

Consensus Analysis for the Panel Responses and the Psychometrics Properties in Script Concordance Testing

Sivaraman Ganesan

Jawaharlal Institute of Post Graduate Medical Education and Research

Shital Bhandary

Patan Academy of Health Sciences

Mahalakshmy Thulasingham (✉ mahalakshmi.dr@gmail.com)

Jawaharlal Institute of Post Graduate Medical Education and Research

Thomas Chacko

Believers Church Medical College Hospital

Zayapragassa Rajan

Jawaharlal Institute of Post Graduate Medical Education and Research

Surya Ravichandran

Jawaharlal Institute of Post Graduate Medical Education and Research

Kalaiarasi Raja

Jawaharlal Institute of Post Graduate Medical Education and Research

Karthikeyan Ramasamy

Jawaharlal Institute of Post Graduate Medical Education and Research

Arun Alexander

Jawaharlal Institute of Post Graduate Medical Education and Research

Sunil Kumar Saxena

Jawaharlal Institute of Post Graduate Medical Education and Research

Research Article

Keywords: Script Concordance testing, Medical Education, medical training assessment, Consensus index

Posted Date: January 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-133349/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Clinical reasoning is an essential attribute in the teaching, learning, and assessment part of medical education for undergraduates. In using the Script Concordance Test (SCT) to foster clinical reasoning, expert panel members' responses are initially created. There is no agreement in optimizing the panel members' responses. Our study aimed to develop and validate an SCT and test the utility of the consensus index and panel response pattern.

Methods: The methodology was an evolving pattern of constructing SCTs, administering them to the panel members, optimizing the panel with response pattern and consensus index. 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.

Results: Our study developed an SCT with 98 items and was administered to 20-panel members. The mean score of the panel members for these 98 items was 79.5 (+/- 4.4 SD). On optimizing with the panel responses, 14 items had a uniform response pattern, and 2 had bimodal response patterns. The consensus index calculated for the 98 item SCT ranged from 25.81 to 100. When the 16 items of bimodal and uniform response pattern were eliminated, the consensus index ranged from 58.65 to 100. We administered this 82 items SCT to 30 undergraduate and ten postgraduate students. The mean score of undergraduate students was 61.1 (+/-7.5 SD), and that of postgraduate students was 67.7 (+/- 6.3 SD), which was statistically significant using an independent t-test. Cronbach's alpha for this 82 item SCT was 0.74. On analysing the item-total correlation, 22 items had a correlation of less than 0.05. Excluding these 22 poor items, the final SCT instrument of 60 items had a Cronbach's alpha of 0.82.

Conclusion: The consensus index can also be used to optimize the items for panel responses in SCT. Our study revealed that a consensus index of above 60 had a good item-total correlation with good internal consistency. Our study also revealed that the panel response clustering pattern could also be used to categorize the items though bimodal and uniform distribution patterns need further differentiation.

Background

Clinical reasoning is an essential attribute in the teaching, learning, and assessment part of medical education for undergraduates. Script Concordance Test (SCT) is a tool to assess clinical reasoning (1,2). Various studies have shown SCT's utility in fostering and assessing clinical reasoning among medical undergraduates (3–5). In using the Script Concordance Test 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 credit. However, SCTs are rarely used in medical education for undergraduates in India.

Despite the presence of guidelines in constructing an SCT (6), there is no uniform agreement in optimizing the panel members' responses. In the study conducted by SH Wan et al., 2015 (3) the author has clustered the panel response and identified four patterns to identify the panel members' agreement. Robert Gagnon et al. (7) analysed the 'outlier' method, 'distance-from-mode' and 'judgment-by-experts' in optimizing the panel responses for Script Concordance Test. The consensus index is a measure of dispersion applied to determine the agreement among any given data on an ordinal scale (8). Though theoretically, it can assess the agreement between panel members' responses, so far, no study has tested the utility of the consensus index in evaluating the panel member response in an SCT.

