Structural equation modelling for a dental health related quality of life framework

Ana Nora Donaldson[1], James Wallace[2], Nelarine Cornellius[2], Nairn Wilson[3], Stephen Dunne[3], Angelo M Passalacqua[3].

[1]Stony Brook University, USA; [2]Bradford University, UK; [3] King's College London, UK

ABSTRACT

Background: Covariance-based (CB) structural equation modelling (SEM), implemented in Lisrel, EQS and AMOS, has been the default SEM approach but partial least square (PLS-SEM) is a new approach which offers, relatively, more flexibility.

Aim: We present the use of PLS-SEM in the context of assessing the factors that drive oral health related quality of life (QOL).

Methods: PLS-SEM is used on baseline data of 147 patients going for tooth extraction at a London dental clinic. QOL was measured using a well-established survey, the OHIP-14. Factors considered are: socio-demographics, oral health, health service availability, dental anxiety, locus of control (self-care and distrust in dentists) and dental-related knowledge, attitudes and behaviours.

Results: Oral health is the most significant direct driver of QOL, followed by dental anxiety. Ignorance (of harmful/beneficial practices) showed a marginal effect. Indirect effects are identified for health service (via oral health), dental anxiety (via oral health) and Ignorance (via dental fear). Mediating relationships are found between age and oral health, and between ignorance and dental anxiety. No significant effect of distrust-dentists, self-care or any other socio-demographic is found.

Conclusion: We verified hypothesized relationships. Under multivariate normality PLS-SEM produces similar results to CB_SEM but PLS-SEM offers more flexibility: it is distribution-free and able to model complex relationships with relatively smaller sample sizes. PLS-SEM is an excellent complement to CB-SEM.

Keywords : SEM ; Partial Least Squares ; PLS-SEM; Quality of life outcomes ; Public Health

1. Introduction

There are two main approaches to structural equation modelling (SEM). The widely used covariance-based structural equation modelling (CB-SEM), as implemented in LISREL, AMOS, EQS, which assumes multivariate normality, looks to explain the covariance matrix of the indicators. The more recent approach, known as partial least squares structural equation modelling **(PLS-SEM**), using proxies (linear combinations of the relevant indicators), maximizes the variance explained by each construct (dependent variable), and estimates the correlations between the latent variables [1]. PLS-SEM does not require multivariate normality. The aim of this paper is to present the PLS approach to SEM (PLS-SEM) in the context of a study assessing the effects of factors for self-assessed oral health related quality of life (OHRQOL), highlighting important methodological considerations in the development of the model. We use Smart-PLS software [2] for this purpose.

2. Methods

2.1 Data sources and target population.

Our case study consists of the baseline data taken on 147 patients referred to a London dental clinic for tooth extraction. The original data was collected as part of a study assessing factors that drive longitudinal changes in OHRQOL [3]. Patients in the study had a mean age of 47 years (95% C.I. 44 to 49), 47% are male. The ethnic distribution was: white (N=76; 55%), black (N=45; 32%) and Asian (N=16; 12%). The education distribution was: GCSE/below (N=67; 50%), A-levels (N=19; 14%) and College/above (N=49; 36).

2.2 The OHIP-14, a measure of self-perceived OHRQOL

Oral health related QOL is self-assessed using the oral health inventory profile (OHIP-14), a construct of 14 items that measure dental problems, grouped into 7 dimensions --two items per dimension [4, 5]. Each item is scored as the *product* of two variables: severity of the problem (VAS 0-4) and extent to which the given problem bothers the patient (VAS 0-3). Therefore, each item results in an ordinal variable ranging 0-12 and, each dimension, taken as the sum of the 2 items loading into it, ranges 0-24, with larger values indicating worse QOL. The dimensions are listed in Table 1 with the descriptive statistics for our sample. The OHIP-14 is usually analysed as an overall aggregate (range 0-168) or by dimension. In this paper we keep the QOL outcome multi-dimensional: an endogenous latent variable *reflectively* formed by these 7 dimensions. (The concept of *reflective* and *formative* constructs is presented in [1]).

