

Generalizabilidad y optimización en la evaluación de la calidad de los Centros de Atención Infantil Temprana (CAIT)*

Generalizability and Optimization for Quality Assessment in Early Intervention Centers

Recibido: 20 de julio de 2013 | Revisado: 4 de abril de 2015 | Aceptado: 4 de abril de 2015

RITA PILAR ROMERO-GALISTEO **
VERÓNICA MORALES-SÁNCHEZ ***
ANTONIO HERNÁNDEZ-MENDO ****

Universidad de Málaga, España

doi:10.11144/Javeriana.upsy14-3.goec

Para citar este artículo: Romero-Galisteo, R. P., Morales-Sánchez, V., & Hernández-Mendo, A. (2015). Generalizabilidad y optimización en la evaluación de la calidad de los Centros de Atención Infantil Temprana (CAIT). *Universitas Psychologica*, 14(3), 1009-1020. <http://dx.doi.org/10.11144/Javeriana.upsy14-3.goec>

* Ha recibido apoyo financiero del Proyecto I+D+I: Observación de la interacción en deporte y actividad física: Avances técnicos y metodológicos en registros automatizados cualitativos-cuantitativos. Secretaría de Estado de Investigación, Desarrollo e Innovación del Ministerio de Economía y Competitividad [Referencia: DEP2012-32124]. También ha recibido apoyo del grupo Evaluación Psicosocial en Contextos Naturales: Deporte y Consumo (SEJ 444), financiado por la Junta de Andalucía (Consejería de Innovación, Ciencia y Empresa). A los Doctores Blanco-Villaseñor y Gálvez-Ruiz por su inestimable ayuda.

** Fisioterapeuta en el Centro de Atención Infantil Temprana de Antequera (Málaga). Doctora por la Universidad de Málaga. Profesora Asociada en el Departamento de Psiquiatría y Fisioterapia. Facultad Ciencias de la Salud de la Universidad de Málaga, España. Contacto principal para la correspondencia. Correo electrónico: rpromero@uma.es

*** Profesora Contratada. Doctora. Dpto. de Psicología Social, Antropología Social, Trabajo Social y Servicios Sociales. Facultad de Psicología. Correo electrónico: vomorales@uma.es

**** Profesor Titular. Dpto. de Psicología Social, Antropología Social, Trabajo Social y Servicios Sociales. Facultad de Psicología. Correo electrónico: mendo@uma.es

RESUMEN

El propósito de este trabajo fue optimizar una nueva herramienta de evaluación de la calidad de servicio percibida en los Centros de Atención Temprana, desarrollada *ad hoc*, que permitiera analizar distintas dimensiones desde la perspectiva de la Teoría de la Generalizabilidad. El objetivo fue desglosar la variabilidad real de la variabilidad del error. Se estimaron los componentes de varianza de las facetas centros, usuarios, ítems y escalas. Se utilizó una estructura de diseño multifaceta parcialmente anidado, analizando distintas categorías independientemente y en interacción, obteniendo excelentes resultados tanto de los índices de fiabilidad como de generalizabilidad. Se realizó una optimización del diseño de medida para optimizar los tamaños muestrales, pudiendo ser considerado un análisis coste-beneficio.

Palabras clave

Atención temprana, análisis de generalizabilidad, evaluación de la calidad, calidad percibida.

ABSTRACT

The purpose of this work was to optimize a new tool for evaluation of quality perceived in the Early Intervention Centers, developed *ad hoc*, allowing to analyze different dimensions of quality service in these kind of centres from the perspective of the Generalizability Theory. The aim was to break down the actual variability of the variability of the error. The variance of the facets centers, users, items and scales components were estimated. We used a structure of multifaceted partially nested design, analyzing different categories independently and in interaction, obtaining excellent results both indexes of reliability and generalizability. Measure design optimization to optimize sample sizes, and can be considered a cost-benefit analysis was conducted.

Keywords

early intervention; generalizability analysis; quality assessment; perceived quality

Introducción

A pesar de la inexistencia de estudios previos que así lo demuestren, con el paso del tiempo, la Atención Temprana (AT) ha experimentado un gran progreso en nuestro país. Los cambios sufridos en los últimos cuarenta años tanto a nivel legislativo como en la organización y la administración, la financiación e incluso en la estructura de la misma, han permitido que se avance en las necesidades de los niños y sus familias (Grupo de Atención Temprana [GAT], 2011). La filosofía, el concepto y los diferentes modelos de AT han ido adaptándose, en parte gracias a la aportación de las nuevas investigaciones sobre el tema (Burger, 2010; Durlak, 2010; Fantuzzo, Perry, & Childs, 2006; Odom et al., 2010). Aunque no siempre la colaboración entre los responsables de las políticas educativas, sociales y sanitarias, los distintos profesionales y las federaciones y asociaciones de padres, han posibilitado que los servicios de AT se aproximen más a las necesidades de las familias y de los niños atendidos en los diferentes centros.

