# تعزيز صدق المحتوى والصدق البنائي بالاستعانة بالمحكمين لأدوات القياس النفسي التربوي (رؤية مستقبلية)

إعداد

أ. د/ مُحَمَّدٌ مُحَمَّدٌ فَتْحْ اللهُ سيد أحمد

أُسْتَاذُ القِيَاسِ وَالتَّقْوِيمِ وَالإِحْصَاءِ النَّفْسِيُّ وَالتَّرْبَوِيُّ

المَرْكَزِ القَوْمِيّ لِلِامْتِحَانَاتِ وَالتَّقْوِيمِ التَّرْبَوِيّ

## مستخلص

تتناول هذه الدراسة الدور الحاسم الذي يلعبه المحكمون المتخصصون في تحسين جودة أدوات القياس النفسي والتربوي من خلال تعزيز صدق المحتوى والصدق البنائي. يتمحور صدق المحتوى حول شمولية الأداة في تغطية جميع جوانب المفهوم أو السلوك المراد قياسه، مما يضمن تمثيل كافة الأبعاد الأساسية للمفهوم. من ناحية أخرى، يركز الصدق البنائي على قدرة الأداة في قياس البناء النظري أو الهيكل المفاهيمي للمفهوم قيد الدراسة، مما يعكس مدى توافق البنود مع الأبعاد النظرية للمفهوم.

في هذه الدراسة، تم التركيز على كيفية تحسين صدق المحتوى والصدق البنائي عبر الاستعانة بالمحكمين المتخصصين، الذين يملكون خبرات نظرية وتطبيقية في مجال القياس النفسي والتربوي. يقوم المحكمون بتحليل البنود من حيث مدى ارتباطها بالمحتوى المستهدف وشموليتها لكل جوانب المفهوم. يُعتبر هذا التحليل ضروريًا لضمان أن كل بند في أداة القياس يسهم في تمثيل جانب من جوانب المفهوم قيد الدراسة. يشمل ذلك تقييم مدى دقة البنود وتوافقها مع المعايير النظرية والعملية المعمول بها.

من بين الأساليب التي تمت مناقشتها في هذه الدراسة لتعزيز صدق المحتوى، تم الاعتماد على مؤشرات كمية مثل مؤشر صدق المحتوى (Content Validity Index) (CVI) ونسبة صدق المحتوى (Content Validity Ratio) (CVR)، والتي تستخدم لقياس مدى اتفاق المحكمين حول مدى ملاءمة البنود وشموليتها. توضح الورقة البحثية أن هذه المؤشرات توفر أدوات كمية موضوعية لتقييم مدى فعالية البنود في تمثيل المفهوم قيد القياس، كما تساهم في تحسين جودة التقييمات من خلال تقليل التأثيرات الذاتية للمحكمين.

بالإضافة إلى ذلك، يتناول البحث كيفية تحسين الصدق البنائي عبر المحكمين المتخصصين، الذين يقومون بتقييم مدى توافق البنود مع البناء النظري. يشمل هذا

التقييم تحليل العلاقة بين البنود والمفاهيم النظرية الأساسية التي تستند إليها الأداة، وهو ما يسهم في ضمان أن الأداة قادرة على قياس الجوانب الأساسية للمفهوم النظري الذي تسعى لدراسته.

كما تُسلط الورقة البحثية الضوء على التحديات التي قد تنشأ عند استخدام المحكمين، مثل الاختلافات في الخلفيات الأكاديمية والتحيزات الشخصية. ولهذا، تقترح الورقة البحثية أن استخدام التكنولوجيا الحديثة، مثل الذكاء الاصطناعي، يمكن أن يُحسن دقة وموضوعية عملية التحكيم، من خلال تقديم تحليل شامل للبنود وتقليل التأثيرات الذاتية للمحكمين. تشير الورقة البحثية إلى أن الذكاء الاصطناعي يمكن أن يُساعد في تقديم توصيات مبنية على تحليل البيانات الكبيرة، مما يسهم في زيادة دقة وموثوقية الأداة.

في النهاية، تقدم الورقة البحثية رؤية مستقبلية لدمج الأساليب التقليدية للتحكيم مع التكنولوجيا الحديثة، حيث يمكن أن تلعب المنصات الرقمية والذكاء الاصطناعي دورًا فعالًا في تحسين جودة أدوات القياس النفسي والتربوي. يتمثل الهدف في تطوير أدوات قياس تتمتع بمستويات عالية من الدقة والموثوقية، مما يعزز مصداقية النتائج ويُسهم في تطوير البحث العلمي في المجالات النفسية والتربوية.

**الكلمات المفتاحية:** صدق المحتوى، الصدق البنائي، المحكمين المتخصصين، الذكاء الاصطناعي، أدوات القياس النفسي، التحكيم.

# Enhancing Content and Construct Validity through Expert Judges for Educational and Psychological Measurement Tools: A Future Vision

**Prof. Mohamed Mohamed Fathallah Sayed Ahmed**

Professor of Psychological and Educational Measurement, Evaluation and Statistics
National Center for Examinations and Educational Evaluation

## Abstract

This study focuses on the crucial role of expert judges in improving the quality of educational and psychological measurement tools by enhancing both content and construct validity. Content validity refers to how comprehensively the measurement tool covers all aspects of the concept or behavior it aims to measure, ensuring that all essential dimensions are represented. Construct validity, on the other hand, examines how well the tool aligns with the theoretical framework underlying the concept being studied, indicating the extent to which the items correspond to the theoretical dimensions.

The study emphasizes how both types of validity can be enhanced through the involvement of expert judges, who have theoretical and practical expertise in educational and psychological measurement. These experts are tasked with evaluating the items based on their relevance and comprehensiveness, ensuring that each item in the measurement tool contributes to representing the concept being measured. This includes evaluating the clarity, theoretical relevance, and practical alignment of the items.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

To strengthen content validity, the study highlights the use of quantitative measures such as the Content Validity Index (CVI) and Content Validity Ratio (CVR), which assess the degree of agreement among expert judges regarding the adequacy and comprehensiveness of the items. These measures provide objective, quantitative tools for evaluating how effectively the items represent the concept under investigation, thereby enhancing the accuracy and reliability of the assessment.

Furthermore, the study addresses how expert judges contribute to improving construct validity by evaluating the extent to which the items correspond to the theoretical structure of the concept. This evaluation includes analyzing the relationship between the items and the underlying theoretical constructs, ensuring that the tool accurately measures the theoretical aspects of the concept.

The study also explores the challenges associated with using expert judges, such as differences in academic backgrounds and personal biases. To address these challenges, the study proposes incorporating modern technology, particularly artificial intelligence (AI), to enhance the accuracy and objectivity of the judging process. AI can analyze large datasets and provide data-driven recommendations, reducing human biases and improving the overall precision of the measurement tool.

In conclusion, the study presents a future vision for integrating traditional expert judging methods with modern technology, where digital platforms and AI can play a critical role in enhancing the quality of educational and psychological measurement tools. The goal is to develop tools with high levels of accuracy and reliability, contributing to the credibility of research findings and advancing the field of educational and psychological research.

**Keywords:** Content validity, construct validity, expert judges, artificial intelligence, psychological measurement tools, judgment process.

## مقدمة:

يُعدّ صدق المحتوى واحدًا من أهم المفاهيم المرتبطة بجودة أدوات القياس في مجال القياس النفسي والتربوي، حيث يمثل مقياسًا لمدى شمولية ودقة الأداة في تغطية جميع جوانب الظاهرة المستهدفة.

حيث يُعد صدق أدوات القياس النفسي والتربوي من العناصر الحيوية التي تضمن دقة البيانات وموثوقية النتائج التي يتم التوصل إليها. يركز صدق الأداة على قدرتها على قياس ما تهدف إلى قياسه بدقة، وهو ما يعزز من ثقة الباحثين والممارسين في استخدام هذه الأدوات لتحقيق أهدافهم التعليمية والنفسية (Messick, 1995). وفي هذا السياق، يُعد كل من صدق المحتوى والصدق البنائي من أبرز أنواع الصدق التي تحظى بأهمية كبيرة في مجال تطوير أدوات القياس. صدق المحتوى يشير إلى مدى شمولية البنود المستخدمة في أداة القياس لتمثيل جميع جوانب المفهوم المراد قياسه، بينما يُعنى الصدق البنائي بمدى تمثيل هذه البنود للبنية النظرية الكامنة وراء المفهوم (Cronbach & Meehl, 1955).

يعتبر الاعتماد على المحكمين المتخصصين أحد الأساليب الأساسية في تحسين جودة أدوات القياس وضمان تحقيق صدق المحتوى والصدق البنائي. يقوم المحكمون بتقديم تقييمات شاملة للبنود من حيث شموليتها وارتباطها بالمحتوى المستهدف ومدى تمثيلها للبنية النظرية للمفهوم. هذا النهج يُسهم في تحسين الأداة من خلال تعديلات مستندة إلى ملاحظات المحكمين، مما يزيد من موثوقية وفاعلية الأداة (Lynn, 2006).

صدق المحتوى والصدق البنائي :يُعرّف صدق المحتوى على أنه مدى قدرة أداة القياس على تغطية جميع جوانب المحتوى المرتبط بالمفهوم المراد قياسه، مما يضمن شمولية البنود وتمثيلها لكافة الأبعاد الأساسية للمفهوم (Nunnally & Bernstein, 1994). ومن خلال هذا التقييم، يسعى الباحثون إلى ضمان أن كل بند يسهم في تغطية جزء جوهري من المحتوى، دون إهمال أي جانب مهم.

في المقابل، يُعنى الصدق البنائي بمدى قدرة الأداة على قياس البناء النظري الكامن خلف المفهوم المستهدف. يشيرا (Cronbach & Meehl (1955 إلى أن الصدق البنائي يتطلب ربط نتائج القياس بالمفاهيم النظرية، مما يسهم في تقديم نتائج دقيقة تُعزز فهم الظاهرة المدروسة.

دور المحكمين في تحسين الصدق: يلعب المحكمون المتخصصون دورًا حاسمًا في تعزيز صدق أدوات القياس. يتطلب هذا الأمر الاستعانة بخبراء قادرين على تقييم البنود وتحليلها بناءً على مدى ارتباطها بالمحتوى النظري والعملي. تُظهر الأبحاث أن استناد عملية التحكيم إلى أسس علمية متينة تُسهم في تقديم أدوات

قياس أكثر دقة وموضوعية، حيث يعمل المحكمون على تعديل البنود لتكون أكثر ارتباطًا بالهدف المراد قياسه (Polit & Beck, 2006). من هنا، تُعتبر الاستعانة بالمحكمين وسيلة فعالة لتحسين الأداة وتقليل التحيزات التي قد تؤثر على النتائج.

