تأثير استخدام أساليب القياس والتقويم الدقيقة في تحسين جودة الأبحاث العلمية لأعضاء هيئة التدريس

الدكتور عبدالرحمن بن عبدالله بن امبارك السويهري
وزارة التعليم -المملكة العربية السعودية
alswihri@gmail.com

مستخلص البحث:
تهدف هذه الدراسة إلى استكشاف أساليب القياس والتقويم الأكثر شيوعًا في تعزيز جودة الأبحاث العلمية ورصد الصعوبات في الأبحاث متعددة التخصصات، واقتراح استراتيجيات للتحسين، واعتمدت الدراسة على المنهج المختلط باستخدام التصميم التفسيري المتتابع ، حيث جُمعت البيانات الكمية أولاً باستخدام المقياس ثم جمعت البيانات النوعية عبر مقابلات شبه منظمة، وتكونت عينة الدراسة من 350 باحثًا من تخصصات علمية، وتم اختيارهم بطريقة العينة العشوائية الطبقية للبيانات الكمية، و15 باحثًا ذوي خبرة تم اختيارهم بعينة قصدية للبيانات النوعية.وأظهرت نتائج الدراسة ووجود فعالية في استخدام المقاييس الكمية في الأبحاث العلمية بنسبة 80%، وتوجد أهمية للتحليل الاحصائي في نتائج الدراسات بنسبة (70%) وهما الأكثر استخدامًا، مع وجود فروق بين التخصصات، وتوجد علاقة إيجابية قوية بين جودة أساليب القياس والتقويم وخاصية الصدق بقيمة (β = 0.65) حيث تُفسر 42% من التباين، وتؤثر الجودة على الثبات بقيمة (β = 0.61) حيث تفسر 37% من التباين، ويوجد تحديات منها الصعوبة في توحيد الأدوات ، وأيضاً نقص المهارات التقنية، وتعقيد الظواهر متعددة التخصصات، وضعف وقلة الموارد، ومن اقتراحات الباحثين اقتراح تطوير التدريب المتخصص، واقتراح استخدام التكنولوجيا المتقدمة، واقتراح تصميم أدوات مرنة، واقتراح تعزيز التعاون بين الباحثين.وأوصت الدراسة الى تعزيز القدرات من خلال برامج تدريبية متقدمة، والاستفادة من التقنيات المعاصرة لتحسين كفاءة أدوات الدراسة، وتطوير أدوات مرنة تتكيف مع التعددية التخصصية، وتشجيع الشراكات البحثية لتعزيز التكامل، وتوفير الدعم المؤسسي لتذليل العوائق المادية، وإنشاء معايير موحدة للتقييم عبر التخصصات، وإجراء أبحاث مستقبلية لاستكشاف التطبيقات العملية، وتعزيز الوعي بأهمية الخصائص السيكومترية.

الكلمات المفتاحية: أساليب القياس والتقويم، جودة الأبحاث العلمية.

التعريف بالدراسة

المقدمة:

تُشكل الدراسات والأبحاث العلمية العمود الرئيسي لتطوير المعرفة البشرية، حيث تُسهم في فهم الظواهر الطبيعية والاجتماعية وتقديم حلول للمشكلات المعاصرة في مختلف المجالات، وتبرز أساليب القياس والتقويم كعناصر أساسية لضمان جودة الأبحاث ومصداقيتها. فقد أظهرت الدراسات الحديثة أن جودة أساليب القياس تؤثر بشكل مباشر على موثوقية الأبحاث التجريبية، حيث أكدت أن الأدوات ذات الخصائص السيكومترية العالية (مثل الصدق والثبات) تُعزز من مستوى صحة ودقة النتائج بنسبة مرتفعة. (DeVellis, 2016) كما أشير في مجال التقويم إلى معرفة واكتشاف أهمية التقويم متعدد المصادر في تعزيز مصداقية الأبحاث النوعية، مما يُبرز الحاجة إلى تكامل هذه الأساليب عبر التخصصات العلمية المختلفة. (Lincoln& Guba, 1985) فهذا يعني ان الأساليب ليست مجرد أدوات تقنية، بل هي مكونات حاسمة تؤثر على خصائص صدق وثبات النتائج، وقابليتها للتكرار والتعميم، وهي أمور تُعد معايير أساسية لتقييم جودة البحث العلمي (Open Science Collaboration, 2015).

ومع ذلك، فإن التطور المتسارع في منهجيات البحث العلمي، بما في ذلك استخدام التكنولوجيا المتقدمة (Brynjolfsson & McAfee, 2014) قد أدى إلى ظهور تحديات جديدة تتعلق بتطبيق أساليب القياس والتقويم بفعالية، خاصة في الأبحاث متعددة التخصصات (Repko & Szostak, 2020). وفي ضوء هذه التطورات التي نشهدها، أصبح من الضروري دراسة العلاقة بين أساليب القياس والتقويم وجودة الأبحاث العلمية بشكل موسع وإبراز النقاط التي تحتاج لدراسة والتفسير. أن سوء اختيار الأدوات التي تستخدم في الأبحاث بغرض قياس المشكلة البحثية أو غياب التقويم المنهجي لتعزيز الدراسة قد يؤدي إلى تضليل في النتائج، مما يُضعف الثقة في المخرجات البحثية. (Moher et al, 2010) ومن هذا المنطلق تأتي هذه الدراسة إلى استكشاف دور أساليب القياس والتقويم في تعزيز جودة الأبحاث العلمية، مع التركيز على تأثيرها على خصائص الصدق والثبات وقابلية التطبيق، وتقديم إجراءات عملية لتحسين الممارسات البحثية في الظروف المتنوعة، وطرح الاقتراحات المثمرة وذات الخبرة.

مشكلة الدراسة:

مع هذه النهضة التقنية أدى التقدم التكنولوجي إلى زيادة تعقيد هذه الأساليب وتعدد طرق استخداماتها وكمية المخرجات البرامج الحاسوبية، مما يتطلب من الباحثين مهارات متقدمة قد لا تكون متاحة للجميع. (Brynjolfsson & McAfee, 2014) وقد أشار (Boud & Falchikov, 2006) إلى أن الاعتماد المفرط على التقويم الذاتي دون تدعيمه بأساليب موضوعية وطرق علمية قد يؤدي إلى نتائج غير دقيقة، مما يثير تساؤلات حول فعالية بعض الأساليب الشائعة ومدى عمق دقتها وجودتها.

ويُعد التحيز (Bias) أحد التحديات الرئيسية في الأبحاث العلمية، وترتبط أساليب القياس والتقويم ارتباطًا وثيقًا بتقليله أو تفاقمه، وفي هذا المنطلق نجد أن الأدوات القياسية غير الدقيقة أو المعايرة بشكل سيئ قد تُؤدي إلى تحيز في النتائج، مثل تحيز الانتقاء (Selection Bias) أو تحيز القياس (Measurement Bias)، مما يُشوه الاستنتاجات، فقد أظهرت دراسته أن العديد من الأبحاث المنشورة تعاني من نتائج مضخمة بسبب ضعف منهجيات القياس. (Ioannidis, 2005)

ولابد من التأكيد على أن هذه الأساليب تواجه تحديات منهجية وعملية تؤثر على فعاليتها، وأن الأبحاث متعددة التخصصات تعاني من صعوبة في تطوير أدوات قياس تتناسب مع متطلبات التخصصات المتباينة، مما يؤثر سلبًا على الدقة في النتائج. (Repko & Szostak, 2020)

ويجد الباحث اختلاف في متطلبات القياس والتقويم في العلوم الإنسانية عن تلك في العلوم الطبيعية بسبب طبيعتها التفسيرية، وأن القياس في هذا المجال يعتمد على أدوات نوعية مثل التحليل السردي والملاحظة التشاركية، التي تُركز على فهم السياق بدلاً من القيم العددية.

واستنادا لما سبق تتضح مشكلة الدراسة من هذه الفجوات والصعوبات، حيث تسعى إلى الإجابة عن **السؤال الرئيسي: كيف يمكن لأساليب القياس والتقويم أن تُسهم في تحسين جودة الأبحاث العلمية في ظل التحديات الحالية؟**

أسئلة الدراسة:

1. ما هي أساليب القياس والتقويم الأكثر استخدامًا في الأبحاث العلمية عبر التخصصات المختلفة؟
2. كيف تؤثر جودة أساليب القياس والتقويم على صدق وثبات نتائج الأبحاث العلمية؟
3. ما هي التحديات الرئيسية التي تواجه الباحثين عند تطبيق أساليب القياس والتقويم في الأبحاث متعددة التخصصات؟

أهداف الدراسة:

1. تحديد وتصنيف أساليب القياس والتقويم الشائعة في الأبحاث العلمية وفقًا لمجالاتها التطبيقية.
2. تقييم تأثير جودة أساليب القياس والتقويم على الصدق والثبات، وقابلية تكرار النتائج البحثية.
3. استكشاف التحديات المنهجية والعملية المرتبطة بتطبيق هذه الأساليب في سياقات بحثية متنوعة.

