الباب الثالث

منهجية البحث

أ- منهج ونوع البحث

اعتمدت الباحثة في هذا البحث المنهج الكمي، وتحديدا البحث التجريبي. ووفقا لسوجيونو (Sugiyono)، فإن المنهج الكمي هو منهج يقوم على الفلسفة الوضعية (Positivism)، ويستخدم لدراسة مجتمع أو عينة معينة. حيث تجمع البيانات باستخدام أدوات بحثية، ثم تحلل تحليلا كميا (إحصائيا) بهدف اختبار الفرضيات التي وضعت مسبقا.[1]

وبناء على هذا المنهج، اختارت الباحثة التصميم شبه التجريبي (Pre-Experimental Design)، وتحديدا نموذج المجموعة الواحدة ذات الاختبار القبلي والبعدي (One Group Pretest–Posttest Design). أما تطبيق هذا التصميم، فيتم من خلال إجراء اختبار قبلي (Pre-test) قبل تقديم المعالجة (التجربة)، ثم إجراء اختبار بعدي (Post-test) بعد تطبيقها. وبذلك تكون نتائج البحث أكثر دقة، حيث تتيح مقارنة مستوى المتعلمين قبل المعالجة وبعدها.

بناء على ما سبق، يعتمد هذا البحث المنهج التجريبي. نظرا لاقتصاره على مجموعة تجريبية واحدة دون الحاجة إلى مجموعة ضابطة. ويهدف هذا التصميم إلى قياس مدى فاعلية وسيلة Educaplay في تنمية مهارة الاستماع لدى المتعلمين. ويمكن تمثيل تصميم البحث (One-Group Pretest-Posttest Design) وفق النموذج الآتي:

الجدول ٣،١ تصميم البحث (One-Group Pretest-Posttest Design)

إختبار قبلي	أنشطة التعليم	اختبار البعدي
O_1	X	O_2

[1] Prof. Dr Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif Dan R&D* (Bandung: Alfabeta, 2022). 110-114

يوضح الجدول (٣،١) تصميم البحث المعتمد، وهو تصميم المجموعة الواحدة مع اختبارين قبلي وبعدي (One Group Pre-test Post-test Design)، والذي يندرج ضمن التصاميم شبه التجريبية. حيث توضيح الرموز فيما يلي:

O_1 : تمثل درجة الاختبار القبلي لقياس مهارة الاستماع (قبل تطبيق المعالجة).

X : تمثل المعالجة أو المتغير المستقل (استخدام وسيلة إيدوكابلاي Educaplay).

O_2 : تمثل درجة الاختبار البعدي لقياس مهارة الاستماع (بعد تطبيق المعالجة)

ب- مجتمع البحث وعينته:

يقصد بمجتمع البحث جميع المفردات التي تمثل مجال التعميم في الدراسة، ولا يقتصر هذا المجتمع على الأفراد فحسب، بل يمتد ليشمل الموضوعات والظواهر محلّ الدراسة، وكافة الخصائص والسمات المرتبطة بها.[٢] أما عينة البحث، فهي جزء ممثل لمجتمع البحث تم اختيارها لتعكس خصائصه بدقة. ويتمثل مجتمع البحث في هذه الدراسة في جميع طلاب الصف السابع المتوسطة الإسلامية الحكومية الثانية بمدينة "كديري"، والبالغ عددهم (٤٧٨) طالبا.

ينما تمثلت عينة البحث في فصل دراسي واحد، وهو الصف السابع (ج)، ويضم (٣٢) طالبا. وقد اختيرت العينة بأسلوب المعاينة العنقودية (Cluster Sampling)، وذلك عبر اختيار فصل كامل يتمتع بخصائص متجانسة نسبيا. ويُعد حجم هذه العينة كافيا ومستوفيا للمتطلبات المنهجية في البحث الكمي القائم على تصميم المجموعة الواحدة باختبار قبلي – بعدي(pre test - post test).

ج- أسلوب جمع البيانات

تعتمد هذه الدراسة على مجموعة من التقنيات لجمع البيانات، وهي: الملاحظة، والاختبار، والتوثيق. وفيما يلي تفصيل لكل أداة:

أ) الملاحظة

[2] Prof. Dr Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif Dan R&D*, oktober (Bandung: Alfabeta, 2022).

