BMJ Open

Lower secondary school students' scientific literacy and their proficiency in identifying and appraising health claims in news media: a secondary analysis using large-scale survey data

Lena Victoria Nordheim [1,2] Kjell Sverre Pettersen,[3] Birgitte Espehaug,[1] Signe Agnes Flottorp,[4,5] Øystein Guttersrud[3,6]

For numbered affiliations see end of article.

Correspondence to
Lena Victoria Nordheim;
lvn@hvl.no

ABSTRACT

Objectives Scientific literacy is assumed necessary for appraising the reliability of health claims. Using a national science achievement test, we explored whether students located at the lower quartile on the latent trait (scientific literacy) scale were likely to identify a health claim in a fictitious brief news report, and whether students located at or above the upper quartile were likely to additionally request information relevant for appraising that claim.

Design Secondary analysis of cross-sectional survey data.

Setting and participants 2229 Norwegian 10th grade students (50% females) from 97 randomly sampled lower secondary schools who performed the test during April–May 2013.

Outcome measures Using Rasch modelling, we linked item difficulty and student proficiency in science to locate the proficiencies associated with different percentiles on the latent trait scale. Estimates of students' proficiency, the difficulty of identifying the claim and the difficulty of making at least one request for information to appraise that claim, were reported in logits.

Results Students who reached the lower quartile (located at −0.5 logits) on the scale were not likely to identify the health claim as their proficiency was below the difficulty estimate of that task (0.0 logits). Students who reached the upper quartile (located at 1.4 logits) were likely to identify the health claim but barely proficient at making one request for information (task difficulty located at 1.5 logits). Even those who performed at or above the 90th percentile typically made only one request for information, predominantly methodological aspects.

Conclusions When interpreting the skill to request relevant information as expressing students' proficiency in critical appraisal of health claims, we found that only students with very high proficiency in science possessed that skill. There is a need for teachers, healthcare professionals and researchers to collaborate to create learning resources for developing these lifelong learning skills.

Strengths and limitations of this study

► The large and representative sample (n=2229) of lower secondary school students who responded to the science achievement test improves the external validity of our findings.

► Estimating students' proficiencies and task difficulties using Rasch modelling, we could compare students' proficiency in science with the difficulty of identifying and appraising a health claim in a fictitious brief news report.

► All achievement test items were piloted twice to ensure a valid and reliable measure of scientific literacy, and the use of a digitalised assessment tool reduced sources of errors.

► We did a secondary analysis of test data collected in 2013, thus a shift in proficiency in subsequent student cohorts may have occurred.

► Using raters to code responses to the open-constructed 'news report' item, there is a potential of misclassifying responses owing to rater subjectivity.

BACKGROUND

News media is a leading source of health and scientific information for the public,[1 2] including adolescents and young people, who frequently encounter and share news and information through digital media.[3 4] According to Eurostat, more than two-thirds of young people access online news media regularly.[4] More than half also deliberately search for health information online, indicating health-related topics to be important for youth, especially for those aged 15 years and above.[5]

Media reports of health research often address preliminary and poorly executed studies as sensational 'breakthroughs', leading to large discrepancies between the claims made and the underlying strength of the evidence.[6–8] The result is confusing and

conflicting claims, for instance, about what to eat and drink to maintain good health—claims that influence peoples' perceptions and actions of health.[9 10] Knowledge about scientific methods and scientific concepts is assumed as a necessity for appraising the reliability of health claims.[11 12] Health literacy initiatives at schools might help develop students' skills in apprising claims, and some suggest that these skills may empower students to make informed decisions about health and well-being over the life course.[13 14]

Some claim that a minimum level of scientific literacy is a prerequisite for developing health literacy.[13 15] The aim of compulsory science education is to develop students' scientific literacy, including the proficiency to design and evaluate scientific inquiry, and gain knowledge about how the procedures of science support or disprove claims.[16] School science may therefore be a key learning area for developing adolescents' proficiency to critically appraise health claims in the media. Importantly, educational frameworks promote media reports of research as important real-life contexts for advancing and assessing students' scientific literacy in terms of evidence appraisal.[16–18]

Without appropriate training, adolescents find it difficult to engage critically with media reports containing scientific content, and this challenge continues as they move from compulsory to higher education.[19–27] Studies indicate that students tend to overestimate the certainty of scientific claims and accept them at face value.[19–21 23 25–27] Moreover, they rely on substitute credibility indicators such as expertise (eg, researchers, journalists) and authors' use of scientific statements and prompts (eg, 'evidence-based' or 'scientifically proven') without any in-depth conceptual understanding.[19 20 28]

The majority of these studies reside within the body of research on scientific literacy, not health literacy. This reflects that critical thinking around science-related claims in media, including the proficiency to appraise the science behind health claims, are underscored themes in models and definitions of health literacy.[29 30] Accordingly, these issues are hardly emphasised in measures and empirical studies of adolescents' and young peoples' health literacy.[31–34]

There has been a call for studies that explore how people's scientific literacy correspond to their proficiency in accomplishing specific tasks associated with their health literacy, such as identifying and appraising health claims (National Academies of Sciences, p. 107 [35]). A relevant question concerns 'what someone who scores in the upper quartile on a science literacy measure can do that someone who scores in the lowest quartile cannot?' (National Academies of Sciences, p. 107 [35]). Our study aims to address this question, using data from a national science achievement test of Norwegian 10th grade students. We explore responses to an item designed as a brief news report of a fictitious scientific study that assessed students' proficiency to identify and appraise a health claim.

