Community Dent Oral Epidemiol 2010; 37: 340–347

Community Dentistry and Oral Epidemiology

Assessing the level of agreement between the self- and interviewer-administered Child-OIDP

Eva Rosel[1], Georgios Tsakos[2], Eduardo Bernabé[2], Aubrey Sheiham[2] and Manuel Bravo[1]

[1]Department of Preventive and Community Dentistry, School of Dentistry, University of Granada, Granada, Spain, [2]Department of Epidemiology and Public Health. University College London, London, UK

Rosel E, Tsakos G, Bernabé E, Sheiham A, Bravo M. Assessing the level of agreement between the self- and interviewer-administered Child-OIDP. Community Dent Oral Epidemiol 2010; 38: 340–347. © 2010 John Wiley & Sons A/S

Abstract – *Objective:* To assess the level of agreement between the self- and interviewer-administered Child version of the Oral Impacts on Daily Performances (Child-OIDP) index. *Methods:* This was a randomised study in 177 children aged 10–13 years from Granada (Spain). All children completed both administration modes of the Child-OIDP; half the sample received the interviewer-administered version first (n = 90), and the other half the self-administered version first (n = 87). This was done to address potential order effects due to the sequential administration of both instruments. The level of agreement between both modes of administration was assessed with the Bland and Altman method for the Child-OIDP score and Kappa for the prevalence of oral impacts. *Results:* The two groups did not differ in their socio-demographic characteristics or self-perceived oral health measures. No order effects were found. There was no significant difference between the two modes of administration in terms of the overall score and prevalence of oral impacts ($P \geq 0.784$ in both cases). The mean difference in Child-OIDP scores was 0.03 (95% CI = −0.29 to 0.35) and the 95% limits of agreement were −6.32 and 4.93. Kappa value for the prevalence of impacts was 0.92. *Conclusions:* The self- and interviewer-administered Child-OIDP had a high level of agreement, irrespectively of whether the overall score or the prevalence of oral impacts was used to describe children's quality of life.

Key words: child; oral health; quality of life; random allocation

Manuel Bravo, Facultad de Odontología. Campus de Cartuja s/n, 18071 Granada, Spain
Tel.: +34 958243804
Fax: +34 958242949
e-mail: mbravo@ugr.es

Submitted 18 November 2008; accepted 23 December 2009

The Child Version of the Oral Impacts on Daily Performances (Child-OIDP) index was designed as an interviewer-administered oral health-related quality of life (OHRQoL) measure (1). Generally speaking, interviews are associated with higher response rates, but they are also time-consuming and relatively costly (2, 3). A self-administered Child-OIDP would be more practical and cost-effective than the originally designed face-to-face interview, provided that adequate response rates and understanding of the content of the questions were maintained. Self-administration would therefore facilitate the application of the Child-OIDP in both clinical practice and population epidemiological surveys. However, the most important question to be addressed before the wide implementation of a self-administered Child-OIDP is whether the data collected by the two modes of administration are comparable.

The first study comparing the OIDP index as interview and questionnaire found that their psychometric properties and scores were not related to the mode of administration (4, 5). That study was on adults and used the same layout for both modes. No changes were made in the interviewer-administered OIDP to facilitate its self-completion and participants followed the interview guidelines without assistance. A recent study in children compared psychometrically the self- and interviewer-administered Child-OIDP and showed that both modes of administration performed similarly, supporting the use of the self-administered version (6). However, that study was conducted in a sample of children referred for orthodontic treatment, with a high

doi: 10.1111/j.1600-0528.2010.00533.x

prevalence of oral impacts and thus may not be representative of the general child population where the prevalence of impacts is expected to be lower.

