

## DISEÑO Y VALIDACIÓN PILOTO DE UNA RÚBRICA DE RETROALIMENTACIÓN INSTRUCCIONAL

### PILOT DESIGN AND VALIDATION OF A RUBRIC OF INSTRUCTIONAL FEEDBACK

Christian Lazcano<sup>1</sup>, Paulo Volante<sup>2</sup>, Claudia Llorente<sup>3</sup>

**Resumen:** La retroalimentación se ha convertido en una práctica habitual en procesos de mejora docente, sobre todo en el contexto de la observación de aula. Sin embargo, se han identificado escasas herramientas que permitan conocer la calidad del feedback entregado a profesores. Por ello, el presente artículo presenta el proceso de diseño y validación piloto de un instrumento para evaluar y apoyar el desarrollo de habilidades de retroalimentación docente con foco instruccional. El diseño metodológico contempló el análisis de la literatura, la validación de contenidos con académicos y líderes escolares, y el análisis de confiabilidad con evaluadores independientes. Los principales resultados validan el instrumento tanto en contenido como en usabilidad, poniendo énfasis en la necesidad de centrar el foco de la retroalimentación en el estudiante, en el desempeño docente y el contenido curricular de la clase. Por otra parte, el testeo piloto de confiabilidad del instrumento mostró resultados concordantes entre evaluadores, medido a través del índice Kappa. Las principales conclusiones apuntan a la pertinencia de la herramienta, especialmente en procesos de desarrollo profesional de líderes escolares, sobre todo por su énfasis en el carácter pedagógico de la retroalimentación.

**Palabras clave:** gestión, liderazgo, liderazgo instruccional, retroalimentación, validez

**Abstract:** Feedback has become a common practice in teacher improvement processes, especially in the context of classroom observation. However, few tools have been identified that allow to know the quality of the feedback given to teachers. Therefore, this article presents the process of design and pilot validation of an instrument to evaluate and support the development of teacher feedback skills with an instructional focus. The methodological design involved literature review, content validation with academics and school leaders, and reliability analysis with independent evaluators. The main results validate the instrument both in terms of content and usability, emphasising the need to focus the feedback on the student, on teacher performance and on the curricular content of the lesson. On the other hand, the pilot reliability assessment of the instrument showed concordant results among evaluators, measured through the Kappa index. The main conclusions point to the relevance of the tool, especially in professional development processes of school leaders, due to its emphasis on the pedagogical dimension of feedback.

**Key Words:** *management, leadership, instructional leadership, feedback, validity*

Recibido: 18.04.2021 / Aceptado: 12.05.2022

<sup>1</sup> Magíster en Educación, Facultad de Educación Pontificia Universidad Católica de Chile, email: [clazcano@uc.cl](mailto:clazcano@uc.cl)

<sup>2</sup> Doctor en Psicología, Facultad de Educación Pontificia Universidad Católica de Chile, email: [pvolante@uc.cl](mailto:pvolante@uc.cl)

<sup>3</sup> Psicóloga, Facultad de Educación Pontificia Universidad Católica de Chile, email: [celloren@uc.cl](mailto:celloren@uc.cl)

## Introducción

En el marco de su rol directivo, cada vez se exige más a los líderes escolares abordar tareas de orden pedagógico por sobre aquellas puramente de gestión. Prácticas como monitorear el currículum, entregar acompañamiento a profesores y fortalecer el desarrollo profesional docente, han sido relevadas por la investigación como parte esencial del trabajo directivo, debido a sus efectos positivos a nivel de sala de clases y más específicamente sobre el aprendizaje de los estudiantes. En esta misma línea, las prácticas de liderazgo con foco instruccional han comenzado ser consideradas como una tarea esencial dentro de los cargos directivos, tal como se desprende de su inclusión en guías de buenas prácticas y en estándares, tanto internacionales (NPBEA, 2015) como en el contexto local de Chile, en documentos como el Marco para la Buena Dirección y el Liderazgo Escolar (2015), y los Estándares Indicativos de Desempeño para los Establecimientos Educativos y sus Sostenedores (Ministerio de Educación de Chile, 2014).

Una de las prácticas, que ha comenzado a instalarse como parte de este énfasis pedagógico en el cargo directivo, es la observación y retroalimentación de clases. Por ejemplo, de acuerdo con los resultados de la encuesta Teaching and Learning International Survey 2018 (TALIS), 75,8% de los profesores en países OECD reporta haber recibido *feedback* de parte de su director o de algún miembro de su equipo directivo en los 12 meses previos a la consulta (OECD, 2020). En Chile, esta cifra es aún más alta, ascendiendo a un 81,6%. En paralelo, en esta misma encuesta un 68,4% de los directores de Chile declaró entregar retroalimentación a sus docentes en base a las observaciones de clases de forma “frecuente” o “muy frecuente”, una cifra también mayor a la del promedio de la OECD. Sin embargo, pese a los esfuerzos y el tiempo dedicados a implementar esta tarea, actualmente es poca la información disponible sobre cuál es el tipo de retroalimentación que realizan directores, directivos y líderes escolares a sus docentes, en términos de su estructura, características y contenido. Por ejemplo, en Chile, el estudio conducido por Quiroga y Aravena (2021) observó que la retroalimentación de directores es efectiva en mostrar el tipo de conductas a mejorar, pero no es capaz de garantizar espacios de reflexión para el docente que la recibe. Si bien este estudio representa un avance en la comprensión del tipo de *feedback* entregado por directivos a docentes, el desafío sigue estando en poder ampliar las muestras de análisis, así como en conocer más sobre el entrenamiento que los líderes escolares reciben en el área.

En este contexto, se hace relevante contar con instrumentos que permitan conocer cuáles son las habilidades de retroalimentación que poseen directores, directivos y líderes escolares, y que, a su vez, apoyen y orienten el desarrollo de las mismas. Esto, entendiendo que la retroalimentación es una práctica frecuente en el marco de la observación de aula, y que ha sido definida como necesaria dentro de los procesos de la mejora de la enseñanza y el aprendizaje. Por ello, el propósito del presente artículo es el de presentar el proceso de diseño y validación piloto de una herramienta (rúbrica) para la evaluación y desarrollo de competencias de retroalimentación con foco instruccional a docentes. El instrumento fue construido en base a antecedentes conceptuales sobre retroalimentación a profesores, para luego ser sometido a distintas etapas de validación.

La relevancia del estudio radica en la necesidad de contar con una herramienta que oriente la retroalimentación a docentes en el contexto de procesos de observación de aula. Esto, entendiendo que la retroalimentación puede movilizar la mejora pedagógica y el aprendizaje

profesional. De esta forma, se propone un instrumento diseñado con un alto nivel de precisión metodológica y conceptual, que puede ser utilizado en contextos de formación de líderes escolares.

