Humanities and Social Sciences Letters
2026 Vol. 14, No. 4, pp. 592-614
ISSN(e): 2312-4318
ISSN(p): 2312-5659
DOI: 10.18488/73.v14i4.5333
© 2026 Conscientia Beam. All Rights Reserved.

Constructing and developing a classroom performance observation card for teachers of gifted students in light of the pentagram strategy: A psychometric study

Rommel Al Ali[1+]
Shoeb Saleh[2]

[1,2]*The National Research Center for Giftedness and Creativity, King Faisal University, Al-Ahsa 31982, Saudi Arabia.*
[1]*Email: ralali@kfu.edu.sa*
[2]*Email: sgsaleh@kfu.edu.sa*

(+ Corresponding author)

Article History
Received: 9 February 2026
Revised: 13 April 2026
Accepted: 7 May 2026
Published: 29 October 2026

Keywords
Classroom performance
Gifted education
Higher-order thinking
Observation card
Pentagram strategy
Psychometric properties
Teaching performance evaluation.

ABSTRACT

This study aimed to construct and psychometrically validate a specialized classroom observation instrument for teachers of gifted students, centered on the Pentagram Strategy framework. Recognizing the scarcity of assessment tools tailored to high-ability instruction, the research employed a methodological-psychometric design to bridge gifted education theory with rigorous measurement standards. The instrument was developed within an integrated framework encompassing structured thinking, higher-order cognitive development, and contemporary evaluation standards. The final version of the observation card comprises five distinct dimensions reflecting the stages of the Pentagram Strategy: Problem Identification (What?), Meaning-Building and Motivation (Why?), Flexible Implementation (How?), Critical and Future-Oriented Thinking (What if?), and Authentic Assessment and Self-Reflection. To evaluate the tool's effectiveness, it was administered to a sample of gifted education teachers and analyzed using advanced statistical techniques. Validity was established through Content Validity and Construct Validity via Exploratory and Confirmatory Factor Analyses (EFA/CFA). Reliability was rigorously confirmed using McDonald's Omega and Composite Reliability indices. Furthermore, the application of the Rasch Measurement Model provided robust evidence of the card's measurement precision and its capacity to differentiate between varying levels of instructional performance. The findings demonstrate that the observation card is a highly reliable, standardized instrument with strong convergent and discriminant validity. Consequently, this study offers a psychometrically sound resource for educational supervisors, researchers, and practitioners, providing a standardized basis for professional development and scientific evaluation within the field of gifted education.

Contribution/Originality: This study provides the first psychometrically validated instrument specifically based on the Pentagram Strategy. By integrating the Rasch Measurement Model with traditional factor analysis, it offers a high-precision, standardized framework for evaluating gifted education, filling a critical methodological gap in professional supervision and classroom performance assessment.

1. INTRODUCTION

The education of gifted students is a specialized pedagogical field requiring instructional practices that transcend traditional methods, given these students' accelerated learning pace, depth of cognitive processing, and advanced skills in analysis, synthesis, and foresight. Literature indicates that the effectiveness of gifted programs depends not only on curricular quality but is also significantly linked to teacher proficiency in employing instructional strategies that foster high-level thinking (VanTassel-Baska, 2018a). Contemporary educational trends emphasize a shift toward

© 2026 Conscientia Beam. All Rights Reserved.

active learning, problem-solving, and cognitive development models, aligning with the demands of the knowledge society and the creative economy (Bonwell & Eison, 1991; Prince, 2004).

Among the modern instructional frameworks that organize thinking processes and promote deep learning is the Pentagram Strategy. This strategy follows a cognitive sequence progressing from problem identification and analysis to the generation, evaluation, and application of alternatives in novel contexts, thereby supporting structured and creative thinking, particularly among gifted learners (El-Sayed, 2021; Salama, 2018).

Classroom observation remains a fundamental tool for assessing teacher performance and analyzing instructional practices, as it facilitates the direct collection of data on interaction patterns and applied teaching strategies (Fraenkel, Wallen, & Hyun, 2019; Shahata & Al-Najjar, 2017). However, many available observation instruments are characterized by excessive generality, failing to account for the cognitive and emotional characteristics of gifted students or lacking a foundation in modern thinking-based strategies such as the Pentagram Strategy (Salama, 2018; VanTassel-Baska, 2018b).

Measurement and evaluation literature underscores that the validity of assessment tools relies on robust psychometric properties, such as validity, reliability, and internal consistency, to ensure the dependability of results for educational decision-making (Crocker & Algina, 2008; DeVellis, 2017). Nevertheless, numerous teacher evaluation tools, especially in gifted education, lack comprehensive psychometric standardization, which limits their effectiveness in diagnosing and developing instructional practices (VanTassel-Baska & Stambaugh, 2021).

In light of this, the research problem emerges from the deficiency in classroom observation tools for teachers of the gifted. These tools often lack precise behavioral indicators reflecting effective instruction for this population and are frequently disconnected from clear strategic frameworks like the Pentagram Strategy. Furthermore, there is a scarcity of standardized Arabic instruments possessing reliable psychometric properties (Al-Harthi, 2020; Danielson, 2013; VanTassel-Baska, 2018a).

Addressing this gap, the current study aims to construct and standardize a classroom performance observation card for teachers of gifted students based on the Pentagram Strategy. It seeks to verify its psychometric properties (validity, reliability, and internal consistency), providing a scientific tool for evaluating and developing teaching performance in gifted education (Crocker & Algina, 2008; DeVellis, 2017).

Consequently, the study's problem is defined by the following primary question: How can a classroom performance observation card for teachers of gifted students be constructed and standardized based on the Pentagram Strategy, and what are its psychometric properties?

This primary question branches into the following sub-questions.

1. What are the theoretical and educational foundations of the Pentagram Strategy, and what are its instructional requirements for teaching gifted students in the classroom?
2. What dimensions and performance indicators should be included in the observation card for teachers of gifted students considering the Pentagram Strategy?
3. What methodological steps are appropriate for developing a classroom observation card for teachers of gifted students based on the Pentagram Strategy?
4. What are the psychometric properties (validity and reliability) of the classroom observation card for teachers of gifted students based on the Pentagram Strategy?

1.1. Significance of Study

1.1.1. Theoretical Significance

The most important part of the study is the contribution of the Pentagram Strategy to literature, framed as an active learning approach where learners organize their thinking, motivate themselves, and connect learning to real-life situations. The researcher situates this strategy in the domain of gifted education, which is often accompanied by high complexity in the required pedagogy. The researcher also contributes to educational measurement and

© 2026 Conscientia Beam. All Rights Reserved.

evaluation by developing the classroom observation card through systematic design, expert review, and adherence to the standardized process with established validity and reliability in the context of Arabic gifted education.

The study also proposes an integrative model that combines a contemporary instructional strategy (the Pentagram Strategy) with measurement and evaluation in gifted education. This creates an opportunity for future research to create additional instruments, such as an attitude scale or a questionnaire on teaching practices, that align with the same strategy or other similar strategies. Moreover, the findings may assist researchers in pinpointing which instructional elements most prominently support the development of higher-order thinking skills in gifted students, as evidenced through the construct validity and factor structure of the instrument.

1.1.2. Applied Significance

Educational supervisors, school administrators, and gifted education program professionals have a practical provision of a standardized observational tool, which is a constructive resource to evaluate teachers against specified criteria that pertain to supportive behavioral and emotional climate, interactional management, thinking-skill promotion, the use of formative assessments, and other characteristics of effective teaching.

The proposed tool also provides an opportunity for teachers to assess their own performance and identify their own strengths and weaknesses to plan for professional development activities using the Pentagram Strategy, thereby improving instructional practices and the behavioral and emotional development of gifted learners.

Additionally, the card offers a supervisor's standardized evaluative tool for teachers of gifted students, minimizing the use of subjective impressions and evaluative summary cards that are so broad they overlook the specifics of the population served, thereby improving evaluative equity and objectivity.

Lastly, the card's results may guide training that aims to strengthen instructional areas and help to frame professional development activities that transform classroom observation from a boring supervisory task into a clear driver of change and improvement.

1.1.3. Previous Studies

Previous studies related to the research topic can be categorized into three primary themes: studies addressing the Pentagram Strategy and thinking-organization strategies; research concerning gifted education and teacher classroom performance; and studies focused on the construction of observation tools and the verification of their psychometric properties.

First, educational literature emphasizes the significance of active learning strategies based on cognitive organization in improving learning outcomes and fostering higher-order thinking skills. Research indicates that these strategies enhance learner engagement, motivation, critical thinking, and decision-making (Bonwell & Eison, 1991; Prince, 2004). Within this framework, the Pentagram Strategy has emerged as an instructional model designed to organize cognitive processes through sequential stages, beginning with problem identification and analysis, followed by generating alternatives, predicting outcomes, and evaluation (Salama, 2018). Empirical studies have confirmed the efficacy of this strategy in developing deep thinking skills; recent findings demonstrate significant improvements in critical reading, deep comprehension, and creative problem-solving among students taught via this strategy compared to traditional methods (Al-Asadi & Al-Fatlawi, 2021; Amoush & Khodair, 2025; El-Sayed, 2021; Kanaan, Muhammad, & Hussein, 2024). These results suggest that the Pentagram framework provides a robust instructional structure capable of fostering deep learning, making it particularly suitable for gifted students who require learning strategies grounded in inquiry, analysis, and creative synthesis.

Second, literature on gifted education highlights the decisive role of a teacher's classroom practices in nurturing student potential. Studies assert that the quality of instruction for the gifted largely depends on the teacher's ability to pose intellectual challenges, cultivate creativity, and provide an inquiry-rich, cognitively interactive environment (VanTassel-Baska, 2018a). However, some studies identify a persistent gap between the theoretical principles of gifted

© 2026 Conscientia Beam. All Rights Reserved.

education and actual classroom practices; instructional patterns for the gifted often do not significantly differ from those provided to general education students, thereby limiting the investment in their latent potential (VanTassel-Baska & Brown, 2007). This context underscores the necessity for precise assessment tools capable of monitoring instructional practices associated with fostering higher-level thinking.

