

# An Outcomes-Based Assessment of Quality of Life in Social Services

Laura Elisabet Gomez · Benito Arias · Miguel Ángel Verdugo · Patricia Navas

Accepted: 23 January 2011 / Published online: 4 February 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** The goal of this article consists of describing the calibration of an instrument to assess quality of life-related personal outcomes using Rasch analysis. The sample was composed of 3,029 recipients of social services from Catalonia (Spain) and was selected using a probabilistic polietapic sample design. Results related to unidimensionality, item and person separation reliability, calibration, items' level of difficulty, response categories and differential item functioning by gender and type of collective are provided and point out the general fit of the data to the model. However, it is recommended to include more difficult items and eliminate one that seems not to be adequate.

**Keywords** Quality of life · Outcomes · Social services · Rasch model · Assessment

## 1 Introduction

This article describes an instrumental study whose goals consisted in calibrating a direct observation quality of life questionnaire and an outcomes-based measurement of services quality of for users of social and human services: the GENCAT<sup>1</sup> Scale (Verdugo et al. 2008a, b). It is widely accepted that quality of life concept is important in social services to implement person-centered programmes and practices, to assess and report personal outcomes, to guide quality improvement strategies, and to improve the effectiveness of those practices and strategies using evidence-based or outcomes-based measurements (Schallock

---

<sup>1</sup> GENCAT is an acronym of “*Generalitat de Catalunya*”; it means “Government of Catalonia” in Catalanian.

---

L. E. Gomez (✉)  
INICO, University of Valladolid, Palencia, Spain  
e-mail: lauragomez@psi.uva.es

B. Arias  
INICO, University of Valladolid, Valladolid, Spain

M. Á. Verdugo · P. Navas  
INICO, University of Salamanca, Salamanca, Spain

et al. 2008a; Verdugo et al. 2010). The quality of life concept has recently begun to be applied in social policies since it has become a reference model for service provision, a basis for evidence-based practices, and a tool to develop quality improvement strategies (Schalock and Verdugo 2002, 2007).

The GENCAT Scale has been developed on the basis of the eight-domain model proposed by Schalock and Verdugo (2002) and subsequent work regarding the model's validation and cross-cultural use (Jenaro et al. 2005; Schalock et al. 2005, 2008b). According to this framework, quality of life is composed of eight domains and core indicators that define operationally each quality of life domain. The indicator measurement results in personal outcomes that can be used for both reporting purposes and guiding organization improvements (Claes et al. 2009; Keith and Bonham 2005; Schalock and Bonham 2003; Schalock et al. 2008a, b). According to this model, individual quality of life is a multidimensional phenomenon composed of core domains that are influenced by personal characteristics and environmental variables; these core domains are the same for all people, although they may vary in relative value and importance, and quality of life domains are assessed on the basis of culturally sensitive indicators (Schalock et al. 2010).

The current approach to the measurement of quality of life is characterized by: (1) its multidimensional nature involving core domains and indicators; (2) the use of methodological pluralism that includes the use of subjective and objective measures; (3) the incorporation of a systems perspective that captures the multiple environments impacting people at the micro, meso, and macrosystem levels; and (4) the increased involvement of persons with risk of social exclusion in the design and implementation processes (Bonham et al. 2004; Verdugo et al. 2005). Moreover, depending on the purpose and the perspective of the instrument used, quality of life related personal outcomes would be used to assess either person's perceived well being on the item (self-report) or the person's life experiences and circumstances (direct observation) (Claes et al. 2009; Schalock et al. 2008a). In this sense, we can speak about objective and subjective measures and measurement instruments, depending on their purpose, content, and respondent. If an evaluator desires to assess personal outcomes and develop person-centered programs, subjective Likert-type scales answered by the client or user of the service should be applied (Bonham et al. 2004; Cummins 2005; Keith and Bonham 2005; Keith and Schalock 2000; Schalock and Felce 2004). In distinction, when the goal is program evaluation, service quality improvement, or to assess organizational changes, it is recommended to use objective questionnaires based on the direct observation of personal experiences and circumstances (Perry and Felce 2005; Schalock and Felce 2004; Schalock et al. 2007; Schalock and Verdugo 2002; Verdugo et al. 2007a; Walsh et al. 2006).

The purpose of this study was to calibrate the GENCAT Scale and evaluate its psychometric properties with the goal of having a tool for carrying out program evaluations, making evidence-based service quality improvements, and assessing and guiding organizational changes. In reference to this purpose, psychometric properties of the measures provided by the Catalanian version of the GENCAT Scale were analyzed using Rasch Rating Scale Model (Andrich 1978; Wright and Masters 1982; Rasch 1960, 1980, 1992). The Rasch Model is often classified under Item Response Theory (IRT) that specifies how people, test items, probes, or others must interact statistically through probabilistic measurement models for linear measures to be constructed from ordinal observations. This model requires the investigation and quantification of accuracy, precision, reliability, construct validity, quality control fit statistics, statistical information, linearity, local dependency and unidimensionality (Linacre 2010). On the other hand, IRT is a system of models that defines one way of establishing the correspondence between latent variables

(e.g., quality of life) and their observed manifestations (e.g., test items). For that, persons and items are located on the same continuum, and it is considered that, for an item to have any utility, it must be able to differentiate among persons located at different points along a continuum.

