

# construcción de escalas de calidad de vida multidimensionales centradas en el contexto: la escala gencat

verdugo, m.a\*., schalock, r.l\*\*., gómez, l.e\*. y arias, b\*\*\* ■ ■ ■

\*instituto universitario de integración en la comunidad, universidad de salamanca, \*\*hastings college, EEUU, e \*\*\* instituto universitario de integración en la comunidad, universidad de valladolid

## resumen

En este artículo se describen las características esenciales del proceso de construcción de escalas de calidad de vida desde una perspectiva multidimensional. Además se presenta, a modo de ejemplo, el proceso seguido para construir la Escala GENCAT. La escala se basa en el modelo de calidad de vida multidimensional de ocho dimensiones propuesto por Schalock y Verdugo. Las principales fases de construcción del instrumento son: a) construir un *pool* de ítems basado en investigación previa, b) consulta a jueces expertos en la población objetivo, c) grupos focales con la participación de personas involucradas (usuarios, familiares y profesionales), d) desarrollo de una escala piloto, e) aplicación a una muestra representativa de la población objetivo, f) análisis de las propiedades psicométricas de la escala, y g) construcción de la escala definitiva.

**Palabras clave:** Calidad de vida. Evaluación. Escalas. Metodología. Estudio teórico.

## summary

*In this article we present the main characteristics of the procedure for developing quality of life questionnaires from a multidimensional perspective. By way of an example, the procedure for developing the GENCAT-questionnaire is also shown. The questionnaire is based on the eight-dimension model proposed by Schalock y Verdugo. The development of the instrument involved seven phases: a) development a pool of questions based on the previous research, b) expert panel about the target populations, c) focus group with the concerned people as participants (users, relatives and service providers), d) development of the field test questionnaire, e) application to a representative sample of the target populations, f) analysis of the psychometric properties of the questionnaire, and g) development of the final questionnaire.*

**Keywords.** Quality of life. Evaluation. Scales. Methodology. Theoretical study.

## Introducción ■ ■ ■

Este artículo presenta las fases que deben incluirse en el proceso de construcción de escalas o cuestionarios multidimensionales de evaluación de calidad de vida contextualizados en una población y cultura concreta. Para ello se describen los pasos dados por el Instituto Universitario de Integración en la Comunidad (INICO) en la construcción de la Escala GENCAT, realizada a petición del Institut Català d'Assistència i Serveis Socials (ICASS) de Cataluña. El proceso de construcción se encuentra ahora en sus últimas fases, analizando las respuestas obtenidas en la aplicación de cerca de 3000 encuestas a una muestra representativa de la población atendida por los servicios sociales catalanes. En este artículo se presentan las fases seguidas para construir el instrumento, procedimiento que posteriormente y de modo similar está siendo también seguido en Holanda, Canadá y otros países.

## Marco de fundamentación teórica ■ ■ ■

Medir los resultados personales relacionados con la calidad de vida de las personas es en la actualidad un proceso conceptual y psicométricamente claro. Los principales usos del concepto de calidad de vida de acuerdo con el modelo de Schalock y Verdugo (2002/2003; en prensa) son de tres tipos: a) como un marco de referencia para la prestación de servicios, b) como un fundamento para las prácticas basadas en la evidencia, y c) como un vehículo para desarrollar estrategias de mejora de la calidad. Y es en este sentido, centrado en el papel que la administración pública puede desarrollar, en el cual se inserta la construcción de la Escala Gencat.

Quien construye un instrumento de

evaluación debe sustentarse siempre en un marco teórico que defina claramente el constructo a medir, así como las conductas observables ligadas a aquél mediante definiciones semánticas. Las cuestiones que hacen referencia a estas conductas son las que deben plasmarse en los ítems del instrumento.

Tal como proponen Schalock y Verdugo (en prensa; Verdugo, 2006) calidad de vida individual es un "estado deseado de bienestar personal que: (a) es multidimensional; (b) tiene propiedades éticas (universales) y émicas (ligadas a la cultura); (c) tiene componentes objetivos y subjetivos; y (d) está influenciada por factores personales y ambientales". El modelo conceptual defendido por los autores durante las dos últimas décadas ha seguido los tres pasos siguientes (Schalock y Verdugo, en prensa): observar y describir el fenómeno estudiado, hacer un mapa conceptual del mismo, y evaluar el modelo (Carlisle y Christensen, 2006; Hughes et al., 1995; Poole et al., 2006; Schalock y Verdugo, 2002, en prensa; Shoemaker et al., 2004; Sutherland y Katz, 2005), estando ahora mismo planteándose el desarrollo de una teoría de calidad de vida (Schalock y Verdugo, en prensa).

La Escala GENCAT está basada en el modelo multidimensional propuesto por Schalock y Verdugo en el cual existen una serie de dimensiones o dominios que cuando se toman en conjunto constituyen el concepto de calidad de vida. Los autores proponen (2002/2003) ocho dimensiones fundamentales: bienestar emocional (BE), relaciones interpersonales (RI), bienestar material (BM), desarrollo personal (DP), bienestar físico (BF), autodeterminación (AU), inclusión social (IS) y derechos (DE). Otros autores proponen modelos en la misma línea que la descrita, variando el número y formula-

ción de las dimensiones, pero manteniendo una propuesta conceptual similar en la cual calidad de vida se entiende como el conjunto global de las valoraciones obtenidas en diferentes dimensiones.

## Proceso de construcción de la escala ■ ■ ■

Una vez contamos con un marco teórico que define el constructo que se pretende medir, el siguiente paso en el desarrollo de un instrumento de evaluación consiste en determinar la finalidad de la evaluación a la cual debe contribuir el instrumento y la población a la que está destinado.

La Escala GENCAT está destinada a evaluar la calidad de vida de usuarios adultos, mayores de 18 años, de los servicios sociales de Cataluña; entre ellos, personas con discapacidad intelectual, personas con discapacidad física, personas con discapacidad sensorial, personas mayores, personas con drogodependencias, personas con VIH, personas con SIDA y personas con problemas de salud mental.

