

CRITERIOS DE EVALUACIÓN INTERPRETADOS POR ESTUDIANTES DE SEGUNDO Y TERCER AÑO MEDIO EN LA ASIGNATURA DE MATEMÁTICAS EN CHILE

ASSESSMENT CRITERIA INTERPRETED BY MIDDLE EDUCATION STUDENTS OF SECOND AND THIRD YEAR IN THE COURSE OF MATHEMATICS IN CHILE

Bastián Peters Rodríguez(*)

Universidad de Viña del Mar

Gloria Contreras

Pontificia Universidad Católica de Valparaíso

Resumen

Este artículo presenta una caracterización de las interpretaciones que estudiantes de enseñanza media otorgan a los criterios de evaluación en la asignatura de Matemáticas. La investigación se asienta en un paradigma cualitativo hermenéutico, en la cual se entrevistaron once estudiantes pertenecientes a dos establecimientos de la Quinta Región de Chile. Toda la información recolectada se ordenó y analizó mediante categorización de nodos libres a través del software NVivo. La interpretación de la información logró describir y facilitar la comprensión de los criterios que los estudiantes identifican y aquellos a los que otorgan mayor valoración, tales como criterios asociados a la calificación o a relaciones interpersonales entre profesor y estudiantes, que en su mayoría dependen de las prácticas de sus profesores, pero también a factores externos como la Prueba de Selección Universitaria.

Palabras clave: Evaluación del estudiante; evaluación formativa; criterio de evaluación; interpretación; docente.

Abstract

This article presents a characterization of the interpretations that secondary education students give to the assessment criteria in the course of mathematics. The research is based on a qualitative hermeneutical paradigm, where eleven students from two schools of the Fifth Region of Chile were interviewed. All the collected information was ordered and analyzed through the categorization of free nodes using the Nvivo software. The interpretation of the information managed to describe and facilitate the understanding of the criteria that students identify and those to which they give greater value, such as criteria associated to the grades or the teacher-student relationship, which mostly depend on the practices of their teacher, but also external factors such as the PSU standardized test.

Keywords: Educational assessment; formative assessment; assessment criteria; interpreting; teacher.

(*) **Autor para correspondencia:**

Bastián Peters Rodríguez

Universidad de Viña del Mar

Escuela de Educación

Formación Inicial

Agua Santa 7055, sector Rodelillo, Viña del Mar

Correo electrónico: bastian.peters@uvm.cl

©2010, Perspectiva Educacional

[Http://www.perspectivaeducacional.cl](http://www.perspectivaeducacional.cl)

RECIBIDO: 06.12.2018

ACEPTADO: 06.08.2019

DOI: 10.4151/07189729-Vol.58-Iss.3-Art.882

1. Introducción

Uno de los elementos fundamentales del proceso de evaluación en cualquier nivel de escolaridad está constituido por los criterios y estándares de evaluación, específicamente por la comunicación con los estudiantes. Faltan investigaciones respecto a la forma en que los criterios se comunican, comprenden, y su impacto en el aprendizaje de los estudiantes. Muchos autores en la literatura inglesa han entendido la importancia de los criterios de evaluación y han desarrollado teoría a partir de distintas líneas y enfoques (Barrie, Brew & McCulloch, 1999; Bell, Mladenovic & Price, 2013; Davinroy, 1995; Ferrara, Goldberg & McTighe, 1995; O'Donovan, Price & Rust, 2008; Price, 2005; Sadler, 1983, 1985, 1989, 2005, 2009b, 2010, 2015).

En el informe de rendimiento de Matemáticas en la prueba PISA de la Organización para la Cooperación y el Desarrollo Económicos (OECD, 2015), Chile se ubica bajo el promedio del puntaje, estando solo sobre Indonesia, Brasil, Perú, Colombia, México y Turquía, de un total de 44 miembros. Esta situación otorga ciertas luces sobre la situación que vivimos en nuestro país.

La siguiente reflexión de Cantoral, Reyes-Gasperini y Montiel (2014) aporta a la comprensión del problema:

La matemática escolar es vista por muchos alumnos como irrelevante y de poca utilidad en sus vidas profesionales. Las miran como una asignatura aburrida, repleta de técnicas y “trucos” difíciles de aprender y basadas en procedimientos adquiridos por repetición memorística. Las Matemáticas, en tanto actividad humana, se plantean por el contrario una gran cantidad de retos interesantes relativos a las ciencias, las tecnologías y la innovación ubicados en la realidad material, orgánica y social que precisan de procesos creativos basados en heurísticas y en el pensamiento crítico, de sofisticados procesos de socialización del saber en el marco de prácticas de referencia diversas. Sin embargo, la enseñanza tradicional, a la que se denomina estructuralista, las presenta como un sistema de razón estructurado lógicamente, como un lenguaje formal. (104)

Picaroni y Loureiro (2010) señalan que los procesos de evaluación en Matemáticas son significados como un conjunto de tareas de reconocimiento de nombres y características, en especial en geometría, en que la aplicación se entiende como ejercitación mecánica descontextualizada y propia del cálculo aislado. Las autoras afirman también que los docentes de Matemáticas suelen tener concepciones instrumentales sobre la evaluación, es decir, que

existe una predominancia de actividades evaluativas en las que los estudiantes ponen en juego saberes procedimentales más que aprendizajes significativos.

La importancia de los criterios de evaluación en la asignatura de Matemáticas viene dada por lo siguiente: los profesores constantemente emiten juicios sobre la calidad de los trabajos de sus estudiantes y ellos entiendan las razones sobre estos juicios, deben tener claridad sobre los criterios utilizados. Según Sadler (1983), cuando los estudiantes identifican los criterios, pueden entender las razones de las decisiones del docente ya que, estos entregan los fundamentos de los juicios realizados. Por otra parte, los estudiantes podrían predecir futuras valoraciones porque, al comprender los criterios de su profesor, asimilan que en las siguientes evaluaciones podría existir una lógica similar.

Este artículo presenta los resultados de una investigación cuyo propósito fue caracterizar las interpretaciones que los estudiantes de segundo y tercer año de enseñanza media, de un establecimiento particular y uno subvencionado de la Quinta Región de Valparaíso, construyen sobre los criterios de evaluación comunicados en Matemáticas. Específicamente, se caracterizan aquellos aspectos, elementos o características que perciben los estudiantes como criterios de evaluación y también se analiza la forma en que participan de la construcción de criterios de evaluación en la asignatura de Matemáticas. La mayor parte de los criterios que han sido analizados pueden ser generalizados en cualquier asignatura.

2. Literatura

2.1. Evaluación formativa, evaluación para el aprendizaje

La evaluación formativa puede concebirse como un proceso al servicio de la enseñanza y se utiliza para ayudar a los alumnos a que autorregulen sus procesos de aprendizaje; es por esta razón que se la considera como evaluación para el aprendizaje (Picaroni, 2011). Según Popham (2013), la evaluación formativa es un proceso planificado en el que tanto los profesores como los alumnos usan la evidencia basada en la evaluación para realizar ajustes en lo que están haciendo. Es importante aclarar que, para ambos autores, la retroalimentación es un eje fundamental en la evaluación formativa y es necesario considerar las formas en que los profesores puedan utilizarla de manera efectiva; para ello, el docente debe encargarse de que los estudiantes se familiaricen con los criterios de evaluación.

