بناء اختبار تحصيلي لمادة علم نفس النمو وفقًا لنظرية القياس التقليدية

منتظر حميد كاظم المرهج

Muntadhar.hameed2302@ircoedu.uobaghdad.edu.iq

أ.م.د هند صبيح رحيم

hind.sabeeh@ircoedu.uobaghdad.edu.iq

جامعة بغداد/ كلية التربية ابن رشد للعلوم الإنسانية

الملخص

هدف البحث إلى بناء اختبار تحصيلي مقنن لمادة "علم نفس النمو" لطلبة الأقسام غير المتخصصة في كليات التربية، واستخلاص خصائصه السيكومترية (الصدق، الثبات، معاملات الصعوبة والتمييز، وفاعلية البدائل) على وفق نظرية القياس التقليدية (CTT). وقد جاء هذا لمعالجة الضعف الملاحظ في المواصفات السيكومترية للاختبارات التي يعدها المعلمون لغياب الأسس العلمية في بنائها.

اعتمد البحث المنهج الوصفي لإتمام إجراءات البحث، وتكون مجتمعه من طلبة المرحلة الثانية في الأقسام غير المتخصصة بكليات التربية بجامعتي بغداد والمستنصرية للعام الدراسي ٢٠٢٤-٢٠٢٥. واعتمد الباحث على عينة طبقية عشوائية قوامها (٤٠٠) طالب وطالبة للتحليل الإحصائي للفقرات.

تضمنت خطوات البناء: تحديد المحتوى (١٠ فصول)، وصياغة (٧٥) هدفاً سلوكياً، وبناء جدول مواصفات (خارطة اختبارية) لضمان صدق المحتوى، حيث وزعت الأهداف على مستويات بلوم المعرفية (التذكر ٤٠%، الفهم ٣٠%، التطبيق ٢٠%، التحليل ١٠%). تكون الاختبار من (٧٥) فقرة موضوعية من نوع الاختيار من متعدد (بأربعة بدائل). وتم التحقق من الصدق الظاهري وصدق المحتوى عبر عرضه على (١١) خبيراً، إذ اتفقوا بنسبة (١٠٠%) على صلاحية جميع الفقرات بعد إجراء تعديلات طفيفة.

بعد إتمام التحليل المنطقي والتحكيم، أصبح الاختبار بصيغته النهائية جاهزاً للتطبيق الميداني على عينة التحليل الإحصائي لاستخراج مؤشرات جودة الفقرات والخصائص السيكومترية النهائية للاختبار، واستنتج الباحث عدة نتائج منها، أن التوزيع ألاعتدالي للدرجات الكلية للأختبار التحصيلي لمادة علم نفس النمو، يشير الى تمثيل عينة البحث لمجتمع البحث. وان الاختبار

التحصيلي لمادة علم نفس النمو هو مناسب لأفراد العينة، وذلك لاتساق قدرات أفراد العينة في تحقيق هدف البحث الحالي الأول. وأخيرا فالخصائص السيكومترية لفقرات الاختبار وللاختبار كاملا، كانت جيدة في عمومها ومقبولة تربويا.

الكلمات المفتاحية: بناء اختبار، الاختبار التحصيلي، علم نفس النمو، نظرية القياس التقليدية.

Construction of an Achievement Test for the "Developmental Psychology" Course According to Classical Test Theory

Muntadher Hameed Kadhim Al-Murhej

Assistant Professor Dr. Hind Subaih Raheem

University of Baghdad College of Education Ibn Rushd for Human Sciences

Abstract

This research aimed to construct a standardized achievement test for the "Developmental Psychology" course, designed for students in non-specialized departments within Faculties of Education. The study also sought to determine its psychometric properties (validity, reliability, difficulty and discrimination indices, and distractor effectiveness) in accordance with Classical Test Theory (CTT). This initiative was undertaken to address the observed weaknesses in the psychometric characteristics of teacher-made tests, which often lack a scientific foundation in their construction.

The study adopted a descriptive methodology. The population comprised all second-stage students in non-specialized departments at the Faculties of Education at the University of Baghdad and Mustansiriyah University for the 2024-2025 academic year. A stratified random sample of (400) students was selected for the purpose of statistical item analysis.

The test construction process involved several steps: defining the content (10 chapters), formulating (75) behavioral objectives, and developing a table of specifications (test blueprint) to ensure content validity. These objectives were distributed across Bloom's cognitive

levels: Remembering (40%), Understanding (30%), Application (20%), and Analysis (10%). The test was composed of (75) objective, multiple-choice items, each with four alternatives. Face validity and content validity were established by submitting the test to a panel of (11) experts. A (100%) agreement was achieved on the suitability of all items following minor revisions.

Following logical analysis and expert validation, the final version of the test was administered to the statistical analysis sample to derive item quality indicators and the test's final psychometric properties. The findings indicated that the normal distribution of the total scores on the Developmental Psychology achievement test suggests the sample is representative of the population. Furthermore, the test was found to be suitable for the sample, demonstrating consistency with the sample members' abilities. In conclusion, the psychometric properties of both the individual items and the test as a whole were found to be generally good and educationally acceptable.

Keywords: Test construction, achievement test, developmental psychology, classical test theory.

الفصل الاول – التعريف بالبحث:

مشكلة البحث:

يسعى علماء القياس والمعنيون بالعلوم التربوية والنفسية للتوصل إلى الموضوعية في قياس سلوك الأفراد واستجاباتهم، فقد بات بناء الأداة التي تمكن الباحث من الوصول إلى التقدير الموضوعي للسلوك من الأهداف المهمة التي ينشدها علماء التربية وعلم النفس (كاظم، ١٩٨٨: ٧). إذ أن الاختبارات التحصيلية ركيزة أساسية من ركائز التقويم التربوي، وإذ ما أحسن بناؤها واستخدامها تساعد في تشخيص العديد من نواحي الضعف والصعوبات في التحصيل الدراسي، كما تفيد نتائجها في اتخاذ قرارات مهمة تمس نمو الطلاب ومستقبلهم بالإضافة إلى أهميتها في الحكم على فاعلية بقية عناصر العملية التعليمية.

ويتضح أن الواقع الحالي للتعليم الجامعي في العراق يعتمد بدرجة كبيرة على اختبارات التحصيل كوسيلة أساسية لتقويم الطلبة في جميع المراحل الدراسية وبمختلف التخصصات (الزاملي واخرون، ٢٠٠٩: ٤٦٧). فالمعلم يقيس تحصيل المتعلمين عن طريق إجراء اختبارات

لهم، وتعتبر العلامات التي يحصلون عليها وصفاً كميا أو تعبيراً رقمياً يمثل مقدار تحصيلهم، وأهم سمات القياس تكمن في تحديد مستوى الخاصية المراد دراستها، عن طريق الأرقام والرموز المستندة إلى قواعد، ثم يتم من خلالها إطلاق الحكم وهذا ما يطلق عليه بالتقويم (المندلاوي والمندلاوي، ٢٠٢١: ١٥). وبمعنى آخر يمكن القول بأن الاختبار التحصيلي قد يلعب دوراً رئيسياً في مختلف البرامج والمراحل التعليمية بدءاً من مرحلة الأساس إلى التعليم الجامعي. ومع ذلك ورغم أهمية وشيوع استخدام الاختبارات التحصيلية من قبل المعلمين إلا أنهم لم يحسنوا هذا الاستخدام لقلة أو عدم كفاية برامج التدريب الخاص بالاختبارات لهؤلاء المعلمين فيما يتعلق بكيفية بناء أو تصميم الاختبارات التحصيلية ونتيجة لذلك فقد اتضح أنها تفتقر إلى المقومات الأساسية للاختبارات ذات المواصفات الجيدة. (المياحي، ٢٠١١: ٨٧).

إذا فالاختبارات التي يقوم المعلم بإعدادها بنفسه في الصف، تتصف بأن معظمها قد لا يحقق الشروط المفروض توافرها في الاختبار الجيد بسبب قلة خبرة العديد من المعلمين بالأساليب العلمية الضرورية لصياغة الاختبارات التحصيلية، وهو ما يشكل مشكلة حقيقية في منظومة التعليم بشكل عام لاعتمادها وحاجتها لأدوات قياس جيدة لقياس وتقويم أداء المتعلمين وقدرتهم وإطلاق الأحكام القيمية عليهم كمتعلمين (الزاملي وآخرون، ٢٠٠٩: ٢٣٧).

ويمكن إيضاح مشكلة البحث اختصارا بالنقاط الآتية:

١.تعاني الاختبارات التحصيلية التي يعدها المعلم من ضعف في مواصفات الاختبار الجيد وبشكل خاص في الخصائص السيكومترية للاختبار.

٢.عدم وجود دراسات عراقية تناولت بناء اختبار تحصيلي لمادة علم نفس النمو لأقسام غير الاختصاص وفق الخطوات العلمية لبناء الاختبارات وأهمها التجريب الميداني وللكشف عن الخصائص السيكومترية للفقرات وللاختبار كاملا.

أهمية البحث Significance of the Research:

ان الاختبارات التحصيلية ركيزة أساسية من ركائز التقويم التربوي فإذا ما أحسن بناؤها واستخدامها ستساعد في تشخيص العديد من نواحي الضعف والصعوبات في التحصيل الدراسي، كما تفيد نتائجها في اتخاذ قرارات مهمة تمس نمو الطلاب ومستقبلهم بالإضافة إلى أهميتها في الحكم على فاعلية بقية عناصر العملية التعليمية. إذ يعد القياس والتقويم ضرورة لا غنى عنها في مختلف العلوم، إذ يشكلان أساسا لتطورها وتقدمها، ويعدان كذلك ركنا محوريا في العملية التربوية والتعليمية، حيث لا يستطيع المعلم أداء دوره بدون امتلاك مهارات القياس والاختبارات التحصيلية بشكل خاص (محاسنة، ٢٠١٣: ٢١). وبما أن فهم السلوك الإنساني يتطلب أدوات كمية دقيقة، فإن الاعتماد على الاختبارات والمقاييس يمثل مدخلاً أساسيا لذلك (الحجامي،

٢٠١١: ٦). وقد اعتمدت حركة القياس النفسي على نظرية القياس التقليدية بوصفها منطلقاً لبناء الاختبارات (سليمان وأبو علام، ٢٠١٢: ٦٨).

وعلى الرغم من الانتقادات التي وجهت للاختبارات التحصيلية التقليدية، فإنها لا تزال الوسيلة الأساسية للتقويم، مما يتطلب إجراء تطبيقاتها لمواكبة التطورات العلمية المختلفة (ميخائيل، ٢٠١٦: ١٢٣). إذ تعتمد نتائجها في اتخاذ قرارات تربوية مصيرية تؤثر على مستقبل الأفراد (سليمان وأبو علام، ٢٠١٠: ١٩١)، وتزداد أهميتها مع تقدم المتعلم في مراحل دراسته (عمر وآخرون، ٢٠٠٩: ٩٧). وتبرز هذه الاعتبارات ضرورة العناية بتحسين جودة الاختبارات التحصيلية من حيث خصائصها السيكومترية، وتحليل استجابات الطلبة على فقراتها بدقة (جاسم، ٢٠٢٣: ٣٦٧). ومن هنا تنبع أهمية هذا البحث في سعيه إلى بناء اختبار تحصيلي في مادة علم نفس النمو لأقسام غير الاختصاص، يتمتع بخصائص سيكومترية عالية الجودة.

وتأسيسا على ذلك، يرى الباحث أهمية نظرية وتطبيقية للبحث:

الأهمية النظرية (Theoretical Significance):

١.يساعد المعلمين على فهم الاسس النظرية التي ترتكز عليها نظرية القياس التقليدية في بنائها للاختبارات التحصيلية.

٢.يوضح البحث الحالي للمعلمين والباحثين والطلبة الأثر الإيجابي للاختبارات التحصيلية ذات الخصائص الجيدة في جميع جوانب العملية التعليمية.

الأهمية التطبيقية (Practical Significance):

١.يساعد المعلمين في المدارس على فهم إجراءات تصميم أسئلة موضوعية متعددة الخيارات بجودة عالية، من خلال الاطلاع والتعرف على الخطوات المنهجية والعلمية لبناء الاختبار التحصيلي في البحث الحالي.

٢.يساعد الباحثين في فهم كيفية تطبيق إجراءات التحليل الإحصائي للاختبار وفقراته.

أهداف البحث Aims of the Research:

١.بناء اختبار تحصيلي لمادة (علم نفس النمو – لغير الاختصاص) لعينة البحث الحالي.

