JoSSEd
Journal of Science and Science Education
ISSN: 2774-1869 eISSN: 2774-2105

Teaching at the Right Level in Science Education: A Systematic Review of Evidence on Conceptual Understanding and Science Process Skills

Qasthalani Dzati Khulwani[1*], Ramlawati[1], Eda Lolo Allo[1]

[1] Graduate Program of Science Education, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Indonesia

Received: August 12, 2026
Revised: October 17, 2026
Accepted: October 28, 2026
Published: October 31, 2026

Corresponding Author:
Qasthalani Dzati Khulwani
Khulwani1501@gmail.com

DOI: 10.29303/jossed.v7i2.16410

Open Access

© 2026 he Authors. This article is distributed under a (CC-BY License)

Abstract: Teaching at the Right Level (TaRL) is an adaptive instructional approach that aligns instruction with students' demonstrated competence rather than solely with age or grade level. This systematic review synthesizes empirical evidence on TaRL in science education, with primary emphasis on conceptual understanding and Science Process Skills (SPS). Literature was searched through Google Scholar, GARUDA, SINTA, and relevant journal portals for studies published between 2016 and 2026. The review followed PRISMA 2020 principles. Study quality was appraised using five criteria covering clarity of objectives, research design, participants and context, instruments and procedures, and adequacy of data analysis and reporting. Twenty-five empirical studies met the eligibility criteria and were synthesized using a narrative-comparative approach. Of these, three studies directly examined conceptual understanding and one directly investigated SPS, while the remaining studies mainly reported general science/IPAS learning outcomes, diagnostic assessment, engagement, or related supporting outcomes. The direct evidence indicated favorable trends in conceptual understanding and SPS, but the evidence base remains limited. Because many included studies used classroom action research or one-group pretest-posttest designs, the findings should be interpreted as promising patterns rather than definitive evidence of effectiveness. Further controlled studies using validated outcome measures are needed.

Keywords: Conceptual understanding; Science education; Science process skills; Systematic review; Teaching at the Right Level

Introduction

Science education is expected to develop students' conceptual knowledge together with their ability to investigate phenomena, interpret evidence, solve problems, and communicate scientifically. Achieving these goals is challenging because students within the same classroom often demonstrate substantial differences in prior knowledge, learning readiness, and pace of learning. Uniform instruction may therefore exceed the readiness of some learners while providing insufficient challenge for others, potentially widening differences in learning achievement.

Teaching at the Right Level (TaRL) addresses this challenge by organizing instruction according to students' demonstrated competence rather than solely according to age or formal grade level. TaRL typically begins with diagnostic assessment, followed by ability mapping, targeted instruction, formative assessment, and flexible regrouping. Evidence from targeted instruction has shown that aligning instructional content with students' current levels can support learning, particularly among students who have not yet mastered expected competencies (Banerjee et al., 2016; Banerji & Chavan, 2020; Wolf et al., 2026).

This principle is particularly relevant to science education because many scientific concepts are hierarchical and depend on prerequisite knowledge. Inadequate understanding of foundational concepts can interfere with students' ability to interpret more complex phenomena and may contribute to persistent misconceptions. Diagnostic assessment is therefore important for identifying prerequisite gaps and misconceptions before instruction (Soeharto, 2021;

How to Cite:
Khulwani, Q. D., Ramlawati, R., & Allo, E. L. (2026). Teaching at the Right Level in Science Education: A Systematic Review of Evidence on Conceptual Understanding and Science Process Skills. *Journal of Science and Science Education*, *7*(2), 69–85. https://doi.org/10.29303/jossed.v7i2.16410

Soeharto et al., 2022). Within a TaRL framework, the results of such assessment can be used to differentiate explanations, task complexity, inquiry support, learning resources, and the level of scaffolding provided to students.

Conceptual understanding is one of the core outcomes of science learning. It extends beyond recalling definitions and requires students to explain, connect, interpret, exemplify, and apply scientific concepts in meaningful contexts. Emerging empirical studies of TaRL in science education have reported favorable changes in conceptual understanding at both junior-secondary and elementary levels (Nurfadilah et al., 2024; Ramlawati et al., 2026; Yulianto et al., 2026). However, the number of studies directly examining conceptual understanding remains limited relative to the broader body of TaRL research on general learning outcomes.

Science Process Skills (SPS) represent another fundamental dimension of science education. These skills include observing, classifying, measuring, predicting, formulating hypotheses, identifying variables, conducting investigations, interpreting data, and drawing evidence-based conclusions. Inquiry-oriented learning has consistently been associated with the development of SPS (Patonah et al., 2020; Numa & Martini, 2022; Suryawanto et al., 2022; Amalia et al., 2023). Within a TaRL framework, inquiry activities can be differentiated according to students' readiness, allowing learners who require greater support to receive more structured guidance while more advanced learners engage in increasingly independent investigation. Direct empirical evidence connecting TaRL with SPS is nevertheless still scarce; one notable example is the TaRL-based Differentiated Science Inquiry study reported by Ramlawati et al. (2025).

Recent research demonstrates that TaRL has been implemented across a growing range of science and IPAS contexts and has been combined with instructional approaches such as Differentiated Science Inquiry, Problem-Based Learning, Discovery Learning, differentiated instruction, and digital learning media. This growing diversity of applications indicates that TaRL is increasingly being used not as a stand-alone teaching model, but as an adaptive instructional framework that determines the level of content, task complexity, and support provided to students.

Several literature reviews have already examined TaRL, but their analytical scope differs from the present review. Ananda et al. (2024) reviewed TaRL as a potential solution for middle-school education and emphasized its principles, implementation strategies, achievement outcomes, and alignment with broader educational reform. Their review was not specifically organized around discipline-specific science outcomes. Shabrina et al. (2024) examined TaRL in relation to mathematics learning outcomes, placing the emphasis on achievement within mathematics rather than on science competencies. More recently, Putri et al. (2025) systematically reviewed TaRL in science learning using 18 studies published between 2018 and 2024. Their synthesis focused broadly on the effectiveness of TaRL across outcomes such as learning achievement, motivation, engagement, and communication skills.

These previous reviews provide important evidence regarding the broader value and implementation of TaRL, but they leave two analytical gaps. First, they do not clearly distinguish between evidence derived from general achievement measures and evidence directly measuring discipline-specific science competencies. Second, they do not explicitly examine how diagnostic assessment functions as a mechanism linking students' initial competence levels to differentiated instructional support and, subsequently, to conceptual and process-oriented outcomes. As a result, it remains unclear how much of the existing TaRL evidence directly supports conceptual understanding and Science Process Skills, rather than broader indicators such as overall achievement, motivation, or classroom participation.

The present review addresses these gaps by organizing the evidence into three analytical layers. The first layer comprises direct evidence, defined as studies that explicitly measure conceptual understanding or Science Process Skills. The second layer comprises supporting outcome evidence, including general science/IPAS learning outcomes, engagement, interest, and related classroom indicators. The third layer concerns implementation mechanisms, particularly the role of diagnostic assessment, ability grouping, differentiated scaffolding, and flexible regrouping in the TaRL process. This structure allows the review to distinguish the strength and type of evidence rather than treating all reported outcomes as equivalent.

Accordingly, the contribution of this review is not limited to selecting conceptual understanding and SPS as two outcome variables. Conceptually, the review positions TaRL as an adaptive instructional mechanism linking diagnostic evidence, instructional differentiation, and science-specific learning outcomes. Methodologically, it separates direct from supporting evidence and evaluates the extent to which different research designs can support claims about TaRL. This distinction is especially important because a substantial proportion of the available studies employ classroom action research or one-group pretest-posttest designs, which are informative for classroom improvement but provide more limited causal evidence than controlled experimental studies.

Based on this framework, the review addresses the following questions:

1. What are the characteristics and methodological patterns of empirical TaRL studies in science education?
2. How is diagnostic assessment used to determine learning levels and differentiated instructional support within TaRL?
3. What direct evidence is available regarding the relationship between TaRL and students' conceptual understanding?
4. What direct evidence is available regarding the relationship between TaRL and students' Science Process Skills?
5. How do general learning outcomes and other supporting indicators complement the direct evidence on conceptual understanding and SPS?
6. What methodological limitations characterize the current TaRL evidence base in science education?

