

The Relationship between Posttreatment Smile Esthetics and the ABO Objective Grading System

Brian J. Schabel^a; James A. McNamara, Jr^b; Tiziano Baccetti^c; Lorenzo Franchi^c; Scott A. Jamieson^d

ABSTRACT

Objective: To evaluate the correlations between the components of the Objective Grading System developed by the American Board of Orthodontics (ABO) and smile esthetics.

Materials and Methods: The clinical photographs of 48 orthodontically treated patients were rated by a panel of 25 experienced orthodontists (15 men and 10 women) and 20 parents of orthodontic patients (eight men and 12 women). One of the investigators, a director of the ABO, scored the posttreatment dental casts and panoramic radiographs of all patients according to the guidelines of the Objective Grading System. The relationship between the occlusal outcome and perceived smile attractiveness of the subjects were evaluated by a Pearson product-moment correlation. Logistic regression was used to determine whether the individual component or total combined scores of the ABO Grading System could predict whether a smile would be considered “attractive” or “unattractive” by the panel of raters.

Results: Extremely weak relationships were found among all factors of the ABO Objective Grading System and perceived smile attractiveness (r values ranging from -0.11 to 0.14 ; $P > .05$). As derived from logistic regression equations, neither total scores nor individual components of the ABO Grading system could predict attractive or unattractive smiles.

Conclusions: This study suggests that additional criteria might be incorporated into the assessment of overall orthodontic treatment outcomes, including variables evaluating the smile.

KEY WORDS: Smile esthetics; American Board of Orthodontics; Digital imaging; Q-Sort

INTRODUCTION

A so-called “successful” treatment result has many meanings in orthodontics. To the patient, success is typically determined predominantly by an esthetic out-

come. According to the Objective Grading System developed by the American Board of Orthodontics (ABO), however, a successful treatment is determined by a functionally placed occlusion that results in beautifully articulated dental casts with parallel roots on the panoramic radiograph.¹ Is the orthodontist at conflict between the occlusal objectives set forth by the ABO and esthetic demands of patients, or by satisfying one objective, does one concurrently achieve the other?

Although the use of the Peer Assessment Rating (PAR) Index² and the Index of Complexity, Outcome, and Need (ICON)³ to evaluate the need for orthodontic treatment has gained popularity in Europe, the orthodontic community in the United States has questioned the value of these occlusal indices. To date, the American Association of Orthodontists does not recognize any index for the purpose of classifying treatment need. The ABO, however, has developed a valid and reliable index that is used specifically to assess treatment outcome. The directors of the ABO developed the Objective Grading System¹ as a means for both the ABO and clinicians to evaluate the Candidate Case Report Examination (CCRE) submitted for

^a Private practice, Santa Cruz, Calif.

^b Thomas M. and Doris Graber Endowed Professor of Dentistry, Department of Orthodontics and Pediatric Dentistry, School of Dentistry; Professor of Cell and Developmental Biology, School of Medicine; and Research Professor, Center for Human Growth and Development, The University of Michigan, Ann Arbor. Private practice of orthodontics, Ann Arbor, Mich.

^c Assistant Professor, Department of Orthodontics, The University of Florence, Florence, Italy; Thomas M. Graber Visiting Scholar, Department of Orthodontics and Pediatric Dentistry, School of Dentistry, The University of Michigan, Ann Arbor.

^d Private practice, Marquette, Michigan; Director, American Board of Orthodontics

Corresponding author: Lorenzo Franchi, DDS, PhD, Department of Orthodontics, Università degli Studi di Firenze, Via del Ponte di Mezzo, 46-48, 50127, Firenze, Italy (e-mail: lorenzo.franchi@unifi.it)

Accepted: September 2007. Submitted: August 2007.

© 2008 by The EH Angle Education and Research Foundation, Inc.

