

# Adapting the European Health Literacy Survey Questionnaire for Fourth-Grade Students in Germany: Validation and Psychometric Analysis

Torsten Michael Bollweg, MA; Orkan Okan, PhD; Alexandra Maria Freĳian, MSc; Janine Bröder, MSc; Olga Maria Domanska, MSc; Susanne Jordan, Dr PH; Dirk Bruland, MSc, PhD; Paulo Pinheiro, MD; and Ullrich Bauer, PhD

## ABSTRACT

**Background:** Until now, children younger than age 13 years have received little attention in research on health literacy. Although some tools assess children's health literacy, no validated tool is available that assesses self-reported health literacy in a systematic and comparable way. The European Health Literacy Survey Questionnaire (HLS-EU-Q) is a valid and reliable measure of adults' self-reported health literacy. It has also been used among adolescents, but it has never been adapted for use with children. We believe it would be worth adapting for younger age groups so that self-reported health literacy could be assessed continuously. **Objective:** This study aimed to quantitatively test an adapted scale based on the HLS-EU-Q developed for German-speaking children age 9 and 10 years. **Methods:** An adapted 26-item HLS-EU-Q scale was given in a paper-and-pencil survey to 907 fourth-grade students in North Rhine-Westphalia, Germany. The psychometric properties of the scale were investigated with item analysis and factor analyses, and both convergent and discriminant validity were assessed. **Key Results:** Of the 26 tested items, 9 were discarded due to poor performance in terms of missing values, item difficulty, and factor structure. This left a 15-item scale with a high internal consistency ( $\alpha = .791$ ) that takes only a short time to administer. The scale, called the HLS-Child-Q15, had a low correlation with functional health literacy ( $r = .107, p < .001$ ), and a moderate correlation with indicators of self-efficacy ( $\rho = .280$  to  $.306, p < .001$ ). The latter indicates adequate discriminant validity, whereas the former points to a need to further investigate convergent validity. **Conclusions:** This is the first study to apply an age-adapted version of the HLS-EU-Q to children. Statistical analyses indicated the successful development of a promising instrument, but further research is needed on its factor structure and validity. This study contributes significantly to the comparative assessment of health literacy across the life course by providing a measurement tool for children age 9 and 10 years. [HLRP: Health Literacy Research and Practice. 2020;4(3):e144-e159.]

**Plain Language Summary:** The European Health Literacy Survey Questionnaire was adapted for German-speaking 9- and 10-year-old children, and 26 adapted items were tested in a written survey of 907 children. Item analysis resulted in a 15-item scale with satisfactory psychometric properties. This scale, the HLS-Child-Q15, shows high internal consistency and can be used to assess self-reported health literacy in German-speaking 9- and 10-year-old children. Nonetheless, further studies are needed to validate these results.

There is consensus among researchers and policy-makers that the promotion of health literacy (HL) at an early age could be foundational for health literacy, overall health, and quality of life throughout the life course (Borzekowski, 2009; Paakkari & Paakkari, 2012; Public Health England, 2015; Schaeffer, Hurrelmann, Bauer, &

Kolpatzik, 2018). In particular, the World Health Organization (WHO) has recognized HL as "a critical determinant of health" and stated that HL "must be an integral part of the skills, and competencies developed over a lifetime, first and foremost through the school curriculum." (WHO, 2017). It could be argued that HL is not important

for children, as it mainly concerns the ability of (adult) patients to understand health care instructions, make appropriate health decisions, and navigate the health care system (Kickbusch & Ratzan, 2001; Ratzan & Parker, 2000). However, in recent years, a public health perspective has been gaining traction in HL research that emphasizes the importance of HL beyond the health care setting. According to a more recent definition, for instance, “health literacy [...] entails people’s knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life [...]” (Sørensen et al., 2012). Although the scope of health-related decisions that are taken by children may be somewhat limited, it has been shown that children do actively access and appraise health information (Fairbrother, Curtis, & Goyder, 2016). Also, even at a young age, children are able to make health decisions that might affect their health: “A 7-year-old may or may not put on a helmet when riding his or her scooter to school. An 11-year-old has a choice when offered to try a cigarette.” (Borzekowski, 2009, p. 287). Accordingly, the promotion of HL early in the life course, as suggested by

the WHO, holds the potential to empower children regarding their health and build a foundation for good HL throughout life.

## STATE OF RESEARCH

The promotion of HL should be informed by evidence, so there is a need for appropriate measurement tools to monitor and evaluate interventions (McCormack, Haun, Sørensen, & Valerio, 2013). Currently, however, there is hardly any reliable or comparable data on the development and distribution of HL among children (defined here as people younger than age 13 years). Although three systematic reviews have identified 18 measurement tools that assess HL in this age group (Guo et al., 2018; Okan et al., 2018; Ormshaw, Paakkari, & Kannas, 2013), these tools differ vastly in their measurement approaches, the components of HL they measure, and the health areas they address (Bollweg & Okan, 2019). Most tools have not been developed specifically for children, and some either lack or do not report adequate psychometric properties (Bollweg & Okan, 2019). Accordingly, there is a need for comparable and validated tools designed to assess children’s HL. This applies particularly to the context of this

---

*Torsten Michael Bollweg, MA, is a Research Fellow, Centre for Prevention and Intervention in Childhood and Adolescence (CPI), Faculty of Educational Science, Bielefeld University. Orkan Okan, PhD, is a Research Fellow and Project Manager, CPI, Faculty of Educational Science, Bielefeld University. Alexandra Maria Frejtan, MSc, is a Research Fellow, CPI, Faculty of Educational Science, Bielefeld University. Janine Bröder, MSc, is a Research Fellow, CPI, Faculty of Educational Science, Bielefeld University. Olga Maria Domanska, MSc, is a Researcher, Department of Epidemiology and Health Monitoring, Robert Koch Institute. Susanne Jordan, Dr PH, is a Senior Researcher and the Deputy Head, Division of Health Behavior, Department of Epidemiology and Health Monitoring, Robert Koch Institute. Dirk Bruland, MSc, PhD, is a Senior Researcher, CPI, Faculty of Educational Science, Bielefeld University. Paulo Pinheiro, MD, is a Senior Researcher, CPI, Faculty of Educational Science, Bielefeld University. Ullrich Bauer, PhD, is a Professor and the Head, Department of Socialisation Research, CPI, Faculty of Educational Science, Bielefeld University.*

©2020 Bollweg, Okan, Frejtan, et al.; licensee SLACK Incorporated. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (<https://creativecommons.org/licenses/by-nc/4.0>). This license allows users to copy and distribute, to remix, transform, and build upon the article non-commercially, provided the author is attributed and the new work is non-commercial.

Address correspondence to Torsten Michael Bollweg, MA, Centre for Prevention and Intervention in Childhood and Adolescence, Faculty of Educational Science, Bielefeld University, Universitaetsstrasse 25, 33615 Bielefeld, Germany; email: [torsten.bollweg@uni-bielefeld.de](mailto:torsten.bollweg@uni-bielefeld.de).

Grant: T.M.B., O.O., J.B., A.M.F., and O.M.D. received grants 01EL1424A and 01EL1424D from Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research) within the frame of the Health Literacy in Childhood and Adolescence research consortium.

Disclosure: The authors have no relevant financial relationships to disclose.