The specific goals of GENCAT's calibration were next: (1) to verify the unidimensionality of each factor; (2) to obtain validity evidence for the instrument through the observed fit of the data to the model, regarding both items and persons; (3) calculate the indices of reliability and separation; (4) estimate the calibration of the items; (5) verify the accuracy of the measurement; and (6) determine whether there is Differential Item Functioning (DIF) by gender and type of collective.

## 2 Methods

### 2.1 Participants

A total of 608 professionals working in 239 centres of social services participated filling in the field-test version of the GENCAT Scale for 3,029 users from Catalonia (Spain). The mean number of evaluated persons per service was 12.67 ( $SD = 7.75$ ) and the mean number of evaluated persons per professionals was five ( $M = 4.98$ ). A probabilistic polietapic sample design was carried out to select the participants with the purpose of guaranteeing the representativity. The sample was composed of two main groups: elder people group (sampling error of 2.43 with 95% confidence and  $p = q$ ) and people at risk of social exclusion (sampling error of 2.91 with 95% confidence and  $p = q$ ).

The requirements for professionals to participate in this study were: (1) to be working in some kind of social service for handicapped adults; and (2) to have been working directly with the client for at least 3 months. The only requirement to apply the scale to a social service user was that this was older than 16 years old. Related to the main socio-demographic characteristics of professionals, most of them were female (85%), had been working with the client for more than 2 years (55.74%), were psychologists (23.01%) and social workers (18.41%), and had been working in social services more than 5 years (52.80%). Regarding the users of the social services evaluated, the sole criteria governing their inclusion in the research were: (1) to be users of some kind of social services attached to the ICASS (Catalan Welfare and Social Services Institute), and have done so for at least 3 months; (2) to be over the age of 18. Concerning to their socio-demographical data, 55.7% were female. Their ages ranged between 16 and 105 ( $M = 64.72$ ;  $SD = 21.34$ ). More than half of sample (57.57%;  $n = 1,711$ ) was older than 60. Actually, the biggest group ( $n = 791$ ) was composed of 81–90 years old people and only 17.39% ( $n = 515$ ) were younger than 41.

The most representative group was the one composed of elder people living in residence settings (44.70%), followed by people with intellectual disabilities (19.35%), physical disabilities (11.72%), mental health (10.33%), and old people in day centres (8.75%). Percents of people with drug dependences and HIV/AIDS ranged from 2.48 to 2.67%.

### 2.2 Instrument

Although there is a Spanish version of the GENCAT Scale (Verdugo et al. 2009), the Catalanian version of the instrument (Verdugo et al. 2008a) was applied in this study.

The GENCAT Scale is a self-administered questionnaire in which professionals answer objective and observable questions about user quality of life based on direct observation of person's life. It is composed of eight scales—that correspond to the eight quality of life domains—and 69 items that are formulated as third person declarative statements and random organized by domains. The answer format is a frequency scale with four options ('never or hardly ever', 'sometimes', 'often', 'always or almost always').

The GENCAT Scale was developed through a systematic and rigorous method (Verdugo et al. 2008b, 2010) and has served internationally as a model for developing other multidimensional quality of life scales focusing on the context (Verdugo et al. 2007b; van Loon et al. 2008).

### 2.3 Procedure

Once the participant services were selected, a letter was sent in order to explain the researching goals and make the participation request. This letter was sent by post and by email. After that, the research team phoned every single service to confirm the post address (since the scales were sent by a courier company) to confirm if they were determined to collaborate. Once their participation was confirmed, the specific number of scales they must complete plus 5 (to be sure of achieving a big enough sample) and an evaluator's guide were sent. About 4,500 scales were sent.

The statistic software that was used to analyze data was WINSTEP 3.68.0 (Linacre 2008; Linacre and Wright 1999).

## 3 Results

### 3.1 Preparing the Data

Before describing the results obtained with the logistical analysis of a parameter, we performed the pertinent prior verifications of the model's data fit: (1) the point-biserial correlations were positive in all cases (.13 and .82); (2) the function of the categories was suitable in all cases; in fact, each category had more than 1,000 observations, comfortably surpassing the recommended minimum number of responses (i.e., 10 observations). Furthermore, the mean measurements for the categories progressively increased in all the dimensions except in Material wellbeing and Rights, in which category 3 was especially noisy; and (3) the study of item misfit confirmed the sitting of the items within the recommended range (Linacre 2002) with the sole exceptions being three items ('*He/she is exposed to exploitation, violence or abuse*' in Rights, '*The service he/she attends caters for their preferences*' in Self-determination and '*He/she has a satisfactory sex life*' in Interpersonal relations, which returned values slightly higher than 2).

