Perceptual Assessment of Science Teaching and Learning

Preliminary Examiner's Manual

Test Development Team:
Bruce. A. Bracken
Kathryn A. Holt
Marvin L. Lee
Christopher J. McCormick
Cristina L. Reintjes
Janice I. Robbins
Tamra L. Stambaugh

The College of William and Mary
School of Education
Center for Gifted Education

CHAPTER 1
INTRODUCTION

The Perceptual Assessment of Science Teaching and Learning (PASTeL; Bracken, Holt, Lee, McCormick, Reintjes, Robbins, & Stambaugh, 208) was designed to assess educators' attitudes toward teaching science and their students' typical response to science instruction. The instrument was developed as part of Project Clarion, a grant awarded to the Center for Gifted Education at the College of William and Mary (Bruce A. Bracken and Joyce Van Tassel-Baska, Co-Principal Investigators) from the United States Department of Education's Jacob K. Javits Program.

Rationale for the PASTeL

The PASTeL was conceptualized as a dependent measure within Project Clarion, a research and implementation study exploring the effects of a concept-based science curriculum on students' science achievement and critical thinking in Kindergarten through third grade. The instrument, however, was designed to be employed in science classes from Kindergarten through twelfth grade. In the final year of Project Clarion implementation, the research team searched for an existing scale to assess teachers' self-perceptions of their science instruction and student learning; however, after a review of the literature, no scale with adequate content coverage or psychometric quality was found. Therefore, the grant's co-principal investigators decided that an instrument should be developed and an author team was created to construct an instrument that would be comprehensive, research based, and applicable to all elementary and secondary grade levels.

The PASTeL's primary purpose is to assess how teachers' perceive their ability to teach, as well as their students' ability to learn, science and how those perceptions are influenced by their Project Clarion curricular experiences. Additionally, the PASTeL developers intentionally designed the assessment to have a wider application to science education in general; that is, the instrument was intended to be useful for investigators of science instruction, beyond the application of the Project Clarion curriculum.

Goals for the PASTeL

During the development of the PASTeL, the author team created and refined the instrument with several goals in mind. The primary goals guiding the PASTeL development and refinement included a desire to create an instrument that would:

- be theoretically and psychometrically sound.
- be comprehensive.
- be brief and easy to administer.
- evaluate important components of science instruction.
- evaluate important attitudes toward science instruction.

Applications of the PASTeL

The PASTeL was designed to be a research-based assessment tool for the evaluation of attitudes toward science instruction and learning among K-12 science teachers. The PASTeL appears well-suited for use as a dependent measure because it was theoretically conceived, rigorously developed, and has been shown to be psychometrically sound, as well as brief and easily administered. The PASTeL can be administered individually or to groups of teachers. Also, the PASTeL can be either hand scored or completed online.

CHAPTER 2
SCALE DEVELOPMENT

The process of developing the PASTeL involved several steps. First, test development began with a literature review of the components of effective science instruction based on the conceptual framework for Project Clarion's curriculum design and differentiation. This literature review resulted in a set of three core categories and four sub-categories that provided the framework or blueprint for scale development. Within this framework, a pool of items was constructed to reflect aspects of effective science instruction across categories. Finally, a structure within which to respond to the different items was created based on two perspectives: (a) respondents' view of their own science instruction ability, and (b) respondents' view of their students' ability to learn science. Each of these components is described in more detail below.

Review of the Literature

The literature in education has identified several critical components of effective teaching practices and science instruction specifically. These components include an integration of teaching science content and the process of scientific investigation while also linking to overarching science concepts. The Integrated Curriculum Model (ICM), a theoretical model of curriculum design for gifted learners, emphasizes this integration of advanced content, higher order thinking processes, and connections to overarching themes and issues as the foundation for curriculum development. The greatest student learning gains have been shown to occur when emphasis is given to the teaching of these dimensions within a given curriculum unit (Van Tassel-Baska, 1986). Expert teachers must have specific pedagogical-content expertise (Shulman, 1987) that includes knowledge of how best to explicate particular concepts (e.g., how to explain gravity), demonstrate and rationalize procedures and methods (Leinhardt &

Greeno, 1986), and correct students' naïve theories and misconceptions (Gardner, 1991).

The knowledge of general science content is obviously important to science education. Students must have a strong understanding of science facts and content on which to build their understanding of science. Richard T. White (1998) wrote that the more complex content is, there is a greater need to integrate it and to show and perceive its unity. The National Science Teachers Association's (2004) position statement on scientific inquiry also noted that scientific inquiry is a powerful tool for understanding science content because students learn how to ask questions and then use evidence to answer them.

