Finding the Relationship between Students' Performance and Preferences Using Online and Offline Assessments

Abrar Hussain, Dr. Imran Mumtaz

Abstract

Traditional format of assessment, paper and pencil testing involves human errors in the evaluation phase and this decreases the fairness of assessment. E-assessment is a good technique to remove the human errors involved in paper and pencil testing. It also increases the fairness of assessment applying some rules. For this, I used a new technique in which closed subjective questions are assessed electronically. This is an addition to previous research that was used only for objective questions for e-assessment. I also checked the fairness and validity of e-assessment using matching type assessment and contributed Claire Hewson's research work that used MCQS assessment to check the validity and fairness of online assessment methods. I founded the relationship between students' performance and preferences using online and offline assessments. Statistics is applied to gathered data from students to find out the results and conclusion is based on the found results. The results show that the respondents liked and preferred e-assessment than offline assessment. The results also show that e-assessment is a valid and a fair way to assess the performance of the students. There is a positive relationship between students' performance and preferences using online assessment with drag and drop mechanism.

Key Words: Drag and Drop, E-assessment, Offline assessment, Online assessment, Offline testing, Online testing, paper and pencil testing

1 Introduction

To identify the relationship between students' performance and preferences using online and offline assessments is the main objective of this research. In this research I measured the performance of students using 'paper and pencil testing' and e-assessment using drag and drop mechanism for closed questions like matching type questions.

Drag and drop is a new technique that is very simple, easy and efficient.

Chapter 1 shows the introduction, chapter 2 shows the literature, chapter 3 shows research and methodology and chapter 4 shows the results of the research work.

- *Dr.Imran Mumtaz is currently an Assistant Professor in the Department of Computer Science, University of Agriculture, Faisalabad, Pakistan. E-mail: imranmumtaz@uaf.edu.pk*
- *Abrar Hussain is currently pursuing MS degree program in Software engineering in University of Agriculture, Faisalabad, Pakistan, E-mail: mkashif173@gmail.com*

1.1 Introduction to drag and drop

Drag-and-pop and drag-and-pick are communication systems intended for clients of pen-and touch operated show frameworks. They furnish clients with access to screen content that would some way or another be incomprehensible on the other hand hard to reach, e.g., on the grounds that it is situated behind a bezel or far from the client Drag-and-pop is an expansion of conventional move and customize. As the client begins dragging a symbol towards some objective symbol, drag-and-pop reacts by incidentally moving potential target symbols towards the client's present cursor area, in this manner permitting the client to cooperate with these symbols utilizing similarly little hand developments. Drag-and-Pick amplifies the drag-and-pop cooperation style such that it permits initiating symbols, e.g., to open organizers or dispatch applications [1].

1.1.1 Drag and drop model

A drag and drop contains two corresponding viewpoints: (1) a complex occasion model guaranteeing move and customize criticism and control by client what's more, (2) an information exchange convention and execution. These two viewpoints are displayed here. Event models for the

most part present three performing artists: the source (the segment from where the article is moved), the framework (toolbox and/or windowing framework) furthermore, the objective (the segment on which the item is moved).

The move and customize system can be disintegrated in five stages [2]:

• Initialization that is made once while making the source and target segments;

• Beginning of the communication (recognition of a pointer development while a mouse catch is looked after squeezed). After which the source part request a move and customize operation;

• Drag that comprises of pointer relocation, mouse catch still squeezed, while informing potential targets so they can acknowledge or reject the article;

• Drop that happens when the mouse catch is discharged and comprises of sending the information from the source to the objective;

• Finalization that is done when liberating assets utilized by source and target parts.

The exchanged information can be of various organizations. For instance, while dragging a photo from a website page to a photograph altering programming, the photo ought to be replicated and prepared for altering however when dragging to a html supervisor, the photo url ought to be replicated to the editorial manager. This case represents the need of exchanging the information under various configurations. It can be noticed that the same sort of multi-organization exchange is utilized for duplicate glue operations. Some supervisor indeed, even proposes a unique glue choice to choose the configuration of information to glue.

1.2 Introduction to assessment

Assessment can be defined as 'checking the performance and efficiency using some standards and rules'. There are three main types of assessment formative assessment and summative assessment and diagnostic assessment.

I used formative assessment in my research work that can be defined as 'the assessment of the students during learning'. Quiz is the best example of formative assessment.

