

## **EL ERROR DE OBSERVACION Y SU INFLUENCIA EN LOS ANALISIS MORFOLOGICOS DE RESTOS OSEOS HUMANOS. DATOS DE VARIACION CONTINUA**

*Sergio I. Perez<sup>1,2</sup>*  
*Paula N. González<sup>1</sup>*  
*Valeria Bernal<sup>1</sup>*  
*Mariano Del Papa<sup>1</sup>*  
*Amelia Barreiro<sup>1</sup>*  
*Carla Negro<sup>1</sup>*  
*Lucas Martínez<sup>1</sup>*

**PALABRAS CLAVE:** Morfometría, Esqueleto craneofacial, Relaciones poblacionales, Error de medición

**RESUMEN:** La utilización de métodos morfométricos se constituyó a lo largo del siglo XX en una de las aproximaciones más empleadas en el campo de la antropología biológica. Junto con el gran desarrollo alcanzado por la misma se generalizó el empleo de bases de datos generadas por otros investigadores con fines comparativos. Por este motivo se ha hecho necesario elaborar estrategias metodológicas que permitan evaluar e incrementar la precisión en las observaciones tomadas por uno o más observadores. El objetivo de este trabajo es evaluar el error intra e interobservador en el relevamiento de variables métricas

---

1 Departamento Científico de Antropología del Museo de La Plata. Facultad de Ciencias Naturales y Museo. Universidad Nacional de La Plata. Paseo del Bosque s/n. 1900 La Plata. Argentina.  
e-mail: iperez@fcnym.unlp.edu.ar, ivanperez@argentina.com

2 Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Buenos Aires. Argentina.

craneofaciales y la influencia de estos en el análisis de relaciones biológicas poblacionales. Con esta finalidad se elaboró un diseño experimental de bloques completos aleatorios con medidas repetidas. Se analizó una muestra de 20 cráneos procedentes del Valle de Río Negro -10 con deformación pseudocircular y 10 con deformación planolámbdica-. Cuatro observadores registraron un conjunto de ocho variables métricas craneofaciales. Las observaciones fueron comparadas mediante pruebas estadísticas de correlación intraclass y ANOVA de medidas repetidas. A su vez, se efectuaron análisis multivariados con el fin de conocer las relaciones entre las muestras de cráneos deformados medidas por los diferentes observadores. Los resultados obtenidos indican que hubo un incremento en la consistencia de las observaciones efectuadas por cada observador, que el error interobservador fue mayor entre los observadores no entrenados y que las inconsistencias en las observaciones afectaron los resultados obtenidos en los análisis multivariados tendientes a conocer las relaciones entre las muestras. *Rev. Arg. Antrop. Biol.* 6(1): 61-75, 2004.

**KEY WORDS:** Morphometric, Craniofacial skeleton, Populations relationships, Measurement error

**ABSTRACT:** The application of morphometric methods became one of the most frequently used approaches in the field of Biological Anthropology throughout the 20th century. Concurrently with this development, the use of databases generated by different researchers became widespread in comparative studies. Therefore, it has become necessary to develop methodological strategies that allow an evaluation of the observations made by one or more observers and increase their accuracy. The aim of this paper is to evaluate the intra- and inter-observer error in the scoring of craniofacial metric variables and its influence on the analysis of biological relationships between populations. For this purpose, an experimental design of random complete blocks with repetitions was applied. A sample of 20 skulls from Rio Negro Valley -10 with pseudo-circular deformation and 10 with plane-lambdaic deformation- was analyzed. Four observers recorded a set of eight craniofacial metric variables. The observations were statistically compared by means of intraclass correlation tests and repeated measures ANOVA. Multivariate analyses were also performed to assess the relationships among the samples of deformed skulls measured by different observers. Our results indicate an increase in the consistency of the observations made by each observer, higher inter-observer error values among non-trained observers, and the influence of the inconsistencies among observations on the results of the multivariate analyses designed for the assessment of inter-sample relationships. *Rev. Arg. Antrop. Biol.* 6(1): 61-75, 2004.

