UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES

CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES
DEPARTAMENTO DE EDUCACIÓN

TESIS

DISEÑO Y VALIDACIÓN DE UN PROTOCOLO DE OBSERVACIÓN PARA EVALUAR LAS ACTIVIDADES DE ENSEÑANZA EN QUINTO GRADO DE PRIMARIA EN LA ASIGNATURA DE MATEMÁTICAS

PRESENTA

Lesly Yahaira Rodríguez Martínez

PARA OBTENER EL GRADO DE DOCTORA
EN INVESTIGACIÓN EDUCATIVA

TUTORA
Dra. María Guadalupe Pérez Martínez

COMITÉ TUTORAL
Dra. Guadalupe Ruiz Cuéllar
Dr. Paul Alberto Hernández Martínez

Aguascalientes, Ags., 22 de febrero de 2018

Carta de aceptación del artículo

El 02/02/2016, a las 10:58 a.m., Kayla Wolfe <kwolfe@igi-global.com> escribió:

Dear Dr. Montoya,

We have just completed the review process of your Edited book proposal. Our external reviewers were very positive about your book proposal and recommended the publication of this book by IGI Global with the title, "**Driving STEM Learning with Educational Technologies**."

Based on the topics and themes to be presented in this title, IGI Global has tentatively placed this title in the **Advances in Educational Technologies and Instructional Design (AETID)** book series. Please note that IGI Global selects book series inclusion based on the topics and themes to be presented in the manuscript. Book series placement is subject to change at the discretion of the publisher.

Shortly, I will ask our contract and agreement facilitator, Ms. Jan Travers, to prepare a contract for this proposed book, using the finalized title given above, and forward it to you for your signature.

Please note that all items of the contract will be formatted in accordance with the proposal submission details and contact information (editor order, contributor listing, etc.). Please advise as soon as possible if there are any editor additions, changes or title discrepancies.*

Upon the completion of the contract formality, our development team will begin working with you toward the materialization of this project. He/She will provide you with all the necessary information and support for producing a quality publication with IGI Global.

All of us at IGI Global look forward to working with you on this project!
Kind regards,

Kayla Wolfe
Managing Editor, Acquisitions

International Publisher of Progressive Information Science and Technology Research
701 E. Chocolate Avenue
Hershey, Pennsylvania 17033-1240, USA
Tel: 717-533-8845 ext. 141 Fax: 717-533-8661
E-mail: Kwolfe@igi-global.com Website: www.igi-global.com
IGI Global is committed to the highest quality standards and excellent service.
This is not just our promise. This is our business.

Capítulo publicado en Handbook: Driving STEM Learning with Educational Technologies

Chapter 9

Assessing Authentic Intellectual Work in Mathematics Tasks

Lesly Yahaira Rodríguez Martínez
Universidad Autónoma de Aguascalientes, Mexico

María Guadalupe Pérez Martínez
Universidad Autónoma de Aguascalientes, Mexico

Adriana Mercado Salas
Universidad Autónoma de Aguascalientes, Mexico

ABSTRACT

This paper reports an analysis of the tasks included in the Mathematical Challenges book. The analysis was based on the proposals of the Authentic Intellectual Work (AIW). The purpose of the study focuses on assessing the potential of the mathematical challenges to promote in-depth and meaningful learning through the connection with different contexts, and other features including purpose, multiple-solution pathways, construction of knowledge and higher order thinking. Participants in this study were 3 elementary school teachers, 2 mathematics specialists and the authors of this paper; they assessed the Mathematical Challenges through a questionnaire based on specific rubrics. The study used a mixed methods approach. The analysis produced two main findings. First, challenges vary in their connections to students' lives according to the context they come from. Second, almost all mathematical challenges are related to the highest levels of others AIW criteria.

INTRODUCTION

Tasks are essential for any learning processes. Students' academic achievement is influenced by what happens in the classroom, by the interaction of teachers and students with the curriculum through learning tasks. However, it has been asserted that many of the tasks teachers offer to their students as part of the educational processes tend to focus on issues that do not stimulate active learning, since they require students to perform isolated exercises (Newmann, King & Carmichael, 2007).

Newmann and Wehlage (1993) note that in most classrooms it is possible to identify two persistent problems in relation to the type of tasks that teachers suggest students carry out: a) tasks that do not

DOI: 10.4018/978-1-5225-2026-9.ch009

UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES

DRA. GRISELDA ALICIA MACÍAS IBARRA
DECANA DEL CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES
P R E S E N T E

Por medio del presente, como Tutor designado de la estudiante **LESLY YAHAIRA RODRÍGUEZ MARTÍNEZ,** con ID 68804, quien realizó la tesis titulada: ***DISEÑO Y VALIDACIÓN DE UN PROTOCOLO DE OBSERVACIÓN PARA EVALUAR LAS ACTIVIDADES DE ENSEÑANZA EN QUINTO GRADO DE PRIMARIA EN LA ASIGNATURA DE MATEMÁTICAS***, y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia, me permito emitir el **VOTO APROBATORIO**, para que ella pueda proceder a imprimirla, así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

A T E N T A M E N T E
"SE LUMEN PROFERRE"

Aguascalientes, Ags., a 22 de febrero de 2018

Dra. María Guadalupe Pérez Martínez
Tutor de tesis

Dra. Guadalupe Ruiz Cuéllar
Integrante Comité Tutorial

Dr. Paul Alberto Hernández Martínez
Integrante Comité Tutorial

c.c.p. Interesado
c.c.p. Secretaría Técnica del Doctorado en Investigación Educativa

UNIVERSIDAD AUTONOMA DE AGUASCALIENTES
CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES

DEC. CCS Y H OF. N° 0261
Asunto: Conclusión de Tesis

DRA. EN ADMÓN. MARÍA DEL CARMEN MARTÍNEZ SERNA
DIRECTORA GENERAL DE INVESTIGACIÓN Y POSGRADO
P R E S E N T E.

Por este conducto le informo que el documento final de Tesis/Trabajo Práctico Titulado: **"DISEÑO Y VALIDACIÓN DE UN PROTOCOLO DE OBSERVACIÓN PARA EVALUAR LAS ACTIVIDADES DE ENSEÑANZA EN QUINTO GRADO DE PRIMARIA EN LA ASIGNATURA DE MATEMÁTICAS"**, presentado por la sustentante **LESLY YAHAIRA RODRÍGUEZ MARTÍNEZ** con **ID. 68804**, egresada del **DOCTORADO EN INVESTIGACIÓN EDUCATIVA**, cumple las normas y lineamientos establecidos institucionalmente para presentar el examen de grado.

Sin más por el momento, aprovecho la oportunidad para enviarle un cordial saludo.

A T E N T A M E N T E
"SE LUMEN PROFERRE"
Aguascalientes, Ags. A 27 de Febrero del 2018

DRA. GRISELDA ALICIA MACÍAS IBARRA
DECANA

c.c.p. Dr. Francisco Javier Pedroza Cabrera. Secretario de Investigación y Posgrado del CCS y H.
c.c.p. Dra. Laura Elena Padilla González. Secretaria Técnica del Doctorado en Inv. Educativa
c.c.p. Mtra. Imelda Jiménez García. Jefa del Depto. De Control Escolar
c.c.p. Mtra. Lesly Yahaira Rodríguez Martínez. Egresada del Doctorado en Investigación Educativa
c.c.p. Archivo

Agradecimientos

A mi tutora, la Dra. Guadalupe Pérez Martínez, por su paciencia, su confianza y sobre todo por guiarme en este arduo pero satisfactorio proceso que significó el doctorado.

Al Mtro. Felipe Martínez Rizo, por sus consejos, su confianza y por su constante apoyo.

Al Dr. Paul Hernández Martínez, por acompañarme durante el proceso del doctorado y guiarme durante una estancia de investigación bajo su tutela.

A la Dra. Guadalupe Ruiz Cuéllar por su apoyo, sus palabras de aliento y sugerencias de mejora.

Al Dr. Horacio Pedroza Zúñiga, por su apoyo, asesoría y sugerencias de mejora.

Al Consejo Nacional de Ciencia y Tecnología (Conacyt), por el apoyo que me brindó para concluir el programa de Doctorado en Investigación Educativa.

A todos los docentes que participaron en este estudio y permitieron videograbar sus clases para obtener los insumos necesarios.

A todos ellos, muchas gracias.

Dedicatorias

A mi madre, una maravillosa mujer y excelente profesionista quien ha sido mi inspiración, mi soporte y la fuerza que me alienta para lograr mis metas.

A mi padre, por darme ánimos para continuar con mis proyectos.

A mis hermanos, Mayra, Jonathan y Devany, los tres fuente de inspiración, cariño y respeto.

A mis amigos, por su tiempo y confianza.

ÍNDICE

Índice de tablas

Índice figuras

Resumen

El propósito de este estudio fue diseñar un protocolo de observación que permitiera recabar información válida y confiable sobre las actividades de enseñanza en la asignatura de matemáticas, entendiendo como “actividades de enseñanza” las acciones y procedimientos que los docentes realizan en el aula para facilitar la construcción de conocimientos de los estudiantes.

Las dimensiones que evalúa el Protocolo de Observación de las Actividades de Enseñanza en Matemáticas (POAEM) son:

a. Claridad de la tarea matemática: comunicación sobre el trabajo que deben realizar los estudiantes y la comprensión que muestran al respecto.
b. Exigencia cognitiva: tipo y nivel de pensamiento requerido para realizar las tareas matemáticas propuestas.
c. Recursos empleados durante la tarea: uso que docentes y estudiantes hacen de los recursos disponibles para promover o fortalecer los aprendizajes.
d. Alternativas de solución: flexibilidad del docente para promover y aceptar una variedad de estrategias o procedimientos para resolver un problema matemático.
e. Preguntas para la reflexión: tipo de preguntas que se generan a partir de la tarea matemática para dar sentido a ideas o relaciones matemáticas importantes.
f. Conexión/aplicación de los conocimientos matemáticos: oportunidad de conectar los aprendizajes matemáticos con la experiencia de los estudiantes, con otras disciplinas o situaciones de la vida cotidiana.
g. Trabajo colaborativo: interacción, negociación y cooperación entre los estudiantes para lograr un propósito común.

Los participantes en el estudio fueron 20 docentes de escuelas primarias públicas, a quienes se videograbó en dos ocasiones mientras impartían la clase de matemáticas. En total se recabaron 40 videograbaciones que a su vez se seccionaron en 111 tareas matemáticas, ya que se utilizó la tarea matemática como unidad básica de análisis de las actividades de enseñanza.

El proceso de construcción del POAEM se documentó con el propósito de ofrecer un referente teórico para otros estudios metodológicos que tengan como objetivo el diseño de instrumentos para evaluar la práctica docente. Asimismo se señalan los análisis estadísticos utilizados para ofrecer evidencias de validez y confiabilidad de la información recabada con el POAEM.

Los resultados muestran que la estructura dimensional del instrumento se agrupa en un factor y no en cuatro como se había propuesto con base en la revisión de literatura. Esta información se corroboró con un Análisis Factorial Confirmatorio a partir del cual se propone una nueva organización dimensional para el instrumento (ver apartado de resultados). El cambio en la estructura dimensional del POAEM mostró que todas las variables, excepto la de claridad de la tarea matemática, se relacionan con el mismo constructo (actividades de enseñanza).

El estudio de generalizabilidad aporta información sobre las distintas fuentes de error en la medición, y mostró que la mayoría de las dimensiones reportaron porcentajes altos de error. Es decir el modelo propuesto para estimar la varianza de las diferentes facetas explicó poco sobre las fuentes de error para cada una de las dimensiones (55%-60%). Sin embargo, las dimensiones en conjunto explicaron cerca del 80% de la varianza, por lo que es posible que el comportamiento de los datos se deba al modelo desbalanceado a partir del cual se realizó el estudio de generalizabilidad, a las ocasiones de observación o a algunos aspectos del diseño del instrumento.

Considerando que a nivel nacional falta recorrer un largo camino para el desarrollo de instrumentos que aporten bases sólidas para obtener información válida y confiable sobre diferentes aspectos de la práctica de enseñanza (López y Mota; 2003 y Ávila *et al*. 2013). Se espera que el proceso que se ha seguido para desarrollar este estudio sirva como un referente para futuras investigaciones que tengan como propósito diseñar instrumentos para evaluar la práctica docente, ya que en la actualidad la política educativa apunta hacia la evaluación del desempeño docente, considerando que su papel y función, hasta ahora irreemplazable, es una pieza clave para lograr la mejora de la calidad de la educación.

Abstract

This study was aimed at designing an observational protocol which gathers valid and reliable information about the teaching activities in mathematics subject, understanding as teaching activities those actions and procedures that teachers perform to facilitate students construction of knowledge.

The dimensions evaluated by the Observation Protocol for Teaching Activities in Mathematics (POAEM) are:

a. Clarity of the task: communication about the work students must to do and the understanding they show about it.
b. Cognitive demand: type and level of thinking required to perform the proposed mathematical tasks.
c. Resources used during the task: use the available resources to promote or strengthen learning.
d. Alternative solutions: teacher's flexibility to promote and accept a variety of strategies or procedures to solve a mathematical problem.
e. Questions for reflection: questions that are generated from the mathematical task to give meaning to important mathematical ideas or relationships.
f. Connection / application of mathematical knowledge: opportunity to connect the mathematical learning with the students' experience, with other disciplines or with everyday life.
g. Collaborative work: Interaction, negotiation and cooperation among students to achieve a common purpose.

Participants in the study were 20 teachers from public elementary schools; all of them were videotaped twice while they were teaching mathematics. A total of 40 videotapes were collected and divided into 111 mathematical tasks, since the mathematical task was used as the basic unit of analysis of teaching activities.

The POAEM construction process was documented with the purpose of offering a theoretical reference for other methodological studies aim to design instruments to assess teaching practice. Likewise, the statistical analyzes used to provide evidence of validity and reliability of the information collected with the POAEM are indicated.

Results shown that the dimensional structure of the instrument is grouped only in one factor and not into four factors as had been proposed based on the literature review. The information was confirmed with a Confirmatory Factor Analysis which proposes a new dimensional organization for the instrument (see the results section). The change in the dimensional structure of the POAEM showed that all variables except clarity of the mathematical task are related to the same construct (teaching activities).

The generalizability study provides information on the different sources of error in a measurement, for this study showed that most of the dimensions reported high percentages of error. The proposed model to estimate the variance of the different facets explained little about the sources of error for each of the dimensions (55%-60%). However, dimensions together accounted for about 80% of the variance, so it is possible that the behavior of data was due to the unbalanced model from which the generalizability study was through, to the observation occasions considered, or to the design of the instrument.

Despite the difficulties encountered during the process of designing the instrument, and on the process of validating the information, the contribution of this study focuses on offering a theoretical reference for other methodological studies that aim to design instruments to assess teaching practice. Through these processes it is expected to pay the referents for the development of instruments to evaluate teaching practice.

Introducción

Este capítulo muestra el propósito general del estudio, los elementos teóricos que lo guían y los alcances que se espera lograr a partir del estudio. El capítulo está dividido en tres apartados. El primero define las actividades de enseñanza y las establece como objeto de estudio; el segundo apartado presenta información sobre las oportunidades de aprendizaje que se pueden generar a partir de las actividades de enseñanza, y, el tercero plantea la importancia de contar con instrumentos que permitan obtener información válida y confiable como principal elemento para llevar a cabo la evaluación de la práctica de enseñanza.

1. Actividades de enseñanza como objeto de estudio

Preparar a los estudiantes para su futuro es uno de los propósitos de todo sistema educativo; por lo tanto, el énfasis en el desarrollo de competencias para la vida tiene que ver con la "formación de un ser universal, competitivo como ciudadano del mundo, responsable, activo y capaz de aprender a lo largo de su vida" (SEP, 2011b, p. 25). Ante dicho propósito, en México, la Reforma Integral de Educación Básica (RIEB), generalizada a partir de 2011, ha propuesto un modelo de enseñanza encaminado a favorecer un aprendizaje que promueva la participación y reflexión continua de los estudiantes a través del diálogo, la colaboración y la construcción de conocimientos.

En este modelo de enseñanza se enfatiza que el docente debe ser un mediador que escucha a sus estudiantes, se interesa por conocerlos y por entender sus necesidades de aprendizaje. Lo anterior involucra la puesta en práctica de una forma de enseñanza compleja, cuyos principios didácticos se fundamentan en lo que el conocimiento científico ha encontrado que puede favorecer un aprendizaje profundo, es decir, a través de una enseñanza que toma en cuenta las características diferenciadas del alumnado; que promueva el descubrimiento del propio conocimiento y favorezca un aprendizaje significativo (Bransford, Brown y Cocking, 1999; Perrenoud, 2012).

El que los estudiantes "desarrollen habilidades superiores del pensamiento para solucionar problemas, piensen críticamente, comprendan y expliquen situaciones desde diversas áreas del saber…" (SEP, 2011b, p. 26) es una preocupación constante que se ve reflejada en los propósitos del currículo, así como en las aportaciones de diferentes estudios que coinciden al señalar que el trabajo intelectual[1] y el pensamiento crítico son la base para que los estudiantes construyan nuevos aprendizajes y logren un entendimiento profundo (Newmann y Wehlage, 1993; Newmann, López y Bryk, 1998; Gulikers, Bastiaens y Kirschner, 2004; Reeves, 2011; McTighe y Wiggins, 2012; Aubuson *et al*., 2014).

Una enseñanza que busca favorecer aprendizajes profundos, por lo tanto, debe apoyarse en actividades de enseñanza con sentido, que impliquen desafíos intelectuales para los estudiantes, así como actividades que resulten relevantes y útiles de acuerdo con sus contextos (MacTighe, Seif y Wiggins, 2004; SEP, 2011a). Algunas de las características que este tipo de actividades comparten tienen que ver con plantear propósitos bien definidos, establecer diferentes niveles de exigencia cognitiva, promover el uso de diferentes herramientas o recursos de apoyo para el aprendizaje, promover la construcción de nuevos conocimientos, permitir múltiples métodos o estrategias de solución, entre otras características que se detallan en el capítulo II de este documento.

Autores como Newmann y Wehlage (1993); Newmann, Marks y Gamoran (1996), Newmann, López y Bryk (1998) y Newmann, Bryk y Nagaoka, (2001) sugieren considerar tres criterios para valorar la calidad de las actividades de enseñanza a partir del trabajo intelectual que promueven: a) *investigación disciplinada*, que se refiere a la puesta en práctica de conocimientos previos y procedimentales para desarrollar nuevos conocimientos; b) *construcción del conocimiento*, que implica producir el conocimiento más que reproducirlo y lograr un significado más que una memorización; y c) *valor más allá del contexto escolar*, que tiene que ver con el significado de los aprendizajes para los estudiantes en contextos diferentes al escolar.

Resultados de investigaciones empíricas sugieren que cuando el desafío intelectual es mayor en las actividades de enseñanza, los estudiantes suelen demostrar un mejor nivel en la construcción de conocimientos, así como en el desarrollo de una investigación disciplinada. Esto significa que los estudiantes se esfuerzan por recuperar conocimientos previos para aplicarlos y ajustarlos a diferentes situaciones, logrando de esta manera

[1] En este documento se entenderá como *trabajo intelectual* toda "actividad mental, no predominantemente perceptual, por la cual se aprende algún aspecto de un objeto o situación basada en el aprendizaje y experiencia pasada" (Reeves, 2011, p. 165).

desarrollar nuevos conocimientos (Newmann, Marks y Gamoran, 1996; Newmann, López y Bryk, 1998).

Los estudios realizados por Newmann y colaboradores, así como los criterios que han sugerido, se han convertido en un parteaguas para investigaciones que han propuesto las actividades de enseñanza como un potencializador de oportunidades de aprendizaje, así como de calidad de la enseñanza. Dichos estudios han adoptado diferentes enfoques de investigación, entre los que se encuentra la intervención (Avery y Freeman, 2002), el diseño de nuevos modelos de enseñanza (McTighe, Seif y Wiggins, 2004; Wiggins y McTighe, 2008), y estudios metodológicos para la obtención de información. Estos últimos a través del diseño de instrumentos que facilitan y favorecen la obtención de la información válida y confiable sobre la calidad de las actividades de enseñanza (Stein, Grover y Henningsen, 1996; Boston y Wolf, 2006; Aubusson, Burke, Shuck, Kearney y Frischknecht, 2014).

En México el acercamiento a la práctica de enseñanza se ha realizado en su mayoría a través de estudios a profundidad, en los que se han empleado entrevistas, cuestionarios con preguntas generales y respuestas abiertas, así como observaciones en aula para dar seguimiento a docentes por diferentes periodos de tiempo. Los estudios cuantitativos o a gran escala son escasos y recientes, algunos de éstos han optado por utilizar las bases de datos que instancias como el Instituto Nacional para la Evaluación de la Educación generan (López y Mota, 2003; Ávila *et al*., 2013; Martínez y Chávez, 2016).

En 2016 el departamento de Educación de la Universidad Autónoma de Aguascalientes (UAA) concluyó una revisión de literatura nacional con el propósito de identificar estudios sobre las prácticas de enseñanza-evaluación en educación básica. Los resultados mostraron que en el área de matemáticas las investigaciones se han centrado en la caracterización de las prácticas de enseñanza, las creencias y concepciones de los docentes, el análisis del dominio que tienen los docentes de los contenidos, y el aprovechamiento de los estudiantes. Por lo que resulta necesario impulsar una línea de trabajo a través del desarrollo de acercamientos e instrumentos que permitan obtener información de una calidad superior a la que han logrado obtener estudios previos (Martínez y Chávez, 2016).

A partir de la información antes mencionada, el propósito de este estudio fue diseñar un protocolo de observación para evaluar las actividades de enseñanza que los docentes de quinto grado de primaria proponen en la asignatura de matemáticas, y obtener evidencias de validez y confiabilidad del instrumento. Este es un estudio metodológico cuyo aporte es el desarrollo de un instrumento que permite estudiar un fenómeno educativo, así como sustentar juicios de valor que se establezcan a partir de la realidad observada con base en criterios e indicadores específicos (ver capítulo II de este documento).

Se pretende que el instrumento permita obtener información sobre algunas características de las actividades de enseñanza que de acuerdo con la teoría representan aspectos fundamentales para promover aprendizajes significativos. Es importante aclarar que los alcances del estudio no incluyen el reporte de información sobre el tipo de actividades de enseñanza que se llevan a cabo como parte de las prácticas docentes en la asignatura de matemáticas; sin embargo, se espera que su diseño y validación apoye el desarrollo de estudios futuros interesados en diferentes aspectos de las actividades de enseñanza en matemáticas.

Para lograr el propósito de este estudio se han establecido elementos teóricos que guían el objeto de estudio. En este caso, las actividades de enseñanza, entendidas como lo que hacen el docente y los estudiantes durante una clase a partir de las tareas matemáticas. Por lo tanto, para ilustrar lo que implica su análisis, a continuación, se señalan tres elementos teóricos que se han tomado como base: a) la enseñanza de las matemáticas como el aspecto más amplio de la práctica docente; b) las actividades de enseñanza que guían todo el proceso de enseñanza (y objeto de estudio en esta investigación); y c) las tareas matemáticas como unidad básica para desarrollo de las actividades de enseñanza y la enseñanza en general (Figura 1).

Figura 1. Esferas de interacción del objeto de estudio
Fuente: Diseño propio a partir de la revisión de literatura

De acuerdo con Hiebert y Grows (2007), la enseñanza de las matemáticas implica una interacción entre el docente y los estudiantes con base en ciertos contenidos matemáticos dirigidos a facilitar los aprendizajes. Al respecto, autores como Cohen, Raudenbush y Ball (2003), así como el National Council of Teachers of Mathematics [NCTM] (2015) señalan que dicha interacción involucra experiencias individuales y colaborativas que fomentan la habilidad de los estudiantes para dar sentido a las ideas matemáticas y su razonamiento; es decir, permite la enseñanza de las matemáticas y se convierte en un panorama que permite a los estudiantes emplear las tareas propuestas por el docente durante la clase.

La actividad de enseñanza tiene que ver con lo que hacen el docente y los estudiantes durante una clase a partir de las tareas matemáticas; es decir, se refiere a las acciones y procedimientos que los docentes realizan en el aula para facilitar la construcción de conocimientos de los estudiantes. Al respecto, Doyle (1984), Stein, Grover y Henningsen (1996) y, Stein y Smith (1998) señalan que una actividad de enseñanza se puede definir a partir de tres de sus componentes: a) los productos que los estudiantes esperan lograr; b) las operaciones que involucran el logro de dichos productos; y c) los recursos disponibles para los estudiantes durante el desarrollo de los productos.

Las tareas matemáticas, en cambio, representan “desde un conjunto de ejercicios rutinarios hasta problemas complejos y desafiantes que enfocan la atención de los estudiantes en una idea matemática particular” (NCTM, 2015, p. 19). Por lo tanto, constituyen la unidad básica a partir de la cual se articula la actividad de enseñanza y por consiguiente la enseñanza en general. No obstante, una tarea matemática no se convierte en actividad de enseñanza hasta que los estudiantes la realizan. Es decir, cuando se pasa de la consigna a la implementación (Griffin, 2009; Swan, 2015).

Bajo esta perspectiva, al referirse a las actividades de enseñanza en este estudio se estará hablando de un sistema de acciones en el que interactúan sujetos (docente y estudiantes), objetos (meta de aprendizaje) y medios (libros, cuadernos u otros), cuya unidad de análisis para captar los diferentes aspectos son las tareas matemáticas propuestas durante la clase o lección.

2. Oportunidades de aprendizaje

De acuerdo con Putnam, Lampert y Peterson (1990); Hiebert y Wearne (1993) y Reevens (2011), una manera de indagar sobre el tipo de aprendizajes que se promueven en el aula es analizar las acciones que realizan los docentes para llevar a cabo la práctica de enseñanza. Al respecto, las actividades de enseñanza juegan un papel fundamental, puesto que representan el tratamiento básico a través del cual se establece la exigencia del trabajo intelectual, las posibilidades de poner en práctica conocimientos previos y procedimientos que ayuden a los estudiantes a desarrollar nuevos conocimientos, así como habilidades intelectuales superiores (Hiebert *et al*., 1997; Aubuson *et al*., 2014 y Hart, Cowhy, Matsko y Sporte, 2015).

Considerando que toda oportunidad de aprendizaje se realiza a través del tipo de actividades de enseñanza y su propósito, diferentes autores señalan la importancia de considerar ciertas características que ayuden a identificar el tipo de actividades que se promueven en el aula, algunas de las cuales tienen que ver con aspectos como el nivel de exigencia cognitiva, contextualización de las situaciones planteadas, y las diferentes estrategias o formas de solución que sugieren a los estudiantes, entre otras características que se describen en el capítulo II de este documento (Newmann y Wehlage, 1993; Stein y Lane, 1996; Newmann, Marks y Gamoran, 1996; Newmann, López y Bryk, 1998; Gulikers, Bastiaens y Kirschner, 2004; Griffin, 2009; McTighe y Wiggins, 2012; Aubusson, Burke, Schuck, Kearney y Frischknecht, 2014; NCTM, 2015).

En la medida en que las actividades de enseñanza cumplan con características que favorezcan el desarrollo de habilidades del pensamiento y de aprendizajes profundo en general, se estarán ofreciendo a los estudiantes oportunidades de aprendizaje útiles y pertinentes, mediante las cuales puedan adquirir herramientas y logren desarrollar competencias que les permitan enfrentar con éxito los desafíos que impone la sociedad actual (Hiebert y Wearne,1993; SEP, 2011a).

Abordar el tema de la práctica docente a través de la valoración del tipo de actividades de enseñanza requiere de criterios e indicadores que permitan identificar sus características, de tal manera que se pueda dar cuenta de aspectos como el trabajo intelectual que las actividades promueven, el potencial para favorecer oportunidades de aprendizaje, analizar de forma general el tipo de instrucción que se promueve en el aula, entre otros elementos que se pueden relacionar con la calidad de la enseñanza.

El Plan de Estudios. Educación Básica 2011, incluyó una serie de sugerencias para llevar a cabo actividades de enseñanza que favorezcan los procesos de enseñanza y aprendizaje. La propuesta enfatizó la movilización de saberes, así como la integración de los aprendizajes esperados como referente de la planificación didáctica. Para esto la SEP (2011) a través del currículo nacional planteo preguntas guía que ayudan a los docentes a reflexionar sobre el propósito de las actividades de enseñanza y el sentido que éstas podrían tener para los estudiantes:

- ¿Qué situaciones resultarán interesantes y desafiantes para que los estudiantes indaguen, cuestionen, analicen, comprendan y reflexionen?
- ¿Cuál es el nivel de complejidad que se requiere para la actividad que se planteará y cuáles son los saberes que los estudiantes tienen?
- ¿Qué aspectos quedarán a cargo de los estudiantes y cuáles será necesario explicar para que puedan avanzar?
- ¿De qué manera pondrán en práctica la movilización de saberes para lograr los aprendizajes y qué desempeños los harán evidentes? (p. 28)

Además de las preguntas que guían la planificación didáctica, el Plan de Estudios. Educación Básica 2011 incluyó principios pedagógicos que enfatizan la importancia de lograr una práctica de enseñanza que tenga como objetivo:

- **Centrar la atención en los estudiantes y en sus procesos de aprendizaje:** para lo cual el docente debe buscar la manera de generar disposición y capacidad en los estudiantes para que sigan aprendiendo a lo largo de la vida, así como que desarrollen habilidades del pensamiento para solucionar problemas y piensen de forma crítica. En este caso, el docente deberá en principio, reconocer la diversidad de capacidades, estilos y ritmos de aprendizaje de sus estudiantes desde la particularidad de las diferentes situaciones y contextos.

- **Planificar para potenciar el aprendizaje:** implica organizar actividades de aprendizaje a partir de diferentes formas de trabajo que involucren situaciones y secuencias didácticas que representen desafíos intelectuales para los estudiantes. Esto requerirá de los conocimientos que el docente tenga sobre lo que aprenderán los estudiantes y cómo lo aprenderán, así como reconocer la capacidad que tienen sus estudiantes para enfrentar los problemas planteados, y lo significativo de los mismos de acuerdo al contexto.
- **Poner énfasis en el desarrollo de competencias, el logro de los estándares curriculares y los aprendizajes esperados:** se debe tener en cuenta los estándares y aprendizajes esperados señalados en el Plan de estudios, y utilizarlos como referentes para planear actividades que permitan a los estudiantes adquirir las herramientas necesarias para aplicar de forma eficiente los conocimientos y habilidades adquiridas.
- **Evaluar para aprender:** la evaluación se plantea como parte constitutiva de la enseñanza y de los aprendizajes de acuerdo con los principios de la evaluación formativa. En este sentido, se espera que el docente encuentre en la evaluación la manera de promover oportunidades de aprendizaje para sus estudiantes (retroalimentación) y a la vez realice ajustes a su práctica de enseñanza de acuerdo con las necesidades de los estudiantes (SEP, 2011a).

Las preguntas y los principios pedagógicos son coincidentes con diferentes modelos de enseñanza efectiva que promueven la implementación de actividades de enseñanza que favorezcan la construcción de conocimientos, la reflexión, la diversidad de métodos o estrategias de solución, el trabajo colaborativo, entre otros (Newmann, López y Bryk, 1998; Aubusson, Burke, Schuck, Kearney y Frischknecht, 2014; NCTM, 2015).

Con la intención de lograr la mejora de la calidad de la educación, el Sistema Educativo Mexicano implementó nuevas políticas para evaluar el desempeño docente, considerando que su papel y función, hasta ahora irremplazable, es una pieza clave para lograr este propósito. No obstante, para valorar su desempeño se requieren evidencias de la actuación del docente en el aula y en otros espacios escolares para compararlas con parámetros e indicadores que determinen una práctica de enseñanza eficaz, lo que después permitiría promover el desarrollo profesional docente, desarrollar modelos de evaluación formativa, así como modelos de asesoría y acompañamiento para profesores, mismos que se enfatizan en la reciente Ley General del Servicio Profesional Docente (INEE, 2015).

Ante este contexto, resulta importante el desarrollo de investigaciones metodológicas, como la que se propone con este estudio, y específicamente de instrumentos que favorezcan el aporte de información válida y confiable sobre las prácticas de enseñanza. Es decir, se requiere de “criterios, procedimientos e instrumentos actualizados o diseñados de acuerdo con las necesidades de calidad y equidad para regular, orientar y evaluar el desempeño docente” (SEP, 2013, p. 35).

Para reportar las diferentes etapas que se desarrollaron a partir de este estudio, el documento de tesis se ha organizado en seis capítulos. El capítulo 1 se refiere a las actividades de enseñanza: su concepto, características, y la propuesta para su evaluación.

El capítulo 2 ofrece una perspectiva general sobre diferentes acercamientos hacia la práctica de enseñanza identificados en la revisión de literatura, lo cuales se tomaron como referente para el diseño del instrumento. El capítulo 3 describe el proceso de construcción del protocolo de observación y la obtención y tratamiento de la información: proceso del trabajo de campo, aspectos éticos que se adoptaron en diferentes momentos del estudio, y las estrategias de atención a los datos recabados.

Los capítulos 4 y 5 presentan información sobre la prueba de funcionamiento del Protocolo de Observación de las Actividades de Enseñanza en Matemáticas (POAEM). El capítulo 6 reporta los resultados obtenidos al respecto. Además de los capítulos antes mencionados, se incluye un apartado con la discusión y conclusiones.

Capítulo I: Las actividades de enseñanza

Este capítulo presenta los fundamentos teóricos a partir de los cuales se han identificado características de las actividades de enseñanza que promueven aprendizajes profundos de acuerdo con modelos de enseñanza para las matemáticas. El capítulo se divide en tres apartados: el primero describe elementos teóricos para definir a las actividades de enseñanza; el segundo establece las características que se tomaron como base para el desarrollo del protocolo de observación; y, el tercero presenta la propuesta para la evaluación de las actividades de enseñanza a partir del POAEM.

1. La actividad de enseñanza

El trabajo más profundo dentro de la práctica de enseñanza es el diseño instruccional. A través de éste, los docentes determinan el tipo de pensamientos a promover de acuerdo con los contenidos, las habilidades intelectuales que pueden desarrollar, y lo que los estudiantes podrían lograr a largo plazo con los conocimientos construidos. Por lo tanto, es a partir de estas acciones e interacciones con los estudiantes durante la práctica docente que se puede determinar el tipo de actividades que se ofrece a los estudiantes como oportunidades de aprendizaje (Reeves, 2011).

La Teoría Histórico-Cultural de la Actividad (CHAT por sus siglas en inglés), cuyos fundamentos se centran en el trabajo del psicólogo Lev Vigotsky, señala que en toda actividad humana la interacción social es fundamental para desarrollar funciones psicológicas superiores a través de herramientas de aprendizaje o procesos mediados. En este sentido, la CHAT proporciona bases teóricas importantes para indagar sobre algunos aspectos de las actividades de enseñanza que los docentes sugieren a los estudiantes durante una clase.

A diferencia de otras nociones en las que una actividad de enseñanza se puede limitar a las tareas que se proponen durante una clase (Stein, Grover y Henningsen, 1996; y, Stein y Smith, 1998), en la CHAT una “actividad” se entiende como una serie de acciones y operaciones que realiza un sujeto sobre un objeto, por lo que involucra una estructura compleja de mediación y acción humana colectiva en la que intervienen diferentes elementos (Wolff y Yew, 2007).

La CHAT puede ser un referente para el análisis de las actividades de enseñanza a partir de algunos de los elementos que la conforman, sin perder de vista que esos elementos representan un todo, y que las funciones de cada uno no están desconectadas de los otros, sino que son parte de la función holística de una actividad. Para los fines de este estudio, se describirán los elementos que conforman las diferentes generaciones de la CHAT, ya que éstas se tomarán como un referente para entender algunos aspectos de las actividades en general y de las actividades de enseñanza en particular.

De acuerdo con Wolff y Yew (2007), la primera generación de la CHAT propone no diferenciar lo individual de lo social, ya que en los procesos cognitivos estos elementos no están separados de la actividad cultural sino que, por el contrario, se complementan. A partir de este supuesto, esta teoría toma como eje tres elementos para abordar los procesos cognitivos: a) los sujetos que participan en una actividad (alumnos, profesor u otros sujetos involucrados), b) los medios o instrumentos que se emplean para realizar la actividad, y c) el objeto o motivo para realizarla, tal como se representa en la siguiente figura (Daniels, 2003; Lucci, 2006; Jaworski, Goodchild, Eriksen y Daland, 2011).

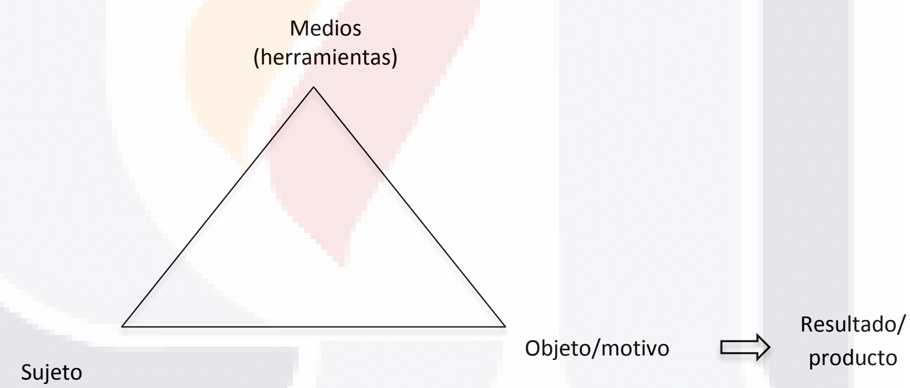


Figura 2. Triángulo de mediación simple
Fuente: Jaworski, Goodchild, Eriksen y Daland (2011). Mediating Mathematics Teaching Development and Pupil´s Mathematics Learning: The life cycle of a Task.

Con base en los elementos de la primera generación, se entiende que una actividad de enseñanza se logra a partir de la interacción que existe entre los sujetos (docente-alumnos, alumno-alumnos, alumnos-otros agentes involucrados) y los medios (libros, cuadernos, revistas, Internet, u otros materiales que apoyen el desarrollo de una actividad), para lograr un objeto o motivo, como es el aprendizaje (Wolff y Yew, 2007, y Jaworski, Goodchild, Eriksen y Daland, 2011).

La segunda generación de la CHAT desarrollada por Aleksandr Luria y A. N. Leontev toma como base los elementos de la primera generación (sujetos, medios y objeto) e incluye elementos relacionados con: las reglas de interacción social, la comunidad y su ambiente (por ejemplo un salón de clases), así como la división de labores en el trabajo colaborativo (el rol que cada sujeto debe tomar en la comunidad). Esta generación enfatiza lo relativo a la participación del sujeto como parte de una comunidad en la que existen responsabilidades. Esto es lo que se conoce como el triángulo de actividad, que se encuentra en la Figura 2 (Daniels, 2003; Wolff y Yew, 2007).

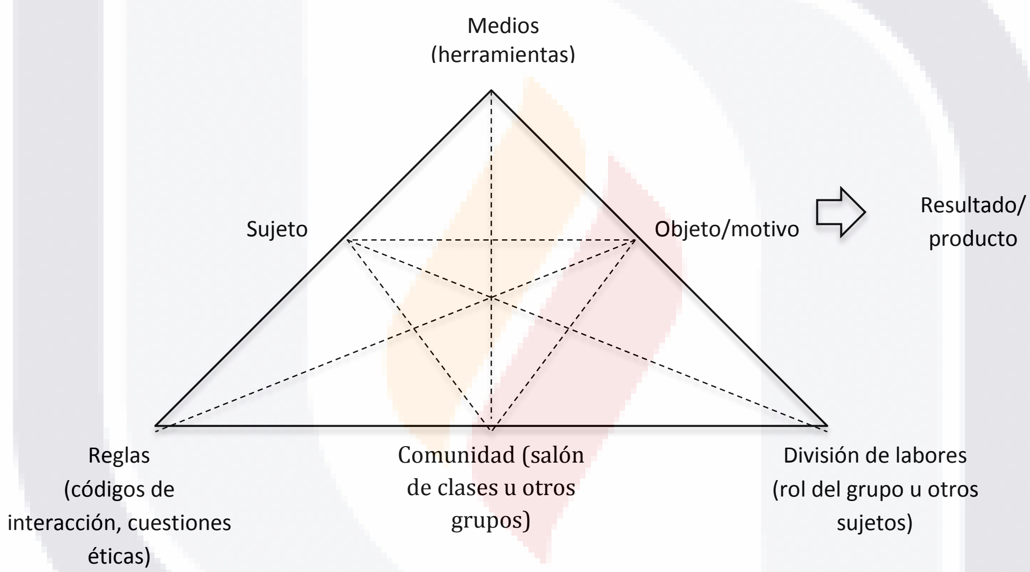


Figura 3. Triángulo de actividad
Fuente: Wolff y Yew (2007). Vigotsky's Neglected Legacy: Cultural Historical Activity Theory. Review of Educational Reseach.

En la segunda generación, el aprendizaje se entiende como un intercambio mutuo entre el sujeto y objeto durante el proceso de una actividad de enseñanza, ya que el aprendizaje se produce durante la expansión de las posibilidades de la acción de los sujetos en la búsqueda de objetos significativos. Por lo tanto, la acción mediada por instrumentos y orientada hacia un objetivo o motivo es una forma de analizar las operaciones que los docentes ejecutan para promover dichos aprendizajes (López, s.f.; Wolff y Yew, 2007).

Lo anterior hace énfasis en lo cognitivo, afectivo y social; aspectos fundamentales de la teoría sociocultural (o socio-histórica) de Vigotsky, a partir de la cual se pueden distinguir tres niveles de conductas: las actividades, que se refieren al objetivo o motivo que guía el comportamiento social y cultural; las acciones, que tienen que ver con la manera en que los sujetos se relacionan con el objeto de aprendizaje para lograr un objetivo; y las operaciones, que realizan los sujetos de forma independiente con base en los medios, condiciones o instrumentos disponibles (López, s.f.).

La tercera generación de la teoría amplía el modelo de la segunda generación hacia los diálogos y redes de sistemas, por lo que se entiende que todo sistema de actividad es parte de una red que en conjunto constituye la sociedad humana. Todos los sistemas de actividades son el resultado de procesos históricos culturales continuos en trabajos progresivos, en los que existe la división de labores entre los sujetos que conforman la sociedad. Se enfatiza la interacción entre dos o más sistemas de actividades como los que se describen en la segunda generación de esta teoría (ver figura 2), priorizando el rol del diálogo, las perspectivas múltiples, las contradicciones y las cuestiones de poder cuando se trata de interactuar en sistemas de actividad como las redes (Wolff y Yew, 2007).

Los elementos descritos en la segunda generación de esta teoría son los que se relacionan con las posturas de otros autores que de igual manera buscan enfatizar la importancia de la interacción social en el desarrollo de las actividades de enseñanza. Por ejemplo, Stein, Grover y Henninggsen (1996) y Boston y Smith (2011) señalan que una actividad de enseñanza comprende tres fases: los materiales curriculares como un medio a través del cual los docentes guían su práctica (herramientas en el caso de la CHAT); el tipo de actividades de enseñanza propuestas a los estudiantes (objetivo de la actividad); y la implementación de dichas actividades (o la actividad en sí, en el caso de la CHAT). En estas fases, la interacción se lleva a cabo entre los sujetos involucrados en la actividad (sujetos a las reglas y la división de labores), a partir de lo cual se obtienen resultados en relación con el objeto, que son los aprendizajes esperados (el producto). La siguiente figura resume el proceso descrito previamente.

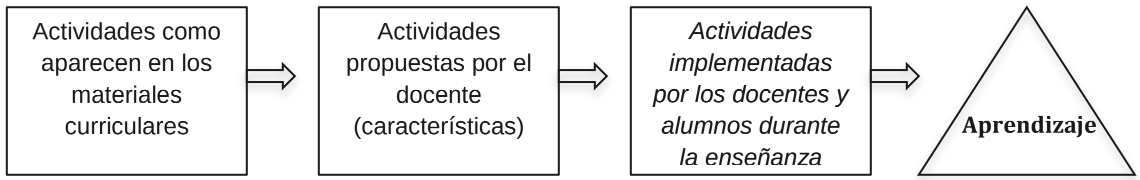


Figura 4. Marco de las actividades de enseñanza en matemáticas
Fuente: Boston y Smith (2011). A task-centric approach to professional development: enhancing and sustaining mathematics teacher's ability to implement cognitively challenging mathematical tasks. ZDM Mathematics Education.

Los elementos del sistema de actividad antes mencionado coinciden con los que sugiere el triángulo de mediación (figura 2), en el sentido de que en una actividad de enseñanza involucra la interacción de los sujetos participantes, docente y estudiantes, que están sujetos a las reglas y la división de labores de la comunidad, los medios o herramientas que se seleccionan para guiar o apoyar el aprendizaje, y el objeto de la actividad, que es el aprendizaje.

Para los propósitos de este estudio, se entenderá como una actividad de enseñanza las acciones que se llevan a cabo durante una clase para lograr un objetivo o aprendizaje. Esto involucra la interacción entre los sujetos (docente-alumnos, alumno-alumnos, alumnos-otros agentes involucrados) y los medios (libros, cuadernos, consignas, revistas, Internet u otros materiales que apoyen el desarrollo de una actividad) para lograr un objetivo: el aprendizaje. No obstante, en este estudio no se pretende dar cuenta sobre el aprendizaje logrado por los estudiantes, o de las características de las actividades de enseñanza que los docentes promueven en el aula, sino aportar evidencias de la validez de la medición con el instrumento diseñado.

2. Características de las actividades de enseñanza en matemáticas

La actividad de enseñanza es la acción directa a través de la cual se promueve cierto tipo de aprendizajes; por lo tanto, es un indicador de las oportunidades de aprendizaje que los docentes ofrecen a los estudiantes durante una clase o lección. Depende de las acciones generadas en el aula el que los estudiantes puedan construir significados y conocimientos de acuerdo con sus propias experiencias (Donovan y Bradsford, 2005; Boston y Wolf, 2006).

Newmann, López y Bryk (1998), Gulikers, Bastiaens y Kirschner (2004), Boston y Wolf, (2006) y Aubusson, Burke, Shuck y Kearney (2014) sugieren favorecer acciones que originen aprendizajes profundos, esto es actividades de enseñanza para promover el desarrollo de habilidades del pensamiento que permitan a los estudiantes aplicar procesos de aprendizaje con flexibilidad y creatividad. Es decir, se requieren condiciones propicias para que los estudiantes encuentren en el aprendizaje de las matemáticas diversas oportunidades para comprometerse de manera activa con el razonamiento, logren sus propios significados y desarrollen una comprensión profunda de la asignatura.

Santos (2003), Llinares (2008) y Verschaffel, Greer y De Corte (2010) identifican tres condiciones fundamentales para favorecer aprendizajes profundos en la enseñanza para las matemáticas: a) el tipo de actividades que se promueven para favorecer los aprendizajes; b) las herramientas que los docentes utilizan para representar las ideas y situaciones problemáticas; y c) los argumentos que docentes y estudiantes utilizan para justificar las conjeturas matemáticas.

