

Likert scale versus the visual analogue scale in evaluating dentofacial aesthetics: a systematic review

Hanis Nabilah Marzuki,* Izzah Zahirah,* May Nak Lau,† Elavarasi Kuppusamy,* Nik Mukhriz Nik Mustapha‡ and Asma Ashari*

Department of Family Oral Health, Faculty of Dentistry, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia*

Department of Paediatric Dentistry and Orthodontics, University of Malaya (UM), Kuala Lumpur, Malaysia†

Centre for Paediatric Dentistry and Orthodontic Studies, Universiti Teknologi MARA (UiTM), 47000 Selangor, Malaysia‡

Background: Numerous studies rating dentofacial aesthetics have been conducted by using the Likert Scale or the Visual Analogue Scale (VAS). The Likert Scale is an assessment tool consisting of items that require respondents to rank a query using graded declarative statements. Alternatively, the VAS is a tool that requires the respondents to place a mark along a horizontal line to rate a particular item.

Objective: This systematic review aimed to identify the validity and reliability of the Likert scale and VAS in evaluating dentofacial aesthetics. A second objective was to identify which scale is preferable for use by clinicians and patients in determining dental aesthetics.

Search methods: With no language limitations, a comprehensive electronic database search was conducted in the PubMed, Cochrane Central Register of Controlled Trials, Scopus, Web of Science, Ovid, and OpenGrey databases using keywords and Mesh terms combinations: ("Likert scale" OR "Likert") AND ("visual analogue scale" OR "VAS") AND ("aesthetic*" OR "facial*" OR "dental*" OR "dentist*"). The selection criteria were set based on the PICO format. Population (P): Laypeople and/or dental clinicians; Intervention (I) and Control (C): Likert Scale and Visual Analog Scale for measuring dentofacial aesthetics; Outcome (O): Validity, reliability, and the patient's or clinician's preference for using the Likert Scale or the Visual Analog Scale for measuring dentofacial aesthetics. The risk of bias was assessed using the STROBE checklist for observational studies and Version 2 of the Cochrane Risk of Bias Tool for Randomized Trials (RoB 2) for randomized controlled trials. The results were summarised qualitatively; no meta-analysis was conducted due to the high level of heterogeneity of the included studies.

Results of the synthesis: Both the Likert Scale and VAS are valid and reliable for scoring dentofacial aesthetics but each have their own advantages in aesthetic evaluation. There were different preferences for the two scoring methods.

Conclusions: Overall, there is insufficient data to draw a conclusion that the VAS or the Likert scale is superior. Either of these scales may be used to rate dentofacial aesthetics.

[Aust Orthod J 2024; 40: 158 - 168. DOI: 10.2478/aoj-2024-0010]

Received for publication: September, 2023

accepted: April, 2024.

Hanis Nabilah Marzuki: hanismarzuki99@gmail.com; Izzah Zahirah: a167920@siswa.ukm.edu.my; May Nak Lau: minalau@um.edu.my; Elavarasi Kuppusamy: elavarasi.k@ukm.edu.my; Nik Mukhriz Nik Mustapha: nikmukhriz@uitm.edu.my; Asma Ashari: asmaashari@ukm.edu.my

Background

The Likert scale is a rating system used in questionnaires and is designed to measure a respondent's attitudes, opinions or perceptions.¹ In

health research studies, the Likert scale is a widely utilised scaling technique whose response could range on a scale point of 5 to 7. Possible ranges of responses for a 5-point Likert scale are rated from "strongly

agree”, “agree”, “neutral”, “disagree” and “strongly disagree”.¹ There are varying viewpoints on the ideal number of response options and the wording of the possible answers has an impact on the responses.² Likert scales are an ordinal-level measure rather than an interval-level measure, in that the response categories have a rank order but the intervals between values cannot be assumed to be equal.³

In 1921, the Visual Analogue Scale (VAS) was first described and referred to as a “graphic rating method”. Respondents indicated their perceived status by placing a mark along a horizontal line at the most relevant position. A VAS is commonly displayed as a 10 cm long horizontal line, anchored by two verbal descriptors at the extremes.⁴ The orientation of a VAS may be either horizontal or vertical, and the choice between the two styles should not impact the survey outcome. However, the orientation of the scale may deviate from the respondent’s likely point of view as a horizontal VAS and a vertical 10-cm VAS have shown different reproducibility results.³

Both Likert scales and VAS have their individual advantages and disadvantages that have been previously described.^{5–10} VAS has been claimed to have better responsiveness than surveys which have used a Likert scale.¹¹ Researchers claim that a Likert scale and a VAS yield a similar result despite being different but comparable in reliability and validity. The disadvantage of a VAS is that it requires more dedication, time and effort to educate respondents who have difficulty in understanding the instructions. The markings of a VAS may also be non-specific and uninterpretable which may be less accurate than using a scale method.

