



METHODOLOGY FOR DEVELOPING SCIENTIFIC LITERACY SKILLS THROUGH THE INTEGRATION OF HPL, STAR LEGACY AND GRASPS

Davletova Hamida Xasan qizi

A. Avloniy National Institute of Pedagogical Skills, Tashkent, Uzbekistan

Abstract

Remembering scientific concepts does not by itself ensure that school students can use them to make reasoned decisions in everyday situations. Scientific literacy requires connected capabilities to explain phenomena, plan investigations, interpret data and evidence, evaluate information sources, communicate conclusions and transfer learning. This article systematises a methodology for developing these skills on the basis of the first and second chapters of the author's dissertation. The contribution is theoretical and design-oriented, combining conceptual analysis, functional comparison of methodological approaches, pedagogical modelling and evidence-oriented task design. The methodology is contextualised for Grade 6 integrated science education. Four interconnected blocks are specified: goal-setting, content, organisational-activity, and diagnostic-outcome. HPL establishes the quality criteria of the learning environment; STAR Legacy organises problem-based inquiry; and GRASPS defines the authentic task, intended product and success criteria. Implementation follows six connected phases: diagnostic activation, problem formulation, investigation and information inquiry, evidence construction and checking, product development and defence, and transfer with reflection. A heat-retaining container task illustrates the connections among a testable question, fair comparison, data representation, an evidence-based recommendation and application in a changed context. Assessment combines a five-dimension analytic rubric with individual process evidence, graduated scaffolding and opportunities to revise work after feedback. The central principle is that neither an engaging experiment nor an attractive product is sufficient



evidence of scientific literacy. A defensible interpretation requires alignment among content, action, evidence, criteria, feedback and transfer. The article presents the conceptual structure and implementation conditions of the methodology without incorporating the experimental results or statistical effectiveness indicators reported in the dissertation's third chapter.

Keywords: Scientific literacy; HPL; STAR Legacy; GRASPS; integrated science education; inquiry-based learning; authentic tasks; formative assessment; scaffolding; transfer.

Introduction

The practical value of school science cannot be reduced to students' ability to repeat information about nature. When interpreting a claim about water quality, energy use, plant growth or weather, learners need to select relevant knowledge, formulate a question that can be investigated and support a conclusion with evidence. Scientific literacy is therefore treated as an educational outcome integrating knowledge, ways of acting, attitudes and reflection [1–5]. Within this account, a skill is an observable action in a learning task, whereas competence concerns the mobilisation of knowledge and skills in context.

The pedagogical problem identified in the dissertation concerns a mismatch between the aim of applying science in life and the reproductive structure of some learning tasks. A student may define a concept correctly but struggle to decide which experimental condition should change or what a result actually supports. Likewise, a successful group presentation does not demonstrate every participant's scientific reasoning equally. These tensions establish the need to connect scientific content, student activity and assessment evidence within a coherent instructional system.

This problem is particularly relevant to Grade 6 integrated science, where earlier experience of observing and describing is extended towards causal explanation, identification of variables and evaluation of evidence sufficiency. International assessment frameworks provide a reference for specifying competencies associated with using knowledge in context [6, 7]. Their classroom use nevertheless requires adaptation of language, information load, resources and the



degree of independence. The present methodology does not directly transplant an international assessment scale into Grade 6 instruction.

HPL (How People Learn), STAR Legacy (Software Technology for Action and Reflection Legacy) and GRASPS address different parts of this problem. HPL aligns learners' prior knowledge, scientific content, assessment and collaboration. STAR Legacy organises activity from an initial challenge to investigation and public presentation. GRASPS specifies for whom, under which circumstances and against which criteria a product is developed [8, 12, 13]. Their integration is consequently understood as coordination of complementary instructional functions rather than the consecutive use of three independent techniques.

The purpose of the article is to systematise the theoretical and methodological solutions developed in Chapters I–II of the dissertation into a coherent methodology for scientific literacy development. Four questions guide the analysis: What is the operational content of the skills? How are HPL, STAR Legacy and GRASPS functionally connected? Through which phases is the task system implemented? What evidence is needed to interpret development? The reported results are conceptual and instructional design outputs rather than a report of a new classroom experiment.

