

Validez y Confiabilidad en Investigación Odontológica

Validity and Reliability in Dental Research

Pedro Christian Aravena^{*,*}; Javier Moraga^{**,***}; Ricardo Cartes-Velásquez^{****} & Carlos Manterola^{****}

ARAVENA, P. C.; MORAGA, J.; CARTES-VELÁSQUEZ, R. & MANTEROLA, C. Validez y confiabilidad en investigación odontológica. *Int. J. Odontostomat.*, 8(1):69-75, 2014.

RESUMEN: El proceso de medición consiste en asignar valores numéricos a características del paciente, un nivel de daño, certeza diagnóstica de la enfermedad, entre otros. Sin embargo, existen fenómenos complejos que incluyen conceptos más amplios que requieren de la creación y validación de instrumentos psicométricos para su registro. El objetivo de esta revisión es presentar los conceptos de validez y confiabilidad de instrumentos de medición aplicados en Odontología. Se expone la definición en salud de variables directas e indirectas, la importancia del uso de instrumentos psicométricos, la definición de Validez y Confiabilidad de un instrumento y sus distintas clasificaciones (validez de contenido, de criterio y de constructo; confiabilidad del instrumento y del observador). Además, se presenta algunos ejemplos del proceso de validación de instrumentos, su metodología y forma de análisis de los resultados. De esta forma, presentamos los tópicos necesarios para comprender el proceso de medición usado en nuestro quehacer clínico o en investigación odontológica, con el fin de identificar los instrumentos adecuados (válidos y confiables) para nuestras necesidades.

PALABRAS CLAVE: validez, confiabilidad, reproductibilidad de resultados, investigación, odontología basada en evidencia, odontología.

INTRODUCCIÓN

La medición se define como proceso de asignar valores a ciertos eventos de la realidad (Kerlinger, 2002). Es un componente esencial en la investigación científica de las diferentes disciplinas ya sea en el área de las ciencias naturales, sociales o de la salud.

En ciencias médicas, la forma de medir un evento de interés se puede realizar de dos maneras: i) a través de medición directa: aquella que podemos comparar directamente con un patrón de referencia (como medir la longitud del saco periodontal con una sonda); ii) o medición indirecta: medidas que no se pueden realizar por comparación directa con el patrón o referencia, sino que deriva de análisis cualitativo y cuantitativo de los de los datos (medir grado de dolor).

Sin embargo, no siempre sabremos con tanta facilidad como medir una variable, hay fenómenos complejos que incluyen conceptos mucho más am-

plios como “conocimiento”, “interés”, “calidad de vida”, “ansiedad preoperatoria”, que requieren mediciones psicométricas (asignar un valor numérico a una característica psicológica de una persona) para lo cual se han diseñado, o se pueden diseñar escalas de medición o cuestionarios. En ello, las primeras áreas en lidiar con este tipo de problemas de la medición fueron psicología y educación intentando generar metodologías y herramientas para evaluar estados subjetivos como la inteligencia y el comportamiento a través del diseño de instrumentos psicométricos (Goldstein & Lersen, 2000), capaces de registrar en un constructo de situaciones un evento de interés en particular.

En las ciencias odontológicas dichos instrumentos son de uso habitual y donde la medición es inherente a la práctica diaria, como por ejemplo la sonda OMS en periodoncia (WHO, 1978); la clasificación de

* Facultad de Medicina, Universidad Austral de Chile, Valdivia, Chile.

** Programas de Magister y Doctorado en Ciencias Médicas, Facultad de Medicina, Universidad de Frontera, Temuco, Chile.

*** Universidad Autónoma de Chile, Santiago, Chile.

**** Facultad de Odontología, Universidad de Concepción, Concepción, Chile.

***** Centro de Investigación en Ciencias Biomédicas, Universidad Autónoma de Chile, Temuco, Chile.

posición de terceros molares descritas por Pell y Gregory o de Winter en cirugía bucal (Almendros-Marqués *et al.*, 2008), o el localizador eléctrico de ápice en endodoncia (Kobayashi, 1995). La pregunta es: ¿estos instrumentos miden lo que realmente se quiere medir? Pues para ello, existen los estudios de validación de escalas que permiten responder estas preguntas.

