International Research Journal of Multidisciplinary Scope (IRJMS),2026;7(1):1161-1174

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2026.v07i01.08619

STEM Scientific Inquiry Rubric for Early Childhood

Fanny de Jesús Rúa Martínez[1]*, Martha Andrea Merchán Merchán[1], Elena Calderón-Canales[2], Beatriz García-Rivera[2]

[1]Universidad Antonio Nariño. Facultad de Educación. Grupo de investigación Conciencia. Calle 22 sur # 12D -81. 111821, Bogotá, Colombia, [2]Instituto de Ciencias Aplicadas y Tecnología, Universidad Nacional Autónoma de México, Circuito Exterior S/N, Ciudad Universitaria, 04510, México City. *Corresponding Author's Email: frua38@uan.edu.co

Abstract

This study presents the design and validation of a rubric to assess scientific inquiry competence in childhood education through a STEM approach. Although notable progress has been made in STEM-oriented pedagogical practices, there remains a lack of validated instruments that rigorously and contextually evaluate inquiry processes at early educational levels. The rubric integrates five core dimensions—formulating questions, planning, collecting data, analyzing, and communicating—organized into nine criteria and 36 descriptors written in language accessible to young learners. Content validity was established through expert judgment using Aiken's V coefficient and 95% confidence intervals, yielding favorable results across indicators of clarity, coherence, sufficiency, and relevance. Subsequently, a pilot study with 113 childhood education students enabled a confirmatory factor analysis (CFA), which demonstrated strong construct validity supported by excellent model fit indices (CFI = 1.000, RMSEA = 0.000). Reliability analysis using Cronbach's alpha produced coefficients above 0.97 across all dimensions, confirming a high level of internal consistency. Overall, the validated rubric constitutes a reliable and pedagogically meaningful tool for identifying inquiry levels and supporting instructional decision-making within STEM-based childhood education. Its methodological rigor, accessible structure, and strong psychometric properties position it as a valuable resource for educators and researchers seeking to strengthen inquiry-driven teaching and learning processes.

Keywords: Assessment, Childhood Education, Rubric Design, Scientific Inquiry, STEM Approach, Validation.

Introduction

Children can develop coherent ideas about natural processes and scientific reasoning from their early years; high-quality science interventions can significantly enhance their learning, sometimes going beyond what is typically emphasized in early childhood curricula (1). Science thus becomes a key element in human development, enabling children to understand their environment, interact with it, and contribute to its transformation (2). From constructivist and sociocultural perspectives, it is recognized that children build knowledge through active exploration, material manipulation, dialogue, and gradual guidance provided by the educational environment, factors that support progression toward higher levels of understanding (3).

Recently, it has been noted that learning progressions in inquiry competence describe how children advance from initial forms of exploration to more sophisticated levels of question formulation, data collection, and explanation of phenomena (4). Likewise, studies in the STEM field focused on inquiry show that experiences such as engineering design projects, educational robotics activities, and problem-solving in authentic contexts promote clear and continuous trajectories for the development of scientific reasoning from an early age (5).

Promoting scientific reasoning in childhood involves strengthening inquiry competence, which allows children to explore, formulate questions, seek explanations, and communicate their findings meaningfully (6). This dynamic process fosters curiosity, critical thinking, and deep understanding of natural phenomena, shaping autonomous and reflective learners (7).

Scientific inquiry develops progressively through interconnected capacities such as problematization, planning, recording data, analysis, and evaluation—core components of scientific thinking from early education (8, 9). These abilities are reinforced through real and contextualized experiences, where children establish relationships, interpret data, and apply knowledge to solve problems (10).

Within this perspective, the STEM (Science,

(Received 11th October 2025; Accepted 04th January 2026; Published 28th January 2026)

Technology, Engineering, and Mathematics) approach acts as a catalyst for inquiry, integrating these disciplines to foster interdisciplinary learning (11). The STEM approach not only strengthens reasoning, hypothesis formulation, and problem-solving but also cultivates transferable competencies essential for the 21st century, including those related to ICT (Information and Communication Technology), programming, and computational thinking (12, 13).

Science, technology, engineering, and mathematics provide opportunities that, when integrated, enhance inquiry, critical thinking, and problem-solving from early ages (14). Its experimental and collaborative nature allows children to explore phenomena, design solutions, use diverse representations, and communicate findings in meaningful contexts, thereby fostering creativity, self-confidence, and cooperative skills (15).

Considering this approach, STEM experiences—by integrating hands-on exploration, gradual support, language, design, and joint problem-solving—create especially favorable conditions for children to develop increasingly complex forms of inquiry and thinking (16). Consequently, assessment tools are needed to accurately document how children articulate skills and knowledge throughout these interdisciplinary experiences (8).

Assessing how children express inquiry skills in learning situations is crucial to document their progress and guide pedagogical practices (17, 18). Therefore, tools that explicitly assess inquiry actions aligned with STEM principles are essential (19). Despite the fundamental role of inquiry skills in early childhood, existing assessment instruments often focus on environmental or teaching factors and provide mainly descriptive information, offering limited guidance on children's actual inquiry processes (20). Scientific inquiry involves children formulating questions, exploring phenomena, planning and conducting investigations, interpreting results, and communicating findings (21). The STEM approach integrates the principles of science, technology, engineering, and mathematics to promote problem-solving, critical thinking, and evidence-based reasoning (22). The linkage between inquiry and STEM establishes a coherent framework to observe, understand, and support the development of scientific competencies in early childhood (23).

Although existing frameworks like the Science Teaching and Environment Rating Scale (STERS), the Preschool Rating Instrument for Science and Mathematics (PRISM), the Classroom Assessment Scoring System (CLASS), the Electronic Quality of Inquiry Protocol (EQUIP), and the Classroom Observation Protocol (COP), contribute to understanding scientific experiences in early education; most focus on environmental or teaching factors rather than directly assessing children's inquiry processes (24–28).

This gap underscores the need for assessment instruments sensitive to children's contexts and capable of evidencing emerging scientific competencies (29). A rubric grounded in the core dimensions of inquiry and aligned with STEM principles can provide teachers with a pedagogical and evaluative tool to recognize and foster these competencies (30).

Therefore, this study aims to design and validate a rubric (RCCI-STEM: Rubric for Assessing Scientific Inquiry Competence and its Relation to STEM Areas) to assess scientific inquiry competence and its connection to STEM areas in early childhood education. Beyond its evaluative role, the rubric seeks to support teachers by offering clear reference points for inquiry development, guiding instructional decisions, and enriching educational practices aligned with 21st-century learning needs.

Methodology

This study employed a descriptive, mixed-methods design aimed at developing and validating a rubric to assess scientific inquiry competence in early childhood education within a STEM framework. The descriptive component facilitated a systematic characterization of the instrument's components, structure, and criteria without variable manipulation (31). The mixed-methods approach combined quantitative data, such as content validity coefficients, with qualitative input from experts, thereby enhancing the analysis and ensuring a thorough and comprehensive evaluation (32).