أما الملاحظة هي تعد الملاحظة أداة لجمع البيانات تعتمد على المشاهدة المباشرة والرصد المنظم للظواهر محل الدراسة.[٣] وفي هذا البحث، طبقت الملاحظة من خلال تتبع الأنشطة الصفية أثناء استخدام وسيلة إيدوكابلاي Educaplay لدى طلبة الصف السابع بمدرسة المتوسطة الإسلامية الحكومية ٢ بكديري. وتستهدف هذه الأداة رصد عملية تطبيق الوسيلة التعليمية، بما يشمل مستويات تفاعل المتعلمين، واستجاباتهم، ومدى مشاركتهم أثناء التعلم، وصولا إلى تقييم فاعلية الوسيلة في تنمية مهارة الاستماع.

ب) الاختبار

يعد الاختبار من الأدوات الجوهرية في البحث الكمي، وهو عبارة عن مجموعة من الأسئلة أو المثيرات التي تقدم للمبحوثين لقياس سمة معينة وتحويلها إلى قيم رقمية.[٤] وقد اعتمد البحث على نوعين من الاختبارات منها: الاختبار القبلي هو يطبق لقياس المستوى المعرفي الأولي للمتعلمين قبل البدء بالمعالجة التجريبية. الاختبار البعدي هو يطبق لقياس مخرجات التعلم بعد تنفيذ المعالجة المتمثلة في استخدام الاختبارات الفيديوية التفاعلية عبر استخدام وسيلة إيدوكابلاي Educaplay، وذلك لمعرفة مدى التطور الحاصل في مهارة الاستماع.

ج) التوثيق

تستخدم تقنية التوثيق لجمع الأدلة المادية والإدارية التي تدعم صدق البحث. وتشمل البيانات الوثائقية في هذه الدراسة: الصور الفوتوغرافية أثناء وسيلة إيدوكابلاي Educaplay وسجلات نتائج الاختبارات، وقوائم حضور الطلاب، بالإضافة إلى كافة الوثائق الرسمية التي توثق مراحل تنفيذ التجربة البحثية.

د– أداة البحث

---

[3] Zuchri Abdussamad, *Metode Penelitian Kualitatif* (Cv. Syakir Media Press, 2021). http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_Sistem_Pembetungan_Terpusat_Strategi_Melestari.

[4] Abigail Soesana et al., *Metodologi Penelitian Kuantitatif*, ed. Abdul Karim (Yayasan Kita Menulis, 2023).

تعرف أدوات البحث بأنها الوسائل والتقنيات التي يعتمد عليها الباحثة في عملية جمع البيانات، وقياس المتغيرات، وتحليلها بما يتفق مع أهداف الدراسة ومشكلتها. وتتعدد هذه الأدوات في البحوث التربوية لتشمل الاختبارات، والاستبانات، وبطاقات الملاحظة، والمقابلات الشخصية.[٥] ونظرا لطبيعة هذا البحث الذي يتبع المنهج الكمي وفق تصميم "ما قبل التجربة" (Pre-Experimental Design)، فقد تمثلت أدوات البحث المستخدمة فيما يلي:

١) أداة الاختبار

يعد الاختبار أداة علمية لجمع البيانات من خلال عمليتي القياس والتقويم. وذلك لتحديد مستويات القدرة، أو المهارة، أو المخرجات التعليمية لدى المتعلمين.[٦] وتستخدم أداة الاختبار في هذا البحث لقياس مدى إتقان المتعلمين لمهارات اللغة العربية المستهدفة. وقد جرى إعداد الاختبار استنادا إلى مؤشرات مهارة الاستماع، والتي تشمل: تمييز الأصوات اللغوية العربية، وفهم المفردات، واستيعاب المعنى العام ومحتوى النص المسموع. ويطبق هذا الاختبار على صورتين: الاختبار القبلي لقياس المستوى المعرفي الأولي للمتعلمين، والاختبار البعدي لقياس مستوى التطور والتحصيل النهائي بعد تنفيذ المعالجة التجريبية. كما تم تحويل كل مؤشر من هذه المؤشرات إلى بنود اختبارية (أسئلة) قدمت عبر وسيلة إيدوكابلاي Educaplay في قالب تدريبات تفاعلية تعتمد على الوسائط الصوتية.