El objetivo es realizar una evaluación objetiva de la calidad de vida de los usuarios de los servicios sociales, que sea sensible a los cambios en el modo de proporcionarlos, en su organización, y en las actividades de mejora que se puedan programar. Por ello, la escala está formada por ítems objetivos y observables por un observador externo, pues de acuerdo con la investigación son los que pueden ser sensibles a los cambios, mientras que los ítems subjetivos son más dudosos en ese sentido (Schallock y Felce, 2004; Verdugo et al., 2005).

La escala ha de ser respondida por profesionales de los servicios sociales que conozcan bien a la persona cuya calidad de vida se quiere evaluar. Por ello, se

establece como requisito para contestar la escala que el profesional trabaje directamente con la persona evaluada, al menos desde hace tres meses, y la conozca bien. Además, debe tener posibilidad de observar a la persona en distintos contextos y durante amplios períodos de tiempo, puesto que tendrá que responder a preguntas sobre la vida de la persona de muy diversa índole (trabajo, familia, ocio, etc.).

Las fases de construcción de la Escala GENCAT han sido las siguientes: 1) construcción de un *pool* de indicadores e ítems, 2) Consulta a jueces expertos sobre cada grupo de población a la que se va a aplicar, 3) Grupos focales incluyendo a distintas personas interesadas, 4) Desarrollo de la escala piloto, 5) Aplicación a una muestra representativa de la población total, 6) Estudios de las propiedades psicométricas de la escala, y 7) Construcción, validación y tipificación de la Escala GENCAT final. A continuación se describen las fases citadas.

## Primera fase: Construcción de un pool de indicadores e ítems ■ ■ ■

Una vez delimitada la finalidad del instrumento y la población a la que está dirigido, llega el momento de definir bien el *dominio*, es decir, el conjunto de indicadores más apropiados para representar bien el nivel de los sujetos en el constructo que se pretende evaluar.

Como se ha mencionado anteriormente, el constructo de calidad de vida se entiende compuesto por ocho dimensiones. Las dimensiones básicas de calidad de vida se definen a su vez a través de indicadores centrales; es decir, a través de percepciones, conductas o condiciones específicas de las dimensiones de calidad de vida (Schallock y Verdugo, 2003).

Para el desarrollo de la Escala GENCAT, en primer lugar se seleccionaron 46 indicadores considerados relevantes para los servicios sociales. Todos ellos fueron derivados empíricamente, seleccionados tras una exhaustiva revisión de las investigaciones publicadas, así como de otras pruebas de evaluación (Gómez, 2005; Gómez y Gómez, 2006; Gómez, Verdugo, y Arias, 2007; Jenaro et al., 2005; Schalock, 1997; Schalock y Verdugo, 2002/2003, en prensa; Schalock et al., 2005, en prensa; Verdugo, 2006; Verdugo, Arias y Gómez, en prensa; Verdugo y Schalock, 2003; Verdugo et al., 2001, 2005).

Estos 46 indicadores se concretaron en más de 300 ítems, algunos originales y otros procedentes de otros instrumentos. Tras una exhaustiva revisión y eliminación de los menos adecuados y los redundantes, el *pool* se redujo a 106 ítems, aquellos que mejor representaban los indicadores seleccionados. Cada dimensión contaba al menos con 10 ítems, siendo AU la que contaba con mayor número, 19 ítems concretamente.

### Segunda fase: Consulta a los jueces expertos ■ ■ ■

Con el objetivo de dotar de validez de contenido al instrumento, la siguiente fase consistió en una consulta a jueces expertos. De este modo, pretendimos comprobar si se concedió en la selección de ítems la importancia requerida a algunas facetas del constructo, y si se omitieron otras que pudieran ser relevantes. De esta manera, se tienen en cuenta los valores y aspiraciones de las personas con discapacidad y los estándares comunitarios para los mismos. Se trata, por tanto, de una primera aproximación al contexto, a las personas objetivo de la evaluación y a la comunidad en la que viven.

Los 106 ítems seleccionados fueron sometidos a una valoración por parte de 73 jueces expertos en calidad de vida y las diferentes poblaciones para las que está dirigida la escala. El 16,4% (N = 12) eran expertos en personas mayores; 11% (N = 8) en personas con problemas de salud mental; 19,2% (N = 14) en personas con discapacidad física; 4,1% (N = 3) en personas con discapacidad auditiva; 8,2% (N = 6) en personas con discapacidad visual; 24,7% (N = 18) en personas con discapacidad intelectual; 11% (N = 8) personas con drogodependencias y 5,5% (N = 4) con VIH/SIDA.

Su tarea consistió en valorar los 106 ítems, ordenados por dimensiones, valorando:

1. *Idoneidad* o grado en que considera que cada ítem es adecuado para medir esa dimensión.
2. *Importancia* para medir la calidad de vida de una persona.
3. *Observabilidad* o grado en que considera que la acción descrita en el ítem es observable y puede ser valorada de forma objetiva por un observador externo.

Además, se les pidió a los jueces que añadieran para cada dimensión otros ítems o indicadores de calidad de vida que no estuvieran incluidos en el *pool* y que consideraran importantes, así como que realizaran las observaciones que consideraran oportunas sobre los ítems propuestos.

Para analizar el acuerdo entre los jueces, se calculó el coeficiente de concordancia ponderado (BN) (Bangdiwala, 1987). El número de jueces participantes (73) superó ampliamente lo habitual en la investigación, en la que suelen incluir-

se entre 6-10 participantes. Teniendo en cuenta el alto número de participantes, el grado de acuerdo encontrado fue muy elevado ( $M = .61$ ), demostrándose así la validez de contenido del instrumento.

