Int Rev Educ (2015) 61:655–671
DOI 10.1007/s11159-015-9519-4

ORIGINAL PAPER

The development and validation of testing materials for literacy, numeracy and digital skills in a Dutch context

Maurice de Greef[1] · Mien Segers[2] · Jan Nijhuis[2] · Jo Fond Lam[3] · Mieke van Groenestijn[4] · Frans van Hoek[3] · Alexander J. A. M. van Deursen[5] · Ella Bohnenn[6] · Marga Tubbing[7]

Published online: 20 November 2015

Abstract Besides work-oriented training, most Dutch adult learning courses of formal and non-formal education focus on three basic skills: literacy, numeracy and problem solving in technology-rich environments. In the Netherlands, the Ministry of Education, Culture and Science recently initiated the development of a new adult education framework concerning literacy, numeracy and digital skills. In order to

✉ Maurice de Greef
ricedegreef@gmail.com

Mien Segers
m.segers@maastrichtuniversity.nl

Jan Nijhuis
jfh.nijhuis@maastrichtuniversity.nl

Jo Fond Lam
jlam@cinop.nl

Mieke van Groenestijn
mieke.vangroenestijn@kpnmail.nl

Frans van Hoek
fhoek@cinop.nl

Alexander J. A. M. van Deursen
a.j.a.m.vandeursen@utwente.nl

Ella Bohnenn
e.bohnenn@xs4all.nl

Marga Tubbing
marga@lezenenschrijven.nl

[1] Vrije Universiteit Brussel, Brussels, Belgium

[2] Maastricht University, Maastricht, The Netherlands

[3] Centrum voor Innovatie van Opleidingen (CINOP), 's-Hertogenbosch, The Netherlands

monitor the progress of literacy, numeracy and digital competencies, it is necessary to develop and validate testing materials for specific competencies. This study validates the testing materials which were developed to assess learners' proficiency in literacy (reading and writing), numeracy and digital skills based on the new Dutch framework. The outcome is that the materials proved valid and can be used in different courses referring to basic skills and adult learning, though there are still some limitations. Besides adult education professionals (such teachers and trainers), policy makers can also use the results of these tests in order to describe and monitor the impact of adult education on the lives of adult learners.

Keywords Basic skills · Adult learning · Testing materials · Literacy · Numeracy · Digital skills

Résumé Conception et validation de documents pour tester les compétences lettrées et numériques dans le contexte néerlandais – En dehors des formations à visée professionnelle, la majorité des cours de l'éducation formelle et non formelle des adultes aux Pays-Bas se concentrent sur trois compétences fondamentales : littératie, numératie et résolution de problèmes dans les environnements fortement informatisés. Le ministère néerlandais de l'éducation, de la culture et des sciences a récemment entrepris la conception d'un nouveau cadre pour l'éducation des adultes portant sur les compétences lettrées et numériques. Afin de suivre les progrès dans ces domaines, il est nécessaire de concevoir et de valider des documents permettant de tester des compétences spécifiques. L'étude présentée ici a procédé à la validation des documents de tests qui ont été conçus pour évaluer le niveau de maîtrise des apprenants en compétences lettrées (lecture et écriture) et numériques ainsi qu'en numératie, conformément à ce nouveau cadre. Selon les résultats de cette étude, les tests s'avèrent valables et peuvent être administrés dans différents cours ayant trait aux compétences de base et à l'apprentissage des adultes, même s'ils présentent encore quelques limites. Outre les professionnels de l'éducation et de la formation des adultes (tels qu'enseignants et formateurs), les décideurs pourront également exploiter les résultats de ces tests, en vue de décrire et de suivre l'impact de l'éducation des adultes sur la vie quotidienne des apprenants.

Background

According to Gadotti (2011) and McCowan (2013), learning should be understood as a human right in everyone's life. Adult education, in particular, can be supportive in stimulating sustainable development (Gadotti 2011; Gartenschlaeger and Hirsch 2015) of communities and nations. In contemporary society there are still groups of

[4] Hogeschool Utrecht, University of Applied Sciences, Utrecht, The Netherlands

[5] University of Twente, Enschede, The Netherlands

[6] Ella Bohnenn Advies, Onderzoek & Coaching, Amsterdam, The Netherlands

[7] Reading & Writing Foundation, The Hague, The Netherlands

marginalised citizens who need this kind of support in order to be and feel included in their societal environment. Reasons for their exclusion include poverty, a low proficiency in literacy, a different cultural background, school dropout, unemployment and age (De Greef et al. 2015). When social inclusion seems out of reach, adult education can make a difference. Vulnerable groups and especially vulnerable adults can be drawn in after joining a programme of adult education (De Greef et al. 2012). Adult education can, for example support learners in improving their language proficiency and digital or financial skills, it helps them to become more assertive and active in nature and sports, makes them feel less isolated and facilitates their meeting other people (ibid.). Such improvements are vital for active participation in social life on a day-to-day basis. According to the Organisation for Economic Co-operation and Development (OECD 2012) and Marieke Buisman et al. (2013), the three basic skills necessary for active and autonomous citizenship are: (1) literacy, (2) numeracy and (3) problem solving in technology-rich environments. Besides offering work-oriented training, the range of adult learning courses in the Netherlands includes many formal and non-formal education opportunities which focus on these three basic skills.

