

Fiabilidad y validez en los potenciales evocados auditivos del tronco cerebral

Franz Zenker* , José Juan Barajas de Prat

Clínica Barajas, Santa Cruz de Tenerife, España

 OPEN ACCESS

PEER REVIEWED

ARTÍCULO ORIGINAL

DOI: 10.51445/sja.auditio.vol9.2025.114

Recibido: 21.02.2025

Revisado: 10.03.2025

Aceptado: 24.06.2025

Publicado: 22.10.2025

Editado por:

Gerard Encina-Llamas

Facultat de Medicina. Universitat de Vic - Universitat Central de Catalunya (UVic-UCC) / Copenhagen Hearing and Balance Center (CHBC) - Rigshospitalet (Dinamarca).

Revisado por:

Miguel Temboury-Gutiérrez

Hearing Systems section. Department of Health Technology. Technical University of Denmark (DTU).

Lidia Roig Canales

Hospital Universitari Sant Joan de Déu, Barcelona.

Vicente Curcio

Universidad del Museo Social Argentino (UMSA), Argentina.

Cómo citar:

Zenker, F. y Barajas de Prat, J.J. (2025). Fiabilidad y validez en los potenciales evocados auditivos del tronco cerebral. *Auditio*, 9, e114. <https://doi.org/10.51445/sja.auditio.vol9.2025.114>

Correspondencia

*Franz Zenker

Calle, Pérez de Rozas 8, 38004 Santa Cruz de Tenerife, Islas Canarias, España.

Correo electrónico: zenker@clinicabarajas.com

Resumen

Los registros electrofisiológicos permiten evaluar la integridad funcional de las estructuras auditivas periféricas y centrales. Entre ellos, Los Potenciales Evocados Auditivos del Tronco Cerebral (PEATC) constituyen una herramienta fundamental para el estudio de la neurofisiología de las vías auditivas. Debido a la baja amplitud y a la inevitable contaminación por ruido tanto biológico como externo, se requiere el uso de procedimientos que mejoren la calidad y definición de los registros obtenidos.

Este estudio tiene como objetivo principal evaluar la fiabilidad y validez de los PEATC. La fiabilidad se analiza mediante el Cociente de Desviación Estándar (SDR: Standard Deviation Ratio); y el Coeficiente de Correlación entre Hemipromedios (CCR: Correlation Coefficient between Replicates); los cuales permiten valorar la consistencia y reproducibilidad de las respuestas. La promediación ponderada mejora la relación señal-ruido (SNR: Signal-to-Noise Ratio) y reduce el tiempo de la evaluación electrofisiológica.

La validez convergente evalúa la correlación entre pruebas electrofisiológicas y conductuales en la medición de un mismo constructo, en este caso, la percepción auditiva. La validez de criterio determina el grado en que una prueba se correlaciona con una variable externa de referencia, como el grado de hipoacusia o el diagnóstico clínico de pérdida auditiva. Se presentan diferentes procedimientos estadísticos para establecer el umbral auditivo a partir de los PEATC. Se postula que, para la estimación de los umbrales auditivos mediante PEATC, cada centro

Implicaciones Clínicas

La precisión de los Potenciales Evocados Auditivos del Tronco Cerebral depende de factores como la relación señal/ruido, la estabilidad de la respuesta neurofisiológica y la variabilidad inter e intraobservador, así como la variabilidad biológica. Para evaluar la calidad de los registros, se emplean estimadores estadísticos que permiten cuantificar estas fuentes de variabilidad. Técnicas como la promediación ponderada mejoran la fiabilidad y reducen el tiempo de prueba. Además, la precisión aumenta cuando cada centro establece modelos de ajuste entre pruebas fisiológicas y conductuales, garantizando estimaciones consistentes y reproducibles de los umbrales auditivos. Estas mejoras en la precisión permiten una caracterización más fiable de la sensibilidad auditiva, lo que facilita el diagnóstico y la selección adecuada de estrategias de intervención, especialmente en pacientes poco colaboradores.

debe establecer sus propios grupos de control. Asimismo, se analizan los procedimientos más utilizados en la actualidad para mejorar la correlación entre los PEATC y las respuestas conductuales.

Introducción

En la práctica clínica, se emplean diversos métodos para evaluar la función auditiva, que van desde procedimientos conductuales, como la audiometría tonal liminal, hasta evaluaciones objetivas, como las Emisiones Otoacústicas (EAOE) y los Potenciales Evocados Auditivos del Tronco Cerebral (PEATC) (Norton et al., 2000; Núñez Batalla et al., 2020; Widen et al., 2005). Estos exámenes deben cumplir con criterios metodológicos que respalden diagnósticos precisos y faciliten la toma de decisiones terapéuticas adecuadas. En este contexto, es crucial responder a tres preguntas clave: ¿mide el componente neurofisiológico registrado realmente lo que pretende medir?, ¿son los registros consistentes y reproducibles en diferentes momentos y condiciones? y, ¿es capaz el examen de detectar cambios clínicamente significativos o diferencias relevantes entre individuos o grupos?

Los PEATC permiten obtener respuestas objetivas asociadas a la sensibilidad auditiva. La respuesta neurofisiológica permite evaluar la funcionalidad de la vía auditiva desde la cóclea y el nervio auditivo hasta el tronco encefálico (Barajas de Prat et al., 2007; Barajas, 1985; Delgado Hernández et al., 2003). La baja amplitud de la señal fisiológica, así como la presencia de artefactos, influye directamente en la calidad de los registros. El estado de alerta del paciente, los movimientos musculares o las perturbaciones eléctricas pueden generar variaciones en los resultados, afectando la precisión de los mismos. Para mejorar su fiabilidad, se emplean procedimientos como el filtrado de la señal y la promediación. Este último, introducido por Dawson (1954), consiste en combinar las respuestas obtenidas tras la presentación repetida del estímulo, minimizando así las interferencias aleatorias de la actividad electroencefalográfica de fondo. El filtrado de la señal elimina las frecuencias no deseadas mediante la aplicación de filtros de paso alto o paso bajo, restringiendo el registro a un rango de frecuencias específico (Doyle & Hyde, 1981). La configuración de este parámetro se establece en función del componente o la actividad neurobiológica que se desea analizar. Por ejemplo, en el espectro de frecuencias,

Palabras clave

Potenciales Evocados Auditivos del Tronco Cerebral, fiabilidad, validez, sensibilidad, especificidad.

los PEATC se encuentran principalmente entre los 500 Hz y los 1500 Hz. Por esta razón, es común aplicar un filtrado de paso de banda en el rango de 30 Hz a 100 Hz para el de paso alto y de 1500 Hz a 3000 Hz para el de paso bajo (Elberling, 1979). Este filtrado permite lograr dos objetivos fundamentales: por un lado, eliminar el ruido de baja frecuencia, el cual está dominado por artefactos fisiológicos como la actividad muscular (ruido miogénico); y, por otro lado, reducir el ruido de alta frecuencia, que generalmente incluye interferencias eléctricas y ruido ambiental.

La validez se refiere a la capacidad de estimar la sensibilidad auditiva a partir de las respuestas electrofisiológicas. Este concepto se apoya en dos indicadores fundamentales: la sensibilidad y la especificidad. La sensibilidad indica la capacidad de la prueba para identificar correctamente los casos positivos, es decir, los sujetos con pérdida auditiva. Por su parte, la especificidad mide la capacidad para detectar los casos negativos, distinguiendo a los sujetos con audición normal.

