

The development of the PAR Index (Peer Assessment Rating): reliability and validity

S. Richmond, W. C. Shaw, K. D. O'Brien, I. B. Buchanan, R. Jones, C. D. Stephens,* C. T. Roberts, and M. Andrews

Department of Oral Health and Development, University of Manchester, and *Department of Child Dental Health, Bristol Dental Hospital, Bristol

SUMMARY The PAR Index has been developed to provide a single summary score for all the occlusal anomalies which may be found in a malocclusion. The score provides an estimate of how far a case deviates from normal alignment and occlusion. The difference in scores between the pre- and post-treatment cases reflects the degree of improvement and, therefore, the success of treatment. Excellent reliability was exhibited within and between examiners (Intraclass Correlation Coefficient, $R > 0.91$). The components of the PAR Index have been weighted to reflect current British orthodontic opinion and is flexible in that the weightings could be changed to reflect future standards and standards currently being achieved in other countries.

The PAR Index offers uniformity and standardization in assessing the outcome of orthodontic treatment.

Introduction

The grading of orthodontic treatment results at study group meetings has been practised for a long time. The idea of individuals grading their own treatment results can be a self-teaching device and improve the quality of future treatment (Hickham, 1975). Unfortunately, the variation in criteria used by different orthodontists makes it difficult to compare results of treatment and the same applies to comprehensive studies which are not limited to the evaluation of specific measurements. Several indices have been developed specifically to assess the success of treatment (Eismann, 1974, 1980; Gottlieb, 1975; Berg, 1979). These indices compare pre- and post-treatment records to register the outcome of orthodontic care. However, the reliability and validity of these indices have never been evaluated. A treatment need index, the Occlusal Index (Summers, 1971) has been used by several authors to assess the outcome of treatment (Pickering and Vig, 1975; Elderton and Clark, 1983, 1984).

The use of indices should ensure uniform interpretation and application of criteria. Although numerous indices have been developed none as yet has been accepted universally. The use of precise criteria is essential,

requiring a quantitative objective method of measuring malocclusion and efficacy of treatment.

To fulfil these criteria the PAR Index (Peer Assessment Rating) was developed to record the malocclusion at any stage of treatment. The index was formulated over a series of six meetings in 1987 with a group of 10 experienced orthodontists (British Orthodontic Standards Working Party). Over two-hundred dental casts representing developmental as well as pre- and post-treatment stages were examined and discussed until agreement was reached regarding the individual features which would be assessed in obtaining an estimate of alignment of occlusion. The working party is shown (Fig. 1) assessing a particular case which has been projected onto four screens allowing all members of the working party to view the same case at one time. The scores derived were directly recorded on a computer database to facilitate rapid multiple examiner comparisons.

Section 1: The PAR Index (Peer Assessment Rating)—description and reliability

The concept is to assign a score to various occlusal traits which make up a malocclusion. The individual scores are summed to obtain an



Figure 1 The British Orthodontic Standards Working Party assessing cases projected on four screens, the results were directly entered onto computer to facilitate intra- and inter-examiner error.

overall total, representing the degree a case deviates from normal alignment and occlusion.¹ The score of zero would indicate good alignment and higher scores (rarely beyond 50) indicating increased levels of irregularity. The overall score is recorded on the pre- and post-treatment dental casts. The difference between these scores represents the degree of improvement as a result of orthodontic intervention and active treatment.

There are 11 components of the PAR Index (Table 1).

Table 1 Components of the PAR Index.

1.	Upper right segment
2.	Upper anterior segment
3.	Upper left segment
4.	Lower right segment
5.	Lower anterior segment
6.	Lower left segment
7.	Right buccal occlusion
8.	Overjet
9.	Overbite
10.	Centreline
11.	Left buccal occlusion

¹Normal occlusion and alignment is defined as all anatomical contact points being adjacent, with a good intercuspal mesh between upper and lower buccal teeth, non-excessive overjet and overbite.

Buccal and anterior segments

The dental arch is divided into three recording segments, left buccal, right buccal and anterior (Fig. 2). Scores are recorded for both upper and lower arches.

Buccal segments. The recording zone is from the mesial anatomical contact point of the first permanent molar to the distal anatomical contact point of the canine.

Anterior segment. The recording zone is from the mesial anatomical contact point of the canine on one side to the mesial anatomical contact point of the canine on the opposite side.

The occlusal features recorded are crowding, spacing, and impacted teeth. Displacements are recorded as the shortest distance between contact points of adjacent teeth parallel to the occlusal plane. The greater the displacement the greater the PAR score. The displacements between first, second, and third molars are not recorded as these contacts are so broad and are extremely variable within the normal range.

An impacted tooth is recorded when the space for this tooth is less than or equal to 4 mm. Impacted canines are recorded in the anterior segment.

Scores for the displacements and impactions are added to give an overall score for each recording zone. The scores for the discrepancies are shown in Table 2.

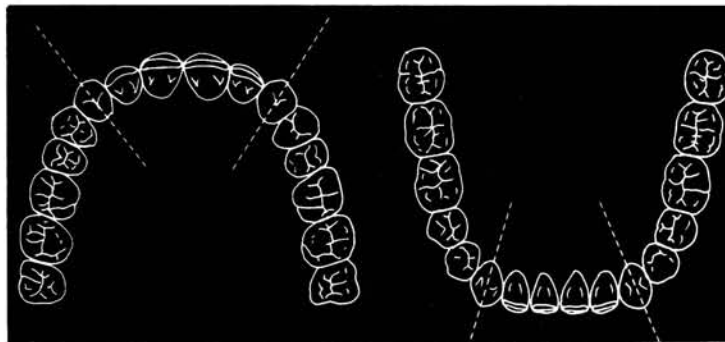


Figure 2 Division of the dental arches into three recording zones: anterior, left, and right.

