Heide, I. van der, Uiters, E., Sorensen, K., Rothlin, F., Pelikan, J., Rademakers, J., Boshuizen, H.
Health literacy in Europe: the development and validation of health literacy prediction models.
European Journal of Public Health: 2016, 26(6), 906-911

Health literacy in Europe: the development and validation of health literacy prediction models

IRIS VAN DER HEIDE[1], ELLEN UITERS[2], KRISTINE SØRENSEN[3], FLORIAN RO¨ THLIN[4], JU¨ RGEN PELIKAN[5], JANY RADEMAKERS[3,6], HENDRIEK BOSHUIZEN[2,7]; ON BEHALF OF THE EPHORT CONSORTIUM

1 Academic Medical Center, University of Amsterdam, Coronel Institute of Occupational Health, PO Box 22700, 1100 DE, Amsterdam, The Netherlands
2 National Institute of Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, MA Bilthoven, 3721, The Netherlands
3 CAPHRI School for Public Health and Primary Care, Maastricht University, PO Box 61, 6200 MD, Maastricht, The Netherlands
4 Gesundheit O¨ sterreich GmbH, Stubenring 6, 1010, Vienna, Austria
5 Ludwig Boltzmann Institut Health Promotion Research, Untere Donaustraße 47, A-1020, Vienna, Austria
6 Netherlands Institute for Health Services Research (NIVEL), PO Box 1568, 3500 BN, Utrecht, The Netherlands
7 Biometrics, Wageningen University, PO Box 16, 6700 AA, Wageningen, The Netherlands

In this publication the terms ‘performance-based’ and ‘self-assessed’ health literacy were used, where in the HEALIT4EU project respectively the terms ‘objective’ and ‘perceived’ health literacy were used to refer to the two types of health literacy.1 Background: Health literacy is an important determinant of health, but national health literacy levels are known for only some European countries. This study aims to examine to what extent national health literacy levels can be estimated based on publicly available census data. Method: Multivariate models were used to predict two types of health literacy on population level. Predictors were selected based on literature, the European Health Literacy Survey (HLS-EU) and the Adult Literacy and Life Skills Survey (ALL). The HLS-EU provides insight into self-assessed health literacy and the ALL into the performance of individuals on health literacy tasks (performance-based health literacy). Dutch HLS-EU and ALL data were used to construct prediction models based on 2/3 of this data, which were validated in the remaining 1/3 of the data and (in case of self-assessed health literacy) in data from seven other European countries. Results: Education is a significant predictor of perceived and performance-based health literacy. Age and working status are significant predictors of performance-based health literacy, whereas gender and income are significant predictors of self-assessed health literacy. Both typologies of health literacy can satisfactorily be predicted within samples

of the Dutch population. The accuracy of estimated self-assessed health literacy varied between the seven other European countries. Conclusion: Prediction models based on publicly available census data can be used for estimating self-assessed and performance-based health literacy on population level. Observed health literacy levels or better prediction models are required when one is interested in ranking European countries.

INTRODUCTION

Health literacy entails the skills to access, understand, appraise and apply information to make health-related decisions.2 Health literacy is recognized as an important determinant of health3–7 and there is a social gradient for health literacy as there is for health.8,9 In the past decade the number of studies indicating that lower health literacy is related to adverse health outcomes and higher healthcare costs has grown exponentially.10,11 Studies show that lower health literacy is associated with a lower mental and physical health status, adverse disease specific outcomes, mortality, more use of healthcare and less use of preventive care.10 Also poorer self-management and less satisfaction with care are associated with a lower level of health literacy.10,12 Although the topic of health literacy has gained attention in European research, policy and practice, only for some European member states knowledge is available concerning the level of health literacy of their population.8 Measures of health literacy require in-person assessment, which is time-consuming and costly.

In many European countries health literacy and national health literacy assessment are not a focus of policy,1 but there are few where it already is (e.g. Austria, Ireland). It is important to have insight into health literacy of populations across Europe, in order to identify countries or regions in which health literacy of the population is relatively low or in order to evaluate the effect of interventions. Unlike many health determinants which are difficult to modify, competencies could be taught and learned. Given the fact that competencies are distributed unevenly within a population, often along socio-demographic characteristics, a competenciesbased approach focusing on health literacy can enrich current understanding of health inequalities. This requires insight into populations' level of health literacy.

