

Multilevel implementation assessment of a national health and life skills curriculum policy: development and content validation

Christine Sundgot-Borgen

`christine.sundgot-borgen@psykologi.uio.no`

University of Oslo

Stephan Daus

The Nordic Institute for the Studies of Innovation, Research and Education

Hildegunn Fandrem

University of Stavanger

Cathrine Pedersen

The Nordic Institute for the Studies of Innovation, Research and Education

Torstein de Besche

The Nordic Institute for the Studies of Innovation, Research and Education

Lene Vestad

University of Stavanger

Thea Steinnes

University of Oslo

Lisbeth M. Brevik

University of Oslo

Oda van Jole

University of Oslo

Neil Humphrey

University of Manchester

Frida Vennerød-Diesen

The Nordic Institute for the Studies of Innovation, Research and Education

Mehdi Poornikoo

University of Oslo

Anne T. Klomsten

Norwegian University of Science and Technology

Ragnhild E. Brandlistuen

The Norwegian Institute of Public Health

Ragnhild B. Nes

University of Oslo
Andreas Ø. Fidjeland
University of Stavanger
Janne M. Vistnes
University of Stavanger
Siri H. Folstad
University of Oslo
Viviana Daza
University of Stavanger
Cecilie Evertsen
University of Stavanger
Geir Skeie
University of Stavanger
Yuan Fang
University of Oslo
Maud Edvoll
University of Oslo
Clara Timpe
University of Oslo
Alf B. Aschim
The Norwegian Association of School Leaders
Eivind Ystrøm
University of Oslo
Espen Røysamb
University of Oslo
Gunnar Bjørnebekk
University of Oslo
Kristina Kymre
University of Oslo
Asgeir K. Kjeldstad
The Norwegian Union of School Employees
Inga Brenne
The Norwegian Network of Health Promoting Communities
Janne M. Øksnes
The Teachers' Union
Lene M. Bakkadal
The Norwegian Union of School Employees
Svein E. Engebretsen
The Teachers' Union

Vigdis Holm

The Norwegian Network of Health Promoting Communities

Egil Nygaard

University of Oslo

Research Article

Keywords: school health promotion, implementation science, curriculum policy, assessment development, content validity, life skills, multilevel assessment

Posted Date: October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11061256/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Curriculum policies addressing health and life skills pose challenges for implementation research because implementation depends on local interpretation and professional judgement rather than uniform content, dosage, or delivery. Existing implementation measures, often developed for standardised interventions, may therefore be poorly suited to capturing how such policies are interpreted and enacted across school-system levels. This study describes the development and content validation of the Health and Life Skills Implementation Assessment (HLS-IA), designed to assess a nationwide curriculum policy across school-owner, school-leader, teacher, and classroom-practice levels.

Methods

Development combined implementation frameworks and reviews of existing measures with stakeholder involvement, two expert workshops, 17 cognitive interviews, and technical and quantitative pilot testing. These iterative procedures examined the relevance, comprehensibility, applicability, and practical feasibility of the questionnaires and informed successive revisions.

Results

The development process resulted in four coordinated HLS-IA questionnaires addressing implementation determinants, processes, and outcomes. The questionnaire targeting classroom practice is administered repeatedly as a lesson-based teacher log and distinguishes planned from enacted practice. Stakeholder involvement shaped both the implementation domains prioritised for assessment and the contextual framing of individual items. Expert appraisal, cognitive interviews, and pilot testing identified problems concerning redundancy, school terminology, interpretation, applicability, response options, and questionnaire routing, leading to the removal, reformulation, and refinement of items and response formats. The first national application in 2025 included responses from approximately 458 schools, demonstrating the practical feasibility of coordinated multilevel data collection while also highlighting recruitment challenges.

Conclusions

Developed through content-validation procedures, the HLS-IA provides a basis for examining how organisational and contextual conditions shape the local enactment of health- and life-skills-related aims in schools and for subsequent psychometric evaluation. The study also offers transferable methodological guidance for assessing complex health-promoting policies implemented through

existing public institutions, particularly when professional judgement, local adaptation, and coordination across system levels are integral to implementation.

1 Background

1.1 Health promotion and life skills in school settings

Schools constitute important settings for population-level health promotion because they reach children and adolescents across socioeconomic and sociocultural backgrounds. In many countries, in addition to traditional academic outcomes, such reforms increasingly emphasise the role of schools in supporting students' health, well-being, and the development of life skills. Within a settings-based approach, schools connect everyday learning with organisational structures, professional practices, and local communities while strengthening students' capacity to manage everyday challenges and make responsible decisions (1, 2). A growing body of research suggests that sustainable impact requires whole-school initiatives characterised by coordinated integration across school subjects, organisational structures, and leadership practices (3, 4). When health-promoting aims are embedded in national curricula, implementation is highly dependent on local interpretation and professional judgement, creating challenges for assessing implementation across policy, organisational, professional, and classroom-practice levels (5, 6).

Reviews suggest that the inclusion of large-scale life skills education initiatives in mainstream education without sustained attention to implementation quality poses the risk of these initiatives being marginalised in overcrowded curricula (1, 2). Successful implementation therefore requires systematic monitoring and structural support. The Norwegian National Curriculum for Knowledge Promotion 2020 (LK20), introduced the interdisciplinary topic Health and Life Skills (HLS), providing an illustrative example of a population-level curriculum policy with explicit health-promoting aims embedded in the ordinary education system (7).

1.2 The interdisciplinary topic of HLS

This study specifically focuses on the implementation of HLS as part of the Norwegian educational reform – a topic implemented with the intention of becoming embedded across most subjects in primary and secondary schools. HLS aims to promote students' physical and mental health, encourage responsible decision-making, and aid in identity development. The topic also seeks to develop students' competence in understanding and managing factors that influence their lives, including a public health perspective, and their skills to critically engage with and shape these factors (7, 8).

Although HLS was not introduced as a formal implementation of the WHO Health-Promoting Schools framework, it shares a settings-based orientation by embedding health and life skills within ordinary educational structures and practices. From a health-promotion perspective, it can therefore be examined as a population-level curricular approach through which health- and life-skills-related aims are enacted in the school system.

Unlike standardised school-based interventions, HLS lacks detailed national implementation guidelines (8). The LK20 reform outlines overarching competencies for students in subject-specific curricula, without specifying content, dosage, or delivery procedures for HLS. School owners- (i.e. representatives of the municipal or county authorities responsible for schools), leaders-, and teachers are granted substantial autonomy in the interpretation and enactment of the topic, consistent with the emphasis of the Norwegian educational policy on local adaptation and stakeholder involvement (9). As such, this broad and flexible framework could promote cocreation among teachers and provide them with innovation opportunities. However, the lack of structured implementation guidelines could contribute to uncertainties regarding the meaning and content of HLS and, therefore, increase teacher workload. Taken together, this flexible framework distinguishes HLS from more clearly specified school-based interventions that are commonly examined in international implementation research. This suggests that insights from intervention research may not fully capture the dynamics of policy-driven implementation, where additional contextual and discretionary factors are likely to shape practice.

1.3 Evaluating curriculum policies with health-promoting aims

Evaluating implementation requires clarity about what constitutes successful implementation. In standardised interventions, implementation can often be assessed against predefined components, dosage, or delivery procedures. In non-standardised curriculum policies addressing health and life skills, however, local interpretation and adaptation may be intended features of implementation rather than deviations from it.

Although school-based mental health initiatives have demonstrated small beneficial effects, on average (10), more recent meta-analyses consistently report substantial heterogeneity, with effectiveness varying across intervention type, target population, outcomes, and study design (11, 12). Whole-school initiatives implemented in real-world conditions often yield smaller effect sizes than interventions tested in controlled settings, reflecting contextual variation and professional discretion in implementation (13). Evidence consistently indicates that the scope and quality of implementation are central determinants of intervention outcomes (13, 14). Implementation research highlights how organisational conditions, leadership, staff capacity, and contextual factors shape implementation outcomes (15, 16), which in turn influence intervention effects (17, 18).

