قائمة المراجع

أحمد الرفاعي غنيم. (١٩٨٣). أثر التغذية الرجعية الفورية على صدق وثبات الاختبارات العقلية ذات الاختيار من متعدد، رسالة دكتوراه غير منشورة، كلية التربية، جامعة الزقازيق.

أحمد الرفاعي غنيم. (١٩٨٥). تطبيقات على ثبات الاختبار. الطبعة الأولى، القاهرة: نهضة الشرق.

أحمد الرفاعي غنيم. (١٩٩١). تعميم معامل ألفا لحساب معامل ثبات المقاييس ذات المفردات غير المتجانسة: مجلة كلية التربية، جامعة عين شمس، ع١٥، ص ص(٢٠٧-٢٣٧).

أحمد الرفاعي غنيم. (١٩٩٥). طريقة جديدة لحساب معامل ثبات المفردات المتجانسة في المقاييس النفسية والتربوية، المجلة المصرية للدراسات النفسية، ع١٢، ص ص (٣٣-٤٠).

أحمد عبادة سرحان. (بدون تاريخ). العينات، معهد الدراسات والبحوث الإحصائية، جامعة القاهرة.

إسماعيل الفقي. (٢٠٠٠). دراسة الخصائص السيكومترية لبعض المقاييس والاختبارات في عينة من البحوث النفسية والتربوية، المجلة المصرية للدراسات النفسية، المجلد العاشر، ع٢٨، ص ص(٢٦١-٢٨٦).

السيد عبدالحميد الدالي. (١٩٥٩). قواعد البحث الاجتماعي بالطريقة الإحصائية. القاهرة: النهضة المصرية.

حسن مصطفى (٢٠٠٠). أثر نوع الأسئلة ومستوى الصعوبة على تقدير الدرجة الحقيقية ومعامل الثبات، رسالة ماجستير غير منشورة، كلية التربية، جامعة المنصورة.

رجاء محمود أبو علام. (١٩٨٧). قياس وتقويم التحصيل الدراسي. الكويت: دار القلم.

رجاء محمود أبو علام. (٢٠٠١). مناهج البحث في العلوم النفسية والتربوية. القاهرة: دار النشر للجامعات.

رجاء محمود أبو علام (٢٠٠٣). التحليل الإحصائي للبيانات باستخدام برنامج SPSS. القاهرة: دار النشر للجامعات.

رمزية الغريب. (١٩٩٦). التقويم والقياس النفسي والتربوي. القاهرة: الأنجلو المصرية.

سعد عبدالرحمن. (١٩٨٣). القياس النفسي. الكويت: مكتبة الفلاح.

سمير كامل عاشور، سامية ابو الفتوح. (١٩٩٤). مقدمة لنظرية العينات، معهد الدراسات والبحوث الإحصائية، جامعة القاهرة.

سمير كامل عاشور، سامية ابو الفتوح (٢٠٠٠). مقدمة في الإحصاء التحليلي، معهد الدراسات والبحوث الإحصائية، جامعة القاهرة.

سيد خير الله، ممدوح الكناني. (١٩٨٧). علم النفس التربوي. كلية التربية، جامعة المنصورة.

صفوت فرج. (١٩٨٠). القياس النفسي. الطبعة الأولى. القاهرة: الأنجلو المصرية.

صفوت فرج. (١٩٩٧). القياس النفسي. الطبعة الثالثة. القاهرة: الأنجلو المصرية.

صلاح أبو ناهية. (١٩٩٤). القياس التربوي. القاهرة: الأنجلو المصرية.

صلاح الدين محمود علام. (١٩٩٧). دليل المعلم في تقويم الطلبة في الدراسات الاجتماعية. القاهرة: دار الفكر العربي.

صلاح الدين محمود علام. (٢٠٠١). الاختبارات التشخيصية مرجعية المحك. القاهرة: دار الفكر العربي.

صلاح الدين محمود علام. (٢٠٠٢). القياس والتقويم النفسي والتربوي، أساسياته وتطبيقاته وتوجهاته المعاصرة، ط١، القاهرة: دار الفكر العربي.

عبدالعزيز القوصي، حسن محمد حسين، محمد خليفة بركات. (١٩٥٦). الإحصاء في التربية وعلم النفس. القاهرة: مكتبة النهضة المصرية.

عبدالهادي السيد، فاروق عثمان. (١٩٩٥). الإحصاء التربوي والقياس النفسي. المنصورة: دار المعارف.

علي حامد الثبيتي. (١٩٩٦). ثبات أدوات القياس في العلوم الإنسانية بين النظرية التقليدية للاختبار ونظرية التعميم. في مستقبل التربية العربية. الطائف.

فاروق عبدالفتاح موسى. (١٩٩٠). القياس النفسي والتربوي للأسوياء والمعوقين. القاهرة: مكتبة النهضة المصرية.

فتحي الزيات. (١٩٨٩). أثر قوة تشتيت البدائل في أسئلة الاختيار من متعدد على صدق الاختبار وثباته. مجلة كلية التربية، جامعة المنصورة. ص ص(٨٥-١٠٨).

فؤاد أبو حطب، سيد عثمان. (١٩٧٣). مشكلات في التقويم النفسي. القاهرة: الأنجلو المصرية.

فؤاد أبو حطب، سيد عثمان، آمال صادق. (١٩٨٧). التقويم النفسي. القاهرة: الأنجلو المصرية.

فؤاد البهي السيد. (١٩٧٩). علم النفس الإحصائي وقياس العقل البشري. الطبعة الثالثة: القاهرة: دار الفكر العربي.

محمد رمضان محمد. (١٩٨٨). الاختبارات التحصيلية والقياس النفسي والتربوي. دبي: دار القلم.

محمد عبدالسلام أحمد. (غير موضح سنة النشر). القياس النفسي والتربوي. القاهرة: مكتبة النهضة المصرية.

محمد منصور الشافعي. (١٩٩٧). أثر طرق معادلة درجات الاختبار وتدريج بنود بنك الأسئلة باستخدام نموذج راش. رسالة دكتوراه غير منشورة، كلية التربية، جامعة المنصورة.

محمود منسى. (١٩٩٤). المدخل إلى علم النفس التعليمي. الطبعة الأولى. الاسكندرية: دار المعرفة الجامعية.

ممدوح الكناني وعيسى الجابر. (١٩٩٥). القياس والتقويم النفسي والتربوي. الكويت. مكتبة الفلاح.

نادية محمود شريف، محمود محمد إبراهيم. (٢٠٠١). مقدمة فى القيـــاس والتقويـــم. معـــهد الدراسات والبحوث التربوية، جامعة القاهرة.

نعمان الموسوى. (١٩٩٧). "تحديد العدد الأمثل لخيارات الفقرة فى اختبار الاختيار من متعدد فى ضوء بعض خصائصه الســـيكومترية، المجلـــة التربويـــة، العـــدد٤٤، مجلـــد١١، ص ص٢٠١-٢٣٧.

Anastasi, A. & Susana, U. (1997). "Psychological Testing" International Edition. (7th Ed). Prentice Hall. International Inc.

Andres, A. M. & Castillo, L. (1990). Multiple Choice Tests: Power, Length And Optimal Number Of Choices Per Item. Journal Of Mathematical And Statistical Psychology, Vol.43. No.I. 57-71.

Annie W. Ward & Mildred Murray- Ward. (1999). Assessment In The Classroom.

Brennan, R. L. (1998). Misconceptions At The Intersection Of Measurement Theory And Practice. Educational Measurement Issues And Practice. Vol.17. No.I.

Brenner, M. H. (1965). Test Difficulty, Reliability And Discrimination As Functions Of Item Difficulty Order. Journal Of Applied Psychology, Vol.48, No.2, 98-100.

Bruno, J. E. & Dirkzwager, A. (1995). Determining The Optimal Number Of Alternatives To Amultiple – Choice Test Item. An Information Theoretical Perspective. Educational And Psychological Measurement, Vol.55, No.6, 959-966.

Budescu, D. V. & Nevo, B. (1985). Optimal Number Of Options: An Investigation Of Assumption Of Proportionality. Journal Of Educational Measurement. Vol.22, No.3, 183-196.

Chase, C. I. (1978). Measurement For Educational Evaluation, London, Addisson Wesely Publishing Co. (2ed).

Chissom, B. C. (1985). An Investigation Of The Relationship Between Item Arrangement And Test Performance. Paper Presented At The Annual Meeting Of The Mid-South Educational Research.

Costin, F. (1970). The Optimal Number Of Alternatives In Multiple-Choice Achievement Test: Some Empirical Evidence For

Mathematical Proof. Educational And Psychological Measurement, Vol.30, 353-358.

Costine, F. (1972). Three Choice Versus Four Choice Items: Implications For Reliability And Validity Of Objective Achievement Test. Educational And Psychological Measurment, Vol.32, 1035-1038.

Cronbach, L. J. (1970). Essential Of Psychological Testing. New York: Harper And Row.

Cronbach, L. J. & Et Al. (1997). Generalizability Analysis For Performance Assessments Of Student Achievement Or School Effectiveness. Educational And Psychological Measurment, Vol.57, No.3, 373-399.

David Thissen. (1996). How Is Reliability Related To The Quality Of Test Scores: What The Effect Of Local Dependence On Reliability. Educational Measurment Issues And Practice, Vol (15), No.1, 22.

David, V. B. (1972). Standard Error Of Estimate In Item Examinee Sampling As A Function Of Test Realiability Variation, In Item Difficulty Indices And Degree Of Skewness In The Normative Distribution. Educational And Psychological Measurement, Vol. 32, No.3, 705-714.

Duncan, E. (1983). An Appropriate Number Of Multiple Choice Item Alternatives: A Difference Of Opinion. Measurement And Evaluation In Guidance. Vol.15, No.4, 283-292.

Ebel, R. L. (1969). Expected Reliability As A Function Of Choices Per Item. Educational And Psychological Measurment. Vol.29, 565-570.

Ebel, R. L. (1975). Can Teachers Write Good True-False Test Items. Journal Of Educational Measurment, Vol.12, No.1, 31-35.

Ebel, R. L. & Williams, B. J. (1979). The Effects Of The Varying The Number Of Alternatives Per Item On Multiple Choice Vocabulary Test Items. New Jeresy: Prentice-Hall INC Englewood Cliffs.

Enders, C. K. (1999). The Effects Of Hetrogenous Item Distributions On Reliability. Applied Measurment In Education, Vol.12, No.2, 133-150.

Fan, X. & Michael, C. (1999). When Inter Rater Reliability Is Obtained From Only Part Of Sample, Paper Presented At The Annual

Meeting Of The American Educational Research Association, Canada: Montreal, Quebec, April 19-23.

Fan, X. & Ping, Y. (2001). Sample Characteristics And Measurment Reliability: An Emprical Exploration, Paper Presented At The Annual Meeting Of The American Education Research Association.

Feldt, L. S. (1985). A Test Of The Hypothesis That Cronbach's Alpha Reliability Coefficent Is The Same For Two Tests Administered To The Same Sample. Psychometrika, Vol.45, No.1, 99-105.

Feldt, L. S. & Ankenmann, R. D. (1998). Appropriate Sample Size For Comparing Alpha Reliability. Applied Psychological Measurment, Vol.22, No.2, 170-178.

Fitzpatrick, A. R. & Yen, W. M. (2001). The Effects Of Test Length And Sample Size On The Reliability And Equating Of Tests Composed Of Constructed – Response Items, Applied Measurment In Education, Vol.14, No.1, 31-57.

Flaugher, R. L. & Melton, R. S. & Myers, C. T. (1968). Item Rearrangement Under Typical Test Conditions. Educational And Psychological Measurment, Vol.28, 813-824.

Frisbie, A. D. & Druva, C. A. (1986). Estimating The Reliability Of Multiple Choice & True – False Tests. Journal Of Educational Measurment, Vol.23, No.2, 99-105.

Gerberich, J. R. & Green, H. A. & Jorgenson. (1962). Measurment And Evaluation In Modern School. New York: David Mckay.

Green, K. & Sax, G. & Micheal, W. B. (1982). Validity And Reliability Of Tests Having Differing Numbers Of Options For Students Of Differing Levels Of Ability. Educational And Psychological Measurment, Vol.42, 239-245.

