ORIGINAL ARTICLE

Early treatment for Class II malocclusion and perceived improvements in facial profile

Kevin O'Brien, Tatiana Macfarlane, Jean Wright, Frances Conboy, Priscilla Appelbe, David Birnie, Stephen Chadwick, Ivan Connolly, Mark Hammond, Nigel Harradine, David Lewis, Simon Littlewood, Catherine McDade, Laura Mitchell, Alison Murray, Julian O'Neill, Jonathan Sandler, Micheal Read, Stephen Robinson, Iain Shaw, and Elizabeth Turbill

Manchester, United Kingdom

Introduction: The aims of this study were to assess whether early Twin-block appliance treatment improves the attractiveness of Class II profiles and to determine the orofacial features of a profile that most influence the perception of attractiveness. **Methods:** Silhouetted profiles of 20 treated patients and 20 untreated controls randomly selected from 174 subjects (ages, 8-10 years) of a randomized, controlled trial into the effectiveness of early Class II treatment were assessed by 30 children (ages, 10-11 years) and 24 teaching staff using a 5-point Likert scale. Independent samples *t* tests were used to compare attractiveness ratings between the treated and untreated groups. Linear regression was used to determine the features defining attractiveness. **Results:** Early orthodontic treatment resulted in improved perceptions of facial profile attractiveness. Profiles were likely to be rated as attractive if the overjet was smaller ($P = 0.001$) and no teeth showed ($P < 0.05$). **Conclusions:** Profile silhouettes of children who had received early orthodontic treatment for Class II malocclusion were perceived to be more attractive by peers than those of children who did not receive treatment. (Am J Orthod Dentofacial Orthop 2009;135:580-5)

There is no doubt that a main goal of orthodontic treatment is to improve the dentofacial attractiveness of our patients. Although these aims are part of our treatment goals, there is currently little evidence from high-quality investigations to underpin these benefits. We addressed this problem in this study by measuring the effect of early functional appliance (phase 1) treatment for Class II malocclusion on the perception of facial profile attractiveness from groups of the children's peers and teachers.

Our literature review was confined to evidence from randomized, controlled trials and systematic reviews. Our electronic literature search found only 1 randomized clinical trial about the influence of functional appliance treatment on the facial profile.[1] This study was in New Zealand, and 42 children (age, 11 years) were randomly allocated to receive treatment with a Harvold activator (12) or a Fränkel appliance (13), with 17 children as untreated controls. Cephalometric radiographs were taken at the start of the study and 18 months later, and silhouettes were constructed from the soft-tissue outlines of lateral cephalograms, by using image manipulation software. Then, groups of 30 fifth-year dental students, 30 art students, and 30 parents of patients undergoing orthodontic treatment rated the attractiveness of the profiles using visual analog scales. The authors concluded that there were no significant differences among the 3 groups and that treatment with functional appliances does not result in more attractive facial profiles.

Although this study provided useful information, only 1 group of raters comprised laypeople, and they might have been attuned to orthodontic treatment, because they were parents of children who were current patients. This is relevant to studies of this type because the perceptions of lay judges are important, since they represent the patients' social groups. Furthermore, innate reactions of schoolmates and teachers are most likely to influence a child's self-esteem and social interactions, and children's attractiveness has been shown to influence teachers' expectations,[2] which in turn correlate with future educational achievement.[3] As a result, we decided to measure the attractiveness of facial profiles using panels of the children's peers and teachers.

This study had the following aims: to determine whether Twin-block functional appliance treatment

From the United Kingdom Class II study group, School of Dentistry, University of Manchester, Manchester, United Kingdom.

Funded by the Medical Research Council (G9410454).

The authors report no commercial, proprietary, or financial interest in the products or companies described in this article.

Reprint requests to: Kevin O'Brien, School of Dentistry, University of Manchester, Higher Cambridge Street, Manchester M15 6FH, United Kingdom; e-mail, Kevin.O'Brien@manchester.ac.uk.

Submitted, June 2007; revised and accepted, February 2008.

0889-5406/$36.00

doi:10.1016/j.ajodo.2008.02.020

when the child is 8 to10 years old (in the transitional dentition) improves the visual appeal of the facial profile as perceived by peers and teachers, and to identify the orofacial features that influenced the raters' perceptions.

