

الفصل الثالث

منهج البحث

أ. تصميم البحث

استُخدم في هذا البحث تصميم شبه تجريبي يتناسب مع طبيعة البحث وأهدافه وفروضه، ويسمح بقياس التغير في مهارة الكلام لدى المجموعة التجريبية، ومقارنة مقدار التحسن بين المجموعة التجريبية والمجموعة الضابطة. وقد أُجري البحث باستخدام فصلين قائمين مسبقاً دون توزيع الطلاب عشوائياً بصورة فردية على المجموعتين. ولذلك، يتيح هذا التصميم دراسة النتائج المرتبطة بالمعالجة التعليمية في حدود خصائص التصميم شبه التجريبي والظروف الفعلية للبحث (Sugiyono, 2022; Creswell & Creswell, 2022).

١. التصميم التجريبي المستخدم

استخدم هذا البحث تصميم المجموعة الضابطة غير المتكافئة (Nonequivalent Control Group Design)، وهو أحد تصميمات البحث شبه التجريبي. ويتضمن هذا التصميم مجموعتين قائمتين

مسبقاً، هما المجموعة التجريبية والمجموعة الضابطة، وتخضع كلتاها للاختبار القبلي والاختبار البعدي.

تكونت المجموعة التجريبية من طلاب الصف الثامن (أ) وعددهم 14 طالباً، في حين تكونت المجموعة الضابطة من طلاب الصف الثامن (ب) وعددهم 12 طالباً. ولم يتم توزيع الطلاب عشوائياً بصورة فردية على المجموعتين، لأن البحث أُجري باستخدام الفصول القائمة فعلياً في المدرسة.

تلقت المجموعة التجريبية عملية التعلم باستخدام وسيلة فيديو المتحركة، في حين تلقت المجموعة الضابطة عملية التعلم دون استخدام هذه الوسيلة. ولا يعني وصف المجموعة الثانية بأنها مجموعة ضابطة أنها لم تتلق عملية تعلم، وإنما تلقت عملية التعلم دون استخدام وسيلة فيديو المتحركة.

ودرست المجموعتان المادة التعليمية نفسها، وهي النص السردي بعنوان الاهتمام بالوقت، وفق أهداف التعلم وإجراءات القياس المحددة نفسها. وبذلك، تمثل الاختلاف الرئيس بين المجموعتين في استخدام وسيلة فيديو المتحركة في المجموعة التجريبية دون المجموعة الضابطة.

ويتيح هذا التصميم قياس التغير داخل كل مجموعة، وكذلك مقارنة مقدار التحسن بين المجموعة التجريبية والمجموعة الضابطة. ومع ذلك، فإن

استخدام فصلين قائمين مسبقاً لا يعني أن المجموعتين متطابقتان تماماً في جميع خصائصهما. ولذلك، استُخدمت بيانات الاختبار القبلي لفحص القدرة الأولية للمجموعتين قبل تفسير النتائج اللاحقة، مع مراعاة طبيعة التصميم شبه التجريبي عند تفسير نتائج البحث.

٢. مخطط تصميم البحث

يمكن تمثيل تصميم المجموعة الضابطة غير المتكافئة المستخدم في هذا البحث في الجدول الآتي:

الجدول ٣.١

مخطط تصميم المجموعة الضابطة غير المتكافئة

المجموعة	الاختبار القبلي	المعالجة	الاختبار البعدي
المجموعة التجريبية	O_1	X	O_2
المجموعة الضابطة	O_3	—	O_4

البيان:

O_1 = الاختبار القبلي للمجموعة التجريبية

O_2 = الاختبار البعدي للمجموعة التجريبية

O_3 = الاختبار القبلي للمجموعة الضابطة

$O_4 =$ الاختبار البعدي للمجموعة الضابطة

$X =$ التعلم باستخدام وسيلة فيديو المتحركة

$=$ — التعلم دون استخدام وسيلة فيديو المتحركة

يوضح المخطط أن المجموعتين خضعتا للاختبار القبلي قبل تنفيذ عملية التعلم، ثم تلقت المجموعة التجريبية التعلم باستخدام وسيلة فيديو المتحركة، في حين تلقت المجموعة الضابطة التعلم دون استخدام هذه الوسيلة. وبعد ذلك، خضعت المجموعتان للاختبار البعدي.

ومن خلال هذا التصميم، أمكن الحصول على بيانات الاختبار القبلي والاختبار البعدي لكل مجموعة، ثم استخدام هذه البيانات لتحليل التغير داخل المجموعتين ومقارنة مقدار التحسن بينهما.

٣. متغيرات البحث والعلاقة بينها

يتضمن هذا البحث متغيرين رئيسيين، هما المتغير المستقل والمتغير التابع،

على النحو الآتي:

أ) المتغير المستقل (X) هو استخدام وسيلة فيديو المتحركة في تعليم اللغة

العربية.

ب) المتغير التابع (Y) هو مهارة الكلام باللغة العربية لدى طلاب الصف الثامن.

ويتمثل استخدام وسيلة فيديو المتحركة في المجموعة التجريبية في تقديم المادة التعليمية من خلال مدخلات سمعية وبصرية مرتبطة بالمحتوى التعليمي، ثم دمجها بأنشطة شفوية منظمة تشمل الاستماع، وملاحظة النموذج والنطق، والتقليد، والسؤال والجواب، وإعادة السرد.

أما مهارة الكلام فتُقاس من خلال الأداء الشفهي للطلاب في خمسة مؤشرات، وهي الطلاقة، والنطق، والمفردات، والقواعد، والفهم. وتتمثل العلاقة التي يختبرها البحث في معرفة ما إذا كان استخدام وسيلة فيديو المتحركة في تعليم اللغة العربية يرتبط بتحسين أكبر في مهارة الكلام لدى طلاب المجموعة التجريبية مقارنة بطلاب المجموعة الضابطة الذين تعلموا دون استخدام هذه الوسيلة.

ويمكن تمثيل العلاقة بين متغيري البحث على النحو الآتي:

استخدام وسيلة فيديو المتحركة $\rightarrow$ (X) مهارة الكلام باللغة العربية (Y)

ونظراً إلى أن البحث يستخدم تصميماً شبه تجريبي، فإن تفسير العلاقة

بين المتغيرين يتم في حدود التصميم المستخدم وأفراد البحث والسياق التعليمي

الذي أُجري فيه البحث، دون افتراض السيطرة الكاملة على جميع العوامل الأخرى التي قد ترتبط بمهارة الكلام.

أما التعريف الإجرائي المفصل لكل من المتغير المستقل والمتغير التابع، فيُعرض في قسم التعريف الإجرائي لمتغيرات البحث.

٤. إجراءات تنفيذ البحث

نُفذت إجراءات البحث من خلال ثلاث مراحل رئيسة، وهي مرحلة الاختبار القبلي، ومرحلة تنفيذ المعالجة التعليمية، ومرحلة الاختبار البعدي، على النحو الآتي:

(أ) مرحلة الاختبار القبلي

في المرحلة الأولى، خضع طلاب المجموعة التجريبية والمجموعة الضابطة للاختبار القبلي لقياس مهارة الكلام باللغة العربية قبل تنفيذ المعالجة التعليمية.

واستُخدمت بيانات الاختبار القبلي للحصول على صورة عن القدرة الأولية للمجموعتين، كما استُخدمت أساساً لحساب مقدار التغير الذي ظهر بعد تنفيذ عملية التعلم.

(ب) مرحلة تنفيذ المعالجة التعليمية

في المرحلة الثانية، درست المجموعة التجريبية والمجموعة الضابطة المادة

التعليمية نفسها، وهي النص السردى بعنوان الاهتمام بالوقت، وفق أهداف

التعلم وإجراءات القياس المحددة نفسها. وتمثل الاختلاف الرئيس بين

المجموعتين في استخدام وسيلة فيديو المتحركة.

تلقت المجموعة التجريبية عملية التعلم باستخدام وسيلة فيديو

المتحركة، ووظفت الوسيلة من خلال أنشطة تعليمية شفوية متدرجة على

النحو الآتي:

عرض الفيديو → الاستماع → ملاحظة النموذج والنطق →

التقليد → السؤال والجواب → إعادة السرد

بدأت عملية التعلم بعرض فيديو المتحركة، ثم استمع الطلاب إلى

المادة اللغوية ولاحظوا النموذج وطريقة النطق. وبعد ذلك، قلد الطلاب

الكلمات والعبارات والجمل الواردة في المادة، ثم شاركوا في السؤال والجواب

الشفهي، وانتهت الأنشطة بإعادة سرد مضمون الاهتمام بالوقت باللغة

العربية.

