

RESEARCH ARTICLE

# Measuring Implementation of Health Promoting School (HPS) Programs: Development and Psychometric Evaluation of the HPS Implementation Questionnaire

GERJANNE VENNEGOOR, MSc<sup>a,b</sup>  PATRICIA VAN ASSEMA, PhD<sup>c,d</sup> IRIS EEKHOUT, PhD<sup>e</sup> JEANETTE LEZWIJN, PhD<sup>f</sup> GERARD MOLLEMAN, PhD<sup>g,h</sup>  
MARIA JANSEN, PhD<sup>i,j,k</sup>

## ABSTRACT

**BACKGROUND:** Implementation of Health Promoting School (HPS) programs can be challenging due to the dynamic school context. Navigating between program fidelity and adaptation, as well as integrating the program, is essential for successful implementation, and consequently, for program effects. As part of an evaluation study in the Netherlands, this study aimed to develop a measurement instrument that differentiates schools according to fidelity, adaptation, and integration of HPS implementation.

**METHODS:** This study presents the development and psychometric evaluation of the 28-item HPS Implementation Questionnaire, covering 7 dimensions: adherence, dose, participant responsiveness, quality of delivery, program differentiation, adaptation, and integration. The questionnaire, to be filled out by school employees, was developed for primary, secondary, secondary vocational, and special needs education, in close collaboration with experts ( $n = 54$ ) in school health promotion.

**RESULTS:** Semi-structured interviews aimed at dimension clarification resulted in a list of 58 items. Items were revised, combined, and/or removed based on quantitative and qualitative feedback by the evaluation study's Community of Practice, 2-round expert consultation, and pre-tests. Psychometric evaluation ( $n = 535$  schools), consisting of calculating Cronbach's  $\alpha$  and confirmatory factor analysis (CFA), confirmed internal consistency ( $\alpha > .72$ ) and the 7-dimension framework.

**CONCLUSION:** The brief yet comprehensive HPS Implementation Questionnaire offers possibilities for research into HPS implementation in various educational sectors and contexts, as well as self-monitoring by individual schools. This study provides first evidence for internal consistency and validity of the questionnaire.

**Keywords:** HPS implementation; questionnaire; Health Promoting School; fidelity; adaptation; integration.

**Citation:** Vennegoor G, Van Assema P, Eekhout I, Lezwijn J, Molleman G, Jansen M. Measuring implementation of Health Promoting School (HPS) programs: development and psychometric evaluation of the HPS implementation questionnaire. *J Sch Health*. 2023; 93: 450-463. DOI: 10.1111/josh.13277

Received on June 14, 2022

Accepted on November 2, 2022

Schools have the potential to support students in adopting healthy behaviors at an early age.<sup>1-3</sup> This may lead to improved health and academic achievements, which can track into adulthood, as well as reduced socioeconomic inequity.<sup>4-7</sup> To exploit the full potential of schools in supporting student health behaviors, effective implementation of health promotion programs is crucial.<sup>8</sup> Numerous countries worldwide have adopted principles of the

World Health Organization's Health Promoting School (HPS) framework,<sup>9,10</sup> which includes a whole-school approach focused on healthy school policies, health education in the curriculum, creating a healthy physical and social environment, engagement with the community, and health services.<sup>1,3,11,12</sup> The framework seems promising, as programs based on its principles show some evidence for improving various health outcomes in students (body mass index,

<sup>a</sup>PhD Candidate, (g.vennegoor@maastrichtuniversity.nl), Health Promotion, Maastricht University, Maastricht, The Netherlands.

<sup>b</sup>PhD Candidate, (g.vennegoor@maastrichtuniversity.nl), Academic Collaborative Center for Public Health Limburg, Heerlen, The Netherlands.

<sup>c</sup>Associate Professor, (p.vanassema@maastrichtuniversity.nl), Health Promotion, Maastricht University, Maastricht, The Netherlands.

physical activity, physical fitness, fruit and vegetable intake, tobacco use, and bullying).<sup>1</sup> However, there has been substantial heterogeneity in reported effects, with some programs showing favorable effects and others none at all.<sup>1</sup>

Heterogeneity in effects of HPS programs may be explained by the dynamics of the school setting, which can be considered a complex adaptive system. In such a system, components (eg, composition of student population, task perception of teachers, and management support) are constantly interacting in an attempt to keep a balance.<sup>13-22</sup> As schools are primarily judged on achieving learning objectives rather than health and wellbeing objectives,<sup>23</sup> introducing health-promoting changes, such as an HPS program, often leads to system instability. Through feedback loops between components, the system searches for a new balance by either integrating or rejecting the change.<sup>14,18,20</sup> This is an unpredictable and non-linear process, in which the same implementation efforts can produce different effects, and vice versa.<sup>14,20</sup> Each school has a unique context and, therefore, reacts differently to health promotion changes.<sup>13-20</sup> As a consequence of these dynamics, it is essential to navigate between fidelity of core elements and adaptation of “peripheral” elements while implementing an HPS program.<sup>8,13,22,24-27</sup> Successful implementation also requires integration of the program into the school system’s routines, norms, and identity (or “adding-in”) rather than solely “adding-on” to core curriculum tasks.<sup>28-31</sup>

Evaluation studies can provide insight into the extent to which implementation dimensions are achieved, thereby contributing to a crucial understanding of what works for which schools, and under which conditions.<sup>32,33</sup> In the Netherlands, such a study is conducted into conditions for effectiveness of the national Healthy School Program (“Programma Gezonde School”).<sup>34</sup> This is a whole-school approach consistent with the HPS framework, aiming for uptake of a healthy lifestyle in the DNA of every school.<sup>35</sup> As part of the evaluation, an efficient instrument is needed to conduct a comprehensive measurement

of the degree of implementation of the approach to health and wellbeing among ±2400 schools which do and do not (yet) work with the program.<sup>34</sup>

Looking at current literature on measurement of implementation, a systematic review on HPS programs demonstrated that studies collecting process evaluation data generally focus on fidelity.<sup>1</sup> Based on previous work, fidelity can be conceptualized by 5 dimensions:<sup>36-41</sup> (1) *adherence* is the extent to which program components were delivered according to underlying theoretical guidelines; (2) *dose* reflects the amount of program content received by the target group; (3) *participant responsiveness* is the degree to which the target group is engaged with the program; (4) *quality of delivery* reflects how the program is delivered; and (5) *program differentiation* is the extent to which the program contains unique elements to achieve effective outcomes.<sup>37</sup> These can either be assessed by direct methods which are most accurate but labor intensive (eg, observation or audiotape), or indirect methods which are less time consuming but prone to social desirability bias (eg, logbook or questionnaire).<sup>40</sup> Most studies on HPS programs that were reviewed assess only the dimension of dose by, for instance, a logbook of the number or content of health promotion sessions.<sup>1</sup> Some studies do assess more dimensions, but operationalization varies considerably.<sup>39,42,43</sup> Limited assessment of fidelity dimensions may also be due to available measurement instruments only focusing on adherence and/or dose.<sup>44-46</sup> Apart from fidelity, in recent evaluation studies, the extent to which adaptations were made to the program is increasingly being assessed.<sup>47-49</sup> Nonetheless, data are often collected qualitatively, without a framework or validated instrument. Finally, the importance of the extent to which the program is integrated into the school system is frequently mentioned,<sup>28,31</sup> but rarely included in study design. This indicates a need for a more consistent and comprehensive assessment of HPS implementation with regard to the hypothesized 7-dimensions framework: all 5 dimensions of fidelity, and the dimensions of adaptation and integration.<sup>31,50</sup>

<sup>d</sup>Associate Professor, (p.vanasema@maastrichtuniversity.nl), Academic Collaborative Center for Public Health Limburg, Heerlen, The Netherlands.

<sup>e</sup>Statistician, (iris.eekhout@tno.nl), Expertise Center Child Health, Netherlands Organization for Applied Scientific Research (TNO), Leiden, The Netherlands.