Table 1. The American Board of Orthodontics Objective Grading System: Point deductions for the individual components

Component	Deduction
Alignment	
< 0.5 mm	0
0.5 to 1 mm	1
> 1 mm	2
Marginal ridge height	
< 0.5 mm	0
0.5 to 1 mm	1
> 1 mm	2
Buccolingual inclination	
< 1 mm	0
1 to 2 mm	1
> 2 mm	2
Occlusal contacts	
0 mm	0
< 1 mm	1
> 1 mm	2
Occlusal relationships	
< 1 mm	0
1 to 2 mm	1
> 2 mm	2
Overjet	
0 mm	0
Less than 1 mm	1
> 1 mm	2
Interproximal contacts	
< 0.5 mm	0
0.5 to 1 mm	1
> 1 mm	2
Root angulation	
Root parallelism	0
Roots are not parallel	1
Contacting adjacent tooth	2

Phase III (clinical examination) of the board-certification process.

The ABO established this grading system for examiners and candidates to evaluate clinical expertise and technical acuity according to the following 8 characteristics measured from the final dental casts and panoramic radiographs: (1) alignment, (2) marginal ridge height, (3) buccolingual inclination, (4) occlusal relationships, (5) occlusal contacts, (6) overjet, (7) interproximal contacts, and (8) root angulation. Post-treatment dental casts and the final panoramic radiographs are inspected and/or measured according to the aforementioned characteristics, and points are deducted (–1 or –2) according to how far the individual teeth deviate from the standards established by the ABO (Table 1). The total point deductions for each of the 8 categories are summed to give each treated case an overall score that determines whether or not the candidate passes the clinical case report portion

of the phase III examination. In general, a case report that loses 30 or more points will be judged incomplete in ABO terminology; a case report that loses 20 or fewer points will pass.¹ Case reports that fall between these ranges are reevaluated and are passed or given a status of incomplete according to the discretion of the Board.

An index such as the ABO grading system is considered valid for assessing posttreatment outcomes in terms of objective occlusal features as appraised by experienced orthodontists. On the other hand, the grading system does not specifically take into account any soft tissue measures. It remains to be tested whether the Objective Grading System is able to capture posttreatment smile esthetics as a fundamental component of orthodontic treatment outcomes.

The objective of the present study, therefore, was to evaluate the correlations between the components of the ABO Objective Grading System and smile esthetics observed after orthodontic treatment. These data will be helpful in understanding whether the criteria established by the ABO result in beautiful occlusions as represented statically in study models and/or beautiful smiles, or whether additional criteria should be included in a more global evaluation of treatment outcome.

MATERIALS AND METHODS

Patient Selection

The subjects enrolled in this study were recruited from the Graduate Orthodontic Clinic during a routine posttreatment appointment (ie, final records or retention check). Each adult subject (≥ 18 years old) reviewed and signed a consent form created in accordance with guidelines of the Health Sciences Institutional Review Board (IRB). Each subject under the age of 18 years reviewed and signed a child's assent form, and a legal guardian reviewed and signed a consent form in accordance with the IRB. Each subject also reviewed and completed a consent form created by the university in accordance with the Health Insurance Portability and Accountability Act for the use and disclosure of protected health information. None of the subjects received compensation for their participation.

To be included in the study, patients had to present with the following characteristics: (1) age ranging from 12 to 20 years; (2) Caucasian ancestry; (3) orthodontic treatment completed within the previous 6 months; (4) absence of missing or malformed teeth; and (5) a complete set of diagnostic posttreatment records, including study models, panoramic radiograph, and intra- and extraoral photographic series.

The protocol used in this investigation required that 48 subjects be recruited to satisfy the design of the Q-sort. A test was performed to determine the power



Figure 1. A standardized smile image using the 3-in \times 5-in template.

of this sample size with respect to correlation tests (Type I error = 0.05). For a bivariate normal distribution and a sample size of 48, a test of $H_0: P = 0$ (ie, the correlation coefficient under the null hypothesis) was found to have a power of 0.80 to detect a linear correlation of $r = 0.38$. Thus, the default sample size for the Q-sort procedure was deemed adequate for purposes of testing for correlation.