التحديات المستقبلية واستخدام التكنولوجيا :مع التطور التكنولوجي، أصبحت هناك توجهات متزايدة نحو استخدام تقنيات الذكاء الاصطناعي والمنصات الإلكترونية لتحسين عملية التحكيم وزيادة دقتها. تشير دراسة (Jones et al., 2020) إلى أن الذكاء الاصطناعي يمكن أن يسهم في تحسين دقة التقييمات من خلال تقديم تحليل شامل للبنود وتقليل التحيزات البشرية. كما تُبرز دراسة (Lee & Kim, 2021) أهمية استخدام المنصات الرقمية التي تتيح للمحكمين التعاون عن بعد، مما يسهم في تحسين الاتساق في التقييمات.

في ظل هذه التطورات، تهدف هذه الورقة البحثية إلى استعراض أهمية الاستعانة بالمحكمين لتحسين صدق المحتوى والصدق البنائي في أدوات القياس النفسي والتربوي، مع تقديم رؤية مستقبلية تعتمد على دمج الأساليب التقليدية مع التكنولوجيا الحديثة. ستتطرق الورقة البحثية إلى كيفية استخدام الذكاء الاصطناعي والتقنيات الحديثة لتعزيز دقة التحكيم وتطوير معايير جديدة لتحسين جودة أدوات القياس في المستقبل.

## المشكلة البحثية:

تعتبر أدوات القياس النفسي والتربوي أحد الأسس الحاسمة في تطوير البحوث والاختبارات التعليمية والنفسية. تعتمد جودة هذه الأدوات على مدى قدرتها على تقديم بيانات دقيقة وموثوقة تُعبر بشكل كامل عن المفاهيم المراد قياسها. يتجلى ذلك من خلال نوعين رئيسيين من الصدق: صدق المحتوى، الذي يشير إلى مدى شمولية البنود المستخدمة في القياس للمفهوم المستهدف (Nunnally & Bernstein, 1994)، والصدق البنائي، الذي يعنى بمدى تمثيل البنود للبنية النظرية الكامنة وراء المفهوم (Cronbach & Meehl, 1955). ومع ذلك، تظل هناك تحديات كبيرة في تحسين هذين النوعين من الصدق، مما يؤثر على دقة النتائج وقابليتها للتعميم.

تواجه أدوات القياس النفسي والتربوي مشكلات متعلقة بالتأكد من شمولية المحتوى وقدرتها على التماشي مع المفاهيم النظرية المستهدفة. من أبرز هذه المشكلات الاعتماد على التقييمات الذاتية لبنود القياس دون الحصول على مراجعات دقيقة من خبراء في المجال (Lynn, 2006). يؤدي هذا الأمر إلى ضعف في شمولية البنود وعدم دقتها، مما يؤثر سلبًا على جودة النتائج المترتبة على استخدام هذه الأدوات. إضافة إلى ذلك، فإن التحيزات الشخصية للمحكمين قد

تؤدي إلى نتائج غير موضوعية، وهو ما يمثل تحديًا آخر للباحثين عند تطوير أدوات القياس (Polit & Beck, 2006).

تُظهر الدراسات السابقة أن التحكيم من قبل خبراء في المجال يعد من أهم الأساليب التي تساهم في تعزيز صدق أدوات القياس (Jones et al., 2020). ومع ذلك، تواجه عملية التحكيم تحديات عديدة، منها الاختلافات في الخلفيات الأكاديمية للمحكمين، وتأثير التحيزات الشخصية على تقييماتهم. كما أن الاعتماد الكبير على الأساليب التقليدية في التحكيم يجعل من الصعب تحسين دقة التقييمات وضمان موضوعيتها (Zamanzadeh et al., 2015).

من هنا، تنبع مشكلة الورقة البحثية من الحاجة الملحة إلى تطوير منهجيات أكثر فاعلية لتعزيز صدق المحتوى والصدق البنائي في أدوات القياس النفسي والتربوي، وذلك من خلال الاستعانة بالمحكمين المتخصصين. بالإضافة إلى ذلك، يبرز التطور التكنولوجي، بما في ذلك استخدام الذكاء الاصطناعي والمنصات الإلكترونية، كأداة واعدة لتحسين عملية التحكيم وزيادة دقتها وموضوعيتها (Jones et al., 2020). على الرغم من هذه الإمكانيات، إلا أن استخدام التكنولوجيا في مجال التحكيم لا يزال محدودًا ويحتاج إلى مزيد من البحث والتطوير لتقديم حلول عملية ومستدامة.

**السؤال الرئيس لمشكلة الدراسة :**كيف يمكن تحسين صدق المحتوى والصدق البنائي في أدوات القياس النفسي والتربوي من خلال الاستعانة بالمحكمين المتخصصين، وما هو الدور الذي يمكن أن تلعبه التكنولوجيا الحديثة في تحسين دقة وموضوعية عملية التحكيم؟

## الأهداف الفرعية للمشكلة البحثية:

1. تحليل مدى فاعلية الاستعانة بالمحكمين المتخصصين في تحسين صدق المحتوى والصدق البنائي.
2. استكشاف التحديات التي تواجه عملية التحكيم وكيف يمكن التغلب عليها.
3. تقديم رؤية مستقبلية لكيفية توظيف التكنولوجيا، مثل الذكاء الاصطناعي، لتحسين دقة وموضوعية عملية التحكيم.

## تساؤلات بحثية:

تسعى هذه الورقة البحثية إلى معالجة مشكلة تعزيز صدق المحتوى والصدق البنائي في أدوات القياس النفسي والتربوي من خلال الاستعانة بالمحكمين المتخصصين، واستكشاف الدور الذي يمكن أن تلعبه التكنولوجيا الحديثة في

تحسين دقة وموضوعية عملية التحكيم. بناءً على ذلك، يتم صياغة التساؤلات البحثية التالية لتوجيه الدراسة وتحقيق أهدافها:

1. **التساؤل الأول :**ما مدى فعالية الاستعانة بالمحكمين المتخصصين في تحسين صدق المحتوى في أدوات القياس النفسي والتربوي؟

يعكس هذا التساؤل الحاجة إلى تقييم دور المحكمين في تحسين مدى شمولية البنود المستخدمة في أدوات القياس لتغطية كافة الجوانب المرتبطة بالمفهوم المدروس. سيُركز على تحليل مدى اتفاق المحكمين حول صلاحية البنود باستخدام معايير الصدق الكمي مثل (Content Validity Index) (CVI)، و (Content Validity Ratio) (CVR) (Polit & Beck, 2006).

2. **التساؤل الثاني :**كيف يمكن تعزيز الصدق البنائي في أدوات القياس النفسي والتربوي من خلال الاستعانة بالمحكمين المتخصصين؟

يهدف هذا التساؤل إلى استكشاف دور المحكمين في تحسين الصدق البنائي لأدوات القياس من خلال تقييم مدى تمثيل البنود للبنية النظرية الكامنة وراء المفهوم. يتطلب ذلك تقييم توافق البنود مع النظرية المستخدمة لقياس المفهوم وتحديد مدى قدرتها على التنبؤ بالسلوكيات والمواقف المرتبطة به (Cronbach & Meehl, 1955).

3. **التساؤل الثالث :**ما التحديات التي تواجه عملية التحكيم باستخدام المحكمين المتخصصين في تعزيز صدق المحتوى والصدق البنائي؟

يناقش هذا التساؤل التحديات التي تواجه الباحثين عند الاستعانة بالمحكمين، بما في ذلك التباينات في تقييمات المحكمين نتيجة لاختلاف خلفياتهم الأكاديمية والمهنية. كما سيتم تناول مشكلة التحيزات الشخصية في التقييمات وطرق معالجتها (Lynn, 2006; Zamanzadeh et al., 2015).

4. **التساؤل الرابع :**ما الدور الذي يمكن أن تلعبه التكنولوجيا الحديثة، مثل الذكاء الاصطناعي والمنصات الرقمية، في تحسين عملية التحكيم وزيادة دقتها وموضوعيتها؟

يركز هذا التساؤل على كيفية الاستفادة من التكنولوجيا الحديثة في تحسين دقة التقييمات وتقليل التحيزات البشرية التي قد تؤثر على نتائج التحكيم. سيتم

استعراض التقنيات المستخدمة حاليًا مثل الذكاء الاصطناعي وتقييم مدى فعاليتها في تحسين كفاءة التحكيم (Jones et al., 2020).

5. **التساؤل الخامس :**كيف يمكن تحقيق التوازن بين التقييمات الكمية والكيفية في عملية التحكيم لتعزيز صدق أدوات القياس النفسي والتربوي؟

يسعى هذا التساؤل إلى استكشاف طرق دمج التحليلات الكمية، مثل مؤشرات صدق المحتوى، مع التقييمات الكيفية التي يقدمها المحكمون لتحسين شمولية البنود ومواءمتها مع الأهداف النظرية للمقياس. سيُركز على تحقيق التوازن بين الأساليب المختلفة لضمان نتائج دقيقة وموثوقة (Nunnally & Bernstein, 1994).

6. **التساؤل السادس :**ما هي المعايير المثلى لتحديد عدد المحكمين المطلوب لتحقيق دقة وموضوعية في عملية التحكيم؟

يستكشف هذا التساؤل العدد الأمثل للمحكمين اللازم لضمان دقة وموضوعية التقييمات، مع مراعاة التكاليف الزمنية والمالية المرتبطة بزيادة عدد المحكمين. يتم استعراض الدراسات التي تناولت تأثير عدد المحكمين على دقة التقييمات وكيفية تحديد العدد الأمثل (Zamanzadeh et al., 2015).

**الأهمية:**

تنبع أهمية هذه الورقة البحثية من الحاجة الملحة إلى تحسين جودة أدوات القياس النفسي والتربوي، خاصة فيما يتعلق بصدق المحتوى والصدق البنائي. تُعد أدوات القياس أحد الأركان الأساسية في البحوث النفسية والتربوية، حيث تعتمد دقة النتائج وموثوقيتها على مدى جودة الأدوات المستخدمة (Nunnally& Bernstein, 1994). ومع تزايد التوجهات نحو الاعتماد على البيانات الكمية والكيفيّة لاتخاذ القرارات في مجالات التعليم والصحة النفسية، يصبح تطوير أدوات قياس دقيقة وموثوقة أمرًا حاسمًا لضمان أن النتائج قابلة للتطبيق والتعميم.