The level of agreement between two instruments is usually tested by collecting repeated data within the same group of individuals (3, 7). However, when two instruments are administered sequentially the order of their presentation may affect responses, a phenomenon known as order effect (8–13). Order effects may occur because participants may experience fatigue or lose concentration towards the end of the administration of the battery of instruments, which increases the probability of misinterpreting or failing to answer some questions. Participants may also produce different patterns of responses to those given to the first instrument because they may be desensitized to or familiarised with a given topic (12, 13). In both scenarios though, participants will score lower in the second instrument. This might be especially relevant when the same instrument is delivered under different modes of administration. In such studies evaluating the performance of alternative administration modes, the issue of order effect should be addressed. Therefore, it has been recommended to administer both instruments in a randomized order (11, 14) or in a counterbalanced manner (8–10, 12, 13) to prevent order effects during validation studies. Furthermore, considerable attention in the literature was given to establishing the most appropriate method to assess agreement between continuous measures (3). In this regard, the use of correlation coefficients to assess the level of agreement has been repeatedly criticised (15–18). Pearson's correlation coefficient is a measure of association but not of agreement, between two instruments. Therefore, a high correlation does not expose systematic bias that can occur when one method provides consistently higher or lower scores than the other (17, 18). Therefore, recommendations exist for more appropriate statistical evaluation methods (7, 15, 16, 19, 20). No previous studies have assessed the agreement between different modes of administration of the Child-OIDP index. Therefore, the objective of this study was to assess the level of agreement between ratings obtained with the self-and interviewer-administered Child-OIDP.

Methods

Study design

Five of the 21 primary schools in the Northern district of Granada city (Andalusia, Spain) were randomly selected. All children (n = 193) registered in the fifth and sixth grades (aged 10–13 years) were visited at school between April and June 2008. Parents of five children refused permission for their participation and seven children were not at school at the time of the study, yielding an initial study sample of 181 children. The self-administered questionnaire (SAQ) and face-to-face interview (FTFI) Child-OIDP were applied to each child. A computer-assisted block randomization method was used to allocate the order of their administration, thereby forming two groups: the FTFI-first group (Group A) and the SAQ-first group (Group B). Three children (1.7%) had missing items in the SAQ Child-OIDP and one child (0.6%) had missing items in both the SAQ and FTFI Child-OIDP. Therefore, they were excluded for analysis, leaving a final sample of 177 children (Group A, n = 90; Group B, n = 87). The sample size per group was established following the broad suggestion by Stewart et al. (21) that between 50 and 200 participants should be used for validation studies. That number is also in the range used in previous studies on children's OHRQoL (22).

The study was approved by the University of Granada Ethics Committee. Parents and teachers were sent information sheets explaining the purpose of the study and giving details on the school visit and parents were informed that they could refuse permission for the participation of their children.

Data collection

Socio-demographic variables were age, sex and socioeconomic status based on parental occupation (23). Three self-perceived oral health measures were also collected for validation purposes: self-perceived oral health status (5-point ordinal scale from very good to very bad); self-perceived oral health problems (3-point ordinal scale from almost none to many); and self-perceived dental treatment need (no/yes). Finally, OHRQoL was assessed by using the Child-OIDP index (1). The Child-OIDP index assesses the impact of oral health on daily life in relation to eight daily performances: eating, speaking, cleaning mouth, sleeping, emotion, smiling, studying and social contact. The frequency of the impact (scale from 0 to 3) and the severity of its effect (scale from 0 to 3) were scored (24), with zero score being assigned if no impact was reported. The performance score (range, 0–9) was estimated by multiplying the corresponding frequency and severity score. The overall Child-OIDP score was the sum

of the eight performance scores (ranging from 0 to 72) multiplied by 100 and divided by 72 to give a percentage score. Then, the prevalence of oral impacts was calculated as the percentage of children with a Child-OIDP score higher than zero.

The Spanish version of the Child-OIDP index was cross-culturally adapted according to published guidelines (25, 26) in 2006–2007 for two Spanish projects: the 2007 Child Oral Health Survey in Navarra (27) and the Evaluation in Andalusia of the Oral Health Capitation Program (see 'Acknowledgements' section), with the validation of the SAQ Child-OIDP for adolescents having been published (28). To facilitate self-completion, some technical terms were simplified and the layout of the questionnaire was made more user-friendly, with clear guiding instructions. The SAQ Child-OIDP was administered to the children as a group in their classroom, supervised by a researcher who addressed queries from the children. For the FTFI Child-OIDP, children were interviewed individually in a separate room. Both instruments were administered on the same day.

Data analysis

The two randomized groups A and B were initially compared according to their socio-demographic characteristics and self-perceived oral health measures. Groups were compared by age, socio-economic status and self-perceived oral health status and oral health problems using the Mann–Whitney test and by sex and self-perceived dental treatment needs using the chi-squared test.