### 3.2 Unidimensionality

The principal components analysis of the eight subscales' residuals gave rise to percentages of between 36 and 58.7% of variance explained by the modelled data. More specifically, an analysis of the subscales Interpersonal relations and Self-determination gave rise to percentages that were slightly lower than the recommended value of 60%, whilst the majority exceeded the commonly used one of 40%. Only the subscales of Physical

wellbeing and Material wellbeing did not reach that value, although they came very close. These results, together with the first tests in each principal components analysis (with eigenvalues below 3.0, which is considered to indicate the existence of a second dimension), led us to confirm the unidimensionality of the eight subscales of quality of life.

### 3.3 Fit of the Data to the Model

Secondly, we analysed the model's data fit. On the one hand, regarding the fit for persons, it is noteworthy that Interpersonal relations is the only subscale that does not have extreme data, whereas Rights and Material wellbeing have more than 600 extreme cases. In addition, we have a perfect fit for persons in Self-determination and Social inclusion in the case of both the infit and the outfit, and in Physical wellbeing in the case of the infit. All the other values ranged between  $-.1$  and  $.1$ . Finally, the values of MNSQ were very close to 1 in all cases (see Table 1). For these reasons, we can conclude that the overall fit for persons shows that the responses are consistent with the response patterns foreseen by the model. On the other hand, regarding the overall fit of the items, the MNSQ values confirmed the items' fit to the RSM in all cases. On this occasion, Emotional wellbeing and Social inclusion were the subscales with a perfect fit and Material wellbeing exceeded the value 11.01, albeit only slightly. Given that all the values fell within the range considered to be acceptable, we confirmed the overall fit of the items to the model. A more detailed analysis of the items' fit furthermore revealed that the more accurate dimensions are Emotional wellbeing, Personal development, Self-determination and Social inclusion. By contrast, the least accurate items are in Material wellbeing, Physical wellbeing, and Rights.

No item presented dependence or determinism (values below 0.60). However, five of the 69 items revealed a lack of fit, noise or high random variability in the data (values substantially higher than 1.5). In brief, most of the parameters for the items in the GENCAT Scale have an acceptable behaviour according to the postulates of the Rasch model, whereby they all have a suitable fit, with the exception of the final item (*'He/she is exposed to exploitation, violence or abuse'* in Rights, whose fit is highly debatable).

### 3.4 Item and Person Separation Reliability

The calculation of reliability involved the use of the item separation reliability index and the person separation reliability index (Fig. 1). Concerning the former, we obtained a separation reliability index for the items equal to 1 in all the subscales, so we can affirm that the items have the utmost reliability. However, the separation reliability indices for people were lower and varied greatly between the subscales: Self-determination was the only dimension that returned a value considered to be acceptable (.82), whereas Material wellbeing returned a reliability coefficient that was so low (.15) that its reliability as regards persons is highly questionable. All the other values ranged between .23 (Physical wellbeing) and .77 (Emotional wellbeing). The results for the separation indices for items and people confirmed the previous results, whereby the separation index for the items exceeds the value of 2.00 in all cases (in fact, it exceeded the value of 18 in all cases and reached a value of 39.19 in Interpersonal relations). However, the separation index for people only exceeded the value of 2.00 in the case of the subscale Self-determination (2.16). The lowest separation indices corresponded to Material wellbeing (0.41) and Physical wellbeing (0.55).