For a curriculum policy addressing health and life skills, such as HLS, these challenges are amplified. As a broadly defined and non-standardised curriculum policy implemented with substantial local autonomy, HLS requires actors across system levels to translate the policy intentions into everyday educational practices. As described in the street-level bureaucracy theory (19, 20), frontline professionals function as de facto policy implementers, translating broad mandates into everyday practice through discretionary decision-making. Such discretionary translation is likely to generate considerable variation in implementation across municipalities, schools, and classrooms. Capturing this variation requires

measurement approaches capable of reflecting multilevel and context-dependent implementation processes.

1.4 The current study

Existing measures have often been developed for more clearly specified or standardised interventions and may be less suitable for complex, multilevel, curriculum policies addressing health and life skills (21, 22). Moreover, there is limited information on how implementation processes in such reforms can be studied at scale. Recent work on Health-Promoting Schools likewise highlights heterogeneity in available implementation instruments and the need for theoretically grounded measures that address implementation processes, outcomes, and a wider range of school actors (23). A multilevel measure tailored to a non-standardised curriculum policy with health-promoting aims may therefore help capture how HLS is interpreted and enacted across the school system. This article describes the development, including content validation, of the HLS Implementation Assessment (HLS-IA). The HLS-IA comprises four coordinated questionnaires targeting school owners, school leaders, teachers, and lesson-level classroom practice; the latter is administered repeatedly as a lesson-based teacher log. The questionnaires were designed to assess implementation determinants, processes, and outcomes across school-system levels in Norway.

This methodological study had three objectives: (1) to describe the theory- and practice-informed development of a multilevel assessment of HLS implementation; (2) to document the content-validation procedures and resulting structure of the assessment; and (3) to describe its national application, including participation patterns relevant to the practical feasibility of multilevel research on school-based policies addressing health and life skills. Based on this work, we discuss methodological considerations for evaluating curriculum policies with health-promoting aims in which implementation depends on professional judgement and local adaptation.

2 Methods

This study is part of two national research initiatives which were established to address the need for systematic knowledge regarding how HLS is implemented across Norwegian schools following the LK20 national reform: *Effects of the Implementation of the Interdisciplinary Topic 'Health and Life Skills' in Schools* (HeaLS) and *Life Skills Education in Norwegian Schools* (LIFE). While LIFE examines the ongoing implementation of life skills education within the LK20 curriculum, HeaLS investigates how variation in HLS implementation relates to outcomes such as students' mental health and academic performance.

2.1 Study design

This methodological instrument-development and content-validation study was informed by the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) (24) and reports the development, content validation, and first national application of the HLS-IA. The overarching study was preregistered on the Open Science Framework (OSF; registration ype2z).

2.2 Project organisation and collaboration

To ensure consistency in measurements across projects and over time, the two independently funded projects collaborated to jointly develop the four HLS-IA questionnaires and align data collection strategies. The two projects have complementary expertise in implementation science, large-scale school surveys, educational research, prevention science, and mixed methods approaches. A dedicated questionnaire development team was established to coordinate and oversee the iterative development and content validation process, including item generation, stakeholder input, piloting, and finalisation of the questionnaires. Importantly, formal practice-based user representation was integrated into the project group. This included representatives from The Association of School Leaders, Teachers' Union, and The Norwegian Union of School Employees, as well as World Health Organisation's (WHO) Healthy Cities Network in Norway. All members of the project group participated in regular meetings and workshops for approximately 1.5 years and contributed to discussions on questionnaire relevance, feasibility, and contextual sensitivity.

The stakeholder advisory group included one teacher and one school leader, each from two different schools, as well as representatives from the Norwegian Directorate for Education and Training (NDET), Norwegian Directorate of Health (NDH), Norwegian Institute of Public Health (NIPH), and Norwegian Council for Mental Health. The stakeholder advisory group contributed contextual knowledge and provided feedback on the recruitment process and dissemination of questionnaires but was not involved in scientific decision-making. A scientific and a national advisory board provided external expertise and ensured alignment with policy and educational sectors, while remaining independent of day-to-day project activities.

2.3 Questionnaire development and content validity

Given the nature of HLS, questionnaire development followed a structured process combining theory-informed item generation with empirical content validation procedures, including expert review, cognitive interviews, and pilot testing (Fig. 1). The overall approach and its key components are outlined below.

[INSERT Fig. 1. here]

2.3.1 Conceptualisation and theoretical framework alignment

The item development process followed an inductive, practice-informed approach. However, the conceptualisation and refinement of constructs were guided by key implementation frameworks.

The HeaLS project built on the Consolidated Framework for Implementation Research (CFIR) (25, 26) and Proctor's taxonomy of implementation outcomes (27). For example, CFIR (25) helped identify determinants across five domains – *innovation characteristics*, the *outer setting*, the *inner setting*, *individuals* (including roles and characteristics), and the *implementation process* – thereby enabling a

systematic analysis of contextual influences on implementation. Specifically, the distinction between outer and inner settings informed our decision to develop items capturing the influence of municipal priorities, policy expectations, and resource conditions (outer setting) on HLS implementation, and of leadership engagement, collaboration structures, and school climate (inner setting) on local HLS enactment. Similarly, attention to the individual's domain guided the inclusion of items targeting teachers' roles and characteristics, including capability, opportunity, and motivation to enact HLS. Proctor's (2011) taxonomy clarified which aspects of implementation should be conceptualised as implementation outcomes (e.g. adoption, feasibility, acceptability, and sustainability) to ensure that items reflected both determinants and manifestations of implementation.

The LIFE project implementation science framework further informed how implementation was conceptualised as an ongoing, multilevel process, combining *implementation process evaluation* (28), *Fixsen's implementation framework* (29, 30), and *Lipsky's theory of street-level bureaucracy* (19). While CFIR (25) structured the identification of contextual determinants, Fixsen's (30) implementation drivers informed the operationalisation of organisational capacity and leadership support. Lipsky's (19) theory provided a lens for understanding how professional discretion shapes the translation of curricular policy into enacted classroom practice through local interpretation and adaptation.

[INSERT Fig. 2. here]

The figure illustrates how four complementary theoretical frameworks (CFIR (25, 26); Fixsen (29, 30); Lipsky (19, 20); Proctor (27)) jointly inform a multilevel, context-sensitive understanding of implementation in school contexts. The overlapping structure reflects their integrative and non-hierarchical use in the present study. Together, these frameworks capture implementation as a dynamic interplay among contextual conditions, organisational capacity, professional practice, and implementation outcomes.

Although distinct, the frameworks complemented each other by sensitising the instrument development process to implementation dynamics across multiple system levels. Together, they supported a combined focus on contextual conditions, organisational processes, and individual enactment of HLS (Fig. 2). Within these theoretically informed domains, specific items were subsequently inductively developed.

2.3.2 Content validity-focused development procedures

2.3.2.1 The practice-informed approach.

- Given the absence of standardised implementation guidelines for HLS and the limited suitability of existing implementation questionnaires for this type of reform, a practice-based approach was applied. This approach was selected to strengthen content validity by grounding item development in real-world implementation practices and constructs perceived as relevant by key stakeholders (31).

2.3.2.2 Defining overarching domains to ensure content coverage.

After identification of key domains relevant to HLS implementation, specific items were iteratively developed and refined through input from the questionnaire development team, the overall project group, and the stakeholder advisory group. To inform the suggested overarching domains, we used semi-structured focus group interviews, with open-ended questions, during stakeholder advisory group meetings. During their initial meetings, they focused on the comprehensiveness of the questionnaires *to be developed* and which constructs they perceived as relevant (i.e. ‘What are the most central implementation topics you would like information about?’). The stakeholder advisory group highlighted the need to assess institutional prioritisation and continuity of HLS at the school level as well as variation in how HLS was perceived, resourced, and enacted across levels (i.e., teachers, school years, and student needs).

