

TNO report

PG/Jeugd 2004.164

**An Internet Based Quality of Life Data Collection
Compared to a Traditional Mailed Survey**

Child Health
Wassenaarseweg 56
P.O. Box 2215
2301 CE Leiden
The Netherlands

www.tno.nl

T +31 71 518 18 18
F +31 71 518 19 15
info-Jeugd@pg.tno.nl

Date	July, 2004
Author(s)	A.G.C. Vogels S.B. Detmar G.H.W. Verrips J.A. Vogelaar
Number of pages	18

All rights reserved.

No part of this publication may be reproduced and/or published by print, photoprint, microfilm or any other means without the previous written consent of TNO.

In case this report was drafted on instructions, the rights and obligations of contracting parties are subject to either the Standard Conditions for Research Instructions given to TNO, or the relevant agreement concluded between the contracting parties. Submitting the report for inspection to parties who have a direct interest is permitted.

© 2004 TNO

Contents

1	Introduction	5
2	Methods	7
2.1	Questionnaire	7
2.2	Mail survey.....	7
2.3	Internet survey.....	7
2.4	Samples	8
2.5	Statistical analyses.....	8
3	Results	9
3.1	Samples compared.....	9
3.2	Scale structure	10
3.3	Differences in mean scores.....	12
4	Discussion and conclusions	15

1 Introduction

Internet based data collection is seen by many as an extremely important venue for the future of social and psychological research. Certainly, many arguments are in favour of this methodology: it allows for the possibility of reaching many people, probably at relatively low costs; people can provide the required data when it suits them, quietly in their home; it allows for reducing the number of missing and incorrect data and so on. At the same time, some doubts about its value seem justified. Depending on the way people are invited to participate in an internet based data collection: who are the people participating, what about people not honest about their identity or background characteristics, such as children pretending to be adults or the other way around. What about the possibility that people, while surfing the net for some fun, just decide to give it a try and give fake answers, out of curiosity about questions being asked or for whatever motive. Both screen and paper have their own restrictions on layout, but they are not the same; does this have effects on the way people react to the questions. Many internet based questionnaires do not offer the possibility to return to previous questions, to check or to correct. Does this have consequences?

There are several ways to collect data via the internet. To indicate a few: sending questionnaires via e-mail, collecting data using web forms, inviting visitors of specific sites to participate in a survey, linking them to a web form, building databases with possible respondents and inviting them –via e-mail or otherwise- to participate by filling in a web form and so on.

This study will compare a data set collected by means of a carefully designed and executed mail survey to a data set collected by means of an open, unsolicited internet survey where visitors of a website for parents were invited to participate. The data sets will be compared to determine the similarity in terms of demographic characteristics of the respondents, scale structure and mean scores. These comparisons were made in order to determine if such a simple internet based methodology is a suitable way to collect reference data for other HRQOL instruments.

2 Methods

2.1 Questionnaire

The TACQOL questionnaires are designed to assess Health Related Quality of Life (HRQOL) among children aged 6 to 15. HRQOL is conceptualized as (the absence of) health related functional problems, weighted by the affective evaluation of the child of such problems when they occur. Following this conceptualisation, first the parent or the child itself is asked whether a specific problem occurs; if so they are asked to what degree the child reacted negatively towards such problems. The answer on the functional problem question is combined with the answer on the emotional evaluation. Combined items are summed in order to calculate scale scores for each of five TACQOL scales: Bodily Complaints, Motor, Autonomy, Cognition and Social. Additionally Positive and Negative Emotions are assessed, based in simple questions on how often the child felt angry, satisfied and so on. Each scale consists of 8 items. Items belonging to specific scales are placed on the same page in the questionnaire. More detailed information was published elsewhere (ref: Manual).

For the purpose of the internet survey the questionnaire was converted to a web form. . During this conversion process, it proved to be impossible to present all 8 questions of a scale on a single computer screen. Therefore one question was asked at a time. The questions on emotional evaluation were shown only when the parent reported the actual occurrence of a specific problem (see figure 1, screen dump).