En la literatura se han identificado diferentes aspectos que coinciden con las condiciones antes mencionadas. Por ejemplo, el NCTM señala que “el aprendizaje de las matemáticas por parte de los estudiantes depende fundamentalmente de lo acontecido en el salón de clases” (2015, p. 8); esto es, del tipo de trabajo intelectual propuesto por los docentes a través de las actividades de enseñanza y de las oportunidades que éstas representan para la construcción de nuevos conocimientos, para estimular el razonamiento e involucrar a los estudiantes de manera significativa con las matemáticas (Picaroni y Loureiro, 2010; Zolkower y Bressan, 2012).

Algunas de las características que diferentes autores sugieren considerar para el desarrollo de las actividades de enseñanza en matemáticas incluyen: a) las metas o propósitos de aprendizaje que guiarán dichas actividades; b) el nivel de complejidad cognitiva que implicarán para los estudiantes; c) la manera en que favorecerán la construcción de conocimientos; d) la diversidad de estrategias de solución; e) el tipo de preguntas que se plantean para fortalecer el aprendizaje; f) la cercanía que las actividades establecen con el contexto de los estudiantes; g) el uso de representaciones matemáticas; y f) el trabajo colaborativo (Wolf, 2006; Llinares, 2008; Picaroni y Loureiro, 2010; Zolkower y Bressan, 2012; Siller *et al*., 2014; NCTM, 2015, y Swan, 2015); sin embargo, dichas características también son aplicables a otras asignaturas, ya que diferentes estudios centrados en la enseñanza en general sugieren implementar características coincidentes con las propuestas anteriormente.

Modelos de enseñanza como la Pedagogía Auténtica, la Enseñanza para la Comprensión y el Modelo de Entendimiento por Diseño, al igual que en el caso de la enseñanza para las matemáticas, sugieren características clave de las actividades de enseñanza, como: a) comunicar claramente los propósitos de aprendizaje; b) ofrecer oportunidades para enfrentar diferentes niveles de exigencia cognitiva; c) promover la construcción de conocimientos a partir de los contenidos o conocimientos previos; d) promover alternativas de solución; e) contextualizar las actividades de enseñanza para que los estudiantes puedan darles sentido a partir de sus propias experiencias; f) promover una investigación disciplinada; g) favorecer el trabajo colaborativo, así como la comunicación de resultados, y otorgar tiempos para una autoevaluación por parte de los estudiantes (Newmann, López y Bryk, 1998; Wiggins,1998; Nicaise, Gibney y Crane, 2000; Oura, 2001; Avery y Freeman, 2002; Lesh y Harel, 2003; Gulikers, Bastiaens y Kirschner, 2004; y Aubusson, Burke, Shuck y Kearney, 2014).

Con base en las características identificadas durante la revisión de literatura, a continuación se presenta la propuesta conceptual a partir de la cual se establecieron siete dimensiones para valorar el constructo “actividades de enseñanza”. La selección de las dimensiones se realizó tomando como base las coincidencias en los criterios propuestos en diferentes modelos de enseñanza para las matemáticas y en general (más de dos asignaturas). La información que se describe en cada punto se centra en una agrupación de criterios de acuerdo con la naturaleza y fundamentos teóricos de las actividades de enseñanza propuestas en los modelos de enseñanza revisados.

a. Propósito o meta de aprendizaje

La meta o propósito de una actividad de enseñanza representa la guía a partir de la cual los docentes deben tomar las decisiones durante la práctica de enseñanza. Dependerá del propósito que se siga durante la clase o lección que los estudiantes tendrán oportunidad de reflexionar sobre los resultados que deben lograr; por lo tanto, las metas de aprendizaje no deberían constituir un misterio y, por el contrario, deberían ser claras para los estudiantes (Wiliam, 2001; Hiebert *et al.,* 2007, y el NCTM, 2015) .

Grabinger y Dunlap (1995); Newmann, López y Bryk (1998); Oura (2001) y Stone (2008) enfatizan la importancia de establecer propósitos de aprendizaje claros durante las actividades de enseñanza, ya que éstos ayudan tanto al docente como a los estudiantes a visualizar qué es lo que deben saber y poder hacer después de trabajar con ciertos contenidos. Por el contrario, si no se establecen propósitos claros, resulta complejo para los docentes identificar si la actividad de enseñanza es apropiada para lograr los aprendizajes esperados.

Wiggins (1998) y Ravela (2009) sugieren utilizar los propósitos de aprendizaje como guía para desarrollar las actividades de enseñanza, de tal manera que en todo momento docentes y estudiantes tengan en cuenta la meta o producto que se pretende alcanzar durante una clase. Lo que se busca es que a partir del propósito los estudiantes logren dar sentido a las propuestas de contenido (por ejemplo, las consignas, la manera de trabajo, u otros aspectos importantes) durante el desarrollo de una actividad.

b. Demanda cognitiva

Para determinar el nivel de demanda cognitiva de las actividades de enseñanza, se han desarrollado diferentes taxonomías que toman como base los objetivos de aprendizaje. Por ejemplo, la taxonomía de Bloom (1956), Perkins (1992), Anderson *et al.* (2001) y Stiggins, Arter y Chappuis (2007) establece niveles de desempeño en los cuales el nivel más bajo tiene que ver con aprender conocimientos, hechos o procesos; el nivel medio-bajo se refiere al desarrollo de habilidades del pensamiento utilizando el conocimiento para tomar decisiones; el nivel medio-alto implica fundamentar ideas o explicar causas; y el nivel alto tiene que ver con discutir resultados y establecer nuevas hipótesis, cuestionar o refutar supuestos; sin embargo, lo que resulta sustancial destacar a partir de este tipo de taxonomías es que independientemente del marco de referencia que los docentes adopten, es importante distinguir entre los objetivos que favorecen un aprendizaje superficial (conocimientos aislados o de repetición sin sentido para los estudiantes), de aquellos que ayudan a promover un aprendizaje profundo (por ejemplo, analizar, explicar, justificar procesos). Por lo tanto, en este estudio se adoptará la taxonomía desarrollada por Stein y Smith (1998), ya que además de ser compatible con los planteamientos de otras taxonomías, se ha diseñado específicamente para aportar información sobre el tipo de tareas matemáticas que conforman la actividad de enseñanza.

La taxonomía de Stein y Smith (1998) comprende cuatro niveles de desempeño de menor a mayor demanda cognitiva, los cuales se describen a continuación:

- **Memorización** (bajo nivel): involucra la reproducción de hechos, reglas, fórmulas o definiciones previamente aprendidos o definidos, se refiere a actividades que requieren la reproducción de material visto con antelación.
- **Procedimientos sin conexiones** (bajo nivel): se centran en la descripción de los procedimientos utilizados, promueven procedimientos específicos y se enfocan en generar respuestas correctas, en lugar de desarrollar la comprensión matemática.
- **Procedimientos con conexiones** (alto nivel): utilizan procedimientos muy generales con el propósito de desarrollar niveles más profundos de comprensión de conceptos e ideas matemáticas. Los alumnos requieren involucrarse con ideas conceptuales que subyacen en los procedimientos para lograr una comprensión.
- **Construcción de las matemáticas** (alto nivel): requieren que los estudiantes exploren y entiendan la naturaleza de los conceptos matemáticos, así como los procesos necesarios para enfrentar la situación problemática, de tal manera que los puedan explicar y justificar (Stein y Smith,1998).

c. Recursos empleados durante la actividad de enseñanza

De acuerdo con el NCTM (2015), el uso de recursos para apoyar la enseñanza de las matemáticas puede ser un auxiliar efectivo para que los estudiantes encuentren sentido a las ideas matemáticas, para razonar matemáticamente, así como para comunicar su pensamiento matemático; sin embargo, tales recursos dependerán del uso que el docente y los estudiantes hagan de materiales disponibles para promover o fortalecer los aprendizajes matemáticos.

Los recursos que comúnmente se emplean durante las actividades de enseñanza pueden comprender artefactos convencionales como el libro de texto, hojas de trabajo, *software*, entre otros, así como el propio conocimiento (el cual no se incluirá para los fines de este estudio). En lo que respecta a los artefactos, cabe señalar que éstos se podrán considerar como recursos sólo cuando el uso que se les da aporte de alguna manera al propósito de la actividad de enseñanza (Gueudet y Trouche, 2008; Cohen, Raudenbush y Ball, 2003). Es decir, es posible que incluso cuando se cuente con algún tipo de artefacto, éste se utilice para fines distintos al propósito de aprendizaje matemático.

De esta manera, es preciso prestar atención a cómo los docentes (y estudiantes) relacionan el uso de algún artefacto con el contenido matemático, de tal manera que lo conviertan en un recurso para apoyar el aprendizaje matemático.

d. Alternativas de solución

La libertad que se ofrece a los estudiantes para buscar y emplear diferentes estrategias o procedimientos para resolver un problema matemático les permite desarrollar autonomía para analizar, comparar, ajustar y seleccionar sus propias estrategias o procedimientos para resolver una tarea matemática (Kramarski, Mevarech y Arami, 2002; Griffin, 2009; Picaroni y Loureiro, 2010 y NCTM, 2015).

De acuerdo con Stein y Smith (1998), Jonassen (2000 y 2011) y el NCTM (2015) los estudiantes que experimentan con diferentes procedimientos y métodos de solución suelen generar con mayor facilidad una conciencia crítica y de razonamiento matemático que les permite considerar restricciones y alternativas en estrategias de solución. Por lo tanto, cuando las actividades de enseñanza promueven múltiples posibilidades de solución también sugieren la justificación del razonamiento matemático, así como del desarrollo de argumentos y contraargumentos.

Wiggins (1998), Kramarski, Mevarech y Arami (2002) y Griffin (2009) enfatizan la importancia de promover tareas matemáticas que permitan a los estudiantes llevar a cabo múltiples posibilidades de solución, en el sentido de dar libertad a los estudiantes para considerar diferentes vías de solución que posteriormente se conviertan en un motor para lograr nuevos aprendizajes al comentar y comparar respuestas de los estudiantes en el trascurso de una clase.

e. Preguntas para la reflexión

Las preguntas para la reflexión son oportunidades que se da a los estudiantes para desarrollar y justificar su propio aprendizaje a partir del pensamiento crítico y reflexivo sobre algún tema o contenido específico. Particularmente promoviendo la reflexión de los estudiantes a través de preguntas deliberadas que les aporten elementos que los ayuden a desarrollar la capacidad de pensar y actuar con flexibilidad a partir de su propio pensamiento y de lo que ya saben, como un puente para la construcción de nuevos conocimientos (Putnam, Lampert y Peterson, 1990; Hiebert y Wearne, 1993; Hiebert *et al.,* 1997 y Donovan y Bransford, 2005; Barreiro y Casetta, 2012).

Schoenfeld (2002), Stain *et al.* (2008) y Jonassen (2011) sugieren plantear preguntas (focalizadas en el contenido o tema tratado) que inciten a los estudiantes al análisis y razonamiento sobre las respuestas, procedimientos, métodos, u otros elementos que se deban emplear para resolver un problema matemático. Además, enfatizan la importancia de que docentes o estudiantes elaboren preguntas que sugieran establecer relaciones entre las ideas clave de contenido abordado y los propósitos de la actividad propuesta.

El NCTM (2015) y Smith (2000) sugieren el planteamiento de preguntas durante la actividad de enseñanza, preguntas que permitan a los estudiantes explicar su razonamiento y aprovechar dicho razonamiento para que se generen nuevas preguntas que permitan fortalecer el propósito de la actividad de enseñanza. Por ejemplo, algunas preguntas que se podrían plantear son las siguientes:

- ¿Por qué ésta es la mejor solución?
- ¿Qué evidencia soporta esas razones?
- ¿Qué soluciones podrían recomendar otros compañeros?
- ¿Qué razones podrían mencionar otros compañeros para soportar sus respuestas?
- ¿Qué evidencia tendrían que aportar? (2011, p. 108)

Lo importante en este aspecto es la comunicación del pensamiento de los propios estudiantes mientras se llevan a cabo las tareas matemáticas que conforman la actividad de enseñanza, ya que de esta manera se posibilita a los estudiantes para lograr un mejor avance para su aprendizaje.

f. Contextualización

Boston y Wolf (2006), Picaroni y Loureiro (2010) y Zolkower y Bressan (2012) señalan que en la enseñanza de las matemáticas es importante adoptar tareas matemáticas desafiantes y originales que se enmarquen dentro de contextos plausibles, de tal manera que los estudiantes puedan reconocer en éstas un significado personal o social, así como su aplicación en algún contexto. Es decir, producir tareas matemáticas que resulten realizables, razonables o imaginables para los estudiantes (Perea y Valdemoros, 2009 y Ravela, 2009).

Newmann, Marks y Gamoran (1996), Nicaise, Gibney y Crane (2000), McTighe y Wiggins (2012), Aubuson *et al*. (2014) sugieren que las tareas matemáticas que se generen durante el proceso de actividad de enseñanza se refieran a situaciones desafiantes como las que las personas comúnmente enfrentan en la vida diaria y en diferentes campos de trabajo, con el propósito de mostrar a los estudiantes escenarios reales en los que el conocimiento adquirido en la escuela es aplicable.

Otros autores como Picaroni y Loureiro (2010) señalan que los escenarios de las tareas matemáticas pueden enmarcarse a partir de diferentes elementos y situaciones durante el desarrollo de la práctica de enseñanza; por lo cual, a partir de información recabada en el aula han establecido un marco de referencia para identificar el tipo de contextualización de las tareas matemáticas que se podría encontrar en la actividad de enseñanza.

Las autoras agrupan cuatro tipos de contextualización que se podrían identificar a partir de las actividades de enseñanza en matemáticas, dicha clasificación va desde un significado personal para los estudiantes, hasta actividades que no sugieren (o requieren un contexto) para relacionar el problema matemático. A continuación, se describen los cuatro niveles de contexto sugeridos por las autoras:

- **Contexto significativo**: los estudiantes pueden encontrar un significado personal o social del conocimiento puesto en juego.
- **Contexto intermatemático**: se refiere a una problematización del conocimiento relacionado con la propia disciplina.
- **Contexto escolar**: se representan situaciones que sólo tienen significado en el ámbito educativo.

- **Actividades de enseñanza sin contexto**: el conocimiento se presenta aislado de las situaciones que podrían darle sentido (Picaroni y Loureiro, 2010).

Además de la contextualización de las actividades de enseñanza sugerida por diferentes autores, el NCTM (2015) promueve el uso y vinculación de representaciones matemáticas que favorezcan el establecimiento de conexiones entre conocimientos matemáticos que ayuden a profundizar el entendimiento de conceptos y procedimientos para la resolución de problemas. Esto tiene que ver con que los estudiantes puedan analizar y reconocer las situaciones problemáticas y establecer conexiones entre las representaciones e ideas matemáticas previas.

En los diferentes casos, la idea que prevalece es la importancia de que los estudiantes comprendan la importancia o aplicación de los conocimientos matemáticos en sus diferentes contextos, lo que incluye el escolar (en la mayoría de los casos), pero sin que sea éste el único contexto o aplicación que se favorezca durante su aprendizaje. Es importante que en la medida de lo posible los estudiantes relacionen los aprendizajes matemáticos con su aplicación más allá de la escuela, ya sea en situaciones cotidianas, de trabajo profesional, o en otros campos de la educación (asignaturas) (Ravela, 2009; McThighe y Wiggins, 2012; y Aubusson, Burke, Shuck, Kearney y Frischknecht, 2014).

g. Trabajo colaborativo

Los seres humanos somos personas sociales que estimulan el aprendizaje mediante la comunicación e intercambio de ideas (Gavilán y Alario, 2010); por lo tanto, una enseñanza que favorezca el trabajo colaborativo entre los estudiantes estará potencializando el aprendizaje que los estudiantes puedan lograr, ya que de acuerdo con diferentes autores, esta forma de trabajo motiva y ayuda a la evolución de las ideas, estrategias y el razonamiento en general (NCTM, 2001; Trigueros, 2009; Verschaffel, Greer y De Corte, 2010 y Barreiro y Casetta, 2012).

A partir del trabajo colaborativo, los estudiantes pueden generar un esfuerzo productivo para lograr ciertos aprendizajes, puesto que al discutir con los pares es posible generar un mejor análisis de la información (desde diferentes perspectivas), comparar ideas e inclusive desarrollar un discurso matemático significativo, con lo cual puedan construir una comprensión compartida (NCTM, 2015). Es preciso diferenciar entre el trabajo colaborativo y la organización que se puede generar dentro de un grupo de trabajo. Este último, por ejemplo, podría implicar el trabajo individual de sus integrantes y compartir espacios para después señalar procesos y resultados obtenidos en cada caso.

Para favorecer el trabajo colaborativo, los docentes deben permitir a los estudiantes construir conocimientos en conjunto, integrando ideas y propuestas de los diferentes participantes; respetando, ante todo, el pensamiento de los demás, y asegurando a la vez que las ideas matemáticas mantengan un lugar primordial en las discusiones que se lleven a cabo (Fernandez y Yoshida, 2004; SEP; 2011a y NCTM, 2015).

De acuerdo con Newmann y Wehlage (1993) y Jaworski, Goodchild, Eriksen y Daland (2011), otros elementos a considerar tienen que ver con la interacción que se promueve entre los propios estudiantes, y entre el docente y estudiantes con las herramientas que se utilizan para lograr un objetivo; así como con la promoción de un ambiente de respeto mutuo entre todos los miembros de la clase.

Se hace énfasis en la importancia de permitir a los estudiantes justificar sus respuestas, hacer distinciones, generalizar ideas, plantear preguntas como intercambios no controlados sino espontáneos, y establecer diálogos entre los participantes. Finalmente, se sugiere que los estudiantes tengan oportunidad de explicar su propio pensamiento, indagar nuevas ideas, y responder a cuestionamientos de sus pares para, en conjunto, construir un entendimiento colectivo (Newmann y Wehlage, 1993).

3. Las actividades de enseñanza en matemáticas: una propuesta para su evaluación

Teniendo en cuenta que una actividad de enseñanza es un sistema que integra diferentes elementos y características que se pueden identificar a partir de las interacciones que existe entre los sujetos: docente-alumnos, alumno-alumno, en este apartado se presenta el modelo teórico a partir del cual se llevará a cabo la valoración de las actividades de enseñanza en quinto grado de primaria en la asignatura de matemáticas.

Las coincidencias encontradas en diferentes estudios sobre las actividades de enseñanza para las matemáticas y la enseñanza en general son los elementos básicos a partir de los cuales se han establecido las dimensiones de referencia para el diseño del protocolo de observación. La información que se buscará obtener a partir de la observación de una clase completa de matemáticas serán los elementos de las tareas matemáticas que conforman la actividad de enseñanza, específicamente:

a. La claridad de la tarea matemática, entendiendo ésta como la forma en que el docente comunica el trabajo que los estudiantes deben desarrollar y cómo indaga si han comprendido lo que se les solicita como trabajo o producto. Esta dimensión se ha incluido ya que, de acuerdo con autores como Newmann, López y Bryk (1998); Oura (2001) y Stone (2008) y el NCTM (2015), es un paso fundamental para llevar a cabo una enseñanza efectiva. Es decir, se requiere tener un propósito claro sobre lo que implicará el trabajo a desarrollar durante la clase (tanto para el docente como para los estudiantes).

b. La exigencia cognitiva de las tareas matemáticas, con la que se identificará el nivel de pensamiento requerido para resolver los problemas matemáticos. Se tomará como base la taxonomía de tareas matemáticas desarrollada por Stein, Grover y Henningsen (1996), porque permite distinguir entre aquellas tareas que favorecen un aprendizaje superficial (conocimientos aislados o de repetición sin sentido para los estudiantes) y aquellas que ayudan a promover un aprendizaje profundo (por ejemplo, analizar, explicar, justificar procesos).

c. Los recursos empleados durante la actividad de enseñanza, para lo que se tendrán en cuenta los artefactos convencionales como el libro de texto, hojas de trabajo, software, entre otros, y el uso que se hace de éstos para convertirlos en recursos para promover el aprendizaje de las matemáticas.

d. Las alternativas de solución que el docente promueve durante la clase para que los estudiantes desarrollen una variedad de estrategias o procedimientos para resolver un problema matemático, específicamente prestando atención a las oportunidades que el docente ofrece a los estudiantes para emplear múltiples posibilidades de solución, sin sesgar el proceso de pensamiento de los estudiantes.

e. Las preguntas para la reflexión. Se buscará establecer criterios para valorar el tipo de preguntas que se plantean durante la clase (ya sea por el docente o los estudiantes) para dar sentido a ideas o relaciones matemáticas importantes.

f. La conexión/aplicación de los conocimientos matemáticos implicará el análisis de la situaciones problemáticas propuestas durante el desarrollo de la actividad de enseñanza, específicamente la relación que las tareas matemáticas sugieren en relación con otras disciplinas o situaciones de la vida cotidiana.

g. El trabajo colaborativo que se promueve durante el desarrollo de la actividad de enseñanza, específicamente la interacción, negociación y cooperación entre estudiantes para lograr un propósito común.

Con base en los aspectos que busca medir cada dimensión, éstas se agruparon en cuatro dominios. El primero se refiere al potencial de la actividad de enseñanza, lo cual está relacionado con los contenidos matemáticos. Este dominio incluye tres dimensiones: claridad de la tarea matemática, exigencia cognitiva, y los recursos empleados para desarrollarla.

El segundo domino se refiere a elaboración de comunicaciones matemáticas, por lo que se centra en las interacciones entre docentes y estudiantes. En este caso se incluyeron dos dimensiones: alternativas para la solución y preguntas para la reflexión.

El tercer dominio tiene que ver la contextualización de la tarea matemática, lo cual se relaciona con la conexión de las tareas con otras disciplinas o aspectos de la vida cotidiana. Este dominio incluye una dimensión: conexión o aplicación de los conocimientos matemáticos.

El cuarto dominio se refiere al trabajo colaborativo, específicamente a las estrategias de trabajo que se generan durante una tarea matemática. Esto es diferente a la organización de los estudiantes, ya que se busca identificar la cooperación para generar una meta o producto en conjunto. Este dominio incluye una dimensión: trabajo colaborativo. La siguiente figura presenta los dominios y dimensiones que conforman el modelo teórico para el diseño del Protocolo de Observación de las Actividades de Enseñanza en Matemáticas.

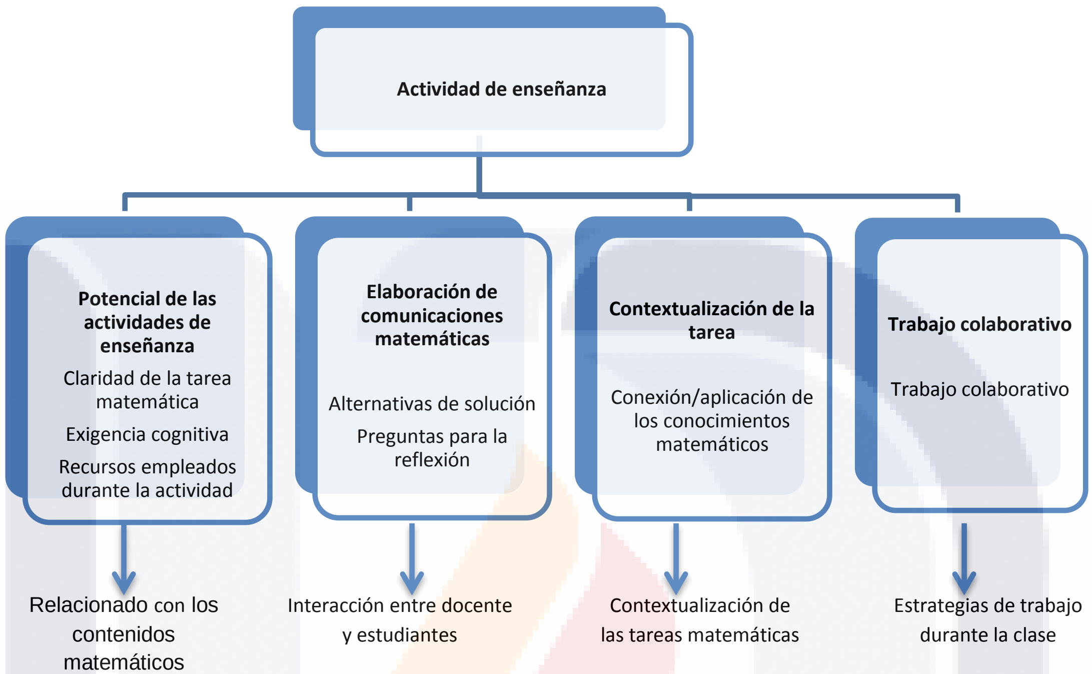


Figura 5. Modelo teórico para evaluar las actividades de enseñanza en matemáticas
Fuente: Diseño propio a partir de la revisión de literatura.

Para evaluar cada una de las dimensiones contenidas en el modelo teórico propuesto, cada dimensión se desarrolló a manera de rúbrica en las que se incluyen los criterios a identificar de acuerdo con cuatro niveles de desempeño en cada caso:

- **Claridad de la tarea matemática:** establece indicadores para identificar la manera en que el docente comunica las tareas matemáticas para llevar a cabo la actividad de enseñanza a través de: los productos o aprendizajes esperados, el lenguaje matemático, y las preguntas o estrategias para favorecer la comprensión de los estudiantes sobre la tarea matemática propuesta (Stein, Grover y Henningsen, 1996; Hiebert, 2003; Boston y Wolf, 2006).

- **Exigencia cognitiva:** Incluye indicadores para identificar tareas matemáticas que favorecen un aprendizaje superficial (conocimientos aislados o de repetición sin sentido para los estudiantes), así como aquellas que favorecen un aprendizaje profundo (por ejemplo, tareas que requieren analizar, explicar, justificar procesos, etc.).

- **Recursos empleados:** Se centra en el uso que el docente y los estudiantes hacen de artefactos y materiales disponibles para convertirlos en recursos que apoyen el aprendizaje de las matemáticas. Los indicadores en esta rúbrica tienen que ver con la manera en que el recurso se relaciona con el contenido que se está trabajando; y cómo el recurso puede estimular a los estudiantes para lograr ciertos desempeños.

- **Alternativas de solución:** Incluye indicadores para identificar las oportunidades que el docente ofrece a los estudiantes para emplear diferentes estrategias o procedimientos de solución, así como para identificar las pistas o sugerencias que pueden sesgar las respuestas de los estudiantes.

- **Preguntas para la reflexión:** Se focaliza en tres indicadores: a) las preguntas que solicitan a los estudiantes describir información; b) preguntas que ayudan a hacer evidente el pensamiento matemático (justificación); y c) preguntas que implican justificar y demostrar el trabajo realizado (NCTM, 2015)

- **Conexión/aplicación de los conocimientos matemáticos:** Los indicadores se focalizan en la clasificación de contextos de tareas matemáticas propuesta por Picaroni y Loureiro (2010), por lo que se busca identificar: experiencias que los estudiantes puedan reconocer por su situación o contexto, la conexión de las situaciones problemáticas con aspectos del mundo que nos rodea; y la relación de ideas o conceptos matemáticos con otros temas o disciplinas.

- **Trabajo colaborativo:** Incluye indicadores relacionados con: el rol activo del docente para promover la interacción de los estudiantes y lograr un aprendizaje o producto común, el tipo de cooperación entre los estudiantes (integración de ideas y propuestas de los participantes), y la negociación que se realiza para llegar a acuerdos comunes sobre ideas, conceptos o procedimientos matemáticos.

Capítulo II: Protocolos e instrumentos para valorar las prácticas de enseñanza en matemáticas

Este capítulo ofrece una perspectiva general sobre los diferentes acercamientos para la medición de la práctica de enseñanza identificados en la revisión de literatura. El primer apartado describe algunos protocolos de observación desarrollados a partir del proyecto MET como un referente en el diseño de instrumentos para medir la práctica efectiva de enseñanza. El segundo apartado describe otros instrumentos con los que se han medido diferentes aspectos de la práctica de enseñanza que no se incluyen en el proyecto MET.

1. Protocolos de observación para medir la práctica de enseñanza en matemáticas

Considerando que las prácticas de enseñanza son un objeto de estudio complejo, en la literatura es posible encontrar diferentes acercamientos que permiten recabar información al respecto. Estos acercamientos generalmente se identifican en tres tipos: información dada por los sujetos (cuestionarios y escalas, autorreportes, bitácoras y diarios, entrevistas, etc.); instrumentos que analizan productos para dar cuenta de dichas prácticas, como los portafolios o el análisis de tareas; y protocolos de observación, que permiten observar situaciones pedagógicas para su análisis (Martínez Rizo, 2013).

Los protocolos de observación se han identificado como uno de los acercamientos más utilizados para obtener información sobre las prácticas docentes (Martínez Rizo, 2013). Al respecto, se puede mencionar el proyecto de Medición de la Efectividad de la Enseñanza (MET por sus siglas en inglés), en el que se aplicaron diferentes protocolos de observación para medir la efectividad de la enseñanza. Dicho proyecto se sustentó en un estudio meticuloso con el que se buscó analizar las medidas de efectividad docente para establecer las características reales de las prácticas efectivas de enseñanza y aportar conocimientos útiles sobre el desempeño docente y el desarrollo profesional (Met Project, 2010a).

El proceso de diseño e implementación de los diferentes protocolos de observación desarrollados a partir del proyecto MET son uno de los referentes más claros para nuevos estudios metodológicos como el que se presenta en este documento. A través del MET se ha documentado el seguimiento de diferentes protocolos de observación e instrumentos que permiten obtener información válida y confiable para identificar prácticas efectivas de enseñanza y establecer referentes para ello. A continuación se describen diferentes protocolos de observación incluidos en dicho proyecto:

Framework For Teaching Protocol (FTT), busca apoyar el desarrollo profesional, la evaluación y tutoría docente. Este protocolo establece cuatro dimensiones: preparación, ambiente de la clase, instrucción/enseñanza y responsabilidades profesionales. A diferencia de otros protocolos de observación, el FFT utiliza sólo una escala para establecer la valoración de la práctica de enseñanza (Met Project, 2010b).

El protocolo **Mathematical Quality of Instruction** (Sistema MQI), diseñado para medir la calidad de las interacciones en las clases de matemáticas entre el docente, los estudiantes y el contenido. Incluye cinco dimensiones: riqueza de las matemáticas; errores e imprecisiones; trabajo con los estudiantes y las matemáticas; participación de los estudiantes en el sentido de decisiones y el razonamiento; y las conexiones entre el trabajo en clase y las matemáticas (Hill, Kapitula y Umland, 2011; Met Project, 2010c).

Del mismo modo, se utilizó el **Classroom Assessment Scoring System** (sistema CLASS), que mide la enseñanza a través de un modelo de interacciones en el aula. El foco de atención son las interacciones entre docentes y alumnos. La medición general de los atributos de enseñanza se evalúan a través de tres dominios: el apoyo emocional, la organización de salón de clase y el apoyo instruccional (Pianta, La Paro y Hamre, 2008; Met Project, 2010e).

2. Instrumentos para medir las actividades de enseñanza en matemáticas

En la literatura se han identificado instrumentos de medición de la práctica docente diferentes a los protocolos de observación, estos permiten recabar información a través de cuestionarios, viñetas, evidencias de trabajo y videograbación de clases como insumos para ofrecen información sobre diferentes aspectos de las *actividades de enseñanza*. A continuación se describen algunos de estos instrumentos:

Instrument for Ranking Classes (IRC). Es un instrumento con el que se busca aportar información sobre la manera en que diferentes factores pueden afectar el progreso de los alumnos durante la enseñanza de las matemáticas. Para este fin, a través del IRC se pueden responder once preguntas, que se agrupan en cuatro rubros: a) las actividades de enseñanza; b) la comunicación en el aula; c) las herramientas empleadas; y d) las relaciones y normas en clase. Estas preguntas se responden de acuerdo con una escala que va de cero a tres puntos (Brown, 2002).

El **Scoop Notebook**. Es un instrumento que permite medir la calidad de la enseñanza de las matemáticas a través de ciertos materiales que se producen en diferentes momentos: a) antes de la clase; b) durante la clase; y c) después de la clase. Para analizar la información, este instrumento se apoya de portafolios, que es un compilado de materiales que los docentes deben proporcionar durante una semana de trabajo. Dichos materiales incluyen las planificaciones u otros materiales que los docentes consideren como insumo de trabajo previo a la clase; materiales generados por los estudiantes durante la clase; y los trabajos propuestos para realizar en casa. Para conformar los portafolios se solicita a los docentes que elijan dos trabajos en cada momento, uno que consideren de alta calidad y otro con baja calidad. Además de los trabajos, este instrumento incluye una bitácora del día de trabajo y reflexiones de los docentes (Borko, Stecher y Kuffner, 2007).

Mathematical Quality of Instruction (MQI). Es un instrumento que se desarrolló como soporte para el *Mathematical Knowledge for Teaching*, con el que se miden los conocimientos matemáticos de los docentes. El propósito del MQI es medir la riqueza de las matemáticas, el trabajo matemático de los estudiantes, los desaciertos e impresiones y el quehacer matemático, de acuerdo con una escala de tres niveles de desempeño en la práctica de enseñanza (Hill et al, 2008).

El **Instructional Quality Assessment (IQA).** Es un instrumento con el que se puede obtener información descriptiva y estadística sobre la naturaleza de la enseñanza y las oportunidades de aprendizaje que los docentes ofrecen a los estudiantes a través de su práctica docente. Su propósito es identificar la calidad de las actividades de enseñanza y si los docentes pueden mantener la demanda cognitiva de dichas actividades durante su implementación, por lo que el instrumento evalúa: las actividades de enseñanza, la implementación de las actividades de enseñanza, las discusiones matemáticas y las expectativas de los docentes a través de rúbricas con una escala de 0-4 puntos (Boston, 2012).

Discrete Choice Experiments (DCE). Es un instrumento cuyo propósito es identificar las decisiones que los docentes toman respecto al tipo de actividades de enseñanza que promueven en la clase de matemáticas. El **DCE** permite adentrarse en las preferencias de los docentes al seleccionar un tipo específico de actividades de enseñanza presentadas en viñetas, mediante las cuales los docentes seleccionan: a) la actividad que puede promover un mejor aprendizaje para los estudiantes; b) los casos en los que los alumnos pueden disfrutar más el proceso de aprendizaje; y c) el caso sería más fácil de planificar (Aubuson *et al*., 2014).

El Instrumento de Evaluación de los Procesos Matemáticos en Educación Infantil (IEPME). Su propósito es determinar la presencia o ausencia de cinco estándares de enseñanza propuestos por el National Council of Teachers of Mathematics (2000): a) la resolución de problemas; b) el razonamiento y demostración; c) las conexiones; d) la comunicación; y, e) la representación de problemas matemáticos que los estudiantes pueden lograr. Para la valoración de estos estándares se hace uso de clases videograbadas y se utilizan siete indicadores con variables dicotómicas para determinar si estos estándares se llevan a cabo durante una clase completa de matemáticas (Alsina y Coronata, 2014).

La revisión de los diferentes protocolos de observación e instrumentos sirvió como referente para el diseño del POEAM, específicamente se identificó que la mayoría de los protocolos e instrumentos se caracterizan por requerir una alta inferencia (juicios de valor) por parte de los observadores, así como por emplear escalas ordinales para clasificar la información. Asimismo, se encontró que el uso de videograbaciones de clase es un práctica común en estos estudios y que el proceso de calificación a través de esta técnica requiere de especial cuidado en la capacitación.

Capítulo III: Diseño del protocolo de observación

En este capítulo se describen los aspectos metodológicos que guían el desarrollo del estudio, específicamente se señala el proceso que se siguió para el diseño del Protocolo de Observación de las Actividades de Enseñanza en Matemáticas (POAEM). El capítulo se divide en dos apartados, el primero establece las variables, dimensiones y criterios para definir el constructo latente a medir. El segundo describe las diferentes etapas que se siguieron durante el proceso de diseño del protocolo de observación.

1. Operacionalización del constructo

Una etapa importante para el diseño del protocolo de observación fue definir el concepto a observar, esto es lo que en la teoría se denomina como el constructo latente a medir. Dichos constructos generalmente son multidimensionales, por lo que resulta complejo realizar su análisis de manera directa; por lo tanto, es necesario definir variables, dimensiones y criterios que permitan establecer de manera operativa o medible la información.

En este estudio, el constructo latente a medir son las actividades de enseñanza en matemáticas, entendiendo éstas como lo que hace el docente y los estudiantes durante la clase a partir de las tareas matemáticas propuestas. Cabe aclarar que en este caso se entiende que “una tarea matemática no se convierte en actividad de enseñanza hasta que los estudiantes la realizan” (Griffin, 2009, p. 32); por lo tanto, la unidad de análisis para valorar estas actividades serán las tareas matemáticas que las conforman.

La operacionalización del constructo se llevó a cabo a partir de algunas variables, dimensiones e indicadores identificados tanto en la revisión de literatura como a partir de elementos que se obtuvieron durante las observaciones de clase de matemáticas previas al diseño del instrumento. De esta manera fue posible traducir el concepto de actividad de enseñanza en hechos observables; es decir, susceptibles de observación y cuantificación.

Las variables consideradas para la operacionalización del constructo son: propósito de la actividad, demanda cognitiva, recursos empleados durante la actividad de enseñanza, alternativas de solución, preguntas para la reflexión, contextualización de las situaciones matemáticas, y el trabajo colaborativo que promueven (ver tabla 1).

Se reconoce que además de las variables mencionadas, existen otros aspectos importantes de las actividades de enseñanza, como: la incertidumbre que generan para los estudiantes, las restricciones o limitaciones que incluyen las tareas matemáticas que las conforman, la disposición productiva que pueden lograr los estudiantes a partir de la actividad de enseñanza, la autoevaluación de los estudiantes, entre otras; sin embargo, las que se emplearon en este estudio son las que se identificaron de manera consistente en la revisión de literatura (Santos, 2003; Llinares, 2008; Verschaffel, Greer y De Corte, 2010; Picaroni y Loureiro, 2010; NCTM 2015; Swan, 2015). La siguiente tabla presenta las variables, dimensiones, valores e indicadores que conforman la operacionalización del constructo “Actividades de enseñanza” y en la siguiente sección se detalla el proceso que se siguió para establecer dichos elementos.

Tabla 1:
Operacionalización

Variable	**Conceptualización**	**Dimensiones**	**Valores**	**Indicadores**
Propósito	Intención general del aprendizaje para identificar los conocimientos y habilidades a desarrollar.	Claridad de la tarea matemática		• Señala el conocimiento o productos esperados. • Plantea preguntas relacionadas con el objetivo o producto esperado. • Estrategias para verificar la comprensión de los estudiantes.
Demanda cognitiva	Tipo y nivel de pensamiento requerido para realizar ciertas tareas matemáticas.	Exigencia cognitiva	Exigencia de bajo nivel (memorización)	• Reproducción de hechos, reglas, fórmulas o definiciones previamente aprendidos o definidos. • No promueven una relación con los conceptos, significados, fórmulas o definiciones aprendidas o reproducidas.
			Exigencia de bajo nivel (procedimientos sin conexiones)	• Utilizan un procedimiento específico o alguno que resulta evidente a partir de la tarea matemática. • No hay relación con conceptos o con el significado subyacente al procedimiento empleado. • No requieren explicaciones o éstas se centran solamente en la descripción del procedimiento utilizado.
			Exigencia de alto nivel (procedimientos con conexiones)	• Enfocan la atención del estudiante en la utilización de procedimientos, con el propósito de desarrollar niveles más profundos de comprensión de los conceptos e ideas matemáticas. • Sugieren procedimientos generales que requieren involucrarse con ideas o procedimientos para lograr una comprensión de las matemáticas. • Requieren un pensamiento complejo y no algorítmico.

			Exigencias de alto nivel (construcción de las matemáticas)	• Demandan que los estudiantes exploren y entiendan la naturaleza de los conceptos matemáticos. • Requieren un esfuerzo cognitivo significativo y pudieran entrañar un nivel de ansiedad para los estudiantes debido a la naturaleza imprescindible de los procesos de solución necesarios.
Recursos empleados durante la actividad de enseñanza	Uso que el docente y los estudiantes hacen de los recursos disponibles para promover o fortalecer los aprendizajes a partir de la tarea matemática propuesta.	Recursos		• Uso de artefactos como medio para promover el aprendizaje. • Coherencia del artefacto o material con el propósito de la tarea matemática. • Apoyo que el artefacto o material ofrece a los estudiantes para desarrollar sus propias estrategias de aprendizaje.
Alternativas de solución	Flexibilidad del docente para promover y aceptar diferentes estrategias o procedimientos para resolver un problema matemático.	Alternativas de solución	Múltiples posibilidades de solución	• La actividad requiere enfrentar y resolver situaciones o problemas poco estructurados • La actividad permite poner en práctica diversas estrategias de solución (evita algún método predeterminado).
			Posibilidades de solución limitadas	• La actividad incluye todos los datos necesarios para abordar la situación propuesta. • La actividad ofrece aspectos (o pistas) y enfatiza la situación planteada. • La actividad promueve un método de solución específico.

Preguntas para la reflexión	Tipo de preguntas que se plantean durante la realización de una tarea matemática (ya sea por el docente o los estudiantes) para dar sentido a ideas o relaciones matemáticas importantes.	Tipo de preguntas	Preguntas generadoras	• Preguntas que implican justificar respuestas, hacer distinciones, generalizar ideas y utilizar el diálogo para construir ideas coherentes que promuevan la mejora de un entendimiento colectivo.
Contextualización	Se refiere a si el contexto de una *situación problemática* ofrece a los estudiantes la oportunidad de conectar los aprendizajes matemáticos con su propia experiencia, con otras disciplinas o con situaciones de la vida cotidiana.	Conexión/aplicación de los conocimientos matemáticos	Contexto significativo	• La actividad permite comunicar información o ideas con un significado personal o público. • La actividad refleja contextos reales, propios de la vida personal o social.
			Contexto intermatemático	• La actividad incluye información que hace referencia a elementos de la realidad para relacionarla con una situación disciplinar.
			Contexto escolar	• La actividad sugiere información que sólo se puede aplicar al contexto escolar. • La actividad plantea una problematización del conocimiento relacionado con la propia disciplina. • La actividad no establece un vínculo con las formas de producción y de uso en situaciones propias de la vida real.
			Actividades de enseñanza sin contexto	• El conocimiento se presenta aislado de las situaciones que podrían darle sentido.
Trabajo colaborativo	Interacción, negociación y cooperación entre estudiantes para lograr un propósito común.	Trabajo colaborativo		• Interacción entre estudiantes (y en algunos casos con el docente) para lograr un propósito común. • Ambiente generado para promover el aprendizaje.

2. Protocolo de Observación de las Actividades de Enseñanza en Matemáticas

El diseño del Protocolo de Observación de las Actividades de Enseñanza en Matemáticas comprendió diferentes etapas de desarrollo a partir de la identificación de las variables y dimensiones que se incluyen en la operacionalización del constructo: a) diseño de un esquema del protocolo de observación como base para el diseño de una primera versión del instrumento; b) revisión de otras estructuras de protocolos de observación; y c) elaboración de tres versiones a partir del esquema inicial, los cuales se fueron ajustando de acuerdo con las observaciones y sugerencias de diferentes especialistas en educación, matemáticas y metodología de diseño de instrumentos.

La primera etapa implicó realizar cuatro observaciones de clase de matemáticas en quinto grado de primaria con el propósito de identificar aspectos importantes a considerar en la observación de las actividades de enseñanza. Algunos de éstos se cotejaron con las variables y dimensiones retomadas de la revisión de literatura, con las cuales se diseñó un esquema general como base para el diseño del instrumento. Este esquema incluyó tres dimensiones adicionales a las que se señalan en la operacionalización del constructo en este documento: a) incertidumbre, b) construcción del conocimiento y c) reflexión de los estudiantes, éstas se reorganizaron en las dimensiones existentes de acuerdo con la naturaleza de los criterios a observar (tabla 1).

La segunda etapa consistió en revisar algunos protocolos de observación en aula validados y empleados en diferentes áreas del campo educativo. Por ejemplo, el Mathematical Quality of Instruction (MQI), el Classroom Assessment Scoring System (CLASS), el Protocol for Language Arts Teaching Observation (PLATO) y el Framework For Teaching protocol (FFT). Todos ellos emplearon clases videograbadas de matemáticas como insumos para probar el protocolo de observación, por lo que se tomaron como un referente para utilizar esta técnica con el POAEM.

Los protocolos de observación mencionados, así como otros instrumentos de observación (Instrument for Ranking Classes, Scoop Notebook, Mathematical Quality of Instruction, Instructional Quality Assessment, Discrete Choice Experiments), permitieron analizar y establecer un diseño para los propósitos del POAEM, así como para el proceso de calificación. Cabe mencionar que el protocolo MQI y el sistema CLASS inspiraron el diseño del POAEM, aunque este último plantea diferencias importantes en relación con los niveles de desempeño incluidos en las rúbricas que lo conforman (ver capítulo III de este documento).

A partir de los instrumentos revisados y del análisis de su diseño y proceso de obtención de información, se desarrolló una primera versión del protocolo de observación en la que se incluyeron algunas variables y dimensiones que se emplearon para elaborar rúbricas. La información se organizó en rúbricas debido a que éstas son herramientas que permiten establecer estándares de desempeño con especificaciones en términos de calidad relativa (Wiggins, 1998; Martínez Rizo, 2012) y considerando que ya han sido probadas y empleadas en diferentes protocolos de observación para valorar algunos aspectos de las prácticas de enseñanza (ver capítulo II de este documento).

La segunda versión del protocolo de observación incluyó cambios en los niveles de desempeño de cada rúbrica, realizados con base en las sugerencias de un panel de especialistas durante el jueceo del instrumento. A partir de dichos cambios, cada una de las rúbricas del instrumento se extendió a cuatro niveles de desempeño (bajo, medio, medio-alto y alto), con el propósito de ofrecer un rango más amplio en las descripciones de cada dimensión y evitar un sesgo hacia el nivel intermedio de la primera versión. Además, se incluyó un diagrama con cuatro dominios generales en los que se organizaron las dimensiones a medir de acuerdo con su naturaleza teórica, para ofrecer una descripción general de la organización del instrumento y lo que pretende medir,.

En esta versión se incluyó una descripción del proceso de observación y calificación. Para este último, se diseñó una hoja de calificación a partir de los dominios y las dimensiones que los integran. Los observadores otorgaron una puntación dependiendo del nivel de desempeño, con un rango de uno a ocho. Las puntuaciones se establecieron como 1 y 2 para el nivel bajo; 3 y 4 para el nivel medio; 5 y 6 para el nivel medio-alto; y 7 y 8 para el nivel alto, con el propósito de que los observadores tuviesen un margen más amplio de puntajes. Asimismo, se anexaron ejemplos para cada nivel de desempeño en las dimensiones propuestas con el propósito de ofrecer a los observadores situaciones concretas en las que se pudiera analizar los criterios incluidos en las diferentes rúbricas.

La tercera versión del protocolo de observación se desarrolló a partir de nuevas sugerencias realizadas por el panel de especialistas durante el segundo jueceo del instrumento. Los cambios realizados con respecto a la segunda versión se centraron en la reorganización de la información del instrumento para mejorar su fluidez y claridad para los observadores. En esta versión se partió de la explicación del propósito del protocolo de observación (valorar lo que hace el docente y los estudiantes durante la clase a partir de las tareas matemáticas que se propone a los estudiantes); se describió la organización del instrumento y se incluyó la definición de las dimensiones a valorar. Otros cambios tuvieron que ver con la hoja de calificación, pues se anexó una opción para cuando las observadoras no pudieran valorar una dimensión específica. Por ejemplo, en trabajo colaborativo se marcó no se puede valorar cuando el docente explícitamente pidió a los estudiantes que el trabajo se hiciera individualmente (ver anexo E).