1. **تحسين جودة أدوات القياس النفسي والتربوي :**تُسهم هذه الورقة البحثية في تحسين جودة أدوات القياس النفسي والتربوي من خلال التركيز على تعزيز صدق المحتوى والصدق البنائي. يُعتبر صدق المحتوى مؤشرًا على مدى شمولية الأداة للمفاهيم المستهدفة، بينما يعكس الصدق البنائي قدرة الأداة على قياس البناء النظري الكامن وراء المفهوم. تحسين هذين النوعين من الصدق يزيد من دقة القياس وموثوقية النتائج التي يمكن الاعتماد عليها في

اتخاذ القرارات المتعلقة بالتعليم والصحة النفسية (Cronbach & Meehl, 1955).

2. **تعزيز دور المحكمين في تحسين أدوات القياس:** تلقي الورقة البحثية الضوء على أهمية الاستعانة بالمحكمين المتخصصين في تعزيز دقة أدوات القياس. يعتمد صدق المحتوى والصدق البنائي بشكل كبير على خبرة المحكمين في تقييم البنود ومدى ارتباطها بالمفاهيم المستهدفة. من خلال تقديم ملاحظاتهم وتحليلاتهم، يمكن للمحكمين تحسين جودة الأداة وضمان أن البنود تغطي جميع الجوانب المتعلقة بالمفهوم النظري (Lynn, 2006)، لذلك تُسهم هذه الورقة البحثية في تطوير آليات فعالة لاستغلال خبرة المحكمين بشكل منهجي وموضوعي.
3. **مواجهة التحديات المتعلقة بالتحكيم:** تُعالج الورقة البحثية التحديات التي تواجه الباحثين في استخدام التحكيم لتحسين صدق أدوات القياس، مثل التحيزات الشخصية واختلاف الخلفيات الأكاديمية للمحكمين. كما تُقدّم حلولاً مبتكرة للتغلب على هذه التحديات، مثل: استخدام الأدوات الكمية لتكميل التقييمات الكيفية التي يقدمها المحكمون. إضافةً إلى ذلك، تُسهم الورقة البحثية في وضع معايير محددة لتقييم صدق المحتوى والصدق البنائي، مما يساعد في تقليل التباينات بين تقييمات المحكمين وزيادة الاتساق في النتائج (Polit & Beck, 2006).
4. **استخدام التكنولوجيا في تحسين عملية التحكيم:** تُعتبر هذه الورقة البحثية رائدة في تقديم رؤية مستقبلية تعتمد على استخدام التكنولوجيا الحديثة لتحسين عملية التحكيم. يُمكن لتقنيات الذكاء الاصطناعي والمنصات الرقمية أن تسهم في تحسين دقة وموضوعية التقييمات، من خلال تحليل البيانات وتقليل الأخطاء البشرية. تُظهر الأبحاث الحديثة أن استخدام الذكاء الاصطناعي في تقييم صدق المحتوى يمكن أن يقلل من التحيزات ويزيد من دقة التقييمات (Jones et al., 2020). وبالتالي، تُعد هذه الورقة البحثية مساهمة هامة في تطوير أدوات القياس من خلال توظيف التكنولوجيا لتحسين عمليات التحكيم.
5. **المساهمة في تطوير الأبحاث النفسية والتربوية:** إن تحسين صدق أدوات القياس يُسهم في تعزيز مصداقية الأبحاث النفسية والتربوية، وبالتالي تحسين جودة النتائج التي يتم التوصل إليها. بفضل الأدوات المحسنة، يُمكن للباحثين تحقيق نتائج أكثر دقة وموثوقية، مما يزيد من الثقة في استنتاجاتهم ويسهم في تطوير المعرفة العلمية في مجالات علم النفس والتربية. تعتمد هذه الأبحاث على أدوات القياس للتأكد من دقة البيانات المجمعة وتفسيرها بشكل صحيح، وهو ما يبرر الحاجة إلى تطوير أدوات ذات صدق عالٍ (Messick, 1995).

6. **تعزيز القدرة على اتخاذ قرارات قائمة على البيانات :**بفضل أدوات القياس المحسنة، ستتمكن المؤسسات التعليمية والصحية من اتخاذ قرارات أكثر استنارة. تعتمد هذه القرارات على دقة البيانات المجمعة، سواء كانت متعلقة بتقييم الطلاب أو بتقديم الرعاية الصحية النفسية. الأدوات ذات الصدق العالي تتيح جمع بيانات دقيقة وموثوقة، مما يُسهم في تحسين الأداء العام للمؤسسات التعليمية والصحية. وبالتالي، تُسهم الورقة البحثية في تحسين القدرة على اتخاذ القرارات القائمة على البيانات المدعومة علميًا (Nunnally & Bernstein, 1994).

**الأهداف:**

تسعى هذه الورقة البحثية إلى معالجة مجموعة من الأهداف المتعلقة بتحسين جودة أدوات القياس النفسي والتربوي من خلال التركيز على تعزيز صدق المحتوى والصدق البنائي باستخدام المحكمين المتخصصين وتوظيف التكنولوجيا الحديثة. بناءً على أهمية الورقة البحثية والتحديات التي تواجه تطوير أدوات القياس، تم تحديد الأهداف التالية:

1. **تحليل فعالية الاستعانة بالمحكمين في تحسين صدق المحتوى والصدق البنائي لأدوات القياس النفسي والتربوي:** يهدف هذا الهدف إلى دراسة الدور الذي يمكن أن يلعبه المحكمون المتخصصون في تحسين جودة أدوات القياس من خلال تقديم ملاحظات حول مدى شمولية البنود وتوافقها مع الأهداف النظرية للمقياس. سيسهم ذلك في تعزيز فهم الباحثين لكيفية الاستفادة من خبرة المحكمين لتحسين الصدق البنائي وصدق المحتوى (Lynn, 2006).
2. **استكشاف التحديات التي تواجه عملية التحكيم وكيفية التغلب عليها:** يهدف هذا الهدف إلى تحديد التحديات التي يواجهها الباحثون عند الاستعانة بالمحكمين لتقييم أدوات القياس، مثل التحيزات الشخصية واختلاف الخلفيات الأكاديمية والمهنية للمحكمين. كما تسعى الورقة البحثية إلى تقديم حلول عملية للتغلب على هذه التحديات من خلال الجمع بين التقييمات الكمية والكيفية (Polit & Beck, 2006).
3. **تقديم إطار عمل لتحسين صدق المحتوى والصدق البنائي باستخدام التكنولوجيا الحديثة:** يركز هذا الهدف على تطوير إطار عمل يستند إلى التكنولوجيا الحديثة، مثل الذكاء الاصطناعي والمنصات الرقمية، لتحسين دقة وموضوعية عملية التحكيم. يهدف هذا الإطار إلى توفير حلول مبتكرة تقلل من الأخطاء البشرية وتزيد من فعالية التقييمات (Jones et al., 2020).

4. **تعزيز دور التحليل الكمي في تقييم أدوات القياس وتطوير معايير لقياس صدق المحتوى والصدق البنائي:** يسعى هذا الهدف إلى تقديم استراتيجيات لدمج التقييمات الكمية مع التقييمات الكيفية، مثل استخدام مؤشرات مثل (Content Validity Index) (CVI) و (Content Validity Ratio) (CVR) لتحسين دقة النتائج التي يتم التوصل إليها من خلال عملية التحكيم. كما يهدف إلى وضع معايير واضحة وموحدة يمكن استخدامها في مجالات القياس النفسي والتربوي لضمان دقة وموضوعية التقييمات (Zamanzadeh et al., 2015).
5. **تحديد العدد الأمثل من المحكمين لتحقيق دقة وموضوعية في عملية التحكيم:** يهدف هذا الهدف إلى استكشاف التأثير الذي يمكن أن يحدثه عدد المحكمين على دقة وموثوقية التقييمات، واقتراح العدد الأمثل للمحكمين لضمان تحقيق أعلى مستويات الدقة دون تحميل الباحثين عبئًا إضافيًا. سيسهم هذا الهدف في توجيه الباحثين حول كيفية اختيار العدد المناسب من المحكمين وفقًا لطبيعة الدراسة وحجم الأداة (Polit & Beck, 2006).
6. **تقديم توصيات لتطوير أدوات القياس النفسي والتربوي بناءً على نتائج الدراسة:** يتمثل هذا الهدف في تقديم مجموعة من التوصيات التي يمكن أن تساعد الباحثين والممارسين في مجال القياس النفسي والتربوي على تطوير أدوات قياس دقيقة وموثوقة. ستعتمد هذه التوصيات على نتائج الدراسة وستتضمن كيفية تعزيز صدق المحتوى والصدق البنائي باستخدام المحكمين والتكنولوجيا الحديثة، بالإضافة إلى تطوير معايير موحدة يمكن تطبيقها في مختلف المجالات التعليمية والنفسية (Nunnally & Bernstein, 1994).

## المصطلحات الإجرائية:

في هذه الورقة البحثية، يتم استخدام مجموعة من المصطلحات الأساسية التي تساهم في توضيح المفاهيم التي تعتمد عليها الورقة البحثية. سيتم تقديم تعريفات إجرائية لهذه المصطلحات لتحديد كيفية استخدامها ضمن سياق الورقة البحثية، وذلك لضمان الفهم المشترك بين القراء والباحثين.

1. **صدق المحتوى:(Content Validity)** يُعرف صدق المحتوى بأنه مدى قدرة أداة القياس على تغطية جميع الجوانب الأساسية للمفهوم أو السلوك المستهدف. في هذه الورقة البحثية، يُقصد بصدق المحتوى قدرة البنود في أدوات القياس النفسي والتربوي على تمثيل المفاهيم المراد قياسها بشكل شامل ودقيق. سيتم تقييم صدق المحتوى من خلال الاستعانة بالمحكمين المتخصصين الذين يقومون بتحليل وتقييم شمولية البنود ومدى مناسبتها لقياس الظاهرة (Polit & Beck, 2006).

