The Second International Conference for Health and Medical Disciplines (ICHMS)
Karbala (Iraq), 23th December, 2023

Effect of Various Educational Levels on Electronic Grading from the Viewpoints of Science Teachers

Hassan Hadi Al Kazzaz[a], A'ida Abood Al-Haydery[b]

[a]Department of Physical therapy techniques - College of Health and Medical Techniques - University of AL-Zahraa for Women -Karbala -Iraq.
[b] Department of Radiology Technologies- College of Health and Medical Techniques- University of AL-Zahraa for Women-Karbala -Iraq.
* Corresponding author, Email: sabahkareem1977@gmail.com

Received: 28 February 2024; Revised: 07 May 2024; Published: 30 June 2024

Abstract

The school evaluation processes have undergone continuous changes to enhance the educational process towards better efficiency, ultimately achieving high-precision and impartial educational outcomes that ensure justice and equality for all students. The electronic corrector emerged as one of these trends due to its new and distinctive ability to compete with manual correction. Being a modern method, it requires an assessment of its quality and efficiency, The current study aimed to how the various educational stages can affect the electronic grading of achievement tests from the viewpoints of science female teachers. A descriptive design was conducted during the period to this end, seventy-five female teachers of science affiliated with Karbala schools, Iraq, were se-lected, and then put into three subgroups (25 teachers each). Drawing on descriptive methodology, The instrument of the study was constructed a 15-item questionnaire has been designed to analyze the data. This paragraph reports the latest resultsThe results of data analysis were processed using SPSS and a one-way ANOVA. The result of the study showed that the F-computed value hit 5.35, while F-tabulated value hit 3.03. Additionally, differences have been found in the results of the three subgroups following the computation of Tukey-value, and the comparison of that value with the mean total of these subgroups to find the most impactful subgroup. This study, also, finds certain differences in the mean results of both primary and preparatory education were the highest of all the subgroups, having reached 16.76 compared to Tukey calculated value which is 12.24. Primary education subgroup, therefore, has made the biggest impact on the other two non-primary subgroups. The recommendation of the study, electronic grading in both intermediate and post-intermediate, but not primary, levels.

Keywords: Electronic grading, Tests, Educational levels, Assessment, Achievement, Effect, Teachers' viewpoints.

1. Introduction

Recently, the volatile events in today's world, be military, economic, psycho-social, political, or health-related, have created a distrust in the achievement of individuals and their must-perform functions, and owing to these changeable events, co-achievement of tasks has, accordingly, emerged to help finish tasks by humans and machines alike.

Later on, such orientation has been extended to education, whereby relevant decision-makers and stake-holders sought to apply different forms of technology within their settings (Hussein,2019,P.102) Fundamental changes in educational sectors were not only deemed urgent and necessary, but they were also a step forward to catch up with the super-fast progressing technology as well as its unprecedented expansions.

However, it is necessary that this change not only be confined to the curricular, syllabus-related, or methodological structures, but also to the very process of evaluation so as to be in line with the cutting-edge educational approaches. Electronic grading, therefore, has been a capstone of these changes having been viewed as an introducer of equality, zero errors, human-free scaling, and zero favoritism (Ashwin,2010.P.119).

Evaluation has seen highly tremendous developments in the last few years that caused consid-erable changes even in the teacher's role. These developments necessitated, thus, organization which can serve high accuracy, impartiality, equity, and quality. Additionally, these would-be developments should address broader educational sectors. No significant changes can be advanced depending on traditional evaluation standards, student-centered practices, area-restricted aspects, and locally-confined processes(Al-Baghdadi,2008,P.208).

Electronic grading, accordingly, emerged to address these developments (Al-Ghamdi,2021,P. 181).

Like any other newly-introduced and emerging practices, electronic grading was criticized and denied through many doubts and objections as to usefulness, selection over traditional grading practices, and possible teacher-initiated reluctance. Opponents of electronic grading based their objections on personal views, inter-teacher controversies, and level-related differences, rather than on concrete and scientifically-proven foundations.