Grier, J. B. (1975). The Number Of Alternatives For Optimum Test Reliability. Journal Of Educational Measurment, Vol.12, No.2, 109-113.

Gronlund, N. E. & Linn R. (1990). Measurment And Evaluation In Testing. (6th Ed). New York: Macmillan Publishing Company.

Guemin, Lee. (2000). Acomparison Of Methods Of Estimating Conditional Standard Errors Of Measurment For Testlet-Based

Test Scores Using Simulation Techniques. Journal Of Educational Measurment, Vol.36, No.2, 91-112.

Guttman, L. (1945). Abasis For Anaryzing Test Retest Reliability. Psychometrika, Vol.10, 255-282.

Hambleton, R. K. & Cook, L. L. (1983). Robustness Of Item Response Models And Effects Of Test Length And Sample Size On The Precision Of Ability Estimates. In D. J. Weiss (Ed), New York: Academic. 31-49.

Hambleton, R. K. & Traub, E. R. (1974). The Effects Of Item Order On Test Performance And Stress. Journal Of Experimental Education Vol. 43, No. 40-45.

Hancock, G. R. & Sax, G. & Micheal, W. B. (1993). Reliability Of Comparaly Written Two Option Multiple- Choice And True-False-Test Items. Educational And Psychological Measurment, Vol.53, 651-658.

Hogben, D. (1973). The Reliability, Discrimination And Difficulty Of Word Knowledge Tests Employing Multiple-Choie Items Containing Three-Four Or Five Alternatives. Australian Journal Of Education, Vol.17, 63-68.

Hopkins, K. D. And J. C. Stanley, 1981, Educational And Psychological Measurement And Evaluation. 6th Ed. Englewood Cliffs, N. J; Prentice-Hall.

Huck, S. W. (1978). Test Performance Under The Condition Of Know Item Difficulty. Journal Of Educational Measurment, Vol.15, No.1.

Huck, S. W. & Bowers, N. D. Item Difficulty Level And Sequence Effects In Multiple-Choice Achievement Tests. Journal Of Educational Measurment, Vol.9, No.2, 1972, 105-111.

Intraprasert, D. (1986). An Inverstigation Of The Reliability Of Five Methods For Detecting Test Item Bias, Phd, University Of North-Taxas.

Jurs, S. G. & Wiersma, W. (1985). Educational Measurment And Testing. The University Of Toledo.

Kendall, M. & Stewart, A. The Advanced Theory Of Statistics, Vol.1, New York: Macmillan Publishers.

Kidder, L. H. (1981). Research Methods In Social Relations. (4th Ed). New York: Holt, Rinehart & Winston.

Kline, P. (2000). Handbook Of Psychological Testing. (2nd Ed). New York: Routledge.

Kopriva, R. J. & Shaw, D. G. (1991). Power Estimates: The Effect Of Dependent Variable Realiability On The Power Of One Factor Anovas. Educational And Psychological Measurment. Vol.51, 585-595.

Krus, D. J. & Helmstadter, G. C. (1993). The Problem Of Negative Reliabilities. Journal Of Educational And Psychological Measurement, Vol. 53, 643-649.

Leon, A. C. Marzuk, P.M. & Portera, L. (1995) More Reliable Outcome Measures Can Reduce Sample Size Requirements, Archgen Psychiatry. Vol. 52, 867-871.

Levin, J. (1986). Note On The Relation Between The Power Of Significance Test And The Reliability Of The Measuring Instrument. Multivariage-Behavioral Research. Vol.21, No.2, 255-261.

Levin, M. V. & Drasqow, F. (1983). The Relation Between Incorrect Option Choice And Estimated Ability. Educational And Psychological Measurment, Vol.43, 675-685.

Linacre, J. M. (1995). Reliability And Separation Nomograms, Rach Measurment Transactions, Vol.9, No.2, 421.

Livingston, A. S. (1990). Reliability Of Test Results Edited By Keeves, J. P. Educational Research Methodology And Measurment. New Jersey: An International Handbook Pergamon Press. 386-392.

Lord, F. M. (1977). Optimal Number Of Choices Per Item. A Comparison Of Four Appraches. Journal Of Educational Measurment. Vol.14, 33-38.

Lord, F. M. & Novick, M. R. (1968). Statistical Theories Of Mental Test Scores. Reading, Ma: Addison-Wesley.

Lund, K. (1953). Test Performance As Related To Order Of Item Difficulty, Anxiety And Intelligence. Unpublished Doctoral Dissertation North Western University.

Macnicol, K. (1966). Effects Of Varying Order Of Item Difficulty In An Unspeeded Verbal Test. Unpublished Manuscript, Educational Testing Service, Princeton, New Jersey.

Maihoff, N. & Meherns, W. A. (1985). Acomparison Of Alternate Choice And True-False Item Forms Used In Classroom Examination. National Councilon Measurment In Evaluation, Chicago: Il. April. 1-3.

Marshal, S. (1990). The Effect Of Employing Different Number Of Response Alternatives On The Reliability And Validity Of Likert Scales For Children In Varied Grade Levels, Phd, John's University.

Marso. R. N. (1970). Test Item Arrangement, Testing Time Limit And Performance. Journal Of Educational Measurment, Vol.7, No.2, 113-118.

Mckeachie, W. J. Pollie, D. & Speisman, J. (1995). Relieving Anxiety In Classroom Examinations. Journal Of Abnormal And Social Psychology, Vol.50, 93-98.

Mehrens, W. A. & Lrvin, J. L. (1978). Measurement And Evaluation In Education And Psychology (2nd Ed). Halt Rinehort And Winstan, New York.

Mehrens, W. A. & Lrvin, J. L. (1980). Measurment And Psychology. (2nd Ed). New York: Holt, Rinehart And Winston, Chicago.

Melnick, S. A. (1993). The Effects Of Item Grouping On The Reliability And Scales Scores Of An Affective Measure. Educational And Psychological Measurment, Vol.53, No.1, 211-216.

Merrill, M. (1993). The Relationship Between Validity, Reliability, And Test Length In Commonly Used Personality Scales. Fairleigh Dickison University.

Messick, S. (1989). Validity In R. L. Linn. (Ed). Educational Measurement. New York: American Council. On Education And Macmillan. 13-103.

Michael, S. T. Sax, G. & William, B. (1991). The Effects Of The Number Of Options Per Item And Student Ability On Test Validity And Reliability. Educational And Psychological Measurement. Vol.51, No.4, 829-837.

Mollenkoph, W. G. (1950). An Experimental Study Of The Effets On Item Analysis Data Of Changing Item Placement And Test Time Limit. Psychometrika, Vol.15, 291-315.

Munz, D. C. & Smouse, A. D. (1968). Interaction Effects Of Item Difficulty Sequence And Achievement – Anxiety Reaction On Academic Performance. Journal Of Educational Psychology, Vol.59, 370-374.

Newman, D. L., Kundert, D. K., Leon, D. S. & Bull, K. B. (1988). Effect Of Varying Item Order On Multiple-Choice Test Scores. Importance Of Statistical And Cognitive Difficulty. Applied Measurment In Education. 22-89-97.

Nitki, A. J. (1983). Educational Tests And Measurement An Introduction, New York: Harcour Brace. Jovanovich.

Novick, M. R. (1966). The Axioms And Principal Results Of Classical Test Theory. Journal Of Mathematical Psychology, Vol.3. 1-18.

Nunnally, J. C. (1978). Psychometric Theory. (2nd Ed). Usa, New York: Macgraw-Hill Company.

Plack, B. S. (1982). Item Arrangement And Knowledge Of Arrangement On Test Scores. Journal Of Experimental Education, 53-58.

Pyrczak, F. (1973). Validity Of The Discrimination Index As Ameasure Of Item Quality. Journal Of Educational Measurment. Vol.10, 227-231.

Qualls-Payne, A. L. (1992). A Comparison Of Score Level Estimates Of The Standard Error Of Measurement, Journal Of Educational Measurment, Vol.29, 213-225.

Rogers, W. T. & Harley, D. (1999). An Emprical Comparison Of Three And Four Choice Items And Tests. Susceptibility To Testwiseness And Internal Consistency Reliability. Educational And Psychological Measurement. Vol.59, No.2, 234-247.

Rose, D. S. (1997). The Relationship Between Sample Size And Measurment Reliability, As Applied To Banding In Continuous Hiring Settings, Depaul- University.

Sax, G. (1980). Principles Of Educational And Psychological Measurment And Evaluation. (2nd Ed). Usa. California: Wods Woth Publishing Co. Inc.

Sax, G. & Cromack, T. R. (1966). The Effects Of Various Forms Of Item Arrangements On Test Performance. Journal Of Educational Measurment. Vol.3, No.4, 309-311.

Schriesheim, C. A. (1980). Item Presentation As An Influence On Questionnaire Validity Afield Experiment. Educational And Psychological Measurment. Vol.40, 175-182.

Schriesheim, C. A. (1981). Leniency Effects On Convergent And Discrimination Validity For Grouped Questionnaire Items. Afurther Investigation. Educational And Psychological Measurment. Vol.41, 1093-1099.

Schurr, K. T. & Henriksen, L. W. (1983). Effects Of Item Sequencing And Grouping In Low Inference Type Questionnaires. Journnal Of Educational Measurement. Vol.20, No.4, 379-390.

Stallings, W. M. & Monk, J. J. (1970). Effects Of Item Order On Test Scores. Journal Of Educational Research. Vol.63, No.10, 463-465.

Steven, G. J. (1981). Reliability Estimation Of Atest With Parts Of Unequal Length. Phd. The University Of Lawa. 10-13.

Steven, T. M. (1990). Reliability And Validity Of Multiple Choice Examination As A Function Of The Number Of Options Per Item And Student Ability. Phd. University Of Washington.

Straton, R. G. & Catts, R. M. (1980) Acomparison Of Two, Three And Four Choice Item Tests Given afixed Number Of Choices. Educational And Psychological Measurment, Vol.23, 357-365.

Swaminathan, J. & Gifford, J. A. (1983). Estimates Of Parameters In The Three Parameter Latent Trait Model In D. J. Weiss. (Ed). New Horizons In Testing. New York: Academic. 13-30.

Symonds, P. M. (1928). Factors Influencing Test Reliability. Journal Of Educational Psychology, Vol.19, 73-87.

Thorndike, R. L. (1976). Educational Measurement, Second Impression, Washington: D. C. American Council On Education. New York. 371-372.

Trevison, M. (1990). Reliability And Validity Of Multiple Choice Examinations As A Function Of The Number Of Options Per Item And Student Ability. Paper Presented At The Annual Meeting Of The American Educational Research Association. April 16-20.

Tversky, A. (1964). On The Optimal Number Of Alternatives At Achoice Point. Journal Of Mathematical Psychology, Vol.1, 389-391.

Wagner, E. (1990). An Emprical Demonstration Of The Stability Of The Maximized Correlation As An Internal – Consistency Reliability Estimate For Tests Of Small Item Size. Educational And Psychological Measurement, Vol.50, No.3.

Wagner, E. Alexander, R. A., Rose, G. & Prospero, K. (1985). Maximizing Split-Half Reliability Estimates For Projective Techniques. Journal Of Personality Assessment, Vol.49, 579-581.

Wagner, E. Green, R. S., Adair, H. E. & Dawson, J. (1987). Maximized Reliability Coefficient For Lnnb Factor Scales. Journal Of Personality Assessment, Vol.9, 132-135.

Weber, M. B. (1978). The Effect Of The Choice Format On Internal Consistency. Paper Presented At The National Council On Measurement In Educational Annual Meeting. Toronto Canada.

Wiersma, W. & Jurs, S. G. (1990). Educational Measurment Testing. (2nd Ed). London: Allyn And Bacon.

Williams, B. J. & Ebel, R. L. (1957). The Effect Of Varying The Number Of Alternatives Per Item On Multiple Choice Vocabulary Test Items. In The 14th Yearbook Of The National Council On Measurement Used In Education East Lansing. Michigan: Michigan State University. 63-65.

Zimmerman, W. S. & Humphreys, L. G. (1953) Item Reliability As A Function Of The Omission Of Misleads. American Psychologist, Vol.8, 460-461.

ملاحق الدراسة

ملحق رقم (١) الأهداف السلوكية للاختبار.

ملحق رقم (٢) الاختبار التحصيلي في صورته النهائية.