In Norway, grade 10 is the final year of compulsory education and most students are 15 years of age. According to the Programme for International Student Assessment (PISA) studies of those aged 15 years, Norwegian students perform slightly above the OECD average in science (OECD, p. 44 [36]) and approximately 80% and 30% perform at or above PISA proficiency level 2 and 4 in science, respectively (OECD, p. 320 [36]). At level 2, students can *typically* 'use common scientific knowledge to identify a valid conclusion from a simple data set' (OECD, p. 68 [36]) and hence identify scientific claims—a prerequisite for appraising claims.[20 21 37] At level 4, students can *typically* 'identify the evidence supporting a scientific claim' and draw on knowledge about scientific procedures (eg, experimental designs) to justify conclusions (OECD, pp. 72–4 [36]). Hence, they can most likely request further evidence when encountering unsupported science-based claims, a hallmark of critical appraisal.[38] Previous studies suggest that students, *if they* request information, usually emphasise methodological aspects of the reported research, the findings as such, and theoretical explanations of the findings.[24 25 27 37 39 40]

Building on knowledge from prior research and applying the national science achievement test of Norwegian 10th grade students as a measure of scientific literacy, we hypothesised that:

1. Students who score at or above the lower quartile on the scientific literacy measure are proficient in identifying a health claim among other competing textual information.
2. Students who score at or above the upper quartile on the scientific literacy measure are proficient in *both* identifying a health claim *and* formulate at least one request for further information relevant for appraising that claim, predominately information about either the research methods applied, the data collected or the underlying mechanisms causing an outcome.

METHODS
Design
We did a secondary analysis of existing data from a large-scale cross-sectional, web-based science achievement test assessing a random sample of the 2013 cohort of 10th grade students in Norway.

Participants
In 2013, the cohort of 10th grade students comprised about 64 000 individuals[41] distributed across 1238 schools.[42] Using random sampling, excluding special schools and international schools, 200 public schools were contacted for consensus of participating in the voluntary student assessment. Eligible schools were selected with a probability-proportional-to-size sampling. No schools selected themselves into the study. All schools were contacted by email and telephone between 20 December 2012 and 6 February 2013. One class at each of 97 schools—a total of 2229 students (50% females), completed the digitalised

assessment during April–May 2013. We estimated the school/class average participation rate as 86%. Owing to technical shortcomings beyond our control, no data on students' socioeconomic status or ethnicity was recorded. The mean final assessment grade in science at each school was available from the Norwegian Directorate for Education and Training. On a scale from 1 to 6, where 6 is best, the sample average grade in science was 4.0—identical to the eligible population average. No experimental manipulations or interventions were part of our study.

Participant and public involvement

Participants were not involved in the development of any part of this study.

Study context

In Norway, the integrated subject 'natural science' is a mandatory subject throughout compulsory education. At the time of the survey (spring 2013), the natural science curriculum was structured into six subject domains: 'body and health', 'diversity in nature', 'the universe', 'phenomena and substances', 'technology and design' and 'the budding researcher'.[43] The latter is a cross-cutting domain to ensure that *knowledge about science as a process* is integrated more systematically throughout science domains.

The national science achievement test assessed students' proficiency in science based on the competence aims in the science curriculum for grade 8–10, with assessment items distributed across the cognitive domains 'knowing' (knowledge of scientific facts, concepts and procedures), 'applying' (apply knowledge to explain phenomena and solve problems) and 'reasoning' (evaluating scientific enquiry and alike). The items were distributed across the science domains and cognitive domains as described in online supplementary file 1, the 2013 assessment emphasised the science domains 'body and health' and 'diversity in nature'.

Test items and the administration procedures

The 54 test items constituted a sufficiently valid and reliable scale for measuring scientific literacy as defined by the Norwegian curriculum. All but the one open-constructed news item, positioned at the end of the assessment test and scored 0–4 points, were dichotomously scored selected-response items. Accordingly, the science test data were analysed against the partial credit parameterisation of the unidimensional Rasch model.[44 45] By sampling items from a bank of prior field-tested items, it was possible to construct a scale with difficulty well-targeted at the population of interest. The 'test reliability' was acceptable (Cronbach's alpha based on completely scored data=0.93; person separation index based on person proficiency estimates=0.92). Measured up against the applied Rasch model all but one item discriminated sufficiently well between students with low and high standing on the latent trait (scientific literacy), and no significant differential item functioning, violations of

unidimensionality or local independence were observed. The one poorly discriminating item was discarded and the analysis was re-run. The students completed the test within 80 min at school using a digitalised assessment tool.