٢.استخراج الخصائص السيكومترية للاختبار الحالي وفقا لنظرية القياس الكلاسيكية والتي تشمل: الصدق (Validity) – الثبات (Reliability) – معامل التمييز (Discrimination Index) – معامل الصعوبة (Difficulty Index) – فاعلية البدائل (Distractor Effectiveness).

حدود البحث Limitation of the Research: يتحدد البحث الحالي بـ:

- طلبة جامعة بغداد والجامعة المستنصرية / كليات التربية / المرحلة الثانية، لكلا الجنسين (ذكور – اناث) للعام الدراسي (٢٠٢٤–٢٠٢٥).

- مفردات مادة (علم نفس النمو لغير الاختصاص) كما وردت في (هيئة العمداء) التابعة لوزارة التعليم العالي والبحث العلمي.

تحديد المصطلحات Limitation of the Terms:

١.الاختبار التحصيلي Achievement Test:

- ايبل (1979) Ebel: اختبار صمم بقصد قياس إلمام الطالب في معرفة محددة أو براعته في مهارات معينة. (الطريري، ١٩٩٧: ٢٧٩).
- الكبيسي (٢٠٠٧): " وسيلة من وسائل القياس التي تستخدم لتدل على معرفة مستوى المتعلمين في مقرر معين" (الكبيسي، ٢٠٠٧: ١٠٧).

٢.الخصائص السيكومترية Psychometric Properties:

- علام (٢٠٠٠): هي الدلائل أو المؤشرات الإحصائية عن مدى جودة المقياس وفقراته، وتوجد خصائص قياسية للمقياس هي صدقه وثباته وحساسيته وشكل التوزيع التكراري للدرجات (رحيم، ٢٠٢١: ١٢).
- مراد وسليمان (٢٠٠٢): الخصائص الضرورية والمتعلقة بالصدق والثبات والمعايير والتي يتم حسابها بعد تجريب الاختبار على عينة ممثلة للمجتمع" (مراد وسليمان، ٢٠٠٢: ٣٥٠).
- عبيدات والجندي (٢٠١٤): "تلك الصفات الضرورية والمتعلقة بمدى فاعلية فقرات الاختبار، وكذلك بالصدق والثبات وما يرتبط بهما من معاملات تمييز ومستويات السهولة والصعوبة في حالة اختبارات التحصيل والقدرات ومعايير تفسير النتائج". (عبيدات والجندي، ٢٠٢٤: ٢٠٦).

الفصل الثاني – إطار نظري ودراسات سابقة:

أولا – الاختبارات التحصيلية (Achievement Tests):

تعد الاختبارات التحصيلية من أهم أدوات القياس والتقويم التربوي، إذ تستخدم لتحديد مدى إتقان الطلبة للمعلومات والمهارات المكتسبة من التعليم النظامي، وقياس فاعلية المناهج وطرائق التدريس (بركات، ٢٠١٨: ١٩١). وقد عرفها Gregory (1992) بأنها اختبارات تقيس التعلم في موضوع معين، وعرفها Mehrens & Lehman (1975) بأنها أداة تقيس مدى اكتساب الفرد للمعرفة نتيجة تعليم منظم. وهي لا تقتصر على قياس المعلومات، بل تمتد لقياس المهارات، الاتجاهات، والسلوك، مما يجعلها أداة شاملة للحكم على الأهداف التعليمية (الطريري، ١٩٩٧: ٢٨١).

تستند الاختبارات التحصيلية إلى ثلاث فرضيات رئيسة:

١.قياس سلوك محدد يمكن صياغته بشكل سلوكي قابل للقياس.

٢. التحقق من صدق المحتوى بحيث يقيس الأهداف التعليمية الدقيقة فقط.

٣. مراعاة الفروق الفردية بين المتعلمين لضمان عدالة التقييم (علام، ٢٠٠٦: ١٢٤).

❖**أهمية الاختبارات التحصيلية:**

- للطلبة: تساعد في تحديد مستوى التحصيل وكشف نقاط الضعف والقوة (بركات، ٢٠١٨: ٨٨).
- للمعلمين: تمكّنهم من تقويم أداء الطلبة وتحسين طرائق التدريس وتقديم تغذية راجعة فعالة
- للنظام التعليمي العام: تتيح تخطيطًا استراتيجيا وتحسينا لجودة التعليم ومتابعة تحقيق الأهداف.
- للمؤسسات التعليمية: تُستخدم في توجيه الطلبة، وتقييم البرامج والمناهج (زاير وآخرون، ٢٠٢٠: ١٧). (عبد الرحمن، ٢٠١١: ٢٢٣).

❖**أهداف وأغراض الاختبارات التحصيلية:**

تهدف الاختبارات التحصيلية إلى قياس مستوى التحصيل الدراسي للمتعلمين، وتحديد مدى استيعابهم للمفاهيم والمعارف المقررة، بالإضافة إلى تشخيص الفروق الفردية بينهم بما يتيح تصنيفهم بدقة ومتابعة تقدمهم الأكاديمي (أبو فودة وبني يونس، ٢٠١٢: ٣٦). وتعد هذه الاختبارات أداة أساسية في منح الدرجات والشهادات للطلبة، كما تسهم بشكل فاعل في التنبؤ بمستوى أدائهم المستقبلي. علاوة على ذلك، تؤدي دوراً محورياً في تحسين جودة التعليم والارتقاء بمستواه، من خلال تقديم مؤشرات دقيقة بشأن مدى تحقق الأهداف التعليمية وفاعلية البرامج الدراسية. كما تعتمد المؤسسات التربوية على نتائج هذه الاختبارات في مراجعة وتطوير المناهج، وتصميم خطط علاجية للمتعثرين، وضمان العدالة في توزيع الفرص التعليمية بين جميع المتعلمين (بركات، ٢٠١٨: ١٩١).

❖**أنواع الاختبارات التحصيلية:** تصنَّف الاختبارات وفق عدد من المعايير، أهمها:

من حيث المرجع التفسيري:

- معيارية المرجع (Norm-referenced): تقارن أداء الفرد بالمجموعة (الكبيسي، ٢٠٠٧: ١٠٤).
- محكية المرجع (Criterion-referenced): تُقاس وفق مستوى أداء محدد (عودة، ١٩٨٥: ٣٨).

من حيث أداة الإعداد:

- اختبارات المعلم (غير رسمية).
- اختبارات مقننة أعدها مختصون (ملحم، ٢٠١٢: ٥٢).

من حيث طبيعة الإجابة:

- اختبارات المقال (Essay Tests): وتعتبر اختبارات تتطلب من الطالب تقديم الإجابة، وتقيس التفكير والتحليل، لكنها تتأثر بعوامل ذاتية (المياحي، ٢٠١١: ١٠١).

• الاختبارات الموضوعية (Objective Tests): وتعتبر اختبارات تتطلب من الطالب اختيار الإجابة، وتشمل فقرات الصواب والخطأ، والاختيار من متعدد، والمزاوجة، وتتميز بالصدق والثبات والموضوعية (العجيلي، ٢٠١٧: ١٠٨).

ثانيا – نظرية القياس التقليدية Classical Test Theory (CTT):

تُعدّ نظرية القياس نموذجاً يستخدمه الباحثون لتفسير العوامل المؤثرة في الدرجة التي يحصل عليها الفرد في الاختبارات، إذ لا تعكس هذه الدرجة بالضرورة مستواه الحقيقي في الصفة المقاسة. وتركز النظرية على تقديم الأدلة التي تدعم الاستنتاجات المستخلصة من نتائج الاختبار وزيادة الثقة بها (المندلاوي والمندلاوي، ٢٠٢١: ٢٩٤). ويعد القياس أداة أساسية في اتخاذ القرارات التربوية والنفسية، لأنه يوفّر بيانات كمية دقيقة تساعد على فهم الظواهر والتنبؤ بها. ويهدف إلى تحقيق الموضوعية بحيث لا تتأثر نتائج الأفراد بالأدوات المستخدمة أو بعينات الاختبار (بركات، ٢٠١٨: ٥٣).

وتشترك نظريات القياس بمسلمات رئيسية، منها: إمكانية قياس أداء الفرد كمياً، وأن الأداء يعدّ دالةً لخصائصه، مع وجود فروق فردية في العلاقة بين الأداء والخصائص (حبيب وكاظم، ٢٠١٨: ١٠).

وتُعد النظرية الكلاسيكية في القياس من أوائل النظريات التي ظهرت في أواخر القرن التاسع عشر، وقد أسسها Spearman (1904) بتأكيده أن درجة الفرد تتكون من درجة حقيقية مضافاً إليها درجة خطأ، لذا سميت بنظرية "الدرجة الحقيقية والخطأ". وتفترض هذه النظرية أن توزيع درجات الأفراد يتبع المنحنى الاعتدالي الطبيعي (اليعقوبي، ٢٠١٣: ٢٦٩)، وأن الخصائص السيكومترية للاختبار تتأثر بخصائص عينة الأفراد والفقرات المكونة له. كما ترى أن درجة الفرد تمثل دالة خطية مطّردة، أي كلما زادت درجته زاد مستوى السمة لديه (حبيب وكاظم، ٢٠١٨: ١٠). وتستمد النظرية الكلاسيكية صلاحيتها من مجموعة فروض ومسلمات علمية تجعلها قابلة للتطبيق والتفسير (بركات، ٢٠١٨: ٥٥).

❖ يرى هامبلتون وزال (Hambleton & Zaal, 1991) وجربجوري (Gregory, 2004) أن تفسير الدرجات في النظرية الكلاسيكية يعتمد على الافتراضات الآتية:

١. الدرجة الظاهرية (O) تتكون من الدرجة الحقيقية (T) مضافاً إليها الخطأ في القياس (O = T + E)

٢. الدرجة التي يحصل عليها الفرد لا تمثل بالضرورة درجته الحقيقية، إذ تتأثر بالظروف الاختبارية.

٣. لا يتم معرفة الدرجة الحقيقية مباشرة، بل يستدل عليها من متوسط الدرجات الناتج عن تكرار التطبيق.

٤.الدرجة الحقيقية للفرد ثابتة لأنها تمثل قدرته المقاسة على فقرات الاختبار.

٥. الارتباط بين الدرجة الحقيقية ودرجة الخطأ يساوي صفر.

٦.الخطأ غير منتظم ولا يتكرر بنفس الصورة في جميع المواقف الاختبارية.

٧. ارتباط درجات الخطأ بين أدوات مختلفة أو صورتين للاختبار نفسه يساوي صفر، مما يدل على عشوائية الخطأ

٨. توجد علاقة عكسية بين الخطأ والدرجة الحقيقية، فكلما قل الخطأ زادت دقة الدرجة

٩.متوسط الأخطاء لجميع الأفراد في الاختبار يساوي صفر. (الطريري، ٢٠١٤: ٣٤–٣٧). (اليعقوبي، ٢٠١٣: ٢٧١). (بركات، ٢٠١٨: ٥٥). (حبيب وكاظم، ٢٠١٨: ١١).

❖الدرجة الحقيقية والدرجة الخطأ:

الدرجة الحقيقية: هي الدرجة الظاهرية للفرد بعد إزالة أخطاء القياس، وتمثل قدرته الفعلية المفترضة، وتعد ثابتة نظرياً لأنها تعبر عن مستوى الفرد الحقيقي في السمة المقاسة. ولا يمكن معرفتها مباشرة، بل يستدل عليها من متوسط الدرجات الناتج عن تكرار تطبيق الاختبار (اليعقوبي، ٢٠١٣: ٢٧١؛ بركات، ٢٠١٨: ١٥٦).

درجة الخطأ: لا يوجد قياس خالٍ من الخطأ، سواء في المجالات الفيزيائية أو النفسية، ويختلف مقدار الخطأ المقبول باختلاف دقة الخاصية المقاسة. تتأثر دقة القياس بأداة الاختبار والظروف المصاحبة له، إذ تنتج الأخطاء من عوامل متعددة قد تكون عشوائية أو منتظمة (Systematic Error)، ويحدد نوع تأثيرها مقدار الخطأ ودقته. (عودة، ١٩٨٥: ١٣٩). ويمكن الإشارة إلى أن درجة الخطأ تأتي من مصادر عدة:

أ – الخطأ المنتظم أو الثابت Systematic Error:

ب – خطأ الاستعمال Using Error:

ج – الخطأ العشوائي Random Error: (اليعقوبي، ٢٠١٣: ٢٧٣).