Third, a body of research has addressed the construction and psychometric validation of educational assessment and observation tools. Educational measurement literature stresses the importance of utilizing advanced analytical models when developing tools for the gifted to ensure estimation accuracy and the ability to discriminate between varying performance levels (Al-Momani & Al-Oweidi, 2020; Nakano & Peixoto, 2023). Other studies have demonstrated the value of structured classroom observation tools in monitoring practices, such as those used in project-based learning environments, which provide vital feedback for teachers and guide professional development programs (Capraro, Capraro, Scheurich, Jones, & Morgan, 2012). Similarly, the Stallings System is a well-recognized model in quantitative classroom observation, relying on systematic coding of teacher-student behaviors to provide reliable data on interaction and time management, yielding high inter-rater reliability when observers are sufficiently trained (World Bank, 2010). Furthermore, recent studies on gifted rating scales emphasize the importance of verifying an instrument's internal structure through Confirmatory Factor Analysis (CFA) and various reliability coefficients to ensure its validity for research and applied purposes (Gifted Rating Scale – School Form, 2025).

Despite the contributions of these studies to instructional strategies, gifted education, and assessment tool development, a review of the literature reveals a research gap: there is a limited availability of specialized tools that integrate classroom performance observation for teachers of the gifted with a clear instructional framework like the Pentagram Strategy. Moreover, there is a scarcity of standardized Arabic instruments that have undergone systematic psychometric validation. Consequently, the current study seeks to fill this gap by constructing and standardizing a classroom performance observation card for teachers of gifted students based on the Pentagram Strategy and verifying its psychometric properties.

1.1.4. Integration of Previous Studies into the Current Study

The design of the classroom observation card for teachers of gifted students, developed in conjunction with the Pentagram Strategy, relied on the methodological, scientific, and empirical insights of prior research. Educational research on the instruction of the gifted developed a framework for the optimal classroom functioning of a teacher, in particular, the teacher's role in the promotion of higher-order thinking, interaction facilitation, and the creation of a positive, intellectually challenging classroom environment. These factors were used to design dimensions and behavioral indicators of the card, which correspond to the cognitive and affective levels of gifted students.

The research on active and guided learning, more specifically research concerning the Pentagram Strategy, enabled the researcher to determine the anticipated order of the instructional sequence within a single classroom. This, in turn, permitted the researcher to align the various items on the card to each of the Phases of the Pentagram Strategy in a way that is observable and measurable while addressing the concerns of previous research, which were associated with loose, overly broad, and vague statements.

Psychometric research within the construct of measurement and evaluation has also shaped the methodology. This included item writing with observable and verifiable behavioral indicators, submission of the card and items for expert review and feedback in the area of content validity, and the application of empirical methods to test for construct validity and reliability. Previous research also illustrated the value of inter-rater reliability concerning observation and evaluation of classroom practice, a principle that helped shape the standardization of the observation tool in this study.

Additionally, the examination of prior research showed a lack of specificity in existing instruments, i.e., they were overly general, misaligned with the gifted education field, and did not incorporate current teaching methodologies.

© 2026 Conscientia Beam. All Rights Reserved.

Accordingly, the current study was better able to target its problem and justify the necessity of designing a proprietary instrument with a well-defined theoretical basis and psychometric documentation.

Informed by the above, the current study sought to translate the previous research into the construction of a standardized observation card, which, due to its combination of theoretical specificity, procedural structure, and psychometric compatibility, will be more effective in assessing the classroom performance of teachers of gifted students in evaluative and research contexts.

1.2. Theoretical Framework

This theoretical framework is based on the premise that the education of gifted students is an econometric model that detects certain teaching practices of high quality and that the results of such practices are best captured in the teacher's own classroom, where she is the critical educational nexus between theory and practice. Consequently, the framework captures the education of gifted learners and the nexus of teacher performance, the Pentagram Strategy as an organizational framework for thinking and learning, the role of classroom observation as an evaluative framework for teaching performance, and the principles and psychometric parameters of the construct and framework of an educational measurement tool.

Gifted education is an area of practice that is built on the assertion that gifted learners are different in the way they think and feel, and that they possess a unique constellation of cognitive and affective attributes such as quick understanding and mentally intensive processing, cognitive agility, heightened problem sensitivity, a readiness to exercise autonomy, and a strong intellectual self-challenge (Clark, 2013). Current models of giftedness, especially Renzulli's Three-ring Conception, largely posit that gifted behaviour is a product of the confluence of above-average, three core dimensions of ability, creativity, and commitment to a given task and the presence of a supportive environment (Renzulli, 2016a).

The success of gifted programs isn't only reliant on the curriculum. Rather, it is more strongly connected to the gaps in teacher training and strategies that promote advanced levels of thinking that encourage richer, more complex, and more productive learning environments (Tomlinson, 2014). Several studies have identified the discrepancies between the theoretical foundations of gifted education and the practical applications in the classroom, specifically that many teachers continue to use stagnant methods that do not meaningfully differ from targeting the general education students, and in doing so, they ignore and undercut the advanced potential of gifted students in their classrooms. Therefore, student achievement becomes the most visible indicator of the quality of what and how they have been taught, evidenced by the kinds of questions posed by the teacher, the learning tasks assigned, the nature of the student-to-student and teacher-to-student interactions, the instructional activities designed, and the assessments administered. This reality underscores the importance of teacher evaluation frameworks that are designed to be as precise and unbiased as possible in order to illuminate teachers' strengths and weaknesses.

The evaluation of how teachers of gifted learners operate in classrooms utilizes a range of modern learning theories, which underline the active learner and the teacher as a facilitator and architect of the learning milieu. Constructivism posits that learning is not concerned with the transference of knowledge, but with the active appropriation of meaning through the engagement with and resolution of learning problems and situations (Piaget, 1972; Vygotsky, 1978). This dynamic is particularly salient in the field of gifted education, as evidence suggests that gifted learners thrive in environments that utilize the inquiry approach, which facilitates exploration, experimentation, construction of hypotheses, and evaluation of competing alternatives.

From this perspective, teacher effectiveness signifies the practical application of the principles of constructivist learning. It is most apparent in the teacher's ability to create intellectually challenging learning opportunities, promote constructive dialogue in the classroom, articulate unresolved problems, and furnish constructive criticism that assists learners in redefining their understanding. An effective teacher of gifted learners is described in the literature as one who successfully strikes a harmony in the role of the facilitator of the learning process in a manner

© 2026 Conscientia Beam. All Rights Reserved.

that is complemented by providing learners adequate opportunities for thoughts and ideas to be autonomous (VanTassel-Baska & Stambaugh, 2006).

The Pentagram strategy refers to an active learning pedagogy where educators can structure learning into five distinct stages: problem identification, problem analysis and interpretation, generation of alternatives, outcome prediction, and the subsequent evaluation and application of the outcomes to a different context. The strategy is grounded in constructivist principles, where the most meaningful learning is said to occur in instances where learners construct their own knowledge and self-regulate their own thinking (Jonassen, 1999).

Research conducted in the Arab world and other countries in the world has proven that the Pentagram Strategy optimally and successfully cultivates organized thinking, problem-solving, comprehension, and creative thinking. Students who have been taught using this strategy have outperformed their peers in terms of analysis, reasoning, and evaluation (Al-Asadi & Al-Fatlawi, 2021). The Pentagram Strategy is particularly aligned with gifted learners, as it utilizes ill-structured, open-ended problems and higher-order thinking tasks. For the Pentagram Strategy to be effective in a classroom, it relies heavily on the teacher to design problem-based pedagogical approaches, facilitate purposeful classroom dialogue, and control the flow of instruction to alternate between problem-solving and outcome analysis and evaluation, as well as outcome ideation. As a result, teacher performance behavioral indicators need to be set for each of the strategy's five stages to facilitate observations and evaluations.

The Pentagram Strategy is noted among teaching practices that try to provide frameworks for tackling and simplifying intricate thinking processes. Its steps correspond to some of the higher-order thinking categories in the revised Bloom's taxonomy, particularly those concerning analysis, evaluation, and creation (Anderson & Krathwohl, 2001). Moving through the various steps demands that learners not just grasp the content, but that they also break down a problem, connect various pieces of information, develop a range of plausible options, and evaluate the possible outcomes of each based on specific criteria. The structured nature of the steps reinforces the importance of the teacher in controlling the thinking that takes place in the classroom as opposed to merely facilitating the 'coverage' of the content. As such, when using the Pentagram Strategy in teaching, the focus of the observation and evaluation of the strategy employed should be on how thinking is directed by the teacher and not on the content that is being 'presented' to the learners. This provides the basis for developing an observation tool based on the strategy and documenting the phases of the strategy in relation to the clearly defined, observable, and measurable actions of the teacher.

Various scholars and practitioners describe structured classroom observation as one of the most effective ways to evaluate teaching, as it provides the most direct representation of teachers' actions and interactions in real-life settings and the most straightforward analysis of their interactions and practices (Gronlund & Brookhart, 2009). Evidence suggests that observation checklists with criteria that are more specific to dimensions of teaching provide more objectivity and are more effective for formative assessment and professional growth than subjective impressions. Nevertheless, most existing observation instruments are too broad and do not address the specific needs of gifted education and modern teaching. Many are still focused on teaching practice, such as classroom control and the clarity of explanation, and neglect cognitive and qualitative aspects of thinking. This suggests the importance of developing such observation tools that are more specialized and can convert teaching frameworks, such as the Pentagram, into observable and actionable teaching behaviors.