Method

This study employed a Systematic Literature Review (SLR) to identify, evaluate, and synthesize empirical evidence on Teaching at the Right Level (TaRL) in science education. The review followed PRISMA 2020 for transparent reporting of identification, screening, eligibility, and study inclusion (Page et al., 2021), while PRISMA-S was used to guide the reporting of the literature-search process (Rethlefsen et al., 2021).

A systematic search was conducted through Google Scholar, GARUDA, SINTA, relevant journal publisher websites, and backward citation searching. The search covered publications from January 2016 to October 2026, with the final search conducted on 15 October 2026. Search terms were used in English and Indonesian and included combinations of "Teaching at the Right Level," "TaRL," "science education," "IPA," "IPAS," "conceptual understanding," "science process skills," "diagnostic assessment," and "science learning outcomes."

All retrieved records were checked for duplicates by comparing titles, authors, publication years, DOI information, and journal details. The remaining records were screened in two stages: title-and-abstract screening followed by full-text eligibility assessment using the predefined inclusion and exclusion criteria.

The screening process was conducted by one reviewer. To reduce subjectivity, potentially eligible articles were reassessed during the full-text stage using the same eligibility criteria. Ambiguous cases were resolved through repeated examination of the article against the predetermined criteria. Although the search covered 2016–2026, the eligible empirical studies were concentrated in 2024–2026, reflecting the recent development of TaRL implementation in Indonesian science and IPAS education.

Search Strategy

A systematic and reproducible literature search was conducted to identify empirical studies on Teaching at the Right Level (TaRL) in science education. The search was performed using Google Scholar, GARUDA (Garba Rujukan Digital), SINTA (Science and Technology Index), and relevant journal publisher websites. The search covered publications from January 2016 to October 2026, and the final search was conducted on 15 October 2026.

Search terms were developed based on the main concepts of the review, namely Teaching at the Right Level, science education, conceptual understanding, Science Process Skills (SPS), diagnostic assessment, and science learning outcomes. Search terms were used in both English and Indonesian to increase the coverage of national and international publications. Boolean operators AND and OR were used where supported by the database.

The principal search strings were:

1. ("Teaching at the Right Level" OR "TaRL") AND ("science education" OR "science learning" OR "IPA" OR "IPAS");
2. ("Teaching at the Right Level" OR "TaRL") AND ("conceptual understanding" OR "pemahaman konsep");
3. ("Teaching at the Right Level" OR "TaRL") AND ("science process skills" OR "keterampilan proses sains" OR "KPS");
4. ("Teaching at the Right Level" OR "TaRL") AND ("diagnostic assessment" OR "asesmen diagnostik");
5. ("Teaching at the Right Level" OR "TaRL") AND ("science learning outcomes" OR "hasil belajar IPA" OR "hasil belajar IPAS").

The following filters were applied whenever available: (1) publication period from 2016 to 2026; (2) journal articles; (3) publications in English or Indonesian; and (4) studies involving elementary/primary or junior-secondary science education. No restriction was initially applied to research design because classroom action research, pre-experimental studies, quasi-experimental studies, mixed-method studies, and qualitative empirical studies were potentially relevant to the review.

For Google Scholar, search results were sorted using the default "relevance" ranking. Because Google Scholar can return a very large number of loosely related records and does not provide stable reproducible result counts across searches, screening was restricted to the first 200 results for each search string, or all available records when fewer than 200 results were returned. The

same search strings were repeated in Indonesian and English where appropriate. Records were examined sequentially according to their relevance-ranked position. Studies appearing repeatedly across different search strings were recorded once and subsequently treated as duplicates during data management.

For GARUDA and SINTA, all records returned by the predefined search terms within the publication-year filter were examined. Journal publisher websites were additionally searched to verify bibliographic information and to identify potentially eligible studies not clearly indexed in the main databases. Backward citation searching of the reference lists of eligible studies was also conducted to identify additional relevant empirical studies.

For each search source, the search date, search string, applied filters, number of records retrieved or examined, and number of potentially eligible articles were documented in a search log. The search results from all sources were then combined before duplicate removal and screening.

Table 1. Literature Search Strategy and Search Log

Source and Search Terms	Search Procedure and Filters	Records Examined
Google Scholar — TaRL AND science education/science learning/IPA/IPAS	Search period: 2016–2026; English and Indonesian; sorted by relevance; first 200 results examined	200
Google Scholar — TaRL AND conceptual understanding/pemahaman konsep	Search period: 2016–2026; English and Indonesian; sorted by relevance; first 200 results examined	200
Google Scholar — TaRL AND science process skills/KPS	Search period: 2016–2026; English and Indonesian; sorted by relevance; first 200 results examined	200
GARUDA — TaRL; Teaching at the Right Level; TaRL IPA; TaRL IPAS	Journal articles published in 2016–2026; all retrieved records examined	68
SINTA — TaRL; Teaching at the Right Level; IPA; IPAS	Journal publications from 2016–2026; all retrieved records examined	42
Journal publisher websites — article title, author, and TaRL-related terms	Relevant journals; targeted verification and supplementary searching; publication period 2016–2026	25
Backward citation searching — reference lists of eligible articles	Relevant empirical studies were screened manually; searching conducted up to 15 October 2026	15
Total		750

The systematic search resulted in 750 records being examined across all search sources. Google Scholar contributed 600 records, consisting of the first 200 relevance-ranked results for each of the three predefined search themes: TaRL in science education, conceptual understanding, and Science Process Skills. In addition, 68 records were examined through GARUDA, 42 through SINTA, 25 through targeted searches of journal publisher websites, and 15 through backward citation searching. All searches covered publications from 2016 to 2026, with the final search conducted on 15 October 2026.

Table 2. Inclusion and Exclusion Criteria

Aspect	Inclusion Criteria	Exclusion Criteria
Publication type	Empirical research articles published in peer-reviewed journals	Review articles, systematic reviews, editorials, books, book chapters, theses/dissertations, conference abstracts, opinion papers, and non-empirical publications
Publication year	Studies published between 2016 and 2026	Studies published before 2016
Language	Articles published in English or Indonesian	Articles published in other languages when sufficient information could not be evaluated
Education level	Studies involving elementary/primary or junior-secondary students	Studies involving senior-secondary students, university students, teachers only, or non-school populations
Educational context	Studies conducted in science, integrated science-social studies (IPAS), or science-education contexts	Studies conducted exclusively in unrelated subject areas
TaRL implementation	Studies that explicitly implemented Teaching at the Right Level (TaRL), either as the main instructional approach or	Studies that did not explicitly implement TaRL

Aspect	Inclusion Criteria	Exclusion Criteria
	integrated with another instructional model or learning medium	
Direct evidence outcomes	Studies that directly measured conceptual understanding and/or Science Process Skills (SPS) as an outcome of TaRL implementation	Studies that did not assess conceptual understanding or SPS were not classified as direct evidence
Supporting evidence outcomes	Empirical TaRL studies reporting general science/IPAS learning outcomes, scientific literacy, engagement, learning interest, diagnostic assessment, or related instructional outcomes that contributed to understanding the implementation or mechanisms of TaRL	Studies reporting outcomes unrelated to science education or not useful for interpreting TaRL implementation
Role in synthesis	Direct-evidence studies were used as the primary basis for synthesizing findings on conceptual understanding and SPS. Supporting-evidence studies were used only to contextualize implementation, diagnostic assessment, instructional mechanisms, and broader learning patterns	Supporting studies were not used to make direct claims regarding the effect of TaRL on conceptual understanding or SPS
Full-text availability	Full-text articles or sufficiently complete methodological and outcome information were available for verification and appraisal	Articles with insufficient bibliographic, methodological, or outcome information
Methodological quality	Studies meeting the predetermined minimum quality-appraisal criteria were retained in the final synthesis	Studies failing to meet the minimum methodological-quality threshold were excluded from the final synthesis

To maintain consistency between the review objectives and the included evidence, eligible studies were classified into two analytical categories. Direct-evidence studies were defined as empirical TaRL studies that explicitly measured conceptual understanding and/or Science Process Skills (SPS). These studies formed the primary evidence base for answering the main review questions. In contrast, supporting-evidence studies included empirical TaRL studies reporting general science/IPAS learning outcomes, scientific literacy, engagement, learning interest, diagnostic assessment, or other related outcomes. These studies were retained because they provided information about the instructional context, implementation mechanisms, and broader learning patterns associated with TaRL, but they were not treated as direct evidence of changes in conceptual understanding or SPS.

