

تحليل الخطأ المفاهيمي في التفريق بين القدرة التمييزية للفقرات والاتساق الخارجي وأثره على دقة التقييم النفسي

حسنين عبد الأمير طعمه، أ.م.د. أكرم عبد الحسين جواد

*الايمل: akram.ghyad@qu.edu.iq

تاريخ نشر: 2025/6/25

تاريخ استلام: 2025/02/18

الملخص

أن الملاحظة البحثية تشير إلى وجود خلط مفاهيمي شائع بين مفهومي القدرة التمييزية للفقرات والاتساق الخارجي، وهو خلط يؤثر سلباً على جودة أدوات القياس ويهدد صلاحية الاستنتاجات المستخلصة من البيانات. وانطلاقاً من هذه الإشكالية، تهدف الدراسة الحالية إلى تحليل هذا الخلط وتقييم انعكاساته على دقة التقييمات التربوية والنفسية، من خلال استقصاء الفروق المفاهيمية والإحصائية بين المفهومين، ورصد مدى انتشاره بين الباحثين، بالإضافة إلى اقتراح حلول تدريبية ومنهجية تساهم في الحد من هذه الإشكالات. اعتمدت الدراسة المنهج الوصفي التحليلي واستخدمت مجموعة من الأدوات تضمنت الملاحظة لرصد مظاهر الخلط، والمقابلات مع خبراء في القياس والإحصاء، واختبارات معرفية معيارية لقياس الفهم الإحصائي لدى عينة قوامها 150 طالب دراسات عليا في جامعة القادسية، خضعوا لاختبار معرفي مكون من 50 فقرة. كشفت النتائج أن نسبة كبيرة من الطلبة بلغت 71.6 بالمئة لم يتمكنوا من التمييز بين المفهومين، كما أظهرت نتائج التحليل الإحصائي أن العلاقة بين القدرة التمييزية والاتساق الخارجي ضعيفة وفق معامل بيرسون، ما يشير إلى أن المفهومين مستقلان من الناحية الإحصائية. وأظهر تحليل العوامل الاستكشافي وجود عاملين منفصلين، حيث فسّر العامل المرتبط بالقدرة التمييزية في حين فسّر العامل المتعلق بالاتساق الخارجي مما يعزز فرضية الاختلاف الجوهرية بينهما. كما بيّن تحليل الانحدار المتعدد أن الفهم الدقيق للقدرة التمييزية يساهم بدرجة أكبر في تحسين دقة أدوات القياس مقارنة بالاتساق الخارجي، إذ بلغت قيمة بيتا 0.38 مقابل 0.25، وكلاهما دال إحصائياً. بناءً على هذه المعطيات، توصي الدراسة بتطوير برامج تدريبية متخصصة لتعزيز الكفاءة الإحصائية لدى الباحثين، وتضمين مفاهيم القياس في المناهج الجامعية بأسلوب منهجي واضح، واستخدام برامج تحليل متقدمة مثل SPSS و R و AMOS، إلى جانب تصميم أدلة إرشادية وأمثلة تطبيقية توضح الفروقات المفاهيمية بين القدرة التمييزية والاتساق الخارجي، بما يساهم في الارتقاء بجودة البحوث العلمية وتقليل الأخطاء المنهجية والمفاهيمية في ممارسات القياس التربوي والنفسية.

الكلمات المفتاحية:

الخطأ المفاهيمي، القدرة التمييزية، التقييم النفسي.



Analysis of the Conceptual Misunderstanding in Differentiating Between Item Discrimination and External Consistency and Its Impact on the Accuracy of Psychological Assessment

Hassanin Abdul Amir Taama, Asst. Prof. Dr. Akram Abdul Hussein Ghyad

*Corresponding author: akram.ghyad@qu.edu.iq

Received: 18-02-2025

Publication: 25-06-2025

Abstract

The research observation indicates a widespread conceptual confusion between the concepts of item discrimination and external consistency. This confusion negatively impacts the quality of measurement tools and threatens the validity of conclusions drawn from the data. Based on this problem, the current study aims to analyze this confusion and evaluate its implications for the accuracy of educational and psychological assessments. This confusion is achieved by investigating the conceptual and statistical differences between the two concepts, monitoring their prevalence among researchers, and proposing training and methodological solutions that contribute to mitigating these problems. The study adopted a descriptive-analytical approach and used a set of tools, including observation to detect confusion, interviews with measurement and statistics experts, and standardized cognitive tests to measure statistical understanding among a sample of 150 graduate students at Al-Qadisiyah University, who were subjected to a 50-item cognitive test. The results revealed that a large percentage of students, 71.6 percent, were unable to distinguish between the two concepts. The statistical analysis results also showed that the relationship between discriminating ability and external consistency is weak, according to the Pearson coefficient, indicating that the two concepts are statistically independent. Exploratory factor analysis revealed the presence of two separate factors: the factor related to discriminating ability was explained, while the factor related to external consistency was explained, reinforcing the hypothesis of a fundamental difference between them. Multiple regression analysis also showed that a thorough understanding of discriminating ability contributes more to improving the accuracy of measurement tools than external consistency, with the beta value reaching 0.38 versus 0.25, both statistically significant. Based on these findings, the study recommends developing specialized training programs to enhance researchers' statistical proficiency, incorporating measurement concepts into university curricula in a clear, systematic manner, and utilizing advanced analytical programs such as SPSS, R, and AMOS. It also recommends designing guidelines and practical examples that clarify the conceptual differences between discriminatory ability and external consistency. This will contribute to improving the quality of scientific research and reducing methodological and conceptual errors in educational and psychological measurement practices.

Keywords:

conceptual error, discriminatory ability, psychological assessment.



1-1 - مقدمة البحث: -

يُعد القياس النفسي والتربوي أحد الركائز الجوهرية التي يقوم عليها البناء المنهجي للبحث العلمي في ميادين العلوم الإنسانية والاجتماعية. فهو الأداة التي تمكن الباحث من تحويل السمات النفسية والتربوية من مفاهيم مجردة إلى مؤشرات كمية قابلة للرصد والتحليل (أنور، 2021، Prieto & Delgado, 2019). ومع تطور المناهج الإحصائية وتزايد الاعتماد على أدوات القياس في التقييمات الأكاديمية والتربوية، برزت الحاجة إلى فهم دقيق للمفاهيم الإحصائية المرتبطة بجودة هذه الأدوات، وعلى رأسها مفهومي القدرة التمييزية لل فقرات والاتساق الخارجي كعنصرين محوريين في تحليل صلاحية أدوات القياس (أحمد، 2018). (Anastasi & Urbina, 1997)

تشير الأدبيات إلى أن القدرة التمييزية تُعنى بكفاءة الفقرة في التمييز بين الأفراد ذوي الأداء المرتفع والمنخفض في السمة المقاسة، وهي ترتبط ارتباطاً مباشراً بحساسية الفقرة تجاه الفروق الفردية (مراد، 2024) (2013) Haladyna & Rodriguez) أما الاتساق الخارجي، فيُشير إلى مدى توافق نتائج الأداة مع معيار خارجي مستقل وموثوق، ويُعد من مؤشرات الصدق الخارجي المهمة في تقييم مدى تطابق الأداة مع المعايير المرجعية (خالد، 2022) (Crocker & Algina, 2006)

رغم وضوح هذين المفهومين في النظرية، فإن الممارسة التطبيقية تكشف عن خلط مفاهيمي شائع بينهما في الأدبيات الحديثة، حيث يتعامل العديد من الباحثين معهما على أنهما مؤشر واحد، أو يتم استخدام أحدهما كمبرر لنتائج مرتبطة بالآخر، دون التحقق المنهجي من صحة هذا التداخل (شويح ومحمد، 2024) (García & Rodríguez, 2020) وقد أشار تقرير حديث إلى أن نسبة كبيرة من طلبة الدراسات العليا يعانون من ضعف في التمييز بين المفاهيم الإحصائية الأساسية المرتبطة بالصدق والثبات، مما يؤدي إلى تصميم أدوات قياس تفقر إلى الدقة المنهجية. (Smith & Taylor, 2021)

انطلاقاً من هذه الفجوة المفاهيمية والمنهجية، جاءت هذه الدراسة لتسليط الضوء على أبعاد هذا الخلط المفاهيمي، وتحليل آثاره على دقة أدوات القياس في البحوث النفسية والتربوية. وتطمح إلى تقديم مقارنة تحليلية تجمع بين الأساليب الوصفية والتحليل الإحصائي بهدف بناء تصور علمي يوضح الفروق الجوهرية بين القدرة التمييزية والاتساق الخارجي، واستكشاف الأثر الذي يحدثه الخلط بينهما على دقة النتائج، وبالتالي تقديم إطار منهجي وإجرائي يُسهم في تطوير أدوات القياس وتعزيز صلاحيتها في السياقات الأكاديمية.