Actualmente, se plantea la necesidad de reflexionar y analizar de la forma más completa posible la realidad de los diferentes Centros de Atención Infantil Temprana (CAIT) en cuanto a la calidad de servicio (Pegenaute, 2009; Pons, 2012). Tanto para las instituciones políticas como para las familias así como para los profesionales que trabajan en el sector, la AT es actualmente motivo de preocupación, por lo que se antoja necesario estudiar, recoger y analizar datos que puedan aportar soluciones. En este sentido, nuestro trabajo refleja la importancia de la evaluación de la calidad percibida por parte de las familias de los niños atendidos en los CAIT (Romero-Galisteo y Morales-Sánchez, 2010a; Romero-Galisteo y Morales-Sánchez, 2010b; Romero-Galisteo, 2011). Llegados a este punto, conviene recordar que la proliferación de este tipo de centros responde al incremento de la demanda tanto por parte de las familias como por la puesta en práctica de las sugerencias y aportaciones de evidencia que los diferentes estudios sobre el tema han ido aportando a lo largo de estos últimos años (Aytch, Castr, & Selz-Campbell, 2004; Essa, 2011; Moraleda, Romero, & Cayetano, 2013).

Hoy día, cualquier sector de actividad productiva ofrece la calidad como un valor intrínseco de sus productos o servicios. Desde esta perspectiva, la calidad de servicio es una de las principales áreas de estudio del comportamiento de los consumidores y usuarios, dado que el rendimiento de las organizaciones de servicios es valorado por las propias personas que adquieren y/o utilizan estos bienes de consumo y/o servicios. Por ello, los estudios de la calidad del servicio y de la satisfacción de los consumidores y usuarios tienen en común la importancia del punto de vista de los clientes, al valorar el producto y el servicio que proporciona la propia organización. Para realizar la gestión adecuada de la calidad, es necesario conocer qué buscan las personas en sus actividades de consumo (Quintanilla, 2002).

La calidad es uno de los elementos estratégicos en que se fundamenta la transformación y la mejora de los sistemas organizativos modernos. En toda clase de organizaciones vinculadas al sector servicios, cuando se analiza la calidad, se obtiene que, a pesar de sus potenciales beneficios, la implantación de una estrategia competitiva de esta naturaleza no está exenta de problemas, habida cuenta de la dificultad que plantea la definición y medida de la misma en el ámbito de la salud. Pese a ello, han sido numerosos los investigadores que han tratado de desarrollar definiciones y diseñar modelos para gestionar adecuadamente la calidad (Urbina, Pérez, & Martínez, 2000).

Son varios los autores que apuestan claramente por una solución que no opte claramente por uno u otro modelo de gestión previamente establecido, sino por la elaboración de uno que integre los ya existentes. En este sentido, una posible futura política para la implementación de procesos de mejora continua podría ser relatada como integrando elementos de los modelos generales de Gestión de la Calidad Total con modelos básicos y específicos propuestos por especialistas en distintos campos (Heras, Cilleruelo, & Iradi, 2008; Hudelson, Cléopas, Chopard, & Perneger, 2008).

La calidad de servicio debiera ser una estrategia clave en la organización de los centros AT en España, por lo que consideramos fundamental establecer un plan de calidad, con una adecuada optimización de los recursos, reducción de costes y una mejora conti-

nua. Ya en 1981, la Oficina Regional para Europa de la Organización Mundial de la Salud (OMS) decidió lanzar un nuevo programa bajo el título: Modelo de programas de asistencia sanitaria y control de calidad. El programa se proponía dos objetivos: la utilización de la combinación más apropiada de servicios en la asistencia a los pacientes con un problema determinado y la calidad de los servicios sanitarios. Aunque la OMS no ha emitido resoluciones explícitas mencionando el control de calidad de la asistencia, se asume la necesidad de actuar en este campo; la calidad está implícita en numerosas recomendaciones referentes al establecimiento de normas y estándares y a la evaluación de los servicios sanitarios. El interés de los Estados miembros por el control de calidad es cada vez mayor. Cualquiera que sea la definición de calidad escogida, la necesidad de evaluar la de los servicios sanitarios, y como tales podemos considerar a la AT, se ha puesto de manifiesto a pesar de las dificultades metodológicas que representa (Vuori, 1989). Estamos, por tanto, ante un cambio no solo en la gestión de los servicios públicos, sino también en la empresa privada. Esta necesidad de cambio debe compatibilizar los objetivos sociales con una calidad orientada a una adecuada gestión y atención al usuario. Para ello, se requiere del conocimiento y desarrollo de unas medidas válidas que permitan una gestión eficiente (Morales-Sánchez, 2009; Morales-Sánchez, 2011; Morales-Sánchez et al., 2013). En este sentido, se quiere dar importancia en la presente investigación al análisis de generalizabilidad.