The National Science Education Standards (NSES, 1996) states that "inquiry is central to science learning" (p.2). Bransford and Donovan (2005) also affirm the importance of scientific inquiry in science instruction, saying that it should be the method for learning content and must go beyond simply memorizing and reproducing the steps of an experiment or setting aside "inquiry time." Instead they recommend students learn science as a process of inquiry using skills such as observation, imagination and reasoning about the phenomena under study, using tools that extend their experiences and organizing data in order to develop new insights. By using inquiry to teach science process skills, students experience the "excitement of actually discovering—and sharing with friends—something that provides a new way of looking at the world" (Bransford & Donovan, 2005, p. 414). Not only does inquiry help students learn science content, but it also motivates them to continue to learn about and explore their world.

In addition to science content and process skills, the teaching of overarching concepts is also important in effective science education. A recent national report underscored the importance of teaching science for deep conceptual understanding at the pre-collegiate level to ensure opportunities for advanced learning (National Research Council [NRC], 2002). The report acknowledged the central role of concept development in the process of learning science deeply, noting the principles of learning with understanding, using meta-cognitive strategies, building on prior knowledge, and creating a community of learners to sustain modeled behavior. Concepts have been shown to help students make inferences and explain complex ideas (Carey, 2000). The American Association for the Advancement of Science (1993) also states the importance of understanding concept-based ideas/themes such as systems, scale, change, and constancy, and their applications in science, mathematics and technology. Research on teaching science concepts highlights the importance of learners' previous experiences as an important feature in conceptual understanding. Studies have shown, for example, that procedural knowledge structures (reasoning patterns) in science predict readiness for instruction in descriptive and theoretical concepts (Johnson & Lawson, 1998; Kwon & Lawson, 2000). Individual concepts also have been shown to exist within complex conceptual systems such that knowledge and understanding of these concepts are deepened by learning related concepts (Mintzes, Wandersee, & Novak, 1998), a process that facilitates transfer of conceptual understanding from one domain to another.

Other important components of science instruction include teaching critical thinking skills, curiosity and creativity. In inquiry, students must combine scientific knowledge and process skills "as they use scientific reasoning and critical thinking to develop understanding" (Lind, 1998, p. 10). In the Delphi Report, a panel of experts defined critical thinking as "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based" (Facione, 1990, p. 2). The report also recommended that critical thinking be explicitly taught at all grade levels. In addition to critical thinking skills, curiosity is also an integral component of science education. Children are "born curious and want to know all about their environment" (Lind, 1998, p. 3) and this curiosity should be encouraged in order to create interest in science and persistence in completing science related tasks.   Developing creativity is also important to science education. Students need to be able to produce new and unique ideas and solutions to scientific problems as well as to develop flexible thinking in order to truly think like scientists. When Torrence examined 142 studies of teaching children to think creatively he found that it is possible to teach students this skill, especially when “deliberate teaching is involved” (Torrence, 1972, p. 203).

Based on this literature review, the PASTeL development team began writing items to assess teachers' attitudes, focusing on each of components of effective science instruction discussed above.

Scale Blueprint

The final form of the PASTeL consists of 50 items arranged in two scales, the *Teaching Scale* and the *Student Learning Scale*. The *Teaching Scale* includes 25 items that assess educators' perceptions about their own science instruction. The *Student Learning Scale* consists of 25 items that assess educators' perceptions about their students' acquisition of science content. Respondents are directed to rate each statement according to the degree to which they agree with the statement, selecting from the following options: “strongly agree,” “agree,” “disagree,” and “strongly disagree.”

Within each scale, three categories of items are represented: content, process, and concept. Each of these categories was then divided into four additional subcategories: proficiency, critical thinking, curiosity and creativity. These categories are described in Table 2.1.

Table 2.1

Description of Item Categories

Category of Science Instruction	**Description of Category**
Content	The knowledge of general science content.
Process	The understanding of fundamental science process skills including *make observations*, *ask questions*, *learn more, design and conduct an experiment*, *create meaning from the experiment*, *and tell others what was found.*
Concept	The development of overarching concepts related to the understanding of the world of science such as patterns, change, systems and cause & effect by determining specific examples and non-examples of concepts, categorizing, establishing generalizations, and applying generalizations to other situations (Taba, 1962).
Proficiency	How teachers perceive their ability to teach and their students' ability to learn science content, process skills and concept development at a range of levels (i.e. knowledge, comprehension, application, analysis, synthesis, and evaluation) (Bloom, 1956).
Critical Thinking	How teachers perceive their ability to teach and their students' ability to learn science content, processes and concepts through the development of critical thinking skills including defining an issue/problem, defining relevant concepts, identifying different points of view, describing evidence/data supporting a point of view, drawing conclusions, and predicting consequences (Paul, 1992).