ثالثًا – الخصائص القياسية (السيكومترية) Psychometric Properties:

تمثل الخصائص السيكومترية الأدلة الكمية/الإجرائية على الصدق والثبات بعد تجريب الاختبار على عينة ممثّلة، وتشمل أيضاً مؤشرات فقرات مثل الصعوبة والتمييز وفاعلية المشتتات. (عبيدات والجندي، ٢٠٢٤: ٢٠٦). وتشمل ما يلي:

❖الصدق Validity:

يشير الصدق إلى مدى قياس الاختبار لما وضع لقياسه واتخاذ قرارات صحيحة بناء على درجاته؛ يتأثر بالأخطاء العشوائية والمنهجية معا. (فرج، ٢٠٠٧: ٢٣٧)

- **أنواعه/أدلته الأساسية:**

١. الصدق الظاهري: يقوم على الفحص الأولي لمحتوى الاختبار للتأكد من ارتباط فقراته بالخاصية المراد قياسها، ويتم ذلك عبر تحكيم الخبراء المتخصصين في المجال والقياس والتقويم لتقييم وضوح الفقرات ودقتها، ويفضل أن تزيد نسبة اتفاقهم على صدق الفقرة عن ٧٥% (العجيلي، ٢٠١٧: ٢٥٤).

٢. صدق المحتوى: يعد تمثيل مفردات الاختبار للمحتوى الدراسي المقرر بدقة عبر تحليل المحتوى وان إعداد جدول المواصفات يضمن توفير صدق المحتوى للاختبار (عبد الرحمن، ٢٠٠٨: ١٩٩).

٣.الصدق المرتبط بمحك: يشير إلى توافر الأدلة التجريبية التي تثبت أن الاختبار يقيس فعلاً الخاصية التي صمم لقياسها، لذا يعرف أحيانا بـ الصدق التجريبي (العجيلي، ٢٠١٧: ٢٥٥).

<u>وينقسم إلى نوعين:</u>

الصدق التلازمي: يستخدم لتشخيص الحالة الراهنة، مثل: هل أحمد مصاب بالفصام العقلي؟

الصدق التنبؤي: يستخدم للتنبؤ بنتائج مستقبلية، مثل: هل من المحتمل أن يصبح أحمد فصامياً؟ (عمر وآخرون، ٢٠٠٩: ١٩٨).

٤. صدق البناء (المفهوم): يشير إلى مدى توافق نتائج الاختبار مع الإطار النظري للمفهوم المقاس، ويتحقق منه بإجراءات تجريبية وإحصائية بين الفقرات. من خلال حساب معامل التمييز وارتباط الفقرات بالدرجة الكلية للاختبار، بهدف تحقيق افتراضاته (افتراض تحقيق الفروق الفردية. وافراض ضرورة قياس الفقرات والاختبار لسمة او خاصية واحدة معينة) ويعد دليلاً على أن الأداة تقيس البعد أو البنية النظرية المستهدفة بدقة. (علام، ٢٠٠٠: ٢١٥). ويعد الصدق مفهوماً تجميعياً متعدد الأدلة وليس رقماً واحداً (Messick)، أما الصدق الظاهري فهو شكل ضعيف من الصدق يعتمد على القبول الاجتماعي (عمر وآخرون، ٢٠٠٩: ١٩١؛ الريماوي، ٢٠١٧: ١٠٣).

- **عوامل مؤثرة في الصدق:**

ان توفر بعض العوامل يحقق علاقة طردية مع الصدق منها ثبات الأداة، الصياغة الجيدة والواضحة للفقرات، طول الاختبار بشكل مناسب، وضوح التعليمات واللغة، ملاءمة العينة للاختبار والعوامل البيئية الجيدة. بينما تمثل عوامل مثل الغش/التخمين او ضعف توافر العوامل السابقة يضعف الصدق ويمثل علاقة عكسية معه (عودة، ١٩٨٥: ١٦٧-١٦٨؛ العجيلي، ٢٠١٧: ٢٦٠-٢٦١؛ ملحم، ٢٠١٢: ٢٧٣-٢٧٤).

❖**الثبات Reliability:**الثبات هو مدى اتساق نتائج الاختبار ودقة قياسه عند تكرار التطبيق عبر الزمن أو الصور المختلفة، ويعد مؤشراً على استقرار الأداة وموثوقيتها في القياس. يقيس

الثبات مقدار الخطأ العشوائي في النتائج، ويعبَّر عنه بمعامل يتراوح بين (٠-١)، أو بالخطأ المعياري للقياس الذي يقدّر تشتت الدرجات حول الدرجة الحقيقية (أبو حطب وآخرون، ١٩٩٣: ١٠١-١٠٩؛ أنستازي وأورينا، ٢٠١٥: ١٣١).

- **العوامل المؤثرة في ثبات الاختبار:**

يتأثر الثبات بعدة عوامل، من أبرزها: عدد الفقرات (كلما زاد العدد ارتفع الثبات)، وتباين أداء المفحوصين، ومستوى صعوبة الفقرات (يفضل أن تكون متوسطة الصعوبة)، إضافة إلى وضوح الصياغة، وتجنب التخمين، والحالة النفسية والصحية للمفحوص أثناء التطبيق. (أبو حطب وآخرون، ١٩٩٣: ١٢٨؛ العجيلي، ٢٠١٧: ٢٦٤)

الفصل الثالث: منهجية البحث وإجراءاته.

منهجية البحث: يمثل منهج البحث الوصفي وصف الظاهرة موضع الدراسة بأتباع منهجية علمية صحيحة وتصوير النتائج التي يتم التوصل اليها على أشكال رقمية معبرة يمكن تفسيرها، وتعد في بعض الأحيان الطريقة الوحيدة التي يمكن اتخاذها للتصدي لمشكلة من المشكلات (شحاته، ٢٠٠٨: ١٩٦). لذلك أعتمد الباحث المنهج الوصفي لتحقيق أهداف بحثه.

مجتمع البحث: يتكون مجتمع البحث من طلبة الصف الثاني للأقسام غير الاختصاص من كل من كلية التربية ابن رشد للعلوم الإنسانية وكلية التربية للبنات في جامعة بغداد وكلية التربية للعلوم الإنسانية في الجامعة المستنصرية للعام الدراسي ٢٠٢٤ / ٢٠٢٥. والجدول (١) يوضح ذلك.

الجدول (١) مجتمع البحث لطلبة الصف الثاني موزعين بحسب الكليات والأقسام العلمية والجنس

الجامعة	الكلية	القسم	الجنس				المجموع	
			ذكور		أناث		عدد	نسبة
			عدد	نسبة	عدد	نسبة		
بغداد	التربية أبن رشد	التاريخ	٣٤	%٤٠	٧٢	%٦٠	١٠٦	%١٥
		الجغرافية	٥٨	%٥٣	٥٢	%٤٧	١١٠	%١٥
		اللغة العربية	٣٥	%٢٩	٨٥	%٧١	١٢٠	%١٦
		اللغة الإنكليزية	١١٥	%٤٣	١٥١	%٥٧	٢٦٦	%٣٦
		اللغة الكردية	٢٢	%٥١	٢١	%٤٩	٤٣	%٦
		علوم القران	٤٢	%٤٧	٤٧	%٥٣	٨٩	%١٢
	المجموع		٣٠٦	%٤٢	٤٢٨	%٥٨	٧٣٤	%٢٦
	التربية للبنات	الاقتصاد المنزلي	-	-	٣٨	%١٠٠	٣٨	%٤
		التاريخ	-	-	٢٠٠	%١٠٠	٢٠٠	%٢٠
		علوم القران	-	-	١٧٩	%١٠٠	١٧٩	%١٨
		اللغة الإنكليزية	-	-	١١٠	%١٠٠	١١٠	%١١

		اللغة العربية	–	–	١٣٥	%١٠٠	١٣٥	%١٤
		الجغرافيا	–	–	١١٢	%١٠٠	١١٢	%١٢
		رياض الأطفال	–	–	٩٣	%١٠٠	٩٣	%١٠
		علم الاجتماع	–	–	٧٦	%١٠٠	٧٦	%٨
		الحاسوب	–	–	٣٠	%١٠٠	٣٠	%٣
	المجموع		–	–	٩٣٥	%١٠٠	٩٧٣	%٣٥
المستنصرية	التربية	علوم القران	٧٨	%٥٣	٦٩	%٤٧	١٤٧	%١٤
		الرياضيات	١١١	%٤٤	١٤٠	%٥٦	٢٥١	%٢٣
		الفيزياء	٨٦	%٣٩	١٣٧	%٦١	٢٢٣	%٢١
		التاريخ	١١٣	%٥٠	١١٢	%٥٠	٢٢٥	%٢١
		الجغرافية	١٠٧	%٤٧	١٢١	%٥٣	٢٢٨	%٢١
	المجموع		٤٩٥	%٤٦	٥٧٩	%٥٤	١٠٧٤	%٣٩
المجموع الكلي			٨٠١	%٢٩	١٩٤٢	%٧١	٢٧٨١	%١٠٠

عينة البحث: تمثل العينة مجموعة جزئية محددة من مجتمع البحث، ويجب أن تكون ممثلة له وتختار بطرق علمية لضمان هذا التمثيل (عودة والخليلي، ١٩٨٨ :١٧١-١٧٢). اعتمد الباحث، لمتطلبات بناء الاختبار، اختيار ثلاث عينات: الأولى هي العينة الطبقية العشوائية، وبلغ عددها (٤٠٠) طالب وطالبة، واختيرت بما يتناسب مع أعدادهم في المجتمع، وتستخدم لحساب خصائص فقرات الاختبار وخصائصه العامة. والثانية هي عينة وضوح الفقرات والتعليمات، وبلغ عددها (٦٠) طالباً وطالبة، والثالثة هي عينة الثبات، وبلغ عددها (١٠٠) طالب وطالبة، وتستخدم لحساب ثبات الاختبار.

أداة البحث: يهدف البحث الحالي بناء اختبار تحصيلي لمادة علم النفس النمو لغير الاختصاص، التي تدرس لطلبة الصف الثاني للأقسام العلمية الواردة في الجدول (١) في كلية التربية ابن رشد وكلية التربية للبنات في جامعة بغداد. وكلية التربية الإنسانية في الجامعة المستنصرية. وقد أتبعت الباحث الخطوات الآتية:

١. تحديد المحتوى الدراسي:

يهدف الاختبار التحصيلي لقياس التحصيل في مادة علم نفس النمو للأقسام العلمية غير الاختصاص والتي تتوزع على أربعة فصول كما ثبتته هيئة العمداء لكليات التربية، والجدول (٢) يوضح ذلك.

الجدول (٢) مفردات مادة علم نفس النمو للأقسام العلمية في كليات التربية

الفصل	المفردات الفرعية	عدد أسابيع التدريس
الأول: علم نفس النمو	• **تعريفه، أهميته.** • **معنى النمو.**	**أسبوعان**

	• المبادئ العامة للنمو. • مراحل النمو.	
الثاني: العوامل المؤثرة في النمو	• العوامل الوراثية. • العوامل البيئية.	أسبوعان
الثالث: مناهج البحث في علم نفس النمو	• أساليب جمع المعلومات. • تصاميم البحوث (طولية، مستعرضة). • أنواع البحوث (تجريبية، ارتباطية).	أسبوعان
الرابع: مرحلة الطفولة	• تعريفها وأهميتها ومراحلها. • * النمو الجسمي. • التطور العقلي واللغوي. • التطور الاجتماعي والانفعالي. • التطور الخلقي.	خمسة أسابيع
الخامس: دور المؤسسات الاجتماعية في التنشئة الاجتماعية للطفل	• الأسرة. • المدرسة. • الأقران. • وسائل الأعلام.	أسبوعان
السادس: المراهقة	• تعريفها وأهميتها ومراحلها. • * النمو الجسمي. • التطور العقلي واللغوي. • التطور الاجتماعي والانفعالي. • التطور الخلقي.	أسبوعان
السابع: المراهق والمجتمع	• المراهق والأسرة. • المراهق والمدرسة. • المراهق والأقران. • المراهق ووسائل الأعلام.	أسبوعان
الثامن: المراهق والمهنة	• أهمية العمل في حياة المراهق. • أهمية اختيار المهنة والعوامل المؤثرة فيها. • توافق المراهق للعمل.	أسبوعان
التاسع: اتجاهات المراهقين وميولهم	• أهمية الميول والاتجاهات. • مصادر اكتساب الاتجاهات والميول. • العوامل المؤثرة في اتجاهات المراهقين وميولهم.	ثلاثة أسابيع
العاشر: بعض مشكلات المراهقين	• التأخر الدراسي. • السلوك العدواني. • جنوح المراهقين.	ثلاثة أسابيع

٢.صياغة الأهداف السلوكية:

صاغ الباحث (٧٥) هدفا سلوكيا مستندا على المادة العلمية لمادة علم نفس النمو للأقسام غير الاختصاص في كليات التربية معتمدا على القواعد العلمية لصياغة الأهداف السلوكية من حيث كونها (أهدافا محددة تحديدا دقيقا، ويكون قابلا للملاحظة والقياس، وتتناول سلوكيات أو استجابات عقلية، ويسترشد بها المعلم في تدريسه اليومي وتعينه على تقويم طلابه وتحقيق ما ينوي قياسه داخل الصف) (داود، ٢٠١٤، ٩١). وتوزعت الأهداف السلوكية بحسب مستويات الأهداف السلوكية لتصنيف بلوم، ويتضح عدد الاهداف متأثرة بالأهمية النسبية للمحتوى، في الخارطة الاختبارية والجدول (٣) يوضح ذلك.