أما المجموعة الضابطة، فقد درست المادة التعليمية نفسها دون

استخدام وسيلة فيديو المتحركة.

ج) مرحلة الاختبار البعدي

في المرحلة الثالثة، خضع طلاب المجموعة التجريبية والمجموعة

الضابطة للاختبار البعدي لقياس مهارة الكلام بعد انتهاء عملية التعلم.

ثم استُخدمت درجات الاختبار القبلي والاختبار البعدي لتحليل

التغير داخل كل مجموعة، وحساب درجات الكسب، ومقارنة مقدار

التحسن بين المجموعة التجريبية والمجموعة الضابطة وفق تقنيات التحليل

الإحصائي المحددة في هذا البحث.

وبناءً على هذا التصميم، أمكن اختبار الفرض الأول المتعلق بالفرق

بين درجات الاختبار القبلي ودرجات الاختبار البعدي لمهارة الكلام لدى

المجموعة التجريبية، واختبار الفرض الثاني المتعلق بالفرق في درجات الكسب

بين المجموعة التجريبية والمجموعة الضابطة، فضلاً عن تحديد مستوى فاعلية

استخدام وسيلة فيديو المتحركة في ترقية مهارة الكلام لدى طلاب الصف

الثامن.

ب. مجتمع البحث وعينته

يُعدُّ تحديد مجتمع البحث وعينته من الجوانب المهمة في البحث الكمي؛ لأنه يحدد الأفراد الذين تتعلق بهم الدراسة والأفراد الذين تُجمع منهم البيانات بصورة فعلية. ويشمل تحديد المجتمع في هذا البحث المجتمع المستهدف والمجتمع المتاح، كما يشمل تحديد العينة وأسلوب اختيارها وحجمها وتوزيعها على المجموعة التجريبية والمجموعة الضابطة.

١. مجتمع البحث

يقصد بمجتمع البحث جميع الأفراد الذين تتوافر فيهم الخصائص المحددة وفق موضوع البحث وأهدافه. وفي هذا البحث، ينقسم مجتمع البحث إلى المجتمع المستهدف والمجتمع المتاح.

(د) المجتمع المستهدف

يتمثل المجتمع المستهدف في جميع طلاب الصف الثامن بالمدرسة المتوسطة الإسلامية المتكاملة جيلورا الفاتح الذين يمثلون الفئة التي يستهدفها البحث في دراسة فاعلية استخدام وسيلة فيديو المتحركة في ترقية مهارة الكلام باللغة العربية.

هـ) المجتمع المتاح

يتمثل المجتمع المتاح في جميع طلاب الصف الثامن المسجلين

فعلياً في المدرسة خلال فترة إجراء البحث، وهم طلاب الصف الثامن

(أ) والصف الثامن (ب)، وعددهم 26 طالباً.

ويتوزع أفراد المجتمع المتاح على الفصلين كما في الجدول الآتي:

الجدول ٣.٢

توزيع مجتمع البحث

عدد الطلاب	الفصل	الرقم
١٤	الصف الثامن (أ)	١
١٢	الصف الثامن (ب)	٢
٢٦		المجموع

يتضح من الجدول ٣.٢ أن عدد طلاب الصف الثامن (أ) بلغ

14 طالباً، في حين بلغ عدد طلاب الصف الثامن (ب) 12 طالباً،

وبذلك بلغ مجموع أفراد المجتمع المتاح 26 طالباً.

ونظراً إلى أن جميع أفراد المجتمع المستهدف الذين يقع عليهم نطاق

البحث كانوا متاحين فعلياً ويمكن إشراكهم في إجراءات البحث، فقد أمكن

استخدام جميع أفراد المجتمع المتاح عينة للبحث.

٢. عينة البحث وأسلوب اختيارها

العينة هي الأفراد الذين تُجمع منهم البيانات بصورة فعلية وفق أهداف البحث وإجراءاته. واستخدم هذا البحث أسلوب العينة الكلية (Total Sampling)، وهو أسلوب يتم فيه إشراك جميع أفراد المجتمع المتاح في البحث عندما يكون عددهم محدوداً ويمكن الوصول إليهم جميعاً (Sugiyono, 2022).

وبناءً على ذلك، بلغ حجم عينة هذا البحث 26 طالباً، أي إن جميع أفراد المجتمع المتاح أصبحوا أفراداً في العينة. ونظراً إلى أن جميع أفراد المجتمع المتاح قد أُدرجوا في العينة، لم تُستخدم معادلة إحصائية خاصة لتحديد حجم العينة؛ لأن البحث لم يقتصر على اختيار جزء من المجتمع، بل شمل جميع أفرادهِ.

وتم تحديد عينة البحث من خلال المراحل الآتية:

(أ) تحديد جميع طلاب الصف الثامن بالمدرسة المتوسطة الإسلامية المتكاملة

جيلورا الفاتح بوصفهم المجتمع الذي يقع ضمن نطاق البحث.

(ب) تحديد الطلاب المسجلين فعلياً في الصفين الثامن (أ) والثامن (ب)

بوصفهم المجتمع المتاح.

ج) التحقق من عدد أفراد المجتمع المتاح، وهم 26 طالباً، يتكونون من 14

طالباً في الصف الثامن (أ) و 12 طالباً في الصف الثامن (ب).

د) إشراك جميع أفراد المجتمع المتاح في عينة البحث باستخدام أسلوب العينة

الكلية. (Total Sampling)

ويمكن بيان توزيع عينة البحث في الجدول الآتي:

الجدول ٣.٣

توزيع عينة البحث

المجموعة	عدد الطلاب	الفصل	الرقم
المجموعة التجريبية	١٤	الصف الثامن (أ)	١
المجموعة الضابطة	١٢	الصف الثامن (ب)	٢
المجموع	٢٦		

يتضح من الجدول ٣.٣ أن جميع أفراد المجتمع المتاح البالغ عددهم 26

طالباً شاركوا في البحث، منهم 14 طالباً في المجموعة التجريبية و 12 طالباً في

المجموعة الضابطة.

٣. تحديد المجموعة التجريبية والمجموعة الضابطة

استخدم هذا البحث فصلين قائمين مسبقاً في المدرسة، ولم يتم توزيع

الطلاب عشوائياً بصورة فردية على المجموعتين. وحُدد الصف الثامن (أ)، وعدد

طلابه 14 طالباً، بوصفه المجموعة التجريبية، في حين حُدد الصف الثامن (ب)، وعدد طلابه 12 طالباً، بوصفه المجموعة الضابطة.

تلقت المجموعة التجريبية عملية التعلم باستخدام وسيلة فيديو المتحركة، في حين تلقت المجموعة الضابطة عملية التعلم دون استخدام هذه الوسيلة. ودرست المجموعتان المادة التعليمية نفسها، وهي النص السردي بعنوان الاهتمام بالوقت، وفق أهداف التعلم وإجراءات القياس المحددة نفسها. ولا يعني تحديد الفصلين بوصفهما مجموعة تجريبية ومجموعة ضابطة أن المجموعتين متطابقتان تماماً في جميع الخصائص؛ لأنهما فصلان قائمان مسبقاً ولم يتم تشكيلهما من خلال التوزيع العشوائي الفردي. ولذلك، استُخدمت بيانات الاختبار القبلي لفحص القدرة الأولية للمجموعتين قبل تفسير النتائج المتعلقة بالمعالجة التعليمية.

ويتفق هذا التحديد مع تصميم المجموعة الضابطة غير المتكافئة (Nonequivalent Control Group Design) المستخدم في البحث، حيث يتم الاحتفاظ بالتكوين الفعلي للفصلين، مع إجراء الاختبار القبلي والاختبار البعدي ومقارنة مقدار التحسن بين المجموعتين.

وبناءً على ما سبق، شملت عينة البحث جميع أفراد المجتمع المتاح البالغ عددهم 26 طالباً باستخدام أسلوب العينة الكلية، مع استخدام الصف الثامن (أ) مجموعةً تجريبية والصف الثامن (ب) مجموعةً ضابطة، دون إجراء توزيع عشوائي فردي للطلاب.

التعريف الإجرائي لمتغيرات البحث

يهدف التعريف الإجرائي إلى تحديد معنى متغيرات البحث بصورة واضحة وقابلة للملاحظة والقياس، حتى لا يحدث اختلاف في تفسير المتغير المستقل والمتغير التابع أثناء تنفيذ البحث وجمع البيانات وتحليلها. ويتضمن هذا البحث متغيرين رئيسين، هما استخدام وسيلة فيديو المتحركة بوصفه المتغير المستقل (X) ، ومهارة الكلام باللغة العربية بوصفها المتغير التابع (Y).

