

# Development and Validation of the Eyelash Satisfaction Questionnaire

Jeff Dang, PhD, MPH; Jason C. Cole, PhD; Somali M. Burgess, PhD;  
Min Yang, MD, PhD; Selena R. Daniels, PharmD, MS;  
and John G. Walt, MBA

Aesthetic Surgery Journal  
2016, Vol 36(2) 221–228  
© 2015 The American Society for  
Aesthetic Plastic Surgery, Inc.  
This is an Open Access article  
distributed under the terms of the  
Creative Commons Attribution Non-  
Commercial License ([http://  
creativecommons.org/licenses/by-nc/  
4.0/](http://creativecommons.org/licenses/by-nc/4.0/)), which permits non-commercial  
re-use, distribution, and reproduction in  
any medium, provided the original  
work is properly cited. For commercial  
re-use, please contact journals.  
permissions@oup.com.  
DOI: 10.1093/asj/sjv154  
[www.aestheticsurgeryjournal.com](http://www.aestheticsurgeryjournal.com)

**OXFORD**  
UNIVERSITY PRESS

## Abstract

**Background:** Patient-reported outcome (PRO) measures have been used to assess treatment benefit in a variety of therapeutic areas and are now becoming increasingly important in aesthetic research.

**Objectives:** The objective of the current study was to develop and validate a new PRO measure (Eyelash Satisfaction Questionnaire [ESQ]) to assess satisfaction with eyelash prominence.

**Methods:** The content of the questionnaire (including conceptual framework and questionnaire items) was generated by review of literature, participant interviews, and expert opinion. Cognitive interviews were conducted to pilot test the questionnaire. Psychometric properties of the questionnaire were examined in a combined sample of participants ( $n = 970$ ) completing Internet- ( $n = 909$ ) and paper-based ( $n = 61$ ) versions. Item- and domain-level properties were examined using modern and classical psychometrics.

**Results:** Content-based analysis of qualitative data demonstrated the presence of 3 distinct domains (Length, Fullness, Overall Satisfaction; Confidence, Attractiveness, and Professionalism; and Daily Routine). Initial confirmatory factor analysis (CFA) results of 23 items revealed insufficient model-data fit (comparative fit index [CFI] of 0.86 and a non-normed fit index [NNFI] of 0.82). A revised model using 9 items (3 per domain) achieved appropriate fit (CFI of 0.99 and NNFI of 0.97). Analyses revealed measurement equivalence across the Internet- and paper-based versions. The 3 ESQ domains had strong internal consistency reliability (Cronbach's  $\alpha$  [range] = 0.919-0.976) and adequate convergent and discriminant validity.

**Conclusions:** The ESQ was found to be a reliable and valid PRO measure for assessing satisfaction with eyelash prominence.

## Level of Evidence: 3

Accepted for publication July 10, 2015.



The use of eyelash enhancement products for aesthetic purposes is widespread. Mascara and other eyelash enhancement products represent a fundamental and sizeable area of aesthetic research.<sup>1</sup> Bimatoprost, originally approved for the treatment of glaucoma, has received US Food and Drug Administration (FDA) approval to be marketed as the first treatment for hypotrichosis of the eyelashes (Latisse; Allergan, Inc., Irvine, CA). Bimatoprost has been shown to stimulate eyelash growth, increasing length, thickness, and darkness.<sup>2</sup>

Alongside traditional measures of efficacy and safety, clinical trials in various therapeutic areas increasingly use patient-reported outcomes (PROs) to assess ease of use,

At the time of this research, Dr Dang was President and Senior Research Scientist at Consulting Measurement Group, Torrance, California. Dr Cole was Director and Senior Scientist and Dr Yang was Senior Scientist at QualityMetric, Inc., Lincoln, Rhode Island. Dr Burgess was Director of Global Health Outcomes Strategy and Research, Dr Daniels was Manager of Global Health Outcomes Strategy and Research, and Mr Walt was Director of Global Strategic Marketing at Allergan, Inc., Irvine, California.