The open-constructed 'news item' was designed to evaluate students' proficiency in identifying and critically appraising a health claim. The item's stem was designed as a brief news report (70 words) that referred to a fictitious study concluding that eating corn regularly reduces the risk of type II diabetes. The content and format of the item was similar to the news brief items in an instrument developed by Korpan *et al*,[38] with no details about the study except being conducted by American scientists. In addition, there was a brief background statement about the rising global prevalence of type II diabetes along with a declaration from a diabetes interest group promoting the study findings. Students were first asked to identify the health claim in the news report, more specifically the conclusion from the fictitious study (the word 'conclusion' was used in the item's question), that is, a regular intake of corn reduces the risk of type II diabetes. Second, they were asked to generate requests for information about the study that they would need to appraise the reliability of the health claim. Students were instructed to write a maximum of one and two sentences for the health claim and requests, respectively. A 250-character limit on students' responses was imposed by the electronic assessment system (beyond our control). Responses to the news item allowed us to assess aspects of students' *functional* and *critical* health literacy,[46] more specifically their comprehension of health information and claims, and their ability to critically appraise claims. The item has been retained for continuous test use and is thus unavailable for publication.

Analysis

We coded responses to the news item using a coding guide of assessment criteria that reflected both credited and non-credited responses with regard to identifying the health claim (first part of the item) and requesting information about the study referred to in the item's stem (second part of the item). The process of coding students' information requests was based on a taxonomy for classifying questions and knowledge about scientific research.[38] See table 1 for an overview of the coding guide, including the taxonomy's main scientific research categories (eg, methods). We continually improved the coding guide during field tests and clarified it by including examples of authentic student responses (see online supplementary file 2 for a complete version of the guide).

One rater coded all student responses and consistency was evaluated by using an additional rater who coded 25% of the responses. Inter-rater agreement (ØG and KSP) for the health claim was 94% and improved to 96% after discussion. Inter-rater agreement (LVN and KSP) for the information requests was 86% and improved to 98% after discussion. The lower initial agreement rate

Table 1 Overview of coding guide for the news item

Part 1: health clam	
Credited response	*A response…*
Complete	providing a complete account of the claim (ie, that a regular intake of corn reduces the risk of type II diabetes).
Mostly complete	providing a mostly complete account of the claim, with some significant words lacking (ie, 'regular' and/or 'type II'), and/or referring to amount of intake (eg, 'much' or 'more' corn).
Non-credited response	*A response…*
Wrong	where the claim relates to the topic of the news report, but is otherwise wrong.
Vague	with no reference to *corn* and/or *type II diabetes.*
Other	which is irrelevant or a 'do not know' response.
Blank	
Part 2: information requests	
Credited response	*A response relating to…*
Methods	*how* the study was conducted, including study design, subjects, procedures and measurements.
Data/statistics	*what* was observed in the reported study, or about statistical tests used to analyse the data.
Theory/agent	*why* the reported effects might have occurred, including questions about the properties of the presumed causal agent and/or possible underlying mechanisms.
Social context	the credentials and bias related to who did the study or funded it and where it was conducted or published.
Relevance	the importance or applicability of the study findings, or the impact of the study.
Related research	whether the findings have been replicated or fit results from previous research.
Ambiguous	the study described in the news report that is ambiguous because it fits under two or more scientific categories.
Non-credited response	*A response…*
Future studies	indicating the need for one or more *future* studies, either in general, or relating specifically to one of the scientific research categories (methods, data, etc).
Disbelief	indicating that the student does not believe that the study has been conducted.
Wrong	relating to the topic of the news report, but is otherwise wrong.
Vague	only vaguely referring to the scientific categories (methods, data, etc).
Other	which is irrelevant or is a 'do not know' response.
Blank	

When coding part 1, the raters applied one variable and used values starting with '1' and '0' to indicate whether the response included an acceptable account of the health claim or not. For part 2, the raters applied eight variables. Seven of these were labelled to reflect the scientific research categories (methods, etc), raters used the values '1' and '0' to indicate whether the response included an acceptable request for information within the specific category. For the eight variable, non-credited responses to part 2, values starting with '0' were used to indicate type of response. Blank responses (part 1 or part 2) were coded with the value '99'.

was mainly owing to interpretation of responses that concerned the need for future studies (see specifications in table 1). As the item's stem explicitly asked students to relate their requests to the specific study presented in the news report, our final decision was not to credit responses that concerned 'future studies'.

Overall, we credited students' responses to the news item according to a 'full credit' (4 points), 'partial credit' and 'no credit' (0 point) system as specified in the scoring guide (table 2). This cumulative scoring guide made it possible to identify a student's skill simply by knowing that student's item score. We considered it unlikely that

students who failed to identify the health claim were able to request information needed to establish the reliability of that claim. Thus, an acceptable account of the claim, as specified in table 1, was a premise for being *credited* on the item.

The software package RUMM2030 was used for Rasch modelling.[47] Using unidimensional Rasch modelling, one may construct a scale and locate each item's threshold(s) on that scale. A dichotomously scored item has one threshold reflecting the difficulty of the item, and a polytomously scored item has $k-1$ thresholds reflecting the difficulty of its k score categories.[48] The news item had

Table 2 News item scoring guide

Credit	Score categories	Type of response
No credit	0	Wrong or vague health claim, irrelevant or blank response.
Partial credit	1	Acceptable account of the health claim.
	2	Acceptable account of the health claim and requests relating to one scientific research category (eg, methods).
	3	Acceptable account of the health claim and requests relating to two unique scientific research categories (eg, methods and data).
Full credit	4*	Acceptable account of the health claim and requests relating to three or more unique scientific research categories (eg, methods, data and theory/agent).

