MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(1), em2380
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/13930

Introduction of Taiwanese literacy-oriented science curriculum and development of an aligned scientific literacy assessment

Tin-Chang Chang [1], Yi-Ming Lyu [1], Huang-Ching Wu [2,3*], Kao-Wei Min [4]

[1] Department of Business Administration, School of Management, Asia University, Taichun, TAIWAN
[2] Science Education Center, National Taiwan Normal University, Taipei City, TAIWAN
[3] Ming-lun Elementary School, Yunlin, TAIWAN
[4] College of Engineering, National Formosa University, Huwei, Yunlin, TAIWAN

Received 24 July 2023 ▪ Accepted 14 October 2023

Abstract
This article reports on the introduction of Taiwanese new literacy-oriented science curriculum reform and the development of a measure of scientific literacy (SL). Curriculum reform has always been received increasing attention from educators in many countries around the world. Meanwhile, trends in science education policy have emphasized the importance of SL as a transferable outcome and the main goal of science education (Fives et al., 2014). It would seem reasonable, therefore, that the new science curriculum guidelines (NSCG) would be developed for grade 3-grade 12 in Taiwan for making progress toward the goal of SL. In this article, the authors (a) discuss the background of science education reform in Taiwan, (b) introduce and describe the features of NSCG, (c) evaluate the relative strengths and limitations of the present assessments, (d) describe a framework for aligning assessment with NSCG, and (e) conduct a pilot study for item analysis. For the 6th grade level, the pilot test reported an acceptable reliability coefficient, high item difficulty and good discrimination value of scientific literacy assessment (SLA). Further revision is necessary to make available a series of validated and reliable items being developed for assessing students' SL at various science learning stages. Hopefully, SLA will finally fill the gap of the assessment part of the current science curriculum reform in terms of guiding educators to precisely evaluate students' SL.

Keywords: scientific literacy, new science curriculum guidelines, scientific literacy assessment

INTRODUCTION

Indeed, educational reform in Taiwan has garnered significant attention from parents, educators, and students (Chang, 2005). The motivation for this reform stems from several shortcomings in the current curriculum. These include reduced student engagement, an excessive emphasis on test scores, and a limited focus on real-world application (Aldridge et al., 1999; Lai et al., 2015; Martin et al., 2016; Wang, 2004). Moreover, the education system is grappling with various challenges, such as a declining birth rate, an aging population, rapid technological advancements, a changing job landscape, and a growing emphasis on global sustainability. These social transformations, together with the up-mentioned curriculum problems, have compelled educators to reevaluate the current curriculum and strive for educational reform at primary and secondary levels.

In 2014, the Ministry of Education (MoE, 2014) convened panels of experts in various fields, including science and language, to initiate a comprehensive and ongoing curriculum reform process. This reform encompasses eight domains, including the arts, health and physical education, language, mathematics, science, technology, social sciences, and extracurricular activities (MoE, 2017). These domains encompass several critical areas of interest, such as gender equity, human rights, environmental conservation, moral and ethical development, rule of law, technology and information, energy and resources, family education, career planning, cultural diversity, outdoor education, international collaboration, and indigenous education. Overall, the

✉ ervine@asia.edu.tw ✉ lyuyiming.au@gmail.com ✉ tarzenlr@yahoo.com.tw **(*Correspondence)**
✉ d0981102@gm.nfu.edu.tw

Contribution to the literature

- This study introduces Taiwan's current science curriculum to provide an overview of the existing science curriculum in Taiwan. It also reveals implementation challenges to identify and discuss the difficulties faced during the implementation of the current curriculum.
- It examines NSCG to analyze and assess the key features of the newly proposed NSCG.
- It develops a sample assessment aligned with science literacy framework to create a sample assessment that aligns with the framework of science literacy, a central goal of the curriculum reform.

new curriculum guidelines aim to provide more flexibility for individual development and promote lifelong learning (MoE, 2017)

One of the central focuses of this article is the introduction of the new science curriculum guidelines (NSCG). Since 2014, a multitude of scientists, science educators, and teachers have collaborated in developing NSCG, which is tailored to meet the contemporary imperative of equipping citizens with essential skills to solve problems, adapt to their environment, and enhance their lives (MoE, 2017). Specifically, NSCG seeks to establish a framework for science literacy (SL) to facilitate students' mastery of scientific inquiry, foster a positive attitude toward science, and instill essential scientific knowledge. In essence, SL is synonymous with the primary goal of this science education reform. Recently, MoE (2018) formally published NSCG in November 2018. NSCG comprehensively outlines science curriculum frameworks, learning content, performance expectations, and customized pedagogical approaches. However, the assessment of the curriculum is only briefly outlined in NSCG. Additionally, in line with recommendations for 12-year science curriculum development (National Academy for Educational Research [NAER], 2014), the current national scientific achievement tests and teachers' formative and summative assessments predominantly gauge how well students grasp school curricula. These assessments often neglect the critical capacity of students to apply scientific skills and knowledge effectively in various real-world situations (Organization for Economic Co-operation and Development [OECD], 2006). Consequently, assessments rooted in SL framework must be designed to provide valid and reliable insights into students' learning. Further details on assessing SL will be presented in this study.

In summary, this study serves as an overview of the introduction to the science curriculum reform in Taiwan. It aims to aid in the development of more suitable scientific testing methods and shed light on the future development of a new science curriculum.