ملحق رقم (٣) نموذج ورقة إجابة الاختبار في صورته النهائية.

ملحق رقم (٤) الفرق بين نسب الإجابة الصحيحة فى المجموعتين (العليا والدنيا) على جميع مفردات الاختبار فى ضوء نتائج الدراسة الأولية.

ملحق رقم (٥) قيم معاملات السهولة والصعوبة لجميع مفردات الاختبار وترتيب المفردات وفقاً لمعاملات الصعوبة.

ملحق رقم (٦) قيم معاملات التمييز لمفردات الاختبار.

ملحق رقم (٧) نتائج تحليل فاعلية المشتتات لمفردات الاختبار.

ملحق رقم (٨) الاختبار الإحصائى الذى تم الاعتماد عليه فى اختبار فروض الدراسة.

ملحق رقم (١)

الأهداف السلوكية للاختبار في صورتها المبدئية

١- يتعرف الطالب على خطوات بناء المقياس.

٢- يميز الطالب بين خطوات القياس.

٣- يتعرف الطالب على مفهوم الموضوعية في مجال القياس.

٤- يميز الطالب بين وظيفتي القياس والتقويم.

٥- يتعرف الطالب على أهم خاصية تميز مستوى النسبة كأعلى مستويات القياس.

٦- يربط الطالب بين بعض إجراءات القياس ومستوى القياس المناسب لكل منها.

٧- يتعرف الطالب على المجالات التي توجد بها أخطاء القياس.

٨- يميز الطالب بين المفاهيم أو السمات التي يسهل قياسها والسمات التي لا يسهل قياسها.

٩- يستنتج الطالب أهم أسباب ارتفاع قيمة معامل ثبات الاختبارات الموضوعية بالمقارنة باختبارات المقال.

١٠- يتعرف الطالب على العلاقة التي تربط التقويم بالأهداف.

١١- يستنتج الطالب الهدف الأمثل من عملية التقويم.

١٢- يتعرف الطالب على أهم الشروط التي يجب مراعاتها عند صياغة الأهداف التربوية الخاصة.

١٣- يرتب الطالب مستويات المجال المعرفي وفقاً لتصنيف بلوم في تسلسلها من الأدنى إلى الأعلى.

١٤- يتعرف الطالب على الأفعال التي يجب استخدامها عند صياغة الأهداف التربوية الخاصة.

١٥- يتعرف الطالب على أهم شرط يجب توافره عند صياغة الهدف السلوكي.

١٦- يربط الطالب بين إجراء القياس الذي يقوم به المعلم ومستوى الهدف الذي يناسب هـذا الإجراء.

١٧- يتعرف الطالب على الأساس الذي تصنف الأهداف التربوية وفقاً له.

١٨- يربط الطالب بين مستوى الهدف التعليمي والمجال المناسب له.

١٩- يتعرف الطالب على الغرض الأساسي من عملية التقويم.

٢٠- يتعرف الطالب على أهم شرط يجب توافره حتى يحقق التقويم فاعليته بالنسبة للمدرس والمتعلم.

٢١- يقيم الطالب مجموعة من الأهداف التعليمية وفقاً لمعايير معينة.

٢٢- يتعرف الطالب على الإجراء المناسب الذي يجب القيام به حتى تسهم نتائج التقويم في تحسين عملية التعلم.

٢٣- يربط الطالب بين مجموعة من الأنشطة التعليمية ومجالات الأهداف التي تنتمي إليـها هذه الأنشطة.

٢٤- يتعرف الطالب على أهم خطوة يجب وضعها في الاعتبار الأول عند بنـاء اختبـار تحصيلي.

٢٥- يتعرف الطالب على أهم خطوة يجب البدء بها عند إعداد برنامج لتقويم الطلبـة فـي المدرسة.

٢٦- يقارن الطالب بين الأسئلة الموضوعية والأسئلة المقالية.

٢٧- يستنتج الطالب الشرط الأساسي الذي يجب توافره حتى يكون السؤال مقياساً صادقـاً للفهم.

٢٨- يتعرف الطالب على أهم خاصية تميز الاختبارات الموضوعية بالمقارنة باختبـارات المقال.

٢٩- يتعرف الطالب على أهم أوجه النقد التي توجه إلى أسئلة المقال.

٣٠- يتعرف الطالب على الميزة الرئيسية لأسلوب الملاحظة المباشرة بالمقارنة بالوسائل الأخرى للتقويم.

٣١- يتعرف الطالب على السمات التي يصعب تقويمها باستخدام مقاييس التقدير.

٣٢- يربط الطالب بين وسيلة التقدير وأهم خصائص الصفة موضوع التقدير.

٣٣- يميز الطالب بين الصفات التي يمكن قياسها قياساً ثابتاً باستخدام مقاييس التقدير والصفات التي يصعب قياسها قياساً ثابتاً باستخدام مقاييس التقدير.

٣٤- يربط الطالب بين النشاط التعليمي ووسيلة القياس المناسبة لهذا النشاط.

٣٥- يتعرف الطالب على المقصود بأثر الهالة.

٣٦- يستنتج الطالب السبب الرئيسي الذي يجعل البعض يعتبر مقاييس التقدير أدوات تقويم ضعيفة.

٣٧- يتعرف الطالب على الأسلوب الأمثل في تحديد نقاط مقياس التقدير والذي يزيد من صدق المقياس.

٣٨- يتعرف الطالب على العامل الرئيسي الذي يؤدي إلى زيادة مصادر الخطأ في مقاييس التقدير.

٣٩- يميز الطالب بين مخرجات التعلم التي يمكن تقويمها بفاعلية باستخدام الواجبات المنزلية ومخرجات التعلم التي يمكن تقويمها بفاعلية باستخدام اختبارات الورقة والقلم.

٤٠- يتعرف الطالب على الأسلوب الأمثل الذي يزيد من فاعلية الدرجات المدرسية للطلاب.

٤١- يقارن الطالب بين الاختبارات محكية المرجع والاختبارات معيارية المرجع.

٤٢- يميز الطالب بين التقويم البنائي والتقويم النهائي.

٤٣- يربط الطالب بين مفهوم صدق الاختبار والتعريف المناسب له.

٤٤- يربط الطالب بين مفهوم الثبات والتعريف المناسب له.

٤٥- يتعرف الطالب على أهمية جدول مواصفات الاختبار التحصيلي.

٤٦- يميز الطالب بين الأفراد الذين يفضل معهم استخدام أسلوب الملاحظة المباشرة.

٤٧- يربط الطالب بين خاصية ثبات القياس ونوع الخطأ المرتبط به.

٤٨- يحسب الطالب قيمة معامل ثبات الاختبار بعد مضاعفة عدد مفرداته.

٤٩- يقيم الطالب مدى صحة أو خطأ قيمة معينة لمعامل الثبات.

٥٠- يستنتج الطالب العلاقة بين معامل ثبات الاختبار والخطأ المعياري للقياس.

٥١- يتعرف الطالب على الخطوة الأولى في إعداد قائمة مراجعه.

٥٢- يستنتج الطالب أفضل تفسير ممكن لمعامل الارتباط الذي حصل عليه الباحث في موقف القياس.

٥٣- يربط الطالب بين مجموعة من الأنشطة أو الإجراءات الخاصة بالاختبار والأهداف المناسبة لها.

٥٤- يربط الطالب بين مجموعة من إجراءات القياس الخاصة بالاختبار والتأثيرات المحتملة لهذه الإجراءات على قيمة معامل ثبات الاختبار.

٥٥- يميز الطالب بين مجموعة من القيم لمعاملات الصدق من حيث قوتها أو ضعفها.

٥٦- يتعرف الطالب على أهم عامل يجب تحقيقه عند اختيار أو بناء وسيلة التقويم.

تابع ملحق (١)

بعض الأهداف السلوكية للمفردات

التي تمت إضافتها للاختبار في صورته النهائية

١- يحسب الطالب قيمة الدرجة المعيارية المناظرة للدرجة الخام إذا علم قيمة المتوسط الحسابي والانحراف المعياري للدرجات.

٢- يميز الطالب بين الهدف المصاغ بطريقة سلوكية والهدف المصاغ بطريقة غير سلوكية.

٣- يحسب الطالب قيمة معامل الصعوبة لسؤال معين.

٤- يختار الطالب وسيلة القياس الأفضل في تقويم أنشطة معينة.

٥- يربط الطالب بين مستوى القياس والخاصية التي تميزه.

تابع ملحق (١)

الأهداف السلوكية والمفردات التي تقابلها والمستوى المعرفي للهدف للاختبار في صورته المبدئية

الهدف السلوكي	المفردات التي تقيسه	مستوى الهدف
-١	١	تطبيق
-٢	٢	فهم
-٣	٣	فهم
-٤	٤	فهم
-٥	٥	تذكر
-٦	٦-٧-٨-٩-١٠	تطبيق
-٧	١١	فهم
-٨	١٢	تطبيق
-٩	١٣	فهم
-١٠	١٤	فهم
-١١	١٥	فهم
-١٢	١٦	تطبيق
-١٣	١٧	تذكر
-١٤	١٨	تذكر
-١٥	١٩	فهم
-١٦	٢٠	تذكر
-١٧	٢١	فهم
-١٨	٢٢	تذكر
-١٩	٢٣	تذكر
-٢٠	٢٤	تطبيق
-٢١	٢٥-٢٦-٢٧-٢٨-٢٩-٣٠-٣١	تطبيق
-٢٢	٣٢	تطبيق
-٢٣	٣٣	تذكر
-٢٤	٣٤	تذكر
-٢٥	٣٥	تذكر
-٢٦	٣٦	تذكر

الهدف السلوكي	المفردات التي تقيسه	مستوى الهدف
-٢٧	٣٧	فهم
-٢٨	٣٨	تذكر
-٢٩	٣٩	تذكر
-٣٠	٤١	تذكر
-٣١	٤٢	تطبيق
-٣٢	٤٣- ٤٤	تذكر
-٣٣	٤٥	تطبيق
-٣٤	٤٦-٤٧-٤٨-٤٩	تطبيق
-٣٥	٥٠	فهم
-٣٦	٥١	فهم
-٣٧	٥٢	فهم
-٣٨	٥٣	فهم
-٣٩	٥٤	تطبيق
-٤٠	٥٥	تطبيق
-٤١	٥٦	تذكر
-٤٢	٥٧-٥٨	تذكر
-٤٣	٥٩-٦٤	تذكر
-٤٤	٦٠-٦٢	تذكر
-٤٥	٦١	فهم
-٤٦	٦٣	تطبيق
-٤٧	٦٥	تطبيق
-٤٨	٦٦	تطبيق
-٤٩	٦٧	تطبيق
-٥٠	٦٩	تذكر
-٥١	٧٠	فهم
-٥٢	٧١	تطبيق
-٥٣	٧٢-٧٣-٧٤-٧٥-٧٥	تطبيق
-٥٤	٧٧-٧٨-٧٩-٨٠	تطبيق
-٥٥	٦٨	فهم
-٥٦	٤٠	تذكر

تابع ملحق (١)

بعض الأهداف السلوكية والمفردات التي تقيسها
ومستوى الهدف للاختبار في صورته النهائية
بعد التعديل فى صياغة بعض المفردات وإضافة مفردات جديدة فى ضوء نتائج الدراسة الأولية

الهدف	المفردة التي تقيسه	مستوى الهدف
١	٢٠	تطبيق
٢	٥٥-٥٦-٥٧-٥٨-٥٩-٦٠	تطبيق
٣	٥٠	تطبيق
٤	٦٨-٦٩-٧٠	تطبيق
٥	٩	فهم

ملحق (٢)
الاختبار التحصلي في صورته النهائية
(٧٠) مفردة في مقرر القياس والتقويم لطلاب الدبلوم الخاص في التربية "السنة الأولى التكميلية"
(كراسة الأسئلة)

تعليمات:

من فضلك حاول أن تقرأ كل فقرة من الفقرات التالية بعناية ثم حدد أي اختيار يجيب على السؤال أو يكمل العبارة. حدد إجابتك على ورقة الإجابة المنفصلة وذلك بتظليل الحـــرف الذي يمثل الإجابة التي تختارها. مثال ذلك إذا كانت الإجابة التي تختارها بالنسبة للسؤال رقم (١) هي (ج) عليك أن تظلل الحرف (ج) في ورقة الإجابة المنفصلة.

