

Validación Psicométrica en Población Española de la Escala The Survey of Autobiographical Memory y Equivalencia Factorial con Población Argentina

Psychometric Validation of The Survey Autobiographical Memory Scale in a Spanish Sample and Factor Equivalence with the Argentine Population

Alejandro Miguel-Alvaro¹, David Paniagua², Soledad Picco³, Iván Sánchez-Iglesias⁴
y Raimundo Aguayo-Estremera⁵

Resumen

La escala *The Survey of Autobiographical Memory* (SAM) es una medida de autoinforme fácil de administrar sobre las capacidades de la Memoria Autobiográfica (MA). Este trabajo tuvo como objetivo obtener evidencias de validez psicométrica del SAM en población española, específicamente de estructura interna y fiabilidad y de relación con otras variables (evidencias de validez convergente). Además, se pretendió analizar el grado de equivalencia factorial de la escala en una muestra de adultos españoles y argentinos. Se realizó un estudio de corte transversal contando con una muestra final de 3306 participantes de entre 18 y 65 años: 3044 personas argentinas y 262 personas españolas. En líneas generales, los resultados permiten concluir que la SAM tiene propiedades psicométricas adecuadas para su utilización con población española y que tiene un grado de equivalencia transcultural entre población argentina y española, ya que se presenta adecuada a nivel de invarianza métrica parcial.

Palabras clave: memoria autobiográfica, validación, medida, escala

Abstract

The Survey of Autobiographical Memory (SAM) scale is an easy-to-administer self-report measure of Autobiographical Memory (AM) abilities. The aim of this study was to obtain evidence of psychometric validity of the SAM in the Spanish population, specifically evidence of internal structure and reliability (factor) and evidence of relationship with other variables (evidence of convergent validity). In addition, another objective was to analyze the degree of factor equivalence of the scale in a sample of Spanish and Argentine adults. A cross-sectional study was carried out with a final sample of 3306 participants between 18 and 65 years of age: 3044 Argentines and 262 Spaniards. In general terms, the results allow us to conclude that the SAM has adequate psychometric properties for its use with the Spanish population and that it has a degree of cross-cultural equivalence between the Argentine and Spanish population, since it is adequate at the level of partial metric invariance.

Keywords: autobiographical memory, validation, measure, scale

¹ MSc. Contratado predoctoral FPU. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosagua, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid. Correo: alemigue@ucm.es.

² MSc. Personal Docente Investigador. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosagua, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid. Correo: davidpan@ucm.es

³ MSc. Contratada predoctoral. Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE - CONICET), Ciudad de Buenos Aires, Argentina. Correo: solepicco@gmail.com

⁴ PhD. Personal Docente Investigador. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosagua, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid. Correo: i.sanchez@psi.ucm.es (Autor de correspondencia)

⁵ PhD. Personal Docente Investigador. Facultad de Psicología, Universidad Complutense de Madrid, España. Campus de Somosagua, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid. Correo: raaguayo@ucm.es

Introducción

La memoria autobiográfica (MA) comprende el recuerdo de episodios ocurridos en la vida de un individuo. Siguiendo a Gilboa et al., (2018) la MA reúne diferentes aspectos mnésicos que subyacen a nuestra experiencia como la recuperación de componentes episódicos, semánticos, espaciales, y prospectivos (simulación de escenarios posibles o futuros). Además, la MA es fundamental para el funcionamiento humano ya que contribuye a la construcción del sentido de sí mismo de un individuo, a su capacidad de mantenerse orientado en el mundo y de perseguir objetivos de forma eficaz (Williams et al., 2022).

Al recordar memorias autobiográficas, el nivel de detalles internos (episódicos) y externos (semánticos) de los eventos varía entre sujetos, junto con otras características como la vivacidad, la emocionalidad, la perspectiva, los detalles sensoriales, etc. (Rubin et al., 2003). Esto implica la existencia de diferencias individuales. En función del sexo, las mujeres tienden a informar de una memoria emocional más específica (episódica) y más rica, y los hombres tienden a tener un mejor rendimiento en las tareas espaciales o de orientación (Karlsson et al., 2019). Además, la investigación es clara al señalar que determinadas características de la MA se relacionan con diversas condiciones psicopatológicas. Por ejemplo, la revisión de Crespo y Fernández-Lansac (2016) muestra que en el caso del Trastorno por Estrés Postraumático los recuerdos del acontecimiento traumático frente a otras memorias se caracterizan por su mayor viveza. Romano et al. (2020) sugirieron que el sesgo de la MA podría considerarse un proceso transdiagnóstico que perpetúa las representaciones negativas y distorsionadas de uno mismo. De hecho, en el ámbito clínico, las técnicas centradas en memorias autobiográficas están experimentando un desarrollo importante en los últimos años (e.g., Miguel-Alvaro et al., 2021). Además, cabe destacar que la autopercepción de baja calidad de la MA se ha relacionado con un mayor riesgo de desarrollar demencia (Mitchell et al., 2014). Por ello, la autopercepción de la capacidad de la MA puede ofrecer una visión única adicional de la experiencia de recordar como un marcador de la salud cognitiva.

La escala *The Survey of Autobiographical Memory* (SAM) fue creada por Palombo et al. (2013) como una medida de autoinforme fácil de administrar sobre las capacidades de la MA. Siguiendo a estos autores, en comparación con el número de tareas de memoria de laboratorio, existen pocas medidas para evaluar la autopercepción del funcionamiento de la memoria. Este hecho es especialmente relevante en el caso de la MA, ya que es un proceso mnésico difícil de medir adecuadamente en el laboratorio dado su carácter naturalista. Además, tal y como hemos comentado, la MA muestra diferencias individuales estables que no son adecuadamente captadas por dichas pruebas de laboratorio (e.g., McDermott et al., 2009).