Literature Review

An operational interpretation of scientific literacy

DeBoer interprets scientific literacy as a broad educational purpose related to historical and social needs [2]. Norris and Phillips examine the relationship between reading and writing scientific texts and using scientific knowledge [3]. Holbrook and Rannikmae emphasise the connection of scientific content with socially meaningful problems, justified choices and responsible action [4]. Together, these positions provide a conceptual basis for not treating familiarity with terminology as a sufficient outcome. In classroom design, this broad purpose has to be translated into observable actions.

Duschl's account of the relationship among conceptual, epistemic and social learning goals clarifies this transition [5]. Students learn what is known about a phenomenon, which evidence supports a conclusion and how that conclusion can be justified to others. The dissertation operationalises this relationship through



content, procedural and epistemic knowledge used in personal, local and global contexts. Searching for scientific information is similarly distinguished from merely locating a statement: learners need to explain its relevance and credibility for a decision [7].

Inquiry activity and graduated scaffolding

Pedaste and colleagues describe inquiry-based learning through relationships among orientation, conceptualisation, investigation, conclusion and discussion [9]. The synthesis by Furtak and colleagues and the methodological account of Hmelo-Silver, Duncan and Chinn indicate why inquiry should not be equated with unsupported freedom [10, 11]. Active student participation does not require the teacher to withdraw from instruction. A question prompt, variables table, partially completed graph or reasoning scaffold can direct attention towards the scientific decision that matters.

Lazonder and Harmsen's analysis of guidance and Bell, Smetana and Binns's classification of inquiry levels support the management of independence through task structure [22, 23]. The dissertation distinguishes supported, guided, independent and authentic-transfer levels of task performance. These levels do not assign students to permanent ability groups. Instead, they specify which parts of the question, method and solution the learner determines in a particular task. Support is reduced in response to stable evidence of the relevant skill rather than according to a fixed timetable.

Alignment of learning environment, authentic tasks and assessment

The learner-, knowledge-, assessment- and community-centred environments of HPL operate together [8]. Eliciting prior conceptions should connect with deepening conceptual knowledge, while collaboration should connect with assessment that makes individual understanding visible. Schwartz and colleagues' work on adaptive instructional design provides a basis for testing and revising ideas within STAR Legacy [12]. Wiggins and McTighe's approach, in turn, directs planning towards the intended result and the evidence that would demonstrate it before activities are selected [13].



Mislevy, Steinberg and Almond explain assessment through relationships among an inference about the learner, observable evidence and the task that elicits that evidence [14]. Work by Black and Wiliam, Hattie and Timperley, and Shute establishes the importance of feedback that informs a next action [15–17]. Completing a project or following instructions is therefore not sufficient evidence of scientific literacy. The scientific claim, relevant data, reasoning, limits of the conclusion and traces of revision need to be considered together.

Integrated content and adaptation to the national context

The integration of scientific practices, crosscutting concepts and core ideas clarifies what interdisciplinary learning entails [18]. The national science curriculum, Sangirova's work on learning projects and Ergasheva's research on literacy monitoring are discussed in the dissertation as local foundations for integrated instruction and assessment [19–21]. This article treats them as complementary contributions to selecting content, organising activity and monitoring development rather than as interchangeable solutions.

The central methodological implication of the dissertation's international comparison is to adapt functional mechanisms rather than copy a national system. Locally meaningful modules, guided inquiry, recurring scientific practices and collaborative lesson analysis are used in this sense. Lesson Study is distinguished as a form of teachers' evidence-informed instructional improvement rather than a student task procedure [29]. Differences in equipment, time and access to information should not remove the connection between a scientific question and an evidence-supported conclusion.

Methodology and Research Design

Research design. The primary material comprises Chapters I and II of Davletova's dissertation. The theoretical and design-oriented analysis aligns the concepts, methodological approaches, pedagogical tensions, model blocks, tasks and assessment instruments presented in those chapters. It is not reported as an independently conducted systematic literature review: the literature discussion is bounded by the sources considered in the dissertation. The participant sample,



experimental stages and numerical results from Chapter III are outside the evidence base of this article.

Analytical procedure. Scientific literacy was first decomposed into observable learning actions and then connected with content, procedural and epistemic knowledge. The third step distinguished the functions of HPL, STAR Legacy and GRASPS; the fourth aligned task types and implementation phases with the target skills. The fifth examined the question–method–data–conclusion relationship in the heat-retaining container example. Finally, each intended outcome was checked for corresponding evidence, criteria, support, feedback and transfer opportunities. This procedure describes how the dissertation material is reorganised for the article.

