International Journal of Research in Medical Sciences
Pathak RR. Int J Res Med Sci. 2025 Nov;13(11):4778-4784
www.msjonline.org
pISSN 2320-6071 | eISSN 2320-6012
DOI: https://dx.doi.org/10.18203/2320-6012.ijrms20253600

Original Research Article

Common parlance English in an Indian medical undergraduate institution: teaching-learning methods for improvement

Rakesh Ranjan Pathak*

Department of Pharmacology, GMERS Medical College, Morbi, Gujarat, India

Received: 12 August 2025
Revised: 02 October 2025
Accepted: 04 October 2025

***Correspondence:**
Dr. Rakesh Ranjan Pathak,
E-mail: rr_pathak@yahoo.com

ABSTRACT

Background: Author's experience/ Capstone Project (T2TM, Harvard University) approval.
Methods: Harrison's Textbook of Internal Medicine was the vocabulary source. COCA corpus was purchased for ranking 'usage frequency' of words. Background (MCQ.1) versus summative (MCQ.2) assessment has already been published – 'teaching learning methods' applied meanwhile is elaborated hereby. After MCQ.1, in sensitisation, 5 glossary options were explained/ offered. Monthly improvement of 1000 COCA ranks were aimed through 5 monthly interim tests (MCQ.T.1-5). MCQ.2 was taken after 6 months. All 7 tests were followed by feedback enucleation. Finally, dictionaries were rated by the participants through 7 questions. Data normality (Shapiro Wilk test) without outliers (Tukey's fence/ QQ plot) was affirmed. Repeated measure ANOVA was applied. After 6 months, 5000 ranks betterment was expected in MCQ.2 apropos MCQ.1 – in outcome, there was no underfitting/ overfitting (no significant bias/ variance).
Results: Sphericity (Mauchly's test) was 0.85 – so adjusted F-tests (Greenhouse Geisser [ε_{GG} = 0.8497]/ Huynh Feldt [ε_{GG} = 0.8888] correction) were used. Effect size (η^2) was 0.0035 (observed) and 0.0201 (partial). Despite monthly increase of usage frequency rank by 1000, participants' 'scores' were not significantly different in the 5 monthly MCQ.Ts. Dictionary market was not congruent with post-awareness choices.
Conclusions: By such a 6-month training, improvement of at least 5000 COCA ranks can be achieved. Dictionaries require in-depth awareness before purchase. Dictionaries D.4 and D.5 could think of lowering the prices.

Keywords: English, COCA, Usage frequency, Teaching-learning method

INTRODUCTION

Based on the Author's experience concerning English language as a medium of teaching-learning communications in the MBBS (Indian medical undergraduate) students of India, a study was conducted under Capstone Project of the author under the tutelage of Harvard University, USA – during the course called T2TM (Training to Teach in Medicine).[1]

Aim and objective

To assess the baseline vocabulary of MBBS students at the given institution.

To sensitise and guide the participants for using various glossary options (dictionaries) to improve their common parlance vocabulary.

To assess the improvement of vocabulary after 6 "monthly teaching/ learning sessions".

To assess final popular perception about glossary options (dictionaries).

METHODS

As per approval (in the Capstone Project and the Institutional Ethics Committee), the study was conducted on the 1st and 2nd year MBBS students of GMERS Medical College at Morbi, Gujarat – 363641 (India).

The worldwide acknowledged Harrison's Textbook of Internal Medicine[1] was adopted as the standardized source of basic vocabulary for medical students. COCA corpus was purchased for ranking usage frequency of the words in all the 7 tests (MCQ.1, MCQ.T.1-5 and MCQ.2).

Highest usage frequency of a word leads to rank one in this COCA corpus – thus increasing number of rankings would be indicative of rareness in common parlance communication. For example, a word at rank 1000 (viz. seat, noun) is much more common than another word at rank 10,000 (viz. Diversify, verb).

The baseline assessment test (MCQ.1) was designed with a range of rank 4000 to 50,000 (3,790 to 51,332 to be exact).[2] MCQ.1 (background analysis) was taken in March 2024 and MCQ.2 (summative) was concluded in September 2024 – in between, MCQ.T.1-5 were the "5 monthly interim tests".

The portion of 'background' versus 'post-training final outcome' assessment questionnaire (called MCQ.1 and MCQ.2 respectively) – techniques and turnouts – have already been discussed and published.[2] Hereby, the 'teaching learning methods,' employed *en passant* for improvement of the participants, would be elaborated.