الجدول ٢،٣ مواصفات أداة الاختبار في ضوء مهارة الاستماع

جوانب مهارة الاستماع	مؤشرات مهارة الستماع
تمييز الأصوات اللغوية	قدرة الطلاب على التمييز بين الأصوات المتقاربة والتنغيم الصوتي في اللغة العربية المسموعة.

[5] Heru Kurniawan, *Pengantar Praktis Penyusunan Instrumen Penelitian* (Sleman: CV. Budi Utama, 2021).1
[6] Supriyadi, *Evaluasi Pembelajaran*, 1st ed. (Pekalongan: PT Nasya Expanding Management, 2020).11-13

استيعاب المفردات الشفوية	قدرة الطلاب على فهم وتحديد معاني المفردات والتراكيب الواردة في المقطع الصوتي.
فهم المعنى الإجمالي (الاستيعاب)	قدرة الطلاب على استخلاص المعلومات الرئيسة، وتحديد الفكرة العامة للمقطع الصوتي المسموع.
الاستجابة للمحتوى الشفوي	قدرة الطلاب على الإجابة عن الأسئلة بدقة بناء على ما استمع إليه في المقطع الصوتي

هـ- أساليب تحليل البيانات

تعد مرحلة تحليل البيانات في البحث الكمي خطوة تالية لعملية جمع البيانات من المبحوثين أو من مصادر الدراسة الأخرى. ويهدف التحليل إلى الإجابة عن تساؤلات البحث واختبار الفرضيات الموضوعة من خلال المعالجات الإحصائية اللازمة. وتعرض بيانات هذا البحث في صورة قيم رقمية (درجات) استمدت من نتائج الاختبارين القبلي والبعدي لمهارة الاستماع المطبقة على مجموعة الدراسة. وتتمثل أساليب التحليل المستخدمة فيما يلي:

١) التحليل الإحصائي الوصفي

يستخدم التحليل الوصفي لوصف البيانات المجموعة وتلخيصها. حيث يتم تحليل درجات إتقان مهارة الاستماع لدى المتعلمين بحساب المتوسط الحسابي (Mean)، والقيمتين العظمى والصغرى، والانحراف المعياري وذلك للحصول على صورة عامة وشاملة لمستوى المتعلمين قبل المعالجة وبعدها.[٧]

صيغة حساب المتوسط الحسابي:

$$\bar{X} = \frac{\sum X}{N}$$

[7] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif Dan R&D*, 2022.

توضيح:

$\bar{X}$: ترمز إلى المتوسط الحسابي

$\sum X$: مجموع درجات المتعلمين

N : عدد أفرادالعينة

٢) اختبار اعتدالية البيانات (اختبار التوزيع الطبيعي)

يهدف اختبار اعتدالية البيانات إلى التحقق مما إذا كانت بيانات البحث تتبع التوزيع الطبيعي (Normal Distribution)، بوصفه شرطا أساسيا لاستخدام الاختبارات الإحصائية المعلمية (البارامترية). وفي هذه الدراسة، يعد اختبار الاعتدالية متطلبا ضروريا. لأن اختبار الفرضياتِ يرتكز على اختبار (t) للعينات المرتبطة (Paired Sample t-test)، والذي يشترط اعتدالية توزيع البيانات.