Most European countries have insight into the socioeconomic and demographic distribution of their population. Since previous research shows that health literacy is closely related to socioeconomic and demographic characteristics,8,13 it has been put forward that health literacy on population level could be estimated by applying prediction models based on these characteristics, when actual health literacy measures are absent. Yet few studies have explored this possibility.14,15 This includes the study of Martin and colleagues,14 who presented a health literacy prediction model, based on data obtained from the American National Assessment of Adult Literacy. Their study shows that gender, age, ethnicity, level of education, income, marital status and time of residence in the United States are significant predictors of health literacy.

To date, no prediction models are available to estimate health literacy on national population level in European countries.

The aim of this study is therefore to examine to what extent national health literacy levels can be validly estimated from socioeconomic and demographic characteristics

that are publicly available. Therefore, two types of health literacy will be estimated: performance-based health literacy and self-assessed health literacy.
This distinction is relevant because performance-based and selfassessed health literacy capture different constructs and may differ in the extent to which they can be predicted by socioeconomic and demographic characteristics. Performance a based measures have the strength to measure an ability directly while the self-assessment based measures have the strength to allow a more comprehensive grab of the intended measurement construct and are rooted in the experiences of the patients. The weaknesses on the other hand are, that performance a based indices tend to measure narrow content16 and self-assessed health literacy measures do not only transport social desirability but also subjective expectations, experiences, perceptions, and concepts of health care.

METHODS

Data sample

Performance-based health literacy

Dutch data from the 2008 Adult Literacy and Life Skills (ALL) survey was used to obtain a measure of performance-based health literacy on population level (see Refs. 17 and 18 for detailed information).
ALL data were gathered through in-person interviews among 5617 respondents from the Netherlands aged 16–65. The ALL measures individuals' performances on reading and problem solving tasks. Based on various health-related tasks included in the ALL, the Health Activities and Literacy Scale (HALS) was developed by the Educational Testing Service to measure health literacy.18 The HALS reflects people's performance on reading and problem solving tasks related to health information.18 Health literacy as assessed by the HALS will therefore be referred to as 'performance-based' health literacy. HALS scores ranged between 0 and 500 points,18 which could be divided into five performance categories reflecting an increasing ability to successfully complete tasks of a given difficulty (see Ref. 4 and Appendix 1). Scores above 275 reflect having adequate health literacy.18 The HALS scores were included both as a continuous outcome (mean scores) and as dichotomous outcome (having adequate versus less than adequate health literacy). The HALS scores were normally distributed.

Self-assessed health literacy

Dutch data from the 2011 European Health Literacy Survey (HLSEU) was used to obtain a measure of self-assessed health literacy on population level (see Ref. 8 for detailed information) by in person interviews. The HLS-EU was conducted in Austria, Bulgaria, Germany, Greece, Ireland, the Netherlands, Poland and Spain, and included _1000 respondents per country aged 15 years or older. The HLS-EU contains an instrument to measure self-assessed difficulties with health-related information,8,19 which will be referred to as 'selfassessed health literacy'. The 47-item HLS-EU Questionnaire (HLSEU- Q), as included in the HLS-EU, was used to measure individuals' perceived difficulties with accessing, understanding, appraising and applying health-related information on a four-point Likert scale, ranging from 1 (very difficult) to 4 (very easy). Index scores were calculated as the sum scores of all

items (see Appendix 1).8 Scores above 33 reflect sufficient or excellent health literacy, referred to as 'adequate health literacy' The HLS-EU-Q index scores, which were approximately normally distributed for all countries in the HLS-EU,9 were included both as a continuous outcome (mean scores) and a dichotomous outcome (adequate vs. less than adequate health literacy).

Demographic and socioeconomic variables

Based on literature1,4,13,14 and the availability in both the ALL and the HLS-EU data, the following possible health literacy predictors were selected for constructing a prediction model: gender, education, work status, income and age. Appendix 1 describes the operationalization of the selected variables in the ALL and the HLS-EU data.