2.3.2.3 Defining topics within domains based on stakeholders’ experiences.

There was a need to refine broad domains into more targeted, practice-relevant questions for each of the four HLS-IA questionnaires (i.e. school owner, school leader, and teacher questionnaire, and the lesson-based teacher log). Hence, we conducted a workshop on prioritising overarching questions using a three-point Likert scale (high, medium, low) to assess the intended *comprehensiveness* of each question and determine to whom each question should be directed in the subsequent item development process. The workshop attendees recommended further item generation to explicitly address contextual positioning, stakeholder perceptions, instructional practices, organisational support, adaptation to student needs, and sustainment over time.

2.3.2.4 Literature review to support item formulation.

Strategic literature searches were conducted to identify existing implementation questionnaires that could be adapted. When suitable questionnaires did not exist, insights from related empirical studies – including qualitative research – were used to inform the development of new questionnaire items. In areas where no relevant questionnaires or single-item formulations existed, new items were developed to capture predefined thematic constructs identified as essential to assess. The questionnaire development process was informed by a broad scoping of approximately 30 existing implementation and school-related questionnaires identified by the questionnaire development team through strategic searches. These were initially screened and categorised as relevant, potentially relevant, or not relevant, before 23 questionnaires were reviewed in greater depth. Fourteen questionnaires were selected as conceptual starting points for item development.

Rather than adopting existing scales, items were substantially adapted, reformulated, or newly developed to reflect the specific characteristics of the Norwegian educational HLS context and the target groups (school owners, school leaders, and teachers). National studies on the evaluation of the LK20

curriculum reform – including Questions to School Norway, The Evaluation of the New National Curriculum: Intentions, Processes, and Practices (EVA2020) (32), and Evaluation of the New Curriculum Reform (EDUCATE) (33) – served as important empirical and contextual reference points for operationalising implementation in the Norwegian LK20 setting and informed item development through existing survey formulations as well as insights into experiences, challenges, and classroom practices of school staff.

The lesson-based teacher log required a somewhat different development approach. Its development was grounded in materials and findings from the EDUCATE project. Specifically, we reviewed a classroom observation protocol for life skills education (5), previously developed teacher logs and empirical findings reported by Siljan et al. (34). In the latter study, log data, classroom video observations, and teacher interviews were integrated to examine how life skills education as a distinct curricular element was conceptualised and enacted in classroom practice, drawing on selected EDUCATE video observations and anonymised interview transcripts accessed with appropriate consent.

These qualitative data informed our understanding of how teachers operationalised life skills education in subject teaching. This empirical material informed the operationalisation of key log dimensions, including (a) the distinction between planned and enacted inclusion of HLS education during teaching; (b) variation in how HLS education was integrated, including both planned thematic work and more situational or emergent inclusion; and (c) the categorisation of student involvement and degree of operationalisation. The HLS-IA log translates selected core dimensions into a feasible self-report format suitable for large-scale data collection.

2.3.2.5 Expert-led evaluation and refinement of item relevance and comprehensibility.

Drafted questionnaires were further evaluated through two workshops (one on the teachers' log; one on the other three questionnaires). Group members of HeaLS and LIFE were organised into small groups and asked to provide structured feedback based on predefined evaluation tasks focusing on item relevance, comprehensibility, and the adequacy of response options. Specifically, items were assessed against the questions: 'Are the questions relevant to the respondents' everyday practice?', 'Are the questions clear and easy to understand?', and 'Do the response options allow for meaningful variation in responses?'. The workshop concluded with a plenary discussion. Oral and written feedback from the workshops were systematically used to revise all questionnaires by removing redundant items, aligning wording with school terminology, clarifying items to reflect the HLS-specific context, and refining response formats and survey logic.

To further refine and quality-assess items included in the questionnaire drafts, a structured, template-based evaluation was conducted. Candidate items were assessed using a standardised template completed by a trained research assistant. The templates supported a systematic appraisal of content validity by documenting construct definitions, target groups, scope limitations, and theoretical or empirical sources. Items were evaluated against predefined criteria addressing relevance, coverage of

construct dimensions, and comprehensibility for the intended respondent group. In addition, item wording and response formats were assessed using established survey methodology principles, including clarity, neutrality of phrasing (35–38), avoidance of hidden assumptions (39, 40), and appropriateness of response options. For the adapted items, the evaluation considered the extent of modifications and available evidence of psychometric quality. Based on the template evaluations, items were retained, revised, or excluded, with decisions documented and, when necessary, discussed within the questionnaire development team. Only items that met predefined quality criteria were taken forward.

2.3.3 Cognitive interviews

Face validity of the questionnaires and teachers' log was assessed through iterative cognitive interviews using think-aloud protocols (41). In line with COSMIN recommendations (24), the interviews were used to assess whether items, instructions, and key terms were understood as intended (comprehensibility) and whether the instruments adequately covered relevant aspects of HLS implementation (comprehensiveness). Interviewees were asked to complete the questionnaires while verbalising their thoughts and placing themselves in a realistic response situation (e.g. available time between classroom lessons). The interviewers documented interpretations, hesitations, and spontaneous comments, and asked follow-up questions when unclear wording, ambiguous concepts, or difficulties with response options were identified. Researchers experienced with cognitive interviewing conducted the interviews between June 2024 and January 2025. Seventeen individual interviews included actors across school levels, recruited through professional networks and social media. Nine interviews focused on the teacher log and eight on the other three questionnaires. Each interview lasted approximately one hour, and participants received a financial compensation (NOK 450–550) to acknowledge and reimburse their time. An organised session with a school leader union assembly meeting provided additional feedback to issues identified in the interviews. Anonymous notes from each interview were made, saved, and shared with the questionnaire development team. Findings from early interviews informed revisions that were tested in subsequent interviews. The cognitive interviews identified items that required revision to ensure applicability across respondents with different levels of HLS experience.

2.3.4 Pilot testing

Following the cognitive interviews at the end of January 2025, the questionnaire development team conducted a technical piloting of the finalised log and questionnaires among the 34 project members as well as additional academic colleagues outside the project who were invited to provide feedback. Each person responded with a pre-defined respondent profile in mind. This was intended to identify issues such as spelling mistakes, routing errors, illogical response choices, and suitability for certain participant backgrounds. Additionally, we estimated completion time and adjusted the questionnaires accordingly.

Next, the questionnaire development team conducted small-scale quantitative pilots to gauge the distribution of response categories and 'don't know' answers as indicators of response option misalignment and completion time and percentage of non-completion as indicators of fatigue. The team also investigated open-ended comments at the end of the questionnaire. In early February 2025, we

collected piloting results from four schools; specifically, 4 school leader questionnaire responses, 56 teacher questionnaire responses, and 46 teacher log entries (Table 1).

Table 1
Overview of the Number of Responses Across Populations.

School year 2024–2025	School owner representatives	School leader representatives	Teacher questionnaire	Teacher log
Population ^a	372	2939	108 000	
Original sampling frame ^b	372	518	16 973	
- Pilot sent	0	4	140	132
- Main invitations sent	372	518	8353	8620
Total sampling frame ^c	372	522	8493	8752
- Pilot responses received	0	4	56	46
- Main responses received	95	322	2156	4566
- Self-registrations received	8	30	6	72
Responses received ^d	103	356	2218	4684
- Duplicate pilot– main ^e	0	2	12	0
- Incompletes (< 30%)	25	35	0	632
Valid responses (≥ 30%) ^f	78	319	2206	4052
Distinct responders	78	319	2206	2148
Distinct organisational units ^g	78	319	415	458

School year 2024–2025	School owner representatives	School leader representatives	Teacher questionnaire	Teacher log
<i>Notes.</i>				
^a See section on ‘Populations’ for definitions. The teacher questionnaire, log population, and original sampling frame involves a shared population.				
^b The original sampling frame comprised respondents from schools with leaders’ positive support and available contact information. Self-registrations were not included here. Teachers were randomised equally into questionnaire and log groups.				
^c The total ‘sampling frame’ consisted of pilot respondents, main invitations, and self-registrations.				
^d The numbers for teacher logs are separate responses, up to five per respondent. Three teachers submitted one more than the four requested logs, all of which were retained.				
^e In the final analysis data frame, only unique responses were retained. Duplicate entries occurred when the same respondent submitted both a pilot and a main survey response (duplicate pilot–main). When deciding which duplicates to exclude, we prioritised questionnaires with the most questions answered; if this still tied two responses, we retained the earliest response based on response date.				
^f Responses with at least 30% item completion were classified as valid; those below this threshold were classified as incomplete. This commonly used threshold provides a rough indication of available respondent data.				
^g The number of nearest distinct units (school owners or schools) among respondents with at least 30% response completion.				