Este estudio analiza la precisión y fiabilidad de los PEATC, considerando criterios como la fiabilidad, la validez, la sensibilidad y la especificidad. Se examinan los fundamentos metodológicos que garantizan una evaluación precisa y reproducible, incluyendo estrategias para minimizar artefactos y mejorar la relación señal/ruido. Asimismo, se destaca el impacto de estos procedimientos en la optimización del diagnóstico auditivo, permitiendo una mejor identificación de la pérdida auditiva y facilitando la toma de decisiones clínicas fundamentadas.

La fiabilidad de la medida: Técnicas de Promediación

Los componentes identificados en el registro electrofisiológico los podemos denominar puntuaciones observadas, que estarían compuestas por dos elementos principales: la puntuación verdadera, que representa la respuesta fisiológica al estímulo, y el error asociado, que corresponde al ruido de fondo del EEG (electroencefalograma). La fiabilidad se define como la proporción de la variabilidad observada atribuible

a la puntuación verdadera. En el caso de los PEATC, la reproducibilidad de los registros es un factor clave de fiabilidad, ya que las pruebas pueden realizarse en condiciones muy variables. Este cociente permite determinar en qué medida los resultados están libres de error y refleja la exactitud de la medición. La siguiente fórmula establece la fiabilidad en relación a la varianza:

$$R = \frac{\sigma_v^2}{\sigma_o^2}$$

Donde:

R: Fiabilidad (valor entre 0 y 1).

σ_v^2 : Varianza de la Puntuación Verdadera (componente deseada de la medida).

σ_o^2 : Varianza total observada, que incluye tanto la varianza verdadera como la varianza del error.

Además, la varianza total observada se descompone de la siguiente manera:

$$\sigma_o^2 = \sigma_v^2 + \sigma_e^2$$

Donde:

σ_e^2 : Varianza del error.

La fórmula establece que la fiabilidad (R) depende directamente de la relación entre la varianza verdadera (σ_v^2) y el error (σ_e^2). Cuando σ_e^2 es pequeña, la fiabilidad se aproxima al valor máximo ($R \approx 1$). Esto ocurre porque, en estas condiciones, casi toda la varianza observada se atribuye a diferencias en las características medidas y no a fluctuaciones aleatorias o errores en la medición. En términos prácticos, esto implica que los resultados del registro son consistentes, estables y reflejan principalmente la respuesta al estímulo presentado. Por el contrario, cuando la varianza de error es grande en relación a la varianza verdadera $\sigma_e^2 \gg \sigma_v^2$, la fiabilidad disminuye. En este caso, las mediciones están tan influenciadas por factores aleatorios o externos que resulta imposible distinguir la señal del ruido de fondo.

Una de las estrategias más efectivas para mejorar la fiabilidad del registro es aumentar el número de promediaciones. Cada barrido adicional reduce progresivamente el impacto del ruido, favoreciendo la detección de la señal. Esta relación se describe mediante la siguiente ecuación, descrita por Hall (2007, p. 95):

$$SNR = \frac{\text{Amplitud de la Señal}}{\text{Amplitud del Ruido}} \times \sqrt{\text{Número de Promediaciones (N)}}$$

La relación señal/ruido (SNR) mejora al aumentar la amplitud de la señal, reducir el ruido o incrementar el número de promediaciones, con una ganancia proporcional a la raíz cuadrada de este último. Sin embargo, esta mejora tiene retornos decrecientes, lo que resulta especialmente relevante en niños pequeños, donde el tiempo de registro es limitado (Cone & Norrix, 2015). La Figura 1 ilustra esta relación, mostrando cómo varía el SNR en función del número de barridos y el nivel de estimulación. El Cociente de Desviación Estándar (SDR) se calcula dividiendo la desviación estándar del promedio del registro (respuesta + ruido) por la del ruido residual. La primera refleja la varianza observada; la segunda, el error. Un SDR alto indica mayor proporción de respuesta frente a ruido. Esta métrica, junto con el cociente de correlación entre hemipromedios (CCR), se seleccionó por su aplicabilidad clínica, sencillez operativa y respaldo en la literatura (Hall, 2007; Picton et al., 1983).

Cociente de Desviación Estándar

El Cociente de Desviación Estándar (por sus siglas en inglés: Standard Deviation Ratio, SDR) es un parámetro utilizado en la evaluación de la calidad del registro (Madsen et al., 2018; Picton et al., 1983). Aunque no equivale directamente al coeficiente de fiabilidad clásica, el SDR permite estimar en qué medida la respuesta evocada es visible por encima del ruido de fondo, facilitando una interpretación más segura de la actividad neurofisiológica. El SDR se calcula como la raíz cuadrada del cociente entre la varianza del promedio del registro (que incluye tanto la respuesta como el ruido) y la varianza del ruido residual:

$$SDR = \sqrt{\frac{\sigma_{\text{señal+ruido}}^2}{\sigma_{\text{ruido}}^2}}$$

Donde:

$\sigma_{\text{señal+ruido}}^2$: Varianza del promedio del registro.

σ_{ruido}^2 : Varianza del ruido residual, estimada mediante la sustracción entre dos hemipromedios del mismo estímulo.

El ruido residual se obtiene a través del procedimiento de referencias cruzadas, que consiste en restar dos hemipromedios de una misma presentación (Schimmel, 1967; Wong & Bickford, 1980). Esta sustracción elimina la respuesta sincronizada, dejando

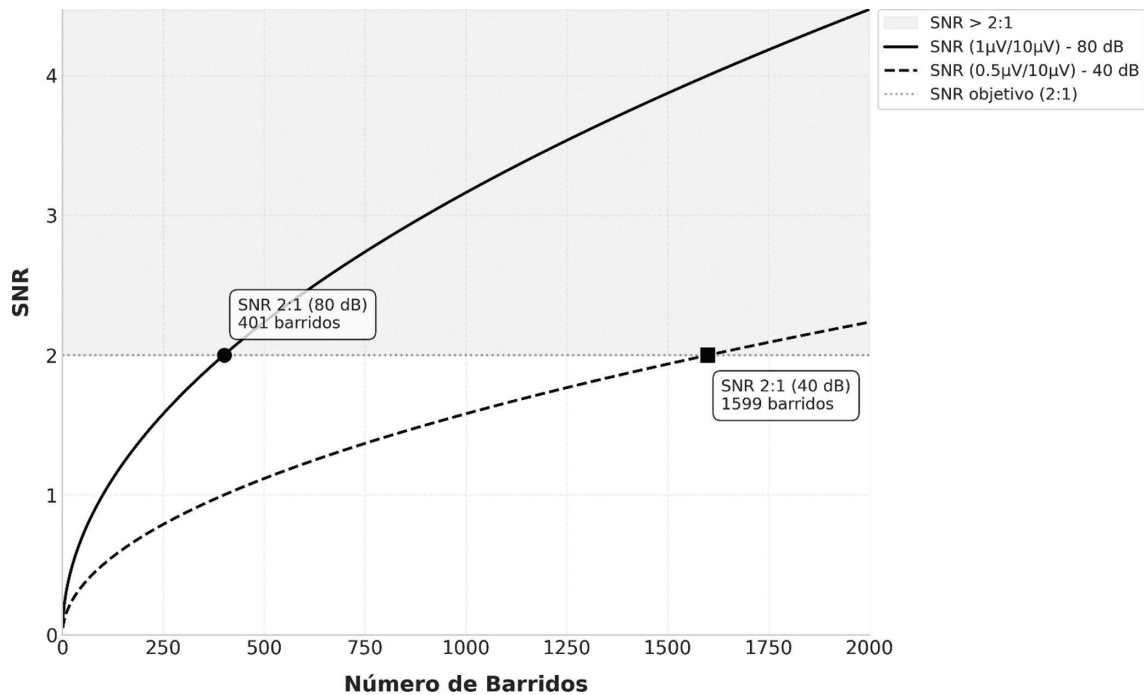


Figura 1. Ejemplo de la evolución del SNR en función del número de barridos en un registro obtenido en un paciente con una pérdida auditiva de 40 dB HL, la amplitud de la onda V a 80 dB nHL es de 1 µV, con un EEG de fondo de 10 µV. Para alcanzar un SNR de 2:1, se requieren aproximadamente 401 promediaciones (línea continua). Sin embargo, a 40 dB nHL, la amplitud de la onda V disminuye a 0.5 µV, lo que incrementa a 1600 las promediaciones para alcanzar el mismo SNR (línea discontinua).

solo el componente aleatorio. Aunque el SDR no representa directamente la relación señal/ruido en decibelios como el SNR, sí actúa como una métrica relativa útil para cuantificar la claridad del registro. Por practicidad, se trabaja con desviaciones estándar en lugar de varianzas, ya que están directamente relacionadas con las amplitudes de los componentes evocados.