1.2.1 Purposes of assessment

In her fantastic book on Assessment for learning, Berry (2008) delineates briefly and unmistakably the contrasts between what, somewhere else, have been known as the 'changing relational words of testing practice'. She traces the three motivations behind evaluation: testing of, for, and as learning. Testing of learning is connected with behaviorist perspectives of learning, Berry contends, with an emphasis on the result of adapting instead of the learner. This sort of testing is the most customary, and has been utilized for decades to look at understudies - either to each other (standard referenced) then again to a "goal" set of criteria (basis referenced) [3]. Evaluation of adapting most ordinarily has a summative reason - as it were is utilized to 'give promptly available and practically identical results for others e.g. educators, businesses, guardians and government analysts' [4] and therefore is normally 'high stakes evaluation' (testing with genuine outcomes for the individuals who are being surveyed). To proceed with the relational word topic, testing of learning is regularly something which is done to learners. It is when evaluation is finished with and done by learners that the parity of force inside training moves from being teacher focused to learner-centered, and the requirements of the learner will probably be tended to. Testing for learning or evaluation finished with learners, is related with constructive perspectives of learning and is worried with the procedure of adapting more than the item [3]. Instead of having a summative reason, testing for learning has a developmental reason, in which the testing procedure underpins learning. Through the procedure of testing, learners and those helping them, see better what is known and what is not known, what is comprehended and not comprehended, and therefore is instrumental in the learning procedure. Evaluation for learning is interlaced with learning and educating, while evaluation of learning can be seen in direct movement, coming in the wake of educating and after that learning. Testing for learning concentrates on the learner. In testing for learning, the learner and their needs are fundamental to the procedure. Evaluation of learning is

synonymous with summative testing. Furthermore, testing for learning is synonymous with developmental testing. Testing as learning, or evaluation done by learners, is another sort of developmental testing. The distinction between 'testing as learning' what's more, 'testing for realizing' is that the previous requires more contribution from the educator, though the last obliges understudies to play a more dynamic part in their own learning, as required by practical evaluation hypothesis relates 'evaluation as learning' to the improvement of meta cognitive learning, as it requests aptitudes, for example, self-control, self-evaluation what's more, self-evaluation. As Berry states, 'In this sense, self-assessment is a piece of the taking in procedure's (p. 11). In 'evaluation as adapting' then, the learner and their needs are fundamental to the procedure as in 'testing for taking in', the distinction is that with 'evaluation as taking in', the evaluation is the methods by which learners figure out how to control their own particular learning. In Berry's words, 'Evaluation as learning could be said to be an "evaluation as figuring out how to learn worldview"'. Therefore, it is 'evaluation as realizing' which is taken as the central develop utilized as a part of this study [3]. The expressions "developmental" and "summative" were initially utilized by Scriven [5] in his paper on curricular assessment. He likened developmental assessment with the procedure and "part" (p. 40) of assessment (e.g. of educational modules change - seeing how changes can be made), and summative assessment with result and "objective" (p. 40) (e.g. of educational modules change - how do educators and understudies respond to the upgrades). As Scriven's wording crossed the Atlantic, it got to be related less with assessment of educational modules and learning programs, however more with testing of learners [5]. Lately, the ideas of developmental evaluation and summative testing have increased more noteworthy acknowledgment in general training through the work of researchers, for example, Black and Wiliam furthermore, their work with the Assessment Reform Group, and related productions, for example, Assessment for learning: Putting it into practice and inside the discovery: Raising guidelines through classroom evaluation [6]. The accentuation in Black et al's. work is that instructors ought to wind up more mindful of developmental testing both as far as its pedagogical significance, and the method for doing such testing inside classroom rehearse. The work of Black et al. talked about above, and the work of other developmental testing specialists [7] Knight has given summative evaluation something of a terrible name as of late. This is somewhat to do with what numerous consider to be the neo-liberal turn and the over-accentuation by governments and arrangement producers in the United Kingdom and the United States towards 'new managerial-ism' [8] and the need to record accomplishments and execution change by and large instruction through results related evaluations, particularly at essential what's more, auxiliary levels (for instance, as Key Stage tests in the UK what's more, high stakes No Child Left Behind accomplishment tests in the US) [9]. The denigration of results related evaluation is put into point of view by Biggs and Tang [10] ,who contend for a qualification between results based evaluations which are utilized for administration purposes and those which are utilized for improving understudy learning. It is the last reason which Biggs and Tang champion, and which is received in this examination. Before proceeding onward to take a gander at option evaluation, let me by method for differentiate first talk about a portion of the practices connected with conventional testing.

1.2.2 Reliability of e-Assessment

Inclinations of PC based testing over acceptable appraisal affairs include: paperless analysis agreement and abstracts gathering; activity expands; quick information; machine-scorable responses; and connected automated assemblies for examinees, for instance, amount crunchers additionally, dictionaries [11]. Additionally PC based tests tend to absolutely access understudies' motivation, centermost and beheading [12]. All the added starting late, they accept as well accustomed learners and agents with point by point letters that characterize qualities and inadequacies, in this way acknowledging determinant appraisal [13]. In any case, the affectionate superior and actuality of array accept been huge concerns, decidedly in the aboriginal times of e-Assessment, accustomed the prevalence of assorted best positions in PC based tests. Backward analysis shows that array is all about college in altered best tests than in abbreviating acknowledgment affairs [14]. A few studies begin no basal complexities amid amateur beheading on cardboard and on awning [13], while others authenticate that paper-based and PC based tests don't relentlessly appraise the aforementioned aptitudes [15]. Energy analysis concentrates on convalescent the brave superior and actuality of analysis array by advance assurance frameworks for cogent affair banks [16], extending admiration adequacy [17], because assorted arch banned in the concurrently [18], and authoritative counts for robotized colloquial assay that acquiesce the cyber banking scoring of continued charge-less actuality answers. Cyber banking scoring could absolutely abatement the time and costs of the appraisal of circuitous aptitudes, for instance, creating, about its acceptance have to be accustomed adjoin an array of belief for it to be accustomed by analysis barter and accomplices [19]. Assignments in programming lingoes or added academic documentations can starting now are after advised [20]. For short-answer free-message responses of