Todo aquel instrumento que se utilice para cuantificar un nivel de daño o de certeza diagnóstica debe basarse en este tipo de estudios, de lo contrario, estaremos usando un aparato que no cuenta con la certificación idónea para su uso y que, a la larga, puede traer consecuencias severas si se está realizando un mal diagnóstico. Por otra parte, no se debe olvidar que los juicios subjetivos como “a mí me parece”, “el buen criterio indica”, “a mí me ha ido bien”, juegan un papel anecdótico en el proceso de la investigación y deberían desaparecer de la práctica habitual de la medicina (Manterola, 2002).

Este proceso de medición lleva implícito los conceptos de validez y confiabilidad (Tabla I) los cuales otorgan a la medición un peso técnico y objetivo. Es por esto que nos parece importante aclarar estos conceptos con el fin de aportar al investigador y al usuario de literatura científica odontológica herramientas adecuadas para su correcto uso e interpretación con el fin de evitar falsas estimaciones, y por ende, sesgos de medición (Nunnally & Berstein, 1999; Zavando *et al.*, 2010).

Concepto de Validez y Confiabilidad

Validez. También denominada exactitud, corresponde al grado en que una medición refleja la realidad de un fenómeno o capacidad de medición o clasificación de un método o instrumento para aquello que fue propuesto o sea que mida o clasifique lo que efectivamente analizamos y no otra cosa (Manterola; Zavando *et al.*; Alarcón & Muñoz, 2008). Un ejemplo de esto es

la necesidad de registrar la calidad de vida que han de tener nuestros pacientes relacionados a la condición de salud oral que éstos presentan. Para ello, se han desarrollado escalas de medición de calidad de vida y que han sido traducidos y usados en países como China, Francia, Alemania, Japón, Somalia, entre otros; sin embargo, el uso de dicho instrumento en la población chilena requiere de un proceso de traducción y de adaptación (López & Baelum, 2006), por lo cual podemos esperar diferencias atribuibles a las discrepancias entre las características genéticas, sociales, económicas, etc., de la población de origen y la nuestra. Esto lleva a la necesidad de generar instrumentos propios o validar los ya existentes para así medir de una forma correcta el fenómeno en nuestro entorno.

Un ejemplo del proceso de validación podemos demostrarlo con un ejercicio del tiro al blanco (Fig. 1). Un rifle bien calibrado puede acertar al blanco varias veces si se apunta correctamente (válido y confiable); Sin embargo, existe la posibilidad que se apunte correctamente, pero todas las balas den en un mismo punto fuera del blanco; esto quiere decir que el instrumento esta mal calibrado. Ahora bien, si se insiste en

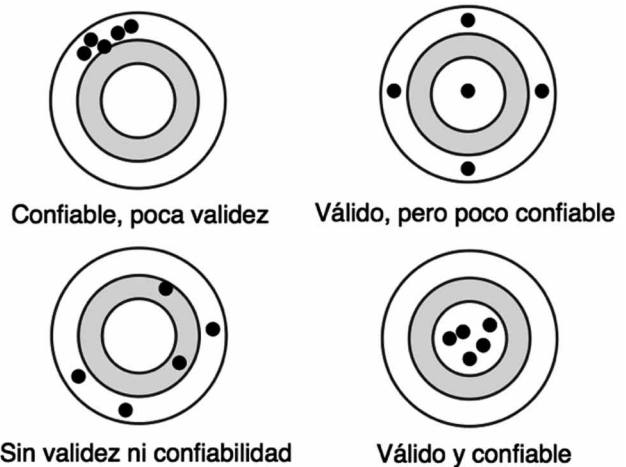


Fig. 1. Diferencia entre validez y confiabilidad. Adaptado de Hulley *et al.* (2007).

Tabla I. Concepto de Validez y confiabilidad de las mediciones en Salud. Adaptado de Hulley *et al.* (2007).

