The German Inpatient Satisfaction Scale: A Large-Scale Survey of Perceived Quality by Inpatients

Winfried Zinn1,2, Sebastian Sauer3,4, and Richard Göllner5

Abstract
A patient satisfaction instrument that is based on a large, representative sample does not yet exist in the German language. The objective of this article was to fill this gap by providing initial validation evidence of the German Inpatient Satisfaction Scale as an instrument for German-speaking countries. We performed an exploratory factor analysis and exploratory structural equation modeling in a cross-sectional design. The instrument was administered to N = 116,325 patients in hospitals in German-speaking countries. The overall response rates ranged from 63% to 98%. We found that a four-factor solution fit the data well. The four factors represented satisfaction with doctors’ care, nursing care, service facilities, and secondary care facilities. Cronbach’s alpha ranged from .72 to .90. The findings may be of practical interest for health care providers for measuring patient satisfaction.

Keywords
patient satisfaction, survey, questionnaire, validation, factor analysis, inpatient

The idea of assessing the quality of care from the patient’s point of view has been a research topic since the 1950s, but it has recently received a surge in interest (Görtz-Dorten, Breuer, Hautmann, Rothenberger, & Döpfner, 2011; Hansen et al., 2010; Parsons, 1951). The importance of this area of research is threefold. First, patient satisfaction is a health care outcome in its own right. Second, satisfied patients are more likely to show better compliance with treatments. Third, from an organizational point of view, understanding the causes, correlates, and consequences of patient satisfaction with health care services can help to improve the quality of these services. In sum, patient satisfaction is an element of medical care of increasing importance (S. J. Williams & Calnan, 1991).

One definition of patient satisfaction can be summarized as a “positive evaluation of distinct dimensions of the health care system” (Linder-Pelz, 1982, p. 578). In more detail, patient satisfaction entails beliefs, evaluations, and reactions to the context, process, and results of a health care provider’s service (Pascoe, 1983). A large number of different approaches that can be used to explain the causes of and the processes behind patient satisfaction have been proposed. For example, discrepancy models have been put forth to explain patient satisfaction as a function of perceived quality and expectancy—the smaller the gap, the higher the resulting satisfaction (B. Williams, 1994).

Although some insights into the processes behind patient satisfaction have been achieved, more questions await answers. To this end, valid measurement instruments are a prerequisite for furthering the understanding of the construct as well as for improving treatment. A number of instruments for assessing patient satisfaction have been developed in English, including, as a prominent example, the Hospital Consumer Assessment of Health-Care Providers and Systems (HCAPS; Giordano, Elliott, Goldstein, Lehrman, & Spencer, 2010), but a need for culturally sensitive adaptations has been identified (Aharony & Strasser, 1993). With regard to German-speaking countries, some patient satisfaction instruments exist (Kleeberg et al., 2005), but most of these have been based on relatively small sample sizes that are most likely not representative of the general patient population.

Furthermore, studies have largely failed to address whether patient satisfaction is homogeneous across different subgroups (e.g., patients in different medical departments). For a given scale, research must confirm that such subgroups share a common understanding of patient satisfaction for the common practice of comparing and combining such scores to be meaningful.

1Applied University for Health and Sports, Berlin, Germany
2Forschungsgruppe Metrik, Bermuthshain, Germany
3FOM University of Applied Sciences, Munich, Germany
4Samueli Institute, Alexandria, VA, USA
5University of Tübingen, Germany

Corresponding Author:
Sebastian Sauer, Institute of Business Psychology, FOM University of Applied Sciences, Arnulfstr. 32, 80335 Munich, Germany.
Email: Sebastian.sauer@fom.de
In light of this, the aim of the present study was to present initial validation evidence for an instrument designed to measure inpatient satisfaction in German-speaking countries. In addition, we tested whether the factor structure of the present instrument would be invariant across different medical departments.

**Current Developments and Research Questions**

The use of patient satisfaction surveys became a routine procedure over the course of the 1990s (Di Palo, 1997). One instrument that was developed in 1995 in the United States and that received a substantial amount of interest was the HCAPS. As it was extended and refined over the years that followed, the instrument increasingly became a widely acknowledged and employed survey instrument (Quigley, Elliott, Hays, Klein, & Farley, 2008). A German version has also been presented (Squires et al., 2012).