2.2 Jueceo del protocolo de observación

Cada una de las tres versiones del protocolo de observación implicó un proceso de jueceo que se llevó a cabo con el apoyo de especialistas en pedagogía, matemáticas y metodología para el diseño de instrumentos de evaluación. La invitación para participar en el estudio se realizó a través de reuniones individuales en las que se explicó el propósito del estudio y las aportaciones que podrían hacer.

La siguiente tabla presenta información general de los jueces participantes de acuerdo con su área de trabajo:

Tabla 2:

Perfil de los jueces participantes

Educación[2]	Matemáticas	Metodología
Licenciada en Educación Primaria con seis años de experiencia. Participó en los diplomados en Evaluación Formativa I y II en la Universidad Autónoma de Aguascalientes.	Ambos obtuvieron el grado de doctor en Ciencias con especialidad en Matemática Educativa. Uno de ellos ha centrado su trabajo en el aprendizaje en la práctica situada de la enseñanza de las matemáticas y el segundo en los procesos metacognitivos en el aula de las matemáticas.	Ambos especialistas obtuvieron su maestría en investigación educativa. Los dos son estudiantes del programa de Doctorado en Ciencias Educativas en la línea de investigación de validación de instrumentos. Además, ambos tienen experiencia en procesos de desarrollo e implementación de instrumentos de evaluación por el Instituto Nacional de Evaluación para la Educación (INEE).

El jueceo de la primera versión del instrumento se llevó a cabo en dos sesiones de cuatro horas. En estas sesiones se explicó el propósito del jueceo, la estrategia de trabajo y se organizó a los jueces en pares, cuidando que especialistas de diferentes áreas[3] pudieran discutir durante la sesión. Los jueces realizaron una lectura de cada uno de los apartados del documento "Protocolo de observación de las actividades de enseñanza en matemáticas", discutieron en pares y, posteriormente, en plenaria las observaciones correspondientes a tres aspectos generales del documento: a) utilidad de la información para los observadores; b) claridad de las ideas; y c) información que podría omitirse en el documento. Para el análisis de las rúbricas de la primera versión del instrumento se solicitó a los participantes que comentaran qué comprendieron de cada dimensión, qué sugerencias podrían dar para mejorar el formato de las rúbricas propuestas y qué sugerencias generales podrían aportar para la mejora del instrumento (en todos los casos justificando las respuestas).

[2] Inicialmente se había convocado a tres docentes con diferentes años de experiencia; sin embargo, por las fechas y carga de trabajo no fue posible contar con el apoyo de todos los docentes contemplados.
[3] 2 especialistas en matemáticas; 2 especialistas en metodología y 1 especialista en educación.

Los acuerdos a los que se llegó involucraron diferentes aspectos, algunos relacionados con el formato, organización de la información, datos generales que convendría incluir, así como cambios que requirieron mayor trabajo. Por ejemplo, ajustes a los criterios propuestos en diferentes rúbricas de acuerdo con las bases teóricas, añadir un nuevo nivel de desempeño para cada rúbrica, incluir información sobre el significado de cada dimensión y lo que se pretende valorar de las actividades de enseñanza.

El segundo jueceo se realizó con la misma estrategia de trabajo, pero con una sesión de cuatro horas. Los acuerdos a los que se llegó se centraron en ajustes a los niveles de las rúbricas propuestas para diferentes dimensiones, además se sugirió segmentar las videograbaciones de clases completas para llevar a cabo la validación del instrumento. La idea central del último punto fue asegurar que los observadores tuviesen las tareas matemáticas como unidad de análisis, y que las tareas identificadas y valoradas fueran las mismas.

Capítulo IV: Insumos para aplicación del POAEM

Este capítulo describe el proceso de obtención de la información para poner a prueba el funcionamiento del protocolo de observación diseñado. El capítulo se divide en cuatro apartados. En el primero se expone la funcionalidad de las videograbaciones como principal insumo para valorar las prácticas de enseñanza; el segundo describe el proceso de gestión y estrategia de contacto con los docentes que participaron en el estudio; el tercero describe el proceso de recolección de las videograbaciones; y, el cuarto apartado se centra en los aspectos éticos para cuidar la integridad de los participantes y de la información que proporcionaron.

1. Uso de la videograbación para valorar las prácticas de enseñanza

Las videograbaciones para valorar la práctica de enseñanza e identificar los diferentes aspectos de la misma son una opción reiterada en diferentes instrumentos y protocolos de observación (ver capítulo II). Estudios como el Classroom Survey Study y el Trends in International Mathematics and Science Study (TIMSS), en 1995, y su réplica en 1999 –TIMSS-R– fueron pioneros en emplear esta técnica para ofrecer una imagen de lo que acontece en las aulas de matemáticas de diferentes países (Stigler, Gallimore y Hiebert, 2000; Stigler y Hiebert, 2004).

Bell *et al.* (2015) señalan que el uso de la videograbación no afecta la obtención de información confiable sobre conductas complejas, ya que la dificultad para identificarlas se puede enfrentar tanto en tiempo real como en tiempo diferido; por lo tanto, los resultados que se obtienen a partir de los protocolos de observación dependen más de los observadores y de su entrenamiento que de la manera en que se obtiene la información.

En el presente estudio, el uso de la videograbación representó un medio fundamental para la obtención de la información. Las videograbaciones de las clases ayudaron a prever complicaciones en la obtención de información considerando que el periodo vacacional del estado de Aguascalientes difiere del resto del país. Se consideraron los compromisos en tiempo y actividades que representaría para los docentes realizar observaciones en tiempo real durante los últimos meses del ciclo escolar. Por lo tanto, el trabajo de campo se adelantó a los meses de abril-junio 2016. Además, las videograbaciones permitieron asegurar la información requerida, buscando contar con la participación de cada uno de los docentes convocados en este estudio (26), así como el número de observaciones en cada caso (dos sesiones en días consecutivos o alternados). Finalmente, las videograbaciones se constituyeron en una técnica que no afectó la obtención de información confiable sobre conductas complejas, ya que de acuerdo con resultados de diferentes estudios éstos dependen más de los observadores y su entrenamiento que de la forma en que se obtiene la información (MET, 2010f; Bell et. al., 2015).

En total, se realizaron 52 videograbaciones de clases completas de matemáticas, correspondientes a 26 docentes participantes (2 clases por docente). Todos los docentes laboran en escuelas públicas en contexto urbano en los municipios de Aguascalientes, Jesús María y San Francisco de los Romo. La distribución de los sujetos participantes se muestra en la siguiente tabla:

Tabla 3:
Distribución de docentes participantes

Municipio	Número de docentes	Número de videograbaciones
San Francisco de los Romo	3	6
Jesús María	2	4
Aguascalientes	21	42
Total	**26**	**52**

Cada una de las 52 videograbaciones realizadas fue segmentada de acuerdo al número de tareas matemáticas identificadas en cada caso. Entendiendo que una tarea matemática “puede abarcar desde un conjunto de ejercicios rutinarios hasta un problema complejo y desafiante que enfoque la atención de los estudiantes en una idea matemática particular” (NCTM, 2015, p. 19).

El proceso de segmentación implicó dos fases: a) una descripción de la información contenida en cada una de las clases videograbadas, y b) el análisis de las descripciones de cada videograbación para determinar el número de tareas matemáticas. La fase de descripción de las videograbaciones se realizó con el apoyo de cinco colaboradoras: dos investigadoras del Departamento de Educación de la Universidad Autónoma de Aguascalientes (UAA), una estudiante de la Maestría en Investigación Educativa, una licenciada en Asesoría Psicopedagógica y la autora de este estudio.

Para realizar la descripción de las videograbaciones, se les explicó que el propósito de la actividad consistía en redactar lo que sucedía durante la clase observada y posteriormente se les entregó un formato para el vaciado de la información con: instrucciones, notas aclaratorias y un apartado de confidencialidad de la información que cada una firmó firmado. Los datos que se registraron en cada formato fueron:

a. Número de actividad propuesta por el docente durante la clase.
b. Minuto y segundo de inicio y final de la actividad.
c. Tema con el que se identifica la actividad.
d. Indicación al inicio de la actividad.
e. Indicación al final de la actividad.
f. Descripción de lo que hace o propone el docente.
g. Trabajo de los estudiantes (qué hacen y cómo se organizan).
h. Observaciones generales: aspectos particulares que apoyen la descripción de la actividad y (si es el caso) el material empleado durante su realización (ver anexo F).

Utilizando el formato de descripción se llevó a cabo un ejercicio individual con una misma videograbación, de tal manera que se pudiera comparar la comprensión respecto a la información solicitada. Se retroalimentó a cada una de las colaboradoras; posteriormente, se designó un número específico de docentes a observar y se acordaron dos fechas de entrega, una de avances y una final. En compensación, se entregó un estímulo monetario para la asistente de investigación y la estudiante de maestría, así como una carta con las actividades realizadas para incluir créditos académicos para esta última. En el caso de las investigadoras, su participación fue voluntaria, ya que mostraron un interés profesional por apoyar en la investigación.

La segmentación de las videograbaciones permitió establecer la unidad de análisis para obtener evidencias de validez del protocolo de observación. Es decir, con la descripción de cada una de las clases videograbadas fue posible identificar el número de tareas matemáticas que las conforman, tomando como base una idea matemática común. En promedio, se identificaron tres tareas matemáticas por clase videograbada.

2. Gestión de participación y agenda de videograbaciones

La obtención de información para probar el instrumento y establecer algunos aspectos de su validez se realizó a partir de la gestión de participación de los docentes y agenda de videograbación de clases de matemáticas con docentes de quinto grado de primaria. Este proceso comprendió tres meses de trabajo, del 14 de marzo al 12 mayo de 2017.

El primer acercamiento a los sujetos de estudio se realizó a través de un muestreo por bola de nieve, utilizando el contacto previo con docentes de primaria que participaron en proyectos de evaluación formativa en la Universidad Autónoma de Aguascalientes (Martínez-Salgado, 2012). A partir de la información disponible, se establecieron dos estrategias de acercamiento. La primera consistió en enviar a través de correo electrónico una invitación para participar en un proyecto de investigación vinculado a dos tesis de doctorado enfocadas al diseño de instrumentos de observación para valorar: a) las actividades de enseñanza que se promueven durante las clases de matemáticas, y b) el tipo de retroalimentación que los docentes ofrecen a los estudiantes a partir de las actividades de enseñanza[4].

La carta invitación incluyó información sobre la participación requerida por parte de los sujetos participantes: videograbación de dos clases completas de matemáticas, el propósito de los estudios, aclarando que no se aportarían resultados sustanciales sobre las prácticas docentes, sino sobre las cualidades de los instrumentos para reflejar de manera adecuada las actividades de enseñanza, información de contacto de las responsables de los estudios, y finalmente se solicitó el apoyo para identificar nuevos docentes que tuvieran interés en participar en el estudio.

La segunda estrategia consistió en contactar a supervisores de zonas escolares a quienes se brindó información sobre los proyectos de investigación, de tal manera que pudieran hacer extensiva la invitación a docentes de quinto grado de sus zonas escolares. A través

[4] Tesis doctoral: Diseño de una guía de observación de prácticas de retroalimentación durante las clases matemáticas en quinto grado de primaria.

de este medio fue posible contar con el apoyo de cinco supervisores, quienes enviaron una carta invitación para los docentes, que incluyó: los propósitos de los estudios, las implicaciones de la participación de los docentes, el carácter voluntario de la participación, y los cuidados que los investigadores tendrían con la información recabada, así como los compromisos adquiridos (ver anexo A).

Con los datos de los docentes que accedieron a participar se organizó una agenda digital en la que se registraron las fechas de grabación de clases de acuerdo con la disponibilidad de los participantes. En la mayoría de los casos fue posible calendarizar las dos sesiones de clase en días consecutivos.

3. Recolección de datos para probar el protocolo de observación

El proceso de videograbación de las clases se realizó durante el periodo del 7 de abril al 8 de junio de 2016. Durante este tiempo se videograbaron aproximadamente a tres docentes por día, en diferentes horarios, en su mayoría en escuelas de turno matutino. Para las videograbaciones se utilizó una cámara Canon HD CMOS PRO como fuente principal, y una cámara SONY HD o Ipod 5 touch-Apple como segunda fuente[5] para contar con información de respaldo en caso de fallas en la fuente principal. Además, se utilizaron micrófonos de solapa que portaron los docentes durante las grabaciones para cuidar la calidad del sonido.

El proceso implicó llegar a la escuela primaria correspondiente al menos 15 minutos antes de la hora acordada, presentarse con el director, quien estaría notificado previamente por el supervisor o el propio docente, y ubicar el salón de clase de quinto grado del docente participante. Una vez localizado el salón de clase, se instalaban el tripié y la cámara principal en un lugar donde se evitara interrumpir el paso de los estudiantes. Una vez instalado el equipo se entregaba a los docentes el micrófono de solapa.

El tiempo de grabación dependió de la duración de la clase completa del docente observado; en promedio setenta minutos, aunque se grabaron clases de más de dos horas. Al finalizar las sesiones, se solicitó a los docentes responder unas breves

[5] No en todas las grabaciones fue posible tener la segunda fuente, ya que se tuvieron problemas para controlar este dispositivo por cuestiones de movilidad en las aulas, así como por problemas técnicos.

preguntas de contextualización en las que se incluyó información sobre el número de estudiantes, número de niños con necesidades educativas especiales, años de servicio como docente, materiales que utiliza para desarrollar su planificación y el propósito de la clase observada (ver anexo B). Al finalizar las videograbaciones de los docentes, se les entregó en agradecimiento, un paquete de libros, producto de investigaciones del departamento de educación de la UAA focalizadas en la evaluación formativa en aulas de primaria y se ofreció compartir un informe de resultados del estudio.

Para organizar la información de las videograbaciones se asignó un nombre a cada uno de los archivos de acuerdo con el número de lista del docente en la agenda de trabajo; iniciales del municipio donde se encuentra la escuela primaria; iniciales del docente participante; número de grabación, especificando si era la primera o segunda sesión; la fecha de grabación, señalando secuencialmente el mes, día y año; y la fuente de videograbación, para identificar si fue a través de la cámara principal o la secundaria, por ejemplo: 02_AGS_ZXZ_02_120416_P (este código representa la segunda sesión grabada).

En el siguiente diagrama se resume cada una de las etapas del proceso de diseño y obtención de información para la prueba piloto del Protocolo de Observación de las Actividades de Enseñanza en Matemáticas (Figura 5).

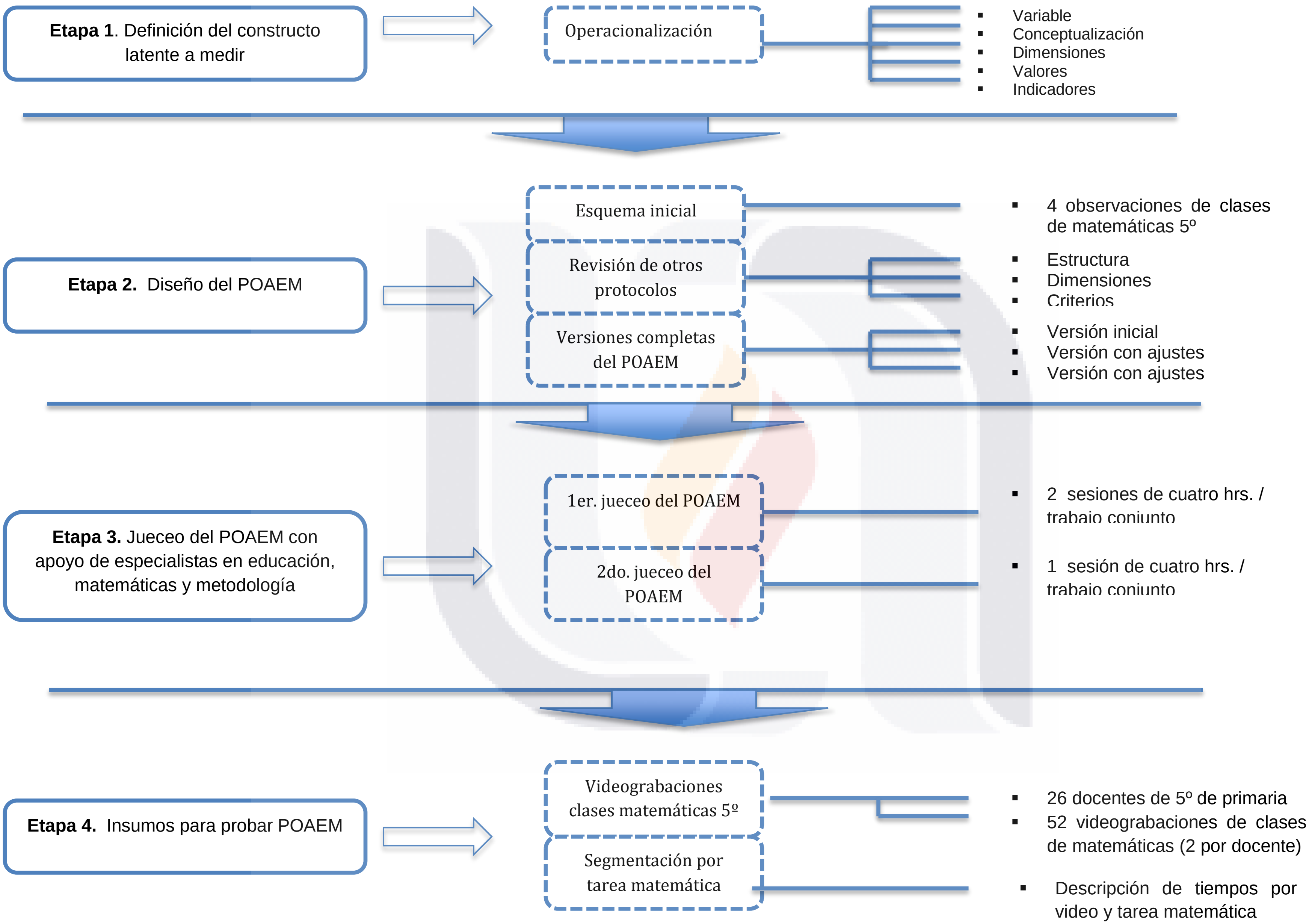


Figura 6. Proceso del diseño del Protocolo de Observación de las Actividades de Enseñanza en Matemáticas
Fuente: Diseño propio.

4. Consideraciones éticas generales

Parte del proceso de recolección de información implicó considerar algunos aspectos éticos que permitieran cuidar la integridad de los participantes y de la información que proporcionaron. Una de las medidas fue diseñar un formato de consentimiento voluntario informado para los docentes, el cual se firmó previo a las videograbaciones. Este formato incluyó, además del propósito del estudio, información sobre la recolección y el tratamiento posterior de la información (ver anexo C).

Se solicitó el apoyo de los docentes participantes para dar parte a los padres de familia sobre las videograbaciones que se realizarían con el consentimiento de las autoridades escolares. No obstante, durante el desarrollo de las videograbaciones surgió la inconformidad de un padre de familia, por lo que fue necesario discutir la situación entre las investigadoras responsables y los directores de tesis, llegando al acuerdo de redactar una carta de dimisión de participación para el docente.

La carta informó sobre la cancelación de la participación del docente y se señaló que los archivos electrónicos no fueron copiados en ningún otro dispositivo, y dado que el docente es el único propietario de las videograbaciones, al firmar la carta de dimisión las coordinadoras de los estudios, así como la institución que representan no son responsables del uso que el docente pueda hacer de los archivos electrónicos entregados. La carta se firmó por las investigadoras responsables y por el docente participante (ver anexo D).

Para el tratamiento posterior de la información se redactó una carta de confidencialidad de la información para los observadores, que incluyó un formato en el que se especifican las videograbaciones y las tareas matemáticas que las conforman. Cada observadora firmó el compromiso de mantener en total confidencialidad la información observada, así como la identidad de los sujetos participantes para cada una de las videograbaciones revisadas.

Capítulo V: Obtención de evidencias de validez y confiabilidad

Este capítulo incluye dos apartados, el primero define los conceptos de validez y confiabilidad como bases teóricas para establecer las evidencias que se ajustan a las particularidades de cada proceso y propósito de interpretación de la información recabada a través del POAEM. El segundo a parado describe el procedimiento que se siguió para obtener dichas evidencias: pilotaje del instrumento, capacitación de observadoras y calificación de videograbaciones.

1. Bases teóricas para evidencias de validez y confiabilidad de la información

La validez se define como el "grado en que la evidencia y la teoría respaldan las interpretaciones de los resultados de un instrumento y el uso que se puede hacer de éstos" (AERA, APA, y NCME 2014, p. 11). Por lo tanto, el proceso de validación involucra la acumulación de evidencias relevantes con las que se puede proporcionar una base científica sólida para llevar a cabo la interpretación, a partir de ciertos resultados que tienen que ver tanto con las declaraciones explícitas sobre el constructo, como con la importancia del uso de esas interpretaciones (Martínez Rizo, 2015).

Existen diferentes fuentes que permiten aportar evidencias de validez sobre las interpretaciones de la información que se obtiene con cualquier instrumento. Estas fuentes representan varias líneas del concepto de validez y se ajustan a las particularidades de cada proceso y propósito de interpretación de resultados o puntajes. Por ejemplo, evidencias basadas en significados e interpretaciones a partir de subgrupos de personas que respondan un instrumento; la estructura interna que se relaciona con la validez de constructo; la relación con otras variables, involucrando evidencia experimental y correlacional; de contenido; y consecuencias de las pruebas y de sus interpretaciones (AERA, APA, y NCME 2014, 2014).

La AERA ha establecido estándares para desarrollar el proceso de validación, considerando diferentes aspectos y circunstancias, algunos de éstos se centran en:

a. Los usos e interpretaciones que se pueden hacer a partir de las puntuaciones, para lo cual se recomienda que el diseñador:
   - Establezca cómo se interpretarán las puntuaciones del instrumento y cómo éstas serán utilizadas. Se recomienda utilizar un lenguaje acorde con los usuarios del instrumento.
   - Establezca una justificación para cada interpretación de los puntajes, así como para el uso que se puede hacer de los mismos. Los diseñadores deben incluir sus propias recomendaciones; sin embargo, será responsabilidad de los aplicadores evaluar la calidad de la evidencia proporcionada y su relevancia en cada situación.

b. Sobre la muestra y la configuración empleada en el proceso de validación:
   - La composición de los sujetos a quienes se aplicará el instrumento se debe describir a detalle, de tal manera que la información que se incluya permita identificar características sociodemográficas y de desarrollo, ya que los resultados estadísticos podrían verse afectados por dichos factores.
   - Cuando el proceso de validación involucra la opinión de jueces expertos, se debe incluir el proceso que se siguió para seleccionar a esos participantes, así como una descripción de su experiencia y cualidades. Además, se deberá incluir cualquier otro procedimiento que haya fortalecido esta etapa (entrenamiento, instrucciones que siguieron, etcétera).

c. Respecto a formas específicas de evidencias de validez:
   - Si las interpretaciones de los puntajes dependen de la relación de los ítems del instrumento es necesario incluir evidencias sobre la estructura interna del instrumento.
   - Lo resultados estadísticos que se incluyan en el proceso de validación deben incluir una clara descripción y justificación que apoyen dicho proceso (AERA, 2014).

Dado que éste es un estudio metodológico con el que se busca aportar evidencias de algunos aspectos de validez, a continuación se enfatizan elementos teóricos de los estándares para desarrollar el proceso de validación propuestos por la AERA. Específicamente, se menciona lo referente a la opinión de jueces expertos y la estructura interna del instrumento como aspectos básicos para establecer las estrategias de atención a los datos y a su interpretación.

1.1 Opinión de jueces expertos: evidencias de validez de contenido

Las evidencias de validez de contenido se pueden obtener mediante las opiniones de expertos, las cuales permiten asegurar que las dimensiones medidas realmente representan las variables y dimensiones de interés. La manera más común de llevar a cabo la validez de contenido es contando con la participación de un panel de jueces expertos, quienes pueden ofrecer una opinión informada sobre el tema (Escobar y Cuervo, 2008; Hernández, Fernández y Baptista, 2010).

Las técnicas para la participación de los jueces pueden ser individuales o grupales. En las técnicas individuales se puede mencionar la de agregados individuales y el método Delphi, en ambas cada juez realiza su evaluación; sin embargo, con el método Delphi se pretende que los jueces lleguen a acuerdos con base en los resultados obtenidos. En cuanto a las técnicas grupales se puede mencionar la nominal y por consenso, las cuales requieren reunir a los expertos para que lleven a cabo el proceso de jueceo en sesiones de trabajo en conjunto (Stewar, Roebber y Bosart, 1997; Escobar y Cuervo, 2008).

En cualquiera de las técnicas, la precisión de los juicios dependerá de las características de los jueces, de su experiencia en el tema y de las características de la tarea que se les encomiende. Para abordar el último aspecto, Escobar y Cuervo (2008) sugieren considerar una serie de puntos que permiten recabar información de manera sistemática:

- Definir el objetivo del juicio de los expertos.
- Considerar la formación académica, experiencia, y reconocimiento de la comunidad por parte de los jueces seleccionados. Aunque no hay un consenso entre autores para determinar el número de jueces, Escobar y Cuervo (2008) sugieren un mínimo de cinco, dos de los cuales deben ser expertos en medición y evaluación y los demás con experiencia en el tema que se esté tratando.
- Explicitar dimensiones e indicadores que están midiendo cada uno de los ítems del instrumento: información que permita a los jueces evaluar la relevancia, suficiencia y pertinencia de cada ítem.

- Especificar el propósito del instrumento: para qué van a ser utilizados los puntajes obtenidos a partir de éste.
- Diseñar plantillas para jueces: diseñar un formato para los jueces de acuerdo con los objetivos de evaluación.

1.2 Estructura interna: validez de constructo

En este caso, la validación se centra en el proceso de acumulación de evidencias que apoyen las inferencias que se realizan a partir de los resultados obtenidos con el instrumento de medición (Pérez, Chacón y Moreno, 2000). Entre las técnicas estadísticas empleadas para llevar a cabo la contrastación de la validez de constructo se puede mencionar el Análisis Factorial, misma que generalmente se lleva a cabo a través de dos modalidades: a) Análisis Factorial Exploratorio (AFE) y b) Análisis Factorial Confirmatorio (AFC). La diferencia más importante entre ambas modalidades tiene que ver con que el AFC se basa en teorías sustantivas y por expectativas, mientras que la primera se centra en los datos para descubrir qué subyace a los mismos (Pérez, Chacón y Moreno, 2000).

En cualquier caso, la evidencia sobre la estructura interna del instrumento se lleva a cabo a través del análisis de factores, ya que éstos permiten identificar cuántas dimensiones integran una variable, así como los ítems que pertenecen a una misma dimensión. A partir del análisis de factores se obtiene mayor información para localizar ítems o reactivos aislados, es decir, que no pertenecen a las categorías que se quiere valorar (Hernández, Fernández y Baptista, 2010).

1.3 Confiabilidad

Brennan (2013) señala que la confiabilidad no es una propiedad de un instrumento en sí mismo, sino una medida de consistencia de los puntajes obtenidos durante todo el proceso de medición. La definición técnica se refiere a entender la confiabilidad a partir de la pregunta "qué tanto error de medición existe en un instrumento, considerando tanto la varianza sistemática como la varianza por el azar" (Quero, 2010, p. 248).

Para estimar la confiabilidad de una medida, Cohe y Swerdlik (2001) y Quero (2010) señalan que es posible emplear la teoría clásica (TC) que divide la varianza de los puntajes observados en dos partes: una sistemática, conocida como varianza del puntaje real y otra aleatoria, también llamada error de varianza. Además, señalan que una alternativa a la TC es la teoría de generalizabilidad (TG), con la que se extienden las bases de la teoría TC, reconociendo y estimando la magnitud de múltiples fuentes de error. Es decir, incorpora fuentes de error potenciales dentro de un diseño de medición para estimar los componentes de varianza relacionados con cada fuente de error detectada en el diseño.

Shavelson y Webb (1991), Brennan (2001) y Martínez Rizo (2015) señalan que para la identificación y estimación de diferentes fuentes de error en una medición, la TG permite reconocer múltiples fuentes de error y estimar cada una de ellas por separado y de forma simultánea; además provee un mecanismo para optimizar la confiabilidad a partir de los resultados obtenidos. Esto es, la TG permite al investigador responder preguntas como: ¿la muestra de tareas o jueces es la mayor fuente de error en la medición?, ¿es posible aumentar la confiabilidad de la medida aumentando el número de tareas o de jueces, o a partir de una combinación de ambos?, entre otras.

Además de las fuentes de error, la TG genera un coeficiente de confiabilidad que refleja el error relativo (coeficiente de generalizabilidad) y el error absoluto (coeficiente de dependencia); el primero apoya decisiones relativas y el segundo las decisiones absolutas. Shavelson y Webb (1991) y Cohen y Swerdlik (2001) sugieren emplear la TG cuando el diseño de medición incluya más de una faceta de error, pues de lo contrario las aportaciones de los análisis no serán redituables. Además, señalan que la confiabilidad de una medición puede tomar varias expresiones al ser estimada, como: coeficientes de precisión, estabilidad, equivalencia, homogeneidad o consistencia interna; sin embargo, un aspecto común es que todos son expresados como coeficientes de correlación.

Aiken (2003); Bell, *et al*. (2015) sugieren que para mediciones relacionadas con aspectos de observación, una manera de aportar evidencia adicional sobre la confiabilidad son los procesos de calibración para valorar si los observadores asignan puntajes que reflejan con mayor precisión la calidad de la información requerida a través de sesiones de evaluación en las que se tenga como referente las puntuaciones de observadores “expertos”. En este sentido, la confiabilidad se puede traducir como la consistencia con que un instrumento está midiendo ciertas características, así como el uso correcto de los datos obtenidos a partir de éste. Por lo tanto, en todo diseño de instrumentos es indispensable reportar la confiabilidad de la medición, así como la manera en que ha sido calculada.

Para el presente estudio, la estrategia de atención y tratamiento de los datos se centró en la aplicación de diferentes análisis estadísticos que permitieran obtener evidencias sobre la validez y especialmente sobre la confiabilidad del POAEM. Para esto se tomó como base la estructura de los datos recabados, el nivel de medición de las variables, el análisis de error de medición, así como el análisis de la calibración de los datos, por lo que las aproximaciones consideradas en cada caso y que se reportan en los resultados se refieren a:

- Promedio por dimensión valorada para determinar la confiabilidad interna, haciendo uso del coeficiente alpha de Cronbach en cada caso (consistencia al interior de los subgrupos).
- Análisis descriptivo de las sesiones de calibración a partir de los porcentajes brutos de las dos observadoras para aportar información adicional y comprender con mayor precisión la etapa de calificación y los resultados obtenidos.
- Estudio de generalizabilidad considerando diferentes facetas de error de acuerdo con la estructura de los datos recabados para el POAEM: docentes, dos jueces, dos ocasiones o días de valoración con el protocolo, y las interacciones entre dichas facetas (la información se detalla en el apartado de resultados).

2. Prueba de funcionamiento del POAEM

La prueba de funcionamiento del protocolo de observación consistió en la aplicación del instrumento en su versión mejorada, posterior al jueceo y prepilotaje del instrumento, con un grupo de sujetos con características similares a la población objetivo. Por las características del estudio no se requirió de una muestra representativa, pero sí contar con el número y tipo de sujetos necesarios para realizar las pruebas estadísticas apropiadas.

Con base en la aplicación de dicha prueba fue posible someter los datos obtenidos a diferentes análisis estadísticos que ayudaron a determinar algunos aspectos de la validez y confiabilidad del protocolo de observación, así como para establecer procedimientos que permitieran fortalecer la aplicación y uso del protocolo de observación en sus siguientes versiones. Los siguientes subapartados describen las etapas que conforman el proceso de la prueba de funcionamiento del Protocolo de Observación de las Actividades de Enseñanza en Matemáticas (POAEM).

2.1 Prepilotaje

En esta etapa se contó con la participación de dos colegas, quienes previamente habían revisado el diseño del POAEM: a) una docente-investigadora especialista en matemáticas y b) una compañera del Doctorado en Investigación Educativa, cuyo trabajo también se centra en diseñar y aportar evidencias de validez de un protocolo de observación (en su caso para valorar la retroalimentación en el aula).

El prepilotaje del instrumento consistió en calificar las videograbaciones de dos clases de un mismo docente, utilizando la versión ajustada del instrumento; los videos utilizados fueron adicionales a los contemplados para la etapa de pilotaje del instrumento. En la mayoría de los casos hubo coincidencias en la calificación para las videograbaciones, si bien no establecieron el mismo puntaje sí coincidieron en el nivel seleccionado. Sólo en la dimensión de recursos y aplicación de los conocimientos se identificaron diferencias de nivel seleccionado, las cuales se discutieron para llegar a un acuerdo.

Las observadoras sugirieron incluir un nivel cero para representar la ausencia de información; sin embargo, considerando los planteamientos teóricos de este estudio, las dimensiones y criterios, se decidió que la ausencia de información se identificara con la opción no se puede valorar (N/V), que se incluyó en la hoja de calificación para cada una de las dimensiones con las que se busca valorar las tareas matemáticas.

Con base en el análisis de la información proporcionada por las participantes, se realizaron ajustes menores al instrumento, se estimó el tiempo de observación y calificación de las videograbaciones, y se incorporaron algunas sugerencias en el manual para los jueces calificadores, así como en la capacitación de los mismos.

2.2 Pilotaje

Para el pilotaje del POAEM (anexo G) se utilizaron 40 videograbaciones de clases de matemáticas correspondientes a 20 docentes de quinto grado de primaria (dos clases por docente). El número de videograbaciones que los observadores calificaron se estableció tomando como referencia el número reportado por otros instrumentos de observación como el CLASS y el CLASS-S (Pianta, La Paro y Hamre, 2008; Gitomer *et al*., 2015).

Considerando la importancia de la labor que los observadores desempeñan a partir del instrumento, se cuidó contar con personas que cumplieran con el perfil sugerido por los diferentes especialistas que apoyaron el jueceo del instrumento: personas con experiencia docente en educación primaria y experiencia en el campo de la investigación educativa. Se optó por dos observadores, ya que resultados de la teoría de la generalizabilidad sugieren que en este tipo de instrumentos la ganancia al aumentar el número de observadores a más de dos es mínina, y por el contrario los costos en recursos humanos, temporales y económicos son más altos (Shavelson, 1991).

A continuación, se describe a detalle la manera en que se llevó a cabo la capacitación para el uso del instrumento, así como las diferentes fases previas al proceso de calificación de las 40 clases de matemáticas videograbadas.

2.2.1 Capacitación para uso del POAEM

La capacitación de las observadoras consistió en tres fases: a) dos sesiones de trabajo en conjunto; b) sesiones de trabajo individual; y c) evaluación y retroalimentación del proceso de capacitación. Cada fase se describe con detalle a continuación.

2.2.1.1 Sesiones de trabajo en conjunto

Para esta fase se realizaron dos sesiones de trabajo en conjunto, cada una con duración de tres horas aproximadamente. Durante la primera sesión se presentó y explicó el objetivo general del proyecto de investigación y el modelo teórico a partir del cual se desarrolló el POAEM; también se resolvieron dudas sobre el propósito del estudio y la participación de las observadoras para poner a prueba el funcionamiento del protocolo.

Durante la sesión, se entregó a las observadoras una versión impresa del POAEM, cuyos apartados se analizaron y revisaron. Se revisaron las descripciones de las siete dimensiones, las rúbricas y los niveles de desempeño que las conforman. Se identificaron los criterios que conforman los niveles de desempeño en las rúbricas, y los elementos que diferencian los niveles en cada dimensión. Además, se mostraron ejemplos para que las observadoras identificaran criterios relacionados con los niveles de desempeño establecidos en las rúbricas analizadas. Finalmente, se revisó la hoja de calificación y se señaló la forma de llenado, en la que las observadoras asignarían un puntaje y justificarían la calificación de acuerdo con los criterios del nivel de desempeño seleccionado. Dicha información se revisó tanto en formato impreso como en formato electrónico[6], ya que este último se emplearía durante el periodo de calificación de las videograbaciones.

En la segunda sesión de trabajo se recapitularon los elementos específicos a valorar en cada una de las dimensiones del POAEM, se discutieron dudas, se explicó el significado del código de los archivos con videograbaciones, y los segmentos que determinan los tiempos de las tareas matemáticas en cada videograbación de las clases. Se realizó un ejercicio de calificación con una tarea matemática, en el que se dio oportunidad de que revisaran los criterios de cada dimensión antes de asignar los puntajes.

La calificación se hizo sobre una de las tareas matemáticas correspondiente al video "10_AGS_JASB_01_240516_P". Se observó el segmento de la primera tarea matemática sin hacer anotaciones, para identificar aspectos generales de la tarea. Posteriormente, se comentó lo observado y se reprodujo la videograbación para que las observadoras participantes, la autora y tutora de la tesis, emitieran un puntaje y su justificación con base en los criterios correspondientes a las diferentes dimensiones.

Los puntajes y justificación de cada participante se pusieron en común para discutir diferencias y coincidencias, de tal manera que se llegara a un acuerdo con base en los fundamentos teóricos de cada dimensión. En la mayoría de los casos hubo coincidencias

[6] La hoja de calificación se construyó utilizando la plataforma de "Google Formularios". Los ítems empleados incluyeron respuestas de opción múltiple para los puntajes de las observadoras (obligatorio), respuesta abierta para la justificación del puntaje (obligatorio), y un apartado de observaciones generales (no obligatorio) en el que las observadoras pudieran incluir su percepción sobre la tarea matemática valorada.

en los niveles, aunque no en el puntaje, a excepción de las dimensiones exigencia cognitiva, conexión/aplicación de los conocimientos y trabajo colaborativo, por lo que fue necesario regresar a la videograbación para discutir sobre los criterios a valorar y llegar a un acuerdo sobre el puntaje. Al final se solicitó a las observadoras su apreciación respecto a la manera de llevar a cabo la calificación de las videograbaciones, las dificultades que habían enfrentado y las sugerencias para el proceso de la calificación. Ambas observadoras coincidieron en la importancia del ejercicio realizado en la primera sesión de capacitación, en el que identificaron y resaltaron elementos que les permitieron diferenciar los niveles de desempeño en cada rúbrica.

2.2.1.2 Sesiones de trabajo individual

La siguiente fase de capacitación del POAEM se centró en un periodo de estudio individual por parte de las observadoras durante cuatro días. Para ello, se les entregó el protocolo de observación impreso, un anexo con ejemplos de situaciones relacionadas con las diferentes dimensiones a valorar, hojas impresas de calificación y la liga de la hoja electrónica, y un documento impreso con los códigos de los videos a revisar, así como los tiempos correspondientes a las tareas matemáticas que los conforman.

Durante el periodo de trabajo individual se solicitó a las observadoras calificar dos videograbaciones de clases de matemáticas correspondientes a un mismo docente. Cada clase videograbada se seccionó previamente de acuerdo con el número de tareas matemáticas identificadas durante la sesión de clase, con la intención de que ambas observadoras calificaran las mismas tareas matemáticas. En total, las dos observadoras calificaron cinco tareas matemáticas.

Las calificaciones otorgadas por cada observadora, así como su justificación, fueron capturadas por las propias participantes en una hoja electrónica diseñada para este fin. Cabe señalar que las videograbaciones utilizadas durante el periodo de trabajo individual se excluyeron de los 40 videos seleccionados de forma aleatoria para la etapa de pilotaje del protocolo de observación.

2.2.1.3 Evaluación y retroalimentación del proceso de capacitación

La evaluación de las calificaciones correspondientes al periodo de estudio individual se llevó a cabo en dos días, y comprendió la revisión y análisis de cada una de las calificaciones que las observadoras dieron a las tareas matemáticas. Las calificaciones se revisaron por dimensión, contrastando los puntajes de cada observadora con las proporcionadas por las calificadoras expertas, la autora de tesis y una investigadora del departamento de educación que apoyó durante el prepilotaje del instrumento.

Con base en la comparación de los puntajes para cada dimensión valorada, se detectaron aquellas en las que las observadoras mostraron diferencias importantes entre sí y en comparación con los observadores expertos. Con base en dicha información, se estableció una estrategia de retroalimentación, que se llevó a cabo de forma conjunta con las dos observadoras, la tutora, y la autora de este estudio.

La sesión de retroalimentación se realizó en una sesión de tres horas. Durante este tiempo se solicitó a las dos observadoras que respectivamente comentaran las dificultades que enfrentaron durante el periodo de calificación individual, de tal manera que fuese posible identificar si existían coincidencias entre las dificultades. Al respecto, se puede mencionar que las dimensiones en las que se trabajó durante la sesión de retroalimentación fueron: claridad de la tarea matemática, recursos empleados durante la actividad de enseñanza, conexión/aplicación de los conocimientos matemáticos y el trabajo colaborativo.

La estrategia de retroalimentación consistió en observar algunos segmentos de las tareas matemáticas calificadas por las observadoras y analizar lo sucedido durante la clase; después de comentar la situación observada, se les preguntó a las observadoras la calificación que otorgarían. Las discusiones permitieron resolver dudas y llegar a acuerdos sobre lo que se debía identificar al momento de realizar la evaluación de la actividad de enseñanza a partir del protocolo de observación. Además, se enfatizó la importancia de las descripciones de las dimensiones a evaluar en cada rúbrica y las notas aclaratorias, como un paso previo a la calificación de las videograbaciones.

2.2.2 Calificación de videograbaciones

Para la calificación de las clases de matemáticas se seleccionaron de manera aleatoria 40 de las 52 videograbaciones recabadas para este estudio. Tomando como base los videos seleccionados, se diseñó un formato con la siguiente información: código del video, por ejemplo, 01_SFR_GFL_01_070416_P); tema tratado en la clase, por ejemplo, equivalencias; tareas matemáticas que conforman el video a observar; tiempos correspondientes a cada una de las tareas matemáticas identificadas, por ejemplo, Tarea1 V0 00:00 – V2 05: 27; una columna para la firma del observador, por cada video, con la que adquirió el compromiso de mantener en confidencialidad la información registrada (ver apartado de consideraciones éticas y anexo E).

A cada una de las observadoras se le entregó un formato con los videos correspondientes, hojas de calificación impresas, y un disco duro con las clases videograbadas de 20 docentes. Una vez entregado el material y firmadas las cartas compromiso por video, se resolvieron dudas sobre la organización de las carpetas electrónicas con las clases de los 20 docentes y las 40 videograbaciones a calificar. Las indicaciones específicas que se dieron a las observadoras fueron:

- **a.** Cumplir con el periodo de calificación establecido para esta etapa, del 26 de enero al 27 de febrero de 2017. El periodo comprendió la calificación de dos videograbaciones por día, considerando la semana de lunes a viernes. Aproximadamente 3 horas diarias.
- **b.** Cumplir estrictamente con la calificación de dos videograbaciones por día. Con esta restricción, se pretendió cuidar que el cansancio de las observadoras no afectara la objetividad de la evaluación de las dimensiones.
- **c.** Capturar los puntajes y su respectiva justificación en la hoja electrónica correspondiente el mismo día de la calificación.
- **d.** Asistir a sesiones de calibración una vez por semana en los días acordados con la coordinadora del estudio.
- **e.** Consultar dudas u observaciones con la coordinadora del estudio. Para esto, se les ofrecieron diferentes medios de contacto, correo electrónico, llamada telefónica o a través de la aplicación de mensajería instantánea para teléfonos inteligentes (WhatsApp).

Capítulo VI: Resultados

Este capítulo se ha organizado en cinco apartados de acuerdo con los análisis previstos para dar cuenta de evidencias de validez y confiabilidad de la información recabada con el POAEM. El primer apartado incluye información descriptiva como contexto de la interpretación de los resultados obtenidos con los diferentes análisis estadísticos realizados. El segundo apartado presenta datos de validez haciendo uso del Análisis Factorial Exploratorio (AFE); los resultados de este análisis aportaron información relevante para identificar la estructura dimensional del instrumento. El tercer apartado es un Análisis Factorial Confirmatorio (AFC) con el cual se estableció *a priori* la estructura dimensional del POAEM tomando como base los resultados de análisis exploratorio; los resultados permitieron conocer el papel de cada dimensión del POAEM en el conjunto global de la estructura unifactorial. Los índices de ajuste con los que se reportaron los resultados del AFC fueron:

- **GFI** (Goodness of Fit Index). Índice de bondad de ajuste del modelo, con valores entre 0 y 1, en donde el valor mínimo requerido de ajuste del modelo es 0.90.
- **CFI** (Comaparative Fit Index). Índice relativo y de mejor comportamiento con valores entre 0 y 1, en donde el valor mínimo requerido para defender el modelo es 0.90.
- **RMSEA** (Root Mean Square Error of Approximation). Es un índice que evalúa el error de aproximación de acuerdo con los valores observados. Un indicador de buen ajuste son valores menores a .08 (Schumacker y Lomax, 1996; Hu y Bentler, 1999; Martínez, García, Sellés, Valero y Soucase, 2012).

El cuarto apartado presenta la estructura dimensional del POAEM con base en los resultados del análisis factorial exploratorio y confirmatorio, así como en la teoría detrás del diseño del protocolo de observación; el modelo resultante sugiere la existencia de un factor relacionado con los aspectos cognitivos para la resolución de la tarea matemática. El quinto apartado presenta un análisis descriptivo de las coincidencias y diferencias en los puntajes de las observadoras, un análisis de las sesiones de calibración realizadas durante el proceso de calificación, así como el coeficiente Alfa de Cronbach para establecer la consistencia al interior de los subgrupos; también se incluye un estudio de generalizabilidad con el que se presentan diferentes facetas de error de acuerdo con la estructura de los datos del POAEM: docentes, jueces, ocasiones, y las interacciones que se pueden establecer entre dichas facetas.

1. Estadísticos descriptivos del POAEM

Para contextualizar la información que se presenta en este capítulo, la tabla 4 presenta las frecuencias y porcentajes de cada uno de los puntajes que las observadoras otorgaron en las dimensiones que conforman el POAEM. Estos datos se incluyeron como referente para la interpretación de los análisis estadísticos que conforman este capítulo, no para establecer inferencias sobre la población o la muestra.

La base de datos que se empleó para obtener las frecuencias y porcentajes por puntaje (0-8) incluye un total de 222 casos, ya que cada observadora calificó 111 tareas matemáticas. Estas tareas matemáticas fueron la unidad de análisis a partir de la cual se valoraron las actividades de enseñanza con base en el POAEM.

Los resultados muestran una concentración de puntajes en los niveles bajos de la mayoría de las dimensiones valoradas, especialmente en la conexión o aplicación de los conocimiento matemáticos y el trabajo colaborativo. Los puntajes del resto de las dimensiones se distribuye entre los niveles 2 y 4 principalmente, sólo en el caso de alternativas de solución y preguntas para la reflexión hay porcentajes considerables en el nivel 6 (20%).

Con el objetivo de obtener evidencias de validez y confiabilidad de la información recabada con el instrumento, a continuación se incluyen tres análisis estadísticos: un análisis factorial exploratorio, un análisis factorial confirmatorio y un estudio de generalizabilidad.

