



INSTITUTO DE  
INVESTIGACIÓN  
Y DESARROLLO  
EDUCATIVO

# Universidad Autónoma de Baja California

## Instituto de Investigación y Desarrollo Educativo



### **“DESARROLLO Y EVIDENCIAS DE VALIDEZ DE UNA RÚBRICA PARA LA AUTOEVALUACIÓN DE PRÁCTICAS ESCOLARES EN ESCUELAS SECUNDARIAS”**

**TESIS**

Que para obtener el grado de

**MAESTRA EN CIENCIAS EDUCATIVAS**

**Presenta**

**Jennifer Gómez Gloria**

Ensenada B. C. México, noviembre de 2016



**“DESARROLLO Y EVIDENCIAS DE VALIDEZ DE UNA RÚBRICA PARA LA AUTOEVALUACIÓN DE PRÁCTICAS ESCOLARES EN ESCUELAS SECUNDARIAS”**

**TESIS**

Que para obtener el grado de  
**MAESTRA EN CIENCIAS EDUCATIVAS**

Presenta

**Jennifer Gómez Gloria**

APROBADO POR:

**Dra. Alicia Alelí Chaparro Caso-López**

Directora de tesis

**Dr. Joaquín Caso Niebla**

Sinodal

**Dr. Luis Lizasoain Hernández**

Sinodal

**Dra. Karla María Díaz López**

Sinodal



Ensenada, B.C. a Noviembre de 2016



ASUNTO: Voto aprobatorio sobre trabajo de tesis de grado de Maestría.

**Mtro. Julio César Cano Gutierrez**

**Subdirector del Instituto de Investigación y Desarrollo Educativo  
Presente.**

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. JENNIFER GÓMEZ GLORIA** para poder presentar la defensa de su examen y obtener el grado de Maestría en Ciencias Educativas, me permito comunicarle que he dado mi VOTO APROBATORIO, sobre su trabajo intitulado:

***“DESARROLLO Y EVIDENCIAS DE VALIDEZ DE UNA RÚBRICA PARA LA AUTOEVALUACIÓN DE PRÁCTICAS ESCOLARES EN ESCUELAS SECUNDARIAS”.***

Esperando reciba el presente de conformidad, quedo de Usted.

Atentamente

Una firma manuscrita en tinta azul que parece decir "Alicia", escrita sobre una línea horizontal.

Dra. Alicia Alelí Chaparro Caso-López

Ensenada, B.C. a Noviembre de 2016



ASUNTO: Voto aprobatorio sobre trabajo de tesis de grado de Maestría.

**Alicia Alelí Chaparro Caso-López**

**Coordinador(a) de la Maestría en Ciencias Educativas  
Presente.**

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. JENNIFER GÓMEZ GLORIA** para poder presentar la defensa de su examen y obtener el grado de Maestría en Ciencias Educativas, me permito comunicarle que he dado mi VOTO APROBATORIO, sobre su trabajo intitulado:

***“DESARROLLO Y EVIDENCIAS DE VALIDEZ DE UNA RÚBRICA PARA LA AUTOEVALUACIÓN DE PRÁCTICAS ESCOLARES EN ESCUELAS SECUNDARIAS”.***

Esperando reciba el presente de conformidad, quedo de Usted.

Atentamente

Una firma manuscrita en tinta azul que parece decir "J. Caso Niebla".

---

Dr. Joaquín Caso Niebla

Ensenada, B.C. a Noviembre de 2016

ASUNTO: Voto aprobatorio sobre trabajo  
de tesis de grado de Maestría.



**Alicia Alelí Chaparro Caso-López**

**Coordinador(a) de la Maestría en Ciencias Educativas  
Presente.**

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. JENNIFER GÓMEZ GLORIA** para poder presentar la defensa de su examen y obtener el grado de Maestría en Ciencias Educativas, me permito comunicarle que he dado mi **VOTO APROBATORIO**, sobre su trabajo intitulado:

***“DESARROLLO Y EVIDENCIAS DE VALIDEZ DE UNA RÚBRICA PARA LA  
AUTOEVALUACIÓN DE PRÁCTICAS ESCOLARES EN ESCUELAS  
SECUNDARIAS”.***

Esperando reciba el presente de conformidad, quedo de Usted.

Atentamente

Una firma manuscrita en tinta azul que parece decir "Luis Lizasoain".

---

Dr. Luis Lizasoain Hernández

Ensenada, B.C. a Noviembre de 2016



ASUNTO: Voto aprobatorio sobre trabajo de tesis de grado de Maestría.

**Alicia Alelí Chaparro Caso-López**

**Coordinador(a) de la Maestría en Ciencias Educativas  
Presente.**

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la C. **JENNIFER GÓMEZ GLORIA** para poder presentar la defensa de su examen y obtener el grado de Maestría en Ciencias Educativas, me permito comunicarle que he dado mi VOTO APROBATORIO, sobre su trabajo intitulado:

***“DESARROLLO Y EVIDENCIAS DE VALIDEZ DE UNA RÚBRICA PARA LA AUTOEVALUACIÓN DE PRÁCTICAS ESCOLARES EN ESCUELAS SECUNDARIAS”.***

Esperando reciba el presente de conformidad, quedo de Usted.

Atentamente

Una firma manuscrita en azul que parece decir "Karla" y "RC2" a la derecha.

---

Dra. Karla María Díaz López

# *Dedicatoria*

*Con amor:*

*A mis padres, por enseñarme a confiar en mí, por esforzarse en brindarme siempre lo mejor y por el gran amor que me demuestran.*

*A mi hermana, porque con su ejemplo me motiva a superarme.*

*A mi hermano, que desde el cielo cuida de mí.*

# *Agradecimientos*

*Gracias a mi familia por su apoyo, confianza y amor.*

*Gracias a Luis Alberto por su tolerancia, paciencia y apoyo incondicional.*

*Gracias a mi compañera y amiga, la Mtra. Maribel Sandoval por compartir los momentos de estrés, cansancio y alegría que vivimos para cumplir nuestra meta.*

*A mi directora de tesis, la Dra. Alicia Chaparro, por su plena confianza, dedicación y buenos consejos.*

*Gracias a mis compañeros y amigos, la Mtra. Genoveva Gutiérrez y al Lic. Noé Mora por su tiempo y apoyo en todo momento.*

*Gracias a la Dra. Karla María Díaz, al Dr. Joaquín Caso y al Dr. Luis Lizasoain quienes como miembros de mi comité de tesis, realizaron valiosas aportaciones.*

*Gracias a mis maestros y compañeros por sus enseñanzas y valiosa realimentación.*

*Gracias al Instituto de Investigación y Desarrollo Educativo, por siempre brindarme lo necesario. En especial a Yesica Espinoza por estar continuamente disponible para atender dudas y necesidades.*

*Gracias a la Universidad Autónoma de Baja California, mi casa de estudios.*

*Gracias al Consejo Nacional de Ciencia y Tecnología, por el apoyo económico otorgado para dedicar tiempo completo a mis estudios.*

# *Resumen*

## RESUMEN

Las investigaciones sobre Eficacia Escolar y Mejora Escolar han aportado información sobre los factores escolares presentes en escuelas con buenos resultados académicos. De esta manera se ha comprobado que las acciones llevadas a cabo por la comunidad escolar (prácticas escolares) inciden positivamente en los procesos educativos. La autoevaluación de las prácticas escolares permite generar conciencia a la comunidad escolar sobre las cuestiones que se deben mejorar para el desarrollo integral de sus estudiantes.

En México, el desarrollo de instrumentos para dicha autoevaluación es incipiente, asimismo, existe necesidad de contar con instrumentos de fácil manejo y con evidencias de validez. Contar con un instrumento de autoevaluación para identificar dichos factores en la práctica escolar contribuiría a la mejora de los procesos educativos. Al respecto, la rúbrica analítica es un instrumento que permite valorar constructos de forma integral. No obstante, existe poca documentación sobre el proceso para elaborarlas, por lo que es necesario procedimientos claros y específicos para sistematizar su elaboración.

El objetivo de la presente investigación fue desarrollar una rúbrica analítica para la autoevaluación de prácticas escolares en escuelas secundarias; misma que se derivó de los hallazgos reportados en el estudio de *Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California* (Chaparro et al., 2015). Estos hallazgos fueron una serie de prácticas escolares, identificadas en escuelas de alto valor añadido, clasificadas con base en factores de Eficacia Escolar y Mejora Escolar.

El procedimiento para la elaboración de la rúbrica consistió dos etapas. La primera con la finalidad de valorar la eficacia de las prácticas escolares mediante la opinión de nueve expertos del Sistema Educativo Estatal. La segunda, con el propósito de establecer los criterios de la rúbrica y estimar el índice de validez de contenido (IVC) de Lawshe (1975) mediante la técnica de jueceo por 12 docentes de educación básica. De esta manera se eliminaron los criterios con un IVC inferior a 0.55 y se conservaron aquellos por encima del mismo.

De acuerdo con los resultados, se obtuvo una rúbrica analítica con 119 criterios que permiten autoevaluar la presencia de prácticas escolares eficaces en escuelas secundarias en 11 categorías: (a) Supervisión, (b) Liderazgo directivo (c) Gestión de recursos, (d) Planeación (e) Formación continua del personal, (f) Evaluación del personal escolar (g) Evaluación de los aprendizajes (h) Metodología didáctica y estrategias de enseñanza (i) Atención a la diversidad (j) Convivencia y clima escolar, y (k) Relación de la escuela con los padres de familia.

Se concluyó, con base a la revisión de la literatura que realizó, que los modelos de Eficacia Escolar y Mejora Escolar son marcos de referencia adecuados para el desarrollo de la rúbrica, puesto que incluyen elementos presentes en la dinámica escolar de escuelas eficaces y de estrategias que promueven la mejora. Asimismo, el complemento de éste marco con evidencias empíricas de prácticas escolares eficaces fortalece la validez de contenido de la rúbrica.

**Palabras clave:** rúbrica analítica; autoevaluación; práctica escolar; educación secundaria

## TABLA DE CONTENIDO

<b>1. Introducción</b> .....	<b>1</b>
1.1. Planteamiento del problema.....	2
1.2. Objetivos.....	5
1.3. Justificación.....	5
<b>2. Revisión de la literatura</b> .....	<b>8</b>
2.1. Eficacia Escolar y Mejora Escolar.....	9
2.2. Buenas prácticas escolares asociadas a los factores de Eficacia Escolar y Mejora Escolar.....	14
2.2.1. Liderazgo escolar.....	15
2.2.2. Gestión y organización.....	16
2.2.3. Recursos.....	17
2.3. Evaluación y autoevaluación como estrategia de mejora.....	21
2.4. Rúbricas.....	25
2.4.1. Tipos.....	28
2.4.2. Procedimientos para su elaboración.....	29
2.5. Validez.....	41
2.5.1. Validez de contenido.....	45
2.5.2. Validez de constructo.....	54
2.5.4. Validez convergente.....	55
<b>3. Contexto de la investigación</b> .....	<b>57</b>
3.1. Investigación del País Vasco: Caracterización y buenas prácticas en los centros escolares de alto valor añadido.....	58
3.2. Investigación mexicana: Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California.....	59
3.2.1. Procedimiento para la selección de las escuelas participantes.....	60
3.2.2. Procedimiento para el diseño y aplicación de entrevistas.....	62
3.2.3. Procedimiento para la selección de las prácticas escolares.....	64
<b>4. Etapas del estudio</b> .....	<b>66</b>
4.1. Etapa 1. Valoración y re-categorización de las prácticas escolares.....	67
4.2.1. Objetivos.....	67

4.2.2. Participantes. ....	67
4.2.3 Materiales.....	68
4.2.4 Procedimiento. ....	68
4.2.5. Análisis de datos. ....	69
Figura 3. Procedimiento para el análisis de contenido.....	70
4.2.6. Resultados. ....	70
4.2. Etapa 2. Integración y validez de contenido de la rúbrica.....	72
4.3.1. Objetivos .....	72
4.3.2. Participantes .....	72
4.3.3. Materiales.....	73
4.3.4. Procedimiento .....	74
4.3.5. Análisis de los datos .....	74
4.3.6. Resultados .....	75
<b>5. Discusión.....</b>	<b>80</b>
5.1. Criterios de la rúbrica analítica. ....	85
5.2. Diseño de los niveles de valoración de la rúbrica analítica.....	92
5.3. Evidencias de validez de contenido.....	93
<b>6. Conclusiones .....</b>	<b>95</b>
<b>7. Limitaciones y sugerencias.....</b>	<b>99</b>
<b>Referencias.....</b>	<b>101</b>
<b>Apéndices .....</b>	<b>116</b>
Apéndice A. Listado de prácticas escolares .....	116
Apéndice B. Captura de valoraciones de los jueces.....	116
Apéndice C. Modas y porcentajes .....	116
Apéndice D. Cuestionario para validación .....	116
Apéndice E. Captura de valoraciones de los jueces (pertinencia de los niveles y de las prácticas.....	116
Apéndice F. Criterios conservados.....	116

## ÍNDICE DE TABLAS

Tabla 1. Investigaciones sobre evidencias de validez de rúbricas.....	43
Tabla 2. Informantes clave del proyecto Caracterización de las buenas prácticas en escuelas secundarias de alto valor añadido de Baja California .....	63
Tabla 3. Total de prácticas por categoría .....	65
Tabla 4. Características de los participantes.....	67
Tabla 5. Criterios para la valoración de las prácticas escolares.....	68
Tabla 6. Total de prácticas por categoría .....	71
Tabla 7. Integración y validez de contenido de la rúbrica.....	72
Tabla 8. Características de los participantes.....	75
Tabla 9 . IVC de niveles de valoración de la rúbrica .....	76
Tabla 10. Criterios eliminados a partir del análisis .....	78

## ÍNDICE DE FIGURAS

<i>Figura 1.</i> Procedimiento para el análisis de contenido de las prácticas escolares.....	70
<i>Figura 2.</i> Formato para valorar la pertinencia de los niveles.....	73
<i>Figura 3.</i> Ejemplo del formato para validar la pertinencia de las prácticas escolares. ...	74
<i>Figura 4.</i> Valores mínimos de Lawshe .....	75

# *1. Introducción*

### 1.1. Planteamiento del problema

La evaluación presenta diversos significados, algunos referidos a la elaboración de juicios sobre la calidad, otros simplemente se refieren a los resultados obtenidos y en ocasiones se considera como un instrumento que permite realizar mejoras e innovaciones (Moreno, 2009). Para el caso de la evaluación educativa, el propósito general es recopilar información mediante procedimientos sistemáticos para realizar juicios sobre la calidad del proceso educativo (García, 2010). De manera que a través de la evaluación es posible verificar el cumplimiento de los programas establecidos por el sistema escolar, valorar el uso que se le da a los resultados de las evaluaciones de los estudiantes (Román, 2011) y estimar de qué manera influyen algunos factores en la calidad de la educación.

En este sentido, se ha demostrado que la comunidad escolar, las instalaciones y el contexto, son factores que influyen significativamente en el desarrollo del proceso educativo (Bellón, Cantero y Ziemer, 2012) y son objeto de evaluación en diversas investigaciones. De manera particular, las prácticas de gestión escolar, los procesos de enseñanza-aprendizaje, los vínculos con la comunidad y la familia, la gestión de la convivencia y el clima escolar, así como el uso que se le da a los recursos, son factores que intervienen en la calidad de la educación, no obstante, son pocas las iniciativas dirigidas a incluir, en la práctica escolar habitual, evaluaciones que integren estos ámbitos (Murillo y Román, 2010).

En México, las prácticas de evaluación educativa se dirigen, mayormente, a evaluar el rendimiento académico de los estudiantes, con ello se aporta información

general del desempeño del alumno, sin embargo, no proporciona información suficiente sobre el contexto en el que funcionan las escuelas (Santiago, McGregor, Nusche, Ravela y Toledo, 2014). En este sentido, la autoevaluación se considera un tipo de evaluación que permite valorar, reflexionar y analizar las propias acciones con el propósito de desarrollar planes para la mejora escolar (Grau y Álvarez ,2005). Asimismo, Murillo (2008) señaló que la autoevaluación escolar influye de manera significativa en el rendimiento académico ya que provee herramientas que orientan la labor de los directivos para transformar sus prácticas escolares (Secretaría de Educación Básica del Estado de Veracruz, 2010) y para fomentar la calidad, la equidad y la eficiencia de la educación (Lizasoain y Angulo, 2014; Schleicher y Zoido, 2013).

Si bien los sistemas educativos estatales mexicanos, proveen a las escuelas materiales e instrumentos de apoyo para fomentar la autoevaluación, no existe un enfoque sistemático y establecido para ello (Santiago et al., 2014). Con relación a lo anterior, Grau y Álvarez (2005) mencionaron que es importante establecer métodos de evaluación en las escuelas ya que permiten fundamentar las decisiones de las autoridades educativas y proporcionan información de la escuela y de los docentes sobre su propio funcionamiento ya que con esto se aporta a la mejora de sus prácticas educativas. Sin embargo, no basta con fomentar la cultura de la autoevaluación escolar. Es necesario proveer a las escuelas de instrumentos de autoevaluación válidos y de fácil manejo para elaborar autodiagnósticos de las diversas áreas que integran la práctica escolar.

De manera particular, las rubricas analíticas, son instrumentos de autoevaluación que permiten realizar evaluaciones precisas, además, permiten especificar los elementos

del objeto de evaluación (en la mayoría de los casos conocimientos y competencias) y el nivel de desempeño (De la Cruz, 2011). Asimismo, se ha reportado que el uso de las mismas ayuda a generar conciencia y reflexión acerca de los logros alcanzados, ya que facilitan la identificación de errores y con los resultados es posible diseñar estrategias que coadyuven a alcanzar objetivos (Andrade, 2005; Reddy y Andrade, 2010).

De acuerdo con las bondades que ofrece el uso de las rúbricas, se considera que el diseño de una rúbrica analítica que integre múltiples dimensiones de la práctica escolar, ayudará a establecer en las escuelas la autoevaluación para la mejora escolar. Se espera que a través de este instrumento, sea posible obtener valoraciones claras y precisas del nivel de eficacia de las acciones llevadas a cabo por los actores escolares; y con ello, permiten identificar las áreas de oportunidad, lo cual servirá como guía para establecer metas y estrategias congruentes con las necesidades.

Cabe señalar que a pesar del frecuente uso de las rúbricas, en el ámbito educativo, los procedimientos para su elaboración y obtención de validez suelen ser soslayados, ya que hay poca claridad con respecto a los criterios que se deben considerar para diseñarlas (Cano-García, 2015; Jonsson y Svingby; 2007; Reddy y Andrade, 2010). Evidencia de ello es la escasez de estudios enfocados a obtener evidencias de validez, los cuales son insuficientes para sistematizar su proceso de elaboración y validación (Reddy y Andrade, 2010). De acuerdo con lo anterior, al ser la rúbrica un instrumento de evaluación, se considera que los procedimientos para su elaboración deben ser sistematizados y sobre todo reportados para lograr que estas se consoliden como un instrumento que asegure validez y confiabilidad en las evaluaciones.

## 1.2. Objetivos

**1.2.1. General.** Desarrollar una rúbrica analítica para la autoevaluación de prácticas escolares en escuelas secundarias.

### 1.2.2. Específicos.

- Determinar la eficacia de las prácticas escolares seleccionadas en la investigación *Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California* (Chaparro et al., 2015) mediante la opinión de expertos, para establecer los criterios de la rúbrica analítica.
- Integrar los criterios y niveles de valoración de la rúbrica analítica.
- Obtener el índice de validez de contenido de Lawshe de la rúbrica desarrollada.

## 1.3. Justificación

Existen prácticas escolares consideradas como factores que inciden positivamente en el rendimiento académico. Éstas se relacionan con las estrategias pedagógicas, el clima escolar, la infraestructura escolar (Backhooff, Bouzas y González-Montesinos, 2008; Hernández, 2009) el liderazgo directivo y la gestión escolar (Muñoz-Repiso, 2000; Lizasoain y Angulo, 2014; Reynolds, Sammons, De Fraine, Townsend y Van, 2011; Volante, 2008). Asimismo, la autoevaluación es considerada como una estrategia de mejora que contribuye al desarrollo de iniciativas a favor del rendimiento académico de los estudiantes.

Las autoridades del sistema educativo mexicano han destacado la importancia que guarda la autoevaluación escolar para diseñar planes de mejora. En la *Revisión de*

la OCDE sobre la evaluación en México (Santiago et al., 2014) se señaló que desde el año 2000 se han llevado a cabo iniciativas para promover la autoevaluación de las escuelas mediante el uso de materiales de orientación, tales como las Guías de apoyo para la autoevaluación (2000); los indicadores de calidad (2003), la descripción de las características de escuelas con altos resultados académicos (2007) y los Estándares de gestión para la Educación Básica propuestos por la SEP (2010).

Asimismo, algunos sistemas educativos estatales cuentan con instrumentos de autoevaluación escolar como los cuestionarios, los formatos de observación y las rúbricas holísticas para elaborar planes de mejora dentro de los programas de escuelas de calidad (Secretaría de Educación Básica del Estado de Baja California Sur, 2010; Secretaría de Educación Básica del Estado de Veracruz, 2010). Estos instrumentos están orientados a la evaluación interna de algunas áreas correspondientes a la Mejora Escolar (gestión escolar, convivencia escolar, participación social/comunitaria, estrategias pedagógicas).

Si bien, existen iniciativas para la práctica de la autoevaluación, de acuerdo con Santiago et al. (2014) la participación de la escuela en programas que fomentan la evaluación escolar es opcional y la autoevaluación, en la mayoría de los casos, se limita a que los supervisores se aseguren de que en las escuelas se cumpla con las obligaciones establecidas por el Sistema Educativo Estatal. Aunado a esto, el equipo directivo no cuenta con el tiempo necesario para realizar autoevaluaciones internas debido a sus múltiples tareas administrativas, dejando de lado las educativas.

## **Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias**

Con base en lo reportado en los documentos de las instituciones educativas, no basta con fomentar la cultura de la autoevaluación. Es necesario proveer a las escuelas de instrumentos de autoevaluación prácticos, de fácil manejo y con evidencias de validez para elaborar autodiagnósticos de las diversas áreas que integran la práctica escolar. A este respecto, se ha establecido que el uso de rúbricas analíticas, en el contexto escolar, favorece la autoevaluación de acciones específicas (Andrade, 2005). Estas son instrumentos que ofrecen ventajas con relación a la precisión de la evaluación de constructos de forma consensuada, ya que al ser instrumentos de medición permiten detallar, de manera anticipada, los aspectos que integran una actividad y los niveles de desempeño que dicha actividad puede alcanzar (De la Cruz, 2011). Sin embargo, aunque su uso es frecuente, los procedimientos reportados para su construcción y obtención de evidencias de validez son insuficientes para sistematizar su proceso de elaboración (Reddy y Andrade, 2010).

De acuerdo con el propósito de esta investigación, se espera que el diseño de esta rúbrica analítica con evidencias de validez para evaluar prácticas escolares, contribuya a facilitar el proceso de autoevaluación escolar. Con ello, los miembros de la comunidad escolar podrán incluir a su práctica la autoevaluación para elaborar estrategias de mejora orientadas a la realidad educativa de cada escuela. Asimismo, se considera que este estudio presenta un aporte metodológico con relación al proceso de desarrollo de la rúbrica analítica, así como también para el proceso de obtención de evidencias de validez.

## *2. Revisión de la literatura*

### 2.1. Eficacia Escolar y Mejora Escolar

La investigación en Efectividad y Mejora Escolar (IEM) se ha consolidado como una línea de investigación que ha influido en políticas y prácticas educativas de diversos países (Reynolds y Teddlie, 2000). Dentro de su enfoque tradicional, se ha centrado hacia el estudio de los factores escuela-aula y el logro académico a través de diferentes enfoques metodológicos (Carrasco, 2008). Si bien, estas dos disciplinas (Eficacia y Mejora Escolar) emergieron de manera simultánea a lo largo del tiempo se ha realizado una diferenciación.