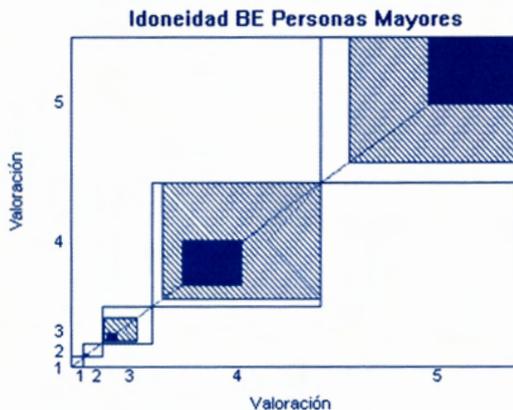
En la Tabla 1 puede observarse el índice de acuerdo en idoneidad, importancia y observabilidad de los ítems de cada una de las ocho dimensiones. Además, se presenta a modo de ejemplo la figura correspondiente al acuerdo encontrado entre los jueces expertos en personas mayores en cuanto a la idoneidad de los ítems en

BE. Los cuadrados totalmente coloreados muestran el acuerdo observado. Dichos cuadrados se inscriben en rectángulos mayores que muestran el acuerdo máximo posible. El acuerdo parcial se consigna incluyendo una contribución ponderada proveniente de las celdas fuera de la diagonal y se representa en el gráfico por medio de rectángulos sombreados. El 5 significa en este caso que los ítems son muy adecuados para medir la dimensión, mientras que el 1 indica que los ítems de esa dimensión son muy poco adecuados.

**Tabla 1. Medias del BN por dimensiones, idoneidad, importancia y observabilidad**

	ID	IM	OB	M
<b>BIENESTAR EMOCIONAL</b>	0,6	0,58	0,57	<b>0,59</b>
<b>RELACIONES INTERPERSONALES</b>	0,57	0,58	0,56	<b>0,57</b>
<b>BIENESTAR MATERIAL</b>	0,54	0,56	0,57	<b>0,56</b>
<b>DESARROLLO PERSONAL</b>	0,51	0,53	0,54	<b>0,53</b>
<b>BIENESTAR FÍSICO</b>	0,68	0,69	0,67	<b>0,68</b>
<b>AUTODETERMINACIÓN</b>	0,69	0,64	0,63	<b>0,65</b>
<b>INCLUSIÓN SOCIAL</b>	0,77	0,76	0,71	<b>0,75</b>
<b>DERECHOS</b>	0,67	0,69	0,61	<b>0,65</b>
<b>M</b>	<b>0,62</b>	<b>0,62</b>	<b>0,6</b>	<b>0,61</b>

**Figura 1. Acuerdo en idoneidad de los ítems de BE en personas mayores**



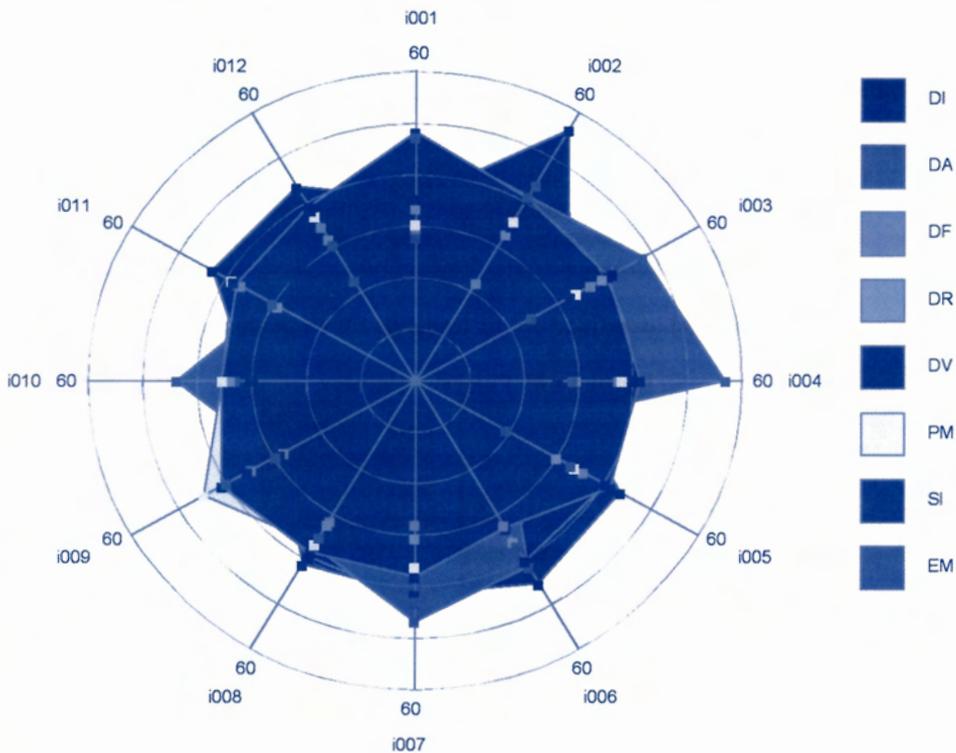
A continuación, se realizó una comparación de rangos de la valoración de los ítems por parte de los jueces en idoneidad, importancia y observabilidad. Se calcularon los rangos promedios de los ítems de cada una de las dimensiones y se compararon atendiendo a la especialidad de los jueces mediante la prueba de

Kruskal-Wallis. A modo de ejemplo, se presenta la tabla de contraste ( $c_2$ ,  $gl$  y nivel de significación asintótica) y el gráfico de los rangos por especialidad e ítem en el caso de la idoneidad en la dimensión BE, señalando los casos en que se han encontrado diferencias significativas ( $p < .05$ ).