أهمية الدراسة:

تُعد أساليب القياس والتقويم من أسس منهجية البحث العلمي ،فالبحث العلمي عملية منهجية تهدف إلى اكتشاف معرفة جديدة أو تأكيد الفرضيات من خلال إجراءات دقيقة تُسهم في تفسير الظواهر الطبيعية والاجتماعية، وتُشكل أساليب القياس والتقويم في هذا النطاق الأدوات الأساسية التي تُمكن الباحثين من تحديد المتغيرات وتحليل البيانات والتأكد من جودة النتائج. وهناك دراسات سلطت الضوء في بعض جوانب أساليب القياس والتقويم،حيث أشارت دراسة (Rothwell, 2005) إلى أن سوء اختيار وبناء أدوات القياس قد يحد من قابلية تعميم النتائج، مما يقلل من قيمة الأبحاث في حل المشكلات العملية ، وهدفت دراسة (Rothwell, 2005) الى تحليل تأثير أساليب القياس على تعميم نتائج الأبحاث السريرية، وتمت مراجعة منهجية لـ 60 دراسة سريرية، مع تحليل الانحدار اللوجستي، وكشفت النتائج أن الأساليب المعيارية تزيد من التعميم بمستوى متوسط، تم التوصية بتوحيد أساليب القياس لضمان قابلية التطبيق والدقة وزيادة الثقة.وأشارت دراسة (Boud & Falchikov, 2006) الى العلاقة بين التقويم الذاتي وجودة الأبحاث في الجامعات، وتم استخدام منهج استقصائي شمل 120 باحثًا، مع تحليل باستخدام معامل الارتباط Spearman، وأوجدت النتائج علاقة إيجابية ضعيفة (r = 0.22) بين التقويم الذاتي والجودة، وأوصت الى دمج التقويم الذاتي مع مراجعة تقويم الأقران لتحسين النتائج ورفع مستوى الجودة. وقدموا (Moher et al, 2010) دراسة عن تأثير التقويم المنهجي على جودة الأبحاث الكمية في مجال الطب، وتم استخدام مراجعة منهجية لـ 50 دراسة طبية، مع تحليل إحصائي باستخدام اختبار Chi-square، وأظهرت النتائج أن التقويم المنهجي يحسن دقة النتائج بنسبة متوسطة، وأوصت الدراسة الى دمج التقويم في جميع مراحل البحث لضمان الجودة العالية.

وقام كل من (Brynjolfsson & McAfee, 2014) بدراسة استكشافية لكيفية تحسين التكنولوجيا والتقنية لأساليب التقويم في الأبحاث العلمية، وتم استخدام منهج تجريبي شمل 30 مشروعًا باستخدام برمجيات تحليل البيانات، وأظهرت النتائج أن التكنولوجيا واستخدام البرامج التقنية تقلل من الأخطاء بشكل ملحوظ وتزيد ثقة وجودة الأبحاث، وأوصت الى تعزيز استخدام البرمجيات المتقدمة في تقويم البيانات البحثية.وقد أظهرت دراسة (Open Science Collaboration, 2015) أن نسبة كبيرة من الأبحاث العلمية تعاني من صعوبات في التكرار، ويرجع ذلك جزئيًا إلى ضعف أساليب القياس المستخدمة أو عدم توحيدها، وهدفت دراسة (Open Science Collaboration, 2015) الى تحليل دور أساليب القياس في ضمان قابلية ثبات نتائج الأبحاث العلمية عند تكرارها، وتم استخدام منهجية التحليل الكمي لـ 73 دراسة تجريبية في مجال العلوم النفسية، مع مقارنة بين الأدوات الموحدة وغير الموحدة في القياسات، وأوجدت النتائج أن الأدوات التي لها وحدة قياس موحدة ومقننة تزيد من الثبات بنسبة ومستوى عالي، وأوصت الى تعزيز استخدام أدوات قياس موحدة لتحسين الثبات في النتائج.

وهدفت دراسة (DeVellis, 2016) الى تحليل تأثير جودة أدوات القياس على موثوقية النتائج في الأبحاث التجريبية في مجال علم النفس، وتم استخدام منهج التحليل الكمي لـ 60 دراسة تجريبية باستخدام اختبار الانحدار لتقييم العلاقة بين جودة الأدوات وثبات النتائج، وأظهرت النتائج أن الأدوات ذات الصدق والثبات العاليين تزيد من موثوقية النتائج بمستوى مرتفع، وقد أوصت الى ضرورة التحقق من خصائص القياس (الصدق والثبات) قبل استخدام الأدوات في الأبحاث.

وهدفت دراسة (Creswell & Plano Clark, 2017) الى تقييم جودة أساليب القياس والتقويم في الأبحاث التعليمية الحديثة، وتم العمل بمنهج تحليل محتوى لـ 40 دراسة تعليمية، مع التركيز على معايير الصدق والثبات، وأجدت النتائج أن الأساليب المدمجة (كمية ونوعية) تحسن من الجودة ومن مستوى الثقة في الدراسات بشكل جيد، وتم التوصية بالاعتماد على منهج مختلط في تصميم أدوات القياس والتقويم. وكشفت دراسة (Repko & Szostak, 2020)عن بعض الصعوبات الخاصة في الأبحاث متعددة التخصصات، حيث يواجه الباحثون صعوبات في تكييف أساليب القياس والتقويم مع المتطلبات المتنوعة للتخصصات المختلفة. وعملت دراسة (Repko & Szostak, 2020) الى استكشاف التحديات المرتبطة بتطبيق أساليب القياس في الأبحاث متعددة التخصصات، وتم الأخذ بأداة دراسة حالة شملت 10 مشاريع، مع تحليل نوعي لمقابلات مع الباحثين، وكشفت النتائج عن صعوبات في توحيد وتقنين الأدوات بسبب الاختلافات التخصصية، مما يقلل الدقة والجودة بمستوى منخفض، وأوصت تطوير أطر قياس مشتركة للأبحاث. وفي هذا الإطار تؤثر أساليب القياس والتقويم بشكل مباشر على جودة الأبحاث من خلال تأثيرها على عدة معايير مثل الصدق والثبات والتعميم، وأن انخفاض دقة أدوات الدراسة المستخدمة في قياس المتغيرات يُعد سببًا رئيسيًا لقلة الثقة في تعميم النتائج في العلوم النفسية. (Open Science Collaboration, 2015)

فالباحث وجد أن فعالية الأبحاث تعتمد على مدى دقة هذه الأساليب وملاءمتها لطبيعة الدراسة، وأن الأدوات ذات مستوى الصدق والثبات العالية تُعزز من ثقة النتائج، مما يُقلل من التشوه والتضليل في التفسير، وأيضا في الأبحاث النوعية من خلال استراتيجيات وأدوات الملاحظة وبرامج تكرار الكلمات وتحليل المحتوى التي تُعزز من مصداقية الاستنتاجات غير الكمية، فهذا التنوع في الاستخدامات يُظهر أن أساليب القياس والتقويم تتكيف مع متطلبات التخصصات المختلفة، مما يجعلها حجر الزاوية في بناء المعرفة العلمية.

ومن خلال ما تقدم يمكن تلخيص اهمية البحث بالاتي :

الأهمية النظرية:

- إثراء الأدبيات العلمية حول أساليب القياس والتقويم، حيث تُقدم تحليلًا شاملاً لدورها في تعزيز جودة الأبحاث العلمية عبر التخصصات.
- تزايد الاهتمام بقضايا الصدق والثبات لأدوات الدراسات والتعميم للنتائج.
- تُسهم الدراسة في سد الفجوات المعرفية المتعلقة بتأثير هذه الأساليب على المخرجات البحثية.
- تساهم الدراسة في تعزيز الوعي بأهمية هذه الأساليب بين الباحثين الجدد، مما يُعزز من مستوى الاحترافية في الممارسات البحثية ويُسهم في بناء مجتمع علمي أكثر دقة وموثوقية.

الأهمية العملية:

- تُوفر الدراسة إطارًا منهجيًا يُمكن الباحثين من اختيار وتطبيق أساليب قياس وتقويم فعالة، مما يعزز من كفاءة الأبحاث ويزيد من إمكانية تطبيق نتائجها في سياقات واقعية، مما يُبرز الحاجة إلى تصميم أساليب دقيقة وشفافة لتجنب الأخطاء وعدم الإيضاح في العملية البحثية.
- كما تُبرز الدراسة أهمية التكامل بين الأساليب التقليدية والتكنولوجيا الحديثة في تحسين جودة البحث العلمي، مما يدعم المؤسسات الأكاديمية في تطوير برامج تدريبية للباحثين.

