

Social desirability in the measuring of patient satisfaction after treatment of coloproctologic disorders

On shortcomings of general bipolar satisfaction scales for quality management

Gerald D. Giebel · Norbert Groeben

Received: 9 May 2007 / Accepted: 31 January 2008 / Published online: 11 March 2008
© Springer-Verlag 2008

Abstract

Background and aims Within the field of medicine, much attention is being paid to quality management, whereby patient satisfaction plays a major role. In order to measure this construct, usually rather general, bipolar rating scales are applied. However, these scales are often susceptible to social desirability biases.

Patients–methods Coloproctological patients were asked to complete a questionnaire with ratings of satisfaction and anxiety at two different points of measurement: One group while in treatment ($N=86$) and a second group approximately 1 year after their treatment ($N=328$).

Results Even when controlling for relevant demographic influences, a clear decrease in intensity of the evaluation ratings is shown when the survey was administered 1 year after release from hospital as compared to during the patients' stay in hospital. For generally formulated scales of patient satisfaction, social desirability constitutes a significant bias.

Conclusion The usual conceptualization of a bipolar continuum of patient satisfaction or dissatisfaction must be renounced. Instead, questionnaires might be constructed in three steps which investigate problem dimensions at a medium level of concreteness. Only in this way can quality management gain credibility and trust within as well as beyond the field of medicine.

Keywords Patient satisfaction · Social desirability · Surgery · Coloproctology

Introduction

Within the field of medicine, quality management is not only en vogue but also a laudable undertaking: It implies a self-critical attitude as a condition for the improvement of medical care. Within quality management, investigating patient satisfaction is of major importance. Here, a bipolar continuum between satisfaction and dissatisfaction is usually assumed, such that dissatisfaction is equivalent to low scores of satisfaction.

It has been repeatedly commented that this conceptualization is inadequate and that patient (dis)satisfaction is more than simply the discrepancy between expectations and experience [1–3]. Instead, patients' ratings are usually also influenced by attribution processes. Therefore, only those negative experiences attributable to derelictions of duty by the medical staff lead to negative satisfaction ratings [4]. Negative experiences, which, for example, can be explained by the fact that the staff is overworked etc., do not inevitably reduce indications of satisfaction [5]. This attributional reformulation of the concept of (dis)satisfaction provides one possible explanation of a much replicated finding: Common measurements reveal extraordinarily high satisfaction scores [6–8]—possibly because negative experiences are attributed to the difficult circumstances with which the staff are faced and excused. As valuable as theoretical explications of the concept of satisfaction might be [9], it is at the same time both useful and necessary to consider methodological problems of the measurement procedures. Aside from potentially influential demographic variables (age, sex, education, marital status of the patient) [7, 10], the majority

G. D. Giebel
Department of Coloproctology, Hospital St. Josef,
Saarbrücken, Germany

N. Groeben (✉)
Department of Psychology, University of Cologne,
Herbert-Lewin-Str. 2,
50931 Cologne, Germany
e-mail: n.groeben@uni-koeln.de

of general problems relevant to the construction and administration of questionnaires also apply. For example, negatively formulated questions might be ignored [11]; orally administered questionnaires lead to artificially higher scores of satisfaction than written ones [12] and the same holds true for surveys which were carried out during the period of treatment as compared to those conducted after termination of treatment [13]. In sum, relevant biases must be considered for the collection as well as interpretation of the data [14]. Among these biases, social desirability is likely to be of major importance [15]. The normal tendency to adapt one's attitudes and judgments to social norms is plausibly even stronger in patients. This can be explained by the fact that they will usually try to be "good patients" and adjust to the medical staff, from which they hope to receive help [16]. This tendency is probably particularly evident during hospitalization, which usually implies a severe illness. Thus, we are in this case not dealing with social desirability as a trait but rather as a state variable. Here, we are concerned with the pressure to conform, which results from the patient's subjective experience of psycho-physical dependency upon the medical staff and can lead to unrealistically high satisfaction ratings [17, 18].

