

RESUMEN

El presente trabajo analiza una estrategia de evaluación a través de rúbricas en la materia de instrumento principal de un programa de estudios superiores en música, a partir de los testimonios de docentes y estudiantes. El estudio corrobora y complementa los hallazgos de la literatura, enfatizando la utilidad pedagógica de este tipo de estrategias, no sólo para la estructuración de la retroalimentación, sino para planeación de la enseñanza; sugiere factores que podrían representar sesgos para la evaluación; y propone avenidas de investigación que exploren la relación entre la percepción de la formación inicial del docente y sus prácticas de evaluación actuales.

Palabras clave: evaluación, rúbricas, ejecución musical, instrumento principal

ABSTRACT

This paper analyses a principal study performance assessment case based on rubrics, in a tertiary education programme, by using staff and students testimonies. The study confirms and complements the literature, emphasizing the pedagogical use of the strategy, not only for feedback provision but also for teaching planning. It also suggests factors, which might bias the assessment. Finally, it points out further research to explore the relationship between teachers' perception of their initial professional training and their current assessment practices.

Keywords: assessment, rubrics, music performance, principal study.

USO DE RÚBRICAS EN LA EVALUACIÓN DE INSTRUMENTO PRINCIPAL ESTUDIO DE CASO EN UN PROGRAMA PROFESIONAL DE ESTUDIOS MUSICALES

*Juan Pablo Correa Ortega**
Universidad Autónoma de Aguascalientes
México

INTRODUCCIÓN Y ANTECEDENTES

Si enseñamos a nuestros estudiantes a entender y dominar “los diferentes elementos de la música, para luego integrarlos en una forma artística coherente, [los maestros] también deberíamos evaluar estos elementos en la ejecución musical [de nuestros estudiantes] y emitir un juicio integrado.”¹

El presente estudio pretende resaltar tres rasgos fundamentales de esta compleja tarea. El primero consiste en su insalvable subjetividad, debido a que la evaluación no es más que la medición de las impresiones del docente generadas a partir del estímulo de la ejecución (Radocy, 1986). El segundo es nuestro deber de asignar calificaciones que informen y acrediten los aprendizajes alcanzados por los alumnos, y les permitan entender su desempeño y lo que requieren para mejorarlo.² Y el tercero, nos remite a los principales propósitos de la evaluación: “(1) recolectar evidencias para informar las decisiones de enseñanza; y, (2) motivar a los estudiantes a continuar aprendiendo”.³

* Correo electrónico: juanpablo_correa@hotmail.com. Artículo recibido el 21 - 10 - 2013 y aprobado por el Comité Editorial el 24 - 10 - 2013.

1 Winter, N. (1993). “Music Performance Assessment: A Study of the Effects of Training and Experience on the Criteria Music Examiners”. *International Journal of Music Education*, p. 37

2 Butler, S., & McMunn, N. (2006). **A teacher’s guide to classroom assessment. Understanding and using assessment to improve student learning**. San Francisco, CA.: Jossey-Bass Teacher.

3 Stiggins, R. (2008). **Assessment Manifesto. A call for the development of balanced assessment systems**. Portland, OR.: ETS.

Este triángulo de factores enfatiza la necesidad de diseñar estrategias que mejoren la validez y confiabilidad de nuestras evaluaciones. La estrategia de evaluación por rúbricas presentada en este estudio, responde a esta necesidad en tanto delimita elementos observables y medibles de la ejecución musical, y define sus niveles de desempeño. Esto la convierte en una herramienta que propicia la estructuración de la retroalimentación del desempeño de los estudiantes, y la planeación de metas y métodos para su mejora.

La estrategia de evaluación fue implementada en las cuatro academias que administran la educación en instrumento principal en un programa de educación profesional en música: piano, guitarra, cuerdas y alientos. La metodología de trabajo para el diseño de las rúbricas consistió en sesiones de reflexión y diálogo al interior de cada academia. Los docentes no recibieron capacitación técnica.

Después de dos años de implementación, la presente investigación plantea conocer las percepciones de docentes y alumnos sobre la estrategia, identificar su impacto en los procesos de enseñanza y aprendizaje de instrumento principal, reconocer variaciones en la implementación y sus causas potenciales, y contribuir a la literatura.