	Validez	Confiabilidad
Definición	Grado en que una variable representa realmente aquello para lo que está destinada	Grado en que una variable otorga casi el mismo valor cuando se mide varias veces
Mejor modo de evaluación	Comparación con un estándar de referencia	Comparación entre medidas repetidas
Amenazado por	Error sistemático (sesgo) al que contribuyen:	Error aleatorio (azar) al que contribuyen:
	Observador	Observador
	Participante	Participante
	Instrumento	Instrumento

apuntar al blanco pero no se logra marcar en el mismo punto repetidamente, es probable que el rifle esté mal calibrado o que no se tenga la capacidad de apuntar correctamente (el instrumento no es válido ni confiable) (Hulley *et al.*, 2007).

Según la American Educational Research Association & National Council on Measurement in Education (1999) el concepto de validez ha presentado una evolución desde 1954 hasta 1999. En una primera fase se diferencian 4 tipos de validez: de contenido, predictiva, concurrente y de constructo (APA, año 1954), que posteriormente se reducen a 3 (APA, año 1974), fusionándose la validez predictiva y concurrente a validez de criterio. En la segunda etapa, más teórica, la validez de constructo se vuelve preponderante ya que esta considera además la interacción con otras variables, en esta etapa aumentan las investigaciones limitadas sólo a este tipo de validez. La fase actual o tercera etapa que corresponde a una fase contextual, se amplía la fase anterior agregando el “uso propuesto” lo que dota de límites éticos a los métodos o instrumentos desarrollados.

A continuación se dará una descripción más detallada de cada uno de los tipos de validez mencionados anteriormente (Tabla II).

Validez de contenido. Este tipo de validez se refiere al grado en que el instrumento representa la totalidad del fenómeno que se pretende medir. Este concepto pretende que la escala presente las siguientes características: Induzca a un mínimo de respuestas sesgadas, sea fácil de entender e interpretar, sea fácil de administrar o aplicar en el trabajo de campo y que posea la capacidad de discriminar (Devlin *et al.*, 1993).

Este tipo de validez puede asociarse con la llamada validez de fachada, pues ésta se relaciona con las características de calidad aparente del instrumento. En ella no se realizan análisis empíricos, sino que se corresponde al juicio subjetivo de un panel de expertos en el área, que garantiza que la medición tiene sentido y es razonable (Pesudovs *et al.*, 2007). Se con-

sidera como la forma más débil de validez, pues en ella no se puede realizar un análisis profundo y exhaustivo de los ítems que deben ser excluidos o incluidos en el instrumento final a utilizar. Se obtiene mediante una inspección por expertos, revisión de la literatura y presentación mediante pilotos a la población blanco o análisis cualitativo de las respuesta de los revisores.

Validez de criterio: Pretende correlacionar las medidas de la escala con otra medición del atributo estudiado, idealmente un estándar de referencia o patrón de oro. Como se menciono anteriormente, este tipo de validez contiene a la validez concurrente y predictiva. En relación con la validez concurrente podemos decir que se correlaciona la nueva escala con un criterio aceptado o estándar de referencia, para lo cual se realiza una medición al mismo tiempo con ambos instrumentos y se analiza su concordancia. En el caso de la validez predictiva se mide o clasifica esperando la ocurrencia de un futuro evento con el cual se puedan establecer asociaciones. En el caso de que no exista un patrón de oro con el cual correlacionar, no se contempla este tipo de validez para el estudio de validación de la escala (Streiner & Geoffrey, 2003).