## INTRODUCCION

La utilización de métodos morfométricos se constituyó a lo largo del siglo XX en una de las aproximaciones más empleadas en el campo de la antropología biológica para el análisis de procesos microevolutivos, adaptativo-funcionales, entre otros. En las décadas de 1960 y 1970 comenzaron a aplicarse métodos estadísticos multivariados a variables métricas (Rohlf y Marcus, 1993), principalmente medidas de distancia lineales, para la descripción de los patrones de variación en forma (*shape*) dentro y entre poblaciones (Buikstra et al., 1990). Esta aproximación ha sido denominada morfometría tradicional (Marcus, 1990) o morfometría multivariada (Blackith y Reyment, 1971). A mediados de la década de 1990, con la aparición de nuevos medios de comunicación (i.e. Internet), se generalizó en los análisis de la variación esquelética el empleo de bases de datos generadas por otros investigadores con fines comparativos (Howells, 1973; 1996; Buikstra et al., 1990; Jantz y Owsley, 2002). A pesar del auge que ha tenido esta práctica en el seno de la antropología, la utilización de datos registrados por distintos observadores fue cuestionada, principalmente por la influencia que podría tener el error de observación o de medición en los resultados obtenidos (Jamison y Zegura, 1974; Utermohle y Zegura, 1982; Saunders, 1989; Guichón et al., 1993; 1996). El error de observación, definido como la diferencia entre medidas repetidas de una misma variable, puede ser de dos tipos, error intra e interobservador. Mientras que el primero aumenta la varianza de las muestras analizadas y por lo tanto incrementa la probabilidad de cometer un error de tipo II, el error interobservador puede resultar en la generación de sesgos sistemáticos en los datos, aumentando la probabilidad de cometer un error de tipo I (Bailey y Byrnes, 1990; Yezerinac et al., 1992). Por lo tanto, la precisión en el registro de los datos - entendida como la cercanía entre medidas repetidas de un mismo objeto- debe ser garantizada con el fin de no incrementar la incertidumbre en los análisis efectuados. En este sentido se ha señalado que la aplicación de diseños experimentales, así como el entrenamiento de los observadores, incrementarían la precisión de las observaciones (Guichón et al., 1993; 1996). Considerando lo expuesto anteriormente, los objetivos de este trabajo son evaluar el error intra e interobservador en el relevamiento de variables métricas craneofaciales, analizar la influencia de ambos tipos de error en el estudio de relaciones biológicas poblacionales y finalmente determinar el efecto del entrenamiento de los observadores sobre el error de medición. Con este fin se analizaron dos muestras procedentes del Valle inferior del Río Negro pertenecientes a diferentes momentos del Holoceno (establecido en base al tipo de deformación craneana artificial,

fechados radiocarbónicos, tipos de entierros, entre otras evidencias; Barrientos y Perez, 2004). Las hipótesis principales a evaluar son: A) las diferencias observadas en las muestras medidas por los diferentes observadores responden a diferencias morfológicas y B) las diferencias encontradas corresponden a diferencias en el entrenamiento de los observadores.

## **MATERIAL Y METODOS**

Para poner a prueba las hipótesis formuladas se elaboró un diseño experimental de bloques completos aleatorios con medidas repetidas (Cochran y Cox, 1997). El diseño implementado consistió en tres series de observaciones efectuadas por cuatro sujetos y espaciadas temporalmente a intervalos regulares de siete días. La aleatorización de los bloques y los intervalos entre las series permiten evitar sesgos y asegurar la independencia de las observaciones. Debido a que uno de los objetivos fue establecer la influencia del entrenamiento de los sujetos sobre el error de medición, se seleccionaron cuatro observadores que no tuvieran experiencia en realizar mediciones sobre restos óseos humanos. Con anterioridad al inicio de la experiencia, un entrenador instruyó a dos de los observadores en el procedimiento de medición de las variables. En cada una de las series, las observaciones se realizaron tomando cada espécimen y registrando el total de las variables mediante el uso de un calibre vernier (0,02 mm de resolución). Los cuatro observadores midieron 8 variables craneofaciales: altura orbital -OBH-, ancho orbital -OBB-, ancho interorbital -DKB-, ancho bifrontal -FMB-, ancho bizigomaxilar -ZMB-, ancho nasal -NLB-, altura nasal -NLH- y altura facial -NPH- (Howells, 1973). La muestra analizada consistió en 20 cráneos de la colección del Departamento Científico de Antropología del Museo de La Plata (FCNyM, UNLP) procedentes del Valle del Río Negro -10 con deformación pseudocircular (SC) y 10 con deformación planolámbdica (PL)-. Los cráneos fueron seleccionados aleatoriamente mediante el empleo del programa Random Number Generator Pro V1.26 (Segobit Software, 2001).