Initial Rubric Design

The purpose of this study was to design and validate a rubric oriented towards the assessment of scientific inquiry competence in childhood education, based on pedagogical experiences grounded in the STEM approach. The instrument design was structured based on a rigorous process that combined revision of specialized literature, identification of precedent referents and the collaboration of experts in the area.

As a starting point, an analysis of instruments used to assess the development of scientific inquiry in early childhood education was conducted. During this process, the Protocol of Observation of STEM Classrooms was identified. The results of the protocol are particularly relevant due to its specific approach to childhood education, its emphasis on the development of particular scientific skills, and its alignment with the STEM framework as an integrative model for fostering scientific thinking from an early age (28). To widen the understanding of the instrument and to know more about its contextual application, direct contact with the authors was established, who authorized access to complementary material that enriched this stage.

Based on this referent, a first version of the rubric was elaborated, oriented towards the analysis of the dimensions that are part of the scientific inquiry process. This initial design was then revised and enriched by the contributions of the Grupo de Cognición y Didáctica de las Ciencias (GCDC, Cognition and Didactics in Sciences Group), from the Instituto de Ciencias Aplicadas y Tecnología (ICAT, Institute of Applied Sciences and Technology) of the Universidad Nacional Autónoma de México (UNAM, National Autonomous University of Mexico).

These contributions led to a conceptual adaptation in which the phases of scientific inquiry are linked to the STEM phases, with an emphasis on the engineering design process. Table 1 illustrates the connections between each stage and the core STEM actions, such as problem identification, solution ideation, conceptual construction, validation, and reflection on the outcomes (33, 34).

Table 1: Conceptual Alignment between Phases of Inquiry and STEM Phases with Emphasis on Engineering Design

Phases of Inquiry	**STEM Phases – Emphasis on Engineering Design**	**Conceptual Actions of Engineering Design**
Preparation / Introduction to Context	Engage – Purpose	Problem identification and goal setting
Idea Inquiry / Prediction Development	Explore / Plan and Design	Solution ideation and conceptual planning
Development	Explain	Conceptual construction, experimentation, and validation
Discussion Phase	Elaborate	Solution refinement and comparison of alternatives
Conclusions	Evaluate	Outcome assessment and reflection on the process

Rubric Design Process

The rubric to assess scientific inquiry competence and its relations with STEM areas (RCCI-STEM) is proposed as a support tool for teacher practice within the classroom. It provides clear and organized references that allow the identification of the development level of scientific inquiry in the students. This tool was not conceived as an isolated instrument, but as a part of a wider evaluative proposal, integrated by an activity of applied structured application, a checklist for the observer, and a complementary rubric that gathers the categories of the inquiry levels. This last phase was subject to expert validation to assure sufficiency, clarity, coherence, and pertinence of the established criteria to assess the level of inquiry shown by students during the pedagogical experience.

RCCI-STEM was structured in five dimensions related to the scientific inquiry process: Formulate Questions and Problematize Situations, Plan and Select Material, Collect and Record Data, Analyze and Provide Conclusions, and Assess and Communicate (35). Each of these dimensions is composed of two analysis criteria, except for the dimension Collect and Record Data, which contains just one, for a total of nine evaluative criteria.

Based on those criteria, performance descriptors were distributed in four levels: initial (1 point), basic (2 points), intermediate (3 points), and advanced (4 points). These descriptors allow for the observation and assessment of the progressive development of the scientific inquiry competence in childhood education students. They were written in a clear and accessible language, according to their cognitive and communicative development level, and allowed to assess core inquiry competencies, as shown in Table 2.

Table 2: Excerpt from the RCCI-STEM Rubric Corresponding to the Dimension "Formulates questions and problematizes situations

Dimension: Formulate Questions and Problematize Situations (FPH)	STEM areas: Science, Engineering	Criteria: Formulate Questions
Inquiry levels		
Advanced (FPH1) Formulate questions about the device and shows an advanced understanding of how different parts contribute to motion. Questions vary and go beyond the observed and can show explanations about the internal functioning and possible improvements. E.g., 'How would it affect changing the type of engine in the device motion?' or 'What would happen if we used a bigger wheel?' 'How would the device move if we add more weight?'		
Intermediate (FPH2) Formulate questions related to the device motion but pinpoints the observable and general aspects. Questions show curiosity and look forward to understanding the relations among components, even though they do not always go deep into technical details. E.g., 'How does the device move when we push it?' or 'What happens if we change the wheel's shape?' 'Does the device move as fast on all surfaces?'		
Basic (FPH3) Formulate simple questions about the device motion, centered on what is evident or in what is observable on the surface. Questions do not aim for a deep explanation or explore possible solutions or variations. E.g., 'Does the device move?' 'Why does the device roll?' Can the device move fast or slow?'		
Initial (FPH4) Does not formulate questions or their questions are very limited, centered only on basic things, and without trying to explore motion on the device. Questions are very general or are not related to the topic of motion. E.g., 'Does the device have wheels?' or 'Does the device work?' 'Can the device be pushed?'		

Dimension: Formulate Questions and Problematize Situations (FPR)	STEM areas: Science, Engineering	Criteria: Proposes Answers (Hypothesis)
Inquiry Levels		
Advanced (FPR1) Propose detailed and evidence-based answers, considers different variables and offers creative and logical solutions. The student justifies their hypothesis on how device modifications would affect its functioning. E.g., 'If we use a more powerful engine, the device could move faster. We could also try different materials for the wheels.'		
Intermediate (FPR2) Propose clear and reasonable answers, but without deep details or an extensive justification. The student limits themself to what seems more obvious or what was learned previously. E.g., 'If we use big wheels, the device could move faster but could not work correctly on rough surfaces.'		
Basic (FPR3) Propose simple answers that do not explore beyond what is evident. Solutions are very straightforward and do not include much testing or reasoning. E.g., 'If we put wheels on it, the device could move. Perhaps we could try to move it by pushing it.'		
Initial (FPR4) Does not propose any hypothesis or solutions. Does not show interest in independently resolving the problem and needs a lot of help to get involved. E.g., when asked how to make the device move, the student responds, "I do not know," or simply does not respond.		

Expert Participants

For the expert judgment validation process, ten evaluators were selected through purposive sampling (36). This strategy was chosen to ensure the participation of professionals with specialized knowledge and proven experience in assessing scientific competencies in early childhood, guaranteeing that their judgments would be well-founded and relevant for instrument validation. Selection criteria included academic background, research experience, and professional expertise in the educational field. All participants held a doctoral degree in Education or advanced knowledge in a STEM area, as well as experience in designing, implementing, or evaluating educational instruments in school contexts, and were affiliated with education faculties of Colombian universities or were teachers in public and private schools in Bogotá.