Validez de constructo: Corresponde al principal tipo de validez. Pretende hacer referencia a la obtención de evidencia empírica que garantice la existencia del constructo en cuestión. Su determinación puede ser realizada estadísticamente mediante: grupos extremos, matriz multirrasgo-multimétodo y análisis factorial. En la determinación de validez de constructo por grupos extremos el comportamiento de la escala se evalúa con sujetos con y sin el evento de interés, de esta forma el grupo que presenta el atributo obtendrá un alto nivel en la escala estudiada, mientras que el grupo que no posee el atributo puntuará muy bajo o cercano a 0; en el caso de la matriz multirrasgo-multimétodo se deben seguir 3 pasos fundamentales: 1) Se seleccionan dos o más formas de medir el constructo, 2) Dos o más constructos que puedan ser medidos por los mismos métodos y 3) Se mide en la muestra todos los constructos con todos los métodos. De los pasos anteriores se obtienen 3 coeficientes:

Tabla II. Tipos de validación, método y forma de análisis de sus resultados. Adaptado de Sánchez & Echeverry (2004).

Validación	Metodología	Análisis de resultados
de Contenido	Revisión de la literatura, opinión de expertos y de la población a medir.	Análisis cualitativo, forma y entendimiento del instrumento.
de Criterio	Comparación con patrón de referencia (estándar de referencia).	Correlación estadística (Coeficiente correlación de Pearson).
de Constructo	Asignación numérica de valores extremos, análisis de variabilidad de respuesta y criterios clínicos.	Análisis factorial; Ecuaciones de regresión.

Coeficiente de fiabilidad: Correlaciones entre las medidas de un solo constructo con el mismo método, estos valores deben ser elevados.

Coeficiente de validez convergente: Correlaciones de un solo constructo con métodos diferentes, estos valores también deberían ser elevados.

Coeficiente de validez discriminante: Correlaciones entre las medidas de diferentes constructos con el mismo método (coeficientes heterorrasgo-monométodo) o de diferentes constructos con diferentes métodos (coeficiente heterorrasgo-multimétodo), estos valores deberían ser los menores de los tres.

Otra forma de determinación es a través del análisis factorial, que corresponde a un análisis multivariado, el cual permite reducir datos y además agruparlos en conjuntos homogéneos de variables que muestran un elevado índice de correlación entre sí.

Para entender mejor el proceso de validación de un instrumento de medición haremos referencia a trabajo realizado en Chile por López & Baelum cuyo objetivo fue adaptar y validar culturalmente al español el cuestionario de medición de impacto de salud bucal (Oral Health Impact Profile; OHIP-sp) con el propósito es medir la percepción subjetiva de la disfunción, el malestar y la discapacidad atribuidas a condiciones orales de la persona, la cual cuenta con 7 áreas de salud oral a evaluar y un rango de puntaje de respuesta de 0 (buena salud oral) a 196 puntos (mala salud oral) (Slade & Spencer, 1994). Para ello, se inició el proceso de validación con la traducción al español y revisión por dos dentistas independientes, cuyo fin fue determinar la comprensión y semántica del instrumento (validez de fachada). Luego, aplicaron el cuestionario a 9133 estudiantes de entre 12 a 21 años de la ciudad de Santiago de Chile. A ellos, se les solicitó exponer sus conductas relacionadas con a salud bucal y la auto-

percepción subjetiva del estado de salud oral (buena, regular o mala). A su vez, recibieron un examen clínico del cual se obtuvo el nivel de inserción clínica en molares e incisivos, pérdida de dientes y presencia de lesiones gingivales ulcerativa necrotizante.

En este ejemplo, para obtener la validez de constructo, los investigadores asociaron la auto-percepción del estado de salud oral con el puntaje total obtenido del cuestionario OHIP-sp (buena, regular y mala) mediante análisis estadístico (correlación *r* de Spearman). Para lograr la validez de tipo discriminante, compararon los resultados clínicos de los sujetos que presentaron mejores y peores condiciones clínicas de salud oral con los puntajes mas bajos y más altos obtenido en los cuestionarios, respectivamente; discriminando mediante análisis estadístico los puntajes de corte que podrian establecer el estado de salud oral real obtenido mediante el cuestionario. De esta forma, los investigadores comprobarían la hipótesis de que los sujetos con pobres resultados de salud oral tendrían una mayor puntuación en el instrumento OHIP-Sp (López & Baelum, 2006).