2. **الصدق البنائي:(Construct Validity)** يشير الصدق البنائي إلى مدى توافق البنود المستخدمة في أداة القياس مع البناء النظري الذي يعتمد عليه المفهوم الذي يتم قياسه. في هذه الورقة البحثية، يُقصد بالصدق البنائي قدرة البنود على تمثيل البنية النظرية للمفهوم النفسي أو التربوي، وسيتم تقييمه من خلال المحكمين الذين يقيمون مدى تمثيل البنود للبنية النظرية (Cronbach & Meehl, 1955).
3. **المحكمين المتخصصين:(Expert Judges)** يُعرف المحكمون المتخصصون بأنهم مجموعة من الخبراء في مجال القياس النفسي والتربوي يتم الاستعانة بهم لتقييم أدوات القياس. في هذه الورقة البحثية، يُقصد بالمحكمين المتخصصين الأفراد الذين يمتلكون خبرة أكاديمية ومهنية في مجال القياس النفسي والتربوي ويشاركون في تقييم البنود من حيث صدق المحتوى والصدق البنائي، ويقدمون توصيات حول تحسين الأداة (Lynn, 2006).
4. **الذكاء الاصطناعي:(Artificial Intelligence - AI)** يُقصد بالذكاء الاصطناعي في هذه الورقة البحثية استخدام الأنظمة الحاسوبية والتقنيات الحديثة التي يمكنها محاكاة الذكاء البشري لتحليل البيانات واتخاذ القرارات. سيتم استخدام تقنيات الذكاء الاصطناعي في تحسين عملية التحكيم وتقييم صدق أدوات القياس من خلال تحليل البيانات وتقليل التحيزات البشرية (Jones et al., 2020).
5. **منصة التحكيم الرقمية:(Digital Judging Platform)** تشير منصة التحكيم الرقمية إلى النظم الإلكترونية التي تستخدم لتسهيل عملية التحكيم من خلال تقديم أدوات رقمية للمحكمين لتقييم البنود. في هذه الورقة البحثية، سيتم استخدام منصات رقمية لتمكين المحكمين من التعاون عن بعد وتقديم تقييمات موضوعية لبنود أدوات القياس، مما يزيد من كفاءة العملية ويقلل من التحيزات الشخصية (Lee & Kim, 2021).
6. **مؤشر صدق المحتوى:(Content Validity Index - CVI)** يُعرف مؤشر صدق المحتوى بأنه مقياس كمي يستخدم لقياس مدى اتفاق المحكمين حول صلاحية البنود في أداة القياس. في هذه الورقة البحثية، سيتم استخدام مؤشر CVI لتحديد مدى شمولية البنود ومدى اتفاق المحكمين على أن كل بند يمثل جزءًا أساسيًا من المفهوم المراد قياسه (Polit & Beck, 2006).
7. **نسبة صدق المحتوى:(Content Validity Ratio - CVR)** يُقصد بنسبة صدق المحتوى (CVR) المقياس الكمي الذي يستخدم لتحديد أهمية كل بند في أداة القياس بناءً على تقييم المحكمين. في هذه الورقة البحثية، سيتم استخدام نسبة (CVR) لتحديد ما إذا كان كل بند من البنود

ضروريًا لقياس المفهوم المستهدف ومدى اتفاق المحكمين حول ضرورته (Lawshe, 1975).

8. **التحكيم:(Judging Process)** يُعرف التحكيم بأنه العملية التي يتم فيها تقييم أدوات القياس من قبل خبراء متخصصين لتحليل مدى دقة وملاءمة البنود المستخدمة. في هذه الورقة البحثية، يُقصد بالتحكيم عملية تقييم البنود من حيث صدق المحتوى والصدق البنائي بواسطة المحكمين المتخصصين، وتحليل مدى شموليتها ودقتها (Lynn, 2006).

## أداة الورقة البحثية:

تتكون أداة الورقة البحثية (التحقق من صدقي المحتوي والبنائي بالاستعانة بالمحكمين) من ثلاث مراحل رئيسة، كل منها يعتمد على تقنيات مختلفة لتحقيق الأهداف المرجوة من الدراسة، وهي مصممة لتحسين صدق المحتوى والصدق البنائي باستخدام المحكمين المتخصصين وتطبيق التكنولوجيا الحديثة.

## المرحلة الأولى: تقييم أبعاد أو مكونات المقياس باستخدام المحكمين المتخصصين (صدق المحتوي)

في هذه المرحلة، تم إعداد قائمة الأبعاد التي سيتم استخدامها في أدوات القياس النفسي والتربوي. تم توزيع هذه الأبعاد على مجموعة من المحكمين المتخصصين الذين لديهم خبرة واسعة في المجال النفسي والتربوي. وقد طُلب منهم تقييم هذه الابعاد من حيث مدى ملاءمتها للمفاهيم النظرية ومدى شموليتها لتغطية مختلف جوانب المفاهيم المدروسة.

- **أداة التقييم المستخدمة** :استبيان تم تصميمه خصيصًا لهذه المرحلة يحتوي على عدة محاور للتقييم:
  1. مدى توافق الأبعاد مع المفاهيم النظرية.
  2. شمولية الأبعاد ومدى تمثيلها لكافة الأبعاد.
  3. وضوح الصياغة ومدى سهولة الفهم.
- **التحليل الكمي** :يتم استخدام مؤشر صدق المحتوى (Content Validity Index - CVI)، ونسبة صدق المحتوى (Content Validity Ratio) (CVR) لقياس مدى اتفاق المحكمين حول صلاحية الأبعاد.

**المرحلة الثانية: التحقق من دقة إعداد وصياغة مفردات المقياس (صدق المحتوي)**

في هذه المرحلة، تم إعداد قائمة البنود (المفردات الاختبارية أو العبارات) التي سيتم استخدامها في أدوات القياس النفسي والتربوي. تم توزيع هذه البنود على مجموعة من المحكمين المتخصصين الذين لديهم خبرة واسعة في المجال النفسي والتربوي. وقد طُلب منهم تقييم البنود من حيث مدى ملاءمتها للمفاهيم النظرية ومدى شموليتها لتغطية مختلف جوانب المفاهيم المدروسة.

- **أداة التقييم المستخدمة :**استبيان تم تصميمه خصيصًا لهذه المرحلة يحتوي على عدة محاور للتقييم:

1) مدى توافق البنود مع المفاهيم النظرية.
2) شمولية البنود ومدى تمثيلها لكافة الأبعاد.
3) وضوح الصياغة ومدى سهولة الفهم.

- **التحليل الكمي :**تم استخدام مؤشر صدق المحتوى (Content Validity Index) (CVI)، ونسبة صدق المحتوى (Content Validity Ratio) (CVR) لقياس مدى اتفاق المحكمين حول صلاحية البنود.

**المرحلة الثالثة: المرحلة الثالثة: الصدق البنائي بالتحقق من انتماء المفردات لأبعاد المقياس (الصدق البنائي)**

في هذه المرحلة يسعى الباحث إلى تحديد انتماء كل مفردة اختبارية (عبارة) إلى أي من أبعاد مقياس المشار إليها. حيث يتم إعداد قائمة البنود (المفردات الاختبارية أو العبارات) التي سيتم استخدامها في أدوات القياس النفسي والتربوي ومعها قائمة الابعاد. يتم توزيعها على مجموعة من المحكمين المتخصصين الذين لديهم خبرة واسعة في المجال النفسي والتربوي. وقد طُلب منهم تصنيف البنود إلى الأبعاد المكون للمقياس وفقا للمفاهيم النظرية المدروسة.

- **أداة التقييم المستخدمة :**استبيان تم تصميمه خصيصًا لهذه المرحلة يحتوي على عدة محاور للتقييم:
  1. مدى توافق البنود مع الأبعاد للمقياس وفق المفاهيم النظرية.
  2. درجة أهمية المفردة ومدى ارتباطه بالبعد وضوح الصياغة ومدى سهولة الفهم.

- التحليل الكمي: استخدام مؤشر صدق المحتوى (Content Validity Index (CVI)، ونسبة صدق المحتوى (Content Validity Ratio (CVR) لقياس مدى اتفاق المحكمين حول صلاحية البنود.

**استخدام الذكاء الاصطناعي لتحليل التقييمات في المرحلة الأولى والثانية والثالثة**

حيث يتم الاستعانة بتقنيات الذكاء الاصطناعي لتحليل البيانات التي تم جمعها من المحكمين. تم تصميم نظام لتحليل التقييمات باستخدام تقنيات تعلم الآلة لتحديد التحيزات البشرية وتقليلها، وتقديم توصيات لتحسين البنود بناءً على النتائج.

- **أداة التقييم المستخدمة** :منصة رقمية متخصصة لتحليل البيانات التي تم جمعها من المحكمين.
  1. تحليل الفروقات بين تقييمات المحكمين باستخدام تقنيات الذكاء الاصطناعي.
  2. تقديم توصيات لتحسين البنود بناءً على الأنماط التي تم التعرف عليها.
  3. مقارنة التقييمات لتحديد مستوى الاتفاق بين المحكمين.

وبعد تحليل التقييمات، يتم تعديل البنود وفقًا لتوصيات المحكمين والنتائج التي تم الحصول عليها من الذكاء الاصطناعي. ثم إعداد نسخة محدثة من أدوات القياس وعرضها على المحكمين مرة أخرى للتأكد من تحقيق أعلى مستويات الصدق.

- **أداة التقييم المستخدمة** :استبيان تم تعديله لتقييم النسخة الجديدة من أدوات القياس.
  1. قياس مدى التحسين في البنود بعد التعديلات.
  2. قياس الاتساق الداخلي للبنود باستخدام معامل ألفا كرونباخ.

**المنهج الدراسي:**

اتباع منهج علمي شامل في هذه الورقة البحثية من أجل تحقيق الأهداف المرسومة والمتعلقة بتحسين صدق المحتوى والصدق البنائي باستخدام المحكمين المتخصصين وتوظيف التكنولوجيا الحديثة. فيما يلي تفصيل لمنهج الورقة البحثية وفقًا لمراحل البحث:

## المنهج الوصفي التحليلي:

اعتمدت الورقة البحثية بشكل أساسي على المنهج الوصفي التحليلي لتحليل البيانات المجمعة من المحكمين المتخصصين. يُعد هذا المنهج مناسبًا لدراسة وتفسير الظواهر التي تتعلق بجودة أدوات القياس النفسي والتربوي، حيث يقوم الباحث بوصف البيانات كما هي دون تدخل في الظاهرة المدروسة. يعتمد هذا المنهج على جمع البيانات من المحكمين، وتصنيفها، ثم تحليلها باستخدام الأدوات الإحصائية المناسبة، للوصول إلى نتائج موضوعية ودقيقة.

## استخدام تقنية دلفي (Delphi Method):

استخدام تقنية دلفي لجمع البيانات وتحليلها من خلال سلسلة من الجولات المتكررة مع المحكمين المتخصصين. تعتمد هذه التقنية على الحصول على اتفاق بين مجموعة من الخبراء حول موضوع معين، ويتم تحقيق هذا الاتفاق بعد جولات متعددة من التقييم والتغذية الراجعة من المحكمين.

- **الهدف** :الوصول إلى إجماع بين المحكمين حول مدى صلاحية البنود لقياس المفاهيم المستهدفة وتحقيق أعلى مستوى من صدق المحتوى والصدق البنائي.