DISADVANTAGES OF CURRENT SCIENCE CURRICULUM

Nine-year compulsory education system has been implemented in Taiwan for over 50 years and was utilized to enhance Taiwanese education levels, which laid a foundation for social prosperity (MoE, 2017). At the same time, problems of the educational system have arisen. "White paper on human-resource development" (American Chamber of Commerce in Taipei [ACCT], 2013) indicated Taiwan's rapid social changes, overextends colleges and universities; declining birth rates are creating a gap between educational training and careers. The result is an imbalance between human-resource supply and demand. The existing curriculum needs to be examined, and its shortcomings recognized. More specifically, there exist three problems awaiting solutions in the current science curriculum:

Reduced Student Engagement

In elementary, as well as junior high schools in Taiwan, building up students' capacities for lifelong learning is one of the basic aspects emphasized. In the current grade 1-grade 9 curriculum system, the core components of this aspect include: active exploration, problem-solving and the utilizing of knowledge and information. For this purpose, grade 1-grade 9 curriculum guidelines suggest that the science curriculum should be implemented in a "learner-centered" manner, which puts learners' interests first and allows them to develop autonomy and skills for lifelong learning. In practice, the traditional method of education, also dubbed "teacher-centered learning," still dominates (Aldridge et al., 1999; Lai et al., 2015). According to the interview results from middle school teachers in Taiwan, the current science curriculum often fails to fully engage students, leading to lower motivation for learning. Hence, students may find it challenging to stay enthusiastic about science due to the curriculum's design.

Overemphasis on Test Scores

Recent international standardized tests results, from trends in mathematics and science (TIMSS), indicate Taiwan's students are generally among the most successful in grades, but their learning attitude and levels of confidence are among the lowest in international rankings (Martin et al., 2016). In addition, results of students' achievements in science cognitive domains showed their scores in "knowing" were much higher than in "applying" and "reasoning" (Martin et al., 2016). These results imply the problems of testing-oriented education, including the diminishment of students' learning interests, as well as the neglect of their development in higher-order thinking.

Limited Real-World Application

One noticeable drawback is the significant gap between what students learn in science classes and how they can apply that knowledge in real-world situations. Real-world problems are difficult to include on paper-and-pencil tests. According to Wang (2004), test-oriented education overemphasizes teaching and testing second-hand information but neglects to educate learners on how to deal with current real-world issues. Thus, students may struggle to see the practical relevance of what they're taught in the curriculum.

A new science curriculum is required to overcome these given shortcomings. For this purpose, a new science curriculum reform was activated in 2014 by the committees of the research and studies of NSCG. The committees convened for over 100 sessions to discuss the details of new curriculum guidelines as well as take in dialogues with various educational levels and the public. Science curriculum reform is still an ongoing process. The introduction of the science curriculum guidelines, released in 2016, is summarized below.

INTRODUCTION OF NEW SCIENCE CURRICULUM GUIDELINES

The primary goal of the new science curriculum is to nurture SL among students. Before we delve into the specifics of the new curriculum guidelines, it's essential to review the existing literature on SL. This review helps us define what SL means and identify its core components, which will serve as the foundation for developing our NSCG and assessments.

To date, there is not a single universally accepted definition of SL. According to the National Science Education Standards released in 1996, SL implies that an individual possesses the ability to address questions stemming from curiosity about everyday experiences. It involves describing, explaining, and predicting natural phenomena, as well as comprehending and validating conclusions drawn from scientific reports (National Research Council [NRC], 1996). Furthermore, a scientifically literate individual can identify the scientific aspects underlying national and local decisions and has the capacity to construct and evaluate arguments based on evidence, applying conclusions from such arguments (NRC, 1996).

On the other hand, PISA's framework defines SL as skill of critically evaluating scientific evidence and claims within social-scientific contexts (OECD, 2001). Fives et al. (2014) characterize SL as ability to understand and meaningfully engage with both scientific processes and scientific information encountered in everyday life. Various other characterizations of SL discussed in the literature encompass a combination of competencies in scientific inquiry, content knowledge, and attitudes toward science (DeBoer, 2000; Fives et al., 2014; Roberts, 2007).

Wenning (2006) proposed a multidimensional view of SL, defining it as the capacity to demonstrate:

(1) Intellectual skills, including problem-solving, reasoning, scientific inquiry, and critical thinking.

(2) Knowledge, including understanding the content of scientific disciplines, the nature of science, and the history of science.

(3) Dispositions, encompassing values, beliefs, assumptions, attitudes, and actions.

These three theoretically defined essential components underpin Taiwan's new science educational guidelines. These guidelines emphasize providing students with opportunities to engage in scientific inquiry and problem-solving, fostering positive attitudes toward science, and acquiring fundamental scientific knowledge for practical daily use.

Overall, NSCG possesses two key characteristics: *Literacy-oriented* and *Progressive & cross-domain integrated*.

Literacy-Oriented

The curriculum design centers around the concept of scientific literacies, aiming for the holistic development of learners' scientific inquiry skills, knowledge and dispositions. However, to provide practical guidance for science teachers, the committee responsible for developing the new science curriculum has offered a list of learning objectives, including:

(1) Developing students' inquiry ability by providing learning opportunities for them to practice thinking ability and problem solving.

(2) Assisting students in cultivating the right attitude towards science and comprehending the nature of science.

(3) Guiding students in acquiring essential scientific concepts, including "components and characteristics of the natural world," "phenomena and mechanisms in the natural world," and "sustainable development of the natural world."