١. المتغير المستقل: استخدام وسيلة فيديو المتحركة

المتغير المستقل في هذا البحث هو استخدام وسيلة فيديو المتحركة في تعليم اللغة العربية . ويُقصد به إجرائياً استخدام مادة تعليمية سمعية بصرية تعرض المحتوى اللغوي باللغة العربية من خلال الجمع بين الصورة المتحركة والصوت

والنص أو السياق المرتبط بالمادة التعليمية، وتُوظف بصورة منظمة في عملية تعلم المجموعة التجريبية.

واستخدمت وسيلة فيديو المتحركة في هذا البحث في تعليم النص السردي بعنوان الاهتمام بالوقت. ولم يقتصر استخدامها على مشاهدة الفيديو، وإنما دُمجت في سلسلة من الأنشطة الشفهية المتدرجة على النحو الآتي:

عرض الفيديو → الاستماع → ملاحظة النموذج والنطق → التقليد → السؤال والجواب → إعادة السرد
ويتمثل التطبيق الإجرائي للمتغير المستقل في الخطوات الآتية:

١. يعرض المعلم فيديو المتحركة المتعلق بمادة الاهتمام بالوقت أمام طلاب المجموعة التجريبية.

٢. يستمع الطلاب إلى المادة اللغوية الواردة في الفيديو مع ملاحظة السياق البصري المرتبط بالمضمون.

٣. يلاحظ الطلاب النموذج اللغوي وطريقة نطق الكلمات والعبارات والجمل.

٤. يقلد الطلاب النطق والنماذج اللغوية الواردة في المادة بصورة شفهية.

٥. يشارك الطلاب في أنشطة السؤال والجواب الشفهي المتعلقة بمضمون المادة.

٦. يعيد الطلاب سرد مضمون الاهتمام بالوقت باللغة العربية وفق قدرتهم.

وبناءً على ذلك، يُقصد باستخدام وسيلة فيديو المتحركة في هذا البحث
المعالجة التعليمية التي تتلقاها المجموعة التجريبية من خلال توظيف الفيديو
السمعي البصري مع أنشطة الاستماع، وملاحظة النموذج والنطق، والتقليد،
والسؤال والجواب، وإعادة السرد.

أما المجموعة الضابطة، فتدرس المادة التعليمية نفسها، وهي الاهتمام
بالوقت، دون استخدام وسيلة فيديو المتحركة. ومن ثم، يتمثل الاختلاف
الرئيس في المعالجة التعليمية بين المجموعتين في استخدام وسيلة فيديو المتحركة في
المجموعة التجريبية وعدم استخدامها في المجموعة الضابطة.

٢. المتغير التابع: مهارة الكلام باللغة العربية

المتغير التابع في هذا البحث هو مهارة الكلام باللغة العربية. ويُقصد بها
إجرائياً قدرة الطالب على التعبير الشفهي باللغة العربية في المهام المحددة في
الاختبار، والتي تشمل الإجابة عن الأسئلة المتعلقة بالمادة وإعادة سرد مضمون
الاهتمام بالوقت.

وتُقاس مهارة الكلام في هذا البحث من خلال خمسة مؤشرات، هي
الطلاقة، والنطق، والمفردات، والقواعد، والفهم، على النحو الآتي:

أ) الطلاقة

يُقصد بالطلاقة قدرة الطالب على التحدث باللغة العربية بصورة

متصلة ومفهومة، مع تقليل التوقف والتردد الذي يعيق انتقال المعنى.

ب) النطق

يُقصد بالنطق قدرة الطالب على نطق الأصوات والكلمات

والعبارات العربية بصورة واضحة ومفهومة بما يساعد المستمع على فهم

الكلام.

ج) المفردات

يُقصد بالمفردات قدرة الطالب على اختيار المفردات العربية المناسبة

واستخدامها في الإجابة والتعبير الشفهي بما يتناسب مع مضمون الموضوع.

د) القواعد

يُقصد بالقواعد قدرة الطالب على استخدام التراكيب النحوية

والنحوية بصورة مناسبة أثناء إنتاج الكلام، بحيث تكون الجمل المستخدمة

قابلة للفهم ولا تعيق الأخطاء النحوية انتقال المعنى بصورة جوهرية.

هـ) الفهم

يُقصد بالفهم قدرة الطالب على فهم مضمون المادة والأسئلة

الموجهة إليه، ثم تقديم استجابة شفوية تتناسب مع السؤال أو الموضوع الذي

يُطلب منه التعبير عنه.

وُقِّم كل مؤشر من المؤشرات الخمسة باستخدام سلم تقدير من ١ إلى

٤، حيث تشير الدرجة الأعلى إلى مستوى أداء أفضل وفق محكات التقدير

المحددة في روبرك مهارة الكلام.

ويبلغ الحد الأدنى للمجموع الخام 5 درجات، في حين يبلغ الحد

الأقصى 20 درجة. ثم تُحوَّل الدرجة الخام إلى درجة نهائية على مقياس من ٠

إلى ١٠٠ باستخدام الصيغة الآتية:

$$\text{الدرجة النهائية} = (\text{مجموع الدرجات المحصّلة} / ٢٠) \times ١٠٠$$

وتُقاس مهارة الكلام قبل تنفيذ المعالجة التعليمية من خلال الاختبار

القبلي، ثم تُقاس مرة أخرى بعد تنفيذ عملية التعلم من خلال الاختبار البعدي .

وتُستخدم درجات الاختبارين لتحديد مقدار التغير في مهارة الكلام لدى

الطلاب، وحساب درجات الكسب، ومقارنة مقدار التحسن بين المجموعة

التجريبية والمجموعة الضابطة.

واعتمدت الدرجة الأساسية المستخدمة في التحليل الإحصائي على

تقييم المقيّم الأول (Rater 1) ، في حين استُخدم تقييم المقيّم الثاني (Rater

2) للتحقق من الثبات بين المقيّمين، وليس لحساب متوسط مشترك يُستخدم

بوصفه الدرجة الأساسية للتحليل.

ويمكن تلخيص التعريف الإجرائي لمتغيرات البحث في الجدول الآتي:

الجدول ٣.٤

التعريف الإجرائي لمتغيرات البحث

طريقة القياس	المؤشرات/الإجراءات	التعريف الإجرائي	المتغير
تطبيق المعالجة على المجموعة التجريبية	عرض الفيديو، الاستماع، ملاحظة النموذج والنطق، التقليد، السؤال والجواب، إعادة السرد	معالجة تعليمية تستخدم الفيديو السمعي البصري في تعليم مادة الاهتمام بالوقت وتدججه بأنشطة شفهية منظمة	استخدام وسيلة فيديو (X) المتحركة
اختبار شفهي وروبرك تحليلي بدرجات ١- ٤ لكل مؤشر	الطلاقة، النطق، المفردات، القواعد، الفهم	قدرة الطالب على أداء مهام الكلام باللغة العربية في الاختبار الشفهي	مهاراة الكلام باللغة العربية (Y)

وبناءً على التعريف الإجرائي السابق، فإن المتغير المستقل في هذا البحث يتمثل في استخدام وسيلة فيديو المتحركة وفق إجراءات المعالجة المحددة، في حين يتمثل المتغير التابع في مهارة الكلام التي تقاس من خلال الأداء الشفهي للطلاب في مؤشرات الطلاقة والنطق والمفردات والقواعد والفهم. ويتيح هذا التحديد الإجرائي قياس المتغيرات بصورة واضحة ومتسقة مع تصميم البحث وأدوات جمع البيانات وتقنيات التحليل المستخدمة.

ج. تقنيات جمع البيانات

تُعدُّ تقنيات جمع البيانات من الإجراءات الأساسية في البحث؛ لأنها تُستخدم للحصول على البيانات اللازمة للإجابة عن أسئلة البحث واختبار فروضه. ويعتمد اختيار تقنيات جمع البيانات على طبيعة المتغيرات، ونوع البيانات المطلوبة، وتصميم البحث المستخدم. (Sugiyono, 2022)

واستخدم هذا البحث ثلاث تقنيات لجمع البيانات، وهي الاختبار الشفهي، والملاحظة، والتوثيق. ويُعدُّ الاختبار الشفهي التقنية الرئيسة للحصول على البيانات المتعلقة بمهارة الكلام، في حين استُخدمت الملاحظة والتوثيق للحصول على البيانات الداعمة المتعلقة بتنفيذ المعالجة وسياق البحث.

١. الاختبار الشفهي

استخدم الاختبار الشفهي بوصفه التقنية الرئيسة لجمع البيانات المتعلقة بمهارة الكلام باللغة العربية. واختير الاختبار الشفهي لأن المتغير التابع في هذا البحث هو مهارة الكلام التي تتطلب قياس أداء الطلاب بصورة مباشرة من خلال إنتاج اللغة شفهيًا.