**Tabla 2. Tabla de contraste de los ítems de la dimensión BE en idoneidad**

	i001	i002	i003	i004	i005	i006	i007	i008	i009	i010	i011	i012
$c^2$ 11,01	12,35	5,41	7,14	5,18	2,28	5,49	1,73	6,04	2,10	2,69	3,26	
$gl$	7	7	7	7	7	7	7	7	7	7	7	7
<b>Sig. asintót.</b>	0,14	0,09	0,61	0,41	0,64	0,94	0,60	0,97	0,54	0,95	0,91	0,86

**Figura 2. Rangos de los ítems de la dimensión BE en idoneidad**

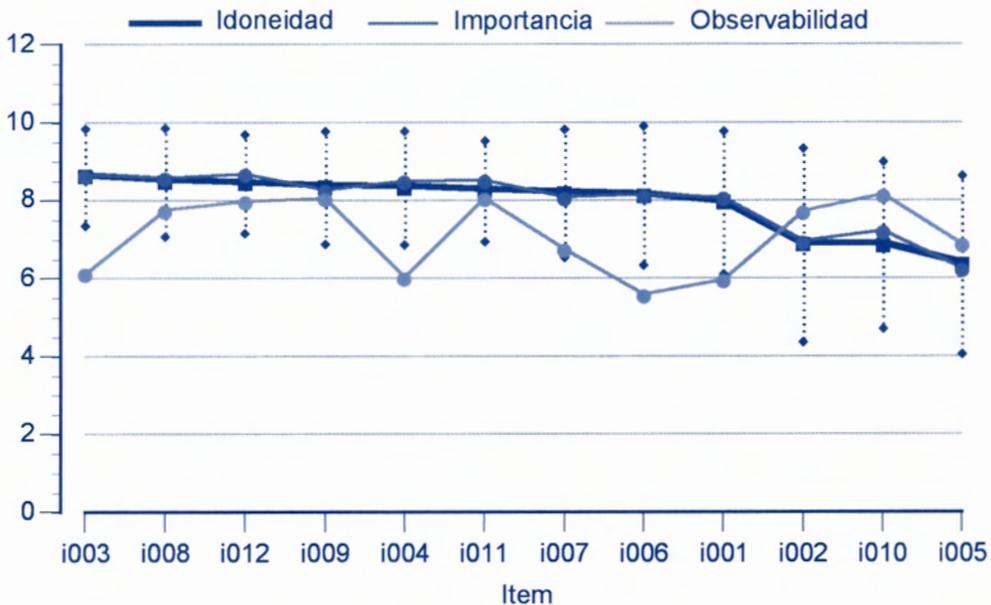


Se encontraron diferencias significativas ( $p < .05$ ) en la valoración de los jueces en cuanto a la idoneidad en 13 ítems, en cuanto a la importancia en 14 ítems y en cuanto a la observabilidad en 4 ítems. Además, observamos correlaciones muy elevadas entre idoneidad e importancia ( $r = .80$ ), y ligeramente más baja con observabilidad ( $r = .60$ ).

El siguiente paso consistió en calcular las puntuaciones promedio otorgadas por los jueces a cada ítem en idoneidad, importancia y observabilidad. Como hemos venido haciendo hasta ahora, presentamos como ejemplo la figura correspondiente a BE. Los ítems se presentan

ordenados de forma decreciente atendiendo a la puntuación alcanzada en idoneidad. Para esta misma variable, se incluye la dispersión de cada ítem. A fin de salvaguardar la validez de contenido, decidimos primar en primer lugar la idoneidad que según los expertos tiene el ítem para medir la dimensión que pretende medir, seguida de la importancia del ítem en la dimensión y, por extensión, en la calidad de vida de la persona. Por otra parte, puesto que el objetivo es construir una escala objetiva, se decidió que la puntuación en observabilidad debería ser superior a 6 puntos para ser aceptable.

**Figura 3. Puntuaciones promedio en idoneidad, importancia y observabilidad de los ítems de la dimensión BE**



Los criterios utilizados para la eliminación de los ítems menos adecuados fueron excluir los ítems con mayor dispersión y en cuya idoneidad, importancia u observabilidad los jueces mostraron mayor grado de desacuerdo global. En consecuencia, se mantuvieron aquellos ítems que presentaron: a) las puntuaciones más elevadas (en torno a 8 ó superiores) y una desviación típica inferior a 2,5 en idoneidad e importancia, y b) puntuaciones razonables (superiores a 6) y una desviación típica inferior a 3 en observabilidad. De este modo, se eliminaron ítems, como por ejemplo, el número 2, 10 y 5 en BE, calificados como fácilmente observables por los jueces, pero que alcanzaron puntuaciones bajas en idoneidad e importancia.

Mediante el proceso anteriormente señalado, se seleccionaron 59 ítems, y además se incluyeron otros 5 nuevos que fueron propuestos por los jueces durante la consulta. De este modo, llegamos al primer borrador de la escala GENCAT, conformada por 64 ítems.

### Tercera fase: Grupos focales ■ ■ ■

A pesar del exhaustivo trabajo realizado mediante la consulta a jueces expertos, decidimos dar un paso más para confirmar la validez de contenido del instrumento construido. Para ello, utilizamos la técnica de *grupos focales*. Los grupos focales requieren de procesos de interacción, discusión y elaboración de unos acuerdos dentro del grupo acerca de la temática propuesta por el investigador, en nuestro caso, los indicadores e ítems de calidad de vida propuestos en el primer borrador de la Escala GENCAT.

Esta técnica puede utilizarse como un método más para dotar de validez de contenido un instrumento de evaluación, puesto que da lugar a conclusiones y

acuerdos entre los participantes -que son producto de la interacción entre las partes involucradas en dicha evaluación-. Se trata de un modo de confirmar si los indicadores e ítems elegidos son importantes para ellos, identificar problemas con los ítems y añadir nuevos ítems o indicadores que podrían haber sido pasados por alto ya no sólo por el equipo de investigación, sino también por los expertos consultados. Es más, la participación en un grupo focal tiene innumerables ventajas tanto desde el punto de vista cognitivo como psicológico. La oportunidad de ser parte de un proceso participativo y tener poder de decisión favorece el compromiso de los participantes con el producto de la investigación. Así, mediante la puesta en marcha de grupos focales se pretendió concretamente que:

- Miembros de los distintos colectivos para los que está dirigida la Escala GENCAT validaran los ítems e indicadores propuestos.
- Los participantes propusieran ítems nuevos si lo consideraban necesario.
- Se implicaran usuarios, familiares y profesionales de los distintos colectivos cuya calidad de vida se pretende evaluar.