2.2. Criterios de evaluación

En el ámbito escolar y en la evaluación del aprendizaje, los criterios serían todas aquellas propiedades que son potencialmente útiles para evaluar a todos los miembros de una clase (Sadler, 1985).

Para efectos de esta investigación se utilizará la siguiente conceptualización propuesta por Sadler (2005):

Un criterio es una propiedad distintiva o característica de cualquier cosa, por el cual su calidad puede ser juzgada o estimada, o por el cual se puede hacer una decisión o clasificación. (Etimología: del griego *Kriterion*: un medio para juzgar). (178-179)

Por otro lado, “estándar” es un nivel de logro mínimo utilizado como punto de referencia al juzgar la calidad del trabajo de un estudiante para asignarle un código (marca, calificación o grado) apropiado (Sadler, 2014). Los criterios y estándares constituyen dos elementos clave en el discurso evaluativo, pues generalmente las evaluaciones deben ser justificadas por referencias a criterios y estándares (Sadler, 1985).

Si bien los criterios y estándares pueden ser especificados previamente por el docente, al establecer un conjunto de ellos para evaluar los trabajos de los estudiantes, pueden emerger ciertos criterios no especificados debido a las diferentes interpretaciones sobre lo que los estudiantes entienden que deben hacer. Es por esta razón que Sadler (1983), expone ciertas dificultades en el uso de los criterios. Por una parte, mientras mayor sea el número de criterios a utilizar, significa que mayor es la cantidad de dimensiones que deben ser atendidas y la evaluación se puede tornar dificultosa. Por otra parte, es necesario considerar también las diferencias entre cada docente, ya que cada uno utiliza distintas formas para emplear los criterios. El mismo autor menciona al menos tres formas: (1) disyuntivamente: es decir, para que un trabajo o “actuación” sea considerada como aceptable, entonces debe alcanzar niveles mínimos (estándares) de calificación en cierto número de criterios propuestos; (2) conjuntivamente: un trabajo que puede ser considerado de calidad o excepcional en alguna dimensión posee errores en las demás áreas porque estas no han sido consideradas; y (3) regla compensatoria simple: trabajos pobres en ciertas características se equilibran con los altos niveles alcanzados en otras.

Los estudiantes deben ser capaces de controlar la calidad de su propio trabajo durante su desarrollo, y para que esto suceda se requiere que comprendan cuán alta debe ser la calidad de

su trabajo, además de poseer cierta destreza evaluativa que les permita comparar con objetividad esta calidad en relación con un estándar más alto (Sadler, 2009a).

Para los profesores de Matemáticas, la comunicación con sus estudiantes no está exenta de dificultades y ambigüedades. Los estudiantes suelen escribir de una manera confusa y vaga (Stonewater, 2002, en Santos & Semana, 2015), lo que puede determinar ciertos cambios en los criterios de evaluación. En el mismo estudio se muestran indicios sobre cómo los profesores de Matemáticas podrían utilizar la escritura para comunicarse con sus estudiantes y propiciar que ellos expresen sus pensamientos mediante sus propias palabras. Así, los estudiantes podrían interpretar correctamente lo que su profesor comunica y a su vez los profesores conseguirían comunicar efectivamente a sus estudiantes lo que esperan de ellos en las tareas propuestas ya que, como señalan Wiggins y McTighe (2005), cada estudiante puede tener significados múltiples sobre los distintos elementos por los cuales serán evaluados. Algo similar señala Sadler (2015), respecto a cómo el docente puede permitir o incluso forzar a sus estudiantes a que construyan juicios de manera holística y luego pedirles que otorguen las razones de estos juicios utilizando sus propias palabras. De esta forma, tanto estudiantes como profesores lograrán unificar los conceptos y el sentido que tomarán los criterios que se utilizarán en Matemáticas.

Por otra parte, aunque se trata de una disciplina considerada objetiva, muchos profesores corrigen y califican las tareas de sus estudiantes desde sus propias creencias y experiencias con cada estudiante (Wang & Cai, 2018).

3. Método

3.1. Diseño de investigación

Se trabaja desde un paradigma cualitativo hermenéutico, ya que este está orientado al estudio en profundidad de la compleja realidad social (Osses, Sánchez & Ibáñez, 2006). La metodología cualitativa, en su sentido más amplio, hace referencia a la investigación que produce datos descriptivos: las propias palabras de las personas, elementos y cualidades que debe tener un estudio y un investigador que se encuentran sumergidos en el paradigma cualitativo.

El diseño metodológico se asienta en una epistemología hermenéutica, ya que para comprender se necesita interpretar el mundo simbólico que le da un significado a lo que las personas dicen y hacen. Como dicen Wiggins y McTighe (2005), el significado se encuentra en el ojo de quien contemple la situación o el momento, y es esta la razón por la que es tan

importante estudiarlo, ya que los significados y patrones que se atribuyen a todos los conocimientos, datos o experiencias, transforman la comprensión de hechos específicos.

3.2. Técnica de recogida de datos

En el contexto de aula, las relaciones que se generan entre estudiantes y profesor están llenas de subjetividades. Las prácticas evaluativas de los docentes están cargadas con experiencias previas, y cada momento de comunicación que surge en estas relaciones interpersonales son significadas o valoradas por los estudiantes, no solo como individuos, sino también como grupo. Es en esta complejidad donde se pretende profundizar, analizando lo que subyace del discurso de los momentos en los que se comunican y comparten los criterios de evaluación en Matemáticas. La técnica utilizada es una entrevista, que, de acuerdo con Taylor y Bogdan (1987), es un relato cara a cara entre el investigador y el informante en donde se pretende comprender las perspectivas del informante respecto de sus vidas, experiencias o situaciones, tal y como lo expresan con sus propias palabras.

La entrevista dirigida al estudiante de educación media tiene por objetivo recoger información profunda sobre las experiencias evaluativas vividas dentro de la asignatura de Matemáticas o fuera de ella, colocando énfasis a los momentos donde se estén comunicando criterios de evaluación y utilizando elementos que controlan los parámetros de la calidad de sus trabajos. Sus ejes temáticos son: reconocimiento de los criterios de evaluación; criterios de evaluación utilizados por los estudiantes, y participación de los estudiantes en el proceso evaluativo. Las entrevistas se llevaron a cabo en los establecimientos, una sala de clases u oficina en donde el estudiante se sintiera cómodo. Para esto el investigador visitó a los estudiantes con anterioridad para que lo conocieran y entendieran en qué consistiría su participación en la investigación. La duración de cada entrevista varía entre los 40 y 62 minutos, comprendiendo que con algunos estudiantes fue más arduo establecer una comunicación fluida.

3.3. Sujetos de estudio y su contexto

Los estudiantes se han seleccionado con base en criterios que le darán validez a la nueva información. En la siguiente tabla se muestra un resumen de la distribución de los estudiantes con sus respectivos establecimientos, profesores, curso y rendimiento en Matemáticas. Este rendimiento es una apreciación directa del docente frente al comportamiento del estudiante tanto en actividades de la asignatura como en sus calificaciones.