MATERIAL AND METHODS

The first stage of the investigation was a small pilot study to generate data that we could use to calculate the sample size. One orthodontist (E.T.) and two lay members (F.C. and J.W.) of the research team assessed a sample of the patients, and sample size was calculated on the basis of numbers needed to show a difference between treated and control subjects of 1 grade (assuming 1 SD) at a 2-way α of 0.05, with power of 80%. This showed that we needed a sample of at least 34 profiles (17 treated, 17 untreated).[4] In addition, for reliability analysis, under the null hypothesis H_0: $\rho_0 = 0$ and alternative hypothesis H_1: $\rho_0 = 0.1$, with $\alpha = 0.05$ and $\beta = 0.2$, we required at least 9 raters assessing the 34 profiles (NQuery Advisor, v 5.0, Statistical Solutions, Saugus, Mass).

Ethical committee approval was obtained.

We randomly selected the patient photographs from the records of 174 Class II subjects from our multi-center, randomized, controlled trial into the effects of early orthodontic treatment for Class II malocclusion.[5] These children were aged 8 to 10 years with overjets of at least 7 mm. Photographs were taken at the start of the study (DC1) and 15 months later (DC2). The photographs were taken with the lips relaxed in natural head position. Half of our subsample had been treated with Twin-block appliances, and half were untreated controls. The subsamples were checked for homogeneity with the parent samples for anteroposterior skeletal discrepancies before and after treatment (*t* test: $P > 0.4$ for both subsamples).

We then converted the DC2 profile photographs of 20 control and 20 treated patients into silhouettes for the study. The profiles were orientated with the alar-tragus line horizontal before silhouetting to minimize the effects of head posture on apparent prognathia or retrognathia.

The assessors were a class of children (ages, 10-11 years; 14 boys, 16 girls) and 24 teachers and teaching assistants (all women) at an elementary school in Manchester, United Kingdom, with a broad ethnic and socioeconomic mix of pupils.

We showed the children and the teaching staff the 40 profiles in separate slide presentations at the school. We also included 5 duplicate slides to assess reliability. Thus, each group of raters assessed 45 slides. Rest slides (cartoons and landscapes) were inserted periodically to rest the assessors' eyes and refresh their concentration. The profiles were arranged in random order, and each was shown for 10 seconds. The teachers and children were asked to grade each profile using a 5-point Likert scale with anchors of 1, really like; 2, quite like; 3, OK; 4, don't like much; and 5, don't like. We stressed that they should record their first impressions of each profile.

We already had cephalometric data for skeletal discrepancy expressed as the difference between maxillary and mandibular length and overjet from the original data analysis of this study using the Pancherz analysis.[5,6]

We also recorded the following soft-tissue variables using a protractor and a ruler from the printed silhouettes: whether the incisor silhouettes were visible (yes/no), lips apart (yes/no), facial convexity (Holdaway angle), nasiolabial and labiomental angles, and lower anterior facial height proportion.

Statistical analysis

Descriptive statistics were used to assess the Twin-block and control subsamples at DC1 on the following data: sex, visibility of anterior teeth, skeletal discrepancy length difference, and overjet.

For the effect of Twin-block appliances on attractiveness, mean grades of all raters for each profile were calculated, and these means were compared between the Twin-block and control groups with independent samples *t* tests. The data were assessed for normality by using the Kolmogorov-Smirnov test.

We also assessed the factors that influence the perception of attractiveness. Associations were then sought between perceived appeal (mean grades of all raters) and skeletal discrepancy and overjet, facial convexity (Holdaway angle), nasolabial angle, labiomental angle, lower anterior facial height, overjet, lip competency, and visibility of teeth. The significant variables were analyzed with stepwise (backward) linear regression ($P < 0.1$).

Intraclass correlation was used to assess intraexaminer and interexaminer reliability.[7] In addition, mean ratings were compared between pupils and adults, and between boys and girls, with paired and independent *t* tests.

RESULTS

There were no differences between the samples for the data at DC1 (Table I), except that there were slightly more boys in the treated group, and more girls in the untreated group.

Table II shows the distribution of scores for all 54 raters and 40 profiles (2160 ratings in all) at DC2. Relatively few (20.9%) were rated as attractive (quite like or really like). Once mean scores (all raters) for each profile were calculated, the central tendency for the

Table I. Description of study subjects and their profiles at baseline (DC1)

Value	*Control group Mean (SD) or n (%)*	*Twin-block group Mean (SD) or n (%)*
Sex		
Male	8 (40%)	14 (70%)
Female	12 (60%)	6 (30%)
Skeletal discrepancy (mm)	0.07 (2.78)	0.44 (3.12)
Overjet (mm)	10.57 (2.15)	9.88 (2.48)

Table II. Distribution of scores (54 raters, 40 profiles)

Descriptor	*Frequency*	*%*
1. Really like	120	5.6
2. Quite like	331	15.3
3. OK	613	28.4
4. Don't like much	618	28.6
5. Don't like	478	22.1
Total	2160	100.0

whole sample was also somewhat toward the less-appealing end of the scale (mean, 3.46; SD, 0.72). There was no reason to believe that the mean scores were not normally distributed (Kolmogorov-Smirnov test, $P > 0.05$), so all calculations were made assuming normality.

