The measurement and valuation of health status using EQ-5D: A European perspective

Evidence from the EuroQol BIOMED Research Programme

Edited by
Richard Brooks
Rosalind Rabin
Frank de Charro

Kluwer Academic Publishers

A C.I.P. Catalogue record for this book is available from the Library of Congress.
ISBN 1-4020-1214-4

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Sold and distributed in North, Central and South America
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

Printed on acid-free paper

Printed in the Netherlands.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

THE EQ-5D SELF-REPORT QUESTIONNAIRE (EQ-5D)

EQ-5D essentially consists of 2 pages, i.e. the EQ-5D descriptive system (often referred to as 'page 2') and the EQ VAS (often referred to as 'page 3'). These pages are combined with a cover page ('page 1') and optional demographic questions.

The descriptive system records the level of self-reported problems on each of the 5 dimensions of the classification (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). Each dimension is divided into 3 levels (no problems, some or moderate problems, extreme problems or unable). This information can be used in the following ways:

(i) As an EQ-5D health profile for individuals or groups, either at a single point in time, or over a period of time; differences in such profiles can be used to describe health outcomes.

(ii) Health states defined by the 5-dimensional descriptive system can be converted into a weighted health state index by applying scores from 'value sets' elicited from general population samples.

A convenient way of identifying health states is employed: by assigning the numbers 1,2,3 to each level in each of the 5 dimensions we can label health states using 5-number descriptors ranging from 11111 to 33333 (a total of 3^5 or 243 states). Thus 12121 is the state ' no problems in walking about, some problems washing or dressing, no problems performing usual activities, moderate pain or discomfort, not anxious or depressed'. 'Unconscious' and 'death' are also included in the valuation exercise to give a total of 245 states addressed by EQ-5D.

The EQ VAS records the respondent's self-rated health status on a graduated (0 -100) visual analogue scale (VAS), where the endpoints are labelled 'best imaginable health state' and 'worst imaginable health state'. This information can be used as a quantitative measure, and differences in this scale can be used as a measure of outcome, as judged by the individual respondents.

THE EQ-5D VALUATION QUESTIONNAIRE

A fuller version of the instrument contains the valuation task (pages 4-7). These pages provide a technique for valuing health states and have been used to value heath states in population studies. Valuations elicited in this way can be applied to health states generated by the EQ-5D descriptive system to produce a weighted health state index. Sets of values are available for a number of countries (see the website for current availability).

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

studies in order to monitor follow-up after a medical intervention. Measurement on the 5 distinct dimensions of the EQ-5D can also be informative in describing the dynamics of HRQoL during treatment and follow-up.

Economic studies

Economic studies have been used alongside clinical studies. Such studies are explicitly concerned with comparisons between alternative medical interventions. Full economic evaluation requires the assessment, measurement and valuation of both inputs to a medical intervention, and the outcomes of these interventions. Outcomes can be valued in monetary terms (in cost-benefit analysis), in 'natural' units such as lives saved (in cost-effectiveness analysis) and in utility terms (in cost-utility analysis). Some analysts are reluctant to use the term 'cost-utility analysis', preferring to include such work under cost-effectiveness analysis. Whichever terminology is used, the effect of an intervention is judged by valuing health status before and after the intervention. This valuation of health status is undertaken by assigning a single numeric value (formally called 'utility') between zero (say, dead) and 1 (say, perfect health) for different health states. The EQ-5D instrument has been expressly developed for application in these types of economic analysis.

Population health studies

Population health studies differ from the other 2 fields of research, as they are not focused on the effectiveness or cost-effectiveness of a particular medical intervention. These studies are predominantly concerned with quantifying the burden of diseases, monitoring disease-specific trends or changes and comparisons between regions and countries. Generic instruments such as the EQ-5D enable such comparisons to be made in HRQoL terms.

Study designs

Different study designs can be identified that are associated with the 3 fields of research mentioned above. The application of the EQ-5D is suitable both for observational studies and RCTs. The choice for a particular study design simply affects the soundness of the comparisons that can be made between different study groups or between measurement at different points of time. Thus the type of study design determines the kind of research questions that can be properly addressed and the associated statistical testing that is appropriate for these purposes.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

With repeated data collection 3 different situations can arise:

(i) Late entry into the study.
(ii) Missing information.
(iii) Drop-out from the study.

Efforts to minimise the proportion of missing data and also the systematic biases that may cause this phenomenon are very important and should be a major concern for any study and for any outcome measure (Curran *et al*, 1998).

Because the EQ-5D descriptive system is based on only 5 independent dimensions, methods of handling missing data for 1 or more of the 5 dimensions are not adequate. In most cases a missing answer on 1 of the dimensions will lead to the complete rejection of the questionnaire. Only in repeated or longitudinal studies can some of the proposed solutions to overcome missing questionnaires be applied (at least partially). To analyse longitudinal HRQoL data, multi-level analysis may be used to overcome this problem (Goldstein, 1995).

2.3 ANALYSIS

Statistical models

Several statistical models are available to analyse and test various assumptions. In general, the analysis of variance (ANOVA) model is suited to test for mean differences between groups. If we are dealing with only 2 groups, ANOVA and the t-test are identical. An extension of this basic statistical model, the analysis of covariance (ANCOVA), may be appropriate when dealing with pre-assessments (see next section), which is the case in many clinical studies. For those situations where responses for the same person are collected on several occasions, the multivariate analysis of variance (MANOVA) model is required. MANOVA is also suitable for testing for mean differences between groups of persons and their repeated measures in a single analysis (Tabachnick and Fidell, 1989).

The same type of statistical analysis can also be performed by applying regression analysis with dummy coding. Because the method of coding, especially in the MANOVA type of analysis, is rather technical and requires substantial knowledge of the underlying statistical theory, this approach is only recommended for suitably trained researchers. However, the advantage of applying regression analysis may be the imputation of various variables to control for covariates. In addition, regression analysis provides a prediction model, which on some occasions may be the purpose of the study. Conventionally, regression analysis was carried out by using the linear optimal least squares (OLS) regression model. A more advanced regression model, that is being used more frequently in quality of life studies and outcomes research, is

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Scores obtained with the descriptive system can be compared with the scores of a reference group (e.g. Table 2.3).

Table 2.3. Frequency distribution (profile) of the EQ-5D descriptive system for comparison of more than 1 population or repeated measurements

EQ-5D profile	Dutch[a] (general population)	Migraine patients[b] (Dutch population)	Migraine[c]
Mobility			
no problems (%)	95.1	83.0	83.0
some problems (%)	4.7	16.3	16.3
confined to bed (%)	0.2	0.7	0.7
Self-Care			
no problems (%)	96.7	97.2	97.2
some problems (%)	2.9	2.8	2.8
unable to (%)	0.3	0.0	0.0
Usual Activities			
no problems (%)	85.8	72.4	72.4
some problems (%)	11.7	26.4	26.4
unable to (%)	2.4	1.2	1.2
Pain/Discomfort			
none (%)	68.0	49.5	49.5
moderate (%)	30.5	46.6	46.6
extreme (%)	1.5	3.9	3.9
Anxiety/Depression			
none (%)	83.6	73.0	73.0
moderate (%)	15.7	24.6	24.6
extreme (%)	0.7	2.5	2.5

a. Essink-Bot *et al*, 1993
b. Essink-Bot *et al*, 1995
c. Criteria International Headache Society. Based on: Essink-Bot *et al*, 1995

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

3
EQ-5D: Modes of administration

Caroline Selai, Idoia Gaminde, Michael Herdman,
Jeffrey Johnson and Rosalind Rabin

3.1 INTRODUCTION

Chapter 1 briefly described the conceptual underpinning and the development of EQ-5D so that the reader will be aware that in developing EQ-5D, the EuroQol Group had a number of specific goals. EQ-5D was originally designed to be a self-administered questionnaire, to be used in postal surveys. There are many instances, however, when self-reported completion of a paper-and-pencil questionnaire is either impractical or not possible. Self-report may not be possible because the respondent is physically incapacitated or cognitively impaired. Administration of paper-and-pencil versions of a questionnaire is sometimes not the most efficient method of data collection and may not be practical in large multi-centre trials. If we alter the mode of administration of a questionnaire, why is this change important?