Image Capture

All photographs were taken by 1 of 2 staff photographers. The extraoral photographic series included photographs of the subject in repose, during smiling, and in profile. For the purpose of the current study, only the extraoral smiling photographs were used. A Canon EF 35 mm SLR camera (Canon USA Inc., Lake Success, NY) was mounted to a frame set at a fixed distance of 36 inches between the lens and the subject. The camera was connected to a 1-strobe lighting source that illuminated the subject indirectly from a flash that reflected off a photographic umbrella. Before taking the smiling image, the photographer instructed the subject to "smile." The reproducibility of the posed smile derived from static photograph has been demonstrated by Ackerman et al.⁴ Each image was captured on Kodak EV-100 slide film (Eastman Kodak Co., Rochester, NY).

The film was developed and the 2-in \times 2-in slides were scanned using the Nikon Super Coolscan 4000 ED (Nikon Inc., Melville, NY) and imported directly into a commercially available image editing software program (Adobe Photoshop 7.0, Adobe Systems Inc., San Jose, CA). Each slide was scanned at maximum dots per inch to enhance the image quality. A 3-in \times 5-in template was created to standardize the size and location of each image. Images were opened in Photoshop, and the template was superimposed on top of the image (Figure 1). The smile images were enlarged until the outer commissures of the lips matched the

vertical tick-marks inset three-quarters of an inch from the border of the template. The smiling images were then positioned so that the maxillary incisal edges coincided with the horizontal line of the template.

After enlarging and positioning the images, the portion of the image outside the template was cropped. The resulting images were edited further in Photoshop using the healing brush tool to remove blemishes, skin irregularities, or other extraneous marks that could influence the rater when evaluating the image. Images were labeled with a unique 4-digit number chosen randomly. Once the editing was complete, each image was compressed to approximately 150 KB and saved as a JPEG file.

Rater Selection

Both orthodontists and parents of orthodontic patients rated the smile images. The orthodontic panel was comprised of 25 full- and part-time orthodontic faculty members recruited from the University of Michigan Graduate Orthodontic Program. This panel consisted of 15 male and 10 female orthodontists whose clinical experience ranged from 1 to 35 years, with an average of 17 years. Their ages ranged from 29 to 64 years, with an average age of 46 years. Sixteen of the panelists were certified by the ABO. Of the 25 orthodontic raters, 13 were graduates of the University of Michigan and two had received their orthodontic specialty training internationally.

The nonorthodontic panel comprised 20 parents whose children were currently undergoing orthodontic treatment in the Graduate Orthodontic Clinic. This panel consisted of 12 women and eight men whose ages ranged from 36 to 52 years, with an average of 43 years. Of the 20 parent panel members, nine had undergone orthodontic treatment themselves.

Because a related study has shown that orthodontists and lay panelists score smile images in a similar way with regard to the level of attractiveness,⁵ the findings of the two groups of raters were pooled ($N = 45$) for the purpose of this study.

The Q-Sort method

The Q-Sort method, originally proposed by Stephenson in 1953,⁶ represents an alternative approach that generates a valid 9-category ordinal ranking of large samples based on a variety of subjective criteria. The Q-sort method uses a progressive forced-choice winnowing of the sample to create a quasnormal distribution, in this instance rating smile images on an esthetic scale from "least pleasing" to "most pleasing." Although originally intended for a sample size of 96, the Q-sort technique has been applied to samples of

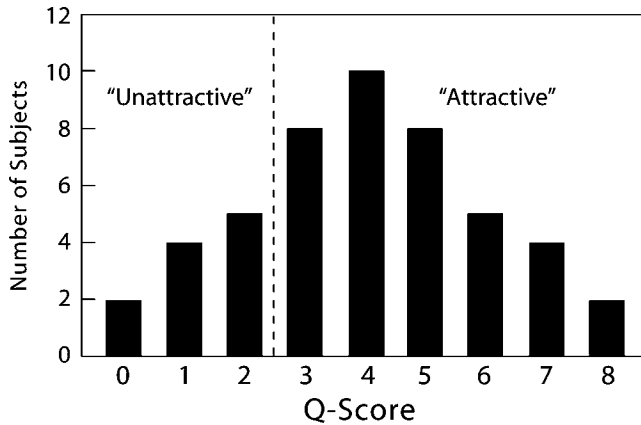


Figure 2. Q-sorting: assignment of scores to the cutoff point used to separate “attractive” from “unattractive” smiles in the Q-sort distribution. Note: a line between columns 2 and 3 was given a numeric score of 2.5.