### Corresponding Author:

Dr Jeff Dang, 15207 Magnolia Blvd, Unit #129, Sherman Oaks, CA 91403, USA.  
E-mail: [jeff.huydinhdang@gmail.com](mailto:jeff.huydinhdang@gmail.com)

tolerability, and overall treatment benefit.<sup>3,4</sup> For aesthetic products, the participant's perspective of treatment benefit has particular importance given the nonreimbursable cost of treatment. To date, few studies have specifically assessed important concepts to participants regarding eyelash appearance, and none have assessed the impact of treatment on perception of self and daily routine.<sup>5</sup>

The lack of PROs that assess treatment from the participant's perspective highlights the need for a brief and simple measure to assess the effect of a treatment designed to enhance or correct hypotrichosis of the eyelashes (inadequate or not enough eyelashes) and related eyelash deficiencies. The objective of this research is to present the methodology and results from qualitative and quantitative research that led to the development and validation of a new PRO measure, the Eyelash Satisfaction Questionnaire (ESQ).

## METHODS

The ESQ was developed in 2 phases: qualitative and quantitative. A clinical study registry and identification number are not available because this study was executed before the requirements for study registration.

### Qualitative Phase

The qualitative phase consisted of a literature review, concept elicitation surveys and focus groups, expert panel discussions, item generation, and cognitive debriefing. Eligible participants were aged 18 years or older who were able to read, write, and speak English. The study was conducted in accordance with the principles of the Declaration of Helsinki and all study procedures were conducted in accordance with the requirements of the institutional review board (Western IRB). All participants provided written informed consent.

#### Literature Review

A comprehensive literature review (ie, gap analysis) using the online databases PSYCHinfo (<http://www.apa.org/pubs/databases/psycinfo/index.aspx>) and PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) was undertaken. The purpose was to identify existing eyelash-related satisfaction instruments and important concepts pertaining to participants' satisfaction related to their eyelashes, including appearance of eyelashes, daily use of mascara, and satisfaction with products that enhance eyelash appearance, to guide the development of the conceptual framework model and instrument. Search terms included eyelash and each of the following: quality of life, health outcomes, growth, satisfaction, health, facial aesthetics, lash self-perception, and satisfaction. The search was conducted in January 2006, was not limited to English-language papers, no cutoff date was applied.

#### Concept Elicitation Surveys

In a proof of concept study conducted between December 2005 and January 2006, treatment outcomes were assessed in participants who wished to grow more prominent eyelashes.<sup>6</sup> This open-label, clinical study enrolled 29 participants who were at least 18 years of age and did not have any uncontrolled systemic disease or any known ocular disease or abnormality. Participants received daily treatment with bimatoprost ophthalmic solution 0.03% for 12 weeks. They were administered a draft 36-item questionnaire derived from a clinician's assessment of the effect of treatment on participants' perceptions about their satisfaction related to their eyelashes and were also asked to provide free-response feedback concerning their treatment. This free-response feedback was used to help structure the interview guide for the concept elicitation focus groups.

#### Concept Elicitation Focus Groups

A total of 4 focus group discussions were conducted in July 2006 (2 groups in Irvine, California; 2 groups in Chicago, Illinois) for concept elicitation interviews among participants from the general population. In each city, 1 group consisted of women aged 20 to 45 years and the other consisted of women aged 46 to 70 years (note: no participants aged between 18 and 19 years and no men were recruited). These participants had previously been identified as being willing to participate in market research and were called to screen and confirm participation. Participants were compensated for their time. All participants were recruited through telephone screening. Participants were required to be able to write, speak, and read English. Discussions were designed to confirm the relevance of the concepts related to satisfaction with the eyelashes identified during the proof of concept study and the literature review, to further develop any additional relevant concepts, and to elucidate the language participants use to describe concepts.

#### Item Generation

Findings of the concept elicitation focus groups were used to refine the draft 36-item questionnaire from the proof of concept study. Content-based analysis involving review of the transcripts and video tapes from the semi-structured discussions was undertaken in order to identify recurring and coherent themes. Modifications were made to the items based on the endorsement from the focus group discussions and subsequently from expert panel discussions.