Por una parte, el movimiento de investigación de Eficacia Escolar surge a mitad de los años 60 a partir del reporte de investigación de Coleman (1966), quien concluyó que “las escuelas tienen poco impacto en el desempeño de un alumno” (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld y York, 1966, citado en Carvallo-Pontón, 2010). Esta conclusión impactó a la comunidad científica de aquel tiempo por lo que surgió el interés por realizar investigaciones que corroboraran lo reportado por Coleman. De acuerdo con algunos estudios de la última década, se ha demostrado que la escuela influye de manera significativa en el rendimiento académico (Lizasoain y Angulo, 2014; Reynolds, Sammons, De Fraine, Townsend y Van, 2011; Román, 2008.; Sheilder y Zoido, 2013). De acuerdo con Murillo (2003), es una línea de investigación pedagógica sólida y sistemática que aporta a la identificación y descripción de los factores que se asocian al rendimiento académico (Murillo, 2003). Por lo que ha sido posible identificar elementos generales que hacen a una escuela eficaz (Muñoz-Repiso et al., 2000).

Las investigaciones sobre Eficacia Escolar se basan en la premisa de que las acciones al interior de la escuela son capaces de mejorar las condiciones de los estudiantes y así compensar características de su contexto (Carvallo-Pontón, 2010). Desde esta perspectiva, nace el término de escuelas eficaces, al respecto, Murillo (2005) las definió como aquellas que consiguen generar un desarrollo integral de todos sus alumnos, más allá de lo esperable; teniendo en cuenta el rendimiento previo, la situación social, económica y cultural de sus familias.

El objeto de estudio en las investigaciones sobre Eficacia Escolar se integra por (a) el estudio de los efectos escolares y sus propiedades científicas, (b) el estudio de los factores de eficacia escolar a través de la utilización del modelo contexto-insumo-proceso-producto, (c) el estudio de la mejora escolar y (d) el estudio de la mejora de la eficacia escolar (Zorrilla, 2008). Siendo este último un movimiento de investigación que surge a raíz de las investigaciones sobre Eficacia Escolar a finales de la década de los 60 el cual ha contribuido a la elaboración de políticas educativas orientadas al cambio y mejoramiento de los centros escolares.

De acuerdo con Zorrilla (2008), el conocimiento sobre los actores, las instituciones y los procesos educativos ha contribuido a la transformación de la educación. Sin embargo, el cambio debe surgir y ser liderado por la comunidad escolar (Murillo, 2003). A este respecto, las investigaciones sobre Mejora Escolar son dirigidas a identificar y describir “los procesos que desarrollan las escuelas que consiguen poner en marcha un proceso de cambio para optimizar su calidad” (Muñoz-Repiso et al. 2000, p. 13). Asimismo, se caracterizan por presentar una orientación práctica hacia el cambio y

mejora de las escuelas con dificultades (Raczynski y Muñoz, 2005) ya que permite identificar elementos que favorecen u obstaculizan los procesos de cambio (Murillo, 2003).

Con base en una revisión de diversas aportaciones conceptuales, Murillo (2003, p.3) definió la Mejora Escolar como “la capacidad de la escuela para incrementar, de forma simultánea, el aprendizaje de los alumnos y el desarrollo de la comunidad escolar”. Además, mencionó que esta se fundamenta en: (a) los cambios centrados en la escuela, (b) la implicación de todo el personal de la escuela en el cambio, (c) la construcción de una comunidad de aprendizaje mediante la inclusión del personal escolar, (d) la información que se obtiene a partir de los datos de la escuela, los alumnos y la literatura, (e) la potencialización del desarrollo continuo de los docentes, (f) la promoción de la capacidad de los estudiantes por aprender. Por último, se centra en (g) el análisis de la enseñanza y el currículo con la finalidad de mejorarlos. Al respecto, cabe señalar que las escuelas eficaces no son el resultado simplista de la suma de estos elementos aislados; estas escuelas se caracterizan por una manera “especial de ser, pensar y actuar, es decir, poseen una cultura de eficacia” (Murillo, 2011).

Los hallazgos derivados de las investigaciones pueden variar dependiendo del contexto sociocultural y del sistema educativo de cada país (Sheilder y Zoido, 2013). Sin embargo, es posible identificar una serie de factores en común que se encuentran presentes en la literatura revisada sobre Eficacia y Mejora Escolar. Hernández-Castilla, Murillo, y Martínez-Garrido (2016, p. 105) sintetizaron los factores relacionados a la Eficacia Escolar en diez:

“a) Sentido de comunidad; b) Liderazgo educativo; c) Clima escolar y de aula; d) Altas expectativas; e) Calidad del currículo y las estrategias de enseñanza; f) Organización del aula; g) Seguimiento y evaluación; h) Desarrollo profesional de los docentes, entendido tanto como actitud hacia el aprendizaje continuo y la innovación; i) Implicación de las familias; e j) Instalaciones y recursos”.

Con relación a las investigaciones sobre Mejora Escolar se reconocen factores esenciales para la mejora de la escuela, entre éstos, Murillo (2003) mencionó cinco: (a) el cambio centrado en la escuela, (b) el cambio dependiente del profesorado; c) el papel determinante de la dirección; (d) la escuela como comunidad de aprendizaje, y (e) el cambio en la forma de enseñar. Asimismo, la interacción entre las dimensiones de (a) liderazgo, (b) guía instruccional, (c) profesional capacitado (d) clima centrado en el estudiante (e) relaciones entre comunidad y familiares y (f) confianza relacional promueven la mejora de las escuelas (Bryk, Sebring, Allensworth, Luppescu, y Easton, citado en Scanlan, 2011).

Otro factor relevante, son las estrategias de Mejora Escolar que se orientan a incrementar la efectividad de las escuelas las cuales se formulan en función de la realidad de cada escuela y su sistema educativo. Al respecto, en el XXII Encuentro de Consejos Escolares Autonómicos y del Estado en Oviedo (Ministerio de Educación, Cultura y Deporte, 2014) se reportaron algunas recomendaciones generales sobre la orientación de las estrategias de mejora, tales como:

- reforzar la calidad y compromiso del profesorado,

- promover en los profesores el dominio de metodologías dirigidas hacia la mejora y eficacia,
- fomentar las relaciones del profesorado tanto con la comunidad escolar como con los familiares,
- desarrollar por parte de los profesores, un currículo que promueva la adquisición de competencias y conocimientos integrales,
- promover la motivación e interés en los alumnos hacia la educación,
- detectar de manera oportuna las dificultades del alumno,
- fomentar el ejercicio de liderazgo efectivo,
- establecer de lazos de apoyo entre la comunidad y la escuela,
- concientizar acerca de la importancia de implementar una cultura evaluativa,
- adaptar los recursos materiales y humanos en función del contexto y las necesidades.

De manera general, mientras la Eficacia Escolar se avoca a la identificación de factores que explican las diferencias de la efectividad entre escuelas, el enfoque de Mejora Escolar se dirige hacia el estudio de los procesos escolares y capacidad de las escuelas para realizar cambios a través del análisis de la trayectoria de mejoramiento de las mismas (Carrasco, 2008). Ambos enfoques -Eficacia Escolar y Mejora Escolar- son ineludibles cuando se trata el tema de la optimización o mejora de los procesos educativos (Muñoz-Repiso et al. 1995). Eficacia Escolar, aporta conocimiento sobre qué implementar en las estrategias de Mejora Escolar; y a su vez, éstas se transforman en herramientas para probar las teorías y modelos propuestos de la Eficacia Escolar (Creemers y Reezigt, 1997).

## 2.2. Buenas prácticas escolares asociadas a los factores de Eficacia Escolar y Mejora Escolar

Las buenas prácticas escolares son acciones eficaces realizadas por los responsables de la educación (Ballesteros y Mata 2009), además, se consideran innovadoras y dignas de ser imitadas (Abdoulaye, 2003; Campaña y Tapia, 2010). De acuerdo con lo reportado por el Fondo Nacional de Desarrollo de la Educación Peruana (2014) la conceptualización de este constructo es amplio, sin embargo, es posible identificar cuatro elementos en común referidas en algunas definiciones:

“(a) una buena práctica está referida a una acción o a una forma de hacer las cosas , que se expresa en un procedimiento o una metodología, (b) implica un modo distinto de hacer las cosas, es decir, la introducción de una novedad, (c) es verificablemente efectiva, es decir, ha mejorado, a través de resultados, la situación previa a la implementación de la “buena práctica”, y existen evidencias de ello, (d) tiene un alto potencial para ser replicable o de servir de ejemplo en otros contextos parecidos” p.28.

Asimismo, de acuerdo con lo reportado en la investigación de Lizasoain y Angulo (2014) las buenas prácticas favorecen la explicación de los buenos resultados obtenidos por los alumnos; además, estas pueden ser transferidas a planes y acciones para la mejor escolar. Se considera que las buenas prácticas pueden ser clasificadas en función de algunos factores de Eficacia Escolar y Mejora Escolar reportados en la literatura, puesto que en el marco de las investigaciones de dichos modelos, se han identificado acciones en común efectivas llevadas a cabo en las escuelas, sin embargo, los factores no se encuentran aislados y suelen estar vinculados entre sí (Cornejo y Redondo, 2007).

Asimismo, al identificar dichos factores permite elaborar intervenciones para incidir positivamente en los aprendizajes y el logro escolar de los estudiantes (Román, 2008).

En virtud de lo anterior, se realizó una clasificación de las prácticas escolares reportadas en investigaciones previas, en función de siete factores: (a) liderazgo escolar, (b) gestión y organización, (c) recursos (d) estrategias de enseñanza, (e) clima escolar, (f) implicación familiar y (g) evaluación escolar.

**2.2.1. Liderazgo escolar.** Se refiere a las acciones de orientación, toma de decisiones, motivación e influencia ejercidas, generalmente, por los miembros de los equipos directivos (Horn y Marfán, 2010; Riveros, 2012). En este sentido Gros, Fernández-Salinero, Martínez y Roca (2013) señalaron que la orientación implica manifestar y precisar los objetivos escolares, mientras que la motivación contribuye al desarrollo de la comunidad escolar mediante el reconocimiento de sus necesidades e intereses ofreciendo espacios para generar compatibilidad entre las características personales con las necesidades e intereses de la escuela. Asimismo, los autores señalaron que las buenas prácticas de liderazgo se caracterizan por autoridad auténtica, la cual implica que es reconocida y respetada por los demás sin sentirse obligados. Se ha demostrado que es esencial la presencia de liderazgo claro de una o varias personas. Generalmente es el director o directora de la escuela quien lidera, sin embargo, depende de su actitud y de cómo es visto por los docentes ya que cuando existe un directivo comprometido, competente, aceptado por la comunidad escolar y capaz de motivar mediante un estilo participativo favorece la ejecución de los procesos eficaces para el cambio escolar (Murillo y Krichensky, 2012).

De acuerdo con Murillo, Barrio y Pérez-Albo (1999) el liderazgo escolar es un factor determinante de la eficacia de las escuelas ya que conlleva al establecimiento de un valor agregado y orienta a la comunidad escolar a obtener resultados esperados (Garay y Uribe, 2006). De manera particular, Murillo (2011) señaló que se ha demostrado mayor grado de eficacia en el liderazgo de tipo pedagógico el cual se distingue por el compromiso de los directivos con la comunidad escolar y es un profesional con altas capacidades. Asimismo, en otras contribuciones, se ha demostrado que las prácticas de liderazgo pedagógico son frecuentes en escuelas con alumnos que presentan altos niveles de rendimiento académico (Franco, 2012; Hernández, 2009; Muñoz-Repiso, 2000; Lizasoain y Angulo, 2014; Beach, 2008). De esta manera, el liderazgo escolar se considera como una variable explicativa de la obtención de mejores resultados educativos (Gros, Fernández-Salineró, Martínez y Roca, 2013).

En los hallazgos reportados por Ritacco-Real (2011), la flexibilidad de los líderes educativos con relación a la forma de organizar la escuela (de acuerdo a las necesidades de los estudiantes) y la ejecución de liderazgo compartido, son buenas prácticas presentes en escuelas con déficit socioeconómico y que aplican estrategias de respuesta para los estudiantes en riesgo de fracaso escolar.

**2.2.2. Gestión y organización.** Implica la gestión de actividades para fomentar el conocimiento en sus estudiantes y mejorar constantemente los procesos de funcionamiento escolar (Garay y Uribe, 2006). Además, las acciones y formas en que se lleva a cabo la gestión y organización al interior de las escuelas han demostrado que inciden en los niveles de logro académico (Hernández, 2009; Lizasoain y Angulo, 2014;

Cabrera, 2006; Muñoz-Repiso et al., 1995). Las tareas de gestión y organización generalmente se atribuyen a los miembros de la dirección escolar, puesto que representan una figura importante dirigir y lograr objetivos escolares (Garay y Uribe, 2006). En las escuelas con altos logros académicos, la gestión es sistematizada y planificada. Se ha demostrado que la manera en la que se gestiona el tiempo dentro de la escuela, influye significativamente en las oportunidades de aprendizaje de los alumnos (Garay y Uribe, 2006; Pérez, 2016). A este respecto, Murillo (2011) señaló que existen seis aspectos a considerar en la gestión del tiempo: (a) número de días lectivos impartidos en el aula, (b) puntualidad con que inician las clases, (c) optimización del tiempo destinado a clase, (e) interrupciones en el tiempo destinado a las tareas de enseñanza y aprendizaje; y (c) flexibilidad para organizar el tiempo.

De acuerdo con López (2010), para lograr un aprendizaje de calidad la gestión escolar debe ser clara ya que impacta de manera positiva en la optimización de recursos, eficiencia y calidad de los procesos al interior de las escuelas. Además, Murillo y Krichensky (2012) señalaron que una organización flexible, capaz de adaptarse a los alumnos, a nuevas necesidades y visiones pedagógicas, favorece a que los procesos de cambio sean exitosos. Ejemplo de ello es la flexibilidad en horarios, la adaptación del calendario escolar y la redistribución de recursos humanos y materiales.

**2.2.3. Recursos.** Según lo reportado en algunas investigaciones, los recursos son insumos con los que cuenta una escuela. Entre estos se encuentran los monetarios, los humanos, los materiales didácticos, la infraestructura, el mobiliario y el tamaño de las aulas (Racynski y Muñoz, 2005). Los recursos generan condiciones adecuadas para el aprendizaje del estudiante. Sin embargo, “ya que se cubre el mínimo necesario, resulta

significativo el uso que se les dé en un sentido pedagógico” (Cornejo y Redondo, 2007, p. 162).

Los recursos son factores que marcan la diferencia en el rendimiento académico de los estudiantes y característico de escuelas eficaces (Murillo, 2008), ya que cuando los materiales educativos son de calidad existe alto desempeño académico (Ruíz de Miguel, 2009). Por lo tanto, se considera que las escuelas deben garantizar más recursos para realizar actividades plasmadas en la planificación de los proyectos escolares, ya que se requiere de espacios y recursos específicos orientados para dicho fin (Murillo y Krichensky, 2012). En el caso de México la desigual con relación a la distribución de los recursos educativos (materiales y humanos) representa un efecto significativo en la calidad y equidad de la educación (Blanco, 2009); lo anterior se relaciona con lo mencionado por Murillo (2003) quien destacó que en los países latinoamericanos los recursos son un factor que incide en la calidad educativa.

**2.2.4. Estrategias de enseñanza.** Pertenecen a factores de nivel aula, específicamente del proceso pedagógico llevado a cabo por el docente (Román, 2008) los cuales se consideran decisivos para la mejora de la eficacia escolar (Cardemil, Maureira y Zuleta, 2012). En esta categoría se incorporan prácticas relacionadas a la utilización de metodologías y estrategias didácticas adecuadas para las asignaturas y los estudiantes, además, incluye elementos sobre la organización física del aula y de las actividades (Murillo, 2008).

Cardemil, Maureira y Zuleta (2012) mencionaron que las estrategias de enseñanza deberán orientarse hacia la mejora de los aprendizajes y al rendimiento de

los estudiantes. Asimismo, los autores señalaron que las prácticas relacionadas con la gestión del tiempo en el aula es un elemento con aporte significativo para la eficacia escolar. En este sentido, en los hallazgos reportados por Intxausti, Joaristi y Lizasoain (2015) la gestión del tiempo dedicado a la enseñanza y al aprendizaje resultó ser un aspecto relevante para los miembros de los equipos directivos escolares. De manera específica, se enfatizó el valor otorgado a las reuniones de comité escolar y la coordinación pedagógica, ya que es ahí en donde se establecen acuerdos sobre la metodología didáctica, recursos e instrumentos de evaluación a implementar.

En virtud de lo anterior, se considera que es posible que los cambios efectivos de una escuela surjan a partir del cambio en la práctica y la perspectiva docente; dichos cambios deberán estar dirigidos a la mejora de la enseñanza, lo cual requiere del esfuerzo de los docentes para asumir y reconocer el papel determinante que tienen en la educación de sus estudiantes (Román, 2008).

**2.2.5. Clima escolar.** Se considera que a través de la buena convivencia, entre los miembros de la comunidad escolar, es posible generar un clima escolar adecuado ya que es indispensable para llevar a cabo el proceso educativo (López de Mesa-Melo, Carvajal-Castillo, Soto-Godoy, Urrea-Roa, 2013). De manera específica, se considera que el clima positivo en el aula es uno de los factores con mayor incidencia en el rendimiento escolar de los estudiantes y está caracterizado por la aceptación de los intereses y las necesidades de los estudiantes por parte del docente, además de la confianza y el establecimiento de reglas claras (Cardemil, Maureira y Zuleta, 2012).

Murillo (2008) señaló que la presencia de estrategias dirigidas a fomentar un clima afectivo y buenas relaciones entre los miembros de la comunidad escolar son elementos presentes en escuelas eficaces. Según lo reportado por Lizasoain y Angulo (2009), en este tipo de escuelas también se pueden encontrar buenas redes de interacción entre la comunidad escolar, mismas que suelen fomentarse por los directivos para satisfacer las necesidades de los estudiantes y los docentes. De forma contraria, se ha evidenciado la presencia de clima escolar “negativo” en escuelas consideradas como ineficaces (Hernández-Castilla, Murillo y Martínez-Garrido, 2014).

En este sentido, se ha demostrado que los estudiantes adquieren mejor los aprendizajes en escuelas con ambientes organizados y tranquilos en donde se sienten seguros, cuidados y atendidos. Asimismo, la presencia de relaciones positivas en el salón de clase influye en el grado de satisfacción de los estudiantes hacia la escuela y en el rendimiento académico (Martínez-Garrido y Murillo, 2016).

**2.2.6. Implicación familiar.** Se entiende como el involucramiento de los padres de familia en una o varias actividades relacionadas con la educación de sus hijos, tales como: participar libremente en la mejora de la escuela, asistir a juntas o ayudar con las tareas (Valdés, Martín, y Sánchez, 2009). La actitud positiva de las familias hacia la colaboración con la escuela, la participación en actividades escolares, extracurriculares y complementarias, son aspectos presentes en las escuelas eficaces (Murillo, 2008) además, se ha comprobado su impacto positivo que tiene en el desarrollo cognitivo, socioafectivo y en el rendimiento académico de los estudiantes (Martínez-Garrido y Murillo, 2016).

Los docentes representan un papel importante para la participación familiar ya que a través de actividades y proyectos son capaces de fomentarla. Además, se ha demostrado que la colaboración de los padres de familia permite que los docentes obtengan más información respecto a sus estudiantes y con ello pueden mejorar su práctica docente (Andrés y Giró, 2016). En este sentido, las relaciones que se establecen entre la familia y la escuela se conciben como una herramienta esencial para mejorar los procesos y resultados educativos por lo que es necesario orientar esfuerzos hacia una colaboración efectiva (Egido, 2015).

**2.2.7. Evaluación escolar.** De acuerdo con la revisión de la literatura, se considera como uno de los factores de eficacia escolar de mayor incidencia en el rendimiento académico. Asimismo, contribuye al desarrollo de políticas educativas acorde a los resultados obtenidos de las mismas. Dada la implicación que tiene este factor en el proceso educativo, se consideró pertinente ampliar se descripción en el siguiente apartado.

### **2.3. Evaluación y autoevaluación como estrategia de mejora**

La evaluación escolar está presente desde la década de los treinta del siglo pasado. Inicialmente, el propósito fue verificar el cumplimiento de objetivos en los programas educativos y posteriormente, la valoración de la utilidad de los resultados de las evaluaciones para la toma de decisiones (Román, 2011). Así, mediante la práctica de la evaluación escolar y con los resultados obtenidos se espera que se sustenten y definan políticas públicas para garantizar la calidad educativa (Ruíz-Cuéllar, 2013). De esta manera, es posible mejorar y transformar los procesos educativos (Schleicher y Zoido, 2013).

Perassi, (2010) señaló que existen evaluaciones caracterizadas por presentar procesos sistematizados y planificados, con lo cual se fundamenta la toma de decisiones. De manera contraria, también se identifican evaluaciones con procedimientos generalizados, los cuales carecen de sistematización y rigor técnico, por lo que la toma de decisiones se basa en juicios de valor y situaciones aisladas. En este sentido, fortalecer los procesos para la evaluación escolar representa un desafío para los sistemas educativos, ya que el análisis es complejo, pues va desde la evaluación de resultados hasta la de los procesos (Ferreira, Salgueiro, y Castagno, 2013). Lo anterior se relaciona a que la educación, entendida como proceso, es integral, puesto que implica la interacción de experiencias e intercambios socioculturales del individuo con otros y con su contexto (Horbath y Gracia, 2014).

En los sistemas educativos de los países de Latinoamérica se ha evidenciado desde el inicio del siglo XXI la práctica de la evaluación de la calidad de la educación. Sin embargo, dichas evaluaciones se consideran limitadas, ya que la conceptualización de calidad educativa es reduccionista, es decir, dirigida a evaluar el desempeño curricular de los estudiantes en ciertas áreas del conocimiento (Murillo y Román, 2010). Asimismo, los instrumentos y herramientas desarrollados por las instituciones evaluativas no contemplan las características contextuales, por lo que el riesgo de caer en el reduccionismo instrumental también representa una limitación para establecer estrategias evaluativas favorecedoras (Poggi, 2013).

No obstante, en América Latina y en México, se reconoce un avance significativo en las prácticas de evaluación educativa, tales como: establecimiento de instituciones

evaluativas y la alineación de objetivos y metas con relación a los resultados de las evaluaciones internacionales (Horbath y Gracia, 2014). De manera particular, Santiago et al.(2014) identificaron ciertas características positivas de la evaluación escolar en México: atención a las políticas de evaluación de las escuelas, énfasis en la capacitación sobre liderazgo escolar y en la autoevaluación escolar. En este sentido, la autoevaluación escolar se considera una práctica que favorece el desarrollo de estrategias y planes para la mejora escolar (Murillo y Krichesky 2015), puesto que para proponer y realizar cambios en los procesos educativos se requiere conocer cuáles son las dimensiones necesarias de ser intervenidas y reconocer los factores que están directamente implicados en la escuela (Román, 2008). Siguiendo con lo anterior, en el contexto mexicano, se han desarrollado modelos de autoevaluación por parte del Instituto Nacional para la Evaluación Educativa (INEE), en los que se propone que la toma de decisiones debe realizarse con base en mecanismos de evaluación manejados en el nivel de cada escuela, es decir, adaptados al contexto escolar (Horbath y Gracia, 2014). En este sentido, Ruíz-Cuéllar (2013) refirió que los mecanismos que se utilizan no proporcionan información suficiente para realizar valoraciones sobre la calidad de la educación y para sustentar acciones de mejora.

Rueda-Beltrán, (2010) señaló que la autoevaluación escolar consiste en la valoración de las acciones de la escuela y de la comunidad escolar (directivos, docentes y estudiantes), con el objetivo de potencializar su desarrollo. Además, mencionó que implica un alto compromiso con los procesos para propiciar la mejora de sus prácticas, y con ello, favorecer la calidad educativa a través de la autorreflexión, puesto que el proceso de autoevaluación implica realizar juicios sobre el propio aprendizaje (Bound,

1995, citado en Sempere Ortells, García Irlles y Marco de la Calle, 2011). Con base en lo anterior, la autoevaluación escolar se puede definir como la participación de los actores escolares en el proceso de realizar juicios o valoraciones sobre su propio nivel de ejecución de prácticas hacia el interior de las escuelas.

