



(RESEARCH ARTICLE)



DO COLLEGE STUDENTS CORRECTLY IDENTIFY THEIR OWN CLASSICAL HUMAN PHENOTYPIC TRAITS? - A CROSS-SECTIONAL OBSERVATIONAL STUDY

Umarani Subburam ^{1,*}, Jeyalalitha Thirupathi ¹, Pavithran Saminathan ¹ and Mahimashree Arumugam ²

¹ PG and Research Department of Zoology, Arulmigu Palaniandavar College of Arts and Culture, Palani-624 601, Tamil Nadu, India.

² Advocate, Palani-624 601, Tamil Nadu, India.

* Corresponding Author

ORCID Details

Pavithran Saminathan: <https://orcid.org/0009-0005-2957-7449>

World Journal of Advanced Research and Reviews, 2026, 32(01), 1604–1613

Publication history: Received on 07 September 2026; revised on 20 October 2026; accepted on 22 October 2026

Article DOI: <https://doi.org/10.30574/wjarr.2026.32.1.2672>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

The present study was undertaken to assess the accuracy of self-reported and directly observed classical human phenotypic traits among undergraduate (UG) and postgraduate (PG) Zoology students. Six traits were selected for investigation: earlobe attachment, tongue rolling, widow's peak, hitchhiker's thumb, clasping of hands, and tongue folding. The study compared students' self-reported responses with direct observations to evaluate reporting accuracy and the degree of agreement between the two methods. The data were analysed descriptively using percentages and performance categories, while Cohen's kappa was used to assess agreement beyond chance. Chi-square-based analysis was also included to examine differences between responses, subject to verification of the statistical calculations and significance labels. The results revealed variation in reporting accuracy across the five academic classes and six traits. Clasping of hands showed high reporting accuracy in I B.Sc (96.6%) and III B.Sc (89.5%), while earlobe attachment also showed relatively high accuracy in I B.Sc (86.2%) and III B.Sc (84.2%). Widow's peak showed 100% accuracy in II M.Sc. In contrast, tongue folding showed low reporting accuracy in I B.Sc (27.6%) and III B.Sc (36.8%). Cohen's kappa analysis indicated substantial agreement for clasping of hands among III B.Sc students ($\kappa = 0.791$) and earlobe attachment among III B.Sc students ($\kappa = 0.617$). However, tongue folding showed only slight agreement among I B.Sc ($\kappa = 0.047$) and III B.Sc students ($\kappa = 0.081$). The findings demonstrate that agreement between self-reporting and direct observation varies according to the trait and academic class, with no single class showing consistently high agreement for all traits. The study highlights the importance of practical demonstrations, standardized observation procedures, and accurate phenotype identification in Zoology education. Since the statistical significance labels require verification, the chi-square findings should be interpreted cautiously. Furthermore, these traits are best treated as classical human phenotypic variations rather than assuming that all follow simple single-gene Mendelian inheritance.

Keywords: Self-Reported Phenotypes, Direct Observation, Reporting Accuracy, Cohen's Kappa, Chi-Square Analysis, Human Phenotypic Variation, Zoology Students.

1 INTRODUCTION

Human beings show considerable variation in their physical and behavioural characteristics. Some of these variations are influenced by genetic factors and can be observed as differences in phenotypic traits among individuals. Classical genetics has traditionally used several easily observable human characteristics to introduce concepts such as dominant and recessive inheritance, phenotype and genotype. College students frequently have some awareness of their own physical characteristics. However, their perception of a particular trait may not always correspond with direct observation. A student may report that he or she can roll the tongue, for example, while actual demonstration may show a different phenotype. Therefore, comparing self-reported information with direct observation provides an interesting way of investigating the reliability of self-identification of phenotypic characteristics.

The present study focuses on six commonly studied human phenotypic characteristics: earlobe attachment, tongue rolling, widow's peak, hitchhiker's thumb, Clasp of hands and tongue folding. Earlobes are commonly classified as free/unattached or attached. In classical genetics teaching, free earlobes have often been presented as a dominant phenotype and attached earlobes as recessive. Tongue rolling refers to the ability to curve the lateral margins of the tongue upward to form a tube-like structure. Traditionally, the ability to roll the tongue has been used as an example of a dominant characteristic. Sturtevant (1) described tongue rolling as an inherited human characteristic, contributing to its later use as a classical example in genetics education. Subsequent research, however, questioned whether tongue rolling is controlled by a single gene. A widow's peak is a V-shaped point in the hairline at the centre of the forehead. It has traditionally been used in elementary genetics as an example of a dominant phenotype. Hitchhiker's thumb refers to marked backward bending or hyperextension of the distal thumb joint. It has traditionally been presented as a simple inherited characteristic, although the degree of thumb flexibility is dominant. Clasp of hands refers to the natural pattern formed when a person interlocks the fingers of both hands, with one thumb/finger position consistently appearing on top. In classical human genetics demonstrations, hand-clasp pattern has been studied as a heritable human phenotypic variation. A person may naturally place the left thumb over the right thumb or the right thumb over the left thumb (dominant) when clasping the hands. Tongue folding refers to the ability to manipulate the tongue into a particular folded (Dominant) or curled configuration. It has traditionally been included among classroom examples of inherited human variation.