La Teoría de la Generalizabilidad (Cronbach, Gleser, Nanda, & Rajaratnam, 1972) permite identificar, medir e implementar estrategias que reduzcan la influencia de las fuentes de error, optimizando distintos diseños de medida y considerando la fiabilidad y la generalizabilidad en términos de coste-beneficio. La Teoría de la Generalizabilidad (TG) asume que existen numerosas fuentes de variación e integra cada una de ellas en una estructura global que permite aplicaciones particulares de la teoría estadística del muestreo. Reconoce además, de manera explícita, las múltiples fuentes de error de medida (participantes, contextos, tratamientos, observaciones, sesiones, ocasiones de medida, etc.), con lo que se puede estimar cada una de ellas así como las diferentes interacciones entre sí. El error de medida no es más que

el efecto de las fluctuaciones debidas a la elección aleatoria de individuos, contextos, tratamientos, observaciones, sesiones, ocasiones de medida, etc., es decir, al muestreo de niveles particulares en cada una de las facetas (variables) del universo de observaciones posibles (Blanco-Villaseñor, 1991, 1992, 1993). Optimizar dicha medida es adaptar el diseño para reducir al máximo la variancia del muestreo debida a estas facetas. En esta investigación, hemos estimado los componentes de variancia de diversas facetas, tales como centros, usuarios escalas e ítems. Al mismo tiempo, la TG ha permitido llevar a cabo un análisis de la variabilidad estimada, a través de los componentes de la variancia, del grado de influencia en el diseño de cada una de las facetas, individualmente o en interacción. La optimización de los recursos representa uno de los aspectos de mayor repercusión actualmente en el ámbito de la gestión, por lo que cada vez son más los estudios que utilizan el análisis de generalizabilidad para el cálculo de un adecuado plan de optimización de la calidad de los servicios (Gálvez & Morales-Sánchez, 2011).

El objetivo principal de este trabajo, en la obtención de un adecuado plan de calidad, es la estimación de las fuentes de variabilidad y estimación de los diseños de medida óptimos para una adecuada evaluación de la calidad en los Centros de Atención Infantil Temprana. Para ello, se han utilizado distintos modelos que posibilitan minimizar la variancia de error y residual, maximizando el control de las otras fuentes de variación. Asimismo, se ha comprobado que la estructura general de diseño se ajusta al Modelo Lineal General (GLM) y que algunos de los diseños de medida propuestos permiten evaluar la calidad de los CAIT de manera eficiente, utilizando el Análisis de los Componentes de Varianza y el Análisis de Generalizabilidad. Asimismo, se ha utilizado la TG para optimizar los diseños de medida relativos al tamaño muestral, optimizando el coste-beneficio.

Método

Participantes

La muestra estuvo formada por 672 participantes, siendo en su mayoría padres y madres de niños que

reciben sesiones de tratamiento en quince CAIT de la provincia de Málaga. La obtención de la misma se realizó utilizando un muestreo de conglomerados. En el momento de la recogida de datos, existían en Málaga 17 CAIT que atendían a un total de 900 familias, por lo que la muestra de participantes se considera representativa. De la misma manera, los centros participantes reflejan la heterogeneidad existente tanto en la provincia de Málaga como en el resto de comunidades autónomas de España, ya que el servicio de AT puede prestarse de manera gratuita tanto en centros educativos como en asociaciones de padres, en centros públicos y privados, en administraciones locales, en centros que a su vez pueden estar subvencionados, conveniados o concertados con la Administración pública, etc. En Andalucía, actualmente, dependen de la Consejería de Salud y Bienestar Social, desde el año 1995.

El género de los participantes es mayoritariamente femenino, siendo 508 mujeres (75.6 %) y 162 hombres (24.1 %), con edades comprendidas entre los 20 y los 80 años, los que componen la muestra. Respecto a la edad de los mismos, el intervalo entre los 30 y 40 años ($DE = 0.84$) es el que más se repite, constituyendo el 58.3 % de la muestra.

Instrumento

En el proceso de recogida de datos, se utilizó el Inventario de Calidad de los Centros de Atención Infantil Temprana en su segunda versión (ICCAIT-v.2.0), de Romero-Galisteo (2011). Dicha herramienta está compuesta por 48 ítems agrupados en 6 escalas: instalaciones del centro, salas de tratamiento y material, atención al usuario, personal especializado, información general e información técnica. Las respuestas se recogen en un continuo de cinco puntos, en el que 1 significa *nada de acuerdo* y 5 *muy de acuerdo*, con el enunciado del ítem. Por último, se consideró importante la recogida de datos sociodemográficos y se añadió un apartado de observaciones y/o sugerencias, cuya información será objeto de futuros análisis.

El ICCAIT-v.2.0 es el cuestionario resultante de la optimización de una primera versión del mismo (ICCAIT-v.1.0), de Romero-Galisteo (2011), en el

que se realizó un estudio piloto inicial con 102 participantes y que estaba compuesto por 6 escalas y un total de 63 ítems. Los resultados mostraron una buena fiabilidad en todas las escalas salvo en la 5, siguiendo el rango establecido por Nunnally (1978). Dicho instrumento muestra una fiabilidad, estimada mediante el coeficiente α de Cronbach, entre 0.41 y 0.884. Posteriormente, esta investigación fue ampliada, aumentando tanto la muestra (pasando de $N = 102$ a $N = 672$) como el número de centros participantes (inicialmente fueron 3 que posteriormente se ampliaron a 15), con lo que se obtuvo una segunda versión de la herramienta que resultó más pragmática y consiguió mejorar sus propiedades psicométricas. Esta segunda versión, además de estar diseñada y validada específicamente en el contexto de la AT, cuenta con una fiabilidad estimada entre 0.322 y 0.885, como puede observarse en la Tabla 1.