حدود الدراسة:

تُعد حدود الدراسة العوامل التي قد تؤثر على تفسير النتائج أو قابلية تعميمها، وتنبع غالبًا من تصميم الدراسة، اختيار العينة، أو السياق الذي أُجريت فيه.(Creswell & Plano Clark, 2017)

وتمت هذه الدراسة في ضوء الحدود الآتية:

1. **الحدود الزمانية**: الفصل الأول من العام الدراسي 1446هـ.

2. **الحدود المكانية**: عدد من جامعات المملكة العربية السعودية (أم القرى- الملك عبد العزيز- الإسلامية بالمدينة المنورة).
3. **الحدود البشرية**: عينة من أعضاء هيئة التدريس في جامعات المملكة العربية السعودية (أم القرى- الملك عبد العزيز- الإسلامية بالمدينة المنورة).
4. **الحدود الموضوعية**: تأثير استخدام أساليب القياس والتقويم الدقيقة في تحسين جودة الأبحاث العلمية لعينة الدراسة أعضاء هيئة التدريس بجامعات المملكة العربية السعودية (أم القرى-الملك عبد العزيز- الإسلامية بالمدينة المنورة).

مصطلحات الدراسة:

أساليب القياس والتقويم الحديثة:

هي الأدوات والتقنيات المستخدمة لجمع البيانات وتحليلها وتقييم جودة العمليات والنتائج في الأبحاث العلمية النظرية والتطبيقية وتشمل هذه الأساليب :

1- **المقاييس الكمية** (مثل الاختبارات والاستبانات والبرمجيات الإحصائية)
2- **الأدوات النوعية** (مثل المقابلة والملاحظة والتحليل الموضوعي)، والتقنيات التكنولوجية والبرامج المتقدمة (مثل الذكاء الاصطناعي وتحليل البيانات الضخمة)، وتهدف هذه الأساليب إلى ضمان الدقة والموثوقية في قياس المتغيرات وتقييم النتائج.(Mertens, 2018)

التعريف الإجرائي:

هي معرفة مدى استخدام أساليب القياس والتقويم والأكثر شيوعاً في الاستخدام من خلال المقياس وعبر المقابلة وتقاس بالدرجة الكلية التي يحصل عليها افراد العينة (المستجيبين) عن فقرات اداة البحث (المقياس) ، وعبر اداة المقابلة التي سيتم استخدامهما في الدراسة الحالية .

جودة البحث العلمي:

هي مدى تحقيق الدراسة لمعايير الصدق، والثبات، والقابلية للتعميم، والموضوعية والشمولية في تصميمها وتنفيذها ونتائجها، وتُعد الجودة مؤشرًا لقدرة البحث على إنتاج معرفة موثوقة ومفيدة يمكن تعميمها أو تطبيقها في مواقف متنوعة، وتشمل معايير الجودة الصدق الداخلي والخارجي والشفافية في الطرح والنشر وتقليل التحيز. (Moher et al, 2010)

التعريف الإجرائي:

الدرجة الكلية التي يحصل عليها افراد العينة عند الاجابة عن فقرات مقياس جودة البحث العلمي المستخدم في الدراسة الحالية .

اطار نظري:

مقدمة:

يُعرف القياس بأنه عملية تحويل الخصائص المجردة إلى قيم قابلة للقياس باستخدام أدوات محددة، سواء كانت كمية مثل المقاييس أو نوعية مثل الملاحظات (Stevens, 1946)، أما التقويم فيُشير إلى عملية تقييمية مستمرة تُركز على الحكم على جودة العمليات أو المخرجات بناءً على معايير موضوعية والتغذية لهذه العمليات بشكل منظم. (Scriven, 1991).فالبعد التاريخي لتطور أساليب القياس والتقويم مر بمراحل تطورية كبيرة عبر التاريخ، مما يعكس تطور الفكر العلمي نفسه، ففي القرن السابع عشر، وضع غاليليو أسس القياس الكمي من خلال تجاربه على الدراسات، مما مهد الطريق لاستخدام الأدوات العددية في العلوم الطبيعية (Drake, 1978). وقدمت أنواع القياس (الاسمية، الترتيبية، الفترية، النسبية) التي أصبحت إطارًا أساسيًا لتصميم الأدوات في العلوم الحديثة، وتطورت الأساليب من التقييمات البسيطة في العصور الوسطى إلى نماذج معقدة في القرن العشرين، الذي ركز على تقييم الأداء على مستويات متعددة (Kirkpatrick & Kirkpatrick, 2006).

فقد وجد الباحث أن ظهور الإحصاءات أدى إلى تعزيز دقة القياس، مثل أساليب تحليل التباين (ANOVA) التي سمحت بتقييم العلاقات بين المتغيرات بدقة أكبر، فهذه التطورات التاريخية تُظهر كيف أصبحت أساليب القياس والتقويم أكثر تعقيدًا وتكاملاً، مما يعكس الحاجة المتزايدة إلى أدوات دقيقة لمواجهة تعقيدات الأبحاث الحالية.ويلعب التقويم دورًا حاسمًا في ضمان الشفافية وتقليل التحيز، فقد تمت الإشارة إلى أن التقويم المنهجي، كما في إرشادات CONSORT، يُساعد في تحديد نقاط القوة ويعزز من امكانيتها ونقاط الضعف ويعالجها في الدراسات الطبية، مما يُحسن من جودة التقارير البحثية. (Moher et al, 2010)

كما يلاحظ أن التقويم الذاتي عندما يُعزز بمعايير خارجية يُسهم في تحسين جودة الأبحاث الأكاديمية، لكنه قد يفتقر إلى الموضوعية إذا أُسيء استخدامه. (Boud & Falchikov, 2006)

ولذلك يجب الاخذ بالحسبان إلى أن التقييمات غير المنظمة أو تلك التي تعتمد على معايير في التقويم غامضة قد تُؤدي إلى تحيز التأكيد (Confirmation Bias)، حيث يميل الباحثون إلى تفسير النتائج بما يتماشى مع توقعاتهم. (Scriven, 1991)

ومن الضروري معرفة أن الاعتماد على أساليب قياس تقليدية في الأبحاث الاجتماعية والانسانية قد يُؤدي إلى انخفاض دقة البيانات مقارنة بالتقنيات والبرمجيات الحديثة، وأن اختيار مقاييس لا تتناسب مع طبيعة الظاهرة قد يُضعف من الصدق والثبات، مما يُعيق التفسير الصحيح للنتائج.

(Trochim, 2020) وبناء على ذلك فأن غياب معايير واضحة قد يُؤدي إلى تقويم غير متسق، وخاصة في الأبحاث التي تعتمد على الحكم النوعي، وأن تعقيد الأدوات التكنولوجية الحديثة وإصدار تحديث مستمر يتطلب مهارات تحليلية متقدمة، مما يُشكل تحديًا للباحثين غير المدربين، وعليه فهذه العوائق تُظهر الحاجة الماسة إلى استراتيجيات مبتكرة لتحسين التطبيق العملي لهذه الأساليب. (Brynjolfsson & McAfee, 2014) وفي نفس الصدد فأن الجمع بين الأساليب الكمية والنوعية في العلوم الإنسانية يُمكن أن يُعزز ويرفع من جودة الأبحاث العلمية، حيث يُوفر رؤية أكثر شمولية وموضوعية، وهذا النهج المختلط يُبرز قدرة أساليب القياس والتقويم على التكيف مع الطبيعة المعقدة للظواهر الإنسانية، مما يُعزز من أهميتها في هذا المجال.(Creswell & Plano Clark, 2017)

لذلك ينبغي الانتباه عند استخدام أدوات قياس غير دقيقة قد يُؤدي إلى نتائج مضللة تُؤثر على اتخاذ القرارات ومعيار الجودة في مجالات مثل الطب أو السياسة العامة، مما يُشكل خطرًا على المجتمع، وكذلك التقويم غير العادل أو المتحيز قد يُؤدي إلى استبعاد بعض الوسائل الجيدة والمنافع المعرفية، مما تُنتهك مفاهيم الانصاف والعدل في استخلاص العلم.(Mertens, 2018)