Priority was given to experts in early childhood teacher training, classroom teachers with experience in teaching scientific competencies during the first years of schooling, and researchers specialized in early education with knowledge in STEM areas, as the literature emphasizes the need to select judges with specific expertise and relevant experience in the study subject (37). This selection ensures content validity, as recommendations and evaluations come from competent professionals, and also guarantees process accuracy by applying standardized criteria to record and analyze their assessments. Participation was voluntary and anonymous, and each expert was provided with detailed information on the study objectives, the instrument to be validated, the evaluation criteria, and the procedure for providing their judgments.

Procedure

The present study was developed within a methodological approach of descriptive analysis with a mixed design. This approach was selected due to the need for characterization and analysis of an educational assessment instrument without variable manipulation, to prioritize the detailed description of its components, structure, and criteria (32). From a descriptive analysis perspective, the objective was to observe and systematically represent the characteristics of the instrument and expert judgment. It has been highlighted that this type of study allows for the description of phenomena as they occur, without direct intervention by the researcher (38).

The mixed design allows for the integration of quantitative and qualitative supplementary data. Quantitative strategies, such as the use of Aiken's V coefficient, were applied to assess content validity, while qualitative contributions were incorporated through written observations of the judges. This approach enriched the analysis and allowed for more substantial adjustments from an integral perspective, consistent with previous studies in the field of educational validation (32, 39).

Content Validity: The content validity of the rubric was estimated through the evaluation of 10 expert judges, who assessed each item based on four criteria: coherence, clarity, sufficiency, and relevance. Ratings were collected using a 4-point Likert Scale (where 4 indicated the highest level of agreement and 1 the lowest). Aiken's V coefficient was then calculated for each item (38, 40). Additionally, 95% confidence intervals were estimated using the Score interval formula, which incorporates the ($C\text{-}1^2$) term in the denominator as recommended by researchers in the past (41).

Pilot Study Participants: Once the content of the instrument had been validated, a pilot study was conducted with a non-probabilistic convenience sample of 113 early childhood education students, drawn from three different groups within the same educational institution, each initially composed of 40 students. Seven students did not participate in the pilot study due to absences during the process or the lack of informed consent. Convenience sampling was used, which allows participants to be selected based on their availability and accessibility (42). The instrument was applied to all students present in the groups, including three students diagnosed with autism and five with attention deficit hyperactivity disorder (ADHD).

Inclusion criteria were: enrollment in early childhood education, regular attendance, and possession of informed consent. The sample included 59 girls (52.2%) and 54 boys (47.8%). Regarding age distribution, most participants were 7 years old (51.3%), followed by 8 years (24.8%), 9 years (14.2%), 6 years (5.3%), and 10 years (4.4%). It is noteworthy that early childhood education generally covers children up to 8 years of age (43). The 21 students (18.6%) older than this exhibited academic delays related to interruptions in their educational trajectory, particularly during the COVID-19 pandemic.

The sample was considered adequate to assess the clarity, feasibility, and initial reliability of the instrument, as it aimed to identify potential difficulties in administration and comprehension, ensuring that the applicability and validity of the instrument were verified before its full implementation (44).

Construct Validity: To assess the construct validity of the instrument, a Confirmatory Factor Analysis (CFA) was conducted using AMOS software. Confirmatory Factor Analysis is a robust statistical technique that contrasts a pre-established theoretical model with empirical data, through the assessment to the latent and observable variables (45, 46). Its application results are particularly relevant within the educational area, as it examines whether the items cluster coherently around the proposed theoretical dimensions, thereby strengthening the construct validity of the instrument (47). In this study, CFA verified the structural robustness of the instrument designed to assess scientific inquiry competence in early childhood within the STEM approach, confirming that the theoretically defined dimensions align with the empirical organization of the data.

Reliability: The reliability of the instrument was estimated using Cronbach's alpha to analyze the internal consistency of its items. This procedure validated the homogeneity of the answers and the coherence among the elements of each one of the dimensions of the rubric. To interpret the obtained values, we followed criteria from specialized literature, which deems coefficients ≥ 0.70 as acceptable and those > 0.80 as optimal (46, 47).

Results and Discussion

Content Validity

Table 3 presents the results of the validation process of the instrument, assessed by expert judges. 144 registers were analyzed, corresponding to 36 items, each one assessed in four criteria: sufficiency, clarity, coherence, and pertinence. All items meet the established thresholds X ≥ 3, V de Aiken ≥ 0.8, and ICI ≥ 0.5, so none were discarded. However, minor wording adjustments were made to some items, as the expert's feedback indicated improvements in the linguistic formulation rather than content flaws. Aiken's V coefficients ranged from 0.83 to 0.93. Additionally, the 95% confidence interval bounds fell between 0.67 and 0.98, reflecting strong inter-judge consensus. The content validity of the instrument is thus supported by these results, consistent with previous recommendations (40).

Table 3: Results of Content Validity by Expert Judges

Items	Validity	X	SD	V	IC-LI	IC-LS	Agreement	Permanence
FPH1	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	No observations
	Clarity	3.7	0.674	0.900	0.838	0.962	Yes	
	Coherence	3.6	0.699	0.867	0.796	0.937	Yes	
	Pertinence	3.5	0.849	0.833	0.756	0.910	Yes	
FPH2	Sufficiency	3.8	0.421	0.933	0.882	0.985	Yes	No observations
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.8	0.699	0.867	0.796	0.937	Yes	
	Pertinence	3.7	0.674	0.900	0.838	0.962	Yes	
FPH3	Sufficiency	3.8	0.421	0.933	0.882	0.985	Yes	Minor wording adjustment
	Clarity	3.5	0.707	0.833	0.756	0.910	Yes	
	Coherence	3.8	0.421	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.674	0.900	0.838	0.962	Yes	
FPH4	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	Prevent negative phrasing: change 'no' to 'has difficulties'
	Clarity	3.5	0.849	0.833	0.756	0.910	Yes	
	Coherence	3.6	0.699	0.867	0.796	0.937	Yes	
	Pertinence	3.5	0.849	0.833	0.756	0.910	Yes	
FPR1	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	No observations
	Clarity	3.7	0.674	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.8	0.632	0.933	0.882	0.985	Yes	
FPR2	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
FPR3	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
FPR4	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	Prevent negative phrasing: change 'no' to 'has difficulties'
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
PNT1	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	Minor wording adjustment
	Clarity	3.4	0.843	0.800	0.717	0.883	Yes	
	Coherence	3.6	0.699	0.867	0.796	0.937	Yes	
	Pertinence	3.5	0.850	0.833	0.756	0.910	Yes	
PNT2	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	Minor wording adjustment
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
PNT3	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
PNT4	Sufficiency	3.7	0.483	0.900	0.838	0.962	Yes	Minor wording adjustment
	Clarity	3.5	0.707	0.833	0.756	0.910	Yes	
	Coherence	3.6	0.516	0.867	0.796	0.937	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	
PNM1	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	No observations
	Clarity	3.5	0.707	0.833	0.756	0.910	Yes	
	Coherence	3.6	0.516	0.867	0.796	0.937	Yes	
	Pertinence	3.5	0.850	0.833	0.756	0.910	Yes	
PNM2	Sufficiency	3.7	0.483	0.900	0.838	0.962	Yes	No observations
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	
PNM3	Sufficiency	3.7	0.483	0.900	0.838	0.962	Yes	Minor wording adjustment
	Clarity	3.5	0.707	0.833	0.756	0.910	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	
PNM4	Sufficiency	3.6	0.516	0.867	0.796	0.937	Yes	