Procedimiento

La administración del cuestionario contó con el consentimiento y aceptación de colaboración de los diferentes centros participantes. Previamente, hubo una primera toma de contacto tanto con un asesor técnico como con la consultora de AT del Servicio de Planificación y Evaluación Asistencial de la Delegación Provincial de Salud de la Junta de Andalucía en Málaga que facilitaron los datos administrativos y manifestaron la necesidad, el interés y la pertinencia de este tipo de estudios.

La recogida de datos fue realizada durante ocho semanas, entre los meses de noviembre y diciembre de 2009, por 37 encuestadores previamente formados para tal fin. Asimismo, los datos fueron tratados de manera confidencial. La cumplimentación del cuestionario se realizó de manera individual por el familiar que acompañaba en ese momento al niño a la sesión de tratamiento correspondiente. En la mayoría de los centros, se llevó a cabo en la sala de espera de los mismos, es decir, sin la presencia de cualquiera de los profesionales que integraban equipos de los distintos CAIT participantes. Los encuestadores estuvieron presentes en todo momento para solucionar posibles dudas en la comprensión de los ítems. Acto seguido, se numeraron los cuestionarios y en ningún momento

TABLA 1
Valores del α Cronbach iniciales y finales en el proceso de optimización del ICCAIT

Escalas	α Cronbach estudio fiabilidad inicial (ICCAIT-v.1.0, 63 ítems, 102 usuarios)	α Cronbach estudio fiabilidad final (ICCAIT-v.2.0, 48 ítems, 672 usuarios)
1	0.665	0.829
2	0.884	0.885
3	0.613	0.738
4	0.687	0.861
5	0.41	0.322
6	0.492	0.86

Fuente: elaboración propia

se mostraron al personal de los centros participantes. El tiempo empleado en rellenar el ICCAIT-v.2.0 fue aproximadamente de doce minutos.

Análisis estadístico

Para el análisis de los componentes de varianza y su precisión de generalización (Mushquash & O'Connor, 2006), se configuraron dos diseños multifaceta de 4 facetas: E (escalas con 6 niveles), U (usuarios con 672), I (ítems con 48) y C (centros con 15). Los modelos propuestos para el análisis fueron: **EI:EU (ítems anidados en las escalas)** y **EI:EUC (ítems anidados en las escalas)**. El procedimiento empleado fue el de Generalized Linear Model (GLM) realizado mediante el paquete estadístico SAS versión 9.1 (Schlotzhauer & Littell, 1997). Además se realizó un análisis de generalizabilidad con el programa EduG versión 6.0 (Cardinet, Jonson, & Pini, 2010), estimándose el porcentaje de variabilidad (varianza explicada) de cada una de las facetas y sus interacciones.

Con objeto de comprobar la fiabilidad test-retest, se dividió la matriz en dos mitades y se comparó el análisis de las mismas por separado, debiendo coincidir en valor los resultados obtenidos.

Resultados

En primer lugar se realizó un análisis de generalizabilidad. Los resultados del ANOVA pueden observarse en las Tablas 2 y 3, donde se muestra el análisis de los componentes de la varianza y el porcentaje de varianza explicada por los 2

modelos multifacéticos propuestos en este estudio: *escalas*ítems*usuarios* (EI:EU), en el que los ítems están anidados en las escalas (Tabla 2) y *escalas*ítems*usuarios*centros* (EI:EUC) en el que, de nuevo, la faceta ítems ha sido anidada en variable escalas (Tabla 3). Este análisis se llevó a cabo con el fin de observar la varianza asumida por cada una de las facetas y sus interacciones. Cabe destacar que las diferencias entre los niveles de las distintas variables son significativas.

En la Tabla 4, se recogen los coeficientes de precisión de generalización para los diferentes planes de medida estudiados (EI/U, U/IE, UE/I, IE/CU y CU/IE). Permite constatar, tal y como reflejan los valores estimados en los coeficientes de generalizabilidad absolutos y relativos, cuándo las facetas *usuarios*, *ítems* y *centros* han sido colocadas en el lugar de la instrumentación dentro del diseño y que los datos sean fiables y generalizables a una población universo (> 0.8), según Casamichana, Castellano, Blanco-Villaseñor y Usabiaga (2012). Sin embargo, estos coeficientes absolutos y relativos de precisión en la generalización indican que se deberían aumentar los niveles dentro de las facetas *escalas*, donde se han anidado los ítems, para incrementar la precisión de la generalización y el índice de fiabilidad.