Coeficiente de Correlación entre Hemipromedios

El Coeficiente de Correlación entre Hemipromedios o (por sus siglas en inglés: Coefficient of Correlation between Replicates, CCR) mide la fiabilidad o consistencia de las respuestas evocadas promediadas a partir de dos hemipromedios (Picton et al., 1983; Wang et al., 2021). Matemáticamente, se expresa como:

$$CCR = \frac{\sum (X_1 - \bar{X}_1)(X_2 - \bar{X}_2)}{\sqrt{\sum (X_1 - \bar{X}_1)^2 \cdot \sum (X_2 - \bar{X}_2)^2}}$$

Donde:

X_1 : Valores del primer hemipromedio.

X_2 : Valores del segundo hemipromedio.

$\bar{X}_1, \bar{X}_2$: Promedios de cada hemipromedio.

Un CCR alto indica una buena consistencia entre los hemipromedios, lo que sugiere que la señal promediada está poco afectada por el ruido. Valores de $CCR > 0.5$ se consideran aceptables, mientras que valores por encima de 0.7 indican un alto nivel de reproducibilidad (Berninger et al., 2014). Valores bajos de SDR o CCR sugieren que el registro está dominado por el ruido, lo que puede deberse a factores técnicos, como mala colocación de electrodos o ruido ambiental excesivo. También puede deberse a condiciones del paciente, como movimientos oculares. En estos casos, es recomendable repetir el registro tratando mejorar las condiciones del estudio. El CCR está estrechamente relacionado con el SNR. Esta relación se expresa matemáticamente como:

$$SNR = \frac{CCR}{1 - CCR}$$

Donde:

CCR: Coeficiente de Correlación entre Hemipromedios.

SNR: Relación Señal/Ruido.

Esta ecuación establece que el CCR no solo mide la consistencia entre hemipromedios, sino que también actúa como un indicador directo de la preponderancia

de la señal sobre el ruido. A medida que el CCR se aproxima a 1, el SNR aumenta de forma no lineal, lo que refleja una señal cada vez más dominante en comparación con el ruido. En la Figura 2 observamos que un CCR de 0.7 resulta en un SNR de 2.3 dB, mientras que un CCR de 0.95 incrementa el SNR a 19 dB, mostrando un crecimiento exponencial. Por el contrario, valores bajos de CCR, cercanos a 0, implican un dominio del ruido sobre la señal, lo que limita la interpretación del registro.

La interpretación conjunta del SDR y el CCR es esencial para identificar registros que requieren cautela en su análisis y, en caso necesario, su repetición. Un SDR elevado acompañado de un CCR bajo sugiere que, aunque la señal promedio es alta en comparación con el ruido, las réplicas no han sido consistentes. Esto puede deberse a la presencia de artefactos o a malas condiciones de registro, lo que compromete la fiabilidad de los resultados.

La combinación ideal es un SDR alto junto con un CCR alto, lo que refleja un registro robusto donde la señal es claramente distinguible del ruido y se mantiene consistente a través de las réplicas. Sin embargo, un CCR alto con un SDR bajo no implica necesariamente un mal registro, sino que puede indicar

simplemente que la respuesta es de baja amplitud, pero fiable. En estos casos, aunque la señal sea débil en comparación con el ruido, su consistencia entre réplicas sugiere que la medición es válida.

Promediación Ponderada

En las técnicas de promediado clásico, se asume que el ruido presente en el registro es estacionario, es decir, que sus propiedades estadísticas, como la media y la varianza, permanecen constantes a lo largo del tiempo (Mühler & von Specht, 1997). Sin embargo, esta suposición rara vez se cumple en registros reales. En entornos clínicos, factores como movimientos del paciente, parpadeos, ruido ambiental o variaciones en la colocación de los electrodos generan un ruido no estacionario que varía en intensidad y características a lo largo del registro. El promediado clásico no tiene en cuenta estas variaciones temporales, lo que afecta la calidad de los registros al incluir en la promediación barridos contaminados por ruido que distorsionan el resultado final. Para abordar esta limitación, se han desarrollado técnicas como la promediación ponderada, que asumen un modelo no estacionario del ruido (Hoke et al., 1984;

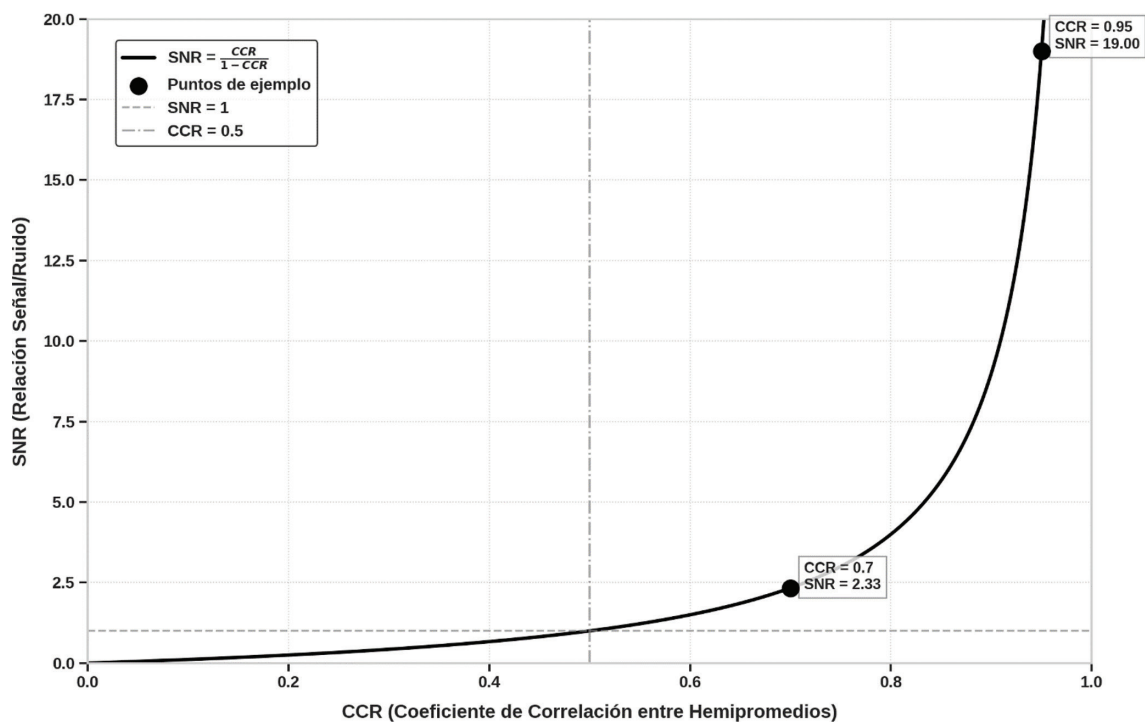


Figura 2. Relación entre el CCR y el SNR. La gráfica ilustra cómo el SNR aumenta de manera exponencial a medida que el CCR se aproxima a 1. Por ejemplo, un CCR = 0.7 corresponde a un SNR = 2.33 dB, mientras que un CCR = 0.95 incrementa el SNR hasta 19 dB. Las líneas de referencia marcan SNR = 1 dB, cuando señal y ruido son equivalentes, y CCR = 0.5, valor intermedio de consistencia.