For interrater reliability, Table III shows that the pupils rated the profiles more harshly than did the adults (mean difference, 0.65 grade; $P < 0.001$). Categorization of the strength of agreement (Landis and Koch[8]) showed moderate agreement between the raters. The mean scores of all raters were plotted against standard deviations (Fig). This figure shows that consensus improved as the profiles became less appealing ($R^2 = 0.35$).

For intrarater reliability, Table IV shows intraclass correlation coefficients for measuring agreement between the first and second ratings for the 5 duplicate profiles in the set of profiles. Categorization of the strength of agreement (Landis and Koch[8]) showed moderate agreement among the raters. There were no significant differences in levels of reproducibility between the rater groups.

The treated profiles were statistically significantly more appealing with a difference in means of 0.61 grade ($P = 0.006$) (Table V); they received more grades of 1 and 2, and fewer grades of 4 and 5.

Univariate analysis of the features that influence perceived attractiveness are shown in Table V. Profiles were less attractive (higher scores) when front teeth were showing ($P = 0.012$). There was negative correlation (-0.49, $P = 0.001$) also for overjet and skeletal discrepancy (-0.33, $P = 0.038$). There was evidence of negative correlation of scores with the labiomental angle (0.28, $P = 0.083$). These 4 variables were included in the linear regression model. There was evidence of colinearity for the skeletal discrepancy length difference (particularly correlation with overjet: 0.40, $P = 0.01$); this variable was excluded from the model. Three other variables (whether teeth showed, overjet, and labiomental angle) remained in the model (Table VI). The model explained 44% of the variation in the attractiveness scores. After we examined residuals, we decided that the models fitted well. The model shows that smaller overjet, no visible teeth, and somewhat smaller labiomental angle were associated with more attractive profiles, whereas larger overjet, visible teeth, and more obtuse labiomental angle resulted in less attractive ratings.

Table III. Reliability of ratings and rater groups assessing 40 profiles

Rater group	*n*	*Intraclass correlation (95% CI)*	*Mean rating (SD)*	P *value (paired t test)*
Pupils	30	0.41 (0.31, 0.54)	3.71 (0.69)	
Adults	24	0.53 (0.43, 0.66)	3.16 (0.81)	<0.001*
All raters	54	0.45 (0.35, 0.57)	3.46 (0.72)	-

*Comparing girls and women.

DISCUSSION

This study suggest that early treatment (phase 1) results in a moderate improvement in the perceived facial profile. Although this difference was statistically significant, the difference in means was not large. This is probably consistent with the findings of other randomized controlled trials of early Class II treatment showing that the effects on skeletal pattern are small; it might not be a surprise that this is reflected in the rating of attractiveness.[9-11]

Our results did not support the findings of the only other randomized controlled trial of the effect of early treatment on facial profile attractiveness.[1] This might be because that study used a small sample size of untreated patients; it could be suggested that this resulted in insufficient statistical power. Furthermore, we measured differences between the 2 groups at the end of treatment, and they evaluated changes in attractiveness. We took this step because it is more relevant to consider the final status of subjects in a randomized controlled trial.

It was interesting and relevant that larger overjets and tooth display (at rest) defined less attractive

