

Imágenes de Canva educativo

Tres procedimientos de baremación para la escala de actitud hacia la exclusividad marital en mexicanos adultos casados de Monterrey

Norms for the attitude toward marital exclusivity scale using three procedures in married individuals in Monterrey, Mexico

José Moral-de-la-Rubia

RESUMEN

En 1987, Weis y Fenton desarrollaron la Escala de Actitud hacia la Exclusividad Marital (ATMES). En 2020, Moral validó esta escala en una muestra de parejas casadas residentes en Monterrey. La distribución de sus puntuaciones se desvió de la normalidad, presentando asimetría positiva y ligera leptocurtosis, por lo que se hizo una baremación de la escala mediante puntuaciones decilares. El objetivo del presente trabajo fue desarrollar baremos de la ATMES para medir el posible efecto de la deseabilidad social sobre las puntuaciones, a través de decatipos y las diez puntuaciones de la proporción áurea en estadística (GRiSTEN), implementado en un script en R, obteniendo normas diferenciadas por sexo; así como comparar los tres sistemas de baremación, el método decilar empleado previamente y los dos procedimientos mencionados. Se realizó un estudio no experimental de corte transversal. Se administró un cuestionario compuesto por ocho escalas, incluida la ATMES, a una muestra de 727 personas mexicanas casadas (52.8 % mujeres y 47.2 % hombres), seleccionadas mediante muestreo de rutas aleatorias en un estudio previo (2015). Dado que se observaron diferencias de tendencia central entre mujeres y hombres, los baremos se desarrollaron de forma independiente para cada sexo. La asociación entre la actitud hacia la infidelidad y la deseabilidad social fue pequeña y similar en ambos sexos, por lo que no fue necesario ajustar la baremación. Los tres sistemas de baremación mostraron una correspondencia ordinal muy alta y produjeron clasificaciones similares de los participantes. Las puntuaciones GRiSTEN constituyeron una alternativa adecuada para la baremación de distribuciones marcadamente asimétricas, al preservar prácticamente la ordenación de los participantes y proporcionar una transformación más acorde con la distribución empírica de los datos.

PALABRAS CLAVE: baremación, actitud, infidelidad marital, sexo, deseabilidad social.

ABSTRACT

In 1987, Weis and Fenton developed the Attitude Toward Marital Exclusivity Scale (ATMES). In 2020, Moral validated this scale in a sample of married couples residing in Monterrey, Mexico. The distribution of ATMES scores deviated from normality, exhibiting positive skewness and slight leptokurtosis; therefore, the scale was normed using deciles. The aim of the present study was to develop additional ATMES norming procedures to assess the potential effect of social desirability on scale scores: sten scores and the ten Golden-Ratio Statistical Transformation Scores (GRiSTEN), implemented using an R script. The study also sought to develop sex-specific norms and compare the three norming procedures: the previously employed decile method and the two procedures proposed in the present study. A non-experimental cross-sectional study was conducted. A questionnaire comprising eight scales, including the ATMES, was administered to a sample of 727 married Mexican adults (52.8 % women and 47.2 % men) selected through randomroute sampling in a previous study conducted in 2015. Because differences in central tendency were observed between women and men, separate norms were developed for each sex. The association between attitudes toward infidelity and social desirability was small and similar across sexes; therefore, no adjustment for social desirability was required in the norming process. The three norming procedures showed very high ordinal correspondence and produced highly similar participant classifications. GRiSTEN scores constituted an appropriate alternative for norming markedly skewed distributions, as they virtually preserved participants' rank ordering while providing a transformation more consistent with the empirical distribution of the data.

KEYWORDS: norming, attitude, marital infidelity, sex, social desirability.

*Correspondencia: jose.morald@uanl.edu.mx/Fecha de recepción: 19 de marzo de 2026/Fecha de aceptación: 21 de octubre de 2026/Fecha de publicación: 30 de octubre de 2026.

Universidad Autónoma de Nuevo León, Facultad de Psicología, Subdirección de Investigación, Dr. Carlos Canseco 110, Mitras Centro, Monterrey, Nuevo León, México, C. P. 64460.

INTRODUCCIÓN

Una actitud positiva hacia la conducta de infidelidad sexual incrementa la probabilidad de incurrir en dicha conducta, lo que puede tener consecuencias muy deletéreas para la relación marital y la estabilidad familiar (Bozoyan y Schmiedeberg, 2023; Rodrigues y col., 2022).

Se puede definir infidelidad entre 2 personas casadas como el acto de incumplir el contrato de exclusividad sexual firmado entre los cónyuges (Rodríguez y Fernández, 2022). No obstante, esta definición sería reduccionista, ya que se suele distinguir una infidelidad sexual y otra emocional, por lo que abarcaría solo el primer subtipo (Estrada-Carmona y col., 2022). Cuando un cónyuge se enamora y fantasea una relación íntima con otra persona, con la que interacciona con frecuencia, sin llegar a la consumación sexual, se incurre en infidelidad emocional (Rokach y Chan, 2023).

Existen varios instrumentos para medir la actitud hacia la infidelidad en la pareja marital (Pandey y col., 2024). Destaca entre ellos una escala breve de 7 ítems creada en Estados Unidos por Weis y Fenton (1987) (Góngora y col., 2022; Moral, 2019), la Escala de Actitud hacia la Exclusividad Marital (ATMES; por sus siglas en inglés: Attitudes Toward Marital Exclusivity Scale) (Weis y col., 1992). Este instrumento es muy utilizado por adoptar una concepción amplia de la infidelidad, abarcando desde la emocional a la física y ya ha sido validada en población mexicana (Moral, 2020).

La ATMES presentó una consistencia interna buena ($\alpha = 0.87$), una estructura unidimensional y buenas propiedades de estabilidad temporal (Weis y col., 1992). Un modelo unifactorial para la ATMES con 7 indicadores (con una correlación entre los residuos de los ítems 6 y 7) mostró un buen ajuste a los datos mediante mínimos cuadrados libres de escala, usándose como datos de entrada la matriz de correlaciones policóricas: $\chi^2[13] = 17.498$, $P = 0.178$; P de Bollen y Stine = 0.061; $\chi^2/gl = 1.346$; valores de 0.999 en el índice de bondad de ajuste (GFI, por sus siglas en inglés: Good-

ness of Fit Index); de 0.997 en el índice de bondad de ajuste ajustado (AGFI, por sus siglas en inglés: Adjusted Goodness of Fit Index); de 0.998 en el índice de ajuste normado (NFI, por sus siglas en inglés: Normed Fit Index); de 0.997 en el índice de ajuste relativo (RFI, por sus siglas en inglés: Relative Fit Index); de 0.999 en el índice de ajuste comparativo (CFI, por sus siglas en inglés: Comparative Fit Index); un valor de 0.022 en la raíz cuadrada del error de aproximación cuadrático medio (RMSEA, por sus siglas en inglés: Root Mean Square Error of Approximation) y de 0.029 en la raíz cuadrada media del residuo estandarizado (por sus siglas en inglés: Standardized Root Mean Square Residual). El modelo de medida arrojó validez convergente de acuerdo a la varianza media extraída (AVE, por sus siglas en inglés: Average Variance Extracted) ($AVE = 0.813 > 0.5$) y confiabilidad interna excelente evaluada mediante los coeficientes omega de McDonald ($\omega = 0.968$) y alfa ordinal (α ordinal = 0.969), cumpliendo el supuesto de tau equivalencia (pesos de medida de 0.862 a 0.932) requerido por el coeficiente alfa (Moral, 2020).

La distribución de las puntuaciones de la ATMES se desvió de la normalidad ($Z_{KS} = 0.230$, $P < 0.001$), presentando asimetría positiva ($g_1/SE(g_1) = 13.758$) junto con ligera leptocurtosis ($g_2/SE(g_2) = 3.204$). Sus puntuaciones de validez y confiabilidad se desvían de la normalidad en la población mexicana (Moral, 2020), al igual que en otras poblaciones (Isma y Turnip, 2019; Labrecque y Whisman, 2017; Weis y Fenton, 1987), de acuerdo a sus puntuaciones decilares para su interpretación (Moral, 2020). Por ello, se optó por baremar la escala mediante puntuaciones percentilares en un rango del 1 a 5, transformando la escala discreta de los ítems en una continua (Moral, 2020). Esto se logra dividiendo la suma de los 7 ítems directos entre 7, es decir, entre el número de ítems, lo que representa la respuesta promedio del participante a los ítems que integran la escala.

La baremación tiene como finalidad facilitar la interpretación de las puntuaciones obteni-

das por una persona al compararlas con las de una muestra normativa. De este modo, el profesional puede determinar si una actitud hacia la exclusividad marital es baja, media o alta con respecto a la población de referencia, en lugar de limitarse a interpretar una puntuación bruta o un promedio de respuestas (Wijsen y col., 2022). Cuando la distribución de las puntuaciones se desvía de la normalidad, el método de baremación elegido influye en la clasificación de los individuos y, por tanto, en la interpretación de sus resultados. En estos casos, resulta conveniente emplear procedimientos que respeten la forma empírica de la distribución y proporcionen categorías interpretativas sencillas y comparables para la práctica psicológica (Ferr, 2025). Por ello, a continuación, se describen los principales métodos de baremación empleados cuando las puntuaciones se apartan de la normalidad, con especial atención al procedimiento GRiSTEN, objeto de este estudio. En psicometría, la baremación de puntuaciones en escalas con distribución normal, $x \in X \sim N(\mu, \sigma^2)$, suele realizarse con las puntuaciones T (media de 50 y desviación estándar de 20), las cuales transforman las puntuaciones estandarizadas para eliminar negativos y decimales ($T = 50 + 20 \times (x - M)/S$, donde M es la media y S la desviación estándar de la muestra normativa) (Kline, 2005).

Cuando las puntuaciones no siguen una distribución normal, una práctica habitual de baremación consiste en obtener eneatis y decatipos (Cohen y col., 2012). En primer lugar, las puntuaciones brutas de la variable X (suma de las puntuaciones en los ítems) se convierten en rangos percentilares medios, que representan el porcentaje acumulado hasta el punto medio del grupo de casos con la misma puntuación. Si F_i es la frecuencia acumulada de las puntuaciones inferiores a x_i , f_i la frecuencia de la puntuación x_i y n el tamaño de la muestra, el rango percentilar medio se calcula como: $P_i = 100 \times [F_i + 0.5 \times f_i]/n$. A continuación, cada rango percentilar medio se transforma en una probabilidad acumulada ($p_i = P_i/100$), a la que se aplica la función

cuantil de la distribución normal estándar (función probit), denotada por $\Phi^{-1}()$, obteniéndose la puntuación normalizada correspondiente: $z_i = \Phi^{-1}(p_i)$.

