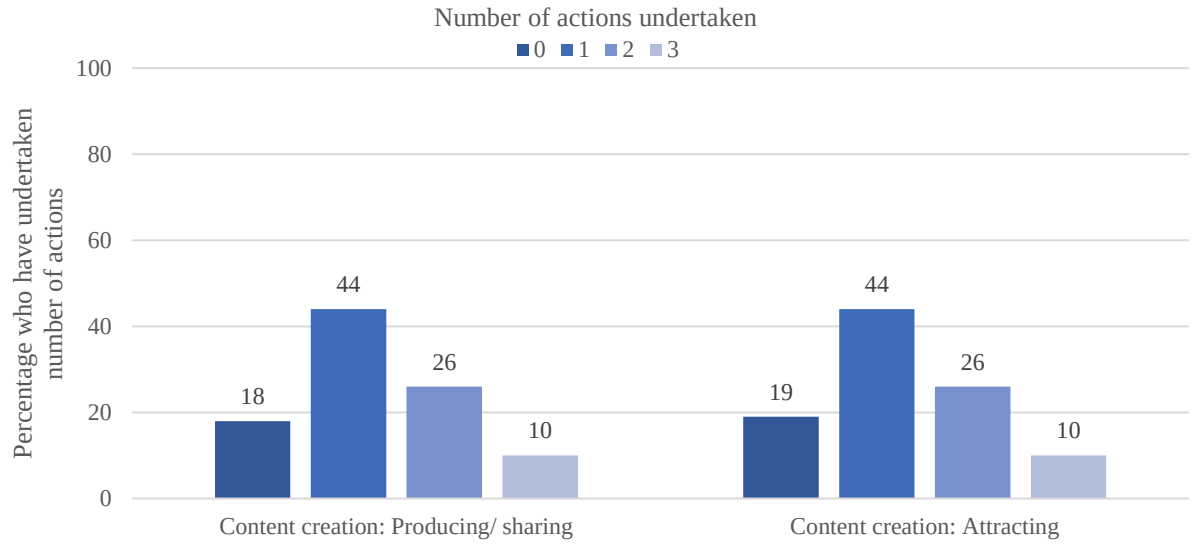
Figure 6. PERCENTAGE OF CHILDREN WHO COMPLETED INFORMATION NAVIGATION AND COMMUNICATION AND INTERACTION TASKS SUCCESSFULLY



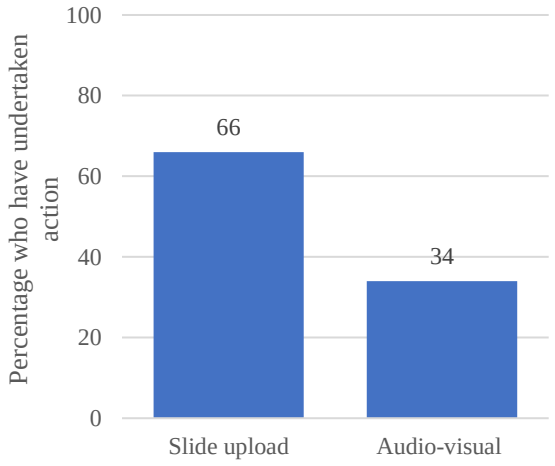
Base: N=143, all participating children

Figure 6 shows that 51% of the children were able to successfully complete all three information navigation tasks. The in-depth analysis revealed that many children faced problems limiting a search to a particular year or restricting the results to news articles. Evaluation (critical processing) was most problematic in relation to identifying fake news (24% succeeded) and recognising a scam (27% succeeded). The success rates for the privacy-related skill tasks were higher, for example 74% of the children would block an unknown sender who asked for their photo and 70% recognised that it was unwise to share a phone number publicly on a social media post. Concerning normative behaviour, 68% of the children noted that insulting language was used in a conversation between classmates.

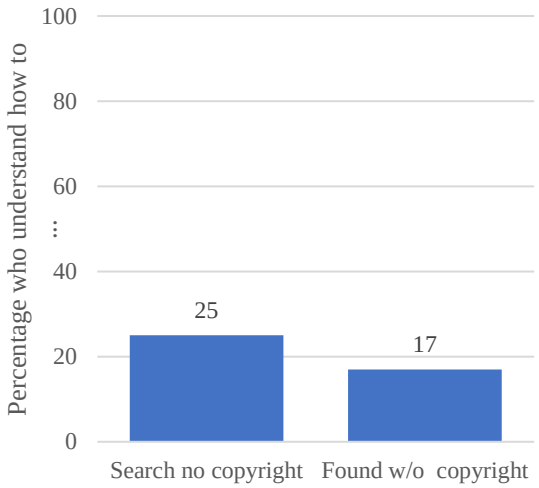
Figures 7a, b, c. **PERCENTAGE OF CHILDREN WHO COMPLETED CONTENT CREATION AND PRODUCTION TASKS SUCCESSFULLY**



a)



b)



c)

Base: N=143, all participating children

Figure 7a shows that 18% did not know how to share a file (classmate's presentation) other than via email and that 19% did not know how to make the banana GIF go viral. Furthermore, 66% created a slide with information on global warming, and 34% were able to add a video to this slide (Figure 7b). Finally, 25% knew how to search for a copyright-free photo, and 17% were actually able to find one (Figure 7c).

To decide which of the remaining items (see Tables 11 and 12) should be included in the final instrument, several performance test indicators were taken into consideration: correlations with successful completion of separate tasks and qualitative indicators around critical evaluation and explanations of actions. Based on the analyses, the decision was made to remove the item "I know how to have the same documents, contacts and apps on all devices that I use". This item is comparable to the item "I know how to store photos, documents or other files in the cloud", but scored worse in terms of correlations. The test furthermore revealed that many children thought the cloud and online drives were different things, for example, they mentioned Google drive as a way to share files, but then also mentioned the cloud. Therefore, examples were added to the item. The final item reads "I

know how to store photos, documents or other files in the cloud (e.g. Google Drive, iCloud)". Second, the results revealed little support for the item "I know how to delete the record of sites which I have visited before", which had already been flagged for deletion because of its skewness in the pilot survey. The decision to remove this item was strengthened by the fact that it loaded onto two factors in the performance test item analysis and thus had lower discriminant validity. Third, the item "I know how to block messages from someone I don't want to hear from" was deleted as it showed a negative correlation with privacy tasks in the performance test, and was also flagged for potential deletion after the pilot survey. Finally, the item "I know which different types of licenses apply to online content" was removed, since this item showed negative correlations with several performance test indicators and it had low variance. The latter was due to almost no children knowing how to search for a license-free photo.

The results for the correlations between performance test indicators and digital skills items were not as consistent as expected based on previous research. However, the results were promising for the digital knowledge items. Most of these items revealed significant correlations in the expected directions with several performance test indicators. Two items, "It is easy to distinguish content produced by bots from that produced by real people" and "Before sharing a picture that clearly shows a friend, I should always ask them for permission first" did not relate linearly with the performance measures and were both deleted. Additionally, the item "When information is backed up on the cloud, it is always encrypted" was removed as almost all children who participated in the performance test answered with "Not sure" and thus did not have much variance and consequently discriminant validity.

The performance tests were the last step in the validation and selection of the digital skills items for the yDSI instrument. The next section describes the longer and shorter versions of the yDSI and how these can be used in survey research.

8 Conclusions: Use and construction of yDSI scales

From the review of the literature, four different dimensions of skills were identified: (1) technical and operational skills; (2) information navigation and processing skills; (3) communication and interaction skills; and (4) content creation and production skills. All of these skill types have functional and critical elements that are operationalised in the items that measure the skills relevant to each dimension. The short final version of the yDSI instrument, which has the items with the best properties and which will be used in the ySKILLS longitudinal panel survey, can be found at the beginning of this report (see Section 2).

This section presents (1) the statistical properties of this short yDSI instrument; (2) the long version of the yDSI; and (3) guidelines on how to create and use composite scores that measure the different underlying dimensions. The use of the longer version is recommended if studies have the space to include all items. The short version is recommended for those studies that have less space. The shorter yDSI was validated through cognitive interviews, pilot surveys and performance tests, and the items on these have high discriminant validity across skills dimensions. It is therefore not recommended to replace items on the shorter version of the yDSI with items from the longer version.

8.1 The properties of the short version of the yDSI scale

The short scale for the yDSI (see Section 2) has good convergent properties. The items that are part of a particular dimension are strongly related to each other and load highly on their dimension when the dimensions are analysed in isolation (see Section 6.4).

The best items for the short scales were selected based on the findings from the cognitive interviews and the data gathered in performance tests as well as the statistical properties of the items and scales in the pilot surveys. The items initially selected for the short scales (see Appendix F) were the same as presented at the beginning of the report with the exception of the items that made up the dimension of the information navigation and processing short scale. On CFA analyses of the statistical properties of the yDSI short scale, the item "I understand what different *icons (* , , , *)* *in* an app or website mean" turned out to have a low loading on the information navigation and processing dimension (r=0.29) and a higher loading on the communication and interaction dimension (r=0.33). The low loading suggests limited convergent validity and the cross-loadings (i.e. high loadings on other skills dimensions) limited discriminant validity of this item. On re-analyses of the pilot survey data, the statistical properties of the scale including the item "I know how to figure out if a website can be trusted" were better. This item was not validated through the performance tests but, considering the lack of issues in the cognitive interviews and the better statistical properties, the decision was made to keep the latter item. All final yDSI short scales have high internal reliability, with Cronbach's alphas between 0.79 and 0.81 (see Table 13).

These are lower than those in the long version (see Section 6.4.1 and Appendix E), but this is to be expected from a shorter scale.

Table 13. FACTOR LOADINGS (ROTATED) FOR THE SHORT yDSI ON THE FOUR CONCEPTUALISED SKILLS DIMENSIONS

	Skills dimension			
Skills item	T&O	INO&P*	C&I	CC&P
I know how to adjust privacy settings	0.45	0.20	0.25	0.18
I know how to turn off the location settings on mobile devices	0.48	0.17	0.31	0.13
I know how to protect a device	0.59	0.05	0.41	–0.01
I know how to store photos, documents or other files in the cloud	0.58	0.20	0.18	0.22
I know how to use private browsing	0.55	0.21	0.17	0.08
I know how to block unwanted pop-up messages or ads	0.43	0.32	0.10	0.22
I know how to choose the best keywords for online searches	0.28	0.39	0.35	0.16
I know how to find a website I have visited before	0.27	0.34	0.35	0.15
I know how to find information on a website no matter how it is designed	0.20	0.44	0.25	0.25
I know how to use advanced search functions in search engines	0.25	0.48	0.13	0.30
I know how to check if the information I find online is true	0.21	0.64	0.17	0.22
I know how to figure out if a website can be trusted	0.20	0.60	0.19	0.26
Depending on the situation, I know which medium or tool to use to communicate with someone	0.22	0.14	0.58	0.13
I know when I should mute myself or disable video in online interactions	0.25	0.13	0.52	0.20
I know which images and information of me it is okay to share online	0.21	0.13	0.58	0.13
I know when it is appropriate and when it is not appropriate to use emoticons, text speak and capital letters	0.14	0.12	0.58	0.17
I know how to report negative content relating to me or a group to which I belong	0.19	0.17	0.49	0.19
I know how to recognise when someone is being bullied online	0.09	0.22	0.40	0.27
I know how to create something which combines different digital media	0.20	0.17	0.12	0.65
I know how to edit existing digital images, music and videos	0.19	0.10	0.11	0.71
I know how to ensure that many people will see what I put online	0.00	0.21	0.17	0.56
I know how to change the things I put online depending on how other people react to it	0.09	0.13	0.21	0.53
I know how to distinguish sponsored and non-sponsored content online	0.15	0.19	0.30	0.42
I know how to reference and use content covered by copyright	0.09	0.29	0.15	0.49
Cronbach's α =	0.79	0.81	0.80	0.79

Base: All participants in the pilot survey, N=4,238. Notes: T&O = technical and operational; INO&P* = information navigation and processing; C&I = communication and interaction; CC&P = content creation and production. * See notes on this dimension below.
Factor loadings calculated using maximum likelihood estimation and varimax rotation based on the average scores (see Section 8.3). Examples have been left out of the item formulation for clarity purposes.

The factor solution presented in Table 13 shows that almost all items uniquely load on their associated skills dimension with high loadings on the assigned dimension only. One exception can be found on the technical and operational skills dimension; the item "I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition)" which loaded onto both the technical and operational (r=0.60) and the communication and interaction (r=0.41) skills dimensions. However, the loading on the second is much lower than on the first. and so this does not cause significant problems for discriminant validity.

The other two exceptions were on the information navigation and processing dimension: "I know how to find a website I have visited before" (r=0.34) and "I know how to choose the best keywords for online searches" (r=0.39) both load lower onto the information navigation and processing dimension and have cross-loadings on the communication and interaction dimensions (r=0.35). The properties of these items improve when the item that was added after the performance tests (i.e., "I know how to figure out if a website can be trusted") is left off the information navigation and processing dimension. The loading of the "I know how to choose the best keywords for online searches" shifts upward to 0.41 and the loading for "I know how to find a website I have visited before" to 0.35. Since six items was the minimum required for the yDSI longitudinal panel survey and since items can be deleted but not added to analyses after fieldwork, the decision was made to keep the item "I know how to figure out if a website can be trusted" on the information navigation and processing dimension. Further testing of convergent and discriminant validity of the items will take place with the larger sample of the ySKILLS longitudinal panel survey with 12- to 17-year-olds.[4]

The convergent and discriminant validity of the final short yDSI scale is clear from the correlations between the dimensions. The dimensions are expected to be correlated because they measure the underlying overall skill construct. However, the correlations are not as high as to cause significant concern around multi-collinearity in multi-variate analyses and display sufficient discriminant validity (see Table 14).