Como ya se mencionó, con los resultados obtenidos de la autoevaluación es posible desarrollar estrategias y planes orientados a la mejora escolar. Al respecto, Perassi, (2010) señaló que es un desafío establecer culturas evaluativas democráticas en las que se vinculen exitosamente estos dos conceptos, y que para ello es necesario una práctica evaluadora con ciertas condiciones; entre estas se destacan: la implicación del sujeto en el proceso de evaluación (rol activo), la reflexión y el debate, concebir la evaluación como una necesidad institucional (perspectiva formativa), la implicación progresiva de los actores escolares (participación y compromiso), la difusión de los resultados obtenidos, considerar las evaluaciones externas como complemento de la autoevaluación y el diseño de estrategias intervención. Además, el proceso de autoevaluación escolar según Ortega, Salazar y Campos (2011) debe ser integral, en el que se involucren dimensiones tales como la pedagógica y sociocultural, de esta manera, podrá contribuir al cambio y mejora de la realidad de la escuela.

De acuerdo con algunos hallazgos empíricos, la autoevaluación escolar también se relaciona con la eficacia escolar, ya que se ha identificado que las escuelas eficaces utilizan los resultados académicos de los estudiantes para el desarrollo de planes de mejora (Lizasoain y Angulo, 2014) y otros cambios orientados a mejorar la calidad, equidad y eficiencia de la educación (Schleicher y Zoido, 2013). Asimismo, las iniciativas

de autoevaluación escolar se han enfocado a evaluar factores identificados en escuelas eficaces, tales como: clima escolar, liderazgo educativo, metas en común de la comunidad escolar, estrategias de enseñanza, recursos educativos, entre otros (Murillo, 2007).

Con relación a los instrumentos que se utilizan en los procesos de autoevaluación Rul y Zaitegi, (2003) mencionaron que pueden ser clasificados según su enfoque metodológico (cuantitativo, cualitativo y mixto) y que existen cinco tipos: las pruebas, los cuestionarios, la observación y la entrevista. De acuerdo con lo reportado en la literatura, es posible identificar el uso de indicadores, cuestionarios, listas de cotejo y encuestas para realizar autoevaluaciones escolares. En el caso particular de México, Santiago et al., (2014) afirmaron que los materiales para la autoevaluación escolar que ofrece el INNE se caracterizan por ser prácticos, de buena calidad y además facilitan obtener información sobre aspectos específicos de la escuela. Además, el autor señaló que la autoevaluación no debe ser un proceso mecánico sustentado solo en elementos cuantitativos, sino que deberá triangularse con elementos contextuales, en este sentido, las rúbricas se consideran instrumentos que favorecen la autoevaluación en el contexto escolar, por lo tanto, en el siguiente apartado se describe su utilidad y características.

### 2.4. Rúbricas

Las rúbrica se definen como un instrumento en el cual se enumeran criterios y niveles de calidad esperados en determinada actividad (Andrade, 2005). En el ámbito educativo se conceptualiza como un conjunto de criterios desde los cuales se juzga, valora y califica un determinado elemento del proceso educativo; por lo que se consideran guías de

puntuaciones precisas utilizadas para la evaluación del aprendizaje, habilidades y productos de los estudiantes (Martínez-Rojas, 2008). Estas suelen ser tablas de datos en las que se desglosan niveles de desempeño con criterios específicos para valorar cada uno (Beyreli y Ari, 2009; Gatica-Lara y Uribarren-Berrueta, 2013). En estos instrumentos se especifica el desempeño general esperado y los niveles de competencia alcanzados en el desarrollo de una habilidad por los aprendices (Mansoor y Grant, 2002, citado en Picón, 2013). En las rúbricas se utilizan escalas cuantitativas o cualitativas asociadas a criterios preestablecidos para evaluar el constructo (Gordillo y Rodríguez, 2006). Así, se obtiene un nivel de desempeño con relación al proceso y resultado de la acción valorada (Lima-Rodríguez, Lima-Serrano, Ponce-González y Guerra-Martín, 2015).

Las rúbricas son instrumentos que ofrecen ventajas con relación a la precisión de la evaluación de constructos de forma consensuada, ya que permiten especificar, de manera anticipada, los aspectos que forman parte de una actividad y el nivel de desempeño a alcanzar (De la Cruz, 2011). Sin embargo, la diversidad de constructos y la falta de precisión de los mismos en ocasiones conlleva a discrepancias entre los evaluadores y a evaluaciones imprecisas que afectan la toma de decisiones (Arias y Maturana, 2005, citado en Picón, 2013); por lo que es importante considerar ciertos aspectos para su elaboración. Al respecto, Gatica-Lara y Uribarren-Berrueta (2013) señalaron que una rúbrica deberá (a) ser coherente con los objetivos educativos que se persiguen, (b) apropiada para el nivel de desarrollo de los estudiantes, y (c) establecer niveles con términos claros.

Los resultados que arrojan las rúbricas permiten generar la toma de conciencia y reflexión acerca de los logros alcanzados y los que faltan por cumplir; ya que facilitan la identificación de errores y permiten diseñar instrucciones para alcanzar metas establecidas (Andrade, 2005). Esto se debe a que es posible valorar áreas subjetivas y complejas mediante criterios que cualifican progresivamente aprendizajes y competencias que van desde niveles incipientes hasta expertos (Díaz-Barriga, 2006; citado en Gatica-Lara y Uribarren-Berrueta, 2013). Sin embargo, para beneficiarse de las ventajas que ofrece este tipo de instrumento, es importante considerar los aspectos relacionados con su validez y fiabilidad. Como lo señaló Andrade (2005), estos aspectos deberán estar alineados a los estándares vigentes de calidad, lo que requiere análisis complejos para obtener evidencias de validez y fiabilidad tales como análisis factorial confirmatorio y alfa de Cronbach.

Otro beneficio o ventajas de las rúbricas, es que pueden ser utilizadas en otros contextos. Al respecto Reddy y Andrade (2010) refirieron que se han utilizado en varias disciplinas: medicina, enfermería, administración, odontología y formación docente. Además, son útiles en la evaluación de grupos de estudiantes tanto grandes como pequeños, mejorar la coordinación entre el profesorado y permiten el desarrollo de liderazgo (Urbieto, Garayalde, y Losada, 2011).

En conclusión, las rúbricas son instrumentos de medición que facilitan la evaluación de conocimientos, habilidades y competencias mediante criterios y niveles de ejecución. Se reconoce que dichos instrumentos son mayormente utilizados en el área de la educación para la evaluación formativa y sumativa de los aprendizajes, Por lo tanto, en el caso de la autoevaluación de prácticas escolares, se considera que puede facilitar

valoraciones claras y precisas del nivel de eficacia de las acciones llevadas a cabo por los actores escolares; y con ello, identificar las áreas en las que requieran mejorar, lo cual servirá como guía para establecer metas y estrategias congruentes con los resultados de las evaluaciones.

**2.4.1. Tipos.** Con base en la revisión de literatura, se encontró que las rúbricas pueden ser dos tipos: holísticas y analíticas. Las rúbricas holísticas evalúan el aprendizaje o competencias desde una perspectiva global (Gordillo y Rodríguez, 2012). En este tipo de rúbrica se asigna una puntuación general a la ejecución, proceso o producto que se está evaluando, por lo que su uso no es recomendable cuando se busca especificar niveles intermedios o bajos de desempeño (Beyreli y Ari, 2009).

Las rúbricas analíticas se centran en áreas concretas de aprendizaje (Gordillo y Rodríguez, 2012), actividad, proceso o producto (Beyreli y Ari, 2009). Mediante éstas se valoran los componentes de desempeño de los estudiantes, se identifican fortalezas y debilidades propiciando mayor comprensión sobre el propio aprendizaje y retroalimentación (Gatica-Lara y Uribarren-Berrueta, 2013). Los componentes se evalúan de forma separada y aunque requiere de mayor tiempo para su aplicación, la puntuación analítica es más amplia, comprensiva y útil (Babin y Harrison, 1999, citado en Beyreli y Ari, 2009). Además, es posible adaptarlas dependiendo el propósito, y utilizarlas como herramientas de autoevaluación (Tierney y Simon, 2004).

Shipman, Roa, Hooten y Wang (2011, citado en Gatica-Lara y Uribarren-Berrueta, 2013) recomendaron el uso de las rúbricas analíticas para valorar habilidades complejas,

obtener información detallada y promover la retroalimentación. Además, describieron tres características de éstas:

- Criterios de evaluación. Son los indicadores que determinarán la calidad del trabajo/actividad de un estudiante; estos reflejan los procesos y contenidos que se consideran importantes para valorar.
- Definiciones de calidad. Proporcionan una explicación detallada de lo que el evaluado debe realizar para demostrar los niveles de eficiencia, para alcanzar un nivel determinado de los objetivos. Estas definiciones deben proporcionar retroalimentación.
- Estrategias de puntuación. Se consideran cuatro niveles: desempeño ejemplar, desempeño maduro, desempeño en desarrollo y desempeño incipiente.

**2.4.2. Procedimientos para su elaboración.** De acuerdo con la revisión de 20 estudios sobre elaboración de rúbricas, realizada por Reddy y Andrade (2010), son pocas las investigaciones en las que se reportan los procedimientos llevados a cabo para la elaboración de las mismas, por lo que ha complicado seguir una metodología.

Tierney y Simon (2004), señalaron que la mayoría de los principios descritos para la elaboración de las rúbricas se han centrado principalmente en los niveles de desempeño y en la claridad de los descriptores; ya que aportan mayor fiabilidad a la interpretación del usuario. Sin embargo describir los criterios de desempeño puede resultar un desafío por su dificultad (García Irlés 2011 et al; Lima-Rodríguez y Lima-Serrano, 2015; Tierney y Simon, 2004). Por lo tanto, los descriptores de cada nivel deben ser tratados con los mismos criterios y atributos de desempeño de acuerdo con la escala

progresiva, la cual deberá ser continua y consistente de un nivel a otro. Así pues, la consistencia en los criterios de desempeño, se interpreta como la referencia a los mismos atributos en los descriptores a través de los niveles de logro (Simon y Forgette-Giroux, 2001).

La complejidad que representa la elaboración de una rúbrica lleva a los interesados a seleccionar el método más adecuado que dé como resultado un instrumento que evalúe con precisión. Por ello, la descripción de constructos, criterios e ítems deben ser realizados por grupos de expertos en el tema de cada investigación.

Considerando la importancia de los aspectos relacionados a la elaboración de criterios y descriptores, diversos autores (Hernández-Rivera, 2012; García Irles et al., 2011; Lima-Rodríguez, 2015; López y Velasco, 2013; Reddy y Andrade, 2010; Urbieto et al., 2011) describieron, en sus investigaciones, los procedimientos que utilizaron para la elaboración de las rúbricas. Algunos de los procesos presentan similitudes en ciertas fases y difieren en otras. Esto se debe a los propósitos y usos que se le otorga a la rúbrica diseñada. A continuación, se describen, de manera sintetizada, los aspectos relevantes de éstos procedimientos.

Con la finalidad de mejorar los procesos de evaluación de la asignatura de Biología molecular, García Irles et al. (2011) desarrollaron una rúbrica de autoevaluación de las competencias adquiridas por los estudiantes después de terminar un curso de tutorías grupales. Para su diseño siguieron los siguientes pasos: (a) revisión del contenido temático a desarrollar, (b) establecimiento de competencias a desarrollar y evaluar, (c) descripción detallada de los criterios de evaluación y asignación del valor

numérico según el nivel de adquisición de las competencias y (d) diseño de la escala de calidad de desempeño de las competencias. Con relación a las ventajas del diseño y la aplicación de la rúbrica, los autores señalaron que además de orientar al alumno respecto a la ejecución y progreso de su trabajo, el uso del instrumento permitió a los docentes realizar una evaluación objetiva y consistente. Sin embargo, no se reportaron procesos estadísticos sobre evidencias de validez del instrumento.

En otro estudio elaborado por Urbieta et al. (2011), señalaron la necesidad de homologar las competencias y la evaluación de las mismas mediante un sistema de evaluación confiable. Para ello, desarrollaron tres rúbricas para la evaluación de los contenidos de tres asignaturas en la Escuela Universitaria de Magisterio de Donostia/San Sebastián. Las rúbricas fueron aplicadas a 90 estudiantes en los grupos de euskera y 25 estudiantes en los grupos de castellano. Para su construcción se llevaron a cabo los siguientes pasos:

1. Análisis de competencias. Se realizó un análisis de las competencias comunes de ambas áreas de conocimiento, posteriormente, se describió cada una.
2. Agrupación e identificación de competencias. Se establecieron los elementos (17) de la rúbrica, comunes a ambas áreas de conocimiento, dichos elementos correspondieron a las asignaturas y a las competencias transversales propuestas.
3. Establecimiento de indicadores. Se realizó, por medio de consenso, una escala descriptiva de cuatro niveles de desempeño: escaso o deficiente, insuficiente y mejorable, suficiente o correcto y excelente o deseable.

4. Explicación de las competencias a los estudiantes y *feed-back* por parte del alumnado. A partir de las sugerencias se reelaboró y consensuaron las características del nivel de exigencia y valoración de las tareas a realizar. Así se consiguió explicitar el nivel de logro de las competencias ligadas tanto al aprendizaje de los contenidos conceptuales, como al logro de los contenidos procedimentales y al logro de los contenidos actitudinales.

De acuerdo con la discusión de resultados, los autores concluyeron que la rúbrica debería ser rediseñada, ya que los estudiantes reportaron confusión respecto a los descriptores. Por lo tanto, sugirieron reducir el número de elementos, incluir más niveles de desempeño sin incrementar su complejidad y clarificar la redacción. Lo anterior, confirma la necesidad de una descripción clara de los elementos que componen las rúbricas, así como de los niveles de desempeño de los mismos.

Hernández-Rivera (2012) elaboró una rúbrica para la evaluación de las competencias de estudiantes del curso de Clínica de Exodoncia y Cirugía en la Universidad de Costa Rica. Para el desarrollo de la rúbrica se elaboró lo siguiente:

1. Establecimiento de las competencias. Se llevaron a cabo entrevistas semiestructuradas al personal docente sobre los objetivos del curso a evaluar.
2. Agrupación de criterios de los docentes. Fueron agrupados en categorías tomando en cuenta las competencias y actitudes que los estudiantes debían desarrollar en el curso.

3. Elaboración de la rúbrica. Se realizó un análisis y se construyeron las características para determinar los niveles de la rúbrica, y con ello, determinar el rendimiento académico de los estudiantes.
4. Construcción de la escala. Se realizó una escala para jerarquizar las competencias a evaluar en los estudiantes según su grado académico y se utilizó un número impar de categorías para cada competencia.

El procedimiento descrito anteriormente refleja las consideraciones generales que conlleva la elaboración de una rúbrica para la evaluación de conocimientos. Si bien, Hernández-Rivera (2012) consideró que estas ofrecen un mecanismo objetivo mediante la estandarización de indicadores, en el procedimiento no se reportan evidencias de validez. Sin embargo, en la primera fase se hace alusión, de manera indirecta, sobre la importancia de un consenso (entre profesionales expertos) para establecer las competencias a evaluar, lo cual tiene relación con la validez de constructo. Así pues, se concluyó que mediante el uso de rúbricas se asegura una evaluación estandarizada de los conocimientos adquiridos por los estudiantes siempre y cuando se unifiquen criterios para obtener una evaluación equitativa.

En otra investigación, Lima-Rodríguez (2015) diseñó y validó el contenido de una rúbrica para evaluar competencias en estudiantes de enfermería a partir de su práctica clínica. En el estudio participaron 11 profesores implicados en la docencia y práctica clínica quienes fueron distribuidos por equipos de trabajo a través de sesiones presenciales y en línea. El procedimiento que se siguió fue el siguiente:

## Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias

1. Búsqueda de información. El equipo de profesores clarificaron la competencia que se deseaba medir y los instrumentos existentes para su evaluación mediante la revisión de base de datos nacionales e internacionales.
2. Selección de competencias. Con base en el punto anterior, se seleccionaron las competencias a evaluar y se establecieron criterios de consecución.
3. Diseño de las rúbricas de evaluación. Para ello se propusieron 65 ítems que respondían a los criterios que el alumno debería cumplir para considerar la consecución de dichos objetivos.
4. Elaboración de escala. A cada ítem se asignó una escala de estimación tipo Likert de tres puntos (*NC: No cumple el criterio; CP: Cumplimiento parcial del criterio; CT: Cumplimiento total del criterio*), que permitía dar un grado o intensidad al enunciado de cada ítem.
5. Validación de contenido. Se conformó un equipo de 20 expertos con conocimientos sobre el tema quienes fueron seleccionados mediante muestreo intencional. Se desarrolló un panel Delphi online como técnica de consenso mediante una encuesta electrónica. Así, cada experto asignó el grado de adecuación de cada ítem, sugirieron modificaciones y propusieron nuevos.
6. Diseño de portafolio de práctica. En este se recogía los trabajos e informes que el estudiante debía realizar durante las prácticas clínicas, incluyendo

las pruebas diseñadas y se daban instrucciones precisas sobre cómo realizarlos.

7. Guía de evaluación (profesores, tutores y estudiantes). En esta guía se explicaban las competencias a adquirir, los objetivos de aprendizaje, indicadores, las pruebas y momentos de evaluación, así como las rúbricas a utilizar en cada caso. Además, se indicaba cómo utilizar el portafolio de prácticas así como los documentos y evidencias de aprendizaje que el alumno debía incluir en el mismo. Tanto el portafolio como la guía de evaluación, se presentaban al inicio del periodo de prácticas, y estaban disponibles en la plataforma de enseñanza virtual.

Lima-Rodríguez concluyó que para construir una rúbrica es necesario realizar una revisión de los objetivos y contenidos que se desea estudiar. Con esto, se podrá definir las tareas de aprendizaje apropiadas así como identificar los criterios de evaluación y establecer criterios de consecución. Además, para el diseño de una rúbrica se requiere la participación de expertos para valorar el contenido. Al respecto, la técnica Delphi como método de consenso ofrece ventajas para realizar el jueceo por expertos.

Las rúbricas, además de ser útiles para evaluar habilidades y conocimientos, también son eficaces para evaluar prácticas. Al respecto, Company, Contero, Otey, y Plumed (2015) construyeron una rúbrica analítica para valorar las buenas prácticas en un entrenamiento de manejo 3D en ingenieros. El instrumento fue desarrollado mediante un panel de expertos quienes orientaron el diseño de los criterios de calidad. Las dimensiones de calidad fueron transformadas en las competencias de la rúbrica, las

cuales fueron valoradas a través de evidencias expresadas como afirmaciones. El procedimiento consistió en cuatro etapas:

#### Primera etapa

1. Organización de conceptos en dimensiones.
2. Consenso de expertos entre el programa y los conceptos elegidos.
3. El programa se organizó de acuerdo a las tareas estructuradas en serie.
4. Desarrollo de mapas de aserción para transmitir con mayor claridad el proceso.

#### Segunda etapa

5. Desarrollo de tareas específicas por uno de los expertos.
6. Otro experto resolvió la rúbrica satisfactoriamente.
7. Los miembros discutieron sobre las deficiencias y fortalezas de los criterios e improvisaron una versión.

#### Tercera etapa

8. Un miembro del equipo desarrolló pruebas para validar cuantitativamente aquellas afirmaciones que son difíciles de validar directamente.
9. Otro miembro utilizó las pruebas para validar las rúbricas.
10. Los dos miembros se reunieron para discutir las deficiencias y fortalezas de las pruebas y las mejoraron.
11. Después de repetir las etapas dos y tres hasta lograr las tareas y las rúbricas asociadas, estas se compilaron en los mapas de afirmaciones y se revisó la idoneidad de los criterios.

#### Cuarta etapa

12. Se solicitó a estudiantes de posgrado comentarios sobre la claridad y uso de las rúbricas.

13. Los comentarios recolectados se utilizaron por los expertos para producir una versión mejorada.

En otro estudio, Blommel y Abate (2007) desarrollaron y validaron una rúbrica para evaluar habilidades de crítica literaria en estudiantes de farmacología. Se realizaron cinco versiones del instrumento, estos se aplicaron a estudiantes y maestros en el periodo 2002-2006 en las jornadas de una revista de farmacología.

El primer paso para su elaboración fue identificar las categorías generales y criterios específicos esperados en la evaluación. Sin embargo, en la prueba de la primera versión se identificaron inconsistencias y falta de comprensión de los niveles de desempeño. Para rectificar las deficiencias, en el segundo paso, se realizó una segunda versión en la cual se amplió la descripción del contenido de las categorías para facilitar la identificación de los criterios de evaluación. En el tercer paso, se utilizaron las observaciones de los evaluadores y una breve encuesta para obtener retroalimentación de los estudiantes para realizar una tercera versión. Finalmente, se realizaron ajustes tanto en la versión cuatro como en la cinco respecto al diseño de criterios, formato y redacción. Además, se evaluó la confiabilidad de la rúbrica mediante el cálculo de los coeficientes de correlación intraclase para cada sección rúbrica de la y la puntuación total.

En conclusión, el desarrollo de una rúbrica apropiada debe proporcionar retroalimentación, debe ser fácil de usar y confiable para los evaluadores. Por lo tanto,

se requieren pruebas piloto para identificar y realizar las modificaciones necesarias para garantizar una evaluación consistente (Blommel y Abate, 2007). Esta afirmación complementa lo referido por otros autores, quienes señalaron la importancia de precisar los constructos y describir los niveles de desempeño con términos claros para evitar discrepancias en las evaluaciones (Arias y Maturana, 2005, citado en Picón, 2013; Gatica-Lara y Uribarren-Berrueta, 2013).

Asimismo López y Velasco (2013) diseñaron y demostraron evidencias de validez de una rúbrica analítica para la evaluación de ensayos como parte de un proyecto de alfabetización. En dicho proyecto se buscaba favorecer la práctica de la escritura en estudiantes de nivel superior y mediante la rúbrica obtener una evaluación objetiva. El procedimiento se llevó a cabo a través de los siguientes pasos:

1. Definición operacional del constructo. De manera colegiada, se realizó la definición del constructo a evaluar, para ello se identificaron ciertos rasgos del nivel de habilidad de los estudiantes, con base en la conceptualización determinada de escritura.
2. Revisión en colegiado de la teoría sobre elaboración de rúbricas. El equipo de alfabetización académica examinó y discutió los procesos sugeridos por Moskal (2003), Boston (2002) y Perlman (2002) para la construcción del instrumento y la conveniencia de diseñar una rúbrica analítica para la evaluación del constructo.
3. Planteamiento de metas y objetivos. Se estableció evaluar la habilidad de los estudiantes para escribir ensayos y como objetivos concretos identificar su

- nivel de habilidad, con respecto a cada uno de los rasgos que constituyen la estructura formal de un ensayo, la estructura de significado en sus niveles micro y macro, la adecuación del texto a la situación comunicativa.
4. Determinación de dimensiones de la escritura a evaluar. Se establecieron las siguientes dimensiones: la estructura formal, la estructura semántica y la adecuación del texto a la situación retórica.
  5. Selección de rasgos de cada dimensión a evaluar. Estos se realizaron con base en las dimensiones previamente descritas y con las teorías correspondientes.
  6. Establecimiento de criterios de ejecución. Para ello se siguieron las siguientes recomendaciones: (a) enlistar los aspectos importantes de la ejecución; (b) limitar el número de elementos, de manera que puedan ser observables; (c) expresar los criterios en términos de características observables; (d) evitar el uso de términos ambiguos y (e) colocar los elementos en orden en el que posiblemente se observen.
  7. Identificación de atributos subyacentes. Fueron planteados de acuerdo con las teorías de Dijk (1978) y Cassany (1997).
  8. Definición conceptual de los términos fundamentales sobre los que se basó el instrumento. Se realizaron con base en Dijk (1978) Halliday y Hasan (1976) y Cassany (1997).
  9. Elaboración de escala valorativa. Se establecieron niveles de logro en un continuo del 0 al 4 sobre la base del nivel de presencia de cada rasgo.

