ORIGINAL RESEARCH

Community health literacy outcome measurement practices: A scoping review of recent interventions

Luke Sawyers[1] | Claire Anderson[1] | Parisa Aslani[2] | Gregory Duncan[3] |
Sobia S. Janjua[1] | Li Shean Toh[1]

[1]Division of Pharmacy Practice and Policy, School of Pharmacy, University of Nottingham, Nottingham, UK

[2]The University of Sydney School of Pharmacy, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[3]Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

Correspondence
Luke Sawyers, Division of Pharmacy Practice and Policy, School of Pharmacy, University of Nottingham, Nottingham, UK.
Email: msxls17@nottingham.ac.uk

Funding information
PhD Studentship (University of Nottingham), Grant/Award Number: N/A

Abstract

Introduction: Evidence suggests that, while a preference for functional Health Literacy (HL) outcome measurement exists, researchers are converging towards more all-encompassing instruments. While this claim is present in the HL field, minimal research has comprehensively explored the state of community HL measurement practices at the direct and proxy level. The almost exclusive focus on direct, as opposed to proxy, community HL measurement indicates a review of progress is needed.

Objective: To identify HL outcome measurement practices for community HL interventions at the direct and proxy level of measurement.

Search Strategy: Medline, PsycINFO, Web of Science, ERIC, Embase, Scopus, CINAHL, ProQuest Dissertations and Theses, Google Scholar and targeted websites were searched.

Inclusion Criteria: Studies were sampled from the general population, included HL as an outcome of interest, involved an intervention aiming to improve HL, were English-text publications and were published ≥2010.

Data Extraction and Synthesis: Study author(s) and publication years, sample characteristics, intervention profiles and direct and proxy instrument and outcome measurement information were extracted. Full-text review retrieved 25 eligible studies.

Main Results: In total, 21 unique direct and 38 unique proxy instruments were extracted. The majority of interventions assessed functional compared to communicative, critical, and other HL domains, with objective instruments more frequently used than subjective or combined objective-subjective types, though more unique subjective HL instruments were extracted overall. The Test of Functional HL in Adults was the most popular instrument, and perceived health, knowledge, behaviors and health intentions were the most frequent proxy outcome measures, with only the Healthy Lifestyle Behavior Scale-II and Patient Activation Measure used across multiple interventions.

Discussion and Conclusions: Direct HL outcome practices endured a unidimensional profile, despite previous suggestions of a convergence towards holistic instruments. This review provides the first overview of proxy HL measurement across community HL interventions, identifying substantial variation in proxy outcome practices.

Patient or Public Contribution: A University-based senior librarian contributed to the development of the search strategy, and reviewed iterations of the strategy until refinement was complete. No further public or patient contribution was made given the review-based nature of the research.

1 | INTRODUCTION

Health literacy (HL) refers to a person's knowledge, competence, and motivation to understand, access, apply and appraise health information to make effective decisions and judgements concerning health promotion, disease prevention and healthcare to promote or maintain quality of life.[1] Although originally considered a person's health-related reading and writing capabilities,[2] HL is now considered more holistically, with various domains comprising a person's HL. While many theories and models exist surrounding the domains encompassing HL, including the expanded model[3] and the structural model,[4] the most prevalent remains Nutbeam's (2000) model, which posits that three domains encompass HL: functional, comprising basic reading and writing skills for everyday functioning; communicative/interactive, concerning more advanced literacy, cognition and social skills to enable active participation, information extraction and derivation of meaning across different communicative modes and circumstances; and critical, considering the appraisal of information—forming the most advanced domain in the model.[5]

HL models have undergone a transition towards multi-dimensionality in recent decades, leading to an increase in HL measurement complexity. With HL now encompassing multiple skill dimensions, further consideration is needed during the measurement phase. However, despite the multifaceted, skills-based nature of HL as a construct,[6] outcome practices appear to align with unidimensional assumptions of HL,[2] with evidence indicating almost a third of HL instruments being based upon common functional literacy screening tools.[7] While not all HL measurements should incorporate the same outcome practices, this divergence may be cause for concern. For instance, McCaffery et al., who conducted an adult education intervention,[8] assessed HL through five dimensions of the Health Literacy Questionnaire (HLQ),[9] along with thermometer, food and medicine label interpretation tasks - with the former assessing communicative and critical HL and the latter measuring Functional HL (FHL). In contrast, Ayaz-Alkaya et al.[10] conducted a healthy lifestyle education programme and utilised two FHL outcome measures—the Adult Health Literacy Scale and Short-form Test of Functional Health Literacy in Adults (S-TOFHLA).[10] Though this does not suggest the varied measurement practice of HL alone is a problem, this contrasts previous suggestions of convergence towards all-encompassing HL measurement.[5]

If HL is not assessed efficiently at the domain or instrument-specific level, data generated through HL research may not represent the construct as intended, leading to misleading conclusions being drawn. The narrow unidimensional HL measures may be incorrectly labelled as holistic HL measurements, thus providing a narrow conceptual interpretation of public HL. For intervention research, poor HL outcome measurement could hinder the reliability and validity of interventions by reducing the evaluative certainty by which HL-promoting techniques are effective, potentially limiting progress in the field.

With estimates indicating 43% of working-age adults in England being unable to comprehend or utilise health-related information,[11] progression in the field is vital. Given the list of negative implications associated with low HL, including lower receipt of mammography screening and influenza vaccinations,[12] all-domain quality of life,[13] non-medication and medication non-adherence[14] and increased mortality risk,[15] enhancing practices in the field is vital for progression. Ensuring interventions appropriately evaluate HL strategies is an important next step, particularly given the suggested convergence towards more holistic, direct HL measurement in recent years.[5] The holistic measurement of HL is an important consideration at the direct level, with all model-based domains important to consider at the commencement of the study design phase. However, a truly holistic depiction of the construct requires consideration beyond model-specific domains alone.

Although direct HL measurements are important, the inclusion of proxy HL measurements may provide further insight into a person's overall HL. To distinguish direct and proxy HL measurement, we define direct HL as any domain(s) from existing HL models or models

applied to HL, and proxy HL as any domain(s) not directly attributable to known HL models or which may have the potential to be useful adjunctive outcomes to determine a person's HL. Research has not yet established clear outcome measurement practices at the proxy level, despite community HL interventions measuring a range of variables, including patient activation,[16] health behaviors,[17] health knowledge,[18] and self-perceived health status,[10] among others. While discussion surrounding direct HL is prevalent,[19] there is minimal discourse around proxy HL measures, and current knowledge surrounding proxy HL measures is minimal.

With community HL interventions being able to elicit behavioral and policy-level change, understanding current HL outcome measurement practices at the direct and proxy domain for community interventions is a necessary consideration, particularly to progress the field of HL overall. This is especially important for community populations, which we define as any nondisclosed disease/condition/illness or otherwise healthy member of the general population—a definition used in a recent community HL intervention review,[19] whose HL considerations are more generalizable than clinical populations due to the unique characteristics presented by clinical samples. Although a recent review has investigated community HL interventions,[19] the criteria for inclusion was restrictive, did not extend into the grey literature, and focused on the intervention methodology more than the outcome measurements implemented. The need for a measurement-oriented overview of recent community HL intervention practices is therefore needed.

Consequently, a scoping review was conducted to identify direct and proxy HL outcome measurement practices for recent community HL interventions. The identification of measurement trends at study and instrument-specific levels was the main focus throughout. Given that existing reviews are now dated[12] or had stringent inclusion criteria,[19] the current literature would benefit from further exploration to determine current measurement practices at the direct and proxy level within community HL intervention research.

2 | MATERIALS AND METHODS

2.1 | Protocol and registration

The scoping review protocol was drafted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses: Scoping Review Checklist and Explanation guidelines,[20] with the final protocol registered with the Open Science Framework on the August 11, 2020 (DOI: 10.17605/OSF. IO/4WRMQ).

2.2 | Inclusion/eligibility criteria

Studies were included if they:

1. Sampled from the general population.
2. Included HL as a significant outcome of interest.

3. Involved an HL intervention aiming to improve person-centered/environmental HL.
4. Used quantitative, qualitative, or mixed-methods approaches.
5. Were English-text publications.
6. Were published ≥2010.

Studies were excluded if they:

1. Sampled from a distinctive subgroup that are not typical members of society, like prison populations.
2. Involved participants with disclosed health conditions, illnesses, or ailments of interest.

Community participants were the primary recruitment focus. Patient populations may have unique characteristics to consider with regard to their HL, whereas community populations possess less distinctive profiles. Given the potential for a generalizable consensus to be attained in community populations, due to their less unique profiles, uncovering current measurement practices in HL interventions provides insight into the degree of existing consensus on HL measurement. In turn, the findings from this review can determine whether existing community HL intervention research would benefit from expert consensus on HL outcome measurement.

There are less factors to consider for community population HL compared with patient population HL, who may have unique characteristics to consider with regard to HL measurement, particularly given their increased exposure to the healthcare system and impairments which may impact HL assessments. Uncovering current measurement practices in HL interventions for community participants subsequently provides more generalizable findings for measurement consensus across community samples than patient/noncommunity samples.

The decision to conduct a scoping review was based on reach, enabling a broader scope of inquiry and promoting a wider overview of the literature.[21] As scoping reviews typically do not include quality/appraisal assessments,[22] this review was centered around identifying community HL intervention outcome practices and trends relative to the interventions and instruments retrieved.

2.3 | Search strategy and procedure

2.3.1 | Traditional literature database search

Medline, PsycINFO, Web of Science, Education Resources Information Center, Embase, Scopus and the Cumulative Index of Nursing and Allied Health Literature was searched. Forward searching was used to scan reference lists of identified articles meeting the inclusion criteria to broaden the scope for extraction.

The search strategy was devised, refined and evaluated by the research team with support from a University Senior Librarian. The following search was used for Medline and translated across the remaining databases (See Table 1):

TABLE 1 Search strategy for Medline

Search line	Input
1	exp Health Literacy/
2	exp "Surveys and Questionnaires"/
3	exp Health Education/
4	communit*.mp.
5	exp Community Participation/
6	general population.mp.
7	public.mp.
8	((health or medicine or medical or medicat* or pharmacotherap*) adj2 literac*).mp.
9	((educat* or behavio?r or ehealth or online or web or internet or complex or prevent* or environ*) adj2 (intervention* or survey* or questionnaire* or program* or curricul* or semina* or session* or workshop*)).mp.
10	1 or 8
11	2 and 3
12	4 or 5 or 6 or 7
13	9 or 11
14	10 and 12 and 13
15	Limit 14 to (English and last 10 years)

2.3.2 | Grey literature search

Grey literature was included to achieve maximum reach regarding existing community-focused HL intervention research. The search was conducted in accordance with recommendations for a systematic grey literature search, including four strategies to minimise the risk of potentially relevant omissions: (1) grey literature databases, (2) customized Google search engines, (3) targeted websites, and (4) consulting with contact experts.[23] The grey literature database search incorporated ProQuest Dissertations and Theses, using a translated version of the Medline database search strategy. Google Scholar was incorporated as the Google search engine of choice (See Table 2), as more general customized Google search engines did not extract relevant results.