١- **يقصد بثبات الاختبار:**

أ - الحصول على نفس النتائج عند إعادة تطبيق الاختبار على مجموعات مختلفة.

ب- الحصول على نفس النتائج عند إعادة تطبيق الاختبار على نفس المجموعة.

جـــ شمولية أسئلة الاختبار لكل المقرر الذي درسه الطلبة في الفصل.

د - قدرة الاختبار على قياس الأهداف التربوية الخاصة.

٢- **ميزة التقويم البنائي على التقويم النهائي هي أن:**

أ - نتائج التقويم البنائي وسيلة لتحسين التعلم.

ب - التقويم البنائي يهتم بتحديد المستوى المطلق للمتعلم مقارنة بزملائه.

جـــ - التقويم البنائي أسهل في إجرائه من التقويم النهائي.

د - التقويم البنائي يستخدم اختبارات سريعة مقارنة بالتقويم النهائي.

٣- **أي التتابعات التالية للأهداف المعرفية تبدأ من المستوى الأدنى إلى المستوى الأعلى؟**

أ- التركيب ـــ التطبيق ـــ التذكر ـــ الفهم.

ب- الفهم ـــ التركيب ـــ التطبيق ـــ التذكر.

جـــ - التذكر ـــ التطبيق ـــ الفهم ـــ التركيب.

د- التذكر ـــ الفهم ـــ التطبيق ـــ التركيب.

٤- **أي المفاهيم التالية يصعب قياسه؟**

أ- التحصيل الدراسي.

ب- الاستعداد الدراسي.

جـــ- المواطنة الصالحة.

د- الميول الميكانيكية.

٥- **يجب صياغة الأهداف التربوية الخاصة في عبارات تبين:**

أ- ما يمكن للمتعلمين القيام به.

ب- عناصر المقرر الدراسية.

ج- طريقة التدريس.

د- مهارات المعلم في التدريس.

٦- **أي الخطوات التالية يجب البدء بها عند إعداد برنامج لتقويم الطلبة في المدرسة؟**

أ- التحقق من صدق وموضوعية الوسائل المستخدمة في التقويم.

ب- تحديد الأهداف المنشودة من التقويم.

ج- تحديد المواقف التي يمكننا أن نجمع منها المعلومات المتصلة بأهداف التقويم.

د- تحديد المجالات التي يجب أن يشملها التقويم.

٧- **يختلف التقويم البنائي عن التقويم النهائي في أن التقويم البنائي:**

أ- يستخدم أساليب أكثر دقة من التقويم النهائي.

ب- أقل استهلاكاً للوقت من التقويم النهائي.

ج- يقوم على الأهداف العامة أكثر من الأهداف الخاصة.

د- يتم على عدة مراحل أثناء العام الدراسي.

٨- **يجب أن يكون الغرض الأساسي من التقويم:**

أ- تحديد من من المتعلمين ينتقل إلى صف أعلى.

ب- إعداد شهادات الفترة للطلاب.

ج- توجيه التعليم والتدريس.

د- المحافظة على المستوى التحصيلي للتلاميذ.

٩- **في أي مستوى للقياس يعني الحصول على الدرجة صفر، الغياب التام للظاهرة موضوع القياس؟**

أ- المستوى الاسمي.

ب- مستوى النسبة.

ج- مستوى المسافة.

د- مستوى الرتبة.

١٠- **حصل باحث على معامل ثبات قدره ١,١٣، ما رأيك في هذه النتيجة؟**

أ- لابد من معرفة نوع الاختبار حتى نحكم على النتيجة.

ب- أن احتمال الحصول على هذه النتيجة احتمال متوسط.

ج- أن معامل الثبات وصل إلى حدة الأقصى.

د- لابد أن الباحث أخطأ في حساب معامل الثبات.

١١- مصطلح "موضوعي" كما يستخدم في القياس يشير إلى:

أ- طريقة تقدير درجات الاختبار.

ب- أهداف الاختبار.

ج- نوع المادة التي يغطيها الاختبار.

د- تقدير درجات أسئلة المقال.

١٢- أخطر نقد يمكن أن يوجه لأسئلة المقال أنها:

أ- تساعد على تكوين عادات استذكار سيئة.

ب- لا تستطيع قياس معظم أهداف التدريس.

ج- يصعب تقدير درجاتها تقديراً موضوعياً.

د- تؤكد على أهمية المهارات اللغوية.

١٣- إذا أعطى الاختبار نفس النتائج بالنسبة لنفس الأشخاص في كل مرة يطبق فيها عليهم فإننا نعتبر هذا الاختبار:

أ- متوازناً.

ب- ثابتاً.

ج- مقنناً.

د- صادقاً.

١٤- أي السمات التالية أصعب في تقويمها باستخدام مقاييس التقدير؟

أ- المواطنة الصالحة.

ب- النظافة.

ج- قيادة الجماعة.

د- وضع خطة للعمل.

١٥- عندما يقوم المعلم بوضع اختبار يركز فيه على استعادة الطلبة للخبرات التعليمية كما مرت بهم أثناء الدرس، فإنه في هذه الحالة يقيس الأهداف المتعلقة:

أ- بالتحليل.

ب- بالتطبيق.

ج- بالفهم.

د- بالتذكر.

١٦- أي الأعمال التالية تمثل استخدام نتائج التقويم لتحسين عملية التعلم؟

أ- إعطاء كل طالب الدرجة التي يستحقها بالضبط.

ب- تجميع التلاميذ الضعاف مع بعضهم البعض لإعطائهم دروساً في التقوية.

ج- إعادة تدريس بعض المبادئ التي أظهر الاختبار ضعف المتعلمين فيها.
د- إعطاء الطلاب نسخة من الاختبار ومناقشة الإجابات في الفصل.

١٧- **أهم ميزة للاختبارات الموضوعية على اختبارات المقال هي أنها:**
أ- أكثر ثباتاً.
ب- أكثر صدقاً.
ج- تقيس أهداف تربوية مفيدة.
د- تحتوي على عينة أكبر لمحتوى المقرر.

١٨- **يقصد بأثر "الهالة":**
أ- تأثير أحد المقدرين على مقدر آخر في مواقف التقويم.
ب- النزعة إلى إعطاء تقدير للفرد أعلى مما يستحق عندما نعرفه.
ج- انتشار الانطباع العام عن الشخص إلى تقدير خصائصه الأخرى.
د- النزعة إلى إعطاء تقديرات عالية للطلبة.

١٩- **ما ميزة الملاحظة المباشرة على ما عداها من وسائل التقويم؟**
أ- يمكن تطبيقها في مواقف الحياة الطبيعية.
ب- اقتصادية وفعالة.
ج- تعرفنا بالدوافع الداخلية للسلوك.
د- يسهل تسجيلها باستخدام مقاييس التقدير.

٢٠- **إذا حصل محمد علي الدرجة ٣٥ في اختبار اللغة العربية وكان المتوسط الحسابي لدرجات فصله ٤١، والانحراف المعياري للدرجات ٣، فأي مما يلي يعبر عن الدرجة المعيارية لهذا الطالب؟**
أ- + ٢,٧.
ب- - ٢.
ج- + ٣.
د- - ٣.

٢١- **تتضح العلاقة بين التقويم والأهداف في:**
أ- أن الأهداف هي المعايير التي يبنى عليها التقويم أحكامه.
ب- أن الأهداف جزء من عملية التقويم.
ج- أنه لابد من صياغة الأهداف قبل إجراء عملية التقويم.
د- أن التقويم جزء من الأهداف.

٢٢- **في أي المجالات التالية يمكن أن نجد أخطاء القياس:**

أ- في مجال التربية.

ب- في مجال العلوم.

ج- في مجال الاجتماعيات.

د- في جميع المجالات.

٢٣- **لكي يكون الهدف السلوكي هدفاً سليماً يجب أن:**

أ- يكون عاماً بحيث يمكن تطبيقه في مختلف مواقف التعليم.

ب- يحدد العناصر الهامة للمادة الدراسية.

ج- يكون قابلاً للتحويل إلى إجراءات خاصة يمكن ملاحظتها.

د- يركز على إعداد المتعلم لأنشطة معينة.

٢٤- **إذا زاد معامل ثبات الاختبار، فإن الخطأ المعياري:**

أ- يظل كما هو.

ب- ينقص.

ج- يزداد.

د- يزداد إذا انخفض الانحراف المعياري فقط.

٢٥- **عند استخدام الدرجات المدرسية لإبلاغ أولياء الأمور بنتائج أبنائهم، فإن هذه الدرجات تصل إلى أقصى درجة من الفاعلية عندما تكون:**

أ- تقويماً عاماً لتحصيل الطالب وشخصيته.

ب- تقويماً لتحصيل الطالب مع أخذ ميوله في الاعتبار.

ج- تقويماً لتحصيل الطالب في ضوء قدراته العقلية.

د- مقياساً صادقاً للدرجة التي حقق بها الطالب الأهداف التربوية.

٢٦- **تتضح العلاقة بين التقويم والقياس في أن:**

أ- القياس والتقويم وجهان لعملية واحدة.

ب- التقويم يؤدي وظيفته بفاعلية عندما يستعين بالقياس.

ج- القياس يوجه نحو صفات لابد من تقويمها.

د- التقويم يسبق القياس في العملية التربوية.

٢٧- **تمتاز الاختبارات محكية المرجع على الاختبارات معيارية المرجع في أنها:**

أ- تحدد الوضع النسبي للطالب مقارنة بزملائه.

ب- أكثر صدقاً وثباتاً من الاختبارات معيارية المرجع.

ج- تحدد مستوى الطالب دون الرجوع إلى محك خارجي.

د- تهتم بقياس الأهداف التربوية أكثر من الاختبارات معيارية المرجع.

٢٨- **أي مما يأتي يجب وضعه في الاعتبار الأول عند بناء اختبار تحصيلي؟**

أ- أهداف المنهج الدراسي.

ب- مستوى تحصيل الطلبة في المدرسة.

ج- أنواع الأسئلة التي يجب استخدامها.

د- الجزء الذي تم تدريسه من المنهج.

٢٩- **أي مما يأتي أكثر انطباقاً على الأسئلة المقالية بالمقارنة بالأسئلة الموضوعية؟**

أ- أفضل لقياس مخرجات التعلم المتعلقة بالتذكر.

ب- أقدر على كشف العيوب في طريقة التعبير.

ج- أكثر ثباتاً من الأسئلة الموضوعية.

د- يمكن استخدامها لكشف المفاهيم العلمية الخاطئة.

٣٠- **تصنف الأهداف التربوية إلى معرفية ووجدانية وحركية على أساس:**

أ- إمكانية اكتسابها من المدرس.

ب- الأهمية الاجتماعية للأهداف.

ج- مدى مناسبتها للمتعلمين.

د- نوع السلوك الذي يظهره المتعلم.

٣١- **أي مخرجات التعلم التالية يمكن تقويمها بفاعلية باستخدام الواجبات المنزلية بدلاً من اختبارات الورقة والقلم؟**

أ- القدرة على بناء خطة بحث لمشروع معين.

ب- المهارة في تفسير مجموعة من البيانات المعطاه.

ج- الحكم على صدق بعض القضايا.

د- المهارة في حل بعض المسائل الحسابية.

٣٢- **عند إعداد قائمة مراجعة لتقويم أداء التلميذ في المعمل يجب علينا أولاً:**

أ- تحديد أسلوب سليم للتقويم.

ب- تعريف المصطلحات الغامضة.

ج- تحليل الأداء المراد تقويمه بعناية لتحديد عناصر العمل.

د- تحديد عدد الفئات التي يمكن عن طريقها التمييز بين مستويات الأداء تمييزاً ثابتاً.

٣٣- **الاختبارات الموضوعية أكثر ثباتاً من اختبارات المقال وذلك بسبب:**

أ- سرعة وسهولة تقدير الدرجات.

ب- احتمال اشتمالها على عينة أكثر تمثيلاً لمحتوى المقرر الدراسي.

ج- استخدام تعليمات مقننة عادة.

د- إمكانية ضبط الظروف التي يستجيب فيها الطالب للاختبارات الموضوعية.

٣٤- **يزداد صدق مقاييس التقدير عند تحديد نقاط المقياس باستخدام:**

أ- الأرقام.

ب- أسماء الصفات.