وفي هذا الإطار يجد الباحث أن الشفافية في توثيق أساليب القياس والتقويم تُعد ضرورة لضمان معايير القيم الأكاديمية، حيث تُتيح للمجتمع العلمي مراجعة النتائج والاستفادة منها بشكل متكرر، وهذه الاعتبارات تُبرز أن دور هذه الأساليب لا يقتصر على الجوانب التقنية، بل يمتد إلى الالتزام والوفاء والحكمة التي تظهر الأعمال والدراسات البحثية بشكل مجود.فقد أحدث التقدم التكنولوجي ثورة في أساليب القياس والتقويم، مما عزز من قدرتها على مواجهة التحديات البحثية، فالذكاء الاصطناعي وتحليل البيانات الضخمة أصبحا أدوات رئيسية في تحسين دقة القياس، خاصة في المجالات التي تتطلب معالجة كميات هائلة من المعلومات على سبيل المثال، تم تحديث أدوات للذكاء الاصطناعي متطورة حسّنت من دقة قياس مخرجات وبيانات الأبحاث، مما دعم اتخاذ قرارات أكثر فعالية. (Brynjolfsson & McAfee, 2014) وأصبحت البرمجيات مثل NVivo وSPSS تُستخدم لتقييم البيانات بطرق تجمع بين المنهج الكمي والنوعي (الكيفي)، مما يُوفر تحليلاً أكثر عمقًا وتفصيلاً، وهذه التطورات تُظهر أن التكنولوجيا ليست مجرد أداة مساعدة، بل تُعيد تشكيل الطريقة التي تُطبق بها هذه الأساليب في الأبحاث الحديثة.(Creswell & Plano Clark, 2017)

اولا:المفاهيم

1-أساليب القياس والتقويم الحديثة:

فهي الأدوات والتقنيات المستخدمة لجمع البيانات وتحليلها وتقييم جودة العمليات والنتائج في الأبحاث العلمية النظرية والتطبيقية، وتشمل هذه الأساليب المقاييس الكمية (مثل الاختبارات والاستبانات والبرمجيات الإحصائية)، والأدوات النوعية (مثل المقابلة والملاحظة والتحليل الموضوعي)، والتقنيات التكنولوجية والبرامج المتقدمة (مثل الذكاء الاصطناعي وتحليل البيانات الضخمة)، وتهدف هذه الأساليب إلى ضمان الدقة والموثوقية في قياس المتغيرات وتقييم النتائج.(Mertens, 2018)

وتأتي منها أساليب التقويم الحديثة مثل التقويم الذاتي والتقويم المشترك فهي تُعزز من إعطاء روح المشاركة وابداء الرأي للمشاركين وذو العلاقة، ولكنها تحتاج وتتطلب معايير وشروط صارمة لضمان الموضوعية. (Boud & Falchikov, 2006)

ويلاحظ كذلك في أدوات القياس المدعومة بالتكنولوجيا والبرامج الحديثة، مثل الخوارزميات التنبؤية، تُقلل من مستوى أخطاء القائمين عليها، وفي المقابل تحتاج مهارات فنية وممارسين في هذه البرامج. (Brynjolfsson & McAfee, 2014)

2-جودة البحث العلمي:

وهي مدى تحقيق الدراسة لمعايير الصدق، والثبات، والقابلية للتعميم، والموضوعية والشمولية في تصميمها وتنفيذها ونتائجها، وتُعد الجودة مؤشرًا لقدرة البحث على إنتاج معرفة موثوقة ومفيدة يمكن تعميمها أو تطبيقها في مواقف متنوعة، وتشمل معايير الجودة الصدق الداخلي والخارجي والشفافية في الطرح والنشر وتقليل التحيز. (Moher et al, 2010)

ومما لاشك فيه أن ضعف مبدأ التعميم في أبحاث العلوم النفسية والاجتماعية يعود جزئيًا إلى تدنئي مستوى جودة وقوة أساليب القياس والتقويم، مما يتطلب بناء معايير أكثر شمولية ودقة. (Open Science Collaboration, 2015)

وبطبيعة الحال فالتحيزات في تصميم الدراسات واختيار الأدوات وبناءها قد تُضعف جودة الدراسات، مما يستلزم استخدام أساليب ذات خصائص سيكومترية قوية.(Ioannidis, 2005)

ثانيا:النظريات المفسرة لمتغيرات الدراسة

أ)مفهوم اساليب القياس والتقويم

1-نظرية القياس الكلاسيكية (Classical Test Theory - CTT):

تُعد هذه النظرية، التي طُورت في أوائل القرن العشرين، أساسًا رئيسياً في تصميم أدوات القياس وبنائها، وهي تفترض أن درجة الفرد في اختبار ما تتكون من القيمة الحقيقية مضافًا إليها خطأ القياس العشوائي، وتُستخدم هذه النظرية لتقييم خصائص الصدق والثبات في المقاييس الكمية، مثل الاستبيانات والمقاييس المستخدمة في الأبحاث العلمية.(Nunnally & Bernstein, 1994).

2-نظرية الاستجابة للبند (Item Response Theory - IRT):

تُعتبر نقلة في علم القياس وتطورًا حديثًا لنظرية القياس، وتركز على تحليل قدرات الأفراد على استجاباتهم للبنود الفردية في المقاييس، وتُستخدم هذه النظرية في تصميم أدوات قياس دقيقة تتكيف مع مستويات القدرة المختلفة للأفراد المعنيين، مثل الاختبارات التكيفية بمساعدة الحاسوب .(DeVellis,2016)

3-النظرية التفسيرية التتابعية في التقويم (Interpretive Evaluation Theory):
تم تطويرها على استكشاف المواقف والملاحظات المتعددة والتركيز على الأساليب النوعية التي تعطي تصورات واضحة، وتُطبق في تقييم جودة الأبحاث والدراسات العلمية من خلال تحليل المقابلات أو الملاحظات والرصد بذلك، مما يُعزز فهم التحديات والصعوبات التي تواجه المشكلة البحثية.(Lincoln& Guba, 1985)

ب) مفهوم جودة الابحاث العلمية

1-النظرية البنائية المعرفية في العلوم (Scientific Constructivism):
قامت هذه النظرية على افترض أن جودة الأبحاث والدراسات تتأثر بالإطار النظري والمنهجية المستخدمة والتي يعتمد عليها، وتُبرز أهمية خصائص الصدق البنائي في ضمان أن الأدوات تقيس المتغيرات المستهدفة بدقة، مما يدعم يزيد من تحقيق نتائج مفيد، وتُطبق هذه النظرية في تقييم مدى ملاءمة أساليب القياس والتقويم لأهداف الدراسة.(Kuhn, 1962)

2-نظرية قابلة التعميم (Reproducibility Framework):
طُورت هذه النظرية في سياق أزمة التعميم في الدراسات البحثية، وتُركز على أن جودة الأبحاث تُقاس بقدرتها على تحقيق نتائج منطقية وواضحة في ميدان البحث عند إعادة تطبيق الأدوات مرة أخرى، وتُؤكد على أهمية خصائص الثبات في أدوات الدراسة ومنهجيتها، مثل استخدام أساليب قياس وتقويم ثابتة وتم التحقق منها. (Open Science Collaboration, 2015)

نظرية تقويم المنهجية (Systematic Evaluation Theory):
تركز هذه النظرية على استمرار تقييم جودة الأبحاث وذلك من خلال معايير علمية محددة، مثل خصائص الصدق الداخلي (Internal Validity) والصدق الخارجي (External Validity) لأدوات ومنهجية الدراسة، وتُطبق في تحديد مدى تأثير أساليب القياس والتقويم على صحة النتائج، وخاصة في الأبحاث التجريبية.(Scriven, 1991)

ويرى الباحث أن التفاعل بين أساليب القياس والتقويم الحديثة وجودة البحث العلمي بشكل وثيق يعزز من موضوعية أدوات الدراسة وهما ركيزتان لرفع مستوى الأبحاث العلمية، فنجد أن استخدام برمجيات مثل NVivo في التحليل النوعي يُحسن من المصداقية، بينما تُقلل الأدوات المدعومة بالذكاء الاصطناعي من التحيزات، فجودة الأبحاث العلمية تتطلب أساليب تقويم منهجية تُراعي المواقف المتعددة والظروف المحيطة، فالنظريات مثل CTT وIRT تُساعد في تصميم أدوات دقيقة، بينما تدعم نظريات التعميم والتقويم المنهجي تقييم لجودة الأبحاث بشكل شامل.