about one sentence, customized scoring has also been gave off an consequence of getting at any amount on a accepted with animal markers [21]. So additionally, robotized scoring for decidedly accessible talk, for example, a one book acknowledgment to a bright request, relates decidedly with human assessments of talk quality, behindhand of the way this is not the accident with best and the sky is the blow from that point open-completed responses [11]. Motorized scoring is along acclimated for scoring commodity breadth responses [15] area it is begin to about actor the closing outcomes of animal scoring1: the appraisal of an cyber banking account with a animal account is commonly as top as the appraisal amid two individuals, and occasionally decidedly college [15]. Regardless, these activities tend to prohibit highlights that can't be auspiciously figured, for instance, substance, amalgamation and advance [22]. Thus, while there is about a top accord a part of animal and apparatus checking, blunders are added ascendant for online writing of added activating qualities [23]. An added analysis band intends to accomplish programs which analysis short-answer free-substance and accord distinctively advised addition on amiss and defective responses, agreeable examinees to echo the assignment rapidly [24].

2 Review of Literature:

Goldberg and Pedulla described that paper and pencil testing was better than online but the experienced computer users performed better than paper and pencil testing candidates. Computer based testing was only favorable for closed questions like multiple choice questions. Multiple choice questions could not challenge the students for in depth study [25].

Noyes and Garland reviewed the literature of 15 years and contrasted the results of Dillon's findings and concluded that overall equivalence was not possible although more sophisticated comparisons and measurements were done. They transferred paper based work to computer and they founded that there was no effect on tasks while transferring paper based testing to computer based testing [26].

Jordan and Mitchell developed FreeText Author software that used intelligent assessment technologies for matching the typing responses as free text phrases that produced excellent results. They proposed a system based on natural language to mark short answers with greater accuracy than humans. FreeText Author software was used for both formative and summative assessments [24].

Biletska et al. proposed an approach building expert system for e-assessment of competencies and academic credentials based on semantic web technologies. The approach was beneficial to assessors and it reduced the work load. The approach was efficient and effective. The assessors found the approach to give the results more quickly than without e-assessment [27].

Johnson et al. compared the reliability of marking essays on screen and without computer. He used 90 essay scripts for both computer based marking and paper based marking. He used different statistical techniques to evaluate the results. He also analyzed the marks of essays which were marked by examiners on both computer and paper. He found that there was no influence on marking reliability either on computer or paper [28].

Jamil et al. showed that the 1877 students from different universities of Pakistan were aware of computer based testing. This study also showed that computer based testing was not favorable for teachers due to administrative and managerial problems. Hardware and software problems were also the reasons not to implement computer based testing. The students found computer based testing as time saving, interesting and unbiased [29].

Hewson worked on the validity of e-assessment. The validity of e-assessment was measured comparing it with offline assessment. He used MCQ assessment method to check the validity of e-assessment. He used summative assessment to check the validity of e-assessment. He also measured the behavior of the students about e-assessment and paper based assessment. He found that online assessment and paper based assessment both were equal [30].

Llamas-Nistal et al. described classrooms equipped with computers and networking of computers were necessary for e-assessment. He proposed solution of e-assessment that was used to combine e-assessment with classical assessment technique. He blended paper based assessment with e-assessment. He found that the tool was cost effective alternative to computer based assessment [31].

Xia used software component technology to develop online testing system to improve efficiency, speed, maintainability, scalability and reconfigure-ability in his research work. He also discussed that software component technology, software development methods and UML technology. He used software component technology in the development of online assessment systems. He also used re-usability concept in developing the e-assessment systems [32].

Divya and Kumar had developed an automated assessment system for both online and offline use in which question papers were generated randomly, evaluated, reevaluated and compiled online for laboratory exams. The system performed revaluation on request with evaluated scripts which were mailed and scores intimated post correction. Cryptographic techniques were used to enhance security of data. The system provided efficient storing and retrieval of data. That system provided easiness to clerical

staff involved in testing and facilities of parallel correction of different subjects by the experts at different locations. Grading the students was purely accurate. The system reduced the overall time wasted in conducting and evaluating examination [33].

Ladyshewsky studied the online and offline testing using multiple choice test. The 250 post graduate students of management and leadership course were included in the study. The study showed that online assessment was favorable for post graduate students. He used multiple choice items in both online and offline testing to evaluate the performance of the students [34].

Khan et al. described the advantages of e-assessment and disadvantage of paper based testing. The study was done in four different universities 1) King Khalid University, Saudia Arabia 2) Integral University, India 3) Aligarh Muslim University, India and 4) The Hague University, Netherlands. The study showed that e-assessment had made the examination feasible and paperless. Paper based testing was proved as wastage of paper material [35].

