

# Questions on Life Satisfaction (FLZ<sup>M</sup>) – A Short Questionnaire for Assessing Subjective Quality of Life

Gerhard Henrich and Peter Herschbach

Institut und Poliklinik für Psychosomatische Medizin, Psychotherapie und Medizinische Psychologie, Klinikum rechts der Isar, Munich, Germany

**Keywords:** Quality of life, life satisfaction, assessment, questionnaire, weighting

**Summary:** A short questionnaire for assessing general and health-related quality of life is presented. Test construction and psychometric evaluation are described and recent findings obtained with the instrument reported and plans for developing additional modules outlined. Questions on Life Satisfaction<sup>Modules</sup> (in the original German: FLZ<sup>M</sup>, Fragen zur Lebenszufriedenheit<sup>Module</sup>) combines three features: economy, modular structure, and individual weighting of items. The questionnaire consists of two 8-item modules, “General Life Satisfaction” and “Satisfaction with Health.” The respondent rates each item twice, once for the subjective importance of the aspects of life or health addressed, and once for the degree of satisfaction in that area. The two ratings are combined into a “weighted satisfaction” score. The total score is the sum of these eight scores.

Norms for Germany and Spain have been established. Data for patients from different diagnostic groups in Germany are also available. The work in progress on international versions (UK English, US English, Dutch, Spanish, and Italian) and on additional modules is discussed.

## Introduction

This paper describes a quality-of-life questionnaire that combines three features: economy, modular structure, and individual weighting of items. The questionnaire consists of a generic module and a health-related module.

Following initial uncertainties in defining the psychological construct of quality of life (QL), there is now a broad consensus in expert circles that two elements are essential, namely, multidimensionality and subjectivity (see Cella, 1998). Multidimensionality means that the definition must cover different relevant aspects or dimensions of QL, at a minimum physical, mental, and social aspects. Subjectivity reflects the everyday experi-

ence that the specific factors determining the quality of a person's life are both highly personal and virtually limitless. Moreover, there are great inter- and intraindividual differences in the perception and evaluation of objective aspects of life or disease. Therefore, it seems impossible to measure QL by assessing a list of specific (objective) factors. Consistent with this view, numerous studies have found only low correlation between objective circumstances and subjective QL. An example that is frequently cited is that certain individuals, among them lottery winners, generally consider themselves much less happy than an outsider would expect (Brickman, Coates, & Janoff-Bulman, 1978). On the other hand, many studies by social psychologists have revealed a high percentage of satisfied individuals, even among persons who clearly have major problems (Ipsen, 1978; Glatzer &

Zapf, 1984; Veenhoven, 1991). With reference to health, the “illness paradox” describes the discrepancy between “objectively” poor health (as assessed by an outside observer) and a positive self-assessment. For example, numerous studies have found that patients with cancer give above-average ratings in regard to their QL (Cassileth, Lusk, & Tenaglia, 1982; Henrich & Herschbach, 1995; Krischke, 1996). The different criteria used by those affected and by outside observers lead to these apparently contradictory results. Subjective QL reflects the difference between an individual’s hopes, expectations, and desires and what he or she considers as reality (Calman, 1984).

If multidimensionality and subjectivity are to be combined in a definition of QL, logic demands weighting the individual dimensions included, particularly if a summary score is to be calculated. It makes a big difference psychologically whether one is dissatisfied with an area of life one regards as unimportant or with an area one regards as important.

There are a great many general and illness-specific questionnaires on QL (McDowell & Newell, 1987; Bowling, 1991; WHO, 1994). In a review of a representative selection of 75 empirical studies on QL using a total of 159 different instruments, Gill and Feinstein (1994) criticized in particular that only six of the articles included a rating of the importance of individual problems (or items), and that this rating was considered in the total score in only three of the six. The issue of importance is addressed in QL questionnaires such as the Quality of Life Index (Ferrans & Powers, 1985) and the Functional Assessment of Cancer Therapy Scale (FACT; Cella et al., 1993). But these instruments are too long to be practical for use with patients with advanced disease, they do not provide direct importance ratings for each item, are too specific for use with healthy people, do not take the importance ratings into account in scoring, or a combination of these things. The Schedule for the Evaluation of Individual Quality of Life (SEIQoL; McGee, O’Boyle, Hickey, O’Malley, & Joyce, 1991; short form: SEIQoL-DW; Hickey, Bury, O’Boyle, Bradley, O’Kelly, & Shannon, 1996) is an interview-based instrument developed to assess QL from the individual’s perspective, which quantifies the relative importance of those areas of life nominated by the respondent.

The FLZ<sup>M</sup> (Fragen zur Lebenszufriedenheit<sup>Module</sup> = Questions on Life Satisfaction<sup>Modules</sup>; Huber, Henrich, & Herschbach, 1988; Roder, Herschbach, Sellschopp, & Siewert, 1991; Henrich & Herschbach, 1995) is an economic questionnaire that operationalizes a subjective concept of QL as described above (Calman, 1984). Moreover, in its development the attempt was made to deal adequately with the problem of the relative importance of individual aspects of QL. The FLZ<sup>M</sup> is designed

to assess both generic aspects of the QL of healthy and ill people, via its general module, and disease- and treatment-specific QL, via specific modules. In the present paper we describe the instrument, outline its development, present recent findings with regard to item and test characteristics as well as normative and comparative data, and discuss work in progress on international versions and additional modules.

## Description of the FLZ<sup>M</sup>

Development of the FLZ<sup>M</sup> began in 1986 and included several phases of data collection with healthy and ill individuals in accordance with the criteria of classic test theory. Based on statistical analyses and feedback from the respondents, the number and wording of the items, the number and wording of the response categories, and the form of the instructions were modified and optimized.