Standard measures of social desirability as a trait variable are not useful for the assessment of the influence of social desirability as a state variable because the results will be affected by the situation in addition to the individual disposition. The assessment during hospitalization is not an option either, because, once more, the answers would be influenced by social desirability. For the same reason, obtaining data across various time measurements from a single sample of patients does not seem optimal because the purpose of the study would be too obvious. The least reactive procedure is the comparison of two different samples at different times.

The comparison of satisfaction ratings in a sample of patients during hospitalization with those in a different sample of patients after their release is thus an indirect but relatively safe method of examining the potential bias which social desirability might create. The unusually high pressure of social desirability during hospitalization should be reduced to a normal rate when patients give judgments based on their memories. It is however also possible that a positive memory bias makes the experiences in hospital appear more pleasant than they were actually perceived within the situation. This tendency to gloss over negative experiences should primarily result in negative emotions being forgotten [19–21]. Such memory distortions would lead to retrospectively higher rates of satisfaction in comparison with ratings given during the stay in hospital. Their effect would therefore be opposite to the hypothesized effect of social desirability. If, nevertheless, the comparison of satisfaction ratings during hospitalization with those of 1 year after release reveals more negative

evaluations for those provided retrospectively, this would strongly suggest that, in the situation of hospitalization, social desirability leads to positively biased appraisals. Because other potential influences work against this hypothesis, this finding would provide a valid argument for the social desirability hypothesis. It is thus possible to examine the effects of social desirability on patients' satisfaction ratings by obtaining these ratings at two different points in time: during and 1 year after treatment in hospital. Evaluations of anxiety levels, however, should not be affected by social desirability caused by the situation: We predict that there will be no difference between anxiety ratings provided during and after hospitalization.

Materials and methods

Between June and September 2003, patients who had undergone a coloproctological operation were asked to complete a survey including ratings of satisfaction and evaluations of anxiety on the last day of their stay in hospital ($N=86$). Data were collected anonymously in order to prevent an artificial increase of social desirability measurements. At the same time, patients who had undergone an operation in the same department in 2002 received the same survey by mail and were asked to complete it, based on their memories of the stay in hospital approximately 1 year earlier ($N=728$; out of these, 48 had died or moved away; rate of returns 48.2%=328). The diagnoses of the patients are in a summarized way listed in ICD 10, so one can see the part of minor and major procedures (see Table 1).

Two scales of satisfaction from a measure developed by Hildebrandt et al. [8] were included in this questionnaire: a scale measuring satisfaction with the medical staff (scale D.1., 11 items), and a scale measuring satisfaction with the nursing staff (scale D.2, six items). These are very carefully developed scales which are frequently used in Germany. They focus upon those aspects of satisfaction for which situational pressure towards social desirability is assumed due to the dependency on medical staff as described above.

Table 1 Diagnoses summarized due to the ICD 10

	Number of patients	Diagnoses
C 18–C 21	39	Carcinoma
D 1–D 37	15	Polyp, intraepithelial carcinoma
I 84	141	Hemorrhoids
K 35–K 59	45	Colitis
K 60–K 61	69	Fissure-in-ano, anal fistula
K 62–K 92	43	Prolapse, peritonitis
L 2–L 73	12	Pilonidal cyst
N 81	45	Rectovaginal fistula, rectocele
R 10	5	Acute abdomen

Answers were obtained on a five-point rating scale (from “very good” to “mediocre” and “very bad”). Self-assessments of anxiety were provided in five comparable answer categories. The items of this measure had been developed specifically for the sample of proctologic patients and belonged to one of the following five theoretically postulated scales (for samples of items see “Appendix”): causes of anxiety, reasons for anxiety reduction, peak of anxiety, postoperative worries, and possibilities of reducing anxiety. In addition, there were three single anxiety items regarding: anxiety during admission, appropriateness of anxiety, and general level of anxiety. Other questionnaires that were used in the study, e.g., the Fragebogen zur Erhebung von Kontrollüberzeugungen zu Krankheit und Gesundheit (Questionnaire for the assessment of control beliefs regarding illness and health) [22], are not relevant to the methodological questions dealt with here.