Se aplica una transformación lineal sobre estas puntuaciones normalizadas: $ei = 5 + 2 \times z_i$ para eneatis y $di = 5.5 + 2 \times z_i$ para decatipos. De este modo, se crean 9 (eneatis) o 10 intervalos de clase (decatipos) de puntuaciones normalizadas, centrados en 0. El 0 es la marca de clase —promedio de los límites: $(-0.5 + 0.5)/2$ — del quinto intervalo de puntuaciones ei con los eneatis, mientras que el 0 es el límite superior del quinto intervalo $[0.5, 0)$ y el límite inferior del sexto intervalo $[0, 0.5)$ de puntuaciones di con los decatipos. La amplitud de los intervalos de clase es de 0.5, salvo la de los dos extremos, los cuales tienen amplitudes indefinidas. Finalmente, los intervalos se convierten en números cardinales: de 1 a 9 en los eneatis y de 1 a 10 en los decatipos. Este cardinal corresponde a la marca de clase de los intervalos con amplitud de 0.5.

La idea original con los eneatis es que reduzcan las puntuaciones en la escala a un dígito (de 1 a 9) y tengan un dígito central (5) para una interpretación fácil e intuitiva. Las puntuaciones de 4 a 6 muestran un nivel medio de la característica medida. Un valor de 3 corresponde a un nivel insuficiente, y de 1 a 2 deficitario, mientras que un valor de 7 revela un nivel destacado, y de 8 a 9 sobresaliente (Ault y Emelianchik-Key, 2024). No obstante, los decatipos son más usados. Las puntuaciones de 4 a 7 indican un nivel medio, de 1 a 3 deficitario, y de 8 a 10 sobresaliente (Gunver, 2022).

Las puntuaciones percentilares representan una excelente opción para la baremación de distribuciones no normales (McKim, 2022).

Una propuesta de baremación más reciente son las 10 puntuaciones de la proporción áurea en estadística (Golden Ratio in Statistics TEN scores o GRiSTEN scores) de Gunver y

col. (2018). Al igual que los decatipos, las puntuaciones GRiSTEN son 10 cardinales, que varían de 1 a 10. De manera similar a los decatipos, puntuaciones de 5 a 6 se pueden interpretar como un nivel medio, de 1 a 2 muy deficitario, 3 a 4 deficitario, de 7 a 8 destacado y de 9 a 10 sobresaliente. Esta clasificación permite que el profesional interprete el resultado de un individuo mediante categorías fácilmente comunicables. Además, se ha propuesto que las puntuaciones GRiSTEN se ven menos afectadas por distribuciones fuertemente asimétricas y podrían proporcionar normas más adecuadas que los decatipos en tales circunstancias (Gunver, 2022).

Este estudio no busca validar nuevamente el instrumento, sino ampliar las alternativas de interpretación normativa, ya que no ha sido evaluada mediante otros baremos, ni se han comparado con otros sistemas de baremación. La reutilización de la base de datos se justifica por su obtención mediante muestreo probabilístico y por la distribución no normal de las puntuaciones, caracterizada por asimetría positiva y ligera leptocurtosis (Moral, 2020), además de su previa evaluación por deciles.

La interpretación de las puntuaciones de la ATMES también requiere considerar posibles fuentes de sesgo de respuesta. Se ha señalado que la deseabilidad social —en sus dimensiones de manejo de la impresión y autoengaño— influye en el autorreporte de conductas y actitudes sexuales en general (King, 2022; Krumpal, 2023) y también en la actitud hacia la infidelidad, aunque con una asociación de magnitud pequeña en este último caso (Bozoyan y Schmiedeberg, 2023). Dado que las actitudes favorables hacia la infidelidad pueden ser socialmente desaprobadas y suscitar respuestas defensivas o normativamente aceptables, resulta pertinente evaluar el grado en que la deseabilidad social afecta las puntuaciones de la ATMES (Bozoyan y Schmiedeberg, 2023).

Además, si dicho efecto alcanza una magnitud relevante, conviene considerar su control en la elaboración de normas interpretativas, con

el fin de aumentar la precisión de la baremación y la validez de las inferencias derivadas de las puntuaciones (Krumpal, 2023), dado que las puntuaciones de la ATMES han mostrado una distribución no normal caracterizada por asimetría positiva (Moral, 2020), y considerando la evidencia de diferencias por sexo en la actitud hacia la infidelidad —más favorable en los hombres que en las mujeres (Pazhoohi, 2022)—.

El objetivo de este estudio fue desarrollar dos baremos de la ATMES, basados en decatipos y puntuaciones GRiSTEN, implementado en R, obteniendo normas diferenciadas por sexo, para determinar el posible efecto de la deseabilidad social sobre las puntuaciones de la escala, así como comparar los tres sistemas de baremación —el método decilar empleado previamente y los dos procedimientos mencionados—.

MATERIALES Y MÉTODOS

Participantes

La muestra analizada en este estudio corresponde a la utilizada en el proceso de validación reportado por Moral (2020), cuyos datos fueron recolectados en 2015. Si bien, la antigüedad de los datos constituye una limitación relevante, se optó por su utilización debido a la alta calidad metodológica del levantamiento original, así como a la consistencia interna y la estabilidad psicométrica previamente documentadas. El estudio no requiere, en sentido estricto, asumir la invarianza temporal del nivel del rasgo, en tanto que las conclusiones se consideran válidas para el momento del muestreo y únicamente hipotéticas respecto al presente.

Los criterios de inclusión fueron: tener al menos 18 años, estar casado con una pareja del sexo opuesto, ser de nacionalidad mexicana, residir en la ciudad de Monterrey, Nuevo León (México), saber leer y escribir, y otorgar el consentimiento informado. Los criterios de eliminación fueron: responder el cuestionario de manera incompleta o informar que la pareja estuvo presente durante su aplicación leyendo o preguntando por las respuestas.

Se obtuvo una muestra de 727 participantes con datos completos, de los cuales 384 (52.8 %) fueron mujeres y 343 (47.2 %) hombres, sin diferencia estadísticamente significativa según la prueba binomial ($P = 0.138$). La mediana y la moda de escolaridad correspondieron al nivel medio superior, variando de primaria a posgrado. La mediana y moda (28.1 %) del ingreso económico familiar mensual en el último año se ubicó en el intervalo de \$ 6 000 a \$ 12 000 pesos mexicanos; 8.4 % tuvo un ingreso menor a \$ 3 000; 26.8 % de \$3 000 a \$ 11 999; 21.0 % de \$ 12 000 a \$ 23 999; 9.9 % de \$ 24 000 a \$ 47 999; y 5.8 % de \$ 48 000 o más.

La media de edad fue de 35.78 años (IC al 95 %: 35.03, 36.52) y el promedio de años de matrimonio fue de 12.30 (IC al 95 %: 11.62 a 12.99). La media, mediana y moda del número de hijos fue de 2. En cuanto a la afiliación religiosa, 82.8 % reportó ser cristiano católico, 9.1 % cristiano no católico, 1.8 % pertenecer a otra religión y 6.3 % no profesar ninguna.

Instrumentos de medida

La ATMES (Weis y Fenton, 1987) consta de 7 ítems directos con cinco categorías de respuesta ordenadas, que se puntúan de 1 = “total rechazo” a 5 = “total aceptación”. Su traducción al español se obtuvo a través de traducción inversa (Moral, 2020). Las puntuaciones se calculan mediante la suma simple de los ítems, variando de 7 a 35, donde valores más altos indican una mayor aceptación de la infidelidad hacia la pareja. En este estudio se utilizó el mismo rango de puntuación de 7 a 35 de la escala que en el estudio original de Weis y Fenton (1987), mientras que en el estudio previo de validación (Moral, 2020) se usó un rango de 1 a 5, lo que permite una interpretación absoluta de las puntuaciones a partir de las etiquetas de respuesta de los valores ordinales de los ítems.

El Inventario Balanceado de Deseabilidad Social al Responder (BIDR, por sus siglas en inglés: Balanced Inventory of Desirable Responding (Paulhus, 1998), en una versión abreviada de 20 ítems, en su validación en México

(BIDR-20, Moral y col., 2012). Los ítems, redactados en sentido de respuestas socialmente deseables, se contestan en una escala Likert de 7 categorías (1 = no cierto a 7 = muy cierto) y se organizan en dos factores de 10 ítems: autoengaño y manejo de la impresión. Las puntuaciones se obtienen mediante la suma de los ítems, de forma que valores más altos indican una mayor tendencia a responder de manera socialmente deseable. En su validación con una muestra de 624 estudiantes universitarios mexicanos (Moral y col., 2012), el modelo de dos factores correlacionados mostró buen ajuste por Mínimos Cuadrados Generalizados: ($\chi^2/\text{gl} = 1.74$), raíz cuadrada del error de aproximación cuadrático medio (RMSEA, por sus siglas en inglés: Root Mean Square Error of Approximation) (RMSEA = 0.03), índice de bondad de ajuste (GFI, por sus siglas en inglés: Goodness of Fit Index) (GFI = 0.95) e índice de bondad de ajuste ajustado (AGFI, por sus siglas en inglés: Adjusted Goodness of Fit Index) (AGFI = 0.94). La correlación entre factores fue de 0.67, lo que respalda su validez discriminante. La consistencia interna fue aceptable para la escala ($\alpha = 0.74$) y para ambos factores ($\alpha = 0.76$ y $\omega = 0.77$ para autoengaño y $\alpha = 0.71$ y $\omega = 0.67$ para manejo de la impresión).

En otro estudio de validación de una versión abreviada del Cuestionario de Alexitimia de Bermond-Vorst (BVAQ-20), en el que se aplicó el BIDR-20 en una muestra de 570 personas casadas (Moral, en prensa), el ajuste del modelo de dos factores correlacionados para el BIDR-20 fue bueno (SRMR = 0.049 (≤ 0.05)) y aceptable por la chi-cuadrada normalizada ($\chi^2/\text{gl} = 3.52 < 5$), el índice de ajuste comparativo (CFI, por sus siglas en inglés: Comparative Fit Index) (0.949), el Índice de Tucker-Lewis (TLI, por sus siglas en inglés: Tucker-Lewis Index) (TLI = 0.942 > 0.90) y RMSEA (0.067, IC 95 % [0.060, 0.074], $p(\text{H}_0: \text{RMSEA} \leq 0.08) \approx 1$), utilizando Mínimos Cuadrados Ponderados con Medias y Varianzas Ajustadas. La confiabilidad de la escala fue muy alta ($\omega = 0.93$). Autoengaño presentó una AVE de 0.46 y una confiabilidad alta ($\omega = 0.89$). De forma semejante, ma-

nejo de la impresión alcanzó una AVE de 0.45 y una confiabilidad alta ($\omega = 0.88$). Ambos factores mostraron validez discriminante de acuerdo a la razón heterotrait-monotrait de correlaciones (HTMT, por sus siglas en inglés: Heterotrait-Monotrait Ratio) ($HTMT = 0.62 < 0.90$). La distribución de la puntuación total y sus dos factores se desviaron de la normalidad.