The Likert scale has been noted for its simplicity of use and ease of comprehension as it uses numerical coding for interpretation, which is preferred compared to the VAS.¹² Children appear to understand the Likert scale better as the system can facilitate communication more easily than the VAS and is able to discriminate between the maximal intensities.² A Likert bias can be reduced by increasing the number of response categories.

Few criteria have been used to compare the different types of VAS and Likert scales, but include the consistency of adjectives and a scale ratings relationship, the magnitude of between-subject variability, test-retest reliability, correlation strength with other measurements, factor analysis

performance, and achieving a uniform response distribution.² The consequences of differences in these criteria have caused variations in the conclusion of the method’s relative value. Grant et al. (1999) ascertained that the VAS had greater sensitivity than the Likert scale but there were no significant differences between the two.¹³

In the dental setting, the Likert scale and VAS have been widely used to measure patient satisfaction regarding treatment, pain, anxiety and more. During orthodontic treatment specifically, patients may evaluate the satisfaction of their changing dentofacial appearance, as in contemporary times, dentofacial aesthetics has become a major concern.¹⁴ Clinically, the commonly used Likert scale and VAS rating scale are important and are applied to determine the patient’s perspective and expectation towards treatment. Many studies have been conducted to evaluate the Likert scale and VAS but the results remain inconclusive. However, a study by Dourado et al. (2021) showed that respondents preferred to use the Likert scale compared to the VAS in rating facial attractiveness.¹⁵

It is apparent that research comparing the benefits and disadvantages of using the VAS and Likert scales have frequently revealed inconsistent results.¹⁶ It has been indicated that the Likert scale is easier to use compared to the VAS while others report the opposite.¹⁶ Therefore, the present systematic review aimed to:

1. identify the validity and reliability of both the Visual Analogue Scale (VAS) and Likert scale in evaluating responses from respondents regarding their perspective of dentofacial aesthetics.
2. identify which scale is preferable for use by clinicians and patients in assessing dentofacial aesthetics.

By comparing the VAS and Likert scales for their advantages and disadvantages in obtaining accurate responses, clinicians can discern the most appropriate scale for incorporation into research questionnaires.

Methods

The present systematic review was conducted in accordance with the Cochrane Handbook for Systematic Reviews of Interventions and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). The review was registered in Prospero with the ID: CRD42023479854.

Table I. Eligibility criteria

Criteria	Inclusion criteria	Exclusion criteria
Population	laypeople and/or dental clinicians	
Intervention	Likert Scale and VAS used in assessing facial aesthetic	
Outcome measures	1.Validity and/or reliability of both scale 2.Patient and/or clinician's preference between Likert Scale and VAS	
Study design	Randomized controlled trial, cross sectional study	Reviews, case reports, letters to editor, and all other non-peer- reviewed articles
Others	All language article	

Criteria for considering studies for this review

The eligibility criteria were based on the PICO format summarised in Table I. The present review included studies that recruited lay people and/or dental clinicians as the study population. The type of intervention was the use of the Likert Scale or VAS for comparison in measuring facial aesthetics. In addition, the present systematic review sought studies that reported outcomes related to the validity, reliability, and patient or clinician's preference between the Likert scale and VAS. The included studies were limited to randomised controlled trials and cross-sectional studies, with no time or language limit. Reviews, case reports, letters to the editor, conference abstracts, meeting abstracts, oral and poster presentations, and all other non-peer-reviewed articles were excluded.

Search methods for identification of studies

The database search and identification of studies were conducted independently by two authors (M.H.N. and A.S.I.Z.) in duplicates, up to the 30 November

2023. Any disagreement between the two authors was resolved by discussion with the third author (A.A.). Relevant studies were screened by reviewing the titles and abstracts of all studies identified from six electronic databases: PubMed, Cochrane Central Register of Controlled Trials, Scopus, Web of Science, Ovid, and OpenGrey. A combination of Keywords and MESH terms: ("Likert scale" OR "Likert") AND ("visual analogue scale" OR "VAS") AND ("aesthetic*" OR "facial*" OR "dental*" OR "dentist*") were applied in the electronic database search, with no time or language limit. Furthermore, the reference lists of all the included studies were screened for suitable additional studies to be included in this review. Corresponding authors were contacted to obtain clarifications when necessary.