<sup>f</sup>Public Health Adviser, (jeanette.lezwijn@proscop.nl), Public Health Service Noord-en Oost Gelderland, Academic Collaborative Center AGORA, Zutphen, The Netherlands.

<sup>g</sup>Professor, (gerard.molleman@radboudumc.nl), Department of Primary and Community Care, Radboud Institute for Health Sciences, Academic Collaborative Center AMPHI, Nijmegen, The Netherlands.

<sup>h</sup>Professor, (gerard.molleman@radboudumc.nl), Department of Healthy Living, Public Health Service Gelderland-Zuid, Nijmegen, The Netherlands.

<sup>i</sup>Professor, (maria.jansen@ggz.nl), Academic Collaborative Center for Public Health Limburg, Heerlen, The Netherlands.

<sup>j</sup>Professor, (maria.jansen@ggz.nl), Department of Health Services Research, Maastricht University, Maastricht, The Netherlands.

<sup>k</sup>Professor, (maria.jansen@ggz.nl), Public Health Service Zuid Limburg, Heerlen, The Netherlands.

Address correspondence to: Gerjanne Vennegoor, PhD Candidate, (g.vennegoor@maastrichtuniversity.nl), Health Promotion, Maastricht University, Maastricht, The Netherlands.

We would like to thank all 54 experts who participated in one or more phases of the development process of the HPS Implementation Questionnaire for their valuable input. We would also like to thank Marije van Koperen, PhD, and Vivian Kruitwagen, PhD, of the Dutch Healthy School Program for reviewing information on the program in the manuscript. This research is funded by a grant from the Netherlands Organization for Health Research and Development (ZonMw; project number 531001113). The funder had no role in designing the study, data collection, analysis, or interpretation.

We aimed to answer the following research question: Can we develop an efficient yet comprehensive measurement instrument that differentiates schools according to their degree of HPS implementation in terms of fidelity, adaptation to the school context, and integration into the school system?

## METHODS

### Setting

This study was undertaken in the Netherlands and was focused on primary, secondary, secondary vocational, and special needs education. In the Dutch Healthy School Program, schools implement health promotion changes in 4 so-called pillars (ie, education to promote knowledge and skills, physical and social school environment, identification of students who need extra attention or referral, and health promotion as part of school policy) for 1 or more of 10 topics (ie, nutrition, physical activity, wellbeing, relationships and sexuality, smoking, alcohol and drug prevention, environment, prevention of hearing damage, media literacy, sleep, and hygiene; the last 2 were added after data collection for the present article). For example, schools include lessons on healthy nutrition in the curriculum, ensure their canteen offer fits the guidelines of the Nutrition Center, perform periodic measurements on nutrition behavior and weight, and include rules on healthy nutrition in school policy.<sup>51</sup> Support is offered by “Healthy School advisers” of regional Public Health Services (PHSs).<sup>35</sup> In 2020, 19% of all primary, secondary, and secondary vocational schools had obtained a program certificate,<sup>52</sup> which means they implemented changes on all 4 pillars for at least 1 topic. An (unknown) higher percentage of schools works with the program, as certificates are not mandatory and are only available for 5 topics. Apart from this national program,

other whole-school approaches and evidence-based interventions are implemented in some Dutch schools.<sup>35,53,54</sup>

### Mixed Methods

The HPS Implementation Questionnaire was developed in 5 phases between August 2019 and January 2021 (Table 1). Using mixed methods in phases 1 to 4, the questionnaire was checked for face and content validity. Purposive sampling<sup>21</sup> among the network of the evaluation study’s consortium was used to select experts employed in school health promotion, presenting differing sectors (education, health promotion, and health sciences), experience, and regions. This sample consisted of 14 employees of PHSs (of whom 11 Healthy School advisers and 3 policy advisers), 17 school employees (of 3 primary, 8 secondary, and 5 secondary vocational schools, and 1 special needs school), 5 employees of national education and health promotion organizations, and 18 health scientists (54 experts in total; Table 1 details the sample per phase). Subsequently, using quantitative methods in phase 5, the questionnaire was checked for internal consistency and construct validity on data collected among a sample of 535 schools.

### Phase 1: Semi-Structured Interviews

Semi-structured 1-hour audio-recorded face-to-face or (video) call individual interviews with experts (n = 13; Table 1) aimed to define the dimensions of fidelity (adherence, dose, quality of delivery, participant responsiveness, and program differentiation<sup>36-41</sup>), adaptation,<sup>26,55</sup> and integration of school health promotion in the Netherlands. For each of the 7 implementation dimensions, the following questions were asked: “when does a school

Table 1. Overview of the Development of the HPS Implementation Questionnaire

Phase	Sample	Content	Number of Items After Phase
Semi-structured interviews	3 PHS employees 6 school employees 4 health scientists	Qualitative data on dimension clarification to generate items	58
Community of Practice	9 PHS employees 3 members of national education and health promotion organizations 2 health scientists	Qualitative written and verbal feedback on an initial version of the questionnaire	51
Expert consultation	14 PHS employees 6 school employees 18 health scientists	Quantitative and qualitative feedback to reduce the number of items as well as revise items and response categories	41 (round 1) 28 (round 2)
Pre-tests	11 school employees	Qualitative feedback to refine final items	28
Psychometric evaluation	3 members of national education councils 535 school employees (1 for each school)	Quantitative data collection to test for internal consistency and construct validity	28

PHS, Public Health Service.

comply with dimension X?"; "what are aspects of compliance with dimension X?"; and "can you think of items to measure compliance with dimension X?" Selective transcripts<sup>56</sup> were analyzed through thematic analysis<sup>57</sup> using Nvivo 12 Pro software. Data were used to generate a list of suggested items for each of the 7 dimensions. Duplicates were removed, and all items were phrased as questions.

### Phase 2: Community of Practice

An initial version of the HPS Implementation Questionnaire was presented to a group of experts (n = 14, 2 of whom also participated in phase 1; Table 1) during a meeting of the evaluation study's Community of Practice, consisting of a network of professionals aiming for a mutual learning process between research and practice.<sup>34,58</sup> Participants provided 15-minute written reader-focused feedback containing plus-minus ratings as well as open suggestions, followed by a 30-minute audio-recorded group discussion of main issues.<sup>59</sup> All feedback were summarized and incorporated in the draft questionnaire through item revision and improvement of the structure.

### Phase 3: Expert Consultation

A 2-round anonymous online expert consultation (n = 38, including 22 of phases 1 and 2; Table 1), inspired by the Delphi Method,<sup>60</sup> was conducted to reduce the number of items as well as to revise items and response categories. Invitations and reminders were sent by email. Per dimension, participants were asked how well each item measured the dimension on a 5-point scale (response options ranging from very badly to very well), and whether they would change anything (open-ended response). Thereafter, an overview of all items for each dimension was presented, and participants indicated which items they recommended for inclusion, as measured on a 5-point scale (response options ranging from definitely out to definitely in). It was also possible to suggest additional items, and to comment on the draft informed consent form and background items.

Data analysis consisted of calculating medians, Inter Quartile Deviations (IQDs; a measure of consensus<sup>61</sup>), and percentages of respondents scoring  $\geq 4$  ([definitely] keep in). Guidelines for retaining items were: median  $\geq 4.00$ , IQD  $\leq 1.00$ , and  $\geq 70.0\%$  of respondents scoring  $\geq 4$ .<sup>61</sup> Qualitative data were used to revise items, response categories, and questionnaire structure, and to inform the reduction process. The final decision to delete an item was based on the combination of quantitative and qualitative data. The same procedure was applied to the second round. Finally, all items were rephrased into a similar sentence structure.

