It's all in the mind: regression analysis of appearance satisfaction, conscientiousness and age based on resting state EEG data

Petra Šurla Toth
University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technologies
Koper, Slovenia
petra.surla@gmail.com

Jakob Sajovic
University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technologies
Koper, Slovenia
University Medical Centre Ljubljana, Division of Stomatology
Ljubljana, Slovenia

Peter Rogelj
University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technologies
Koper, Slovenia

Abstract

Satisfaction with appearance is an important part of psychosocial well-being, especially in young people with cleft lip and/or palate. Facial differences can affect how they see themselves and how they interact with others. This study evaluated whether resting-state electroencephalography (EEG) connectivity contains sufficient information to estimate satisfaction with appearance. It also tested whether prediction accuracy differs between psychological and objective personal characteristics. Appearance satisfaction, conscientiousness, and age were considered as three target variables with different expected levels of prediction difficulty.

Resting-state EEG recordings from 54 adolescents with cleft lip and/or palate were analysed. The seven-minute recordings were collected with eyes closed using 32 EEG channels and divided into 10-second segments. Functional and effective connectivity features were calculated using four methods: phase-locking value (PLV), weighted phase lag index (wPLI), complex Pearson correlation coefficient (CPCC), and Granger causality (GC). The 10% most relevant features were selected using minimum redundancy maximum relevance method, and linear support vector regression models were evaluated using two data-splitting approaches.

In the intersubject analysis, participants in the training and test sets were completely separated, whereas in the intrasubject analysis, segments from the same participant could occur in both sets. The intrasubject models achieved positive test-set R^2 values for all three target variables, with CPCC providing the best overall performance. However, all intersubject test-set R^2 values were negative. These results suggest that resting-state EEG contains information related to the target variables, but the models did not perform well on new participants. These findings show the importance of participant-independent testing when evaluating EEG-based prediction models.

Keywords

EEG, cleft lip and/or palate, satisfaction with appearance, CHASQ, conscientiousness, brain activity, functional connectivity, PLV, wPLI, CPCC, GC, intersubject analysis, intrasubject analysis

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Information Society 2026, Ljubljana, Slovenia
© 2026 Copyright held by the owner/author(s).
https://doi.org/10.70314/is.2026.cogni.75

1 Introduction

Young people with cleft lip and/or palate may experience difficulties with speech, hearing, and feeding, as well as emotional and social challenges related to their appearance. Satisfaction with appearance can affect self-image and social life. However, it cannot be explained only by the presence or severity of a visible facial difference. Previous studies [8] have shown that dissatisfaction with appearance is linked to emotional and social difficulties, but it is still not clear which factors influence why some young people are more satisfied with their appearance than others. This suggests that appearance satisfaction is shaped by a combination of psychological, social, and developmental factors [8].

Because clinical outcomes do not fully reflect a person's own experience, patient-reported outcome measures can provide important additional information. One such measure is the Cleft Hearing, Appearance and Speech Questionnaire (CHASQ), which allows young people with cleft lip and/or palate to report their satisfaction with appearance, hearing, and speech. This also raises the question of whether individual differences in appearance satisfaction can be reflected in measurable patterns of brain activity.

To explore this question, electroencephalography (EEG) can be used as a non-invasive method for measuring electrical brain activity with high temporal resolution [2]. Brain connectivity analysis can further describe relationships between signals recorded from different brain regions. Different connectivity measures can therefore capture different aspects of these relationships [5]. Phase-locking value (PLV) measures phase synchronisation [1], while weighted phase lag index (wPLI) is less sensitive to volume conduction and noise [17]. Complex Pearson correlation coefficient (CPCC) provides information about complex relationships between analytic EEG signals [14], while Granger causality (GC) measures directed predictive relationships between time series [12].

The main research question of this study was whether functional and effective brain-connectivity features derived from resting-state EEG can be used to estimate appearance satisfaction, conscientiousness, and age.

The three target variables were selected to represent different expected levels of prediction difficulty. Appearance satisfaction was considered the most complex target because it is a subjective psychological construct affected by several developmental and psychosocial factors. Conscientiousness represented an intermediate level as a personality trait, while age was considered the simplest because it can be measured directly.

Accordingly, the following hypotheses were evaluated:

- **H1:** Appearance satisfaction can be estimated from functional and effective connectivity features derived from resting-state EEG.
- **H2:** Conscientiousness can be estimated from functional and effective connectivity features derived from resting-state EEG.
- **H3:** Age can be estimated from functional and effective connectivity features derived from resting-state EEG.