The targeted websites consisted of the following:

1. The Centers for Disease Control and Prevention (CDC; http://www.cdc.gov).
2. The United States Department of Health and Human Services (DHHS; https://www.hhs.gov/).
3. United Nations Educational, Scientific and Cultural Organisation (https://en.unesco.org/).
4. Public Health England (https://www.gov.uk/government/organisations/public-health-england).
5. World Health Organisation (http://www.who.int).

TABLE 2 Search strategy for Google Scholar

Search line	Input
1	allintitle: "Health literacy" AND intervention
2	allintitle: "Health literacy" AND community OR public AND intervention

TABLE 3 Search strategy for targeted websites

Search line	Input
1	Health literacy AND intervention
2	Health literacy AND community AND intervention
3	Health literacy AND public AND intervention

6. Australian Government Department of Health (https://www.health.gov.au/).

Given that the CDC and DHHS are specifically recommended for public health grey literature intervention searching, they were included.[24] In reference to consulting with contact experts, consultations with two HL experts yielded the above series of websites as appropriate resources (See Table 3 for the search strategy).

2.3.3 | General search details

Filters used across all search strategies included publication range (2010–2020) and publication language (British/American-English). For Google Scholar and targeted websites, the first five pages of items were extracted per each search line, being retrieved irrespective of relevance. This was done across each search line. If <5 pages of items were identified by a search line, the items across the pages identified were retrieved. Page filters were modified to contain 10 references per page for Google Scholar and targeted websites where possible.

2.3.4 | Review process and data charting

One reviewer worked independently to screen eligible abstracts for full-text review. After the initial screening, two reviewers worked independently to screen full-text items for inclusion. Once reviewed, both members discussed the decisions made and verified the screening accuracy. A third reviewer was available to make final decisions on any items in the event that discrepancies arose between the first and second reviewers. The researchers resolved disagreements that could not be resolved by consultation with the third reviewer through discussion with the remaining members of the team.

The data charting form was developed by the principal investigator and refined with feedback from the team. Once the

prospective data charting form draft was ready, suggestions by Levac et al. were implemented.[25] The form was iteratively managed, and underwent a process of calibration via pilot testing across 5–10 papers from the sample to ensure data extraction was consistent and relevant (See Appendix 1). The scoping review was managed through EndNote.

2.4 | Synthesis of results and analysis

Measures were extracted and categorized into direct or proxy measures. Direct instruments were further categorized by instrument type, and were cataloged into the following types: objective, referring to performance-based measures; subjective, involving self-report data; or objective and subjective, taking inspiration from a recent HL measurement review.[26] Frequencies were computed for instruments and studies in accordance with the above categories, and additionally for the domains assessed at the direct and proxy level. Direct measures were categorized according to Nutbeam's (2000) HL model, denoting measures which assessed functional, communicative and critical HL. This is due to this model being the most abundant for guiding HL measurement, and was expected to form the majority of outcomes retrieved. To account for additional HL model domains, measures directly assessing HL in reference to domains outside of Nutbeam's model were categorized as 'other' direct HL domains,[5] and those unable to be categorized were labelled 'unidentified' due to insufficient or unavailable information to determine categorization.

Proxy measures were categorized via Braun and Clarke's six-step thematic analysis[27] to logistically manage the volume of measures extracted. Qualitative themes were generated by the principal investigator, which were then cross-checked by the team.

3 | RESULTS

3.1 | Screening and extraction

The scoping review yielded 270 items after the initial screening, with full-text review identifying 25 items for synthesis (see Figure 1).

3.2 | Data charting table

A plethora of study characteristics and outcome practices were retrieved, including author and publication details, listed under the study column; sample characteristics, sizes, and nationality, identified within the sample and sample size column; intervention delivery, type, and material covered, extracted in the intervention column; direct and proxy instrument names, provided under the namesake columns, and the instrument domains assessed at the direct and proxy level, listed under the direct and proxy outcomes column (see Table 4).

3.3 | Direct health literacy outcome and instrument frequencies

A total of 21 unique direct HL measures were extracted, including measures which assessed outcomes pertaining to Nutbeam's[19] HL model and measures that utilized other models (See Figure 2).

The Test of Functional HL in Adults (TOFHLA) was the most frequently extracted direct HL instrument, with short-form ($n = 3$[10,28,29]) and normal variants ($n = 5$[31–35]) identified. Four direct HL instruments were used multiple times, with the Newest Vital Sign (NVS; $n = 2$[17,36]), HL for Iranian Adults Questionnaire (HELIA; $n = 2$[37,38]) and European HL Short Scale (HLS-EU-Q16; $n = 2$[39,40]) identified.

3.3.1 | Categorized direct health literacy outcome and instrument frequencies

Frequencies were charted across six formats Figure 3.

Individual domains extracted
The majority of individual domain direct HL outcome instruments assessed FHL ($n = 5$: NVS; TOFHLA; Short Assessment HL-English; Interpreting Thermometer, Food and Medicine Label Tasks; Study-specific Functional Literacy Survey: Water-related Basic Knowledge). One instrument assessed communicative HL alone ($n = 1$; Study-specific Interactive Literacy – Sharing with Family: Newsletter Activities), and no instruments assessed critical HL in isolation.

At the individual level, FHL was assessed substantially more ($n = 18$[17,10,28–44]) than communicative ($n = 9$[37–41,43–46]) or critical HL ($n = 9$[37–41,43–46]) across the studies sampled.

Combined domains extracted
For combined direct HL domains, communicative and critical HL ($n = 3$: Communicative and Critical HL Scale; Five Dimension HLQ; Non-specific Communicative and Critical HL Questionnaire), were the most prevalent direct HL instruments from Nutbeam's (2000) model domains.

No instruments assessed functional and communicative HL or functional and critical HL in combination.

The most frequent combined domain measures across the studies were functional, communicative and critical HL evaluations ($n = 6$[37–41,44]). Alternative combined Nutbeam domain measures in the studies were communicative and critical HL measures ($n = 3$[43,45,46]). No investigations implemented measures assessing functional and communicative HL or functional and critical HL.

Direct assessment instrument type frequencies
The majority of direct HL instruments were subjective ($n = 9$: HELIA; Deductive Thematic Analysis of Lesson Transcripts; HLS-EU-Q16; Communicative and Critical HL Scale; Five Dimension HLQ; Study-specific Functional Literacy Survey: Water-Related Basic Knowledge; Functional, Interactive and Critical Literacy: Behavior and Attitude Survey; Lenartz's German HL Questionnaire [LGHLQ]; Rapid Assessment of HL Questionnaire [RAHL]), with objective instruments second in frequency
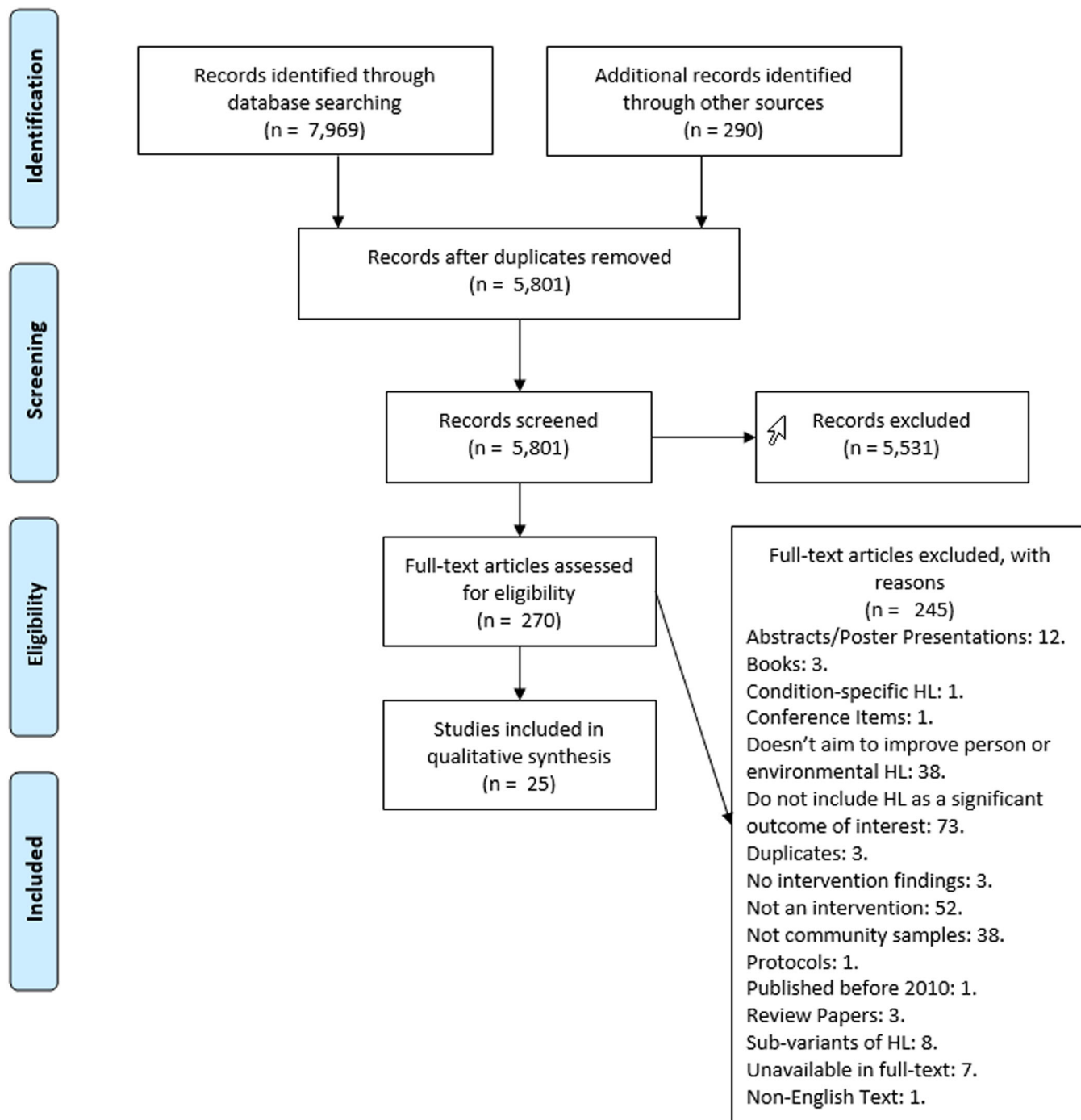
FIGURE 1 PRISMA flow chart for the data screening and extraction process.