Literacy

According to the results of the Programme for the International Assessment of Adult Competencies (PIAAC),[1] 20 per cent of the working-age population (i.e. aged 16–65) in the European Union (EU) have low literacy skills (European Commission 2013). These citizens experience problems in reading and writing in everyday life. For example, reading an instruction leaflet included with their medicaction or writing a message to a colleague can be difficult for them to accomplish. In other words, their reading and writing problems are blocking their full participation in day-to-day societal activities. This constraint also has an impact on their daily work; e.g., they are unable to read safety instructions or write reports. César Guadalupe and Manuel Cardoso (2011) stress that measurement of basic skills like literacy is becoming increasingly important, especially since the commitments made by participants from all over the world at the end of the Sixth International Conference on Adult Education (CONFINTEA VI) in 2009, documented in the *Belém Framework for Action* (UIL 2010). Several international studies have been conducted in recent years to investigate the current status of literacy and its implications for different countries worldwide. These studies include, for example, UNESCO's Literacy Assessment and Monitoring Programme (LAMP)[2] and the aforementioned Programme for the International Assessment of Adult Competencies (PIAAC). These are complemented

[1] PIAAC is a cyclical survey of adult skills conducted in 24 participating countries (more than half of them in Europe) by the Organisation for Economic Co-operation and Development (OECD). The first cycle was run in 2008–13; the second one covers 2012–16 and the third one is scheduled to run 2014–18. For more information, see http://www.oecd.org/site/piaac/ [accessed 16 October 2015].

[2] LAMP, initiated in 2003 by the UNESCO Institute for Statistics (UIS), is designed to obtain differentiated information (avoiding rigid categories of literate/illiterate) required to effectively plan and implement literacy programmes. For more information, see http://www.uis.unesco.org/literacy/Pages/lamp-literacy-assessment.aspx [accessed 22 October].

by national studies which underline the impact of language courses on an increase of writing, reading and literacy skills (Reder 2009, 2010, 2012; DoLNZ 2010).

Numeracy

Besides literacy, competencies in numeracy are also important for using and processing information in everyday life. Over the past thirty years, several studies and descriptions of numeracy have been published, starting with the Crowther report (1959),[3] followed by the Cockcroft report (1982), in which the word *numeracy* was used. Numeracy was given particular attention in the United Kingdom, the USA and in Australia (ibid.). The first international study introducing standardised testing[4] was conducted by the International Association for the Evaluation of Educational Achievement (IEA) between 1959 and 1962. This was followed by research into numeracy skills, regarded as being part of literacy (so-called *quantitative literacy*), in the International Adult Literacy Survey (IALS), conducted by the OECD in the years 1990 to 1996 (Darcovich et al. 1997). This survey was then in turn succeeded by the Adult Literacy and Lifeskills (ALL) survey, run from 2003 to 2006 (Houtkoop et al. 2012), in which numeracy was described as "the knowledge and skills required to effectively manage the mathematical demands of diverse situations" (Gal et al. 2005). In the Netherlands, a first national-level study into numeracy of adults facing low literacy was carried out by Mieke van Groenestijn (2002). In her study, she described numeracy as *a dynamic concept* which, in addition to a range of mathematical and other knowledge skills, also includes the ability to accommodate and adjust flexibly to new demands in a continuously changing society.

In the more recent PIAAC study (Buisman et al. 2013, p. 19; OECD 2012), numeracy refers to "the ability to access, use, interpret and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life" (OECD 2012, p. 33).

All these definitions highlight the importance of numeracy to a wide range of skills and knowledge used in everyday life. Numeracy concerns dealing with numbers in real life situations and requires more than an understanding of basic mathematical operations. Additionally, it requires the ability to compute and interpret dimensions like proportions, measurements and statistics.

Problem solving in technology-rich environments

The use of the Internet does not benefit every citizen (equally) in daily life and work situations (Van Deursen et al. 2014). Although the use of most types of Internet activities has increased over the last years among all segments of the population, for

[3] This journal published a short note at the time, entitled "Developments in English education in 1959: The year of the Crowther Report". According to its author Thomas Jarman, "This authoritative document recommended the raising of the school-leaving age to 16 and compulsory part-time education in county colleges up to 18" (Jarman 1960, p. 231).

[4] It was called the "Pilot twelve-country study"; for more information, see http://www.iea.nl/pilot_twelve-country_study.html [accessed 16 October 2015] and Foshay et al. 1962.

most capital-enhancing activities, men, younger people, higher-educated people and people with higher than average incomes are prominent (ibid.). Furthermore, the Internet seems to provide an increasing number of capital-enhancing opportunities for those with higher education and income, thus reinforcing their already strong positions in society (ibid.). Among other reasons, one explanation is that there are differences in digital skills related to gender, age and educational background (Hargittai 2002; Helsper and Eynon 2013; Van Deursen and Van Dijk 2014). When accounting for different types of digital skills, recent studies have shown that both operating skills (in terms of the technology itself, e.g., the Internet) and processing skills (in terms of the content provided by the technology) are important (Van Dijk and Van Deursen 2014). An increasing number of people feel that they need more formal adult education to acquire and improve their digital skills (ibid.). However, participation in these courses is not particularly widespread; in developed countries on average only between five and ten per cent of people above school age have ever enrolled in a computer course. For low-skilled adults, however, special requirements like personal guidance of a teacher, learning in the classroom and at home and self-directed learning seem to be important (De Greef and Bohnenn 2011).