La ATMES y el BIDR-20 se aplicaron junto con otros 7 instrumentos de medida y preguntas sociodemográficas en forma de cuestionario de autorreporte, al formar parte de un proyecto más amplio sobre infidelidad marital. Los otros instrumentos fueron: la Escala de valoración de la relación (RAS, por sus siglas en inglés: Relationship Assessment Scale) (Hendrick y col., 1998), el Índice de Satisfacción Sexual (ISS) (Hudson, 1992), la Escala de Engrandecimiento Marital (MAS, por sus siglas en inglés: Marital Aggrandizement Scale) (O'Rourke y Cappeliez, 2002), la Escala de Afecto Positivo y Negativo (PANAS, por sus siglas en inglés: Positive and Negative Affect Schedule) (Moral, 2011), la Escala de Búsqueda de Sensaciones Sexuales (SSSS, por sus siglas en inglés: Sexual Sensation Seeking Scale) (Moral, 2018) y la Escala de Conducta Infidel (ECI) (Romero y col., 2008). El cuestionario autoaplicado iniciaba con el consentimiento informado. Seguían los datos sociodemográficos, entre los cuales no aparecía nombre, dirección, teléfono o correo electrónico, para garantizar el anonimato y confidencialidad de los datos.

Procedimiento

Se realizó un estudio descriptivo-correlacional con un diseño *ex post facto* transversal. Se empleó un muestreo de rutas aleatorias, seleccionándose de forma al azar 81 calles a partir de la Guía Roji (2014). En cada calle se recolectaron 10 unidades (con consentimiento informado y datos completos). Se pretendía que fueran 5 mujeres y 5 hombres independientes (uno por casa).

El encuestador tocaba el timbre o puerta de la casa. Si esta era abierta, el entrevistador se

presentaba como ayudante de investigación e informaba del objetivo del estudio (la dinámica de la vida marital o de convivencia mediante un cuestionario), del responsable del proyecto y de los criterios de inclusión. Si la persona deseaba participar voluntariamente y satisfacía los criterios de inclusión, se dejaba un cuestionario para recogerlo media hora después o cuando lo indicase la persona. Se instruía que debía contestarlo a solas, sin que su pareja leyese las preguntas o respuestas. Al recogerlo se preguntaba si la pareja había estado presente a la hora de contestar y había visto sus respuestas. En tal situación, se descartaba el caso. De no ser así, se revisaba que estuviera todo el cuestionario contestado. Si había preguntas vacías, se pedía amablemente que se respondiesen y se explicaba que era necesario para que el análisis de los datos fuera confiable.

El tamaño muestral de 727 participantes (eliminados los 80 participantes en unión libre de la muestra total de 807) fue adecuado, ya que para estimar la media de la ATMES se requería una muestra mínima de 710 participantes, considerando una desviación estándar esperada de 7.61 (Moral, 2020; Weis y Fenton, 1987) dentro de un rango de 7 a 35, con un error absoluto de estimación de 0.56 y un IC del 95 %, aplicando la fórmula para la estimación bilateral de una media en poblaciones infinitas o muy grandes ($N > 100\ 000$).

Consideraciones éticas

La investigación contó con la aprobación del Comité de Doctorado en Filosofía con Orientación en Psicología de la Facultad de Psicología de la Universidad Autónoma de Nuevo León, México que también fungía como Comité de Ética en 2014. Se solicitó el consentimiento expreso informado de los participantes, se garantizó el anonimato de sus respuestas, no se solicitó ningún dato de identificación personal y se informó sobre la identidad del responsable del estudio, a quien se podía contactar por correo electrónico para cualquier cuestión suscitada por el cuestionario. De este modo, el estudio cumplió con

las normas éticas para investigación con participantes humanos, de la Declaración de Helsinki de la Asociación Médica Mundial (AMM, 2024), de la Sociedad Mexicana de Psicología (SMP, 2010) y de la Asociación Americana de Psicología (APA, por sus siglas en inglés: American Psychological Association) (APA, 2017).

Análisis de datos

Los análisis se iniciaron comprobando si las puntuaciones en ATMES se desvían de la normalidad. Para ello, se representó su distribución mediante un diagrama de cajas y bigotes. La asimetría ($H_0: \sqrt{\beta_1(X)} = 0$) se contrastó con la prueba de D'Agostino (1970); la mesocurtosis ($H_0: \beta_2(X) = 3$), con la prueba de Anscombe y Glynn (1983); y la normalidad ($H_0: X \sim N(\mu, \sigma^2)$), con la prueba de Shapiro y Francia (1972), utilizando la estandarización de Royston (1993), así como con la prueba omnibus de D'Agostino y col. (1990). Los análisis se realizaron con el programa R Core Team, 2025, versión 4.5.2. El nivel de significación se fijó en 0.05.

Comprobada la no normalidad de la distribución debido a la asimetría, las puntuaciones GRiSTEN se calcularon siguiendo el procedimiento propuesto por Gunver (2022) para la muestra total. Para obtener estas puntuaciones, se requiere calcular la media GRiS, denotada por O y las dos desviaciones gris (D_{GRIS}) independientes, a cada lado de la media GRiS, denotadas como D_{Izq} y D_{Der} . Sea X una variable cuantitativa continua o discreta con muchos valores distintos (al menos 10), de la cual se tiene una muestra aleatoria de tamaño grande ($n \geq \min(200, 4/5 \times N)$) y representativa de la población: $x = \{x_1, x_2, \dots, x_n\}$. Se ordenan los n datos muestrales en sentido ascendente, $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(i)} \leq \dots \leq x_{(n-1)} \leq x_{(n)}$, donde $i = 1, 2, \dots, n$. Se obtiene la mediana de los n datos muestrales: $mdn(x)$.

Se prosigue calculando las puntuaciones diferenciales con respecto a la mediana: $x_i - mdn(x)$. Se obtiene el ponderador de cada puntuación diferencial (m_{ci}), cuya formulación presenta dos

variantes, como se muestra en la Ecuación 1, donde $\phi = (1 + \sqrt{5})/2 \approx 1.618$.

$$m_{ci} = \begin{cases} \frac{1}{\phi} + 2 \times \frac{i-1}{n-1} & x_i < mdn(x) \\ 1 + \phi - 2 \times \frac{i-1}{n-1} & x_i \geq mdn(x) \end{cases} \quad (1)$$

Se denomina desviación ponderada desde la mediana (dpm) al producto de la puntuación diferencial con respecto a la mediana y su correspondiente ponderador: $m_{ci} \times (x_i - mdn(x))$. Véase Ecuación 2.

$$dpm_i = m_{ci} \times (x_i - mdn(x)) \quad (2)$$

La desviación desde la mediana (DM) es el cociente entre la suma de las n desviaciones ponderadas desde la mediana (numerador) y la suma de sus n ponderadores (denominador). Véase Ecuación 3.

$$DM = \frac{\sum_{i=1}^n dpm_i}{\sum_{i=1}^n m_{ci}} = \frac{\sum_{i=1}^n [m_{ci} \times (x_i - mdn(x))]}{\sum_{i=1}^n m_{ci}} \quad (3)$$

La *media GRiS*, que se denota por $O(x)$, es la suma de la mediana y la DM (Ecuación 4).

$$O(x) = mdn(x) + DM = mdn(x) + \frac{\sum_{i=1}^n [m_{ci} \times (x_i - mdn(x))]}{\sum_{i=1}^n m_{ci}} \quad (4)$$

Una vez obtenida la $O(x)$, se pueden calcular las D_{Izq} y D_{Der} . Se empieza con las puntuaciones diferenciales con respecto a la $O(x)$: $x_i - O$. Se sigue con la obtención de los ponderadores de estas puntuaciones diferenciales (d_{ci}), como se muestra en la Ecuación 5, donde k = número de datos muestrales menores a la $O(x)$.

$$k = \sum_{i=1}^n 1_{x_i < O(x)} = \#(x_i < O(x))$$

$$d_{ci} = \phi - \begin{cases} \frac{i-1}{k-1} & x_i < O(x) \\ 1 - \frac{i-k-1}{n-k-1} & x_i \geq O(x) \end{cases} \quad (5)$$

La desviación ponderada desde la media GRiS (DMO) es el producto de la puntuación diferencial con respecto a la $O(x)$ y su correspondiente ponderador (Ecuación 6).

$$DMO_i = d_{ci} \times (x_{(i)} - O(x)) \quad (6)$$

La D_{Izq} es el cociente entre la suma de las k desviaciones ponderadas desde la media GRiS negativas ($x_{(i)} < O(x)$) y la suma de sus correspondientes ponderadores. Su valor es negativo. La D_{Der} es el cociente entre la suma de las k desviaciones ponderadas desde la media GRiS positivas ($x_{(i)} > O(x)$) y la suma de sus correspondientes ponderadores. Su valor es positivo. Las desviaciones desde la $O(x)$ nulas quedan excluidas ($x_{(i)} = O(x)$). Para más detalles, véase la Ecuación 7.

$$D_{GRIS} = \begin{cases} D_{Izq}(x) = \frac{\sum_{i=1}^k [d_{ci} \times (x_{(i)} - O(x))]}{\sum_{i=1}^k d_{ci}} & x_{(i)} < O(x) \\ D_{Der}(x) = \frac{\sum_{i=k+1}^n [d_{ci} \times (x_{(i)} - O(x))]}{\sum_{i=k+1}^n d_{ci}} & x_{(i)} > O(x) \end{cases} \quad (7)$$

Finalmente, la variable continua (o discreta con muchos valores) se discretiza (o transforma) en puntuaciones GRiSTEN (de 1 a 10) mediante la suma de la media GRiS (O) y la D_{GRIS} (a la izquierda o a la derecha) ponderada. Para una escala directa al rasgo (mayor puntuación significa mayor nivel en el rasgo), los límites inferiores (inclusivos) de las puntuaciones GRiSTEN de 1 a 4 se determinan multiplicando la D_{GRIS} a la izquierda ($D_{Izq} < 0$) por los factores 2, 1.5, 1 y 0.5, respectivamente. La puntuación GRiSTEN de 5 toma como límite superior el valor 0. En el extremo opuesto, los límites superiores (excluyentes) para las puntuaciones GRiSTEN de 6 a 9 se calculan ponderando la D_{GRIS} a la derecha ($D_{Der} > 0$) por los factores 0.5, 1, 1.5 y 2, respectivamente. Este último valor constituye, a su vez, el límite inferior de la puntuación GRiSTEN de 10. En el caso de una escala inversa al rasgo (menor puntuación significa mayor nivel en el rasgo), los criterios de ponderación se aplican de forma invertida (Tabla 1).