Step one consisted of a factor analysis to test whether the two factors of satisfaction (with medical and with nursing staff) could be replicated for our sample. The comparability of the two samples (during versus after hospitalization) regarding demographic variables (age, sex, education, type of illness, etc.) was tested using *t*-tests for independent samples.

A multivariate analysis of variance was computed in order to test the central methodological question: whether satisfaction ratings and anxiety ratings are reduced in the second observation period. Due to the fact that the sample of patients who completed the survey while hospitalized was much smaller than the sample of patients who completed it after their release from hospital, some of the demographic characteristics differed between the two subsamples. Therefore, in the following analyses, the variables which revealed significant differences between the two groups were controlled statistically with regard to their influence on the dependent variable (satisfaction ratings and anxiety ratings). For this purpose, an analysis of covariance was computed, in which the proportion of variance accounted for by the covariates was partitioned out. This ensures that the remaining variance can be explained by the influence of the independent variable. With a large sample size, even numerically rather minor differences can yield a statistically significant result. Therefore, in addition to the test of significance, a statistical power analysis (Eta square) was computed, in order to estimate whether the effects were substantial or not. The structure of the results was additionally tested, using a discriminatory analysis and calculating the correspondence of the observed group membership with a predictive model. This allows an examination of whether the differences (within the dependent variables, satisfaction ratings, and anxiety ratings) which were significant in the analysis of covariance also emerge as the main predictors of the observed group membership.

The calculations were computed using the statistical package SPSS 11.0.

Results

In order to create a measure of satisfaction, we used items drawn from the factors “satisfaction with medical staff” and “satisfaction with nursing staff” from the comprehensive measure of Hildebrandt et al. [8]. By means of a confirmatory factor analysis (principle component analysis with varimax rotation; $N=354$), we tested whether it is possible to replicate these factors for our sample. Two factors were extracted, which accounted for 51.76% of the variance (see Table 2).

As Table 2 reveals, the structure of factor loadings was successfully replicated. All items with the exception of one (No. 7) loaded onto the postulated factors, the factor loadings being between 0.401 and 0.821. For further calculations concerning the two measures of satisfaction, the factor values (for each subject) were entered, in which single items for each factor were weighted, according to their factor loading.

Following theoretical and conceptual considerations, the proposed dimensional differentiation of the anxiety scale was maintained, and a factor analysis was not computed for these items. The internal consistency, i.e., the reliability of the scales (Cronbach’s alpha: causes of anxiety $\alpha=0.827$; reasons for anxiety reduction $\alpha=0.796$; peak of anxiety $\alpha=0.593$ postoperative worries $\alpha=0.796$; possibilities of reducing anxiety $\alpha=0.835$), was sufficient for our purpose, which was a methodological comparison of the two groups. It was not necessary to reduce the number of items inductively. The ratings of all items on each scale were summated, and the mean was computed in order to obtain separate scale values.

Data had been collected from the sample of patients still in hospital treatment, however, only for the purpose of controlling for the factor of social desirability. Accordingly, this sample was ultimately much smaller than the posttreatment sample (86 hospitalized patients as compared to 328 released patients). Given such large differences between the sample sizes, it is clear that differences in demographical variables etc. might occur. Therefore, we tested the samples for their comparability regarding the following variables: sex, age, education, and type of illness. The variable “education” was measured by the level of school education (five levels). Concerning the variable “type of illness,” a distinction was made between (less serious) operations on the rectum and (more serious) colon operations. The differences between the two samples were tested with *t*-tests for independent samples, and significant differences regarding the variables age, sex, and type of illness were revealed (cf. Table 3).