These three learning objectives are defined as "learning outcomes" within this curriculum framework, while the three themes mentioned above represent " learning content." The learning outcomes related to "inquiry skills" and "attitudes and essence of science" for learners at this stage, while in the "learning content" section, we'll present specific scientific knowledge content. This approach provides a clear and practical definition of scientific literacies to guide both science instruction and assessment. **Table 1** provides a summary of this framework, which also serves as a reference for assessors in designing assessments aligned with SL.

Table 1. Framework of "learning outcomes" (aimed at 5th and 6th grades)

Events	Sub-events	
Scientific cognition	Corresponding to critical learning content, divided into six levels, including **knowledge, comprehension, application, analysis, evaluation,** & **creation.**	
Inquiry ability	**Thinking ability**	**Imagination & creativity:** Detecting changes in natural laws from vital phenomena & exploring possible causes through scientific knowledge. **Reasoning & argumentation:** Connecting observed natural features to prior knowledge, identifying causal factors, & formulating and comparing explanations. **Critical thinking:** Recording & classifying data & checking correctness against existing knowledge. **Modeling:** Creating models to describe inquiry results while acknowledging their limitations.
	Problem-solving	**Observing & identifying:** Observing daily life, natural world, & resources to identify problems & formulate inquiry questions. **Planning & executing:** Conducting scientific inquiry, identifying variables, & performing controlled experiments. **Analyzing & finding:** Analyzing data, constructing graphs, formulating explanations, & developing new knowledge. **Discussing & communicating:** Presenting results through various means, using scientific terminology, & providing constructive feedback.
Attitude toward science & nature of science	**Interest in science:** Satisfying curiosity, enjoying learning, & collaborating during scientific exploration. **Habit of scientific thinking & inquiry:** Applying scientific knowledge to understand daily-life phenomena & addressing questions. **Nature of science:** Recognizing that scientific knowledge is human-developed & evolves based on related evidence.	

Table 2. Framework of "learning content"

Themes	Topics
Components & characteristics of natural world	Components & characteristics of materials Phase & transfer of energy Structures & functions of materials Structures & functions of creatures System of materials' global environment
Phenomenon & mechanism of nature world	Evolution & adaptation History of earth Active planet Balance & creation' reactions of materials Phenomenon & mechanism of nature world Biology & environment
Sustainable development of nature world	Science, technology, society, & humanities Resources & sustainable development

The focus of learning outcomes encompasses the understanding of scientific concepts, inquiry skills, and attitudes towards science and its essence. This curriculum at this learning stage is organized by integrating these three aspects in a suitable manner, considering the characteristics of learners' physical and mental development as well as societal and life needs.

Limited by space, **Table 1** only provides the contents of learning outcomes suitable for 5th and 6th grade students. Further details regarding the coherence and sophistication of this learning system will be discussed later on. The learning contents are the summation of systematic science knowledge, in the current era, suitable for students to learn and apply. These learning contents also act as prior knowledge for science inquiry and problem-solving. The summary of the learning contents is listed in **Table 2**.

While the information listed in **Table 2** may initially seem like a broad overview of scientific knowledge, it is actually introduced in accordance with the intellectual maturity of the students. For example, "components and characteristics of materials" will cover concepts like "the material world," which includes both living organisms such as animals and non-living objects like stones for 3rd and 4th graders. Meanwhile, 5th and 6th graders will delve into more advanced topics, such as "materials consist of particles that are constantly in motion (e.g., vaporizing, condensing, and effusing)." In summary, NSCG aims to nurture students' SL by implementing a curriculum that focuses on crucial learning outcomes and content. Within NSCG framework, equal emphasis is placed on students' cognitive, emotional, and conceptual growth in science education. Another distinctive feature of NSCG, which underscores coherence of learning experiences and content, will be discussed next.

Progressive & Cross-Domain Integrated

NSCG strengthens the learning transition from elementary school up to high-school and focuses on the cross-field curriculum design. This progressive method of curriculum implementation, which finds its root in

Table 3. An example of progressive learning in "imagination & creativities"

Learning stage	Expected learning performances
Grade 3- grade 4	Students observe nature's laws through guided instruction and use their imagination and curiosity to understand and describe natural phenomena.
Grade 5-grade 6	Students identify changes in natural laws from vital phenomena, imagine potential causes on scientific knowledge, & recognize same evidence can result from different experiment approaches.
Grade 7-grade 9	Students explore variations that may occur when changing observation or experimentation methods. They engage in self-exploration/group discussions to generate innovative models/results.
Grade 10-grade 12	Students personally identify observation-based problems, imagine multiple hypotheses, and conduct experiments, or design new ones, individually or collaboratively to test these hypotheses.

present experience as well as emphasizes learning by doing, will ensure comprehensive learning among all kinds of students. **Table 3** lists an example of progressive learning performances set for students, grade 3 to grade 12, to achieve.

The content of **Table 3** outlines the current learning expectations for students at different grade levels. It emphasizes the alignment between these expectations and students' cognitive development. The paragraph also highlights the increasing sophistication of learning expectations as student progress from elementary to high school. It mentions the emphasis on hands-on practice in elementary school, the introduction of scientific knowledge and abstract thinking in junior high school, and the focus on microscopic methods and advanced scientific concepts in senior high school. Finally, it underscores the aim of ensuring that students learn age-appropriate material through this progressive curriculum guideline.