Accordingly, the synthesis did not consider all included studies as equivalent evidence. Findings concerning conceptual understanding and SPS were derived primarily from direct-evidence studies, whereas supporting studies were analyzed separately to explain how diagnostic assessment, ability grouping, differentiated scaffolding, instructional models, and related classroom processes may contribute to TaRL implementation in science education.

Study Selection and PRISMA Flow

The study-selection process followed the PRISMA 2020 framework and consisted of identification, duplicate checking, title and abstract screening, full-text eligibility assessment, and final inclusion. A total of 50 potentially relevant publications were compiled through targeted searches of Google Scholar, GARUDA, SINTA, relevant journal publisher websites, and backward citation searching. All retrieved records were entered into a single screening database containing the article title, author(s), publication year, journal name, DOI or URL, and source of retrieval.

Duplicate checking was conducted before screening. Records were compared based on article title, author names, publication year, DOI, and journal information. When the same publication was retrieved from more than one source, only one record was retained. In the final verified dataset, no duplicate records remained after this checking procedure; therefore, 50 unique records proceeded to title and abstract screening.

During title and abstract screening, each record was assessed against the predefined eligibility criteria. Twenty-five records were excluded because they did not report an empirical implementation of Teaching at the Right Level (TaRL) in elementary or junior-secondary science education. The excluded records mainly consisted of methodological references, review articles, studies of differentiated or adaptive learning that did not explicitly implement TaRL, general Science Process Skills studies without a TaRL intervention, diagnostic assessment studies outside the TaRL context, and methodological publications such as PRISMA 2020 and PRISMA-S.

The remaining 25 articles were retrieved in full text and assessed systematically for eligibility. Full-text assessment focused on the following criteria: (1) the study involved elementary/primary or junior-secondary students; (2) the educational context was science, IPAS, or science education; (3) TaRL was explicitly implemented as the main instructional

approach or integrated with another instructional model or medium; (4) the study employed an empirical research design; (5) the study reported outcomes relevant to conceptual understanding, Science Process Skills, general science/IPAS learning outcomes, diagnostic assessment, engagement, or related supporting outcomes; and (6) sufficient methodological and outcome information was available for appraisal.

Each full-text article was checked against the same eligibility checklist. Articles meeting all mandatory criteria were retained. Studies that directly measured conceptual understanding and/or Science Process Skills were classified as direct-evidence studies, whereas studies reporting general science/IPAS learning outcomes, diagnostic assessment, engagement, interest, or related outcomes were classified as supporting-evidence studies.

After full-text assessment, all 25 articles in the candidate set met the predefined eligibility criteria and were therefore retained for qualitative synthesis. No study was excluded at the full-text stage. This decision was based on explicit application of the eligibility checklist to each article rather than on title or abstract information alone.

Table 3. PRISMA 2020 Study Selection

PRISMA Stage	Number of Records
Records identified through targeted searches and citation tracking	50
Duplicate records removed after title, author, year, DOI, and journal comparison	0
Records screened by title and abstract	50
Records excluded after title and abstract screening	25
Full-text reports assessed for eligibility	25
Full-text reports excluded after eligibility assessment	0
Studies included in the qualitative synthesis	25

Table 4. Reasons for Exclusion at the Title and Abstract Screening Stage

Reason for Exclusion	Description
Non-empirical publication	Review article, methodological paper, editorial, or conceptual paper
No explicit TaRL implementation	Differentiated or adaptive instruction was discussed, but TaRL was not implemented
Outside science-education context	Study was conducted in mathematics, language, or another unrelated subject
Science Process Skills study without TaRL	SPS was examined, but no TaRL intervention was used
Diagnostic assessment outside TaRL	Diagnostic assessment was investigated independently of a TaRL intervention
Methodological reference	PRISMA, PRISMA-S, or other methodological guidance
Ineligible education level	Study involved populations outside elementary or junior-secondary education
Insufficient relevance	The article did not provide evidence relevant to the review questions

The 25 records excluded during screening were therefore used, where appropriate, as background or methodological references in the Introduction, Method, or Discussion, but they were not included as empirical studies in the qualitative synthesis.

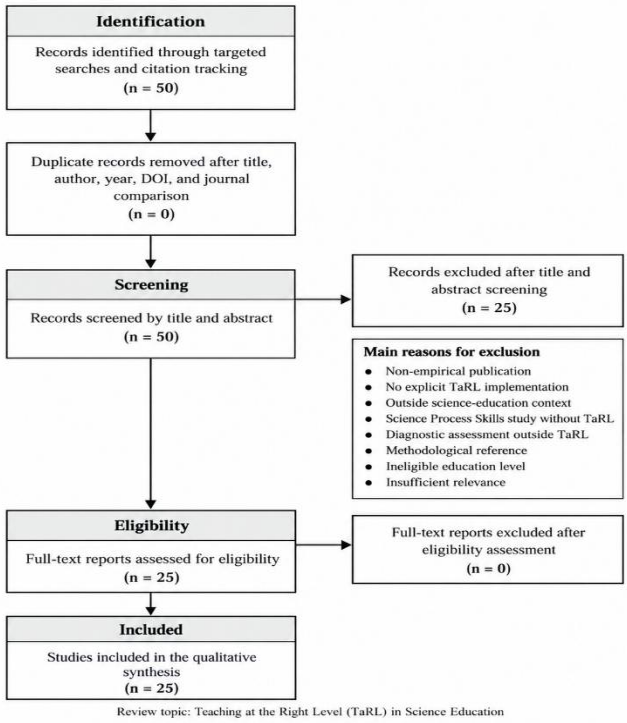


Figure 1. PrISMA 2020 Flow Diagram

Result and Discussion

The verified evidence set contained 25 empirical studies directly relevant to TaRL in science or IPAS education at elementary and junior-secondary levels (see Table 5). The studies were concentrated in Indonesia and published primarily between 2024 and 2026.

Quantitative Distribution of the Included Studies

The 25 included studies were relatively evenly distributed across the three most recent publication years, with eight studies published in 2024 (32%), nine in 2025 (36%), and eight in 2026 (32%). This pattern indicates that empirical applications of Teaching at the Right Level (TaRL) in science education have increased particularly during the most recent years of the review period.

In terms of educational level, 14 studies (56%) were conducted at the elementary or primary-school level,

while nine studies (36%) involved junior-secondary students. The educational level was not sufficiently specified in two studies (8%). Thus, the available evidence is more concentrated in elementary/IPAS contexts than in junior-secondary science education.

Table 5. Characteristics and Evidence Profile of the 25 Included Empirical Studies