٣.جدول المواصفات (الخارطة الاختبارية):

والذي يقوم بتحديد الموضوعات الرئيسة للمادة الدراسية وأهميتها، ويربطها بالأهداف السلوكية وعدد الأسئلة لكل هدف. يستند الجدول إلى نوع المحتوى، عمر المتعلم، مدة التدريس وزمن الاختبار، ويحدد الوزن النسبي للأهداف حسب رأي المعلم أو خبراء المجال، ليكون المجموع ١٠٠%. (عمر واخرون، ٢٠٠٩: ٤١٣). يمكن الاستفادة من الترتيب التنازلي لنسب الأهداف وفق تصنيف بلوم، حيث تبدأ بقدرة التذكر وتتصاعد نحو الفهم والتطبيق والتحليل. (اليعقوبي، ٢٠١٣: ٩٦). ويمكن تطبيق إجراءات جدول المواصفات بالشكل الآتي:

- تحديد الوزن النسبي لكل فصل أو مفردة رئيسة من الكتاب (المحتوى) بالقانون الآتي: الوزن النسبي للمفردة = الوقت المستغرق في تدريس المفردة × ١٠٠ / الوقت الكلي
- تحديد الوزن النسبي لكل هدف من الأهداف السلوكية بتصنيف بلوم بالقانون الآتي: الوزن النسبي للهدف = عدد الأهداف لكل مستوى × ١٠٠ / العدد الكلي للأهداف
- تحديد عدد الفقرات الاختبارية لكل مفردة، كالآتي: عدد الفقرات الاختبارية لكل مفردة = الوزن النسبي للمفردة × عدد الأسئلة الكلي / ١٠٠
- تحديد عدد الفقرات الاختبارية لكل مستوى من مستويات الأهداف، كالآتي: عدد الفقرات الاختبارية لكل هدف = الوزن النسبي للهدف × عدد الأسئلة الكلي / ١٠٠
- تحديد عدد الفقرات الاختبارية لكل مفردة ومستوى للأهداف، كالآتي: عدد الفقرات الاختبارية لكل مفردة ومستوى للأهداف = الوزن النسبي للمفردة × الوزن النسبي لمستوى الهدف × عدد الأسئلة الكلي والجدول (٣) يوضح ذلك.

الجدول (٣) جدول المواصفات للأختبار التحصيلي لمادة علم نفس النمو

الفصل	الوقت	الأهمية	الأهداف السلوكية				المجموع
	اسبوع	النسبية	التذكر ٤٠%	الفهم ٣٠%	التطبيق ٢٠%	التحليل ١٠%	
الأول	٢	٨%	٢	٢	١	١	٦
الثاني	٢	٨%	٢	٢	١	١	٦

الثالث	٢	%٨	٢	٢	١	١	٦
الرابع	٥	%٢٠	٦	٥	٣	١	١٥
الخامس	٢	%٨	٢	٢	١	١	٦
السادس	٢	%٨	٢	٢	١	١	٦
السابع	٢	%٨	٢	٢	١	١	٦
الثامن	٢	%٨	٢	٢	١	١	٦
التاسع	٣	%١٢	٣	٣	٢	١	٩
العاشر	٣	%١٢	٣	٣	٢	١	٩
المجموع	٢٥	%١٠٠	٢٥	٢٥	١٤	١١	٧٥

٤.صياغة فقرات الأختبار:

صاغ الباحث (٧٥) فقرة موضوعية من نوع الاختيار من متعدد، تضمنت كل فقرة من أصل للفقرة تتبعها أربعة بدائل واحدة منها صائبة والبقية خاطئة بالاستناد على الخارطة الاختبارية واختيرت نوع الفقرات هذه تحديداً لكونها تمتلك القدرة على تغطية محتوى المادة الدراسية، وخلوها من ذاتية التصحيح، كما ترتفع معاملي صدق الأختبار وثباته، فضلا عن قلة تأثرها بتخمين الإجابة مقارنة بالاختبارات القائمة على اختيار الطالب للإجابة (ميخائيل، ٢٠٢١: ٣٢٦).

٥.تعليمات الاختبار:

إن تعليمات الأختبار لابد أن تكون منسجمة مع المستوى العلمي للمتعلمين، وأن تتميز بوضوحها وقابليتها للفهم وسهلة التنفيذ، وينبغي اللجوء لإعطاء أمثلة توضيحية حول كيفية الإجابة (مراد وسليمان، ٢٠٠٢: ٢٠٦). وبناء على ذلك، تم صياغة التوجيهات والتعليمات الخاصة بكيفية الإجابة على الفقرات الاختبارية جميعها دون ترك أي منها، وتتضمن مدة الإجابة والمعلومات الخاصة بالطالب متمثلة بالجنس والشعبة.

٦.تصحيح إجابات الأختبار:

انطلاقا من كون الأختبار بصيغته الأولية يتكون من (٧٥) فقرة اختبارية، فقد أعتمد التصحيح بوضع درجة واحدة لكل فقرة اختبارية كانت إجابة الطالب عنها صحيحة وصفر لكل فقرة اختبارية كانت إجابة الطالب عنها خاطئة أو تركت الفقرة دون جواب أو عند اختيار الطالب لأكثر من بديل، حيث اعد مفتاح شفاف لتصحيح الإجابة لجميع الفقرات الاختبارية

٧.التحليل المنطقي لفقرات الأختبار:

يستعمل هذا التحليل للتأكد من مناسبة الأختبار الذي وضع من أجله، حيث يتركز على المظهر العام للأختبار بعرض الفقرات الاختبارية وكيفية صياغتها على من مجموعة من الخبراء للتأكد من مدى وضوحها وموضوعيتها (علام، ٢٠٠٠: ١٩٠). لذلك عرض الباحث الفقرات الاختبارية بصيغتها الأولية مع الأهداف السلوكية التي قام بأعدادها على مجموعة من الخبراء

المتخصصين في القياس والتقويم وعلم نفس النمو وعلم النفس التربوي والذي بلغ عددهم (١١) خبير، للتأكد من سلامة الفقرات وملائمتها للأغراض المحددة ووضوح صياغة الفقرات وتجانس البدائل. وقد كانت نتيجة التحليل المنطقي لفقرات الأختبار أتفاق الخبراء بنسبة (١٠٠%) على صلاحية جميع الفقرات الاختبارية عدى بعض التعديلات البسيطة التي اجراها الباحث، وبذلك أصبح الأختبار جاهزا لأجراء التحليل الإحصائي لفقرات الأختبار.

٨.وضوح تعليمات الأختبار وفقراته:

فمن الضروري التحقق من وضوح تعليمات وفقرات الاختبار ومدى فهم العينة لها (فرج، ١٩٨٠: ١٦٠) ولذا طُبقت تجربة الوضوح على عينة عشوائية طبقية خارج عينة التحليل الإحصائي، حيث طُلب من الطلبة قراءة التعليمات والفقرات وطرح الاستفسارات. أشارت نتائج التجربة إلى وضوح الفقرات والتعليمات وبدائل الإجابة للاختبارين، مما أدى إلى اعتمادهما كما هما، كما حدد الزمن المستغرق للإجابة بتراوح بين (٣٨-٤٣) دقيقة، بمتوسط حسابي قدره (٤٠) دقيقة.

الجدول (٤) عينة وضوح تعليمات الاختبار بحسب الكلية والجنس

الجامعة	الكلية	القسم	الجنس		المجموع
			ذكور	إناث	
بغداد	التربية ابن رشد	التاريخ	١٠	١٠	٢٠
المستنصرية	التربية	علوم القران	١٠	١٠	٢٠
المجموع			٢٠	٢٠	٤٠

٩.التحليل الإحصائي لفقرات الاختبار:

تعد صياغة الفقرات وأحكام المحكمين مهمة في بناء الاختبارات، لكنها لا تغني عن التجريب الميداني والتحليل الإحصائي للفقرات (علام، ٢٠٠٠: ٢٦٧)، الذي يهدف لزيادة صدق وثبات الاختبار عبر اكتشاف واستبعاد الفقرات شديدة الصعوبة أو السهولة أو غير المميزة، وذلك بتقدير معاملات الصعوبة والتمييز (Anastasi, 1982: 192). ويؤكد النص على أهمية حجم العينة، مشيرا إلى أن الحجم الأمثل لعينة التحليل هو ١٠٠ فرد في كل مجموعة من المجموعتين المتطرفتين (٢٧% من العينة الكلية) (Anastasi, 1988: 23). وبناء على ذلك، تكونت عينة التحليل الإحصائي للدراسة من (٤٠٠) طالب وطالبة من كليات تربوية مختلفة، واختيرت بأسلوب العينة الطبقية العشوائية وفقاً لمتغيرات الكلية والقسم والجنس ومثلت نسبة (١٤.٤%) من مجتمع البحث.

الجدول (٥) عينة التحليل الإحصائي لطلبة الصف الثاني موزعة بحسب الكلية والقسم والجنس

الجامعة	الكلية	القسم	الجنس		المجموع
			ذكور	أناث	
بغداد	التربية ابن رشد	التاريخ	25	27	52
		اللغة العربية	25	27	52

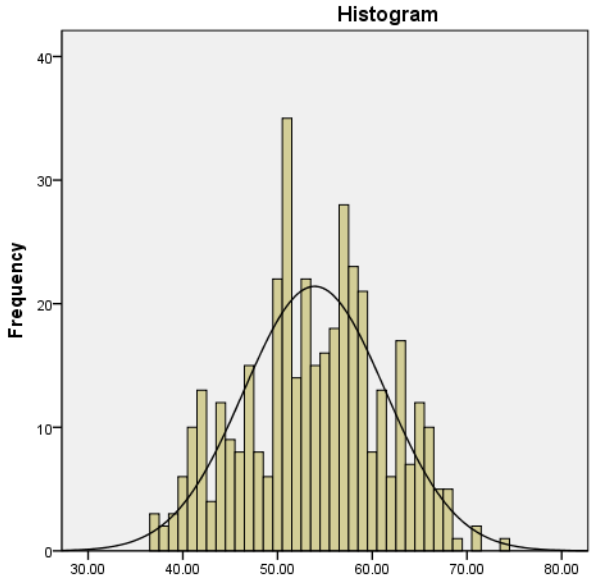
	المجموع		50	54	104
	التربية للبنات	اللغة العربية	-	70	70
		اللغة الإنكليزية	-	70	70
	المجموع		-	140	140
المستنصرية	التربية	علوم القران	36	42	78
	الإنسانية	الفيزياء	36	42	78
	المجموع		72	84	156
المجموع الكلي			122	278	400

إذ تم تطبيق الاختبار وحساب المؤشرات الإحصائية لعينة التحليل الإحصائي، واتضحت نتائج درجات أفراد عينة التحليل الإحصائي للأختبار التحصيلي. الجدول (٨) يوضح ذلك.

الجدول (٨) المؤشرات الإحصائية لعينة تحليل الفقرات

عدد أفراد العينة	الوسط الحسابي	الوسيط	المنوال	الانحراف المعياري	الالتواء	التفلطح	اقل قيمة	أعلى قيمة
٤٠٠	٥٩.٤٥٧	٥٨.٠٠٠	٥٩.٠٠٠	٥.٢٠٨	٠.٠٩٢	٠.٣٣١	٢٢	٧١

وتم التأكد من شكل التوزيع التكراري للدرجات للاختبار التحصيلي فهو مؤشر على دقة الاختبار في قياس الخاصية أو السمة التي تتوزع في المجتمع توزيعاً اعتدالياً، فكلما كان التوزيع للدرجات اعتدالياً أو قريباً من الاعتدالية كان ذلك مؤشراً على صحة تمثيل العينة للمجتمع الذي تعود إليه ودقة الاختبار في قياسه للسمة، وأن عينة الأفراد التي يجري عليها الاختبار يعتمد عليها شكل التوزيع التكراري وكذلك نوع الاختبار، وعلى الصفة التي تقيسها (خيري، ١٩٧٠: ٢١٨). والشكل (١) يوضح ذلك.