Cuando el límite de un intervalo excede el mínimo o el máximo teórico de la escala, se trunca en dicho valor. Los intervalos subsecuentes (hacia la izquierda desde O , para el mínimo, o hacia la derecha desde O , para el máximo) quedan colapsados en ese mismo punto y se les asigna una puntuación GRiSTEN pro-

medio, calculada a partir de las puntuaciones GRiSTEN correspondientes a los intervalos colapsados en uno o ambos límites. En estos casos, se recomienda modificar la tabla eliminando los intervalos colapsados en ambos extremos y asignando la puntuación GRiSTEN promedio al intervalo colapsado en uno de ellos.

Tras detallar las fórmulas para calcular las puntuaciones GRiSTEN, se muestra un script en R para su cómputo en una muestra normativa. El script puede ejecutarse localmente con RStudio o en el entorno R, así como en línea mediante la plataforma, snippets. Se requiere $R \geq 2.14.0$. Con la muestra de ejemplo que trae el script (de 900 datos), el tiempo de ejecución en línea es inferior a un minuto, que es el tiempo máximo permitido por la plataforma Snippets.

Cabe advertir que, en muestras muy pequeñas o en distribuciones extremadamente asimétricas (por ejemplo, $x \leftarrow c(7, 7, 7, 7, 7, 7, 7, 35)$), puede ocurrir que el número de observaciones menores que la media GRiS (k) sea igual o menor que 1, o igual o mayor que $n - 1$, lo que produce una división por 0 en los ponderadores d_{ci} . En estos casos, el cómputo de las D_{GRIS} no puede realizarse mediante el procedimiento general, por lo que el script detiene la ejecución y muestra un mensaje de error.

```
# ----- Inicio
del script -----
--
# Indique la opción que aplica al Tipo de Escala
analizada:
# "directa" (mayor puntuación en X indica mayor
nivel del rasgo)
# "inversa" (menor puntuación en X indica
mayor nivel del rasgo)
tipo_escala <- "directa"
MIN <- 7 # Proporcione el mínimo teórico de la
variable X
MAX <- 35 # Proporcione el máximo teórico de
la variable X
# -----
-----
---
```

■ Tabla 1. Transformación de las puntuaciones de X a puntuaciones GRiSTEN.
Table 1. Transformation of X scores into GRiSTEN scores.

Puntuación GRiSTEN	Puntuaciones en X (directa al rasgo)	Puntuaciones en X (inversa al rasgo)
1	$x \geq O + 2 \times D_{Izq}$	$x \geq O + 2 \times D_{Der}$
2	$O + 2 \times D_{Izq} > x \geq O + 1.5 \times D_{Izq}$	$O + 2 \times D_{Der} > x \geq O + 1.5 \times D_{Der}$
3	$O + 1.5 \times D_{Izq} > x \geq O + D_{Izq}$	$O + 1.5 \times D_{Der} > x \geq O + D_{Der}$
4	$O + D_{Izq} > x \geq O + 0.5 \times D_{Izq}$	$O + D_{Der} > x \geq O + 0.5 \times D_{Der}$
5	$O + 0.5 \times D_{Izq} > x \geq 0$	$O + 0.5 \times D_{Der} > x \geq 0$
6	$0 > x \geq O + 0.5 \times D_{Der}$	$0 > x \geq O + 0.5 \times D_{Izq}$
7	$O + 0.5 \times D_{Der} > x \geq O + D_{Der}$	$O + 0.5 \times D_{Izq} > x \geq O + D_{Izq}$
8	$O + D_{Der} > x \geq O + 1.5 \times D_{Der}$	$O + D_{Izq} > x \geq O + 1.5 \times D_{Izq}$
9	$O + 1.5 \times D_{Der} > x \geq O + 2 \times D_{Der}$	$O + 1.5 \times D_{Izq} > x \geq O + 2 \times D_{Izq}$
10	$O + 2 \times D_{Der} > x$	$O + 2 \times D_{Izq} > x$

Nota. Puntuación GRiSTEN = las 10 puntuaciones de la proporción áurea en estadística, O = media GRiS o de la proporción áurea en estadística, D_{Izq} = desviación GRiS a la izquierda (su valor es negativo), D_{Der} = desviación GRiS a la derecha (su valor es positivo).

```
# Datos de entrada
# Muestra de ejemplo con 900 datos discretos
para ejecutar el script
set.seed(123) # Se fija una semilla para que la
muestra aleatoria sea reproducible.
x <- 7 + rbinom(n = 900, size = 28, prob = 0.3) #
Distribución binomial (n = 28, π = 0.3)
# Sustituya la muestra de ejemplo por sus datos.
Escriba o pegue los datos separados por comas.
# Los valores perdidos se codifican como NA
# Para activar un código quite el símbolo de
numeral que antecede y para desactivarlo se
añade.
# x <- c(14, 17, 15, ..., 15, 12, 17)
# x <- na.omit(x) # Elimina valores perdidos
(NA) si los hubiera
# -----
# Tamaño muestral y mediana
n <- length(x)
mdn <- median(x)
# Ordenación de los datos
xs <- sort(x)
i <- seq_len(n)
# -----
# Desviación ponderada respecto a la mediana
(DPM)
phi <- (1 + sqrt(5)) / 2
```

```
m_c <- ifelse(xs < mdn,
1 / phi + 2 * (i - 1) / (n - 1),
1 + phi - 2 * (i - 1) / (n - 1))
dpm <- (xs - mdn) * m_c
# Desviación media respecto a la mediana (DM)
DM <- sum(dpm) / sum(m_c)
# -----
# Media GRiS
O <- mdn + DM
# -----
# Desviación ponderada respecto a la media
GRiS (DMO)
k <- sum(xs < O) # Número de observaciones por
debajo de la media GRiS
if (k <= 1 || k >= n - 1) {stop("El cálculo de las
desviaciones GRiS no puede realizarse.\n",
"Esto se debe a que el número de observaciones
menores que la media GRiS (k = ", k, ")
es menor o igual a 1 o mayor o igual a n-1 que
corresponde a ", n - 1, " con estos datos.\n",
"Este caso puede presentarse con un tamaño
muestral muy reducido ",
"o una distribución extremadamente
asimétrica.")}
d_c <- ifelse(xs < O,
phi - (i - 1) / (k - 1),
phi - 1 + (i - k - 1) / (n - k - 1))
```

```

DMO <- (xs - O) * d_c
# -----
# Desviaciones GRiS
idx_left <- xs < O
Dizq <- sum(DMO[idx_left]) / sum(d_c[idx_left])
idx_right <- xs > O
DDer <- sum(DMO[idx_right]) / sum(d_c[idx_right])
# -----
# Límites de los intervalos GRiSTEN acotados al
mínimo y máximo teóricos
L1 <- max(MIN, O + 2 * Dizq)
L2 <- max(MIN, O + 1.5 * Dizq)
L3 <- max(MIN, O + 1 * Dizq)
L4 <- max(MIN, O + 0.5 * Dizq)
L5 <- O
L6 <- min(MAX, O + 0.5 * DDer)
L7 <- min(MAX, O + 1 * DDer)
L8 <- min(MAX, O + 1.5 * DDer)
L9 <- min(MAX, O + 2 * DDer)
# -----
# Puntuaciones GRiSTEN directas
GRiSTEN_directas <- ifelse(x <= L1, 1, ifelse(x <=
L2, 2,
ifelse(x <= L3, 3, ifelse(x <= L4, 4,
ifelse(x <= L5, 5, ifelse(x <= L6, 6, ifelse(x <= L7, 7,
ifelse(x <= L8, 8, ifelse(x <= L9, 9, 10))))))))))
# Puntuaciones inversas
GRiSTEN_inversas <- 11 - GRiSTEN_directas
# -----
# Ajuste por intervalos colapsados (corrección
por rango medio)
GRiSTEN_directas_adj <- GRiSTEN_directas
lower_bounds <- c(MIN, L1, L2, L3, L4, L5, L6,
L7, L8, L9)
upper_bounds <- c(L1, L2, L3, L4, L5, L6, L7, L8,
L9, MAX)
groups <- list()
trunc_min_idx <- which(lower_bounds == MIN)
if (length(trunc_min_idx) > 1)
{groups[[length(groups) + 1]] <- trunc_min_idx}
trunc_max_idx <- which(upper_bounds ==
MAX)
if (length(trunc_max_idx) > 1)
{groups[[length(groups) + 1]] <- trunc_max_idx}

for (g in groups) {mid_rank <- mean(g)
for (rank_idx in g) {GRiSTEN_directas_
adj[GRiSTEN_directas == rank_idx] <- mid_
rank}}
GRiSTEN_inversas_adj <- 11 - GRiSTEN_
directas_adj
# -----
# Puntuaciones finales
GRiSTEN_final <- if (tipo_escala == "directa")
GRiSTEN_directas_adj else GRiSTEN_
inversas_adj
# -----
# Resultados
# Controlar el número de decimales. Se redondea
la salida a 4.
rnd <- function(z, k = 3) round(z, k)
cat("Tamaño muestral: n =", n, ".\n")
cat("Mínimo muestral: min =", rnd(min(x)),
".\n")
cat("Máximo muestral: max =", rnd(max(x)),
".\n")
cat("Mediana muestral: mdn =", rnd(mdn), ".\n")
cat("Media GRiS: O =", rnd(O), ".\n")
cat("Desviación GRiS a la izquierda: D_izq =",
rnd(Dizq), ".\n")
cat("Desviación GRiS a la derecha: D_der =",
rnd(DDer), ".\n")
# -----
# Tabla de intervalos
if (tipo_escala == "directa") {
intervalos <- c(paste0("[", rnd(MIN), ", ", rnd(L1),
"]"), paste0("(", rnd(L1), ", ", rnd(L2), ")",
paste0("(", rnd(L2), ", ", rnd(L3), ")", paste0("(",
rnd(L3), ", ", rnd(L4), ")",
paste0("(", rnd(L4), ", ", rnd(L5), ")", paste0("(",
rnd(L5), ", ", rnd(L6), ")",
paste0("(", rnd(L6), ", ", rnd(L7), ")", paste0("(",
rnd(L7), ", ", rnd(L8), ")",
paste0("(", rnd(L8), ", ", rnd(L9), ")", paste0("(",
rnd(L9), ", ", rnd(MAX), ")",
cat("\nCuando varios intervalos colapsan
(comparten el mismo límite),
las puntuaciones GRiSTEN se ajustan
automáticamente mediante rangos medios.\n")
cat("\nTabla. Correspondencia entre las
puntuaciones GRiSTEN y los intervalos de una