La precisión de las observaciones fue evaluada mediante análisis de correlación intraclase y de ANOVA de medidas repetidas (Zar, 1999). Estas pruebas posibilitan la evaluación de la concordancia y la presencia de posibles sesgos sistemáticos, respectivamente, en el registro de las variables (Bernal et al., 2004). Asimismo, con el fin de poner a prueba las hipótesis formuladas más arriba se efectuaron, sobre los resultados de la última serie, análisis de componentes principales y discriminante y se calcularon las distancias  $D^2$  de Mahalanobis (Mahalanobis, 1925; 1930; 1936; Manly, 1994) entre los dos grupos de cráneos

deformados medidos por los cuatro individuos. Se realizó, además, un análisis de escalamiento multidimensional (Kruskal, 1964a; b) sobre las distancias  $D^2$  de Mahalanobis. Finalmente, se efectuó una prueba de Mantel (Mantel, 1967; Smouse y Long, 1992; Manly, 1994; Waddle, 1994) para comparar la matriz de distancias  $D^2$  de Mahalanobis con dos matrices (Tabla 1) que describen las expectativas generadas por las dos hipótesis formuladas anteriormente.

**Tabla 1**

Matrices de 0 y 1 que describen las hipótesis A (a) y B (b) formuladas para comparar con las distancias  $D^2$  de Mahalanobis

**a**

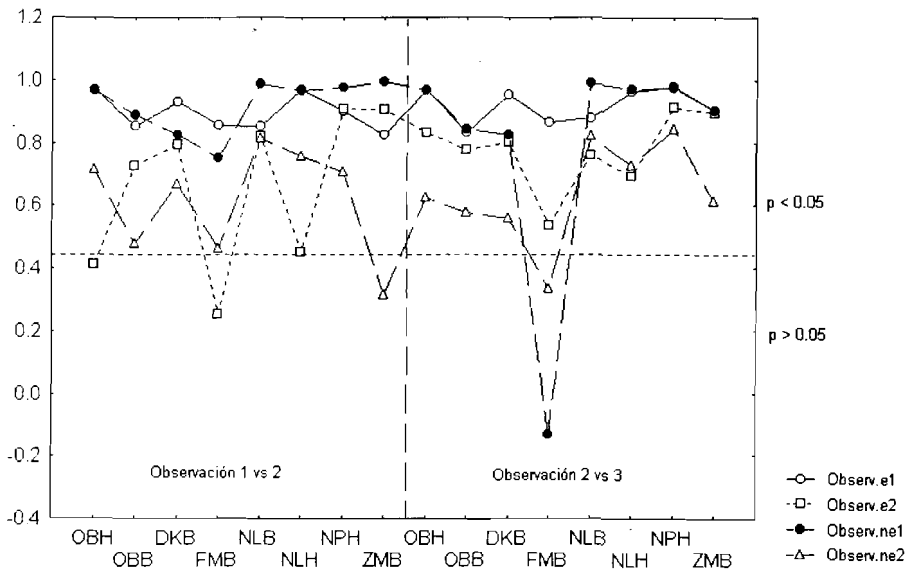
	SCe1	PLe1	SCne1	PLne1	SCe2	PLe2	SCne2	PLne2
SCe1	0							
PLe1	1	0						
SCne1	0	1	0					
PLne1	1	0	1	0				
SCe2	0	1	0	1	0			
PLe2	1	0	1	0	1	0		
SCne2	0	1	0	1	0	1	0	
PLne2	1	0	1	0	1	0	1	0

**b**

	SCe1	PLe1	SCne1	PLne1	SCe2	PLe2	SCne2	PLne2
SCe1	0							
PLe1	1	0						
SCne1	1	1	0					
PLne1	1	1	1	0				
SCe2	0	1	1	1	0			
PLe2	1	0	1	1	1	0		
SCne2	1	1	1	1	1	1	0	
PLne2	1	1	1	1	1	1	1	0

