



Developing And Standardizing A Psychometric Scale to Evaluate Professional And Teaching Performance of Physical Education Teachers at Open Education College In Light of Academic Accreditation Standards

¹Falah Kareem Jeleeb Muhammad*

*Corresponding Author, E-mail: Falahkrarem@gmail.com

¹Open Educational College, Al-Shatra Branch

Abstract

The study aimed to develop and standardize a scale for evaluating performance of male and female physical education teachers during lessons, to identifying level of performance of male and female physical education teachers during lesson by building a scale to evaluate this performance. Research problem was that current scale does not objectively evaluate performance of male and female teachers, as it relies on limited items within an annual teacher evaluation form whose items do not cover all aspects of performance evaluation during lesson. The researcher used descriptive approach with a survey method to suit nature of problem, in addition to describing research community and its sample and determining axes of scale, which included four axes, in light of opinions of experts and specialists. (59) items were identified and presented to a group of experts to show their validity. Exploratory experiment was carried out, main experiment construction experiment. Statistical analyses were also carried out for items of scale in terms of their discriminatory ability, internal consistency, and scientific basis of scale. Scale was standardized so that data could be statistically processed, standards derived, and levels determined. Results showed performance level of male and female physical education teachers at Open Educational College that it was an unacceptable level. Performance level of female physical education teachers during lesson is better than performance level of male teachers in all aspects of study. most important recommendation was to use current scale in evaluating performance of physical education teachers.

Keywords: Psychometric Scale, Professional Performance, Teaching Performance, Physical Education Teachers.

Introduction

scientific and technological progress witnessed in modern era has cast its shadow on various fields of life and brought about significant changes and developments. Main pillar of this progress and development is man, with enormous and limitless energies and capabilities, if means to advance and develop are provided. Caring for human capabilities is one of important principles upon which modern education in current era is based.

Physical education is one of these fields that has benefited greatly from this development, due to its effective role in development in its various dimensions. physical education teacher is one of most prominent figures through whom educational goals are achieved, due to his great influence in raising students and preparing them in a way that makes them capable of serving their country with dedication, sincerity, and high enthusiasm.

Physical education has a status and importance no less than other curriculum lessons. Through it, many educational and learning goals can be achieved through various activities, whether physical, skill-based, or social, which will greatly contribute to preparing and equipping students in a way that enables them to serve their country and nation with high efficiency. In

order to prepare a successful physical education teacher capable of achieving educational goals, it is necessary for him to master a set of performance skills that together constitute ability to implement different teaching methods. (Marwan, 2001, 22)

Teacher's performance includes a set of teaching skills that require continuous practice and mastery to ensure their contribution to bringing about desired change in students' behavior. There is no doubt that evaluation is a fundamental pillar of educational process, therefore, evaluating performance of male and female teachers during lesson is one of important and necessary requirements for improving it.

Importance of research is evident through building and standardizing a scale to evaluate performance of male and female physical education teachers during lesson. This enables us to identify this level and determine its strengths and weaknesses, thus diagnosing and correcting errors to improve and develop teacher's performance in a way that serves educational process and contributes to advancing it, and to provide those concerned with school sports affairs with a clear picture of performance of male and female teachers during physical education lesson in order to take necessary measures to develop and improve it.

Contemporary higher education is undergoing rapid qualitative transformations imposed by requirements of quality and academic accreditation. This necessitated that educational institutions review their administrative and teaching systems to align with established standards, ensuring efficient performance and achieving desired educational outcomes. Open Education College is one such academic institution striving to enhance its institutional and professional performance, particularly in field of training physical education teachers.

This specialization is unique because it combines theoretical and practical aspects. By examining field situation and approved performance evaluation mechanisms It is observed that evaluation of professional and teaching performance of physical education teachers is often conducted using traditional tools or general forms that are not based on a rigorous scientific psychometric framework, nor do they directly align with academic accreditation standards adopted by higher education institutions. Furthermore, these tools may lack necessary psychometric properties of validity, reliability, and objectivity, thus limiting accuracy of their results and weakening their reliability in making developmental decisions.