Table 2 Displacement scores.

Score	Discrepancy
0	0 mm to 1 mm
1	1.1 mm to 2 mm
2	2.1 mm to 4 mm
3	4.1 mm to 8 mm
4	greater than 8 mm
5	impacted teeth

If there is potential crowding in the mixed dentition, average mesio-distal widths are used to calculate the space deficiency (Table 3).

Buccal occlusion

The buccal occlusion is recorded for both left and right sides. The fit of the teeth is scored with respect to the three planes of space. The recording zone is from the canine to the last molar, either first, second, or third. All discrepancies are recorded when the teeth are in occlusion.

The antero-posterior, vertical and transverse irregularities are summed for each buccal occlusion (Table 4).

Overjet

Positive overjet as well as teeth in cross-bite are recorded (Table 5). The recording zone is from the left to right lateral incisors. The most prominent aspect of any one incisor is recorded. When recording the overjet the ruler is held parallel to the occlusal plane and radial to the line of the arch. It is not uncommon to see two upper laterals in cross-bite as well as an increased overjet on the central incisors. In this situation if the overjet were 4 mm, the score would be 3 for the cross-bite and 1 for the positive overjet (4 in total).

Overbite

Records the vertical overlap or open bite of the anterior teeth. Overbite is recorded in relation to the coverage of the lower incisors or the degree of open bite (Table 6). The recording zone includes the lateral incisors. The tooth with the greatest overlap is recorded.

Centreline

Records the centreline discrepancy in relation to the lower central incisors (Table 7). If a lower

Table 3 Mixed dentition crowding assessment using average mesio-distal widths.

Upper		
canine	8 mm	Total = 22 mm (impaction <= 18 mm)
1st premolar	7 mm	
2nd premolar	7 mm	
Lower		
canine	7 mm	Total = 21 mm (impaction <= 17 mm)
1st premolar	7 mm	
2nd premolar	7 mm	

Table 4 Buccal occlusion assessments. (Temporary developmental stages and submerging deciduous teeth are excluded.)

Score	Discrepancy
Antero-posterior	
0	Good interdigitation Class I, II and III
1	Less than half unit discrepancy
2	Half a unit discrepancy (cusp to cusp)
Vertical	
0	No discrepancy in intercuspation
1	Lateral open bite on at least two teeth greater than 2 mm
Transverse	
0	No cross-bite
1	Cross-bite tendency
2	Single tooth in cross-bite
3	More than one tooth in cross-bite
4	More than one tooth in scissor bite

Table 5 Overjet measurements.

Score	Discrepancy
Overjet	
0	0-3 mm
1	3.1-5 mm
2	5.1-7 mm
3	7.1-9 mm
4	greater than 9 mm
Anterior cross-bites	
0	No discrepancy
1	One or more teeth edge to edge
2	One single tooth in cross-bite
3	Two teeth in cross-bite
4	More than two teeth in cross-bite

central incisor has been extracted the measurement is not recorded.

The PAR ruler

A ruler has been designed to make measurement easier (Fig. 3). The information briefly summarizes the recording features of the index and facilitates quick assessments, and allows contact points to be viewed through the ruler.

The ruler is shown recording the displacement between the mesial contact points of the incisors (Fig. 4). The line designated 2 is slightly short of the contact points; therefore, the next longer line would be used and the contact point score between the central incisors would be 3. Figure 5

Table 6 Overbite measurements. Cross-bites including the canines are recorded in the anterior segment.

Score	Discrepancy
Open bite	
0	No open bite
1	Open bite less than and equal to 1 mm
2	Open bite 1.1-2 mm
3	Open bite 2.1-3 mm
4	Open bite greater than or equal to 4 mm
Overbite	
0	Less than or equal to one third coverage of the lower incisor
1	Greater than one-third, but less than two-thirds coverage of the lower incisor
2	Greater than two-thirds coverage of the lower incisor
3	Greater than or equal to full tooth coverage

Table 7 Centreline assessments.

Score	Discrepancy
0	Coincident and up to one-quarter lower incisor width
1	One-quarter to one-half lower incisor width
2	Greater than one-half lower incisor width

ANT-POST	
0	None
1	< 1/2 unit dis
2	= 1/2 unit dis
TRANSVERSE	
0	None
1	Xbite tend > = 1t
2	1 tooth in xbite
3	> 1 tooth in xb
4	> 1 tooth in sb
VERTICAL	
0	None
1	openb 2t > 2mm
CENTRELINE	
0	< = 1/4
1	1/4 - 1/2
2	> 1/2
OVERBITE	
0	0 - 1/3 open b
1	1/3 - 2/3 -
2	> 2/3 -
3	> = FTC -
4	→
CONTACT Pt	
0	-
1	-
2	-
3	-
4	→
5	Impacted tooth
THE PAR INDEX Manchester	
OVERJET	
4	> 2t xb
3	2 t xb
2	1 t xb
1	e to e
0	

Figure 3 The PAR ruler.

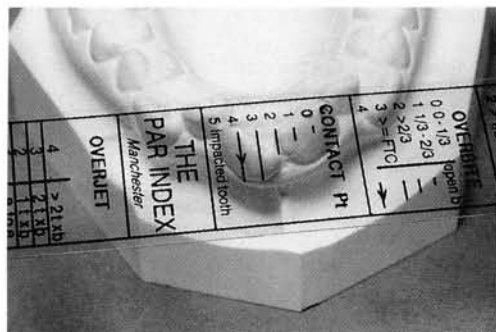


Figure 4 The PAR ruler being used to record the displacement of contact points between the central incisors.