	Clarity	3.4	0.699	0.800	0.717	0.883	Yes	Minor wording adjustment
	Coherence	3.6	0.516	0.867	0.796	0.937	Yes	
	Pertinence	3.5	0.850	0.833	0.756	0.910	Yes	
RD1	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
RD2	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
RD3	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	Omit the connector 'Whether or'
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
RD4	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	Minor wording adjustment
	Clarity	3.3	0.823	0.767	0.679	0.854	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.5	0.850	0.833	0.756	0.910	Yes	
ACD1	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ACD2	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ACD3	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ACD4	Sufficiency	3.7	0.483	0.900	0.838	0.962	Yes	No observations
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	
ACC1	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	No observations
	Clarity	3.5	0.850	0.833	0.756	0.910	Yes	
	Coherence	3.6	0.699	0.867	0.796	0.937	Yes	
	Pertinence	3.5	0.850	0.833	0.756	0.910	Yes	
ACC2	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ACC3	Sufficiency	3.8	0.422	0.933	0.882	0.985	Yes	Minor wording adjustment
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ACC4	Sufficiency	3.6	0.699	0.867	0.796	0.937	Yes	Minor wording adjustment
	Clarity	3.5	0.850	0.833	0.756	0.910	Yes	
	Coherence	3.6	0.699	0.867	0.796	0.937	Yes	
	Pertinence	3.4	0.843	0.800	0.717	0.883	Yes	
ECH1	Sufficiency	3.7	0.483	0.900	0.838	0.962	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ECH2	Sufficiency	3.6	0.516	0.867	0.796	0.937	Yes	Minor wording adjustment
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	
ECH3	Sufficiency	3.6	0.516	0.867	0.796	0.937	Yes	Minor wording adjustment
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	
ECH4	Sufficiency	3.7	0.483	0.900	0.838	0.962	Yes	No observations
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ECE1	Sufficiency	3.7	0.483	0.900	0.838	0.962	Yes	Substitute 'Reflect' with 'Demonstrate'
	Clarity	3.7	0.675	0.900	0.838	0.962	Yes	
	Coherence	3.8	0.422	0.933	0.882	0.985	Yes	
	Pertinence	3.7	0.675	0.900	0.838	0.962	Yes	
ECE2	Sufficiency	3.6	0.516	0.867	0.796	0.937	Yes	No observations
	Clarity	3.6	0.699	0.867	0.796	0.937	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	
ECE3	Sufficiency	3.6	0.516	0.867	0.796	0.937	Yes	Minor wording adjustment
	Clarity	3.5	0.707	0.833	0.756	0.910	Yes	
	Coherence	3.7	0.483	0.900	0.838	0.962	Yes	
	Pertinence	3.6	0.699	0.867	0.796	0.937	Yes	

ECE4	Sufficiency	3.5	0.707	0.833	0.756	0.910	Yes	Prevent negative phrasing: change 'no' to 'has difficulties'
	Clarity	3.5	0.707	0.833	0.756	0.910	Yes	
	Coherence	3.6	0.699	0.867	0.796	0.937	Yes	
	Pertinence	3.5	0.850	0.833	0.756	0.910	Yes	

Note: X: Mean, SD: Standard deviation, V: Aiken's V coefficient, LCI: Lower confidence interval (95%), UCI: Upper confidence interval (95%).

Based on the information given by the judges regarding sufficiency, clarity, coherence, and pertinence criteria on each item, minor wording adjustments were made. Particularly, observations related to the use of negative phrasing were pointed out, replacing them with affirmative formulations, per the guidelines for the elaboration of clear and precise statements in validation instruments (48). This decision aims to promote more direct and respectful expression of the capabilities of the students. Other comments were directed towards the substitution of verbs as 'reflect' with 'demonstrate', as well as removing unnecessary connectors. Table 4 presents a synthesis of the more representative observations from judges during the validation process.

Table 4: Validation's Representative Observations

Item	Original Wording	Judge's Comments	Restructuration
FPH4	Does not formulate questions or their questions are too limited, centered on the basics, and without exploring device motion. Questions are too general or aren't related to the motion topic. E.g., 'Does the device have wheels?' or 'Does the device work?' 'Can the device be pushed?	Jz9: I suggest changing 'no' to 'has difficulties?	Has difficulties formulating questions about the device's motion. Their questions are too general, focusing on basic aspects rather than exploring motion in depth. E.g., 'Does the device have wheels?' 'Does the device work?' 'Can the device be pushed?' 'Is it heavy or light?
ECE1	The questions they formulated for their classmates reflect attentive listening to others' explanations, as they are relevant for exploring ideas more deeply. Furthermore, they conduct reflective comparisons of work samples, identifying similarities and differences, which enhances their work process.	Jz8: I suggest changing the word 'reflect' to 'demonstrate' because the goal is for the student to articulate what they are internalizing	The questions they formulated for their classmates demonstrate attentive listening to others' explanations, as they are relevant for exploring ideas more deeply. Furthermore, they conduct reflective comparisons of work samples, identifying similarities and differences, which enhances their work process.
PNM4	They do not know which materials to choose by themselves and depend completely on the help of the teacher or classmates to select and understand the use of materials during the device construction.	Jz5: Suggestion: delete the first words and start the sentence with: 'Depends completely on the help…'	Depends completely on the help of the teacher or classmates to select and understand the use of materials during the device construction. They struggle to choose materials on their own.

Based on the quantitative and qualitative analysis made by the expert judges, specific observations were identified that gave rise to minor adjustments in the wording of some items. In general terms, the instrument assessment demonstrated that more than 80% of the judge's considered sufficiency, clarity, coherence, and pertinence criteria suitable, which indicates the items are understandable, formulated with an appropriate syntax, maintain logical congruence with the assessed dimensions, and respond to the instrument's purpose (38). These observations optimized the results, particularly in aspects like negative phrasing, selection of more precise verbs, and the suppression of unnecessary connectors (48). Likewise, judges highlighted the instrument's relevance to assess the inquiry process within STEM education in childhood (49). As a result of the content validation process and the theoretical reinforcement of the rubric, a final version was consolidated, comprising 36 items organized into five dimensions, each with its respective descriptors. Table 5 presents these dimensions, which were defined a priori and subsequently evaluated in terms of construct validity.