Partiendo de 102 ítems compuestos a partir de otros instrumentos sobre MA y la literatura científica, Palombo et al. (2013) crearon el SAM de 26 ítems utilizando métodos multivariantes para condensar las respuestas de una muestra de 598 adultos sanos (de 18 a 65 años de edad) a cuatro dimensiones, conservando los ítems fiables para formar las subescalas Episódica, Semántica, Espacial y Futura. La escala Episódica (8 ítems) se refiere a los eventos de memoria contextualizados (es decir, "qué", "dónde" y "cuándo"), la escala Semántica (6 ítems), a la memoria para el conocimiento abstracto y los hechos (v.g., nombres de familiares), la escala Espacial (6 ítems), a la memoria de lugares y la capacidad de navegar por ellos (v.g., encontrar una ruta alternativa para ir al trabajo), y la escala Futura (6 ítems), a la capacidad de simular mentalmente acontecimientos en el futuro dentro de un contexto específico (v.g., imaginar unas vacaciones futuras). Por tanto, el SAM evalúa la capacidad autoinformada de recordar o simular instancias específicas del pasado o del futuro personal dentro de un contexto espaciotemporal, junto con los detalles o hechos externos asociados.

La adaptación al español de la SAM por Picco et al. (2020) incluyó una evaluación psicométrica con población argentina para comprobar su estructura factorial. Replicaron los análisis de Palombo et al. (2013) con una muestra de 1879 adultos sanos e identificaron cuatro dimensiones similares a las halladas por los autores originales. Además, Picco et al. (2020) probaron varios modelos de análisis factorial confirmatorio y

encontraron que los modelos que contenían cuatro factores, o cuatro factores más un factor general adicional (SAM-total) exhibieron ajustes adecuados. Por otro lado, Setton et al., (2022) llevaron a cabo una evaluación psicométrica crítica del SAM. Contando con una muestra de 209 adultos y 193 ancianos sanos estadounidenses, hallaron evidencias de fiabilidad del SAM como medida de autoinforme de la capacidad percibida de recuerdo. Además, observaron una alta consistencia interna en las subescalas con la excepción de la Semántica. Las pruebas de independencia entre las subescalas mostraron que los ítems de Episódica y Semántica tenían una pobre correspondencia con las respectivas subescalas, y se observó una buena correspondencia entre los ítems de las escalas Futura y Espacial. Además, la SAM mostró correlaciones bajas con el rendimiento de la MA medido a través de la Entrevista Autobiográfica (Levine et al., 2002). No obstante, este resultado se asoció con la vivacidad de los eventos medidos con el mismo instrumento. Estos autores concluyeron recomendando cautela en el uso e interpretación del SAM como medida de la MA a la espera de más investigación psicométrica sobre ella. Por último, Gaviria-Castaño et al. (2018) estudiaron la estructura factorial del SAM con población colombiana, observando que la mejor configuración era la de cuatro factores y obteniendo índices de fiabilidad aceptables.

Considerando estos antecedentes, destacamos que hasta la fecha no se han registrado estudios psicométricos en España, ni hay trabajos que analicen modelos de medida simultáneos entre muestras de diferentes países. Por ello, este estudio tuvo como objetivo obtener evidencias de validez psicométrica del SAM en población española en las siguientes características: estructura interna y fiabilidad y evidencias de relación con otras variables (validez convergente). Además, se analizó el grado de equivalencia factorial de la escala en una muestra de adultos españoles y argentinos. A partir de estos resultados se esperó encontrar evidencias de validez en ambos países que permitieran medir las diferencias individuales de la autopercepción de la calidad de la MA para la posible generación de diseños de investigación más exhaustivos en torno a la MA.

Método

Participantes

Este estudio cuenta con una muestra recogida para el estudio de las evidencias de validez en población española junto con los datos de los participantes recogidos en Argentina en la adaptación de Picco et al. (2020). Los criterios de inclusión fueron que las personas hubieran nacido y vivido en España o Argentina respectivamente; como criterio de exclusión se eliminaron las personas que fueran menores de edad o tuvieran más de 65 años. Por ello, se excluyó a 37 participantes españoles. La muestra argentina ya estaba depurada bajo estos criterios.

La muestra final se compuso de 3306 participantes: 3044 personas argentinas y 262 personas españolas. La muestra española contó con 178 mujeres (67.9%) y 84 hombres (32.1%), con una media de edad de 35.2 años ($DT=20.5$). En cuanto al nivel de estudios 71 personas (27.1%) terminaron educación secundaria y 191 personas (19.9%) estudios universitarios. La muestra argentina contó con 2431 mujeres (79.9%), 607 hombres (19.9%) y 5 personas (0.2%) que no se incluían en estas categorías de género. La media de edad fue de 32.2 años ($DT=10.4$); nivel de estudios: 8 personas (0.3%) terminaron estudios primarios, 349 personas (11.5%) educación secundaria y 2320 personas estudios universitarios (76.2%).

Instrumentos

The Survey of Autobiographical Memory (SAM)

Esta escala es un instrumento de autoinforme que se utiliza para medir las capacidades de MA autopercibidas (Palombo et al., 2013). Se puntúa en una escala Likert de cinco puntos, desde muy en desacuerdo hasta muy de acuerdo. Las puntuaciones de las subescalas Episódica, Semántica, Espacial, Futura, y la puntuación global SAM-total se calcularon de acuerdo con el protocolo original. Se utilizó la versión en castellano de Picco et al. (2020). A mayor puntuación mejores capacidades autopercibidas.