The literature surrounding educational measurement details the psychometric principles that outline the characteristics of a measurement tool (or the evaluative tool) for it to be considered high in quality. The first is content validity. An expert review demonstrates that a measurement item reflects the domain of the performance to be measured (Crocker & Algina, 2008). Construct validity is a measurement of the extent to which the tool is a reflection of a set of constructs. It is also evaluated through exploratory or confirmatory factor analyses. The operational definition of the term reliability is the degree to which measurement results are replicable and consistent. The other measurement types that each of the review tools has are internal consistency, test-retest reliability, and inter-rater

© 2026 Conscientia Beam. All Rights Reserved.

reliability. Without reliability, the tool is not useful for any supervisory or educational decisions (DeVellis, 2017). Also, as a tool in gifted education, the review item is evaluative of whether or not the components are built to differentiate among the various dimensions or levels of teacher performance.

Based on the above, the theoretical framework for this study combines three elements: the teaching practices pertaining to gifted education, the Pentagram Strategy as a framework for structured thinking, and observation tools for the classroom aligned with psychometric criteria. These factors make it possible to ascertain unambiguous behavioral components and indices related to teaching performance. Furthermore, behavioral components reinforce the reasoning for the methodology and the statistical techniques utilized in the construction and standardization of the observation card, which serves to enhance teaching practices and the overall quality of gifted education.

1.3. Behavioral Dimensions of Classroom Performance in Light of the Pentagram Strategy

1. Stage (What?): Advanced Problem Identification and Conceptual Framing

This stage involves creating what is known as an intellectually demanding Learning Situation and pinpointing the foundational concepts that serve as the entry point for learning. Studies in gifted education show that gifted students need multi-layered problems that are open-ended and complex. Such problems encourage the students to devise their own solutions rather than relying on direct and conventional answers (Maker, Jo, & Muammar, 2008; VanTassel-Baska & Brown, 2007). From the description of this stage, one can assess a teacher's capability to develop instructional goals that incorporate some elements of the revised Bloom's taxonomy in the affective and cognitive domains, integrate problems in realistic and/or research-and-forecast-based ramifications, and use models, abstractions, and concepts at the cognitive level of gifted students. The stage also involves the scaffolding of, in the language of problem-based learning (PBL), students expanding dimensions of the problem and posing complex questions, an aspect that is empirically supported by the problem-posing research, which underscores problem definition and quality as paramount to promote deep learning (Hmelo-Silver, 2004).

2. Stage (Why?): Intellectual Motivation and Meaning Making

This stage pertains to the breakdown and construction of meaning from a given problem, through analysis and justification of its constituent elements and the relationships between them. The literature suggests that gifted learners' intellectual motivation is fostered to the greatest extent when they are provided with a high degree of freedom in selecting and comparing different approaches to analysis and in expressing themselves through evidence-based dialogues of a scientific nature (Renzulli, 2016b; Tomlinson, 2014). The description of effective performance at this stage of the classroom involves the provision of diverse analysis frameworks, systematic decomposition of the problem into smaller constituent parts, the use of various learning and sophisticated analysis tools, the encouragement of scientific dialogues, and the exchange of ideas. The literature on active learning and critical thinking indicates that these practices are important for fostering a deeper understanding and minimizing shallow and rote learning (Prince, 2004).

3. Stage (How?): Flexible Implementation and Constructing Explanations

This stage illustrates the probable outcomes and the associated phenomena and develops explanatory frameworks that involve causal and inferential reasoning. Gifted scholars are reported to demonstrate exceptional ability in processing and handling varied explanations and in requiring varied classroom experiences, where they can re-examine, provide justifications, and defend their position scientifically (VanTassel-Baska, 2018b). A teacher's achievements at this stage are seen in the instances where students are assisted to deliver coherent and logical explanations, connect cause and effect, demonstrate cause-and-effect reasoning, and demonstrate conceptual clarity in precise and unequivocal meaning. These approaches are consistent with the constructivist view that characterizes explaining as a reorganization of knowledge, rather than an end product of the learning process (Chi, 2009).

© 2026 Conscientia Beam. All Rights Reserved.

4. Stage (What If? What Are the Consequences?): Critical and Futures Thinking

This stage signifies the height of critical and future thinking. In this case, students will need to anticipate the outcomes of differing alternatives and assess the consequences of each, make responsible decisions, and analyze the impact of their decisions (short/long term). According to the literature, thinking about the future and creating scenarios as a gifted education practice is viewed as a means of exercising cognitive flexibility, probabilistic reasoning, and the ability to deal with ambiguity and uncertainty (Beghetto & Kaufman, 2014). The teacher's performance in this stage is exhibited through the use of sophisticated and complex "what if" scenarios, instruction in which students consider the "what if" alternatives in a positive and ethically/professionally aligned prediction coupled with a decision, and the ability to offer spontaneous, even radical proposals. The practice of creative thinking supports the notion that these behavioral tendencies foster originality and cognitive flexibility (Guilford, 1967).

5. Stage (Evaluation and Feedback): Authentic Assessment and Reflective Thinking

The main focus of the last stage is the authentic assessment of the learning outcomes and the self-reflection on the mental processes and the transfer of learning to different situations. As stated in the literature on educational assessments, the most effective assessments for gifted learners should be formative and authentic, criteria-based, and should help students learn rather than just give conclusive judgments (Danielson, 2013; Wiggins, 1998). In this stage, teachers help students self-assess their answers, give feedback to improve subsequent performances, correlate the learning to actual and/or professional situations, and foster self-direction and the habits of continuous and lifelong learning.

Table 1 contains the dimensions of the Pentagram Observation Card for assessing the classroom practice of teachers of gifted students. Each dimension corresponds to one of the stages of the Pentagram Strategy and is accompanied by the key behavioral performance indicators for that dimension. This descriptive framework aims to clarify the conceptual foundations of the dimensions and to identify observable and measurable classroom practices that support the effective use of the card in evaluating instructional performance, diagnosing strengths and areas for improvement, and guiding professional development programs for teachers of gifted learners.

Given the theoretical framework, the Pentagram Strategy seemed to be sufficiently backed by theory and practice, particularly constructivist learning theory and current, active learning models of gifted education, which stress meaning-making and the acquisition of advanced cognitive processes. The framework illustrates the Pentagram Strategy as a cognitive tool that guides learning and problem exploration through structured, progressive phases, from problem definition to evaluation and application, closely mirroring the cognitive and affective profiles of gifted learners. Furthermore, the framework illustrated that using this strategy to teach gifted students demands a form of practice that is instructional, problem-focused, complex, and open-ended; nurtures scientific dialogue; fosters analysis, interpretation, and prediction; supports authentic assessment; encourages constructive feedback; and maintains balance. Therefore, the framework addressed the first research question by defining the theories and models that inform the Pentagram Strategy and outlining the expectations it creates for practice in gifted education.

The theoretical framework also addressed the second research question by identifying the dimensions of classroom performance for teachers of gifted students, along with the behavioral indicators associated with each dimension in accordance with the Pentagram Strategy. Five core dimensions of classroom performance were derived: advanced problem and concept identification (What?), intellectual engagement and meaning-making (Why?), flexible implementation and interpretation (How?), critical and future-oriented thinking (What if?), and authentic evaluation and self-reflection. These dimensions were operationalized through observable behavioral indicators reflecting the teachers' role in managing the instructional environment and activating high-level thinking among gifted learners, thereby providing a scientific and methodological foundation for designing and standardizing the classroom observation card.

© 2026 Conscientia Beam. All Rights Reserved.

Table 1. Description of the pentagram observation dimensions and classroom performance indicators for teachers of gifted students.

Dimension	Dimension Description	Key Performance Indicators
Stage 1 (What?): Identifying the Problem and Core Concepts	This dimension focuses on the teacher's ability to present an advanced instructional problem, define its essential concepts, and link it to multiple knowledge domains in ways that align with the characteristics of gifted learners.	• Formulating high-level learning objectives that reflect analytical and creative thinking. • Presenting a complex, multidimensional instructional problem. • Connecting the problem to other disciplinary areas. • Situating the problem within a real-world, research-based, or future-oriented context. • Using advanced conceptual maps and knowledge models.
Stage 2 (Why?): Intellectual Engagement and Meaning-Making	This dimension emphasizes fostering intellectual motivation among gifted learners and developing a deep understanding of the problem through analysis, comparison, and reasoned justification.	• Guiding students to select appropriate analytical strategies. • Deconstructing the problem into its main and subsidiary components. • Comparing multiple solution alternatives. • Justifying analytical choices using logical or scientific evidence. • Supporting scientific dialogue and exchange of perspectives.
Stage 3 (How?): Flexible Implementation and Building Interpretation	This dimension concerns enabling students to interpret phenomena and outcomes scientifically, constructing coherent explanatory models and causal relationships within a rich and flexible learning environment.	• Interpreting results in a logical and scientific manner. • Linking causes to effects. • Justifying opinions using clear evidence. • Revisiting assumptions in light of new information. • Recognizing the existence of multiple possible interpretations.
Stage 4 (What if? / What are the Results?): Critical and Future Thinking	This dimension focuses on promoting critical and future-oriented thinking by encouraging students to anticipate consequences, assess alternatives, and make responsible decisions.	• Presenting hypothetical scenarios for prediction. • Evaluating risks and opportunities associated with each alternative. • Using future-oriented scenarios. • Weighing alternatives against defined criteria. • Enhancing cognitive flexibility and acceptance of multiple possibilities.
Stage 5 (Evaluation and Feedback): Authentic Assessment and Self-Reflection	This dimension targets the authentic assessment of learning outcomes, fostering students' self-reflection on thinking processes, and supporting continuous learning and improvement based on feedback.	• Evaluating solutions using clear criteria. • Encouraging self-reflection on thinking and performance. • Using feedback to improve future learning. • Transferring learning to new contexts. • Supporting independent planning for subsequent learning.

