

الرؤية المستقبلية لتطوير مجال القياس والتقويم والإحصاء النفسي والتربوي (الجزء الثاني)

الأستاذ الدكتور

محمد محمد فتح الله سيد أحمد

أستاذ القياس والتقويم والإحصاء النفسي والتربوي
المركز القومي لامتحانات والتقويم التربوي
ورئيس مجلس أمناء المؤسسة القومية للبحوث
والاستشارات والتدريب (NRCT)

المستخلص

تهدف هذه الورقة إلى تقديم رؤية مستقبلية لتطوير مجال القياس والتقويم والإحصاء النفسي والتربوي في الجامعات المصرية، باعتباره أحد الركائز الأساسية في منظومة التعليم والبحث العلمي. تعتمد الرؤية على منهجية التخطيط الاستراتيجي، حيث مرت بعدة مراحل تضمنت تحليل الوضع الراهن، تحديد التحديات القائمة، وصياغة معايير ومؤشرات قياس الأداء لضمان تحقيق الأهداف المرجوة. كما تسلط الورقة الضوء على الاتجاهات البحثية الحديثة في مجال القياس والتقويم، بما في ذلك الاختبارات التكيفية المحوسبة، نموذج المعادلة البنائية، التقويم البديل، والتقييم الإلكتروني، مع التركيز على رقمنة القياس والتقويم من خلال بنوك الأسئلة الإلكترونية ونظرية الاستجابة للمفردة. بالإضافة إلى ذلك، تقترح الورقة عددًا من الإستراتيجيات لتعزيز جودة التعليم، تطوير المناهج، ودعم البحث العلمي، مع التأكيد على أهمية التعاون الدولي والشراكة المجتمعية في تحقيق هذه الأهداف. وخلصت الورقة إلى أن تبني هذه الرؤية سيسهم في تحقيق العدالة والموضوعية في التقييم، وتحسين جودة التعليم، وتعزيز مكانة التخصص على المستويين المحلي والدولي.

الكلمات المفتاحية: القياس والتقويم، الإحصاء النفسي والتربوي، التخطيط الاستراتيجي، الاختبارات التكيفية المحوسبة، نموذج المعادلة البنائية، التقويم البديل، رقمنة التقييم، بنوك الأسئلة الإلكترونية، نظرية الاستجابة للمفردة، جودة التعليم، البحث العلمي.

ملحوظة: تم نشر الجزء الأول من هذه الرؤية في مجلة كلية التربية جامعة العريش أكتوبر 2021 المجلد 9، العدد 28 - الصفحة 19-55. وقد قدمت تلك الورقة إلى لجنة الترقيات لعلم النفس التربوي والصحة النفسية عام 2020.

رابط الجزء الأول: https://foej.journals.ekb.eg/article_196309.html

ISSN: 3009-612X

E. ISSN: 3009-6146

الترقيم الدولي الموحد للطباعة

الترقيم الدولي الموحد الإلكتروني

Future Vision for Developing the Field of Psychological and Educational Measurement, Evaluation, and Statistics (Part Two)

Prof. Mohamed Mohamed Fathallah Sayed Ahmed

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

Chairman of the Board of Trustees of the National Research, Consulting and Training Corporation (NRCT)

Abstract

This paper aims to present a future vision for developing the field of measurement, evaluation, and psychometric statistics in Egyptian universities, as it represents a fundamental pillar in the educational system and scientific research. The proposed vision follows a strategic planning methodology, progressing through various stages, including analyzing the current status, identifying existing challenges, and formulating standards and performance indicators to ensure the achievement of desired objectives. The paper also highlights recent research trends in the field of measurement and evaluation, such as computerized adaptive testing, structural equation modeling, alternative assessment, and digital assessment, with a particular focus on digitizing evaluation through electronic item banks and the item response theory. Additionally, the paper proposes several strategies to enhance education quality, develop curricula, and support scientific research, emphasizing the importance of international collaboration and community partnerships in achieving these goals. The findings indicate that adopting this vision will contribute to achieving fairness and objectivity in evaluation, improving education quality, and enhancing the field's status at both local and international levels.

Keywords: Measurement and Evaluation, Psychometric Statistics, Strategic Planning, Computerized Adaptive Testing, Structural Equation Modeling, Alternative Assessment, Digital Assessment, Electronic Item Banks, Item Response Theory, Education Quality, Scientific Research.

مقدمة:

ملخص الجزء الأول من الرؤية المستقبلية لتطوير مسار القياس والتقويم والإحصاء النفسي والتربوي:

استهدف الجزء الأول إلى وضع رؤية مستقبلية لتطوير مجال القياس والتقويم والإحصاء النفسي والتربوي في الجامعات المصرية، باعتباره أحد الركائز الأساسية في علم النفس التربوي. وقد اعتمد الباحث منهجية التخطيط الإستراتيجي في إعداد هذه الرؤية، حيث مرت بعدة مراحل متكاملة تبدأ بوصف وتحليل الواقع الحالي للمجال، متبوعة بتحديد أبرز التحديات التي تواجهه، وانتهاءً بوضع آليات تنفيذ واضحة ومؤشرات قياس الأداء لضمان تحقيق الأهداف المرجوة.

يتمتع مجال القياس والتقويم بأهمية كبرى في المنظومة التعليمية، حيث يساهم في تحسين مخرجات التعلم من خلال تقديم أدوات دقيقة لقياس أداء الطلاب، وتحديد مدى تحقيقهم لنواتج التعلم المستهدفة. كما يلعب دوراً رئيسياً في رسم السياسات التربوية، وصياغة خطط تطوير المناهج، وتقييم كفاءة المعلمين، مما يعزز من جودة التعليم بوجه عام. ومع ذلك، يواجه المجال عدداً من التحديات، من أبرزها عدم وجود أقسام أكاديمية مستقلة له في الجامعات المصرية، ونقص البرامج الدراسية المتخصصة على مستوى الدراسات العليا، بالإضافة إلى غياب التوحيد في المصطلحات العلمية المستخدمة داخل المجال.

ولتجاوز هذه العقبات، اقترح الجزء الأول من الرؤية مستقبلية ارتكاز تعزيز دور القياس والتقويم النفسي والتربوي على دعم البحث العلمي وتحقيق التميز الأكاديمي. وتتمثل هذه الرؤية في وضع خطط تطويرية تتضمن إنشاء أقسام مستقلة للمجال داخل كليات التربية، وتحديث المناهج بما يتوافق مع المستجدات العلمية والتكنولوجية، وتوسيع نطاق الدراسات التطبيقية لضمان ارتباط الأبحاث باحتياجات المجتمع التعليمي. كما تدعو إلى تبني أساليب تقييم متطورة تعتمد على الذكاء الاصطناعي والاختبارات الإلكترونية وبنوك الأسئلة، لضمان دقة وموضوعية عمليات التقييم والقياس.

وفي ضوء هذه الرؤية، توصي الورقة بضرورة تبني مشروع وطني شامل يهدف إلى تطوير أدوات القياس والتقويم في المؤسسات التعليمية، مع توفير بيئة داعمة للبحث العلمي في هذا المجال. كما تدعو إلى عقد مؤتمرات علمية متخصصة لمناقشة أبرز القضايا والتحديات، وتقديم حلول مبتكرة لتعزيز كفاءة نظم التقييم. إضافة إلى ذلك، تشدد على أهمية توحيد المصطلحات العلمية المستخدمة في المجال، من خلال التنسيق بين الخبراء والأكاديميين، بما يساهم في توحيد المفاهيم وتعزيز التواصل العلمي الفعال.

ختاماً، تمثل هذه الرؤية خارطة طريق واضحة لتطوير مجال القياس والتقويم والإحصاء النفسي والتربوي في مصر، بما يضمن تحقيق أعلى معايير

الجودة في التعليم والتعلم. ومن خلال تطبيق هذه الإستراتيجيات، يمكن تعزيز دور القياس والتقويم في تحسين الأداء الأكاديمي، وضمان تحقيق العدالة والموضوعية في تقييم القدرات والمهارات التعليمية، مما يسهم في بناء نظام تعليمي أكثر كفاءة وفاعلية.

الجزء الثاني من الرؤية المستقبلية مسار القياس والتقويم والإحصاء النفسي والتربوي:
المرحلة الرابعة: معايير ومؤشرات الرؤية الإستراتيجية المقترحة وخطة تحقيقها.

المعيار الأول: جودة عمليتي التعليم والتعلم في البحث العلمي لمجال القياس والتقويم النفسي والتربوي بالكليات والمراكز البحثية.