Cuestionario de Salud del Paciente-2 (PHQ-2; Kroenke et al., 2003)

Este cuestionario es ampliamente utilizado para el cribado de la depresión; tiene dos ítems

sobre la frecuencia de la depresión y la anhedonia en las últimas dos semanas, puntuando cada respuesta de 0 ("nada") a 3 ("casi todos los días"). En el estudio original (Kroenke et al., 2003) se estableció una puntuación de 3 como punto de corte óptimo para fines de cribado. El PHQ-2 mostró buenas propiedades psicométricas, incluyendo una alta sensibilidad (83%) y especificidad (92%). Al ser una herramienta de dos ítems se calculó la fiabilidad con el coeficiente de Spearman-Brown ($r_{xx}=.60$) como se sugiere en Eisinga, Grotenhuis y Pelzer (2013).

Satisfaction with Life Scale (SWLS; Pavot & Diener, 1993)

Esta escala consta de cinco ítems para evaluar la satisfacción vital general. Solicita a los individuos que valoren su grado de acuerdo con cada ítem en una escala Likert de 7 puntos, desde "absolutamente en desacuerdo" hasta "absolutamente de acuerdo". Las puntuaciones en esta escala van de 5 a 35 puntos, y puntuaciones más altas indican niveles más altos de satisfacción vital. Investigaciones anteriores han encontrado alfas de Cronbach oscilando entre .79 y .89 (Pavot y Diener 2008). Se utilizó la versión española de la escala cuyo valor del alfa de Cronbach fue .88 (Vázquez et al. 2013).

Procedimiento

Para la obtención de los datos con muestra española se seleccionaron las escalas con la que se estudiaría la relación convergente (síntomatología depresiva, PHQ-2, y satisfacción con la vida, SWLS), se generó un cuestionario sociodemográfico preguntando por las variables género, nivel de estudios y si la persona había nacido y crecido en España. Se procedió a la implementación en la aplicación informática *Microsoft Forms*. La recogida de muestra fue no probabilística a través de redes sociales. Se contactó con los autores del artículo de adaptación a la población argentina (Picco et al., 2020) para proponerles participar en el presente estudio aportando la muestra argentina y revisión del manuscrito.

Análisis de Datos

Todos los análisis se realizaron con R Studio 4.0.4 de R (R Core Team, 2021), usando el paquete *lavaan* (Rosseel, 2012) para los análisis

factoriales. Se partió de la matriz de correlaciones policóricas. Como método de estimación de parámetros se usó WLSMV, dado el número de categorías de respuesta y los índices de asimetría y curtosis. No había presencia de valores perdidos. Tanto el código como los datos se encuentran disponibles a petición al autor de correspondencia.

Los modelos puestos a prueba fueron los siguientes: (a) 1-factor: modelo unidimensional donde todos los ítems saturan en un único factor; (b) 3-factores: modelo de tres factores que tiene en cuenta la relación entre la memoria episódica y semántica (Greenberg & Verfaellie, 2010). Se especificó que los ítems de dichos factores (Episódica y Semántica) saturaban en un único factor mientras que el resto de ítems saturaban en los restantes: Espacial y Futuro; (c) 4-factores ortogonal: modelo de cuatro factores donde las correlaciones se forzaban a ser nulas; (d) 4-factores oblicuo: Modelo de cuatro factores donde se permiten las correlaciones entre los factores; (e) Segundo Orden: modelo de segundo orden donde los cuatro factores quedan recogidos en un factor de orden superior.

Para valorar el ajuste de cada modelo individualmente se tuvieron en cuenta los siguientes indicadores: el estadístico χ^2 , el índice de ajuste comparativo (*CFI*), el Tucker-Lewis Index (*TLI*), Root Mean Square Error of Approximation (*RMSEA*) y Standardized Root Mean-Square (*SRMR*). Para los índices *CFI* y *TLI* se consideran como un ajuste adecuado del modelo valores superiores a .90 (Schumacher et al., 1996), mientras que para el *RMSEA* se considera un ajuste razonable valores inferiores a .08 (Browne & Cudeck, 1993). Para determinar el mejor modelo estadístico se usó la prueba de diferencias de ji-cuadrado.

Tras la selección del modelo factorial se calculó el índice de discriminación ítem-test corregido, y la fiabilidad de las escalas utilizando el coeficiente alfa de Cronbach; considerando las limitaciones de este coeficiente (Dunn, et al., 2014; Green & Yang, 2009; Revelle & Zinbarg, 2009; Sijtsma, 2009; Toro, et al., 2022; Trizano-Hermosilla & Alvarado, 2016) también se calculó el coeficiente omega (McDonald, 1999).

