Developing Script Concordance Test Items in Otolaryngology to Improve Clinical Reasoning Skills: Validation using Consensus Analysis and Psychometrics

Abstract

Background: Script concordance testing is widely practiced to foster and assess clinical reasoning. Our study aimed to develop script concordance test (SCT) in the specialty of otolaryngology and test the validation using panel response pattern and consensus index. **Materials and Methods:** The methodology was an evolving pattern of constructing SCTs, administering them to the panel members, and optimizing the panel with response patterns and consensus index. The SCT's final items were chosen to be administered to the students. **Results:** We developed 98 items of SCT and administered them to 20 panel members. The mean score of the panel members for these 98 items was 79.5 (standard deviation [SD] = 4.4). The consensus index calculated for the 98-item SCT ranged from 25.81 to 100. Sixteen items had bimodal and uniform response patterns; the consensus index improved when eliminated. We administered the rest 82 items of SCT to 30 undergraduate and ten postgraduate students. The mean score of undergraduate students was 61.1 (SD = 7.5) and that of postgraduate students was 67.7 (SD = 6.3). Cronbach's alpha for the 82-item SCT was 0.74. Excluding the 22 poor items, the final SCT instrument of 60 items had a Cronbach's alpha of 0.82. **Conclusion:** Our study revealed that a consensus index above 60 had a good item-total correlation and be used to optimize the items for panel responses in SCT, necessitating further studies on this aspect. Our study also revealed that the panel response clustering pattern could be used to categorize the items, although bimodal and uniform distribution patterns need further differentiation.

Keywords: *Consensus index, medical education, medical training assessment, script concordance testing*

Sivaraman Ganesan,
Shital Bhandary[1],
Mahalakshmy
Thulasingam[2],
Thomas Vengail
Chacko[3],
Z. Zayapragassarazan[4],
Surya Ravichandran,
Kalaiarasi Raja,
Karthikeyan
Ramasamy,
Arun Alexander,
Lokesh Kumar
Penubarthi

*Departments of ENT,
[2]Preventive and Social Medicine
and [4]Medical Education,
Jawaharlal Institute of
Postgraduate Medical Education
and Research, Puducherry,
[3]Department of Community
Medicine, Believers Church
Medical College Hospital,
Thiruvalla, Kerala, India,
[1]Department of Public Health
and Medical Education, Patan
Academy of Health Sciences,
Lalitpur, Nepal*

*Address for correspondence:
Dr. Mahalakshmy Thulasingam,
Department of Preventive and
Social Medicine, Jawaharlal
Institute of Postgraduate
Medical Education and
Research, Puducherry, India.
E-mail: mahalakshmi.dr@
gmail.com*

Access this article online

Website:
https://journals.lww.com/IJAB

DOI:
10.4103/ijabmr.ijabmr_604_22

Quick Response Code:

Introduction

Clinical reasoning is an essential skill for clinical practice; hence, it is crucial to incorporate clinical reasoning in the teaching, learning, and assessment of medical undergraduates. Clinical reasoning assessment methods can be categorized into nonworkplace-based, simulated clinical environments and workplace-based assessments.[1] Script concordance test (SCT) is designed to assess clinical reasoning.[2,3] SCT induces clinical reasoning by hypothetico-deductive model. Using the SCT model, expert panel members' responses are initially created based on which the students are evaluated and scored. The degree of concordance of the student's response to that of the expert panel decides the student's score. SCTs have been used in various institutes for fostering and assessing clinical reasoning

among medical undergraduates.[4-6] Despite the presence of guidelines in constructing an SCT,[7] there is no uniform agreement in optimizing the panel members' responses. In the study by Wan, 2015, the author clustered the panel response and identified four patterns to determine the panel members' agreement.[4] Gagnon *et al.*[8] analyzed the "outlier" method, "distance-from-mode," and "judgment-by-experts" in optimizing the panel responses for SCT. Few inherent vulnerabilities of SCT have been identified in the literature noting the credit anomaly in which the extreme answer response, such as −2 and +2, had lesser credit on average and the flaw in aggregate scoring of the panel responses.[9,10]

Another commonly used statistical tool to determine the degree of agreement of a given data on an ordinal scale is the

consensus index.[11] So far, no study has tested the utility of the consensus index in evaluating the panel member response in an SCT. Our study aims to develop and pilot a script concordance testing model for ENT undergraduates in the domains of diagnosis, investigation, and management. We also collated the panel responses and tested the utility of the consensus index and panel response pattern to assess the agreement between the panel members in evaluating the SCT.

Materials and Methods

The present study was conducted in the Department of ENT, in a tertiary care teaching hospital, to develop the script concordance testing to foster clinical reasoning among undergraduate medical students. The institutional ethics committee approved the study (JIP/IEC/2014/9/460). The methodology was designed to be an evolving pattern of constructing SCTs, administering them to the panel members, and analyzing the panel with a response pattern and consensus index, based on which the SCT's final items were chosen to be administered to the students. Item-total correlation and Cronbach's alpha were calculated from the students' scores.