After start-up sensitization, and background assessment (MCQ.1) - there was a feedback enucleation-cum-orientation class of 1 hour. Therein, the relevance of defining vocabulary was explained and blueprint of 6 month's schedule was detailed. Ensuing 5 interim formative assessments, 5 "monthly teaching/ learning sessions in the feedback form" and final summative MCQ.2 (6 months after MCQ.1) were explained. Thus, baseline MCQ.1 + interim MCQ.T.1-5 + summative MCQ.2 – each and every test was to be followed by enucleation sessions (mistakes→ reasoning and resolution).

Based on the top selling score of 3 book-stores at Ahmedabad, Gujarat, India, various 'international' glossary options (dictionaries) were explained as follows:

D.1. The Oxford Advanced Learner's Dictionary (10th edition, 2020).

D.2. The Concise Oxford English Dictionary (12th edition, 2019).

D.3. The Longman Dictionary of Contemporary English (6th edition, 2020).

D.4. The Merriam Webster's Collegiate Dictionary (11th edition, 2003).

D.5. The Collin's Cobuild Advanced Learner's Dictionary (10th edition, 2023).

The exemplary vocabulary test items in MCQ.T.1 to MCQ.T.5 were segregated in 5 sections viz. 1. processes and functions, 2. ideas and thoughts, 3. Look-alike sound-alike (LASA) words, 4. Non-medical terminology, and 5. Rare glossary items. There were 60 words in each section to begin with (overall 300 words) - from which, a total of 50 were taken in increasing order of usage frequency rank.

From 300 words chosen to start, after stratified random sampling (5 strata, each of 60 words, having nearly equal usage frequency in the COCA corpus – and successive inter-strata gap of around 1000 ranks), 50 went to the 5 interim tests.

Thus, 50 exemplary glossary items (Table 1) overall – 10 in each of the five categories, and 10 in each test too – were identified from Harrison's Textbook of Internal Medicine[1] and quantified in the terms of COCA ranking (of usage frequency).

MCQ.T.1 indicates first interim examination organised 1 month after MCQ.1 and "streak.1 (4980)" indicates that the word 'streak' is in category 1 (viz. 'processes and functions', as explained before) having the COCA rank of 4980. This was around 1000 ranks above the least rank in MCQ.1.[2]

Noticeably, all inflexional forms (for example, 'see' and 'saw' from the same verb) or homographic (for example saw as 1. a cutting instrument and as 2. the past tense of see) forms are not equally common in COCA usage frequency.

For example, 'run' as verb has COCA rank of 202 but as a noun, the rank is 1225. 'Knot' as a noun has the COCA rank of 5743 - but as a verb, the rank is 16454. While enlistment for our all 7 tests (MCQ.1, MCQ.T.1-5 and MCQ.2), these differentiations were strictly followed.

The first test (background test MCQ.1) and the final test (summative test MCQ.2) were based on each other item-wise – for each COCA rank in MCQ.1 nearly 5000 higher ranking word was there in MCQ.2. (range of 3,790-51,332 in MCQ.1 was upgraded to 8,956-55,308 in MCQ.2).[2]

Meanwhile, during training (MCQ.T.1-5), this range was broken in 5 step for 5 categories cited above (range of 51,332 – 3,790 = 47542 in MCQ.1 was divided into 5

category steps of nearly 9,500 – each step with 2 words in that category, and temporally monthly upgradation of 1000 ranks). Thus, in MCQ.T.1 (first interim test), 3790 (lowest COCA frequency of MCQ.1) + 1000 = 4790 was the target COCA rank to start in MCQ.T.1and 51,332 + 1000 = 52,332 was target rank for the last item.

As one word in MCQ.1 was beyond 60,000 ranking of COCA – in each further interim tests (MCQT.1-MCQ.T.5) one such word was included. After enucleation session of the final summative session MCQ.2, a survey of participants' opinion about glossary instruments (i.e. various dictionaries) was also done.

In category 3 (LASA words) enlisted in Table 1, c. indicates the supposed alternative, with which the test item could be confused. For example, in MCQ.T.1, category 3, tempered (COCA rank 33,100) was included because, it could be confused with tampered.

