

Non-sequential Answers in Written Examinations and Errors Committed by Examiners in Checking Them

Kamal Krishna De¹, Subhas Chandra Bhat², Ujjwal Paul³, Niranjana Maity⁴

¹Retd. Principal, ²Associate Professor, ^{3,4}Assistant Professor

¹David Hare Training College & IASE, Kolkata-700 019, India

^{2,3,4}Government College of Education & CTE, Banipur, PIN-743233, India

Abstract: Some students while answering questions and sub-questions do not follow sequence of questions and answers. While checking such answer scripts examiners commit errors and take longer time in evaluation. The present investigators developed questions and their answers in both sequential and non-sequential form. These were given a group of examiners including male and female, humanities and science post graduate pre-service teachers (B.Ed.). Subject matter was more and more based on General Knowledge of class VIII standard. It was found that (i) in same time frame the non-sequenced answers give rise to more checking errors (ii) no. of errors arising in checking sequenced and non-sequenced answers are highly correlated. (iii) the checking errors so committed do not depend significantly on examiners' gender and subject of study in higher education.

Keywords: Sequential and non-sequential answers, MCQ-its stems and options for answer, Mental Work Load

Introduction: Orderly arrangement in any work paves the way of smooth performance. In the universe there are law and order everywhere, which the scientists reveal from time to time. But this order or system or sequence can't be always maintained strictly in earthly works like examination in schools, colleges and in other activities for different reasons. One such disorder is often found in the answers of the students against different questions in examinations. Examiners allege that many students of secondary examinations do not write answers to respective questions in proper place and in proper order. This type of mismatch in sequence of answers with the sequence of the questions breeds many difficulties particularly for the examiners who are eager to check the answers within a stipulated time frame (or examiner's hours). In checking these types of answers the examiners are stressed and take more time for evaluation. If extra time can't be allowed for some reasons, errors in evaluation arise. This is ominous both for students and examiners for obvious reasons.

In the ordered answers whether wrong or right, the answers are placed near the question number and it might demand less attention, less memory and less mental load for evaluation. But the disorder in writing answers might put extra mental load due to searching of the corresponding answers requiring extra time. Megaw and Sharples (in Kantowitz, 2000) showed that Mental Work Load (MWL) can be increased until a point where task performance will suddenly drop, as when the demands of the task exceed the person's resources available, performance errors can happen. These maximum resources vary by both individual (e.g. cognitive span) and situational (e.g. alertness, vigilance, and fatigue) factors.

During the checking of answer scripts in final board examinations, it has been often observed that if same number of scripts are allotted to the examiners by their head examiner to complete their works within a stipulated period of time, two things generally happen: (i) some of the examiners submit examined answer scripts in time where there are sizeable number of errors in checking (ii) some take more time to avoid or minimize such errors.

Non-sequenced answers occur in different manners. Some of them as shown below:

(i) A segment of answer of one question appears with the segment of answer of some other question. This arouse confusions among examiners who want finish their jobs as quickly as possible. Such haphazard ways of writing answers are mostly found in the answer papers of students of rural and backward areas. Perhaps the students of these areas don't rigorously practise how to answer questions in proper way.

(ii) These non-sequenced answers mostly occur in the case of answering short answer type questions. An answer to question no. 3(a) might appear after 7(c). A part of a long answer might appear after a part of another long or short answer. When MCQ are answered on a formatted sheet there is no scope misplacement of answers. But if the answer book is meant for all types of questions, the answer to MC Questions might be sandwiched between other answers. In this way even, evaluation of a single answer-script might become a herculean task. This is definitely a hurdle in the process reliability of evaluation.

(iii)The dislocation of answers may also occur in general for the following cases:

(a) In describing a figure (geometrical and others) candidates are sometimes compelled to use different pages. This creates problems in evaluation. (b) When due to sudden decision a candidate plans to change or extend any answer or a part of it already written.

It has been frequently observed, by the senior investigators, from the scrutiny of examined scripts that the examiners commit two types of errors: (i) overlook the mistakes of the candidates and award marks in favour of the candidates & (ii) overlook correct answer and deprive the candidates.

Researches on different facets of examination:

Researches on different facets of examination have been extensively undertaken for the development and uses like:

Transformation of Examination into evaluation; Development of type of questions; Framing of questions; Introducing semester, unit test, open book system examination; Increasing validity & reliability of the questions, discrimination and difficulty levels of questions; Choice Based Credit System (CBCS) etc. While wide researches have been reported on examinations, little has been done on the non-sequencing answers of the students in board examinations.