Research problem lies in absence of a standardized psychometric scale, developed according to sound scientific principles, that measures professional and teaching performance of physical education teachers at Open Education College in light of academic accreditation standards and objectively and comprehensively reflects their actual performance.

Research Objectives to developing and standardizing a performance evaluation scale for physical education teachers during lessons at College of Education. Identifying performance level of physical education teachers during lesson.

Research hypothesis there are differences in level of performance during lesson between male and female physical education teachers.

Study terms

Performance Level: Analyzing an individual's performance in all aspects related to them, including psychological, physical, technical, intellectual, and behavioral qualities. (Shawqi, 1992, 15)

Performance Evaluation: A periodic process aimed at measuring strengths and weaknesses in teacher's efforts and behaviors during a physical education lesson to achieve educational and pedagogical goals. (Mu'ayyad, 2000, 167).

Research Methodology

the researcher used descriptive approach with a survey method because it suited nature of research problem. Information and data were collected on research community, which consists of male and female physical education teachers in southern governorates, in cooperation with Educational Planning Department in General Directorate of Education and

Directorate of Sports and Scouting Activity. Number of male and female teachers reached (362), representing total research community, as shown in Table (1).

Table 1. Number of individuals in research population and sample is shown according to variables of gender and experience.

| Sex | Categories (Years of Service) | | | | | | | | | | | | Total | | |
|--------|-------------------------------|---------|-------|------------|---------|-------|-------------|---------|-------|--------------|---------|-------|--------|---------|-------|
| | 1-5 years | | | 6-10 years | | | 11-15 years | | | 16 and above | | | Sample | Society | % |
| | Sample | Society | % | Sample | Society | % | Sample | Society | % | Sample | Society | % | | | |
| Male | 50 | 76 | 13.81 | 40 | 16 | 11.04 | 35 | 52 | 9.66 | 42 | 63 | 11.60 | 167 | 252 | 46.13 |
| Female | 21 | 32 | 5.80 | 16 | 23 | 4.41 | 19 | 29 | 5.24 | 18 | 26 | 4.97 | 74 | 110 | 20.44 |
| Total | 71 | 108 | 19.61 | 56 | 84 | 15.46 | 54 | 81 | 14.91 | 60 | 89 | 16.57 | 241 | 362 | 66.57 |

Table 2. Number of individuals in research sample and their percentage are shown

| Categories | Construction sample | | Standardization Sample | | Sex |
|--------------|---------------------|-------|------------------------|-------|-------------------------|
| | Number | Ratio | Number | Ratio | |
| 1-5 | 22 | 6.07 | 28 | 7.73 | Teacher |
| | 9 | 2.48 | 12 | 3.31 | Female Teacher |
| 6 – 10 | 17 | 4.69 | 23 | 6.35 | Teacher |
| | 7 | 1.93 | 9 | 2.48 | Female Teacher |
| 11 - 15 | 15 | 4.14 | 20 | 5.52 | Teacher |
| | 8 | 2.20 | 11 | 3.03 | Female Teacher |
| 16 and above | 18 | 4.97 | 24 | 6.62 | Teacher |
| | 8 | 2.20 | 10 | 2.76 | Teacher |
| Total | 104 | 28.72 | 137 | 37.84 | Male and Female teacher |

Field research procedures

Steps for scale building: In order to achieve first objective of research, which is to build a scale for evaluating performance of physical education teachers during lesson, the researcher followed following steps:

Defining scale's axes: To develop scale, criteria for evaluating teachers' performance during physical education lessons must be identified. After reviewing numerous studies and literature in fields of teaching methods, testing and measurement, educational psychology, and sports psychology, the researcher identified four criteria for evaluating performance during physical education lessons: Teacher's personality focus. Administrative and organizational axis. Educational axis. Applied axis.

the researcher prepared and distributed questionnaire forms to survey opinions of a group of experts and specialists in field of teaching methods, testing and measurement, general psychology and sports psychology regarding validity of axes and addition, deletion or modification of any axis. After collecting and processing data, the researcher used (Chi^{-2}) test to identify valid axes from others. results showed validity of three axes and rejection of one axis, which is applied axis, because calculated value of (Chi^{-2}) was less than its tabulated value of (3.84) at a degree of freedom of (1) and a significance level of (0.05), as shown in Table (3).