Statistical analysis

As there was only one variable (same training to all participant, without any co-variable) and samples were paired (each single participant was measured 7 times - out of which, 5 were interim analyses). ANOVA was chosen.[3]

Repeated measures ANOVA (RMA) was chosen – which is also referred to as 'within-subject ANOVA' or 'ANOVA for paired samples' because same participant was measured more than 2 times (5 times to be exact). Significance level was fixed as 0.05. Bonferroni correction was *en passant* to reduce false positives.[3]

Each of the 5 monthly interim tests (MCQ.T.1, MCQ.T.2, MCQ.T.3, MCQ.T.4, and MCQ.T.5) was containing 2 items from each of the 5 categories chosen in an increasing order of COCA usage ranking – thus each successive test was covering 10 glossary items with nearly 1000 ranks higher than the previous test. After 6 months, summative MCQ.2 was organised.

RESULTS

Tukey's fence (K=1.5) ruled out any outlier, as also evident in QQ plot in figure.1. Shapiro Wilk test affirmed normality of data distribution (α=0.001125).[3] For all the 5 outcomes of interim tests, an RMA under fixed effect model was run.

Table 1: Words used in interim tests (MCQ.T1-5) with their category and COCA rank (in parenthesis).

MCQ.T.1	MCQ.T.2	MCQ.T.3	MCQ.T.4	MCQ.T.5
Archive.1 (14471)	Recoil.1 (15418)	Intoxicate.1 (16460)	Fence.1 (17459)	Invert.1 (18698)
Feign.1 (14538)	Dislodge.1 (15615)	Sweeten.1 (16753)	Disparage.1 (17571)	Droop.1 (18712)
Aberrant.2 (23985)	Colloquial.2 (25126)	Catchment.2 (25925)	Receding.2 (26953)	Ectopic.2 (27988)
Empiricism.2 (24112)	Lubrication.2 (25212)	Expedited.2 (25988)	Ablation.2 (27207)	Flavoring.2 (28176)
Tempered.3 (33100) c. tampered	Inventory.3 (34497) c. invention	Menthol.3 (35312) c. methanol	Insidiously.3 (36522) c. inside	Lagged.3 (37096) c. legged
Splice.3 (33571) c. slice	Tampering.3 (34791) c. tempering	Ironical.3 (35776) c. iron	Tabular.3 (36757) c. tubular	Quarry.3 (37269) c. query
Dinoflagellate.4 (42466)	Distillate.4 (43806)	Enantiomer.4 (44400)	Heterogenous.4 (45705)	Buffering.4 (46364)
Colloidal.4 (42637)	Palindrome.4 (43821)	Avidity.4 (44446)	Leonine.4 (45754)	Dimorphic.4 (46734)
Gasping.5 (51976)	Likertscale.5 (52833)	Epistasis.5 (54098)	Volar.5 (54824)	Disambiguate.5 (56202)
Integument.5 (> 60,000)	Cribriform.5 (> 60,000)	Reassortment.5 (>60,000)	Psychedelic.5 (>60,000)	Terpene (>60,000)

Mauchly's test of sphericity indicated that assumption of sphericity has been violated ($p<0.001$, $W=0.537$, $\chi^2=52.432$). That's why, for repeated measure ANOVA, p was 0.1356 originally – which became p=0.1467 after Greenhouse Geisser correction ($\varepsilon_{GG}=0.8497$). As sphericity was around 0.85, Huynh Feldt correction could also be used, and then $\varepsilon_{HF}=0.8888$.[3]

The Repeated Measures ANOVA test indicated that there is a non-significant difference in the dependent variable

between the different groups. There was F (3.4, 292.28) = 1.76, p=.147, with a mean of 5.33 for MCQ.T.1, 5.31 for MCQ.T.2, 5.17 for MCQ.T.3, 5.44 for MCQ.T.4, 5.22 for MCQ.T.5.3.

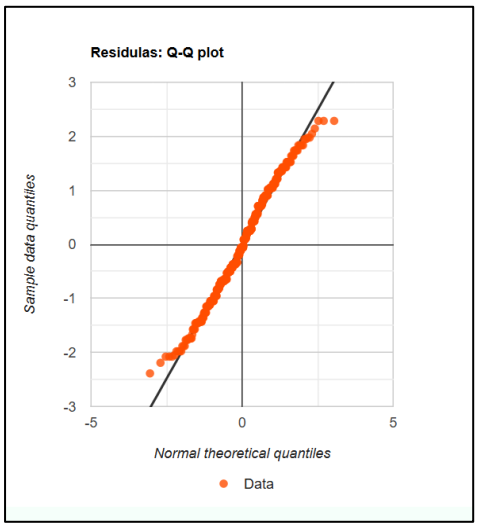

Figure 1: QQ plot (visual plot for normality).