Psychometric properties of the Spanish version of the Child-OIDP were assessed in terms of construct validity, internal consistency and test-retest reliability. For construct validity, SAQ and FTFI Child-OIDP scores were compared between groups with different oral health statuses, as defined by self-perceived oral health, oral health problems and dental treatment needs. Internal consistency of the SAQ and FTFI questionnaires was assessed by the Cronbach's alpha coefficient. Test-retest reliability was assessed by calculating the intraclass correlation coefficient (ICC) for the repeated administration of each instrument in a sub-sample of the participants.

The effect of the order of administration of both instruments was tested by considering four data sets (two sets for each child): set A_{SAQ}, SAQ Child-OIDP scores of group A (n = 90); set B_{SAQ}, SAQ Child-OIDP scores of group B (n = 87); set A_{FTFI}, FTFI Child-OIDP scores of group A (n = 90); and set B_{FTFI}, FTFI Child-OIDP of group B (n = 87). Child-OIDP scores and prevalences of impacts were compared within groups (A_{SAQ} versus A_{FTFI} and B_{SAQ} versus B_{FTFI}) using the Wilcoxon signed-rank test for scores and chi-square test for prevalences and between groups (A_{SAQ} versus B_{FTFI} and A_{SAQ} versus B_{FTFI}) using the Mann–Whitney test for scores and McNemar test for prevalences.

After demonstrating no order effects (see 'Results'), differences in performances and overall scores between the SAQ and FTFI Child-OIDP were assessed in the whole sample (n = 177) with the Wilcoxon signed-rank test. The prevalence of oral impacts was compared between both instruments using the McNemar chi-square test. Throughout the manuscript, we have reported means and SDs of the Child-OIDP score in the tables for ease of understanding, however we have used nonparametric tests for the comparisons between groups as the Child-OIDP scores were not normally distributed.

The level of agreement between the SAQ and FTFI Child-OIDP scores was established using the Bland and Altman method (15, 16, 19) for continuous scales. First, the unknown 'true' Child-OIDP scores were estimated by determining the average of both modes of administration. Then, differences between SAQ and FTFI modes were calculated for each child to estimate the systematic bias. The SD of such differences was calculated, to estimate the random error. Next, differences between the two instruments were plotted against their average. Because differences between the SAQ and FTFI Child-OIDP were not normally distributed, the non-parametric version of the Bland and Altman method was used (16), calculating the 2.5th and 97.5th percentiles to estimate the 95% limits of agreement between SAQ and FTFI Child-OIDP scores. Finally, the level of agreement between both modes of administration was also estimated for the overall prevalence of oral impacts through the Kappa statistic.

Results

One hundred and seventy-seven children, 85 males (48.0%) and 92 females (52.0%), with a mean (±SD) age of 11.1 ± 0.8 years participated in the study. There were no statistically significant differences between groups A (FTFI Child-OIDP first) and B (SAQ Child-OIDP first) in terms of their socio-demographic characteristics or self-perceived oral health statuses (Table 1).

Table 1. Comparison of socio-demographic characteristics and self-perceived measures between the two halves of the study population (n = 177); group A, who completed the FTFI[a] Child-OIDP first followed by the SAQ[b] Child-OIDP and group B, who completed the SAQ Child-OIDP first followed by the FTFI Child-OIDP

Variable	Group A (n = 90)	Group B (n = 87)	P-value
Age in years (mean ± sd)	11.2 ± 0.8	11.0 ± 0.7	0.179[c]
Range	10–13	10–13	
Sex (% male)	53.3	42.5	0.198[d]
Socio-economic status (%)			
High	3.5	2.4	0.655[c]
Medium–high	9.3	12.9	
Medium	26.8	18.8	
Medium–low	52.3	57.7	
Low	8.1	8.2	
No response (n)	4	2	
Self-perceived oral health status (%)			
Very good	35.0	31.2	0.686[c]
Good	40.0	42.9	
No good not bad	21.2	22.0	
Bad	2.5	2.6	
Very bad	1.3	1.3	
No response (n)	10	10	
Self-perceived oral health problems (%)			
Almost none	46.7	47.4	0.421[c]
Few	33.8	44.9	
Many	19.5	7.7	
No response (n)	13	9	
Self-perceived dental treatment need (%)			
No	46.3	50.0	0.752[d]
Yes	53.7	50.0	
No response (n)	10	7	

[a]FTFI: face-to-face interview.
[b]SAQ: self-administered questionnaire.
[c]Mann–Whitney test.
[d]Chi-square test (with continuity correction).