Curiosity	How teachers perceive their ability to teach and their students' ability to learn science content, processes and concepts through the development of curiosity and interest in the world of science.
Creativity	How teachers perceive their ability to teach and their students' ability to learn science content, processes and concepts through the development of creative thinking as demonstrated by fluency, flexible thinking, elaboration, and originality.

Comment [bab1]: Remove open box at bottom of previous page

On both the *Teaching Scale* and the *Student Learning Scale*, 24 items were included for the proficiency category, 14 items for the critical thinking category, six items for the creativity category and six items for the curiosity category. Table 2.2 classifies items into the seven categories by their item numbers on the two scales.Table 2.2

Deleted: ¶
¶

PASTeL Items Classified by Category

	Items on Teaching Scale			**Items on Student Learning Scale**		
Item Category	Content	Process	Concept	Content	Process	Concept
Proficiency	1,2,3,4,5	11,12,13, 14, 15	21, 22	1,2,3,4,5	11,12,13, 14, 15	21, 22
Critical Thinking	7,8,9	17, 18, 19	24	7,8,9	17, 18, 19	24
Curiosity	6	16	23	6	16	23
Creativity	10	20	25	10	20	25

Development of Categories and Items

The initial framework for the PASTeL was based on the model used to create the science curriculum for Project Clarion. Based on this conceptualization, discussed in the literature review section above, the developers created three overarching categories: *content, process,* and *concept.* Four subcategories were then classified across these three larger groups: *proficiency, critical thinking, curiosity, and creativity.* These subcategories were then divided further to aid in item development. The *proficiency, critical thinking, and creativity* categories were subdivided to ensure that each aspect of that category was equally sampled across items. The *proficiency* category was divided into six subcategories: *knowledge, comprehension, application, analysis, synthesis and evaluation (Bloom, 1956).* The *critical thinking* category was divided into seven subcategories: *purpose, concept, point of view, assumptions, evidence, inference, and implications/consequences* (Paul, 1992). Finally, the *creativity* category was separated into four subcategories: *fluency, flexibility, originality, and* elaboration. Based on this framework and a review of additional literature related to science instruction research and recommendations, the PASTeL development team began writing items within the three categories and four subcategories.

For each category identified, members of the development team drafted items individually, and then the team collaboratively reviewed, discussed and edited items. Items were reviewed for applicability to the category, reflection of the literature on effective science instruction, appropriateness across grades K-12, and general clarity and wording. Additional considerations for item development and editing included the following:

- Items should focus on a single key idea.
- Items should be brief and easily read.
- Items should make limited use of educational jargon or “trendy” language.

Piloting the PDQ

Selected items were collected into a pilot form of the instrument, which included 128 items, with 64 items on the *Teaching Scale* and 64 items on the *Student Learning Scale.*

Following the piloting of the PASTeL, items in each category were reviewed for their contributions to the overall content requirements of the blueprint and the reliability of the scales and the total instrument. Items were retained based on both their statistical contributions and their reflection of the theoretical framework. Based on the results of the team review, the test was revised and reduced in length. The wording of two chosen items (item 23, *Teaching Scale;* item 24, *Student Learning Scale*) were changed to make them more clear and understandable.

The decisions made based on the PASTeL pilot analysis resulted in the 50-item scale described earlier. This final 50-item instrument went through a final tryout on a larger sample of teachers, and the results are discussed in chapter 4.

CHAPTER 3
ADMINISTRATION AND SCORING

This chapter presents the procedures for administering, completing, and scoring the PASTeL in either group or individual administration formats.

Materials Needed: To complete the PASTeL, participants need only a pen or pencil, the PASTeL *Teaching Scale*, and the PASTeL *Student Learning Scale* forms.

Time: The administration of the PASTeL is NOT timed. Participants may take as long as needed to complete the forms; however, based on scale piloting and tryouts, typical scale completion should require approximately five minutes.

Administration:

The PASTeL *Teaching Scale* may be administered any time researchers or educators wish to learn what professionals think about how they generally teach science; the *Student Learning Scale* should be administered to learn what professionals think about how their students typically respond to science instruction.