The comparison of self-reported and directly observed phenotypes is important because participants may not always classify their own characteristics consistently. A person may misunderstand the definition of a phenotype, use a different threshold for classifying it, or perform a task differently when directly observed. These factors can produce disagreement even when the characteristic is readily visible. Studies of human morphogenetic traits among university students, including research by Azimi-Garakani and Beardmore (2) and the study of Nigerian undergraduates published in *Scientific African* (3), demonstrate the continuing use of student populations to investigate the distribution and classification of such characteristics.

The important feature of the present study is not simply determining how common each phenotype is. Instead, it investigates whether our APAC college students can accurately identify their own phenotypic characteristics. Therefore, two observations will be obtained from every participant. They are Self-reported phenotype and Directly observed phenotype. The two observations will then be compared to determine the level of agreement and the percentage of correct self-identification.

The present study aims to assess the accuracy of self-reported identification of six classical inherited phenotypic traits among UG and PG Zoology students by comparing their self-reported responses with direct phenotypic observation. Selected six traits are Earlobe attachment, Tongue rolling, Widow's peak, Hitchhiker's thumb, Clasp of Hands and Tongue folding.

2 MATERIALS AND METHODS

Basic materials like Participant consent form, Structured questionnaire, Data collection sheet, Pen/pencil, Clipboard, Calculator/computer, Spreadsheet software such as Microsoft Excel or Google Sheets Mobile phone/camera only if participants provide specific consent for photographs were used. For individual observations, Direct observation is used for all the 6 Mendelian traits. The study population (92) will consist of undergraduate (77) and postgraduate (15) Arulmigu Palaniandavar college of Arts and Culture, Palani students. Students may be selected from different years of study to obtain a broader sample.

2.1 Data collection procedure

A two-stage procedure should be followed.

- Stage I - Self-reporting: Each participant will be given a questionnaire. The participant will independently record their answers.
- Stage II - Direct observation: After completion of the questionnaire, the investigator will independently assess each trait. The investigator's observations should be recorded separately so that the original self-reported response does not influence the observation.

2.2 Assessment of the six traits

The following table showed that the assessment study for Self-Reporting and Direct Observation.

Table 1: Assessment Questions and Observational Criteria for Human Phenotypic Traits

S.N	Mendelian Traits	Question	Dominant	Recessive	Investigator Observation
1.	Earlobe Attachment	How would you describe your earlobe?	Free	Attached	Free/ Attached
2.	Tongue rolling	Can you roll the sides of your tongue upward to form a tube-like shape?	Rolling	Non-Rolling	Able to roll/ Unable to roll
3.	Widow's peak	Do you have a noticeable V-shaped point at the centre of your frontal hairline?	Present	Absent	Present/ Absent
4.	Hitchhiker's thumb	Can your thumb bend noticeably backward at the thumb joint?	Straight	Bend	Marked backward flexibility/ No marked backward flexibility
5.	Clasping of Hands	Clasp your hands naturally by interlocking your fingers, without thinking about which thumb should be on top	Right over Left	Left over Right	Record the Position of Upper thumb
6.	Tongue folding	Can you fold or curl your tongue into the specified shape shown by the investigator?	Able to Fold	Not able to Fold	Able/ Unable

Possible statistical tests used are Percentage agreement, Correct identification percentage, Incorrect identification percentage and Cohen's kappa (κ) to quantify agreement between categorical self-report and observed phenotype. Chi-Square goodness of fit test to check whether the observed patterns occur equally from the direct observation counts.

Reliable assessment requires statistical methods that distinguish overall agreement from agreement beyond chance. Cohen (4) introduced Cohen's kappa as a coefficient for measuring agreement between two classifications while accounting for chance agreement. Landis and Koch (5) subsequently discussed methods for evaluating observer agreement with categorical data. These methods are relevant to the present study because self-reported and directly observed phenotypes form paired categorical observations for each participant.

McNemar (6) developed a method for evaluating differences between correlated proportions. This approach is particularly relevant when the same participants provide both self-reported and directly observed responses. Whereas Cohen's kappa evaluates agreement beyond chance, McNemar's test assesses whether discordant responses occur in different directions. Using these methods together can provide a more complete understanding of reporting accuracy and systematic discrepancies.

2.3 Statistical Analysis

The collected data were statistically analysed using percentage analysis, the chi-square (χ^2) test, and Cohen's kappa coefficient.