### Phase 4: Pre-Tests

Thirty-minute interactive (video) call pre-test interviews (n = 14, one of whom also participated in phase 2; Table 1) were aimed at resolving final issues regarding wording, layout, length, or structure of the questionnaire. Participants filled out the questionnaire while the researcher encouraged them to describe their thoughts, following the "think aloud" method.<sup>62</sup> Lastly, participants were asked how well the questionnaire assessed the degree of implementation, in terms of fidelity, adaptation, and integration, at their school. All comments were recorded by the researcher. Qualitative data were reviewed prior to the next pre-test, and subsequent changes were made. The iterative process was continued until no new issues emerged.

### Phase 5: Psychometric Evaluation

Psychometric evaluation was performed on data collected among a sample of schools from various geographical locations between September and December 2020. All schools from the 7 participating PHS regions in the evaluation study (approximately 2400 out of 9000 primary, secondary, secondary vocational, and special needs schools in the country) received an invitation to fill out the online questionnaire. For every school, the invitation was directed to 1 employee who was most knowledgeable about the school's approach to student health and wellbeing (according to the respondent's own judgment). PHS employees distributed a (personalized) email or postcard and reminder to all schools, handed out postcards during regular school visits, invited employees to participate during regular phone calls, added an invitation to their newsletter, and/or posted a message on social media.

Using SPSS 24 software (V.24.0; Armonk, NY, USA), internal consistency was assessed by Cronbach's  $\alpha$  with a reliability threshold of  $\geq 0.7$ .<sup>63</sup> Additionally, the scale's Cronbach's  $\alpha$  if an item were to be removed was evaluated. To assess construct validity for the hypothesized 7-dimension framework, confirmatory factor analysis (CFA) was conducted using the "lavaan" package in R statistical software.<sup>64,65</sup> Three main models were analyzed: model 1: 7 factors (1 for each dimension), model 2: 1 factor (implementation), and model 3: a nested model containing 7 factors and 1 overall factor. Variations to the final model were explored to identify possibilities for improvement. Models were evaluated according to the  $\chi^2$  goodness-of-fit statistic, Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Squared Residual (SRMR). Indicators of acceptable model fit were:  $\chi^2 > .05$ , CFI  $> 0.90$ , TLI  $> 0.90$ , RMSEA  $< 0.08$ , and SRMR  $< 0.08$ .<sup>66-68</sup> To compare the models,  $\chi^2$  difference tests were conducted.

Standardized factor loadings were evaluated by significance and magnitude ( $\geq .30$  was considered acceptable).<sup>69</sup> Significance level was set at  $p < .05$  for all analyses.

After psychometric evaluation, the revised questionnaire was translated from Dutch to English by a bilingual translator.

## RESULTS

### Phase 1: Semi-Structured Interviews

All participants described adherence using the 4 pillars of the Dutch Healthy School program (education, physical and social environment, extra attention or referral if needed, and policy<sup>35</sup>) and recommended items on each pillar to be included. Other suggestions were focused on active coordination by a coordinator or workgroup, and regular evaluation with various parties in the school.

A broadly recognized aspect for measuring dose was the extent of student exposure in all layers of the school system. Communication to students, parents, employees, and general public was considered another key aspect by many participants. Awareness of employees about their exemplar function, frequency of discussions among employees, and policy enforcement were also mentioned by some.

To measure participant responsiveness, all participants advised assessing active involvement, support, and/or ownership among students, parents, management, employees, and external advisers. Some also mentioned the importance of regular activities to increase interaction.

According to the majority of participants, main aspects of quality of delivery were related to knowledge and experience of teachers, external professionals,

and external advisers. Other suggestions included the onboarding of new employees, teacher training, and teacher characteristics (eg, motivation, creativity, and health behavior).

To define program differentiation, many participants recommended including the reason to start working on health and wellbeing, as well as the most significant moments that determined the current situation.

Adaptation of the approach to the school context was considered relevant by all participants for various dimensions, for example, adherence (eg, choosing suitable education materials) and quality of delivery (eg, acquiring relevant expertise). Some participants recommended including a general item on adaptation.

Integration of the approach into the school system was considered essential by the majority of participants. Many participants mentioned that the approach should be automatic, meaning it is executed routinely. Furthermore, many advised including items on alignment with school vision and norms.

Examples of suggested items by participants for each dimension can be found in Table 2. Interview data resulted in an initial version of the questionnaire containing 58 items divided over 7 subscales reflecting the 7-dimension framework (Table 3).

### Phase 2: Community of Practice

Written feedback by Community of Practice members on the initial version of the questionnaire mostly contained practice-oriented suggestions for revising multiple items and response categories, for instance explaining why some schools could not answer the item, or pointing out difficult terms. Subsequent discussion revealed issues concerning matching wording to the respondents' language as well as reducing the number of items. Discussion, comments, as well as

Table 2. Examples of Suggested Items During Semi-Structured Interviews

Dimension	Example(s) of Suggested Items by Participants	Interviewee
Adherence	"Is the school environment designed for a healthy lifestyle?" "Is someone responsible for the progress?" "How, when, and how often do evaluations take place?"	Interviewee 1, health scientist Interviewee 3, health scientist Interviewee 5, PHS employee
Dose	"To what extent are students and teachers confronted with the same message on multiple locations?"	Interviewee 3, health scientist
Participant responsiveness	"To what extent can you hold each other's behavior to account?" "To what extent and in which phase are various parties involved?" "How often do you speak with parents?"	Interviewee 8, school employee Interviewee 10, school employee Interviewee 7, school employee
Quality of delivery	"Is sufficient knowledge available in the school and among external advisers?" "Is it possible to have peer supervision?"	Interviewee 2, PHS employee Interviewee 3, health scientist
Program differentiation	"Looking back, what was the tipping point?" "At what percentage of your school potential is it currently performing?" "Can you adapt your approach to your school's situation?"	Interviewee 6, health scientist Interviewee 7, school employee Interviewee 3, health scientist
Adaptation	"When an opportunity arises, do teachers have a natural reflex to discuss health and wellbeing with students?" "Is it part of the system, the way we work, the way we think?"	Interviewee 7, school employee Interviewee 6, health scientist

PHS, Public Health Service.