Design criteria. Internal alignment is examined through preservation of scientific content, a necessary context, observable action, age-appropriate support, individual accountability, safe participation, feedback use and transfer to a changed situation. This check concerns the logical relationship between concepts and instructional decisions. It does not substitute for empirical effectiveness, inter-rater reliability or statistical evidence of a causal effect.

Research Results

Result 1: an operational structure of skills and the pedagogical model

The first result expresses scientific literacy through five connected areas of activity. They are separated for diagnosis and feedback but function together in an authentic task. When recommending a heat-retaining design, for example, a student explains the phenomenon, selects a comparison method, interprets measurements and identifies the limits of the conclusion. Table 1 connects these areas with observable performance and evidence sources.



Table 1. Operational structure of scientific literacy skills

Activity area	Knowledge focus	Observable indicator	Performance evidence
Scientific explanation	Concepts, models and cause–effect relations	Selects a relevant scientific idea; distinguishes observation from interpretation.	Explanation, diagram or model; initial and revised response.
Investigation design	Question, hypothesis, variables, method and safety	Formulates a testable question and specifies fair-comparison conditions.	Investigation plan, variables table and method rationale.
Data and evidence interpretation	Records, units, graphs, trends and limitations	Represents data accurately and selects evidence relevant to the conclusion.	Original records, graph and evidence-supported conclusion.
Information evaluation and decision	Sources, claims, credibility and alternatives	Checks supporting evidence; justifies a choice and identifies uncertainty.	Source comparison and justified recommendation.
Communication, reflection and transfer	Reasoning, revision and changed contexts	Answers with evidence and adapts a method to changed conditions.	Individual defence, revision explanation and transfer task.

These activities are situated within a four-block model. The goal-setting block specifies the scientific and practical performance to be developed. The content block connects scientific concepts with ways of investigating and justifying knowledge. The organisational-activity block coordinates context, inquiry, collaboration and instructional support. The diagnostic-outcome block interprets evidence from responses, processes and products to inform the next pedagogical decision. The blocks thus organise scientific reasoning rather than the transmission of a predetermined answer.

The central reasoning cycle is problem–question–hypothesis–investigation–data–evidence–decision–reflection. Its recursive character matters: new information may change a hypothesis, inconsistent data may prompt a methodological revision, and an audience question may qualify a conclusion. The outcome is therefore not only a product but also an explanation of which idea changed and why. Figure 1 shows the model blocks and their feedback relationship.

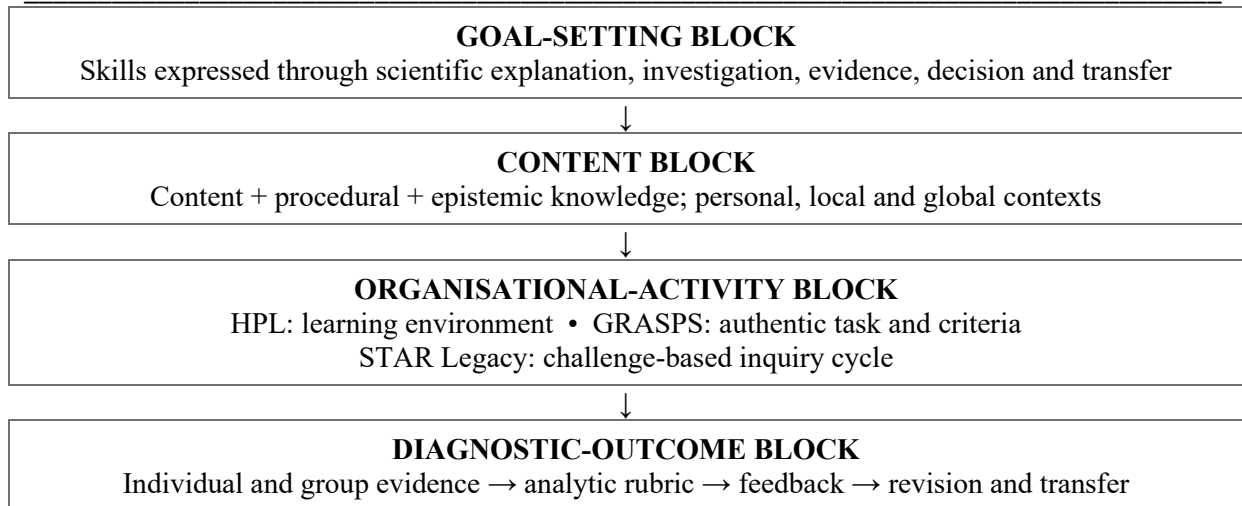


Figure 1. Pedagogical model blocks and feedback relationship (adapted from).

