

EVALUACION DE LA PERFORMANCE EN LAS PRÁCTICAS DE PRODUCCION CANTADA: ¿CRITERIOS OBJETIVOS O SUBJETIVIDAD INFORMADA?

Valles, Mónica y Martínez, Isabel

Universidad Nacional de La Plata

La performance musical es un proceso complejo (Rusell 2010) que pone en juego la integración de conductas de diferente índole e involucra aspectos tales como lo afectivo, lo cognitivo, lo psicomotor y lo metacognitivo (Gutierrez 2005). Constituye una *“manifestación evidente de procesos psicológicos, fisiológicos y neurológicos que un individuo realiza en relación al sonido musical, eventos o fenómenos”* (Boyle y Radocy 1987, p. 4). Según Mc Pherson y Thompson (1998), la evaluación de la ejecución musical es *“un proceso mediante el cual un individuo intenta equilibrar y sintetizar las diferentes cualidades de la actuación de otro individuo con el objetivo de promover un juicio que puede estar representado por un rango, un grado o una valoración cualitativa”* (p.12) y sirve tanto para describir un rendimiento como para guiar el mejoramiento del mismo.

Si bien estas cualidades pueden ser medidas, parece difícil que una evaluación resulte absolutamente objetiva. La evaluación de la ejecución es inherentemente subjetiva ya que cualquier juicio sobre una performance musical se basa en sensaciones procesadas en el cerebro de un juez y aunque los oyentes acuerden sobre el rendimiento en aspectos particulares de la performance, tales como la precisión de las notas, la adecuación de los *tempi*, matices del fraseo o la calidad del sonido, la valoración final resulta de las decisiones de los individuos funcionando como ‘sujetos’, por lo tanto, son subjetivas (Boyle y Radocy *op.cit.*). La naturaleza de lo que se intenta evaluar determina en gran medida los criterios de evaluación y teniendo en cuenta el componente de ‘sentimiento’ de la música, resulta ingenuo pensar en una evaluación totalmente objetiva (Prieto Alberola 2001). La evaluación de un producto musical, como la de cualquier otro producto artístico, se halla fuertemente influenciada por las características individuales de quien está evaluando, de manera que *“el valor de una obra de arte depende de cuál sea la perspectiva de cada crítico...hay factores que pueden inducir al crítico al error: su capacidad para comprender la obra, el interés que le suscite y su valoración de la originalidad”* (Young 1997, citado por Prieto Alberola *op. cit.*). Mientras que la subjetividad en la evaluación de una performance es insoslayable dado el alto componente sentimental de la música, la necesidad de una base objetiva no puede ser negada (Gutierrez 2005). La información subjetiva tanto como la objetiva son esenciales en la obtención de evaluaciones válidas y confiables de la performance musical.

La evaluación del desempeño musical es un tema polémico. Una de las grandes discusiones al respecto se centra en si es o no apropiado evaluar una performance musical en componentes separados. Esto da lugar a controversias entre optar por una evaluación de tipo global o utilizar una evaluación puntual es decir, parcializada en categorías.

Una evaluación global se basa en la impresión general del evaluador. Desde este acercamiento, una performance musical es juzgada como una entidad y o bien se le asigna una puntuación o bien se la clasifica en alguna de entre varias categorías disponibles, o en relación a otros trabajos. En contraste, una evaluación puntual implica un juicio basado en datos detallados o específicos de la ejecución de manera que un evaluador puede seguir una lista de verificación (o control) y asignar una puntuación, o hacer comentarios referidos a los diferentes aspectos de una ejecución tales como la articulación, el ritmo, la entonación, estilo, dinámicas, etc. (Boyle y Radocy *op.cit.*).

En tanto algunos consideran que parcializar en categorías una evaluación entorpece la percepción de la música como totalidad, algunos estudios (Burnsed, Hinkle, King, 1985; Mills, 1987; Bergee, 1995; Saunders & Holahan, 1997; Zdzinski & Barnes, 2002; Russell, 2007)¹ mostraron que la evaluación de la performance en componentes separados no incide

¹ Citados por Rusell 2010

en la evaluación global de la calidad del desempeño. Desde un punto de vista pedagógico, la evaluación de factores musicales separados puede resultar útil con fines diagnósticos para promover el mejoramiento de estrategias de enseñanza y aprendizaje (Russell, *op. cit.*). Mc Pherson (1998) propone un modelo de evaluación de la performance musical compuesto por una multiplicidad de factores musicales y no musicales interrelacionados que intervienen tanto en el resultado de la performance como en el juicio que un evaluador hace de la misma. Según este modelo, el determinante más importante en cualquier proceso de evaluación es la calidad de la interpretación musical en referencia al 'acto real de ejecución', pero existen muchas otras influencias que impactan en una interpretación y en su evaluación por parte de los jueces. Uno de los factores interactuantes es el instrumento de evaluación. Desde la perspectiva de Mc Pherson, el propósito de la evaluación es fundamental tanto para la selección de un instrumento de evaluación como para la de los criterios subyacentes y éstos varían según la evaluación tenga un fin pedagógico, de investigación o de clasificación de intérpretes en un concurso.