### *Objetivo*

El propósito del artículo es mostrar el proceso piloto de diseño y validación de una rúbrica para la evaluación de competencias de retroalimentación, con el fin de ser utilizada en procesos de observación y retroalimentación de profesores. En específico, el objetivo del instrumento es orientar la retroalimentación ejercida por líderes escolares a sus docentes, tomando como referencia el marco conceptual del liderazgo instruccional. Se espera que la rúbrica pueda ser utilizada en el contexto de procesos de mejora de la enseñanza y el aprendizaje, y de formación de directivos que proveen retroalimentación a docentes.

Para lograr los objetivos, en primer lugar se presenta una revisión de antecedentes clave sobre retroalimentación, particularmente cuando esta se enmarca en procesos de mejora de la enseñanza y el aprendizaje. En segundo lugar, se identificarán las dimensiones que dan forma al instrumento. A continuación, se presentará un proceso de validación de contenido realizado por medio de consultas a líderes y docentes escolares y a un grupo de expertos, para, finalmente, testear la confiabilidad del instrumento mediante la concordancia entre evaluadores (Kappa de Cohen).

### *Antecedentes*

#### *Liderazgo Escolar e Instruccional*

Una de las prioridades centrales del estudio del liderazgo escolar, y en particular, del liderazgo instruccional, ha sido la de describir las prácticas que logran influir positivamente sobre la enseñanza de los docentes y sobre los logros de aprendizaje de sus estudiantes (Volante, 2010; Pérez-García et al, 2018). Si bien el enfoque del liderazgo instruccional (*instructional leadership* en la literatura anglosajona) comparte características con otros modelos de liderazgo, se distingue por su énfasis a pedagógico, orientado a la mejora de la enseñanza y el aprendizaje que ocurre dentro del aula, como también alineando y motivando a los profesores y a otros miembros de la organización escolar (Hallinger y Heck, 2002; Hallinger y Murphy, 1985; Robinson et al, 2008).

En concreto, los denominados “líderes instruccionales” realizan acciones como coordinar y monitorear el currículum y la enseñanza, promover el aprendizaje profesional de los profesores, monitorear el progreso de los estudiantes y establecer un clima adecuado para la enseñanza, entre otras prácticas (Robinson et al., 2008). Lo anterior es concordante con lo planteado en el modelo de liderazgo distribuido y su énfasis en ampliar redes de influencia para el logro de objetivos académicos compartidos (Harris, 2009); y con la idea de que más que prácticas aisladas de liderazgo, éstas funcionan como prácticas que estimulan la acción de otros (Leithwood y Riehl, 2005).

### *Retroalimentación Para El Mejoramiento Escolar*

La retroalimentación puede ser definida como la información entregada por un agente (profesor, director, formador) en relación al desempeño o al entendimiento de otra persona, con el propósito de movilizar cambios (Hattie y Temperley, 2007). En otras palabras, es un proceso interpersonal que puede ocurrir en distintos niveles dentro de las escuelas (por ejemplo entre docentes, de profesores a sus estudiantes, de líderes a docentes) y que tiene como propósito primordial lograr transformaciones en aspectos previamente definidos.

La retroalimentación como herramienta de mejoramiento escolar ha sido reconocida por la investigación como un componente central para influenciar los procesos de enseñanza y aprendizaje (Hattie y Temperley, 2007, Darling-Hammond, 2013). Ya sea de profesores a líderes, de líderes a profesores o de estudiantes a profesores, la retroalimentación se presenta como uno de los cinco factores que más influyen en el logro de los estudiantes (Hattie, 1999).

Se ha reportado que los directores más efectivos son hábiles proporcionando a sus profesores una retroalimentación centrada en la mejora de la enseñanza (Stronge, 2006; U.S. Department of Education Office of Innovation and Improvement, 2004). La retroalimentación ayuda a los docentes a tomar conciencia sobre sus prácticas que no cumplen con estándares predefinidos, lo que les permite iniciar acciones correctivas (Ivancevich y MacMahon, 1982), cerrando la brecha entre el nivel actual de desempeño y el esperado (Hattie y Temperley, 2007). Además, tiene implicancias concretas en el desarrollo de prácticas docentes. Por ejemplo, Lizasoain, Tourón y Sobrino (2015), utilizando datos del cuestionario docente Talis 2013, reportaron una mayor prevalencia de prácticas innovadoras centradas en el estudiante como agente de aprendizaje en profesores que recibían una mayor cantidad y calidad en su retroalimentación. Por su parte, Tuytens y Devos (2011) reportan un mayor involucramiento de los docentes en actividades de desarrollo profesional luego de recibir retroalimentación de parte de sus líderes escolares.

### *Retroalimentación Formativa y Auto-retroalimentación*

La literatura enfatiza que la retroalimentación, utilizada como herramienta de mejora, debe poseer un carácter formativo y constructivo (Ovando, 2005). El propósito de la retroalimentación es mejorar el desempeño y promover aprendizajes, por lo que la retroalimentación formativa buscaría modificar el pensamiento o comportamiento del profesional involucrado con- el propósito de mejorar su práctica de enseñanza (Shute, 2008). La retroalimentación formativa se entiende como la entrega, en forma útil y a tiempo, de información relevante sobre el desempeño (Pajak, 2001; Goldring, et al, 2013), utilizando un sentido de formación y desarrollo constante, y por lo tanto, minimizando el carácter evaluativo.

La perspectiva constructiva de la retroalimentación plantea el uso de información relevante referente al desempeño en el proceso de enseñanza-aprendizaje. Esta información ayuda a introducir modificaciones, corregir errores o involucrarse en el desarrollo profesional que mejore los procesos de enseñanza y aprendizaje (Ovando, 2005).

Por otro lado, las prácticas de retroalimentación por sí mismas no conducen a una mejora duradera en las prácticas docentes. Estas mejoras se darán solo cuando el individuo se auto-

dirija, gestione, monitoree y busque la auto-modificación (Glickman, 2002), lo que implica que la calidad de la retroalimentación debe tener tanto un carácter formativo como auto-reflexivo para los profesores. Hallazgos como los de Ivancevich y McMahon (1982) en los ochenta ya daban cuenta de la importancia de la auto-retroalimentación por sobre incluso la retroalimentación externa, siendo la satisfacción intrínseca y compromiso con el trabajo superior a la retroalimentación externa. Adicionalmente, es sabido que si los profesores no encuentran sentido y utilidad a la retroalimentación que se les está dando, las prácticas que se intenten instalar no producirán cambios duraderos (Kinicki et al., 2004). Esta visión sin duda viene a complementar la idea de que la retroalimentación es relevante cuando el sujeto es capaz de tomar conciencia de su propia necesidad de cambio. Con esto, la retroalimentación debe ser vista como un proceso continuo y que transforma al líder en un coach que utiliza la retroalimentación como una de sus herramientas de mejora, incluso siendo capaz de mejorar su desarrollo como directivo influyendo en las prácticas instruccionales de los profesores y mejorando la confianza con los docentes (Goff et al., 2014).