Tabla 1
Distribución y características de los estudiantes

ESTABLECIMIENTO	PROFESORES	CURSO	RENDIMIENTO	ESTUDIANTES
SUBVENCIONADO	Profesor 1	3º	alto	Entrevista 1
			medio	Entrevista 2
			bajo	Entrevista 3
			medio	Entrevista 4
	Profesor 2	2º	bajo	Entrevista 5
			alto	Entrevista 6
			medio	Entrevista 7
			alto	Entrevista 8
PARTICULAR	Profesor 3	2º	bajo	Entrevista 9
			alto	Entrevista 10
			medio	-
	Profesor 4	3º	bajo	Entrevista 11

Fuente: Elaboración propia.

3.4. Sobre los establecimientos

En Chile las escuelas se clasifican, de acuerdo con su dependencia administrativa y financiera, en municipales (establecimientos escolares públicos de propiedad y financiamiento estatal, administrados por las municipalidades del país); particulares subvencionadas (de propiedad y financiamiento privado, pero que reciben un financiamiento estatal por cada alumno matriculado y que asiste a clases) y particulares pagadas (de propiedad, administración y financiamiento privados) (Roco, 2010). Los resultados de la principal prueba estandarizada que se aplica anualmente en el país (Sistema de Medición de la Calidad de la Educación, SIMCE) arrojan diferencias significativas entre las escuelas de distintas dependencias, lo que lleva a suponer que los procesos de enseñanza, aprendizaje y evaluación, entre otros, se desarrollan con características distintivas.

Dado lo anterior, uno de los primeros criterios fue escoger dos establecimientos con diferente administración escolar, es decir particular, particular subvencionado o municipal. Teniendo en consideración lo anterior, la medida que se tomó fue escoger un establecimiento particular y otro particular subvencionado con alto índice de vulnerabilidad, para lograr una mayor heterogeneidad. Ambos poseen niveles socioeconómicos distintos y, además, son diferentes en cuanto al género, ya que uno es mixto y el otro solo de varones. Ambos pertenecen a la región de Valparaíso, comuna de Quillota.

3.5. Sobre los participantes

Se trata de once estudiantes de ambos sexos, de 2º y 3º medio cuya elección fue realizada a partir de un criterio exigido a los profesores: los 3 estudiantes de cada curso debían ser distintos –de acuerdo a la percepción de los docentes– en cuanto a las habilidades y rendimiento que estos poseen en Matemáticas: alguien con facilidades, uno con ciertas dificultades y otro con mucha dificultad para las Matemáticas, sin que la variable rendimiento y calificación sea tan gravitante en su decisión, ya que por ejemplo, podríamos considerar a un estudiante que tenga dificultades para las Matemáticas pero buenas notas o viceversa. La elección no debía fundamentarse solo en las notas de sus estudiantes, sino también en su participación o actuaciones durante el desarrollo de las clases, momentos claves para la comprensión de los criterios de evaluación.

Los profesores del área de Matemáticas son cuatro, dos de cada establecimiento, y cada uno de ellos a cargo de un 2º y un 3º medio.

3.6. Sobre los cursos escogidos

Se debe considerar que este tipo de estudio podría ser realizado en cualquier curso y en cualquier disciplina, ya que los criterios de evaluación siempre existirán pues los profesores deben juzgar las actuaciones de sus estudiantes.

La siguiente tabla presenta los criterios de inclusión y de exclusión sobre los cursos escogidos.

Tabla 2

Criterios de inclusión y exclusión en los cursos escogidos

Criterios de inclusión	Criterios de exclusión
Profesor de educación media	Cursos de enseñanza básica
Cursos de 3º medio	Cursos de 4º medio
Cursos de 2º medio	Cursos de 1º medio

Fuente: Elaboración propia.

4. Resultados

Las entrevistas realizadas a los estudiantes fueron grabadas y luego transcritas. El análisis del discurso se realizó con el método de categorización libre mediante la ayuda del software NVivo, el cual permitió almacenar la información de manera ordenada, creando cada entrevista como un caso y escogiendo cuatro niveles de clasificación para todas las entrevistas, según los criterios de selección de la unidad de estudio:

- Administración escolar (Establecimiento particular y subvencionado)
- Curso (2º y 3º)
- Rendimiento (Alto, medio y bajo)
- Sexo (Hombre y mujer)

Durante una primera aproximación se realizó un primer nivel de categorías (“nodos libres”), que surgen desde los discursos revisados en las transcripciones. El proceso formal para la construcción de los nodos libres es resumir el texto expuesto en cada entrevista, sin cambiar el sentido de lo que está expresando; sin embargo, cuando ya se tiene un mayor manejo sobre los contenidos de las entrevistas –debido a varias lecturas– se puede definir un nodo libre como un primer nivel de categoría o interpretación del investigador. Estos nodos se agrupan en categorías, desde donde emergen los núcleos de la información. Los núcleos corresponden al eje medular que agrupa las categorías, y algunos de estos nodos se van repitiendo, lo cual permite una cristalización del discurso. Según Conde (2009), la cristalización se refiere al nivel de consistencia propia y de materialización expresiva más o menos codificada que sostiene un discurso social, esto quiere decir que mientras más cristalizado sea un discurso social, mayor aprobación genera su participación.

Con el programa es posible fusionar los nodos que parecen estar relacionados en un nivel más amplio con otros, creando así las categorías madres y bases del árbol categorial. Dependiendo de la clasificación realizada a las entrevistas, el software permite comparar los discursos de los estudiantes mediante distintos atributos. Además, se pueden realizar búsquedas de palabras que sean frecuentes, crear nubes de palabras y realizar gráficos descriptivos o de frecuencias con la información.

En el siguiente párrafo se definirán los conceptos fundamentales de esta:

- a) **Núcleo:** eje de análisis que agrupa las categorías desde las cuales se desprenden los nodos libres. Corresponden a grupos de categorías levantadas desde los discursos que conforman el corpus.
- b) **Categorías:** aquí se agrupan los nodos libres. Estas categorías también pueden corresponder a nodos libres dentro de los cuales se subsumen más nodos.
- c) **Nodos libres:** partes del discurso que son resumidas durante un primer nivel de categorización.

Si bien, el tipo de análisis se llama categorización libre ya que se fundamenta en la libertad del investigador para escoger términos y conceptos que pueden ser frases identificadas en ciertos segmentos del discurso, el objetivo principal es organizar esto en árboles categoriales, en donde se enfoca el análisis y la triangulación de la información obtiene protagonismo. Esta se configura desde la interpretación del investigador sobre el discurso de los informantes y la fundamentación en conjunto con la literatura estudiada, lo que correspondería a una doble interpretación o doble hermenéutica.

Cada entrevista transcrita es importada en el software, sistematizando las once entrevistas mediante casos. Las entrevistas se abren como documento de lectura y se puede seleccionar parte del texto mediante un nodo, escogiendo un nombre que puede ser un resumen de lo que señala el estudiante en esa parte del discurso o incluso llevarse a un nivel más interpretativo, creando un primer nivel de categorías. A través de este proceso, el árbol categorial comienza a tomar forma definiendo los elementos que lo componen, como se especificó anteriormente.