*No student made more than three unique requests for scientific information. Accordingly, full credit on the item was set to 4 score points.
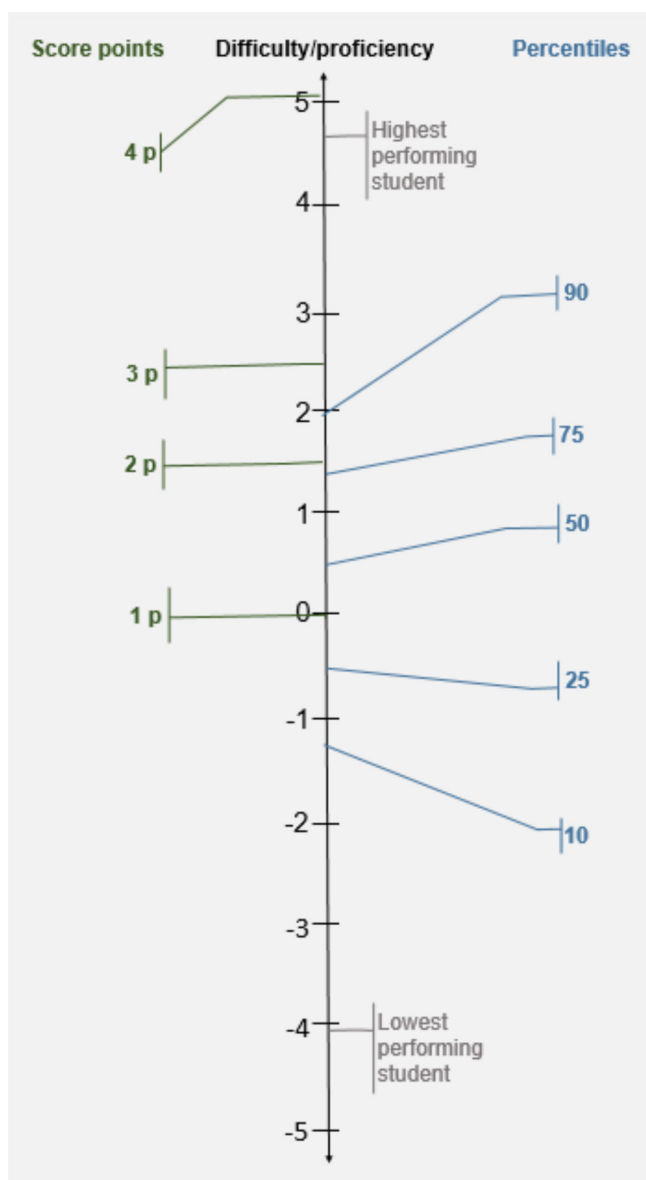

Figure 1 Difficulty of each score on the news item (left side) and the proficiencies associated with different percentiles (right side) on the scientific literacy scale.

five score categories (table 2), and four ordered thresholds reflecting the difficulties of each score. We located the four thresholds on the left side of the scale in figure 1. The scale was made up of observable behaviours—the specific achievements associated with each threshold of the news item described in table 2. These observed achievements were governed by the students' proficiency in science (scientific literacy)—the underlying but unobservable latent trait. On the right side of the scale (figure 1), we located the person (student) proficiencies associated with the 10th percentile, the quartiles and the 90th percentile. The possibility of locating item thresholds (difficulties) and person proficiencies on the same logit-scale, is a benefit of using Rasch modelling. We used the information in figure 1 to test both our hypotheses.

Missing data were handled using pairwise maximum likelihood estimation for the item location estimates—a so-called full information method. During field trials, items displaying 'differential item functioning' (DIF) for central person factors were revised or discarded. DIF means that for example males and females or minority and majority students with the same proficiency estimate have different probabilities of responding correctly. Hence, items displaying DIF are biased as gender and/or cultural background significantly influences students' responses. An example is an item assessing how hormones influence the menstruation cycle. This item probably uniformly favours females at all proficiencies along the latent trait scale.

RESULTS

Two-thirds (64%) of the students identified the health claim, of whom only half gave a complete account of it (table 3). Figure 1 shows that the difficulty associated with identifying the health claim was 0.0 logits (score point 1), that is, it equals the mean difficulty of the test items in the national achievement test, which was set to 0.0. Accordingly, the average scientific literate student was likely able to identify the health claim in the brief news story. Students who reached the lower quartile on the scientific

Table 3 Proportion of students who identified the health claim of the (news item part 1) and the proportions who requested different types of information for appraising that claim (part 2)

	N (of 2229)	%
Part 1: health claim		
Credited responses	1420	64
Complete	710	32
Mostly complete	710	32
Non-credited responses	809	36
Wrong, vague or other	415	18
Blank	394	18
Part 2: information requests		
Credited responses	652*†	29†
Methods	376	17
Data/statistics	189	
Theory/agent	146	7
Social context	57	3
Relevance	12	<1
Related research	9	<1
Ambiguous	31	1
Non-credited responses	1577	71
Future studies	365	16
Disbelief	79	4
Wrong, vague or other	618	28
Blank	515	23

*Comprise all students who made one or more information requests, including the 50 students who were not credited on the news item part 1 (health claim).