Missing values

Most variables that were included in our analyses contained few missing values (<5%). The income variable contained 18.6% missing values in the HLS-EU data and 20.2% in the ALL data, which is common in Dutch survey research. Furthermore one item from the HLS-EU-Q which did not apply to retired people contained 25.8% missings ('how easy would you say it is to: find out about efforts to promote your health at work?'). Multiple imputations by chained equations was used to handle the presence of missing values in the study data,20–22 delivering five imputed datasets. The outcomes of the individual analyses were pooled incorporating the uncertainty due to the missing values.23 The imputations were done in R 2.14.0, with use of the mice package.20,23

Statistical analysis

Step 1: Constructing health literacy prediction models

Since the Netherlands is the only European country that participated in both the ALL and the HLS-EU, data from the Netherlands was used to develop and cross-validate health literacy prediction models for performance- based and self-assessed health literacy. To construct the prediction models, two random samples were drawn: one containing two-third of the Dutch ALL data (N=3745) and one containing twothird of the Dutch HLS-EU data (N=682). Based on each of these two random samples, two types of prediction models were developed: one predicting mean health literacy scores and one predicting the proportion of respondents having adequate health literacy. To predict mean health literacy scores and having adequate health literacy, respectively backward linear and logistic regression analyses, were applied. Predictors that were not significantly associated with health literacy (P>0.05) were excluded step by step (excluding the least significant variable first) until a model was created including only significant predictors. All analyses were conducted in SAS by use of the Mianalyze procedure.

Step 2: Cross validation of the prediction models in Dutch data

The constructed prediction models were validated by applying the regression equations obtained from Step 1 on the remaining onethird of the Dutch ALL (N= 1872) and HLS-EU (N= 341) data.

Pearson's correlation coefficients were calculated to indicate the association between predicted and observed mean health literacy scores. ROC curves were calculated to indicate the degree of correspondence between the observed and predicted proportion of respondents having adequate health literacy.

Step 3: Validation of the prediction model for selfassessed health literacy in other European countries

Data from the other countries that participated in the HLS-EU (Austria, Bulgaria, Germany, Greece, Ireland, Poland and Spain) were used to further validate the prediction model for mean selfassessed health literacy. Since ALL data including performance-based health literacy was not available for other European counties, this validation step could only be made for self-assessed health literacy.
However, mean performance-based health literacy was predicted for the other European countries, to obtain a first insight into how this estimation relates to observed and estimated self-assessed health literacy.

[TABLE 1]
By applying the regression equations obtained in Step 1, self-assessed as well as performance a based mean health literacy
was predicted for each of the countries. Pearson's correlation coefficients were used to estimate the relationship between predicted and observed mean health literacy scores.

RESULTS

Sample characteristics and health literacy

Table 1 describes the characteristics of the Dutch ALL and HLS-EU samples. The mean age was 44 (SD:13) in the ALL sample and 53 (SD:19) in the HLS-EU sample. With respect to level of education, relatively many high educated people were included in both samples (37.1% tertiary educated in the ALL and 34.8% in the HLS-EU), compared with the other educational categories. Most respondents belonged to the working population (68.5% in the ALL sample and 49.7% in the HLS-EU sample). The difference between the percentages concerning occupational status is likely due to the fact that the HLS-EU includes 15-year old respondents and respondents older than 65 (33.6%) while the ALL includes respondents of maximum 65 years of age. Mean income per month was E4379 in the ALL sample and between E1850 and E2950 in the HLS-EU sample. Table 1 also shows that mean health literacy scores differ by gender, age, level of education, income and working status.

Constructing prediction models of health literacy

Predicting performance-based health literacy

Table 2 provides information on the initial and the final models to predict performance-based health literacy (using ALL data). Age, level of education and working status were significant predictors of mean performance-based health literacy; being older, having attained (pre-)primary or lower level of education (compared with tertiary or upper secondary education) and not working or being retired (compared with working or being a student) were significantly associated with having lower performance-based health literacy. Older age and having attained

(pre-) primary or lower secondary education were significant associated with lower odds of having adequate performance-based health literacy.