Acknowledgment: We thank Frank Faulbaum, PhD, Dawid Bekalarczyk, Dipl-Soz-Wiss, and Marc Danullis, Dipl-Soz, at the Social Sciences Survey Centre (SUZ, Duisburg, Germany) for consultation during the instrument development process and for coordinating the process of data collection and entry. We also thank Kristine Sørensen, PhD, Global Health Literacy Academy (Denmark); Luis Saboga Nunes, PhD, Universidade NOVA de Lisboa (Portugal); and Christiane Firnges, MPH, Frauenhauskoordination e.V. (The Association of Women’s Shelters), Berlin (Germany), for providing feedback in the instrument development process. Lastly, we thank the following graduate assistants at the Centre for Prevention and Intervention in Childhood and Adolescence, Bielefeld University for their support in the recruitment process, verification of data entry, and drafting of information materials: Katharina Kornblum, MA; Sandra Kirchhoff, MA; Sophie Langer, MA; Sandra Schlupp, MA; and Juri Kreuz, BA.

Received: February 7, 2019; Accepted: August 26, 2019

doi:10.3928/24748307-20200428-01

study (i.e., Germany), because only two of the identified measurement tools are available in German (Schmidt et al., 2010; Wallmann, Gierschner, & Froböse, 2012). This study was conducted to address this gap in research by developing and testing a measurement tool that assesses generic HL in fourth-grade elementary school students in Germany.

### **ADAPTING THE HLS-EU-Q**

We chose the European Health Literacy Survey Questionnaire (HLS-EU-Q) as the starting point for questionnaire development because its validity and reliability has been confirmed in a range of studies in different countries and different settings (Amoah, Phillips, Gyasi, Koduah, & Edusei, 2017; Duong et al., 2017; Nakayama et al., 2015; Pelikan & Ganahl, 2017; Sørensen et al., 2015; Toçi, Burazeri, Sørensen, Kamberi, & Brand, 2015). The HLS-EU-Q is a measure of self-reported general HL that assesses participants' perceived difficulty in accessing, understanding, appraising, and applying health information in the contexts of health care, disease prevention, and health promotion (Sørensen et al., 2012). As such, it permits a broad measurement of different aspects related to HL instead of just measuring specific HL-related skills in specific contexts only, such as reading comprehension in medical contexts (Sharif & Blank, 2010). However, it needs to be stressed that self-reported general HL is a subjective indicator and does not assess actual skills. Instead, the scope of this measurement covers the perceived ease or difficulty in dealing with health information. This can indicate perceived (individual and system-related) barriers and actual problems in accessing, understanding, or using health information, but it can also indicate in-depth knowledge and an awareness of the ambiguities and lack of evidence to be found in certain health topics.

The HLS-EU-Q has been used with both adults and adolescents but never with children (Pelikan & Ganahl, 2017); therefore, it seems particularly advantageous to adapt this measurement tool for a younger age group in order to obtain comparable and continuous assessments of HL across the lifespan.

## **METHOD**

### **Study Design**

We conducted a questionnaire development and validation study, including cognitive pretests, and a quantitative pilot study. This article focuses only on the quantitative study and explores the quality of the instrument regarding its psychometric properties. The questionnaire development process as well as the results of cognitive testing are reported elsewhere (Bollweg et al., 2020). The pilot study was carried

out as a cross-sectional survey in school classes using a written, self-administered questionnaire.

### **Sample and Data Collection**

The target group consisted of children in fourth-grade attending elementary school in North Rhine-Westphalia, Germany, and data were collected only from children. Fourth grade is usually the last year of elementary education in Germany, and the children are age 9 and 10 years. We chose this age group because at this age children can be expected to have the necessary language skills to participate in a written survey, and they also represent an under-researched group with regard to HL. Moreover, starting mostly in fifth grade, children are allocated to the different tracks of the German school system, each featuring specific curricula and areas of focus. Thus, one aim of this study was to perform a baseline measure of children's HL before differences in educational tracks affect the acquisition of HL.

Participants were recruited via schools in an area of about 100 km (roughly 62 miles) around the city of Duisburg, North Rhine-Westphalia, Germany. Data were collected between November 2016 and May 2017. Our cooperation partner, the Social Sciences Survey Center (SUZ, Duisburg, Germany), was responsible for data collection and data entry. Both trained staff and classroom teachers were present during the survey, and the children received standardized instructions. The study was conducted over the course of two school lessons (90 minutes in total).

### **Measures**

Self-reported HL as measured by an adapted HLS-EU-Q scale constitutes the key focus of this study. Additional indicators of functional HL and self-efficacy were assessed to determine the convergent and discriminant validity of the HL scale. The scales described here were embedded in a larger questionnaire containing a total of 116 items addressing the determinants and outcomes of HL.

### **Self-Reported Health Literacy**

We measured self-reported HL, operationalized as the perceived difficulty in accessing, understanding, appraising, and applying health-related information in the contexts of health care, disease prevention, and health promotion (Sørensen et al., 2012), with adapted HLS-EU-Q items (see HLS-EU Consortium, 2012, for the original items). We had developed 26 adapted items in a previous phase of this study, described in greater detail in Bollweg et al. (2020). The original HLS-EU-Q

item format and response categories were retained with slight changes in wording. Thus, each item is worded: “How easy or difficult is it for you to . . . ?” and rated on a 4-point scale with the points being 1 (*very difficult*), 2 (*fairly difficult*), 3 (*fairly easy*), and 4 (*very easy*). We added “don’t know” as an additional response category, because this option can also be recorded when using the original HLS-EU-Q in personal interviews, although this category is usually not offered explicitly (Sørensen et al., 2013).

### Functional Health Literacy

We assessed functional HL (FHL) in order to determine the convergent validity of HL by developing a 12-item, cloze procedure reading comprehension test based on the Test of FHL in Adults (TOFHLA; Parker, Baker, Williams, & Nurss, 1995). This test was based on an informational text on vaccination published specifically for children age 8 to 13 years (Zentrum für Kinder- und Jugendmedizin Heidelberg, 2011). In line with the TOFHLA, words in the text were replaced by blanks and participants had to fill in the blank by picking out one of four options (**Figure A**).

### Self-Efficacy

We assessed the discriminant validity of the HL scale by measuring self-efficacy with three items: “I can trust in my knowledge and abilities,” “I can find a solution for most problems,” and “If I make an effort, I will succeed” on 4-point scales ranging from 1 (*not true at all*) to 4 (*absolutely true*) (all translated from German). Items were selected to indicate different aspects of self-efficacy and were based on scales developed by Reinders, Mangold, and Varadi (2005) and Schwarzer and Jerusalem (1999).

### Demographic Variables

Data were collected from children only, which is why standard indicators of socioeconomic status (i.e., parents’ education level, income, and occupation) could not be recorded reliably. Thus, we assessed socioeconomic status with the latest version of the Family Affluence Scale (FAS-III) (Torsheim et al., 2016), a self-report screener developed for the Health Behaviour in School-Aged Children study. The FAS includes six items on material assets at home, such as “does your family own a car, van or truck?” (*no, yes, one, or yes, two or more*), or “how many bathrooms are in your home?” (*none, one, two, or more than two*). Further, we assessed migration background indirectly as the language spoken with both parents, aggregated as “only German with both parents,” “German and other language,” and “no German at all.” We asked for age (years of life) and sex (female/male) with single items.

### Administration Time

Participants were asked to write down the current time at several points in the questionnaire to allow an estimation of administration time.