2. METHODS

2.1. Research Design

This study focused on the observation card for the classroom performance of teachers of gifted students in light of the pentagram strategy through a descriptive quantitative survey study approach. This method enables the description of the phenomenon to a greater extent and increases the extent to which the findings may be generalized because standardized data can be collected and analyzed from a large population.

A comprehensive instrument, an observation card designed to assess the classroom performance of teachers of gifted students in light of the Pentagram strategy, was developed for this study. The dimensions of the card were aligned with the stages of the Pentagram model, yielding five core domains: advanced problem and conceptual definition (What?); intellectual motivation and meaning-making (Why?); flexible implementation and explanation-building (How?); critical and future-oriented thinking (What if? What are the consequences?); and authentic evaluation and self-reflection (Evaluation).

© 2026 Conscientia Beam. All Rights Reserved.

The research procedures commenced after obtaining ethical approval from the King Faisal University Institutional Review Board. The instrument was initially administered in a pilot study to examine its clarity, reliability, and validity. It was subsequently applied to the main study sample. The resulting data were analyzed using appropriate descriptive and inferential statistical methods to derive the study's findings.

Developing an appropriate research design is a central component of scientific inquiry, as it enables the systematic selection of suitable methods, data-collection procedures, and analytical techniques. Figure 1 presents an overview of the study's design and the sequential procedures followed.

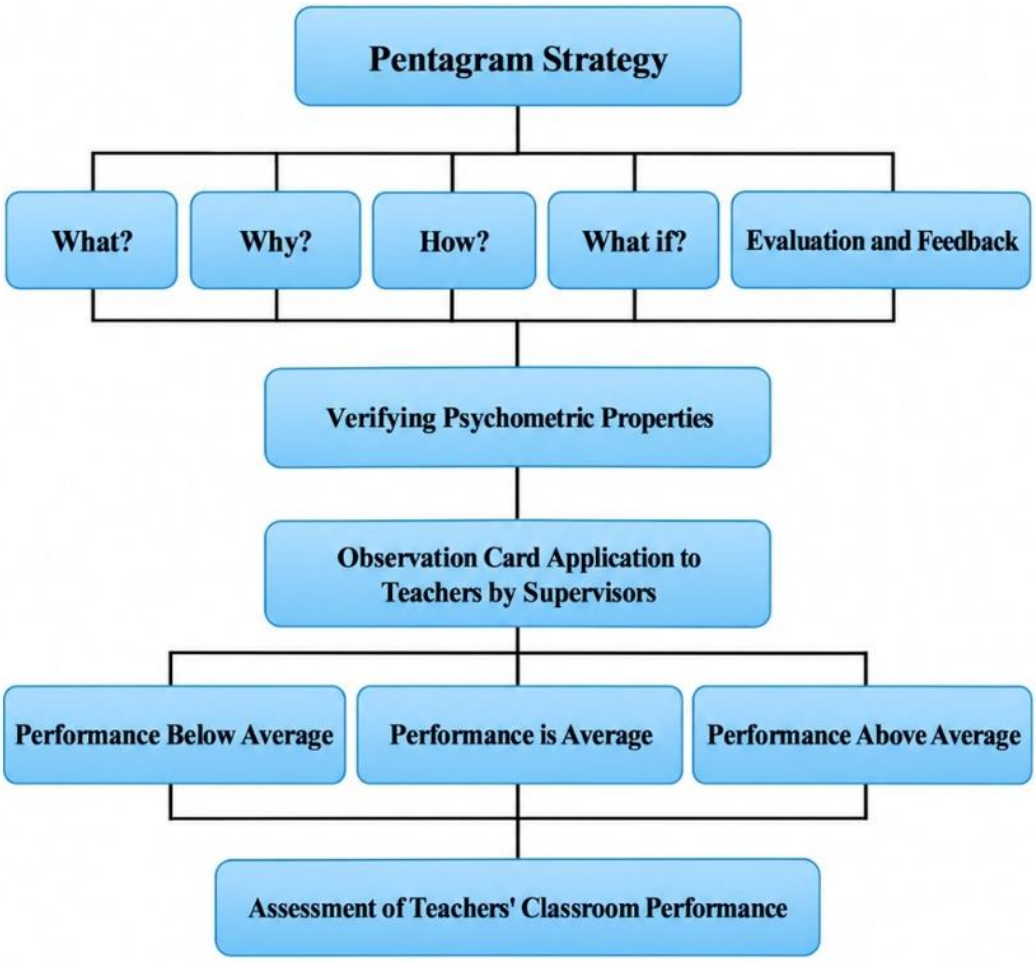

Figure 1. Tool construction procedures.

2.2. Population and Sampling

The study population comprised all teachers of gifted students within government schools under the jurisdiction of the General Department of Education in the Eastern Province, Saudi Arabia, during the academic year 2025/2026. The sample included 128 teachers purposively selected from 32 schools across the districts of Dammam, Dhahran, Khobar, and Qatif. These schools were specifically chosen because they host dedicated programs and classes for gifted students.

To ensure high-quality data, the sample was stratified by school level (14 elementary, 10 middle, and 8 high schools). Furthermore, 16 educational supervisors were recruited to conduct the observations. These supervisors were selected based on rigorous criteria, including holding a postgraduate degree in Gifted Education and a minimum of 10 years of professional field experience, ensuring the validity of the evaluation of teaching practices using the designated observation card.

© 2026 Conscientia Beam. All Rights Reserved.

2.3. Steps for Developing a Classroom Observation Card for Teachers of Gifted Students Based on the Pentagram Strategy

A series of systematic methodological steps was followed to develop a classroom observation card for teachers of gifted students based on the Pentagram Strategy. These steps contributed to constructing the fundamentals of the study and the accurate representation of the design aligned with the targeted instructional practices. The development process relied on theoretical literature on gifted education, constructivist learning, and the Pentagram Strategy and attempted to extrapolate the requirements of the strategy into observable and measurable behavioral events in the classroom. It also aimed at contemporary standards in measurement and educational evaluation to design a tool with high validity and reliability that could be used in various educational contexts, measurement and evaluation of gifted education teachers, and professional development of teachers of gifted students. The development process followed these steps.

The first step consisted of a review of educational and psychological literature about the Pentagram Strategy, gifted education, and performance-based assessment in order to develop a holistic theoretical and conceptual framework. This framework clarified the philosophy of the strategy, the 5 stages (What? Why? How? What if? Evaluation), and the theory of gifted children, pedagogical strategies to be used with them, and conceptualization of the theory of classroom performance as an observable and measurable behavioral event (Hattie, 2012; Tomlinson, 2017; VanTassel-Baska & Stambaugh, 2006). This framework informed the description of the dimensions and indicators of the card to be aligned with higher-order thinking, deep learning, and authentic assessments.

The card's dimensions were established based on the stages of the Pentagram Strategy and the theoretical framework. Advanced problem and concept identification (What?), intellectual engagement and meaning-making (Why?), flexible implementation and interpretation (How?), critical and future-oriented thinking (What if? Expected results), and authentic evaluation and self-reflection (Evaluation) are the five components. Marzano (2007) and Paul and Elder (2014) analyze the formulation of these dimensions. These dimensions point to the factors of the comprehensiveness, integration, and alignment of each dimension to a cohesive set of classroom practices that characterize the teacher and the facilitator, enriched learning environments.

Performance indicators for each dimension were developed based on analyses of effective teaching practices, a review of existing observation tools, and adherence to clear behavioral formulations using observable action verbs, incorporating cognitive progression and depth in performance (Moskal & Leydens, 2000; Nitko & Brookhart, 2014). A five-point rating scale (Not evident, Little, Moderate, High, Very high) was adopted to enhance accuracy in scoring and enable quantitative analysis of classroom performance (Brookhart, 2013; Stevens & Levi, 2013).

2.4. Procedures for Administering the Observation Card

To provide the authors and participating supervisors with the tools needed for uniform and proper use of the observation card, a workshop was held. Step-by-step instructions on the use of the tool, as well as strategies for overcoming obstacles in its use, were provided to the participants. Then, both supervisors and authors did pilot implementations to confirm the clarity of procedures and to ensure proper management of the observation card application.

Once the observation card was designed and its validity and reliability were established, for a month, ten supervisors applied the observation card to assess teachers' educational practices in the gifted education classrooms. During this time, each supervisor made observations of five teachers and, on average, did ten observations of each teacher. For each observed teaching practice, the supervisor assigned a score ranging from 1 to 5. Because teachers may employ specific practices selectively depending on the nature and instructional requirements of each lesson, it was neither expected nor appropriate for all practices to appear in a single class session.

Consequently, the final score for each teacher was calculated based on the average score of the practices actually observed and implemented, rather than on the average across all ten observations. For instance, if a teacher

© 2026 Conscientia Beam. All Rights Reserved.

demonstrated a particular practice three times during the semester, the average score for that practice was computed as the sum of the three scores divided by three. An illustrative example of this calculation is presented in Table 2.

Table 2. Observation score card.