Cada uno de los cinco grupos focales -uno por colectivo implicado, a excepción de los grupos sobre drogodependencias y SIDA/VIH, que se unieron en uno sólo - estuvo formado por un moderador y 8-14 personas, entre ellas, usuarios de servicios sociales del ICASS, familiares de éstos y profesionales que trabajaban con ellos. Los participantes fueron seleccionados y convocados por el propio ICASS y el INICO.

En el desarrollo de los grupos focales se dispuso de una sala independiente en

cada uno, con una configuración en forma de "U". Cada uno de los participantes disponía de una escarapela de identificación para facilitar la comunicación. El tiempo estipulado para la discusión fue dos horas de duración. Para realizar la tarea, los participantes contaban con una lista que incluía los indicadores de calidad de vida y los ítems del primer borrador de la Escala GENCAT. El moderador, además de esta lista, disponía de una guía de discusión con las cuestiones a tratar, y unas hojas de registro para anotar sistemáticamente los comentarios, aportaciones y conclusiones del grupo.

La tarea se realizó dimensión por dimensión. Tras leer detenidamente los ítems de los que constaba cada dimensión, se discutía sobre las siguientes preguntas estímulo:

1. ¿Considera que el indicador/ítem afecta a la calidad de vida de este colectivo?
2. ¿Considera que medir ese indicador/ítem es importante para ese colectivo?
3. ¿Cree que los ítems que miden este indicador son observables para un evaluador externo?
4. ¿Piensa que falta algún indicador/ítem importante para medir esta dimensión?

Puesto que la información obtenida fue recogida de forma sistemática, se realizó a continuación un análisis cualitativo de la misma para cada grupo, y después un análisis global de las respuestas de todos los grupos. Las conclusiones obtenidas se pueden resumir en lo siguiente:

1. Todos los ítems e indicadores propuestos fueron considerados como impor-

tantes, idóneos y observables por los participantes, ratificando así los resultados obtenidos en la consulta a jueces de la fase anterior. Por lo tanto, no se planteó eliminar algún ítem.

2. Se propuso la reformulación y/o matización de algunos ítems en todos los grupos dirigidas a mejorar la expresión y precisión de lo evaluado.
3. Se propusieron algunos ítems nuevos en los grupos, pero ninguno era exclusivo de uno de los colectivos sino que eran aplicables a la población en general. Por ello, se incorporaron a la escala.

---

#### Cuarta fase: Escala piloto ■ ■ ■

---

Los grupos focales no solo confirmaron la validez de contenido del instrumento, sino que además sirvieron para reformular y hacer las matizaciones oportunas a aquellos ítems que podían resultar confusos. Así, el siguiente y último paso para llegar a la escala piloto consistió en reformular todos los ítems de la siguiente forma:

1. Teniendo en cuenta la opinión de los grupos focales en la reformulación de todos ellos.
2. Ponerlos en formato declarativo, de modo que puedan responderse en una escala de frecuencia.
3. Evitando el adverbio "no" y otras expresiones que incluyeran términos de negación -puesto que confunde a quien cumplimenta la escala- y cualquier otra expresión que implique frecuencia -puesto que entra en conflicto con el sistema de respuesta.
4. Reformularlos de modo que la mitad de los ítems de cada dimensión tuvie-

ran valencia negativa y la otra mitad valencia positiva, con el fin de evitar tendencia en la respuesta.

El instrumento construido proporciona un perfil objetivo de la calidad de vida de los usuarios de servicios sociales de acuerdo al modelo de ocho dimensiones propuesto por Schalock y Verdugo (2002/2003, en prensa). Es aplicable a cualquier usuario de los servicios con una edad mínima de 18 años. Se trata de un cuestionario autoadministrado con un formato de respuesta tipo Likert, en el que el profesional responde a cuestiones observables de la calidad de vida del usuario. El tiempo de cumplimentación es de 10 a 15 minutos, y aproximadamente unos 20 minutos si sumamos el tiempo de lectura del manual, donde se facilita toda la información necesaria para poder responder de forma autónoma.

La escala piloto GENCAT consta de 69 ítems que miden diferentes indicadores de calidad de vida y que conforman las 8 dimensiones citadas: BE (8), RI (10), BM (8), DP (8), BF (8), AU (9), IS (8) y DE (10). Los ítems están enunciados en tercera persona, con formato declarativo, y ordenados de forma aleatoria dentro de su correspondiente dimensión.

Todos los ítems pueden responderse con un formato de respuesta de frecuencia de cuatro opciones: en "*Nunca o casi nunca*", "*A veces*", "*Frecuentemente*", "*Siempre o casi siempre*". No obstante, para aquellos ítems a los que puede resultar difícil contestar con esta escala de frecuencia, se da la instrucción de responder pensando en la siguiente escala Likert de cuatro puntos: "*Totalmente de Acuerdo*", "*Acuerdo*", "*En desacuerdo*", "*Totalmente Desacuerdo*". Todos ellos son ítems observables, concretos y fáciles de comprender. No obstante, en la escala se incluye una pequeña matización para

algunos de ellos, que se deben leer detenidamente antes de responder.

Los ítems vienen precedidos por unas breves instrucciones y por dos fichas de datos sociodemográficos. Una correspondiente a los datos del profesional que rellena la escala y otra para los datos de la persona cuya calidad de vida se evalúa.

El cuestionario no ha sido diseñado para generar un índice global de calidad de vida, sino para proporcionar 8 índices, uno por cada una de las dimensiones. De este modo, para cada una de las 8 dimensiones, los ítems son codificados, agregados y transformados en una escala que tiene un recorrido desde 1 (el peor estado de calidad de vida para ese ítem) hasta 4 (el mejor estado de calidad de vida para ese ítem). Así la puntuación directa mínima para una dimensión será 1 multiplicado por el número total de ítems que la conforman, y la puntuación directa máxima será dicho número de ítems multiplicado por cuatro: BE (PDmin=8; PDmax=32), RI (PDmin=10; PDmax=40), BM (PDmin=8; PDmax=32), DP (PDmin=8; PDmax=32), BF (PDmin=8; PDmax=32), AU (PDmin=9; PDmax=36), IS (PDmin=8; PDmax=32) y DE (PDmin=10; PDmax=40).