In addition to the progressive approach, there is a strong emphasis on cross-disciplinary learning referred to as "science-inquiry and practical courses." These courses provide students with opportunities to apply and utilize their scientific knowledge and skills effectively. In elementary schools, the primary focus is on cross-disciplinary curriculum design. In junior high schools, there is a predominant emphasis on departmental teaching (biology, physics, chemistry, and earth science), with cross-disciplinary integration serving as supplementary. Similarly, in senior high schools, the main approach is departmental teaching (biology, physics, chemistry, and earth science), alongside the provision of cross-disciplinary curriculum options that allow students to engage in interdisciplinary learning experiences.

NECESSITIES OF NEW SCIENCE CURRICULUM GUIDELINES

There are several issues with the current state of science education in Taiwan, necessitating a reform of the science curriculum. Firstly, to address the problem of low student engagement in science learning, NSCG places a strong emphasis on SL, which is viewed as a means of effectively engaging and motivating students in their science studies (Smith et al., 2012). Secondly, unlike the current test-focused science education system, NSCG considers students' attitudes toward science and their higher-order thinking skills, incorporating these aspects as crucial learning outcomes within the framework of the new science curriculum. In addressing the challenge of bridging the gap between classroom learning and real-world application, NSCG promotes active engagement through hands-on practice and project-based exploration, fostering connections between learners and their surroundings through collaborative problem-solving endeavors. Lastly, through the successful implementation of the new 3science curriculum, we aim to raise students' awareness of global sustainability. The next task to be undertaken will be the measurement of students' achievements with respect to these NSCG. Through this, we can evaluate the consequence of this curriculum reform and develop information that will allow for continued and sustainable improvements in student learning.

DEVELOPING AN ASSESSMENT ACCORDING TO NEW SCIENCE CURRICULUM GUIDELINES

The assessment of SL is important but neglected in NSCG to a certain extent. In this Taiwanese new science curriculum reform, the purpose of NSCG fit the current trends in science education policy that emphasize the importance of SL as a transferable outcome of science education. NSCG provides a concrete framework of curriculum to guide the edition of the learning materials such as the scientific textbooks. Meanwhile, these guidelines also inform teachers the related pedagogies to teach the targeted skills and contents mentioned in the essential literacies. However, NSCG do not make it clear to address on how to assess students' proficiency in using SL skills to solve confronting problems in real-world situation.

Currently, several measures for SL are available. Fives et al. (2014) published a validate assessment for measuring middle school students' scientific literacies. Gormally et al. (2012) developed a test of SL skills to provide reliable measurement of undergraduates' proficiencies in using SL skills and the impact of the curricula reform. Kuo (2015) developed a multimedia-based scientific inquiry assessment to assess over 1,000 8th and 11th grades, and it offered precise measuring results. These scientific literacy assessments (SLAs) may provide a solid referencing basis for the development of a ready means to assess competencies mentioned in the framework of NSCG.

Indeed, Taiwan's most currently adopted scientific tests draws on some degree of complex knowledge of specific science disciplines. In addition, students' attitude toward science are usually not included or neglected in most measures (Tsai & Kuo, 2008). For example, please refer to the following item description:

> "The electrically neutral magnesium atom (^{25}Mg) loses two electrons to become a magnesian ion, please compare the size of proton number, neutron number and electron number in a magnesian ion."

This item was selected form the "basic scholastic assessment", which is administrated annually in Taiwan to assess nationwide junior graduates' learning results. The item described is largely information-dependent and students must have sufficient scientific information to respond accurately to it. We admit that it is important to assess student understanding of specific science topics or information, since the students are not possible to be scientifically literate if they do not possess any scientific knowledge (Roberts, 2007). However, in the past few years, the scientific assessments focused mostly on testing students' understanding of scientific contents of specific fields, such as asking students the contents of Newton's law of universal gravitation. In addition, most students were overburdened with the wide scope of the examinations that covered the basic contents of main findings in biology, chemistry, physics and earth sciences. These, field specific and content-focused, testing methods have both influenced teachers and students. In this situation, teachers tended to catch up the progress of curriculums by engaging in surficial/conceptual level coverage of a wide range of science topics. Hence, students are merely allowed to leisurely experience processes of science and focus deeply on a few central scientific concepts (Lambert, 2006). Therefore, the new scientific-literacy-oriented curriculum reform is emerging in response to shift learning focus from field specificity to cover materials that transcend specific fields. In addition, the focus also shifts from emphasizing on acquiring knowledge of scientific findings, principles and laws to cultivate individuals' abilities to use scientific information in real-world situations beyond the classroom.

Coincident with this shift is a requirement in finding ways to assess students' development of SL skills. The assessments must closely align with the scope of the new science curriculum standards and are capable of measuring students' learning outcomes of "the essential literacies", which covered both "the critical learning performances" and "the critical learning contents". Measurement of students' motivation for, attitude toward and beliefs about science are necessary. DeVellis (1991) outlined a set of specific guidelines for the measurement development. The first suggested step is to define clearly the construct being measured. Hence, we start from referring back to the framework of "the critical learning performances and contents" mentioned above with a detailed manner to generate several components of SL to assess.