(n = 6: NVS; TOFHLA, Study-specific Internet Searching Task; Study-specific Critical Appraisal Task; Short Assessment HL-English; and Interpreting Thermometer, Food and Medicine Label Tasks). A subset of instruments was objective and subjective (n= 2: 2008 CCHLQ; Study-specific Interactive Literacy – Sharing with Family: Newsletter Activities).

Objective instruments were the most commonly used in the studies (n = 13[16,17,10,28–36,42,43]), with subjective instruments less frequently used across the studies (n = 10[37–41,43–45,47,48]). Few objective and subjective instruments were used (n = 2[44,49]).

3.3.2 | Other and unidentified frequencies

In total, five direct instruments were categorized as 'other' due to the measures being based on models not utilizing Nutbeam's HL model (Study-specific Internet Searching Task; Study-specific Critical Appraisal Task; 2008 CCHLQ; LGHLQ; RAHL). In sum, four studies used "other" direct instruments as HL outcome measures.[16,47–49] Alternative models were: the Zarcadoolas et al. model,[3] utilised once[16]; the knowledge-attitude-practice model, used once[49]; the

TABLE 4   Data charting table of intervention characteristics and outcome practices

Study	Sample and sample size	Intervention	Direct instruments	Proxy instruments	Direct and proxy outcomes
Austvoll-Dahlgren et al.[16]	Sample: Norwegian parents with children <4 years old (baseline n = 96). Study aim: evaluate impact of web intervention vs. no intervention.	Delivery: online web portal. Type: education. Material: health info tasks and TPB-based self-report survey. Int. aim: improve generic public HL.	S-S internet health info searching task; S-S critical appraisal task (DISCERN tool); TPB survey.	Honeycomb model feedback; PAM.[a]	Direct: other HL. Proxy: satisfaction; patient activation.
Ayaz-Alkaya et al1.[10]	Sample: Turkish women from family health centers in Ankara (baseline n = 42). Study aim: identify whether healthy lifestyle education int. improved HL and healthy lifestyle behaviors.	Delivery: in-person. Type: education. Material: home healthy lifestyle-related visits from trained public health nurses. Int. aim: improve HL and health behaviors.	Adult health literacy scale; S-TOFHLA.	HLSBS-II/HPLP-II; PHS.	Direct: unknown,[b] FHL. Proxy: health behaviors/lifestyle; health status and self-perceived health.
Bayati et al.[10]	Sample: Iranian health ambassadors (baseline n = 256). Kazeroon residents on the Persian integrated health system. Study aim: determine effect of int. on HL and health behaviors in Kazeroon health network ambassadors.	Delivery: in-person/self-study. Type: education. Material: based on the Ministry of Health and Treatment's Self-care in Minor Morbidities book. Int. aim: improve HL/health behaviors of health network ambassadors.	HELIA/HL inventory for adults.	HLSBS-II/HPLP-II.	Direct: FCC HL. Proxy: health behaviors/lifestyle.
Bolton et al[28] (Two study designs).	Case-control study: Sample: baseline int. (n = 68) and ctrl. (n = 67); both pregnant women at 22 weeks gestation). Community evaluation study: Sample: baseline (n = 61) mothers from the PACT programme. Sampled around London, England. Overall study aim: provide a new service to benefit maternal mental health, HL and social capital, and engage local populations to reduce statutory services contact.	Delivery: in-person. Type: education. Material: health education (parent-focused healthcare topics) and social support (parents/children socialising and parent-led workshops) intervention. Int. aim: apply developmental health science findings, social support and health education to improve health outcomes of new mothers and their children.	Case-control study: NVS-UK.	Case-control study: adapted social capital integrated questionnaire; ASQ-3; ASQ:SE. Community evaluation study: Arizona social support interview schedule. Both studies: PHQ-9; GAD-7.	Direct: FHL. Proxy: case-control study: social capital; infant outcomes/physical development. Community evaluation study: social capital/support. Both studies: depression; anxiety.
Bruselius-Jensen et al.[29]	Sample: 5th (n = 4), 6th (n = 4) and 7th (n = 4) grade classes across 4 Danish public primary	Delivery: in-person. Type: education.	Deductive thematic analysis of lesson transcripts.	None.	Direct: FCC HL. Proxy: N/A.

(Continues)

TABLE 4 (Continued)

Study	Sample and sample size	Intervention	Direct instruments	Proxy instruments	Direct and proxy outcomes
	schools; pupils (n = 281) and teachers (n = 9). Baseline sample. Study aim: observe the IMOVE int.	Material: processes supporting being physically active, while improving application/use of statistics. Int. aim: improve physical activity-related HL and statistics skills.			
Carolyn et al.[30]	Sample: African-American adult churchgoers (baseline n = 321) from African Methodist Episcopal churches in Florida. Could understand/read at sixth-grade level. Study aim: examine impact of Health-Smart church int. on health outcomes of African-American adult churchgoers.	Delivery: in-person (church-based). Type: education and behavioral. Material: health empowerment and health promotion. Int. aim: empower people to overcome barriers to health-promoting behaviors.	NVS.	HLSBS-II/HPLP-II; health-smart behavior frequency scale; bodyweight scales; sphygomomanometer.	Direct: FHL. Proxy: health behaviors and lifestyle; health behaviors engagement; weight; blood pressure (Diastolic/Systolic).
Chervin et al.[31]	Sample: adults (n = 2412) in education centers (n = 6) predominantly serving minority groups. Teachers (n = 21) and center directors (n = 6) also took part. Sample represents baseline exposed to int. Mid-Atlantic America. Study aim: determine efficacy of infusing HL instruction in adult education on adult HL.	Delivery: in-person. Type: professional development grant to enable adult HL education curriculum interventions. Material: 'Study Circles', allowing adult education professionals to learn/discuss novel research for classroom use. Int. aim: increase HL in adult learners.	S-TOFHLA.	Self-efficacy assessment; skills learned essays; phone interviews; unspecified support received surveys.	Direct: FHL. Proxy: self-efficacy in using HL; HL project experience; capacity to teach HL for center directors and center teachers.
Cook[42]	Sample: African-American faith center adult parishioners (n = 14) around west Chicago. Baseline (no drop-out reported) Study aim: implement HIV stigma reduction/HL int. at a cross-generational African-American faith center.	Delivery: in-person (faith center). Type: education. Material: in-person adult HIV/AIDS and HL education. Int. aim: to shift stigma about HIV/AIDS and expand HL skills and HIV knowledge.	Short assessment HL-English.	Stereotypes about AIDS questionnaire; HIV knowledge questionnaire; focused observation questions.	Direct: FHL. Proxy: AIDS-related stereotypes; HIV knowledge; disease comprehension and reflection.
de Buhr et al.[32]	Sample: primary (n = 12), high (n = 13) and integrated primary-high schools (n = 3) schools in Hessen (n = 10) and	Delivery: in-person Type: education and behavioral. Material: nurses given vocational training/intensive education to	HLSAC; HLS-EU-Q16.	HLSAC.	Direct: other HL; FCC HL. Proxy: health behaviors.

TABLE 4 (Continued)

Study	Sample and sample size	Intervention	Direct instruments	Proxy instruments	Direct and proxy outcomes
	Brandenburg (n = 18), Germany. Baseline sample (children n = 2773; parents n = 3978; teachers n = 420). Study aim: examine observed changes in HL from a school int.	provide primary care/health education to schools. Int. aim: for nurses to address healthcare needs in schools and impact health-related outcomes.			
Fernández-Gutiérrez, Bas-Sarmiento & Poza-Méndez.[33]	Sample: immigrant adults in Spain (n = 93). Baseline sample. Study aim: evaluate a mobile health intervention for improving cognitive and social skills to improve access and use of health services in migrants.	Delivery: mobile phone-based. Type: education. Material: health education promoting phone health, awareness, education and navigation of Andalusian sociomedical system. Int. aim: to promote access and facilitate navigation of the sociomedical system and promote/maintain wellness.	HLS-EU-Q16; HLS-APP-Q14; S-S practical HL tests.	S-S satisfaction, usability, functionality and applicability for mobile device app survey.	Direct: FCC HL; unknown.[c] Proxy: app intervention evaluation.
Grebner.[34]	Sample: adults ≥18 years of age from central Illinois (n = 86) at baseline. Study aim: determine whether participant-matched learning styles improve HL education.	Delivery: in-person. Type: education. Material: tailored to learning styles. Material based on patient radiologic exam and Medicaid information scenario. Int. aim: determine whether information displayed with participant-matched learning styles improved HL.	S-TOFHLA.	VARK questionnaire.	Direct: FHL. Proxy: VARK learning.
Ishikawa et al.[35]	Sample: Japanese adult community members (baseline n = 67). Study aim: evaluate a HL community programme.	Delivery: in-person. Type: education. Material: Japan healthcare system/healthcare issues, active patient role development, communication and patient collaboration. Int. aim: improve knowledge of healthcare policy, systems and issues in Japan, patient roles/relationships with healthcare providers and interpersonal skills.	CCHLS.	Abbreviated five-item measure of patient trust in the medical profession; open question with thematic analysis.	Direct: Comm and Crit HL. Proxy: trust in the medical profession; programme learning.

(Continues)

TABLE 4 (Continued)

Study	Sample and sample size	Intervention	Direct instruments	Proxy instruments	Direct and proxy outcomes
Khaleghi et al.[36]	Sample: second year students aged 18–25 from Islamic Azad University, Shahr Rey Branch, Iran; (baseline n = 120). Study aim: evaluate HL-based training via social networking to improve health quality of life.	Delivery: in-person/social media. Type: education. Material: physical and psychological health education via buzz groups, networking and brainstorming. Int. aim: address physical and psychological health education needs.	TOFHLA.	SF-12.	Direct: FHL. Proxy: health-related quality of life.
Liu et al.[37]	Sample: Chinese nursing home residents ≥60 years (baseline n = 263). Study aim: explore the efficacy of teach-back for improving nursing home resident HL.	Delivery: in-person. Type: education. Material: teach-back following "66 Indicators for Chinese Citizens' HL" brochure. Int. aim: improve older adult HL.	2008 Chinese citizens HL questionnaire.	Teach-back assessment index interviews (int. only).	Direct: other HL. Proxy: intervention material knowledge.
McCaffery et al.[8]	Sample: basic education students from New South Wales (≥16 years of age), graded as level two learners via the Australian Core Skills Framework (baseline n = 308; int. = 167, ctrl. = 141). Study aim: evaluate HL adult education int. for low literacy and numeracy adults.	Delivery: in-person (TAFE colleges). Type: education. Material: health promotion, wellbeing and shared decision-making themes. Used real-world scenarios with reading, writing, speaking, listening and numeracy in the health context. Int. aim: improve adult HL.	Interpreting thermometer, food, and medicine label tasks; five dimension HLQ.	Health confidence S-S scale; PAM; S-S student satisfaction scale; 12-item curriculum measure; self-report diet and physical activity measure.	Direct: FHL; Comm and Crit HL. Proxy: confidence in health skills; patient activation; student intervention satisfaction; health knowledge; health behaviors.
Panahi et al.[38]	Sample: second/third year undergraduate students in Shahid Beheshti University of Medical Sciences dormitories (baseline n = 130). Study aim: develop an extended health belief model with HL elements to assess whether a smoking prevention int. is effective.	Delivery: instant messaging (Telegram). Type: education. Material: HL and Health Belief Model-based education; covered perceived benefits, barriers, harms, self-efficacy, perceived susceptibility, severity and benefits of physical activity for smoking. Int. aim: impact smoking prevention behaviors.	HELIA/HL inventory for adults.	Study-specific 46-item questionnaire; 8-item smoking knowledge and associated adverse events questionnaire.	Direct: FCC HL. Proxy: susceptibility, severity, barriers, benefits, self-efficacy, cues to action and smoking preventive behaviors; smoking knowledge/adverse effects.
Simonds et al.[39]	Sample: American fourth-grade children from an elementary school near the Crow	Delivery: in-person (elementary school and surrounding areas of the Crow reservation).	Functional literacy survey: water-related basic knowledge; interactive	Qualitative interviews.	Direct: FHL; Comm HL; FCC HL.