### *Habilidades de Retroalimentación*

La literatura sobre la retroalimentación para la mejora de la enseñanza ha descrito diversas buenas prácticas. Es recomendable que el proceso comience reforzando positivamente al docente, es decir, entregándole información sobre cuáles fueron sus prácticas exitosas (Blase y Blase, 2004; Danielson y McGreal, 2000). Sin embargo, también se sugiere evitar que los profesores solo reciban comentarios positivos, ya que el foco de la retroalimentación debe ser constructivo, esperando movilizar la mejora (Thurlings et al., 2013). De otro modo, si la retroalimentación es percibida como negativa, existe el riesgo que el evaluado tienda a protegerse, perdiendo el sentido instruccional del proceso (Kanfer y Ackerman, 1989).

Por otra parte, es importante que la retroalimentación sea descriptiva, específica y basada en información observable, proporcionando al docente evidencias que sustenten lo que se va a discutir y a la vez evitando emitir juicios de opinión (Blase y Blase, 2004; Danielson y McGreal, 2000; Feeney, 2007). Es recomendable que quien retroalimenta se centre en uno o dos focos (Marshall, 2005), como una forma de hacer abordable el contenido.

Cuando la retroalimentación se produce en el contexto de procesos de observación de salas de clases, se sugiere que el tema a abordar sea lo que se conoce como el “triángulo instruccional” (Cohen et al., 2003; Müller et al., 2014), es decir, la actividad de los estudiantes, el desempeño del profesor y el contenido curricular de la clase, y las interacciones entre estos elementos. Si por algún motivo, quien retroalimenta decide utilizar elementos externos a la clase para argumentar puntos específicos, debe disponer de evidencia para informar a los profesores (Lochmiller, 2016).

Por lo tanto, al momento de realizar retroalimentación es clave reportar al profesor cuáles son sus tareas y áreas a desarrollar que tendrían más impacto en el aprendizaje de los estudiantes (U.S. Department of Education Office of Innovation and Improvement, 2004; Ovando, 2005). Esto promoverá procesos de aprendizaje y planificación que conducen a cambios en el docente (Khachatryan, 2015).

Los profesores perciben que los evaluadores con demasiado conocimiento en sus propias áreas de contenido tienden a imponer sus percepciones sobre lo que significa una enseñanza

efectiva (Lochmiller, 2015). Por ello, es relevante dejar de lado los juicios y sustentar los comentarios en evidencia de la clase (Blase y Blase, 2004; Marshall, 2005), dejando de lado las expectativas propias sobre lo que representa una buena experiencia de enseñanza.

El desarrollo de un diálogo profesional de alto nivel, en el marco de la retroalimentación, ayuda a crear un lenguaje común sobre la enseñanza y el aprendizaje en la escuela (Kersten y Israel, 2005). Además, contribuye al uso de información que permita corregir y mejorar los procesos de enseñanza y aprendizaje (Ovando, 2005). Finalmente, se espera que quien retroalimenta genere espacios para promover la reflexión del docente sobre su propio aprendizaje (Danielson y McGreal, 2000; Feeney, 2007). Se ha estudiado que la retroalimentación conduce a la mejora una vez que el individuo se auto-dirige, auto-gestiona, auto-monitorea y auto-modifica (Glickman, 2002). Se sugiere que la puesta en marcha de procesos de auto-mejora, tenga una mirada a largo plazo, que establezca un plan de acción con metas orientadas a la mejora (Hattie, 1999).

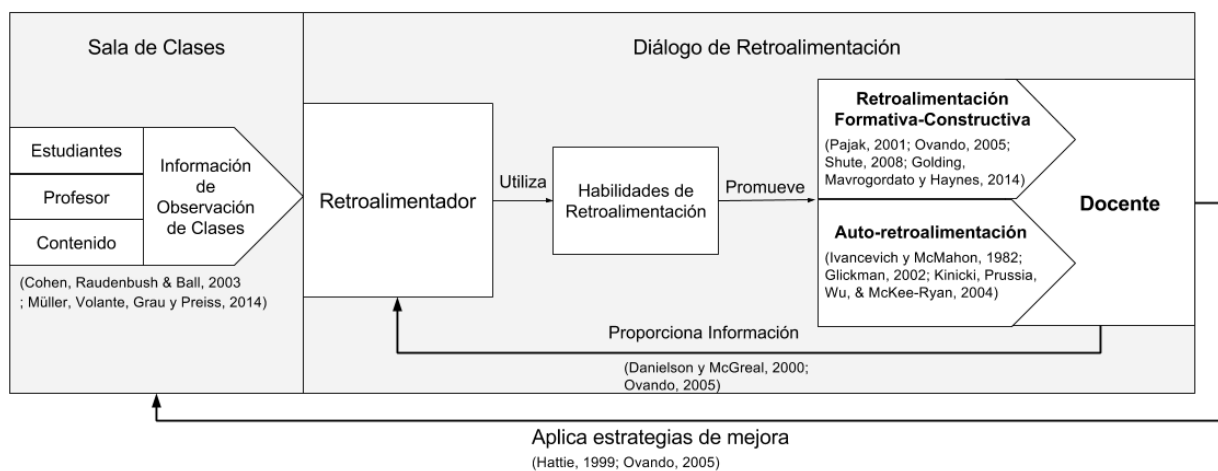
### Metodología

A continuación, se expone la estrategia metodológica utilizada durante el proceso de diseño y validación piloto de la rúbrica de retroalimentación. Para ello, se describen los procedimientos de selección de dimensiones e ítems, validez de contenido y confiabilidad.

#### *Definición de Dimensiones de Retroalimentación.*

Tomando en consideración la evidencia sobre buenas prácticas de retroalimentación, se identificó un set de dimensiones relevantes para llevar a cabo un proceso efectivo. Un líder educativo que busca influir en la mejora del desempeño de otros profesionales tendrá en consideración que requiere habilidades para poder llevar a cabo este proceso de diálogo e influencia con otros docentes.

**Figura 1.** Modelo de retroalimentación instruccional. Elaboración propia.





Las dimensiones fueron discutidas y validadas por el equipo de investigación, siguiendo la teoría de acción que se observa en la Figura 1. Esta teoría de acción posee un foco instruccional y releva la interacción entre el desempeño docente, el contenido curricular y la actividad de los estudiantes. De acuerdo al modelo, mientras realiza la observación de clases, el líder escolar recoge y selecciona información relevante relacionada con el triángulo instruccional. A continuación, elabora una retroalimentación con sentido formativo-constructivo y auto-reflexivo, la que es entregada (de forma oral o escrita) al profesor observado. El docente retroalimentado, posteriormente, proporciona información al líder y, finalmente, aplica estrategias de mejora que potencien su enseñanza, dando inicio a un nuevo ciclo de retroalimentación.