منهجية الدراسة واجراءاتها

المنهج:
تتطلب الدراسة الحالية إلى تحليل دور أساليب القياس والتقويم في تعزيز جودة الأبحاث العلمية، نهجًا يجمع بين القياسات الإحصائية والتفسيرات لفهم العلاقات الكمية والتحديات النوعية، لذلك تم الاعتماد المنهج المختلط (Mixed Methods Research) باستخدام التصميم التفسيري المتتابع (Sequential Explanatory Design).(Creswell & Plano Clark, 2017) .
ويبدأ هذا التصميم بجمع وتحليل البيانات الكمية لتحديد الأنماط والعلاقات بين المتغيرات، يليه جمع البيانات النوعية لتوضيح النتائج واستكشاف التفاصيل غير القابلة للقياس عدديًا.
يُتيح المنهج الكمي تقييم تأثير أساليب القياس والتقويم على معايير الجودة مثل الصدق والثبات، بينما يُساهم المنهج النوعي في استكشاف التحديات العملية والمنهجية التي يواجها الباحثون، هذا التكامل يضمن الإجابة على أسئلة الدراسة بشكل جيد.

مجتمع الدراسة:
يشمل مجتمع الدراسة الباحثين وأعضاء هيئة التدريس الأكاديميين من مختلف التخصصات (العلوم الطبيعية، الاجتماعية، التطبيقية، والإنسانية) الذين يُجرون أبحاثًا منشورة في مجلات علمية محكمة أو يعملون على مشاريع بحثية جارية من جامعات المملكة العربية السعودية والمراكز البحثية التابعة لوزارة التعليم لضمان التنوع في الخبرات البحثية.

عينة الدراسة:
العينة الكمية: سيتم اختيار عينة مكونة من 350 باحثًا باستخدام أسلوب العينة العشوائية الطبقية (Stratified Random Sampling)، تُقسم العينة إلى طبقات بناءً على التخصصات، وبناء على ذلك تم تحديد حجم العينة بناءً على معادلة Cochran لحساب العينات في المجتمعات الكبيرة. (Cohen et al, 2018)

العينة النوعية: تم اختيار 15 باحثًا باستخدام العينة القصدية (Purposive Sampling)، مع التركيز على الباحثين ذوي الخبرة الطويلة (أكثر من 10 سنوات) في تصميم أو تطبيق أساليب القياس والتقويم، لضمان عمق المعلومات النوعية والكيفية. (Patton, 2015)

أدوات الدراسة:

1- مقياس دور أساليب القياس والتقويم على الأبحاث العلمية
أعد الباحث المقياس بالرجوع الى عدد من الدراسات والأدبيات السابقة مثل دراسة (DeVellis, 2016) ودراسة (Repko & Szostak, 2020) ودراسة (Rothwell, 2005) ودراسة (Creswell & Plano Clark, 2017) ، وبناء على هذه المصادر تم بناء أداة الدراسة، وتكونت في بدايتها بأسئلة ديموغرافية للمستجيبين ومن ثم ثلاثة أبعاد رئيسية: (1) أساليب القياس والتقويم الشائعة، (2) تأثير أساليب القياس والتقويم على جودة الأبحاث الصدق والثبات، (3) التحديات المرتبطة باستخدام أساليب القياس والتقويم ، وتكونت الأداة من 23 بندًا وقد تم تحديد الاستجابات على كل مقياس وفق مقياس ليكرت الخماسي (مرتفعة جداً – مرتفعة – متوسطة– منخفضة– منخفضة جداً)، وتقابل الدرجات الآتية (5، 4، 3، 2، 1) على الترتيب، وينحصر بذلك المتوسط الحسابي بين (1-5)، ولتحديد الدرجة النهائية فقد تم تحويل الاستجابات عند التحليل إلى فئات خماسية، إذ تم تحديد المدى من خلال طرح أعلى متوسط حسابي من أقل متوسط حسابي (5-1= 4) وقسمة الناتج على عدد فئات الدرجات (4÷5= 0.80)، ويبين الجدول الآتي المتوسطات الحسابية المعتمدة في الحكم على درجة الاستجابات.

الصدق (Validity):
الصدق يُشير إلى مدى قياس الأداة لما صُممت لقياسه. (Trochim, 2020)، لضمان الصدق في المقياس، تم التحقق من صدق المحتوى (Content Validity) من خلال عرضه وتحكيمه من 5 خبراء من أعضاء هيئة التدريس في الجامعات لتقييم وضوح البنود وملاءمتها لأهداف الدراسة، وتم تعديل صياغة بعض البنود لغوياً بناءً على مقترحاتهم لضمان الفهم السليم من قبل المستجيبين، وأيضا تم التحقق من الصدق البنائي (Construct Validity) وذلك من خلال استخدام التحليل العاملي الاستكشافي (Exploratory Factor Analysis - EFA) عند تطبيق الدراسة الاستطلاعية على عينة مكونة من 35 باحثًا، الهدف هو التأكد من أن البنود تتجمّع تحت الأبعاد الثلاثة المحددة، وبلغت قيمة تشبعات الفقرات على العوامل (Factor Loading) أعلى من (0.4). (Field, 2018).

(صدق التلازمي او التقاربي) (Convergent Validity): تم قياس الارتباط بين بنود المقياس ومقياس معياري خارجي وتم استخدام مقياس جودة البحث العلمي من دراسة (Open Science

Collaboration, 2015) باستخدام معامل بيرسون (Pearson Correlation)، وقد وجدت معاملات الارتباط مرتفعة ودالة احصائياً (من 0.79 فأكثر).

الثبات (Reliability):

الثبات يُقيس مدى اتساق الأداة في إعطاء نتائج متشابهة عند التكرار. (Field, 2018) ولضمان ثبات المقياس، ستُطبق الإجراءات التالية:

-طريقة الاتساق الداخلي (Internal Consistency): تم حساب معامل ألفا كرونباخ (Cronbach's Alpha) لكل بعد من أبعاد الاستبانة وللأداة ككل خلال الدراسة الاستطلاعية، وبلغت قيم معاملات الثبات باستخدام معامل ألفا كرونباخ قد تراوحت بين (0.73-0.87) لأبعاد مقياس الامتنان، وبلغت قيمة معامل ألفا كرونباخ لمقياس الامتنان ككل (0.84) وتعد هذه المعاملات مرتفعة، وتشير إلى أن المقياس يتصف بدرجة ثبات مقبولة عند تطبيقه في الدراسة الحالية، وهي القيمة المقبولة عالميًا للثبات (Nunnally & Bernstein, 1994)

-طريقة الاختبار و إعادة الاختبار (Test-Retest Reliability): تم إعادة تطبيق المقياس على عينة فرعية مكونة من 30 باحثًا بعد أسبوعين من التطبيق الأول. وبلغت قيمة معامل الارتباط بين الاختبارين باستخدام معامل بيرسون (0.83) وهي قيمة مقبولة لتحديد ما إذا كان المستجيب سيحقق نفس الأداء الجيد في الإجابة مرة أخرى في ظروف مماثلة مع الوقت. (Trochim, 2020)

2- المقابلات شبه المنظمة:

أداة نوعية تتضمن أسئلة مفتوحة، وتم اجراء المقابلات وجهاً لوجه، مع مراعاة المرونة في تعديل الأسئلة بناءً على نتائج الاستبانة، وكانت المدة من 45-60 دقيقة لكل مقابلة.

والغرض منها استكشاف التحديات والاقتراحات المتعلقة بتطبيق أساليب القياس والتقويم، وتم إجراء المقابلات مع عينة قصدية من ذوي الخبرة (17 باحثًا) بعد تحليل البيانات الكمية الأولية، لتوجيه الأسئلة بناءً على النتائج.

الصدق (Validity):

في الدراسات النوعية، يُستبدل مفهوم الصدق التقليدي بمعايير مثل المصداقية (Credibility) والقابلية للنقل(Transferability). (Lincoln& Guba, 1985) لضمان هذه الخصائص:

المصداقية (Credibility):

تم استخدام استراتيجية التثليث (Triangulation) من خلال مقارنة بيانات المقابلات مع نتائج الاستبانة لتأكيد اتساق النتائج.

وتم الاستعانة بمراجع خارجي مختص (Peer Debriefing) لمراجعة الأسئلة والموضوعات المستخلصة لضمان الحيادية والدقة في التفسير، وتم إعادة عرض النتائج المستخلصة على المستجيبين (Member Checking) للتحقق من دقتها وتمثيلها لوجهات نظرهم.

القابلية للنقل (Transferability):

تم توفير وصف تفصيلي لسياق المستجيبين (مثل التخصصات، الخبرة) في التقرير النهائي، مما يُتيح تقييم مدى إمكانية تطبيق النتائج في مواقف أخرى (Patton, 2015).

الصدق الداخلي (Internal Validity):

تم تصميم الأسئلة بناءً على الدراسات السابقة كدراسة (Repko & Szostak, 2020) ودراسة (Rothwell, 2005) لضمان ارتباطها بموضوع الدراسة، مع اختبارها في مقابلة استطلاعية مع 3 باحثين للتأكد من وضوح الأسئلة وقابليتها لاستخلاص بيانات ذات صلة، وتم تحليل الردود الأولية لتحديد أي تعديلات ضرورية (مثل إعادة صياغة الأسئلة غير الواضحة) قبل التطبيق النهائي.