قائمة (1) مؤشرات الأداء وأدلتها وآليات التنفيذ للمعيار الأول

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
وضع معايير للخريجين المتخصصين في القياس والتقويم والإحصاء النفسي والتربوي بمرحلة البكالوريوس والدراسات العليا	عقد ورش عمل للعاملين بالتخصص يشارك فيها أفراد من المؤسسات التربوية للخروج بمجموعة من المعايير لطالب البكالوريوس والدراسات العليا.	عدد ورش العمل.	لجنة صياغة المعايير.	موازنة لجنة القطاع
الاستفادة من المعايير العالمية في صياغة معايير محلية للخريجين.	وجود قائمة بالمعايير.	لجنة صياغة المعايير.	موازنة لجنة القطاع	
توظيف المعايير في بناء مقررات القياس والتقويم والإحصاء النفسي والتربوي.	عدد المقررات التي بنيت في ضوء المعايير.	الأقسام العلمية	موازنة الأقسام العلمية	
تشكيل لجان لمراجعة محتوى الكتب الجامعية بالتخصص وتطويرها في ضوء النظريات والتوجهات الحديثة وتخليصها من القديم واعتباره مقدمة للجديد.	عدد المقررات المطورة.	لجنة المقررات بالأقسام العلمية	موازنة لجنة القطاع	
الأخذ في الاعتبار رؤية ورسالة الكلية في إعداد الطالب المعلم	استطلاع رأي	لجنة المقررات	موازنة الكلية	

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
ضوء التطور المعرفي	النافع لنفسه ووطنه.	الطلاب والمؤسسات المشغلة.	بالقسم	
	تقييم المقررات بالقسم دوريا وعلى مستوى الأقسام بالجامعات المختلفة في ضوء نتائج التقييم السنوي للطلاب	عدد المقررات التي تم تقييمها.	لجنة المقررات بالأقسام العلمية المختصة	موازنة القسم
	الاستعانة بالمناهج التعليمية المناظرة في التخصص عالميا وإقليميا.	عدد المراجع الأجنبية بالكتب.	أعضاء هيئة التدريس	موازنة الكلية
	تدعيم مكتبة الكلية بالكتب الحديثة في مجال القياس والتقييم والإحصاء النفسي والتربوي.	عدد الكتب الحديثة بالتخصص.	القسم مع الكلية	موازنة الكلية
	دعم رقمنة القياس والتقييم والإحصاء النفسي والتربوي بتحويل المقررات إلى مقررات إلكترونية يمكن تحميلها ومشاهدتها وتبادلها بين الجامعات.	عدد المقررات الإلكترونية	عضو هيئة التدريس بكل مقرر	موازنة الكلية
	تشجيع تأليف الكتب الحديثة في التخصص	عدد الكتب المؤلفة في كل العام	الجامعة	موازنة الجامعة
تحسين استراتيجيات التعلم والتعليم في ضوء نواتج التعلم المستهدفة	عقد ورش عمل لتحديد استراتيجيات التعلم المناسبة لتحقيق نواتج التعلم المختلفة.	عدد ورش العمل	القسم العلمي	موازنة القسم
	تحقيق التوافق بين التخصص الدقيق لعضو هيئة التدريس والمقررات التي يقوم بتدريسها.	استطلاع رأي أعضاء هيئة التدريس.	القسم العلمي	موازنة الكلية
	تحديد نواتج التعلم المطلوبة بدقة واختيار الإستراتيجيات التي	مصفوفة نواتج التعلم	القسم العلمي	موازنة الكلية

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
	تحققها.			
التوظيف الأمثل لتكنولوجيا لتطوير أدوات القياس والتقويم والإحصاء النفسي والتربوي بالمعامل	عمل دراسات باحتياجات المعامل من الأجهزة والاختبارات في ضوء الاتجاهات المعاصرة	عدد الأجهزة الحديثة	لجنة المعامل بكل قسم	موازنة الكلية
	إنشاء معمل مركزي لتخصص القياس والتقويم والإحصاء النفسي والتربوي يلبي احتياجات الباحثين.	خطة إنشاء المعمل المركزي.	لجنة على مستوى التخصص.	موازنة الكلية
	تزويد المعامل بالبرامج والأدوات الحديثة	عدد البرامج الحديثة	القسم	موازنة الكلية
	تحويل الاختبارات الورقية إلى اختبارات إلكترونية.	عدد اختبارات إلكترونية.	القسم	موازنة الكلية
	توفير الصيانة الدورية للأجهزة والبرامج.	عدد مرات الصيانة	القسم	موازنة الكلية

المعيار الثاني: تطوير مهارات أعضاء هيئة التدريس ومعاونيهم وطلاب الدراسات العليا.

قائمة (1) مؤشرات الأداء وأدلتها وآليات التنفيذ للمعيار الثاني

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
تطوير العمل بأقسام علم النفس التربوي لتخصص القياس والتقويم والإحصاء	وضع خطة إستراتيجية لكل قسم	نسبة المتحقق من الخطة إلى المستهدف منها	القسم العلمي	موازنة القسم
	دمج أقسام وتخصصات القياس والتقويم والإحصاء النفسي والتربوي المتناظرة بالجامعة الواحدة (تربية،	عدد الأقسام المتناظرة في الجامعة	المجلس الأعلى للجامعات.	موازنة المجلس الأعلى للجامعات

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
النفسي والتربوي.	أداب، تربية نوعية، رياض أطفال، معاهد)			
	تكوين لجنة من الحكماء بكل قسم وذلك لتقديم المشورة وإمداد أفراد أسرة القسم بالخبرة والأفكار المتنوعة.	قرار تشكيل اللجنة	القسم	موازنة القسم
	توفير قاعدة بيانات عن أعضاء هيئة التدريس بالتخصص ونشرها	رابط قاعدة البيانات	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني
	تعيين المعيدين لفترة اختبار بعقود مؤقتة للتأكد من كفاءته عدد المعيدين علميا وأخلاقيا فالتحصيل وحدة ليس مؤشر الكفاءة	عدد المعيدين بعقود مؤقتة.	القسم العلمي	موازنة الكلية
	تعيين المعيدين على تخصصات فرعية (اختبارات، إحصاء، نظريات القياس الحديثة، قياس ذوي الإعاقات إلخ)	عدد المعيدين وتخصصاتهم.	القسم العلمي	موازنة الكلية
توفير نظام فعال للتدريب يسهم في تنمية الموارد البشرية في القياس والتقويم والإحصاء النفسي	عقد مؤتمر عام للعاملين بالتخصص لمناقشة أفكار ومتطلبات أعضاء هيئة التدريس لتطوير التخصص	عدد المشاركين في المؤتمر	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني
	عقد اجتماع مع طلاب الدراسات العليا لمناقشة معوقات إجراء البحث العلمي وطرق تذليلها.	عدد طلاب الدراسات العليا المشاركين.	القسم	موازنة الكلية
	عمل دراسة لتحديد الاحتياجات التدريبية للعاملين بالتخصص.	وجود قائمة بالاحتياجات.	القسم	موازنة القسم

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
والتربوي.	تحديد مجموعة من الدورات التدريبية للعاملين في التخصص نابعة من احتياجاتهم الفعلية.	عدد الحقائق التدريبية في التخصص.	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني
	عقد اتفاقيات لتدريب أعضاء هيئة التدريس مع الجامعات المتميزة في التخصص.	عدد الاتفاقات التدريبية وتبادل الخبرات	القسم	موازنة القسم
	تفعيل الإرشاد الأكاديمي للطلاب وتخصيص ساعات له بجداول هيئة التدريس.	تقييم الطلبة للإرشاد الأكاديمي.	القسم	موازنة القسم
اختبار أفضل المتقدمين للقسم من الطلاب بمرحلتى البكالوريوس والدراسات العليا والعمل على رفع كفاءتهم	السعي لإنشاء مدارس تدريبية ملحقه بكل كلية تحت إشراف أساتذة من التخصص لتكون مجالا خصبا لتجريب النظريات الحديثة القياس والتقويم والإحصاء النفسي والتربوي.	قرار إنشاء المدرسة	الكلية مع الجامعة	موازنة الجامعة
	إنشاء جمعية علمية لطلاب الدراسات العليا تختص بنشر النشاط العلمي للطلاب في القياس والتقويم والإحصاء النفسي والتربوي.	قرار اعتماد الجمعية	القسم العلمي	دعم مؤسسات المجتمع المدني
	عقد دورات تدريبية لطلبة الدراسات العليا في التخصص	عدد الدورات التدريبية	القسم	موازنة الكلية
	عقد لقاءات حوارية دورية للطلاب الذين تم تسجيلهم لعرض تقدمهم في البحث والمشكلات التي تواجههم.	عدد الطلاب الذين قدموا سيمينارات	القسم	موازنة القسم

المعيار الثالث: نشر ثقافة القياس والتقويم والإحصاء النفسي والتربوي.