Table 3. Quantitative Results of the Expert Consultation

Item*	Round 1										Round 2														
	How well does this item measure dimension X? <sup>†</sup>					Should this item be included? <sup>‡</sup>					How well does this item measure dimension X? <sup>†</sup>					Should this item be included? <sup>‡</sup>									
	N	Mdn	Q1	Q3	IQD	% ≥4	N	Mdn	Q1	Q3	IQD	% ≥4	N	Mdn	Q1	Q3	IQD	% ≥4	N	Mdn	Q1	Q3	IQD	% ≥4	
Subscale adherence																									
Q1 Use of sources	24	4.00	3.00	4.00	0.50	54.2	24	4.00	3.25	5.00	0.88	75.0	19	4.00	3.00	5.00	1.00	73.7	17	4.00	4.00	4.50	0.25	88.2	
Q2 Long-term vision	24	4.00	3.00	4.00	0.50	54.2	24	4.00	3.25	5.00	0.88	75.0	19	4.00	3.00	5.00	1.00	73.7	17	4.00	4.00	4.50	0.25	88.2	
Q3 Mapping urgent problems	24	4.00	3.00	4.00	0.50	66.7	24	4.00	3.00	4.00	0.50	70.8	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	5.00	0.50	82.4	
Q4 School policy in school plan	24	4.00	3.25	4.00	0.38	75.0	24	4.00	4.00	5.00	0.50	79.2	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	5.00	0.50	88.2	
Q5 Employees responsible	24	3.00	2.00	3.75	0.88	25.0	24	3.00	2.00	4.00	1.00	29.2	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q6 Atmosphere during evaluation	24	3.00	2.00	4.00	1.00	41.7	24	3.00	2.25	4.00	0.88	33.3	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q7 Positive framing	24	3.50	2.25	4.00	0.88	50.0	24	4.00	2.00	4.00	1.00	54.2	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q8 Coherence of the approach	24	4.00	3.00	4.00	0.50	70.8	24	4.00	3.00	4.00	0.50	70.8	19	4.00	3.00	4.00	0.50	70.8	17	4.00	4.00	4.00	0.00	0.00	
Q9 Care plan available	24	4.00	3.25	4.00	0.38	75.0	24	4.00	4.00	5.00	0.50	79.2	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q10 Action plan available	24	4.00	3.00	4.00	0.50	70.8	24	4.00	3.25	4.75	0.75	75.0	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q11 School policy on website	24	4.00	4.00	4.00	0.00	91.7	24	4.00	3.25	4.00	0.38	75.0	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q12 One or more workgroups present	24	4.00	3.00	4.00	0.50	70.8	24	4.00	3.00	4.00	0.50	66.7	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q13 Recurrent measurement of student health	24	4.00	4.00	4.00	0.00	91.7	24	4.00	4.00	4.75	0.38	83.3	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q14 School environment	24	4.00	3.00	4.00	0.50	70.8	24	4.00	4.00	4.00	0.00	79.2	19	4.00	3.00	4.00	0.50	73.7	17	4.00	4.00	4.00	0.00	0.00	
Q15 Use of existing teaching materials	24	4.00	3.00	4.00	0.50	70.8	24	4.00	4.00	4.00	0.00	79.2	19	4.00	3.00	4.00	0.50	68.4	17	4.00	4.00	4.00	0.00	0.00	
Q16 Employees are able to identify problems	24	4.00	2.25	4.00	0.88	54.2	24	4.00	3.25	4.00	0.38	75.0	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q17 Rules of conduct available	24	4.00	3.25	4.00	0.38	75.0	24	4.00	4.00	4.00	0.00	79.2	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q18 Part of school policy	24	4.00	3.25	4.00	0.38	75.0	24	4.00	4.00	4.00	0.00	79.2	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q19 One or more coordinators	24	4.00	3.25	4.00	0.38	75.0	24	4.00	4.00	4.75	0.38	91.7	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q20 Sufficient hours available	24	4.00	4.00	4.00	0.00	87.5	24	4.00	4.00	5.00	0.50	79.2	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	
Q21 Sufficient budget available	24	4.00	4.00	4.00	0.00	79.2	24	4.00	3.00	5.00	1.00	70.8	19	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.00	0.00	

Subscale adherence

Q1 Use of sources ●

Q2 Long-term vision ●

Q3 Mapping urgent problems ■

Q4 School policy in school plan ■

Q5 Employees responsible ■

Q6 Atmosphere during evaluation ■

Q7 Positive framing ■

Q8 Coherence of the approach ■

Q9 Care plan available ▲

Q10 Action plan available ▲

Q11 School policy on website ▲

Q12 One or more workgroups present ▲

Q13 Recurrent measurement of student health

Q14 School environment

Q15 Use of existing teaching materials

Q16 Employees are able to identify problems

Q17 Rules of conduct available

Q18 Part of school policy

Q19 One or more coordinators

Q20 Sufficient hours available

Q21 Sufficient budget available

Table 3. Continued

Item*	Round 1										Round 2													
	How well does this item measure dimension X? <sup>†</sup>					Should this item be included? <sup>‡</sup>					How well does this item measure dimension X? <sup>†</sup>					Should this item be included? <sup>‡</sup>								
	N	Mdn	Q1	Q3	IQD	% ≥ 4	N	Mdn	Q1	Q3	IQD	% ≥ 4	N	Mdn	Q1	Q3	IQD	% ≥ 4	N	Mdn	Q1	Q3	IQD	% ≥ 4
Q22 Annual evaluation	24	4.00	3.00	4.00	0.50	70.8	24	4.00	4.00	4.75	0.38	83.3	17	4.00	4.00	4.00	0.00	0.00	17	5.00	4.00	5.00	0.50	94.1
Q23 Making good use of feedback	24	4.00	4.00	4.00	0.00	79.2	24	4.00	4.00	4.00	0.00	83.3	17	4.00	4.00	5.00	0.00	0.00	17	4.00	4.00	5.00	0.50	88.2
Q24 Use of educational activities	24	4.00	3.00	4.00	0.50	66.7	24	4.00	4.00	4.75	0.38	79.2	17	4.00	3.00	4.00	0.88	58.8	17	4.00	3.50	5.00	0.75	76.5
Q26 Awareness among employees	24	4.00	3.25	4.00	0.38	75.0	24	4.00	3.25	4.75	0.75	75.0	17	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.75	76.5
Q27 Activities aimed at employees	24	4.00	4.00	4.00	0.00	83.3	24	4.00	4.00	5.00	0.50	79.2	17	4.00	4.00	4.00	0.00	0.00	17	4.00	4.00	4.00	0.75	76.5
Q28 Approach is visible	24	4.00	3.00	4.00	0.50	58.3	24	3.50	2.00	4.00	1.00	50.0	17	4.00	3.00	4.50	0.88	58.8	17	4.00	3.50	5.00	0.75	76.5
Q29 Coherent communication	24	4.00	4.00	4.00	0.00	79.2	24	4.00	3.25	5.00	0.88	75.0	17	4.00	4.00	5.00	0.50	100.0	17	4.00	4.00	5.00	0.50	82.4
Q30 Regular agenda item—students	24	4.00	3.00	4.00	0.50	62.5	24	4.00	3.00	4.00	0.50	66.7	17	4.00	3.50	4.00	0.25	76.5	17	4.00	4.00	5.00	0.50	100.0
Q31 Regular agenda item—parents	24	4.00	3.25	4.00	0.38	75.0	24	4.00	3.00	4.00	0.50	62.5	17	4.00	4.00	5.00	0.50	88.2	17	4.00	4.00	5.00	0.50	94.1
Q32 Regular agenda item—external parties	24	4.00	3.00	4.00	0.00	79.2	24	4.00	3.00	4.75	0.88	70.8	17	4.00	4.00	4.50	0.25	100.1	17	4.00	4.00	5.00	0.50	88.2
Q33 Regular agenda item—employees	24	4.00	4.00	4.00	0.00	79.2	24	4.00	3.25	5.00	0.88	75.0	17	4.00	4.00	5.00	0.50	100.0	17	4.00	4.00	5.00	0.50	82.4
Q34 Reach of (almost) all students	24	4.00	3.00	4.00	0.50	58.3	24	4.00	3.25	4.75	0.75	75.0	17	4.00	4.00	4.50	0.25	82.4	17	4.00	4.00	5.00	0.50	94.1
Q35 Employees give a good example	24	4.00	3.00	4.00	0.50	70.8	24	4.00	3.00	4.75	0.88	70.8	17	4.00	4.00	4.50	0.25	82.4	17	4.00	4.00	5.00	0.50	82.4
Q36 Employees adhere to code of conduct	24	4.00	3.00	4.00	0.50	62.5	24	4.00	3.00	4.00	0.50	66.7	17	4.00	3.50	4.00	0.25	76.5	17	4.00	4.00	5.00	0.50	100.0
Q37 Active communication to students, employees, parents	24	4.00	4.00	4.00	0.00	79.2	24	4.00	4.00	5.00	0.50	91.7	17	4.00	4.00	5.00	0.50	88.2	17	4.00	4.00	5.00	0.50	94.1
Q38 Frequency of contact	23	4.00	3.00	4.00	0.50	73.9	23	4.00	3.00	5.00	1.00	73.9	17	4.00	4.00	4.50	0.50	100.1	17	4.00	4.00	5.00	0.50	88.2
Q39 Perceived support	23	4.00	4.00	5.00	0.50	91.3	23	4.00	4.00	5.00	0.50	95.7	17	4.00	4.00	4.50	0.25	82.4	17	5.00	4.00	5.00	0.50	94.1