The respondent fills out the FLZ<sup>M</sup> questionnaire, which presently consists of two parts, or modules, “General Life Satisfaction” (FLZ<sup>M</sup>-“General LS”) and “Satisfaction with Health” (FLZ<sup>M</sup>-“Health”). Two modules related to specific illnesses (gastrointestinal problems and growth-hormone deficiency) are under development. The general module can be used in all cases and should be the first page of any future disease-specific modules. Each module consists of a single sheet of paper containing both instructions and items. The FLZ<sup>M</sup> modules have been evaluated for ease of comprehension, and (judging from the small number of missing data; see Table 1) they appear to be easily understood. They can be completed in a few minutes even by the elderly and by those with acute or severe illness.

The instructions state that the ratings refer to the past four weeks. This is to differentiate life satisfaction (LS) from the psychological construct of “mood,” which describes a momentary state and which can vary over the course of the day. The respondent is first asked to rate the “subjective importance” of a given area of life (“How important [is item X] for your overall satisfaction?”). In this way the principle of individual weighting is realized. Then the respondent is asked about the degree of satisfaction in that area. All responses are given on a five-point scale.

The scale used for rating satisfaction was initially a symmetrical four-point scale. It is now asymmetrical, with two negative and three positive responses to choose from. The reason for this change was that in the original version many responses were in the two available positive categories. The change increased the degree of differentiation possible in the positive range. A category

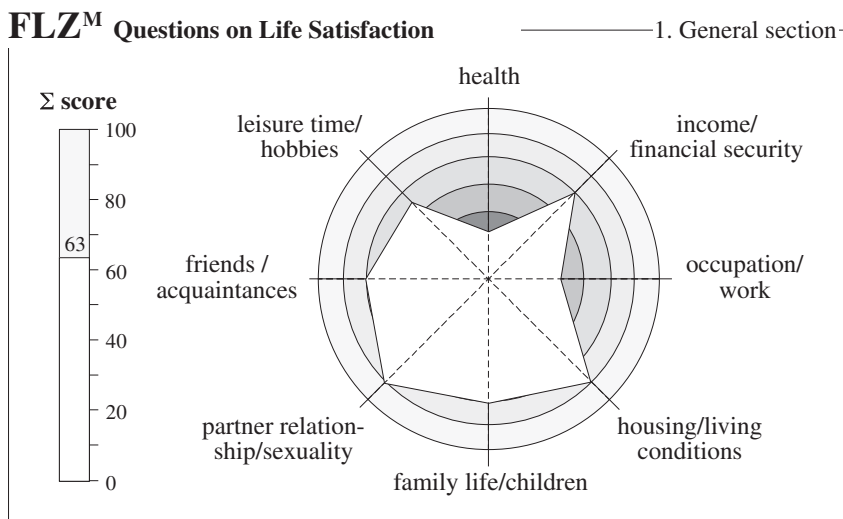


Figure 1. Graphical representation of a subject's response on the FLZ<sup>M</sup> (the total score on the left has been transformed to a range of 0 to 100, where 100 = maximum life satisfaction).

“neither satisfied nor dissatisfied” was omitted intentionally to force a positive or negative decision. The scale for importance, which was also initially a four-point scale (“not important” to “very important”) was expanded to include the category “extremely important,” resulting in a distribution of responses that was less skewed to the left.

The module FLZ<sup>M</sup>-“General LS” covers eight areas of life that are usually relevant to some degree for everyone in the Western world (“friends/acquaintances,” “leisure time/hobbies,” “health,” “income/financial security,” “occupation/work,” “housing/living conditions,” “family life/children,” and “partner relationship/sexuality”). It can therefore be used with very different groups of subjects, making comparisons possible. The eight items are the result of principal components analysis and reliability analyses of an initial 16-item version of the module, which was developed after a study of the literature and interviews with experts and patients. This fixed item list still seems restrictive compared to a judgment analysis method like the SEIQoL, but it contains the aspects of life postulated as being most important to quality of life (see, for example, Table 2 of McGee et al., 1991), provides the opportunity to indicate the relative importance of each, and, finally, is much easier to administer.

The second module, FLZ<sup>M</sup>-“Health,” includes the eight dimensions “physical condition/fitness,” “ability to relax/stay on an even keel,” “energy/zest for life,” “mobility (e. g., walking, driving),” “vision and hearing,” “freedom from anxiety,” “freedom from aches and pains” and “independence from help/care.” The initial version of this module consisted of 13 items.

In the evaluation of the responses, the ratings for importance and satisfaction are combined to yield information about “weighted satisfaction” (wS), which varies between -12 and +20. The weighting formula is: wS =

importance rating  $\times [(2 \times \text{satisfaction rating}) - 3]$ , provided both ratings are made on scales ranging from 0 to 4. The measure of global LS (with reference to the content of a given module) is the sum of the wS values. Combinations with “unsatisfied” yield a negative value for wS, which leads to a reduction in the total score. In other words, dissatisfaction in any area of life reduces LS as a whole. Figure 1 shows a way of representing the FLZ<sup>M</sup> results graphically, which has proven useful, especially for individual cases.

## Psychometric Evaluation

The final determination of item and test characteristics was based on three representative samples in Germany (total  $N = 7796$ ; age: mean = 46.0 years,  $SD = 17.6$ ; gender: 47.0% male; marital status: 62.4% married; employment status: 40.6% full-time). Data were collected in 1991, 1994, and 1995 by two commercial institutes specializing in social-science research. The general population consisted of all individuals who were German citizens and at least 18 years of age when the samples were selected. The respondents were instructed to complete the questionnaire themselves. The psychometric evaluation also included numerous groups of patients and healthy subjects who were participating in studies being conducted by the authors or their colleagues at other institutions.