Result 2: functional integration of HPL, STAR Legacy and GRASPS

The second result assigns distinct instructional functions to the three approaches. HPL governs learning-environment quality throughout the activity. Learner-centredness elicits prior conceptions and needs; knowledge-centredness connects activity with core ideas; assessment-centredness makes evidence of learning available; and community-centredness supports respectful scrutiny of ideas and evidence [8]. When one of these perspectives is missing, an engaging activity may remain disconnected from deeper understanding.

GRASPS specifies the Goal, Role, Audience, Situation, Product/Performance and Standards/Success criteria from the beginning of the task [13]. Students therefore receive more than an instruction to build a model: they understand whose problem the model addresses and how the scientific basis of the solution must be demonstrated. Making criteria explicit in advance reduces the risk of valuing the visual appeal of a presentation above its evidence.

STAR Legacy organises action within this authentic task. Challenge introduces the problem; Generate Ideas records initial conceptions; Multiple Perspectives introduces alternative explanations and information sources; Research and Revise connects investigation with reconsideration; Test Your Mettle provides an interim check; and Go Public requires the result to be justified to an audience [12]. When the interim check reveals insufficient evidence, the student returns to inquiry.



Readiness for presentation therefore depends on checking the evidential basis, not merely on having a completed product.

Reducing the integration to a linear sequence of HPL followed by STAR Legacy and then GRASPS would narrow its meaning. HPL criteria remain active, GRASPS goals and criteria orient the work, and STAR Legacy structures the process. This preserves the relationship among content, action, evidence, criteria, feedback and transfer. A digital platform may support this relationship but is not a prerequisite: printed sources, tables, simple equipment and group discussion can implement the same core mechanism.

Result 3: a task system and six-phase implementation methodology

The third result connects separate exercises into a task chain in which one output prepares the next. An initial prediction informs an investigation question, the plan guides data collection, data support a conclusion, and the conclusion informs a recommendation and transfer. Task design starts with the evidence needed for the target skill rather than with a text or worksheet format [14]. A task specification identifies the context, core scientific idea, target skill, question, resources, student decision, expected evidence and criteria. Table 2 summarises the principal functions in the system.

Table 2. A system of tasks oriented towards scientific literacy

Task type	Leading action	Expected evidence	Next instructional step
Diagnostic prediction	Express an initial idea and its reason.	Prediction, question and initial model.	Identify a misconception and the necessary support.
Explanation and modelling	Connect cause and effect with a scientific idea.	Explanation, model and its limitations.	Clarify a concept or test an alternative account.
Investigation design	Align the question, variables and method.	Plan and fair-comparison conditions.	Check safety and measurement procedures.
Experiment and data	Observe, measure, tabulate and graph.	Original records, trends and exceptions.	Address measurement or interpretation errors.
Multiple-source information	Compare claims and supporting evidence.	Credibility rationale and missing information.	Complete evidence and check relevance to the decision.
Argument and GRASPS product	Justify an evidence-supported solution to an audience.	Product, recommendation and individual conclusion.	Revise against criteria and defend again.
Transfer and reflection	Adapt the method to changed conditions.	New plan and explanation of a revised decision.	Choose the next task and support level.



Phase 1 – diagnostic activation. In the heat-retaining container situation, students predict which material will keep a drink warm for longer and explain their reasoning. The teacher does not turn this prediction into a final grade; it is used to identify the existing conception and the claim that needs investigation. HPL's learner-centred perspective starts here but continues through later adjustments to support. Each student's initial response is retained for comparison with the revised explanation.

Phase 2 – problem formulation. The question "Which container is better?" is refined into a comparison of water temperature after a specified interval, for example, 10–15 minutes, with equal starting temperature and volume. Students identify the factor being compared and the conditions that must remain constant. In GRASPS terms, the goal is to keep a drink warm at a school event without an electric heater; the role is a research team proposing a solution; the audience is the event organisers; and the product is a container model with an evidence-based recommendation. Scientific adequacy, safety, resource use and justification are made explicit in advance.