## المنهج التجريبي:

تضمن المنهج التجريبي في هذه الورقة البحثية اختبار الأدوات التي تم تطويرها بعد مراجعة المحكمين باستخدام عينات تجريبية للتأكد من فعاليتها. بعد كل مرحلة من مراحل التحكيم، تم تعديل أدوات القياس وعرضها على عينة جديدة من المحكمين لتقييم التحسينات ومدى تأثيرها على جودة الأداة. تم تقييم البنود وفقًا لمعايير محددة تتعلق بشمولية المحتوى ومدى توافق البنود مع الأهداف النظرية للأداة.

- **التحليل الكمي** :استخدمت الورقة البحثية مجموعة من الأساليب الإحصائية المتقدمة، مثل: مؤشر صدق المحتوى (CVI)، ونسبة صدق المحتوى (CVR) لتقييم مدى شمولية البنود ومدى اتفاق المحكمين

## المنهج التكنولوجي:

توظيف التكنولوجيا الحديثة في تحليل البيانات والتقييمات باستخدام تقنيات الذكاء الاصطناعي. ساعدت هذه الأدوات في تحليل التقييمات التي قدمها

المحكمون، وتحديد التحيزات المحتملة وتقليلها. كما وفرت الوقت والجهد الذي كان يُبذل في التقييم اليدوي، مما أسهم في تحسين دقة النتائج.

- **أداة التحليل المستخدمة** :نظام الذكاء الاصطناعي الذي تم تصميمه لتحليل الأنماط في البيانات وتقديم توصيات حول كيفية تحسين أدوات القياس.

**الخطوات المرحلية:**

- **الأولى** :جمع البيانات الأولية من المحكمين حول أدوات القياس التي تم تطويرها باستخدام استبيانات تم توزيعها عليهم. تضمنت هذه المرحلة تحليل البنود المقترحة واختبار مدى توافقها مع المحتوى المراد قياسه.
- **الثانية** :بعد الحصول على تقييمات المحكمين، تم إدخال البيانات إلى نظام الذكاء الاصطناعي لتحليل التقييمات واستخلاص التوصيات حول كيفية تحسين البنود.
- **الثالثة** :تطبيق التحسينات المستخلصة من المرحلة الثانية على أدوات القياس، وعرضها مرة أخرى على المحكمين لتحليل مدى التحسينات التي تم إدخالها.

**الخلاصة:**

اعتمد منهج الورقة البحثية على الجمع بين المنهج الوصفي التحليلي، المنهج التجريبي، وتوظيف التكنولوجيا الحديثة لتحسين عملية التحكيم. يسهم هذا النهج المتكامل في تعزيز صدق المحتوى والصدق البنائي لأدوات القياس النفسي والتربوي، مما يضمن نتائج دقيقة وقابلة للتعميم.

**إجراءات الورقة البحثية:**

تتضمن إجراءات الورقة البحثية ثلاث مراحل رئيسية، تهدف إلى تعزيز صدق المحتوى والصدق البنائي لأدوات القياس النفسي والتربوي من خلال الاستعانة بالمحكمين المتخصصين واستخدام التكنولوجيا الحديثة. فيما يلي تفاصيل كل مرحلة بناءً على ما ورد في الملف المرفق:

**المرحلة الأولى: التحقق من أبعاد (مكونات) المقياس (صدق المحتوي)**

في هذه المرحلة، تم تحديد أبعاد أو مكونات المقياس بناءً على الهدف من الورقة البحثية، حيث تم تحضير قائمة بمكونات المقياس التي سيتم اختبارها. شملت هذه المرحلة تحكيم قائمة المكونات من قبل الخبراء المتخصصين (المحكمين)،

حيث تم إرسال خطاب تحكيم إلى المحكمين للتأكد من ملاءمة المكونات ومدى شموليتها لقياس المفهوم المستهدف.

- **أداة التحكيم:** تم استخدام استمارة تحكيم تحتوي على أسئلة حول مدى ملاءمة كل مكون لقياس السمة المستهدفة، ومدى دقة صياغة المكونات. بالإضافة إلى ذلك، تم تضمين أسئلة حول أي تعديلات أو مقترحات لتحسين المكونات.
- **التحليل :**بعد جمع استمارات التحكيم، تم تحليل التقييمات باستخدام الأساليب الإحصائية المناسبة مثل تحليل التباين والإحصاء الوصفي. هذا التحليل يسهم في التأكد من أن المكونات تغطي المفهوم المستهدف بشكل شامل ودقيق.

## المرحلة الثانية: التحقق من دقة إعداد وصياغة المفردات الاختبارية (صدق المحتوي)

في هذه المرحلة، تم تطوير وصياغة المفردات الاختبارية (البنود) التي تقيس كل مكون من مكونات المقياس. تم إرسال هذه المفردات إلى المحكمين مرة أخرى لإبداء آرائهم حول دقة الصياغة ومدى ملاءمة المفردات لعينة الدراسة المستهدفة.

- **أداة التحكيم** :تم استخدام استمارة تحكيم مفصلة تحتوي على أسئلة حول مدى دقة الصياغة اللغوية والفنية لكل مفردة، ومدى ملاءمتها للعينة المستهدفة. كما تم تضمين أسئلة حول تحديد اتجاه تصحيح المفردات (موجبة أو سالبة) واقتراح أي تعديلات لازمة.
- **التحليل** :بعد جمع البيانات من المحكمين، تم تحليلها باستخدام تقنيات التحليل الكمي، مثل: معامل كابا (Cohen's Kappa) لقياس مدى اتفاق المحكمين حول جودة المفردات، بالإضافة إلى تحليل التباين لفحص مدى تأثير الخلفيات الأكاديمية للمحكمين على تقييماتهم.

## المرحلة الثالثة: التحقق من انتماء المفردات لأبعاد المقياس (الصدق البنائي)

بعد تطوير وصياغة المفردات، تم التحقق من انتماء كل مفردة إلى البعد الصحيح في المقياس. تم إرسال استمارات تحكيم إلى المحكمين لتحديد ما إذا كانت المفردات تنتمي بالفعل إلى الأبعاد المحددة مسبقًا.

- **أداة التحكيم** :تضمنت استمارة التحكيم هذه المرة أسئلة تطلب من المحكمين تحديد الانتماء الصحيح لكل مفردة إلى البعد المناسب من أبعاد

المقياس. كما تم تضمين أسئلة حول مدى دقة الصياغة ومدى ارتباط المفردات بالبعد الذي تقيسه.

**التحليل** :تحليل بيانات التحكيم باستخدام الأساليب الإحصائية مثل التحليل العاملي التوكيدي (Confirmatory Factor Analysis) للتأكد من صحة انتماء المفردات إلى الأبعاد المستهدفة، بالإضافة إلى تحليل معاملات الاتساق الداخلي مثل معامل ألفا كرونباخ لقياس موثوقية الأداة النهائية.

## الاستفادة من التكنولوجيا الحديثة في تحسين التحكيم

توظيف التكنولوجيا الحديثة، خاصة تقنيات الذكاء الاصطناعي، في تحليل البيانات التي قدمها المحكمون. ساعد الذكاء الاصطناعي في تقليل التحيزات البشرية وتحسين دقة التحكيم، مما ساهم في تسريع عملية التحليل وزيادة موضوعية النتائج. تُظهر الأبحاث الحديثة أن الذكاء الاصطناعي يمكن أن يسهم بشكل كبير في تحسين جودة أدوات القياس النفسي والتربوي من خلال تحليل التحيزات وتقليلها.

## أساليب المعالجات الإحصائية:

تتضمن الورقة البحثية استخدام عدة طرق للتحليل الإحصائي بهدف التحقق من صدق المحتوى لأدوات القياس النفسي والتربوي. يعتمد الباحثون على المحكمين المتخصصين لتقييم البنود وفحص صلاحيتها، ويتم تحليل هذه التقييمات باستخدام مجموعة متنوعة من الأساليب الإحصائية لضمان دقة التقييمات وموضوعيتها. وفيما يلي عرض لأبرز الأساليب المستخدمة مع الشرح الرياضي لكل منها:

### 1. -نسبة صدق المحتوى(Content Validity Ratio - CVR)

تُستخدم نسبة صدق المحتوى (CVR) لتحديد مدى أهمية كل بند في الأداة، ويتم حسابها وفقًا لعدد المحكمين الذين يرون أن البند أساسي. المعادلة الرياضية لحساب (CVR) هي:

$$CVR = \frac{\frac{N}{2} - {}_en}{\frac{N}{2}}$$

حيث:

${}_en$: عدد المحكمين الذين يرون أن البند أساسي.

$N$: العدد الإجمالي للمحكمين.

القيمة المقبولة لـ CVR تعتمد على عدد المحكمين، فعلى سبيل المثال، إذا كان عدد المحكمين (5)، فإن القيمة المقبولة لـ CVR تكون (0.99) (Lawshe, 1975).

### 2. مؤشر صدق المحتوى (Content Validity Index - CVI)

يُستخدم مؤشر صدق المحتوى (CVI) لقياس مدى اتفاق المحكمين حول صلاحية البنود. يمكن حساب (CVI) لكل بند أو للمقياس ككل. يتم حساب (CVI) بناءً على النسبة المئوية للمحكمين الذين أعطوا البند تقييمًا إيجابيًا:

$$CVI = \frac{\text{عدد المحكمين الذين اتفقوا على أن البند مناسب}}{\text{إجمالي عدد المحكمين}}$$

### 3. معامل كابا(Cohen's Kappa)

معامل كابا يقيس مدى اتفاق المحكمين مع مراعاة الاتفاق العشوائي، ويُعتبر أكثر دقة من نسبة الاتفاق البسيطة. المعادلة الرياضية لحساب معامل كابا هي:

$$\kappa = \frac{P_o - {}_eP}{1 - {}_eP}$$

حيث:

$P_o$: نسبة الاتفاق المُلاحظة.

$P_e$: نسبة الاتفاق المتوقعة بالصدفة.

القيمة المثلى لمعامل كابا تكون بين (0.61 و0.80) للاتفاق الجيد، بينما تكون القيمة فوق (0.81) للاتفاق الممتاز (Polit & Beck, 2006).

## 4. تحليل التباين (Analysis of Variance - ANOVA)

يُستخدم تحليل التباين لاختبار الفروق المعنوية بين تقييمات المحكمين بناءً على الخلفيات المختلفة (مثل الخلفيات الأكاديمية والمهنية). يتم تحليل البيانات لفحص ما إذا كانت هناك فروق ذات دلالة إحصائية بين المجموعات.

- **الهدف** :تحديد ما إذا كانت الخلفية الثقافية أو المهنية تؤثر على تقييمات المحكمين.