En el intento de garantizar la equidad de la evaluación es importante atender a la correspondencia entre los juicios de diferentes evaluadores y/o el grado de coherencia con que un juez determinado evalúa un conjunto de actuaciones en diferentes ocasiones (confiabilidad) y verificar el grado en que el instrumento mide efectivamente lo que se pretende medir (validez). No obstante, más allá del cuidado que pueda ponerse en este sentido, los diversos factores involucrados en torno al proceso de evaluación pueden incidir en la confiabilidad de un instrumento, y "*en general se reconoce que en algunos contextos de evaluación la confiabilidad es pobre*" (Mc Pherson *op. cit.*). Cuestiones como la homogeneidad o heterogeneidad en la calidad de los evaluados o el llamado 'efecto de orden' (por el cual los ejecutantes que interpretan cerca del inicio de la evaluación tienen más probabilidades de obtener mejores puntuaciones que quienes lo hacen cuando el proceso de evaluación está más avanzado) parecen tener alta incidencia en el resultado del proceso de evaluación.

Este trabajo se centra en el estudio de los procedimientos de evaluación utilizados para estimar el rendimiento de los estudiantes en las prácticas de lectura a primera vista en las cátedras Audioperceptiva 1 y 2 de la Facultad de Bellas Artes de la UNLP.

El desarrollo de competencias vinculadas a la producción musical es uno de los ejes que sustentan el trabajo en dichas cátedras. Las prácticas desarrolladas en el área de producción apelan a habilidades relacionadas con la ejecución rítmica, el canto individual con y sin acompañamiento, el canto a voces y la elaboración de arreglos instrumentales para el acompañamiento de líneas melódicas.

Para evaluar cada una de estas instancias se utilizan planillas en las que el evaluador debe estimar distintos ítems que describen componentes de la habilidad de ejecución puesta en juego y luego estimar una calificación final global. Se supone que la atención puesta en cada ítem debería contribuir a la concreción de la estimación final. La aplicación sostenida del citado instrumento de evaluación en la prueba de lectura cantada a primera vista y el análisis de los resultados de su uso en el interior de la asignatura evidenció diferencias entre los evaluadores al calificar una misma performance e incertidumbres con respecto al alcance de algunos de los ítems de evaluación, lo que parecía dar cuenta de problemas de confiabilidad y validez del instrumento. En un trabajo anterior (Valles, García y Tobio, 2010) se acometió un análisis preliminar del instrumento de testeo, cuyos resultados indicaron la necesidad de obtener mayores precisiones acerca de los procesos reflexivos que los evaluadores ponen en juego a la hora de evaluar el desempeño, cuestión que constituye uno de los propósitos principales de este trabajo.

Objetivos

- Estudiar la actividad reflexiva del docente como proceso que informa sobre la calidad de la evaluación.
- Analizar correspondencias intra e inter evaluadores entre estimaciones puntuales y globales de la performance.

En los informes escritos solicitados a los E se les pidió que analizaran su propio proceso evaluativo respondiendo a las siguientes cuestiones:

- 1) Dificultades encontradas para realizar la evaluación.
- 2) Descripción del significado que para el E revestía cada ítem a los fines de ser considerado como indicador de un atributo de la performance a ser evaluado.
- 3) Criterios de rendimiento según los cuales estimaron la magnitud de la calificación para cada ítem.
- 4) Pertinencia relativa de los ítems utilizados.
- 5) Toda otra consideración relativa al proceso evaluativo.

Las respuestas fueron sometidas a un análisis de contenido del que se desprendieron los patrones de respuesta que resultaron más relevantes.