Confiabilidad. La confiabilidad, también denominada precisión, corresponde al grado con que los puntajes de una medición se encuentran libres de error de medida. Es decir, al repetir la medición en condiciones constantes estas deberían ser similares. Este concepto se relaciona con la estabilidad del instrumento en sí mismo, independiente del individuo quien lo aplique (observador) y del momento en que es aplicado (tiempo). Tabla III.

Confiabilidad del instrumento: La confiabilidad de los componentes del instrumento se obtiene mediante la correlación que presentan sus ítems entre sí mismos y el concepto para el cual fue creado. Ejemplo de esta sentencia es la medición de la calidad de vida de los pacientes luego de una cirugía oral (Shugars *et al.*, 1996) ¿Qué elementos son importantes a consi-

Tabla III. Tipos de análisis de confiabilidad, metodo y análisis de sus resultados. Adaptado de Sánchez & Echeverry (2004).

Confiabilidad	Metodología	Análisis de resultados
del Instrumento	Ajustar el número de ítems, dominios y establecer su consistencia Interna	Prueba KR20 en ítems dicotómicos; Alpha de Cronbach en aquellos ítems con más de dos opciones de respuesta
según Tiempo	Prueba test-retest	Análisis de respuesta de sujetos: Correlación de Pearson Análisis de respuesta de sujetos, característica de población y riesgo de sesgo: Correlación Intraclase
según Observador	Interobservador Intraobservador	Correlación de Pearson Índice Kappa

derar? Se deben incluir todos los constructos que definen el término de "calidad de vida" como son la función oral (poder masticar, hablar, comer); función doméstica (dormir bien, comunicarse, asistir al trabajo, tener vida social); dolor y otros síntomas (inflamación, sangrado, hematoma, náuseas, entre otras). ¿Cómo se eligen los ítems que incluye un cuestionario o escala? Esto se logra mediante el análisis de conglomeración y de representación de las preguntas (ítems) de cada una de las ideas más globales (dominios) mediante análisis estadísticos, como son la correlación de Pearson y de Consistencia Interna mediante el Alpha de Cronbach.

Entre los tipos de confiabilidad del instrumento podemos nombrar:

Confiabilidad interinstrumento: Correlación de los resultados generados por diferentes instrumentos en mediciones repetidas en un grupo de sujetos.

Confiabilidad intrainstrumento: Correlación de los resultados obtenidos por un mismo instrumento realizada en mediciones repetidas en un grupo de sujetos. Confiabilidad del observador: Establece la correcta medición obtenida del instrumento asociado al número de observadores (investigadores o encuestadores) que recopilan o responden el instrumento aplicado.

Confiabilidad interobservador: Se evalúa correlacionando los puntajes obtenidos por distintos observadores en un mismo periodo de tiempo, indicando de esta forma el grado de acuerdo respecto de la medida observada corrigiendo por el azar. Esto da cuenta del grado de reproducibilidad del instrumento al ser aplicado por diferentes sujetos (Alarcón & Muñoz; Streiner & Geoffrey).

Confiabilidad intraobservador: Se evalúa correlacionando los puntajes obtenidos por el mismo observador en dos momentos de tiempo, da cuenta de la estabilidad temporal que presenta el instrumento. Este tipo de confiabilidad puede verse afectada no solo por el azar, sino también por las condiciones del sujeto en estudio (Streiner & Geoffrey).

Confiabilidad según el tiempo (test-retest): Otorga la seguridad al instrumento, que al ser aplicada en momentos distintos en una misma condición, debería otorgar los mismos resultados. ¿Cómo se obtiene? Se aplica a la misma muestra la escala en 2 momentos distintos, distanciados en un periodo de tiempo, con el fin de que el evaluador o evaluado genere un sesgo en la

respuesta. Estos resultados se comparan mediante el test de correlación de Pearson o el test de Correlación Interclase (CCI), el primero asume que toda la variabilidad de respuestas se debe al sujeto, por lo que a veces no se ocupa; mientras que el segundo considera otras fuentes de variabilidad propio de los observadores (Streiner & Geoffrey).