Phase 3 – investigation and information inquiry. Groups plan a comparison using identical containers, a thermometer, a timer and available wrapping materials. Cardboard, cloth, foil, cotton or paper may be considered, but the best option is not announced beforehand. Safety, starting conditions, recording procedures and opportunities for repetition are checked. Students record observations in a time–temperature table. They connect the explanation with an appropriate text or model without presenting a statement from that source as their own measurement.

Phase 4 – evidence construction and checking. The table is transformed into a graph or comparative explanation, and the conclusion follows a claim–evidence–scientific reasoning–limitation structure [24]. Students identify the result that supports their recommendation, consider any unusual value and state what the investigation does not demonstrate. Questions such as "Were the initial conditions equal?", "How many factors changed at once?" and "Is this evidence sufficient?"



structure the interim check. The example specifies an instructional design; it does not provide classroom measurements or a predetermined best solution.

Phase 5 – product development and defence. The group presents its container or wrapping model, measurement records and recommendation for use. The choice of material and the practical limits of the solution are explained. In addition to the group product, each student prepares a short individual conclusion or answers an evidence-focused question. Feedback from the teacher and peers is followed by an opportunity to revise the recommendation or explanation. Assessment thereby functions as a means of improving scientific quality rather than only delivering a final judgement.

Phase 6 – transfer and reflection. The situation changes to protecting a cold product from external heat or choosing a clothing material. Rather than copying the previous solution, the learner explains which scientific idea remains relevant, which conditions have changed and what further evidence is needed. Reflection is focused through questions such as “Which evidence changed my view?” and “What will I do differently next time?” These six phases represent the implementation methodology in the dissertation. They connect with the named STAR Legacy stages but are not simply a translation or an identical classification of them.

The system extends across content areas by returning to the same skill in different topics. Controlling variables can be revisited through seed germination, graph interpretation through changing shadow length, and source comparison through claims about local air quality. Such recurrence is not repetition of a topic; it makes the use of a scientific practice in new content visible.

Result 4: pedagogical conditions and evidence-based assessment

The fourth result connects implementation conditions with assessment evidence. The scientific problem, instructional support and feedback must become visible through what students actually do. Table 3 systematises the conditions described in the dissertation as operational mechanisms and records. They are requirements



for implementing the methodology; determining whether they have been met requires examination of actual classroom materials.

Table 3. Pedagogical conditions for implementing the methodology

Pedagogical condition	Operational mechanism	Documented evidence
Meaningful, bounded problem	The context requires a scientific decision; time, resources and safety are explicit.	Task specification, question and constraints.
Coherent interdisciplinary content	Concepts contribute to a common problem and product.	Concept map and integrated explanation.
Differentiated scaffolding	Common expectations are retained; prompts, tables and instructions vary.	Record of support and independently completed actions.
Collaboration with individual accountability	Individual conclusions accompany group roles.	Group product, personal explanation and question response.
Feedback followed by revision	Feedback identifies a specific next action and time is provided to revise.	Initial and revised versions; reason for the change.
Safe, equitable participation	Suitable equipment, understandable resources and alternative expression routes are available.	Safety instructions, participation and resource records.
Reflection and transfer	Conditions change; retained ideas and necessary adaptations are explained.	New task, individual reflection and next learning target.

Broad developmental criteria are distinguished from the rubric for a particular task. The model's cognitive, motivational, creative, practice-oriented and reflective-personal criteria describe development broadly. The task rubric assesses five performance dimensions: scientific explanation and modelling; investigation design; data and evidence; decision and argumentation; and communication, reflection and transfer. Interest in inquiry may be a motivational indicator, for example, but it cannot establish the scientific adequacy of an incorrectly designed investigation.

Each dimension is interpreted through four descriptors. At the initial level, the student describes the event or repeats a procedure without establishing a relevant scientific basis. At the partially developed level, some actions are correct, but gaps remain in the method, interpretation or reasoning. At the mostly correct



level, the substantive task is completed, although an alternative explanation or limitation remains insufficiently developed. At the independent and justified level, the learner connects evidence, scientific reasoning and the boundaries of the conclusion and explains the necessary adaptation. These descriptors refer to the quality of a particular performance, not to fixed qualities of the person.