Notwithstanding all the controversies mentioned-above, the move towards a technology-based, automated, teacher-friendly, and level-compatible evaluation has been a turning-point for any success. Electronic grading, thus, should be seriously considered.

This study, accordingly, draws on the pros and cons of electronic grading, (un)applicability of electronic grading, and how far electronic grading can be encouraged or discouraged, in order to formulate a researchable problem. Given these theoretical premises, this paper formulates a re-search question: How far do various educational levels affect electronic grading of achievement tests from the viewpoints of female teachers of science?

To verify this research question, a generalized hypothesis has been proposed, which stated; there are no statistically significant differences between the primary and intermediate subgroups on one hand and the preparatory subgroups on the other hand in the electronic grading of the achievement tests.

This study, therefore, aims to find how various educational levels affect the electronic grading of achievement tests from the viewpoints of female teachers of science. To this end, three sub-hypotheses have been proposed;

1) There are no statistically significant differences at 0.05 in the mean of primary teachers' answers and the mean of intermediate teachers' answers regarding the electronic grading of achievement tests.

2) There are no statistically significant differences at 0.05 in the mean of primary teachers' answers and the mean preparatory teachers' primary teachers' regarding the electronic grading of achieve-ment tests

3) There are no statistically significant differences at 0.05 in the mean intermediate teachers' pri-mary teachers' and the mean of preparatory teachers' primary teachers' answers regarding the electronic grading of achievement tests.

As for data, it consisted of human samples; female science teachers based in the Karbala schools, Iraq, geo-educational samples; primary, intermediate, preparatory levels, and temporal samples; the school term 2023-2024.

The study, in general, based its key assumptions on the highly important as well as effective roles of evaluation following these must-do practices;

1. Electronic grading is being debated simultaneously with technology-focusing educational direc-tives, and thus, highlighting technological advance in order to keep pace with education-upgrading digitalization in teaching in general and grading in particular.
2. The effects of such evaluation may help stake-holders and decision-makers determine the re-quirements necessary for the proper application of electronic grading.

3. Electrnic grading may help teaching staff advance their IT skills.
4. Research into electronic evaluation may help decision-makers to decide whether to proceed in electronic grading or not. This can be achieved by further research in this area
5. Electronic evaluation in general can be promoted in educational settings through technolo-gy-aware practices.
6. Electronic evaluation can be plainly promoted through the elaboration on advantages and disad-vantages of technology-assisted grading.
7. Electronic evaluation can draw decision-makers' attention to the shortcomings that may arise as teaching staff may not possess IT literacy that render them eligible for that type of grading.
8. Electronic evaluation can indicate to decision-makers how satisfied the teachers are with this type of grading.

2. Methodology

Being preferred in most analytical methodologies, suitable for research procedures, a reliable interpreter of findings, and an effective explainer and analyzer of meta-data (gathering, interpreting, analyzing, tabulating), a descriptive approach has been followed in This study (Amaabed,2021,P.41).

The data is a randomized sample of urban schools based in central Karbala, Iraq, of which 25 female (primary/ intermediate/ preparatory) teachers of science have been selected, as detailed in Table (1) below.

Table (1) Distribution of data

Subgroup	No.
Primary education	25
Intermediate education	25
Preparatory education	25

As far as data gathering is concerned, a 15-item questionnaire has been designed following Likert five-item scale (Al-Ghaith,2021,P.84). These items involved; Strongly agree (5), Agree (4), Neutral (3), Disagree (2), and Strongly disagree (1)