Table 3. Continued

Item*	Round 1										Round 2													
	How well does this item measure dimension X?†					Should this item be included?‡					How well does this item measure dimension X?†					Should this item be included?‡								
	N	Mdn	Q1	Q3	IQD	% ≥4	N	Mdn	Q1	Q3	IQD	% ≥4	N	Mdn	Q1	Q3	IQD	% ≥4	N	Mdn	Q1	Q3	IQD	% ≥4
Q40 Active involvement	23	4.00	4.00	5.00	0.50	78.3	23	4.00	4.00	5.00	0.50	82.6	17	4.00	4.00	5.00	0.50	88.2	17	5.00	4.00	5.00	0.50	94.1
Subscale quality of delivery																								
Q41 Agreements on task division ●	24	3.50	3.00	4.00	0.50	50.0	24	3.00	2.00	4.00	1.00	33.3												
Q42 Part of annual employee performance appraisals ■	24	4.00	2.00	4.00	1.00	54.2	24	4.00	3.00	4.00	0.50	58.3												
Q43 External supporters provide good information ■	24	3.00	3.00	4.00	0.50	45.8	24	3.50	3.00	4.00	0.50	50.0												
Q44 Teachers can exchange experiences ▲	24	3.00	2.25	4.00	0.88	41.7	24	3.00	2.00	4.00	1.00	33.3												
Q45 Enthusiasm of employees ▲	24	4.00	4.00	4.75	0.38	83.3	24	4.00	4.00	5.00	0.50	83.3												
Q46 Number of partner organizations ▲	24	4.00	3.00	4.00	0.50	66.7	24	4.00	3.00	4.00	0.50	66.7												
Q47 Expertise of teachers	24	4.00	3.00	4.00	0.50	58.3	24	4.00	3.00	4.00	0.50	62.5												
Q48 Expertise of external professionals	24	4.00	4.00	5.00	0.50	91.7	24	4.00	4.00	5.00	0.50	91.7												
Q49 Notifying new employees	24	4.00	3.00	4.00	0.50	62.5	24	4.00	3.00	4.00	0.50	70.8												
Q50 Regular contact with external supporter(s)	24	4.00	3.00	4.00	0.50	58.3	24	4.00	3.00	4.00	0.50	62.5												
Q51 Sufficient ownership present	24	4.00	3.00	4.00	0.50	58.3	24	4.00	3.00	4.00	0.50	62.5												
Subscale program differentiation																								
Q52 Unique elements	23	4.00	2.00	4.00	1.00	52.2	23	3.00	3.00	4.00	0.50	36.0												
Q53 Adaptation to school context	24	3.00	2.00	4.00	1.00	29.2	24	3.00	3.00	4.00	0.50	36.0												
Q54 Part of nonm	25	3.00	2.00	4.00	1.00	40.0	25	3.00	3.00	4.00	0.50	36.0												
Q55 Automatic inclusion in new activities ■	25	4.00	3.00	4.00	0.50	60.0	25	4.00	3.00	5.00	1.00	72.0												
Q56 In line with school vision	25	4.00	3.50	4.00	0.25	76.0	25	4.00	4.00	5.00	0.50	84.0												
Q56 In line with school vision	25	4.00	3.50	4.00	0.25	76.0	25	4.00	4.00	5.00	0.50	84.0												



Table 3. Continued

Item*	Round 1										Round 2													
	How well does this item measure dimension X? <sup>†</sup>					Should this item be included? <sup>‡</sup>					How well does this item measure dimension X? <sup>†</sup>					Should this item be included? <sup>‡</sup>								
	N	Mdn	Q1	Q3	IQD	% ≥ 4	N	Mdn	Q1	Q3	IQD	% ≥ 4	N	Mdn	Q1	Q3	IQD	% ≥ 4	N	Mdn	Q1	Q3	IQD	% ≥ 4
<i>Q57 Healthy choices are made easy</i>	25	4.00	3.50	4.00	0.25	76.0	25	4.00	3.00	5.00	1.00	68.0	17	4.00	4.00	4.50	0.25	88.2	17	4.00	4.00	5.00	0.50	88.2
<i>Q58 Self-evident character</i>	25	4.00	3.00	4.00	0.50	64.0	25	4.00	3.00	5.00	1.00	60.0	17	4.00	3.50	4.00	0.25	76.5	17	4.00	3.00	5.00	1.00	70.6
Average		3.90	3.10	4.10	0.50	66.8		3.90	3.30	4.50	0.60	70.1		4.00	3.80	4.40	0.30	83.7		4.20	3.80	4.90	0.60	86.3

Mdn, Median; Q1, 25th Percentile; Q3, 75th Percentile; IQD, Interquartile Deviation; % ≥ 4, Percentage of Participants Scoring ≥ 4.

Items in italics are part of the final questionnaire. ● Deleted before round 1 (based on community feedback in phase 2), ■ Deleted after round 1, ▲ Deleted after round 2, ◆ Added after round 2.

\*Summaries of each item.

<sup>†</sup>Measured on a 5-point scale (1 = very badly; 5 = very well).

<sup>‡</sup>Measured on a 5-point scale (1 = definitely out; 5 = definitely in).

<sup>§</sup>Split up into 4 items in round 2.

<sup>||</sup>Split up into 2 items in round 1.

<sup>¶</sup>Combined item in round 1.

<sup>#</sup>Data are missing.

plus-minus ratings were informative for combining, regrouping, and/or removing some items. A revised version of the questionnaire, containing 51 items, was entered into phase 3 (Table 3).

### Phase 3: Expert Consultation

In total, 25 of the 38 invited experts completed round 1 of the expert consultation (response rate overall: 66%; PHS employees: 100%; school employees: 67%; health scientists: 39%). In the second round, 23 experts participated (response rate overall: 61%; PHS employees: 86%; school employees: 50%; health scientists: 44%). Of those who responded, 2 (8%) provided partial answers in round 1, and 6 (26%) in round 2.

Qualitative data contained 666 comments in the first, and 340 in the second round. In round 1, the majority of comments were suggestions for rephrasing the item or response categories (in whole or in part) to better suit (the language of) the target group. Comments for example concerned differences between primary and secondary education, and difficult terms such as *positive framing*, *norms*, and *coherence*. Participants recommended providing a more detailed formulation of items, and examples, as some items were considered unclear. Additional suggestions led to inclusion of 2 new items. In round 2, comments were focused on the order of items, consistent item structure and response categories, and minor rephrasing. Participants also frequently advised on which items could be combined or removed. One item was added after this round.

Table 3 presents an overview of all quantitative data. In rounds 1 and 2 respectively, favorable medians were reached on both questions for 82.4% and 100.0% of the items, IQDs for 100.0% and 100.0%, and percentage of participants scoring ≥ 4.00 for 49.0% and 90.2%. On average, rephrased items in round 2 had more favorable scores on all criteria.

Using both quantitative and qualitative data, it was decided to delete 14 items after the first, and 12 after the second round. Some items scored particularly low and this was confirmed by qualitative data (eg, Q6, Q29, and Q54). A few items with lower scores were retained, as suggestions were made for major rephrasing (eg, Q44, Q45, and Q58). For other items, participants advised removing them, although scores were average (eg, Q5, Q27, and Q32). Finally, duplicate items were removed (eg, Q3, Q28, and Q38) or combined (eg, Q4, Q12, and Q46). The 28 items that were retained were rephrased based on qualitative data.

### Phase 4: Pre-Tests

Pre-test sessions led to minor modifications in wording of the 28 items and response categories, and adding supplementary explanation for difficult terms

(eg, for *support staff* and *ownership*). Furthermore, a few response categories were added (eg, *not applicable*) and the order of items was altered (eg, more difficult items were moved down). Minor changes were made in the functionality of the online questionnaire (eg, access to information buttons). All participants confirmed proper assessment of the degree of implementation at their school by the questionnaire. Quantitative data were consistent with comments concerning the degree of implementation in the school. All participants considered the questionnaire to be the right length (response time varied between 10 and 15 minutes).