Predicting self-assessed health literacy

Table 3 provides information on the initial and the final models to predict self-assessed health literacy (using HLS-EU data). Level of education, monthly household income and gender were significant predictors of mean self-assessed health literacy and of having adequate self-assessed health literacy. Having attained (pre-)primary or lower secondary education (compared with tertiary or upper secondary education), having a lower monthly household income and being male were significantly associated with lower self-assessed health literacy and with lower odds of having adequate self-assessed health literacy.

[TABLE 2]

[TABLE 3]

[TABLE 4]

Cross-validation of the prediction models in Dutch data

Table 4 shows observed and predicted mean health literacy scores per decile, which indicate that performance-based as well as selfassessed health literacy can accurately be predicted for subgroups of the same population. With respect to predicting adequate health literacy, ROC curves suggest that adequate performancebased health literacy can better be predicted than adequate selfassessed health literacy: area under the curve was respectively 0.74 and 0.58 (data not shown).24

Validation of the prediction model for self-assessed health literacy in seven European countries

Table 5 shows the observed mean self-assessed health literacy scores per country as well as the predicted mean self-assessed and predicted mean performance-based health literacy scores. The predicted self-assessed health literacy scores per country do not accurately resemble the observed scores, variation of observed scores between countries is considerably wider (4,8 points) than variation of predicted scores (1,6 points). Furthermore, the predicted mean self-assessed health literacy scores do not produce an adequate ranking of countries, compared with the ranking based on observed scores.

DISCUSSION

Development of the prediction models

Our results indicate based on Dutch data that education is a significant predictor of both self-assessed and performance-based health literacy. Age and working status are significant predictors of performance- based health literacy, where gender and income are significant predictors of self-assessed health literacy. A clarification for finding different predictors for the two types of health literacy can partly be found in the literature. Literature suggests for instance that being in labour force provides learning opportunities that may positively affect people's skills to derive meaning from texts.25 It might therefore positively affect people's performance on health

literacy tasks, but not necessarily one's self-assessed difficulty with the understanding and use of health-related information.

[TABLE 5]

Furthermore, recent research shows that age-related cognitive decline can impact one's ability to fulfill health literacy tasks.26,27

This could explain why age seems stronger related to performancebased health literacy than to self-assessed health literacy, which is not only dependent on personal skills, but also on e.g. social support.28 The lack of an association between gender, income and performance- based health literacy was salient. Both variables were significant predictors of performance-based health literacy in the prediction models that were developed by Martin and colleagues.14 An implication for future research is to examine whether the differences between the present study and the study of Martin and colleagues can be attributed to for instance sample differences and/or population differences (e.g. prevalence of socioeconomic inequalities in the population).

Validation of the prediction models

Overall, the selected predictors explained more variation in performance- based health literacy than in self-assessed health literacy. This indicates that socioeconomic and demographic characteristics may be more suitable for predicting individuals' performances on healthrelated tasks than individuals' perceptions with respect to the accessibility, understandability, appraisability and applicability of health-related information. This study showed that mean performance- based health literacy, mean self-assessed health literacy and having adequate performance-based health literacy, could satisfactorily be predicted in the Dutch population. The proportion of people having adequate self-assessed health literacy could less accurately be predicted. This difference might be related to the fact that the HALS has more established cut-of points for levels of health literacy compared to the HLS-EU-Q. Additionally, selfassessed health literacy might be better predicted by psychological and social factors than socioeconomic and demographic factors.

Furthermore, this study indicates that the developed prediction models are not suitable for calculating health literacy estimates for other countries than the Netherlands and for drawing comparisons between countries. Mean differences between countries were considerably reduced and the ranking of countries based on the predicted self-assessed health literacy scores did not reflect the ranking based on actual self-assessed health literacy scores. For countries to get information on the distribution of their population's health literacy and for true bench marking, national health literacy surveys or better prediction models remain important. An implication for future research is to confirm whether performance-based health literacy can be better predicted for other countries then selfassessed health literacy. Performance-based health literacy seems a more robust construct than self-assessed health literacy and might therefore be better predicted by the use of publicly available census data.