### Statistical Analysis

We subjected all items to an item analysis (i.e., an evaluation of missing values, item difficulty, variance, and discrimination). Missing values occur when participants do not give a valid answer, such as when they skip a questionnaire item or select *don’t know* instead of selecting one of the valid response options that are provided (e.g., *very difficult* or *fully agree*). A low frequency of missing values is desirable, as this indicates that participants understand the question and are able to respond using the options that are provided. Although the literature provides no hard cut-offs for missing values, we took a value of more than 8% as an indicator of comprehension problems. Item difficulty refers to the percentage of participants choosing the correct response, whereby an item difficulty parameter of 100% indicates that each and every participant picked the correct response. However, when applying rating scales (such as for agreement), there is no correct response. Thus, item difficulty in this context refers to the percentage of participants that choose the maximum possible response option (e.g., *fully agree, very easy, all of the time*). Item difficulty parameters between 20% and 80% are recommended to identify items that are able to discriminate between people with differing levels of the respective trait (Schinka, Velicer, & Weiner, 2003, p. 431). Item variance, too, is concerned with how well an item can differentiate between different respondents. If item variance is zero, all respondents chose the same (not necessarily the maximum or minimum) response, and no differences in the trait under investigation can be observed. Although no hard cut-offs for item variance have been suggested, one recommendation is to select items with higher variance (Kelava & Moosbrugger, 2008). Item discrimination (corrected item-total correlation) relates to whether scores on an individual item correspond to scores on an overall scale. For instance, participants that score highly on the overall HL scale should also score highly on all individual HL items. If this is the case, the individual HL items were able to discriminate between participants with a high and low overall level of HL. For item discrimination, a value of at least .300 is suggested (Streiner, Norman, & Cairney, 2015, p. 84).

We investigated the factor structure of the HL instrument (i.e. the underlying common themes) and different contents (i.e., factors) reflected by its items with exploratory factor analysis (EFA). We extracted factors with principal axis factoring to uncover the underlying latent constructs. We also chose

oblique rotation (oblimin) to allow correlations between factors. Rotation is a standard procedure used to make differences between factors more prominent and to highlight which items belong to which factor. We used Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) test of sampling adequacy to verify that the data were adequate for conducting EFA. We chose a  $p$  value of less than .05 and a KMO coefficient of more than .50 to indicate meaningful relationships between the variables of interest (Kline, 1994). We suppressed factor loading coefficients of less than .32 to allow for an easier interpretation of factors (Costello & Osborne, 2005).

We examined internal consistency (i.e. the degree of similarity between the HL items) with Cronbach's alpha and the Spearman-Brown split-half reliability coefficient. Values of .700 or higher indicate sufficient internal consistency (Streiner, 2003), indicating that the individual items measure the same construct. We tested the factor structure implied by the underlying theoretical model with confirmatory factor analysis (CFA), and we evaluated the model fit with the following indices: the relative chi-squared degree of freedom ( $\chi^2/df$ ), the normed-fit index (NFI), the relative fit index (RFI), the incremental fit index (IFI), the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA). In essence, CFA tests whether our assumptions about which items belong to which factors are probable or not. Thresholds for acceptable model fit were 5 (upper bound) for  $\chi^2/df$ , .07 (upper bound) for RMSEA, and .900 to .950 (lower bound) for NFI, RFI, IFI, CFI, and TLI (Hooper, Coughlan, & Mullen, 2008). CFA was carried out in SPSS Amos, version 25 (Arbuckle, 2017); all other analyses were carried out using SPSS Statistics, version 25. We could not calculate the standardized root-mean-square residual (SRMR) in SPSS Amos due to missing data. For correlations between metric variables, we calculated Pearson's correlation coefficient (denoted as  $r$ ); and for correlations between ordinal variables, Spearman's rank correlation coefficient (denoted as  $\rho$ ). We removed no cases from the dataset, and we did not impute missing values because we used the frequency of missing values as an indicator of item comprehensibility.

### Ethics Approval, Consent for Participation, and Funding

This study was approved by the Bielefeld University Ethics Board (Reference No 2016-141-R), as well as the Data Protection Officer at Bielefeld University. Parents or legal guardians provided informed written consent for all participants. Participation was voluntary and participants were in-

formed that all information would be treated confidentially. No incentives were used.

This work was carried out within the Health Literacy in Childhood and Adolescence Consortium ([www.hlca-consortium.com](http://www.hlca-consortium.com)) and funded by the German Federal Ministry of Education and Research (Grant numbers 01EL1424A and 01EL1424D).

## RESULTS

### Sample

A total of 200 schools were invited to participate, of which 32 accepted (16%). Schools had between one and four fourth-grade classes, and the number of students in fourth grade ranged from 19 to 100 (mean = 47.13; standard deviation [ $SD$ ] = 16.73) per school. Of the 1,537 fourth-grade students attending the participating schools, parental consent was obtained for 907 students (59%). About one-half of the sample were girls (53.5%), and most respondents were age 9 (47.3%) or 10 (45.5%) years. Two-thirds of the sample (63.8%) reported speaking only German with their parents. The FAS mean score was 8.64 (range: 0-13;  $SD$ : 2.35). **Table 1** reports the sample characteristics.

### Item Analysis

**Table 2** presents an overview of the 26 items tested in this survey and the statistics from the item analysis.

**Missing values.** The frequency of missing values ranged from 3.1% (items 20 and 24) to 20.2% (item 18), with a mean of 7.68% ( $SD$  = 11.79%) per item. Missing values include "don't know" as well as nonresponse (i.e., item skipped). On average, 1.5% ( $SD$  = 6%) of responses to the 26-item HL scale were nonresponses, and 6.2% ( $SD$  = 10.19%) were "don't know." Nine items had a missing rate greater than 8%. These items are marked in **Table 2**. Notably, six of the items with more than 8% missing values assessed the dimension "appraising health information."

**Item difficulty.** Item difficulty parameters ranged between 58.7% (item 18) and 88.66% (item 24). Eleven items had an item difficulty parameter above 80%. No items were observed in the "difficult" answer spectrum (item difficulty parameter < 20%). Accordingly, 15 items had "medium" difficulty. However, in general, items tended toward the "easy" response spectrum, with an average item difficulty of 76.06%.

**Variance.** Standard deviations ranged from .597 (item 24) to .897 (item 18) (response range, 1-4), whereas the average  $SD$  for all items was .755. One item had an  $SD$  below .6, three items had an  $SD$  between .6 and .7, 14 items had an  $SD$  between .7 and .8, and 8 items had an  $SD$  above .8.

**Discrimination.** Item discrimination indices (corrected item–total correlations [ITC]) ranged between .255 and .654, with an average ITC of .446. Only one item showed an ITC below .3. Seven items showed an ITC between .3 and .4, 14 items showed an ITC between .4 and .5; and four items showed an ITC greater than .5.

### Exploratory Factor Analysis

We performed EFA to evaluate the latent factor structure of the adapted HLS-EU-Q items. Bartlett’s test of sphericity indicated significant correlations,  $\chi^2(325) = 2320.34, p < .001$ , and a KMO coefficient of .915 indicated high strength of relationships between variables. Six factors with eigenvalues greater than 1.0 were extracted that accounted for 48.18% of the variance. In the rotated model, 9 of 26 items did not show factor loadings greater than .32 on any factor, and two factors were defined by one item only. Because it was eventually impossible to interpret factors, we fixed the number of factors to three and four, respectively, in line with the underlying conceptual model encompassing three health domains and four action areas (Sørensen et al., 2012). However, both solutions were unsatisfactory because the hypothesized factors did not correspond with the respective items.

### Item Selection

In our first step, we excluded nine items with more than 8% missing values. These items are noted in **Table 2**. In our second step, we inspected item difficulty parameters and excluded one item (item 24) for being too easy (difficulty: 88.66%) and having the lowest variance of all items ( $SD = .597$ ). However, we retained other items that exceeded the recommended maximum item difficulty of .80 because they otherwise performed reasonably well. We re-examined ITC after item selection, and all items had an ITC above .300 (range: .314 to .555). In total, we excluded 10 items in the item analysis.