Table 2
Findings and revisions across stages of HLS-IA development

Development stage	Sources of input	Main findings	Resulting revisions or decisions
Stakeholder-informed domain development	Six stakeholder advisory-group members representing school practice, education, health and mental-health perspectives	Institutional prioritisation, continuity, resources, variation across school levels, instructional practices, adaptation to student needs and sustainment were identified as important areas of assessment.	These areas were prioritised when defining questionnaire domains and generating items for the different system levels.
Review of existing measures	Approximately 30 questionnaires screened; 23 reviewed in greater depth; 14 selected as conceptual starting points	Existing measures did not adequately capture the non-standardised, multilevel and context-dependent implementation of HLS.	Existing items were substantially adapted or reformulated, and new context-specific items were developed.
Development of the lesson-based teacher log	Observation protocols, previous teacher logs, classroom observations and teacher interviews from EDUCATE	Relevant dimensions included planned versus enacted HLS, planned versus emergent inclusion, student involvement and degree of operationalisation.	These dimensions were translated into a repeated self-report format suitable for national data collection.
Expert workshops and structured item appraisal	Two workshops and template-based appraisal of candidate items	Redundancy, unclear or unsuitable terminology, hidden assumptions, construct-coverage limitations and problems with response options and survey logic were identified.	Items were retained, revised or excluded; wording was aligned with school terminology, and response formats and routing were refined.
Cognitive interviews	17 interviews: nine concerning the teacher log and eight concerning the other questionnaires	Problems concerning interpretation, applicability and response options were identified. Some items could not be answered meaningfully by respondents with limited HLS experience.	Items, instructions and response options were revised. Changes following early interviews were tested in subsequent interviews.
Technical and quantitative pilot testing	34 project members and invited academic colleagues; four schools, four school-leader responses, 56 teacher responses and 46 teacher-log entries	Testing examined technical errors, routing, response-option alignment, applicability, completion time, non-completion and open-ended feedback.	Technical errors were corrected, response options and routing were refined, and questionnaire length and completion time were reviewed.
Final assessment	Evidence synthesised across all stages	The development process supported coverage of determinants, processes and outcomes across system levels and differentiation	Four coordinated HLS-IA questionnaires were finalised for national administration.

Development stage	Sources of input	Main findings	Resulting revisions or decisions
		between planned and enacted classroom practice.	

Table 3
Teacher log responses by session and cumulative number of completed sessions.

Log session	#1	#2	#3	#4	#5 ^a	Total
Valid logs ($\geq 30\%$) ^b	1550	1096	759	644	3	4052
Incomplete logs ($< 30\%$) ^b	241	171	140	80	0	632
Total number of log sessions	1	2	3	4	5	Total
Respondents with valid logs ($\geq 30\%$)	1 142	419	278	307	2	2148
<i>Notes.</i>						
^a Three teachers misunderstood and submitted one more than the four requested logs, which were retained.						
^b Responses with at least 30% item completion were classified as valid; those below this threshold were classified as incomplete. This commonly used threshold provides a rough indication of available respondent data.						

2.3.5 Data synthesis and reporting

Evidence generated through workshops, template-based item evaluations, cognitive interviews, and pilot testing was used iteratively to inform questionnaire revision. The material was not analysed as an independent qualitative dataset intended to generate themes or test hypotheses. Findings from the development process are therefore reported descriptively, focusing on documented categories of revisions and the resulting assessment structure. National application is described using response rates, log-completion patterns, and comparisons between respondents and the relevant sampling frames.

2.3.6 Measures

At this stage, the HLS-IA questionnaires (Additional file 1, an Excel file with one worksheet for each instrument) were presented at the domain level as part of the development and content validation phase. Psychometric evaluation of the scale structure and measurement properties represents a subsequent stage in the development and evaluation of the HLS-IA, preregistered in the Open Science Framework (OSF) (Sundgot-Borgen et al., 2026).

The HLS-IA includes item-level variables targeting *implementation determinants* (e.g. resources, organisational climate, planning, and capability) (25), *implementation outcomes* (e.g. adoption, acceptability, feasibility, and sustainability) (27), and *implementation processes* (e.g. training, leadership

engagement, and facilitative administration) (25, 29). The questionnaires include Likert-type items assessing e.g., perceived extent, agreement, or frequency (e.g. 1 = strongly disagree to 5 = strongly agree). Additional categorical items are included where relevant (e.g., availability of resources, types of implementation activities).

Teacher logs provide event-based measures of planned and enacted classroom practices. Teachers were asked to complete a short log immediately after each of four separate lessons, reporting planned learning goals, planned teaching approaches, actual HLS implementation, adaptations, time use, and perceived teaching challenges. The teacher log comprised brief checklists and short Likert-format ratings designed to capture proximal variation in enacted practice.

Across log and questionnaires, additional background variables (school level, teachers' subject affiliation, municipality characteristics, formal education/professional qualifications, age, and gender) are included to enable subsequent subgroup analyses and psychometric evaluation across relevant subgroups.

2.4 Participants

2.4.1 Populations

We targeted three main populations within Norway: (A) public-school-owner representatives, including the 15 administrative counties responsible for upper-secondary schools (A1), the 357 municipalities responsible for primary and lower-secondary schools (A2), and the Svalbard administrative area; (B) public- and private-school-leader representatives, including 384 upper-secondary (B1) and 2555 lower-secondary and primary schools (B2), registered in the National School Registry in 2024–2025; (C) teachers, including approximately 108,000 upper-secondary, lower-secondary, and primary school teachers in 2024 (42). Due to lack of centralised, updated, and well-defined databases of teachers, we asked school leaders to provide the contact information for all their staff involved in teaching or educational activities, broadly defined, excluding the school leader.

2.4.2 Sample sizes and power calculations

The two projects address multiple research questions across different populations. While certain analyses require aggregation of responses to the school level, others require high response rates in the overall teacher sample. Accordingly, large samples of both schools and teachers were required, including schools with responses from both teachers and a school-leader representative.

Power calculations for intervention effects are reported in the respective analytical articles. As the present article focuses on methodological aspects of questionnaire development, including content validation, the key consideration is whether the available sample will be adequate for subsequent psychometric analyses. A priori simulation studies suggest that multilevel factor analyses require approximately 100 clusters, with 5–10 respondents per cluster. Based on prior evidence, we expected a teacher questionnaire response rate of 44% (43) and participation from fewer than 30% of all schools.

2.4.3 Data collection

The questionnaire development team was granted access to updated contact information for population (A) school owners and (B) school leaders through NIFU's database. School owners and school leaders were recruited between March 2024 and May 2025.

The school owners were contacted directly after distribution of corresponding questionnaires. School-owner recipients were permitted to forward the questionnaire to other school-owner representatives. Similarly, school leaders were invited by email and phone and were first asked to approve their school's participation in both the school leader and teacher questionnaires. This step was essential for practical and ethical reasons, as school leaders act as gatekeepers for external requests to teachers and provide access to teacher contact information.

Invited participants received an email containing a personalised link to the electronic questionnaire. However, due to low response rates among school leaders and slightly concerning teacher participation rates, procedures were changed to permit teachers to participate by directly self-registering (from January 2025) through an open web link distributed among schools, and through Facebook groups, advertisements on the projects' websites, and in teachers' magazines.