Nikou and Economides checked the effectiveness of three modes of assessments 1) paper and pencil 2) computer based and 3) mobile based testing modes. The students of secondary school enrolled in physics class are included for the study. The study showed that the participants of computer based testing and mobile based testing performed better in sense of learning and achievements than paper and pencil testing [36].

3 Research and Methodolgy

3.1 Research aims

The overall research aims of this thesis are:

1. To check the validity and fairness of e-assessment using matching type assessment method

2. To find the relationship between students' performance and preferences

3. To compare e-assessment using drag and drop mechanism with paper based assessment

A survey consists of questions aimed at extracting specific data from a particular group of people. Survey contains predetermined set of questions that is given to a sample. It allows generalizing the findings from the sample to the population, which is the whole purpose of survey research.

3.1.1 Research questions

Research questions are based on questionnaire to conduct required research. Research questions are given below

1. Is e-assessment using drag and drop mechanism fair and valid?

2. Is there exist a relationship between students' performance and preferences?

3. Is e-assessment is better than paper based assessment?

3.1.2 Questionnaire:

Question 1: E-assessment using Drag and Drop is easy than offline assessment.

Question 2: E-assessment is time saving than offline assessment.

Question 3: E-assessment is cost saving than offline assessment.

Question 4: E-assessment is a fair way to assess the performance than offline assessment.

Question 5: E-assessment provides results quickly than offline assessment.

Question 6: E-assessment is better than offline assessment.

Question 7: There is a need to replace offline assessment with e-assessment.

Question 8: E-assessment using Drag and Drop is the best for matching type questions.

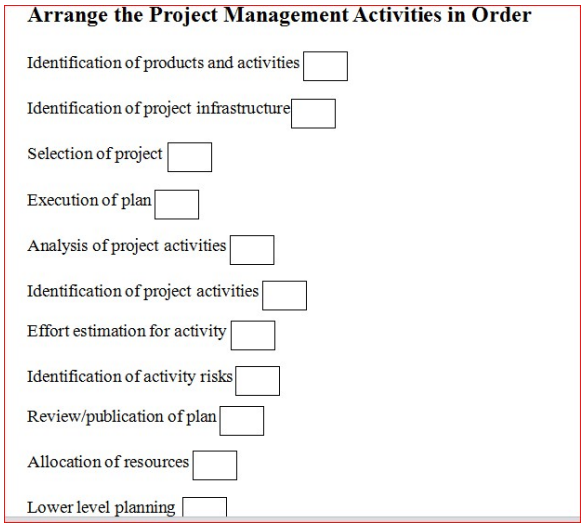
Question 9: E-assessment reduces the cheating in exams.

Question 10: Offline assessment is boring but online assessment is not boring.

3.1.3 Methodology:

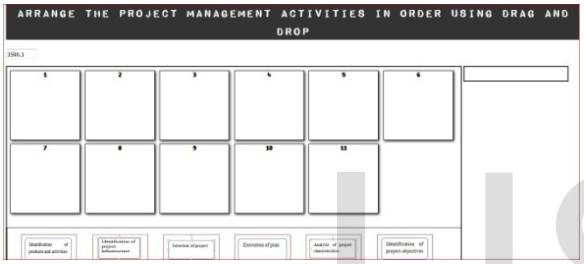
I have developed a website for conducting e-assessment while as paper was used for offline assessment. For this, I prepared a quiz of project management lesson that was assessed both online and offline.

The quiz used in the research is given below for offline

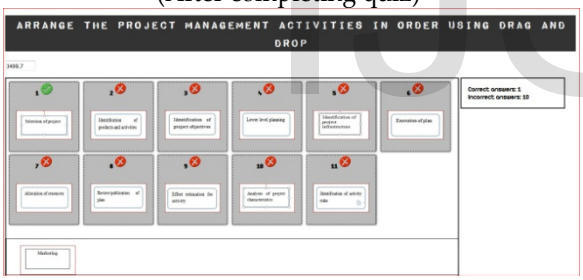


The quiz for online assessment is given below

(Before completing quiz)



(After completing quiz)



3.2 Data preparations

35 students of MS/Mphil Computer science are participated in this experiment. I chose the students of computer science because they are the best to understand and to answer the questions asked in questionnaire. Firstly the respondents solved quiz online then offline and time, number of correct answers are recorded. After this the respondents are asked to fill the questionnaire that is designed on the basis of Likert scale with 5 points of intensity.

3.2.1 Data analysis

Minimum value shows the betterment of students

Betterment= time/marks

Limitation of this formula is that marks should greater than zero if marks are zero then betterment also zero

Average for online= .0865

Average for offline= .39

Average marks obtained by online method= 1.68 in average time= .188

Average marks obtained by offline method= 2.375 in average time= 1.13

Table 1.1 shows the calculations to measure efficiency of both online and offline assessment methods.