Small differences in the wording, layout and/or administration of a questionnaire can influence the results. In between-group comparisons, for example, it is essential to ensure that the conditions under which the questionnaire is administered are as similar as possible in the different groups. In longitudinal studies, where the same questionnaire is administered on several occasions to the same group, every effort must be made to ensure that the conditions of administration are as similar as possible on each administration. Where this is not the case, the onus is on the researcher to prove, empirically, that the use of alternative formats or different modes of administration make no substantial difference to the results obtained.

This chapter provides recommendations for researchers wishing to collect clinical data using the EQ-5D descriptive system and the EQ VAS. Recommendations for researchers wishing to collect EQ-5D valuation data are not included here: they are provided in Chapter 4. This chapter has two aims:

(i) To give a brief overview of the literature on alternative modes of administration (i.e. anything different from pencil-and-paper self-report) relevant to EQ-5D.
(ii) To make recommendations for researchers intending to collect EQ-5D data using any of the alternative modes of administration. Scripts to be used for telephone administration, for interviews and proxy-report are included as appendices.

R. Brooks et al. (eds.),
The Measurement and Valuation of Health Status Using EQ-5D: A European Perspective, 21–27.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

between responses from patients and those from their proxies for the more directly observable dimensions of EQ-5D, and that proxy agreement was less good for the more subjective dimensions. They concluded that, in general, information about some dimensions obtained from a proxy may be sufficiently valid and unbiased to be useable in most types of trials and surveys.

Dementia

The assessment of quality of life of patients with dementia is complex. Despite the many methodological issues, there are several HRQoL assessment techniques in development. Studies have shown that patients with mild-to-moderate dementia can rate their own HRQoL but that proxies are needed for patients with severe dementia. EQ-5D was administered in 2 studies to both patients with dementia and their carers (Coucill *et al*, 2000; Coucill *et al*, 2001; Selai, 2000a; Selai *et al*, 2001). Overall, the findings generally supported the results reported in the literature i.e. better agreement for observable items and poorer agreement for pain/discomfort and anxiety/depression. Two particular findings emerged from these studies. First, in both studies there was very poor agreement for the usual activities dimension. Qualitative data revealed that both patients and carers were asking, "what is usual?" Since dementia leads to a gradual loss of abilities, it was not clear what 'usual' activities meant. Second, it was found that differences emerged not only because the patients and proxies had different perceptions of the degree of problem (e.g. patient ticked 'moderate' but proxy ticked 'extreme'). In addition, there were differences in interpretation of the scope of the question. For example, some patients thought the question about mobility referred to 'ability to walk'. Some of the proxies also thought it referred to this but some proxies said they were thinking about orientation (e.g. patient's ability to find their way home from the shops, an ability involving a number of cognitive skills). So some of the discrepancies arose from differences in interpretation of the question. The methodological issues are currently the subject of debate in the HRQoL literature. We recommend that reporting of experimental design, methods and results be explicit and detailed.

Additional methodological considerations:

Differences may occur according to who the proxy is (spouse, health professional, etc.) and whether he or she lives with the patient or just visits. There is also a big difference according to the instructions given to the proxy (not often explicitly reported). For instance, is the proxy reporting either (a) how he or she (i.e. the proxy), judges the HRQoL of the patient or (b) how he or she (i.e. the proxy) thinks the patient would judge his or her own HRQoL? In the case of dementia, there are problems due to denial and loss of insight. At a certain point of (severe) dementia the proxy might say that he/she thinks that the patient has an extreme problem in the area of social life but that he/she recognises that the patient has lost insight and seems

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

These data may be treated as 5 separate items that constitute a profile of HRQoL and are usually represented graphically. When taken together these same data uniquely define 1 of the 243 possible EQ-5D health states. For reference purposes, this health state is labelled by combining the 5 separate digits that indicate the level of problem recorded for each dimension. Hence the health state 11232 is defined by responses indicating no problem with dimensions 1 and 2 (mobility and self-care), a moderate problem with dimensions 3 and 5 (usual activities and anxiety/depression) and an extreme problem with dimension 4 (pain/discomfort). In its most sophisticated form, this EQ-5D health state can be transformed into a single index score, once the rules for determining the relative value for each dimension/level have been determined.

Taken together, the 5 dimensions of the EQ-5D descriptive system can be manipulated in 4 formats:

(i) Single digit representation of level of problem on each dimension.
(ii) Collective representation of level of problem as a profile.
(iii) Encoded representation of profile as a unique health state.
(iv) Conversion of health state to corresponding weighted index.

The second source of information on HRQoL within EQ-5D is generated through the directly-assessed rating made on the standardised visual analogue scale that complements the descriptive record of problems dealt with above. This global rating yields a score in the range 0 to 100 (respectively the worst and best imaginable health states).

A number of assumptions regarding value can be introduced for each of these formats. The EuroQol Group has adopted an explicit position about the legitimacy of these assumptions but this does not preclude the possibility that others may wish to take a counter view. The key feature of these value assumptions is that they are reasonably transparent. Others who may wish to interpret EQ-5D differently will need to make clear any alternate value assumptions of their own. The basic data element of EQ-5D is the coded level of problem on each dimension. The numerals 1 to 3 are in fact no more than numeric labels with no intrinsic arithmetic properties. The numerals might just as well be substituted by the characters x, y and z. The numeric labels correspond to the hierarchical structure of the level of problem within each dimension (level 3 < level 2 < level 1). The use of the label as a cardinal score was never intended. To do so in this way would be to force a value system on to the data in ways that run counter to the express intentions of the Group. Interpreting the level codes as numbers implies for example, that within each dimension a level 2 problem lies midway between no problem and extreme problem. Furthermore, since the same assumption would be applied to each dimension in turn, the application of this value transformation implies that the value associated with a given level of problem is identical across all dimensions. These assumptions are dangerously appealing to the uninformed user but are to be generally avoided. Where they are encountered it should be

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

ular tendency for respondents who are older and/or less well educated, to omit a rating for 'dead' so that these vulnerable population subgroups are at additional risk of being censored as a consequence of the transformation process. In selecting the most appropriate source of transformed EQ-5D values it might be prudent for the user to take account of any possible biases introduced by this route.

4.5 QUESTIONS OF AGGREGATION

The source of reference values and the methods used to obtain them are two central issues in assessing the robustness of any measure of health status that claims to be a preference-based instrument. However, one further issue is relevant. Supposing that values have been obtained from an appropriate reference source then these will have to be represented by some sort of average. The aggregate form of these data might be one of several different measures of central tendency. The choice of such measures is often seen as being driven exclusively by statistical considerations, for example, where distributions are non-normal then the median would be preferred to the mean. There are related secondary aspects of such a choice in that appropriate statistical tests would be non-parametric with consequent limitations for the generalisation of results. In the area of preference measurement the choice of average measure is not entirely a statistical one - other perspectives can be legitimately applied. If the measurement of values for health states were treated as akin to a ballot of the sort encountered as part of the electoral process then the median value would be the natural corresponding choice. In some circumstances a case might even be made for the mode. The form of measure of central tendency may well be influenced by other factors, notably where incomplete sets of values have been used to interpolate values for health states. Since this type of modelling is common practice within the EuroQol Group it is recommended that end-users take note of any constraints that might arise as a result of the methodology adopted in estimating values for the complete set of EQ-5D health states.