The new Dutch adult education framework of literacy, numeracy and digital skills

In the Netherlands, the Ministry of Education, Culture and Science recently initiated the development of a new framework entitled *Standaarden en eindtermen volwasseneneducatie* [Standards and goals of adult education] which concerns adult literacy, numeracy and the use of digital skills (CINOP 2012). The development of this framework involved the construction of new levels of competencies and the elaboration and adjustment of existing reference levels of language and arithmetic (Commissie Meijerink 2009) for adult education to meet the needs of today's society. In order to accomplish everyday tasks at home or at work, citizens must be proficient on specific levels of literacy, numeracy and digital skills. This new framework has defined 4 to 5 proficiency levels (reflecting 4 official levels and 1 "entrance level" for literacy) for the Dutch educational system concerning literacy and numeracy and possibly digital skills. The "Language for life" programme of the Dutch Reading & Writing Foundation[5] facilitated the opportunity to develop new testing materials for adult learners who need to increase their literacy, numeracy and digital skills. This programme connects several providers of formal and non-formal education in literacy (besides numeracy and digital skills) and contributes to a more accessible educational system for adults in the Netherlands. Furthermore, in co-operation with providers of adult education, the programme facilitates training sessions and courses for adults who are struggling with reading and writing.

[5] *Stichting Lezen & Schrijven*, the Dutch Reading & Writing Foundation, offers advice to companies, health care organisations and government agencies who are or want to become involved in addressing literacy-related problems among the Dutch population. For more information, see http://www.readingandwriting.eu/[accessed 19 October 2015].

With the new Dutch adult education framework in place since 2010, policy makers as well as educators, trainers and teachers, but also the adult learner him- or herself, now needed to be able to assess whether participation in the newly-designed courses was leading to any improvement. A new coherent system of testing materials to monitor individual learners' progress in the three aforementioned basic skills therefore had to be developed and validated.

Research question

The development of the new testing materials was carried out by the Dutch Reading & Writing Foundation in cooperation with a number of partners such as Maastricht University, the University of Twente and CINOP (the national Centre for the Innovation of Education and Training in the Netherlands), as well as two of the authors of this paper, Ella Bohnenn and Mieke van Groenestijn. They were designed to monitor progress at the first two levels of the new Dutch adult education framework and were earmarked to become part of the Dutch "Language for life" programme. In the study presented here, these testing materials are now validated. Our research question was: Are these testing materials suitable to measure adult learners' progress in literacy, numeracy and digital skills?

Methodology

Sample

Our sample included 926 participants enrolled in adult education courses at 23 learning centres of formal and non-formal education in 103 different groups throughout the different provinces of the Netherlands. The courses taught basic skills in formal and non-formal settings of adult education. Participants were selected at random and asked to complete one or, if possible, more (up to four) tests. They were asked to do the reading test, if possible in combination with the numeracy test or the writing test, and if possible in combination with the digital skills test. Table 1 presents an overview of participants' main characteristics per test. As can be seen in Table 1, some groups are rather small, so comparison was limited for some variables such as age, for example, where a needed distribution of equal ranges of 25 years for each group would not have been possible.

Procedure

Each learning centre was asked to select groups of learners to take the tests. Selection criteria included certain basic levels of reading, writing and numeracy. Participants therefore had to accomplish a conditional test before they were selected for inclusion in our sample. All groups of learners completed two tests during one or two different sessions, except for five groups who only completed the writing test and two groups who only completed the reading test. The participants of each group

Table 1 Socio-demographic characteristics of participants (N = 926)

Characteristic	Kind of test			
	Reading (N = 528)	Writing (N = 398)	Numeracy (N = 447)	Digital skills (N = 294)
Gender				
Male	37.2	38.7	35.5	35.2
Female	62.8	61.3	64.5	64.8
Nationality				
Native	20.8	21.4	18.5	20.1
Foreign	79.2	78.6	81.5	79.9
Age				
0–5	7.2	9.9	6.8	12.6
26–45	53.3	57.7	56.6	56.5
46–65	37.6	29.1	34.3	28.8
66–100	1.9	3.3	2.3	2.1
Highest level of education				
Primary school	21.0	18.1	22.1	18.5
Secondary school	23.1	21.4	21.6	28.1
Further education on level of middle class	24.7	23.7	23.9	20.6
Higher education/University	19.3	20.3	19.8	18.1
Other school systems	11.9	16.4	12.6	14.6
Total years of education				
5 years or less	19.6	18.3	20.1	18.7
6 to 10 years	23.8	24.3	23.3	26.6
11 to 15 years	35.7	34.0	36.7	32.7
More than 16 years	21.0	23.3	19.9	21.9
Job				
Paid work	46.6	39.3	40.0	36.3
Voluntary work	15.9	15.0	16.6	15.8
Both paid and voluntary work	4.2	2.6	4.0	1.4
None	33.3	43.2	36.5	46.4

Note Missing values are excluded in percentage calculations

completed the test in the classroom after instruction by and in attendance of a trained test leader. A team of 14 test leaders were trained to conduct the tests according to standard procedures in 103 groups of learners.

Instrument development

Reading test

Literacy proficiency was divided into two tests: a reading test and a writing test. The reading test is based on the "reading skills" section of the new Dutch adult

education framework for literacy, numeracy and digital skills (CINOP 2012). The test covers Entry level and Level 1F (reflecting the levels of functional literacy and primary school) described in this framework and measures four dimensions of reading. The test was designed to be completed in one hour and initially consisted of 61 items (version A) or 54 items (version B). After the analysis, 7 items were deleted (more on this in the Results section below). Besides this, two parallel versions of the reading test were developed for the pre-test and post-test.