Limitations Only for the Netherlands data were available on comprehensive performance- based health literacy, therefore the prediction model for performance-based health literacy could only be validated for the Netherlands.

Second, the HLS-EU data show a rather specific picture for the Netherlands compared with the other countries which took part in this study.8 For instance, the associations between determinants and health literacy in the Netherlands somewhat

differs from associations in other countries.8,9 Furthermore, in the HLS-EU study the Netherlands had a considerably lower response rate (36%, probably due to first contacting survey participants by telephone) and the best health literacy results compared with the other countries.8,9 This suggests that population health literacy levels may be more accurately predicted by countryspecific models.
Third, data availability restricted the range of potentially important characteristics that could be included in the prediction model. It is likely that there are unmeasured characteristics that contribute to health literacy that were not included in the model, such as quality of education, language proficiency, place of residence, social resources and ethnicity. We refrained from including ethnicity in the models, since the variables we had at our disposal did not provide an adequate reflection of ethnic minorities in the Netherlands. Moreover, the composition of ethnic minority groups in the Netherlands is different from that in other countries, so the predictive value for health literacy in other countries would have been limited.

CONCLUSION

Prediction models based on publicly available census data can be used for estimating health literacy on population level. The choice for a prediction model depends on whether one is interested in estimating the overall mean level of health literacy in a population or the proportion within a population that is expected to have adequate health literacy. Moreover, the choice depends on whether one aims for estimating performance-based or selfassessed health literacy. The accuracy of predicted self-assessed health literacy in other European countries, based on a Dutch prediction model, varies per country. Actually observed health literacy levels or better prediction models based on more and different data are required when one is interested in ranking countries.
Funding This research is based on Work Package 3 of the HEALIT4EU research project, executed under the EU Health Programme (2008–13) in the framework of contract no. 20146201 with the Consumers, Health and Food Executive Agency (Chafea) acting under the mandate of the European Commission. The content of this article represents the views of the contractor (the EPHORT consortium) and is its sole responsibility; it can in no way be taken to reflect the views of the European Commission and/or
Chafea or any other body of the European Union. The full HEALIT4EU report is available via http://ec.europa.eu/health/
health_policies/docs/2015_health_literacy_en.pdf.

CONFLICTS OF INTEREST:

None declared.

KEY POINTS _

Self-assessed and performance-based health literacy require distinct prediction models.
_ Mean self-assessed health literacy and mean performancebased health literacy can adequately be predicted within the Dutch population.
_ The accuracy of predicted mean self-assessed health literacy in seven other European countries varies between countries.

_ Predicted national self-assessed health literacy levels cannot be used for ranking countries.
_ Therefore it will be necessary to measure observed health literacy of populations in different countries.