4.1. Árbol categorial

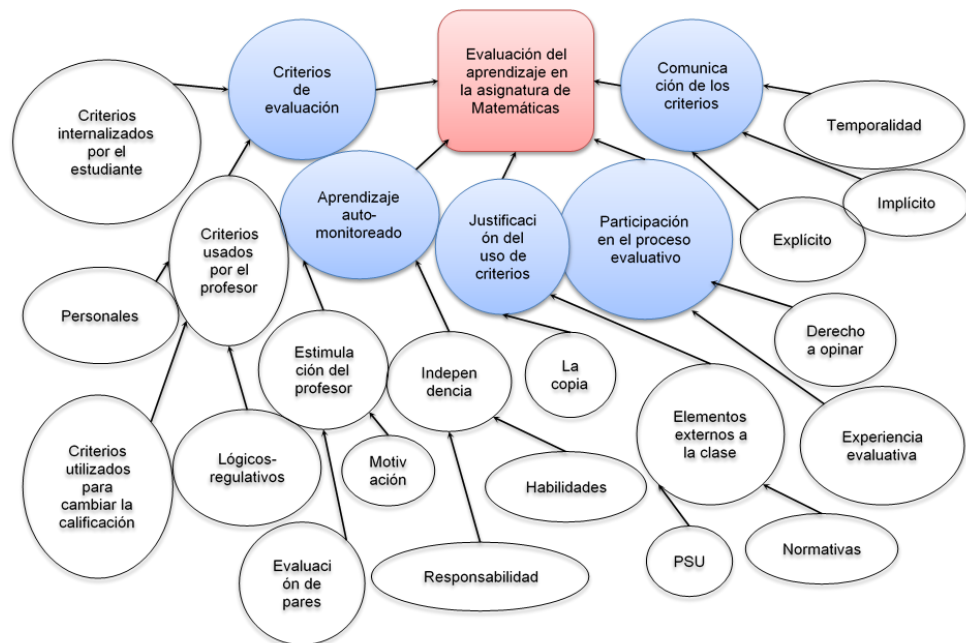


Figura 1. Árbol categorial.

Fuente: Elaboración propia. Apoyo NVivo.

El núcleo de este árbol categorial es *Evaluación del aprendizaje en la asignatura de Matemáticas*, y se utilizó color rojo para distinguirlo, color azul para las categorías analizadas y color blanco para todos los nodos libres, donde algunos ya comprenden un primer nivel de interpretación.

Es necesario explicar que cuando se plantea realizar una entrevista en profundidad que recorre ciertos ejes temáticos, indudablemente el investigador tendrá cierta incidencia en el posible discurso del estudiante ya que, a pesar de ser un instrumento no específico como una encuesta, va dirigida hacia elementos específicos, como los criterios de evaluación. Los nodos y categorías por lo tanto, emergen desde el discurso de los estudiantes que recorre los ejes temáticos que el investigador ha propuesto desde la literatura y experiencia, sin embargo, estas categorías engloban la mayor cantidad de referencias realizadas por los estudiantes e incluso algunos tópicos que surgen espontáneamente desde estos mismos.

Debido a motivos de extensión, las siguientes discusiones serán presentadas con pocas citas y no todas las referencias utilizadas en la investigación.

4.2. Criterios de evaluación

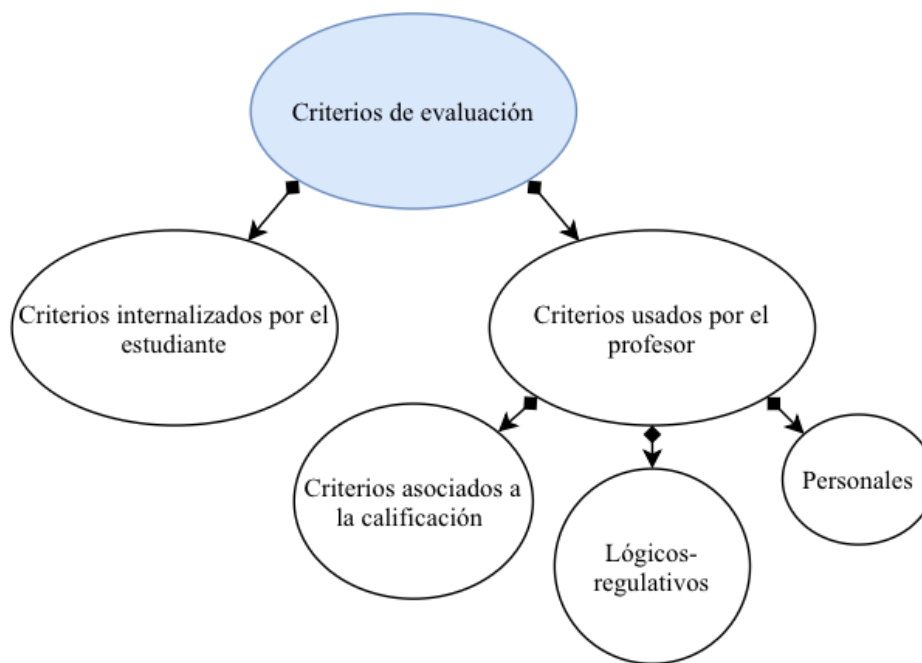


Figura 2. Árbol categorial 1.
Fuente: Elaboración propia.

Esta categoría presenta 11 fuentes (cada entrevista guardada en el programa) y un total de 245 referencias (nodos o categorías creadas a partir del discurso) dentro del programa NVivo. Cabe señalar que todos los estudiantes mencionaron al menos una cosa sobre esta categoría, ya que incluye a todas las fuentes y además es la categoría con mayor número de referencias dentro del programa.

Aquí aparecen los elementos identificados por los estudiantes que permiten juzgar la calidad de los trabajos, actuaciones o ejercicios en la asignatura de Matemáticas y que, además, se pueden utilizar para tomar decisiones calificativas. También se evidencian partes del discurso en donde los estudiantes señalan que el profesor utiliza un criterio para mejorar las calificaciones: esfuerzo, participación, asistencia, ejercicios resueltos. Por otra parte, también se logró identificar una clasificación de los criterios más utilizados por los estudiantes y sus docentes: criterios utilizados para cambiar la calificación, criterios personales y criterios lógicos-regulativos.

Una gran diferencia entre los establecimientos se observa cuando los estudiantes del establecimiento particular señalan que sus profesores no otorgan décimas extras para mejorar

la calificación; en cambio, en el establecimiento subvencionado los dos profesores otorgan décimas como una práctica habitual dentro del proceso de evaluación, con el fin de mejorar la calificación, como vemos en el siguiente extracto:

[Be: estudiante]

Be: Hay unos trabajos que hace en la clase anterior a la prueba, que son en uno puede ser en grupo (se permite grupal o individual), pero es lo mismo que una prueba (los trabajos), o sea la misma clase de ejercicios, que son más complicados de lo normal y todo y da unos cinco ejercicios más menos y al término de la clase hay que entregarlo y te da un puntaje en décimas, cada respuesta buena es una décima y cosas así.

(Entrevista 4)

Para Sadler (2009a), lo que hacen estos docentes se llama inflación de la nota, que se define como “un cambio ascendente en el promedio de la nota de los estudiantes en un periodo extendido de tiempo sin un correspondiente aumento en los logros de los estudiantes” (Rosovsky & Hartley, 2000, en Sadler, 2009a, 9).