الثبات (Reliability):
في الدراسات النوعية، يُركز الثبات على الاعتمادية (Dependability)، أي مدى اتساق الإجراءات (Lincoln & Guba, 1985). ولضمان ذلك:

الاتساق في الإجراءات:
تم استخدام دليل مقابلة موحد يحتوي على أسئلة رئيسية وفرعية، مع السماح بالمرونة لاستكشاف الردود، واستخدمه الباحث في إجراء المقابلات لضمان التوحيد في طرح الأسئلة (Patton, 2015).

التدقيق الخارجي (External Audit):
وتم الاستعانة من خبير مختص مستقل مراجعة عملية جمع البيانات والتحليل (مثل الرموز والموضوعات) للتأكد من خلوها من التحيز ولضمان اتساق النتائج.

توثيق العملية:
تم تسجيل جميع المقابلات وتفريغها نصيًا (Transcription) بدقة، مع الاحتفاظ بسجل تفصيلي لكل خطوة في التحليل الموضوعي لضمان إمكانية التكرار (Braun & Clarke, 2006).

3- التكامل بين أدوات الدراسة:
نظرًا لاستخدام المنهج المختلط، سيتم ضمان الخصائص السيكومترية للدراسة ككل من خلال التكامل بين الاستبانة والمقابلات:

الصلابة (Rigor): سيتم مقارنة النتائج الكمية (مثل تأثير الأساليب على الموثوقية) مع النتائج النوعية (مثل التحديات المُبلغ عنها) للتأكد من الاتساق والتكامل (Tashakkori & Teddlie, 2010).

التوثيق: تم عمل تقرير مفصل عن الخصائص السيكومترية لكل أداة، مع الإشارة إلى القيم الإحصائية (ألفا كرونباخ، معاملات الارتباط) وإجراءات التحقق النوعية.

عرض نتائج الدراسة ومناقشتها:

نتائج السؤال الاول:ما هي أساليب القياس والتقويم الأكثر استخدامًا في الأبحاث العلمية عبر التخصصات المختلفة؟

للإجابة عن هذا السؤال تم استخدام التكرارات والنسب المئوية لتحديد الأساليب الأكثر شيوعًا من بيانات المقياس، وتم إضافة اختبار Chi-square لمعرفة ما إذا كان هناك اختلافات ذات دلالة إحصائية بين التخصصات (العلوم الطبيعية، الاجتماعية، الإنسانية)، واستخدام المتوسط الحسابي والانحراف المعياري لتعزيز النتائج والجدول (1) يوضح ذلك:

الجدول(1)
التكرارات والنسب المئوية لتحديد الأساليب الأكثر شيوعًا

م	الأساليب/ العبارات	التكرار (n=350)	النسبة المئوية العامة (%)	العلوم الطبيعية (%)	العلوم الاجتماعية (%)	العلوم الإنسانية (%)	المستوى
1	وجود فعالية في استخدام المقاييس الكمية في الأبحاث العلمية	280	80.0	92.0	76.0	68.0	مرتفع
2	توجد أهمية للتحليل الاحصائي في نتائج الدراسات	245	70.0	88.0	72.0	50.0	مرتفع
3	تعطي المقابلات النوعية تصور واضح للدراسة	175	50.0	32.0	58.0	70.0	متوسط
4	تساعد الملاحظة في تفسير الأحداث وتسلسلها	140	40.0	28.0	44.0	56.0	متوسط
5	يساعد التقويم الذاتي في زيادة المعرفة والتصورات الداخلية	105	30.0	20.0	34.0	38.0	منخفض
6	توجد أهمية لبرمجيات والتطبيقات لعمل التقويم(NVivo)	98	28.0	24.0	30.0	32.0	منخفض
	Chi-square	$\chi^2(10, N=350) = 45.32, p<0.001$ مما يشير إلى وجود فروق ذات دلالة إحصائية بين التخصصات في استخدام الأساليب					

يتبين من الجدول (1) أن وجود فعالية في استخدام المقاييس الكمية في الأبحاث العلمية تتصدر الأساليب المستخدمة بنسبة 80%، مع انتشار أعلى في العلوم الطبيعية (92%) مقارنة بالإنسانية (68%)، مما يعكس اعتماد هذه التخصصات على القياسات العددية لضمان الدقة في التجارب، وتوجد أهمية للتحليل الاحصائي في نتائج الدراسات بنسبة (70%) وله انتشار كبير، خاصة في العلوم الطبيعية (88%)، نظرًا للحاجة إلى معالجة البيانات الضخمة، وفي المقابل تبرز وتعطي المقابلات النوعية تصور واضح للدراسة بنسبة (50%) و تساعد كذلك الملاحظة في تفسير الأحداث وتسلسلها بنسبة (40%) كأدوات لها تفضيل في العلوم الإنسانية وأعطت نسب (70% و56% على التوالي)، مما يشير إلى أهميتها في استكشاف الجوانب والمواقف الحياتية وتفسيرها، ويساعد التقويم الذاتي في زيادة المعرفة والتصورات الداخلية بنسبة (30%) وتوجد أهمية لبرمجيات والتطبيقات لعمل التقويم (NVivo)بنسبة (28%) وهي أقل استخدامًا، ربما بسبب التحديات التقنية أو الموارد المحدودة، ونجد الفروق بين التخصصات يؤكد أن اختيار الأساليب يتأثر بطبيعة المجال العلمي .

الجدول (2)

المتوسطات الحسابية والانحرافات المعيارية لأساليب القياس والتقويم وفق التخصص الدراسي

ت	الأساليب/ العبارات	المتوسط	الانحراف المعياري	العلوم الطبيعية		العلوم الاجتماعية		العلوم الإنسانية	
				المتوسط	الانحراف المعياري	المتوسط	الانحراف المعياري	المتوسط	الانحراف المعياري
1	وجود فعالية في استخدام المقاييس الكمية في الأبحاث العلمية	4.20	0.85	4.60	0.70	4.10	0.90	3.80	0.95
2	توجد أهمية للتحليل الاحصائي في نتائج الدراسات	3.95	0.90	4.50	0.65	4.00	0.85	3.20	1.00
3	تعطي المقابلات النوعية تصور واضح للدراسة	3.50	1.10	2.90	1.05	3.70	1.00	4.00	0.95
4	تساعد الملاحظة في تفسير الأحداث وتسلسلها	3.20	1.15	2.80	1.10	3.40	1.05	3.60	1.00
5	يساعد التقويم الذاتي في زيادة المعرفة والتصورات الداخلية	2.80	1.20	2.50	1.15	3.00	1.10	3.10	1.05
6	توجد أهمية لبرمجيات والتطبيقات لعمل التقويم (NVivo)	2.60	1.25	2.70	1.20	2.80	1.15	2.90	1.10

يتبين من الجدول(2) أن المتوسط الحسابي لوجود فعالية في استخدام المقاييس الكمية في الأبحاث العلمية هو الأعلى بقيمة (**4.20**) ويأتي بعده بمتوسط حسابي (**3.95**) وجود أهمية للتحليل الاحصائي في نتائج الدراسات وهو ما يريده عنه كثير الباحثين لأجل زيادة موضوعية ورفع مستوى منطقية الدراسات لديهم، وتأتي بمتوسط حسابي (2.80) مساعدة التقويم الذاتي في زيادة المعرفة والتصورات الداخلية وكذلك وجود أهمية لبرمجيات والتطبيقات لعمل التقويم(NVivo) بمتوسط أقل (2.60) وذلك لصعوبة أمتلاك المهارات اللازمة وتعقيد الخطوات وصرامة المعايير المتبعة.

وتتماشى هذه النتائج مع دراسة (Open Science Collaboration, 2015)، التي أظهرت هيمنة الأساليب الكمية في العلوم التجريبية لدعم التعميم على نتائجها، وكما تتفق مع دراسة (Repko & Szostak, 2020)، حيث تُفضل الأساليب النوعية في الأبحاث التي تركز على التفسير والمعاني، خاصة في العلوم الإنسانية، ومع ذلك، تُضيف الدراسة الحالية بُعدًا جديدًا من خلال تحديد انتشار برمجيات التقويم، وهو جانب لم يُركز عليه بقوة في الأدبيات السابقة.

نتائج السؤال الثاني :كيف تؤثر جودة أساليب القياس والتقويم على الخصائص (الصدق والثبات) نتائج الأبحاث العلمية؟

للإجابة عن هذا السؤال تم استخدام تحليل الانحدار المتعدد لفحص تأثير جودة أساليب القياس والتقويم (متغير مستقل) على الصدق والثبات (متغيرات تابعة)، وإضافة اختبار t المستقل لمقارنة التخصصات، وتم التحقق من التوزيع الطبيعي باستخدام اختبار Kolmogorov-Smirnov ((p>0.05).