جدول (3) المعيار الثالث وآلية تنفيذه ومؤشرات الأداء والمسؤول عن التنفيذ آلية التنفيذ ومصادر التمويل

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
تدعيم البنية التحتية للبحث العلمي لأدوات وآليات القياس والتقويم والإحصاء النفسي والتربوي	تشكيل هيئة قومية أو نقابة ترعى البحوث المتخصصة في القياس والتقويم وتنظيمها.	قرار تشكيل اللجنة	الوزارة المختصة	دعم مؤسسات المجتمع المدني
	تشكيل لجنة من المتخصصين في مجال القياس والتقويم والإحصاء النفسي والتربوي لوضع معايير محددة للخطط العلمية المقدمة للبحث بمرحلتى الماجستير والدكتوراه.	قرار تشكيل اللجنة	القسم	دعم مؤسسات المجتمع المدني
	حصر أدوات القياس التي تقيس نفس السمة ودمجها وتكوين فرق متخصصة لتقنينها وضبطها.	عدد المقاييس المطورة	لجنة على مستوى التخصص	دعم مؤسسات المجتمع المدني
	عقد ورش عمل لتحديد المصطلحات في التخصص وعمل قاموس للمصطلحات في القياس والتقويم والإحصاء النفسي والتربوي.	وجود كتيب إرشادي بالمصطلحات	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني
دعم التميز البحثي بتخصص القياس والتقويم والإحصاء النفسي والتربوي.	عقد عدة ورش عمل وندوات للمتخصصين لوضع خطة بحثية معتمدة في مجال التخصص.	عدد ورش العمل لهذا الغرض.	لجنة على مستوى التخصص	دعم مؤسسات المجتمع المدني
	حصر للرسائل والأبحاث بالتخصص وتصنيفها للتعرف على الفجوات فيها.	قوائم بالرسائل والأبحاث وتصنيفها.	لجنة بكل قسم	دعم مؤسسات المجتمع المدني
	تشجيع البحوث المشتركة في مقابل البحوث الفردية	نسبة البحوث المشتركة إلى الفردية	المجلس الأعلى للجامعات	موازنة المجلس الأعلى للجامعات

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
	تشجيع البحوث البيئية من تخصصات مختلفة ليست تربوية فقط بتطبيقات منهجية القياس والتقويم في تخصصات طبية وهندسية وزراعية وتجارية وغيرها	عدد البحوث البيئية	القسم العلمي	موازنة الجامعة
بناء قاعدة بيانات بحثية متطورة لتخصص القياس والتقويم والإحصاء النفسي والتربوي. تغطي مختلف فروع التخصص	تبني مشروع يشترك فيه لجأن من جميع الأقسام لتجميع البحوث ورسائل الماجستير والدكتوراه لبناء قاعدة بيانات موحدة للتخصص على مستوى الجامعات.	وجود قاعدة بيانات بالبحوث بالتخصص.	لجنة على مستوى التخصص مع مختصين	دعم مؤسسات المجتمع المدني
	تصميم موقع إلكتروني على غرار المواقع العالمية المتخصصة وجود في مجال التخصص وتحديثه بصفة مستمرة على غرار موقع (ETS).	وجود رابط الموقع.	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني
تشجيع البحوث التطويرية والمسحية للقياس والتقويم النفسي والتربوي	توجيه الباحثين لإعادة تقنين أدوات القياس الرئيسية العالمية والقومية وغيرها.	عدد البحوث والرسائل المعدة لهذا الغرض.	القسم	دعم مؤسسات المجتمع المدني
	تضمين الخطة البحثية دراسات مسحية وتقويمية للبحوث التي أجريت في التخصص على فترات زمنية محددة	عدد البحوث المسحية والتقويمية	القسم	دعم مؤسسات المجتمع المدني
	توفير دعم مادي للمشروعات البحثية التطويرية	عدد المشروعات المدعمة	القسم مع الكلية	موازنة الكلية
دعم الشراكة العالمية في البحث	عمل شراكات دولية مع أقسام القياس والتقويم والإحصاء النفسي والتربوي والمؤسسات بالجامعات العالمية.	عدد الشراكات الدولية.	القسم	موزنة القسم

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
العلمي للقياس والتقويم النفسي والتربوي.	ترجمة البحوث العربية للغة الأجنبية ونشرها لدعم التواجد العربي على شبكة المعلومات الدولية.	عدد البحوث المترجمة والمنشورة.	كل عضو هيئة تدريس بالتخصص.	دعم مؤسسات المجتمع المدني
زيادة التعاون والتنسيق بين مراكز القياس والتقويم بالجامعة ووحداتها بالكليات وبين الجامعات.	الأخذ بنظام الإشراف المشترك بين الأقسام في الجامعات المختلفة.	عدد رسائل الإشراف المشترك.	القسم	موازنة القسم
	دعوة أساتذة التخصص من جامعات مختلفة للمشاركة عبر تطبيقات الانترنت في سيمينارات أقسام في جامعات أخرى.	عدد الأساتذة المشاركين	القسم	دعم مؤسسات المجتمع المدني
	انتداب أساتذة التخصص للتدريس بجامعات مختلفة.	عدد الأساتذة المتدربين	القسم	موازنة الكلية
	توحيد المقررات في مجال التخصص بحيث يكون إعداد الطالب متشابه في كل الكليات.	نسبة الاتفاق في المقررات	المجلس الأعلى للجامعات	موازنة المجلس الأعلى للجامعات
	تشجيع البحوث المشتركة بين أكثر من عضو من أعضاء هيئة التدريس وأكثر من قسم.	عدد البحوث المشتركة	المجلس الأعلى للجامعات	موازنة المجلس الأعلى للجامعات
	العمل بنظام المجموعات البحثية	عدد المجموعات البحثية.	القسم	موازنة القسم
توظيف المؤتمرات العلمية لتطوير القياس والتقويم والإحصاء النفسي والتربوي.	تشكيل لجنة مشتركة خاصة بالمؤتمرات على مستوى العاملين بالتخصص.	قرار تشكيل اللجنة.	اللجنة العمومية للتخصص	موازنة الكلية
	وضع خطة عامة للمؤتمرات على مستوى أقسام علم النفس التربوي تخصص قياس وتقويم نفسي وتربوي بالجامعات.	وجود خطة للمؤتمرات.	لجنة المؤتمرات بالتخصص	دعم مؤسسات المجتمع المدني
	تشكيل لجان لتحويل توصيات المؤتمرات إلى إجراءات تنفيذية.	قرار تشكيل اللجنة.	الجمعيات الأهلية العلمية	دعم مؤسسات المجتمع

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
	تخصيص محور في كل المؤتمرات لتقويم الوضع الحالي للتخصص ومتطلبات التطوير.	نسخة من منشورات المؤتمر	الجمعيات الأهلية العلمية المتخصصة	المدني
	عقد مؤتمر واحد شامل للتخصص كل عام يتم تدويره في كل الأقسام يعالج قضايا ملحة.	عدد المؤتمرات في التخصص	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني

المعيار الرابع: توفير مصادر التمويل البحث العلمي والمشاريع والبرامج القومية.

جدول (4) المعيار الرابع وآلية تنفيذه ومؤشرات الأداء والمسؤول عن التنفيذ ومصادر التمويل

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
تفعيل دور الوحدات ذات الطابع الخاص بالتخصص لدعم البحث العلمي وخدمة المجتمع..	تأسيس وحدات ذات طابع خاص بكل أقسام علم النفس التربوي لخدمة الباحثين والمجتمع مثل: - مركز لتقييم ذوي صعوبات التعلم، مركز للقياس والتقويم والتحليل الإحصائي، مركز للاستشارات التربوية.	عدد المراكز بكل قسم	القسم مع الجامعة	موازنة الجامعة
تطوير شراكات بحثية مع الجهات المانحة	التقدم للجهات المانحة الدولية للبحوث العلمية مثل صندوق العلوم والتنمية التكنولوجية وأكاديمية البحث العلمي والتكنولوجيا ووزارة التعليم العالي من خلال وحدة إدارة مشروعات	عدد المشروعات المقدمة	أعضاء هيئة التدريس بالتخصص	موازنة المجلس الأعلى للجامعات

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
	التميز وبرامج الاتحاد الأوروبي.			
	عقد ورش عمل تعريفية بالجهات المانحة للمشاريع	عدد الورش التعريفية	لجنة بالقسم	موازنة المجلس الأعلى للجامعات
	عقد ورش عمل في كيفية كتابة المشاريع البحثية والتقدم بها للجهات المانحة.	عدد الورش المتعلقة.	مكتب المنح بالجامعة	موازنة المجلس الأعلى للجامعات
	تشجيع مشاركة طلاب الدراسات العليا والهيئة المعاونة في المشاريع البحثية المتخصصة.	عدد الطلاب المشاركين	القسم	موازنة الكلية
إنشاء كراسي علمية في التخصص	دعوة رجال الأعمال للإسهام في تمويل البحوث العلمية في التخصص بعمل كرسي بحثي يتناول قضية من قضايا التخصص مثل بنوك الأسئلة Computer-Adaptive Test (CAT) واختبارات الاستعدادات	عدد الكراسي العلمية	القسم	موازنة الكلية
	توعية أفراد المجتمع بما يمكن أن يقدمه مجال القياس والتقويم والإحصاء النفسي والتربوي لخدمة المجتمع	عدد اللقاءات التعريفية	القسم	موازنة الكلية
	عقد ورشة عمل لتحديد أبرز مشكلات الميدان والمنتشرة بالمدارس أو المجتمع لدراساتها في إطار مشاريع قومية.	عدد ورش العمل	القسم	موازنة الجامعة
تبني فكرة المشاريع والحملات القومية التي تتناول				

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
قضايا تهم المجتمع	عقد ورش عمل للقيام بحملات قومية وقائية لنشر ثقافة القياس والتقويم والإحصاء النفسي والتربوي وفلسفته لدى فئات المجتمع.	عدد ورش العمل	القسم	موازنة الجامعة

المعيار الخامس: تحقيق شراكة مجتمعية فاعلة لجودة القياس والتقويم والإحصاء النفسي والتربوي.