Rubric descriptors should not be converted automatically into school grades or a single composite score. The dissertation's overall levels of excellent, good, satisfactory and unsatisfactory are not treated as identical to its task-specific formative descriptors numbered 1–4. A skill profile is retained first; where a final score is required, criteria and weights must be specified in advance for the purpose of the task [1, 25]. This prevents strength in scientific explanation from concealing difficulty in data interpretation.

Questions, hypotheses, investigation plans, measurement records, graphs, source comparisons, individual conclusions, products and revision histories are considered together. Feedback does not end with a general judgement such as “good work”; it identifies a next action, for example adding units, clarifying a controlled condition or selecting evidence relevant to a conclusion [15–17, 28]. Without an opportunity to apply that advice, the formative cycle remains incomplete. The teacher distinguishes conceptual, procedural, epistemic and communicative difficulties and selects support accordingly.

Discussion

The principal contribution lies not in multiplying method labels but in connecting existing approaches through explicit functions. HPL coordinates the environment, GRASPS clarifies purposes and evidence, and STAR Legacy structures student activity. Together, they make visible the task in which content is used, the evidence supporting its use and the criteria against which it is interpreted. This relationship prevents the entire methodology from being equated with one experiment or project lesson.

The methodology distinguishes surface realism from scientific authenticity. Assigning a “researcher” role or producing a colourful model does not itself demonstrate literacy. Authenticity appears in the investigability of the question, the fit between question and method, the evidence supporting the conclusion and



the relationship between the product and audience needs. Similarly, a complex digital tool is not automatically superior to a simple experiment. Physical and virtual laboratories are selected for the activity and evidence they support, with explicit discussion of how a model simplifies the phenomenon [26].

Differentiation concerns adaptation of the route to a common scientific expectation rather than reduction of that expectation [27]. Shared support makes goals, criteria and safety clear to everyone; optional support supplies key words, tables or graph templates; and temporary intensive assistance helps analyse a difficult stage collaboratively. Students may express reasoning orally, in writing, through tables or through models while preserving scientific accuracy. Group assistance nevertheless does not replace evidence of individual reasoning.

The “5 T” interpretation discussed in Chapter II–Ta’lim, Tarbiya, Tadqiqot, Tijoratlashtirish and Tizimli takomillashtirish—provides a complementary substantive perspective. These refer to education, value formation, research, commercialisation and systematic improvement. In the dissertation’s school-level didactic interpretation, “commercialisation” means turning a scientific idea into a useful solution or product rather than engaging pupils in commercial transactions. Knowledge, responsible resource use, investigation, a justified product and revision in the container task can exemplify this interpretation. It neither replaces the functions of HPL–STAR Legacy–GRASPS nor constitutes a separate empirical finding.

Practical Implications

Implementation can begin with a small number of explicit indicators within one topic. Formulating a testable question, specifying fair-comparison conditions and selecting evidence relevant to a conclusion can be observed in a single task chain. A later chain can strengthen source evaluation or transfer. Mapping goals, evidence, and criteria for each task helps limit unnecessary activity while preserving its scientific focus.

Materials can offer minimal analogue, available-laboratory and digital versions for the same scientific goal. The analogue version relies on printed texts, simple containers, measuring tools and recording sheets. Digital use is justified by a function such as representing data or modelling a process that is difficult to



observe directly. Differences in access to technology should not exclude a student from evidential reasoning. Safety instructions and meaningful participation remain necessary in every version [1, 26].

Teacher collaboration can review scientific accuracy, language difficulty, completion time, plausible responses and assessment criteria. Open tasks should allow a scientifically justified solution that differs from the model answer. Joint discussion of the same student work and the reasons for assigning a descriptor can support assessment moderation. Retaining initial responses, amended plans and reasons for revision alongside the final product in a portfolio provides a clearer basis for the next support decision.

Limitations and Future Research

The limitations follow from the source and research boundaries: only the theoretical and methodological material of Chapters I–II is synthesised. No new participant study is conducted within this article, and optimal task duration, criterion weights and effect sizes are not established. The model and examples are consequently not conclusive evidence of pedagogical effectiveness. This delimitation does not deny the research reported elsewhere in the dissertation; it identifies the material on which this article rests.

Further research needs to examine task content, comprehensibility, cognitive load, feasibility under different resource conditions and consistency of rubric use. Repetition of an identical task should be distinguished from transfer to a new context. Agreement among group products, individual responses and teacher observations also requires investigation. Empirical comparisons should consider equivalence of content and instructional time, prior preparation and variation in teacher support.