FRAMEWORKS FOR ASSESSING SCIENTIFIC LITERACY

The frameworks are similar to curriculum guidelines that setting out goals for an educational system and organize the subject domain into categories based on some organizing principles (Moseley et al., 2005). In addition, the frameworks usually guide and inform science assessors and educators about the science learning domain (Kind, 2013). In this article, the framework for assessing SL can be a matrix of **Table 1** and **Table 2** mentioned before. Since these two tables have clearly set out goals for the students to achieve to become scientifically literate, the educator needs to conclude these learning goals as the constructs of the assessment. In this manner, the framework developed appears to be very similar to the traditional content-behavior matrix (see **Table 4**). The behavior dimension consists of two constructs that represent critical learning performances, including scientific cognition and inquiry abilities. The dimension name is "the critical learning performances". The constructs and name of content dimension are in consistent with "critical learning contents".

Specifically, the "attitude toward science and nature of science" become an "overarching" dimension in this framework because of the content-free nature of this

Table 4. Scientific literacy assessment framework

			Learning contents		
			Components & characteristics of nature world	Phenomenon & mechanism of nature world	Sustainable development of nature world
Learning performances	Scientific cognition	Knowledge comprehension application analysis evaluation creation	Competencies	Competencies	Competencies
	Inquiry abilities	Thinking ability Problem-solving	Competencies Competencies	Competencies Competencies	Competencies Competencies
Attitude toward science & nature of science					

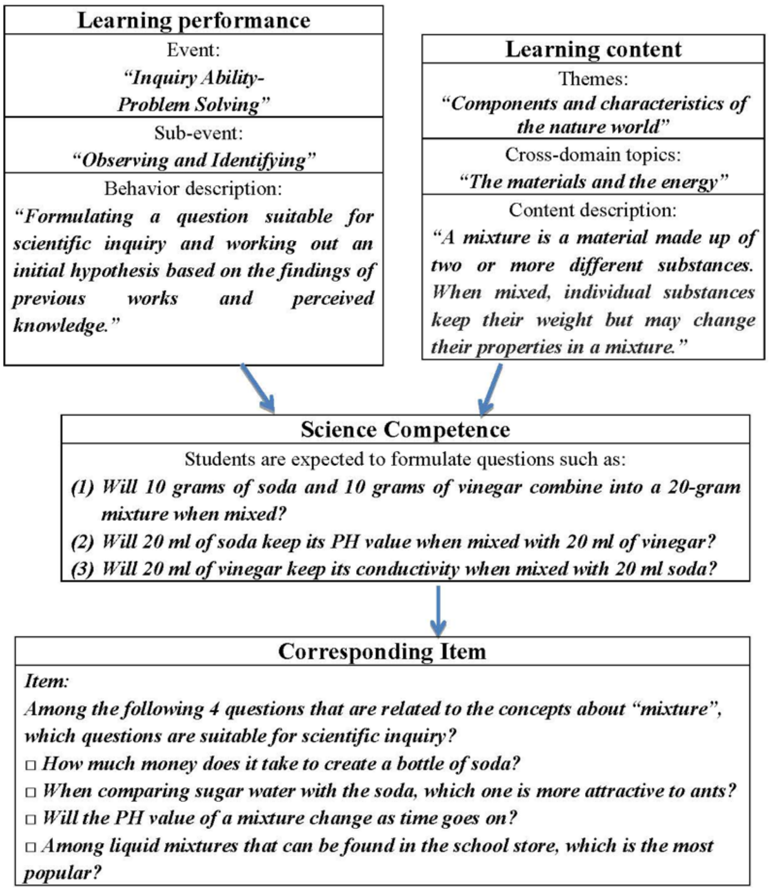


Figure 1. In a situation aiming at 5th & 6th grade, this concept map shows how science competency is organized by combining conceptual knowledge & science practice (it also provides an example of how a corresponding item is generated) (Source: An item selected from SLA)

construct. Since this construct concerns more of the individuals dispositions like interest in science or habit of scientific thinking, it should act like a "cause" to influence students' competency to "do science"(Kind, 2013). Designing a scale according to this construct will help to reveal the perceptions of the respondents. Hopefully, this arrangement will explain more explicitly the relationships between the dimensions.

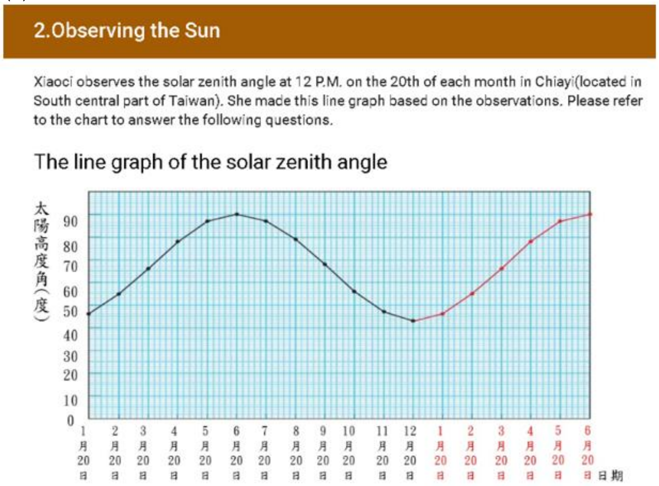
The *competency* adopted is to represent the combination of performance and content dimensions as the meaning of "successful performance of science task or activity".

This is very much like the achievement level dimension named "*performance expectations*" used in the framework of the *national assessment of educational progress*. The combination of science contents and science behavior will generate performance expectations, and assessment items can be developed based on these performance expectations (NAGB, 2008). Similarly, the "competencies" used in our framework represent a description of students' expected and observable critical learning performances that are embedded in the critical learning contents. Corresponding item will be generated based on the competency. Since NSCG stress the application of students' developed scientific literacies in dealing with daily situations involving science and technology (MoE, 2016), we will arrange item similarly to corresponding task students should be able to handle in everyday life. We provide an example in **Figure 1**.