جدول (5) المعيار وآلية تنفيذه ومؤشرات الأداء والمسؤول عن التنفيذ ومصادر التمويل

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
الوفاء بمتطلبات سوق العمل بإعداد خريجين متميزين تتلاءم قدراتهم ومهاراتهم مع متطلبات عملهم.	تحليل للمهارات المطلوبة للخريج للعمل في مجال التخصص ووضع المقررات لتحقيق هذه المهارات	قوائم تحليل للمهارات	لجنة المقررات.	موازنة لجنة القطاع
	تطوير التدريب الميداني وعمل معامل للتدريس المصغر في التخصص.	عدد معامل التدريس المصغر في التخصص	القسم	موازنة الكلية
	تقويم أداء العاملين بالتخصص من معلمين أو معيدين ومدرسين مساعدين على فترات زمنية.	نتائج تقييم الأداء	لجنة التقويم بالتخصص	موازنة الكلية
	تعميم نظام المرشد ل مجال القياس والتقويم النفسي التربوي في مختلف مراحل التعليم	عدد المرشدين النفسيين	وزارة التربية والتعليم.	موازنة وزارة التربية والتعليم صندوق دعم المشروعات

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
	مشاركة أفراد من المؤسسات التربوية والمجتمعية في مجالس وسيمنارات القسم.	عدد المشاركين بالمجالس والسيمنارات	القسم	موازنة وزارة التربية والتعليم صندوق دعم المشروعات
توجيه إمكانات خبراء القياس والتقويم والإحصاء النفسي والتربوي في البحث العلمي والاستشارات والتدريب لحل مشكلات المجتمع والمساهمة الفعالة في تحقيق أهدافه.	عقد ورش عمل بين العاملين بالمجال والمؤسسات التربوية للتعرف على المشكلات في الميدان.	عدد ورش العمل	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني
	فتح فرص عمل لخريجي التخصص بالشركات والمصانع لتعليم منهجية اعداد أدوات القياس وتطبيقها وتفسير النتائج لقياس مشكلات العمل والرضا عنه.	عدد العاملين بالشركات والمصانع	وزارة القوى العاملة	موازنة صندوق دعم المشروعات
	دعوة المؤسسات التربوية بالمجتمع للمشاركة في وضع الخطة البحثية للتخصص من خلال ورش عمل وندوات	عدد المشاركين من مؤسسات المجتمع.	القسم	موازنة صندوق دعم المشروعات
	تحفيز أعضاء هيئات التدريس والهيئة المعاونة للتصدي لحل مشاكل المجتمع عن طريق اعداد أدوات لقياس الظواهر بدقة ووضع ضوابط التطبيق ومنهجية تفسير النتائج	عدد البحوث التي شارك فيها تخصصات مختلفة	المجلس الأعلى للجامعات (لجان الترقية)	موازنة المجلس الأعلى للجامعات
	فتح قنوات للتواصل مع وزارة التربية والتعليم والوزارات الأخرى ذات الصلة	عدد المشاركات مع الوزارة	القسم	موازنة صندوق دعم المشروعات

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
	توجيه نظر الباحثين إلى دراسة موضوعات مختلفة تخدم توجيه تخدم المجتمع مثل: اعداد أدوات تكشف واقع الامية في مصر وكذا تقييم واقع منظومة التعليم والتعلم في كافة المراحل الدراسية وكذا الكشف عن بعض الظواهر ارتفاع درجات الطلاب في الشهادات العامة وظواهر الغش بأنواعه وغيرها.	عدد البحوث التي تناولت قضايا مجتمعية.	القسم	موازنة القسم
تعزيز دور تخصص القياس والتقويم والإحصاء النفسي والتربوي في الحياة الثقافية والاجتماعية لمختلف قطاعات المجتمع التعليمي.	تنظيم معارض علمية لنشر الوعي العلمي والثقافة العلمية المتخصصة في المجتمع.	صور من المعارض.	القسم	موازنة الكلية
	طبع دوريات أو نشرات لمد المهتمين بالمجال بالأسس العلمية للتعامل مع المشكلات القياس وأدواته وتطوره.	نسخة من النشرات والدوريات.	القسم	موازنة الكلية
	اقتراح دورات تدريبية لفئات المجتمع المختلفة.	عدد الدورات المقترحة.	القسم	موازنة الكلية
	توعية المجتمع بالأبحاث والتخصصات القائمة في القسم وأوجه الاستفادة منها.	عدد اللقاءات التعريفية	القسم	موازنة الكلية
	عقد ورش عمل دورية على شبكة المعلومات الدولية بين المتخصصين والعاملين	عدد ورش العمل.	لجنة عليا على مستوى التخصص	موازنة المجلس الأعلى للجامعات
	تفعيل مواقع التواصل الاجتماعي بين المتخصصين والعاملين	استطلاع رأي العاملين.	لجنة عليا على مستوى التخصص	موازنة المجلس الأعلى

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
				للجامعات
	التنسيق مع وسائل الإعلام المقروءة والمرئية والمسموعة لنشر الثقافة الصحيحة لفلسفة القياس والتقويم، وتحسين صورة متخصصي القياس والتقويم والإحصاء النفسي والتربوي بوسائل الإعلام	عدد اللقاءات التليفزيونية والمقالات المتخصصة.	لجنة عليا على مستوى التخصص	موازنة المجلس الأعلى للجامعات
تحفيز منظمات المجتمع للمشاركة الفعالة في تطوير سياسات وخطط تخصص القياس والتقويم والإحصاء النفسي والتربوي المستقبلية.	عقد ورش عمل بين المتخصصين وأفراد من مؤسسات المجتمع لمناقشة خطط القسم واحتياجات المجتمع	عدد ورش العمل.	لجنة المقررات بالقسم	موازنة الكلية
تصميم موقع خاص بالتخصص وإبريد إلكتروني مشترك خاص بالعاملين بالمجتمع لتداول أي مقترحات وتعميمات (مكاتبات / وإعلانات) وكذلك المستقبلية تبادل التواصل في المناسبات المختلفة.	عقد شراكات للتعاون مع المؤسسات التربوية لإجراء البحوث الميدانية والتطبيقية والقيام بالاستشارات لمواجهة القضايا التربوية	عدد بروتوكولات التعاون	عدد بروتوكولات التعاون	موازنة صندوق دعم المشروعات
تعميق الهوية العربية الإسلامية	عقد مؤتمر يتعلق بالهوية العربية الإسلامية لتخصص القياس والتقويم	نشرة المؤتمر	الجمعيات الأهلية العلمية	دعم مؤسسات المجتمع

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
لتخصص القياس والتقويم والإحصاء النفسي والتربوي.	والإحصاء النفسي والتربوي.		المتخصصة	المدني
	تضمين الخطة البحثية للتخصص محورا للهوية العربية الإسلامية القياس والتقويم والإحصاء النفسي والتربوي.	نسخة من الخطة.	لجنة إعداد الخطة	موازنة القسم
	تشجيع الدراسات النقدية لتوجهات البحوث الغربية ومدى ملاءمتها للهوية العربية الإسلامية.	عدد الدراسات النقدية	القسم	موازنة الكلية
	تشجيع البحوث التطبيقية لمعالجة مشكلات واقعية من الميدان التربوي والمجتمع.	عدد البحوث التطبيقية	القسم	موازنة الكلية

المعيار السادس: تعزيز المكانة الدولية للقياس والتقويم النفسي والتربوي.