## Item Characteristics

Table 1 gives item characteristics and total scores for the two modules “General LS” (upper part, 1994 sample)

Table 1. Score distribution and item characteristics of the German FLZ<sup>M</sup> modules (General LS, *N* = 2562; Health, *N* = 2226).

	M.d.	Mean	SD	Floor	Ceiling	IC
FLZ <sup>M</sup> General LS						
Friends/acquaintances	1.0%	8.1	6.3	0.0%	10.3%	0.51
Leisure time/hobbies	1.2%	6.3	6.3	0.1%	7.5%	0.54
Health	0.8%	8.1	7.5	1.7%	14.6%	0.56
Income/financial security	1.0%	6.5	7.3	1.8%	9.7%	0.57
Occupation/work	3.9%	5.5	7.3	2.0%	7.6%	0.55
Housing/living conditions	1.1%	8.3	6.4	0.6%	9.6%	0.57
Family life/children	1.7%	9.8	6.9	0.2%	19.8%	0.55
Partner relationship/sexuality	2.1%	7.9	7.7	1.0%	16.4%	0.52
Total score	1.1%	60.5	37.3	0.0%	1.6%	0.82
FLZ <sup>M</sup> Health						
Physical condition/fitness	0.3%	8.1	7.0	1.1%	10.9%	0.74
Ability to relax/stay on an even keel	0.5%	7.4	6.5	0.4%	7.8%	0.65
Energy/zest for life	0.5%	9.1	6.5	0.3%	13.6%	0.70
Mobility (e. g., walking, driving)	0.7%	9.1	7.0	0.2%	15.3%	0.64
Vision and hearing	0.4%	11.0	7.0	0.6%	24.5%	0.64
Freedom from anxiety	1.0%	8.1	6.7	0.8%	10.8%	0.62
Freedom from aches and pains	0.4%	9.1	7.4	0.9%	16.5%	0.74
Independence from help/care	0.5%	12.5	6.7	0.5%	31.2%	0.66
Total score	0.4%	74.4	41.5	0.0%	3.4%	0.89

M.d. = Missing data, IC = Internal consistency (items: part-whole correlation; total score: Cronbach's  $\alpha$ )

and "Health" (lower part, 1995 sample). Included are the percentage of missing data, mean and standard deviation, floor and ceiling effects, and part-whole correlation of the items with the scale value, or with Cronbach's  $\alpha$  for the whole scale.

Overall, the frequency of missing data on the FLZ<sup>M</sup>-"General LS" is very low. The item means for wS are relatively close to the middle of the range of possible values (-12 to +20). Floor effects are negligible. Ceiling effects are relatively low considering the fact that this was a representative sample of the population. The part-whole correlations between items and scale are between 0.50 and 0.60, which is in the desired range. Evaluation of test-retest reliability is in progress.

For the module "Health," the percentage of missing data is even lower, and as expected the ceiling effect for this (mainly healthy) sample is higher. The scales are more homogeneous, as evidenced by the high part-whole correlations. Evaluation of test-retest reliability is in progress.

## Test Characteristics

### Reliability

Internal consistency as a measure of the reliability of the scales can be considered to be high (lines in Table 1 labeled "Total score"). Cronbach's  $\alpha$  is 0.82 and 0.89, respectively. The test-retest reliability was evaluated in a small sample of medical students (*N* = 45; time lag: 1

week); the  $r_{tt}$  for the total scores was satisfactory (0.87 and 0.85, respectively).

### Content Validity

Content validity can be assumed to exist because of the assumptions on which the modules are based and the way the modules were developed.

### Convergent Validity

So far the FLZ<sup>M</sup> has been used with various samples of healthy subjects and patients in a total of 29 studies. In some of these studies internationally established measures of QL and other subjective measures of well-being have been used in addition. Table 2 shows Pearson product-moment correlations between such measures and the FLZ<sup>M</sup> total scores "General LS" and "Health." The table includes the name of the instrument used for validation, the correlation coefficients  $r$ , a brief description of the sample used and the sample size.

The FLZ<sup>M</sup> total score "General LS" has the highest correlations with those instruments that assess mainly psychological aspects of well-being, such as the General Well-Being Schedule (GWB; DuPuy, 1984), the Beck Depression Inventory (BDI; Beck, Ward, Mendelsohn, Mock, & Erbaugh, 1961) and the trait anxiety of the State-Trait-Anxiety-Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970). Here, the correlation coefficients lie between 0.48 and 0.63. The correlations are smaller with questionnaires that assess mainly physical

Table 2. Pearson product-moment correlations between the FLZ<sup>M</sup> modules and other internationally established well-being measures in different samples.