#### Expert Panel Discussions

Following the concept elicitation focus groups, a total of 6 expert panel focus group discussions were conducted in January 2007 (3 groups in Los Angeles, California; 3 groups in New York, New York) to help refine the conceptual framework of the ESQ. In addition, coherence and comprehensiveness of the draft items were evaluated as a

preparation step for cognitive debriefing in participants. Each expert panel group consisted of 6 to 8 clinicians (all dermatologists, plastic surgeons, or ophthalmologists).

### **Further Testing with Concept Elicitation Focus Groups**

Following the expert panel discussions, 2 additional focus groups were conducted in New York, New York, using a 2-stage approach. The first stage was designed to elicit spontaneous responses from participants. The second was used to pilot test the draft ESQ using cognitive interviews (described below).

### **Cognitive Interviews**

The draft ESQ was examined for face validity and comprehensibility in 2 focus groups using the semi-structured cognitive debriefing interviews for testing the draft ESQ and discussing participants' reactions to the items.<sup>7</sup> Following each set of cognitive interviews, recommendations for changes to the language and the instrument format were discussed.

## **Quantitative Phase**

### **Psychometric Evaluation**

The measurement properties of the draft ESQ were evaluated using 2 convenience samples (Internet and paper samples) designed to provide psychometric justification for the final instrument items, assess the psychometric properties of the draft ESQ, and evaluate the effect of mode of administration (Internet vs paper-and-pencil). The Internet samples comprised volunteers who agreed to participate in online market research (Analytica; New York, NY); the paper-based sample comprised volunteers from a market research database (Caughlin Research; Phoenix, AZ). Both groups were willing to answer questions about their eyelashes. The draft ESQ underwent modern and classical psychometric analyses to evaluate item- and domain-level properties.

### **Modern Psychometrics**

Modern psychometric analyses were used to guide scale construction. Confirmatory factor analysis (CFA) was conducted to assess the latent structure of the questionnaire, (ie, the extent of correspondence of the proposed scoring system with the latent constructs). Initial CFA was conducted using the data from the proof of concept study questionnaire.<sup>6</sup> A split-sample validation methodology was used with CFA, by randomly separating the Internet sample into development (one-third) and validation (two-thirds) subsamples. Mplus latent variable software version 5.1 (Muthén & Muthén; Los Angeles, CA) was used to perform multivariate normal CFA.<sup>8-10</sup> The CFA was assessed also using a weighted least squares mean and variance adjusted (WLSMV) estimation with categorical indicators to ensure similar validity was achieved even if the assumption of multivariate normality was violated.<sup>8</sup>

Measurement invariance was subsequently assessed using multiple group confirmatory factor analysis (MG-CFA). The invariance models assessed were (in order of increasing stringency) (1) Configural, (2) Threshold, (3) Loading/Discrimination, (4) Residual, and (5) Factor. Significance was assessed using chi-squared difference tests. Given the number of comparisons involved, a conservative *P*-value threshold was adopted ( $P < .01$ ).

Following measurement invariance, an analysis of differential item functioning (DIF) was performed through ordinal logistic regression methodology to explore possible item bias resulting from different modes of questionnaire administration (ie, Internet-based and paper-based versions). The magnitude of DIF was quantified using a pseudo- $r^2$  difference measure.<sup>11,12</sup> Because multiple tests were performed, a 2-part criterion for considering the DIF of an item was applied: statistical significance (a conservative  $P < .01$  was adopted) and magnitude of DIF ( $r^2$  difference [ $\Delta r^2$ ] of at least 0.035).<sup>13,14</sup>

An item response theory model was then developed for evaluation of item fit and scale monotonicity. A maximum marginal likelihood estimation procedure<sup>15</sup> in PARSCALE 4.1 (Scientific Software International, Inc., Skokie, IL, USA) was used to examine item properties.<sup>16</sup> Item information functions were calculated using SAS V9.1 (SAS Institute Inc., Cary, NC, USA) for each item by comparing expected and observed frequency distributions.<sup>17,18</sup>

### **Classical Psychometrics**

The measurement properties (internal consistency and test-retest reliability, construct validity) of the ESQ was assessed by classical psychometric analyses to provide results that could be compared with similar measures and against current set criteria.<sup>19-21</sup>

## **RESULTS**

### **Qualitative Findings**

#### **Literature Review**

The literature review revealed a lack of published papers concerning the development and/or validation of eyelash-related satisfaction instruments. No relevant articles on eyelash satisfaction or health outcomes were identified at the time of the search. Therefore, development of items for the ESQ was guided almost exclusively by participant and expert input.