ويستخدم في هذه الدراسة اختبار شابيرو—ويلك (Shapiro- Wilk) لاختبار اعتدالية بيانات الاختبارين القبلي والبعدي. وقد وقع الاختيار على هذا الاختبار نظرا لأن حجم العينة أقل من (٥٠) مشاركا، حيث يعد اختبار "Shapiro -Wilk" الأنسب والأكثر حساسية في الكشف عن اعتداليةِ التوزيع في العينات الصغيرة. أجريت المعالجة الإحصائية لاختبار الاعتدالية باستخدام برنامج (SPSS). وتتمثل صيغة معادلة "Shapiro -Wilk" فيما يلي:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

توضيح:

W : قيمة إحصاء Shapiro Wilk

$x_{(i)}$: البيانات المرتبة من أصغر قيمة إلى أكبر قيمة

$\bar{x}$: المتوسط الحسابي للبيانات

a_i : ثوابت Shapiro Wilk

n : عدد أفراد العينة

أما معايير اتخاذ القرار في اختبار الاعتدالية، يتم الحكم على طبيعة توزيع البيانات بناء على قيمة الدلالة الإحصائية (p-value) هي:

١) إذا كانت قيمة الدلالة ($p > ٠,٠٥$) فإن البيانات تتبع التوزيع الطبيعي (normal distribution) وهو ما يمسح باستخدام الاختبارات البارامترية.

٢) إذا كانت قيمة الدلالة ($P \leq ٠,٠٥$) فإن البيانات لا تتبع التوزيع الطبيعي، وفي هذه الحالة يتم للجوء إلى الإختبارات غير البارامترية.[٨]

٣) اختبار الفرضية (اختبار ويلكوكسون للرتب الموقعة)

يعد اختبار ويلكوكسون للرتب الموقعة (Wilcoxon Signed-Rank Test) من الأساليب الإحصائية غير المعلمية التي تستخدم لتحديد الفروق بين مجموعتين من البيانات المرتبطة، أو الناتجة عن قياسين متكررين على المجموعة ذاتها مثل نتائج الاختبارين القبلي (Pre-test) والبعدي (Post-test) للمشاركين أنفسهم. ويستخدم هذا الاختبار كبديل لاختبار (Paired Sample t-test) في حال عدم استيفاء البيانات لشرط التوزيع الطبيعي.

وتعتمد الية إجراء اختبار ويلكوكسون على حساب الفروق بين درجات القياسين، ثم ترتيب هذه الفروق تصاعديا ومنحها رتبا محددة. وذلك للتحقق مما إذا كان التغير الحاصل ذا دلالة إحصائية معنوية أم لا.[٩] في هذا البحث، استخدم اختبار ويلكوكسون للرتب الموقعة (Wilcoxon Signed-Rank Test) للتحقق من وجود فرق ذي دلالة إحصائية بين درجات الاختبارين القبلي (Pre-test) والبعدي (Post-test) في مهارة الاستماع لدى الطلبة عقب تقديم المعالجة التجريبية في العملية التعليمية.

---

[8] Theodoros Kyriazos and Mary Poga, "Dealing with Multicollinearity in Factor Analysis : The Problem , Detections , and Solutions," 2023, 404–24, https://doi.org/10.4236/ojs.2023.133020.

[9] Dika Rahayu Widiana, Imaniah Sriwijayasih, and Irma Rustini Aju, "The Use of the Wilcoxon Signed Rank Test in Analyzing the Difference in Test Scores Before and After Digital Marketing Training ." 8, no. 2 (2025): 22–28, https://doi.org/10.35991/jtm.v8i2.73.

وقد اعتمد هذا الاختبار استنادا إلى نتائج اختبار التوزيع الطبيعي التي أظهرت عدم اتباع البيانات للتوزيع الطبيعي؛ مما تعذر معه استخدام الاختبارات الإحصائية المعلمية في اختبار الفرضية، والاستعاضة عنها بالاختبارات الإحصائية غير المعلمية. وأما الصيغة الإحصائية لاختبار ويلكوكسون فيمكن التعبير عنها كما يأتي:

$$T = \min(T^+, T^-)$$

التوضيح:

T : قيمة الإحصاء في اختبار ويلكوكسون للرتب الموقعة (Wilcoxon Signed-Rank Test)

T^+ : مجموع الرتب للفروق ذات الإشارة الموجبة

T^- : مجموع الرتب للفروق ذات الإشارة السالبة

وفي تحليل البيانات باستخدام برنامج (IBM SPSS)، جرى اختبار الفرضية من خلال الاعتماد على قيمة الدلالة الإحصائية (Asymp. Sig. 2-tailed) في اختبار ويلكوكسون. وقد تمثلت معايير اتخاذ القرار في هذا البحث فيما يأتي:

١) رفض الفرضية الصفرية (H0): إذا كانت قيمة الدلالة الإحصائية (Sig) أقل من أو تساوي (٠,٠٥)، ففي هذه الحالة تُرفض الفرضية الصفرية وتقبل الفرضية البديلة (Ha) مما يشير إلى وجود فرق دال إحصائيا بين نتائج الاختبارين القبلي (Pre-test) والبعدي (Post-test).