4.6 APPLICATION

It is assumed here that EQ-5D has been collected with a view to measuring health status (HRQoL), sometimes for individuals or groups of individuals, and sometimes at several points of time. When used to measure health outcomes it is useful to note that this can be conducted in several frameworks, by different professionals with different objectives. Whilst it may reasonably be argued that a distinction between clinical and economic evaluation is artificial, it remains the case that health care professionals use health outcomes data in forms that are different to those adopted by health economists. Nothing is lost by retaining the option to adopt different values with which to make the conversion from ordinal ratings to a weighted index. This then is the entry point for the flow chart shown in Figure 4.1.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

in terms of a cost per unit change in self-assessed health status. Table 4.1 summarises the options in selecting EQ-5D weights for economic studies.

Table 4.1. Summary of choice factors in selecting EQ-5D weights for economic studies

	Cost-utility analysis	Cost-effectiveness analysis	Cost consequences
Format	Index	Index	Profile or Index
Utility weights	Yes	Optional	Optional
Source	General population	General population Population subgroups Self-assessment	General population Population subgroups Self-assessment
Normalised scores	Yes	Optional	Optional
0 –100 scale	No	Optional	Optional
Suggested default source	UK TTO A1	EQ-net VAS	Optional
own VAS	No	Optional	Optional

4.8 EQ-5D IN NON-ECONOMIC STUDIES

Health status measurement is a key activity in the planning and delivery of care. In addition to the role that EQ-5D has in evaluation, as indicated in the previous section, it is well suited to generating information for use in many other settings including clinical decision-making, informing public health policy and in clinical studies. For example, the issue may simply be to compare the health status of an individual with normative data based on population averages for someone of similar background. Such reference values are available for laboratory tests and it is wholly unexceptional to report say, a patient's blood serum levels and at the same time indicate the 'normal' range. Management or treatment decisions may stem from such information. In using EQ-5D to compare the health of patient groups or local populations with such normative reference values, we can establish variations that can impact on the formation of policy. Where such activity is put into place more systematically, then health authorities can utilise the information to monitor progress in achieving national targets for health improvements. Comparisons *between* provider units in different locations or between consultant specialties *within* a single institution can result in

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

5
Test-retest performance of EQ-5D

Susan Macran

5.1 INTRODUCTION

In order to assess the usefulness of a measure there is a need to determine its measurement (or psychometric) properties. This usually involves examining its reliability (the extent to which the results it provides are reproducible) and its validity (the extent to which it measures what it is intended to measure). The reliability and validity of a measure can be assessed in a number of different ways and the book by Streiner and Norman (1995) provides a readable guide. This chapter reports on the exploration of 1 aspect of reliability, the test-retest reliability or the reproducibility of scores, for the EQ-5D descriptive system and the EQ VAS. Although 2 previous studies have examined the test-retest reliability of health state values elicited using the EQ-5D VAS (Van Agt *et al,* 1994; Dolan *et al*, 1994), until now there have been no published findings on the reliability of the EQ-5D descriptive system and the EQ VAS.

An important requirement of the EQ-5D and indeed all health outcome measures, is that the measurements they produce are sensitive to change. When assessing the performance of a measure users need to know the extent to which any change reflected by a measure is due to 'real' change or due to some sort of measurement error. Such a characteristic is usually assessed by calculating the test-retest reliability of a measure. Test-retest reliability calculates the consistency of a measure over time, and is most commonly assessed by administering a measure to the same group of individuals on 2 occasions separated by a time period which is short enough to assume that the individuals' health will not have changed. This time interval is probably the most important consideration when assessing the test-retest reliability of a measure. If the interval is too short, respondents may remember what they answered on the first occasion, while if it is too long, the characteristic that is being measured may have changed. As the most appropriate time interval is dependent both on the task and the nature of the characteristic being measured there are no standard rules for determining this period.

Test-retest reliability is summarised by calculating a reliability coefficient, which defines the ratio of the variability in the 2 scores for a respondent to the total variance and thus the proportion of the total variance in the scores which is due to 'true' difference. Reliability coefficients range from 0 to 1.0, where 0 indicates no reliability and 1.0 indicates perfect reliability (no measurement error). While there are a number of methods for calculating such coefficients, the intra-class correlation coefficient (ICC)

R. Brooks et al. (eds.),
The Measurement and Valuation of Health Status Using EQ-5D: A European Perspective, 43–54.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Table 5.3 examines the relationship between a change in EQ VAS scores between time 1 and time 2 and change in the number of problems reported on the EQ-5D dimensions. Respondents who experienced a change of more than 3 points in their EQ VAS scores between time 1 and time 2 were more likely to experience a change in the number of problems reported on the EQ-5D dimensions than those who experienced no change or a change of 3 points or less on the EQ VAS. Those who experienced a negative change on the EQ VAS were more likely to change on 2 or more dimensions than those who experienced a positive change.

Table 5.3. Percentage change in EQ-5D dimensions between time 1 and time 2 by change in EQ VAS time 1-time 2 (MVH survey)

	No change[a]	Better[b]	Worse[b]
No change EQ-5D dimensions	73	59	49
Change on 1 dimension	17	30	23
Change on > 1 dimension	10	11	28
N	81	92	47

a. Change of 3 points or less
b. Change of more than 3 points

31 cases (14%) reported that they had a serious illness at time 1. Table 5.4 shows that these individuals were far more likely than the rest of the sample to experience a change in their level of reported problems on the EQ-5D dimensions between time 1 and time 2: 70% compared to 33% experienced a change on at least 1 dimension. Table 5.4 also shows that these sample members were also more likely to experience a change (both positive and negative) in their EQ VAS scores between time 1 and time 2. 97% of those with a serious illness experienced a change of more than 3 points on the EQ VAS compared with only 72% of those with no serious illness.

9 cases reported that they had experienced a new serious illness since they first completed EQ-5D at time 1. However, this number was too small for any meaningful analysis.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

more likely to experience changes on more than 1 dimension. In the MVH sample those who experienced no change in their VAS scores were much less likely to experience a change in the number of problems reported on the dimensions. Also, in contrast to the MVH sample, those who experienced a positive change in their VAS scores were more likely to experience a change on 1 or more of the dimensions than those who experienced a negative change in their scores.

Table 5.7. Percentage change on EQ-5D dimension between time 1 and time 2 by change in the EQ VAS time 1-time 2 (rheumatoid arthritis study)

	No change[a]	Better[b]	Worse[b]
No change EQ-5D dimensions	34	21	33
Change on 1dimension	45	37	33
Change on > 1dimension	21	42	34
N	58	90	73

a. Change of 3 points or less
b. Change of more than 3 points

At time 2, sample respondents were asked whether they felt their arthritis had become better, worse or stayed the same over the past 3 months. 93 individuals reported no change in their arthritis, 57 said that their arthritis had improved, and 73 reported that their arthritis had become worse. Table 5.8 presents the reliability coefficients (kappa and ICC) for EQ-5D for respondents who said that their arthritis had not changed. Compared to the coefficients for the whole sample reported in Table 5.9, the coefficients had improved, especially those for the usual activities and the pain/discomfort dimensions. The anxiety/depression dimension still exhibited the worst reliability.

Table 5.8. Reliability coefficients for sample members who reported no change in their arthritis over 3 months (rheumatoid arthritis study)

EQ-5D dimension	Reliability coefficient
Mobility	0.67 (0.09)[a]
Self-Care	0.71 (0.06)[b]
Usual Activities	0.70 (0.07)[b]
Pain/Discomfort	0.67 (0.09)[b]
Anxiety/Depression	0.43 (0.08)[b]
EQ VAS	0.82 (0.74-0.88)[b]

a. Kappa coefficient (standard error)
b. ICC (95% CI)

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

6
Construction of the EQ-net VAS and TTO databases

Tom Weijnen, Martin Nieuwenhuizen,
Arto Ohinmaa and Frank de Charro

6.1 INTRODUCTION

One of the objectives of the EQ-net project was to create an integrated database containing valuation data gathered previously in different European countries and in different studies. This database would enable secondary cross-national and comparative data analysis to be undertaken.