Las medias de las puntuaciones de cada factor (y el SAM-total) se correlacionaron con las escalas

Tabla 1. Estadísticos Descriptivos de los Ítems en Ambas Muestras

Ítem	Muestra española					Muestra argentina				
	<i>M</i>	<i>D</i>	<i>DT</i>	<i>As</i>	<i>Cu</i>	<i>M</i>	<i>D</i>	<i>DT</i>	<i>As</i>	<i>Cu</i>
Episódica 1	3.64	.73	1.21	-.62	-.69	3.64	.73	1.29	-.48	-1.11
Episódica 2	3.45	.69	1.17	-.42	-.89	3.74	.74	1.26	-.61	-.95
Episódica 3	3.70	.75	1.07	-.72	-.23	3.58	.72	1.24	-.61	-.69
Episódica 4	2.80	.56	1.35	.10	-1.30	3.12	.62	1.43	-.19	-1.34
Episódica 5	3.31	.66	1.23	-.33	-.97	3.48	.70	1.30	-.47	-.96
Episódica 6	3.50	.77	1.15	-.61	-.56	3.40	.68	1.25	-.42	-.89
Episódica 7	2.39	.48	1.28	.46	-1.05	2.63	.53	1.35	.22	-1.28
Episódica 8	3.17	.63	1.17	-.37	-.98	3.07	.62	1.24	-.23	-1.08
Semántica 1	3.66	.73	1.15	-.70	-.42	3.64	.73	1.08	-.67	-.20
Semántica 2	3.33	.67	1.27	-.28	-1.12	3.35	.67	1.30	-.22	-1.21
Semántica 3	3.00	.60	1.41	.01	-1.36	2.90	.58	1.50	.07	-1.47
Semántica 4	3.06	.61	1.31	-.15	-1.14	3.01	.60	1.43	-.06	-1.36
Semántica 5	3.64	.73	1.11	-.46	-.81	3.64	.83	1.22	-.51	-.89
Semántica 6	3.47	.70	1.28	-.44	-.95	3.44	.69	1.34	-.45	-1.03
Espacial 1	3.07	.61	1.46	-.15	-1.38	3.25	.65	1.45	-.30	-1.26
Espacial 2	3.67	.73	1.34	-.79	-.63	3.86	.77	1.30	-.97	-.28
Espacial 3	2.95	.59	1.39	.02	-1.31	3.18	.64	1.39	-.12	-1.30
Espacial 4	4.00	.80	1.30	-1.10	-.10	3.96	.80	1.34	-1.01	-.35
Espacial 5	3.98	.80	1.25	-1.05	-.13	4.24	.85	1.11	-1.46	1.16
Espacial 6	4.33	.87	.90	-1.70	3.12	4.31	.86	.94	-1.64	2.68
Futura 1	3.82	.77	1.03	-.87	.23	3.83	.77	1.14	-.89	.06
Futura 2	3.76	.75	1.08	-.92	.27	3.62	.72	1.18	-.71	-.30
Futura 3	3.72	.74	1.06	-.91	.35	3.38	.68	1.21	-.45	-.73
Futura 4	4.11	.82	.94	-1.29	1.74	4.03	.80	1.02	-1.20	1.12
Futura 5	4.06	.82	.90	-1.25	1.90	3.90	.78	1.08	-1.04	.56
Futura 6	4.14	.83	1.04	-1.19	.72	4.07	.81	1.18	-1.06	-.03

Nota. *D*: Índice de dificultad; *As*: Asimetría; *Cu*: Curtosis.

Tabla 2. Resultados de los Análisis Factoriales con la Muestra Española

	χ^2 (<i>gl</i>);	<i>p</i>		$\Delta\chi^2$ (Δ <i>gl</i>)	<i>p</i>	CFI	RMSEA	SRMR	TLI
Modelo 1	1484.89 (299)	<i>p</i> <.01				.690	.123	.147	.663
Modelo 3	386.10 (296)	<i>p</i> <.01	M1 - M3	132.31 (3)	<i>p</i> <.01	.976	.034	.069	.974
Modelo 4 oblicuo	321.06 (293)	<i>p</i> =.13	M3 - M4 Obl	19.22 (3)	<i>p</i> <.01	.993	.034	.069	.974
Modelo 4 ortogonal	1344.81 (299)	<i>p</i> <.01	M4 Ort - M4 Obl	73.34 (6);	<i>p</i> <.01	.727	.116	.129	.703
MSO	324.55 (295)	<i>p</i> =.11	MSO - M4 Obl	0.56 (2)	<i>p</i> =.75	.992	.020	.064	.991

Nota. *gl*=grados de libertad; $\Delta\chi^2$ (Δ *gl*)=Cambio en χ^2 (cambio en *gl*); M1=Modelo unidimensional; M3=Modelo de tres factores; M4 Obl=Modelo 4 oblicuo; M4 Ort=Modelo 4 ortogonal; MSO=Modelo segundo orden.

de depresión (PHQ-2) y satisfacción con la vida (SWLS) y con la edad. Además, se calcularon contrastes de medias entre las puntuaciones promedio de cada factor (y el total) con género.

Para finalizar se puso a prueba la invarianza (base, débil, fuerte y estricta) del modelo 4 oblicuo en ambas poblaciones teniendo en cuenta los criterios de χ^2 y el criterio de cambio de CFI (Cheung & Rensvold, 2002).

Para estimar el tamaño de la muestra requerido para este estudio actual, se realizó un análisis a priori del tamaño muestral usando un software online (Soper, 2022). El tamaño de la muestra se calculó usando un tamaño esperado del efecto de $r=.30$, con $\alpha=.05$ y un nivel de potencia estadística $1-\beta=.80$, siendo 5 el número de variables latentes y 26 el número de variables observadas. El tamaño muestral mínimo para

detectar el efecto planteado es $N=150$, un valor menor a nuestro tamaño de muestra.