	Observation Score												
Practice	**Obs.1**	**Obs.2**	**Obs.3**	**Obs.4**	**Obs.5**	**Obs.6**	**Obs.7**	**Obs.8**	**Obs.9**	**Obs.10**	**Mean**	**St. Dev.**	**Practice level**
Practice1	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	-	-	-
Practice 2													
Practice 3													
Practice 4 . . . . Practice 50													

2.5. Verifying the Psychometric Properties of the Observation Card Using the Rasch Measurement Model

The validation of the scale proceeded through several steps. In the first, one of the content-validation exercises involved a panel of thirteen experts in curriculum and instruction, gifted education, and educational measurement from Saudi universities. Following their suggestions, some ambiguous items were revised while a few grammatical errors were corrected. The card was then evaluated by means of a pilot study with 50 teachers of gifted students, and this resulted in the gathering of more responses. The responses generated in this pilot study and the qualitative feedback that accompanied them were analyzed and used for the scale's revision. The evaluation of construct validity was done primarily with the Rasch model. This psychometric model provides a unique, objective, linear approach, which is why it is used for this kind of validation in psychological and educational contexts (AlAli & Al-Barakat, 2024; Hair, Black, Babin, & Anderson, 2019). Interpreting results and analyses in Rasch measurement relies mostly on:

2.5.1. Model-Data Fit

2.5.1.1. Item Fit Statistics

Interpreting results and analyses in Rasch measurement relies mostly on the evaluation of the model-data fit stemming from infit and outfit statistics. An important explanation is that infit refers to the weighted fit of an item or person and, as such, represents the information-weighted mean-square residual, which is sensitive to unexpected responses that are located near the measure's ability or difficulty level.

- MNSQ (Mean Square Residual Statistics)

An interpretation from the practical side is that mean square (MNSQ) values estimate the observed-to-the-model-predicted variance ratio. An infit MNSQ value within the range of 0.4 to 1.5. The validity of the scale was measured using MNSQ values for the infit, and the results showed that the scale had an appropriate degree of validity. Scale validity scores according to MNSQ values fall within the safe limits (see Appendix A).

- Zstd (Standardized Fit Statistics)

A Zstd value within the range of -2.0 and +2.0 is a common guideline and acceptance. It has a suitable standardized fit statistic (Zstd) value, which should be between -2 and 2, as shown in Appendix A.

2.5.1.2. Person Fit Statistics

- Outfit MNSQ and Infit MNSQ

An infit MNSQ value within the range of 0.4 to 1.5. Scale validity scores according to MNSQ values fall within the safe limits (see Appendix A).

© 2026 Conscientia Beam. All Rights Reserved.

2.5.2. Point Measure Correlation (PTMEA-CORR)

Assessing item polarity through the point-measure correlation (PTMEA-CORR) is crucial for determining whether scale items effectively capture the intended latent construct(s). Depending on the context, positive correlations may suggest that an item has reasonably good functionality, whereby higher/positive responses relate to higher levels of the latent trait of the respondents. In contrast, negative point-measure correlations indicate cause for concern because in such situations, the item functions to the detriment of the overall construct cohesiveness of the scale. Such items will operationally need some form of action, such as modification or exclusion, to maintain the instrument's unidimensionality. According to AlAli and Al-Barakat (2022), point-measure correlations between 0.2 and 1.0 can be considered a signal of sufficiently productive measurement. In the current case, all items correlated within this boundary (see Appendix A).

2.5.3. Item Dimensionality Analysis

Two principal criteria were employed in the Rasch analysis with respect to evaluating variance. These are: the raw variance explained by measures should be more than 40%, and the first contrast of unexplained variance should be no more than 15%. The raw variance explained by measures that is reflected in Appendix B is 53.2, while the first contrast of unexplained variance is 5.3. These figures meet the criteria. Therefore, the data obtained from the observation card verify the presence of unidimensionality; as such, they are consistent with the Rasch model (see Appendix B).

2.5.4. Calibration and Rating Scale Functioning

The Rasch analysis shows how calibration verifies the functioning of the polytomous rating scale. With more confidence and using an interval-level approach, it would be able to tell whether the categories employed by respondents were consistent and logical on a scale of 1 to 5. Data from the observation card were utilized to analyze rating scale calibration and functionality, and the results are captured in Appendix C. Rating categories and response patterns were analyzed, and confirmed the layering of the tiers. As stated in Appendix C, the response frequency that was recorded through the rating categories was tiered, and the distribution reflected the response levels calibrated to the probabilities. Tier 3 was, as was most probable, rated 3 by 9 respondents (30%). Rating tier 4 was nearly as probable, recorded 9 selections (29%), and tier 5 was selected 6 times (19%) to be selected. Tier 2 recorded only 5 selections (16%), and tier 1 was probably the lowest with only 2 respondents (6%). This distribution of responses indicated the participants could progressively demonstrate and differentiate between the levels of the scale. Each step in tiered progression also constitutes a reasonable calibration that each tier contained a proprietary distinct meaning or characteristic for the respondents. This results in the rating being operational.

2.5.5. Person and Item Reliability Using the Rasch Measurement Model

When evaluating the reliability of a measure using the Rasch model, both person reliability and item reliability require consideration. Measures with reliability parameters of over 50% are considered acceptable, while person and item separation parameters of over 2 are considered acceptable (AlAli & Al-Barakat, 2022; AlAli & Al-Barakat, 2024). For this study, the scale's reliability was assessed from the perspective of both item and person reliability, with the scale, at a minimum, appearing to possess a sufficient level of reliability with respect to its items (see Appendix D).

2.6. Data Analysis

In the process of answering the research question, while assessing the Psychometric Property of Construct Validity and completing a Psychometric Evaluation, Analysis of Reliability using McDonald's Omega and Composite Reliability, and Convergent and discriminant validity were conducted.

© 2026 Conscientia Beam. All Rights Reserved.

To provide insight into measurement, the Rasch Model was used with Winsteps. This was chosen as it allows for the simultaneous estimation of the person's ability and the items' parameters of difficulty and discrimination. In addition, person and item reliability coefficients are provided to assess the reliability of the estimates (AlAli & Al-Barakat, 2022).

Factor validity was investigated through a sequential approach employing both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), conducted using SPSS and Amos, respectively. CFA, a specialized form of Structural Equation Modeling (SEM), was utilized to test the hypothesized factor structure derived from the EFA and theory. This method is a powerful tool for verifying the latent construct relationships, validating the instrument's dimensional structure, and confirming factor loadings, thereby playing a critical role in evaluating construct validity (AlAli & Al-Barakat, 2024).

3. RESULTS

To answer the fourth question, which is: what are the psychometric properties (validity and reliability) of the classroom observation card for teachers of gifted students based on the Pentagram Strategy?

There are steps to developing an Observation Card. The construct validity of the Observation Card was verified using different methods.

3.1. Person and Item Reliability Using the Rasch Measurement Model

To assess both person reliability (the consistency of supervisors' ratings using the observation card) and item reliability, the Rasch measurement model was employed using Winsteps software, version 3.68.2. The Rasch model was selected because it provides estimates for both person ability and item difficulty, along with discrimination indicators.

Person reliability was calculated to evaluate the stability and consistency of supervisors' evaluations obtained through the observation card. Item reliability was also assessed to determine the internal consistency and stability of the items comprising the card. According to established criteria, reliability values must exceed 0.50 to be considered acceptable, and a person separation index greater than 2.0 indicates adequate discrimination (AlAli & Al-Barakat, 2022).

The values for person reliability in Table 3 are also noticeable, standing at an impressive 0.90 and 0.92. Correspondingly, the person and item separation indices also surpass 2.0, at 3.17 and 3.58. This indicates that the card for observations is reliable according to the Rasch model and that the card is reliable, stable, and psychometrically sound.

Table 3. Person and item separation and reliability for the observation card.

	Score	Count	Measure	Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	163.4	11.2	1.71	0.24	1.09	-0.7	1.08	-0.6
S.D.	37.2	0.0	2.19	0.16	0.70	3.6	0.81	1.8
Real RMSE	0.61							
Adj. SD.	2.36							
Separation	3.17							
Person reliability	0.90							
Mean	189.2	46.0	0.00	0.06	1.00	-0.2	1.07	0.1
S.D	32.3	0.0	0.32	0.0	0.19	1.7	0.42	1.8
Real RMSE	0.11							
Adj. SD.	0.29							
Separation	3.58							
Item reliability	0.92							

© 2026 Conscientia Beam. All Rights Reserved.

3.2. Indicators and Coefficients of Construct Validity

McDonald's omega and composite reliability (CR) are industry benchmarks for measuring the reliability of observational tools. In Table 4, McDonald's omega ranges between 0.916 and 0.953, and CR between 0.914 and 0.953. These metrics exceed the industry standard of 0.70, which indicates that the observational card is consistent.

Additionally, the average variance extracted (AVE) values, between 0.586 and 0.715, exceed the threshold of 0.50, confirming that convergent validity is present. For discriminant validity, the final column of Table 4 demonstrates that each square root of the AVE value is greater than the corresponding factor loadings, satisfying the condition needed. Taken together, these findings affirm that the observation card is both reliable and valid in terms of construct validity.

Table 4. Indicators and coefficients of construct validity.

Constructs	Items	McDonald's ω	CR	AVE	$\sqrt{AVE}$
(What?): Identifying the problem and core concepts	10	0.923	0.920	0.602	0.776
(Why?): Intellectual engagement and meaning-making	10	0.922	0.918	0.695	0.834
(How?): Flexible implementation and building interpretation	10	0.916	0.914	0.594	0.771
(What if? What are the results?): Critical and future thinking	10	0.936	0.934	0.715	0.846
(Evaluation and feedback): Authentic assessment and self-reflection	10	0.956	0.953	0.586	0.766

3.3. Construct Validity Based on Factor Analysis

Construct validity of the observation card was examined through both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), using SPSS and AMOS statistical software. For the EFA, principal components analysis was conducted on participants' responses to the 50 items included in the final version of the observation card. Oblique rotation using the Promax method was applied to the extracted factors, with retention based on the criterion of eigenvalues greater than 1. Table 5 presents the results of the EFA using the Promax rotation method.

Table 5. Results of exploratory factor analysis using the Promax method.