School owners and leaders were each asked to complete one questionnaire. Teachers were randomly assigned to one of two groups: C1: completed the teacher questionnaire once, and C2: completed the lesson-based teacher log following each of four classroom lessons, in addition to background questions. Screening questions were included in the teacher logs and surveys to filter respondents and ensure that school leaders, teachers, and administrative staff responded to the correct questionnaires.

2.4.4 Participant outreach and recruitment strategies

To aid recruitment, a multipronged outreach strategy was employed to increase awareness of and interest in the study, with particular emphasis on school owners and leaders (Additional file 2). First, information regarding the project was disseminated among all potential participants through professional magazines, podcasts, videos, newsletters distributed via trade unions and municipal channels, project webpages, as well as social media. In addition, school owners and school leaders were invited to a series of digital information meetings.

Second, targeted invitation emails were distributed to communicate the study purpose and practical relevance for schools. Research staff made follow-up calls to nonresponsive schools and directly contacted large school owners to facilitate centralised recruitment. School owners and school leaders already involved in a qualitative research study of the LIFE project were contacted and encouraged to participate in the questionnaire study.

Finally, participation was encouraged through a combination of individual and school-level incentives, opportunities for professional development, and access to aggregated feedback at school-owner and school levels (Additional file 2). Multiple reminders were sent, and school leaders were informed about

their school's overall response rate to enable local encouragement of participation while maintaining individual confidentiality.

2.4.5 Informed consent

Written informed voluntary consent was electronically obtained via the first page of each online questionnaire, where participants confirmed that they had read the study and privacy information before proceeding. Participants were informed regarding the purpose of the study, what participation entailed, incentives, data handling and sharing, the legal basis for data processing, their rights (withdrawal, access, correction, deletion, and complaint), and contact details for the project leaders and data protection officer.

2.4.6 Data security and complex data sharing among projects and institutions

Data were collected through Qualtrics and stored on servers located in the European Union (EU). Technical metadata were deleted upon downloading, and identifiable information was separately stored from questionnaire responses and linked only through generic IDs, used for reminders, response monitoring, and potential follow-up. Data were stored on secure, appropriate data storage platforms.

To support longitudinal follow-up, the HLS-IA was administered again in spring 2026, one year after the first national data collection, using the same questionnaires. The 2025 and 2026 data were collected by NIFU and the University of Stavanger, respectively. De-identified questionnaire data and linkage identifiers were transferred between institutions to enable respondent-level linkage across data collections. All transfers followed national standards for encrypted file transfer, ensuring privacy protection and regulatory compliance. The present study reports the development and first national application of the HLS-IA in 2025; longitudinal analyses are outside its scope.

2.5 Generative AI tools

During the preparation of this work the authors used ChatGPT (OpenAI, GPT-5.3) in order to assist with language editing and improve the clarity of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the article.

3. Results

3.1 Resulting multilevel assessment and documented revisions

The development process resulted in four coordinated HLS-IA questionnaires aligned across school-owner, school-leader, teacher, and classroom-practice levels. Table 2 summarises the sources of input, main findings, and resulting revisions or decisions across the stages of questionnaire development and content validation.

[Insert Table 2. Findings and revisions across stages of HLS-IA development here]

Across the development stages, stakeholder involvement shaped both the implementation domains prioritised for assessment and the wording and contextual relevance of individual items. The review of existing measures showed that substantial adaptation was needed to assess a non-standardised curriculum policy implemented across multiple system levels. Expert appraisal, cognitive interviews, and pilot testing identified problems concerning redundancy, school terminology, interpretation, applicability, response options, and questionnaire routing. These findings informed successive revisions, including the removal, reformulation, and refinement of items and response formats. The resulting questionnaires address implementation determinants, processes, and outcomes at the item and domain levels. The lesson-based teacher log additionally distinguishes planned from enacted classroom practice. Psychometric evaluation of the scale structure and measurement properties constitutes a subsequent stage in the development and evaluation of the HLS-IA.

3.2 National application and participation patterns

Table 1 provides an overview of the populations, sampling frames, and the final total valid respondents for the first data collection. Response rates were calculated as the number of valid responses to main invitations divided by the number of main invitations sent. The response rate was highest for school leaders (61.6%) and similar for school owners (21.0%) and teacher questionnaires (25.7%). Response rates for the teacher logs were influenced by the fact that while teachers were asked to complete the log for four lessons, certain teachers submitted fewer and three teachers submitted five logs. The log response rate was 24.9%, based on the number of distinct teachers with at least one valid log ($\geq 30\%$), whereas only 4% of distinct teachers submitted four valid logs (Table 3).

With data from 458 schools, the national application generated a dataset suitable for subsequent psychometric evaluation of the HLS-IA.

[INSERT Table 1. Overview of the Number of Responses Across Populations]

[INSERT Table 3. Teacher log responses by session and cumulative number of completed sessions]

Assessing the representativeness of respondents relative to the population is notoriously difficult, as we lack variables linking response propensity to the outcome of interest. What we do have are demographic characteristics of municipalities and schools: language form, region, and municipal centrality for all respondent groups, as well as school size, type, and ownership type for schools. In Additional file 3, the tables show no statistically significant differences between school-owner and school-leader respondents and the corresponding population distributions on the examined characteristics; teacher-questionnaire and teacher-log respondents did not differ significantly from the corresponding populations in region, school size, or private/public ownership. Compared to the population, the schools represented in the teacher samples included proportionally more upper-secondary schools and fewer

primary/lower-secondary schools, more in urban areas, and for the teacher logs – a higher proportion of schools using Bokmål (one of Norway's two written standards and the most widely used form of written Norwegian). Cohen's w indicated that the observed deviations were small, although selection bias cannot be excluded.

Moreover, 310 teachers were invited to complete both the teacher logs and the teacher questionnaire for the purpose of linking responses across the logs and questionnaires. Of these, 134 answered both the teacher questionnaire and at least one log, of which 120 completed all 4 logs.

4. Discussion

This study addressed a methodological challenge at the intersection of health-promotion and education research: how to assess implementation when a curriculum reform does not prescribe uniform content, dosage, or delivery procedures. Using HLS as a national case of a curriculum policy addressing health and life skills, we developed and content-validated the HLS-IA, a coordinated multilevel assessment spanning policy, organisational, professional, and classroom-practice levels. The study contributes by presenting the HLS-IA and documenting its theory-informed development and content-validation process, while also illustrating broader design considerations for evaluating settings-based policies in which local interpretation and professional judgement are integral to implementation. Further research is needed to examine the instrument's structural validity, reliability, and measurement invariance before scale-level interpretations can be recommended.

4.1 Capturing HLS implementation: the need for tailored measures

Several characteristics of HLS are central for interpreting the development process of this study. The cross-curricular topic HLS is mandatory and nationwide, infused within and across most school subjects and all school years. Additionally, teachers are expected to implement HLS in their classrooms without a manual, predefined dosage or standardised delivery procedures. Therefore, implementation is shaped by cocreation, professional judgement, and everyday teaching practices (5, 6). As such, HLS cannot be treated or investigated as a bounded intervention but rather as an ongoing and context-dependent implementation process (22). Thus, rather than viewing implementation as the delivery of a predefined programme, the questionnaires were designed to capture implementation as how schools and teachers interpret and translate curricular intentions into everyday school environments and classroom practices (34, 44, 45). By aligning the way implementation is conceptualised with the nature of this curriculum policy with health-promoting aims, these questionnaires were designed to better reflect implementation in the context of HLS than measures originally developed for manualised school-based interventions.

4.2 Measuring implementation in curriculum policies with health-promoting aims

Variation in implementation should not automatically be understood as weak fidelity. In a curricular reform such as HLS, variation is built into the design. School staff are expected to interpret, cocreate, and integrate broad curricular intentions into their existing practices (9). Therefore, evaluating implementation against a standard of uniform delivery would risk misinterpreting legitimate professional judgement and contextual adaptation as implementation failure.