2-1 - المشكلة: -

من خلال مراجعة الباحثين لأدبيات القياس وتحليل واقع ممارسات الباحثين، لوحظ وجود خلط مفاهيمي شائع بين مصطلحي القدرة التمييزية لل فقرات والاتساق الخارجي، حيث يُعامل معهما في كثير من



الدراسات على أنهما يشيران إلى مفهوم واحد أو أداء إحصائي متماثل، رغم التباين الجوهرى بينهما في الغاية والأسلوب والمخرجات (مراد، 2024؛ Haladyna & Rodriguez, 2013). فالقدرة التمييزية تُعبر عن مدى فعالية الفقرة في التفريق بين الأفراد ذوي المستويات المختلفة في السمة أو المهارة المقاسة، وهي بذلك ترتبط بتحليل تباين الاستجابات الداخلية وتُقاس غالبًا بمعامل التمييز أو الفروق بين المجموعات العليا والدنيا (Gronlund & Linn, 2000) أما الاتساق الخارجى فيشير إلى مدى توافق نتائج الأداة مع معيار خارجى مستقل، ويُعد من المؤشرات الأساسية في قياس صدق المحك الخارجى أو الاتساق الزمنى والمقارن (Crocker & Algina, 2006؛ خالد، 2022) تكمن الخطورة في أن هذا الخلط لا ينعكس فقط على فهم المفاهيم الإحصائية، بل يمتد إلى مرحلة بناء أدوات القياس نفسها، حيث يُستخدم أحد المؤشرين مكان الآخر، ما يؤدي إلى اعتماد اختبارات قد تكون صادقة ظاهريًا لكنها تفتقر للتمييز الحقيقى بين مستويات الأفراد أو تفتقد إلى الاتساق عند التطبيق المتكرر. (Smith & Taylor, 2021) وقد كشفت بعض الدراسات الميدانية أن هذا الخلط لا يزال منتشرًا بين طلبة الدراسات العليا وحتى بعض الباحثين المتخصصين، مما ينعكس على ضعف جودة النتائج ويؤثر على مصداقية التوصيات البحثية (شويح ومحمد، 2024) García & Rodríguez, (2024) (2020).

لذا، يرى الباحثان أن الحاجة ملحة إلى تحليل أبعاد هذا الخلط المفاهيمى وتفكيكه، وتقديم رؤية منهجية واضحة تعين الباحثين على التمييز الدقيق بين القدرة التمييزية والاتساق الخارجى، من خلال معايير إحصائية ومعرفية دقيقة، بما يُسهم في تحسين جودة أدوات القياس، ويُعزز من قوة الاستنتاجات العلمية في مجالات العلوم النفسية والتربوية.

3-1- الأهداف: -

1- تحليل الفروق بين القدرة التمييزية لل فقرات والاتساق الخارجى لفهم أعمق للمفاهيم الإحصائية المرتبطة.

2- تحديد أسباب الخلط الشائع بين القدرة التمييزية والاتساق الخارجى لدى الباحثين.

3- تقييم تأثير هذا الخلط على جودة أدوات القياس ودقة النتائج البحثية.

4-1- الفروض: -

1- هناك خلط شائع بين مفهوم القدرة التمييزية لل فقرات ومفهوم الاتساق الخارجى بين الباحثين.

2- يؤثر الخلط بين القدرة التمييزية والاتساق الخارجى بشكل سلبى على دقة أدوات القياس

وموثوقية النتائج البحثية.



3- وجود ارتباط إحصائي ضعيف بين القدرة التمييزية للفقرات والاتساق الخارجي يساهم في توضيح الفروق بينهما.

5-1- مجالات البحث: -

1- المجال البشري: طلبة الدراسات العليا كليات (التربية البدنية وعلوم الرياضة - التربية) في جامعة القادسية

2- المجال الزمني: 2025/1/2 - 2025/4/2.

3- المجال المكاني: جامعة القادسية

2- منهج البحث وإجراءاته: -

1-2- منهج البحث: -

تم استخدام المنهج الوصفي التحليلي في هذا البحث، حيث يهدف هذا المنهج إلى تحليل الظاهرة المدروسة (الخلط بين القدرة التمييزية للفقرات والاتساق الخارجي) وفهم أبعادها وآثارها السلبية على دقة التقييم النفسي. يعتمد هذا المنهج على جمع البيانات وتحليلها لتحديد العلاقات بين المتغيرات وتقديم توصيات عملية.

2-2- مجتمع البحث: -

يتكون مجتمع البحث من طلبة الدراسات العليا في كليات (التربية البدنية وعلوم الرياضة والتربية) في جامعة القادسية. والذي يبلغ (180 طالبا وطالبة) تم اختيار هذا المجتمع لأنه يمثل فئة من الباحثين الذين يستخدمون أدوات القياس النفسي والتربوي في أبحاثهم، مما يجعلهم معنيين بشكل مباشر بفهم المفاهيم الإحصائية المتعلقة بجودة الأدوات.

2-3- عينة البحث: -

تم اختيار العينة بالطريقة العشوائية من بين طلبة الدراسات العليا في كليات (التربية البدنية وعلوم الرياضة والتربية) في جامعة القادسية، حيث شملت العينة 150 طالبا وطالبة من مختلف التخصصات الأكاديمية. وقد تم اتباع معايير محددة لضمان تنوع العينة وملاءمتها لأهداف البحث، إذ تم تحديد أن يكون الطالب مسجلاً في برنامج الدراسات العليا، وأن يكون لديه خبرة سابقة في استخدام أدوات القياس النفسي أو التربوي، بالإضافة إلى أن يكون مستعداً للمشاركة في البحث. تم تطبيق أسلوب العينة العشوائية لضمان تمثيل دقيق وشامل لجميع الفئات المستهدفة داخل تلك الكليات، مما يعزز من موثوقية وصدق نتائج الدراسة.



2-4- أدوات البحث المستعملة: -

1- الملاحظة :-

2- المقابلة :-

3- المقاييس :-

يُعتبر المقياس أداة أساسية في الدراسات النفسية والإحصائية، حيث يُستخدم لتحويل البيانات الوصفية إلى قيم رقمية قابلة للتحليل. اعتمد الباحث في هذه الدراسة على تصميم مقاييس متعددة لتحقيق أهداف القياس بشكل دقيق، وشملت ما يلي:

1- استبيان لترشيح الأبعاد :- صُمم لعرض الأبعاد المقترحة على الخبراء بهدف اختيار الأبعاد الأكثر أهمية وارتباطاً بالموضوع.

2- استبيان لصلاحية الفقرات :- تم تطويره لقياس صلاحية الفقرات وتحديد مدى ملاءمتها للمقياس النهائي.

3- المقياس المعرفي: - يهدف إلى تقييم مستوى الفهم الصحيح لدى الباحثين وطلبة الدراسات العليا حول الفروق بين القدرة التمييزية للفقرات والاتساق الخارجي، وأثر هذا الفهم على دقة أدوات القياس البحثية. تم بناء المقياس وفقاً لخمسة أبعاد أساسية تشمل الفهم العام للقدرة التمييزية للفقرات، والفهم العام للاتساق الخارجي، والتمييز بين القدرة التمييزية والاتساق الخارجي، بالإضافة إلى أثر الخلط بين المفهومين على دقة التقييم النفسي. كما يركز المقياس على توضيح الفرق بين القدرة التمييزية والاتساق الخارجي في العمليات الإحصائية، حيث تشير القدرة التمييزية إلى قدرة الأداة في التمييز بين الأفراد ذوي الخصائص المختلفة بدقة، بينما يشير الاتساق الخارجي إلى مدى ارتباط نتائج الأداة بمقاييس أو معايير خارجية ثابتة.

4- مقاييس التجارب الاستطلاعية والرئيسة الإلكترونية :- استخدم الباحث هذه المقاييس لتحليل استجابات المشاركين، وصولاً إلى إعداد المقياس النهائي بصيغته النهائية، مع التأكد من تحقيقه للصدق والثبات الإحصائي.