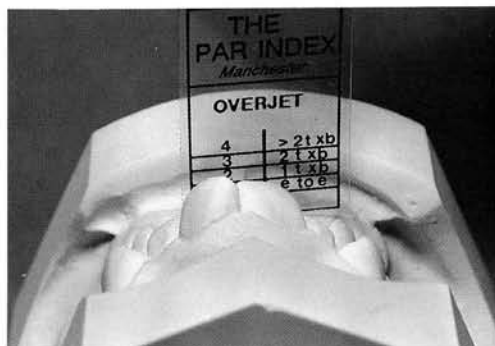


Figure 5 The PAR ruler being used to assess overjet: the overjet recorded would score 2.

shows the ruler being used to record the overjet and in this case would score 2.

Conventions

A series of conventions have been drafted to standardize the approach in assessing occlusions (Appendix 1).

Case analysis

To illustrate the use of the PAR Index, pre- and post-treatment dental casts are shown in Fig. 6;

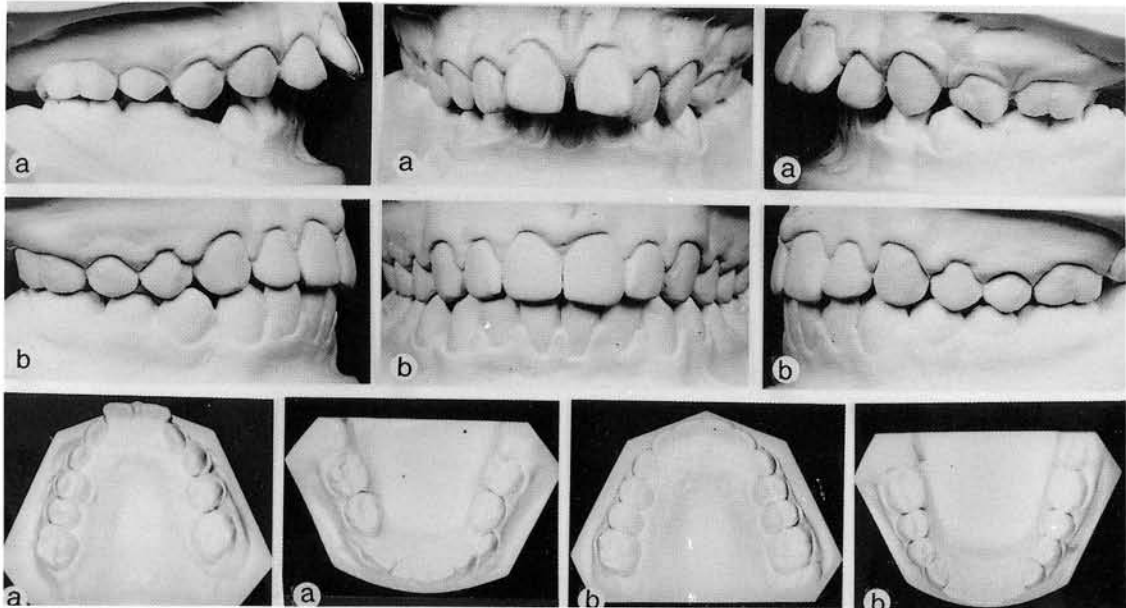


Figure 6 Cases analysis: (a) pre-treatment and (b) post-treatment dental casts illustrate the reduction in PAR score from 21 to 3.

the derivation of the individual component scores are recorded in Table 8. The initial score of 21 has been reduced to 3, a reduction of 18 points.

The case illustrated is a Class II, division 1 malocclusion with an increased overjet greater than 10 mm, and an increased and complete overbite. The molar relationship is half a unit

Class II on the right and more than a full unit Class II on the left side. The upper left second premolar is impacted.

As a result of orthodontic treatment the overjet and overbite have been reduced to less than 3 mm. The impacted second premolar has been accommodated. The buccal segments have been improved although the buccal teeth are

Table 8 Derivation of PAR scores (Fig. 6).

PAR components	Before treatment	Total	After treatment	Total
Upper right segment	(6-5),0, (5-4),0, (4-3),0	0	(6-5),0, (5-4),0, (4-3),1	1
Upper anterior segment	(3-2),0, (2-1),2, (1-1),1, (1-2),2, (2-3),0	5	(3-2),0, (2-1),0, (1-1),0, (1-2),0, (2-3),0	0
Upper left segment	(6-5),0, (5-4),0,(4-3),1, and (impaction 3)5	6	(6-5),0, (5-4),0, (4-3),0	0
Lower right segment	(6-5),0, (5-4),0 (4-3),0	0	(6-5),0, (5-4),0, (4-3),0	0
Lower anterior segment	(3-2),0, (2-1),1, (1-1)0, (1-2),0, (2-3),0	1	(3-2),0 (2-1),0, (1-1),0, (1-2),0, (2-3),0	0
Lower left segment	(6-5),0, (5-4),0, (4-3),0	0	(6-5),0, (5-4),0, (4-3),0	0
Right buccal occlusion	Cusp to cusp	2	Less than 1/2 unit Class II	1
Overjet	Overjet greater than 10 mm	4	Overjet within 3 mm	0
Overbite	Overbite within the normal range	0	Normal overbite	0
Centreline	Centreline >1/4 width of lower incisor	1	Centreline within 1/4 width	0
Left buccal occlusion	Cusp to cusp	2	Less than 1/2 unit Class II	1
Total		21		3

Notation of teeth in brackets, followed by the PAR score for each contact point.

not intercusating fully. The overall alignment has been improved although the distal of the upper right canine is slightly more buccal compared to the upper left canine.

A single score has been derived for the pre- and post-treatment dental casts representing the deviation from normal. The size of reduction in PAR score reflects the degree of improvement and success of treatment.

Reliability

Four examiners (two orthodontically qualified and two postgraduate students) were trained in the use of the PAR Index over a series of meetings.