Según Sadler (2009a), este tipo de comportamientos compromete la integridad de la nota, ya sea en su dimensión intrínseca o extrínseca. La extrínseca tiene relación con el uso que se les da a las notas, la medida en que se puede confiar y basarse en ellas para la toma de decisiones. La dimensión intrínseca o mérito, se refiere a cuán bien una nota representa lo que se supone que debe representar: por una parte, los alumnos merecen que su trabajo sea calificado estrictamente de acuerdo con su calidad, además, merecen que sus notas tengan valores que se puedan comparar entre los cursos sin exhibir calificación más difícil o fácil, por otro, las notas de los estudiantes deben mantener su valor en el tiempo para que puedan compararse con diferentes instituciones (Sadler, 2009a).

La inflación de notas se relaciona con los criterios personales que los estudiantes han logrado identificar, ya que aquí aparecen criterios como la participación, buen comportamiento y otras características que permiten al estudiante mejorar sus calificaciones. Estos criterios no tienen relación con aprendizajes en Matemáticas, al parecer responden a objetivos transversales dentro del currículo y están completamente asociados al establecimiento subvencionado, ya que para estos estudiantes es una situación común el hecho de inflar las notas mediante criterios personales, como el buen comportamiento e incluso llevarse bien con el profesor. Esto les permite obtener mejores calificaciones y lo valoran positivamente, como podemos apreciar en el fragmento:

aprendizaje alcanzado por el estudiante, por lo cual posee integridad cuando se acerca a la verdadera representación. Para ellos sería una falta en la integridad de la nota.

4.3. Aprendizaje auto-monitoreado

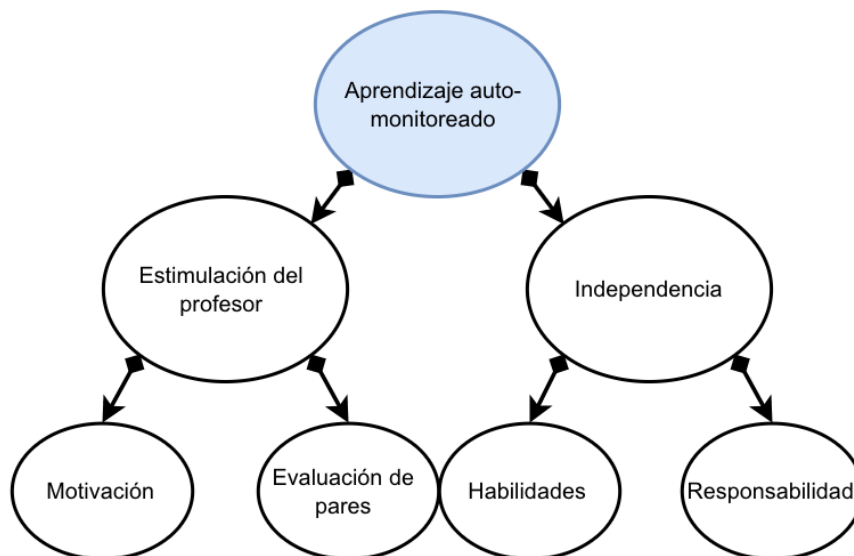


Figura 4. Árbol categoría 2.

Fuente: Elaboración propia.

Aquí encontramos 11 fuentes y un total de 90 referencias y se refiere a aquellas partes del discurso donde los estudiantes describen acciones o actividades que realizan en clases o fuera de ella, mostrando que pueden llevar un aprendizaje de calidad sin la presencia del profesor. Por otro lado, esta incluye todos los momentos donde el profesor coacciona el desarrollo del auto-monitoreo en sus estudiantes, de forma intencionada o involuntaria por parte del profesor.

Respecto del grado de independencia que poseen los estudiantes sobre su propio aprendizaje, debemos señalar que aquellos con mejor rendimiento en Matemáticas poseen un mayor control, a diferencia de aquellos con bajo rendimiento, quienes para lograr estudiar necesitan de compañía. Como plantean Jorba y Sanmartí (1993), cada persona posee un sistema personal de aprendizaje que se ha ido formando de manera autónoma, sin embargo, desde la literatura, entendemos que esta autonomía puede ser adquirida. Sadler (2009b), por su parte, explica que el ser auto-monitoreado implica que los estudiantes deben ser destetados de la dependencia continua en la retroalimentación externa, sin importar su origen o naturaleza. Sadler (1998, 2005), señala que el auto-monitoreo es una habilidad que puede ser adquirida y que incluso existen estudiantes que lograron ser auto-monitoreados sin considerarlo como un objetivo a alcanzar.

[C: estudiante]

C: Yo creo que igual es importante repasar como todos los días un poco, para... porque, así como que se va, va memorizando lo que tienes que hacer y así cuando llegas al ejercicio como que tienes que leerlo y plantear... y como entender bien lo que te preguntes, y ahí solo aplicas lo que aprendiste, lo que memorizaste, y ahí...

(Entrevista 8)

Estos estudiantes reciben una estimulación en retrospectiva por parte de sus docentes y, según Wiliam (2009), la retroalimentación más efectiva es aquella que focaliza su atención prospectiva más que retrospectivamente. Se menciona la retroalimentación en estos momentos, puesto que existen ciertas intervenciones con efectos positivos que tienen por objetivo aumentar la motivación hacia alguna tarea. Sobre esto, el autor señala que, si no se realiza de buena forma, podría problematizar el aprendizaje futuro, dado que el estudiante podría llegar a ser dependiente de una retroalimentación continua.

Los estudiantes con mejor rendimiento de ambos establecimientos describen situaciones en que el profesor los incentiva o mediante su propia iniciativa, ayudan a sus compañeros en clases o fuera de ellas. Algunos docentes, sobre todo del establecimiento subvencionado, intentan activar a los estudiantes con mejores rendimientos para aportar al aprendizaje de sus compañeros. No es seguro si es que los docentes pretendían en algún momento desarrollar estas habilidades, sin embargo y como señala Sadler (1998), el primer paso para lograr que un estudiante desarrolle la habilidad de auto-monitorearse, es comparar su trabajo con el de sus compañeros, sobre todo con un ejemplo de mayor calidad, para que así puedan mejorar su trabajo o mantener su calidad. Gran parte de los estudiantes menciona que el trabajar con sus compañeros les permite comparar sus resultados e identificar los errores, sin embargo, no es una práctica que el profesor determine, si no que en su mayoría lo hacen por motivación propia:

[E: entrevistador. Ao: estudiante]

Ao: Sí po, o por ejemplo con mis compañeros también.

E: tus compañeros te ayudan para...

Ao: Sí, además los comparo con ellos de repente...en qué me equivoqué y en qué me faltó.

(Entrevista 3)

4.4. Justificación del uso de criterios

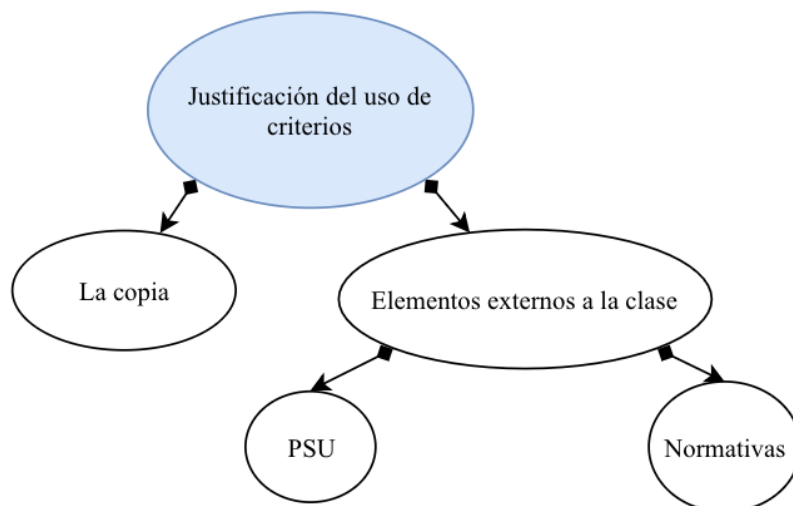


Figura 5. Árbol categoría 3.