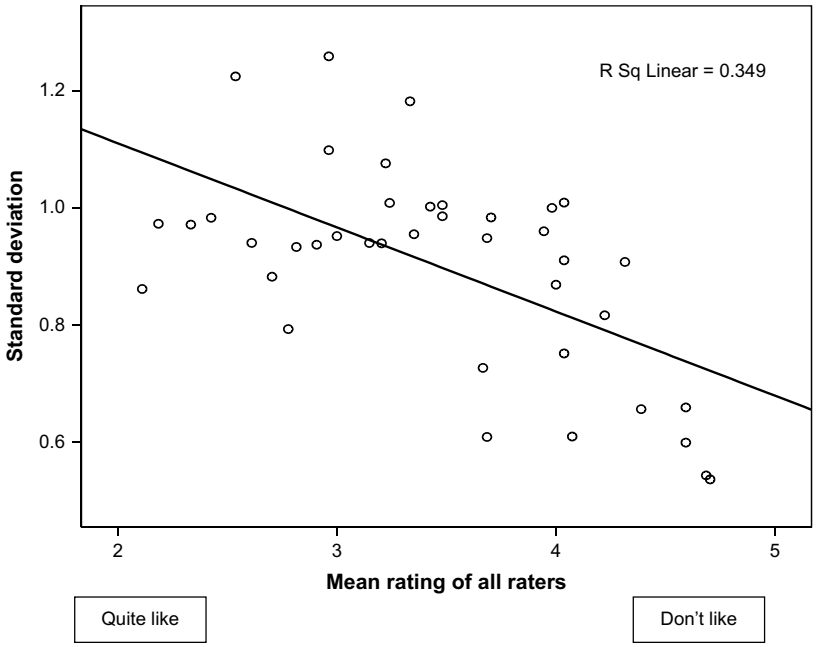

Fig. Standard deviation by mean rating of appeal (54 raters, 40 profiles).

Table IV. Reproducibility (intraclass correlation) by rater group assessing 5 profiles on 2 occasions

Rater group	*n*	*Intraclass correlation mean (SD)*	P *value (t test)*
Pupils	28	0.46 (0.38)	
Adults	24	0.51 (0.31)	0.414
All raters	54	0.48 (0.35)	-

profiles. Importantly, correction of this visible feature of Class II malocclusion is of value and could be considered a benefit of early orthodontic treatment for Class II malocclusion. Paradoxically, it was surprising that we found an association between more obtuse labiomental angles and perceived unattractiveness in the regression model. This might have been because the effect was modified by other variables, and, because there was no effect of the labiomental angle in the univariate analysis, we concluded that this was not relevant.

We did not attempt to evaluate whether the peers' perceptions of facial attractiveness had an influence on the children's self esteem because of sample-size constraints. However, our investigation into the effect of early Twin-block treatment on self-esteem concluded that early treatment resulted in improved self-esteem.[12] Although it is tempting to conclude that this was due to a change in appearance, this must be considered conjecture and certainly requires further investigation.

Table V. Clinical, radiographic, and psychosocial variables and perceived attractiveness

Characteristic	*n*	*Unattractiveness mean (SD) or correlation coefficient*	P *value (t test or correlation*)*
Treatment			
Control	20	3.77 (0.55)	0.006
Twin-block	20	3.16 (0.75)	
Sex			
Male	22	3.30 (0.74)	
Female	18	3.66 (0.66)	0.115
Front teeth showing			
No	23	3.22 (0.67)	0.012
Yes	17	3.79 (0.68)	
Lips together			
No	24	3.55 (0.74)	0.355
Yes	16	3.33 (0.69)	
Overjet	40	−0.492	0.001
Skeletal discrepancy length difference	40	−0.330	0.038
Labiomental angle	40	0.278	0.083
Facial convexity-Proffit's angle	40	−0.209	0.197
Lower facial height (%)	40	0.120	0.459

*To test null hypothesis that correlation coefficient equals zero.

We used profile silhouettes because they had been used previously by O'Neill et al,[1] and we wanted our results to be comparable. It could be suggested that this is a simplistic approach because people rarely see

Table VI. Results for linear regression (dependent variable: mean rating)

	Coefficients				*95% CI for β*	
Variable	*β*	*SE*	*t*	*Significance*	*Lower bound*	*Upper bound*
Constant	.996	.675	1.475	0.149	−.373	2.365
Overjet at DC2	−.075	.021	−3.546	0.001	−.118	−.032
Front teeth showing at DC2	.403	.187	2.156	0.038	.024	.783
Labiomental angle	.014	.005	2.762	0.009	.004	.024

$R^2 = 0.44$

themselves from the side, and most interaction is face to face. However, the perception of attractiveness varies with the angle from which a face is seen,[1,13] but orthodontic treatment can cause changes principally in the anteroposterior and vertical aspects, which are readily assessed in profile.[14] Photographs are a reliable way of assessing appearance,[15] but eyes,[16,17] cheeks, hair, complexion or makeup,[17] coloring,[18] and facial expression[19] can dominate perception. Viewing profiles in silhouette minimizes these confounding effects and, importantly, provides anonymity for subjects.[20]

The moderate R^2 value for the regression (0.44) agrees with recent findings that much is unexplained about perceived attractiveness, which can be influenced by features unlikely to be affected by orthodontic treatment (ie, shape of nose or forehead).[13]

There was a high degree of parity between the treated and untreated samples at DC1 for all data except sex, where there were slight preponderances of boys in the treated group (14:6) and girls in the control group (12:8). This was a chance occurrence; the samples were not stratified for sex, because a randomized controlled trial has shown it to have no influence on change in skeletal pattern, so its influence on perceived attractiveness was expected to be slight.[5] Furthermore, sex was not obvious in these silhouetted profiles of prepubertal subjects. It was impossible to obtain the children's perceptions of the pretreatment views of the patients in the sample because of the need to keep the data-gathering session short, since a child's attention span is limited. Nevertheless, the sample was selected from a large well-run clinical trial, and the study groups did not differ with respect to morphologic variables such as skeletal discrepancy and overjet.