McKearney et al., 2023). Este enfoque divide el registro en subconjuntos discretos, denominados bloques, permitiendo identificar y excluir segmentos específicos afectados por picos de ruido excesivo.

En este modelo, cada registro se representa como la suma de la señal fija $s(t)$ y un componente de ruido no estacionario $c_i(t)$, que a su vez es el producto de un ruido estacionario $\eta_i(t)$ y un factor multiplicativo $c_i(t)$ que varía lentamente:

$$x_i(t) = s(t) + c_i(t) \cdot \eta_i(t)$$

Donde:

$x_i(t)$ es el registro promedio.

$s(t)$ es la señal de interés.

$c_i(t)$ es un factor que modula la amplitud del ruido.

$\eta_i(t)$ es un ruido estacionario con medida cero y varianza unitaria.

En las técnicas de promediación ponderada, se emplean estadísticos específicos, como la varianza del ruido residual y la potencia promedio de la señal, para optimizar el SNR (Davila & Mobin, 1992; Kumaragamage et al., 2016). La varianza del ruido residual mide la dispersión de las amplitudes del ruido dentro de cada bloque de datos, identificando aquellos segmentos con mayor contaminación de artefactos o ruido de fondo. Por su parte, la potencia promedio de la señal evalúa la intensidad relativa de la respuesta fisiológica al estímulo dentro de cada bloque. El procedimiento comienza con la división de los datos en bloques discretos y la estimación de estos parámetros en cada uno. A continuación, se asignan pesos a los bloques, de manera que aquellos con menor varianza de ruido y mayor potencia de señal tengan mayor influencia en el promedio final.

El análisis en bloques facilita la aplicación de herramientas estadísticas que permiten detectar automáticamente la respuesta. Esto permite reducir el número de promedios necesarios para obtener un registro fiable, ya que este puede detenerse antes de lo establecido al alcanzar un valor de SNR pre establecido. Esto optimiza el tiempo requerido para las pruebas y minimiza la carga sobre el paciente, especialmente en poblaciones pediátricas o difíciles de evaluar. Por ejemplo, en recién nacidos, las técnicas ponderadas pueden evitar el uso de sedación (Cone & Norrix, 2015), como se representa esquemáticamente en la Figura 3, donde se ilustra el proceso de segmentación del registro, asignación de pesos

y contribución diferencial de los bloques en función del nivel de ruido.

Validez convergente. La estimación del audiograma tonal

En el ámbito del diagnóstico, los PEATC son herramientas fundamentales para evaluar la funcionalidad de la vía auditiva y detectar alteraciones estructurales, como lesiones en el nervio auditivo o el tronco cerebral (Barajas, 1985). Además, estos registros permiten estimar umbrales auditivos en poblaciones donde la audiometría convencional no es factible, como en recién nacidos o en pacientes con capacidad limitada de colaboración (Barajas et al., 1981).

En ambas situaciones el concepto de validez adquiere una gran importancia clínica. En el primero caso, resulta esencial determinar la capacidad de los PEATC para contribuir al diagnóstico de patologías como la neuropatía auditiva (Starr & Rance, 2015), tumores del ángulo pontocerebeloso, como los meningiomas (Barajas de Prat et al., 2007), o retrasos en la conducción nerviosa asociados a la desmielinización en esclerosis múltiple (Barajas, 1982). En el segundo caso, particularmente en los programas de cribado auditivo universal neonatal, la validez se refiere al grado en que los resultados obtenidos se correlacionan con la sensibilidad auditiva evaluada mediante métodos conductuales (Widen et al., 2000).

Existen diferentes enfoques para establecer la validez de criterio entre estas pruebas, incluyendo procedimientos analíticos como la sustracción de umbrales y los modelos de regresión, así como indicadores diagnósticos como la sensibilidad y la especificidad.

Método de las diferencias entre umbrales

Este método de calibración biológica emplea sujetos con audición normal como grupo control para establecer la relación entre respuestas electrofisiológicas y perceptuales. Para ello, se registran los umbrales auditivos con los mismos estímulos utilizados en la evaluación electrofisiológica, como clics o tonos puros por impulsos. A partir de estos datos, se define el nivel de referencia de 0 dB nHL, equivalente al promedio