Our study aimed to develop and validate a Script concordance testing model in the domains of diagnosis, investigation, and management for ENT undergraduates. We also collated the panel responses, 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 carried out in the Department of ENT, in Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India, a tertiary care teaching hospital to develop the Script Concordance testing to foster clinical

reasoning among undergraduate medical students. The Institutional Ethics Committee of JIPMER approved the study (JIP/IEC/2014/9/460). All methods were carried out in accordance with the guidelines and regulations of JIPMER. The methodology was designed to be an evolving pattern of constructing SCTs, administering them to the panel members, analysing the panel with 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. (6) for the construction of script concordance tests. The case scenarios were made by two specialists in ENT who had more than five years' experience in ENT practice and undergraduate teaching. The authors of this study designed an SCT comprising 26 clinical scripts, each clinical script with three to six items, thus a total of 98 items. The developed 98-item SCT is attached as Supplementary File 1. These items were made at the standard of the undergraduate curriculum. Each script of the SCT was designed to reflect the common ENT conditions the undergraduate students face and learn during their clinical postings. The scripts included rhinology, otology, and laryngology. The scripts were developed to promote clinical reasoning in the domains of diagnosis, investigation, and management for an undergraduate student.

Construction of the Expert panel

To achieve the highest possible reliability for our study, we set up a panel comprising 20 members (9). Each panel member was a certified ENT specialist who had a work experience of more than five years and had undergone undergraduate curriculum pedagogy training. The 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 were 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, for each item, a credit score was calculated 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 to the other nonmodal response to get this proportional credit scoring. This method of aggregate scoring has a better construct validity than consensus scoring (10) as well as better reliability and validity coefficients (11).

$$\text{Credit score for each response of an item } x = \frac{\text{Number of panel members selecting the response}}{\text{Number of panel members selecting the modal response}}$$

For example, on item number 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 the 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. SH Wan et al. (3) classified the panel response patterns into four types: ideal response, uniform response, bimodal, and outlier response. When we attempted to classify our panel responses

in a similar manner, we found an additional pattern, which we labelled as partial ideal response, which is elaborated below (**Fig. 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 (**Fig 1a**). When there was an equal spread in the number of members choosing all the five options, it was classified as uniform divergence responses (**Fig 1b**). A discrete outlier response (**Fig 1c**) was labelled 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 (**Fig 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, and we labelled this as *partial ideal response pattern* (**Fig 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. Analysing the panel response patterns to identify the uniform and bimodal response patterns was a time-consuming process. So, 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,

$$Cns(X) = 1 + \sum_{i=1}^n$$

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 -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 (8).

SCT Administration to Students

Around 30 undergraduate (UG) students and 10 postgraduate (PG) 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 in a pre-designated hall for this study, and their responses were to be marked in the answer sheet. No time limit was set to complete the SCT. However, all the participants were required to answer all the 82-item SCT.

Scoring of students

Based on the scoring grid, credits were awarded to the students for all the 82 items, and 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 to yield the result as a percentage. 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 standard deviation (SD). These scores of the students were compared with the responses of the experts using a *t*-test. A p-value of less than 0.05 was taken as significant. The reliability of the test was calculated by Cronbach's alpha. 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.

Item-total correlation is the correlation between the score on a particular question and the collective score on all the remaining questions. In short, questions with low item-total correlations do not produce responses that are consistent with the remainder of the test (12). 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 the Cronbach's alpha reliability coefficient was recalculated after the deletion of such items.

Results

The panel members' summary statistics were already discussed in the materials and methodology in 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 (+/- 4.4 SD). All panel members had a score within two standard deviations from the mean.

Optimization of Panel members responses

On analysing the response patterns of the panel members, we noticed five types of response patterns as described above (**Fig 1**). Based on the above classification, we noticed that 37 items showed an ideal response pattern, another 37 items showed a partial ideal response pattern, and 8 items showed a discrete outlier pattern. The rest of the items had 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 items 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 to the panel response patterns to 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. (**Table 5**)

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 less than 60, it had 2-bimodal, 13-uniform, and 1-partial ideal response pattern. Setting a consensus index cut-off at 60 was able to identify 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 cut-off 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 (+/-7.5 SD), and the 10 postgraduate students had a mean score of 67.7 (+/- 6.3 SD). (**Table 6**)

Item total Correlation

Item-total Correlation was done with Pearson's Correlation. We categorised the 82-item SCT into 3 groups - $r < 0.05$ = poor item; $0.05 < r < 0.2$ = fair item, $r > 0.2$ = good item. 22 items were $r < 0.05$, 15 items had $0.05 < r < 0.2$ and 44* items had $r > 0.2$.

*In the analysis, the undergraduate scores for one item (Item no.6) was found to have a Mean score of 1 with SD 0 as all the students responded with the modal panel responses (+1 and +2).