CONCLUSIONS

1. Children with Class II malocclusion, treated with Twin-blocks in the mixed dentition, had profiles that were generally perceived as more attractive than those of an untreated cohort, by both peers and teachers. However, these differences were small.
2. Subjects whose profiles were perceived as more attractive tended to have smaller overjets, no visible teeth, and slightly more acute labiomental angles.

We thank the pupils and the staff of Etchells School, Stockport, Cheshire, United Kingdom.

REFERENCES

1. O'Neill K, Harkness M, Knight R. Ratings of profile attractiveness after functional appliance treatments. Am J Orthod Dentofacial Orthop 2000;118:371-6.
2. Frude N, Kenealy P, Shaw WC. Influence of children's physical attractiveness on teacher expectations. J Soc Psychol 1988;128: 373-83.
3. Kenealy P, Frude N, Shaw WC. Teacher expectations as predictors of academic. J Soc Psychol 1991;131:305-6.
4. Donner A, Eliasziw M. Sample size requirements for reliability studies. Stat Med 1987;6:441-8.
5. O'Brien K, Wright J, Conboy F, Sanjie Y, Mandall N, Chadwick S, et al. Effectiveness of early orthodontic treatment with the Twin-block appliance: a multicenter, randomized, controlled trial. Part 1: dental and skeletal effects. Am J Orthod Dentofacial Orthop 2003;124:234-43.
6. Pancherz H. The mechanism of Class II correction in Herbst appliance treatment. A cephalometric investigation. Am J Orthod Dentofacial Orthop 1982;82:104-13.
7. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420-8.
8. Landis JR, Koch GC. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
9. Tulloch JFC, Phillips C, Proffit WR. Benefits of early Class II treatment: progress report of a two-phased randomized clinical trial. Am J Orthod Dentofacial Orthop 1998;113:62-72.
10. Keeling SD, Wheeler TT, King GJ, Garvan CW, Cohen DA, Cabassa S, et al. Anteroposterior skeletal and dental changes after early Class II treatment with bionators and headgear. Am J Orthod Dentofacial Orthop 1998;113:40-50.
11. Ghafari J, Shofer FS, Jacobsen-Hunt U, Markowitz DL, Laster LL. Headgear versus functional regulator in the early treatment of Class II Division 1 malocclusion. Am J Orthod Dentofacial Orthop 1998;113:51-61.
12. O'Brien K, Wright J, Conboy F, Chadwick S, Connolly I, Cook P, et al. Effectiveness of early orthodontic treatment with the Twin-block appliance: a multicenter, randomized, controlled trial. Part 2: psychosocial effects. Am J Orthod Dentofacial Orthop 2003; 124:488-95.

13. Phillips C, Griffin T, Bennett E. Perception of facial attractiveness by patients, peers and professionals. Int J Adult Orthod Orthognath Surg 1995;10:127-35.
14. Dongieux J, Sassouni V. The contribution of mandibular positioned variation to facial esthetics. Angle Orthod 1980;50: 334-9.
15. Howells DJ, Shaw WC. The validity and reliability of ratings of dental and facial attractiveness for epidemiologic use. Am J Orthod 1985;88:402-8.
16. Kerr WJS, O'Donnell JM. Panel perception of facial attractiveness. Br J Orthod 1990;17:299-304.
17. Foster EJ. Profile preferences among diversified groups. Angle Orthod 1973;43:34-40.
18. Cochrane SM, Cunningham SJ, Hunt NP. A comparison of the perception of facial profile by the general public and 3 groups of clinicians. Int J Adult Orthod Orthognath Surg 1999;14:291-5.
19. Barrer JG, Ghafari J. Silhouette profiles in assessment of facial aesthetics: a comparison of cases treated with various orthodontic appliances. Am J Orthod 1985;87:385-91.
20. Johnston DJ, Hunt O, Johnston CD, Burden DJ, Stevenson M, Hepper P. The influence of lower facial vertical proportion in facial attractiveness. Eur J Orthod 2005;25:349-54.