	FLZ <sup>M</sup> General LS	FLZ <sup>M</sup> Health	Samples	
			N	Description
GWB General Well-Being Schedule	0.63		57	Ulcer patients
SF-36 MOS Health Scale			3102	Pain patients (chron.)
Mental component summary	0.42	0.64		
Physical component summary	0.10	0.40		
NHP Nottingham Health Profile			3210	Pain patients (chron.)
Scale "Energy"	-0.32	-0.51		
Scale "Pain"	-0.16	-0.46		
Scale "Emotional reaction"	-0.43	-0.55		
Scale "Sleep"	-0.17	-0.38		
Scale "Social isolation"	-0.40	-0.39		
Scale "Physical mobility"	-0.14	-0.46		
Karnofsky Performance Scale	0.16	0.35	365	Cancer patients
	0.16		198	Cancer patients
	0.27		83	Cancer patients
SCL-90-R Symptom Checklist				
Scale "Somatization"	-0.34		2503	General population (Germany)
	-0.06	-0.44	81	Urological patients
Scale "Depression"	-0.40		2503	General population (Germany)
	-0.33	-0.59	81	Urological patients
BDI Beck Depression Inventory	-0.51	-0.55	180	Psychosomatic patients
STAI State-Trait-Anxiety Inventory				
Trait anxiety	-0.48	-0.54	1486	Pain patients (chron.)

problems, such as the "physical component summary" of the SF-36 MOS Health Scale (Ware et al., 1995), some scales of the Nottingham Health Profile (NHP; Hunt, McEwen, & McKenna, 1986), and "functional status" as rated by a physician on the Karnofsky Performance Scale (Karnofsky, Abelman, Craver, & Burchenal, 1948).

As expected, satisfaction with health as represented by the total score on FLZ<sup>M</sup>-"Health" has a higher correlation with the physical- and health-related QL measures than does "General LS." However, the correlations of 0.38 to 0.64 with the global scales of the SF-36 and the scales of the NHP, which correspond to a common variance of only 14% to 41%, are not very high. The reason for this is that these widely used scales, and also the EORTC for patients with cancer (Aaronson et al., 1993), assess mainly the functional status and complaints of the patients without taking into account a subjective evaluation and weighting. The correlations of the FLZ<sup>M</sup>-"Health" items with the subscales of the SF-36 and the NHP (not shown) provide an additional indicator of its convergent validity. The highest correlations are between the scales with similar content, e. g., "ability to relax/stay on an even keel" and "mental health" (SF-36;

$r = 0.57$ ), "energy/zest for life" and "vitality," "mental health" (SF-36;  $r = 0.58$  and  $0.60$ ), and "emotional reaction" (NHP;  $r = 0.57$ ); and "mobility" and "physical functioning" (SF-36;  $r = 0.60$ ) and "physical mobility" (NHP;  $r = 0.54$ ). The association with the physician-rated Karnofsky Performance Scale is low. The module FLZ<sup>M</sup>-"Health" also has a relatively high correlation with depression and anxiety (scale "Depression" on the Symptom Checklist SCL-90-R (Derogatis, 1977), BDI, and STAI) because it contains items related to these constructs (e. g., "ability to relax," "energy," "freedom of anxiety").

#### Discriminant Validity

Additional evidence for the construct validity of the FLZ<sup>M</sup> is that certain groups for whom differences in LS must be assumed to exist have different FLZ<sup>M</sup> scores. An example is a sample of women looking for work, whose LS as measured with the FLZ<sup>M</sup> was lower than that of the (female) normative sample ( $N = 1054$ , Western Germany, 1991) (mean of 32.6 vs. 62.1; Mann-Whitney  $U$ -test:  $z$  corrected for ties =  $-21.8$ ,  $p = .000$ ), but whose  $wS$

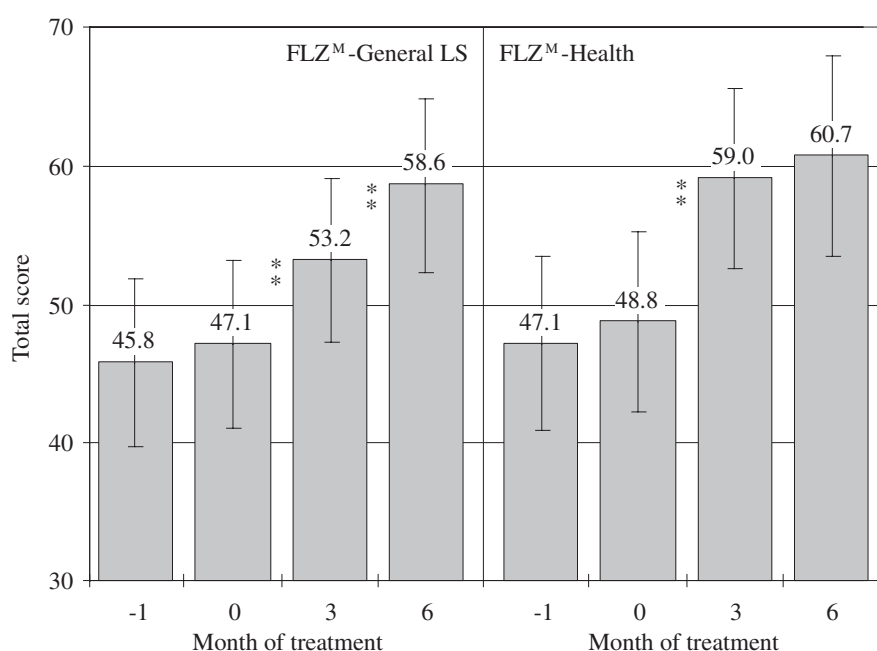


Figure 2. Responsiveness of the FLZ<sup>M</sup>-total scores during growth hormone replacement therapy in patients with growth hormone deficiency ( $N = 120$ ; months -1 to 0 = baseline without treatment; Wilcoxon matched-pairs signed-ranks test: \*\* $p \leq 0.01$ ; 95% confidence interval of the means are indicated).

Table 3. Normative data for the FLZ<sup>M</sup> total scores (representative samples of the German population).