Tabla 4:
Frecuencias y porcentajes por puntaje y dimensión

	0		**1**		**2**		**3**		**4**		**5**		**6**		**7**		**8**		**TOTAL**	
Dimensión	**F**	**%**	**F**	**%**	**F**	**%**	**F**	**%**	**F**	**%**	**F**	**%**	**F**	**%**	**F**	**%**	**F**	**%**	**F**	**%**
Claridad			7	**3.2**	11	**5.0**	31	**14.0**	108	**48.6**	15	**6.8**	36	**16.2**	6	**2.7**	8	**3.6**	222	100
Exigencia			31	**14**	54	**24.3**	13	**5.9**	57	**25.7**	11	**5**	41	**18.5**	10	**4.5**	5	**2.3**	222	100
Recursos	12	**5.4**	1	**0.5**	12	**5.4**	7	**3.2**	152	**68.5**	13	**5.9**	14	**6.3**	8	**3.6**	3	**1.4**	222	100
Alternativas			34	**15.3**	41	**18.5**	15	**6.8**	38	**17.1**	26	**11.7**	45	**20.3**	10	**4.5**	13	**5.9**	222	100
Preguntas			16	**7.2**	27	**12.2**	29	**13.1**	63	**28.4**	33	**14.9**	43	**19.4**	7	**3.2**	4	**1.8**	222	100
Conexión	64	**28.8**	7	**3.2**	41	**18.5**	19	**8.6**	48	**21.6**	3	**1.4**	21	**9.5**	16	**7.2**	3	**1.4**	222	100
Trabajo	90	**40.5**	16	**7.2**	19	**8.6**	7	**3.2**	18	**8.1**	17	**7.7**	47	**21.2**	4	**1.8**	4	**1.8**	222	100

2. Análisis Factorial Exploratorio

Con la intención de dar cuenta de la agrupación de las dimensiones propuestas a partir del modelo teórico del POAEM, se realizó un Análisis Factorial Exploratorio (AFE) con el que se busca identificar el número y composición de los factores o variables latentes necesarias para explicar la varianza común del conjunto de ítems analizados. Es decir, para el AFE se parte del supuesto que “las variables observadas son indicadores de cierto número de factores o variables latentes comunes… y el ítem es una manifestación de ese factor” (Lloret, Ferretes, Hernández y Tomás, 2014, pp. 1152).

Para este análisis se utilizaron dos bases de datos, una con 111 casos aleatorios de un total de 222 (total de tareas matemáticas calificadas por ambas observadoras) y otra con un promedio por docente con 20 casos (total de docentes participantes). Para obtener la primera base de datos se utilizó el programa estadístico SPSS con el que se realizó la selección de los casos de forma aleatoria. El propósito de este tipo de base fue que se tuviera información disponible para llevar a cabo tanto el AFE como el Análisis Factorial Confirmatorio (AFC) con dos bases de datos distintas pero con la misma cantidad de casos.

Para crear la base de datos con el promedio por docente (20 casos) se sumó la calificación de cada tarea valorada en los dos días que se observó a cada docente y el resultado se dividió entre el número total de tareas calificadas en los dos días. El procedimiento se repitió para cada uno de los 20 docentes participantes. El propósito de incluir los análisis con ambas bases de datos, por tarea y por docente, es aportar más elementos que permitan confirmar la información obtenida a partir de los análisis estadísticos propuestos.

A continuación se presentan los resultados obtenidos a partir del AFE con cada una de las bases de datos utilizadas. En ambos casos se utilizó una matriz de correlaciones producto-momento de Pearson, varianza-covarianza,, con la que se evalúa la relación lineal entre dos variables continuas y preferentemente de distribución normal. Se empleó esa matriz debido a que los ítems del POAEM son ordinales, por lo que la continuidad de las variables es permite que la distribución de los datos se aproxime a la normalidad (Lloret, Ferretes, Hernández y Tomás, 2014).

El primer paso para realizar el análisis factorial fue comprobar el grado de adecuación para dicho análisis a partir de la prueba KMO (Kaiser, Meyer y Olkin), que permite conocer si el tamaño de la muestra es adecuado, así como a través de la prueba de esfericidad de Bartlett que nos verifica si la matriz de correlaciones es una matriz identidad, lo que indicaría que el modelo factorial es inadecuado.

Para la prueba KMO se consideró como valor estadístico aceptable un valor mayor a 0.50 y en el caso de la esfericidad de Bartlett, se buscó rechazar la Ho, por lo que se consideró como valor apropiado un alfa menor o igual a 0.05 de acuerdo con Hair, Anderson, Tatham y Black (2005). El punto de saturación para determinar las variables que cargan en los diferentes factores fue un valor igual o mayor a .40 de acuerdo con los criterios de Bandalos y Finney (2010).

La tabla 5 incluye información de la prueba KMO y esfericidad de Bartlett para las dos bases de datos utilizadas. En ambos casos los valores obtenidos son adecuados para realizar el AFE, ya que la medida de adecuación KMO fue mayor a .50 y el valor de significancia en la prueba de esfericidad fue menor a .05. Esta última significa que la matriz de correlación es adecuada para realizar el análisis factorial.

Tabla 5:
Prueba de KMO y prueba de Bartlett

(N= 111) KMO y prueba de Bartlett		**.816**	**(N= 20) KMO y prueba de Bartlett**		**.752**
Prueba de esfericidad de Bartlett	Chi-cuad	204.295	Prueba de esfericidad de Bartlett	Chi-cuad	68.727
	gl	21		gl	21
	Sig.	.000		Sig.	.000

Una vez comprobada la factibilidad para realizar el análisis factorial con los datos, se empleó el programa *R Factor para análisis factoriales* mediante el paquete SPSS, con el que es posible realizar distintas técnicas de extracción, rotación y selección de factores. El método de extracción utilizado fue el de factorización de ejes principales, con una rotación oblicua (Quartimin), ya que de acuerdo con Basto y Pereira (2012) este tipo de rotaciones genera soluciones más confiables y además permite simplificar y clarificar la estructura de los datos.

La tabla 6 presenta las cargas factoriales para las siete variables que conforman el POAEM utilizando las dos bases de datos. Los resultados de este modelo factorial muestran que en ambos casos el comportamiento de las cargas excluye la variable claridad del primer factor. Con la base N=111 porque hay una mayor saturación en el factor 2 y con N=20 porque las cargas tienen valores muy similares en ambos factores. Esto muestra la necesidad de realizar pruebas estadísticas que ofrezcan mayor información sobre el número de factores a retener.

Tabla 6:
Análisis Factorial Exploratorio (todas las variables)

Ítem	(111 casos)		(20 casos)	
	F1	F2	F1	F2
Claridad	.449	**.682**	**.605**	**.646**
Exigencia	**.746**	-.034	**.767**	-.364
Recursos	**.658**	.282	**.732**	.402
Alternativas	**.668**	-.508	**.700**	-.545
Preguntas	**.769**	-.018	**.886**	.043
Conexión	**.554**	-.497	**.811**	-.237
Trabajo	**.788**	.205	**.837**	.155

La tabla 7 presenta el porcentaje de varianza para los dos modelos de AFE incluyendo las siete variables que integran el POAEM. Para el primer modelo (N=111) se observa que la varianza que se explica con la solución factorial final es de 60%, mientras que la solución final del segundo modelo (N=20) explica el 74.5%. Sin embargo, para determinar el número correcto de factores a retener a partir del AFE se utilizaron diferentes métodos: regla de Kaiser (autovalores), análisis de paralelo, coordinación óptima, factor de aceleración, y prueba de Velicer (Figura 7).

Tabla 7:
Porcentaje de varianza explicada y autovalores del POAEM (todas las variables)

AFE (111 Casos)				AFE (20 Casos)			
Factor	Autovalores	% varianza	Acumulada	Factor	Autovalores	% varianza	Acumulada
1	3.155224	45.074629	45.074629	**1**	4.125481	58.935444	58.935444
2	1.091879	15.598273	60.672901	**2**	1.090235	15.574792	74.510236
GFI: .899				**GFI: .959**			

Los resultados de los diferentes métodos de extracción de factores para ambas bases de datos son coincidentes con los datos que se muestran en la figura 7. En este caso la regla de Kaiser (autovalores) retiene dos factores, mientras que el análisis paralelo establece el punto de corte en un factor, al igual que la coordinación óptima, la prueba Velicer, y el factor de aceleración.

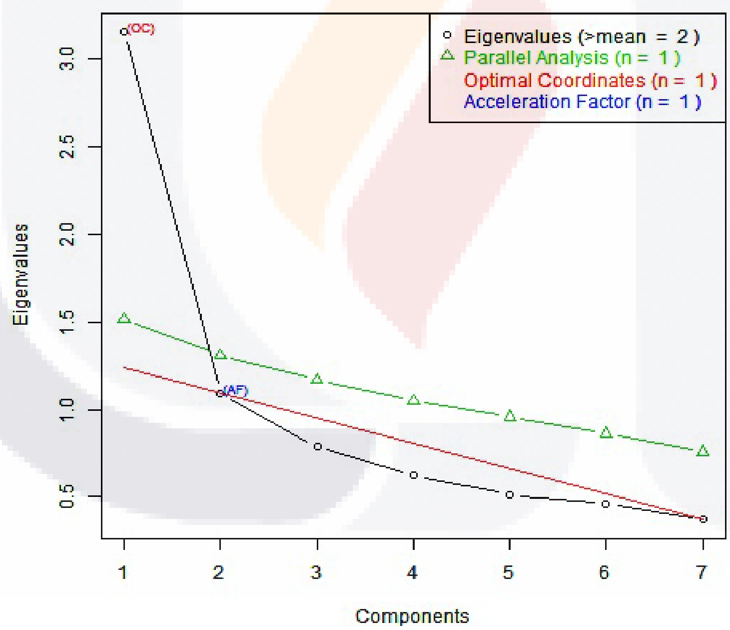


Figura 7. Estimación del número de factores a partir de cuatro métodos (todas las variables)
Fuente: Elaboración propia a partir de la base de datos

La bondad de ajuste para modelos del AFE y los respectivos métodos de extracción fue de .899 para el modelo con N=111 y .959 para el modelo con N=20 de acuerdo con el *Goodness of Fit Index* (GFI). Este índice considera los valores mayores a .90 como buen ajuste del modelo y los valores mayores a 0.95 como muy buen ajuste (Basto y Pereira, 2012). Con base en lo anterior, el mejor ajuste se identificó con el modelo N=20. Sin embargo, para obtener mayor información sobre las cargar factoriales y la retención de factores se realizó el mismo procedimiento con las dos bases de datos excluyendo la variable claridad, considerando que dicha variable podría estar midiendo aspectos distintos a los que miden el resto de las variables.

La tabla 8 muestra las cargas factoriales con los dos modelos para el AFE excluyendo la variable claridad. Los resultados del primero modelo (N=111) sugieren un factor en el que se agrupan las seis variables de este modelo. El modelo dos (N=20) sugiere dos factores, aunque los valores más grandes de las cargas factoriales se concentran en el primer factor, sólo en el caso de la variable recursos hay una carga considerable en el factor dos, sin embargo la mayor carga permanece en el primer factor.

Tabla 8:
Análisis Factorial Exploratorio (Sin variable Claridad)

Ítem	**(111 casos)**	**(20 casos)**	
	F1	**F1**	**F2**
Exigencia	.743	.788	-.391
Recursos	.644	.696	**.501**
Alternativas	.710	.750	-.456
Preguntas	.779	.882	.176
Conexión	.573	.828	-.193
Trabajo	.777	.833	.368

La tabla 9 presenta el porcentaje de varianza para los dos modelos de AFE con seis variables, sin la variable claridad. Los resultados para el primer modelo (N=111) explica el 50% de la varianza con la solución factorial final. Es decir, en comparación con el modelo que incluyó la variable claridad, el modelo que la excluye explica mayor proporción de varianza. La misma situación se observa con el segundo modelo (N=20), ya que explica hasta el 63% con el primer factor y 77% con dos factores, aunque el segundo factor no es factible por las cargas factoriales negativas o positivas muy bajas que genera.

Tabla 9:
Porcentaje de varianza explicada y autovalores del POAEM (Sin variable claridad)

AFE (111 Casos)				AFE (20 Casos)			
Factor	Autovalores	% varianza	Acumulada	Factor	Autovalores	% varianza	Acumulada
1	3.009643	50.160711	**50.160711**	1	3.824550	63.742505	**63.742505**
2	0.923707	15.395109	65.555820	2	0.814505	13.575087	77.317591
GFI: .899				**GFI: .959**			

Con base en la información obtenida con los modelos excluyendo la variable claridad se realizó el procedimiento para determinar el número correcto de factores a retener, utilizando los métodos: regla de Kaiser (autovalores), análisis de paralelo, coordinación óptima, factor de aceleración, y prueba de Velicer (Figura 8). Los resultados para ambas bases de datos (N=111 y N=20) coinciden en que se debe retener un factor para las seis variables, tal como se muestra en la figura 8, ya que en todos los casos los métodos retienen un factor.

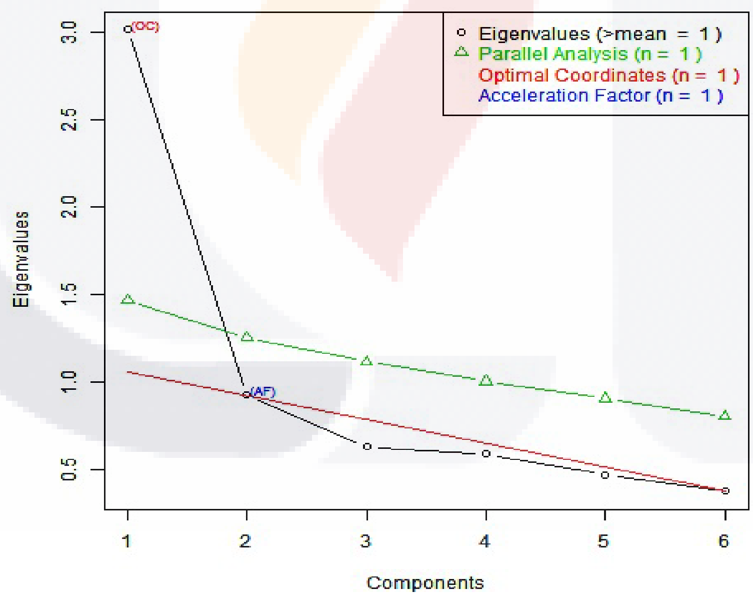


Figura 8. Estimación del número de factores a partir de cuatro métodos (Sin variable claridad)
Fuente: Elaboración propia a partir de la base de datos

La bondad de ajuste para los dos modelos de AFE que excluyen la variable claridad reporta los mismos valores que los modelos anteriores, esto es un valor GFI = .899 para el modelo con N=111 y GFI= .959 para el modelo con N=20. Los resultados en ambos casos apuntan hacia una estructura unifactorial para el POAEM. Sin embrago, con el propósito de tener mayor certeza de esta decisión se contrastó el modelo unifactorial con un Análisis Factorial Confirmatorio (AFC) para conocer el papel que juega cada una de las dimensiones en el conjunto global de la estructura propuesta en el modelo (varianza total explicada por el factor, varianza y saturación de los ítems).

3. Análisis Factorial Confirmatorio

El AFC que se presenta a continuación se realizó con el programa estadístico R y las librerías "Psych" y "Sem". Este análisis contrasta el modelo unifactorial que surgió a partir del AFE con la base de datos aleatorios N=111 (de un total de 222) excluyendo la variable claridad. Los resultados obtenidos muestran que la estructura dimensional para el POAEM se concentra en un factor tal como se señaló con el factorial exploratorio, de acuerdo con los valores obtenidos para el índice de bondad de ajuste, el índice de ajuste comparativo y el error aproximado obtenido.

La tabla 10 concentra la información sobre el valor estimado para las cargar factoriales de cada una de las seis variables que conforman el modelo a probar, el error correspondiente, así como los resultados para diferentes índices de ajuste. En todos los casos se cumple con los supuestos para aceptar el modelo unifactorial propuesto, ya que los índices GFI= 0.96 y CFI= 0.96 presentan un excelente ajuste, y el índice de error aproximado RMSEA=0.06 es poco menor al valor de referencia. Esto significa que la estructura unifactorial para el POAEM señalada con el AFE se confirma con el AFC.

Tabla 10:
Estimación de parámetros

	Estimado	**Error estándar**	**Valor - z**	**Pr(>|z|)**	
Lam1	1.0447533	0.1880305	5.556297	2.755583e-08	exigencia <--- f1
Lam2	0.8386852	0.1470194	5.704590	1.166232e-08	recursos <--- f1
Lam3	1.3707566	0.2052748	6.677666	2.427779e-11	alternativas <--- f1
Lam4	0.9411524	0.1608615	5.850700	4.895075e-09	preguntas <--- f1
Lam5	1.1781422	0.2373110	4.964550	6.886051e-07	conexion <--- f1
Lam6	1.3976662	0.2587482	5.401647	6.603176e-08	trabajo <--- f1
e1	2.3391224	0.3732836	6.266342	3.696292e-10	exigencia <--> exigencia
e2	1.4081542	0.2277144	6.183861	6.255261e-10	recursos <--> recursos
e3	2.4143934	0.4399225	5.488224	4.059943e-08	alternativas <--> alternativas
e4	1.6588692	0.2720637	6.097356	1.078374e-09	preguntas <--> preguntas
e5	3.9310656	0.6001656	6.549968	5.754954e-11	conexión <--> conexión
e6	4.4971421	0.7085337	6.347111	2.193956e-10	trabajo <--> trabajo

Model Chisquare = 13.21541 Df = 8 Pr(>Chisq) = 0.1530988

RMSEA = 0.06525331 90% CI: (NA, 0.1352021)

CFI = 0.9630051

La figura 9 resume las cargas factoriales, el error estimado y los índices de ajuste presentados en la tabla 9 para llevar a cabo el AFC con el modelo unifactorial propuesto. La dirección de las flechas en la figura indica que es el factor el que genera las puntuaciones en el ítem, por lo que el grado en que su relación con el factor es mayor, el ítem mide adecuadamente el factor al que pertenece (Herrero, 2010; Fernández, 2015). En este caso, la mayoría de los ítems presentan cargas factoriales altas, por lo que la relación entre el constructo y las variables muestra valores apropiados.

La puntuación en cada ítem es generada por un factor latente, que es la variable no observada y explica la variabilidad de las puntuaciones; sin embargo, no explica la totalidad de la variabilidad de las respuestas. La parte no explicada se denomina como error de medida (E). Al igual que con otros análisis, este tipo de varianza puede deberse a diferentes circunstancias, como la naturaleza del instrumento, el constructo que se intenta medir, u otros aspectos no relacionados con cuestiones aleatorias.

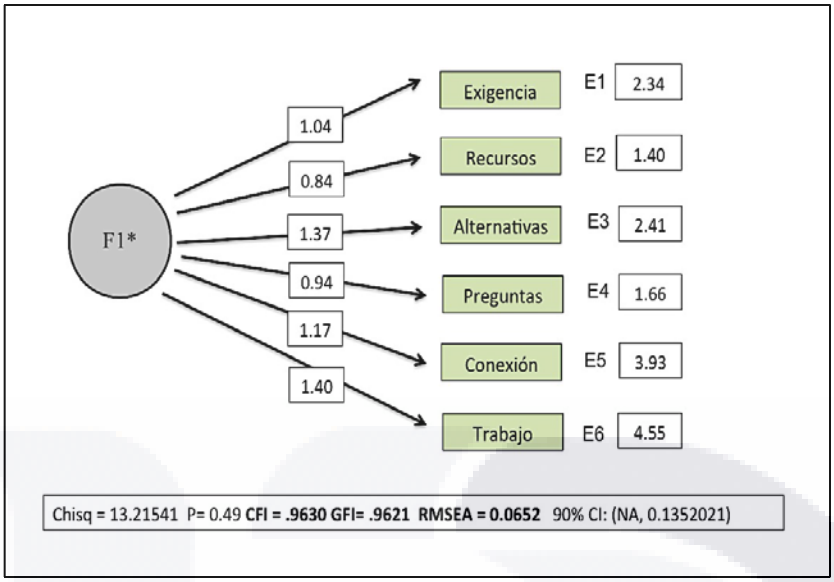


Figura 9. Modelo unifactorial.
Fuente: Diseño propio a partir de la estimación de parámetros.

Tomando como base la información que aportaron el AFE y el AFC es posible mencionar que la estructura dimensional del POAEM es unifactorial. Este factor concentra seis de las siete variables que se habían incluido en el modelo teórico del instrumento, y en conjunto miden aspectos cognitivos relacionados con las actividades de enseñanza en matemáticas. Es decir, con este factor es posible medir algunos aspectos sobre el nivel de pensamiento que requieren poner en juego los estudiantes para lograr un razonamiento de las situaciones matemáticas, cómo establecen diferentes estrategias de solución, los recursos que emplean para ello, y cómo construyen el conocimiento en conjunto (Stein y Smith, 1998; Barreiro y Casetta, 2012; NCTM, 2015). La siguiente tabla describe las seis variables que integran este factor.

Tabla 11:
Estructura unifactorial del POAEM: Aspectos cognitivos

Exigencia cognitiva	Tipo y nivel de pensamiento requerido para realizar una tarea matemática.
Alternativas de solución	Flexibilidad del docente para promover y aceptar una variedad de estrategias y procedimientos para resolver un problema matemático.
Recursos empleados	Uso y aplicación de artefactos coherentes con el propósito de la tarea y su aporte para promover o fortalecer los aprendizajes matemáticos.
Preguntas para la reflexión	Tipo de preguntas que se plantean durante una tarea para dar sentido a ideas o relaciones matemáticas importantes.
Conexión/ aplicación del conocimiento	Contexto de una situación problemática para conectar los aprendizajes matemáticos con la experiencia, otras disciplinas o con situaciones de la vida cotidiana.
Trabajo colaborativo	Interacción, negociación y cooperación entre los estudiantes para desarrollar una tarea matemática (nivel de exigencia cognitiva).

Es importante mencionar que la estructura dimensional resultante es distinta a la que se había propuesto a partir de la teoría. La figura 10 presenta la estructura dimensional previa a los análisis factoriales con cuatro factores.

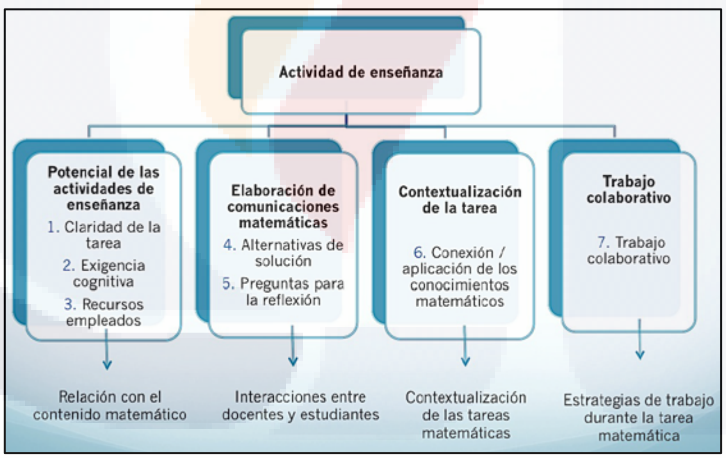


Figura 10. Modelo teórico para evaluar las actividades de enseñanza
Fuente: Diseño propio a partir de la revisión de literatura.

El modelo teórico propuesto fue refutado después de llevar acabo los análisis estadísticos correspondientes, ya que todas las variables se agrupan en un factor a excepción de la variable claridad. Con esta última variable se pretendía obtener información sobre la manera en que los docentes comunican el trabajo a los estudiantes y cómo indagan si han comprendido lo que se les solicitó como trabajo o producto. Aspectos distintos de los que se intenta indagar con las seis variables restantes.

4. Confiabilidad de las dimensiones

Para contextualizar los diferentes análisis estadísticos que guían este apartado, a continuación se presenta un análisis descriptivo con los puntajes que las observadoras dieron a las tareas matemáticas valoradas con el POAEM. Para ello, se contabilizaron las diferencias y coincidencias de los puntajes en cada dimensión. Se estableció como puntaje cero las discrepancias entre ambas observadoras, y como 1 las coincidencias exactas para obtener un conteo total de acuerdos y de su porcentaje. Además se calculó el acuerdo relativo, ampliando el margen de error a ±1 punto, ya que la escala de calificación para las rúbricas del POAEM van de 1-8 puntos, distribuidos en dos puntajes por nivel de desempeño (bajo 1,2; medio 3,4; medio-alto 5,6; y alto 7,8).

Las Tablas 14-20 presentan las frecuencias de los puntajes para ambas observadoras, así como un gráfico de dispersión en cada caso. Las tablas y los gráficos incluyen un rango de 0-8 puntos por observadora, ya que en los conteos se incluyó las respuestas *no aplica* y no se puede valorar. De acuerdo con la información obtenida, las dimensiones con mayor concentración de acuerdos son: recursos empleados durante la tarea matemática (62) y claridad de la tarea (42). Por el contrario, las dimensiones con menos coincidencias son: alternativas de solución (20), exigencia cognitiva (22) y preguntas para la reflexión (22).

Para explorar las circunstancias que pudieron influir en los acuerdos y desacuerdos de las observadoras, se muestra a continuación información sobre cuatro sesiones de calibración que se llevaron a cabo durante el proceso de calificación con el POAEM. La calibración consistió en monitorear el trabajo de calificación realizado por las observadoras e identificar las dimensiones en que requerían mayor entrenamiento. Los elementos a revisar en cada sesión se obtuvieron a partir de la comparación de las calificaciones otorgadas por ambas observadoras, así como con las asignadas por la autora de la tesis, considerada en este proceso como una calificadora experta. El proceso que se siguió en cada sesión se describe a continuación:

1. Se explicó a las observadoras el trabajo de análisis de las calificaciones realizado previo a las sesiones de calibración y el propósito del mismo: monitorear las calificaciones otorgadas y detectar dimensiones que requirieran aclaraciones.
2. Se preguntó a las observadoras sobre las dimensiones en las que tuvieron mayor dificultad (como punto de comparación con dimensiones detectadas previamente). Además se pidió a las observadoras explicar el porqué de sus dificultades.
3. Se revisó y discutió en conjunto la descripción de las rúbricas en las que se detectó mayor dificultad.
4. Se revisó y discutió la sección de ejemplos por dimensión y niveles de desempeño de las rúbricas detectadas con mayor dificultad.
5. Se analizaron diferentes fragmentos de videograbaciones para ejemplificar criterios de las rubricas del POAEM, así como para discutir los puntajes y calificaciones otorgadas por las observadoras y llegar a acuerdos sobre los criterios observados.

La Tabla 12 muestra la distribución de dimensiones que se trabajó con cada observadora durante las sesiones de calibración.

Tabla 12:
Dimensiones revisadas / Sesión de calibración

Sesión 1	
Observador 1	**Observador 2**
Claridad de la tarea matemática	Exigencia cognitiva
Recursos empleados	Recursos empleados
	Alternativa de solución
Sesión 2	
Observador 1	**Observador 2**
Conexión/ aplicación de los conocimientos	Conexión/ aplicación de los conocimientos matemáticos
Exigencia cognitiva	Trabajo colaborativo
Alternativas de solución	Alternativas de solución
Sesión 3	
Observador 1	**Observador 2**
Preguntas para la reflexión	Exigencia cognitiva
Trabajo colaborativo	Trabajo colaborativo
Alternativas de solución	Alternativas de solución
Conexión/ aplicación del conocimiento	Conexión/ aplicación del conocimiento
Sesión 4	
Observador 1	**Observador 2**
Alternativas de solución	Alternativas de solución
Exigencia cognitiva	Recursos empleados

Tomando como base las frecuencias en los puntajes de ambas observadoras, así como los gráficos de dispersión, a continuación se presenta información relevante para cada dimensión y se detallan las dificultades que las observadoras mencionaron durante la etapa de calibración, así como las estrategias que se generaron para apoyar un mejor desempeño durante el proceso de calificación.

Claridad de la tarea matemática.

En esta dimensión se pidió a las observadoras valorar la manera en que los docentes comunican el trabajo que los estudiantes deben realizar y cómo indagan si los estudiantes comprendieron lo que se les solicitó, sea una meta específica o un producto. Los resultados de las frecuencias muestran que en esta dimensión los puntaje se centraron en 4,4 y 3,4, ambos correspondientes al nivel medio de esta rúbrica. Esto es “El docente plantea la tarea matemática y señala a los estudiantes los productos que espera logren. Enfatiza conceptos o ideas matemáticas importantes para el desarrollo de la tarea”.

Algunos desacuerdos importantes se identificaron con puntajes de 0,4; 0,5 y 8,4, lo que sugirió problemas con la comprensión del propósito de la rúbrica, ya que en los tres casos había elementos para identificar aspectos relacionados con la comunicación del trabajo a realizar. Para ofrecer un panorama general de los resultados obtenidos, la tabla 13 presenta la distribución de las calificaciones de ambas observadoras y la figura 11 un diagrama de dispersión de esas calificaciones. En ambos casos es posible observar que los niveles extremos, bajo y alto, causaron mayor dificultad a las observadoras.

Tabla 13:
Claridad de la tarea matemática

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	Total
0					4	1				5
1					1					1
2		1	1							2
3			3	1	23	2				29
4			3	1	29	2	3			38
5					4		1			5
6			2		5	4	9			20
7					3	1	1			5
8					1		2	1	2	6
Total		**1**	**9**	**2**	**70**	**10**	**16**	**1**	**2**	**111**

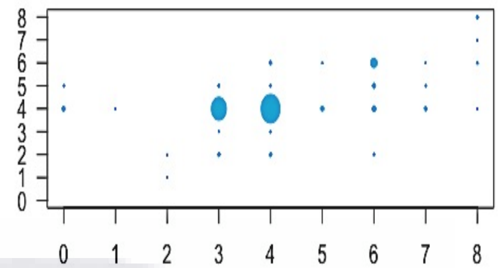


Figura 11. Dispersión de puntajes. Claridad de la tarea

Con base en los comentarios de las observadoras sobre esta dimensión, a continuación se mencionan algunos ejemplos de las dificultades que se presentaron durante el proceso de calificación.

Una de las situaciones que se presentó durante la identificación de los niveles bajos de desempeño, al respecto una de las observadoras seleccionó la opción no se puede valorar (puntaje cero) para varias de las tareas designadas en la primera semana de calificación. Algunas de las anotaciones que hizo la observadora para justificar esta opción fueron:

> El docente no da una instrucción, va induciendo lo que se verá en la clase
> No hay una explicación. Parte de un trabajo previo, se trata de corregir
> El docente no da indicaciones para ninguna tarea, sondea conocimientos previos sobre el tema
> En el video no se observa la instrucción, aunque más adelante menciona que están marcando las partes del círculo

Para abordar esta situación se revisaron algunos fragmentos de videograbaciones previamente calificadas por las observadoras y se analizó la descripción de la dimensión a valorar. Además, se aclaró que la claridad de la tarea es una dimensión dinámica, por lo que se podría identificar en diferentes momentos de la tarea observada, no sólo al inicio. Por lo tanto, se enfatizó la importancia de diferenciar entre el aprendizaje esperado que frecuentemente se menciona al inicio de una tarea y la identificación de la forma en que el docente comunica a los estudiantes lo que deben hacer para desarrollar algún producto o tarea, y cómo indaga si han comprendido lo que se les ha solicitado.

Los siguientes dos fragmentos de videos calificados también corresponden a situaciones en donde alguna de las observadoras marcó la opción *No se puede valorar*. Estos fragmentos se retomaron durante la sesión de calibración para identificar criterios incluidos en la rúbrica correspondiente y señalar a las observadoras que existen elementos para emitir una calificación correspondiente al nivel medio de acuerdo con los criterios de la rúbrica claridad de la tarea matemática.

Video: 01_SFR_02_T1 (Nivel medio)

“Sistema métrico decimal”

> D[7]: Muy bien, vamos a empezar… vamos a retomar lo que vimos el día de ayer.. ¿Qué fue lo que vimos el día de ayer?
> AA: El sistema métrico decimal…
> D: El día de ayer vimos el sistema métrico decimal de medidas, específicamente de longitud; porque son tres los que vamos a ver… longitud, peso y volumen… En este caso estamos viendo lo de longitud… y el día de ayer vimos lo que significa cada una de las abreviaturas que manejamos en nuestro sistema. Nuestro sistema se basa principalmente en el metro… el metro es el punto de partida para obtener los submúltiplos y múltiplos… ¿Quién me quiere recordar cuales son los múltiplos (del metro)?
> A: yo, yo…
> AA: hectómetro, decámetro… (contestan varios estudiantes a la vez)
> D: A ver… (pregunta a uno de los estudiantes)…
> AA: contestan varios al mismo tiempo…

El fragmento muestra una situación en la que la tarea matemática es continuación de otra realizada en la clase anterior. Por lo que se puede observar que el docente retoma elementos y enfatiza conceptos o ideas matemáticas importantes para el desarrollo de la nueva tarea, es decir hay elementos que coinciden con el nivel medio-bajo de esta rúbrica.

Otra de las situaciones en que una de las observadoras marcó la opción no se puede valorar tiene que ver con el siguiente ejemplo en el que el docente plantea una tarea sin enfatizar conceptos o ideas matemáticas importantes, pero en la que se dan indicaciones e información como preámbulo de lo que realizarán los estudiantes a partir de la tarea

[7] D: Docente
A: Un alumno o alumna hablando
AA: Todos los estudiante o la mayoría

matemática. Por lo que existen elementos para utilizar la rúbrica y calificar el nivel de desempeño de acuerdo con los criterios que la integran.

02_SFR_01_T1 (Nivel bajo)
“Gráfica de barras”

> D: Por favor sacamos nuestro cuadernillo de trabajo y libro de texto… aparte del libro de matemáticas vamos a sacar nuestro libro de geografía… Porque vamos a iniciar con nuestro tema que es “Gráfica de barras”… y el propósito… la intención didáctica es que analicemos los datos que contienen las gráficas.... y de igual forma vamos a analizar qué datos tienen las gráficas… Muy bien, ¿alguien me puede responder qué es una gráfica?... levantando su mano
> A: Es donde se ponen porcentajes de algo que se hace, por ejemplo de una encuesta
> D: ¿Alguien que me de otra definición?,
> A: Es a qué le dedicamos más tiempo…
> D: a ver.. gráfica, palabra gráfica… a ver.. ¿qué es una gráfica?...

Exigencia cognitiva.

Esta dimensión implicó un proceso complejo de calificación dada su naturaleza teórica. Las observadoras tuvieron que identificar el tipo de pensamiento requerido para realizar la tarea matemática de acuerdo con la descripción de los niveles de desempeño de la rúbrica (taxonomía de tareas matemática). Los acuerdos para esta dimensión se concentraron en los puntajes 2,2 y 4,4, es decir, en el nivel bajo y medio de la rúbrica. Esto es, “La tarea es algorítmica, requiere de procedimientos específicos para trabajar alguna idea o concepto matemático. Los estudiantes deben identificar, reconocer, o asociar algún procedimiento determinado”.

La tabla 14 concentra los puntajes de ambas observadoras y la figura 12 un diagrama de dispersión de esos puntajes. Los resultados muestran que hubo un gran número de desacuerdos entre las observadoras. Esto pudo deberse a la naturaleza de la dimensión (nivel de pensamiento), ya que en algunos casos las observadoras confundieron dificultad de la tarea matemática con exigencia cognitiva.

Tabla 14:
Exigencia cognitiva

	0	1	2	3	4	5	6	7	8	Total
0		2	1							3
1		3	1		6		6			16
2		5	8	1	6		7			27
3		1	1		7		2			11
4		1	9		6		4		1	21
5			3	1	5		2			11
6			2		3		5		1	11
7			2		3		3		1	9
8							1	1		2
Total		**12**	**27**	**2**	**36**		**30**	**1**	**3**	**111**

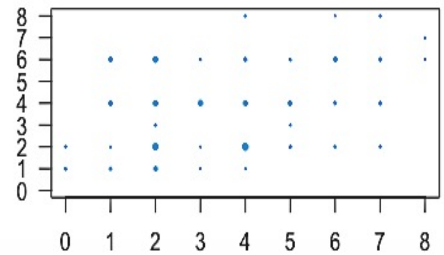


Figura 12. Dispersión de puntajes. Exigencia cognitiva

Los diferentes problemas detectados durante el proceso de calificación se abordaron en las sesiones de calibración con nuevos ejercicios, especialmente porque las observadoras estaban valorando con niveles bajos la mayoría de las tareas matemáticas. Por lo tanto, se solicitó que justificaran la selección de sus puntajes en algunas de las tareas revisadas en conjunto. En una de las ocasiones, una observadora mencionó:

> Le di un puntaje de uno porque considero que al tener que realizar un problema matemático que cumplía con ciertos aspectos matemáticos que la maestra les está pidiendo se considera una actividad de no mucha exigencia cognitiva, porque implica que los estudiantes sigan ciertos pasos ya establecidos…

14_AGS_01_T3 (Nivel alto)
"Elaborar un problema matemático"

> D: Bien, en esta hojita van ustedes a elaborar un problema… fíjense bien... El resultado debe ser en fracción y podemos utilizar también como datos fracciones con numerador del 1 al 9, denominador, podemos manejar del 2 al 9, ¿sí? Adelante y vamos a plantear un problema en esta hojita.. pero también va ir dibujo…

El análisis y discusión permitió llegar al acuerdo de que este tipo de tareas realmente requieren que los estudiantes pongan en juego diferentes habilidades intelectuales que les permitan demostrar cierto dominio del concepto y noción de fracción, así como otros elementos para plantear una situación problemática adecuada al tema y a las restricciones propuestas por la docente. Por lo tanto, de acuerdo con los criterios descritos en la rúbrica de esta dimensión, la tarea requiere un pensamiento complejo, ya que los estudiantes deben diseñar y representar conceptos, algoritmos o procedimientos para establecer soluciones y lograr un entendimiento de la situación propuesta, que se refleja en el nivel alto de esta rúbrica.

Otro aspecto importante fue identificar cuándo las tareas propuestas por los docentes implican un proceso de mecanización o repetición. Para ello, se seleccionaron ejemplos en las videograbaciones para detectar los criterios que se incluyen en el nivel bajo de esta rúbrica "La tarea se limita a promover la reproducción de hechos, reglas, formulas, o definiciones previamente aprendidas".

El siguiente fragmento se discutió con la intención de identificar el tipo de tarea matemática que propone el docente a los estudiantes. Específicamente se analizó el nivel de exigencia cognitiva que representa esta tarea para los estudiantes. Se acordó que por la dirección del docente y por los ejercicios que implican repetir un patrón, este tipo de tareas se apega a los criterios que se describen en el nivel bajo de la rúbrica.

04_JM_01_T2 (Nivel bajo)

"Sucesiones numéricas"

D: Ahí están las sucesiones niños (7, 21_,189,_), si se fijan tenemos dos líneas ¿qué significa eso niños?
A: Que está incompleta…
D: Que está incompleta, ¿que faltan los qué niños?
A: Los números…
D: Dos cifras verdad, pero ustedes las van a encontrar… ¿Cómo las encontrarían? A ver…
A: primero tendría que saber cómo multiplico para que me de el resultado de 21 y luego encontrar el otro…
D: Dice.. que primero multiplicar para que le de el 21 y poder encontrar… ¿Cuál?.. quién encontró ya, ese número… 7X3?
AA: 21
D: Entonces el 21 niños, ¿por qué lo vamos a multiplicar?
AA: por 3
D: dijimos que lo íbamos a multiplicar 21X3… multiplíquenlo ustedes en la calculadora niños…
A: Profe, ya tengo la siguiente…
D: ¿Qué les salió? ¿63?
AA: sí…
D: Ahora vamos a multiplicar el 63 por?
AA: por tres
D: Bien… ¿El patrón cuál es niños?
AA: 3…
D: Bien, vamos con la siguiente sucesión (6, 36, 216_)… (anota otro ejercicio en el pizarrón)… ¿Quién y encontró cuál es el patrón?... Anota en el pizarrón otras sucesiones con el mismo patrón.

Este fragmento necesitó un proceso de análisis y discusión con una de las observadoras, ya que a pesar de que se habían revisado los criterios de la rúbrica, así como otros ejemplos, la observadora señalaba que en este caso en particular, había elementos para aumentar el nivel de la exigencia cognitiva, empleando como justificación la siguiente:

> … los estudiantes deben resolver problemas y analizar un procedimiento para determinar si su respuesta es correcta, yo lo que veo es que la hojita tenía más problemas, al final se centran en la revisión de uno… pero al final no sabemos si los problemas eran similares

Para abordar este caso, al igual que con otros ejemplos se analizó por segunda ocasión la videograbación correspondiente. A continuación se presenta una fragmento de la segunda trascripción:

13_AGS_01_T2 (Nivel bajo)

"Problemas de reparto"

> D: Les voy a entregar una hoja, la cual van a resolver de forma individual (entrega)… Pegamos la hoja y contestamos… (después de un tiempo). A ver ¿quién nos quiere compartir?
> A: Yo… (un estudiante)
> D: ¿A ver primero qué encontraste en tu equipo?, Ok! Vamos a ver…
> AA: sí… (responden en coro)
> D: A ver.. ¿primer problema?
> A: Francisco tiene tres pasteles y los quiere repartir entre cuatro personas, ¿qué cantidad de pastel le toca a cada quién?
> … Los siguientes problemas hacen referencia al reparto de gelatinas y chocolates

Los elementos que permitieron llegar a un acuerdo tienen que ver con las situaciones problemáticas que revisaron los estudiantes después de la plenaria, ya que se trataba de situaciones similares a la revisada previamente, en la que se planteaban problemas de reparto con gelatinas, chocolates y pasteles. Por lo tanto, se trataba de actividades repetitivas. Después de analizar la videograbación, la misma observadora mencionó:

> Ok, sí finalmente sigue reparto y reparto bajo el mismo concepto, o sea no les pide a lo mejor ahora que lo hagan con suma de fracciones… Entonces podría quedar más o menos una nota, en que si las situaciones son repetitivas le pega a la exigencia cognitiva, ¿verdad?

Recursos empleados durante la actividad de enseñanza.

La valoración de esta dimensión involucró el reconocimiento del uso que el docente y los estudiantes hacen de los artefactos o recursos disponibles para promover los aprendizajes matemáticos (libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, etc.). Específicamente de la coherencia de éstos con el propósito de la tarea, el análisis o discusión que favorecen y si permiten comprobar procedimientos y estrategias. Al respecto se puede observar que la mayoría de los acuerdos en esta dimensión se encuentran en los puntajes 4,4, representados en el diagrama de dispersión con un círculo mayor que el resto de la distribución de puntajes (Figura 13). No obstante, al igual que en otras dimensiones hay desacuerdos con puntajes extremos, como 0,5 y 4,8, aunque en esta dimensión con menor frecuencia que las anteriores (Tabla 15).

Tabla 15:
Recursos empleados

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	Total
0	3				1	1				5
1			1							1
2										
3			1	1	1					3
4	4		8	3	57	5	7	3	1	88
5			1		3		2			6
6					1		1	1		3
7			1			1	1			3
8					1			1		2
Total	**7**		**12**	**4**	**64**	**7**	**11**	**5**	**1**	**111**

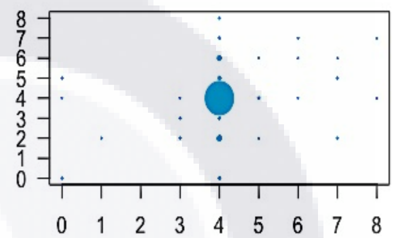


Figura 13. Dispersión de puntajes. Recursos empleados

La dificultad que se detectó con ambas observadoras para valorar esta dimensión se centró en la identificación del uso de recursos como el libro de texto, cuaderno y pizarrón, ya que en estos casos las observadoras seleccionaron la opción *no aplica*. El tipo de anotaciones que hicieron para justificar dicha opción fueron:

> El docente no utiliza ningún recurso, todo es mental
> Aunque retoman un ejercicio del libro… es sólo la revisión de lo hecho anteriormente
> La maestra pide que saquen el libro pero no lo utilizan para esta tarea
> Solicita que saquen la libreta pero sólo anotan algunos ejercicios para resolver

Para abordar estas situaciones se analizó la descripción de los niveles de desempeño, enfatizando los aspectos relacionados con el uso de artefactos o recursos como medios de información. Por ejemplo:

01_SFR_02_T1(Nivel medio)
“Medidas de longitud”

D: el día de ayer me fui con una duda… en el cuadrito que llenamos (señala cuadro en el pizarrón), saquen su cuadrito en su libro…
A: Ese…
D: ¿Cuál es la pregunta que nos plantearon sobre la altura del bote? Está en la siguiente página (los estudiantes revisan su libro de texto).
AA: Si es mayor o menor que un metro…
D: Y le respondimos que es menor… pero de acuerdo al cuadro (señala el pizarrón), ¿qué altura tiene el bote en metros?
A: ¡Ah! ¡No!.. tiene 4.5 metros..
AA: ¿En metros?... ¿metros?
D: ¿Cuánto le pusimos ahí? (señala uno de las partes del cuadro dibujado en el pizarrón)...
A: 0.0435
D: ¿serían metros verdad? Convertidos a decímetros cuántos sería?
AA: 0.435
D: y convertidos a centímetros ¿cuánto sería? … 4.35 (anota el docente en el pizarrón).. Ok. Esta es mi duda… ¿quién les dijo que pusieran un punto aquí (señala la cifra en el cuadro del pizarrón)?
AA: Nadie…
D: A ver… porque la altura del bote sería de 43 milímetros o sea 4 centímetros… O sea, el bote sí es menor que un metro, pero imagínense de 4 centímetros en una oficina… ¿hay botes de 4 centímetros?
AA: No…
D: ¿Podremos corregir esto?... Sí, le quitamos el punto y partimos de aquí …

La tarea matemática se desarrolló a partir de un problema propuesto en el libro de texto. Sin embargo, el docente utilizó el pizarrón para hacer anotaciones en un cuadro de referencia para los estudiantes, que utilizaron para discutir en plenaria. No obstante, la función que se le dio al pizarrón fue mostrar y concentrar información. Si se hubiera utilizado otro artefacto para construir el bote de basura con las medidas que se discuten en plenaria tal vez habría permitido a los estudiantes llegar a una noción clara y confirmar las medidas que fueron objeto de la discusión descrita, esto habría hecho que el artefacto fuera de nivel alto.

Se revisaron otras situaciones en las que se utilizaron artefactos diversos: el pizarrón, cuaderno y libro de texto. En la revisión se buscó ayudar a las observadoras a identificar cuando los artefactos se emplean como elementos informativos o de reforzamiento para el propósito matemático o cuando permiten otras funciones (analizar, discutir, evaluar o comparar ideas). Después de analizar diferentes ejemplos con las videograbaciones una de las observadoras mencionó:

Creo que ya quedó claro… aunque sea el libro sí puede cambiar de nivel de uso siempre y cuando cumpla con ciertas características…

Esta aclaración permitió que ambas observadoras identificaran el uso informativo de los artefactos. Esta dimensión sólo se abordó en una de las cuatro sesiones de calibración, ya que la coincidencia de calificaciones entre las dos observadoras y la observadora experta fue más constante que en otras de las dimensiones valoradas. Algunas de las anotaciones de las observadoras después de la calibración se señalan a continuación:

> Utilizan el libro de geografía, el cuaderno, y un diccionario… pero sólo se emplean como medios de información
> Libro de texto para analizar información y realizar ejercicios matemáticos
> Libreta para registrar el ejercicio que van hacer de tarea
> Pizarrón como medio informativo y de reforzamiento

Alternativas de solución.