Study	Participants / Research Design	Outcome, Evidence Type, and Main Finding
Nurfadilah et al. (2024)	Grade IX, junior secondary; Classroom Action Research, 3 cycles	Conceptual understanding – Direct evidence. Conceptual understanding increased across cycles; exact cycle scores were not available in the extracted dataset.
Ramlawati et al. (2025)	55 Grade VIII students; non-equivalent control-group design	SPS and critical thinking – Direct evidence. The intervention group showed more favorable SPS and critical-thinking outcomes. This represents the principal direct evidence for SPS in the review.
Ramlawati et al. (2026)	34 Grade VIII students; one-group pretest-posttest	Conceptual understanding – Direct evidence. Mean score increased from 6.31 to 16.12.
Murniningsih et al. (2026)	Science-education context; qualitative descriptive study	Diagnostic assessment – Supporting evidence. Diagnostic assessment was used to inform differentiation; no standardized effect estimate was reported.
Panjaitan & Puspita (2025)	50 Grade VIII students; quasi-experimental design	Science learning outcomes – Supporting evidence. Posttest mean was 91.73 for the intervention group and 68.53 for the control group.
Fatma & Listiawan (2025)	32 Grade VIII students; pre-experimental design	Science learning outcomes – Supporting evidence. Mean score increased from 65.53 to 83.43, with a reported N-gain of 0.5064.
Rifqi et al. (2025)	Grade VIII; pre-experimental design	Science learning outcomes – Supporting evidence. Learning outcomes improved following Discovery Learning–TaRL implementation; no exact effect estimate was available in the extracted dataset.
Millah et al. (2024)	Grade VII; Classroom Action Research, 2 cycles	Interest and science learning outcomes – Supporting evidence. Improvement was reported across intervention cycles.
Ahmad et al. (2024)	28 Grade VIII students; Classroom Action Research, 2 cycles	Science learning outcomes – Supporting evidence. Learning outcomes increased across cycles.
Lukman et al. (2026)	48 Grade VIII students; Classroom Action Research, 3 cycles	Science learning outcomes – Supporting evidence. Progressive improvement was reported across three intervention cycles.
Yulianto et al. (2026)	60 Grade V students; quasi-experimental design	Conceptual understanding and engagement – Direct evidence. Posttest mean was 82.7 for the intervention group and 71.4 for the comparison group.
Sastrawati et al. (2025)	29 Grade V students; Classroom Action Research	IPAS learning outcomes – Supporting evidence. 86.2% mastery was achieved in Cycle II.
Patajangi et al. (2025)	20 Grade IV students; Classroom Action Research, 2 cycles	IPAS learning outcomes – Supporting evidence. Learning outcomes increased across cycles.
Wahyuni et al. (2024)	19 Grade V students; Classroom Action Research	IPAS learning outcomes – Supporting evidence. Improvement was reported following PBL–TaRL implementation.
Setiadi et al. (2025)	Grade V; mixed-methods study	Learning outcomes – Supporting evidence. Mean score increased from 63.67 to 78.25.
Azlina et al. (2025)	Grade V; pre-experimental design	IPAS learning outcomes – Supporting evidence. More favorable post-intervention outcomes were reported; exact statistics were not available in the extracted dataset.
Sinaga et al. (2026)	41 Grade IV students; quasi-experimental design	IPAS learning outcomes – Supporting evidence. A statistical difference was reported ($p = .017$).
Mirwani et al. (2026)	45 Grade V students; quasi-experimental design	IPAS learning outcomes – Supporting evidence. The intervention group showed more favorable learning outcomes; no exact effect estimate was available in the extracted dataset.
Sartiga et al. (2026)	Grade V; Classroom Action Research	IPAS learning outcomes – Supporting evidence. Learning outcomes increased across intervention cycles.
Novena (2024)	Grade V; qualitative descriptive study	Learning process/outcomes – Supporting evidence. Findings indicated better alignment between instruction and students' learning readiness; no effect estimate was reported.

Study	Participants / Research Design	Outcome, Evidence Type, and Main Finding
Afandi et al. (2024)	School students; empirical study	Interest and learning outcomes – Supporting evidence. Favorable changes in learning interest and outcomes were reported; the extracted information did not provide a standardized effect estimate.
Anggraini et al. (2024)	Grade IV; quantitative study	IPAS learning outcomes – Supporting evidence. Favorable learning outcomes were reported; detailed effect estimates were not available in the extracted dataset.
Meinawati et al. (2024)	Grade IV; classroom empirical study	Learning outcomes – Supporting evidence. Improvement was reported following TaRL implementation; no standardized effect estimate was available in the extracted dataset.
Ervina et al. (2025)	30 Grade V students; Classroom Action Research	IPAS learning outcomes – Supporting evidence. Learning outcomes increased across intervention cycles.
Nihayah & Tirtoni (2026)	Elementary level; empirical study	Goal attainment – Supporting evidence. Diagnostic grouping within TaRL was associated with attainment of learning objectives; detailed quantitative evidence was not available in the extracted dataset.

The methodological profile was dominated by classroom-based and non-randomized designs. Nine studies (36%) employed Classroom Action Research, five studies (20%) used quasi-experimental or non-equivalent control-group designs, and four studies (16%) employed pre-experimental or one-group pretest-posttest designs. Two studies (8%) used qualitative descriptive designs, one study (4%) employed mixed methods, and four studies (16%) were classified as other empirical, quantitative, or classroom-based studies for which the specific methodological design was not sufficiently detailed in the extracted information. This distribution is important when interpreting the findings because only a minority of the included studies used controlled comparative designs.

Regarding the focal outcomes of this review, three studies (12%) directly examined conceptual understanding, whereas only one study (4%) directly assessed Science Process Skills (SPS). Thus, only four of the 25 studies (16%) provided direct evidence related to the two primary outcomes. The remaining 21 studies (84%) were classified as supporting evidence because they primarily examined general science/IPAS learning outcomes, diagnostic assessment, engagement, interest, goal attainment, or other implementation-related indicators. This distribution demonstrates that the evidence base directly addressing conceptual understanding remains limited and that direct evidence concerning SPS is even more restricted.

Among studies for which quantitative evidence was available in the extracted dataset, Ramlawati et al. (2026) reported an increase in conceptual-understanding scores from 6.31 to 16.12. Panjaitan and Puspita (2025) reported posttest means of 91.73 for the intervention group and 68.53 for the comparison group, while Fatma and Listiawan (2025) reported an increase from 65.53 to 83.43, with an N-gain of 0.5064. Yulianto et al. (2026) reported posttest means of 82.7 and 71.4 for the intervention and comparison groups, respectively. Setiadi et al. (2025) reported an increase from 63.67 to 78.25, Sastrawati et al. (2025) reported 86.2% mastery in Cycle II, and Sinaga et al. (2026) reported a statistical difference of $p = .017$.

For studies in which effect sizes, exact test statistics, confidence intervals, or complete pretest-posttest values were not available in the extracted information, findings were described conservatively as favorable changes or improvements rather than as statistically significant effects. Consequently, numerical findings across studies should not be directly compared without considering differences in research design, sample size, measurement instruments, scoring scales, and intervention procedures.

Table 6. Quantitative Distribution of the Included Studies

Characteristic	Category	n	%
Publication year	2024	8	32
	2025	9	36
	2026	8	32
Education level	Elementary/primary	14	56
	Junior secondary	9	36
	Not sufficiently specified	2	8
Research design	Classroom Action Research	9	36
	Quasi-experimental/non-equivalent control	5	20
	Pre-experimental/one-group	4	16
	Qualitative	2	8
	Mixed methods	1	4
	Other/insufficiently specified	4	16
Evidence type	Direct: conceptual understanding	3	12
	Direct: SPS	1	4
	Supporting evidence	21	84

Table 7. Thematic Classification of the Included Studies

Analytical Dimension	Theme / Category	Representative Studies	Number of Studies	Synthesis
Learning outcomes	Conceptual understanding	Nurfadilah et al. (2024); Ramlawati et al. (2026); Yulianto et al. (2026)	3	Direct evidence was limited to three studies. The available findings indicated favorable changes in conceptual understanding, but the small number of studies and variation in research design limit the strength of the conclusion.
Learning outcomes	Science Process Skills (SPS)	Ramlawati et al. (2025)	1	Direct evidence was available from only one study integrating Differentiated Science Inquiry with TaRL. The finding suggests potential benefits for SPS, but it is insufficient to establish a consistent effect of TaRL on SPS.
Learning outcomes	General science/IPAS learning outcomes	Panjaitan & Puspita (2025); Fatma & Listiawan (2025); Rifqi et al. (2025); Millah et al. (2024); Ahmad et al. (2024); Lukman et al. (2026); Sastrawati et al. (2025); Patajangi et al. (2025); Wahyuni et al. (2024); Setiadi et al. (2025); Azlina et al. (2025); Sinaga et al. (2026); Mirwani et al. (2026); Sartiga et al. (2026); Anggraini et al. (2024); Meinawati et al. (2024); Ervina et al. (2025)	17	Most studies reported favorable learning patterns after TaRL implementation. These studies were treated as supporting evidence rather than direct evidence for conceptual understanding or SPS.
Learning outcomes	Engagement, interest, and goal attainment	Millah et al. (2024); Yulianto et al. (2026); Afandi et al. (2024); Nihayah & Tirtoni (2026)	4	These outcomes provided supporting information regarding learner participation, motivation, and attainment, but were not considered primary evidence for the two focal science competencies.
Intervention characteristics	Diagnostic assessment and ability grouping	Murniningsih et al. (2026); Nihayah & Tirtoni (2026)	2	Diagnostic assessment was used to identify students' current competence and inform grouping or differentiated support. It functioned as an implementation mechanism rather than a primary learning outcome.
Intervention characteristics	TaRL integrated with inquiry-oriented instruction	Ramlawati et al. (2025)	1	TaRL was combined with Differentiated Science Inquiry to adjust inquiry support according to student readiness. Evidence remains limited to a single study.
Intervention characteristics	TaRL integrated with Problem-Based Learning	Panjaitan & Puspita (2025); Wahyuni et al. (2024); Ervina et al. (2025); Sinaga et al. (2026)	4	TaRL was used to adapt the level of support within PBL activities. The studies generally reported favorable learning outcomes, although most did not directly measure conceptual understanding or SPS.
Intervention characteristics	TaRL integrated with Discovery Learning	Rifqi et al. (2025)	1	TaRL was used to differentiate support within discovery-