In order to determine the examiner reliability a stratified subsample of 38 cases was selected from a random sample of 2000 treated and untreated cases. The dental casts were measured on two occasions 8 weeks apart. Differences in measurements were estimated by calculating the Intraclass Correlation Coefficient of Reliability. This provides a single summary of reliability based on a comparison between ratings and comparison between subjects (Barkto, 1966; Fleiss *et al.*, 1979; Fleiss, 1981, 1986). These statistics were calculated between the four examiners separately and as a group.

Results

Table 9 shows excellent reliability within ($R > 0.95$) and between examiners ($R = 0.91$).

The intra-examiner lower 95 per cent confidence limits are generally greater than 0.93 and for inter-examiner, 0.87. The paired *t*-test was used to examine for bias. Examiners 2 and 3 both indicated bias between first and second recordings. However, the differences between the means were less than 1.6 PAR points (means: Examiner 2—21.90, 22.97; Examiner 3—19.05, 20.58). An analysis of variance detected bias between the four examiners, Examiner 3, scoring the lowest and Examiner 4 the highest. The difference in the means being first recording, 3.66 and for the second recording 2.81 PAR points. The reliability coefficient and confidence limits were high indicating excellent reliability. The measure of reliability is large and the effect of bias would mean that little damage is done to the precision of measurement.

Discussion

In general, the PAR Index indicated excellent intra- and inter-examiner agreement. Lower levels of agreement were achieved for some of the individual PAR components, particularly the upper left and right buccal segments ($R = 0.76$) compared to the corresponding segments in the lower arch. This may be due to the increased size and the differing morphology of the maxillary teeth, which result in broad contact points and make accurate displacement measurements more difficult than in the lower arch.

It is not surprising that the left and right

Table 9 Calibration exercise: Intraclass Correlation Coefficient of Reliability for the four examiners. Intra- and inter-examiner

Component	Intra-examiner				Inter-examiner
	1	2	3	4	
Upper right	0.88	0.92	0.93	0.90	0.76
Upper anterior	0.94	0.98	0.98	0.95	0.90
Upper left	0.84	0.87	0.84	0.87	0.76
Lower right	0.93	0.95	0.93	0.92	0.87
Lower anterior	0.83	0.96	0.97	0.91	0.87
Lower left	0.96	0.97	0.91	0.86	0.84
Right buccal occlusion	0.79	0.89	0.87	0.89	0.75
Overjet	0.95	0.99	0.87	0.94	0.91
Overbite	0.74	1.00	0.82	0.89	0.77
Centreline	0.76	0.89	0.90	0.83	0.81
Left buccal occlusion	0.83	0.91	0.96	0.77	0.75
Total score	0.96	0.98	0.96	0.95	0.91
Lower 95 per cent confidence limit	0.94	0.97	0.94	0.93	0.87

buccal occlusions also show lower measures of agreement ($R=0.75$) as this depends on the correct registration of the models which can easily 'move' in the hand, and that the line of the buccal occlusion forms a curve and the recording depends on the angle to which the models are viewed.

However, the overall inter-examiner agreement of $R=0.91$ is excellent and compares very favourably to caries calibration exercises (Mitropoulos and Downer, 1987). The use of the ruler with a brief summary contained within, eliminates cross-references to notes of complicated criteria.

Timing trials have been undertaken on 10 sets of mixed and permanent dentitions (Buchanan, 1990). The average time taken to record the overall PAR score for pre- and post-treatment models is 6 minutes. Direct computer entry tends to reduce this time even further.

The difference between the pre- and post-treatment scores reflects the success or degree of improvement. As the score tends towards 0, the deviation from normal is less. Obviously, a score of 0 is not always achievable because of the complexity of the case, but generally a measure of 10 or less indicates an acceptable alignment and occlusion, and 5 or less suggests an almost ideal occlusion.

The case illustrated (Fig. 6) recorded a change in PAR score of 18 and indicates substantial improvement. However, not all cases achieve this degree of improvement.

Section II: The PAR Index—validation

Evaluation of the validity of an index involves enquiring into whether the index measures what it purports to measure. Many studies have used consensus of several experienced orthodontists in order to compare with the results of the index (Salzmann, 1968; Summers, 1971; Malmgren, 1980; Jenny *et al.*, 1983). In general, the process of validation involves the comparison of a subjective measure against a more objective measure of the characteristic.

The primary objective of the validation exercise was to compare a definitive measure of malocclusion using the indices with mean subjective dental opinion, in terms of deviation from normal for a representative sample of pre- and post-treated cases.

The PAR Index is made up of a number of

sub-components, each measuring distinct occlusal traits. Perhaps of primary contention amongst the profession would be the importance that should be attributed to any one component relative to another. For example, is an open bite equivalent in importance to a cross-bite or overjet? The validation study provides an opportunity to determine the relative weight that should be applied to each component.

Design of the validation study

The examiners. Seventy-four dentists were invited to participate, representing the various groups carrying out orthodontic treatment in England and Wales, 22 Consultant orthodontists, 22 specialist practitioners, two members of staff employed by the Dental Practice Board of England and Wales and two junior hospital staff members, all possessing specialist orthodontic qualifications, and 15 general dental practitioners and 11 community dentists without orthodontic qualifications.

Each member of the panel of 74 examiners was asked to assess a representative sample of dental casts with respect to deviation from normal occlusion on a nine point scale (Fig. 7). For each dental cast a mean score was obtained from the panel of 74 examiners representing consensus opinion.

The representative sample. A broad range of cases was collected representing treatments undertaken within the general dental and hospital services in England Wales. In all 272 cases were collected consisting of 68 cases from each of the following sources:

1. non-prior approval cases from the Dental Practice Board (one upper removable appliance treatment);
2. prior approval cases from the Dental Practice Board (more than one removable appliance or fixed appliances);
3. treated cases from Manchester and Bristol Dental Hospitals;
4. untreated cases selected from a study of 1000 Cardiff school children taken over a 4-year period 1981–1984 (Shaw *et al.*, 1986).