Participants are instructed to read the following directions at the beginning of the assessment:

> *Please rate the following statements according to how well the statement applies to you. There are no right or wrong answers, but it is important that you rate each statement according to how you honestly feel. Be sure to be honest with yourself as you consider the statement you are rating. To make your answer, simply circle the letters that correspond with your feelings towards the statements. Each statement should be rated as:*
>
> *Strongly Disagree (SD)*
> *Disagree (D)*
> *Agree (A)*
> *Strongly Agree (SA)*

Participants also are instructed to read the following directions at the top of the *Teaching Scale*:

> *Please focus your ratings on how you generally teach science.*

Participants then complete the 25 items on the *Teaching Scale.* They are then instructed to read the following directions at the top of the *Student Learning Scale*:

> *Please focus your ratings on how your students typically respond to science instruction.*

Participants then complete the 25 items on the *Student Learning Scale*. Participants may complete the form at their own pace, and should submit the form when they have finished.

Scoring the PASTeL

Scoring the Teaching Scale and *Student Learning Scale.* Both of the PASTeL scales are scored according to the following format: items that are endorsed as Strongly Agree are credited with 4 points; 3 points are given to Agree; 2 points for Disagree; and, 1 point for Strongly Disagree. The item weights are summed for each 25-item scale, resulting in a total *Teaching Scale* score and a total *Student Learning Scale* score. If a respondent omits one or more items, the total score should be prorated. Prorating can be accomplished as follows:

Step 1: Divide the Scale item raw score sum by the number of items completed
Step 2: Multiply that dividend by 25

Example: If the respondent completed 23 items, which summed to a Total Raw Score of 72, the prorated score would be (72/23 = 3.13) * (25) = 78.26 or 78

PASTeL Scoring Worksheet

	Total Raw Score
Teaching Scale	☐
Student Learning Scale	☐

Total *Teaching* Raw Score		Total *Student Learning* Raw Score		Total Instrument Raw Score
☐	+	☐	=	☐

Pilot Analyses

During the PASTeL pilot study, a total of 72 forms were completed by science teachers (70 female, 2 male) of Kindergarten, first, second, and third grades. The means and standard deviations for the three PASTeL scales for this sample were as follows:

	Mean	**SD**
Teaching Scale	79.36	8.34
Student Learning Scale	71.32	8.51
Total Instrument Raw Score	150.78	15.79

CHAPTER 4

TECHNICAL ADEQUACY

During the piloting of the PASTeL a total of 34 teachers (8 males, 25 females, 1 gender not specified) responded to the original 128 PASTeL items. Coefficient alpha reliabilities were calculated for each of the two scales (*Teaching* and *Student Learning*), and those items that contributed least to scale reliability were omitted. Based on these preliminary reliability and content analyses, the PASTeL was reduced to 50 items, with 25 items on the *Teaching Scale* and 25 items on the *Student Learning Scale.*

Reliability

Test reliability represents the proportion of variance that results from meaningful variation in test scores, as opposed to score variation that results from error. The authors of the PASTeL sought to reduce measurement error by ensuring that the PASTeL was as easy as possible for examinees to take and as brief as possible to reduce the risk of an adverse environmental setting. They also sought to ensure that the test was sufficiently well-written and objectively scored to produce reliable estimates of respondents' attitudes toward science instruction. These were the goals the authors hoped to achieve during the development of the PASTeL.

Internal consistency refers to the extent to which items within a scale are positively correlated and contribute to the reliable variation of scores. One would expect that items drawn from the same content areas within an instrument would be positively correlated to a moderate degree. Bracken (1987) suggested (Wasserman & Bracken, 2003 elaborated) that tests intended for research applications should minimally be reliable at a level of .70, and preferably .80.

This level of reliability (i.e. .70 to .80) was the goal that guided PASTeL item selection, retention, and test refinement.

Internal Consistency. The principal issue in PASTeL internal consistency is the extent to which individual items contribute meaningfully to the total scale score. To produce a measure with high internal consistency, all of the PASTeL items should correlate positively and moderately with each other and their combined total score.

Table 4.1 presents estimates of internal consistency (i.e., coefficient alpha) for the 50 item PASTeL *Teaching Scale*, *Student Learning Scale*, and Total Instrument scores for the preliminary sample of 34 respondents. The scale and total test reliability are all greater than the .70 to .80 criterion set by Bracken (1987) and Wasserman & Bracken (2003) for research instruments. Such strong preliminary estimates of internal consistency suggest that researchers can expect the PASTeL to provide quite consistent estimates of teachers' attitudes toward teaching science and their students' typical response to science instruction. Consistent with the scale reliabilities, the 50-item PASTeL Total Instrument reliability is .95. Based on these preliminary total sample estimates of internal consistency, examiners can expect to use the PASTeL as a highly dependent measure in research with considerable confidence.