10. Redacción de descriptores de cada nivel de la escala para cada rasgo. Se definieron con claridad y precisión la diferencia entre un nivel de ejecución y el siguiente, respetando el tipo de medida establecida para cada rasgo.
11. Aplicaciones piloto. Debido a la extensión de la rúbrica, fue necesario cuatro aplicaciones piloto. En cada una, se evaluaron los ensayos escritos por 9 estudiantes en cuatro materias diferentes, lo que dio una suma de 36 textos por aplicación piloto y un total de 144 por las cuatro ocasiones.
12. Modificación de ítems a partir de la discusión de puntos de desacuerdo entre los evaluadores. Se realizó con base en el análisis de los puntos de desacuerdo entre los evaluadores lo cual puso de manifiesto la imprecisión de algunos descriptores.
13. Realización de pruebas estadísticas para establecer propiedades psicométricas del instrumento y obtener evidencias de validez y confiabilidad. se llevaron a cabo las siguientes pruebas: (a) análisis discriminante del método Lambda de Wilks, (b) análisis factorial de componentes principales con rotación Varimax y (c) análisis de confiabilidad (*alpha* de Cronbach).

Con base en las conclusiones reportadas, se consideró que el instrumento es confiable para conocer el grado de habilidad de los estudiantes para la elaboración de ensayos. Las evidencias de validez y confiabilidad que se obtuvieron sustentan la pertinencia de utilizar el instrumento para disminuir el nivel de subjetividad de los evaluadores de este tipo de textos (López y Velazco, 2013).

### 2.5. Validez

Las rúbricas son instrumentos de evaluación utilizados con frecuencia dentro del salón de clases para la evaluación y la autoevaluación de competencias, habilidades o conocimientos de manera sistemática (Reddy y Andrade, 2010). La utilidad y características de las rúbricas se relacionan con las propiedades de los test referidas en los *Standards for Educational and Psychological Tests* (SEPT) de 1999. En éstos se describe que un test es “un instrumento evaluativo o procedimiento sistemático con el que se obtiene una muestra de la conducta de los examinados en un dominio especificado y posteriormente es evaluada y puntuada usando un procedimiento estandarizado” (p.3). Martínez-Arias (2014) señaló que el uso de los resultados de las evaluaciones mediante estos instrumentos, implica consecuencias para la población evaluada ya que a partir de los resultados se toman decisiones. Por ende, su construcción deberá estar regulada por los estándares vigentes de la *American Educational Research Association* (AERA), *American Psychological Association* (APA) y *National Council on Measurement Education* (NCME).

La validez representa un elemento fundamental de los instrumentos de evaluación ya que aportan evidencias que sustentan las inferencias derivadas a partir de los resultados del instrumento (Martinez-Arias, 2014). De acuerdo con los SEPT, la validez se entiende como “el grado en que la evidencia y teoría soportan la interpretación de las puntuaciones del test para el uso pretendido” (AERA, APA y NCME, 2014, p. 11) además es definido como un concepto unitario, la validez del constructo.

Andrade (2005) señaló que las evidencias de validez son parte esencial de la construcción de las rúbricas y se relacionan con la claridad en la descripción de los constructos a evaluar y de los niveles de desempeño. Asimismo, existen diversas fuentes de evidencia de validez como métodos para determinarla al respecto Martínez-Arias (2014, p. 31-33) los sintetizó en: (a) contenido del test, (b) procesos de respuesta, (c) estructura interna del test, (d) relaciones con otras variables y (c) consecuencias derivadas del uso de los test.

En la mayoría de los casos son los docentes quienes diseñan estos instrumentos y deben hacerlo con base en lo que les interese evaluar del desempeño de sus estudiantes y la intención del uso de los resultados (Mertler, 2001). Pese a que las rúbricas son parte de la práctica cotidiana dentro del salón de clase, los procedimientos para su elaboración y obtención de fuentes de validez suelen ser soslayados. De acuerdo con Boston (2002) la validez de las rúbricas debe establecerse al igual que los demás instrumentos de evaluación, es decir, en términos de contenido, constructo, criterio y pruebas consecuentes.

De acuerdo con lo reportado por algunos autores (Cano-García, 2015; Jonsson y Svingby; 2007; Reddy y Andrade, 2010), quienes han realizado revisiones detalladas sobre estudios enfocados al uso y diseño de rúbricas, son pocos los estudios enfocados a obtener evidencias de validez de las mismas. En la revisión de Jonsson y Svingby (2007) identificaron 25 de 75 estudios analizados, posteriores al año 2000, en los cuales se reportaron diversas fuentes de validez, siendo los estudios sobre validez de contenido y externa los de mayor número, 10 y 15 respectivamente. Por su parte, Reddy y Andrade

(2010) identificaron tres estudios relacionados a la validez, sin embargo, estos no incluyen la validez de contenido, de constructo ni de criterio. Finalmente, Cano-García (2015, p. 1) concluyó que es necesario “acompañar el uso de las rúbricas con procesos de formación y con análisis de fiabilidad y validez del instrumento”.

Debido a la escasa literatura sobre evidencias de validez de las rúbricas, en la presente investigación se identificaron algunos estudios en los cuales se reportan de manera directa e indirecta métodos para la obtención de validez. De tal manera que se realizó una revisión en bases de datos especializadas (WOS, EBSCO, SCIELO) con los siguientes criterios de búsqueda: rúbricas, *rubrics*, validez y rúbricas, *validity and rubrics*. A partir de la búsqueda se seleccionaron nueve investigaciones (Contreras Valenzuela et al., 2013; Blommel y Abate, 2007; García- Ros, 2011; García-Ros et al., 2012; Lima-Rodríguez, Lima-Serrano, Ponce-González, y Guerra-Martín, 2015; López y Velasco, 2013; Musial et al., 2007; Verano-Tacoronte, González-Betancor, Bolívar-Cruz, Fernández-Monroy, y Galván-Sánchez, 2016). El criterio para la selección fue que en éstos se reportaran los procedimientos para obtener fuentes de validez. Asimismo, estos estudios se clasificaron por fuente de validez (de contenido, de constructo, convergencia y predictiva) tal como se muestra en la tabla 1. Cabe señalar que esta clasificación es con base en la definición de validez expuesta por los autores de los estudios analizados.

Tabla 1  
*Investigaciones sobre evidencias de validez de rúbricas*

Evidencia de validez	Técnica	Estudio
Contenido	Selección y definición de las competencias a evaluar mediante grupo de expertos y aplicación de pruebas piloto.	Musial et al., 2007. Developing a Scoring Rubric for Resident Research Presentations: A Pilot Study.
	Técnica de jueceo por expertos y pruebas piloto	Blommel y Abate, 2007. Rubric to assess critical literature evaluation skills.

## Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias

	Elaboración de criterios mediante grupo de expertos.	García-Ros et al., 2012. Diseño y utilización de rúbricas en la enseñanza universitaria: una aplicación en la titulación de Psicología.
	Juicio de expertos a través de protocolo de validación.	Gordillo y Vázquez, s.f. Validez y fiabilidad de la rúbrica.
	Juicio de expertos y dos aplicaciones piloto.	Contreras Valenzuela et al., 2013. IOC, un instrumento para cualificar desempeño docente en aula: Su generación y validación.
	Revisión teórica y participación de expertos para especificar el constructo, los criterios de evaluación y la definición de los descriptores.	López y Velasco, 2013. La construcción y evidencias de validez de una rúbrica analítica para la evaluación de ensayos.
	Consenso a través de <i>Panel Delphi</i>	Lima-Rodríguez, Lima-Serrano, Ponce-González, y Guerra-Martín, 2015. Diseño y validación de contenido de rúbricas para evaluar las competencias prácticas en estudiantes de Enfermería.
	Revisión teórica y participación de expertos.	Verano-Tacoronte, González-Betancor, Bolívar-Cruz, Fernández-Monroy, y Galván-Sánchez, 2016. Valoración de la competencia de comunicación oral de estudiantes universitarios a través de una rúbrica fiable y válida.
Constructo	Análisis factorial de componentes principales y rotación Varimax.	López y Velasco, 2013. La construcción y evidencias de validez de una rúbrica analítica para la evaluación de ensayos.
	Análisis factorial de componentes principales (AF) y rotación Promax y análisis factorial exploratorio.	Muñoz y Valenzuela, 2015. Características Psicométricas de una Rúbrica para Evaluar Expresión Escrita a Nivel Universitario.
Predictiva	Análisis discriminante de <i>Lambda Wilks</i>	López y Velasco, 2013. La construcción y evidencias de validez de una rúbrica analítica para la evaluación de ensayos.
Convergente	Se determinó el grado de acuerdo o convergencia a través del estadístico <i>kappa</i> y se estimó el coeficiente de Correlación de <i>Pearson</i> .	García-Ros, 2011. Análisis y validación de una rúbrica para evaluar habilidades de presentación oral en contextos universitarios.

Fuente. Elaboración propia.

Con base en la revisión de los estudios enlistados en la tabla anterior, se identificaron diferencias con relación al orden metodológico para obtener las fuentes de validez. Por lo tanto se realizó una descripción, de manera sintetizada, de los aspectos considerados sobresalientes en los procedimientos establecidos en los estudios de acuerdo con la fuente de validez que se obtuvo.

**2.5.1. Validez de contenido.** Musial et al. (2007) diseñaron una rúbrica para evaluar el desempeño de médicos residentes en presentaciones orales. Para obtener evidencias de validez de la rúbrica consideraron el estándar 1.7 (SEPT, 2014), en el cual se establece que “cuando la validez es obtenida mediante la opinión y juicios de expertos tanto su calidad como experiencia profesional debe ser probada y descrita como parte del procedimiento”. Con base en lo anterior, seleccionaron a un equipo multidisciplinario de expertos para diseñar las competencias a evaluar con base lo establecido por la *Accreditation Council for Graduate Medical Education (ACGM)*. Posteriormente las competencias se establecieron mediante consenso de los expertos, quienes elaboraron la escala de calificación después de cinco sesiones de análisis y correcciones. Finalmente la rúbrica se realizó una prueba piloto para estimar la confiabilidad de la misma. Con base en los resultados preliminares reportados, los autores concluyeron que la rúbrica presentó evidencias de validez de contenido a través de la participación del equipo multidisciplinario quienes seleccionaron y definieron las competencias a evaluar de manera detallada. Por lo que puede ser utilizada en el contexto de las evaluaciones orales. Sin embargo, reconocieron que la rúbrica no es válida a nivel general y que no presenta validez predictiva, por lo que se considera necesario estudios posteriores para extender su uso en otros contextos.

En otro estudio, Blommel y Abate (2007) desarrollaron una rúbrica para valorar, en estudiantes de farmacología, habilidades relacionadas con la evaluación crítica literaria. Para ello los autores consideraron pertinente que para el procedimiento de diseño, validez de contenido y confiabilidad se requería de la participación de expertos y estudiantes para realimentar el procedimiento. Así pues, realizaron cinco estudios

piloto en los cuales se aplicó la rúbrica; posteriormente, dos expertos y los estudiantes participantes realimentaron con sus opiniones y sugerencias en relación al uso del instrumento, la claridad de los ítems y los niveles de desempeño para asegurar la validez de la rúbrica. Por último, se realizaron adecuaciones al instrumento y se aplicaron pruebas piloto, las cuales derivaron en cinco versiones del instrumento. A continuación se describen algunas de las observaciones y ajustes más relevantes realizadas en cada versión:

*Versión 1.* Se identificaron deficiencias relacionadas al orden lógico de los criterios y a los niveles de desempeño, ya que los números resultaban difíciles de cuantificar.

*Versión 2.* Se ampliaron las descripciones y el contenido de las categorías para facilitar su identificación y algunas de ellas se dividieron para su mejor comprensión. Además, se incluyó para los evaluadores un *Check box* al lado de cada ítem para realizar observaciones. Posterior a las adecuaciones, el instrumento se utilizó por los estudiantes, quienes señalaron presentar dificultades en diferenciar entre algunas categorías.

*Versión 3.* En esta versión se incluyó una encuesta de opinión para recabar la realimentación de los estudiantes, en los resultados, se observó que los estudiantes no entendían con claridad algunas categorías de la rúbrica.

*Versión 4.* Miembros de la facultad desarrollaron una serie de preguntas para ayudar a los estudiantes a identificar las consideraciones clave que se debían tomar en cuenta para la redefinición de las categorías.

*Versión 5.* Se elaboró la rúbrica final con base en la realimentación de los estudiantes y de los evaluadores.

Blommel y Abate (2007) mencionaron que para diseñar una rúbrica se requiere de múltiples pruebas piloto, seguidas de modificaciones para mejorar y redefinir el instrumento, lo anterior con el objetivo de lograr una buena evaluación.

García-Ross et al., (2012) presentaron el diseño de tres rúbricas para evaluar competencias de titulación en estudiantes de psicología. Asimismo, reportaron el procedimiento para obtener evidencias de validez de contenido para cada una de las rúbricas:

*Rúbrica para evaluar competencias de indagación científica.* Se elaboró la estructura de la rúbrica a través de la revisión de competencias establecidas en las materias del plan curricular de la licenciatura en psicología. Posteriormente, se llevaron a cabo tres sesiones de trabajo con un equipo de expertos (profesores de la universidad) en las cuales se discutió la estructura, la organización y el contenido específico considerado en la rúbrica hasta establecer la versión final de la misma.

*Rúbrica para evaluar habilidades de presentaciones orales en contexto universitario.* Para el desarrollo de los criterios de evaluación y descripciones de niveles de ejecución, se consideraron los resultados y conclusiones de investigaciones previas

sobre presentaciones orales (García-Ros y Pérez-González, 2011; García-Ross, 2011, citado en García-Ross et al. ,2012 p. 10).

*Rúbrica de competencias básicas en el practicum de titulación.* Se realizaron sesiones de trabajo con un grupo de expertos en la práctica de titulación con el objetivo de (a) discutir y se seleccionar las competencias específicas y generales de la rúbrica, (b) definir los criterios e indicadores, (c) especificar los niveles y los estándares de ejecución para cada uno de los criterios y (d) realizar una propuesta de estructura, de organización y de contenidos de las memorias que facilitarían la inclusión e integración de los criterios.

De acuerdo con la revisión del estudio de García-Ross et al. (2012), se enfatiza la participación de jueces expertos como técnica principal para obtener de validez de contenido. A manera de conclusión, reportaron que es importante la unificación de criterios sobre qué y cómo evaluar las competencias, para ello, las rúbricas se consideran instrumentos válidos, por los expertos (profesores), para garantizar la evaluación de dichas competencias. No obstante, sostienen que es necesario realizar análisis sobre confiabilidad intrajueces e interjueces de las rúbricas.

Siguiendo con la técnica de jueceo por expertos para obtener evidencias de validez de contenido, Gordillo y Vázquez, s.f. realizaron un protocolo de validación para asegurar la validez de una rúbrica de evaluación de aprendizajes en estudiantes universitarios. El protocolo se integró por “con ocho criterios para comprobar el grado de ajuste de la rúbrica en función de una escala de puntuación que iba de 0 (mínima) a 5 (máxima)” (Gordillo y Vázquez, s.f. p.4). Los criterios fueron: (a) pertinencia, (b)

coherencia, (c) estructuración, (d) claridad, (e) suficiencia, (f) aplicabilidad, (g) cobertura y (h) oportunidad.

Los resultados de las puntuaciones obtenidas para cada criterio se expresaron en porcentajes de frecuencia. De tal manera que, los ocho criterios obtuvieron puntuaciones altas (entre cuatro y cinco puntos), siendo los criterios de pertinencia y oportunidad los mejor valorados (cinco puntos) con 71.4 % y 85.7% respectivamente. En conclusión, los autores afirmaron que con base en las puntuaciones obtenidas del juicio de los expertos, la rúbrica es válida para evaluar los aprendizajes de los estudiantes en el contexto universitario. De acuerdo con el procedimiento, se considera que la metodología utilizada contribuye a obtener datos precisos sobre la opinión de los jueces participantes, sin embargo, falta especificar el análisis de los datos de manera detallada.

Por su parte, Contreras -Valenzuela et al.(2013) realizaron adaptaciones de la rúbrica *Tsang-Hester Observation Rubric*, la cual es utilizada para evaluar el desempeño docente en el aula. Como primer paso, la rúbrica fue revisada por un experto en el idioma inglés, de esta manera se aseguró la calidad de la traducción y el sentido de los criterios y niveles de la rúbrica, de acuerdo con la versión original. Después, la versión traducida fue adaptada por un grupo de expertos de acuerdo con los siguientes criterios: “a) que el lenguaje utilizado fuese comprensible en el contexto nacional, b) que lo expresado en los niveles de las rúbricas de cada criterio fuese realmente susceptible de ser observado” (Contreras -Valenzuela et al.,2013,p.6).

Posterior a las adaptaciones realizadas a la rúbrica, se procedió a validar su contenido. Para ello se realizaron dos evaluaciones por expertos y dos pruebas piloto:

*Primera evaluación por jueceo de experto.* Para evaluar la primera versión, la rúbrica fue revisada por un experto en evaluación docente para confirmar la coherencia, y pertinencia de las dimensiones, los criterios y niveles de desempeño. A partir de las sugerencias, se elaboró la segunda versión.

*Primera aplicación piloto.* De manera voluntaria, participaron seis docentes novatos a quienes se les evaluó su desempeño en el aula a través de la segunda versión de la rúbrica. Las evaluaciones fueron realizadas por dos investigadores del estudio, posterior a las evaluaciones, se realizó un análisis de la experiencia sobre el uso de la rúbrica con relación a su uso y las dificultades identificadas. Así, se concluyó, que era necesario incluir herramientas complementarias tales como: entrevista, matriz para notas de campo y hoja de registro de las puntuaciones para cada dimensión de la rúbrica. Además, se realizaron ajustes a la precisión del lenguaje, desagregación de componentes e incorporación de una nueva dimensión. A partir de estas modificaciones, se obtuvo la tercera versión de la rúbrica.

*Segunda evaluación por experto.* Un experto realizó observaciones a cerca de la comprensión, la pertinencia de los ajustes realizados (al instrumento original) de acuerdo con el contexto nacional y por último revisó la calidad de la escala de los niveles descritos en la rúbricas. Después de atender las sugerencias y comentarios, se elaboró la cuarta versión.

*Segunda aplicación piloto.* En esta aplicación participaron siete evaluadores, los cuales fueron capacitados previamente en el uso de la rúbrica. Posteriormente, observaron una clase grabada en video (la misma para cada evaluador) y evaluaron el

desempeño de docente con la cuarta versión de la rúbrica. Después se realizaron ajustes con base en las observaciones de los evaluadores sobre la coherencia y comprensión de la rúbrica. Por último, se realizó la versión final del instrumento.

De acuerdo con las conclusiones de los autores el procedimiento que se realizó para adaptar la rúbrica, permitió establecer que es válida para el contexto en el cual se llevó a cabo la investigación. En síntesis, es posible afirmar que éste procedimiento presenta un orden lógico y detallado que permite comprender el proceso metodológico que implica obtener evidencias de validez de contenido de una rúbrica.

López y Velasco (2013) diseñaron una rúbrica para la evaluación de ensayos de estudiantes universitarios. Si bien, en su estudio no se reportan evidencias de validez de contenido, la rúbrica sí la presenta, de manera implícita. Esta afirmación se sustenta con base en la revisión del procedimiento, en el cual se describe como parte de la metodología, para la construcción de la rúbrica, la definición del constructo a evaluar de acuerdo con la revisión teórica realizada y el establecimiento de criterios de evaluación, niveles de desempeño y descriptores con base en la opinión de un grupo de expertos. Así pues, aunque en este estudio se reporta validez de constructo y validez predictiva, es necesario mencionar y detallar, en primera instancia, el procedimiento para la obtención de validez de contenido. Ya que de acuerdo con lo señalado en los SEPT (2014), el primer paso en el proceso de validez de un instrumento es especificar el constructo que se desea medir (AERA, APA y NCME, 2014).

En otro estudio, Lima-Rodríguez et al. (2015), utilizaron la técnica de consenso por expertos (20) a través del Panel Delphi *on-line* para obtener validez de contenido de

un sistema de evaluación de competencias de estudiantes de enfermería, integrado por ocho rúbricas. Para la selección de los expertos, se tomó como criterio de inclusión que estos fueran enfermeros/as con más de cinco años de experiencia profesional y con conocimiento sobre los temas de las rúbricas. Así, el procedimiento fue el siguiente:

1. Informar a los expertos seleccionados sobre los objetivos del estudio y solicitud de consentimiento para su participación.
2. De manera individual, cada experto designó el grado de adecuación de cada ítem (elementos que debía contener el portafolio de trabajo), de acuerdo a una puntuación del 1 al 5 (1 menor adecuación, 5 mayor adecuación). Además, podían sugerir modificaciones o proponer otros para ser validados.
3. Una vez obtenidas las respuestas del punto anterior, se realizó un análisis descriptivo de las variables ordinales calculando las medidas de tendencia central y los porcentajes de las puntuaciones obtenidas del jueceo.
4. Se realizaron tres rondas de valoración por el panel de expertos en las cuales se estimaron y rechazaron los ítems con base en criterios de estimación y desestimación. Para desestimar un ítem se consideró: media menor a 3.5, y mediana menor a 3. Para estimar un ítem: porcentaje de valoraciones (4-5) mayor o igual a 80 %. Si no cumplían este criterio, se consideraron los que tenían una desviación típica igual o inferior a 0,90 (63,64 % de la desviación típica uniforme).
5. Los ítems rechazados en una ronda, se modificaron con las aportaciones de los expertos y pasaban a la siguiente ronda, hasta alcanzar los criterios de estimación. Además, posterior a cada ronda, se envió a cada miembro del panel

de expertos, un informe con los resultados alcanzados hasta ese momento para su realimentación.

6. Finalmente, la validación de contenido se aprobó por el Comité Ético de Experimentación de la Universidad de Sevilla una vez que se cumplió con lo requerido.

En las conclusiones reportadas se mencionó que la realimentación controlada de los expertos como procedimiento para obtener validez de contenido contribuyó a mejorar la calidad de los ítems, no obstante, a partir de la revisión realizada, como parte de la presente investigación, se consideró que los pasos a seguir no se especificaron de manera clara y precisa, por lo que su comprensión resulta complejo.

Al igual que en otras investigaciones, Verano-Tacoronte et al., (2016) realizaron una revisión teórica para la elaboración de una rúbrica para evaluar presentaciones orales en universitarios. Sin embargo, una de las características sobresalientes fue que se incluyeron atributos observables en los criterios de evaluación y en los niveles de ejecución para facilitar y clarificar a los evaluados lo que se espera de ellos. Por lo tanto, la validez de contenido se obtuvo a partir de la teoría y de lo observable. Posterior a la elaboración de la rúbrica ésta fue sometida a revisión por un grupo de expertos. Si bien resulta interesante la incorporación de atributos observables para garantizar la validez del contenido de la rúbrica, el procedimiento no se detalla a profundidad.

De acuerdo con la revisión de los estudios para obtener evidencias de validez de contenido de las rúbricas, se concluye que la técnica de jueceo por expertos y la revisión teórica son utilizadas de manera frecuente. Sin embargo, en algunos estudios no se describe de manera detalla, ni en su totalidad los procedimientos.

**2.5.2. Validez de constructo.** López y Velasco (2013) además de reportar evidencias de validez de contenido de la rúbrica para evaluar ensayos, también obtuvieron validez de constructo. Para ello realizaron un análisis factorial mediante el método de componentes principales con rotación varimax. De acuerdo con las conclusiones, a partir del análisis se arrojaron cuatro factores (estructuras formales y semántica global, adecuación, estructura lógica del documento y elementos relativos al tema), los cuales no correspondieron con exactitud a la distribución de los atributos de la rúbrica. No obstante, se consideró que la rúbrica “responde operativamente al constructo definido teóricamente” (López y Velasco, 2013, p. 16).