٢) قبول الفرضية الصفرية (H0): إذا كانت قيمة الدلالة الإحصائية (Sig) أكبر من (٠,٠٥)، ففي هذه الحالة تقبل الفرضية الصفرية وترفض الفرضية البديلة (Ha) مما يشير إلى عدم وجود فرق دال إحصائيا بين نتائج الاختبارين القبلي (Pre-test) والبعدي (Post-test).وبذلك، يستخدم اختبار ويلكوكسون للتحقق مما إذا كان للمعالجة التجريبية أثر في تنمية مهارة الاستماع لدى الطلبة من عدمه.

٤) اختبار كسب الدرجات (N-Gain Score)

إلى جانب اختبار الفرضية باستخدام اختبار "ويلكوكسون" (Wilcoxon)، استخدمت الباحثة في هذه الدراسة تحليل كسب الدرجات (N-Gain) لقياس مستوى تطور قدرات المتعلمين بعد تقديم المعالجة التعليمية. ويستخدم هذا التحليل لتقييم مقدار الزيادة في نواتج التعلم بناء على المقارنة بين نتائج الاختبار القبلي (Pre-test) والاختبار البعدي (Post-test).

وتهدف عملية حساب (N-Gain) إلى تحديد مدى فاعلية العملية التعليمية من خلال رصد مستوى تحسن قدرات المتعلمين بعد تلقيهم المعالجة التعليمية المحددة. ومن خلال هذا التحليل، تمكنت الباحثة من الوقوف على مستوى التطور في نتائج التعلم بعد تطبيق الوسيلة التعليمية المستخدمة في هذه الدراسة.[١٠] أما الصيغة الرياضية المتبعة لحساب (N-Gain) فهي كما يأتي:

$$N\text{-}Gain = \frac{Skor\ Posttest - Skor\ Pretest}{Skor\ Maksimal - Skor\ Pretest}$$

التوضيح:

الاختبار البعدي (Post-test): هو الدرجة التي حصل عليها المتعلمون بعد تلقي المعالجة التعليمية.

الاختبار القبلي (Pre-test): هو الدرجة التي حصل عليها المتعلمون قبل البدء في المعالجة التعليمية.

الدرجة لعظمى (Ideal Score): هي أعلى درجة ممكنة يمكن أن يحققها المتعلمون في الاختبار (وهي ١٠٠ في هذه الدراسة).

ثم تفسر قيمة (N-Gain) المستخرجة استنادا إلى معايير "مؤشر الكسب" (Gain Index) التي وضعها ريتشارد هيك (Richard R. Hake)، كما هو موضح في الجدول (٣,٣) الآتي:

[10] Hendra Rustantono et al., "Pengembangan Media Membaca Karpaca Untuk Meningkatkan Kemampuan Membaca Siswa Kelas 1 Sekolah Dasar," *Primed (Primary Education Journal)* 3, no. 2 (2023): 76.

الجدول ٣،٣ معايير تفسير قيمة كسب الدرجات (N-Gain)

القيمة	الفئة
N-Gain أقل من ٠،٣ (g<٠،٠٣)	منخفض
N-Gain تتراوح ما بين ٠،٠٣ و ٠،٠٧ (٠،٧ >g ٠،٣)	متوسط
N-Gain اكبر من ٠،٧ (g > ٠،٧)	مرتفع

يبين الجدول (٣،٣) معايير مؤشر الكسب (Gain Index)؛ حيث يستخدم هذا التفسير لتحديد مستوى تطور نواتج تعلم المتعلمين بعد تقديم المعالجة التعليمية. فإذا وقعت قيمة (N-Gain) ضمن "الفئة المرتفعة"، فإن ذلك يشير إلى أن مستوى تحسن نواتج التعلم لدى المتعلمين يُعد مرتفعا. أما إذا وقعت هذه القيمة ضمن "الفئة المتوسطة" أو "المنخفضة"، فإن مستوى التطور في نتائج التعلم يصنف تبعا لذلك ضمن الفئة المقابلة (المتوسطة أو المنخفضة).