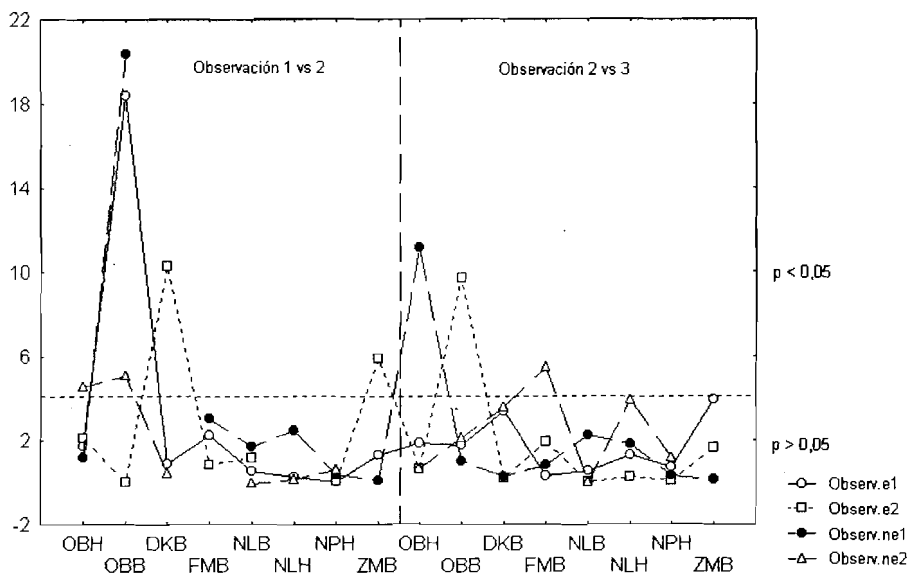
## RESULTADOS Y DISCUSION

Los resultados obtenidos indican un incremento en la precisión de las medidas efectuadas por cada sujeto a través del tiempo (Figuras 1 y 2), tendencia que es más clara en los observadores entrenados (e1 y e2). En las series del observador entrenado 1 (e1) y el no entrenado 1 (ne1), los resultados del análisis de correlación intraclase son superiores a 0,8 para la mayoría de las variables relevadas (Figura 1). En cambio, en las series del individuo entrenado 2 (e2) si bien se incrementa la consistencia, los valores del coeficiente son inferiores a 0,8 para casi todas las variables analizadas (Figura 1). Finalmente, en las variables registradas por el observador no entrenado 2 (ne2) los valores del coeficiente de correlación intraclase son bajos a lo largo de toda la experiencia (Figura 1). El análisis de ANOVA de medidas repetidas permite ver que en todos los observadores disminuyen las diferencias a lo largo de las sucesivas series de medición (Figura 2). Es importante señalar que, para facilitar la representación gráfica de los resultados del análisis de ANOVA de medidas repetidas, se eliminaron algunos de los valores de F de las variables DKB y NLH porque eran muy elevados, por lo tanto no están representados en la Figura 2.



**Figura 1**

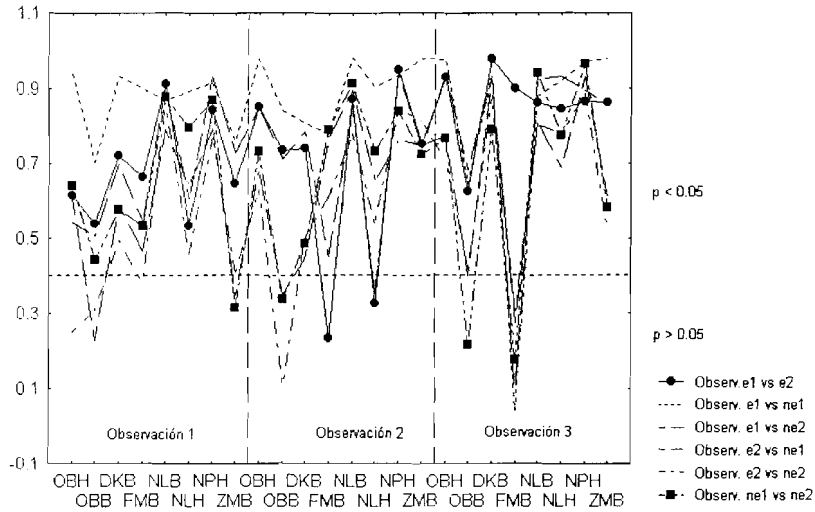
Resultados del análisis de correlación intraclase efectuado sobre las series 1 vs. 2 y 2 vs. 3 de los cuatro observadores.