Table 14. CORRELATIONS BETWEEN SKILLS DIMENSIONS FOR THE SHORT yDSI

		(1)	(2)	(3)	(4)
(1)	Technical and operational skills	1			
(2)	Information navigation and processing skills	0.65	1		
(3)	Communication and interaction skills	0.57	0.60	1	
(4)	Content creation and production skills	0.45	0.57	0.51	1
(5)	Programming skills*	0.19	0.24	0.07	0.27

Note: * Programming skills is a single item indicator.
Base: All participants in the pilot survey, N=4,238

The strongest correlations can be found between the technical and operational and the information navigation and processing dimensions (r=0.65; see Table 14). There is a surprisingly low correlation between technical and operational and content creation and production skills (r=0.45); this is lower than found in previous research. This higher discriminatory validity might be caused by the inclusion of new content creation and production items that are based on the critical evaluation of content (including safety) as well as more everyday content creation practices. The justification of including

[4] It is notable that the information navigation and processing dimension unexpectedly caused the most problems in all validations (i.e., cognitive interviews, survey pilots and performance tests). This might be due to the expanded multi-dimensional nature of the measure that was not incorporated into existing research. WP6 of the ySKILLS project will take a deep dive into misinformation in particular; this might also be useful to further test the items on this dimension.

programming as a separate dimension is clear from the low correlations between this indicator and the four skills dimensions.

8.2 Long version of the yDSI scales

Most surveys will only have space for the short version of the yDSI (see Section 2). In this section, there is a longer version for researchers who have space in their questionnaires and who want to have a more inclusive range of skills measures.[5]

The questions and answer scales for the long version of the yDSI are the same as for the short yDSI scale.

8.2.1 Long version of the yDSI digital skills instrument

The question used for the digital skills items in the long version of the yDSI is:

> Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones or computers. Reply thinking about **how true this would be of you** if you had to do it now, on your own. **If you do not understand** what the question is asking, tick the box "I do not understand what you mean by this".

The following should be added to the instructions of the communication and interaction and content creation and production skills item blocks only (see Tables 17 and 18).

> Sometimes there are various examples given; only select "Very true of me" if all of the examples apply to what you do or know.

The answer scale used for the digital skills items in the long version of the yDSI is:

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	I do not want to answer (99)

Note I: The answer categories should be presented in this order and the scores on the Likert scale (1 through 5, 66 and 99) should not be presented to the participants; these are only included for coding and analyses.
Note II: For analyses, the "I do not understand what you mean by this" answer category is part of the skills scale and should be converted to zero because it indicates a lack of knowledge as well as a lack of skill, and thus ranks below not having a skill (see Section 8.3).

[5] Since the long version includes all the items from the short version, the correlations between the composite scores for the dimensions are very high. Including more items will reduce error in measurement and allow researchers to obtain an even deeper insight into the variety of digital skills that their respondents have.

Table 15.	FINAL LONG VERSION OF THE yDSI – TECHNICAL AND OPERATIONAL SKILLS ITEMS
Skills dimension	Item
Technical and operational skills	I know how to adjust privacy settings
	I know how to turn off the location settings on my mobile devices
	I know how to recognise whether a WIFI network is safe and secure
	I know how to connect devices to each other using Bluetooth or wireless connections
	I know how to have the same documents, contacts and apps on all devices that I use
	I know how to protect a mobile device (e.g. with a PIN, a screen pattern or a finger print)
	I know how to store photos, documents, contacts or other files in the cloud (e.g. Google Drive, iCloud)
	I know how to delete the record of sites which I have visited before
	I know how to use private browsing
	I know how to block unwanted pop-up messages or ads
Programming*	I know how to use programming language (e.g. XML, Python, Java, C++)

Note: * Programming is included as a single item; it does not load onto the skills dimensions as the other items do, but is considered important in the literature and interventions, and is thus included.

Table 16.	FINAL LONG VERSION OF THE yDSI – INFORMATION NAVIGATION AND PROCESSING SKILLS ITEMS
Skills dimension	Item
Information navigation and processing skills*	I know how to choose the best keywords for online searches
	I know how to change how I search for things online, for example, if I do not obtain the result I want
	I know how to find a website I have visited before
	I know how to find information on a website no matter how it is designed
	I understand what different icons (   ) in an app or website mean
	When I have a question, I am able to find information online that is relevant to answering it
	I know how to use advanced search functions in search engines
	I know how to check if the information I find online is true
	I know how to figure out if a website can be trusted

Note: See Section 8.1 for a comment on the statistical properties of the items on this dimension and how they should be used in future research.

The items in black in tables 15 and 16, should always be included when measuring digital skills. The items in red were removed after the cognitive interviews, the pilot surveys and performance tests to create a shorter scale but can be used for surveys that have more space and are able to include a longer version of the yDSI.

Table 17.	FINAL LONG VERSION OF THE yDSI – COMMUNICATION AND INTERACTION SKILLS ITEMS
Skills dimension	Item
Communication and interaction skills	Depending on who I want to communicate with, I know which medium or tool to use (make a call, send a WhatsApp message, send an email, etc.)
	I know when to mute myself or disable video in online interactions
	I know how to block messages from someone I don't want to hear from
	I know which images and information of me it is OK to share online
	I know when I should not post pictures or videos of others online
	I know how to change who I share content with (e.g. just friends, friends of friends or make it public)
	I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis) or text speak or capital letters
	I know how to make my comments and behaviours appropriate to the online situation
	I know how to report negative content relating to me or a group to which I belong
	I know how to recognise when someone is being bullied online

Table 18.	FINAL LONG VERSION OF THE yDSI – CONTENT CREATION AND PRODUCTION SKILLS ITEMS
Skills dimension	Item
Content creation and production skills	I know how to create something which incorporates different digital media (images, music, video, GIFs) I know how to edit existing online images, music and videos
	I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people
	I know how to ensure that many people will see what I put online
	I know how to use filters and other tools to make a photo or video look more attractive
	I know how to change the things I put online depending on how other people react to it
	I know how to distinguish sponsored and non-sponsored content
	I know when I am allowed to use content covered by copyright
	I know which different types of licenses apply to online content

The items in black in tables 17 and 18 should always be included when measuring digital skills. The items in red were removed after the cognitive interviews, the pilot surveys and performance tests to create a shorter scale but can be used for surveys that have more space and are able to include a longer version of the yDSI.

8.2.2 Long version of the yDSI digital knowledge instrument

The question for the digital knowledge items on the long version of the yDSI is:

> To what extent are the following statements about technologies such as the internet and mobile phones **true or not true**? If you are not sure, please let us know.

This question is accompanied by the following answer scale:

Definitely not true (1)	Definitely true (2)	I'm not sure (3)	I do not want to answer (99)

Note: The answer categories should be presented in this order and the scores on the scale (1 through 3 and 99) should not be presented to the participants; these are only included for coding and analyses.

The digital knowledge skills items for the long version of the yDSI are presented in Table 19.

Table 19. FINAL LONG VERSION OF THE yDSI – DIGITAL KNOWLEDGE ITEMS

Skills dimension	Item
Information navigation and processing	The first search result is always the best information source
	Everyone gets the same information when they search for things online
	When information is backed up on the cloud, it is always encrypted
Communication and interaction	Whether I like or share a post can have a negative impact on others
	The first post I see on social media is the last thing that was posted by one of my contacts
	A post that a friend shared with you is more trustworthy than other posts
	Before sharing a picture that clearly shows a friend, I should always ask them for permission first
	If someone has posted something online without making it private, it is okay to share or forward it without asking
	Negative comments you make online about people are less hurtful than saying them face-to-face
Content creation and production	Using hashtags increases the visibility of a post
	Companies pay ordinary people to use their products in videos and content they create
	Companies use the information you post on your social media profile to market their products and services

Note: The items in black should always be included when measuring digital skills. The items in red were removed after the cognitive interviews, the pilot surveys and performance tests to create a shorter scale but can be used for surveys that have more space and are able to include a longer version of the yDSI.

8.3 Guidelines for the creation of composite scales for analysis in the survey

The items and scales that were designed in this report can be used in different ways for analysis in projects interested in measuring digital skills amongst young people. However, before creating composite scales, the "I don't know what you mean by this" (IDKWYM) option has to be converted to 0. IDKWYM in skills measurement does not mean "I do not know whether or not I have the skill"; instead, it represents the answer "I do not know what this skill is that you are referring to". This interpretation was validated through cognitive interviews (van Deursen et al., 2016).[6] The score 0 is given to IDKWYM because this answer option represents a lack of knowledge as well as a lack of skill and thus comes lower on the skills level scale than not having the skill but knowing what it is (i.e. "Not at all true of me" answer category). The IDKWYM option should still be presented at the end of the scale in the survey to prevent social desirability bias and coded with a higher double-digit number (e.g. 66) before conversion to avoid confusion amongst researchers (van Deursen et al., 2016). After conversion, the items can be subjected to traditional Likert scale scoring with a scale from 0 to 5 (0 = "IDKWYM", 1 = "Not at all true of me" and 5 = "Very true of me").

There are three alternative ways of designing composite scales, both as an assessment of digital skills in general, and of each skills dimension separately:

1. The **average of the scores** on the items in each dimension can be calculated, resulting in a composite score of the average skill level from 0 to 5 for each dimension (or for an overall digital skill level). Scores for each item can also be added up, resulting in a score from 0 to 30 for each skills dimension (and 0 to 5 for programming). This would then result in a scale for overall level of digital skills from 0 to 125 (overall skills includes programming).
2. Composite scores of **high skill levels** can also be created for each dimension counting the number of items for which respondents report the highest level of skill (i.e., "Very true of me") thus creating scores from 0 to 6 out of 6 for each skills dimension, and from 0 to 25 out of 25 for overall digital skills (overall skills includes programming).
3. For better understanding, the **proportion of skills at a high level** can be calculated. This is achieved by dividing the high skill score by the number of items in the dimension (i.e., 6 for dimensions or 25 for the overall score). A person would have a score of 0% if they had none of the skills at a high level (i.e., no scores of 5, "Very true of me") and a score of 100% if they had 6 skills at a high level for the separate skills dimensions or 25 skills at a high level for the overall skills score (i.e., they answered "Very true of me" to all skills items). Obviously, the proportion calculation makes no sense for the single item programming skill.

The composite score options 2 and 3 have the added advantage that knowledge-based items can also be incorporated. A correct answer for these items would count as 1 (as seen in Shadel, Pak, & Sauer, 2014). This would bring up the maximum composite scores of the dimensions of information navigation and processing, communication and interaction and content creation and production to 8 (or divided by 8), and the maximum total score to 30 (or divided by 30). In addition, knowledge-based items can be kept as a separate score out of 6 (or divided by 6) using this same scoring method as indicated under option (4).[7] For ySKILLS the decisions on which composite scores to create will be made on the basis of the analysis of the longitudinal panel survey as part of this study.

[6] This answer category is not the same as an "I don't know option" in an opinion survey or standard agree/disagree Likert scales, which reflects a person not having an opinion (rather than a low opinion). In these surveys it is used to prevent people with no opinion on a matter selecting the neither agree or disagree or neutral option.

[7] The literature on such knowledge scores and questionnaires is predominantly found in clinical fields, and addresses the testing of patients or general population knowledge of symptoms or side-effects of specific health conditions of diseases (e.g. White et al., 2006). As this strand of literature is not related or relevant to anything pertaining to the ySKILLS project, interested parties are invited to conduct their own investigation on the matter.

Acknowledgements
The authors thank the reviewers Verónica Donoso, Natalia Waechter, Leen d'Haenens and Willem Joris for their critical reading and useful suggestions that helped improve and clarify this report.

References

ACARA (Australian Curriculum Assessment and Reporting Authority) (2020). National Assessment Program – ICT literacy. Available at: https://acara.edu.au/.

Aesaert, K., & van Braak, J. (2014). Exploring factors related to primary school pupils' ICT self-efficacy: A multilevel approach. *Computers in Human Behavior*, *41*, 327–41. doi:10.1016/j.chb.2014.10.006.

Aesaert, K., van Braak, J., van Nijlen, D., & Vanderlinde, R. (2015). Primary school pupils' ICT competences: Extensive model and scale development. *Computers & Education*, *81*, 326–44. doi:10.1016/j.compedu.2014.10.021.

Aesaert, K., Voogt, J., Kuiper, E., & van Braak, J. (2017). Accuracy and bias of ICT self-efficacy: An empirical study into students' over- and underestimation of their ICT competences. *Computers in Human Behavior*, *75*, 92–102. doi:10.1016/j.chb.2017.05.010.

Alkan, M., & Meinck, S. (2016). The relationship between students' use of ICT for social communication and their computer and information literacy. *Large-Scale Assessments in Education*, *4*(15). doi:10.1186/s40536-016-0029-z.

Areepattamannil, S., & Khine, M. S. (2017). Early adolescents' use of information and communication technologies (ICTs) for social communication in 20 countries: Examining the roles of ICT-related behavioral and motivational characteristics. *Computers in Human Behavior*, *73*, 263–72. doi:10.1016/j.chb.2017.03.058.

Balea, B. (2016). The role of smartphones in increasing digital and social inequalities among Romanian children. *Journal of Comparative Research in Anthropology and Sociology*, *7*(2), 1–20. Available at: www.ceeol.com/search/article-detail?id=552530

Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, *57*(2), 218–59. doi:10.1108/EUM0000000007083.

Broos, A., & Roe, K. (2006). The digital divide in the playstation generation: Self-efficacy, locus of control and ICT adoption among adolescents. *Poetics*, *34*(4–5), 306–17. doi:10.1016/j.poetic.2006.05.002.

Burton, R. F. (2004). Multiple choice and true/false tests: Reliability measures and some implications of negative marking. *Assessment & Evaluation in Higher Education*, *29*(5), 585–95. doi:10.1080/02602930410001689153.

Burton, R. F. (2005). Multiple-choice and true/false tests: Myths and misapprehensions. *Assessment & Evaluation in Higher Education*, *30*(1), 65–72. doi:10.1080/0260293042003243904.

Byrne, B. M. (2001). *Structural equation modeling with AMOS*. Mahwah, NJ: Lawrence Erlbaum Associates.

Byrne, B. M., & Stewart, S. M. (2006). The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling*, *13*(2), 287–321.

Byrne, B. M., Shavelson, R., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–66.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modelling*, *14*(3), 464–504.

Chepp, V., & Gray, C. (2014). Foundations and new directions. In K. Miller, S. Willson, V. Chepp, & J. L. Padilla (eds) *Cognitive interviewing methodology* (pp.7–14). Hoboken, NJ: John Wiley & Sons.

Christoph, G., Goldhammer, F., Zylka, J., & Hartig, J. (2015). Adolescents' computer performance: The role of self-concept and motivational aspects. *Computers & Education*, *81*, 1–12. doi:10.1016/j.compedu.2014.09.004.