Table 3. It demonstrates validity of proposed axes

| No. | Proposed themes | Validity | | Chi ⁻² value | Sig. type |
|-----|-----------------------------------|----------|--------------|-------------------------|-----------|
| | | Suitable | Not suitable | | |
| 1 | Teacher's personality | 13 | zero | 13 | Sig. |
| 2 | Administrative and organizational | 13 | zero | 13 | Sig. |
| 3 | Educational | 13 | zero | 13 | Sig. |
| 4 | Applied | 8 | 5 | 0.692 | Insig. |

In light of results achieved, psychological axis was added after it was suggested by one of experts. It was then presented to remaining experts and gained their approval, so axes became as follows: Teacher's personality focus. Administrative and organizational axis. Educational axis. Psychological axis.

Determining relative importance and percentage of study's axes

Determining relative importance of study's axes is extremely useful for understanding whether one axis is more important than others in evaluating performance of physical education teachers during lessons, and for determining appropriate number of items for each axis based on its relative importance. the researcher determined relative importance of each axis, which was presented in a questionnaire using a five-point scale (1–5) It indicates lowest and highest degree of importance, according to a group of experts and specialists. In field of teaching methods, general psychology, testing and measurement, and sports psychology, and in light of results of opinions of experts and specialists, relative importance of each axis of study was determined, as shown in Table (4).

Table 4. It shows relative importance of study's themes according to opinions of experts and specialists.

| No. | Axes | Total axis scores | Relative importance |
|-------|-----------------------------------|-------------------|---------------------|
| 1 | Administrative and organizational | 39 | 86.66 % |
| 2 | Educational | 37 | 82.22 % |
| 3 | Teacher's personality | 34 | 75.55 % |
| 4 | Psychological | 33 | 73.33 % |
| Total | | 143 | 317.76 % |

Table 5. It shows percentage of relative importance and number of items for each axis of study

| No. | Axes | Relative importance | Percentage | Paragraphs number |
|-----|------|---------------------|------------|-------------------|
|-----|------|---------------------|------------|-------------------|

| | | | | |
|---|-----------------------------------|-------|---------|----|
| 1 | Administrative and organizational | 86.66 | 27.27 % | 16 |
| 2 | Educational | 82.22 | 25.87 % | 15 |
| 3 | Teacher's personality | 75.55 | 23.77 % | 14 |
| 4 | Psychological | 73.33 | 23.07 % | 14 |

Exploratory application of scale

the researcher conducted a pilot study on a sample of male and female physical education teachers at Open Education College after scale was distributed to observers to monitor teachers. sample consisted of (15) male and female teachers, representing a percentage of (4.14) of total research population, with 9 male and 6 female teachers, for period from 5/12/2025 to 15/12/2025. Purpose of study was as follows: Determining clarity of scale items for observers. Knowing extent to which items of scale cover aspects of evaluating performance of male and female teachers during a physical education lesson.

scale instructions are clear. possibility of observing items of scale within time allotted for physical education lesson. scale items were linguistically sound, easy to understand, and accurately formulated. It became clear from this that all items were understandable, clear, observable, and measurable during physical education lesson.

Main Experiment

After completing pilot application procedures, the researcher proceeded to conduct main experiment construction experiment by applying scale to construction sample, which consisted of (104) male and female teachers, with (72) male teachers and (32) female teachers, who constituted a percentage of 28.72% of total research population for period from 20/12/2025 to 30/12/2025. It should be noted here that observers, assessments are recorded on scale items through their field visits to male and female teachers and their observation during physical education lesson. After collecting forms, data for individuals of construction sample were entered using scale scoring key tool by which examiner reveals answers that indicate presence of result that is measured. (Amer, 2002, 54) and arranging them in tables in preparation for their statistical analysis.

Statistical analysis of performance evaluation scale items

Statistical analysis of scale items is an important and necessary requirement in process of building it in order to obtain good items that achieve goal of building scale. the researcher followed following procedures after process of correcting forms and emptying their data.