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
(What?): Identifying the problem and core concepts	4.276	85.528	85.528
(Why?): Intellectual engagement and meaning-making	0.295	5.893	91.421
(How?): Flexible implementation and building interpretation	0.212	4.248	95.669
(What if? What are the results?): Critical and future thinking	0.145	2.908	98.576
(Evaluation and feedback): Authentic assessment and self-reflection	0.071	1.424	100.000

Table 5 presents the results of the exploratory factor analysis using the Promax rotation method. The analysis identified five components, all of which have eigenvalues greater than 1. The first component, Identifying the Problem and Core Concepts, shows the strongest contribution with an eigenvalue of 4.276, accounting for 85.528% of the explained variance, substantially higher than the remaining components. The second component, Intellectual Engagement and Meaning-Making, has an eigenvalue of .295 and explains 5.893% of the variance, followed by Flexible Implementation and Building Interpretation with an eigenvalue of .212 (4.248%). Critical and Future Thinking and Authentic Assessment and Self-Reflection constitute the fourth and fifth components, respectively, and both stand above the eigenvalue barrier with figures of .145 (2.908%) and .071 (1.424%) each. Such large discrepancies between the first component's eigenvalue (4.276) and the others attest that the first factor is far more dominant. The first component's high variance also suggests that it is fundamental in the overarching construct. Despite the

© 2026 Conscientia Beam. All Rights Reserved.

emergence of several factors with eigenvalues greater than one, the large discrepancy between the first and the other eigenvalues indicates that a dominant dimension is present, accounting for the variance in the data.

In order to test the factorial construct validity of the tool, the observation card's final version was applied to the study sample; then a CFA was performed regarding the items within the CFA-determined dimensions. The model was built to assess 50 items across the five dimensions. It was important for the model to demonstrate that each of the 50 items fulfilled the 0.40 threshold of positive loading suggested by some researchers (AlAli & Al-Barakat, 2022). For instance, in the 2, all of the items exhibited positive loadings of more than 0.40 within the dimensions to which each of the items was assigned. Thus, it affirmed that every item had a significant contribution to the construct of each dimension. Hence, the observation card was confirmed to have a substantial five-dimensional construct, with each dimension containing 10 items.

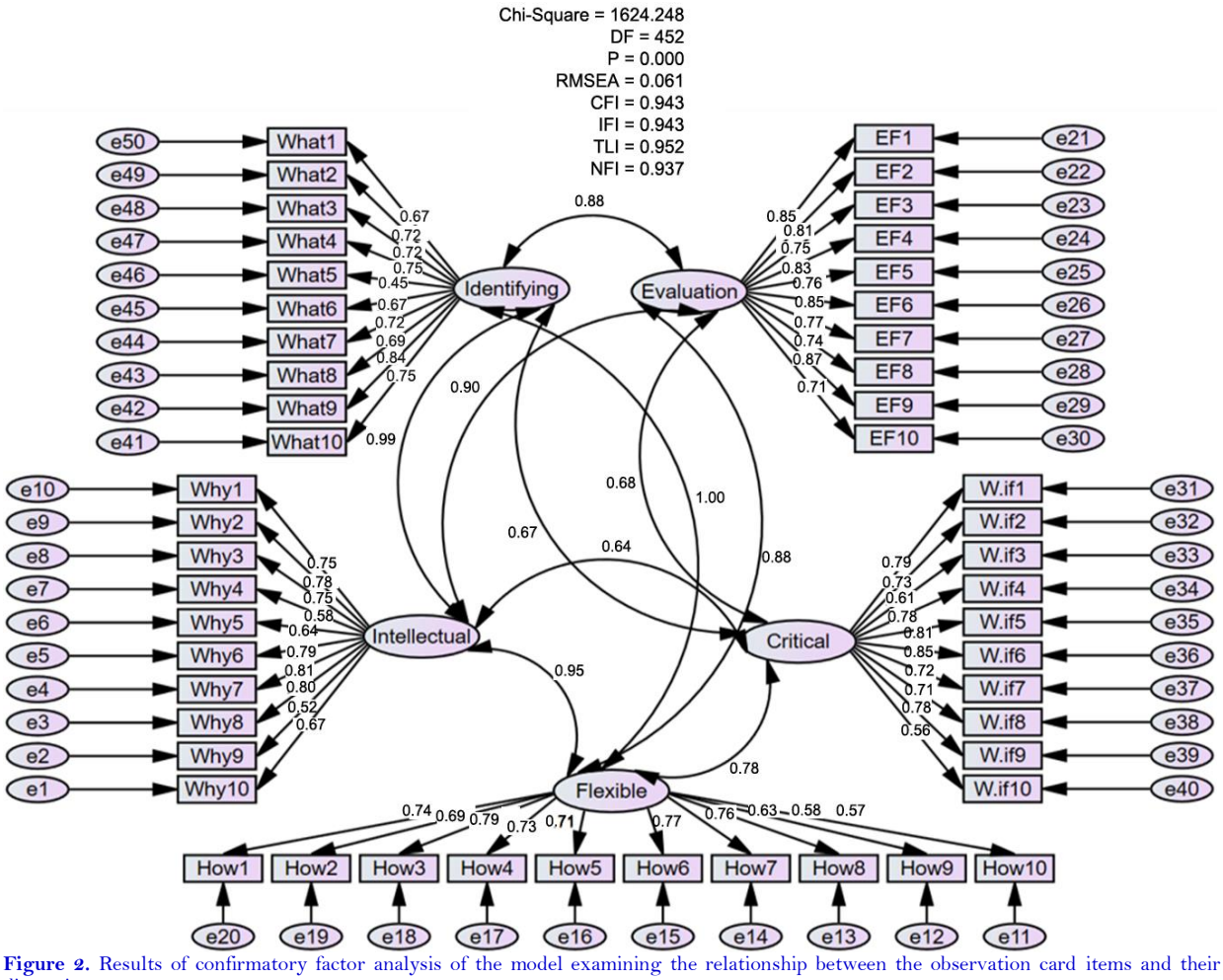

Figure 2. Results of confirmatory factor analysis of the model examining the relationship between the observation card items and their dimensions.

In Figure 2, results from the confirmatory factor analysis (CFA) are shown, where all items have shown high loading values on the dimensions to which they are assigned. These values indicate a good association between the items and the specific construct that they are designed to measure. The figure also shows that the five dimensions of the observation card have a high correlation with each other. This indicates that there are five dimensions meaningful of the same latent variable.

In Table 6, the results also show compliance with construct validity criteria as outlined by AlAli and Al-Barakat (2022). This also shows that the estimated model is adequate. Yes, the chi-square statistics are high, and this is expected with large sample sizes; however, the other fit indices do show the model reasonably describes the data. An RMSE value of 0.061 is good enough as it is below 0.08 and indicates an acceptable fit. The fit indices show that the

© 2026 Conscientia Beam. All Rights Reserved.

model exceeds the 0.90 threshold for adequate fit, with the CFI at 0.943, TLI at 0.952, and NFI at 0.937. For the parsimonious fit index, the value is also acceptable with a chi-square value of 3.59, which is < 5 making it also reasonable to accept the model.

These indicators affirm the stability and validity of the models associating the observation card items with the respective dimensions.

Table 6. Results of confirmatory factor analysis of the model examining the relationship between the observation card items and their dimensions.

Category	Indicators of internal construct validity	Level of acceptance	Indexes in the proposed model
Absolute fit	χ^2	$p > 0.05$	Significant
	RMSE	RMSE < 0.08	0.061
	CFI	CFI > 0.90	0.943
Incremental fit	TLI	TLI > 0.90	0.952
	NFI	NFI > 0.90	0.937
Parsimonious fit	χ^2/df	χ^2/df < 5.0	χ^2/df = 3.59 < 5.0

4. DISCUSSION

The current study's primary goal is to analyze the psychometric evidence of the reliability of the classroom observation card assessing the teachers of gifted learners using the Pentagram Strategy. The results converged to show that the instrument exhibits high reliability and strong construct validity, justifying its use for research and practice.

Applying the Rasch measurement model, the evidence of high reliability for both persons and items indicates that there is consistency and stability in the supervisors' evaluations across the various indicators of the observation. A reliability level of 0.90 and above for persons indicates that the card is effective in assessing varying levels of teaching performance. Similarly, the person and item separation indices that are above the recommended 2.0 threshold demonstrate that there is adequate measurement. The result is consistent with prior research that notes the significance of using Rasch in developing observation tools in increasing the level of objectivity and minimizing evaluator bias in assessing performance (AlAli & Al-Barakat, 2022). The sufficient Rasch indices also suggest that the items are sufficiently targeted to the latent construct they are measuring and that the indicators function uniformly across varying performance levels of the teachers.

The results from the analyses of construct reliability and internal consistency strengthen the psychometric validation of the observation card. For the five dimensions in the study, McDonald's omega and composite reliability showed values greater than the threshold of 0.70, indicating strong reliability and consistency for each dimension. Also, AVE values for all dimensions exceeded the threshold of 0.50, thus validating the convergent construct of each dimension. The discriminant validity was confirmed since the correlation between the dimensions was less than the square root of the AVE. These results indicate that although the dimensions of the Pentagram Strategy are interconnected, each dimension provides a distinct, definable, and measurable dimension of effective teaching practices for educating the gifted.

The further support of construct validity is backed by both exploratory and confirmatory factor analyses. Exploring factor analyses showed a five-factor structure, which is similar to the theoretical components of the Pentagram Strategy, where all factors exceeded the eigenvalue criterion. With respect to the results, the first factor, which identifies the problem and the core concepts, accounted for the highest amount of explained variance, making the first factor central to the classroom practice structure for gifted education. This was especially prominent due to the framing of the problem, along with the concept formation, in the development of higher-order thinking, which is a principle that a lot of gifted education literature highlights.