Construction of the scripts

The study used the guidelines put forth by Fournier *et al.*[7] for the construction of SCTs. The case scenarios were made by two specialists in ENT who had more than 5 years of experience in ENT practice and undergraduate teaching. The authors of this study designed the SCT comprising 26 clinical scripts, each clinical script with two to six items, thus a total of 98 items. These items were made at the standard of the undergraduate curriculum. Each script of the SCT was designed to reflect the common ENT conditions for undergraduate students. The scripts were developed to promote clinical reasoning in the domains of diagnosis, investigation, and management for an undergraduate student. An example of a developed scenario in our study is shown in Table 1. The developed scripts are given as Supplementary file (available in web version).

Construction of the expert panel

To achieve the highest possible reliability for our study, we set up a panel comprising 20 members.[12] Each member of the expert panel was a certified ENT specialist who had a work experience of more than 5 years and a minimum of 3 years of teaching experience to undergraduate students. After getting informed consent for the study, SCT was mailed to the panel members. The experts took the SCT independently, and their responses were recorded. After collecting all the panel members' responses, for each item, the number of panel members marking the respective responses was aggregated, as shown in Table 1. In this way, for all 98 items, the aggregated responses were recorded.

Construction of the scoring grid

From the responses obtained from the panel, a credit score was calculated for each item corresponding to the proportion of panel members who have chosen the same response. The credit scores ranged from 0 to 1. The maximal score of 1 was given to the response having the maximal number of panel members (modal response) agreeing to it. A partial credit score was calculated for any nonmodal panel member response. The number of panel members who gave the modal response was taken as the denominator and divided by the other nonmodal response to get this proportional credit scoring. This aggregate scoring method has better construct validity than consensus scoring[13] and better reliability and validity coefficients.[14]

Credit score for each response of an item

$$x = \frac{\text{Number of panel members selecting the response}}{\text{Number of panel members selecting the modal response}}$$

For example, on item 46 from Table 1, 14 members chose "+1" on the Likert scale. Hence "+1" is the modal response. The response "+1" for that item, a score of 1 (14/14), is calculated [Table 2]. Only three members chose "−1" as their response. The response "−1" gets only a partial credit score of 0.21 (3/14). Similarly, the response "+2" gets a partial credit score of 0.14 (2/14), and so on.

In this way, the credit scores for all 98 items in the 26 clinical scripts were tabulated to create the scoring grid [Table 3].

Panel response patterns

For each item, the pattern of expert panel responses was analyzed. Wan[4] classified the panel response patterns into four types: ideal response, uniform response, bimodal, and outlier response. When we attempted to classify our panel responses similarly, we found an additional pattern, which we

Item number	If you were thinking of	And then you find	Response				
			−2	−1	0	+1	+2
45	Hemophilia	History of nose picking	0	0	0	5	15
46	Foreign body nose	History of foul-smelling discharge mixed with blood on the left side	0	3	1	14	2
47	Nasopharyngeal angiofibroma	History of left-side nasal obstruction	17	3	0	0	0
48	Septal deviation	History of recurrent knee swelling	10	3	1	3	3

Table 1: An example of responses by the 20 membered expert panel on a clinical script

−2: Ruled out or almost ruled out; −1: Less likely; 0: Neither more nor less likely; +1: More likely; +2: Certain or almost certain

Table 2: Methodology to calculate the credit for item number 46 from [Table 1]

	−2	−1	0	+1	+2	Calculation of credits
Number of panel members who selected the response for item number 46	0	3	1	14	2	Step I: Identify the answer chosen by most panel members (+1)
The mechanism for the creation of scores	0/14	3/14	1/14	14/14	2/14	Step II: Dividing by the number of members who selected the modal response (*n*=14)
Credit score for item number 46	0	0.21	0.07	1	0.14	Step III: Credit score to be awarded to the student choosing the response

Table 3: Creation of the credits for students for item number 45–48

Items number	−2	−1	0	+1	+2
45	0	0	0	0.33	1
46	0	0.21	0.07	1	0.14
47	1	0.18	0	0	0
48	1	0.3	0.1	0.3	0.3

labeled as the partial ideal response [Figure 1]. We noticed that the panel members were split in choosing the extremes of the options available for some questions. We called this a bimodal response [Figure 1a]. When there was an equal spread in the number of members choosing all five options, it was classified as uniform responses [Figure 1b]. A discrete outlier response [Figure 1c] was labeled when there were one or more responses beyond a nil response. The ideal response pattern meant a close convergence with some variation limited within ≤3 options [Figure 1d]. We noticed the fifth pattern in which there was relatively close convergence with some variation limited by four options chosen by the panel members. We labeled this as a partial ideal response pattern [Figure 1e].

We eliminated the items showing uniform and bimodal patterns (as elaborated in the Results section), and the SCT to be administered to the students had 82 items. Analyzing the panel response patterns to identify the uniform and bimodal response patterns was time-consuming. Hence, we looked for a much simpler tool to do the same. We tested the use of the consensus index to achieve this purpose.