Calculating scientific basis of scale

Validity: Validity is an important criterion for test quality, as it indicates degree of accuracy with which a measurement tool measures trait or phenomenon it was designed to measure. (Muhammad, 177 2006) And there are many types of truth, and the researcher used following methods to verify truth of his scale.

Face Validity: A test is valid if its appearance indicates that it is a valid test, such as if its form is reasonable and its items indicate their connection to measured behavior. This is not true validity as most studies and research indicate, but rather formal validity. apparent validity of scale was verified by presenting its axes and items to a group of experts and specialists to determine its validity.

Content Validity: Content validity aims to determine extent to which test or scale represents aspects of trait, characteristic, or ability to be measured, and whether test or scale measures a specific aspect of this phenomenon or measures all of it. (Muhammad, 1988, 258). To verify this type of validity, the researcher presented scale's axes to a group of experts and specialists to determine their validity, calculate relative importance of each area, and

demonstrate their ability to measure performance during lesson. Formulated scale items precisely and presented them to another group of experts and specialists to demonstrate their validity and their connection to axis they represent, items that did not obtain acceptable percentage of agreement from experts were excluded.

Construct Validity: Construct validity is also called hypothetical construct validity, and it refers to ability of a scale to measure a hypothetical construct, psychological concept, or specific trait. (Muhammad, 1994, p. 98)

First: two-group approach: Ability of scale items to distinguish between individuals with higher and lower levels is evidence of construct's validity. the researcher verified this by calculating discriminatory power of scale items using two extreme groups, where weak items were excluded and items with high discriminatory power were retained.

Second: Internal consistency: the researcher verified validity of internal consistency through following: Calculating correlation coefficient between item score and total score of sub-axis to which it belongs. Calculating correlation coefficient between sub-axis score and total scale score.

Stability: the researcher followed these procedures:

First: Half division: split-half test is an indicator of internal consistency of test. A sub-score is obtained for each of two halves, and then correlation coefficient between these two halves is calculated. correlation coefficient is an indicator of reliability of half of test, which is corrected using Spearman-Brown formula. the researcher divided items of sub-axes of scale into two halves. first half included items with odd numbers, and second half included items with even numbers. After that, simple Pearson correlation coefficient was extracted for sum of scores of two halves of test. Since these values represent reliability coefficients of half of test for sub-axes, the researcher corrected them using Spearman-Brown equation to obtain reliability of test as a whole, as shown in Table (6).

Table 6. It shows reliability coefficients of split-half test with correction factor

| No. | Sub-themes | Stability before correction | Stability after correction |
|-----|-----------------------------------|-----------------------------|----------------------------|
| 1 | Administrative and organizational | 0.631 | 0.77 |
| 2 | Educational | 0.753 | 0.85 |
| 3 | Teacher's personality | 0.672 | 0.80 |
| 4 | Psychological | 0.789 | 0.88 |

Second: Cronbach's Alpha Equation: This method is used to extract reliability coefficient in cases where rating is related to type of response of sample members, which falls within multiple levels, such as response to scale items on a rating scale such as (very good - good - average - weak - very weak). Cronbach found that this coefficient is an indicator of equivalence, i.e., it gives good estimated values for equivalence coefficient, in addition to internal consistency or homogeneity. (Salah, 2000, pp. 165-166) reliability coefficient was calculated using Cronbach's alpha equation, and results are as shown in Table (7).

Table 7. It shows test reliability coefficients for scale domains using Cronbach's alpha equation.

| No. | Axis | Stability coefficient |
|-----|-----------------------------------|-----------------------|
| 1 | Administrative and organizational | 0.81 |
| 2 | Educational | 0.76 |
| 3 | Teacher's personality | 0.75 |
| 4 | Psychological | 0.82 |

Third: Standard error: standard error of measurement is a fundamental factor in estimating and interpreting test and measurement results ' as it is closely related to concept of reliability (Salah, 2000, p. 179). It is an accepted scientific fact that it is impossible to assert that a study is free from errors, as every work is accompanied by a set of errors, even if to a small degree, as a result of lack of complete control and precise control of test situations due to presence of many variables that affect these situations, in addition to errors in application and observation (Abdul Rahman, 1985, 41). standard error was calculated and results were as shown in Table (8).