Table 1. Summaries of the QOL (OHIP-14) dimensions for the study sample

Dimension	Mean	Std. Dev.	Min	Max
QOL1. Functional Limitation	2.1	3.8	0	21
QOL2. Physocal Pain	9.7	6.7	0	24
QOL3. Psychological Discomfort	6.5	6.3	0	24
QOL 4. Physical Disability	4.9	5.7	0	24
QOL 5. Psychological Disability	5.8	5.5	0	24
QOL 6. Social Disability	3.9	4.8	0	24
QOL 7. Handicap	4.1	5.2	0	24

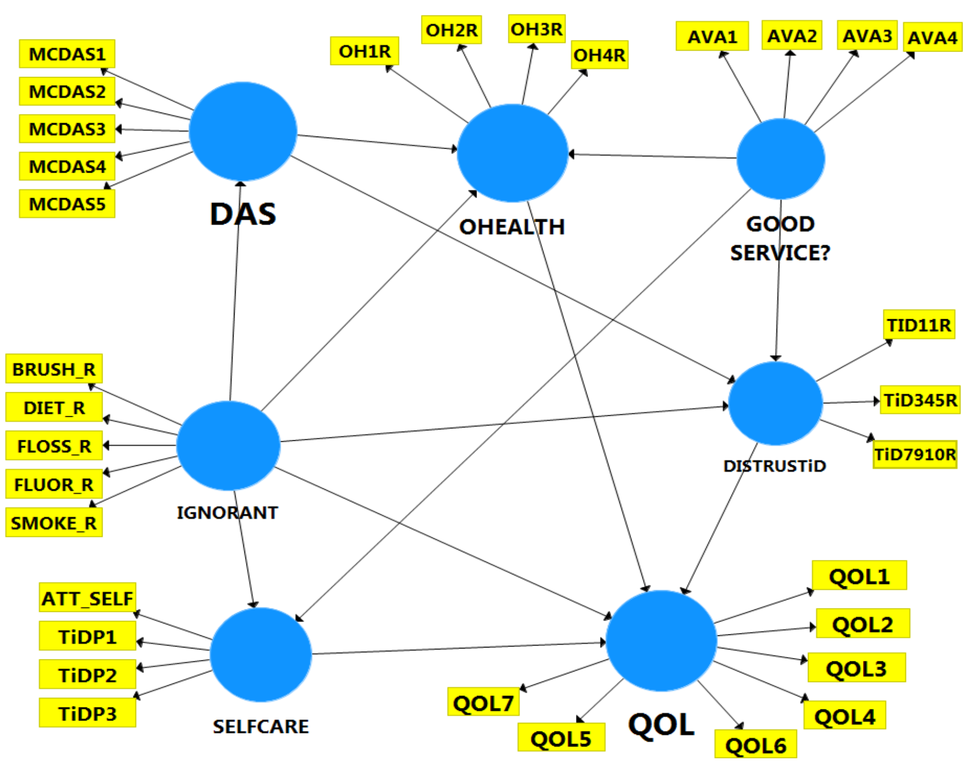
2.3 Potential prognostic factors shaping the self-perceived OHRQOL

In addition to the usual demographics (age, gender, ethnicity, education level), other factors considered for OHRQL were: perception of own *oral health*, *knowledge* of oral health harming or beneficial practices, oral health related behaviours, perception of availability of *good dental services*, and *attitudes* as considered in the research literature; in particular we consider the attitudes proposed in a qualitative framework by Gregory et al [6] who proposed a battery of items grouped in seven dimensions to reflect attitudes and behavior that are relevant to patients' ratings of their own OHRQOL: *Normative* self-perception, *Attribution of Control* of oral health (to self, others, values or dentist), *Trust* in dentistry and Trust in dental products, good dental service *Accessibility/Availability*, *Commodity* product view of dentistry, *Authenticity* preference over artificial beauty, and *Character* bias in judging oral health of others. *Dental anxiety* was also considered.