The selection of questions from a question paper in Board examinations is a privilege to the students but to maintain sequence of answers in conformity with the question numbers is important (i) for the convenience and interest of both candidates and examiners, (ii) for reliability of checking answers. Sometimes the students are instructed to leave some pages blank for the answer of a selected question, for attempting later. Authorities of Board examination time to time give instructions to the candidates regarding maintaining sequence of answers. But the present investigators think that only such instructions can't eradicate the maladies of situation. A critical study on this issue is, therefore, necessary to estimate the difficulties of checking the non-sequenced of answers of the candidates and its consequences.

Delimitations of the study:

(i) For the purpose of the study only MCQ test has been used. Stems of the questions have been put serially on a paper in a column but the corresponding groups of options have been placed haphazardly in an adjacent column in order to demonstrate the haphazard answers of the candidates. Examiners (also the respondent of the test) were to correctly identify the group of options commensurate with a stem, and then identify the distractors. Distractors were more than one and similarly correct options could be more than one for some questions.

(ii) Checking errors of an examiner: When an examiner fails to identify errors or distractors in a MCQ by putting cross marks (X) over them, the examiner commits a Checking error.

(iii) Percentage (%) of checking errors of an examiner (\bar{x}) in a test = $\frac{\text{No. of checking errors of an examiner in a test}}{\text{Total no. of errors (distractors) existing in the test}} \times 100$

(iv) Mean of % of checking errors = $\frac{\text{Total } (\bar{x})}{\text{Total no. of respondents}}$

= $\frac{\text{Total \% of checking errors committed by examiners}}{N}$

(v) Standard deviation of % of checking errors = SD of the (\bar{x})

Objectives of the Study:

(i) To compare the number of errors of checking when the answers are sequential and when not.

(ii) To find the correlation between errors of evaluation in sequential and non-sequential answers.

(iii) To find whether the errors committed by the examiners during checking answers depend on their subjects of study in higher education.

(iv) To find whether the errors committed by the examiners during checking answers depend on their genders.

Hypotheses: Since, the literature on the problem selected is severely limited, the investigators made the following action hypotheses on the basis of their professional experience:

H₁: There exists significant difference between the errors arising while examining the sequential and non-sequential answers of a question paper.

H₂: There exists significant correlation between errors occurring while examining the sequential and non-sequential answers.

H₃: The errors committed by the examiners depend upon their subjects of study in higher education.

H₄: The errors committed by the examiners depend upon their genders

Methodology

A set of MCQs was developed writing stems of the questions in one column and their options in an adjacent column. For some of the questions the sequence of answers was kept normal and for some the sequence of answers was broken.

Representation of sequential and non-sequential answers:

48 MCQs type questions were developed on general knowledge of class VIII standard. The stems of the questions were written serially in one column and their group of options (including distractors and correct options) in an adjacent column. Number of distractors varied within 0-3 and correct options within 1-3. If the examiner, fails to identify a distractor with cross(X) mark over it, the examiner commits one checking error and so on. There are total 84 distractors in the entire test. The respondents in this study were examiners having knowledge of evaluation.

48 MCQs were divided into 4 parts (A and A₁) & (B and B₁). Each part has its separate serials. A and A₁ parts are materially same in all respects except the serial numbers. of the stems and the positions of the corresponding distractors. The same techniques were followed in the case of B and B₁. In A the options were written side by side of their stems in two adjacent columns and in A₁ the options were not at all written side by side with their stems but in rather haphazard manner. While checking the answers of A there might be some checking errors. Similarly, there might be some checking errors in A₁. The investigators wanted to see whether there is any significant difference between the numbers of those errors. The items of B&B₁ were framed and used in the same way as A &A₁. The use of B&B₁ serves as triangulation of the case A &A₁.

With respect to a stem, a respondent has to identify the corresponding set of options and put cross over the distractors. It was applied on 5 qualified examiners for finalising number of items and time required for checking the answers. Initially 32 minutes were allotted but 80% of the examiners completed their works in 75% of time allotted. So, 24 minutes were finally allotted for the test.

The question-answer sets were given to the respondents one after another in the sequence A, A₁, B & B₁. Other particular of the items can be had from the table 1.

Table 1. Description of the Tool

Parts of questions	A	A ₁	B	B ₁
No. of Questions	08	08	16	16
Serial No.	1-8	9-16	17-32	33-48
Allotted time	04min.	04min.	08 min	08min.
Maximum checking errors	14	14	28	28

A and A₁ parts are materially same in all respects except the serial numbers. Similarly B & B₁.