وأُجري الاختبار الشفهي على المجموعة التجريبية والمجموعة الضابطة في مرحلتين، هما الاختبار القبلي (Pretest) والاختبار البعدي (Posttest).

أ) الاختبار القبلي

أُجري الاختبار القبلي قبل تنفيذ المعالجة التعليمية على طلاب المجموعة التجريبية والمجموعة الضابطة. ويهدف هذا الاختبار إلى الحصول على بيانات القدرة الأولية للطلاب في مهارة الكلام قبل تنفيذ عملية التعلم.

كما استُخدمت درجات الاختبار القبلي أساساً لتحليل القدرة الأولية للمجموعتين وحساب مقدار التغير بين القياس القبلي والقياس البعدي. ولا يُفترض من إجراء الاختبار القبلي أن المجموعتين متكافئتان تلقائياً، وإنما يُفحص ذلك من خلال التحليل الإحصائي المناسب.

ب) الاختبار البعدي

أُجري الاختبار البعدي بعد انتهاء عملية التعلم على المجموعة التجريبية والمجموعة الضابطة. ويهدف هذا الاختبار إلى الحصول على بيانات مهارة الكلام بعد تنفيذ عملية التعلم، ثم مقارنة النتائج بدرجات الاختبار القبلي لمعرفة مقدار التغير الذي حدث لدى الطلاب. واستخدمت إجراءات القياس نفسها في المجموعتين حتى تكون البيانات الناتجة قابلة للمقارنة. وتكونت مهام الاختبار الشفهي من أنشطة تتعلق بمادة الاهتمام بالوقت، وتشمل الإجابة الشفهية عن الأسئلة وإعادة سرد مضمون المادة باللغة العربية. وقُومَ أداء الطلاب في الاختبار الشفهي من خلال خمسة مؤشرات، وهي:

٢) الطلاقة

٣) النطق

٤) المفردات

٥) القواعد

٦) الفهم

واستخدم في تقييم الأداء روبرك تحليلي بسلم تقدير من 1 إلى ٤ لكل مؤشر. أما تفاصيل شبكة الاختبار ومحكات التقدير وطريقة حساب الدرجات فتُعرض بصورة أكثر تفصيلاً في قسم أدوات البحث. وبذلك، تمثل درجات الاختبار القبلي والاختبار البعدي البيانات الكمية الرئيسة المستخدمة في تحليل التغير في مهارة الكلام ومقارنة مقدار التحسن بين المجموعة التجريبية والمجموعة الضابطة.

٢. الملاحظة

استخدمت الملاحظة لجمع البيانات المتعلقة بمدى تنفيذ استخدام وسيلة فيديو المتحركة في عملية التعلم لدى المجموعة التجريبية. وأُجريت الملاحظة أثناء جلسات المعالجة التعليمية باستخدام ورقة ملاحظة معدة وفق خطوات التعلم المحددة. وتولى الملاحظ متابعة تنفيذ الأنشطة التعليمية للتحقق من أن استخدام وسيلة فيديو المتحركة تم وفق الإجراءات المخطط لها.

وركزت الملاحظة على تنفيذ المراحل الأساسية لاستخدام الوسيلة، ولا سيما عرض الفيديو، واستماع الطلاب إلى المادة، وملاحظة النموذج والنطق، والتقليد، والسؤال والجواب، وإعادة السرد.

وأُجريت الملاحظة خلال أربع جلسات للمعالجة التعليمية في المجموعة التجريبية. واستخدمت نتائجها بوصفها بيانات داعمة لبيان مدى تنفيذ المعالجة التعليمية، ولم تُستخدم درجات الملاحظة بوصفها درجات لقياس المتغير التابع أو بديلاً عن درجات الاختبار الشفهي.

ومن ثم، فإن وظيفة الملاحظة في هذا البحث هي دعم التحقق من أن المعالجة المتمثلة في استخدام وسيلة فيديو المتحركة قد نُفذت وفق الخطوات المحددة، في حين يبقى الاختبار الشفهي المصدر الرئيس للبيانات المتعلقة بمهارة الكلام.

٣. التوثيق

استخدم التوثيق للحصول على البيانات والوثائق التي تدعم تنفيذ البحث وتكمل البيانات التي تم الحصول عليها من الاختبار والملاحظة.

وشملت الوثائق المستخدمة في هذا البحث ما يأتي:

أ) بيانات المدرسة المتعلقة بالسياق العام لمكان البحث.

ب) بيانات عدد طلاب الصف الثامن وتوزيعهم على الصنفين الثامن (أ) والثامن (ب).

ج) قوائم الطلاب المشاركين في المجموعة التجريبية والمجموعة الضابطة.

د) خطط الدروس أو وحدات التعلم المستخدمة في تنفيذ البحث.

هـ) المادة التعليمية المستخدمة، وهي الاهتمام بالوقت.

و) الوثائق المتعلقة بوسيلة فيديو المتحركة المستخدمة في المجموعة التجريبية.

ز) أوراق الاختبار الشفهي وروبك التقدير وورقة ملاحظة تنفيذ التعلم.

ح) وثائق التحقق من صلاحية المادة والوسيلة وأدوات البحث.

ط) سجلات درجات الاختبار القبلي والاختبار البعدي والوثائق الأخرى ذات الصلة بعملية البحث.

ي) الصور التوثيقية المتعلقة بتنفيذ البحث.

واستخدمت بيانات التوثيق بوصفها بيانات داعمة للتحقق من هوية

أفراد البحث، وسياق تنفيذ البحث، والأدوات والمواد المستخدمة، والإجراءات

التي تم تنفيذها. ولم تُستخدم بيانات التوثيق بديلاً عن نتائج الاختبار الشفهي

في قياس مهارة الكلام.

وبناءً على ما سبق، اعتمد جمع البيانات في هذا البحث على الاختبار الشفهي بوصفه المصدر الرئيس للبيانات الكمية لمهارة الكلام، والملاحظة بوصفها وسيلة للتحقق من تنفيذ المعالجة التعليمية، والتوثيق بوصفه مصدراً للبيانات والوثائق الداعمة. ويساعد التكامل بين هذه التقنيات على تنظيم عملية جمع البيانات بما يتفق مع تصميم البحث وأهدافه ومتغيراته.

د. أدوات البحث

تُعَدُّ أدوات البحث وسائلَ منظمةً يستخدمها الباحث للحصول على البيانات اللازمة وفق أهداف البحث ومتغيراته وإجراءاته. ويتطلب إعداد أدوات البحث تحديد المتغيرات ومؤشراتها، واختيار نوع الأداة وسلم التقدير المناسب، وإعداد شبكة الأداة ومحكات التقدير، ثم التحقق من صلاحيتها قبل استخدامها في جمع البيانات. (Sugiyono, 2022)

واستخدم هذا البحث أداتين رئيسيتين، هما الاختبار الشفهي المصحوب بروبرك التقدير التحليلي لقياس مهارة الكلام باللغة العربية، وورقة ملاحظة تنفيذ المعالجة التعليمية للتحقق من مدى تنفيذ استخدام وسيلة فيديو المتحركة في المجموعة التجريبية. واستخدم سلم تقدير رباعي من 1 إلى 4 في روبرك مهارة الكلام وورقة الملاحظة، مع اختلاف وظيفة كل أداة وفق نوع البيانات المراد الحصول عليها.

١. مراحل تطوير أدوات البحث

تم إعداد أدوات البحث وتطويرها بصورة منظمة استناداً إلى أهداف البحث ومتغيراته ومؤشراته. ومرت عملية تطوير الأدوات بالمراحل الآتية:

ب) تحديد التعريف المفاهيمي للمتغيرات

استند إعداد أدوات البحث إلى التعريف المفاهيمي للمتغيرين الذي تم عرضه في الدراسة النظرية. ويُفهم استخدام وسيلة فيديو المتحركة مفاهيمياً بوصفه استخدام وسيلة تعليمية سمعية بصرية تجمع بين الصور المتحركة والصوت والمحتوى اللغوي في تقديم المادة التعليمية بما يتيح للطلاب الاستماع إلى النموذج اللغوي وملاحظته في سياق ذي معنى. أما مهارة الكلام باللغة العربية فتُفهم مفاهيمياً بوصفها قدرة المتعلم على استخدام اللغة العربية شفهاً للتعبير عن المعنى بصورة مفهومة ومناسبة من خلال توظيف العناصر اللغوية اللازمة في الأداء الشفهي.