2.2 Mail survey

For the original mail survey a strategy was used that is known to provide highly representative samples and good quality data (ref naar peilingen en zo). Eight departments for regional preventive Youth Health Care (YHC) distributed all over the Netherlands participated in this data collection. They were asked to draw a sample of children aged between 6 and 12 years, stratified by age and gender. Samples were based on YHC registers, containing all children living in their region. All parents in the sample were sent a letter, by YHC, asking them to participate in the study, and to answer the questionnaire. Parents were asked to return the anonymous questionnaires directly to the researchers at TNO Prevention and Health. TNO Prevention and Health informed YHC about the registration codes of the questionnaires received. Non responding parents were sent a reminder once. Due to the complete registers of YHC and its high status as a medical institution among parents, a fair response rate could be achieved (71%) and the sample proved to be representative, apart from the fact that parents from non-Dutch origin were underrepresented. Details on data collection and sampling are presented elsewhere (Vogels et al 2000).

2.3 Internet survey

The TACQOL web-questionnaire was placed on the site of 'Ouders Online' ('Parents Online'), a site attracting many parents as visitors. On the site's opening page an invitation was placed, asking visitors with children between 0.5 and 11 years of age to participate in the study. Visitors doing so were linked to a page with information on the aim of the study. When they agreed to participate they were asked to answer the questions that would follow, keeping one of their children in mind. The first question

asked for the age of the child. This question was used as a screening question: when the age of the child fell between 6 and 11, they were linked to the TACQOL based web form; parents of younger children were linked to another web form. Parents with older children were informed that participation in the study was not possible. Visitors interested in the study's results were invited to send an e-mail to TNO Prevention and Health. Apart from that no token was offered. The questionnaires were available on this website for a period of 30 days. Three quarters of all respondents filled in the questionnaire in the first 4 days.

2.4 Samples

The original mail survey resulted in 1790 returned questionnaires. Due to missing data not all of these could be used in the analyses. For the current analysis 1741 respondents were available. More details of this group will be given when comparing the two samples.

In the internet survey 784 persons started the questionnaire. 69 of them quit after the first three questions (age, gender and the first health problem presented). Then with each question some more people quit. All respondents answering at least 12 of the last 16 items were included in the analyses (n=545). This criterion is similar to the treatment of missing data when calculating scale scores, where a maximum of 25% of missing data is allowed (Vogels et al, 2000).

2.5 Statistical analyses

Similarity of the samples was assessed using the following variables: who filled in the questionnaire (mother, foster mother, other female guardian, father, foster father, other male guardian, all bi-variate), origin of the father and the mother (Dutch, Western Non-Dutch, Non-Western, Unknown), level of education (no education/only primary education, secondary education, primary level, secondary education secondary level, higher education), does the child involved suffer from a chronic condition or long term disease, age of the child (in years: 6/7, 8/9, 10/11). Similarity was assessed first bi-variately, using Chi square. Then all variables with a significant Chi square were entered as predictors in a logistic regression analysis with the data collection methodology (Internet No/Yes) as the variable to predict.

To determine the similarity of the scale structure, first Cronbach's alphas were calculated for all scales, for the mail survey data set and the internet data set separately. Then, also separately for both sets, (adjusted) item scale correlation coefficients were calculated; the numbers of items with low adjusted item scale correlation coefficients (<.35) were counted, as were number of items with which were correlated more to an other scale score than the one to which it was supposed to belong.

To determine possible differences in mean scale scores Multivariate Analyses of Variance (Manova) were done with scale scores as dependent and data collection method as independent variables. As the comparison of the two samples showed some significant differences on many background characteristics, these analyse were repeated with these variables as factors too.

In testing a significance level of .05 was used,

3 Results

3.1 Samples compared

Table 1 shows the results of the comparison of the two samples. Bi-variate tests showed significant differences between the two data set on all variables, except for gender.

Table 1: A comparison between the two data sets; bi-variate chi2 and multivariate significance of Wald's index, Odds Ratios 95% and 95% Confidence Interval.