Both the SAQ and FTFI Child-OIDP had acceptable psychometric properties. For construct validity, the SAQ and FTFI questionnaires were able to discriminate between groups defined by three self-perceived oral health measures (Table 2). Child-OIDP scores significantly increased with each higher level of self-perceived oral health status, oral health problems and dental treatment needs. Besides, all three self-perceived measures were positively correlated with Child-OIDP score, both in SAQ (correlations ranged from 0.22 to 0.27) and FTFI administration mode (from 0.19 to 0.23). For internal consistency, Cronbach's alpha values were 0.53 for both the SAQ and FTFI Child-OIDP. Finally, a randomly selected sample of 22 children completed the SAQ and FTFI Child-OIDP at baseline and again 1 week later for test-retest reliability; the mean (±SD) Child-OIDP scores for SAQ and FTFI, were respectively, 5.81 ± 8.36 and 5.24 ± 7.88 at baseline and 5.37 ± 7.93 and 4.74 ± 7.66 after 1 week. The intraclass correlation coefficient was 0.98 for SAQ and 0.97 for FTFI Child-OIDP.

Table 3 shows the comparisons in performance and overall Child-OIDP scores between and within the data sets. No significant differences between groups A and B were found in overall SAQ Child-OIDP scores or performance scores (A_{SAQ} versus B_{SAQ}) or in overall FTFI Child-OIDP scores or performance scores (A_{FTFI} versus B_{FTFI}). Likewise, no statistically significant differences in performance or overall Child-OIDP scores were found when the two modes of administration were compared in the same children, i.e. within group A (A_{SAQ} versus A_{FTFI}) and within group B (B_{SAQ} versus B_{FTFI}). Similar findings were obtained when the prevalence of oral impacts was used instead of Child-OIDP scores to describe OHRQoL (results not shown).

Following the above demonstration of no order effect, SAQ and FTFI Child-OIDP scores and prevalences of oral impacts were compared in the whole sample (n = 177). The overall SAQ Child-OIDP (mean ± SD) was 4.33 ± 6.41 and 4.30 ± 6.27 for FTFI Child-OIDP, with identical distributions in terms of quartiles (Q1 = 0.0, Q2 = 1.39 and Q3 = 6.94). There was no significant difference between the overall SAQ Child-OIDP and FTFI Child-OIDP scores or in any of the eight performances scores between the two administration modes (Table 4). Likewise, there was no significant difference in the overall prevalence of oral impacts between the SAQ and FTFI Child-OIDP (55.4% and 55.9% respectively) or in the prevalences of impacts for each performance. Differences between overall SAQ and FTFI Child-OIDP scores were plotted against mean scores following the non-parametric Bland and Altman approach (Fig. 1). The median difference was 0.00 (95% limits of agreement: −6.32, +4.93). In addition, the level of agreement between the SAQ and FTFI Child-OIDP for the overall prevalence of oral impacts was high (Kappa = 0.92).

Discussion

This study assessed the level of agreement between the SAQ and FTFI administration modes of the Child-OIDP index. Due to the sequential administration of both instruments, children were randomized to two possible orders of presentation (SAQ first and FTFI second or vice versa). After

Table 2. Association of self-perceived measures with both Child-OIDP modes of administration and comparison of overall Child-OIDP scores for each category of self-perceived measures (n = 177)

Variable	n[a]	SAQ[b] Mean ± SD	SAQ[b] rho[d]	FTFI[c] Mean ± sd	FTFI[c] rho[d]
Self-perceived oral health status (%)					
Very good	52	2.78 ± 4.48	0.22[e]	3.21 ± 4.81	0.19[e]
Good	65	3.78 ± 6.58		3.46 ± 6.06	
No good not bad/bad/very bad[e]	40	6.42 ± 7.65	0.007	6.46 ± 7.90	0.017
(P-value)		0.032[f]		0.040[f]	
Self-perceived oral health problems (%)					
Almost none	73	2.55 ± 4.85	0.27	2.78 ± 4.90	0.23
Few	61	4.83 ± 6.76		4.51 ± 6.85	
Many	21	7.47 ± 8.43	0.001	7.74 ± 7.75	0.003
(P-value)		0.002[f]		0.004[f]	
Self-perceived dental treatment need (%)					
No	77	2.74 ± 5.47	0.26	2.78 ± 5.41	0.23
Yes	83	5.37 ± 6.81		5.37 ± 6.74	
(P-value)		0.001[g]	0.001	0.004[g]	0.004

[a]For 'No response' see Table 1.
[b]SAQ: self-administered questionnaire.
[c]FTFI: face-to-face interview.
[d]Spearman's rank correlation.
[e]Calculated considering all original categories.
[f]Kruskal–wallis test was used for comparison.
[g]Mann–Whitney test was used for comparison.