Table 4.1

PASTeL Scale and Total Assessment Reliability Coefficients for 34 Educators

Scale	**Internal Consistency (Coefficient Alpha)**	**Number of Items**
Teaching Scale	0.91	25
Student Learning Scale	0.94	25
Total Instrument	0.95	50

Standard Error of Measurement. The standard error of measurement (SEM) of a scale is directly proportional to the reliability of the instrument. From its computational formula, it is apparent that as a scale's reliability increases, its *Standard Error of Measurement* (SEM) decreases. This functional relationship further indicates the importance of reliability and its contribution to the confidence an examiner has that an examinee's obtained scores are representative of the individual's theoretical "true scores." Because the PASTeL *SEM*s are quite small (approximately 2.5 and 2.08 raw score points for the 25-item *Teaching Scale* and *Student Learning Scale*, respectively, and 3.53 for the 50-item *Total Instrument*), the individual scales' estimated true scores lie in fairly tight bands of confidence.

Scoring Reliability. Because the PASTeL is objectively scored by hand, scoring reliability is not an issue that needed to be investigated. Scoring reliability is assured on the PASTeL, except to the extent that there exists the possibility of occasional clerical or mechanical scoring errors—minor or occasional scoring errors that exist in all objectively scored scales.

Validity

A scale is considered valid to the extent to which it measures what its authors claim that it measures. The PASTeL is an assessment that its authors propose assesses teachers' attitudes toward science instruction and their students' ability to learn science. Given this basic definition of validity and the focus of the PASTeL, preliminary PASTeL validity is demonstrated through content validity.

Content Validity. The PASTeL was based theoretically on a corpus of literature describing recommended components of science instruction. The scale blueprint depicted in Chapter 2 illustrates the extent to which the scale items were constructed to match the components of effective science instruction.

References

Comment [CLR2]: References that are highlighted were pulled from the document titled: "Project Clarion: An Integrative Curriculum Scale-Up to Promote Scientific Conceptual Understanding in Promising Young Children" I don't have the reference list from this document—I need someone with access to the G drive to find it and fill in the references.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy.* New York: Oxford University Press.

Bloom, Benjamin. S., (Ed.). 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals; Handbook I, Cognitive Domain.* New York: David McKay.

Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*, *5*, 313-326.

Bransford, J. D., & Donovan, M. S. (2005). Scientific inquiry and how people learn. In Bransford, J. D., & Donovan, M. S. (Eds.), *How students learn: History, mathematics and science in the classroom*. (397-420). Washington, DC: National Academies Press.

Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology. 21 (*1*),* 13-19.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction: Executive summary*: The delphi report. Millbrae, CA: The California Academic Press.

Gardner, 1991

Johnson & Lawson, 1998

Kwon & Lawson, 2000

Leinhardt, G., & Greeno, J. G. (1986). The Cognitive Skill of Teaching. *Journal of Educational Psychology*. *78*(2), 75-95.

Lind, K. K. (1998 February). *Science in early childhood: Developing and acquiring fundamental concepts and skills.* Paper presented at the Forum on Early Childhood Science, Mathematics, and Technology Education, Washington, D.C.

Mintzes, Wandersee, & Novak, 1998

National Research Council. 1996. *National Education Standards.* Washington, DC: National Academy Press.

National Research Council [NRC], 2002

National Science Teachers Association. (2004). *NSTA position statement: Scientific inquiry.* Arlington, VA.

Bracken, B. A., Holt, K. A., Lee, M. L., McCormick, C. J., Reintjes, C. L., Robbins, J. I, & Stambaugh, T. L. (2008). *Perceptual Assessment of Science Teaching and Learning (PASTeL).* Williamsburg, VA: Center for Gifted Education, The College of William & Mary.

Paul, R. (1992). Critical Thinking: What, Why, and How. *New Directions for Community Colleges*. *20*(1), 3-24.

Shulman, 1987

Taba, H. (1962). *Curriculum: Theory and Practice.* New York: Harcourt, Brace.

Torrence, E. P. (1972). Teaching for creativity. *The Journal of Creative Behavior (6)*, 114-143.

Van Tassel-Baska, J. (1986). Effective curriculum and instructional models for talented students. *Gifted Child Quarterly, (30),* 164-169.

Wasserman, J. D., & Bracken, B. A. (2003). Psychometric considerations of assessment procedures. In J. Graham and J. Naglieri (Eds). *Handbook of assessment psychology* (pp. 43-66). New York: Wiley.

White, R. T. (1994). Dimensions of content. In Fensham, P. J., Gunstone, R. F., & White, R. T. (Eds.), *The content of science: A constructivist approach to its teaching and learning.* Routledge.