ج) تحديد التعريف الإجرائي للمتغيرات

حُدّد المتغير المستقل إجرائياً في صورة معالجة تعليمية تستخدم وسيلة فيديو المتحركة في تعليم مادة الاهتمام بالوقت من خلال عرض الفيديو،

والاستماع، وملاحظة النموذج والنطق، والتقليد، والسؤال والجواب، وإعادة السرد.

أما المتغير التابع، وهو مهارة الكلام باللغة العربية، فقد حُدد إجرائياً من خلال أداء الطلاب في الاختبار الشفهي وفق خمسة مؤشرات، هي الطلاقة، والنطق، والمفردات، والقواعد، والفهم، باستخدام سلم تقدير من 1 إلى 4 لكل مؤشر.

(د) إعداد شبكة الاختبار الشفهي

أُعدت شبكة الاختبار الشفهي استناداً إلى التعريف الإجرائي لمهارة الكلام ومؤشراتها الخمسة. وتهدف الشبكة إلى تحقيق الاتساق بين المتغير المقاس، ومؤشراته، ومهام الاختبار، ومجال الأداء المقاس، وسلم التقدير المستخدم.

(هـ) إعداد مهام الاختبار وروبرك التقدير

بعد إعداد شبكة الاختبار، صيغت مهام الاختبار الشفهي في صورة الإجابة الشفهية عن الأسئلة المتعلقة بمضمون مادة الاهتمام بالوقت وإعادة سرد مضمونها باللغة العربية.

كما أُعد روبرك تحليلي (Analytic Rubric) يتضمن محكات واضحة لمستويات الأداء في كل مؤشر من مؤشرات مهارة الكلام، باستخدام درجات من 1 إلى 4.

(و) إعداد ورقة ملاحظة تنفيذ المعالجة التعليمية

أُعدت ورقة الملاحظة استناداً إلى خطوات المعالجة التعليمية المحددة في خطة التعلم باستخدام وسيلة فيديو المتحركة. واشتملت الورقة على بنود متابعة مدى تنفيذ المراحل الرئيسة للمعالجة التعليمية في المجموعة التجريبية.

(ز) التحقق من الاتساق المفاهيمي وتحكيم الأداة

تمت مراجعة أداة الاختبار الشفهي للتأكد من اتساقها مع التعريف المفاهيمي والتعريف الإجرائي لمهارة الكلام ومؤشراتها. ثم عُرضت المادة وأداة الاختبار الشفهي وروبرك التقدير على المختصين من خلال تحكيم المختصين (Expert Judgment) للتحقق من صدق المحتوى وملاءمة الأداة للغرض الذي أُعدت من أجله.

(ح) تعديل الأداة واعتماد صورتها النهائية

أُجريت التعديلات اللازمة على الأداة في ضوء ملاحظات

المختصين، ثم اعتمدت صورتها النهائية واستخدمت في جمع بيانات الاختبار

القبلي والاختبار البعدي.

أما ثبات التقدير بين المقيمين، فقد تم التحقق منه بعد حصول

المقيمين على درجات الأداء الشفهي للطلاب، باستخدام معامل الارتباط

داخل الفئة (ICC)، كما سيأتي بيانه في قسم ثبات التقدير بين المقيمين.

٢. أداة الاختبار الشفهي وشبكته

استخدم الاختبار الشفهي بوصفه الأداة الرئيسة لقياس مهارة الكلام

باللغة العربية لدى طلاب المجموعة التجريبية والمجموعة الضابطة. وأجري

الاختبار في مرحلتين، هما الاختبار القبلي (Pretest) قبل تنفيذ المعالجة

التعليمية والاختبار البعدي (Posttest) بعد انتهاء عملية التعلم.

ارتبطت مهام الاختبار الشفهي بالنص السردي بعنوان الاهتمام

بالوقت، وتضمنت الإجابة الشفهية عن الأسئلة المتعلقة بمضمون المادة وإعادة

سرد مضمونها باللغة العربية.

واستخدمت إجراءات القياس والمؤشرات وسلم التقدير ومحكات

التصحيح نفسها في الاختبار القبلي والاختبار البعدي ولدى المجموعة التجريبية

والمجموعة الضابطة، بما يجعل البيانات الناتجة قابلة للمقارنة.

وتُعرض شبكة الاختبار الشفهي في الجدول الآتي:

الجدول ٣.٥

شبكة الاختبار الشفهي لمهارة الكلام

الرقم	مؤشر مهارة الكلام	مجال الأداء المقاس	شكل المهمة	سلم التقدير
١	الطلاقة	استمرارية الكلام وتقليل التوقف والتردد الذي يعيق انتقال المعنى	الإجابة الشفهية وإعادة السرد	١-٤
٢	النطق	وضوح نطق الأصوات والكلمات والعبارات العربية	الإجابة الشفهية وإعادة السرد	١-٤
٣	المفردات	ملاءمة المفردات وكفايتها للتعبير عن المضمون	الإجابة الشفهية وإعادة السرد	١-٤
٤	القواعد	ملاءمة التراكيب والقواعد المستخدمة في الكلام	الإجابة الشفهية وإعادة السرد	١-٤
٥	الفهم	فهم السؤال والمضمون وتقديم استجابة شفهية مناسبة	الإجابة الشفهية وإعادة السرد	٤-١

وبناءً على الجدول السابق، يبلغ الحد الأدنى للمجموع الخام 5

درجات، في حين يبلغ الحد الأقصى 20 درجة.

٣. سلم التقدير ومحكات تقويم مهارة الكلام

استخدم في تقويم أداء الطلاب روبرك تحليلي (Analytic

Rubric) بسلم تقدير من 1 إلى 4 لكل مؤشر. وتشير الدرجة 4 إلى أعلى

مستوى من الأداء، في حين تشير الدرجة 1 إلى أدنى مستوى من الأداء وفق

المحكات المحددة.

ويتيح الروبرك التحليلي تقويم كل مؤشر من مؤشرات مهارة الكلام

بصورة مستقلة وباستخدام المحكات نفسها لجميع الطلاب.

وتتمثل محكات التقدير فيما يأتي:

الجدول ٣.٦

محكات تقدير مهارة الكلام باللغة العربية

الدرجة ١: ضعيف	الدرجة ٢: مقبول	الدرجة ٣: جيد	الدرجة ٤: ممتاز	المؤشر
يتحدث بصورة متقطعة جداً ويصعب متابعة كلامه	يكثر التوقف والتردد بما يؤثر في استمرارية الكلام	يتحدث بصورة جيدة مع بعض التوقف أو التردد المحدود، ويبقى المعنى واضحاً	يتحدث بصورة متصلة وواضحة، مع توقف أو تردد قليل جداً لا يؤثر في المعنى	الطلاقة

النطق	ينطق الأصوات والكلمات والعبارات بوضوح، والأخطاء قليلة جداً ولا تؤثر في الفهم	توجد بعض أخطاء النطق، ولكن الكلام يبقى مفهوماً	توجد أخطاء متكررة في النطق تؤثر أحياناً في الفهم	تكثر أخطاء النطق إلى درجة تعيق فهم الكلام
المفردات	يستخدم مفردات مناسبة وكافية للتعبير عن المضمون	يستخدم مفردات مناسبة عموماً مع بعض القصور أو التكرار	مفرداته محدودة ويكثر التكرار أو الاستخدام غير المناسب	مفرداته محدودة جداً ويجد صعوبة في تكوين الاستجابة
القواعد	يستخدم تراكيب لغوية مناسبة، والأخطاء القليلة لا تعيق انتقال المعنى	توجد بعض الأخطاء في التراكيب، ولكن المعنى يبقى واضحاً	تكثر الأخطاء في التراكيب وتؤثر جزئياً في وضوح المعنى	تكثر الأخطاء في التراكيب إلى درجة تعيق انتقال المعنى
الفهم	يفهم السؤال والمضمون جيداً ويقدم استجابة مناسبة ويعيد السرد بصورة مترابطة	يفهم المضمون الرئيس ويقدم استجابة مناسبة عموماً مع نقص بسيط	يفهم جزءاً من السؤال أو المضمون وتكون استجابته ناقصة أو تحتاج إلى توجيه	يظهر فهماً محدوداً ولا يستطيع تقديم استجابة مناسبة

وتُحسب الدرجة الخام للطالب من مجموع الدرجات التي يحصل عليها
في المؤشرات الخمسة. ثم تُحوَّل الدرجة الخام إلى درجة نهائية من ١٠٠
باستخدام الصيغة الآتية:

$$\text{الدرجة النهائية} = (\text{مجموع الدرجات المحصّلة} / ٢٠) \times ١٠٠$$

وبناءً على هذه الصيغة، تتراوح الدرجات النهائية الممكنة من 25 إلى
١٠٠ درجة.