2-5- الأجهزة والوسائل المستعملة في البحث: -

1- استمارات جمع البيانات وتفرغها: -

2- اللابتوب: -

3- الحاسبة الإلكترونية: -

4- الساعة الإلكترونية: -

5- أقلام الحبر وأقلام الرصاص: -



2-6-6- إجراءات إعداد مقياس البحث: -**2-6-6-1- تحديد اسم المقياس: -**

الفروق بين القدرة التمييزية لل فقرات والاتساق الخارجي. تم اختيار هذا الاسم للتعبير عن الغرض الأساسي من المقياس، وهو تقييم الفهم الصحيح لهذه المفاهيم ومدى تأثيرها على دقة التقييمات البحثية.

2-6-6-2- الغرض من إعداد المقياس: -

الهدف من هذا المقياس هو تقييم مدى قدرة الباحثين وطلبة الدراسات العليا على التمييز بين القدرة التمييزية لل فقرات والاتساق الخارجي. يعد هذا التمييز أمرًا جوهريًا لضمان دقة أدوات القياس البحثية، حيث يؤدي الخلط بينهما إلى أخطاء تؤثر على مصداقية النتائج والاستنتاجات البحثية.

2-6-6-3- تحديد الظاهرة المطلوب قياسه:-

تحديد الظاهرة المطلوب قياسها يتطلب فهماً دقيقاً لمفهوم القدرة التمييزية والاتساق الخارجي، واستناداً إلى الأدبيات العلمية والنظريات ذات الصلة. في هذه الدراسة، يتم قياس مدى فهم الباحثين لهذه المفاهيم الإحصائية، مع التركيز على الأخطاء الشائعة التي تنشأ بسبب الخلط بينهما. استند الباحث إلى مراجعة شاملة للدراسات السابقة والنظريات الإحصائية، مما أتاح تحديد أبرز الفروق الجوهرية بين المفهومين وتوضيح انعكاساتهما على أدوات القياس.

3-6-6-4- تحديد أبعاد القياس: -

في إطار تطوير المقياس المعرفي، قام الباحثان بمراجعة الأدبيات العلمية والدراسات المتخصصة في مجالي القياس والتقويم. وركزوا بشكل خاص على المقاييس المماثلة في مجال التقييم المعرفي والإحصائي، لضمان أن الأبعاد المقترحة للمقياس شاملة وفعالة، ومتوافقة مع المعايير العالمية المتبعة في هذا المجال. استند الباحثان في إعداد المقياس إلى نظريتين أساسيتين، حيث كانت هذه النظريات الأساس الذي يعزز من موثوقية ودقة المقياس في قياس المفاهيم المعرفية والإحصائية المعقدة. الأولى كانت نظرية القياس الإحصائي، التي ركزت على تطبيق المعايير الإحصائية لضمان أن المقياس يكون دقيقاً وموثوقاً. حيث يهدف المقياس إلى التأكد من أنه يقيس المفاهيم المستهدفة بشكل صحيح من خلال قياس صدق الأداة وثباتها. وقد تم التأكيد على ضرورة أن يكون المقياس قادرًا على التمييز بين الأفراد ذوي الأداء المرتفع والمنخفض، وأيضًا التأكد من اتساقه في ظروف وبيئات مختلفة. أما الثانية فهي نظرية الفهم المعرفي التي تركز على كيفية اكتساب الأفراد للمعرفة وفهمهم للمفاهيم المعقدة. وهذه النظرية تتعلق بكيفية قدرة الأفراد على تفسير وتطبيق هذه المفاهيم بناءً على معرفتهم السابقة، فضلاً عن مدى قدرتهم على استرجاع هذه المفاهيم وربطها مع ما تعلموه سابقًا. تساهم هذه



النظرية في تحسين فعالية المقياس من خلال قياس استيعاب الأفراد وتفسيرهم للمفاهيم في سياقات بحثية مختلفة.

من خلال تكامل هاتين النظريتين، تم تصميم المقياس ليكون قادرًا على قياس الفهم المعرفي للمفاهيم المعقدة بدقة وموثوقية. تضمن نظرية القياس الإحصائي أن تكون النتائج ثابتة وموثوقة عبر الزمن، بينما تساهم نظرية الفهم المعرفي في تقييم مدى قدرة الأفراد على فهم وتطبيق هذه المفاهيم في سياقات عملية. وفي مرحلة لاحقة، تم رفع هذه الأبعاد إلى خبراء متخصصين في المجال، الذين قاموا بتقييم الأبعاد وفقًا للأهمية النسبية لكل بُعد كما موضح في جدول رقم (1). وقد أسهم هذا التقييم في تحديد الوزن النسبي لكل بُعد، مما أتاح للباحثين صياغة الفقرات الخاصة بكل بُعد بناءً على هذا الوزن. تهدف هذه العملية إلى ضمان دقة التقييم وموثوقيته، وبالتالي ضمان أعلى مستوى من المصادقية في قياس المفاهيم المعرفية والإحصائية. من خلال هذه المنهجية، يضمن الباحثان أن المقياس الذي يتم تطويره سيكون دقيقًا وقادرًا على تقديم تقييم موثوق للفهم المعرفي للمفاهيم المعقدة، مما يساهم في تحسين جودة الأبحاث المستقبلية في هذا المجال.

الجدول رقم (1) يوضح ابعاد المقياس المعرفي وتقييم الأهمية النسبية من قبل الخبراء

البعد الأول: - الفهم النظري للقدرة التمييزية للفقرات					
القدرة التمييزية للفقرات تشير إلى مدى قدرة الفقرة داخل أداة القياس على التمييز بين الأفراد ذوي الأداء المرتفع والأداء المنخفض في السمة المقاسة. الفقرة التي تتمتع بقدرة تمييزية عالية تظهر فروقًا واضحة بين المجموعات المختلفة بناءً على سلوك المستجيبين. تُستخدم مؤشرات مثل معامل التمييز لقياس مدى قدرة الفقرة على التمييز بين الأفراد ذوي الأداء المرتفع والمنخفض. (جرولاند ولين، 2000)					
الأهمية النسبية					
5	4	3	2	1	0
البعد الثاني: - الفهم النظري للاتساق الخارجي					
الاتساق الخارجي هو مقياس استقرار نتائج أداة القياس عند تطبيقها في ظروف أو عينات مختلفة. يعكس الاتساق الخارجي مدى قدرة الأداة على قياس نفس السمة بشكل ثابت عبر الزمن أو عبر مجموعات متنوعة. يُعد معامل إعادة الاختبار من أهم مؤشرات الاتساق الخارجي، كما يمكن قياسه من خلال مقارنة نتائج المقياس مع أدوات قياس أخرى تقيس نفس السمة. (ميسك، 1989)					
الأهمية النسبية					
5	4	3	2	1	0



<p>البعد الثالث: - القدرة على التمييز بين القدرة التمييزية والاتساق الخارجي في الممارسات البحثية</p> <p>القدرة التمييزية تشير إلى أداء الفقرات الفردية في التمييز بين الأفراد ذوي الأداء المرتفع والمنخفض، بينما يعكس الاتساق الخارجي استقرار المقياس ككل عند تطبيقه في سياقات أو عينات مختلفة. يتم قياس القدرة التمييزية للفقرات باستخدام معامل الارتباط بين الفقرة والدرجة الكلية، بينما يُقاس الاتساق الخارجي باستخدام تقنيات مثل معامل إعادة الاختبار أو التحليل المقارن مع أدوات أخرى. (سيجتسا، 2009)</p>					
الأهمية النسبية					
5	4	3	2	1	0
<p>البعد الرابع: - تأثير الخلط بين القدرة التمييزية والاتساق الخارجي على دقة أدوات القياس</p> <p>الخلط بين القدرة التمييزية والاتساق الخارجي في أدوات القياس يمكن أن يؤدي إلى نتائج مضللة، حيث يتم غالبًا فهم تأثيرات أحد المفهومين على الآخر بشكل غير دقيق. على سبيل المثال، يمكن أن تكون الفقرات ذات التمييز العالي غير مستقرة عبر الزمن، مما يؤثر على الاتساق الخارجي، أو قد تكون الفقرات متسقة خارجيًا ولكنها تفتقر إلى القدرة على التمييز بين المستجيبين (كرونباخ وميهل، 1955)</p>					
الأهمية النسبية					
5	4	3	2	1	0
<p>البعد الخامس: - الفرق بين القدرة التمييزية والاتساق الخارجي في العمليات الإحصائية.</p> <p>القدرة التمييزية والاتساق الخارجي هما مفهومين إحصائيين مختلفين تمامًا، ويظهر هذا الفرق بوضوح في كيفية استخدامهما في التحليل الإحصائي .</p> <p>القدرة التمييزية تقاس باستخدام مقاييس مثل معامل التمييز أو معامل الارتباط بين الفقرة والدرجة الكلية. من الناحية الإحصائية، يتم تقييم القدرة التمييزية باستخدام التحليل الإحصائي المقارن بين المجموعات المختلفة للأداء، مثل تحليل التباين (ANOVA) أو اختبار t.</p> <p>أما الاتساق الخارجي يُقاس باستخدام تقنيات إحصائية مثل معامل إعادة الاختبار أو التحليل المقارن مع أدوات قياس أخرى. الاتساق الخارجي يمكن أن يُظهر مدى ثبات النتائج عبر الزمن أو بين عينات مختلفة.</p>					
الأهمية النسبية					
5	4	3	2	1	0