Table 5: Instrument Dimensions, Associated Criteria, and Number of Items by Performance Level

Dimension	Associated Criteria	General Descriptor (Quispe, 2023)	Performance Levels	Items
Formulate Questions and Problematize Situations	FPH: Formulate questions FPR: Propose answers (hypothesis)	Formulate questions based on observations and propose possible answers or hypotheses.	Advanced, Intermediate, Basic, Initial	8
Plan and Select Materials	PNT: Plan the work PNM: Select materials	Organize the work in advance and choose pertinent materials to complete it.	Advanced, Intermediate, Basic, Initial	8
Collect and Record Data	RD: Write or draw what has to be done.	Record through drawings and text what they do during the inquiry process.	Advanced, Intermediate, Basic, Initial	4

Analyze and Provide Conclusions	ACD: Check the data ACC: Formulate conclusions	Examine the information obtained and extract conclusions based on the data.	Advanced, Intermediate, Basic, Initial	8
Evaluate and Communicate	ECH: Explain what was done ECE: Listen to classmates	Reflect on what was done and share their experience, considering different perspectives.	Advanced, Intermediate, Basic, Initial	8

Construct Validity

The findings from the confirmatory factor analysis demonstrate that the theoretical model has an excellent fit to the data, supporting its construct validity. The chi-square test was not significant (χ^2 (10) = 6.814, *p* = 0.743), suggesting no relevant differences between the observed and estimated covariance matrices. The ratio of chi-square to degrees of freedom (χ^2/df = 0.681) was below the recommended threshold (< 2.0), further reinforcing the model's adequacy (46, 47).

The absolute fit indices also support this conclusion: the Root Mean Square Residual (RMR = 0.005) was minimal the Goodness-of-Fit Index (GFI = 0.985) and its adjusted version (AGFI = 0.947) exceeded the benchmark value of 0.90. Although the Parsimony Goodness-of-Fit Index (PGFI = 0.274) was low, this result is common in models with simple structures and does not compromise overall fit quality (46, 50).

Additionally, the RMSEA index showed a value of 0.000, indicating a perfect fit. The incremental fit indices were also outstanding: NFI = 0.995, RFI = 0.987, IFI = 1.002, TLI = 1.006, and CFI = 1.000, all well above the minimum threshold of 0.90, confirming the robustness of the theoretical model concerning empirical data (47, 50).

These indicators—summarized in Table 6—demonstrate a strong model fit, consistent with methodological standards recommended in specialized literature (47, 50).

Table 6: Model Fit Indices for Confirmatory Factor Analysis (CFA)

Index	Value	Recommended Criterion	Interpretation
χ^2 (gl = 10)	6.814	p > 0.05	Adequate Fit (p = 0.743)
χ^2/gl	0.681	< 2.00 o < 3.00	Excellent fit
RMR	0.005	< 0.05	Very good fit
GFI	0.985	≥ 0.90	Excellent fit
AGFI	0.947	≥ 0.90	Good fit
PGFI	0.274	> 0.50 preferred	Low due to parsimony
RMSEA	0.000	< 0.05 excellent; < 0.08 acceptable	Perfect fit
NFI	0.995	≥ 0.90	Perfect fit
RFI	0.987	≥ 0.90	Excellent fit
IFI	1.002	≥ 0.90	Excellent fit
TLI	1.006	≥ 0.90	Excellent fit
CFI	1.000	≥ 0.95	Excellent fit

The model represented in Figure 1 demonstrates the theoretical structure of the instrument, consisting of five latent dimensions: Formulate Questions and Problematize Situations, Plan and Select Materials, Analyze and Provide Conclusions, Evaluate and Communicate, and Collect and Record Data. The last one is cross-cutting and is not graphically represented because it is constituted by one observable criterion, which articulates implicitly with the other dimensions during the inquiry process. Such a saturation level suggests that items measure precisely the theoretical constructs previously defined (45, 46).

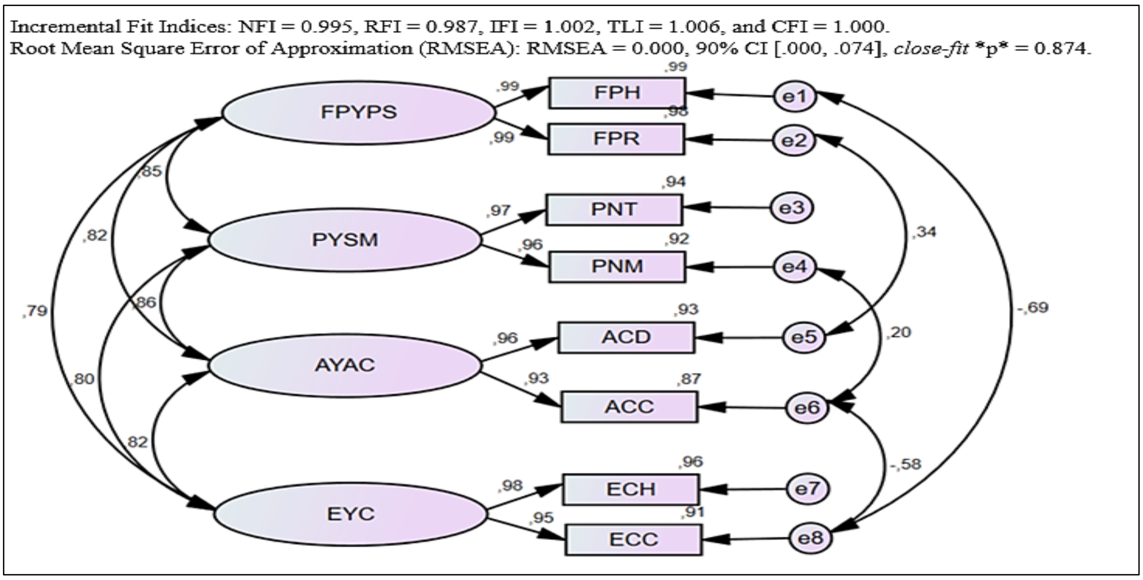

Figure 1: Confirmatory Factor Analysis (CFA) Structural Model with Fit Indices

Additionally, elevated correlations among latent dimensions were observed with values that range from 0.79 to 0.86. This high correlation was expected in a context such as childhood education, where the cognitive processes linked to scientific inquiry are deeply interdependent (51). For example, question formulation is strongly related to action planning, selecting adequate materials, interpreting results, and communicating findings. These interconnections reinforce the model's holistic character and demonstrate that, while dimensions were defined as distinct categories, they operate in an integrated manner within pedagogical practice (52).

The visual absence of the Collect and Record Data dimension from the graphical mode reflects purely technical representation criteria. It's single item, not being part of the cluster of indicators, cannot be directly included without distorting the covariance matrix. Nevertheless, its role is fundamental within the inquiry process because data recording occurs simultaneously with other cognitive actions. From this perspective, its representation was addressed complementarily without compromising the model's structural coherence or the obtained fit indices (47).