ج- عبارات وصفية للسلوك.

د- مصطلحات معرفة مثل: "دائماً ــ أحياناً ــ نادراً".

٣٥- **الهدف الأمثل الذي يجب أن يحققه المعلمون من التقويم هو معرفة:**

أ- التلاميذ المتفوقين وغير المتفوقين في المادة الدراسية.

ب- أوجه النقص في أساليب التقويم التربوي.

ج- مدى تحقيق التلاميذ للمستوى اللازم في المادة الدراسية.

د- أنواع التغيرات التي حدثت في سلوك التلاميذ.

٣٦- **أي الأفعال التالية يجب استخدامه عند صياغة الأهداف التربوية الخاصة؟**

أ- يكتب.

ب- يفهم.

ج- يستقبل.

د- يتعلم.

٣٧- **إذا أردنا قياس قدرة التلميذ على الفهم، فأي مما يأتي لابد من توفره حتى يكون السؤال مقياساً صادقاً للفهم؟**

أ- أن يكون الموقف المتضمن في السؤال مبنياً على رسم بياني.

ب- أن يكون السؤال مركباً.

ج- أن يكون الموقف المستخدم في السؤال جديداً على التلميذ.

د- أن يكون السؤال شاملاً لمعظم عناصر الموضوع.

٣٨- **من أهم الأشياء التي يحققها بناء جدول مواصفات الاختبار التحصيلي:**

أ- تحديد طول الاختبار.

ب- صدق الاختبار.

ج- موضوعية الأسئلة.

د- تنوع الأسئلة.

٣٩- **إذا أردنا بناء مقياس لقياس اتجاهات الطلبة نحو العلم فإن الخطوة الأولى التي يجب القيام بها هي:**

أ- عمل جدول مواصفات للمقياس.

ب- تحديد أهداف العلم.

ج- تحديد الدرجات التي تقدر بها كل فقرة من فقرات المقياس.

د- تعريف المقصود "باتجاهات الطلبة".

٤٠- وضع باحث اختباراً في القدرة العددية لطلبة المرحلة الثانوية، وقام بتطبيق هذا الاختبار على عينة من الطلبة في بداية العام الدراسي، وفي نهاية العام حصل على درجات نفس الطلبة في امتحان الرياضيات. وبلغ معامل الارتباط بين اختبار القدرة العددية وامتحان الرياضيات ٠,٦٠، أي التفسيرات التالية أفضلها لمعامل الارتباط الذي حصل عليه الباحث؟

أ- يبلغ معامل صدق اختبار القدرة العددية بالطريقة التنبؤية ٠,٦٠.

ب- يبلغ معامل ثبات اختبار القدرة العددية ٠,٦٠.

ج- يبلغ معامل صدق اختبار القدرة العددية بالطريقة التلازمية ٠,٦٠.

د- يبلغ معامل ثبات اختبار الرياضيات ٠,٦٠.

٤١- يرتبط الخطأ المعياري للقياس ارتباطاً مباشراً:

أ- بصدق القياس.

ب- بثبات القياس.

ج- بموضوعية القياس.

د- بتحيز القياس.

٤٢- أي الصفات التالية يمكن قياسها قياساً ثابتاً باستخدام مقاييس التقدير؟

أ- النظافة.

ب- القدرة على قيادة الجماعة.

ج- المواطنة الصالحة.

د- التوافق المدرسي.

٤٣- أي الأهداف التالية لا تنتمي لنفس المجال الذي تنتمي إليه الأهداف الأخرى؟

أ- يذكر.

ب- يحلل.

ج- يؤلف.

د- يستقبل.

٤٤- يفضل استخدام قوائم المراجعة عندما تكون الصفة التي نرغب في تقديرها:

أ- متعددة التقدير.

ب- ثنائية الاستجابة.

ج- متغيرة.

د- ثابتة.

٤٥- **يرتبط صدق الاختبار ارتباطاً مباشراً:**

أ- بدقة تمثيل الاختبار للسمة التي وضع لقياسها.

ب- بالموضوعية في تقدير درجات الطلاب.

ج- بقياس الاختبار لأهداف تربوية متنوعة.

د- بمدى جودة وفاعلية المفردات التي يتضمنها الاختبار.

٤٦- **درست ثلاث خطوات عامة للقياس، وثاني هذه الخطوات هي تحديد:**

أ- الوحدات التي تبين النتائج.

ب- الخاصية التي نهتم بها.

ج- إجراءات إظهار الصفة.

د- الصفة التي نريد قياسها.

٤٧- **تميل معظم مقاييس التقدير إلى أن تكون أدوات تقويم ضعيفة لأن:**

أ- قياس السلوك قياساً دقيقاً عملية غير ممكنة.

ب- المصطلحات المستخدمة بها عادة ما تكون غامضة.

ج- صدقها منخفض.

د- المقدرين نادراً ما يتفقون فيما بينهم.

٤٨- **أهم عامل يجب تحقيقه عند اختيار أو بناء وسيلة للتقويم أن تكون الوسيلة:**

أ- فعالة.

ب- موضوعية.

ج- شاملة لجميع الأهداف.

د- صادقة.

٤٩- **من الضروري حتى يحقق التقويم فاعليته بالنسبة للمدرس والمتعلم:**

أ- أن تكون عملية التقويم منظمة.

ب- إبلاغ المتعلمين بنتائج التقويم فوراً.

ج- أن يتعاون المدرس والمتعلم في عملية التقويم.

د- استخدام أساليب تقويم متنوعة.

٥٠- **قام مدرس رياضيات بتطبيق اختبار تحصيلي على (٦٠) طالب، وقد بلغ عدد الطلاب الذين أجابوا على السؤال الأول من الاختبار إجابة صحيحة (٣٠) طالب، فكم يبلغ معامل صعوبة هذا السؤال؟**

أ- ٧٥%.

ب- ٦٠%.

ج- ٥٠%.

د- ٣٠%.

تعليمات:

استخدم البدائل التالية لتحديد مستوى المقياس المناسب في الفقرات من ٥١ – ٥٤:

أ- الموازين الاسمية.

ب- موازين الرتبة.

ج- موازين المسافة.

د- موازين النسبة.

٥١- إعطاء رقم لكل كلية من كليات الجامعة المختلفة.

٥٢- تزن سمية ٦٠ كيلو جراماً.

٥٣- حصل أحمد على ٢٥ درجة في اختبار اللغة العربية.

٥٤- مستوى تعليم حامد ٣.

تعليمات:

في الأسئلة من ٥٥ ــ ٦٠ المطلوب منك تقويم كل هدف وفقاً للمعايير التالية:

أ- إذا كان الهدف مصاغاً في عبارات تعبر عن سلوك ظاهر.

ب- وجود أكثر من ناتج تعليمي في عبارة الهدف.

ج- إذا كان الهدف يحدد موضوعات التعلم بدلاً من ناتج التعلم.

د- إذا كان الهدف مصاغاً في عبارات غير قابلة للملاحظة المباشرة.

٥٥- يفهم معنى قصيدة من الشعر.

٥٦- يكتب ما يملى عليه كتابة صحيحة خالية من الأخطاء.

٥٧- يشرح مفهوم التعلم ويذكر شروطه الأساسية.

٥٨- يخرج باستنتاجات سليمة من البيانات المعطاه له.

٥٩- يتقن التعبير الشفوي والتعبير الكتابي.

٦٠- قوانين الكهرباء.

تعليمات:

أعد اختبار مكون من ٥٠ سؤال في اللغة العربية لطلبة المرحلة الثانوية. وأثناء إعداد الاختبار للاستخدام العام قام واضعوا الاختبار بعدة خطوات نذكرها في الأسئلة من ٦١ ــ ٦٤ وتهتم هذه الخطوات بتحقيق مجموعة من الأهداف وهي:

أ- تحقيق الصدق بطريقة المحتوى.

ب- تحقيق الصدق بطريقة التمايز الطرفي.

ج- تحقيق استقرار درجات الاختبار.

د- تحقيق التناسق الداخلي للاختبار.

ضع علامة في ورقة الإجابة المنفصلة على الحرف الذي يمثل الهدف الذي تحققه كل خطوة من الخطوات من ٦١ ــ ٦٤.

٦١- بعد تطبيق الاختبار تم تقسيمه إلى نصفين وأجرى معامل الارتباط بين النصفين.

٦٢- قام واضعوا الاختبار ببناء جدول مواصفات قبل كتابة أسئلة الاختبار.

٦٣- أعيد إجراء الاختبار على عينة من ٢٠٠ طالب وطالبة بعد مرور أسبوعين على الإجراء الأول.

٦٤- عرض الاختبار على سبعة من المحكمين لأخذ رأيهم في الاختبار.

تعليمات:

طبق اختبار في الاجتماعيات مكون من ٤٠ سؤال من نوع الاختيار من متعدد على طلبة الصف الثالث الإعدادي. وكان زمن الاختبار ٣٠ دقيقة. وكان معامل الثبات بطريقة التجزئة النصفية ٠,٨٥ ما الأثر المحتمل الذي يمكن أن تحدثه التغيرات المبينة في الأسئلة من ٦٥-٦٧ على قيمة معامل الثبات المذكور. حدد إجابتك على ورق الإجابة المنفصلة وفقاً لما يلي:

أ- إذا كان التغير يؤدي إلى زيادة معامل الثبات.

ب- إذا كان التغير يؤدي إلى خفض معامل الثبات.

ج- إذا لم يكن من الممكن تحديد هل يزيد الثبات أو ينقص.

٦٥- حسب معامل الثبات باستخدام طريقة الصور المتكافئة بعد تطبيق صورتين من الاختبار على طلبة الصف الثالث الإعدادي.

٦٦- زيد عدد أسئلة الاختبار إلى ٨٠ سؤالاً وزيد زمن الاختبار إلى ساعة ونصف.

٦٧- تغير الاختبار ليضم ٢٠ سؤالاً من أسئلة الإكمال و٢٠ سؤالاً من نوع الاختيار من متعدد.

تعليمات:

تخير مما يأتي وسيلة القياس الأفضل في تقويم مخرجات التعلم الموجودة بالأسـئلة من ٦٨-٧٠:

أ- الملاحظة المباشرة.

ب- قوائم المراجعة.

ج- الاستبيان (الاستفتاء).

د- مقاييس التقدير.

٦٨- تقويم قدرة الطالب على استخدام المجهر.

٦٩- تقويم الطلاب لأداء المعلم في نهاية الفصل الدراسي.

٧٠- تشخيص نوع صعوبة النطق التي يعاني منها الطفل.