**Figura 2**

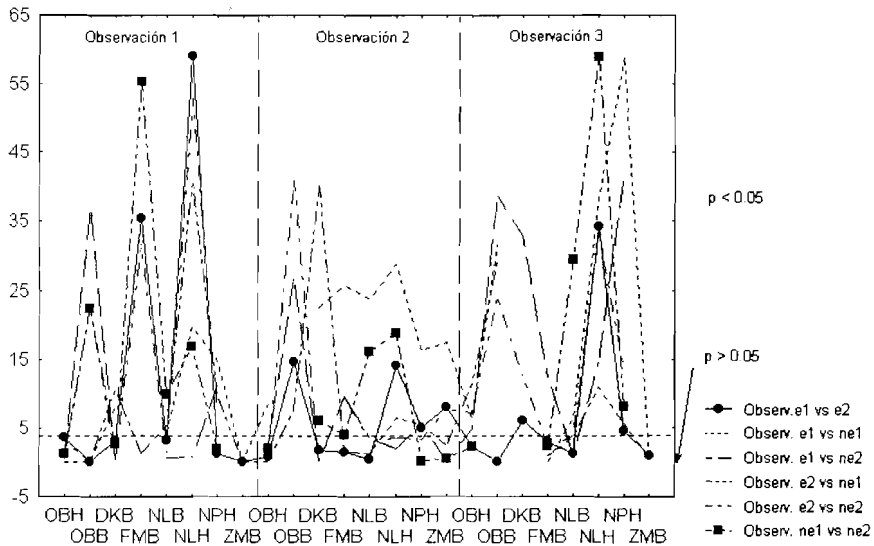
Resultados del análisis de ANOVA de medidas repetidas efectuado sobre las series 1 vs. 2 y 2 vs. 3 de los cuatro observadores.

El análisis de las series de medición realizadas por los distintos observadores indica que hay diferencias significativas entre estas a lo largo de la experiencia (Figuras 3 y 4), sin embargo, las inconsistencias son menores entre los observadores entrenados. Sólo en las mediciones de la última serie realizada por los observadores entrenados se alcanzaron valores elevados y significativos de concordancia en todas las variables (Figura 3). La prueba de ANOVA de medidas repetidas permite ver diferencias significativas ( $p < 0,05$ ) entre los observadores entrenados y los no entrenados para la mayoría de las variables, esta tendencia se mantiene a través de las series de medición (Figura 4). Los valores de F obtenidos al comparar las observaciones efectuadas por los individuos no entrenados se incrementan en las sucesivas series; en la última observación difieren significativamente en seis de las ocho variables analizadas (Figura 4). En los individuos entrenados, en cambio, los valores de F son bajos para la mayoría de las variables, excepto para la altura nasal (NLH) que presenta diferencias significativas a lo largo de toda la experiencia (Figura 4).



**Figura 3**

Resultados del análisis de correlación intraclass efectuado entre los cuatro observadores para las tres series registradas.



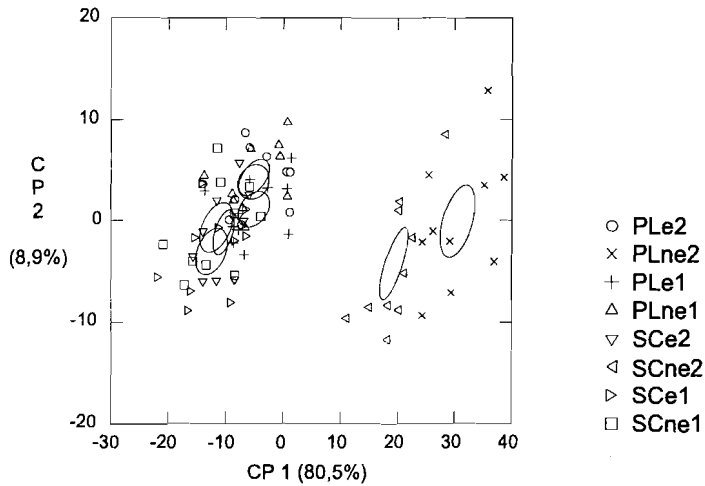
**Figura 4**

Resultados del análisis de ANOVA de medidas repetidas efectuado entre los cuatro observadores para las tres series registradas.