Cortesi, S. C., Hasse, A., Lombana-Bermudez, A., Kim, S., & Gasser, U. (2020). *Youth and digital citizenship+ (plus): Understanding skills for a digital world*. Available at: https://cyber.harvard.edu/publication/2020/youth-and-digital-citizenship-plus.

Cronbach, L. J. (1941). An experimental comparison of the multiple true-false and multiple multiple-choice tests. *Journal of Educational Psychology*, *32*(7), 533–43. doi:10.1037/h0058518.

Department for Education (2018). *Essential Digital Skills Framework.* Available at: www.gov.uk/government/publications/essential-digital-skills-framework.

Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, *26*(1), 1–22. doi:10.3102/01623737026001001.

DiSTO (2020). From Digital Skills to Tangible Outcomes. Available at: www.lse.ac.uk/media-and-communications/research/research-projects/disto

Durndell, A., & Haag, Z. (2002). Computer self efficacy, computer anxiety, attitudes towards the Internet and reported experience with the Internet, by gender, in an East European sample. *Computers in Human Behavior*, *18*(5), 521–36.

EU Kids Online (2019). Enhancing knowledge of European children's online opportunities, risks and safety. Available at: www.lse.ac.uk/media-and-communications/research/research-projects/eu-kids-online

European Commission (2019). *The 2018 International Computer and Information Literacy Study (ICILS): Main findings and implications for education policies in Europe*. Available at: https://ec.europa.eu/education/resources-and-tools/document-library/the-2018-international-computer-and-information-literacy-study-icils-main-findings-and-implications-for-education-policies-in-europe_en

European Commission (2020a). DigComp. Available at: https://ec.europa.eu/jrc/en/digcomp

European Commission (2020b). Mapping Digital Inclusion landscape to support Cohesion and Integration (MEDICI). Available at: https://medici-project.eu/

Eurostat (2019). *Community survey on ICT usage in households and by individuals.* Available at: https://circabc.europa.eu/sd/a/28e0bbb8-af94-455d-93de-94985f62ade7/Model%20Questionnaire%202019%20v%201.1%20-%20after%20WG.pdf

Frisbie, D. A. (1973). Multiple choice versus true-false: A comparison of reliabilities and concurrent validities. *Journal of Educational Measurement*, *10*(4), 297–304. Available at: www.jstor.org/stable/1434001

Gastelu, C. A. T. (2013). Digital skills in Mexican children. In L. Zhang, X. Li, & J. Chen (eds) *Proceedings of the 2013 International Conference on Information, Business and Education Technology* (Vol. 26, pp.1062–5).

Gecas, V. (1989). The social psychology of self-efficacy. *Annual Review of Sociology*, *15*(8), 291–316.

Global Kids Online (2020). Available at: http://globalkidsonline.net/.

Grimm, P. (2010). Social desirability bias. In J. Sheth & N. Malhotra (eds) *Wiley International Encyclopedia of Marketing* (part 2). New York: Wiley & Sons.

Gui, M., & Argentin, G. (2011). Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. *New Media & Society*, *13*(6), 963–80. doi:10.1177/1461444810389751.

Haddon, L., Cino, D., Doyle, M. A., Livingstone, S., Mascheroni, G., & Stoilova, M. (2020). *Children's and young people's digital skills: A systematic evidence review*. Leuven, KU Leuven: ySKILLS. Available at: https://zenodo.org/record/4274654#.X-pMceSWysc

Hargittai, E. (2005). Survey measures of web-oriented digital literacy. *Social Science Computer Review*, *23*(3), 371–9.

Helsper, E. J. (2017). The social relativity of digital exclusion: Applying relative deprivation Theory to digital inequalities. *Communication Theory*, *27*(3), 223–42. doi:10.1111/comt.12110.

Helsper, E. J. (2021). *Digital disconnect: The social causes and consequences of digital inequalities.* London: Sage.

Helsper, E. J., & van Deursen, A. J. A. M. (2018). ICT skills for the future. In ITU (International Telecommunications Union), *Measuring the Information Society* (pp.21–50). Geneva: ITU.

Herde, C. N., Lievens, F., Solberg, E. G., Harbaugh, J. L., Strong, M. H., & Burkholder, G. J. (2019). Situational judgment tests as measures of 21st century skills: Evidence across Europe and Latin America. *Journal of Work and Organizational Psychology–Revista De Psicologia Del Trabajo Y De Las Organizaciones*, *35*(2), 65–74. doi:10.5093/jwop2019a8.

Hinostroza, J. E., Matamala, C., Labbe, C., Claro, M., & Cabello, T. (2015). Factors (not) affecting what students do with computers and internet at home. *Learning Media and Technology*, *40*(1), 43–63. doi:10.1080/17439884.2014.883407.

Hohlfeld, T. N., Ritzhaupt, A. D., & Barron, A. E. (2013). Are gender differences in perceived and demonstrated technology literacy significant? It depends on the model. *Etr&D–Educational Technology Research and Development*, *61*(4), 639–63. doi:10.1007/s11423-013-9304-7.

Huang, K. T., Cotten, S. R., & Rikard, R. V. (2017). Access is not enough: The impact of emotional costs and self-efficacy on the changes in African-American students' ICT use patterns. *Information Communication & Society*, *20*(4), 637–50.

ITU (International Telecommunication Union) (2017). *The ICT Development Index (IDI): Conceptual framework and methodology*. Available at: www.itu.int/en/ITU-D/Statistics/Pages/publications/mis2017/methodology.aspx

Jobe, J. B., & Mingay, D. J. (1989). Cognitive research improves questionnaires. *American Journal of Public Health*, *79*(8), 1053–55. doi:10.2105/ajph.79.8.1053.

Kaarakainen, M. T. (2019). ICT intentions and digital abilities of future laborm entrants in Finland. *Nordic Journal of Working Life Studies*, *9*(2), 105–26.

Kim, H. S., Kil, H. J., & Shin, A. (2014). An analysis of variables affecting the ICT literacy level of Korean elementary school students. *Computers & Education*, *77*, 29–38. doi:10.1016/j.compedu.2014.04.009.

Kim, J., & Lee, W. (2013). Meanings of criteria and norms: Analyses and comparisons of ICT literacy competencies of middle school students. *Computers & Education*, *64*, 81–94. doi:10.1016/j.compedu.2012.12.018

King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing*, *17*(2), 79–103.

Kolle, S. R. (2017). Global research on information literacy: A bibliometric analysis from 2005 to 2014. *Electronic Library*, *35*(2), 283–98. doi:10.1108/el-08-2015-0160.

Larson, K. E., & Bradshaw, C. P. (2017). Cultural competence and social desirability among practitioners: A systematic review of the literature. *Children and Youth Services Review*, *76*, 100–11. doi:10.1016/j.childyouth.2017.02.034.

Lau, W. W. F., & Yuen, A. H. K. (2015). Factorial invariance across gender of a perceived ICT literacy scale. *Learning and Individual Differences*, *41*, 79–85. doi:10.1016/j.lindif.2015.06.001.

Lazonder, A. W., Walraven, A., Gijlers, H., & Janssen, N. (2020). Longitudinal assessment of digital literacy in children: Findings from a large Dutch single-school study. *Computers & Education*, 1*43*. doi:10.1016/j.compedu.2019.103681.

Le, A. V., Do, D. L., Pham, D. Q., Hoang, P. H., Duong, T. H., Nguyen, H. N., et al. (2019). Exploration of youth's digital competencies: A dataset in the educational context of Vietnam. *Data*, *4*(2). doi:10.3390/data4020069.

Lee, M. F. (2018). *Digital skills measurement: A study on the Malaysian youth*. In IEEE Conference on e-Learning, e-Management and e-Services (IC3e).

Li, Y., & Ranieri, M. (2010). Are "digital natives" really digitally competent? A study on Chinese teenagers. *British Journal of Educational Technology*, *41*(6), 1029–42. doi:10.1111/j.1467-8535.2009.01053.x.

Looker, E. D., & Naylor, T. D. (2010). *Digital diversity: Youth, equity, and information technology*. Waterloo, ON: Wilfried Laurier University Press.

Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C., & Moe, A. (2018). Internet source evaluation: The role of implicit associations and psychophysiological self-regulation. *Computers & Education*, *119*, 59–75. doi:10.1016/j.compedu.2017.12.009.

Moto, S., Ratanaolarn, T., Tuntiwongwanich, S., & Pimdee, P. (2018). A Thai junior high school students' 21st century information literacy, media literacy, and ICT literacy skills factor analysis. *International Journal of Emerging Technologies in Learning*, *13*(9), 87–106. doi:10.3991/ijet.v13i09.8355.

NCES (National Center for Education Statistics) (2018). *International Computer and Information Literacy Study: 2018 results*. Available at: https://nces.ed.gov/surveys/icils/

Net Children Go Mobile. (2020). Questionnaire available at: https://centridiricerca.unicatt.it/osscom.

Nygren, T., & Guath, M. (2019). Swedish teenagers' difficulties and abilities to determine digital news credibility. *Nordicom Review*, *40*(1), 23–42. doi:10.2478/nor-2019-0002.

OECD (2020a). Programme for International Student Assessment (PISA). Available at: https://pisadataexplorer.oecd.org/ide/idepisa/

OECD (2020b). Programme for the International Assessment of Adult Competencies (PIAAC). Available at: www.oecd.org/skills/piaac/

Polizzi, G. (2020a). Digital literacy and the national curriculum for England: Learning from how the experts engage with and evaluate online content. *Computers & Education*, 103859. doi:10.1016/j.compedu.2020.103859.

Polizzi, G. (2020b). Digital literacy in theory and practice: Learning from how experts and advocates engage in civic life (PhD). London: London School of Economics and Political Science.

Ponte, C. (2019). Challenging online situations reported by Italian and Portuguese children in 2018. *Revista Mediterranea Comunicacion–Journal of Communication*, *10*(2), 165–78. doi:10.14198/medcom2019.10.2.3.

Porat, E., Blau, I., & Barak, A. (2018). Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. *Computers & Education*, *126*, 23–36. doi:10.1016/j.compedu.2018.06.030.

Richter, T., Naumann, J., & Groeben, N. (2001). The computer literacy inventory (INCOBI): An instrument for the assessment of computer literacy and attitudes toward the computer in university students of the humanities and the social sciences. *Psychologie in Erziehung und Unterricht*, *48*(1), 1–13.

Ridolfo, H., & Schoua-Glusberg, A. (2011). Analyzing cognitive interview data using the constant comparative method of analysis to understand cross-cultural patterns in survey data. *Field Methods*, *23*(4), 420–38. doi:10.1177/1525822x11414835.

Robinson, L. H., & Thoms, K. J. (2001). A longitudinal study of college student computer knowledge. *Journal of Computer Information Systems*, *42*(1), 9–12.

Rodríguez-de-Dios, I., Igartua, J. J., & González-Vázquez, A. (2016). Development and validation of a digital literacy scale for teenagers. Paper presented at the Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality, Salamanca, Spain. doi:10.1145/3012430.3012648.

Rodríguez-de-Dios, I., van Oosten, J. M. F., & Igartua, J. J. (2018). A study of the relationship between parental mediation and adolescents' digital skills, online risks and online opportunities. *Computers in Human Behavior*, *82*, 186–98. doi:10.1016/j.chb.2018.01.012.

Saranto, K., & Hovenga, E. J. S. (2004). Information literacy – What it is about? Literature review of the concept and the context. *International Journal of Medical Informatics*, *73*(6), 503–13. doi:10.1016/j.ijmedinf.2004.03.002.

Scherer, R., Rohatgi, A., & Hatlevik, O. E. (2017). Students' profiles of ICT use: Identification, determinants, and relations to achievement in a computer and information literacy test. *Computers in Human Behavior*, *70*, 486–99. doi:10.1016/j.chb.2017.01.034.

Schmittlein, D. C., & Morrison, D. G. (1983). Measuring miscomprehension for televised communications using true-false questions. *Journal of Consumer Research*, *10*(2), 147–56. Available at: www.jstor.org/stable/2488920

Selber, S. A. (2004). Reimagining the functional side of computer literacy. *College Composition and Communication*, *55*(3), 470–503. doi:10.2307/4140696.

Shadel, D., Pak, K., & Sauer, J. (2014). *Caught in the scammer's net: Risk factors that may lead to becoming an internet fraud victim*. Available at: www.aarp.org/content/dam/aarp/research/surveys_statistics/econ/2014/Caught-Scammer-Net-Risk-Factors-Internet-Fraud-Victims.doi.10.26419%252Fres.00076.001.pdf

Shank, D. B., & Cotten, S. R. (2014). Does technology empower urban youth? The relationship of technology use to self-efficacy. *Computers & Education*, *70*, 184–93. doi:10.1016/j.compedu.2013.08.018.

Siddiq, F., Hatlevik, O. E., Olsen, R. V., Throndsen, I., & Scherer, R. (2016). Taking a future perspective by learning from the past – A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educational Research Review*, *19*, 58–84. doi:10.1016/j.edurev.2016.05.002.

Silvera, D., Martinussen, M., & Dahl, T. I. (2001) The Tromsø Social Intelligence Scale, a self-report measure of social intelligence. *Scandinavian Journal of Psychology*, *42*(4), 313–19. doi:10.1111/1467-9450.00242.

Smith, H., & Pettigrew, T. (2015) Advances in relative deprivation theory and research. *Social Justice Research*, *28*(1), 1–6. doi:10.1007/s11211-014-0231-5.

Sorgo, A., Bartol, T., Dolnicar, D., & Boh Podgornik, B. (2017). Attributes of digital natives as predictors of information literacy in higher education. *British Journal of Educational Technology*, *48*(3), 749–67. doi:10.1111/bjet.12451.

Spenner, K. I. (1990). Skill: Meanings, methods, and measures. *Work and Occupations*, *17*(4), 399–421. doi:10.1177/0730888490017004002.

Tondeur, J., Sinnaeve, I., van Houtte, M., & van Braak, J. (2011). ICT as cultural capital: The relationship between socioeconomic status and the computer-use profile of young people. *New Media & Society*, *13*(1), 151–68. doi:10.1177/1461444810369245.