REFERENCES

1 Heijmans M, Uiters E, Rose T, et al. Study on sound evidence for a better understanding of health literacy in the European Union. Utrecht: Netherlands Institute for Health Services Research, 2015.
2 Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80.
3 Pleasant A, Kuruvilla S. A tale of two health literacies: public health and clinical approaches to health literacy. Health Promot Int 2008;23:152–9.
4 van der Heide I, Wang J, Droomers M, et al. The relationship between health, education, and health literacy: results from the Dutch Adult Literacy and Life Skills Survey. J Health Commun 2013;18:172–84.
5 Bennett IM, Chen J, Soroui JS, White S. The contribution of health literacy to disparities in self-rated health status and preventive health behaviors in older adults. Ann Fam Med 2009;7:204–11.
6 Howard DH, Sentell T, Gazmararian JA. Impact of health literacy on socioeconomic and racial differences in health in an elderly population. J Gen Intern Med 2006;21:857–61.
7 Schillinger D, Barton LR, Karter AJ, et al. Does literacy mediate the relationship between education and health outcomes? A study of a low-income population with diabetes. Public Health Rep 2006;121:245–54.
8 HLS-EU Consortium. Comparative Report on Health Literacy in Eight EU Member States. The European Health Literacy Project 2009–2012. Vienna: Ludwig Boltzmann Institute for Health Promotion Research, 2012. (Second Revised and Extended Version, Date July 22th, 2014), Online Publication: http://lbihpr.lbg.ac.at.w8.netzwerk.com/sites/files/lbihpr/attachments/neu_rev_hls-eu_report_2015_05_13_lit.pdf
9 Sørensen K, Pelikan JM, Ro¨thlin F, et al. HLS-EU Consortium. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). Eur J Public Health 2015;25:1053–8.
10 Berkman ND, Sheridan SL, Donahue KE, et al. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155:97–107.
11 Eichler K, Wieser S, Bru¨gger U. The costs of limited health literacy: a systematic review. Int J Public Health 2009;54:313–24.
12 Easton P, Entwistle VA, Williams B. Health in the 'hidden population' of people with low literacy. A systematic review of the literature. BMC Public Health 2010;10:459.
13 van der Heide I, Rademakers J, Schipper M, et al. Health literacy of Dutch adults: a cross sectional survey. BMC Public Health 2013;13:179.
14 Martin LT, Ruder T, Escarce JJ, et al. Developing predictive models of health literacy. J Gen Intern Med 2009;24:1211–6.
15 Hanchate AD, Ash AS, Gazmararian JA, et al. The Demographic Assessment for Health Literacy (DAHL): a new tool for estimating associations between health literacy and outcomes in national surveys. J Gen Intern Med 2008;23:1561–6.
16 Jordan JE, Osborne RH, Buchbinder R. Critical appraisal of health literacy indices revealed variable underlying constructs, narrow content and psychometric weaknesses. J Clin Epidemiol 2011;64:366–79.
17 Rudd RE. Health literacy skills of U.S. adults. Am J Health Behav 2007;31:S8–18.
18 Educational Testing Service. Literacy and health in America 2004. Available at: http://www.ets.org/research/pic 19 Sørensen K, Van den Broucke S, Pelikan JM, et al. HLS-EU Consortium. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). BMC Public Health 2013;13:948.

20 R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2008.
21 Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.
22 Klebanoff MA, Cole SR. Use of multiple imputation in the epidemiologic literature.
Am J Epidemiol 2008;168:355–7.
23 Van Buuren S. Flexible Imputation of Missing Data. Boca Raton: Chapman & Hall/ CRC Press, 2012.
24 Ebell MH. Evidence-Based Practice for the Health Professions: A Free Online Course.
Available at: http://ebp.uga.edu/courses/Chapter%204%20-%20Diagnosis%20I/ 8%20-%20ROC%20curves.html (16 July 2015, date last accessed) 25 von Wagner C, Steptoe A, Wolf MS, Wardle J. Health literacy and health actions: a review and a framework from health psychology. Health Educ Behav 2009;36:860–77.
26 Baker DW, Gazmararian JA, Sudano J, et al. Health literacy and performance on the Mini-Mental State Examination. Aging Ment Health 2002;6:22–9.
27 Kobayashi LC, Wardle J, Wolf MS, von Wagner C. Cognitive Function and Health Literacy Decline in a Cohort of Aging English Adults. J Gen Intern Med 2015;30:958–64.
28 Lee SY, Arozullah AM, Cho YI. Health literacy, social support, and health: a research agenda. Soc Sci Med 2004;58:1309–21.

TABLES

Tabel 1 Sample characteristics and distribution of health literacy (HL) scores based on two-third of the Dutch ALL ($N=3745$) and Dutch HLS-EU ($N=682$) samples