La identificación y síntesis de dimensiones de retroalimentación efectivas fue un primer paso en la construcción del instrumento. Estas dimensiones se configuraron a través del agrupamiento de prácticas de retroalimentación relevantes en la literatura. En cada dimensión se anidaron diferentes prácticas, las que fueron operacionalizadas en ítems.

El instrumento originalmente diseñado constó de 10 ítems distribuidos en 6 dimensiones, cada una representativa de una habilidad detectada en la literatura sobre retroalimentación. En cuanto a la escala de medición, se utilizan 4 niveles (del 1 al 4): 1: “Insatisfactorio”, 2: “Básico”, 3: “Suficiente” y 4: “Sobresaliente”.

#### *Validación de Contenido con Expertos Profesionales.*

La primera versión del instrumento fue sometida a una validación de contenido con docentes y líderes educativos en ejercicio en el sistema escolar chileno, todos ellos con experiencia en proporcionar y recibir retroalimentación en procesos de observación de clases en sus escuelas. El grupo estuvo conformado por 39 personas, 28 mujeres y 11 hombres. Un 60% eran directivos al momento de ser consultados y un 40%, docentes con experiencia previa en cargos de liderazgo. Un 35% pertenecía a colegios de tipo municipal (es decir, pertenecientes al sistema público), 40% a colegios subvencionados (de financiamiento mixto) y un 15% a colegios particulares. En total, los 39 líderes educativos promediaban 12 años de experiencia de trabajo en colegios.

A cada participante del grupo se le entregó una copia en papel de la rúbrica, para que la revisaran con detalle, y se les hizo una serie de preguntas de forma escrita destinadas a evaluar la calidad del instrumento. Estas estaban encabezadas por la pregunta: “¿Se trata de una buena rúbrica para evaluar la retroalimentación de la actividad de enseñanza?”. Las respuestas fueron codificadas, realizándose un análisis según su frecuencia. Esto permitió determinar qué dimensiones e ítems fueron considerados como más pertinentes al momento de evaluar la calidad de la retroalimentación y cuáles presentaban problemas y, en general, qué percepción tuvieron del instrumento. Esta fase ayudó a realizar ajustes al instrumento antes de enfrentarlo a un juicio de expertos.

#### *Validación de Contenidos con Juicio de Expertos Académicos.*

El siguiente paso fue realizar una consulta a expertos en el área del liderazgo escolar y la observación y retroalimentación de clases, con el propósito de evaluar la pertinencia del

contenido de las dimensiones e ítems del instrumento, además de la coherencia y calidad de estos. Esta validación contó con la participación de tres académicos de la Facultad de Educación de una universidad chilena (dos mujeres y un hombre), quienes se desempeñaban como docentes en un programa de formación directiva con énfasis en liderazgo para la mejora pedagógica y un profesional participante del programa, con amplia experiencia en trabajo con líderes educativos.

El procedimiento se realizó de manera individual y sin contacto entre evaluadores. Los resultados fueron compilados en una ficha de validación, siguiendo una versión adaptada del modelo propuesto por Corral (2009). Esta ficha permitió calcular porcentajes de acuerdo respecto a diferentes dimensiones de evaluación (claridad de la redacción, coherencia interna, inducción a la respuesta (sesgo), lenguaje adecuado y pertinencia del ítem), y registrar observaciones cualitativas que permitieran expresar comentarios para cada ítem y pregunta cerrada. La decisión de modificar los ítems se hizo bajo el criterio de acuerdo porcentual para cada una de las preguntas cerradas, codificadas con 0 y 1 (Sí y No), sometiendo a revisión todos los ítems que presentaran menos del 75% de acuerdo, lo que es igual a decir que hay inconsistencia entre 3 de los 4 evaluadores. Por otra parte, al revisar la estructura de los ítems se tomó en consideración las observaciones cualitativas. Este proceso de validación no presentó diferencias sustantivas con el ya realizado con los expertos profesionales, por lo que se decidió utilizar la versión ajustada a partir de los comentarios de los expertos académicos como la final.

#### *Evaluación de Confiabilidad con Evaluadores Independientes*

Finalizado el proceso de validación de contenido, el instrumento modificado fue sometido a una evaluación de confiabilidad. Para ello, se solicitó a evaluadores entrenados que observaran grabaciones en video de retroalimentaciones orales realizadas por líderes escolares chilenos, y que las evaluaran utilizando la rúbrica.

En concreto, se convocó a 3 evaluadores mediante un concurso abierto y patrocinado por un proyecto de investigación (FONDEF ID14I10139: Prototipo y validación de una plataforma virtual de competencias directivas de liderazgo instruccional para contextos escolares basado en métodos de análisis de desempeño directivo). Los evaluadores tuvieron 2 sesiones de inducción al instrumento (con un total de 5 horas), durante las que se les explicaron las dimensiones e ítems de la herramienta y se les entrenó en su uso. Para ello, se les solicitó que observaran 4 retroalimentaciones videogradas, realizadas por líderes escolares de establecimientos escolares chilenos. Luego de cada video, cada evaluador debió calificar de forma independiente la calidad de la retroalimentación realizada por el líder escolar, utilizando la rúbrica. Al completar la evaluación de los primeros 4 videos, se encontraron niveles de concordancia aceptables entre los tres evaluadores. Por ello, se procedió a examinar la confiabilidad del instrumento por medio de 10 nuevos videos, que los evaluadores debieron observar y calificar de manera independiente.

El procedimiento estadístico para estimar consistencia y acuerdo entre evaluadores se hizo mediante el coeficiente Kappa (Cohen, 1960), puesto que permite establecer un ajuste de efectos dados por el azar en una proporción de la concordancia observada. Los valores



deseados debían ser de 0,7 a 1; lo que indicaría un nivel de concordancia entre evaluadores: “considerable” (0,61-0,80) y “casi perfecto” (0,81-1) (Landis y Koch, 1977; El Emam, 1999).

## Resultados

A continuación, se presentan los resultados de los pasos descritos en la metodología, vinculados a los objetivos inicialmente formulados.

### *Selección Habilidades de Retroalimentación*

Luego de analizar la bibliografía sobre retroalimentación, se definieron 6 grandes habilidades, presentadas en la Tabla 1. Cada una de estas habilidades (dimensiones) anidan a los ítems que dan forma al instrumento. Cada ítem describe prácticas específicas de retroalimentación.