Resultados

Antes de explorar los distintos modelos factoriales se calcularon los estadísticos descriptivos para ambas muestras (Tabla 1). Se calcularon correlaciones de Pearson entre los índices de dificultad ($r=.94$, $p<.001$), desviación típica ($r=.88$, $p<.001$), asimetría ($r=.93$, $p<.001$) y curtosis ($r=.90$, $p<.001$) de ambas muestras. Estos resultados indican una alta relación en el centro, dispersión y forma de las puntuaciones de los ítems en las dos muestras.

Tabla 3. Índices de Fiabilidad y Correlaciones Entre Factores en Ambas Muestras

		Episódica $\alpha=.85$ $\omega=.88$	Semántica	Espacial
Semántica	$\alpha=.85$ $\omega=.88$.298*		
Espacial	$\alpha=.83$ $\omega=.88$.115	.096	
Futuro	$\alpha=.87$ $\omega=.92$.189*	.108*	.075

Nota. Las correlaciones del triángulo superior corresponden a la muestra argentina, las del triángulo inferior a la española. $p<.050$, ** $p<.010$

Tabla 4. Pesos Factoriales e Índices de Discriminación de los ítems

Ítems	Muestra española		Muestra argentina	
	r_{jxc}	λ	r_{jxc}	λ
Episódica 1	.58	.66	.51	.57
Episódica 2	.55	.59	.43	.49
Episódica 3	.43	.47	.39	.44
Episódica 4	.58	.61	.51	.53
Episódica 5	.67	.73	.64	.75
Episódica 6	.73	.80	.66	.75
Episódica 7	.51	.59	.48	.54
Episódica 8	.60	.65	.52	.60
Semántica 1	.41	.49	.32	.40
Semántica 2	.42	.55	.28	.38
Semántica 3	.41	.47	.38	.51
Semántica 4	.43	.46	.41	.51
Semántica 5	.36	.45	.26	.37
Semántica 6	.56	.76	.42	.60
Espacial 1	.72	.79	.72	.75
Espacial 2	.75	.85	.75	.79
Espacial 3	.48	.52	.48	.37
Espacial 4	.75	.84	.75	.75
Espacial 5	.69	.76	.69	.65
Espacial 6	.25	.27	.25	.10
Futuro 1	.67	.70	.67	.74
Futuro 2	.68	.71	.67	.75
Futuro 3	.68	.73	.64	.72
Futuro 4	.57	.65	.52	.54
Futuro 5	.74	.80	.62	.68
Futuro 6	.69	.78	.62	.68

Resultados de los Análisis Factoriales y Fiabilidad

Los resultados de los análisis factoriales en la muestra española y sus comparativas se presentan en la Tabla 2. Como se puede observar tanto el modelo de 4 factores correlacionados como el modelo de segundo orden tienen un buen nivel de ajuste y son significativamente mejores modelos que los demás.

Las correlaciones entre factores en la muestra española fueron: ($r=.30$, $p<.001$) entre los factores Episódica y Semántica, ($r=.19$; $p<.001$) entre los factores Episódica y Futura y ($r=.11$; $p<.001$) entre los factores Semántica y Futura (Tabla 3).

Como se puede ver en la Tabla 4, para ambos países todos los pesos factoriales e índices de

discriminación de los ítems eran altos. Sin embargo, estos pesos son mejores en la muestra española, donde sólo el ítem 6 presenta pesos menores de .30.

Relaciones con otras Variables

Se correlacionaron los totales de los cuatro factores y el total del test con diferentes variables (Tabla 5). El nivel de sintomatología depresiva (PHQ-2) presentó una relación negativa con la escala Episódica ($r=-.17$, $p<.001$), Semántica ($r=-.21$, $p<.001$) y SAM-total ($r=-.23$, $p<.001$). No se han encontrado evidencias de que el resto de los factores se relacionen con el nivel de ansiedad, de depresión ni con la satisfacción con la vida medida mediante el SWLS.

Tabla 5. Correlación con Otras Variables en la Muestra Española

	Episódica	Semántica	Espacial	Futuro	Total
Depresión (PHQ-2)	-.174**	-.210**	-.112	-.089	-.228*
Satisfacción con la vida (SWLS)	.089	.104	.098	-.031	.106
Edad	-.049	-.239**	.058	-.134*	-.125*

* $p < .050$, ** $p < .010$.

Tabla 6. Invarianza Factorial del SAM Entre Población Española y Argentina

	χ^2 (gl)	p	Comparación	$\Delta\chi^2$ (Δ gl)	p	CFI	RMSEA	SRMR	TLI
Base	2665.40 (586)	$p < .001$.945	.046	.048	.939
Débil	2738.90 (608)	$p < .001$	base-débil	24.97 (22)	$p = .298$.944	.046	.049	.94
Fuerte	2805.12 (630)	$p < .001$	débil-fuerte	46.93 (22)	$p = .001$.942	.046	.049	.941
Fuerte-p*	2765.36 (627)	$p < .001$	débil-parcial	27.18 (19)	$p = .100$.943	.045	.049	.941
Estricta	2916.66 (656)	$p < .001$	parcial-fuerte	51.53 (29)	$p < .001$.940	.046	.050	.941

Nota. Fuerte-p = Invarianza fuerte parcial. * La invarianza parcial libera los pesos de los ítems 3 y 4 de SAM-Semántica y 6 de SAM-Espacial.