Una vez definidos los Planes de Medida, se pasará a la fase de optimización, donde se varía el plan original para lograr una optimización de cada una de las facetas incluidas en el estudio (Blanco-Villaseñor & Oliva-Millán, 2010). Para cada Plan de Medida se realizan las modificaciones apropiadas en las facetas consideradas instrumentos de

TABLA 2

Análisis de la variancia del modelo de tres facetas: EI:EU

Fuente	Suma Cuadr.	gl	Media Cuad.	Aleatorizado	Componentes			DE
					Mixto	Corregido	%	
E	3371	5	674.2	0.04788	0.04788	0.04788	10.4	0.02616
I:E	3099	240	12.9125	0.03754	0.03754	0.03754	8.2	0.00349
U	3369	335	10.05672	0.03263	0.03263	0.03263	7.1	0.00316
EU	3401	1675	2.03045	0.04226	0.04226	0.04226	9.2	0.00171
IU:E	23954	80400	0.29794	0.29794	0.29794	0.29794	65.0	0.00149
Total	37194	82655					100%	

Fuente: elaboración propia

TABLA 3

Análisis de la variancia del modelo de cuatro facetas: EI:EUC

Fuente	Suma Cuadr.	gl	Media Cuad.	Aleatorizado	Componentes			DE
					Mixto	Corregido	%	
E	3371	5	674.2	0.00272	0.00272	0.00272	13.6	0.00149
C	467	14	33.35714	0.00028	0.00028	0.00028	1.4	0.00012
I:E	3093	282	10.96809	0.0021	0.0021	0.0021	10.5	0.00018
U:C	3175	5025	0.63184	0.00175	0.00175	0.00175	8.7	0.00004
EC	432	70	6.17143	0.00035	0.00035	0.00035	1.8	0.00006
EU:C	3248	25125	0.12927	0.0025	0.0025	0.0025	12.5	0.00002
CI:E	1471	3948	0.37259	0.00108	0.00108	0.00108	5.4	0.00002
IU:EC	13042	1417050	0.00920	0.0092	0.0092	0.0092	46.1	0.00001
Total	28299	1451519					100%	

Fuente: elaboración propia

medida. Tal y como se observa en las Tablas 5, 6 y 7 donde, a modo de ejemplo, se exponen en los resultados obtenidos del Plan de Optimización de los diseños de medida: EI/U, en el que la faceta usuarios interacciona como faceta de instrumentación e ítems y escalas como facetas de generalización o diferenciación; U/IE donde ítems y escalas interaccionan como facetas de instrumentación y usuarios como faceta de generalización o diferenciación e IE/CU donde centros y usuarios interaccionan como facetas de instrumentación e ítems y escalas como facetas de generalización o diferenciación.

En la Tabla 5, se presenta un diseño en el que se pretende optimizar la faceta U. En el presente estudio, se partió de 336 valores del criterio U, la mitad del total de la muestra, lo que comporta un total de 82 656 observaciones y supone unos coeficientes tanto de generalizabilidad como de fiabilidad de 0.99. Los resultados obtenidos implementando

340, 350 y 360 valores de U (aumentando con ello el número de observaciones a 83 640, 86 100 y 88 560, respectivamente) ofrecen igualmente excelentes coeficientes de generalizabilidad.

En la Tabla 6, la faceta de optimización es E (escalas). En el planteamiento de la investigación, se partió de 6 escalas, que equivalen a 82 656 observaciones con unos coeficientes de fiabilidad y generalizabilidad asociados de 0.8 y 0.67, respectivamente. No obstante, la TG permite de nuevo conocer los respectivos coeficientes de generalizabilidad absolutos y relativos correspondientes a otros valores de E. Aun así, para 10 E, con un total de 137 760 observaciones, solo se obtiene unos coeficientes de fiabilidad y generalizabilidad de 0.87 y 0.77, respectivamente.

En el último modelo implementado (Tabla 7), se muestra la optimización de la faceta C. Mediante la TG se han averiguado los distintos valores

TABLA 5
Optimización del diseño de medida EI/U

	Est. Generalizabil.		Opción 1		Opción 2		Opción 3	
	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.
E	6	INF	6	INF	6	INF	6	INF
I:E	48	INF	48	INF	48	INF	48	INF
U	336	INF	340	INF	350	INF	360	INF
Observ.	82 656		83 640		86 100		88 560	
Coef_G rel.	0.98829		0.98842		0.98875		0.98906	
redondeado	0.99		0.99		0.99		0.99	
Coef_G abs.	0.98718		0.98733		0.98768		0.98802	
redondeado	0.99		0.99		0.99		0.99	
Varianza de Error Relativo	0.00101		0.001		0.00097		0.00094	
Medida estandarizada del error relativo	0.03182		0.03163		0.03118		0.03074	
Varianza del Error absoluto	0.00111		0.0011		0.00107		0.00104	
Medida estandarizada del Error Absoluto	0.03331		0.03311		0.03264		0.03218	

Fuente: elaboración propia

TABLA 6
Optimización del diseño de medida U/IE incrementando el número de niveles de las escalas

	Est. Generalizabil		Opción 1		Opción 2		Opción 3		Opción 4	
	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.
E	6	INF	7	INF	8	INF	9	INF	10	INF
I:E	48	INF	48	INF	48	INF	48	INF	48	INF
U	336	INF	336	INF	336	INF	336	INF	336	INF
Observ.	82 656		96 432		110 208		123 984		137 760	
Coef_G rel.	0.7981		0.8218		0.84053		0.85569		0.86822	
redondeado	0.8		0.82		0.84		0.86		0.87	
Coef_G abs.	0.66568		0.69907		0.72639		0.74917		0.76844	
redondeado	0.67		0.7		0.73		0.75		0.77	
Varianza del Error Relativo	0.00825		0.00707		0.00619		0.0055		0.00495	
Medida estandarizada del Error relativo	0.09085		0.08411		0.07868		0.07418		0.07037	
Varianza del Error Absoluto	0.01639		0.01405		0.01229		0.01092		0.00983	
Medida estandarizada del Error Absoluto	0.12801		0.11851		0.11086		0.10452		0.09915	