UNESCO (2020). *Digital Kids Asia Pacific (DKAP)*. Available at: https://bangkok.unesco.org/content/digital-kids-asia-pacific-insights-childrens-digital-citizenship

van Deursen, A. J. A. M. (2020). Digital inequality during a pandemic: Quantitative study of differences in COVID-19-related internet uses and outcomes among the general population. *Journal of Medical Internet Research*, *22*(8). doi:10.2196/20073.

van Deursen, A. J. A. M., & Helsper, E. J. (2017). Collateral benefits of internet use: Explaining the diverse outcomes of engaging with the Internet. *New Media & Society*, *20*(7), 2333–51. doi:10.1177/1461444817715282.

van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2010). Measuring internet skills. *International Journal of Human–Computer Interaction*, *26*(10), 891–916. doi:10.1080/10447318.2010.496338.

van Deursen, A. J. A. M., Helsper, E. J., & Eynon, R. (2014). *Measuring digital skills: From digital skills to tangible outcomes*. Project report. London: LSE. Available at: www.lse.ac.uk/media-and-communications/assets/documents/research/projects/disto/Measuring-Digital-Skills.pdf

van Deursen, A. J. A. M., Helsper, E. J., & Eynon, R. (2016). Development and validation of the Internet Skills Scale (ISS). *Information Communication & Society*, *19*(6), 804–23. doi:10.1080/1369118x.2015.1078834.

van Deursen, A. J. A. M., van Dijk, J. A. G. M., & Peters, O. (2012). Proposing a survey instrument for measuring operational, formal, information, and strategic internet skills. *International Journal of Human–Computer Interaction*, *28*(12), 827–37. doi:10.1080/10447318.2012.670086.

van Deursen, A. J. A. M., Helsper, E. J., Eynon, R., & van Dijk, J. A. G. M. (2017). The compoundness and sequentiality of digital inequality. *International Journal of Communication*, *11*, 452–73. Available at: https://ijoc.org/index.php/ijoc

van Dijk, J. A. G. M., & van Deursen, A. J. A. M. (2014). *Digital skills: Unlocking the information society*. New York: Palgrave Macmillan.

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. *Sage Open*, *10*(1). doi:10.1177/2158244019900176.

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*. doi:10.1016/j.chb.2017.03.010.

Vandenberg, R. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, *5*(139), 139–58.

Vekiri, I. (2010). Socioeconomic differences in elementary students' ICT beliefs and out-of-school experiences. *Computers & Education*, *54*(4), 941–50. doi:10.1016/j.compedu.2009.09.029.

Vuorikari, R., Punie, Y., Carretero Gomez, S., & Van Den Brande, G. (2016). *DigComp 2.0*. Brussels: European Commission. Available at: https://ec.europa.eu/jrc/en/digcomp

Walther, B., Hanewinkel, R., & Morgenstern, M. (2014). Effects of a brief school-based media literacy intervention on digital media use in adolescents: Cluster randomized controlled trial. *Cyberpsychology, Behavior, and Social Networking*, *17*(9), 616–23. doi:10.1089/cyber.2014.0173.

Wamuyu, P. K. (2017). Bridging the digital divide among low income urban communities. Leveraging use of community technology centers. *Telematics and Informatics*, *34*(8), 1709–20. doi:10.1016/j.tele.2017.08.004.

White, R., Walker, P., Roberts, S., Kalisky, S., & White, P. (2006). Bristol COPD Knowledge Questionnaire (BCKQ): Testing what we teach patients about COPD. *Chronic Respiratory Disease*, *3*(3), 123–31. doi:10.1191/1479972306cd117oa.

Williams-Diehm, K. L., Miller, C. R., Sinclair, T. E., & Wronowski, M. L. (2018). Technology-based employability curriculum and culturally diverse learners with disabilities. *Journal of Special Education Technology*, *33*(3), 159–70. doi:10.1177/0162643417749933.

Willson, S., & Miller, K. (2014). Data collection. In K. Miller, S. Willson, V. Chepp, & J. L. Padilla (eds) *Cognitive interviewing methodology* (pp.15–34). Hoboken, NJ: John Wiley & Sons.

Yoshida, M. (2018). An investigation of the social network system competencies of high school students in Japan. *International Journal of Emerging Technologies in Learning*, *13*(5), 4–18. doi:10.3991/ijet.v13i05.8101.

Zhong, Z. J. (2011). From access to usage: The divide of self-reported digital skills among adolescents. *Computers & Education*, *56*(3), 736–46. doi:10.1016/j.compedu.2010.10.016.

Appendices (English versions only)

A. Details of adjustments made to the skills survey instrument after partner discussion

Table 20.	TECHNICAL AND OPERATIONAL SKILLS ITEMS ADJUSTMENTS
Items	Comments on changes and justification for not making suggested changes
I know how to adjust privacy settings	- This item has been extensively tested - The phrasing "on the devices and platforms I use" is probably too specific – want to know it in general, in all possible contexts
I know how to turn off the location settings on my mobile devices	
I know how to connect to a safe and secure WIFI network	- This is a slightly more advanced technical skill by integrating "safe and secure" rather than just connecting to any network - This is not double-barrelled in practice since "safe and secure" are used together in everyday language and interpreted as such
I know how to connect devices to each other (e.g. pairing devices, screen mirroring, wireless connection to a printer)	Changed the item by focusing solely on connecting devices to each other and by revising the corresponding examples
I know how to have the same documents, contacts, and apps on all devices that I use	Kept the wording simple and avoid less common words such as "synchronise"
I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition)	"Protect" instead of "lock", because of translatability
I know how to store photos, documents or other files in the cloud	
I know how to remove apps from a mobile device	Asking only about removing because this is a more advanced skill than installing; if they are able to remove they should be able to install
I know how to delete the record of sites which I have visited before	
I know how to use private browsing	This is a newly added item after a suggestion to include "private browsing"
I know how to block unwanted pop-up messages or ads	
I know how to use a programming language (e.g. XML, Python)	

Table 21.	INFORMATION NAVIGATION AND PROCESSING SKILLS ITEMS ADJUSTMENTS
Items	Comments on changes and justification for not making suggested changes
I know how to choose the best keywords for online searches	
I know how to change how I search for things online, for example if I do not obtain the result I want	Removed because the item seems unclear. Choosing keywords is already reflected in the previous one. This will be checked through performance tests
I know how to use advanced search functions in search engines	Added to replace the item before
I know how to find a website I have visited before	Kept the phrasing because this item has been tested before. In English and Spanish it is not problematic – mark as potentially difficult in translation
I know how to navigate my way through a website with many pages	Removed because the item seems archaic and does not apply to mobile apps very well If asked it should not be related to the design of apps or websites but to whether they can no matter what the design "Navigating" could be viewed as just opening new windows and links
I know how to find information on a website no matter how it is designed	Very loosely adapted from previously tested item ("I find the design of many websites to be confusing"). Reverse coding did not work here and thus adapted to be positive testing
I understand what different icons (, , ,) on apps or websites mean	
I know how to find information on a website no matter how it is designed	Item added to replace the previous item on website design
I know how to use advanced search functions in search engines	Item added to replace the previous item on additional searches
Sometimes I end up on websites without knowing how I got there	
When I have a question, I am able to find information online that is relevant to answering it	
I know how to incorporate different online sources when writing/producing a text of my own	Item causes confusion and is captured under content creation skills Content implies referencing but doesn't say this explicitly – may not be appropriate for the youngest children Referencing might not be relevant in relation to wellbeing
I know how to check if the information I find online is true	Kept this item because this one has been extensively tested If you ask whether they cross-check, they will know that this is what they should be doing – better to check through critical skills items and performance tests
I know how to figure out if a website can be trusted	

Table 22.	COMMUNICATION AND INTERACTION SKILLS ITEMS ADJUSTMENTS
Items	Comments on changes and justification for not making suggested changes
Depending on the situation, I know which medium/tool to use to communicate with someone (make a call, send a WhatsApp message, send an email, etc.)	Rephrased item because the situation or context matters
I know how to remove people from my contact lists	Kept the word "how" instead of "when" because this is a more functional social skill Kept the word "contact lists" because this is supposed to be device- and platform-independent This item has been tested
I know when I should mute myself or disable video in online interactions	This is about a more critical skill of knowing under which circumstances (thus not how, but when)
I know how to block messages from someone I don't want to hear from	
I know which images and information of me it is OK to share online	Social desirability is circumvented by using a scale and not "yes/no" answers
I know when I should not post pictures or videos of others online	Avoided the wording "when I am allowed to" to make sure the item is about social rules and not about the legislative system These items are not concerned with uses or outcomes – as long as they know that what they are doing is "wrong" then they are skilled even if they still go ahead and do this
I know how to change who I share content with (e.g. just friends, friends of friends or make it public)	Kept the word "how" instead of "when" because this is intended to measure a more functional social skill. This does risk the item being technical – to be tested in terms of how it groups in the pilot survey
I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters	Text speak is still part of WhatsApp and instant messaging. If there is a problem in translation, this has to be removed We specified text speak by including some examples
I know how to make my comments and behaviours appropriate to the online situation	
I know how to report negative content relating to me or a group to which I belong	
I know how to recognise when someone is being bullied online	If more than one cyberbullying item is asked, this "recognise" one should be asked first
I know who to turn to when someone I know is being bullied or harassed online	

Table 23.	CONTENT CREATION AND PRODUCTION SKILLS ITEMS ADJUSTMENTS
Items	Comments on changes and justification for not making suggested changes
I know how to create something which combines different digital media (photo, music, video, GIF)	This item is not about sharing with others, but about the ability to create digital content
I know how to edit existing digital images, music and videos	Kept the word "digital" because "online" is more associated with using a device that is connected to the internet or information that is created by others – too limited
I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people	Removed the word "or" to avoid a triple-barrelled item Changed the word "audience" into "specific groups of people" to align with everyday language but kept the word "reach" instead of "communicate" because otherwise it might be related more to communication (interpersonal communication) than content distribution (one to many communication).
I know how to ensure that many people will see what I put online	The word "put" is better in this case – otherwise the item would be narrowly social media-related Question should be about the skill (knowing how to) and not the activity – whether they actually want to reach as many people as possible
I know how to use filters and other tools to make a photo or video look more attractive	Kept the phrasing "filters and other tools" because this is a more basic, functional skill than editing which we already ask about – to weed out in the pilot survey which item is better in terms of distinctiveness "More attractive" is closer to everyday language than "more appealing" – although it should not (only) be about the attractiveness of the person in the photo or video but of the video
I know how to change the things I put online depending on how other people react to it	The word "put" is better in this case – otherwise the item would be narrowly social media-related Check for overlap with social skills and posting in pilot survey
I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post)	This item is not just about adverts but also product placement; added examples but not sure they work Will be checked in performance tests
I know how to reference and use content covered by copyright	Checked in performance tests?
I know which different types of licenses apply to online content	In pilot survey this one will be checked against previous one to determine which has most distinctiveness If we give examples, then the question lacks validity

Table 24.	DIGITAL KNOWLEDGE ITEMS ADJUSTMENTS	
Skills dimension	Statement	Comment on changes justification
Information navigation and processing	Everyone gets the same information when they search for things online	Comments on changes justification for not making suggested changes
	The first search result is always the best information source	
	The lock icon means a website can be trusted	
	A post that a friend shared with you is more trustworthy than other posts	
	Online cookies protect my information and activities online from being shared with other companies or organisations	
	When information is backed up on the cloud, it is always encrypted	Data literacy item added "Always" is key here
Communication and interaction	Whether I like or share a post can have a negative impact on others	This gets at looking at whether online behaviour is seen as less harmful than offline behaviour
	The first post I see on social media is the last thing that was posted by one of my contacts	This gets at personalisation/ algorithmic bias
	If someone has posted something online without making it private, it is ok to share or forward it without asking	This is highly contested in academia – there is probably no right answer? We might want to change this to: "If someone has shared something with you without asking you to keep it a secret or without making it private it is okay to forward it to other people without asking"
	Before sharing a picture that clearly shows a friend, I should always ask them for permission first	Included the suggestion to specify the item by using the phrasing "before sharing a picture that clearly shows a friend" Changed the direction of the skill action to others sharing about "me" instead of "others" sharing about "you" since that does not refer to a skill or knowledge of the person but of others
	Negative comments you make online about people are less hurtful than saying them face-to-face	Included the suggestion to rephrase the item to "less" instead of "just as" to avoid ambivalence. This also adds a "false" item Specified comments in general to be about "negative comments" in particular
Content creation and production	I can freely use an image published with a creative commons license for commercial purposes	Included the suggestion to change wording. Instead of specifying a particular creative commons license, we made this a "false" statement
	Using hashtags increases the visibility of a post	This might be an educational item – in that on reading it they realise for the first time that this is the case
	Companies pay ordinary people to use their products in videos and content they create	This is about detecting advertising and promotion in non-commercial content
	Companies use the information you post on your social media profile to market their products and services	Kept the phrasing because we did not consider it necessary to add "to you" and it should be more general. They use information to market to others like them as well
	It is easy to distinguish content produced by bots from that produced by real people	Added to have an even number of content creation, communication and information items and get more at algorithmic literacy This might be a social as much as a content production creation skill

B. Question and answer formulation and items tested in cognitive interviews and pilot surveys (including source and adaptation notes)

The initial question and answer formulation for the initial long version of the yDSI digital skills items was:

Please indicate **how true the following statements are of you** when thinking about how you use the internet and technologies such as mobile phones.

If you have never done this, then reply thinking about how true this would be of you **if you had to do it now and by yourself**.

If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

Not at all true of me	Not very true of me	Neither true nor untrue of me	Mostly true of me	Very true of me	I do not understand what you mean by this	I do not want to answer

The initial question and answer formulation for the initial long version of the yDSI digital knowledge items was:

To what extent are the following statements about technologies such as the internet and mobile phones **true or not true**?

If you are not sure, please let us know.

Definitely not true	Mostly not true	Definitely true	I'm not sure	I do not want to answer

Tables 25–29 report on the digital skills and digital knowledge items in the initial long version of the yDSI.

Tables 25–28 with digital skills items also indicate the main source of inspiration for each item, as well as in how many studies it was repeated, and whether:

- The item from the original source is untouched, including terminology, phrasing and examples (O).
- The item from the original source has been slightly rephrased to comply with the best practice guidelines (R).
- The item from the original source has been majorly rephrased, including any or all of the following: changes in terminology; changes in structure (e.g. positive versus negative phrasing); and cutting or adding elements (R*).
- It is a new item (specified if based on the existing self-assessment checklist) (N).