جدول (6) المعيار السادس وآلية تنفيذه ومؤشرات الأداء والمسؤول عن التنفيذ ومصادر التمويل

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
دعم أنشطة التعاون الدولي والعلاقات الثقافية مع الجامعات العالمية والمنظمات الدولية والمراكز البحثية	عمل شراكات بحثية مع المراكز البحثية العالمية.	عدد الشراكات	القسم	دعم مؤسسات المجتمع المدني
	دعوة أساتذة بارزين في التخصص لحضور ندوات وسيمينارات التخصص.	عدد الأساتذة الحضور في الندوات	القسم	دعم مؤسسات المجتمع المدني
	تنظيم مؤتمرات مشتركة مع جامعات عالمية وإقليمية متميزة.	عدد المؤتمرات المشتركة.	لجنة المؤتمرات بالتخصص.	دعم مؤسسات المجتمع المدني

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
	تفعيل الشراكة مع جمعيات وروابط علم النفس في الوطن العربي	عدد اتفاقيات الشراكة	الجمعيات الأهلية العلمية المتخصصة	دعم مؤسسات المجتمع المدني
توفير منح دراسية للهيئة المعاونة بالتخصص في الجامعات المتقدمة	السعي لتخصيص منح دراسية للتخصص لدى وزارة التعليم العالي.	عدد المنح بالتخصص.	القسم	موازنة وزارة التعليم العالي
	عقد لقاءات مع الهيئة المعاونة لإكسابهم مهارات التواصل مع الجامعات المتقدمة.	عدد اللقاءات المتعلقة بمهارات التواصل	خبراء من مكتب المنح	موازنة وزارة التعليم العالي
	التوسع في نظام الإشراف المشترك بين الأقسام في الداخل والخارج.	عدد الرسائل بنظام الإشراف المشاركين.	القسم	موازنة وزارة التعليم العالي
تشجيع المشاركة في البحوث والمؤتمرات الدولية.	دعوة المتخصصين لحضور المؤتمرات الدولية ونشر مواعيد انعقادها.	عدد أعضاء هيئة التدريس المشاركة	القسم	موازنة الكلية
	فتح قنوات للتواصل مع أساتذة من التخصص في جامعات عالمية لعمل بحوث مشتركة.	عدد البحوث المشتركة..	لجنة التواصل بالتخصص	دعم مؤسسات المجتمع المدني
تشجيع النشر في الدوريات العالمية	وضع حوافز للنشر العلمي في مجلات متخصصة وعالمية	عدد الحاصلين على جوائز نشر.	وكيل الكلية للدراسات العليا	موازنة الكلية

المؤشرات	آلية التنفيذ	الشواهد والأدلة	المسؤول عن التنفيذ	مصادر التمويل
المتميزة	الاشتراك في المجلات العلمية الدولية في مجال التخصص.	عدد الاشتراكات في المجلات	الكلية	موازنة الكلية
	جعل النشر بمجلات عالمية متطلب من متطلبات الترقية	عدد البحوث المطلوبة للترقية	المجلس الأعلى للجامعات (لجنة الترقيات)	موازنة المجلس الأعلى للجامعات

المرحلة الخامسة: التوجهات البحثية الحديثة والتي ينبغي الاهتمام بها:
يري الباحث أنه ينبغي التعرض بشيء من التفصيل إلى التوجهات البحثية الحديثة والتي ينبغي الاهتمام بها والتعريف بأساسيات القياس والتقويم بما يمكن من استخدامها بكفاءة ودراية في تقييم وتقويم المتعلمين.

1- الاختبارات المحبوبة /التكيفية/ الموائمة المحوسبة Computerized Adaptive Testing (CAT):

والذي يطلق عليه أيضا " الاختبار المفصل " حيث يتم تفصيل وعرض مفردات لكل ممتحن تبعا لمستوي قدراتهم وبذلك يمتحن كل ممتحن بمفردات مختلفة عن الآخرين مراعاة للفروق الفردية بين الممتحنين. حيث يتم عرض مفردات لكل فرد تبعا لمستوي قدرته في المفردات الأولية، ومن ثم يتعرض كل فرد لمفردات مختلفة عن الآخرين لاختلاف قدراتهم، ويتم تحديد مستوى كل فرد من خلال مفهوم "العمر القاعدي" الذي يتحدد من الاختبارات الأولية التي يتعرض له كل فرد وفي ضوء العمر القاعدي يتعرض الفرد لمفردات مختلفة عن الأفراد الآخرين وذلك لاختلاف العمر القاعدي لكل فرد عن الآخرين، ومن ثم يتم اختزال الـ(129) اختبار فرعي في عدد قليل من الاختبارات التي يمكن تطبيقها في زمن ساعة ونصف. (Ranger& Kuhn, 2011)

وبالتالي فإن هناك مميزات للاختبارات التكيفية المحوسبة عن الاختبارات التقليدية التي كانت تهتم بعرض جميع مفردات الاختبارات بنفس الطريقة على جميع الممتحنين وفيما يلي جدول (1) يوضح الفرق بينهما.

جدول (7) المقارنة بين الاختبارات التقليدية والاختبارات التكيفية المحوسبة

الاختبارات التكيفية المحوسبة	الاختبارات التقليدية	وجه المقارنة
يؤدي كل ممتحن اختبار مختلف عن الآخر	كل المشتركين يؤديون نفس الاختبار	تركيب الاختبار
تحدد في ضوء إجابة كل فرد	تحدد في ضوء متوسط إجابات المشتركين	صعوبة الاختبار
طول الاختبار يختلف باختلاف المشتركين يفضل الاختبارات القصيرة	طول الاختبار واحد لكل المشتركين يفضل الاختبارات الطويلة	طول الاختبار
يختلف الزمن لكل مشترك	زمن واحد لكل المشتركين	زمن الاختبار
يحتاج لوقت قصير	يحتاج لوقت طويل	تنظيم الاختبار
نعم	لا	فورية النتائج

يلاحظ من الجدول أن انتقاء المفردات من بنك الأسئلة في الاختبارات التكيفية المحوسبة يعتمد على قدرات المشتركين التي تم تحديدها من المفردات الأولى التي تم عرضها في بداية الدقائق الأولى من الاختبار لتحديد قدرات كل ممتحن وبالتالي يتم استدعاء مفردات تتناسب مع كل ممتحن حسب قدراته كما أن زمن الاختبار بالتالي يختلف باختلاف تلك المفردات التي تظهر لكل ممتحن نظرا لاختلاف طول الاختبار الذي يتناسب أيضا مع قدرات كل ممتحن على حدة (Tian, 2007 et al).

خطوات بناء الاختبارات التكيفية المحوسبة: يتم من خلال اتباع الخطوات التالية:

بناء مستودع مفردات Item Pool، اختيار أحد نماذج نظرية الاستجابة للمفردة الاختبارية المناسب، تحديد إستراتيجية انتقاء وعرض العناقيد، اختيار أسلوب تقدير القدرة المناسب، تحديد قاعدة إنهاء الاختبار (Wainer, Kaplan, Lewis, 1991).

وقد تطورت العديد من التطبيقات لبنوك الأسئلة والاختبارات التكيفية المحوسبة اعتمادا على برنامج R والذي يستند إلى لغة البرمجة R ويوجد العديد من البرمجيات التي تعتمد على لغة البائسون python التي تعد اللغة الأساسية في مجال الذكاء الاصطناعي نظرا لسهولة استخدامها ولتوافر الأكواد بسهولة حتى أن النسخ

الحديثة من برنامج SPSS تعتمد لغة البايثون والعديد من برمجيات الذكاء الاصطناعي في مجال التربية والتربية الخاصة والعلاج النفسي ويتضمن الاختبار التكيفي المحوسب لكل مفردة، والاختبار التكيفي المحوسب متعدد المراحل. (محمد بشير زينة، 2017)

وفي الاختبار التكيفي المحوسب لكل مفردة يتم تكيف مفردات الاختبار لتناسب مستوى الطالب مما يعطي دقة أكبر في القياس بالرغم من محدودية عدد المفردات. حيث تقدم للطالب المفردات تدريجيًا وفق مستوى قدرته، حيث تؤثر إجابته عن مفردة ما على مستوى المفردات التالية التي يختارها الحاسب، ففي حال الإجابة الصحيحة عن السؤال يختار الحاسب سؤالاً أصعب قليلاً من الأول، وبالعكس إذا كانت الإجابة خطأ يختار الحاسب سؤالاً أقل صعوبة. (Zahng, B., & Cindy, M., W. 2008)

وفي الاختبار التكيفي المحوسب متعدد المراحل Multistage Adaptive Testing (MST) يطبق الاختبار على عدة مراحل اختبارية حيث يبدأ جميع الطلاب بمجموعة مفردات متوسطة الصعوبة يطلق عليها الاختبار التوجيهي، ثم يتم توجيه كل طالب إلى مجموعة أخرى من المفردات (تم تجميعها بناءً على خصائص إحصائية ووصفية معينة) وفق استجابته على الاختبار التوجيهي، وهكذا حتى التوصل إلى أفضل تقدير لقدرة الطالب في ضوء دالة المعلومات المستهدفة. وتمتاز هذه الطريقة بأن الطالب يمكن أن يراجع استجابته في كل مرحلة اختبارية. (Keng, 2008). وتمتاز الاختبارات التكيفية في أنها تعطي دقة في القياس على متصل القدرة؛ وذلك لطبيعتها التكيفية التي تحبك الاختبار على مستوى الطالب (Zahng, B., & Cindy, M., W. 2008). وتهتم الدراسات المعاصرة بالمقارنة بين تصميمات الاختبارات التكيفية متعددة المراحل، من حيث محكات البدء والإنهاء لكل تصميم، وعدد المراحل، والخصائص السيكمترية للصور الاختبارية المستخدمة وغيرها من المتغيرات التي تؤثر على جودة تقدير قدرة المفحوص.