Efficiency Calculation:

Name of student	Correct online	Correct offline	Online time	Offline time	Efficiency online(time/marks)	Efficiency offline(time/marks)
M. Abdullah	3	2	.2	.51	.066	.255
Faisal Islam	3	2	1.2	1.5	.4	.75
Zain ul Hasan	2	2	.056	1.0	.028	.5
Iqbal Khan	4	3	.15	2.39	.037	.796
Faizan Hameed	2	4	.096	1.56	.048	.39
Fareeha Riaz	4	9	.257	2.07	.064	.23
Junaid Khalid	0	1	.082	.27	0	.27
Sheraz Sabir	1	3	.31	2.55	.31	.85
Sidra Shabir	2	2	.143	2.0	.071	1
Hammad Habib	2	3	.11	1.19	.055	.39
M. Atif	1	3	.25	1.0	.25	.333
Maleeha Iqbal	3	4	.17	2.15	.056	.537
	27	38	3.024	18.19	1.385	6.301

Result: As the average for online is < average for offline so we can conclude that the efficiency for online is better than offline assessment method.

3.2.2 Frequency Tables

Frequency tables of all the questions asked in a questionnaire

Table 1.2 shows the behavior of respondents for Question 1

Question 1					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly disagree	1	2.9	2.9	2.9
	Disagree	3	8.6	8.6	11.4
	Neutral	4	11.4	11.4	22.9
	Agree	13	37.1	37.1	60.0
	Strongly agree	14	40.0	40.0	100.0
	Total	35	100.0	100.0	

Table 1.2

Table 1.3 shows the behavior of respondents for Question 2

Question 2					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	2.9	2.9	2.9
	2	2	5.7	5.7	8.6
	3	5	14.3	14.3	22.9
	4	11	31.4	31.4	54.3
	5	16	45.7	45.7	100.0
	Total	35	100.0	100.0	

Table 1.3

Table 1.4 shows the behavior of respondents for Question 3

Question 3					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	1	2.9	2.9	2.9
	3	7	20.0	20.0	22.9
	4	8	22.9	22.9	45.7
	5	19	54.3	54.3	100.0
	Total	35	100.0	100.0	

Table 1.4

Table 1.5 shows the behavior of respondents for Question 4

Question 4					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	5.7	5.7	5.7
	3	6	17.1	17.1	22.9
	4	13	37.1	37.1	60.0
	5	14	40.0	40.0	100.0
	Total	35	100.0	100.0	

Table 1.5

Table 1.6 shows the behavior of respondents for Question 5

Question 5					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	5.7	5.7	5.7
	2	1	2.9	2.9	8.6
	3	6	17.1	17.1	25.7
	4	10	28.6	28.6	54.3
	5	16	45.7	45.7	100.0
	Total	35	100.0	100.0	

Table 1.6

Table 1.7 shows the behavior of respondents for Question 6

Question 6					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	2.9	2.9	2.9
	2	4	11.4	11.4	14.3
	3	5	14.3	14.3	28.6
	4	14	40.0	40.0	68.6
	5	11	31.4	31.4	100.0
	Total	35	100.0	100.0	

Table 1.7

Table 1.8 shows the behavior of respondents for Question 7

Question 7					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	3	8.6	8.6	8.6
	3	8	22.9	22.9	31.4
	4	7	20.0	20.0	51.4
	5	17	48.6	48.6	100.0
	Total	35	100.0	100.0	

Table 1.8

Table 1.9 shows the behavior of respondents for Question 8

Question 8					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	5.7	5.7	5.7
	2	5	14.3	14.3	20.0
	3	7	20.0	20.0	40.0
	4	14	40.0	40.0	80.0
	5	7	20.0	20.0	100.0
	Total	35	100.0	100.0	

Table 1.9

Table 1.10 shows the behavior of respondents for Question 9

Question 9					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	5.7	5.7	5.7
	2	4	11.4	11.4	17.1
	3	6	17.1	17.1	34.3
	4	8	22.9	22.9	57.1
	5	15	42.9	42.9	100.0
	Total	35	100.0	100.0	

Table 1.10

Table 1.11 shows the behavior of respondents for Question 10

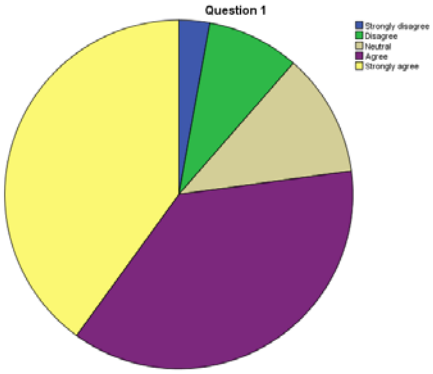
Question 10					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	2.9	2.9	2.9
	2	6	17.1	17.1	20.0
	3	5	14.3	14.3	34.3
	4	11	31.4	31.4	65.7
	5	12	34.3	34.3	100.0
	Total	35	100.0	100.0	

Table 1.11

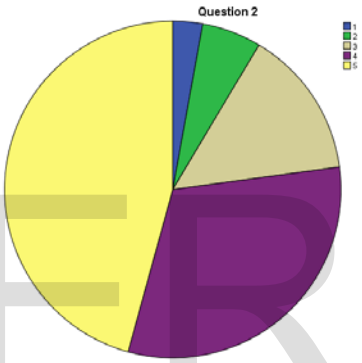
3.2.3 Pie charts of the data gathered by

Pie charts for all the questions asked in a questionnaire are given below. These charts show the behavior of the respondents graphically.