Coming the other way round once again, observed and partial effect sizes (η2) were 0.0035 and 0.0201 respectively (indicating that differences of mean were negligibly small). As sample size of 87 was large enough, ω2 (necessary for small sample, because therein, η2 overestimates the effect size) was not obligatory to calculate the effect size.

Thus, despite monthly increase of usage frequency rank by 1000, the difference in the five monthly MCQ.Ts was not significant. The same conclusion is reaffirmed by F-statistic (=1.7643) being within the 95% region of acceptance (-∞ to 2.5302) 3. Scores of participants in MCQ.T.1- MCQ.T.5 are enlisted in Table 2.

As per 4th objective, after summative test (MCQ.2), post awareness perception of the participants about their glossary instruments (i.e. dictionaries) was collected. Based on basic 7 questions, the choice outcome from participants is summarized in Table 3.

In the starting, most (79 out of 87) participants didn't know about learner series of dictionaries explaining the words live - their uses, usage precautions, minor differences between apparent synonyms, methodical mentoring for special settings like official emails/ housing/ gardening/ garments/ market/ kitchen etc and precautions about common errors. In our glossary instruments, D.1, D.3 and D.5 were learner dictionaries.

Table 2: Test-score (out of 10) in each of 5 interim tests (MCQ.T.1- MCQ.T.5).

Interim test	Test-score (out of 10)
MCQ.T.1	7, 7, 4, 6, 5, 5, 7, 7, 3, 3, 5, 4, 5, 4, 5, 6, 6, 6, 5, 5, 3, 5, 5, 8, 5, 6, 2, 3, 4, 5, 4, 5, 6, 6, 5, 5, 6, 6, 7, 7, 2, 6, 3, 6, 6, 5, 7, 5, 5, 6, 5, 9, 5, 6, 8, 7, 7, 3, 6, 4, 4, 5, 5, 6, 5, 6, 6, 5, 6, 4, 5, 6, 8, 7, 7, 6, 7, 5, 4, 3, 4, 5, 4, 6, 6, 5, 5 (n=87)
MCQ.T.2	8, 6, 4, 7, 6, 5, 7, 6, 3, 3, 6, 3, 5, 5, 4, 5, 7, 5, 6, 4, 4, 5, 4, 8, 4, 7, 2, 4, 4, 5, 3, 6, 7, 5, 4, 4, 7, 6, 8, 8, 3, 6, 2, 6, 7, 4, 6, 5, 4, 6, 4, 10, 6, 6, 8, 8, 7, 2, 6, 5, 5, 4, 6, 6, 4, 5, 6, 4, 6, 3, 4, 7, 9, 6, 6, 5, 6, 6, 5, 2, 5, 4, 5, 5, 6, 6, 5 (n=87)
MCQ.T.3	8, 8, 3, 6, 5, 6, 6, 6, 4, 2 ,6, 4, 5, 4, 5, 6, 5, 7, 6, 5, 4, 6, 4, 9, 4, 5, 2, 2, 3, 6, 3, 4, 6, 5, 4, 4, 5, 5, 7, 6, 1, 6, 2, 6, 6, 4, 6, 5, 6, 7, 5, 9, 4, 6, 7, 6, 7, 2, 7, 5, 3, 6, 4, 5, 4, 7, 6, 4, 6, 3, 6, 5, 7, 8, 8, 5, 7, 6, 4, 2, 5, 5, 5, 5, 6, 4, 6 (n=87)
MCQ.T.4	7, 8, 3, 5, 6, 6, 7, 8, 4, 3, 4, 5, 5, 5, 6, 6, 7, 5, 4, 5, 2, 6, 4, 9, 5, 5, 3, 2, 4, 6, 4, 6, 7, 6, 6, 6, 5, 7, 8, 6, 1, 7, 4, 6, 6, 4, 8, 5, 5, 7, 4, 9, 5, 6, 7, 8, 7, 3, 5, 5, 5, 5, 6, 5, 4, 7, 6, 4, 7, 3, 6, 7, 8, 7, 7, 5, 6, 4, 5, 2, 5, 4, 3, 5, 7, 6, 6 (n=87)
MCQ.T.5	8, 8, 5, 6, 4, 5, 7, 6, 2, 3, 5, 4, 6, 4, 4, 5, 7, 6, 6, 4, 3, 5, 6, 7, 5, 6, 3, 3, 4, 5, 3, 6, 6, 7, 5, 5, 6, 7, 6, 7, 1, 7, 3, 5, 6, 4, 6, 6, 5, 6, 5, 8, 4, 5, 7, 8, 6, 2, 7, 3, 3, 5, 6, 5, 6, 5, 6, 5, 6, 4, 5, 5, 8, 8, 8, 5, 7, 4, 4, 3, 5, 4, 3, 7, 5, 4, 4 (n=87)