Fuente: elaboración propia

de los coeficientes de fiabilidad y generalizabilidad que se obtendrían de haber empleado menos centros. Así, en un nivel inicial de 15 C habrían de realizarse 1 451 520 observaciones, obteniéndose unos valores excelentes tanto del índice de fiabilidad como del índice de generalizabilidad, coincidentes en 0.98. Al disminuir, por ejemplo, el nivel de la faceta centros a 10, se conseguirá reducir significativamente el número de observaciones hasta un valor de 947 520, permaneciendo prácticamente invariables los índices de fiabilidad

y generalizabilidad con valores de 0.97 y 0.96, respectivamente.

Discusión

Se considera que la presente investigación es un primer intento de profundizar en el ámbito de la calidad de servicio en Atención Temprana, facilitando estudios posteriores así como abriendo nuevas líneas de investigación. La propuesta permite el análisis de la calidad percibida por parte de los principales usuarios

TABLA 7

Optimización del diseño de medida IE/CU disminuyendo el número de niveles de los centros

	Est. Generaliz.		Opción 1		Opción 2		Opción 3		Opción 4		Opción 5	
	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.	Nivel	Univ.
E	6	INF	6	INF	6	INF	6	INF	6	INF	6	INF
C	15	INF	14	INF	13	INF	12	INF	11	INF	10	INF
I:E	48	INF	47	INF	47	INF	47	INF	47	INF	47	INF
U:C	336	INF	336	INF	336	INF	336	INF	336	INF	336	INF
Observ.	1 451 520		1 326 528		1 231 776		1 137 024		1 042 272		947 520	
Coef. G rel.	0.98009		0.9787		0.9771		0.97524		0.97305		0.97043	
redondeado	0.98		0.98		0.98		0.98		0.97		0.97	
Coef. G abs.	0.97637		0.97473		0.97284		0.97064		0.96806		0.96498	
redondeado	0.98		0.97		0.97		0.97		0.97		0.96	
Rel. Err. Var.	0.0001		0.0001		0.00011		0.00012		0.00013		0.00015	
Medida estandarizada del Error Relativo	0.00989		0.01024		0.01063		0.01106		0.01155		0.01212	
Abs. Err. Var.	0.00012		0.00012		0.00013		0.00015		0.00016		0.00017	
Medida Estandarizada del Error Absoluto	0.0108		0.01118		0.0116		0.01207		0.01261		0.01323	

Fuente: elaboración propia.

de los Centros de Atención Temprana y, aunque las aplicaciones de la TG tienen que ver habitualmente con el cálculo de la fiabilidad –intraobservador e inter-observadores–, mediante el coeficiente de correlación intraclass (CCI) y de la validez en la construcción de sistemas de categorías para la observación directa del comportamiento (Blanco-Villaseñor & Morales-Sánchez, 2010; García-García, Hernández-Mendo, Serrano, & Morales-Sánchez, 2013; Usabiaga, Castellano, Blanco-Villaseñor, & Casamichana, 2013), también puede utilizarse para la optimización de los tamaños de muestra ideales que permiten inferir de las mismas a las poblaciones pertinentes, constituyendo un estudio de coste-beneficio que aporta mayor flexibilidad y parsimonia en las diferentes ocasiones de medida. Optimizar dicha medida es adaptar el diseño para reducir al máximo la variancia del muestreo debida a las facetas utilizadas, en el presente caso: usuarios, centros, ítems y escalas. Esta investigación abre nuevas vías para profundizar en los estudios sobre Atención Temprana, haciendo posible el análisis de los aspectos de la misma desde una perspectiva contextualizada y secuenciada que se ajusta de forma más pertinente a la realidad de la misma en cuanto a la percepción de la calidad por parte de las familias se refiere.

En este trabajo, teniendo en cuenta los resultados que se han obtenido en el estudio de generalizabili-

dad, ha servido para comprobar que se partió de un diseño de investigación adecuado en cuanto a las facetas analizadas (centros, ítems, escalas y usuarios), ya que los modelos propuestos (EI:EU y EI:EUC) arrojan valores tanto del índice de fiabilidad como del índice de generalizabilidad excelentes, aunque con la posibilidad de disminuir el número de observaciones y tener la misma precisión de generalización en los análisis posteriores. Pero además ha servido para poder comprobar, en los diseños en los que la escala interacciona como faceta de instrumentación, que escalas e ítems son fiables pero no generalizables debido a que son indicadores de distintos constructos. Según Blanco-Villaseñor, Castellano-Paulis y Hernández-Mendo (2000), esta propuesta de anidación de facetas trae consigo un incremento en la relación coste-beneficio, en el sentido de que se cubre con un mayor índice de precisión en la generalización de resultados de una investigación. Se coincide en que siempre se contará con un menor número de componentes de error que permitirá dicha precisión. Constatar además, que los dos modelos propuestos son significativos. Asimismo, en el Análisis de Componentes de Varianza, coinciden en el valor del error, por lo que se demuestra la normalidad, linealidad y homocedasticidad de los mismos, siendo ambos significativos.