Table 25.	INITIAL LONG LIST OF TECHNICAL AND OPERATIONAL SKILLS ITEMS AND THEIR SOURCE (USED IN COGNITIVE INTERVIEWS AND PILOT SURVEYS)		
Item	Main source	Annotation	Number of studies it appears in
I know how to adjust privacy settings	DiSTO (2014)	O	4
I know how to turn off the location settings on my mobile devices	DiSTO (2014)	O	3
I know how to connect to a safe and secure WIFI network	DiSTO (2014)	R*	4
I know how to connect devices to each other (e.g. pairing devices, screen mirroring, wirelessly connecting to a printer)	Lau & Yuen (2015)	R*	2
I know how to have the same documents, contacts, and apps on all devices that I use	NCGM (2014)	R*	3
I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition)	NCGM (2014)	R	2
I know how to store photos, documents or other files in the cloud		N (based on self-assessment items of DigComp and EDSF)	0
I know how to remove apps on a mobile device	DiSTO (2014)	R*	6
I know how to delete the record of sites which I have visited before	NCGM (2014)	R	2
I know how to use private browsing		N	0
I know how to block unwanted pop-up messages or ads	DiSTO (2014)	R	2
I know how to use programming language (e.g. XML, Python)	DiSTO (2014)	O	6

Notes: O = original; R = rephrased (small adjustment); R* = rephrased (large adjustment); N = new
NCGM = Net Children Go Mobile. DISTO = From Digital Skills to Tangible Outcomes

Table 26.	INITIAL LONG LIST OF TECHNICAL AND OPERATIONAL SKILLS ITEMS AND THEIR SOURCE (USED IN COGNITIVE INTERVIEWS AND PILOT SURVEYS)

Item	Main source	Annotation	Number of studies it appears in
I know how to choose the best keywords for online searches	Global Kids Online (2016)	O	5
I know how to find a website I have visited before	Global Kids Online (2016)	O	2
I understand what different icons ( ,  ,  , ) in an app or website mean	Porat et al. (2018)	R*	3
Sometimes I end up on websites without knowing how I got there	DiSTO (2014)	O	2
When I have a question, I am able to find information online that is relevant to answering it	Lau & Yuen (2015)	R*	4
I know how to find information on a website no matter how it is designed	DiSTO (2014)	R*	3
I know how to use advanced search functions in search engines	DiSTO (2014)	N (based on an item that was deemed too ambiguous to rephrase)	0
I know how to check if the information I find online is true	Global Kids Online (2016)	O	10
I know how to figure out if a website can be trusted	Global Kids Online (2016)	R	4

Notes: O = original; R = rephrased (small adjustment); R* = rephrased (large adjustment); N = new.
DISTO = From Digital Skills to Tangible Outcomes

Table 27.	INITIAL LONG LIST OF COMMUNICATION AND INTERACTION SKILLS ITEMS AND THEIR SOURCE (USED IN COGNITIVE INTERVIEWS AND PILOT SURVEYS)

Item	Main source	Annotation	Number of studies it appears in
Depending on who I want to communicate with, I know which medium/tool to use (make a call, send a WhatsApp message, send an email, etc.)	Rodríguez-de-Dios et al. (2018)	R	1
I know how to remove people from my contact lists	DiSTO (2014)	R	3
I know when to mute myself or disable video in online interactions		N	0
I know how to block messages from someone I don't want to hear from	NCGM (2014)	R	2
I know which images and information of me it is OK to share online	DiSTO (2014)	R	5
I know when I should not post pictures or videos of others online	Global Kids Online (2016)	R*	1
I know how to change who I share content with (e.g. just friends, friends of friends or make it public)	DiSTO (2014)	R	2
I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters	DiSTO (2014)	N (item was taken from another component and was heavily modified)	1
I know how to make my comments and behaviours appropriate to the online situation	DiSTO (2014)	R*	1
I know how to report negative content relating to me or a group to which I belong	Global Kids Online (2016)	O	2
I know who to turn to when someone I know is being bullied or harassed online	DKAP (2019)	R*	1
I know how to recognise when someone is being bullied online	DKAP (2019)	R*	1

Notes: O = original; R = rephrased (small adjustment); R* = rephrased (large adjustment); N = new NCGM = Net Children Go Mobile. DISTO = From Digital Skills to Tangible Outcomes. DKAP =Digital Kids Asia Specific.

Table 28.	INITIAL LONG LIST OF CONTENT CREATION AND PRODUCTION SKILLS ITEMS AND THEIR SOURCE (USED IN COGNITIVE INTERVIEWS AND PILOT SURVEYS)		
Item	**Main source**	**Annotation**	**Number of studies it appears in**
I know how to create something which combines different digital media (photo, music, video, GIF)	DKAP (2019)	R*	4
I know how to edit existing digital images, music and video	Global Kids Online (2016)	R	5
I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people		N	0
I know how to ensure that many people will see what I put online		N	0
I know how to use filters and other tools to make a photo or video look more attractive		N	0
I know how to change the things I put online depending on how other people react to it		N	0
I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post)		N	0
I know how to reference and use content covered by copyright		N (based on self-assessment items of DigComp)	0
I know which different types of licenses apply to online content	DiSTO (2014)	O	2

Notes: O = original; R = rephrased (small adjustment); R* = rephrased (large adjustment); N = new.
DISTO = From Digital Skills to Tangible Outcomes.

Table 29.	INITIAL LONG LIST OF DIGITAL KNOWLEDGE ITEMS (USED IN COGNITIVE INTERVIEWS AND PILOT SURVEYS)
Skills dimension	Item
Information navigation and processing	Everyone gets the same information when they search for things online
	The first search result is always the best information source
	The lock icon means a website can be trusted
	A post a friend shares with me is more trustworthy than other posts
	It is easy to distinguish content produced by bots from that produced by real people
	Online cookies protect my information and activities online from being shared with other companies or organisations
Communication and interaction	Whether I like or share a post can have a negative impact on others
	The first post I see on social media is the last thing that was posted by one of my contacts
	Before sharing a picture that clearly shows a friend, I should always ask them for permission first
	If someone has posted something online without making it private, it is okay to share or forward it without asking
	Negative comments you make online about people are less hurtful than saying them face-to-face
Content creation and production	Using hashtags increases the visibility of a post
	Companies pay ordinary people to use their products in videos and content they create
	I can freely use an image published with a creative commons license for commercial purposes
	Companies use the information you post on your social media profile to market their products and services
	When information is backed up on the cloud, it is always encrypted

C. Instructions for probing for cognitive interviews on skills questions

The items are divided by dimension (technical and operational, information navigation and processing, communication and interaction, content creation and production) and into two blocks (A and B) per dimension. Each interview only asks one block per child so that each child will see only half of the questions.

Q3.1 **A** Please indicate **how true the following statements are of you** when thinking about how you use the internet and technologies such as mobile phones.

If you have never done this, then reply thinking about how true this would be of you **if you had to do it now and by yourself**.

If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for the interview
I know how to adjust privacy settings (Q3.1_1)	What examples can they give of privacy settings? • Do they understand the difference between a safe and an unsafe network when they answer 4 or 5? • Does someone who can connect but does not know if a network is safe respond 3 or 4 on the scale?
I know how to connect to a safe and secure WI-FI network (Q3.1_3)	
I know how to connect devices to each other (e.g. pairing devices, screen mirroring, wirelessly connecting to a printer) (Q3.1_4)	Check whether they understand the examples
I know how to store photos, documents or other files in the cloud (Q3.1_7)	
I know how to delete the record of sites which I have visited before (Q3.1_9)	Do they understand the terminology "record of sites"? Perhaps use the word "history"?
I know how to use programming language (e.g. XML, Python) (Q3.1_12)	

Q3.1 **B** Please indicate **how true the following statements are of you** when thinking about how you use the internet and technologies such as mobile phones.

If you have never done this, then reply thinking about how true this would be of you **if you had to do it now and by yourself**.

If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for the interview
I know how to turn off the location settings on mobile devices (Q3.1_2)	Is this technical or privacy (interactional) related?
I know how to have the same documents, contacts and apps on all devices that I use (Q3.1_5)	
I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition) (Q3.1_6)	Does someone who knows how to do some of the examples respond 3 or 4 on the scale and not 5?
I know how to remove apps from a mobile device (Q3.1_8)	What do they understand by removing apps – just making sure they don't run in the background or removing them completely?
I know how to use private browsing (Q3.1_10)	Can they describe what "private browsing" is when they answer 4 or 5?
I know how to block unwanted pop-up messages or ads (Q3.1_11)	

Q4.1 **A** Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones.

If you have never done this, then reply thinking about how true this would be of you **if you had to do it now and by yourself**.

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for interviews
I know how to find a website I have visited before (Q4.1_2)	Do they relate this to browsing history, keyword searches or to remembering the website's address? Do they make a distinction between websites they go to often and a website they went to once and might want to go back to again?
I understand what different icons (e.g. , , ,) on apps or websites mean (Q4.1_3)	Do they experience difficulties in relating this to icons in general? Perhaps specify by using "icons that are commonly found on apps and websites"? Are these the right icons to include? Are others mentioned as something they don't know about?
When I have a question, I am able to find information online that is relevant to answering it (Q4.1_5)	Do they answer this in relation to relevance of the information or just any answer to their question?
I know how to use advanced search functions in search engines (Q4.1_7)	What do they consider advanced search functions? (e.g. Google Scholar, Boolean search?)
I know how to figure out if a website can be trusted (Q4.1_9)	Check how they figure out if a website can be trusted by asking for examples Check whether they understand the difference between "trusted information" and "a trusted website" when they answer the same to both items

Q4.1 **B** Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones.

If you have never done this, then reply thinking about how true this would be of you **if you had to do it now and by yourself**.

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for interviews
I know how to choose the best keywords for online searches (Q4.1_1)	
I understand what different icons (e.g. , , ,) on apps or websites mean (Q4.1_3)	Do they experience difficulties in relating this to icons in general? Perhaps specify by using "icons that are commonly found on apps and websites"? Are these the right icons to include? Are others mentioned as something they don't know about?
Sometimes I end up on websites without knowing how I got there (Q4.1_4)	
I know how to find information on a website no matter how it is designed (Q4.1_6)	What do they understand by "how to find" – is this about navigating or about "ugliness" or lack of user friendliness of the website?
I know how to check if the information I find online is true (Q4.1_8)	Do they mention things like cross-checking across websites when they answer 4 or 5?

Q5.1 **A** Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones.

If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for the interviews
Depending on the situation, I know which medium/tool to use to communicate with someone (make a call, send a WhatsApp message, send an email, etc.) (Q5.1_1)	• Do they interpret the question as referring to the person, the task or interaction at hand? • Do they prefer the language of tool or medium (i.e., how do they interpret either)?
I know how to remove people from my contact lists (Q5.1_2)	
I know when I should not post pictures or videos of others online (Q5.1_6)	• Are they able to distinguish sharing content of "me" and "others"? Do they answer the question differently from the previous one? • Do they link this to asking permission and/or to their understanding of what is appropriate and not appropriate to post (about themselves) online?
I know how to make my comments and behaviours appropriate to the online situation (Q5.1_9)	• Do they give examples and which do they give in relation to situations? • What do they mean by "appropriate"?
I know how to recognise when someone is being bullied online (Q5.1_11)	
I know how to report negative content relating to me or a group to which I belong (Q5.1_10)	• How do they understand reporting? • To whom are they reporting, and which groups do they have in mind?

Q5.1 **B** Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones.

If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for the interviews
I know when I should mute myself or disable video in online interactions (Q5.1_3)	• Do they understand the words "mute" and "interactions"? • Do they separate this from the functional skill of knowing how to do this?
I know how to block messages from someone I don't want to hear from (Q5.1_4)	Is it necessary to include "I don't want to hear from"? We might use the item without it but then it becomes a technical skill. This is not about deleting a person entirely
I know which images and information of me it is OK to share online (Q5.1_5)	Check whether it might be considered "silly" to answer "Not at all true of me" to this question
I know how to change who I share content with (e.g. just friends, friends of friends or make it public) (Q5.1_7)	Do they interpret this as how or when? What is better to ask in terms of a more advanced skill? Where are they more likely to not answer 5 on the scale?
I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters (Q5.1_8)	• Are the examples of text speak appropriate for young people? Which ones do they use? • Do those who know when for all of them answer 5 and others who know when for some (i.e., smileys but not capital letters) answer not 5 ("Very true of me"). What do those who don't understand one of them do with this question? Do they answer "Don't understand what you mean by this" or less than 5?
I know who to turn to when someone I know is being bullied or harassed online (Q5.1_12)	Is the question too specific by focusing on "who to turn to" and not on "what to do" in general? Other actions such as flagging, reporting or writing a supportive comment are also possible

Q6.1 **A** Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones.

If you have never done this, then reply thinking about how true this would be of you **if you had to do it now and by yourself**.

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for interview
I know how to create something which combines different digital media (photo, music, video, GIF) (Q6.1_1)	• Do they relate this to "creating digital content" instead of "sharing digital content"? • Is "create something" too vague? • Is it clear that this is integrating and combining different media types (i.e., multi-media)?
I know how to ensure that many people will see what I put online (Q6.1_4)	Do they actually want their stuff to be seen by as many people as possible? Is this something that complicates the answer – e.g. if they definitely know how to do it but are not interested – they should answer 5
I know how to use filters and other tools to make a photo or video look more attractive (Q6.1_5)	Which examples are given of "other tools" they use?
I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post) (Q6.1_7)	• Which examples do they give of "sponsored content"? • Do they relate the content to product placement or solely to advertising?
I know which different types of licenses apply to online content (Q6.1_9)	Is "licenses" a term that the youngest people understand? That is, if they know what licenses are but do not understand this term, do they answer 4 or 5 or "I don't understand what you mean by this"?

Q6.1 **B** Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones.

If you have never done this, then reply thinking about how true this would be of you **if you had to do it now and by yourself**.

Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)

	Notes for interview
I know how to edit existing digital images, music and video (Q6.1_2)	• Do those who know how to edit only one or two (e.g. only images or videos) of these answer something other than 5? • Do those who know how to do this in an advanced way (e.g. using advanced editing software) answer 5 on this? • Are they confused by "existing" digital images as opposed to their own images?
I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people (Q6.1_3)	• Is "reach" is a term that young people understand? • Which groups of people do they have in mind? Does this correspond to audiences in media and communications terms?
I know how to change the things I put online depending on how other people react to it (Q6.1_6)	• Do they give examples of situations in which they change their post because of people's reactions? • Is "the things I put online" too broad?
I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post) (Q6.1_7)	• Which examples do they give of sponsored content? • Do they relate the content to product placement or solely to advertising?
I know how to reference and use content covered by copyright (Q6.1_8)	

Q7.1 and Q7.2 **A** To what extent are the following statements about technologies such as the internet and mobile phones **true or not true**?