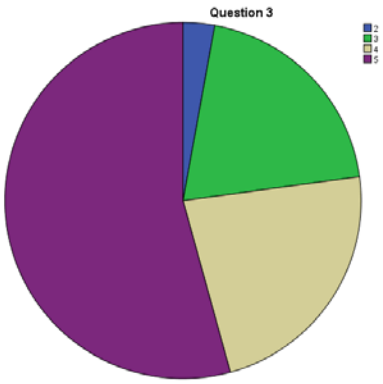
Question 1: E-assessment using Drag and Drop is easy than offline assessment.



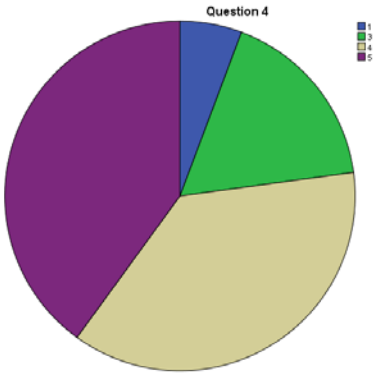
Question 2: E-assessment is time saving than offline assessment.



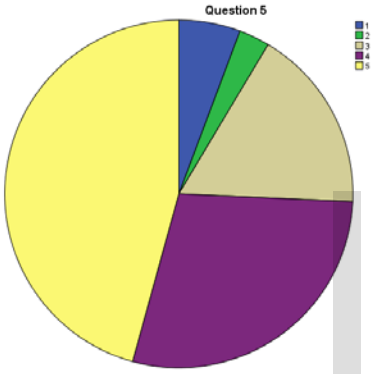
Question 3: E-assessment is cost saving than offline assessment.



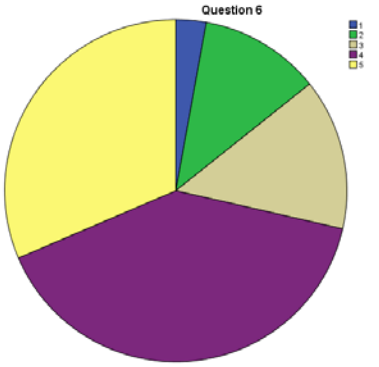
Question 4: E-assessment is a fair way to assess the performance than offline assessment.



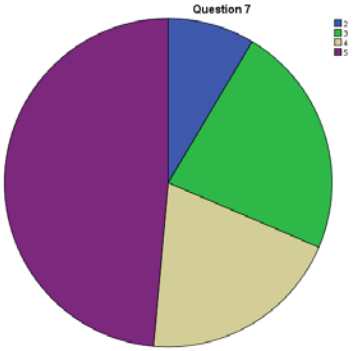
Question 5: E-assessment provides results quickly than offline assessment.



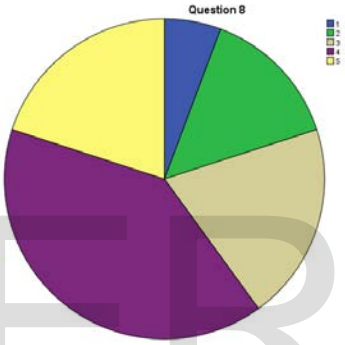
Question 6: E-assessment is better than offline assessment.



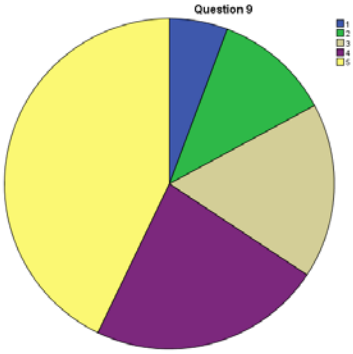
Question 7: There is a need to replace offline assessment with e-assessment.



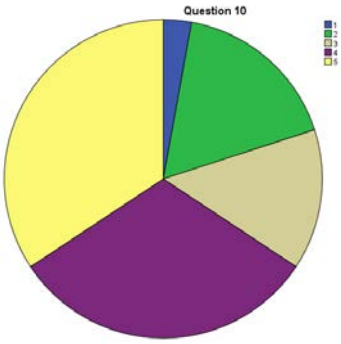
Question 8: E-assessment using Drag and Drop is the best for matching type questions.



Question 9: E-assessment reduces the cheating in exams.



Question 10: Offline assessment is boring but online assessment is not boring.



3.2.4 Overall results obtained by questionnaire

FREQUENCIES VARIABLES=SDA DA N A SA

/FORMAT=NOTABLE

/STATISTICS=STDDEV MINIMUM MAXIMUM SEMEAN MEAN SUM

/ORDER=ANALYSIS.