Fuente: Elaboración propia.

Esta categoría posee 8 fuentes y 61 referencias dentro del programa. Los estudiantes justifican el uso de un criterio de evaluación, es decir, son las razones de los estudiantes por las cuales ciertas características son consideradas como criterios. Estas justificaciones no solo se manifiestan en torno a los criterios, si no que además pueden justificar el actuar del profesor y su diseño evaluativo.

“La copia”, hace referencia al discurso de los estudiantes donde mencionan el hecho de copiar como una justificación o razón para que el profesor utilice algún criterio, y así evitar esta copia. Este es un pensamiento reiterativo entre los estudiantes:

[C: estudiante]

C: Sí, por ejemplo, él (el profesor) nos pide el desarrollo, revisa bien siempre si desarrollamos los ejercicios porque no tenemos que, como llegar a la pregunta solo porque sí, porque sino él puede pensar que estamos copiando.

(Entrevista 8)

Manifiestan que el desarrollo es importante porque así pueden demostrar al profesor que no han copiado la respuesta. Este tipo de criterios tiene relación con los discutidos anteriormente, en donde revisamos que la importancia de utilizar criterios del tipo lógico o regulativo en el desarrollo de los ejercicios puede ser una forma de asegurar que el estudiante maneja en cierto

grado los contenidos, aunque el resultado final estuviese malo. Sin embargo, desde la percepción de los estudiantes, estos criterios tienen relación con una forma de control.

Los estudiantes reconocen que han visto a compañeros copiar durante las pruebas o intentar engañar al profesor mediante una escritura confusa para luego copiar el resultado correcto, algo similar expone Stonewater (2002) respecto a la escritura. Por otra parte, los estudiantes señalan que sus profesores les han indicado que en pruebas de selección múltiple deben realizar el desarrollo.

Por último, la Prueba de Selección Universitaria (PSU) y las normativas dispuestas en el establecimiento (control del proceso evaluativo), también influyen en las justificaciones de los estudiantes, pero como elementos externos a la clase, tal como se refleja en esta cita:

[A: estudiante]

A: Eh... no, encuentro que es lo que tiene que hacer, porque bueno, en la PSU no nos van a estar dando puntos por desarrollo.

(Entrevista 7)

Les parece aceptable que su profesor no revise el desarrollo ya que en la PSU tampoco lo hacen. Sobre esto, una alumna hace referencia al formato de pruebas que deben utilizar los profesores, lo que indirectamente tiene relación con la PSU, ya que en gran parte de los establecimientos escolares en Chile se ha modificado el formato de sus procedimientos evaluativos utilizando en mayor medida pruebas de selección múltiple para preparar a sus estudiantes, un ejercicio que solo tensiona a los docentes y los climas escolares (Casassus, 2009).

Algunos estudiantes del establecimiento particular refieren a la dirección del colegio y al jefe de UTP (Unidad Técnico-Pedagógica) como factores que influyen en el proceso de evaluación. Expresan, por ejemplo, que el jefe de UTP o coordinador pedagógico es el culpable de no poder optar a décimas extras. Esto quiere decir que ciertos tipos de criterios asociados a la calificación son regulados por autoridades externas a la clase y, por lo tanto, el profesor no tiene total autonomía para utilizar los criterios que a él le parezcan pertinentes. Como señala Contreras (2014), la evaluación escolar es un proceso complejo en donde confluyen una variedad de aspectos, algunos de ellos están asociados a la escuela y tienen relación con normas, regulaciones, criterios, procedimientos, exigencias administrativas y condiciones de trabajo en el aula.

4.5. Participación en el proceso evaluativo



Figura 6. Árbol categoría 4.

Fuente: Elaboración propia.

Encontramos un total de 10 fuentes y 31 referencias y se evidencia el grado y tipo de participación que tienen los estudiantes en los distintos momentos del proceso evaluativo, como en la construcción de criterios de evaluación y también en la frecuencia con que los estudiantes se nutren de experiencia evaluativa en los momentos de la clase.

En el siguiente fragmento un estudiante describe un momento vivido por su curso, en donde los compañeros propusieron al profesor cambiar el formato de la prueba:

[A: estudiante. E: entrevistador]

A: Eh, sí porque él... hace poco, fue en la última prueba no más, que dijo que iba a ser con alternativas y que la última iba a ser con desarrollo.

E: Ya.

A: Y en el momento fue como que no, que el desarrollo no porque igual se equivocaban caleta, pero iba a ser la prueba de puro desarrollo y le dijo: "ya con alternativas".

(Entrevista 1)

Los estudiantes de ambos establecimientos no tienen una participación significativa y responden de manera similar al ejemplo expuesto. Una práctica que puede ser común en las clases de Matemáticas tiene relación con la resolución de ejercicios o problemas en la pizarra. Esto es una forma de participar en la clase pero que no apunta a una experiencia evaluativa en su totalidad, sin embargo, que los estudiantes puedan juzgar las respuestas de sus compañeros

o que el estudiante deba explicar las razones sobre la forma en que realizó el ejercicio, promueve a los estudiantes como recursos de enseñanza para otros alumnos (Wiliam, 2009).

Según Wiliam y Thompson (2007), para que la evaluación para el aprendizaje se lleve a cabo, debe ser construida a partir de cinco estrategias. Dos apuntan directamente al estudiante como protagonista y participante del proceso de evaluación: una es la autoevaluación y la otra es la evaluación de pares. Lamentablemente no evidenciamos procesos de autoevaluación en esta investigación. Por otra parte, los estudiantes valoran positivamente los momentos en donde comparten con sus compañeros y pueden retroalimentarse entre sí.

4.6. Comunicación de criterios de evaluación

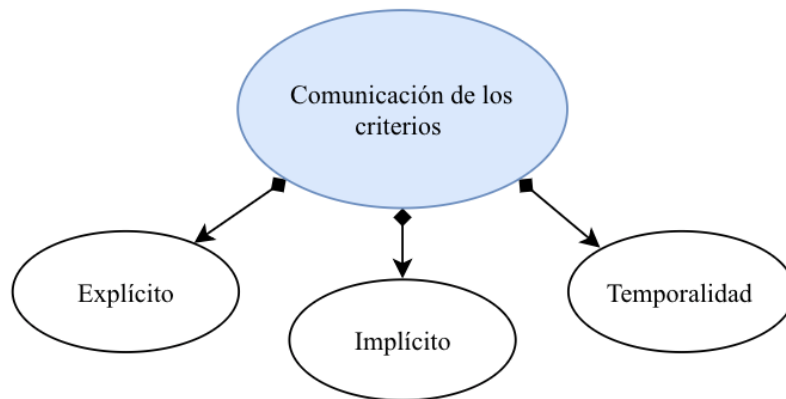


Figura 7. Árbol categoría 5.
Fuente: Elaboración propia.