Table 2 Rotated component matrix of satisfaction items

	Component	
	Satisfaction with physicians	Satisfaction with nursing care
N. 1: professional competence of physicians	0.497	0.425
N. 2: professional competence of nursing staff		0.677
N. 3: frequency and length of doctor's rounds	0.762	
N. 4: personal contact during doctor's rounds	0.801	
N. 5: information about health condition and therapies	0.821	
N. 6: thoroughness of medical examinations	0.779	
N. 7: avoidance of pain		0.396
N. 8: punctuality of medication	0.401	
N. 9: appropriateness of examinations and therapies	0.451	
N. 10: information about treatment and operations	0.737	
N. 11: participation in decisions	0.737	
N. 12: friendliness of nursing staff		0.721
N. 13: sufficient number of nursing staff		0.658
N. 14: promptness of nursing staff in cases of need		0.800
N. 15: anxiety reduction through physicians	0.661	0.358
N. 16: anxiety reduction through nursing staff		0.680
N. 17: cooperation between doctors and nursing staff	0.519	0.571

These differences between the two subsamples must be statistically controlled when evaluating the original differences of interest, namely differences in satisfaction ratings and anxiety ratings.

Table 4 shows these differences (satisfaction ratings and anxiety ratings during treatment in hospital and after release from hospital). A multivariate analysis of variance reveals significant effects for three factors, namely satisfaction with medical staff, satisfaction with nursing staff, and reasons for anxiety reduction. The effects occur in the expected direction: The intensity of the ratings is reduced for the group in which the data were collected after their release from hospital, as compared to the group in which the

survey was completed during the stay in hospital, when higher rates of satisfaction and anxiety reduction were indicated. For all other anxiety scales, there were no significant differences between the two groups.

At the beginning of this article, the methodological question was posed whether there is a danger of satisfaction ratings being biased due to the influence of social desirability. Before it is possible to draw a sufficiently reliable conclusion based on this data set, two problems must be taken into account. First of all, differences in satisfaction ratings and anxiety ratings between the two measurement times might (partially) be due to the variability among the two samples, as described above. This can be

Table 3 Descriptive values of the two groups and results of the *t*-tests

Variable	Group	<i>N</i>	Mean	Average deviation	<i>t</i>	<i>df</i>	<i>p</i>	Average difference
Type of illness ^a	After release from hospital	323	1.38	0.486	2.844	145.900	0.005 ^b	0.15
	During hospitalization	84	1.23	0.421				
Sex	After release from hospital	325	0.45	0.498	-2.399	407	0.017 ^c	-0.15
	During hospitalization	84	0.60	0.494				
Age	After release from hospital	324	58.55	14.075	2.626	407	0.009 ^b	4.57
	During hospitalization	85	53.98	15.093				
Education	After release from hospital	323	2.87	0.059	0.110	404	0.912	0.01
	During hospitalization	83	2.86	0.123				

Coding of sex: 0=female, 1=male; type of illness: 1=not severe (rectum), 2=severe (colon); age: interval scale; education: ordinal, 1=without degree, 2=Volks-, Hauptschulabschluss (lowest educational qualification in Germany, following 9 years of school), 3=Realschule, Mittlere Reife (medium-level educational qualification after 10 years of school), 4=(Fach-)Abitur (highest high-school qualification, qualifying for admission to a university and rewarded after 12 or 13 years of school), 5=(Fach-)Hochschulabschluss (university degree)

^a Welch's correction of the degrees of freedom because the Levene test detected unequal variances

^b Significant at a level of $\alpha=0.01$

^c Significant at a level of $\alpha=0.05$

tested by controlling for variables in which a significant difference exists—here: age, sex, and type of illness—in an analysis of covariance. However, it is possible that not all relevant differences between the two subsamples have been identified and controlled for. There may well be other variables which correlate with proneness to situational social desirability. In this respect, the lack of difference in the level of education provides relative safety because differences regarding proneness to social desirability are primarily due to education [23, 24]. Therefore, we may assume that the computed analysis of covariance controls for the relevant sample differences. Secondly, relatively small differences can become significant in large samples. Thus, the substantiality of the effects must be tested using a measure of effect size (qua explanation of variance); in this case, the partial Eta square was used [25]. Table 5 shows the results of the corresponding multivariate analysis of covariance, in which the three major differences between the two measurement times (satisfaction with medical staff and nursing staff as well as reasons for anxiety reduction) yielded significant results and a substantial effect size.

According to Cohen [26], the single effects—being between 0.018 (satisfaction with medical staff) and 0.046

(reasons for anxiety reduction) up to 0.076 (satisfaction with nursing staff)—must be classified as small effects. The overall value of 0.141, however, corresponds to a large effect.