An important part of the study was comparing intersubject and intrasubject evaluation. This comparison showed whether patterns found in EEG data could also be used for new participants or were mainly specific to individual participants.

2 Methodology

The data were obtained from a previous study conducted at the University Medical Centre Ljubljana [11]. The sample included 54 adolescents with cleft lip and/or palate, including 23 females and 31 males. Participants were between 11.67 and 18.31 years old, with a mean age of 14.72 years.

Our three target variables were included in the analysis. Appearance satisfaction was measured using a 15-item version of the CHASQ questionnaire, where higher scores indicated greater satisfaction. Conscientiousness was obtained from the VMR-OM personality questionnaire, while age was used as the third target variable.

For each participant, seven minutes of resting-state EEG with closed eyes were analysed. EEG was recorded using a g.Tec system with 32 wet electrodes positioned according to the 5–10 electrode placement system, with a sampling frequency of 500 Hz [10].

The recordings were processed in MATLAB [3] and divided into non-overlapping 10-second segments. Connectivity matrices were calculated independently for each segment using PLV, wPLI, CPCC, and GC.

PLV and wPLI produced 496 connectivity features. CPCC included both the absolute value and imaginary component and therefore produced 992 features. GC measures directed relationships between channels and therefore also produced 992 features.

Because each participant had a single value for each target variable, the corresponding appearance satisfaction, conscientiousness, or age value was assigned to all EEG segments belonging to that participant. Since the number of connectivity features was large, the minimum redundancy maximum relevance (MRMR) algorithm was used to rank them [13]. The top 10% of features were retained, and feature selection was performed only on the training data.

After feature selection, linear support vector regression models were trained. The data were standardised and model hyperparameters were optimised [7, 4]. For model evaluation, the data were divided into 70% training and 30% testing subsets using two different approaches: intersubject and intrasubject analysis.

In the *intersubject analysis*, participants were divided between the training and test sets. All segments from the same participant were therefore kept in only one set. This approach was used to test whether the model could generalise to new participants who were not included in the training data.

In the *intrasubject analysis*, individual EEG segments were randomly divided between the training and test sets. As a result, segments from the same participant could appear in both sets. This approach was used to evaluate the model's performance on unseen EEG segments.

Finally, model performance was evaluated using root mean square error (RMSE) [6], the coefficient of determination (R^2), and adjusted R^2 [15].

3 Results

Table 1 shows the test results for all three target variables, four connectivity methods, and both analysis approaches.

A clear difference was observed between the two analysis approaches. In the intrasubject analysis, all test-set R^2 values were positive. CPCC consistently achieved the strongest overall performance across all three target variables. For appearance satisfaction, CPCC achieved RMSE = 14.59, $R^2 = 0.734$, and adjusted $R^2 = 0.688$. For age, the corresponding values were RMSE = 1.08, $R^2 = 0.622$, and adjusted $R^2 = 0.554$. For conscientiousness, the results were RMSE = 1.04, $R^2 = 0.768$, and adjusted $R^2 = 0.728$.

The intersubject analysis showed different results. All test-set R^2 and adjusted R^2 values were negative for all three target variables and all four connectivity methods. Negative R^2 values mean that the model performed worse than simply predicting the mean value of the target variable for all cases. This indicates that the models were not able to reliably estimate the target variables for new participants. For appearance satisfaction, wPLI produced the lowest test RMSE of 26.62 and the least negative R^2 of -0.205. For conscientiousness, wPLI again had the lowest RMSE of 2.86 and the least negative R^2 of -1.027. For age, wPLI produced the lowest RMSE of 2.09, while GC obtained the least negative R^2 of -0.545.

CPCC showed the best overall performance in the intrasubject analysis, while wPLI performed best in the intersubject analysis. GC showed the poorest performance on new participants.

4 Discussion

The main finding of this study was the clear difference between the intersubject and intrasubject analyses. In the intrasubject analysis, the models achieved positive R^2 values for all three target variables. However, when participants in the training and test sets were completely separated, all test-set R^2 values were negative.

The positive intrasubject results show that the EEG segments contain information related to appearance satisfaction, conscientiousness, and age. Still, these patterns were not stable enough to generalise to new participants. For this reason, the intersubject analysis was more important for evaluating the hypotheses. Since all test-set R^2 values were negative, none of the three hypotheses was confirmed in the intersubject analysis. **H1**, which assumed that appearance satisfaction could be estimated from resting-state EEG connectivity features, was not confirmed. **H2**, concerning the estimation of conscientiousness, and **H3**, concerning age, were also not confirmed.