### Second Exploratory Factor Analysis

We re-examined the factor structure of the shortened adapted HL scale in a second EFA. Bartlett’s test of sphericity indicated significant correlations,  $\chi^2(120) = 1441.413$  and  $p < .001$ , and a KMO coefficient of .870 indicated a high strength of relationships between variables. Three factors were extracted, explaining 39.78% of variance. The factor structure computed without any constraints is displayed in **Table 3**. The three factors had eigenvalues of 4.09, 1.22, and 1.06, and explained 25.6%, 7.62%, and 6.6% of variance, respectively. It should be noted that the three extracted factors did not correspond to the three theoretical factors (health

**TABLE 1**  
**Sample Characteristics (N = 907)**

Characteristic	N (%)
Sex	
Female	483 (53.3)
Male	408 (45.0)
Missing	16 (1.8)
Age (years)	
8	6 (0.7)
9	429 (47.3)
10	413 (45.5)
11	44 (4.9)
12	5 (0.6)
Missing	10 (1.1)
Language spoken with parents	
Only German	579 (63.8)
German and other languages	226 (24.9)
No German at all	82 (9.0)
Missing	20 (2.2)
Family affluence scale	
0-5	86 (9.5)
6	81 (8.9)
7	93 (10.3)
8	140 (15.4)
9	144 (15.9)
10	142 (15.7)
11	103 (11.4)
12	70 (7.7)
13	27 (3.0)
Missing	21 (2.3)

care, disease prevention, and health promotion). Also, fixing the number of factors to four did not yield a factor structure resembling the four action areas of the underlying theoretical model (access, understand, appraise, and apply health information).

Whereas Factor 3 seemed to circle around the topics of medication and vaccination, Factor 1 was defined most strongly by items on healthy nutrition. However, the remaining items loading on Factor 1 pointed to a broader construct than nutrition alone by also addressing such topics as traffic education or communication with parents. Factor 2 is defined by only one item (item 3). Correlations were  $r = .232$  between Factors 1 and 2;  $r = .592$  between Factors 1 and 3; and  $r = .272$  between Factors 2 and 3. The low correlations with Factor 2 point to the multidimensionality of the scale.

TABLE 2

## Item Analysis

Question	"How easy or difficult is it for you to..."	Don't know (%)	Nonresponse (%)	Missing total (%)	Mean	Difficulty (%)	SD	Variance	Excluded in step	ITC
1	find out how to recover quickly when you have a cold?	5.5	0.9	6.4	2.95	65.06	.791	.626		.383
2	find information about a cold, sore throat, or coughing?	8.9	0.9	9.8 <sup>a</sup>	2.91	63.57	.806	.650	1	.445
3	find out what you can do so that you don't get too fat or too thin?	6.3	1.1	7.4	3.39	79.80	.810	.656		.377
4	find out what you can do to avoid getting a cold in the winter?	2.2	1.4	3.6	3.42	80.56	.707	.500		.419
5	find out how you can best relax?	4.1	1.5	5.6	3.39	79.75	.748	.559		.242
6	find out which food is healthy for you?	3.1	1.4	4.5	3.42	80.56	.707	.500		.419
7	understand when and how you should take your medicine when you are ill?	5.2	0.8	6.0	3.10	69.95	.849	.729		.447
8	understand what your doctor says to you?	2.4	1.2	3.6	3.28	75.93	.730	.533		.430
9	understand why you sometimes need to see the doctor even though you are not ill?	4.1	1.3	5.4	3.54	84.62	.738	.545		.397
10	understand why you need vaccinations?	3.3	1.0	4.3	3.46	82.03	.787	.620		.386
11	understand what your parents tell you about your health?	3.6	1.4	5.0	3.24	74.83	.678	.459		.474
12	understand why you need to relax sometimes?	3.7	1.2	4.9	3.43	81.15	.749	.562		.509
13	judge what helps or does not help to get rid of a cold?	10.0	1.4	11.4 <sup>b</sup>	2.91	63.60	.761	.579	1	.588
14	judge the truth of what the doctor tells you in order for you to get well again?	7.1	1.7	8.8 <sup>a</sup>	3.32	77.47	.727	.529	1	.431
15	judge whether you can trust the media when they warn you about risks to your health?	16.6	1.6	18.2 <sup>a</sup>	2.82	60.60	.832	.692	1	.388

**TABLE 2** (continued)  
**Item Analysis**

Question	"How easy or difficult is it for you to..."	Don't know (%)	Nonresponse (%)	Missing total (%)	Mean	Difficulty (%)	SD	Variance	Excluded in step	ITC
16	judge whether what may happen to you later if you start smoking is true?	9.6	1.8	11.4 <sup>a</sup>	3.57	85.53	.800	.639	1	.451
17	judge what helps a lot for you to stay healthy and what does not help much?	5.8	1.7	7.5	3.15	71.63	.723	.523		.646
18	judge how where you live (neighborhood, district, street) is connected to your health?	18.1	2.1	20.2 <sup>a</sup>	2.76	58.70	.897	.805	1	.438
19	judge how your behavior (exercise and diet) is connected to your health?	7.5	2.1	9.6 <sup>a</sup>	3.35	78.33	.756	.572	1	.541
20	do what your parents tell you to do so that you can get well again?	1.9	1.2	3.1	3.43	81.11	.654	.428		.385
21	take your medicine in the way you're told to?	2.9	1.3	4.2	3.40	80.02	.744	.554		.477
22	call an ambulance in an emergency?	6.7	1.6	8.3 <sup>a</sup>	3.33	77.56	.859	.738	1	.470
23	stick to what you have learned in road safety lessons?	6.1	1.6	7.7	3.48	82.52	.633	.400		.479
24	decide when you need to wash your hands?	1.1	2.0	3.1	3.66	88.66	.597	.357	2	.382
25	have a healthy diet?	2.2	2.1	4.3	5.41	80.34	.701	.491		.482
26	take part in deciding whether healthy food is served at school?	13.3	2.4	15.7 <sup>a</sup>	3.21	73.73	.812	.659	1	.381

Note. All items translated from German. <sup>a</sup>More than 8% missing. ITC = item-total correlation.



**TABLE 3**  
**Exploratory Factor Analysis**

Question	How easy or difficult is it for you to...	Factor		
		1	2	3
16	have a healthy diet?	.701	-	-
5	find out which food is healthy for you?	.613	-	-
15	stick to what you have learned in road safety lessons?	.433	-	-
10	understand what your parents tell you about your health?	.427	-	-
13	do what your parents tell you to do so that you can get well again?	.409	-	-
11	understand why you need to relax sometimes?	.399	-	-
12	judge what helps a lot for you to stay healthy and what does not help much?	.344	-	.309
7	understand what your doctor says to you?	.331	-	-
4	find out how you can best relax?	.287	-	-
2	find out what you can do so that you don't get too fat or too thin?	.264	-	-
3	find out what you can do to avoid getting a cold in winter?	-	.752	-
6	understand when and how you should take your medicine when you are ill?	-	-	.532
9	understand why you need vaccinations?	-	-	.459
8	understand why you sometimes need to see the doctor even though you are not ill?	-	-	.407
1	find out how to recover quickly when you have a cold?	-	-	.362
14	take your medicine in the way you're told to?	-	-	.338

Note. Coefficients smaller than .250 were suppressed for illustrative purposes. Items translated from German.

Therefore, we dropped item 3, and thus also Factor 2, from the scale. **Table 4** presents the final selection of items and also reports the original HLS-EU-Q items on which the adapted items were based.

### Confirmatory Factor Analysis

We retested the factor structure with CFA because we expected, in line with the HLS-EU-Q framework (Sørensen et al., 2012), that the items would assess the perceived difficulty of accessing, understanding, appraising, and applying health information (four factors) in the contexts of health care, disease prevention, and health promotion (three factors). Accordingly, we tested a four- and a three-factor model. However, we could not estimate the parameters of the four-factor model because it was underspecified due to the “appraise” factor being reflected by only one item. Thus, *post hoc* modifications had to be made, and the single “appraise” item was assigned to the most closely related factor, which was “understand.” **Table 5** reports the fit indices for all models.