The dental casts were divided into four groups (A, B, C, and D) consisting of 40 pre- and post-treatment pairs of dental casts including 16 cases which were randomly selected from the

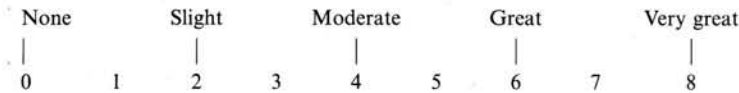


Figure 7 Nine-point scale to assess deviation from normal.

four sources and duplicated to allow them to be added to each group to enable examiner reliability assessments to be undertaken. Each of the four groups was presented as a randomized stratified sample. Three-hundred-and-twenty dental casts were examined. To reduce order bias, four sequences were devised to enable the dental casts to be examined in each session (Table 10). Frequent breaks were organized to reduce fatigue. Booklets were devised and issued in sequence so that an equal number of readings would be collected. Six validation study days were arranged in Manchester and Bristol.

Assessment of the dental casts using the indices. The four trained and calibrated examiners each scored one group using the PAR Index.

Relationship between the mean subjective ratings of the panel of 74 examiners and the PAR Index. For every dental cast the mean scores for the panel of 74 examiners were derived for deviation from normal occlusion. In order for an index to be valid the linear correlation, r , between it and the mean subjective score must be large. If this is so, the index will provide a valid substitute for the average opinion for the panel of 74 examiners, being a good predictor of their mean scores.

Results

The PAR Index. The raw total of the PAR components had a correlation, r , of 0.74 with the panel of 74 examiners mean subjective scores for deviation from normal occlusion. Direct summing of the sub-components might not provide the best index, as the profession might place greater importance on certain aspects. In order to determine the weights that should be placed on each component to best predict average opinion, multiple regression of the sub-components of PAR against the panel's mean score was carried out. Variables (components of PAR) were added using a forward stepwise procedure, provided a significant improvement in fit was achieved at each stage. The regression coefficients provided weights for each of the components. These were multiplied by a suitable constant and rounded to integers for the convenience of the user. The regression coefficients were derived on one half of the sample (by taking alternate cases) and the PAR components were multiplied in the same ratio as their partial regression coefficients and applied to the other half of the sample. Some of the components did not appear to have any predictive power and consequently were excluded from the weighted PAR Index (Table 11). The Pearson Correlation Coefficient, r , was 0.87 for the regression model and 0.85 for the weighted PAR

Table 10 Design of the validation study.

	Session				
	1st	2nd	3rd	4th	5th
Time	9.30–10.15	10.30–11.45	12.00–13.15	14.00–15.15	15.30–16.45
	Number of models				
	32 pairs	40	40	40	40
Booklet 1	B	A1–40	A41–80	B1–40	B41–80
Booklet 2	D	C1–40	C41–80	D1–40	D41–80
Booklet 3	C	B41–80	B1–40	C40–80	C1–40
Booklet 4	A	D41–80	D1–40	A41–80	A1–40

Table 11 Partial regression coefficients, unweighted and weighted integers for individual PAR components.

PAR components	Unweighted	Partial reg. coeff	Weighted
Buccal segments	1	0	0
Upper anterior	1	0.15	1
Lower anterior	1	0.11	1
R and L buccal occlusion	1	0.13	1
Overjet	1	0.65	6
Overbite	1	0.24	2
Centreline	1	0.47	4
r^2	0.47	0.75	0.72
r	0.68	0.87	0.85

Index). There were no significant differences between the correlation values for the regression model and the corresponding weighted index ($P > 0.1$). The weighted version gave a statistically significant higher correlation with the mean deviation from normal occlusion than the unweighted PAR ($P < 0.001$). The weighted regression line is shown in Fig. 8. The coefficient of determination, r^2 , the measure of prediction accuracy and strength of linear association was 0.72 for the weighted index.

Differences in weightings were noticed between the four groups (Table 12). The weightings for the centre line were much greater for consultants and general dental practitioners than for specialist practitioners and community dentists. The buccal segments (displacement measurements) were not weighted for any of

the four groups of examiners. The weighting for upper anteriors was 2 for the specialists and 1 for the other groups.

The mean weightings were taken to represent the collective opinion of the panel of the 74 dentists.

Reliability of the weighted PAR Index. The inter-examiner Intraclass Correlation Coefficient of Reliability for the total score for the four examiners on the 38 cases was 0.93 with the lower 95 per cent confidence limit of 0.90. The weighted index shows a slight improvement over the unweighted version (unweighted $R = 0.91$, lower 95 per cent confidence limit 0.87).

Case analysis. A case is shown in Fig. 9 and the unweighted and weighted PAR scores presented in Table 13. The case is a Class II

