

# LAS RÚBRICAS COMO INSTRUMENTO DE EVALUACIÓN DEL DESEMPEÑO MUSICAL EN UNA UNIVERSIDAD DE MÉXICO. UN ESTUDIO EXPLORATORIO

RUBRICS AS A RESOURCE FOR EVALUATING MUSICAL PERFORMANCE IN A UNIVERSITY OF MEXICO. AN EXPLORATORY STUDY

Raúl W. Capistrán Gracia

---

1.- Doctor en Educación Musical; Universidad Autónoma de Aguascalientes, [raul.capistran@edu.uaa.mx](mailto:raul.capistran@edu.uaa.mx).  
<https://orcid.org/0000-0002-4594-8570>

---

*Recibido: 26 de enero de 2023*

*Aceptado: 08 de abril de 2023*

## Resumen

En este artículo se presentan los resultados de una investigación que tuvo como propósito conocer, desde la perspectiva de los estudiantes de una licenciatura en música en el centro de México: a) el nivel de conocimiento de la rúbrica que se utiliza para evaluar su desempeño musical; b) su grado de utilidad como instrumento de evaluación; c) su grado de utilidad como fuente de retroalimentación y; d) el nivel de objetividad y justeza de las evaluaciones. La investigación fue de corte cuantitativo y de tipo exploratorio. La recogida de datos se realizó a través de una encuesta creada por el propio investigador. En el estudio participaron 84 de los 100 estudiantes matriculados. Los resultados indican que el programa educativo enfatiza una evaluación del aprendizaje y no para el aprendizaje; que existen áreas de oportunidad muy importantes que también podrían estar presentes en programas académicos similares y que esas debilidades deben ser subsanadas lo más pronto posible a fin de promover procesos de enseñanza-aprendizaje más integrales.

**Palabras clave:** educación musical; interpretación musical; desempeño musical; rúbricas de evaluación; evaluación para el aprendizaje

## Abstract

This article presents the results of a study whose purpose was to know from the perspective of students enrolled in a Bachelor of Music degree in central Mexico: a) the level of knowledge of the rubric that is used to evaluate their musical performance; b) its degree of usefulness as an evaluation instrument; c) its degree of usefulness as a source of feedback and d) the level of objectivity and fairness of evaluations. The research method employed was quantitative and exploratory in nature. The data collection was carried out through a survey created by the researcher. The study involved the participation of 84 out of the 100 students enrolled. Results indicate the educational program emphasizes an evaluation of learning and not for learning and show there are very important areas of

opportunity that could also be present in similar academic programs. Therefore, those weaknesses must be corrected as soon as possible in order to promote more holistic teaching-learning processes.

**Key words:** music education; music interpretation; performance quality; evaluation rubrics; evaluation for learning

La evaluación de los aprendizajes es una tarea intencional, sistemática, continua, permanente y reflexiva. Es una acción de gran relevancia que tiene como propósito principal determinar los aprendizajes alcanzados por el estudiante y estimar la calidad de los mismos. A través de la evaluación se obtiene información crucial que permite al maestro identificar sus propias debilidades, valorar las metodologías empleadas, ponderar las actividades de aprendizaje que ha implementado, y optimizar su labor docente. Del mismo modo, proporciona retroalimentación al estudiante acerca de sus fortalezas y áreas de oportunidad, promueve su espíritu de superación y le suministra elementos a partir de los cuales puede mejorar su desempeño académico. Como Stanley, Brooker y Gilbert (2002, p. 46) explican:

El propósito esencial de la evaluación educativa es mejorar el aprendizaje. La evaluación proporciona retroalimentación crucial a estudiantes y maestros acerca del progreso en relación al logro de los objetivos de enseñanza.

Por todo lo anterior, en el ámbito educativo-musical se han tratado de crear e implementar instrumentos de evaluación confiables y válidos, que tomen en cuenta los criterios y parámetros involucrados en la ejecución musical, y permitan llevar a cabo una evaluación más objetiva y justa que refleje el desempeño del estudiante y proporcione información veraz que le posibilite mejorar.

Por cerca de 12 años, al finalizar cada semestre el Departamento de Música de una universidad en el centro de México ha empleado rúbricas analíticas para evaluar el desempeño en la ejecución de instrumentos musicales por parte de los estudiantes de la Licenciatura en Música. Sin embargo, poco se sabe de la eficiencia y efectividad con que se han implementado, mucho se ignora respecto a cualquier logro obtenido a través del uso de este instrumento de evaluación, y menos aún se conoce acerca de la percepción que los estudiantes tienen respecto a la objetividad y justeza de las evaluaciones y a la utilidad de las rúbricas como fuente de retroalimentación para mejorar su desempeño. Con el propósito de llenar ese vacío, el autor llevó a cabo un estudio de tipo cuantitativo y carácter exploratorio cuyos resultados se presentan en este artículo.

### **La evaluación al desempeño musical. Aportaciones teóricas y empíricas**

En un examen de instrumento musical, el sinodal debe tomar en consideración, entre otros aspectos, la afinación, la precisión rítmica, la técnica, el fraseo, la dinámica, la agógica, las articulaciones (legato, staccato, portato, etc.), las respiraciones, e incluso, el desenvolvimiento escénico; debe valorar todos esos elementos al trasluz de ciertos criterios de calidad y finalmente,

emitir un juicio sobre su mérito. En otras palabras, la evaluación del desempeño musical involucra ponderar la calidad de muchos elementos que suceden al mismo tiempo (Saunders & Holahan, 1997). Para dificultar las cosas, se debe añadir que la interpretación musical, como todo arte escénico, involucra aspectos subjetivos que muchas veces dificultan la emisión de valoraciones justas y objetivas. Como explican Chase, Ferguson y Hoey (2014), analizar y evaluar la práctica y el producto artístico puede ser una tarea difícil. Evaluar el trabajo de artistas e intérpretes es un ejercicio que tradicionalmente es de naturaleza subjetiva y siempre se ha basado en el juicio de expertos.