Conclusion

Scientific literacy involves more than knowing scientific concepts. It is expressed in their use to explain phenomena, design investigations, evaluate data and sources, make evidence-supported decisions and transfer learning. Analysis of the dissertation's first and second chapters shows how this outcome can be designed



through interconnected goal-setting, content, organisational-activity and diagnostic-outcome blocks.

HPL–STAR Legacy–GRASPS integration coordinates the environment, process and authentic outcome. HPL informs the conditions for learning throughout; GRASPS specifies the task purpose and evidence; and STAR Legacy organises inquiry and revision. The six-phase task chain makes it possible to trace the learner’s progress from an initial conception to an evidence-supported product and transfer.

The practical value of the methodology lies in adapting support while retaining scientific expectations, combining collaboration with individual accountability, interpreting process and product evidence through an analytic rubric, and following feedback with revision. Its central design principle is that an engaging activity alone is insufficient: coherence among content, action, evidence, criteria, feedback and transfer provides the methodological basis for purposeful scientific literacy development.

References

1. Davletova H.X. Tabiiy-ilmiy savodxonlikni shakllantirish – zamonaviy umumiy o‘rta ta’limning dolzarb pedagogik muammosi // Maktab ta’limi: muammolar, izlanishlar, yechimlar. – 2026. – 1-son. – B. 121–129.
2. 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.
3. Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87(2), 224–240.
4. Holbrook, J., & Rannikmae, M. (2009). The meaning of scientific literacy. *International Journal of Environmental & Science Education*, 4(3), 275–288.
5. Duschl, R. A. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, 32(1), 268–291.
6. OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. Paris: OECD Publishing. DOI: 10.1787/b25efab8-en.



-
7. OECD. (2026). PISA 2025 Assessment and Analytical Framework. Paris: OECD Publishing. DOI: 10.1787/86c36975-en.
 8. Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). How People Learn: Brain, Mind, Experience, and School. Expanded edition. Washington, DC: National Academy Press.
 9. Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14, 47–61.
 10. Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, 82(3), 300–329.
 11. Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99–107.
 12. Schwartz, D. L., Lin, X., Brophy, S., & Bransford, J. D. (1999). Toward the development of flexibly adaptive instructional designs. In C. M. Reigeluth (Ed.), *Instructional-Design Theories and Models* (Vol. II, pp. 183–213). Mahwah, NJ: Lawrence Erlbaum Associates.
 13. Wiggins, G., & McTighe, J. (2005). *Understanding by Design*. Alexandria, VA: ASCD.
 14. Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1(1), 3–62.
 15. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
 16. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
 17. Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.
 18. National Research Council. (2012). *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press. DOI: 10.17226/13165.
-



-
19. O‘zbekiston Respublikasi Xalq ta’limi vazirligi, Respublika ta’lim markazi. (2021). Umumiy o‘rta ta’limning milliy o‘quv dasturi: Tabiiy fanlar, 1–6-sinflar. Toshkent.
 20. Sangirova, Z. B. (2022). Umumta’lim maktablarida STEAM yondashuvi asosida o‘quv-loyiha ishlarini tashkil etish metodikasi (tabiiy fanlarni o‘qitish misolida). Pedagogika fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi. Chirchiq.
 21. Ergasheva, M. T. (2023). O‘quvchilarning tabiiy-ilmiy savodxonligi monitoringida baholash dasturlaridan foydalanish metodikasi (PISA, TIMSS misolida). Pedagogika fanlari doktori (DSc) dissertatsiyasi. Toshkent.
 22. Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, 86(3), 681–718.
 23. Bell, R. L., Smetana, L., & Binns, I. (2005). Simplifying inquiry instruction: Assessing the inquiry level of classroom activities. *The Science Teacher*, 72(7), 30–33.
 24. McNeill, K. L., & Krajcik, J. S. (2011). *Supporting Grade 5–8 Students in Constructing Explanations in Science: The Claim, Evidence, and Reasoning Framework for Talk and Writing*. Boston, MA: Pearson.
 25. 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.
 26. de Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science*, 340(6130), 305–308.
 27. Tomlinson, C. A. (2014). *The Differentiated Classroom: Responding to the Needs of All Learners* (2nd ed.). Alexandria, VA: ASCD.
 28. Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers’ informal formative assessment practices and students’ understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44(1), 57–84.
 29. Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of Lesson Study. *Educational Researcher*, 35(3), 3–14.