The item provided in **Figure 1** focuses mainly on measuring students' competency to formulate a research question with regard to the concept of "mixture". It is worth to noting that one item is suggested to measure two things, including one of the content topics and one of the cognitive domains (Martin et al., 2015). Therefore, a single item will not be developed to assess all the performances and contents described. In this manner, our sample item does not assess the students' abilities to form hypotheses or their knowledge of the properties of mixtures, although these abilities and knowledge are also included in their competencies. More items will be generated for possible inclusion of these parts.

In conclusion, developing a framework of SLA will allow the assessor to confirm if the purposes of the assessment match up with the goals of the curriculum. In addition, the framework offers a rationale for the science domain helps assessors maintain construct validity when developing items (Kind, 2013). Generating an item pool for each specification presented in the framework is the following steps, and it has been initially completed. In order to check the items for accuracy and clarity, we started a pilot test of SLA, and

Table 5. Representative items for measuring competencies of (1) building a model to describe result of an experiment about sound & (2) analyzing data from line graphs of solar zenith angle

Item	Item description
(1) 	(Before answering questions, students were asked to wear headphones & watch an animation about "an experiment of sounding fishing lines, to make sound" for about 50 seconds.) *After watching experiment in animation, please pick a correct model to describe experimental results:* *(a) The thicker the fishing line, the lower-pitched the tone.* *(b) The longer the fishing line, the lower-pitched the tone.* *(c) The looser the fishing line, the lower-pitched the tone.* *(d) All of the above.*
Learning performances	Inquiry abilities-thinking abilities-modeling
Learning contents	Phenomenon & mechanism of nature world-sound
Competences	Students are able to build a model to describe result of an experiment about sound
(2) 	*After reading line graph, please analyze relationship between solar zenith angle & seasons to select a correct description from following:* *(a) The season with the highest solar zenith angle is spring; autumn is the season with the lowest solar zenith angle.* *(b) The season with the highest solar zenith angle is summer; autumn is the season with the lowest solar zenith angle.* *(c) The season with the highest solar zenith angle is summer; spring is the season with the lowest solar zenith angle.* *(d) The season with the highest solar zenith angle is summer; winter is the season with the lowest solar zenith angle.*
Learning performances	Inquiry abilities-problem-solving-**analyzing**
Learning contents	Components & characteristics of nature world-**solar zenith angle**
Competences	Students are able to analyze data from graphs to find relationship between solar zenith angle & seasons

more details will be listed in the next paragraph. Finally, building up a framework of the assessment is the first step, before finishing the development of a valid, reliable assessment suitable for measuring the learning outcomes of this new science curriculum reform. There is still a very long way to go.

INITIAL PILOT TESTING SCIENTIFIC LITERACY ASSESSMENT

The authors developed a pool of 34 questions for possible inclusion in the final edition of SLA. There are 27 multiple-choice questions with four possible answers and seven short-answer questions. At least one question was generated for each construct presented in the framework. **Table 5** listed two representative items for assessing students' competencies of building a model and analyzing the data from line graphs. A team of four reviewers consisting of one professor and three elementary school teachers, all with a background in science education reviewed the items for the alignment with the targeted competences. Each of these reviews had at least five years of teaching experience of science education and had a good understanding of inquiry and problem-solving abilities.

An initial pilot test consisting of 34 questions was administered to 104 6th grade elementary school students in south part of Taiwan. The test was administrated one week before their graduation. This time was chosen to confirm that the test takers have completed all science courses in elementary school. The overall mean score for this test was 14.4 (42.35%) with a standard deviation of 5.4 and a standard error of measurement of 0.53.

The discrimination index (D) ranged from 0.01 to 0.55 with 34 items demonstrating Ds of 0.35. There were nine items with D values below 0.30. Using guidelines for evaluating items based on the D values (Reynolds et al., 2006), items with a D of 0.40 were considered as very strong, 0.30-0.40 as good, and below 0.30 as needing revision, to discriminate participants from the top and bottom percentiles. Therefore, in the pilot test, items with *D*s lower than 0.30 were then be deleted or revised based on some criteria, including the clarity of item descriptions, the feedback from the participants and the distribution of responses to each distractor.

On the other hand, KR20 reliability coefficient for pilot test was 0.78, which as an acceptable value of alpha ranging from 0.70 to 0.95 (Tavakol & Dennick, 2011). This value implied all the items in the test measured the same construct named SL to an acceptable extent. However, since SL in the assessment framework was defined to contain more than one construct including thinking abilities and problem-solving abilities, in principle, the reliability coefficient should be calculated for each of these constructs to prove the internal consistency of the concluded items.

Finally, the mean item difficulty was 0.44, which mean on average 44.00% of students completing an item gave the correct response. The ideal mean difficulty for the test should not deviate much from a value of 0.625, specifically for four-response multiple-choice questions (Wenning, 2006). Since the present pilot test contained seven short-answer questions, if only calculated the multiple-choice questions in the pilot test, the mean item difficulty was 0.47, which was still a lower value.

Overall, through initial finding of the above data for item analysis, the author needs to revise some items, and perhaps to create new items for the next pilot test to make sure the quality of SLA can be enhanced.