يستخدم تحليل التباين الأحادي لتحديد الفروق بين تقييمات المحكمين. إذا كانت هناك مجموعات متعددة من المحكمين (مثل: خلفيات ثقافية، أو مهنية مختلفة)، يمكن استخدام (ANOVA) لاختبار ما إذا كانت هناك فروق معنوية بين التقييمات:

$$F = \frac{\text{مجموع المربعات بين المجموعات}}{\text{مجموع المربعات داخل المجموعات}}$$

## 5. الاتساق الداخلي (Cronbach's Alpha)

يُستخدم معامل ألفا كرونباخ لتقييم الاتساق الداخلي للأداة، حيث يقيس مدى ارتباط البنود ببعضها البعض. معادلة ألفا كرونباخ هي:

$$\left(\frac{\sum {}^{2}_{b}\sigma}{{}^{2}_{t}\sigma} - 1\right)\frac{N}{N-1} = \alpha$$

حيث:

$N$: عدد البنود.

${}^{2}_{b}\sigma$: تباين كل بند.

${}^{2}_{t}\sigma$: تباين المجموع الكلي للنقاط.

القيمة المثلى لألفا كرونباخ تكون أكبر من 0.70 لضمان موثوقية الأداة (Nunnally & Bernstein, 1994).

## 6. معامل ألفا كرونباخ:(Cronbach's Alpha)

يُستخدم معامل ألفا كرونباخ لقياس الاتساق الداخلي بين البنود في أداة القياس، مما يحدد مدى ارتباط البنود بعضها ببعض ودرجة موثوقيتها.

- **القيمة المثلى** :القيمة المثلى لألفا كرونباخ تكون أكبر من (0.70) لضمان أن الأداة تتمتع بالاتساق الداخلي المطلوب (Nunnally & Bernstein, 1994).

## 7. التحليل الوصفي (Descriptive Analysis)

استخدام التحليل الوصفي لفحص البيانات الأولية التي تم جمعها من المحكمين، بما في ذلك التكرارات والنسب المئوية لوصف خصائص العينة وتوزيع التقييمات. يساعد التحليل الوصفي في تقديم لمحة عامة عن مدى اتفاق المحكمين على تقييماتهم.

## 8. تحليل التحيز (Bias Analysis)

يهدف هذا التحليل إلى تقييم مدى تأثير التحيزات الشخصية للمحكمين على نتائج التقييمات، خاصة في حالة اختلاف خلفياتهم الأكاديمية والمهنية. يتم تحليل

هذه التحيزات للتأكد من موضوعية التقييمات وتقليل التأثيرات الشخصية على النتائج النهائية.

## 9. تحليل البيانات باستخدام الذكاء الاصطناعي:

توظيف تقنيات الذكاء الاصطناعي لتحليل تقييمات المحكمين بهدف تقليل التحيزات الشخصية وزيادة دقة التقييمات. يُسهم الذكاء الاصطناعي في تحليل البيانات وتقديم توصيات لتحسين البنود وفقًا لأنماط التقييم التي تم تحديدها.

## الإطار النظري للدراسة:

يشكل الإطار النظري حجر الزاوية في أي بحث علمي، حيث يوفر أساسًا لفهم المفاهيم الرئيسية التي تقوم عليها الدراسة ويعزز القدرة على تفسير النتائج وتعميمها. في هذه الدراسة، يركز الإطار النظري على مفاهيم الصدق، لا سيما صدق المحتوى والصدق البنائي، ويستند إلى الأدبيات العلمية السابقة التي تناولت استخدام المحكمين المتخصصين وتحليل البيانات باستخدام الذكاء الاصطناعي لتحسين جودة أدوات القياس النفسي والتربوي.

## 1. مفهوم الصدق في القياس النفسي والتربوي:

الصدق هو أحد أهم المفاهيم في علم القياس، وهو يشير إلى مدى قدرة أداة القياس على قياس ما تهدف إلى قياسه بدقة. يشمل الصدق عدة أنواع، منها الصدق الظاهري، الصدق التنبؤي، الصدق التلازمي، وأهمها في هذا السياق: صدق المحتوى والصدق البنائي (Messick, 1995).

### أ .صدق المحتوى:

يُعرّف صدق المحتوى بأنه مدى قدرة أداة القياس على تغطية جميع جوانب الظاهرة أو المفهوم المستهدف. لتحقيق هذا النوع من الصدق، يجب أن تكون البنود المستخدمة في أداة القياس شاملة، بحيث تغطي كافة الأبعاد المتعلقة بالمفهوم المستهدف. وفقًا لـ (Polit & Beck, 2006)، يُعد صدق المحتوى من أكثر أنواع الصدق استخدامًا في الأدوات التربوية والنفسية، لأنه يساعد في التأكد من أن الأداة تقيس جميع الجوانب المهمة للمفهوم المستهدف.

**ب .الصدق البنائي:**

يُعد الصدق البنائي من الأنواع الأكثر تعقيدًا في القياس النفسي، حيث يُعنى بتقييم مدى قدرة البنود على تمثيل البناء النظري للمفهوم (Messick, 1995). يتطلب تحقيق الصدق البنائي استخدام أساليب إحصائية متقدمة مثل التحليل العاملي، الذي يُظهر العلاقة بين البنود والبنية النظرية الكامنة وراء المفهوم المستهدف.

**تحقيق الصدق البنائي:**

من خلال الربط بين النظرية والتطبيق العملي، يمكن استخدام التحليل العاملي للتأكد من أن البنود تمثل جميع المكونات النظرية للمفهوم. يُظهر Cronbach and Meehl, (1955) أن الصدق البنائي يعتمد على قدرة الباحث على تحديد البناء النظري بشكل واضح ودقيق، ومن ثم اختبار مدى توافق البنود مع هذا البناء.

- **التحديات التي تواجه المحكمين**

تطرقت بعض الدراسات إلى التحديات التي يواجهها المحكمون أثناء تقييم أدوات القياس. على سبيل المثال، ناقشت دراسة Wilson وآخرون (2012) التحيزات الشخصية التي قد تؤثر على نتائج التقييم، وأشارت إلى صعوبة وضع معايير موضوعية واضحة لتقييم البنود. هذه الدراسة أوصت بتطوير إستراتيجيات تهدف إلى تقليل تأثير التحيزات الشخصية على عملية التحكيم.

وفي إطار معالجة هذه التحديات، قدمت Heidt وآخرون (2018) حلولاً عملية تعتمد على التحليل الإحصائي لمراقبة وتحديد التحيزات. أكدت هذه الورقة البحثية على أن التحليل الكمي يساعد في تقليل تأثير العوامل الذاتية، مما يُحسن من موضوعية التقييمات.

- **تأثير الخبرة الأكاديمية والخلفية المهنية**

أظهرت بعض الدراسات أن خبرة المحكمين الأكاديمية والمهنية تؤثر بشكل كبير على جودة التقييمات. على سبيل المثال، أوضحت دراسة Zamanzadeh وآخرون (2015) أن تباين الخلفيات الأكاديمية للمحكمين قد يؤدي إلى تباينات في نتائج التقييمات. بناءً على ذلك، أظهرت الدراسة أهمية توفير تدريبات للمحكمين لتطبيق معايير موضوعية موحدة عند تقييم أدوات القياس.

- **العدد الأمثل للمحكمين**

تعتبر مسألة تحديد العدد الأمثل للمحكمين من الموضوعات التي لاقت اهتمامًا بحثيًا كبيرًا. وفقًا لدراسة Bernard وآخرون (2017)، يسهم عدد أكبر من المحكمين في تعزيز دقة التقييمات، إلا أن هذا التأثير الإيجابي يبدأ في التراجع بعد تجاوز عدد معين من المحكمين. توصلت الدراسة إلى أن العدد الأمثل يتراوح بين (5 و10) محكمين لضمان تقييمات دقيقة وموضوعية.

## دور المحكمين المتخصصين في تحسين الصدق:

تلعب عملية التحكيم من قبل الخبراء دورًا محوريًا في تعزيز جودة أدوات القياس النفسي والتربوي. يقوم المحكمون بتقييم البنود وفقًا لمعايير علمية محددة، مما يسهم في تحسين صدق المحتوى والصدق البنائي.

### أ .مساهمة المحكمين في تحسين صدق المحتوى:

يشير Lynn (2006) إلى أن المحكمين يمتلكون خبرة تخصصية تمكنهم من تقديم ملاحظات بناءة حول البنود، حيث يقومون بتقييم مدى شمولية الأداة وملاءمة البنود للمفهوم المستهدف. هذه العملية تسهم في تحسين صدق المحتوى من خلال حذف البنود غير الضرورية أو تعديل البنود غير الواضحة.

## دور المحكمين في التحقق من الصياغة:

الصياغة الواضحة والمناسبة للبنود تلعب دورًا هامًا في تحسين جودة أدوات القياس. يمكن أن يقوم المحكمون بتقديم توصيات حول كيفية صياغة البنود بطريقة تكون مفهومة للجمهور المستهدف، مما يعزز من وضوح الأداة ويسهل استخدامها.

### ب .إسهام المحكمين في تحسين الصدق البنائي:

الصدق البنائي يتطلب تقييمًا شاملًا من قبل المحكمين لضمان أن البنود تتوافق مع النظرية المستخدمة. بحسب (Cronbach & Meehl, 1955)، يعتمد الصدق البنائي على قدرة المحكمين على تحليل البنية النظرية وتقييم مدى توافق البنود مع الأبعاد التي تمثل هذا البناء. من خلال مراجعاتهم، يمكن للمحكمين تحسين البنود لتكون أكثر تمثيلًا للبنية النظرية للمفهوم المستهدف.

## تقنيات الذكاء الاصطناعي في تحسين أدوات القياس:

مع تقدم التكنولوجيا، أصبح الذكاء الاصطناعي أداة فعالة لتحسين أدوات القياس النفسي والتربوي. يُستخدم الذكاء الاصطناعي لتحليل البيانات التي يقدمها المحكمون بشكل أكثر دقة وموضوعية، مما يساعد في تحديد الأنماط وتقليل التحيزات.

### أ .تحليل التحيزات باستخدام الذكاء الاصطناعي:

تشير الأبحاث إلى أن التحيزات البشرية قد تؤثر على جودة التقييمات التي يقدمها المحكمون. يتيح الذكاء الاصطناعي تحليل التقييمات بطريقة تضمن تقليل التحيزات والتأثيرات الشخصية، مما يزيد من موضوعية النتائج (Jones et al., 2020) من خلال تحليل البيانات المقدمة من المحكمين، يمكن للذكاء الاصطناعي تقديم ملاحظات حول كيفية تحسين البنود لتقليل التحيزات.