If you are not sure, please let us know.

Definitely true (1)	Definitely not true (2)	Not sure (88)	I don't understand what you mean by this (66)

	Notes for interviews
Everyone gets the same information when they search for things online (Q7.1_1)	• Do they understand "the same information"? Do they understand that this refers to a personalised search? • Does the item need to be specified to a search engine?
The lock icon means a website can be trusted (Q7.1_3)	Do they mention other ways in which you can determine how a website's trustworthiness can be checked?
The first post I see on social media is the last thing that was posted by one of my contacts (Q7.1_6)	Do they say this depends?
Negative comments you make online about people are less hurtful than saying them face-to-face (Q7.1_8)	Do they relate the question to situations in general or to their own experience? (We want it to be the former)
Before sharing a picture that clearly shows a friend, I should always ask them for permission first (Q7.2_1)	Does the answer depend on who the friend is and their judgement in the past?
I can freely use an image published with a creative commons license for commercial purposes (Q7.2_3)	• All should be asked what a creative commons license is • Are they familiar with the creative commons terminology or would they use a different word to describe this kind of license? For example, do they use Google image search terminology? • What do they understand by "commercial purposes"?
Companies pay ordinary people to use their products in videos and content they create (Q7.2_5)	Who do they consider "ordinary people"? Are (non-celebrity) influencers mentioned?
When information is backed up on the cloud, it is always encrypted (Q7.2_7)	Do those who do not understand what the cloud or encryption is answer "I do not understand what you mean by this"?

Q7.1 and Q7.2 **B** To what extent are the following statements about technologies such as the internet and mobile phones **true or not true**?

If you are not sure, please let us know.

Definitely true (1)	Definitely not true (2)	Not sure (88)	I don't understand what you mean by this (66)

	Notes for interviews
The first search result is always the best information source (Q7.1_2)	• Do they mention the difference between advertised links and general search results? • Do they mention that you should cross-check across several websites?
A post a friend shares with me is more trustworthy than other posts (Q7.1_4)	How do they interpret "other posts"?
Online cookies protect my information and activities online from being shared with other companies or organisations (Q7.1_5)	
Whether I like or share a post can have a negative impact on others (Q7.1_7)	If they say this is not true, what is their reasoning?
If someone has posted something online without making it private, it is okay to share or forward it without asking (Q7.2_2)	• Do they understand "private" in this context? • Do they understand this as sharing between friends or with a broader audience?
Using hashtags increases the visibility of a post (Q7.2_4)	Does "hashtags" need to be specified? Perhaps it is better to use "existing hashtags"?
Companies use the information you post on your social media profile to market their products and services (Q7.2_6)	
It is easy to distinguish content produced by bots from that produced by real people (Q7.2_8)	If they do not know what a bot is, they should answer "I don't know what you mean by this"

D. Descriptives: Digital skills items for full sample based on pilot survey

Table 30. DESCRIPTIVE STATISTICS FOR TECHNICAL AND OPERATIONAL SKILLS ITEMS ACROSS ALL COUNTRIES

Item	Total N	DWTA	Mean	SDV	Skewness	Kurtosis
I know how to adjust privacy settings	2,246	17	4.15	1.20	–1.72	2.63
I know how to turn off the location settings on mobile devices	2,244	21	4.40	1.13	–2.35	5.34
I know how to connect to a safe and secure WI-FI network	2,235	20	4.37	1.08	–2.25	5.34
I know how to connect devices to each other (e.g. pairing devices, screen mirroring, wirelessly connecting to a printer)	2,240	22	4.33	1.04	–1.97	4.07
I know how to have the same documents, contacts and apps on all devices that I use	2,245	27	4.17	1.15	–1.73	2.87
I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition)	2,247	19	4.54	0.99	–2.87	8.66
I know how to store photos, documents or other files in the cloud	2,242	22	4.21	1.20	–1.84	3.06
I know how to remove apps from a mobile device	2,242	22	4.62	1.03	–3.29	10.64
I know how to delete the record of sites which I have visited before	2,260	29	4.54	1.04	–2.83	8.06
I know how to use private browsing	2,246	28	4.37	1.19	–2.19	4.28
I know how to block unwanted pop-up messages or ads	2,246	31	3.99	1.25	–1.25	0.86
I know how to use programming language (e.g. XML, Python)	2,243	30	2.42	1.51	0.41	-1.08

Notes: N = Total number of responses; DWTA = Don’t want to answer; SDV = standard deviation.

Table 31.	DESCRIPTIVE STATISTICS FOR INFORMATION NAVIGATION AND PROCESSING ITEMS ACROSS ALL COUNTRIES					
Item	Total N	DWTA	Mean	SDV	Skewness	Kurtosis
I know how to choose the best keywords for online searches	2,380	20	4.12	1.02	1.54	2.78
I know how to find a website I have visited before	2,383	21	4.38	0.98	2.20	5.76
I understand what different icons (e.g , , ,) on apps or websites mean	2,380	22	4.30	1.02	1.91	3.97
Sometimes I end up on websites without knowing how I got there	2,382	35	2.58	1.38	0.35	–0.98
When I have a question, I am able to find information online that is relevant to answering it	2,380	19	4.16	1.18	1.78	2.88
I know how to find information on a website no matter how it is designed	2,436	31	4.05	1.12	1.65	3.09
I know how to use advanced search functions in search engines	2,432	34	3.98	0.12	1.37	1.69
I know how to check if the information I find online is true	2,384	27	3.89	1.16	1.30	1.73
I know how to figure out if a website can be trusted	2,380	25	3.96	1.11	1.31	1.83

Notes: N = Total number of responses; DWTA = Don't want to answer; SDV = standard deviation.

Table 32.	DESCRIPTIVE STATISTICS FOR COMMUNICATION AND INTERACTION ITEMS ACROSS ALL COUNTRIES					
Item	Total N	DWTA	Mean	SDV	Skewness	Kurtosis
Depending on the situation, I know which medium/tool to use to communicate with someone (make a call, send a WhatsApp message, send an email, etc.)	2,226	21	4.47	0.98	–2.43	6.40
I know how to remove people from my contact lists	2,250	23	4.61	0.92	–3.04	1.00
I know when I should mute myself or disable video in online interactions	2,247	25	4.44	1.03	–2.39	6.04
I know how to block messages from someone I don't want to hear from	2,240	24	4.51	1.04	–2.76	7.94
I know which images and information of me it is OK to share online	2,237	27	0.45	1.00	–2.59	7.44
I know when I should not post pictures or videos of others online	2,246	25	4.49	1.04	–2.72	7.76
I know how to change who I share content with (e.g. just friends, friends of friends or make it public)	2,227	28	4.48	0.10	–2.60	7.27
I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters	2,250	23	4.44	1.00	–2.41	6.45
I know how to make my comments and behaviours appropriate to the online situation	2,259	35	4.40	1.03	–2.43	6.63
I know how to report negative content relating to me or a group to which I belong	2,238	27	4.31	1.11	–2.10	4.60
I know how to recognise when someone is being bullied online	2,236	32	4.06	1.15	–1.56	2.52
I know who to turn to when someone I know is being bullied or harassed online	2,258	33	3.82	1.25	–1.03	0.51

Notes: N = Total number of responses; DWTA = Don't want to answer; SDV = standard deviation.

Table 33.	DESCRIPTIVE STATISTICS FOR CONTENT CREATION AND PRODUCTION ITEMS ACROSS ALL COUNTRIES					
Item	Total N	DWTA	Mean	SDV	Skewness	Kurtosis
I know how to create something that combines different digital media (photo, music, video, GIF)	2,300	25	3.78	1.24	–0.96	0.26
I know how to edit existing digital images, music and video	2,301	26	3.89	1.22	–1.15	0.80
I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people	2,298	27	3.79	1.24	–1.11	0.85
I know how to ensure that many people will see what I put online	2,312	34	3.64	1.29	–0.84	0.07
I know how to use filters and other tools to make a photo or video look more attractive	2,308	34	4.10	1.18	–1.60	2.33
I know how to change the things I put online depending on how other people react to it	2,304	34	3.76	1.30	–1.16	0.90
I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post)	2,297	36	4.09	1.16	–1.62	2.60
I know how to reference and use content covered by copyright	2,328	43	3.67	1.31	–0.94	0.24
I know which different types of licenses apply to online content.	2,324	36	3.33	1.40	–0.57	0.54

Notes: N = Total number of responses; DWTA = Don’t want to answer; SDV = standard deviation.

E. Factor analyses: Digital skills items based on pilot survey

Table 34. TECHNICAL AND OPERATIONAL SKILL ITEMS FACTOR LOADINGS IN EACH COUNTRY

Item	All	UK	ES	FI	GE	IT	NL	PL	PT
I know how to adjust privacy settings	0.61	0.74	0.62	0.63	0.53	0.42	0.65	0.74	0.60
I know how to turn off the location settings on mobile devices	0.66	0.54	0.69	0.61	0.66	0.60	0.72	0.81	0.59
I know how to connect to a safe and secure WI-FI network	0.70	0.73	0.67	0.68	0.70	0.58	0.67	0.79	0.70
I know how to connect devices to each other (e.g. pairing devices, screen mirroring, wirelessly connecting to a printer)	0.71	0.70	0.64	0.66	0.71	0.63	0.71	0.75	0.77
I know how to have the same documents, contacts, and apps on all devices that I use	0.67	0.75	0.54	0.68	0.66	0.69	0.72	0.72	0.68
I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition)	0.73	0.47	0.70	0.63	0.76	0.71	0.79	0.79	0.69
I know how to store photos, documents or other files in the cloud	0.71	0.53	0.73	0.70	0.78	0.58	0.70	0.79	0.68
I know how to remove apps from a mobile device	0.68	0.26	0.60	0.65	0.75	0.58	0.72	0.72	0.75
I know how to delete the record of sites which I have visited before	0.72	0.72	0.67	0.67	0.73	0.60	0.77	0.79	0.68
I know how to use private browsing	0.66	0.68	0.55	0.59	0.67	0.58	0.66	0.77	0.62
I know how to block unwanted pop-up messages or ads	0.60	0.49	0.60	0.63	0.61	0.64	0.52	0.67	0.69
I know how to use programming language (e.g. XML, Python)	0.24	0.14	0.17	0.24	0.19	0.32	0.18	0.29	0.32

Table 35.	INFORMATION NAVIGATION AND PROCESSING SKILLS ITEMS FACTOR LOADINGS IN EACH COUNTRY								
Item	All	UK	ES	FI	GE	IT	NL	PL	PT
I know how to choose the best keywords for online searches	0.71	0.65	0.72	0.65	0.70	0.73	0.67	0.75	0.70
I know how to find a website I have visited before	0.66	0.51	0.56	0.64	0.66	0.68	0.76	0.69	0.69
I understand what different icons (e.g , , ,) on apps or websites mean	0.66	0.59	0.66	0.59	0.72	0.54	0.69	0.74	0.61
Sometimes I end up on websites without knowing how I got there	–0.01	–0.10	–0.14	–0.20	–0.02	–0.05	0.01	0.10	0.09
When I have a question, I am able to find information online that is relevant to answering it	0.57	0.67	0.73	0.59	0.09	0.67	0.72	0.73	0.62
I know how to find information on a website no matter how it is designed	0.71	0.73	0.65	0.65	0.71	0.63	0.70	0.79	0.74
I know how to use advanced search functions in search engines	0.67	0.66	0.62	0.72	0.63	0.66	0.62	0.70	0.74
I know how to check if the information I find online is true	0.73	0.71	0.79	0.72	0.64	0.70	0.72	0.77	0.77
I know how to figure out if a website can be trusted	0.72	0.70	0.69	0.68	0.68	0.71	0.69	0.75	0.74

Table 36.	COMMUNICATION AND INTERACTION SKILLS ITEMS FACTOR LOADINGS IN EACH COUNTRY								
Item	All	UK	ES	FI	GE	IT	NL	PL	PT
Depending on the situation, I know which medium/tool to use to communicate with someone (make a call, send a WhatsApp message, send an email, etc.)	0.70	0.66	0.72	0.58	0.70	0.70	0.68	0.74	0.77
I know how to remove people from my contact lists	0.72	0.66	0.82	0.51	0.72	0.72	0.68	0.73	0.83
I know when I should mute myself or disable video in online interactions	0.69	0.67	0.74	0.54	0.63	0.65	0.73	0.75	0.82
I know how to block messages from someone I don't want to hear from	0.68	0.60	0.67	0.48	0.61	0.75	0.71	0.77	0.77
I know which images and information of me it is OK to share online	0.71	0.72	0.75	0.66	0.65	0.63	0.71	0.71	0.81
I know when I should not post pictures or videos of others online	0.70	0.59	0.72	0.68	0.67	0.71	0.61	0.76	0.83
I know how to change who I share content with (e.g. just friends, friends of friends or make it public)	0.73	0.61	0.68	0.66	0.73	0.70	0.66	0.78	0.86
I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters	0.71	0.71	0.69	0.73	0.75	0.51	0.73	0.76	0.65
I know how to make my comments and behaviours appropriate to the online situation	0.71	0.61	0.79	0.52	0.73	0.55	0.74	0.76	0.74
I know how to report negative content relating to me or a group to which I belong	0.66	0.72	0.62	0.57	0.69	0.62	0.58	0.77	0.58
I know how to recognise when someone is being bullied online	0.57	0.59	0.43	0.59	0.63	0.53	0.47	0.59	0.59
I know who to turn to when someone I know is being bullied or harassed online	0.46	0.66	0.41	0.48	0.43	0.29	0.48	0.52	0.46

Table 37.	CONTENT CREATION AND PRODUCTION SKILLS ITEMS FACTOR LOADINGS IN EACH COUNTRY								
Item	All	UK	ES	FI	GE	IT	NL	PL	PT
I know how to create something which combines different digital media (photo, music, video, gif)	0.69	0.71	0.79	0.68	0.63	0.69	0.66	0.71	0.77
I know how to edit existing digital images, music and video	0.72	0.72	0.81	0.72	0.66	0.70	0.68	0.74	0.78
I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people	0.74	0.79	0.62	0.68	0.74	0.69	0.70	0.83	0.81
I know how to ensure that many people will see what I put online	0.69	0.67	0.69	0.71	0.63	0.68	0.68	0.73	0.71
I know how to use filters and other tools to make a photo or video look more attractive	0.65	0.48	0.72	0.66	0.55	0.66	0.66	0.69	0.75
I know how to change the things I put online depending on how other people react to it	0.68	0.66	0.71	0.75	0.62	0.69	0.66	0.72	0.69
I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post)	0.61	0.62	0.52	0.52	0.72	0.52	0.69	0.61	0.62
I know how to reference and use content covered by copyright	0.68	0.68	0.70	0.63	0.69	0.70	0.66	0.73	0.68
I know which different types of licenses apply to online content	0.64	0.73	0.72	0.74	0.53	0.62	0.58	0.68	0.63

F. Reasoning for deletion or modification of items after piloting and cognitive interviews

This appendix includes a detailed justification of why certain items were adjusted after the cognitive interviews and pilot surveys.