2- نمذجة المعادلة البنائية (الهيكلية) Structural Equation Modeling (SEM):

تعتبر المعادلة البنائية مكوناً رئيساً في التحليلات الإحصائية متعددة المتغيرات. وأطلق عليها مسميات عديدة من أهمها نمذجة المعادلة البنائية، ونمذجة بنية التغاير، كما أطلق عليها العلاقات البنائية الخطية، وهي إطار تحليلي إحصائي يسمح للباحثين باختبار الفروض وتحديد واختبار النماذج بين المتغيرات المشاهدة والكامنة، كما تعرف بأنها أسلوب يستخدم لتحديد وتقدير نماذج العلاقات الخطية بين المتغيرات.

وتتضمن المعادلة البنائية نماذج عدة تتضمن تحليل المسار، والتحليل العاملي التوكيدي، والتحليل التمييزي، وتحليل التباين المتدرج، وتحليل الانحدار المتعدد

وغيرها (Schumacker & Lomax, 1996) ويتمثل الهدف من نموذج المعادلة البنائية في اختبار صحة الفروض حول العلاقات بين المتغيرات المقاسة والمتغيرات الكامنة ومن ثم يستخدم هذا الأسلوب في التحليل العاملي التوكيدي للتحقق من الصدق البنائي للمقاييس المختلفة التي يتم بناؤها في ضوء أطر نظرية معينة، كما تستخدم في تحليل المسار وتحليل الانحدار (عبد الناصر السيد عامر، 2004). وأصبحت نمذجة المعادلة البنائية SEM تقنية تحليل إحصائي قياسية في ذخيرة الباحثين في العلوم السلوكية والتعليمية والنفسية والاجتماعية. وتهدف الي تحديد مدى مطابقة النموذج النظري للبيانات الميدانية، أي المدى الذي يتم فيه تأييد النموذج النظري بواسطة بيانات العينة (Hair et al., 2018).

جدول رقم (8) مقارنة بين إحصاءات الجيل الأول والجيل الثاني

نوع الإحصاء		الفترة
التوكيدية	الاستكشافية	1st Generation الجيل الأول (1900-1990)
ANOVA Logistic Regression Multiple Regression	Exploratory Factor Analysis (EFA) Cluster Analysis Multidimensional Scaling	
التوكيدية	الاستكشافية	2nd Generation الجيل الثاني (1990-Now)
SEM (CB-SEM)	PLS-SEM (VB-SEM)	
AMOS (SPSS) LISREL SAS EQS, MPLUS SIMPLIS, PRELIS SePATH	Smart PLS Warp PLS Visual PLS Spad PLS PLS Graph PLS Gui	البرامج المستخدمة

جدول رقم (9) مقارنة بين إحصاءات النمذجة البارامترية والنمذجة اللابارامترية

تقييم النموذج	المدخل الإحصائي	هدف البحث	أسلوب التقدير	
مؤشرات المطابقة	الإحصاء البارامترى الانحدار، التباين، اختبار ت، F-Test	توكيدى	MLE Maximum Likelihood Estimator الارجحية القصى	SEM النمذجة البارامترية
بدون مؤشرات مطابقة	احصاء لابارامترى Wilcoxon Mann-Whitney Kruskal Wallis	استكشافى	OLS Ordinary Least Square المربعات الصغرى الموزونة	النمذجة اللابارامترية PLS_SEM

3- التحقق من صدق البنية العاملية للمقاييس النفسية والتربوية في ضوء مدخل نمذجة المعادلة البنائية:

تزايد مؤخرًا استخدام المعادلة البنائية في العلوم الاجتماعية والسلوكية، وساعد على انتشارها توافر كثير من المراجع، وبرامج الحاسب الآلي المتخصصة ومن أهمها MPLUS ، LISREL ، AMOS ، EQS (عزت عبد الحميد، 2016). ونظرًا لما توفره نماذج المعادلة البنائية من دلالات للصدق والثبات وطبيعة العلاقات بين المتغيرات فقد تزايد مؤخرًا استخدام نماذج المعادلة البنائية التي تتضمن تحليل المسار، والتحليل العاملي التوكيدي، والتحليل التمييزي، وتحليل التباين المتدرج، وتحليل الانحدار المتعدد وغيرها في دراسة العلاقات بين المتغيرات النفسية والتربوية وكذلك التحقق من البنية العاملية للمقاييس، واستهدفت عدد من الدراسات استخدام التحليل العاملي التوكيدي في التحقق من صدق البنية العاملية للمتغيرات النفسية والتربوية حيث وظفت تلك الدراسات ما يوفره النموذج من مؤشرات للملاءمة ودلالات للصدق والثبات في توفير مقاييس نفسية وتربوية تتمتع بخصائص سيكومترية مناسبة (Lin & Huang, 2016)

كما ظهرت دعوات باستخدام التحليل العاملي الاستكشافي للبندود بديلا للتحليل العاملي الاستكشافي للاختبار في حال المقاييس الرتيبة وذلك باستخدام برمجية R وبرنامج factor لصعوبة اداء ذلك من خلال برنامج SPSS، كما تطورت نقاط بحثية عديدة في مجال النمذجة والتحليل البعدي للنمذجة سواء المقرة

بالأرجحية القصوى أو المقدرة وفق نظرية بيز Bayes Theorem لمعالجة مشكلات التحيز القائمة على التحيز المعرفي أو التحيز الثقافي أو التحيز الجنسي والعنصري بحيث تسهم المداخل الحديثة في معالجة تلك المشكلات (صلاح علام، 2005) كما ظهرت العديد من النقاط التي تناقش نماذج التحليل العاملي التوكيدي المائل والتحليل العاملي التوكيدي المتعامد من ناحية ومقارنة نماذج التحليل العاملي التوكيدي الانعكاسية مقابل نماذج التحليل العاملي التوكيدي التكوينية، ومن ثم دراسات التحليل العاملي الاستكشافي والتوكيدي ما تزال تشغلان الساحة العلمية ومازال يظهر الجديد القديم من القضايا التي تبدو على الساحة في لباس علمي جديد وتحت مفاهيم علمية جديدة، كما توجد حاجة ملحة لتطوير مجالات مناهج البحث وكتابة التقارير بما يحقق جودة البحث العلمي من خلال تحقيق التكامل بين الدلالة الإحصائية والدلالة العملية في تقارير الأبحاث وتناول أحجام التأثير وقوة الاختبار الإحصائي لتفادي الأخطاء من الدرجة الأولى والأخطاء من الدرجة الثانية مع تناول النقاط الحديثة في تناول مناهج البحث وتكاملها مع الإحصاء (محمد بشير زينة، 2017).

4- التقويم البديل: Alternative Evaluation

يعتمد على الافتراض القائل بأن المعرفة يتم تكوينها وبناءها بواسطة المتعلم، حيث تختلف تلك المعرفة من سياق لآخر، وتقوم فكرة هذا النوع من التقويم على إمكانية تكوين صورة متكاملة عن المتعلم في ضوء مجموعة من البدائل بعضها أو جميعها.