A frequent problem encountered with many satisfaction instruments is item skewness, which thus limits the researcher’s ability to distinguish between patient satisfaction levels when using the instrument (Sitzia & Wood, 1997). For example, many respondents tend to choose good and very good as the answer to a given item. Administered answer options should be designed to allow for enough differentiation in the light of such response patterns. One current remedy is to include a response category above very good, such as excellent (Young, Meterko, & Desai, 2000).

**Patient Satisfaction in Different Health Care Systems**

As mentioned above, the health care system has a large influence on the satisfaction of the patients and what the patients can or must be asked about. Most patients in the United States have to pay for their treatment, and therefore it would be appropriate to ask them how they evaluate the price-performance ratio. In German-speaking countries, the treatment is mostly covered by the mandatory health insurance. Therefore, different health care systems need different instruments, and the results from one country cannot be transferred to another (Aharony & Strasser, 1993).

**Patient Satisfaction in German-Speaking Countries**

What is known about the situation in German-speaking countries? For example, the Munich Patient Satisfaction Scale (MPSS-24) is a 24-item scale that was developed with exploratory factor analysis in three small samples (n = 85, n = 161, n = 91; Möller-Leimkühler et al., 2002); one principal factor was found. The authors reported a high internal consistency as well as a good reliability; convergent validity was found to be satisfactory. In addition, it has been utilized in a number of studies, for example, to evaluate medical care (Clever, Jin, Levinson, & Meltzer, 2008; Hannemann-Weber, Kessel, Budych, & Schultz, 2011; Hojat et al., 2011; Zickmund, Hillis, Barnett, lppolito, & LaBrecque, 2004).

Although the need for a validated language-specific patient satisfaction questionnaire is obvious, it appears that there is a paucity of measures that were developed with a rigorous methodology for German-speaking countries. It was the aim of this article to help fill this gap. We thus hypothesized that the instrument under investigation in the present article, the German Inpatient Satisfaction Scale (GISS), would demonstrate adequate psychometric properties.

**Method**

**Procedure and Sample**

Hospital staff members in German-speaking countries (Germany, Switzerland, and Austria) collected the data during the patients’ stay at the respective hospital. Care was taken so that the questionnaire was distributed after lunch and collected before dinner.

In sum, 143 hospitals participated, out of which 115 (80%) were from Germany, six (4%) came from the German-speaking parts of Switzerland, and 22 (16%) stemmed from Austria. It should be noted that the results apply mainly for Germany. In total, 24% of the hospitals were in private ownership, and the hospitals are spread all over Germany, Switzerland, and Austria. Some of them belong to same, and some to different clinic groups. The sample was not drawn randomly, but the hospitals were approached for participation in the study in the course of business partnership with the first author at that time.

As an incentive to participate in the study, each hospital received feedback about their average level of patient satisfaction in comparison with the respective mean. Hospitals were allowed to use this feedback for internal quality management purposes.

Overall, our sample size was N = 116,325 patients. The mean number of patients per hospital was 813 (SD = 775). Data were collected between 2005 and 2009. We took care to access a representative sample of all different somatic medical departments. The inclusion criteria were an age of 18 or older, status as an inpatient in the hospital under investigation when the survey was administered, and the ability to understand and respond to the items. An exclusion criterion was admission to a psychiatric or psychosomatic ward. Patients were informed about the aims of the study and filled out the survey anonymously.

To gauge the degree to which our sample is representative for the general population, we compared age and sex distribution of our sample with the general patients’ population based on statistics from the Statistisches Bundesamt (2015). Over and above, the figures for the age distribution for our sample and the overall population were similar (15–44 years: 21% vs. 17%; 45–64 years: 31% vs. 25%; 65 years and older: 48% vs. 59%). As to sex, the figures were identical (GISS sample male: 47%; general population 47%). Thus, some support for the notion
that our sample was representative is present. However, more characteristics would be needed to back up this claim.

