



A STUDY OF EFFECT OF STRUCTURED QUESTION PAPER ON TEACHERS EVALUATION - A COMPARATIVE STUDY

Dr Ravindra Y M

Assistant Professor, Department of Community Medicine, Shridevi Institute of Medical Sciences and Research Hospital.

Dr Raju Hanumant Patil*

Assistant Professor, Prakash Institute of Medical Sciences and Research Centre, Urun-Islampur, Maharashtra. *Corresponding Author

ABSTRACT A test was conducted with 2 question papers: Paper I and Paper II. Paper I was unstructured question paper and Paper II was structured question paper. Once collected, all the answer scripts were coded. All the 60 answer scripts (Paper I -30 scripts and Paper II-30 scripts) were evaluated by all the 3 teachers twice at an interval of one week. Marks were tabulated analysed for intra-rater and inter-rater variability for structured and unstructured question papers because when previous university question papers were analysed for the pattern of questions, it was found that majority of the question papers were not having structured questions. Hence it was decided to conduct this study.

KEYWORDS :

Introduction:

Written assessment methods, especially essay questions are the most commonly used assessment tools to assess cognitive domain. Essay questions can be long essay questions, short essay questions or short answer questions. Essay questions even though very commonly used assessment tools, are considered to be poor in the reliability of grading. Different examiners may vary in their grading of the same or similar answers (inter-rater reliability) and the same examiner can vary significantly in his/her grading consistency depending on many factors (intra-rater reliability)¹⁻². Structured essay questions with mark distribution for each component will help to improve the reliability of evaluation³⁻⁹. It is a well known fact that structured essay questions will improve the reliability of essay questions. But when previous university question papers of almost all the subjects were analysed for the pattern of questions, it was found that majority of the question papers were not having structured questions. Few question papers were partially structured i.e only long essay questions were structured and short essay questions, short answer questions were found to be unstructured questions. It was also observed that in many question papers long essay questions even though were structured, marks distribution was not indicated for the sub questions. Hence it was decided to conduct this study so as to emphasize the importance of structured question papers in improving the reliability and the necessity to strictly implement the same in the summative assessments.

Aim and Objectives:

To compare inter-rater and intra-rater reliability of structured versus unstructured question papers.

Materials & Methods:

60 first year MBBS students, who were willing, selected for the study.

This study was done in the Department of PSM, Shridevi Institute of Medical Sciences and Research Hospital. Tumkur Three teaching faculty with varied teaching experience from the Department of Anatomy were included in the study as examiners.

A test was conducted with 2 question papers: Paper I and Paper II. Paper I was unstructured question paper and Paper II was structured question paper. Once collected, all the answer scripts were coded. All the 60 answer scripts (Paper I -30 scripts and Paper II-30 scripts) were evaluated by all the 3 teachers twice at an interval of one week. Marks were tabulated analysed for intra-rater and inter-rater variability for structured and unstructured question papers.

Statistical Analysis:

Student t test (two tailed, dependent) has been used to find the significance of study parameters on continuous scale with in each group.

Data will be analysed using

Pearson correlation between study variables is performed to find the degree of relationship or inter-rater reliability/Intra-rater reliability.

Significant figures

+ Suggestive significance (P value: 0.05<P<0.10)

* Moderately significant (P value:0.01<P 0.05)

** Strongly significant (P value: P0.01)

Statistical software: The Statistical software namely SAS 9.2, SPSS 15.0, Stata 10.1, MedCalc 9.0.1 ,Systat 12.0 and R environment ver.2.11.1 were used for the analysis of the data and Microsoft word and Excel have been used to generate graphs, tables etc.