The problem with generalisation was not limited to appearance satisfaction. Similar results were found for conscientiousness and age. This suggests that the main difficulty was not only the complexity of the target variables, but also the ability of the models to find patterns that remain stable across different participants.

This is also seen in the results for conscientiousness, which are similar to those reported by Korjus et al. [9]. Their study found that Big Five personality traits could not be reliably predicted from resting-state EEG [9]. Although a different approach was

Table 1: Test results of the regression models

Target	Method	Intersubject			Intrasubject		
		RMSE	R^2	Adj. R^2	RMSE	R^2	Adj. R^2
CHASQ	PLV	37.75	-1.423	-1.624	21.22	0.457	0.412
	wPLI	**26.62**	**-0.205**	**-0.305**	24.87	0.208	0.144
	CPCC	40.32	-1.765	-2.259	**14.59**	**0.734**	**0.688**
	GC	86.34	-11.675	-13.940	17.96	0.590	0.518
Age	PLV	2.61	-1.794	-2.026	1.27	0.481	0.437
	wPLI	**2.09**	-0.634	**-0.769**	1.64	0.110	0.035
	CPCC	2.40	-1.309	-1.721	**1.08**	**0.622**	**0.554**
	GC	29.88	**-0.545**	-0.821	21.08	0.441	0.341
Conscientiousness	PLV	3.67	-2.411	-2.694	1.31	0.616	0.584
	wPLI	**2.86**	**-1.027**	**-1.195**	1.84	0.237	0.175
	CPCC	3.87	-2.717	-3.381	**1.04**	**0.768**	**0.728**
	GC	86.82	-10.321	-12.344	17.55	0.597	0.527

used, the results were similar. This suggests that conscientiousness is difficult to predict from EEG, especially when the model is tested on new participants.

The findings for age showed a different pattern. Vandenbosch et al. [16] successfully predicted age from resting-state EEG in children and adolescents [16]. Their study included 1538 participants and a wider age range, while our study included a much smaller group and a narrower age range. These differences may have made age prediction more difficult in our analysis.

Among the connectivity methods, CPCC showed the best overall performance in the intrasubject analysis. One possible reason is that CPCC includes both the absolute value and the imaginary component of complex connectivity, which provides more information about relationships between EEG signals [14]. However, these results did not generalise well to new participants. In the intersubject analysis, wPLI performed best, while GC showed the largest errors and the poorest generalisation to new participants.

One of the main limitations of this study was the relatively small number of participants compared with the large number of EEG connectivity features. Although MRMR was used to reduce the number of features, differences between participants were still difficult to model.

Future studies should first repeat the analysis with a larger sample, because a larger dataset is necessary for more reliable intersubject modelling. Once enough data are available, other feature-selection methods and machine-learning models could be tested. Different EEG segment lengths and modified connectivity estimation methods could also be explored.

5 Conclusion

This study examined whether functional and effective connectivity features from resting-state EEG can be used to estimate appearance satisfaction, conscientiousness, and age in adolescents with cleft lip and/or palate. Four connectivity methods and two different evaluation approaches were compared.

The results showed that model performance strongly depended on how the data were divided. Intrasubject analysis produced positive results for all three target variables, with CPCC showing the best overall performance. Still, the models did not generalise successfully to new participants, as all intersubject test-set R^2 values were negative.

The proposed hypotheses were therefore not supported in the intersubject analysis. Nevertheless, the positive intrasubject results indicate that resting-state EEG contains patterns related to the investigated characteristics. These patterns, however, were not stable enough to generalise reliably to new participants.

An important contribution of this study is the comparison between intersubject and intrasubject analysis. The results show that good performance on new EEG segments from previously seen participants does not necessarily mean that the model will also perform well on completely new participants. This difference is important when interpreting the results of EEG-based prediction models.

Overall, the results provide a useful basis for further research on the relationship between brain connectivity and individual characteristics. Although the hypotheses were not confirmed, the study provides a better understanding of the limitations of using resting-state EEG connectivity features for prediction. Further research in this area requires careful study design, larger samples, and strict evaluation of model generalisation.

References

[1] Sergul Aydore, Dimitrios Pantazis, and Richard M. Leahy. 2013. A note on the phase locking value and its properties. *NeuroImage*, 74, (July 2013), 231–244. doi:10.1016/j.neuroimage.2013.02.008.

[2] Andrea Biasiucci, Benedetta Franceschiello, and Micah M. Murray. 2019. Electroencephalography. eng. *Current biology: CB*, 29, 3, (Feb. 2019), R80–R85. doi:10.1016/j.cub.2018.11.052.