Finally, we tested whether, complementarily, the variables of interest are also significant predictors of the two samples (during hospitalization vs. after release from hospital). The corresponding statistical test carried out by means of discriminatory analysis leads to results reported in Table 6.

The prediction was indeed significant ($z=5.194$; $p<0.001$). The correspondence between the cross-validated prediction model and the observed group membership is, with 78.9%, by all means satisfactory (Tables 6 and 7). The variables, confirmed by the analysis of variance, were confirmed to be the major predictors.

All in all, the results clearly reveal a reduction of intensity in the investigated satisfaction, even when controlling for relevant confounding demographic variables. The results of the model applying analysis of variance, as well as of the model using discriminatory analysis, are consistent, and both reveal a significant decrease between the first and the second measurement time. Thus, the

Table 4 Descriptive values of anxiety and satisfaction scores; results of multivariate analysis of variance ($N = 331$; Pillai's trace method) and the corresponding univariate analyses of variance

Variable	Group	<i>N</i>	Mean	Average deviation	SS	<i>df</i>	MS	F	<i>p</i>
Univariate models									
A. 1: anxiety during admission	After release from hospital	325	2.898	1.167	0.046	1	0.046	0.034	0.854
	During hospitalization	85	2.894	1.272					
A. 2: causes of anxiety	After release from hospital	327	2.803	0.856	0.246	1	0.246	0.354	0.552
	During hospitalization	86	2.872	0.900					
A. 3: reasons for anxiety reduction	After release from hospital	325	3.252	0.782	9.660	1	9.660	15.744	0.000 ^a
	During hospitalization	86	3.606	0.861					
A. 4: peak of anxiety	After release from hospital	326	2.509	0.930	1.713	1	1.713	1.994	0.159
	During hospitalization	86	2.667	0.943					
A. 5: post-operation worries	After release from hospital	307	3.087	0.890	0.092	1	0.092	0.122	0.728
	During hospitalization	85	3.189	0.880					
A. 6: possibilities for anxiety reduction	After release from hospital	320	2.399	0.849	0.626	1	0.626	0.906	0.342
	During hospitalization	83	2.531	0.874					
A. 7: appropriateness of anxiety	After release from hospital	319	3.413	0.795	1.126	1	1.126	1.865	0.173
	During hospitalization	86	3.476	0.890					
A. 8: general anxiety	After release from hospital	325	2.772	0.904	0.877	1	0.877	1.062	0.303
	During hospitalization	86	2.872	1.049					
Satisfaction with physicians	After release from hospital	277	0.062	1.038	4.105	1	4.105	4.115	0.043 ^b
	During hospitalization	77	-0.225	0.813					
Satisfaction with nursing care	After release from hospital	277	0.148	0.991	30.589	1	30.589	33.810	0.000 ^a
	During hospitalization	77	-0.533	0.840					
Multivariate model									
Stationary–poststationary ^c								5.497	0.000 ^a

^a Significant at a level of $\alpha=0.01$

^b Significant at a level of $\alpha=0.05$

^c Value = 0.147; hypothesis $df=10$; error $df=320$

Table 5 Multivariate analysis of covariance (N=323; Pillai's trace method) and corresponding univariate results

	SS	df	MS	F	p	Eta square
Multivariate model						
During–after hospitalization ^a				5.084	0.000 ^b	0.141
Univariate models						
Anxiety 1	0.098	1	0.098	0.077	0.782	0.000
Anxiety 2	0.043	1	0.043	0.065	0.798	0.000
Anxiety 3	9.146	1	9.146	15.285	0.000 ^b	0.046
Anxiety 4	2.962	1	2.962	3.505	0.062	0.011
Anxiety 5	0.003	1	0.003	0.005	0.942	0.000
Anxiety 6	0.941	1	0.941	1.440	0.231	0.005
Anxiety 7	1.187	1	1.187	2.000	0.158	0.006
Anxiety 8	1.609	1	1.609	2.002	0.158	0.006
Satisfaction with physicians	4.802	1	4.802	4.907	0.027 ^c	0.015
Satisfaction with nursing care	23.286	1	23.286	26.096	0.000 ^b	0.076

^a Value=0.141; hypothesis $df=10$; error $df=309$

^b Significant at a level of $\alpha=0.01$

^c Significant at a level of $\alpha=0.05$

pattern of results can be considered both stable and substantial.