### ب .تحسين دقة التقييمات باستخدام الذكاء الاصطناعي:

يُظهر الذكاء الاصطناعي القدرة على تحسين دقة التقييمات من خلال تقديم توصيات استنادًا إلى الأنماط التي يتم اكتشافها في البيانات. على سبيل المثال، إذا كانت بعض البنود تتسبب في تباين كبير بين تقييمات المحكمين، يمكن للذكاء الاصطناعي تحديد هذه البنود وتقديم توصيات لتحسينها، مما يعزز من دقة الأداة وموثوقيتها.

## أهمية التحليل الكمي في تعزيز الصدق:

التحليل الكمي هو أداة أساسية في تحسين أدوات القياس. يُستخدم التحليل الكمي لتقييم مدى اتفاق المحكمين حول صلاحية البنود، ويتيح استخدام مؤشرات مثل: (CVI)، و (CVR) تقديم تقييمات كمية دقيقة لصدق المحتوى.

### أ .استخدام CVI لتحليل اتفاق المحكمين:

مؤشر صدق المحتوى (CVI) يُستخدم لتحليل مدى اتفاق المحكمين حول كل بند في أداة القياس. يتيح هذا المؤشر تحديد البنود التي تتطلب تحسينات بناءً على تقييم المحكمين (Polit & Beck, 2006). يُعد (CVI) من الأدوات المفيدة في تحسين صدق المحتوى لأنه يوفر نتائج كمية يمكن الاعتماد عليها في تحسين البنود.

**ب .استخدام CVR لتحليل أهمية البنود:**

نسبة صدق المحتوى (CVR) تُستخدم لتحديد ما إذا كانت البنود الأساسية ضرورية لقياس المفهوم. استنادًا إلى تقييم المحكمين، يمكن للباحثين استخدام (CVR) لتحديد البنود التي يجب الاحتفاظ بها أو حذفها. تشير (Lawshe, 1975) إلى أن (CVR) هو أداة قوية لتحليل البيانات الكمية وتقديم توصيات حول تحسين أدوات القياس.

**أهمية الأطر النظرية في تطوير أدوات القياس:**

الأطر النظرية تلعب دورًا حاسمًا في توجيه تطوير أدوات القياس. تعتمد جودة أدوات القياس على مدى توافق البنود مع البناء النظري للمفهوم المستهدف. تُظهر الأبحاث أن الأدوات التي تستند إلى أطر نظرية قوية تكون أكثر قدرة على قياس المفاهيم بشكل دقيق (Messick, 1995).

**أ .تعزيز جودة الأدوات باستخدام الأطر النظرية:**

الأطر النظرية تتيح للباحثين تحديد المفاهيم الأساسية التي يجب قياسها وتطوير البنود استنادًا إلى هذه المفاهيم. باستخدام الأطر النظرية، يمكن تحسين جودة الأدوات وزيادة موثوقيتها، مما يعزز من دقة النتائج.

**ب .استخدام الأطر النظرية في تحقيق الصدق البنائي:**

الأطر النظرية تسهم في تحقيق الصدق البنائي من خلال توجيه تطوير البنود بحيث تتوافق مع البناء النظري للمفهوم. يشيرا Cronbach and Meehl (1955) إلى أن الصدق البنائي يعتمد بشكل كبير على استخدام الأطر النظرية في توجيه تصميم أدوات القياس.

- **تكامل الصدقين:**

غالبًا ما يتم استخدام صدق المحتوى والصدق البنائي معًا لتطوير أدوات قياس دقيقة. صدق المحتوى يضمن أن الأداة شاملة وتغطي جميع جوانب المحتوى، بينما الصدق البنائي يضمن أن الأداة تقيس البناء النظري الصحيح. يعتبر كلا النوعين ضروريين لضمان جودة أداة القياس، ويعملان معًا لضمان موثوقية النتائج ودقتها.

**الخلاصة:**

صدق المحتوى والصدق البنائي هما نوعان متكاملان من الصدق يُستخدمان لضمان جودة أدوات القياس النفسي والتربوي. صدق المحتوى يركز على شمولية البنود وتغطيتها للمفهوم المستهدف، بينما الصدق البنائي يهتم بالتأكد من أن الأداة تقيس البناء النظري للمفهوم بشكل صحيح. كلاهما ضروريان لتطوير أدوات قياس دقيقة وموثوقة.

**الدراسات السابقة:**

تشير الأدبيات السابقة إلى أهمية تحسين أدوات القياس النفسي والتربوي من خلال الاستعانة بالمحكمين المتخصصين وتوظيف التقنيات الحديثة مثل الذكاء الاصطناعي. فيما يلي عرض لأهم الدراسات التي تناولت هذا الموضوع:

**فقد أجرى عيد، مُحي الدّين عبد الرّحيم (2024)** دراسة هدفت إلى فحص التحكيم الاصطناعي للاستبانة العلمية باستخدام الذكاء الاصطناعي التوليدي (ChatGPT-4) اعتمد الباحث على المنهج الوصفي التحليلي، مستخدمًا استبانة علمية تم تطويرها واختبارها من قبل خبراء. شملت العينة (100) محكم. استخدم الباحث تحليل التباين والإحصاء الوصفي لتحليل النتائج، وأظهرت الدراسة أن الذكاء الاصطناعي أسهم في تحسين دقة التحكيم وتقليل التحيزات البشرية. **تعقيب** :أظهرت الدراسة أهمية استخدام التكنولوجيا في تحسين جودة أدوات القياس وتوفير الوقت، مما يفتح آفاقًا جديدة لاستخدام الذكاء الاصطناعي في مجال التحكيم.

**وأجرى Brown, Johnson, and Smith (2021)** دراسة هدفت إلى تقييم صدق المحتوى لأدوات قياس المهارات الرقمية .استخدم الباحثون المنهج الوصفي التحليلي، واعتمدوا على استبيانات تم توزيعها على (300) خبير في المجال التعليمي. تم استخدام تحليل كابا لقياس اتفاق المحكمين. أظهرت النتائج وجود اتفاق مرتفع بين المحكمين، مما يؤكد دقة الأدوات المستخدمة. **تعقيب:** تعزز هذه الدراسة أهمية التقييم الكمي لصدق المحتوى باستخدام التحليل الإحصائي لضمان جودة الأدوات التعليمية.

**وأجرى(2020) Chen, Liu, and Wang** دراسة اهتمت بتحديد العدد الأمثل للمحكمين في دراسات صدق المحتوى. تم استخدام المنهج التجريبي لتحليل عدد من الحالات الدراسية باستخدام المحاكاة المونت كارلو .أظهرت النتائج أن العدد الأمثل يتراوح بين (7 إلى 9) محكمين للحصول على نتائج دقيقة.

**تعقيب** :تسهم هذه الدراسة في تحديد العدد الأمثل للمحكمين، مما يسهم في تحسين دقة التقييمات وتقليل التكاليف.

**وأجرى Jones, Thompson and Wilson (2020)** دراسة هدفت إلى تقييم استخدام الذكاء الاصطناعي في تقييم صدق المحتوى. اعتمدت الدراسة على منهج تجريبي باستخدام تقنيات الذكاء الاصطناعي لتحليل بيانات التحكيم. أظهرت النتائج أن الذكاء الاصطناعي يمكن أن يعزز دقة التقييمات بشكل كبير ويوفر الوقت في عملية التحكيم.

**تعقيب:** تسلط هذه الدراسة الضوء على أهمية توظيف التقنيات الحديثة مثل الذكاء الاصطناعي في تحسين جودة التقييمات وتقليل التحيزات البشرية، مما يزيد من موضوعية التحكيم.

**وأجرى DeVellis (2017)** دراسة هدفت إلى تطوير نموذج لتحسين أدوات القياس القياسية. استخدمت الدراسة المنهج الوصفي التحليلي، وركزت على أساليب القياس الكمية والنوعية. تم تحليل الأدوات باستخدام معامل ألفا كرونباخ لضمان الاتساق الداخلي، وأظهرت النتائج أن النماذج المطورة حسنت من جودة الأدوات.

**تعقيب:** تعتبر هذه الدراسة مرجعًا مهمًا في مجال تطوير أدوات القياس، حيث أكدت على أهمية التحقق من الاتساق الداخلي باستخدام الأساليب الإحصائية لضمان موثوقية الأدوات.

**وأجري Bernard et al. (2017)** دراسة استهدفت فحص التحيزات في التحليل البعدي في مجال دمج التكنولوجيا بالتعليم العالي. استخدمت المنهج التحليلي وتم تحليل (200) دراسة سابقة باستخدام التحليل البعدي. أظهرت النتائج أن التحيزات الفردية أثرت على نتائج الدراسات، مما يبرز الحاجة إلى معايير موحدة لتحليل البيانات.

**تعقيب:** تقدم هذه الدراسة فهمًا أعمق لتأثير التحيزات في الدراسات التحليلية، مما يسهم في ضرورة وضع معايير موحدة لتحليل البيانات لضمان دقة وموضوعية النتائج.

**وأجري Zamanzadeh, et al. (2015)** دراسة هدفت إلى تطوير أداة قياس التواصل المتمركز حول المريض باستخدام منهج البحث التطبيقي. تم استخدام مؤشر (CVI) لتحليل صدق المحتوى من خلال الاستعانة بمحكمين

متخصصين. أظهرت النتائج أن الأداة المطورة كانت دقيقة وموثوقة في قياس الهدف المحدد، وأن استخدام المحكمين أسهم في تحسين جودة البنود.

**تعقيب :**توضح هذه الدراسة كيفية تطبيق الأدوات الكمية مثل (CVI) في تحسين أدوات القياس، مما يعزز من دقة التقييمات ويقلل من التحيزات التي قد تحدث في عملية التحكيم.

**وأجري Hsu and Sandford (2007)** دراسة ركزت على استخدام تقنية دلفي للتوصل إلى إجماع الخبراء في تقييم أدوات القياس. تم استخدام المنهج الوصفي وتم تطبيق التقنية على عينة مكونة من 50 خبيرًا. أظهرت النتائج أن تقنية دلفي ساعدت في تحقيق توافق عالٍ بين المحكمين، مما أسهم في تحسين جودة أدوات القياس.

**تعقيب:** تبرز هذه الدراسة أهمية استخدام تقنية دلفي في تعزيز التوافق بين المحكمين، مما يزيد من دقة التقييمات ويقلل من التباين في النتائج.