Un total de 169 referencias, obtenidas desde 11 fuentes. Considera todas las acciones realizadas por el profesor o los estudiantes en las cuales se están compartiendo criterios de evaluación. Se han evidenciado dos formas de comunicación: explícita e implícita. La primera hace referencia a la comunicación de criterios predefinidos por parte del profesor, es decir, cuando el docente de manera consciente juzga la calidad de los trabajos mediante criterios establecidos o señala a los estudiantes, de manera directa, cuáles son los criterios por los que serán o están siendo evaluados. La comunicación implícita hace referencia a la comunicación involuntaria de criterios, aquellos criterios que sin ser comunicados de forma directa logran ser identificados por los estudiantes.

Al parecer, la comunicación de criterios se realiza de manera más precisa durante las clases anteriores a la prueba, en que el profesor comparte criterios más exactos o mejor definidos sobre lo que desea que sus estudiantes logren en la prueba. Algo importante que se debe

mencionar es la disposición de los estudiantes durante la clase anterior a la prueba, percibida por ellos mismos como un momento en que están más interesados y con una mayor participación:

[E: entrevistador. A: estudiante]

E: Ah ya, entiendo. Ya, enfocándonos nuevamente en Matemáticas, ¿las clases antes de las pruebas cómo son? Así la clase justo antes de la prueba o un poquito antes de la prueba.

A: Ah, todos están atentos po, están todos preguntando, la clase siempre el profesor hace siempre una clase antes de la prueba en donde solo hay preguntas, no hay materia.

(Entrevista 7)

En general la temporalidad, según lo expresado por los estudiantes, cumple con la secuencia mostrada en la figura 8, en donde lo café son momentos de comunicación moderada de criterios, en rojo se encuentra el punto más alto y en celeste el de menor comunicación:

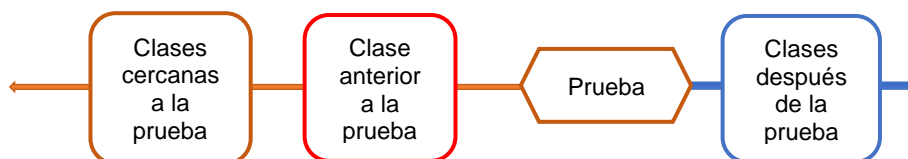


Figura 8. Línea temporal de la comunicación de criterios.

Fuente: Elaboración propia.

5. Discusión y conclusiones

Esta investigación pretendió caracterizar las interpretaciones que los estudiantes construyen sobre los criterios de evaluación comunicados en Matemáticas. Específicamente se centró en aquellos aspectos que los estudiantes consideran como criterios de evaluación y también en su percepción acerca del grado de participación en la construcción de criterios.

La mayor parte de los criterios identificados y usados por los estudiantes son del tipo lógicos y personales, sobre todo aquellos que están relacionados con la calificación. Los primeros están asociados a cadenas de razonamiento válidas cuyo resultado es replicable (Sadler, 1983), y se consideran buenos o errados, no existiendo un grado intermedio de calidad; por ejemplo, realizar erróneamente operaciones básicas como suma y resta, obteniendo malo el resultado. Los criterios personales tienen que ver con características de comportamiento o buena relación entre estudiante y profesor (Wang & Cai, 2018), los que generan la inflación de la nota. Los

estudiantes del establecimiento subvencionado reconocen el uso de estos criterios de evaluación por parte de sus profesores como una práctica frecuente, ya que les otorgan décimas para mejorar las calificaciones finales, práctica que agradecen. En el establecimiento particular, en cambio, aunque los estudiantes reconocen que estos aspectos podrían ser criterios, consideran que otorgar ciertas décimas por realizar trabajos es un reconocimiento al esfuerzo demostrado, señalan que sus profesores no usan estos criterios en sus evaluaciones. Atribuyen esta situación a regulaciones y normas de su colegio, y a otros factores externos a este, como la presión de la Prueba de Selección Universitaria (PSU), en la cual dichos criterios no se usan.

Por otra parte, aquellos estudiantes con mejor rendimiento poseen mayor claridad de los elementos con los cuales son juzgadas sus actuaciones, tienen mayor capacidad de predecir futuros resultados en sus evaluaciones y, además, cada uno de ellos es parte del proceso evaluativo, ayudando a sus compañeros con más dificultades. Este hallazgo corrobora lo que señala la teoría, sobre cómo el auto-monitoreo y la internalización de criterios sirven para mejorar el aprendizaje de los estudiantes.

Los estudiantes con menor rendimiento demuestran menor preocupación por estudiar, a diferencia de los de mayor rendimiento, sin embargo, ambos reconocen la necesidad de una continua retroalimentación por parte del profesor o de sus compañeros para sentirse seguros sobre su aprendizaje. Pueden identificar los mismos criterios que sus compañeros de mejor rendimiento, pero sus valoraciones suelen ser distintas, ya que muestran mayor preocupación por los criterios asociados a la calificación.

Los estudiantes de ambos establecimientos consideran que la razón de utilizar criterios para corregir el desarrollo de un problema (y no solo fijarse en el resultado del ejercicio) es para prevenir la copia. Esta situación tiene relación con la forma tradicional de pensar, en que se considera a la evaluación como un proceso técnico de verificación de resultados de aprendizaje (Contreras & Prieto, 2008).

Una de las grandes diferencias entre los estudiantes del establecimiento subvencionado en comparación al particular, son las referencias hacia la Prueba de Selección Universitaria (PSU). Los estudiantes del colegio particular realizan muchas referencias hacia la PSU, en cambio los del establecimiento subvencionado jamás lo hicieron. Esto en parte se podría explicar por la fuerte segregación socioeconómica que representa la PSU (Revista Docencia, 2009) y la diferencia de capital cultural entre ambos establecimientos.

La teoría (Popham, 2013; Sadler, 1983, 1989, 2009a, 2009b; Wiliam, 2009) señala que los estudiantes deben adquirir experiencia evaluativa de forma tal que comprendan lo que se espera de ellos. No obstante, en este caso la participación de los estudiantes se reduce a reclamos sobre las notas o a dar sugerencias sobre el formato de la prueba (selección múltiple). No hay espacios o momentos en que los estudiantes puedan participar en la construcción de criterios de evaluación, aportando con sus ideas y experiencias. Si bien puede ser difícil que aporten criterios del tipo lógico, ya que existe un corpus de conocimiento que está validado por una comunidad científica en matemática educativa, los estudiantes podrían opinar sobre las formas de resolución que utilizarán para resolver ejercicios, ya que, para resolver una ecuación, se podría realizar un proceso mental, un desarrollo algebraico, una estimación, etc. Por otra parte, podrían opinar con propiedad respecto de cuáles criterios personales son más pertinentes para considerar.

En esta investigación, los estudiantes valoran positivamente la evaluación cuando esta se encuentra asociada a calificaciones, especialmente si es para mejorarlas, situación que da cuenta de una perspectiva instrumental sobre dicho proceso.