Analytical Dimension	Theme / Category	Representative Studies	Number of Studies	Synthesis
				oriented learning. Evidence was limited to one empirical study.
Intervention characteristics	TaRL integrated with differentiated instruction or other media	Ahmad et al. (2024); Mirwani et al. (2026); Sartiga et al. (2026); Afandi et al. (2024)	4	TaRL was combined with differentiated instruction, Wordwall, mind mapping, or other classroom supports to accommodate differences in learner readiness.
Educational context	Junior-secondary science	Nurfadilah et al. (2024); Ramlawati et al. (2025); Ramlawati et al. (2026); Panjaitan & Puspita (2025); Fatma & Listiawan (2025); Rifqi et al. (2025); Millah et al. (2024); Ahmad et al. (2024); Lukman et al. (2026)	9	Junior-secondary studies covered science topics such as conceptual understanding, SPS, ecology, compounds, simple machines, Earth structure, light, and waves.
Educational context	Elementary/primary IPAS and science	Yulianto et al. (2026); Sastrawati et al. (2025); Patajangi et al. (2025); Wahyuni et al. (2024); Setiadi et al. (2025); Azlina et al. (2025); Sinaga et al. (2026); Mirwani et al. (2026); Sartiga et al. (2026); Novena (2024); Anggraini et al. (2024); Meinawati et al. (2024); Ervina et al. (2025); Nihayah & Tirtoni (2026)	14	Elementary studies primarily examined IPAS or integrated science learning and mostly focused on general achievement, implementation, or learner participation rather than direct measurement of SPS.
Educational context	Context insufficiently specified	Murniningsih et al. (2026); Afandi et al. (2024)	2	The available extracted information did not provide sufficiently specific educational-level details; these studies were retained only where their relevance to TaRL implementation could be verified.

The thematic synthesis used a hybrid deductive-inductive approach. The main analytical dimensions were established deductively from the review questions and the conceptual framework of the study. These dimensions included: (1) learning outcomes, particularly conceptual understanding and Science Process Skills; (2) intervention characteristics, including diagnostic assessment, ability grouping, instructional integration, and differentiated scaffolding; and (3) educational context, particularly elementary/primary and junior-secondary science education.

Within these predefined dimensions, subthemes were developed inductively through repeated reading and comparison of the included studies. For example, the integration of TaRL with Differentiated Science Inquiry, Problem-Based Learning, Discovery Learning, differentiated instruction, Wordwall, and mind mapping emerged from the empirical characteristics of the included studies rather than being predetermined before article screening.

This analytical structure was used to avoid treating different types of evidence as equivalent. Studies directly assessing conceptual understanding or SPS were synthesized as direct evidence, whereas studies reporting general learning outcomes, engagement, interest, diagnostic assessment, or related implementation variables were used as supporting evidence.

Science Process Skills

Evidence concerning Science Process Skills was substantially more limited than evidence concerning general science learning outcomes. Among the included studies, only Ramlawati et al. (2025) directly assessed SPS within a TaRL-based instructional intervention. In that study, TaRL was integrated with Differentiated Science Inquiry, allowing the level of inquiry guidance and instructional support to be adjusted according to students' readiness.

Although the study reported favorable SPS outcomes, this evidence should be interpreted cautiously. Because the finding is derived from a single study in which TaRL was combined with an inquiry-oriented instructional model, the observed outcome cannot be attributed to TaRL alone. The contribution may reflect the interaction between diagnostic grouping, differentiated support, and inquiry-based learning.

Accordingly, the present review does not conclude that TaRL has established effectiveness in improving

Science Process Skills. Rather, the available evidence suggests a potentially promising relationship that requires confirmation through additional controlled studies in different science topics, educational levels, and learning contexts using validated SPS instruments.

Table 8. Reported and Synthesized Stages of TaRL Implementation in Science Education

Stage	Procedures Reported in the Reviewed Studies	Diagnostic / Instructional Evidence	Influence on Grouping and Instruction	Evidence Status
Diagnostic assessment	Several studies reported an initial assessment before instruction to identify students' current level of understanding or learning readiness.	The reviewed studies referred to diagnostic tests, prior-learning assessments, classroom assessment results, or initial achievement measures. However, not all studies provided detailed descriptions of the diagnostic instrument, number of items, validity, or scoring procedure.	Assessment results were used to identify differences in student competence and to determine the level of instructional support required.	Empirically reported, although instrument details were uneven across studies
Ability mapping / grouping	Students were grouped according to demonstrated competence rather than solely according to grade level.	Grouping was generally based on diagnostic or initial-assessment results. Some studies described categories of learner readiness or achievement levels, but reporting of cut-off scores varied.	Students with similar learning needs were placed in groups to facilitate targeted instruction and differentiated support.	Empirically reported in several studies
Differentiated instructional planning	Teachers adjusted learning activities according to students' readiness or competence level.	Variations included differences in task complexity, worksheets, learning materials, instructional support, or level of inquiry guidance.	Lower-readiness groups generally received greater scaffolding, while higher-readiness groups were given more complex or more independent learning tasks.	Empirically reported pattern synthesized across studies
Targeted instruction	Science or IPAS learning was delivered according to the competence profile of each group.	TaRL was implemented either as the main adaptive framework or integrated with other approaches such as Differentiated Science Inquiry, Problem-Based Learning, Discovery Learning, differentiated instruction, or learning media.	Instructional activities were adjusted to match the students' current learning level rather than applying identical tasks to the entire class.	Empirically reported
Formative assessment	Some studies reported ongoing assessment during or after learning activities to monitor progress.	Evidence included classroom tests, observation, learning-task completion, cycle evaluation, or post-instruction assessment. Detailed formative-assessment procedures were not consistently reported across all studies.	Assessment results were used to determine whether students had achieved the targeted level and whether additional support was required.	Reported in some studies; not consistently documented
Flexible regrouping	Regrouping based on student progress is consistent with the TaRL principle, but explicit descriptions of repeated regrouping were limited in the reviewed studies.	Evidence of repeated movement between groups was not consistently reported.	Where progress monitoring was used, grouping could theoretically be adjusted according to subsequent performance.	Primarily a synthesized/recommended implementation principle rather than a consistently reported empirical procedure

Diagnostic Assessment in TaRL Implementation

Diagnostic assessment emerged as an important component of TaRL implementation, but its operationalization varied across the included studies. In several studies, initial assessment was used to determine students' prior knowledge, readiness, or current level of achievement before instructional grouping. Murniningsih et al. (2026), for example, emphasized the

role of assessment in identifying students' learning levels and informing differentiated instruction, while Nihayah and Tirtoni (2026) linked diagnostic assessment with the determination of appropriate learning targets.

The types of diagnostic evidence reported across the studies included initial achievement tests, prior-learning assessments, classroom assessment results, and teacher-based identification of student readiness. However, the reporting quality was inconsistent. Not all studies provided detailed information about the diagnostic instrument, such as the number and format of items, scoring procedures, validity evidence, reliability, or the cut-off criteria used to assign students to particular competence groups. This limitation reduces the extent to which the grouping procedures can be replicated across contexts.

Where diagnostic assessment was explicitly described, its main instructional function was to support ability mapping. Students were grouped according to demonstrated competence or learning readiness, and instructional activities were then differentiated accordingly. The differentiation could involve variations in task difficulty, worksheet complexity, learning resources, amount of scaffolding, level of inquiry guidance, or teacher support. Thus, diagnostic assessment functioned less as a grading mechanism and more as a decision-making tool for determining the level and form of instruction provided to each learner group.

Importantly, the reviewed studies did not report all stages of TaRL implementation with equal consistency. Diagnostic assessment, grouping, and differentiated instruction were more frequently described, whereas repeated formative assessment and flexible regrouping were less explicitly documented. Consequently, the implementation sequence presented in this review should not be interpreted as a standardized procedure followed identically by all studies. Rather, it represents a synthesis of the recurring practices identified across the literature.