```

```
escala con puntuaciones directas al rasgo\n")
} else {
intervalos <- c(paste0("[", rnd(MAX), ", ",
rnd(L9), "]"), paste0("(", rnd(L9), ", ", rnd(L8),
"]"),
paste0("(", rnd(L8), ", ", rnd(L7), "]"), paste0("(",
rnd(L7), ", ", rnd(L6), "]"),
paste0("(", rnd(L6), ", ", rnd(L5), "]"), paste0("(",
rnd(L5), ", ", rnd(L4), "]"),
paste0("(", rnd(L4), ", ", rnd(L3), "]"), paste0("(",
rnd(L3), ", ", rnd(L2), "]"),
paste0("(", rnd(L2), ", ", rnd(L1), "]"), paste0("(",
rnd(L1), ", ", rnd(MIN), "]"))
cat("\nCuando varios intervalos colapsan
(comparten el mismo límite),
las puntuaciones GRiSTEN se ajustan
automáticamente mediante rangos medios.\n")
cat("\nTabla. Correspondencia entre las
puntuaciones GRiSTEN y los intervalos de una
escala con puntuaciones inversas al rasgo\n")
}
tabla <- data.frame('Puntuaciones GRiSTEN' =
1:10,
'Intervalo en X' = intervalos, check.names =
FALSE)
print(tabla, row.names = FALSE)
cat("\nPuntuaciones GRiSTEN:\n")
print(GRiSTEN_final)
# ----- Fin
del script -----
---
```

Antes de establecer baremos distintos para mujeres y hombres, se contrastó la diferencia de medias entre los dos sexos mediante la prueba t de Welch (1947) para dos muestras independientes con varianzas heterogéneas, utilizando su implementación bootstrap (Efron y Narasimhan, 2020). Los IC se estimaron mediante el método de percentiles corregidos por sesgo y aceleración (BCa), con remuestreo con reposición a partir de la muestra original (2 000 muestras del mismo tamaño). Para garantizar la reproducibilidad de los resultados se fijó una semilla (2 000 000). La homogeneidad de varianzas se evaluó con la prueba de Levene. Los análisis se realizaron con el programa SPSS versión 27. El tamaño del efecto se estimó mediante el estadístico e insesgado de Aoki (2020), cuyo cálculo se rea-

lizó con Microsoft® Excel® para Microsoft 365 MSO (versión, 2602). Este se interpreta del mismo modo que el estadístico d de Cohen (1988): valores de $e < 0.20$ indican un tamaño del efecto trivial, entre 0.20 y 0.49 pequeño, entre 0.50 y 0.79 medio y ≥ 0.80 grande (Aoki, 2020). Dado que la diferencia de medias fue significativa, se calcularon las puntuaciones GRiSTEN para cada una de las submuestras por sexo.

Con fines comparativos, la asociación entre los sistemas de baremación basados en deciles, decatipos y puntuaciones GRiSTEN se evaluó mediante el coeficiente tau-b de Kendall, apropiado para variables ordinales con un elevado número de empates (Berry y Johnston, 2023). Además, para valorar la similitud en la clasificación de los participantes entre cada par de sistemas de baremación, se calcularon el porcentaje de coincidencia exacta (%exacta) y de coincidencia exacta o con una diferencia máxima de una categoría (% $\mp$), como indicadores descriptivos del grado de concordancia entre los procedimientos (Vanbelle y col., 2024).

Para valorar el sesgo por deseabilidad social en las puntuaciones de la ATMES, se calculó el coeficiente de correlación de Pearson (r). Dado el incumplimiento del supuesto de normalidad bivariada, el IC del 95 % se estimó mediante el método bootstrap de percentiles corregidos por sesgo y aceleración (BCa), a partir de 2 000 muestras obtenidas con reemplazo de la muestra original. Se consideró que la correlación era estadísticamente significativa cuando el IC no incluía el valor 0. Conforme a Cohen (1988), valores de $|r|$ entre 0 y 0.09 se consideraron triviales, entre 0.1 y 0.29 pequeños, entre 0.30 y 0.49 medios, entre 0.5 y 0.69 grandes, entre 0.70 y 0.89 muy grandes y entre 0.9 y 1 unitarios.

Si la varianza compartida (r^2) entre la puntuación total de la ATMES y la puntuación total del BIDR-20, o la de cualquiera de sus dos factores, resultaba igual o superior al 9 % ($r > 0.3$; Cohen, 1988), se planteó que las puntuaciones de la ATMES se ajustarían median-

te un modelo de regresión lineal simple. En este modelo, la puntuación total de la ATMES se especificaría como variable dependiente y la puntuación total del BIDR-20, o la del factor que presentara el mayor r^2 , como variable predictora. Los baremos se elaborarían a partir de las puntuaciones ajustadas estimadas por dicho modelo.

Además, estas correlaciones se compararon entre ambos sexos desde una implementación bootstrap, calculando un intervalo de confianza al 95 % para la diferencia entre las correlaciones (d) por el método BCa (Lakens, 2017). El tamaño del efecto se estimó por el estadístico Q de Cohen (1988). Valores de $Q < 0.20$ indican un tamaño del efecto trivial, de 0.20 a 0.49 pequeño, de 0.50 a 0.79 medio y ≥ 0.80 grande (Cohen, 1988). Cuando la magnitud de la relación entre ambas variables difiera significativamente entre hombres y mujeres y el tamaño del efecto sea al menos pequeño ($Q \geq 0.20$), se estimarán modelos de regresión independientes para cada sexo para obtener las puntuaciones ajustadas.

RESULTADOS

Distribución de ATMES

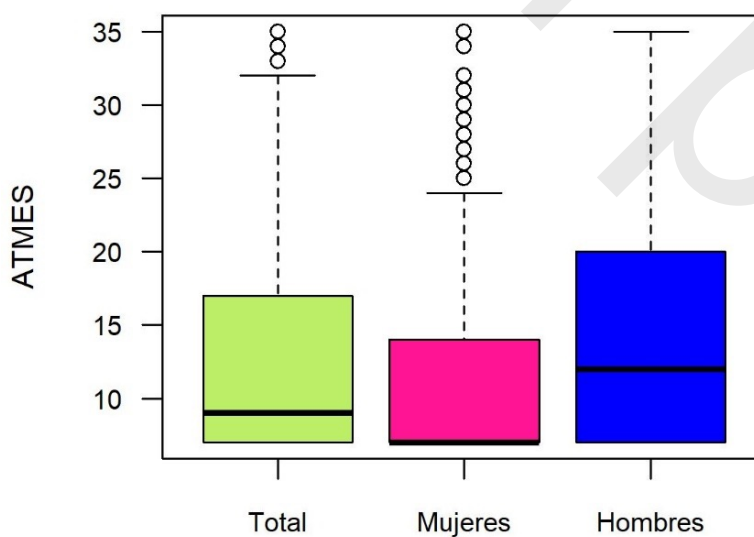
Las distribuciones de las puntuaciones en ATMES, de la muestra total y las submuestras

por sexo, presentaron no normalidad y asimetría positiva, o sesgo hacia la cola derecha (Figura 1) significativa en las tres muestras, según la prueba de D'Agostino (1970). Asimismo, mostraron leptocurtosis significativa o colas pesadas en la muestra total y en la submuestra de mujeres, mientras que en la submuestra de hombres se sostuvo la hipótesis nula de mesocurtosis. El supuesto de normalidad fue rechazado tanto por la prueba de Shapiro y Francia (Royston, 1993) como por la prueba de D'Agostino y col. (1990) en las tres muestras (Tabla 2).

Diferencia de medias

La media de los hombres ($m = 14.344$, IC BCa al 95 % [13.501, 15.203]; sesgo bootstrap = -0.001; error estándar bootstrap = 0.428) fue significativamente mayor en un contraste bilateral (P bootstrap = 0.001) que la de las mujeres ($m = 11.474$, IC BCa al 95 % [10.754, 12.183]; sesgo bootstrap = 0.011; error estándar bootstrap = 0.356).

La diferencia media sin asumir igualdad de varianzas fue de 2.870 (IC BCa al 95 % [1.797, 3.962]; sesgo bootstrap = -0.012; error estándar bootstrap = 0.561). El tamaño del efecto del sexo sobre el nivel de actitud fue pequeño (e insesgada de Aoki = 0.381).



■ Figura 1. Gráfico de caja de puntuaciones ATMES en la muestra total y por sexos.

Figure 1. Boxplot of ATMES scores in the total sample and by sex.

■ Tabla 2. Pruebas estadísticas para simetría, mesocurtosis y normalidad.
Table 2. Statistical tests for symmetry, mesocurtosis, and normality.

Prueba	Estadístico	Total	Mujeres	Hombres
Shapiro y Francia (Royston, 1993)	W'	0.780	0.698	0.852
	Z	10.574	9.620	7.806
	P	< 0.001	< 0.001	< 0.001
D'Agostino (1970)	$\sqrt{b_1}$	1.249	1.708	0.877
	Z	11.047	9.983	5.919
	P	< 0.001	< 0.001	< 0.001
Anscombe y Glynn (1983)	b_2	3.568	5.127	2.713
	Z	2.640	4.766	-1.102
	P	0.008	< 0.001	0.270
D'Agostino y col. (1990)	K^2	129	122.384	36.246
	P	< 0.001	< 0.001	< 0.001

Nota. Prueba de normalidad de Shapiro y Francia (Royston, 1993): W' = estadístico de contraste, Z = valor de W' logarítmicamente transformado y estandarizado bajo la hipótesis nula de normalidad, P = probabilidad en la cola derecha de una distribución normal estándar. Prueba de asimetría de D'Agostino (1970): $\sqrt{b_1}$ = medida de asimetría basada en tercer momento central estandarizado, Z = valor de $\sqrt{b_1}$ transformado y estandarizado bajo la hipótesis nula de simetría, P = probabilidad bilateral en una distribución normal estándar. Prueba de curtosis de Anscombe y Glynn (1983): b_2 = medida de curtosis basada en el cuarto momento central estandarizado, Z = valor de b_2 transformado y estandarizado bajo la hipótesis nula de mesocurtosis, P = probabilidad bilateral en una distribución normal estándar. Prueba ómnibus de normalidad de D'Agostino y col. (1990): K^2 = estadístico de contraste bajo la hipótesis nula de normalidad, P = probabilidad en la cola derecha de una distribución chi-cuadrado con dos grados de libertad.