La Escala GENCAT, se encuentra en la actualidad en proceso de validación. Una vez validada, se facilitarán también puntuaciones estandarizadas con los valores de las normas poblacionales para facilitar su interpretación.

---

Quinta fase: Aplicación a una muestra representativa de la población total ■ ■ ■

---

Un aspecto esencial del proceso de construcción de una escala es la aplicación de la versión piloto inicial a una

muestra representativa de población, cuidadosamente seleccionada, que permita posteriormente obtener las propiedades psicométricas adecuadas y realizar la selección final de los ítems para la escala definitiva.

### **Selección de la muestra**

A la hora de construir un buen instrumento, es siempre fundamental realizar un buen muestreo. Las investigaciones mediante encuestas, escalas o cuestionarios dominan el panorama de la investigación social. Sin embargo, a pesar de la extrema importancia del muestreo en cualquier procedimiento científico, se encuentran aún muchas deficiencias en la utilización de las muestras por parte de los investigadores.

Es habitual en las ciencias sociales asumir supuestos como el supuesto de uniformidad (dado que las características poblacionales son más o menos constantes en toda su extensión, cualquier porción es representativa del total) o el supuesto de disposición aleatoria (la población no es uniforme en cuanto a los valores de las variables evaluadas, pero se distribuye de forma aleatoria en la población). En las ciencias sociales se considera poco adecuado asumir dichos supuestos puesto que no se debe olvidar que el objeto de estudio en este caso son personas. Es por este motivo por el que se debe prestar especial atención a la hora de escoger la muestra utilizada a la hora de desarrollar y validar un instrumento como el aquí presentamos.

A continuación, sintetizamos el proceso seguido para muestrear la población a la que la Escala GENCAT fue aplicada con el objetivo de que dicha muestra fuera representativa. El muestreo se realizó a nivel de ámbitos geográficos y a nivel de comarcas, y para los dos grupos de pobla-

ción que se detallan a continuación.

*A. Muestra de personas en desventaja social: personas con discapacidad intelectual, personas con discapacidad física, personas con discapacidad sensorial, personas con problemas de salud mental, personas con drogodependencias, personas con SIDA y personas con VIH.*

Observamos una gran heterogeneidad en el número de centros que proporcionan servicios a los distintos colectivos por ámbitos geográficos en Cataluña. Con el objeto de acotar el error de muestreo se aplicó un diseño muestral estratificado con afijación de la muestra uniforme y proporcional. La afijación de la muestra se compone de un número determinado de unidades para cada ámbito y un número proporcional dependiendo del número total de centros en cada ámbito geográfico. Para garantizar que el error de muestreo no fuera elevado, el número de muestras uniforme para cada estrato geográfico que se propuso fue de 90 unidades.

El número de unidades muestrales que se ha considerado idóneo para este estudio es de 1100. Con el objetivo de garantizar que el error de muestreo no sea elevado, se han asignado un total de 195 encuestas para el estudio del colectivo de drogodependientes, 110 para el colectivo con SIDA, 457 para el colectivo de personas con discapacidad física. En el caso del colectivo de personas con discapacidad intelectual, se asignó un número fijo de 7 escalas para cada uno de los centros seleccionados y 15 para el resto de los colectivos. El error de muestreo para el total de los colectivos fue del 2,93%.

Aunque la escala se diseñó para evaluar la calidad de vida también de personas con discapacidad auditiva y visual, dichas poblaciones no fueron incluidas en el

muestreo dado que el ICASS no contaba con ningún centro o servicio dedicado exclusivamente a ellas. Los participantes de estos grupos fueron incluidos en otros grupos.

*B. Muestra de personas mayores (centros de día y residencias)*

En este segundo diseño muestral, al igual que en el caso anterior, también es evidente la desproporción que existe en los ámbitos geográficos, por lo que en el reparto de la muestra se procedió a asignar una parte fija y otra proporcional. La afijación de la muestra se compone de un número determinado de unidades para cada ámbito y un número proporcional dependiendo del número total de residencias y de centros de día en cada ámbito geográfico. Para garantizar que el error de muestreo no fuera elevado, el número de muestras uniforme para cada estrato geográfico que se propuso fue de 90 unidades.

Al igual que en el grupo anterior, el número de unidades muestrales que se ha considerado idóneo es de 1100, 713 en residencias y 397 en centros de día. El error de muestreo para el total de residencias y centros de día en Cataluña es del 2,91%. Si se consideran sólo las residencias el error se eleva al 3,62. Para los centros de día el error de muestreo es del 4,74%.

La distribución de las encuestas por comarcas se ha determinado asignando a cada una de ellas un número fijo de 7 encuestas, repartiéndose el resto proporcionalmente de acuerdo al número de residencias de cada comarca.

Respecto a la distribución de los centros de día por comarcas, también se ha procedido a repartir un número fijo de encuestas a cada comarca, que en este

caso ha sido de 5 unidades, procediéndose a repartir el resto, como en el caso anterior, de forma proporcional al número de centros de día de cada comarca.

**Aplicación de la Escala**

Una vez seleccionados los centros participantes en el proceso de validación de la escala (N = 350) es muy importante aplicar un riguroso protocolo de administración. En nuestro caso, en primer lugar se envió una carta desde el ICASS explicando los objetivos del estudio y rogando la participación a cada uno de los centros. Dicha carta se envió tanto por correo postal como por correo electrónico, con el fin de asegurar que todos ellos la recibieran al menos por uno de los medios.