**Tabla 1:** Habilidades de retroalimentación extraídas desde la literatura. Elaboración propia.

Habilidades (Dimensiones)	Referencias
H1. Enfocar la conversación en torno a puntos específicos y observables (triángulo instruccional)	Cohen et al., 2003; Blase y Blase, 2004; Marshall, 2005; Feeney, 2007; Müller et al., 2014
H2. Reforzar las buenas prácticas realizando críticas constructivas	Danielson y McGreal, 2000; Blase y Blase, 2004 y Thurlings et al., 2014
H3. Centrarse en los detalles instruccionales del profesor	U.S Department of Education Office of Innovation and Improvement, 2004, Ovando, 2005 y Khachatryan, 2015
H4. Suspender las experiencias sin fundamento sobre la instrucción del profesor	Blase y Blase, 2004; Marshall, 2005 y Lochmiller, 2016
H5. Dar oportunidad para que el profesor entregue información relevante	Danielson y McGreal, 2000; Kersten y Israel, 2005 y Ovando, 2005
H6. Promover la auto-reflexión dentro de la retroalimentación llegando a acuerdos de mejora con el docente	Hattie, 1999; Glickman, 2002; Ovando, 2005 y Feeney, 2007

La primera versión del instrumento estuvo constituida por 6 dimensiones, cada una representativa de una habilidad de retroalimentación. Estas, en total, anidaron 10 prácticas, representadas por un ítem cada una. Las habilidades fueron ordenadas en una secuencia lógica de inicio, desarrollo y cierre, homologando los momentos de una conversación de retroalimentación. Esta versión fue sometida a prueba en el proceso de validación de contenido con líderes del sistema escolar.

*Validación de contenido con Expertos Profesionales*

Una vez construida la primera versión del instrumento, fue sometido a un proceso de validación con líderes escolares y docentes. En cuanto a su calidad (Tabla 2), un 25,6% del grupo consideró que la rúbrica era un buen instrumento para evaluar retroalimentación. Un 56,4% aprobó su calidad, pero recomendó realizar modificaciones con el propósito de mejorar aspectos puntuales.

**Tabla 2:** Evaluación de calidad de la rúbrica. Elaboración propia.

Categoría de agrupación	Frecuencia	Porcentaje
Es una buena rúbrica para evaluar la retroalimentación	10	25,6
Es una buena rúbrica para evaluar retroalimentación, aunque deben realizarse modificaciones	22	56,4
No es buena rúbrica para evaluar la retroalimentación	3	7,7
No deja claro una postura frente al instrumento	1	2,6
No contesta	3	7,7
<b>Total</b>	<b>39</b>	<b>100</b>

**Tabla 3:** Problemáticas detectadas y sugerencias – Primera versión de la rúbrica. Elaboración propia.

Propuestas de cambio	Frecuencia	Porcentaje
La descripción de los niveles es extensa	8	28,6
Faltan plazos y tiempo para los acuerdos de mejora	8	28,6
Mejorar redacción de algunos ítems	2	7,1
Cuesta leer descripciones	1	3,6
Complementar la rúbrica con una lista de cotejo	1	3,6
Falta evaluar el clima de aula	1	3,6
La habilidad 3 debería ir al inicio de la rúbrica	1	3,6
Las debilidades deberían agruparse entorno al triángulo	1	3,6

Focalizarse en el uso de evidencia	1	3,6
Aumentar indicadores con detalle en lo instruccional	1	3,6
Falta saber acerca de cómo se siente el profesor	1	3,6
Agregar mayor detalle al ítem 6	1	3,6
“Suspender las creencias” no es una frase apropiada para la	1	3,6
<b>Total</b>	<b>28</b>	<b>100</b>

Respecto a las problemáticas encontradas y posibles propuestas de cambio (ver Tabla 3), es posible observar que la extensión de la descripción de cada nivel de desempeño es el problema más mencionado, junto con la falta de una mención a plazos y tiempos para los acuerdos de mejora (dentro de la dimensión nº 6). Estas sugerencias buscaron simplificar la lectura del instrumento e incorporar elementos que aseguren la ejecución de acciones de mejora.

Con estos resultados, se tomó la decisión de modificar el instrumento en base a las dos propuestas de mejora más frecuentes. De esta forma, (a) se acortó la descripción de los niveles de desempeño en todos los ítems, y (b) se incorporó un ítem adicional dentro de la dimensión 6 (Promover la auto-reflexión dentro de la retroalimentación llegando a acuerdos de mejora con el docente), que recogiera el establecimiento de acuerdos de mejora en plazos acotados entre quien retroalimenta y quien recibe la retroalimentación. El nuevo ítem (11) se definió como: “Se fija un plan de trabajo conducente a mejorar las prácticas docentes de aula en plazos definidos y consensuados”.

La nueva versión de la rúbrica siguió componiéndose por 6 habilidades, pero ahora con 11 ítems en total, uno más que la versión presentada a los líderes escolares. La nueva estructura de la rúbrica es mostrada en la Tabla 4.

**Tabla 4:** Habilidades de retroalimentación e ítems de la rúbrica. Elaboración propia

Etapa de retroalimentación	Habilidad	Ítem
Inicio	H1: Enfocar la conversación en torno a puntos específicos y observables	1. Captura la atención y confianza del docente a través de un trato respetuoso como parte de la preparación de proceso de retroalimentación
		2. Se asegura de explicitar el objetivo de la conversación en función de analizar la situación de enseñanza, aprendizaje de los estudiantes y propuesta de mejora
		3. Se establece una conversación que gira en torno a los elementos específicos y observables de la clase

	H2: Reforzar las buenas prácticas y proponer críticas constructivas	4. Comunica tempranamente y de forma ordenada las fortalezas y debilidades observadas en la clase
		5. Las fortalezas que se plantean en la retroalimentación son centrales a los núcleos pedagógicos del triángulo instruccional
		6. Las debilidades son comunicadas como críticas constructivas, con un claro énfasis en el desarrollo docente
Desarrollo	H3: Centrarse en los detalles instruccionales del profesor	7. El foco de la retroalimentación se centra en dialogar sobre las prácticas instruccionales del profesor en el aula
	H4: Suspender las experiencias sin fundamento sobre la instrucción del profesor	8. El retroalimentador suspende sus creencias de lo que representa una buena clase mientras dialoga con el docente
	H5: Dar oportunidad para que el profesor entregue información relevante sobre su enseñanza	9. El diálogo durante la retroalimentación es promovido por el retroalimentador con preguntas para que el docente entregue información relevante de la clase y sus prácticas
Cierre	H6: Promover la auto-retroalimentación en los acuerdos de mejora con el docente	10. El retroalimentador promueve instancias para que el docente elabore su propio plan de mejora guiando el proceso de auto-retroalimentación
		11. Se fija un plan de trabajo conducente a mejorar las prácticas docentes de aula en plazos definidos y consensuados

*Validación de Contenido con Expertos Académicos.*

El proceso de validación con expertos académicos mostró que un 100% de ellos consideró que el instrumento era aplicable (aunque atendiendo a algunos comentarios). La tabla 5 permite apreciar mayor detalle, mostrando los porcentajes de acuerdo obtenidos en cada dimensión de la ficha de validación (Claridad en la redacción, Coherencia interna, Inducción a la respuesta, Lenguaje adecuado, Mide lo que pretende). Al ser cuatro los evaluadores expertos, cada cuartil representa un individuo. Se aprecia que el ítem 11 (“Se fija un plan de trabajo conducente a mejorar las prácticas docentes de aula en plazos definidos y consensuados”), agregado tras el proceso la validación con líderes escolares, logra total porcentaje de acuerdo en todas las dimensiones salvo la de redacción. Pese a eso, se levanta como un ítem robusto.