Together, the CFA graphic model confirms the structural solidity of the instrument and empirically validates the items grouping in coherent and conceptually strong dimensions. This evidence backs up the instrument as a valid and reliable tool to assess the development of inquiry thinking in educational contexts based on a childhood STEM educational approach (37).

Reliability Using Cronbach's Alpha

The instrument's reliability was estimated through Cronbach's alpha coefficient to determine the internal consistency of its items. For this analysis, 144 ratings were considered, corresponding to the assessment of the 36 rubric's items across four criteria: sufficiency, clarity, coherence, and pertinency, rated by 10 expert judges. The reliability analysis of the complete instrument — comprising 144 items distributed across five dimensions—yielded an overall Cronbach's alpha of 0.996, indicating excellent internal consistency (53,54). As shown in Table 7, the RCCI-STEM rubric demonstrated excellent internal consistency both overall and across its dimensions. This value demonstrates that the instrument's items are highly coherent with one another, ensuring measurement stability and precision.

Table 7: Cronbach's alpha per Dimension

Dimension	Cronbach's alpha	No. of Items
FPH - FPR	0.979	32
PNT - PNM	0.980	32
RD	0.974	16
ACD - ACC	0.986	32
ECH - ECE	0.989	32
Total	0.996	144

Additionally, Cronbach's alpha based on standardized elements was 0.997, which reaffirms the instrument's high reliability, even, if possible, differences in scale are considered among items. These results highly surpass the minimum threshold of 0.70, commonly accepted as an indicator of adequate internal consistency (54). This value falls within a range that reflects strong item homogeneity (53). Consequently, these findings empirically support the instrument's stability, coherence, and precision, solidifying its validity and reliability for assessing scientific inquiry processes in childhood education within a STEM approach (55).

As shown in Table 6, the dimension Collect and Record Data (RD), composed of 16 items, presented a value α = 0.974, considered excellent despite having fewer items compared to the other dimensions. This result confirms that, even with a smaller number of items, adequate internal cohesion is maintained, and reliability is not compromised when a reduced set of items is well-designed and conceptually aligned (56).

In terms of the dimension "Analyze and Provide Conclusions" (ACD-ACC) an alpha of 0.986 was obtained, while the dimension "Evaluates and Communicates" (ECH-ECE) reached the highest value with α = 0.989, which confirms the strongness of both dimensions as compounds of the instrument.

Finally, when analyzing the entire instrument (144 items) an overall coefficient of α = 0.996 was obtained, which confirms the exceptional reliability of the rubric as a whole, as it far exceeds the recommended threshold of 0.70 (46, 57). These results support the psychometric robustness of the RCCI-STEM rubric and its ability to assess coherently and accurately the associated dimensions to the scientific inquiry process in childhood education (58).

Conclusion

The results of the validation process of the instrument demonstrate a psychometric robust structure, supported by both the expert judgement and confirmatory factor analysis. The grouping of items into five theoretically defined dimensions was empirically confirmed through high factor loadings and excellent global model fit indices (CFI = 1.000, TLI = 1.006, RMSEA = 0.000), providing strong evidence of construct validity (45, 54).

The instrument's internal consistency is exceptional, as evident by an overall Cronbach's alpha of 0.996 with values exceeding 0.97 in each of its dimensions. These results amply surpass the minimum threshold (α ≥ 0.70) and fall within the range considered excellent, ensuring stability, precision, and measurement coherence (54, 56).

Furthermore, it is reaffirmed that when an instrument reaches these levels, the data obtained can be interpreted with high confidence and reliability.

From an educational perspective, the validated instrument is a relevant tool for fostering teacher reflection by rigorously assessing levels of scientific inquiry competence in childhood within the STEM approach. This assessment provides the teacher key information to identify strengths and difficulties in the development of scientific skills, thereby guiding more informed, relevant, pertinent and contextualized pedagogical decision-making. By understanding their students' inquiry levels, teachers can adjust instructional strategies, plan more challenging learning experiences, personalize support processes, and promote active knowledge construction from an early age (59, 60).

In summary, the instrument not only meets high psychometric standards in terms of validity and reliability but also positions itself as a valuable resource for pedagogical innovation, educational research, and the improvement of teaching practices. It has potential for application in national and international contexts where the development of scientific skills from early childhood is an educational priority (49).

Despite these strengths and contributions, it is important to acknowledge certain internal limitations of the study. The age range of the participants included children older than the typical upper limit for early childhood education, which may have affected certain observed performances and patterns (61). In addition, the presence of students with specific conditions, such as autism or Attention-Deficit/Hyperactivity Disorder (ADHD), required careful management during the administration of the instrument; having additional support staff would allow for more individualized attention to the needs of all participants (62). Furthermore, the sample size and the use of a non-probabilistic convenience sampling method limit the generalizability of the findings to other populations (42). Nevertheless, these limitations can be addressed in future research through larger and more diverse samples, as well as stronger support strategies during instrument administration, thereby ensuring the validity and applicability of the results in similar contexts.

Abbreviations

ACC: Make conclusions, ACD: Review the data, CFA: Confirmatory Factor Analysis, FPH: Formulate Questions, FPR: Proposes Answers (Hypothesis), PNM: Choose materials, PNT: Plan the work, RCCI-STEM: Rubric to Assess Scientific Inquiry Competence and Its Relation with STEM Areas, RD: Write or draw what you do, STEM: Science, Technology, Engineering, and Mathematics.

Acknowledgment

The authors express their sincere gratitude to the research team from the Instituto de Ciencias Aplicadas y Tecnología (ICAT-UNAM) for their valuable conceptual and methodological contributions throughout the development of this study.

Author Contributions

All authors have sufficiently contributed to the study and agreed with the results and conclusions. Fanny de Jesús Rúa Martínez: Conceptualization, conducting the research, and drafting the original manuscript; Martha Andrea Merchán Merchán: Supporting the conceptual framework, methodological design, and supervision; Elena Calderón-Canales: Conceptual advising, review, and editing; Beatriz García-Rivera: Conceptual advising, review, and editing.

Conflict of Interest

The authors declare that there are no actual or potential conflicts of interest regarding the research, authorship, or publication of this article.

Declaration of Artificial Intelligence (AI) Assistance

The study employed AI assisted technologies in the process of translating from Spanish to English.

Ethics Approval

This study received ethical approval from the University Ethics Committee, in accordance with Act No. 3 dated May 14, 2024. Participation was voluntary and informed consent was obtained as appropriate. For students involved in the pilot study, informed consent was provided by their parents or legal guardians, and the educational institution formally authorized its implementation. Confidentiality was ensured, and data were used exclusively for academic and research purposes.

Funding

No funding sources.