DISCUSSION

SLA appears to be a crucial component in evaluating the effectiveness of NSCG in Taiwan. Discussion about how SLA was constructed under NSCG, how NSCG affected SLA, how SLA was designed and applied, and what the results of SLA imply in relation to NSCG is summarized below. Meanwhile, the novelty of this study compared to other related studies is also addressed.

Construction of Scientific Literacy Assessment Under New Science Curriculum Guidelines

A previous study reveals that science education reform in Taiwan is a complex and multifaceted process, while emphasizing the crucial role of assessments in driving and evaluating reform's effectiveness (Wu et al., 2018). SLA was designed to align with the goals and objectives outlined in NSCG. NSCG appears to serve as a foundational document that sets the curriculum goals and categorizes the subject domain into various categories or organizing principles. SLA have used these categories and principles as the basis for its assessment framework. Specifically, SLA framework comprises two main dimensions: "learning performances" and "learning contents," which reflect the structure of NSCG.

Impact of New Science Curriculum Guidelines on Scientific Literacy Assessment

NSCG have a significant impact on shaping the content and structure of SLA. It not only provides goals and guidelines for science education but also serves as a reference point for SLA's content. NSCG's emphasis on SL, scientific cognition, and inquiry abilities seems to have influenced the development of SLA framework, which focuses on assessing these learning performances.

Design & Application of Scientific Literacy Assessment

SLA is designed as an assessment tool to measure students' SL and their ability to apply scientific knowledge and skills in real-life situations, as emphasized in NSCG. SLA framework includes various constructs and dimensions, such as "attitude toward science and nature of science," which is considered an overarching dimension due to its content-free nature. This dimension reflects NSCG's emphasis on nurturing students' interest in science and scientific thinking.

Implications of Scientific Literacy Assessment Results & Relation to New Science Curriculum Guidelines

The results of the initial pilot testing of SLA provide valuable insights into the effectiveness of the current science curriculum in Taiwan. The mean score of 14.4 (42.35%) suggests that there may be room for improvement in students' scientific literacies, as it indicates that on average, they answered less than half of the questions correctly. These results highlight the need to introduce NSCG.

D values are used to assess the quality of the assessment items. Items with lower *D* values (below 0.30) are considered in need of revision. This suggests that certain items may not effectively discriminate between high-performing and low-performing students. Addressing these issues is crucial for improving the quality of SLA.

KR20 reliability coefficient of 0.78 indicates an acceptable level of internal consistency for the overall test. However, it is noted that SL, as defined in NSCG, contains multiple constructs, including thinking abilities and problem-solving abilities. Further analysis is needed to assess the internal consistency of these constructs individually.

In conclusion, SLA is intricately linked to NSCG and serves as a tool to assess whether the curriculum goals outlined in NSCG are being met. The results of SLA pilot test suggest that there may be areas for improvement in science education, and adjustments may be needed to align more closely with NSCG's objectives. Ongoing assessment and refinement of SLA will be essential to ensure that it accurately measures students' SL and their ability to apply science in everyday life, in line with the goals of NSCG.

CONCLUDING REMARKS

While the concept of SL has become an internationally well-organized educational catchphrase and science learning goals, Taiwan's educational institutions follow the trend and carry out a comprehensive scientific curriculums reform. The reform aims at solving the problems arising from the present science educational systems such as

(1) students' low motivations for science learning,

(2) favoring of testing-oriented education, and

(3) not taking proper care of students' hands-on doing abilities.

NSCG are thereby developed in response to the current science education needs. NSCG underline the importance of teaching not just the specific content knowledge but also fulfilling the need to allow students opportunities to develop functional understandings and appreciations of science. Meanwhile, the progressive and cross-field ways of curriculum design will ensure comprehensive and age-appropriate learning among all kinds of students. In this paper, we have undertaken a comprehensive examination of the state of science education in Taiwan, with a particular focus on the recent efforts towards curriculum reform. Our objectives included providing an overview of the existing science curriculum, revealing the challenges it posed, analyzing NSCG, and developing a sample assessment aligned with the framework of SL.

Taiwan's education system has long been lauded for its success, but it has faced its share of challenges. Reduced student engagement, an overemphasis on test scores, and limited real-world application of knowledge have been identified as shortcomings. These issues have been compounded by broader societal changes, including a declining birth rate, aging population, and rapid technological advancements. To address these challenges, MoE (2014) initiated a comprehensive curriculum reform process in 2014, encompassing various domains, including science.

Central to this reform is NSCG, which places a strong emphasis on SL as a means to engage and motivate students. NSCG seeks to instill scientific inquiry skills, foster positive attitudes toward science, and impart fundamental scientific knowledge. The curriculum framework outlines three critical dimensions: inquiry skills, content knowledge, and attitudes towards science. One notable feature of NSCG is its progressive and cross-domain integrated approach, ensuring a seamless transition from elementary to high school. It emphasizes hands-on practice, scientific inquiry, and the application of scientific skills in real-world contexts.

To measure the success of this reform, we embarked on the development of a sample assessment aligned with SL framework. We created a pool of 34 questions encompassing various constructs of SL. The initial pilot test, administered to 6th grade students, revealed valuable insights. While the overall mean score was 14.4, indicating room for improvement, *D* varied across items. Items with *D* values below 0.30 were flagged for revision. The pilot test demonstrated acceptable reliability (KR20 coefficient=0.78) and highlighted the need for further refinement.