Desgraciadamente, aún en la actualidad las evaluaciones al desempeño musical en muchas instituciones de educación a nivel superior se llevan a cabo de manera global (Navarro, 2013). Así, después de escuchar la interpretación, el sinodal suele dar una calificación numérica que debería reflejar la apreciación que tuvo respecto a la calidad total de la ejecución del estudiante. Como el lector podrá imaginar, los resultados de las evaluaciones algunas veces motivan acaloradas discusiones (Navarro, 2013) y no menos agudas críticas (Chacón, 2012). Por todo lo anterior, desde hace varias décadas se han tratado de crear rúbricas de evaluación que tomen en consideración los diversos factores y parámetros involucrados en la ejecución musical. De ese modo, al impulsar su implementación en contextos educativos, se ha esperado lograr evaluaciones más justas, objetivas y congruentes, que proporcionen retroalimentación de utilidad para que los estudiantes, al igual que los docentes, puedan mejorar su desempeño.

De acuerdo con Díaz-Barriga (2006), estos instrumentos tienen la función de guiar el proceso de evaluación, al establecer niveles de dominio de manera progresiva, que pueden ser representativos del grado de desempeño que un estudiante exhibe al momento de ejecutar una actividad o realizar una producción determinada, de ahí que sean ideales para determinar niveles de competencia. Para Arends (2012), básicamente existen dos tipos de rúbricas: la holística y la analítica. En la primera, se selecciona el descriptor (con su respectivo puntaje) que mejor refleje la ejecución de una actividad o la realización de una producción. Algunos de estos instrumentos pueden incluir niveles de logro muy elementales que ayudan al sinodal para que decida el grado de calidad alcanzado.

Por otro lado, Cano (2015) define la rúbrica analítica como una matriz con propósitos de evaluación que tiene dos ejes; en uno se listan los criterios de ejecución, mientras que en el otro se despliegan los niveles y descriptores de logro. Así, corresponde al juez seleccionar el nivel y descriptor de logro que mejor refleje el desempeño del estudiante en un criterio específico.

Las rúbricas, al igual que cualquier otro instrumento de evaluación, deben cumplir con tres requisitos indispensables: deben ser confiables, es decir, deben arrojar resultados consistentes; deben ser válidas, en otras palabras, deben evaluar los aspectos para los que se supone fueron diseñadas y, finalmente, deben ser justas, o sea, deben ofrecer a todos los estudiantes las mismas posibilidades de obtener buenos resultados y no discriminar a alguien por alguna razón en particular (Arends, 2012).

Un aspecto crucial que debe ser tomado en cuenta al utilizar una rúbrica de evaluación, es que los estudiantes deben estar familiarizados con los criterios y niveles y descriptores de logro antes del examen (Arends, 2012), y de hecho deben entrenar diligentemente para alcanzarlos (Díaz-Barriga, 2006) pues, como explica Cano (2015), estos deben ser conocidos, transparentes y difundidos. Y es que el desconocimiento de estos ha causado que algunos estudiantes perciban la evaluación como un evento arbitrario e incluso riesgoso. Más aún, ese conocimiento permite que cada estudiante se convierta en evaluador de su propio trabajo, lo que constituye una competencia indispensable para promover la metacognición.

Según Moreno-Olivos (2016), las rúbricas no solo representan un instrumento de evaluación, sino que, cuando se implementan de manera informada, seria y profesional, y en el marco de una cultura de evaluación para el aprendizaje, obligan al docente a definir de manera puntual y clara los objetivos del mismo y a establecer criterios precisos para evaluarlos; lo impulsan a conectar esos objetivos con los contenidos y las actividades de la clase; favorecen una comunicación ágil y eficiente de los resultados de aprendizaje que se esperan; facilitan la entrega de retroalimentación precisa y clara; promueven la autorregulación del aprendizaje por parte de los estudiantes y fomentan la metacognición.

En el ámbito de la educación general, Cano (2015) asevera que diversas investigaciones en torno al uso de rúbricas han revelado buenos índices de validez y de confiabilidad, y en general, beneficios importantes derivados de su uso. Sin embargo, esa autora destaca vacíos o áreas de oportunidad importantes en la investigación, como la existencia de pocos estudios que aborden el desarrollo mismo de las rúbricas, y el hecho de que los estudios realizados han dejado fuera aspectos como la validez de contenido y el diseño de los constructos y de los criterios.

En el área de la educación musical también se han llevado a cabo estudios que han permitido dilucidar la validez y confiabilidad de las rúbricas. Así, por ejemplo, Ciorba y Smith (2009) analizaron la efectividad de una rúbrica utilizada para evaluar el desempeño musical en una universidad en el medio oeste de los Estados Unidos de América. Los resultados obtenidos indicaron que había un nivel de acuerdo de moderado a alto entre los jueces y que la rúbrica de evaluación medía eficazmente el rendimiento de los estudiantes en el área de la ejecución musical individual. Por su parte, Latimer, Bergee y Cohen (2010) llevaron a cabo un estudio con la finalidad de determinar la confiabilidad de una rúbrica utilizada en Kansas, USA, para evaluar los grupos musicales que participaban en festivales. Esta investigación también tuvo como propósito conocer las percepciones de los docentes con respecto a su utilidad. Los resultados indicaron que la consistencia interna de la rúbrica era moderadamente alta, al igual que la confiabilidad de las puntuaciones totales, y que la confiabilidad de las dimensiones oscilaba entre moderadamente baja y moderada. Por otro lado, los jueces y los directores de los ensambles coincidieron en que la rúbrica representaba un mejor instrumento para justificar las calificaciones, y proveía descripciones más detalladas de lo que constituía

una ejecución aceptable, en comparación con los recursos de evaluación que habían utilizado anteriormente.