**Tabla 5:** Porcentaje de acuerdo de la ficha de validación. Elaboración propia

Indicador	Claridad en la redacción	Coherencia interna	Inducción a la respuesta (Sesgo)	Lenguaje adecuado	Mide lo que pretende
Ítem 1	100%	100%	25%	100%	75%
Ítem 2	75%	50%	25%	100%	100%
Ítem 3	75%	50%	0%	100%	25%
Ítem 4	100%	25%	0%*	67%*	75%
Ítem 5	100%	100%	0%	100%	75%
Ítem 6	100%	25%	0%	100%	25%
Ítem 7	100%	75%	0%*	100%	75%
Ítem 8	50%	25%	0%	75%	67%*
Ítem 9	25%	75%	0%	100%	75%
Ítem 10	100%	100%	0%	100%	100%
Ítem 11	75%	100%	0%	100%	100%
Promedio rúbrica	82%	66%	5%	95%	72%
Desviación estándar	0,25%	0,32%	0,10%	0,12%	0,26%

\*Contiene un valor perdido

En términos generales, las características mejor evaluadas de la rúbrica son su “Lenguaje adecuado” y su “Claridad en la redacción”, en tanto que la “Coherencia interna” y “Mide lo que pretende” presentan los promedios más bajos. En la variable “Sesgo” era esperable obtener un bajo porcentaje, lo que reflejaría un acuerdo positivo en esta categoría. Al revisar la evaluación de cada ítem de la rúbrica, se sugiere revisar todos, con excepción del 10 y el 11.

Una segunda parte de la ficha de validación entrega porcentajes de concordancia para aspectos generales del instrumento. Un 100% de los evaluadores cree que el instrumento tiene instrucciones claras y precisas para responder el cuestionario; 50% cree que los ítems corresponden a etapas de inicio, desarrollo y cierre; 67% (contiene un valor perdido) señala que los ítems permiten cumplir el logro de los objetivos declarado en cada uno; 75% declara que la distribución es lógica y secuencial y 75% señala que el número de ítems es suficiente para recoger la información.

Con estos resultados, se procedió a realizar mejoras en los ítems que tuvieran menos de un 75% de acuerdo en alguno de los aspectos evaluados. En el caso de sesgo, los que tuvieran más de un 25%. Especial atención recibieron los ítem 3 y 6 en la categoría “Mide lo que pretende”,

los ítems 4, 6 y 8 en “Coherencia interna” y el ítem 9 en “Claridad en la redacción”, los cuales fueron reformulados de manera tanto en redacción como en contenido.

**Tabla 6:** Observaciones de mejora de expertos, agrupadas por categorías (ítems que requieren cambios en la categoría están marcados en negrita). Elaboración propia.

Categorías de mejora	Ítem										
	1	2	3	4	5	6	7	8	9	10	11
Mejorar la estructura de la progresión de niveles de desempeño	1	2	3	4	5	6	7	8	9	10	11
Redactar en positivo niveles de	1	2	3	4	5	6	7	8	9	10	11
Mejorar problemas de redacción en los niveles de desempeño	1	2	3	4	5	6	7	8	9	10	11
Reordenar para responder a estructura inicio, desarrollo, cierre.	1	2	3	4	5	6	7	8	9	10	11
Incorporar elementos de forma a algún nivel de desempeño	1	2	3	4	5	6	7	8	9	10	11
Incorporar elementos instruccionales a algún nivel de desempeño	1	2	3	4	5	6	7	8	9	10	11

La tabla 6 agrupa en categorías los comentarios realizados por el grupo de expertos. En cuanto a las áreas a modificar, más de la mitad de los ítems requirió cambios en los niveles de desempeño, ya sea en su forma o en cuanto a la necesidad de agregar elementos instruccionales. Prácticamente todos los ítems sufrieron algún tipo de modificación. Los ítems 8 y 1 son los que recibieron un mayor número de comentarios, lo que supuso una mayor reestructuración. De hecho, las sugerencias realizadas al ítem 8 implicaron cambiar el nombre de la habilidad, desde “Suspender las experiencias sin fundamento sobre la instrucción del profesor” a “Realizar comentarios sobre la instrucción docente utilizando información detallada y argumentada”, y del ítem desde “El retroalimentador suspende sus creencias de lo que representa una buena clase mientras dialoga con el docente” a “El retroalimentador realiza comentarios al profesor sobre su instrucción utilizando información detallada de lo observado en clase”. Sin embargo, pese al cambio realizado al ítem y su estructura interna, el sustento teórico detrás de este no sufrió modificaciones.

En síntesis, los expertos académicos validaron el contenido, aunque sugirieron modificaciones. Dado que en su mayoría los cambios se concentraron en aspectos formales, el contenido teórico-empírico del instrumento no se vio enjuiciado.



### Confiabilidad del instrumento

Como parte de la medición de confiabilidad, 3 evaluadores independientes evaluaron retroalimentaciones videograbadas. En primer lugar, se realizó una ronda de entrenamiento con 4 videos, cada uno de los cuales fue observado en 2 ocasiones. Como muestra la tabla 7, en general los niveles de concordancia fueron bajos en la primera observación de cada video. Esta situación efectivamente comienza a mejorar al segundo intento con los mismos videos iniciales, alcanzándose niveles de concordancia sobre 0.7 (significativos), lo que llevó a iniciar la evaluación formal de confiabilidad.

**Tabla 7:** Matriz de resultados Kappa. Elaboración propia

Etapa	Nº Videos	Nº Intentos	Evaluador 1 - Evaluador 2	Evaluador 1 - Evaluador 3	Evaluador 2 - Evaluador 3
Entrenamiento	Video 1	1	0,258	0,241	0,154
	Video 1	2	0,863**	0,738**	0,871**
	Video 2	1	-0,236	0,602**	-0,279
	Video 2	2	0,738**	0,866**	0,864**
	Video 3	1	0,298	0,402*	0,267
	Video 3	2	0,756**	0,874**	0,876**
	Video 4	1	0,5**	0,621**	0,633**
	Video 4	2	0,864**	0,864**	0,732**
Aplicación formal	Video 5	1	0,744**	1**	0,744**
	Video 6	1	0,872**	0,872**	1**
	Video 7	1	0,864**	0,861**	0,857**
	Video 8	1	1**	0,864**	0,864**
	Video 9	1	0,876**	0,758**	0,878**
	Video 10	1	0,722**	0,722**	1**
	Video 11	1	0,863**	0,855**	0,728**
	Video 12	1	0,861**	0,853**	0,707**
	Video 13	1	0,871**	0,866**	0,732**

	Video 14	1	1**	0,841**	0,841**
	Video 15	1	0,878**	0,878**	0,758**

\* $p < .05$ , \*\* $p < .01$

En esta etapa de aplicación formal del instrumento, como se aprecia en la misma tabla 7, los niveles de concordancia interevaluador no bajaron de 0.7, siendo el nivel más bajo el que se registra entre los evaluadores 2 y 3 en el video 12 ( $Kappa=0.707$ ). Estos resultados indican que la rúbrica obtuvo un nivel de confiabilidad aceptable luego que los evaluadores pasen por un breve período de entrenamiento.