**Table 1** The GENCAT Scale calibration

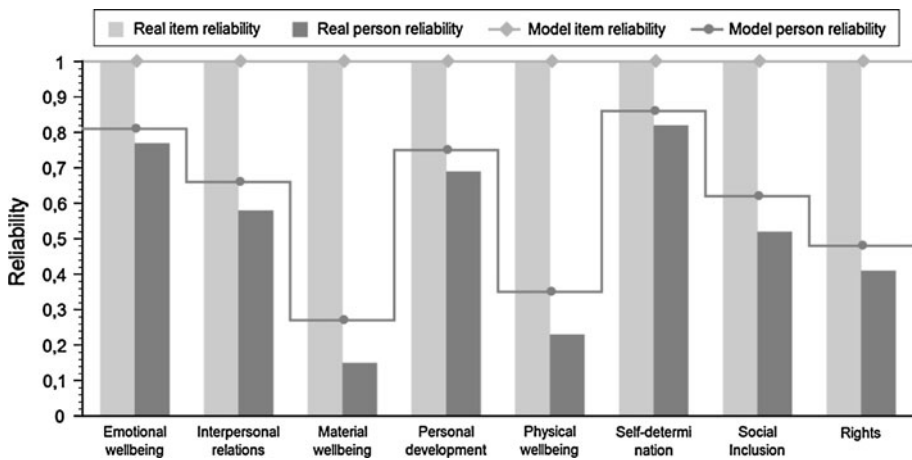
| Items and domains   | Measure | Model S.E. | Infit MNSQ | Outfit MNSQ |
|---|---------|------------|------------|-------------|
| <i>Emotional wellbeing</i>  |         |            |            |             |
| 8 He/she is motivated when performing some kind of activity   | .71     | .03        | 1.31       | 1.31        |
| 6 He/she is satisfied with themselves   | .57     | .03        | .79        | .77         |
| 3 He/she is happy and in a good mood  | .47     | .03        | .59        | .62         |
| 1 He/she is satisfied with their present life   | .40     | .03        | .77        | .75         |
| 4 He/she expresses feelings of helplessness or insecurity   | -.03    | .03        | 1.07       | 1.16        |
| 5 He/she shows symptoms of anxiety  | -.46    | .03        | 1.05       | 1.08        |
| 2 He/she shows symptoms of depression   | -.75    | .03        | .94        | .94         |
| 7 He/she has problems of conduct  | -.92    | .03        | 1.48       | 1.55        |
| <i>Interpersonal relations</i>  |         |            |            |             |
| 9 Most of the people with whom they interact are in a similar situation to their own  | 1.95    | .02        | 1.18       | 1.31        |
| 8 He/she says they feel loved by the people who are important to them   | 1.36    | .02        | 1.11       | 1.47        |
| 5 He/she says they feel undervalued by their family   | .83     | .02        | 1.36       | 1.36        |
| 1 He/she does things they enjoy with other people   | .33     | .02        | .78        | .79         |
| 7 He/she gets on well with their colleagues at work   | .19     | .02        | .70        | .67         |
| 6 He/she finds it difficult to start up a relationship with a potential partner   | -.19    | .02        | .71        | .73         |
| 4 He/she has a negative view of their friendships   | -.79    | .02        | .99        | .91         |
| 2 The relations with his/her family are as they would like them to be   | -1.03   | .03        | .95        | .94         |
| 3 He/she complains about a lack of close friends  | -1.14   | .03        | .95        | .99         |
| 10 He/she has a satisfactory sex life   | -1.52   | .04        | 1.88       | 2.10        |
| <i>Material wellbeing</i>   |         |            |            |             |
| 7 He/she does not earn enough to be able to afford luxuries   | 1.47    | .02        | 1.16       | 1.18        |
| 4 He/she is unhappy with where they live  | .46     | .03        | 1.17       | 1.29        |
| 6 He/she has enough money to cover their basic needs  | .30     | .03        | .85        | .76         |
| 3 He/she has the material possessions they need   | .29     | .03        | .73        | .69         |
| 1 Where he/she lives stops them from leading a healthy life (noise, fumes, odours, gloom, lack of ventilation, damage, inaccessibility, etc.) | -.24    | .04        | 1.60       | 1.37        |
| 8 Where he/she lives has been adapted to their needs  | -.41    | .04        | .98        | .90         |
| 2 His/her workplace complies with rules on health and safety  | -.86    | .05        | 1.13       | .33         |
| 5 Where he/she lives is clean   | -1.01   | .05        | 1.11       | .34         |
| <i>Personal development</i>   |         |            |            |             |
| 2 He/she has access to new technologies (Internet, mobile phone, etc.)  | 1.14    | .02        | 1.53       | 1.23        |
| 7 He/she is involved in the drafting of their own individual programme  | .77     | .02        | 1.23       | 1.21        |
| 4 He/she finds it difficult to effectively deal with the problems they have to face   | .22     | .02        | .79        | .82         |