Table 3. Comparison of performances and overall scores depending on the version administered and the order of administration of the Child-OIDP[a]

Performances and overall score	SAQ version A_{SAQ} (n = 90) Mean ± SD	SAQ version B_{SAQ} (n = 87) Mean ± sd	FTFI version A_{FTFI} (n = 90) Mean ± SD	FTFI version B_{FTFI} (n = 87) Mean ± sd	Comparison (P-values) A_{SAQ} versus B_{SAQ}[b]	A_{FTFI} versus B_{FTFI}[b]	A_{SAQ} versus A_{FTFI}[c]	B_{SAQ} *vs* B_{FTFI}[c]
1. Eating	0.73 ± 1.36	0.79 ± 1.45	0.77 ± 1.41	0.84 ± 1.52	0.906	0.925	0.691	0.670
2. Speaking	0.36 ± 1.21	0.20 ± 0.78	0.37 ± 1.21	0.26 ± 0.99	0.606	0.654	0.317	0.361
3. Cleaning mouth	0.37 ± 1.14	0.24 ± 0.86	0.31 ± 0.93	0.20 ± 0.52	0.652	0.814	0.201	0.500
4. Sleeping	0.21 ± 1.02	0.26 ± 0.98	0.27 ± 1.11	0.24 ± 0.98	0.901	0.504	0.144	0.180
5. Emotion	0.44 ± 1.37	0.22 ± 0.69	0.37 ± 1.29	0.23 ± 0.71	0.493	0.843	0.109	0.317
6. Smiling	0.81 ± 2.03	1.13 ± 2.20	0.80 ± 1.91	1.11 ± 2.16	0.235	0.302	0.800	≈1
7. Studying	0.10 ± 0.48	0.07 ± 0.30	0.11 ± 0.48	0.07 ± 0.30	0.977	0.784	0.317	≈1
8. Social contact	0.19 ± 0.81	0.11 ± 0.62	0.17 ± 0.80	0.08 ± 0.53	0.148	0.170	0.180	0.317
Overall Child-OIDP	4.46 ± 6.80	4.20 ± 6.01	4.38 ± 6.42	4.21 ± 6.14	0.982	0.781	0.726	0.909

[a]The data sets are: Set A_{SAQ}, SAQ Child-OIDP scores of group A (see Table 1); set B_{SAQ}, SAQ Child-OIDP scores of group B; set A_{FTFI}, FTFI Child-OIDP scores of group A; and set B_{FTFI}, FTFI Child-OIDP of group B.
[b]Mann–Whitney test.
[c]Wilcoxon signed-rank test for paired groups.

corroborating that order effects did not influence children's reports of their sociodental impacts, we found that there was a high level of agreement between the SAQ and FTFI Child-OIDP. This finding was irrespective of using the Child-OIDP score or the prevalence of oral impacts to describe the sociodental impacts attributed to oral conditions among these children.

As this study assessed the same OHRQoL measure under alternative modes of administration, it was thought that children could be tired, distracted or familiarized with the topics when completing the second administration. Therefore, the order of administration of the SAQ and FTFI Child-OIDP was randomized to take into account potential order effects. However, the order of administration had no effect on children's responses. This may be due to the relatively short time required to complete each instrument (6 min on average), which could have reduced the likelihood of differences

Table 4. Comparison of scores and prevalence of impacts between the Self-administered (SAQ) and interviewer-administered (FTFI) Child-OIDP in the whole sample (n = 177)