Characteristics	ALL data	Performance-based HL[a]		HLS-EU data	Self-assessed HL[b]	
	Mean (SD)/%	Mean (95% CI)	% having adequate hl	Mean (SD)/%	Mean (95% CI)	% having adequate hl
Gender						
Male	43.9	271.4 (269.8–273.2)	48.0	49.3	36.2 (35.4–36.9)	67.4
Female	56.1	269.8 (268.3–271.3)	45.3	50.7	37.4 (36.7–38.0)	72.9
Age	44 (13)			53 (19)		
16–24	8.4	276.2 (272.8–280.7)	53.4	8.5	34.5 (33.1–36.1)	54.8
25–39	29.2	278.6 (277.0–280.9)	56.5	18.0	37.5 (36.5–38.5)	78.9
40–49	24.8	275.1 (273.0–277.2)	51.4	14.1	36.8 (35.5–38.1)	71.9
50–65	37.6	259.9 (257.9–261.5)	33.9	25.8	37.1 (36.1–38.0)	71.1
65+	–	–	–	33.6	36.6 (35.7–37.5)	68.0
Level of education						
Pre-primary or primary	6.4	230.4 (226.7–237.0)	11.9	8.0	33.5 (31.2–35.7)	52.2
Lower secondary	24.8	254.2 (252.4–256.5)	24.2	26.3	36.0 (35.7–37.5)	63.7
Upper secondary non-tertiary	31.8	271.9 (269.5–272.9)	46.2	30.9	37.4 (35.1–36.9)	74.5
Tertiary	37.1	287.1 (285.9–288.9)	67.6	34.8	37.5 (36.9–38.1)	75.4
Total household income per month	4379 (5791)			–		
Less than €1350		–		19.4	35.5 (34.2–36.5)	61.6
€1350 to under €1850	–	–	–	17.4	35.6 (34.4–36.8)	64.8
€1850 to under €2400	–	–	–	18.0	38.2 (37.0–39.5)	74.1
€2400 to under €2950	–	–	–	14.2	36.7 (35.3–37.7)	69.4
€2950 to under €3600	–	–	–	14.3	36.9 (35.6–37.9)	72.8
€3600 to under €4400	–	–	–	7.7	36.9 (35.7–39.1)	74.3
€4400 or more	–	–	–	9.0	38.8 (37.8–40.7)	84.7
Working status						
Working	68.5	275.3 (274.1–276.7)	51.9	49.7	37.4 (36.7–38.0)	74.7
Not working[c]	13.7	253.1 (249.7–256.1)	28.8	14.6	35.9 (34.5–37.3)	64.4
Retired	11.1	256.1 (252.4–258.9)	28.1	28.9	36.6 (35.6–37.5)	69.1
Student	6.6	281.0 (277.6–286.0)	57.8	6.8	34.7 (33.0–37.0)	54.1

a: Based on the HALS: rage 0–500,
b: Based on the HLS-EU-Q: range 0–50,
c: Including those who state to: be unemployed, conduct unpaid household work, be permanently disabled, to be a full-time homemaker, parent or caregiver, to be inactive.

Table 2 HL prediction models based on two-third of the Dutch ALL dataset ($N = 3745$)

	Predicting mean performance-based HL				Predicting having adequate performance-based HL			
	Initial model		Final model		Initial model		Final model	
	B (95% CI)		B (95% CI)		log odds (95% CI)		log odds (95% CI)	
Constant	**296.22**	**(290.37 to 302.07)**	**299.06**	**(294.79 to 303.33)**	**1.28**	**(0.74 to 1.82)**	**1.59**	**(1.24 to 1.94)**
Gender								
Male	0.67	(−1.91 to 3.24)			−0.02	(−0.23 to 0.19)		
Age	**−0.41**	**(−0.51 to −0.31)**	**−0.40**	**(−0.50 to −0.30)**	**−0.03**	**(−0.03 to −0.02)**	**−0.03**	**(−0.04 to -0.02)**
Level of education (ref. tertiary or upper secondary)								
Pre-primary or primary	**−44.71**	**(−49.90 to −39.51)**	**−45.60**	**(−50.86 to −40.34)**	**−2.15**	**(−2.62 to −1.68)**	**−2.34**	**(−2.78 to −1.90)**
Lower secondary	**−22.75**	**(−25.99 to −19.51)**	**−22.99**	**(−25.72 to −20.25)**	**−1.31**	**(−1.67 to −0.96)**	**−1.42**	**(−1.68 to −1.18)**
Total household income per month	0.0004	(−0.0003 to 0.001)			0.00005	(−0.0001 to 0.0002)		
Working status (ref. working)								
Not working	**−10.66**	**(−14.71 to −6.62)**	**−11.02**	**(−14.56 to −7.48)**	−0.45	(−0.77 to −0.13)		
Retired	**−6.65**	**(−11.25 to −2.05)**	**−7.34**	**(−12.51 to −2.16)**	−0.41	(−0.81 to −0.01)		
R^2	0.24		0.23		0.21		0.19	

a: All significant associations are printed in bold.