Esta dimensión buscó identificar la flexibilidad del docente para proveer y aceptar una variedad de estrategias o procedimientos para resolver un problema matemático. El foco de atención fueron las oportunidades que el docente dio a los estudiantes para emplear múltiples posibilidades de solución sin sesgar el proceso de pensamiento de los estudiantes. A pesar de que esta rúbrica incluyó indicadores y ejemplos de los diferentes niveles de desempeño, obtuvo los resultados más bajos de acuerdo entre las observadoras, tal como se muestra en la tabla 16 y en la figura 14 .

Los resultados también reportan desacuerdos importantes, tales como: 0,7; 1,8; 3,6; 4,8. Por lo que en diferentes sesiones de calibración fue necesario retomar casos en que las observadoras calificaron como niveles bajos algunas tareas cuyas características coincidían con los criterios de niveles altos y viceversa. Para abordar estas discrepancias se seleccionaron fragmentos de videograbaciones previamente calificadas para mostrar y discutir con las observadoras situaciones particulares en relación con los criterios incluidos en la rúbrica.

Tabla 16:
Alternativas de solución

	0	1	2	3	4	5	6	7	8	Total
0			3					1		4
1	1	4	5		5		4		1	20
2		1	7	1	1					10
3		1	2		3		3			9
4		1	2	2	2	3	3		2	15
5	1	1	3	2	5	1	2		1	16
6			8	1	3	6	5	2		25
7			1		2		1		1	5
8					2		2	2	1	7
Total	**2**	**8**	**31**	**6**	**23**	**10**	**20**	**5**	**6**	**111**

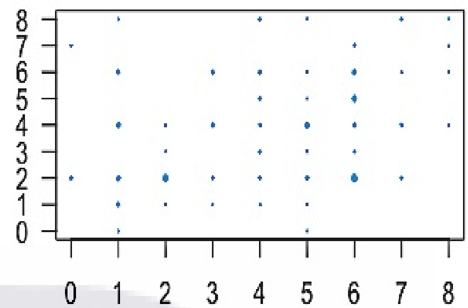


Figura 14. Dispersión de puntajes. Alternativas de solución

Uno de los casos revisados fue cuando las observadoras asignaron una calificación del nivel alto a una tarea en la que la docente planteaba preguntas en relación con conceptos matemáticos y las opciones de respuesta eran múltiples. Sin embargo, el tipo de tarea en realidad no estaba promoviendo la implementación de estrategias o procedimientos para resolver un problema matemático, sino que se centraba en la selección de una de las opciones de respuesta dadas por la docente en un lapso de treinta segundos para corroborar información aprendida previamente, tal como se muestra en la siguiente trascripción:

08_AGS_02_T3
“El círculo y la circunferencia”

> D: Entonces vamos.. yo les voy a ir leyendo las preguntas y les voy a dar las opciones de respuesta, ¿sí? Pregunta número uno, les voy a poner la opción hasta que termine de leer (Los estudiantes tiene un control electrónico que envía la respuesta seleccionada a la computadora de la maestra). Dice… ¿Es la línea que pasa por en medio de un círculo y lo divide en dos partes iguales? A) diámetro, B) fracción, C) radio
> A: Maestra, ¿cuál dijo que era el diámetro?
> (La docente repite las opciones de respuesta… Espera que todos los controles hayan enviado sus respuestas. Después de treinta segundos señala el porcentaje de respuestas correctas y la opción correcta, la “A”)
> AA: ¡Uh! (La mayoría de los estudiantes festeja).

Una de las observadoras comentó que al ver la videograbación por primera vez consideró que la opción de respuesta múltiple daba oportunidad a los estudiantes de elegir. Sin embargo, después de comentar esta situación en una de las sesiones de calibración y de revisar nuevamente lo que se busca valorar con esta rúbrica: “Se refiere a la flexibilidad del docente para promover y aceptar una variedad de estrategias o procedimientos para resolver un problema matemático” la observadora señaló:

> Bueno yo lo veía como esa que tuvieran varias opciones, pero yo lo veía en el sentido de opciones de respuesta… y si somos claros aquí dice que es una variedad de estrategias o procedimientos… entonces ahí está la clave…

En otra tarea matemática en la que una docente promueve la implementación de diferentes estrategias y procedimientos de solución, las observadoras asignaron calificaciones en niveles bajo y medio. En la discusión de este caso, una de las observadoras mencionó que calificó a la docente en ese nivel ya que:

> La maestra permite que los estudiantes expongan los diferentes procedimientos que emplean para solucionar un mismo problema… Pero al final verifica el resultado mostrando un procedimiento a los alumnos…

A través de la revisión de este caso, se demostró que durante el desarrollo de la tarea la docente tiene la apertura para que los estudiantes expongan las diferentes estrategias y procedimientos empleados, y da tiempo suficiente para que compartan entre compañeros, corrijan y comprueben que en los diferentes casos obtuvieron el mismo resultado.

13_AGS_01_T2
“Problemas de reparto”

> D: A ver… vamos a ver, se me hace que no se imaginaron los pasteles.. A ver quién nos quiere compartir..
> AA: yo, yo, yo…(los alumnos voluntariamente levantan la mano para participar)
> D: A ver… Pasa (nombre de alumno) al pizarrón... Primero, ¿qué encontraste en tu equipo?, ¿hubo diferencias?
> A: No.
> D: En la primera no, ok. ¿En la primera que trabajaron encontraron diferencias en alguna? (pregunta al grupo)
> AA: Sí, sí… (algunos alumnos responden afirmativamente y otros negativamente)
> D: Entonces, ahorita vamos a ver cómo resolvió su compañero y el problema y luego vemos quién lo hizo diferente a él… ¿De acuerdo?
> AA: Sí… sí.. (festeja la mayoría de los estudiantes)
> D: Bien, a ver quién lo hizo como (nombre del alumno)… Bueno, quién quiere compartir que los resolvió de manera diferente… Pasa (indica a otro alumno).. A ver, ¿cómo lo resolviste?
> (Pasa otro estudiante al pizarrón)
> D: A ver, cuando haces una división, ¿qué debes poner dentro de la galera?
> AA: Lo que tengo…
> D: Lo que vas a repartir… y, ¿qué vas a repartir?... ¿Son cuatro pasteles?
> AA: No… (responden a coro)
> D: ¿Cuál fue el primer error?... (permite que el estudiante responda)

(los estudiantes empiezan a comentar y a opinar)
D: a ver déjenlo que piense y después le ayudamos…
(El alumno que está trabajando en el pizarrón modifica la posición de los números en la división)
D: ¡Ah! Muy bien, ¡ya viste el primer error! … quién lo resolvió como el compañero… e ¿hicieron lo mismo?
AA: No…
D: A ver, su compañera dice que ella lo resolvió de otra manera y lo va a compartir…
(Pasa una alumna al pizarrón y realiza el procedimiento empleado)
D: A ver, explíqueles a sus compañeros… A ver, fíjense su compañero lo hizo de una forma y su compañera de otra y los dos lo tiene bien… Los circulitos que hizo su compañera son los niños, entonces nos está explicando que primero los dos pasteles los dividió en medios y a cada niño le dio un medio, y le sobraba un pastel… y el último pastel lo dividió en cuatro partes y a cada niño le dio un medio más un cuarto y ella los sumo… Un medio más un cuarto, ¿cuánto te da?
AA: Tres cuartos…
D: A ver, fíjense ya llevamos diferentes maneras de resolverlo, ¿alguien que lo haya tenido diferente?.. bueno vamos a ver, hubo alguna otra forma de resolverlo, ya sea que lo tuvieran bien o mal, ¿no? Vamos a ver, esta fue la situación… Francisco tiene tres pasteles redondos… Hay problemas que se pueden resolver de hasta tres o más maneras..

Durante las cuatro sesiones de calibración se analizaron y discutieron diferentes ejemplos para identificar los diferentes niveles de desempeño propuestos en la rúbrica. Sin embargo, los resultados muestran que al igual que la dimensión de exigencia cognitiva y el trabajo colaborativo; identificar las alternativas de solución en las tareas fue complejo para las observadoras. En este caso, se observó dificultad para identificar cuándo los docentes generan espacios u oportunidades para que los estudiantes compartan y comparen sus propios procedimientos.

Preguntas para la reflexión.

Para esta dimensión se solicitó a las observadoras identificar el tipo de preguntas que el docente o los estudiantes plantean para promover la reflexión durante las tareas matemática. Los acuerdos identificados en este caso se distribuyen entre los cuatro niveles de desempeño de la rúbrica. Por lo tanto, no se observa una concentración hacia algún nivel en específico. Los desacuerdos entre las observadoras muestran diferencias importantes, con puntajes como: 1,6; 2,5; 3,6, 6,2, aunque en menor medida que otras dimensiones (Tabla 17, Figura 15).

Tabla 17:
Preguntas para la reflexión

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	Total
0										
1	1	2	3	1	3	1	1			12
2			1	2	6	1				10
3			6	1	9	1	8			25
4		1	4		8	8	9			30
5			2		5	2	8	2		19
6			1		1	1	6			9
7			1				1	1		3
8							1	1	1	3
Total	**1**	**3**	**18**	**4**	**32**	**14**	**34**	**4**	**1**	**111**

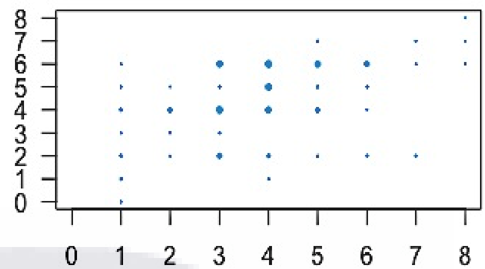


Figura 15. Dispersión de puntajes. Preguntas para la reflexión

En esta dimensión se detectaron problemas con preguntas que permitían a los estudiantes justificar o demostrar ideas o procedimientos matemáticos, así como para identificar las acciones del docente para animar a los estudiantes a comentar y comprobar respuestas.

Uno de los ejemplos que se retomó para explicar la descripción del nivel alto de la rúbrica se refiere al tipo de preguntas que se plantean durante una tarea en la que los estudiantes deben emplear la escala gráfica u otros artefactos para obtener la distancia real aproximada de un trayecto entre diferentes cerros de México. Considerando que los estudiantes tenían diferentes opciones para medir la distancia: estambre, escala gráfica u otros. La docente preguntó en cada equipo: “¿Por qué seleccionaron ese trayecto?” y “¿Cómo llevarían a cabo el procedimiento para obtener la distancia real aproximada?”

La trascripción del ejemplo que se presenta a continuación muestra que el tipo de preguntas que planteó la docente permitió a los estudiantes justificar su trabajo y sus respuestas:

26_AGS_02_T2
“La escala gráfica: Determinar la distancia real aproximada de un trayecto”

> D: ¿De cuál cerro salieron?, y ¿hacia dónde se dirigen? (pregunta la docente a un equipo de cuatro estudiantes)
> A: Al cerro pelón…
> D: ¿Por qué decidieron irse a ése y no al cerro de San Andrés?

AA: Porque se nos hace que el trayecto va ser un poco más derecho…
D: ¿Y queremos que el trayecto sea más derecho? ¿O queremos recorrer la mayor distancia?
A: La mayor distancia…
D: Entonces… ¿qué están utilizando para comparar cuál es la mayor distancia?, ¿qué necesitan comparar?... yo también me podría ir del cerro la Guadalupana y decir, es que es más derecho cerro la Catedral.. yo te puedo decir, es que es más derecho… ¿voy a tomar esa decisión?, ¿en qué tengo que fijarme?
A: Si es mayor o menor la cantidad de kilómetros.
D: ¿Y cómo voy hacer para saber si es mayor cantidad el cerro la Guadalupana o el cerro pelón?
AA: Midiendo…
D: Entonces… ¿cómo van a medir?

(Luego pregunta a otro equipo de estudiantes)

D: A ver ¿por cuál están partiendo?
AA: Dos estudiantes muestran el trayecto en el mapa…
D: ¿Por qué al cerro San Andrés y no al cerro la Catedral?, por ejemplo
A: Porque queríamos hacer el mayor número de kilómetros…
D: ¿Y cómo saben que ahí hay mayor número de kilómetros?, ¿sólo viéndolo?, ¿qué está utilizando?

D: ¿Hacia dónde se fueron?
A: Empezamos desde aquí y nos vamos a ir así (muestra en el mapa)
D: ¿Por qué?
AA: Del cerro la Guadalupana al cerro la Catedral.
D: ¿Qué consideraron para decidir ese trayecto?
A: Porque como dice que tenemos que hacer más recorrido, pensamos que de aquí, luego nos pasamos al cerro la catedral, luego al cerro Vicente Barrancas, y de ahí nos vamos al cerro…
AA: El guacal… y de ahí nos vamos al cerro Prieto
D: ¿y consideran que de ahí ya es la mayor distancia?, ¿Por qué?
A: Sí…
D: ¿Cómo comprobaron que era la mayor distancia?
A: Primero pensamos que podía ser así… pero luego pensamos que era la mayor distancia…
D: ¿Nada más con pensar?
AA: Primero hicimos razonar a (nombre del compañero)… porque él quería ir así (muestra en el mapa) hasta llegar al cerro del prieto…
D: ¿Pero, qué es lo que estamos viendo que les permite a ustedes valorar las distancias?
AA: Los kilómetros.
D: Y, ¿en qué momento lo están considerando? Ya pensaron una forma, pero, ¿ya lo comprobaron?, ¿qué están utilizando para comprobar las distancias?, ¿por qué?

(La docente pasó con los diferentes equipos haciendo preguntas sobre los trayectos, cómo los seleccionaron y por qué. Cuestionó sobre los artefactos, como estambres o cintas métricas, que los estudiantes estaban utilizando).

Es preciso mencionar que esta dimensión sólo se abordó en una de las cuatro sesiones de calibración, ya que las diferencias entre las observadoras fueron de uno o dos puntos, por lo que se optó por sugerir que revisaran los ejemplos que se incluyeron en el apartado de ejemplos por dimensión y nivel de desempeño del POAEM.

Conexión/ aplicación de los conocimientos matemáticos.

En Conexión/aplicación de los conocimientos, el total de acuerdos fue bajo, con una concentración notoria de ambas observadoras hacia la opción *no aplica*, ya que la primera condición para determinar un nivel de desempeño fue identificar si la tarea proponía una situación matemática, y no ejercicios de mecanización sin contexto. Es decir, el objetivo de la calibración fue identificar si las situaciones problemáticas estaban vinculadas con la experiencia de los estudiantes o en ámbitos más allá del contexto escolar.

Los resultados muestran coincidencias en diferentes niveles de desempeño (tabla 18, figura 16). Sin embargo, también se observan diferencias importantes entre los puntajes 0,2; 0,3 y 0,4, es decir, en algunos casos, una de las observadoras no identificó situaciones de contexto o aplicación del conocimiento en las tareas matemáticas mientras la otra marcó calificaciones del nivel bajo y medio.

Tabla 18:
Conexión de los conocimientos

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	Total
0	20		1		6					27
1	1	1								2
2	7	2	2	13	10		4			38
3	5				1					6
4	5				11					16
5					1		1			2
6					1		2			3
7		1			2	1	9	1		14
8							2	1		3
Total	**38**	**4**	**3**	**13**	**32**	**1**	**18**	**2**		**111**

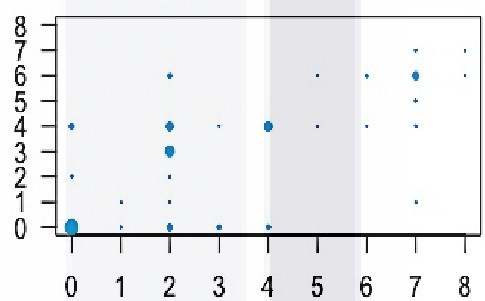


Figura 16. Dispersión de puntajes. Conexión de los conocimientos

A partir de algunos comentarios de las observadoras se revisaron ejemplos en las videograbaciones para identificar situaciones problemáticas contextualizadas. Para ello se tomaron como referencia los criterios de la opción *no aplica*, en los que se especifica que una tarea matemática que plantea ejercicios algorítmicos o de mecanización no se considerará como situación o problema contextualizado. A continuación se incluyen algunos de los comentarios de las observadoras, para mostrar los problemas más recurrentes que enfrentaron con esta rúbrica:

> No logro visualizar que los docentes desarrollen un contexto… generalmente utilizan el libro para poner las actividades". "En ningún momento hay un contexto real de la situación, por eso mis calificaciones son de no aplica o *no se puede valorar*.
>
> Creo que no tengo tanta dificultad como para identificar lo que está pasando con la conexión, pero si luego para brincar de la baja a la media baja… de repente digo no aplica, entre no aplica definitivamente porque están viendo seriaciones, pero luego hay otra que por ejemplo… si repartimos cuatro manzanas entre cinco niños ¿cuánto me da? … bueno no es un contexto educativo, escolar porque no habla de libretas, o de maestros… pero entonces cómo puede cambiar porque repartimos cuatro manzanas entre cuatro niños o cuatro alumnos… ¿entonces ya es contexto escolar porque habla de alumnos o sigue siendo un contexto fuera de su realidad…?
>
> Hay muchas cosas que sirven así como sólo para problematizar… la tarea establece una problematización relacionada con la disciplina, pero por ejemplo eso de las manzanas, son los ejemplos típicos que se dan en clase… ¿cómo determinar si no está siendo una conexión clara con su realidad o si tratan, pero se queda en lo matemático? … como una simulación de la realidad pero falsa…

El siguiente fragmento de una videograbación muestra que la tarea matemática no plantea un problema, sino que se repasan las características de diferentes cuerpos geométricos.

07_AGS_01_T2 (No aplica/ no hay contextualización)

Video corto "Prismas y pirámides"

> D: vamos a poner atención, no anoten… (la docente proyecta un video)… Muy bien, ¿qué observamos?, ¿qué nos faltó comentar? (como tarea previa revisaron conceptos de prismas y pirámides)… ¿Qué escuchaste? (pregunta a un estudiante)
> A: Que tienen dos bases…
> A: Que sus caras son rectangulares…
> A: Qué por medio de la base sé el número de caras que tiene…
> D: Ahora vamos a ver qué dice de las pirámides… (pregunta sobre características de las pirámides)
> A: Sus caras son triangulares…
> A: Terminan en punta… se unen todas sus caras en una cúspide…
> D: Bien, ahora vamos a observarlas (les muestra cuerpos geométricos de plástico), y ustedes me van a decir si es pirámide o si es prisma… Levantan su mano, o bueno me lo dicen todos… (muestra un cuerpo)
> AA: ¡Prisma! (responden los estudiantes en coro)… Continúan identificando los cuerpos geométricos durante la tarea matemática…

El fragmento de la tarea antes descrita se utilizó para trabajar con ambas observadoras respectivamente. Después de revisar el fragmento, una observadora hizo un comentario en el que ella misma clarificó la idea que se buscó trasmitir para esta dimensión a partir de la sesión de calibración:

> Podría ser… si les da una hojita en la que diga obtén el total de caras… ¿esa sería una situación problemática?... No, verdad más bien si lo aplica como, la pirámide de Egipto tiene tantas caras.. ahí sí estaría planteado un problema… ¿verdad?

Después de esta aclaración, se revisó la sección de ejemplos por dimensión y niveles de desempeño del POAEM y se discutieron los criterios que conforman la rúbrica para llegar a acuerdos sobre los niveles de desempeño y los aspectos que permiten diferenciar entre un nivel y otro. Por ejemplo, tareas que plantean un problema en el que el contexto se centra en aspectos de la vida cotidiana, situaciones prácticas o aplicables, se ubican en los niveles altos de conexión; mientras que las situaciones que se centran en aspectos de la vida escolar o disciplinar se consideran en nivel medio. El nivel bajo se identificó con situaciones que simulan una conexión o utilizan objetos para tratar de contextualizar algún algoritmo previamente aprendido; algunos de los objetos que suelen utilizarse con este fin son manzanas, gelatinas, galletas, moños, entre otros.

Trabajo colaborativo.

La indicación para emplear esta rúbrica y valorar el trabajo colaborativo en una tarea matemática fue excluir aquellas tareas en que los docentes solicitaron realizar un trabajo individual. A partir de esta selección se pidió a las observadoras prestar especial atención a la interacción, negociación y cooperación que existía entre los estudiantes, así como entre docente y estudiantes durante el desarrollo de una tarea matemática, esto incluyó actividades en plenaria.

La tabla 19 y la figura 17 presentan los puntajes de ambas observadoras. Los resultados muestran que la mayoría de los acuerdos se concentraron en la opción *no aplica*, mientras que los desacuerdos más notorios se identifican en los puntajes: 0,7; 0,6 y 1,6. Es decir, hubo situaciones en que una de las observadoras calificó como tarea individual, mientras que la otra observadora calificó con valores altos de colaboración. Estas situaciones se abordaron durante las diferentes sesiones de calibración.

Tabla 19:
Trabajo colaborativo

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	Total
0	27	3	2		8		17			57
1	1	2	3		1		1			8
2	1	1	1	2			5			10
3							3			3
4			1	1	1	2				5
5	3		1	1	2		8			15
6	2		1				4			7
7	1				1		1			3
8							1	1	1	3
Total	**35**	**6**	**9**	**4**	**13**	**2**	**40**	**1**	**1**	**111**

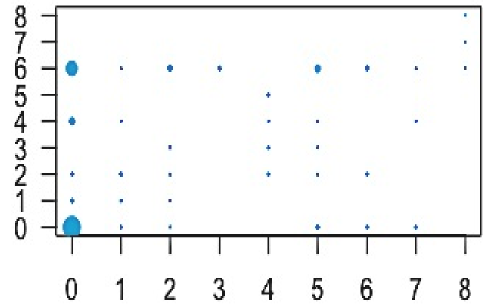


Figura 17. Dispersión de puntajes. Trabajo colaborativo

Las dificultades para determinar los niveles de desempeño con esta rúbrica surgieron con tareas matemáticas en las que el trabajo colaborativo se realizó en plenaria. Esto de acuerdo con comentarios de una de las observadoras:

> Se me dificulta la del trabajo colaborativo… de repente cuando hacen la plenaria, pero que no indica… no dice vamos a trabajar colaborativamente, no lo hace explicito, pero en grupo tratan de armar una definición o algo… y así otros videos donde igual… pero finalmente no sé distinguir hasta donde están desarrollando un conocimiento o no… Por ejemplo en un video retomaban conocimientos previos y entre todos arman una nueva definición, ahí sí lo estoy tomando como un trabajo colaborativo.. Pero me ha costado identificar si es o no es… Para diferenciar me estoy fijando en si arman o desarrollan algo en conjunto, entonces ahí sí lo tomo como trabajo colaborativo

Durante esta sesión de calibración se llegó al acuerdo que las tareas matemáticas que se realizan en plenaria y en las que se observa que los estudiantes aportaban ideas y discutían con el docente para construir una idea, concepto o producto, se valorarían como trabajo colaborativo de nivel medio-alto, ya que son elementos que coinciden con la descripción de ese nivel en la rúbrica: interactuar constantemente con los estudiantes para sugerir que compartan sus ideas o procedimientos con los que construyen o fortalecen algún conocimiento matemático.

Otra situación discutida en esta dimensión fueron tareas en donde los estudiantes se reunieron en equipos, sin embargo, durante el desarrollo de la tarea no se identificaron elementos del trabajo colaborativo, ya que al interior de los equipos los estudiantes trabajaban individualmente. En este tipo de situaciones una de las observadoras optó por seleccionar la opción *no se puede* valorar argumentando:

> Yo le había puesto ahí que la maestra hace énfasis en que trabajen en equipo, que se ayuden… Pero yo no podía valorar si realmente llegaban a la conclusión o desarrollo del resultado en equipo…

Lo anterior llevó a revisar ejemplos para relacionarlos con los criterios y descripción del nivel de la rúbrica. En casos como el que se describirá más adelante, las observadoras reconocieron que se trataba de situaciones en las que la interacción de la docente con los estudiantes se limita a verificar que estén desarrollando el trabajo solicitado. Es decir, no hay evidencias de que los estudiantes han compartido ideas o procedimientos para construir o fortalecer algún conocimiento matemático.

08_AGS_01_T2
"Multiplicación de números decimales por naturales"

> D: Van a contestar en equipo la página 167 y 168… si hay alguna duda levantan su mano y me acerco a su equipo… empezamos a contestar esas dos páginas. Veinte minutos…
> (Los estudiantes se reúnen en equipos de cinco integrantes y la docente resuelve dudas individuales sobre los procedimientos empleados en los diferentes equipos, da rondas para comprobar que estén trabajando)
> A: Un estudiante levanta la mano y menciona a la docente que ya ha terminado…
> D: La docente responde al alumno: "Estás en un equipo, faltan los demás… entonces todavía no acaban" … ¿Listos?, vayan sacando su pluma porque vamos a poner palomita si la tuvimos bien y a corregir si la tuvimos mal… ¿ya tiene cada equipo la persona que va a decir la respuesta?... van a leer el problema, la situación y su respuesta y cómo la obtuvieron… ya se acabó el tiempo…

A pesar de las sesiones de calibración que se realizaron durante el proceso de calificación para esta dimensión, una de las observadoras comentó:

> Esta es una de las dimensiones más difíciles al momento de verla, porque… a veces es muy difícil separar cuando sí están construyendo y de cuando sólo están compartiendo información… A lo mejor ahorita con lo que sólo reportan resultados, porque aunque digan siéntense en parejas, siéntense en equipos hay muchas situaciones en las que están trabajando individual aunque ella diga es equipo, háganlo así… y siguen… así que cuesta encontrar dónde van… incluso los propios docentes revisan en individual

Acuerdos brutos y relativos por dimensión.

Para complementar la información antes descrita, la tabla 20 presenta los puntajes de acuerdo bruto y relativo entre ambas observadoras. Este último considerando ± 1 punto de diferencia para las calificaciones otorgadas en cada dimensión, ya que la escala de las rúbricas del POAEM propone niveles graduales con dos puntajes por nivel (1,2; 3,4; 5,6; 7,8). Este rango implica un margen de error del 12.5%.

Los resultados obtenidos dejan ver que el porcentaje bruto más alto se encuentra en la dimensión recursos empleados con 56%, seguido por la dimensión claridad con 38%, y conexión y trabajo colaborativo con 33% respectivamente. El resto de las observaciones se mantiene por debajo del 30%. Sin embargo, esta situación cambia con respecto al porcentaje relativo (±1 punto), ya que los resultados aumentan considerablemente en cada caso, especialmente en las dimensiones alternativas de solución, en el que se pasa de 18% a 51%; preguntas para la reflexión, en el que se triplica el porcentaje; y exigencia cognitiva, en el que aumenta al doble del porcentaje bruto.

Tabla 20:
Acuerdo entre observadoras por dimensión

Dimensión/Acuerdo	**Frecuencias de acuerdo bruto**	**% de acuerdo bruto**	**Frecuencias de acuerdo (+-1)**	**% de acuerdo (+-1)**
Claridad de la tarea	42	**38**	83	**75**
Exigencia cognitiva	22	**20**	51	**46**
Recursos empleados	62	**56**	81	**73**
Alternativas de solución	20	**18**	57	**51**
Preguntas para la reflexión	22	**20**	67	**60**
Conexión/aplicación	37	**33**	66	**59**
Trabajo colaborativo	36	**33**	61	**55**

En conjunto, los diagramas de dispersión, porcentajes brutos y relativos, así como los comentarios registrados revelan que las observadoras tuvieron mayor facilidad para calificar las rúbricas de claridad de la tarea matemática y los recursos empleados para ésta. En el caso de la claridad de la tarea matemática ambas observadoras coincidieron en que los criterios para identificar el producto solicitado o la meta a alcanzar facilitó la calificación, las dificultades surgieron cuando las tareas matemáticas eran continuación de otra, ya que las metas o productos no eran explicitas.

En cuanto a los recursos empleados durante la tarea matemática, las observadoras señalaron que identificar artefactos distintos al libro de texto resultó sencillo, sin embargo la confusión se presentó con el uso del libro de texto, ya que no tenían claro si éste se podría considerar otro artefacto para la realización de la tarea matemática. Esta situación se clarificó con los ejercicios realizados durante las sesiones de calibración.

Respecto a la dimensión de conexión o aplicación del conocimiento, las observadoras mencionaron que discriminar las tareas que no incluían un contexto fue un procedimiento sencillo. Sin embargo, distinguir entre los niveles medios y altos de esta rúbrica resultó complejo durante el proceso de calificación. La misma situación se observó con la rúbrica de trabajo colaborativo, ya que identificar las tareas que exigían un trabajo individual fue fácil, no obstante cuando se trató de trabajo en plenaria las observadoras tuvieron problemas para distinguir entre los niveles medios y altos de interacción.

En la rúbrica sobre preguntas para la reflexión se registraron acuerdos entre ambas observadoras para diferentes niveles de desempeño, en este caso las observadoras señalaron que la dificultad surgió para identificar los niveles altos de desempeño, por lo que en las sesiones se calibración se identificaron diferentes ejemplos sobre el tipo de preguntas que pueden llevar a los estudiantes a la argumentación de sus respuestas.

Las rúbricas para valorar la exigencia cognitiva de las tareas matemáticas y las alternativas de solución implicaron un proceso complejo de identificación de indicadores, ya que esto involucró el nivel de pensamiento requerido en una tarea matemática, así como identificar las oportunidades que los estudiantes tuvieron para aplicar múltiples métodos de solución. Al respecto ambas observadoras coincidieron en que los niveles de las rúbricas resultaron complicados de aplicar, lo cual se confirma con los porcentajes de acuerdo más bajos reportados en la tabla 20.

4.1 Consistencia interna: Alfa de Cronbach

El coeficiente de homogeneidad o consistencia interna Alfa de Cronbach permite calcular la correlación de cada reactivo o ítem con cada uno de los otros considerados en un instrumento (Cohen y Swerdlik, 2001). En la literatura se sugiere emplear este coeficiente con ítems policotómicos (respuestas múltiples), considerando que el valor que más se aproxime a 1 indicará mayor consistencia interna del instrumento. Autores como George y Melley (2003) proponen los siguientes valores para este coeficiente: excelente para valores mayores a 0.90; bueno de 0.80 a 0.90; aceptable de 0.70 a 0.79; cuestionable de 0.60 a 0.69; pobre de 0.59 a 0.50; e *inaceptable* menor a 0.50.

La tabla 21 presenta los valores obtenidos para el coeficiente Alfa de Cronbach con las dos bases de datos utilizadas para realizar los análisis factoriales: tareas matemáticas con N=111 y docentes con N=20. Los resultados obtenidos en cada caso sugieren que la consistencia interna de los ítems es aceptable para la base N=111 y buena para la base N=20 según los criterios de George y Melley (2003). Sin embargo en este último caso debe tenerse en cuenta que la cantidad de casos es limitada.

Tabla 21:
Estadísticos de fiabilidad

Coeficiente	Tarea matemática (111 casos)	Docente (20 casos)
Alfa de Cronbach	.785	.869

Para complementar la información anterior, las siguientes tablas presentan datos generales sobre el aporte que cada uno de los ítems hace al constructo de acuerdo con el coeficiente Alfa de Cronbach si se suprime alguno de los elementos. La tabla 22 presenta el modelo excluyendo la variable claridad, en donde los resultados muestran que todos los ítems hacen una aportación considerable.

Los resultados muestran que la varianza explicada se mantiene en el mismo rango de porcentaje entre los diferentes elementos 55%-57%, sólo en el caso de la variable trabajo hay una disminución mayor si se elimina el elemento (47%) y un aumento si se elimina la variable recursos (64%). No obstante, todas las correlaciones de los ítems con el elemento son mayores a .40 (punto de saturación) y el coeficiente Alfa de Cronbach es mayor con todo el conjunto de variables de este modelo (.785), por lo que resulta una mejor decisión mantener todos los ítems.

Tabla 22:
Aporte para el constructo por variable (sin claridad)

	Media de escala si el elemento se ha suprimido	Varianza de escala si el elemento se ha suprimido	Correlación total de elementos corregida	Alfa de Cronbach si el elemento se ha suprimido
Exigencia	17.58	55.283	.604	.737
Recursos	17.13	64.838	.483	.771
Alternativas	17.09	54.683	.568	.745
Preguntas	17.04	57.853	.634	.736
Conexión	18.53	56.724	**.423**	.785
Trabajo	18.45	47.632	.599	.741

La tabla 23 presenta el modelo incluyendo la variable claridad. Los resultados muestran que la varianza explicada se mantiene en un rango mayor al del modelo anterior con 55%-77% si se elimina algún elemento. En este modelo hay un aumento considerable si se elimina la variable claridad, ya que el coeficiente Alfa de Cronbach que se puede obtener es de 0.78. Esto podría deberse a que la correlación de la variable claridad con el elemento está por debajo del punto de saturación (.40). Por lo tanto, la información obtenida con este análisis es coincidente con los resultados del AFE y el AFC, por lo tanto, la confiabilidad del instrumento aumenta si se excluye la variable claridad.

Tabla 23:
Aporte para el constructo por variable (con claridad)

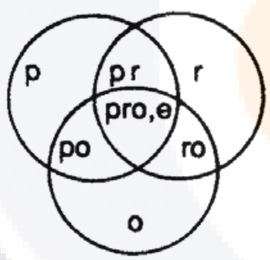
	Media de escala si el elemento se ha suprimido	Varianza de escala si el elemento se ha suprimido	Correlación total de elementos corregida	Alfa de Cronbach si el elemento se ha suprimido
Exigencia	21.90	64.126	.617	.732
Recursos	21.45	74.213	.506	.761
Alternativas	21.41	65.099	.528	.750
Preguntas	21.36	67.342	.629	.735
Conexión	22.86	66.506	**.413**	.777
Trabajo	22.77	55.431	.627	.730
Claridad	21.16	**77.555**	**.312**	.785

4.2 Estudio de Generalizabilidad

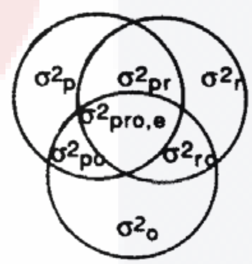
Para ofrecer mayor información sobre los puntajes obtenidos con el POAEM y sobre las diferentes facetas de error en la medición (no sólo del error de los jueces), se realizó un estudio de confiabilidad basado en los componentes de varianza. El estudio de generalizabilidad o estudio G se basa en la confiabilidad del comportamiento de mediciones estimando múltiples fuentes de error de forma independiente.

Los resultados dan información sobre la confiabilidad para diferentes tipos de interpretación, ya que el universo de medición es definido por el conjunto de condiciones posibles en que se podría tomar la medida y al que se pretende generalizar el puntaje observado. Es decir, por la mayor fuente de variación o facetas que se determinan a partir de las posibles fuentes de error en las muestras de una medida, por ejemplo: ítem, juez, ocasión, segmentos, etcétera (Shavelson, 1991; Shavelson y Webb, 1991; Brennan, 2001).

Las facetas o fuentes de error para un estudio G se establecen de acuerdo con la estructura de los datos de una medida. Por ejemplo, para un diseño **P x R x O**, en donde "P" se refiere a persona, "R" a juez y "O" a la ocasión en que se lleva a cabo una medida, se pueden establecer siete fuentes de varianza, así como las interacciones que se pueden establecer entre éstas, tal como se muestra en la siguiente figura:



Fuentes de varianza



Componentes de varianza

Figura 18. Diagrama para dos facetas, anidado p x r x o.

Fuente: Shavelson, R. J., y Webb, N. M. (1991). Generalizability Theory. A Primer. Thousand Oaks, California: SAGE Publications, Inc.

Frecuentemente, los diseños propuestos cumplen con el mismo número de facetas anidadas para cada nivel. Es decir, "una faceta i está anidada con otra j, cuando cada nivel de i aparece en un solo nivel de j, y cuando hay más de un nivel de I para cada nivel de j", por lo que existe el mismo número de intersecciones entre todas las facetas incluidas en un diseño (Martínez, 2015, p. 47). Un ejemplo gráfico se establece a partir del siguiente cubo en el que se incluyen las facetas y se señalan las interacciones en cada caso.

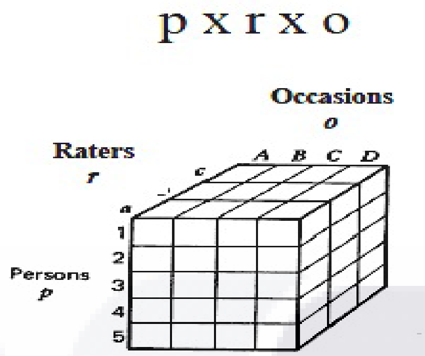


Figura 19. Diseño anidado con dos facetas
Fuente: Shavelson, R. J., y Webb, N. M. (1991). *Generalizability Theory. A Primer*. Thousand Oaks, California: SAGE Publications, Inc.

A diferencia del diseño típico anidado en el que se tiene el mismo número de facetas en un diseño, los datos recabados con el POAEM sugieren un diseño *anidado desbalanceado* con tres facetas, ya que de acuerdo con Shavelson y Webb (2006, p. 610) "un diseño desbalanceado surge cuando el número de niveles de facetas anidadas varía en algún nivel, produciendo un número desigual de niveles". En este caso, para el diseño **P x R x O x S:O,** en donde **"**P" se refiere al docente, "R" al juez, O[8] a la ocasión y S:O al segmento dentro de ocasión, se tiene el mismo número de niveles para las facetas P, R y O; pero no para la faceta "S:O". Sin embargo, los resultados obtenidos a partir del diseño permiten estimar la magnitud del error asociado a las facetas de medición (componentes de varianza) que se describen a continuación:

σ^2_p = varianza real de las medias de las personas (docentes)
σ^2_r = diferencias constantes promedio entre jueces (severidad)
σ^2_o = diferencias constantes promedio entre ocasiones
σ^2_{pr} = inconsistencia de los jueces al evaluar a distintos maestros
σ^2_{po} = inconsistencia del puntaje de los maestros en distintas ocasiones
σ^2_{ro} = inconsistencia del puntaje promedio por juez en distintas ocasiones
$\sigma^2_{s:o}$ = inconsistencia del puntaje promedio por segmento dentro de ocasión
$\sigma^2_{ps:o}$ = inconsistencia del puntaje de los maestros por segmento dentro de ocasión
$\sigma^2_{rs:o}$ = inconsistencia del puntaje promedio por juez en distintos segmentos dentro de ocasión
σ^2_{pro} = inconsistencia de los jueces al evaluar distintos maestros en distintas ocasiones
$\sigma^2_{prs:o,e}$ = triple interacción y varianza residual intracelda (error)

[8] Ocasión se refiere a los días en que se realizaron las observaciones (dos días/dos ocasiones).

La tabla 24 presenta los resultados para cada una de las siete dimensiones que conforman el POAEM, así como para el conjunto de dimensiones incluyendo y excluyendo la dimensión claridad. La información que se analizó en cada caso incluye la proporción de varianza obtenida con el programa SPSS (ANOVA con efectos aleatorios), así como el porcentaje correspondiente a la varianza para cada una de las facetas del diseño **P x R x O x S:O,** utilizando las fórmulas de cuadrados medios esperados E[MS].

Tabla 24:

Varianza por dimensión para el Protocolo de Observación de las Actividades de Enseñanza en Matemáticas

	Claridad de la tarea matemática		**Exigencia cognitiva**		**Recursos empleados**		**Alternativas de solución**		**Preguntas reflexión**		**Conexión/ aplicación**		**Trabajo colaborativo**		**SUMA + CLARIDAD**		**SUMA SIN CLARIDAD**	
	σ^2	**%**	**σ^2**	**%**	**σ^2**	**%**	**σ^2**	**%**	**σ^2**	**%**	**σ^2**	**%**	**σ^2**	**%**	**σ^2**	**%**	**σ^2**	**%**
Docente (P)	.228	10.0%	.420	10.1%	.198	9.4%	.817	16.8%	.495	15.7%	2.798	49.9%	.945	11.4%	23.085	29.5%	19.497	27.7%
Día (O)	.000[9]	0.0%	.000	0.0%	.000	0.0%	.000	0.0%	.000	0.0%	.057	1.0%	.000	0.0%	-.725	-0.9%	-.627	-0.9%
Segmento(Día); (S:O)	.033	1.4%	.000	0.0%	.000	0.0%	.000	0.0%	.202	6.4%	.000	0.0%	.307	3.7%	.000	0.0%	.000	0.0%
Juez (R)	.000	0.0%	.000	0.0%	.000	0.0%	.000	0.0%	.196	6.2%	.000	0.0%	.597	7.2%	-.306	-0.4%	.107	0.2%
Docente * Día (PO)	.000	0.0%	.000	0.0%	.000	0.0%	.000	0.0%	.111	3.5%	.059	1.0%	.000	0.0%	-1.902	-2.4%	-.677	-1.0%
Docente * Segmento (Día); (PS:O)	.830	36.4%	.919	22.1%	.598	28.5%	1.080	22.2%	.572	18.2%	.525	9.4%	1.654	20.0%	17.170	21.9%	15.163	21.5%
Docente * Juez (PR)	.191	8.4%	.904	21.7%	.429	20.4%	.816	16.8%	.583	18.5%	.384	6.8%	1.070	12.9%	24.258	31.0%	22.231	31.6%
Día * Juez (OR)	.017	0.7%	.000	0.0%	.000	0.0%	.000	0.0%	.000	0.0%	.000	0.0%	.074	0.9%	.763	1.0%	.529	0.8%
Juez * Segmento(Día); (RS:O)	.000	0.0%	.000	0.0%	.000	0.0%	.015	0.3%	.000	0.0%	.000	0.0%	.000	0.0%	-1.539	-2.0%	-1.432	-2.0%
Docente * Día * Juez (POR)	.000	0.0%	.014	0.3%	.214	10.2%	.166	3.4%	.000	0.0%	.359	6.4%	.000	0.0%	.347	0.4%	-.427	-0.6%
Docente * Juez * Segmento(Día), Error (PRS:O, e)	.981	**43.1%**	1.908	**45.8%**	.660	**31.4%**	1.961	**40.4%**	.984	**31.3%**	1.424	**25.4%**	3.627	**43.8%**	17.098	**21.9%**	16.002	**22.7%**
Total Varianza	2.279	100.0%	4.165	100.0%	2.100	100.0%	4.855	100.0%	3.143	100.0%	5.606	100.0%	8.276	100.0%	78.249	100.0%	70.366	100.0%

[9] Los resultados del estudio G incluyeron algunos valores negativos, por lo que se optó por adoptar la recomendación de Shavelson y Webb (2005) en la que señalan que "los valores negativos pueden deberse a un error de muestreo o una mala especificación del modelo (modelo desbalanceado). Sin embargo, una solución es sustituir el valor negativo por el cero cuando éstos son pequeños en magnitud relativa" (p 610).

Con base en los resultados del estudio G es posible explicar el porcentaje de varianza en cada una de las siete dimensiones controlando las once facetas de este diseño (**P x R x O x S:O**). Al respecto, la faceta relacionada con el desempeño de los docentes (P) presenta porcentajes de varianza pequeños para las dimensiones claridad, exigencia cognitiva y recursos (aproximados a 10%). Sin embargo, hay un aumento considerable para las dimensiones de preguntas (16%), trabajo colaborativo (12%) y especialmente en conexión de los conocimientos (50%). Esto es, el porcentaje de varianza en el desempeño docente varía en gran medida entre una dimensión y otra, ya que la práctica de enseñanza puede cambiar por múltiples circunstancias entre un docente a otro. Por ejemplo, por el número de estudiantes que atiende, horario de la clase de la clase, estado de ánimo de los estudiantes o del docente, etc.

La faceta que aborda las diferencias constantes promedio entre ocasiones de observación (O) reporta valores mínimos en todas las dimensiones (0% a 1%), esto sugiere que este diseño no aporta información suficiente sobre la varianza por ocasión. Por lo tanto, sería importante considerar un nuevo diseño que contemple un mayor número de ocasiones de observación, de tal manera que sea posible contar con información relevante al respecto.

Las diferencias constantes promedio entre los jueces u observadoras (R) mantiene un porcentaje de varianza pequeño en la mayoría de las dimensiones, ya que los resultados de la dimensión claridad, exigencia, recursos, alternativas, y conexión reportan un porcentaje de 0%, sólo en el caso de preguntas (6%) y trabajo (7%) indican porcentaje altos. Lo anterior sugiere que no hay diferencias importantes entre la forma de calificar de una observadora y otra en la mayoría de las dimensiones, excepto en el caso de preguntas y trabajo colaborativo.

Las diferencias de los puntajes otorgados a los docentes en los diferentes segmentos dentro de cada ocasión (PS:O) reportan porcentajes de varianza altos y distintos entre las dimensiones. El porcentaje más alto se concentra en la dimensión claridad (36%). Esto significa que en esta dimensión la práctica de los docentes varió de un segmento a otro en gran medida. Una situación similar se registró con la dimensión recursos (28%), seguida de alternativas (22%), exigencia (22%) y trabajo (20%). Sólo las dimensiones preguntas (18%) y conexión (9%) reportaron un porcentaje menor en comparación con el resto de las dimensiones.

Las diferencias entre los jueces al calificar a los docentes (PR) es otra faceta que reporta porcentajes de varianza importantes. Por ejemplo, la dimensión exigencia y recursos reportan los porcentajes con poco más del 20% en cada caso, lo que sugiere que especialmente en estas dimensiones los jueces pudieron calificar de manera distinta dependiendo el docente a evaluar. La misma situación aplica para las dimensiones preguntas (18%), alternativas (17%) y trabajo (12%) con un porcentaje similar. Sólo en el caso de conexión (7%), y claridad (8%) se reportan porcentajes más bajos.

El resto de las facetas: puntaje de los docentes en distintas ocasiones (PO), puntaje promedio por juez en distintas ocasiones (OR), puntaje promedio por juez en distintos segmentos dentro de cada ocasión (RS:O) y la diferencia entre los jueces al evaluar a los distintos docentes en las distintas ocasiones (POR) reportaron porcentajes de varianza mínimos (0%-1%), lo que sugiere que no hay diferencias significativas en la calificación otorgadas a los docentes en las distintas ocasiones de observación o en los diferentes segmento evaluados por los distintos jueces.

La varianza residual (PRS:O,e) de cada dimensión permite resumir el porcentaje total de varianza que no es posible explicar con este diseño. Al respecto se observa que hay porcentajes altos en cuatro de las dimensiones. Esta son exigencia (46%), trabajo (44%), claridad (43%) y alternativas (40%). El resto de las dimensiones reportan porcentajes medios, como es el caso de recursos (32%), preguntas (31%) y conexión (25%). Sin embargo, para obtener mayor información sobre el total de varianza para el POAEM se calculó el porcentaje de varianza que en conjunto aportan todas las dimensiones del instrumento, así como el total de varianza excluyendo la variable claridad tomando como base los resultados del AFE.

Los resultados en conjunto para el diseño **P x R x O x S:O** incluyendo todas las dimensiones, muestran que el mayor porcentaje de varianza se concentra en las diferencias entre los jueces al calificar a los distintos docentes (PR) con 31%, seguido de la variación entre el desempeño de los docentes (P) con 30% y la variación de la práctica de los docentes entre segmentos (PS:0) con 22%. En conjunto este diseño reporta una varianza residual (PRS:O,e) de 22%, por lo cual el diseño con todas las dimensiones explica el 78% de la varianza. Es decir, la magnitud del error asociado a las facetas de medición es mucho menor que con las reportadas de forma individual.