3 RESULTS AND DISCUSSION

The present study builds on these areas of research by comparing self-reported and directly observed responses for six classical human phenotypic traits among I M.Sc, II M.Sc, I B.Sc, II B.Sc, and III B.Sc Zoology students. Reporting percentages were used to describe the patterns of agreement, Cohen's kappa was used to evaluate agreement beyond chance, and chi-square-based analysis was considered for examining response differences. The study provides an educational assessment of how consistently students identify and report observable human phenotypes. It also emphasizes the importance of standardized phenotype definitions and careful interpretation of classical human genetics examples.

Table 2: Percentage of Self-Reported (incorrectly) and Direct observed (correctly) Mendelian traits among UG and PG Zoology students

CLASS	MENDELIAN TRAITS											
	Ear Lobe		Tongue Rolling		Widow's Peak		Hitchhiker's Thumb		Clasping of Hands		Tongue Folding	
	SR %	DO%	SR %	DO%	SR %	DO%	SR %	DO%	SR %	DO%	SR %	DO%
I M.Sc	42.9	57.1	14.3	85.7	42.9	57.1	28.6	71.4	28.6	71.4	57.1	42.9
II M.Sc	25.0	75.0	50.0	50.0	0	100	37.5	62.5	25.0	75.0	25.0	75.0
I B.Sc	13.8	86.2	48.3	51.7	10.3	89.7	37.9	62.1	03.4	96.6	72.4	27.6
II B.Sc	41.4	58.6	37.9	62.1	24.1	75.9	27.6	72.4	20.7	79.3	34.5	65.5
III B.Sc	15.8	84.2	36.8	63.2	42.1	57.9	26.3	73.7	10.5	89.5	63.2	36.8

Foot Note: SR – Self Reporting; DO – Direct Observation.

The present study was undertaken to evaluate the accuracy of self-reported (SR) and directly observed (DO) classical human phenotypic traits among undergraduate and postgraduate Zoology students. Six traits were investigated: earlobe attachment, tongue rolling, widow's peak, hitchhiker's thumb, clasping of hands and tongue folding. The results revealed variation in the percentages recorded by self-report and direct observation across the five academic classes: I B.Sc., II B.Sc., III B.Sc., I M.Sc. and II M.Sc. Fleiss *et al* (7) and Agresti, (8) analysed that the statistical methods for rates and proportions and a standard reference for categorical data analysis and proportions.

Earlobe attachment: In I B.Sc., the self-reported percentage was 13.8%, whereas direct observation recorded 86.2%. In II B.Sc., the corresponding values were 41.4% and 58.6%. III B.Sc. showed 15.8% and 84.2%, while I M.Sc. recorded 42.9% and 57.1%. In II M.Sc., the values were 25.0% and 75.0%, respectively. These differences indicate that the recorded classifications varied considerably across classes. However, the percentages alone cannot establish whether students incorrectly identified their own traits; individual paired responses are required to determine actual agreement. Shaffer *et al.* (9) observed that the Multiethnic GWAS reveals polygenic architecture of earlobe attachment and they revealed that earlobe attachment is the important evidence for polygenic inheritance.

Tongue rolling: The SR and DO percentages also varied across the classes. I B.Sc. recorded 48.3% SR and 51.7% DO, II B.Sc. recorded 37.9% and 62.1%, III B.Sc. recorded 36.8% and 63.2%, I M.Sc. recorded 14.3% and 85.7%, and II M.Sc. recorded 50.0% and 50.0%. The differences suggest variation in the recorded frequencies, but they do not independently demonstrate the accuracy of self-identification. Standardized instructions and consistent observation are important when assessing this trait.

Widow's peak: The SR and DO percentages were 10.3% and 89.7% in I B.Sc., 24.1% and 75.9% in II B.Sc., 42.1% and 57.9% in III B.Sc., and 42.9% and 57.1% in I M.Sc. In II M.Sc., the reported values were 0% and 100%. This variation may reflect differences in the classification of the hairline, including uncertainty in identifying a widow's peak. Hair styling, hairline shape and observer judgment may affect classification. A clear operational definition should therefore be used.

Hitchhiker's thumb: I B.Sc. recorded 37.9% SR and 62.1% DO, while II B.Sc. recorded 27.6% and 72.4%. III B.Sc. showed 26.3% and 73.7%, I M.Sc. showed 28.6% and 71.4%, and II M.Sc. showed 37.5% and 62.5%. These results demonstrate variation in the recorded categories across classes. Because thumb flexibility exists along a continuum, a standardized method of observation is essential to reduce subjective classification.

Clasping of hands: This trait showed 3.4% SR and 96.6% DO in I B.Sc., 20.7% and 79.3% in II B.Sc., 10.5% and 89.5% in III B.Sc., 28.6% and 71.4% in I M.Sc., and 25.0% and 75.0% in II M.Sc. The high DO percentages in several classes indicate

that one observed hand-clasping pattern was more frequent than the other recorded category. The percentages do not, however, establish that direct observation was more accurate than self-report. Students should be asked to clasp their hands naturally without being shown either position beforehand.