Scores and prevalences	SAQ	FTFI	P-value[a]
1. Eating			
Performance score, mean ± SD	0.76 ± 1.40	0.80 ± 1.47	0.544
Prevalence of impacts (%)	31.1	29.9	0.727
2. Speaking			
Performance score, mean ± sd	0.28 ± 1.02	0.32 ± 1.11	0.281
Prevalence of impacts (%)	9.0	10.2	0.500
3. Cleaning mouth			
Performance score, mean ± SD	0.31 ± 1.01	0.25 ± 0.76	0.173
Prevalence of impacts (%)	12.4	14.1	0.250
4. Sleeping			
Performance score, mean ± sd	0.24 ± 1.00	0.25 ± 1.04	0.345
Prevalence of impacts (%)	9.0	8.5	≈1
5. Emotion			
Performance score, mean ± SD	0.33 ± 1.10	0.30 ± 1.05	0.201
Prevalence of impacts (%)	13.0	11.9	0.500
6. Smiling			
Performance score, mean ± sd	0.97 ± 2.12	0.95 ± 2.04	0.814
Prevalence of impacts (%)	21.5	23.2	0.250
7. Studying			
Performance score, mean ± SD	0.08 ± 0.40	0.09 ± 0.40	0.317
Prevalence of impacts (%)	5.6	6.2	≈1
8. Social contact			
Performance score, mean ± sd	0.15 ± 0.72	0.12 ± 0.68	0.109
Prevalence of impacts (%)	6.2	4.5	0.250
Overall Child-OIDP			
Oveall score, mean ± SD	4.33 ± 6.41	4.30 ± 6.27	0.784
Prevalence of impacts (%)	55.4	55.9	≈1

[a]Wilcoxon signed-rank test for paired groups for performance/overall scores, and McNemar test for prevalence of impacts, comparing for each performance and overall score the percentage of category = 0 between the two modes (SAQ and FTFI).

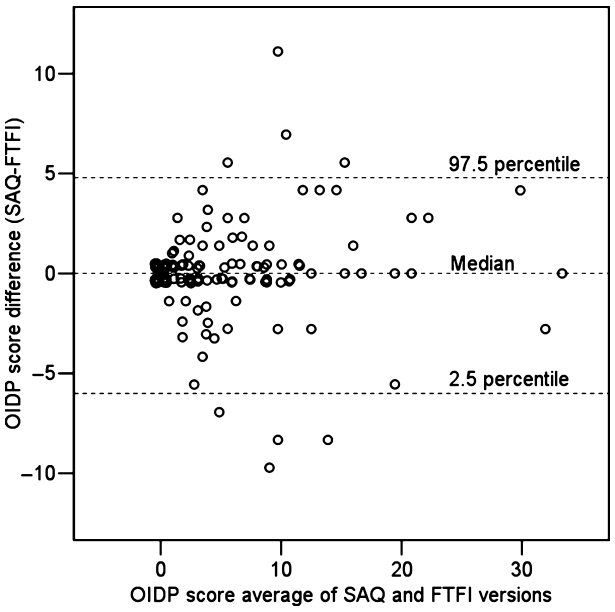

Fig. 1: Agreement in Child-OIDP scores: difference (SAQ-FTFI) versus average of values with 95% limits of agreement. Slight random variations were introduced to differentiate among circles (children) with the same values.

produced by fatigue or distraction. In this regard, it has been suggested that order effects are larger in longer questionnaires (9).

Regarding the level of agreement between the SAQ and FTFI Child-OIDP, there were no statistically significant differences between the two modes of administration, either in scores or prevalence of oral impacts (Table 4). Using the Bland and Altman method (15, 16), high levels of agreement were found between the SAQ and FTFI Child-OIDP when the overall score was used to describe children's oral impacts on their quality of life. The mean (±SD) difference between the two instruments was almost zero and 95% of the differences lay between −6.32 and +4.93 on a scale of 0 to 100. Further studies on the responsiveness and especially the minimally important difference for the Child-OIDP index are needed to clarify the clinical relevance of the variations observed. According to the Kappa value, a high level of agreement was also found when the prevalence of oral impacts was used to describe children's oral impacts on daily performances. This reinforces the

respective findings for the Child-OIDP score, particularly as Kappa is not affected by the prevalence of the event studied (29). Previous studies comparing different modes of administration for HRQoL measures have been contradictory (30–32), although many researchers found that a better health status was recorded using an interviewer-administered versus a SAQ (14, 31, 33–35). Differences in the length, text adaptation and application of questionnaires and in the characteristics of study populations may explain discrepancies in published findings. As both the SAQ and FTFI Child-OIDP were administered on the same day, our findings could be partly due to the fact that children may have matched their responses between the two administrations. However, previous research has shown that the consecutive administration of two instruments tends to produce lower scores in the second measure because of desensitization or familiarization with the questions (8–13). More importantly, the randomization in the order of presentation for the SAQ and FTFI Child-OIDP allowed us to compare scores for the two questionnaires at baseline (namely, before the second administration and thus, independently of any attempt to match responses or recall bias). Indeed, the difference in baseline scores between the SAQ and FTFI Child-OIDP was nonsignificant and of similar magnitude to the differences within groups (A_{SAQ} versus A_{FTFI} and B_{SAQ} versus B_{FTFI}, respectively), therefore providing evidence that children did not try to match their responses across the sequential administration of the two questionnaires.