٤. أداة ملاحظة تنفيذ المعالجة التعليمية

استخدمت ورقة الملاحظة للتحقق من مدى تنفيذ المعالجة التعليمية
باستخدام وسيلة فيديو المتحركة في المجموعة التجريبية. ولم تُستخدم ورقة
الملاحظة لقياس مهارة الكلام بوصفها المتغير التابع، وإنما استخدمت للحصول
على بيانات داعمة تتعلق بمدى تنفيذ المعالجة التعليمية.
وأعدت ورقة الملاحظة استناداً إلى خطوات المعالجة التعليمية المحددة في
خطة التعلم. وتكونت من 10 بنود، واستخدم لكل بند سلم تقدير من 1 إلى
٤، ولذلك بلغ الحد الأقصى للدرجة في كل جلسة 40 درجة.

واستخدمت ورقة الملاحظة خلال أربع جلسات للمعالجة التعليمية في
المجموعة التجريبية.

وتتمثل مواصفات أداة الملاحظة فيما يأتي:

الجدول ٣.٧

مواصفات أداة ملاحظة تنفيذ المعالجة التعليمية

البيان	العنصر
التحقق من مدى تنفيذ المعالجة التعليمية باستخدام وسيلة فيديو المتحركة	الغرض من الأداة
المجموعة التجريبية	المجموعة التي تمت ملاحظتها
بنود ١٠	عدد البنود
٤-١	سلم التقدير
درجة في كل جلسة ٤٠	الحد الأقصى للدرجة
أربع جلسات	عدد جلسات الملاحظة
بيانات داعمة لبيان مدى تنفيذ المعالجة، وليست درجة لمهارة الكلام	وظيفة البيانات

وركزت الملاحظة على المراحل الرئيسة لتنفيذ المعالجة التعليمية، ومنها عرض الفيديو، والاستماع إلى المادة، وملاحظة النموذج والنطق، والتقليد، والسؤال والجواب، وإعادة السرد.

وتُحسب نسبة تنفيذ المعالجة باستخدام الصيغة الآتية:

$$\text{نسبة التنفيذ} = (\text{الدرجة المحصّلة} / ٤٠) \times ١٠٠\%$$

واستخدمت نتائج الملاحظة لتوضيح مدى تنفيذ المعالجة التعليمية، في

حين بقي الاختبار الشفهي الأداة الرئيسة للحصول على بيانات مهارة الكلام.

٥. صدق أداة الاختبار الشفهي وروبرك التقدير

تم التحقق من صدق أداة الاختبار الشفهي وروبرك التقدير باستخدام

صدق المحتوى (Content Validity) من خلال تحكيم المختصين
(Expert Judgment).

وشملت عملية التحقق من الصدق مرحلتين، هما التحقق من الاتساق

المفاهيمي للأداة والتحكيم المتخصص لمحتواها.

ط) التحقق من الاتساق المفاهيمي للأداة

تم التحقق من الاتساق المفاهيمي من خلال مطابقة أداة الاختبار

مع التعريف المفاهيمي والتعريف الإجرائي لمهارة الكلام.

وبناءً على ذلك، تم التأكد من أن المؤشرات الخمسة، وهي

الطلاقة، والنطق، والمفردات، والقواعد، والفهم، تتفق مع الجوانب المحددة

لقياس مهارة الكلام في هذا البحث، وأن مهام الإجابة الشفهية وإعادة

السرد تتناسب مع طبيعة الأداء الشفهي المراد قياسه.

ويُعدُّ هذا الإجراء مرحلةً للتأكد من الاتساق بين المفهوم النظري

والمغير الإجرائي وبناء الأداة، وليس اختباراً إحصائياً مستقلاً للصدق.

ي) التحكيم المتخصص وصدق المحتوى

عُرِضَت المادة وأداة الاختبار الشفهي وروبك التقدير على

المختصين للمراجعة والتحكيم قبل استخدام الأداة في جمع البيانات.

وركزت عملية التحكيم على الجوانب الآتية:

(١) ملاءمة مادة الاهتمام بالوقت لمستوى طلاب الصف الثامن وأهداف

التعلم.

(٢) ملاءمة المهام الشفهية لقياس مهارة الكلام.

(٣) ملاءمة مؤشرات الطلاقة، والنطق، والمفردات، والقواعد، والفهم للمتغير

المراد قياسه.

(٤) وضوح محكات التقدير ومستويات الدرجات من ١ إلى ٤.

(٥) ملاءمة المحتوى واللغة المستخدمة في الأداة.

(٦) وضوح تعليمات الاختبار وإجراءات التقويم.

وحُسِبَت نسبة صلاحية الأداة استناداً إلى درجات استمارة

التحكيم باستخدام الص ز ظوظعم كظ ([رظنخومزكم»ظي ٢٣ ثقة

الآتية:

نسبة الصلاحية = (مجموع الدرجات المحصّلة / مجموع الدرجات القصوى)

$\times 100\%$

واستخدمت النسبة الناتجة، إلى جانب الملاحظات النوعية للمختصين، أساساً لمراجعة الأداة وإجراء التعديلات اللازمة قبل استخدامها في جمع البيانات.

أما النتائج العددية الفعلية لعملية التحكيم فتعرض في باب نتائج البحث، حتى يبقى هذا القسم مخصصاً لشرح إجراءات التحقق من صدق الأداة دون خلطها بنتائج البحث.

٦. ثبات التقدير بين المقيمين

نظراً إلى أن الاختبار الشفهي يعتمد على تقدير أداء الطلاب بواسطة المقيمين، فقد تم التحقق من ثبات عملية التقدير باستخدام الثبات بين المقيمين (Inter-Rater Reliability).

وقام بتقويم أداء الطلاب مقيمان بصورة مستقلة باستخدام روبرك التقدير والمؤشرات والمحكات نفسها. وكان الباحث هو المقيم الأول (Rater

1)، في حين كان معلم اللغة العربية هو المقيم الثاني. (Rater 2)

وقام كلا المقيمين بتقويم أداء الطلاب في الاختبار القبلي والاختبار البعدي وفق المؤشرات الخمسة، وهي الطلاقة، والنطق، والمفردات، والقواعد، والفهم.

وبعد اكتمال عملية التقدير، حُسب الثبات بين المقيمين باستخدام

معامل الارتباط داخل الفئة (Intraclass Correlation

Coefficient/ICC) بواسطة برنامج IBM SPSS Statistics

الإصدار ٢٣.

واستخدم نموذج Two-Way Mixed Effects مع تعريف

Absolute Agreement؛ لأن المقيمين المستخدمين في هذا البحث

مقيمان محددان، ولأن الغرض من التحليل هو التحقق من مدى الاتفاق بينهما

في إعطاء الدرجات للأداء نفسه.

ونظراً إلى أن الدرجة التي اعتمدت بوصفها البيانات الأساسية في جميع

التحليلات الإحصائية هي درجة المقيم الأول (Rater 1)، فإن قيمة الثبات

الرئيسية التي يعتمد عليها في تفسير ثبات التقدير هي Single Measures

ICC.

أما درجات Rater 2 فلم تُدمج مع درجات Rater 1، ولم

يُحسب متوسط درجتي المقيمين بوصفه الدرجة الأساسية للتحليل، وإنما

استخدمت درجات Rater 2 حصراً للتحقق من الثبات بين المقيمين.

وعليه، تعتمد التحليلات الإحصائية الرئيسة المتعلقة بالاختبار القبلي والاختبار البعدي ودرجات الكسب (Gain Score) و N-Gain وحجم الأثر (Effect Size) على درجات Rater 1، في حين تستخدم درجات Rater 2 للتحقق من موثوقية عملية تقدير الأداء الشفهي.

أما نتائج حساب معامل الثبات بين المقيمين فتُعرض في باب نتائج البحث، بينما يقتصر هذا القسم على بيان إجراءات حساب الثبات والنموذج الإحصائي المستخدم.

وبناءً على ما سبق، تم تطوير أدوات البحث من خلال تحديد التعريف المفاهيمي والإجرائي للمتغيرات، وإعداد شبكة الاختبار ومهامه وروبرك تقديره، وإعداد ورقة الملاحظة، والتحقق من الاتساق المفاهيمي وصدق المحتوى من خلال تحكيم المختصين، وإجراء التعديلات اللازمة قبل استخدام الأداة، ثم التحقق من ثبات تقدير نتائج الاختبار بين المقيمين باستخدام معامل ICC بعد إجراء عملية التقويم. وبذلك تتسق أدوات البحث مع أهداف البحث

ومتغيراته وتقنيات جمع البيانات وإجراءات التحليل الإحصائي المستخدمة.