### Phase 5: Psychometric Evaluation

The sample for psychometric evaluation consisted of 535 schools (22% of ±2400 invited schools), divided over 365 primary schools (22% of those invited), 102 secondary schools (28%), 25 secondary vocational schools (13%), and 43 special needs schools (28%). Respondents of 117 schools (22%) provided partial responses, including 85 (16%) who stopped answering at the verification question on knowledgeability about the school's approach toward health and wellbeing. Analyses were thus performed on 450 schools.

CFA showed low factor loading (standardized  $\beta < .3$ ) in all 3 models for Q40, sub-item *preventive youth health care*. Upon removal, goodness-of-fit indices improved for all models and Cronbach's  $\alpha$  of both the subscale participant responsiveness and full scale increased. For the 3 revised models,  $\chi^2$  tests showed no perfect model fit (Table 4). The most favorable results were shown for model 1 (7 factors) on all fit measures, and the least favorable for model 2 (1 factor). While acceptable absolute fit indices (RMSEA and SRMR) were reached for model 1 and partly for model 3 (nested), no acceptable comparative fit indices (CFI and TLI) were reached (Table 4). Model 1 had a significantly better fit than model 2 ( $\Delta\chi^2$  (19) = 550.02,  $p < .001$ ) and model 3 ( $\Delta\chi^2$  (14) = 151.62,  $p < .001$ ). Additionally, model 3 fit better than model 2 ( $\Delta\chi^2$  (5) = 398.40,  $p < .001$ ). Modification indices did not indicate considerable improvements to model fit. Model 1 was chosen as the final model. Factor loadings for model 1 of topic scores *drug use*, *relationships & sexuality*, *(indoor) environment*, and *media use* were below the threshold and/or not significant. All other

Table 5. Internal Consistency of Subscales and Full Scale

Scale	Number of (sub) items	Cronbach's $\alpha$
Subscale adherence	13	0.74
Subscale dose	7	0.89
Subscale participant responsiveness	9	0.84
Subscale quality of delivery	5	0.74
Subscale program differentiation	1	*
Subscale adaptation	1	*
Subscale integration	3	0.73
Full scale	39	0.90

\*Subscale consists of 1 item.

factor loadings were positive and significant (see Table S1). Variations to model 1 (excluding integration and/or adaptation, and excluding program differentiation and adaptation [and integration]) decreased model fit.

Finally, Cronbach's  $\alpha$ 's were calculated for the revised scale. All subscales ( $\alpha > .72$ ) as well as the full scale ( $\alpha = .90$ ) were internally consistent (Table 5). Reliability of subscales could be improved by deleting topic score *drug use* ( $\alpha = .75$ ), Q40 sub-item *management* ( $\alpha = .85$ ), and Q50 ( $\alpha = .75$ ). The full scale could be slightly improved ( $\leq .01$ ) by removing topic scores *drug use*, *relationships & sexuality*, and *media use*. Although results of psychometric evaluation for these 3 topics would imply removing them, they were retained to include all 8 topics of the Dutch Healthy School Program. Q40 sub-item *management* and Q50 were also retained, as CFA showed acceptable factor loadings.

### DISCUSSION

This study aimed to develop an instrument which differentiates schools in the degree of HPS implementation. Together with a large group of experts in school health promotion, a 28-item questionnaire was developed in a 5-phase process of item generation, refinement, and reduction. This practice-oriented phased approach was essential to enable a suitable instrument to be developed. The final questionnaire offers an extensive yet concise assessment covering 7 dimensions (adherence, dose, participant responsiveness, quality of delivery, program differentiation,

Table 4. Goodness-of-Fit Indices of the Tested Models

	$\chi^2$	df	$\chi^2/df$	CFI	TLI	RMSEA	SRMR
Model 1	2638.789*	683	3.86	0.743	0.721	0.080	0.075
Model 2	3188.807*	702	4.54	0.673	0.655	0.089	0.080
Model 3	2790.404*	697	4.00	0.725	0.708	0.082	0.077

\* $p < .001$ .

n = 450 schools, model 1 = 7 factors (1 for each dimension), model 2 = 1 factor (implementation), model 3 = nested model containing 7 factors and 1 overall factor.

df, Degrees of Freedom; CFI, Comparative Fit Index; TLI, Tucker-Lewis Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Squared Residual.

adaptation, and integration), each measured by 1 to 12 items.

Using data collected from a large sample of schools, both the subscales and full scale were shown to be internally consistent. Additionally, CFA confirmed the hypothesized 7-dimension framework. One sub-item, concerning active involvement of preventive youth health care, was removed based on psychometric evaluation, since it did not explain the factor participant responsiveness nor the overall factor implementation. This is an interesting finding in the Dutch context, where involvement of preventive youth health care in school health promotion is considered essential, but is not always firmly in place.<sup>70-72</sup>

### Fidelity, Adaptation, and Integration

The present study showed that it is possible to combine an evaluation of the 5 dimensions of fidelity, and the 2 dimensions of adaptation and integration into 1 instrument. Even though all dimensions were successfully operationalized, in line with previous literature,<sup>37,39</sup> generating and refining the item to measure program differentiation proved to be difficult. For adaptation, development was also challenging, resulting in 1 general item. Interestingly, analyses on the 7-factor model indicated that both single-item factors did improve model fit. Moreover, in the development process, participants frequently mentioned that items belonging to other dimensions could also provide an indication of the level and type of adaptations.

Adding factor integration to the model also enhanced fit. In the present study, it was defined as “the extent to which the approach is part of school routines, norms, and identity.” Previous studies have suggested similar dimensions, mainly as an indicator of program sustainability.<sup>8,73</sup> While integration may be conditional for sustaining the program, the authors argue that it is more related to the extent to which the program is “added-in” to the school system (rather than “added-on”).<sup>29,47</sup> Further research is needed to accumulate evidence for the relation with other dimensions, and indicators of integration.

Based on the results of this study, it is recommended to include assessment of fidelity, adaptation, and integration in future studies on HPS implementation.

### Applications of the HPS Implementation Questionnaire

The HPS Implementation Questionnaire provides an opportunity to obtain more insight into the degree of implementation of HPS programs—a key item on the research agenda.<sup>31,50</sup> This may help prevent type III errors and reduce heterogeneity in effects.<sup>1,8,37,74</sup> Given the estimated response time of 10 to 15 minutes, the measure is an interesting alternative to existing

options focusing only on adherence and/or dose.<sup>44,45</sup> It provides a comprehensive way of assessing many dimensions, preventing choosing only some due to, for instance, time constraints.<sup>32,43</sup>

The questionnaire is suitable for data collection (both online and offline) among multiple schools, aiming to differentiate between schools with varying degrees of implementation (eg, in a district, municipality, or country), possibly combined with an investigation of contextual factors to explore what works for which schools, and under which conditions.<sup>34</sup> Secondly, it could be applied as a self-monitoring tool for individual schools to assess the current situation and identify points of improvement. However, more research is needed to determine reference values.

Pre-tests ensured suitability for Dutch primary, secondary, secondary vocational, and special needs schools which do and do not (yet) work with the national Healthy School Program. The questionnaire was translated into the English language to enable its use in other countries. Although it was developed in and for the Dutch context, the design of the national program is to a great extent consistent with the international HPS framework and questions do not focus on the Dutch educational system. It is, therefore, expected to be suitable for evaluating other HPS programs in other countries as well. However, item invariance was not assessed, so it is recommended to conduct a linguistic and cross-cultural evaluation before using the questionnaire in other contexts and/or sectors (eg, child care or universities).<sup>75</sup> Minor wording changes may be needed to ensure optimal fit (eg, addition/removal of health promotion topics or changing context- or sector-specific terms). The authors consider these changes acceptable as they do not compromise the validity and psychometric properties as established in the current study.