La homogeneidad de varianza no se sostuvo por la prueba de Levene (F [1, 725] = 17.885, P < 0.001). La variabilidad fue mayor en hombres (s = 7.978, IC BCa 95 % [7.453, 8.473]; sesgo bootstrap = -0.018, error estándar bootstrap = 0.280) que en mujeres (s = 6.998, IC BCa al 95 % [6.297, 7.638]; sesgo bootstrap = -0.009; error estándar bootstrap = 0.358).

Sesgo de respuesta en relación con la deseabilidad social

La correlación entre la ATMES y la puntuación total de deseabilidad social fue significativa, pero de magnitud pequeña (r = -0.185, IC del 95 % BCa [-0.247, -0.121]), por lo que no se consideró suficiente para justificar un ajuste de las puntuaciones ni la elaboración de baremos corregidos por este sesgo en la muestra conjunta. Asimismo, la correlación con el componente de manejo de la impresión fue significativa, pero también pequeña (r = -0.234, IC del 95 % BCa [-0.297, -0.172]), mientras que la asociación con el componente de autoengaño fue significativa, aunque trivial (r = -0.065, IC del 95 % BCa [-0.134, -0.002]).

En términos de varianza compartida, la deseabilidad social total explicó el 3.4 % de la variabilidad de las puntuaciones de la ATMES, el manejo de la impresión el 5.5 % y el autoengaño apenas el 0.4 %.

No se observaron diferencias significativas entre las correlaciones de mujeres y hombres (d_r). En todos los casos, el tamaño del efecto (Q) fue trivial (Q < 0.20): para la puntuación total del BIDR-20, d_r = 0.047 (IC BCa del 95 % [-0.085, 0.184]) y Q de Cohen = 0.049 (IC del 95 % [-0.098, 0.195]); para el factor de autoengaño, d_r = 0.019 (IC BCa del 95 % [-0.115, 0.158]) y Q = 0.019 (IC del 95 % [-0.127, 0.165]); y para el factor de manejo de la impresión, d_r = 0.051 (IC BCa del 95 % [-0.087, 0.179]) y Q = 0.054 (IC del 95 % [-0.093, 0.200]).

Decatipos

En la Tabla 3 se presentan la media aritmética y la desviación estándar de la muestra total y de las submuestras por sexo requeridas para la transformación de las puntuaciones.

ciones brutas en decatipos, así como la correspondencia entre cada decatipo y su intervalo de puntuaciones en la ATMES. Cada decatipo (valor discreto) representa la marca de clase de un intervalo de puntuaciones continuas D (media 5.5 y desviación estándar 2), equivalente a un intervalo de puntuaciones Z (media 0 y desviación estándar 1), que se omitió en la tabla por simplicidad. Por ejemplo, en la muestra total, el decatipo 7 corresponde al intervalo Z [0.5, 1), al intervalo D [6.5, 7.5) y al intervalo de puntuaciones ATMES [16, 21].

En la submuestra de mujeres no existe ninguna puntuación bruta cuyo valor Z se sitúe entre -0.5 y 0 (equivalente al intervalo D [4.5, 5.5]), por lo que el decatipo 5 no tiene puntuaciones ATMES asociadas, mientras que el decatipo 4 solo comprende el valor mínimo de la escala (7). Aunque los intervalos en Z y D son simétricos por construcción, los intervalos correspondientes de la ATMES no lo son, reflejando la marcada asimetría positiva de la distribución. En consecuencia, los decatipos 1, 2 y 3 carecen de puntuaciones asociadas en las tres muestras, y la cola inferior de la distribución es más corta en mujeres que en hombres, debido a la ausencia o menor amplitud de los intervalos de puntuaciones bajas.

Puntuaciones GRiSTEN

En la Tabla 3 también se presentan los valores de la $O(x)$, que constituye el origen del centrado de las puntuaciones de la ATMES (O), y la D_{Izq} y D_{Der} , correspondientes a la muestra total y a las submuestras por sexo. Estos parámetros permiten definir los intervalos de puntuaciones de la ATMES asociados a cada puntuación discreta GRiSTEN. En la muestra total, el límite inferior del tercer intervalo teórico (5.097, 7.385] se truncó en el mínimo teórico de la escala (7), quedando como (7, 7.384], mientras que los dos primeros intervalos colapsaron en el único valor posible [7, 7], situación que también ocurrió en las submuestras por sexo. En consecuencia, la tabla se simplificó eliminando los intervalos anteriores al pri-

mero truncado, al que se asignó una puntuación GRiSTEN promedio de $(1 + 2 + 3)/3 = 2$.

Comparaciones entre deciles, decatipos y puntuaciones GRiSTEN

La asociación ordinal entre los sistemas de baremación fue muy alta en la muestra total y en ambos sexos ($\tau_b = 0.948 - 0.974$), indicando que los tres procedimientos preservan prácticamente el mismo orden de los participantes. Sin embargo, la coincidencia exacta en las categorías asignadas fue variable (2.9 % a 55.7 %) (Tabla 4).

DISCUSIÓN

Se confirmó que las puntuaciones de la ATMES se apartan de la normalidad debido a una asimetría positiva, mediante pruebas robustas y potentes que no habían sido aplicadas previamente al estudio de esta escala (Moral, 2020). En conjunto, los resultados indicaron que la mayoría de los participantes muestra un bajo nivel de aceptación de la conducta infiel, mientras que solo una minoría presenta actitudes claramente favorables, lo cual es consistente con los resultados de Bozoyan y Schmiedeberg (2023). Esta distribución hace poco apropiado el empleo de baremos simétricos basados en la media y la desviación estándar y justifica el uso de procedimientos de baremación que preserven la forma de la distribución, como los percentiles o las puntuaciones GRiSTEN.

Como era esperable, los hombres obtuvieron puntuaciones medias más altas que las mujeres. Esta diferencia es consistente con estudios previos y puede interpretarse desde diversas perspectivas propuestas en la literatura, tanto socioculturales como biológicas. Entre las primeras, se ha señalado la persistencia de normas de género que sancionan con mayor severidad la infidelidad femenina y muestran más permisividad hacia la conducta sexual masculina (Gregersen, 2022; Pazhoohi, 2022). Entre las segundas, se ha propuesto la influencia de diferencias sexuales en el impulso sexual y en factores hormonales relacionados con la conducta sexual (Jev-

■ Tabla 3. Correspondencia de los decatipos y las puntuaciones GRiSTEN con las puntuaciones de la ATMES en la muestra total y las submuestras por sexo, y estadísticos descriptivos.
Table 3. Correspondence between ATMES scores and STEN and GRiSTEN scores in the total sample and sex-specific subsamples, with descriptive statistics.

		Total	Mujeres	Hombres
Tamaño muestral	<i>n</i>	727	384	343
Mínimo	<i>min</i>	7	7	7
Máximo	<i>max</i>	35	35	35
Decil 1	<i>D1</i>	7	7	7
Decil 2	<i>D2</i>	7	7	7
Decil 3	<i>D3</i>	7	7	7
Decil 4	<i>D4</i>	7	7	9
Decil 5	<i>D5</i>	9	7	12
Decil 6	<i>D6</i>	12	9	15
Decil 7	<i>D7</i>	16	12	19
Decil 8	<i>D8</i>	19	16	21
Decil 9	<i>D9</i>	25	22	27
Media aritmética	<i>m</i>	12.828	11.474	14.344
Desviación estándar	<i>s</i>	7.607	6.998	7.978
Decatipos	1 $(-\infty, 1.5)$	-	-	-
	2 $[1.5, 2.5)$	-	-	-
	3 $[2.5, 3.5)$	-	-	-
	4 $[3.5, 4.5)$	[7, 7]	[7, 7]	[7, 7]
	5 $[4.5, 5.5)$	[8, 8]	-	[8, 11]
	6 $[5.5, 6.5)$	[9, 15]	[8, 11]	[12, 18]
	7 $[6.5, 7.5)$	[16, 21]	[12, 18]	[19, 22]
	8 $[7.5, 8.5)$	[22, 27]	[19, 25]	[23, 28]
	9 $[8.5, 9.5)$	[28, 33]	[26, 32]	[29, 33]
	10 $[9.5, \infty)$	[34, 35]	[34, 35]	[34, 35]
Estadísticos GRiS	<i>O</i>	11.961	9.442	13.692
	<i>D_{Izq}</i>	-4.576	-2.298	-5.686
	<i>D_{Der}</i>	10.140	10.922	10.097
Puntuaciones GRiSTEN	2*	[7, 7.384]	[7, 7.144]	[7, 8.006]
	4	(7.384, 9.673]	(7.144, 8.293]	(8.006, 10.849]
	5	(9.673, 11.961]	(8.293, 9.442]	(10.849, 13.692]
	6	(11.961, 17.031]	(9.442, 14.903]	(13.692, 18.741]
	7	(17.031, 22.101]	(14.903, 20.364]	(18.741, 23.789]
	8	(22.101, 27.171]	(20.364, 25.825]	(23.789, 28.838]
	9	(27.171, 32.241]	(25.825, 31.286]	(28.838, 33.887]
	10	(32.241, 35]	(31.286, 35]	(33.887, 35]

Nota. El decil 5 corresponde a la mediana. *O* = media GRiS, *D_{Izq}* = desviación GRiS a la izquierda y *D_{Der}* = desviación GRiS a la derecha. *Debido a que los dos primeros intervalos estaban colapsados en el valor 7 [7, 7], que corresponde al valor mínimo de ATMES (MIN = 7), y dado que este valor está incluido en el tercer intervalo [7, 7.384], dicho valor mínimo se asignó al tercer intervalo —que aparece como el primero en la tabla— y se le otorgó una puntuación GRiSTEN promedio de $(1 + 2 + 3)/3 = 2$.

■ Tabla 4. Comparaciones entre deciles, decatipos y puntuaciones GRiSTEN.
Table 4. Comparisons between deciles, STEN scores, and GRiSTEN scores.