Cronbach's Alpha

The Cronbach's alpha of the 82 items 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 cut off of the consensus index. We tested on different consensus indices to learn how it influences the Cronbach's alpha and the number of items on SCT. (Table 7)

Discussion

In our study, we developed 98 items of SCT. We found that 14 of them had a uniform response, and 2 of them had bimodal response patterns. The consensus index calculated for the 98 item SCT ranged from 25.81 to 100. When we eliminated the 16 items of bimodal and uniform response pattern, the consensus index ranged from 58.65 to 100. We administered this 82 items SCT to 30 undergraduate and 10 postgraduate students. The mean score of undergraduate students was 61.1 (+/-7.5 SD), and that of postgraduate students was 67.7 (+/- 6.3 SD), which was statistically significant using independent t-test. Cronbach's alpha for this 82 item SCT was calculated to be 0.74. On analysing the item-total correlation, we found 22 items had a correlation of less than 0.05. Excluding these 22 poor items, the final SCT instrument of 60 items had a Cronbach's alpha of 0.82.

The currently available methods to evaluate an undergraduate medical student include Case presentation, Objective structured clinical examination (OSCE), and Multiple Choice Questions (MCQ). These strategies have their own shortfalls of being cumbersome, time-consuming, and resource-intensive. Script Concordance Testing (SCT) offers an innovative approach for the same while addressing these shortcomings. It can be used to assess undergraduate students' clinical reasoning and analytical knowledge as it is a simple yet effective instrument.

In this study, we found the panel response pattern and the consensus index to be good drivers for designing a Script Concordance testing tool. Based on our observation, we noticed the five patterns as described above. The study by SH Wan et al. collated the panel responses into four types (3). We additionally described a partial ideal response as it seemed necessary as the response clustering was robust in either of the polarities. They have described ideal as a pattern with relatively close convergence with some variation. They meant the responses should be clustered within three contiguous responses on either side, including the zero. At the same time, this leaves out a meaningful pattern that was interestingly found in our study, which was the response pattern, including the fourth contiguous response. Still, the last response is very meagre, which forced us to classify it as 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 analysing the Likert scale's ordinal pattern used in Script Concordance Testing (8). The consensus index is the closest measure of capturing the collective opinion of the panel member, which maybe 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. It is also clear that the range of consensus index matches the panel response clusters. The relationship between the consensus index and the response clusters may be evident from the table will require a further detailed statistical analysis. The optimal cut-off for the consensus index to differentiate a bad and good item may be the prospects of future study and will require a bigger sample size.

From the undergraduate student's response perspective, we calculated the item-total correlation for each item. We noticed 22 items were poorly correlating 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 Script Concordance Test. Even with initially administered 82 items, the Cronbach's Alpha was 0.74. The Cronbach's alpha we got is comparable to other studies like 0.80 in Kamyar Iravani et al. (4), 0.73 in Aloysius Humbert et al. (11).

It is also clear that the mean scores of the undergraduate and the Postgraduate students are significantly different, proving the ability of the SCT to discriminate the proficiency levels of the students. As could also be noted, the maximum score of the undergraduate student and the postgraduate student was very close, showing no statistical difference. It shows that a top-performing undergraduate student will reason out well at the post graduate's proficiency level. Similar findings of SCT scores differentiating the participants based on their expertise level have been noted in the literature (13–15).

Robert Gagnon (7) concluded that optimizing done by an independent review by another three experts to remove the deviant answers ('judgment-by-experts') is superior of the three. They also concluded that distance from mode is also a practical and efficient method. The 'distance-from-mode' method described by them is more or less similar to the pattern response clustering theme described by Wan S (3) and also as used in our study. In our study, an item scoring the 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. Finding the appropriate cut off value for the consensus index that can be used in script concordance testing universally may be the future prospects.

The 5-point Likert scale used in Script Concordance Testing has been tested with 3-point and 7-point scores in various studies. 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 scale together and the negative scale, we found no difference in the final Cronbach's alpha or any change in the item-total correlation as a whole. Though this study has a limited sample size from which such conclusions of usage of 3-point or 5-point or 7-point Likert scale could not be made, 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. The use of 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 0 credit though his thinking was in the right direction (16). Similarly, the 5-point Likert scale was found to be better than the 3-point Likert scale in various studies (6,11).

The number of panel members used in our study was 20. This is in accordance with studies (9), which showed no statistically significant advantage of using panel members more than 20.