TABLE 4 (Continued)

Study	Sample and sample size	Intervention	Direct instruments	Proxy instruments	Direct and proxy outcomes
	reservation (baseline $n = 44$ across two or more int. program components; 9–13 years old). Study aim: feasibility evaluation of environmental HL int. for children.	Type: behavioral and education. Material: water-related environmental knowledge. Int. aim: enhance environmental functional, interactive and critical HL in children.	literacy – sharing with family: newsletter activities; functional, interactive, and critical literacy: behavior and attitude survey.		Proxy: intervention experience.
Soto Mas et al.[40]	Sample: Hispanic/Latino adults ≥18 years of age who were able to read/write in Spanish. America (Baseline $n = 49$). Study aim: explore feasibility of different community settings for improving adult HL via an English language programme.	Delivery: in-person; elementary school ($n = 19$), large chain hotel ($n = 16$) or community church ($n = 14$). Type: education. Material: HL and ESL education curriculum at sample sites. Dialog, role-play and interactive skill-development activities; encouraged to engage in healthy extracurricular activities. Int. aim: improve English language proficiency and HL in Spanish-speaking Hispanic adults.	TOFHLA.	None.	Direct: FHL. Proxy: N/A.
Soto Mas et al.[41]	Sample: US-Mexico border college students enrolled in High School Equivalency/Migrant Access Programme, able to read and write in Spanish and ≥21 years of age (baseline $n = 156$). Study aim: explore whether basic adult instruction improves HL in Spanish-speaking immigrants.	Delivery: in-person (local community college). Type: education. Material: connected life science content with health and disease; discussed disease-specific/ general health information. Int. aim: improve HL in Spanish-speaking immigrants.	TOFHLA.	None.	Direct: FHL. Proxy: N/A.
Soto Mas et al.[42]	Sample: adults ≥21 years from the local community (Texas area) with no intervention experience, low-to-intermediate English proficiency, able to read/ write/speak basic English and read/write Spanish (baseline $n = 181$).	Delivery: in-person (community colleges) Type: education. Material: HL/ESL education curriculum. Health behavior theory, HL research and practice, sociocultural literacy and communication theories.	TOFHLA.	None.	Direct: FHL. Proxy: N/A.

(Continues)

TABLE 4 (Continued)

Study	Sample and sample size	Intervention	Direct instruments	Proxy instruments	Direct and proxy outcomes
	Study aim: test feasibility of ESL instruction for improving HL in Spanish-speaking adults.	Guided by 'Health for Heart' programme. Int. aim: familiarize low-to-moderate English proficiency Spanish-speaking adults with literacy demands in health settings.			
Soto Mas et al.[40]	Sample: American Spanish-speaking Hispanic adults ≥21 years, able to read/write in Spanish, no prior relevant intervention experience, and had low-to-intermediate English proficiency (baseline n = 181). Study aim: explore HL curriculum on cardiovascular health behaviors in Spanish-speaking adults.	Delivery: in-person. Type: education. Material: HL and ESL education curriculum. Focused on personal skills, health-related vocabulary, and how lifestyle can impact chronic disease. Int. aim: improve English proficiency and develop HL and cardiovascular disease preventive knowledge/skills.	TOFHLA.	Spanish cardiovascular health questionnaire; CELSA.	Direct: FHL. Proxy: cardiovascular health behaviors; English proficiency.
Stassen et al.[43]	Sample: students aged 18–25 with project agreements between a German university and vocational schools (Baseline n = 495). Study aim: understand whether web-based int. in schools improves structural HL model competencies.	Delivery: online and/or in-person (vocational schools). Type: education. Material: general health information, clarifying misinformation, nutrition, check-ups, quizzes, quick recipes, and motivation topics, focusing on everyday vocational student working life. School health day for in-person segment. Int. aim: strengthen competencies regarding a healthy lifestyle.	Lenartz's German HL questionnaire.	None.	Direct: other HL. Proxy: N/A.
Tsai et al.[44]	Sample: Southeast Asian women who immigrated to Taiwan due to marriage to a Taiwanese man (baseline n = 223). Study aim: describe and evaluate a PBL HL int. for Southeast Asian immigrant women.	Delivery: in-person. Type: education. Material: PBL HL education, with structured problem group learning facilitated by faculty tutor/coach (experienced PBL health educator) with a	Non-specific communicative and critical HL questionnaire.	Non-specific health empowerment questionnaire; 10-item five-point scale navigation self-efficacy scale; non-specific healthcare utilisation questionnaire.	Direct: Comm and Crit HL. Proxy: health empowerment; navigation efficacy; healthcare utilization.

TABLE 4 (Continued)

Study	Sample and sample size	Intervention	Direct instruments	Proxy instruments	Direct and proxy outcomes
		co-coach (medical translation immigrant women). Int. aim: promote competencies regarding access, comprehension and use of health information, empowerment, and use of health services.			
Zhuang et al.[45]	Sample: community residents in Shenzhen, China ≥18 years of age, and lived in Shenzhen for a minimum of six months (baseline n = 6413). Study aim: explore SMS health education for improving HL.	Delivery: in-person/instant messaging-based. Type: education. Material: conventional health education via bulletin boards, posters and lectures. Int. aim: improve adult HL.	Rapid assessment of HL questionnaire.	None.	Direct: other HL. Proxy: N/A.

Note: Under the sub-heading *Delivery*, within the wider *Intervention* column, we reported the location of the intervention being conducted where possible. We reported this information when available to provide further context, with the inconsistent presentation being due to studies failing to report study location information. The sample values were also reported as baseline values under the *Sample* sub-heading of the wider *Sample and Sample Size* column (and inferred when not directly reported in the studies), as more studies reported baseline values than during/post-intervention sample values, and as such baseline values were reported to ensure consistency within the table presentation. The presentation order for direct and proxy outcomes reflects the order in which direct and proxy instruments are presented.

Abbreviations: ASQ-3, Ages and Stages Questionnaire; ASQ:SE, Ages and Stages Questionnaire: Social-Emotional; CCHLS, Communicative and Critical HL Scale; CELSA, Combined English Language Skills Assessment; Comm HL, Communicative Health Literacy; Crit HL, Critical Health Literacy; Ctrl, Control; ESL, English as a Second Language; FCC, Functional, Communicative and Critical; FHL, Functional Health Literacy; GAD-7, Generalized Anxiety Disorder Questionnaire; HELIA, Health Literacy for Iranian Adults; HL, Health Literacy; HLQ, Health Literacy Questionnaire; HLSAC, Health Literacy for School-Aged Children; HLS-APP-Q14, Health Literacy App Questionnaire; HLSBS-II, Healthy Lifestyle Behavior Scale-II; HLS-EU-Q16, European HL Short Scale; HPLP-II, Health-Promoting Lifestyle Profile-II; Int., Intervention; N/A, Not Applicable; NVS, Newest Vital Sign; PACT, Parents and Communities Together; PAM, Patient Activation Measure; PBL, Problem-Based Learning; PHQ-9, Patient Health Questionnaire; PHS, Perception of Health Scale; SF-12, Short Form-12 Questionnaire of Life Quality; SMS, Short Message Service; SS, Study-Specific; S-TOFHLA, Short-form Test of Functional Health Literacy in Adults; TOFHLA, Test of Functional Health Literacy in Adults; TPB, Theory of Planned Behavior; VARK, Visual, aural, read/write and kinesthetic.

[a]Austvoll-Dahlgren et al.[16] originally listed the PAM as a measure of civic and science literacy, however, we categorized instruments based on their intended use, which led to the categorization of the PAM as a proxy measure of patient activation rather than of science and civic literacy identified in the original article.

[b]The Adult Health Literacy Scale was described in minimal detail within Ayaz-Alkaya et al.,[10] and the original instrument development paper was used to support the categorization of the measure as a direct or proxy measurement. However, the original paper by Sezer and Kadıoğlu (2014) was non-English, and subsequently uninterpretable. The instrument was subsequently identified as an unknown direct health literacy measurement.