No.	Item	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1	Electronic graders can be easily handled and controlled					
2	My school suffers from cognitive obstacles to electronic grading					
3	My school suffers from physical obstacles to electronic grading					
4	Electronic grading is quality in time management and effort shortening					
5	Electronic grading fosters learning and teaching outcomes accurately					
6	Electronic grading requires hard-to-provide technologies and equipment					
7	Electronic grading requires additional efforts					
8	Electronic grading provides satisfaction for students and parents					
9	Electronic grading will replace traditional, manual grading					
10	Electronic grading provides teaching staff with new technologies					
11	Human-independent grading provides trusted results					
12	Electronic grading fosters impartiality and equity in scores					
13	Electronic grading is compatible with the established curriculi					
14	Electronic grading fits my level-specific tests					
15	Electronic grading is universal for any applicable tests					

Data analysis necessary measure before a given research tool has been applied is to ensure the reliability of that tool (Al-Mandlawi,2016,P.31). To that purpose, the 15-item questionnaire has been reviewed by university professors majored in learning technologies, teaching methods, and pedagogical psychology. The comments, notes, and modifications sent by those experts have added to the significance an development of that questionnaire. Additionally, the questionnaire items, having been tentatively tested as to stability by being replied to and re-replied to by the same 30 female teachers (10 each educational level) with a 14-day period in-between, achieved (0.70) (Bahi,2010,P.182). This tentative test of the questionnaire items showed that stability following the Pearson correlation coefficient was 0.96, which indicates that the questionnaire is valid. To answer the research questions thoroughly, appropriate statistical means have been a must (Al-Anzi,2021,P.382). These means included SPSS, a one-way ANOVA, Pearson correlation factor, and Tukey test.

This study, as it has been elaborated on earlier, attempts to identify how various educational levels affect electronic grading. To fill this gap, a hypothesis has been proposed, stating; there are no statistically significant differences in the primary, intermediate, and preparatory subgroups the electronic grading of achievement tests

The procedure that requires comparing these means was (one-way ANOVA) is the most appropriate for this procedure(Haddou,2019,P.427). it leads to the significance of differences between more than one group, so if the computed (F) value is greater than the tabulated (F) value, it indicates significant differences between them(Al-Rimawi,2017,P.127). And the researcher explained the details of the correction program (Remark Office) for the three groups and clarified its operation mechanism, starting with explaining the concept of electronic correction to them, which is a process of storing students' answer sheets through the scanning process performed by the device connected to the program, then matching what has been stored with the answer keys provided by the program (Ahmed,2022,P.357). Therefore, the (electronic correction process) mostly consists of two main corners, which are: (Al-Jabali,2016,P.132).

First - Scanner: It is a device of various types, shapes, and properties responsible for transferring data with the (scanning) feature, storing it, and sending it later to the second corner of the correction process (the program). Therefore, it is preferred to be characterized by speed, image quality, and storage capacity (Yass,2013,P.126).

Secondly- The electronic program: It is a software application designed with specific specifications that enable it to match students' answers with model templates by retrieving stored data from the (scanner device) and performing the comparison process and extracting the result using algorithms designed for this purpose (Zaytoun,2005,P.155).

Then the researcher addressed the most important types of electronic correction programs that have multiplied and diversified as a result of the competition of the companies that manufacture them, which have made sure that their product is the best, most accurate, and can be accessible to all groups. It is certain that the evaluation is up to the user and not to the producing company, and perhaps the most famous of these programs are (ZipGrade, Datalink, and Remark Office) that the researcher chose within her research(https://edutec4all.medu.s.). This program was developed by (Gravic Company) in the United States of America with a long history in this field. The basic principle of this program is based on the technology of verifying marks through imaging by scanning the answer and entering it into the program, which compares it with the provided answer models and then outputs the models directly(https://remarkomrsoftware.com/ar/blog/autograding).

After the researcher summarized this information for the three samples, she began to explain the working mechanism as follows:

First- Answer sheets are scanned optically through a command issued from the programs to the attached device.

Second- A comparison is made between the collected data and the answer models stored in the program.

Third- After collecting the data, the device compares it with the answer models stored in the program.

Fourth- The program reflects light on specific areas in the answer sheet that are supposed to have specific markings to match what is in the templates of the model answers, so it monitors them.