Se encontró una relación negativa de la edad con la escala Semántica ($r = -.24$, $p < .001$), Futura ($r = -.13$, $p = .031$) y SAM-total ($r = -.13$, $p < .043$). Las diferencias de género exploradas solo fueron significativas en el factor de memoria episódica, $t(213.69) = 6.7$, $p < .001$, donde los hombres obtuvieron puntuaciones significativamente superiores, $M_{mujeres} = 20.6$ y $M_{hombres} = 24.9$; y en Futura, $t(143.43) = 2.1$, $p = .04$, donde las mujeres obtuvieron puntuaciones significativamente superiores, $M_{mujeres} = 24.0$ y $M_{hombres} = 22.7$.

Invarianza Factorial entre Población Española y Argentina

Se realizó la invarianza factorial del modelo de 4 factores correlacionados, comparando la muestra argentina y la española. Los resultados del contraste realizado para la invarianza factorial de la escala se presentan en la Tabla 6. El modelo base de la invarianza (que prueba la estructura factorial por separado para ambos grupos) tuvo un buen ajuste. El siguiente paso puso a prueba la invarianza débil (métrica). El modelo muestra adecuados índices de ajuste, similares al modelo base; el incremento en χ^2 resultó ser no significativo y el incremento de CFI menor a .001. Por tanto la escala presenta evidencia de, al menos, invarianza débil (métrica). En el siguiente paso el modelo de invarianza fuerte no se cumple: el incremento en χ^2 fue significativamente distinto de cero y el incremento de CFI superior a .001. Por ello se exploraron los índices de modificación, donde se decidió liberar los pesos factoriales de los ítems 3 y 4 del factor Semántica y el peso del Ítem 6 del factor Espacial. Se puso a prueba la invarianza fuerte (escalar) parcial; el incremento

en χ^2 resultó ser no significativo y el incremento de CFI menor a .001. Se puede asumir que la escala presenta evidencia de invarianza fuerte (escalar) parcial. El último paso, la invarianza estricta (medias latentes) no se cumple, tanto en el incremento de en χ^2 fue significativamente distinto de cero y el incremento de CFI superior a .001.

Discusión

El objetivo del presente estudio fue obtener evidencias de validez y fiabilidad del SAM en población española, y analizar el grado de equivalencia factorial de la escala en una muestra de adultos españoles y argentinos.

Los resultados del análisis factorial confirmatorio en la muestra española sugieren la adecuación de la estructura teórica de cuatro factores y uno general propuesta por Palombo et al. (2013). La correlación más alta y significativa entre dos factores fue entre Episódica y Semántica ($r = .30$, $p < .001$). No obstante, dicha correlación no es elevada en ninguna de las dos poblaciones. En ambas, el ajuste del modelo de cuatro factores es significativamente mejor que el de tres factores. Estas dos circunstancias entran en contradicción con el trabajo de Setton et al. (2022) en el que recomiendan no considerar los factores Episódica y Semántica de manera independiente. De hecho, dicho estudio exhibe que el ajuste de la solución de tres factores para el SAM era casi idéntico al ajuste de la solución de cuatro factores. No obstante, es coherente que ambas escalas sean las que correlacionen con más fuerza si se toma en cuenta que hay autores que han propuesto que

estos dos constructos mnésicos no son tan diferenciales como se concibió clásicamente (Tulvin, 1972). La memoria semántica se desarrolla tempranamente (e.g., Newcombe et al., 2007) y proporciona un andamiaje a partir del cual pueden emerger los procesos episódicos (Irish & Piguet, 2013) una vez que también se han desarrollado el sentido del tiempo subjetivo y el yo (v.g., Martín-Ordas et al., 2014). Futuros estudios deberán arrojar luz sobre si estas dos subescalas deben ser interpretadas como medidas independientes y específicas (tal y como se deriva de los análisis del presente trabajo y del de Picco et al., 2020) o por el contrario formar parte de un mismo factor (Setton et al., 2022).

En lo concerniente a los análisis de fiabilidad de la SAM con muestra española, encontramos que los ítems poseen una consistencia interna adecuada ya que la contribución de cada enunciado a la medición del constructo obtiene niveles adecuados. Únicamente el Ítem 6 (perteneciente al factor Espacial) aumentaría el alfa de la escala al ser eliminado (además de tener un bajo índice de discriminación y peso factorial en ambas muestras). Se torna de interés que futuras investigaciones aborden las problemáticas de este ítem, valorar si es un problema en la redacción, si al eliminarlo se pierde parte del contenido del factor y si tras su eliminación la escala sigue obteniendo buenos índices psicométricos.

Respecto a la relación del SAM con otras variables el nivel de sintomatología depresiva (PHQ-2) presentó correlaciones negativas con las escalas Episódica, Semántica y SAM-total (las dos últimas por encima de .20). Aunque en general las correlaciones tienen valores bajos, este dato es congruente con la ampliamente contrastada relación entre depresión y sobregeneralización de la MA. Se ha observado en gran cantidad de estudios que a los sujetos que padecen problemas depresivos les cuesta más recuperar recuerdos específicos y detallados del pasado (Williams et al., 2007). Por otro lado, el hecho de que no se hallaron relaciones significativas entre las puntuaciones del SAM y el cuestionario de satisfacción con la vida (SWLS) nos puede estar indicando que la autopercepción sobre la capacidad de la MA medida mediante el SAM es independiente de la satisfacción que