Por otra parte, (Muñoz & Valenzuela, 2015) validaron el constructo de una rúbrica para evaluar la expresión escrita a nivel universitario a través del análisis y el contraste de la estructura factorial por componentes principales, con relación a los fundamentos teóricos en los cuales se basó el diseño de la rúbrica. De acuerdo con lo reportado, el constructo “escritura académica” se concibió como una estructura doble. Asimismo, este constructo se caracterizó por aspectos locales y globales del texto. Así, en los aspectos locales se identificaron elementos tales como: la ortografía acentual y puntual, el vocabulario y la coherencia entre las frases). Mientras que en el aspecto de orden global se encontraron: la unidad temática, la organización del texto, la introducción, el desarrollo del tema, la conclusión y la coherencia entre párrafos. Así, se estableció que la relación entre ambos aspectos da como resultado un texto académico caracterizado por una escritura coherente y cohesiva.

Finalmente, concluyeron que derivado de los análisis para determinar las propiedades psicométricas de la rúbrica, específicamente la validez de constructo, la rúbrica para evaluar la expresión escrita a nivel universitario es un instrumento válido, confiable y con altos niveles de discriminación.

**2.5.3. Validez predictiva.** Con relación a la validez predictiva en las rúbricas, López y Velasco (2013) consideraron pertinente obtener evidencias de validez predictiva de la rúbrica para evaluar ensayos a través del análisis discriminante –Lambda Wilks-. En análisis se realizó con base en las categorías establecidas del nivel de escritura de la rúbrica. Estos fueron: (a) inicial, (b) en desarrollo y (c) logrado. A través del análisis, emergieron cuatro variables que distinguen a los escritores de cada nivel: 1) síntesis de los argumentos (en la conclusión), 2) cohesión lineal, 3) sustento bibliográfico (en el cuerpo o desarrollo) y 4) identificación del significado global del texto.

De acuerdo con los coeficientes estandarizados, la síntesis de los argumentos y el sustento bibliográfico resultaron ser las dos variables con mayor peso para distinguir a los escritores iniciales de los escritores más avanzados. En los resultados se reportó que el instrumento total obtuvo un 91% de clasificación correcta, es decir, un nivel de predicción alto. Basado en esto, concluyeron que es pertinente utilizar la rúbrica para reducir el nivel de subjetividad de los evaluadores de este tipo de textos.

**2.5.4. Validez convergente.** En el estudio de García-Ros (2011) se analizó validez convergente de una rúbrica para evaluar habilidades de presentación oral en contextos universitarios. El análisis de convergencia se realizó entre la evaluación sumativa de los docentes y las evaluaciones entre pares a través de la rúbrica. Así, la

rúbrica se utilizó para evaluar una serie de presentaciones orales, las cuales se evaluaron por profesores y estudiantes. Posteriormente el grado de acuerdo en la aplicación de ambos evaluadores se estimó mediante dos procedimientos (conservador y laxo) a través del estadístico *kappa*. Asimismo se calculó el nivel de asociación entre las valoraciones globales de los profesores y por los pares (estudiantes) mediante el coeficiente de *correlación de Pearson*. De acuerdo con las conclusiones, el grado de acuerdo entre las valoraciones fue significativa.

En los procedimientos analizados hasta el momento, se observa que existe poca especificación acerca de las fuentes de evidencias de validez que se obtienen de las rúbricas, además éstas no se reportan de manera directa. Asimismo, fue posible identificar que la mayoría de dichos procedimientos se enfocan en obtener evidencias de validez de contenido mediante la participación de expertos y apegándose a teorías que sustentan el constructo que se desea evaluar.

Así pues, en algunos estudios existe ambigüedad con relación a la descripción y manejo de las variables involucradas en el proceso de obtención de validez. En consecuencia, es necesario reportar de manera clara y detallada estos procedimientos para contribuir de manera significativa a la utilización de la rúbrica como un instrumento de evaluación de calidad y con propiedades psicométricas que respalden los resultados que emergen de éstas.

### *3. Contexto de la investigación*

Este estudio forma parte de la investigación *Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California* (Chaparro et al., 2015) cuyo propósito fue identificar las prácticas escolares presentes en escuelas secundarias que obtuvieron buenos resultados académicos en la Evaluación Nacional de Logros Académicos en Centros Escolares (ENLACE) del 2010 y 2011. Para ello se realizó una réplica del estudio *Caracterización y buenas prácticas en los centros escolares de alto valor añadido* (Lizasoain et al., 2012) llevado a cabo en la Universidad del País Vasco. Dado que estas dos investigaciones enmarcan el presente estudio se consideró pertinente describir, de manera breve, cada uno de los proyectos. Cabe señalar que se realizó una descripción más específica de la investigación *Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California* (Chaparro et al., 2015) puesto que los resultados derivados de la misma fueron el insumo para desarrollar la presente rúbrica.

### **3.1. Investigación del País Vasco: Caracterización y buenas prácticas en los centros escolares de alto valor añadido**

Con base en lo reportado por Lizasoain et al. (2012) el objetivo general del proyecto fue

“Conocer la verdadera aportación de la escuela al conjunto de aprendizajes y logro de competencias que se miden en las pruebas de la Evaluación Diagnóstica para después elaborar un catálogo de buenas prácticas que puedan ser transferibles a otros centros para contribuir así a la mejora del sistema educativo del País Vasco”(pág. 7).

Para cumplir con el propósito, se llevó a cabo un análisis estadístico a través de técnicas de regresión múltiple multinivel para así determinar el valor añadido (lo que aporta la escuela) a partir del rendimiento medio de las escuelas en las Evaluaciones Diagnósticas realizadas en 2009 y 2010. En la segunda fase se recopiló información a través de entrevistas, grupos de discusión, observación y cuestionarios a personal de la comunidad escolar de las escuelas seleccionadas (inspectores de zona, equipos directivos, profesores, entre otros) (Lizasoain et al., 2012). Dicha información se analizó y se transcribió selectivamente en función de 14 categorías: (a) Percepción global del centro, (b) Proyecto, planes y formación (c) Metodología didáctica y material de enseñanza (d) Atención a la diversidad (e) Seguimiento del alumnado, orientación y tutoría (f) Evaluación del alumnado (g) Gestión de los tiempos (h) Liderazgo y equipo directivo (i) Modelos de gestión y organización, el centro como organización (j) Coordinación (k) Sentimiento de pertenencia al centro (l) Evaluación de los docentes, del propio centro, de programas y actividades y uso de la ED (m) Clima y convivencia y (n) otros. De manera general, se esperaba que al identificar las buenas prácticas de las mejores escuelas estas podrían incorporarse en proyectos de mejora escolar dirigidas a todas las escuelas a través de un *catálogo de buenas prácticas* (Angulo y Lizasoain, 2014).

#### **3.2. Investigación mexicana: Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California**

El objetivo de la investigación *Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California* (Chaparro et al., 20015) fue identificar las prácticas escolares de escuelas secundarias de alto valor añadido en Baja California. Para ello se llevaron a

cabo entrevistas semiestructuradas a 111 informantes clave (ver tabla 2) de 30 escuelas seleccionadas previamente con base en los puntajes obtenidos en La Evaluación Nacional de Logros Académicos en Centros Escolares (ENLACE) del 2010 y 2011. La selección de las escuelas participantes y de los informantes clave se realizó en la primera etapa de su investigación llamada *Identificación de escuelas de alto y bajo residuo, con base en puntajes de ENLACE y variables de contexto* (Martínez, Chaparro Lizasoain, Caso, y Urias, 2014). Cabe señalar que estos dos estudios constituyeron una réplica de la investigación realizada en el País Vasco.

**3.2.1. Procedimiento para la selección de las escuelas participantes.** Se utilizó una muestra de estudiantes, docentes y directores escolares derivada de la *Estrategia Evaluativa Integral* realizada en 2010 y 2011 por la Unidad de Evaluación Educativa de la Universidad Autónoma de Baja California en 85 escuelas secundarias. La muestra se integró por 6762 estudiantes de la asignatura de Matemáticas y 7036 de Español, 279 docentes de Matemáticas, 285 docentes de Español y 71 directivos. Posteriormente la muestra se redujo a 65 planteles de acuerdo con los procedimientos estadísticos realizados.

Se consideraron las variables de contexto recuperadas de los cuestionarios de contexto aplicados en la *Estrategia Evaluativa Integral 2010* (Caso, Díaz, Chaparro y Luzanilla, 2011, citado en Martínez et al., 2012) y 2011 (Caso, Chaparro, Díaz y Luzanilla, 2012, citado en Martínez et al., 2012) entre las que se encuentran: oportunidades de aprendizaje con que cuenta el estudiante, clima escolar, estrategias de aprendizaje, autorregulación académica, oportunidades de aprendizaje que brinda el docente y oportunidades de aprendizaje que aporta el director (Martínez et al., 2012).

Asimismo, se aplicaron tres modelos jerárquicos lineales de dos niveles (estudiante y escuela-turno): Modelo de ganancia residual, Modelo brutal y Modelo contextualizado sin ganancia; siendo este último el que permitió la selección de las escuelas. En cada modelo se introdujeron las variables contextuales como covariables de efectos fijos en el primer nivel y se consideraron tres criterios:

1. Que de las ocho medidas obtenidas, se dispusiera al menos de seis medidas para la escuela.
2. Hacía referencia al número de veces que la escuela ocupó las posiciones superiores o inferiores en los residuos. Esto se tomó como una medida de valor añadido de la escuela. En este sentido, para las escuelas de residuo alto y positivo se estableció como criterio de selección que estas situaran en el tercio superior de las puntuaciones y para la selección de las escuelas de residuo negativo, el criterio se aplicó de manera inversa, es decir, aquellas que se encontraban en el tercio inferior de las puntuaciones.
3. El último criterio se determinó en torno al percentil medio que ocupaban las escuelas en función de la distribución de los residuos. Así, las escuelas de residuo positivo y negativo seleccionadas debían estar ubicadas en el tercio superior inferior e inferior de la distribución respectivamente. Para calcular el percentil medio de cada escuela se aplicó la siguiente ecuación:

$$\textit{Percentil} = \frac{\sum_{j=1}^m \frac{\textit{Rango}_{ij}}{n_j}}{m}$$

En donde:

m= Número de modelos en los que estaba presente la escuela (*i*)

$Rango_i$  = posición directa que ocupa la escuela  $i$  en el modelo ( $j$ ) en función del residuo obtenido

$n_i$  = Número de escuelas presentes en el modelo ( $j$ )

De esta manera, toda escuela seleccionada debía ubicarse de acuerdo a los residuos obtenidos en un percentil inferior o igual a 33 o superior o igual a 66. Posterior a la aplicación de estos criterios de selección, fueron elegidas 30 escuelas, 16 consideradas con alto valor añadido (residuo alto) y 14 con bajo residuo (Martínez et al., 2012).

**3.2.2. Procedimiento para el diseño y aplicación de entrevistas.** De acuerdo con lo señalado por Chaparro et al. (2015), posterior a la selección de las escuelas se diseñaron guiones de entrevistas para obtener información de las escuelas a través de entrevistas dirigidas a informantes clave. Se consideró pertinente adaptar los guiones de las entrevistas utilizadas en la investigación Caracterización y buenas prácticas en los centros escolares de alto valor añadido (Lizasoain et al., 2012), así como las categorías utilizadas para el análisis de la información. Las modificaciones realizadas fueron respecto al idioma y características del Sistema Educativo Mexicano, y de manera específica, para las categorías también se consideraron factores de Eficacia Escolar y Mejora Escolar reportados en la literatura. De esta manera, se establecieron 14 categorías: (a) Percepción general de la escuela, (b) Entorno e historia, (c) Gestión y organización, (d) Equipo directivo, (e) Recursos, (f) Programas, planes y proyectos, (g) Formación y compromiso del personal de la escuela, (h) Evaluación, (i) Evaluación del aprendizaje de los alumnos, (j) Metodología didáctica y materiales de enseñanza, (k)

Atención a la diversidad, (l) Convivencia y clima escolar, (m) Relación bidireccional entre la escuela, la familia, la comunidad y redes de apoyo y, (n) Otras. Además, se incluyeron registros de observación en los cuales se incluyeron aspectos tales como la dinámica y organización del personal, infraestructura, aspecto físico de la escuela, ambiente y convivencia.

*Selección de los informantes clave.* El criterio establecido para su participación fue que estos tuvieran al menos seis meses de antigüedad en la zona escolar o en la escuela. Con relación a los informantes de la supervisión escolar la muestra se conformó de supervisores de zona, inspectores y asesores técnico pedagógicos (ATP). Para el caso de los miembros de equipos directivos participaron: directores, subdirectores, prefectos, coordinadores, docentes y orientadores.

Tabla 2  
Informantes clave del proyecto *Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California*

Supervisión Escolar		Equipo Directivo	
Puesto	Participantes	Puesto	Participantes
Inspector	13	Director	27
Supervisor	11	Subdirector	21
Asesor Técnico Pedagógico	24	Asistente de dirección	1
Jefe de dpto.	1	Prefecto	2
		Coord. de área	5
		Docente	3
		Orientador	2
		Auxiliar de laboratorio	1
Subtotal	49	Subtotal	62
Total de participantes 111			

*Fuente.* Elaboración propia con base en Chaparro et al., (2015)

*Entrevistas a informantes clave.* Se llevaron a cabo capacitaciones para entrevistadores y observadores por parte del equipo de investigación de la Universidad del País Vasco. Posteriormente se visitaron las escuelas seleccionadas y se realizaron las entrevistas a los informantes clave, mismas que se llevaron a cabo en diadas (entrevistador y observador). Asimismo, se solicitó autorización para ser audio grabadas (Chaparro et al., 2015).

*Transcripción de entrevistas.* Se realizaron transcripciones completas de cada una de las entrevistas realizadas, a diferencia de la transcripción selectiva realizada en el la investigación del País Vasco. Posteriormente se analizó la información derivada de las transcripciones para obtener las prácticas escolares, este procedimiento se describe en el siguiente apartado.

**3.2.3. Procedimiento para la selección de las prácticas escolares.** De acuerdo con lo reportado, la información obtenida de las entrevistas se organizó en las 14 categorías establecidas previamente: (a) Percepción general de la escuela, (b) Entorno e historia, (c) Gestión y organización, (d) Equipo directivo, (e) Recursos, (f) Programas, planes y proyectos, (g) Formación y compromiso del personal de la escuela, (h) Evaluación, (i) Evaluación del aprendizaje de los alumnos, (j) Metodología didáctica y materiales de enseñanza, (k) Atención a la diversidad, (l) Convivencia y clima escolar, (m) Relación bidireccional entre la escuela, la familia, la comunidad y redes de apoyo y, (n) Otras (Chaparro et al., 2015). Posteriormente, se realizó un análisis selectivo de contenido de la información de cada categoría con el propósito de seleccionar las prácticas escolares y sintetizarlas en un enunciado. No obstante, solo se consideraron

11 categorías, ya que estas son las de mayor presencia en lo reportado por la literatura de Eficacia Escolar y Mejora Escolar. Como resultado se obtuvo un total de 955 prácticas clasificadas en las 11 categorías (Apéndice A) (ver tabla 3). Los resultados derivados del proyecto antes mencionado fueron los insumos para el establecimiento de los criterios que conforman la rúbrica analítica desarrollada en la presente investigación.

Tabla 3  
*Total de prácticas por categoría*

Categoría	Total
1. Gestión y organización de la escuela	89
2. Equipo directivo	187
3. Recursos	45
4. Programas, planes y proyectos	50
5. Formación y compromiso del personal de la escuela	80
6. Evaluaciones escolares	84
7. Evaluación del aprendizaje del alumno	116
8. Metodología didáctica y materiales de enseñanza	66
9. Atención a la diversidad	16
10. Clima escolar y convivencia	78
11. Familia y comunidad	144

*Fuente.* Elaboración propia con base en (Chaparro et al., 2015)

# *4. Etapas del estudio*

El procedimiento para el desarrollo de la rúbrica se llevó a cabo en dos etapas. Para cada una se establecieron objetivos, materiales y procedimiento. Asimismo, se obtuvieron resultados específicos para cada una de ellas.

#### **4.1. Etapa 1. Valoración y re-categorización de las prácticas escolares**

**4.2.1. Objetivos.** (a) Valorar, mediante la opinión de expertos, las 955 prácticas escolares para determinar su eficacia y (b) re-categorizar las prácticas escolares a partir de las valoraciones de los expertos.

**4.2.2. Participantes.** Se convocó a 10 expertos del Sistema Educativo Estatal de Baja California de la delegación de Ensenada, de los cuales acudieron 9. El criterio de participación fue experiencia mínima de 10 años como docente, miembro de equipo directivo o supervisor escolar en escuelas secundarias. La media de años de experiencia estimada es de 17.22. Las características de los participantes se muestran en la tabla 4.

Tabla 4  
*Características de los participantes*

No. ID	Edad	Género	Puesto actual	Años de experiencia en educación básica	Grado de estudios
1	53	Femenino	Directora	10	Maestría en educación
2	39	Femenino	Directora	10	Maestría en Pedagogía
3	48	Masculino	Director	18	Especialidad en lengua y literatura
4	52	Masculino	Director	28	Doctorado en Ciencias de la Educación
5	52	Femenino	Inspectora	28	Doctorado en Educación
6	46	Femenino	Docente	16	Maestría en pedagogía
7	45	Masculino	Asesor Técnico Pedagógico	17	Maestría
8	56	Femenino	Inspectora	12	Doctorado en Pedagogía crítica y educación popular
9	39	Femenino	Docente	16	Maestría

Fuente. Elaboración propia.

### 4.2.3 Materiales.

**Manual del participante.** Se elaboró para contextualizar la actividad a desarrollar por los participantes, así como para describir y ejemplificar las instrucciones. El manual contenía una descripción del proyecto, el objetivo de la participación y la descripción de la dinámica de trabajo que se realizaría en las sesiones. Asimismo, se incluyeron las definiciones de los criterios de valoración de las prácticas escolares (ver tabla 5).

Tabla 5  
*Criterios para la valoración de las prácticas escolares*

Muy deficiente -2	Deficiente -1	Neutral 0	Eficiente 1	Muy eficiente 2
Acción que puede perjudicar significativamente los procesos de enseñanza aprendizaje en el aula o en la escuela y/o el desarrollo de los miembros de la comunidad escolar	Acción que puede obstaculizar los procesos de enseñanza aprendizaje en el aula o en la escuela y/o el desarrollo de los miembros de la comunidad escolar	Acción sin efecto positivo o negativo	Acción que puede contribuir a los procesos de enseñanza aprendizaje en el aula o en la escuela y/o el desarrollo de los miembros de la comunidad escolar	Acción que puede contribuir significativamente a los procesos de enseñanza aprendizaje en el aula o en la escuela y/o el desarrollo de los miembros de la comunidad escolar

*Fuente.* Elaboración propia.

**Hojas de registro.** Se integraron por dos listados que contenían las 955 prácticas escolares y la escala de valoración para cada una que los participantes utilizaron para determinar el nivel de eficacia de cada una.

**4.2.4 Procedimiento.** Se realizó en dos sesiones de trabajo, de cinco horas cada una.

*Sesión 1.* En plenaria, a los jueces se les dio una breve introducción sobre el proyecto de investigación y se explicó el propósito de su participación. Posteriormente, se describió la organización de la sesión, se entregaron los manuales de trabajo con las hojas de registro; y se indicó a los participantes, marcar con una “X” en las hojas de

registro, el nivel de eficacia de cada práctica escolar de acuerdo con las definiciones de cada nivel y con base en su experiencia. Asimismo, se les solicitó anotar las observaciones. Posteriormente el moderador mencionó cada una de las prácticas escolares en plenaria y a su vez los jueces realizaban las valoraciones correspondientes.

Cabe señalar que una vez que se inició la valoración del primer bloque de prácticas, los participantes expresaron dudas y confusiones respecto a la manera de valorar. Por lo tanto, se detuvo la valoración y se explicaron nuevamente las indicaciones, una vez que los participantes expresaron comprender las instrucciones, se continuó con la valoración.

*Sesión 2.* En esta sesión se llevó a cabo el mismo procedimiento de valoración seguido en la sesión 1. Los jueces valoraron 452 prácticas contenidas en el segundo listado. Al finalizar la sesión, el moderador solicitó a los participantes expresar las observaciones y comentarios finales sobre la valoración que realizaron.

**4.2.5. Análisis de datos.** Las valoraciones de las prácticas escolares de cada uno de los jueces se capturaron en formato Excel y se estimó la moda de cada valoración así como el porcentaje de frecuencia de cada una (Apéndice B). Después se realizó un filtro de los porcentajes superiores al 70% de frecuencia de las valoraciones y se obtuvo un total de 205 prácticas superiores al 70%. En seguida, se realizó una revisión y depuración del listado de las 205 prácticas a partir de ésta se eliminaron aquellas prácticas repetidas, no generalizables y las que no pertenecían a escuelas públicas. De este análisis se obtuvo un total de 154 prácticas escolares (Apéndice C). Asimismo, se identificaron discrepancias entre la categorización original y las características de las prácticas

obtenidas. Por lo tanto, se consideró conveniente realizar un análisis de contenido para re-categorizar las prácticas.

*Re-categorización de las prácticas escolares.* Las nuevas categorías se establecieron a partir del análisis de contenido de las 154 prácticas escolares con relación a los principales factores de Eficacia Escolar y Mejora Escolar reportados en la revisión de la literatura. Para ello, se analizaron las prácticas agrupadas en las categorías originales. El criterio que se consideró para el análisis fue si el contenido de cada práctica pertenecía a la categoría o si era necesario ubicarla en otra. De esta manera emergieron grupos de prácticas que por su contenido fue necesario renombrar (figura 2).

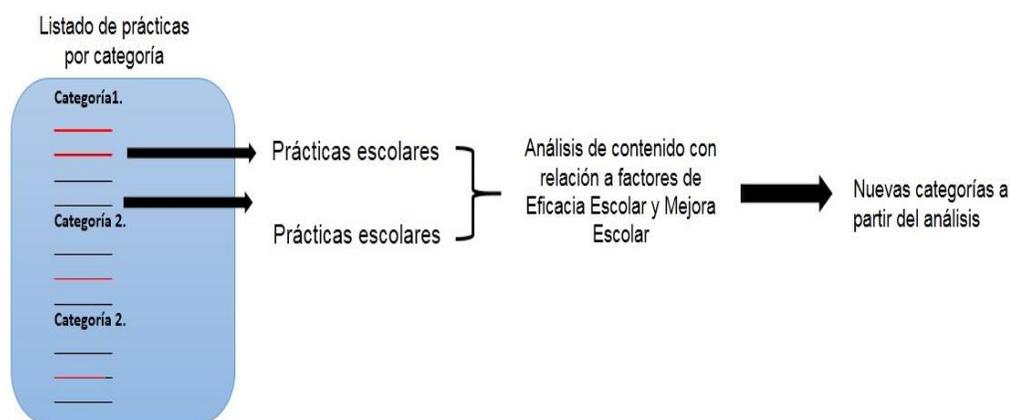


Figura 1. Procedimiento para el análisis de contenido de las prácticas escolares.

**4.2.6. Resultados.** Se obtuvieron 11 categorías en las cuales se clasificaron las 154 prácticas (ver tabla 6). Las categorías fueron: (a) Supervisión, (b) Liderazgo directivo, (c) Gestión de recursos, (d) Planeación, (e) Formación continua del personal escolar, (f) Evaluación del personal escolar, (g) Evaluación de los aprendizajes, (h)

**Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias**

Metodología didáctica y Estrategias de enseñanza (i) Atención a la diversidad, (j) Convivencia y clima escolar y (k) Relación de la escuela con los padres de familia. El total de prácticas en cada categoría fue:

**Tabla 6**  
**Total de prácticas por categoría**

Categoría	Total
1. Supervisión	13
2. Liderazgo directivo	10
3. Gestión de recursos	
a. Materiales	10
b. Humanos	8
4. Planeación	
a. Escolar	7
b. Clase	20
5. Formación continua del persona escolar	9
6. Evaluación del personal escolar	7
7. Evaluación de los aprendizajes	9
8. Metodología didáctica y estrategias de enseñanza	7
9. Atención a la diversidad,	12
10. Convivencia y clima escolar	21
11. Relación de la escuela con los padres de familia	19

Fuente: Elaboración propia.