Table 3 HL prediction models based on two-third of the Dutch HLS-EU dataset ($N = 682$)

	Predicting mean self-assessed HL				Predicting having adequate self-assessed HL			
	Initial model		Final model		Initial model		Final model	
	B (95% CI)		B (95% CI)		log odds (95% CI)		log odds (95% CI)	
Constant	**32.50**	**(30.01 to 34.99)**	**34.04**	**(32.00 to 36.07)**	−0.10	(-0.97 to 0.77)	0.10	(−0.63 to 0.82)
Gender								
Male	**1.46**	**(0.48 to 2.44)**	**1.42**	**(0.45 to 2.39)**	**0.38**	**(0.02 to 0.73)**	**0.35**	**(0.01 to 0.70)**
Age	**0.04**	**(0.01 to 0.08)**			0.005	(−0.01 to 0.02)		
Level of education (ref. tertiary or upper secondary)								
Pre-primary or primary	**−3.36**	**(−5.29 to -1.44)**	**−3.38**	**(−5.27 to −1.50)**	**−0.80**	**(−1.42 to −0.18)**	**−0.81**	**(−1.42 to -0.21)**
Lower secondary	−1.11	(−2.26 to 0.03)	**−1.15**	**(−2.28 to −0.02)**	**−0.41**	**(−0.81 to −0.01)**	**−0.42**	**(−0.82 to −0.03)**
Total household income per month	**0.29**	**(0.01 to 0.57)**	**0.33**	**(0.06 to 0.61)**	**0.12**	**(0.02 to 0.22)**	**0.13**	**(0.02 to 0.23)**
Working status (ref. working)								
Not working	**−1.78**	**(−3.39 to −0.16)**			−0.36	(−0.92 to 0.20)		
Retired	−1.40	(−3.07 to 0.28)			−0.18	(−0.77 to 0.41)		
R^2	0.06		0.05		0.06		0.06	

a: All significant associations are printed in bold.

Table 4 Observed and predicted mean HL scores overall and per decile based on Dutch data

Per decile	Observed mean performance-based HL 269.79[a]	Predicted mean performance-based HL 270.64[a]	Observed mean self-assessed HL 36.49[b]	Predicted mean self-assessed HL 36.79[b]
1	234.04	235.01	32.53	33.82
2	252.38	250.37	35.90	35.40
3	259.66	259.99	35.21	36.04
4	262.96	267.06	35.75	36.28
5	271.73	273.95	37.30	36.66
6	277.65	277.77	36.17	37.17
7	281.65	280.45	37.30	37.50
8	283.00	282.83	37.44	37.86
9	288.63	285.64	38.44	38.40
10	285.63	289.03	38.98	39.06

a: $r = 0.50$ ($P < 0.0001$),
b: $r = 0.23$ ($P < 0.0001$).

Table 5 Associations between observed and predicted mean health literacy (hl) scores in seven European countries

	Mean self-assessed HL			Mean performance-based HL	
Country	Observed	Predicted	*r* (CI)	Predicted	*r* (CI)[a]
Austria	31.9	37.0	0.12 (0.05–0.18)	274.3	0.18 (0.12–0.23)
Bulgaria	29.7	36.8	0.33 (0.27–0.38)	271.0	0.35 (0.30–0.40)
Greece	33.5	36.1	0.44 (0.38–0.49)	264.0	0.48 (0.43–0.52)
Spain	32.5	35.6	0.24 (0.18–0.29)	257.0	0.28 (0.23–0.34)
Ireland	34.7	36.8	0.23 (0.17–0.29)	268.1	0.15 (0.09–0.21)
Poland	33.3	37.2	0.26 (0.20–0.32)	268.3	0.29 (0.23–0.34)
Germany	34.5	36.8	0.24 (0.18–0.30)	264.1	0.15 (0.09–0.21)

a: correlation between observed self-assessed HL and predicted performance-based HL.