Data collection and analysis

Selection of studies

After database searches and the identification of potential studies, the full texts of identified studies were obtained. The selection of studies was conducted independently by two authors (M.H.N. and A.S.I.Z.) in duplicates, and any disagreement was resolved by discussion with the third author (A.A.) by reviewing the full texts.

Data extraction and management

Data were extracted from the full text of the selected studies using a standardised data extraction sheet that was designed specifically for the present systematic review. The findings related to study methodology, study population, intervention characteristics, and outcome measure(s) were extracted and are summarised in Table III.

Risk of bias assessment in the included studies

All risk of bias assessments were conducted by two authors (M.H.N. and A.S.I.Z.) independently, in duplicates. Any disagreement between these two authors was resolved by discussion with the third author (A.A.).

The risk of bias assessment of the cross-sectional study was conducted using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist. STROBE is a guideline for

the report of observational research such as cohort, case-control, and cross-sectional studies. The checklist consists of 22 items that are divided into six categories: (1) the article's title and abstract, (2) introduction, (3) methods, (4) results, (5), discussion, (6) other information. The guideline helps authors to ensure that selected research papers were of high quality.¹⁷ The risk of bias assessment of a randomised controlled trial was assessed using Version 2 of the Cochrane Risk of Bias Tool for Randomised Trials (RoB 2) as shown in Table II. RoB 2 is structured into a series of fixed bias regions that focus on different aspects of the study design, conduct, and reporting. Within each domain, a series of questions ("signalling questions") aimed to gather information about study characteristics relevant to the risk of bias. Based on the answers to the signalling questions, the algorithm generates suggested judgments regarding the risk of bias attributable to each domain. Verdicts may be at 'low' or 'high' risk of bias, or express 'some degree of concern'.¹⁸

Result/review findings

Study selection

Figure 1 shows the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram of this systematic review. The six electronic database searches identified 517 potentially relevant articles. After removing 30 duplicates, 487 papers were screened. A total of 449 articles were excluded as they did not meet the inclusion and exclusion criteria. The full texts of the remaining 38 articles were successfully retrieved and assessed for eligibility. A total of 33 papers were subsequently excluded as they did not match the eligibility criteria. Finally, a total of five research papers were included in the systematic review (one randomised controlled trial and four cross-sectional studies).

Risk of bias

Table II and Figure 2 show the risk of bias results of the articles included in the present systematic review.

Study characteristics

Table III shows the study methodology, population, intervention characteristics, and outcome measure(s) of the included articles.

Research outcomes

All of the studies included in the present systematic review presented high heterogeneity which enabled a qualitative review of the use of the Likert scale and VAS in evaluating dentofacial aesthetics. The research outcomes from the papers are summarised in Table IV.

Synthesis of result

Reliability and Validity between the Likert Scale and Visual Analogue Scale

Based on Hatch et al. (2017), the test-retest reliability of the Likert scale and VAS was 0.69 and 0.75, respectively, for overall facial attractiveness ratings which indicated overall acceptable reliability.¹⁹ Aşık and Kök (2021) found that, for dental midline deviation and smile attractiveness, the scores of both the Likert scale and VAS were consistent.²⁰ The intraclass correlation coefficient values for the VAS and Likert scales were 0.805 and 0.76, respectively.²⁰

Reliability can also be measured by using an intra-examiner error analysis. In a study evaluating facial pleasantness, the weighted kappa test statistic was applied to the Likert Scale and intraclass correlation coefficient for the VAS to determine intra-examiner reliability. The result showed moderate intra-examiner agreement for the VAS (0.69) and Likert scale (0.70).¹⁵ An additional study conducted using intra-examiner reliability testing on a VAS showed acceptable results of an intraclass correlation coefficient which was over 0.90.²¹

A study conducted by Fudalej et al. (2017) evaluating the nasolabial aesthetics of patients presenting with a unilateral cleft lip and palate (CLP) showed that the VAS was more consistent than the aesthetic index using a Likert scale in rating nasal shape ($P = 0.001$).²² Fudalej et al. (2017) further compared three methods of rating nasolabial appearance in CLP patients and showed a 5-point aesthetic index was unsatisfactory for the aesthetic rating of the nasolabial region. The results from a 5-point aesthetic index, which was based on a Likert scale, was considered to be unreproducible and generated variable results.²²

When relating gender to validity and reliability scoring, Aşık and Kök (2021) reported that there were no statistically significant differences between genders and ages, observed in both the VAS and