Consensus index

The consensus index reflects the agreeability among the panel members for each item, and it is calculated using the following formula,

$$\text{Cns}\,(X) = 1 + \sum_{(i=1)}^{n} p_i \log_2 (1 - \frac{|X_i - \mu_X|}{d_X})$$

Where μ_X is the mean of item X, and d_X is the width of X, $d_X = X_{max} - X_{min}$.

The consensus index takes a value ranging from 0 to 100, with complete disagreement being 0 and 100 being the entire agreement. As an ordinal measure of the panel members' scoring, the consensus index is argued to be superior to mean and standard deviation (SD).[11]

Script concordance test administration to students and scoring

Thirty undergraduate students and ten postgraduate students of ENT volunteered for this study. Informed consent was taken after explaining their role in this study. The SCT was provided in printed format to the students, and their responses were marked on the answer sheet. No time limit was set to complete the SCT. Based on the scoring grid, credits were awarded to the students for all 82 items. The total marks of each student and mean marks scored by all students in each item were calculated. Each student's total credit score was calculated and converted to a 100-point score. The mean marks scored by all students were calculated separately for undergraduate and postgraduate students.

Statistical analysis

With the confidence interval set at 95%, the students' mean score for each item was calculated along with its SD. These scores of the students were compared with the responses of the experts using a *t*-test. $P < 0.05$ was taken as statistically significant. Cronbach's alpha was calculated to determine the reliability of the test. Pearson's correlation was used to ascertain the item-total correlation. IBM SPSS software (Version 17; SPSS Inc., Chicago, Illinois, USA) was used for statistical analysis.

We used the item-total correlation to identify questions with low values ($r < 0.05$). The items with low item-total correlation were discarded, and Cronbach's alpha reliability coefficient was recalculated after the deletion of such items.[15]

Results

The panel members' summary statistics were already discussed in the materials and methodology of the SCT evolution.

Panel members' response summary

The grid score was applied to the panel members' responses. The expert panel members had a mean score of 79.5 (SD = 4.4). All panel members had a score within two SDs from the mean.

Optimization of panel members' responses

On analyzing the response patterns of the panel members, we noticed five types of response patterns, as described above [Figure 1]. Based on the above classification, we
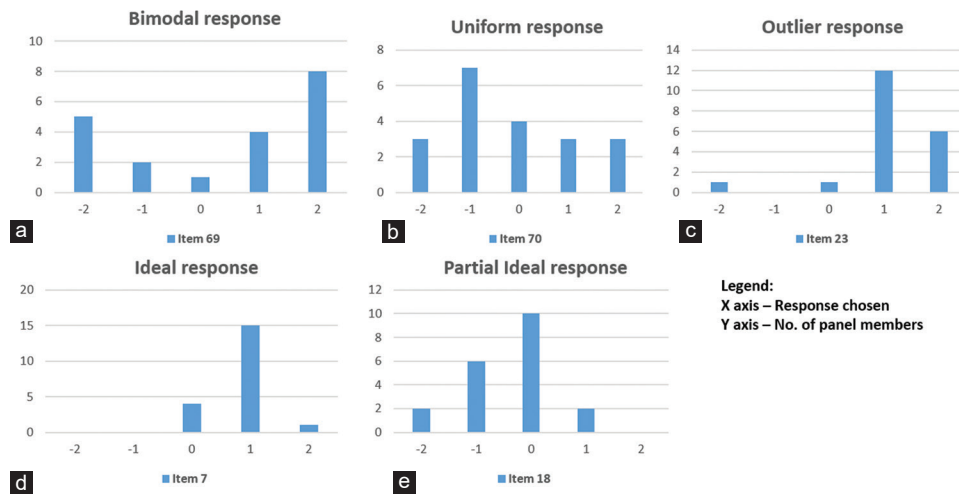
Figure 1: Various response patterns from panel members for individual items, a: bimodal response, b: uniform response, c: outlier response, d: ideal response, e: Partial Ideal Response

noticed that 37 items showed an ideal response pattern, another 37 items showed a partial ideal response pattern, and eight items showed a discrete outlier pattern. The rest of the items had a uniform ($n = 14$) or bimodal response patterns ($n = 2$). These 16 items (2 bimodal and 14 uniform response pattern items) were deleted, and the 82-item SCT was developed for administering to the students.

Over the 98-item SCT on calculating the consensus index for each item, we found that our study's consensus index ranged from 25.8 to 100. We tried to identify if there is any relation between the panel response patterns and the consensus index [Table 4]. We noticed that an item with low agreeability among the expert panel members (low consensus index) is an item that is poorly constructed or overtly confusing.

Panel response pattern and consensus index

The responses recorded from the 20-membered expert panels to the original 98-item SCT and the response pattern type identified were sorted in the descending order of each item's consensus index. In this study, we analyzed the panel members' response patterns and identified 16 items having bimodal or uniform response patterns. Interestingly, when we analyzed the items with a consensus index of <60, they had 2-bimodal, 13-uniform, and 1-partial ideal response pattern. Setting a consensus index cutoff at 60 identified 15 of the 16 items with bimodal or uniform response patterns. In this way, we propose using a consensus index to improve the quality of the items in the SCT.