A continuación, el equipo de investigación se puso en contacto telefónico con cada centro, con varios objetivos:

1. Confirmar que habían recibido la carta del ICASS. En el caso de que no la hubieran recibido, se procedió a enviar la carta de nuevo a la nueva dirección que nos facilitaban.
2. Confirmar la dirección postal, puesto que les mandamos las escalas por mensajería urgente.
3. Informarles con detalle del proyecto de investigación que se estaba llevando a cabo: objetivos, participantes, método de selección de centros, características de la evaluación y número de usuarios que debían evaluar en cada caso.
4. Confirmar si estaban dispuestos a participar en la investigación. En el caso de respuesta afirmativa, les enviamos las escalas y se establecía el correspondiente plazo para devolvernos las escalas, también por mensajería urgente. Si por el contrario, la respuesta era negativa,

el centro en cuestión era sustituido por otro de las mismas características, es decir, que atendiera al mismo tipo de población y que se encontrara en la misma comarca. En todos los casos en los que fue posible, se sustituyeron los centros que no quisieron colaborar por otro de las mismas características de la misma localidad. Debemos señalar, que fueron muy pocos los centros que rechazaron la participación. Tan sólo 18 de los 351, y entre las razones que ofrecieron destaca la sobrecarga de trabajo, la escasez de personal –la evaluación se realizó en verano y la mayor parte del personal se encontraba de vacaciones- y el estar participando en otros proyectos de investigación en ese momento.

5. Ponernos a su disposición en el caso de que tuvieran alguna duda, comentario o sugerencia sobre la escala.

Una vez confirmada la participación de todos los centros contemplados en el muestreo, enviamos a cada centro:

1. Un *manual* con las instrucciones necesarias para completar la escala, la descripción del proyecto de investigación, objetivos, descripción del concepto de calidad de vida, y explicación paso por paso del procedimiento de aplicación, confidencialidad de los datos proporcionados de los usuarios y el anonimato de los profesionales, y un número de teléfono y una dirección de correo electrónico para que pudieran realizar las consultas, comentarios o sugerencias que estimaran oportunos.
2. El número concreto de *escalas* que debían rellenar. En el caso de los centros que no tuvieran que evaluar a todos sus usuarios, se incluyeron 5 escalas más para asegurar que se alcan-

zaría el número de participantes que se consideró idóneo para validar la escala (N = 2200), en el caso de no obtener respuesta de alguno de los centros. En total, se enviaron un total de 4.500 escalas, de las cuales recibimos aproximadamente 3.000.

La mayor parte de los centros devolvieron las escalas en el plazo acordado. Sólo unos pocos se demoraron en el plazo, y en este caso, volvimos a ponernos en contacto telefónicamente con ellos para confirmar que seguían interesados en colaborar con nosotros y acordar un nuevo plazo. En la mayor parte de los casos, el plazo no se demoró más de una semana.

De este modo, el periodo de aplicación duró desde junio hasta la primera semana de septiembre de 2007, aunque la mayor parte de las escalas se recibieron durante los dos primeros meses.

Llegados a este punto, debemos hacer especial hincapié en la importancia que tuvo el contacto telefónico en la alta tasa de respuesta alcanzada en el estudio. Frecuentemente, en investigaciones de este tipo, suele solicitarse la colaboración a los centros por correo electrónico o correo postal. Sin embargo, el contacto telefónico, como ocurrió en este caso, es determinante para que muchos de los participantes se decidan a participar. El contacto telefónico requiere de mucho más tiempo y esfuerzo que otros medios pero asegura la obtención de una mayor tasa de respuesta.

Las fases finales de la Escala consistirán en analizar psicométricamente los datos para comprobar la fiabilidad y validez de la medida, seleccionando los ítems y estructura finales. Asimismo, se incluirán las instrucciones de administración y de puntuación.

## Conclusiones ■ ■ ■

El instrumento construido, con una fundamentación psicométrica contextualizada y sólida, se espera pueda ser utilizado en los servicios sociales de Cataluña para evaluar e incrementar la calidad de vida de los usuarios. Probablemente, el instrumento también pueda ser de utilidad en otros contextos próximos y de población similar.

La prueba de evaluación de la calidad de vida construida no debe ser utilizada como criterio para asignar más o menos recursos, pero sus resultados sí que pueden servir para conocer y evaluar los compromisos de mejora de las organizaciones y de los profesionales. En este sentido, conviene recordar lo que decíamos al principio de este artículo sobre el concepto de calidad de vida, y para lo cual el instrumento estaría indicado:

- a) es un marco de referencia para la prestación de servicios,
- b) es un fundamento para las prácticas basadas en la evidencia, y
- c) es un vehículo para desarrollar estrategias de mejora de la calidad.

Que lo dicho sea realidad va a depender no solo del esfuerzo investigador sino de la voluntad e iniciativa y capacidad de

innovación de profesionales, organizaciones y de la propia administración pública. En este sentido, las estrategias de mejora de la calidad deben abordar aspectos importantes de cambio y mejora organizacional, introducción de prácticas innovadoras y basadas en los datos, formación-reciclaje de profesionales, y otras.

La experiencia acumulada de evaluación de programas en las organizaciones dedicadas a los servicios sociales, particularmente en lo referente a las personas con discapacidades intelectuales y del desarrollo, nos indica que son necesarias varias fases para lograr el cambio en los servicios sociales, y que se pueden resumir en (Schallock, Bonham y Verdugo, en prensa):

*Pasar de la incertidumbre al interés por el cambio, y para ello se precisa de visión, esperanza, y respuestas potenciales sobre como mejorar resultados personales.*

*Pasar del interés por el cambio al compromiso, donde se necesita un modelo conceptual de calidad de vida, un apoyo institucional previo, y una involucración personal en el proceso de cambio.*

*Pasar del compromiso a la acción, lo cual requiere conocer estrategias específicas, aceptar el valor positivo del cambio, y desarrollar un sentido de competencia personal que afecta al cambio.*

Bangdiwala, K. (1987). Using SAS software graphical procedures for the observer agreement chart. *Proceedings of the SAS Users Group International Conference*, 12: 1083-1088.