Table 3: Participants' opinion about vocabulary tools (Dictionaries).

Seven opinion questions	Cumulative score of various dictionaries (as per serial number in the text)				
	D.1	**D.2**	**D.3**	**D.4**	**D.5**
Best for scientific terminology	2	12	19	53	1
Best for definitions	2	31	4	49	1
Best for rapid reading	2	2	45	3	35
Best for written English	36	2	18	4	27
Best for explaining actions and ideas	27	1	8	8	43
Best value for money	42	6	26	7	6
Best one as free mobile app	23	2	7	19	36

Secondly, only one of the participants was versed about defining vocabulary – the basic 2, 3, or by the maximum 5 thousand words, most commonly used in communication, and used by the dictionaries to explain (definiens) the other millions of words in entries (definiendum). During the orientation, both these concepts were inculcated in all the participants.

DISCUSSION

Statistically, for repeated measure ANOVA, all the three conditional assumptions were adjusted. Firstly, the data set was normally distributed without outliers (Shapiro wick test as well as QQ plot).

Moreover, as there was a significant deviation from sphericity (Mauchley's test) – Fischer's exact test was avoided and Greenhouse-Geisser correction and Huynd Feldt correction were used. And finally, as each and every participant was performing separately and simultaneously – no contamination of response was possible and independence of data within the group could be safely assumed.

When all the formative assessments (MCQ.T.1-5) turned out to be satisfactory, final summative test (MCQ.2) was aimed to assess improvement of at least 5000 COCA ranks (calculated 6000 rank improvement in 6 months was not targeted in the summative test because of a thumb rule to adjust overfitting of data, if any).[4,5]

Coming the other way round, in psychology, 'law of diminishing return' explains that despite constant inputs, learning progression is not rectilinear – but a hyperbolic concavity instead (after a certain experience, learning doesn't seem to improve as fast). And this is also proven in medical education.[4,5] Thus a slightly lower target of improvement of 5000 ranks in 6 months was justifiable.

Instead of improvement of 6000 ranks, only 5000 ranks betterment was expected in the summative test (MCQ.2) - but due to this slightly lower targeting, there was no significant variance (or bias) leading to overfitting (or underfitting).[2]

Coming to the linguistic part, historically, Longman's Dictionary of Contemporary English started with a defining vocabulary of 2000 – using only these 2000 words, all other words in the dictionary were defined. For better explanation, Oxford Advanced Learner's Dictionary used 3000 words in their defining vocabulary – which extended to 5000 for academic circle.

As all Indian medical undergraduates and their teachers use English as the medium of communication, these words in defining vocabulary were supposed to be known to one and all. This concept was imparted to all the participants of this study during orientation session.

Medical terminology was intentionally omitted, as the target of the study was 'common parlance' English. All the 5 dictionaries cited above were offered as problem solving instruments (left to the participant's discretion), and a monthly improvement of 1000 COCA ranks was aimed – assessed by 5 monthly interim tests (MCQ.T.1-5).

Before MCQ.T.4, participants did enjoy a full month of vacation, during which, they didn't have any regular course burden. As they could concentrate better on this co-curricular activity, average 473/87 = 5. 4368 was the best.

To the contrary, just after MCQ.T.3, the participants had semester end exams – and that's why the result of 3rd episode was minimal (450/87 = 5.1724). Thus, when we visually analyse the arithmetic mean of outcome of 5 interim tests, slightly worse performance in 3rd test and better performance in 4th tests are explicable.

Despite all this seemingly ups-and-downs in outcome, the performance of the participants overall was not "statistically significantly different" across the continuous growth – which indicated a satisfactory completion of aim and objective of the study.