We need more studies of SCT to determine the permissible cutoff value for the consensus index among the panel members for each SCT item.

Students response summary

On administering the 82-item SCT to the participants, the 30 undergraduate students had a mean score of 61.1 (SD = 7.5), and the ten postgraduate students had a mean score of 67.7 (SD = 6.3) [Table 5].

Table 4: Comparing the consensus index and the panel response pattern

Panel response pattern	Number of responses (n=98)	Consensus index (range)
Ideal responses	37	67.82–100
Partial ideal responses	37	58.65–85.22
Discrete outlier responses	8	60.67–73.7
Uniform responses	14	32.28–61.37
Bimodal responses	2	25.81–58.95

Table 5: Scores of panel members, UG and PG students

	Panel	UGs	PGs
n	20	30	10
Items tested	98	82	82
Mean score	79.5	61.1*	67.7*
SD	4.4	7.5	6.3
Median	79.4	61.0	68.6
Minimum	72.6	43.2	54.5
Maximum	87.5	77.0	75.4

*P-value between UGs and PGs=0.039 (statistically significant using independent Samples t-test). SD: Standard deviation

Item-total correlation

Item-total correlation was done with Pearson's correlation. We categorized the 82-item SCT into three groups: $r < 0.05$ = poor item; $0.05 < r < 0.2$ = fair item, and $r > 0.2$ = good item. 22 items had $r < 0.05$, 15 items had $0.05 < r < 0.2$, and 44 item had $r > 0.2$. Item 6 was found to have mean score of 1 (SD = 0) for undergraduates, as all undergraduate students selected the modal panel response.

Cronbach's alpha

The Cronbach's alpha of the 82-item SCT administered for the students was 0.74. Among these 82 items, 22 items had a poor item-total correlation (<0.05). Excluding those

poor items, the final SCT instrument of 60 items had a Cronbach's alpha of 0.82.

Relationship of consensus index with Cronbach's alpha

We extrapolated our study and tried to identify a permissible cutoff of the consensus index. We tested different consensus indices to learn how it influences Cronbach's alpha and the number of items on SCT [Table 6].

Discussion

In our study, we developed 98 items of SCT. On analyzing the panel responses, we found that 14 had a uniform response, and 2 had bimodal response patterns. The consensus index calculated for the 98-item SCT ranged from 25.8 to 100. We eliminated the 16 items of bimodal and uniform response pattern, leading to 82-item SCT. The consensus index range of these 82 items improved to 58.6–100. The mean score of the 82-item SCT was significantly different between undergraduate and postgraduate students. The 22 items with poor item-total correlation (<0.05) in the student responses were deleted. Excluding these 22 poor items, the final SCT instrument of 60 items had a Cronbach's alpha of 0.82.

Our study found that the mean scores of the undergraduate and postgraduate students are significantly different, proving the ability of SCT to discriminate the students' proficiency levels. As could also be noted, the maximum score of the undergraduate and postgraduate students was very close, showing no statistical difference. It shows that a top-performing undergraduate student will reason well at the postgraduate level. Similar findings of SCT scores differentiating the participants based on their expertise level have been noted in the literature.[16-19]

The study by Wan collated the panel responses into four types.[4] We additionally described a partial ideal response as it seemed necessary as the response clustering was robust in either of the polarities. They described the ideal as a pattern with relatively close convergence with some variation and responses to be clustered within three contiguous responses. At the same time, this leaves out a meaningful pattern that was interesting in our study: the response pattern, including the fourth contiguous response. Still, the last response is very meager, which forced us to classify it as a partial ideal.

The partial ideal category also fulfilled the necessity of close convergence with minimal variation.

In our study, we calculated the consensus index for the responses from the panel members. Whereas the consensus index was not calculated in other studies, studies show that the consensus index may be appropriate or even superior in analyzing the Likert scale's ordinal pattern used in script concordance testing.[11] The consensus index is the closest measure of capturing the collective opinion of the panel member, which may be effectively used in optimizing the items. The items with a low consensus index reflect more variation and need further modification before administering to the students. The range of the consensus index matches the panel response clusters. The relationship between the consensus index and the response clusters may be evident from the table and will require a further detailed statistical analysis. The optimal cutoff for the consensus index to differentiate bad and good items may be the prospect of future study and need a bigger sample size.

We calculated each item's item-total correlation from the undergraduate student's response perspective. We noticed 22 items were poorly correlated to the total score despite the items having an agreeable response from the panel members. Such poor items were identified and excluded to improve the strength of the SCT. The Cronbach's alpha of the final 60 items was 0.82, which indicates a very good internal consistency of the SCT. Even with 82 items initially administered, Cronbach's alpha was 0.74. The Cronbach's alpha we got is comparable to other studies like 0.80 in Iravani *et al*.[5] and 0.73 in Wilson *et al*.[14]

Gagnon *et al*.[8] concluded that optimizing done by an independent review by another three experts to remove the deviant answers ("judgment by experts") is superior to the three methods (outlier, distance-from-mode and judgment-by-experts). They also concluded that "distance-from-mode" is a practical and efficient method. The "distance-from-mode" method they described is more or less similar to the pattern response clustering theme related by Wan[4] and also used in our study. In our study, an item scoring a consensus index of more than 60 seems to optimize the panel responses. The results of the consensus index are also able to shadow the response clustering pattern. The consensus index is easy to calculate and provides an objective mathematical value for optimizing the panel responses. The future prospects may be finding the appropriate cutoff value for the consensus index that can be used in script concordance testing universally.