Discussion

The two scales of satisfaction (regarding the medical and the nursing staff) reveal the expected decrease of rating scores after release from hospital, compared to the results of measurement during hospitalization. This reduction indicates a bias caused by effects of social desirability because potential memory distortions would have had the reverse

effect. The possibility that the disease reoccurred or worsened in the period after release from hospital is highly unlikely for the case of coloproctological operations, even given cancerous diseases. Therefore, it is implausible that events subsequent to release are responsible for the decreased satisfaction scores. Furthermore, such complications concerning the illness would be attributed to the physicians and the satisfaction ratings would thus decrease even more for the medical staff than for the nursing staff. However, the contrary is the case (cf. Tables 4 and 5) so that the present satisfaction ratings can almost certainly be traced back to social desirability bias.

Anxiety ratings, however, were not expected to be influenced by social desirability to different extents at the two times of observation. This proved true for all anxiety scales, with the exception of the scale “reasons for anxiety reduction.” This result does not imply that the social desirability bias is less influential. A precise examination of the items shows that the scale is, after all, a scale measuring anxiety reduction. Asking about reduction is, de facto, inquiring about positive emotionality, for which the social desirability bias should also be of relevance. Therefore, due to

Table 6 Discriminatory analysis for separating the groups during and after hospitalization (N=331)

	Standardized canonical discriminatory coefficient	Structure matrix (rank by size)
A. 1	0.221	-0.025 (10)
A. 2	0.369	0.079 (8)
A. 3	-0.276	-0.528 (2)
A. 4	-0.556	-0.188 (4)
A. 5	0.157	-0.046 (9)
A. 6	-0.223	-0.127 (7)
A. 7	-0.088	-0.182 (5)
A. 8	-0.136	-0.137 (6)
Satisfaction with physicians	0.250	0.270 (3)
Satisfaction with nursing care	0.779	0.773 (1)

Standardized canonical discriminatory coefficients and structure matrix

Test of significance for discriminant function with Wilk's statistics: Wilk's lambda=0.853; chi-square=51.360; $df=10$; $p=0.000$ (significant at a level of $\alpha=0.001$)

Table 7 Cross-validated table of group membership (from probability of sample size estimated a priori)

	Predicted group membership		
	During hospitalization	After release from hospital	Total
During hospitalization	12	62	74
After release from hospital	8	249	257

the reverse coding, which implicitly asks about positive emotionality, the results of the scale measuring “reasons for anxiety reduction” in fact provide additional evidence of a social desirability bias. Because of the choice of experimental design, the estimation of the extent of this bias is rather conservative as its dynamics cannot be utterly neglected at the second observation time (after release from hospital). The pattern of results was obtained in a specialized clinic for coloproctology. This clinic does not however, to our knowledge, significantly differ from other surgical hospitals. Therefore, external validity of the results can be assumed, at least for stays in surgical hospitals. Thus, it can be stated that social desirability leads to a stable and substantial bias for ratings of satisfaction and emotion obtained with general scales.

Conclusion

The methodological consequences of this conclusion are considerably far reaching. First of all, the conceptualization of a bipolar continuum of patient satisfaction and dissatisfaction, which has already been criticized from a theoretical point of view [27], should be renounced. If the self-critical approach towards an improvement of medical care as described at the start of the present study is taken seriously, then putting this into practice cannot simply entail a theoretical elaboration of the concept of patient *dissatisfaction* but rather must also include methodological considerations.

And this should be achieved—at least within a first step—by not only using general rating scales because these are particularly prone to bias factors such as social desirability. Rather, a multistage approach should be favored, just as it is used for other attitude and evaluation questionnaires: By first of all using qualitative measurement steps and starting with individual cases, relevant aspects of judgment, problem areas, etc. can be determined, as was done at least rudimentarily through the application of group discussions [28–30].