### Strengths and Limitations

A major strength of the present study is the involvement of a large sample of practical and scientific health promotion experts in the development process. Therefore, all relevant aspects of implementation were most likely covered. Moreover, the initial list of generated items was refined in several phases using different methods. This allowed for optimal tailoring to the respondent group. Another strength is the assessment of psychometric properties, which is often lacking in similar studies.<sup>43</sup>

Regarding the first 4 phases of the development process, the main limitation concerns the use of a purposive sample of experts. However, in each phase, participants from all educational sectors, various geographical locations and levels of experience were included to ensure representativeness. Impact of Covid on the development process is limited, as phases 1

and 2 were conducted before the first wave in the Netherlands, phase 3 at the end of the first, and phase 4 between the first and second. Regarding psychometric evaluation, the response rate among invited schools was lower compared to other studies focusing on high school principals<sup>76,77</sup> and primary and secondary school- and workplace canteen staff.<sup>78</sup> This is possibly due to limited time as a result of the second Covid wave in the Netherlands. Furthermore, respondents judged themselves whether they were most knowledgeable about the school's approach toward student health and wellbeing. Inter-rater reliability was not assessed, and it is likely that differences in answers would exist between employees of 1 school based on their level of involvement. Therefore, it is recommended to carefully plan a strategy for data collection in order to reach the right respondent group. Future research is needed to investigate whether assessment by employees with a different level of involvement, and by multiple employees with the same level of involvement, actually leads to different results. Moreover, the questionnaire is a self-reported instrument, which is prone to social desirability bias. Although bias is expected to be limited as questions are answered on the school level, it may still lead to an overestimation of the degree of implementation. Additionally, as the final CFA model showed no optimal fit indices, it is recommended for future studies to repeat this analysis in order to verify that these results hold. Finally, not all relevant types of validity (eg, discriminative and convergent) were assessed in the present study, but these will be examined in follow-up studies.

## IMPLICATIONS FOR SCHOOL HEALTH

The HPS Implementation Questionnaire assesses schools on the fidelity, adaptation, and integration of their approach to student health and wellbeing. The brief yet comprehensive questionnaire offers possibilities for crucial research into HPS implementation in various educational sectors and contexts. This study provides preliminary evidence for the internal consistency and validity of the questionnaire. Future research is needed to confirm and extend these results.

## Human Subjects Approval Statement

Written informed consent was obtained prior to participation in any phase of the study. Ethical clearance was provided by the Ethics Review Committee of the Faculty of Health, Medicine, and Life Sciences of Maastricht University (FHML-REC/2020/082.01).

## Conflict of Interest

The authors declare that they have no competing interests.

## REFERENCES

- Langford R, Bonell CP, Jones HE, et al. The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. *Cochrane Database Syst Rev.* 2014;4(4):CD008958.
- Bonell C, Parry W, Wells H, et al. The effects of the school environment on student health: a systematic review of multi-level studies. *Health Place.* 2013;21:180-191.
- Lewallen TC, Hunt H, Potts-Datema W, Zaza S, Giles W. The whole school, whole community, whole child model: a new approach for improving educational attainment and healthy development for students. *J Sch Health.* 2015;85(11):729-739.
- Suhrcke M, De Paz Nieves C. *The Impact of Health and Health Behaviours on Educational Outcomes in High-Income Countries: A Review of the Evidence.* Copenhagen: World Health Organization, Regional Office for Europe; 2011.
- Fiscella K, Kitzman H. Disparities in academic achievement and health: the intersection of child education and health policy. *Pediatrics.* 2009;123(3):1073-1080.
- Conti G, Heckman J, Urzua S. The education-health gradient. *Am Econ Rev.* 2010;100(2):234-238.
- Marmot MAJ, Goldblatt P, Boyce T, McNeish D, Grady M, Geddes I. *Fair Society, Healthy Lives.* London, UK: Institute of Health Equity; 2010.
- Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol.* 2008;41(3-4):327-350.
- World Health Organization and the United Nations Educational, Scientific and Cultural Organization. *Making every School a Health-Promoting School: Global Standards and Indicators for Health-Promoting Schools and Systems. Country Case Studies.* Geneva: World Health Organization and the United Nations Educational, Scientific and Cultural Organization; 2021.
- Vilaça T, Darlington E, Rosário R, et al. *SHE Mapping Report. Lessons Learnt from Policies and Practices of SHE Member Countries.* Haderslev: Schools for Health in Europe Network Foundation (SHE); 2019.
- World Health Organization. *Ottawa Charter for Health Promotion.* Copenhagen: Regional Office for Europe; 1986.
- World Health Organization and the United Nations Educational, Scientific and Cultural Organization. *Making every School a Health-Promoting School: Global Standards and Indicators for Health-Promoting Schools and Systems.* Geneva: World Health Organization and the United Nations Educational, Scientific and Cultural Organization; 2021.
- Darlington EJ, Violon N, Jourdan D. Implementation of health promotion programmes in schools: an approach to understand the influence of contextual factors on the process? *BMC Public Health.* 2018;18(1):163.
- Egan M, McGill E, Anderson de Cuevas R, et al. *NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 1: Introducing Systems Thinking.* London: National Institute for Health Research School for Public Health Research; 2019.
- Plsek PE, Greenhalgh T. The challenge of complexity in health care. *BMJ (Clinic Res ed).* 2001;323(7313):625-628.
- Rosas SR. Systems thinking and complexity: considerations for Health Promoting Schools. *Health Promot Int.* 2015;32(2):301-311.
- Keshavarz N, Nutbeam D, Rowling L, Khavarpour F. Schools as social complex adaptive systems: a new way to understand the challenges of introducing the Health Promoting Schools concept. *Soc Sci Med.* 2010;70(10):1467-1474.
- Dooley KJ. A complex adaptive systems model of organization change. *Nonlinear Dyn Psychol Life Sci.* 1997;1(1):69-97.
- Turunen H, Sormunen M, Jourdan D, von Seelen J, Buijs G. Health Promoting Schools—a complex approach and