La misma situación se observa con el diseño excluyendo la dimensión claridad, los resultados reportan porcentajes poco menores a los obtenidos con todas las dimensiones. En este caso, el porcentaje de varianza más alto se concentra de igual forma en las diferencias entre los jueces al calificar a los distintos docentes (PR) con 32%, seguido de la variación entre el desempeño de los docentes (P) con 28% y la variación de la práctica de los docentes entre segmentos (PS:O) con 22%. En conjunto este diseño reporta una varianza residual (PRS:O,e) de 23%, por lo cual el diseño excluyendo la dimensión claridad explica 77% de la varianza. Es decir 1% menos que el diseño con todas las dimensiones. No obstante, la magnitud del error asociado a las facetas de medición se mantiene baja en relación con los errores reportados de forma individual.

En resumen los resultados del diseño **P x R x O x S:O** para el POAEM explican en mayor medida la varianza en conjunto con todas las dimensiones que de forma independiente. Sin embargo, otro aspecto que se debe considerar dentro de este estudio es el coeficiente G de confiabilidad. La interpretación para este coeficiente se centra en los valores que más se acercan a 1, ya que esto indicará una mayor confiabilidad de los datos. La tabla 25 concentra los valores del coeficiente G para cada una de las dimensiones, en conjunto, así como su respectivo porcentaje de varianza.

Tabla 25:
Coeficiente de generalizabilidad por dimensión para el POAEM

Dimensión	Coeficiente G	% varianza (PRS:O,e)
Claridad de la tarea	0.3484	43.1%
Exigencia cognitiva	0.3126	45.8%
Recursos empleados	**0.2833**	**31.4%**
Alternativas de solución	0.4585	40.4%
Preguntas para la reflexión	0.4466	31.3%
Conexión/aplicación	**0.8185**	**25.4%**
Trabajo colaborativo	0.4027	43.8%
Suma+ Claridad	**0.5661**	**21.9%**
Suma - Claridad	**0.5422**	**22.7%**

Los resultados de la tabla 25 muestran coeficientes de confiabilidad bajos para todas las dimensiones, excepto para conexión/aplicación en la que se reporta un valor de 0.81. Esta situación podría deberse al diseño desbalanceado empleado para realizar el estudio de generalizabilidad, ya que estos modelos incluyen descomposiciones diferentes de las sumas totales de cuadrados sin una base obvia para elegir entre ellas, por lo tanto pueden generar una variedad de maneras en que los cuadrados medios pueden ajustarse para otros efectos en el modelo (Shavelson y Webb, 2010, p.610).

Otro aspecto importante que puede estar relacionado con los resultados obtenidos es el propio diseño de las rúbricas que conforman el instrumento, ya que los coeficientes en conjunto (todas las rúbricas) fueron más altos que los obtenidos en cada una de las rúbricas. No obstante, los coeficientes alfa de Cronbach y coeficiente G con el conjunto de rúbricas (0.56, 0.54) no reportan un valor suficiente para lograr una buena consistencia interna. Por lo tanto, para otro estudio será necesario lograr un diseño equilibrado (con el mismo número de facetas).

Discusión y conclusiones

El propósito de este estudio fue diseñar un protocolo de observación que permitiera evaluar las actividades de enseñanza, para esto a partir de la revisión de literatura se definieron siete dimensiones para medir el constructo, estas dimensiones son: claridad de la tarea matemática, exigencia cognitiva, recursos empleados, alternativas de solución, preguntas para la reflexión, conexión/aplicación de los conocimientos y trabajo colaborativo. Después del diseño se realizaron diferentes análisis para obtener evidencias de validez de la información recabada con el propósito de determinar si la teoría y las evidencias empíricas respaldan las inferencias realizadas a partir de dicha información.

Las evidencias de validez se centraron en juicios de valor realizados por un comité de expertos que participó en diferentes sesiones de trabajo para discutir elementos conceptuales y operacionales del instrumento, así como un Análisis Factorial Exploratorio (AFE) y un Análisis Factorial Confirmatorio (AFC) para establecer la estructura dimensional del instrumento. Las evidencias de confiabilidad tienen que ver con la concordancia inter-jueces (porcentaje de acuerdo), con los coeficientes de confiabilidad Alfa de Cronbach y coeficiente G para determinar la consistencia interna, así como un estudio de generalizabilidad para identificar las fuentes de error en la medición.

Las sesiones que se realizaron con el comité de expertos permitieron realizar ajustes a las diferentes versiones del POAEM, así como asegurar en cierta medida que las dimensiones definidas para medir el constructo actividades de enseñanza estuvieran relacionadas con elementos conceptuales adecuados, de acuerdo con la opinión y valoración de personas con experiencia en el tema.

En cuanto a la estructura dimensional del instrumento, se encontró que ésta es distinta a la que se había establecido a partir de la revisión de literatura, ya que la mayoría de las dimensiones del POAEM se agruparon en un factor y no en cuatro como se anticipaba. Además, la estructura factorial resultante excluyó a la variable claridad después de los análisis factoriales, es decir que esta variable en realidad no parece medir aspectos relacionados con el constructo actividades de enseñanza. Este resultado se confirmó con el coeficiente Alfa de Cronbach, ya que la consistencia interna aumentó cuando se eliminó esa dimensión.

Lo anterior llama la atención puesto que autores como Newmann, López y Bryk (1998), Stone (2008), y el NCTM (2015) enfatizan la importancia de establecer claramente los propósitos o productos de las actividades de enseñanza, ya que esto permite dirigir el trabajo del docente y los estudiantes estableciendo qué es lo que deben saber y poder hacer después de trabajar con ciertos contenidos. De acuerdo con la literatura sobre la enseñanza de las matemáticas, se infiere que el establecimiento de los propósitos o productos es un requisito indispensable para el logro de los aprendizajes.

Además, los resultados de la rúbrica claridad de la tarea sorprenden dado que durante el proceso de calificación fue justamente esta dimensión la que obtuvo mayor coincidencia entre las observadoras. Por lo tanto, parecen necesario realizar investigaciones adicionales para entender la razón por la cual esta dimensión no parece contribuir a la medición de las actividades de enseñanza; en estas investigaciones se tendrá que revisar la forma en que fue operacionalizada la variable y sus diferentes niveles de desempeño dentro de la rúbrica. Asimismo, será necesario complementar la revisión de literatura y documentos orientadores de las prácticas docentes en matemáticas en México, para distinguir si esta dimensión pertenece a otros constructos de las prácticas docentes.

Con respecto a la concordancia inter-jueces se encontró que las observadoras tuvieron dificultad para comprender los criterios de las siguientes dimensiones: exigencia cognitiva, alternativas de solución y trabajo colaborativo, ya que éstas involucran aspectos complejos de medir. En el caso de la exigencia cognitiva la observación y valoración se centra en el nivel de pensamiento requerido en las tareas matemáticas propuestas. Es decir, un docente puede aumentar la demanda cognitiva permitiendo que los estudiantes exploren y entiendan la naturaleza de procedimientos o conceptos matemáticos o disminuirla presentando ejemplos similares, explicando los procedimientos a seguir, etc. (Stein y Smith, 1998; Jonassen, 2011; NCTM, 2015).

Las alternativas de solución, implican valorar cómo los estudiantes experimentan con diferentes procedimientos y métodos para generar una conciencia crítica y de razonamiento matemático. En este caso, lo que se debe identificar son las oportunidades que los docentes dan a los estudiantes para enfrentar cierto grado de incertidumbre ante el tipo de procedimiento o estrategia a seguir. No obstante, esta dimensión se relaciona en cierta manera con el nivel de exigencia cognitiva, ya que se requiere de tareas de mayor exigencia para tener oportunidades de establecer diferentes métodos de solución (Griffin, 2009; Picaroni y Loureiro, 2010).

El trabajo colaborativo es una dimensión que requiere valorar la interacción entre docentes y estudiantes, y entre estudiantes y estudiantes, más allá de la organización del trabajo. Esto también involucra el trabajo que se realiza entre el docente y todo el grupo de estudiantes (plenaria), por lo que se debe tener especial cuidado en identificar cuándo dichas interacciones en realidad aportan a la construcción de los conocimientos, metas o productos conjuntos (NCTM, 2015).

Algunas de las dificultades que se identificaron en las dimensiones antes mencionadas de acuerdo con las observadoras, tienen que ver con la diferenciación entre niveles de desempeño, así como para identificar los procesos cognitivos; especialmente en los niveles bajos y altos de las rúbricas. Por ejemplo, en exigencia cognitiva, los ejercicios de repetición o mecanización (nivel bajo) y los aspectos que implican poner en juego diferentes habilidades intelectuales y demostrar cierto dominio de conceptos matemáticos (nivel alto) fueron marcados por las observadoras como nivel medio o medio bajo. Estas dificultades se podrían abordar en el proceso de capacitación asegurando en la medida de los posible que se logre la comprensión de los niveles de desempeño que integran esta rúbrica (taxonomía de tareas matemáticas).

En cuanto a las alternativas de solución, se identificó que las observadoras calificaron como nivel de desempeño bajo cuando al final de una tarea algún docente confirmó el resultado de un problema matemático, inclusive cuando previamente el docente ofreció a los estudiantes oportunidades para explorar diferentes formas de solucionar un problema (nivel alto). Es decir, a pesar de contar con evidencias que coincidían con el nivel alto de la rúbrica, las observadoras se guiaron sólo por un rasgo al final de la tarea para determinar la calificación, suponiendo que confirmar un resultado es sinónimo de limitar las alternativas de solución, alejándose así del propósito de esta dimensión: identificar las oportunidades que el docente ofrece a los estudiantes para emplear múltiples posibilidades de solución.

Respecto al trabajo colaborativo, la mayor dificultad se identificó en la valoración de situaciones donde todo el grupo junto con el docente desarrollaban una idea, procedimiento o producto. En estos casos, la observación requería diferenciar cuando las aportaciones de los estudiantes servían para lograr alguna meta o producto en conjunto, de aquéllas en las que el trabajo era conducido por el docente.

Los resultados en las dimensiones señaladas sugieren la necesidad de revisar el planteamiento de los criterios incluidos en las rúbricas, así como aclarar la relación que puede existir entre éstas. En la capacitación parece conveniente incorporar sesiones en las que los observadores u observadoras revisen y analicen los aspectos teóricos que fundamentan cada una de las dimensiones e incorporar casos desafiantes para asegurar la comprensión de los criterios.

En cuanto al estudio de generalizabilidad, se encontró que la mayoría de las dimensiones reportaron porcentajes altos de error. Es decir, el modelo propuesto para estimar la varianza de las diferentes facetas explicó poco sobre las posibles fuentes de error en la medición para cada una de las dimensiones (55%-60%). Sin embargo, esta situación cambió cuando se estimó la varianza del conjunto de las dimensiones que conforman el POAEM, ya que de esta manera el modelo explicó cerca del 80% de la varianza.

Es posible que el comportamiento de los datos del estudio de generalizabilidad esté relacionado con: a) el modelo desbalanceado a partir del cual se realizó el estudio, b) las ocasiones de observación consideradas para el estudio, y c) el diseño del instrumento. En el primer caso, el modelo desbalanceado puede tener repercusiones importantes para el análisis, ya que se trató de un modelo en el que no fue posible contar con el mismo número de niveles por faceta, ya que la cantidad de segmentos o tareas dentro de cada ocasión varió entre 2 y 4. Por lo tanto, convendría que en futuros estudios se probara el protocolo de observación con un modelo balanceado de segmentos por ocasión y se contrastara el comportamiento de los resultados con respecto a los obtenidos en el modelo desbalanceado.

En las ocasiones de observación, los resultados sugieren que con un mayor número de días de observación la confiabilidad de los datos aumentaría en cada dimensión, ya que con el diseño actual el porcentaje de varianza en esta faceta fue mínima (0%-1%). De acuerdo con esta versión del instrumento, se podría sugerir que 4 días de observación serían suficientes para aumentar la confiabilidad de la información obtenida, sin embargo es necesario realizar un estudio de decisión que aporte información precisa al respecto. Los resultados del estudio de generalizabilidad también pueden estar relacionados con el diseño del instrumento, en particular con lo que se ha mencionado en párrafos anteriores sobre la dificultad de la calificación de tres dimensiones: exigencia cognitiva, alternativas de solución y trabajo colaborativo. Por lo tanto, después de que se realicen modificaciones al instrumento será necesario obtener nuevas evidencias de validez de la información utilizando las videograciones que se tienen recabadas.

Los resultados obtenidos en general eran esperables, en el sentido que la medición del objeto de estudio es un proceso complejo. Algunos instrumentos internacionales como el CLASS, que estudia la complejidad de las interacciones en el aula identificando aspectos relacionados con la práctica pedagógica, el manejo de las actividades y el clima emocional (Pianta, La Paro y Hamre, 2008) y el MQI, que busca medir la calidad de las interacciones en las clases de matemáticas entre el docente, los estudiantes y el contenido (Hill, Kapitula y Umland, 2011), siguen buscando la manera de mejorar sus mediciones. En el primer caso en relación con la estructura dimensional del instrumento y en el segundo a través de estudios de generalizabilidad. Por lo tanto, se entiende que este tipo de estudios requiere de procesos de rediseño y verificación constante. En este caso, para el POAEM será necesario hacer modificaciones que implicarán: la revisión de los criterios incluidos en las rúbricas considerando los hallazgos de las sesiones de calibración; elaborar un nuevo diseño para el estudio de generalizabilidad, que considere el número de días de observación, y una cantidad única de tareas por ocasión para cada docente.

Se espera que los resultados de este estudio y su proceso de desarrollo sirva como referente para el desarrollo de instrumentos que busquen evaluar las prácticas docentes, ya que como se mencionó en capítulos anteriores los estudios metodológicos son poco frecuentes en nuestro país, especialmente aquellos que buscan medir las prácticas de enseñanza en matemáticas en el contexto mexicano (López y Mota, 2003; Ávila *et al.*, 2013). A continuación se exponen las aportaciones específicas de esta versión del instrumento, con respecto a: la medición de las prácticas de enseñanza y usos futuros para el POAEM; la observación y videograbación de las prácticas; y las líneas de investigación futuras que se podrían abordar.

1. Medición de las prácticas de enseñanza y usos futuros para el POAEM

Por tratarse de un estudio metodológico, el POAEM no reporta resultados sobre el desempeño de los docentes, ya que los datos que se recabaron se emplearon como insumo para realizar las pruebas de validez y confiabilidad de la información. Además, es importante señalar que por el propósito que subyace al instrumento y por el requerimiento de una capacitación previa para su utilización, los principales usuarios del POAEM deberán ser investigadores que estén interesados en estudiar las actividades de enseñanza en matemáticas. Lo anterior partiendo de la necesidad del desarrollo de instrumentos que apoyen la investigación a nivel nacional, así como en la aspiración de lograr una evaluación con instrumentos de mejor calidad.

Algunos usos que se podrían considerar para el POAEM tienen que ver con ofrecer retroalimentación de la practica de enseñanza, a través de la devolución de los observadores a los actores. Sin embargo, esto tendrá que ser el propósito de un nuevo proyecto de investigación que además deberán considerar la obtención de evidencias de validez de la información generada para tal fin. Asimismo, este es un instrumento que podría ser utilizado por actores educativos como los supervisores escolares, quienes como parte de sus actividades deben realizar observaciones de la práctica docente con base en aspectos que les permitan regresar información valida y confiable a los docentes.

Es imprescindible señalar que esta versión del POAEM sólo se ha puesto a prueba considerando dos observadoras distintas al actor que está siendo observado, quienes han pasado por un proceso específico de entrenamiento. Para otras formas de uso del instrumento será necesario repensar y probar nuevos procesos de implementación e identificar las que resulten más adecuadas. En cualquier caso, en principio será necesario realizar los ajustes señalados para el POAEM y obtener nuevas evidencias de la validez y confiabilidad de la información que se recabe.

2. Las observadoras y las videograbaciones en el proceso de validación del POAEM

A continuación se mencionan dos elementos importantes en el desarrollo de este estudio. El primero tiene que ver con las observadoras que realizaron la calificación de las actividades de enseñanza y el segundo con las videograbaciones como insumos para probar la información recabada con el instrumento. A partir de la experiencia con el desarrollo del POAEM se encontró que se requieren cuidados importantes para la selección de las observadoras, así como para la recolección de la información a través de las videograbaciones.

Para seleccionar a las observadoras se estableció como perfil deseado que tuvieran experiencia en educación primaria. Las personas seleccionadas fueron capacitadas para utilizar el protocolo de observación en sesiones individuales y en sesiones en las que se trabajó conjuntamente con las dos observadoras. Los resultados obtenidos a partir de los análisis estadísticos sugieren que los cuidados tanto en la selección, capacitación, así como en el seguimiento que se realizó durante la calificación de las actividades de enseñanza permitieron que la variación atribuida al observador fuera reducida de acuerdo con los resultados del estudio de generalizabilidad.

En cuanto a las videograbaciones, en el levantamiento de los datos se tomaron decisiones que favorecieron su uso, dentro de ellas se encuentra: haber ofrecido una explicación de la presencia de las cámaras a los estudiantes, pues así ellos se familiarizaron rápidamente con los equipos de grabación; haber contactado a cada uno de los docentes para explicarles el proyecto, proporcionar un documento para que informaran su consentimiento de participación en el proyecto, permitió que tuvieran claridad sobre los alcances de la investigación y los usos que se darían a las videograbaciones, facilitando su disposición para participar en el proyecto y promoviendo que reaccionaran con naturalidad durante las videograbaciones de sus clases.

Por otro lado, algunas condiciones de las aulas dificultaron las videograbación, en particular el tamaño del salón y el número de alumnos en los grupos, puesto que grupos numerosos en salones reducidos dejaban poco espacio para el acomodo del equipo de videograbación de tal manera que no se obstruyera el paso a los estudiantes y se lograran tomas que permitieran identificar las acciones del docente, de los alumnos, las interacciones entre ellos, y los recursos utilizados.

Algunas sugerencias para futuros estudios que utilicen videograbaciones son: a) establecer el actor al que se dará prioridad durante la videograbación, por ejemplo, mantener siempre las tomas con el docente; b) determinar qué recursos se utilizarán para cuidar la calidad de la información, lo cual comprenden el tipo de cámara de videograbación y la calidad de imagen; c) especificar las herramientas a emplear para la recolección, almacenamiento, análisis y reporte de la información, por ejemplo, dispositivos de almacenamiento portátiles, fijos o virtuales; y d) las consideraciones éticas con respecto a la recolección, uso y almacenamiento de la información, por ejemplo, la responsabilidad de los investigadores para guardar el anonimato de los participantes, así como cuidar que la información obtenida no se filtre en plataformas públicas que pudieran afectar la integridad de los participantes (Derry et al, 2010).

3. Líneas de investigación futuras

En México se han implementado nuevas políticas educativas para evaluar el desempeño docente, considerando que su papel y función, hasta ahora irreemplazable, es una pieza clave para lograr la mejora de la calidad de la educación. Por ejemplo, la Ley General del Servicio Profesional Docente, establece nuevos mecanismos de reclutamiento, selección, permanencia y promoción del personal a través de la valoración de su desempeño. Sin embargo, para lograr una evaluación adecuada se requiere de evidencias de la actuación de los docentes en el aula y en otros espacios escolares, para compararlas con parámetros e indicadores que determinen una práctica de enseñanza eficaz (INEE, 2015).

Ante este contexto, se vuelve necesario establecer “criterios, procedimientos e instrumentos diseñados de acuerdo con las necesidades de calidad y equidad para regular, orientar y evaluar el desempeño docente” (SEP, 2013, p. 35). No obstante, a nivel nacional falta recorrer un largo camino para el desarrollo de instrumentos que aporten bases sólidas para obtener información válida y confiable sobre diferentes aspectos de la práctica de enseñanza (López y Mota; 2003 y Ávila *et al*. 2013). Por lo tanto, es importante desarrollar estudios metodológicos centrados en la elaboración de instrumentos que permitan obtener información de mejor calidad tanto en el campo de las matemáticas como en otras asignaturas.

La investigación educativa reconoce que el diseño de instrumentos es un proceso complejo y laborioso que requiere de varias etapas para ajustar las diferentes versiones, de suerte que en algún momento sea posible sustentar generalizaciones. Además se señala que para medir un constructo tan complejo como las prácticas de enseñanza se requiere establecer enfoques claros y sistemáticos para recolectar, analizar y usar la información que se recabe. Es decir, es imprescindible definir claramente qué se va a evaluar, para qué se va a evaluar, cómo se va a evaluar, y de qué manera se utilizarán los resultados que se obtengan (Stiggins et.al. 2007).

Considerando lo anterior, valdría la pena pensar en estudios metodológicos-longitudinales que permitan dar seguimiento para probar y mejorar los instrumentos y lograr mediciones cada vez más confiables. Por ejemplo, si se tuviera la oportunidad de seguir a un grupo de docentes por un periodo largo de tiempo (años) videograbando sus clases y empleando esta información para realizar diferentes pruebas y ajustes al POAEM, no sólo se estarían midiendo ciertos aspectos de la práctica de enseñanza (dimensiones), sino que se tendrían elementos cada vez más confiables para ofrecer información que favorezca la mejora de la práctica docente a través de una retroalimentación clara y precisa.

Asimismo se podría emplear la información de contexto que se recabó con este estudio: número de estudiantes que atiende el docente, niños con Necesidades Educativas Especiales (NEE), tipo de NEE, años de experiencia, materiales que toma como referencia para las planeaciones, etc. para plantear preguntas de investigación que permitan contextualizar las prácticas de enseñanza a partir de las condiciones que enfrentan los docentes para llevar a cabo su labor.

Además de las línea de investigación futuras, es importante mencionar dos aportes que este estudio ha generado a partir de su desarrollo. El primero tiene que ver con establecer dimensiones específicas para valorar las actividades de enseñanza, identificadas a partir de la revisión de literatura, las cuales son coincidentes con instrumentos internacionales para la evaluación de la práctica docente (por ejemplo, MQI, CLASS, y IQA), y con algunas de las orientaciones del modelo educativo para la educación obligatoria vigente al momento del levantamiento de datos.

El segundo aporte es la documentación del proceso de diseño y proceso de obtención de evidencias de validez y confiabilidad de la información recabada con el POAEM, ya que sirve como referente para otros estudios metodológicos que tengan como propósito diseñar instrumentos para evaluar las prácticas docentes. No obstante, se debe aclarar que para dar continuidad a este estudio será necesario realizar los ajustes propuestos para el POAEM, de tal manera que se logre una versión mejorada para realizar nuevas pruebas de validez y confiabilidad de la información que se recabe con otra versión del instrumento. Además, será necesario tener en cuenta elementos adicionales para que eventualmente se pueda ofrecer mayor información a los docentes que participen en la evaluación, tales como un sistema de calificación general por docente y un protocolo de retroalimentación.

Referencias

Alsina, A., y Coronata, C. (2014). Los procesos matemáticos en las prácticas docentes: diseño, construcción y validación de un instrumento de evaluación. *Edma 0-6: Educación Matemática en la Infancia, 3*(2), 23-36.

Aubuson, P., Burke, P, Schuck, S. y Kearney, M. (2014). Teachers choosing rich taks: The moderating impac of technology on student learning, enjoyment, and preparation. En *Educational Reseacher*, Vol. 43, No. 5, pp. 219-229

Avery, P. y Freeman, C. (1999). Authenctic instrution and assessment. Social Education. Recuperado de http://www.eeraonline.org/journal/files/2002/JRE_2002_04_Avery.pdf

Ávila et. al. (2013). Una década de investigación educativa en conocimientos disciplinares en México. Matemáticas, ciencias naturales, lenguaje y lenguas extranjeras 2002-2011. México: ANUIES-COMIE

Bell, C., Qi, Y., Croft, A., Leusner, D., McCaffrey, D., Gitomer, D. y Pinta, R. (2015). Improving observation Score Quality. Challenges in observer thinking, Chapter 3 …

Borko, H., Stecher, B. M., y Kuffner, K. (2007). *Using artifacts to characterize reform-oriented instruction: The Scoop Notebook and rating guide*. Los Angeles, California: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Boston y Smith (2011). A task-centric approach to professional development: enhancing and sustaining mathematics teacher's ability to implement cognitively challenging mathematical tasks*. ZDM Mathematics Education*. Recuperado de http://link.springer.com/article/10.1007%2Fs11858-011-0353-2#/page-1

Boston, M. (2012). Assessing Instructional Quality in Mathematics. *The Elementary School Journal,* Vol. 113 (1), pp. 76-104.

Boston, M. (2012). Assessing instructional quality in mathematics. *The Elementary School Journal, 113*, 76-104.

Boston, M. y Wolf, M. (2006). Assessing academic rigor in mathematics instruction: The development of instructional quality assessment toolkit. Reporte técnico 672. Recuperado de http://www.cse.ucla.edu/products/reports/r672.pdf

Brennan, R. L. (2001). Generalizability Theory. New York: Springer-Verlag.

Brown, M. (2002). Researching Primary Numeracy. In A. D. Cockburn y E. Nardi (Eds.), *Proceedings of the Annual Meeting of the International Group for the Psychology of Mathematics Education* (pp. 15-30). Norwich, England.

Cohen, D., Raudenbush, S. y Ball, D. (2003). Resources, Instruction, and Research. *Educational Evaluation and Policy Analysis*, Vol. 25, No. 2, pp. 119-142

Daniels, H. (2003). Vygotsky y la Pedagogía. Barcelona: Paidós

Diario Oficial de la Federación (2013). *Lineamientos iniciales específicos para llevar a cabo la evaluación del ingreso al Servicio Profesional Docente en Educación Básica y Educación Media Superior y lineamientos iniciales específicos para llevar a cabo la evaluación para la promoción a cargos con funciones de Dirección (Directores) en Educación Media Superior, para el ciclo escolar 2014-2015. LINEE-02-2014.* Recuperado de http://www.inee.edu.mx/images/stories/2014/Normateca/2014/Lineamientos_Iniciales_Espec%C3%ADficos_inee.pdf

Diario Oficial de la Federación. (2014). Acuerdo número 717 por el que se emiten los lineamientos para formular los Programas de Gestión Escolar: Autor.

Donovan, M. y Bransford, J. (2005). How Students Learn: History, Mathematics and Science in the Classroom. National Reseach Council, Committe on How people Learn: A targeted repport for teachers. Washington, D.C: National Academies Press

Doyle, W., y Carter, K. (1984). Academic tasks in classrooms. Curriculum Inquiry, 14, 129-149.

Fernández y Yoshida (2004). Lesson Study: A japanese Approach to improving mathematics teaching and learning. Mahwah, New Jersey: Laurence Erlbaum Associates.

Fernández, A. (2015). Aplicación del análisis factorial confirmatorio a un modelo de medición del rendimiento académico en lectura. Ciencias Económicas 33 (2) 39-66. Recuperado de https://revistas.ucr.ac.cr/index.php/economicas

Gitomer, D., Bell, C., Qi, Y., McCaffrey, D., Hamre, B. y Pianta, R. (2015).The instructional challenge in improving teaching quality: lessons from a classroom observation protocol (en prensa).

Grabinger, R. S., y Dunlap, J. C. (1995). Rich environments for active learning: A definition. *Research in Learning Technology*, *3*(2). Recuperado de http://www.researchinlearningtechnology.net/index.php/rlt/article/view/9606

Griffin Pete. (2009). What makes a rich task? Association of teachers of mathematics. for mathematics educators primary, secondary and higher. Recuperado de http://www.atm.org.uk/write/MediaUploads/Journals/MT212/Non-Member/ATM-MT212-32-34.pdfGulikers, J. T., Bastiaens, T. J., y Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, *52*(3), 67–86.

Grossman, P., Greenberg, S., Hammerness, K., Cohe, J. Alston, C. y Brown, M. (2009). Development of the Protocol for Language Arts Teaching Observation (PLATO). Paper presented at the annual meeting of the American Educational Research Association, San Diego California.

Guilkers, J., Bastiaens, T. y Kirschner, P. (2004). A five-Dimensional Framework for authentic assessment. ETRyD Vol. 52, No. 3. Pp. 67-86

Hart, H., Cowhy, J., Matsko, K. y Sporte, S. (2015). Demands for Authentic Intellectual Work in Teacher Assignments. Documento presentados en el congreso anual de la American Educational Reseach Association. Recuperado el 04 de agosto de 2015 en AERA online paper repository.

Hernández, Fernández y Baptista (2010). Metodología de la investigación (quinta edición). Empresa Editora El comercio S.A: Perú.

Herrero, Juan. (2010). El Análisis Factorial Confirmatorio en el estudio de la Estructura y x Estabilidad de los Instrumentos de Evaluación: Un ejemplo con el Cuestionario de Autoestima CA-14. *Psychosocial Intervention*, *19*(3), 289-300. Recuperado de http://scielo.isciii.es/scielo.php?script=sci_arttextypid=S113205592010000300009 ylng=esytlng=es

Hiebert, J. Carpenter, T., Fennema, E., Fuson, K., Wearne, D. Murray, H. Oliver, A., y Humen, P. (1997). Making sense: Theaching and learning mathematics with understanding. Portsmouth, NH: Hienemann.

Hiebert, J. y Wearne, D. (1993). Instructional tasks, classroom discourse, and students´ learning in second-grade arithmetic. American Educational Reseach Journal 30, no 2. pp. 393-425

Hill, H. C., Blunk, M. L., Charalambos, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., y Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An Exploratory Study. *Cognition and Instruction, 26*, 430-511.

Hill, H., Kapitula, L. y Umland, K. (2011). A validity argument approach to evaluating value-added score. American Educational Reseach Journal, 48 (3). 794-831.

Hill, Heather, Charalambos,Y., Kraft, M. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, 41 (2), 56-64.

Jaworski, B., Goodchild, S., Eriksen, S. y Daland, E. (2011). Mediating Mathematics Teaching Development and Pupil´s Mathematics Learning: The life cycle of a Task

Jonassen, D. (2000). Toward a design theory of problem solving. Educational Technology: Research y Development, 48 (4), pp. 63-85.

Kramarski, B., Mevarech, Z. R., y Arami, M. (2002). The effects of metacognitive Instruction on solving mathematical authentic tasks. *Educational studies in mathematics*, *49*(2), 225–250.

Llinares, S. (2008). Aprendizaje del estudiante para profesor de matemáticas y el papel de los nuevos instrumentos de comunicación. *III Encuentro de Programas de Formación Inicial de Profesores de Matemáticas.* Santa Fe de Bogotá, Colombia: Universidad Pedagógica Nacional.

López y Mota, Á. (2003). Saberes científicos, humanísticos y tecnológicos: procesos de enseñanza y aprendizaje. México: Consejo Mexicano de Investigación Educativa

Lucci, M. A. (2006). La propuesta de Vigotsky: La psicología socio-histórica. *Revista* de currículum y formación del profesorado, Vol. 10, No. 2, pp. 1-11.

Martínez-Slgado (2012). El muestreo en investigación cualitativa. Principios básicos y algunas controversias. Recuperado de http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232012000300006

Martínez, F. (2012). Procedimientos para el estudio de las prácticas docentes. Revisión de la literatura. RELIEVE. Recuperado de http://www.uv.es/RELIEVE/v18n1/RELIEVEv18n1_1.pdf

Martínez, F. (2012). *Las matrices de valoración o rúbricas.* Recuperado de http://www.youtube.com/watch?v=oLCxt8BhRcU

Martínez, F. (2013). *La evaluación, parte integral de la práctica docente* (Diapositivas de PowerPoint). Recuperado de: https://www.youtube.com/watch?v=zh9Jhpc5ZVI

Martínez, F. (2014). El estudio de las prácticas docentes. En G. Ruiz (Coord.), Acercamientos empíricos las prácticas de evaluación en el aula en Educación Básica (pp. 17-71). México: Universidad Autónoma de Aguascalientes.

Matínez, J.F; Taut, S. y Schaaf, (2016). Classroom observation for evaluating and improving teaching: An international perspective. *Studies in Educational Evaluation* 49, 15-29.

McTighe, J. y Wiggins, G. (2012). Understanding by design framework. Consultado en: http://www.ascd.org/ASCD/pdf/siteASCD/publications/UbD_WhitePaper0312.pdf, el 15 de mayo de 2015.

McTighe, J., Seif, E. y Wiggins, G. (2004). You can teach for meaning. En Educational Leadership. 62 (1), pp. 26-31.

Mendoza, J. y Garza, J. (2009). La medición en el proceso de investigación científica: Evaluación de contenido y confiabilidad. México: UANL.

MET Project (2010f). *Validation Engine for Observational Protocols*. Bill y Melinda Gates Foundation.

MET project. (2010a). Overview: *Teacher Observation Rubrics*. Measures of Effective Teaching. Recuperado de http://www.ocmboces.org/tfiles/folder896/METrubricexamples.pdf

MET project. (2010b). *Danielson's Framework for Teaching for Classroom Observations.* MET project Research paper. Bill y Melinda Gates Foundation. Recuperado de http://metproject.org/resources/Danielson%20FFT_10_29_10.pdf

MET project. (2010c). *The MQI Protocol for Classroom Observations.* MET project Research paper. Bill y Melinda Gates Foundation. Recuperado de http://metproject.org/resources/MQI_10_29_10.pdf

MET project. (2010d). *The PLATO Protocol for Classroom Observations.* MET project Research paper. Bill y Melinda Gates Foundation. Recuperado de http://www.metproject.org/resources/PLATO_10_29_10.pdf

MET project. (2010e). *The CLASS protocol for Classroom Observations.* MET project Research paper. Bill y Melinda Gates Foundation. Recuperado de http://metproject.org/resources/CLASS_10_29_10.pdf

Mohr, A. (2000). Análise do conteúdo de saúde em livros didáticos. Ciencia y Educacao, 6(2), pp.89-106

National Council of Teachers of Matematics. (2015). De los principios a la acción para garantizar el éxito matemático para todos. México: Libros S.A.

National Council of Teachers of Mathematics. (2000). Principles and standars for school mathematics. Reston: NCTM

Newemann, F. y Wehlage, G. (1993). Five standards of authentic instruction. En *Educational Leadership*. 50 (7), pp. 8-12.

Newmann, F. Bryk, A. y Nagaoka, J. (2001). Authentic Intellectual Work and Standardized Test: Conflicto r coexistence?. Chicaho: Consortium on chicago School Research.

Newmann, F., Marks, H. y Gamoran, A. (1996). Authentic pedagogy and studentperformance. *American Journal of Education,* 104 (4), pp. 280-312.

Newmann, F.M., Lopez, G. y Bryk, A.S. (1998) *The quality of intellectual work in Chicago schools: A baseline report*. Consortium on Chicago School Research, Chicago.

Nicaise, M., Gibney, T., y Crane, M. (2000). Toward an understanding of authentic learning: student perceptions of an authentic classroom. *Journal of Science Education and Technology*, *9*(1), 79–94.

Oura, G. K. (2001). Authentic task-based materials: Bringing the real world into the classroom. Recuperado de https://www.jrc.sophia.ac.jp/pdf/research/bulletin/ki21/gaio.pdf, el 17 de mayo de 2015.

Pedroza, H. (2016). Desarrollo de una estrategia de evaluación formativa para el profesorado de educación preescolar (Tesis doctoral). Ensenada B.C: México.

Pérez, J. A., Chacón, S. y Moreno, R. (2000). Validez de constructo: El uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. Recuperado de http://www.psicothema.com/pdf/601.pdf

Perkins, D. (1995). *La escuela inteligente*. España: Editorial Gedisa

Pianta, R., La Paro, K. y Hamre, B. (2008). *Classroom Assessment Scoring System. Manual K-3.* Baltimore, Maryland: Paul H. Books Publishing Co.

Picaroni, B. y Loureiro, G. (2010). ¿Qué matemática se enseña en aulas de sexto año de Primaria en escuelas de Latinoamérica?. *Páginas de Educación*, Vol. 3 (1). Recuperado de: https://revistas.ucu.edu.uy/index.php/paginasdeeducacion/article/view/657

Picaroni, B. y Loureiro, G. (2010). Qué matemáticas se enseña en aulas de sexto año de x primaria en escuelas de Latinoamérica. En *Páginas de Educación.* Vol. 3, pag. 29-60.

Putnam, R., Lampert, M. y Peterson, P. (1990). En C.B. Cazden, (Ed.), *Review of reseach in education*. Vol. 16 pp. 57-150. Washington, DC: American Educational Reseach

Ravela, P. (2009). Consignas, devoluciones y calificaciones: problemas de la evaluación en aulas de educación primaria en América Latina. *Páginas de Educación*, Año 2 N°: 49-89.

Rositas, J.M. (2006). Factores críticos de éxito en la gestión de calidad y su grado de presencia e impacto en la industria manufacturera mexicana. Monterrey. México: UANL.

Santos, M. (2003). Hacia una instrucción que promueva los procesos de pensamiento matemático. En E. Filloy (coord). *Matemática Educativa: Aspectos de la investigación actual*, pp. 314-332. FCE: México.

Schoenfeld, A. (2000). A highly interactive discourse structure. Social constructivist teaching, 9, pp. 131-169

Secretaría de Educación Pública. (1993). Plan de Estudios. Educación Básica. Primaria. México:SEP

Secretaría de Educación Pública. (2009). Plan de Estudios. Educación Básica. Primaria. México:SEP

Secretaría de Educación Pública. (2011). Plan de Estudios. Educación Básica. México: SEP

Secretaría de Educación Pública. (2011a). Programas de estudio 2011. Guía para el maestro. Educación Básica Primaria. Sexto Grado. México:SEP

Secretaría de Educación Pública. (2011b). Plan de Estudios. Educación Básica. México: SEP

Secretaría de Educación Pública. (2013). Un Modelo de Gestión para la Supervisión Escolar. Modulo V. Consultado el 09 de agosto de 2015 en: http://basica.sep.gob.mx/pec/pdf/dprograma/MatGestModulo5.pdf

Secretaría de Educación Pública. (2017). Modelo Educativo para la Educación Obligatoria. México: SEP.

Shavelson, R. J., y Webb, N. M. (1991). *Generalizability Theory. A Primer*. Thousand Oaks, California: SAGE Publications, Inc.

Siller, H., Gotz, T., Bracke, M. y Bock, W. (2014). A new concept for the difficulty classification of modeling problems, en proceso.

Solano-Flores, G., Li, M. (2009). Generalizability of cognitive interview-based measures across cultural groups. *National Council on Measurement in Education*, 9-17.

Stallings, J. A. (1986). Effective use of time in secondary reading programs. In J. Hoffman (Ed.), *Effective teaching of reading: Research and practice* (pp. 85-106).

Stein, M. y Lane, S. (1996). Instructional tasks ans the development of students capacity to think and reason: An analysis of the relationship between teaching and learning in a reform Mathematics Project. En Educational Reseach and Evaluation. No. 1 pag. 50-80.

Stein, M., Grover, B. y Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysisi of mathematical task used in reform classroom. *American Educational Research Journal, 33* 455-488.

Stein, M., Smith, M., Henningsen, M., y Silver, E. (2009). Implementing standards-based mathematics instructions: A casebook for professional development. New York: Teachers College Press.

Stewar, T., Roebber, P. y Bosart, L. (1997). The importance of the task in analyzing expert judgment. Organizational Behavior and Human Decision process, 69 (3), 205-219.

Stigler, J. W., y Hiebert, J. (2004). Improving Mathematics Teaching. *Educational Leadership, 61*(5), 12-17.

Stigler, James, Gallimore, R., Hiebert, J. (2000). Using Video Surveys to Compare Classrooms and Teaching Across Cultures: Examples and Lessons from the TIMSS Video Studies. *Educational Psychologist,* Vol. 35 (2): 87-100.

Stone, M. (2008). ¿Qué es la enseñanza para la comprensión?. En Stone, M. (Ed.). *La enseñanza para la comprensión. Vinculación entre la investigación y la práctica*. Buenos Aires: Paidós

Swan, M. (2015). Designing tasks and lesson that develop conceptual understanding, strategic competence and critical awareness. Center for Research in Mathematics Education: University of Nottingham.

Taut, S., Manzi J. y Molina, A. (2013). *La evaluación de las prácticas docentes.* Pensamiento Educativo. Revista de Investigación Educacional Latinoamericana, 2013, 50 (1), 1-3. Recuperado de http://www.eduglobal.cl/2013/05/21/revista-pensamiento-educativo-la-evaluacion-de-las-practicas-docentes/

Trigueros, M. (2009). ¿Qué hemos aprendido de la enseñanza de las matemáticas a través de la investigación?. ¿qué retos enfrentamos?. En De Alba, A. y Glazman, R. (Coord.). ¿Qué dice la investigación educativa? pp. 27-69. México: COMIE

Verschaffel, Greer y De Corte (2010). Mathematics learning. International Handbook of Mathematics Education, Vol. 3 pp. 401-405.

Wiggins, G. (1989). The futility of trying to teach everything of importance. En *Educational Leadership,* pp. 44-59.

Wiggins, G. (1998) Rúbricas para la evaluación. Selección y traducción por J. VIÑAS y P. Ravela. En *Educative Assessment. Designing Assessments to Inform y Improve Student Performance.* San Francisco, Jossey-Bass, Cap. 7.

Wiggins, G. y McTighe, J. (2008). Put understanding first. En Educational Leadership. 65 (8), pp. 36-41.

Wolff, M. R. y Yew J. L. (2007). Vygotsky´s Neglected Legacy: Cultural Historical Activity Theory. *Review of Educational Reseach,* Vol. 77, No. 2, pp. 186-232.

Bibliografía

Andrade, Heidi G. (s/f). Cuando la Valoración es instrucción y la instrucción es valoración: utilizando matrices analíticos para promover el pensamiento complejo y la comprensión. En Herland, Lois y Sh. Veenema, *Proyecto Cero*, pp.91-99.

Archbald, D. y Newmann, F. (1988). Beyond standardized testing: Assessing authentic academic achievenment in secondary schools. Reston, VA: National Association of secondary school principals.

Ávila Baray, H. L. (2006). *Introducción a la metodología de la investigación.* Edición electrónica. Recuperado de http://www.eumed.net/libros/2006c/203/

Ávila, A; (1996). Los usos reconocidos de los textos de matemáticas. *Revista Mexicana de Investigación Educativa,* (1). Recuperado de http://www.redalyc.org/articulo.oa?id=14000204

Ball, D., Hoover, M., y Phelps, G. (2008). Content knowledge for teaching: What makes it special?. En Journal of Teachers Education No. 59, pag. 389-407.

Barrera, F. y Santos, M. (2000). Hacia un documento de referencia del currículo para el nivel medio superior. En sociedad matemática mexicana. Reporte parcial, Conacyt: México.

Bell, A. (1993). Some experiments in diagnostic teachinng. Educational Studies in Mathematics, vol. 24. No. 1

Bransford, J. D., Brown, A. L., Cocking, R. R. (Eds.) (1999). How People Learn: Brain, Mind, Experience, and School: Expanded Edition. DC: National Academy Press.

Bransford, J. y Stain, B. (1984). The IDEAL problem solver: A guide for improvig thinking, learning, and creativity. New York: W.H. Freeman

Caixeta de Castro Lima, M. y De Souza Silva, P. (2010). Criterios que profesores de química como orientadores da escolha do libro didáctico. Ensaio- Pesquisa Educacao en Ciencias, 12 (2), pp. 121-135.

Cassab, M. Martins, I. (2008). Significacoes de profesores de ciencias a respeito do livro didáctico. Ensaio- Pesquisa Educacao en Ciencias, 10 (1), pp. 97-116.

Centro Virtual Cervantes (2016). *Actividad de enseñanza/aprendizaje*. Recuperado de http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/activaprendizaje.htm

Centro Virtual Cervantes (2016). *Actividad de enseñanza/aprendizaje*. Recuperado de http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/activaprendizaje.htm

Chiappetta, E. Setha, G., Fillman, D. (1993). Do middle school life science textbooks privide a balance of scientific literacy themes? Journal of Research in Science Teaching, 30 (7). Pp. 787-797.

Cohen, R. Yarden, A. (2010). How the curriculum Guideline. The cell is to be studied Longitudinally is expressed in six Israeli Junior-High School Textbooks. Journal of Science Education and Technology, 19 (3), pp. 276-292.

Cooper, J. (1999). *Estrategias de enseñanza. Guía para una mejor enseñanza*. México: Limusa Noriega Editors

Cooper, J. (1999). *Estrategias de enseñanza. Guía para una mejor enseñanza*. México: Limusa Noriega Editors

Cortina, J. L. (2006). Las mediciones de la calidad del aprendizaje matemático en México: ¿qué nos devela la prueba PISA 2003 y cómo podemos responder?. *Educación Matemática,* 18(1) 161-176. Recuperado de http://www.redalyc.org/articulo.oa?id=40518107

Cross, D., Hudson, R., Adefone, O., Lee, M., Rapacki, L. Y Pérez, A. (2012). Success made orobable: African-american Girl´s exploration in statistics through. *Journal of Urban Mathematics Educations* Vol. 5 No. 2, pp. 55-86

Definición.org (2016). *Definición de actividad*. Recuperado de http://www.definicion.org/actividad

Definición.org (2016). *Definición de actividad*. Recuperado de http://www.definicion.org/actividad

Doyle, K. (2007). The teacher, the tasks: their rol in students? Mathematical literacy. In Watson, Jane and Beswick, Kim Eds. *Proceedings 30th anual conference of mathematics Educational Reseach Group of Australasia-Mathematics: Essential Research. Essential Practice*, pages pp.246-254, Hobart, Tasmania.

Escudero, I. y Sanchez, V. (2008). A Mathematics Teachers’ perspective and its relationship to practice. *International Journal of Science and Mathematics Education,* 6(1) 87-106.

Fan, L. y Zhu, Y. (2007). Representation of problema-solving procedures: Acomparative look at China, Singapore and US mathematics books. Recuperado de http://link.springer.com/article/10.1007/s10649-006-9069-6#/page-1

Filloy, E, Rojano, T. y Rubio, G. (2001). Proposition concerning the resolution of arithmetical-algebraic problems. En Sutherland, S., Rojano T. Bell, A. y Lins, R. (Eds.) Perspetives on school algebra, Dordrecht, kluwer, pp. 155-175.

Gardner, H. (2000). La educación de la mente y el conocimiento de las disciplinas, Barcelona: Paidós, pag. 136-157.

Gavilán, J., García, M., y Llinares, S. (2007). Una perspectiva para el análisis de la práctica del profesor de matemáticas. Implicaciones metodológicas. *Enseñanza de las Ciencias*, 157-170.

Gobierno de la República (2014). *Reforma Constitucional en Materia Educativa.* Recuperado de http://www.presidencia.gob.mx/reformaeducativa/#leyes-secundarias

Goffree, F. (2000). Principios y paradigmas de una educación matemática realista. Matemáticas y educación. Retos y cambios desde una perspectiva internacional. Barcelona, Graó, vol. 9 pp. 151-167.