الملحق (٣)

نموذج ورقة الإجابة المنفصلة للاختبار في صورته النهائية

اسم الطالب: ---------- القسم: ----------

١	أ	ب	ج	د	٢٤	أ	ب	ج	د	٤٧	أ	ب	ج	د
٢	أ	ب	ج	د	٢٥	أ	ب	ج	د	٤٨	أ	ب	ج	د
٣	أ	ب	ج	د	٢٦	أ	ب	ج	د	٤٩	أ	ب	ج	د
٤	أ	ب	ج	د	٢٧	أ	ب	ج	د	٥٠	أ	ب	ج	د
٥	أ	ب	ج	د	٢٨	أ	ب	ج	د	٥١	أ	ب	ج	د
٦	أ	ب	ج	د	٢٩	أ	ب	ج	د	٥٢	أ	ب	ج	د
٧	أ	ب	ج	د	٣٠	أ	ب	ج	د	٥٣	أ	ب	ج	د
٨	أ	ب	ج	د	٣١	أ	ب	ج	د	٥٤	أ	ب	ج	د
٩	أ	ب	ج	د	٣٢	أ	ب	ج	د	٥٥	أ	ب	ج	د
١٠	أ	ب	ج	د	٣٣	أ	ب	ج	د	٥٦	أ	ب	ج	د
١١	أ	ب	ج	د	٣٤	أ	ب	ج	د	٥٧	أ	ب	ج	د
١٢	أ	ب	ج	د	٣٥	أ	ب	ج	د	٥٨	أ	ب	ج	د
١٣	أ	ب	ج	د	٣٦	أ	ب	ج	د	٥٩	أ	ب	ج	د
١٤	أ	ب	ج	د	٣٧	أ	ب	ج	د	٦٠	أ	ب	ج	د
١٥	أ	ب	ج	د	٣٨	أ	ب	ج	د	٦١	أ	ب	ج	د
١٦	أ	ب	ج	د	٣٩	أ	ب	ج	د	٦٢	أ	ب	ج	د
١٧	أ	ب	ج	د	٤٠	أ	ب	ج	د	٦٣	أ	ب	ج	د
١٨	أ	ب	ج	د	٤١	أ	ب	ج	د	٦٤	أ	ب	ج	د
١٩	أ	ب	ج	د	٤٢	أ	ب	ج	د	٦٥	أ	ب	ج	د
٢٠	أ	ب	ج	د	٤٣	أ	ب	ج	د	٦٦	أ	ب	ج	د
٢١	أ	ب	ج	د	٤٤	أ	ب	ج	د	٦٧	أ	ب	ج	د
٢٢	أ	ب	ج	د	٤٥	أ	ب	ج	د	٦٨	أ	ب	ج	د
٢٣	أ	ب	ج	د	٤٦	أ	ب	ج	د	٦٩	أ	ب	ج	د
										٧٠	أ	ب	ج	د

الملحق (٤)

الفرق بين نسب الإجابات الصحيحة في المجموعتين (العليا والدنيا) على جميع مفردات الاختبار في ضوء نتائج الدراسة الأولية(**)

المفردة	المجموعة	بدائل الإجابة				الفرق بين النسبتين
		أ	ب	ج	د	
١	العليا ٤٢	٢٦	١٣*	٣	صفر	٠,٠٧
	الدنيا ٤٢	٢٦	١٠	٢	٤	
٢	العليا ٤٢	٦	١٨	٩	٩*	٠,١٢
	الدنيا ٤٢	٤	٢٧	٧	٤	
٣	العليا ٤٢	٢	صفر	٣٧*	٣	٠,٥
	الدنيا ٤٢	٦	٥	١٦	١٤	
٤	العليا ٤٢	٢٦*	٢	٥	٩	٠,٢
	الدنيا ٤٢	١٩	٦	١٠	٧	
٥	العليا ٤٢	٢	١	٩	٣٠*	٠,١٢
	الدنيا ٤٢	٣	٧	٧	٢٥	
٦	العليا ٤٢	٣٢*	٩	صفر	١	٠,١٢
	الدنيا ٤٢	٢٧	١٤	صفر	١	

(**) استفادت الباحثة من استخدام الطريقة اليدوية التقليدية في حساب الفروق بين نسب الإجابات الصحيحة في المجموعتين (العليا والدنيا)، في تحديد المشتتات غير الفعالة والتي تم اختيارها بنسبة أعلى في المجموعة العليا بالمقارنة بالمجموعة الدنيا والتي تسببت في انخفاض قيم معاملات التمييز لبعض مفردات الاختبار، حيث تم تعديل هذه المشتتات، وإعادة إجراء الدراسة الأولية مرة أخرى، وقد تحسنت قيم معاملات التمييز في المرة التالية(***).

(*) تشير إلى البديل الصحيح لكل مفردة.

(***) ملحق (٦).

تابع ملحق (٤)

المفردة	المجموعة	بدائل الإجابة				الفرق بين النسبتين
		أ	ب	ج	د	
٧	العليا	٩	٦	٨	*١٩	٠,٢
	الدنيا	١١	٩	١٠	١٢	
٨	العليا	صفر	*٤٠	١	١	٠,٤٣
	الدنيا	٥	٢٢	٥	١٠	
٩	العليا	٣	٨	*١٩	١٢	٠,٢٦
	الدنيا	٥	١٦	٨	١٣	
١٠	العليا	صفر	صفر	*٧	٣٥	-٠,٠٢٤
	الدنيا	١	٦	٨	٢٧	
١١	العليا	٣	١١	صفر	*٢٨	٠,١٩
	الدنيا	٦	١٣	٣	٢٠	
١٢	العليا	*٣٣	٣	١	٤	٠,٠٧١
	الدنيا	٣٠	٢	١	٩	
١٣	العليا	٣	١١	١٠	*١٨	٠,١
	الدنيا	٥	١٦	٧	١٤	
١٤	العليا	٥	*٣٢	٣	٢	٠,٣٦
	الدنيا	١٣	١٧	٦	٦	
١٥	العليا	صفر	٥	*٢١	١٦	٠,٣٣
	الدنيا	١	١٠	٧	٢٤	
١٦	العليا	*٤١	صفر	١	صفر	٠,٤٨
	الدنيا	٢١	١٢	١	٨	
١٧	العليا	صفر	*٤١	صفر	صفر	٠,٤٥
	الدنيا	٤	٢٢	٦	١٠	
١٨	العليا	٩	٢	*٢٥	٦	٠,٥
	الدنيا	١٣	٩	٤	١٤	
١٩	العليا	صفر	*٣١	٣	٨	٠,٣٣
	الدنيا	١١	١٧	٦	٨	

تابع ملحق (٤)

المفردة	المجموعة	بدائل الإجابة				الفرق بين النسبتين
		أ	ب	ج	د	
٢٠	العليا	٢٨*	٤	٩	١	٠,١٤
	الدنيا	٢٢	١٤	٥	١	
٢١	العليا	١٢	١	١	٢٨*	٠,٣٨
	الدنيا	١٦	٨	٦	١٢	
٢٢	العليا	٤	١٨	١٤*	٦	٠,٢٤
	الدنيا	٥	٣١	٤	٢	
٢٣	العليا	٤	٣٢*	صفر	٦	٠,٢١
	الدنيا	٩	٢٣	٢	٨	
٢٤	العليا	٦*	٢	١٣	٢١	٠,٠٤٧
	الدنيا	٤	٧	٦	٢٥	
٢٥	العليا	١٧	صفر	٥*	٢٠	٠,٠٤٨
	الدنيا	٢٢	٢	٣	١٥	
٢٦	العليا	١٩*	٣	١٤	٦	٠,١٩
	الدنيا	١١	١١	١٤	٦	
٢٧	العليا	٣	١	١٢	٢٦*	٠,٣٨
	الدنيا	١٠	٩	١٢	١٠	
٢٨	العليا	٦	٣٤*	صفر	٢	٠,٢
	الدنيا	٣	٢٧	٩	٣	
٢٩	العليا	٣٩*	٢	١	صفر	٠,٣٦
	الدنيا	٢٤	٥	٧	٥	
٣٠	العليا	٣٢	٢	٢	٦*	-٠,٠٤٨
	الدنيا	٢٠	٣	١١	٨	
٣١	العليا	٣	٢١*	٣	١٥	٠,٠٤٨
	الدنيا	٩	١٩	٥	٨	
٣٢	العليا	٣	٦	٣١*	٢	٠,٢٨
	الدنيا	١٠	٨	١٩	٥	

تابع ملحق (٤)

المفردة	المجموعة	بدائل الإجابة				الفرق بين النسبتين
		أ	ب	ج	د	
٣٣	العليا	٢	١٠	٢٥	٥*	-٠,٠٢٤
	الدنيا	٤	٥	٢٧	٦	
٣٤	العليا	٢٧*	٥	١	٩	٠,٢٦
	الدنيا	١٦	٨	٨	١٠	
٣٥	العليا	٣	٣٣*	٣	٣	٠,١٢
	الدنيا	٢	٢٨	٣	٩	
٣٦	العليا	١٠	٦	٢٥*	١	٠,٢١
	الدنيا	١٩	٢	١٦	٥	
٣٧	العليا	١٧*	٢	٢٢	١	٠,١٤
	الدنيا	١١	٢	٢٨	صفر	
٣٨	العليا	٦	٢	صفر	٣٤*	٠,٤٣
	الدنيا	١٢	١٠	٤	١٦	
٣٩	العليا	صفر	٩	٣٣*	صفر	٠,٣٣
	الدنيا	٧	١٠	١٩	٦	
٤٠	العليا	٧*	٢	١	٣٢	٠,١
	الدنيا	٤	٣	٦	٢٩	
٤١	العليا	٢	١١	٢٧*	٢	٠,١٢
	الدنيا	٨	٥	٢٢	٧	
٤٢	العليا	٢	٣٣*	١	٦	٠,٣٦
	الدنيا	٤	١٨	٥	١٥	
٤٣	العليا	١٣*	١٣	١٠	٦	٠,٢٤
	الدنيا	٣	١٦	١٢	١١	
٤٤	العليا	صفر	٢	١٨*	٢٢	٠,١٤
	الدنيا	٨	٤	١٢	١٨	
٤٥	العليا	١٨	٣	١	١٨*	٠,٣٨
	الدنيا	٢٤	٨	٨	٢	

تابع ملحق (٤)

المفردة	المجموعة	بدائل الإجابة				الفرق بين النسبتين
		أ	ب	ج	د	
٤٦	العليا	١٥	٤*	٢٣	صفر	٠,١٧-
	الدنيا	١٠	١١	٢٠	١	
٤٧	العليا	٧	٨	٢٧*	صفر	٠,٠٩٥
	الدنيا	١٠	٧	٢٣	٢	
٤٨	العليا	٢٠*	٣	١٩	صفر	٠,٠٩٥
	الدنيا	١٦	٧	١٩	صفر	
٤٩	العليا	٩	٢٥*	٨	صفر	٠,٤
	الدنيا	١٣	٨	٢١	صفر	
٥٠	العليا	١	٨	صفر	٣٣	٠,٤
	الدنيا	١	١٩	٦	١٦	
٥١	العليا	٨*	١٦	١١	٧	٠,٢
	الدنيا	٣	١٦	١٠	١٣	
٥٢	العليا	١١	١	١٦*	١٤	٠,٠٥
	الدنيا	٩	٢	١٤	٧	
٥٣	العليا	٢	٧	١٦	١٧*	٠,٠٢
	الدنيا	٢	٦	١٨	١٦	
٥٤	العليا	٣	٢١*	٢	١٦	٠,٠١٩
	الدنيا	٥	١٣	٧	١٧	
٥٥	العليا	٦	٢	٨	٢٦*	٠,٢
	الدنيا	٧	٥	١١	١٩	
٥٦	العليا	٣٠*	٢	٣	٦	٠,٣٨
	الدنيا	١٤	٩	١٣	٦	
٥٧	العليا	١	٣٤*	صفر	٧	٠,١٩
	الدنيا	٣	٢٦	٣	١٠	
٥٨	العليا	٣	١	٢	٣٦*	٠,٢١
	الدنيا	٣	٦	٦	٢٧	

تابع ملحق (٤)

المفردة	المجموعة	بدائل الإجابة				الفرق بين النسبتين
		أ	ب	ج	د	
٥٩	العليا	٣	١١*	١٨	١٠	٠,١٤
	الدنيا	٤	٥	٢٢	١١	
٦٠	العليا	٣٤*	١	٣	٤	٠,٣٨
	الدنيا	١٨	٣	٢	١٨	
٦١	العليا	١	١٢	٩	٢٠*	٠.٣٣
	الدنيا	٨	١٢	١٤	٦	
٦٢	العليا	٢	٣٨*	١	١	٠,٢٦
	الدنيا	١١	٢٧	٣	١	
٦٣	العليا	٦	٩	٢٧*	صفر	٠,١٢
	الدنيا	١٥	٣	٢٢	٢	
٦٤	العليا	٤	٨	١٣	١٧*	-٠,٠٥
	الدنيا	٦	٧	١٠	١٩	
٦٥	العليا	٤	١٤*	٨	١٦	٠,١٩
	الدنيا	٤	٦	٩	٢٣	
٦٦	العليا	٢٤	٥	٥*	٨	٠,٠٢٤
	الدنيا	١٥	٤	٤	١٩	
٦٧	العليا	٣٦*	١	٢	٣	٠,٤٣
	للدنيا	١٨	١٣	٥	٤	
٦٨	العليا	٣٥	٢	٢	٣*	-٠,٠٧١
	الدنيا	١٩	٧	١٠	٦	
٦٩	العليا	١	٣١*	٤	٥	٠,٤
	الدنيا	٤	١٤	٦	١٦	
٧٠	العليا	٣	٦	٩	٢٤*	٠,٣٨
	الدنيا	٩	١٤	١١	٨	
٧١	العليا	٤	١٠	١٩*	٧	٠,٣٦
	الدنيا	١٧	١٠	٤	٩	

تابع ملحق (٤)