Table 9. Studies Directly Addressing Conceptual Understanding

Study	Level	Intervention	Result
Nurfadilah et al. (2024)	Grade IX	TaRL	Conceptual understanding improved
Ramlawati et al. (2026)	Grade VIII	DSI-TaRL	Mean 6.31 → 16.12
Yulianto et al. (2026)	Grade V	TaRL	82.7 experimental vs 71.4 control

Evidence on SPS is substantially more limited. Ramlawati et al. (2025) reported a significant positive effect of DSI-TaRL on SPS and critical thinking.

Table 10. Differentiated Inquiry Support

Readiness Level	Inquiry Form	Typical Support
Foundational	Demonstrated/ structured inquiry	Teacher modelling, procedures, prompts
Developing	Guided inquiry	Selected decisions made by students
Advanced	Independent/ self-directed inquiry	Greater autonomy in planning and analysis

Table 11. Integration of TaRL with Instructional Models and Media

Integration	Representative Studies	Comparative Synthesis and Evidence Limitation
Differentiated Science Inquiry (DSI) + TaRL	Ramlawati et al. (2025; 2026)	This integration differentiates inquiry guidance, scaffolding, and task complexity according to student readiness. Ramlawati et al. (2025) reported favorable SPS and critical-thinking outcomes, while Ramlawati et al. (2026) reported an increase in conceptual-understanding scores from 6.31 to 16.12. However, the available studies do not establish that DSI + TaRL is superior to TaRL alone, and the contribution of DSI cannot be separated clearly from that of TaRL.
Problem-Based Learning (PBL) + TaRL	Panjaitan & Puspita (2025); Wahyuni et al. (2024); Sinaga et al. (2026); Ervina et al. (2025)	PBL + TaRL adapts problem complexity and instructional support to students' competence levels. Panjaitan and Puspita (2025) reported posttest means of 91.73 vs 68.53, while Sinaga et al. (2026) reported $p = .017$. The evidence suggests favorable learning patterns, but most studies compared the integrated approach with conventional instruction rather than standalone TaRL.
Discovery Learning + TaRL	Rifqi et al. (2025)	This integration supports concept discovery while adapting the amount of guidance to students' readiness. Improved science outcomes were reported, but the evidence is limited to a single study and does not show whether the combination performs better than TaRL alone.
Differentiated Learning + TaRL	Ahmad et al. (2024); Novena (2024); Afandi et al. (2024)	The integration adapts content, process, learning resources, and scaffolding to learner readiness. The studies reported favorable patterns in outcomes or engagement, but the conceptual overlap between differentiated learning and

Integration	Representative Studies	Comparative Synthesis and Evidence Limitation
		TaRL makes the unique contribution of each component difficult to determine.
Wordwall + TaRL	Mirwani et al. (2026)	Wordwall provides interactive practice within competence-based differentiation. Favorable IPAS learning outcomes were reported, but the evidence is limited to one study and does not isolate the added contribution of Wordwall beyond TaRL itself.
Mind Mapping + TaRL	Sartiga et al. (2026)	Mind mapping supports visual organization of scientific concepts within differentiated instruction. Improvement was reported across classroom-action cycles, but the evidence comes from a single CAR study and does not allow a direct comparison with standalone TaRL.

Across the included studies, the overall direction of findings was generally favorable; however, the strength of the evidence varied substantially across research designs. Studies employing quasi-experimental or non-equivalent control-group designs provide relatively stronger comparative evidence because they include a reference condition against which changes in the intervention group can be interpreted. In contrast, classroom action research and one-group pretest-posttest studies are more useful for documenting classroom improvement and feasibility than for establishing causal effects. Improvements observed across intervention cycles or between pretest and posttest scores may also be influenced by maturation, repeated testing, teacher effects, increased familiarity with the material, or other uncontrolled classroom factors. Consequently, positive findings from these designs should be interpreted as indications of potential benefit rather than definitive evidence that TaRL alone caused the observed improvement.

The magnitude and meaning of reported changes also differed across studies. Some studies provided numerical evidence, such as changes in mean scores, mastery percentages, N-gain values, or statistical significance, whereas others reported improvement only in descriptive terms. For example, Ramlawati et al. (2026) reported an increase in conceptual-understanding scores from 6.31 to 16.12, while Panjaitan and Puspita (2025) reported posttest means of 91.73 and 68.53 for the intervention and comparison groups. Fatma and Listiawan (2025) reported an increase from 65.53 to 83.43 with an N-gain of 0.5064. These findings provide more interpretable evidence than studies that only describe outcomes as "improved" or "positive." Nevertheless, direct comparison across studies remains difficult because the instruments, score ranges, intervention conditions, and outcome definitions were not standardized.

Differences in intervention duration may also contribute to variation in findings. Classroom action research studies typically implemented TaRL across several cycles, allowing teachers to revise instruction after each cycle, whereas pre-experimental and quasi-experimental studies may have applied the intervention over shorter or more fixed periods. Longer interventions may provide greater opportunities for students to adapt to competence-based grouping and differentiated support, but they may also introduce additional influences such as teacher reflection, repeated practice, and cumulative exposure. Because intervention duration was not reported consistently across all studies, the review could not determine whether longer implementation produced stronger outcomes.

Sample size is another source of heterogeneity. Several studies involved relatively small classroom samples, whereas others included larger groups or comparison classes. Small samples are common in classroom-based educational research and are useful for examining implementation in authentic settings, but they limit statistical power and generalizability. In addition, a finding obtained from a single class may reflect contextual characteristics of that class, teacher, school, or topic rather than the broader effect of TaRL. Future research should therefore include larger multi-class or multi-school samples where feasible.

Variation in measurement instruments also affects the comparability of the evidence. Conceptual understanding was assessed using different instruments and scoring systems, while Science Process Skills were directly assessed in only a very small number of studies. Some studies focused on general achievement tests, others on mastery percentages, engagement, interest, or diagnostic assessment. This heterogeneity means that favorable findings across studies do not necessarily represent the same underlying learning construct. The use of validated and clearly reported instruments is therefore essential for building a stronger cumulative evidence base.

The conceptual-understanding pattern can be interpreted through the principles of adaptive instruction and scaffolding. Science learning often requires students to integrate new information with prerequisite concepts. When instructional tasks substantially exceed students' current knowledge, cognitive demands may become too high and meaningful integration may be disrupted. Conversely, tasks that are substantially below students' competence may provide insufficient challenge. TaRL attempts to

reduce this mismatch by using diagnostic information to align task difficulty, teacher support, pacing, and expected independence with students' demonstrated readiness. From a scaffolding perspective, support can be increased for learners who require more guidance and gradually reduced as competence develops.

This interpretation also helps explain why TaRL is frequently combined with other instructional models. Problem-Based Learning, Discovery Learning, Differentiated Science Inquiry, Wordwall, mind mapping, and differentiated instruction provide different forms of learning activity, whereas TaRL primarily determines the level at which those activities are delivered. In this sense, TaRL may be understood as an adaptive instructional layer rather than a substitute for inquiry, PBL, Discovery Learning, or educational technology. However, the reviewed literature does not yet demonstrate that integrated approaches are superior to standalone TaRL because most studies did not include a direct TaRL-only comparison condition.

The evidence concerning Science Process Skills requires particular caution. Only one study directly examined SPS within a TaRL-based intervention, and that study combined TaRL with Differentiated Science Inquiry. Although favorable SPS outcomes were reported, the contribution of TaRL cannot be isolated from the contribution of inquiry-oriented learning. The current evidence therefore supports only a tentative proposition that competence-based differentiation may facilitate process-oriented science learning. Additional controlled studies are required to examine specific SPS indicators, including observation, classification, measurement, hypothesis formulation, identification of variables, experimental design, data interpretation, and evidence-based conclusion drawing.

Diagnostic assessment represents another critical component of the TaRL mechanism. The instructional logic of TaRL depends on the accuracy with which students' current competence is identified. If diagnostic instruments do not adequately represent prerequisite knowledge or if grouping thresholds are poorly defined, students may be assigned to inappropriate instructional levels. Such misclassification may weaken the intended alignment between learner readiness and instructional support. The reviewed studies indicate that diagnostic assessment was frequently used to guide grouping and differentiation, but detailed information about instrument validity, reliability, scoring procedures, and cut-off criteria was not consistently reported. This methodological weakness should be addressed in future TaRL studies.