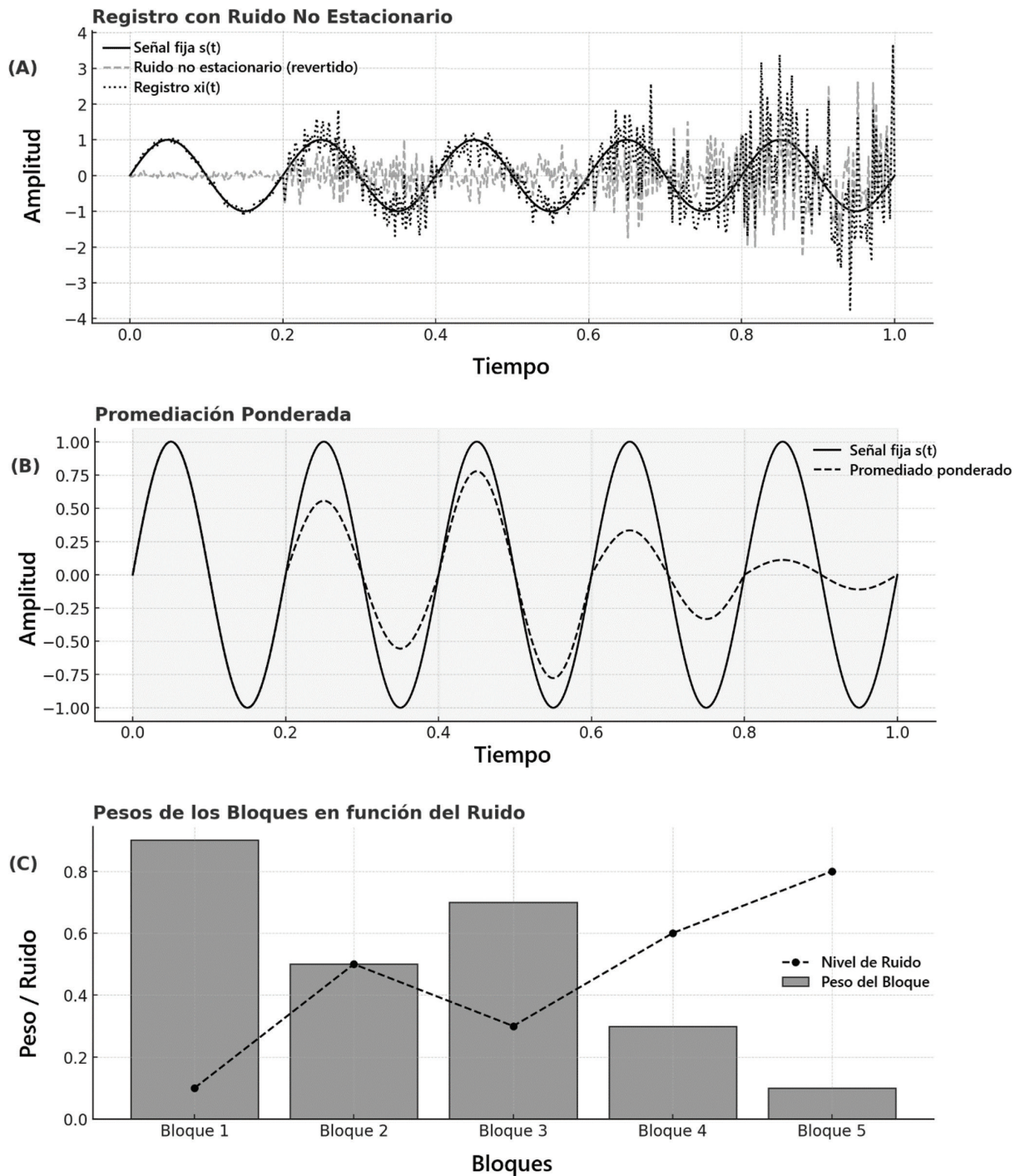


Figura 3. Representación esquemática del principio de promediación ponderada, basada en una simulación temporal de una respuesta ficticia. (A) Registro simulado con contaminación por ruido no estacionario, en el que una señal fija $s(t)$ se ve afectada por un componente de ruido $x_i(t)$ de intensidad variable. (B) Implementación de la técnica de promediación ponderada, en la que el registro se segmenta en bloques y se asignan coeficientes de ponderación en función del SNR de cada segmento. Los bloques con menor interferencia y mayor coherencia con la señal de interés contribuyen de manera más significativa al promedio final. (C) Distribución de los pesos asignados a los bloques en función del nivel de ruido, ilustrando la estrategia de optimización del aporte relativo de cada segmento en el proceso de estimación de la señal subyacente.

de los umbrales medidos en sujetos normoyentes (Bagatto et al., 2010; Gorga et al., 2006; Vander Werff et al., 2009). Por ejemplo, si el umbral conductual promedio para los estímulos de los PEATC-IT en un

grupo normoyente es de 30 dB HL a 500 Hz, este valor se establece como referencia y se considera igual a 0 dB nHL. Si la respuesta electrofisiológica mínima observable para esta frecuencia es de 25 dB nHL (5

dB por debajo del criterio de 30 dB nHL), se aplica una corrección de -5 dB para ajustar la estimación, alineando los resultados electrofisiológicos con los conductuales. Diversos estudios han cuantificado estos factores de corrección para distintas frecuencias, como se resume en la Tabla 1.

Este método destaca por su simplicidad y facilidad de implementación en entornos clínicos, ya que se basa en referencias constantes, eliminando la necesidad de análisis complejos. Además, su aplicación es ampliamente generalizable, permitiendo el uso de valores de referencia uniformes para cualquier

paciente, siempre que se mantengan la misma configuración de estímulos y calibración.

Sin embargo, este enfoque presenta limitaciones. La dependencia de un grupo normoyente asume una relación constante entre los umbrales electrofisiológicos y conductuales, lo cual no se cumple en individuos con pérdida auditiva. Además, en hipoacusias severas o profundas, la falta de sincronización neural y la reducción de amplitud en la respuesta pueden afectar la precisión de la estimación, incrementando el margen de error en la predicción (Chalak et al., 2013; Gorga et al., 2006; McCreery et al., 2015).

Tabla 1. Los valores en la tabla representan los factores de corrección aplicables a los resultados fisiológicos de los PEATC-IT para estimar los umbrales auditivos de la audiometría tonal en normoyentes, expresados en dB nHL. Las diferencias promedio se calcularon mediante el método de sustracción. Los datos corresponden a adultos, excepto los valores propuestos por Bagatto, obtenidos en población pediátrica.

Estudio	500 Hz	1000 Hz	2000 Hz	4000 HZ
Purdy et al., 1989	9,8	9,4	5,3	8,3
Stapells, 2000	20,4	--	13,4	11,8
Vander Werff et al., 2009	20,5	--	10,0	9,6
Bagatto et al., 2010	15	10	5	0

Modelos de regresión para la estimación de umbrales

Este enfoque emplea modelos de regresión lineal para estimar la relación entre los umbrales electrofisiológicos y los conductuales, permitiendo ajustar las diferencias sistemáticas entre ambos (Gorga et al., 2006; McCreery et al., 2015). En lugar de asumir una correspondencia fija, el modelo identifica patrones de variación según la frecuencia y el grado de pérdida auditiva, lo que facilita la aplicación de factores de corrección específicos.

El modelo está diseñado para identificar y cuantificar estas diferencias a través de una representación matemática que las relaciona con las medidas obtenidas en los PEATC. La regresión lineal se ajusta a los datos mediante ecuaciones específicas que reflejan cómo las diferencias cambian en función del umbral electrofisiológico observado. De esta forma, se logra una corrección más precisa que la obtenida con métodos que aplican factores constantes. La ecuación utilizada en este modelo es:

$$Y_{conductual} = X_{fisiológico} + (mx + b)$$

- Donde:
- $Y_{conductual}$: es el umbral conductual predicho, obtenido tras aplicar la corrección.
 - $X_{fisiológico}$: es el umbral fisiológico observado, que sirve como punto de referencia inicial.
 - x : es el valor del umbral fisiológico observado en el término $mx + b$, que influye en las diferencias predichas.
 - m : es la pendiente, que refleja cómo varían las diferencias entre los umbrales en función del umbral fisiológico (x).
 - b : es la intersección, que describe la diferencia promedio entre los umbrales cuando $x=0$.

Esta ecuación modela las discrepancias sistemáticas entre umbrales, permitiendo calcular el umbral conductual como la suma del umbral fisiológico observado $X_{fisiológico}$ y la corrección predicha ($mx + b$). Este enfoque resulta especialmente útil en la prescripción de audífonos y en el diagnóstico temprano de pérdidas auditivas significativas, ya que las correcciones adaptativas ayudan a evitar errores críticos en las estimaciones. Sin embargo, su precisión puede verse limitada en pérdidas auditivas profundas, donde la respuesta es menos robusta y más difícil

de modelar. Además, en neonatos y niños pequeños, pueden ser necesarios ajustes adicionales para mejorar la predicción. En este contexto, la obtención de medidas en oído real mediante sonda microfónica, y en particular la medición de la diferencia del oído real con el acoplador (Real Ear to Coupler Difference, RECD), es fundamental para reducir la variabilidad inter e intraindividual en los registros audiométricos a estas edades y optimizar la precisión de las estimaciones (Bagatto et al., 2005; Zenker Castro, 2011). La Tabla 2 resume los modelos de regresión propuestos por McCreery et al. (2015), así como los factores de

corrección derivados en función de la frecuencia y el nivel del umbral fisiológico.

Un aspecto importante de este enfoque es su capacidad para ajustarse al grado de pérdida auditiva del paciente. Estudios como el de McCreery et al. (2015) y Gorga et al. (2006) han demostrado que las diferencias entre los umbrales electrofisiológicos y conductuales no son lineales, especialmente en casos de pérdida auditiva significativa. Por ejemplo, se observó que los umbrales fisiológicos tienden a subestimar los umbrales conductuales en pacientes con hipoacusias severas, lo cual resalta la necesidad de aplicar correcciones dinámicas.

Tabla 2. Modelos de las diferencias entre los umbrales fisiológicos y conductuales, calculados a partir de ecuaciones de regresión lineal propuestos por McCreery et al (2015). Se presentan los factores de corrección por frecuencias asociados al nivel del umbral fisiológico.