2.4 Initial theoretical model

Path models should be underpinned by theory –from the literature or logical reasoning. The path diagram for the theoretical model based on the Gregory et al framework [6], is presented in Figure 1.

Figure 1. Theoretical model



All the latent variables (in circles) are reflectively measured by the corresponding indicators in the following manner:

- *QOL,* an endogenous construct reflectively measured by the seven OHIP-14 indicators (dimensions) of worse OHRQOL.
- *Oral Health,* an endogenous construct reflectively measured by self-assessing indicators for worse oral health and symptom (bleeding gums severity).
- *Health Service,* an exogenous construct reflectively measured by indicators of good dental services accessibility and availability.
- *Ignorance,* an exogenous construct reflectively measured by indicators of ignorance of effects five practices: brushing, flossing, fluoridation, sweets in diet and smoking.
- *Self-care,* an endogenous construct reflectively measured by indicators of attributing control of dental care to self and the trust and benefits of dental products.
- *Distrust-in-Dentists (DISTRUSTiD),* an endogenous construct reflectively measured by indicators of distrust in their dentists approach to treatment and care.
- *DAS,* an endogenous construct reflectively measured by the five indicators of the MCDAS dental anxiety instrument scoring severity of dental anxiety under five dental related scenarios.

To sum up, in our latent variables, larger values indicated:

- Worse: QOL (OHRQOL), OHEALTH (self-perceived oral health), DAS (dental anxiety), IGNORANT (ignorance of the harms and benefits of widely known practices and policies) and DISTRUSTiD (distrust in dentist), and
- Better: HEALTH SERVICE (better availability and accessibility of dental services) and better SELF-CARE.

The model building process, in the usual way, consist of the following steps:

1. apply PLS path algorithm to theoretical model to obtain an intermediate model;

2. evaluate intermediate model quality. If model is valid, it is declared as the final model. Validity here refers to being logical and complying with benchmark values for psychometric properties (convergence and discriminant validity, internal consistency, etc).

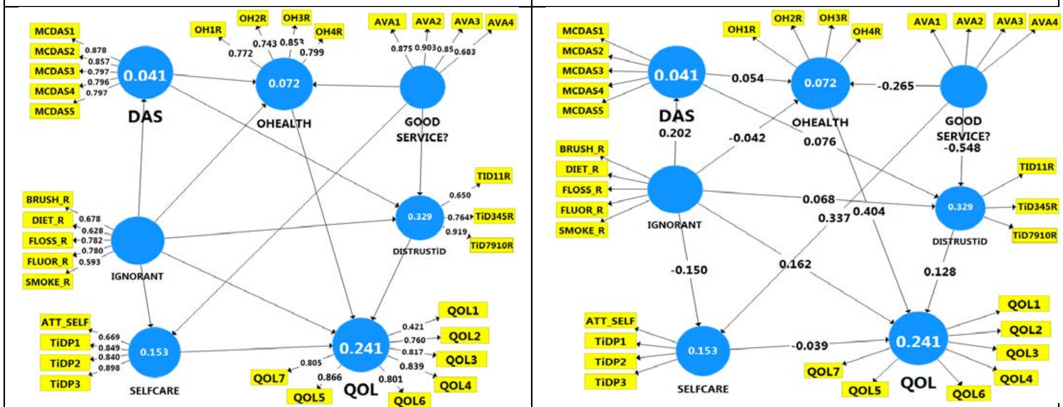
3. If the model is not confirmed, adjust intermediate model as necessary and return to 1.

1. Results

Figure 2 shows a model obtained in an intermediate stage. The path coefficients (loadings) of the outer model and, inside the circles, the R-square for the endogenous latent variables, are exhibited in Figure 2a. Figure 2b exhibits the inner path model.