En cuanto al Plan de Optimización realizado, el trabajo previo de estimación de las fuentes de variancia permitió una puesta a punto de los diseños de medida adaptados a las decisiones consideradas en la investigación principal. Se puede afirmar, por tanto, que se ha tomado un número adecuado de centros. No obstante, es pertinente señalar que en el modelo EI:EUC existe la posibilidad de reducir considerablemente el número de observaciones, sin perder precisión de fiabilidad y generalización en posteriores análisis.

Respecto a los diseños de medida generados a partir de estos dos modelos, vale destacar que en los diseños de medida en los que los usuarios interaccionan como faceta de instrumentación, se debe aumentar y no disminuir los niveles, aunque ello suponga un aumento del número de observaciones. Esto fue debido a la conveniencia de dividir la matriz inicial en dos mitades, para comprobar así la fiabilidad test-retest, puesto que los resultados obtenidos, comparando el análisis de las 2 mitades por separado, coincidían en valor.

Estos análisis servirán, en cualquier caso, para diseñar futuras investigaciones más amplias, incorporando las modificaciones oportunas y que permitan alcanzar la mayor precisión de generalización y la optimización en términos de coste-beneficio.

Referencias

- Aytch, L. S, Castro, D. C., & Selz-Campbell, L. (2004). Early Intervention Services Assessment Scale (EISAS) - conceptualization and development of a program quality self-assessment instrument. *Infants and Young Children*, 17(3), 236-246.
- Blanco-Villaseñor, A. (1991). La teoría de la generalizabilidad aplicada a diseños observacionales. *Revista Mexicana de Análisis de la Conducta*, 14(3), 23-64.
- Blanco-Villaseñor, A. (1992). Aplicaciones de la teoría de la generalizabilidad en la selección de diseños evaluativos. *Bordón*, 43(4), 431-459.
- Blanco-Villaseñor, A. (1993). Fiabilidad, precisión, validez y generalización de los diseños observacionales. En M. T. Anguera (Ed.), *Metodología observacional en la investigación psicológica* (Vol. 2, pp. 151-261). Barcelona: PPU.
- Blanco-Villaseñor, A., Castellano-Paulis, J., & Hernández-Mendo, A. (2000). Generalizabilidad de las observaciones de la acción del juego en el fútbol. *Psicothema*, 12(2), 81-86.
- Blanco-Villaseñor, A., & Morales-Sánchez, V. (2010). Teoría de la generalizabilidad: investigaciones aplicadas. En M. J. Blanca, R. Alarcón & D. López-Montiel (Coords.), *Libro de Actas del XI Congreso de Metodología de las Ciencias Sociales y de la Salud* (p. 142). Málaga: CEES-IFV, UMA-Tecnolex y Asociación Malagueña de Estudios e Investigaciones Sociales.
- Blanco-Villaseñor, A., & Oliva-Millán, C. (2010). Generalización de la acción de éxito en fútbol. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 5(2), 283-296.
- Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly*, 25(2), 140-165.
- Cardinet, J., Jonson, S., & Pini, G. (2010). *Applying Generalizability Theory using EduG*. New York: Routledge.
- Casamichana, D., Castellano, J., Blanco-Villaseñor, A., & Usabiaga, O. (2012). Estudio de la percepción subjetiva del esfuerzo en tareas de entrenamiento en fútbol a través de la teoría de la generalizabilidad. *Revista de Psicología del Deporte*, 21(1), 35-40.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: John Wiley.
- Durlak, J. A. (2010). The importance of doing well in whatever you do: A commentary on the special section, "Implementation research in early childhood education". *Early Childhood Research Quarterly*, 25(3), 348-357.
- Essa, E. L. (2011). *Introduction to early childhood education. Annotated student's edition* (6th ed.). Belmont, CA: Wadsworth.
- Fantuzzo, J., Perry, M. A., & Childs, S. (2006). Parent satisfaction with Educational Experiences Scale: A multivariate examination of parent satisfaction with early childhood education programs. *Early Childhood Research Quarterly*, 21(2), 142-152.