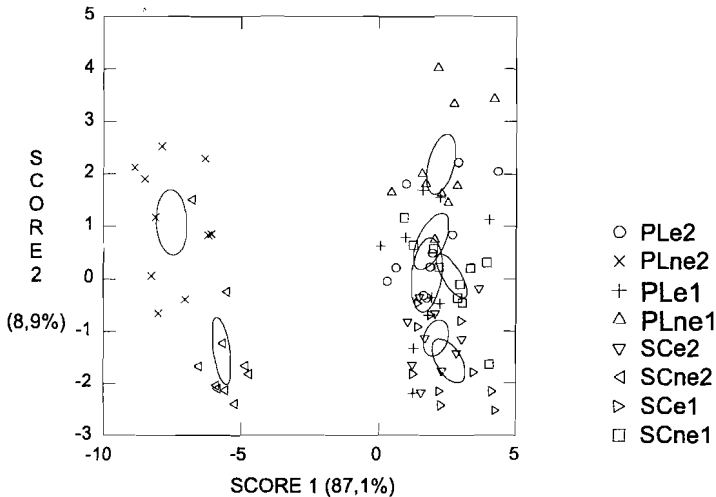
Los resultados hasta aquí obtenidos permiten señalar que a medida que se incrementa la precisión de cada observador disminuye la existente entre observadores, excepto para los individuos entrenados en los cuales se incrementan tanto la precisión intra como interobservador. En las primeras observaciones los sujetos definen los criterios de medición y presentan error principalmente aleatorio. Después de este período, alcanzan una mejor comprensión de la definición de las variables y tienen más consistencia en las mediciones. Sin embargo, los individuos que no han recibido entrenamiento establecen criterios de medición diferentes y en consecuencia aparecen errores sistemáticos entre los resultados de sus observaciones.

Como se discutió mas arriba, el error intraobservador aumenta la varianza de las muestras analizadas y por lo tanto aumenta la probabilidad de cometer un error de tipo II. A diferencia de este, el error interobservador introduce sesgos sistemáticos en la generación de los datos que pueden incrementar la probabilidad de cometer un error de tipo I. Sin embargo, no es claro en qué medida estos diferentes tipos de error afectan los análisis morfométricos, en particular, aquellos relacionados con el estudio de las relaciones interpoblacionales. En este sentido, los análisis multivariados efectuados ofrecen diferente tipo de información acerca de cómo el error intra e interobservador pueden influir en los estudios morfométricos. En el análisis de componentes principales (Figura 5) el primer componente, que explica un 80,5% de la varianza total, separa tres de los observadores (e1, e2 y ne1) del observador no entrenado 2. Las diferencias entre las muestras -con deformación pseudocircular y planolámbdica- se encuentran fundamentalmente en el segundo componente principal (8,9%). Los datos relevados por el individuo ne2 son los que exhiben mayor varianza en las muestras y mayor diferencia entre ellas. Los análisis discriminante y de escalamiento multidimensional (Figuras 6 y 7) presentan resultados similares entre sí. Las muestras medidas por los individuos que recibieron entrenamiento se agrupan de la manera esperada, es decir, en base a las diferencias morfométricas. Del mismo modo que en el análisis de componentes principales el *score* 1 del análisis discriminante explica gran parte de la variación total (87,1%). A diferencia del análisis de componentes principales los resultados de los dos individuos no entrenados difieren de lo esperado; las dos muestras medidas por ne2 y la muestra de cráneos con deformación planolámbdica medida por ne1 no se acercan a ninguna de las muestras analizadas, mientras que la muestra de cráneos con deformación pseudocircular medida por ne1 se acerca a las planolámbdicas de los entrenados.



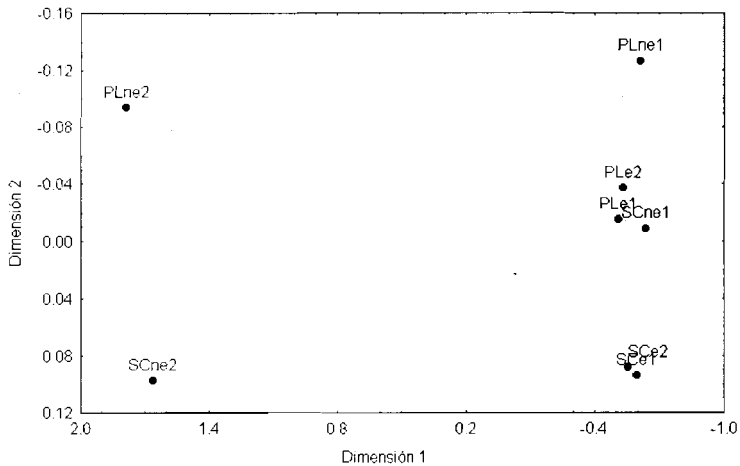
**Figura 5**

Resultados del análisis de componentes principales efectuado sobre la última serie de los cuatro observadores (PLe= muestra de individuos con deformación planolámbdica medida por un observador entrenado; PLne= muestra de individuos con deformación planolámbdica medida por un observador no entrenado; SCe= muestra de individuos con deformación pseudocircular medida por un observador entrenado; SCne= muestra de individuos con deformación pseudocircular medida por un observador no entrenado).