**Questionnaire**

To develop the GISS’s initial item set, we included results from focus groups, expert interviews, published research, patient feedback, as well as the clinical practices of the authors and their colleagues. This process resulted in an initial set of 36 satisfaction items. In addition, we asked the patients how important the main aspects of the hospital (e.g., doctors, nursing staff, eating) were to them. In a second step, the instrument received a revision, which resulted in a reduction in the number of items. Later, the instrument was extended by including additional items as a result of analyses of patients’ qualitative comments (Zinn, 2010, p. 177).

The present study examined the 28 satisfaction items from the GISS. Each item stem was formulated as a semantic differential with five response options (1 = the best I ever experienced, 2 = very good, 3 = good, 4 = acceptable, 5 = bad). All items were later reverse-coded so that they would reflect satisfaction rather than dissatisfaction.

**Analysis**

We used *Mplus* 5.2 (Muthén & Muthén, 2007) for all analyses. Alpha was set at .01. The primary objective was to analyze the psychometric quality of the items by means of exploratory principal axis factor analysis (EFA). To determine the number of factors, we used Horn’s parallel analysis (Glorfeld, 1995) and Velicer’s minimum average partial (MAP) test (O’Connor, 2000). Horn’s parallel analysis compares the number of extracted factors with the number of factors that would have been extracted if the data had been completely random. The focus of the MAP test lies on the relative amounts of systematic variance remaining in the correlation matrix after extracting increasing numbers of factors. For EFA, we decided to use a varimax rotation. In addition, more basic criteria (the screeplot “elbow”; the Kaiser criterion, that is, the number of eigenvalues > 1) were considered. Following, descriptive statistics were calculated to describe means, standard deviations, and reliabilities of derived factor scores. In addition, the large sample size allowed for a more fine-grained analysis on the level of medical departments. The following 10 departments were included for this purpose: surgery, internal medicine, gynecology, orthopedics, ophthalmology, otolaryngology (ear/nose/throat), urology, neurology, oncology, and other. We used full information maximum likelihood (FIML) estimation in *Mplus* (Version 7.1) to deal with missing data. In the methodological literature on missing data (Schafer & Graham, 2002), there is a growing consensus that FIML estimations (and other procedures, for example, multiple imputation) are preferable to casewise or listwise deletion.

In addition to factor analysis, to test whether the factor solution held across all subgroups of medical departments, we used an exploratory structural equation modeling approach (ESEM). The ESEM model is estimated separately for each group, and some parameters can be constrained to be invariant across those groups. In the present study, we used multigroup ESEM tests of full measurement invariance of EFA factors across all groups. We thus tested whether the factor structure was identical in all subgroups. To test for the appropriateness of this invariance assumption, we used common fit indices for this purpose; that is, we used the comparative fit index (CFI), the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), and the chi-square test statistic to evaluate the goodness of fit. The CFI ranges from 0 to 1, with values greater than .90 and .95 typically taken to reflect acceptable and excellent fits to the data, respectively. RMSEA values of less than .05 and .08 reflect close and reasonable fits, respectively; values between .08 and .10 reflect a moderate fit, and values greater than .10 are generally considered unacceptable. Given normally distributed outcomes and a large sample size, the cutoff value for the SRMR should be close to .07 (Hu & Bentler, 1995). Prior to analysis, we examined the intraclass correlation coefficient (ICC) of all items. The ICC represents the proportion of variance that can be attributed to higher level units (i.e., hospitals).

**Results**

The overall response rate per department for each hospital ranged from 63% to 98%. The average age was 59.1 years (SD = 18.4); 53% of the patients were female. The mean length of stay was 7.6 days (SD = 8.2). It is informative to compare these figures with German averages: 53% of inpatients in German hospitals are female, and the mean duration is 7.4 days (Statistisches Bundesamt, 2015).

All items were checked for nonnormality prior to the analyses. Skewness and kurtosis in the data were small (for all indicators, skewness ranged from 0.04 to 0.58, kurtosis from 0.27 to 0.73). A robust Maximum Likelihood (MLR) estimation was employed as the estimation method for the analysis. Due to only small skewness and kurtosis in the data, (largely) unbiased model fit estimates could be expected (West, Finch, & Curran, 1995).