48 to rank items according to complex, subjective criteria.⁷

Panelists were asked to apply the Q-sort technique to rate the attractiveness of the series of 48 subjects on the smile images captured with clinical photography. Rating sessions for the parents coincided with consecutively scheduled orthodontic appointments, while the rating sessions for the orthodontists occurred sporadically over a 2-month period.

Each panelist was asked to select and organize groups of images in a specific order (Figure 2). From the 48 images, the two least and most attractive smiles were identified and placed into their respective columns. The raters then selected the four least and most attractive smiles, followed in a similar manner by the five and then the eight smile photographs from each extreme, leaving the 10 smiles presumably of neutral attractiveness. Each rater was also asked to survey the distribution and establish a cutoff point between two columns separating “unattractive” from “attractive” smiles.

The assignment of scores to each Q-sort distribution is depicted in Figure 2. The X-axis represents the score assigned to each photograph in a given group; and the Y-axis, the number of subjects. Each of the nine groups of images were given a score ranging from 0, for the two images with the least perceived attractiveness, to 8, for the two images with the greatest perceived attractiveness. The scores each subject received from the various judges were averaged to generate each subject’s overall Q-score for the smile captured with clinical photography.

The cutoff point separating the unattractive and attractive images was also given a numeric value. For example, a line drawn between columns 2 and 3 on the Q-sort distribution was given a score of 2.5 (Figure

2), which represented the esthetic boundary for that panel member. The cutoff points received from the various judges were averaged to generate the overall demarcation between attractive and unattractive images.

ABO Grading System Assessment

One of the investigators (Dr Jamieson, a calibrated grader and a director of the ABO) scored the posttreatment dental casts and panoramic radiographs of all subjects according to the guidelines of the ABO Objective Grading System.¹ Specifically, study models were evaluated according to the following characteristics: alignment, occlusal contacts, marginal ridges, occlusal relationships, buccolingual inclination, overjet, and interproximal contacts; panoramic radiographs were evaluated for root parallelism. The individual component and combined scores generated from the ABO Objective Grading System of each subject served as a measure of each subject’s occlusal outcome.

Statistical Analysis

Standard descriptive statistics (means, standard deviations, and ranges) were calculated for age and for Q-sort scores, esthetic boundary cutoff scores, and components of the ABO Objective Grading System. The relationship between the occlusal outcome and perceived smile attractiveness of the subjects was evaluated by means of a Pearson product-moment correlation. Specifically, the individual component and total combined scores of the ABO Grading System were evaluated against the average combined Q-scores of smiles captured with clinical photography. Logistic regression was used to evaluate whether the individual component or total combined scores of the ABO Grading System could predict whether a smile would be attractive or unattractive.

The Type I error rate for all statistical tests was set at 0.05. All statistical tests were performed with the aid of a statistical software program (SPSS; Statistical Package for the Social Science for Windows, version 12.0, Chicago, Ill).

RESULTS

Standard descriptive statistics were calculated for the average Q-scores of the various images and the esthetic boundary scores (Table 2). Average Q-scores were calculated by combining the results of the 45 raters for each subject. Esthetic boundary scores represent the average cutoff point specified as a demarcation between unattractive and attractive photographs during the Q-sort assessment. Because of the

Table 2. Descriptive statistics for average Q-sort and esthetic boundary scores of the various raters for each image type

Smile Photo Variable and Rater	Mean	SD	Range	
			Minimum	Maximum
Average Q-sort score (range 0–8) ^a				
Orthodontist	4	1.4	0.5	6.7
Parent	4	1.4	0.7	7.1
Average esthetic boundary (range 0–8)				
Orthodontist	3.2	1.3	0.5	6.5
Parent	3.7	1.7	1.5	6.5

^a The normal distribution of the Q-sort results in a mean of 4 when subjects are combined.