Con respecto a 1) la información obtenida indica que:

- Evaluar separadamente los aspectos de la performance (esto es, ítem por ítem), genera un esfuerzo de atención a aspectos escindidos de la performance que da por resultado la pérdida de la noción de totalidad de la obra.
- Evaluar separadamente los aspectos de la performance implica disponer de más tiempo para completar cada evaluación individual.
- Utilizar la misma escala (1 a 10) para puntuar aspectos de naturaleza diferente dificulta la evaluación. Mientras algunos componentes de la performance serían susceptibles de ser medidos más 'objetivamente' (por ejemplo cantidad de grupos rítmicos, cantidad de alturas correctos), otros parecen ser de naturaleza más subjetiva (por ejemplo la interpretación).
- Del mismo modo, cuando el ítem a evaluar contiene información global y puntual a la vez, como ocurre con *centro tonal/altura*, donde se demanda al E que aplique un criterio holístico para estimar la estabilidad relativa de la configuración tonal y al mismo tiempo que este criterio tome en cuenta la evaluación puntual de la producción sucesiva de las alturas, el E tiene la impresión de que debe considerar un monto de información que resulta compleja.
- Evaluar el ítem *interpretación* presentó una dificultad para determinar el límite entre una interpretación aprobada y una desaprobada y decidir la calificación en el contexto de una evaluación de lectura a primera vista. Esto lleva a interrogarse acerca de lo que debería esperarse de la interpretación como ejecución expresiva en tal situación.

En relación a 2), el análisis de las descripciones de los ítems provistas por los E, dio cuenta de dos modalidades diferentes de entender el significado de los mismos. En tanto algunos entendieron a los ítems como un 'producto' susceptible de ser medido según dos categorías opuestas, otros lo otorgaron la significación de un 'proceso'. Así, por ejemplo el ítem *continuidad*, fue definido por uno de los E como "Si se detiene o no" (E3), en tanto otro consigna que "se refiere a que el discurso no se fragmente, no se vuelva atrás repitiendo pasajes, sino que la música tenga un discurrir, un transcurrir fluido de principio a fin: cuanto más se fragmente o se interrumpa, más se desdibuja el discurso musical" (E4).

Con respecto a 3), los criterios utilizados por los E para estimar la magnitud de las calificaciones resultaron coherentes con el enfoque según el cual definieron los ítems (2). Tomando nuevamente la *continuidad* como ejemplo, mientras el E3 utilizó el criterio 'poca-mucha', el E4 evaluó este ítem en términos de la inteligibilidad de la performance como comunicación discursiva y consigné que "Desaprobé un ítem cuando la cantidad de errores no me permitía entender o configurar auditivamente la melodía" (E4).

Respecto de 4), las categorías son consideradas pertinentes con las salvedades que se desprenden de las dificultades encontradas en el uso de la planilla (ver punto 1).

En cuanto a 5), otras consideraciones realizadas por los E se vincularon con que todos los ítems son tomados en cuenta con el mismo valor intrínseco cuando posiblemente la experiencia previa de los estudiantes no es suficiente para alcanzar grados equivalentes de

resolución en todos ellos por igual. Tal sería el caso del ítem *interpretación* donde el desarrollo musical que se requiere para resolver rápidamente una performance expresiva en una prueba de lectura a primera vista sea probablemente insuficiente en la población de alumnos examinada.

Por otra parte se sugirió la posibilidad de realizar ajustes en la estimación de la puntuación por ítem de acuerdo al grado de dificultad que cada ítem evaluado tiene en la totalidad de la obra. Se propuso además la inclusión de un ítem que contemple la capacidad de detección y resolución de dificultades por parte del alumno durante la performance, esto es, que de cuenta del uso de habilidades de reflexión en la acción (Schön 1992).

(ii) Análisis del acuerdo inter e intra E en la evaluación de la performance

En primer lugar se estimó el grado de acuerdo alcanzado por los E respecto de si aprobar o desaprobar al evaluado (el puntaje mínimo de aprobación en las pruebas de lectura a primera vista es de 6 puntos). Los resultados muestran que se alcanzó un 100% de acuerdo inter E en cinco de los diez casos. De los cinco restantes, en cuatro el acuerdo fue del 80%, en tanto que en el quinto caso se obtuvo un 60% de acuerdo.

A continuación se estimó el acuerdo inter E tanto en las calificaciones por ítem (*continuidad, centro tonal/ alturas, tempo/ritmo e interpretación*) como en la calificación global (*nota*). Se calculó el coeficiente de correlación r de Pearson. El análisis mostró un acuerdo parcial entre E. Por ejemplo, *continuidad, centro tonal/ alturas y nota* mostraron acuerdos entre los E 1 y 2 (0,79; 0,81 y 0,83 respectivamente, siendo $p < 0,01$), los E 3 y 4 (0,80; 0,79 y 0,90 respectivamente, siendo $p < 0,01$) y los E 3 y 5, (0,78 siendo $p < 0,01$; 0,74 y 0,67 siendo $p < 0,05$). Los E 4 y 5 muestran un grado de acuerdo en *centro tonal/ alturas y 'nota'* (0,63 y 0,74 respectivamente, siendo $p < 0,05$)