- a major means to health improvement. *Health Promot Int*. 2017;32(2):177-184.
20. Naaldenberg J, Vaandrager L, Koelen M, Wagemakers A-M, Saan H, de Hoog K. Elaborating on systems thinking in health promotion practice. *Glob Health Promot*. 2009;16(1):39-47.
  21. Patton MQ. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. Los Angeles, CA: SAGE Publications; 2014.
  22. Bartelink N, Bessems K. Health Promoting Schools in Europe. In: *State of the Art*. Maastricht: Schools for Health in Europe; 2019.
  23. World Health Organization. Promoting health through schools: report of a WHO expert committee on comprehensive school health education and promotion. *World Health Organ Tech Rep Ser*. 1997;870:i-93.
  24. Kremser W. Phases of school health promotion implementation through the lens of complexity theory: lessons learnt from an Austrian case study. *Health Promot Int*. 2010;26(2):136-147.
  25. Dusenbury L, Brannigan R, Hansen WB, Walsh J, Falco M. Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. *Health Educ Res*. 2005;20(3):308-313.
  26. Baumann AA, Cabassa LJ, Stirman SW. Adaptation in dissemination and implementation science. In: *Dissemination and Implementation Research in Health: Translating Science to Practice*. 2nd ed. Oxford, UK: Oxford University Press; 2017.
  27. Patton MQ. Developmental evaluation. In: *Applying Complexity Concepts to Enhance Innovation and Use*. New York, NY: Guilford; 2011.
  28. Inchley J, Muldoon J, Currie C. Becoming a Health Promoting School: evaluating the process of effective implementation in Scotland. *Health Promot Int*. 2007;22(1):65-71.
  29. Bentsen P, Bonde AH, Schneller MB, Danielsen D, Bruselius-Jensen M, Aagaard-Hansen J. Danish 'add-in' school-based health promotion: integrating health in curriculum time. *Health Promot Int*. 2020;35(1):e70-e77.
  30. Macnab AJ, Gagnon FA, Stewart D. Health Promoting Schools: consensus, strategies and potential. *Health Educ*. 2014;114(3):170-185.
  31. Gugglberger L. A brief overview of a wide framework—Health Promoting Schools: a curated collection. *Health Promot Int*. 2021;36(2):297-302.
  32. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350:h1258.
  33. Moore GF, Evans RE, Hawkins J, et al. From complex social interventions to interventions in complex social systems: future directions and unresolved questions for intervention development and evaluation. *Evaluation*. 2018;25(1):23-45.
  34. Vennegoor G, Vonk L, van Assema P, et al. Onderzoeksopzet evaluatie Gezonde School: onder welke condities doet het ertoe? *Tijdschr Gezondheidsw*. 2020;98(2):25-33.
  35. van Koperen M, Kruitwagen V, Westhuis A, Sobels M. Gezonde School: landelijke samenwerking voor een gezonde jeugd. *Tijdschr Gezondheidsw*. 2020;98(2):62-65.
  36. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin Psychol Rev*. 1998;18(1):23-45.
  37. Dusenbury L, Brannigan R, Falco M, Hansen WB. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res*. 2003;18(2):237-256.
  38. Mihalic S. The importance of implementation fidelity. *Emot Behav Disord Youth*. 2004;4(4):83-105.
  39. Schaap R, Bessems K, Otten R, Kremers S, van Nassau F. Measuring implementation fidelity of school-based obesity prevention programmes: a systematic review. *Int J Behav Nutr Phys Act*. 2018;15(1):75.
  40. Allen JD, Shelton RC, Emmons KM, Linnan LA. Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination. In: *Dissemination and Implementation Research in Health: Translating Science to Practice*. 2nd ed. Oxford, UK: Oxford University Press; 2017.
  41. Linnan L, Steckler A. *Process Evaluation for Public Health Interventions and Research*. San Fransisco, CA: Jossey-Bass; 2002.
  42. Lambert JD, Greaves CJ, Farrand P, Cross R, Haase AM, Taylor AH. Assessment of fidelity in individual level behaviour change interventions promoting physical activity among adults: a systematic review. *BMC Public Health*. 2017;17(1):765.
  43. Toomey E, Hardeman W, Hankonen N, et al. Focusing on fidelity: narrative review and recommendations for improving intervention fidelity within trials of health behaviour change interventions. *Health Psychol Behav Med*. 2020;8(1):132-151.
  44. Safarjan E, Buijs G, de Ruiter S. *SHE Rapid Assessment Tool: A Companion Document for the SHE Online School Manual*. Utrecht: CBO; 2013.
  45. Dix KL, Green MJ, Tzoumakis S, et al. The Survey of School Promotion of Emotional and Social Health (SSPESH): a brief measure of the implementation of whole-school mental health promotion. *Sch Ment Health*. 2018;11:294-308.
  46. Dadaczynski K, Hering T. Health Promoting Schools in Germany. Mapping the implementation of holistic strategies to tackle NCDs and promote health. *Int J Environ Res Public Health*. 2021;18(5):2623.
  47. Bartelink NHM, van Assema P, Jansen MWJ, et al. Process evaluation of the healthy primary school of the future: the key learning points. *BMC Public Health*. 2019;19(1):698.
  48. Driessen-Willems MD, Bartelink NHM, Bessems KMHH, Kremers SPJ, Kintzen C, van Assema P. Co-creation approach with action-oriented research methods to strengthen "Krachtvoer"; a school-based programme to enhance healthy nutrition in adolescents. *Int J Environ Res Public Health*. 2021;18(15):7866.
  49. Bonde AH, Stjernqvist NW, Sabinsky MS, Maingal HT. Process evaluation of implementation fidelity in a Danish health-promoting school intervention. *BMC Public Health*. 2018;18(1):1407.
  50. Langford R, Bonell C, Komro K, et al. The Health Promoting Schools framework: known unknowns and an agenda for future research. *Health Educ Behav*. 2016;44(3):463-475.
  51. Programma Gezonde School. *Criteria en Vragenlijst Themacertificaat Voeding VO/VSO*. Utrecht: Programma Gezonde School; 2022.
  52. Programma Gezonde School. *Evaluatie programma Gezonde School 2017-2020*. Utrecht: Programma Gezonde School; 2020.
  53. Willeboordse M, Jansen MW, van den Heijkant SN, et al. The healthy primary school of the future: study protocol of a quasi-experimental study. *BMC Public Health*. 2016;16:639.
  54. Busch V, van Opdorp PAJ, Broek J, Harmen IA. Bright spots, physical activity investments that work: JUMP-in: promoting physical activity and healthy nutrition at primary schools in Amsterdam. *Br J Sports Med*. 2018;52(20):1299-1301.
  55. Chambers DA, Norton WE. The adaptome: advancing the science of intervention adaptation. *Am J Prev Med*. 2016;51(4 Suppl 2):S124-S131.
  56. Emerson RM, Fretz RI, Shaw LL. *Writing Ethnographic Fieldnotes*. Chicago: University of Chicago Press; 1995.
  57. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101.
  58. Wenger E. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge: Cambridge University Press; 1998.
  59. Schriver KA. Evaluating text quality: the continuum. From text-focused to reader-focused methods. *IEEE Trans Prof Commun*. 1989;32(4):238-255.

60. Ziglio E, Adler M, eds. *Gazing into the Oracle. The Delphi Method and its Application to Social Policy and Public Health*. London: Jessica Kingsley Publishers; 1996.
61. Rayens MK, Hahn EJ. Building consensus using the policy Delphi method. *Policy Polit Nurs Pract*. 2000;1(4):308-315.
62. Ericsson KA, Simon HA. Verbal reports as data. *Psychol Rev*. 1980;87(3):215-251.
63. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334.
64. Rosseel Y. Lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48(2):1-36.
65. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2019.
66. Hu L-t, Bentler PM. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Methods*. 1998;3(4):424-453.
67. Hu L-t, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6(1):1-55.
68. Chen F, Curran P, Bollen K, Kirby J, Paxton P. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociol Methods Res*. 2008;36:462-494.
69. Stevens J. *Applied Multivariate Statistics for the Social Sciences*. 2th ed. Hillsdale: Lawrence Erlbaum Associates; 1992.
70. Neppelenbroek S, Ekkelboom J, de Meij J, Schonenberger M. De GGD als spil in de Gezonde School-aanpak. *Tijdschr Gezondheidsw*. 2020;98(2):82-85.
71. De JM. Gezonde School: de vrijblijvendheid voorbij?! *Tijdschr Gezondheidsw*. 2020;98(2):23-24.
72. Boot NMWM, de Jongh DM, Leurs MTW, de Vries NK. Gezonde School als methode voor GGD'en bij de invoering van schoolgezondheidsbeleid. *Tijdschr Gezondheidsw*. 2011;89(4):222-228.
73. Frigge VK, Nanney MS, Harnack L, Haggemiller M, Pratt R. Using theory to evaluate the implementation and integration of an expanded school breakfast program in rural midwestern high schools. *J Nutr Educ Behav*. 2019;51(3):277-286.
74. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci*. 2007;2(40). <https://implementationscience.biomedcentral.com/articles/10.1186/1748-5908-2-40#citeas>.
75. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186-3191.
76. Rozema AD, Mathijssen JJP, Jansen MWJ, van Oers JAM. Sustainability of outdoor school ground smoking bans at secondary schools: a mixed-method study. *Eur J Public Health*. 2018;28(1):43-49.
77. Dadaczynski K, Rathmann K, Hering T, Okan O. The role of school leaders' health literacy for the implementation of Health Promoting Schools. *Int J Environ Res Public Health*. 2020;17(6):1855.
78. Lassen AD, Christensen LM, Spooner MP, Trolle E. Characteristics of canteens at elementary schools, upper secondary schools and workplaces that comply with food service guidelines and have a greater focus on food waste. *Int J Environ Res Public Health*. 2019;16(7):1115.

## SUPPORTING INFORMATION

The following Supporting Information is available for this article:

**Table S1.** Factor loadings for model 1.

**Data S1.** The HPS implementation questionnaire.

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