Results:

Table 1: Paper I (Unstructured question paper)

	First Valuation Mean±SD	Second valuation Mean±SD	Difference	t value	P value	Inter-rater P value	
Examiner 1	13.97±3.11	15.63±3.69	1.667	-3.226	0.003**	Examiner 1-2	0.396
Examiner 2	13.57±3.35	13.03±3.18	0.533	1.743	0.092+	Examiner 1-3	0.007**
Examiner 3	12.60±3.32	12.25±3.29	0.350	1.497	0.145	Examiner 2-3	0.010**

Table 2: Paper II (Structured question paper)

	First Valuation Mean±SD	Second valuation Mean±SD	Difference	t value	P value	Inter-rater P value	
Examiner 1	13.18±4.10	14.00±4.50	0.817	1.688	0.102	Examiner 1-2	0.082+
Examiner 2	12.43±4.55	11.83±3.73	0.600	2.090	0.045*	Examiner 1-3	0.522
Examiner 3	12.95±4.44	12.67±4.30	0.283	1.517	0.140	Examiner 2-3	0.170

Table 3: Pearsons Correlation in Paper I & Paper II

Intra-rater reliability	r value		Inter-rater reliability	r value	
	Paper I	Paper II		Paper I	Paper II
Examiner 1: 1 st vs 2 nd valuation	0.666	0.814	Examiner 1 vs Examiner 2	0.692	0.866
Examiner 2: 1 st vs 2 nd valuation	0.869	0.947	Examiner 1 vs Examiner 3	0.685	0.896

Examiner 3: 1 st vs 2 nd valuation	0.925	0.973	Examiner2 vs Examiner3	0.835	0.900
--	-------	-------	------------------------------	-------	-------

Discussion:**Literature review:**

1)Richard J Hift Should essays and other open-ended type questions retain a place in written summative assessment in clinical medicine. BMC Medical Education 2014, 14:249.

Author in this article mentioned that written assessment methods are easier to organize, compared to practical assessment methods. Hence written assessment tools such as essay questions are commonly used to assess knowledge. Author also states that all assessments, especially summative one should meet an adequate standard in terms of quality. Quality of assessment methods can be measured by a number of parameters, Principal among these are the reliability or consistency which means that a result would not change with retesting under similar conditions. Essay questions are poor in their reliability, as the evaluation is influenced by many external factors.

2)Nithyaanandam Suneetha, Joseph Mary, Sriram Sangeetha, Acharya Nibedita, Priya Yamini, Vasu Usha Reliability of an Essay Question – Effect of Using an Answer Key. Journal of Research in Medical Education & Ethics 2015(5); 2:134-137.

The aim of the study was to evaluate reliability of essay and short essay questions and to evaluate the role of key answers in increasing the reliability of essay questions. The study concludes by stating essay questions are not reliable tools but using answer key for evaluation helps in improving the reliability.

Conclusion:

Structuring of Essay questions will help to improve inter and intra-rater reliability of its Grading

References:

1. Richard J Hift Should essays and other open-ended type questions retain a place in written summative assessment in clinical medicine. BMC Medical Education 2014, 14:249
2. Nithyaanandam Suneetha, Joseph Mary, Sriram Sangeetha, Acharya Nibedita, Priya Yamini, Vasu Usha Reliability of an Essay Question – Effect of Using an Answer Key. Journal of Research in Medical Education & Ethics 2015(5); 2:134-137.
3. New York State Education Department (February 2012). Guidance on New York State's Annual Professional Performance Review Law and Regulations. Albany, NY: Author.
4. New York State Education Department (March 2012). Guidance on the New York State District Wide Growth Goal-Setting Process: Student Learning Objectives. Albany, NY: Author.
5. New York State Union of Teachers. Integrated Teacher Evaluation and Development System: A Plan for Teacher Evaluation and Development. Albany, NY: Author.
6. NY: Author. New York State Union of Teachers (August 2012). NYSUT's Teacher Practice Rubric Aligned with the New York State Teaching Standards. Albany, NY: Author.
7. New York State Union of Teachers (July 2011). NYSUT's Teacher Practice Rubric Aligned with the New York State Teaching Standards. Albany, NY: Author.
8. New York State Union of Teachers. Teacher Evaluation and Development Implementation Guide. Albany, NY: Author.
9. Plattsburgh City School District (January 2013). Annual Professional Performance Review Plan 2012– 2013. Retrieved from <http://usny.nysed.gov/rtrt/teachers-leaders/plans/docs/plattsburgh-appriplan.pdf> on June 10, 2013.