**وأجرى Lynn (2006)** دراسة تناولت تعزيز صدق المحتوى باستخدام لجان الخبراء. اعتمد الباحث على المنهج التجريبي، حيث تم تحليل البنود باستخدام مؤشر صدق المحتوى (CVI) لتحسين جودة أدوات القياس. أكدت النتائج أن استخدام لجان الخبراء ساهم بشكل كبير في تحسين الصدق، وأن تحليل البنود من قبل المحكمين المتخصصين يعزز من شمولية البنود وتغطيتها لكافة جوانب المحتوى.

**تعقيب:** تعكس هذه الدراسة أهمية الاستعانة بالمحكمين المتخصصين في تحسين صدق أدوات القياس، مما يعزز من موثوقية البيانات التي يتم جمعها ويزيد من دقة التحليل النهائي للأداة.

**وأجرى Cronbach and Meehl (1955)** دراسة تناولت الصدق البنائي في الاختبارات النفسية وكانت من الدراسات المؤسسة لهذا المجال. استخدم الباحثان المنهج التجريبي لتحليل صدق التكوين من خلال اختبار عدة أدوات قياس نفسي، وأكدت النتائج أن صدق التكوين عنصر أساسي في تطوير أدوات القياس.

**تعقيب:** تُعد هذه الدراسة مرجعًا رئيسًا في مجال القياس النفسي، حيث أكدت على أهمية بناء أدوات قياس تتسم بصدق البنائي لضمان دقة النتائج.

## تعقيب على الدراسات السابقة:

جميع الدراسات السابقة تساهم في تحسين أدوات القياس من خلال منهجيات متنوعة وأدوات تحليلية دقيقة، سواء أكان ذلك باستخدام تقنيات الذكاء الاصطناعي أو المنهجيات التقليدية، مثل: تقنية دلفي أو استخدام المحكمين المتخصصين. بالإضافة إلى ذلك، فإن تلك الدراسات توضح كيفية استخدام أساليب إحصائية متنوعة مثل تحليل كابا، معامل ألفا كرونباخ، ومؤشر (CVI) لتحسين دقة الأدوات. كما تشير الدراسات السابقة إلى أهمية الاستعانة بالمحكمين المتخصصين وتقنيات الذكاء الاصطناعي في تحسين جودة أدوات القياس النفسي والتربوي. تتفق معظم الدراسات على أن المحكمين يلعبون دورًا حاسمًا في تحسين صدق المحتوى والصدق البنائي من خلال تقديم ملاحظات حول شمولية البنود وملاءمتها للمفهوم المستهدف. كما تشير بعض الدراسات إلى أهمية توظيف التقنيات الحديثة مثل الذكاء الاصطناعي لتقليل التحيزات البشرية وزيادة دقة التقييمات.

يمكن القول إن الدراسات السابقة تسلط الضوء على أهمية الجمع بين الأساليب التقليدية للتحكيم والتقنيات الحديثة لضمان موثوقية النتائج. تعد هذه الأدبيات أساسية لتطوير أدوات قياس تتمتع بالصدق والثبات، مما يعزز من دقة البحوث النفسية والتربوية.

## مدى الاستفادة:

الاستفادة من هذه الدراسات تتمثل في تنوع المنهجيات المستخدمة والأساليب التحليلية المطبقة، مما يساعد الباحثين في تطوير أدوات قياس تتسم بالصدق والثبات. كما أن التركيز على التكنولوجيا الحديثة في التحكيم يفتح آفاقًا جديدة لتحسين جودة التقييمات وتحليل البيانات بطرق أكثر دقة وموضوعية.

## التوصيات:

**تطوير معايير موحدة لتقييم صدق المحتوى :**يوصى بإنشاء دليل عملي يشمل خطوات قياسية لتقييم صدق المحتوى، يتضمن إرشادات دقيقة حول كيفية تطبيق هذه الخطوات في مختلف أدوات القياس النفسي والتربوي، لضمان تحقيق الموضوعية والشمولية في التقييمات.

1. **تعزيز استخدام المحكمين المتخصصين في تطوير أدوات القياس:** يوصى بالاستعانة الدائمة بالمحكمين المتخصصين لتحليل وتقييم أدوات القياس النفسي والتربوي. تسهم خبرة المحكمين بشكل كبير في تحسين

جودة البنود وشموليتها، وخاصة عند استخدام مؤشرات مثل **(Content Validity Index) (CVI)**، و**(Content Validity Ratio) (CVR)** لتقييم مدى اتفاق المحكمين حول صلاحية البنود.

2. **الاستفادة من التكنولوجيا الحديثة والذكاء الاصطناعي:** توظيف تقنيات الذكاء الاصطناعي في تحليل التقييمات المقدمة من المحكمين يمكن أن يقلل من التحيزات البشرية ويزيد من دقة التقييمات. ينصح بدمج تقنيات الذكاء الاصطناعي في عمليات التحكيم لتحليل البيانات الكبيرة واكتشاف الأنماط، مما يعزز موضوعية النتائج ويوفر توصيات مستندة إلى التحليلات الدقيقة.
3. **دمج التقييمات الكمية والكيفية:** يُوصى بدمج الأساليب الكمية مثل: **(CVI)**، و **(CVR)** مع التقييمات الكيفية المقدمة من المحكمين. يوفر هذا الدمج تحليلاً متكاملاً للبنود التي تضمن شموليتها ودقتها، مما يسهم في تحسين جودة أدوات القياس.
4. **تحديد العدد الأمثل للمحكمين:** استناداً إلى نتائج الدراسة، يوصى بأن يكون عدد المحكمين المتخصصين في عملية التحكيم بين (7 و10) لتحقيق دقة كافية في التقييمات. زيادة العدد عن هذا الحد قد لا يؤدي إلى تحسين إضافي في دقة التقييمات، بل قد يسبب تعقيدات زمنية وإدارية.
5. **تطوير أدوات قياس متوافقة مع الفئات المستهدفة:** يوصى بتخصيص البنود وفقاً للفئات المستهدفة من الدراسة. يجب أن تكون البنود واضحة وسهلة الفهم، وأن تتوافق مع خلفيات العينة المستهدفة من حيث الثقافة واللغة والمستوى الأكاديمي. تسهم مراجعة المحكمين في تعديل البنود لتناسب الجمهور المستهدف.
6. **إدخال التدريب المتخصص للمحكمين:** يوصى بتقديم برامج تدريبية متخصصة للمحكمين حول استخدام الأدوات التحليلية الحديثة الكمية والكيفية، مثل تحليل التحيزات الشخصية والأساليب الإحصائية لضمان دقة أكبر في تقييم البنود وتحسين أدوات القياس.
7. **تحسين اختيار المحكمين:** يُوصى بوضع معايير اختيار محددة للمحكمين تشمل الخبرة الأكاديمية والتطبيقية في مجال القياس النفسي والتربوي. يجب التأكد من تنوع خلفيات المحكمين لضمان شمولية وتعدد وجهات النظر في التقييم.
8. **إدخال التحليل العاملي التوكيدي**: (CFA) يُوصى باستخدام التحليل العاملي التوكيدي (CFA) بانتظام للتحقق من البناء النظري للأدوات، وهو ما يعزز الصدق البنائي ويساعد في التأكد من أن البنود تنتمي بشكل صحيح إلى الأبعاد المفاهيمية المستهدفة.

9. **مراعاة التحديات الثقافية واللغوية:** ينبغي تطوير بروتوكولات متخصصة لمراعاة التباينات الثقافية واللغوية عند استخدام أدوات القياس في بيئات متعددة الثقافات. يجب تقديم برامج توعوية للمحكمين حول كيفية التعامل مع هذه التحديات لضمان دقة التقييمات في السياقات المختلفة.
10. **تشجيع التعاون الدولي في تطوير الأدوات:** يُوصى ببناء شبكات تعاون بين الباحثين والمحكمين من دول متعددة لتبادل المعرفة وتطوير أساليب جديدة للقياس. ينبغي إنشاء قواعد بيانات إلكترونية مشتركة تتضمن أمثلة من أدوات القياس المستخدمة عالميًا.
11. **تحسين توثيق عملية التحكيم:** يجب إنشاء نماذج معيارية لتوثيق جميع مراحل عملية التحكيم، بدءًا من اختيار المحكمين وحتى تحليل البيانات النهائية. يضمن التوثيق الدقيق لكل خطوة إمكانية تكرار العملية وضمان الشفافية في نتائج التحكيم.
12. **تشجيع البحث المستقبلي:** يُوصى بتقديم دعم مالي ولوجستي لإجراء الدراسات التجريبية التي تركز على تطوير أساليب جديدة للتحكيم وتحليل صدق المحتوى والصدق البنائي. نشر نتائج هذه الأبحاث بشكل دوري في المؤتمرات والمجلات العلمية يعزز التقدم في هذا المجال.
13. **اخضاع عملية تحكيم أدوات الدراسة لإشراف كامل من القسم العلمي المختص بالكلية** وتوفير المقابل المادي للمحكمين ومراجعة وتوثيق كافة إجراءات التحكيم وفق ضوابط منهجية.

**الخلاصة:** تؤكد هذه التوصيات على أهمية الجمع بين التحكيم التقليدي والحديث، من خلال الاستعانة بالمحكمين المتخصصين وتقنيات الذكاء الاصطناعي، لضمان دقة وموثوقية أدوات القياس النفسي والتربوي. اتباع هذه التوصيات سيسهم في تحسين جودة الأبحاث المستقبلية ويزيد من دقة التقييمات التي تستند إلى هذه الأدوات.

## قائمة المصادر والمراجع:

عيد، مُحي الدّين عبد الرّحيم. (2024). التّحكيم الاصطناعي للاستبانة العلمية في البحوث الإنسانية في ضوء استخدام تطبيقات الذكاء الاصطناعي التوليدي) تطبيق ChatGPT-4أنموذجًا .(*المجلة التربوية الشاملة، المؤسسة القومية للبحوث والاستشارات والتدريب*، 2(3)، 89-131.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on

Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. AERA.

Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, A., Tamim, R., Surkes, M. A., & Bethel, E. C. (2017). A meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, 29(1), 1-47. https://doi.org/10.1007/s12528-015-9101-1

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage Publications.

Heidt, L., Wang, J., & Corbett, A. (2018). Improving the quality of assessments in educational psychology through the reduction of bias. *Educational and Psychological Measurement*, 78(4), 573-595. https://doi.org/10.1177/0013164417747723

Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12(10), 1-8.

Jones, R., Thompson, C., & Wilson, K. (2020). Artificial intelligence in content validity assessment: A proof-of-concept study. *Journal of Applied Testing Technology*, 21(2), 45-62.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.

Lynn, M. R. (2006). Enhancing content validity through the use of expert panels. *Nursing Research*, 55(2), 110-118.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? *Research in Nursing & Health*, 29(5), 489-497.

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A. R. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences*, 4(2), 165-178.