There are two other findings worthy of discussion. The first finding relates to the extent of missing data in each mode of administration. Even though the SAQ Child-OIDP led to slightly more missing items than the FTFI Child-OIDP, the number of missing responses was very low with both modes of administration in this study (1.7% and 0.6% of the cases, respectively). Because missing item responses are one of the shortcomings of self-administered questionnaires (2, 3), this should be kept in mind when using the SAQ Child-OIDP in large epidemiological studies. The second finding relates to the psychometric properties of both administration modes. The SAQ and FTFI Child-OIDP were found to have acceptable psychometric properties. Since this finding agrees with those previously reported in a clinical sample (6), it seems that the mode of administration has no effect on the psychometric properties of the Child-OIDP index in either clinical or school-based samples. However, the psychometric properties of the SAQ Child-OIDP index need to be comprehensively tested in a large population-based sample of children.

Finally, the present findings are valid for the use of the Child-OIDP index in general child population and cannot necessarily be extrapolated to other groups. Future research should focus on other age groups and patient populations as well as on other administration modes, such as internet or telephone interviews. Following a general recommendation (36), longitudinal studies using this instrument are also warranted to assess the effects of treatment on OHRQoL.

In conclusion, having first demonstrated that the order of administration of the SAQ and FTFI Child-OIDP had no effects on children's report of oral impacts on quality of life, this study showed that the level of agreement between both modes of administration of the Child-OIDP index was very high. This finding was irrespective of using the overall Child-OIDP score or the prevalence of oral impacts to describe the oral impacts on children's quality of life.

Acknowledgements

This study was supported by the public research project #PI051172, Evaluation in Andalusia of the Oral Health Child Capitation Program (Fondo de Investigación Sanitaria, Spain). Requests for the Spanish version of the Child-OIDP index, in both formats (SAQ and FTFI), can be obtained from the authors. Eduardo Bernabé was supported by the Alßan Programme, the European Union Programme of High Level Scholarships for Latin America (Scholarship N° E06D1000352PE).

References

1. Gherunpong S, Tsakos G, Sheiham A. Developing and evaluating an oral health-related quality of life index for children; the CHILD-OIDP. Community Dent Health 2004;21:161–9.
2. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Oxf) 2005;27:281–91.
3. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. New York: Oxford University Press; 2003.
4. Robinson PG, Gibson B, Khan FA, Birnbaum W. A comparison of OHIP 14 and OIDP as interviews and questionnaires. Community Dent Health 2001;18:144–9.
5. Robinson PG, Gibson B, Khan FA, Birnbaum W. Validity of two oral health-related quality of life