الشكل (١) الأعمدة البيانية لدرجات الطلبة على الاختبار التحصيلي.

ومن مراجعة المؤشرات الإحصائية لدرجات أفراد عينة التحليل الإحصائي للاختبار متمثلة بمقاييس النزعة المركزية ومقاييس التشتت وقيمة التفلطح والالتواء، وعند مقارنتها بالقيم المعيارية للتوزيع الاعتدالي تبين أنها قريبة من التوزيع الاعتدالي، وهذا يعني ان المنحنيين يميلان قليلاً باتجاه الدرجات العالية وان التفلطح والالتواء اعلى قليلاً لما يمثله في التوزيع الاعتدالي (خيري، ١٩٧٠ :١٩٥). وللكشف عن مستوى صعوبة الفقرات وقوة تمييزها وفعالية البدائل الخاطئة، وفيما يأتي توضيح للخصائص الإحصائية للفقرات:

اولاً: صعوبة الفقرة: للحصول على أدق قياس في المقاييس معيارية المرجع، تعد درجة صعوبة الفقرة خاصية سيكومترية أساسية تقاس بنسبة الإجابات الصحيحة أو الخطأ، ويجب ترتيب الفقرات من الأسهل للأصعب على أن يكون مدى الصعوبة حول نسبة ٠.٥٠ (الريماوي، ٢٠١٧: ٧١-٧٢). ولحساب معامل صعوبة فقرات الاختبار التحصيلي، اعتمد الباحث على المجموعتين المتطرفتين (العليا والدنيا)، حيث تم تطبيق معادلة الصعوبة التي تتضمن جمع عدد الإجابات الصحيحة في المجموعتين وقسمته على العدد الكلي لأفراد هاتين المجموعتين. والجدول (٩) يوضح ذلك.

ثانيا: القوة التمييزية للفقرات: تهدف درجة التمييز إلى إظهار الفروق بين مستويات الأفراد، حيث يعد السؤال جيدا إذا خدم غرضه في التمييز بين الطلبة المتميزين والضعاف (المياحي، ٢٠١١: ١٧٧-١٧٨). فالفقرة المميزة هي التي تمتلك القدرة على التفريق بين الأفراد ذوي المستويات العليا والضعاف في السمة المقاسة (دوران، ١٩٨٥: ١٢٥). لحساب التمييز، رتب الباحث درجات العينة (٤٠٠ طالب وطالبة) من الأعلى إلى الأقل، وحدد المجموعتين المتطرفتين (العليا والدنيا) باعتماد نسبة (١٤%) في كل مجموعة، حيث ضمت كل منها (١٠٨) طالباً وطالبة. تم حساب معامل التمييز لكل فقرة بطرح عدد الإجابات الصحيحة في المجموعة الدنيا من عدد الإجابات الصحيحة في المجموعة العليا، ثم قسمة الناتج على نصف مجموع عدد الأفراد في المجموعتين المتطرفتين. وكما في الجدول (٩).

الجدول (٩) قيم معاملات الصعوبة ومعاملات التمييز لفقرات الاختبار

ت الفقرة	مج	إجابات صحيحة	معامل الصعوبة	معامل التمييز	ت الفقرة	مج	إجابات صحيحة	معامل الصعوبة	معامل التمييز	ت الفقرة	مج	إجابات صحيحة	معامل الصعوبة	معامل التمييز
١	العليا	٧٤	٠.٥٣	٠.٣١	٢٦	العليا	٧٧	٠.٥٠	٠.٤٣	٥١	العليا	٥٩	٠.٣٥	٠.٣٩
	الدنيا	٤١	متوسطة	جيدة		الدنيا	٣١	متوسطة	جيدة جدا		الدنيا	١٧	متوسطة	جيدة
٢	العليا	٥٧	٠.٣٦	٠.٣٣	٢٧	العليا	٤٠	٠.٢٧	٠.٢٠	٥٢	العليا	٧٨	٠.٥٦	٠.٣١
	الدنيا	٢١	متوسطة	جيدة		الدنيا	١٨	صعبة	حدية		الدنيا	٤٤	متوسطة	جيدة
٣	العليا	٤٥	٠.٢٦	٠.٣١	٢٨	العليا	٦٩	٠.٤٢	٠.٤٤	٥٣	العليا	٩١	٠.٦٩	٠.٣١
	الدنيا	١١	صعبة	جيدة		الدنيا	٢٢	متوسطة	جيدة جدا		الدنيا	٥٧	متوسطة	جيدة
٤	العليا	٨١	٠.٥٦	٠.٣٨	٢٩	العليا	٧٥	٠.٤٤	٠.٥٠	٥٤	العليا	٨٩	٠.٦٥	٠.٣٥
	الدنيا	٤٠	متوسط	جيدة		الدنيا	٢١	متوسطة	جيدة جدا		الدنيا	٥١	متوسطة	جيدة
٥	العليا	٤٠	٠.٢٥	٠.٢٣	٣٠	العليا	٦٤	٠.٣٩	٠.٣٩	٥٥	العليا	٦٩	٠.٤٨	٠.٣٢
	الدنيا	١٥	صعبة	حدية		الدنيا	٢١	متوسطة	جيدة		الدنيا	٣٤	متوسطة	جيدة

٦	العليا	٧٤	٠.٤٨	٠.٤١	٣١	العليا	٨٦	٠.٥٧	٠.٤٥	٥٦	العليا	٧٢	٠.٥١	٠.٣١
	الدنيا	٣٠	متوسطة	جيدة		الدنيا	٣٧	متوسطة	جيدة جدا		الدنيا	٣٨	متوسطة	جيدة
٧	العليا	١٠٦	٠.٧٠	٠.٥٦	٣٢	العليا	٩٠	٠.٤٧	٠.٧٣	٥٧	العليا	٨٨	٠.٦٢	٠.٣٩
	الدنيا	٤٥	سهلة	جيدة		الدنيا	١١	متوسطة	جيدة جدا		الدنيا	٤٦	متوسطة	جيدة
٨دة جد	العليا	٥٧	٠.٣٦	٠.٣٤	٣٣	العليا	١٠١	٠.٧٤	٠.٣٩	٥٨	العليا	٦١	٠.٤٠	٠.٣٢
	الدنيا	٢٠	متوسطة	جيدة		الدنيا	٥٩	سهلة	جيدة		الدنيا	٢٦	متوسطة	جيدة
٩	العليا	٨٨	٠.٦٩	٠.٢٥	٣٤	العليا	٨٣	٠.٥٠	٠.٥٥	٥٩	العليا	٩٥	٠.٧٣	٠.٣١
	الدنيا	٦١	متوسطة	حدية		الدنيا	٢٤	متوسطة	جيدة جدا		الدنيا	٦٢	سهلة	جيدة
١٠	العليا	٦٤	٠.٣٩	٠.٣٩	٣٥	العليا	٦٥	٠.٣٦	٠.٤٩	٦٠	العليا	٤٨	٠.٢٨	٠.٣٢
	الدنيا	٢١	متوسطة	جيدة		الدنيا	١٢	متوسطة	جيدة جدا		الدنيا	١٣	صعبة	جيدة
١١	العليا	٧٢	٠.٤٤	٠.٤٦	٣٦	العليا	٩٨	٠.٧٥	٠.٣٢	٦١	العليا	٦٢	٠.٣٦	٠.٣٦
	الدنيا	٢٢	متوسطة	جيدة		الدنيا	٦٣	سهلة	جيدة		الدنيا	٢٣	متوسطة	جيدة
١٢	العليا	٥٥	٠.٣٨	٠.٢٥	٣٧	العليا	٧٤	٠.٤٥	٠.٤٧	٦٢	العليا	٨١	٠.٥٩	٠.٣١
	الدنيا	٢٨	متوسطة	حدية		الدنيا	٢٣	متوسطة	جيدة جدا		الدنيا	٤٧	متوسطة	جيدة
١٣	العليا	٣٥	٠.٢١	٠.٢٣	٣٨	العليا	٦٦	٠.٤٥	٠.٣٢	٦٣	العليا	٩٧	٠.٧٢	٠.٣٥
	الدنيا	١٠	صعبة	حدية		الدنيا	٣١	متوسطة	جيدة		الدنيا	٥٩	سهلة	جيدة
١٤	العليا	٣٦	٠.٢٠	٠.٢٦	٣٩	العليا	٥٢	٠.٣٨	٠.٢١	٦٤	العليا	٦٧	٠.٤٧	٠.٣٠
	الدنيا	٨	صعبة	حدية		الدنيا	٢٩	متوسطة	حدية		الدنيا	٣٥	متوسطة	جيدة
١٥	العليا	٨٩	٠.٥٧	٠.٥١	٤٠	العليا	٩٨	٠.٧١	٠.٣٨	٦٥	العليا	٤٤	٠.٢٦	٠.٣٠
	الدنيا	٣٤	متوسطة	جيدة		الدنيا	٥٧	سهلة	جيدة		الدنيا	١٢	صعبة	جيدة
١٦	العليا	٦٤	٠.٤٩	٠.٢١	٤١	العليا	٦٤	٠.٤١	٠.٣٧	٦٦	العليا	٩٦	٠.٧٢	٠.٣٤
	الدنيا	٤١	متوسطة	حدية		الدنيا	٢٤	متوسطة	جيدة		الدنيا	٥٩	سهلة	جيدة
١٧	العليا	٨٦	٠.٦١	٠.٣٨	٤٢	العليا	٥٥	٠.٣٣	٠.٣٦	٦٧	العليا	٨٣	٠.٥٨	٠.٣٨
	الدنيا	٤٥	متوسطة	جيدة		الدنيا	١٦	متوسطة	جيدة		الدنيا	٤٢	متوسطة	جيدة
١٨	العليا	٩٣	٠.٦٩	٠.٣٥	٤٣	العليا	٩٦	٠.٧٦	٠.٢٦	٦٨	العليا	٦٩	٠.٤٧	٠.٣٣
	الدنيا	٥٥	متوسطة	جيدة		الدنيا	٦٨	سهلة	حدية		الدنيا	٣٣	متوسطة	جيدة
١٩	العليا	٤٨	٠.٢٩	٠.٣١	٤٤	العليا	٧٤	٠.٤٤	٠.٤٩	٦٩	العليا	٥٨	٠.٣٧	٠.٣٤
	الدنيا	١٤	صعبة	جيدة		الدنيا	٢١	متوسطة	جيدة جدا		الدنيا	٢١	متوسطة	جيدة
٢٠	العليا	١٠١	٠.٧١	٠.٤٤	٤٥	العليا	٤٩	٠.٢٨	٠.٥٦	٧٠	العليا	٩٢	٠.٦٦	٠.٣٨
	الدنيا	٥٤	سهلة	جيدة		الدنيا	١١	صعبة	جيدة جدا		الدنيا	٥١	متوسطة	جيدة
٢١	العليا	٦٧	٠.٤٩	٠.٢٧	٤٦	العليا	٨١	٠.٥٦	٠.٣٨	٧١	العليا	٦٨	٠.٤٦	٠.٣٣
	الدنيا	٣٨	متوسطة	حدية		الدنيا	٤٠	متوسطة	جيدة		الدنيا	٣٢	متوسطة	جيدة
٢٢	العليا	٩٦	٠.٥٩	٠.٥٩	٤٧	العليا	٨٣	٠.٦٢	٠.٣٠	٧٢	العليا	٥٩	٠.٣٨	٠.٣٣
	الدنيا	٣٢	متوسطة	جيدة جدا		الدنيا	٥١	متوسطة	جيدة		الدنيا	٢٣	متوسطة	جيدة
٢٣	العليا	٨٨	٠.٥٣	٠.٥٧	٤٨	العليا	٩٥	٠.٧٢	٠.٣١	٧٣	العليا	٦٨	٠.٤٧	٠.٣١
	الدنيا	٢٦	متوسطة	جيدة جدا		الدنيا	٦٢	سهلة	جيدة		الدنيا	٣٤	متوسطة	جيدة
٢٤	العليا	٤١	٠.٢٧	٠.٢٠	٤٩	العليا	٧٥	٠.٥٨	٠.٢٢	٧٤	العليا	٥٥	٠.٣٥	٠.٣١
	الدنيا	١٩	صعبة	حدية		الدنيا	٥١	متوسطة	حدية		الدنيا	٢١	متوسطة	جيدة
٢٥	العليا	٧٦	٠.٤٩	٠.٤٢	٥٠	العليا	٨٣	٠.٦١	٠.٣١	٧٥	العليا	٦١	٠.٤٤	٠.٣٥
	الدنيا	٣٠	متوسطة	جيدة جدا		الدنيا	٤٩	متوسطة	جيدة		الدنيا	٢٣	متوسطة	جيدة

ولقد أعتمد على معيار الحكم على مستوى صعوبة الفقرات المعتمد في معظم الدراسات السابقة والتي تناولت بناء الاختبارات التحصيلية والذي قسم مستويات الصعوبة الى خمسة مستويات وهي (٠.١٩ فأقل صعبة جدا، ٠.٢٠ – ٠.٢٩ صعبة،٠.٣٠ – ٠.٦٩ متوسطة الصعوبة، ٠.٧٠ – ٠.٧٩ سهلة، ٠.٨٠ فأكثر سهلة جدا، والمدى المعتمد لمستويات الصعوبة يتراوح بين

(٠.٢٠ – ٠.٨٠) (حبيب وكاظم، ٢٠١٨: ١٢). وفي ضوء مدى الصعوبة أعلاه وزعت فقرات الاختبار التحصيلي بحسب مستوى صعوبتها والنسب المئوية لها، والجدول (١٠) يوضح ذلك.