Un ejemplo actual en la odontología es la difusión del nuevo método de diagnóstico de lesiones de caries dental llamado ICDAS (International Caries Detection and Assessment System) (2013). Un ejemplo claro para observar su validez y precisión diagnóstica es comparar la nueva clasificación de lesión de caries obtenida mediante una inspección visual con cortes histológicos del diente (patrón de oro) o radiografía digital. Luego, se realiza el grado de acuerdo entre los examinadores, la sensibilidad, la especificidad, la exactitud, el valor predictivo positivo y negativo y la validez de dichas mediciones. En este caso, el uso de ICDAS presenta una adecuada precisión diagnóstica, siendo útil para la práctica clínica actual en el diagnóstico del grado de avance de daño de caries dental (Mitropoulos *et al.*, 2010).

Otro ejemplo que aplica el proceso de confiabilidad es el estudio de Castrejón-Pérez *et al.* (2010). Al igual que la validación de la calidad de salud oral realizada en Chile por López & Baelum (2006), este grupo de investigadores adaptó culturalmente y validó una nueva versión del OHIP-sp en adultos mayores de la Ciudad de México. Se estableció la confiabilidad interobservador mediante el diagnóstico de la salud oral realizada por dos dentistas en 136 adultos mayores. Luego, se comparó los resultados registrados por ambos investigadores y se estableció un alta correlación entre sus mediciones mediante el índice kappa (0,9). A su vez, se aplicó el nuevo instrumento en dos tiempos distintos con una diferencia de 3 semanas, en el cual se comprobó que la repetibilidad del instrumento fue elevada (CCI= 0,877), por lo que se comprueba la posibilidad de usar dicho instrumento independiente del tiempo o condición local en cual se aplica (Castrejón-Pérez *et al.*).

DISCUSIÓN

El conocimiento de los conceptos de validez y confiabilidad, de sus diversos componentes obliga a los investigadores y clínicos del área odontológica a ser críticos con los métodos de medición utilizados.

Lo primero es fijarnos que el instrumento usado en nuestro quehacer clínico está siendo utilizado para el mismo objetivo para el cual fue diseñado. Así mismo, si se presentase más de una opción de medición (uso de nuevos localizadores de ápice o radiografías digitales, entre otras) debemos compararlas mediante sus características y rendimiento (validez y fiabilidad del instrumento). Cuando un reporte o equipo no menciona dichos valores dentro de sus pruebas científicas válidamente publicadas, es muy probable que no esté validado.

Es importante considerar que un instrumento de medición usado para registrar conceptos psicométricos o directamente estados clínicos que definen como validación una traducción al idioma empleado o una revisión por pares, no es un instrumento íntegramente validado (Sánchez & Echeverry, 2004). Es perentorio que se consideren todos los puntos comentados a lo largo de este manuscrito, con el fin de discriminar en el análisis interino de los artículos en el que se compruebe la validez del instrumento usado para el registro de las variables de interés.

ARAVENA, P. C.; MORAGA, J.; CARTES-VELÁSQUEZ, R. & MANTEROLA, C. Validity and Reliability in Dental Research. *Int. J. Odontostomat.*, 8(1):69-75, 2014.

ABSTRACT: The measurement process consists in assigning numerical values to the patient's characteristics, level of injury, diagnostic certainty of the disease, among others. However, there are complex phenomena that include broader concepts, which require the creation and validation of psychometric instruments for its measurement. The aim of this review is to present the concepts of validity and reliability of the instruments used in Dentistry. It exposes health definition of direct and indirect variables, the importance of using psychometric instruments, the definition of Validity and Reliability of an instrument and its different classifications (content validity, criterion and construct; reliability of the instrument and the observer). Furthermore, it presents some examples of the validation process of an instrument, its methodology and analysis of the results. Thus, we present the topics necessary to understand the measurement process used in our clinical work or dental research in order to identify appropriate instruments, considered valid and reliable for our needs.

KEY WORDS: validity, reliability, reproducibility of results, research, evidence-based dentistry, dentistry.