The EQ-5D offers a standardised framework to compare values for health states elicited with VAS and TTO methods in different European countries. This framework basically consists of the valuation of health states, which are defined by the descriptive system of the EQ-5D, together with a standardised set of questions on respondent characteristics.

The standard valuation method of the EuroQol Group has been to use the EQ-5D valuation questionnaire in a postal survey. With this questionnaire a set of 18 states shown in Table 6.1 can be valued. This set comprises 15 states: 13 unique health states defined by the EQ-5D, plus 'dead' and 'unconscious'. 11111, 33333 and dead are each valued twice.

For various reasons valuation studies have departed from the standard postal VAS method. For example, the set of states to be valued may have differed from the standard set, some studies have been interview-based rather than postal, and in some studies the respondent was asked to rank the health states before the VAS valuation.

This chapter describes the construction of the EQ-net databases. Separate databases were constructed for VAS and TTO valuations. 11 studies from 6 countries were included in the VAS database. The TTO database includes 3 studies. The main features of the 11 studies in the VAS and TTO databases are summarised in sections 6.2 - 6.4. Section 6.5 contains information on the background variables in the databases. Section 6.6 reports on the number of exclusions and inconsistencies. An analysis is undertaken on the shares of exclusions by respondent characteristics and sample features in section 6.7.

R. Brooks et al. (eds.),
The Measurement and Valuation of Health Status Using EQ-5D: A European Perspective, 55–79.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Finland

The data from Finland originates from the Finnish EuroQol valuation study (Finland). The Finnish valuation questionnaire was sent by post in November 1992 to a sample of 4000 persons over 16 years of age chosen randomly from the computerised population registry (Ohinmaa *et al*, 1996). The sample was divided into 17 sub-samples. These sub-samples each received a different questionnaire. The data included in the EQ-net VAS database were from 11 sub-samples totalling 2530 persons. After 2 reminders 1630 questionnaires were returned from the 11 sub-samples, a response rate of 64.4%. Some health states were common to every questionnaire, while some states were unique to various sub-samples. The 18 states, including unconscious, twice 11111, twice 33333, and twice death, were part of a total of 49 health states valued.

United Kingdom: Frome IV

The Frome IV study (UK_FR) was 1 of a series of studies carried out in a general practice population in Frome, UK (Gudex, 1992). The purpose of the study was to see whether an extra dimension of energy/tiredness should be included in the EQ-5D descriptive system. This postal survey was carried out in June-August 1991. Questionnaires were sent out to people chosen randomly from the computerised register of the Frome general practice. Respondents were required to be aged 16 years and over. Five different versions of the questionnaire were posted to 340 people, who received each version of the questionnaire. Questionnaires differed with respect to the wording of the mobility dimension and with respect to the number of dimensions. In the EQ-net VAS database, data from the standard version of the questionnaire, with the 18 states, were included. From the 340 questionnaires, 122 (35.9 %) were returned. No reminders were sent.

United Kingdom: MVH

The second UK data set (UK_MVH) included in the EQ-net VAS database is from the MVH study (Dolan, 1995). This was a representative sample of the non-institutionalised adult population of England, Scotland and Wales (Williams, 1995). Anticipating a response rate of 55% and a sample size of 3235, 6080 addresses were selected from the Postcode Address file (Gudex *et al*, 1997). Of the 6080 addresses, 756 were outside the scope of the survey, being non-residential, empty/derelict, untraceable or even not yet built. This resulted in a total of 5324 addresses. 3395 interviews were completed, a response rate of 63.7%. Trained interviewers conducted the interviews between August and November 1993. Each interview comprised: completion of page 2 (the descriptive system), ranking of states, VAS rating of states, TTO rating of states and provision of personal background data.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Respondents first ranked 15 health states and then valued these states. Of these 15 health states, 13 states were randomly chosen from a subset of 43 EQ-5D health states used in the UK MVH study. Every respondent valued the states 11111, 33333, unconscious, and dead.

6.5 BACKGROUND VARIABLES

In this section a description is provided of the background variables in the integrated database. The background variables of the 3 TTO-based studies are the same as those in the EQ-net VAS database and in the EQ-net TTO database because the data on these variables came from the same questionnaires. The background variables can be divided into sample features and respondent characteristics. The sample features refer to attributes of the sample as a whole. These features were added to the database on the basis of the information in the publications on the 11 valuation studies in the EQ-net database. Respondent characteristics describe the attributes of the respondents. Data on these characteristics were elicited from the questionnaires returned in the different samples. The standard EQ-5D questionnaire contains a set of background variables. However, for various reasons, the questionnaires in many valuation studies departed from the standard set. The respondent characteristics also include the variables describing the respondent's scores on the EQ-5D descriptive system and the rating of the respondent's own health on the EQ VAS.

Sample features

The following variables describe the sample features of the EQ-net databases. Table 6.2 summarises the sample features for each study in the database.

Time of data collection: The time of data collection is represented by 2 variables indicating the year and season of the main data collection.

Sample method: This refers to the postal versus interview-based samples.

RIT variable: The ranked, interview-based study protocol was originally developed for the UK MVH study where the respondents carried out VAS and TTO valuations of health states. The respondent ranked the health states before the EQ-5D VAS valuation task. In the VAS and TTO valuation tasks the health states were assumed to have a duration of 10 years, whereas on the standard EQ-5D VAS valuation questionnaire the duration is 1 year. The RIT variable (Ranked, Interview-based, Ten years duration) is a binary variable which indicates whether the data comes from 1 of the 3 interview-based TTO-studies.

Participation of interviewer: When an interviewer was present a distinction was made between an active and a passive interviewer. An active interviewer was an inter-

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

this was categorised as a low level of education. Other types of schools or education the respondent had finished were classified as medium level education.

Filling in the questionnaire: The standard questionnaire includes a question about the difficulty respondents experienced with filling in the questionnaire, which can be answered by ticking the boxes labelled 'easy', 'fairly difficult' or 'difficult'. The original information was re-coded into a binary variable with the categories 'easy' and 'difficult'. The variable 'difficult' consists of the original categories 'fairly difficult' and 'difficult'.

Own health and EQ VAS: The respondent characteristics also include the variables describing own health of the respondent. These variables are the scores on the EQ-5D descriptive system (mobility, self-care, usual activities, pain/discomfort, anxiety/depression), and the rating of own health on the EQ VAS. A distinction was made between those respondents with a median score of ≥ 85 and those with a score of < 85.

The analysis of the respondent characteristics in the VAS database shows that there were at least some variations in every background variable between the studies (Table 6.3). The same is also true of the TTO database (Table 6.4). The analysis in subsequent chapters handles these variations using a variety of statistical approaches.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

6.6 EXCLUSIONS AND INCONSISTENCIES

The criteria for exclusions were explained in section 6.2. Table 6.5 shows the number of exclusions and inconsistencies in the EQ-net VAS database. The database consists of records from 8709 respondents spread over 11 different studies. 897 respondents were excluded by applying criteria (i) (all states given the same value) and (ii) (less than 3 states valued). 501 respondents returned questionnaires that contained more than 3 inconsistencies (criterion (iii)). Exclusions were high in the samples from Finland (29.4%), Germany 1994 (51.1%), the Netherlands (40.5%), and Sweden (33.4%). Exclusions were absent in the Catalonian Cat_Cor and the German TTO samples. Exclusions were low in the Catalonian Cat_Hos data (0.5%), the UK MVH data (1.5%), and the Spanish Navarra data (11.0%) These were all interview-based samples. The low number of exclusions in the l'Hospitalet del Llobregat study has to be interpreted with caution, because respondents were selected *a priori*.