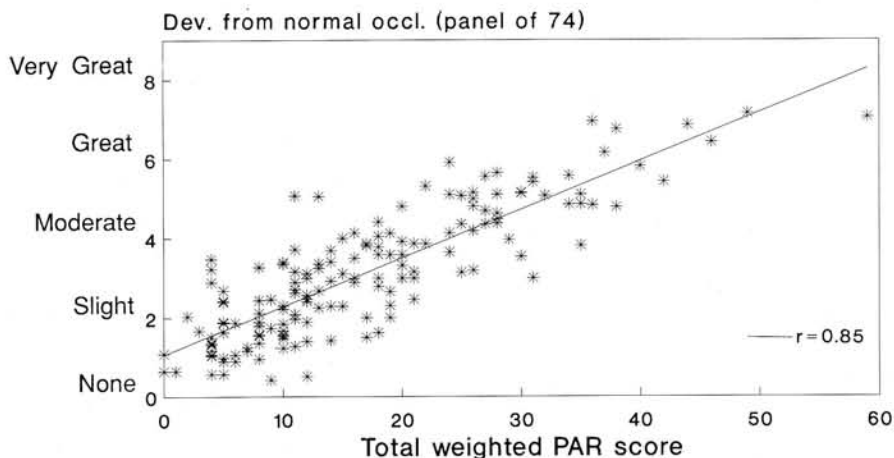
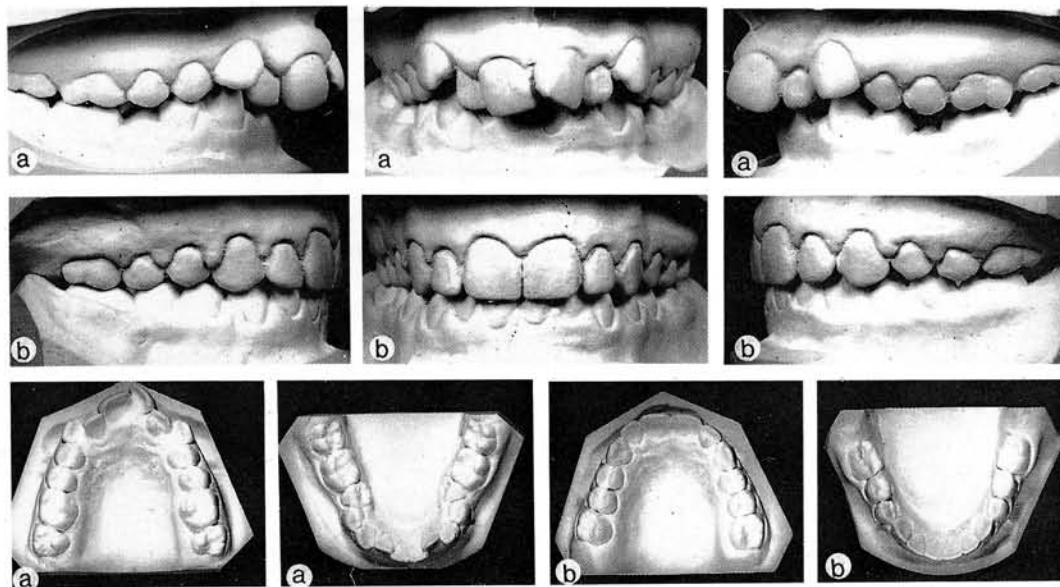
**Figure 8** Correlation between the panel of 74 examiners mean subjective scores for deviation from normal and the weighted PAR Index.

Table 12 Weightings derived for each of the groups of examiners.

	PAR components					
	Upper anterior	Lower anterior	Buccal occlusion	Overjet	Overbite	Centreline
Consultants	1	1	1	5	3	6
Specialists	2	1	1	4	1	-0.54
G.D.P.	1	1	1	5	2	4
Community	1	1	1	4	2	-0.25
All	1	1	1	6	2	4

**Figure 9** Case analysis: pre-treatment (a) and post-treatment (b) dental casts illustrate the reduction in PAR score from 45 to 2.

division 1 malocclusion with marked upper and lower incisor crowding. The overjet is increased over 10 mm, and the overbite is increased and complete. The molar relationship is three-quarter Class II on the right side and a quarter Class II on the left side. The lower second premolars are in lingual cross-bite.

The post-treated case shows that the overjet and overbite have been fully reduced. The buccal occlusions have been improved, but the buccal segments are not interdigitating fully. Good alignment has been achieved in both arches. The PAR scores were 36 for the pre-treatment and 2 for the post-treatment dental casts. The weightings have increased the initial

score by 5 to 41, but the finished score stays at 2. The major contribution to the initial weighted PAR score was from the overjet.

In the previous case (Fig. 6) the PAR score when weighted rose by 17 to 38 for the pre-treatment model and reduced by 1 to 2 for the post-treatment models.

The reduction in weighted PAR scores for the case in Fig. 6 (36 PAR points) and Fig. 9 (43 PAR points) represents the degree of improvement and the success of treatment.

Reliability of the panel of 74 examiners. Reliability within and between the panel of 74 examiners was tested on 16 sets of dental casts

Table 13 Derivation of PAR scores (Fig. 9).

PAR components	Before treatment	Total		After treatment	Total	
		UN	W		UN	W
Upper right segment	(6-5),0, (5-4),0, (4-3),2	2	-	(6-5),0, (5-4),0, (4-3),0	0	-
Upper anterior segment	(3-2),2, (2-1),1, (1-1),2, (1-2),0, (2-3),1	6	6	(3-2),0, (2-1),0, (1-1),0, (1-2),0, (2-3),0	0	0
Upper left segment	(6-5),0, (5-4),1, (4-3),1	2	-	(6-5),0, (5-4),0, (4-3),0	0	-
Lower right segment	(6-5),2, (5-4),2, (4-3),0	4	-	(6-5),0, (5-4),0, (4-3),0	0	-
Lower anterior segment	(3-2),1, (2-1),1, (1-1),0, (1-2),1, (2-3),1	4	4	(3-2),0, (2-1),0, (1-1),0, (1-2),0, (2-3),0	0	0
Lower left segment	(6-5),2, (5-4),3, (4-3),0	5	-	(6-5),0, (5-4),0, (4-3),0	0	-
Right buccal occlusion	Crossbite 45 and < cusp to cusp	3	3	Less than cusp to cusp	1	1
Overjet	Overjet greater than 10 mm	4	24	Overjet within 3 mm	0	0
Overbite	Overbite greater than 2/3	2	4	Normal overbite	0	0
Centreline	Centreline with 1/4 width of lower incisor	0	0	Centreline within 1/4 width	0	0
Left buccal occlusion	Cross-bite, cusp to cusp	4	4	Less than cusp to cusp	1	1
Total		36	45		2	2

UN—unweighted; W—weighted. Tooth notation in brackets, followed by the PAR score.

which had been duplicated and randomized within the groups. The chance corrected measure of agreement was used (O'Connell and Dobson, 1984).