**Table 1** continued

| Items and domains  | Measure | Model S.E. | Infit MNSQ | Outfit MNSQ |
|--|---------|------------|------------|-------------|
| 3 The work they do enables them to learn new skills  | .18     | .02        | .83        | .83         |
| 5 He/she does their work competently and responsibly   | .03     | .02        | .68        | .67         |
| 1 He/she finds it difficult to cope with everyday situations   | -.13    | .02        | .81        | .91         |
| 8 He/she lacks motivation at work  | -.58    | .02        | 1.03       | 1.02        |
| 6 The service he/she attends caters for their personal development and the learning of new skills            | -1.63   | .03        | 1.28       | 1.20        |
| <i>Physical wellbeing</i>  |         |            |            |             |
| 4 His/her state of health allows them to lead a normal life  | 1.21    | .02        | .88        | .83         |
| 7 His/her health problems cause them pain and discomfort   | 1.05    | .02        | .76        | .77         |
| 1 He/she finds it difficult to sleep   | .50     | .02        | .95        | .95         |
| 3 He/she has healthy eating habits   | .08     | .03        | .93        | .93         |
| 2 Technical aids are available if he/she needs them  | -.28    | .03        | 1.25       | 1.23        |
| 8 He/she finds it difficult to access healthcare resources (preventive care, GP, at home, in hospital, etc.) | -.29    | .03        | 1.83       | 1.62        |
| 5 He/she maintains good personal hygiene   | -.39    | .03        | 1.03       | 1.03        |
| 6 The service he/she attends supervises the medication they take   | -1.88   | .07        | 1.64       | 1.36        |
| <i>Self-determination</i>  |         |            |            |             |
| 9 He/she chooses who they live with  | 1.34    | .03        | 1.55       | 1.36        |
| 1 He/she has personal targets, goals and interests   | .81     | .02        | 1.23       | 1.30        |
| 8 He/she organises their own life  | .52     | .02        | .58        | .55         |
| 5 Other people decide upon his/her personal life   | -.03    | .02        | .62        | .62         |
| 6 Other people decide how he/she spends their money  | -.12    | .02        | .93        | .89         |
| 2 He/she decides how to spend their free time  | -.20    | .02        | .75        | .77         |
| 7 Other people decide what time he/she goes to bed   | -.23    | .02        | 1.17       | 1.08        |
| 4 He/she defends their ideas and opinions  | -.39    | .02        | .90        | .88         |
| 3 The service he/she attends caters for their preferences  | -1.71   | .03        | 1.57       | 2.06        |
| <i>Social inclusion</i>  |         |            |            |             |
| 7 The only friends he/she has are the ones who attend the same service                                       | 1.07    | .02        | 1.21       | 1.22        |
| 1 He/she frequents communal areas (public swimming pools, cinemas, theatres, museums, libraries, etc.)       | .95     | .02        | 1.24       | 1.22        |
| 5 His/her friends provide support whenever it is needed  | .50     | .02        | .76        | .77         |
| 2 His/her family provides support whenever needed  | .09     | .02        | 1.23       | 1.28        |
| 3 There are physical, cultural or social barriers that hinder his/her social inclusion                       | -.21    | .02        | .90        | .90         |
| 4 He/she lacks the necessary support for taking an active part in everyday life in their community           | -.47    | .02        | .80        | .80         |
| 6 The service he/she attends encourages them to take part in community activities                            | -.65    | .02        | .91        | .91         |
| 8 He/she is rejected or discriminated against by others  | -1.28   | .03        | .90        | .90         |

**Table 1** continued

| Items and domains  | Measure | Model S.E. | Infit MNSQ | Outfit MNSQ |
|--|---------|------------|------------|-------------|
| <i>Rights</i>  |         |            |            |             |
| 4 He/she finds it difficult to defend their rights when these are violated   | 1.34    | .02        | 1.01       | 1.04        |
| 3 He/she has information on their basic rights as a citizen  | 1.19    | .02        | .91        | .88         |
| 7 One or more of his/her legal rights have been impaired (citizenship, vote, legal processes, respect for their beliefs, values, etc.) | .84     | .02        | 1.33       | 1.17        |
| 1 His/her family violates their privacy (reading their letters, entering without knocking, etc.)                                       | .66     | .03        | 1.18       | 1.16        |
| 5 The service he/she attends respects their privacy  | -.13    | .03        | .69        | .76         |
| 2 He/she is treated with respect in their environment  | -.15    | .03        | .89        | 1.05        |
| 6 The service he/she attends respects their possessions and their ownership rights   | -.56    | .04        | .88        | .81         |
| 8 The service he/she attends respects and defends their rights (confidentiality, information on their rights as a user...)             | -.66    | .05        | .86        | .73         |
| 9 The service respects the privacy of his/her information  | -.85    | .05        | .94        | .75         |
| 10 He/she is exposed to exploitation, violence or abuse  | -1.67   | .08        | 2.10       | 2.12        |

**Fig. 1** Item and person separation reliability

### 3.5 Calibration

The next step involved the calibration of the items (Table 1). Generally speaking, we can conclude that there is equilibrium in all the scales regarding the number of difficult items (above 0 logits) and easy items (below 0 logits). Regarding the distribution along the continuum of the dimension of quality of life they are evaluating, Interpersonal relations is the one with the greatest range and best distribution of items together with Social inclusion.



On the other hand, the dimensions with the largest gaps between items are Material wellbeing, Personal development, Physical wellbeing, Self-determination and Rights. Finally, Emotional wellbeing is the subscale whose items are grouped into a smaller range, albeit evenly distributed without large gaps between them.

### 3.6 Items' Level of Difficulty

An analysis of the suitability of the items' level of difficulty for the sample confirms these results, highlighting the almost perfect adaptation of the items in Interpersonal relations. For Emotional wellbeing, by contrast, it would be advisable to include not only easier items but also, and above all, more difficult ones. It is highly advisable to include more difficult items in Material wellbeing, Physical wellbeing and Rights, as most of those included are too easy for the participants. In Personal development and Self-determination it would be convenient above all to include more difficult items, although there is a noticeable lack of one or two in the easier levels. Finally, in Social inclusion it would be appropriate to include more difficult items, whilst one item (*'He/she is rejected or discriminated against by others'*) is too easy for the participants.