REVISIÓN DE LITERATURA

Las experiencias reportadas en la literatura, parten del supuesto de que un mayor grado de validez y confiabilidad puede ser logrado al establecer criterios claros que definan niveles de ejecución musical, los cuales, aunados a escalas de medición, puedan cuantificar el nivel de desempeño de un estudiante.⁴

El uso de escalas estandarizadas sobre diferentes aspectos de la ejecución musical, ligadas a escalas de medición tipo Likert se reportan en: ej., Abeles⁵ y Bergee⁶. Saunders & Holahan⁷ argumentaron que éstas no describen el nivel de desempeño, disminuyendo su validez y eficacia pedagógica. Así, proponen escalas de criterios específicos. Finalmente, ventajas análogas a las de estas

4 Stanley, M., Brooker, R., & Gilbert, R. (2002). "Examiner Perceptions of Using Criteria in Music Performance Assessment". *Research Studies in Music Education*, p.46-56

5 Abeles, H. (1973). "Development and Validation of a Clarinet Performance Adjudication Scale". *Journal of Research in Music Education*.

6 Bergee, M. (1993). "A Comparison of Faculty, Peer, and Self-Evaluation of Applied Brass Jury". *Journal of Research in Music Education*

7 Saunders, C., & Holahan, J. (1997). "Criteria-specific rating scales in the evaluation of high school instrumental performance". *Journal of Research in Music Education*.

últimas fueron descritas para las rúbricas; ej. Ciorba & Smith⁸ y Latimer.⁹

El objetivo de esta revisión de literatura no es, sin embargo, hacer un análisis comparativo de estas estrategias; sino enfocarse en los factores que puedan influir su uso, y en el impacto de estas prácticas en los procesos de enseñanza y aprendizaje. Es por esto que el modelo de McPherson & Thompson (1998) resulta pertinente. En éste, se plantea la interacción de factores involucrados en el proceso de evaluación de la ejecución musical. Factores contextuales, como: el propósito de la evaluación o el tipo de instrumento; factores musicales, como el tipo de repertorio o el desempeño de instrumentistas acompañantes; y, factores no musicales, susceptibles de convertirse en sesgo, como el orden de intervención de los intérpretes o su aspecto físico; se corresponden directa o indirectamente con los hallazgos de la investigación. Así mismo, algunas de las características del evaluador señaladas por McPherson juegan un papel importante en el presente estudio: la experiencia –particularmente la asociada a su etapa estudiantil y laboral-, la familiaridad con el intérprete (alumno), y la capacitación en el uso de instrumentos de evaluación.

La percepción del evaluador con respecto a la pertinencia de instrumentos basados en escalas o criterios es también mencionado en el modelo. La disyuntiva se centra en cómo alcanzar la validez y confiabilidad sin fracturar la integralidad del fenómeno de la ejecución musical. Ej., mientras Mills¹⁰ argumenta en favor de una evaluación holística que represente la realidad compleja de la ejecución musical, otros como Saunders & Holahan¹¹ argumentan a favor de la validez y confiabilidad del uso de criterios específicos.

Otro elemento específico, que no ha sido ampliamente comentado, es el grado variable de subjetividad intrínseco de los diferentes parámetros de la ejecución musical evaluados. Radocy¹² toca tangencialmente este problema al admitir que “algunos criterios ofrecen un campo más amplio de acuerdo que otros.” Éste es sugerido como una aportación del presente estudio a la discusión genral sobre la delimitación de aspectos a evaluar.

No obstante la profundidad y precisión de su modelo, McPherson &

8 Ciorba, C., & Smith, N. (2009). “Measurement of Instrumental and Vocal Undergraduate Performance Juries Using a Multidimensional Assessment Rubric”, *Journal of Research in Music Education*, pp. 5-15

9 Latimer, M., Bergee, M., & Cohen, M. (2010). “Reliability and Perceived Pedagogical Utility of a Weighted Music Performance Assessment Rubric”, *Journal of Research in Music Education*, pp.168-183.