Age (years)	Mean	SD	N	Male			Female		
				Mean	SD	N	Mean	SD	N
FLZ <sup>M</sup> General LS									
≤ 25	62.54	37.07	276	61.36	38.78	151	63.95	34.99	125
26 to 35	62.16	36.06	534	63.70	34.74	270	60.59	37.37	264
36 to 45	63.53	38.11	434	62.80	38.65	205	64.19	37.69	229
46 to 55	59.87	36.15	403	59.89	34.47	182	59.86	37.54	222
56 to 65	63.05	39.52	460	65.15	40.43	224	61.05	38.61	236
> 65	51.83	35.71	427	56.97	36.57	164	48.60	34.84	262
Total	60.49	37.31	2534	62.02	37.28	1197	59.13	37.31	1338
FLZ <sup>M</sup> Health									
≤ 25	92.47	38.40	281	92.41	37.99	160	92.55	39.09	121
26 to 35	88.86	41.50	466	93.52	44.03	229	84.36	38.45	237
36 to 45	82.67	38.44	365	80.86	37.78	175	84.35	39.07	189
46 to 55	69.59	34.42	373	73.15	35.41	173	66.49	33.32	199
56 to 65	62.24	37.61	378	62.35	39.85	176	62.14	35.63	202
> 65	50.66	41.86	356	55.22	41.03	130	48.03	42.19	226
Total	74.39	41.54	2218	77.80	41.96	1044	71.36	40.94	1174

for the area “occupation/work” was disproportionately low (deviation of  $-1.27$  SD; Mann-Whitney  $U$ -test:  $z$  corrected for ties =  $-30.4$ ,  $p = .000$ ).

Another example is the comparison of representative samples from Eastern and Western Germany in 1991. As might be expected considering the economic situation at that time, the wS differs most in the areas “income/financial security” (Mann-Whitney  $U$ -test:  $z$  corrected for ties =  $-22.8$ ,  $p = .000$ ) and “occupation/work” (Mann-Whitney  $U$ -test:  $z$  corrected for ties =  $-12.3$ ,  $p = .000$ ). But there are also marked differences in “leisure time/hobbies,” “housing/living conditions” and “health,” in each case with higher levels in western Germany. These

differences are not (chiefly) a consequence of “reality,” but rather of the very large discrepancy in Eastern Germany between expectations and (perceived) reality (Henrich, Herschbach, & von Rad, 1992).

#### Sensitivity to Change

The FLZ<sup>M</sup> modules have properties that are conducive to sensitivity to change (number of items, number of possible responses per item, wide range of the scale for wS and total score, item-by-item evaluation). The FLZ<sup>M</sup> has been used in a number of therapy studies and has proven to be a helpful tool for following progress (Herschbach,

Table 4. Score distribution and item characteristics of the Spanish FLZ<sup>M</sup> modules (General LS and Health, *N* = 896).

	M.d.	Mean	SD	Floor	Ceiling	IC
FLZ <sup>M</sup> General LS						
Friends/acquaintances	1.9%	7.4	6.3	0.1%	10.1%	0.31
Leisure time/hobbies	2.6%	5.2	6.2	0.6%	4.9%	0.42
Health	2.6%	9.5	8.3	3.2%	22.9%	0.41
Income/financial security	2.7%	3.3	7.2	5.6%	4.5%	0.52
Occupation/work	4.0%	5.1	7.7	3.7%	8.8%	0.47
Housing/living conditions	3.1%	7.9	6.6	0.8%	11.8%	0.49
Family life/children	3.5%	14.7	6.9	0.7%	55.1%	0.41
Partner relationship/sexuality	4.4%	10.5	8.3	1.8%	30.2%	0.42
Total score	2.9%	64.0	33.9	0.0%	0.3%	0.74
FLZ <sup>M</sup> Health						
Physical condition/fitness	2.0%	4.1	6.3	1.4%	4.8%	0.56
Ability to relax/stay on an even keel	3.0%	4.8	6.7	2.3%	6.1%	0.56
Energy/zest for life	2.1%	9.1	7.2	0.7%	19.5%	0.59
Mobility (e. g., walking, driving)	2.7%	9.2	7.3	1.2%	19.6%	0.56
Vision and hearing	2.8%	10.1	8.0	2.1%	26.9%	0.51
Freedom from anxiety	2.6%	7.2	7.0	1.4%	12.1%	0.50
Freedom from aches and pains	1.9%	8.3	8.7	5.1%	21.2%	0.60
Independence from help/care	2.6%	13.0	7.7	1.0%	44.7%	0.54
Total score	2.5%	65.8	40.0	0.0%	1.4%	0.83

M.d. = Missing data, IC = Internal consistency (items: part-whole correlation; total score: Cronbach's  $\alpha$ )

Henrich, & Oberst, 1994). Figure 2 demonstrates the responsiveness of the FLZ<sup>M</sup>-total scores during growth-hormone replacement therapy in patients with growth-hormone deficiency (unpublished observations). The *Z*-values (of the Wilcoxon matched-pairs signed-rank test) for the difference between months 0 and 6 as an indicator of the responsiveness of the FLZ<sup>M</sup>-total scores ("General LS": 4.78; "Health": 4.05) are higher than seven of the eight SF-36 scores obtained in the same sample (range: 0.56 to 5.33) and comparable to those on the physical and mental component summary (2.07 and 4.06, respectively). Several other studies are in progress; a psychometric evaluation of the data has yet to be done.

## Normative Data

German norms are available for the FLZ<sup>M</sup> module "General LS" from two different points in time (1991 and 1994) and for the module "Health" for one (1995). The total scores in the representative samples are normally distributed. The range of possible values is from -96 to 160, with negative values indicating a predominance of "dissatisfaction." Table 3 includes normative data for the samples from 1994 and 1995 by age and gender. The FLZ<sup>M</sup> manual contains the values for the individual items. Normative data for a representative sample of the Spanish population were collected in October and November 1998 (see Table 4).