Writing test

The writing test is based on the "writing skills" section of the new Dutch adult education framework (CINOP 2012). Like the reading test, the writing test covers Entry level and Level 1F described in this framework and measures ten dimensions of writing. Students were given one hour to complete the writing test, which consisted of 112 items (in both version A and version B). After the analysis, 27 items which did not fit the model were deleted for statistical reasons. Two parallel versions of the writing test were also developed for the pre-test and post-test.

Numeracy test

The numeracy test was also based on the new Dutch adult education framework (CINOP 2012), but also on the results of the European Grundtvig project entitled "InBalance", which was run in six countries from 2009 to 2011, namely: Austria, Finland, the United Kingdom, Hungary, the Netherlands and Spain. One of the outcomes was a common database with real-life numeracy problems for adult education, based on a 5-level numeracy framework.[6] In one European study, this database was complemented with a numeracy test (in two parallel versions) in order to validate the levels in the database. Our participants were given one hour to complete the numeracy test which consisted of 14 items (in both version A and version B). Besides this, we updated and validated both tests in the Netherlands on a larger scale among 489 participants.

Digital skills test

Like the first three tests, the digital skills assessment was based on the new Dutch adult education framework (CINOP 2012). There are some practical limitations concerning the environment in which digital skills can be assessed: It is not possible to conduct the test one-on-one, not everyone has access to a computer and the test should last no longer than 30 minutes. Due to the combination of these environmental factors, we were unable to realise a more advanced method of testing, such as computer-based testing with adaptability functions (which adjust the

[6] The European "InBalance" project aimed to "establish European standards in numeracy […] by supporting adult tutors, producing more effective resources & disseminating best practice […] to improve the quality of adult education tutors. This [was hoped to be] achieved by producing a European Numeracy Framework (ENF) which [would] assist adult tutors in developing high quality learning materials" (EACEA 2009, p. 3).

level of difficulty of later questions to the number of correct answers a participant gave to earlier questions) and the use of software simulations. Taking all these limitations into account, a paper-based self-assessment questionnaire with 27 items was developed instead of a test.

Testing and refinement

Before applying the test, we ran a pre-pilot among a small sample of participants at an adult education centre in the South-Eastern part of the Netherlands to investigate whether the items and the text in the assignments were understandable for adult learners and whether the test could be completed within an acceptable timeframe (of 1 hour per test for literacy and numeracy and 30 minutes per test for digital skills). After the pre-pilot, some items were deleted or adjusted in order to increase usability, understandability and consistency of the four tests.

Statistical methods of analysis

The items for the reading test were analysed and calibrated according to Item Response Theory (IRT), and the items for the writing test were analysed by using Classical Test Theory (CTT) due to the composition of the test. In order to validate the numeracy tests, two parallel versions (set A and set B) with anchored items[7] were distributed among learners in adult education programmes. Each set consisted of 12 items plus 2 anchor items. Furthermore, an Analysis of variance (ANOVA) test was used to measure the homogeneity between and within the levels of items.

Finally, in order to develop an instrument for measuring digital skills, a Principal Component Analysis (PCA) was conducted in order to reduce information into a set of weighted linear combinations of items (which were not categorical, but ratio responses): a factor analysis with the maximum likelihood extraction procedure was conducted in SPSS 15.0.[8] In order to confirm whether all items provided a good data fit, a Confirmatory Factor Analysis (CFA) was carried out in EQS.[9]

Results

Reading

In order to compose the final reading tests, two booklets were used to collect at least 260 observations per item. Afterwards, based on content and item-analysis, seven

[7] Anchored items are items which are common to two different test forms, and are thus "anchored" in the respective other form. The purpose is facilitating test form equation.

[8] SPSS, which stands for Statistical Package for the Social Sciences, is a software programme designed for statistical analysis.

[9] EQS, also designed for statistical analysis, is a "structural equations" software programme.

Table 2 Statistical figures reading test Versions A and B

Statistical figures	Version A	Version B
N	269	260
Number of items	61	54
Average raw score	38.54	36.17
Standard deviation	14.99	10.43
Percentage correct	63%	67%
Minimum – maximum	0–61	7–64
Reliability (Cronbach's alpha)	0.96	0.93
Average theta	0.117	0.180
Variance estimated theta	0.098	0.065
Variance true theta	0.088	0.057
Reliability estimated theta	0.91	0.88

items were excluded from the final test: three items were excluded due to insufficient space for the answer and insufficient clear instruction for putting the symbol of an X in a scheme, and four other items were excluded based on low or negative achievement (Rasch unit or RIT) values or extreme probability (p) values. Eventually, the final model calibrated with an IRT model (i.e. One Parameter Logistic model [OPLM]) had r1c = 362.567 and p = 0.6430 (see Table 2). Based on this model, we composed two parallel versions of the reading test.

Writing

Compared with the reading tests, the items of the final writing tests were tested by using two booklets in order to collect at least 357 observations for each item. Using Classical Test Theory, we conducted an item analysis (RIT-value <.30 and/or p-value <.10 or >.90). As a result, 27 items were excluded from the final test versions. Both versions have a reliability (Cronbach's alpha) of 0.96 (see Table 3).