Comparación	Muestra total			Mujeres			Hombres		
	τ_b	%exacta	% $\neq$	τ_b	%exacta	% $\neq$	τ_b	%exacta	% $\neq$
Deciles vs Decatipos	0.961	14.9	47.5	0.969	13.0	93.0	0.956	23.0	61.2
Deciles vs GRiSTEN	0.971	5.4	85.7	0.974	2.9	83.6	0.948	55.7	91.5
Decatipos vs GRiSTEN	0.959	33.3	50.6	0.969	28.6	43.0	0.956	43.7	61.5

Nota. τ_b = coeficiente de correlación por rangos tau-b de Kendall, %exacta = porcentaje de coincidencia exacta y % $\neq$ = porcentaje de coincidencia exacta o con una diferencia máxima de una categoría.

toxic-Todorovic, 2022; Maxwell y col., 2023; Sahithya y Kashyap, 2022). Sin embargo, el presente estudio no permite discriminar entre estas explicaciones ni establecer relaciones causales. En consecuencia, dichas interpretaciones deben considerarse únicamente como hipótesis que futuras investigaciones podrán contrastar.

Los baremos construidos mediante puntuaciones GRiSTEN reflejan adecuadamente la forma asimétrica de la distribución observada (Tabla 3). La mayor concentración de participantes en puntuaciones bajas explica que los intervalos correspondientes a los niveles superiores sean más estrechos, mientras que los inferiores sean más amplios, especialmente entre las mujeres. Este patrón probablemente refleja características de la distribución poblacional, aunque tampoco puede descartarse la influencia de factores de medición, como el contenido de los ítems, efectos suelo o posibles sesgos de medición. Dado que dichos aspectos no fueron evaluados mediante análisis específicos de funcionamiento diferencial de los ítems o de invarianza, esta interpretación debe considerarse provisional. En consecuencia, cuando se dispone de información sobre el sexo de la persona evaluada resulta recomendable emplear los baremos específicos para mujeres y hombres. Cuando dicha información no esté disponible, los baremos derivados de la muestra total constituyen una alternativa razonable.

Los tres procedimientos de baremación mostraron una correspondencia ordinal muy alta, como reflejan los coeficientes tau-b de Ken-

dall superiores a 0.94 en todas las comparaciones (Tabla 4). Estos resultados indican que los deciles, los decatipos y las puntuaciones GRiSTEN preservan prácticamente el mismo orden de los participantes. Aunque la coincidencia exacta entre categorías fue variable según el par de procedimientos y el sexo, la mayoría de las discrepancias se mantuvo dentro de una categoría de diferencia. Cabe destacar que la coincidencia exacta entre deciles y GRiSTEN fue particularmente baja en mujeres (2.9 %), pero considerablemente mayor en hombres (55.7 %), pese a que la asociación ordinal fue muy alta en ambos grupos. Este patrón muestra que una elevada correspondencia en la ordenación de los participantes no implica necesariamente una coincidencia elevada en las categorías asignadas, debido a los distintos criterios de transformación y asignación empleados por cada procedimiento, sin que ello implique cambios sustanciales en la posición relativa de los participantes. En consecuencia, los tres sistemas permiten realizar una interpretación ordinal consistente de las puntuaciones, aunque no deben considerarse intercambiables. Entre ellos, las puntuaciones GRiSTEN presentan la ventaja de aproximar la distribución transformada a la normalidad sin alterar apreciablemente la ordenación de los individuos. Desde una perspectiva aplicada, el procedimiento GRiSTEN permite obtener categorías interpretativas más acordes con la distribución real de las puntuaciones de la ATMES, reduciendo el riesgo de clasificar de forma poco precisa a los participantes cuando la distribución es marcadamente asimétrica. En consecuencia, facilita una interpretación normativa más precisa de

los resultados tanto en investigación como en la práctica profesional.

Una aportación adicional del estudio es la disponibilidad de un script en R para calcular puntuaciones GRiSTEN. Aunque previamente Gunver y col. (2018) habían desarrollado un procedimiento en Matlab, la implementación en R facilita su utilización al tratarse de un software libre ampliamente utilizado en investigación psicológica y estadística (Staples, 2023). No obstante, la evidencia empírica acerca de la utilidad práctica de las puntuaciones GRiSTEN frente a los métodos tradicionales de baremación todavía es limitada. El presente estudio muestra que los distintos procedimientos producen clasificaciones muy similares, aunque no idénticas, pero no permite establecer cuál resulta más útil para la interpretación clínica o para la toma de decisiones. Para responder a esta cuestión será necesario incorporar criterios externos, como diagnósticos, desenlaces clínicos o estudios longitudinales, objetivos que exceden el alcance del presente trabajo.

Como cabía esperar en un constructo socialmente sensible (King, 2022; Krumpal, 2023), la actitud hacia la infidelidad mostró una asociación negativa con la deseabilidad social. Sin embargo, el tamaño de esta relación fue pequeño y no alcanzó el criterio previamente establecido para justificar ajustes en la baremación. Además, la magnitud de las asociaciones fue similar en mujeres y hombres. En conjunto, estos resultados sugieren que la influencia de la deseabilidad social sobre las puntuaciones de la ATMES es reducida, en concordancia con investigaciones previas.

Un aspecto que debe considerarse al interpretar los baremos es su vigencia temporal. Los datos utilizados para su construcción fueron recolectados en 2015, por lo que existe un intervalo de aproximadamente una década entre la obtención de la muestra y la publicación de los resultados. Durante este periodo, las actitudes hacia las relaciones de pareja y la infidelidad podrían haber experimentado cam-

bios asociados con transformaciones sociales, culturales y generacionales. Por ejemplo, modificaciones en las normas de género, en las concepciones sobre las relaciones de pareja y en la aceptación social de distintas formas de comportamiento sexual que podrían alterar la distribución de las puntuaciones de la ATMES. Sin embargo, el presente estudio no incluye datos obtenidos en diferentes momentos que permitan determinar si se ha producido un cambio temporal en estas actitudes. Por tanto, no es posible establecer que los baremos hayan perdido vigencia, pero tampoco asumir que la distribución observada en 2015 representa sin cambios a la población actual. En este sentido, los baremos deben considerarse normas derivadas de una muestra recolectada en 2015 y su utilización en población contemporánea debería acompañarse de cautela, particularmente cuando las decisiones individuales tengan consecuencias relevantes. Estudios posteriores con muestras recientes permitirían evaluar la estabilidad temporal de la distribución de las puntuaciones y determinar la necesidad de actualizar los baremos.

Entre las limitaciones del estudio debe señalarse que la muestra estuvo integrada exclusivamente por personas casadas en relaciones heterosexuales residentes en Monterrey, con edades correspondientes a la adultez joven y media. En consecuencia, los resultados no son directamente aplicables a personas en otros tipos de relación de pareja, a parejas homosexuales o poligámicas, a adolescentes, adultos mayores o población rural.

Asimismo, aunque los baremos podrían resultar útiles en contextos urbanos con características socioculturales similares, su generalización a otras regiones de México o de Latinoamérica no se ha evaluado. Por ello, cualquier extrapolación debe considerarse provisional y ser confirmada mediante investigaciones independientes en las poblaciones de interés.

Además, todas las variables fueron evaluadas mediante un único cuestionario de autorreporte aplicado en un mismo momento, lo que in-

introduce la posibilidad de sesgo de método común. En consecuencia, parte de las asociaciones observadas entre las variables podría estar influida por características compartidas del procedimiento de medición más que por la relación sustantiva entre los constructos evaluados.

CONCLUSIONES

El script desarrollado en R es de fácil aplicación y permite obtener puntuaciones GRiSTEN a partir de muestras aleatorias amplias y representativas de la población. Este procedimiento proporciona una baremación adecuada para distribuciones no normales y asimétricas, frecuentes en investigación psicológica. En la ATMES se confirmó una distribución con asimetría positiva y diferencias en tendencia central entre mujeres y hombres, lo que justificó la elaboración de baremos diferenciados por sexo, preferibles a los de la muestra total cuando se dispone de esta información. La asociación entre actitud hacia la infidelidad y deseabilidad social fue pequeña y similar en ambos sexos, por lo que no fue necesario ajustar la baremación. Los baremos basados en deciles, decatipos y puntuaciones GRiSTEN mostraron una correspondencia ordinal muy alta, preservando prác-

ticamente el mismo orden de los participantes, aunque no deben considerarse intercambiables debido a las diferencias en la asignación de categorías derivadas de los distintos procedimientos de transformación. Entre ellos, las puntuaciones GRiSTEN presentan la ventaja de proporcionar una transformación más acorde con la distribución empírica de los datos sin alterar apreciablemente la ordenación de los individuos. Los baremos obtenidos se desarrollaron para adultos jóvenes y de mediana edad en matrimonios heterosexuales residentes en Monterrey. Su utilización en otras poblaciones urbanas mexicanas o latinoamericanas constituye una extrapolación que requiere confirmación empírica y debe interpretarse con cautela, particularmente considerando que los datos fueron recolectados en 2015.

AGRADECIMIENTOS

El estudio fue financiado a través del Programa de Apoyo a la Investigación Científica y Tecnológica de la Universidad Autónoma de Nuevo León (proyecto CS454-14).

DECLARACIÓN DE CONFLICTO DE INTERESES

Los autores declararon no tener conflictos de intereses de ningún tipo.