Table 3. Descriptive statistics for point deductions of the individual components and total scores from the American Board of Orthodontics (ABO) Objective Grading System

Component of the ABO Grading System	Mean	SD	Range	
			Minimum	Maximum
Alignment	3.4	1.8	0	8
Occlusal contacts	1.4	2.6	0	15
Marginal ridges	4.6	2.6	0	12
Occlusal relationship	1.8	2.7	0	10
Buccolingual inclination	3.6	2.6	0	12
Overjet	1.8	2.2	0	8
Interproximal contacts	0.2	0.6	0	2
Root angulation	2.5	1.8	0	8
Total score	19.3	7.6	5	42

ordinal nature and normal distribution of the Q-sort, the mean Q-sort score is 4 when scores of all subjects are combined. Q-scores and esthetic boundary scores ranged from 0 to 8.

Standard descriptive statistics were also calculated for the various components of the ABO Objective Grading System for each subject that participated in the study. The means, standard deviations, and ranges of the point deductions of the individual components and the total scores are summarized in Table 3.

Pearson correlation coefficients were calculated to characterize the relationship between various components of the ABO Grading System and average Q-scores of smiles captured with clinical photography (Table 4). It should be pointed out that a perfect occlusal outcome according to the guidelines of ABO Objective Grading System would receive a score of zero; therefore, a perfect correlation between optimal occlusion and a beautiful smile would result in a correlation of -1 . Extremely weak positive and negative relationships were found between all factors of ABO Objective Grading System and perceived smile attractiveness (r values ranging from -0.11 to 0.14 ; $P > .05$).

The coefficients for the logistic regression equations

Table 4. Pearson Correlation between the components of the American Board of Orthodontics (ABO) Objective Grading System and average Q-sort scores of smiles captured with clinical photography

Component of the ABO Grading System	Average Q-sort Score of Smiles Captured with Clinical Photography	
	Correlation (r)	P
Alignment	.14	.34
Occlusal contacts	-.11	.46
Marginal ridges	.13	.39
Occlusal relationships	.11	.46
Buccolingual inclination	-.06	.67
Overjet	-.05	.74
Interproximal contacts	.08	.57
Root angulation	-.07	.62
Total score	.03	.84

Table 5. Logistic regression: Regression coefficients (B) and probabilities (P) for individual components and total scores with “attractive” and “unattractive” smiles captured with clinical photograph as the dependent variable

Component of the ABO Grading System	B	SE (B)	P
Alignment	-0.02	0.98	.92
Occlusal contacts	-0.12	0.88	.30
Marginal ridges	0.01	1.00	.98
Occlusal relationships	0.06	1.06	.64
Buccolingual inclination	0.02	1.02	.86
Overjet	0.17	1.18	.29
Interproximal contacts	0.29	1.34	.63
Root angulation	0.03	1.03	.85
Total score	0.01	1.01	.82

with “attractive” and “unattractive” as the dependent variable and the individual component and total scores of the ABO Objective Grading System as the predictors are summarized in Table 5. Neither total scores nor individual components of the ABO Grading system could predict attractive or unattractive smiles.

DISCUSSION

As part of an ongoing attempt to define and standardize the criteria for attaining Diplomate status, the ABO established a grading system for dental casts and panoramic radiographs.¹ This grading system was devised as a means of providing clinicians with an objective index to evaluate treatment outcome. In that the mission of the ABO, stated in 1997,⁸ is “to establish the highest standards of clinical excellence in orthodontics,” the present study sought to examine the relationship between occlusal outcomes (according to the guidelines of the Objective Grading System) and posttreatment smile esthetics.

Some remarks concerning the methodology used need to be emphasized before discussing the findings of the current study. The Q-sort technique was used

as a previous study indicated a higher reliability for this method than the Visual Analog Scale when analyzing the esthetics of the smile.⁸ The same investigation reported that orthodontists and parents of orthodontic patients agree with respect to attractive and unattractive smiles. Therefore, the results of the 2 panels of raters were pooled in the current study for the evaluation of perceived smile attractiveness.