Carlisle, P. R. y Christensen, C. M. (2006). *The cycles of theory building in management research*. Boston: School of Management-Boston University and Harvard Business School.

Gómez, L. (2005). *Construcción de una escala de calidad de vida para adultos con discapacidad intelectual*. Trabajo de grado sin publicar, Universidad de Salamanca, Salamanca.

Gómez, L.E. y Gómez, M. (2006). Análisis sobre la importancia y uso de los indicadores de calidad de vida en personas con discapacidad intelectual, familiares y profesionales. *INTEGRA*, 9(23), 4-5.

Gómez, L.E., Verdugo, M.A. y Arias, B. (2007). Aplicación del Modelo de Calidad de Vida en los Servicios Sociales de Cataluña. *INTEGRA*, 10(27), 4-5.

Hughes, C., Hwang, B., Kim, J-H., Eisenman, L.T., y Killian, D. J. (1995). Quality of life in applied research: A review and analysis of empirical measures. *American Journal on Mental Retardation*, 99 (6), 623-641.

Jenaro, C., Verdugo, M.A., Caballo, C., Balboni, G., Lachappele, Y., Otbrebski, W. y Schalock, R. L. (2005). Cross-cultural study of person-centered quality of life domains and indicators: A replication. *Journal of Intellectual Disability Research*, 49

(part 10), 734-739.

Poole, D. L., Duvall, D. y Wofford, B. (2006). Concept mapping key elements and performance measures in a state nursing home-to-community transition project. *Evaluation and Program Planning*, 29 (1), 10-22.

Schalock, R.L. (1997). *Quality of life: Vol I: Conceptual and measurement issues*. Washington, DC: American Association on Mental Retardation.

Schalock, R. L., Bonham, G.S. y Verdugo, M.A. (en prensa). The Concept of Quality of Life as a Framework for Program Planning, Implementation, Evaluation, and Improvement. *Evaluation and Program Planning*.

Schalock, R. L. y Felce, D. (2004). Quality of life and subjective well-being: conceptual and measurement issues. In E. Emerson, C. Hatton, T. Thompson, and T. R. Parmenter (Eds.), *International handbook of applied research in intellectual disabilities* (pp. 261-279). London: John Wiley and Sons.

Schalock, R. y Verdugo, M.A. (2002/2003). *Quality of life for human service practitioners*. Washington, DC: American Association on Mental Retardation [Traducido al castellano por M.A. Verdugo y C. Jenaro. *Calidad de vida. Manual para profesionales de la educación, salud y servicios sociales*. Madrid: Alianza Editorial].

Schalock, R. y Verdugo, M.A. (en prensa). Quality of Life Theory: Its Development and Importance in the Field of Intellectual Disabilities. *Journal of Policy and Practice in*

*Intellectual Disabilities.*

Schalock, R., Verdugo, M.A., Bonham, G. S., Fantova, F. y Van Loon, J. (en prensa). Enhancing Personal Outcomes: Organizational Strategies, Guidelines, and Examples. *Journal of Policy and Practice in Intellectual Disabilities.*

Schalock, R. L., Verdugo, M.A., Jenaro, C., Wang, M., Wehmeyer, M., Xu, J., y Lachapelle, Y. (2005). A cross-cultural study of quality of life indicators. *American Journal on Mental Retardation*, 110, 298-311.

Shoemaker, P. J., Tankard, J.W. Jr., y Lasorsa, D. L. (2004). *How to build social science theories*. Thousand Oaks, CA: Sage Publications.

Sutherland, S. y Katz, S. (2005). Concept mapping methodology: A catalyst for organizational learning. *Evaluation and Program Planning*, 28 (3) 257-269.

Verdugo, M.A. (2006). *Cómo mejorar la calidad de vida de las personas con discapacidad*. Salamanca: Amarú.

Verdugo, M.A., Arias, B. y Gómez, L.E. (2006). Escala integral de medición subjetiva y objetiva de la calidad de vida en personas con discapacidad intelectual. En M.A. Verdugo (Dir.), *Como mejorar la calidad de vida de las personas con discapacidad. Instrumentos y estrategias de evaluación* (pp 417-448). Salamanca: Amarú.

Verdugo, M.A., Arias, B. y Gómez, L.E. (en prensa). *Escala Integral de Calidad de Vida*. Salamanca: INICO.

Verdugo, M.A., Arias, B. y Martin, J.C. (2006). Validación mediante análisis

multivariante del modelo de Schalock y Verdugo (2002). *Informe de Investigación sin publicar*. INICO, Universidad de Salamanca, Salamanca.

Verdugo, M.A. y Schalock, R. L. (2003). *Cross-cultural Survey of QOL Indicators*. Salamanca, Spain: Institute for Community Integration-Faculty of Psychology, University of Salamanca.

Verdugo, M.A., Arias, B. y Gómez, L.E. (en prensa). *Escala Integral de Calidad de Vida*. Salamanca: INICO.

Verdugo, M.A., Arias, B. y Martin, J.C. (2006). Validación mediante análisis multivariante del modelo de Schalock y Verdugo (2002). *Informe de Investigación sin publicar*. INICO, Universidad de Salamanca, Salamanca.

Verdugo, M.A. y Schalock, R. L. (2003). *Cross-cultural Survey of QOL Indicators*. Salamanca, Spain: Institute for Community Integration-Faculty of Psychology, University of Salamanca.

Verdugo, M.A., Schalock, R. L., Keith, K.D., y Stancliffe, R. (2005). Quality of life and its measurement: Important principles and guidelines. *Journal of Intellectual Disability Research*, 49 (10), 707-717.

Verdugo, M.A., Schalock, R., Wehmeyer, M., Caballo, C. y Jenaro, C. (2001). *Cuestionario de Indicadores transculturales de calidad de vida: evaluación de su importancia y uso*. Salamanca: INICO, Universidad de Salamanca