We found no consistently satisfactory fit for either of the models. Although the fit indices  $\chi^2/df$  and RMSEA indicated an acceptable model fit for all three models, no acceptable

values were observed for NFI, RFI, or TLI. The best model fit indices were obtained for Model 2, which assumed that the 15 selected items reflect three different action areas (access, understand, apply health information). This model scored best in all model fit indices. It needs to be noted, however, that the values for NFI, RFI, and TLI were too low to indicate an adequate fit for this model. Moreover, the values for IFI and CFI indicated a sufficient model fit ( $\geq .9$ ) but fell short of indicating an excellent one ( $\geq .95$ ).

### Internal Consistency

An alpha coefficient of .791 indicated a satisfactory internal consistency for the 15-item HL scale. Additionally, the split-half reliability was high ( $r = .771$ ).

### HL Scores

We computed HL mean scores for participants who gave valid responses to at least 12 of the 15 items, which corresponds to a maximum missing rate of 20%. Accordingly, we calculated a mean score for 819 of the 907 participants (90.3%). We did not follow the usual practice in studies using the HLS-EU-Q of transforming scores. This made it easier to interpret mean scores because they related directly to the response format of

**TABLE 4**  
**Final Selection of Items for the HLS-Child-Q15**

Question	How easy or difficult is it for you to . . .	Adapted HLS-EU-Q item	Mean	SD	Missing (%)	ITC
1	find out how to recover quickly when you have a cold?	2	2.95	.791	6.4	.357
2	find out what you can do so that you don't get too fat or too thin?	20	3.39	.810	7.4	.341
3	find out how you can best relax?	33	3.39	.748	5.6	.319
4	find out which food is healthy for you?	32	3.42	.707	4.5	.407
5	understand when and how you should take your medicine when you are ill?	8	3.10	.849	6.0	.418
6	understand what your doctor says to you?	5	3.28	.730	3.6	.355
7	understand why you sometimes need to see the doctor even though you are not ill?	23	3.54	.738	5.4	.343
8	understand why you need vaccinations?	22	3.46	.787	4.3	.323
9	understand what your parents tell you about your health?	37	3.24	.678	5.0	.431
10	understand why you need to relax sometimes?	40	3.43	.749	4.9	.452
11	judge what helps a lot for you to stay healthy and what does not help much?	-	3.15	.723	7.5	.541
12	do what your parents tell you to do so that you can get well again?	13	3.43	.654	3.1	.379
13	take your medicine in the way you're told to?	14	3.40	.744	4.2	.430
14	stick to what you have learned in road safety lessons?	-	3.48	.633	7.7	.434
15	have a healthy diet?	-	3.41	.701	4.3	.466

Note. Items translated from German. HLS-Child-Q15 = Health Literacy Survey Questionnaire for Children.

**TABLE 5**  
**Confirmatory Factor Analysis**

Model	Items	Factors	$\chi^2/df$	NFI	RFI	IFI	TLI	CFI	RMSEA
1	1-15	1	2.933	.857	.810	.901	.866	.899	.046 (95% CI [.040, .053])
2	1-4 5-11 12-15	Access Understand Apply	2.789	.869	.819	.912	.867	.910	.044 (95% CI [.038, .051])
3	1, 5, 6, 12, 13 2, 7, 8, 11, 14 3, 4, 9, 10, 15	HC DP HP	2.882	.865	.812	.907	.869	.905	.046 (95% CI [.039, .052])

Note. CFI = comparative fit index; CI = confidence interval; df = degrees of freedom; DP = disease prevention; HC = health care; HP = health promotion; IFI = incremental fit index; NFI = normed-fit index; RFI = relative fit index; RMSEA = root mean square error of approximation; TLI = Tucker-Lewis index;  $\chi^2/df$  = relative chi-square.

1 (*very difficult*) to 4 (*very easy*). These mean scores revealed that participants used most of the response range (from 1.86 to 4.00). The resulting frequency distribution was negatively

skewed (skewness =  $-.639$ ; standard error of skewness =  $.084$ ), and a Kolmogorov-Smirnov one-sample test indicated that scores were not normally distributed ( $p < .001$ ). The mean

score for the whole sample was 3.34 ( $SD = .37$ ). The correlation between the 15-item scale and the original 26-item set was very high ( $r = .948, p < .01$ ).

### Convergent Validity

We examined the correlation between HL and FHL to determine the convergent validity. For the newly developed cloze procedure reading comprehension test, we obtained scores of 0 to 12. The distribution of FHL scores was negatively skewed (skewness = -1.52; standard error of skewness = .084). There was also a ceiling effect with 16.3% of participants achieving the maximum score. A correlation coefficient of  $r = .107$  ( $p < .001$ ) indicated a small but significant statistical relationship between HL and FHL. This supported the assumption that FHL and self-reported HL are related but distinct components of HL.

### Discriminant Validity

We assessed discriminant validity by inspecting the correlation between HL mean scores and self-efficacy scores. For the individual self-efficacy indicators items, we found moderate correlations ( $\rho = .280$  to  $.306, p < .001$ ). This finding does not support the assumption of discriminant validity, because we hypothesized that HL would be less strongly correlated with self-efficacy than with FHL.

### Administration Time

It took participants 10 minutes ( $SD = 4$  minutes) to fill out the first two sections of the questionnaire containing 9 items on “sources of health information” and the 26 described HL items. Hence, we assumed that the 15-item HL scale could be administered in under 10 minutes.

## DISCUSSION

In this study, a German-language 15-item scale based on HLS-EU-Q items, the HLS-Child-Q15, was developed and tested on fourth-grade students in North Rhine-Westphalia, Germany. It is the first study to explore the feasibility of adapting the HLS-EU-Q for children. The questionnaire developed in this study provides a novel tool that links up with previous research efforts and focuses on an age group that has not yet received much attention in this field of research. The HLS-Child-Q15 is multidimensional in that it captures different health-related topics such as “health-related communication,” “nutrition,” “health care,” and “medication adherence.” However, at the same time, it is, statistically speaking, internally consistent because it regards these topics as indicators of just one construct: chil-

dren’s perceived difficulty in dealing with health information. The questionnaire shows a high internal consistency, takes only a short time to administer, and allows an evaluation of the self-reported HL of children age 9 and 10 years. By providing a feasible tool, this study facilitates the building of an evidence base on children’s HL, which will enable the development of targeted interventions to improve HL and to mitigate disparities in HL at an early age. However, this study was also able to highlight critical aspects regarding the use of an adapted version of the HLS-EU-Q among children, which we hope will inform further advancements in the measurement of children’s HL. In the following text we will discuss our findings regarding item analysis, factor structure, and validity.

### Item Analysis

Item analysis showed that several items led to a rather high number of missing values—in one case up to 20.2%. In our understanding, this points to problems of understanding and helps identify those items that might be unsuitable for the target group. For instance, it seems that items such as “judge how where you live (neighborhood, district, street) is connected to your health?” or “judge how your behavior (exercise and diet) is connected to your health?” might be too complex and abstract. Also, we identified items that do not seem excessively complex or abstract, but still led to a high number of missing values. One example is “find information about a cold, sore throat, or coughing?” In this case, we think that children understand the item but cannot relate to it. In other words, finding information on this topic is not relevant to children’s everyday lives because it is most probably something that parents deal with.

Interestingly, all but one item from the area “appraise health information” were discarded in the item selection process due to poor performance. Hence, this area is not reflected adequately in the HLS-Child-Q15. Hence, future studies will need to investigate whether and how this action area can be covered by satisfactorily performing items. Our experience during the item development process indicated that developing items within the domain of appraising health information is particularly challenging because hardly any insights are published on the instances in which children engage in critically appraising health information. Accordingly, our findings highlight that further studies are needed to explore in what everyday situations children critically appraise what specific health information. Such studies will be particularly helpful to enable the development of tools aiming to assess and promote critical HL among children.