Gómez *et al*. (2014). El trabajo por proyectos en educación primaria en México: análisis de las propuestas curriculares en la reforma educativa. En *Revista Citecsa*. Vol 5 (8). Recuperado de http://revista.unipaz.edu.co/ojs/index.php/revcitecsa/index

Gómez, B. (2000). Los libros de texto de matemáticas. Recuperado de http://www.sinewton.org/numeros/numeros/43-44/Articulo14.pdf

González, Espino, G., y González, S. (2006). La enseñanza de las matemáticas en las escuelas de primarias de México (Distrito Federal) durante el Porfiriato: programas de estudio, docentes y prácticas escolares. *Educación matemática*, 18 (3), pp. 39-63.

Hill, H. (2010). Mathematical Quality of Instruction (MQI). Coding Tool. University of Michigan, Learning Mathematics for Teaching.

Hill, H. C., Ball, D. L., Blunk, M. L., Goffney, I. M., y Rowan, B. (2007). Validating the ecological assumption: the relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives, 5*(2-3), 107-118.

Hill, P. y Rowe, K. (1998) Modeling student progress in studies of educational effectiveness. *School Effectiveness and School Improvement,* 9(3), pp. 310–333.

http://www.ocmboces.org/tfiles/folder896/METrubricexamples.pdf

Instituto Nacional para la Evaluación de la Educación (2015). Evaluación del desempeño de docentes. Recuperado de http://www.inee.edu.mx/index.php/component/content/article/557-dialogos-con-x docentes/1720-lo-que-debemos-saber-de-la-evaluacion

Jackson, K., Garrison, A., Wilson, J. Gibbons, L. y Shahan, E., (2013). Exploring relationships between setting up complex and opportunities to learn in concluding whole-class Discussions in Middle-grades Mathematics instruction. *Journal for research in mathematics Educations*, Vol. 44, No. 4, pp.646-682.

Johnsen, E. (1993). Textbooks in the kaleidoscope. A critical survey of literatura and reseach on Eductional Texts. Olso: Scandinavian University Press.

Jonassen, D. (1997). Instructional Design Models for Well_Structured and Ill-Structured Problem-Solving Learning Outcomes. *ETRyD* Recuperado de http://www.webkelley.com/HBS/ID%20Models%20for%20Well-Structured.pdf

Kulikowich, J. y DeFranco, T. (2003). Philosophy´s rol in characterizing the nature of educational pychology and mathematics. *Educational Psychologist*, 38 (3), 147-156.

Lamb, S. y Fullarton, S. (2002) Classroom and school factors affecting mathematics achievement: A comparative study of Australia and the United States using TIMMS. *Australian Journal of Education,* 46(2), pp. 154–171.

Larios Lozano, M. d. C. (2001). Los libros de texto gratuitos. *Revista Mexicana de Investigación Educativa,* 6(12). Recuperado de http://www.redalyc.org/articulo.oa?id=14001202

Llinares y Krainer, S. K. (2006). Mathematics (student) teachers and teacher educators as learnes. En A. P. Gutiérrez y Boero, *Handbook of research on the psychology of mathematics education: past, present and future.* (págs. 429-459). Rotterdam: The Netherlands: Sense Publishers.

Martínez Losada, C., García Barros, S. y Rivadulla López, J. (2009). Qué saben los alumnos de primaria y secundaria sobre los sistemas materiales. Cómo lo tratan los textos escolares. Revista Electrónica de Enseñanza de la Ciencias, 8 (1), pp. 137-155.

Martínez Rizo, F. (2014). Capítulo 5. *Evaluación de los docentes, mejora profesional y la educación.* pp. 1-52 en prensa.

McClain, K., Cobb, P. y Bowers, J. (1998). A contextul investigation of three-digit addition and subtraction. In L.J. Morrow y M.J. Kenney (Eds.). *The teaching and learning of algorithms in school mathematics*. Reston: NCTM

Occelli, M. y Valeiras, N. (2013). Los libros de texto de ciencias como objeto de investigación: una revisión bibliográfica. *Enseñanza de las ciencias*, 31 (2). Recuperado de http://www.raco.cat/index.php/Ensenanza/article/viewFile/285774/373774

Organización para la Cooperación y el Desarrollo Económicos. (2012). El Programa de PISA de la OCDE. Qué es y para qué sirve. Recuperado de http://www.oecd.org/pisa/39730818.pdf

Parra, B. (1991). La resolución de problemas en la construcción de esquemas de razonamiento. Matemática educativa. México: Cinvestav

Perrenoud, P. (2012). *Cuando la escuela pretende preparar para la vida. Desarrollar competencias o enseñar otros saberes.* Barcelona: Editorial Graó.

Pianta, R. y Hamre, B. (2009). Conceptualization, Measurement and Improvement of Classroom Processes: Standardized Can Leverage Capacity. *Educational Research*, Vol. 38 (2): 109-119

Pianta, R., La Paro, K., y Hamre, B. (2008). Classroom Assessment Scoring System. Baltimore, MD: Paul H. Brookes.

Ponte y Chapman, J. (2008). Preservice mathematics teachers knowledge and development. En English, *Handbook of international research in mathematics education* (p. 225-263). New York: NY: Routledge.

Pozo, J. y Gómez Crespo, M. (1998). *Aprender y enseñar ciencias*. Madrid: Morata. pp. 36-41.

Professional Support and Curriculum Directorate (2003). *Quality teaching in NSW public schools. An annotated bibliography*. New South Wales: Department of Education and Training

Quilez, J. (2009). Análisis de los errores que presentan los libros de texto universitarios de química general al tratar la energía libre de Gibbs. *Enseñanza de las ciencias*, 27 (3), pp. 317-330.

Real Academia Española (2016). *Diccionario de la Lengua Española*. Recuperado de http://dle.rae.es/?id=0chgoNb

Real Academia Española (2016). *Diccionario de la Lengua Española*. Recuperado de http://dle.rae.es/?id=0chgoNb

Rezat, S. (2009). The utilization of mathematics textbooks as instruments for learning. En V. Durand-Guerrier, S. Soury-Lavergne, y F. Arzello (Eds.) *Proceedings of CERME6*, Lyon France. Recuperado de http://ife.ens-lyon.fr/publications/edition-electronique/cerme6/wg7-22-rezat.pdf

Rowe, K. y Rowe, K. (2002) *What matters most: Evidence-based findings of the key factors affecting the educational experiences and outcomes for girls and boys throughout their primary and secondary schooling*. ACER, Melbourne. Consultado en: http: //www.acer.edu.au

Saíz, M. (2008). La enseñanza del volumen en los libros de texto mexicanos: un siglo de educación en México. *Revista de didáctica de las matemáticas*, 48, pp. 76-93

Santos-Trigo, L.M. (2007). *Resolución de problemas matemáticos*. Fundamentos cognitivos. México: Trillas.

Saucedo, M. y Hermosillo, A., (2004). Los libros de texto en la clase de matemáticas. En Alicia Ávila (Dir.). *La reforma realizada. La resolución de los problemas como vía del aprendizaje en nuestras escuelas*. SEP-SEBYN, pp. 165-214.

Secretaría de Educación Pública (2012). Desafíos matemáticos. Libro para el maestro. Sexto grado de primaria. México: SEP

Secretaría de Educación Pública. (2013). Un Modelo de Gestión para la Supervisión Escolar. Modulo V. Recuperado de http://basica.sep.gob.mx/pec/pdf/dprograma/MatGestModulo5.pdf

Shavelson, R. y Webb, N. (2006). Generalizability theory. Recuperado de: https://web.stanford.edu/dept/SUSE/SEAL/Reports_Papers/methods_papers/G%20Theory%20AERA.pdf

Shield, M. y Dole, S. (2012). *Assessing the potencial of mathematics textbooks to promote deep learning*. Recuperado de http://link.springer.com/article/10.1007/s10649-012-9415-9#/page-1

Sierpinska, A. (1994). *Understanding in Mathematics*. London: Falmer

Silva, F., Marinho, K. Y Compiani, M. (2006). Las imágenes geológicas y geocientíficas en los libros didácticos de ciencias. *Enseñanza de las ciencias*, 24 (2), pp. 207-217.

Sowder, J. (2007). The mathematical education and development of teachers. En L. K.F., *Second handbook of research on mathematics teaching and learning.* Charlotte, NC: Information Age Publishing.

Stylianides, G. (2009). Reasoning and proving in school mathematics textbooks. Mathematical Thinking and Learning, Vol.11 (4), pp. 258-288. Recuperado de http://dx.doi.org/10.1080/10986060903253954

Tarr, J., Chávez, O., Reyes, R. y Reys, B. (2006). From the written to the enacted curricula: The intermediary role of middle school mathematics teachers in shaping students´s opportunity to learn. *School Science and Mathematics, 106 (4),* 191-201.

Thomson, S. y Fleming, N. (2004). *Summing it up: mathematics achievement in Australian schools in TIMSS 2002*. Melbourne: Australian Council for Educational Reseach.

Vasconcelos, S. y Souto, E. (2003). O livro didático de ciencias no esino fundamental proposta de criterios para análise de conteúdo zoológico. Ciencia y Educacao, 9 (1), pp. 93-104.

Villalobos, J; (2003). El docente y actividades de enseñanza / aprendizaje: algunas consideraciones teóricas y sugerencias prácticas. *Educere,* (7) 170-176. Recuperado de http://www.redalyc.org/articulo.oa?id=35602206

Villalobos, J; (2003). El docente y actividades de enseñanza / aprendizaje: algunas consideraciones teóricas y sugerencias prácticas. *Educere,* (7) 170-176. Recuperado de http://www.redalyc.org/articulo.oa?id=35602206

Weiss, E., Pasley, J., Smith, P., Banilower, E. y Heck, H. (2003). Looking inside the classroom: A study of K-12 mathematics and science education in the United States. Chapel Hill, N.C: Horizon Research.

Wenzelburger, E. (1990). ¿Cómo enseñar la matemática para mañana?. *Educación matemática,* Vol. 2. No. , pp. 48-51.

Wolfgang, B. Bracke, M. y Kreckler, J. (2015). Taxonomy of modelling tasks. Paper in process. University of Kaiserslautern, Alemania.

Anexo A

SOLICITUD DE PARTICIPACIÓN PARA DOCENTES DE 5º DE PRIMARIA

Estimado maestro (a):

De la manera más atenta solicitamos su apoyo para participar en el proceso de dos proyectos de investigación que se están llevando a cabo como parte del programa de Doctorado en Investigación Educativa del Departamento de Educación, en la Universidad Autónoma de Aguascalientes. En ambos casos el propósito es el diseño y validación de **instrumentos de observación** que permitan ofrecer información sobre algunos aspectos de la práctica docente en matemáticas, específicamente sobre: a) las *actividades de enseñanza* que se promueven en el aula y b) el tipo de *retroalimentación* que se ofrece a los estudiantes durante el desarrollo de estas actividades.

El propósito en ambos casos es desarrollar instrumentos que permitan (como fin último) ofrecer información a los docentes para apoyar la mejora de sus prácticas de enseñanza. No obstante, para la etapa que involucra el desarrollo y validación de estos instrumentos es necesario contar con información preliminar que permita realizar los ajustes pertinentes a los mismos. Por lo cual solicitamos su apoyo para videograbar dos clases de matemáticas (sesiones completas) con su grupo, con el único propósito de obtener los insumos necesarios para hacer dichos ajustes.

Algunas de las especificaciones que se considerarán para llevar a cabo las videograbaciones son:

- Contar con la participación de docentes que laboren en escuelas primarias públicas de *contexto urbano* de los municipios de: Jesús María, San Francisco de los Romo y/o Aguascalientes.
- Entregar un *formato de consentimiento voluntario informado* para los docentes participantes, en el cual se especificarán los propósitos del los proyectos de investigación y se incluirá una cláusula sobre el cuidado que se tendrá con la información recabada (videograbaciones). Esto es, la confidencialidad y no divulgación de los videos, atendiendo los principios éticos para el tratamiento de la información.

Una vez que se tenga una versión final de los instrumentos, se podrá ofrecer a los docentes participantes retroalimentación y asesoría con base en los elementos de los instrumentos desarrollados. Es decir, a partir de las videograbaciones de sus propias clases, los docentes podrán analizar los aspectos que se buscó identificar y valorar a través de los instrumentos. Para ello, el equipo de investigadores los apoyará en el uso de las rúbricas y otros elementos que se están considerando como parte de los instrumentos de observación. De esta manera, se espera poder retribuir en parte la aportación que nos ofrecen al participar en ambos estudios.

Además de lo antes mencionado, el equipo de investigadores se compromete a proporcionar una constancia de participación en los proyectos así como un juego de libros que puedan aportar más elementos para el desarrollo de su práctica docente (actividades de enseñanza y evaluación formativa).

Esperamos poder contar con su colaboración.

Para mayor información favor de contactar a las titulares de los proyectos de investigación:

M.I.E. Lesly Yahaira Rodríguez Martínez
yahaira-07@hotmail.com

M.I.E. Adriana Mercado Salas
chaams_ags@yahoo.com.mx

Universidad Autónoma de Aguascalientes

Departamento de Educación
9107400
Ext. 313

Anexo B

Contextualización de la práctica de enseñanza

Inicio de la clase_________ Final de la clase________

1. ¿Cuántos alumnos atiende en su grupo? ____
2. NEE (cuántos, qué tipo de NEE) ____
3. ¿Cuántos niños son los que regularmente participan (interesan) en la clase de matemáticas?________
4. ¿Cuánto tiempo tiene laborando en esta escuela primaria? (en caso que sea su primer año, comenzó a trabajar desde el inicio del ciclo escolar) ____
5. ¿Cuántos años tiene de servicio como profesor de educación primaria?_________________________
6. ¿Qué materiales toma como referencia para llevar a cabo la planificación de las clases?

   Libro de desafíos matemáticos ☐

   Libro de texto (diferente a DM) ☐

   Especifique:_____________________________________

   Guía de aprendizaje ¿cuál? ____

   Otro ____

Le voy a hacer un par de preguntas para contextualizar la clase que observamos el día de hoy:

7. ¿Con qué tema estuvo relacionada esta clase? ¿A qué bloque pertenece?

   ____

8. ¿Contestó las actividades o problemas antes de plantearlos al grupo?___________________________

   Sí, ☐ ¿Identifico una o varias formas de solución?____

   No ☐

9. ¿Qué propósito/aprendizaje esperado perseguía en esta clase?

   ____

10. ¿Considera que se cumplió el propósito de la clase? ¿Por qué?

    ____

Sólo en caso que haya iniciado un nuevo tema durante esta clase, podría indicar cuál fue y qué propósito persigue

11. ¿Qué nuevo tema introdujo durante la clase? ¿A qué bloque pertenece
12. ¿Qué propósito/aprendizaje esperado perseguía el nuevo tema?
13. ¿Considera que se cumplió el propósito del nuevo tema? ¿Por qué?

Anexo C

UNIVERSIDAD AUTONOMA DE AGUASCALIENTES

CARTA DE CONSENTIMIENTO VOLUNTARIO INFORMADO

Mediante la presente doy **mi consentimiento** para participar en los proyectos de investigación, coordinados por la M.I.E. Yahaira Rodríguez y M.I.E. Adriana Mercado, cuyos propósitos son el diseño y validación de **instrumentos de observación** enfocados en dos aspectos de la práctica docente: a) las *actividades de enseñanza* que se promueven en el aula y b) el tipo de *retroalimentación* que se ofrece a los estudiantes.

Entiendo que fui elegido (a) para participar en estos proyectos por ser docente de 5° de primaria de una escuela pública perteneciente al municipio de [Aguascalientes, Jesús María o San Francisco de los Romo]. Mi participación consistirá en permitir la videograbación de dos clases de matemáticas (sesiones completas) con mi grupo, entregar copia de la planeación de dichas clases y responder a un breve cuestionario que permita contextualizar mi trabajo.

Los resultados de esta investigación, a partir de mi participación y la de otros profesores, servirán para dar evidencia sobre algunos aspectos de validez y confiabilidad de los instrumentos desarrollados. Por lo tanto, dado que los estudios persiguen objetivos que se refieren a los instrumentos y no a la evaluación de mi práctica, **autorizo para que mi práctica sea filmada y audiograbada, y doy fe que estoy participando de manera voluntaria y que la información que se obtenga será tratada de manera anónima y confidencial**.

De igual manera, manifiesto haber recibido información sobre los proyectos e implicaciones de mi participación (y que puedo obtener más información en caso de ser necesario), y ser consciente de que puedo desistir de participar en dicho proceso, con previo aviso a las investigadoras coordinadoras. Finalmente, confirmo que el equipo de investigadores se compromete a otorgarme una copia de los instrumentos una vez finalizadas las investigaciones, como una herramienta que apoye la mejora de mi práctica docente, así como una copia del video de mis clases (si así lo requiero).

Aguascalientes, Ags., ___ de _____ de _____.

____

Nombre y firma del participante

____

Nombre y firma del Investigador 1. Nombre y firma del investigador 2.

Universidad Autónoma de Aguascalientes
Tel. 9107400
Ext. 313

Anexo D

UNIVERSIDAD AUTONOMA DE AGUASCALIENTES

CARTA DE DIMISIÓN DE PARTICIPACIÓN

Mediante la presente hacemos de su conocimiento la aceptación de la solicitud hecha por **Nombre del docente** para retirar su participación en los proyectos de investigación, coordinados por las estudiantes del Doctorado en Investigación Educativa de la UAA, Lesly Yahaira Rodríguez Martínez y Adriana Mercado Salas. Dichos proyectos de investigación tienen como propósitos el diseño y validación de **instrumentos de observación** enfocados en dos aspectos de la práctica docente: a) las *actividades de enseñanza* que se promueven en el aula y b) el tipo de *retroalimentación* que se ofrece a los estudiantes.

Entendemos que su decisión fue tomada por motivos personales. Ante esto le entregamos las videograbaciones realizadas sobre su práctica durante los días 17 y 18 de mayo de 2016. Confirmamos que estos archivos electrónicos no han sido copiados en ningún otro dispositivo .Dado que usted será la única propietaria de los videos, a partir del momento en que se firme esta carta, no nos hacemos responsables del uso que se haga de los archivos electrónicos.
Manifestamos nuestro agradecimiento por su tiempo y quedamos a sus órdenes para cualquier aclaración o duda.

Aguascalientes, Ags., 01 de junio de 2016.

M.I.E. Lesly Yahaira Rodríguez Martínez

yahaira-07@hotmail.com

M.I.E. Adriana Mercado Salas

chaams_ags@yahoo.com.mx

Nombre del docente

Nombre y firma del participante

Anexo E

COMPROMISO DE CONFIDENCIALIDAD EN EL MANEJO DE LA INFORMACIÓN

La calificación de clases videograbadas en diferentes escuelas primarias del estado de Aguascalientes implicará mantener contacto con la actividad de enseñanza de diferentes docentes que aceptaron participar en el proyecto “Protocolo de Observación de las Actividades de Enseñanza en Matemáticas”, bajo la condición de mantener su confidencialidad y utilizar la información para los propósitos de este proyecto de investigación. Por lo tanto, el observador participante deberá apegarse a este compromiso y mantener en total confidencialidad la identidad de las personas observadas, así como de la información que estos actores manejan durante las clases.

Señalado lo anterior y al firmar la presente yo ______________________________________ me comprometo a mantener la confidencialidad de los docentes participantes y hacer uso adecuado de la información. Lo anterior implicará apegarme a las indicaciones que la coordinadora del proyecto señale respecto a la manera de emplear las videograbaciones para determinar las calificaciones correspondientes.

____

Firma del observador

Fecha:

ORGANIZACIÓN DE TAREAS MATEMÁTICAS POR VIDEOGRABACIÓN DE CLASE

COMPROMISOS ESPECÍFICOS PARA EL TRATAMIENTO DE LA INFORMACIÓN:

- No descargar los archivos (videos) proporcionados en equipos de computo personales o diferentes a los autorizados por la coordinadora del proyecto (equipos resguardados por la Universidad Autónoma de Aguascalientes).
- No hacer copia total o parcial de la información de las videograbaciones con las que estarán trabajando.
- No comentar la información observada en las videograbaciones más allá de las reuniones con el equipo de trabajo.
- No publicar información sobre sus calificación u otros elementos observados en las videograbaciones en ningún medio o redes sociales.

No.	CODIGO DE VIDEO	TEMA	TAREA MATEMÁTICA	PERIODO DE TIEMPO	FIRMA DEL OBSERVADOR
1	01_SFR_GFL_01_070416_P	Equivalencias	1	V0 00:00 - V0 14:42	
			2	V0 14:43 - V3 18:04	
2	01_SFR_GFL_02_080416_P	Equivalencias/conversiones	1	V0 00:51 - V0 05:06	
			2	V0 05:06 - V0: 21:38	
			3	V0 21:40 - V3 09:00	
3	02_SFR_JCRL_01_110416_P	Gráficas	1	V0 00:00 - V0 10:04	
			2	V0 10:05 - V2 08:17	
			3	V2 08:18 - V3 21:44	
			4	V3 21:45 - V4 05:16	
4	02_SFR_JCRL_02_120416_P	Gráficas	1	V0 00:15 - V1 03:35	
			2	V1 03:37 - V3 13:48	
			3	V3 13:49 - V4 15:58	
5	04_JM_ASJ_01_120516_P	Reparto	1	V0 00:02 - V0 14:19	
			2	V0 14:20 – V2 16:00	
			3	V2 16:01 - V2 17:56	
6	04_JM_ASJ_02_190416_P	Sucesión /Progresión	1	V0 00:10 - V1 08:37	
			2	V1 08:38 - V1 17:19	
			3	V1 17:20 - V2 11:23	
			4	V2 11:24 -V4 10:33	
7	05_SFR_JAS_01_310516_P	Fracciones	1	V0 00:10 - V1 09:50	
			2	V1 09:50 - V2 20:15	
	05_SFR_JAS_02_310516_P	Reparto/ Recta numérica	1	V0 00:00 - V2 05:39	

8			2	V2 05:40 - V3 03:23	
9	07_AGS_CSTG_01_250516_P	Prismas y pirámides	1	V0 00:22 – V0 04:54	
			2	V0 04:55 – V0 15:20	
			3	V0 15:21 – V1 20:39	
10	07_AGS_CSTG_02_250516_P	Aristas, vértices y caras	1	V1 00:00 – V2 01:04	
			2	V2 01:05 – V3 20:10	
			3	V3 20:11 – V4 20:50	
11	08_AGS_CSTG_01_240516_P	Multiplicación de números decimales por números naturales	1	V0 01:27 - V0 13:27	
			2	V0 13:28 - V2 18:59	
12	08_AGS_CSTG_02_240516_P.	Circunferencia, Círculo, radio y diámetro.	1	V0 00:00 - V0 16:55	
			2	V0 17:18 - V0 20:20	
			3	V0 20:21 – V2 1:04	
13	09_AGS_SJLEM_01_190516_P	Problemas de reparto	1	V0 00:00 – V0 12:21	
			2	V0 12:22 – V1 20:24	
			3	V120:25 – V3 00 :24	
			4	V3 00:25 – V3 07:04	
14	09_AGS_SJLEM_02_190516_P	Números decimales	1	V0 00:57 – V0 19:44	
			2	V0 19:45 – V3 07:56	
15	11_AGS_MVR_01_260516_P	Sucesiones numéricas	1	V0 00:00 - V0 10:26	
			2	V1 00:08 - V2 01:26	
			3	V2 01:27 - V3 02:15	
16	11_AGS_MVR_02_260516_P	Multiplicación con punto decimal	1	V0 00:20 – V0 12:11	
			2	V0 12:12 - V1 18:09	
17	13_AGS_PRA_01_240516_P	Reparto	1	V0 00:00 - V1 01:17	
			2	V1 01:18 - V2 08:55	
			3	V2 08:56 - V4 00:08	
18	13_AGS_PRA_02_240516_P	Reparto	1	V0 00:04 – V0 13:54	
			2	V0 13:55 – V1 18:13	
			3	V1 18:14 – V2 10:17	
			4	V2 10:18 – V4 09:56	
19	14_AGS_NGG_01_1020616	Fracciones	1	V1 01:41 – V2 14:08	
			2	V2 14: 09 – V3 9:39	

			3	V4 00:27 – V4 20:50	
20	14_AGS_NGG_02_1020616_P **(Una tarea 5 se menciona pero no se lleva a cabo)**	Series numéricas	1	V1 00:47 - V1 08:35	
			2	V0 09:10 – V1 04:55	
			3	V1 04:57 – V2 13:00	
			4	V2 13:01 – V4 08:10	
21	17_AGS_MJLL_01_310516_P	La circunferencia	1	V1 00:16 – V2 02:20	
			2	V2 02:21 – V4 11:00	
22	17_AGS_MJLL_02_310516_P	Radio y diámetro	1	V1 00:00 – V1 09:23	
			2	V1 09:24 – V2 19:04	
			3	V2 19:05 – V3 00: 53	
			4	V3 00:54 – V3 23:35	
23	19_AGS_TICJ_01_240516_P	Fracciones	1	V1 01:14 – V1 12:05	
			2	V1 16:39 – V2 12:34	
			3	V2 12:35 – V2 20:40	
24	19_AGS_TICJ_02_240516_P	Fracciones	1	V0 03:00 – V0 17:38	
			2	V0 17:40 – V2 06:41	
25	20_AGS_ADBA_01_170516_P	Fracciones	1	V0 01:12 - V0 19:52	
			2	V0 21:05 - V2 20:15	
26	20_AGS_ADBA_02_170516_P	Series numéricas	1	V2 06:28 – V3 00:50	
			2	V3 00:51 – V5 14:18	
27	21_AGS_MLCA_01_250516_P	Círculo y circunferencia	1	V0 01:36 - V1 10:15	
			2	V1 10:16 - V1 17:50	
			3	V1 17:51 - V3 19:48	
			4	V4 03:19 – V6 09:58	
28	21_AGS_MLCA_02_250516_P	Círculo y circunferencia.	1	V0 01:56 - V1 02:35	
			3	V1 02:36 - V2 21:20	
29	24_AGS_BRH_01_180516_P	Numeración maya	1	V0 01:00 - V0 20:07	
			2	V0 20:08 - V2 00:30	
			3	V2 00:31 - V2 11:20	
30	24_AGS_BRH_02_180516_P	Relaciones y equivalencias de fracciones	1	V0 00:14 - V0 08:33	
			2	V0 08:34 - V0 19:43	
			3	V0 19:44- V1 13:28	

31	25_AGS_ACCS_01_200516_P	Cuerpos geométricos	1	V0 02:23 - V0 10:10	
			2	V0 10:11 - V4 00:09	
32	25_AGS_ACCS_02_200516_P	Cuerpos geométricos	1	V0 00:00 - V0 18:22	
			2	V0 18:23- V1 03:21	
			3	V1 03:25 - V1 18:34	
33	26_AGS_MEPM_01_200516_P	Escala gráfica	1	V0 01:11 - V0 15:04	
			2	V0 15:05 - V3 07:39	
			3	V3 07:40 – V3 10:47	
34	26_AGS_MEPM_02_200516_P	Escala gráfica	1	V1 00:00 - V1 13:44	
			2	V1 13:45 - V4 03:54	
35	27_AGS_SJAG_01_230516_P	Reparto	1	V0 00:12 - V1 12:58	
			2	V1 13:00 - V2 18:00	
			3	V2 20:39 - V5 03:15	
			4	V5 03:20 - V7 11:22	
36	27_AGS_SJAG_02_230516_P	Reparto	1	V0 00:24 - V1 20:44	
			2	V1 20:46 - V2 06:53	
37	28_AGS_RESR_01_230516_P	División con fracciones	1	V0 00:42 - V0 11:33	
			2	V0 12:45 - V2 09:24	
			3	V2 09:25 – V4 03:52	
			4	V6 00:57 –V6 19:43	
38	28_AGS_RESR_02_230516_P	División con fracciones	1	V0 05:25 - V1 06:52	
			2	V1 06:53 - V2 14:43	
39	29_AGS_GERR_01_300516_P	Números decimales	1	V13 01:01 - V14 02:37	
			2	V14 2:38 - V16 00: 06	
40	29_AGS_GERR_02_300516_P	Números decimales	1	V0 00:15 - V1 00:32	
			2	V1 00:33 – V2 18:43	

Anexo F

FORMATO PARA SEGMENTAR VIDEOGRABACIONES

Instrucciones: Utilice este formato para identificar y describir las diferentes actividades de enseñanza que el docente propone a los estudiantes durante la clase de matemáticas. Para organizar la información enumere las actividades que el docente señala a lo largo de la clase y complemente dicha información describiendo lo que los estudiantes realizan con base en dichas actividades. Asegúrese de incluir los datos que se solicitan en el concentrado de información (Formato 1).

Notas:

1. El observador sólo debe describir las actividades que suceden durante la clase, por lo cual no es necesario establecer una clasificación de las actividades de enseñanza observadas.
2. Para establecer los tiempos en cada actividad de enseñanza será necesario señalar en primer lugar la letra “V” (Video), seguido del número del video observado correspondiente al código de videograbación que se describirá; después anotar el número de minuto y segundo de inicio y final. Por ejemplo: Actividad 1 (V1 15:21 – V2 21:56).

Confidencialidad en el manejo de la información

Considerando que la realización de este trabajo implica entrar en contacto con la actividad de enseñanza de diferentes docentes que aceptaron ser videograbados durante el desarrollo de dos de sus clases de matemáticas; bajo la condición de mantener su confidencialidad y utilizar la información para los propósitos de este proyecto de investigación. Se solicita a los observadores participantes apegarse a este compromiso y mantener en total confidencialidad la identidad de las personas observadas, así como las información que estos actores manejan durante la realización de las clases.

Por lo tanto, se le solicita de la manera más atenta firmar esta carta compromiso con la que manifiestan estar de acuerdo con lo antes señalado. Es decir, hacer un uso apropiado de la información, así como mantener la confidencialidad de los sujetos que aparecen en las videograbaciones correspondientes.

____

Nombre completo y firma del observador

CONCENTRADO DE INFORMACIÓN DE ACTIVIDADES DE ENSEÑANZA								
ETIQUETA DE VIDEOGRAVACIÓN: 07_AGS_CSTG_01_250516_P INICIALES DEL OBSERVADOR: LYRM								
NO. DE ACTIVIDAD	**INICIO (minutos: segundo)**	**FINAL (minutos: segundo)**	**TEMA O CONTENIDO**	**INDICACIÓN AL INICIO DE LA ACT.**	**INDICACIÓN AL FINAL DE LA ACT.**	**ACTIVIDAD DE ENSEÑANZA** (Qué hace o propone el docente)	**TRABAJO DE LOS ESTUDIANTES** (Qué hacen y cómo se organizan para trabajar)	**OBSERVACIONES** Señale aspectos particulares que apoyen la descripción de la actividad e indique (si es el caso) el material empleado durante su realización.
ACTIVIDAD 1								
ACTIVIDAD 2								
ACTIVIDAD 3								
ACTIVIDAD 4								
ACTIVIDAD 5								
ACTIVIDAD 6								
ACTIVIDAD 7								

Anexo G

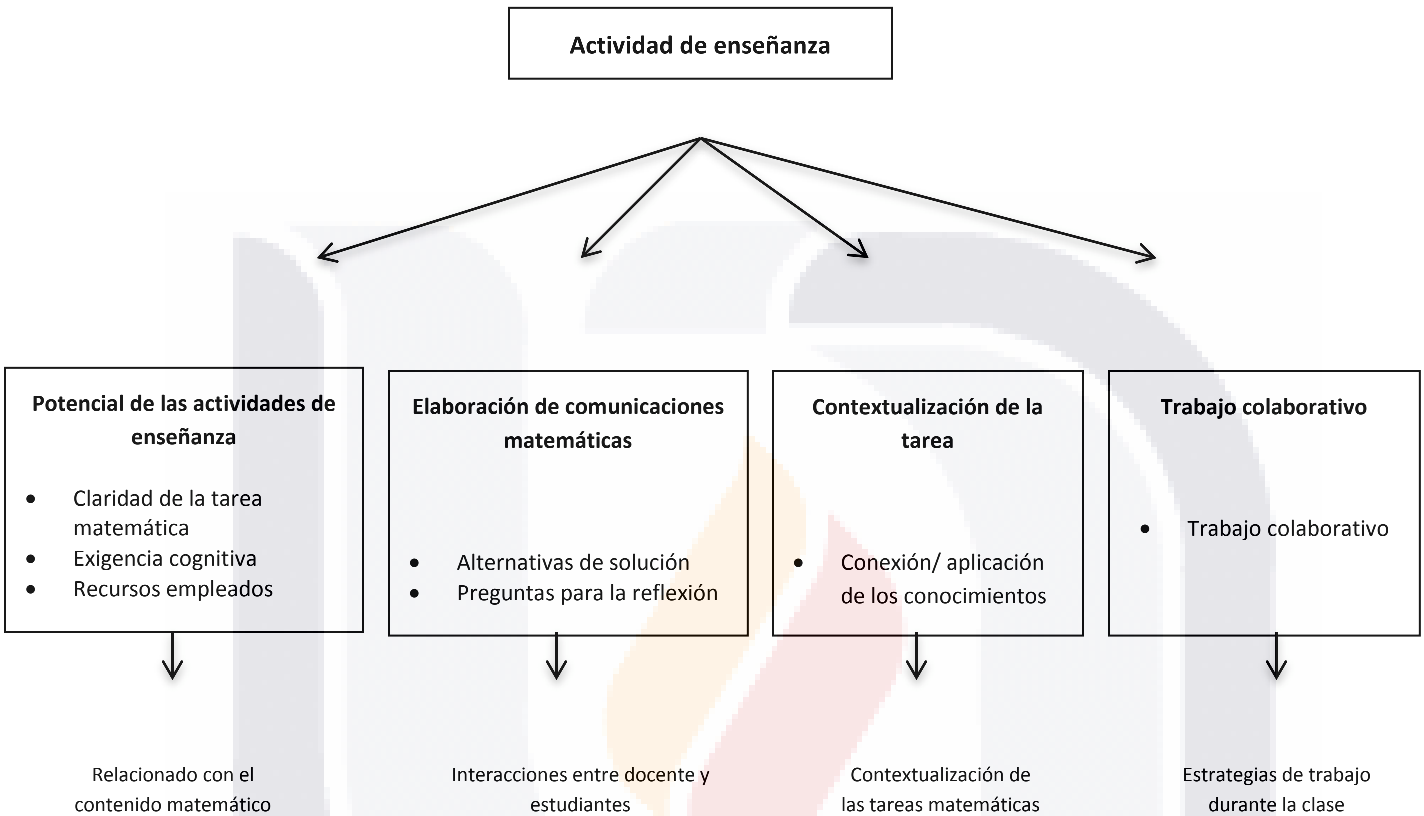
PROTOCOLO DE OBSERVACIÓN DE LAS ACTIVIDADES DE ENSEÑANZA EN MATEMÁTICAS

Introducción

El *Protocolo de Observación de las Actividades de Enseñanza en Matemáticas* (POAEM) tiene como propósito valorar lo que hace el docente y los estudiantes durante la clase a partir de las tareas matemáticas. Se prestará especial atención a los productos que se espera los estudiantes logren, los recursos disponibles para desarrollar dichos productos, y las operaciones que involucran ese proceso (Stein, Grover y Henningsen, 1996; Boston y Wolf, 2006). Para ello, este protocolo de observación evaluará los siguientes dominios:

- **a.** E*l potencial de las actividades de enseñanza,* relacionado con algunos aspectos cognitivos.
- **b.** L*a elaboración de comunicaciones matemáticas,* que tiene que ver con las interacciones que se promueven entre los estudiantes para establecer diferentes estrategias de solución y preguntas para favorecer la reflexión.
- **c.** L*a contextualización de la tarea*, que se refiere a la conexión o aplicación de los conocimientos matemáticos a partir de la tarea matemática propuesta.
- **d.** El *trabajo colaborativo*, entendiendo éste como la interacción, negociación y cooperación entre los estudiantes para lograr un propósito común (ver figura 1).

Figura 1. Dominios de las actividades de enseñanza



Cada uno de los dominios antes señalados incluyen diferentes dimensiones que constituyen los ejes de la observación. La valoración de las actividades de enseñanza se hace a través de rúbricas que han sido diseñadas para evaluar cada una de las dimensiones. A continuación se describen las dimensiones que integran cada dominio, su importancia para valorar las actividades de enseñanza y los criterios que incluyen las rúbricas correspondientes.

a) Potencial de la actividad de enseñanza (tres dimensiones):

1. **Claridad de la tarea matemática.** Se refiere a la manera en que el docente comunica el trabajo que los estudiantes deben desarrollar y cómo indaga si los estudiantes han comprendido lo que se les solicita como trabajo o producto.

Stein, Grover y Henningsen (1996), Boston y Wolf (2006) y Boston (2012) señalan que a partir de las tareas matemáticas los estudiantes son expuestos a información que los compromete con ciertos conceptos o ideas matemáticas y por lo tanto es fundamental que comprendan lo que deben realizar para que tengan mayor posibilidad de lograr los aprendizajes esperados.

Para evaluar esta dimensión, se utilizará una rúbrica de cuatro niveles de desempeño. La rúbrica permite identificar la calidad de tres indicadores: a) los productos o aprendizajes esperados, b) el lenguaje matemático, y c) las preguntas o estrategias para favorecer la comprensión de los estudiantes sobre la tarea matemática propuesta (Stein, Grover y Henningsen, 1996; Hiebert, 2003; Boston y Wolf 2006; y NCTM, 2015).

2. **Exigencia cognitiva.** Se refiere al tipo y nivel de pensamiento que requieren los estudiantes para resolver ciertas tareas matemáticas.

   De acuerdo con el NCTM “una tarea matemática puede abarcar desde un conjunto de ejercicios rutinarios hasta un problema complejo y desafiante que enfoque la atención de los estudiantes sobre una idea matemática particular” (2015, p. 19). Sin embargo, no todas las tareas ofrecen las mismas oportunidades de aprendizaje, por lo que es importante distinguir el tipo de habilidades cognitivas que éstas requiere por parte de los estudiantes [10] . Por ejemplo, reproducir de memoria hechos, reglas o fórmulas representan una baja exigencia cognitiva, mientras que explorar y establecer relaciones de conceptos matemáticos implican una exigencia mayor. Es decir, mientras más habilidades cognitivas debe poner en juego el estudiante, la tarea requerirá una mayor exigencia cognitiva.

   Para evaluar la dimensión de exigencia cognitiva, se utilizará una rúbrica con indicadores que permiten distinguir las tareas que favorecen un aprendizaje superficial (conocimientos aislados o de repetición sin sentido para los estudiantes) de aquellas que ayudan a promover un aprendizaje profundo (por ejemplo, analizar, explicar o justificar procesos).[11]

3. **Recursos empleados durante la actividad de enseñanza.** Se refiere al uso que el docente y los estudiantes hacen de materiales disponibles para promover o fortalecer los aprendizajes matemáticos.

   En esta dimensión se buscará identificar la disponibilidad de los recursos y su funcionalidad. Se incluye la funcionalidad debido a que se considera que en los procesos de aprendizaje no sólo influye el acceso a los recursos sino el uso que se hace de éstos, de acuerdo con el propósito de la tarea matemática (Gueudet y Trouche, 2008; Cohen, Raudenbush y Ball, 2003, p.122).

---

[10] En caso que una tarea matemática incluya más de un ejercicio y estos impliquen diferentes niveles de exigencia cognitiva, se deberá calificar y reportar la de mayor nivel.
[11] Se utilizará la taxonomía de tareas matemáticas desarrollada por Stein, Grover y Henningsen (1996).

Esta dimensión se evalúa con una rúbrica que permite valorar los siguientes indicadores: a) cómo se involucra el recurso con el contenido que se está trabajando; y b) la movilización de incentivos[12] que se promueven para favorecer el desempeño de los estudiantes.

b) Elaboración de comunicaciones matemáticas (dos dimensiones)

4. **Alternativas de solución.** Se refiere a la flexibilidad del docente para promover y aceptar una variedad de estrategias o procedimientos para resolver un problema matemático. Específicamente, las oportunidades que el docente ofrece a los estudiantes para emplear múltiples posibilidades de solución, sin sesgar el proceso de pensamiento de los estudiantes.

   El NCTM (2015) señala que para emplear múltiples posibilidades de solución, se requieren problemas poco estructurados que faciliten diversas formas de abordarlos y permitan a los estudiantes emplear diferentes estrategias de solución. Con base en ello, la rúbrica para valorar esta dimensión se ha centrado en dos indicadores: a) las oportunidades que se dan a los estudiantes para emplear diferentes estrategias o procedimientos de solución, y b) las pistas o sugerencias que podrían sesgar las respuestas de los estudiantes.

5. **Preguntas para la reflexión.** Se refiere al tipo de preguntas que se platean durante la clase (ya sea por el docente o los estudiantes) para dar sentido a ideas o relaciones matemáticas importantes.

   De acuerdo con diferentes autores (Walsh y Sattes, 2005; Jonassen, 2011; y NCTM, 2015) una enseñanza eficaz de las matemáticas debe promover preguntas deliberadas que apoyen la reflexión de los estudiantes sobre el pensamiento matemático. Además, señalan que el tipo de preguntas y el modelo de cuestionamiento son dos aspectos importantes para que los estudiantes logren dar un sentido a las matemáticas, así como favorecer mayores progresos en el razonamiento.

   Las preguntas que se generan durante una clase de matemáticas representan oportunidades para que los estudiantes justifiquen sus respuestas, hagan distinciones, generalicen ideas, y utilicen el diálogo para construir ideas coherentes que promuevan la mejora de un entendimiento colectivo (Newmann y Wehlage, 1993).

---

[12] Los incentivos, se refieren a estimular el trabajo de los estudiantes mediante el uso del conocimiento y las habilidades en cierto tipo de tareas académicas.

Para evaluar esta dismensión la rúbrica se centra en tres indicadores: a) las preguntas que solicitan a los estudiantes describir información, b) preguntas que ayudan a hacer evidente el pensamiento matemático (justificación), y c) preguntas que implican justificar y demostrar el trabajo realizado (NCTM, 2015).

c) Contextualización de la tarea (una dimensión)

6. **Conexión/ aplicación de los conocimientos matemáticos.** Se refiere a la medida en que el contexto de una *situación problemática* ofrece a los estudiantes la oportunidad de conectar los aprendizajes matemáticos con su propia experiencia, con otras disciplinas o con situaciones de la vida cotidiana como una forma de reconocer los aprendizajes matemáticos por su relevancia y utilidad. Cabe aclarar que no todas las tareas matemáticas plantean situaciones problemáticas, por lo cual en esta dimensión se excluirán aquellos problemas o tareas de mecanización que no comprendan una contextualización.

Autores como Boston y Wolf (2006), Picaroni y Loureiro (2010) y Zolkower y Bressan (2012), señalan que un aspecto importante para fortalecer la enseñanza de las matemáticas es que los estudiantes puedan reconocer en su aprendizaje un significado personal o social, así como la aplicación de ese aprendizaje en diferentes contextos mediante tareas matemáticas que resulten realizables, razonables o imaginables para los estudiantes. Al respecto, el NCTM señala que para lograr dicho propósito los docentes deben promover “condiciones que ayuden a desarrollar tareas que promuevan experiencias significativas y un sentido de identidad que motive el compromiso con el aprendizaje de las matemáticas” (2015, p. 18).

La rúbrica con la que se valorará esta dimensión considera tres indicadores: a) experiencias que los estudiantes puedan reconocer por su situación o contexto, b) conexión de las situaciones problemáticas con aspectos del mundo que nos rodea, y c) relación de las ideas o conceptos matemáticos con otros temas o disciplinas.[13]

d) Trabajo colaborativo (una dimensión)

7. **Trabajo colaborativo.** Se refiere a la interacción, negociación y cooperación entre estudiantes para lograr un propósito común. Un aspecto importante de esta dimensión es la interacción que se genera durante una tarea matemática, más que la manera en que se lleva a cabo la organización de los estudiantes.

Autores como Fernández y Yoshida (2004), Trigueros (2009), SEP (2011a), Barreiro y Casetta, (2012), NCTM (2015) señalan que un elemento que puede favorecer el trabajo colaborativo en el aula, es que los docentes adopten un rol activo para promover la interacción entre los estudiantes mediante diferentes formas de trabajo (pares, pequeños grupos, todo el grupo), que motive la colaboración y el esfuerzo productivo para abordar ideas o procedimientos matemáticos.

[13] **Esta clasificación se basa en los planteamientos de en Picaroni y Loureiro (2010).**

Para evaluar esta dimensión se utilizará una rúbrica para identificar tres indicadores: a) el rol activo del docente para promover la interacción de los estudiantes y lograr un aprendizaje o producto común, b) el tipo de cooperación entre los estudiantes (integración de ideas y propuestas de los participantes), y c) la negociación que para llegar a acuerdos comunes sobre ideas, conceptos o procedimientos matemáticos (Fernández y Yoshida, 2004, Verschaffel, Greer y De Corte, 2010, Barreiro y Casetta, 2012; y NCTM, 2015).

PROCESO DE OBSERVACIÓN Y CALIFICACIÓN DE LAS VIDEOGRABACIONES

Para realizar la observación y calificación de las videograbaciones será necesario que los observadores participantes tomen previamente el curso de capacitación para el uso de este instrumento, que consistirá en entrenamiento con el acompañamiento de un experto, así como de sesiones de reforzamiento individual a partir de ejercicios de calificación con videograbaciones de prueba.

La unidad de análisis que se empleará para el uso de este instrumento serán las tareas matemáticas que conforman las actividades de enseñanza. Por lo tanto, a partir de las tareas matemáticas identificadas previamente los observadores deberán seguir el siguiente proceso:

1. Registrar los datos generales del video a observar (fecha, código de video, código observador) en hojas de calificación correspondientes a cada una de las tareas matemáticas que lo conforman.
2. Observar la videograbación de acuerdo con los tiempos señalados para cada una de las tareas matemáticas.
3. Registrar en la sección de observaciones información relacionada con los criterios de las rúbricas (y el nivel seleccionado) para justificar puntuaciones otorgadas, así como información adicional que el observador considere necesaria para justificar su valoración.
4. Registrar las puntuaciones por dimensión y tarea matemática para la videograbación correspondiente en hoja electrónica.

POTENCIAL DE LA ACTIVIDAD DE ENSEÑANZA			
I. CLARIDAD DE LA TAREA MATEMÁTICA. Se refiere a la manera en que el docente comunica el trabajo que los estudiantes deben desarrollar y cómo indaga si los estudiantes han comprendido lo que se les solicita como trabajo o producto. **NOTA:** a. En este caso se busca identificar la manera en que el docente busca dejar en claro el trabajo que los estudiantes deben realizar, no necesariamente que explicite el objetivo o aprendizaje esperado. b. Evite calificar como tareas matemáticas aquellas que No estén relacionadas con contenidos matemáticos.			
BAJO 1, 2	MEDIO 3, 4	MEDIO-ALTO 5, 6	ALTO 7, 8
El docente no enfatiza conceptos o ideas matemáticas importantes para el desarrollo de la tarea. Si varios estudiantes preguntan qué deben hacer, el docente pide que vuelvan a leer las instrucciones o él mismo repite las instrucciones de trabajo mencionadas previamente.	El docente plantea la tarea matemática y señala a los estudiantes los productos que espera logren a partir de ésta. Enfatiza conceptos o ideas matemáticas importantes para el desarrollo de la misma.	El docente plantea la tarea matemática y señala a los estudiantes los productos que espera logren. Utiliza un lenguaje matemático para enfatizar algunos conceptos o ideas importantes y reafirma las indicaciones dadas con otras palabras o con ejemplos. Sin embargo, no verifica la comprensión de los estudiantes a través de preguntas o estrategias que los ayuden a clarificar lo que deben realizar a partir de la tarea propuesta.	El docente plantea la tarea matemática y señala a los estudiantes los productos que espera logren. Utiliza un lenguaje matemático para enfatizar algunos conceptos o ideas importantes y verifica la comprensión de los estudiantes a partir de preguntas relacionadas con el objetivo o producto esperado (qué deben hacer), o a través de estrategias para que los estudiantes clarifiquen lo que deben lograr (lluvia de ideas, que expliquen con sus palabras, que comenten la tarea entre pares, entre otras).