Conclusions

This study clearly revealed that the consensus index could also be used to item analyze the script concordance test. It also provides a clear-cut numerical value so that it is objective in further analysis. Our study revealed that a consensus index of above 60 had a good item-total correlation with good internal consistency. Further studies are required to find out the exact cut-off values for the consensus index. Our study also revealed that the panel response clustering pattern could also be used to categorize the items though bimodal and uniform distribution patterns need further differentiation.

Our study also showed Cronbach's alpha and item-total correlation were useful in detecting the psychometric properties of the Script Concordance Test.

The developed Script Concordance test (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 during their academic years.

Declarations

Ethics approval and consent to participate:

The study was conducted in Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India and was approved by the Institute ethics committee of JIPMER (JIP/IEC/2014/9/460). Written informed consent was taken from the participants.

Consent for publication:

Not applicable.

Availability of data and materials:

The data set supporting the conclusions of this article can be obtained from the author on request.

Competing interests:

The authors declare that they have no competing interests.

Funding:

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.

Authors' contributions:

Conceptualisation: SG, MT; Methodology: SG, MT, TC, SB, KLR, KTR; Formal analysis: SG, MT, SR, SB; Supervision: AA, SK. The first draft of the manuscript was written by MT, SG, and then revised by SR, SB, TC, ZP. All authors read and approved the final manuscript.

Acknowledgements:

We would like to acknowledge PSG FAIMER, JSAC, and JIPMER IEC.

References

1. Charlin B, Tardif J, Boshuizen H. Scripts and Medical Diagnostic Knowledge: Theory and Applications for Clinical Reasoning Instruction and Research. *Academic medicine: journal of the Association of American Medical Colleges*. 2000 Mar 1;75:182–90.
2. Charlin B, Van der Vleuten C. Standardized Assessment of Reasoning in Contexts of Uncertainty The Script Concordance Approach. *Evaluation & the health professions*. 2004 Oct 1;27:304–19.
3. Wan S. Using the script concordance test to assess clinical reasoning skills in undergraduate and postgraduate medicine. *Hong Kong medical journal = Xianggang Yi Xue za Zhi / Hong Kong Academy of Medicine*. 2015 Aug 28;21.
4. Irvani K, Amini M, Doostkam A, Dehbozorgian M. The validity and reliability of script concordance test in otolaryngology residency training. *Journal of advances in medical education & professionalism*. 2016 Apr 1;4:93–6.
5. Humbert A, Johnson M, Miech E, Friedberg F, Grackin J, Seidman P. Assessment of clinical reasoning: A Script Concordance test designed for pre-clinical medical students. *Medical teacher*. 2011 Jun 1;33:472–7.

6. Fournier JP, Demeester A, Charlin B. Script Concordance Tests: Guidelines for Construction. *BMC medical informatics and decision making*. 2008 Feb 1;8:18.
7. 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? *Advances in health sciences education : theory and practice*. 2011 Feb 1;16:601–8.
8. Tastle WJ, Wierman MJ. Consensus and dissent: A measure of ordinal dispersion. *International Journal of Approximate Reasoning*. 2007 Aug 1;45(3):531–45.
9. Gagnon R, Charlin B, Coletti M, Sauv e 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? *Medical education*. 2005 Apr 1;39:284–91.
10. 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. *Teaching and learning in medicine*. 2002 Feb 1;14:150–6.
11. Wilson A, Pike G, Humbert A. Analyzing Script Concordance Test Scoring Methods and Items by Difficulty and Type. *Teaching and learning in medicine*. 2014 Apr 1;26:135–45.
12. Faucher C, Dufour-Guindon M-P, Lapointe G, Gagnon R, Charlin B. Assessing clinical reasoning in optometry using the script concordance test: Assessing clinical reasoning in optometry *Faucher, Dufour-Guidon, Lapointe, Gagnon, and Charlin*. *Clin Exp Optom*. 2016 May;99(3):280–6.
13. 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. *American journal of surgery*. 2012 Apr 1;203:530–4.
14. Wan M, Tor E, Hudson J. Construct validity of script concordance testing: progression of scores from novices to experienced clinicians. *International Journal of Medical Education*. 2019 Sep 20;10:174–9.
15. Subra J, Chicoulaa B, Stillmunk s A, Mesth  P, Oustric S, Roug  Bugat M-E. Reliability and validity of the script concordance test for postgraduate students of general practice. *European Journal of General Practice*. 2017 Oct 2;23:208–13.
16. Bland A, Kreiter C, Gordon J. The Psychometric Properties of Five Scoring Methods Applied to the Script Concordance Test. *Academic medicine : journal of the Association of American Medical Colleges*. 2005 May 1;80:395–9.