6. Referencias Bibliográficas

- Barrie, S., Brew, A., & McCulloch, M. (1999, septiembre). Qualitatively different conceptions of criteria used to assess student learning. *Australian Association for Research in Education (AARE) Conference*. Conferencia en Melbourne, Australia. Recuperado desde <http://www.aare.edu.au/99pap/bre99209.htm>
- Bell, A., Mladenovic, R., & Price, M. (2013). Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assessment & Evaluation in Higher Education*, 38(7), 769-788. doi:10.1080/02602938.2012.714738
- Cantoral, R., Reyes-Gasperini, D., & Montiel, G. (2014). Socioepistemología, Matemáticas y Realidad. *Revista Latinoamericana de Etnomatemática*, 7(3), 91-116.
- Casassus, J. (2009): Actuales escenarios educativos: una mirada desde la democratización de la educación, en Otra escuela es posible, la Educación puede cambiar, Documento de Trabajo, Colegio de Profesores de Chile A.G., Santiago, Chile.
- Conde, F. (2009). *Análisis sociológico del sistema de discursos* (1ª ed.). Madrid: Centro de investigaciones sociológicas.
- Contreras, G. (2014). La Retroalimentación en Educación Superior. Conceptos, principios y estrategias para la práctica. En J. M. Garrido, A. Arenas, & D. Contreras (Coords.). *Mejorando las prácticas de evaluación de los aprendizajes en la docencia universitaria. Análisis y experiencias* (pp. 161-175). Valparaíso: Ediciones Universitarias de Valparaíso.
- Davinroy, K. (1995). "How Does My Teacher Know What I Know?" *Third Graders' Perceptions of Math, Reading, and Assessment*. National Center for Research on Evaluation, Standards, and Student Testing, U.S Department of Education. Los Angeles: The Regents of the University of California.
- Ferrara, S., Goldberg, G., & McTighe, J. (1995, abril). Ways in Which teachers Communicate Targets, Criteria, and Standards for Performance to Their

- Students. *American Educational Research Association Meeting*. Conferencia en San Francisco, California. Recuperado desde <https://marces.org/mdarch/pdf/M032003.pdf>
- Jorba, J., & Sanmartí, N. (1993). La función pedagógica de la evaluación. *Aula de Innovación Educativa*, 20, 20-30.
- Lukas, J., & Santiago, K. (2004). Evaluación de centros escolares de educación secundaria del País Vasco. *REDIE. Revista Electrónica de Investigación Educativa*, 6(2), 1-23.
- O'Donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: a nested hierarchy of approaches. *Teaching in Higher Education*, 13(2), 205-217. doi:10.1080/13562510801923344
- Organización para la Cooperación y el Desarrollo Económicos. (2015). *Mathematics performance (PISA) (indicator)*. doi:10.1787/04711c74-en
- Osses, S., Sánchez, I., & Ibáñez, F. (2006). Investigación cualitativa en educación: hacia la generación de teoría a través del proceso analítico. *Estudios pedagógicos*, 32(1), 119-133. doi:10.4067/S0718-07052006000100007
- Picaroni, B. (2011). Prácticas de evaluación en las aulas de primaria en ocho países de Latinoamérica. *IV Congreso Nacional y III Encuentro Internacional de Estudios Comparados en Educación. ¿Hacia dónde va la Educación en la Argentina y en América Latina? Construyendo una nueva agenda*. Congreso en Buenos Aires, Argentina. Recuperado desde <http://www.saece.com.ar/docs/congreso4/trab77.pdf>
- Picaroni, B., & Lourerio, G. (2010). ¿Qué Matemáticas se Enseña en Aulas de sexto año de primaria en Escuelas de Latinoamérica? *Páginas en Educación*, 3(3), 29-60. doi:10.22235/pe.v3i1.657
- Popham, W. J. (2013). *Evaluación trans-formativa. El poder transformador de la evaluación formativa*. Madrid: Narcea.
- Price, M. (2005). Assessment standards: the role of communities of practice and the scholarship of assessment'. *Assessment & Evaluation in Higher Education*, 30(3), 215-230. doi:10.1080/02602930500063793

- Revista Docencia. (2009). Estandarización educativa en Chile: un peligroso hábito. *Autor*, (38), 4-17. Recuperado desde <http://mapeal.cippec.org/wp-content/uploads/2014/05/Estandarización-educativa-en-Chile.-Un-peligroso-habito..pdf>
- Roco, R. (2010, septiembre-octubre). Caracterización de los establecimientos educacionales en Chile: la necesidad de nuevas consideraciones. *Primer Congreso Interdisciplinario de Investigación en Educación, CIAE, Universidad de Chile y CEPPE, Pontificia Universidad Católica de Chile*. Congreso en Santiago, Chile.
- Sadler, D. R. (1983). Evaluation and the improvement of academic learning, *Journal of Higher Education*, 54(1), 60-79. doi:10.2307/1981645
- Sadler, D. R. (1985). The Origin and Function of Evaluative Criteria. *Educational Theory*, 35(2), 285-297. doi:10.1111/j.1741-5446.1985.00285.x
- Sadler, D. R. (1989). Formative assessment and the design of instructional assessment. *Instructional science*, 18(2), 119-144.
- Sadler, D. R. (1998, julio). Letting students into the secret: Further steps in making criteria and standards work to improve learning. *Annual Conference for State Review Panels and District Review Panel Chairs, Bardon Professional Centre*. Conferencia en Brisbane, Australia.
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Studies in Higher Education*, 30(2), 175-194. doi:10.1080/0260293042000264262
- Sadler, D. R. (2009a). Grade integrity and the representation of academic achievement. *Studies in Higher Education*, 34(7), 807-826. doi:10.1080/03075070802706553
- Sadler, D. R. (2009b). Transforming hoistic assessment and grading into a vehicle for complex learning. En G. Joughin (Ed.). *Assessment, learning and judgement in higher education* (45-63). Dordrecht, The Netherlands: Springer.
- Sadler, D. R. (2014). The futility of attempting to codify academic achievement standards. *High Educ*, 67, 273-288. doi:10.1007/s10734-013-9649-1

- Sadler, D. R. (2015). Backwards assessment explanations: Implications for teaching and assessment practice. En D. Lebler, S. Harrison, & G. Carey (Eds.). *Assessment in Music: From Policy to Practice* (9-19). Cham, Switzerland: Springer.
- Santos, L., & Semana, S. (2015). Developing mathematics written communication through expository writing supported by assessment strategies. *Educational Studies in Mathematics*, 88(1), 65-87. doi:10.1007/s10649-014-9557-z
- Stonewater, J. (2002). The Mathematics Writer's Checklist: The Development of a Preliminary Assessment Tool for Writing in Mathematics. *School Science and Mathematics*, 102(7), 324-334. <https://doi.org/10.1111/j.1949-8594.2002.tb18216.x>
- Taylor, S. and Bogdan, R. (1987). *Introducción a los métodos cualitativos de investigación*. 3rd ed. Barcelona: Paidós, pp.1-12.
- Wang, N., & Cai, J. (2018). An Investigation of How Teachers Score Constructed-Response Mathematics Assessment Tasks. *Journal of Research in Education*, 28(1), 1-29.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development ASCD.
- William, D. (2009). Una síntesis integradora de la investigación e implicancias para una nueva teoría de la evaluación formativa. *Archivos de Ciencias de la Educación*, 3(3), 15-44.
- William, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? En C. A. Dwyer (Ed.). *The future of assessment: Shaping teaching and learning* (53-82). Mahwah, NJ: Lawrence Erlbaum Associates.