In our research, when we analyzed converting the 5-point Likert scale to a 3-point Likert scale by fusing the responses on the positive and negative scales, we found no difference in the final Cronbach's alpha or any change in the item-total correlation as a whole. Although this study has a limited sample size from which such conclusions of usage of a 3-point or 5-point, or 7-point Likert scale

Table 6: Relation of the consensus index and Cronbach's alpha with the respective number of items of the script concordance test

Consensus index	Number of eligible items	Cronbach's alpha with number of similar items
>58.64*	82	0.74
>60	81	0.74
>70	59	0.76
>80	28	0.96

*58.65 was the minimum consensus index of 82-item SCT administered to the students. SCT: Script concordance test

could not be made, the authors felt that the 5-point Likert scale is essential to force the student to choose either of the responses on the positive or negative aspect. However, when it is need to simplify the process in clinical setting such as rounds, using SCT in clinical 3-point Likert scale may be useful.[20] Using a 7-point Likert scale can be problematic as, for example, when the panel members mark-2 as their response, a student marking-3 might get zero credit, though his thinking was in the right direction.[21]

Conclusion

Our study revealed that a consensus index above 60 had a good item-total correlation with good internal consistency. Further studies are required to determine the exact cutoff values for the consensus index. Our study also revealed that the panel response clustering pattern could be used to categorize the items, although bimodal and uniform distribution patterns need further differentiation.

Our study also showed that Cronbach's alpha and item-total correlation were useful in detecting the psychometric properties of the SCT. The developed SCT tool has high internal reliability and can be used in the routine assessment of undergraduate students. In the future, such SCT can be integrated into the curriculum of medical students.

Ethical statement

The study was approved by the institutional ethics committee of Jawaharlal Institute of Postgraduate Medical Education and Research (JIP/IEC/2014/9/460).

Financial support and sponsorship

This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sector. This project was done under education innovation project as a part of PSG-FAIMER Regional Institute Medical Education Fellowship Program.

Conflicts of interest

There are no conflicts of interest.

References

1. Daniel M, Rencic J, Durning SJ, Holmboe E, Santen SA, Lang V, *et al*. Clinical reasoning assessment methods: A scoping review and practical guidance. Acad Med 2019;94:902-12.
2. Charlin B, Tardif J, Boshuizen HP. Scripts and medical diagnostic knowledge: Theory and applications for clinical reasoning instruction and research. Acad Med 2000;75:182-90.
3. Charlin B, van der Vleuten C. Standardized assessment of reasoning in contexts of uncertainty: The script concordance approach. Eval Health Prof 2004;27:304-19.
4. Wan SH. Using the script concordance test to assess clinical reasoning skills in undergraduate and postgraduate medicine.