Table 6.6 shows the exclusions and inconsistencies in the EQ-net TTO database. Exclusions as a result of incomplete or unreliable data were small. Only 1.3% of the respondents were excluded on this basis. There were no respondents in the German data who valued less than 3 states. Total exclusions in the TTO database were 9.3%. The Catalonian data had the lowest share of total exclusions (5.7%).

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

6.7 EFFECT OF EXCLUSIONS ON RESPONDENT CHARACTERISTICS

Table 6.7 presents the effects of the exclusions on the respondent characteristics and on the missing values in the VAS database. For all the respondent characteristics, except the variables smoker, working status and employed, the exclusions were high in the missing values. A Chi-Square test for independence was conducted. For all the background variables, except experience of serious illness in the family (Exp. illness family) independence between the background variables and exclusions had to be rejected ($p < 0.000$). Males, older people, respondents who experienced serious illness themselves or in caring for others, those with an EQ VAS < 85, non-smokers, unemployed respondents, those who found it difficult to fill in the questionnaire and, finally, those respondents with a low level of education, all had a higher share of exclusions.

Table 6.7. Respondent characteristics by exclusions in the EQ-net VAS databases

		N			%	
		Not excl.	Excluded	Total	Not excl.	Excluded
Sex	Male	3352	694	4046	82.8	17.2
	Female	3948	575	4523	87.3	12.7
	Total	7300	1269	8569	85.2	14.8
	Missing	11	129	140	7.9	92.1
Age	≤ 24	825	42	867	95.2	4.8
7 groups	25-34	1595	76	1671	95.5	4.5
	35-44	1204	111	1315	91.6	8.4
	45-54	1105	165	1270	87.0	13.0
	55-64	1027	223	1250	82.2	17.8
	65-74	1016	371	1387	73.3	26.7
	75+	527	274	801	65.8	34.2
	Total	7299	1262	8561	85.3	14.7
	Missing	12	136	148	8.1	91.9
Exp. illness yourself	No	4713	430	5143	91.6	8.4
	Yes	1771	390	2161	82.0	18.0
	Total	6484	820	7304	88.8	11.2
	Missing	827	578	1405	58.9	41.1
Exp. illness family	No	2451	268	2719	90.1	9.9
	Yes	3965	395	4360	90.9	9.1
	Total	6416	663	7079	90.6	9.4
	Missing	895	735	1630	54.9	45.1

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

The EQ-net VAS 18 state database is used in Chapter 7 in the analysis of the influence of the background variables on the values for these states. In Chapter 8, the extended EQ-net VAS database is used in the estimation of a European set of weights for the EQ-5D. The EQ-net TTO database is studied in Chapter 9.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

effect of sex is uncertain, because a weak effect was found in an earlier study and therefore sex was not included (Björk and Norinder 1999).

Methods of comparing valuations

The analysis is based on 18 VAS valuations of EQ-5D states. Only 4 of the 11 studies provide information on all 18 valuations. Mean and median VAS values were computed for each study and over all the studies.

The relationship between mean VAS valuations and key background variables was first examined graphically and separately for each variable. To explore in a more rigorous way whether observed differences in valuations could be explained by study and respondent characteristics, a weighted linear regression analysis was carried out for each state with the valuation of that state as the dependent variable. To correct for the differences in sample sizes, weights (defined as 1/n, where n = the number of respondents in a study after exclusions) were applied. When only 1 set of values for 11111, 33333 and Dead (11111A, 33333A, DeadA) was included, a total of 15 regressions was performed. First, the respondent characteristics and the study feature RIT were entered as explanatory variables. Then the dummies for the standard studies were entered to see whether, after controlling for respondent characteristics and RIT, the different studies had independent explanatory power. Sweden was used in the regression analyses as a comparator or 'norm', because it was closest to the average study.

In order to investigate the similarities in the VAS valuations from the 11 studies, a singular value decomposition (SVD) was carried out on the median VAS scores. This examines to what extent the median scores of the studies can be represented by common scores, except for random error. It is an exploratory technique, looking for the best representation of the median scores for each study, using the common score and a study-specific proportionality constant. This constant shows to what extent (a given percentage) the common scores should be increased or decreased to represent the valuations for a particular study.

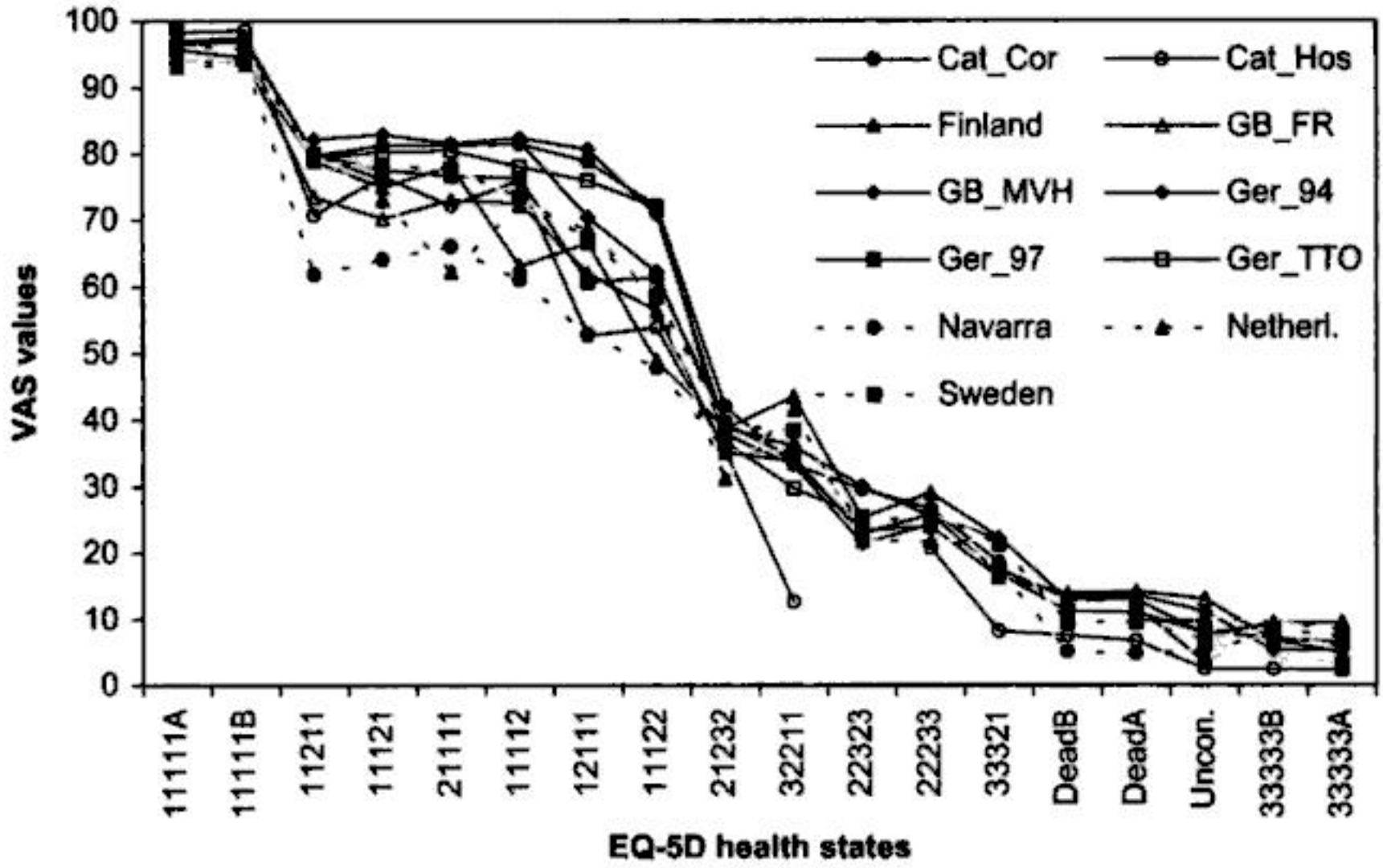
If the set of common scores and the study constants do not represent the matrix of scores satisfactorily, the technique shows whether a second component can explain the residuals. This could be the case if participants in some studies (for example, with a particular cultural background) use different valuation strategies than those in other studies (with other cultural backgrounds). In principle, 11 components will always result in a perfect solution. However, it is expected that fewer components will be necessary to represent the data. The first component is, in general, similar to the average values of health state scores (median values averaged over the studies), but not equal to these values. The values on this component are weighted means of the study scores, the weights depending on how well a study is represented by the common