[c]We were unable to gain access to the Fernández-Gutiérrez[34] full-text paper, and were unable to categorize the outcomes assessed for the HLS-APP-Q14, which appears to be a study-specific intervention knowledge measurement. Given our lack of confidence in the categorization, we categorized the HLS-APP-Q14 as an unknown direct HL measurement.
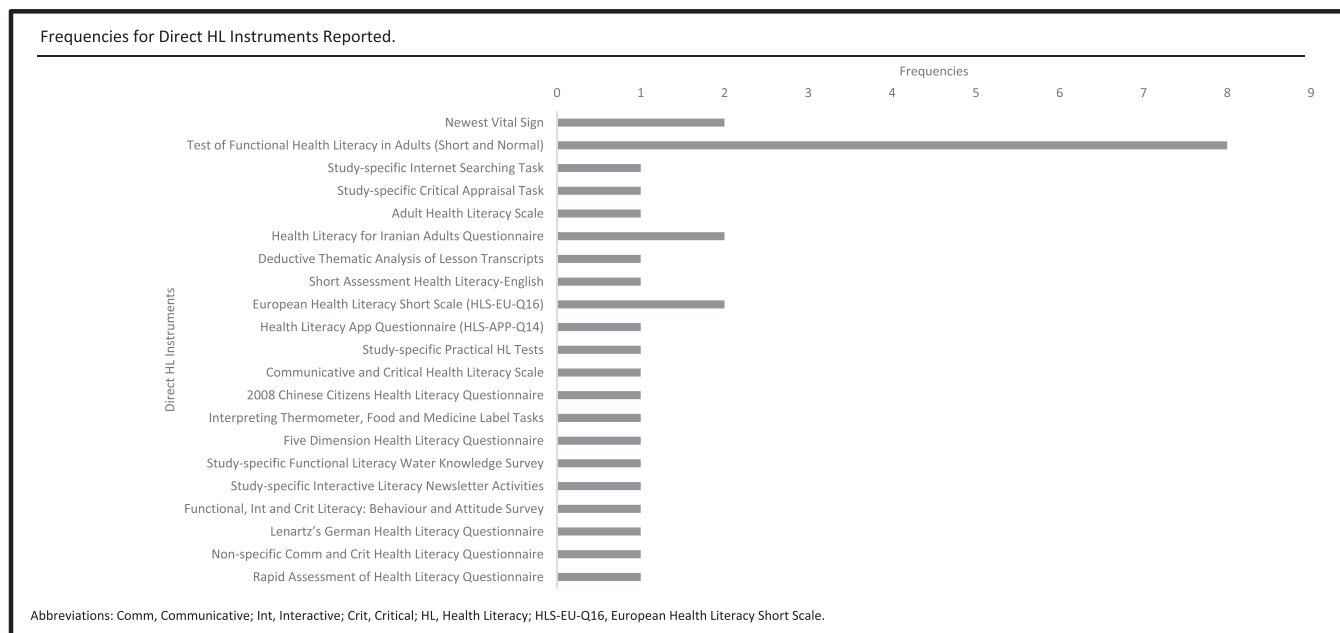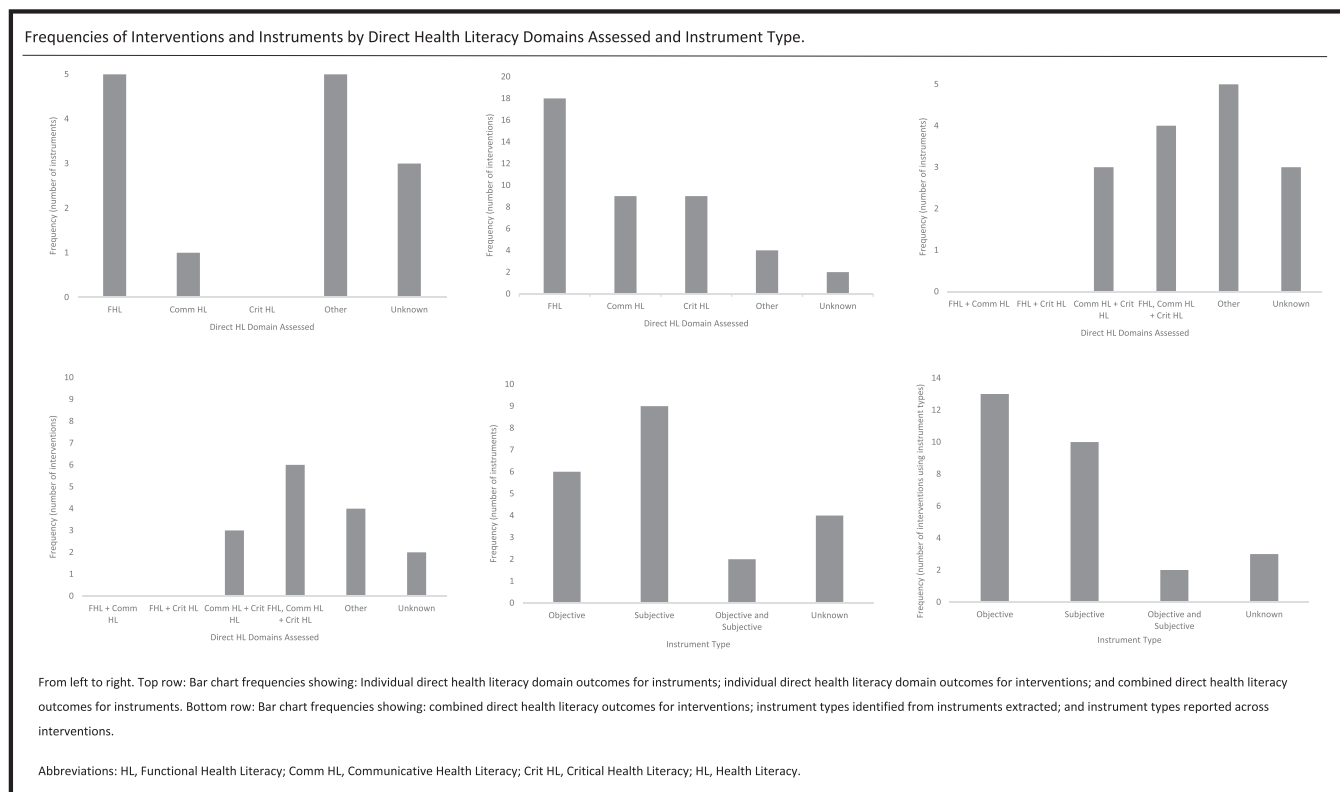
Frequencies for Direct HL Instruments Reported.

Abbreviations: Comm, Communicative; Int, Interactive; Crit, Critical; HL, Health Literacy; HLS-EU-Q16, European Health Literacy Short Scale.

FIGURE 2 Frequencies for direct HL instruments reported.

Frequencies of Interventions and Instruments by Direct Health Literacy Domains Assessed and Instrument Type.

From left to right. Top row: Bar chart frequencies showing: Individual direct health literacy domain outcomes for instruments; individual direct health literacy domain outcomes for interventions; and combined direct health literacy outcomes for instruments. Bottom row: Bar chart frequencies showing: combined direct health literacy outcomes for interventions; instrument types identified from instruments extracted; and instrument types reported across interventions.

Abbreviations: HL, Functional Health Literacy; Comm HL, Communicative Health Literacy; Crit HL, Critical Health Literacy; HL, Health Literacy.

FIGURE 3 Frequencies of interventions and instruments by direct health literacy domains assessed and instrument type.

structural model of HL,[4] implemented once[47]; and systems theory, used once.[48] For the 'other' direct HL measures, 11 domains comprised the factors assessed: Science literacy; knowledge; beliefs; behaviors; skills; self-perception; proactive approach to health;

dealing with health information; self-control; self-regulation; and communication and cooperation.

For instruments assessing direct HL with individual or combined outcomes (i.e., FHL alone or functional, communicative, and critical HL in

conjunction), a small proportion of instruments assessed unidentifiable outcomes (n = 3: AHLS; Health Literacy App Questionnaire [HLS-APP-Q14]; Study-specific Practical HL Tests), and a subset of studies used these unidentifiable outcome measures (n = 2[10,40]). Four instruments were categorized as unknown instrument types (n = 4: AHLS; HLS-APP-Q14; Study-specific Practical HL Tests; Non-specific Communicative and Critical HL Questionnaire) compared to those categorized as subjective, objective or objective, and subjective instrument types, and three studies used instruments with unidentifiable types (n = 3[10,40,46]).

3.4 | Proxy health literacy outcomes

3.4.1 | Proxy instrument frequency analysis

In summary, 38 unique proxy HL instruments were extracted, with thematic analysis retrieving the following measurement themes: anthropometric, developmental, and physiological characteristics; perceived health, knowledge, behaviors, and health intentions; perceptions of healthcare, usage and patient experiences; intervention experience-based evaluations; psychosocial, general and non-health factors; and miscellaneous (n = 1).

One proxy HL outcome tested the capacity to teach HL for education center directors, but this was not tested in the primary sample within which HL improvement was sought.[28] The measure was not an assessment of active participants in the intervention, and was listed separately—forming the miscellaneous group—and did not feature in further data representations in Figure 4.

The Healthy Lifestyle Behavior Scale-II (HLSBS-II)/Health-Promoting Lifestyle Profile-II (HPLP-II) was the most common proxy HL instrument

utilized across community HL interventions (n = 3[17,10,37]). The second most frequently implemented instrument was the PAM (n = 2[16,44]), and the remaining 36 instruments were utilized once.

3.4.2 | Proxy outcome categorization

Proxy measures were categorized into one of five measurement categories (see Figure 5).

3.5 | Combined direct and proxy measures

Two instruments measured direct and proxy HL simultaneously. The study-specific theory of planned behavior questionnaire (n = 1[16]) considered "other" HL components like civic and science literacy at the direct level, and beliefs about behavioral intentions at the proxy level. The HL for school-aged children questionnaire (n = 1[39]), which assessed "other" HL components (theoretical knowledge, practical knowledge, critical thinking, self-awareness, and citizenship), measured health behaviors at the proxy level. Both measures were subjective instrument types.

4 | DISCUSSION

4.1 | Key findings

Our investigation identified a myriad of direct and proxy HL outcome instruments, with 21 unique direct and 38 unique proxy measures retrieved. The studies frequently used functional measures (18/25) as
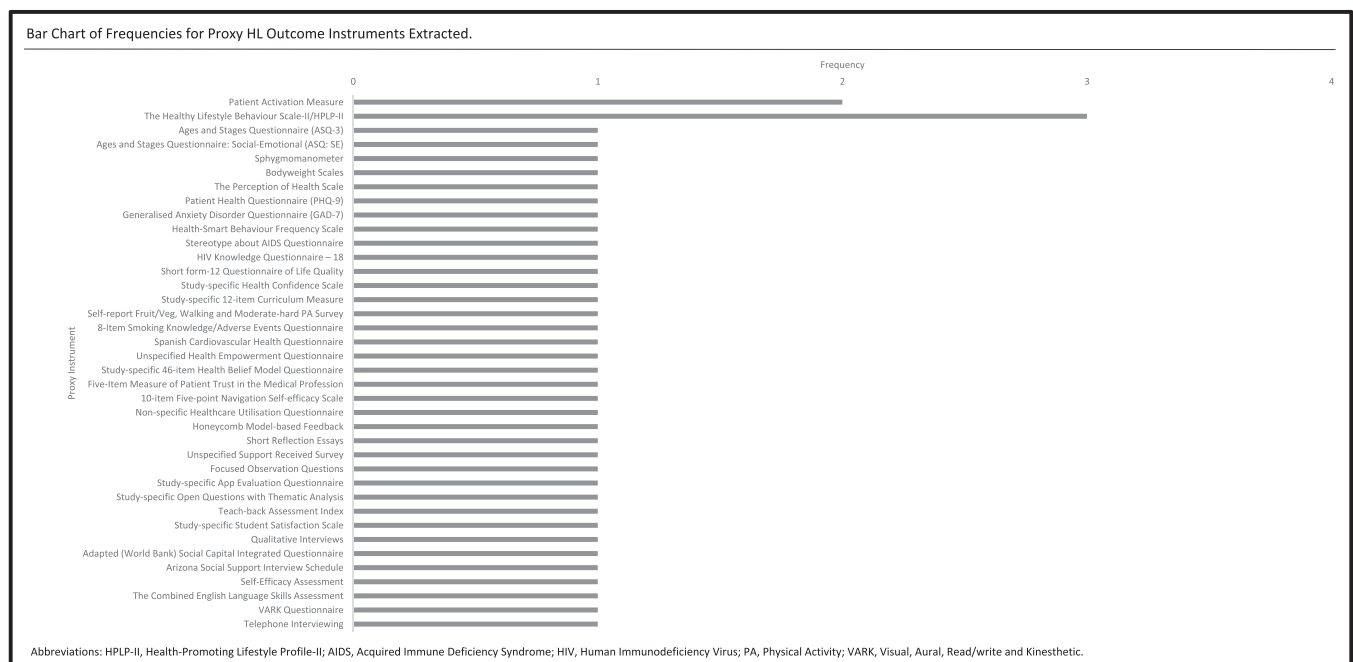