و- تأكيد صحة البيانات

يقصد بالصدق (Validity) مدى دقة الأداة وملاءمتها لقياس المتغيرات التي صممت من أجلها، وبناء عليه يحكم على الأداة بأنها صادقة. وفي البحث العلمي، يعد صدق البيانات معيارا جوهريا لتحديد مدى مطابقة النتائج الفعلية لما يقدمه الباحث من خلال المعالجات الإحصائية. وتشمل إجراءات التحقق من صلاحية الأداة في هذا البحث "الصدق" و"الثبات"، وذلك وفق التفصيل الاتي:

أ) صدق المحتوى (Content Validity):

يشير صدق المحتوى إلى مدى ملاءمة بنود الاختبار وصياغتها لتغطية كافة المؤشرات المراد قياسها.[11] حيث يركز على مدى التوافق بين الأسئلة ومؤشرات التعلم

[11] Soesana et al., *Metodologi Penelitian Kuantitatif*.

الخاصة بمهارة الاستماع. ولتحقيق ذلك، عرضت الأداة على مجموعة من المحكمين (Expert Judgment)، حيث تم إشراك خبيرين في اللغة العربية لتقييم مدى ملاءمة كل بند من بنود الاختبار لمؤشرات الأداء وأهداف القياس مهارة الاستماع.[١٢] وقد تم هذا التقييم باستخدام مقياس رباعي التدرج، وفقا لما يلي:

الجدول ٣،٤ مقياس رباعي التدرج

مقياس التقييم	الفئة
١	غير مناسب
٢	أقل ملائمة
٣	مناسب
٤	مناسب جدا

استنادا الى الجدوال ٣،٤ يدل على مقياس رباعي التدرج تحلل نتائج تقييم الخبراء باستخدام معامل أيكن (Aiken's V) للتحقق من معامل صدق المحتوى لكل بند من بنود الاختبار وذلك وفق الصيغة التالية[١٣]:

$$V = \frac{\sum s}{\mathrm{n(c-1)}}$$

توضيح:

V : معامل صدق المحتوى لنبد السؤال

s : القيمة الناتجة عن طرح أدنى درجة في المقياس من الدرجة التي منحها الخبير $(s = r - l_o)$

r : الدرجة التي يمنحها الخبير للبند

l_o : أدنى درجة في مقياس التقييم

[12] Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan* (Jakarta: Bumi Aksara, 2021).
[13] Nur Hidayah, "Educaplay as Teaching Media Inn Vitual Clasess," in *The 3*, vol. 1, 2023, 75–82.

n : عدد الخبراء (المحكمين)

c : عدد فئات مقياس التقييم (أعلى درجة في المقياس)

تتراوح قيمة معامل أيكن (Aiken's V) بين ٠ و ١ حيث يشير اقتراب القيمة من ١ إلى ارتفاع مستوى صدق المحتوى للبند. وفي هذا البحث، اعتمدت المعايير الآتية للحكم على صدق البنود:

١) إذا كانت قيمة V ≥ ٠،٨٠ فإن البند يعد صادقا (Valid).

٢) إذا كانت قيمة V < ٠،٨٠، فإن البند يعد غير صادق أو بحاجة إلى تعديل جذري.

وبناء على ذلك، فإن البنود التي يثبت صدقها تدرج ضمن أداة البحث النهائية، أما البنود التي لم تستوف المعايير، فيعاد النظر فيها إما بالتعديل والتحسين وفقا لمقترحات الخبراء، أو بالحذف إذا لم تكن ملائمة.