We also found that some items were too easy (item difficulty  $\geq 80\%$ ), which means that participants often used the high end of the response spectrum (*fairly easy* or *very easy*) instead of using the whole range of response options. This resulted in a negatively skewed distribution of HL mean scores, meaning that the HL scale does not differentiate well in the upper spectrum of HL mean scores. Thus, more difficult items need to be developed to assess the whole range of levels of HL. However, we think that the present HL scale sufficiently captures the variability in HL scores because we observed no significant ceiling effect, and most studies focus on the effects and emergence of limited HL (i.e., scores at the lower end of the spectrum). We found a rather high mean score of 3.34 in this study, which translates to a perceived ease of dealing with health-related information among participants. Although this score is higher than mean scores in other studies, the general trend is comparable to reports on using different versions of the original HLS-EU-Q in different samples (e.g., Berens, Vogt, Messer, Hurrelmann, & Schaeffer, 2016; Duong et al., 2017; Levin-Zamir, Baron-Epel, Cohen, & Elhayany, 2016; Sørensen et al., 2015; Sukys, Cesnaitiene, & Ossowsky, 2017).

### Factor Structure

In this study, we have examined whether the items of the questionnaire that we have developed reflect just one underlying construct (i.e., if the scale is unidimensional), or if they reflect multiple different constructs (i.e., if it is multidimensional). Cronbach's alpha and the split-half reliability coefficient both indicate that the 15-item HL scale has a high consistency, which is associated with unidimensionality. It has to be noted, however, that ITCs are mostly in the moderate range (.300 to .500). This might imply that although the scale is sufficiently consistent, it is not necessarily unidimensional. This is also indicated by EFA and CFA that both suggest at least two latent factors within the HL scale. Future studies should investigate how far these latent factors conform to the factors outlined in the theoretical model of HL, as described by Sørensen et al. (2012), or whether they suggest a different theoretical model. Although, to a certain extent, CFA supports the assumption that different action areas are addressed by the items (i.e., accessing, understanding, and applying health information), EFA indicates a different factor structure that is related to health topics (e.g., vaccination/medication and nutrition) rather than action areas.

In this study, all HL items were considered to be part of one scale, although three or four subscales are usually reported in studies using the original HLS-EU-Q. As we

have shown, the "usual" HLS-EU-Q factor structure could not be fully replicated in our adaptation of the questionnaire. This might be due partly to the low number of items used, which is why no subscales could be derived that were internally consistent on their own. However, it is also plausible that the individual items actually do address a very similar construct and not several distinct domains of HL. This, again, is supported by the finding that the model fit indices for the one-factor CFA model are similar to those for the three-factor model. In a similar vein, Paakkari, Torppa, Kannas, and Paakkari (2016) developed a questionnaire to assess HL among adolescents based on five domains of HL that, nonetheless, were summarized under one factor in the final questionnaire (Paakkari et al., 2016). In summary, more research is needed on the factor structure of the HLS-Child-Q15, and also on the HLS-EU-Q and HL in general. We hope that our findings on the factor structure of the HL scale will stimulate further research on the relationship between the different components of HL while also providing impulses for the development of a theory of HL in childhood.

### Validity

In this study, we assessed convergent and discriminant validity by comparing the HL scale with measures of FHL and self-efficacy. Although we hypothesized that HL would relate more strongly to FHL than to self-efficacy, we observed the opposite. This indicates that the operationalization of self-rated HL we applied relates more closely to self-efficacy than previously assumed. This might be due to the notion of "perceived difficulty" inherent in every item of the HL scale, which is why it seems reasonable to control for self-efficacy when applying this scale. We were unable to confirm convergent validity through comparisons with the applied measure of FHL. However, we do not attribute this to shortfalls in questionnaire development but rather to a content-wise and conceptual mismatch between our measure of generic self-reported HL and the chosen indicator. Whereas generic HL and FHL are often referred to in the same context (Nutbeam, 2000; Sørensen et al., 2012), they constitute different aspects of a broader model of HL that are not equivalent. In our case, it seems plausible in hindsight that the perceived difficulty in using health information (self-reported HL), and health-related reading comprehension skills (FHL as measured by the adapted TOFHLA used here) do not correlate highly. Accordingly, it seems fruitful to explore convergent validity with more similar tools. One promising approach might be to test the questionnaire presented here against other questionnaires developed specifically to assess generic HL among young

people (e.g. Brown, Teufel, & Birch, 2007; Paakkari et al., 2016; Schmidt et al., 2010; Teufel, Vrtis, & Felder-Puig, 2019; Yu, Yang, Wang, & Zhang, 2012). Further, the presented findings suggest that it might be worthwhile scrutinizing existing definitions and concepts of HL regarding a more pronounced separation of FHL and self-reported HL. Lastly, it will be important for future studies to examine to what extent self-report HL measures are able to predict health-related outcomes (i.e., if they have predictive validity) when compared to performance-based measures.

## STUDY LIMITATIONS

There are several limitations to this study. For instance, the present questionnaire is an age-adapted version of the HLS-EU-Q that was developed originally for use in adults. Although we have adapted the questionnaire specifically for the target group of fourth-grade students in Germany (for more details on the development process see Bollweg et al. (2020), it was developed neither from scratch to meet their specific needs and idiosyncrasies, nor with their active participation. Accordingly, the questionnaire might be appropriate but not necessarily optimal for use in this age range. This indicates a potential for future studies on the development of HL questionnaires not only for, but also with, children. Moreover, there was little margin for item exclusion because of the rather low number of HL items ( $n = 26$ ) included in the quantitative pilot study. This was because the HL scale was included in a broader questionnaire containing 116 items. Accordingly, there was limited space for including a broader item pool. However, more leeway with respect to item exclusion and scale development could have been achieved by including a larger number of items reflecting the different dimensions of self-reported generic HL.

Moreover, social desirability may have distorted participants' responses, particularly because the survey was conducted within the classroom in the presence of the class teacher. Future research should investigate whether and to what extent the HLS-Child-Q15 scale is prone to social desirability bias. Another limitation is the use of unvalidated measures. Specifically, we used indicators of FHL and self-efficacy that have not been validated in the target group. Although validated, sophisticated assessments of reading skills are available for students in Germany, none were appropriate for use in this study because they either require a long administration time or can be used only in a fee-based license system. Accordingly, a new indicator of functional literacy had to be derived to allow for a time- and cost-efficient assessment. Also, because no brief test of self-efficacy was avail-

able in German for 9- and 10-year-old children, we used a compilation of three indicator items. However, it would be preferable to use validated measures.

Moreover, we determined the reliability of the questionnaire with only a one-time assessment. However, it is also important to investigate test stability (retest reliability) by conducting at least one follow-up survey. We were unable to do this due to financial and time constraints. Particularly with respect to the development of HL over the life course, it would seem worthwhile to conduct several repeated assessments. Finally, we would like to point out that all reported findings are based on a German questionnaire tested on German-speaking children. English translations of items are presented here for illustrative purposes but have undergone neither a professional back-translation process nor an evaluation of psychometric properties, and both of these would be necessary before using the questionnaire in English.

## CONCLUSION

This is the first study using an age-adapted version of the HLS-EU-Q for children. Based on a statistical analysis of 26 items tested in a survey of 907 fourth-grade students, a well-performing 15-item scale, the HLS-Child-Q15, was developed. This scale has a high internal consistency and takes only a short time to screen self-reported generic HL in German-speaking children age 9 and 10 years. However, further studies need to explore the factor structure of the questionnaire and its convergent validity, and it also needs to be validated in different languages and settings. Finally, this study contributes significantly to the comparative assessment of HL across the life course by providing a promising measurement tool for children age 9 and 10 years.