†The total sum of requests will exceed 652 as 154 students made requests relating to more than one unique scientific category.

literacy scale were *not* likely to identify the claim, as their proficiency (−0.5 logits) was 0.5 logits below the difficulty estimate of score point 1. Hence, hypothesis 1 was weakened as students' skills were much poorer than we expected based on our interpretation of PISA results.

Less than one-third of the students (29%) made one or more information requests about the reported study relevant to appraise the health claim (table 3). Figure 1 indicates that the difficulty associated with score point 2 (1.5 logits)—identifying the claim and making one request for information, was rather close to the proficiency associated with the upper quartile (1.4 logits). Therefore, hypothesis 2 was strengthened. However, students located even at or below the 90th percentile (2 logits) were not likely to score >2 points on the news item, that is, identifying the health claim and making more than one request for information.

A few responses (n=115) exceeded the 250-character limit and were thus truncated by the assessment system. By coding and analysing these responses, we concluded that the technical deficit might have constrained the

opportunities to achieve a higher score (ie, to be credited for further information requests) for 31 students only.

Characteristics of students' information requests

As shown in table 3, and in line with hypothesis 2, the most frequent requests were related to how the study had been conducted (methods), the data collected (data/statistics) and the theoretical explanations of the results (theory/agent). The requests across these topics varied in level of detail (table 4). More than half of the requests about data/statistics were rudimentary. In comparison, all requests about theory/agent were specific, for example, concerning what active ingredient in corn actually caused the preventive effect. Methods was the only topic where several students made more than one request about specific features of the topic. Nearly half of these requests concerned the study participants, primarily the sample size (126 of 230 requests). Less frequent were requests about design, including the control of confounding variables and use of control groups (33 of 60 requests). As these requests belonged to the same unique scientific research category (methods), they were credited only once (table 2).

Seventy-one per cent of students provided a response that was either blank or otherwise disapproved, and were thus assigned a 'non-credit' category. The average proficiency estimate for this group was 0.0 logits, which equals the difficulty of identifying the health claim. Students who made suggestions for the conduct of future studies, rather than making requests for information related to the study reported in the news story (and thus were not credited), performed somewhat better on the achievement test (average proficiency 0.64 logits).

DISCUSSION

We assessed how 10th grade students' levels of scientific literacy corresponded to their proficiency in identifying and critically appraising health claims in the news media—two essential aspects of health literacy. The findings weakened our first hypothesis, as only the average scientific literate students aged 15 years leaving compulsory school was able to identify a clearly stated health claim in a rather simple news report. Students performing at the upper quartile of the scientific literacy measure were barely proficient at identifying and appraising the claim, namely making a request for evidence needed to determine the reliability of the claim. Accordingly, our second hypothesis was strengthened.

About half of the invited schools participated, and the average student participation rate at these schools was at an acceptable 86%. Our analyses indicated that the sample school average grade and gender distribution matched the population distributions at grade 10, thus indicating generalisability of our findings. Although data on socioeconomic status and ethnicity was unavailable for this study, previous studies have found these factors to predict science proficiency in Norwegian students aged

Table 4 Characteristics of the requests related to methods, data and theory (specifications of table 3)

Type of request	Examples of students' responses	No. of requests	%
Methods		**471**	**100**
Rudimentary	How was the study conducted?	80	17
Design	Did they use a control group? How long did the study last?	60	13
Agent delivery (procedure)	How much corn is necessary to eat? How often?	100	21
Participants	How many were tested? Who participated in the study? What was their eating habits?	230	49
Measures	Did they measure the participants' blood sugar?	1	>1
Data/statistics		**191**	**100**
Rudimentary	We need the results.	111	58
Absolute nature of data*	For how many did this (not) work?	32	17
Comparative nature of data*	What is the reduction in risk? How large was the effect?	43	23
Duration of effect	Does it work over time?	5	3
Theory/agent		**149**	**100**
Identification	We need to know what corn is composed of (nutrients).	2	1
Agent mechanisms	What ingredient in corn prevents type II diabetes? Why does corn prevent type II diabetes?	117	79
(Side) effects	Is there any side effect of eating corn?	23	15
Alternative agents	Do other corn products have this effect?	7	5

*For the dependent variable (here: diabetes risk).

15 years.[36 49 50] For socioeconomic status, however, the relationship with proficiency is relatively weak compared with most other countries.

Previous studies of students' evaluations of scientific claims have mostly been conducted at upper secondary school level and above, and on smaller, mostly self-selected samples of students.[24 25 27 37 39 40] Thus, to our knowledge, this is the first study investigating students' critical appraisal of science-based health claims in the context of a large student sample at lower secondary school level. Furthermore, while students' scientific health knowledge in important areas such as chronic and infectious diseases and their knowledge of sources to science-based health information has previously been explored,[51 52] we have addressed a call for research on the utility of scientific literacy for critical appraisal of health claims.[35] The analysis of responses to the news item provided useful information about *how* proficient students were (ie, the levels of scientific literacy necessary to involve in identifying and appraising health claims in news reports of science), and *what* kind of knowledge they possessed and applied when approaching such claims (ie, which responses earned credits and which did not). Such in-depth knowledge of students' thinking around a topic or task is an important outcome of secondary analyses of individual test items used in large-scale surveys, as it may formatively inform and develop teachers' practices.[53]