tenga la persona respecto a los eventos de su vida pasada. En lo concerniente a la relación entre las puntuaciones del SAM y la edad, los resultados muestran que las personas de más edad tienden a tener puntuaciones más bajas en la escala. Este hecho es coherente con el conocimiento en torno a los efectos de la edad en el empeoramiento de la memoria declarativa de los sujetos (e.g., Melendez et al., 2020). Finalmente, en lo referente al sexo biológico de los participantes, se hallaron diferencias significativas en las puntuaciones Episódica (siendo superiores las de los hombres) y en Futura (siendo superiores las de las mujeres). La investigación previa afirma que las mujeres recuerdan los acontecimientos pasados con mucha más rapidez que los hombres y que estos recuerdos son más detallados y emocionalmente intensos (Andreano et al., 2009). Por ello el dato en la escala Episódica difiere con el estado de la cuestión. Una posible explicación estribaría en el hecho de que, tal y como afirman Setton et al. (2022), el SAM únicamente es una medida válida sobre la autopercepción del desempeño de la MA y no sobre el desempeño real en esta. Es decir, puede ser que los varones confíen más en la exactitud de sus recuerdos que las mujeres. Este aspecto no es baladí si tomamos en cuenta los resultados del meta-análisis de Williams et al. (2014) en el cual se halló que las personas con quejas subjetivas de memoria pero sin déficits objetivos tienen un mayor riesgo de desarrollar deterioro cognitivo leve y demencia. Es decir, el SAM se puede constituir como un marcador de salud cognitiva ya que podría ser capaz de captar esas quejas subjetivas de memoria. Futuras investigaciones deberán clarificar este aspecto.

Por otro lado, respecto al análisis de la invarianza factorial de la SAM (considerando la muestra española y la argentina), el primer modelo contrastado mostró que el número de factores y sus ítems correspondientes fueron los mismos en ambas poblaciones. El segundo modelo implementado arrojó como resultado que las cargas factoriales de los ítems eran invariantes en ambas muestras, lo que conocemos como invarianza parcial, que permite una mayor flexibilidad que los modelos de invarianza completa (Byrne et al., 1989). Por otro lado, no existieron evidencias de invarianza escalar. Este aspecto supone que existan dificultades a la hora

de realizar comparaciones válidas entre los promedios de los ítems y, por lo tanto, siendo más exhaustivos, entre las comparaciones de las medias de las variables latentes de los factores de la escala.

Los resultados de esta investigación exponen las características psicométricas del SAM en población española y el grado de estabilidad transcultural de la escala en adultos sanos argentinos y españoles. La principal limitación del presente estudio estriba en el tipo de muestreo, ya que en ambas poblaciones fue no probabilístico. Teniendo en cuenta el número de ítems por factor y sus altas cargas factoriales, consideramos adecuado el tamaño muestral para el AFC según los criterios recopilados por Kyriazos (2018), sin embargo existen diferencias en el tamaño muestral entre las poblaciones españolas y argentinas que podrían afectar a los resultados de la invarianza. En este sentido sería de interés generar estudios similares considerando otras poblaciones hispanohablantes de latinoamérica asegurando que el tamaño muestral no pueda constituir un problema para este tipo de análisis en ambas poblaciones. En líneas generales, los resultados permiten concluir que la SAM tiene propiedades psicométricas adecuadas para su utilización con población española y que tiene un grado de equivalencia transcultural entre población argentina y española ya que se presenta adecuada a nivel de invarianza métrica parcial.

Agradecimientos

Queremos agradecer a Rodrigo S. Fernández y María E. Pedreira su participación en la captación de la muestra argentina.

Referencias

- Andreano, J. M., & Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learning y Memory*, *16*(4), 248-266. <https://doi.org/10.1101/lm.918309>
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Testing structural equation models*, Edited by: Bollen, K. A. and Long, J. S. 136-62. Sage.
- Byrne, M., Shavelson, R., & Muthén, B. (1989). Testing for equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456-466. <https://doi.org/10.1037/00332909.105.3.456>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Crespo, M., y Fernández-Lansac, V. (2016). Memory and narrative of traumatic events: A literature review. *Psychological Trauma: Theory, Research, Practice, and Policy*, *8*(2), 149-156. <https://doi.org/10.1037/tra0000041>
- Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399-412. <https://doi.org/10.1111/bjop.12046>
- Eisinga, R., Grotenhuis, M. T. & Pelzer, B. (2013) The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown?. *Int J Public Health* *58*, 637-642. <https://doi.org/10.1007/s00038-012-0416-3>
- Ferrando, P. J., Lorenzo-Seva, U., Hernández-Dorado, A., & Muñiz, J. (2022). Decálogo para el análisis factorial de los ítems de un test. *Psicothema*, *34*, 7-17. <https://doi.org/10.7334/psicothema2021.456>
- Gaviria-Castaño, G., Dominguez-Lara, S., & Tamayo-Agudelo, W. (2018). Factorial structure of The Survey Autobiographical Memory (SAM) in a sample of Colombian population. *Revista Mexicana de Neurociencia*, *19*(1), 23-36.
- Gilboa, A., Rosenbaum, R. S., & Mendelsohn, A. (2018). Autobiographical memory: From experiences to brain representations. *Neuropsychologia*, *110*, 1-6. <https://doi.org/10.1016/j.neuropsychologia.2017.12.020>
- Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, *74*(1), 155-167. <https://doi.org/10.1007/S11336-008-9099-3>
- Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology.