The overall score for the 28 satisfaction items was between 2.10 and 2.96 (on a scale ranging from 1 = best I ever experienced to 5 = bad). The ICC coefficients for the items ranged from .06 to .13, so that 6% to 13% of the variance in patients’ ratings was explained by the difference in hospitals. These results indicate that a relatively small percentage of item variances was attributable to variations among hospitals.

**EFA**

An examination of the screeplot indicated a one-factor solution with an eigenvalue of 12.91 (see Figure 1). By contrast, the Kaiser criterion favored a four-factor solution as four
factors had eigenvalues > 1 (eigenvalues: Factor 1: 12.91; Factor 2: 1.63; Factor 3: 1.31; Factor 4: 1.15).

On a more fine-grained level, Horn’s parallel analysis (see Figure 1) and the MAP test were employed on the basis of indicator correlations. Both parallel analysis and the MAP test indicated four factors. Given these results, we decided to force four factors in the factor analysis.

The four-factor solution showed a clearly interpretable structure. The results can be found in Table 1. For better readability, values above .3 are printed in bold. The four factors explained 61% of the variance (Factor 1: 19%; Factor 2: 18%; Factor 3: 12%; Factor 4: 12%). The first factor could tentatively be called satisfaction with medical doctors’ care. The second factor described satisfaction with nursing care; the third factor indicated satisfaction with service facilities (e.g., food, patient rooms, cafeteria), and the fourth factor captured satisfaction with secondary care facilities (e.g., physiotherapy, X-rays). Some of the items were excluded due to cross-loadings or relatively small factor loadings. After dropping seven items, 21 of the initial 28 items remained, showing a clear factor structure (see Table 1).

**Internal Consistencies of Factor Scores**

On the basis of the established factor solution, we further analyzed the psychometric properties of the factor scores. The internal consistency of the first factor (satisfaction with medical doctors’ care) was high. For the second scale (satisfaction with nursing care), the reliability was again high. Excluding any one item increased alpha by only .01. The alpha value for the third scale (service facilities) was acceptable again; excluding any of the items did not increase alpha. For the fourth scale (secondary care services), alpha was again good. Excluding any one item did not increase alpha. The correlations between the scales are shown in Table 2. The Pearson correlation coefficients ranged from .43 to .66.

**Factor Structure in Subgroups**

Next, we were interested in whether the factor structure would be invariant across different medical departments. To compare patient satisfaction across different departments, it is necessary that the factor structure be identical or at least similar across the subgroups (see Table 3 for sample sizes in different medical departments).

For eye patients, the factor structure differed from the rest of the medical departments; thus, we removed eye patients from the ESEM. As the factor score results showed moderate but substantial correlations, we decided to use an oblimin rotation for the ESEM analysis. The nine-group model with no invariance constraints provided a good fit to the data, \( \chi^2(2448) = 88495.1, \) RMSEA = .05, CFI = .94, Tucker–Lewis index (TLI) = .92, and SRMR = .03. These results supported the configural invariance of the four proposed factors, meaning that the same factor structure was able to fit the data for each group. Next, we constrained the factor loadings, factor intercepts, and item uniqueness to be invariant across the nine groups. To do this, 84 new constraints were added. A lack of support for this model would suggest that the measurement model was not comparable across the nine groups. The differences in the fit indices were small, \( \chi^2(3744) = 106438.0, \) RMSEA = .046, CFI = .927, TLI = .934, and SRMR = .041. Although the chi-square value for the configural invariance model was significantly smaller than that of the full invariance model (due to the large sample size), the chi-square/df ratio was substantially smaller. The small changes in the fit indices supported interpretations of invariance (Cheung & Rensvold, 2001) and provided clear evidence for the comparability of factor solutions across all of the patient groups.

Taken together, the results of the ESEM showed that the factor structure was similar across all medical departments with the exception of ophthalmology. Ophthalmology patients seemed to have a different understanding of the meaning of patient satisfaction compared with all other somatic departments (including neurology). For ophthalmology patients, we found that several items loaded on two factors.

In sum, we conclude that the results of the ESEM analysis could be taken as additional support for the validity of the instrument.