**Figura 6**

Resultados del análisis discriminante efectuado sobre la última serie de los cuatro observadores (PLe= muestra de individuos con deformación planolámbdica medida por un observador entrenado; PLne= muestra de individuos con deformación planolámbdica medida por un observador no entrenado; SCe= muestra de individuos con deformación pseudocircular medida por un observador entrenado; SCne= muestra de individuos con deformación pseudocircular medida por un observador no entrenado).



**Figura 7**

Resultados del análisis de escalamiento multidimensional efectuado sobre las distancias  $D^2$  de Mahalanobis calculadas sobre la última serie de los cuatro observadores (PLe= muestra de individuos con deformación planolámbdica medida por un observador entrenado; PLne= muestra de individuos con deformación planolámbdica medida por un observador no entrenado; SCe= muestra de individuos con deformación pseudocircular medida por un observador entrenado; SCne= muestra de individuos con deformación pseudocircular medida por un observador no entrenado).

En la Tabla 2 se presentan los valores de los pesos de cada variable en los dos primeros *scores* para los análisis de componentes principales y discriminante, se puede observar que en ambos análisis las variables que tienen mayor peso son las mismas que presentaron las mayores diferencias en el ANOVA de medidas repetidas (v.g. ZMB y NPH; Figura 4). Las diferencias encontradas entre los resultados de los tres análisis multivariados se deben a que el análisis de componentes principales se basa en el cálculo de los vectores de máxima variación en la muestra, en tanto el análisis discriminante y las distancias de Mahalanobis describen vectores de máxima separación entre grupos. Por lo tanto, los resultados de los dos últimos análisis están más influenciados por la presencia de diferencias debidas al error de medición en los conjuntos de datos. Esto mismo permite explicar por qué, contrariamente a lo esperado en un modelo de ANOVA, las muestras con mayor varianza (las que fueron medidas por ne2) son también las que más se diferencian. Finalmente, a partir de los resultados anteriores, se puede señalar que las diferencias observadas entre los conjuntos de datos se deben a diferencias en

el entrenamiento de los observadores. El análisis de Mantel (10000 remuestreos) confirma estos resultados. La correlación entre la matriz de distancias  $D^2$  y la matriz que representa la hipótesis A es baja y no significativa ( $r = -0,30890$ ;  $p = 0,0513$ ), en tanto que la correlación con la matriz que representa la hipótesis B es elevada y altamente significativa ( $r = -0,90268$ ;  $p = 0,0000$ ).

**Tabla 2**

Pesos para los dos primeros componentes principales (a) y funciones canónicas discriminantes (b)

**a**

	1	2
OBH	-0.134	0.561
OBB	-0.774	0.796
DKB	1.065	0.512
NLH	1.348	1.267
NLB	0.334	1.635
NPH	1.412	2.777
FMB	2.428	3.637
ZMB	15.376	-0.966

**b**

	1	2
OBH	0.152	0.024
OBB	0.398	0.437
DKB	0.226	-0.432
NLH	0.047	-0.776
NLB	0.169	0.672
NPH	0.256	0.531
FMB	-0.110	0.303
ZMB	-1.114	0.199

## CONCLUSIONES

Los resultados obtenidos indican que la aplicación de diseños experimentales permite controlar la variación debida al error intraobservador, en consecuencia puede reducirse la varianza introducida en el registro de los datos de manera que no influya sobre los análisis estadísticos realizados. Por otro lado, el error

interobservador sólo puede ser controlado en aquellos diseños en los cuales los observadores reciben un mismo entrenamiento. Estos resultados son consistentes con otros presentados anteriormente (Jamison y Zegura, 1974; Utermohle y Zegura, 1982; Saunders, 1989; Guichón et al., 1993; 1996) e indicarían la conveniencia de no incluir variables medidas por más de un observador en un mismo análisis estadístico. A partir de los resultados obtenidos en este trabajo se puede afirmar, además, que la presencia de inconsistencias intra e interobservador puede modificar completamente los resultados de los análisis multivariados. El grado y tipo de influencia variará según las técnicas estadísticas empleadas y el número de variables medidas que presenten diferencias significativas entre las observaciones.

## **AGRADECIMIENTOS**

Al Dr. Héctor M. Pucciarelli por facilitar el acceso a las colecciones del Departamento Científico de Antropología del Museo de La Plata (FCNyM. UNLP) a su cargo. Esta investigación fue realizada con fondos del Subsidio de Inicio de Carrera N° 14116-111 de la Fundación Antorchas, dirigido por el Dr. G. Barrientos.