Figure 2. Intermediate model

a. Measurement (outer) model, with path coefficients and R-square's	b. Structural Model with R-square's and path coefficients



3.1 Evaluating model quality.

Evaluation of the model is pursued in a systematic manner, checking at each stage, separately, the measurement (outer) model and then the structural (inner) model. The measurement model, if reflectively formed, is checked first for:

i. internal consistency, checking the composite reliability (CR) for each construct is between 0.70 to 0.90;
ii. indicator reliability, checking the standardized outer loading for each individual indicator is at least 0.708 (i.e. at least 50% is of the indicator variability contributes to the construct);
iii. convergent validity, checking the average variance explained (AVE) for each construct is at least 0.90 (AVE is the mean of the squared loadings of the indicators associated with the construct); and
iv. discriminant validity, checking that the square root of the AVE should be larger than all correlations of indicators not loading onto the latent construct being examined (i.e. the Fornell-Larcker criterion) and also checking the heterotrait-montrait (HTMT) ratio of correlations.

After reliability and validity are established, the structural model is checked. The PLS-SEM algorithm maximizes the explained variance for each endogenous construct hence covariance-based measures are not helpful for this task; instead the focus is on the predictive capability of the model. The structural model is therefore checked for:

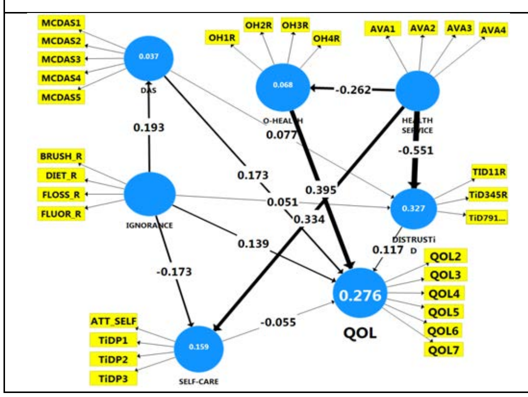
i. coefficients of determination (R-square),
ii. predictive relevance (Q-square),
iii. size and significance of path coefficients (this is achieved by bootstrapping the model and using the bootstrapped SE to estimate a t-value for each path and assess the significance; and
iv. effect sizes (f-square and q-square).

3.2 The final model for self-perceived OHRQOL

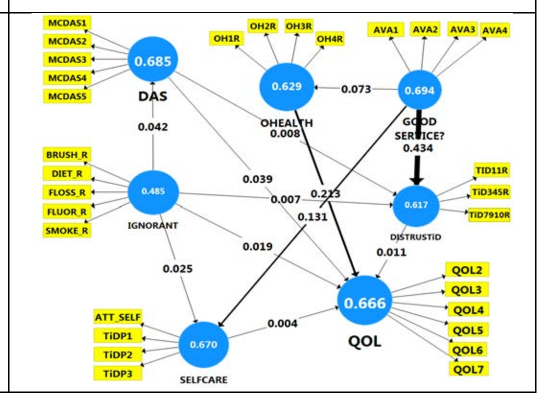
The final structural model, has the dimension QOL1 excluded and is presented in Figure 3. Figure 3a shows the path coefficients over the arrows and the coefficients of determination (R-sq) inside the circles of the endogenous latent variables. (The path coefficients are also presented in Table 4). Figure 3b shows the f-sq effect size over the arrow and its strength depicted by the thickness of the arrow; the AVE is also shown inside the circles of the latent variables.

Figure 3. Final model

a. With R-sq and path coefficients



b. With AVE and effect size



3.3 Quality of the OHRQOL measurement model

Figure 1b shows the model obtained in an intermediate stage of the process. The coefficient of determinations are all low. Several indicators showed reliabilities (path coefficients) below 0.708 (AVA4 in Good Health Service, TID11R in Distrust in Dentists and Brushing, Diet and Smoking in Ignorance) but they were not below 0.60 and their exclusion did not increase the quality of the structural model. Therefore, only the first indicator of QOL, with a reliability of only 0.42, was discarded; the effect of this was to increase the coefficient of determination for QOL from 0.241 to 0.273.