Intra-examiner: There was a poor level of agreement between the panel of 74 examiners when assessing deviation from normal occlusion (Kappa Statistic: 0.39–0.87; Fig. 10).

Inter-examiner: Agreement between examiners in the five practitioner groupings was assessed using the Kappa Statistic (Table 14). There was only moderate agreement between examiners when assessing deviation from normal occlusion indicating a division in assessments between examiners. The decisions of the general dental practitioners and the community

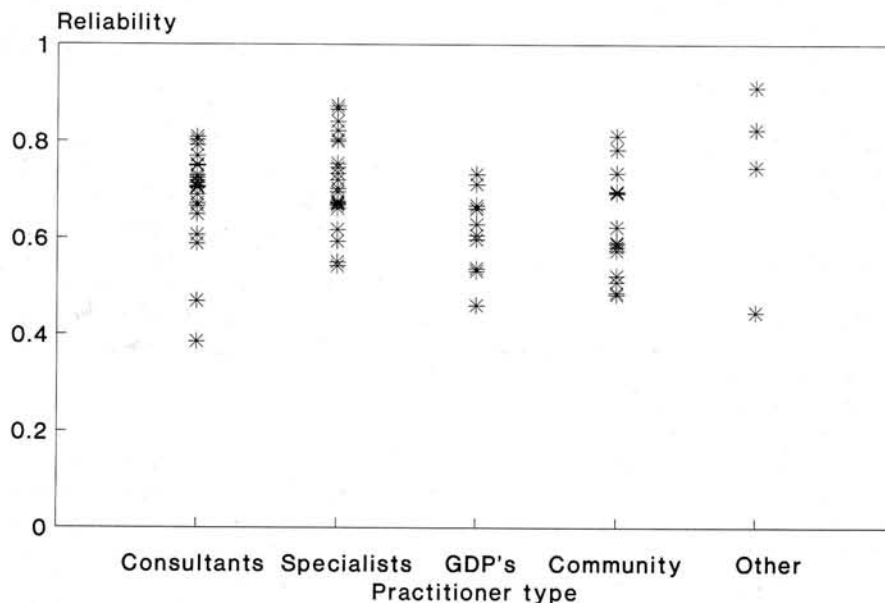


Figure 10 Intra-examiner reliability for the panel of 74 examiners deviation from normal occlusion ratings.

Table 14 Inter-examiner agreement for the five groups of practitioners (Kappa statistic).

	Consultants	Specialists	GDP's	Community	Other
Deviation from normal	0.52	0.57	0.50	0.44	0.59

dentists showed the poorest agreement with other practitioners in their respective groups.

Discussion

The primary objective was to validate the PAR Index. Two-hundred-and-seventy-two dental casts were assessed by a panel of 74 examiners representing all interested parties. The mean scores were obtained from the panel of 74 examiners to reflect contemporary orthodontic opinion.

The components of the PAR Index were weighted statistically to reflect consensus opinion. Several of the individual components were not included in the weighted index as they did not have any predictive power. One possible explanation for the exclusion of the buccal segments (i.e. spacing, crowding, and impactions) may be that the contact points between the buccal teeth were so variable. If the displacement of a tooth was severe, a cross-bite would result and, therefore, be recorded in a score for left or right buccal occlusion. However, there were a small number of cases with impacted premolars and it could be that the panel of 74 dentists missed these impactions or that there were too few cases to produce a significant contribution. Premolar impactions would not be recorded in the weighted PAR Index although impacted canines would, as these are scored in the anterior segment.

Residual extraction spaces are more likely to be accepted in the U.K. than in America (Elderton and Clark, 1983) and it may be argued that the panel of 74 examiners was willing to accept extraction spaces.

The overjet, overbite, and centreline discrepancy were weighted highly (6, 2, and 4, respectively). The weighted factor of 6 for overjet appears high. An overjet of 6 mm would have a PAR score of 2, and a weighted PAR score of 12. In fact, the overjet score is similar to the overjet in millimetres multiplied by a factor of 2. The range of weightings and the equivalent weighting for an overjet in millimetres is shown in Table 15. The overjet score also includes teeth in cross-bite.

Table 15 Comparison of weighted PAR scores for overjet with the equivalent weighting for the overjet in mm.

Overjet (mm)	PAR score	Weighted PAR score	Equivalent weighting for overjet (mm)
0	0	0	0.0
1	0	0	0.0
2	0	0	0.0
3	0	0	0.0
4	1	6	1.5
5	1	6	1.2
6	2	12	2.0
7	2	12	1.7
8	3	18	2.3
9	3	18	2.0
10	4	24	2.4
11	4	24	2.2
12	4	24	2.0
13	4	24	1.9
14	4	24	1.7
15	4	24	1.6

The weighting for centreline discrepancy was higher for consultants and general dental practitioners than the specialist and community dentists. It would seem that the consultants and general dental practitioners were more likely to notice differences in the centrelines.

There was a poor level of agreement between the panel of the 74 examiners when assessing deviation from normal. This division highlights the need for an objective and reliable index.

The weightings which have been derived reflect current British orthodontic opinion. The PAR Index is flexible in that the weightings could be changed to reflect future standards and the standards that are currently being achieved in other countries.