In this respect, this study conceptualised variation as analytically meaningful rather than problematic. By combining implementation frameworks, we integrated determinants from CFIR (25, 26), with implementation drivers described by Fixsen (30) to differentiate among contextual conditions, organisational support, and individual enactment. Contextual conditions primarily reflect determinants located in the CFIR outer setting domain, whereas organisational support captures features of the inner setting domain and implementation drivers related to leadership and organisational capacity. Individual enactment reflects the roles and characteristics of the individual domain as well as enacted classroom practices captured through the teacher logs (34). This integration was designed to support future analyses of whether differences in reported practices reflect structural constraints, limited capacity, or deliberate pedagogical choices; the present study does not test these explanations.

Simultaneously, integrating multiple frameworks increases conceptual and analytical complexity and places greater demands on construct specification and interpretation. Methodologically, however, this approach shifts the focus from compliance with predefined components to how policy intentions are translated into practice under different conditions (46). Such an approach may be particularly relevant for evaluating large-scale curriculum policies with health-promoting aims embedded in school systems that rely on professional discretion rather than standardised procedures.

For health-promotion research, the study illustrates the methodological challenges involved in assessing such policies as part of existing institutional structures rather than as bounded interventions. The HLS-IA links information across policy, organisational, professional, and classroom-practice levels and may thereby support examination of how contextual conditions shape the enactment of health- and life-skills-related aims in schools. The assessment is not intended to prescribe a uniform model of implementation, but to make local interpretation and variation empirically examinable.

4.3 Cross-sector collaboration and stakeholder involvement

The development process brought together expertise from implementation science, educational research, prevention science, and public health, alongside stakeholders from schools, education authorities, health authorities, and civil society. This cross-sector collaboration strengthened the

relevance and usability of the HLS-IA and helped align scientific constructs with the realities of enacting curricular aims related to health and life skills through ordinary school structures (31).

Coordination across the two research initiatives also reduced duplicate data collection and enabled information to be aligned across system levels and over time. This is particularly relevant to research on health and life skills in school settings, where implementation spans policy, organisational, professional, and classroom contexts.

4.4 Strengths and limitations

This study has several strengths. First, we are not aware of previous studies reporting comparable questionnaire development for non-standardised curriculum reforms across multiple levels of the education system. Employing several theoretical frameworks (CFIR, Proctor, Fixsen, and Lipsky) to inform the item generation – with perspectives of implementation determinants, process, and outcomes – strengthened the ability to capture the complex nature of such an initiative (19, 25-27, 30). Second, we combined practice and research-based expertise in the item generation process, which was aided by the collaboration among different institutions. Third, the use of cognitive interviews and quantitative piloting provided complementary evidence relevant to content validity and practical application. Finally, the coordinated design and repeated national administration of the same questionnaires in 2025 and 2026 provide a basis for longitudinal analyses and further evaluation of the HLS-IA's measurement properties over time.

The response rate was generally low, particularly among school owners and teachers, which may affect representativeness and statistical power. Although the observed deviations from the population were small in magnitude, selection bias cannot be excluded. Lastly, there were recruitment challenges in this study, which likely reflect broader contextual constraints in contemporary school-based research.

The study documents the theory-informed development of the HLS-IA, evidence concerning content relevance and comprehensibility, and practical experience from its national application. Its scope is limited to instrument development and content validation; psychometric evaluation of the scale structure and measurement properties represents a subsequent stage in the development and evaluation of the HLS-IA. Accordingly, the assessment is presented at the item and domain levels rather than as a set of fully validated scales. Analyses of implementation levels, relationships across system levels, and associations with student outcomes fall outside the scope of the present study. The principal transferable contribution lies in the development approach and its implications for assessing curriculum policies with health-promoting aims, as the HLS-specific items may require contextual adaptation before application in other educational systems.

4.5 Implications for public health research, policy and practice

The four HLS-IA questionnaires developed in this study (Additional file 1) may support future analyses of implementation in cross-curricular and whole-school initiatives that rely on professional judgement rather than standardised procedures. Although this study does not report empirical implementation outcomes, the development process yielded several methodological insights. First, combining theory-informed frameworks with an inductive, practice-informed approach helped capture implementation across multiple system levels. Second, stakeholder involvement shaped both item wording and the aspects of implementation prioritised for assessment. Third, iterative workshops and cognitive interviews were critical for identifying problems concerning interpretation, applicability, and response options. Finally, the lesson-based teacher log demonstrated the importance of distinguishing planned and enacted practice and assessing the feasibility of repeated, real-time data collection.

Together, these insights may inform future measurement development, school-based reflection and decisions concerning implementation support. Although developed in Norway, the HLS-IA items may require adaptation elsewhere; however, the methodological challenges are relevant to education systems in which broad curricular priorities are interpreted and enacted locally. Participation patterns also illustrate the challenges of conducting national, multilevel research in school settings. Survey fatigue and gatekeeping may limit recruitment even when strategies are developed with stakeholders and adapted to different system levels (47). Future studies should therefore match recruitment strategies to institutional capacity and consider whether more resource-intensive, relationship-based approaches are needed to sustain research–practice collaboration (48).

5. Conclusion

Evaluating curriculum policies addressing health and life skills requires measures that reflect intended local discretion, align information across system levels, and distinguish general implementation conditions from enacted classroom practice. The HLS-IA offers a multilevel assessment developed through content-validation procedures for this purpose and illustrates both the feasibility and the challenges of coordinated national data collection. Psychometric properties were not examined in the present study; accordingly, the HLS-IA is presented at the item and domain levels, and no scale-level interpretations are made. More broadly, the study provides a transparent example of how implementation assessment may be designed when health- and life-skills-related aims are embedded in existing education systems and professional judgement and contextual adaptation are inherent features of implementation.

Abbreviations

LK20

Norwegian National Curriculum for Knowledge Promotion 2020

CFIR

Consolidated Framework for Implementation Research

COSMIN

Consensus–based Standards for the Selection of Health Measurement Instruments

NDET

Norwegian Directorate for Education and Training

NDH

Norwegian Directorate of Health

NIPH

Norwegian Institute of Public Health

WHO

World Health Organization

OSF

Open Science Framework

NIFU

Nordic Institute for Studies in Innovation, Research and Education

EU

European Union

AI

artificial intelligence

Declarations

Ethics approval and consent to participate

The data collection conducted in 2025 was evaluated by the internal research ethics committee at the Department of Psychology at the University of Oslo (ref. 33569637). In accordance with requirements from the funder, the European Research Council, the HeaLS-project has an independent ethical advisor who reports to the council. Furthermore, on assignment from the University of Oslo, the data protection services at Sikt – the Norwegian Agency for Shared Services in Education and Research, has assessed that the processing of personal data is in accordance with the data protection regulations (ref. 103488). All participants provided written informed consent electronically before completing the questionnaires.

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work is funded by the European Research Council (Grant No. 101088481), the Research Council of Norway (Grant No. 288083, 314601, and 343159), the Foundation DAM (2021/FO347381), and the Department of Psychology, University of Oslo. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

Conceptualization: All authors; Methodology: All authors; Validation: All authors; Investigation: EN, CSB, SD, CP, TB, NH, MP, FD, AF, LV, JMV, and TKS; Data curation: CT, KK, and OJ; Formal analysis: TB and SD; Project administration: EN and HF; Funding acquisition: EN and HF; Writing—original draft: CSB, SD, EN, HF, LV, and TKS; Writing—review and editing: All authors. All authors contributed to defining and refining the focus, content, and methodological design of the HLS-IA questionnaires. This included iterative assessment of item relevance, comprehensibility, practical applicability, and organisational relevance. All authors critically reviewed and approved the final manuscript and agree to be accountable for all aspects of the work.

Acknowledgements

We thank all participating school owners and schools for their time and engagement in this study. We are grateful to the school owners, school leaders, and teachers who contributed to the stakeholder advisory group, as well as to those who participated in cognitive interviews and pilot testing. We also sincerely thank the school staff who completed the questionnaire and teacher logs; their contributions were essential to the study.

Preregistration

The overarching study was preregistered on the Open Science Framework (OSF) (ref: ype2z).