Table 2. Construct Reliability & Validity

	Cronbach's Alpha	CR	AVE	R-sq
HEALTH SERVICE	0.85	0.9	0.69	
Dental Anxiety (DAS)	0.89	0.912	0.69	0.04
DISTRUSTiD	0.68	0.83	0.62	0.32
IGNORANCE	0.74	0.82	0.49	
O-HEALTH	0.81	0.87	0.63	0.07
QOL	0.9	0.92	0.67	0.27
SELFCARE	0.83	0.89	0.67	0.15

3.4 Quality of the OHRQOL structural model

- Goodness of fit measures like the Normed fit index (NFI) and the Comparative fit index (CFI), are widely used in CB-SEM but are not applicable in the PLS context as the model does not have distributional requirements. Instead, given that ours is a purely reflective model, the RMS_Theta is recommended [7]. This is the root mean squared residual covariance matrix of the outer model residuals --the differences between corresponding elements of the observed and predicted covariance matrix.

- The RMS_Theta assesses the degree to which the outer model residuals correlate. Zero represents a perfect fit and the cut-off value for RMS_Theta to suggest a good fit or lack of fit is 0.12 [8]. The RMS_Theta in our model was 0.14 which is above this minimum widely accepted for a good fit.

- The average variance explained (AVE), see Table 2, is used to check if there is convergent validity (AVE should be at least 0.5). The AVE for most of the constructs is well above 0.50 so convergent validity is ensured (Ignorance has an AVE of 0.49).

- From the Fornell & Larcker criterion [9], the square root of AVE is used to check discriminant validity. Fornell and Larcker argued that, for acceptable discriminant validity, the sqrt(AVE) needs to be greater than the correlation of the construct with any of the remaining constructs - and this relationship should hold for each construct. As seen in Table 3, discriminant validity is indicated for the final fitted model.

- Table 4 presents the results of the validation of the paths with bootstrapping t-statistics, showing the statistical significance for the paths.

Table 3. Discriminant Validity

	HEALTH SERVICE	DAS	DISTRUST iD	IGNORANCE	O-HEALTH	QOL	SELF-CARE
HEALTH SERVICE	0.833						
DAS	-0.048	0.83					
DISTRUSTiD	-0.563	0.11	0.785				
IGNORANCE	-0.169	0.2	0.177	0.697			
OHEALTH	-0.262	0.05	0.123	0.013	0.793		
QOL	-0.227	0.23	0.237	0.197	0.428	0.82	
SELFCARE	0.363	0.04	-0.555	-0.206	-0.092	-0.18	0.82

Table 4. Validation of the path coefficients by bootstrapping

	Path Coefficient (O)	Sample Mean	Standard Deviation (SD)	t =|O/SD|	Sig P
DAS -> DISTRUSTiD	0.077	0.073	0.073	1.05	0.30
DAS -> QOL	0.173	0.175	0.067	2.58	0.01
Distrust Dentists -> QOL	0.117	0.11	0.076	1.54	0.12
Health Service -> DISTRUSTiD	-0.551	-0.554	0.055	10.04	0.0000
Health Service -> Oral Health	-0.262	-0.26	0.084	3.10	0.002
Health Service -> SELFCARE	0.334	0.339	0.081	4.15	0.0000
IGNORANCE -> DAS	0.193	0.197	0.108	1.79	0.07
IGNORANCE -> DISTRUSTiD	0.051	0.059	0.079	0.64	0.52
IGNORANCE -> QOL	0.139	0.134	0.076	1.82	0.07
IGNORANCE -> SELFCARE	-0.173	-0.183	0.105	1.64	0.10
Oral Health -> QOL	0.395	0.404	0.065	6.08	0.0000
SELF-CARE -> QOL	-0.055	-0.071	0.079	0.71	0.48

3.5 Interpretation of the final model

Using Figure 3a and Table 4, the following direct effects are observed:

- Self-perception of *ORAL HEALTH* is the most significant direct driver of *QOL* (P<0.001).
- A significant direct effect was also found for *DAS* (P=0.01).
- *IGNORANCE* (of harmful/beneficial practices) shows a weak borderline significant direct effect (P=0.07) on QOL.