Tongue folding: I B.Sc. recorded 72.4% SR and 27.6% DO, II B.Sc. recorded 34.5% and 65.5%, III B.Sc. recorded 63.2% and 36.8%, I M.Sc. recorded 57.1% and 42.9%, and II M.Sc. recorded 25.0% and 75.0%. This trait showed considerable variation across the five classes. Differences may be associated with the way students interpret the task or perform the movement. Therefore, identical instructions should be provided to all participants. Ahmed *et al.* (10) published that the tongue rolling, tongue folding, hand clasping, hitchhiker's thumb, and other visible characteristics are the inheritance patterns and potential associations of selected human morphogenetic traits.

Consequently, the study is best interpreted as an investigation of students' recognition of classical human phenotypic variations rather than a direct test of Mendelian inheritance. Overall, the findings highlight the importance of distinguishing self-reported classifications from directly observed phenotypes in genetics education.

Table 3: Assess the accuracy of Mendelian traits among UG and PG Zoology students

CLASS	MENDELIAN TRAITS											
	Ear Lobe		Tongue Rolling		Widow's Peak		Hitchhiker's Thumb		Clasping of Hands		Tongue Folding	
	AR%	PT	AR%	PT	AR%	PT	AR%	PT	AR%	PT	AR%	PT
I M.Sc	57.1	LA	85.7	R	57.1	LA	71.4	A	71.4	A	42.9	W
II M.Sc	75.0	A	50.0	LA	100	P	62.5	LA	75.0	A	75.0	A
I B.Sc	86.2	R	51.7	LA	89.7	R	62.1	LA	96.6	P	27.6	W
II B.Sc	58.6	LA	62.1	LA	75.9	A	72.4	A	79.3	A	65.5	LA
III B.Sc	84.2	R	63.2	LA	57.9	LA	73.7	A	89.5	R	36.8	W

Foot note: AR- Accurate Reporting (DO); PT- Performance Tier (Below 50% = Worse [W]; 50-69% = Low accuracy [LA]; 70-79% = Acceptable [A]; 80-89% = Reliable [R]; 90-94% = Highly accurate [HA]; 95-100% = Perfect [P]).

Table 3 presents the accuracy of reporting six classical human phenotypic traits-earlobe attachment, tongue rolling, widow's peak, hitchhiker's thumb, clasping of hands, and tongue folding-among I B.Sc, II B.Sc, III B.Sc, I M.Sc, and II M.Sc Zoology students. The accuracy percentages were classified into performance tiers ranging from worse to perfect. The results reveal considerable variation in reporting accuracy among traits and academic classes. McHugh (11) and Sim *et al.* (12) were over-viewed the Interpreter reliability of kappa interpretation. Kappa statistic is useful for explaining sample-size considerations and interpreting agreement.

3.1 Earlobe Attachment

Earlobe attachment showed relatively high accuracy in several classes. I B.Sc students recorded 86.2%, classified as reliable, followed by III B.Sc students at 84.2%, also classified as reliable. II M.Sc students achieved 75.0%, indicating acceptable performance, whereas II B.Sc students recorded 58.6%, indicating low accuracy. I M.Sc students showed 57.1%, also classified as low accuracy. These findings suggest that earlobe attachment was reported accurately by many undergraduate students, particularly those in I and III B.Sc. However, the variation between classes indicates that recognition of this trait was not consistent across all groups. Shaffer *et al.* (9) conducted a multi ethnic genome-wide association study involving 74,660 individuals and identified multiple genetic loci associated with earlobe attachment. Their findings strongly supported a polygenic basis for this characteristic.

3.2 Tongue Rolling

Tongue rolling showed variable levels of accuracy across the five classes. I M.Sc students achieved the highest accuracy at 85.7%, classified as reliable. III B.Sc students recorded 63.2%, II B.Sc students 62.1%, and I B.Sc students 51.7%; all three were classified as low accuracy. II M.Sc students recorded the lowest value, 50.0%, also indicating low accuracy. Azimi-Garakani and Beardmore (2) investigated tongue-rolling phenotypes among university undergraduates and reported differences between student groups. These studies demonstrate the need to distinguish observable phenotype frequencies from proof of a specific mode of inheritance. These results suggest that the ability to identify or report tongue-rolling status was inconsistent among students. The findings also highlight the importance of distinguishing between a student's self-reported ability and the ability directly observed during assessment.

3.3 Widow's Peak

Accuracy in reporting widow's peak varied considerably. II M.Sc students achieved 100%, classified as perfect, while I B.Sc students recorded 89.7%, classified as reliable. II B.Sc students achieved 75.9%, indicating acceptable performance. In contrast, I M.Sc students recorded 57.1% and III B.Sc students 57.9%, both indicating low accuracy. The perfect score in II M.Sc suggests complete agreement under the study's assessment method for this class, whereas the lower values in I M.Sc and III B.Sc indicate that recognition or reporting of the trait may require further clarification. However, the small sample size in the postgraduate classes should be considered when interpreting these differences.