Clinical trial number

Not applicable

References

1. Hvalby L, Guldbrandsen A, Fandrem H. Life skills in compulsory education: A systematic scoping review. *Educ Sci*. 2024;14(10):1112.
2. United Nations Children's Fund Evaluation office [Unicef]. Global Evaluation of Life Skills Education Programmes. New York: United Nations Children's Fund Evaluation office; 2012.
3. Gugglberger L. A brief overview of a wide framework—health promoting schools: a curated collection. *Health Promot Int*. 2021;36(2):297–302.

4. Goldberg JM, Sklad M, Elfrink TR, Schreurs KM, Bohlmeijer ET, Clarke AM. Effectiveness of interventions adopting a whole school approach to enhancing social and emotional development: a meta-analysis. *Eur J Psychol Educ.* 2019;34(4):755–82.
5. Brevik LM, Gudmundsdottir GB, Doetjes G, Barreng R. Å observere fagfornyelsen i klasserommet. Observasjonsprotokoller for livsmestring, utforskning og digital kompetanse (EDUCATE-rapport 1) [Observing the curriculum reform in the classroom. Observation protocols for life skills, inquiry-based learning, and digital competence (EDUCATE-report 1)]. Institutt for lærerutdanning og skoleforskning. Universitetet i Oslo: University of Oslo; 2023.
6. Evertsen I, Brevik LM. Life skills education in secondary language classrooms: Empathy, communication and interpersonal relations. *J Curriculum Stud.* 2025;57(2):164–83.
7. Norwegian Directorate for Education and Training [NDET]. Folkehelse og livsmestring [Health and life skills] Oslo2020 [Available from: <https://www.udir.no/lk20/overordnet-del/prinsipper-for-laring-utvikling-og-danning/tverrfaglige-temaer/folkehelse-og-livsmestring/>]
8. Ministry of Education and Research. Core curriculum–values and principles for primary and secondary education. The National Curriculum for the Knowledge Promotion2017. p. 21.
9. Opplæringsutvalget. Ny opplæringslov [New Education Act]. Oslo: Norwegian Ministry of Education and Research. 2019. Report No.: NOU 2019: 23.
10. Clarke A, Sorgenfrei M, Mulcahy J, Davie P, Friedrich C, McBride T. Adolescent mental health. A systematic review on the effectiveness of school-based interventions. 2021.
11. Cilar L, Štiglic G, Kmetec S, Barr O, Pajnkihar M. Effectiveness of school-based mental well-being interventions among adolescents: A systematic review. *J Adv Nurs.* 2020;76(8):2023–45.
12. Lemberger-Truelove ME, Li D, Kim H, Hill DD, Dickson R, Kang Z. School Mental Health Interventions for Adolescents: A Meta-Analysis of Effectiveness and Relevant Moderators. *Adolescents.* 2026;6(1):6.
13. Wigelsworth M, Lendrum A, Oldfield J, Scott A, ten Bokkel I, Tate K, et al. The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: a meta-analysis. *Camb J Educ.* 2016;46(3):347–76.
14. Schultes M-T. An introduction to implementation evaluation of school-based interventions. *Eur J Dev Psychol.* 2023;20(1):189–201.
15. Kim B, Sullivan JL, Ritchie MJ, Connolly SL, Drummond KL, Miller CJ, et al. Comparing variations in implementation processes and influences across multiple sites: What works, for whom, and how? *Psychiatry Res.* 2020;283:112520.
16. Weare K, Nind M. Mental health promotion and problem prevention in schools: what does the evidence say? *Health Promot Int.* 2011;26:129–69.
17. Williams NJ, Beidas RS. Annual Research Review: The state of implementation science in child psychology and psychiatry: a review and suggestions to advance the field. *J Child Psychol Psychiatry.* 2019;60(4):430–50.

18. Allen P, Pilar M, Walsh-Bailey C, Hooley C, Mazzucca S, Lewis CC et al. Quantitative measures of health policy implementation determinants and outcomes: a systematic review. *Implement Sci.* 2020;15(1).
19. Lipsky M. *Street-level bureaucracy: Dilemmas of the individual in public service.* Russell sage foundation; 2010.
20. Lipsky M. *Street-level bureaucracy: Dilemmas of the individual in public services.* New York: Russell Sage Foundation; 1980.
21. Nilsen P, Bernhardsson S. Context matters in implementation science: a scoping review of determinant frameworks that describe contextual determinants for implementation outcomes. *BMC Health Serv Res.* 2019;19.
22. Hahnrahts MT, Willeboordse M, Van Schayck OC. Challenges in evaluating implementation and effectiveness in real-world settings: evaluation proposal for school-based health-promoting intervention. *Health Promot Int.* 2023;38(1):daac185.
23. Delbosq S, Bessems K, Branda M, Dadaczynski K, Vennegoor G, Leksy K, et al. A scoping review of instruments measuring the implementation of the Health-Promoting School approach. *Health Promot Int.* 2026;41(4):124.
24. Terwee CB, Prinsen CA, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res.* 2018;27(5):1159–70.
25. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. *Implement Sci.* 2022;17(1):75.
26. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.* 2009;4(1):50.
27. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm policy mental health mental health Serv Res.* 2011;38(2):65–76.
28. Humphrey N, Lendrum A, Ashworth E, Frearson K, Buck R, Kerr K. *Implementation and process evaluations (IPE) for interventions in education settings: A synthesis of the literature.* The University of Manchester; 2016.
29. Fixsen DL, Blase KA. Active implementation frameworks. *Handbook on implementation science* 2020. pp. 62–87.
30. Fixsen DL, Naoom SF, Blase KA, Friedman RM, Wallace F. *Implementation research: A synthesis of the literature.* Florida, USA: National Implementation Research Network; 2005.
31. DeVellis RF. *Measurement in the Broader Research Context. Scale development: Theory and applications.* 4 ed. Sage; 2017.
32. University of Oslo. *The Evaluation of the New National Curriculum: Intentions, Processes, and Practices (EVA2020)* Oslo: University of Oslo; 2022 [Available from:

<https://www.uv.uio.no/english/research/projects/eva2020/index.html>

33. University of Oslo. Evaluation of the new curriculum reform (EDUCATE) Oslo: University of Oslo. 2021 [Available from: <https://www.uv.uio.no/ils/english/research/projects/educate/index.html>]
34. Siljan HH, Magnusson CG, Goldschmidt-Gjerløw B. Livsmestring i norskfaget: Hvordan forstår og tilrettelegger norsklærere for livsmestringsarbeid i klasserommet? [Life Skills Education in Norwegian Language Arts: How Do Teachers Understand and Approach Life Skills in Their Instruction?]. *Nordic J Lit Res*. 2024;10(3).
35. Aithal A, Aithal P. Development and validation of survey questionnaire & experimental data—a systematical review-based statistical approach. *Int J Manage Technol Social Sci (IJMTS)*. 2020;5(2):233–51.
36. Artino AR Jr, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide 87. *Med Teach*. 2014;36(6):463–74.
37. Brancato G, Macchia S, Murgia M, Signore M, Simeoni G, Blanke K, et al. Handbook of recommended practices for questionnaire development and testing in the European statistical system. Italian National Institute of Statistics; 2006.
38. Irwing P, Booth T, Hughes DJ. The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development. Wiley; 2018.
39. Saris WE, Gallhofer IN. Design, Evaluation, and Analysis of Questionnaires for Survey Research. Wiley; 2007.
40. Stats NZ. A guide to good survey design. Stats NZ; 2025. Report No.: 978-1-991307-72-9.
41. Padilla J-L, Leighton JP. Cognitive Interviewing and Think Aloud Methods. In: Zumbo BD, Hubley AM, editors. Understanding and Investigating Response Processes in Validation Research. 69. Switzerland: Springer Nature; 2017. pp. 211–28.
42. Statistics Norway. Ansatte i barnehage og skole [Employees in kindergartens and schools] 2025 [Available from: <https://www.ssb.no/utdanning/barnehager/statistikk/ansatte-i-barnehage-og-skole>]
43. Wu M-J, Zhao K, Fils-Aime F. Response rates of online surveys in published research: A meta-analysis. *Computers Hum Behav Rep*. 2022;7:100206.
44. Furberg A, Rødnes KA, Silseth K, Arnseth HC, Kvamme OA, Lund A, et al. Fagfornyelsen i møte med klasseromspraksiser: Læreres planlegging og gjennomføring av undervisning om tverrfaglige tema. Sciences FoE: University of Oslo; 2023. Report No.: 6.
45. Gunnulfsen AE, Jensen R, Hall JB, Abrahamsen HN, Kleveland H, Olsen RV. Underveis med nytt læreplanverk. Det utdanningsvitenskapelige fakultet, Universitetet i Oslo; 2022. Report No.: 5.
46. Ogden T, Fixsen DL. Implementation science. *Z Psychol*. 2015;222:4–11.
47. Ghafourifard M. Survey Fatigue in Questionnaire Based Research: The Issues and Solutions. *J caring Sci*. 2024;13(4):214–5.
48. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol*. 2007;17(9):643–53.