بعد أن ثبتت صلاحية الأداة من حيث صدق المحتوى من خلال تحكيم الخبراء (Expert Judgment) وتحليلها باستخدام معادلة "أيكن" (Aiken's V)، تم إجراء اختبار الصدق التجريبي لفقرات الاختبار؛ للوقوف على مستوى صدق كل فقرة استنادا إلى نتائج إجابات الطلبة. وقد أجري هذا الاختبار باستخدام معامل ارتباط "بيرسون" (Product Moment Pearson) بين درجة كل فقرة والدرجة الكلية للاختبار. أما الصيغة الرياضية المستخدمة فهي كما يأتي:

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

التوضيح:

r_{xy} = معامل الارتباط بين درجة فقرة الاختبار والدرجة الكلية

X = درجة فقرة الاختبار

Y = الدرجة الكلية

N = عدد المستجيبين (أفراد العينة)

$\sum XY$= مجموع حاصل ضرب درجات فقرات الاختبار في الدرجات الكلية

$\sum X$ = مجموع درجات كل فقرة من فقرات الاختبار

$\sum Y$ = مجموع الدرجات الكلية

$\sum X^2$= مجموع مربعات درجات فقرات الاختبار

$\sum Y^2$= مجموع مربعات الدرجات الكلية

وقد أجري هذا الاختبار باستخدام برنامج (SPSS) من خلال استخراج قيمة معامل الارتباط المصحح (Corrected Item-Total Correlation). وتعد فقرة الاختبار صادقة (صالحة) إذا كانت قيمة (r المحسوبة) أكبر من قيمة (r الجدولية) عند مستوى دلالة (٠,٠٥). كما يستخدم معامل ارتباط "بيرسون" (Product Moment Pearson) للتحقق من مدى دقة الأداة وقدرتها على قياس متغيرات البحث بشكل سليم.

ب) اختبار ثبات الأداة (Reliability Test)

يستخدم اختبار الثبات لتحديد مدى اتساق أداة القياس واستقرار نتائجها عند إعادة تطبيقها على المجموعة نفسها في ظروف مشابهة، فالأداة ذات الثبات العالي هي التي تعطي نتائج متقاربة ومستقرة. وفي هذا البحث، أجري اختبار الثبات على بنود الاختبار التي اجتازت مرحلة "صدق المحتوى" واعتمدت من قبل المحكمين.

وقد استخدم معامل ألفا كرونباخ (Cronbach's Alpha) للتحقق من الثبات، نظرا لملاءمته للاختبارات التي تعتمد درجاتها على تدرج متعدد (غير ثنائي)، وكأداة فعالة لقياس الاتساق الداخلي لنتائج تعلم المتعلمين وفقا لـ اريكونطا (Arikunto). وتتم عملية الحساب وفق الصيغة التالية:

$$r_{11} = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_b^2}{\sigma_t^2}\right)$$

توضيح:

r_{11} : معامل ثبات الأدة

k : عدد فقرات الاختبار

$\sum\sigma^2_\beta$: مجموع تباينات درجات كل فقرة من فقرات الاختبار

σ^2_t : تبيان الدرجة الكلية

بعد استخراج قيمة معامل الثبات، يتم تفسير النتيجة للحكم على جودة الأداة وفقا للتصنيفات التالية حسب معايير اريكونطا (Arikunto) :

الجدول ٣،٥ معايير تفسير معامل الثبات

قيمة معامل الثبات	الفئة
≥٠،٨٠	عالي جدا
٠،٦٠ — ٠،٧٩	عالي
٠،٤٠ — ٠،٥٩	متوسط
٠،٢٠ — ٠،٣٩	منخفض
<٠,٢٠	منخفض جدا

استنادا إلى الجدول (٣,٥) أعلاه، الذي يوضح معايير تفسير معامل الثبات (Reliability Coefficient) وتعد الأداة ذات ثبات عال (موثوقة) إذا وقعت قيمة معامل الثباتِ ضمن الفئات (مرتفع) أو (مرتفع جدا) وفقا للتصنيف المعتمد.[١٤]

[14] Arikunto, *Dasar-Dasar Evaluasi Pendidikan*.