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Figure 7.1 shows that differences between the total mean and median values emerged at the lower end of the VAS (severe states). The mean VAS values of the moderate health states with 1 or 2 dimensions at level 2 and the other dimensions at level 1 varied considerably across studies. For state 32211 the Cat_Hos study was clearly an outlier



	mean	median	SD		mean	median	SD
11111A	97.1	100	7.4	32211	33.6	30	21.1
11111B	95.3	100	8.9	22323	26.1	25	18.0
11211	77.7	80	15.7	22233	24.9	20	16.5
11121	77.6	80	15.9	33321	18.5	15	15.9
21111	76.4	80	16.1	DeadB	10.3	4	18.0
11112	75	80	18.0	DeadA	10.0	1	18.5
12111	69.1	70	20.0	Uncon.	8.2	2	15.5
11122	59.0	60	18.9	33333B	7.5	3	13.0
21232	37.3	35	18.2	33333A	5.9	2	11.1

Figure 7.1. Mean values of standard states by study and overall

Figure 7.1 does, however, suggest that there was a common pattern in mean valuations across studies, although there were also systematic differences. A large proportion of the variation between studies was evidently due to RIT. The valuations from these studies appeared to be higher, especially for the less serious states. This can be seen clearly from Figure 7.2

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Table 7.3. Weighted linear regression of VAS values on background variables for all studies (coefficients with $p < 0.01$ depicted)

	11111A	11211	11121	21111	11112	12111	11122	21232	32211	22323	22233	33321	DeadA	Uncon.	33333A
(Constant)	97.65	72.11	74.03	71.39	73.41	60.68	55.56	35.59	32.72	22.36	21.76	16.85	9.13	8.44	6.05
RIT study	2.82	4.79	7.17	7.72	7.51	16.02	14.88	2.66		5.23					-2.83
Age 25 - 34	-0.96		-2.60												
Age 35 - 44		2.89													
Age 45 - 54		5.26	3.81	6.21	5.03	4.18	4.89							-2.55	
Age 55 - 64		6.60	4.02	4.86	4.42	4.46	5.57	4.91			4.46			-3.00	1.92
Age 65 - 74	-1.14	5.75		4.26	3.44					7.46	5.08			-2.85	2.34
Age ≥ 75	-1.63	6.72									5.42				2.18
EQ VAS < 85	-2.74	-5.08	-4.51	-4.87	-4.96	-5.32	-3.81		-3.02						
Edu_med		3.75	2.18	2.17		3.31	294	597	1150	95	3344	181	92	339	267
Edu_high				2.29	-2.23	3.08			2.62			2.99	1.95		
R^2	0.07	0.06	0.07	0.08	0.07	0.13	0.09	0.02	0.02	0.05	0.02	0.01	0.01	0.02	0.02
Adjusted R^2	0.07	0.06	0.06	0.08	0.07	0.13	0.09	0.02	0.02	0.04	0.02	0.01	0.00	0.01	0.02
F	54.29	24.35	27.52	34.54	31.36	58.15	33.91	7.17	6.66	12.25	6.62	4.61	3.62	9.75	17.83
P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Table 7.6 shows that a representation using only the first comman component does not show consistent deviations between the studies: only some isolated states have a substantial deviation from the value predicted by the first left and right eigenvector

Table 7.6. Absolute differences between the original scores and the representation by the first component

Study	Cat_Cor	Cat_Hos	Finland	UK_FR	UK_MVH	Ger_94	Ger_97	Ger_TTO	Navarra	Netherl.	Sweden	Mean State
11211	1.21	3.92	3.00	1.65	0.22	0.23	1.21	0.11	9.44	4.32	0.46	2.34
11111A	6.15	6.35	1.45	2.27	7.41	1.05	0.19	5.29	12.03	1.63	0.77	4.05
21232	5.03	0.26	2.27	1.42	2.45	3.05	2.42	4.65	6.87	6.10	2.05	3.32
11122	3.44	3.72	11.17	0.65	7.65	7.86	2.42	8.98	5.16	0.11	3.19	4.30
11121	0.98	5.88	1.79	6.56	0.02	0.01	1.00	1.66	4.63	0.88	0.24	2.15
22233	0.76	1.38	5.73	2.09	0.52	2.57	2.21	4.04	7.91	1.89	3.01	2.92
33333A	1.40	1.23	3.72	0.77	0.59	1.33	1.31	1.39	1.16	3.74	0.67	1.57
33321	2.13	10.11	4.26	0.61	2.67	1.31	0.11	1.99	0.80	4.53	3.74	2.93
21111	1.00	0.89	2.80	2.45	0.00	0.03	1.01	1.68	0.38	10.87	0.26	1.94
11111B	6.10	6.40	1.49	2.31	7.36	1.01	0.23	5.24	12.07	1.17	0.72	4.01
Uncon.	1.39	1.22	1.28	3.73	3.60	1.32	1.31	0.62	1.15	0.75	1.32	1.61
12111	6.76	14.62	2.69	1.75	10.88	0.27	8.87	7.35	10.70	3.85	0.47	6.20
11112	1.72	6.53	11.54	0.89	2.73	5.72	1.69	1.89	9.02	0.22	0.06	3.82
32211	5.17	21.03	7.68	1.95	0.59	0.52	3.07	4.88	10.85	8.24	6.61	6.42
33333B	0.71	3.27	1.59	1.12	0.75	3.53	1.51	0.68	1.93	1.65	1.48	1.66
22323	7.34	5.01	0.82	0.65	5.07	1.57	1.30	2.53	16.28	0.46	1.51	3.87
DeadA	1.34	4.77	0.37	4.66	3.71	1.51	1.56	2.63	3.04	2.31	1.48	2.49
DeadB	0.37	5.14	1.02	1.02	0.42	1.66	3.39	0.34	3.62	0.05	2.15	1.74
Mean	2.94	5.65	3.59	2.03	3.15	3.15	1.93	3.11	6.50	2.93	1.68	3.19

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Our analyses show that the study features and respondent characteristics, as they were measured, had an effect on valuations, although the proportion of variation they explain remains very low. This suggests that we may not have been able to measure these characteristics properly, there may be important omitted explanatory variables, and/or that there is a large proportion of random variation. Many studies included a wider range of potential explanatory variables, but the set used was common to all studies. More stringent exclusion criteria than those applied would reduce the random variation and thus increase the proportion of variation explained, but maximising this proportion is not of primary importance in itself.

From the variables considered, the study characteristic referred to, as 'RIT' appears to have had the clearest and most consistent effect on valuations. This feature resulted in generally higher valuations than the standard approach. With respect to respondent characteristics, the respondent's own health status as rated on the EQ VAS exhibited a rather consistent effect so that poorer health status tended to be associated with lower valuations, other things being equal, especially where less serious health states are concerned. Age and education also seemed to have some effect: higher age tended to be associated with higher valuations, but for education there was no systematic pattern.