3.4 Hitchhiker's Thumb

The accuracy of reporting hitchhiker's thumb ranged from 62.1% to 73.7%. III B.Sc students recorded the highest value at 73.7%, followed by II B.Sc at 72.4%; both were classified as acceptable. I M.Sc students achieved 71.4%, also indicating acceptable performance. II M.Sc students recorded 62.5%, while I B.Sc students recorded 62.1%; both were classified as low accuracy according to the study's stated performance tiers. However, research discussed by McDonald (13) in *Myths of Human Genetics* indicates that thumb flexibility varies along a continuum and does not fit a straightforward single-gene model. The study of these traits is therefore useful for demonstrating the distinction between simplified classroom models and the complexity of human phenotypic variation. Overall, the results indicate moderate but inconsistent accuracy in reporting this trait. Clear and standardized criteria for distinguishing thumb flexibility and shape may help improve consistency during assessment.

3.5 Clasping of Hands

Clasping of hands showed high accuracy in some undergraduate classes. I B.Sc students achieved 96.6%, classified as perfect, and III B.Sc students recorded 89.5%, classified as reliable. II B.Sc and II M.Sc students both performed at an acceptable level, with 79.3% and 75.0%, respectively. I M.Sc students recorded 71.4%, also indicating acceptable performance. These findings suggest that clasping of hands was one of the more consistently reported traits, particularly among I B.Sc students. The results may reflect the ease with which this trait can be demonstrated and observed using a consistent procedure. Martin (14), in a study of twins, reported no convincing evidence for a simple genetic basis for tongue rolling or hand clasping.

3.6 Tongue Folding

Tongue folding showed the lowest and most variable accuracy among the six traits. II M.Sc students achieved 75.0%, classified as acceptable, followed by II B.Sc students at 65.5% and I M.Sc students at 42.9%. III B.Sc students recorded 36.8%, while I B.Sc students recorded 27.6%. Under the study's classification, I M.Sc, III B.Sc, and I B.Sc were in the worse category because their values were below 50%. The low values in these classes suggest that tongue folding may be difficult for students to identify or report consistently. A practical demonstration with clear instructions may improve recognition of the observed phenotype. Nevertheless, the results should not be interpreted as evidence of poor genetic knowledge alone, because the task may also involve differences in understanding the trait or performing the action.

Table 4: Analysis of Chi-Square test for Mendelian traits among UG and PG Zoology students

CLASS	MENDELIAN TRAITS											
	Ear Lobe		Tongue Rolling		Widow's Peak		Hitchhiker's Thumb		Clasping of Hands		Tongue Folding	
	χ^2 Value	Sig	χ^2 Value	Sig	χ^2 Value	Sig	χ^2 Value	Sig	χ^2 Value	Sig	χ^2 Value	Sig
I M.Sc	0	NS	1.181	NS	0	NS	1.470	NS	0.365	NS	0	NS
II M.Sc	0.320	S	0	S	3.556	S	0	S	0.5	S	0.320	S
I B.Sc	0.250	NS	0.643	NS	0	NS	9.091	S	0	NS	19.048	S
II B.Sc	0.167	NS	0.112	NS	4.415	S	5.623	S	8.650	S	1.001	NS
III B.Sc	0	NS	0.571	NS	0.125	NS	0.800	NS	0.500	NS	0.081	NS

Foot note: Sig- Significance, S- Significant, NS- Not Significant.

Among the six traits, clasping of hands and earlobe attachment generally showed higher accuracy, particularly in I B.Sc and III B.Sc students. Widow's peak also showed high accuracy in I B.Sc and II M.Sc. In contrast, tongue folding had the lowest accuracy in I B.Sc and III B.Sc, indicating that this trait may require greater attention during practical teaching.

Tongue rolling showed inconsistent performance across the classes, while hitchhiker's thumb generally produced low-to-acceptable accuracy.

No single academic level performed consistently best across all six traits. For example, I B.Sc students performed particularly well in clasping of hands and widow's peak but poorly in tongue folding. Similarly, II M.Sc students achieved perfect accuracy for widow's peak but only low accuracy for tongue rolling. These differences suggest that reporting accuracy is trait-specific rather than being determined solely by the academic level of the students.

Table 4 presents the chi-square analysis of six classical human phenotypic traits among undergraduate and postgraduate Zoology students. The traits investigated were earlobe attachment, tongue rolling, widow's peak, hitchhiker's thumb, clasping of hands, and tongue folding. The analysis aimed to identify differences between self-reported and directly observed phenotypes across five academic classes. The results showed variation among traits and classes, suggesting that discrepancies between reported and observed responses were not uniform.

3.7 Earlobe Attachment

Earlobe attachment showed no statistically significant difference in I M.Sc, I B.Sc, II B.Sc, or III B.Sc according to the significance labels in the table. The reported chi-square values for these classes were 0, 0.250, 0.167, and 0, respectively. II M.Sc recorded a chi-square value of 0.320, although its significance label requires verification. Overall, the results suggest that the differences between self-reported and directly observed earlobe attachment may not be statistically significant in most classes. However, the interpretation should be confirmed using the original paired data and the appropriate test.