Álvarez-Díaz, Muñiz-Bascón, Soria-Aleman, Veintimilla-Bonet y Fernández-Alonso (2021) realizaron un estudio que involucró el diseño y validación de una rúbrica analítica para evaluar las ejecuciones en el marco de un concurso de música. Un panel de tres expertos creó la rúbrica, la cual consistió en cinco criterios de revisión y tres niveles de puntuación con sus respectivos descriptores. Para validar la rúbrica, dos paneles de jueces la utilizaron para calificar una muestra de grabaciones. Los investigadores examinaron la dimensionalidad, las fuentes de error y la confiabilidad entre evaluadores, así como la consistencia interna de los puntajes, pero no encontraron efectos diferenciales entre los evaluadores; y las diferencias observadas respecto a la consistencia interna de cada juez no fueron significativas. Por el contrario, la rúbrica demostró ser confiable y proporcionar resultados objetivos, tanto en términos de consistencia interna como de nivel de acuerdo entre los evaluadores.

Por supuesto, aun cuando las rúbricas han representado un avance en la calidad de las evaluaciones, también ofrecen algunas desventajas. La principal de ellas radica en que el sinodal o juez se limita a manifestar qué tan de acuerdo está con el descriptor. Como consecuencia, la evaluación resultante no explica integralmente por qué una ejecución es excepcional o deficiente. Además, no todo se puede comprimir en criterios, descriptores y niveles de logro, por lo que es posible que algunos factores queden fuera. Más aún, en muchos casos, existen otras alternativas para efectuar una buena evaluación. Finalmente, una buena rúbrica es difícil de diseñar, requiere del involucramiento de expertos y debería pasar por un proceso de validación que dé cuenta de la suficiencia, claridad, coherencia y relevancia de los criterios y descriptores, así como de su consistencia interna y de su confiabilidad, antes de ser implementada. A pesar de todas esas desventajas, diversos investigadores y académicos concuerdan en que los beneficios superan por mucho las debilidades que pudieran tener.

### **Preguntas que orientaron la investigación**

Desde la perspectiva de los estudiantes:

1. ¿Cuál es su nivel de conocimiento con respecto a la rúbrica que se utiliza para evaluar su desempeño musical?
2. ¿Cuál es el grado de utilidad de la rúbrica como instrumento de evaluación?
3. ¿Cuál es el grado de utilidad de la rúbrica como fuente de retroalimentación?
4. ¿Qué tan objetivas y justas son las evaluaciones?

### **Metodología**

La investigación fue de tipo exploratorio y de corte cuantitativo. La recogida de datos se realizó a través de un cuestionario creado por el propio

investigador, el cual fue integrado por preguntas generales y 12 ítems con escalas tipo Likert organizados en cuatro dimensiones: a) “conocimiento de la rúbrica” (3 ítems); b) “utilidad de la rúbrica para proporcionar retroalimentación” (2 ítems); c) “involucramiento de los docentes y los estudiantes en procesos de reflexión después de la evaluación a partir de la rúbrica y la grabación” (5 ítems); d) “congruencia y justeza de las evaluaciones” (2 ítems).

Cinco académicos con nivel de doctorado y experiencia en el ámbito de la investigación en educación musical participaron en el proceso de validación por juicio de expertos. Los jueces respondieron un cuestionario propuesto por Escobar-Pérez y Cuervo-Martínez (2008) diseñado para determinar la suficiencia, claridad, coherencia y relevancia de cada uno de los ítems. Para determinar el índice de validez de contenido y la concordancia entre los jueces se utilizó el estadístico de Hernández Nieto, el cual arrojó un resultado de 0.915 (excelente). Posteriormente, se llevó a cabo una prueba piloto a través de la cual se identificaron algunas pequeñas problemáticas en la claridad que fueron corregidas.

En el estudio participaron 84 de los 100 estudiantes matriculados en el programa educativo: 16 de 2º semestre, 20 de 4º semestre, 15 de 6º semestre, 14 de 8º semestre, y 19 de 10º semestre, lo que implicó un 84% de la totalidad. La muestra estadística tomada fue por conveniencia. 49 participantes fueron hombres y 35 fueron mujeres, con edades que oscilaron entre los 18 y los 31 años. La media fue de  $M = 22.21$  años y la desviación típica de  $DT = 3.04$  años.

Para establecer el grado de confiabilidad del instrumento, se llevó a cabo un estadístico  $\alpha$  de Cronbach, el cual dio un resultado de  $\alpha = 0.849$  (confiabilidad alta). Por otro lado, la prueba de Kolmogorov-Smirnov determinó que los datos recabados no cumplían con el supuesto de normalidad, por lo que los cálculos estadísticos inferenciales se realizaron a través de pruebas de tipo no paramétrico.

## Procedimiento

El cuestionario fue enviado a los estudiantes a través de *Google Forms* y estuvo disponible por un periodo de cuatro semanas. Dos correos electrónicos adicionales fueron enviados para invitar a los estudiantes a responder la encuesta. Los datos fueron procesados con el programa *SPSS 2.0*. Así, se explicó que los resultados serían publicados como cifras estadísticas y que en todo momento se garantizaría la confidencialidad de los datos personales.