Table II. Risk of bias (RoB) with STROBE checklist

Paper	Study	STROBE	
		Score	Paper
Perception of dental midline deviation and smile attractiveness by eye-tracking and aesthetic ratings	Aşik & Kök (2021) ¹³	78.5%	Good
Effects of objective 3D measures of facial shape and symmetry on perceptions of facial attractiveness	Hatch et al. (2017) ¹²	75.3%	Good
Likert scale vs visual analog scale for assessing facial pleasantness	Dourado et al. (2021) ⁸	85.18%	Excellent
Correlation coefficients of three self-perceived orthodontic treatment need indices	Eslamipour et al. (2017) ¹⁴	85.18%	Excellent

the Likert scale evaluations of the effects of midline deviation on smile aesthetic perception.²⁰ Daurado et al. (2021) and Eslamipour et al. (2017) also reported that, regardless of gender and age, the evaluators were able to similarly assess using either method.^{15,21}

Preferred scale among laypeople and dental practitioners

In a study conducted to assess facial pleasantness, Dourado et al. (2021) showed that 75% of the examiners, consisting of laypeople, senior orthodontists, and senior oral and maxillofacial surgeons, preferred the Likert scale over the VAS.

The claims made by the respective examiners were that the Likert scale had a better representation of their opinions and was easier and simpler to use.¹⁵

However, a further study reported that the VAS had a higher sensitivity (44.6%) compared to the Likert Scale (22.3%), namely the 'Aesthetic Components of The Index Orthodontic Treatment Need' and 'Oral Aesthetic Subjective Index Scale'.

Discussion

The VAS and Likert Scales have been widely used in evaluating dentofacial aesthetics and much research has been conducted on comparing the Likert Scale and VAS as scoring systems.^{10,12,23–26} Considerations in choosing between these scales include the subjective nature inherent in both, potential variations in individual interpretations, and respondent comfort with either format. The choice may also hinge on the study objectives, with the VAS being preferred for a more detailed and continuous measurement. Cultural factors, individual preferences, and the research tradition within the field should also influence this decision. Therefore, the present systematic review was conducted to identify the reliability, validity, and preference of clinicians and patients for both the Likert scale and VAS in the evaluation of dentofacial aesthetics. However, due to the heterogeneity of the included studies, a meta-analysis could not be conducted.

Previous research studies have indicated diverse researcher preferences regarding the suitability of the Likert scale or VAS for research evaluation. The preference between Likert scales and VAS among clinicians and laypeople is shaped by distinct factors tied to their professional backgrounds and familiarity with different rating systems. Clinicians, accustomed to precise measurements in their professional training, may gravitate toward the VAS due to its continuous nature, which enables finer gradations. Conversely, laypeople, who often encounter Likert scales in non-clinical assessments, may find them more accessible and user-friendly due to their simplicity and ease of comprehension. However, the present systematic review found that both clinicians and laypersons preferred the Likert scale over the VAS, as it was perceived that the Likert scale had a better representation of their opinions, and was easier and simpler to use to assess facial pleasantness.¹⁵