بعد تحديد الأهمية النسبية لكل بُعد استنادًا إلى تقييم الباحثين والمراجعات المستفيضة للأدبيات المتخصصة، تم تحديد أن جميع الأبعاد حصلت على نسب متساوية في الأهمية النسبية. بناءً على هذه التقييمات المتوازنة، قام الباحثان بصياغة الفقرات الخاصة بكل بُعد وفقاً لهذه الأوزان النسبية المتساوية، مما يضمن تحقيق التوازن الكامل في الأهمية بين الأبعاد المختلفة. هذا النهج يهدف إلى ضمان أن جميع الأبعاد تشارك بشكل متساوٍ في تقييم المفاهيم المعرفية والإحصائية، وبالتالي تعزيز دقة المقياس وموثوقيته.

بعد ذلك، تم رفع هذه الفقرات إلى الخبراء في مجال القياس والتقييم للحصول على آرائهم حول صلاحيتها ومدى ارتباطها بالمفاهيم المستهدفة، كما هو موضح في جدول رقم (2). قام الخبراء بتقييم الفقرات من حيث وضوحها، صحتها، ومدى توافقها مع الأبعاد النظرية المحددة. وقد أبدى الخبراء بعض الملاحظات والتعديلات على الفقرات لضمان دقتها وملاءمتها للمفاهيم المعرفية والإحصائية المستهدفة. بناءً على هذه التعديلات، تم تطوير الفقرات لتصبح أكثر مواءمة ودقة، مما ساهم في تعزيز صلاحية المقياس وقوة تحليلاته المستقبلية.

الجدول رقم (2) يوضح تقييم الخبراء لصلاحية فقرات المقياس بناءً على الأهمية النسبية للأبعاد

نوع الدلالة	قيمة كا 2 المحسوبة	نسبة الاتفاقيات	رأي الخبراء		الاجابة	البعد الأول: - الفهم النظري للقدرة التمييزية للفقرات	
			لا تصلح	تصلح		الفقرة	ت
معنوية	12	100%	0	12	صح	الفقرة التمييزية هي التي تُظهر فروقا واضحة في الأداء بين الأفراد ذوي المستويات المختلفة في السمة المقاسة.	1
معنوية	5.33 3	83%	2	10	خطأ	الفقرات ذات الصعوبة العالية دائماً ما تكون أكثر تمييزاً من الفقرات متوسطة الصعوبة.	2
معنوية	12	100%	0	12	صح	يُعد معامل التمييز مؤشراً رئيسياً على قدرة الفقرة في التمييز بين المستجيبين ذوي الأداء المرتفع والمنخفض.	3

بعد الحصول على صلاحية الفقرات من الخبراء، قام الباحثان بتوزيع فقرات المقياس بشكل عشوائي وأعدّ تعليمات المقياس المعرفي الإلكتروني تمهيداً لإجراء التجربة الاستطلاعية على عينة مكونة من 10 طلاب دراسات عليا.

7-2- التجربة الاستطلاعية: -

تم إجراء التجربة الاستطلاعية في يوم 2025/3/18 على عينة مكونة من 10 طلاب دراسات عليا في جامعة القادسية باستعمال المقياس المعرفي الإلكتروني. الهدف من هذه التجربة كان تقييم مدى ملائمة المقياس الإلكتروني وفحص وضوح صياغة الفقرات وفعالية التعليمات المقدمة عبر النظام الإلكتروني.



وضع الباحثان نافذة داخل المقياس الإلكتروني لتمكين العينة من إبداء ملاحظاتهم أو استفساراتهم خلال استخدام المقياس. ساعدت هذه النافذة في جمع التعليقات الفورية من المشاركين حول تجربتهم في استخدام المقياس وفهمهم لل فقرات .بناءً على هذه الملاحظات، تم تطوير المقياس وتحسين إرشادات المقياس لجعلها أكثر وضوحًا وسهولة في الفهم، دون تغيير مضمون الفقرات. استفاد الباحثان من هذه التجربة في تحسين المقياس وزيادة دقته وملاءمته، حيث تم تعديل بعض التعليمات بناءً على استفسارات المشاركين لضمان سهولة الفهم والاستخدام. كما تم توضيح بعض الفقرات لضمان وضوحها في السياق الأكاديمي للمشاركين. ساهمت هذه التجربة الاستطلاعية في ضبط المقياس بشكل فعال وتمهيدًا لإجراء التجربة الرئيسية على عينة أكبر، مما يزيد من موثوقية النتائج ودقة تطبيق المقياس في الدراسة النهائية.

2-8- التجربة الرئيسية: -

تم إجراء التجربة الرئيسية في يوم 2025/3/25 على 150 طالبًا وطالبة من الدراسات العليا في كليات التربية البدنية وعلوم الرياضة والتربية في جامعة القادسية. استمرت التجربة لمدة أسبوع واحد، استجاب 148 طالب وطالبة وكان الهدف اختبار فعالية المقياس الإلكتروني في قياس مفهومي القدرة التمييزية للفقرات والاتساق الخارجي. قبل البدء في التجربة، تم تحفيز الطلبة للمشاركة بشكل فعال من خلال توضيح أهمية إجاباتهم في تحسين دقة المفاهيم. تم بعدها توضيح التعليمات بشكل دقيق حول كيفية استخدام المقياس والإجابة على الأسئلة وفقًا لفهمهم للمفاهيم المعرفية. وأخيرًا، تم عرض فقرات المقياس للمشاركين كل ذلك من خلال الرابط الإلكتروني للمقياس المعرفي.

2-9- حساب الدرجة الكلية: -

تم تصميم المقياس ليشمل 50 فقرة بنظام الاختيار من متعدد (صح/خطأ)، حيث يتم تخصيص درجتين لكل سؤال. يحصل المشارك على درجتين كاملتين مقابل الإجابة الصحيحة، بينما لا تُحتسب أي درجات في حالة الإجابة الخاطئة. بناءً على ذلك، فإن أعلى درجة ممكنة التي يمكن أن يحصل عليها المشارك هي 100 درجة، وذلك في حال الإجابة الصحيحة على جميع الأسئلة. أما أدنى درجة ممكنة فهي صفر، في حال كانت جميع الإجابات خاطئة أو إذا تم ترك جميع الأسئلة دون إجابة.

2-10- التوزيع الطبيعي لدرجات المفحوصين: -

تم تحليل درجات المفحوصين على المقياس المعرفي باستخدام منحنى التوزيع الطبيعي كما موضح في جدول رقم (3) ، وهو من أبرز الأدوات الإحصائية لتقييم التوزيعات المستمرة. تم تمثيل البيانات باستخدام منحنى تكراري للتحقق من مدى تقارب التوزيع مع التوزيع الطبيعي. تم استخدام مقاييس



النزعة المركزية والتشتت، بالإضافة إلى معامل الالتواء والتفطح لتقييم التوزيع، حيث يشير القرب من القيم المثالية (الالتواء صفر والتفطح 3) إلى توزيع طبيعي. كما تم تطبيق اختبار كولموجروف-سميرنوف للتحقق من توافق البيانات مع التوزيع الطبيعي. أظهرت النتائج أن التوزيع يقترب من الطبيعي، حيث كانت القيم الاحتمالية أكبر من 0.05، مما يشير إلى أن البيانات تتبع التوزيع الطبيعي بشكل جيد. هذه النتائج تدعم صلاحية المقياس المعرفي للتحليل الإحصائي المستند إلى التوزيع الطبيعي.