#### **Concept Elicitation Focus Groups**

A total of 32 participants in 4 focus groups of 8 participants each were interviewed to elicit concepts.

#### **Item Reconciliation**

A content analysis of the focus groups was used to modify the draft items in the 36-item questionnaire originally generated by clinicians to improve clarity and precision and to

incorporate new concepts consistently mentioned by the participants. Only items relevant to the concepts were retained (based on identified conceptual categories and initial CFA analyses). The process of item generation led to a final set of 23 items included in the draft ESQ for administration in the cognitive interviews. Items were grouped into 3 domains based on their conceptual meaning to represent coherent clinically meaningful constructs: (1) Length, Fullness, and Overall Satisfaction (LFOS), (2) Confidence, Attractiveness, and Professionalism (CAP), and (3) Daily Routine (DR).

### Expert Panel Discussions

Expert panel discussions in California and New York were conducted with a total of 41 clinicians. The clinicians' feedback confirmed findings from the concept elicitation surveys and interviews regarding the importance of participants' perceptions of physical appearance, feelings about appearance, and the burden of daily usage of eyelash products such as mascara. Vocabulary and anchor improvements per clinicians were incorporated into the draft ESQ administered to the validation sample.

### Further Testing with Concept Elicitation Focus Groups

Two concept elicitation focus groups were conducted in 16 participants (7 participants in 1 focus group and 9 participants in the other focus group). New items were added for comprehensiveness of concepts and incorporation of additional constructs that were identified through participant input.

### Cognitive Interviews

Sixteen participants participated in the cognitive interviews to provide feedback on the clarity of each item. Assessment of item redundancy was also undertaken in order to ensure that the questionnaire had minimal responder burden. Cognitive debriefing was considered to be complete as these groups had no recommended changes.

## Quantitative Findings

### Psychometric Evaluation

A total of 928 participants completed the Internet-based version of the ESQ (Internet sample) and 68 participants completed the paper-based version of the questionnaire (paper sample). Male respondents ( $n = 26$ ) were removed from the total sample due to their small number and lack of representation during focus group interviews, resulting in an Internet sample of 909 participants and a paper sample of 61 participants. Sociodemographic characteristics of the combined validation samples are shown in Table 1. The descriptive statistics for all questionnaire items are provided in Table 2.

### Modern Psychometrics

CFA results based on all 23 items confirmed monotonicity of all items but provided unacceptable fit with the data, with a comparative fit index (CFI) of 0.86 and a non-normed fit index (NNFI) of 0.82. A revised model was developed based on the initial CFA and content-based analysis, resulting in 9 items as indicators of the 3 ESQ domains: LFOS domain (Items 1, 2, and 4), CAP domain (Items 16, 18, and 19), and DR domain (Items 9, 10, and 11). Results of the revised model showed good model fit, with a CFI of 0.99 and NNFI of 0.97 (Figure 1). The 9 items representing the 3 domains of the conceptual framework of the ESQ were assessed for their psychometric properties in the analyses that follow. These findings indicate that scoring of the conceptual domains is limited to the 9 items in the final validated conceptual model. Note that 14 residual items were retained in the questionnaire based on participants' input regarding relevance and importance to overall satisfaction related to eyelashes (see Discussion for further details).

MG-CFA confirmed equivalency among the Internet- and paper-based versions of the ESQ. Differential item functioning was not found for any of the 9 items representing the ESQ domains, indicating that there were no differences in the way participants responded to the questions based on the mode of administration. Therefore, all remaining psychometric analyses used pooled data from the 2 ESQ versions.