المفردة	المجموعة	بدائل الإجابة				الفرق بين النسبتين
		أ	ب	ج	د	
٧٢	العليا	٢٢*	١٢	٥	١	٠,٢٨
	الدنيا	١٠	١٥	١٣	٣	
٧٣	العليا	١٢	١٥*	١٢	٣	٠,٠٢٤
	الدنيا	١٠	١٦	٩	٧	
٧٤	العليا	٢١*	٣	٣	١٥	٠,٢٦
	الدنيا	١٠	٥	٦	٢١	
٧٥	العليا	٩	١٦	١٠*	٧	٠,٠٧١
	الدنيا	١٧	١٠	٧	٨	
٧٦	العليا	٢	٦	٢٩*	٥	٠,٥
	الدنيا	١٠	١٧	٨	٦	
٧٧	العليا	١٢	٢	٢٨*	صفر	٠,٣٣
	الدنيا	١٢	١٥	١٤	صفر	
٧٨	العليا	٢٤*	١٠	٧	١	٠,٣٨
	الدنيا	٨	١٣	٢٠	صفر	
٧٩	العليا	٦	١٤	٢٠*	١	٠,٣٣
	الدنيا	١٧	١٦	٦	٢	
٨٠	العليا	٨	٢*	٢٥	٤	-٠,٢١
	الدنيا	٦	١٤	١٧	٥	

الملحق (٥)

قيم معاملات السهولة والصعوبة لجميع مفردات الاختبار في ضوء نتائج الدراسة الأولية

المفردة	معامل السهولة	معامل الصعوبة	المفردة	معامل السهولة	معامل الصعوبة
١	٧٤ %	٢٦ %	٣٦	٤٤ %	٥٦ %
٢	٦٠ %	٤٠ %	٣٧	٣٢ %	٦٨ %
٣	٧٢ %	٢٨ %	٣٨	٤٢ %	٥٨ %
٤	٧٦ %	٢٤ %	٣٩	٥٢ %	٤٨ %
٥	٧٢ %	٢٨ %	٤٠	٢٦ %	٧٤ %
٦	٥٢ %	٤٨ %	٤١	٤٦ %	٥٤ %
٧	٧٤ %	٢٦ %	٤٢	١٨ %	٨٢ %
٨	٧٨ %	٢٢ %	٤٣	٦٦ %	٣٤ %
٩	٤٨ %	٥٢ %	٤٤	٣٤ %	٦٦ %
١٠	٦٠ %	٤٠ %	٤٥	٦٦ %	٣٤ %
١١	٦٦ %	٣٤ %	٤٦	٤٠ %	٦٠ %
١٢	٥٢ %	٤٨ %	٤٧	٢٦ %	٧٤ %
١٣	٧٤ %	٢٦ %	٤٨	٢٦ %	٧٤ %
١٤	٥٨ %	٤٢ %	٤٩	١٦ %	٨٤ %
١٥	٨٠ %	٢٠ %	٥٠	٧٤ %	٢٦ %
١٦	٨٠ %	٢٠ %	٥١	٥٠ %	٥٠ %
١٧	٦٢ %	٣٨ %	٥٢	٦٤ %	٣٦ %
١٨	٤٤ %	٥٦ %	٥٣	٣٠ %	٧٠ %
١٩	٦٨ %	٣٢ %	٥٤	٤٢ %	٥٨ %
٢٠	٦٦ %	٣٤ %	٥٥	٥٤ %	٤٦ %
٢١	٤٢ %	٥٨ %	٥٦	٥٦ %	٤٤ %
٢٢	٦٦ %	٣٤ %	٥٧	٦٢ %	٣٨ %
٢٣	٨٢ %	١٨ %	٥٨	٣٠ %	٧٠ %
٢٤	٧٩ %	٢٤ %	٥٩	٥٠ %	٥٠ %
٢٥	٤٦ %	٥٤ %	٦٠	٤٠ %	٦٠ %
٢٦	٤٦ %	٥٤ %	٦١	٤٤ %	٥٦ %
٢٧	٣٢ %	٦٨ %	٦٢	٤٠ %	٦٠ %
٢٨	٤٨ %	٥٢ %	٦٣	٥٢ %	٤٨ %
٢٩	٥٢ %	٤٨ %	٦٤	٣٦ %	٦٤ %
٣٠	٦٦ %	٣٤ %	٦٥	٤٨ %	٥٢ %
٣١	٤٢ %	٥٨ %	٦٦	٣٦ %	٦٤ %
٣٢	٦٤ %	٣٦ %	٦٧	٤٦ %	٥٤ %
٣٣	٤٢ %	٥٨ %	٦٨	٢٠ %	٨٠ %
٣٤	٤٢ %	٥٨ %	٦٩	٢٤ %	٧٩ %
٣٥	٤٨ %	٥٢ %	٧٠	٦٢ %	٣٨ %

تابع الملحق (٥)

ترتيب مفردات الاختبار وفقاً لمعاملات الصعوبة (من السهل إلى الصعب) بهدف إعداد صورتي الاختبار (أ، ب)

المفردة	معامل الصعوبة	المفردة	معامل الصعوبة	المفردة	معامل الصعوبة
٢٣	٨٢ %	١٤	٥٨ %	٥٤	٤٢ %
١٥	٨٠ %	٥٦	٥٦ %	٤٦	٤٠ %
١٦	٨٠ %	٥٥	٥٤ %	٦٠	٤٠ %
٨	٧٨ %	٦	٥٢ %	٦٢	٤٠ %
٤	٧٦ %	١٢	٥٢ %	٦٤	٣٦ %
٢٤	٧٦ %	٢٩	٥٢ %	٦٦	٣٦ %
١	٧٤ %	٣٩	٥٢ %	٤٤	٣٤ %
٧	٧٤ %	٦٣	٥٢ %	٢٧	٣٢ %
١٣	٧٤ %	٥١	٥٠ %	٣٧	٣٢ %
٥٠	٧٤ %	٥٩	٥٠ %	٥٣	٣٠ %
٣	٧٢ %	٩	٤٨ %	٥٨	٣٠ %
٥	٧٢ %	٢٨	٤٨ %	٤٠	٢٦ %
١٩	٦٨ %	٣٥	٤٨ %	٤٧	٢٦ %
١١	٦٦ %	٦٥	٤٨ %	٤٨	٢٦ %
٢٠	٦٦ %	٢٥	٤٦ %	٦٩	٢٤ %
٢٢	٦٦ %	٢٦	٤٦ %	٦٨	٢٠ %
٣٠	٦٦ %	٤١	٤٦ %	٤٢	١٨ %
٤٣	٦٦ %	٦٧	٤٦ %	٤٩	١٦ %
٤٥	٦٦ %	١٨	٤٤ %		
٣٢	٦٤ %	٣٦	٤٤ %		
٥٢	٦٤ %	٦١	٤٤ %		
١٧	٦٢ %	٢١	٤٢ %		
٥٧	٦٢ %	٣١	٤٢ %		
٧٠	٦٢ %	٣٣	٤٢ %		
٢	٦٠ %	٣٤	٤٢ %		
١٠	٦٠ %	٣٨	٤٢ %		

الملحق (٦)

قيم معاملات التمييز لمفردات الاختبار

(Corrected Item Total Correlation)

المفردة	معامل التمييز	المفردة	معامل التمييز	المفردة	معامل التمييز
١	٠,٣٦	٢٦	٠,٣٤	٥١	٠,٣١
٢	٠,٣٧	٢٧	٠,٢١	٥٢	٠,٤٨
٣	٠,٤٥	٢٨	٠,٣٦	٥٣	٠,٣٩
٤	٠,٤٤	٢٩	٠,٢٥	٥٤	٠,١٨
٥	٠,٥٣	٣٠	٠,٤٨	٥٥	٠,٤٤
٦	٠,٢٧	٣١	٠,٤٢	٥٦	٠,٤٨
٧	٠,٣٨	٣٢	٠,٤٩	٥٧	٠,٤٥
٨	٠,٤٦	٣٣	٠,٢٦	٥٨	٠,٢٩
٩	٠,٤٣	٣٤	٠,٤٤	٥٩	٠.٣١
١٠	٠,٤٣	٣٥	٠,٤٥	٦٠	٠،٤٥
١١	٠,٣٢	٣٦	٠,٤٢	٦١	٠,٤٩
١٢	٠,٢٦	٣٧	٠,٤٤	٦٢	٠,٢٩
١٣	٠,٤٣	٣٨	٠,٢١	٦٣	٠,٥١
١٤	٠,٥٩	٣٩	٠,٢٤	٦٤	٠,٤٨
١٥	٠,٢٩	٤٠	٠,٤٥	٦٥	٠,٤٩
١٦	٠,٢٩	٤١	٠,٤٧	٦٦	٠,٢١
١٧	٠.٤٩	٤٢	٠,٤١	٦٧	٠,٥٥
١٨	٠,٤٢	٤٣	٠,٣٨	٦٨	٠,٥٢
١٩	٠,٣٦	٤٤	٠,٥٤	٦٩	٠,٤٨
٢٠	٠,٣٤	٤٥	٠,٤٠	٧٠	٠,٣٠
٢١	٠,٥٦	٤٦	٠,٢٦		
٢٢	٠,٣٢	٤٧	٠,٣٨		
٢٣	٠,٣٦	٤٨	٠,٣٠		
٢٤	٠,٣٦	٤٩	٠,٢٩		
٢٥	٠,٢٤	٥٠	٠,٤٢		

ملحق (٧)

نتائج تحليل فاعلية المشتتات لجميع مفردات الاختبار

رقم المفردة	تكرارات اختيار البدائل لجميع أفراد العينة				متروك	المجموع
	أ	ب	ج	د		
١	٥	*٣٧	٤	٣	١	٥٠
٢	*٣٠	٣	٨	٩	–	٥٠
٣	٢	٤	٨	*٣٦	–	٥٠
٤	٦	٢	*٣٨	٤	–	٥٠
٥	*٣٦	٥	٦	٢	١	٥٠
٦	٩	*٢٦	٧	٨	–	٥٠
٧	٢	٣	٨	*٣٧	–	٥٠
٨	٢	٤	*٣٩	٥	–	٥٠
٩	٦	*٢٤	١٣	٦	١	٥٠
١٠	١١	١	٧	٣٠	١	٥٠
١١	*٣٣	٩	٥	٣	–	٥٠
١٢	٧	١٤	*٢٦	٢	١	٥٠
١٣	٢	*٣٧	٤	٦	١	٥٠
١٤	*٢٩	٣	٩	٨	١	٥٠
١٥	٢	٥	٣	*٤٠	–	٥٠
١٦	٣	٥	*٤٠	٢	–	٥٠
١٧	١٠	٦	٣	*٣١	–	٥٠
١٨	٤	١٨	*٢٢	٥	١	٥٠
١٩	*٣٤	٤	٤	٨	–	٥٠
٢٠	٥	*٣٣	٦	٥	١	٥٠
٢١	*٢١	١٠	١٥	٤	–	٥٠
٢٢	٢	٣	١٢	*٣٣	–	٥٠
٢٣	٤	٢	٤١	٣	–	٥٠

(*) تشير إلى أن هذا البديل يمثل الإجابة الصحيحة عن المفردة.