Table 8. It shows standard errors for subdomains of scale

| No. | Axes | Standard deviation | Retail stability coefficient | Standard error of segmentation | Alpha coefficient of stability | Standard error |
|-----|-----------------------------------|--------------------|------------------------------|--------------------------------|--------------------------------|----------------|
| 1 | Administrative and organizational | 1.23 | 0.77 | 0.589 | 0.81 | 0.536 |
| 2 | Educational | 1.21 | 0.86 | 0.452 | 0.76 | 0.592 |
| 3 | Teacher's personality | 1.21 | 0.80 | 0.541 | 0.75 | 0.605 |
| 4 | Psychological | 1.14 | 0.88 | 0.394 | 0.82 | 0.484 |

Objectivity of scale

Objectivity is defined as degree of agreement among those evaluating score. Since observation form prepared by the researcher is primary measurement tool in this study, based on observers' ratings of its items, the researcher had to ensure objectivity of scale, i.e., consistency of observers' ratings. To verify this, the researcher photographed four teachers and arranged for several observers to observe each teacher individually and record their ratings on observation form. After collecting forms and extracting data, the researcher used Kendall's coefficient of agreement, a correlation coefficient used to calculate agreement of experts. This coefficient is used to identify relationship between three or more groups of ranks that are established and ranked based on ratings of three or more observers of a group of individuals according to a specific characteristic or trait. (Muhammad Nasr, 2002, p. 340) value of Kendall's coefficient of agreement between observers reached (0.84), which gives a good indication of objectivity of scale.

Standardization of scale

Standardizing scale means construction, correction, and interpretation of test or measurement instrument results must be based on specific rules that unify and precisely define test materials, method of administration, instructions for answering, and method of correction or recording scores. (Salah al-Din, 29, 2000) A standardized scale is a scale that has specific and clear instructions regarding its application and method of calculating its results. It has been applied to a sample that is representative of original population of research, and if it is applied to another sample similar to construction sample, it gives same results or almost similar results. (Ahmed Mohamed, 1996, 37)

the researcher followed these steps in standardizing scale, which represents final stage of its development.

Sample Standardization

standardization sample included male and female physical education teachers in Open Educational College in southern governorates, totaling (137) male and female teachers, with (95) male teachers and (42) female teachers, and they constituted a percentage of 37.84 % of total research population.

Applying scale

the researcher began applying scale form consists of (46) items on individuals of rationing sample during period between 1/1/2026 and 20/1/2026 according to same previous conditions and instructions. After collecting forms, data for individuals of rationing sample was extracted and arranged in tables in preparation for its statistical analysis.

Statistical methods

the researcher relied on statistical package SPSS to extract following statistical methods: mean. Standard deviation. Skewness Coefficient. Independent Samples Test (t) test. Pearson's simple correlation coefficient. Alpha - Cronbach's equation. modified standard score (T. SCORE). Chi⁻² test. Spearman - Brown equation.

Results and Discussing

In order to clarify information that the researcher extracted, he presented it in form of tables, then analyzed it and discussed it immediately after presenting it.

Presenting, analyzing, and discussing performance level results of physical education teachers

Table 9. It shows means and standard deviations of performance level of male and female physical education teachers.

| No. | Study areas | Mean | Standard deviation |
|-----|-----------------------------------|-------|--------------------|
| 1 | Administrative and organizational | 49.97 | 6.11 |
| 2 | Educational | 37.07 | 3.01 |
| 3 | Teacher's personality | 39.95 | 3.88 |
| 4 | Psychological | 39.30 | 3.04 |
| 5 | Total | 166.3 | 16.05 |
| 6 | Overall score on scale | 230 | |

Mean shows performance level of male and female physical education teachers at Open Education College. By examining Table (9), we observe that arithmetic mean of research sample of male and female teachers was (166.3) and its standard deviation was (16.05), indicating that performance level of physical education teachers is average, as score of 166.3 corresponds to benchmark score of (49.650), which falls within average level. This level is unacceptable and below expectations, considering that physical education is one of important subjects through which physical education teachers can achieve multiple objectives. Therefore, they must be at a high level of performance to achieve desired goals of physical education at this important stage of students' academic lives.