POTENCIAL DE LA ACTIVIDAD DE ENSEÑANZA

II. **EXIGENCIA COGNITIVA**. Se refiere al tipo y nivel de pensamiento requerido para realizar ciertas tareas matemáticas. Por ejemplo reproducir de memoria hechos, reglas o fórmulas son aprendizajes de baja exigencia cognitiva, mientras que explorar y establecer relaciones de conceptos matemáticos representa una mayor exigencia. Es decir, mientras más habilidades cognitivas deba poner en juego el estudiante, la tarea implicará una mayor exigencia cognitiva.

NOTAS:

a. La exigencia cognitiva es distinta a la dificultad de una tarea matemática. La *dificultad* que los estudiantes pueden tener para llevar a cabo la tarea matemática se refiere a las circunstancias que los estudiantes enfrentan ante cierto problema matemático. Esto es, una tarea con baja exigencia cognitiva puede resultar poco o muy difícil para un estudiante de acuerdo con su situación particular. Por ejemplo, si pedimos a un niño de quinto grado de primaria que defina que es un “Axioma” probablemente le resulte una tarea difícil aunque se trate de un nivel de exigencia bajo, ya que es una tarea que implica conocer o memorizar un concepto.
b. Una tarea matemática puede incluir varios ejercicios con diferentes niveles de exigencia cognitiva, si es el caso califique aquella que implique mayor exigencia para los estudiantes.

BAJO 1, 2	MEDIO 3, 4	MEDIO-ALTO 5, 6	ALTO 7, 8
La tarea se limita a promover la reproducción de hechos, reglas, fórmulas o definiciones previamente aprendidas. Implican la reproducción exacta de conceptos o procedimientos (mecanización).	La tarea es algorítmica, requiere de procedimientos específicos para trabajar alguna idea o concepto matemático, pero sin promover un significado de ideas o conceptos matemáticos. En este caso, los estudiantes deben identificar, reconocer o asociar algún procedimiento determinado.	La tarea propuesta sugiere procedimientos generales que se relacionan con ideas o conceptos matemáticos sobre los que debe reflexionar. Los estudiantes requieren analizar, comparar, clasificar, o evaluar información para lograr una comprensión de la situación propuesta.	La tarea requiere un pensamiento complejo (no algorítmico), en el que los estudiantes deben integrar elementos para resolver el problema matemático. Los estudiantes deben explorar, diseñar, representar, demostrar o generalizar conceptos, algoritmos o procedimientos para establecer soluciones y lograr un entendimiento de la situación propuesta.

POTENCIAL DE LA ACTIVIDAD DE ENSEÑANZA

III. RECURSOS EMPLEADOS DURANTE LA ACTIVIDAD DE ENSEÑANZA. Se refiere al uso que el docente y los estudiantes hacen de los recursos disponibles para promover o fortalecer los aprendizajes a partir de la tarea matemática propuesta.

NOTAS:

a. Con esta rúbrica no se espera identificar el tipo de materiales que se emplean durante una clase de matemáticas, sino cómo los docentes y estudiantes utilizan los recursos disponibles en beneficio de los aprendizajes matemáticos.

b. No es necesario que el docente utilice en todo momento el recurso seleccionado para desarrollar la actividad de enseñanza. Por ejemplo, el docente puede sólo entregar a los estudiantes algún artefacto y dar instrucciones para que realicen la tarea matemática.

BAJO 1, 2	MEDIO 3, 4	MEDIO-ALTO 5, 6	ALTO 7, 8
El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros. Sin embargo, estos no están alineados con un propósito de aprendizaje matemático. Es decir, se utilizan como medio para realizar actividades recreativas o complementarias a la tarea matemática.	El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros. Éstos son coherentes con el propósito de la tarea propuesta, sin embargo sólo se emplean como medio de información o reforzamiento para el propósito matemático. No se toman como base para promover un análisis, discusión evaluación o comparación de ideas o procedimientos.	El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros, para promover el aprendizaje de las matemáticas. Éstos son coherentes con el propósito de la tarea propuesta, ya que se toman como base para llevar a cabo el análisis, discusión, evaluación o comparación de las ideas, procedimientos o estrategias de solución empleadas para realizar la tarea matemática. Sin embargo, no se utilizan para llegar a una comprobación de las estrategias o procedimientos empleados.	El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros, para promover el aprendizaje de las matemáticas. Éstos son coherentes con el propósito de la tarea propuesta, ya que se toman como base para llevar a cabo el análisis, discusión, evaluación o comparación de las ideas, procedimientos o estrategias de solución empleadas para realizar la tarea matemática. Además, permiten a los estudiantes comprobar sus propios procedimientos o estrategias de solución y autocorregir los procedimientos empleados para realizar la tarea matemática.

ELABORACIÓN DE COMUNICACIONES MATEMÁTICAS			
IV. ALTERNATIVAS DE SOLUCIÓN. Se refiere a la **flexibilidad del docente para promover y aceptar** una variedad de estrategias o procedimientos para resolver un problema matemático. **NOTAS:** a. A diferencia de la dimensión "demanda cognitiva" que enfatiza el nivel de pensamiento requerido para solucionar ciertas situaciones problemáticas, el foco de atención en esta dimensión serán las oportunidades que el docente ofrece a los estudiantes para emplear múltiples posibilidades de solución, sin sesgar el proceso de pensamiento de los estudiantes. b. Durante la realización de las tareas matemáticas, el docente resolverá algunas dudas de los estudiantes para que realicen el trabajo solicitado. Es importante prestar atención si las respuestas del docente ayudan a los estudiantes para continuar con su propio proceso o si les proporciona las respuestas específicas, limitando así el desempeño de los estudiantes.			
BAJO 1, 2	MEDIO 3, 4	MEDIO-ALTO 5, 6	ALTO 7, 8
Los estudiantes no tienen oportunidad de aplicar múltiples estrategias de solución, puesto que se les solicita seguir (o reproducir) un procedimiento, fórmula o método de solución específico (explicado previamente por el docente o señalado en la tarea matemática propuesta).	El docente solicita a los estudiantes que apliquen algún procedimiento específico. Consulta si el problema se podría resolver de otra manera y escucha las respuestas de los estudiantes. Sin embargo, al final se enfatiza o promueve un procedimiento específico.	El docente permite que los estudiantes indaguen sus propios procedimientos. Sin embargo, éstos no se retoman para compartir los diferentes procedimientos empleados. Sólo se comprueba que se haya llegado al resultado (generalmente al final de la tarea).	El docente permite que los estudiantes indaguen y propongan sus propios procedimientos. Además, solicita a los estudiantes que compartan y comparen los diferentes procedimientos empleados para comprobar que se pueden tomar varios caminos para llegar a un resultado (durante el desarrollo o al final de la tarea).

ELABORACIÓN DE COMUNICACIONES MATEMÁTICAS			
V. PREGUNTAS PARA LA REFLEXIÓN. Se refiere al tipo de preguntas que se platean durante la realización de una tarea matemática (ya sea por el docente o los estudiantes) para dar sentido a ideas o relaciones matemáticas importantes. **NOTAS** **a.** Para determinar el nivel de desempeño en esta rúbrica, es importante prestar especial atención al tipo de preguntas que se generan a partir de la tarea matemática y qué tipo de pensamiento favorece: justificación y comprobación de respuestas, o la explicación de un procedimiento. **b.** Aunque diferentes tipo de preguntas resultan importantes en la enseñanza de las matemáticas de acuerdo con su propósito, en este caso se busca identificar aquellas que favorezcan la reflexión de los estudiantes. **c.** Las preguntas se pueden plantear durante el desarrollo de una tarea matemática o al final de la misma. Es decir, el momento en que se planteen las preguntas no es relevante para los propósito de esta dimensión, sino determinar si se realizan y qué tipo de preguntas son las que se identifican.			
BAJO 1, 2	MEDIO 3, 4	MEDIO-ALTO 5, 6	ALTO 7, 8
No hay preguntas que permitan a los estudiantes explicar, justificar o demostrar una estrategia o procedimiento de solución. El docente da indicaciones o guía a los estudiantes para aplicar algún procedimiento, lo revisa y señala si es correcto o incorrecto.	El docente plantea preguntas que llevan a los estudiantes a describir los pasos de las estrategias o procedimientos de solución que emplearon para resolver algún problema matemático. No se solicita una justificación o demostración de la respuesta obtenida.	El docente plantea preguntas a través de las cuales solicita a los estudiantes que expliquen y justifiquen las estrategias o procedimientos empleados para resolver algún problema matemático. El docente hace uso de sus propios argumentos y demostraciones para reforzar alguna idea o procedimiento matemático específico.	El docente promueve la comprensión de ideas o procedimientos matemáticos mediante preguntas que requieren que los estudiantes justifiquen y demuestren sus respuestas. Las preguntas son planteadas por el docente o por los propios estudiantes (se les anima para comentar y comprobar sus respuestas). El docente evita dar sus propios argumentos sobre la manera de resolver el problema matemático y retoma (en algunos casos) el razonamiento de los estudiantes para plantear nuevas preguntas.

TIPO DE SITUACIONES MATEMÁTICAS

VI. **CONEXIÓN/ APLICACIÓN DE LOS CONOCIMIENTOS MATEMÁTICOS.** Se refiere a la medida en que el contexto de una *situación problemática* ofrece a los estudiantes la oportunidad de conectar los aprendizajes matemáticos con su propia experiencia, con otras disciplinas o con situaciones de la vida cotidiana, como una manera de reconocer los aprendizajes matemáticos por su relevancia y utilidad.

NOTA:

a. El foco de atención será identificar si las **situaciones problemáticas** propuestas por los docentes tienden o no a vincular los conocimientos o procedimientos matemáticos con su aplicación en diferentes circunstancias o contextos. Aunque se reconoce que no en todos los casos es posible establecer *situaciones problemáticas* para los contenidos matemáticos, en este instrumento se buscará identificar el tipo de contexto que se emplea en las tareas matemáticas que apliquen.

b. En los casos en que una tarea matemática sugiera ejercicios de mecanización o problemas matemáticos sin contexto se deberá macar la opción "No aplica" y señalar en las sección de *observaciones* qué tipo de problemas se identificaron durante la tarea revisada.

BAJO 1, 2	MEDIO 3, 4	MEDIO-ALTO 5, 6	ALTO 7, 8
La tarea matemática propuesta a los estudiantes hace referencia a situaciones que no resultan prácticas o aplicables a las experiencias de los estudiantes o situaciones cotidianas. No se establece una conexión clara entre las matemáticas que están aprendiendo (ideas o procedimientos) y su aplicación o relevancia en diferentes contextos o situaciones.	La tarea matemática propuesta a los estudiantes se centra en una situación problemática en la que se incluye información relacionada con el ámbito académico o el contexto escolar. Sin embargo, la tarea establece una problematización en la que se enfatizan sólo cuestiones relacionadas con la disciplina.	La tarea matemática propuesta a los estudiantes se centra en una situación problemática en la que se incluye información sobre diferentes temas o disciplinas. La tarea establece una problematización en la que es posible vislumbrar la relación de los conocimientos matemáticos para fortalecer otras disciplinas o campos de trabajo.	La tarea matemática propuesta a los estudiantes se centra en una situación problemática que incluye información o datos que los estudiantes pueden reconocer por su experiencia o por situaciones de la vida cotidiana, así como por una conexión explicita entre diferentes temas o disciplinas. Los estudiantes pueden establecer un vínculo claro entre las matemáticas que están aprendiendo y su aplicación o relevancia en diferentes contextos o situaciones.

TRABAJO COLABORATIVO

VII. TRABAJO COLABORATIVO. Se refiere a la interacción, negociación y cooperación entre estudiantes para lograr un propósito común.

NOTAS:

a. Es importante prestar atención a la estrategia de trabajo que se establece entre el docente y los estudiantes (y entre estudiantes y estudiantes). Es decir la interacción que se genera durante una tarea matemática, más que sólo identificar si se promueve el trabajo en equipo (binas, pequeños grupos, todo el grupo).
b. Es importante observar el tipo de interacción, negociación y/o cooperación que realizan el docente y los estudiantes (y estudiantes y estudiantes) en los diferentes momentos de la tarea matemática.

BAJO 1, 2	MEDIO 3, 4	MEDIO-ALTO 5, 6	ALTO 7, 8
El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Sin embargo, su interacción con los estudiantes se limita a verificar que estén desarrollando el trabajo solicitado. No hay evidencias de que los estudiantes han compartido ideas o procedimientos para construir o fortalecer algún conocimiento matemático, puesto que no hay un momento de socialización o discusión por equipos o en plenaria.	El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Su interacción con los estudiantes es ocasional, pero cuando lo hace pide que compartan ideas o procedimientos que apoyen la construcción o fortalecimiento de algún conocimiento matemático. Sin embargo no hay una discusión en plenaria sobre cómo se abordó (o se está abordando) algún problema matemático.	El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Interactúa constantemente con los estudiantes para sugerir que compartan sus ideas o procedimientos con los que construirán o fortalecerán algún conocimiento matemático. Docente y estudiantes intercambian ideas y en plenaria o equipos discuten cómo se abordó (o se está abordando) algún problema matemático.	El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Explica y promueve la interacción que se espera de los miembros de un mismo equipo para que compartan sus ideas o procedimientos, así como los resultados con los que construirán o fortalecerán algún conocimiento matemático. Los estudiantes son los principales actores en la discusión (en equipos o plenaria) sobre cómo se abordó (o se está abordando) algún problema matemático.

CRITERIOS DE SELECCIÓN DE PUNTAJES PARA CALIFICACIÓN DE DIMENSIONES

Instrucciones: A continuación se describen los criterios a partir de los cuales se sugiere seleccionar el puntaje correspondiente al nivel de desempeño identificado en las observación de una clase. Analice cuidadosamente las diferencias entre un puntaje y otro para que a partir de éste pueda seleccionar el puntaje adecuado al nivel de desempeño identificado.

NOTA: El observador deberá apegarse a la descripción de cada nivel de desempeño para identificar el puntaje correspondiente. Es decir, si señala el nivel medio, y la descripción del puntaje "4" indica que coincide con todos los criterios señalados en el nivel de desempeño valorado, entonces deberá marcar el puntaje "4", ya que es el que se apega a la descripción de lo observado.

CRITERIOS DE SELECCIÓN DE PUNTAJES							
BAJO		**MEDIO**		**MEDIO-ALTO**		**ALTO**	
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
La descripción del nivel bajo coincide con toda la descripción de los criterios que lo conforman	La descripción del nivel bajo coincide con la descripción, excepto en uno de los criterios que lo conforman	La descripción del nivel medio coincide con la descripción, excepto en uno de los criterios que lo conforman	La descripción del nivel medio coincide con toda la descripción de los criterios que lo conforman	La descripción del nivel medio-alto coincide con la descripción, excepto en uno de los criterios que lo conforman	La descripción del nivel medio-alto coincide con toda la descripción de los criterios que lo conforman	La descripción del nivel alto coincide con la descripción, excepto en uno de los criterios que lo conforman	La descripción del nivel alto coincide con toda la descripción de los criterios que lo conforman

HOJA DE CALIFICACIÓN POR TAREA MATEMÁTICA

Instrucciones: Marque con una "X" el puntaje correspondiente a las diferentes dimensiones evaluadas en cada una de las tareas matemáticas señaladas para clase observada. Asegúrese de anotar en la sección de "observaciones" información que le permita justificar el nivel de desempeño seleccionado.

NOTAS:

a. **Para calificar cada una de las tareas matemáticas que se identifican en la videograbación observada deberá apegarse a los criterios establecidos en las diferentes dimensiones valoradas, así como tomar el curso previo de capacitación.**
b. **En las opciones de puntuación se incluye "No se puede valorar" (N/V), si es el caso marque esta opción y señale en las observaciones las razones por las que no fue posible valorar dicha dimensión en la tarea matemática correspondiente.**

Fecha de observación _____________ **Número de tarea** ___________________
Código de videograbación ___________ **Iniciales del observador** _____________

La clase es continuación de la clase observada anteriormente **Sí** ☐ **No** ☐

POTENCIAL DE LA ACTIVIDAD DE ENSEÑANZA		
CLARIDAD DE LA TAREA MATEMÁTICA	Observaciones	N/V 1 2 3 4 5 6 7 8
EXIGENCIA COGNITIVA	Observaciones	N/V 1 2 3 4 5 6 7 8
RECURSOS EMPLEADOS DURANTE LA ACTIVIDAD DE ENSEÑANZA **NO APLICA ( )**	Observaciones	N/V 1 2 3 4 5 6 7 8
ELABORACIÓN DE COMUNICACIONES MATEMÁTICAS		
ALTERNATIVAS DE SOLUCIÓN	Observaciones	N/V 1 2 3 4 5 6 7 8
PREGUNTAS PARA LA REFLEXIÓN	Observaciones	N/V 1 2 3 4 5 6 7 8
TIPO DE SITUACIONES MATEMÁTICAS		
CONEXIÓN/ APLICACIÓN DE LOS CONOCIMIENTOS MATEMÁTICOS **NO APLICA ( )**	Observaciones	N/V 1 2 3 4 5 6 7 8
TRABAJO COLABORATIVO		
TRABAJO COLABORATIVO **NO APLICA ( )**	Observaciones	N/V 1 2 3 4 5 6 7 8

POAEM: EJEMPLOS POR DIMENSIÓN Y NIVEL DE DESEMPEÑO

En este apartado se incluye una descripción general de las dimensiones que conforman el *Protocolo de Observación de las Actividades de Enseñanza en Matemáticas* (POAEM). Además, se presentan ejemplos de situaciones relacionadas con cada uno los niveles de desempeño de las rúbricas que conforman este instrumento. Dichos ejemplos ofrecerán una idea de lo que se podría observar en algunas clases de matemáticas. Sin embargo, no se deberán tomar como aspectos de comparación, ya que la evaluación y calificación tendrá que estar sujeta a los criterios que se establecen en cada nivel de desempeño de las rúbricas que conforman el protocolo de observación.

1. **CLARIDAD DE LA TAREA MATEMÁTICA.** Se refiere a la manera en que el docente comunica el trabajo que los estudiantes deben desarrollar y cómo indaga si han comprendido lo que se les ha solicitado como trabajo o producto.

Nivel Alto: El docente plantea la tarea matemática y señala a los estudiantes los productos que espera logren. Utiliza un lenguaje matemático para enfatizar algunos conceptos o ideas importantes y verifica la comprensión de los estudiantes a partir de preguntas relacionadas con el objetivo o producto esperado (qué deben hacer), o a través de estrategias para que los estudiantes clarifiquen lo que deben lograr (lluvia de ideas, que expliquen con sus palabras, que comenten la tarea entre pares, entre otras).

> Ejemplo: Durante una clase el docente solicita a los estudiantes que revisen su libro de texto y ubiquen la tarea 78, con la que trabajarán el tema “préstamos con intereses”. Para introducir a los estudiantes a la tarea, pide que lean la situación problemática (de manera individual o en grupo) y posteriormente plantea a todo el grupo la pregunta "¿qué es lo que debemos hacer para obtener la información que nos están solicitando?" Con las respuestas de los estudiantes anota varias ideas en el pizarrón y va enfatizando los conceptos matemáticos que los propios estudiantes mencionan y en los que deberán centrarse; el porcentaje. Posteriormente pide a los estudiantes que en equipos elaboren una tabla en la que tendrán que calcular y comparar el interés mensual de las diferentes cantidades, para en algún momento de la clase exponerla a todo el grupo y poder discutir los procedimientos propuestos por los diferentes equipos.

Nivel Medio-alto. El docente plantea la tarea matemática y señala a los estudiantes los productos que espera logren. Utiliza un lenguaje matemático para enfatizar algunos conceptos o ideas importantes y reafirma las indicaciones dadas con otras palabras o con ejemplos. Sin embargo, no verifica la comprensión de los estudiantes a través de preguntas o estrategias que los ayuden a clarificar lo que deben realizar a partir de la tarea propuesta.

> Ejemplo: Durante una clase el docente solicita a los estudiantes que revisen su libro de texto y ubiquen la tarea 78, con la que trabajarán el tema “préstamos con intereses”. Pide a los estudiantes que lean la situación problemática (de manera

individual o en grupo). Posteriormente, retoma el tema y comenta que cuando se pide dinero prestado en un banco, éste nos presta cobrándonos cierta cantidad adicional por la cantidad solicitada, lo cual se conoce como interés y se determina a partir de cierto porcentaje (tema visto anteriormente). Luego explica que lo mismo sucede con otras instancias que se dedican a prestar dinero, y sugiere que para realizar la tarea propuesta, elaboren una tabla en la que calculen y comparen el interés mensual de diferentes cantidades, la cual deberán exponer a todo el grupo en algún momento de la clase.

Nivel medio. El docente plantea la tarea matemática y señala a los estudiantes los productos que espera logren a partir de ésta. Enfatiza conceptos o ideas matemáticas importantes para el desarrollo de la misma.

Ejemplo: Durante la clase el docente escribe en el pizarrón el tema “El interés” y pide a los estudiantes que escriban en su cuaderno el concepto que les dictará. Con base en la información proporcionada, el docente explica como el banco y otras instancias manejan el interés en los prestamos. Después dibuja en el pizarrón una tabla con dos columnas, una para diferentes cantidades de dinero y otra para el interés correspondiente al 4% de cada cantidad. Posteriormente, solicita a los estudiantes que copien la información en su cuaderno y les pide que de forma individual calculen el interés en cada caso, ya que una vez que hayan hecho los cálculos, en plenaria completarán la tabla, para revisar los resultados de diferentes compañeros.

Nivel Bajo: El docente no enfatiza conceptos o ideas matemáticas importantes para el desarrollo de la tarea. Si varios estudiantes preguntan qué deben hacer, el docente pide que vuelvan a leer las instrucciones o él mismo repite las instrucciones de trabajo mencionadas previamente.

Ejemplo : Durante una clase el docente solicita a los estudiantes que vayan a la pagina 145 de su libro de texto y lean en plenaria la tarea 78, con el tema “préstamos con intereses”. Una vez que se ha leído la tarea, el docente da la indicación de que se organicen en parejas para resolver la tarea. Cuando varios alumnos preguntan qué deben hacer con el problema (ya sea al docente o a otros compañeros) el docente pide a uno o varios alumnos que lean las instrucciones de la tarea en el libro y posteriormente reitera las indicaciones de organización en parejas para realizar el trabajo solicitado.

2. **EXIGENCIA COGNITIVA.** Se refiere al tipo y nivel de pensamiento requerido para solucionar ciertas tareas matemáticas. Por ejemplo reproducir de memoria hechos, reglas o fórmulas son aprendizajes de baja exigencia cognitiva, mientras que explorar y establecer relaciones de conceptos matemáticos representa una mayor exigencia. Es decir, mientras más habilidades cognitivas deba poner en juego el estudiante, la tarea requerirá una mayor exigencia cognitiva.

Nivel Alto: La tarea requiere un pensamiento complejo (no algorítmico), en el que los estudiantes deben integrar elementos para resolver el problema matemático. Los estudiantes deben explorar, diseñar, representar, demostrar o generalizar conceptos, algoritmos o procedimientos para establecer soluciones y lograr un entendimiento de la situación propuesta.

> Ejemplo: Tomando como base el tema de fracciones equivalentes (revisado previamente), el docente solicita a los estudiantes que en equipos diseñen dos situaciones problemáticas relacionadas con el tema, en los que podrán incluir ejemplos gráficos de productos reales o matemáticos. En seguida, aclara que las situaciones problemáticas serán intercambiadas entre los diferentes equipos para solucionarlas. Por lo que cada equipo deberá exponer a los compañeros el procedimiento de solución de la situación problemática que les haya tocado, así como si encontraron alguna relación entre los procedimientos seguidos para resolver ambas situaciones (estrategias). Los estudiantes de los diferentes equipos exponen sus respuestas por turnos, mientras el resto de los compañeros discuten para llegar a acuerdos sobre los diferentes procedimientos.

Nivel Medio-alto: La tarea propuesta sugiere procedimientos generales que se relacionan con ideas o conceptos matemáticos sobre los que debe reflexionar. Los estudiantes requieren analizar, comparar, clasificar, o evaluar información para lograr una comprensión de la situación propuesta.

> Ejemplo: Durante la sesión de clase, el docente dibuja en el pizarrón cuatro barras (rectángulos) del mismo tamaño y las divide de manera diferente. Después, solicita a los estudiantes que hagan lo mismo en su cuaderno y de manera individual, determinen las fracciones que corresponden a cada barra (rectángulo) y analicen cuáles de esas fracciones son equivalentes entre sí y por qué. Para conocer las respuestas de los estudiantes el docente pide a diferentes estudiantes que den sus respuestas y las va anotando en el pizarrón mientras los demás, dan su opinión al respecto. Finalmente llegan a acuerdos sobre porque algunas de las fracciones son equivalentes.

Nivel medio: La tarea es algorítmica, requiere de procedimientos específicos para trabajar alguna idea o concepto matemático, pero sin promover un significado de ideas o conceptos matemáticos. Los estudiantes deben identificar, reconocer o asociar algún procedimiento determinado.

> Ejemplo: Tomando como base el tema de fracciones equivalentes (revisado previamente), el docente solicita a los estudiantes que encuentren tres fracciones equivalentes por cada una de las que el ha propuesto en el pizarrón.

Nivel bajo: La tarea se limita a promover la reproducción de hechos, reglas, fórmulas o definiciones previamente aprendidas. Implican la reproducción exacta de conceptos o procedimientos (mecanización).

Ejemplo: El docente dicta o escribe en el pizarrón el concepto de fracción equivalente y solicita a los estudiantes que lo escriban en su cuaderno. Posteriormente anota algunas fracciones como ejemplo en el pizarrón y las explica. Para reforzar su explicación anota otras fracciones similares a las que utilizó como ejemplos y pide a los estudiantes que las resuelvan.

3. **RECURSOS EMPLEADOS DURANTE LA ACTIVIDAD DE ENSEÑANZA.** Se refiere al uso que el docente y los estudiantes hacen de los recursos disponibles para promover o fortalecer los aprendizajes a partir de la tarea matemática propuesta.

Nivel Alto: El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros, para promover el aprendizaje de las matemáticas. Éstos son coherentes con el propósito de la tarea propuesta, ya que se toman como base para llevar a cabo el análisis, discusión, evaluación o comparación de las ideas, procedimientos o estrategias de solución empleadas para realizar la tarea matemática. Además, permiten a los estudiantes comprobar sus propios procedimientos o estrategias de solución y autocorregir los procedimientos empleados para realizar la tarea matemática.

Ejemplo: Organizados en equipos el docente entrega a los estudiantes un pedazo de hilo y una regla para que puedan medir la longitud del hilo. Después pide que con el hilo midan la circunferencia y el diámetro de tres objetos que tengan entre sus pertenecías y completen una tabla con los datos: objeto, circunferencia, diámetro, y cociente de la circunferencia y el diámetro. Con las respuestas de algunos equipos el docente solicita a los estudiantes que comenten cómo son los resultados de sus cocientes, para lo que la mayoría de los equipos señala que son muy parecidos, excepto uno en el que los resultados son muy diferentes.

El docente solicita a los miembros de ese equipo que compartan con el resto del grupo el procedimiento que han seguido para realizar la tarea, y mientras exponen su procedimiento; miembros de otros equipos piden la palabra para hacerles ver que se han confundido, y señalan que los datos para hacer el cociente debe ser la circunferencia entre el diámetro y no al contrario, ya que la circunferencia es más extensa. Aclarado lo anterior y con los nuevos cálculos del equipo, el resto de los estudiantes señala que los datos que todos han obtenido son muy parecidos. El docente aprovecha esta apreciación de los estudiantes, y les pregunta ¿por qué creen que esto ha sucedido? Uno de los equipos señala que es debido a que el diámetro cabe tres veces en la circunferencia y pedacito más, lo que comprueba pasando al frente a medir diferentes objetos con el pedazo de hilo que le proporcionó el docente.

Nivel Medio-alto: El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros, para promover el aprendizaje de las matemáticas. Éstos son coherentes con el propósito de la tarea propuesta, ya que se toman como base para llevar a cabo el análisis, discusión, evaluación o comparación de las ideas, procedimientos o estrategias de solución empleadas para realizar la tarea matemática. Sin embargo, no se utilizan para llegar a una comprobación de las estrategias o procedimientos empleados.

> Ejemplo: Organizados en equipos el docente entrega a los estudiantes un pedazo de hilo y una regla para que puedan medir su longitud. Después pide que con el hilo midan la circunferencia y el diámetro de tres objetos que tengan entre sus pertenecías y completen una tabla con los datos: objeto, circunferencia, diámetro, y cociente de la circunferencia entre el diámetro. Con las respuestas de algunos equipos el docente solicita a los estudiantes que comenten cómo son los resultados de sus cocientes, para lo que los estudiantes señalan que son muy parecidos. El docente aprovecha la apreciación de los estudiantes, y les pregunta ¿por qué creen que esto ha sucedido? Uno de los equipos señala que es debido a que el diámetro cabe tres veces en la circunferencia y un pedacito más; el resto del grupo está de acuerdo y el docente confirma la información explicando porqué.

Nivel medio: El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros. Éstos son coherentes con el propósito de la tarea propuesta, sin embargo sólo se emplean como medio de información o reforzamiento para el propósito matemático. No se toman como base para promover un análisis, discusión evaluación o comparación de ideas o procedimientos.

> Ejemplo: Después de una explicación en plenaria sobre la circunferencia y el diámetro (y de señalar algunos ejemplos), el docente solicita a los estudiantes que se reúnan en equipos y les entrega un pedazo de hilo, una regla para que midan su longitud y una hoja de trabajo con información para complementar. Pide que midan algunos objetos con forma circular teniendo en cuenta lo que les ha señalado previamente con los ejemplos y con base en ello complementen la información que se solicita en la hoja de trabajo. Indica que al terminar la tarea entreguen la hoja con los nombres de los integrantes de equipo.

Nivel bajo: El docente sugiere algún artefacto: libro de texto, hojas de trabajo, calculadora, software, objetos manipulables, entre otros. Sin embargo, estos no están alineados con un propósito de aprendizaje matemático. Es decir, se utilizan como medio para realizar actividades recreativas o complementarias a la tarea matemática.

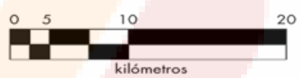
> Ejemplo: Después de una explicación en plenaria sobre la circunferencia y el diámetro (y de señalar algunos ejemplos), el docente entrega una hoja con círculos de diferente medidas y pide a los estudiantes que con diamantina y pegamento

(materiales de tarea) señalen los elementos de los círculos que se solicitan: en algunos casos sólo señalar la circunferencia o el diámetro y en otros casos rellenar todo el circulo.

4. **ALTERNATIVAS DE SOLUCIÓN.** Se refiere a la flexibilidad del docente para promover y aceptar una variedad de estrategias o procedimientos para resolver un problema matemático.

Nivel Alto: El docente permite que los estudiantes indaguen y propongan sus propios procedimientos. Además, solicita a los estudiantes que compartan y comparen los diferentes procedimientos empleados para comprobar que se pueden tomar varios caminos para llegar a un resultado (durante el desarrollo o al final de la tarea).

Ejemplo: El docente solicita a los estudiantes que se reúnan en equipos y les entrega un mapa a escala de la localidad. Solicita que de los diferentes lugares públicos que se incluyen en el mapa (museos, parques, instituciones públicas) elijan el que les gustaría visitar. Pide a los estudiantes que identifiquen la ruta más corta y más larga para trasladarse de la escuela al lugar seleccionado y que escriban las indicaciones. Además, solicita que calculen la distancia real aproximada para ambas rutas utilizando la escala gráfica que se incluye en el mapa.



Para realizar la tarea propuesta, el docente pide a los estudiantes que comenten con los compañeros de equipo cómo podrían emplear la escala gráfica o si podrían emplear otro método para determinar las respectivas distancias. Menciona que deberán compartir con el resto del grupo los procedimientos empleados para discutir y comparar los diferentes procedimientos. Si, detecta que los estudiantes tienen dificultad los apoya haciendo preguntas relacionadas con la tarea matemática, pero sin sugerir un procedimiento o respuesta para resolverla.

Nivel Medio-alto: El docente permite que los estudiantes indaguen sus propios procedimientos. Sin embargo, éstos no se retoman para compartir los diferentes procedimientos empleados. Sólo se comprueba que se haya llegado al resultado (generalmente al final de la tarea).

Ejemplo: El docente solicita a los estudiantes que se reúnan en equipos y les entrega un mapa a escala de la localidad. Posteriormente les pide que elijan uno de los lugares públicos que se incluyen en éste (museos, parques, instituciones públicas) y que calculen la distancia real aproximada de la escuela al lugar elegido siguiendo la ruta más corta y más larga utilizando la escala gráfica que se incluye en el mapa.

Señala que la gráfica está en kilómetros y que deberán estimar la distancia real aproximada en cada caso y anotar en su cuaderno los pasos que siguieron para lograrlo. Pide que al terminar lo señalen para revisar que la respuesta sea correcta. Si los estudiantes tienen dudas, el docente los apoya con respecto a la unidad de medida de la escala, pero sin sugerir un procedimiento específico.

Nivel Medio: El docente solicita a los estudiantes que apliquen algún procedimiento específico. Consulta si el problema se podría resolver de otra manera y escucha las respuestas de los estudiantes. Sin embargo, al final se enfatiza o promueve un procedimiento específico.

Ejemplo: El docente solicita a los estudiantes que se reúnan en equipos y les entrega un mapa a escala de la localidad. Posteriormente pide que identifiquen cuál de los dos lugares marcados (Parque México y el Museo Ferrocarrilero) está más cerca de la escuela. Solicita que comprueben su respuesta calculando la distancia real aproximada tomando como base la escala gráfica que se incluye en el mapa.

Sugiere a los estudiantes considerar las diferentes medidas que se incluyen en la escala (5, 10 y 20 km). Explica la respuesta en plenaria y pregunta a los estudiantes si creen que hay otras formas de emplear la escala gráfica u otro método para determinar la distancia aproximada. Sin embargo, no hay una comprobación o comparación de otros posibles procedimientos.

Nivel Bajo: Los estudiantes no tienen oportunidad de aplicar múltiples estrategias de solución, puesto que se les solicita seguir (o reproducir) un procedimiento, fórmula o método de solución específico (explicado previamente por el docente o señalado en la tarea matemática propuesta).

Ejemplo: El docente explica a los estudiantes lo que es una escala gráfica y utiliza un mapa de la localidad para dar un ejemplo. Luego, en conjunto con el grupo eligen un lugar del mapa para calcular la distancia real aproximada a partir de la escala dada. Posteriormente, solicita a los estudiantes que se reúnan en equipos e identifiquen en el mapa el “Parque México” y el “Museo Ferrocarrilero”; para calcular la distancia de la escuela a ambos lugares, siguiendo las rutas marcadas. Si los estudiantes tiene dudas sobre como calcular las distancias, el docente vuelve a explicar con otro ejemplo el procedimiento propuesto.

5. **PREGUNTAS PARA LA REFLEXIÓN.** Se refiere al tipo de preguntas que se platean durante la clase (ya sea por el docente o los estudiantes) para darle sentido a ideas o relaciones matemáticas importantes.

Nivel Alto: El docente promueve la comprensión de ideas o procedimientos matemáticos mediante preguntas que requieren que los estudiantes justifiquen y demuestren sus respuestas. Las preguntas son planteadas por el docente o por los propios estudiantes (se les anima para comentar y comprobar sus respuestas). El docente evita dar sus propios argumentos sobre la manera de resolver el problema matemático y retoma (en algunos casos) el razonamiento de los estudiantes para plantear nuevas preguntas.

> Ejemplo: Durante una exposición sobre la interpretación de una gráfica circular, el docente cuestiona a los estudiantes ¿por qué se empleó tal procedimiento y no otro?, ¿cómo están seguros de qué obtuvieron una respuesta correcta? Y ¿cómo pueden comprobarlo?. Además, solicita a los estudiantes que señalen las dudas que tiene respecto al procedimiento empleado por algunos de los compañeros (que han pasado al pizarrón), de tal manera que los expositores puedan justificar sus respuestas y con ejemplos demuestren el resultado obtenido.

Nivel Medio-alto: El docente plantea preguntas a través de las cuales solicita a los estudiantes que expliquen y justifiquen las estrategias o procedimientos empleados para resolver algún problema matemático. El docente hace uso de sus propios argumentos y demostraciones para reforzar alguna idea o procedimiento matemático específico.

> Ejemplo: Durante el desarrollo de una tarea matemática en la que los estudiantes deben interpretar una gráfica circular, el docente solicita a los estudiantes que expliquen a los compañeros (o para el docente) ¿cuál fue el procedimiento que siguieron?, y ¿por qué lo realizaron de esa manera? (es la más adecuada, existen otras, etc.). Sin embargo, no pide a los estudiantes que comprueben las respuestas dadas, en su lugar el propio docente hace la comprobación del procedimiento que siguieron los estudiantes y enfatiza algunos aspectos o conceptos matemáticos importantes.

Nivel Medio: El docente plantea preguntas que llevan a los estudiantes a describir los pasos de las estrategias o procedimientos de solución que emplearon para resolver algún problema matemático. No se solicita una justificación o demostración de la respuesta obtenida.

> Ejemplo: Durante el desarrollo de una tarea matemática en la que los estudiantes deben interpretar una gráfica circular, el docente solicita a los estudiantes que comenten con los compañeros o con todo el grupo ¿cuales son los pasos que siguieron para responder las preguntas de la tarea. Por ejemplo: para la pregunta ¿qué número de paletas representa el 30% en la gráfica?, señalan que fue necesario en principio conocer el número total de paletas, el cual representará el 100% (80), para después multiplicar el 30% por 80 paletas y dividir el resultado entre l00%, dando un total de 24 paletas (otra manera es por deducción). El docente confirma o corrige algunos de los procedimientos empleados.

Nivel Bajo: No hay preguntas que permitan a los estudiantes explicar, justificar o demostrar una estrategia o procedimiento de solución. El docente da indicaciones o guía a los estudiantes para aplicar algún procedimiento, lo revisa y señala si es correcto o incorrecto.

> Ejemplo: Durante el desarrollo de una tarea matemática en la que los estudiantes deben interpretar una gráfica circular, el docente pasa a los equipos o con los estudiantes de manera individual para asegurarse que están logrando el trabajo propuesto. El docente se asegura de indicar a los estudiantes si van por un camino correcto o si podrían emplear otra estrategia que les ayude a resolver la tarea matemática.

6. **CONEXIÓN/ APLICACIÓN DE LOS CONOCIMIENTOS MATEMÁTICOS.** Se refiere a la medida en que el contexto de una *situación problemática* ofrece a los estudiantes la oportunidad de conectar los aprendizajes matemáticos con su propia experiencia, con otras disciplinas o con situaciones de la vida cotidiana, como una manera de reconocer los aprendizajes matemáticos por su relevancia y utilidad.

Nivel Alto: La tarea matemática propuesta a los estudiantes se centra en una situación problemática que incluye información o datos que los estudiantes pueden reconocer por su experiencia o por situaciones de la vida cotidiana, así como por una conexión explicita entre diferentes temas o disciplinas. Los estudiantes pueden establecer un vínculo claro entre las matemáticas que están aprendiendo y su aplicación o relevancia en diferentes contextos o situaciones.

> Ejemplo: Durante una clase, el docente pregunta a los estudiantes si en alguna ocasión han escuchado sobre los descuentos de productos, y si es así en dónde recuerdan haberlo escuchado. Posteriormente, da oportunidad a los estudiantes para que mencionen algunos ejemplos, pero una vez aclarado el punto, entrega a los estudiantes algunas hojas de periódicos y les pide que en equipo localicen dos productos que tengan descuento, con base en los cuales deberán responder dos preguntas: a) ¿a cuanto dinero corresponde el porcentaje de descuento aplicado al costo total de cada producto? y b) si se aplicara un 15% de descuento extra al señalado en cada producto ¿cuál sería el costo total a pagar en cada caso?. Para ampliar el tema de porcentajes, el docente solicita a los estudiantes que en equipo comenten en que otras situaciones utilizamos los porcentajes y establezcan dos ejemplos para compartir con todo el grupo (datos de la población, datos deportivos, etc.).

Nivel Medio-alto: La tarea matemática propuesta a los estudiantes se centra en una situación problemática en la que se incluye información sobre diferentes temas o disciplinas. La tarea establece una problematización en la que es posible vislumbrar la relación de los conocimientos matemáticos para fortalecer otras disciplinas o campos de trabajo.

> Ejemplo: Durante una clase en donde se trabaja el tema de porcentajes, el docente plantea a los estudiantes una tarea en la que los estudiantes deben analizar una tabla que concentra información sobre la división política de México y su extensión territorial total y por estado en kilómetros cuadrados. Con base en dicha información, el docente solicita a los estudiantes que seleccionen dos estados de nuestro país y determinen el porcentaje que su extensión representa del total del territorio del país.

Nivel Medio: La tarea matemática propuesta a los estudiantes se centra en una situación problemática en la que se incluye información relacionada con el ámbito académico o el contexto escolar. Sin embargo, la tarea establece una problematización en la que se enfatizan sólo cuestiones relacionadas con la disciplina.

> Ejemplo: Para trabajar el tema de porcentajes, el docente solicitó a los estudiantes que durante el receso (en una sesión anterior) preguntaran a tres estudiantes de otros grados y grupos, cual de los siguientes productos prefieren comprar en la cooperativa como almuerzo: a) tacos de guisado, b) torta de jamón, c) papas fritas o e) pizza. Con los resultados que cada niño señaló, el docente dibujó una tabla en la que concentraron la información, con un total de 102 encuestados. A partir de la información concentrada solicitó a los estudiantes que en equipos elaboraran una gráfica circular para determinar el porcentaje correspondiente a cada producto de acuerdo con el total de niños encuestados.

Nivel Bajo: La tarea matemática propuesta a los estudiantes hace referencia a situaciones que no resultan prácticas o aplicables a las experiencias de los estudiantes o situaciones cotidianas. No se establece una conexión clara entre las matemáticas que están aprendiendo (ideas o procedimientos) y su aplicación o relevancia en diferentes contextos o situaciones.

> Ejemplo: Durante una clase el docente escribe en el pizarrón una tabla en la que se incluye tres columnas, la primera señala un porcentaje, la segunda la formula n/100, y la última indica una fracción. Luego, solicita a los estudiantes que la anoten en su cuaderno y resuelva el siguiente problema " si la mitad de una cantidad es 50%, qué parte de esa cantidad es 10%, 25% y 75%". Los estudiantes deben completar una tabla con lo que se espera que los estudiantes comprendan la razón n/100, sin embargo la situación no incluye información que los estudiantes puedan vincular con alguna experiencia o situación de la vida cotidiana.

7. **TRABAJO COLABORATIVO.** Se refiere a la interacción, negociación y cooperación entre estudiantes para lograr un propósito común.

Nivel Alto: El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Explica y promueve la interacción que se espera de los miembros de un mismo equipo para que compartan sus ideas o procedimientos, así como los resultados con los que construirán o fortalecerán algún conocimiento matemático. Los estudiantes son los principales actores en la discusión (en equipos o plenaria) sobre cómo se abordó (o se está abordando) algún problema matemático.

> Ejemplo: El docente pide a los estudiantes que reunidos en equipos elaboren un problema sobre reparto (tema visto previamente), que implique usar y comparar fracciones. Para ello, entrega una hoja de papel bond en las que escribirán la situación problemática que expondrán al resto del grupo. Durante el trabajo, el docente supervisa la participación de los estudiantes en los equipos, y sugiere que cada integrante comente una idea, para que todos puedan discutir y elegir cuales resultan más apropiadas para plantear el problema solicitado. Además, pide a los estudiantes que le expliquen cómo han organizado el trabajo en el equipo: quienes serán los encargados de escribir, quienes presentarán los problemas al resto de los compañeros, etc.
>
> Durante la presentación en plenaria, el docente alienta al resto de los estudiantes para que planteen sus dudas sobre los problemas y señalen alguna forma de resolverlos, de tal manera que todos los estudiantes tengan oportunidad de aportar sus ideas y logren acuerdos sobre los procedimiento o ideas matemáticas substanciales.

Nivel Medio-alto: El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Interactúa constantemente con los estudiantes para sugerir que compartan sus ideas o procedimientos con los que construirán o fortalecerán algún conocimiento matemático. Docente y estudiantes intercambian ideas y en plenaria o equipos discuten cómo se abordó (o se está abordando) algún problema matemático.

> Ejemplo: El docente pide a los estudiantes que reunidos en equipos elaboren un problema sobre reparto (tema visto previamente), que implique usar y comparar fracciones. Con ello, promueve la participación de los estudiantes (en los equipos y para todo el grupo) mediante preguntas que los ayudan explicar sus ideas o propuestas para establecer los problemas solicitados, mismos que intercambiarán con otros equipos. Para cerrar la actividad, el docente pide a dos equipos que expongan los problemas que diseñaron, y los toma como base para discutir en plenaria cómo se pueden resolver y qué coincidencias identifican en los procesos empleados.

Nivel Medio: El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Su interacción con los estudiantes es ocasional, pero cuando lo hace pide que compartan ideas o procedimientos que apoyen la construcción o fortalecimiento de algún conocimiento matemático. Sin embargo no hay una discusión en plenaria sobre cómo se abordó (o se está abordando) algún problema matemático.

> Ejemplo: El docente pide a los estudiantes que reunidos en binas resuelva un problema sobre reparto que implica usar y comparar fracciones. Da oportunidad para que los estudiantes comenten y decidan la forma de solución, por lo que la mayoría de los estudiantes parece estar comentando sus ideas. Sin embargo, sólo en algunos casos el docente pide a los estudiantes que comenten las ideas que han discutido y las estrategias que decidieron emplear. Después indica para todo el grupo que conforme terminen el trabajo lo señalen para pasar a revisar a sus lugares.

Nivel Bajo: El docente solicita a los estudiantes que trabajen en conjunto (pares, pequeños grupos, todo el grupo) para desarrollar alguna idea, procedimiento o producto común. Sin embargo, su interacción con los estudiantes se limita a verificar que estén desarrollando el trabajo solicitado. No hay evidencias de que los estudiantes han compartido ideas o procedimientos para construir o fortalecer algún conocimiento matemático, puesto que no hay un momento de socialización o discusión por equipos o en plenaria.

Ejemplo: El docente pide a los estudiantes que reunidos en binas resuelva un problema sobre reparto que implica usar y comparar fracciones. Sin embargo, los estudiantes trabajan de forma individual dentro del equipo, por lo que no se identifica una discusión de ideas o propuestas, sino una comparación de resultados o en algunos casos la explicación de un estudiante al otro para resolver el problema sugerido por el docente.

TESIS TESIS TESIS TESIS TESIS