- Journal of the International Neuropsychological Society*, 16(5), 748-753.
<https://doi.org/10.1017/S1355617710000676>
- Irish, M., & Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. *Frontiers in Behavioral Neuroscience*, 7(April), 27.
<https://doi.org/10.3389/fnbeh.2013.00027>
- Karlsson, K. P., Sikström, S., Jönsson, F. U., Sendén, M. G., & Willander, J. (2019). Gender differences in autobiographical memory: Females latently express communality more than do males. *Journal of Cognitive Psychology*, 31(7), 651-664.
<https://doi.org/10.1080/20445911.2019.1659281>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 41(11), 1284-92.
<https://doi.org/10.1097/01.MLR.0000093487.78664.3C>
- Kyriazos, T. (2018) Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, 9, 2207-2230.
<https://doi.org/10.4236/psych.2018.98126>
- Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, 17(4), 677-689.
<https://doi.org/10.1037/0882-7974.17.4.677>
- Martin-Ordas, G., Atance, C. M., & Caza, J. S. (2014). How do episodic and semantic memory contribute to episodic foresight in young children? *Frontiers in Psychology*, 5, 1-11.
<https://doi.org/10.3389/fpsyg.2014.00732>
- McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia*, 47(11), 2290-2298.
<https://doi.org/10.1016/j.neuropsychologia.2008.12.025>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates, Inc.
- Meléndez, J. C., Agustí, A. I., & Satorres, E. (2020). ¿Existe un sesgo de sobregeneralización y un efecto de positividad relacionado con la edad? *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 2(55).
<https://doi.org/10.21865/RIDEP54.1.01>
- Miguel-Alvaro, A., Guillén, A. I., Contractor, A. A., & Crespo, M. (2021). Positive memory intervention techniques: A scoping review. *Memory*, 29(6), 793-810.
<https://doi.org/10.1080/09658211.2021.1937655>
- Mitchell, A. J., Beaumont, H., Ferguson, D., Yadegarfar, M., & Stubbs, B. (2014). Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: Meta-analysis. *Acta Psychiatrica Scandinavica*, 130(6), 439-451.
<https://doi.org/10.1111/acps.12336>
- Newcombe, N. S., Lloyd, M. E., & Ratliff, K. R. (2007). Development of episodic and autobiographical memory: A cognitive neuroscience perspective. *Advances in Child Development and Behavior*, 35, 37-85.
<https://doi.org/10.1016/B978-0-12-009735-7.50007-4>
- Palombo, D. J., Williams, L. J., Abdi, H., & Levine, B. (2013). The survey of autobiographical memory (SAM): A novel measure of trait mnemonics in everyday life. *Cortex*, 49(6), 1526-1540.
<https://doi.org/10.1016/j.cortex.2012.08.023>
- Pavot W., & Diener, E. (1993). The affective and cognitive contest of self reports measures of subjective well-being. *Social Indicators Research*, 28, 1-20.
<http://dx.doi.org/10.1007/BF01086714>
- Pavot W., & Diener, E. (2008). The Satisfaction with Life Scale and the emerging construct of life satisfaction. *Journal of Positive Psychology*, 3, 137-152.
<http://dx.doi.org/10.1080/17439760701756946>
- Picco, S., Pedreira, M. E., & Fernández, R. S. (2020). Psychometric validation of the survey of autobiographical memory: Confirmatory factor analysis and network analysis. *Memory*, 28(8), 1037-1050.
<https://doi.org/10.1080/09658211.2020.1812662>
- R Core Team (2021). R: A language and environment for statistical computing. R

- Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*(1), 145-154. <https://doi.org/10.1007/S11336-008-9102-Z>
- Romano, M., Ma, R., Moscovitch, M., & Moscovitch, D. A. (2020). Autobiographical memory bias. In *Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms* (pp. 183-202). American Psychological Association. <https://doi.org/10.1037/0000150-011>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory y Cognition*, *31*(6), 887-901. <https://doi.org/10.3758/BF03196443>
- Schumacher, R. E., & Lomax, R. G. (1996). *A beginner's guide to SEM*. Manwah.
- Setton, R., Lockrow, A. W., Turner, G. R., & Spreng, R. N. (2022). Troubled past: A critical psychometric assessment of the self-report Survey of Autobiographical Memory (SAM). *Behavior Research Methods*, *54*(1), 261-286. <https://doi.org/10.3758/s13428-021-01604-7>
- Sijtsma, K. (2009). Reliability beyond theory and into practice. *Psychometrika*, *74*(1), 169-173. <https://doi.org/10.1007/S11336-008-9103-Y>
- Soper, D. S. (2022). *A-priori Sample Size Calculator for Structural Equation Models* [Software]. Available from <https://www.danielsoper.com/statcalc>
- Toro, R., Peña-Sarmiento, M., Avendaño-Prieto, B. L., Mejía-Vélez, S., & Bernal-Torres, A. (2022). Análisis empírico del Coeficiente alfa de Cronbach según opciones de respuesta, muestra y observaciones atípicas. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, *63*(2), 17-30. <https://doi.org/10.21865/RIDEP63.2.02>
- Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, *7*, 769. <https://doi.org/10.3389/fpsyg.2016.00769>
- Tulving, E. (1972). Episodic and semantic memory. In Tulving, E. & Donaldson, W. (Eds.), *Organization of memory* (pp. 381-403). Academic Press.
- Vázquez, C., Duque, A., & Hervás, G. (2013). Satisfaction with Life Scale in a representative sample of Spanish adults: Validation and normative data. *The Spanish Journal of Psychology*, *16*, E82. <https://doi.org/10.1017/sjp.2013.82>
- Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., y Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin*, *133*(1), 122-148. <https://doi.org/10.1037/0033-2909.133.1.122>
- Williams, S. E., Ford, J. H., & Kensinger, E. A. (2022). The power of negative and positive episodic memories. *Cognitive, Affective, y Behavioral Neuroscience*. <https://doi.org/10.3758/s13415-022-01013-z>