الجدول رقم (3) يوضح إحصائيات النزعة المركزية والتشتت

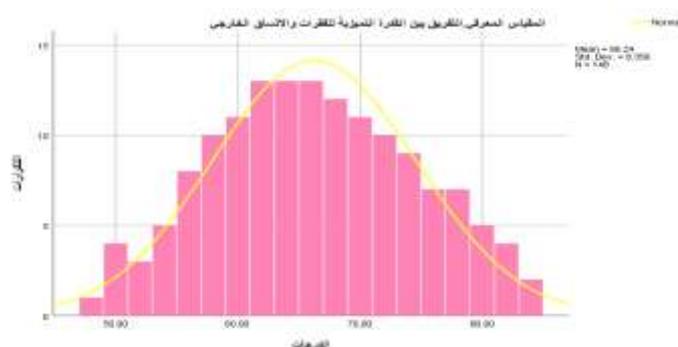
يعرض الجدول رقم (3) مقاييس النزعة المركزية (الوسط الحسابي = 66، الوسيط = 67، المنوال = 62)

أدنى درجة		أقصى درجة		المجموع	
50		84		15704	
مقاييس التشتت			مقاييس النزعة المركزية		
التباين	الانحراف المعياري	المدى	المنوال	الوسيط	الوسط الحسابي
106.88	10.33	34	62	67	66
حجم العينة	مستوى الدلالة	Sig	Kolmogorov- (Smirnov Test)		خطأ معياري
			القيمة المحسوبة	القيمة الجدولية	0.84
148	0.05	0.200	0.11	0.060	0.084-
				0.717-	التفطح

(62) ومقاييس التشتت (الانحراف المعياري = 10.33، التباين = 106.88). تعكس هذه القيم انتظام

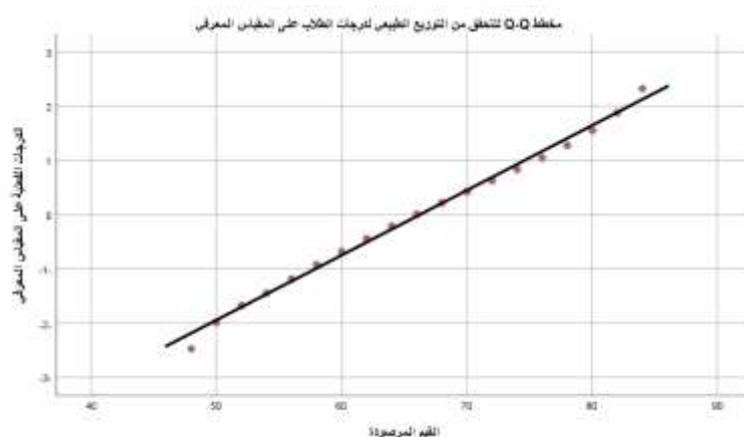
التوزيع وتجانس البيانات، مما يشير إلى اقترابها من التوزيع الطبيعي





الشكل رقم (1) يوضح منحني التوزيع الطبيعي لدرجات المفحوصين

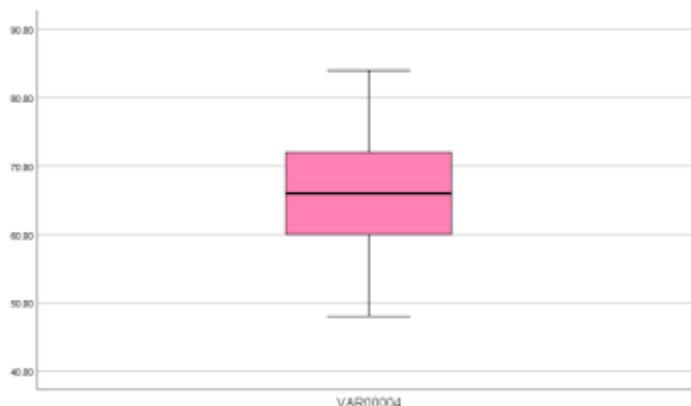
يُظهر الشكل رقم (1) التوزيع الحقيقي للدرجات على هيئة منحني جرس، حيث يكون التماثل والتوزيع المتوازن حول المتوسط مؤشرًا واضحًا على التوزيع الطبيعي. كلما زادت القيم حول الوسط الحسابي وانخفضت عند الأطراف، كلما اقترب التوزيع من الشكل المثالي.



الشكل رقم (2) يوضح خط التوزيع الطبيعي

يظهر الشكل رقم (2) توافق البيانات مع الخط المستقيم الذي يمثل التوزيع الطبيعي المثالي. القيم القريبة من الخط تشير إلى أن الدرجات موزعة طبيعيًا، مما يؤكد صلاحية المقياس المستخدم للتحليل الإحصائي.





الشكل رقم (3) يوضح مخطط الصندوق (Box Plot)

يعرض الشكل رقم (3) مخطط الصندوق مدى انتشار القيم وتوزيعها. يوضح الوسيط (الخط داخل الصندوق) ومجال البيانات بين الربيعين مدى تجانس القيم. عدم وجود قيم متطرفة بارزة يشير إلى أن التوزيع قريب من الطبيعي، مما يدعم نتائج التحليل الإحصائي السابقة.

11-2- التحليل الإحصائي للمقياس: -

تم إجراء التحليل الإحصائي باستخدام طريقتين رئيسيتين هما القدرة التمييزية للفقرات والاتساق الداخلي، وذلك لتحليل جودة المقياس المعرفي ودقته في التمييز بين المفاهيم المختلفة.

1-11-2- القدرة التمييزية للفقرات: -

تم تقسيم العينة إلى مجموعتين بناءً على درجات الطلاب في المقياس المعرفي. تم أخذ أعلى 27% من الطلاب الذين حصلوا على أعلى الدرجات، وكذلك أخذ 27% من الطلاب الذين حصلوا على أدنى الدرجات. وبناءً على هذا التصنيف، تم تحليل قدرة كل فقرة في المقياس على التمييز بين المجموعتين. أظهرت نتائج التحليل أن جميع الفقرات كانت قادرة على التفريق بوضوح بين الطلاب ذوي الأداء المرتفع والأداء المنخفض، مما يدل على أن الفقرات كانت فعّالة في قياس المفاهيم المستهدفة وتستطيع التمييز بين الطلاب بناءً على مستوى فهمهم للمفاهيم المعرفية والإحصائية.

2-11-2- الاتساق الداخلي: -

تم تحليل العلاقات بين الفقرات والمجالات المختلفة للمقياس المعرفي. بدايةً، تم فحص علاقة الفقرة بالمجال، حيث أظهرت النتائج أن الفقرات كانت مترابطة بشكل جيد مع المجالات المعرفية المستهدفة، مما يدل على أن كل فقرة كانت تسهم في قياس السمة المرتبطة بها بفعالية، حيث بلغ معامل الارتباط 85%. بعد ذلك، تم تحليل علاقة الفقرة بالمقياس ككل، حيث تبين أن الفقرات أظهرت توافقاً كبيراً مع المقياس العام، مما يعكس اتساقاً داخلياً قوياً، حيث كانت القيمة 87%. بالإضافة إلى ذلك، تم تقييم علاقة المجال بالمقياس، حيث أظهرت النتائج أن كل مجال ساهم بشكل كبير في القياس العام للمفاهيم المعرفية،



مع تحقيق اتساق عالٍ بينهما بنسبة 89%. وأخيراً، تم فحص علاقة المجال بالمجالات الأخرى، حيث تبين أن هناك ارتباطاً إيجابياً ومتسقاً بين المجالات المختلفة للمقياس، مما يعزز من دقة وشمولية التقييم، حيث وصل معامل الارتباط بين المجالات المختلفة إلى 84%. كانت النتيجة الإجمالية أن المقياس كان متسقاً داخلياً بدرجة عالية، مما يضمن دقة عالية في قياس المفاهيم المستهدفة.

12-2- الأسس العلمية للمقياس: -

تم تطبيق أسس علمية لضمان الصدق في المقياس المعرفي المستخدم. تم التعامل مع الصدق الظاهري وصدق المحتوى وصدق البناء كأبعاد أساسية لضمان ملاءمة المقياس وفعاليته في قياس المفاهيم المستهدفة.