**Discussion**

The overall goal of the present study was to present initial validation evidence for an instrument measuring inpatient satisfaction in German-speaking countries (the GISS). In addition, we tested whether the factor structure of the GISS would remain invariant across different medical departments. The particular strength of the study was its large-scale scope, which allowed us to derive a representative picture of German-speaking inpatients across most of the main medical departments.
Table 1. Varimax Rotated Factor Loadings of the Items on the Four Factors.

| No. | Item                                                                 | Factors |       |       |       |       |       |       |       |       |       |       |
|-----|----------------------------------------------------------------------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |                                                                      | Doctors | Nurses | Service | Secondary care | $h^2$  |       |       |       |       |       |       |
| 12  | I feel that the medical doctors adequately inform me about my treatment | 0.80    | 0.24  | 0.17   | 0.19   | 0.76  |       |       |       |       |       |       |
| 11  | The doctors answer my questions during their ward rounds in an informative and friendly manner | 0.77    | 0.26  | 0.17   | 0.20   | 0.73  |       |       |       |       |       |       |
| 13  | The doctors are friendly                                             | 0.67    | 0.31  | 0.19   | 0.22   | 0.63  |       |       |       |       |       |       |
| 10  | Diagnoses are transmitted with a lot of empathy                       | 0.59    | 0.36  | 0.20   | 0.26   | 0.59  |       |       |       |       |       |       |
| 26  | I am well informed about the possible complications of my condition that I may face after I leave the hospital | 0.49    | 0.28  | 0.30   | 0.28   | 0.48  |       |       |       |       |       |       |
| 15  | My medical care has been successful so far                            | 0.42    | 0.29  | 0.21   | 0.29   | 0.39  |       |       |       |       |       |       |
| 14  | My pain has been alleviated efficiently                              | 0.38    | 0.37  | 0.20   | 0.30   | 0.41  |       |       |       |       |       |       |
| 27  | On the basis of all of the experiences I have had so far, my overall judgment of the hospital is positive | 0.42    | 0.42  | 0.46   | 0.27   | 0.63  |       |       |       |       |       |       |
| 28  | I will recommend the hospital                                       | 0.40    | 0.29  | 0.34   | 0.29   | 0.59  |       |       |       |       |       |       |
| 25  | I am well prepared to do what I need to do after I leave the hospital | 0.39    | 0.29  | 0.34   | 0.29   | 0.43  |       |       |       |       |       |       |
| 20  | My family is well informed                                          | 0.37    | 0.30  | 0.30   | 0.30   | 0.41  |       |       |       |       |       |       |
| 6   | The daily support of the nursing staff is excellent                   | 0.24    | 0.75  | 0.21   | 0.17   | 0.69  |       |       |       |       |       |       |
| 8   | The nursing staff members are very friendly                           | 0.24    | 0.71  | 0.20   | 0.16   | 0.62  |       |       |       |       |       |       |
| 4   | The nursing staff members are Considerate of the things that worry me or make me sad | 0.26    | 0.70  | 0.19   | 0.21   | 0.64  |       |       |       |       |       |       |
| 5   | The nursing staff members are open to my suggestions for how services could be improved | 0.24    | 0.66  | 0.21   | 0.22   | 0.59  |       |       |       |       |       |       |
| 7   | The nursing staff members keep me well informed about hospital procedures | 0.30    | 0.62  | 0.22   | 0.22   | 0.57  |       |       |       |       |       |       |
| 9   | The nursing staff members respect my privacy                          | 0.27    | 0.56  | 0.23   | 0.26   | 0.51  |       |       |       |       |       |       |
| 3   | The general personnel are very friendly upon admission to the hospital | 0.24    | 0.44  | 0.23   | 0.28   | 0.38  |       |       |       |       |       |       |
| 21  | Patient rooms are excellent                                          | 0.14    | 0.17  | 0.73   | 0.12   | 0.59  |       |       |       |       |       |       |
| 22  | All facilities are very hygienic                                      | 0.16    | 0.22  | 0.70   | 0.15   | 0.59  |       |       |       |       |       |       |
| 24  | The food is good                                                      | 0.15    | 0.17  | 0.40   | 0.21   | 0.26  |       |       |       |       |       |       |
| 23  | The cafeteria is good                                                 | 0.12    | 0.12  | 0.34   | 0.24   | 0.20  |       |       |       |       |       |       |
| 1   | Signposting and directions are very clear                             | 0.18    | 0.19  | 0.30   | 0.30   | 0.25  |       |       |       |       |       |       |
| 18  | My experiences with the X-ray department have been good              | 0.17    | 0.17  | 0.17   | 0.77   | 0.67  |       |       |       |       |       |       |
| 19  | My experiences with other medical services such as ECG or endoscopy have been good | 0.23    | 0.21  | 0.17   | 0.71   | 0.62  |       |       |       |       |       |       |
| 16  | Waiting times are not too long                                        | 0.23    | 0.19  | 0.22   | 0.54   | 0.43  |       |       |       |       |       |       |
| 17  | My experiences with the physiotherapy service are good               | 0.17    | 0.22  | 0.14   | 0.44   | 0.29  |       |       |       |       |       |       |
| 2   | The forms are easy to understand                                     | 0.22    | 0.27  | 0.27   | 0.35   | 0.32  |       |       |       |       |       |       |