Frecuencia (Hz)	Ecuación de regresión	Factores de corrección
500	$Y = -0.22x + 5.90$	20 dB: 5, 40 dB: -3, 60 dB: -7, 80 dB: -12
1000	$Y = -0.13x + 8.32$	20 dB: 5, 40 dB: 3, 60 dB: 0, 80 dB: -2
2000	$Y = -0.14x + 7.31$	20 dB: 5, 40 dB: 2, 60 dB: -1, 80 dB: -4
4000	$Y = -0.16x + 9.32$	20 dB: 6, 40 dB: 3, 60 dB: 0, 80 dB: -3

Validez de criterio. Sensibilidad y especificidad

La validez de criterio es un indicador de que las pruebas diagnósticas utilizadas identifican tanto la presencia como la ausencia de una pérdida auditiva. Este indicador adquiere especial relevancia en contextos clínicos como el cribado neonatal, donde una detección temprana es esencial para evitar repercusiones negativas en el desarrollo del lenguaje y la comunicación.

La sensibilidad mide la capacidad de una prueba para identificar correctamente a los individuos con pérdida auditiva (verdaderos positivos). Este indicador se calcula mediante la fórmula:

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \times 100$$

Donde:

VP (Verdaderos Positivos). Son los individuos con pérdida auditiva que han sido correctamente identificados por la prueba.

FN (Falsos Negativos). Son los individuos con pérdida auditiva que no han sido detectados por la prueba.

Una alta sensibilidad minimiza los falsos negativos, lo que es crucial en el cribado neonatal. Por ejemplo, si un protocolo de PEA tiene una sensibilidad del 95 %, significa que el 95 % de los pacientes con pérdida auditiva serán correctamente identificados, mientras que el 5 % restante podría clasificarse erróneamente como normoyentes.

La especificidad evalúa la capacidad de una prueba para identificar correctamente a los individuos con audición normal (verdaderos negativos). Una alta especificidad establece que pocas personas sin pérdida auditiva sean clasificadas incorrectamente como hipoacúsicas. Este indicador se calcula mediante la ecuación:

$$\text{Especificidad} = \frac{VN}{VN + FN} \times 100$$

Donde:

VN (Verdaderos Negativos). Son los individuos con audición normal que han sido correctamente identificados por la prueba.

FN (Falsos Positivos). Son los individuos con audición normal que fueron erróneamente diagnosticados con hipoacusia.

Zenker et al. (2013) analizaron el grado de concordancia entre tres centros que emitieron diagnósticos de hipoacusia basados en los PEATC. Los resultados evidenciaron una baja correlación en la clasificación de los casos entre los distintos centros, con un índice de correlación intraclase ($\rho = 0.37$), como se muestra en la Figura 4. Este valor indica una concordancia baja entre los evaluadores, lo que sugiere una alta variabilidad diagnóstica. En

términos clínicos, esto significa que la interpretación de los PEATC no fue consistente entre los centros, lo que podría llevar a diferencias en la identificación de la pérdida auditiva dependiendo del lugar donde se realice la prueba. En este estudio, la baja concordancia estuvo asociada con una ausencia de validación concurrente entre pruebas electrofisiológicas y conductuales, lo que derivó en una tasa elevada de errores diagnósticos.

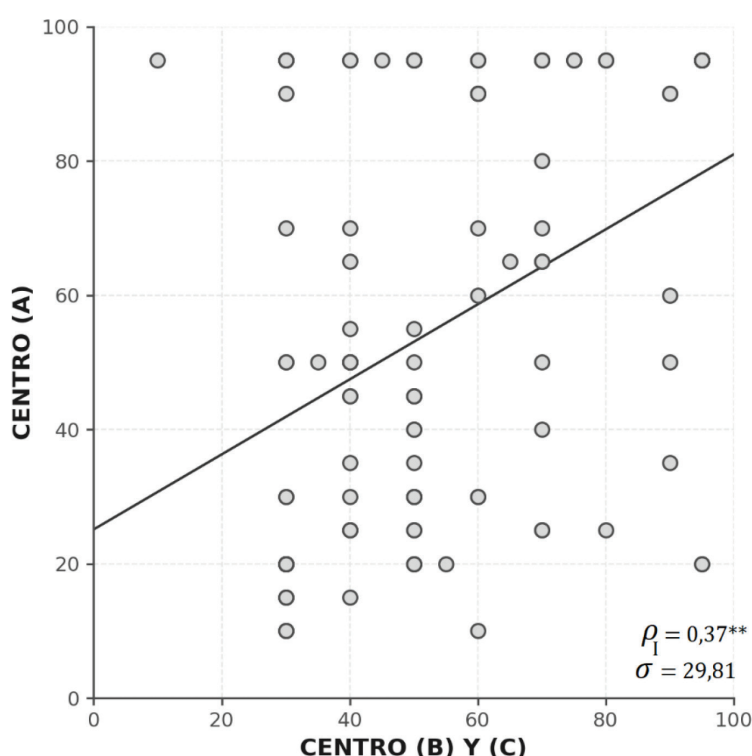


Figura 4. Representación gráfica de la Correlación Intraclase (CCI) en el diagnóstico de pérdida auditiva basado en los PEATC de tres centros. Se observa una elevada variabilidad entre los diagnósticos de los centros evaluados ($\sigma = 29,81$) y una baja concordancia ($\rho = 0.37$), lo que sugiere que el diagnóstico depende significativamente del centro de evaluación (Zenker et al., 2013).

Discusión

La evaluación de la fiabilidad y validez de los registros de PEATC permite determinar su precisión y confirmar su utilidad en el diagnóstico clínico. En este estudio, hemos analizado cómo distintos factores influyen en la calidad de los registros, incluyendo el SNR, la reproducibilidad de las mediciones y la precisión en la estimación de umbrales auditivos a partir de respuestas neurofisiológicas. La obtención de estos parámetros permite optimizar la metodología de adquisición e interpretación de los PEATC, minimizando la variabilidad y mejorando su precisión como herramienta diagnóstica.

En esta revisión se propone el uso de criterios estadísticos al uso, el SDR y el CCR, para evaluar la fiabilidad de los registros. Valores elevados indican registros estables y reproducibles, mientras que discrepancias entre ambos pueden reflejar la presencia de artefactos o ruido no estacionario. Estos indicadores, disponibles en la mayoría de los equipos comerciales, permiten una evaluación más objetiva de la calidad del registro y contribuyen a una interpretación más precisa.

Dado que la fiabilidad es un prerrequisito para la validez, la implementación de criterios estadísticos que evalúen el resultado de los trazados no solo optimiza la consistencia, sino que también fortalece

su capacidad diagnóstica. Al minimizar la variabilidad intra e interobservador, estos indicadores permiten que la estimación de umbrales auditivos a partir de respuestas neurofisiológicas sea más precisa y reproducible. Esto es especialmente relevante en contextos clínicos donde la interpretación subjetiva sigue desempeñando un papel clave, ya que la combinación de estos parámetros con modelos predictivos basados en la regresión lineal o logística mejoran la objetividad de los PEATC, reduciendo la dependencia de la experiencia del examinador y aumentando su validez.

La optimización del procedimiento de registro es un aspecto clave en la mejora de la fiabilidad. Se ha demostrado que las técnicas de promediación ponderada son más efectivas que los métodos tradicionales, ya que consideran la naturaleza no estacionaria del ruido y asignan mayor peso a los segmentos con mejor SNR. Este enfoque no solo mejora la detectabilidad de las respuestas, sino que también reduce el tiempo total de registro sin comprometer la precisión diagnóstica. En entornos clínicos pediátricos, esta reducción del tiempo es especialmente relevante, ya que permite obtener mejores registros, facilitando el estudio de la población poco colaboradora.