## Discusión

La literatura muestra acuerdo en la importancia de la retroalimentación en contextos escolares. En concreto, el presente artículo se sitúa en el marco de la retroalimentación con foco en la mejora pedagógica realizada a profesores, específicamente en el contexto de procesos de observación de clases. Desde la literatura y evidencia, se ha descrito que la retroalimentación que consigue promover cambios significativos en la práctica docente, es aquella que se enfoca en críticas constructivas, que dirige la conversación hacia aspectos observables y específicos relacionados con el triángulo instruccional (desempeño del profesor, actividad del estudiante y contenido de la clase), la que evita juicios no fundamentados, provee oportunidades al docente para complementar el *feedback* entregado y promueve la reflexión, movilizandolos acuerdos de mejora. Sin embargo, pocas veces los líderes escolares cuentan con oportunidades de formación específica que les permitan desarrollar sus habilidades de retroalimentación a profesores. Bajo esta perspectiva, el presente artículo presenta el proceso de creación y validación de una herramienta para evaluar y desarrollar competencias de retroalimentación de líderes escolares hacia profesores.

En este contexto, se detallan las distintas fases de la construcción del instrumento. En un primer momento, la etapa de diseño permitió compilar evidencia en torno al tema, identificando habilidades específicas de retroalimentación que, en conjunto, tienen el propósito de contribuir al mejoramiento de la calidad de enseñanza. La revisión de los antecedentes permitió afirmar que las prácticas de retroalimentación centradas en lo formativo (Goff et. al., 2014) y lo constructivo (Ovando, 2005) representan el núcleo central del proceso de retroalimentación con foco en la mejora de prácticas pedagógicas. Las fases de validación permitieron confirmar la relevancia de las habilidades seleccionadas y, una vez introducidos cambios en la redacción de algunos ítems y en el orden en que se muestran en el instrumento, reafirmar que se trata de una herramienta relevante para el uso en el contexto del desarrollo de habilidades de retroalimentación. Más en detalle, cabe destacar que el proceso de validación de expertos permitió eliminar posibles sesgos entre evaluadores que resultarían perjudiciales para el instrumento, lo que es parte central de la idea detrás del método de agregados individuales (Corral, 2009). De esta forma, el análisis sustantivo del instrumento da cuenta que la rúbrica propuesta en este estudio cumple con los criterios de rigor de contenido, que permiten evaluar la práctica de retroalimentación, incorporando un conjunto de habilidades relevantes y específicas.

En cuanto a la confiabilidad, los resultados (coeficiente “Kappa”) muestran que los niveles de acuerdo óptimos entre evaluadores aparecen luego de recibir un entrenamiento breve y al poco tiempo de utilizar la rúbrica. Esto implica que una vez que el instrumento es conocido por los evaluadores, se trata de una herramienta confiable (en su fase piloto), lo que permitiría proyectar su uso para efectos de capacitación y formación.

Estos resultados son particularmente útiles en el contexto actual, donde se considera que la práctica pedagógica, es decir, la enseñanza y el aprendizaje, no es un proceso “privado” ni aislado de cada sala de clases. Cada vez es más frecuente que líderes y docentes se involucren en procesos de observación y retroalimentación docente, aunque en ocasiones sin recibir formación ni orientaciones específicas sobre cómo llevar a cabo estos procesos, de forma que produzcan un verdadero impacto en la práctica en el aula. La “desprivatización” de la enseñanza exige apertura, flexibilidad y disposición a la mejora a los profesores que muestran cómo realizan el ejercicio de la docencia. En respuesta, es esperable que reciban retroalimentación rigurosa, centrada en los elementos centrales de la clase, acotada y útil. Por ello, es importante que existan herramientas que sirvan como un apoyo en el desarrollo de habilidades de retroalimentación, y que permitan que esta tenga un foco definido.

El artículo posee limitaciones, en primer lugar, en términos de que solo se reporta la creación del instrumento y el proceso de validación piloto. Aún es necesario probar su uso en muestras de mayor tamaño, con el fin de validar estadísticamente sus ítems y dimensiones. Por otro lado, para efectos de la investigación solo se testeó su aplicación para evaluar retroalimentaciones simuladas. Esto en sí mismo no constituye una limitante, dado el énfasis de uso en instancias de formación de directores. Sin embargo, en una siguiente etapa, es importante conocer cómo se comporta cuando es utilizada en procesos “reales” de retroalimentación. Un próximo paso, va en la línea de evaluar si el instrumento puede promover efectos reales en la práctica de la retroalimentación, es decir, si a través de su uso es posible impulsar mejoras en la calidad del *feedback*, en cualquiera de las dimensiones que son evaluadas: enfocar la observación en torno a puntos específicos y observables, reforzar buenas prácticas y proponer críticas constructivas, centrarse en los detalles instruccionales del profesor, suspender las experiencias sin fundamento sobre la instrucción del profesor, dar oportunidad para que el profesor entregue información relevante sobre su enseñanza y promover la auto-retroalimentación en los acuerdos de mejora con el docente.

Las proyecciones de uso del instrumento se circunscriben en el marco de programas de entrenamiento específicos en habilidades de retroalimentación para la mejora pedagógica, o más generales, de desarrollo profesional para líderes escolares. También se proyecta el uso para reforzar la implementación de comunidades de práctica o comunidades de aprendizaje profesional, donde la retroalimentación sea un eje central de las posibilidades de cambio y mejora. Si bien para su óptimo uso, la aplicación del instrumento requiere de un proceso de entrenamiento, este es breve y focalizado. Este marco de uso se hace especialmente relevante en el contexto de políticas actuales, donde los procesos de observación y retroalimentación son considerados como un componente importante en los procesos de mejoramiento escolar. Para finalizar, este estudio representa un avance en el esfuerzo por producir herramientas validadas para el ejercicio de la mejora pedagógica y del liderazgo instruccional, desarrollando recursos para la investigación y para la práctica de los líderes escolares.