In the next step, these can be translated into quantitative research instruments for general survey studies. One method which appears to be particularly useful for the qualitative and heuristic research phase is the critical incident technique [31]. In this method, participants are asked, in the course of a narrative interview, to describe particularly impressive (in this case, evoking dissatisfaction) situations [32, 33]. This method has been applied within the field of medicine, in order to develop and carry out training for nursing staff [34, 35]. So far, the potential of this technique for research on patient dissatisfaction has not been fully recognized [36]. As soon as a multitude of situations evoking dissatisfaction have been systematically assessed in this manner, it will be possible to group them

into typical categories, by means of a systematic content analysis [37, 38]. In a third step, these categories could be used to develop questionnaires, which could then be administered in representative survey studies. Questionnaires developed according to these three steps will no longer target unspecific and general evaluation ratings. Instead, they will inquire about types of situations at a medium level of abstraction and concreteness. For this type of item, the influence of common bias factors, such as social desirability, should be less severe.

It is our opinion that only when researchers of quality management within the field of medicine can be convinced to apply such an ambitious procedure, combining qualitative and quantitative research methods, will this domain also be able to optimize its potential for change and, furthermore, gain credibility and trust beyond the field of medicine.

Appendix: sample of items

1. Causes of anxiety: My anxiety was caused primarily by the fact that
 - (a) I did not know what to expect
 - (b) I was worried about the severity of my illness...
2. Reasons for reduction of anxiety: My anxiety was reduced by
 - (a) the friendly admission
 - (b) conversation with the nurses...
3. Peak of anxiety: My anxiety was at its highest
 - (a) during the admission
 - (b) after the briefing before the operation...
4. Postoperative worries: After the operation, my greatest concern was
 - (a) the removal of swabs and drainages
 - (b) whether I would receive a sufficient amount of pain killers...
5. Possibilities of reducing anxiety: It would have considerably helped me to reduce my anxiety if
 - (a) I had received a brochure with information, which I could have read in peace
 - (b) the conversations with the physicians had been more detailed...

References

1. Coyle J, Williams B (1999) Seeing the wood for the trees: defining the forgotten concept of patient dissatisfaction in the light of patient satisfaction research. *Leadersh Health Serv* 12:i–ix