Digital skills assessment items (deletions and modifications)

Technical and operational skills dimension

I know how to connect devices to each other (e.g. pairing devices, screen mirroring, wirelessly connecting to a printer)

- ⇨ This item caused confusion among respondents in quite a few countries, so was initially rephrased as "I know how to connect devices to each other using Bluetooth or wireless connections" to get rid of confusing terms.
- ⇨ This item is conceptually similar to the next one ("I know how to have the same documents, contacts and apps on all devices that I use"), so we elected to keep the one that did not require any changes.

I know how to remove apps from a mobile device

- ⇨ This item was one of the few that didn't load well in every country. Specifically, it did not load well in the UK, meaning that it was not a translation issue. Because of this, we elected to remove this item.

I know how to delete the record of sites which I have visited before

- ⇨ Participants in the cognitive interviews suggested using the term "history", but aside from this, the item wasn't particularly problematic.
- ⇨ However, since it is conceptually similar to the next one ("I know how to use private browsing"), we may consider removing this item in the final version of the survey.
- ⇨ This item was deleted for the above-mentioned reasons after too many items remained in the technical and operational dimension following factor analyses.

I know how to recognise whether a WIFI network is safe and secure

- ⇨ This item was rephrased after the cognitive interviews, as participants focused on their ability to connect to a Wifi network rather than on whether they knew if that network was safe and secure.
- ⇨ This item was deleted, as it caused confusion, and it was not deemed to capture a skill that would add value to the discussion of digital skills in the broader context of the ySKILLS project.

I know how to have the same documents, contacts and apps on all devices that I use

- ⇨ This item was lightly rephrased after the cognitive interviews to include contacts in the list of things that children should know how to share across devices.
- ⇨ It was tentatively kept in before looking at the factor analysis, but as there were enough other items remaining, it was decided that all items that included changes would be removed.
- ⇨ In addition, another item was conceptually similar to this one.

Information navigation and processing skills dimension

Sometimes I end up on websites without knowing how I got there

- ⇨ This item was deleted as it didn't load in any of the countries.

When I have a question, I am able to find information online that is relevant to answering it

- ⇨ This item doesn't load in Germany, which may indicate a translation issue.
- ⇨ Since other navigation items loaded nicely across all countries, this item was deleted.

I know how to figure out if a website can be trusted

- ⇨ This item is conceptually similar to the previous one ("I know how to check if the information I find online is true").
- ⇨ There were no specific statistical issues or points of concern in the cognitive interviews. This item was deleted because the previous one was deemed more directly pertaining to information navigation and to the performance tests.
- ⇨ This item was later reintroduced after it became clear that it had better statistical properties on the short scale (see Section 8.1).

I understand what different icons (e.g. , , ,) on apps or websites mean

- ⇨ This item was kept for validation in performance tests and had good properties in the cognitive interviews and in the CFA including only the information navigation and processing dimension. It was later removed from the short yDSI because of low loadings and cross-loadings on the communication and interaction dimension (see Section 8.1).

Communication and interaction skills dimension

I know how to remove people from my contact lists

- ⇨ This item was the most skewed out of all of them, so it was deemed too easy, and was deleted.

I know how to block messages from someone I don't want to hear from

- ⇨ This item was also strongly skewed, but less so than the previous one. It remained within the considered pool for factor analysis.
- ⇨ After seeing that there was enough room to delete an additional item from the communication and interaction dimension, this item was deleted.

I know when I should not post pictures or videos of others online

- ⇨ This item is conceptually similar to the previous one ("I know which images and information of me it is okay to share online"), but it was slightly more skewed, so we elected to delete this one.

I know how to change who I share content with (e.g. just friends, friends of friends or make it public)

- ⇨ This item is very skewed, so we deleted it and keep the next one ("I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters").

I know how to make my comments and behaviours appropriate to the online situation

- ⇨ This item is very skewed.
- ⇨ Cognitive interviews revealed that there is too much room for interpretation, so we elected to delete this item.

I know who to turn to when someone I know is being bullied or harassed online

- ⇨ This item is conceptually similar to the previous one ("I know how to recognise when someone is being bullied online"), so we deleted this one because the phrasing "who to turn to" is less directly indicative of skill.

Content creation and production skills dimension

I know how to use different types of content (e.g. images, videos, music, text) to reach specific groups of people

- ⇨ Conceptually similar to other items in the creation dimension, but less clear (second part of the item caused some confusion in the cognitive interviews).
- ⇨ Delete this because of clarity.

I know how to use filters and other tools to make a photo or video look more attractive

- ⇨ Conceptually similar to other items in creation dimension, but too specific, so we elected to delete this one.

Digital knowledge items (outright deletions)

The lock icon means a website can be trusted
Deleted because answering patterns between participants who answered everything else wrong or right was confusing.
Online cookies protect my information and activities online from being shared with other companies or organisations
Deleted because of inconsistent answering patterns.
It is easy to distinguish content produced by bots from that produced by real people
Initially deleted because of ambiguity in what is the correct answer and inconsistent answering patterns, but it was later restored. Pending checks with performance tests to decide whether to keep or delete it in the full survey.

G. Performance tasks

Welcome

Great that you are participating in this research!

After two short questions, you will be presented with a few statements and then a series of tasks. Please read the instructions carefully every time you see a new task.

Q1.1 How would you describe yourself? I am...
Male
Female
Non-binary/Gender-queer
Other:

Q1.2 How old are you?

Q2.1 Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones or computers. Reply thinking about **how true this would be of you** if you had to do it now, on your own.

If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

Sometimes there are various examples given; only select "Very true of me" if all of the examples apply to what you do or know.

	Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)
I know how to adjust privacy settings							
I know how to turn off the location settings on mobile devices							
I know how to recognise whether a WIFI network is safe and secure							
I know how to have the same documents, contacts, and apps on all devices that I use							
I know how to delete the record of sites which I have visited before							
I know how to protect a device (e.g. with a PIN, a screen pattern, a finger print, facial recognition)							
I know how to store photos, documents or other files in the cloud							
I know how to use private browsing							
I know how to block unwanted pop-up messages or ads							
I know how to use programming language (e.g. XML, Python)							

Q2.2 Please indicate how true the following statements are of you when thinking about how you use the internet and technologies such as mobile phones or computers. Reply thinking about **how true this would be of you** if you had to do it now, on your own.

If you do not understand what the question is asking, tick the box "I do not understand what you mean by this".

	Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)
I know how to choose the best keywords for online searches	○	○	○	○	○	○	○
I know how to find a website I have visited before	○	○	○	○	○	○	○
I understand what different icons (e.g. , , ,) on apps or websites mean	○	○	○	○	○	○	○
I know how to find information on a website no matter how it is designed	○	○	○	○	○	○	○
I know how to use advanced search functions in search engines	○	○	○	○	○	○	○
I know how to check if the information I find online is true	○	○	○	○	○	○	○

Q2.3 Please indicate how true the following statements are of you when thinking about how you use the Internet and technologies such as mobile phones or computers. Reply thinking about **how true this would be of you** if you had to do it now, on your own.

If you do not understand what the question is asking, tick the box 'I do not understand what you mean by that'.

Sometimes there are various examples given, only select very true of me if all of the examples apply to what you do or know.

	Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)
Depending on the situation, I know which medium/tool to use to communicate with someone (make a call, send a WhatsApp message, send an email, etc.)	○	○	○	○	○	○	○
I know when I should mute myself or disable video in online interactions	○	○	○	○	○	○	○
I know how to block messages from someone I don't want to hear from	○	○	○	○	○	○	○
I know which images and information of me it is OK to share online	○	○	○	○	○	○	○
I know when it is appropriate and when it is not appropriate to use emoticons (e.g. smileys, emojis), text speak (e.g. LOL, OMG) and capital letters	○	○	○	○	○	○	○
I know how to report negative content relating to me or a group to which I belong	○	○	○	○	○	○	○
I know how to recognise when someone is being bullied online	○	○	○	○	○	○	○

Q2.4 Please indicate how true the following statements are of you when thinking about how you use the Internet and technologies such as mobile phones or computers. Reply thinking about **how true this would be of you** if you had to do it now, on your own.

If you do not understand what the question is asking, tick the box 'I do not understand what you mean by that'.

	Not at all true of me (1)	Not very true of me (2)	Neither true nor untrue of me (3)	Mostly true of me (4)	Very true of me (5)	I do not understand what you mean by this (66)	Don't want to answer (99)
I know how to create something which combines different digital media (photo, music, video, gif)	○	○	○	○	○	○	○
I know how to edit existing digital images, music and video	○	○	○	○	○	○	○
I know how to ensure that many people will see what I put online	○	○	○	○	○	○	○
I know how to change the things I put online depending on how other people react to it	○	○	○	○	○	○	○
I know how to distinguish sponsored and non-sponsored content online (e.g. in a video or social media post)	○	○	○	○	○	○	○
I know how to reference and use content covered by copyright	○	○	○	○	○	○	○
I know which different types of licenses apply to online content.	○	○	○	○	○	○	○

Q2.5 To what extent are the following statements about technologies such as the internet and mobile phones true or not true?

If you are not sure, please let us know.

	Definitely true (1)	Definitely not true (2)	I am not sure (88)	I don't understand what you mean by this (99)
The first search result is always the best information source	○	○	○	○
Everyone gets the same information when they search for things online	○	○	○	○
It is easy to distinguish content produced by bots from that produced by real people	○	○	○	○
Before sharing a picture that clearly shows a friend, I should always ask them for permission first	○	○	○	○
The first post I see on social media is the last thing that was posted by one of my contacts	○	○	○	○
Whether I like or share a post can have a negative impact on others	○	○	○	○
Using hashtags increases the visibility of a post	○	○	○	○
Companies pay ordinary people to use their products in videos and content they create	○	○	○	○
When information is backed up on the cloud, it is always encrypted	○	○	○	○

Q3.1

What follows now are a series of tasks that you are asked to complete.

Try to find the answers or solve the problem. If you can’t figure it out, don’t try too long and move to the next task.

In some tasks you are asked to do things by opening up a new window. After you have looked for the answer or tried to find the solution for the task, you should come back to this survey page, give your answer and move on to the next task.

Please use only the computer you are on right now to find the answers and solutions and don’t use your mobile or another device.

Q3.2

Netflix is a very popular streaming service that allows members to watch a wide variety of TV shows, movies, documentaries, and more.

Please **open a new window and use a search engine** such as Google or Bing to find out who founded the streaming platform.

The founders of Netflix are ...

Q3.3

In 2018, Netflix released their first interactive film.

Please open a new window and use a search engine such as Google or Bing to find out what the name of this movie was.

This time you only want to search for *news items* published in *2018*.

The name of the movie is ...

Q3.4

A popular movie on Netflix is *Jurassic Park*. In reality dinosaurs used to live in the Mesozoïc era. This era contains three periods.

Please open a new window and search the internet to find out what the names of these three periods were.

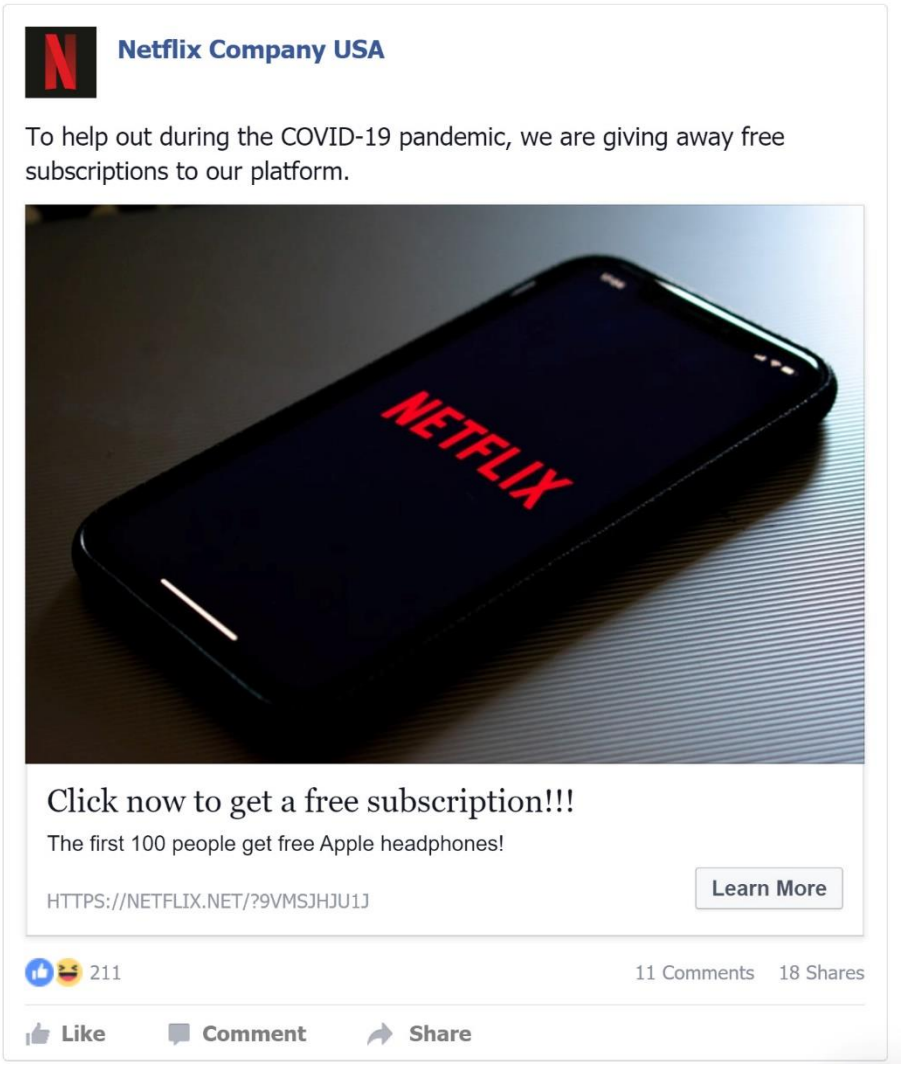
The names of these three periods are …

Q4.1 In what follows, we present you with four messages.