## Resultados

Como se muestra en la Figura 1, en la dimensión de “conocimiento de la rúbrica”, los resultados indican que a un 19% de los estudiantes “nunca” se les ha explicado para qué sirve la rúbrica de evaluación, mientras que a un 17.9% solo “rara vez”. Por su parte, 22.6% de los estudiantes declararon que “algunas veces” su maestro les había proporcionado información al respecto, mientras que 17.9% había recibido la información “frecuentemente” y 22.6% “siempre”. La media de este ítem fue  $M = 2.07$  y la desviación típica fue  $DT = 1.42$ .

**Figura 1.**

*Grado en que los maestros explican a sus estudiantes para que sirve la rúbrica*



Del mismo modo, 2.4% comprenden “nada” los criterios de desempeño en ejecución musical que enlista la rúbrica, un porcentaje similar aseguraron hacerlo “muy poco”, 26.2% afirmaron comprenderlos “algo”, mientras que 13.1% de los estudiantes no conocen este instrumento. Por otro lado, 40.5% afirmaron comprenderlos “bastante” y 15.5% “mucho” (ver Tabla 1).

Como se muestra en la Tabla 1, en lo relativo al nivel de comprensión de los niveles y descriptores de logro, es decir, de los parámetros de evaluación que se establecen en la rúbrica, hubo cambios nominales importantes en la mayoría de las categorías. Por supuesto, el porcentaje de estudiantes que manifestaron no conocer la rúbrica fue el mismo, sin embargo, 3.6% afirmaron comprender “nada” los niveles y descriptores de logro, 11.9% declararon comprenderlos “muy poco” y 35.7% “algo”. Por otro lado, 25% manifestaron comprenderlos “bastante” y 10.7% aseguraron hacerlo “mucho”. En otras palabras, mientras que las categorías “nada”, “muy poco” y “algo” tuvieron un incremento, las categorías “bastante” y “mucho” mostraron un descenso, lo que refleja menor grado de comprensión de los niveles y descriptores de logro en comparación con el grado de comprensión de los criterios.

**Tabla 1.**

*Comprensión de los criterios y niveles y descriptores de logro de la rúbrica*

Categorías	Comprensión de los criterios		Comprensión de los niveles de logro y descriptores	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
No conozco la rúbrica	11	13.1	11	13.1
Nada	2	2.4	3	3.6
Muy poco	2	2.4	10	11.9
Algo	22	26.2	30	35.7
Bastante	34	40.5	21	25.0
Mucho	13	15.5	9	10.7

La dimensión de “conocimiento de la rúbrica” obtuvo una media de  $M = 2.19$  ( $DT = 1.13$ ) (ver Tabla 2). A fin de comprobar si existían diferencias estadísticas significativas entre las medias en función del semestre cursado por

los estudiantes, se realizó una prueba de Kruskal Wallis. Los resultados indicaron que no existen tales diferencias ( $p > 0.05$ ).

**Tabla 2.**

*Medias correspondientes al conocimiento general de la rúbrica*

Semestre que cursa	Media	N	DT
2º semestre	2.00	16	1.491
4º semestre	2.32	20	0.693
6º Semestre	1.50	15	1.118
8º Semestre	2.57	14	1.016
10º semestre	2.47	19	1.111
Total	2.19	84	1.137

Como puede verse en la Figura 2, En la dimensión “utilidad de la rúbrica para proporcionar retroalimentación”, los resultados indicaron que un 15.5% de los estudiantes consideraba que los sinodales “nunca” utilizaban los criterios, niveles y descriptores de logro de la rúbrica para proporcionar retroalimentación; el 8.3% manifestaron que solo lo hacían “rara vez” y el 20.2% declararon que “algunas veces”. Por otro lado, el 22.6% indicaron que los sinodales “frecuentemente” se referían a ellos y solo el 11.9% aseguraron que lo hacían “siempre”. Finalmente, un 21.4% afirmó que “nunca había sido llamado para recibir retroalimentación”

**Figura 2.**

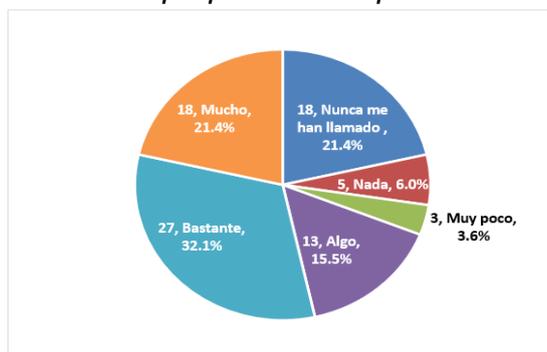
*Uso de criterios, niveles y descriptores de logro al proporcionar retroalimentación*



En relación a la utilidad de la retroalimentación, un 6% declaró que le había servido “nada”, el 3.6% manifestó que le había servido “muy poco” y el 15.5% “algo”. Por otro lado, el 32.1% señaló que la retroalimentación le había servido “bastante” y el 21.4% indicó que le había servido “mucho”. Por supuesto, el mismo 21.4% afirmó que “nunca había sido llamado para recibir retroalimentación” (ver Figura 3).