Most physical education teachers possess natural aptitudes and qualities that exceed their job requirements. several obstacles hinder them, including a lack of experience, nature of curriculum and teaching methods, an inability to handle administrative matters, declining health, pressure to obtain awards, and limited resources. All these impediments prevent them from fully utilizing their skills and natural abilities. Therefore, understanding these factors and overcoming them through situational analysis, identifying available resources, preparing necessary tools for implementation, providing them with latest developments in their field, and making concerted efforts to support them will constitute a means to achieve this goal. Useful for improving their performance and thus achieving best results using latest methods. (Mohamed Daoud, 42, 1994)

From this we can deduce that physical education teachers are not interested in developing their teaching skills, nor are they keeping up with and familiar with modern teaching methods and techniques. the researcher attributes this to lack of a real and tangible incentive that drives teacher to strive to achieve a high level of performance during lesson, such as

providing necessary and essential equipment and tools to produce a successful physical education lesson and preparing special squares and fields for practicing various and multiple activities, in addition to lack of financial and moral incentives such as financial allocations that help teacher to perform his work satisfactorily, absence of academic promotion and letters of appreciation, in addition to neglecting physical education lesson and not giving it importance it deserves like rest of curriculum lessons, especially by school principals. All of this has led to a decrease in desire and motivation of physical education teachers, and consequently it has been negatively reflected in their level of performance during physical education lesson.

Presenting, analyzing, and discussing results of differences in performance levels between male and female physical education teachers according to study's axes:

In order to identify differences in performance level between male and female physical education teachers, according to study's axes, following was highlighted in Table (10).

Table 10. Study presents values of t - test among male and female physical education teachers according to study's themes.

| No. | Study areas | Male teachers | | Female teachers | | (t) value | Sig. type |
|-----|-----------------------------------|---------------|------|-----------------|------|-----------|-----------|
| | | M. | St.d | M. | St.d | | |
| 1 | Administrative and organizational | 48.84 | 5.43 | 51.11 | 7.31 | 2.04 | Sig. |
| 2 | Educational | 35.42 | 2.43 | 38.73 | 3.45 | 6.49 | Sig. |
| 3 | Teacher's personality | 38.41 | 4.08 | 41.5 | 3.05 | 6.16 | Sig. |
| 4 | Psychological | 38.40 | 3.09 | 40.19 | 4.97 | 3.65 | Sig. |

value of (t) at degrees of freedom ($n+n- 2$) = 153 and a significance level of 0.05 = 1.980

results of (t) test between male and female teachers indicate that there are significant differences in level of performance between male and female teachers according to study axes, as calculated t-value is greater than tabulated value of (1.980) at a degree of freedom of (135) and a significance level of (0.05), which indicates significance of differences in level of performance between male and female teachers, in favor of higher arithmetic means female physical education teachers.

Presenting, analyzing, and discussing results of determining performance levels for physical education teachers

To determine performance evaluation scale levels for (137) male and female teachers in standardization sample, the researcher used normal distribution curve, which is a theoretical distribution for collected data and is based on premise that different traits, characteristics, and abilities are distributed normally among group of people, and that extreme values of this characteristic appear among people in a limited way, while people are concentrated in middle. (Muhammad Hassan, 1988 , p. 145)

Six standard levels were identified that occupied space under curve and were distributed to right and left of mean in different proportions, and Figure (1) illustrates this.

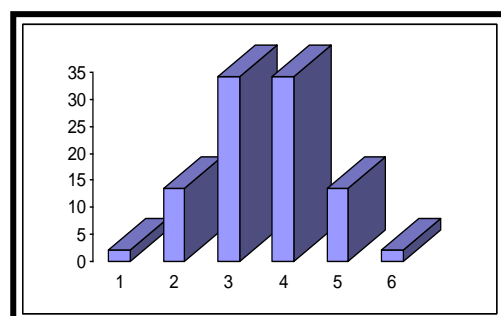


Figure 1. It shows percentages of standard levels

levels "Very Good" and "Poor" each achieved a percentage of (2.14) of area under normal distribution curve. levels "Good" and "Very Poor" each achieved a percentage of (13.59) of that area, while percentage achieved by levels "Average" and "Poor" was (34.13) of aforementioned area. Based on these levels, positions of adjusted standardization population scores, as per Table (11), were determined on area under curve.