FIGURE 4 Bar Chart of frequencies for proxy HL outcome instruments extracted.
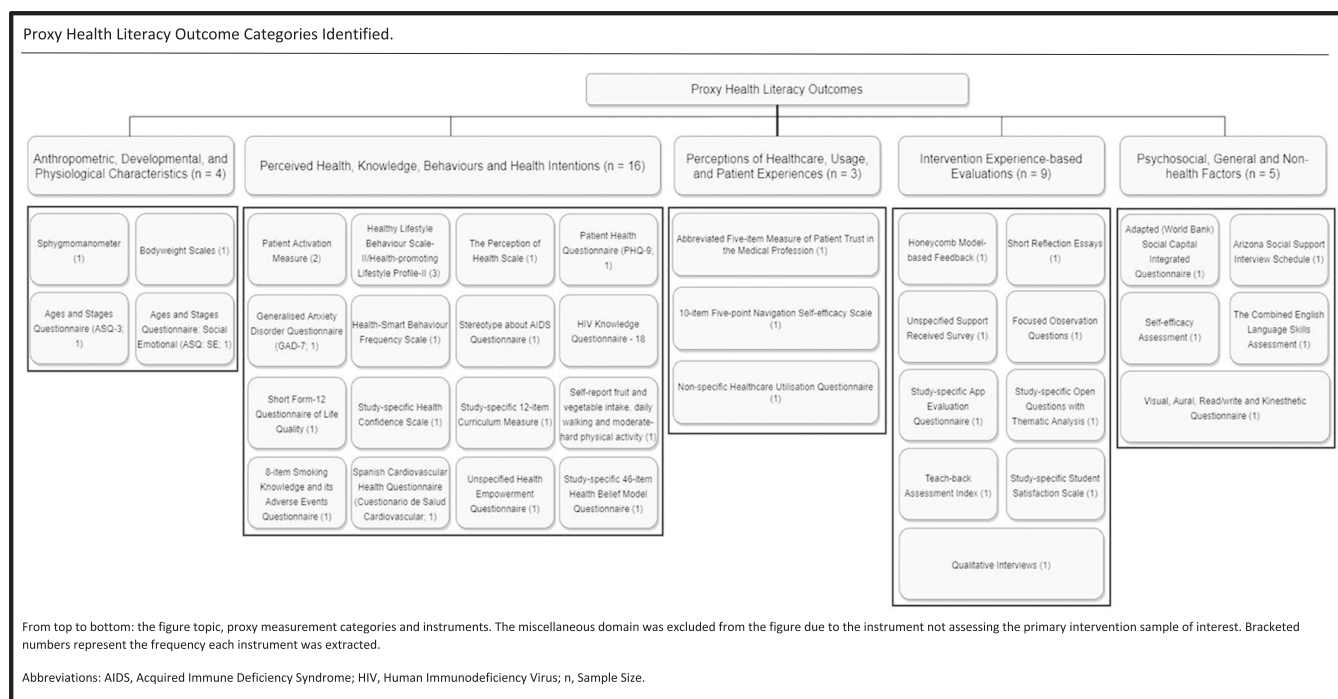
Proxy Health Literacy Outcome Categories Identified.

From top to bottom: the figure topic, proxy measurement categories and instruments. The miscellaneous domain was excluded from the figure due to the instrument not assessing the primary intervention sample of interest. Bracketed numbers represent the frequency each instrument was extracted.

Abbreviations: AIDS, Acquired Immune Deficiency Syndrome; HIV, Human Immunodeficiency Virus; n, Sample Size.

FIGURE 5 Proxy health literacy outcome categories identified.

opposed to communicative (9/25) and critical (9/25) measures. Not all studies incorporated measures which evaluated Nutbeam's model (4/25), with conceptual variance observed. The instruments retrieved followed a similar trend, with FHL instruments more prevalent (5/21) than communicative (1/21) and critical (0/21) instruments when considered as individual direct HL domains. For combined direct HL instruments, although no functional and communicative (0/21) or functional and critical (0/21) instruments were extracted, several communicative and critical (3/21) and functional, communicative and critical (4/21) instruments were present. Similar to the intervention frequencies, a variety of instruments measured domains not included in Nutbeam's model (5/21), but a unidimensional trend towards functional measurement became apparent, with the TOFHLA, a measurement of FHL, being the most frequently extracted direct HL instrument overall (8/21).

The gap in frequency was much greater between the TOFHLA and other popular direct measures, with the HELIA and HLS-EU-Q16 used two times, respectively. This frequency gap indicated no such convergence towards holistic HL measurement for direct HL, that is, measurement practices appear to be more unidimensional as opposed to being holistic, which is the opposite of a recent suggestion.[5] This is supported by evidence demonstrating that, from the 18 studies assessing FHL, only six measured communicative and critical HL as well, with three assessing communicative and critical HL alone. The direct practices of current community HL interventions appear to indicate that, while a degree of holistic instruments are available and are actively being used, the measurement preference remains one-dimensional.

At the proxy level, 38 unique instruments were extracted and categorized into: anthropometric, developmental and physiological characteristics ($n = 4$); perceived health, knowledge, behaviors and health intentions ($n = 16$); perceptions of healthcare, usage and patient experiences ($n = 3$); intervention experience-based evaluations ($n = 9$); and psychosocial, general and nonhealth factors ($n = 5$). The majority of proxy instruments were utilized individually across the studies, excluding the PAM and HLSBS-II/HPLP-II, which were utilized three and two times, respectively.

Across both direct and proxy HL measures, a significant degree of measurement variance was reported, demonstrating variation and potential disagreement towards outcome measurement at the domain, instrument, direct, and proxy level within present-day community HL intervention research.

4.2 | Functional domain frequency

The prioritization of FHL is expected, reflecting existing evidence in the field. Around one in every three HL instruments were previously suggested to be based on popular functional measures,[7] with a functional preference noted across adult[50] and child[51] populations. This preference appears to form standard practice, despite previous indications of a convergence towards all-encompassing HL measurement.[5] While this is not problematic if interventions intend to focus exclusively on FHL, this review indicates that, while a large portion of studies incorporated FHL-specific outcomes which reflect their proposed intervention aims,[29,32-35,43,44] others utilised interventions

with aims relating to the improvement of HL in general—implying an intended improvement across multiple HL levels—but only incorporated direct outcomes targeting FHL.[17,28,42] These practices indicate that, while the majority of studies successfully apply FHL outcomes in reference to their intervention aims, research may struggle to identify appropriate direct HL evaluations beyond the functional level, ignoring the assessment of potentially relevant HL skills.

Given the frequent use of the HL term in titles and key word sections of many interventions, these unidimensional practices infer that HL may be seen as a label for established interventions and outcomes[19] rather than as the multidimensional, skills-based construct that it is,[6] with communicative and critical components frequently overlooked. This may explain why investigations referring to HL as a general concept only assessed FHL, as they indirectly attribute HL as more of a descriptor rather than a construct. An individual's skill in one HL domain does not represent their overall HL, yet this approach infers the opposite is true, which may lead to inaccurate representations of public HL.

One problem with the label perspective is that it does not explain the preference for FHL outcomes observed, as only a subset of studies did not successfully match FHL outcomes with their intervention goals. This preference may be a consequence of the hierarchical nature of Nutbeam's model.[5] To build upon an individual's communicative or critical HL, they must first possess adequate FHL, and because of this interventions may attempt to focus exclusively on FHL to understand, identify and improve a person's foundational HL before they can explore higher levels of the construct. In this regard, FHL acts as an important catalyst to enable access to higher order HL skills in the public, and is a fundamental level for people to self-manage their health. The scope for public HL change may be most opportune at the functional level, and may contribute to the long-standing FHL prioritization at both the domain and instrument level.

Although a functional prioritization for direct HL outcomes is appropriate when the intervention aims are concerned with FHL, this review identifies a problem with HL outcome practices when considered with respect to community HL intervention aims. More specifically, some studies incorporated measurements that do not necessarily evaluate the proposed intervention HL-oriented goals,[17,28,42] indicating a misrepresentation with the construct as opposed to a preference for evaluating FHL. Although the majority of studies appropriately applied direct HL measurements relative to their intervention aims, a minority did not, and there may consequently be scope for an expert-led measurement framework to clarify HL assessment at the direct level to simplify the HL-related evaluation process.

Such a framework could also consider the instruments used to measure the outcomes. Currently, the functionally-focused TOFHLA was the most frequently implemented direct HL instrument, with more holistic instruments like the HELIA and HLS-EU-Q16 being incorporated substantially less. Given that popular functional instruments like the TOFHLA and NVS are relatively simple to administer or have short-form variants available,[52] functional measures may be more appealing to implement for already complex interventional research. Alternatively, this may be because the TOFHLA/NVS and associated variants most appropriately assessed the intervention goals. While this was the case for the majority of investigations sampled,[29,32–35] one study utilized the NVS which did not evaluate the intervention aims as effectively as other measures available.[17] In this instance, the All Aspects of Health Literacy Scale (AAHLS)[53] may have been more appropriate due to the wider scope for HL assessment—FHL, communicative HL, and critical HL—compared with the unidimensional NVS and the similar length of administration provided by the AAHLS. The AAHLS has been available for several years and indicates that, while FHL is an essential domain to consider, there may be instances where all-encompassing instruments like the AAHLS are more appropriate for assessing general HL-focused interventions. The development of an expert-led measurement framework may subsequently help alleviate confusion regarding the application of appropriate direct HL instruments for general HL-focused community interventions.