Hong Kong Med J 2015;21:455-61.
5. Iravani K, Amini M, Doostkam A, Dehbozorgian M. The validity and reliability of script concordance test in otolaryngology residency training. J Adv Med Educ Prof 2016;4:93-6.
6. Humbert AJ, Johnson MT, Miech E, Friedberg F, Grackin JA, Seidman PA. Assessment of clinical reasoning: A Script Concordance test designed for pre-clinical medical students. Med Teach 2011;33:472-7.
7. Fournier JP, Demeester A, Charlin B. Script concordance tests: Guidelines for construction. BMC Med Inform Decis Mak 2008;8:18.
8. Gagnon R, Lubarsky S, Lambert C, Charlin B. Optimization of answer keys for script concordance testing: Should we exclude deviant panelists, deviant responses, or neither? Adv Health Sci Educ Theory Pract 2011;16:601-8.
9. Lineberry M, Kreiter CD, Bordage G. Threats to validity in the use and interpretation of script concordance test scores. Med Educ 2013;47:1175-83.
10. Lineberry M, Hornos E, Pleguezuelos E, Mella J, Brailovsky C, Bordage G. Experts' responses in script concordance tests: A response process validity investigation. Med Educ 2019;53:710-22.
11. Tastle WJ, Wierman MJ. Consensus and dissention: A measure of ordinal dispersion. Int J Approx Reason 2007;45:531-45.
12. Gagnon R, Charlin B, Coletti M, Sauvé E, van der Vleuten C. Assessment in the context of uncertainty: How many members are needed on the panel of reference of a Script Concordance Test? Med Educ 2005;39:284-91.
13. Charlin B, Desaulniers M, Gagnon R, Blouin D, van der Vleuten C. Comparison of an aggregate scoring method with a consensus scoring method in a measure of clinical reasoning capacity. Teach Learn Med 2002;14:150-6.
14. Wilson AB, Pike GR, Humbert AJ. Analyzing script concordance test scoring methods and items by difficulty and type. Teach Learn Med 2014;26:135-45.
15. Faucher C, Dufour-Guindon MP, Lapointe G, Gagnon R, Charlin B. Assessing clinical reasoning in optometry using the script concordance test. Clin Exp Optom 2016;99:280-6.
16. Nouh T, Boutros M, Gagnon R, Reid S, Leslie K, Pace D, *et al*. The script concordance test as a measure of clinical reasoning: A national validation study. Am J Surg 2012;203:530-4.
17. Wan MS, Tor E, Hudson JN. Construct validity of script concordance testing: Progression of scores from novices to experienced clinicians. Int J Med Educ 2019;10:174-9.
18. Subra J, Chicoulaa B, Stillmunkès A, Mesthé P, Oustric S, Rougé Bugat ME. Reliability and validity of the script concordance test for postgraduate students of general practice. Eur J Gen Pract 2017;23:208-13.
19. Omega A, Wijaya Ramlan AA, Soenarto RF, Heriwardito A, Sugiarto A. Assessing clinical reasoning in airway related cases among anesthesiology fellow residents using Script Concordance Test (SCT). Med Educ Online 2022;27:2135421.
20. Tayce JD, Saunders AB. The Use of a Modified Script Concordance Test in Clinical Rounds to Foster and Assess Clinical Reasoning Skills. J Vet Med Educ 2022;49:556-9.
21. Bland AC, Kreiter CD, Gordon JA. The psychometric properties of five scoring methods applied to the script concordance test. Acad Med 2005;80:395-9.

Supplementary File

1. A 16 year old female presents to ENT OPD with fever for two days and on examination you find a membrane over tonsil

If you were thinking of	And then you find	This diagnosis becomes:
1.Diphtheria	Greyish white membrane difficult to remove	-2 -1 0 +1 +2
2.Vincent angina	Curdy white patch	-2 -1 0 +1 +2
3.Membranous tonsillitis.	An exudative membrane easy to remove	-2 -1 0 +1 + 2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

2. A 6-year-old previously healthy boy presents with acute onset of fever of 39°C (102°F), severe throat pain that is exacerbated by swallowing, headache, and malaise.

If you were thinking of	And then you find	This diagnosis becomes:
4.Acute epiglottitis	His tonsils are symmetrically enlarged and red, with purulent exudate.	-2 -1 0 +1 +2
5.Acute Laryngotracheobronchitis	No respiratory difficulty	-2 -1 0 +1 +2
6.Peritonsillar abscess	Marked Odynophagia with muffled and thick speech	-2 -1 0 +1 +2
7.Acute retropharyngitis	Neck stiffness	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

3. A 3-year-old girl child presents with acute onset of high grade fever with minimal respiratory distress

If you were thinking of	And then you find	This diagnosis becomes:
8.Acute epiglottitis	Child being Toxic with drooling of saliva	-2 -1 0 +1 +2
9.Adenoiditis	Voice change	-2 -1 0 +1 +2
10.Acute tonsillitis	Difficulty in swallowing	-2 -1 0 +1 +2
11.Diphtheria	Hoarse voice, croupy cough	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

4. A 3-year-old boy was presented with a history of sudden onset respiratory difficuilty.

If you were thinking of	And then you find	This diagnosis becomes:
12.Acute episode of asthma	No previous episodes	-2 -1 0 +1 +2
13.Foreign body ingestion	Difficuilty swallowing	-2 -1 0 +1 +2
14.Foreign body aspiration	Playing with colourful interlocking plastic bricks	-2 -1 0 +1 +2
15.Foreign body aspiration	Past history of fever for 4-5 days	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"

1 "More likely"
2 "certain or almost certain

5. A 57-year-old man presents with a 6-month history of hoarseness of voice

If you were thinking of	And then you find	This diagnosis becomes:
16.Cancer larynx	History of chronic cough with hemoptysis	-2 -1 0 +1 +2
17.Tuberculosis of larynx	History of odynophagia	-2 -1 0 +1 +2
18.Singer's nodule	History of smoking	-2 -1 0 +1 +2
19.Ortners syndrome	History of mitral stenosis	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

6. A 45-year-old man presents with a 3-month history of painless, enlarging, left-sided neck mass

If you were thinking of	And then you find	This diagnosis becomes:
20.Thyroid mass	Moving with deglutition	-2 -1 0 +1 +2
21.Tuberculosis	Granulation over the posterior glottis (cord and arytenoid)	-2 -1 0 +1 +2
22.Cancer esophagus	Left vocal cord palsy	-2 -1 0 +1 +2
23.Supraclavicular lymph node	Abdominal lump on examination	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