**Figura 3.**

*Utilidad de la retroalimentación proporcionada por los sinodales*



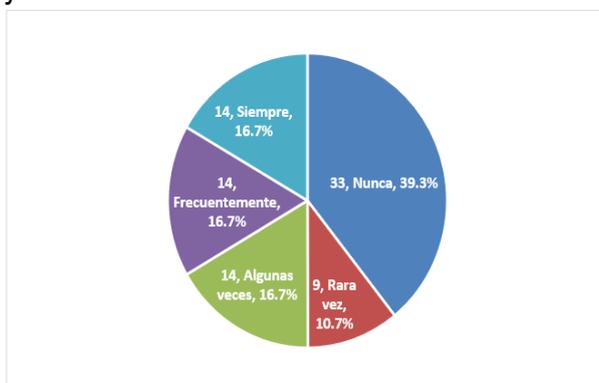
Debido a que los datos estadísticos no tienen una distribución normal, el autor realizó un estadístico Rho de Spearman para determinar el grado de dependencia entre las variables “utilización de los criterios, niveles y descriptores de logro al dar retroalimentación” y “utilidad de la retroalimentación”, el cual dio un resultado de  $Sp = 0.674$  con un valor de  $p = 0.000$ , lo que indica que las variables son dependientes, que existe una correlación positiva moderada fuerte, y que, a mayor utilización de los criterios, niveles y descriptores de logro por parte de los sinodales al dar retroalimentación, existiría un mayor nivel de utilidad de esta para los estudiantes.

En lo que respecta a la dimensión “involucramiento de los docentes y los estudiantes en procesos de reflexión después de la evaluación a partir de la rúbrica y la grabación, para mejorar la práctica y el desempeño musical”, un abrumador 86.9% afirmó que “nunca” había recibido una copia de ese documento. Del resto de los estudiantes, el 3.6% afirmaron que contar con la rúbrica les había sido “algo” de utilidad y el 9.5% aseguraron que les había servido “bastante”.

Como se muestra en la Figura 4, los resultados también indicaron que el 39.3% de los estudiantes “nunca” habían revisado los resultados que contiene la rúbrica de evaluación junto con sus maestros, ni habían reflexionado de manera conjunta sobre ellos para mejorar su práctica y su desempeño musical. Del mismo modo, un 10.7% afirmaron que “rara vez” lo habían hecho y un 16.7% señalaron que “algunas veces”. Por otro lado, el 16.7% afirmaron involucrarse “frecuentemente” en este tipo de reflexión, mientras que un porcentaje igual aseveraron hacerlo “siempre”.

**Figura 4.**

*Involucramiento de docentes y estudiantes en procesos de reflexión a partir de los resultados reflejados en la rúbrica*



Este programa educativo tiene la norma de videgrabar los exámenes de instrumento musical. Así, se conserva una evidencia que permite llevar a cabo una revisión de examen en caso necesario. Este material puede ser solicitado por el docente como recurso para la autoevaluación y la reflexión con el estudiante. Por supuesto, el estudiante y el docente también tienen el derecho de efectuar una grabación personal durante el examen con ese mismo propósito. De acuerdo con los resultados, un alto porcentaje de los estudiantes (78.6%) afirmó que sus maestros “nunca” han solicitado la grabación escolar, ni ellos mismos (53.6%) han procurado grabar su desempeño durante el examen. Como se muestra en la Tabla 3, el resto de las categorías obtuvo porcentajes muy bajos.

**Tabla 3.**

*Cuadro comparativo entre solicitud de grabación escolar y grabación personal*

	Solicitud de grabación escolar		Grabación personal	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Nunca	66	78.6%	45	53.6%
Rara vez	4	4.8%	14	16.7%
Algunas veces	5	6.0%	11	13.1%
Frecuentemente	5	6.0%	6	7.1%
Siempre	4	4.8%	8	9.5%

**Fuente:** elaboración propia.

Al agrupar las variables de la dimensión “involucramiento de los docentes y los estudiantes en procesos de reflexión después de la evaluación a partir de la rúbrica y la grabación”, se obtuvo una media de  $M = 1.03$  ( $DT = 1.13$ ) lo que indicaría que estos lo hacen apenas por encima de la categoría “rara vez” (ver Tabla 4). A fin de comprobar si existían diferencias estadísticas significativas entre las medias de los cinco semestres cursados por los

estudiantes, se realizó una prueba de Kruskal Wallis a través de la cual se determinó que no había tales diferencias ( $p > 0.05$ ).

**Tabla 4.**

*Medias correspondientes a la dimensión “involucramiento de los docentes y los estudiantes en procesos de reflexión después de la evaluación a partir de la rúbrica y la grabación”*

Semestre	Media	Frecuencia	DT
2º semestre	1.43	16	1.227
4º semestre	0.86	20	1.299
6º Semestre	1.00	15	1.436
8º Semestre	0.97	14	0.831
10º semestre	0.96	19	0.792
Total	1.03	84	1.136

**Fuente:** elaboración propia.

Finalmente, como puede verse en la Tabla 5, en la dimensión “congruencia y justeza de las evaluaciones”, el mismo 39.3% que afirmaron no involucrarse con su maestro en procesos de reflexión posteriores a la evaluación, aseguraron no haber visto la rúbrica después del examen, por lo que no podían saber que tan congruente había sido la selección de los niveles y descriptores de logro con la percepción que ellos tenían de su ejecución. El 3.6% señalaron que había “poca” congruencia y el 20.2% indicaron que había “algo” de congruencia. Por otro lado, el 29.8% aseguraron que los niveles de logro y descriptores seleccionados eran “bastante” congruentes con la percepción que ellos tenían de su ejecución, mientras que el 7.1% aseguraron que eran “muy” congruentes.