3.8 Tongue Rolling

Tongue rolling showed variation in the reported chi-square values across the academic classes. I M.Sc recorded a value of 1.181, while I B.Sc, II B.Sc, and III B.Sc recorded values of 0.643, 0.112, and 0.571, respectively. These were labelled not significant in the table. II M.Sc recorded a value of 0, but it was labelled significant. This label is inconsistent with the usual interpretation of a chi-square statistic and should be checked. The findings suggest that self-reported tongue-rolling ability may differ from observed performance in some students, but reliable conclusions about statistical significance require verification of the calculations.

3.9 Widow's Peak

Widow's peak showed different chi-square values among the classes. I M.Sc and I B.Sc recorded values of 0, and III B.Sc recorded 0.125; these were labelled not significant. II B.Sc recorded a value of 4.415, labelled significant, while II M.Sc recorded 3.556, also labelled significant. These results suggest possible differences between self-reported and directly observed widow's peak phenotypes in some classes. Nevertheless, the significance labels and calculations should be checked against the original response counts before drawing firm conclusions.

3.10 Hitchhiker's Thumb

Hitchhiker's thumb showed relatively high reported chi-square values in I B.Sc and II B.Sc, at 9.091 and 5.623, respectively. Both were labelled significant in the table. In contrast, I M.Sc and III B.Sc recorded values of 1.470 and 0.800, respectively, and were labelled not significant. II M.Sc recorded a value of 0, although it was labelled significant. The pattern suggests that discrepancies between self-reported and observed thumb shape may be more pronounced in some undergraduate classes. However, the significance labels for II M.Sc and the underlying calculations require verification.

3.11 Clasping of Hands

Clasping of hands produced chi-square values of 0.365 in I M.Sc, 0 in I B.Sc, 8.650 in II B.Sc, and 0.500 in III B.Sc. The table labels the first, second, and fourth values as not significant, while the II B.Sc value is labelled significant. II M.Sc recorded a value of 0.5 but was labelled significant, which should be checked. The results as displayed suggest a possible difference in reporting consistency between classes, particularly in II B.Sc. However, statistical significance cannot be established confidently until the calculations and labels are corrected.

3.12 Tongue Folding

Tongue folding showed high chi-square values in I B.Sc and II M.Sc, reported as 19.048 and 0.320, respectively. Both were labelled significant in the table. II B.Sc recorded 1.001 and III B.Sc recorded 0.081, both labelled not significant. I M.Sc recorded 0 and was labelled not significant. The high value reported for I B.Sc suggests a marked difference between self-reported and directly observed tongue-folding responses. However, the II M.Sc result requires particular attention because a chi-square value of 0.320 would not ordinarily be statistically significant with one degree of freedom.

The table indicates variation in the reported chi-square statistics among the six traits and five academic classes. Hitchhiker's thumb and tongue folding showed comparatively high values in I B.Sc, while widow's peak, hitchhiker's thumb, and clasping of hands showed higher values in II B.Sc. In contrast, several traits in III B.Sc had low reported chi-square values and were labelled not significant.

These patterns may indicate that the consistency between self-reported and directly observed phenotypes differs by trait and academic class. However, the results do not establish that one academic level has greater genetic knowledge than another. Differences may also arise from how students understand the trait, how the observation is performed, and how the phenotype categories are defined.

Table 5: Assess the Cohen's Kappa analysis for Mendelian traits among UG and PG Zoology students

CLASS	MENDELIAN TRAITS											
	Ear Lobe		Tongue Rolling		Widow's Peak		Hitchhiker's Thumb		Clasping of Hands		Tongue Folding	
	Ck	SA	Ck	SA	Ck	SA	Ck	SA	Ck	SA	Ck	SA
I M.Sc	0.087	S	0.696	SB	0.087	S	0.720	SB	0.462	M	0.125	S
II M.Sc	0.467	M	0	PR	1.000	P	-0.200	LC	0.500	M	0.467	M
I B.Sc	0.521	M	-0.180	LC	0.664	SB	0.326	F	0	H	0.047	S
II B.Sc	0.147	S	0.140	S	0.438	M	0.470	M	0.593	SB	0.190	S
III B.Sc	0.617	SB	0.131	S	0.106	S	0.296	F	0.791	SB	0.081	S

Foot note: Ck- Cohen's kappa, SA- Strength of Agreement. (P- Perfect, H- High, M- Moderate, SB- Substantial, S- Slight, PR- Poor, F- Fair, LC-Less than chance).

Table 5 presents the results of Cohen's kappa analysis used to assess the agreement between self-reported and directly observed phenotypes for six classical human traits among undergraduate and postgraduate Zoology students. The traits examined were earlobe attachment, tongue rolling, widow's peak, hitchhiker's thumb, clasping of hands, and tongue folding. The kappa values revealed considerable variation in agreement across traits and academic classes, indicating that the accuracy of self-reporting was not consistent among students.