Table 11. It shows standard levels and their percentages for performance evaluation scale for physical education teachers

| Standard levels | Very good 2.14 | | Good 13.59 | | Average 34.13 | | Weak 34.13 | | Very weak 13.59 | | Bad 2.14 | | Total 137 |
|------------------------------|-------------------|------|---------------|-------|------------------|-------|---------------|-------|--------------------|-------|-------------|------|--------------|
| | No | % | No | % | No | % | No | % | No | % | No | % | |
| Performance evaluation scale | 4 | 2.91 | 18 | 13.13 | 51 | 37.22 | 38 | 27.73 | 20 | 14.95 | 6 | 4.37 | 137 |

table shows that there is a clear difference between percentages of standard levels achieved by members of regulatory community. At (very good) level, standardization sample achieved a percentage of (2.91), which is higher than percentage specified for this level under curve, which is (2.14). At (good) level, standardization sample achieved a percentage of (13.13), which is lower than percentage specified for this level under curve, which is (13.59). At average level, standardization sample achieved a percentage of (37.22), which is higher than percentage specified for this level under curve, which is (34.13). At weak level, standardization sample achieved a percentage of (27.73), which is lower than percentage specified for this level under curve, which is (34.13). At very weak level, standardization sample achieved a percentage of (14.59), which is higher than percentage specified for this level under te curve, which is (13.59). As for poor level, standardization sample achieved a percentage of (4.37), which is higher than percentage specified for this level under curve, which is (2.14).

By presenting and analyzing above results, it is shown that highest percentage of legalization community members were distributed within (average) level, which is an unacceptable level, while rest of them were distributed within levels of (weak), (very weak), (good), (poor), and (very good).

Conclusions

Reaching conclusion developing a performance evaluation scale during lesson for male and female physical education teachers in College of Education. It was found through identifying performance level of male and female physical education teachers at Open Educational College that it was an unacceptable level. Performance level of female physical education teachers during lesson is better than performance level of male teachers in all aspects of study. best level of performance during lesson for male and female physical education teachers was within administrative and organizational axis.

Recommendations

In light of findings, the researcher recommends following: Supervisors of physical education use current scale to evaluate performance of physical education teachers at Open Educational College. Providing male and female physical education teachers at Open Educational College with handouts containing most important teaching skills necessary to

develop their performance, in addition to modern teaching methods and techniques and latest developments in laws of sports games, so that they can review them periodically.

References

- 1 - Marwan Abdul Majeed, and Muhammad Jassim Al-Yasiri: Modern Trends in Physical Education Teaching Methods . 1st ed. 2001. p. 22.
- 2- Shawqi Shamikh: Human Behaviors and Their Impact on Business Management, Baghdad, Dar Al-Hikma, 1992, p. 15.
- 3- Mu'ayyad Saeed Al-Salem, Adel Harhoush Saleh: Human Resources Management, Baghdad University Press, 2000, p. 167.
- 4- Amer Saeed Jassim Al-Khaikani: Building a scale for aggression against athletes, standardizing it for football players, determining its levels, and comparing it according to playing positions, PhD thesis, unpublished, University of Baghdad, 2002, p. 54.
- 5- Muhammad Hassan Alawi, Muhammad Nasr Al-Din Radwan: Measurement in Physical Education and Sports Psychology , Cairo, Dar Al-Fikr Al-Arabi, 1988, p. 258.
- 6- Muhammad Shehata Rabie: Personality Measurement , Cairo, Dar Al-Maarefa, 1994, p. 98.
- 7- Abdul Rahman Ahmed: Leadership and Training in Psychology and Education, Alexandria, University Knowledge House, 1985, p. 41.
- 8- Ahmed Mohamed Khater, Fahmy Al-Beik: Measurement in Field of Sports, 4th Edition, Cairo, Dar Al-Kitab Al-Hadith, 1996, p. 37.