7. A 7-year-old girl presents with frequent nosebleeds, worse on the left

If you were thinking of	And then you find	This diagnosis becomes:
24.Hemophilia	History of nose picking	-2 -1 0 +1 +2
25.Foreign body nose	History of foul smelling discharge mixed with blood on the left side	-2 -1 0 +1 +2
26.Nasopharyngeal angiofibroma	History of left side nasal obstruction	-2 -1 0 +1 +2
27.Septal deviation	History of recurrent knee swelling	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

8. A 30 year male presents with nasal obstruction

If you were thinking of	And then you find	This diagnosis becomes:
28.Septal deviation	Decreased fogging on the left side	-2 -1 0 +1 +2
29.Antrochoanal polyp	Mass in the left side with bleeds on probing	-2 -1 0 +1 +2
30.Inverted papilloma	Roomy nasal cavity	-2 -1 0 +1 +2
31.rhinosporidiosis	leafy, polypoidal mass, pink to purple in color with surface studded with white spots	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

9. A 50year old male present with epistaxis

If you were thinking of	And then you find	This diagnosis becomes:
32.Cancer nasopharynx	Unilateral SOM	-2 -1 0 +1 +2
33.Nasopharyngeal angiofibroma	Mass in nasopharynx	-2 -1 0 +1 +2
34.Foreign body nose	Bilateral enlarged lymphnodes	-2 -1 0 +1 +2
35.Atrophic rhinitis	Bilateral nasal obstruction	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

10. A 15 year male anosmia presents to you

If you were thinking of	And then you find	This diagnosis becomes:
36.Ethmoidal polyp	Unilateral nasal obstruction	-2 -1 0 +1 +2
37.Deviated nasal septum	Inferior turbinate hypertrophy	-2 -1 0 +1 +2
38.Nasal bone fracture	Watery discharge when the patient leans forward	-2 -1 0 +1 +2
39.rhinoscleroma	Granulomatous lesion in the nose	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

11. A 35 year female with nasal symptoms underwent Diagnostic nasal endoscopy

If you were thinking of	And then you find	This diagnosis becomes:
40.Inverted papilloma	Friable bleeding mass in the nasal cavity	-2 -1 0 +1 +2
41.Nasal polyposis	Minimal polypoidal change in the left middle meatus	-2 -1 0 +1 +2
42.Chronic sinusitis	Mucopurulent discharge in sphenoethmoidal recess	-2 -1 0 +1 +2
43.Saddle nose	Spur on the left side septum	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

12. A 8 year old male child with nasal discharge

If you were thinking of	And then you find	This diagnosis becomes:
44.Adenoiditis	Bilateral retracted tympanic membrane with air fluid level	-2 -1 0 +1 +2
45.Antrochoanal polyp	Posterior rhinoscopy polyp visualized in choana	-2 -1 0 +1 +2
46.Rhinolith	History of foreign body removal	-2 -1 0 +1 +2
47.CSF rhinorrhoea	Purulent nasal discharge, Sneezing, nasal stuffiness	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

13. A 60 year old male presenting with headache underwent x-ray PNS

If you were thinking of	And then you find	This diagnosis becomes:
48. Frontal Mucocele	Presence of scalloping with haziness in frontal sinus	-2 -1 0 +1 +2
49. Chronic sinusitis	Bony septum deviated to right	-2 -1 0 +1 +2
50. Malignancy of nose and PNS	Haziness extending outside maxillary sinus	-2 -1 0 +1 +2
51. Deviated septum	Heterogeneity in the maxillary sinus	-2 -1 0 +1 +2
52. Actue sinusitis	Air fluid level in the maxillary sinus	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

14. A 40 year old male with left ear discharge

If you were thinking of	And then you find	This diagnosis becomes:
53. ASOM	facial nerve palsy	-2 -1 0 +1 +2
54. Otitis externa	History of throat pain	-2 -1 0 +1 +2
55. Ramsay hunt syndrome	painful vesicles in the ear	-2 -1 0 +1 +2
56. CSOM	History of ear discharge for 2 months	-2 -1 0 +1 +2
57. Malignant otitis externa	The patient is uncontrolled diabetic	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

15. A 23 year old male with ear discharge

If you were thinking of	And then you find	This diagnosis becomes:
58. ASOM	Watery discharge	-2 -1 0 +1 +2
59. Furuncle of EAC	Serous discharge	-2 -1 0 +1 +2
60. CSOM (TTD)	Scanty fowl smelling discharge	-2 -1 0 +1 +2
61. Glomus tumor	Blood stained discharge	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

16. A 30 year female with hard of hearing

If you were thinking of	And then you find	This diagnosis becomes:
62. otosclerosis	No family history and intact TM	-2 -1 0 +1 +2
63. CSOM	A perforated tympanic membrane	-2 -1 0 +1 +2
64. Menieres diseases	no tinnitus or vertigo	-2 -1 0 +1 +2
65. Otitis externa	Wax in the ear	-2 -1 0 +1 +2

Judgement type: Diagnosis
-2 "Ruled out or almost ruled out"
-1 "Less likely"
0 "neither more nor less likely"
1 "More likely"
2 "certain or almost certain