3.13 Earlobe Attachment

The agreement for earlobe attachment varied across the five classes. III B.Sc students showed the highest agreement, with a kappa value of 0.617, indicating substantial agreement. I B.Sc students recorded a kappa value of 0.521, indicating moderate agreement, followed by II M.Sc students ($\kappa = 0.467$), also showing moderate agreement. II B.Sc students exhibited slight agreement ($\kappa = 0.147$), while I M.Sc students recorded the lowest value ($\kappa = 0.087$), indicating slight agreement. These results suggest that III B.Sc students demonstrated comparatively better consistency between self-reported and observed earlobe phenotypes. However, the differences between classes indicate that the recognition and reporting of this trait were not uniform.

3.14 Tongue Rolling

Tongue rolling showed considerable variation in agreement between self-reporting and direct observation. I M.Sc students achieved a kappa value of 0.696, indicating substantial agreement. II B.Sc and III B.Sc students showed slight agreement, with values of 0.140 and 0.131, respectively. I B.Sc students recorded a negative kappa value of -0.180, indicating agreement below that expected by chance. II M.Sc students had a kappa value of 0, suggesting no agreement beyond chance. These findings indicate that students' self-reported ability to roll the tongue did not consistently match their observed performance. The results highlight the need for clear instructions and standardized practical assessment when evaluating this trait.

3.15 Widow's Peak

Widow's peak demonstrated the greatest variation in agreement across academic classes. II M.Sc students achieved a kappa value of 1.000, indicating perfect agreement between self-reporting and direct observation in the recorded sample. I B.Sc students showed substantial agreement ($\kappa = 0.664$), while II B.Sc students demonstrated moderate agreement ($\kappa = 0.438$). In contrast, III B.Sc students recorded slight agreement ($\kappa = 0.106$), and I M.Sc students also showed slight agreement ($\kappa = 0.087$). The perfect agreement observed among II M.Sc students indicates complete consistency in their recorded responses, whereas the lower values in I M.Sc and III B.Sc suggest discrepancies in identifying or reporting this phenotype. Nevertheless, the sample sizes, particularly in the postgraduate classes, should be considered when interpreting these differences.

3.16 Hitchhiker's Thumb

The kappa values for hitchhiker's thumb ranged from negative agreement to substantial agreement. I M.Sc students showed substantial agreement, with a kappa value of 0.720. II B.Sc students recorded moderate agreement ($\kappa = 0.470$), while I B.Sc and III B.Sc students demonstrated fair agreement, with values of 0.326 and 0.296, respectively. II M.Sc students recorded a negative kappa value of -0.200 , indicating agreement below that expected by chance. These results suggest that students differed in their ability to identify and report thumb shape consistently. Standardized criteria for distinguishing bent and straight thumb positions may improve the consistency of observations.

3.17 Clasping of Hands

Clasping of hands showed comparatively strong agreement in several classes. III B.Sc students achieved the highest kappa value (0.791), indicating substantial agreement. II B.Sc students also demonstrated substantial agreement ($\kappa = 0.593$). I B.Sc students recorded a kappa value of 0, indicating no agreement beyond chance, while I M.Sc and II M.Sc students showed moderate agreement, with values of 0.462 and 0.500, respectively. These results suggest that clasping of hands was reported more consistently in III B.Sc and II B.Sc than in the other classes. Differences in the distribution of the two hand-clasping patterns may influence kappa values, so the results should be interpreted alongside the raw agreement percentages and the underlying response counts.

3.18 Tongue Folding

Tongue folding showed generally low agreement in several classes. II M.Sc students recorded the highest kappa value for this trait (0.467), indicating moderate agreement. II B.Sc students showed slight agreement ($\kappa = 0.190$), while I M.Sc, I B.Sc, and III B.Sc students recorded kappa values of 0.125, 0.047, and 0.081, respectively, all indicating slight agreement. The particularly low values among I B.Sc and III B.Sc students suggest that their self-reported tongue-folding ability often differed from direct observation. This may reflect difficulties in identifying the phenotype, differences in how students interpreted the task, or inconsistencies in performing the action during assessment. Clear instructions and a standardized observation procedure may help improve the reliability of future assessments.

Across the six traits, no single academic class showed consistently high agreement for every phenotype. III B.Sc students demonstrated substantial agreement for earlobe attachment and clasping of hands, but only slight agreement for tongue rolling, widow's peak, and tongue folding. I M.Sc students showed substantial agreement for tongue rolling and hitchhiker's thumb but slight agreement for earlobe attachment, widow's peak, and tongue folding. II M.Sc students achieved perfect agreement for widow's peak but had no agreement beyond chance for tongue rolling and negative agreement for hitchhiker's thumb.