### 3.7 Response Categories

Regarding the adaptation of the response categories to the sample, we observe that, although the extreme categories are the most probable ones in all cases, the four response options are suitable in all the subscales, except for Interpersonal relations, Material wellbeing and Rights, in which option 2 ('sometimes' or 'often' according to the item's valence) is not very suitable. Furthermore, the information curves for the categories reveal that the extreme options provide the most information ('never or hardly ever' in first place, followed by 'always or almost always') in Interpersonal relations, Material wellbeing and Rights. Emotional wellbeing is the only subscale in which the intermediate options ('sometimes' or 'often') provide more information than the extreme ones. In Self-determination and Social inclusion, the category that provides the most information is 'often', followed by 'never or hardly ever'. Finally, the category in Social inclusion and Physical wellbeing providing the most information is 'never or hardly ever', with the least being provided by 'sometimes'.

### 3.8 Differential Item Functioning

Differential item functioning (DIF) was analysed by *gender* and by group or *condition*. On one hand, the analysis of the DIF between men and women confirmed solely the differential function of one item: i39 *'He/she maintains good personal hygiene'* in Physical wellbeing, which was much easier for women than for men. However, if we consider its content, we cannot affirm that there is a bias in favour of one gender or the other. On the other hand, the analysis of the DIF confirmed only the differential function of 10 of them: *'He/she has problems of conduct'* in Emotional wellbeing, which was more difficult for people in a situation of social disadvantage; *'Most of the people with whom he/she interacts are in a similar situation to their own'* in Interpersonal relations, which was more difficult for elderly people; *'He/she has access to new technologies'* in Personal development, which was more difficult for elderly people; *'Technical aids are available if he/she needs them'*; *'He/she has healthy eating habits'* and *'He/she maintains good personal hygiene'*,

which were more difficult for people at social disadvantage; '*He/she has personal targets, goals and interests*' and '*He/she chooses who they live with*' in Self-determination, the former is more difficult for the elderly whilst the latter is for people at social disadvantage; '*He/she frequents communal areas*' in Social inclusion, was much more difficult for the elderly; and in Rights '*One or more of his/her legal rights has been impaired*' was more difficult for people at social disadvantage. This meant that Physical wellbeing and Self-determination were the most problematic subscales in this sense. We thus note that 6 of the 10 items specified are more difficult for people at risk of social exclusion, whereby there could be a certain bias for this collective in these items.

#### 4 Discussion

It is worth stressing what, in general terms, we believe to be the main contribution of the analyses made from the perspective of the IRT: not only do they corroborate the results obtained through Classical Test Theory (CTT), but they also act as a complement by shedding some light on their possible interpretations or explanations. There follows an overview of some of the analyses' more specific conclusions: (1) the data fitted the model, with the only exception of the item 10 in Rights, that is highly recommended to be eliminated; (2) the highest point-biserial correlations (whose interpretation is similar to the alpha coefficient in the CTT) were recorded in Self-determination and Emotional wellbeing; the lowest appeared in Physical wellbeing; (3) the dimensions Rights and Material wellbeing recorded the most extreme scores (i.e., very high scores); (4) most of the items are considered to be very accurate: 59 (85.5%) have a measurement error of .02 to .03; (5) the most accurate items (i.e., whose observed score is closest to the person's true score in the construct evaluated) are in the dimensions Self-determination, Emotional wellbeing, Personal development and Social inclusion (in descending order); on the other hand, the least accurate ones are (in ascending order) in Rights, Physical wellbeing, Material wellbeing and Interpersonal relations; (6) there is a major discrepancy between the excellent reliability or separation of the items and the moderate or low reliability of the individuals: this means it would be convenient to include a number of more difficult items in these dimensions in order to raise their level of accuracy (not just simply to increase their number, but rather to adjust the items' difficulty to the level of competence); (7) the calibration of the items shows that the dimensions Interpersonal relations, Self-determination and Rights cover a greater range of personal results compared to all the others: this phenomenon could indeed be explained by the greater number of items included in them; (8) the dimension Interpersonal relations is the one with the greatest suitability in terms of the difficulty of the items for the sample; (9) the response categories are generally suitable: nevertheless, the extreme categories ('never or hardly ever' and 'always or almost always') are the ones used most compared to the intermediate ones. The option 'often' is used much more than 'sometimes' (i.e., for the items of positive valence; and the opposite occurs for those of negative valence). The latter, furthermore, does not seem to be appropriate for Interpersonal relations, Material wellbeing and Rights, given that there are few opportunities for it to be used in these dimensions; (10) as a whole, the items and dimensions provide the most information in intermediate ability levels, whilst the level of information drops in the lower levels; (11) there is no differential item functioning (DIF) by gender with the exception of '*He/she maintains good personal hygiene*' in Physical wellbeing which, a priori and strictly applying the assumptions of the Rating Scale Model, should be deleted due to a possible bias in favour of women; that is, it appears to measure men and

women differently, whereby it might reflect a content more closely associated with social aspects of men and women instead of with attributes of Physical wellbeing; and (12) there is a differential function in 10 of the 69 items (14.5%) by condition or collective, with the sole exception of Material wellbeing, where there do not appear to be any biases. We thus find that 6 of the 10 items specified are more difficult for the people at risk of social exclusion and the other four are easier for them than for elderly people. This points to a possible bias in favour of one collective or the other. Nevertheless, a qualitative analysis of the items with differential function does not permit us to single out any explanation or characteristic that might be detrimental to one of the groups and which is not directly related to those people's quality of life, so we conclude that they are not biased items.