The example of the subset (A₁) of the question shown below where the groups of options were not juxtaposed to their respective stems.

Table 2: An example of disorderly arrangement of answers by subset (A₁) of the question

Questions	Disordered options
1. We drink	*2311, 2401, 24001
2. The moon is smaller than	*Akbar, Sersah, Alauddin
3. 49 ² =	*Kerala, Kashmir, Darjeeling
4. Non-Mughal emperor is	*kg, litre, metre, ft , joule
5. Units of length are	*tea, butter, coffee, biscuit
6. Winter garments are abundantly available in	*Observation, opinion, love
7. Synonyms of 'View' are	* 15, 17, 21
8. number straight lines in A/AA/AAA	*stars, satellites, planets

Design of the study:

Population of the study: Any graduate with knowledge in Bengali & English having preliminary experience of testing and evaluation can act as an examiner of the MCQs and hence respondent in the study. Such persons are easily available in Teachers' Training (B.Ed.) colleges.

Sample: A 'convenient sample' of size 38 was selected for the study, who were considered as examiner as well as respondents in the study.

One group pre-test-post-test pre-experimental design was used for the study.

Administration of test:

(i)The nature of the tests and subtests were stated before the respondent cum examiner in the study, (ii) technique of how to identify the errors were explained by demonstration with a sample question paper.

Data Collection: The MCQ test was given to the respondents (also called as examiners) for estimating the errors checking answer papers. In the 1st cycle A was given as pre-test question, A₁ was given in the post test. In the 2nd cycle B & B₁ were given in the same way. Time for checking the answer of each sub-test was properly maintained

The errors of each examiner (respondent) in each part of the questions were counted and converted in terms of % of maximum possible errors in the sub-test.

Presentation of data

Table 3: Frequency of checking errors in each sub-test

Frequency → Classes of checking errors in % ↓	f _A	f _{A₁}	f _B	f _{B₁}
32-39		1		
24-31	1	3	3	2
16-23	4	4	0	3
8-15	12	12	7	10
0-7	21	18	28	23
N=	38	38	38	38

Distributions are highly and positively skewed.

Table 4: Descriptive statistics of evaluation errors

Subtests	A	A ₁	B	B ₁
Checking errors				
Mean % of errors	6.974	9.658	5.421	7.158
Median of % of errors	6.500	8.000	3.000	4.000
SD of % of errors	7.058	9.157	7.504	8.209

Analysis

(i)To find the significance of difference between the errors arising due to checking sequential and non-sequential answers.

Table 6: Paired Sample t- tests

Pair	Mean Checking Errors in %	SD in%	t	df	Sig.
1. A&A ₁	6.974 & 9.658	7.058 & 9.157	2.8	37	S
2. B&B ₁	5.421 & 7.158	7.504 & 8.209	2.6	37	S

S=significant

The difference of errors in A&A₁ are significant 0.05 level.

The difference of errors in B&B₁ are significant 0.05 level.

These two results conclusively show that there exists significant difference between the mean % errors in checking orderly and disorderly answers. So, hypothesis H₁ was retained.

(ii) To find the correlation between errors of evaluation in sequential and non-sequential answers

Table 5: Coefficient of correlation (r) between the number of checking errors committed in sequential & non-sequential answers, df =36

Sub-tests	A ₁	B ₁	Sig.
A	0.771		S
B		0.870	S

S=significant

The numbers of errors in checking orderly written answers in A and disorderly written answers in A₁ are significantly & positively correlated. Similar findings are with for B & B₁. This goes to show that the two errors are positively and significantly correlated. So, hypothesis H₂ is retained.

(iii) To find whether the errors committed by the examiners depend on their subject of study in higher education?

Table 6 a: Effect of the Subjects of Study of the Examiners in Higher Education on Checking errors while checking the Part A₁ (which is jumbled)

Checking Errors →	Mean in A ₁ in %	N	df	t	Significance level at 0.05 level
Subjects ↓			36	2.97	S
Humanities ↓	12.50	26			
Science	4.00	12			

S=significant

Table 6 b: Effect of the Subjects of Study of the Examiners in Higher Education on Checking errors while checking part B₁ (which is jumbled)

Checking Errors →	Mean in B ₁ in %	N	df	t	Significance level at 0.05 level
Subject ↓			36	1.5	NS
Humanities ↓	8.75	26			
Science	5.57	12			

NS= not significant

Table 6a & 6b show no persistent significant difference exists between the number of errors committed by humanities and science-based examiners. Hence the hypothesis H₃ is rejected.