En cuanto a la justeza de las evaluaciones, un 56% de los estudiantes consideró que eran “bastante” justas y un 16.7% que eran “muy” justas. Por otro lado, el 20.2% de los estudiantes señalaron que eran “algo” justas, el 3.6% afirmaron que “muy poco” y un porcentaje similar aseveraron que “nada”.

**Tabla 5.**

*Congruencia y justeza de las evaluaciones*

	Congruencia		Justeza	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Nunca he visto la rúbrica de evaluación	33	39.3%	No aplica	No aplica
Nada	0	0%	3	3.6%
Muy poco	3	3.6%	3	3.6%
Algo	17	20.2%	17	20.2%
Bastante	25	29.8%	47	56.0%
Mucho	6	7.1%	14	16.7%

## Discusión

El análisis de los resultados indica que existe un área de oportunidad importante en lo que respecta al conocimiento de la rúbrica por parte de los estudiantes, ya que, como se vio en la revisión bibliográfica, los criterios de evaluación, así como los niveles y descriptores de logro de este instrumento deben ser debidamente difundidos por los docentes (Cano, 2015) y deben ser conocidos de manera clara y precisa por todos los estudiantes antes del examen (Arends, 2012), de tal manera que estos puedan practicar sus ejercicios, estudios y obras de manera diligente y tener un buen desempeño (Díaz-Barriga, 2006).

Por lo que respecta a la utilidad de este instrumento para proporcionar retroalimentación, el hecho de que alrededor de una quinta parte de los estudiantes afirmara que nunca ha sido llamado para recibir retroalimentación, es un factor relevante, ya que muchas veces los estudiantes después de su examen abandonan la sala con enormes dudas respecto a la calidad de su desempeño, dudas válidas que requerirían de retroalimentación efectiva para ser disueltas y, sobre todo, para fundamentar las acciones que el estudiante debe tomar para mejorar, pues la retroalimentación correctiva genera resultados muy positivos, sobre todo, en los estudiantes que tienen un menor nivel de desempeño (Day, Blankenstein, Westenberg & Admiraal, 2018). Este aspecto podría subsanarse si los estudiantes recibieran la rúbrica con la retroalimentación por escrito, pero la mayoría jamás ha recibido una copia.

Uno de los hallazgos más preocupantes de este estudio está representado por la falta de conexión entre la terminología establecida en la rúbrica (criterios de evaluación y niveles y descriptores de logro) y el vocabulario utilizado por los maestros para proporcionar retroalimentación, ya que despierta dudas sobre la validez de las evaluaciones. Y es que, es inevitable preguntarse: ¿Realmente la rúbrica evalúa lo que se supone que debería de evaluar? Si es así ¿Por qué la retroalimentación no se basa en los criterios, niveles de logro y descriptores? Si la rúbrica no es adecuada, ¿Por qué los directivos y docentes no se han involucrado en un proceso de mejoramiento de la misma? Así, es crucial que el sinodal no solo ofrezca retroalimentación, sino que también debe esforzarse en utilizar la terminología establecida en la rúbrica para que ese conocimiento sea todavía más útil.

En cuanto al involucramiento de los docentes y los estudiantes en procesos de reflexión después de la evaluación a partir de la rúbrica y la grabación, los resultados indican que existe bastante espacio para mejorar. Y es que, por ejemplo, la revisión de una videograbación afina el criterio de autoevaluación del estudiante (Daniel, 2001), promueve la reflexión sobre el proceso de aprendizaje (Boucher, Dube & Creech, 2017) y lo ayuda a escoger la estrategia más pertinente para superar los desafíos de las obras (Hallam, 2001; McPherson & Zimmerman, 2002; Nielsen, 1999). Sin embargo, aunque los resultados indican que los estudiantes prefieren gestionar sus propias grabaciones del examen, en lugar de solicitar una copia de la grabación institucional, es evidente que se hace muy poco. Lo anterior es alarmante, ya que, al prescindir de un recurso tan valioso, como lo es la videograbación de la

ejecución musical durante el examen, el autor podría aventurar que el criterio o filosofía predominante tanto del estudiante como del docente, parecería ser el de la evaluación del aprendizaje y no para el aprendizaje.

Finalmente, un porcentaje importante de los estudiantes no pudo afirmar si los niveles y descriptores de logro que son seleccionados en la rúbrica por los jueces son congruentes o no con la percepción que ellos tienen de su ejecución, pues nunca han visto la rúbrica de evaluación después del examen. Este resultado hace que la objetividad de la evaluación sea bastante cuestionable. En primer lugar, porque todos los estudiantes deberían conocer en detalle la rúbrica antes del examen, y todos deberían poder revisarla después y; en segundo lugar, porque se esperaría un alto nivel de congruencia entre las percepciones que tienen los estudiantes acerca de su propia ejecución y los niveles y descriptores de logro seleccionados por los sinodales.

Lo más sorprendente es que, a pesar de lo anterior, los resultados reflejan que los estudiantes consideran que las calificaciones son justas. Estos últimos resultados parecen reforzar más la idea de que los alumnos, al igual que los docentes, se inclinan hacia una filosofía de la “evaluación del aprendizaje” más que “para el aprendizaje”. En otras palabras, los estudiantes parecen satisfechos con las evaluaciones y, sobre todo, parece que las evaluaciones representan un simple escollo a vencer para continuar con la carrera. Más aún, los resultados sugieren que una parte importante de la información derivada de la evaluación es poco utilizada, y que las acciones que se deberían emprender, tales como explicar el contenido y funcionamiento de las rúbricas, entregar una copia de las mismas al final del examen, involucrarse en procesos de análisis crítico y reflexivo y diseñar o seleccionar estrategias de autorregulación y práctica efectiva, son dejadas de lado en muchas ocasiones.