## REFERENCES

- Amoah, P. A., Phillips, D. R., Gyasi, R. M., Koduah, A. O., & Edusei, J. (2017). Health literacy and self-perceived health status among street youth in Kumasi, Ghana. *Cogent Medicine*, 4(1), 271. <https://doi.org/10.1080/2331205X.2016.1275091>
- Arbuckle, J. L. (2017). *Amos*. Chicago, IL: IBM SPSS.
- Berens, E.-M., Vogt, D., Messer, M., Hurrelmann, K., & Schaeffer, D. (2016). Health literacy among different age groups in Germany: Results of a cross-sectional survey. *BMC Public Health*, 16(1), 1151. <https://doi.org/10.1186/s12889-016-3810-6> PMID:27829401
- Bollweg, T. M., & Okan, O. (2019). Measuring children's health literacy: Current approaches and challenge. In O. Okan, U. Bauer, D. Levin-Zamir, P. Pinheiro & K. Sørensen (Eds.), *International handbook of health literacy. Research, practice and policy across the lifespan* (pp. 83-97). Bristol, UK: Policy Press.
- Bollweg, T. M., Okan, O., Pinheiro P., Bröder, J., Bruland, D., Frejtian, A. M., . . . Bauer, U. (2020). Adapting the European Health Literacy Survey for fourth-grade students in Germany: Questionnaire development and qualitative pretest. *HLRP: Health Literacy Research and Practice*, 2020, 4(2), e119-e128 <https://doi.org/10.3928/24748307->

- Borzekowski, D. L. G. (2009). Considering children and health literacy: a theoretical approach. *Pediatrics*, *124*(Suppl. 3), S282-S288. <https://doi.org/10.1542/peds.2009-1162D> PMID:19861482
- Brown, S. L., Teufel, J. A., & Birch, D. A. (2007). Early adolescents perceptions of health and health literacy. *The Journal of School Health*, *77*(1), 7-15. <https://doi.org/10.1111/j.1746-1561.2007.00156.x> PMID:17212754
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1-9. <https://doi.org/10.7275/jyj1-4868>
- Duong, T. V., Aringazina, A., Baisunova, G., Nurjanah., Pham, T. V., Pham, K. M., . . . Chang, P. W. (2017). Measuring health literacy in Asia: Validation of the HLS-EU-Q47 survey tool in six Asian countries. *Journal of Epidemiology*, *27*(2), 80-86. <https://doi.org/10.1016/j.je.2016.09.005> PMID:28142016
- Fairbrother, H., Curtis, P., & Goyder, E. (2016). Making health information meaningful: Children's health literacy practices. *SSM - Population Health*, *2*, 476-484. <https://doi.org/10.1016/j.ssmph.2016.06.005> PMID:29349163
- Guo, S., Armstrong, R., Waters, E., Sathish, T., Alif, S. M., Browne, G. R., & Yu, X. (2018). Quality of health literacy instruments used in children and adolescents: A systematic review. *BMJ Open*, *8*(6), e020080. <https://doi.org/10.1136/bmjopen-2017-020080> PMID:29903787
- HLS-EU Consortium (2012). Comparative report on health literacy in eight EU member states (second extended and revised version. July 22th, 2014). Retrieved from [https://cdn1.sph.harvard.edu/wp-content/uploads/sites/135/2015/09/neu\\_rev\\_hls-eu-report\\_2015\\_05\\_13\\_lit.pdf](https://cdn1.sph.harvard.edu/wp-content/uploads/sites/135/2015/09/neu_rev_hls-eu-report_2015_05_13_lit.pdf).
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modeling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53-60.
- Kelava, A., & Moosbrugger, H. (2008). Deskriptivstatistische evaluation von items (itemanalyse) und testwertverteilungen. [Evaluating items and test score distributions in descriptive statistics] In H. Moosbrugger & A. Kelava (Eds.), *Springer-Lehrbuch. Testtheorie und Fragebogenkonstruktion* (pp. 73-98). Heidelberg, Germany: Springer. [https://doi.org/10.1007/978-3-540-71635-8\\_4](https://doi.org/10.1007/978-3-540-71635-8_4)
- Kickbusch, I., & Ratzan, S. C. (2001). Health literacy: Making a difference in the USA. *Journal of Health Communication*, *6*(2), 87-88. <https://doi.org/10.1080/108107301750254439> PMID:11405081
- Kline, P. (1994). *An easy guide to factor analysis*. London, England: Routledge.
- Levin-Zamir, D., Baron-Epel, O. B., Cohen, V., & Elhayany, A. (2016). The association of health literacy with health behavior, socioeconomic indicators, and self-assessed health from a national adult survey in Israel. *Journal of Health Communication*, *21*(Suppl. 2), s61-s68. <https://doi.org/10.1080/10810730.2016.1207115>
- McCormack, L., Haun, J., Sørensen, K., & Valerio, M. (2013). Recommendations for advancing health literacy measurement. *Journal of Health Communication*, *18*(Suppl. 1), 9-14. <https://doi.org/10.1080/10810730.2013.829892> PMID:24093340
- Nakayama, K., Osaka, W., Togari, T., Ishikawa, H., Yonekura, Y., Sekido, A., & Matsumoto, M. (2015). Comprehensive health literacy in Japan is lower than in Europe: A validated Japanese-language assessment of health literacy. *BMC Public Health*, *15*(1), 505. <https://doi.org/10.1186/s12889-015-1835-x> PMID:26001385
- Nutbeam, D. (2000). Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promotion International*, *15*(3), 259-267. <https://doi.org/10.1093/heapro/15.3.259>
- Okan, O., Lopes, E., Bollweg, T. M., Bröder, J., Messer, M., Bruland, D., Pinheiro, P. (2018). Generic health literacy measurement instruments for children and adolescents: a systematic review of the literature. *BMC Public Health*, *18*(1), 166. <https://doi.org/10.1186/s12889-018-5054-0> PMID:29357867
- Ormshaw, M. J., Paakkari, L., & Kannas, L. (2013). Measuring child and adolescent health literacy: a systematic review of literature. *Health Education*, *113*(5), 433-455. <https://doi.org/10.1108/HE-07-2012-0039>
- Paakkari, L., & Paakkari, O. (2012). Health literacy as a learning outcome in schools. *Health Education*, *112*(2), 133-152. <https://doi.org/10.1108/09654281211203411>
- Paakkari, O., Torppa, M., Kannas, L., & Paakkari, L. (2016). Subjective health literacy: development of a brief instrument for school-aged children. *Scandinavian Journal of Public Health*, *44*(8), 751-757. <https://doi.org/10.1177/1403494816669639> PMID:27655781
- Parker, R. M., Baker, D. W., Williams, M. V., & Nurss, J. R. (1995). The test of functional health literacy in adults: A new instrument for measuring patients' literacy skills. *Journal of General Internal Medicine*, *10*(10), 537-541. <https://doi.org/10.1007/BF02640361> PMID:8576769
- Pelikan, J. M., & Ganahl, K. (2017). Measuring health literacy in general populations: primary findings from the HLS-EU Consortium's health literacy assessment effort. *Studies in Health Technologies and Informatics*, *240*, 34-59. <https://doi.org/10.3233/978-1-61499-790-0-34>
- Public Health England (2015). *Local action on health inequalities. Improving health literacy to reduce health inequalities: Practice resource*. Retrieved from <https://www.gov.uk/government/publications/local-action-on-health-inequalities-improving-health-literacy>
- Ratzan, S., & Parker, R. M. (2000). Introduction. In C.R. Selden, M. Zorn, S.C. Ratzan, R.M. Parker (Eds.), *National Library of Medicine current bibliographies in medicine: Health literacy: January 1990 through October 1999*. (pp. v-vii). Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.
- Reinders, H., Mangold, T., & Varadi, E. (2005). *Freundschaftsbeziehungen in interethnischen Netzwerken. Skalendokumentation des Längsschnitts 2003-2004. Frient-Projektbericht Nr. 8* [Friendships in interethnic networks. Scale documentation of the 2003 - 2004 longitudinal wave. Frient project report no. 8]. Mannheim, Germany: Universität Mannheim.
- Schaeffer, D., Hurrelmann, K., Bauer, U., & Kolpatzik, K. (2018). *National action plan health literacy. Promoting health literacy in Germany*. KomPart.
- Schinka, J. A., Velicer, W. F., & Weiner, I. B. (2003). *Research methods in psychology. Handbook of psychology*. Hoboken, NJ: Wiley.
- Schmidt, C. O., Fahland, R. A., Franze, M., Splieth, C., Thyrian, J. R., Plachta-Danielzik, S., Kohlmann, T. (2010). Health-related behaviour, knowledge, attitudes, communication and social status in school children in Eastern Germany. *Health Education Research*, *25*(4), 542-551. <https://doi.org/10.1093/her/cyq011> PMID:20228152
- Schwarzer, R., & Jerusalem, M. (Eds.). (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen* [Scales assessing teacher and student characteristics:]. Berlin, Germany: R. Schwarzer.
- Sharif, I., & Blank, A. E. (2010). Relationship between child health literacy and body mass index in overweight children. *Patient Education and Counseling*, *79*(1), 43-48. <https://doi.org/10.1016/j.pec.2009.07.035> PMID:19716255
- Sørensen, K., Pelikan, J. M., Röthlin, F., Ganahl, K., Slonska, Z., Doyle, G., . . . Brand, H., & the HLS-EU Consortium. (2015). Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). *European Journal of Public Health*, *25*(6), 1053-