(iv) To find whether the errors committed by the examiners during checking answers depend on their genders?

Table 7a: Effect of Gender of Examiners on Checking errors of answers of A₁

Checking Errors →	Mean in B ₁ in %	N	df	t	Significance level at 0.05 level
Gender ↓			36	1.33	NS
Female	8.33	24			
Male	12.35	14			

Table 7b: Effect of Gender of Examiners on Checking errors of answers of B₁

Checking Errors →	Mean in B ₁ in %	N	df	t	Significance level at 0.05 level
Gender ↓			36	.63	NS
Female	6.96	24			
Male	8.64	14			

Table 7 a & b show no significant difference between checking errors of examiners on the basis of their genders. Hence the hypothesis H₄ is rejected.

Findings:

1. The difference between the number of checking errors in sequential and non-sequential answers are significant.
2. The errors in checking the sequenced and non-sequenced answers are positively and significantly correlated.
3. No persistent significant difference exists between the number of errors committed by humanities and science-based examiners.
4. No significant difference between checking errors of examiners exists on the basis of their genders.

Limitations: (i) The samples of respondent examiners were less. For better results it was necessary to broaden the size of the sample.

(ii) The sample was not drawn randomly.

Discussion:

(i) A study like this is more relevant if executed in real situation. But the examinations conducted in real situation are confidential and highly time consuming.

(ii) The respondent (examiners) on the request of the present investigators, arranged the parts of the question as A, A₁, B & B₁ in the order of increasing difficulty (A being the easiest part where stems and options of MCQ items are laid side by side and errors are less; B₁ is most difficult where number of items, disorder between stems and options, and checking errors are more). But their performance (in terms of the number of checking errors) showed that in order of increasing difficulties the parts should be arranged as B, B₁, A & A₁. Therefore, perceived difficulties of the examiners and practical difficulties are not identical.

In an interview the examiners mostly reported that in the initial stage of checking the answers they were psychologically tensed. That’s why they cut a sorry figure in the so-called easy parts of the question paper. Gradually they got adjusted and performed well even in the relatively difficult parts of the question paper (table 4). This situation frequently tallies with the actual situation of examining answer scripts. In the first instalment of assigned scripts, errors of examiners were more in comparison to latter instalments when they seemed to be conversant with the process.

(iii) Non-sequence answers arise in all types of questions. MCQ-answers are highly sequenced when the answers are recorded on formatted sheets. But when the answers are to be recorded on plain papers of answer books, sometimes the answers are jumbled. Keeping these facts in mind, only MCQ items have been considered for the present investigation. To demonstrate non-sequenced answers, the stems of MCQs and the corresponding groups of options were typed haphazardly on a plane sheet.

(iv) To prepare the question paper, General Knowledge of class VIII standard was considered rather than a particular academic subject, to avoid any peculiar effect due to the particular subject.

(v) The distribution of scores (here the number of checking errors) of the study was bi-modal and for this SD was large in some cases.

(vi) To avoid misplaced answers and to pave the way for reliable checking of answers, spaces for the answers to different questions are to be demarcated on the answer book clearly and sequentially.

(vii) If (vi) is not possible, for some reasons, candidates might be advised to leave sufficient space for each question or the sub-questions so that, in case of need, they can use the space.

Conclusions: The evaluation of the answer scripts by the examiner is one of the crucial factors for justifying the students' performances in academic achievement. This needs special attention with the sufficient time that might minimise the errors in evaluation process. Errors in evaluation are not at all desirable. Many suitable steps have, so far, been taken but the examiners' hours still remain very small. Therefore, further research is needed in this field for minimising the errors in evaluation process of the examiners.

References:

- Harper, A.E. (Jr) & Misra, V.S. (1976). Researches on Examinations in India. NCERT, New Delhi.
- Harper, A.E. (Jr.) & Erika S Harper (1990). Preparing Objective Examinations. Prentice Hall of India, New Delhi.
- Kantowitz, B.H. (2000). Attention and Mental Workload Sage Journals
<https://journals.sagepub.com/doi/pdf>
- NCERT (1987). Third Survey of Researches in Education (1978-83), New Delhi
- NCERT (1991). Fourth Survey of Researches in Education (1983-88), New Delhi
- https://www.mathsgenie.co.uk/resources/65_sequences.pdf (retrieved on 02.05.2024)
- <https://www.quora.com/How-do-I-present-answers-in-an-exam> (retrieved on 02.04.2024)