10 Mills, J. (1991). “Assessing Musical Performance Musically”. *Educational Studies*, pp. 173-181.

11 Saunders, C., & Holahan, J. (1997). “Criteria-specific rating scales in the evaluation of high school instrumental performance”. *Journal of Research in Music Education*, pp. 259-272.

12 Radocy, R. (1986). “On Quantifying the Uncountable in Musical Behavior”. *Bulletin of the Council for Research in Music Education*, p.26

Thompson¹³ no abordan la experiencia del evaluador desde el punto de vista de su formación y experiencia profesional. Este aspecto será sugerido como una avenida de investigación.

Finalmente, los posibles usos pedagógicos de dichas prácticas evaluativas tampoco son incluidas en su modelo. Latimer *et al.* (2010, p. 179), encontraron que ‘evaluadores y directores de escuela [de un festival musical de educación media] perciben en el uso de rúbricas un mayor potencial pedagógico si se comparan con otras estrategias que no usaron rúbricas.’ La percepción y el uso de las rúbricas como herramientas para los procesos de enseñanza y aprendizaje es una idea fundamental para el presente trabajo.

MÉTODO

Después de realizar una revisión de literatura preliminar que señalara posibles rutas de análisis, el investigador realizó un grupo de enfoque por cada uno de los tres grupos generacionales de estudiantes que habían sido evaluados con esta estrategia. En diciembre del mismo año, se realizaron entrevistas semiestructuradas a los cuatro coordinadores de las academias mencionadas, así como al jefe del departamento de música, y se revisaron las rúbricas que se habían usado desde el inicio de la implementación. Finalmente, se realizaron 6 entrevistas más, entre los profesores de las diferentes academias, con el fin de corroborar y profundizar en los resultados parciales arrojados en la primera etapa.

HALLAZGOS Y DISCUSIÓN

Las rúbricas varían en aspectos y descriptores según la academia. Cada aspecto tiene una ponderación que puede cambiar de acuerdo al tipo de examen o al nivel. Se partió de la idea de que “cada instrumento requiere ser evaluado de manera diferente,”¹⁴ y esta estrategia de trabajo ha resultado conveniente, en opinión de todos los entrevistados. No obstante las diferencias, todas incluyen un punto destinado al trabajo en clase que evalúa únicamente el maestro titular. Éste aspecto se define como altamente subjetivo ya que ningún profesor demuestra claridad sobre sus descriptores y los testimonios señalan interferencia de las relaciones personales docente-estudiante. Sin embargo, su inclusión se propone como una aporte a la discusión sobre el uso de criterios.

13 McPherson, G., & Thompson, W. (1998). “Assessing Music Performance: Issues and Influences”. *Research Studies in Music Education*, pp.12-24.

14 McPherson, G., & Thompson, W. (1998). “Assessing Music Performance: Issues and Influences”. *Research Studies in Music Education*, p.13.

Los testimonios indicaron que sólo dos de las academias continuaron realizando revisiones semestrales a sus rúbricas. Esto se pudo constatar en los documentos del archivo del departamento de música. La percepción de utilidad pedagógica de las rúbricas en estas dos academias fue alta. Por el contrario, de las otras dos, en una se encontró poco útil, y en la otra no hubo una opinión uniforme.

Sobre el proceso de diseño, los docentes reportaron dificultades al definir los aspectos, debido a su interrelación: ej., afinación y calidad sonora, o ritmo, articulación y limpieza. Esta interrelación, no en el diseño, pero en la aplicación, fue reportada por Latimer, quienes sugieren estudios relacionados a la pertinencia de aspectos específicos y su delimitación de acuerdo a las diferentes dimensiones de la ejecución musical.

En conexión con esto, se observó un cambio en la adición, fusión o substracción de aspectos, con respecto a las primeras rúbricas empleadas, así como la intención de agregar aspectos no contemplados hasta el momento. Ej., la fusión de limpieza, ritmo y control del pulso en alientos, la fusión de interpretación y memoria en guitarra, la inclusión del concepto de estilo dentro del aspecto interpretación –a pesar de que originalmente se acordó no incluirlo por su alta subjetividad intrínseca-, o la propuesta de incluir desenvolvimiento escénico. Estos hallazgos sugieren que la metodología empleada promueve la concientización de lo que están enseñando y/o lo que quisieran enseñar. No obstante, las fusiones descritas también pueden interpretarse como una falta de claridad, evidenciando la necesidad de sistemas que propicien una definición válida y confiable de ejecución de calidad, y guíen a estudiantes y profesores en su búsqueda.