Las categorías con mayor agrupación de prácticas fueron: Convivencia y clima escolar, Planeación de clase y Relación con padres de familia. Por otra parte, las categorías con menor cantidad de prácticas fueron: Planeación escolar, Metodología didáctica y estrategias de enseñanza y Gestión de recursos humanos. Así pues, las prácticas escolares derivadas de este análisis son los criterios que se integraron en la presente rúbrica.

## 4.2. Etapa 2. Integración y validez de contenido de la rúbrica

**4.3.1. Objetivos.** (a) Integrar la rúbrica para la evaluación de las prácticas escolares y (b) validar el contenido de la rúbrica mediante jueceo por expertos.

**4.3.2. Participantes.** Se consideró pertinente la colaboración de 12 jueces expertos, ya que según lo establecido por Lawshe (1975), para obtener la validez de contenido este número de jueces permite ampliar el criterio de valoración. El criterio de participación fue que éstos tuvieran diez años, o más, de experiencia en educación básica. En la tabla 7 se describen las características de los participantes.

Tabla 7  
Características de los participantes

No. ID	Edad	Género	Puesto actual	Años de experiencia en educación básica	Grado de estudios
1	49	Femenino	Docente y Coordinadora de del departamento de acreditación y certificación	20	Maestría en educación
2	40	Femenino	Coordinadora de formación continua y posgrado	17	Doctorado en educación
3	56	Masculino	Docente	26	Maestría en Pedagogía
4	35	Femenino	Docente	10	Maestría en Ciencias Educativas
5	56	Femenino	Docente	16	Maestría en Educación
6	38	Femenino	Docente	14	Maestría en Educación
7	43	Femenino	Docente y Coordinador académico	15	Maestría en Educación
8	40	Femenino	Directora	14	Maestría en Educación
9	45	Femenino	Directora	10	Maestría en Educación
10	46	Masculino	Docente	20	Licenciatura en Ciencias Educativas
11	33	Femenino	Docente	12	Licenciatura en Ciencias Educativas

Fuente. Elaboración propia.

### 4.3.3. Materiales

**Cuestionario de validación.** Se elaboró un cuestionario para la validación de la rúbrica, el cual se integró en archivo Excel (Apéndice D) y se conformó por dos secciones. En la primera se solicitó a los participantes que validaran la pertinencia de los niveles de valoración establecidos para la evaluación de las prácticas escolares (figura 2). La consigna para los jueces fue: determinar si la definición de los niveles era pertinente y que realizaran, si consideraban necesario, observaciones o sugerencias.

DEFINICIÓN	La definición ¿Es pertinente? ¿Considera necesario modificar la definición?	Observaciones o sugerencias
<b>Nivel 0 (nulo).</b> La práctica no se ha incorporado en la planificación u organización de las actividades escolares, y nunca se ha realizado.		
<b>Nivel 1 (insuficiente).</b> La práctica se ha incorporado en la planificación u organización de la escuela; sin embargo nunca se ha realizado.		
<b>Nivel 2 (aceptable).</b> La práctica se ha incorporado en la planificación u organización de la escuela; y se realiza ocasionalmente		
<b>Nivel 3 (deseable).</b> La práctica está establecida dentro de las planeaciones de las actividades escolares y se realiza con cierta frecuencia.		

Figura 2. Formato para valorar la pertinencia de los niveles.

En la segunda sección se solicitó que determinaran para cada una de los criterios si resultaban pertinentes para valorar la categoría en la que estaban clasificados. En la figura 3 se ilustra la consigna dada a los jueces y el formato utilizado.

<b>Actividad 2.</b> A continuación se presenta un listado las prácticas integradas en 11 categorías. Lea la definición de la categoría, a continuación revise cada práctica y escriba "SI" si considera que la práctica es pertinente para valorar la categoría, en caso contrario escriba "NO", y anote en la tercera columna las observaciones o sugerencias.		
<b>Categoría 1 Supervisión.</b> Acciones de la comunidad escolar dirigidas a supervisar: actividades escolares, de orden y disciplina, tanto a los estudiantes, como a los a docentes.		
PRÁCTICA	¿La práctica es pertinente para valorar la categoría?	Observaciones o sugerencias
1. Los prefectos vigilan y llevan a cabo el control de entrada de los alumnos.		
2. Se supervisa constantemente a los alumnos que están en riesgo de reprobado y a los que están reprobados.		

Figura 3. Ejemplo del formato para validar la pertinencia de las prácticas escolares.

**4.3.4. Procedimiento.** Los cuestionarios para la validación se enviaron a los 12 participantes vía correo electrónico. Se les solicitó que devolvieran el cuestionario contestado en un plazo máximo de una semana por la misma vía. Cabe señalar que el plazo se extendió a dos semanas en la mayoría de los casos. Una vez que se recibieron los cuestionarios, las respuestas para cada una de las actividades se capturaron en un archivo de Excel (Apéndice E) y se procedió a analizar los datos obtenidos.

**4.3.5. Análisis de los datos.** Con la finalidad de estimar la validez de contenido se realizó un análisis de las valoraciones realizadas por los jueces expertos. Para ello se utilizó el índice de validez de contenido (IVC) de Lawshe (citado en Arregui, Chaparro y Cordero, 2016), de esta manera se estimó la pertinencia de los niveles de valoración y los criterios de cada categoría que integran la rúbrica. La fórmula del IVC es:

$$IVC = \frac{n_e - N/2}{N/2}$$

## Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias

En dónde:  $n_e$  es la cantidad de jueces que valoraron el ítem como esencial y N el total de jueces participantes.

Posteriormente se descartaron los criterios que no cumplieron con los valores mínimos establecido por Lawshe cuando participan 12 jueces expertos (.56).

Número de Panelistas	Valor mínimo
5	.99
6	.99
7	.99
8	.75
9	.78
10	.62
11	.59
12	.56
13	.54
14	.51
15	.49
20	.42
25	.37
30	.33
35	.31
40	.29

Figura 4. Valores mínimos de Lawshe (Arregui, Chaparro y Cordero, 2016).

**4.3.6. Resultados.** De acuerdo con en el análisis de validez de contenido, en la tabla 8 se muestra el IVC de cada nivel de valoración así como el total de jueces a favor de la pertinencia en cada uno.

Tabla 8  
*IVC de niveles de valoración de la rúbrica*

Nivel	Jueces a favor	Porcentaje	IVC
Nivel 0	9		0.50
Nivel 1	10		0.66
Nivel 2	10		0.66
Nivel 3	9		0.50

Fuente. Elaboración propia.

Para los niveles 0 y 3 de la rúbrica se obtuvo un IVC de 0.50 con lo cual se determinó que éstos no cumplen con el IVC mínimo establecido por Lawshe cuando participan 12 jueces expertos. De manera contraria, para los niveles 1 y 2 se obtuvo un IVC de 0.66 con lo cual se aseguró la pertinencia de los mismos. Asimismo, se recuperaron las sugerencias y observaciones consideraras de mayor relevancia realizadas por los jueces para la modificación de los cuatro niveles:

- Descripción ambigua y confusa de los niveles de valoración.
- Incorporación de dos acciones en la definición de los criterios de valoración.
- Falta de especificidad en las descripciones.
- Mejorar la redacción.

Del mismo modo, se obtuvieron los IVC de los criterios incorporados en cada categoría de la rúbrica y con base en los resultados, se eliminaron aquellos con un IVC inferior a 0.56. A partir del análisis se eliminaron 36 criterios (ver tabla 9).

Tabla 9  
*Criterios eliminados a partir del análisis*

No.	Criterios	Jueces en desacuerdo	IVC
Categoría 1. Supervisión			
3	El director realiza rondas por las aulas para supervisar el orden y disciplina.	3	0.50
5	Cuando faltan miembros del equipo directivo se comisiona a otro personal para cubrir las actividades	4	0.33
6	El director implementa estrategias de supervisión a los docentes, considerando el contexto.	4	0.33
11	Diariamente, la dirección lleva un registro de los temas que impartirá cada docente.	3	0.50

**Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias**

Categoría 3. Gestión de recursos		
3.1 Materiales		
2	Los maestros llevan su propio cañón, equipo de cómputo y materiales a la escuela.	4 0.33
4	La aportación mensual del padre de familia se utiliza para la atender las necesidades de la escuela.	6 0
7	Los docentes se encargan de comprar material para sus clases.	5 0.16
3.2 Humanos		
3	A falta de coordinador académico el subdirector apoya con la supervisión de los trabajos extra, de los alumnos con bajo aprovechamiento.	3 0.5
6	El director pide a los vecinos apoyo de vigilancia los días en que la escuela no está trabajando.	4 0.33
8	La escuela gestiona apoyo de ex alumnos que están estudiando pedagogía o psicología con pláticas para impartir pláticas a los estudiantes.	5 0.16
Categoría 4. Planeación		
4.1 Escolar		
2	El director elige los programas y los proyectos en los que la escuela participará de acuerdo a las necesidades del alumnado.	3 0.50
4	El director en conjunto con el subdirector elabora los objetivos escolares con base en un diagnóstico.	3 0.50
5	El directivo toma en cuenta las actividades sugeridas por los padres de familia respecto a los objetivos escolares.	3 0.50
7	Los padres de familia complementan la planeación de las actividades escolares.	3 0.50
Categoría 6. Evaluación del personal escolar		
2	Se realizan exámenes para evaluar al personal docente.	4 0.33
5	Los docentes se autoevalúan.	3 0.50
Categoría 8. Metodología didáctica y estrategias de enseñanza		
5	Diariamente los alumnos reciben una firma de los docentes para dar una tarea por concluida.	4 0.33
10	Los docentes promueve actividades altruistas con los alumnos (p.ej. Visitar orfanatorios)	4 0.33
Categoría 9. Atención a la diversidad.		
2	Los docentes imparten asesorías los sábados o en contra turno a los alumnos con bajo aprovechamiento.	4 0.33
3	Los docentes acuden a los domicilios de los alumnos que no han asistido a la escuela.	7 0.16
Categoría 10. Convivencia y clima escolar.		
7	Los docentes demuestran aprecio a sus alumnos.	3 0.50
Categoría 11. Relación con los padres de familia.		
4	Los padres de familia firman un documento comprometiéndose a asistir cuando sea requerido.	4 0.33
5	Los padres de familia apoyan a la escuela a la hora de entrada y salida, para la supervisión de los alumnos.	3 0.50
6	El director se adecua a los horarios de los padres de familia para garantizar su asistencia.	3 0.50
12	El director usa los medios electrónicos para mantener informados a los padres de familia.	4 0.33
14	El director entrega oficios a los padres de familia para que, en su trabajo, les den permiso de asistir a las juntas que se les convocan.	3 0.50

17	La escuela solicita apoyo al consejo de participación social cuando se da un problema serio con los padres de familia (p.ej. Padres que se comportan violentamente, que maltratan a sus hijos, etc.).	4	0.33
----	---	---	------

*Fuente.* Elaboración propia

De manera general, se conservaron la mayoría de los criterios de cada categoría. Lo cual indica que los criterios seleccionados representan a la categoría asignada y son pertinentes para valorar cada categoría. De manera específica, en las categorías de Gestión de recursos humanos, Planeación escolar y Evaluación del Personal Escolar se eliminaron el mayor número de criterios con relación al número original contenidos en cada categoría de la rúbrica. En Gestión de Recursos Humanos se descartaron cuatro de los ocho criterios establecidos, para la categoría de Planeación escolar se eliminaron cuatro de siete criterios y en Evaluación del personal escolar se conservaron cuatro de siete criterios. Con relación a las sugerencias y comentarios de los jueces, señalaron que algunos de los criterios no son congruentes con la realidad de las escuelas secundarias públicas; asimismo, recomendaron mejorar la redacción para mayor comprensión. Finalmente, de los 155 criterios sometidos al proceso de validez de contenido, se conservaron 119 criterios con IVC aceptable (ver tabla 10). De tal manera que, los criterios conservados son los mismos que integran la rúbrica de la presente investigación (ver Apéndice F).

Tabla 10  
*Criterios con IVC aceptable por categoría.*

Categoría	Total de criterios	Total de criterios con IVC aceptable
1. Supervisión	13	9
2. Liderazgo directivo	10	10
3. Gestión de recursos		
a. Materiales	10	7

**Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias**

b. Humanos	8	4
4. Planeación		
a. Escolar	7	3
b. Clase	10	10
5. Formación continua del personal	9	9
6. Evaluación del personal escolar	7	4
7. Evaluación de los aprendizajes	9	9
8. Metodología didáctica y estrategias de enseñanza	13	11
9. Atención a la diversidad,	12	10
10. Convivencia y clima escolar	21	20
11. Relación de la escuela con los padres de familia	19	13
	Total	155
		119

*Fuente.* Elaboración propia.

## *5. Discusión*

De acuerdo con los resultados obtenidos de la presente investigación, se determina que el objetivo general *desarrollar una rúbrica analítica para la autoevaluación de las prácticas escolares en escuelas secundarias*, se cumplió satisfactoriamente. De tal manera que, en el presente capítulo se describe la interpretación de los resultados derivados de cada una de las etapas con relación a los objetivos específicos y con la revisión de literatura realizada en el capítulo II.

Según lo reportado por algunos autores en estudios sobre procedimientos para la elaboración de rúbricas (Arias y Maturana, citado en Picón 2013; Blommel y Abate, 2007, Gatica-Lara y Uribarren-Berrueta, 2013; López y Velasco, 2013) la definición del constructo a evaluar comprende el primer paso para el desarrollo de este tipo de instrumentos ya que de ello depende, en gran medida, la precisión de su evaluación. Para ello, se sugiere el uso de referentes teóricos y conceptuales, así como la participación de expertos. En el caso de la presente investigación, la definición del constructo se realizó con base en las definiciones conceptuales de prácticas escolares y buenas prácticas escolares, asimismo, se realizó análisis de contenido de entrevistas y jueceo por expertos.

Para la definición del constructo, en la investigación *Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California* (Chaparro et al., 20015) se realizó una selección de prácticas escolares mediante análisis de contenido de transcripciones de entrevistas realizadas a miembros de la comunidad escolar de escuelas secundarias. Para ello, de manera colegiada, se definió el concepto de práctica

escolar como toda acción llevada a cabo por los miembros de la comunidad escolar (docentes, equipo directivo, personal administrativo).

Este primer paso en el procedimiento, concuerda con los procesos seguidos en las investigaciones de Contero, Otey y Plumed (2015), Hernández-Rivera (2013) y, López y Velazco (2013) en las cuales se recomienda la colaboración de expertos para definir, de manera consensuada, el constructo que se desea evaluar. Con base en esta definición, se realizó la selección de las prácticas escolares de las transcripciones de entrevistas realizadas a los miembros de la comunidad escolar y posteriormente se clasificaron en categorías previamente establecidas.

Cabe señalar que, las definiciones de dichas categorías se establecieron en el proyecto antes mencionado, con base en investigaciones previas sobre Eficacia Escolar y Mejora Escolar. A partir del procedimiento se seleccionaron 955 prácticas escolares clasificadas en 11 categorías: (a) Gestión y organización de la escuela, (b) Equipo directivo, (c) Recursos, (d) Programas, planes y proyectos, (e) Formación y compromiso del personal de la escuela, (f) Evaluaciones escolares, (g) evaluación del aprendizaje del alumno (h) Metodología didáctica y materiales de enseñanza, (i) Atención a la diversidad (j) Clima escolar y convivencia; y (k) Familia y comunidad (Chaparro et al., 20015).

El proceso que se siguió concuerda con el procedimiento realizado en la investigación de Hernández-Rivera (2013). En su estudio, el autor estableció las competencias a evaluar de su rúbrica mediante el análisis de entrevistas semiestructuradas a personal docente. Después, mediante consenso, se clasificaron

según la dimensión a la cual los expertos consideraban para formular los criterios de la rúbrica. Para el caso de esta investigación, la selección de las prácticas escolares de las transcripciones, así como la clasificación, se consideró la primera aproximación para establecer, posteriormente, los criterios de la rúbrica.

Como segundo objetivo específico se estableció *determinar la eficacia de las prácticas escolares mediante la opinión de expertos, para establecer los criterios de la rúbrica*. Éste propósito se cumplió satisfactoriamente y se considera que fue la segunda aproximación para la elaboración de los criterios que forman parte la rúbrica.

Colaboraron nueve expertos en educación básica y se elaboraron cuatro niveles de valoración para que determinaran la eficacia de las prácticas escolares seleccionadas previamente. El diseño de los niveles se realizó con base en el concepto de escuelas eficaces propuesto por Murillo (2005), y en las sugerencias realizadas por Blommel y Abate (2007), López y Velasco (2013), Urbieta et al. (2011), quienes señalaron que para elaborar los niveles de desempeño de una rúbrica se debe asegurar que éstos sean descripciones precisas y claramente diferentes entre cada nivel. Al respecto, los expertos participantes en esta investigación concluyeron que los niveles de valoración fueron pertinentes para determinar la eficacia de las prácticas escolares. Posterior a la valoración de los jueces expertos y al análisis de los datos, se conservó el 21.46% (205) del 100% (955) de las prácticas escolares. Tomando como criterio el 70% de acuerdo entre los jueces con relación a la eficacia de cada practica escolar. De tal manera que, se descartaron más de la mitad (78.53%). Asimismo, se recuperaron los comentarios de los jueces con relación a la pertinencia de las prácticas escolares valoradas, los comentarios son:

## **Desarrollo y evidencias de validez de una rúbrica para la autoevaluación de prácticas escolares en escuelas secundarias**

- Algunas de las prácticas escolares seleccionadas por los analistas no se consideraron parte de la realidad educativa de las escuelas secundarias públicas.
- Se identificaron prácticas escolares que no corresponden a acciones ejecutadas por miembros de la comunidad escolar.
- Las prácticas escolares de la categoría (d) programas, planes y proyectos, en su mayoría, ya no se realizan debido al cambio en la Reforma Educativa.
- Existen enunciados que fueron identificados como prácticas escolares, sin embargo, es contexto escolar.

Se considera que éstos resultados son consecuencia de las características de la muestra original, ya que algunas de las transcripciones de entrevistas pertenecían a escuelas secundarias privadas y no se realizó una diferenciación entre éstas y las públicas. Ya que si bien una buena práctica escolar puede ser replicada, el contexto determina en qué medida. Asimismo, se considera que la selección realizada por los analistas presenta ciertas deficiencias al observarse la inclusión de fragmentos que corresponden a contexto escolar y prácticas que no corresponden a la definición establecida en la investigación.

En una nueva revisión y con la finalidad de conservar solo aquellas prácticas generalizables al contexto de escuelas secundarias públicas se realizó una depuración de las 205 prácticas escolares con base en dos criterios: aquellas repetidas y las no generalizables al contexto de escuelas secundarias públicas. De esta manera, se eliminó el 24.87 % y se obtuvo un total de 154 prácticas escolares.

### 5.1. Criterios de la rúbrica analítica

Para establecer los criterios de la rúbrica analítica se seleccionaron las 154 prácticas escolares derivadas del análisis de datos de la etapa uno. Además, se consideró pertinente realizar una nueva categorización de las prácticas seleccionadas con base en la revisión de la literatura de las investigaciones sobre Eficacia Escolar y Mejora Escolar en las que se han identificado elementos característicos de las escuelas eficaces. De esta manera, se establecieron 11 categorías integradas por prácticas escolares. Estas prácticas se establecieron como los criterios que conformaron la primera versión de la rúbrica analítica.

**Categoría 1. Supervisión.** Se define como las acciones de la comunidad escolar dirigidas a supervisar actividades escolares, de orden y disciplina, tanto a los estudiantes, como a los docentes. Dicha definición se basa en las características de gestión escolar identificadas en las investigaciones de Hernández (2009), Lizasoain y Angulo (2014). Al respecto, los autores señalaron que las escuelas eficaces presentan organización sistematizada y planificada. Asimismo, esta categoría se relaciona con algunos elementos presentes en la dimensión de clima escolar referidos en las investigaciones de Hernández-Castilla et al. (2016), Murillo (2005,2008) y Sebring et al. (2009) puesto que se integran aspectos de orden y disciplina, los cuales, según estos estudios, favorecen la dinámica escolar. Así, la categoría se integró por nueve criterios clasificados de acuerdo con la definición elaborada, algunos ejemplos son: *Los prefectos vigilan y llevan a cabo el control de entrada de los alumnos, El equipo directivo lleva un registro de las observaciones que hace de los docentes en el aula.*

**Categoría 2. Liderazgo directivo.** Para su definición se tomó como referencia la definición de liderazgo educativo propuesta por Horn y Marfán (2010). Los autores señalaron que éste se caracteriza por acciones llevadas a cabo, la mayoría de las veces, por directivos escolares, quienes fomentan objetivos de mejora compartidos entre la comunidad escolar, además, de que logran alinear los recursos pedagógicos, financieros y humanos en pro de dichos objetivos. Lo anterior hace referencia a características de liderazgo con orientación pedagógica, el cual, se relaciona con el manejo de los procesos al interior de la escuela de manera efectiva y eficaz (Scanlan, 2011), además, de ser un factor determinante para la eficacia escolar (Franco, 2012; Hernández-Castilla et al, 2016; Izaguirre, 2000; Lizasoain y Angulo, 2014). Con base en lo anterior, se consideró pertinente incluir esta dimensión como categoría de la rúbrica analítica, la cual se definió como: acciones del personal directivo dirigidas hacia la organización, seguimiento y manejo de actividades relacionadas a la escuela, docentes y alumnos. Esta categoría se integra por 10 criterios: *El director trabaja de manera colaborativa con todo el equipo directivo, El director da preferencia al cumplimiento de los objetivos orientados a la educación de los alumnos.*

**Categoría 3. Gestión de recursos.** Se consideró como punto de referencia lo señalado por Racynski y Muñoz (2005) quienes afirmaron que los recursos son factores relacionados a la eficacia escolar, los cuales pueden ser: humanos, materiales y de infraestructura. Con base en lo expuesto y con relación al contenido de los criterios, se optó por realizar una subcategorización en recursos materiales y recursos humanos con diez y ocho criterios respectivamente. Asimismo, se decidió enfocar la categoría a la gestión de los recursos, ya que si bien, la presencia de recursos en las escuelas influye

en el rendimiento académico de los estudiantes, éste no es un factor determinante (L. Lizasoain, mayo 2016); no obstante, la forma en la que dichos recursos se utilizan es más significativa. En este sentido, las prácticas escolares relacionadas gestión de los recursos (humanos y materiales) puede ser un factor determinante para la eficacia escolar más que la simple presencia de los recursos. Esta categoría se define como: Acciones de la comunidad escolar dirigidas a solicitar o generar recursos económicos, materiales o humanos para la escuela y se integra por dos subcategorías: recursos humanos y materiales, con 7 y 4 criterios respectivamente (ej. *La comunidad escolar ha adecuado y construido con recursos propios instalaciones de la escuela, La escuela gestiona apoyo de ex alumnos que están estudiando pedagogía o psicología con pláticas para a los estudiantes*).

**Categoría 4. Planeación.** Esta categoría se subcategorizó en planeación escolar y planeación de clase debido a que los criterios clasificados en esta categoría hacen referencia a éstas dos dimensiones. Lo anterior, coincide con los hallazgos de Román (2008) , quien afirmó que la planeación dentro del aula es un factor propio del proceso pedagógico y que además, influye en la eficacia escolar. Asimismo, puntualizó que las comunidades escolares que realizan cambios dirigidos a la mejora escolar, se caracterizan por organizarse, planificar y priorizar objetivos. Con base en lo anterior, la presente categoría se define como: acciones del personal escolar relacionadas con la planeación y organización de actividades (escolares y de aula). Al respecto, para la subcategoría de planeación escolar se incluyeron 3 criterios (ej. *La planeación anual escolar se elabora al principio del año con base en las necesidades de la escuela, con el seguimiento del director, subdirectora y docentes*) y para la de planeación de clase 10

(ej. *El equipo directivo supervisa que los docentes planifiquen con base en los aprendizajes de los alumnos*).