Se analizó la correlación entre la calificación de *nota* con el resto de los ítems (figura 2). Los resultados muestran lo siguiente:

- E1: el ítem que mejor correlaciona es *centro tonal/alturas* y el que menos correlaciona es *interpretación*.
- E2: todos los ítems correlacionan en un grado similar.
- E3: todos los ítems correlacionan pero en distinta medida. El que muestra la correlación más alta es *interpretación*. Le siguen *tempo/ritmo, continuidad y centro tonal/alturas*.
- E4: la correlación más alta se produce con *centro tonal/alturas* y la más baja con *interpretación*. Se observó una correlación negativa con el ítem *tempo/ritmo*. En el informe escrito este E informó que, según su criterio, la dificultad de los fragmentos musicales que los alumnos debieron leer a primera vista, se vinculaba más al aspecto melódico que al rítmico, el que resultaba de un grado bajo de dificultad.
- E5: las correlaciones más altas se observan con *centro tonal/alturas y tempo/ritmo* y la más baja con *continuidad*. No es posible estimar la correlación para el ítem *interpretación* ya que este E consignó la puntuación más baja (1) en todos los casos evaluados. En su informe manifestó no tener certeza sobre qué es lo esperable de una interpretación en una lectura a primera vista y según su criterio, no hubo una ejecución expresiva en ninguno de los casos evaluados razón por la cual optó por otorgar la calificación más baja. Así mismo informó que según su criterio los ítems más importantes para evaluar la ejecución de una lectura a primera vista son la estabilidad del centro tonal y la continuidad de la ejecución ya que de ello depende una buena performance de las alturas, el tempo y el ritmo.

Correlaciones						
		Continuidad E1	CTAlturasE1	Tempo RitmoE1	InterpretaciónE1	NotaE1
NotaE1	Correlación de Pearson	,837**	,925**	,893**	,849**	1
	Sig. (bilateral)	,003	,000	,001	,042	
	N	10	10	10	10	10

** La correlación es significativa al nivel 0,01 (bilateral).
* La correlación es significativa al nivel 0,05 (bilateral).

Correlaciones						
		Continuidad E2	CTAlturasE2	Tempo RitmoE2	InterpretaciónE2	NotaE2
NotaE2	Correlación de Pearson	,947**	,905**	,981**	,943**	1
	Sig. (bilateral)	,000	,000	,000	,000	
	N	10	10	10	10	10

** La correlación es significativa al nivel 0,01 (bilateral).

Correlaciones						
		Continuidad E3	CTAlturasE3	Tempo RitmoE3	InterpretaciónE3	NotaE3
NotaE3	Correlación de Pearson	,935**	,869**	,946**	,954**	1
	Sig. (bilateral)	,000	,001	,000	,000	
	N	10	10	10	10	10

** La correlación es significativa al nivel 0,01 (bilateral).

Correlaciones						
		Continuidad E4	CTAlturasE4	Tempo RitmoE4	InterpretaciónE4	NotaE4
NotaE4	Correlación de Pearson	,882**	,956**	-,097	,782**	1
	Sig. (bilateral)	,001	,000	,790	,008	
	N	10	10	10	10	10

* La correlación es significativa al nivel 0,05 (bilateral).
** La correlación es significativa al nivel 0,01 (bilateral).

Correlaciones						
		Continuidad E5	CTAlturasE5	Tempo RitmoE5	InterpretaciónE5	NotaE5
NotaE5	Correlación de Pearson	,835**	,934**	,905**	,835**	1
	Sig. (bilateral)	,003	,000	,000	,000	
	N	10	10	10	10	10

** La correlación es significativa al nivel 0,01 (bilateral).
a. No se puede calcular porque al menos una variable es constante.

Figura 2. Correlación entre la calificación de nota y el resto de los ítems en cada uno de los evaluadores representados por la letra E y los números 1 a 5.

Discusión

Los resultados del presente estudio arrojan luz acerca de la riqueza y complejidad de los procesos reflexivos que se ponen en juego en el evaluador al momento de juzgar una performance musical. Asimismo ponen de manifiesto que el instrumento de evaluación es sólo un eslabón dentro del proceso de evaluación y que se halla interrelacionado con una diversidad de factores tanto musicales como no musicales que intervienen en dicho proceso y que determinan la actuación tanto del evaluado como la del evaluador.