*Items were excluded for psychometric reasons (i.e., either low factor loading or cross-loadings on additional factors). ECG = electrocardiogram. Factor loadings greater than .30 were shown in bold face indicating items’ main loading.

In summary, our data spoke in favor of a four-factor solution of inpatient satisfaction as measured by the GISS. We found initial support for the validity of the instrument: The four factors were clearly interpretable and reflected the main domains of patients’ experiences in a hospital. A substantial amount of variance was explained by the four factors, thus indicating that the model was able to account for most of the variability in the data. The first factor described satisfaction with medical doctors’ care, the second factor captured satisfaction with nursing care, the third factor could be explained as satisfaction with service facilities, and the fourth factor captured satisfaction with secondary care facilities.

The analysis of the subgroups provided an additional validation of the factor structure because most somatic departments showed the same factor structure. This was confirmed not only by individual exploratory factor analyses, but additional
Table 2. First-Order Correlations for the Four Factor scores.

|        | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|--------|----------|----------|----------|----------|
| Doctors|          |          |          |          |
| Factor 1| 1        | .66      | .48      | .53      |
| Factor 2|          | 1        | .50      | .49      |
| Factor 3|          |          | 1        | .43      |
| Factor 4|          |          |          | 1        |
| M      | 2.49     | 2.41     | 2.78     | 2.75     |
| SD     | 0.65     | 0.63     | 0.65     | 0.71     |
| Cronbach's α | .89 | .90 | .72 | .80 |

Table 3. Sample Sizes per Medical Department.

| Medical department | Sample size | Percentage of total sample |
|---------------------|-------------|---------------------------|
| Internal medicine   | 34,214      | 29                        |
| Surgery             | 28,062      | 24                        |
| Other               | 17,396      | 15                        |
| Gynecology          | 11,479      | 10                        |
| Orthopedics         | 6,734       | 6                         |
| Urology             | 5,308       | 5                         |
| Neurology           | 4,574       | 4                         |
| Oncology            | 3,518       | 3                         |
| Otolaryngology      | 3,140       | 3                         |
| Ophthalmology       | 1,900       | 2                         |

psychometric methodology (ESEM) was also able to back up this claim—with the exception of patients in ophthalmology departments. For patients in ophthalmology departments, 14 items loaded on two or more factors, suggesting that such patients seem to have a different understanding of their satisfaction and what it means. One plausible post hoc explanation for this finding is that such patients are in a special situation as their visual perception is damaged. Thus, their perception of what satisfaction means to them is different on an elementary perceptual level. It is a novel finding that independent research may be needed for ophthalmology patients but not for other somatic patients. Clearly, more research is needed before definitive conclusions can be drawn.

The fact that the item responses did not show strong skewness is particularly encouraging because positive skewness (i.e., many responses well above a medium level of satisfaction) is a problem for satisfaction questionnaires as many procedures, including factor analysis, rely on the assumption of normality (Gavra, 1997; Montanari & Viroli, 2010; Peterson & Wilson, 1992).