Concerning table.3, the outcome was unexpected. As most of the participants were formerly unacquainted about the glossary options – after thorough awareness, their choices, with and without price constraints (question 6 and 7 respectively), were surprisingly different from that reported by the current sale trend of dictionaries in the book market.

By the end, the participants were well-versed with various dictionary options and defining vocabulary (hence comfortable with international dictionaries) had changed their choices - which would change their personal use choice as well as change the market sale pattern of dictionaries in future. Otherwise, heretofore, dictionary purchasers are not well aware of the content utility and purchases are mostly based on brand popularity-cum-cost criteria.

On personal interaction with the participants, it was found that concerning question.1, students did feel that D.4 was the best – but availability of maximum words (technical too) at cheaper price in D.3, with 'colloquial explanation', was also a concern among 'easy going' users. Concerning Q.2, other terminologies (like legal, economic etc.) too came into play, and precision of D.2 and D.4 were unmatched.

Concerning Q.3, for rapid reading through magazines, novels, newspapers etc., widest coverage with workably good explanation in D.3 was favoured. D.5 could be the best – but for its bulk and price! Concerning Q.4, grammatical improvement was the best in inducing confidence of correctness in D.1, but due to lucidity and ease of understanding, D.5 was also a 'close second' option.

Concerning Q.5, obviously lucidity of explanation was the first criteria for the participant (D.5 scoring best) but explanation by contrast, and "king's English" level standard of D.1 was a nearby runner up for 'academic upper class'.

Concerning Q.6 and 7, the final choice was obvious – if free, the choice was for lucidity and ease of understanding (D.5); while if paid, neither D.4 nor D.5 fared well, despite their best total score.

Limitations

The COCA corpus has differentiated a word by parts of speech (noun/ verb/ adjective etc) and assigned different usage frequency – but homographs are not differentiated as per context. This is a widespread limitation of AI based Boolean search.

For example, knot in the sense of 1. string interlacement 2. Unit to measure the speed of a ship - can't be differentiated on AI/ Boolean based digital searches and such secondary segregation in usage frequency in a specific contextual sense was not feasible in this research too.

2. In all the MCQ tests, serial number of answer options were randomised - so that no a) b) c) d) pattern could be traced for hit and trial attempts. But still, as there was no negative marking, trend for tentative answering couldn't be controlled – there is no way to ascertain/ deny that some answers by the participants were correct by chance

CONCLUSION

By such a 6-month training of English vocabulary – sensitization/ glossary orientation/ baseline MCQ.1, 5 monthly MCQs, and summative MCQ.2, and each test followed by enucleation sessions – improvement of vocabulary of Indian medical undergraduates by at least 5000 COCA ranks can be achieved.

While sale of English dictionaries in the Indian book market is currently decided by pricing/ public opinion, purchase of dictionary should ideally be based on personal utilisation skill and thorough exposure to the options. D.4/ D.5 was not preferred for purchase because of exorbitant (Indian scale) prices – they can think of special/low priced Indian or South Asian editions.

ACKNOWLEDGEMENTS

This work was originally proposed and accepted as author's Capstone Project during T2TM (Training to Teach in Medicine) course from Harvard University, USA (year 2024). Although university authorities have conferred the copyright to the author – for the meticulous word-to-word verification and upgrade by the veterans in methodology, and above all, for the fee waiver for the full course by the Harvard University, the author hereby expresses a heartfelt deep gratitude.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. Kasper D, Fauci A, Hauser S, Longo D, Jameson J, Loscalzo J. Harrison's principles of internal medicine. 21st edn. McGraw Hill; 2022
2. Pathak RR. Common parlance English in an Indian medical undergraduate institution: current scenario and ranking the improvement. Int J Res Med Sci. 2025;13:2054-8.
3. Repeated Measures ANOVA Calculator Available at: https://www.statskingdom.com/repeated-anova-calculator.html. Accessed on 15 April 2024.
4. Thompson BM, Rogers JC. Exploring the learning curve in medical educat.ion: using self-assessment as a measure of learning. Academic Medicine. 25008;83(10):S86-S88.
5. Yeolekar M, Yeolekar A. Learning curve in medical education: Revisited. Somaiya Medical Journal. 2015;2(1):40-4.

Cite this article as: Pathak RR. Common parlance English in an Indian medical undergraduate institution: teaching-learning methods for improvement. Int J Res Med Sci 2025;13:4778-84.