Numeracy

For the reliability of the items of the numeracy test, Cronbach's alpha is 0.76 for set A and 0.80 for set B. Second, we conducted an ANOVA test in order to analyse the homogeneity of the four levels of the items. The result for set A was F = 16.314, for set B it was F = 17.907 and for the total set it was F = 33.042. This means that the sets of items have a proper level indication.

Digital skills

Based on the results of the Principal Component Analysis, it can be concluded that the digital skills test (besides the socio-demographic characteristics) consists of one scale (named digital skills) with a Cronbach's alpha of 0.98. This scale contains

Table 3 Statistical figures writing test Versions A and B

Statistical figures	Version A	Version B
N	392	357
Number of items	54	49
Mean	49 SD 23	41 SD 19
Min–max RIT value	0.24–0.89	0.23–0.89
Average RIT-value	0.56	0.57
Average p-value	0.54	0.59
Standard deviation	0.18	0.17
Min–max p-value	0.14–0.94	0.13–0.94
Reliability (Cronbach's alpha)	0.96	0.96

three components referring to the three different levels of digital skills, namely Entry level (Cronbach's alpha is 0.90), Basic level 1 (Cronbach's alpha is 0.95) and Basic level 2 (Cronbach's alpha is 0.97). Furthermore, the Confirmatory Factor Analysis showed that all items of the factors provide a good data fit, with χ^2 of 379.75 by N = 294, Comparative Fit Index (CFI) = 0.973, a Root Mean Square Error of Approximation (RMSEA) of 0.069 and a Standard Root Mean Square Residual (RMR) of 0.024.

Conclusions and discussion

Usability of the developed tests

Given the rapidly changing demands of contemporary society, basic competencies with a focus on literacy, numeracy and digital skills are becoming increasingly important. Improvements are especially necessary for people who lack these skills and are therefore unable to fully participate in social life. Such improvements can be accomplished by joining courses of adult education. We have validated the four tests discussed in this paper and found that they prove to be appropriate instruments to monitor adult learners' progress in acquiring reading, writing, numeracy and digital skills, even if some limitations still apply which need to be rectified (more on these below). Moreover, the combination of these four tests constitutes a coherent instrument for assessing adult learners' overall progress in the acquisition of basic skills.

The signatories of the *Belém Framework for Action* committed to

> developing literacy provision that is relevant and adapted to learners' needs and leads to functional and sustainable knowledge, skills and competence of participants empowering them to continue as lifelong learners whose achievement is recognised through appropriate assessment methods and instruments (UIL 2010, p. 6, section 11d).

Though numeracy and digital skills are not specifically mentioned, they can be taken to be included in a broader concept of "functional" literacy. The signatories also committed to

> establishing appropriate coordination mechanisms, such as *monitoring* committees involving all stakeholders active in adult learning and education (ibid., p. 7, section 12d – emphasis added)

In particular, initiatives in adult literacy should be reviewed systematically and progress should be reported (ibid., p. 6, section 11g). Moreover, greater support should be lent

> to systematic interdisciplinary research in adult learning and education, complemented by knowledge management systems for the collection, analysis and dissemination of data and good practice (ibid., p. 9, section 16f).

In order to ensure a sustainable system of monitoring progress in adult education, validated instruments are needed. Increasingly, testing materials or validated questionnaires are required in order to obtain data reflecting the outcome of adult education efforts and to stimulate policy makers to invest in lifelong learning based on the achieved and analysed results.

It seems that with the efforts of the 6th and 7th Framework Programmes for Research,[10] comparable European sources like the Continuing Vocation Training Survey (CVTS), the Adult Education Survey (AES) and the European Union Labour Force Survey (EU LFS) have been realised.[11] Furthermore the interim results of the OECD's ongoing PIAAC survey and analysis systems for adult education like the European Lifelong Learning Index (ELLI)[12] have made it possible to analyse basic skills levels in the context of lifelong learning, making needs and processes more transparent. Despite the availability of these instruments and databases, substantiated facts concerning the impact and outcome of adult education are still thin on the ground. According to Jon Carpentieri (2013), policy makers need more evidence in this matter. Due to the fact that for example literacy programmes have to be

[10] The European Union (EU) has been running a series of "Framework programmes for research and technological development" since 1984. These framework programmes (FPs) are a funding instrument; the current programme (2014–2020) is called "Horizon 2020". The main objective of the sixth programme (FP6; 2002–2006) was to contribute to the creation of the European Research Area (ERA); the main objective of the seventh programme (FP7; 2007–2013; with some FP7-projects still running) was also collaborative European research.

[11] The Continuing Vocational Training Survey (CVTS) is part of the EU statistics on lifelong learning. For more information, see http://ec.europa.eu/eurostat/web/microdata/continuing-vocational-training-survey [accessed 22 October 2015]. The Adult Education Survey (AES) is a household survey which is also part of the EU statistics on lifelong learning. For more information, see http://ec.europa.eu/eurostat/web/microdata/adult-education-survey [accessed 22 October 2015]. For more information on the European Union Labour Force Survey (EU LFS), see http://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey [accessed 22 October 2015].

[12] The European Lifelong Learning Indicators project (ELLI), spearheaded by the Bertelsmann Foundation, was launched in 2008 to raise awareness for and monitor the state of lifelong learning in Europe. It is based on the four pillars of learning (Delors et al. 1996): Learning to know, learning to do, learning to live together and learning to be. For more information, see Bertelsmann Stiftung 2010.

subsidised by a welfare state in which public and competing claims are vying for attention, the evaluation of education programmes is crucial (ibid.). Carpentieri underlines the indispensability of longitudinal evaluation for supporting funders' decision making (ibid.). According to the European Basic Skills Network (EBSN),[13] research-based and efficient policy recommendations are needed, particularly in the field of basic skills (EBSN 2014).