### Classical Psychometrics

Internal consistency reliability of the domains was measured with 3 indices: Cronbach's alpha ( $\alpha$ ), a Spearman-Brown adjusted  $\alpha$  to a 10-item domain ( $\alpha_{10}$ ), and the average inter-item correlation ( $r_{ij}$ ). Cronbach's  $\alpha$  values for the LFOS, CAP, and DR domains were found to range from excellent (0.925 for LFOS and 0.926 for CAP) to fair (0.772 for DR). None of the domains had an improvement in Cronbach's  $\alpha$  upon removal of an item, suggesting that each item was sufficiently reliable in its respective scale. A summary of domain-level psychometrics is presented in Table 3.

The average item-total correlations for the LFOS, CAP, and DR domains were 0.85, 0.84, and 0.61, respectively. Equality of item-total correlations for each domain was determined to be appropriate for each item on each domain, with little deviation. This finding suggests that the items measure the domain that they form to a similar degree without requiring special weighting. Corrected item-total correlations were considered to be sufficient if they were greater than or equal to 0.40. The variances of items in each domain were found to be similar, indicating that each set of items could be appropriately summed to represent their respective domain score. In addition, each item had a sufficiently higher correlation with its own domain as opposed to its correlation with other domains. This finding

**Table 1. Sociodemographic Characteristics of the Validation Samples**

	Internet Sample (n = 909)		Paper Sample (n = 61)	
	No.	%	No.	%
<b>Marital status</b>				
Single	264	29.0	NC	NC
Married	494	54.3	NC	NC
Otherwise	151	16.7	NC	NC
<b>Race</b>				
Black	25	2.8	0	0
Asian	22	2.4	5	8.2
Caucasian	833	91.8	52	85.2
Hispanic	0	0	2	3.3
Native American, other, or mixed	27	3.0	2	3.3
Declined to answer	2	0.2	0	0
<b>Education</b>				
High school or less	233	25.7	NC	
Some college	331	36.4		
Bachelor's or associate's degree	288	31.7		
Higher education	56	6.2		
Declined to answer	1	0.3		
<b>Annual income</b>				
<\$35,000	321	35.3	NC	
\$35,000 but < \$65,000	328	36.2		
\$65,000 but < \$100,0000	130	14.3		
>\$100,000	92	10.1		
Declined to answer	38	4.1		

NC, data were not collected for this sample.

indicates that there were no items which measure multiple domains. Item-level psychometrics are presented in Table 4.

## DISCUSSION

In the advent of a new medication to treat hypotrichosis of the eyelashes, a PRO measure was warranted to assess factors of importance in eyelash appearance. Participant interview data and expert opinion informed the development of the conceptual framework. The main concepts within the framework were related to the physical attributes of eyelashes (length and fullness), subjective attributes of eyelash health (confidence, attractiveness, and professionalism),

and perceptions of the degree of bother associated with making eyelashes presentable (daily routine). The results from the modern and classical psychometric analyses provided substantial evidence for the ESQ being an adequate and well-developed PRO measure.

The ESQ was developed using rigorous qualitative and quantitative methods to be in accordance with the FDA recommendations outlined in the FDA Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.<sup>19,20</sup> The FDA recommends that a PRO measure should be shown in a sequential process to be (1) valid in terms of content, (2) reliable, (3) valid in terms of construct, (4) able to detect

Table 2. Descriptive Statistics for the Eyelash Satisfaction Questionnaire

ESQ Item	Validation	
	Range	Mean (SD)
Q1. Length satisfaction	1-5	2.70 (1.14)
Q2. Fullness satisfaction	1-5	2.97 (1.11)
Q3. Darkness satisfaction	1-5	2.77 (1.13)
Q4. Overall satisfaction	1-5	2.78 (1.09)
Q5. How often receive compliments	1-5	3.74 (1.13)
Q6. Rate eyelash length	1-5	2.91 (0.82)
Q7. Rate eyelash fullness/thickness	1-5	3.16 (0.76)
Q8. Rate eyelash color	1-5	2.77 (0.96)
Q9. Time applying mascara	1-5	3.43 (1.04)
Q10. Time removing mascara	1-5	2.96 (1.19)
Q11. Hassle with eyelashes	1-5	3.06 (1.15)
Q12. Can go out in public w/o mascara	1-5	2.48 (1.27)
Q13. Worry about mascara smearing	1-5	2.68 (1.17)
Q14. Eyes look tired without mascara	1-5	2.86 (1.18)
Q15. Eyelashes naturally attractive	1-5	3.17 (1.09)
Q16. Feel confident about looks	1-5	3.28 (1.06)
Q17. Confident to go out in public	1-5	3.16 (1.06)
Q18. Look professional	1-5	3.39 (1.07)
Q19. Feel attractive	1-5	3.34 (1.05)
Q20. Eyelashes look healthy	1-5	2.72 (1.01)
Q21. Eyes look vibrant	1-5	3.51 (1.0)
Q22. Eyelashes look full	1-5	3.51 (1.1)
Q23. Feel beautiful	1-5	3.41 (1.1)
LFOS Domain score <sup>a,b</sup>	3-15	8.45 (3.11)
CAP Domain score <sup>a,b</sup>	3-15	10.00 (2.97)
DR Domain score <sup>a,b</sup>	3-15	9.45 (2.80)