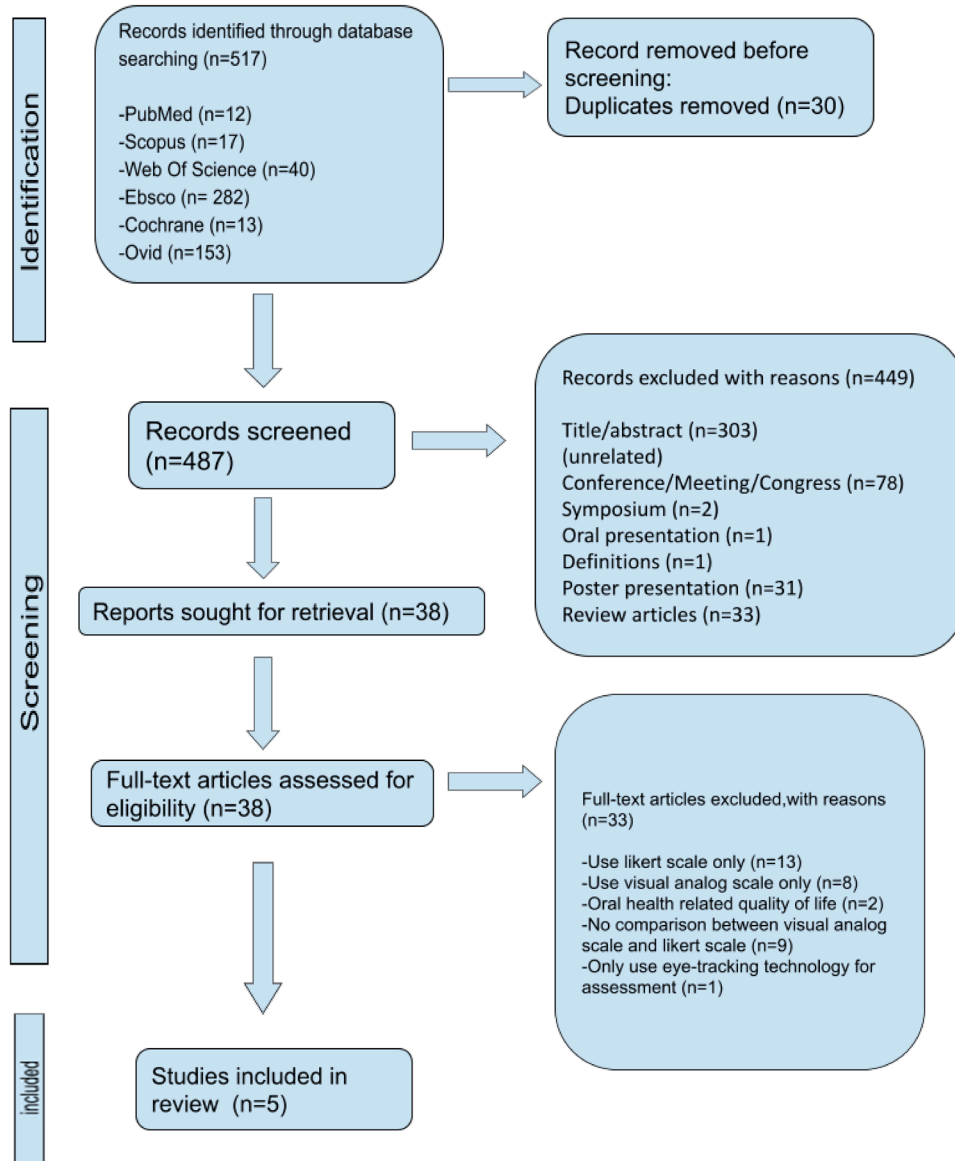


Figure 1. Study selection according to the Preferred reporting items for systematic reviews and meta-analysis (PRISMA) flowchart.

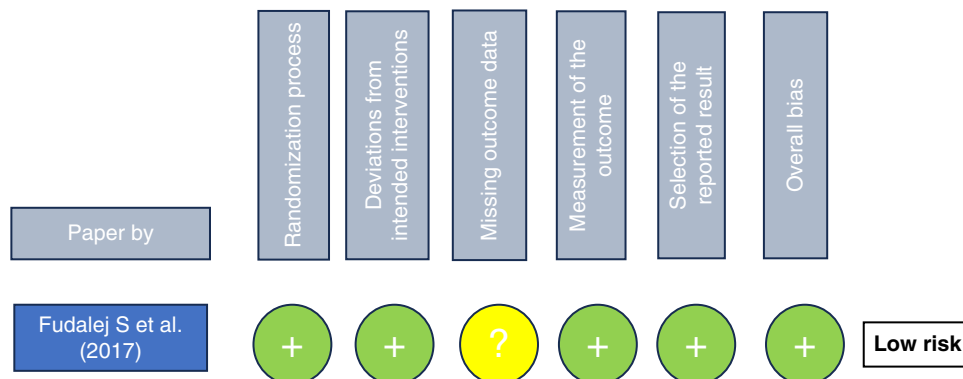


Figure 2. Risk of bias according to the Cochrane Risk of Bias Tool for Randomised Trials (RoB 2).

Table III. Study methodology, population, intervention characteristics, and outcome measure(s)

Study	Study Methodology	Study Population	Intervention Characteristics	Outcome Measure(s)
Fudalej et al. (2017) ¹⁵	Randomized controlled trial	Eight junior orthodontic residents without any experience in treatment of CLP. Age: 25 to 31 years old (four male and four female).	Extra-oral images of 60 non-syndromic complete unilateral cleft lip and palate (CLP) patients were taken from the frontal and profile views and were cropped. Photographs (N = 60) were divided into three groups (A, B and C) equally. Afterwards, three subgroups were formed within each group (A1, A2, A3; B1, B2, B3; and C1, C2, C3). The order of images in subgroups were arranged using random numbers generated online. Images from subgroups designated 1 were rated with VAS, images from subgroups designated 2 were rated with esthetic index (5-point Likert scale), and images from subgroups designated 3 were rated with reference score which used numerical scale from 0 to 200 and a reference photograph	Reliability was measured by using intraclass correlation coefficient (ICC) in assessing nasolabial appearance.
Aşık & Kök (2021) ¹³	Cross-sectional	195 participants (52 laypersons, 50 patient's relatives, 51 dentists, and 42 orthodontists) from the Necmettin Erbakan University Faculty of Dentistry, Department of Orthodontics. Gender: 102 females, 93 males. Age groups: 20-30, 30-40, 40-50 and over 50 years.	Photographs of a female frontal posed smile were cropped to include only mouth and teeth produced a base image. Dental midline was adjusted 1, 2, 3, and 4mm to the left and right sides using Adobe Photoshop CC 2014 software resulting in 8 modified smile images. 11 images with eight modified smile photographs, one base photograph and 2 repeating photographs were randomly listed. Eye tracking device was used to detect participants' eyes on the screen. Volunteers evaluated each image from an aesthetic point of view using survey forms consisting of VAS and Likert Scales	Effect of dental midline deviation on the perception of smile aesthetics by orthodontist, dentists, patient relatives, and layperson. Reliability of eye tracking data, VAS and Likert scale was measured using intraclass correlation coefficients (ICC). Age and gender effect towards VAS and Likert scale scores in assessing dental midline.