هـ. تقنيات تحليل البيانات

تهدف تقنيات تحليل البيانات إلى تنظيم البيانات التي تم جمعها ومعالجتها وتحليلها إحصائياً من أجل الإجابة عن أسئلة البحث واختبار فروضه. ونظراً إلى أن هذا البحث يستخدم المنهج الكمي بتصميم شبه تجريبي يضم مجموعة تجريبية ومجموعة ضابطة، فقد تم تحليل البيانات باستخدام الإحصاء الوصفي والإحصاء الاستدلالي واختبارات افتراضات التحليل الإحصائي بما يتناسب مع طبيعة البيانات وأهداف البحث (Sugiyono, 2022; Creswell & Creswell, 2022).

وأُجريت التحليلات الإحصائية باستخدام برنامج IBM SPSS Statistics الإصدار ٢٣. واعتمدت جميع التحليلات الإحصائية الرئيسة على درجات المقيم الأول (Rater 1) بوصفها البيانات الأساسية للبحث، في حين استخدمت درجات المقيم الثاني (Rater 2) للتحقق من ثبات التقدير بين المقيمين فقط.

١. التحليل الإحصائي الوصفي

استخدم التحليل الإحصائي الوصفي لوصف بيانات مهارة الكلام لدى طلاب المجموعة التجريبية والمجموعة الضابطة في الاختبار القبلي والاختبار البعدي.

وشمل التحليل الوصفي حساب عدد البيانات (N) ، والمتوسط الحسابي (Mean) ، والانحراف المعياري (Standard Deviation) ، والدرجة الدنيا (Minimum) ، والدرجة العليا (Maximum). وعُرضت نتائج التحليل الوصفي في جداول إحصائية بما يتناسب مع طبيعة البيانات وحجم العينة.

كما حُسبت درجة الكسب (Gain Score) لكل طالب لمعرفة مقدار التغير بين الاختبار القبلي والاختبار البعدي باستخدام الصيغة الآتية:

$$\text{Gain Score} = \text{درجة الاختبار البعدي} - \text{درجة الاختبار القبلي}$$

وتمثل درجة الكسب الفرق بين القياسين البعدي والقبلي لكل طالب، واستخدمت في وصف مقدار التحسن وفي المقارنة بين المجموعة التجريبية والمجموعة الضابطة.

٢. اختبار افتراضات التحليل الإحصائي

قبل إجراء الاختبارات الإحصائية الاستدلالية، تم فحص خصائص البيانات من خلال اختبار التوزيع الطبيعي واختبار تجانس التباين لتحديد الاختبار الإحصائي المناسب للتحليلات التي تتطلب تلك الافتراضات.

ك) اختبار التوزيع الطبيعي

استخدم اختبار شابيرو-ويلك (Shapiro-Wilk

Test) للتحقق من التوزيع الطبيعي للبيانات؛ نظراً إلى ملاءمته لفحص طبيعية التوزيع في العينات الصغيرة المستخدمة في هذا البحث. وشمل اختبار التوزيع الطبيعي بيانات الاختبار القبلي والاختبار البعدي ودرجات الكسب في المجموعة التجريبية والمجموعة الضابطة. واعتمد مستوى الدلالة $\alpha = 0.05$ وفق المعايير الآتية:

١) إذا كانت قيمة $\text{Sig.} > 0.05$ ، فلا يوجد دليل إحصائي كافٍ لرفض افتراض التوزيع الطبيعي، ولذلك يمكن التعامل مع البيانات بوصفها مستوفيةً لافتراض التوزيع الطبيعي.

٢) إذا كانت قيمة $\text{Sig.} \leq 0.05$ ، فيُرفض افتراض التوزيع الطبيعي، وتُعد البيانات غير مستوفية لهذا الافتراض.

ل) اختبار تجانس التباين

استخدم اختبار ليفين (Levene's Test) للتحقق من

تجانس التباين بين المجموعة التجريبية والمجموعة الضابطة.

واعتمد مستوى الدلالة $\alpha = 0.05$ وفق المعايير الآتية:

(١) إذا كانت قيمة $\text{Sig.} > 0.05$ ، فلا يوجد دليل إحصائي كافٍ

لرفض افتراض تجانس التباين، ولذلك يمكن التعامل مع تباين المجموعتين بوصفه مستوفياً لافتراض التجانس.

(٢) إذا كانت قيمة $\text{Sig.} \leq 0.05$ ، فيُرفض افتراض تجانس التباين، ويُعد

تباين المجموعتين غير مستوفٍ لافتراض التجانس.

واستخدمت نتائج اختبار التوزيع الطبيعي واختبار تجانس التباين

أساساً في تحديد الاختبار الإحصائي المناسب للمقارنة بين المجموعتين.

٣. فحص الفرق في القدرة الأولية بين المجموعة التجريبية والمجموعة الضابطة

قبل تفسير النتائج المتعلقة بالمعالجة التعليمية، تم فحص الفرق في القدرة

الأولية بين المجموعة التجريبية والمجموعة الضابطة باستخدام درجات الاختبار القبلي.

ونظراً إلى أن بيانات الاختبار القبلي في المجموعتين استوفت افتراض

التوزيع الطبيعي، كما استُوفي افتراض تجانس التباين بينهما، استخدم اختبار t

لعينتين مستقلتين (Independent Samples t-Test)

لفحص وجود فرق ذي دلالة إحصائية بين متوسط درجات الاختبار القبلي للمجموعتين.

واعتمدت معايير اتخاذ القرار الآتية:

(١) إذا كانت قيمة $\text{Sig. (2-tailed)} > 0.05$ ، فلا يوجد فرق ذو دلالة

إحصائية بين درجات الاختبار القبلي للمجموعتين.

(٢) إذا كانت قيمة $\text{Sig. (2-tailed)} \leq 0.05$ ، فيوجد فرق ذو دلالة

إحصائية بين درجات الاختبار القبلي للمجموعتين.

ولا يهدف هذا الاختبار إلى إثبات التطابق أو التكافؤ التام بين

المجموعتين في جميع الخصائص، وإنما إلى فحص ما إذا كان هناك فرق ذو دلالة

إحصائية بين درجتهما القبلية قبل تفسير النتائج المتعلقة بالمعالجة التعليمية.

٤. اختبار الفرض الأول: الفرق بين الاختبار القبلي والاختبار البعدي

للمجموعة التجريبية

يتعلق الفرض الأول بفحص التغير في مهارة الكلام لدى طلاب

المجموعة التجريبية بين الاختبار القبلي والاختبار البعدي بعد استخدام وسيلة

فيديو المتحركة.

ولأن المقارنة في هذا الجزء تتعلق بقياسين مترابطين للطلاب أنفسهم،
ولأن توزيع درجات الفروق لم يستوف افتراض التوزيع الطبيعي ولم يُفترض تماثل
توزيع الفروق، استخدم اختبار الإشارة (Sign Test) لفحص اتجاه التغير
بين درجات الاختبار القبلي والاختبار البعدي.

ويعتمد اختبار الإشارة على مقارنة عدد الحالات التي ارتفعت درجاتها
بعد المعالجة بعدد الحالات التي انخفضت درجاتها، دون اشتراط التوزيع الطبيعي
أو تماثل توزيع الفروق.

وصيغ الفرضان الإحصائيان على النحو الآتي:

H₀: لا يوجد فرق ذو دلالة إحصائية بين درجات الاختبار القبلي ودرجات

الاختبار البعدي لمهارة الكلام باللغة العربية لدى طلاب المجموعة التجريبية بعد

استخدام وسيلة فيديو المتحركة.

H_a: يوجد فرق ذو دلالة إحصائية بين درجات الاختبار القبلي ودرجات

الاختبار البعدي لمهارة الكلام باللغة العربية لدى طلاب المجموعة التجريبية بعد

استخدام وسيلة فيديو المتحركة.

واعتمد مستوى الدلالة $\alpha = 0.05$ ، وتُحدد نتيجة الاختبار وفق

المعايير الآتية:

(١) إذا كانت قيمة $\text{Exact Sig. (2-tailed)} \leq 0.05$ ، فيُرفض

H_{01} ويُدعم H_{a1} .

(٢) إذا كانت قيمة $\text{Exact Sig. (2-tailed)} > 0.05$ ، فلا يُرفض

H_{01} .

وعند تفسير نتائج الاختبار، تؤخذ في الاعتبار أعداد الفروق الموجبة، والفروق السالبة، والحالات المتعادلة (Ties) إلى جانب قيمة الدلالة الإحصائية.