References

1. Roy G, Sikder S, Letts W. Understanding the process of scientific literacy development among children in the early years through play and intentionality. Res Sci Educ. 2025;55(6):1–22. https://doi.org/10.1007/s11165-025-10291-9
2. Alali R, Al-barakat A. Young children's attitudes toward science learning in early learning grades. Asian Educ Dev Stud. 2025;13(4):340–55. https://doi.org/ 10.1108/AEDS-02-2024-0036
3. Kwarteng M. Practical application of Piaget's cognitive theory and Vygotsky's sociocultural theory in classroom pedagogy. J Stud Educ. 2025;15(2):68–91. https://doi.org/10.5296/jsse.v15i2.22703
4. Björnhammer S, Gyllenpalm J, Lundegård I. A study of learning progression during open inquiry in year 9: learning to do and learning about scientific inquiry. Sci Educ. 2025;34(6):1–26. https://doi.org/10.1007/s11191-025-00705-5
5. Larkin K, Lowrie T. Teaching approaches for STEM integration in pre- and primary school: a systematic qualitative literature review. Int J Sci Math Educ. 2023;21(Suppl 1):11–39. https://doi.org/10.1007/s10763-023-10362-1
6. Fridman R, Eden S, Spektor-Levy O. Nascent inquiry, metacognitive, and self-regulation capabilities among preschoolers during scientific exploration. Front Psychol. 2020; 11:1–16. https://doi.org/10.3389/fpsyg.2020.01790
7. Hidi SE, Renninger KA. On educating, curiosity, and interest development. Behavioral Sciences. 2020;35 :99–103. https://doi.org/10.1016/j.cobeha.2020.08.002
8. García-Carmona A. From inquiry-based science education to the approach based on scientific practices. Sci Educ. 2020;29(2):443–63. https://doi.org/10.1007/s11191-020-00108-8
9. Lazonder AW, Harmsen R. Meta-analysis of inquiry-based learning: effects of guidance. Rev Educ Res. 2016;86(3):681–718. https://doi.org/10.3102/0034654315627366
10. García-Rodeja I, Barros S, Sesto V. Inquiry-based activities with woodlice in early childhood education. Educ Sci. 2024; 14:710. https://doi.org/10.3390/educsci14070710
11. Halawa S, Lin T-C, Hsu Y-S. Exploring instructional design in K-12 STEM education: a systematic literature review. Int J STEM Educ. 2024; 11:43. https://doi.org/10.1186/s40594-024-00503-5
12. Cantlon JF, Becker KT, Delong CM. Computational thinking during a short, authentic, interdisciplinary

STEM experience for elementary students. J STEM Educ Res. 2024;7(3):425–43. https://doi.org/10.1007/s41979-024-00117-0
13. Khalid IL, Syahrir M, Fadzil H. A systematic review: Digital learning in STEM education. J Adv Res Appl Sci Eng Technol. 2025;1(1):98–115. https://doi.org/10.37934/araset.51.1.98115
14. Uyulan V, Aslan S. Development of students' critical thinking skills with STEM activities in early childhood science education. Anadolu Univ J Educ Fac. 2024;8(3):971-99. https://doi.org/10.34056/aujef.1464605
15. Monteira SF. Creative approaches for inclusive STEM learning in early years. Early Child Educ J. 2025;53(8):1–14. https://doi.org/10.1007/s10643-025-01922-x
16. Yalçın V, Erden Ş. The effect of STEM activities prepared according to the design thinking model on preschool children's creativity and problem-solving skills. Thinking Skills and Creativity. 2021; 41:100864. https://doi.org/10.1016/j.tsc.2021.100864
17. Pedaste M, Baucal A, Reisenbuk E. Towards a science inquiry test in primary education: development of items and scales. Int J STEM Educ. 2021;8(1):19. https://doi.org/10.1186/s40594-021-00278-z
18. Johnston K, Kervin L, Wyeth P. STEM, STEAM and makerspaces in early childhood: A scoping review. Sustainability. 2022;14(20):1–20. https://doi.org/10.3390/su142013533
19. Care E, Luo R. Assessment of transversal competencies: Policy and practice in the Asia-Pacific region. UNESCO; 2016. https://unesdoc.unesco.org/ark:/48223/pf0000246590.locale=en
20. Wolf S, Jukes MCH, Yoshikawa H, *et al*. Examining the validity of an observational tool of classroom support for children's engagement in learning. Early Child Educ J. 2025;53(4):1325–39. https://doi.org/10.1007/s10643-024-01731-8
21. Kinyota M. Reconciling teachers' views and practices with early graders' ability to engage in scientific inquiry. Sci Educ Int. 2023;34(4):274–82. https://doi.org/10.33828/sei.v34.i4.3
22. English LD. STEM education K-12: Perspectives on integration. Int J STEM Educ. 2016;3(1):1–8. http://dx.doi.org/10.1186/s40594-016-0036-1
23. Chen YL, Tippett CD. Project-based inquiry in STEM teaching for preschool children. Eurasia J Math Sci Technol Educ. 2022;18(4):1–15. http://doi.org/10.29333/ejmste/11899
24. Brenneman K. Assessment for preschool science learning and learning environments. Early Child Res Pract. 2011;13(1):1–9. https://ecrp.illinois.edu/v13n1/brenneman.html
25. Brenneman K, Stevenson-Garcia J, Kwanghee J, *et al*. The Preschool Rating Instrument for Science and Mathematics (PRISM). In: Society for Research on Educational Effectiveness (SREE). 2011. https://files.eric.ed.gov/fulltext/ED528753.pdf
26. La Paro K, Pianta R, Stuhlman M. The classroom assessment scoring system: findings from the prekindergarten year. *Elem Sch J.* 2004;104(5):409–26. https://www.jstor.org/stable/3202821
27. Marshall J, Horton R, White C. EQUIPping teachers. Sci Teach. 2009;76(4):46–53. https://alapex.org/mediawiki/images/d/de/2_Article_EQUIPping_Teachers_for_Inquiry.pdf
28. Milford T, Tippett C. The design and validation of an early childhood STEM classroom observational protocol. Int Res Early Child Educ. 2015;6(1):24–37. https://eric.ed.gov/?id=EJ1150965
29. Trina N, Monsur M, Cosco N, *et al*. Tools for assessing the STEAM learning affordances and quality of outdoor learning environments of childcare centers: A systematic review. Early Child Educ J. 2025;1-7. https://doi.org/10.1007/s10643-025-01911-0
30. Reynders G, Lantz J, Ruder SM, *et al*. Rubrics to assess critical thinking and information processing in undergraduate STEM courses. Int J STEM Educ. 2020;7(1):9. https://doi.org/10.1186/s40594-020-00208-5
31. Mejeh M, Hagenauer G, Gläser-zikuda M. Mixed methods research on learning and instruction: meeting the challenges of multiple perspectives and levels within a complex field. Forum Qual Soc Res. 2023;24(1):1–24. https://doi.org/10.17169/fqs-24.1.3989
32. Creswell JW, Plano Clark VL. Designing and conducting mixed methods research. SAGE; 2018. https://bayanbox.ir/view/236051966444369258/9781483344379-Designing-and-Conducting-Mixed-Methods-Research-3e.pdf
33. Urdanivia Alarcon DA, Talavera-Mendoza F, Rucano Paucar FH, Cayani Caceres KS, Machaca Viza R. Science and inquiry-based teaching and learning: a systematic review. Front Educ. 2023; 8:1170487. https://doi.org/10.3389/feduc.2023.1170487
34. Froschauer L. Bringing STEM to the elementary classroom. NSTA, editor. Arlington, Virginia; 2016. https://my.nsta.org/resource/103773/bringing-stem-to-the-elementary-classroom-e-book
35. Pedaste M, Mäeots M, Siiman LA, *et al*. Educational Research Review Phases of inquiry-based learning: Definitions and the inquiry cycle. Educ Res Rev. 2015; 14:47- 61. http://dx.doi.org/10.1016/j.edurev.2015.02.003
36. Etikan I, Musa SA, Alkassim RS. Comparison of convenience sampling and purposive sampling. Sci Publ Gr. 2016;5(1):1–4. https://doi.org/10.11648/j.ajtas.20160501.11
37. Camacho-Tamayo E, Bernal-Ballén A. Validation of an instrument to measure natural science teachers' self-perception about implementing STEAM approach in pedagogical practices. Educ Sci. 2023;13(8):764. https://doi.org/10.3390/educsci13080764
38. Fernández-Gómez E, Martín-Salvador A, Luque-Vara T, *et al*. Content validation through expert judgement of an and habits of pregnant women. Nutrients. 2020;12(4):1136. https://doi.org/10.3390/nu12041136
39. Hasselgreen A, Khalifa H, Weir CJ, *et al*. Second language assessment and mixed methods research. Sage Publication. 2017. https://www.cambridgeenglish.org/Images/735110-studies-in-language-testing-volume-43.pdf
40. Penfield RD, Giacobbi JPR. Applying a score confidence interval to Aiken's item content-relevance index. Meas Phys Educ Exerc Sci. 2009;8(4):213–25.