As the main conclusion of this work, we should note that this study is a first and unprecedented approach to the evaluation of the quality of life of the users of social services. We can affirm that the GENCAT Scale has sufficient evidence regarding its validity for evaluating the quality of life of users of social services in Catalonia and it stands as the only instrument so far available that is sensitive to those intervention programmes designed to improve personal results (and therefore the quality of life of users) and so can be of considerable use when applied in the services and in the development and assessment of programmes. Indeed, the results forthcoming comfortably outperform those provided by other instruments designed to evaluate quality of life. Nonetheless, we do not dismiss other kinds of complementary evaluation; quite the contrary, we consider it highly recommendable to apply it together with, for example, subjective evaluations of quality of life (e.g., INTEGRAL Scale; Verdugo et al. 2009).

**Acknowledgments** Work on this research was partially supported by a grant (EDU/894/2009) given to the Excellence Group of Research (GR197) by the Junta de Castilla y León (Spain). It was also founded by the Catalan Welfare and Social Services Institute (ICASS, Generalitat of Catalonia) and the Science and Technology Ministry (R&D Projects, 2009) (PSI2009-10953).

## References

- Andrich, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, 2, 581–594. doi:10.1177/014662167800200413.
- Bonham, G. S., Basehart, S., Schalock, R. L., Marchand, C. G., Kirchner, N., & Rumenap, J. M. (2004). Consumer-based quality of life assessment: The Maryland Ask Me! Project. *Mental Retardation*, 42, 338–355. doi:10.1352/0047-6765(2004)42<338:CQOLAT>2.0.CO;2.
- Claes, C., van Hove, G., van Loon, J., Vandeveld, S., & Schalock, R. L. (2009). Quality of life measurement in the field of intellectual disabilities: Eight principles for assessing quality of life-related personal outcomes. *Social Indicator Research*. doi:10.1007/s11205-009-9517-7.
- Cummins, R. A. (2005). Moving from the quality of life concept to a theory. *Journal of Intellectual Disability Research*, 49, 699–706. doi:10.1111/j.1365-2788.2005.00738.x.
- Jenaro, C., Verdugo, M. A., Caballo, C., Balboni, G., Lachapelle, Y., Otbrebski, W., et al. (2005). Cross-cultural study of person-centered quality of life domains and indicators: A replication. *Journal of Intellectual Disability Research*, 49, 734–739. doi:10.1111/j.1365-2788.2005.00742.x.
- Keith, K. D., & Bonham, G. S. (2005). The use of quality of life data at the organization and systems level. *Journal of Intellectual Disability Research*, 49, 799–805. doi:10.1111/j.1365-2788.2005.00755.x.
- Keith, K. D., & Schalock, R. L. (2000). *Cross-cultural perspectives on quality of life*. Washington, DC: American Association on Mental Retardation.
- Linacre, J. M. (2002). What do infit and outfit, mean square and standardized mean? *Rasch Measurement Transactions*, 12(2), 878. Retrieved from <http://www.rasch.org/rmt/rmt162f.html>.
- Linacre, J. M. (2008). *Winsteps. Rasch measurement computer program* (computer software). Chicago, IL: Winsteps.com.