REFERENCIAS BIBLIOGRÁFICAS

- Alarcón, A. & Muñoz, S. Medición en salud: Algunas consideraciones metodológicas. *Rev. Med. Chile*, 136(1):125-30, 2008.
- Almendros-Marqués, N.; Berini-Aytés, L. & Gay-Escoda, C. Evaluation of intraexaminer and interexaminer agreement on classifying lower third molars according to the systems of Pell and Gregory and of Winter. *J. Oral Maxillofac. Surg.*, 66(5):893-9, 2008.
- American Educational Research Association & National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, American Psychological Association, 1999.
- Castrejón-Pérez, R. C.; Borges-Yáñez, S. A. & Irigoyen-Camacho, M. E. Validación de un instrumento para medir el efecto de la salud bucal en la calidad de vida de adultos mayores mexicanos. *Rev. Panam. Salud Pública*, 27(5):321-9, 2010.
- Devlin, S. J.; Dong, H. K. & Brown, M. Selecting a scale for measuring quality. *Market Res.*, 5(3):12-7, 1993.
- Goldstein, G. & Lersen, M. *Handbook of psychological Assessment*. 3rd ed. Kidlington, Elsevier, 2000.
- Hulley, S.; Cummings, S.; Browner, W.; Grady, D. & Newman, T. *Planificación de las mediciones: precisión y exactitud*. En: Diseño de las investigaciones clínicas. 3a ed. New York, Editorial Lippincott Williams & Wilkins, 2007.
- ICDAS. *ICDAS Foundation, International Caries Detection and Assessment System, 2013*. Disponible en: <https://www.icdas.org/>
- Kerlinger, F. *Investigación del comportamiento. Métodos de investigación en ciencias sociales*. México DF, McGraw-Hill, 2002.
- Kobayashi, C. Electronic canal length measurement. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.*, 79(2):226-31, 1995.
- López, R. & Baelum, V. Spanish version of the oral health impact profile (OHIP-Sp). *BMC Oral Health*, 6:11, 2006.
- Manterola, C. El proceso de medición con variables cualitativas y su aplicación en cirugía. *Rev. Chil. Cir.*, 54(3):307-15, 2002.
- Mitropoulos, P.; Rahiotis, C.; Stamatakis, H. & Kakaboura, A. Diagnostic performance of the visual caries

classification system ICDAS II versus radiography and micro-computed tomography for proximal caries detection: an in vitro study. *J. Dent.*, 38(11):859-67, 2010.

Nunnally, J. C. & Bernstein, I. H. *Teoría Psicométrica*. 2a ed. México D.F., McGraw-Hill Interamericana, 1999.

Pesudovs, K.; Burr, J. M.; Harley, C. & Elliott, D. The development, assessment, and selection of questionnaires. *Optom. Vis. Sci.*, 84(8):663-74, 2007.

Sánchez, R. & Echeverry, J. Validación de Escalas de Medición en Salud. *Rev. Salud Pública*, 6(3):302-18, 2004.

Shugars, D. A.; Benson, K.; White, R. P. Jr.; Simpson, K. N. & Bader, J. D. Developing a measure of patient perceptions of short-term outcomes of third molar surgery. *J. Oral Maxillofac. Surg.*, 54(12):1402-08, 1996.

Slade, G. D. & Spencer, A. J. Development and evaluation of the oral health impact profile. *Community Dent. Health*, 11(1):3-11, 1994.

Streiner, D. L. N. & Geoffrey, R. *Health Measurement Scales: A Practical Guide to Their Development and Use*. London, Oxford University Press, 2003.

WHO. *World Health Organization Technical Report Series 621. Epidemiology, etiology, and prevention of periodontal diseases*. Geneva, WHO, 1978.

Zavando, D.; Suazo, I. & Manterola, C. Validez en la investigación imagenológica. *Rev. Chil. Radiol.*, 16(2):75-9, 2010.

Dirección para Correspondencia:

Pedro Aravena

Rudloff 1640

Valdivia

CHILE

Teléfono: 063- 2221205

Email: paravena@uach.cl

Recibido: 07-09-2013

Aceptado: 03-12-2013