**Categoría 5. Formación continua del personal escolar.** Se integraron nueve criterios (ej. *El equipo directivo incluye actividades de acompañamiento y pláticas dirigidas a los docentes, Los docentes asisten a los cursos de actualización a los que son convocados*) que se relacionan con el factor de desarrollo profesional referido por Hernández-Castilla (2016) quien señaló que éste se caracteriza por una actitud hacia el aprendizaje continuo y que además, corresponde a factores de Eficacia Escolar. Asimismo, se incluyeron criterios vinculados a la capacitación profesional (Serbing et al., 2009) y a la promoción de la adquisición de conocimientos y competencias (Ministerio de Educación, Cultura y Deporte, 2014) ya que éstos elementos se consideran factores de mejora escolar. Con base en lo anterior, la categoría se define como: actividades encaminadas a la actualización profesional (cursos, conferencias, capacitaciones, etc.) del personal escolar.

**Categoría 6. Evaluación del personal escolar.** Se tomaron como base los factores de mejora escolar mencionados por Murillo (2003), quien afirmó que los cambios para la mejora de una escuela se centran en la información derivada de los datos tanto de los alumnos, como de la comunidad escolar en general. Dicha información se obtiene mediante evaluaciones y seguimiento de las mismas, además, de acuerdo con Hernández- Castilla (2016) existe evidencia empírica de que estos dos elementos son factores que determinan la eficacia de las escuelas puesto que, gracias a los resultados de las evaluaciones, orienta la toma de decisiones (Román, 2011). De acuerdo con las características de contenido de las prácticas escolares (criterios) y con base en las

aportaciones antes mencionadas, se consideró pertinente definir la categoría como las acciones dirigidas hacia la evaluación de las actividades y desempeño del personal escolar. Esta categoría se integra por cuatro criterios (ej. *Los objetivos escolares se evalúan a partir de los resultados académicos de cada bimestre*).

**Categoría 7. Evaluación de los aprendizajes.** Se definió como las acciones y estrategias dirigidas a evaluar el aprendizaje de los alumnos. En este sentido, se ha demostrado que las escuelas eficaces utilizan los resultados académicos de los estudiantes para desarrollar planes de mejora (Lizasoain y Angulo, 2014), además de que es posible transformar los procesos educativos (Schleicher y Zoido, 2013). Por consiguiente, al incluir esta categoría en la rúbrica se espera que los resultados derivados de la misma promuevan adecuaciones pertinentes a las estrategias de evaluación de los aprendizajes. Los criterios que conforman esta categoría son nueve (ej. *El docente aplica una evaluación diagnóstica de los alumnos para conocer su nivel de aprovechamiento, así como sus fortalezas y debilidades*).

**Categoría 8. Metodología didáctica y estrategias de enseñanza.** De acuerdo Murillo (2003) los cambios dirigidos a la manera de enseñar los contenidos curriculares son propios de los factores asociados a la mejora escolar. Asimismo, Hernández-Castilla (2016) afirmó que la calidad de las estrategias de enseñanza es un factor que se relaciona con la eficacia de las escuelas. Con base en lo anterior, se define esta categoría como las acciones relacionadas con el diseño, adecuación y organización de métodos didácticos y estrategias para el aprendizaje de los alumnos. Al interior de esta dimensión se clasificaron once criterios (ej. *Los docentes tratan de innovar constantemente sus estrategias de enseñanza*).

**Categoría 9. Atención a la diversidad.** Se incluyeron 10 criterios (ej. *Los docentes investigan estrategias para trabajar con alumnos con NEE*) con base lo expuesto por Gairín (2012) quien señaló que una escuela que atiende la diversidad se caracteriza por compensar las desigualdades de origen de los estudiantes y provee oportunidades educativas y desarrolla estrategias de mejora permanente considerando la realidad social. Asimismo, de acuerdo con lo reportado en el XXI Encuentro de Consejos Escolares Autonómicos del Estado de Oviedo (Ministerio de Educación, Cultura y Deporte, 2014), detectar de manera oportuna las dificultades del alumno se relaciona con factores de mejora escolar, puesto que, según García (2002) el sistema educativo debe garantizar que los estudiantes logren los objetivos académicos independientemente de sus características sociales y personales. De acuerdo con lo anterior, esta categoría se conceptualiza como las acciones y estrategias orientadas a la atención de estudiantes con algún tipo de necesidad (necesidades educativas especiales, problemas de aprendizaje, problemas de conducta, etc.).

**Categoría 10. Convivencia y clima escolar.** Para la selección de los criterios 20 (ej. *Al inicio del ciclo escolar, la escuela organiza actividades de integración para los estudiantes y los nuevos docentes*) que la integran, se utilizaron hallazgos referidos por Murillo (2005, 2008) quien afirmó que las buenas relaciones entre los miembros de la comunidad escolar tanto un ambiente libre de conflictos son elementos presentes en las escuelas eficaces. Además, de acuerdo con Mesa-Melo et al (2013) con ello se favorece el desarrollo académico. Asimismo, con base en las aportaciones de Lizasoain y Angulo (2014), en las escuelas con alumnos que obtuvieron resultados académicos mayor a lo esperado, se encuentran buenas prácticas escolares relacionadas a la resolución de

conflictos y existe una percepción de “buen” clima escolar. De tal manera, la categoría 10 se define como las acciones y estrategias dirigidas hacia la promoción de la convivencia escolar para generar un clima adecuado para el desarrollo de todos los miembros de la comunidad escolar.

***Categoría 11. Relación de la escuela con los padres de familia.*** Se conceptualiza como las acciones del personal escolar para fomentar y mantener comunicación adecuada con los padres de familia de los estudiantes. Para esta categoría se seleccionaron 13 criterios que se relacionan con lo referido en investigaciones previas sobre esta dimensión (ej. *El director invita a los padres de familia que tiene alguna duda, visiten el salón de clases, hagan cita con el docente o con el mismo*). Así, se consideraron criterios vinculados a las acciones llevadas a cabo por la comunidad escolar para implicar a las familias en actividades escolares, ya que como reportó Hernández-Castilla, Murillo y Martínez-Garrido (2016), la participación familiar tanto las actitudes positivas de colaboración (Murillo, 2008) son factores presentes en las escuelas eficaces.

De acuerdo con el procedimiento para el establecimiento de los criterios, así como para su clasificación, se identifican semejanzas a los procesos llevados a cabo en algunas investigaciones. En el caso de los estudios de Company, Contero, Otey y Plumed (2015), Hernández-Rivera (2012), Lima-Rodríguez (2015) y Urbiera et al. (2011), se identificaron competencias y criterios a evaluar, posteriormente se agruparon de acuerdo al área del conocimiento (categorías). De manera similar, en la presente investigación se seleccionaron los criterios para valorar y después se clasificaron en categorías relacionadas con los factores de eficacia escolar y mejora escolar. No obstante, en primera instancia, se partió de categorías previamente establecidas para obtener las

primeras aproximaciones de los criterios. En el caso de los procedimientos llevados a cabo por Blommel y Abate (2007); y López y Velasco (2013), los autores establecieron, antes que los criterios, las dimensiones que deseaban evaluar con la rúbrica, sin embargo, no se reportó una segunda categorización.

## **5.2. Diseño de los niveles de valoración de la rúbrica analítica.**

Para diseñar los niveles de valoración de la rúbrica analítica se llevó a cabo un consenso con el equipo de investigación siguiendo las sugerencias expuestas por Blommel y Abate (2007), López y Velasco (2013), Urbietta et al. (2011) para la elaboración de las descripciones de los niveles de desempeño de una rúbrica. Asimismo, se tomó como referencia lo señalado por Shimpan, Roa, Hooten y Wang (2012) quienes afirmaron que los niveles deben explicar detalladamente lo que se debe hacer para demostrar, los niveles de eficiencia, y con ello proporcionar retroalimentación. Además, recomendaron que la estrategia de puntuación debe considerar cuatro niveles. Para el propósito de esta investigación, se diseñaron cuatro niveles para la presencia de las prácticas escolares.

- Nivel 0 (nulo)
- Nivel 1 (insuficiente)
- Nivel 2 (aceptable)
- Nivel 3 (deseable)

De esta manera, se lograron establecer niveles caracterizados por una puntuación analítica, ya que de acuerdo con Beyreli y Ari (2009) se considera útil para evaluar criterios de manera separada. Lo cual es característico de las rúbricas analíticas.

### 5.3. Evidencias de validez de contenido

Para el objetivo específico tres, se propuso *obtener la validez de contenido de la rúbrica mediante el índice de validez de Lawshe (1975)*. De acuerdo con la interpretación de los resultados obtenidos el objetivo se cumplió satisfactoriamente. En este estudio la validez se conceptualiza de acuerdo como “el grado en que la evidencia y la teoría soportan la interpretación de las puntuaciones del test para el uso pretendido” (AERA, APA y NCME, 2014, p.11). Asimismo, se atendieron las sugerencias señaladas por Andrade (2005) para obtener la validez de las rúbricas. En este sentido, el autor refirió que para asegurar la validez de estos instrumentos, los constructos deben ser descritos claramente, así como los niveles de desempeño, que en este caso, son los niveles de valoración. De acuerdo con Martínez-Arias (1995) la validez de contenido es el grado en que los ítems (en este caso criterios) de un instrumento, representan adecuadamente lo que se desea evaluar. Por lo tanto, este tipo de validez se basa en la definición precisa del dominio y en el juicio de los expertos con relación al grado con que éste dominio es evaluado a través de los criterios.

Con base en la revisión de la literatura, las evidencias de validez de contenido son las que mayor se reportan en los estudios de elaboración de rúbricas, así como el uso de técnicas de jueceo por expertos (Blommel y Abate, 2007; Contreras-Valenzuela et al, 2013; Musial et al, 2007; García-Ros et al, 2012; Gordillo y Vázquez, 2012; López y Velasco, 2013; Lima-Rodríguez et al, 2015; Verano-Tacoronte, 2016). Con relación a lo anterior, de acuerdo con el estándar 1.7 establecido en los SEPT (2014) cuando la validez se consigue mediante la opinión y juicios de expertos, debe comprobarse dicha experiencia, así como describirse en el procedimiento.

Atendiendo a los puntos señalados, para el proceso de validez de contenido de la rúbrica analítica, tanto los criterios previamente clasificados en las 11 categorías, así como los niveles de valoración, se integraron en una matriz para ser validados por un grupo de expertos. Para ello, se optó utilizar el índice de validez de contenido de Lawshe (1975) puesto que se ha comprobado su efectividad para obtener este tipo de validez cuando participan un grupo de expertos (Arregui, Chaparro y Cordero, 2016). De acuerdo con el análisis de los datos, posterior al jueceo de los expertos, los niveles de valoración 0 y 3 no se consideran pertinentes, puesto que obtuvieron índices de validez inaceptables (0.50 respectivamente) con relación al IVC aceptable cuando participan 12 jueces (0.56). Aunado a que los jueces, enfatizaron una descripción ambigua de los mismos. De manera contraria, para los niveles de valoración 1 y 2 se obtuvieron niveles aceptables (0.66), no obstante, los jueces sugirieron mejorar la redacción de los mismos. De esta manera logra una segunda aproximación a los niveles de valoración, los cuales necesitan ser mejorados.

Con relación a la validez de contenido de los criterios, los jueces valoraron la pertinencia de éstos con relación a la categoría asignada. De acuerdo con el análisis de los datos del 100% de los criterios (154) se eliminó el 23 % (36). Con lo anterior, se determinó que la rúbrica analítica presenta categorías y criterios adecuados que permiten la valoración de las prácticas escolares en las dimensiones relacionadas con la Eficacia Escolar y la Mejora Escolar.

## *6. Conclusiones*

Se obtuvo una rúbrica analítica integrada por 119 criterios (Apéndice F) formulados a partir de prácticas escolares presentes en la dinámica escolar de escuelas secundarias públicas valoradas eficaces y once categorías que corresponden a elementos relacionados a los modelos de Eficacia Escolar y Mejora Escolar referidos en investigaciones previas. Además, se obtuvieron índices aceptables de validez de contenido tanto de los criterios como de las categorías de la rúbrica. Por consiguiente, se establece que la presente rúbrica analítica, permite autoevaluar la presencia de prácticas escolares eficaces en escuelas secundarias en once categorías: (a) Supervisión, (b) Liderazgo directivo (c) Gestión de recursos, (d) Planeación (e) Formación continua del personal, (f) Evaluación del personal escolar (g) Evaluación de los aprendizajes (h) Metodología didáctica y Estrategias de Enseñanza (i) Atención a la diversidad (j) Convivencia y clima escolar, y (k) Relación de la escuela con los padres de familia.

Se considera que el principal aporte de esta investigación es de tipo metodológico. Ya que si bien existen estudios en los que se reportan procedimientos para la elaboración de rúbricas, éstos son escasos con relación a la frecuencia con que es utilizada en el campo de la educación. Por consiguiente, la poca literatura al respecto ha dificultado seguir una metodología para el desarrollo de las rúbricas (Reddy y Andrade, 2010). Con base en lo reportado en otras investigaciones para la elaboración de las rúbricas es importante la colaboración de expertos para definir el constructo a evaluar, así como para diseñar los criterios y los niveles de desempeño. Los cuales deberán ser formulados de manera clara y específica para asegurar una evaluación

precisa del constructo, además, se parte de un análisis teórico y conceptual para diseñarlos.

Para la elaboración del presente instrumento se consideraron las recomendaciones antes mencionadas tanto para la definición del constructo, los criterios y los niveles de desempeño. No obstante, se consideró que para obtener una rúbrica que permita valorar las prácticas escolares en escuelas secundarias, es necesario complementar lo que se reporta en la literatura con evidencias empíricas del contexto al cual está dirigido el instrumento. Ya que si bien se generalizan elementos y prácticas escolares asociados a la Eficacia Escolar y Mejora Escolar éstos dependen de la realidad educativa y del contexto de las escuelas.

Por lo anterior, la elaboración de los criterios se realizó a partir de análisis de contenido de entrevistas sobre prácticas escolares realizadas a miembros de la comunidad escolar (equipo directivo, supervisores, inspectores y asesores técnicos pedagógicos) de una muestra de escuelas secundarias de Baja California. Posteriormente, los resultados derivados del análisis se categorizaron en función de los elementos relacionados a los modelos de Eficacia Escolar y Mejora Escolar.

De manera general se determinó que los modelos de Eficacia Escolar y Mejora Escolar son marcos de referencia adecuados para el desarrollo de una rúbrica para la autoevaluación de prácticas escolares puesto que incluyen elementos presentes en la dinámica escolar de escuelas consideradas eficaces y de estrategias llevadas a cabo que promueven la mejora continua. Asimismo, el complemento de éste marco de

referencia con evidencias empíricas de prácticas escolares eficaces fortalece la validez de contenido de la rúbrica.

## *7. Limitaciones y sugerencias*

De acuerdo con el procedimiento llevado a cabo en la presente investigación, se identifican algunas limitaciones y sugerencias que deben considerarse para futuros estudios en esta área.

- Alguna de la información contenida en las transcripciones de entrevistas utilizadas para el análisis de contenido no se relacionaba a la realidad actual del sistema escolar y en consecuencia se descartaron diversas prácticas escolares, las cuales pueden ser de utilidad para otro tipo de estudios. Por ello, se sugiere utilizar información reciente para que esta sea congruente con el contexto actual.
- Se considera que una de las limitaciones de mayor importancia fue la poca disposición de tiempo y los recursos económicos para realizar pruebas piloto del instrumento y así obtener retroalimentación para realizar los ajustes pertinentes y obtener otras evidencias de validez (constructo).

### **7.1. Recomendaciones**

- Mejorar la capacitación a los analistas para lograr una diferenciación adecuada entre práctica escolar y contexto escolar.
- Realizar aplicaciones piloto para recibir retroalimentación de los usuarios con relación a los criterios de la rúbrica, niveles de eficacia y dimensiones y así hacer adecuaciones pertinentes a la rúbrica.
- Elaborar guía de aplicación que acompañe a la rúbrica. Ya que de acuerdo con estándares, un instrumento de evaluación debe expresar claramente su propósito e instrucciones de aplicación.

## Referencias

- Abdoulaye, A. (2003). *Conceptualisation et dissemination des “bonnes pratiques” en éducation: essai d’une approche internationale à partir d’enseignements tirés d’un projet*. Recuperado del sitio de internet de UNESCO:[http://www.ibe.unesco.org/fileadmin/user\\_upload/HIV\\_and\\_AIDS/publications/abdoulaye.pdf](http://www.ibe.unesco.org/fileadmin/user_upload/HIV_and_AIDS/publications/abdoulaye.pdf)
- Aguerrondo, I., Blanco, R., Calvo, G., Carillo, S. (2008). *Eficacia escolar y factores asociados en América Latina y el Caribe*. Recuperado del sitio de internet de UNESCO: <http://unesdoc.unesco.org/images/0015/001528/152895e.pdf>
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education & Joint Committee on Standards for Educational and Psychological Testing (U.S). (2014). Standards for educational and psychological testing.
- Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *Collage Teaching*, 53 (1), 27–31. Recuperado de <http://www.jstor.org/stable/27559213>
- Andrés, S. & Giró, J. (2016). El papel y la representación del profesorado en la participación de las familias en la escuela. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 19(1), 61-71. doi: <http://dx.doi.org/10.6018/reifop.19.1.245461>
- Arregui, E. I.G, Chaparro Caso, A.A., y Cordero, A.G. (2016). El IVC de Lawshe en el análisis de la validez de contenido de un instrumento para la evaluación del desempeño docente. Guadalajara, México: Editorial Universidad de Guadalajara
- Backhoff, E., Bouzas, A., González-Montesinos, M. J., Contreras, C., Hernández, E. y García, M. (2008). *Factores asociados al aprendizaje de estudiantes de 3º de primaria en México*. Recuperado del sitio de Internet del Instituto Nacional para

- la Evaluación de la Educación, México:  
[http://www.senado.gob.mx/comisiones/educacion/Escuela\\_Libre\\_Violencia/docs/Factoresescolarestercero2008.pdf](http://www.senado.gob.mx/comisiones/educacion/Escuela_Libre_Violencia/docs/Factoresescolarestercero2008.pdf)
- Ballesteros, B., & Mata, P. (2009). Approaching the Concepts of “Educational Achievement” and “Best Practice”: Delphi methodology as a tool for building consensus. *Research in Comparative and International Education*, 4(4), 356–365.
- Beach, P. V. (2008). Influencia de la dirección escolar en los logros académicos. In Innovación y cambio en las organizaciones educativas. *Universidad de Deusto; Deustuko Unibertsitatea*, 1009-1029
- Bellón, E. M. E., Cantero, J. M. M., y Ziemer, M. F. (2012). La autoevaluación y el diseño de planes de mejora en centros educativos como proceso de investigación e innovación en Educación Infantil y Primaria. *Revista Electrónica Interuniversitaria de Formación Del Profesorado*, 15 (1), 145–155. Recuperado de <http://www.redalyc.org/articulo.oa?id=217024398012>
- Beyreli, L., & Ari, G. (2009). The Use of Analytic Rubric in the Assessment of Writing Performance—Inter-Rater Concordance Study. *Educational Sciences: Theory and Practice*, 9(1), 105–125.
- Blanco, B. E. (2009). La desigualdad de resultados educativos: aportes a la teoría desde la investigación sobre eficacia escolar. *Revista mexicana de investigación educativa*, 14(43), 1019-1049. Recuperado de [http://www.scielo.org.mx/scielo.php?pid=S1405-66662009000400003&script=sci\\_arttext](http://www.scielo.org.mx/scielo.php?pid=S1405-66662009000400003&script=sci_arttext)

- Blommel, M. L., & Abate, M. A. (2007). A rubric to assess critical literature evaluation skills. *American Journal of Pharmaceutical Education*, 71(4). Recuperado de <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1959203/>
- Boston, C. (2002). Understanding Scoring Rubrics: A Guide for Teachers. *ERIC Cleringhouse on Assesment and evaluation*. Recuperado de <http://eric.ed.gov/?id=ED471518>
- Cabrera, O.M. (2006). Dirección y eficacia escolar, una relación fundamental. REICE: Revista Electrónica Iberoamericana sobre la Calidad, Eficacia y Cambio en Educación, 4 (4), 1-10.
- Campaña, K., & Tapia, H. (2010). Concepto y características de las buenas prácticas. Recuperado de <http://www.saladeprofes.cl/se-dice/449-concepto-y-caracteristicas-de-las-buenas-practicas.html>
- Cano García, E. (2015). Las rúbricas como instrumento de evaluación de competencias en Educación Superior ¿Uso o abuso? *Profesorado*, 19(2), 265-280. Recuperado de <http://hdl.handle.net/10481/37376>
- Cardemil, C., Maureira, F., & Zuleta, J. (2012). Resiliencia y eficacia escolar. Cuaderno de educación, 45. Recuperado de <http://www.cide.cl/documentos/articulo%20resiliencia%2045final2x.pdf>
- Carrasco, A. (2008). Investigación en efectividad y mejora escolar: ¿nueva agenda? REICE: *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*. Recuperado de <https://repositorio.uam.es/handle/10486/661141>
- Carvallo-Pontón, M. (2010). Eficacia escolar: antecedentes, hallazgos y futuro. *MAGIS. Revista Internacional de Investigación En Educación*, 3 (5), 199–214. Recuperado de <http://www.redalyc.org/pdf/2810/281023476012.pdf>

- Chaparro, A., Gutiérrez, G., Gómez-Gloria, J., Murillo, O., Arregui, I., Lizasoain, L., y Caso, J. (2015). Prácticas escolares en escuelas secundarias de alto valor añadido en Baja California. *UEE RT 15-01*. Ensenada, México. Universidad Autónoma de Baja California.
- Creemers, L.S., Reezigt G. (1997) Effective School Improvement - Ingredients for Success. The Results of an International Comparative Study of Best Practice Case Studies. Recuperado de [https://www.rug.nl/staff/b.p.m.creemers/effective\\_school\\_improvement\\_ingredients\\_for\\_success.pdf](https://www.rug.nl/staff/b.p.m.creemers/effective_school_improvement_ingredients_for_success.pdf)
- Company, P., Contero, M., Otey, J., & Plumed, R. (2015). Approach for developing coordinated rubrics to convey quality criteria in MCAD training. *Computer-Aided Design*, 63, 101–117. doi:10.1016/j.cad.2014.10.001
- Contreras Valenzuela, I., Rittershaussen Klaunig, S., Correa Molina, E., Solís Zañartu, C., Núñez Vega, C., & Vásquez Lara, N. (2013). IOC, un instrumento para cualificar desempeño docente en aula: Su generación y validación. *Estudios Pedagógicos (Valdivia)*, 39(2), 85–96. Recuperado de [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-07052013000200006](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-07052013000200006)
- Cornejo Chávez, R., y Redondo Rojo, J. M. (2007). Variables y factores asociados al aprendizaje escolar: una discusión desde la investigación actual. *Estudios pedagógicos (Valdivia)*, 33(2), 155-175. Recuperado de <http://dx.doi.org/10.4067/S0718-07052007000200009>
- Cruz de la, G. (2011). La construcción y aplicación de rúbricas: una experiencia en la formación de psicólogos educativos. *OBSERVAR*. 5, 21-41. Recuperado de <http://www.odas.es/site/new.php?nid=37>
- De Miguel, C. R. (2009). Las escuelas eficaces: un estudio multinivel de factores explicativos del rendimiento escolar en el área de matemáticas. *Revista de*