The confirmatory factor analysis reasonably justified the proposed model. All the items and their respective dimensions passed the minimum acceptable loading factor criterion. The model explains the factor heavily relied upon

© 2026 Conscientia Beam. All Rights Reserved.

and shows good fit indices (CFI, TLI, NFI, RMSEA, and χ^2/df), which means the proposed five-factor framework is capable of representing the five observed factors. Although the chi-square test was significant at $p<0.05$, a statistically significant chi-square is common for mid- to large-sized samples; the model fit indices and the incremental (CFI and TLI) and parsimonious (NFI) fit indices, along with the model stability and appropriateness to the proposed framework by the authors (AlAli & Al-Barakat, 2024), are justified.

Collectively, the results validate that the observation card captures a unified measure of what constitutes effective teaching of gifted students and of a constellation of interlinked but analytically separable instructional components aligned with the Pentagram Strategy. The results of the Rasch analysis, reliability coefficients, and factor-analytic techniques all point to a greater degree of psychometric confidence regarding the observation card.

4.1. Practical and Research Implications

From a practical standpoint, the validated observation card aids educators, supervisors, and program evaluators in gifted education. The card's psychometric discriminative capacity makes it appropriate for constructive feedback and planning for the professional development continuum for instructional supervision and coaching. Moreover, the card dimensions and the Pentagram Strategy aligned provide a framework for teachers on instruction, problem identification, cognitive complexity, and the flexible use of the strategies, critical and futures thinking, and authentic assessment.

The instrument contributes to the field of gifted education and, more specifically, to the observation tool that is psychometrically validated. The instrument's robust psychometric evidence supports empirically driven investigations about education quality, educational interventions, and the professional development of teachers. The use of classical test theory and the Rasch model for validation of the instrument is a novel approach and enhances the existing literature in educational research on observation and assessment tools.

Lastly, the card may also be useful for cross-cultural validation studies and longitudinal studies on the evolution of specific teaching practices. With the construction of a measurement tool that is both theoretically sound and empirically valuable, this study aims to improve teaching quality and accountability in gifted education programs, thus promoting adaptive and constructive learning environments for gifted students.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU261344].
Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of King Faisal University, protocol code KFUREC-2025-Oct, ETHICS3787, and the date of approval was 19 October 2025.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.
Disclosure of AI Use: This study was originally prepared by the authors. An AI tool (Grammarly) was utilized to enhance readability and language. After using this tool, the authors carefully reviewed and edited the content as necessary, and they take complete responsibility for the content, interpretation, and scientific integrity of the research.

REFERENCES

Al-Asadi, M. H. R., & Al-Fatlawi, K. R. H. (2021). The effect of the Pentagram strategy in deep comprehension of literature among fifth-grade students. *Journal of Research in Teacher Development, 5*(3), 77–95.

Al-Harthi, K. A. (2020). Teaching strategies for gifted learners in inclusive classrooms. *Journal of Gifted Education Research, 14*(2), 55–72.

© 2026 Conscientia Beam. All Rights Reserved.

Al-Momani, H., & Al-Oweidi, A. (2020). The psychometric characteristics of the Renzulli scale of behavioral characteristics (3rd edition) in the detection of gifted students in the age group (12-18) in Jordan. *Journal for the Education of Gifted Young Scientists*, *8*(1), 105-132. https://doi.org/10.17478/jegys.656673

AlAli, R., & Al-Barakat, A. (2022). Using structural equation modeling to assess a model for measuring creative teaching perceptions and practices in higher education. *Education Sciences*, *12*(10), 690. https://doi.org/10.3390/educsci12100690

AlAli, R. M., & Al-Barakat, A. A. (2024). Constructing and developing a scale for assessing language teachers' performance in integrating reflective thinking skills within primary reading learning environments. *Forum for Linguistic Studies*, *6*(6), 194-210. https://doi.org/10.30564/fls.v6i6.7744

Amoush, S., & Khodair, R. (2025). Effectiveness of the pentagram strategy (PG) in enhancing critical reading skills among tenth-grade female students in Jordan. *Educational Process: International Journal*, *16*, e2025266. https://doi.org/10.22521/edupij.2025.16.266

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.

Beghetto, R. A., & Kaufman, J. C. (2014). Classroom contexts for creativity. *High Ability Studies*, *25*(1), 53-69. https://doi.org/10.1080/13598139.2014.905247

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom*. ASHE-ERIC Higher Education Report No. 1. Washington, DC: George Washington University, School of Education and Human Development.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria, VA: ASCD.

Capraro, M. M., Capraro, R. M., Scheurich, J. J., Jones, M., & Morgan, J. (2012). A teacher observation instrument for PBL classroom instruction. *Journal of STEM Education*, *13*(2), 39–43.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Clark, B. (2013). *Growing up gifted* (8th ed.). Boston, MA: Pearson.

Crocker, L. M., & Algina, J. (2008). *Introduction to classical and modern test theory*. Mason, OH: Cengage Learning.

Danielson, C. (2013). *The framework for teaching: Evaluation instrument*. Princeton, NJ: The Danielson Group.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Los Angeles, CA: Sage.

El-Sayed, A. H. (2021). The effectiveness of pentagram-based instruction in developing higher-order thinking skills. *International Journal of Education and Learning Systems*, *6*, 45–58.

El-Sayed, M. (2021). Thinking maps and higher-order thinking skills. *Arab Journal of Educational Research*, *41*(1), 87–112.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). New York, NY: McGraw-Hill Education.

Gronlund, N. E., & Brookhart, S. M. (2009). *Grading and assessment in the student-centered classroom*. Boston, MA: Pearson.

Guilford, J. P. (1967). *The nature of human intelligence*. New York, NY: McGraw-Hill.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Boston, MA: Cengage Learning.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London, England: Routledge.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*(3), 235-266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Jonassen, D. H. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 215–239). Mahwah, NJ: Lawrence Erlbaum Associates.

Kanaan, J. A. Z., Muhammad, F. H., & Hussein, Q. M. (2024). The effect of using the pentagram strategy in learning some grips in freestyle wrestling for students. *Journal of Studies and Researches of Sport Education*, *34*(3), 361–372.

Maker, C. J., Jo, S., & Muammar, O. M. (2008). Development of creativity: The influence of varying levels of implementation of the DISCOVER curriculum model, a non-traditional pedagogical approach. *Learning and Individual Differences*, *18*(4), 402–417. https://doi.org/10.1016/j.lindif.2008.03.003

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: ASCD.

© 2026 Conscientia Beam. All Rights Reserved.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research, and Evaluation, 7*(10), 1–6.

Nakano, T. C., & Peixoto, E. M. (2023). Item analysis of the teacher's gifted rating screening scale. *Estudos de Psicologia (Campinas), 40*, e200056. https://doi.org/10.1590/1982-0275202340e200056

Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students* (7th ed.). Boston, MA: Pearson Education.

Paul, R., & Elder, L. (2014). *Critical thinking: Tools for taking charge of your learning and your life* (3rd ed.). Boston, MA: Pearson Education.

Piaget, J. (1972). *The principles of genetic epistemology (W. Mays, Trans.)*. London, England: Routledge & Kegan Paul.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223-231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Renzulli, J. S. (2016a). *The enrichment triad model: A guide for developing defensible programs for the gifted and talented*. Mansfield Center, CT: Creative Learning Press.

Renzulli, J. S. (2016b). The three-ring conception of giftedness. *Gifted Child Quarterly, 60*(4), 233–245.

Salama, A. M. (2018). Pentagram thinking framework: A pedagogical approach for enhancing critical and systemic thinking. *International Journal of Education and Learning Systems, 3*, 1–10.

Shahata, H., & Al-Najjar, Z. (2017). *Active learning: Concept and applications*. Cairo: Egyptian-Lebanese House.

Stevens, D. D., & Levi, A. J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning* (2nd ed.). Sterling, VA: Stylus Publishing.

Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). Alexandria, VA: ASCD.

Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms* (3rd ed.). Alexandria, VA: ASCD.

VanTassel-Baska, J. (2018a). Curriculum issues: Theory and practice. *Gifted Child Quarterly, 62*(3), 295–309.

VanTassel-Baska, J. (2018b). *Curriculum planning and instructional design for gifted learners*. Waco, TX: Prufrock Press.

VanTassel-Baska, J., & Brown, E. F. (2007). Toward best practice: An analysis of the efficacy of curriculum models in gifted education. *Gifted Child Quarterly, 51*(4), 342-358. https://doi.org/10.1177/0016986207306323

VanTassel-Baska, J., & Stambaugh, T. (2006). *Comprehensive curriculum for gifted learners* (3rd ed.). Boston, MA: Allyn & Bacon.

VanTassel-Baska, J., & Stambaugh, T. (2021). *Teaching gifted learners*. London: Routledge.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.

World Bank. (2010). *The stallings classroom observation system: Measuring teaching practices in developing countries*. Washington, DC: World Bank.

Appendices

Appendix A. Item Fit Analysis for observation card.

Items	Measure	Infit		Outfit		Pt-measure	
		MNSQ	ZSTD	MNSQ	ZSTD	CORR	EXP
What 5	0.34	1.72	1.9	1.89	1.7	0.57	0.55
Why 1	0.07	1.77	1.8	1.83	1.6	0.53	0.69
What 6	0.14	1.46	1.7	1.81	1.8	0.56	0.64
What if 1	0.38	1.12	0.7	1.74	1.2	0.64	0.69
Why 8	0.28	1.37	1.6	0.98	-0.7	0.63	0.70
Why 9	0.19	1.51	1.7	1.78	1.7	0.61	0.64
What 8	0.78	1.48	-1.3	1.56	1.8	0.52	0.61
EV 9	0.16	1.39	-0.6	0.87	-1.6	0.56	0.63
How 2	0.54	1.14	-0.5	0.98	-0.9	0.62	0.69
EV5	0.42	1.21	0.4	0.89	-0.8	0.53	0.60
EV 4	0.59	1.20	-1.3	1.47	-0.8	0.64	0.67
What if 6	0.36	0.97	-0.9	1.02	-0.9	0.65	0.66
What 9	0.79	0.70	0.6	1.11	0.5	0.72	0.69
How 8	0.17	1.12	0.8	0.86	-1.5	0.65	0.67

© 2026 Conscientia Beam. All Rights Reserved.