In other words, a rising number of organisations seem to underline the importance of monitoring the outcome and impact of adult education programmes. With the validation of a combination of four tests (assessing reading, writing, numeracy and digital skills), this study aims to contribute to the development and fine-tuning of the testing materials necessary to analyse and describe the outcome and impact of adult education processes. According to several evaluation studies on the impact of adult education, reading and writing seem to be two especially important skills to describe the outcome of adult education (De Greef et al. 2013). Taking the holistic approach of the PIAAC study into account, numeracy and digital skills also appear to be important outcome measures, alongside reading and writing, for describing the impact of adult education processes (Buisman et al. 2013). This perception also corroborates the holistic description of basic skills of the European Basic Skills Network (EBSN 2014). Being a validation of the combination of four tests, this study fits the concept of basic skills as used in adult education programmes. According to the results of our validation, the high figures for Cronbach's alpha in particular demonstrate a good reliability of the tests in the context of adult education settings. The combination of the tests is suitable for measuring the holistic impact of adult education in the course of longitudinal research. Each of the four tests gives insight into an individual learner's progress. Alternatively, it can be used as a diagnostic instrument, due to the fact that the tests were designed to show ongoing progress at several points throughout the learning process.

Suggestions for future improvement

There are still some limitations which have to be taken into account when using these tests to monitor the progress of adult learning. First of all, the test for assessing learners' progress in digital skills is based on self-reports. Therefore one should be aware of the fact that the result of this test will differ from the results of the other tests.

Second, we neglected oral communication and listening, which are both important basic skills and not integrated in the four tests. Though reading and writing can be tested with the same tests for native and non-native citizens, this will not work for oral communication and listening. In most cases, native citizens have a higher level at the start of the course than non-native citizens have.

Third, using the developed tests does not reveal much information about the success factors which cause learners' progress. If the results are to be used in order to adjust the courses or improve the learning environment, additional tests or

[13] The European Basic Skills Network (EBSN), which originated in the European Commission's Working Group for the Implementation of the Action Plan for Adult Learning, was officially founded in 2010. For more information, see http://www.basicskills.eu/ [accessed 23 October 2015].

methodologies will be necessary, due to the fact that these tests only give insights into the progression of the learners and are by no means diagnostic.

Fourth, according to studies by Barry Cooper and Máiréad Dunne (1998, 2000), the results of testing by using "real-life" items may be influenced by learners' socio-cultural backgrounds. Therefore it seems necessary to analyse the influence of participants' socio-cultural background characteristics to ascertain whether learning results might differ among different groups of adult learners.

Finally, as a research team we are aware of the fact that we have only realised the first steps concerning statistical analyses and more steps are necessary to define the levels of the different tests. The next step (especially for the reading and writing tests) will be the involvement of different experts to determine the different levels of reading and writing.

Future development of these testing materials should aim to test all basic skills (including oral communication and listening) and analyse coherence with the elements of the learning environment to determine which elements of the learning environment will influence the learning outcome. Furthermore, new ways of testing digital skills (besides self-reports) should be integrated to obtain results which are comparable with the results for other basic skills like literacy and numeracy. The Dutch Reading & Writing Foundation will be involved in the development of the new items, the realisation of adjustments, etc. In order to ensure that these tests are more than a one-off initiative and to realise this ongoing process, they are currently setting up co-operation with different experts in the near future.

Need for future research

Further research is necessary to strengthen the testing materials and to realise an instrument which can be used in different contexts of adult education. First of all, in order to confirm the construct validity of these tests, a Confirmatory Factor Analysis (CFA) should be conducted by using EQS 6.1.[14] This would ensure that the testing materials can be used for the analysis and description of the impact of adult education in terms of content. Second, taking congruent validity into account, it might be possible to test participants' learning outcomes using these tests in conjunction with a test of self-perception focusing on the same contents. For example, the developed Social Inclusion after Transfer (SIT) instrument of De Greef et al. (2010) analyses and describes the impact of adult education on social inclusion (including language and digital skills) based on self-perception. If one uses both tests, and the results are comparable, it seems conceivable that both tests might provide a valuable contribution to the description and analysis of the impact of adult education in different basic skills (including literacy). Third, an extra analysis of subgroups of adults (joining learning environments of adult education) might describe whether the impact differs per subgroup of adult education and indicate whether these tests are usable for the different target groups engaged in lifelong learning. For example, in the Netherlands in some cases foreigners and natives seem

[14] As already mentioned in footnote 9, EQS is a "structural equations" software programme designed for statistical analysis. EQS 6.1 is its 2012 version.

to have different needs and levels in learning the Dutch language. Differences in starting levels are likely to depend on learners' cultural and social backgrounds. Therefore, in order to use the four developed tests adequately for specialised groups in adult education, extra analyses on subgroups will be needed.