REFERENCIAS

- Ault, H. R. & Emelianchik-Key, K. (2024). Test selection, scoring, and statistics. In C. S. Gill, A. Torres, & K. Emelianchik-Key (Eds.), *Appraisal, assessment, and evaluation for counselors: a practical guide* (pp. 123-148). Springer Publishing. <https://doi.org/10.1891/9780826189134.0006>
- AMM, Asociación Médica Mundial (2024). Declaración de Helsinki de la AMM – Principios éticos para las investigaciones médicas en seres humanos. [En línea]. Disponible en: <https://www.wma.net/es/polices-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/>. Fecha de consulta: 24 de marzo de 2026.
- Anscombe, F. J. & Glynn, W. J. (1983). Distribution of the kurtosis statistic b_2 for normal samples. *Biometrika*, 70(1), 227-234. <https://doi.org/10.1093/biomet/70.1.227>
- Aoki, S. (2020). Effect sizes of the differences between means without assuming variance equality and between a mean and a constant. *Heliyon*, 6, e03306. <https://doi.org/10.1016/j.heliyon.2020.e03306>
- APA, American Psychological Association (2017). Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2016). [En línea]. Disponible en: <https://www.apa.org/ethics/code/index.aspx>. Fecha de consulta: 24 de marzo de 2026.
- Berry, K. J. & Johnston, J. E. (2023). Measures of ordinal association I. In *Statistical methods:*

- connections, equivalencies, and relationships* (pp. 379-434). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-41896-9_9
- Bozoyan, C. & Schmiedeborg, C. (2023). What is infidelity? A vignette study on norms and attitudes toward infidelity. *The Journal of Sex Research*, 60(8), 1090-1099. <https://doi.org/10.1080/00224499.2022.2104194>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second edition). Routledge. <https://doi.org/10.4324/9780203771587>
- Cohen, R. J., Swerdlik, M., & Sturman, E. (2012). *Psychological testing and assessment* (Eight edition). McGraw-Hill Higher Education.
- D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g_1 . *Biometrika*, 57(3), 679-681. <https://doi.org/10.1093/biomet/57.3.679>
- D'Agostino, R. B., Belanger, A., & D' Agostino, Jr. R. B. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4), 316-321. <https://doi.org/10.1080/00031305.1990.10475751>
- Efron, B. & Narasimhan, B. (2020). The automatic construction of bootstrap confidence intervals. *Journal of Computational and Graphical Statistics*, 29(3), 608-619. <https://doi.org/10.1080/10618600.2020.1714633>
- Estrada-Carmona, S., Pérez-Aranda, G. I. y Catzín-López, E. A. (2022). Infidelidad, deseo de infidelidad e inteligencia sexual en el suroeste de México. *Calidad de Vida y Salud*, 15(2), 29-42.
- Ferr, H. (2025). The normal distribution is not normal in psychological data: Moving beyond parametric dogma. *PLOS Mental Health*, 2(8), e0000403. <https://doi.org/110.1371/journal.pmen.0000403>
- Góngora, J. R., Rodríguez, J. A. R., Rodríguez, J. C. R. y Muñante, G. S. P. (2022). Construcción de un cuestionario de infidelidad y validación en jóvenes españoles. *Know and Share Psychology*, 3(1), 67-86. <https://doi.org/10.25115/kasp.v3i1.7897>
- Gregersen, E. (2022). Human sexuality in cross-cultural perspective. In D. Byrne & K. Kelley (Eds.), *Alternative approaches to the study of sexual behavior* (pp. 87-102). Psychology Press. <https://doi.org/10.4324/9781315792637>
- Guía Roji (2014). *Guía de calles Monterrey y su área metropolitana* (Veintitresava edición). Guía Roji.
- Gunver, M. G. (2022). Norm-referenced scoring on real data: a comparative study of GRiSTEN and STEN. *Sage Open*, 12(2). <https://doi.org/10.1177/21582440221091253>
- Gunver, M. G., Senocak, M. S., & Vehid, S. (2018). To determine skewness, mean and deviation with a new approach on continuous data. *Ponte: International Scientific Researches Journal*, 73(2), 64-79. <https://doi.org/10.21506/j.ponte.2018.2.5>
- Hendrick, S. S., Dicke, A., & Hendrick, C. (1998). The relationship assessment scale. *Journal of Social and Personal Relationships*, 15, 137-142. <https://doi.org/10.1177/0265407598151009>
- Hudson, W. W. (1992). *The WALMYR assessment scales scoring manual*. WALMYR Publishing.
- Isma, M. N. P. & Turnip, S. S. (2019). Personality traits and marital satisfaction in predicting couples' attitudes toward infidelity. *Journal of Relation-ships Research*, 10, e13. <https://doi.org/10.1017/jrr.2019.10>
- Jevtovic-Todorovic, V. (2022). Testosterone: much more for the brain than a sex hormone. *British Journal of Anaesthesia*, 128(3), 411-413. <https://doi.org/10.1016/j.bja.2022.01.002>
- King, B. M. (2022). The influence of social desirability on sexual behavior surveys: A review. *Archives of Sexual Behavior*, 51(3), 1495-1501. <https://doi.org/10.1007/s10508-021-02197-0>
- Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation* (Third edition). Sage.
- Krumpal, I. (2023). Social desirability bias and context in sensitive surveys. In *Encyclopedia of quality of life and well-being research* (pp. 6527-6532). Springer International Publishing. https://doi.org/10.1007/978-3-031-17299-1_4086
- Labrecque, L. T. & Whisman, M. A. (2017). Attitudes toward and prevalence of extramarital sex and descriptions of extramarital partners in the 21st century. *Journal of Family Psychology*, 31(7), 952-957. <https://doi.org/10.1037/fam0000280>
- Lakens, D. (2017). Equivalence tests: A practical

- primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355-362. <https://doi.org/10.1177/1948550617697177>
- Maxwell, A. M., Brucar, L. R., & Zilverstand, A. (2023). A systematic review of sex/gender differences in the multi-dimensional neurobiological mechanisms in addiction and their relevance to impulsivity. *Current Addiction Reports*, 10(4), 770-792. <https://doi.org/10.1007/s40429-023-00529-9>
- McKim, C. (2022). Percentile ranks/scores. Routledge. https://doi.org/10.4324/9780367198459-R_EPRW158-1
- Moral, J. (2011). La Escala de afecto positivo y negativo (PANAS) en parejas casadas mexicanas. *Revista CIENCIA Ergo Sum*, 18(2), 1-9. <https://cienciaergosum.uaemex.mx/article/view/7409>
- Moral, J. (2018). Validación de la Escala de Búsqueda de Sensaciones Sexuales en mexicanos casados o cohabitantes. *Revista Evaluar*, 18(3), 49-72. <https://doi.org/10.35670/1667-4545.v18.n3.22224>
- Moral, J. (2019). Frecuencia de relaciones sexuales y número de parejas concurrentes entre personas casadas o cohabitantes de Monterrey, Nuevo León, México. *Perspectivas Sociales*, 21(1), 115-145. <https://perspectivassociales.uaenl.mx/index.php/pers/article/view/94>
- Moral, J. (2020). Validación de una escala para medir actitud hacia la infidelidad en personas mexicanas casadas de Monterrey. *Revista Costarricense de Psicología*, 39(1), 107-132. <https://doi.org/10.22544/rcps.v39i01.06>
- Moral, J. (2026). Evidence of convergent validity of the BVAQ-20 in relation to the TAS-20 among Mexican adults and its association with social desirability. *Psychology*, 17(10), 941-967. <https://doi.org/10.4236/psych.2026.1710047>
- Moral, J., García, C. H. y Antona, C. J. (2012). Traducción y validación del Inventario Balanceado de Deseabilidad Social al Responder en una muestra probabilística de estudiantes universitarios mexicanos. *Revista de Psicología GEPU*, 3(2), 20-32.
- O'Rourke, N. & Cappeliez, P. (2002). Development and validation of a couples measure of biased responding: The Marital Aggrandize-ment Scale. *Journal of Personality Assessment*, 78(2), 301-320. https://dx.doi.org/10.1207/S15327752JPA7802_06
- Pandey, K., Narain, S., & Jahan, A. (2024). Development and Validation of the 14-Item Intentions Toward Infidelity Scale. *Deviant Behavior*, 45(10), 1333-1346. <https://doi.org/10.1080/01639625.2023.2294082>
- Paulhus, D. L. (1998). *Manual for balanced inventory of desirable responding: Version 7 (BIDR-7)*. Multi-Health Systems.
- Pazhoohi, F. (2022). Cultural differences and similarities in the nature of infidelity. In T. DeLecce & T. K. Shackelford (Eds.), *The Oxford handbook of infidelity* (pp. 349-374). Oxford University Press. <https://doi.org/10.1093/oxfordhob/9780197502891.013.28>
- Rodrigues, D. L., Aybar-Camposano, G. A., & Lopes, D. (2022). Stigmatization of consensual non-monogamous partners: Perceived endorsement of conservation or openness to change values vary according to personal attitudes. *Archives of Sexual Behavior*, 51(8), 3931-3946. <https://doi.org/10.1007/s10508-022-02368-7>
- Rodríguez, C. L. R. y Fernández, M. J. A. (2022). Acuerdos matrimoniales e infidelidad. Ayer y hoy. *GLOSSAE. European Journal of Legal History*, 2022(19), 474-549. <https://www.glossae.eu/glossaeojs/article/view/500>
- Rokach, A. & Chan, S. H. (2023). Love and infidelity: causes and consequences. *International Journal of Environmental Research and Public Health*, 20(5), 3904. <https://doi.org/10.3390/ijerph20053904>
- Romero, A., Cruz, C. y Díaz-Loving, R. (2008). Propuesta de un modelo bio-psico-socio-cultural de infidelidad sexual y emocional en hombres y mujeres. *Psicología Iberoamericana*, 16(2), 14-21. <https://doi.org/10.48102/pi.v16i2.281>
- Royston, J. P. (1993). A toolkit of testing for non-normality in complete and censored samples. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 42(1), 37-43. <https://doi.org/10.2307/2348109>
- Sahithya, B. R. & Kashyap, R. S. (2022). Sexual addiction disorder—A review with recent updates. *Journal of Psychosexual Health*, 4(2), 95-101.

- <https://doi.org/10.1177/26318318221081080>
- Shapiro, S. S. & Francia, R. S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67(337), 215-216. <https://doi.org/10.1080/01621459.1972.10481232>
- SMP, Sociedad Mexicana de Psicología (2010). Código ético del psicólogo (Quinta edición). Tri-l-las.
- Staples, T. L. (2023). Expansion and evolution of the R programming language. *Royal Society Open Science*, 10(4), 221550. <https://doi.org/10.1098/rsos.221550>
- Vanbelle, S., Engelhart, C. H., & Blix, E. (2024). A comprehensive guide to study the agreement and reliability of multi-observer ordinal data. *BMC Medical Research Methodology*, 24(1), 310. <https://doi.org/10.1186/s12874-024-02431-y>
- Weis, D. L. & Fenton, J. R. (1987). Marital exclusivity and the potential for future marital conflict. *Social Work*, 32(1), 45-49. <https://dx.doi.org/10.1093/sw/32.1.45>
- Weis, D. L., Rabinowitz, B., & Ruckstuhl, M. F. (1992). Individual changes in sexual attitudes and behavior within college-level human sexuality courses, *The Journal of Sex Research*, 29(1), 43-59. <https://dx.doi.org/10.1080/00224499209551633>
- Welch, B. L. (1947). The generalization of "Student's" problem when several different population variances are involved. *Biometrika*, 34(1-2), 28-35. <https://doi.org/10.2307/2332510>
- Wijisen, L. D., Borsboom, D., & Alexandrova, A. (2022). Values in psychometrics. *Perspectives on Psychological Science*, 17(3), 788-804. <https://doi.org/10.1177/17456916211014183>