## Comparative Data

For the module "General LS," in addition to the data for the representative samples of the general population there are comparative data for 29 groups of patients (total *N* = 7654) and 12 groups of healthy subjects (total *N* = 5425). For the module "Health" there are data for 22 groups of patients (total *N* = 6970) and seven groups of healthy subjects (total *N* = 1537).

Figure 3 shows the mean total score on the FLZ<sup>M</sup>-"General LS" for the normative samples from the Eastern and Western parts of Germany and for a number of different samples. The samples consisted of patients with psychiatric or psychosomatic disorders, with acute or chronic physical disorders, and with various types of problems related to their psychosocial situation (women looking for employment, nurses working with patients with cancer or Parkinson's disease and relatives of such patients). For this figure samples from different studies and different diagnostic groups were combined. The figure shows that the patients with psychiatric problems had the poorest QL. For those with chronic physical illness, the values vary widely, but overall they were higher than what one might expect. In patients whose cancer was in remission QL was very high, sometimes even higher than in the general population. In contrast, those groups of people with problems in the psychosocial area had relatively low values.

A global conclusion that can be drawn from Figure 3 is that the "objective severity of illness" does not corre-

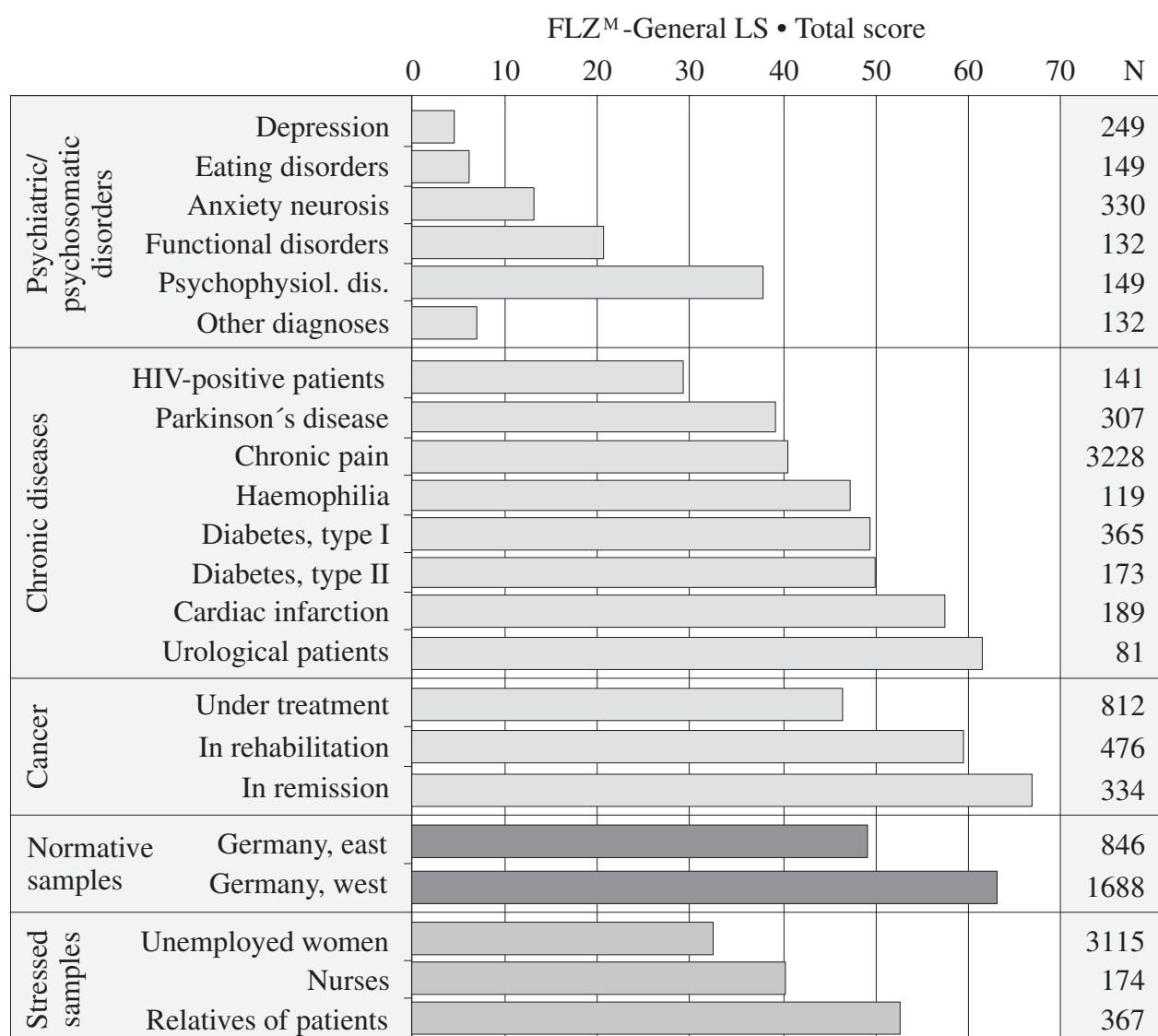


Figure 3. FLZ<sup>M</sup>-General LS; mean total score in different samples of ill and healthy persons.

late very highly with QL. This is consistent with the concept of QL outlined in the Introduction.

### International Versions of the FLZ<sup>M</sup>

Both modules of the FLZ<sup>M</sup>, initially developed in German, will soon also be available in UK English, US English, Dutch, Spanish, and Italian. For each language, development of the final version involved translation and back-translation by two independent bilingual translators and reviews by the test authors, trials with the translated questionnaire and discussions about problems in filling it out (Bullinger, Anderson, Cella, & Aaronson, 1993). The cross-cultural psychometric evaluation is in progress.