Indirect effects on QOL are also identified for *DENTAL SERVICE* (*Betha*=-0.19), DAS (*Betha* =0.01) and *IGNORANCE* (*Betha* =0.049). No significant effects of *DISTRUSTiD* (P=0.12) or *SELFCARE (P=0.48)* on QOL were found.

The total effects are calculated using the additive-multiplicative model. This is to say, one looks at every possible path from construct A to construct B. The effect for every possible path from construct to construct (these are the indirect effects) is calculated using the multiplicative model (i.e. multiplying the coefficients for each path) and then add all of these together (additive model). Note that the Total contains the direct effect. For example, the indirect effect (-0.19) of *Dental Services on QOL* is obtained as the sum of the effects resulting from the paths:

- (Good Dental Service →Distrust in Dentists → QOL). This resulting effect is -0.548 X 0.128=-0.07
- (Good Dental service →ORAL HEALTH → QOL). This resulting effect is -0.26 X 0.40 =- 0.107
- (Good Dental service →Selfcare → QOL). This resulting effect is 0.337 X -0.039 = -0.013.

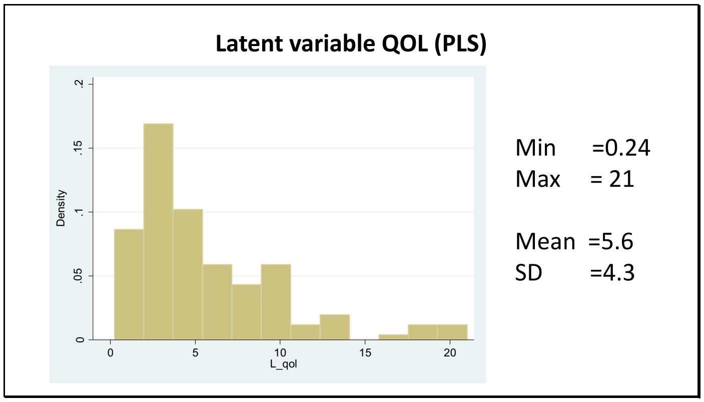
The indirect effect of *IGNORANCE* ON *QOL* (0.049) is found by adding the effects through *DAS* (0.03339), *DISTRUSTiD* (0.00585) and *SELF-CARE* (0.00952). Adding this indirect effect to the direct effect (0.139), the total effect of *IGNORANCE* on *QOL* is 0.19.

In the same manner one can check that there is an indirect effect of *IGNORANCE* on *DISTRUSTiD* (0.015) which, by the way, has no impact on QOL.

3.6 The estimated OHRQOL latent variable

The PLS latent variable QOL, based on 147 observations, ranged from 0.225 to 21, with a mean of 5.6 (SD=4.3). The histogram for this is exhibited in Figure 5.

Figure 5



The PLS-SEM results were consistently supported by traditional multivariate linear regression: the direct effects of oral health (Coeff=3.2; 95% CI 2 to 4.4; $P<0.001$) and DAS (Coeff=0.81; 95% CI 0.17 to 1.5; P=0.01) on QOL were manifest.

On simple linear regression, no statistically significant effects on QOL were found for gender (P=0.75); marital status ($F_{2,124}$=0.21; P=0.81); education (F(2,130)=0.47; P=0.62); Brush daily frequency >1 (P=0.37).

By the approach advocated by Baron & Kenny [10], age and ignorance on QOL mediate the effects of Oral health and Dental Anxiety on QOL. Significant effects of Age (*Betha*=-0.06; 95% CI -0.11 to -0.01; P=0.03) and IGNORANCE (*Betha* =1.19; 95% CI -0.002 to 2.6; P=0.05) are suggested but, when the latent variables O-HEALTH and DAS are present, the significances for age and IGNORANCE are somewhat reduced: (age *Betha* =-0.06; 95% CI -0.13 to 0; P=0.06) and (ignorance *Betha* =1.25; 95% CI -0.10 to 2.6; P=0.07).