El citado instrumento contempla dos miradas de la performance: puntual y global. Si bien la evaluación puntual podría hacer perder de vista la performance como totalidad, ésta se equilibra con la inclusión de la nota global que refleja una visión más holística de la ejecución. Para Boyle y Radocy (*op. cit.*) “es posible un balance entre ambos tipos de acercamiento donde la evaluación de aspectos particulares de la performance funcione como una guía más que como categorías que deben ser cuantificadas”. Se estima que la consideración por ítems organiza la tarea del evaluador al momento de dar feedback al estudiante acerca del resultado de la tarea evaluada, porque trae a la conciencia del evaluador aquellos aspectos de la performance que deben ser mejorados.

Si bien se observó consenso entre evaluadores respecto de la pertinencia de los ítems de la planilla vigente, las diferencias encontradas en los informes indicarían la necesidad de discutir con mayor profundidad y definir con mayor precisión el alcance de cada uno de ellos.

La mayor dificultad que ofrecen algunos de los ítems para ser puntuados, podría vincularse con el hecho de utilizar una misma metodología de puntuación para ítems de diferente naturaleza. Así, sería posible distinguir entre ítems que corresponden al conteo de eventos de la superficie musical (*alturas y ritmo*) y otros de naturaleza más compleja y/o estructural (*centro tonal, tempo, continuidad e interpretación*). Se considera que la elección de

diferentes metodologías de puntuación según la naturaleza del ítem facilitarían la calificación en cada uno de ellos.

El análisis aquí presentado muestra que la subjetividad del evaluador es un factor determinante en la estimación del desempeño de los estudiantes durante las pruebas de lectura a primera vista. Los evaluadores que participaron de este trabajo son profesores de las asignaturas anteriormente mencionadas, por lo que comparten la perspectiva desde la cual se aborda la evaluación de la lectura a primera vista. A pesar de ello, el análisis de los datos mostró una correlación parcial entre sus desempeños lo que podría ser interpretado como un indicador del carácter altamente subjetivo de la evaluación. Si bien la perspectiva clásica de la evaluación supone que la utilización de instrumentos altamente pautados garantiza un criterio de objetividad, los resultados del estudio –en particular los que emergen de las relaciones entre los informes escritos de los evaluadores y los criterios utilizados para la asignación de puntajes- muestran que la actividad de evaluar es más bien una tarea de índole subjetiva informada por una serie sofisticada de presupuestos ideas y valores que alimentan la reflexión en la acción y acerca de la acción al momento de estimar el desempeño del estudiante.

Los resultados permitieron advertir dos perspectivas diferentes acerca de del proceso de evaluación: un modo de reflexión en términos de resultado y otro en términos de proceso. Éste último informa un modo de evaluación centrado en la performance como proceso de producción en el tiempo. Se considera que el desarrollo de este modo reflexivo es el que debería promoverse en el interior del grupo de evaluadores si es que se quiere alcanzar una subjetividad informada y un criterio de verdad entre evaluadores como comunidad de expertos.

Referencias

- Boyle, J.D. y Radocy, E. R. (1987) *Measurement and Evaluating of Musical Experiences*. New York: Schirmer Books.
- Gutierrez, A. (2005) The Development of the Music Performance Assessment Tool (MPAT) For College Music Majors in the Piano, Strings, Voice and Guitar Departments. http://www.ssc.edu.ph/sscweb_content/Gutierrez%20Music%20Research%20for%20Europe%20Conf%202005.PDF (Página consultada entre el 28 de Julio y el 20 de Agosto de 2011).
- McPherson, G. and Thompson, W., (1998) Assessing Music Performance: Issues and Influences. *Research studies in Music Education*, Vol 10 N° 1, pp. 12 a 24.
- McPherson, G (1995) The Assessment of Musical Performance: Development and Validation of Five New Measures. *Psychology of Music*, Vol 23 N° 2, pp. 142 a 161.
- Prieto Alberola, R. (2001) La evaluación de las actividades musicales. *Contextos educativos: Revista de educación*, N° 4, pp. 329 a 340
- Russell, B. (2010) *The Empirical Testing of Assessment Paradigm*. Electronic Theses and Dissertations, University of Miami. http://scholarlyrepository.miami.edu/cgi/viewcontent.cgi?article=1386&context=oa_dissertations&sei-redir=1#search=%22assessment%20musical%20performance%22. (Página consultada entre el 2 de Julio y el 22 de Agosto de 2010).
- Schön, D. (1992) *La formación de profesionales reflexivos*. Paidós, Barcelona. 1992.
- Valles, M.; García, G. y Tobio P. (2010) ¿Es la evaluación un índice confiable de la competencia de ejecución? En F. Shifres (Ed.) *Actas de las II Jornada de Desarrollo Auditivo en la Formación del Músico Profesional*, pág. 20 a 25. Buenos Aires: SACCoM.