### **Alcances y limitaciones**

Aunque se trata de un tema ampliamente estudiado desde hace varias décadas, los resultados que se han obtenido a través de esta investigación, las observaciones vertidas en diversos coloquios y congresos de educación musical y la experiencia, indican que es crucial seguir abordándolo. El autor recomienda a otros investigadores realizar estudios de este tipo en los que se incluya también las percepciones de los docentes, y, por supuesto, que el estudio sea replicado en otros contextos educativos a nivel superior.

### **Conclusiones**

Los resultados de este estudio parecen indicar que los estudiantes y docentes se inclinan hacia una cultura de la evaluación del aprendizaje. En ese sentido, desde la perspectiva del investigador, es crucial impulsar y promover un cambio hacia la evaluación para el aprendizaje. Evaluar no es solo un medio para asignar una calificación y acreditar o no una asignatura del semestre. Las evaluaciones al desempeño musical en una institución de nivel superior deben representar una oportunidad invaluable para que el estudiante demuestre su capacidad interpretativa, en un contexto muy cercano a la realidad que vivirá

como futuro profesional de la música. Así, debe constituirse en una vivencia que, si bien puede ser estresante, debe ser predominantemente constructiva. Una experiencia en la que pueda tocar frente a un grupo de expertos que harán su mejor esfuerzo para determinar la calidad de su ejecución, identificar las fortalezas, detectar las áreas de oportunidad y, en lo posible, sugerir técnicas o estrategias para mejorar. En otras palabras, el examen debe representar una evaluación para el aprendizaje, un peldaño del andamiaje sobre el cual el estudiante puede continuar construyendo su carrera.

Más aún, no se debe olvidar que el ejecutante es un ser humano quien, al mostrar sus logros musicales, también está exhibiendo sus aspiraciones y exponiendo su dimensión afectiva. No es de sorprender que Prieto Alberola (2002, p. 329) afirmara: “Esto representa una gran responsabilidad para los jurados y tribunales evaluadores, ya que están en juego, a veces, las ilusiones personales y otras su futuro profesional, con las consiguientes consecuencias económicas, sociales y afectivas”. En ese sentido, las rúbricas representan un recurso valioso para promover evaluaciones justas, válidas, confiables, que reflejan procesos de valoración que no solo arrojan una calificación representada por un número, sino que manifiestan el compromiso serio, responsable, crítico y reflexivo del evaluador.

Por tanto, es vital tener consciencia de las áreas de oportunidad y fortalezas en la ejecución de los estudiantes y poder articularlas de manera objetiva, coherente y concisa, para beneficio del propio estudiante y de su maestro, algo que desgraciadamente muchas veces no ocurre (Prieto-Alberola, 2001; Moreno-Olivos, 2016). En ese sentido Winter (1993) explica que, del mismo modo que se enseña a los estudiantes a entender y dominar los diferentes elementos de la música para luego integrarlos en una forma artística y coherente, los maestros también deberían ser capaces de evaluar esos elementos en la ejecución musical y emitir un juicio lo más objetivo y justo posible.

Para que las rúbricas puedan ser utilizadas de manera idoneidad, se requiere de docentes capacitados, informados y con un alto nivel humanista y ético. Se necesitan profesionales que consideren los exámenes como una experiencia integradora en la que los estudiantes y los docentes vean en cada detalle la posibilidad para aprender y mejorar. En pocas palabras, se necesitan maestros que crean en evaluaciones para el aprendizaje. Las oportunidades abundan. Por ejemplo, día con día se evalúa a los estudiantes o se les pide que se autoevalúen después de una lección. ¿Por qué no utilizar la rúbrica de evaluación del examen final a lo largo del semestre? De esa manera, se lograría una mayor congruencia entre lo que se hace en clase y lo que se evalúa; el estudiante conocería en detalle los criterios, niveles de logro y descriptores; afinaría su capacidad para autoevaluarse y mejoraría sus procesos metacognitivos. Después de todo, la evaluación también debe ser continua y formativa. También se podría diseñar y establecer un protocolo de examen, de tal manera que la cultura de la evaluación para el aprendizaje pueda establecerse y consolidarse. Si bien, el estudio indica áreas de oportunidad muy importantes, desde la perspectiva del investigador estas no

son difíciles de subsanar. Aun cuando implican un gran cambio de mentalidad, los beneficios lo valen.

Para finalizar, el autor desea advertir que si bien, la temática de las rúbricas como instrumento de evaluación ha sido ampliamente abordada desde la década de los 70, las aportaciones empíricas y teóricas podrían no haber ejercido influencia en los mecanismos de evaluación de programas que siguen prácticas arraigadas en modelos educativo-musicales conservatorianos del siglo XIX (Navarro, 2013; Carbajal, 2016, 2017). En ese sentido, los resultados de este artículo podrían representar la “punta del *iceberg*” de una problemática que podría estar presente en más programas académico-musicales a nivel superior en México, como lo han sugerido los comentarios sobre el tema que se han hecho en los últimos años en coloquios y congresos de educación musical llevados a cabo en este país. Por lo anterior, se vuelve crucial continuar estudiando este tema, pues los resultados de investigaciones futuras pueden revelar de manera fehaciente las áreas de oportunidad que se deben atender e impulsar acciones para subsanarlas.