CAP, confidence, attractiveness, and professionalism; DR, daily routine; ESQ, Eyelash Satisfaction Questionnaire; LFOS, length, fullness, and overall satisfaction; SD, standard deviation. Green cells, LFOS items; Blue cells, CAP items; Orange cells, DR items. <sup>a</sup>Based on linear combination of items in each domain as reported in the finalized factor structure. <sup>b</sup>A low score in the LFOS and CAP domains is indicative of a high degree of satisfaction, whereas in the DR domain a high score indicates a high degree of satisfaction.

clinical change, and (5) able to establish a threshold for treatment benefit. This article reports on the first 3 steps of this process.

The content validity of the ESQ was established through qualitative research (concept elicitation surveys and interviews, item generation, and cognitive interviews). During

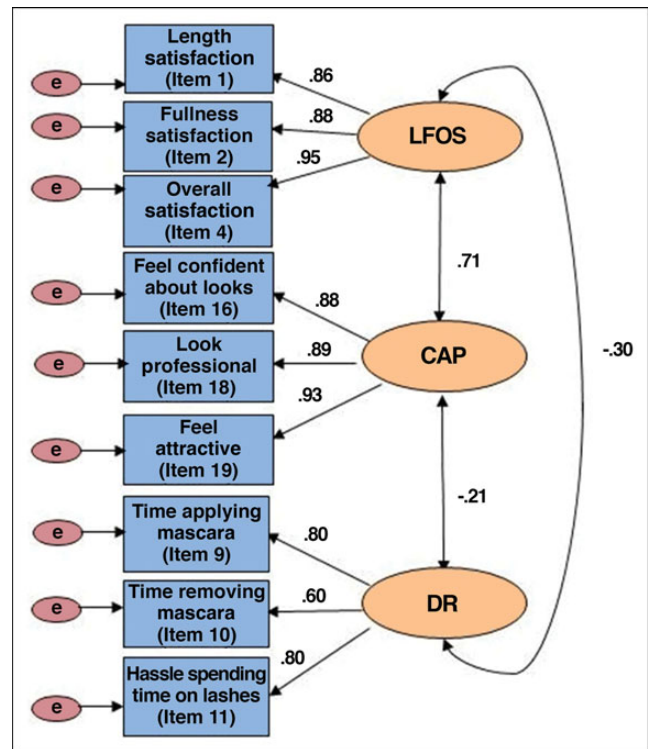


Figure 1. Conceptual framework and factor structure (final model) of Eyelash Satisfaction Questionnaire (N = 970). CAP, confidence, attractiveness, and professionalism; DR, daily routine; LFOS, length, fullness, and overall satisfaction.

psychometric evaluation, measurement with the ESQ was shown to improve when reduced to a short-form version of the ESQ (referred to as the 9-item ESQ [ESQ-9]). Using modern psychometric methods, the ESQ-9 obtained strong model fit in the CFA models and demonstrated measurement equivalence among Internet- and paper-based versions from MG-CFA. From a classical psychometric approach, the ESQ-9 exceeded recommended criteria for measurement. It is noted that the discriminant validity of the DR domain was nominal, potentially due to the unavailability of eyelash-specific discriminant measures. Additionally, average inter-item and item-total correlations were not as high as in the other domains, possibly because the measurement concepts in the DR domain were marginally related; despite this, the domain still satisfied psychometric threshold criteria and was deemed reliable.