Finally, it has to be borne in mind that the analysis of patient satisfaction was conducted at the level of individual patients. The present study showed that about 6% to 13% of variance could be attributed to differences among hospitals. Even though such differences can be regarded as small, further research should explicitly account for possible clustering effects including cluster-adjusted standard errors (Murray, Varnell, & Blitstein, 2004) and the modeling of patient ratings at different levels of analysis. For instance, it would be very interesting to see if the derived factor structure is also suitable to assess the quality of care at the level of departments or hospitals. An extension of the present analysis in terms of multilevel factor analysis is highly promising and points to an important aspect for future research.

Some shortcomings of the present study need to be borne in mind. For example, any questionnaire needs to be validated with regard to more than just its internal structure (e.g., by using EFA). In addition, external criteria need to be considered. For example, is patient satisfaction as measured by the GISS positively associated with “hard” quality indicators such as time between patient admission and surgery? There are certainly a substantial number of external criteria that may serve as (convergent or divergent) indicators of the external validity of the GISS.

The analysis of patient satisfaction was conducted at the level of individual patients. Further research should examine if the derived factor structure (satisfaction with doctors, nurses, service, and secondary care) is also suitable to assess the quality of care at the level of departments or hospitals. The fact that systematic variations in patient satisfaction were found for all survey items (ICCs ranged between .06 and .13) points to an important aspect for future research.

For the time being, we conclude that the results of the present study provide initial support that the GISS provides an adequate measure of inpatient satisfaction on four dimensions: doctors’ care, nursing care, service facilities, and secondary care. From a practical point of view, care providers may consider using the GISS for satisfaction surveys. More research is needed to substantiate knowledge about the value of the instrument; for example, the nested structure of the data should be analyzed in future studies. Moreover, other German-speaking countries than Germany should be given more emphasis in future studies to provide insight whether patient satisfaction differs in those countries.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

Aharony, L., & Strasser, S. (1993). Patient satisfaction: What we know about and what we still need to explore. Medical Care Review, 50, 49-79.

Cheung, G. W., & Rensvold, R. B. (2001). The effects of model parsimony and sampling error on the fit of structural equation models. Organizational Research Methods, 4, 236-264.

Clever, S. L., Jin, L., Levinson, W., & Meltzer, D. O. (2008). Does doctor-patient communication affect patient satisfaction with
Di Palo, M. T. (1997). Rating satisfaction research: Is it poor, fair, good, very good, or excellent? *Arthritis Care & Research*, 10, 422-430.

Gavra, T. G. A. (1997). *Improving your measurement of customer satisfaction: A guide to creating, conducting, analyzing, and reporting customer measurement programs*. Wilwaukee, WI: ASQ Press.

Giordano, L. A., Elliott, M. N., Goldstein, E., Lehrman, W. G., & Spencer, P. A. (2010). Development, implementation, and public reporting of the HCAHPS survey. *Medical Care Research and Review, 67*, 27-37.

Glorfeld, L. W. (1995). An improvement on Horn’s Parallel Analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement, 55*, 377-393.

Görtz-Dorten, A., Breuer, D., Hautmann, C., Rothenberger, A., & Döpfner, M. (2011). What contributes to patient and parent satisfaction with medication in the treatment of children with ADHD? A report on the development of a new rating scale. *European Child & Adolescent Psychiatry, 20*(2), 1-11.

Hannemann-Weber, H., Kessel, M., Budych, K., & Schultz, C. (2011). Shared communication processes within healthcare teams for rare diseases and their influence on healthcare professionals’ innovative behavior and patient satisfaction. *Implementation Science, 6*(1), Article 40.

Hansen, L. K., Vincent, S., Harris, S., David, E., Surafudheen, S., & Kingdon, D. (2010). A patient satisfaction rating scale for psychiatric service users. *The Psychiatrist, 34*, 485-488.

Hojet, M., Louis, D. Z., Maxwell, K., Markham, F. W., Wender, R. C., & Gonnella, J. S. (2011). A brief instrument to measure patients’ overall satisfaction with primary care physicians. *Family Medicine, 43*, 412-417.

Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: SAGE.

Kleeberg, U. R., Tews, J. T., Ruprecht, T., Höing, M., Kuhlmann, A., & Runge, C. (2005). Patient satisfaction and quality of life in cancer outpatients: Results of the PASQOC* study. *Supportive Care in Cancer, 13*, 303-310.

Linder-Pelz, S. (1982). Toward a theory of patient satisfaction. *Social Science & Medicine, 16*, 577-582.

Möller-Leimkühler, A. M., Dunkel, R., Müller, P., Pukies, G., De Fazio, S., & Lehmann, E. (2002). Is patient satisfaction a unidimensional construct? *European Archives of Psychiatry & Clinical Neuroscience, 252*, 19-23.

Montanari, A., & Viroli, C. (2010). A skew-normal factor model for the analysis of student satisfaction towards university courses. *Journal of Applied Statistics, 37*, 473-487.

Murray, D. M., Varnell, S. P., & Blitstein, J. L. (2004). Design and analysis of group-randomized trials: A review of recent methodological developments. *American Journal of Public Health, 94*, 423-432.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user’s guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

O’Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer’s MAP test. *Behavior Research Methods Instruments & Computers, 32*, 396-402.

Parsons, T. (1951). *The social system*. London, England: Psychology Press.

Pascoe, G. C. (1983). Patient satisfaction in primary health care: A literature review and analysis. *Evaluation and Program Planning, 6*, 185-210.

Peterson, R., & Wilson, W. (1992). Measuring customer satisfaction: Fact and artifact. *Journal of the Academy of Marketing Science, 20*, 61-71.

Quijley, D. D., Elliott, M. N., Hays, R. D., Klein, D. J., & Farley, D. O. (2008). Bridging from the Picker Hospital Survey to the CAHPS Hospital Survey. *Medical Care, 46*, 654-661.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177.

Sitzia, J., & Wood, N. (1997). Patient satisfaction: A review of issues and concepts. *Social Science & Medicine, 45*, 1829-1843.

Squires, A., Bruyneel, L., Aiken, L. H., Den Heede, K., Brzostek, T., Busse, R., . . . Sermeus, W. (2012). Cross-cultural evaluation of the relevance of the HCAHPS survey in five European countries. *International Journal for Quality in Health Care, 24*, 470-475.

Statistisches Bundesamt [Federal Statistical Office]. (2015). *Fachserie 12 Gesundheitswesen Reihe 6.2.1 [Series 12 Healthcare 6.2.1]*. Wiesbaden, Germany.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Los Angeles, CA: SAGE.

Williams, B. (1994). Patient satisfaction: A valid concept? *Social Science & Medicine, 38*, 509-516.

Williams, S. J., & Calnan, M. (1991). Convergence and divergence: Assessing criteria of consumer satisfaction across general practice, dental and hospital care settings. *Social Science & Medicine, 33*, 707-716.

Young, G. J., Meterko, M., & Desai, K. R. (2000). Patient satisfaction with hospital care: Effects of demographic and institutional characteristics. *Medical Care, 38*, 325-334.

Zickmund, S., Hillis, S. L., Barnett, M. J., Ippolito, L., & LaBrecque, D. R. (2004). Hepatitis C virus-infected patients report communication problems with physicians. *Hepatology, 39*, 999-1007. doi:10.1002/hep.20132

Zinn, W. (2010). *Person-factors, patient satisfaction, standardization approaches, hospital, patient satisfaction measurement, the German-speaking, Germany*. Weingarten, Germany: University of Education-Weingarten. (In German)

Author Biographies

Winfried Zinn started as a nurse before he studied psychology. The main research interests are patient satisfaction, methods (particularly predictive modeling) and patient safety management.

Sebastian Sauer has worked as a consultant focusing on healthcare industry. Currently, he works as a research professor for Industrial Psychology at FOM University of Applied Sciences in Munich. His research interests include psychological methods including predictive modeling, and mindfulness.

Richard Göllner is a post-doctoral researcher at the Hector Research Institute of Education Sciences and Psychology in Tübingen. His main research interests concern the assessment of contextual attributes as determinants of individuals’ emotional, social and personal development.