There were some limitations to this study. First, the test data were collected 6 years ago (2013), thus we acknowledge that a possible shift in students' knowledge might have occurred. Nevertheless, our study is timely due to a major revision of the curriculum that will be implemented from autumn 2020. Second, to avoid response dependence between similar items, and accordingly violations of local independence in the data, the test comprised only one of the news item developed and field-tested. This prohibited us from evaluating whether variations in text dimensions (eg, the claim's plausibility and how familiar the students were with the health topic) could have influenced students' information request—dimensions previously reported to impact on students' critical engagement with news reports.[37 39] For the same reason, and because the news item did not include any embedded attitudinal items, we were unable to assess whether important personal factors (eg, interest in the health topic, belief in the claim, scientific attitude)[24 25 37 39 40] could have affected students' requests. Third, to make students respond shortly and 'on task', they were encouraged to write only two brief sentences. This might have constrained their opportunities to make several requests for evidence regarding the claim, although our analysis of incomplete responses due to the limit of 250 characters indicated this was probably not the case. Moreover, our findings resemble previous studies with regard to both the number and type of requests made.[24 25 37 39] Finally, while it is common practice to retain test items for re-use, this implies a lack of transparency in the test data used for this study.

Being able to correctly identify the nature of information included in media reports of science is a

prerequisite for critical appraisal. High school and university students exposed to news reports containing a variety of scientific features, often confuse conclusion statements with statements about the results (data) and explanations (theory).[20 21 37] In comparison, the news item in our study was less complex, including only a few statements beside the health claim (study's conclusion). Still, only students being proficient at or above the average in science managed to identify the claim, often providing an incomplete account of it. Therefore, the underlying problem seems to be the same as noted in previous studies, namely a lack of training in *reading* media reports of science.[20 21 37] As previously noted, we were unable to explore whether the students perceived the claim as plausible or not. The former might be the case as uncertainty about a claim's plausibility has been found to provoke more methods questions.[37 39] For instance, students' occasional requests about procedures indicate that they believed in the claim, and thus simply wanted information about how much, how often and how long the intake of corn should be to see the reported effect on diabetes risk. A further observation was the few requests about the social context of the research (eg, the American scientists' affiliation), perhaps suggesting that students regarded science and scientists as authoritative and accordingly that the claim was plausible. This has also been noted in previous studies of students across educational levels.[19–21 23 25–27]

Our findings are of concern as they illustrate students' functional and critical health literacy at the end of compulsory school. Almost three-quarters of the 10th grade students were unable to identify *and* critically appraise the health claim in a brief news report. Even the highest performing students mostly requested only one scientific research category of an optimum of six broad categories. Despite curricular mandates to develop scientific literacy and critical appraisal skills important for health literacy, it seems like students' actual skills are underdeveloped and not taught in a way that improve their appraisal of health claims as assessed by the news item. This is consistent with findings from a qualitative study where science teachers reported that opportunities for teaching critical appraisal during inquiry-based activities, such as online health information seeking or small-scale experiments, are lost in the need to emphasise factual knowledge on health topics.[54] Importantly, teachers do not acknowledge the relevance of teaching critical appraisal, or they lack methods to teach them. Thus, there is clearly a prospect for cross-sector collaboration between healthcare and education professionals and researchers to work together to enable laypersons to appraise health claims, as pointed out by Sharples *et al*.[14] Research is still scarce as to which interventions best improve students' ability to appraise health claims.[55 56] However, a recent cluster-randomised controlled trial shows promising effects of a cross-disciplinary developed intervention aimed at teaching primary school children to appraise claims about treatment effects.[57]

Our study has identified specific areas that require attention in further development and evaluation of interventions—areas that align with important key concepts lay people need to know to assess health claims.[11] It was encouraging that the students in our study—*when* they employed scientific criteria—were sensitive to methodological information, which often is lacking in media reports of health research.[7] However, students requested only a limited range of methodological evidence, with little attendance to details about the study design, such as the use of control groups or control of confounding variables. This was noteworthy given the news report's assertiveness in claiming a causal relationship between the intake of corn and the reduced risk of type II diabetes, mirroring the many misleading media reports that fail to differentiate association from causation.[6 7] Science instruction should therefore develop students' knowledge of good and weak designs for establishing a cause-effect relationship, including the design of controlled studies, the importance of fair comparisons, the principles of randomisation and blinding and proper and improper ways of reporting outcomes (eg, absolute vs relative risk). Existing evidence suggests that such knowledge is better gained through teacher-guided investigations that allow students to reflect on adequate and inadequate experimental strategies, rather than through student-led hands-on or virtual experiments.[58] Importantly, teachers need to make explicit the link between experiments, critical reading, and appraisal of health claims in the new(s) media. In our study, several students suggested the conduct of a future study rather than requesting information about the reported study, often involving themselves in doing so (eg, "we have to test a number of people"). This perhaps supports the notion that teachers enforce experimentation and hands-on activities without linking relevant learning outcomes to reading and critically appraising science presented in out-of-school contexts, including media reports.[59 60] Finally, students hardly requested related research supporting or disproving the claim. Accordingly, teaching could sensitise students to the limitations of single studies, introducing the idea of systematic reviews.