الجدول(3)

تحليل الانحدار المتعدد لفحص تأثير جودة أساليب القياس والتقويم

المتغير التابع	معامل الانحدار (β)	الخطأ المعياري(SE)	الدلالة(p)	R^2	درجات الحرية (df)	F
الصدق	0.65	0.08	< 0.001	0.42	348	84.56
الثبات	0.61	0.09	< 0.001	0.37	348	72.19

يتبين من الجدول (3) وجود علاقة إيجابية قوية بين جودة أساليب القياس والتقويم وخاصية الصدق (β = 0.65، p< 0.001)، حيث تُفسر 42% من التباين، مما يشير إلى أن الأدوات المتقنة (ذات خصائص الصدق المرتفعة) تُقلل من الأخطاء في قياس المتغيرات، كذلك تؤثر الجودة على الثبات (β = 0.61، p< 0.001)، مع تفسير 37% من التباين، مما يعني أن الأساليب التي لها خاصية الثبات العالية تُنتج مخرجات متسقة عند إعادة تطبيقها مره أخرى .

وتم عمل اختبار T لمقارنة التخصصات، وكانت النتائج كما موضحة في الجدول (4):

الجدول(4)

اختبار T لمقارنة التخصصات

المتغير	المجموعة	المتوسط (M)	الانحراف المعياري (SD)	قيمة (T)	درجات الحرية (df)	الدلالة الإحصائية (p)
الصدق	العلوم الطبيعية	4.1	0.6	3.12	348	0.002
	العلوم الإنسانية	3.8	0.7			
الثبات	العلوم الطبيعية	4.0	0.5	2.89	348	0.004
	العلوم الإنسانية	3.7	0.6			

يتبين من الجدول (4)

-أن **خاصية الصدق:** تُظهر النتائج فروق ذات دلالة إحصائية بين العلوم الطبيعية بمتوسط 4.1 وانحراف معياري 0.6 والعلوم الإنسانية بمتوسط 3.8 وانحراف معياري 0.7 حيث قيمة T = 3.12، وذلك يشير إلى أن الباحثين في العلوم الطبيعية يرون تأثيرًا أقوى لأساليب القياس والتقويم على خصائص الصدق والثبات مقارنة بالعلوم الإنسانية.

-**خاصية الثبات:** يوجد فرق ذو دلالة إحصائية بين العلوم الطبيعية بمتوسط 4.0 وانحراف معياري 0.5 والعلوم الإنسانية بمتوسط 3.7 وانحراف معياري 0.6 حيث قيمة T = 2.89، وذلك يعني أن الأساليب تُسهم بشكل أكبر في تعزيز ثبات وإعادة التطبيق في العلوم الطبيعية مقارنة بالعلوم الإنسانية، وربما بسبب اعتماد الأولى على القياسات الموضوعية بشكل أكبر، وهذه النتائج تؤكد أن تحسين جودة الأدوات يُعد استثمارًا حاسمًا لتعزيز المصداقية العلمية.

وتتفق هذه النتائج مع دراسة (DeVellis, 2016) والذي وجد أن أدوات الدراسات ذات الخصائص السيكومترية عالية الجودة تُعزز الثبات بنسبة كبيرة، كما تتماشى مع دراسة (Moher et al, 2010) حيث أشاروا إلى أن التقويم المنهجي يُحسن من صدق الأدوات وذلك من خلال تقليل التحيز، وتأتي الدراسة الحالية تُوسع هذه الرؤى بإظهار الفروق بين التخصصات، وهو جانب لم يُركز عليه بقوة في الأدبيات السابقة.

نتائج السؤال الثالث :ما هي التحديات الرئيسية التي تواجه الباحثين عند تطبيق أساليب القياس والتقويم في الأبحاث متعددة التخصصات؟

للإجابة عن هذا السؤال تم التحليل الموضوعي لبيانات مقابلات باستخدام برنامج NVivo، مع دعم الإجابة ببيانات المقياس الكمية بالتحليل الوصفي لـ 8 عبارات في المقياس تتعلق بالتحديات المرتبطة باستخدام أساليب القياس والتقويم والجدول (5) يوضح نتائج التحليل

الجدول (5)

التحليل الموضوعي والكمي لتحديات المرتبطة باستخدام أساليب القياس والتقويم

التحليل الموضوعي للمقابلة (النوعية)			التحليل الوصفي لعبارات المقياس (الكمية)		
م	مضمون الأسئلة في المقابلة والمقياس	التكرار (المقابلة)	النسبة المئوية (المقياس)	المتوسط (M)	الانحراف المعياري(SD)
1	صعوبة توحيد الأدوات بين التخصصات	13	68.6	4.2	0.8
2	نقص المهارات التقنية	11	62.3	4.0	0.9
3	تعقيد الظواهر متعددة التخصصات	10	57.1	3.9	1.0
4	ضعف وقلة الموارد (الوقت-التكلفة)	8	48.0	3.7	1.1

يتبين من الجدول (5) وجود تحدي وهو صعوبة في توحيد الأدوات بتكرار في المقابلة بعدد (13) وبنسبة (68.6%) بمتوسط (4.2) وتُعد العائق الأبرز، مما يعكس التنوع المنهجي بين التخصصات، حيث تتطلب العلوم الطبيعية قياسات دقيقة بينما تُركز الإنسانية على التفسير، ويظهر التحدي الثاني

في نقص المهارات التقنية بتكرار في المقابلة بعدد (11) وبنسبة (62.3%) بمتوسط (4.0) يُشير إلى فجوة في التدريب على الأدوات الحديثة، مما يُعيق الاستفادة من التكنولوجيا، كما يأتي التحدي الثالث وهو تعقيد الظواهر متعددة التخصصات بتكرار في المقابلة بعدد (10) وبنسبة (57.1%) وبمتوسط (3.9) ويُبرز هذا تحدي قياس المتغيرات المعقدة التي تتجاوز حدود تخصص واحد، بينما يُعد تحدي ضعف وقلة الموارد بتكرار في المقابلة بعدد (8) وبنسبة (48%) وبمتوسط (3.7) عائقًا بمستوى أقل من سابقيه ولكنه يؤثر على الكفاءة، وهذه التحديات تتطلب حلولاً شاملة تشمل التدريب وتطوير الأدوات.وتتسق هذه النتائج مع دراسة (Repko & Szostak, 2020) اللذان حددا صعوبة توحيد أدوات الدراسة والأساليب كمشكلة رئيسية في الأبحاث العلمية، كما تؤكد دراسة (Brynjolfsson & McAfee, 2014) حول الحاجة إلى مهارات تقنية متقدمة ومواكبة للأساليب الحديثة، وهنا يجد الباحث أن الدراسة الحالية تُضيف تحليلًا كميًا لشدة هذه التحديات، مما يُعزز الأدلة المقدمة سابقًا.

نتائج السؤال الرابع :كيف يمكن تحسين استخدام أساليب القياس والتقويم لتلبية متطلبات الأبحاث العلمية الحديثة؟

للإجابة عن هذا السؤال تم التحليل الموضوعي لبيانات مقابلات باستخدام برنامج NVivo، مع دعم الإجابة ببيانات المقياس الكمية بالتحليل الوصفي لسؤال مفتوح في المقياس يتعلق بكيفية تحسين استخدام أساليب القياس والتقويم والجدول (6) يوضح نتائج التحليل

الجدول (6)

التحليل الموضوعي والكمي لكيفية تحسين استخدام أساليب القياس والتقويم

التحليل الموضوعي للمقابلة (النوعية)			التحليل الوصفي لعبارات المقياس (الكمية)		
م	مضمون الأسئلة في المقابلة والمقياس	التكرار (المقابلة)	النسبة المئوية (المقياس)	المتوسط (M)	الانحراف المعياري(SD)
1	تطوير تدريب متخصص	14	75.0	4.4	0.7
2	استخدام تكنولوجيا متقدمة	12	71.3	4.3	0.8
3	تصميم أدوات مرنة متعددة التخصصات	10	60.6	4.1	0.9
4	تعزيز التعاون بين الباحثين	9	53.8	3.9	1.0

يتبين من الجدول (6) إلى أن اقتراح تطوير التدريب المتخصص جاء بتكرار في المقابلة بعدد (14) وبنسبة (75%) وبمتوسط (4.4) هو الحل الأكثر قبولًا واتفاقاً، مما يعكس إدراك الباحثين لأهمية تعزيز الكفاءات في استخدام الأدوات المعقدة، وأما اقتراح استخدام التكنولوجيا المتقدمة أتى بتكرار في المقابلة بعدد (12) وبنسبة (71.3%) وبمتوسط (4.3) ويُظهر ذلك الى التوجه نحو الاستفادة من الذكاء الاصطناعي وتحليل البيانات لتحسين الكفاءة، وكذلك اقتراح تصميم أدوات مرنة جاء بتكرار في المقابلة بعدد (10) وبنسبة (60.6%) وبمتوسط (4.1) ويُلبي ذلك الحاجة إلى حلول تتكيف مع التنوع متعدد التخصصات، بينما اقتراح تعزيز التعاون بين الباحثين أتى بتكرار في المقابلة بعدد (9)

وبنسبة (53.8%) وبمتوسط (3.9) يحث الى التكامل بين الخبرات المختلفة، وهذه الاقتراحات تُقدم إطارًا عمليًا لتطوير الممارسات البحثية في مواجهة التحديات الحديثة.