Please read them carefully and explain for each what you think it is trying to do and what else you note about the post.

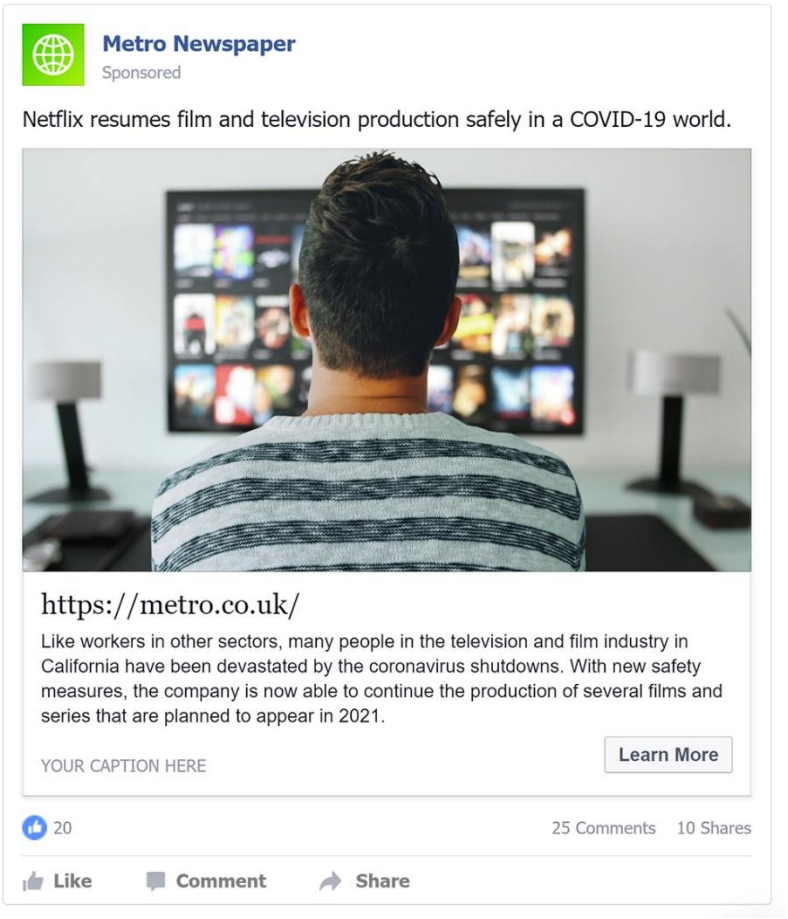
What do you think this message is trying to accomplish?

Q4.2 Message 2



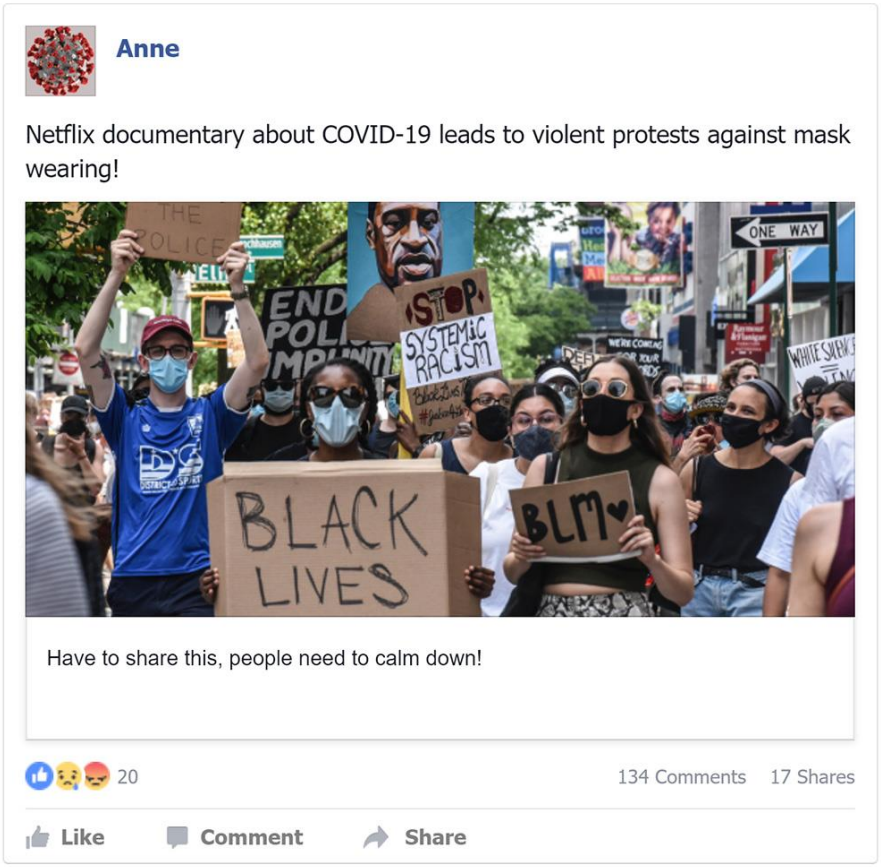
What do you think the people who posted the above message are trying to do?

Q4.3 Message 3



What do you think the purpose of this message is?

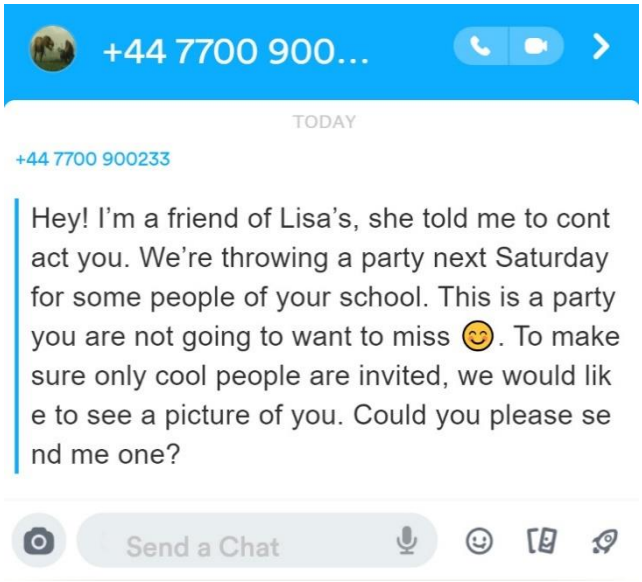
Q4.4 Message 4



What is Anne doing by posting this message? What do you think about the message?

Q5.1 In the next questions you will see a few messages.
Read them and then let us know what you think about them and how you would react to them.

Someone sends you the following message:

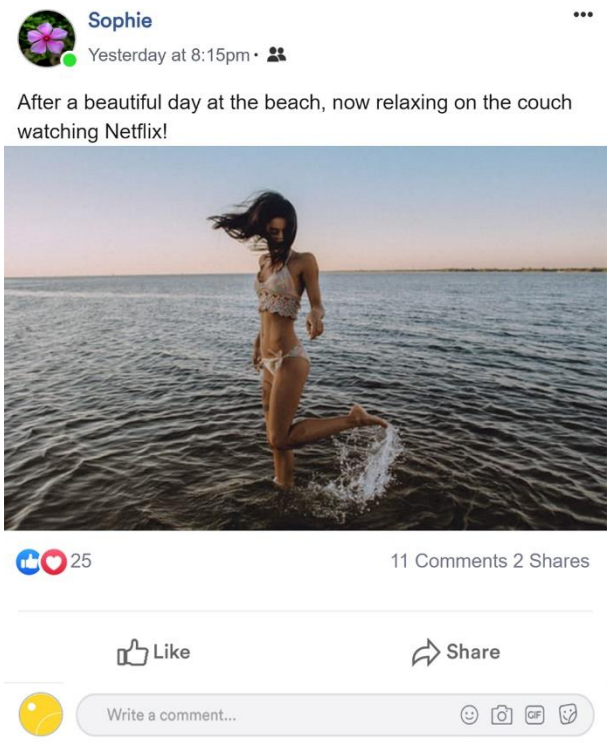


What would you do when you receive this message? Explain below why:

Q5.2 The following posts appear on your page:

Post 1

Post 2



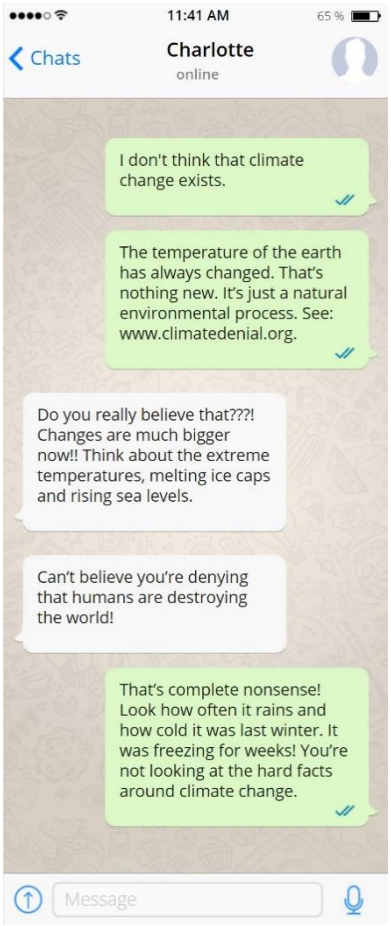
Q5.3 Could you explain for each post why it is appropriate or inappropriate to post them on social media like that?

Post 1 by Lucas:

Post 2 by Sophie:

Q6.1 Below are parts of two chats between classmates about climate change. Please read them both.

Chat 1



Chat 2

Do you think anything in these two chats is problematic?

If so, please describe what seems problematic to you below.
If not, please describe below why nothing is problematic for you below.

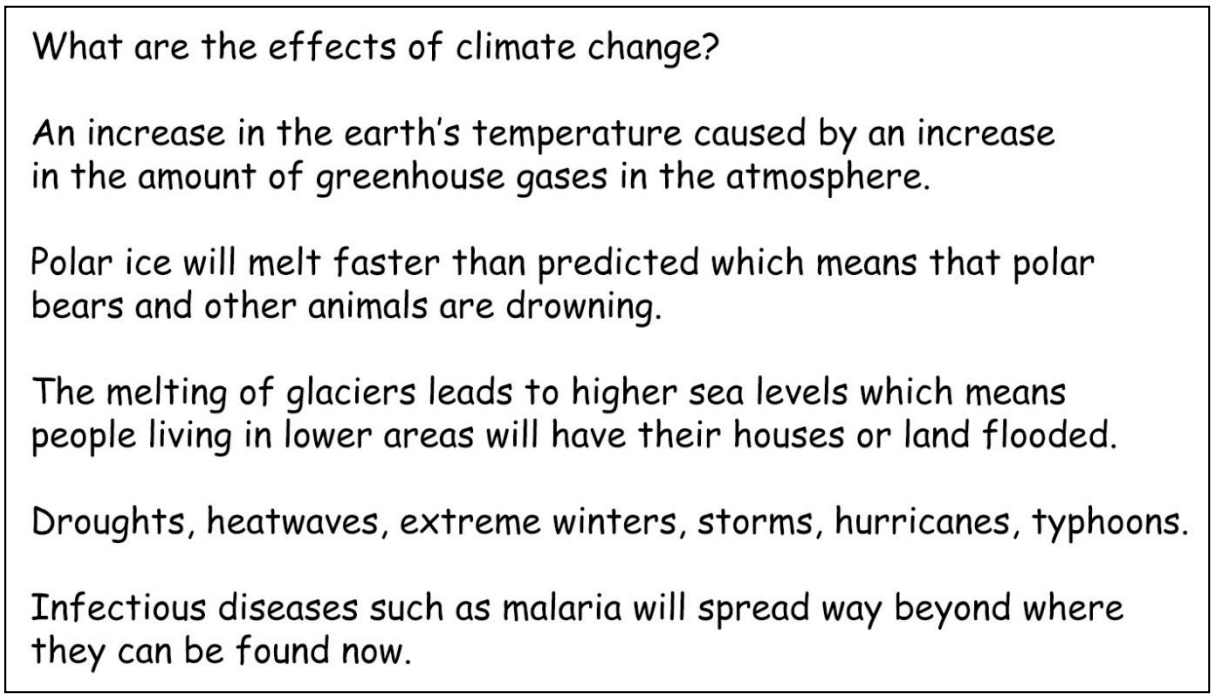
Q7.1 Below is an animated GIF image:

You would like to this image **to go viral** on social media (e.g. Facebook, Instagram, Snapchat, Pinterest, TikTok, etc.).

Please describe how you would do that:

Q7.2
You are preparing a group presentation on the effects of climate change for school. One of your fellow group members sends you the following slide.

Are there other ways than e-mail to share this slide with your group members? If so, please describe how you would do this:



Q7.3 You want to improve the slide. Please describe below how you would do this:

Q7.4 Please take 5 minutes to actually create an improved slide.

If you had not thought of this already, please add a video of an animal to the presentation that you can download from pixabay animal videos.

After you have added the video, please upload the new slide here:

Q7.5
You are also going to try and find an image to add to your presentation about climate change.

You would like an image that contains polar bears and melting ice. **Make sure that you are allowed to use the image freely (i.e. there is no copy right).**

Please open a new window, search the internet and find an image that fits the description above.

Then upload the image here:

Q8.1
You finished all the tasks!

Thank you so much for taking the time to participate. We are very interested in your opinion about the questions you answered and the tasks you completed. Did you find them difficult? Were they easy to do? Were they fun?

If you have anything that you would like to tell us about them, please write it down in the box below.