After controlling for these factors the origin of the study did not have a consistent effect on valuations with the exception of the studies from Spain (Cat_Hos and Navarra), which produced significantly lower valuations. This raises the question whether the Spanish values really differ or whether the results can be explained by some subtle features of Spanish studies (e.g. respondent selection) that are not reflected in our variables. One feature that might explain, at least partially, the results from Cat_Hos is related to an earlier translation of the questionnaire from the UK English used in the study. This indicates the importance of checking the translations carefully for international comparisons.

The SVD analyses show how similar the median valuations of the studies for the standard states are if the EQ-5D model restrictions described in Chapter 8 are omitted. These analyses suggest that all studies are rather well represented by the solution of the first component. However, the median values for studies differ in the multiplicative constant. For some studies, especially the RIT studies (UK_MVH, Cat_Cor and Ger_TTO), the common scores must be increased by a given percentage. For other studies (especially Navarra and Cat_Hos) the overall values must be reduced by a given percentage. For Navarra and Cat_Hos it was also found that there were (minor) inconsistencies, expressed by different factors. However, there was no indication that these differences were related to specific response strategies, nor were they of the same kind.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Table 8.1 presents the results from the transformation for the standard states based on linear regressions. The mean and median of the transformed values have been re-transformed using formula (2) for the reverse of the transformations. Transformation and re-transformation does not affect the median values, because the rank order between them does not change. The transformation is effective in reducing the skewness of the health states at both ends of the scale. For these health states the absolute differences between the mean and the median values become smaller. For dead and unconscious the transformed mean approaches the median values.

Table 8.1. Effects of the transformation procedures on the mean, median and measure for skewness (0 : a normal distribution; > 1 and < -1 : a skewed distribution)
Standard states only

	Raw values				Re-transformed values			
	Mean	Median	Difference mean-median	Skewness	Mean	Median	Difference Mean-median	Skewness
11111	97.5	100.0	-2.5	-4.3	99.4	100.0	-0.6	-1.4
11211	78.0	80.0	-2.0	-1.2	82.1	80.0	2.1	0.3
11121	77.8	80.0	-2.2	-1.3	81.7	80.0	1.7	-0.1
21111	76.7	80.0	-3.3	-1.4	80.2	80.0	0.2	-0.5
11112	75.3	80.0	-4.7	-1.2	79.5	80.0	-0.5	-0.3
12111	69.4	71.0	-1.6	-0.8	73.2	71.0	2.2	-0.6
11122	59.3	60.0	-0.7	-0.3	60.8	60.0	0.8	-0.3
21232	36.9	35.0	1.9	0.4	34.0	35.0	-1.0	-0.8
32211	32.9	30.0	2.9	0.4	27.4	30.0	-2.6	-0.8
22323	25.3	23.0	2.3	0.9	19.5	23.0	-3.5	-0.9
22233	24.3	20.0	4.3	1.0	19.4	20.0	-0.6	-1.1
33321	17.9	15.0	2.9	1.4	12.3	15.0	-2.7	-0.6
33333	5.3	2.0	3.3	4.1	1.8	2.0	-0.2	0.5
Dead	10.0	2.0	8.0	2.9	2.6	2.0	0.6	1.1
Uncon.	8.0	2.0	6.0	3.2	2.3	2.0	0.3	0.9
Total (absolute values)			48.6	24.8			19.7	10.3

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

VAS value for a particular health state, the RIT dummy variable always has the value of 1. In addition the coefficients expressing the interaction can be added to the coefficients of the N2 and N3 variables. Per health state, the dummies for N2 or N3 and their interaction dummies have the same value of 0 or 1.

If the estimated VAS value of, for example, the health state 12113 is calculated, the dummies in the model will have the following scores:

$$VAS_{12113} = (C + c_1.1) + (c_2.1 + c_4.1) + (c_3.1 + c_5.1) + c_6.0 + c_7.0 + c_8.1 + c_9.0 + c_{10}.0 + c_{11}.0 + c_{12}.0 + c_{13}.0 + c_{14}.0 + c_{15}.1$$

8.3 RESULTS

Influence of number of inconsistencies

Table 8.3 shows the number of health states per number of inconsistencies. When the maximum possible inconsistencies per respondent of 14 are allowed, the data consists of a total of 101583 health states. When no inconsistencies per respondent are allowed, 49% of the health states will be included. Beyond a level of 5 inconsistencies per respondent, the influence on the number of health states becomes small. The exclusion criterion of no more than 3 inconsistencies leads to the exclusion of 12.2% of the health states.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Table 8.5. Exclusions and inconsistencies (Continued)

	Ger_97 Germany 1997	Ger_TTO Germany TTO	Navarra	Netherl. Netherlands	Sweden
N	124	339	300	993	542
Exclusions due to incomplete data	19 (15.3%)	-	-	330 (33.2 %)	114 (21.0 %)
Due to more than 3 inconsistencies	11 (8.9%)	6 (1.8 %)	34 (11.3 %)	132 (13.3 %)	59 (10.9 %)
Number of respondents not excluded	94 (75.8%)	333 (98.2%)	266 (88.7%)	531 (53.5%)	369 (68.1%)

Background variables

Chapter 7 showed that some background variables have considerable influence on the VAS valuations. The sample characteristic 'RIT design study', and the respondent characteristics 'VAS score < 85', 'Age ≥ 65', 'Education medium and Education high', which had an influence on the VAS valuations, were tested for their impact on the regression results.

Table 8.6 shows the distribution of these respondent characteristics in the regressions. The table consists of 2 parts. The lower part describes the distribution of the background variables after exclusions. There was a clear effect of the exclusions on the distribution of background variables. The largest effects can be found in the data sets where exclusions were high. These were the samples from Finland, the German 1994 and 1997 samples, the Netherlands sample and the sample from Sweden. In the Finnish data, the effect of exclusions was especially high on the variable 'Age ≥ 65': the percentage of respondents older than 64 dropped from 35.64% to 21.00%. The same effect could be found in the other samples with a high number of exclusions. The percentages for 'VAS score < 85' also went down in the samples from Finland, Germany 1994, Germany 1997, the Netherlands, and Sweden. However, the percentages for medium level of education and high level of education rose in these samples due to the exclusions. The overall effects of the exclusions were that respondents were healthier and more educated, which may have been caused by the higher exclusions of respondents older than 64.

In Chapter 7 we reported the outcome of an OLS regression analysis with only the background variables as explanatory factors for the VAS valuations. We now report the outcomes of OLS analysis in which the EQ-5D dimensions and background variables are included. Results are given for the transformed, rescaled and transformed-rescaled VAS values.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Table 8.8. Results from OLS regression on the VAS values. Model without background variables (linear regression)

(Adjusted)	Raw VAS		Transformed VAS		Rescaled VAS		Transformed Rescaled VAS	
Constant	97.50	***	4.54	***	100	***	3.71	***
At least one 2 or 3	-17.48	***	-3.01	***	-21.00	***	-1.65	***
At least one 3	-15.33	***	-0.26	***	-19.15	***	-0.63	***
Mobility = 2	-5.57	***	-0.28	***	-5.58	***	-0.16	***
Mobility = 3	-16.69	***	-1.25	***	-18.72	***	-0.43	***
Self-Care = 2	-8.60	***	-0.43	***	-10.74	***	-0.35	***
Self-Care = 3	-12.61	***	-0.92	***	-14.87	***	-0.41	***
Usual Activities = 2	-2.97	***	-0.16	***	-2.49	***	-0.06	***
Usual Activities = 3	-7.44	***	-0.67	***	-7.87	***	-0.18	***
Pain/Discomfort = 2	-7.22	***	-0.33	***	-8.63	***	-0.26	***
Pain/Discomfort = 3	-12.60	***	-1.00	***	-14.72	***	-0.37	***
Anxiety/Depression = 2	-6.61	***	-0.38	***	-7.66	***	-0.27	***
Anxiety/Depression = 3	-10.17	***	-0.90	***	-11.74	***	-0.29	***
R^2	0.725		0.719		0.572		0.623	
Adjusted R^2	0.725		0.719		0.572		0.623	

*: $p < 0,05$. **: $p < 0,01$. ***: $p < 0,001$

Non-linear regression analysis

The 'CANALS' model (Gifi, 1990) from the SAS-programme 'TRANSREG' was used for the non-linear regression analysis In CANALS it is assumed that the nature of all variables is categorical, i.e. the 'score' on a particular variable is regarded as a classification into a category of that variable. This implies that the score is just a category name. Further restrictions can be added: the underlying measurement scale can be regarded as purely nominal (just a name, without any order between the categories), ordered, or metric (no transformation).