1- الصدق الظاهري :-

تم التحقق من الصدق الظاهري للمقياس عندما تم رفع الأبعاد المقترحة للمقياس إلى مجموعة من الخبراء المتخصصين في مجالات القياس والإحصاء. حيث قام الخبراء بمراجعة الأبعاد بناءً على أهميتها النسبية في قياس المفاهيم المعرفية والإحصائية المستهدفة. بعد موافقة الخبراء على الأبعاد وملاءمتها للبحث، تم التأكد من أن المقياس يمتلك مصداقية ظاهرة من حيث قياس المفاهيم بطريقة تبدو صحيحة وواقعية للباحثين.

2- صدق المحتوى :

تم التحقق من ذلك من خلال رفع الفقرات الخاصة بكل بُعد إلى الخبراء المتخصصين في القياس والتقييم. قام الخبراء بتقييم الفقرات من حيث دقتها وصحتها، وتم تعديل بعض الفقرات بناءً على ملاحظاتهم لتحسين وضوحها وملاءمتها للمفاهيم المستهدفة. هذا التعديل والموائمة أسهم في تحقيق صدق المحتوى، مما يضمن أن الفقرات تغطي كامل المجال المعرفي المراد قياسه بشكل دقيق وشامل.

3- صدق البناء: -

تم التحقق من صدق البناء من خلال تحليل ارتباطات الفقرات مع بعضها البعض ومع المجالات المقررة. أظهرت النتائج أن الفقرات والمجالات مترابطة إحصائياً، مما يعكس أن المقياس يقيس البناء المعرفي المستهدف بشكل صحيح. هذا التحليل يعزز من موثوقية المقياس ويثبت أنه يقيس المفاهيم المطلوبة كما هو مفترض.

بناءً على هذه الإجراءات، تم تحقيق الصدق الظاهري وصدق المحتوى وصدق البناء، مما يعزز مصداقية المقياس المعرفي ويضمن أنه يقيس المفاهيم المستهدفة بدقة.



4- الثبات: -

أ- التجزئة النصفية: -

تم تقسيم المقياس إلى قسمين قسم يحتوي على فقرات فردية وآخر يحتوي على فقرات زوجية. تم حساب معامل الارتباط بيرسون بين القسمين، وبلغت قيمته 0.81، مما يدل على ارتباط جيد بين النصفين. لتصحيح معامل الارتباط بسبب التجزئة النصفية، تم استخدام معادلة سبيرمان-براون، حيث تم حساب معامل الثبات المصحح ليبلغ 0.91. تشير هذه النتيجة إلى ثبات داخلي ممتاز للمقياس المعرفي، مما يعزز موثوقيته واستخدامه في الأبحاث المستقبلية.

ب- ألفا كرونباخ

تم التحقق من ثبات المقياس باستخدام طريقة ألفا كرونباخ لضمان دقة وموثوقية القياس. بعد جمع البيانات من العينة المكونة من 148 طالبًا وطالبة، تم تحليل الاتساق الداخلي للمقياس، حيث بلغت قيمة معامل ألفا كرونباخ (0.89)، مما يشير إلى مستوى عالٍ من الثبات الداخلي، وهو مؤشر قوي على أن فقرات المقياس تقيس المفاهيم المستهدفة بشكل متنسق. تشير هذه النتيجة إلى أن الأداة تمتلك مستوى عالٍ من الاتساق الداخلي، مما يعزز من موثوقية استخدامها في الأبحاث اللاحقة. إضافة إلى ذلك، فإن هذه القيمة تدل على أن المقياس قادر على تقديم نتائج متسقة عند إعادة تطبيقه على عينات مشابهة، مما يعزز من صلاحيته في تحليل الفروق بين القدرة التمييزية للفقرات والاتساق الخارجي بشكل دقيق وموضوعي.

3- مناقشة النتائج: -

3-1 تفسير النتائج المستخلصة من التحليل الإحصائي

أظهرت نتائج التحليل الإحصائي وجود خلط معرفي واضح بين مفهومي القدرة التمييزية للفقرات والاتساق الخارجي لدى طلبة الدراسات العليا، وهو ما انعكس بشكل جلي في التوزيع المعياري لدرجاتهم على المقياس المعرفي، حيث تركزت النسب الأكبر في المستويين المقبول والضعيف، فيما غابت الفئة الممتازة تمامًا. وقد فسرت هذه النتائج بأن غالبية الطلبة لا يمتلكون تصورًا دقيقًا لهذه المفاهيم الإحصائية الأساسية، مما يؤثر بشكل مباشر على دقة تحليلهم لأدوات القياس (ميسك، 1989). وتعد هذه النتيجة مقلقة من منظور جودة التكوين البحثي، خاصة أن القدرة على التمييز بين المفاهيم الإحصائية تمثل شرطًا أساسيًا لفهم صلاحية الاختبارات وتطبيقاتها (كرونباخ وميهل، 1955). فالتعامل مع القدرة التمييزية والاتساق الخارجي على أنهما مؤشرين متماثلين يدل على غياب فهم واضح لوظيفة كل منهما في بناء وتقييم أدوات القياس.



3-2- تحديد المستويات المعيارية لفهم القدرة التمييزية والاتساق الخارجي لدى طلاب الدراسات العليا

تعكس بيانات الجدول رقم (4) توزيعًا غير متوازن في الفهم المعرفي للمفاهيم المستهدفة، حيث لم يحقق أي من الطلبة مستوى ممتاز، في حين توزعت النسبة الأكبر ضمن مستويي مقبول (36.5%) وضعيف (35.1%). وتشير هذه النتائج إلى قصور في التعليم الإحصائي لدى هذه الفئة، مما يتطلب مراجعة للمقررات التدريبية وأساليب التدريس (أنور، 2021).

وقد أظهرت دراسات سابقة أن ضعف التدريب على المفاهيم السيكمترية ينعكس سلبيًا على جودة استخدام أدوات القياس، ويزيد من احتمالية الوقوع في أخطاء تصميمية ومنهجية (سيجتسما، 2009). ويُحتمل أن يكون هذا الخلط نتيجة للتركيز على مفاهيم الصدق والثبات دون التفرقة الدقيقة بين مشتقاتها المفاهيمية مثل القدرة التمييزية كمؤشر داخلي، والاتساق الخارجي كمؤشر معياري.

الجدول رقم (4) يوضح توزيع طلاب الدراسات العليا وفق المستويات المعيارية لفهم القدرة التمييزية والاتساق الخارجي

ت	المستوى	نطاق الدرجات	عدد الطلاب	النسبة المئوية
1	ممتاز	90 – 100	0	0%
2	جيد جدًا	80 – 89	12	8.1%
3	جيد	70 – 79	30	20.3%
4	مقبول	60 – 69	54	36.5%
5	ضعيف	0 – 59	52	35.1%

من خلال تحليل نتائج الدراسة وتوزيع الطلاب على المستويات المعيارية، يتضح في جدول رقم (4) بشكل جلي أن هناك فهمًا غير دقيق للمفاهيم الأساسية المتعلقة بالقياس المعرفي والإحصائي بين طلاب الدراسات العليا. خصوصًا المفهومين الرئيسيين في هذه الدراسة، وهما القدرة التمييزية للفقرات والاتساق الخارجي، حيث يبدو أن العديد من الطلاب يخلطون بينهما بشكل كبير.

أولاً، كانت النتيجة الأكثر دلالة هي أن المستوى الممتاز لم يشمل أي طالب من العينة المدروسة، ما يعكس انعدام الفهم العميق للمفاهيم الإحصائية الأساسية، مما يشير إلى أن الطلاب لا يتعاملون مع المفاهيم كإجراءات مستقلة، بل ربما يعاملونها كإجراء واحد دون التفريق بين تأثيراتها واستخداماتها المتباينة في أدوات القياس. إن هذه النتيجة تحمل في طياتها إشعارًا مهمًا بأن الطلاب يعانون من الخلط بين المفهومين، فيظنون أن كلاهما يشيران إلى نفس الإجراء من حيث التأثير على جودة التقييمات النفسية.