Items	Measure	Infit		Outfit		Pt-measure	
		MNSQ	ZSTD	MNSQ	ZSTD	CORR	EXP
EV 1	0.22	1.12	0.7	1.14	-1.4	0.67	0.65
Why 5	0.24	0.09	0.6	1.06	-1.1	0.67	0.65
EV 2	0.34	1.07	-0.8	0.82	1.2	0.67	0.63
What1	0.27	1.04	-0.7	1.36	-0.4	0.64	0.65
How 9	0.15	0.99	1.2	0.69	1.6	0.68	0.72
Why 3	0.13	1.02	1.4	0.71	1.2	0.68	0.69
What 3	0.04	1.02	0.8	0.84	1.1	0.61	0.66
How 10	0.10	1.01	1.3	1.46	0.4	0.60	0.68
What 7	0.08	0.99	0.2	1.04	0.8	0.70	0.66
How 7	0.26	0.81	0.3	1.23	0.7	0.67	0.64
What if 8	0.34	0.87	0.1	0.93	1.0	0.67	0.64
What if 9	0.27	0.91	0.6	0.79	1.7	0.71	0.64
Why 2	0.15	0.99	0.7	0.81	0.1	0.65	0.63
What 2	0.11	0.95	0.2	0.99	-1.1	0.62	0.68
What if 10	0.35	0.87	0.3	1.49	0.6	0.63	0.67
Why10	0.35	0.91	0.6	0.91	-1.2	0.68	0.67
What if 3	0.26	0.87	0.4	1.00	0.7	0.67	0.68
EV 8	0.27	0.93	0.3	0.85	-0.5	0.65	0.66
EV 3	0.62	0.96	0.9	1.05	0.6	0.65	0.69
What 10	0.57	0.94	1.2	0.83	1.2	0.68	0.64
Why 7	0.26	0.86	0.3	0.85	1.2	0.72	0.67
EV 6	0.18	0.83	0.5	0.87	-1.1	0.66	0.68
How 3	0.45	0.94	-1.3	0.74	-1.4	0.68	0.64
EV 7	0.24	0.91	-1.5	0.66	-0.7	0.68	0.69
Why 4	0.35	0.86	0.7	0.85	-0.3	0.68	0.67
How 1	0.45	0.76	-0.4	0.91	-1.6	0.70	0.66
What 4	0.26	1.13	0.6	0.89	-1.6	0.74	0.68
EV 10	0.42	0.87	1.8	0.73	1.4	0.71	0.67
What if 7	0.40	0.78	0.7	1.32	-1.6	0.73	0.73
What if 5	0.12	0.90	-1.1	0.91	-0.5	0.70	0.62
How 4	0.13	0.91	-1.2	0.67	-1.6	0.71	0.64
What if 2	0.42	0.75	-1.3	0.87	-1.3	0.76	0.66
How 5	0.51	0.88	-1.6	0.95	-0.9	0.77	0.63
How 6	0.06	0.90	-1.3	0.63	-1.8	0.74	0.66
Why 6	0.28	0.67	-1.8	0.77	-1.4	0.71	0.71
What if 4	0.37	0.76	-1.6	0.72	-1.7	0.70	0.67

Appendix B. Item dimensionality of the observation card.

	Empirical			Modeled
Total raw variance in observations	93.2	100%		100%
Raw variance explained by measures	53.4	53.2%		51.3%
Raw variance explained by persons	26.3	27.6%		26.7%
Raw variance explained by items	21.6	23.6%		21.7%
Raw unexplained variance (total)	50.0	48.6%	100%	47.1%
Unexplained variance in the 1st contrast	5.3	5.2%	11.0%	
Unexplained variance in the 2nd contrast	5.1	5.1%	10.0%	
Unexplained variance in the 3rd contrast	3.7	3.6%	7.9%	
Unexplained variance in the 4th contrast	3.7	3.6%	7.7%	
Unexplained variance in the 5th contrast	3.3	3.3%	6.3%	

Appendix C. Calibration scaling analysis of observation card.

Category Label	Observed Count %	Observed Average	Sample Expect	Infit MNSQ	Outfit MNSQ	Structure Calibration	Category Measure
1	2(6%)	1.82	-1.66	1.82	1.93	None	(-3.44)
2	5(16%)	-0.35	-.18	1.38	1.61	-2.30	-1.37
3	9(30%)	0.38	0.57	1.04	1.30	-0.38	0.27
4	9(29%)	1.13	1.18	0.67	0.53	1.05	1.54
5	6(19%)	1.85	1.90	1.27	1.26	1.64	3.08

© 2026 Conscientia Beam. All Rights Reserved.

Appendix D. Person and item separation and reliability for the observation card.

	Score	Count	Measure	Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	124.1	50.0	0.80	0.21	1.06	-0.4	1.13	-0.4
S.D	38.6	0.0	1.31	0.05	0.63	3.4	0.70	3.4
Real RMSE	0.19							
Adj. SD.	1.14							
Separation	4.09							
Person reliability	0.92							
Mean	173.7	150.0	0.02	0.21	1.00	0.00	1.10	0.1
S.D	8.3	0.0	0.33	0.01	0.21	1.1	0.66	1.6
Real RMSE	0.18							
Adj. SD.	0.27							
Separation	3.06							
Item reliability	0.72							

Appendix E. Observation card.

	Level of achievement				
	Very High	High	Moderate	Low	Not achieved
Dimension One: What? Stage advanced problem and concept identification					
Formulates lesson objectives using action verbs that reflect higher-order thinking skills appropriate for gifted students.					
Presents a complex, multi-dimensional instructional problem that does not lend itself to a single direct solution.					
Connects the lesson topic to concepts or issues from other disciplinary domains in a clear and meaningful manner.					
Introduces the instructional problem within a real-world, research-based, or future-oriented context that requires advanced processing.					
Identifies the core concepts of the lesson and clearly distinguishes them from subordinate concepts during instruction.					
Guides students in defining the dimensions and boundaries of the problem using structured, guiding questions.					
Allows students to reframe, extend, or propose alternative issues related to the original problem.					
Encourages students to pose higher-order questions and formulate initial hypotheses or preliminary explanations.					
Employs advanced cognitive models, abstractions, and concept maps appropriate for the intellectual level of gifted learners.					
Emphasizes precise use of scientific or disciplinary terminology when defining the problem and its core concepts.					
Dimension Two: Why? Stage intellectual motivation and meaning-making					
Guides students in selecting a variety of strategies to analyze the problem in ways that align with its nature.					
Trains students to deconstruct the problem into its main and subsidiary components in a systematic manner.					
Encourages students to compare multiple alternatives for addressing the problem.					
Enables students to justify their analytical choices using logical or scientific evidence.					
Links the steps of analysis to critical and logical thinking skills.					
Uses multiple learning resources to support in-depth analysis of the problem.					
Directs students to employ advanced analytical tools (models, matrices, diagrams).					
Provides opportunities for scientific dialogue and the exchange of perspectives among students.					
Supports students' autonomy in choosing appropriate methods of analysis.					
Consider individual differences in the depth of analytical reasoning among gifted learners.					
Dimension Three: How? Stage flexible implementation and enrichment					
Guides students to interpret phenomena or results related to the problem using logical, scientifically grounded explanations.					
Encourages students to justify their opinions and conclusions with clear evidence.					

© 2026 Conscientia Beam. All Rights Reserved.

	Level of achievement				
	Very High	**High**	**Moderate**	**Low**	**Not achieved**
Connects cause and effect when discussing the instructional problem.					
Supports the development of relationships among concepts within a coherent explanatory framework.					
Provides opportunities for students to defend their viewpoints according to defined scientific criteria.					
Cultivates students' awareness of the multiple possible interpretations of the problem.					
Directs students to revisit their prior assumptions in light of new evidence.					
Promotes students' use of causal and inferential reasoning.					
Relates explanations to the real-world, professional, or future context of the problem.					
Enhances conceptual precision and clarity of meaning in students' explanations.					
Dimension Four: What If? What are the outcomes? Stage critical and future-oriented thinking					
Presents hypothetical situations that require students to predict alternative outcomes.					
Encourages students to anticipate the short- and long-term consequences of different decisions.					
Guides students in evaluating the risks and opportunities associated with each alternative.					
Supports future-oriented thinking when addressing educational or professional issues.					
Provides opportunities for students to propose innovative and non-traditional solutions.					
Connect prediction with ethically and professionally responsible decision-making.					
Encourages students to use scenario-based analysis in evaluating alternatives.					
Trains students to weigh competing options according to defined criteria.					
Allows for revisiting decisions in light of new hypothetical outcomes.					
Strengthens students' cognitive flexibility and acceptance of multiple possibilities.					
Dimension Five: Assessment and feedback stage authentic evaluation and self-reflection					
Guides students to evaluate their solutions or decisions using clear, explicit criteria.					
Encourages students to engage in self-reflection on their thinking processes and performance.					
Connects learning outcomes to real-world or professional applications.					
Enables students to transfer their learning to new contexts.					
Supports the use of feedback to enhance future performance.					
Guides students to revise their solutions based on assessment results.					
Encourages students to document their learning experiences in an organized manner.					
Promotes a culture of continuous learning and self-improvement.					
Links assessment to standards of quality and professional competence.					
Supports students' autonomy in planning subsequent learning steps.					

Views and opinions expressed in this article are the views and opinions of the author(s), Humanities and Social Sciences Letters shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.

© 2026 Conscientia Beam. All Rights Reserved.