4.3 | Communicative and critical health literacy omission

While communicative and critical HL were measured more consistently than previous evidence indicates,[19] they were assessed substantially less than the functional domain. Prior suggestions for the disparity pointed towards the lack of comprehensive HL instruments available, with the AAHLS, the HLQ,[9] the Functional, Communicative and Critical Health Literacy Scale (FCCHLS)[54] and the European Health Literacy Scale (EU-HLS) being unavailable at the time of an earlier review.[19] While this may have been the case previously, the majority of sampled interventions in this review were able to access these measures, yet only three incorporated the HLQ[43] and EU-HLS.[39,40] Given that only three studies utilized communicative and critical HL measures[43,45,46] and si studies assessed functional, communicative, and critical domains,[37–41,44] the aforementioned instruments were likely seldom implemented for direct HL evaluations because the sampled interventions intended to target FHL more than communicative and/or critical HL. While there were studies which did not comprehensively investigate HL which may have benefitted from communicative and/or critical evaluations,[17,28,43] the majority of studies aiming to target communicative and/or critical HL domains successfully applied instruments evaluating these components. However, one study proposing an intervention to improve functional, communicative, and critical HL only included communicative and critical HL measures, omitting FHL.[46] When considering studies using interventions aiming to target functional, communicative, and critical HL, while we observed a greater number of studies which failed to incorporate communicative and critical direct HL outcomes where general HL improvement was sought, studies did not exclusively omit higher level HL domains and are capable to elicit narrow conceptual measurement practices beyond the functional level as well, though this was less prevalent.

Again, the preference for FHL measures may stem from their ease of implementation. However, the preference for functional-only measures cannot be explained in full by this. The AAHLS, a recently recommended instrument for assessing all domains of Nutbeam's (2000) model,[5,19] takes approximately 7 min to administer,[53] whereas the S-TOFHLA, the most commonly implemented instrument in this review – which assessed FHL alone, can take around 12 min to administer.[52] Given that 5 of the 8 studies used the normal TOFHLA, which can take approximately 22 min to administer,[55] the prioritisation towards FHL measures may not be a consequence of the ease of implementation which these assessments are traditionally associated with. A lack of awareness regarding holistic instruments like the AAHLS may contribute to the overuse of popular unidimensional measures like the TOFHLA and NVS. However, while this suggestion is neither supported nor disputed by this review, instrument type may explain the functional preference for measurement frequently observed in the literature.

4.4 | Instrument type and functional trend

The focus towards functional instruments like the TOFHLA may be a consequence of community HL intervention research prioritising instrument type in addition to the domain assessed. While less unique objective instruments were identified (n = 6) compared to subjective instruments (n = 9), the frequency by which interventions used objective instruments was greater than that of subjective instruments, with 13 interventions using objective HL instruments[16,17,10,28–36,42,43] versus the 10 interventions incorporating subjective instruments.[37–41,43–45,47,48] Though the difference between these frequencies was minimal, a preference emerged for objective instruments.

Instrument type can have a strong impact on the information retrieved, and should be applied in the appropriate context. Objective measures, for example, provide estimates for an individual's true HL skills, as their performance-based nature acts like a test, providing empirically grounded information. Subjective measures, on the other hand, are useful for determining population needs, including whether the current healthcare system is supporting the population appropriately, and are simpler to undertake for the participant due to their low cognitive demand.[56] However, both objective and subjective measures come with inherent limitations, with objective instruments using a test-based, potentially stigma-inducing approach and subjective instruments being unable to accurately determine whether a person's response denotes their true HL.[56]

While this review does not dispute that objective instruments, like the TOFHLA, may contribute to a valid public HL measurement, subjective instruments could act as more holistic evaluations of community HL interventions because of their wide-ranging profiles. This is particularly important to consider, as objective instruments tend to assess direct HL unidimensionally, with none of the 6 objective instruments extracted investigating functional, communicative and critical HL. Prioritizing FHL may therefore unintentionally

restrict the measurement scope. While providing an objective reality for community HL intervention evaluations is important, future research should be cautious when implementing objective instruments in isolation due to their unidimensional focus, as HL remains a multidimensional construct, and should only be used when the intervention focus considers FHL, and not HL generally.

4.5 | "Other" direct health literacy measures

HL outcome measurement stemmed beyond the Nutbeam levels, with five instruments across four interventions assessing alternative model/theory-driven domains. Different approaches to direct HL measurement included: the Zarcadoolas et al. model of HL[16,57]; an applied variation of the Knowledge-Attitude-Practice model[49]; the structural model of HL[4,58]; and an interpretation of systems theory.[48] Variation is an expected consequence of the myriad interpretations of HL as a construct, with evidence suggesting up to 250 different definitions of HL having co-existed in recent times.[59] Although unique attributes like science literacy contribute towards a well-rounded understanding of the public's HL, knowledge-based outcomes form ever-present requirements of the Nutbeam HL model domains. As a consequence, using knowledge as a direct HL outcome in isolation makes for a challenging interpretation of the domains in which HL change can occur. Using functional, communicative, critical, science literacy, or other nonoverlapping dimensions, on the other hand, subsequently adds conceptual clarity to the evaluation process when HL is concerned. HL interventions intend to improve HL by understanding mechanisms inducing HL skill changes, and knowledge instruments alone may not act as suitable evaluations of this.

4.6 | Proxy measurement inconsistency

At the proxy level, substantial variation at the instrument and domain level was observed, with 38 unique proxy HL instruments extracted and five broad outcome measurement categories identified. Although perceived health, knowledge, behaviors, and health intentions was the most frequently extracted proxy HL category, a range of outcome measures was retrieved. Though the proxy fluctuation measurement was expected given the wide conceptual scope for HL, the lack of consistency surrounding the frequency by which instruments were utilized for similar measurement domains suggests greater consensus may be achievable for proxy HL measurement.

Only two proxy instruments were implemented across multiple interventions: the HLSBS-II/HPLP-II and the PAM, reported three[17,10,37] and two times[16,43] respectively. While various proxy measurements were retrieved, only a fraction of domains utilized the same instruments. Moreover, many factors have demonstrated a shared association with HL, including physical activity,[60] dietary quality,[61] medical treatment adherence,[14] medical service usage,[62] and cognitive functioning.[63] However, although some of these variables were considered in a subset of interventions

sampled,[17,10,37,43,46] others were not, including adherence and cognitive functioning.

While we do not propose that all variables should utilize the same instruments and that all interventions should use the same proxy domains, there may be scope for a framework to better guide direct and proxy outcome measurement for community HL interventions. For example, one intervention assessed FHL alone and utilized an education intervention with 12 units of study—one being medication management—to improve HL. Although at the direct HL level the functional domain was considered, at the proxy level no further variables were assessed.[32] Incorporating an adherence or medication knowledge measure could have provided a proxy medication management indication of HL improvement at the functional level, suggesting more comprehensive proxy HL measurement may be plausible.

The low frequencies extracted from the 38 unique proxy instruments identified, with only two instruments implemented across multiple interventions, further support the notion of an expert-led framework to foster greater depth of HL-related measurement for community interventions. While direct HL measures have model-based frameworks providing a rough overview of the relevant HL dimensions, proxy HL measures have no such guiding mechanism. This review subsequently highlights that current outcome practices for community HL interventions may benefit from a framework guiding HL measurement at the direct and proxy level.

4.7 | Future research recommendations and research limitations

This review identified a unidimensional focus on HL outcome measurement and outlines the importance of defining HL as a multifaceted construct, recommending unidimensional direct HL measurement only where a unidimensional component is of interest, and not the construct as a whole. Determining the use of objective, subjective, or objective and subjective instruments for HL measurement is another important consideration, and should be decided by reflection upon the intervention design and intent, with objective tools providing true, unidimensional evaluations and subjective instruments providing broader self-reports of direct HL. Secondly, while the proxy level of HL remains unclear, future research should consider adopting proxy measures via guidance from existing association-based HL research while awaiting the development of an expert-led conceptual outcomes framework.

Although this review provides a fruitful, informative overview of HL measurement practices for community HL interventions on a broad scale, some limitations emerged. First, the findings do not evaluate the quality of HL interventions, primarily due to scoping reviews typically not including quality or appraisal elements.[22] Ensuring that interventions are being evaluated appropriately is arguably a greater consideration in the short-term, and understanding existing community HL intervention outcome practices will foster an efficient and knowledgeable appraisal moving forward. Additionally,

only English-text studies were included, leading to potentially relevant intervention omissions.

5 | CONCLUSION

This scoping review provides important evidence regarding the trends in HL outcome measurement across current community HL interventions. Despite previous suggestions of a potential convergence towards a more holistic HL measurement practice,[7] the functional prioritization of direct HL outcomes remains prevalent. Recommendations to consider HL as more than a functional skill[6,50] have done little to elicit meaningful change in outcome practices over the years. Although multidimensional direct HL instruments exist, such as the AAHLS,[53] which can be implemented faster than the most frequently extracted instrument from this review (TOFHLA/S-TOFHLA), interventions continue to implement unidimensional HL measures frequently. One explanation for the continued implementation of unidimensional FHL measures may be due to the preference for objectivity, with objective instruments more prevalent than their subjective counterparts. While this could explain the high TOFHLA/S-TOFHLA frequency, a combination of instrument types may yield more holistic direct HL measurement processes without the need for a sacrifice in objectivity.

For proxy HL, a preference for self-reported health, health behaviors, health knowledge, and intervention experience measurements emerged. The generation of 38 unique proxy HL instruments, of which only two were applied across multiple interventions, suggests that community HL interventions could benefit from in-depth literature consults to guide proxy HL association measurement inclusions presently. However, the lack of homogeneity surrounding community HL intervention measurement at the direct and proxy level indicates the potential for an expert-led outcomes framework to be developed. Such advancement may help alleviate confusion regarding the most appropriate dimensions to consider for community HL interventions. In providing a framework, a more consistent, all-encompassing, rigorous and reliable measurement practice may be in reach, promoting the standard for community HL intervention evaluations moving forward.

improve the research process. We would particularly like to thank our senior librarian at the University of Nottingham, who helped provide the framework for a comprehensive and improved search strategy aligned with the objectives of this review. This investigation was partially funded by the principal investigator's PhD Studentship (L.S.), provided by the University of Nottingham.

DATA AVAILABILITY STATEMENT

TRANSPARENCY STATEMENT

I confirm that this manuscript is an honest, accurate and transparent account of the findings reported from this review; that no important parts of the study have been omitted; and that any discrepancies from the review have been highlighted and discussed.