Potential publication bias should also be considered. Most included studies reported favorable outcomes, while few reported null or unfavorable findings. This pattern may reflect genuine benefits of TaRL, but it may also result from a greater likelihood that studies reporting improvement are submitted, accepted, or made visible in indexed sources. Because the present review did not conduct a statistical assessment of publication bias and the available studies were heterogeneous in design and outcome measurement, the predominance of positive findings should not be interpreted as conclusive evidence of effectiveness.

The geographical concentration of the evidence further limits generalization. Most of the included studies were conducted in Indonesia, particularly within elementary IPAS and junior-secondary science contexts. The findings are therefore closely connected to the Indonesian curriculum, classroom structure, teacher practices, and implementation of differentiated learning. Although the underlying principle of matching instruction to demonstrated competence may be relevant across educational systems, the present evidence should not be generalized directly to international contexts with different curricula, assessment systems, class sizes, teacher preparation, or institutional resources. Cross-national studies are needed to determine whether similar patterns emerge in other educational settings.

From a practical perspective, the findings suggest that TaRL implementation in science education requires more than simply dividing students into groups. Effective implementation depends on a sequence of diagnostic assessment, meaningful competence mapping, differentiated task design, targeted instruction, progress monitoring, and appropriate adjustment of support. Teachers also need access to diagnostic instruments that are sufficiently sensitive to identify prerequisite knowledge and misconceptions. In science classrooms, differentiation may involve variations in worksheet complexity, inquiry guidance, problem difficulty, laboratory support, learning media, pacing, and the degree of student independence.

At the same time, competence-based grouping should remain flexible rather than becoming a permanent label for students. Progress should be monitored regularly, and instructional groups should be adjusted when evidence indicates that students have developed new competencies. This principle is particularly important for avoiding unintended tracking effects and ensuring that TaRL remains a responsive instructional strategy rather than a fixed classification system.

Overall, the reviewed literature indicates a promising but methodologically uneven evidence base. Direct evidence for conceptual understanding is limited, and direct evidence for Science Process Skills is particularly scarce. General science/IPAS learning outcomes provide useful supporting evidence, but they

should not be treated as equivalent to direct measures of the two focal competencies. Future research should prioritize controlled comparative designs, larger and more diverse samples, clearly reported intervention duration, validated outcome instruments, explicit diagnostic-assessment procedures, and direct comparisons between standalone TaRL and integrated TaRL approaches.

Conclusion

This systematic literature review synthesized 25 empirical studies on Teaching at the Right Level (TaRL) in science education. Overall, the evidence indicates that TaRL is being used as an adaptive instructional framework in which diagnostic assessment informs competence-based grouping, differentiated tasks, scaffolding, and targeted instruction. However, the strength of evidence varies considerably across the included studies.

The strongest available evidence comes from a limited number of quasi-experimental or non-equivalent control-group studies. In contrast, a substantial proportion of the literature consists of classroom action research, pre-experimental, and one-group pretest-posttest studies. These designs provide useful information about classroom implementation and improvement but offer more limited support for causal conclusions.

Evidence directly related to conceptual understanding remains limited. Only a small number of studies explicitly measured this outcome, although the available findings generally showed favorable changes. Direct evidence concerning Science Process Skills (SPS) is even more limited and is primarily represented by one study integrating TaRL with Differentiated Science Inquiry. Consequently, the current literature does not yet provide sufficient evidence to conclude that TaRL consistently improves SPS across different science topics and educational settings. Most of the remaining studies focused on general science/IPAS learning outcomes, engagement, diagnostic assessment, or other supporting indicators.

Based on these findings, future research should prioritize controlled comparative designs, including quasi-experimental studies with well-matched control groups and, where feasible, randomized designs. Studies should employ validated and clearly reported instruments for conceptual understanding and specific SPS indicators, such as observing, classifying, measuring, formulating hypotheses, identifying variables, designing investigations, interpreting data, and drawing evidence-based conclusions. Effect sizes, confidence intervals, and complete statistical results should also be reported to facilitate comparison across studies.

Future studies should additionally examine longer-term outcomes by including delayed posttests, retention measures, or longitudinal follow-up. Larger multi-class and multi-school samples are also needed to strengthen generalizability. Further research should compare standalone TaRL with TaRL integrated with inquiry, Problem-Based Learning, Discovery Learning, differentiated instruction, and educational technology to determine whether these combinations provide additional benefits beyond competence-based differentiation alone.

Accordingly, TaRL is best interpreted as a promising adaptive instructional framework rather than as an intervention with conclusively established effects across all science-learning outcomes. Its potential contribution appears strongest in supporting differentiated instruction based on diagnostic information, while stronger evidence is still needed for conceptual understanding and, especially, Science Process Skills.

Acknowledgments
The authors would like to express their appreciation to the researchers whose published studies provided the empirical and theoretical basis for this systematic literature review. The authors also acknowledge the academic support provided by the Graduate Program of Science Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Indonesia, during the preparation of this manuscript.

Author Contributions
Conceptualization; Q. D. K.; methodology.; R.; validation; formal analysis; E.L.A.; investigation.; Q.D.K; resources; R.; data curation: E.L.A.; writing—original draft preparation. Q.D.K.; writing—review and editing: Q.D.K.; visualization: E.L.A All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare no conflict of interest.

References

Afandi, R. A., Ningtyas, N. S., Susiyawati, E., & Pratiwi, P. (2024). The effectiveness of differentiated learning using the TaRL (Teaching at the Right Level) approach for improving learning interest and learning outcome. *Jurnal Pijar Mipa, 19*(4), 657–662. https://doi.org/10.29303/jpm.v19i4.6860

Ahmad, D. A., Palloan, P., & Mardawiah. (2024). Penerapan pendekatan Teaching at the Right Level

(TaRL) melalui pembelajaran berdiferensiasi terhadap hasil belajar IPA peserta didik kelas VIII SMP Negeri 7 Makassar. *Jurnal Pemikiran dan Pengembangan Pembelajaran, 6*(2).

Amalia, H., Ramlawati, R., & Hasanuddin, H. (2023). Pengaruh E-LKPD berbasis inkuiri terbimbing terhadap keterampilan proses sains peserta didik kelas VIII SMPN 8 Makassar. *Celebes Science Education, 2*(3). https://doi.org/10.35580/cse.v2i3.53691

Anggraini, E. Y. N., Subekti, E. E., & Lestari, A. B. (2024). Pengaruh pendekatan TaRL terhadap hasil belajar pada mata pelajaran IPAS kelas IV SDN Sawah Besar 01 Semarang. *Education Journal: Journal Educational Research and Development, 8*(2). https://doi.org/10.31537/ej.v8i2.1940

Azlina, N., Aisy, P. R., & Mas'ula, S. (2025). Efektivitas pendekatan Teaching at the Right Level (TaRL) dalam meningkatkan hasil belajar siswa kelas V sekolah dasar. *Journal of Innovation and Teacher Professionalism, 3*(3), 728–738. https://doi.org/10.17977/um084v3i32025p728-738

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., & Walton, M. (2016). *Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India* (NBER Working Paper No. 22746). National Bureau of Economic Research. https://doi.org/10.3386/w22746

Banerji, R., & Chavan, M. (2020). A twenty-year partnership of practice and research: The Nobel laureates and Pratham in India. *World Development, 127*, 104788. https://doi.org/10.1016/j.worlddev.2019.104788

Ervina, Saparini, & Mindarti, P. (2025). Penerapan pendekatan Teaching at the Right Level terintegrasi model Problem Based Learning dalam meningkatkan hasil belajar materi IPAS kelas V.A SDN 55 Palembang. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 10*(2). https://doi.org/10.23969/jp.v10i02.27750

Fatma, A. F., & Listiawan, T. (2025). Efektivitas pendekatan Teaching at the Right Level terhadap hasil belajar IPA siswa kelas VIII. *Journal of Innovation and Teacher Professionalism, 3*(2), 287–294. https://doi.org/10.17977/um084v3i22025p287-294

Lukman, K. N., Ramlawati, R., & Rosdiana, R. (2026). Implementing Teaching at the Right Level to improve science learning outcomes: A classroom action research study. *MISOOL: Jurnal Pendidikan Dasar, 8*(1), 159–167. https://doi.org/10.47945/misool.v8i1.3290

Meinawati, R., Wicaksono, V. D., Firnanda, Z. I., Nina, P. L., & Sunipah, A. (2024). Pengaruh penerapan pendekatan Teaching at the Right Level (TaRL) terhadap hasil belajar peserta didik pada materi fotosintesis kelas 4 di SDN Kalisari II Surabaya. *Jurnal Kajian Ilmu Pendidikan, 5*(2), 272–280.