Pearson correlations were found to be extremely poor between the 8 components of the ABO Objective Grading System and average Q-scores of smiles captured with clinical photography by the combined panel of orthodontists and parents of orthodontic patients (r values ranging from -0.11 to 0.14). No relationship could be inferred between the total scores derived from the Objective Grading System and the Q-scores of smiles captured with clinical photography ($r = 0.03$; $P = .84$). Logistic regression also revealed that neither individual components nor total scores derived from the ABO Objective Grading System were capable of predicting whether a smile would be deemed attractive or unattractive. The combined results of the logistic regression and the weak correlations between the components of the ABO Objective Grading System and smile esthetics demonstrate that beautifully articulated dental casts with parallel root relationships do not necessarily result in beautiful smiles. These results are not surprising, however, in that the current Objective Grading System does not evaluate any soft tissue parameters, either intraorally or extraorally.

Although no published study has examined the relationship between the ABO Objective Grading System and posttreatment smile esthetics, numerous investigators have speculated about the interaction between occlusal outcomes and esthetics. It has been suggested that perfect occlusion does not necessarily result in desirable dentofacial features.^{4,9} Others, however, have stated that ideal occlusion and facial beauty are interdependent.¹⁰⁻¹² The results of the present analysis suggest that ideal posttreatment occlusion and posttreatment smile esthetics are not related. Considering that the smile is what most laypeople use to judge treatment success,¹³ orthodontists should strive to establish optimal occlusal, facial, and smile esthetic outcomes that would most benefit the individual patient.¹⁴

CONCLUSIONS

- The present study did not detect any correlation between components of the ABO Objective Grading System and the esthetics of the smile.

- None of the individual components or total scores of the ABO Objective Grading System can be used as a predictor of an attractive or unattractive smile after orthodontic treatment.
- In that the directors of the ABO continually attempt to refine the criteria for what is considered a successfully treated orthodontic outcome, the future inclusion of such variables as soft tissue measures, especially of smile esthetics, may be beneficial.

ACKNOWLEDGMENTS

Support for this project was derived from monies made available through the Thomas M. and Doris Graber Endowed Professorship, The University of Michigan. Additional funding was made available through the LeGro Fund at the same institution.

REFERENCES

1. Casco JS, Vaden JL, Kokich VG, et al. Objective grading system for dental casts and panoramic radiographs. *Am J Orthod Dentofacial Orthop.* 1998;114:589-599.
2. Richmond S, Shaw WC, O'Brien KD, et al. The development of the PAR index (Peer Assessment Rating): reliability and validity. *Eur J Orthod.* 1992;14:125-139.
3. Daniels C, Richmond S. The development of the Index of Complexity, Outcome, and Need (ICON). *J Orthod.* 2000;27:149-162.
4. Ackerman JL, Ackerman MB, Brensinger CM, Landis JR. A morphometric analysis of the posed smile. *Clin Orthod Res.* 1998;1:2-11.
5. Schabel BJ, McNamara JA Jr, Franchi L, Baccetti T. Q-Sort Assessment vs Visual Analog Scale in the evaluation of smile esthetics. *Am J Orthod Dentofacial Orthop.* In press.
6. Stephenson W. *The Study of Behavior.* Chicago, IL: The University of Chicago Press; 1953.
7. Coy K, Speltz ML, Jones K. Facial appearance and attachment in infants with orofacial clefts: a replication. *Cleft Palate Craniofac J.* 2002;39:66-72.
8. American Board of Orthodontics. *Information for the Candidates.* 4th ed. St Louis, MO: American Board of Orthodontics; 1997.
9. Riedel RA. Esthetics and its relation to orthodontic therapy. *Angle Orthod.* 1950;20:168-178.
10. Angle EH. Classification of malocclusion. *Dent Cosmos.* 1899;41:248-264.
11. Downs WB. Analysis of the dento-facial profile. *Angle Orthod.* 1956;26:191-212.
12. Steiner CC. Cephalometrics in clinical practice. *Angle Orthod.* 1959;29:8-29.
13. Espeland LV, Stenvik A. Perception of personal dental appearance in young adults: relationship between occlusion, awareness, and satisfaction. *Am J Orthod Dentofacial Orthop.* 1991;100:234-241.
14. Ackerman JL, Proffit WR, Sarver DM. The emerging soft tissue paradigm in orthodontic diagnosis and treatment planning. *Clin Orthod Res.* 1997;2:49-52.