2. Conclusion

The partial least square approach to structural equation modelling yields consistent and reasonable models while allowing smaller samples required (in relation to CB-SEM). It deals with complex models (formative, reflective, mediating and moderating effects) with ease. In contrast to CB-SEM, PLS-SEM is a non-parametric method in the sense that does not need to assume multivariate normality. Easily deals with exploratory data analysis. By the nature of the method, it readily produce estimators of the constructs which allow further investigations.

PLS-SEM is an excellent complementary approach to CB-SEM when: sample sizes are small, no normality is suspected, there are many indicators, need to explore indirect relationships (e.g. mediations, moderations, etc.), need to accommodate complex relationships (e.g. *formativ*ely and *reflectively* formed constructs). Smart-PLS is a very versatile, fun and easy to use tool.

In relation to our case study, we found that, not surprisingly, self-perception of oral health is the most significant direct driver of oral health related quality of life (OHRQOL or QOL for short in this paper), followed by dental anxiety. Similar results are found in the original study based on the same patient population [3], in covariance-based SEM of the 2008 Adult Dental Survey for England and Wales [11] and other NHS (London based) patient populations [12].

Indirect effects were manifest for availability of a good dental health service and ignorance of basic harmful and beneficial habits and policies. Age had a mediating effect with self-perception of oral health but no other socio demographic variable and none of the attitudes considered in Gregory's framework showed to have any effect.

4. References

[1] Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modeling. 2nd Edition. Thousand Oaks: Sage.

[2] Ringle CM, Wende S, Will A (2005). “SmartPLS 3.0.” University of Hamburg, URL http://www.smartpls.de/.

[3] Passalacqua, A.M., Dunne, S., Newton, T., Wilson, N., Donaldson, A.N. Responsiveness of the OHIP-14 and Factors Relevant to Oral Health Related Quality of Life. 88th General Session & Exhibition of the IADR. Barcelona, Spain. July 14-17, 2010

[4] Locker, D., Berka, E., Jokovic, A., & Tompson, B. Does self-weighting of items enhance the performance of an oral health-related quality of life questionnaire? *Community Dentistry And Oral Epidemiology. 2007; 35*(1), 35-43.

[5] Locker, D., Quiñonez, C. Functional and Psychosocial Impacts of Oral Disorders in Canadian Adults: A National Population Survey. JCDA. www.cda-adc.ca /jcda September 2009, Vol. 75, No. 7. https://www.cda-adc.ca/jcda/vol-75/issue-7/521.pdf (accessed on 10/09/2016 21:34 ET).

[6] Gregory, J., Gibson, B., Robinson, P. G. (2005). Variation and change in the meaning of oral health related quality of life: a grounded systems approach. *Soc Sci Med, 60*(8), 1859-1868.

[7] Lohmöller, J.-B. (1989). Latent Variable Path Modeling with Partial Least Squares. Physica: Heidelberg.

[8] Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Hair, J. F., Hult, G. T. M., & Calantone, R. J. 2014. Common Beliefs and Reality about Partial Least Squares: Comments on Rönkkö & Evermann (2013)**.** *Organizational Research Methods* 17(2): 182-209.

[9] Fornell, C. G., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

[10]. Baron, R. M. and Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research – Conceptual, Strategic and Statistical Considerat*ions. Journal of Personality and Social Psychology.* Vol. 51(6), pp. 1173–1182.

[11] Donaldson, N., Everitt, B., Newton, T., Steele, J., Sherriff, M., Bower, E. The effects of social class and dental attendance on oral health. ***Journal of Dental Research. January 2008.*** vol. 87 no. 1 **60-64**

[12] Milgrom, P., Newton, T., Boyle, C., Heaton, L., Donaldson, N. The Effects Of Dental Anxiety And Irregular Attendance On Referral For Dental Treatment Under Sedation

Within The National Health Service In London. ***Community Dentistry and Oral Epidemiology***. 2010; 38: 453-459.