Así mismo, se identificaron dos aspectos que presentan una alta subjetividad intrínseca y dificultad para delimitar sus descriptores: la calidad de sonido y la interpretación. Sobre el primero los docentes argumentaron que depende de tres factores: la escuela de formación, la obra y del concepto estilístico empleado. Además, consideran difícil evaluar el sonido de un instrumento diferente al de su especialidad y encontrar una terminología adecuada para sus descriptores. Sobre la interpretación, argumentaron a cerca de su evidente complejidad, y que se puede confundir con el estilo, caracterizado anteriormente como altamente subjetivo.

Sobre las variaciones en el uso de las rúbricas, se encontraron dos extremos: los docentes que las usan en la práctica, y los que no. Estas dos variaciones guardan relación con dos perfiles de docentes caracterizados: a) no siente haber recibido atención integral como estudiante, su maestro se centraba en cuestiones técnicas, nunca tuvo claridad sobre qué y cómo lo evaluaban, no recibía retroalimentación o era limitada, ha suplido estas carencias de manera autodidacta, y declara un interés en cubrir en su enseñanza las deficiencias de

su formación; b) el perfil opuesto. El presente estudio no puede definir causas posibles de la variación observada en el uso de las rúbricas; sin embargo se halló que, los profesores con el perfil a) usan las rúbricas y las han encontrado útiles, no sólo para la evaluación y retroalimentación, sino para la planeación de su enseñanza. Por el contrario, los profesores del perfil b) no las usan, y continúan apegados a la definición de criterios personales y a una concepción holística de la evaluación. Sólo se encontró una excepción de una maestra de perfil b) quien ha encontrado las rúbricas de utilidad, como los profesores a).

Con respecto a la calificación, todos coincidieron en que les era difícil seguir la rúbrica para asignar una nota, y, como sus estudiantes, no la consideran factor promotor del aprendizaje. Este hallazgo es identificado como un área oportunidad importante de la estrategia, en tanto representa un desconocimiento de uno de los rasgos fundamentales de la evaluación caracterizados en la introducción de esta ponencia. Así mismo, en cuanto a los efectos de la capacitación para la evaluación con criterios¹⁵, las dificultades generales descritas que experimentan los docentes pueden representar una consecuencia de la falta de capacitación sistemática. Un hallazgo revelador fue que ningún docente pudo definir rúbrica de manera integral. Sus definiciones demostraron una falta de reflexión profunda sobre el concepto.

Finalmente, en cuanto a la percepción de los estudiantes, la mayoría cree que esta estrategia es útil sólo para los docentes. No demuestran consciencia del efecto que, a juzgar por los testimonios de sus profesores, esta estrategia esta teniendo en los procesos de enseñanza y aprendizaje. No obstante, los que reportaron un conocimiento más profundo de las rúbricas, coincidieron con ser alumnos de los profesores con el perfil a).

Por otro lado, los estudiantes revelaron preocupaciones a cerca de factores que corroboran o complementan el modelo de McPherson y Thompson. Se preguntan hasta qué punto las largas jornadas de evaluación afectan la percepción de los docentes. Advierten sobre el sesgo potencial al escuchar primero alumnos de semestres avanzados, capaces de ofrecer ejecuciones de mayor nivel, y luego los de semestre inferiores. Reportan haber experimentado detrimentos importantes en su ejecución a causa del desempeño del pianista acompañante, la cantidad y dificultad del repertorio, y las malas condiciones del instrumento. Y alertan sobre el efecto de los cambios en el sínodo dentro o entre sesiones de evaluación, cuestionando la confiabilidad de algunos instrumentos.

15 Winter, N. (1993). "Music Performance..."p.34-39

CONCLUSIONES

En lo general, esta investigación verifica los retos de la evaluación de aprendizajes en ejecución musical: resalta su subjetividad, su impacto en los procesos de enseñanza y aprendizaje, y la necesidad de reconocer la función pedagógica de la calificación.