- Gálvez, P., & Morales-Sánchez, V. (2011). Evaluación de la calidad en los programas municipales deportivos: generalizabilidad y optimización de diseños de medida. *Cuadernos de Psicología del Deporte*, 11(2), 123-130.
- García-García, O., Hernández-Mendo, A., Serrano-Gómez, V., & Morales-Sánchez, V. (2013). Aplicación de la teoría de la generalizabilidad a un análisis de tensiomiografía en ciclistas profesionales de ruta. *Revista de Psicología del Deporte*, 22(1), 53-60.
- Grupo de Atención Temprana. (2011). *La realidad actual de la atención temprana en España*. Madrid: Real Patronato sobre Discapacidad & Ministerio de Sanidad, Política Social e Igualdad.
- Heras, I., Cilleruelo, E., & Irandi, J. (2008). Quality management and quality of care in nursing homes. *International Journal of Health Care Quality Assurance*, 21(7), 659-670.
- Hudelson, P., Cléopas, A., Kolly, V., Chopard, P., & Perneger, T. (2008). What is quality and how is it achieved? Practitioner's views versus quality models. *Quality and Safety in Health Care*, 17, 31-36. doi:10.1136/qshc.2006.021311
- Morales-Sánchez, V. (2009). Evaluación de la calidad en organizaciones deportivas: análisis de generalizabilidad. *Revista de Psicología General y Aplicada*, 62(1-2), 99-109.
- Morales-Sánchez V. (2011). Evaluación de la calidad en organizaciones de servicios. En M. J. Blanca, R. Alarcón & D. López-Montiel (coords), Libro de Actas del XII Congreso de Metodología de las Ciencias Sociales y de la Salud (19-22).
- Morales-Sánchez, V., Berrocal, M. A., Morquecho, R., & Hernández-Mendo, A. (2013). Evaluación de la calidad en el área de educación física en un centro de enseñanza secundaria y bachillerato. *Revista Iberoamericana de psicología del ejercicio y el deporte*, 8(2), 411-427.
- Moraleda, E., Romero, M. J., & Cayetano, M. J. (2013). La parálisis cerebral como una condición dinámica del cerebro: un estudio secuencial del desarrollo de niños hasta los 6 años de edad. *Universitas Psychologica*, 12(1), 119-127.
- Mushquash, C., & O'Connor, B. P. (2006). SPSS and SAS programs for generalizability theory analyses. *Behavior Research Methods*, 38(3), 542-557.
- Nunnally, J. C. (1978). *Psychometric theory* (2a. ed.). New York: Mc Graw-Hill.
- Odom, S. L., Fleming, K., Diamond, K., Lieber, L., Hanson, M., Butera, G., ... Marquis, J. (2010). *Early Childhood Research Quarterly*, 25(3), 314-328.
- Pegenaute, F. (enero, 2009). *Expectativas de los profesionales*. Trabajo presentado en la Jornada de Presentación del Manual de Estándares de Calidad y de Acreditación de los Centros de Atención Infantil Temprana, Escuela Andaluza de Salud Pública, Granada.
- Pons, A. (marzo, 2012). *Modelo de investigación en Atención Temprana*. Trabajo presentado en las IX Jornadas de Atención Temprana de Andalucía: calidad e investigación en atención temprana, Asociación Interprofesional de Atención Temprana de Andalucía, Sevilla.
- Quintanilla, I. (2002). *Psicología social del consumidor*. Valencia: Promolibro.
- Romero-Galisteo, R. P., & Morales-Sánchez, V. (2009). Estimación de diseños para evaluar la calidad en los centros de atención infantil temprana(CAIT). En M. J. Blanca, R. Alarcón & D. López-Montiel (coords), Libro de Actas del XI Congreso de Metodología de las Ciencias Sociales y de la Salud (pp. 149-153). Málaga
- Romero-Galisteo, R. P., & Morales-Sánchez, V. (2010a). Atención Temprana y calidad de servicio. Propuesta de una herramienta de medida. *Efdeportes*, 15(145), 1-11. Recuperado de <http://www.efdeportes.com/efd145/atencion-temprana-y-calidad-del-servicio.htm>
- Romero-Galisteo, R. P., & Morales-Sánchez, V. (2010b). Calidad y Atención Temprana: breve revisión teórica. *Efdeportes*, 15(145). Recuperado de <http://www.efdeportes.com/efd145/calidad-y-atencion-temprana.htm>
- Romero-Galisteo, R. P.(2011). Evaluación de la calidad percibida en los centros de atención infantil temprana de la provincia de Málaga (Tesis Doctoral). Recuperado del Servicio de Publicaciones de la Universidad de Málaga (SPICUM).
- Romero-Galisteo, R. P. & Morales-Sánchez, V. (2011). Inventario de calidad en los centros de atención infantil temprana: análisis factorial exploratorio. En M. J. Blanca, R. Alarcón & D. López-Montiel

- (coords), Libro de Actas del XII Congreso de Metodología de las Ciencias Sociales y de la Salud. Donostia.
- Romero-Galisteo, R. P. & Morales-Sánchez, V. (2012). Calidad percibida por los usuarios de los centros de atención temprana. En M. J. Blanca, R. Alarcón & D. López-Montiel (coords), libro de actas del XIV Congreso Nacional de Fisioterapia. Madrid.
- Schlotzhauer, S. D., & Littell, R. C. (1997). *SAS Sistem for Elementary Statical Analysis*. Cary, NC: SAS Institute Inc.
- Urbina, O., Pérez, M., & Martínez, A. (2000). Flexibilidad organizativa y relación entre JIT y calidad total. *Alta Dirección*, 210(35), 74-84.
- Usabiaga, O., Castellano, J., Blanco-Villaseñor, A., & Casamichana, D. (2013). La Teoría de la Generalizabilidad en las primeras fases del método observacional aplicado en el ámbito de la iniciación deportiva: calidad del dato y estimación de la muestra. *Revista de Psicología del Deporte*, 22(1), 103-109.
- Vuori, H. V. (1989). *El control de calidad en los servicios sanitarios. Conceptos y metodología*. Barcelona: Masson.