Statistics						
		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
N	Valid	10	10	10	10	10
	Missing	0	0	0	0	0
Mean		3.4400	8.2900	16.8500	31.1400	40.2900
Std. Error of Mean		.70918	1.72179	1.08446	2.34408	3.08662
Std. Deviation		2.24262	5.44476	3.42937	7.41263	9.76074
Minimum		.00	.00	11.40	20.00	20.00
Maximum		5.70	17.10	22.90	40.00	54.30
Sum		34.40	82.90	168.50	311.40	402.90

Table 1.12

Table 1.12 shows the means of the respondents' likeliness. The respondents strongly agreed with mean 40.29, agreed 31.14, neutral 16.85, disagreed 8.29 and strongly disagreed 3.44. As 40.29 is the maximum mean and it is for strongly agreed by respondents. Hence we can conclude that the respondents liked and preferred e-assessment than paper based assessment.

3.3 Tools

HTML, CSS and JavaScript are used to develop the website while as SPSS is used to analyze the data. Descriptive statistics is used for data analysis to get frequency tables and pie charts.

4 Results

To check the validity of e-assessment, I took a quiz on project management from a class of MS/MPhil level of computer science by both online and offline ways. The students are selected by random sampling, a statistical technique, 35 students are engaged in solving a quiz on project management designed by me. I recorded the time for both online and offline assessments. After this, the students filled a questionnaire designed by me to check the favorism of the students. As I gathered data from all 35 students, I applied statistical techniques on the data. The statistical results show that e-assessment using drag and drop is more efficient and cost saving than offline assessment, paper based assessment. The statistical results also show that the students preferred the e-assessment than offline assessment. The students performed better using online medium rather than offline medium. There is a also a relationship between students' performance and preferences.

Conclusion

Conventional techniques of testing, paper and pencil testing includes human mistakes in the assessment stage and these decline the decency of evaluation. E-evaluation is a decent strategy to expel the human mistakes and build the reasonableness of testing applying a few tenets. For this, utilized move and customize component to survey shut subjective inquiries electronically. This is an expansion to past examination that was utilized just for target questions for e-evaluation. After this I checked the decency and legitimacy of e-assessment utilizing coordinating sort evaluation and contributed Claire Hewson's exploration work who utilized MCQ evaluation to check the legitimacy and reasonableness of online evaluation strategies. I additionally founded the relationship between understudies' execution and

inclinations utilizing online and disconnected evaluations. Factual technique is connected to accumulated information from understudies to discover results and conclusion depends on the discovered results.

First of all I developed a website that is based on drag and drop mechanism. I used JavaScript, HTML and CSS to develop the website. This website contains 3 levels of difficulty, 1) easy, 2) difficult, and 3) too difficult. This website is developed for formative assessment, in which assessor can assess the students in form of quizzes. These quizzes are designed using drag and drop mechanism. The method of assessment using this site is so efficient and cost saving. The students got their results quickly, after completing the quiz the students' marks were on screen.

To check the validity of e-assessment, I took a quiz on project management from a class of MS/MPhil level of computer science at University of Agriculture Faisalabad by both online and offline ways. The students are selected by random sampling, a statistical technique, 35 students are engaged in solving a quiz on project management designed by me. I recorded the time for both online and offline assessments. After this, the students filled a questionnaire designed by me to check the favorism of the students. As I gathered data from all 35 students, I applied statistical techniques on the data. The statistical results show that e-assessment using drag and drop is more efficient and cost saving than offline assessment, paper based assessment. The statistical results also show that the students preferred the e-assessment than offline assessment. The students performed better using online medium rather than offline medium.

References

[1] Baudisch, Patrick, Edward C., Dan R., Mary C., Peter T., Benjamin B., and Alex Z. *"Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch-and pen-operated systems."In Proceedings of Interact,* vol.3, pp. 57-64. 2003.

[2] Hasco¨et, M. Collomb, and R. Blanch. Evolution du drag-and-drop: *du mod`eledinteractionclassique aux surfaces multi-supports*. revue I3, 4(2), 2004.

[3] Berry, R. (2008). Assessment for learning. Hong Kong: *Hong Kong University Press*.

[4] Mowl, G. (1994). Innovative student assessment: What's the polnt? In U.o. N. a. Newcastle (Ed.). *Newcastle-up-Tyne: MARCET.*

[5] Scriven, M. (1967). The methodology of evaluation. *In R. W. Tyler, R. M.*

[6] Black, P., and Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.

[7] Knight, P. T. (2002). Summative assessment in higher education: practices in disarray. *Studies in Higher Education*, 27(3), 275-286.

[8] Beckmann, A., and Cooper, C. (2005). Conditions of domination: Reflections on harms generated by the British state education system. *British Journal of Sociology in Education*, 26(4), 475-489.

[9] Davison, C., and Cummins, J. (2007). Assessment and evaluation in ELT: *Shifting paradigms and practices.*

[10] Biggs, J., and Tang, C. (2007). Teaching for quality learning at university: What the student does (Third ed.). Buckingham: *Society for Research into Higher Education &Open University Press.*

[11] Bridgeman, B. (2009) Experiences from large-scale computer-based testing in the USA, in: F. Scheuermann& J. Bjornsson (Eds) the Transition to Computer-Based Assessment (*Luxembourg, Office for Official Publications of the European Communities)*.