In conclusion, this study serves as an essential step towards the development of a new science curriculum and corresponding assessments in Taiwan. This study's novelty lies in its comprehensive analysis of Taiwan's science curriculum reform, its development of a context-specific assessment aligned with SL goals, and its emphasis on multidimensional assessment. It contributes to the broader field of science education by offering insights into curriculum reform and assessment practices in Taiwan while aligning with global trends in science education reform. As the curriculum reform continues, ongoing evaluation and refinement will be crucial to ensure the success of this transformative endeavor.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study was approved by the institutional ethics committee of Kang Hsuan Educational Publishing Group (Approval code: 234704166). All participating students were provided with a detailed explanation of the study's purpose and potential implications, and written informed consents were obtained from the participants.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

ACCT. (2013). *2013 Taiwan white paper*. www.amcham.com.tw

Aldridge, J. M., Fraser, B. J., & Huang, I. T. C. (1999). Investigating classroom environment in Taiwan and Australia with multiple research methods. *Journal of Educational Research, 93*, 48-62. https://doi.org/10.1080/00220679909597628

Chang, C. Y. (2005). Taiwanese science and life technology curriculum standards and earth systems education. *International Journal of Science Education, 27*(5), 625-638. https://doi.org/10.1080/0950069042000323764

DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching, 37*(6), 582-601. https://doi.org/10.1002/1098-2736(200008)37:6<582::AID-TEA5>3.0.CO;2-L

DeVellis, R. (1991). *Scale development: Theory and applications*. SAGE.

Fives, H., Huebner, W., Birnbaum, A. S., & Nicolich, M. (2014). Developing a measure of scientific literacy for middle school students. *Science Education, 98*(4), 549-580. https://doi.org/10.1002/sce.21115

Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a test of scientific literacy skills (TOSLS): Measuring undergraduates' evaluation of scientific information and arguments. *Life Sciences Education, 11*, 364-377. https://doi.org/10.1187/cbe.12-03-0026

Kuo, C. Y., Wu, H.-K., Jen, T. H., & Hsu, Y. S. (2015). Development and validation of a multimedia-based assessment of scientific inquiry abilities. *International Journal of Science Education, 37*(14), 2326-2357. https://doi.org/10.1080/09500693.2015.1078521

Lai, H. R., Chou, W. L., Miao, N. F., Wu, Y.P., Lee, P. H., & Jwo, J. C. (2015). A comparison of actual and preferred classroom environments as perceived by middle school students. *Journal of School Health, 85*, 388-97. https://doi.org/10.1111/josh.12263

Lambert, J. (2006). High school marine science and scientific literacy: The promise of an integrated science course. *International Journal of Science Education, 28*(6), 633–654. https://doi.org/10.1080/09500690500339795

Martin, M. O., Mullis, I. V. S., & Hooper, M. (2016). Methods and procedures in TIMSS 2015. *Boston College, TIMSS & PIRLS International Study Center*. http://timssandpirls.bc.edu/publications/timss/2015-methods.html

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). TIMSS 2015 International results in science. *TIMSS*. http://timss2015.org/wp-content/uploads/filebase/full%20pdfs/T15-International-Results-in-Science.pdf

MoE. (2014). *The general curriculum guidelines for twelve-year compulsory education*. https://www.naer.edu.tw/eng/

MoE. (2016). *The science curriculum guidelines for twelve-year compulsory education (the draft)*. https://www.naer.edu.tw/eng/

MoE. (2017). *The general curriculum guidelines for twelve-year compulsory education*. https://www.naer.edu.tw/eng/

MoE. (2018). *The science curriculum guidelines for twelve-year compulsory education*. https://www.naer.edu.tw/eng/

NAER. (2014). *The suggestions for curriculum development for twelve-year compulsory education*. National Academy for Educational Research.

National Assessment Governing Board. (2008, September). *Science framework for the 2009 National Assessment of Educational Progress*.

NRC. (1996). *National science education standards*. National Academies Press.

OECD. (2001). *Knowledge and skills for life: First results from PISA 2000*. OECD Publishing.

OECD. (2006). *Assessing scientific, reading, and mathematical literacy*. OECD Publishing.

Reynolds, C. R., Livingston, R. B., & Wilson, V. (2006). *Measurement and assessment in education*. Pearson.

Roberts, D. A. (2007). Scientific literacy/science literacy. In S. K. Abell, & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 729-780). Erlbaum.

Smith, K. V., Loughran, J., Berry, A., & Dimitrakopoulos, C. (2012). Developing scientific literacy in a primary school. *International Journal of Science Education, 1*(1), 127-152. https://doi.org/10.1080/09500693.2011.565088

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tsai, C.-C., & Kuo, P.-C. (2008). Cram school students' conceptions of learning and learning science in Taiwan. *International Journal of Science Education, 30*, 353-375. https://doi.org/10.1080/09500690701191425

Wang, Z. (2004). *An antidote to modern test-oriented education: Toward a constructive post-modern education* [Paper presentation]. The Forum for Integrated Education and Educational Reform.

Wenning, C. J. (2006). Assessing nature-of-science literacy as one component of scientific literacy. *Journal of Physics Teacher Education Online, 3*(4), 3-10.

Wu, H.-K., Kuo, C. Y., Jen, T.-H., & Hsu, Y. S. (2015). What makes an item more difficult? Effects of modality and type of visual information in a computer-based assessment of scientific inquiry abilities. *Computers & Education, 85*, 35-48. https://doi.org/10.1016/j.compedu.2015.01.007