Table III. Continued

Study	Study Methodology	Study Population	Intervention Characteristics	Outcome Measure(s)
Hatch et al. (2017) ¹²	Cross-sectional	Ten university students and staff (5 females and 5 males) identified from convenience sampling approach	Frontal and lateral views of images of 313 adults in Iowa were presented to the raters. The rates evaluated the overall facial attractiveness on a 5-point Likert scale and a VAS	Reliability of both scales by using test-retest reliability in measuring overall facial attractiveness. Correlation between Likert Scale and VAS.
Dourado et al. (2021) ⁸	Cross-sectional	90 evaluators divided into 3 groups equally: Lay people (20-67 years old) Orthodontist (27-45 years old) Oral and maxillofacial surgeon (26-51 years old).	Evaluation on facial pleasantness was done on 10 adult patients' photographs that underwent orthodontic treatment for different facial discrepancies compiled in an album. Each evaluator will assess the album at 2 different times 15-30 days apart.	Reliability of Likert scale and VAS was measured through intra examiner error analysis. Correlation between both scales was measured using the Spearman correlation coefficient. Compare preference of both scales from the perspective of orthodontist, oral maxillofacial surgeons and laypeople.
Eslamipour et al. (2017) ¹⁴	Cross-sectional	993 freshman students from Isfahan University were randomly selected and examined to assess the dental health component (DHC) index of the Index of orthodontic treatment need (IOTN) Examination was done by two examiners.	Questionnaire divide into three sections: Demographic, index of oral aesthetic subjective index scale (OASIS) using 7-point Likert scale, VAS and IOTN.	Determination of most reliable self-perceived indices by using intraclass coefficients. To assess sensitivity and specificity of self-perceived indices (Aesthetic component, OASIS and VAS) in comparison to DHC as normative index in young adults.

Table IV. Summary of research outcomes

Study	Title	Outcomes
Aşık & Kök (2021) ¹³	Perception of dental midline deviation and smile attractiveness by eye-tracking and aesthetic ratings	The ICC was 0.805 for the VAS, and 0.760 for the Likert Scale No statistically significant difference between the genders VAS and Likert scores were consistent with the eye-tracking data. No statistically significant difference between the age groups.
Hatch et al. (2017) ¹²	Effects of objective 3D measures of facial shape and symmetry on perceptions of facial attractiveness	The test-retest reliability for overall facial attractiveness ratings was 0.69 on the Likert-scale and 0.75 on the VAS.
Dourado et al. (2021) ⁸	Likert scale vs visual analog scale for assessing facial pleasantness	Intra examiner errors were 0.70 and 0.69 for the Likert scale and VAS, respectively. Good correlation was observed between both the scales. Laypeople presented the lowest correlation (moderate) among the 3 groups. 75% of the examiners preferred the Likert scale. VAS requires some time to measure. Regardless of sex, the evaluators were able to assess similarly using the 2 methods.
Eslamipour et al. (2017) ¹⁴	Correlation coefficients of three self- perceived orthodontic treatment need indices	The intraclass coefficient for each examiner was over 0.9. VAS had the highest sensitivity (44.6%). VAS is an easy and understandable index for all age groups and social levels. Age and gender of the patient, and also the economic status of the family had no significant correlation with these indices.
Fudalej et al. (2017) ¹⁵	Comparison of three methods of rating nasolabial appearance in cleft lip and palate	The VAS method was in turn more reproducible than Likert scale. Results imply that a 5-point esthetic index is not optimal for esthetic rating of the nasolabial region in CLP. Recommend the use of reference. photographs along with the VAS. Likert-type scales seem to produce the most variable results