Fifth- The student must shade or circle the correct answer symbol clearly so that the optical scan can detect it, otherwise the answer will be considered wrong even if it is correct.

Sixth- The results are extracted in the form of a table and stored in the program quickly and accurately (Zidane,2022,P.717).

And the preparation of the question paper is considered part of the (electronic correction) methodology, as it is the basis for its healthy work, and preparing it outside the specifications causes major obstacles in the results. Therefore, it must be comprehensive in the data and consistent in the models. The questions should be modified to create deception or education instead of placing signals. The answer keys should be clear, easy to read and highlight, and emphasize not to pierce the paper. Also, part of the (working methodology) is building model answers and feeding them into programs, obligating correction within the school by a special committee and not taking the papers out of the school. A review committee should also be formed to audit abnormal cases, such as a student changing an answer, leaving part of the shading on the wrong answer, and receiving a (zero), which is then adjusted to a full grade as determined for it(Abdulmajeed,2021,P.72) . Then the researcher began to clarify the negatives of the software and its flaws, whether in manufacturing or programming, which can be overcome with experience and commitment to instructions. Among these flaws is that some complications may occur when compiling data, especially with the abundance of texts or the absence of numbering. Also, it is possible that a portion of the data may be lost or not stored correctly when conducting a scan as a result of jamming answer sheets or entering them in a reverse or distorted manner. In addition, there is an increase in financial cost, limited to specific papers, and correcting it to a specific test pattern. The positive side lies in its speed, accuracy, objectivity, and ease of use, and its results are free from errors, issuing comprehensive reports on the student's level(http://www.accproject.org/main/english/et/et72)

The researcher clarified to the committee the three things that tests should be characterized by when corrected according to the electronic corrector as follows:

First- Define the goal: The first step required of the teacher is to clearly and in detail define the goal to facilitate his understanding of the most important characteristics of the test and its variables, whether for the purpose of distributing students or identifying weaknesses... etc (Naif,2017,P.64).

Second- Define educational levels: The teacher must derive the objectives from the core of the prescribed scientific content and realize the cognitive levels required, whether limited to (knowledge or recall) or should include higher levels (Abdul Raouf,2017,P.126).

Thirdly- Determining the scientific content: The teacher here specifies all the topics and axes presented to the students that he wishes to evaluate by finding the (relative weight) for each axis or part of the specified content.

Fourth- Building (specification table): It is the structure that represents the mechanism of distributing questions according to the curriculum and behavioral purposes.

Fourth- Test formulation: This process is not an easy task; the teacher is required to adhere to the ministry's instructions in terms of structure and the necessity to adhere to the scientific content, and comply with the (specification table) that has been prepared, so it requires alertness and focus from him.

Fifth- Test printing: This is a very important stage as it is the final image of the test that will be shown to the student, so it must be clear in words, free of errors, and of varying difficulty levels. Also, the designated time and final mark for the test must be fixed, along with any other instructions such as permissions and diverse use (Abu Al-Am,2009,P.277).

The researcher has clarified the program in detail for the three groups and is now ready for scientific application.

3. Review of Findings

After she applied the first application and then requested them to implement the correction process. After that, the questionnaire was presented to the groups for the purpose of filling it out, and after returning the questionnaire, the researcher analyzed the questionnaire results using (SPSS) soft-ware. The computed value of (F) was (5.35), while the tabulated value of (F) was (3.03), indicating differences between the research groups, as shown in Table No (2).

Table (2) A one-way ANOVA analysis

Source of variation	Degree of freedom	Sum of squares	Means of squares	F-value		Significance
				Computed	Tabulated	
Inter-subgroup	2	351	175	5.35	3.03	0.05
Intra-subgroup	72	2364	31			
Total	74	27156				

Abdulaziz (2014) indicated that a one-way ANOVA analysis only determines inter-subgroup without any subgroup-specific preference(Abdulaziz,2014,P.105). The study, thus, has to use a Tukey-test which serves that purpose. In the same vein, Tukey-test calculates differences in the inter-subgroup means, then compares these differences with Tukey-test values. If the said differences are greater than means, there are significant differences for the higher means subgroup(Al-Fil,2022,P.41). After the computation of Tukey-test values, the study finds(Accept the first and third hypothesis and reject the second hypothesis as shown in Table No (3).