## **BIBLIOGRAFIA CITADA**

Bailey RC y Byrnes J (1990) A new, old method for assessing measurement error in both univariate and multivariate morphometrics studies. *Syst. Zool.* 39:124-130.

Barrientos G y Perez SI (2004) Was there a population replacement during the late middle Holocene in the southeastern Pampas of Argentina? Discussing its archaeological evidence and its paleoecological basis. En Zárate M, Neme G y A Gil (eds): *Quaternary International (Special Volume) Ocupaciones Humanas Durante el Holoceno Medio en el Extremo Sudamericano*. INQUA, Pergamon Press. En prensa.

Bernal V, González P, Perez I y Del Papa M (2004) Evaluación del error intraobservador en bioarqueología. *Intersecciones en Antropología* 5:129-140.

Blackith RE y Reyment RA (1971) *Multivariate Morphometrics*. London, Academia London Press.

Buikstra JE, Frankenberg SR y Konigsberg LW (1990) Skeletal biological studies in American physical anthropology: recent trends. *Am. J. Phys. Anthropol.* 82:1-7.

Cochran WG y Cox GM (1997) *Diseños Experimentales*. México, Editorial Trillas.

Guichón R, Frascaroli F y Muñoz S (1996) Diseños de experimentos para el estudio de la variación interobservador: una aplicación en antropología biológica. *Arqueología: Solo Patagonia*. Ponencias de las II Jornadas de Arqueología de la Patagonia, pp.207-213.

Guichón R, Neder S y Orellana L (1993) Algunas consideraciones sobre los diseños de experimentos y estudios observacionales en antropología biológica. *Palimpsesto. Revista de Arqueología* 3:53-61.

Howells WW (1973) *Cranial variation in man. A study by multivariate analysis of patterns of difference among recent human populations*. Papers of Peabody Museum of Archaeology and Ethnology Harvard University Vol. 67. Cambridge, Mass.

Howells WW (1996) Howell's craniometric data on the Internet. *Am. J. Phys. Anthropol.* 101:441-442.

Jamison P y Zegura SL (1974) A univariate and multivariate examination of measurement error in anthropometry. *Am. J. Phys. Anthropol.* 40:197-204.

Jantz E y Owsley D (2002) Variation among early North American crania. *Am. J. Phys. Anthropol.* 114:146 -155.

Kruskal JB (1964a) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrics* 29:1-27.

Kruskal JB (1964b) Nonmetric multidimensional scaling: a numerical method. *Psychometrics* 29:115-129.

Mahalanobis PC (1925) Analysis of race mixture in Bengal. *J. Asiatic Soc. Bengal* 23:301-333.

Mahalanobis PC (1930) On test and measures of group divergence. *J. Asiatic Soc. Bengal* 26:541-588.

Mahalanobis PC (1936) On the generalized distance in statistics. *Proc. Nat. Inst. Sci. India* 2:49-55

Manly BFJ (1994) *Multivariate Statistical Methods*. Ed. Chapman & Hall.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209-220.

Marcus LF (1990) Traditional morphometrics. En Rohlf FJ y Bookstein FL (eds): *Proceedings of the Michigan Morphometrics Workshop*. Special Publication Number 2. Ann Arbor, University of Michigan. Museum of Zoology, pp.77-122.

Rohlf F y Marcus L (1993) A revolution in morphometrics. *Tree* 8:129-132.

Saunders SR (1989) Nonmetric skeletal variation. En Iscan MY y Kennedy KAR (eds): *Reconstruction of Life from the Skeleton*. New York, Alan R Liss, Inc., pp.95-108.

Segobit Software (2001) Programa Random Number Generator Pro V1.26. <http://www.segobit.com>

Smouse PE y Long JC (1992) Matrix correlation analysis in anthropology and genetics. *Am. J. Phys. Anthropol.* 35:187-213.

Utermohle CJ y Zegura SL (1982) Intra and interobserver error in craniometry: A cautionary tale. *Am. J. Phys. Anthropol.* 57:303-310.

Waddle DM (1994) Matrix correlation test support a single origin for modern humans. *Nature* 368:452-454.

Yezerinac SM, Loughheed S y Handford P (1992) Measurement error and morphometric studies: statistical power and observer experience. *Syst. Biol.* 41:471-482.

Zar JH (1999) *Biostatistical Analysis*. New York, Prentice Hall.