الجدول (١٠) توزيع فقرات الأختبار التحصيلي لمادة علم نفس النمو وفقاً لمستويات الصعوبة

عدد الفقرات	مدى صعوبة الفقرات	مستوى صعوبة الفقرات								متوسط الصعوبة
		٠.١٩ فأقل صعبة جداً		٠.٢٠– ٠.٢٩ صعبة		٠.٣٠–٠.٦٩ متوسطة الصعوبة		٠.٧٠– ٠.٧٩ سهلة		٠.٨٠ فأكثر سهلة جداً	
		العدد	%	العدد	%	العدد	%	العدد	%	العدد	%	
٧٥	٠.٢٠ – ٠.٧٦	–	–	١٠	%١٣	٥٥	%٧٤	١٠	%١٣	–	–	٠.٤٩

ويتضح من الجدول (١١) أن فقرات الأختبار لم تتضمن أي فقرة واحدة صعبة جداً أو فقرة سهلة جدا، وتراوحت مديات صعوبة فقرات للأختبار التحصيلي من إعداد الباحث بين (٠.٢٠ – ٠.٧٦) وبمتوسط صعوبة قدره (٠.٤٩).

وأعتمد محك الحكم على تمييز الفقرات الذي يستند الى أربعة مستويات هي (٠.١٩ فأقل تمييز ضعيف تحذف الفقرة، بين ٠.٢٠ – ٠.٢٩ تمييز حدي لابد من تحسين الفقرة، ٠.٣٠ –٠.٣٩ تمييز جيد يفضل تحسين الفقرة، ٠.٤٠ فأكثر تمييز جيد جدا (حبيب وكاظم، ٢٠١٨: ١٣). وفي ضوء معيار التمييز المبينة وزعت فقرات الأختبار التحصيلي لمادة علم نفس النمو لطلبة الصف الثاني بحسب مستوى تمييزها والنسب المئوية لها، والجدول (١١) يوضح ذلك.

الجدول (١١) تصنيف فقرات الأختبار لمادة علم نفس النمو وفقاً لمحك التمييز

عدد الفقرات	مدى تمييز الفقرات	مستوى تمييز الفقرات								متوسط التمييز
		٠.١٩ فأقل تمييز ضعيف		٠.٢٠–٠.٢٩ حدي		٠.٣٠–٠.٣٩ تمييز جيد		٤٠ فأكثر تمييز جيد جدا		
		العدد	%	العدد	%	العدد	%	العدد	%	
٧٥	٠.٢٠ – ٠.٧٣	–	–	١٢	%١٦	٤٢	%٥٦	٢١	%٢٨	٠.٣٦

ويتضح من الجدول (١١) أن فقرات الأختبار المكون من (٧٥) فقرة لم تتضمن أي فقرة ذات تمييز ضعيف، وتراوحت مديات تمييزها بين (٠.٢٠ – ٠.٧٣) وبمتوسط صعوبة قدره (٠.٣٦) وهذا يقع ضمن مستوى الفقرات تمييز (الحدي، الجيد، الجيد جدا).

ثالثا: صدق الفقرات: Validity Items: يكاد يجمع معظم المتخصصين في ميدان القياس النفسي على ضرورة توافر الصدق في فقرات المقاييس النفسية، والشيء الذي يزيد من أهمية الصدق التجريبي للفقرة، والذي يعد المحك الداخلي الافضل لقياس صدق الفقرة هو حساب الدرجة الكلية للمقياس من خلال حساب معامل الارتباط بين درجات كل فقرة ودرجة المحك

(Anastasis,1976:211) ولكون صدق الاختبار يعتمد على صدق فقراته، ويعد الصدق مؤشرا مهما لقياس المفهوم المراد قياسه والذي تقيسه الدرجة الكلية (عبد الرحمن، ٢٠٠٨: ٤١٣)، لذا قام الباحث بحساب معامل الصدق من خلال استعمال معامل ارتباط (بوينت بايسيريال – الاعتباطي) بين الدرجة الكلية للاختبار ودرجات فقراته والجدول (١٢) يوضح ذلك.

الجدول (١٢) قيم معاملات الصدق لفقرات الاختبار التحصيلي

ت الفقرة	معامل الصدق	ت الفقرة	معامل الصدق	ت الفقرة	معامل الصدق	ت الفقرة	معامل الصعوبة	ت الفقرة	معامل الصدق	ت الفقرة	معامل الصعوبة
١	٠.٤٢٢	١٣	٠.٣٢٦	٢٦	٠.٥٢٣	٣٨	٠.٥٠٢	٥١	٠.٣٣٦	٦٣	٠.٤٥١
٢	٠.٣٦٧	١٤	٠.٣٥٤	٢٧	٠.٣٢٦	٣٩	٠.٣٦٤	٥٢	٠.٥٢٣	٦٤	٠.٦٢٤
٣	٠.٣٢٤	١٥	٠.٤٢٠	٢٨	٠.٣٠١	٤٠	٠.٤٧٥	٥٣	٠.١٢٨	٦٥	٠.٤١٥
٤	٠.٣٨١	١٦	٠.٣٦٢	٢٩	٠.٤٥٧	٤١	٠.٣٢٤	٥٤	٠.٣٢٥	٦٦	٠.٣٠٢
٥	٠.٣٢٣	١٧	٠.٣٨٩	٣٠	٠.٥٢٣	٤٢	٠.٢٣٥	٥٥	٠.٤١٢	٦٧	٠.٣٢٦
٦	٠.٢٣٥	١٨	٠.٣٣٠	٣١	٠.٣٦٢	٤٣	٠.٢٦٧	٥٦	٠.٣٢٣	٦٨	٠.٣٣٩
٧	٠.٣٤١	١٩	٠.٣٢٤	٣٢	٠.٣٩٩	٤٤	٠.٣٦٤	٥٧	٠.٣٦٥	٦٩	٠.٣٢٤
٨	٠.٤٥٧	٢٠	٠.٤١١	٣٣	٠.٤٢١	٤٥	٠.٢٦٨	٥٨	٠.٤٤٧	٧٠	٠.٣٢٦
٩	٠.٣٦٢	٢١	٠.٣٢٦	٣٤	٠.٤٣٩	٤٦	٠.٦١٥	٥٩	٠.٥١٢	٧١	٠.٣٢٥
١٠	٠.٣٩٩	٢٢	٠.٥٤٨	٣٥	٠.٣٢٥	٤٧	٠.٣٢٤	٦٠	٠.٥٤٧	٧٢	٠.٣١٤
١١	٠.٤٢٥	٢٣	٠.٣٢٨	٣٦	٠.٤٢٨	٤٨	٠.٤٤٥	٦١	٠.٦١٢	٧٣	٠.٣٢٥
١٢	٠.٣٣٣	٢٤	٠.٤٨٨	٣٧	٠.٣٥١	٤٩	٠.٣٢٥	٦٢	٠.٣٢٥	٧٤	٠.٢٢٥
		٢٥	٠.٣٢٥			٥٠	٠.٦٢٣			٧٥	٠.٣٩٩

***القيمة الحرجة لمعامل الأرتباط عند مستوى دلالة (٠.٠٥) ودرجة حرية (٣٩٨) تساوي (٠.٠٩٨).**

يتبين من الجدول (١٢) أن قيم معاملات الارتباط المحسوبة بين درجة الفقرة والدرجة الكلية للاختبار كانت أكبر من القيمة الحرجة لمعامل الارتباط وعند مستوى الدلالة (٠.٠٥)، ودرجة حرية (٣٩٨)، لذلك فأن علاقة جميع الفقرات بالدرجة الكلية للاختبار دالة إحصائياً، وذلك دليل على إن فقرات كل اختبار تتسق فيما بينها في قياس قدرة واحدة.

رابعا – فعالية البدائل الخاطئة Effectiveness of Distractors: عند استخدام فقرات الاختيار من متعدد، ينبغي على الباحث فحص إجابات العينة لجميع البدائل بهدف تحديد فعالية كل بديل ومدى قدرته على التمييز بين المتفوقين والضعفاء تحصيلياً (المياحي، ٢٠١١: ١٨١). ويتم ذلك من خلال مقارنة إجابات المجموعتين العليا والدنيا، إذ يعد البديل الخاطئ فعالاً إذا جذب عدداً أكبر من أفراد المجموعة الدنيا دون أن يخطئ فيه الجميع، ويحسب معامل فعالية

البدائل بالطريقة نفسها التي يحسب بها معامل التمييز، سواء منطقياً أو إحصائياً (الإمام وآخرون، ١٩٩٠: ١١٣). وقد تم حساب عدد الإجابات لكل بديل في المجموعتين العليا والدنيا في الاختبار التحصيلي، كما يوضح الجدول (١٣).

الجدول (١٣) فعالية البدائل الخاطئة للأختبار التحصيلي لمادة علم نفس النمو من إعداد الباحث