[3] Arnaud Delorme and Scott Makeig. 2004. EEGLAB: an open-source toolbox for analysis of EEG dynamics. *Journal of neuroscience methods*, 134, (Apr. 2004), 9–21. doi:10.1016/j.jneumeth.2003.10.009.

[4] David Fagbuyiro. 2023. Understanding the support vector machine (SVM) model. en. (June 2023). Retrieved July 26, 2026 from https://medium.com/@davidfagb/understanding-the-support-vector-machine-svm-model-c8eb9bd54a97.

[5] Matthew Feigelis and Deanna Greene J. 2022. Functional connectivity in the Gilles de la Tourette syndrome. en-US. In *International Review of Movement Disorders*. Vol. 4. Academic Press, (Jan. 2022), 103–125. doi:10.1016/bs.irmvd.2022.04.001.

[6] Jim Frost. 2023. Root Mean Square Error (RMSE). en-US. (May 2023). Retrieved May 14, 2026 from https://statisticsbyjim.com/regression/root-mean-square-error-rmse/.

[7] Eda Kavlakoglu. 2023. What Is Support Vector Machine? | IBM. en. (Dec. 2023). Retrieved May 14, 2026 from https://www.ibm.com/think/topics/support-vector-machine.

[8] Sarah N. Kelly and Joanna Shearer. 2020. Appearance and Speech Satisfaction and Their Associations With Psychosocial Difficulties Among Young People With Cleft Lip and/or Palate. *The Cleft Palate-Craniofacial Journal*, 57, 8, (Aug. 2020), 1008–1017. doi:10.1177/1055665620926083.

[9] Kristjan Korjus, Andero Uusberg, Helen Uusberg, Nele Kuldkepp, Kairi Kreegipuu, Jüri Allik, Raul Vicente, and Jaan Aru. 2015. Personality cannot

be predicted from the power of resting state EEG. eng. *Frontiers in Human Neuroscience*, 9, 63. doi:10.3389/fnhum.2015.00063.

[10] R. Oostenveld and P. Praamstra. 2001. The five percent electrode system for high-resolution EEG and ERP measurements. eng. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 112, 4, (Apr. 2001), 713–719. doi:10.1016/s1388-2457(00)00527-7.

[11] Jakob Sajovic. 2024. *Povezava možganske aktivnosti, morfologije obraza ter psiholoških značilnosti otrok s shizo med 12. in 16. letom starosti*. Doktorska disertacija. Medicinska fakulteta, Univerza v Ljubljani, Ljubljana. https://repozitorij.uni-lj.si/Dokument.php?id=196100%5C&lang=slv.

[12] Ali Shojaie and Emily B. Fox. 2022. Granger Causality: A Review and Recent Advances. *Annual review of statistics and its application*, 9, 1, (Mar. 2022), 289–319. doi:10.1146/annurev-statistics-040120-010930.

[13] Alihan Suiçmez and Cengiz Tepe. 2022. Comparison of Different Feature Selection Methods Using Support Vector Machine for Estimating Eye Angles. In *2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. ISSN: 2770-7962. (Oct. 2022), 413–416. doi:10.1109/ISMSIT56059.2022.9932725.

[14] Zoran Šverko, Miroslav Vrankić, Saša Vlahinić, and Peter Rogelj. 2022. Complex Pearson Correlation Coefficient for EEG Connectivity Analysis. *Sensors (Basel, Switzerland)*, 22, 4, (Feb. 2022), 1477. doi:10.3390/s22041477.

[15] The Investopedia Team. 2025. R-squared vs. adjusted r-squared: what's the difference? Updated May 4, 2025; accessed May 14, 2026. (2025). https://www.investopedia.com/ask/answers/012615/whats-difference-between-rsquared-and-adjusted-rsquared.asp.

[16] Marjolein M. L. J. Z. Vandenbosch, Dennis van 't Ent, Dorret I. Boomsma, Andrey P. Anokhin, and Dirk J. A. Smit. 2019. EEG-based age-prediction models as stable and heritable indicators of brain maturational level in children and adolescents. eng. *Human Brain Mapping*, 40, 6, (Apr. 2019), 1919–1926. doi:10.1002/hbm.24501.

[17] Martin Vinck, Robert Oostenveld, Marijn van Wingerden, Franscesco Battaglia, and Cyriel M. A. Pennartz. 2011. An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *NeuroImage*, 55, 4, (Apr. 2011), 1548–1565. doi:10.1016/j.neuroimage.2011.01.055.