This study has several limitations. The ESQ is not considered valid for participant groups that were not represented in the development process (eg, men, who were omitted because of small sample sizes). Furthermore, participant perceptions of eyelash prominence are influenced by their cultural environment, and the ESQ was developed solely in a US population. However, cross-national equivalence studies were later completed to establish the reliability and validity for the ESQ in the United Kingdom, China, and

**Table 3. Domain-Level Psychometrics of Eyelash Satisfaction Questionnaire**

Domain	Reliability			Validity Fixed Effects		
	$\alpha$	$\alpha_{10}$	$r_{ij}$	$Z_{convergent}$	$Z_{discriminant}$	$Z_{diff} (P)$
LFOS	0.925	0.976	0.802	0.15	0.04	0.11 (<.05)
CAP	0.926	0.976	0.802	0.15	0.01	0.14 (<.05)
DR	0.772	0.919	0.530	0.16	0.13	0.03 (NS)

$\alpha$ , coefficient alpha;  $\alpha_{10}$ , coefficient alpha with a Spearman-Brown correction to a 10-item scale; CAP, confidence, attractiveness, and professionalism; DR, daily routine; LFOS, length, fullness, and overall satisfaction; NS, not significant;  $r_{ij}$ , average inter-item correlation;  $Z_{diff}$ , Fisher’s z difference of average  $Z_{convergent} - Z_{discriminant}$ .

**Table 4. Item Total Correlations for Eyelash Satisfaction Questionnaire**

ESQ Item	ESQ Domain			Correlations			Variance	$\alpha_{removed}$
	LFOS	CAP	DR	LFOS vs CAP, $Z_{diff} (P)$	LFOS vs DR, $Z_{diff} (P)$	CAP vs DR, $Z_{diff} (P)$		
1	0.83	0.55	-0.22	0.55 (<.001)	0.95 (.001)	-	1.30	0.91
2	0.83	0.61	-0.24	0.47 (<.001)	0.94 (<.001)	-	1.23	0.90
4	0.88	0.65	-0.28	0.59 (<.001)	1.08 (<.001)	-	1.19	0.86
16	0.60	0.83	-0.20	0.50 (<.001)	-	0.99 (<.001)	1.12	0.90
18	0.59	0.84	-0.17	0.56 (<.001)	-	1.06 (<.001)	1.14	0.89
19	0.64	0.86	-0.18	0.54 (<.001)	-	1.12 (<.001)	1.10	0.88
9	-0.24	-0.16	0.66	-	0.55 (<.001)	0.63 (<.001)	1.08	0.64
10	-0.11	-0.11	0.54	-	0.49 (<.001)	0.50 (<.001)	1.42	0.77
11	-0.33	-0.23	0.63	-	0.40 (<.001)	0.50 (<.001)	1.30	0.67

*Note:* The first 3 columns indicate corrected item-total correlations for the correlations in colored boxes (ie, items with their respective total scores).  $\alpha_{removed}$ , alpha if item removed (compare with alpha for CAP = 0.926, LFOS = 0.925, DR = 0.772); CAP, confidence, attractiveness, and professionalism; DR, daily routine; ESQ, Eyelash Satisfaction Questionnaire; LFOS, length, fullness, and overall satisfaction.  $Z_{diff}$ , difference between Fisher’s z transformation of the correlations.

Japan; the ESQ has also been translated and linguistically validated in Swedish and Russian languages. Additionally, since the electronic and paper questionnaires were not equal in number, no comparability statements can be made.

**CONCLUSIONS**

The ESQ can provide critical information about the effectiveness of eyelash hypotrichosis treatment from the participant’s perspective. Rigorous qualitative and quantitative methodologies were used to develop this instrument, which complies with FDA standards.<sup>19,20</sup> Evidence supports the reliability and validity of the ESQ for assessing participant satisfaction with eyelash prominence.