- educación*, (348), 355-376. Recuperado de [http://www.revistaeducacion.mec.es/re348/re348\\_15.pdf](http://www.revistaeducacion.mec.es/re348/re348_15.pdf)
- Egido Gálvez, I. Las relaciones entre familia y escuela. Una visión general. *Ministerio de Educación*. Recuperado de <http://www.mecd.gob.es/dctm/cee/encuentros/23encuentro/23encuentroceaedocumentobase2015.pdf?documentId=0901e72b81cba426>
- Ferreira, H. A., Salgueiro, A. M., y Castagno, F. (2013). Autoevaluación institucional: desafíos de articular procesos y resultados en el sistema educativo de la Provincia de Córdoba, Argentina (2008-2011). *Educación Y Desarrollo Social*, 7(1), 30–40.
- Fondo Nacional de Desarrollo de la Educación Peruana. (2014). *Marco de la Innovación y las buenas prácticas Educativas en el Perú*. Recuperado del sitio de internet de la FONDEP: <http://www.fondep.gob.pe>
- Gairín, J. (2012). La organización y atención a la diversidad en centros de educación secundaria en Iberoamérica. Santiago de Chile: Santillana.
- Garay, S., y Uribe, M (2006). Dirección escolar como factor de eficacia y cambio: Situación de la dirección escolar en Chile. *REICE: Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 4 (4), 39-64.
- García-Cabrero. (2010). Modelos teóricos e indicadores de evaluación educativa. *Sinéctica*, (35), 1-17. Recuperado de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1665-109X2010000200005&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-109X2010000200005&lng=es&tlng=es).
- García-Ros, R. (2011). Análisis y validación de una rúbrica para evaluar habilidades de presentación oral en contextos universitarios. *Electronic Journal of Research in Educational Psychology*, 9(3), 1043–1062. Recuperado de <http://www.redalyc.org/pdf/2931/293122852004.pdf>

- García-Ros, R., Fuentes, M. C., González, E., Molina, G., Moya, L., Natividad, L., & Sánchez, P. (2012). Designing and using rubrics in higher education: An innovation project in the psychology degree. *Electronic Journal of Research in Educational Psychology*, 10(3), 1477–1492. Recuperado de <http://docplayer.es/11106571-Designing-and-using-rubrics-in-higher-education-an-innovation-project-in-the-psychology-degree.html>
- Gatica-Lara, F., y Uribarren-Berrueta, T. del N. J. (2013). ¿Cómo elaborar una rúbrica? *Investigación En Educación Médica*, 2(5), 61–65. Recuperado de [http://riem.facmed.unam.mx/sites/all/archivos/V2Num01/10\\_PEM\\_GATICA.PDF](http://riem.facmed.unam.mx/sites/all/archivos/V2Num01/10_PEM_GATICA.PDF)
- Gordillo, J. J. T., y Vázquez, E. A. H. (2012) Validez y fiabilidad de la rúbrica. Recuperado de <http://gtea.uma.es/congresos/CDROM/comunicaciones/carpeta1/6-juan-jesus-torres-gordillo.pdf>
- Grau, S., y Álvarez, J. D. (2005). La autoevaluación de un centro educativo como medida para mejorar su calidad educativa. Recuperado de <http://rua.ua.es/dspace/bitstream/10045/12104/1/Article3.pdf>
- Gros, B., Fernández-Salineró de Miguel, C., Martínez, Martín, M., y Roca Casas, E. (2013). El liderazgo educativo en el contexto escolar.
- Hernández, C. (2009). Prácticas prometedoras de desempeño Institucional asociadas al nivel de logro Académico de las escuelas. Lecciones desde Escuelas secundarias con mejores niveles de Rendimiento en el estado de chihuahua. X Congreso Nacional De Investigación Educativa área 13: política y gestión. Recuperado de [http://www.comie.org.mx/congreso/memoriaelectronica/v10/pdf/area\\_tematica\\_13/ponencias/0023-F.pdf](http://www.comie.org.mx/congreso/memoriaelectronica/v10/pdf/area_tematica_13/ponencias/0023-F.pdf)
- Hernández-Castilla, R., Murillo, F. J., y Martínez-Garrido, C. (2016). Factores de ineficacia escolar. REICE. *Revista Iberoamericana Sobre Calidad, Eficacia Y*

- Cambio En Educación*, 12(1). Recuperado de <https://revistas.uam.es/index.php/reice/article/view/2867>
- Hernández-Rivera, P. (2012). ¿Por qué es importante establecer una rúbrica de evaluación? El caso del curso Clínica de Exodoncia y Cirugía. *Revista Educación*, 36(1). Recuperado de <http://revistas.ucr.ac.cr/index.php/educacion/article/view/452>
- Horbath, J. E., y Gracia, M. A. (2014). Educational evaluation in México. *Revista de Relaciones Internacionales, Estrategia y Seguridad*, 9(1), 59–85.
- Horn Kupfer, A., & Marfán Sánchez, J. (2010). Relación entre liderazgo educativo y desempeño escolar: Revisión de la investigación en Chile. *Psicoperspectivas. Individuo Y Sociedad*, 9(2). doi: 10.5027
- Intxausti, N., Joaristi, L., & Lizasoain, L. (2015). Educational leadership as best practice in highly effective schools in the Autonomous Region of the Basque County (Spain). *Educational Management Administration & Leadership*. doi: 1741143214558570
- Irlles, M. G., Ortells, J. S., de la Calle, F. M., & de la Sen Fernández, M. L. (2011). La rúbrica de evaluación como herramienta de evaluación formativa y sumativa. *IX Jornades de xarxes d'investigació en docència universitària: Disseny de bones pràctiques docents en el context actual*, 144. Recuperado de <http://web.ua.es/va/ice/jornadas-redes-2011/documentos/posters/184446.pdf>
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2 (2), 130–144. doi: 10.1016/j.edurev.2007.05.002
- Lima-Rodríguez, J. S., Lima-Serrano, M., Ponce-González, J. M., y Guerra-Martín, M. D. (2015). Diseño y validación de rúbricas para evaluarlas competencias

- prácticas en estudiantes de enfermería. 29 (1), 119-133. Recuperado de <http://www.ems.sld.cu/index.php/ems/rt/printerFriendly/368/23396+3>
- Lizasoain, L., & Angulo, A. (2014). Buenas prácticas de escuelas eficaces del País Vasco. Metodología y primeros resultados. *Participación Educativa*. 3 (4), 17-27  
Recuperado de [http://ntic.educacion.es/cee/revista/pdfs/n4art\\_lizasoain\\_angulo.pdf](http://ntic.educacion.es/cee/revista/pdfs/n4art_lizasoain_angulo.pdf)
- Lizasoain, L., Intxausti, N., Azpillaga, V., Etxeberria, F., Del Frago, R., y Joaristi, L. (2012). Detección, caracterización y buenas prácticas de centros escolares de alto valor añadido de la comunidad autónoma vasca. En XV Congreso Nacional y V Iberoamericano de Pedagogía. Burgos. Recuperado de <http://www.mecd.gob.es/dctm/cee/encuentros/xxiiencuentro/informepvbbppcentrosaltovalor.pdf?documentId=0901e72b8181d808>
- López, P. (2010). Variables asociadas a la gestión escolar como factores de calidad educativa. *Estudios pedagógicos (Valdivia)*, 36(1), 147-158. Recuperado de [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-07052010000100008](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-07052010000100008)
- López de Mesa-Melo, C., Carvajal-Castillo, C. A., Soto-Godoy, M. F., Urrea-Roa, P. N. (2013). Factores asociados a la convivencia escolar en adolescentes. *Educ. Educ.* (16), 3, 383-410. Recuperado de <http://www.scielo.org.co/pdf/eded/v16n3/v16n3a01.pdf>
- López, M. L., & Velasco, V. (2013). La construcción y evidencias de validez de una rúbrica analítica para la evaluación de ensayos. *REVALUE*, 1(2), 1-21.  
Recuperado de <http://revalue.mx/revista/index.php/revalue/article/viewArticle/35>
- Martínez-Arias, R. (2014). Normas técnicas de los test psicológicos y educativos. *Educación y Orientación: La revista de copoe*. (1), 28-35. Recuperado de

- [http://www.copoe.org/revista-copoe-educar-orientar/Educacion-Orientar\\_COPOE\\_n1\\_Noviembre2014.pdf](http://www.copoe.org/revista-copoe-educar-orientar/Educacion-Orientar_COPOE_n1_Noviembre2014.pdf)
- Martínez, F., Chaparro, A., Lizasoain, L., Caso, J. y Urias, E. (2014). *Caracterización de las buenas prácticas en escuelas secundarias de alto valor añadido de baja California: Etapa I. Identificación de Escuelas de Alto y Bajo residuo* (UEE Reporte Técnico 14). Recuperado del sitio de internet de Universidad Autónoma de Baja California, Unidad de Evaluación Educativa:  
<http://uee.uabc.mx/uee/nivel1.php?p=estudiosYproyectos/reportesTecnicos>
- Martínez-Garrido, C. y Murillo Torrecilla, F. Javier (2016). Factores de eficacia escolar en Venezuela. Areté. *Revista Digital del Doctorado en Educación de la Universidad Central de Venezuela*, 2 (1), 73 – 90. Recuperado de [http://190.169.94.12/ojs/index.php/rev\\_arete/article/view/10771](http://190.169.94.12/ojs/index.php/rev_arete/article/view/10771)
- Martínez-Rojas, J. G. (2008). Construcción y uso de rúbricas de evaluación. *Avances en Medición*, 6, 129-138. Recuperado de <http://www.uaem.mx/sites/default/files/facultad-de-medicina/descargas/construccion-y-uso-de-rubricas-de-evaluacion.pdf>
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25), 1-10.
- Ministerio de Educación, Cultura y Deporte (2014). Las escuelas de éxito. Características y Experiencias. *XXII Encuentro de Consejos Escolares Autonómicos y del Estado*. Recuperado de <http://ntic.educacion.es/cee/jpa/22encuentro/index.html>
- Moreno-Olivos, T. (2009). La evaluación del aprendizaje en la universidad: tensiones, contradicciones y desafíos. *Revista mexicana de investigación educativa*, 14(41), 563-591. Recuperado de [http://www.scielo.org.mx/scielo.php?pid=S1405-66662009000200010&script=sci\\_arttext&tlng=en](http://www.scielo.org.mx/scielo.php?pid=S1405-66662009000200010&script=sci_arttext&tlng=en)

- Muñoz, C., & Valenzuela, J. (2015). Características Psicométricas de una Rúbrica para Evaluar Expresión Escrita a Nivel Universitario. *Formación Universitaria*, 8(6), 75–84. doi: 10.4067/S0718-50062015000600010
- Muñoz-Repiso, M., Cerdán, J., Murillo, F.J., Calzón, J., Castro, M., Egido, I., García, R. y Lucio-Villegas, M. (1995). Calidad de la educación y eficacia de la escuela. Estudio sobre la gestión de los recursos educativos. *Madrid: CIDE*. Recuperado de [https://www.uam.es/personal\\_pdi/stmaria/jmurillo/documentos/Calidad.pdf](https://www.uam.es/personal_pdi/stmaria/jmurillo/documentos/Calidad.pdf)
- Muñoz-Repiso, M., Murillo, F., Barrio, R, Brioso, M. J, Hernández, M. y Pérez-Albo, M. (2000). La mejora de la eficacia escolar: un estudio de casos. *Ministerio de Educación*, (144). Recuperado de [http://books.google.com/books?hl=en&lr=&id=6ONaLOVjIIEC&oi=fnd&pg=PA1&dq=%22Ander+Deuna+Ikastola+\(Sopelana,+Pa%C3%ADs%22+%22clave+del+marco+de+evaluaci%C3%B3n+de+los+programas%22+%22y+la+atenci%C3%B3n+a+la+diversidad+en+el+Centro%22+%22Caso+5:+Las+nuevas+ense%C3%B1anzas+de+Educaci%C3%B3n%22+&ots=Af3bxe9J-f&sig=JDTk9hrAiw\\_wc14M2VeZbfbPk](http://books.google.com/books?hl=en&lr=&id=6ONaLOVjIIEC&oi=fnd&pg=PA1&dq=%22Ander+Deuna+Ikastola+(Sopelana,+Pa%C3%ADs%22+%22clave+del+marco+de+evaluaci%C3%B3n+de+los+programas%22+%22y+la+atenci%C3%B3n+a+la+diversidad+en+el+Centro%22+%22Caso+5:+Las+nuevas+ense%C3%B1anzas+de+Educaci%C3%B3n%22+&ots=Af3bxe9J-f&sig=JDTk9hrAiw_wc14M2VeZbfbPk)
- Murillo, F., Barrio, R. y Pérez-Albo, M. (1999). La dirección escolar. Análisis e investigación. *Ministerio de Educación y Cultura, CIDE*. 136. Recuperado de: [https://www.uam.es/personal\\_pdi/stmaria/jmurillo/recursos/Documentos/Murillo%20\(1999\).pdf](https://www.uam.es/personal_pdi/stmaria/jmurillo/recursos/Documentos/Murillo%20(1999).pdf)
- Murillo, F. J. (2003). El movimiento teórico-práctico de mejora de la escuela. Algunas lecciones aprendidas para transformar los centros docentes. *REICE: Revista Electrónica Iberoamericana Sobre Calidad, Eficacia Y Cambio En Educación*. 1 (2), 1-22 Recuperado de <https://repositorio.uam.es/handle/10486/660688>
- Murillo, F. (2005). El movimiento de investigación de Eficacia Escolar. La investigación sobre Eficacia Escolar en Iberoamérica. Revisión internacional del estado del arte. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en*

- Educación*, 1 (2). Recuperado de <http://www.ice.deusto.es/RINACE/reice/vol1n2/Rec3.htm>
- Murillo, F. J. (2008). Hacia un modelo de eficacia escolar. Estudio multinivel sobre los factores de eficacia de las escuelas españolas. *REICE: Revista Electrónica Iberoamericana Sobre Calidad, Eficacia Y Cambio En Educación*, 6 (1), 4–28. Recuperado de <http://www.rinace.net/arts/vol6num1/art1.pdf>
- Murillo, F.J. (2011). Mejora de la Eficacia Escolar en Iberoamérica. *Revista Iberoamericana de Educación*, 55, 49-83. Recuperado de [http://rieoei.org/rie\\_contenedor.php?numero=rie55](http://rieoei.org/rie_contenedor.php?numero=rie55)
- Murillo, F. J., y Krichesky, G. J. (2015). Mejora de la escuela: medio siglo de lecciones aprendidas. *REICE: Revista Electrónica Iberoamericana Sobre Calidad, Eficacia Y Cambio En Educación*, 13(1), 69–102. Recuperado de [https://www.uam.es/personal\\_pdi/stmaria/jmurillo/arts/reice/vol13num1\\_art5.pdf](https://www.uam.es/personal_pdi/stmaria/jmurillo/arts/reice/vol13num1_art5.pdf)
- Murillo, F. J., y Román, M. (2010). Retos en la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de Educación*, (53), 97–120. Recuperado de <http://www.ceppe.cl/articulos-politicas-educativas/708-murillo-fj-y-roman-m-2010-retos-en-la-evaluacion-de-la-calidad-de-la-educacion-en-america-latina-revista-iberoamericana-de-educacion-53-pp-97-120>
- Musial, J. L., Rubinfeld, I. S., Parker, A. O., Reickert, C. A., Adams, S. A., Rao, S. & Shepard, A. D. (2007). Developing a Scoring Rubric for Resident Research Presentations: A Pilot Study. *Journal of Surgical Research*, 142(2), 304–307. doi: 10.1016/j.jss.2007.03.060
- Ortega, E. M., Salazar, D. M., y Campos, A. L. (2011). Proposal for self-evaluation in educational organizations. *Gestión de la Educación*, 1–27.
- Perassi, Z. (2010). ¿En qué medida la evaluación colabora con la mejora escolar? *Revista Iberoamericana de Educación*, 54(4), 8. Recuperado de

- <https://dialnet.unirioja.es/servlet/articulo?codigo=3698298&orden=305911&info=link>
- Pérez, A. E. R. (2016). Tiempo de aprender: el aprovechamiento de los periodos en el aula. *Revista Mexicana de Investigación Educativa*, 21(69), 611. Recuperado de <http://www.redalyc.org/pdf/140/14045395011.pdf>
- Picón, E. (2013). La rúbrica y la justicia en la autoevaluación. *Íkala, Revista de Lenguaje y Cultura*, 18(3), 79-94. Recuperado de <http://www.scielo.org.co/pdf/ikala/v18n3/v18n3a6.pdf>
- Poggi, M. (2013). Hacia la construcción de nuevas estrategias de evaluación de la calidad educativa en América Latina. 33–46. Recuperado del sitio web de UNESCO <http://unesdoc.unesco.org/images/0017/001776/177648S.pdf>
- Racynski, D y Muñoz, G. (2005). Efectividad Escolar y cambio educativo en condiciones de pobreza en Chile. *Ministerio de Educación*. 15-23. Recuperado de <http://www.gse.upenn.edu/pdf/Efectividad%20Escolar.pdf>
- Ritacco-Real, M. (2011). El liderazgo de los centros educativos y las buenas prácticas ante el fracaso escolar y la exclusión social en la comunidad de Andalucía. *Revista de Investigación en Educación*, 9 (1), 157-167.
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435–448. doi: 10.1080/02602930902862859
- Reynolds, D., Sammons, P., & Townsend, T. (2011). A State of the Art Review. *Educational Effectiveness Research (EER)*. Recuperado de [http://www.icsei.net/icsei2011/State\\_of\\_the\\_art/State\\_of\\_the\\_art\\_Session\\_A.pdf](http://www.icsei.net/icsei2011/State_of_the_art/State_of_the_art_Session_A.pdf)

- Reynolds, D., & Teddlie, C. (2000). 12 The Future Agenda for School Effectiveness Research. *The international handbook of school effectiveness research*, 322.
- Román, M. (2011). Autoevaluación: estrategia y componente esencial para el cambio y la mejora escolar. *Revista Iberoamericana de Educación*, (55), 107–136. Recuperado de [http://www.rinace.net/riee/numeros/vol6-num2/Riee\\_6,2.pdf](http://www.rinace.net/riee/numeros/vol6-num2/Riee_6,2.pdf)
- Román, M. (2008). Estrategias e instrumentos para la mejora de la eficacia de la escuela. *Cuadernos de Educación*, 1(9), 1-18. Recuperado de [http://portales.mineduc.cl/usuarios/convivencia\\_escolar/doc/201103070159510.Planes\\_de\\_Mejoramiento\\_Estrategias\\_e\\_Instrumentos\\_para\\_la\\_mejora\\_de\\_la\\_eficacia\\_de\\_la\\_Escuela.pdf](http://portales.mineduc.cl/usuarios/convivencia_escolar/doc/201103070159510.Planes_de_Mejoramiento_Estrategias_e_Instrumentos_para_la_mejora_de_la_eficacia_de_la_Escuela.pdf)
- Rueda-Beltrán, M. (2010). Autoevaluación y autonomía. *Perfiles Educativos*, 32(130), 1–6.
- Ruiz Cuéllar, G. (2012). La Reforma Integral de la Educación Básica en México (RIEB) en la educación primaria: desafíos para la formación docente. *REIFOP*, 15 (1), 51-60. Recuperado de [http://www.aufop.com/aufop/uploaded\\_files/articulos/1335398629.pdf](http://www.aufop.com/aufop/uploaded_files/articulos/1335398629.pdf)
- Rul, J., y Zaitegi, N. (2003). Autoevaluación del centro educativo: taller de evaluación, modelo de gestión evaluativa Servicio Central de Publicaciones del Gobierno Vasco. Recuperado del sitio de internet de la Administración de la Comunidad Autónoma del País Vasco: Departamento de Educación, Universidades e Investigación: [http://www.hezkuntza.ejgv.euskadi.eus/contenidos/informacion/dig\\_publicaciones\\_innovacion/es\\_liderazg/adjuntos/200003c\\_Pub\\_EJ\\_autoevaluacion\\_centro\\_c.pdf](http://www.hezkuntza.ejgv.euskadi.eus/contenidos/informacion/dig_publicaciones_innovacion/es_liderazg/adjuntos/200003c_Pub_EJ_autoevaluacion_centro_c.pdf)
- Santiago, P., McGregor, I., Nusche, D., Ravela, P., y Toledo, D. (2014). *Revisiones de la OCDE sobre la Evaluación en Educación*. Recuperado del sitio de internet del

- Instituto Nacional Para La Evaluación de La Educación (INEE):  
<http://www.snte.org.mx/pdf/evaluacionEs.pdf>
- Salvat, B. G., Miguel, C. F. S., Martín, M. M., y Casas, E. R. (2013). El liderazgo educativo en el contexto del centro escolar. *Liderazgo y educación*, 173-192. Servicio de Publicaciones.
- Scanlan, M. (2011). Review of Organizing Schools for Improvement. UCEA Review. Recuperado de [http://epublications.marquette.edu/cgi/viewcontent.cgi?article=1133&context=edu\\_fac](http://epublications.marquette.edu/cgi/viewcontent.cgi?article=1133&context=edu_fac)
- Shipman, D., Roa, M., Hooten, J., & Wang, Z. J. (2012). Using the analytic rubric as an evaluation tool in nursing education: The positive and the negative. *Nurse Education Today*, 32(3), 246-249. Recuperado de [http://www.nurseeducationtoday.com/article/S0260-6917\(11\)00094-3/abstract](http://www.nurseeducationtoday.com/article/S0260-6917(11)00094-3/abstract)
- Schleicher, A., Zoido, P. (2013). La transparencia y la mejora de los sistemas educativos. *Participación educativa*. 2(3), 45-81. Recuperado de <https://sede.educacion.gob.es/publiventa/la-transparencia-y-la-mejora-de-los-sistemas-educativos-el-caso-de-pisa/educacion-transparencia-administrativa/16169>
- Sempere Ortells, J. M., García Irles, M., Marco de la Calle, F. M., & Sen Fernández, M. L. D. L. (2011). La rúbrica de evaluación como herramienta de evaluación formativa y sumativa. *IX Jornades de Xarxes D'investigació En Docència Universitària: Disseny de Bones Pràctiques Docents En El Context Actual*. Recuperado de <http://web.ua.es/va/ice/jornadas-redes-2011/documentos/posters/184446.pdf>
- Simon, M. & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. *Practical Assessment, Research & Evaluation*, 7(18), 103–121. Recuperado de <http://pareonline.net/getvn.asp?v=7&n=18>

- Tierney, R. & Simon, M. (2004). What's still wrong with rubrics: focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*, 9 (2), 1–10. Recuperado de [http://www.asu.edu/courses/asu101/asuonline/temp/whats\\_still\\_wrong\\_with\\_rubrics.pdf](http://www.asu.edu/courses/asu101/asuonline/temp/whats_still_wrong_with_rubrics.pdf)
- Urbietta, J. M. E., Garayalde, K. A., y Losada, D. (2011). Diseño de rúbricas en la formación inicial de maestros/as. *Revista de Formación e Innovación Educativa Universitaria*, 4(3), 156-169.
- Valdés Cuervo, Ángel Alberto, Martín Pavón, Mario José, y Sánchez Escobedo, Pedro Antonio. (2009). Participación de los padres de alumnos de educación primaria en las actividades académicas de sus hijos. *Revista electrónica de investigación educativa*, 11(1), 1-17. Recuperado de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1607-40412009000100012&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1607-40412009000100012&lng=es&tlng=es).
- Verano-Tacoronte, D., González-Betancourt, S. M., Bolívar-Cruz, A., Fernández-Monroy, M. y Galván- Sánchez, I. (2016). Valoración de la competencia de comunicación oral de estudiantes universitarios a través de una rúbrica fiable y válida. *Revista Brasileira de Educação*, 21(64), 39–60.  
doi:10.1590/S1413-24782016216403
- Zorrilla M. F (2008). Reflexiones a partir de la educación básica en México. Sistema de información de tendencias educativas en América Latina. Recuperado de <http://www.siteal.iipe-oei.org>

## Apéndices

**Apéndice A.** Listado de prácticas escolares

**Apéndice B.** Captura de valoraciones de los jueces

**Apéndice C.** Modas y porcentajes

**Apéndice D.** Cuestionario para validación

**Apéndice E.** Captura de valoraciones de los jueces (pertinencia de los niveles y de las prácticas.

**Apéndice F.** Criterios conservados.

Disponibles en:

<https://drive.google.com/open?id=0Bz5uRsoVXpHEVXBpeUVLTExlaG8>