Tables

Due to technical limitations, tables are only available as a download in the Supplemental Files section.

Figures

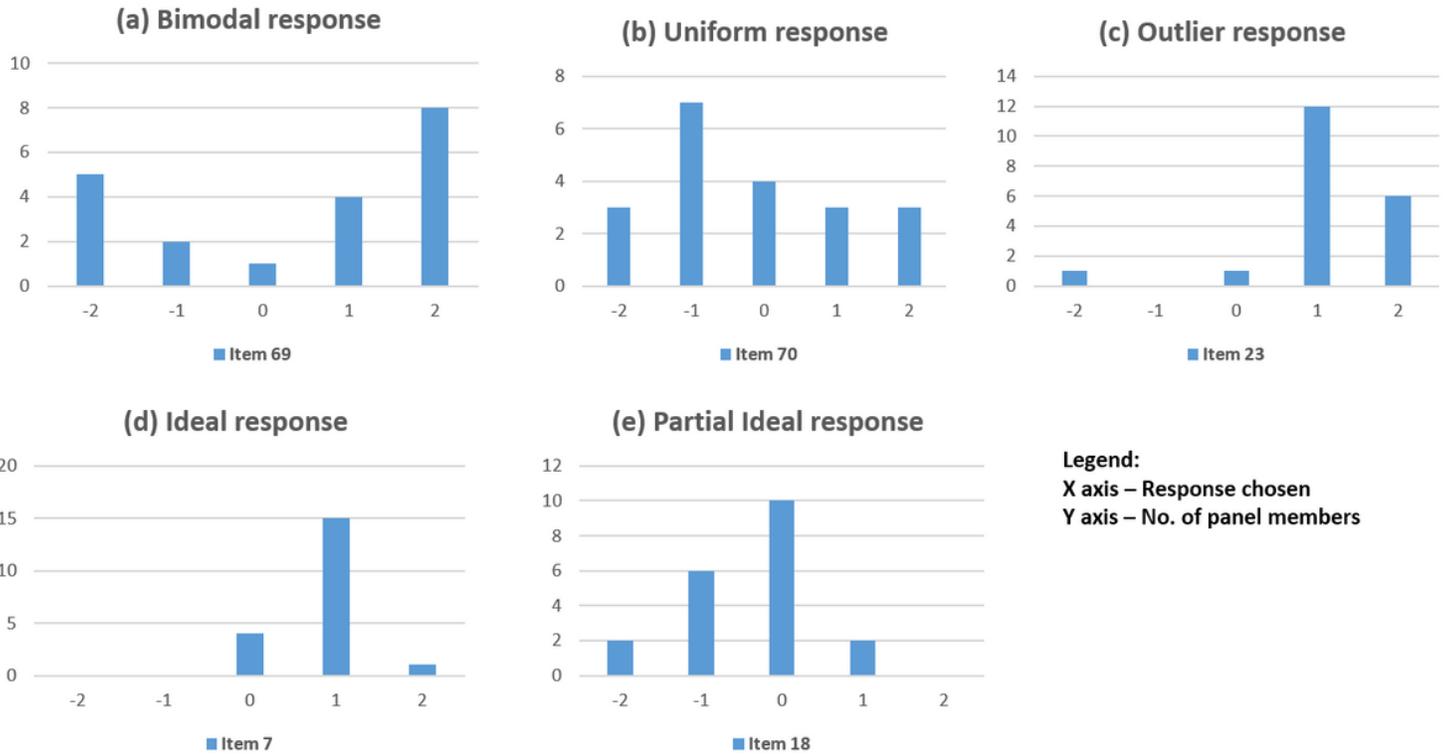


Figure 1

Various response patterns from panel members for individual items

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1Anexampleofresponsesbythe20memberedexpertpanelonaclinicalscript.docx](#)
- [Table2Methodologyto calculatethecreditforitemnumber46fromTable1.docx](#)
- [Table3CreationofthecreditsforstudentsforItemno.4548.docx](#)
- [Table4Comparingtheconsensusindexandthepanelresponsepattern.docx](#)
- [Table5Responsesofthepanelresponsepatternandconsensusindexofeachitem.docx](#)
- [Table6.ScoresofPanelmembersUGandPGstudents.docx](#)
- [Table7.TableshowingtherelationoftheconsensusindexandtheCronbachsalphawiththerespectivenumberofitemsoftheSCT.docx](#)
- [SupplementaryFile1.doc](#)