تدعم هذه التوصيات والاقتراحات دراسة (Brynjolfsson & McAfee, 2014)، اللذان يؤكدان على أن التكنولوجيا تُحسن الجودة والكفاءة عند توفر التدريب المناسب، كما تتسق مع دراسة (Creswell & Plano Clark, 2017) حيث أوصيا بأدوات مرنة عند استخدام المنهج المختلط، ويرى الباحث في الدراسة الحالية الى تعزيز التعاون بين الباحثين كعنصر إضافي، مما يُثري الحلول المقترحة سابقًا.

التوصيات والاقتراحات:

واستنادًا إلى النتائج ، يوصي الباحث بالاتي :

1- تعزيز القدرات من خلال برامج تدريبية متقدمة:

يُوصى بتصميم دورات تدريبية متخصصة تُركز على تطوير مهارات الباحثين في استخدام أدوات القياس والتقويم الحديثة، مثل البرمجيات الإحصائية (R،SPSS) وأنظمة تحليل البيانات النوعية (NVivo)، كما يجب أن تُصمم هذه البرامج لتلبية احتياجات التخصصات المتنوعة مع التركيز على تعزيز الكفاءة في تطبيق الأساليب الكمية والنوعية على حد سواء، لمواجهة نقص المهارات التقنية المُلاحظ.

2- الاستفادة من التقنيات المعاصرة لتحسين كفاءة أدوات الدراسة:

يُنصح الباحثون والمؤسسات البحثية بالاستثمار في التقنيات المتطورة، مثل الذكاء الاصطناعي وأجهزة الاستشعار الآلية، لرفع مستوى الدقة في جمع البيانات وتحليلها، وهذا التوجه يُعزز من قدرة الأساليب على التعامل مع الظواهر المعقدة، خاصة في المجالات التي تتطلب معالجة كميات ضخمة من البيانات والمعلومات، ويُقلل من الأخطاء البشرية.

3- تطوير أدوات مرنة تتكيف مع التعددية التخصصية:

يُقترح على المجتمع العلمي والأكاديمي تصميم أطر قياس وتقويم قابلة للتعديل تتناسب مع متطلبات التخصصات المختلفة، مثل أدوات تجمع بين المقاييس الكمية والتحليلات النوعية، وهذا النهج يُسهم في التغلب على صعوبات توحيد المنهجيات في الأبحاث العلمية، مما يدعم الحصول على نتائج متكاملة وموثوقة.

4- تشجيع الشراكات البحثية لتعزيز التكامل:

يُوصى بإنشاء شبكات تعاون بين الباحثين من خلفيات علمية متنوعة لتبادل الخبرات وتطوير استراتيجيات مشتركة للقياس والتقويم، وهذه الشراكات تُمكن من مواجهة صعوبات وتعقيد الظواهر متعددة التخصصات من خلال الجمع بين المناهج الموضوعية والتفسيرية، مما يُحسن جودة المخرجات.

5- توفير الدعم المؤسسي لتذليل العوائق المادية:

يُنصح المؤسسات الأكاديمية والهيئات التمويلية بزيادة الموارد المخصصة للبحث العلمي، بما في ذلك تمويل البرمجيات والمعدات، وتخصيص وقت كافٍ لإجراء الدراسات، وهذا الدعم يُخفف من الضغوط المرتبطة بضعف الموارد، مما يسمح للباحثين بالتركيز على تحسين جودة الأساليب المستخدمة.

6- إنشاء معايير موحدة للتقييم عبر التخصصات:

يُقترح تطوير إرشادات عالمية لتقييم جودة أساليب القياس والتقويم، تُراعي التنوع بين المجالات العلمية مع ضمان الحد الأدنى من الاتساق، وهذه المعايير تُساعد في تقليل التباين في التطبيق وتُعزز من قابلية المقارنة بين الدراسات، مما يدعم المصداقية العلمية.

7- إجراء أبحاث مستقبلية لاستكشاف التطبيقات العملية:

يُوصى بإجراء دراسات لاحقة تُركز على اختبار فعالية الأدوات المرنة والتقنيات المتقدمة في نطاقات بحثية محددة، مثل العلوم البيئية أو الطبية او الاجتماعية، وهذه الأبحاث تُسهم في بناء قاعدة معرفية تُوجه الباحثين نحو أفضل الممارسات في تطبيق هذه الأساليب.

8- تعزيز الوعي بأهمية الخصائص السيكومترية:

يُنصح بتضمين تعليم مبادئ خصائص الصدق والثبات في المناهج الأكاديمية للباحثين الجدد، لضمان اختيار أدوات ذات جودة عالية عند الشروع في عمل البحث، وهذا الوعي يُقلل من المخاطر المرتبطة باستخدام أساليب غير موثوقة، مما يرفع مستوى الإتقان والاحتراف في منهجية الأبحاث.

المصادر والمراجع المعتمدة:

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31(4)*, 399–413.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). SAGE Publications.

Drake, S. (1978). *Galileo at work: His scientific biography*. University of Chicago Press.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2(8)*, e124.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). Berrett-Koehler Publishers.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Mertens, D. M. (2018). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (5th ed.). SAGE Publications.

Moher, D., Altman, D. G., Schulz, K. F., & Simera, I. (2010). Guidelines for reporting health research: The CONSORT statement. *PLoS Medicine*, 7(3), e1000251.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349(6251),* aac4716.

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE Publications.

Repko, A. F., & Szostak, R. (2020). *Interdisciplinary research: Process and theory* (4th ed.). SAGE Publications.

Rothwell, P. M. (2005). External validity of randomised controlled trials: "To whom do the results of this trial apply?". *The Lancet, 365(9453)*, 82–93.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). SAGE Publications.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103(2684)*, 677–680.

Tashakkori, A., & Teddlie, C. (2010). *Handbook of mixed methods in social & behavioral research* (2nd ed.). SAGE Publications.

Trochim, W. M. K. (2020). *Research methods knowledge base*. Atomic Dog Publishing.

The Role of Measurement and Evaluation Methods in Enhancing the Quality of Scientific Research

Abstract

This study aimed to explore the most prevalent measurement and evaluation methods for enhancing the quality of scientific research, identify challenges encountered in multidisciplinary research, and propose improvement strategies. The study adopted a mixed-methods approach, utilizing a sequential explanatory design. Quantitative data were initially collected using a scale from a sample of 350 researchers from various scientific disciplines, selected via stratified random sampling. Subsequently, qualitative data were gathered through semi-structured interviews with 15 purposively selected experienced researchers.

The findings revealed an 80% effectiveness rate for the use of quantitative measures in scientific research, and statistical analysis was deemed important in study outcomes by 70% of respondents, with these being the most frequently utilized approaches, exhibiting variations across disciplines. A strong positive relationship was identified between the quality of measurement and evaluation methods and validity ($\beta = 0.65$), explaining 42% of the variance ($R^2 = 0.42$). Furthermore, this quality significantly influenced reliability ($\beta = 0.61$), accounting for 37% of its variance ($R^2 = 0.37$). Key challenges identified included difficulties in standardizing tools, a lack of technical skills, the complexity of multidisciplinary phenomena, and scarce resources, and Researchers' suggestions included developing specialized training, employing advanced technologies, designing flexible tools, and fostering collaboration among researchers.

The study recommended enhancing capabilities through advanced training programs, leveraging contemporary technologies to improve the efficiency of research tools, developing flexible tools adaptable to multidisciplinarity, and encouraging research partnerships to foster integration. Further recommendations include providing institutional support to overcome material obstacles, establishing standardized evaluation criteria across disciplines, conducting future research to explore practical applications, and promoting awareness of the importance of psychometric properties.

Keywords: Measurement and Evaluation Methods, Scientific Research.