تابع ملحق (٧)

نتائج تحليل فاعلية المشتتات لجميع مفردات الاختبار

رقم المفردة	تكرارات اختيار البدائل لجميع أفراد العينة				متروك	المجموع
	أ	ب	ج	د		
٢٤	٢	*٣٨	٧	٣	–	٥٠
٢٥	٨	٩	٩	*٢٣	١	٥٠
٢٦	١٠	*٢٣	١١	٥	١	٥٠
٢٧	٨	٧	*٢٦	٩	–	٥٠
٢٨	*٢٤	١٠	٤	١٢	–	٥٠
٢٩	١٦	٢٦	٥	٢	١	٥٠
٣٠	٢	٧	٨	٣٣	–	٥٠
٣١	٢١	٣	٥	٢٠	١	٥٠
٣٢	٥	٥	*٣٢	٨	–	٥٠
٣٣	١١	*٢١	٥	١٣	–	٥٠
٣٤	*٢١	٤	٢١	٤	–	٥٠
٣٥	٤	٨	١٤	*٢٤	–	٥٠
٣٦	*٢٢	١٣	١٠	٤	١	٥٠
٣٧	٥	٨	*١٦	٢١	–	٥٠
٣٨	٩	*٢١	٨	١٢	–	٥٠
٣٩	١٠	١٠	٣	*٢٦	١	٥٠
٤٠	*١٣	١٥	١٦	٦	–	٥٠
٤١	١٠	*٢٣	٧	٩	١	٥٠
٤٢	*٩	١١	٦	٢٠	٤	٥٠
٤٣	٥	٧	٥	*٣٣	–	٥٠
٤٤	١٢	*١٧	١٦	٤	١	٥٠
٤٥	٣٣	٨	٤	٥	–	٥٠
٤٦	٦	٤	٢٠	١٨	٢	٥٠
٤٧	١٧	١٣	٨	١١	١	٥٠

تابع ملحق (٧)

نتائج تحليل فاعلية المشتتات لجميع مفردات الاختبار

رقم المفردة	تكرارات اختيار البدائل لجميع أفراد العينة				متروك	المجموع
	أ	ب	ج	د		
٤٨	٦	٨	٢٣	*١٣	–	٥٠
٤٩	١٢	*٨	١٤	١٦	–	٥٠
٥٠	٢	٥	*٣٧	٥	١	٥٠
٥١	*٢٥	١٦	٥	٤	–	٥٠
٥٢	١	٦	١١	٣٢	–	٥٠
٥٣	١٠	٩	*١٥	١٦	–	٥٠
٥٤	١٢	*٢١	١٠	٥	٢	٥٠
٥٥	٦	٣	١٤	*٢٧	–	٥٠
٥٦	*٢٨	٧	١١	٢	٢	٥٠
٥٧	٤	*٣١	١٠	٤	١	٥٠
٥٨	*١٥	١١	٧	١٤	٣	٥٠
٥٩	١١	*٢٥	١٢	٢	–	٥٠
٦٠	١٠	٣	*٢٠	١٥	٢	٥٠
٦١	٥	١٥	٧	*٢٢	١	٥٠
٦٢	*٢٠	٨	٧	١٥	–	٥٠
٦٣	٨	١٣	٢٦	٣	–	٥٠
٦٤	*١٨	١٤	١٠	٧	١	٥٠
٦٥	١٣	١١	*٢٤	١	١	٥٠
٦٦	*١٨	١٩	٣	١٠	–	٥٠
٦٧	١٦	٨	*٢٣	٣	–	٥٠
٦٨	*٢٨	١٠	٥	٧	–	٥٠
٦٩	٦	٤	٢٨	*١٢	–	٥٠
٧٠	*٣١	٩	٣	٦	١	٥٠

ملحق (٨)

معادلة الاختبار الإحصائى الذى تم الاعتماد عليه فى اختبار فروض الدراسة

"اختبار الدلالة الإحصائية للفروق بين معاملات الارتباط"

$$ى = \frac{٠,٥ \text{ لو } \frac{١ + ر_{١}}{١ - ر_{١}} - ٠,٥ \text{ لو } \frac{١ + ر_{٢}}{١ - ر_{٢}}}{\sqrt{\frac{١}{ن_{١} - ٣} + \frac{١}{ن_{٢} - ٣}}}$$

حيث إن: لو = اللوغاريتم الطبيعى (الأساس e).

٠,٥ = قيمة ثابتة.

ر١ = معامل ثبات الاختبار الأول.

ر٢ = معامل ثبات الاختبار الثانى.

ن١ = حجم العينة الأولى.

ن٢ = حجم العينة الثانية.

مأخوذة ببعض التصرف عن (سمير كامل عاشور، سامية أبو الفتوح، ٢٠٠٠).

ملخص الدراسة

مشكلة الدراسة:

تتمثل مشكلة الدراسة الحالية في معالجة تأثير مجموعة من العوامل أو المتغيرات على القيمة العددية لمعامل الثبات وهذه المتغيرات منها ما يرتبط بالاختبار نفسه، ومنها ما يتعلق بعينة المختبرين وتتضمن:

١- عدد المفردات التي يتألف منها الاختبار (طول الاختبار).

٢- موقع المفردة من الاختبار (تسلسل أو ترتيب عرض المفردات داخل الاختبار).

٣- عدد بدائل المفردة في أسئلة الاختيار من متعدد.

٤- حجم العينة التي يطبق عليها الاختبار.

فروض الدراسة:

في ضوء مشكلة البحث والتساؤلات التي يطرحها، والدراسات والبحوث التي تم الاطلاع عليها، صيغت فروض الدراسة الحالية على النحو التالي:

١- يختلف تقدير قيمة معامل ثبات الاختبار بالتغير في طول الاختبار (عدد مفرداته).

٢- يختلف تقدير قيمة معامل الثبات بتغير موقع المفردات من الاختبار (ترتيب المفردات داخل الاختبار).

٣- يختلف تقدير قيمة معامل الثبات بتغير عدد بدائل مفردات الاختيار من متعدد.

٤- يختلف تقدير قيمة معامل الثبات بالتغير في حجم عينة المختبرين.

منهج البحث:

أ- العينة:

تألفت عينة الدراسة الأولية من (١٥٦) طالب وطالبة من طلاب الدبلوم العام في التربية (نظام العامين)، وعينة الدراسة الأساسية من (١١٠) طالب وطالبة من طلاب الدبلوم الخاص في التربية (السنة التكميلية).

ب- أدوات الدراسة:

تكونت أدوات الدراسة من أربعة اختبارات تحصيلية من نوع الاختيار من متعدد في مادة القياس والتقويم المقررة على طلاب الدبلوم العام في التربية (نظام العامين)، والدبلوم الخاص (السنة التكميلية).

ج- إجراءات الدراسة:

قامت الباحثة بإجراء عملية التطبيق الفعلي لأدوات الدراسة تمهيداً لاختبار فروض الدراسة، وذلك وفقاً للخطوات التالية:

١- تم إجراء الدراسة الأولية للاختبار، حيث تم تطبيق الاختبار التحصيلي في صورته المبدئية على عينة تتألف من (١٥٦) طالب وطالبة من طلاب الدبلوم العام في التربية (نظام العامين)، بهدف إجراء التحليل الإحصائي لمفردات الاختبار.

٢- حسبت قيم معاملات الصعوبة ومعاملات التمييز، وفاعلية المشتتات لجميع مفردات الاختبار، تمهيداً لإعداد الاختبار في صورته النهائية.

٣- في ضوء نتائج الدراسة الأولية، تم إعداد أربعة صور اختبارية وهي (أ، ب، ج، د) بهدف اختبار أثر متغيرات الدراسة (عدد البدائل، موقع المفردة من الاختبار، عدد المفردات) على قيمة معامل الثبات.

٤- تم تطبيق الصور الاختبارية الأربعة (أ، ب، ج، د) على عينة تتألف من (١١٠) طالب وطالبة من طلاب الدبلوم الخاص في التربية (السنة التكميلية)، حيث تم تعيين صور الاختبار على أفراد العينة عشوائياً.

٥- حسبت قيم معامل الثبات (ألفا) لجميع صور الاختبار، في ضوء متغيرات الدراسة الحالية (عدد المفردات، ترتيب المفردات داخل الاختبار، عدد بدائل المفردة، حجم العينة).

٦- تم تطبيق اختبار الدلالة الإحصائية للفروق بين معاملات الارتباط في عينتين، على اعتبار أن معامل الثبات هو في حقيقته معامل ارتباط الاختبار بنفسه، وذلك تمهيداً لاختبار فروض الدراسة.

نتائج الدراسة:

١- توجد فروق دالة إحصائياً بين قيمتي معامل الثبات المحسوبة لصورتي الاختبار (أ، د) عند مستوى الدلالة (٠,٠٥) نتيجة لزيادة عدد مفردات الاختبار من (٤٠) مفردة إلى (٧٠) مفردة.

٢- توجد فروق دالة إحصائياً بين قيمتي معامل الثبات المحسوبة لصورتي الاختبار (أ، ب) نتيجة اختلاف أسلوب ترتيب مفردات الاختبار من حيث مستوى الصعوبة (من السهل إلى الصعب ــ ومن الصعب إلى السهل)، وكانت الفروق نتيجة لترتيب مفردات الاختبار من السهل إلى الصعب.

٣- توجد فروق دالة إحصائياً بين قيمتي معامل الثبات المحسوبة لصورتي الاختبار (أ، ج) نتيجة اختلاف عدد بدائل المفردة، من ثلاثة إلى أربعة بدائل للمفردة.

٤- لا توجد فروق دالة إحصائياً بين قيمتي معامل الثبات المحسوبة للاختبار نتيجة اختلاف حجم عينة المختبرين.

Cairo University
Institute of Education Studies
Educational Psychology Department

Examining the Effect of the Number of Items, Item Arrangement Sample Size, and Number of Alternatives on the Value of Measurement Reliability Coefficient

M.A thesis in Education
(Educational Psychology)

Submitted by
Reda Samir Awad

Demonstrator of Educational Psychology at the Institute of Education Studies Cairo University

Supervised by

Prof. Ragaa Mahmoud Abo-Alaam
Professor Non-emeritus of educational Psyhology- Instudies of educational studies
Cairo University

Prof. Nadia Mahmoud Sherif
Head of the educational Psyhology Department Instudies of educational studies
Cairo University

Dr. Sawsan Ebrahim Abo-El Ela
Lecturer of Psyhology- Instudies of educational studies
Cairo University

2004

Summary of The Study

Problem of The Study:

This study examined the effect of a group of factors or variables, some of which are related to the test and others are related to the testees, on the value of the test reliability coefficient.

These Variables Include:

1- The number of items included in the test (length of the test).

2- The arrangement of test items.

3- The number of multiple choice item alternatives; and

4- The size of the sample subjected to the test.

Hypotheses of The Study:

In the light of the problem and questions of the study, and the previous studies reviewed, the following hypotheses were formed:

1- Estimating the value of the test reliability coefficient will be affected by the length of the test (the number of its items).

2- Estimating the value of the test reliability coefficient will be affected by the arrangement of test items.

3- Estimating the value of the test reliability coefficient will be affected by multiple choice item alternatives.

4- Estimating the value of the test reliability coefficient will be affected by the number of testees subjected to the test (sample size).

Sample of The Study:

The primary sample of the study consisted of (15) students enrolled in the General Diploma of Education (the two – year system), while the main sample of the study consisted of (110) students enrolled in the Special Diploma of Education (the complementary – year system).

Tools of The Study:

The study used four multiple-choice achievement tests that were based on the Education Measurement and Evaluation Syllabus Studied by the Students enrolled in the General Diploma of Education (the two-year system) and the Special Diploma in Education (the complementary year system).

Procedures of The Study:

To test the hypotheses of the study, the researcher adopted the following procedures:

1- Conducting the pilot study of the test to make the statistical analysis of the test items; the achievement test in its intial form was administered to (156) students enrolled in General Diploma of Education (the two year system).

2- Calculating the coefficient of difficulty, coefficient of discrimination and effectiveness of distractors of all the items included in the test in order to develop (prepare) the test in its final form.

3- In the light of the pilot study, four equivalent forms of the test (A, B, C and D) were developed to examine the effect of the variables included in the study, the number of items, item arrangement and number of alternatives on the value of test reliability coefficent.

4- Administering the four equivalent forms of the test (A, B, C and D) randomly to a sample of (110) students enrolled in the Special Diploma of Education (the complementary – year system).

5- Calculating the values of the reliability coefficient for all the test forms using the Alpha Cronbach equation in the light of the variables of the study (the number of the items, item arangement, number of alternatives, and sample size).

6- Administering the statistical significance test to determine the differences among the coefficents of correlation in two samples, given that the coefficient of reliability is in reality the same as the coefficient of the test correlation, so that the hypoheses of the study can be tested.

Results of The Study:

1- There was a significant difference between the value of the reliability coefficient of the test form (A) at that of the test form (D) in favour of the form (A), consisted of (70) items due to the length of the test.

2- There was a significant difference between the value of the reliability coefficient of the test form (A) at that of the test form (B) in favour of the form (A), whose items were arranged form the easiest to the most difficulty, due to the item arrangement.

3- There was a significant difference between the value of the reliability coefficient of the test form (A) at that of the test form (C) in favour of the form (A), which has four alternatives, due to the number of alternatives.

4- There was no significant difference between the two values of the reliability coefficient of the test due to the sample size.