Desde una perspectiva clínica, la validez de criterio de los PEATC es un factor esencial para su aplicación en el diagnóstico de la hipoacusia. La sensibilidad y especificidad de estos registros dependen de la precisión con la que pueden estimar los umbrales auditivos en comparación con la audiometría tonal conductual. En este estudio, hemos analizado el uso de modelos de regresión lineal y el método de las diferencias entre umbrales, concluyendo que estos permiten ajustar las estimaciones a partir de respuestas electrofisiológicas y minimizar errores sistemáticos.

Finalmente, aunque este estudio ha identificado múltiples estrategias para mejorar la fiabilidad y validez de los PEATC, existen algunas limitaciones que deben ser consideradas. La variabilidad interlaboratorios en la configuración del equipo y los parámetros de registro puede afectar la comparación de los resultados. Además, la interpretación de los PEATC sigue dependiendo en gran medida de la experiencia del examinador, lo que resalta la necesidad de desarrollar protocolos estandarizados que incorporen criterios estadísticos objetivos. Futuras investigaciones deberían centrarse en la integración de nuevas tecnologías, como la inteligencia artificial y el procesamiento avanzado de señales, para optimizar la calidad del

registro y proporcionar herramientas objetivas que complementen la interpretación clínica, mejorando así la precisión diagnóstica.

Conclusiones

La incorporación de indicadores objetivos de calidad en la interpretación de los PEATC es clave para su uso clínico. La fiabilidad y la validez son dos de los más relevantes, ya que garantizan registros reproducibles y mediciones precisas de los umbrales auditivos. La fiabilidad, evaluada mediante SDR y CCR, permite identificar registros estables y minimizar la variabilidad, mientras que la validez asegura que los umbrales estimados reflejen con precisión la realidad audiológica del paciente.

Más allá de la aplicación de protocolos estandarizados, el uso de criterios cuantitativos mejora la interpretación clínica al reducir la influencia de artefactos y la subjetividad del examinador. Además, su implementación contribuye a disminuir la variabilidad interlaboratorios, facilitando la comparabilidad de resultados entre diferentes centros y optimizando la fiabilidad de los PEATC como herramienta diagnóstica. La integración de estos indicadores en la práctica clínica optimiza el diagnóstico de la pérdida auditiva y facilita la toma de decisiones, especialmente en poblaciones en las que estos registros constituyen la única evidencia de la sensibilidad auditiva.

Referencias

- Bagatto, M., Moodie, S., Scollie, S., Seewald, R., Moodie, S., Pumford, J., & Liu, K. P. R. (2005). Clinical Protocols for Hearing Instrument Fitting in the Desired Sensation Level Method. *Trends in Amplification*, 9(4), 199–226. <https://doi.org/10.1177/108471380500900404>
- Bagatto, M., Scollie, S. D., Hyde, M., & Seewald, R. (2010). Protocol for the provision of amplification within the Ontario Infant hearing program. *International Journal of Audiology*, 49(SUPPL. 1), 70–80. <https://doi.org/10.3109/14992020903080751>
- Barajas, J. J. (1982). Evaluation of ipsilateral and contralateral brainstem auditory evoked potentials in multiple sclerosis patients. *Journal of the Neurological Sciences*, 54(1), 69–78. [https://doi.org/10.1016/0022-510X\(82\)90219-2](https://doi.org/10.1016/0022-510X(82)90219-2)
- Barajas, J. J. (1985). Brainstem Response Audiometry as Subjective and Objective Test for Neurological Diagnosis. *Scandinavian Audiology*, 14(1), 57–62. <https://doi.org/10.3109/01050398509045923>

- Barajas, J. J., Olaizola, F., Tapia, M. C., Alarcon, J. L., & Alaminos, D. (1981). Audiometric Study of the Neonate: Impedance Audiometry. Behavioural Responses and Brain Stem Audiometry. *International Journal of Audiology*, 20(1), 41-52. <https://doi.org/10.3109/00206098109072681>
- Barajas de Prat, J., Zenker Castro, F., & Fernández Belda, R. (2007). Potenciales Evocados Auditivos. In C. Suárez & L. M. Gil Carcedo (Eds.), *Tratado de Otorrinolaringología y Patología de Cabeza y Cuello* (Vol. 6, Issue 5, pp. 1133 - 1155). Editorial Médica Panamericana, S.A.
- Berninger, E., Olofsson, Å., & Leijon, A. (2014). Analysis of click-evoked auditory brainstem responses using time domain cross-correlations between interleaved responses. *Ear and Hearing*, 35(3), 318-329. <https://doi.org/10.1097/01.aud.0000441035.40169.f2>
- Chalak, S., Kale, A., Deshpande, V. K., & Biswas, D. A. (2013). Establishment of normative data for monaural recordings of auditory brainstem response and its application in screening patients with hearing loss: A cohort study. *Journal of Clinical and Diagnostic Research*, 7(12), 2677-2679. <https://doi.org/10.7860/JCDR/2013/6768.3730>
- Cone, B., & Norrix, L. W. (2015). Measuring the advantage of kalman-weighted averaging for auditory brainstem response hearing evaluation in infants. *American Journal of Audiology*, 24(2), 153-168. https://doi.org/10.1044/2015_AJA-14-0021
- Davila, C. E., & Mobin, M. S. (1992). Weighted Averaging of Evoked Potentials. *IEEE Transactions on Biomedical Engineering*, 39(4), 338-345. <https://doi.org/10.1109/10.126606>
- Dawson, G. D. (1954). A summation technique for the detection of small evoked potentials. *Electroencephalography and Clinical Neurophysiology*, 6(C), 65-84. [https://doi.org/10.1016/0013-4694\(54\)90007-3](https://doi.org/10.1016/0013-4694(54)90007-3)
- Delgado Hernández, J., Zenker Castro, F., & Barajas de Prat, J. J. (2003). Normalización de los Potenciales Evocados Auditivos del Tronco Cerebral I: Resultados en una muestra de adultos normoyentes. *Auditio: Revista Electrónica de Audiología*, 2(1), 13-18. <https://doi.org/10.51445/sja.auditio.vol2.2003.0020>
- Doyle, D. J., & Hyde, M. L. (1981). Analogue and digital filtering of auditory brainstem responses. *Scandinavian Audiology*, 10(2), 81-89. <https://doi.org/10.3109/01050398109076166>
- Elberling, C. (1979). Letter to the editor: Auditory electrophysiology: Spectral analysis of cochlear and brain stem evoked potentials. *Scandinavian Audiology*, 8(1), 57-64. <https://doi.org/10.3109/01050397909076302>
- Gorga, M. P., Johnson, T. A., Kaminski, J. K., Beauchaine, K. L., Garner, C. A., & Neely, S. T. (2006). Using a combination of click-and toneburst-evoked auditory brainstem response measurements to estimate pure-tone thresholds. *Ear Hear*, 27(1), 60-74. <https://doi.org/10.1097/01.aud.0000194511.14740.9c>
- Hall, J. W. I. (2007). Improving the Signal-to-Noise Ratio (SNR). In *New Handbook of Auditory Evoked Responses* (1st ed., pp. 95-96). Pearson Education.
- Hoke, M., Ross, B., Wickesberg, R., & Lütkenhöner, B. (1984). Weighted averaging - theory and application to electric response audiometry. *Electroencephalography and Clinical Neurophysiology*, 57(5), 484-489. [https://doi.org/10.1016/0013-4694\(84\)90078-6](https://doi.org/10.1016/0013-4694(84)90078-6)
- Kumaragamage, C. L., Lithgow, B. J., & Moussavi, Z. K. (2016). Investigation of a new weighted averaging method to improve SNR of electrocochleography recordings. *IEEE Transactions on Biomedical Engineering*, 63(2), 340-347. <https://doi.org/10.1109/TBME.2015.2457412>
- Madsen, S. M. K., Harte, J. M., Elberling, C., & Dau, T. (2018). Accuracy of averaged auditory brainstem response amplitude and latency estimates. *International Journal of Audiology*, 57(5), 345-353. <https://doi.org/10.1080/14992027.2017.1381770>
- McCreery, R. W., Kaminski, J., Beauchaine, K., Lenzen, N., Simms, K., & Gorga, M. P. (2015). The Impact of Degree of Hearing Loss on Auditory Brainstem Response Predictions of Behavioral Thresholds. *Ear and Hearing*, 36(3), 309-319. <https://doi.org/10.1097/AUD.0000000000000120>
- McKearney, R. M., Bell, S. L., Chesnaye, M. A., & Simpson, D. M. (2023). Optimising weighted averaging for auditory brainstem response detection. *Biomedical Signal Processing and Control*, 83(August 2022), 104676. <https://doi.org/10.1016/j.bspc.2023.104676>
- Mühler, R., & von Specht, H. (1997). Reduction of Background Noise in Human Auditory Brainstem Response by Means of Classified Averaging. In *Acoustical Signal Processing in the Central Auditory System* (pp. 599-604). Springer US. https://doi.org/10.1007/978-1-4419-8712-9_56
- Norton, S. J., Gorga, M. P., Widen, J. E., Folsom, R. C., Sininger, Y., Cone-Wesson, B., Vohr, B. R., Mascher, K., & Fletcher, K. (2000). Identification of Neonatal Hearing Impairment: Evaluation of transient evoked otoacoustic emission, distortion product otoacoustic emission, and auditory brain stem response test performance. *Ear and Hearing*, 21(5), 508-528. <https://doi.org/10.1097/00003446-200010000-00013>
- Núñez Batalla, F., Jáudenes Casaubón, C., Sequí Canet, J. M., Vivanco Allende, A., Zubicaray Ugarteche, J., & Olleta Lascarro, I. (2020). Actualización de los programas de detección precoz de la sordera infantil: recomendaciones CODEPEH 2019 (Niveles 2, 3 y 4: diagnóstico, tratamiento y seguimiento) | Update in early detection of pediatric hearing loss: 2019 CODEPEH recommendations. *Revista Española de Discapacidad*, 8(1), 219-246. <https://doi.org/10.5569/2340-5104.08.01.13>
- Picton, T., Linden, R., Hamel, G., & Maru, J. (1983). Aspects of Averaging. *Seminars in Hearing*, 4(04), 327-340. <https://doi.org/10.1055/s-0028-1094195>
- Purdy, S. C., Houghton, J. M., Keith, W. J., & Greville, K. A. (1989). Frequency-Specific Auditory Brainstem Responses: Effective Masking Levels and Relationship to Behavioural Thresholds in Normal Hearing Adults. *International Journal of Audiology*, 28(2), 82-91. <https://doi.org/10.3109/00206098909081613>