## Referencias

- Álvarez-Díaz, M., Muñoz-Bascón, L. M., Soria-Aleman, A., Veintimilla-Bonet, A. & Fernández-Alonso, R. (2021). On the design and validation of a rubric for the evaluation of performance in a musical contest. *International Journal of Music Education*, 39(1), 66-79.  
<https://doi.org/10.1177/0255761420936443>
- Arends, R. I. (2012). *Learning to teach*. McGraw Hill.
- Boucher, M., Dube, F. & Creech, A. (2017). The effect of video feedback on the self-assessment of a music performance by college level classical guitarists. En G. Hughes (Ed.), *Ipsative assessment and learning gains: Case studies from international practitioners* (pp. 197-220). Palgrave Macmillan.
- Cano, E. (2015). Las rúbricas como instrumento de evaluación de competencias en educación superior: ¿uso o abuso? *Profesorado. Revista de Currículum y Formación de Profesorado*, 19(2), 265-280.  
<https://www.redalyc.org/pdf/567/56741181017.pdf>
- Chacón, L. A. (2012). ¿Qué significa “evaluar” en música? *Revista Electrónica Complutense de Investigación en Educación Musical*, 9, 1-25.  
[http://dx.doi.org/10.5209/rev\\_REC1.2012.v9.42805](http://dx.doi.org/10.5209/rev_REC1.2012.v9.42805)
- Carbajal-Vaca, I. S. (2016). Del modelo de conservatorio al modelo universitario: la experiencia de transición en el departamento de música de la Universidad de Guadalajara. En A. Barrera & D. Pérez (compiladores), *Modelos educativos alternativos*, Tomo 1 (pp. 40-52). Universidad Autónoma de Nayarit.
- Carbajal-Vaca, I. S. (2017). Educación musical superior: el desarrollo de competencias profesionales en músicos universitarios. En *Memoria Electrónica del Congreso Nacional de Investigación Educativa COMIE* (pp. 1-12). COMIE.

- Chase, D. M., Ferguson, J. L. & Hoey, J. J. (2014). *Assessment in Creative Disciplines; Quantifying the Aesthetic*. Champaign, IL: Common Ground Publishing.
- Ciorba, C. R. & Smith, N. Y. (2009). Measurement of Instrumental and Vocal Undergraduate Performance Juries Using a Multidimensional Assessment Rubric. *Journal of Research in Music Education*, 57(1), 5–15. <https://doi.org/10.1177/0022429409333405>
- Daniel, R. (2001). Self-assessment in performance. *British Journal of Music Education*, 18(3), 215-226. [https://doi: 10.1017/S0265051701000316](https://doi:10.1017/S0265051701000316)
- Day, I. N. Z., Blankenstein, F. M., Westenberg, M. & Admiraal, W. (2017). A review of the characteristics of intermediate assessment and their relationship with student grades. *Assessment & Evaluation in Higher Education*, 43(6), 908-929. <https://doi:10.1080/02602938.2017.1417974>
- Díaz-Barriga, F. (2006). *Enseñanza situada: Vínculo entre la escuela y la vida*. McGraw-Hill Interamericana
- Escobar-Pérez, J. & Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en Medición*, 6, 27–36.  
[https://www.researchgate.net/publication/302438451\\_Validez\\_de\\_contenido\\_y\\_juicio\\_de\\_expertos\\_Una\\_aproximacion\\_a\\_su\\_utilizacion](https://www.researchgate.net/publication/302438451_Validez_de_contenido_y_juicio_de_expertos_Una_aproximacion_a_su_utilizacion)
- Hallam, S. (2001). The development of meta-cognition in musicians: Implications for education. *The British Journal of Music Education*, 18(1), 27-39. <https://doi:10.1017/S0265051701000122>
- Latimer, M. E., Bergee, M. J., & Cohen, M. L. (2010). Reliability and perceived pedagogical utility of a weighted music performance assessment rubric. *Journal of Research in Music Education*, 58(2), 168–183. <https://doi.org/10.1177/0022429410369836>
- McPherson, G. E. & Zimmerman, B. J. (2002). Self-regulation of musical learning: A social cognitive perspective. En R. Colwell y C. Richardson (Eds.). *The new handbook of research on music teaching and learning: a project of the Music Educators National Conference* (pp. 327-347). Oxford University Press.
- Moreno-Olivos, T. (2016). *Evaluación del aprendizaje y para el aprendizaje: reinventar la evaluación en el aula*. Universidad Autónoma Metropolitana.
- Navarro, J. (2013). La evaluación musical en una escuela universitaria mexicana: opiniones del alumnado. *Revista Electrónica de LEEME*, 32, 19-52. <https://ojs.uv.es/index.php/LEEME/article/view/9850/9271>
- Nielsen, S. (1999). Learning strategies in instrumental music practice. *British Journal of Music Education*, 16(3), 275-291. <http://doi:10.1017/S0265051799000364>
- Prieto-Alberola, R. (2001). La evaluación de las actividades musicales. *Contextos Educativos*, 4, 329-340. <https://publicaciones.unirioja.es/ojs/index.php/contextos/article/view/500>
- Saunders, T. C. y Holahan, J. M. (1997). Criteria-specific rating scales in the evaluation of high school instrumental performance. *Journal of Research in Music Education*, 45(2), 259-272. [https://doi: 10.2307/3345585](https://doi:10.2307/3345585).

- Stanley, M., Brooker, R. y Gilbert, R. (2002). Examiner perceptions of using criteria in music performance assessment. *Research Studies in Music Education*, 18(1), 46-56. [https://doi: 10.1177/1321103X020180010601](https://doi.org/10.1177/1321103X020180010601).
- Winter, N. (1993). Music Performance assessment: A study of the effects of training and experience on the criteria used by music examiners. *International Journal of Music Education*, 22(1), 34-39. <https://doi.org/10.1177/025576149302200106>