I B.Sc students demonstrated substantial agreement for widow's peak, moderate agreement for earlobe attachment, and fair agreement for hitchhiker's thumb, but their agreement was poor or absent beyond chance for other traits. II B.Sc students showed substantial agreement for clasping of hands and moderate agreement for widow's peak and hitchhiker's thumb, while tongue folding and several other traits showed weaker agreement.

These findings indicate that agreement between self-reported and directly observed phenotypes is trait-specific and cannot be explained by academic level alone. Practical experience may help students recognize certain traits, but further investigation would be required to establish whether academic progression is associated with improved reporting accuracy.

4 CONCLUSION

The present study demonstrated considerable variation in the accuracy of reporting six classical human phenotypic traits among undergraduate and postgraduate Zoology students. The comparison between self-reported and directly observed responses revealed that some traits were identified relatively consistently, whereas others showed substantial discrepancies.

Clasping of hands and earlobe attachment generally showed higher reporting accuracy in several classes. Widow's peak also demonstrated high accuracy in selected groups, including 100% in II M.Sc. However, tongue folding emerged as a major area of difficulty, particularly among I B.Sc and III B.Sc students, where the reporting accuracy was 27.6% and 36.8%, respectively. Cohen's kappa results further supported the finding of weak agreement for tongue folding in these two classes.

The study also showed that no single academic class performed consistently best for every trait. Therefore, reporting accuracy appears to depend on the particular phenotype being assessed rather than academic level alone. Clear definitions, standardized observation procedures, and practical demonstrations may help improve the consistency of phenotype identification.

The chi-square results should be interpreted cautiously until the calculations and significance labels are verified using the original paired observations. Similarly, Cohen's kappa should be considered alongside raw agreement percentages and sample sizes, particularly in the smaller postgraduate classes.

In conclusion, this study emphasizes the importance of comparing self-reported responses with direct observation in practical Zoology education. Such activities can help students recognize discrepancies in phenotype identification and develop a more systematic understanding of human phenotypic variation. Nevertheless, the six traits should be described as classical human phenotypic traits rather than assumed to be straightforward single-gene Mendelian traits. The findings assess agreement in reporting and observation; they do not independently establish students' genetic knowledge or the inheritance pattern of each trait.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Disclosure of conflict of interest

The authors declare no conflicts of interest regarding this manuscript.

Author's contribution

This work was carried out in collaboration among the authors. Author Dr. Umarani Subburam, designed the study, wrote the protocol, and wrote the first draft of the manuscript. Author Dr. Umarani Subburam, Dr. Jeyalalitha Thirupathi, Dr. Pavithran Saminathan, and Mrs. Mahimashree Arumugam, did the experimental work, performed statistical analysis and managed the analyses of the study. All authors read and approved the final manuscript.

REFERENCES

- [1] Sturtevant AH. A new inherited character in man. *Proc Natl Acad Sci U S A*. 1940; 26(2): 100-2. doi:10.1073/pnas.26.2.100.
- [2] Azimi-Garakani C, Beardmore JA. An association between tongue-rolling phenotypes and subjects of study of undergraduates. *Journal of Biosocial Science*. 1979 Apr; 11(2):193-9.
- [3] Adekoya KO, Fakorede ST, Ogunkanmi AL, Amusa OD, Sifau MO, Oboh B. Inheritance pattern and association studies of some human morphogenetic traits among Nigerian undergraduate students. *Scientific African*. 2020 Sep 1;9:e00508.
- [4] Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46. doi:10.1177/001316446002000104.
- [5] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74. doi:10.2307/2529310.
- [6] McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*. 1947; 12(2): 153-7. doi:10.1007/BF02295996.
- [7] Fleiss JL, Levin B, Paik MC. *Statistical methods for rates and proportions*. 3rd ed. Hoboken (NJ): Wiley; 2003.
- [8] Agresti A. *Categorical data analysis*. 3rd ed. Hoboken (NJ): Wiley; 2013.
- [9] Shaffer JR, Li J, Lee MK, Rosenboom J, Orlova E, Adhikari K. Multiethnic GWAS reveals polygenic architecture of earlobe attachment. *Am J Hum Genet*. 2017; 101(6):913-24. doi:10.1016/j.ajhg.2017.10.001.
- [10] Ahmed MA, Ullah MA, Farooq A, Ali Q, Abbas A, Rehman AU. Investigating the inheritance patterns and potential associations of selected human morphogenetic traits. *Bulletin of Biological and Allied Sciences Research*. 2024 Oct 2;2024(1):77-77.
- [11] McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-82.
- [12] Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85(3):257-68.
- [13] McDonald JH. *Myths of human genetics* [Internet]. Newark (DE): University of Delaware; 2011 [cited 2026 Oct 18]. Available from: <https://udel.edu/~mcdonald/>
- [14] Martin NG. No evidence for a genetic basis of tongue rolling or hand clasping. *J Hered*. 1975; 66(3):179-80. doi:10.1093/oxfordjournals.jhered.a108608.