	Mail survey	Internet survey	Significance chi2; bi- variate ¹	Mutivariate log. reg.; significance overall effect ²	Odds ratio Internet vs mail ³	OR 95 CI
Child: gender						
Boy	50	54		Excluded		
Girl	50	46				
Child: Age group						
			.000	.000	Rc	
6/7 yrs	37	52			0.677	0.527-0.870
8/9 yrs	30	28			0.441	0.338-0.576
10/11 yrs	33	21				
Child: chronic condition/ prolonged disease						
Yes					1.809	1.421-2.304
Respondent (not exclusive)						
Mother	88	94	.000			
Foster mother	1	1	-			
Female carer	0	0	-			
Father	15	5	.000	.000	0.289	0.187-0.447
Foster father	0	0	-			
Male carer	0	0	-			
Unknown	0	1				
Father highest Educational level completed						
None/primary	6	4		.003		
Secondary, prim level	34	24				
Secondary, sec level	28	18				
Higher Education	28	53				
Unknown	3	1				
Father: origin						
Dutch	90	90	.000	.000	rc	
Western, non Dutch	1	5			3.498	1.842-6.643
Non Western	7	5				

Unknown	2	1			
Mother highest Educational level completed			.000	.000	
None/primary	6	1			Rc
Secondary, prim level	44	21			
Secondary, sec level	30	27		2.392	1.052-5.441
Higher Education	19	50		6.763	2.935-15.584
Unknown	1	1			
Mother: origin			.000	.000	
Dutch	92	95			Rc
Western, non Dutch	1	4			
Non Western	6	1		0.173	0.058-0.518
Unknown	1	1		10.731	2.003-57.493
Total	1741	545			

1. Only shown when < .05; -=not presented because of number of expected cell frequencies below 5
2. Only shown when < .05; excluded = excluded from logistic regression
3. Only shown when confidence interval does not include 1; Rc= reference category

Compared to the mail survey the Internet data set contained more children in the youngest age group (6 or 7 years) and more children with a chronic condition or long lasting disease. The Internet questionnaires were less often answered by fathers and more often by mothers. The level of education was higher in the Internet data set, both for fathers and for mothers. The percentage of parents born in the Netherlands is comparable, but the number of fathers and mothers from non Western origin (mainly Turkish and Moroccan) is even smaller in the internet data set than in the mail survey. The multivariate analyses showed that all these differences are independent from each other, except 'Mother as respondent'. Father's level of education had a significant overall effect, but the calculated odds ratios all included 1, also when other categories were used as reference for the calculation. The independent effect of this variable, therefore, may be neglected, probably because its strong association with educational level of the mother, which showed a very strong effect. The Odds Ratio for Origin of mother-unknown is large, but the large Confidence Interval in combination with the small absolute numbers indicates that this effect may be considered as doubtful.

3.2 Scale structure

Table 2 presents the Cronbach's alphas of the TACQOL-PF scales in the two data sets. Having a chronic disease or a long lasting disease is known to have an impact on the way the questionnaires are answered and this may have an impact on Cronbach's alpha. As the internet data set contained much more children with such conditions, Cronbach's alphas were calculated, both for the complete data sets and for those without a chronic condition in both data sets.

Table 2: TACQOL – PF scales Cronbach's alpha in the mail and internet survey; complete data sets and those without a chronic condition or long lasting disease.

	Complete data sets		Without chronic condition	
	Mail	Internet	Mail	Internet
Bodily complaints	0,70	0,62	0,69	0,58
Motor	0,79	0,81	0,78	0,74
Autonomy	0,70	0,80	0,66	0,74
Cognition	0,85	0,87	0,85	0,84
Social	0,67	0,73	0,63	0,69
Positive emotions	0,84	0,89	0,82	0,87
Negative emotions	0,71	0,69	0,70	0,66

The mean absolute difference for the complete data sets is 0.05; for those without a chronic condition 0.06. The biggest difference is found for Autonomy in the complete sets and Bodily Complaints in those without a chronic condition. The differences do not seem to be systematic; sometimes the indices are higher in the mail survey, sometimes it is the other way around.