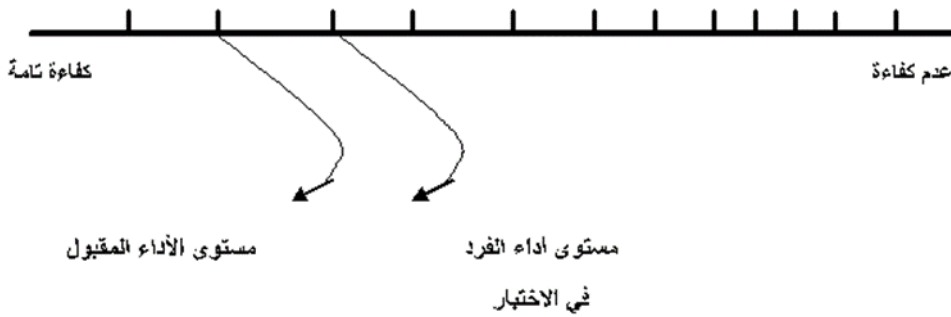
وتقوم فكرة التقويم البديل على الاعتقاد بأن تعلم الطالب وتقدمه الدراسي يمكن تقييمها بواسطة أعمال ومهام تتطلب انشغالا نشطاً مثل البحث والتحري في المشكلات المعقدة، والقيام بالتجارب الميدانية، والأداء المرتفع. من أبرز أدوات التقويم البديل وأكثرها تقويم الأداء: Performance Evaluation: حقائق التقويم (ملفات التعلم) (الأعمال) أو البورتفوليو Portfolios

ويستخدم التقييم البديل غالباً في البرامج التربوية الخاصة مع الأطفال الذين لديهم محدودية في قدراتهم على فهم الكلمات والمصطلحات والعبارات الواردة في التقييم أو الذين قدراتهم وإمكانياتهم للاستجابة اللفظية والحس حركية محدودة أو الأفراد الذين لا يستطيعون التعامل مع المهارات المعقدة لمعالجة المعلومات التي يتطلبها الاختبار أو الأفراد غير القادرين على ضبط سلوكهم أثناء موقف التقييم ولا يستطيعون (أو يجدون صعوبة) في فهم ما يسألون عنه في موقف التقييم. (صلاح علام، 2007).

مفهوم تقويم الأداء:

يتطلب تقويم الأداء Performance Assessment، أو التقويم القائم على الأداء Performance - Assessment based، أن يظهر المتعلم بوضوح،

أو يبرهن، أو يقدم أمثلة أو تجارب أو نتائج أو غير ذلك تتخذ دليلاً على تحقيقه مستوى تربوياً Standard، أو هدفاً تعليمياً معيناً. ويمكن استخدام هذه الاختبارات في تقييم أنواع مختلفة من التفاعلات سواء مع الأفراد، أو الأجهزة والآلات والمعدات، أو المواقف. ولعل اختبارات الأداء الميكانيكي، واختبارات قيادة السيارات، واختبارات الكفاءة في تشغيل الحاسبات الآلية، واختبارات الأداء الموسيقي، واختبارات الأداء الحركي، تعد أمثلة لهذه الاستخدامات. وترتبط فكرة مستويات الأداء بالاختبارات محكية المرجع وليست الاختبارات مرجعية المعيار فدرجة الفرد في اختبار ما يصعب تفسيرها دون تحديد معيار أو محك أداء معين (صلاح علام، 2007).



شكل رقم (2) متصل الأداء

مسميات مستويات الأداء:

وفى إطار ذلك يتم وضع مواصفات مستويات الأداء من خلال ورش عمل بمشاركة جميع المعنيين ثم الاتفاق على عددها من (3-5).
مثال: أربعة مستويات مسمياتها: (متعثر - نام - مرض - متقدم)

التوصيف العام لمستويات الأداء كما يلي:

1. متعثر: Falling Down: يمكنه تذكر بعض المعلومات الأولية ويعجز عن حل المشكلات البسيطة. "وهو يحتاج إلى مجهود وتحسين كبير جداً".
2. نام: Growing: يظهر فهماً جزئياً للمحتوي يمكنه من حل المشكلات البسيطة النمطية. "وهو يحتاج إلى تحسين".
3. مرض Satisfactory: يظهر فهماً جيداً للمحتوي يتمثل في حل مشكلات مركبة. "الحد الأدنى المطلوب تحقيقه".
4. متقدم "Advanced": فهو المستوي المطلوب تحقيقه وهو أعلى من المستوى المرضي.

يقدم أدلة وبراهين تؤيد آراء علمية متنوعة ومختلفة ويتعامل مع مواقف علمية غير مألوفة، يظهر فهما شاملا ومتعمقا وناقدا للمحتوي ويعطي حلول متنوعة للمشكلات غير النمطية، يحل مشكلات مألوفة بطرق غير مألوفة، يظهر مهارات تواصل فعالة وفهم متعمق يمكنه من ابتكار أفكار وحلول غير تقليدية لمدي واسع من الموضوعات، يحل أكبر قدر من المشكلات في وقت قصير بأقل قدر من الأخطاء مع إبداء الرأي في كل حل. (المركز القومي للامتحانات والتقويم التربوي، 2015).

التقويم البديل للطلاب ذوي الإعاقات:

يعد قياس تحصيل الطلاب ذوي الاحتياجات الخاصة وسمات شخصية من القضايا الحديثة والتي ينبغي أن تكون احد محاور الاهتمام والتطوير لعلم القياس والتقويم والإحصاء النفسي والتربوي ، حيث إن دمج هؤلاء الطلاب في التقويم يعني أنه سيعاد في نماذج القياس القديمة لتحديد كيفية التأكد من صدق نموذج القياس، والتطرق إلى التقييمات البديلة على أساس معايير التحصيل البديلة (AA- AAS) لذوي الاحتياجات الخاصة وما بعدها، وكيف يمكن تطوير معايير التحصيل البديلة عبر الزمن شكلا ومحتوى، وإجراءات محكات التصحيح، ومواصفات مستويات التحصيل، والتطورات الجديدة للمعايير البديلة وتقويمها، والدروس المستفادة منها، والتحديات المستمرة في تطوير وتنفيذ المعايير البديلة. (صلاح علام، 2004).

5- قضايا تتعلق بتقدير الدرجات للمفردات الاختبارية (أسئلة المقال).

من عيوب المفردات الاختبارية إنتاج الاستجابة محدودية الموضوعية والعدالة في تقدير درجات إجابات المتعلمين وطول الوقت اللازم لتقدير الدرجات (التصحيح) حيث يختلف تقدير المصححين في إعطاء الدرجة، إضافة إلى أن صياغتها تقتصر في كثير من الأحيان للوضوح مما قد يؤدي في بعض الأحيان إلى اختلاف المتعلمين في فهم المقصود منها، وهذا يؤدي إلى عدم توفيق البعض في التوصل إلى الإجابة المطلوبة ليس لضعف مستواهم في المادة، وإنما لعدم وضوح المطلوب من سؤال المقال. قد يكون لدى بعض الطلاب مهارة لغوية في التعبير الكتابي، وفي طريقة عرض وربط وتنظيم المعلومات، فيؤثر هذا على المصحح، مما يجعله يعطي درجة عالية جدا لموضوع قد لا يتضمن أفكارا جوهرية ذات قيمة، بمعنى أن الإطار الذي توضع فيه الإجابة، قد يؤثر على نوعية الدرجة بصرف النظر عن سلامة الموضوعات ودقتها. بل قد يختلف المصحح نفسه مع نفسه من وقت لآخر.

للمفردات الاختبارية دور مهم في تقويم نواتج تعلم المتعلمين، وتحديد مستوى تحقيقهم لمعايير الأداء الجيد؛ لذا تعد مهارة صياغتها أحد أهم المعايير التي ينبغي أن تكون ضمن معايير جودة التقويم التربوي، حتى لا يعزو مستوى أداء

المتعلم إلى عوامل لا دخل له بها كعدم وضوح تعليمات المفردة أو عدم وضوح المطلوب منه، أو عدم مراعاة الشروط الأخرى للصياغة الجيدة للمفردة. وتعتبر المفردات الاختبارية إنتاج الاستجابة (المقال) من أقدم أنواع المفردات وأكثرها شيوعاً واستعمالاً فهي مجموعة من المثيرات المقننة التي تقدم للمتعلم كي تستثير تفكيره، ليقيم استجابته عليها لتكون مؤشراً ودليلاً على حدوث عملية التعلم، ومن خلال تلك الاستجابات نحكم على مدى امتلاك المتعلم للسمة أو القدرة المقاسة بدقة وموضوعية. (محمد فتح الله، 2020)

وتواجه المفردات الاختبارية إنتاج الاستجابة عدة عيوب من أهمها أنها تفسح مجالاً واسعاً لتدخل العوامل الذاتية في التصحيح وظهور التباين والاختلاف بين المصححين في عملية تقدير الدرجات، بل يظهر أيضاً عند المصحح نفسه من وقت إلى آخر، وهذا ما يؤدي إلى خفض مستوى ثبات تقدير الدرجات. فضلاً عما يتطلبه عادة وقتاً طويلاً وجهداً كبيراً.

ومن أهم التطورات الحديثة في علم القياس والتقويم مقياس التقدير Rating Scale وهي أداة بسيطة تظهر مستوى مهارات المتعلم، حيث تخضع كل فقرة لتدريج من عدة فئات أو مستويات، حيث يمثل أحد طرفية انعدام وجود الصفة التي نقيدها بشكل ضئيل. ويمثل الطرف الآخر تمام أو كمال وجودها، وما بين الطرفين التدريج لمستوى وجودها.