- Linacre, J. M. (2010, January 16). *WINSTEPS & facets Rasch software*. Retrieved from <http://www.winsteps.com/index.htm>.
- Linacre, J. M., & Wright, B. D. (1999). *A user's guide to WINSTEPS, BIGSTEPS, MINISTEP Rasch model computer programs*. Chicago, IL: MESA Press.
- Perry, J., & Felce, D. (2005). Factors associated with outcomes in community group homes. *American Journal on Mental Retardation*, *110*, 121–135. doi:10.1352/0895-8017(2005)110<121:FAWOIC>2.0.CO;2.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: MESA Press.
- Rasch, G. (1992). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: MESA Press.
- Schalock, R. L., & Bonham, G. S. (2003). Measuring outcomes and managing for results. *Evaluation & Program Planning*, *26*, 229–235.
- Schalock, R. L., & Felce, D. (2004). Quality of life and subjective well-being: Conceptual and measurement issues. In E. Emerson, C. Hatton, T. Thompson, & T. R. Parmenter (Eds.), *International handbook of applied research in intellectual disabilities* (pp. 261–279). London: Wiley.
- Schalock, R. L., Gardner, J. F., & Bradley, V. J. (2007). *Quality of life for people with intellectual and other developmental disabilities. Applications across individuals, organizations, communities, and systems*. Washington, DC: American Association on Intellectual and Developmental Disabilities.
- Schalock, R. L., Keith, K. D., Verdugo, M. A., & Gomez, L. E. (2010). Quality of life model development in the field of intellectual disability. In R. Kober (Ed.), *Quality of life for people with intellectual disability* (pp. 17–32). New York: Springer.
- Schalock, R. L., & Verdugo, M. A. (2002). *Quality of life for human service practitioners*. Washington, DC: American Association on Mental Retardation.
- Schalock, R. L., & Verdugo, M. A. (2007). El concepto de calidad de vida en los servicios y apoyos para personas con discapacidad intelectual. *Siglo Cero*, *38*, 21–36.
- Schalock, R. L., Bonham, G. S., & Verdugo, M. A. (2008a). The conceptualization and measurement of quality of life: Implications for program planning and evaluation in the field of intellectual disabilities. *Evaluation and Program Planning*, *31*(2), 181–190. doi:10.1016/j.physleth.2003.10.071.
- Schalock, R. L., Verdugo, M. A., Bonham, G. S., Fantova, F., & Van Loon, J. (2008b). Enhancing personal outcomes: Organizational strategies, guidelines, and examples. *Journal of Policy and Practice in Intellectual Disabilities*, *5*, 18–28. doi:10.1111/j.1741-1130.2007.00135.x.
- Schalock, R. L., Verdugo, M. A., Jenaro, C., Wang, M., Wehmeyer, M., Xu, J., et al. (2005). A cross-cultural study of quality of life indicators. *American Journal on Mental Retardation*, *110*, 298–311. doi:10.1352/0895-8017(2005)110[298:CSOQOL]2.0.CO;2.
- Van Loon, J., van Hove, G., Schalock, R. L., & Claes, C. (2008). *Personal outcomes scale*. Middleburg, Holland: Arduin Steichlich.
- Verdugo, M. A., Arias, B., Gómez, L. E., & Schalock, R. L. (2008a). *Formulari de l'Escala Gencat de Qualitat de vida. Manual d'aplicació de l'Escala Gencat de Qualitat de vida*. Barcelona: Departamento de Acción Social y Ciudadanía, Generalitat de Cataluña.
- Verdugo, M. A., Arias, B., Gómez, L. E., & Schalock, R. L. (2008b). *Escala GENCAT. Informe sobre la creació d'una escala multidimensional per avaluar la qualitat de vida de les persones usuàries dels serveis socials a Catalunya*. Barcelona: Departamento de Acción Social y Ciudadanía, Generalitat de Cataluña.
- Verdugo, M. A., Arias, B., Gómez, L. E., & Schalock, R. L. (2009a). *Formulario de la Escala GENCAT de Calidad de vida. Manual de aplicación de la Escala GENCAT de Calidad de vida*. Barcelona: Departamento de Acción Social y Ciudadanía, Generalitat de Cataluña.
- Verdugo, M. A., Arias, B., Gómez, L. E., & Schalock, R. L. (2010). Development of an objective instrument to assess quality of life in social services: Reliability and validity in Spain. *International Journal of Clinical and Health Psychology*, *10*(1), 105–123. Retrieved from [http://www.aepc.es/ijchp/articulos\\_pdf/ijchp-343.pdf](http://www.aepc.es/ijchp/articulos_pdf/ijchp-343.pdf).
- Verdugo, M. A., Gómez, L. E., & Arias, B. (2007a). La Escala INTEGRAL de calidad de vida: desarrollo y estudio preliminar de sus propiedades psicométricas. *Siglo Cero*, *38*, 37–56.
- Verdugo, M. A., Gómez, L. E., Arias, B., & Schalock, R. L. (2009b). *Escala Integral de Calidad de Vida*. CEPE: Madrid.
- Verdugo, M. A., Schalock, R. L., Gómez, L. E., & Arias, B. (2007b). Construcción de escalas de calidad de vida multidimensionales centradas en el contexto: la Escala GENCAT. *Siglo Cero*, *38*, 57–72.

- Verdugo, M. A., Schalock, R. L., Keith, K., & Stancliffe, R. (2005). Quality of life and its measurement: Important principles and guidelines. *Journal of Intellectual Disability Research*, *49*, 707–717. doi: [10.1111/j.1365-2788.2005.00739.x](https://doi.org/10.1111/j.1365-2788.2005.00739.x).
- Walsh, P., Erickson, E., Bradley, V., Moseley, C., & Schalock, R. L. (2006). *Supported accommodation services for people with intellectual disabilities: A review of models and instruments used to measure quality of life in various settings*. Dublin: National Disability Authority.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.