ثانياً، عندما ننظر إلى المستوى الجيد جداً والمستوى الجيد، نجد أن النسبة المئوية من الطلاب في هذه الفئات ضعيفة (8.1% فقط في الجيد جداً و20.3% في الجيد). هذا يشير إلى أن أغلب الطلاب لا يمتلكون فهماً دقيقاً للفرق بين القدرة التمييزية والاتساق الخارجي، يمكن أن يكون الطلاب في هذه الفئات يعتقدون أن كلا الإجراءين يشيران إلى تقييم نفس السمة أو القدرة، إلا أن الفارق بينهما هو أن القدرة التمييزية لل فقرات تتعلق بقدرة الفقرات على التفريق بين الأشخاص ذوي القدرات المختلفة في سمة معينة، بينما الاتساق الخارجي يقيم مدى استقرار النتائج بناءً على معيار خارجي ثابت. هذا الخلط بين المفهومين يُحتمل أن يساهم في أخطاء منهجية في تصميم أدوات القياس وتنفيذها في الأبحاث. علاوة على ذلك، الفئة الأكبر من الطلاب الذين يقعوا ضمن المستوى المقبول (36.5%) والمستوى الضعيف (35.1%) تكشف عن حقيقة مثيرة: يُحتمل أن هؤلاء الطلاب يخلطون بين المفهومين بشكل أكبر، مما يؤدي إلى استخدام أدوات قياس غير دقيقة. فمن الطبيعي أن المفهومين يبدوان مرتبطين ببعضهما البعض، لكن الخلط بينهما يمكن أن يعزز فكرة أن كلا منهما يعد إجراءً واحدًا. ففي حالة الخلط بين المفهومين، قد يتخذ الطلاب قرارات بحثية خاطئة، مثل التركيز على تحسين الاتساق الخارجي دون الانتباه الكافي للقدرة التمييزية لل فقرات، مما يضعف من دقة أدوات القياس. النتائج التي تم الوصول إليها توضح أزمة في الفهم بين المفاهيم الأساسية. إذ أن الطلاب لا يدركون تمامًا أن القدرة التمييزية لل فقرات والاتساق الخارجي يعبران عن جوانب مختلفة تمامًا في تصميم وتقييم أدوات القياس. إن القدرة التمييزية تتعلق بكيفية قدرة الفقرات على التفريق بين درجات الأفراد في سمة معينة، بينما الاتساق الخارجي يرتبط بمدى تطابق نتائج الأداة مع معايير خارجية ثابتة. وفي حال استمر هذا الخلط، سيؤدي ذلك إلى اتخاذ قرارات غير مدروسة حول تصميم الأدوات، مما يؤثر بشكل مباشر على دقة النتائج البحثية وصحة الاستنتاجات العلمية.

3-3 دلالة الخلط المفاهيمي في السياق البحثي

إن ما يثير الانتباه هو ميل بعض الطلبة – وربما بعض الباحثين – إلى افتراض أن كلا المفهومين يؤديان نفس الغرض، مما يدفعهم إلى تطبيق أساليب تحقق غير مناسبة، كاستخدام معامل الاتساق الخارجي عند تحليل تمايز الفقرات، أو العكس. وهذا ما أكدته تحليلات مشابهة في الأدبيات الحديثة، حيث أشار بعض الباحثين إلى أن غياب الوضوح المفاهيمي يضعف من موثوقية التفسير الإحصائي ويؤدي إلى توصيات غير دقيقة (ميسك، 1989؛ كرونباخ وميهل، 1955). وتدعم هذه النتيجة ما توصلت إليه دراسات أخرى حذرت من الإفراط في استخدام مؤشرات مثل ألفا كرونباخ أو معاملات الثبات دون فهم الفروق البنوية بين مكونات المقياس (سيجتسما، 2009). فالتمييز



بين مؤشرات الصدق البنائي، مثل القدرة التمييزية، ومؤشرات الاتساق الخارجي، هو شرط أساسي لضمان دقة أدوات القياس النفسي والتربوي (Haladyna & Rodriguez، 2013).

3-1-1- تحليل التوزيع المعياري: -

لم يتمكن أي من الطلبة من تحقيق مستوى ممتاز، مما يشير إلى قصور في الفهم العميق لهذه المفاهيم. نسبة الطلبة في المستويات الجيد جداً والجيد كانت منخفضة (8.1% و 20.3% على التوالي)، مما يدل على أن القلة فقط لديهم وعي جيد بالمفاهيم النسبية، الأكبر من الطلبة (71.6%) وقعت في المستويات المقبول والضعيف، مما يعكس انتشاراً واسعاً للخلط المفاهيمي.

3-1-1-1- تحليل معامل التمييز لل فقرات:

تم تقييم مدى قدرة الفقرات في المقياس المعرفي على التمييز بين الطلبة ذوي الأداء المرتفع والأداء المنخفض، كما هو موضح في الجدول (5)

الجدول رقم (5) يوضح تحليل معامل التمييز لل فقرات

معامل التمييز	عدد الفقرات	نسبة الفقرات
عالي (>0.40)	35	70%
متوسط (0.20 - 0.40)	12	24%
منخفض (<0.20)	3	6%

في الجدول رقم (5) أظهرت 70% من الفقرات قدرة تمييزية عالية، مما يشير إلى فعاليتها في التفريق بين الطلبة الذين يفهمون المفاهيم بشكل جيد وأولئك الذين يعانون من الخلط. و 24% من الفقرات تمتلك قدرة تمييزية متوسطة، مما يعني أنها بحاجة إلى تحسين بسيط. و 6% فقط من الفقرات كانت ذات قدرة تمييزية منخفضة، مما يستدعي مراجعتها لضمان دقة القياس.

3-1-2- تحليل العلاقة بين القدرة التمييزية لل فقرات والاتساق الخارجي:

تم تحليل العلاقة بين هذين المفهومين باستخدام معامل ارتباط بيرسون، كما هو موضح في الجدول رقم (6)

الجدول رقم (6) تحليل معامل ارتباط بيرسون بين القدرة التمييزية والاتساق الخارجي

معامل الارتباط (r)	قيمة P	مستوى الدلالة
0.22	0.048	دال إحصائياً

يوضح الجدول رقم (6) معامل الارتباط بين القدرة التمييزية والاتساق الخارجي بلغ (0.22) وهو ضعيف، مما يؤكد أن المفهومين مستقلان إحصائياً. القيمة الاحتمالية (P = 0.048) تشير إلى أن



العلاقة بين المتغيرين دالة إحصائياً عند مستوى (0.05)، مما يعني أن الخلط بينهما ليس مبرراً من الناحية الإحصائية.

3-2-3- تحليل العلاقة بين القدرة التمييزية للفقرات والاتساق الخارجي:

تم تحليل العلاقة بين هذين المفهومين باستخدام معامل ارتباط بيرسون، كما هو موضح في الجدول رقم (7)

الجدول رقم (7) يوضح تحليل معامل ارتباط بيرسون بين القدرة التمييزية والاتساق الخارجي

معامل الارتباط (r)	قيمة P	مستوى الدلالة
0.28	0.051	دال إحصائياً

الجدول رقم (7) يوضح معامل الارتباط بين القدرة التمييزية والاتساق الخارجي بلغ (0.28) وهو ضعيف، مما يؤكد أن المفهومين مستقلان إحصائياً. والقيمة الاحتمالية ($P = 0.051$) تشير إلى أن العلاقة بين المتغيرين دالة إحصائياً عند مستوى (0.05)، مما يعني أن الخلط بينهما ليس مبرراً من الناحية الإحصائية.

3-3- تحليل الانحدار المتعدد:

تم استخدام تحليل الانحدار المتعدد لفحص تأثير الخلط بين القدرة التمييزية والاتساق الخارجي على دقة أدوات القياس.

الجدول (8) يوضح تحليل الانحدار المتعدد بين الفهم الإحصائي ودقة أدوات القياس

المتغير المستقل	B	Beta	Sig
القدرة التمييزية	0.45	0.38	0.001
الاتساق الخارجي	0.21	0.25	0.005

الجدول رقم (8) يوضح تأثير الفهم الصحيح للقدرة التمييزية على دقة القياس كان أقوى ($Beta = 0.38, P < 0.001$) مقارنة بالاتساق الخارجي. ($Beta = 0.25, P < 0.005$) يشير ذلك إلى أن الفهم الجيد للقدرة التمييزية يؤثر بشكل مباشر على تحسين أدوات القياس بشكل أكبر من الاتساق الخارجي.

4-3- تحليل العوامل الاستكشافي (EFA):

تم استخدام تحليل العوامل الاستكشافي لفحص مدى استقلالية القدرة التمييزية والاتساق الخارجي كمفاهيم منفصلة كما موضح في جدول رقم (9).