The development and validation of these four tests assessing adult learners' progress in basic skills, namely reading, writing, numeracy and digital skills, constitutes a first attempt in the realisation of a holistic testing instrument with which to analyse and describe the impact of adult education. According to the *Belém Framework for Action* (UIL 2010) and the results of previous studies, this seems to be a desirable step to strengthen facilitation of adult education. However, further research is necessary to ensure that the developed testing instruments can bridge the gap between different adult education stakeholders and policy makers who need to optimise the provision of learning environments of adult education. With the development of the four different tests, Dutch professionals and policy makers now have the ability to analyse and describe the progression of adult learners concerning elementary skills for participation in social life. The tests have been proved valid and can be used in basic skills and adult learning courses. Besides professionals in the field of adult education (such as teachers and trainers), policy makers can also use these results in order to describe and monitor the impact adult education can have on the lives of adult learners. This constitutes a contribution to one of the commitments of the *Belém Framework for Action* and to a variety of agendas of both European and international networks and organisations supporting the realisation and increase of quality in adult education.

Acknowledgements This research has been possible due to the financial support of the Dutch Reading & Writing Foundation in the Netherlands. Thanks to their support, the research team was able to develop and validate the four tests assessing adult learners' progress in reading, writing, numeracy and digital skills as part of the Dutch *Language for Life* programme.

References

Bertelsmann Stiftung (2010). ELLI indext in brief. Gütersloh: Bertelsmann Foundation. Accessed 23 October 2015, from https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Presse/imported/downloads/xcms_bst_dms_32063_32087_2.pdf.

Buisman, M., Allen, J., Fouarge, D., Houtkoop, W. & Van der Velden, R. (2013). *PIAAC: Kernvaardigheden voor werk en leven* [PIAAC: Basic skills for work and life]. 's-Hertogenbosch: Expertisecentrum Beroepsunderwijs (ECBO).

Carpentieri, J. D. (2013). Evidence, evaluation and the "tyranny of effect size": A proposal to more accurately measure programme impacts in adult and family literacy. *European Journal of Education, 48*(4), 543–556.

CINOP (Centrum voor Innovatie van Opleidingen). (2012). *Standaarden en eindtermen VE* [Standards and terms of outcome of adult education (AE)]. 's-Hertogenbosch: CINOP.

Cockcroft, W. H. (1982). *Mathematics counts: Report of the Committee of Inquiry into the Teaching of Mathematics in Schools.* London: Her Majesty's Stationery Office. Accessed 16 October 2015, from http://www.educationengland.org.uk/documents/cockcroft/cockcroft1982.html.

Commissie Meijerink. (2009). *Referentiekader Taal en Rekenen: de referentieniveaus* [Framework of reference of literacy and numeracy: Levels of references]. Enschede: Doorlopende leerlijnen Taal en Rekenen.

Cooper, B., & Dunne, M. (1998). Anyone for tennis? Social class differences in children's responses to national curriculum mathematics testing. *The Sociological Review, 46*(1), 115–148.

Cooper, B., & Dunne, M. (2000). *Assessing children's mathematical knowledge: Social class, sex and problem-solving*. Buckingham: Open University Press.

Crowther, G., et al. (1959). 15 to 18. A report of the Central Advisory Council for Education (England) (Vol. I). London: Her Majesty's Stationery Office. Accessed 16 October 2015, from http://www.educationengland.org.uk/documents/crowther/crowther1959-1.html.

Darcovich, N., Jones, S., Kirsch, I. S., Levin, H. M., Murray, T. S., Rubenson, K., et al. (1997). *Literacy skills for the knowledge society: Further results from the International Adult Literacy Survey (IALS)*. Paris: OECD.

De Greef, M., & Bohnenn, E. (2011). *Ondersteuning en certificering van digitaal leren voor laagopgeleiden* [Support and certification of digital learning for the low-skilled]. 's-Hertogenbosch: Artéduc.

De Greef, M., Segers, M., & Nijhuis, J. (2013). *Feiten en cijfers geletterdheid* [Facts and figures literacy]. Den Haag: Stichting Lezen & Schrijven.

De Greef, M., Segers, M., & Verté, D. (2010). Development of the SIT, an instrument to evaluate the transfer effects of adult education programs for social inclusion. *Studies in Educational Evaluation, 36*(1), 42–61.

De Greef, M., Segers, M., & Verté, D. (2012). Understanding the effects of training programs for vulnerable adults on social inclusion as part of continuing education. *Studies in Continuing Education, 34*(3), 357–380.

De Greef, M., Verté, D., & Segers, M. (2015). Differential outcomes of adult education on adult learners' increase in social inclusion. *Studies in Continuing Education, 37*(1), 62–78.

Delors, J., et al. (1996). Learning: The treasure within. Report to UNESCO of the International Commission on Education for the Twenty-First Century. Paris: UNESCO.

DoLNZ (Department of Labour New Zealand). (2010). Upskilling partnership programme: Evaluation report. Wellington: Department of Labour New Zealand.

EACEA (Education, Audiovisual and Culture Executive Agency). (2009). InBalance progress report: Public part. Brussels: EACEA. Accessed 22 October 2015, from http://eacea.ec.europa.eu/LLp/project_reports/documents/grundtvig/multilateral_projects_2009/grundtvig-504006-nl.pdf.

EBSN (European Basic Skills Network). (2014). *EBSN Annual report 2014–2015*. Budapest: EBSN.

European Commission. (2013). *The Survey of Adult Skills (PIAAC): Implications for education and training policies in Europe.* Brussels: European Commission. Accessed 15 October 2015, from http://ec.europa.eu/education/policy/strategic-framework/doc/piaac_en.pdf.