The first available data from a country other than Ger-

many are from a representative sample in Spain ( $N = 896$ ; age: mean = 42.4 years,  $SD = 16.5$ ; gender: 48.4% male; marital status: 63.6% married). They were collected in 1998 as a sample of all Spanish citizens between 20 and 70 years old. Table 4 gives item characteristics and total scores for the two modules "General LS" (upper part) and "Health" (lower part), equivalent to Table 1 for Germany.

The item and test characteristics are largely comparable to the German results. There are some differences in means, which can be explained plausibly by differences in sociocultural values and in mentality (e. g., in "family life/children"). The two outliers in the ceiling effects are analogous to the differences in means: In "family life/children" 55.1% and in "partner relationship/sexuality" 30.2% of the Spanish sample have the maximum value (i. e., the combination of "extremely important"



and “very satisfied”). With two exceptions, the part-whole correlations between items and scale are between 0.40 and 0.50.

For the module “Health,” the highest value for the ceiling effect is 44.7%, for the item “independence from help/care” (compared to 31.2% in Germany). Again, the part-whole correlations for the Spanish sample are lower than those for the German sample. The somewhat lower homogeneity of the Spanish scales is reflected in Cronbach’s  $\alpha$ , which is 0.74 for the “General LS” module and 0.83 for the “Health” module. These are acceptable values for generic and unspecific health-related QL modules.

A definitive evaluation of the suitability of the FLZ<sup>M</sup> for use in other countries and for cross-cultural QL assessment will be possible after analysis of these data and the data being collected in the other countries has been completed.

## Discussion

We set out to develop the FLZ<sup>M</sup> because we were convinced that a subjective construct of QL consisting of several dimensions needs to include weighting for the relative importance of each dimension for the individual concerned. Clearly, the rating on a dimension that is of little importance to an individual should not contribute the same amount to the overall QL score as the rating on a dimension that is especially important.

Over a period of more than 10 years this concept was operationalized, modified, and subjected to psychometric evaluation. The resulting FLZ<sup>M</sup> is a standardized, economical questionnaire consisting of two modules “General Life Satisfaction” and “Satisfaction with Health,” with eight items each. The two modules are conceived of as measures of general QL and health-related QL, respectively. The instrument can be used with both individuals and groups and for both cross-sectional and longitudinal assessments. It is now, or soon will be, available in German, UK English, US English, Dutch, Spanish, and Italian. Other modules related to specific illnesses are under development.

Over the past several years the FLZ<sup>M</sup> modules have been made available to many colleagues for use in their own studies, with the proviso that the test authors be allowed to use the raw data for the psychometric evaluation of the instrument and as comparative data. From these studies and our own we now have data from more than 7500 patients from various diagnostic groups and from more than 5000 healthy individuals. Additionally, norms for a representative sample of the general (German and Spanish) population have been established.

The psychometric evaluation of the FLZ<sup>M</sup> demonstrated a high level of internal consistency. With regard to construct validity, the adequate correlations with established measures of QL could be demonstrated. The module “General LS” correlates best with scales that address psychological well-being. The module “Health” correlates better than the module “General LS” with scales addressing physical problems, and also with those addressing their psychological components such as depression and anxiety. This is consistent with the findings of several other authors (Watson & Clark, 1984; Salovey & Birnbaum, 1989; Watson & Pennebaker, 1989) that mood, depression, or negative affectivity have a major influence on the perception of symptoms and physical well-being.

Overall, the scientific studies and clinical experience with the FLZ<sup>M</sup> have shown that weighting of the individual items for their importance to the respondent is an effective way to incorporate the concept of the subjectivity of QL into the QL instrument, and it is well-accepted by the respondents.

## Acknowledgment

We are grateful to all of our colleagues who shared their data with us.

## References

- Aaronson, N.K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N.J., Filiberti, A., Flechtner, H., Fleishman, S.B., de Haes, J.C.J.M., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P.B., Schraub, S., Sneeuw, K., Sullivan, M., & Takeda, F. for the EORTC Study Group on Quality of Life (1993). The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, 85(5), 365–376.
- Beck, A.T., Ward, C.H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571.
- Bowling, A. (1991). *Measuring health – A review of quality of life measurement scales*. Philadelphia: Open University Press.
- Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, 36(8), 917–927.
- Bullinger, M., Anderson, R., Cella, D., & Aaronson, N. (1993). Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Quality of Life Research*, 2, 451–459.
- Calman, K.C. (1984). Quality of life in cancer patients – an hypothesis. *Journal of Medical Ethics*, 10, 124–127.
- Cassileth, B.R., Lusk, E.J., & Tenaglia, A.N.A. (1982). Psycho-