Dourado et al. reported that some participants required 5 to 10 minutes or longer to comprehend the VAS.¹⁵ Nevertheless, the VAS was considered to have greater sensitivity compared to the Likert Scale. This finding is in agreement with Eslamipour et al. (2017) who asserted that the VAS has higher sensitivity and serves as an accessible and comprehensible index across diverse social strata.²¹ Furthermore, Fudalej et al. (2017) chose the VAS over the Likert scale in an assessment of the nasolabial region ratings in cleft lip and palate (CLP) research, as it was argued that the Likert scale lacked reproducibility and yielded variable results within the scoring system. Conversely, the VAS method was deemed more reproducible and

preferable by the researchers. However, this claim is not substantiated nor supported by the present systematic review.²² From the perspectives of data collection and analysis, the Likert scale is more straightforward than the VAS because the latter requires additional time to measure the extent of the generated line.

Aesthetics is commonly regarded as both subjective and universal, often assuming that experts comprehend it better than individuals. However, interventions and treatments might be validated based on individual preferences. When evaluating their own aesthetic preferences and necessities, patients should be regarded as experts. The assessment of any deformity by a

professional provides another rationale for treatment, although it may not always be essential if the individual does not perceive the deformity as affecting their quality of life. What appears as a discrepancy to dental professionals might be perceived as normal by the layperson. It is crucial to exercise caution against medicalising the spectrum of normalcy, by advocating treatment in the absence of pathology and without the patient's self-perceived need.

The included studies in this systematic review recruited either university students, young adults, and/or the elderly as their study subjects. The current systematic review found that age groups did not affect the results of the dentofacial aesthetics evaluation on either the Likert scale or the VAS. However, none of the included studies recruited children in the evaluation of dentofacial aesthetics. Children have been reported to prefer the Likert scale and find it easier to complete compared to the VAS.²⁷ The relation between age and preference for VAS and Likert scales is multifaceted, involving cognitive, linguistic, and experiential factors. Researchers and educators should consider these factors when selecting assessment tools for different age groups, and aim for formats that align with the developmental stage and preferences of the target population. Further studies on the preference of individuals from different age groups on the tool used to evaluate dentofacial aesthetics are warranted.

Limitations

A limitation of this review was that the reviewed papers were taken from many different areas and parts of the world. Therefore, the epidemiology of the conducted research might affect the research results. The level of knowledge and the examiners involved in all the research papers were varied. Preferred scoring systems cannot be concluded as the results from each of the papers were varied.

The papers that have been reviewed were also limited. There were a lack of resources comparing the Likert Scale and VAS from a dentofacial aesthetics perspective. Only five papers were identified for the comparative use of the assessment scales. More online databases could have been explored to find additional related papers. The difficulty in finding the appropriate MESH words is a possible reason for the inability to obtain optimal search papers.

Furthermore, some papers had no direct comparison between the Likert scale and VAS which made the interpretation of these papers difficult.

Aspects of aesthetic evaluation were performed differently between the included research papers which generated inconsistent and variable results. From the present review, it was concluded that the Likert Scale and VAS could be used interchangeably based on the suitability of the scoring system for reliable results of the aesthetics evaluation.

Lastly, the heterogeneity of the included studies precluded the possibility of conducting a meta-analysis.

Conclusion

Both the Likert Scale and VAS are valuable tools for the evaluation of dentofacial aesthetics. While both systems have consistently yielded favourable results in objective evaluations of dentofacial aesthetics across the included studies, reaching a consensus on the superiority of one method over the other proved challenging due to heterogeneity which precluded a quantitative analysis and meta-analysis. Therefore, further research is indicated to decisively identify the strengths and weaknesses of these two methods in various applications and considerations within dentofacial aesthetics. Additionally, there is an urgent need for the ongoing development and evaluation of dentofacial aesthetic questionnaires, given the relatively limited current research and literature in this domain.

Conflict of interest

The authors declare that there is no conflict of interest.

Ethical clearance

This study was approved by the UKM Research Ethics Committee [JEP-2022-311].

Corresponding author

Asma Ashari, Department of Family Oral Health, Faculty of Dentistry, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia Email: asmaashari@ukm.edu.my

May Nak Lau, Department of Paediatric Dentistry and Orthodontics University of Malaya (UM),

50603 Kuala Lumpur, Malaysia Email: minalau@um.edu.my

Nik Mukhriz Nik Mustapha, Centre for Paediatric Dentistry and Orthodontic Studies, Universiti Teknologi MARA (UiTM), 47000 Selangor, Malaysia Email: nikmukhriz@uitm.edu.my

Acknowledgement

The authors would like to thank all of the people who were involved in this study.