measures. Community Dent Oral Epidemiol 2003;31:90–9.
6. Tsakos G, Bernabé E, O'Brien K, Sheiham A, de Oliveira C. Comparison of the self-administered and interviewer-administered modes of the Child-OIDP. Health Qual Life Outcomes 2008;6:40.
7. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17:571–82.
8. McColl E, Eccles MP, Rousseau NS, Steen IN, Parkin DW, Grimshaw JM. From the generic to the condition–specific? Instrument order effects in Quality of Life Assessment. Med Care 2003;41:777–90.
9. Cheung YB, Wong LC, Tay MH, Toh CK, Koo WH, Epstein R et al. Order effect in the assessment of quality of life in cancer patients. Qual Life Res 2004;13:1217–23.
10. Cheung YB, Lim C, Goh C, Thumboo J, Wee J. Order effects: a randomised study of three major cancer-specific quality of life instruments. Health Qual Life Outcomes 2005;3:37.
11. Childs AL, The Submacular Surgery Trials Research Group. Effect of order of administration of health-related quality of life interview instruments on responses. Qual Life Res 2005;14:493–500.
12. Jensen PS, Watanabe HK, Richters JE. Who's up first? Testing for order effects in structured interviews using a counterbalanced experimental design. J Abnorm Child Psychol 1999;27:436–9.
13. Lucas CP. The order effect: reflections on the validity of multiple test presentations. Psychol Med 1992;22:197–202.
14. Cheung YB, Goh C, Thumboo J, Khoo KS, Wee J. Quality of life scores differed according to mode of administration in a review of three major oncology questionnaires. J Clin Epidemiol 2006;59:185–91.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
16. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60.
17. Hebert JR, Miller DR. The inappropriateness of conventional use of the correlation coefficient in assessing validity and reliability of dietary assessment methods. Eur J Epidemiol 1991;7:339–43.
18. Schmidt ME, Steindorf K. Statistical methods for the validation of questionnaires – discrepancy between theory and practice. Methods Inf Med 2006;45: 409–13.
19. Altman DG, Bland JM. Measurement in medicine – the analysis of method comparison studies. Statistician 1983;32:307–17.
20. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30–46.
21. Stewart AL, Hays RD, Ware JE Jr. Methods of constructing health measures. Measuring functioning and well-being. In: Stewart AL, Ware JE editors. The medical outcomes study approach. Durham and London: Duke University Press; 1992; 67–85.
22. Locker D, Jokovic A, Stephens M, Kenny D, Tompson B, Guyatt G. Family impact of child oral and oro-facial conditions. Community Dent Oral Epidemiol 2002;30:438–48.
23. Grupo de trabajo de la Sociedad Española de Epidemiología y de la Sociedad Española de Medicina de Familia y Comunitaria. Una propuesta de medida de la clase social. Aten Primaria 2000;25:350–63.
24. Gherunpong S, Tsakos G, Sheiham A. The prevalence and severity of oral impacts on daily performances in Thai primary school children. Health Qual Life Outcomes 2004;2:57.
25. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–32.
26. Acquadro C, Conway K, Giroudet C, Mear I. Linguistic validation manual for patient-reported outcomes (PRO) instruments. Lyon: Mapi Research Institute; 2004.
27. Cortés-Martinicorena FJ, Bravo-Pérez M, Artázcoz-Osés J, Tsakos G, Rosel-Gallardo E. Estudio sobre la salud dental de los niños y adolescentes de Navarra (4ª edición). Necesidades normativas, percibidas y calidad de vida en relación a su salud oral. Pamplona: Fundación Miguel Servet. Gobierno de Navarra, 2008.
28. Cortés-Martinicorena FJ, Rosel-Gallardo E, Artázcoz-Osés J, Bravo M, Tsakos G. Adaptation and validation for Spain of the Child-Oral Impact on Daily Performance (C-OIDP) for use with adolescents. Med Oral Patol Oral Cir Bucal 2010;15:e106–11.
29. Vach W. The dependence of Cohen's kappa on the prevalence does not matter. J Clin Epidemiol 2005;58:655–61.
30. García M, Rohlfs I, Vila J, Sala J, Pena A, Masiá R et al. Comparison between telephone and self-administration of Short Form Health Survey Questionnaire (SF-36). Gac Sanit 2005;19:433–9.
31. Lyons RA, Wareham K, Lucas M, Price D, Williams J, Hutchings HA. SF-36 scores vary by method of administration: implications for study design. J Public Health Med 1999;21:41–5.
32. Wu AW, Jacobson DL, Berzon RA, Revicki DA, Van der Host C, Fichtenbaum CJ. The effect of mode of administration on medical outcomes study ratings and EuroQol scores in AIDS. Qual Life Res 1997;6: 3–10.
33. Weinberger M, Oddone EZ, Samsa GP, Landsman PB. Are health-related quality-of-life measures affected by the mode of administration? J Clin Epidemiol 1996;49:135–40.
34. Cook DJ, Guyatt GH, Juniper E, Griffith L, McIlroy W, Willan A et al. Interviewer versus self–administered questionnaires in developing a disease-specific, health-related quality of life instrument for asthma. J Clin Epidemiol 1993;46:529–34.
35. McHorney CA, Kosinski M, Ware JE Jr. Comparisons of the costs and quality of norms for the SF-36 health survey collected by mail versus telephone interview: results from a national survey. Med Care 1994;32:551–67.
36. Slade GD, Reisine ST. The child oral health impact profile: current status and future directions. Community Dent Oral Epidemiol 2007;35(Suppl. 1):50–3.