- logical comparison of patients with malignant melanomas and other dermatological diseases. *American Academy of Dermatology*, 7, 742–746.
- Cella, D. (1998). Quality of life. In J.C. Holland (Ed.), *Psychooncology* (pp. 1135–1146). New York, Oxford: Oxford University Press.
- Cella, D.F., Tulsky, D.S., Gray, G., Sarafian, B., Linn, E., Bonomi, A., Silberman, M., Yellen, S.B., Winicour, P., Brannon, J., et al. (1993). The Functional Assessment of Cancer Therapy scale: Development and validation of the general measure. *Journal of Clinical Oncology*, 11(3), 570–579.
- Derogatis, L.R. (1977). *SCL-90-R, administration, scoring, and procedures manual for the R(evised) version*. Baltimore: John Hopkins University School of Medicine.
- DuPuy, H.J. (1984). PGWB – The psychological well-being index. In N.K. Wenger, M.E. Mattson, C.D. Furberg, & J. Ellison (Eds.), *Assessment of quality of life in clinical trials of cardiovascular therapies* (pp. 170–183). New York: LeJacq Publishers.
- Ferrans, C.E., & Powers, M.J. (1985). Quality of life index: Development and psychometric properties. *Advances in Nursing Science*, 8(1), 15–24.
- Gill, T.M., & Feinstein, A.R. (1994). A critical appraisal of the quality of quality-of-life measurements. *Journal of the American Medical Association*, 272(8), 619–626.
- Glatzer, W., & Zapf, W. (Eds.). (1984). *Lebensqualität in der Bundesrepublik: Objektive Lebensbedingungen und subjektives Wohlbefinden* [Quality of life in the Federal Republic of Germany: Objective life situation and subjective well-being]. Frankfurt: Campus.
- Henrich, G., & Herschbach, P. (1995). Fragen zur Lebenszufriedenheit (FLZ) – ein Gewichtungsmo- dell [Questions on Life Satisfaction (FLZ) – A weighting model]. In R. Schwarz, J. Bernhard, H. Flechtner, T. Küchler, & C. Hürny (Eds.), *Lebensqualität in der Onkologie II* [Quality of life in oncology II] (pp. 77–93). München: W. Zuckschwerdt Verlag.
- Henrich, G., Herschbach, P., & von Rad, M. (1992). “Lebensqualität” in den alten und neuen Bundesländern [“Quality of life” in the original and new states of the Federal Republic of Germany]. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 42(1), 31–32.
- Herschbach, P., Henrich, G., & Oberst, U. (1994). Lebensqualität in der Nachsorge. Eine Evaluationsstudie in der Fachklinik für Onkologie und Lymphologie, Bad Wildungen-Reinhardshausen [Quality of life in aftercare: An evaluation study in the Treatment Center for Oncology and Lymphology, Bad Wildungen-Reinhardshausen]. *Praxis der Klinischen Verhaltensmedizin und Rehabilitation*, 28, 241–251.
- Hickey, A.M., Bury, G., O’Boyle, C.A., Bradley, F., O’Kelly, F.D., & Shannon, W. (1996). A new short form individual quality of life measure (SEIQoL-DW): Application in a cohort of individuals with HIV/AIDS. *British Medical Journal*, 313, 29–33.
- Huber, D., Henrich, G., & Herschbach, P. (1988). Measuring the quality of life: A comparison between physically and mentally chronically ill patients and healthy persons. *Pharmacopsychi- atry*, 21, 453–455.
- Hunt, S.M., McEwen, J., & McKenna, S.P. (1986). *Measuring health status*. London: Croom Helm.
- Ipsen, D. (1978). Das Konstrukt Zufriedenheit [The construct satisfaction]. *Soziale Welt*, 29(1), 44–53.
- Karnofsky, D.A., Abelmann, W.H., Craver, L.F., & Burchenal, J. H. (1948). The use of the nitrogen mustards in the palliative treatment of carcinoma. *Cancer*, 1, 634–656.
- Krischke, N.R. (1996). *Lebensqualität und Krebs* [Quality of life and cancer] (Vol. 2). München: Quintessenz MMV Medizin Verlag.
- McDowell, I., & Newell, C. (1987). *Measuring health: A guide to rating scales and questionnaires*. New York: Oxford University Press.
- McGee, H.M., O’Boyle, C.A., Hickey, A., O’Malley, K., & Joyce, C.R.B. (1991). Assessing the quality of life of the individual: The SEIQOL with a healthy and a gastroenterology unit population. *Psychological Medicine*, 21(3), 749–759.
- Roder, J.D., Herschbach, P., Sellschopp, A., & Siewert, J.R. (1991). Quality-of-life assessment following oesophagectomy. *Theoretical Surgery*, 6, 206–210.
- Salovey, P., & Birnbaum, D. (1989). Influence of mood on health-relevant cognitions. *Journal of Personality and Social Psychology*, 57, 539–551.
- Spielberger, C.D., Gorsuch, R.L., & Lushene, R.E. (1970). *STAI, Manual for the State-Trait-Anxiety-Inventory*. Palo Alto: Consulting Psychologist Press.
- Veenhoven, R. (1991). Ist Glück relativ? Überlegungen zu Glück, Stimmung und Zufriedenheit aus psychologischer Sicht [Thoughts on happiness, mood, and satisfaction from a psychological perspective]. *Report Psychologie*, 7, 14–20.
- Ware, J.E., Jr., Kosinski, M., Bayliss, M.S., McHorney, C.A., Rogers, W.H., & Raczek, A. (1995). Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: Summary of results from the Medical Outcomes Study. *Medical Care*, 33(4 Suppl.), AS264–279.
- Watson, D., & Clark, L.A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96, 465–490.
- Watson, D., & Pennebaker, J. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review*, 96, 234–254.
- WHO. (1994). *Quality of life assessment. An annotated bibliography* (Vol. WHO/MNH/PSF/94.1). Geneva: WHO. Division of Mental Health.

---

Dr. Gerhard Henrich  
 Institut und Poliklinik für Psychosomatische Medizin,  
 Psychotherapie und Medizinische Psychologie der  
 Technischen Universität München  
 Langerstr. 3  
 D-81675 Munich  
 Germany  
 Tel. +49 89 4140 4317  
 Fax +49 89 4140 4845  
 E-mail G.Henrich@lrz.tu-muenchen.de

---