Author affiliations
[1]Centre for Evidence-Based Practice, Western Norway University of Applied Sciences, Bergen, Norway
[2]Faculty of Medicine and Dentistry, Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway
[3]Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway
[4]Norwegian Institute of Public Health, Oslo, Norway
[5]Institute of Health and Society, University of Oslo, Oslo, Norway
[6]Faculty of Mathematics and Natural Sciences, Norwegian Centre for Science Education, University of Oslo, Oslo, Norway

Contributors LVN, KSP and ØG developed a revised version of the news brief item and the coding guide, and coded the student responses. ØG administered the survey, managed the data handling, conducted the Rasch analysis and constructed the achievement scale. LVN wrote the preliminary draft of the paper

ORCID iD
Lena Victoria Nordheim http://orcid.org/0000-0001-7370-1121

REFERENCES

1 European Commission. *Scientific research in the media (Special Eurobarometer 282)*. Brussels: European Commission, 2007.

2 Funk C, Gottfried J, Mitchell A. *Science news and information today*. Washington, DC: Pew Research Center, 2017.

3 Anderson M, Jiang J. *Teens, social media and technology 2018*. Washington, DC: Pew Research Centre, 2018.

4 Eurostat. *Being young in Europe today*. Luxembourg: Publications Office of the European Union, 2017.

5 Park E, Kwon M. Health-related Internet use by children and adolescents: systematic review. *J Med Internet Res* 2018;20:e120.

6 Haber N, Smith ER, Moscoe E, *et al*. Causal language and strength of inference in academic and media articles shared in social media (CLAIMS): a systematic review. *PLoS One* 2018;13:e0196346.

7 Schwitzer G. How do US journalists cover treatments, tests, products, and procedures? An evaluation of 500 stories. *PLoS Med* 2008;5:e95.

8 Sumner P, Vivian-Griffiths S, Boivin J, *et al*. The association between exaggeration in health related science news and academic press releases: retrospective observational study. *BMJ* 2014;349:g7015.

9 Matthews A, Herrett E, Gasparrini A, *et al*. Impact of statin related media coverage on use of statins: interrupted time series analysis with UK primary care data. *BMJ* 2016;353:i3283.

10 Nagler RH. Adverse outcomes associated with media exposure to contradictory nutrition messages. *J Health Commun* 2014;19:24–40.

11 Austvoll-Dahlgren A, Oxman AD, Chalmers I, *et al*. Key concepts that people need to understand to assess claims about treatment effects. *J Evid Based Med* 2015;8:112–25.

12 Schwitzer G. A guide to reading health care news stories. *JAMA Intern Med* 2014;174:1183–6.

13 Grace M, Bay JL. Developing a pedagogy to support science for health literacy. *Asia-Pacific Forum on Science Learning and Teaching*. 2011;12(2). Available: https://www.eduhk.hk/apfslt/v12_issue2/foreword/foreword7.htm [Accessed 4 Sep 2019].

14 Sharples JM, Oxman AD, Mahtani KR, *et al*. Critical thinking in healthcare and education. *BMJ* 2017;357:j2234.

15 Harrison JK. Science education and health education: locating the connections. *Stud Sci Educ* 2005;41:51–90.

16 OECD. *PISA 2015 Assessment and analytical framework*. Paris: OECD Publishing, 2017.

17 Millar R, Osborne J, (eds). *Beyond 2000: science education for the future*. London: King's College London, 1998.

18 National Research Council. *A framework for K-12 science education*. Washington, DC: The National Academies Press, 2012.

19 Kolstø SD. 'To trust or not to trust,…'-pupils' ways of judging information encountered in a socio-scientific issue. *Int J Sci Educ* 2001;23:877–901.

20 Norris SP, Phillips LM. Interpreting pragmatic meaning when reading popular reports of science. *J Res Sci Teach* 1994;31:947–67.

21 Norris SP, Phillips LM, Korpan CA. University students' interpretation of media reports of science and its relationship to background knowledge, interest, and reading difficulty. *Public Underst Sci* 2003;12:123–45.

22 Oliveras B, Márquez C, Sanmartí N. The use of newspaper articles as a tool to develop critical thinking in science classes. *Int J Sci Educ* 2013;35:885–905.

23 Oliveras B, Márquez C, Sanmartí N. Students' attitudes to information in the press: critical reading of a newspaper article with scientific content. *Res Sci Educ* 2014;44:603–26.

24 Pettersen S. Critical thinking in Norwegian upper secondary biology education: the cases of complementary-alternative-medicine and health claims in the media. *Nordic Studies in Science Education* 2005;1:61–71.

25 Pettersen S, Solberg J. Students of health sciences' evaluation of media reports of health research: a Norwegian study. In: Lewis J, Magro A, Simonneaux L, eds. *Biology education for the real world: student - teacher - citizen; Proceedings of the IVth ERIDOB Conference; 2002 Oct 22-26*. Toulouse-Auzeville: ENFA, 2003: 293–306.

26 Phillips LM, Norris SP. Interpreting popular reports of science: what happens when the reader's world meets the world on paper? *Int J Sci Educ* 1999;21:317–27.

27 Ratcliffe M. Evaluation of abilities in interpreting media reports of scientific research. *Int J Sci Educ* 1999;21:1085–99.

28 Cusack L, Desha LN, Del Mar CB, *et al*. A qualitative study exploring high school students' understanding of, and attitudes towards, health information and claims. *Health Expect* 2017;20:1163–71.