Table (3) Tukey-Test Calculations

Subgroup	Inter-mean differences	Tukey-tabulated values	Comparison with Tukey-calculations	Significance (0.05)
Primary-Intermediate	17.16-64.24=7.92	12.24	12.24 greater than 7.92	Insignificant
Primary-Preparatory	72.16-55.40=16.76		12.24 less than 16.76	Significant
Intermediate-Preparatory	64.24-55.40=8.84		12.24 greater than 8.84	Insignificant

.

4. Discussion of Results

The results of the statistical analysis showed that the application of electronic correction in the preparatory and intermediate stages is possible, easy to implement, and with few difficulties. The researcher believes that this is due to the skills and knowledge that this age group possesses in electronic transactions, their motivation to use modern technological products, as well as their psychological balance and acceptance of any change in the educational pattern, without fear. Additionally, the ease of communication between the person responsible for this process and the students, and identifying the obstacles that appear during the actual application, as this age group can easily and accurately convey problems. Whereas, the elementary stage itself is a critical and spe-cific stage in terms of integration and interaction within the normal framework, so how about a change they

are not accustomed to Moreover, age groups in elementary education, especially in the early stages, require specific educational practices. Also, it's difficult to evaluate the tests deter-mined at these levels electronically, and the basic requirements of electronic assessment, such as registration papers, which elementary school teachers and students are not accustomed to, may lead to many errors and unsatisfactory results. Additionally, electronic correction may impose specific test patterns that may not be suitable for students at this stage. Furthermore, the transition from manual traditional assessment to electronic assessment in elementary education requires specific material and local requirements, which most elementary schools cannot provide.

4. Conclusions

The most important thing that the researcher concluded through her research is:

1. The implementation of electronic grading has encouraged teaching staff, especially veterans accustomed to traditional teaching methods, to embrace digital education techniques, whether in terms of methods, tests, or grading.
2. Directing the attention of educational decision-makers to the importance of measurement and evaluation at all academic levels, not just at the university
3. There's an increase in the utilization of question banks in achievement tests due to the requirements of electronic grading.
4. Diversity in formulating questions and not relying solely on essay questions.
5. Tests with electronic grading lead to an increase in students' skill in quick responses and time utilization.

4. Recommendations

At the end of the research, the researcher recommends the following:

1- Embedding electronic grading programs in middle and high school education stages to correct achievement tests.
2- Constructing questions according to the (specifications table) and (question bank) to obtain a comprehensive and reliable test.
3- Intensifying courses for teaching staff on the mechanism and programs of electronic grading and to familiarize them with its types and features.
4- Conducting multiple workshops for all educational stages to expand the staff's knowledge of modern measurement and evaluation methods and systems.
5- Adopting other research studies to examine the impact of job service differences on the adoption of electronic grading.

6- Adopting other research studies to examine the impact of professional specialization on the adoption of electronic grading.

References:

[1] Hussein, G.; and Khalid. (2019). Evaluation of journalistic performance. Egypt: Arab Publishing.

[2] Ashwin, Paul; and Lamis. (2010). Changing Higher Education: Evolution of Learning and Teaching. Saudi Arabia: Al-Obeikan Publishing.

[3] Al-Baghdadi, Manar. (2010). Developing Education in Light of Experiences of Some Countries. Egyptian Book House for Publishing.

[4] Al-Ghamdi, Fawzia. (2021). The reality of applying electronic correction and its impact on academic achievement in the subject of Social Studies. Saudi Arabia: International Journal of Educational Sciences, 1(64), pp. 181-226, DOI: 10.21608/IJEPS.2021.242041.