The essential feature of CANALS is that in one and the same analysis optimal transformations (with regard to the linear model) are found for the dependent and independent variables, together with the best fitting linear regression model on the transformed variables. The outcome is a solution of optimally transformed variables (independent as well as dependent) together with regression weights for the variables. Given the optimal transformations, the model is an ordinary (multiple) linear regression model.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

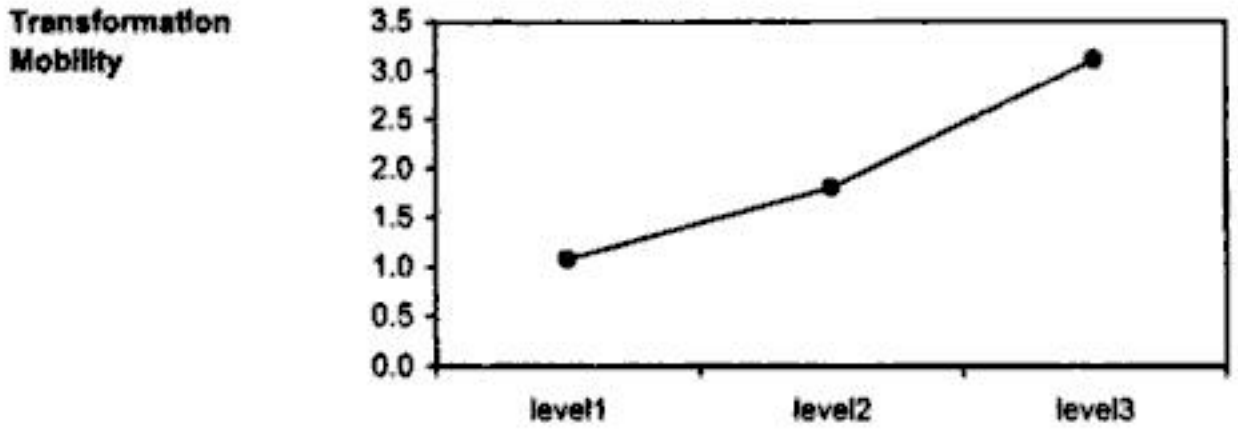


Figure 8.1. Transformations of the categories on the EQ-5D dimensions from CANALS analysis with linear restriction on the VAS values

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

In the second series of analyses, a dummy for the RIT-effect was added, together with 2 dummies for the interaction of RIT with the N2 and N3 effects. In previous analyses it was found that the addition of the interaction terms improved the regression solution significantly: the RIT effect in the higher region of VAS values was larger than in the lower region.

From the 3 analyses including the RIT-effects the constant fell in all analyses to the value found in the previous solution for level I+II+III, including the study effect. The variance component for studies fell considerably (from 7.01 to 4.13), indicating that the major difference between studies was with respect to the RIT-effect.

From the fixed effects RIT-parameters it can be seen that there was a general but moderate RIT-effect (1.96): RIT values were on average higher. The RIT*N2-effect shows that a considerable extra increase (8.30) is necessary if at least one 2 or 3 is scored on 1 dimension, but that an even bigger figure should be subtracted (-11.48) if a 3 is scored on at least 1 dimension. These effects combined suggest that the difference in constants is indeed multiplicative as found in the SVD-analysis reported in Chapter 7, instead of additive as assumed in an ordinary regression analysis.

If the first and second series of analyses are compared, the parameters for the EQ-5D dimensions are only slightly different. However, the parameters for N2 and N3 are much more influenced by the addition of RIT-effects. This can be explained by the addition of the interaction parameters for RIT and N2 and N3. The column with the aggregated coefficients presents the simplified I+II+III level model for RIT-studies, in which the coefficients for RIT and the constant were added together, and in which the coefficients expressing the interactions were added to the coefficients of the N2 and N3 variables.

The goodness-of-fit, as measured by R^2 was 0.745. All regression weights were significant at $p < 0.001$. For all the dimensions of EQ-5D, the weights for a score on level 2 are smaller than the weights for a score on level 3. The largest effect in the non-aggregated model is found for any deviation from full health, indicated by the N2 variable (-19.52). For the aggregated coefficients the N3 variable, which indicates at least one 3 on the descriptive system, together with the variable which indicates interaction of N3 with RIT, showed the largest influence (-20.06). When the first 3 appears on the descriptive system the predicted VAS value goes down by 33 points regardless of the scores on each different dimension.

Figure 8.4 and Annex 8.2 show the estimated VAS values for the observed health states, compared to the mean and the median values of these health states. The health states have been ordered according to the mean VAS values. For the mild health states the estimated values are very close to the median values and higher than the mean values; for the worst health states the estimated values are lower than the mean values. Figure 8.5 shows the difference between the observed mean and median val-

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

8.4 DISCUSSION

Data from respondents who valued all the health states the same or who valued less than 3 states were excluded from the analysis. These exclusions had an effect on the distribution of the background variables. Respondents aged over 65 had a higher probability of being excluded, as did respondents with a VAS value of less than 85, and respondents with low education. This had the effect that the data on which the regressions were performed was from respondents who were younger, healthier, and more educated. In the data from Finland and the Netherlands, 36 % of the respondents were older than 64. After exclusions, the share of respondents older than 65 dropped to 21 % for Finland and 24 % for the Netherlands. These shares were more in line with the overall share in the database of respondents aged over 65.

Because of the restriction of the range of the raw VAS values to between 0 and 100, skewed response distributions were expected, and found, at the end of the scale. This was expected to limit linear regression analysis because the values to be included in this kind of analysis are assumed to be normally distributed. In order to solve this problem, a logistic transformation was applied. The raw VAS values were transformed, with 100.5 and minus 0.5 as the upper and lower limits in the transformation procedure. These produced the lowest skewness. This transformation was effective in reducing the skewness of the values for the health states at both ends of the scale. For these health states the absolute difference between the mean and the median values became smaller. In particular, for dead and unconscious the transformed mean approached the median values. The regression on the transformed values, however, resulted in a much lower R^2 compared to the regression on the non-transformed data. The explanation is that although the transformation solved the problem of skewed response distributions, it also changed the nature of the linear additive model into a kind of multiplicative model. So it can be assumed that the underlying structure of the VAS values is indeed additive and that the transformation applied made the situation worse. It was therefore decided to apply a non-linear regression analysis in order to check this assumption. It emerged that the best transformation was linear over most of the scale, with exceptions at the ends of the scale. The improvement in the model (in terms of the R^2) was, however, small. Thus it was decided to maintain the use of the linear regression model.

In order to take individual and study differences other than the RIT-model into account, a multi-level (random effects) model was used. This revealed that 28.3% of the total error variance was due to individual differences and 1.6% was due to study differences. However, this had only a small effect on the parameter estimates in the model. The fixed parameters can be compared directly with the OLS-parameters. A weighted OLS model was used in which the weights expressed the differences in the total number of valuations per study. This revealed that the estimated values of the health states, calculated with the fixed parameters, did not deviate much from those calculated with the parameters of the weighted OLS regression. The average absolute

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Printed in the United States
54331LVS00002B/266