- Schimmel, H. (1967). The (+) reference: accuracy of estimated mean components in average response studies. *Science*, 157(784), 92–94. <https://doi.org/10.1126/science.157.3784.92>
- Stapells, D. R. (2000). Threshold Estimation by the Tone-Evoked Auditory Brainstem Response: A Literature Meta-Analysis. *Journal of Speech Language Pathology and Audiology*, 24(2), 74–83.
- Starr, A., & Rance, G. (2015). Auditory neuropathy. In *Handbook of Clinical Neurology* (1st ed., Vol. 129). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-62630-1.00028-7>
- Vander Werff, K. R., Prieve, B. A., & Georgantas, L. M. (2009). Infant Air and Bone Conduction Tone Burst Auditory Brain Stem Responses for Classification of Hearing Loss and the Relationship to Behavioral Thresholds. *Ear & Hearing*, 30(3), 350–368. <https://doi.org/10.1097/AUD.0b013e31819f3145>
- Wang, H., Li, B., Lu, Y., Han, K., Sheng, H., Zhou, J., Qi, Y., Wang, X., Huang, Z., Song, L., & Hua, Y. (2021). Real-time threshold determination of auditory brainstem responses by cross-correlation analysis. *IScience*, 24(11), 103285. <https://doi.org/10.1016/j.isci.2021.103285>
- Widen, J. E., Folsom, R. C., Cone-Wesson, B., Carty, L., Dunnell, J. J., Koebse, K., Levi, A., Mancl, L., Ohlrich, B., Trouba, S., Gorga, M. P., Sininger, Y. S., Vohr, B. R., & Norton, S. J. (2000). Identification of Neonatal Hearing Impairment: Hearing status at 8 to 12 months corrected age using a visual reinforcement audiometry protocol. *Ear and Hearing*, 21(5), 471–487. <https://doi.org/10.1097/00003446-200010000-00011>
- Widen, J. E., Johnson, J. L., White, K. R., Gravel, J. S., Vohr, B. R., James, M., Kennalley, T., Maxon, A. B., Spivak, L., Sullivan-Mahoney, M., Weirather, Y., & Meyer, S. (2005). A Multisite Study to Examine the Efficacy of the Otoacoustic Emission/Automated Auditory Brainstem Response Newborn Hearing Screening Protocol. *American Journal of Audiology*, 14(2), S205–S216. [https://doi.org/10.1044/1059-0889\(2005/022\)](https://doi.org/10.1044/1059-0889(2005/022))
- Wong, P. K., & Bickford, R. G. (1980). Brain stem auditory evoked potentials: The use of noise estimate. *Electroencephalography and Clinical Neurophysiology*, 50(1–2), 25–34. [https://doi.org/10.1016/0013-4694\(80\)90320-x](https://doi.org/10.1016/0013-4694(80)90320-x)
- Zenker Castro, F. (2011). Controversias en el diagnóstico y adaptación audiológica en pérdidas auditivas severas en niños. *Audiol Hoy*, 8(1), 65–71.
- Zenker Castro, F., Estrada Alonso, M. M., Larumbe Zabala, E., & Barajas De Prat, J. J. (2013). Fiabilidad diagnóstica de los potenciales evocados auditivos del tronco cerebral en una muestra de pacientes en edad infantil. *Actas Del IX Congreso de La Asociación Española de Audiología*, 153–159.

Conflicto de interés

Los autores declaran no tener ningún conflicto de interés.

Contribución de los autores

FZ: conceptualización, borrador, revisión, edición del manuscrito, supervisión y visualización. JJB: conceptualización, revisión y edición del manuscrito.

Financiación

Esta investigación no ha recibido financiación externa.

Agradecimientos

Dedicado a la memoria de D. Manuel Expósito, profesional de referencia en la obtención de Potenciales Evocados Auditivos en la Clínica Barajas, quien formó al primer autor en esta técnica y fue responsable de realizar los registros del primer programa de cribado auditivo neonatal de España, llevado a cabo en 1988 en Santa Cruz de Tenerife.

Declaración del uso de Inteligencia Artificial

Durante la preparación de este manuscrito, los autores utilizaron ChatGPT 4o para mejorar la claridad, precisión del lenguaje y elaboración de las figuras. Posteriormente, revisaron y ajustaron el contenido, asumiendo plena responsabilidad por la versión final.

Oficina Editorial

Corrección: Rita López

Producción: GlauX Publicaciones Académicas