٥. تحليل الفرق بين الاختبار القبلي والاختبار البعدي للمجموعة الضابطة

إضافة إلى اختبار الفرض الأول، تم تحليل التغير في درجات المجموعة الضابطة بين الاختبار القبلي والاختبار البعدي للحصول على معلومات داعمة حول التغير الذي حدث لدى الطلاب الذين تعلموا دون استخدام وسيلة فيديو المتحركة.

واستخدم اختبار الإشارة (Sign Test) أيضاً لتحليل التغير داخل

المجموعة الضابطة؛ لأن البيانات تمثل قياسين مترابطين للطلاب أنفسهم، ولا

يتطلب هذا الاختبار افتراض التوزيع الطبيعي أو تماثل توزيع الفروق.

واعتمد مستوى الدلالة $\alpha = 0.05$ وفق المعايير الآتية:

أ) إذا كانت قيمة $\text{Exact Sig. (2-tailed)} \leq 0.05$ ، فيوجد تغير

ذو دلالة إحصائية بين درجات الاختبار القبلي والاختبار البعدي للمجموعة الضابطة.

ب) إذا كانت قيمة $\text{Exact Sig. (2-tailed)} > 0.05$ ، فلا يوجد

دليل إحصائي كافٍ على وجود تغير ذي دلالة إحصائية بين القياسين. ولا يمثل هذا التحليل فرضاً رئيساً مستقلاً في البحث، وإنما استخدم بوصفه تحليلاً داعماً لتوضيح اتجاه التغير الذي حدث في المجموعة الضابطة، بما يساعد على تفسير المقارنة بين مقدار التحسن في المجموعتين.

٦. اختبار الفرض الثاني: الفرق في درجات الكسب بين المجموعتين

يتعلق الفرض الثاني بمقارنة مقدار التحسن في مهارة الكلام بين المجموعة التجريبية والمجموعة الضابطة.

واستخدمت درجات الكسب (Gain Score)، وهي الفرق بين

الدرجة البعدية والدرجة القبليّة لكل طالب.

ونظراً إلى أن درجات الكسب في المجموعتين لم تستوف افتراض التوزيع

الطبيعي، استخدم اختبار مان-ويتني U (Mann-Whitney U

(Test) لمقارنة توزيع درجات الكسب بين المجموعة التجريبية والمجموعة

الضابطة.

وصيغ الفرضان الإحصائيان على النحو الآتي:

لمهارة (Gain Score) لا يوجد فرق ذو دلالة إحصائية في درجات الكسب: H_{02}

الكلام باللغة العربية بين طلاب المجموعة التجريبية الذين تعلموا باستخدام وسيلة فيديو

المتحركة وطلاب المجموعة الضابطة الذين تعلموا دون استخدام هذه الوسيلة

لمهارة (Gain Score) يوجد فرق ذو دلالة إحصائية في درجات الكسب: H_{a2}

الكلام باللغة العربية بين طلاب المجموعة التجريبية الذين تعلموا باستخدام وسيلة فيديو

المتحركة وطلاب المجموعة الضابطة الذين تعلموا دون استخدام هذه الوسيلة

واعتمدت معايير اتخاذ القرار الآتية:

(أ) إذا كانت قيمة $Asymp. Sig. (2-tailed) \leq 0.05$ ، فيرفض H_{02} ويُدعم

H_{a2} .

(ب) إذا كانت قيمة $Asymp. Sig. (2-tailed) > 0.05$ ، فلا يُرفض H_{02} .

ويمثل هذا الاختبار التحليل الرئيس للمقارنة الإحصائية بين مقدار التحسن لدى

المجموعة التجريبية والمجموعة الضابطة.

٧. تحليل الكسب المعياري (N-Gain)

استخدم تحليل الكسب المعياري-N (Normalized Gain)

(Gain) بوصفه تحليلاً مكماً لتوضيح مقدار التحسن النسبي في مهارة الكلام

مع مراعاة الدرجة القصوى الممكنة.

وفي هذا البحث، حُسبت قيمة N-Gain لكل طالب على حدة

باستخدام الصيغة الآتية:

$$N-Gain = \frac{(100 - \text{درجة الاختبار القبلي})}{(\text{درجة الاختبار البعدي} - \text{درجة الاختبار القبلي})}$$

(الاختبار القبلي)

وبعد حساب قيمة N-Gain الفردية لكل طالب، حُسب متوسط

قيم N-Gain الفردية لأفراد كل مجموعة لتقديم وصف عام لمستوى التحسن

النسبي في المجموعة التجريبية والمجموعة الضابطة.

واسترشد هذا البحث بحدود تصنيف الكسب المعياري التي قدمها

(Hake 1998)، على النحو الآتي:

قيمة N-Gain	مستوى التحسن
$N-Gain < 0.30$	منخفض
$0.30 \leq N-Gain < 0.70$	متوسط
$N-Gain \geq 0.70$	مرتفع

واستخدم تحليل N-Gain بوصفه تحليلاً مكماً لتفسير مستوى التحسن النسبي، ولم يستخدم بديلاً عن اختبار الفرض الإحصائي المتعلق بالفرق في درجات الكسب بين المجموعتين. (Hake, 1998)

٨. تحليل حجم الأثر (Effect Size)

إلى جانب اختبار الدلالة الإحصائية، حُسب حجم الأثر (Effect Size) لتوضيح حجم الفرق المرتبط بالمقارنة بين درجات الكسب للمجموعة التجريبية والمجموعة الضابطة.

ونظراً إلى أن المقارنة الرئيسة لدرجات الكسب بين المجموعتين تمت باستخدام اختبار Mann-Whitney U، حُسب حجم الأثر r اعتماداً على القيمة المعيارية Z الناتجة عن اختبار Mann-Whitney باستخدام الصيغة الآتية:

$$r = |Z| / \sqrt{N}$$

حيث:

r = حجم الأثر

$|Z|$ = Mann-Whitney U الناتج عن اختبار Z القيمة المطلقة لإحصاء

N = العدد الكلي لأفراد المجموعتين المشمولين في المقارنة

واسترشد تفسير حجم الأثر بالحدود المستخدمة لتفسير قيمة r في

الاختبارات غير البارامترية، على النحو الآتي: (Peres, 2026)

تفسير حجم الأثر	قيمة $ r $
دون حد الأثر الصغير	$ r < 0.10$
صغير	$0.10 \leq r < 0.30$
متوسط	$0.30 \leq r < 0.50$
كبير	$ r \geq 0.50$

ولا يُفسر حجم الأثر في هذا البحث بوصفه دليلاً على علاقة سببية مطلقة؛ لأن البحث يستخدم تصميماً شبه تجريبي دون توزيع عشوائي فردي للطلاب، وإنما يستخدم لتوضيح حجم الفرق في مقدار التحسن بين المجموعتين في حدود عينة البحث وتصميمه وظروف تنفيذه.

٩. معايير اختبار الفروض وتفسير النتائج

اعتمد مستوى الدلالة الإحصائية $\alpha = 0.05$ في الاختبارات

الإحصائية المستخدمة في هذا البحث.

وبصورة عامة، يستند اتخاذ القرار إلى مقارنة قيمة الدلالة الإحصائية

بمستوى $\alpha = 0.05$ ؛ فإذا كانت قيمة الدلالة مساوية لـ 0.05 أو أقل منها،

عُدَّت النتيجة ذات دلالة إحصائية، أما إذا كانت أكبر من ٠.٠٠٥، فلا يوجد

دليل إحصائي كافٍ لرفض الفرض الصفري المرتبط بالاختبار.

ولا يعتمد تفسير نتائج استخدام وسيلة فيديو المتحركة على الدلالة

الإحصائية وحدها، وإنما يتم بصورة متكاملة من خلال التغير داخل المجموعة

التجريبية، والتحليل الداعم للتغير داخل المجموعة الضابطة، والمقارنة بين درجات

الكسب للمجموعتين، ومتوسط N-Gain، وحجم الأثر.

وبناءً على ما سبق، تم تنظيم تحليل البيانات في هذا البحث بصورة

متدرجة تبدأ بالتحليل الإحصائي الوصفي، ثم اختبار افتراضات التحليل

الإحصائي، وفحص الفرق في القدرة الأولية بين المجموعة التجريبية والمجموعة

الضابطة، وتحليل التغير داخل كل مجموعة باستخدام اختبار الإشارة، ومقارنة

درجات الكسب بين المجموعتين باستخدام اختبار Mann-Whitney

U، ثم تحليل N-Gain وحجم الأثر. ويتفق هذا التسلسل مع طبيعة

التصميم شبه التجريبي وأسئلة البحث وفروضة، ويساعد على تفسير النتائج في

حدود البيانات وإجراءات البحث المستخدمة.