## Referencias

- Blase, J., & Blase, J. (2004). *Handbook of instructional leadership: How successful principals promote teaching and learning* (2nd ed.). Corwin Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational Psychology Measures*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119-142. <https://doi.org/10.3102/01623737025002119>
- Corral, Y. (2009). Validez y confiabilidad de los instrumentos para la recolección de datos. *Revista Ciencias de la Educación*, (33), 228-247. <https://dialnet.unirioja.es/servlet/articulo?codigo=5362681>
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Association of Supervision and Curriculum Development.
- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. Teachers College Press.
- El Emam, K. (1999). Benchmarking Kappa: Interrater agreement in software process assessment. *Empirical Software Engineering*, 4(2), 113–133. <https://doi.org/10.1023/A:1009820201126>
- Feeney, E. J. (2007). Quality feedback: The essential ingredient for teacher success. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 80(4), 191-198. <https://doi.org/10.3200/TCHS.80.4.191-198>
- Glickman, C. (2002). *Leadership for learning: How to help teachers succeed*. Association of Supervision and Curriculum Development.
- Goff, P., Guthrie, J. E., Goldring, E., & Bickman, L. (2014). Changing principals' leadership through feedback and coaching. *Journal of educational administration*, 52 (5), pp. 682-704. <https://doi.org/10.1108/JEA-10-2013-0113>
- Goldring, E., Mavrogordato, M.C. and Haynes, K.T. (2013). *Multi-source principal evaluation data: how principals approach, interpret and use teacher feedback regarding their leadership effectiveness*. [Paper]. Annual Meeting of the American Education Research Association, San Francisco, CA, USA.
- Hallinger, P. & Heck, R. (2002). What do you call people with visions? The Role of Vision, Mission and Goals in School Leadership and Improvement en K. Leithwood & P. Hallinger (eds.), *Second International Handbook of Educational Leadership and Administration* (pp. 9-40). Kluwer Academic
- Hallinger, P. & Murphy, J. (1985). Assessing the Instructional Management Behavior of principals. *Elementary School Journal*, 86(2), 212-247. <https://doi.org/10.1086/461445>
- Harris, A. (2009). *Distributed Leadership: Different Perspectives*. Springer.
- Hattie, J. (1999). Influences on student learning, [Clase inaugural, Professor of Education], University of Auckland, Auckland, NZ.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>

- Ivancevich, J. M., & McMahon, J. T. (1982). The effects of goal setting, external feedback, and self-generated feedback on outcome variables: A field experiment. *Academy of Management Journal*, 25(2), 359-372. <https://doi.org/10.5465/255997>
- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/ aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74(4), 657-690. <https://doi.org/10.1037/0021-9010.74.4.657>
- Kersten, T. A., & Israel, M. S. (2005). Teacher evaluation: principals' insights and suggestions for improvement. *Planning and Changing*, 36(1-2), 47-67 .
- Khachatryan, E. (2015). Feedback on Teaching From Observations of Teaching What Do Administrators Say and What Do Teachers Think About It?. *NASSP Bulletin*, 99(2), 164-188. <https://doi.org/10.1177/0192636515583716>
- Kinicki, A. J., Prussia, G. E., Wu, B. J., & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees' response to performance feedback. *Journal of Applied Psychology*, 89(6), 1057–1069. <https://doi.org/10.1037/0021-9010.89.6.1057>
- Landis J, Koch G. (1997). The measurement of observer agreement for categorical data. *Biometrics*, 33(1) 159-74. <https://doi.org/10.2307/2529310>
- Leithwood, K. & Riehl, C. (2005). What do we already know about educational leadership? en W. Firestone & C. Riehl, *A New Agenda for Research in Educational Leadership*. Teachers College Press.
- Lizasoain, L., Tourón, J., & Sobrino, Á. (2015). La evaluación del profesorado español y el impacto del feedback en las prácticas docentes. Análisis de TALIS 2013. *Revista Española de Pedagogía*, año LXXIII (262), 465-482. <https://www.jstor.org/stable/24711323>
- Lochmiller, C. R. (2016). Examining Administrators' Instructional Feedback to High School Math and Science Teachers. *Educational Administration Quarterly*, 52(1), 75-109. <https://doi.org/10.1177/0013161X15616660>
- Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, 86(10), 727-735. <https://doi.org/10.1177/003172170508601004>
- Ministerio de Educación. (2014). *Estándares indicativos de desempeño para los establecimientos educacionales y sus sostenedores*. <https://hdl.handle.net/20.500.12365/587>
- Müller, M., Volante, P., Grau, V., & Preiss, D. D. (2014). Desarrollo de Habilidades de Observación en la Formación de Liderazgo Escolar a Través de Videos de Clases. *Psykhē*, 23(2), 1-12. <http://dx.doi.org/10.7764/psykhe.23.2.713>
- National Policy Board for Educational Administration, NPBEA (2015). Professional standards for educational leaders 2015. Reston, VA: Author.
- OECD (2020), TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals, TALIS, OECD Publishing, Paris, <https://doi.org/10.1787/19cf08df-en>
- Ovando, M. N. (2005). Building instructional leaders' capacity to deliver constructive feedback to teachers. *Journal of Personnel Evaluation in Education*, 18(3), 171-183. <https://doi.org/10.1007/s11092-006-9018-z>

- Pajak, E. (2001). Clinical Supervision in a Standards-Based Environment Opportunities and Challenges. *Journal of Teacher Education*, 52(3), 233-243. <https://doi.org/10.1177/0022487101052003006>
- Pérez-García, Purificación, Bolívar, Antonio, García-Garnica, Marina, & Caracuel, Alfonso. (2018). Adaptación española de la escala de liderazgo pedagógico Vanderbilt Assessment of Leadership in Education (VAL-ED). *Universitas Psychologica*, 17(1), 171-183. <https://dx.doi.org/10.11144/javeriana.upsy17-1.aeel>
- Quiroga, M., & Aravena, F. (2021). La fuga. Estrategia de directores en situaciones de retroalimentación difíciles. *Perfiles educativos*, 43(172), 8-25. <https://doi.org/10.22201/iisue.24486167e.2021.172.59149>
- Robinson, V., Lloyd, C. & Rowe, K. (2008). The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types. *Educational Administration Quarterly*, 44(5), 635-674. <https://doi.org/10.1177/0013161X08321509>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189. <https://doi.org/10.3102/0034654307313795>
- Stronge, J. H. (2006). *Evaluating Teaching: A Guide to Current Thinking and Practice*. (2nd Ed.). Corwin.
- Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning theory perspective. *Educational Research Review*, 9, 1-15. <https://doi.org/10.1016/j.edurev.2012.11.004>
- Tuytens, M., & Devos, G. (2011). *The voice of the teacher: feedback as a hindrance or as an incentive?* [Conference paper] The biennial conference of the International Study Association on Teachers and Teaching. Braga, Portugal
- U.S. Department of Education Office of Innovation and Improvement (2004). *Innovative Pathways to School Leadership*. U.S. Department of Education.
- Volante, P. (2010). *Influencia instruccional de la dirección escolar en los logros académicos* [Tesis Doctoral]. Pontificia Universidad Católica de Chile.