https://doi.org/10.1207/s15327841mpee0804_3
41. Aya-Roa K, Beltrán-Campos V, Hernández J, García-Campos M, Ramírez-Gómez X, Nuñez-Colín C. Design and content validation of the Transcend instrument for family caregivers. Rev Cuid. 2025;16(2). https://doi.org/10.15649/cuidarte.4595.
42. Golzar J, Tajik O, Noor S. Convenience Sampling. Int J Educ Lang Stud. 2022;1(2):72–7. https://doi.org/10.22034/ijels.2022.162981
43. UNESCO. Tashkent declaration and commitments to action for transforming early childhood care and education. World Conf Early Child Care Educ. 2022. https://unesdoc.unesco.org/ark:/48223/pf0000384045
44. Muasya JN, Mulwa PK. Pilot study, a neglected part of qualitative and quantitative research process: evidence from selected PhD thesis and dissertations. Sci Publ Gr. 2023;8(4):115–23. https://doi.org/10.11648/j.her.20230804.11
45. Brown TA. Confirmatory factor analysis for applied research. Guilford Press; 2015. https://www.guilford.com/books/Confirmatory-Factor-Analysis-for-Applied-Research/Timothy-Brown/9781462515363
46. Kline RB. Principles and practice of structural equation modeling. Guilford Press; 2023. https://www.researchgate.net/publication/379694768
47. Hair JJ, Black WC, Babin BJ, Anderson R. Multivariate data analysis. Cengage; 2019. https://au.cengage.com/c/multivariate-data-analysis-8e-hair-babin-anderson-black/9781473756540/
48. Stefana A, Damiani S, Granziol U, *et al*. Psychological, psychiatric and behavioral sciences measurement scales: best practice guidelines for their development and validation. Front Psychol. 2025; 15:1494261. https://doi.org/10.3389/fpsyg.2024.149426
49. National Research Council. A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. National Academies Press; 2012. https://doi.org/10.17226/13165
50. Byrne BM. Structural equation modeling with Amos: Basic concepts, applications and programming. Routledge; 2016. https://doi.org/10.4324/9781315757421
51. Minner DD, Jurist A, Century J. Inquiry-based science instruction -What is it and does it matter? Results from a research synthesis Years 1984 to 2002. J Res Sci Teach. 2010;47(4):474–96. https://doi.org/10.1002/tea.20347
52. Campbell C, Jobling W, Howitt C. Science in early childhood. 2021. p. 91–104. https://api.pageplace.de/preview/DT0400.9781108871853_A45556181/preview-9781108871853_A45556181.pdf
53. McNeish D. Thanks coefficient Alpha, We'll Take it from here. Psychol Methods. 2017;23(3):412–33. https://psycnet.apa.org/doi/10.1037/met0000144
54. Taber K. The use of Cronbach's Alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018;48(6):1273-96. https://doi.org/10.1007/s11165-016-9602-2
55. Darman DR, Suhandi A, Kaniawati I, *et al*. Development and validation of Scientific Inquiry Literacy Instrument (SILI) using Rasch measurement model. Educ Sci. 2024;14(3):322. https://doi.org/10.3390/educsci14030322
56. Tavakol M, Dennick R. Making sense of Cronbach' s alpha. Int J Med Educ. 2011; 2:53–5. https://doi.org/10.5116/ijme.4dfb.8dfd
57. Kline RB. Psychometric theory. In: Journal of Psychoeducational Assessment. McGraw-Hill; 1999. p. 275–80. https://doi.org/10.1177/073428299901700307
58. Ghazali N. A Reliability and validity of an instrument to evaluate the school-based assessment system: A pilot study. Int J Eval Research Educ. 2016;5(2):148–57. https://ijere.iaescore.com/index.php/IJERE/article/viewFile/4533/2267
59. Maric D, Fore GA, Nyarko SC, Varma-Nelson P. Measurement in STEM education research: A systematic literature review of trends in the psychometric evidence of scales. Int J STEM Educ. 2023;10(1):39. https://doi.org/10.1186/s40594-023-00430-x
60. Kelley TR, Knowles JG. A conceptual framework for integrated STEM education. Int J STEM Educ. 2016;3(1):11. https://doi.org/10.1186/s40594-016-0046-z
61. Baxter J, Lam J, Povey J, *et al.* Family dynamics over the life course. 2022. 37–56. https://link.springer.com/book/10.1007/978-3-031-12224-8
62. Udhnani M, Becker L, Lecavalier L. Performance of autism screening and diagnostic instruments among children with ADHD: A systematic review. J Autism Dev Disord. 2025;1–12. https://doi.org/10.1007/s10803-025-06857-1

How to Cite: Martínez FdJR, Merchán MAM, Calderón-Canales E, García-Rivera B. STEM Scientific Inquiry Rubric for Early Childhood. Int Res J Multidiscip Scope. 2026;7(1):1161-1174.
DOI: 10.47857/irjms.2026.v07i01.08619