References

- Jamieson S. Likert scale [Internet]. 2023 [cited 2024 Jan 4]. Encyclopedia Britannica. Accessed October 31, 2023. <https://www.britannica.com/topic/Likert-Scale>.
- Mohammad Yusof N, Karupiah K, Md Tamrin S, Rasdi I, Mohamad Jamil PAS. Likert Scale vs. Visual Analogue Scale on vehicle seat discomfort questionnaire: a review. *Malaysian J Med Health Sci*. 2019;15:159–65.
- Jamieson S. Likert scales: how to (ab)use them. *Med Educ*. 2004;38:1217–8.
- Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health*. 1990;13:227–36.
- McLeod A, Pippin S, Wong JA. Revisiting the Likert scale: can the fast form approach improve survey research. *Int J Behav Accounting Finance*. 2011;2:310.
- Maurer TJ, Pierce HR. A comparison of likert scale and traditional measures of self-efficacy. *J Appl Psychol*. 1998;83:324–9.
- Taherdoost H. Designing a questionnaire for a research paper: a comprehensive guide to design and develop an effective questionnaire. *Asian J Managerial Sci*. 2022;11:8–16.
- Heller GZ, Manuguerra M, Chow R. How to analyze the Visual Analogue Scale: Myths, truths and clinical relevance. *Scand J Pain*. 2016;13:67–75.
- Rosas S, Paço M, Lemos C, Pinho T. Comparison between the Visual Analog Scale and the Numerical Rating Scale in the perception of esthetics and pain. *Int Orthod*. 2017;15:543–60.
- Alkadi L, Masuadi E, Mohamed TA, Mohamud M, Farook F. A Likert Scale Versus a Visual Analogue Scale and the participant response: a cross sectional study. *J Int Dent Med Res*. 2022;15:255–62.
- Vickers AJ. Comparison of an ordinal and a continuous outcome measure of muscle soreness. *Int J Technol Assess Health Care*. 1999;15:709–16.
- Akad K, Solmaz D, Sari I, Onen F, Akkoc N, Akar S. Performance of response scales of activity and functional measures of ankylosing spondylitis: numerical rating scale versus visual analog scale. *Rheumatol Int*. 2013;33:2617–23.
- Grant S, Aitchison T, Henderson E, Christie J, Zare S, McMurray J, et al. A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest*. 1999;116:1208–17.
- Ghasemi M, Sharifi R, Tahani B. Satisfaction with dental appearance and willingness to improve dental aesthetic among patients attending the dental school of Isfahan university of medical sciences. *jdm*. 2021;34:102–13.
- Dourado GB, Volpato GH, de Almeida-Pedrin RR, Pedron Oltramari PV, Freire Fernandes TM, de Castro Ferreira Conti AC. Likert scale vs visual analog scale for assessing facial pleasantness. *Am J Orthod Dentofacial Orthop*. 2021;160:844–52.
- Hasson D, Arnetz BB. Validation and Findings Comparing VAS vs. Likert Scales for psychosocial measurements. *Int Electron J Health Educ*. 2005;8:178–92.
- Cuschieri S. The STROBE guidelines. *Saudi J Anaesth*. 2019;13: S31–4.
- Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
- Hatch CD, Wehby GL, Nidey NL, Moreno Uribe LM. The effects of objective 3D measures of facial shape and symmetry on perceptions of facial attractiveness. *J Oral Maxillofac Surg*. 2017;75:1958.
- Aşık S, Kök H. Perception of dental midline deviation and smile attractiveness by eye-tracking and aesthetic ratings. *Australasian Orthod J*. 2021;37:187–96.
- Eslamipour F, Riahi FT, Etemadi M, Riahi A. Correlation coefficients of three self-perceived orthodontic treatment need indices. *Dent Res J (Isfahan)*. 2017;14:37.
- Fudalej SA, Desmedt D, Bronkhorst E, Fudalej PS. Comparison of three methods of rating nasolabial appearance in cleft lip and palate. *Cleft Palate Craniofac J*. 2017;54:400–7.
- Jaeschke R, Singer J, Guyatt GH. A comparison of seven-point and visual analogue scales. Data from a randomized trial. *Control Clin Trials*. 1990;11:43–51.
- Kremer E, Hampton Atkinson J, Ignelzi RJ. Measurement of pain: patient preference does not confound pain measurement. *Pain*. 1981;10:241–8.
- Phan NQ, Blome C, Fritz F, Gerst J, Reich A, Ebata T, et al. Assessment of pruritus intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronic pruritus. *Acta Derm Venereol*. 2012;92:502–7.
- Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs*. 2005;14:798–804.
- Van Laerhoven H, Van Der Zaag-Loonen HJ, Derkx BHF. A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatr*. 2004;93:830–5.