الفقرة	أ	ب	ج	د	الفقرة	أ	ب	ج	د	الفقرة	أ	ب	ج	د
١	-	٠,٠١٢	٠,٠٠١	٠,٠٠٥	٢٦	٠,٠٢٠	٠,٠١٠	٠,٠١٢	-	٥١	٠,٠٠١	-	٠,٠٠٢	٠,٠١٢
٢	٠,١٠١	٠,٠٤١	٠,٠٢٣	-	٢٧	٠,٠٣٠	-	٠,٠٢١	٠,٠٢٢	٥٢	٠,٠٢٠	٠,٠٢٠	-	٠,٠٣٢
٣	٠,٠١٢	٠,٠٣٤	-	٠,٠٢٢	٢٨	-	٠,٠١٠	٠,٠١٢	٠,٠٢٣	٥٣	٠,٠١٠	٠,٠١٢	٠,٠٣٢	-
٤	٠,٠١٣	-	٠,٠٣٣	٠,٠٢٥	٢٩	٠,٠٢١	-	٠,٠٢٥	٠,٠١١	٥٤	٠,٠٠٥	٠,٠١٠	-	٠,٠٢٥
٥	٠,٠١١	٠,٠٥٢	-	٠,٠٤٢	٣٠	٠,٠١٢	-	٠,٠١١	٠,٠٢٣	٥٥	٠,٠٢١	٠,٠١٠	٠,٠٢٣	-
٦	٠,٠٣١	٠,٠٢٢	-	٠,٠٦٢	٣١	-	٠,٠٣٠	٠,٠٢١	٠,٠١٠	٥٦	٠,٠٣٢	٠,٠٠٩	-	٠,٠٢١
٧	-	٠,٠٤٢	٠,٠٠٢	٠,٠١٢	٣٢	-	٠,٠١٢	٠,٠٢٢	٠,٠٢١	٥٧	٠,٠٢٣	٠,٠٤١	٠,٠٠٨	-
٨	٠,٠١٧	٠,٠٣٤	٠,٠٢٦	-	٣٣	٠,٠١٢	-	٠,٠١١	٠,٠٠١	٥٨	٠,٠١٠	٠,٠٢٥	-	٠,٠١٢
٩	٠,٠٢١	٠,٠٥١	-	٠,٠٤٥	٣٤	٠,٠٢٠	٠,٠١٢	٠,٠٣٢	-	٥٩	٠,٠٢١	٠,٠٢٢	-	٠,٠٢١
١٠	٠,٠٠٤	٠,٠٢١	-	٠,٠٣٠	٣٥	-	٠,٠١١	٠,٠١٠	٠,٠٠٥	٦٠	-	٠,٠٢٤	٠,٠٠٦	٠,٠٤١
١١	٠,٠١٢	٠,٠٣٣	-	٠,٠٢٥	٣٦	٠,٠٢٥	٠,٠٢٣	٠,٠١٤	-	٦١	٠,٠١٠	٠,٠٢١	٠,٠٣٤	-
١٢	-	٠,٠٢٢	٠,٠٣٤	٠,٠٣٨	٣٧	٠,٠٢٢	-	٠,٠١٠	٠,٠٠٤	٦٢	٠,٠٢٢	-	٠,٠١٠	٠,٠٢٤
١٣	-	٠,٠٢١	٠,٠٣٠	٠,٠١٥	٣٨	٠,٠٣٢	٠,٠٢٤	-	٠,٠١٥	٦٣	٠,٠١٥	-	٠,٠٠٦	٠,٠١٢
١٤	٠,٠١١	٠,٠٢١	٠,٠٣٥	-	٣٩	٠,٠٤٠	٠,٠٢٠	-	٠,٠٢٣	٦٤	٠,٠٢٣	٠,٠٣٠	٠,٠١٢	-
١٥	٠,٠١٥	-	٠,٠٢٠	٠,٠٢٤	٤٠	-	٠,٠١١	٠,٠١٥	٠,٠٣٢	٦٥	-	٠,٠٢٠	٠,٠٣٣	٠,٠٢٥
١٦	٠,٠١١	-	٠,٠١٢	٠,٠٢١	٤١	٠,٠١٢	٠,٠٢٤	٠,٠١٠	-	٦٦	٠,٠١١	-	٠,٠١٥	٠,٠٢٣
١٧	٠,٠٢٣	٠,٠٢٤	-	٠,٠٣١	٤٢	٠,٠٠٣	-	٠,٠١٤	٠,٠٢٥	٦٧	-	٠,٠١٢	٠,٠٠٥	٠,٠٢٣
١٨	٠,٠٢٢	٠,٠٣١	-	٠,٠١٢	٤٣	٠,٠٢٣	-	٠,٠٠٢	٠,٠٣١	٦٨	٠,٠١٢	٠,٠٢١	-	٠,٠٣٤
١٩	٠,٠٣٢	٠,٠٢١	-	٠,٠١١	٤٤	٠,٠٢٥	-	٠,٠١٢	٠,٠٠٥	٦٩	٠,٠١٠	٠,٠٠٨	-	٠,٠٢٣
٢٠	-	٠,٠٢٠	٠,٠١١	٠,٠٣٢	٤٥	-	٠,٠٢٦	٠,٠١٠	٠,٠٢٥	٧٠	٠,٠١٠	-	٠,٠١٢	٠,٠٣٠
٢١	٠,٠١١	٠,٠٢٣	٠,٠٢٤	-	٤٦	٠,٠١١	٠,٠٠٦	-	٠,٠١٠	٧١	٠,٠٢٢	٠,٠١٥	-	٠,٠٠٧
٢٢	٠,٠٣٠	٠,٠٠٤	-	٠,٠٥٥	٤٧	-	٠,٠٢٢	٠,٠١٠	٠,٠٣٠	٧٢	-	٠,٠١٢	٠,٠٠٥	٠,٠١٢
٢٣	-	٠,٠٢٠	٠,٠٣٠	٠,٠١٢	٤٨	٠,٠٢١	٠,٠٢٢	٠,٠٤٥	-	٧٣	٠,٠٣٢	٠,٠١٠	٠,٠٢٥	-
٢٤	٠,٠٥٠	٠,٠٢٢	٠,٠١٥	-	٤٩	٠,٠١٠	-	٠,٠٢١	٠,٠٠١	٧٤	٠,٠٢١	٠,٠٢٢	-	٠,٠٠٤
٢٥	٠,٠١١	٠,٠٢٠	٠,٠٤٠	-	٥٠	-	٠,٠٠٢	٠,٠١٥	٠,٠٢٢	٧٥	-	٠,٠٢١	٠,٠٣٥	٠,٠٠٩

يتبين من الجدول (١٣) أن جميع بدائل فقرات الأختبار التحصيلي فعالة، فهي بدائل قد جذبت أفراد المجموعة الدنيا أكثر مما جذبت أفراد المجموعة العليا، وهذا ما ينبغي أن تكون عليه البدائل الخاطئة للحكم على أنها فعالة،

١٠. **الخصائص السيكومترية للأختبار التحصيلي:** لزيادة دقة الاختبارات فقد ركز المختصون في القياس التربوي والنفسي على حساب الخصائص السيكومترية لها، وأنصب الأهتمام على خاصيتي الصدق والثبات والتي يتوجب توافرها في الاختبارات (عبد الرحمن،

١٩٩٨: ١٦٠). حيث إن الصدق يشير الى أن يقيس الأختبار ما وضع لأجل قياسه في حين يشير الثبات الى استقرار الأجوبة بفاصل زمني والتجانس الداخلي التكافؤ، ومن خلالهما يتم تقييم جودة وكفاية ودقة إجراءات القياس في البحث العلمي، (الغانم ،٢٠٢٢: ٢٤٣). وقد أتبع الباحث لحساب الصدق والثبات الإجراءات الآتية:

أولا – صدق الاختبار: تدعم الاستدلالات التي يمكن استخلاصها من درجات الأختبار والتي تمثل أدلة على صدق الأختبار (كروكر والجينا، ٢٠٠٩: ٢٩١)، والمقياس الصادق هو الذي يقيس فعلا ما وضع لأجله، ويرتبط الصدق بالهدف أو بالأهداف المتوقعة من أداة القياس تحقيقها (النبهان، ٢٠٠٤: ٢٧٢)، وقد حقق الباحث صدق الاختبار على النحو الاتي:

أ. الصدق الظاهري: ويشير الى قدرة الأختبار ظاهريا على قياس الخاصية المراد قياسها، ويتأكد المختصون من هذا النوع من الصدق بتقييم المصداقية التفسيرية باستخدام لجنة من المحكمين الخبراء، وقد قام الباحث بعرض الأختبار التحصيلي على مجموعة من المحكمين الخبراء وطلب من لجنة الخبراء ابداء حكمهم على مدى صلاحية كل فقرة من فقرات الاختبار من حيث التعديل أو الحذف، واعتمد الباحث على نسبة موافقة (١٠٠%)، ولم تحذف اي فقرة من فقرات الاختبار.

ب. صدق المحتوى: يعد تمثيل مفردات الاختبار للمحتوى الدراسي المقرر بدقة عبر تحليل المحتوى وان إعداد جدول المواصفات والذي قام به الباحث، يضمن توفير صدق المحتوى للاختبار (عبد الرحمن، ٢٠٠٨: ١٩٩).

ت. صدق البناء (التكوين الفرضي): شير النص إلى مفهوم صدق البناء (Construct Validity)، وهو قدرة الاختبار على التنبؤ بالتوقعات النظرية للسمة أو القدرة المقاسة، مفترضا أن الأداة تستند إلى نظرية تضع توقعات وافتراضات حول السمة (البطش وابو زينة ،٢٠٠٧: ١٣٢). لتحقيق صدق البناء، يجب أن يظهر الاختبار قدرته على إبراز الفروق المتوقعة، مثل اختلاف الأفراد في امتلاك الخاصية أو وجود فروق في الأداء بين الجماعات العمرية. تعد مؤشرات تحليل الفقرات كاستخراج معاملات الصعوبة والتمييز وفعالية البدائل الخاطئة دلائل جيدة على صدق البناء واتساق مكونات الاختبار في قياس نفس السمة (الامام واخرون، ١٩٩٠: ١٣٢-١٣٣). ويظهر تحقق هذا الافتراض أيضا من خلال أن تكون معاملات ارتباط درجات الفقرات بالدرجة الكلية دالة إحصائيا، مما يشير إلى تجانس فقرات الاختبارات في قياس المفهوم المراد.

ثانياً – ثبات الاختبار: يشير ثبات الأختبار إلى درجة الاستقرار أو الاتساق في الدرجات المتحققة على أداة القياس مع الزمن، والواقع أن عملية التوصل لمؤشرات عن ثبات الاختبار أو المقياس تعتبر أسهل بكثير من محاولة التوصل لمؤشرات عن صدق المقياس. فكلما اقتربت الدرجات الملاحظة للمفحوصين من الدرجات الحقيقية على السمة أو القدرة المقاسة كان

الاختبار ثابتا، وعندما تقترب قيمة معامل الثبات من (صفر) يشير إلى ضعف الثبات في الدرجات التي تتمخض عن أداة القياس، وكلما اقترب معامل الثبات من (١) دل ذلك على انخفاض تباين الخطأ، أي ارتفاع قيمة الثبات، وعلى أن الاختلافات التي يقيسها الاختبار هي اختلافات حقيقية تمثل الفروق الفردية الحقيقية في السمة أو القدرة التي يفترض الاختبار قياسها (البطش وابو زينة ،٢٠٠٧: ١٣٣). وقام الباحث بحساب الثبات للأختبار التحصيلي بعد اختيار عينة بصورة طبقية عشوائية تتألف من (١٠٠) طالب وطالبة، والجدول (١٤) يوضح ذلك.

الجدول (١٤) حجم عينة الثبات موزعة بحسب الكلية والقسم والجنس

الجامعة	الكلية	القسم	الجنس		المجموع
			ذكور	إناث	
بغداد	التربية ابن رشد	جغرافية	٢٥	٢٥	٥٠
المستنصرية	التربية	رياضيات	٢٥	٢٥	٥٠
المجموع			٥٠	٥٠	١٠٠

إذ قام الباحث بحساب الآتي:

أ. **طريقة "كودر ريتشاردسون ٢٠" (Kuder & Richardson 20)**: تمتاز هذه الطريقة بتناسقها وإمكانية الوثوق بنتائجها على حساب البيانات بين درجات جميع فقرات الاختبارات على اعتبار أن الفقرة عبارة عن مقياس قائم بذاته، ويشير معامل الثبات إلى اتساق أداء الفرد أي التجانس بين فقرات الاختبار (عودة، ١٩٨٨: ٣٥٤). ولاستخراج الثبات بهذه الطريقة، طبقت معادلة "كودر ريتشاردسون ٢٠" على درجات أفراد العينة، وكانت قيمة معامل الثبات للأختبار تساوي (٠.٨٤٥) وهو معامل ثبات جيد عند مقارنته بقيم معاملات الثبات الواردة في الدراسات السابقة.

ب. **معادلة هويت Hoyt:** وهي من الاساليب التي تعتمد على تجزئة التباين الكلي لدرجات الأفراد إلى مصادر ثلاثة للتباين حيث ترجع إلى الأفراد والفقرات وتباين الخطأ، فهي احدى طرائق استخراج الثبات المستعملة في الاختبارات ومن الأساليب الشائعة والمناسبة في تقدير الثبات. وأن فكرة تحليل ثبات الاختبار يمكن ان تقوم على تحليل التباين لعلاقات المفحوصين وعلى جميع فقرات الاختبار. (علام، ٢٠٠٠: ١٦٨). واستعمل الباحث تحليل التباين الثنائي (بدون تفاعل) لدرجات عينة الإفراد، ومن ثم استعملت معادلة هويت (Hoyt) من نتائج تحليل التباين والتي يمكن ان تعتمد على تباين الخطأ التباين بين الإفراد، حيث كانت معاملات ثبات الاختبار كما هو في الجدول (١٥).

الجدول (١٥) نتائج تحليل التباين الثنائي بدون تفاعل لدرجات عينة الثبات

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات
بين الإفراد	١٥١.٣٦٢	٩٩	١.٥٢٨

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات
بين الفقرات	٥١٦.٨٩٠	٧٤	٦.٩٨٥
الخطأ	٢٣٣٦.٩٩٤	٧٣٢٦	٠.٣١٩

وبهذا فقد بلغت قيمة معامل ثبات الأختبار (٠.٧٩١)، التي تم حسابها بطريقة تحليل التباين، وكانت كلها معاملات جيدة عند مقارنتها بالدراسات السابقة التي تناولت ثبات الاختبار التحصيلي.

١١. **الصيغة النهائية للأختبار:** يتكون الأختبار التحصيلي لمادة علم نفس النمو لطلبة الأقسام العلمية غير الاختصاص من (٧٥) فقرة صيغت بأسلوب الاختيار من متعدد مكون من أصل فقرة تتبعها أربعة بدائل، أحدها صحيحة والباقي خاطئة، وقد أجري تحليل منطقي للفقرات وتم التأكد من وضوح تعليمات الأختبار وفقراته وتم التحليل الاحصائي لمعاملات صعوبة الفقرات وتمييزها وفعالية البدائل الخاطئة، كما تم التحقق من صدق الأختبار وثباته.