29 Bröder J, Okan O, Bauer U, *et al*. Health literacy in childhood and youth: a systematic review of definitions and models. *BMC Public Health* 2017;17:361.

30 Sørensen K, Van den Broucke S, Fullam J, *et al*. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health* 2012;12:80.

31 DeWalt DA, Hink A. Health literacy and child health outcomes: a systematic review of the literature. *Pediatrics* 2009;124(Suppl 3):S265–74.

32 Fleary SA, Joseph P, Pappagianopoulos JE. Adolescent health literacy and health behaviors: a systematic review. *J Adolesc* 2018;62:116–27.

33 Guo S, Armstrong R, Waters E, *et al*. Quality of health literacy instruments used in children and adolescents: a systematic review. *BMJ Open* 2018;8:e020080.

34 Stellefson M, Hanik B, Chaney B, *et al*. eHealth literacy among college students: a systematic review with implications for eHealth education. *J Med Internet Res* 2011;13:e102.

35 National Academies of Sciences, Engineering, Medicine. *Science literacy: concepts, contexts, and consequences*. Washington, DC: The National Academies Press, 2016.

36 OECD. *PISA 2015 results (Volume I): excellence and equity in education*. Paris: OECD Publishing, 2016.

37 Korpan CA. *Science literacy: what do students know and what do they want to know?* Toronto: Canadian Council on Learning, 2009.

38 Korpan CA, Bisanz GL, Dukewich TL, *et al*. *Assessing scientific literacy: a taxonomy for classifying questions and knowledge about scientific research. technical report No. 94-1*. Edmonton, Canada: University of Alberta, 1994.

39 Korpan CA, Bisanz GL, Bisanz J, *et al*. Assessing literacy in science: evaluation of scientific news briefs. *Sci Educ* 1997;81:515–32.

40 Leung JSC, Wong ASL, Yung BHW. Evaluation of science in the media by non-science majors. *Int J Sci Educ B Commun Public Engagem* 2017;7:219–36.

41 Grunnskolens informasjonssystem (The information system for primary and secondary schools) [program]. Oslo: Norwegian Directorate for education and training, 2018. Available: https://gsi.udir.no/ [Accessed 4 Sep 2019].

42 Statistics Norway. Table 1. Primary and lower secondary schools. Pupils in primary and lower secondary school. Oslo: Statistics Norway, 2018. Available: https://www.ssb.no/en/utdanning/statistikker/utgrs [Accessed 4 Sep 2019].

43 Norwegian Directorate for Education and Training. *Natural science subject curriculum (NAT1-01)*. Oslo: The Directorate, 2006.

44 Masters GN. A rasch model for partial credit scoring. *Psychometrika* 1982;47:149–74.

45 Rasch G. *Probabilistic model for some intelligence and achievement tests*. Copenhagen: Danish Institute for Educational Research, 1960.

46 Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int* 2000;15:259–67.

47 RUMM Laboratory Pty Ltd. RUMM2030 [program] 2011.

48 Andrich D. A rating formulation for ordered response categories. *Psychometrika* 1978;43:561–73.

49 OECD. *PISA 2006 Volume II: Data/Données*. Paris: OECD Publishing, 2007.

50 Turmo A. Scientific literacy and socio-economic background among 15-year-olds—a Nordic perspective. *Scand J Educ Res* 2004;48:287–305.

51 Ghaddar SF, Valerio MA, Garcia CM, *et al*. Adolescent health literacy: the importance of credible sources for online health information. *J Sch Health* 2012;82:28–36.

52 Khajouei R, Salehi F. Health literacy among Iranian high school students. *Am J Health Behav* 2017;41:215–22.

53 Angell C, Kjaernslie M, Lie S. Exploring students' responses on free-response science items in TIMSS. In: Shorrocks-Taylor D, Jenkins EW, Angell C, eds. *Learning from others: international comparisons in education*. Hingham, MA, USA: Kluwer Academic Publishers, 2000.

54 Nordheim L, Pettersen KS, Flottorp S, *et al*. Critical appraisal of health claims: science teachers' perceptions and practices. *Health Educ* 2016;116:449–66.

55 Cusack L, Del Mar CB, Chalmers I, *et al*. Educational interventions to improve people's understanding of key concepts in assessing the effects of health interventions: a systematic review. *Syst Rev* 2018;7:68.

56 Nordheim LV, Gundersen MW, Espehaug B, *et al*. Effects of school-based educational interventions for enhancing adolescents abilities in critical appraisal of health claims: a systematic review. *PLoS One* 2016;11:e0161485.

57 Nsangi A, Semakula D, Oxman AD, *et al*. Effects of the informed health choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial. *Lancet* 2017;390:374–88.

58 Schwichow M, Croker S, Zimmerman C, *et al*. Teaching the control-of-variables strategy: a meta-analysis. *Developmental Review* 2016;39:37–63.

59 Barton ML, Jordan D. *Teaching reading in science: a supplement to "teaching reading in the content areas teacher's manual"*. 2nd edn. Aurora, Colorado: McREL, 2001.

60 Ekborg M, Ottander C, Silfver E, *et al*. Teachers' experience of working with socio-scientific issues: a large scale and in depth study. *Res Sci Educ* 2013;43:599–617.