Table 3 presents the number of items with an adjusted item-scale product moment correlation coefficient of less than .35 and the number of items correlating higher with an other scale than the one to which it is supposed to belong,

Table 3: Number of items with andjusted product moment correlation coefficient of less than .35 and the number of items correlated more to an other scal than its own, in the two data sets, for the seven TACQOL PF scales

	No of items with an ajusted PM Correlation coefficient less than .35		No of items loading highest on other than proper factor	
	Mail	Internet	Mail	Internet
Bodily complaints	0	5	0	1
Motor	0	0	1	2
Autonomy	3	0	1	1
Cognition	0	0	0	0
Social	2	3	1	3
Positive emotions	0	0	0	0
Negative emotions	0	2	0	1
Total	5	10	3	8

In both counts, the number of problematic items higher in the internet data set as in the mail data set. In the internet data set the distinction between the scales Bodily complaints, Motor and Autonomy seems to be less well clear than in the mail data set. The problem items in the scale Social show higher associations with Negative emotions; this is less a problem, as the two emotion scales were not supposed to be independent from the other scales (Ref: manual).

Table 4 presents the Product Moment Correlation Coefficients between scale scores for the two data sets. Scale in the internet data set were less independent, with a mean percentage of variance shared with an other scale of 15%, compared to 11% in the mail survey. Main problems are with the scales Motor, Autonomy and Social.

Table 4: Interscale correlation coefficients and mean (bi-variate) shared variance; mail and internet survey

	Motor	Autonomy	Cognition	Social	Positive Emotions	Negative Emotions	Mean shared variance
MAIL							
Bodily complaints							
Motor	0,38	0,26	0,26	0,27	0,26	0,31	0.09
Autonomy		0,53	0,31	0,33	0,35	0,25	0.14
Cognition			0,25	0,29	0,25	0,22	0.10
Social				0,32	0,29	0,29	0.08
Positive emotions					0,44	0,48	0.13
Negative emotions						0,39	0.11
							0.11
INTERNET							
Bodily complaints							
Motor	0,37	0,29	0,28	0,23	0,30	0,26	0.08
Autonomy		0,66	0,43	0,34	0,36	0,21	0.18
Cognition			0,48	0,42	0,38	0,23	0.19
Social				0,38	0,34	0,31	0.14
Positive emotions					0,59	0,48	0.18
Negative emotions						0,45	0.17
							0.12

3.3 Differences in mean scores

A Manova with data set as independent variable and the seven TACQOL-PF scales as dependent variables resulted in a significant multivariate effect (Wilks's Lambda < .001, partial eta squared = 0.11). Data set showed significant effects on each of the scales individually (all alpha's < .001), with the largest effects on Autonomy and Social (partial eta squared = .079, resp. .058).

Such effects may be caused by the differences between the two samples, presented in table 1. Therefore, a new Manova was done to determine which of these background characteristics on which the samples showed significant (multivariate) significances, had an impact on the scale scores. Significant multivariate effects were found for all these variables (significance Wilks's Lambda .013 or less), except for origin of the father. The background characteristics with a significant effect were then entered in a third MANOVA, together with method of data collection. In this analysis, too, data set had a significant multivariate effect (significance Wilks's Lambda < .001, partial eta squared = .085).

Table 5 presents the means in both samples. The table also presents the results of the bi-variate analysis of variance with internet as independent variable.

To facilitate the assessment of the relative importance of difference in data collection methodology, compared to the other variables in this analysis, adjusted R² are presented, one for the analysis with data collection methodology as well as background characteristics as independent factors and one for the analysis with only background characteristics as independent factor.

Table 5: Mean scale scores in the two data sets, results of bivariate ANOVAs with internet as independent factor, and explained variance of two models: a) only background characteristics as factor and b) background characteristics and data collection methodology as factor

Dependent Variable	Data set	Mean	F	Df	Significance	Adjusted R ² Without data set	Adjusted R ² With data set
PBOD	Mail	27.1	41.92	1	< .001	.067	.084
	Internet	25.7					
PMOT	Mail	30.8	39.52	1	< .001	.057	.073
	Internet	29.7					
PAUT	Mail	31.2	95.34	1	< .001	.051	0.89
	Internet	30.0					
PCOG	Mail	29.0	72.15	1	< .001	.036	.066
	Internet	27.7					
PSOC	Mail	29.9	150.65	1	< .001	.047	.107
	Internet	27.9					
PPOS	Mail	14.8	38.52	1	< .001	.054	.069
	Internet	14.2					
PNEG	Mail	11.5	32.13	1	< .001	.042	.056
	Internet	10.8					

The means on all scales scores are significantly lower in the internet data set. The differences between the two R²'s indicate that, compared to the other variables, data set is a very relevant predictor. Adding data set to the model increases the explaining power of the model for all scales; for Social, the explained variance is more than doubled.