Figures

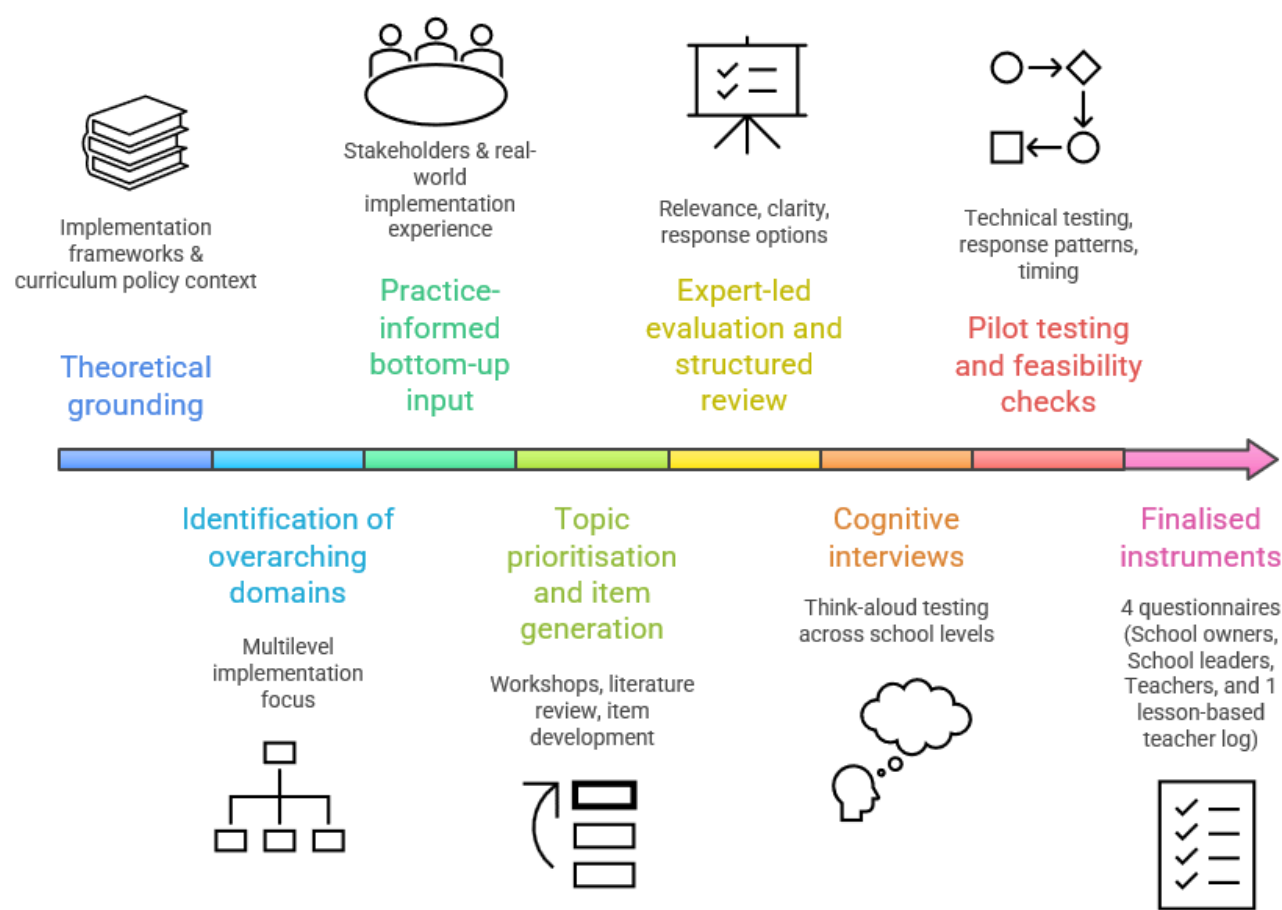


Figure 1

Overview of the Questionnaire Development Process.

The figure summarises the theory-, practice-, and stakeholder-informed stages of questionnaire development, content validation, cognitive testing, and pilot testing leading to the four final HLS-IA questionnaires.

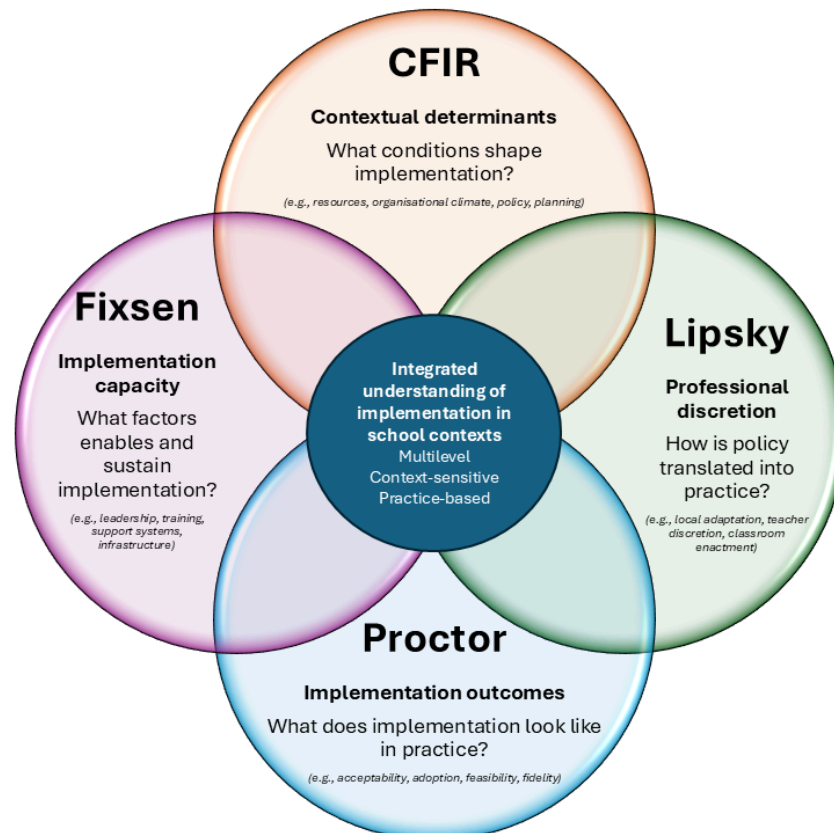


Figure 2

Integrated Conceptual Model of Implementation in School Contexts.

The figure illustrates how four complementary theoretical frameworks (CFIR (25, 26); Fixsen (29, 30); Lipsky (19, 20); Proctor (27)) jointly inform a multilevel, context-sensitive understanding of implementation in school contexts. The overlapping structure reflects their integrative and non-hierarchical use in the present study. Together, these frameworks capture implementation as a dynamic interplay among contextual conditions, organisational capacity, professional practice, and implementation outcomes.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1HLSIAquestionnaires.xlsx](#)
- [Additionalfile2Recruitmentstrategies.docx](#)
- [Additionalfile3Representativityanalysis.xlsx](#)
- [STROBEchecklistcrosssectionalupdated.docx](#)