En lo particular, aporta a la literatura sobre los factores que condicionan este tipo de evaluaciones, señalando dificultades en la definición de aspectos; subjetividad notoria en la apreciación de sonido, interpretación y estilo; necesidad de incluir del aspecto desempeño en clase; y el efecto de: la duración de las sesiones de examen, las condiciones del instrumento, la ejecución del pianista acompañante, y la cantidad y dificultad del repertorio. Así mismo, los hallazgos respaldan las conclusiones de estudios que señalan el uso de criterios específicos y rúbricas como herramientas idóneas para organizar y comunicar la retroalimentación, y añaden que, para un cierto perfil de docente, pueden ser además útiles en la planeación de su enseñanza.

La investigación no puede establecer relaciones causales entre los perfiles docentes caracterizados y las variaciones en el uso de las rúbricas, pero sirve como base para sugerir estudios sistemáticos que exploren relaciones entre la experiencia estudiantil del docente, su ideología y su práctica evaluativa actual.

La revisión de experiencias en el diseño y validación de rúbricas, estudios sobre la capacitación para la evaluación, y hallazgos como la el efecto de la variación del sínodo, sugieren que el caso investigado presenta áreas de acción importantes. Es probable que las dificultades que los docentes aún siguen experimentando en el diseño de rúbricas, y las inconsistencias que presenta la estrategia esté, en gran parte, relacionado con la falta de una capacitación técnica inicial y una subsiguiente reflexión y revisión continua.

Finalmente, la falta de consciencia de los estudiantes respecto al efecto en la enseñanza y aprendizaje reportado por sus profesores, es un signo que alerta sobre la brecha existente entre las acciones pedagógicas del docente y la comprensión de las mismas por parte de los estudiantes. Mantener las reglas del juego ocultas para nuestros alumnos restringe su aprendizaje.¹⁶ Sólo la comunicación directa de nuestras creencias, acciones y propósitos como docentes pueden aminorar esta brecha.

16 Brown, S. (2004). "Assessment for Learning". Learning and Teaching in Higher Education , pp.81-89.

BIBLIOGRAFÍA

- Latimer, M., Bergee, M., & Cohen, M. (2010). "Reliability and Perceived Pedagogical Utility of a Weighted Music Performance Assessment Rubric", *Journal of Research in Music Education*, 168-183.
- Ciorba, C., & Smith, N. (2009). "Measurement of Instrumental and Vocal Undergraduate Performance Juries Using a Multidimensional Assessment Rubric", *Journal of Research in Music Education*, 5-15.
- Abeles, H. (1973). "Development and Validation of a Clarinet Performance Adjudication Scale". *Journal of Research in Music Education*, 246-255.
- Bergee, M. (1993). "A Comparison of Faculty, Peer, and Self-Evaluation of Applied Brass Jury". *Journal of Research in Music Education*, 19-27.
- Brown, S. (2004). "Assessment for Learning". *Learning and Teaching in Higher Education*, 81-89.
- Butler, S., & McMunn, N. (2006). **A teacher's guide to classroom assessment. Understanding and using assessment to improve student learning.** San Francisco, CA.: Jossey-Bass Teacher.
- McPherson, G., & Thompson, W. (1998). "Assessing Music Performance: Issues and Influences". *Research Studies in Music Education*, 12-24.
- Mills, J. (1991). "Assessing Musical Performance Musically". *Educational Studies*, 173-181.
- Radocy, R. (1986). "On Quantifying the Uncountable in Musical Behavior" . *Bulletin of the Council for Research in Music Education*, 22-31 .
- Saunders, C., & Holahan, J. (1997). "Criteria-specific rating scales in the evaluation of high school instrumental performance". *Journal of Research in Music Education*, 259-272.
- Stanley, M., Brooker, R., & Gilbert, R. (2002). "Examiner Perceptions of Using Criteria in Music Performance Assessment". *Research Studies in Music Education*, 46-56.
- Stiggins, R. (2008). **Assessment Manifesto. A call for the development of balanced assessment systems.** Portland, OR.: ETS.

Winter, N. (1993). "Music Performance Assessment: A Study of the Effects of Training and Experience on the Criteria Music Examiners". *International Journal of Music Education*, 34-39.