ORCID

Luke Sawyers http://orcid.org/0000-0002-8700-5335
Parisa Aslani http://orcid.org/0000-0002-7976-8236
Sobia S. Janjua https://orcid.org/0000-0002-4560-0208
Li Shean Toh https://orcid.org/0000-0002-9228-6266

REFERENCES

1. 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(1):1-13.
2. American Medical Association AHCotCoSA. Health literacy: report of the Council on Scientific Affairs. *JAMA*. 1999;281(6):552-557.
3. Zarcadoolas C, Pleasant A, Greer DS. Understanding health literacy: an expanded model. *Health Promot Int*. 2005;20(2):195-203.
4. Soellner R, Lenartz N, Rudinger G. Concept mapping as an approach for expert-guided model building: the example of health literacy. *Eval Program Plann*. 2017;60:245-253.
5. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int*. 2000;15(3):259-267.
6. Hibbard J. Patient activation and health literacy: what's the difference? How do each contribute to health outcomes. *Stud Health Technol Inform*. 2017;240:251-262.
7. Altin SV, Finke I, Kautz-Freimuth S, Stock S. The evolution of health literacy assessment tools: a systematic review. *BMC Public Health*. 2014;14(1):1-13.
8. McCaffery KJ, Morony S, Muscat DM, et al. Evaluation of an Australian health literacy program delivered in adult education settings. *HLRP: Health Lit Res Practice*. 2019;3(3):S42-S57.
9. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health*. 2013;13(1):1-17.
10. Ayaz-Alkaya S, Terzi H, Işık B, Sönmez E. A healthy lifestyle education programme for health literacy and health-promoting behaviours: a pre-implementation and post-implementation study. *Int J Nurs Pract*. 2020;26(2):e12793.
11. Rowlands G, Protheroe J, Winkley J, Richardson M, Seed PT, Rudd R. A mismatch between population health literacy and the complexity of health information: an observational study. *Br J Gen Pract*. 2015;65(635):e379-e386.

12. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med*. 2011;155(2):97-107.
13. Panagioti M, Skevington SM, Hann M, et al. Effect of health literacy on the quality of life of older patients with long-term conditions: a large cohort study in UK general practice. *Qual Life Res*. 2018;27(5):1257-1268.
14. Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. *Patient Educ Couns*. 2016;99(7):1079-1086.
15. Bostock S, Steptoe A. Association between low functional health literacy and mortality in older adults: longitudinal cohort study. *BMJ*. 2012;344:344-e1602.
16. Austvoll-Dahlgren A, Bjørndal A, Odgaard-Jensen J, Helseth S. Evaluation of a web portal for improving public access to evidence-based health information and health literacy skills: a pragmatic trial. *PLoS One*. 2012;7(5):e37715.
17. Tucker M, Kang S, Ukonu NA, et al. A culturally sensitive church-based health-smart intervention for increasing health literacy and health-promoting behaviors among black adult churchgoers. *J Health Care Poor Underserved*. 2019;30(1):80-101.
18. Cook L. HIV stigma reduction and health literacy education program with a cross-generational populaton in an African American faith-based church. *J Public Health (Bangkok)*. 2020:1-18.
19. Nutbeam D, McGill B, Premkumar P. Improving health literacy in community populations: a review of progress. *Health Promot Int*. 2018;33(5):901-911.
20. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467-473.
21. Peterson J, Pearce PF, Ferguson LA, Langford CA. Understanding scoping reviews: definition, purpose, and process. *J Am Assoc Nurse Pract*. 2017;29(1):12-16.
22. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Lib J*. 2009;26(2):91-108.
23. Godin K, Stapleton J, Kirkpatrick SI, Hanning RM, Leatherdale ST. Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. *Syst Rev*. 2015;4(1):1-10.
24. Turner AM, Liddy ED, Bradley J, Wheatley JA. Modeling public health interventions for improved access to the gray literature. *J Med Libr Assoc*. 2005;93(4):487-494.
25. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci*. 2010;5(1):1-9.
26. Okan O, Lopes E, Bollweg TM, et al. Generic health literacy measurement instruments for children and adolescents: a systematic review of the literature. *BMC Public Health*. 2018;18(1):1-19.
27. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101.
28. Chervin C, Clift J, Woods L, Krause E, Lee K. Health literacy in adult education: a natural partnership for health equity. *Health Promot Pract*. 2012;13(6):738-746.
29. Grebner LA. Learning style needs and effectiveness of adult health literacy education. *Dissertation Abstracts Int Sec A: Humanities Soc Sci*. 2015;75(8-A(E)).
30. Carolyn MT, Kang S, Ukonu NA, et al. A culturally sensitive Church-Based Health-Smart intervention for increasing health literacy and health-promoting behaviors among black adult Churchgoers. *Journal of Health Care for the Poor and Underserved*. 2019;30(1):80-101. doi:10.1353/hpu.2019.0009
31. Khaleghi M, Shokravi FA, Peyman N, Moridi M. Evaluating the effect of educational interventions on health literacy through social networking services to promote students' quality of life. *Korean J Fam Med*. 2019;40(3):188-193.

32. Soto Mas F, Cordova C, Murrietta A, Jacobson H, Ronquillo F, Helitzer D. A multisite community-based health literacy intervention for Spanish speakers. *J Community Health*. 2015;40(3): 431-438.

33. Soto Mas F, Jacobson HE, Olivárez A. Adult education and the health literacy of hispanic immigrants in the United States. *J Latinos Educ*. 2017;16(4):314-322.

34. Soto Mas F, Ji M, Fuentes BO, Tinajero J. The health literacy and ESL study: a community-based intervention for Spanish-speaking adults. *J Health Commun*. 2015;20(4):369-376.

35. Soto Mas F, Schmitt CL, Jacobson HE, Myers OB. A cardiovascular health intervention for spanish speakers: the health literacy and ESL curriculum. *J Commun Health*. 2018;43(4):717-724.

36. Bolton D, Khazaezadeh N, Carr E, et al. Evaluation of a community-led intervention in south London: how much standardization is possible? *Int J Environ Res Public Health*. 2020;17(7):2523.

37. Bayati T, Dehghan A, Bonyadi F, Bazrafkan L. Investigating the effect of education on health literacy and its relation to health-promoting behaviors in health center. *J Educ Health Promot*. 2018; 7:127-128.

38. Panahi R, Ramezankhani A, Tavousi M, Niknami S. Adding health literacy to the health belief model: effectiveness of an educational intervention on smoking preventive behaviors among university students. *Iran Red Crescent Med J*. 2018;20(2):e13773.

39. de Buhr E, Ewers M, Tannen A. Potentials of school nursing for strengthening the health literacy of children, parents and teachers. *Int J Environ Res Public Health*. 2020;17(7):2577.

40. Fernández-Gutiérrez M, Bas-Sarmiento P, Poza-Méndez M. Effect of an mhealth intervention to improve health literacy in immigrant populations: a quasi-experimental study. *Comput Inform Nurs*. 2019;37(3):142-150.

41. Bruselius-Jensen M, Bonde AH, Christensen JH. Promoting health literacy in the classroom. *Health Educ J*. 2017;76(2):156-168.

42. Cook L. HIV stigma reduction and health literacy education program with a cross-generational populaton in an African American faith-based church. *J Public Health (Germany)*. 2020;29(5):1089-1106.

43. McCaffery KJ, Morony S, Muscat DM, et al. Evaluation of an Australian health literacy program delivered in adult education settings. *Health Lit Res Pract*. 2019;3(3 Suppl):S42-S57.

44. Simonds VW, Kim FL, LaVeaux D, Pickett V, Milakovich J, Cummins J. Guardians of the living water: using a health literacy framework to evaluate a child as change agent intervention. *Health Educ Behav*. 2019;46(2):349-359.

45. Ishikawa H, Yamaguchi I, Nutbeam D, et al. Improving health literacy in a Japanese community population a pilot study to develop an educational programme. *Health Expect*. 2018;21(4):814-821.

46. Tsai T-I, Lee S-YD, Yu W-R. Impact of a problem-based learning (PBL) health literacy program on immigrant women's health literacy, health empowerment, navigation efficacy, and health care utilization. *J Health Commun*. 2018;23(4):340-349.

47. Stassen G, Grieben C, Sauzet O, Frobose I, Schaller A. Health literacy promotion among young adults: a web-based intervention in German vocational schools. *Health Educ Res*. 2020;35:87-98.

48. Zhuang R, Xiang Y, Han T, Yang G-A, Zhang Y. Cell phone–based health education messaging improves health literacy. *Afr Health Sci*. 2016;16(1):311-318.

49. Liu Y-B, Li Y-F, Liu L, Chen Y-L. Effectiveness of the teach-back method for improving the health literacy of senior citizens in nursing homes. *Japan J Nurs Sci*. 2018;15(3):195-202.

50. Guzys D, Kenny A, Dickson-Swift V, Threlkeld G. A critical review of population health literacy assessment. *BMC Public Health*. 2015;15(1):1-7.

51. Guo S, Armstrong R, Waters E, et al. Quality of health literacy instruments used in children and adolescents: a systematic review. *BMJ Open*. 2018;8(6):020080.

52. Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. *Patient Educ Couns*. 1999;38(1):33-42.

53. Chinn D, McCarthy C. All Aspects of Health Literacy Scale (AAHLS): developing a tool to measure functional, communicative and critical health literacy in primary healthcare settings. *Patient Educ Couns*. 2013;90(2):247-253.

54. Finbråten HS, Guttersrud Ø, Nordström G, Pettersen KS, Trollvik A, Wilde-Larsson B. Validating the functional, communicative, and critical health literacy scale using rasch modeling and confirmatory factor analysis. *J Nurs Meas*. 2018;26(2):341-363.

55. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults. *J Gen Intern Med*. 1995;10(10):537-541.

56. Nguyen TH, Paasche-Orlow MK, McCormack LA. The state of the science of health literacy measurement. *Inf Serv Use*. 2017;37(2): 189-203.

57. Zarcadoolas C, Pleasant A, Greer DS. *Advancing Health Literacy: A Framework for Understanding and Action*. Vol 45. John Wiley & Sons; 2009.

58. Stassen G, Grieben C, Sauzet O, Fröböse I, Schaller A. Health literacy promotion among young adults: a web-based intervention in German vocational schools. *Health Educ Res*. 2020;35(2):87-98.

59. Malloy-Weir LJ, Charles C, Gafni A, Entwistle V. A review of health literacy: definitions, interpretations, and implications for policy initiatives. *J Public Health Policy*. 2016;37(3):334-352.

60. Geboers B, de Winter AF, Luten KA, Jansen CJ, Reijneveld SA. The association of health literacy with physical activity and nutritional behavior in older adults, and its social cognitive mediators. *J Health Commun*. 2014;19(Sup 2):61-76.

61. Cha E, Kim KH, Lerner HM, et al. Health literacy, self-efficacy, food label use, and diet in young adults. *Am J Health Behav*. 2014;38(3):331-339.

62. Haun JN, Patel NR, French DD, Campbell RR, Bradham DD, Lapcevic WA. Association between health literacy and medical care costs in an integrated healthcare system: a regional population based study. *BMC Health Serv Res*. 2015;15(1):1-11.

63. Geboers B, Uiters E, Reijneveld SA, et al. Health literacy among older adults is associated with their 10-years' cognitive functioning and decline-the Doetinchem Cohort Study. *BMC Geriatr*. 2018;18(1):1-7.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.