Foshay, A. W., Thorndike, R. L., Hotyat, F., Pidgeon, D. A., & Walker, D. A. (1962). *Educational achievements of thirteen-year-olds in twelve countries: Results of an international research project, 1959–1961*. Hamburg: UNESCO Institute for Education.

Gadotti, M. (2011). Adult education as a human right: The Latin American context and the ecopedagogic perspective. *International Review of Education, 57*(1–2), 9–25.

Gal, I., Van Groenestijn, M., Manly, M., Schmitt, M. J., & Tout, D. (2005). Adult numeracy and its assessment in the ALL survey: A conceptual framework and pilot results. In S. T. Murray, Y. Clermont, & M. Binkley (Eds.), *Measuring adult literacy and life skills: New frameworks for assessment* (pp. 137–191). Ottawa: Statistics Canada.

Gartenschlaeger, U., & Hirsch, E. (2015). *Adult education in an interconnected world: Cooperation in lifelong learning for sustainable development.* Festschrift in honour of Heribert Hinzen. Bonn: DVV International. Accessed 15 October 2015, from http://rur.pascalobservatory.org/sites/default/files/ipe_71_web.pdf.

Guadalupe, C., & Cardoso, M. (2011). Measuring the continuum of literacy skills among adults: educational testing and the LAMP experience. *International Review of Education, 57*(1–2), 199–217.

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday, 7*(4). Accessed 15 October 2015, from http://firstmonday.org/article/view/942/864.

Helsper, E., & Eynon, R. (2013). Distinct skill pathways to digital engagement. *European Journal of Communication, 28*(6), 671–696.

Houtkoop, W., Allen J., Buisman, M., Fouarge, D., & Van der Velden, R. (2012). *Kernvaardigheden in Nederland: Resultaten van de Adult Literacy and Life Skills Survey (ALL)* [Basic skills in the Netherlands: Results of the Adult Literacy and Life Skills Survey (ALL)]. 's Hertogenbosch: ECBO.

Jarman, T. L. (1960). Developments in English education in 1959: The year of the Crowther Report. *International Review of Education, 6*(1), 231–234.

McCowan, T. (2013). *Education as a human right: Principles for a universal entitlement to learning*. London: Bloomsbury Academic.
OECD (Organisation for Economic Co-operation and Development). (2012). *Literacy, numeracy and problem solving in technology-rich environments: Framework for the OECD survey of adult skills*. Paris: OECD Publishing. Accessed 15 October 2015, from http://www.oecd.org/site/piaac/PIAAC%20Framework%202012–%20Revised%2028oct2013_ebook.pdf.
Reder, S. (2009). The development of literacy and numeracy in adult life. In S. Reder & J. Bynner (Eds.), *Tracking adult literacy and numeracy: Findings from longitudinal research* (pp. 59–84). New York: Routledge.
Reder, S. (2010). Participation, life events and the perception of basic skills improvement. In J. Derrick, J. Field, P. Lavender, S. Meyer, U. Howard, & T. Schuller (Eds.), *Remaking adult learning: Essays on adult education in honour of Alan Tuckett* (pp. 31–55). London: Institute of Education.
Reder, S. (2012). *The longitudinal study of adult learning: Challenging assumptions*. Montreal: The Centre for Literacy.
UIL (UNESCO Institute for Lifelong Learning). (2010). *Confintea VI: Belém Framework for Action*. Hamburg: UIL.
Van Deursen, A. J. A. M., & Van Dijk, J. A. G. M. (2014). Internet skill levels increase, but gaps widen: A longitudinal cross-sectional analysis (2010–2013) among the Dutch population. *Information, Communication and Society, 18*(7), 782–797.
Van Deursen, A. J. A. M., Van Dijk, J. A. G. M., & Ten Klooster, P. M. (2014). Increasing inequalities in what we do online. A longitudinal cross-sectional analysis of Internet activities among the Dutch population (2010 to 2013) over gender, age, education, and income. *Informatics and Telematics, 32*(2), 259–272.
Van Dijk, J. A. G. M., & Van Deursen, A. J. A. M. (2014). *Digital skills: Unlocking the information society*. New York: Palgrave Macmillan.
Van Groenestijn, M. (2002). *A gateway to numeracy. A study of numeracy in adult basic education*. Utrecht: CD β-Press, Centrum voor Didactiek van Wiskunde, Universiteit Utrecht.

The authors

Maurice de Greef is a visiting professor working in the field of low-skilled and illiterate learners at Vrije Universiteit Brussel in Belgium.

Mien Segers is a professor in the field of corporate learning at Maastricht University in The Netherlands.

Jan Nijhuis is a researcher and project coordinator at Maastricht University in The Netherlands.

Jo Fond Lam is a senior manager in education at CINOP, the national Centre for the Innovation of Education and Training in The Netherlands.

Mieke van Groenestijn is an Emeritus Professor of numeracy at Hogeschool Utrecht, University of Applied Sciences, in The Netherlands.

Frans van Hoek is a senior manager in education at CINOP, the national Centre for the Innovation of Education and Training in The Netherlands.

Alexander J. A. M. van Deursen is an assistant professor at the University of Twente, Enschede, in The Netherlands.

Ella Bohnenn is an advisor and researcher concerning education at Ella Bohnenn Advies, Onderzoek & Coaching in Amsterdam, The Netherlands.

Marga Tubbing is an advisor in education and educational processes at the Dutch Reading & Writing Foundation in The Hague, The Netherlands.