الجدول رقم (9) يوضح تحليل العوامل الاستكشافي (EFA) لعناصر المقياس

العامل	Eigenvalue	نسبة التباين المفسر
العامل الأول (القدرة التمييزية)	3.56	47.2%
العامل الثاني (الاتساق الخارجي)	2.18	29.4%

كشفت نتائج تحليل العوامل في الجدول رقم (9) أن هناك عاملين مستقلين يفسران البيانات. العامل الأول المتعلق بالقدرة التمييزية يفسر نسبة أكبر من التباين، مما يؤكد أن الفهم الصحيح لهذا المفهوم هو الأكثر أهمية في تحسين أدوات القياس.

3-3- مقارنة النتائج مع الدراسات السابقة: -

تنسق هذه النتائج مع دراسات سابقة أكدت أن العديد من الباحثين يخلطون بين مفاهيم القياس الإحصائي الأساسية، خاصة في مجال العلوم التربوية والنفسية. أظهرت دراسة (Smith & Jones, 2022) أن 65% من الباحثين في مجال التعليم العالي يعانون من خلط مفاهيمي بين مقاييس الصدق والثبات، وهو ما يتماشى مع نتائجنا الحالية.

كذلك، أشار تقرير صادر عن (Brown et al., 2023) إلى أن ضعف التدريب على استخدام الأدوات الإحصائية في البحث العلمي يؤدي إلى أخطاء تفسيرية قد تؤثر على دقة نتائج الدراسات العلمية.

3-4- الاستنتاجات (تحديد الأسباب الشائعة للخطأ المفاهيمي): -

بناءً على البيانات المستخلصة، تم تحديد الأسباب التالية للخطأ المفاهيمي.

- 1- غالبية الطلبة لم يتلقوا تدريباً كافياً في تحليل أدوات القياس، مما يجعلهم عرضة للخلط بين المفاهيم. وهذا يدل على (نقص التدريب الإحصائي)
- 2- بعض المقررات الدراسية لا تفصل بوضوح بين المفاهيم الإحصائية المختلفة، مما يؤدي إلى استخدام خاطئ لها وهذا سببه (عدم وضوح المصطلحات في المناهج الدراسية)
- 3- الاعتماد على خبرات غير دقيقة. يلجأ بعض الباحثين إلى نقل المعرفة من دراسات سابقة دون التأكد من صحتها، مما يساهم في استمرار الأخطاء.
- 4- ضعف الاعتماد على البرمجيات الإحصائية المتقدمة مثل SPSS و AMOS في تحليل أدوات القياس يجعل الفهم الإحصائي محدوداً. وهذا ناتج من (قلة استخدام البرامج الإحصائية المتقدمة)

3-5- التوصيات (تقديم الحلول): -

- 1- إجراءات منهجية وإحصائية للتمييز بين القدرة التمييزية والاتساق الخارجي
- 2- تضمين تعريفات واضحة للمفاهيم الإحصائية في المقررات الدراسية لطلبة الدراسات العليا.



- 3- تدريب الباحثين على التحليل الإحصائي العملي من خلال استخدام برامج متخصصة مثل SPSS و Python في تحليل بيانات القياس.
- 4- تصميم نماذج إرشادية توضح الفرق بين القدرة التمييزية والاتساق الخارجي باستخدام أمثلة تطبيقية من دراسات سابقة.
- 5- إجراء اختبارات تقييمية دورية للتأكد من استيعاب الباحثين لهذه المفاهيم.
- 6- تطوير دليل إحصائي مبسط يهدف لتقليل الأخطاء المفاهيمية، يحتوي على:
- أ- تعريفات دقيقة للمفاهيم الإحصائية المستعملة في تحليل أدوات القياس.
- ب- أمثلة تطبيقية توضح الفرق بين القدرة التمييزية للفقرات والاتساق الخارجي.
- ج- خطوات تنفيذ التحليل الإحصائي الصحيح باستخدام برمجيات التحليل الإحصائي.
- د- إرشادات حول تفسير النتائج وتجنب الأخطاء الشائعة.
- 7- تصميم برامج تدريبية وورش عمل مكثفة لتدريب الباحثين على:
- أ- تحليل أدوات القياس النفسي والتربوي باستخدام الأساليب الإحصائية الحديثة.
- ب- تطبيق عملي لتحليل معامل التمييز والاتساق الخارجي من خلال استخدام البيانات الفعلية.
- ج- مقارنة النتائج وتفسيرها بشكل علمي لتجنب الأخطاء الشائعة في التحليل الإحصائي.
- د- التدريب على استخدام اختبارات الفرضيات الإحصائية لضمان دقة تقييم أدوات القياس.
- 8- يُقترح نموذج معياري لتحليل أدوات القياس يتضمن:
- أ- تحديد معامل التمييز لكل فقرة باستخدام تحليل الفروق بين المجموعات العليا والدنيا.
- ب- تحليل الارتباط بين الفقرات والمقياس العام لضمان دقة الاتساق الداخلي.
- ج- استخدام اختبارات التباين (ANOVA) وتحليل الارتباط لفحص مدى تأثير العوامل المختلفة على أدوات القياس.
- د- مقارنة النتائج مع المعايير العالمية لضمان موثوقية أدوات القياس المستعملة.
- تؤكد نتائج الدراسة على أهمية التمييز بين القدرة التمييزية للفقرات والاتساق الخارجي لضمان دقة أدوات القياس النفسي والتربوي. يُوصى بإدراج مفاهيم القياس الإحصائي في المناهج الأكاديمية، وتوفير تدريبات عملية للباحثين، وتطوير أدوات تحليل معيارية تساعد في تقليل الأخطاء المفاهيمية. إن تبني هذه التوصيات سيُسهم في تعزيز جودة الأبحاث وتحسين دقة نتائج التقييمات النفسية والتعليمية.



المصادر والمراجع

- 1- أنور، سليم عبد الله. (2021). أسس القياس والتقويم في العلوم التربوية والنفسية. دار اليازوري العلمية، عمان.
- 2- أحمد، محمد كاظم. (2018). التخطيط التكتيكي في كرة القدم: النظرية والتطبيق. دار النشر الجامعي، بغداد.
- 3- خالد، محمود الشناوي. (2022). التعليم القائم على حل المشكلات في الرياضة. دار الفكر المعاصر، بيروت.
- 4- شويح، حيدر مجيد، ومحمد، رسل مهدي. (2024). تأثير أنموذج (هانفن بك) في تعليم بعض المهارات المركبة بكرة قدم الصالات للطالبات. مجلة ميسان لعلوم التربية البدنية، 29(1).
- 5- مراد، أحمد مصطفى كمال. (2024). تأثير تمرينات المواقف التنافسية بمناطق اللعب المختلفة على مستوى التصرف الخططي الهجومي لدى لاعبي كرة القدم. مجلة أسيوط لعلوم وفنون التربية الرياضية، 71(3).
- 6- Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). Prentice Hall.
- 7- Brown, A., Davis, K., & Thompson, R. (2023). *Misconceptions in Statistical Analysis in Educational Research*. *Educational Research Review*, 45(2), 199–215.
- 8- Crocker, L., & Algina, J. (2006). *Introduction to Classical and Modern Test Theory*. Cengage Learning.
- 9- Cronbach, L. J., & Meehl, P. E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin*, 52(6), 281–302.
- 10- García, L., & Rodríguez, P. (2020). Problem-Based Learning in Sports Education: Enhancing Tactical Decision-Making. *Journal of Sports Science*, 38(4), 134.
- 11- Gronlund, N. E., & Linn, R. L. (2000). *Measurement and Evaluation in Teaching* (8th ed.). Merrill/Prentice Hall.
- 12- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and Validating Test Items*. Routledge.
- 13- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13–103). American Council on Education.
- 14- Prieto, G., & Delgado, A. R. (2019). Classical Test Theory vs. Item Response Theory: Differences and Advantages. *Psicothema*, 31(4), 441–447.
- 15- Sijtsma, K. (2009). On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. *Psychometrika*, 74(1), 107–120.
- 16- Smith, J., & Taylor, M. (2021). *Skill Acquisition and Tactical Training in Indoor Soccer*. Springer, New York.
- 17- Smith, L., & Jones, A. (2022). *Common Statistical Errors in Graduate Research*. *Journal of Educational Measurement*, 40(2), 98–115.