Conclusions

The PAR Index provides a single summary score for all the occlusal anomalies and may be used for all types of malocclusions, treatment modalities and extraction/non-extraction cases. The score provides an estimate of how far a case deviates from normal and the difference in

scores for pre- and post-treated cases reflects the perceived degree of improvement and therefore the success of treatment.

Excellent reliability within and between four trained examiners has been demonstrated using the PAR Index. The PAR Index offers uniformity and standardisation in assessing the outcome of orthodontic treatment. The components of the PAR Index have been weighted to reflect current British dental opinion more closely.

Address for correspondence

Dr Stephen Richmond
Department of Oral Health and Development
Coupland III Building
University of Manchester
Higher Cambridge Street
Manchester M15 6FH

Acknowledgements

The authors are grateful to the British Orthodontic Standards Working Party whose members were actively involved in developing the PAR Index and for the Panel of 74 examiners who gave up their valuable time to record the 320 dental casts. The authors would also like to acknowledge the financial support of the Dental Practice Board.

References

- Barkto J J 1966 The intraclass correlation coefficient as a measure of reliability. *Psychology Report* 19: 3–11
- Berg R 1979 Post-retention analysis of treatment problems and failures in 264 consecutively treated cases. *European Journal of Orthodontics* 1: 55–68
- Buchanan I D 1990 The PAR versus the Summers index. Presentation at the British Orthodontic Conference, Torquay 1990
- Eismann D 1974 A method of evaluating efficiency of orthodontic treatment. *Transactions of the European Orthodontic Society*, 223–232
- Eismann D 1980 Reliable assessment of morphological changes results from orthodontic treatment. *European Journal of Orthodontics* 2: 19–25
- Elderton R J, Clark J D 1983 Orthodontic treatment in general dental service assessed by the occlusal index. *British Journal of Orthodontics* 10: 178–186
- Elderton R J, Clark J D 1984 An investigation of treatment in the General Dental Service for patients with Class II division I malocclusions. *British Journal of Orthodontics* 11: 2–8
- Fleiss J L 1981 *Statistical methods for rates and proportions*, 2nd ed. Wiley, London

- Fleiss J L 1986 *The design and analysis of clinical experiments*. Wiley, London
- Fleiss J L, Slakter M J, Fischmann S L, Park M H, Chilton N W 1979 Inter examiner reliability in caries trials. *Journal of Dental Research* 58: 604–609
- Gottlieb E L 1975 Grading your orthodontic treatment results. *Journal of Clinical Orthodontics* IX: 156–161
- Hickham J H 1975 Directional edgewise orthodontic approach. *Journal of Clinical Orthodontics* IX: 143–154
- Jenny J, Cons N C, Kohout F J 1983 Comparison of SASOC, a measure of dental aesthetics, with three orthodontic indices and orthodontic judgement. *Community Dentistry and Oral Epidemiology* 11: 236–241
- Malmgren O 1980 Studies on the need and demand for orthodontic treatment. *Swedish Dental Journal Supplement* 6
- Mitropoulos C M, Downer M C 1987 Inter-examiner agreement in the diagnosis of dental caries among officers of the reference service. *British Dental Journal* 162: 227–228
- O'Connell D L, Dobson A J 1984 General observer agreement measures on individual subjects and groups of subjects. *Biometrics* 40: 973–983
- Pickering E A, Vig P 1975 The occlusal index used to assess orthodontic treatment. *British Journal of Orthodontics* 2: 47–51
- Salzmann J A 1968 Handicapping malocclusion assessment to establish treatment priority. *American Journal of Orthodontics* 54: 749–765
- Shaw W C, Addy M, Dummer P M H, Ray C, Frude N 1986 Dental and social effects of malocclusion and effectiveness of orthodontic treatment: a strategy for investigation. *Community Dentistry and Oral Epidemiology* 14: 60–64
- Summers C J 1971 The occlusal index: a system for identifying and scoring occlusal disorders. *American Journal of Orthodontics* 59: 552–566

Appendix 1: Conventions for the PAR Index

General

1. All scoring is accumulative.
2. There is no maximal cut-off level.
3. The occlusion should be scored disregarding functional displacement (this cannot be determined from dental casts alone).
4. The contact points between first, second, and third molars are not recorded. The contact points between molars are so variable, however, severe deviations will produce a cross-bite, and will be noted in the buccal occlusions.
5. If the contact point displacement is as a result of poor restorative work (restorations or crowns), the displacement is not recorded.

6. Contact points between deciduous teeth are not recorded.
7. Extraction spaces are not recorded if the patient is to receive a prosthetic replacement. However, if space closure is intended, the distance between adjacent teeth should be noted.

Canines

1. Where there are missing canines, displacements resulting from discrepancies between the mesial contact point to the first premolar and the distal of the lateral incisor should be recorded in the anterior segment.
2. Canine cross-bites should be recorded in the overjet section.
3. Contact points between the canines and premolars are scored as follows: the distal contact point of the canine to the midpoint on the mesial surface of the adjacent premolar. (These contact points are so variable. When untreated normal occlusions were assessed this relationship seemed to be the most acceptable.)

Impactions

If a tooth is unerupted and displaced from the line of the arch either buccally or palatally

due to insufficient space, this is regarded as an impaction. However, if the tooth is erupted and displaced, the displacement score is recorded.

Incisors

1. If there is agenesis of the upper incisor or the tooth has been lost due to trauma or caries the procedure is as follows:
 - (a) if the space is maintained (for a prosthesis), the distance between adjacent teeth is not recorded;
 - (b) if the space is to be closed, the distance between adjacent teeth is recorded.
2. When recording an overjet, if the tooth falls on the line the lower grade is recorded.
3. If a lower incisor has been extracted or is missing, the centreline is not recorded.

Molars

1. Contact points between first and second molars are not recorded.
2. If the first molars have been extracted, the contact point of the second molar is recorded.