17. A 25 year old female presented to EMS with trauma to the midface

If you were thinking of	And then you find	This investigation becomes:				
66.x-ray nasal bone	External nasal frame work normal with Tenderness and crepitus over the dorsum of nose	-2	-1	0	+1	+2
67.x-ray PNS	Left eye enophthalmos and she is 3 month pregnant	-2	-1	0	+1	+2
68.beta2 transferrin	few drops or a stream of fluid gushes down when bending forward or straining	-2	-1	0	+1	+2

Judgement type: investigation
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

18. A 20 year male patient presenting with hot potato voice and a oropharyngeal mass

If you were thinking of doing	And then you find	This investiation becomes:				
69.Biopsy	Unilateral tonsillar enlargement	-2	-1	0	+1	+2
70.Needle aspiration	Bulge in the posterior 1/3 of tongue	-2	-1	0	+1	+2
71.x- ray STN	Retropharyngeal bulge	-2	-1	0	+1	+2

Judgement type: investigation
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

19. A 5 year child presented with history of respiratory distress.

If you were thinking of doing	And then you find	This invesigation becomes:				
72.Direct laryngoscopy	rapidly progressing history with high grade fever	-2	-1	0	+1	+2
73.x-ray chest	History suggesting aspiration	-2	-1	0	+1	+2
74.x-ray nasal bone	history of croupy cough	-2	-1	0	+1	+2

Judgement type: investigation
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

20. A 19-month-old boy presents with a swinging fever and irritability. The parents report that the child has been pulling at his throat and ears, which they think indicates that the child is in pain.

If you were thinking of	And then you find	This treatment becomes:				
75.Incision and drainage	x- ray soft tissue neck lateral view Thumb sign	-2	-1	0	+1	+2
76.Steroids	x-ray STN showing steeple's sign	-2	-1	0	+1	+2
77.Antibiotics	Membrane Over the Tonsil	-2	-1	0	+1	+2
78.Ampicillin	Paul bunnel test positive	-2	-1	0	+1	+2

Judgement type: treatment
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

21.   A 5 year child presented to EMS with history of Foreign body intake

If you were thinking of doing	And then you find	This treatment becomes:
79.Bronchoscopy	Unilateral decrease in lung sound and x- ray shows hyperinflation on same side	-2 -1 0 +1 +2
80.Oesophagoscopy	Pooling of saliva in right piriform sinus on tele laryngoscopy	-2 -1 0 +1 +2

Judgement type: treatment
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

22. A 50 year male presented with difficulty in swallowing after the party session.

If you were thinking of doing	And then you find	This treatment becomes:
81.Oesophagoscopy	History of choking with aphonia in the party itself in your presence	-2 -1 0 +1 +2
82.Heimleichs maneuver	x-ray radio opaque foreign body at the level of C4-C5 of prevertebral shadow	-2 -1 0 +1 +2

Judgement type: treatment
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

23. A 45 year male with cancer larynx

If you were thinking of doing	And then you find	This treatment becomes:
83.Tracheostomy	Noisy breathing	-2 -1 0 +1 +2
84.Total laryngectomy	Lung metastasis	-2 -1 0 +1 +2
85.Radiation	T1 larynx	-2 -1 0 +1 +2

Judgement type: treatment
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

24. A 10 year with chronic tonsillitis presents to the OPD

If you were thinking of doing	And then you find	This diagnosis becomes:
86.Managing conservatively	OSAS	-2 -1 0 +1 +2
87.Tonsillectomy	<5 episodes in 1 year	-2 -1 0 +1 +2
88.Tonsillectomy	Patient is a Diphtheria carrier	-2 -1 0 +1 +2
89.Adenotonsillectomy	No history snoring and impedence normal	-2 -1 0 +1 +2

Judgement type: treatment
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

25. A 25 year old male with nasal obstruction and haziness of left maxillary sinus

If you were thinking of	And then you find	This treatment becomes:
90.Proof puncture	Air fluid level in the maxillary sinus	-2 -1 0 +1 +2
91.biopsy	Septal bulge indicating hematoma	-2 -1 0 +1 +2
92.Cald well luc	Suspicion of malignancy in CT scan	-2 -1 0 +1 +2
93.polypectomy	Bilateral Pale polyps in DNE responding to nasal steroid	-2 -1 0 +1 +2
94.FESS	Purulent discharge in the middle meatus	-2 -1 0 +1 +2

Judgement type: treatment
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"

26. A 30 year male with chronic sinusitis is due for treatment

If you were thinking of	And then you find	This treatment becomes:
95.antibiotic	Polyp in the middle meatus extending to the choana	-2 -1 0 +1 +2
96.Nasal steroid spray	Purulent discharge in the middle meatus	-2 -1 0 +1 +2
97.antihistamine	No feature of allergic rhinitis	-2 -1 0 +1 +2
98.Oral steroid	Is diabetic	-2 -1 0 +1 +2

Judgement type: treatment
-2 "Completely or almost completely unnecessary"
-1 "Less useful"
0 "Neither more nor less useful"
1 "More useful"
2 "Completely or almost completely necessary"