1058. <https://doi.org/10.1093/eurpub/ckv043> PMID:25843827
- Sørensen, K., van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., & Brand, H., & the (HLS-EU) Consortium Health Literacy Project European. (2012). Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health*, *12*(1), 80. <https://doi.org/10.1186/1471-2458-12-80> PMID:22276600
- Sørensen, K., van den Broucke, S., Pelikan, J. M., Fullam, J., Doyle, G., Slonska, Z., . . . Brand, H., & the HLS-EU Consortium. (2013). Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BMC Public Health*, *13*(1), 948. <https://doi.org/10.1186/1471-2458-13-948> PMID:24112855
- Streiner, D. L. (2003). Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*(1), 99-103. [https://doi.org/10.1207/S15327752JPA8001\\_18](https://doi.org/10.1207/S15327752JPA8001_18) PMID:12584072
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford, England: Oxford University Press. <https://doi.org/10.1093/med/9780199685219.001.0001>
- Sukys, S., Cesnaitiene, V. J., & Ossowsky, Z. M. (2017). Is health education at university associated with students' health literacy? Evidence from cross-sectional study applying HLS-EU-Q. *BioMed Research International*, *2017*, 8516843. <https://doi.org/10.1155/2017/8516843> PMID:29130048
- Teuffl, L., Vrtis, D., & Felder-Puig, R. (2019). QUIGK-K: Quiz zur Erhebung von Gesundheitskompetenz bei Kindern. *Prävention Und Gesundheitsförderung*, *3*, 282. <https://doi.org/10.1007/s11553-019-00749-w>
- Toçi, E., Burazeri, G., Sørensen, K., Kamberi, H., & Brand, H. (2015). Concurrent validation of two key health literacy instruments in a South Eastern European population. *European Journal of Public Health*, *25*(3), 482-486. <https://doi.org/10.1093/eurpub/cku190> PMID:25395401
- Torsheim, T., Cavallo, F., Levin, K. A., Schnohr, C., Mazur, J., Nilclausen, B., & Currie, C., & the FAS Development Study Group. (2016). Psychometric validation of the Revised Family Affluence Scale: A latent variable approach. *Child Indicators Research*, *9*(3), 771-784. <https://doi.org/10.1007/s12187-015-9339-x> PMID:27489572
- Wallmann, B., Gierschner, S., & Froböse, I. (2012). Gesundheitskompetenz. Was wissen unsere schüler über gesundheit? [Health literacy. What do our students know about health?]. *Prävention und Gesundheitsförderung*, *7*(1), 5-10. <https://doi.org/10.1007/s11553-011-0322-1>
- World Health Organization. (2017). Shanghai declaration on promoting health in the 2030 Agenda for Sustainable Development. *Health Promotion International*, *32*(1), 7-8. <https://doi.org/10.1093/heapro/daw103> PMID:28180270
- Yu, X., Yang, T., Wang, S., & Zhang, X. (2012). Study on student health literacy gained through health education in elementary and middle schools in China. *Health Education Journal*, *71*(4), 452-460. <https://doi.org/10.1177/0017896911430548>
- Zentrum für Kinder- und Jugendmedizin Heidelberg. (2011). Medizin-fuer-kids.de - Was ist eine impfung? [Medicine for children: What is a vaccination?]. Retrieved from <http://www.medizin-fuer-kids.de/kinderarztpraxis/kinderkrankheiten/impfung.htm#>

Gleich siehst du Sätze, in denen Wörter fehlen. Wo eine Lücke im Satz ist, siehst du auch immer vier Wörter. Wähle für jede Lücke das Wort aus, das am besten passt! Hierzu machst du einfach einen **Kreis um den Buchstaben vor dem richtigen Wort**.

Hier ist ein Beispiel:

In jede \_\_\_\_\_ kommt ein \_\_\_\_\_

a. Karte	a. Mensch
b. Sonne	<b>b. Wort</b>
<b>c. Lücke</b>	c. Tier
d. Kurve	d. Schiff

Alles klar? Los geht's!

Was ist eine Impfung?

Gegen viele \_\_\_\_\_ kann man sich impfen \_\_\_\_\_.

a. Unfälle	a. haben
b. Krankheiten	b. müssen
c. Wartezimmer	c. können
d. Katzen	d. lassen

Die Impfung \_\_\_\_\_ eine ganz \_\_\_\_\_ Erkrankung in deinem Körper.

a. riecht	a. leckere
b. isst	b. dunkle
c. verursacht	c. laute
d. stört	d. leichte

Meistens \_\_\_\_\_ du gar \_\_\_\_\_ davon, ...

a. tust	a. nichts
b. merkst	b. viel
c. fliegst	c. keines
d. schmeckst	d. alles

... weil der \_\_\_\_\_ mit diesen wenigen Keimen leicht fertig \_\_\_\_\_.

a. Lehrer	a. geht
b. Körper	b. kann
c. Zeh	c. wird
d. Finger	d. macht

Von nun an bist du eine bestimmte \_\_\_\_\_ vor dieser Krankheit \_\_\_\_\_.

a. Person	a. gegangen
b. Namen	b. geschützt
c. Zeit	c. fertig
d. Summe	d. gewaschen

\_\_\_\_\_ lange, \_\_\_\_\_ ist bei jeder Krankheit unterschiedlich.

a. Wo	a. das
b. Was	b. der
c. Wer	c. die
d. Wie	d. er

Figure A. Cloze procedure test that was used to assess functional health literacy. This test was developed based on the Test of Functional Health Literacy in Adults (Parker, Baker, Williams, & Nurss, 1995). A text on vaccination was used as the blueprint of this test that had been specifically developed for children age 8 to 13 years (Zentrum für Kinder- und Jugendmedizin Heidelberg, 2011).