Millah, A. I., Widodo, W., Haryono, E. D. P., & Grahani, N. P. S. (2024). Peningkatan minat dan hasil belajar IPA materi ekologi dan keanekaragaman hayati Indonesia dengan pendekatan TaRL (Teaching at the Right Level). *LEARNING: Jurnal Inovasi Penelitian Pendidikan dan Pembelajaran, 4*(3). https://doi.org/10.51878/learning.v4i3.3123

Mirwani, Sayidiman, & Nurhaedah. (2026). Pengaruh penerapan pendekatan Teaching at the Right Level (TaRL) dan media Wordwall terhadap hasil belajar IPAS siswa SD Inpres Palompong Kabupaten Gowa. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 11*(1), 460–475. https://doi.org/10.23969/jp.v11i1.43292

Murniningsih, M., Widyawati, A., Trisniawati, T., Awani, N. C., & Maharani, C. A. (2026). How is assessment with the Teaching at the Right Level (TaRL) approach implemented in science learning? *Tarbiyah: Jurnal Ilmiah Kependidikan, 14*(2), 413–434. https://doi.org/10.18592/tarbiyah.v14i2.17922

Nihayah, D., & Tirtoni, F. (2026). Penerapan Teaching at the Right Level (TaRL) berbasis asesmen diagnostik untuk meningkatkan ketercapaian tujuan pembelajaran (KKTP) di sekolah dasar. *Jurnal Ilmiah Pendidikan Dasar, 6*(1). https://doi.org/10.37081/jipdas.v6i1.4576

Novena, B. P. (2024). Implementasi pendekatan Teaching at the Right Level (TaRL) pada mata pelajaran IPA materi siklus air kelas V. *Pedagogi: Jurnal Penelitian Pendidikan, 11*(1), 10–18. https://doi.org/10.25134/pedagogi.v11i1.9683

Numa, N. A. S., & Martini, M. (2022). Application of guided inquiry student worksheet on additive materials to improve the science process skills of junior high school students. *Jurnal Pijar MIPA, 17*(4), 424–429. https://doi.org/10.29303/jpm.v17i4.3535

Nurfadilah, A., Saenab, S., & Timung, H. (2024). Penerapan pendekatan Teaching at the Right Level dalam peningkatan pemahaman konsep IPA peserta didik kelas IX.8 SMPN 3 Makassar. *Jurnal Pemikiran dan Pengembangan Pembelajaran, 6*(2), 1374–1381.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting

systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Panjaitan, H. S. A., & Puspita, R. D. (2025). Efficacy of Teaching at the Right Level approach combined with the Problem-Based Learning model on science learning outcomes at middle school. *Jurnal Pendidikan Sains, 13*(3), 154–162. https://doi.org/10.17977/jps.v13i32025p154

Patajangi, A. P. N., Mahmud, R., & Muslim, S. (2025). Penerapan pendekatan Teaching at the Right Level (TaRL) untuk meningkatkan hasil belajar IPAS siswa kelas IV SD. *Lempu PGSD, 2*(2), 126–132. https://doi.org/10.70713/lempu.v2i2.4303

Patonah, S., Sajidan, Cari, & Rahardjo, S. B. (2020). The effectiveness of STEM-based science learning in improving students' science process skills. *Jurnal Penelitian Pendidikan Sains, 9*(2), 1812–1818. https://doi.org/10.26740/jpps.v9n2.p1812-1818

Ramlawati, R., Sari, N. I., Kusumawati, R., Yesin, M., Ilmi, N., & Arsyad, A. A. (2025). The effect of Differentiated Science Inquiry learning model based on Teaching at the Right Level on students' critical thinking and science process skills. *Jurnal Pendidikan IPA Indonesia, 14*(1), 1–16. https://doi.org/10.15294/jpii.v14i1.19479

Ramlawati, R., Sari, N. I., Yunus, S. R., Faisah, N., & Alqadri, Z. (2026). Students' conceptual understanding of the DSI model within the TaRL approach. *Jurnal Inovasi Pendidikan IPA, 12*(1), 580–590. https://doi.org/10.21831/jipi.v12i1.82342

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA Statement for reporting literature searches in systematic reviews. *Systematic Reviews, 10*, 39. https://doi.org/10.1186/s13643-020-01542-z

Rifqi, A. M., Sutrisno, S., & Lestari, L. P. (2025). Pengaruh model pembelajaran Discovery Learning berpendekatan TaRL (Teaching at the Right Level) terhadap hasil belajar siswa pada materi senyawa kelas VIII SMP. *Jurnal Integrasi dan Harmoni Inovatif Ilmu-Ilmu Sosial, 5*(5). https://doi.org/10.17977/um063.v5.i5.2025.4

Sartiga, A. M. E., Manafe, Y. Y., & Ratu, K. T. R. A. (2026). Penggunaan media mind mapping menggunakan pendekatan Teaching at the Right Level (TaRL) dengan materi perpindahan kalor untuk meningkatkan hasil belajar IPAS di kelas 5 SD Negeri Kelapa Lima. *Jurnal Ilmiah Pendidikan Dasar, 6*(1), 66–71. https://doi.org/10.37081/jipdas.v6i1.3812

Sastrawati, E., Nurjanah, Kurnia, D., Evidawati, Saputra, M. D., Sari, P. A., & Wiyanpuri, R. S. (2025). Peningkatan hasil belajar IPAS peserta didik kelas V melalui pendekatan TARL. *Jurnal Ilmiah Profesi Pendidikan, 10*(1), 557–562. https://doi.org/10.29303/jipp.v10i1.2973

Setiadi, N., Azizah, M., & Antoko, L. (2025). Pengaruh pendekatan TaRL (Teaching at the Right Level) terhadap hasil belajar pada mata pelajaran IPAS siswa kelas V. *Jurnal Pendidikan Tambusai, 8*(3), 51441–51447.

Sinaga, R. S., Manurung, I. F. U., Nurmayani, Rozi, F., & Ratno, S. (2026). Pengaruh model Problem Based Learning berbasis pendekatan TaRL (Teaching at the Right Level) terhadap hasil belajar IPAS siswa kelas IV sekolah dasar. *Primary Education Journals, 6*(1), 414–420. https://doi.org/10.36636/primed.v6i1.7593

Soeharto, S. (2021). Development of a diagnostic assessment test to evaluate science misconceptions in terms of school grades: A Rasch measurement approach. *Journal of Turkish Science Education, 18*(3), 351–370. https://doi.org/10.36681/tused.2021.78

Soeharto, S., & Csapó, B. (2022). Exploring Indonesian student misconceptions in science concepts. *Heliyon, 8*(9), e10720. https://doi.org/10.1016/j.heliyon.2022.e10720

Suryawanto, A. M., Subali, B., & Haryani, S. (2022). Implementation of guided inquiry model with MURDER strategy to improve elementary school students' science process skills. *Journal of Primary Education, 11*(3), 313–326. https://doi.org/10.15294/jpe.v11i3.72385

Wahyuni, L. S., Sukmanasa, E., & Dhiani, A. N. (2024). Penerapan pendekatan Teaching at the Right Level model Problem Based Learning untuk meningkatkan hasil belajar peserta didik pada mata pelajaran IPAS di kelas V SDN Kedung Badak 2. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 9*(2), 4572–4586. https://doi.org/10.23969/jp.v9i2.14243

Yulianto, A. A., Rianti, W., & Ayu, C. (2026). Pengaruh pembelajaran Teaching at the Right Level terhadap keaktifan dan pemahaman konsep siswa MI Aulia Cendekia Pekanbaru. *Journal of Education Research, 7*(1), 134–141. https://doi.org/10.37985/jer.v7i1.3423