2. Dozier AM, Harriet JK, Ingersoll GL, Homberg S, Schultz AW (2001) Development of an instrument to measure patient perception of the quality of nursing care. *Res Nurs Health* 24:506–517
3. Sitzia J, Wood N (1997) Patient satisfaction: a review of issues and concepts. *Soc Sci Med* 45:1829–1843
4. William B, Coyle J, Healy D (1998) The meaning of patient satisfaction: an explanation of high reported levels. *Soc Sci Med* 47:1351–1359
5. Dubé L, Bélanger M-C, Trudeau E (1996) The role of emotions in health care satisfaction. *J Health Care Mark* 16:45–51
6. Hall J, Dorman M (1988) Meta-analysis of satisfaction with medical care: description of research domain and analysis of overall satisfaction levels. *Soc Sci Med* 27:637–644
7. Hall J, Dorman M (1990) Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis. *Soc Sci Med* 30:811–818
8. Hildebrandt H, Nickel S, Trojan A (1999) Planung, Methodik und Durchführung der DAK-Patientenbefragung. (Survey of DAK patients: planning, method, and procedure). In: Schupeta E, Hildebrandt H (eds) *Patientenzufriedenheit messen und steigern (Measuring and increasing patient satisfaction)*. Asgard, Sankt Augustin, pp 56–81
9. Calnan M (1988) Towards a conceptual framework of lay evaluation of health care. *Soc Sci Med* 27:927–933
10. Tucker JL (2002) The moderators of patient satisfaction. *J Manag Med* 16:48–66
11. Stump T, Dexter P, Tierney W, Wolinsky FD (1995) Measuring patient satisfaction with physicians among older and diseased adults in a primary care municipal outpatient setting: an examination of three instruments. *Med Care* 33:958–972
12. Health Services Research Group (1992) A guide to direct measures of patient satisfaction in clinical practice. *Can Med Assoc J* 46:1727–1731
13. Harpole L, Orav J, Hickey M, Posther K, Brennan TA (1995) Patient satisfaction in the ambulatory setting: influence of data collection methods and sociodemographic factors. *J Gen Intern Med* 11:431–434
14. Fitzpatrick R (1991) Survey of patient satisfaction: I—important general considerations. *BMJ* 302:887–889
15. Paulhus DL (2002) Socially desirable responding: the evolution of a construct. In: Braun HL, Jackson DN (eds) *The role of constructs in psychological and educational measurement*. Erlbaum, Mahwah, pp 49–69
16. Edwards C, Staniszewska S, Crichton N (2004) Investigation of the ways in which patients' reports of their satisfaction with healthcare are constructed. *Sociol Health Illn* 26:159–183
17. Currie V, Harvey G, West E, McKenna H, Keeney S (2005) Relationship between quality of care, staffing levels, skill mix and nurse autonomy: literature review. *J Adv Nurs* 51:73–82
18. Edwards C, Staniszewska S (2000) Accessing the user's perspective. *Health Soc Care Commun* 8:417–424
19. Ross M, Buehler R (1994) Creative remembering. In: Neisser U, Fivush R (eds) *The remembering self: construction and accuracy in the self-narrative*. Cambridge University Press, New York, pp 205–235
20. Rubin DC (1996) *Remembering our past*. Cambridge University Press, New York
21. Schacter DL (1995) *Memory distortions: how minds, brains, and societies reconstruct the past*. Harvard University Press, Cambridge
22. Lohaus A, Schmitt GM (1989) Fragebogen zur Erhebung von Kontrollüberzeugungen zu Krankheit und Gesundheit (KKG). Handanweisung (Questionnaire for the assessment of control beliefs regarding illness and health). Hogrefe, Göttingen
23. Fisher RJ, Katz JE (2000) Social-desirability bias and the validity of self-reported values. *Psychol Market* 17:105–120
24. Schmidt C, Möller J, Reibe F, Güntert B, Kremer B (2003) Patientenzufriedenheit in der stationären Versorgung: Stellenwert, Methoden und Besonderheiten (Patient's satisfaction with their care in hospital: its ranking, methods and special features). *Dtsch Med Wochenschr* 128:619–624
25. Kerlinger FN (1964) *Foundations of behavioral research*. Holt, Rinehart and Winston, New York
26. Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Erlbaum, Hillsdale
27. Turris SA (2005) Unpacking the concept of patient satisfaction: a feminist analysis. *J Adv Nurs* 50:293–298
28. Hyrkäs K, Paunonen M (2000) Patient satisfaction and research-related problems (part 2). Is triangulation the answer? *J Nurs Manag* 8:237–245
29. Schneider H, Palmer N (2002) Getting to the truth? Researching user views of primary health care. *Health Policy Plan* 17:32–41
30. Staniszewska S, Henderson L (2005) Patients' evaluations of the quality of care: influencing factors and the importance of engagement. *J Adv Nurs* 49:530–537
31. Flanagan JC (1954) The critical incident technique. *Psychol Bull* 51:327–358
32. Byrne M (2001) Critical incident technique as a qualitative research method. *AORN J* 74:536–539
33. Dunn WR, Hamilton DD (1986) The critical incident technique—a brief guide. *Med Teach* 8:207–215
34. Dachelet CZ, Wemett MF, Garling EJ, Craig-Kuhn K, Kent N, Kitzman HJ (1981) The critical incident technique applied to the evaluation of the clinical practicum setting. *J Nurs Educ* 20:15–31
35. Newble DI (1983) The critical incident technique: a new approach to the assessment of clinical performance. *Med Educ* 17:401–403
36. Kamppainen JK (2000) The critical incident technique and nursing care quality research. *J Adv Nurs* 32:1264–1271
37. Krippendorff K (1980) *Content analysis: an introduction to its methodology*. Sage, Thousand Oaks
38. Neuendorf KA (2001) *The content analysis guidebook*. Sage, Thousand Oaks