[4] Amaabed, Kamal; Magdy; Ashraf; Mahmoud. (2021). Training program based on integrated learning to develop digital educational material production skills for elementary school teachers. Egypt: Adult Education Journal, 3(3), pp. 41-60, DOI: 10.21608/altc.2021.203500

[6] Al-Ghaith, Amal. (2021). The impact of mathematics teachers' practice in the middle stage of alternative assessment methods in distance education during the COVID-19 pandemic. Palestine: Journal of Educational and Psychological Sciences, 14(5), pp. 84-112, DOI: 10.26389/AJSRP.S221120.

[7] Al-Mandlawi, Nibras. (2016). Using scientific research methods in university theses. Iraq:

Journal of the College of Arts, Al-Mustansiriya University, 73(1), pp. 1-31, DOI: 10.35167/muja.v0i73.220.

[8] Bahi, Mustafa; and Mona. (2010). Glossary of statistical terms. Egypt: Anglo Library Publishing.

[9] Al-Anzi, Abdullah. (2021). The impact of differences in response scoring points in Likert scales on the psychometric properties of psychological measures: The self-efficacy scale as a model. Iraq: Al-Fateh Journal of Educational and Psychological Research, 3(25), pp. 382-408, DOI: 10.23813/FA/87/15.

[10] Haddou, Adel. (2019). Using one-way multivariate analysis of variance in the impact of database on the areas of banking services in some Iraqi banks. Iraq: Journal of Administration and Economics, University of Baghdad, 112(25), pp. 427-441 DOI: 10.33095/jeas.v25i112.1671.

[11] Al-Rimawi, Omar. (2017). Building and designing psychological tests and measures.

[12] Ahmed, Mustafa. (2020). Education Renaissance Strategy, PMBIC Publishing.

[13] Al-Jabali, Hamza. (2016). E-learning is an introduction to computing, Dar Al-Ilm Publishing

[14] Yass, Najlaa. (2013). Digitization and its techniques in Arab libraries. Egypt: Al-Arabi for Publishing.

[15] Zaytoun, Hassan. (2005). A New Vision in E-Learning, Dar Al-Nashr.

[16] School Facilities Management in the Saudi Ministry of Education, 2023, https://edutec4all.medu.sa, 6/31/2023.

[17] American Graphics Company, 2023, https://remarkomrsoftware.com/ar/blog/autograding 7/22/2023

[18] Zidane, Emad. (2022). The Impact of the Auto-Correction Program (Remark Office) on Developing Courses in the Department of Libraries and Information, Egypt: Journal of Humanities and Literary Studies, 26(4), pp. 717-753 DOI: 10.21608/shak.2022.247665,

[19] Abdulmajeed, Asim et al. (2021) Question Banks and Electronic Correction. Qatar: Dar Al-Watad for Publishing.

[20] Green, P. (2000). Optical Scanning System. Retrieved April 24, 2012 from: http://www.accproject.org/main/english/et/et72

[21] Naif, Amatanius. (2017). Constructing Psychological and Educational Tests and Scales and Their Techniques. Jordan: Dar Al-A'sar for Publishing.

[22] Abdul Raouf, Tarek; and others. (2017). Scales and Tests: Design, Preparation. Jordan: Al-Tantheem Arab Group for Publishing.

[23] Abu Al-Am, Raja. (2009). Measurement and Evaluation in Humanities. Egypt: Dar Al-Kitab Al-Hadith for Publishing.

[24] Abdul Aziz, Barkat. (2014). Introduction to Statistical Analysis. Egypt: Dar Al-Masriah Al-Lubnaniah for Publishing

[25] Al-Fil, Helmi. (2022). Statistical Data Analysis Using SPSS "Theorizing, Application, Interpretation. Egypt: Scientific Journal, Faculty of Education, Alexandria University, 42(14), pp. 182-196, DOI: 10.21608/SJSW.2022.246039.