**Acknowledgment**

The authors thank Geoffrey C. Hammond, PhD, of NMAHS Mental Health Adult Program, Mt. Hawthorn, Western Australia, for his contributions to data analysis, statistical support, and data interpretation.

**Disclosures**

The Consulting Measurement Group (Dr Dang) and QualityMetric, Inc., (Drs Cole and Yang) were contracted by Allergan, Inc. for this research. Dr Burgess, Dr Daniels, and Mr Walt were employees at Allergan, Inc., at the time of this research.

**Funding**

Funding by Allergan, Inc. supported focus group-related costs, including participant compensation and travel. The sponsor and coauthors participated in the study design, statistical analysis, and interpretation of results.

**REFERENCES**

1. Woodward JA, Haggerty CJ, Stinnett SS, Williams ZY. Bimatoprost 0.03% gel for cosmetic eyelash growth and enhancement. *J Cosmet Dermatol.* 2010;9(2):96-102.
2. Cohen JL. Enhancing the growth of natural eyelashes: the mechanism of bimatoprost-induced eyelash growth. *Dermatol Surg.* 2010;36(9):1361-1371.

3. Flay BR, Biglan A, Boruch RF, et al. Standards of evidence: Criteria for efficacy effectiveness and dissemination. *Prev Sci*. 2005;6(3):151-175.
4. Bottomley A, Jones D, Claassens L. Patient-reported outcomes: Assessment and current perspectives of the guidelines of the Food and Drug Administration and the reflection paper of the European Medicines Agency. *Eur J Cancer*. 2009;45(3):347-353.
5. Kwak TJ, Lee SM, Cho WG. The character of eyelashes and the choice of mascara in Korean women. *Skin Res Technol*. 2002;8(3):155-163.
6. Yoelin S, Walt JG, Earl M. Safety, effectiveness, and subjective experience with topical bimatoprost 0.03% for eyelash growth. *Dermatol Surg*. 2010;36(5):638-649.
7. Willis GB. Cognitive interviewing revisited: A useful technique, in theory? In: Presser S, Rothgeb JM, Couper MP, et al. *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken, NJ: John Wiley & Sons Inc; 2004.
8. Muthen LK, Muthen BO. *Mplus User's Guide*. 5th ed. Los Angeles, CA: Muthen & Muthen; 1998-2007.
9. Bollen KA, Stine RA. Bootstrapping goodness-of-fit measures in structural equation models. *Socio Methods Res*. 1992;21:205-229.
10. Nevitt J, Hancock GR. Evaluating small sample approaches for model test statistics in structural equation modeling. *Multivariate Behav Res*. 2004;39:439-478.
11. Nagelkerke NJ. A note on a general definition of the coefficient of determination. *Biometrika*. 1991;78:691-692.
12. Zumbo BD. *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (ordinal) Item Scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense; 1999.
13. Bjorner JB, Kosinski M, Ware JE Jr. Calibration of an item pool for assessing the burden of headaches: An application of item response theory to the Headache Impact Test (HIT). *Qual Life Res*. 2003;12(8):913-933.
14. Jodoin MG, Gierl MJ. Evaluating Type 1 error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Appl Meas Educ*. 2001;14:329-349.
15. Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*. 1981;46:443-459.
16. du Toit M. *IRT From SSI*. Lincolnwood, IL: Scientific Software International, Inc; 2003.
17. Orlando M, Thissen D. Likelihood-based item-fit indices for dichotomous item response theory models. *Appl Psychol Measure*. 2000;24:50-64.
18. Orlando M, Thissen D. Further investigation of the performance of S-X squared: An Item Fit Index for Use with Dichotomous Item Response Theory Models. *Appl Psychol Measure*. 2003;27:289-298.
19. US Department of Health and Human Services and Food and Drug Administration. Final Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>. Accessed October 31, 2014.
20. US Department of Health and Human Services and Food and Drug Administration. Draft Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2006. <http://www.fda.gov/OHRMS/DOCKETS/98fr/06d-0044-gdl0001.pdf>. Accessed October 31, 2014.
21. Aaronson N, Alonso J, Burnam A, et al. Assessing health status and quality of life instruments: attributes and review criteria. *Qual Life Res*. 2002;11(3):193-205.