4 Discussion and conclusions

This study compared two data sets, collected with the TACQOL Parent Form; one collected by means of a traditional and tested mail survey and one in which data were collected via the internet. In the last procedure visitors of the site Ouders Online (Parents Online) were invited to participate in the study.

The two sets differed markedly on demographic and other background characteristics of the responding parents and the children about whom they reported. Children in the internet data set were younger and suffered more often from chronic conditions. The responding parent was more often the mother, had a higher level of education, was more often of Western, but non Dutch origin and less often of non Western origin. The internet data set reflected the conceptualized data structure less well than the mail survey data set. Reported HRQOL was on average lower in the internet data set than in the mail survey. The effects of data collection methodology in this study were comparable or larger to the effects of chronic conditions, as reported in the manual (Vogels et al, 2000). The differences in background characteristics which were included in this study were insufficient to explain the differences in HRQOL scales.

How to explain these differences and what does it mean as for the suitability of this internet data collection procedure for collecting reference data for other HRQOL instruments?

As mentioned before, the layout of the pencil and paper questionnaire differed from that of the web form; in the former all items belonging to a specific scale were placed on one page whereas in the latter each questions appeared separately on the screen. The former layout may induce a specific mind set in the respondents, focussing them on a specific domain. This may explain the fact that scales in the internet data set were less independent of each other. However, it is an unlikely explanation for the differences in mean scale scores which we found. Furthermore, other TACQOL studies than the original reference study, also showed problems in the TACQOL scale structure (data collection among 12-15 year old children (to be published shortly and unpublished analyses on TACQOL scores in a sample of children born preterm or with a very low birth weight). So, probably, the different layout is not a major issue.

The differences in back ground characteristics between the two samples were, as such, not surprising. Even in a country such as the Netherlands with a high rate of internet connections, internet remains a relatively new technique, which has not yet reached all levels of the population. It is likely that ethnic minorities and lower social economic levels will be less often connected to the internet. What is more surprising is that the differences in background characteristics were proven insufficient to explain the differences in mean scale scores. This is not only surprising, it is troublesome too. More or less correct estimates of population indices can be constructed using biased samples, by weighting for relevant factors. This is only possible, however, when such factors are known and measured. The standard set which was included in this study is, overall, quite irrelevant in this respect. Due to the lack of theory building and theory testing in the field of HRQOL research it is difficult to define which other factors should be included to allow for an adequate correction or weighting of the data collected.

Clearly, parents visiting the site of Ouders Online are a quite selective sample from the population. One might guess that the site attracts parent who are more concerned about their child's well being or upbringing, or who are more conscious in their upbringing. These parents may, rightly or wrongly, more sensitive to possible problems confronting their child and therefore they may report their child having a lower HRQOL. If this is plausible, a solution may be found in trying to find more and other sites, attracting

visitors with other concerns, where invitations to participate in the study could be placed. Whether this is a sufficient solution, however, remains a question, as it seems very probable that on such sites, too, especially concerned parents will accept such an invitation.

Based on the results of this study we feel therefore that the fast and simple design of internet based data collection is not very suitable for building reliable reference data sets, representative for the population as a whole. Such a methodology may only be used if relevant factors affecting the relevant dependent variable are known, when they can be measured reliably and when the (multivariate) distribution of such factors in the population is known. Only under such conditions, probable differences between sample and population distributions can be corrected for.

In other situations more sophisticated methodologies must be used. One of these methodologies is using samples from registers with e-mail addresses and quality background data. Such registers, currently available at many marketing research institutes, offer the opportunity to invite people to participate, actively and allow for a better control of response and non response. Such a methodology could combine some of the advantages of traditional mail surveys with the advantages promised by computer and internet: large response, fast procedures, lack of item non response and, possibly, low cost.

References

Vogels, T, Verrips GHW, Koopman HM, Theunissen NCM, Fekkes M, Kamphuis RP. TACQOL Manual Parent Form and Child Form. Leiden: Leiden Center for Child Health and Pediatrics LUMC-TNO, 2000 LCCHP-01.