الوسائل الإحصائية: استعمل الباحث الحقيبة الإحصائية للعلوم الاجتماعية SPSS لاستخراج:

١.معامل ارتباط (بوينت بايسيريال) لحساب ارتباط درجة الفقرة بالدرجة الكلية للاختبارات

٢.معاملات الصعوبة، والتمييز، وفعالية البدائل الخاطئة.

٣.تحليل التباين الثنائي (بدون تفاعل) لمعرفة التباين بين الأفراد وبين الفقرات، وتباين الخطأ (المتبقي) لغرض حساب الثبات.

٤.الاختبار التائي لعينتين مستقلتين (t-test): وذلك لغرض التعرف على دلالة الفرق بين متوسطي درجات الأفراد الذين حصلوا على أعلى درجات والأفراد الذين حصلوا على أدنى الدرجات.

٥.معادلة هويت لاستخراج الثبات وفق نظرية القياس الكلاسيكية.

الفصل الرابع: عرض النتائج ومناقشتها:

الهدف الاول: بناء اختبار تحصيلي لمادة (علم نفس النمو – لغير الاختصاص) لعينة البحث الحالي.

أتبع الباحث جميع الإجراءات والخطوات العملية اللازمة لبناء الاختبارات التحصيلية لتحقيق الهدف الأول ويتبين ذلك من الخطوات الواردة في الفصل الثالث، إذ بدأً بتحديد المحتوى الدراسي وصياغة الأهداف السلوكية وفق تصنيف بلوم، تبعتها إجراءات التحليلات المنطقية ممثلة بالتحقق من صلاحية تعليمات الاختبار التحصيلي وفقراته ثم التحليل الإحصائي للتحقق من الخصائص السيكومترية لفقرات الاختبار وقد تبين ان جميع الفقرات تتمتع بخصائص سيكومترية جيدة، وتم الأبقاء على جميع الفقرات. وبذلك تم تحقيق هذا الهدف.

الهدف الثاني: استخراج الخصائص السيكومترية للاختبار الحالي وفقا لنظرية القياس الكلاسيكية.

قام الباحث بحساب الخصائص السيكومترية لفقرات الاختبار وللاختبار كاملا، وكانت خصائص جيدة في عمومها ومقبولة تربويا، وكما هو واضح في الخطوة (٩-١٠) من خطوات بناء الاختبار التحصيلي. وبذلك تم تحقيق هذا الهدف

الاستنتاجات: بعد التحقق من إجراءات بناء الاختبار التحصيلي لمادة علم نفس النمو وما توصل إليه من نتائج بناء الاختبار وحساب خصائصه السيكومترية، يستنتج ما يأتي:

١.التوزيع ألاعتدالي للدرجات الكلية للأختبار التحصيلي لمادة علم نفس النمو، يشير الى تمثيل عينة البحث لمجتمع البحث.

٢.إن الاختبار التحصيلي لمادة علم نفس النمو هو مناسب لأفراد العينة (طلبة الصف الثاني للأقسام غير المتخصصة)، وذلك لاتساق قدرات أفراد العينة في تحقيق هدف البحث الحالي الأول.

٣.الخصائص السيكومترية لفقرات الاختبار وللاختبار كاملا، كانت جيدة في عمومها ومقبولة تربويا.

التوصيات: في ضوء نتائج التحليل التي تم توصل لها وتفسيرها، يمكن للباحث ان يوصي بالآتي:

١. استعمال الباحثون في القياس والتقويم وعلم نفس النمو، للاختبار التحصيلي الحالي وعدم الاقتصار على العينة الحالية طلبة الأقسام العلمية غير الاختصاص في كليات التربية، وإنما يمكن استعمال الاختبار الحالي مع طلبة أقسام الأختصاص، مع بعض التعديلات الضرورية لتكييفه للعينة الجديدة.

٢. ضرورة قيام الدارسون والباحثون بأستعمال الاختبار التحصيلي لمادة علم نفس النمو في الدراسات الميدانية التي يكون جزء من أهدافها التعرف على النمو الجسمي والعقلي والانفعالي والأجتماعي للطلبة.

٣.أجراء مراجعة للأختبار التحصيلي لمادة علم نفس النمو لأقسام غير الأختصاص من المتخصصين بالقياس النفسي لتطوير الأختبار وتضمين فقرات مصاغة بأساليب أخرى، بإدراج فقرات أختبارات أخرى بنيت لقياس المادة الدراسية نفسها.

المقترحات: استكمالا للإفادة المرجوة لتطوير البحث الحالي ووفقا لنتائجه، يقترح الباحث الاتي:

١.إجراء دراسة بإضافة متغير آخر يتعلق بخصائص العينة مثل (خصائص توزيع العينة او حجم العينة) باستعمال تصميم تجريبي أحادي العامل او أي تصميم عاملين، وبحسب مستويات المتغيرين.

٢.إعادة إجراءات البحث ذاتها وباستعمال مادة مقرر دراسي اخرى، ومقارنة نتائجها مع نتائج البحث الحالي.

٣.إجراء دراسة لنفس المادة باستخدام اجراءات نظرية القياس الحديثة (IRT) في بناء الاختبار التحصيلي ومقارنة نتائج الخصائص السيكومترية للاختبارين.

المصادر العربية:

١.أبو حطب، فؤاد وعثمان، سيد أحمد وصادق، آمال (١٩٩٣). **التقويم النفسي**. مكتبة الأنجلو المصرية.

٢.أبو فودة، باسل خميس وبني يونس، نجاتي أحمد (٢٠١٢). **الاختبارات التحصيلية**. عمان: دار المسيرة.

٣.الإمام، مصطفى محمود وعبد الرحمن، أنور حسين والعجلي، صباح حسين (١٩٩٠). **التقويم والقياس**. جامعة بغداد، كلية التربية – ابن رشد.

٤.أنستازي، آن واورينا، سوزانا (٢٠١٥). **القياس النفسي**. ترجمة صلاح الدين محمود علام. عمان: دار الفكر.

٥.بركات، زياد (٢٠١٨). **القياس والتقويم النفسي والتربوي بين النظرية والتطبيق**. فلسطين: جامعة القدس المفتوحة.

٦.البطش، محمد وليد وأبو زينة، فريد كامل (٢٠٠٧). **مناهج البحث العلمي: تصميم البحث والتحليل الإحصائي**. عمان: دار المسيرة.

٧.حبيب، صفاء طارق وكاظم، بلقيس حمود (٢٠١٨). **نظريتي القياس الحديثة والتقليدية**. عمان: الدار المنهجية للنشر.

٨.الحجامي، بلقيس حمود (٢٠٠٩). **بناء اختبار محبوك في مادة علم النفس التربوي على وفق الإستراتيجية ثنائية المرحلة في نظرية السمات الكامنة**. اطروحة دكتوراه غير منشورة.

٩.خيري، السيد محمد. (١٩٧٠). **الاحصاء في البحوث النفسية والتربوية والاجتماعية**، القاهرة: دار النهضة العربية.

١٠.داود، أحمد عيسى (٢٠١٤): **أصول التدريس**. عمان: دار يافا العلمية للنشر والتوزيع،

١١.دوران، رودي (١٩٨٥). **أساسيات التقويم في تدريس العلوم**. ترجمة محمد سعيد صبار وآخرين. الأردن: دار الأمل.

١٢.رحيم، هند صبيح (٢٠١٢) **دراسة مقارنة في الخصائص القياسية لمقياس دافعية التعلم المبني وفقاً لنظرية القياس التقليدية ونظرية السمات الكامنة**. اطروحة دكتوراه غير منشورة.

١٣. الريماوي، عمر طالب (٢٠١٧). **بناء الاختبارات والمقاييس النفسية والتربوية**. عمان: دار أمجد للنشر والتوزيع.

١٤. الزاملي، علي عبد جاسم والصارمي، عبد الله بن محمد وكاظم، علي مهدي (٢٠٠٩). **مفاهيم وتطبيقات في التقويم والقياس التربوي**. مكتبة الفلاح.

١٥. زاير، سعد علي والطائي، نعمة دهش وجاسم، وسن عباس (٢٠٢٠). **بنك الاختبارات اللغوية**. عمان: دار الرضوان.

١٦. سليمان، امين علي محمد وابو علام، رجاء محمود (٢٠١٢). **القياس والتقويم في العلوم النفسية**. القاهرة: دار الكتاب الحديث.

١٧. شحاته، حسن .(2008) .**مناهج البحث التربوي** .القاهرة: دار العالم العربي.

١٨. الطريري، عبد الرحمن بن سليمان (١٩٩٧). **القياس النفسي والتربوي**. الرياض: مكتبة الرشيد.

١٩. عبد الرحمن، أحمد محمد (٢٠١١). **تصميم الاختبارات: أسس نظرية وتطبيقات عملية**. عمان: دار أسامة.

٢٠. عبد الرحمن، سعد (٢٠٠٨). **القياس النفسي: النظرية والتطبيق**. القاهرة: هبة النيل. ط٥.

٢١. عبيدات، سلامة، والجندي، إيمان (٢٠٢٤). **البحوث التربوية والنفسية: أسسها ومناهجها**. عمان: دار الثقافة.

٢٢. العجيلي، عبد الكريم (٢٠١٧). **القياس والتقويم التربوي والنفسي**. بغداد: دار الحكمة.

٢٣. علام، صلاح الدين محمود (٢٠٠٦). **الاختبارات والمقاييس التربوية والنفسية**. عمان: دار الفكر للنشر والتوزيع.

٢٤. عمر، محمود أحمد وفخرو، حصة عبد الرحمن والسبيعي، تركي وتركي، آمنة عبد الله (٢٠٠٩). **القياس النفسي والتربوي**. عمان: دار المسيرة.

٢٥. عودة، احمد سليمان (١٩٨٥). **القياس والتقويم في العملية التدريسية**. دار الامل للطباعة والنشر

٢٦. عودة، أحمد سليمان والخليلي، خليل يوسف (١٩٨٨). **الإحصاء للباحث في التربية والعلوم الإنسانية**. دار الفكر للنشر والتوزيع، عمان، الأردن.

٢٧. الغانم، ستار (٢٠٢٢). **القياس النفسي**. عمان: دار أمجد.

٢٨. فرج، صفوت (٢٠٠٧). **القياس النفسي**. القاهرة: مكتبة الأنجلو المصرية.

٢٩. كاظم، أميمة محمد (١٩٨٨). **القياس الموضوعي للسلوك**. الكويت: مؤسسة الكويت للتقدم العلمي.

٣٠. الكبيسي، عبد الواحد (٢٠٠٧). **القياس والتقويم**. عمان: دار جرير.

٣١. كروكر، ليندا والجينا، جيمز (٢٠٠٩). **مدخل إلى نظرية القياس التقليدية والمعاصرة**. ترجمة زينات يوسف دعنا. عمان: دار الفكر.

٣٢. محاسنة، إبراهيم محمد (٢٠١٣). **القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة**. عمان: دار جرير.

٣٣. مخائيل، امطانيوس (٢٠٢١). **القياس والتقويم في التربية الحديثة**. جامعة دمشق.

٣٤. مخائيل، امطانيوس نايف (٢٠١٦). **بناء الاختبارات، والمقاييس النفسية، والتربوية، وتقنينها**. عمان: دار الإعصار العلمي.

٣٥. مراد، صلاح أحمد وسليمان، أمين علي (٢٠٠٢). **الاختبارات والمقاييس في العلوم النفسية والتربوية – خطوات إعدادها وخصائصها**. الكويت: دار الكتاب الحديث.

٣٦. ملحم، سامي محمد (٢٠١٢). **القياس والتقويم في التربية وعلم النفس**. عمان: دار المسيرة.

٣٧. المندلاوي، ضياء عبد الخالق والمندلاوي، علاء عبد الخالق (٢٠٢١). **القياس والتقويم في التربية وعلم النفس**. بغداد: مكتبة الأمير.

٣٨. المياحي، جعفر عبد كاظم (٢٠١١). **القياس النفسي والتقويم التربوي**. عمان: دار كنوز المعرفة.

٣٩. المياحي، جعفر عبد كاظم (٢٠١١). **القياس النفسي والتقويم التربوي**. عمان: دار كنوز المعرفة.

٤٠. النبهان، موسى (٢٠١٣). **أساسيات القياس في العلوم السلوكية**. عمان: دار الشروق.

٤١. اليعقوبي، حيدر (٢٠١٣). **التقويم والقياس في العلوم التربوية والنفسية**. بغداد: مركز المرتضى.

المصادر الأجنبية:

1. Anastasi, A. (1976). **Psychological testing** (5th ed.). New York: Macmillan.

2. Anastasi, A. (1982). **Psychological testing** (6th ed.). New York: Macmillan.

3. Anastasi, A. (1988). **Psychological testing** (6th ed., reprint). New York: Macmillan.