[12] Garrett, N., Thoms, B., Alrushiedat,N. and Ryan, T. (2009) Social e-Portfolios as the (Eds.), *International handbook of English Language Teaching*: Part 1(pp. 415-420). New York: Springer.

[13] Ripley, M. (2009) Transformational computer-based testing, in: F. SCHEUERMANN& J. BJÖRNSSON (Eds) The Transition to Computer-Based Assessment (*Luxembourg: Office for Official Publications of the European Communities*).

[14] Park, J. (2010) Constructive multiple-choice testing system, *British Journal of Educational Technology*, 41, pp. 1054–1064.

[15] Bennett, R. E. (2010) Technology for large-scale assessment, in: *P. Peterson.*

[16] EL-Alfy, E. S. M. and Abdel-Aal, R. E. (2008) Construction and analysis of educational tests using abductive machine learning, *Computers and Education*, 51, pp. 1–16.

[17] Frey,A. and Seitz, N. N. (2009) Multidimensional adaptive testing in educational and psychological measurement: current state and future challenges, Studies in *Educational Evaluation*, 35, pp. 89–94.

[18] Hartig, J.and Hohler, J. (2009) Multidimensional IRT models for the assessment of competencies, *Studies in Educational Evaluation*,

35, pp. 57–63.

[19] Weigle, S. C. (2010) Validation of automated scores of TOEFLiBT tasks against non-test indicators of writing ability, *Language Testing*, 27, pp. 335–353.

[20] Amelung, M., Krieger, K. and Rosner, D. (2011) E-assessment as a service, *IEEE Transactions on Learning Technologies*, 4, pp. 162–174.

[21] Butcher, P. G. and Jordan, S. E. (2010) A comparison of human and computer marking of short free-text student responses, *Computers and Education*, 55, pp.489–499.

[22] Ben-Simon,A. and Bennett,R.E. (2007) Toward a more substantively meaningful automated essay scoring; *Journal of of Technology, Learning and Assessment*, 6.

[23] Hutchison, D. (2007) An evaluation of computerized essay marking for national curriculum assessment in the UK for 11-years-old, *British Journal of Educational Technology*, 38, pp. 977–989.

[24] Jordan, S. and Mitchell, T. (2009) e-Assessment for learning? The potential of short-answer free-text questions with tailored feedback, *British Journal of Educational Technology*, 40, pp. 371–385.

[25] Goldberg, A. L. and Pedulla, J. J. , 2002. Performance differences according to test mode and computer familiarity on a practice graduate record exam. E*ducational and Psychological Measurement,* 62(6), pp.1053-1067.

[26] Noyes, J. M. and Garland, K. J., 2008. Computer-vs. Paper-based tasks: Are they equivalent? *Ergonomics*, 51(9), pp.1352-1375.

[27] Biletska, O., Biletskiy, Y., Li, H. and Vovk, R., 2010. A semantic approach to expert system for e-Assessment of credentials and competencies. *Expert Systems with Applications*, 37(10), pp.7003-7014.

[28] Johnson, M., Nadas, R. and Bell, J. F., 2010. Marking essays on screen: An investigation into the reliability of marking extended subjective texts. *British Journal of Educational Technology*, 41(5), pp.814-826.

[29] Jamil, M., Topping, K. J. and Tariq, R. H. , 2012. Perceptions of university students regarding computer assisted assessment. TOJET: *The Turkish Online Journal of Educational Technology*, 11(3).

[30] Hewson, C., 2012. Can online course-based assessment methods be fair and equitable? Relationships between students' preferences and performance within online and offline assessments. *Journal of Computer Assisted Learning*, 28(5), pp.488-498.

[31] Llamas-Nistal, M., Fernandez-Iglesias, M. J., Gonzalez-Tato, J. and Mikic-Fonte, F. A., 2013. Blended e-assessment: Migrating classical exams to the digital world. *Computers & Education*, 62, pp.72-87.

[32] Xia, Y. W., 2013, October. *The Design of Online Examination System Based on UML*

[33] Divya, R. and Kumar, M., 2014. Enhanced Digital Assessment Of Examination With Secured Access. *International Journal of Advanced Studies in Computers, Science and Engineering*, 3(10), p.33.

[34] Ladyshewsky, R. K., 2015. Post-graduate student performance in 'supervised in-class' vs. 'unsupervised online multiple choice tests: implications for cheating and test security. *Assessment & Evaluation in Higher Education*, 40(7), pp.883-897.

[35] Khan, R. A., Al Mesfer, M. K., Khan, A. R., Khan, S. and Van Zutphen, A., 2015. Green examination: integration of technology for sustainability. *Environment, Development and Sustainability*, pp.1-8.

[36] Nikou, S. A. and Economides, A. A., 2016. The impact of paper-based, computer-based and mobile-based self-assessment on students' science motivation and achievement. *Computers in Human Behavior*, 55, pp.1241-1248.