ومن مزاياه: توفير جهد المعلم ووقته وتمتعه بدرجة من الموضوعية والثبات، وتحديد به بشكل واضح مواطن القوة والضعف في أداء المتعلم، ومدى ما أحرزه من تقدم، كما ويستخدم في تقويم أنواع مختلفة ومتعددة من أداء المتعلمين (ملحق 2)

وفي ضوء هذا التطور (مقياس التقدير) يتم وضع آليات لتقدير الدرجات للمفردات الاختبارية إنتاج الاستجابة تحقق الموضوعية والعدالة وارتفاع ثبات المقدرين وعلاج مشكلة ذاتية المقدرين.

خطوات إعداد مقاييس التقدير لدليل التصحيح:

تحليل المهارة أو المهمة المطلوبة في السؤال إلى مجموعة من المهام الأصغر، أو إلى مجموعة من المعارف المحددة في السؤال، ثم ترتيب المعارف المكونة للمهارة حسب تسلسل حدوثها، ثم اختيار التدريج المناسب على مقياس التقدير لتقدير مدى إنجاز المهارة.

مقياس التقدير Rating Scale:

أداة بسيطة تظهر مستوى مهارات المتعلم، حيث تخضع كل فقرة لتدريج من عدة فئات أو مستويات، حيث يمثل أحد طرفيه انعدام وجود الصفة التي نقيدها بشكل ضئيل. ويمثل الطرف الآخر تمام أو كمال وجودها، وما بين الطرفين التدريج لمستوى وجودها. ومن مزاياه: توفير جهد المعلم في تقدير الدرجات

لإجابات المتعلمين ووقته وتمتعه بدرجة من الموضوعية والثبات، وتحديدته بشكل واضح مواطن القوة والضعف في أداء المتعلم، ومدى ما أحرزه من تقدم، كما ويستخدم في تقويم أنواع مختلفة ومتعددة من أداء المتعلمين. (محمد فتح الله، 2020).

6- قضية التعامل مع القيم المفقودة في البيانات: Missing Data/Values

تتعرض جميع الاختبارات على اختلاف مستوى ودرجة أهميتها إلى قيام بعض المستجيبين بترك بعض المفردات دون الاستجابة عليها نهائياً. أو قد يحدث أثناء عملية إدخال البيانات يدوياً، وتؤدي ذلك إلى نتائج ومؤشرات إحصائية متحيزة. كما تأتي أهمية دراسة القيم المفقودة في كونها تعد دالة لعدة خصائص للمفحوص، منها ما يتعلق بخصائصه الشخصية، ومستوى قدرته، وكذلك المتغيرات الديمجرافية (De Ayala, Plake & Impara., 2001).

وفي هذا الإطار فقد ازداد اهتمام الباحثين بالبحث عن الطريقة الأكثر ملائمة لمعالجة البيانات المفقودة قبل استكمال إجراءات التحليل الإحصائي، وذلك للحصول على استدلالات أكثر دقة عن عينة ومجتمع الدراسة.

ومع التقدم الواسع في مجال الحاسب الآلي من ناحية وتقدم البرامج الإحصائية وظهور اتجاهات حديثة في مجال القياس كنظرية الاستجابة للمفردة، ومن ناحية أخرى فقد ازداد الاهتمام بدراسة مشكلة القيم المفقودة وآليات معالجتها. ووجدت العديد من الطرق الإحصائية للتعامل مع البيانات المفقودة، منها الطرق القائمة على الحذف وذلك بحذف الحالة أو الحالات التي تتضمن استجابات بها بيانات مفقودة، وهي الطريقة المفترضة في الكثير من البرامج الإحصائية، ومنها الطرق القائمة على التعويض وتهدف إلى اكتمال البيانات قبل تحليلها إما باعتبار القيمة المفقودة استجابة خاطئة، أو تعويض القيمة المفقودة بقيمة مقدرة من خلال القيم الملاحظة، أو اعتبار أن المفحوص لم يتعرض للمفردة التي لم يستجب عليها، أي وكان هذه المفردة غير موجودة بالاختبار (Peugh & Enders, 2004).

7- قضية معالجة القيم المتطرفة: Outliers

هي قيم غير طبيعية وتوزيعها مختلف عن باقي القيم ضمن العينة الواحدة، ووجود القيم الشاذة في البيانات يؤدي إلى استنتاجات مضللة وغير دقيقة وبعيدة عن الواقع بالنسبة للظواهر التي يتم دراستها، فكلما كانت البيانات دقيقة ونقية ستعكس نتائج دقيقة عن الظاهرة المدروسة (زايد صالح، 2018) وبعد موضوع القيم المتطرفة في البيانات الإحصائية وأثر وجودها وانعكاساته على نتائج ودقة التحليلات الإحصائية موضع اهتمام العديد من الباحثين في معظم بحوثهم. وزاد الاهتمام بموضوع القيم المتطرفة لأنه عادة ما يتم التعامل معها بالتجاهل، وهذا يؤدي إلى تقديرات ذات كفاءة أقل وتحيز في النتائج وتأثير عليها، لأن وجود القيم المتطرفة التي تختلف بشكل جوهري عن بقية البيانات يؤدي إلى اختلافات كبيرة

وتغيرات جوهرية وتشويه في النتائج الممثلة لبيانات العينة، وعليه يتطلب الأمر البحث والتحقق في هذه المشاهدات غير العادية والبحث عن طرائق للسيطرة على تأثيرها.

وهناك دراسات عديدة قام بها الباحثون لدراسة موضوع القيم المتطرفة واختلفت آراؤهم حول موضوع القيم المتطرفة، إلى أن القيم المتطرفة هي مشاهدات غير طبيعية وتوزيعها مختلف عن باقي المشاهدات ضمن العينة الواحدة، ووجودها بالبيانات يؤدي إلى استنتاجات بعيدة عن الواقع بالنسبة للظواهر التي يتم دراستها، ويتم تحديد القيم المتطرفة في البيانات والكشف عنها، طريقة Tukey للكشف عن القيم المتطرفة، الانحراف المعياري Standard Deviation، الدرجة المعيارية Z-Score ومن أساليب التعامل مع القيم المتطرفة بعد أن تم تحديد القيم المتطرفة والكشف عنها وتشخيصها في البيانات وبيان أسباب شذوذها وأصنافها يتم معالجتها بإحدى الطرائق وهي طريقة الحذف، والاحتفاظ بالقيم المتطرفة، واستبدال القيم المتطرفة في البيانات حيث يتم التعامل مع القيم الشاذة في البيانات باستخدام الأساليب الإحصائية المناسبة، واستبدال القيم المتطرفة بطرائق الاستبدال المختلفة التي من أهمها طريقة الوسط الحسابي المبتور Trimmed Mean، الوسط الحسابي التعويضي Winsorized Mean، زيادة حجم العينة. (إخلاص العطيان، 2018).

8- قضية انتهاك افتراضات نظرية الاستجابة للمفردة الاختبارية:

يعرف الانتهاك Violation على أنه تجاوز مفردات الاختبار بشرط أو افتراض محدد ينبغي أن يتوافر بالاختبار، ومعدل الانتهاك: هو نسبة عدد المفردات المنتهكة لافتراضات نظرية الاستجابة للمفردة الاختبارية إلى العدد الكلي للمفردات المكونة للاختبار (محمد الشافعي، 2014)

الافتراضات الأساسية لنماذج نظرية الاستجابة للمفردة: أي نموذج رياضي يتضمن مجموعة من الافتراضات تكون بمثابة دليل عملي أثناء تطبيق النموذج على بيانات الاختبار ومن أهم الافتراضات:

أ- أحادية البعد Unidimensionality: يعني أن تكون جميع المفردات الاختبارية متجانسة أي تقيس خاصية واحدة فقط وأن هناك قدرة كامنة وحيدة هي التي تفسر أداء الفرد في الاختبار.

ب- الاستقلال الموضوعي Local independence: احتمال الإجابة الصحيحة للفرد على مفردة الاختبار يكون مستقلاً عن أجابته على أي مفردة أخرى في ذلك الاختبار.

ج- التحرر من عامل السرعة في الإجابة Speedness: أن نماذج الاختبارات التي تكون ملائمة لنماذج الاستجابة للمفردة لا يتم إعدادها تحت شرط السرعة أي أن إخفاق بعض الأفراد في الإجابة على بعض مفردات الاختبار يرجع إلى انخفاض قدراتهم وليس إلى عامل السرعة في إجاباتهم (صلاح علام، 2005).

د- منحنى خاصية المفردة (Item Characteristic Curve (ICC):

هو دالة رياضية تربط بين احتمال إجابة المفحوص على الفقرة إجابة صحيحة وبين السمة أو القدرة التي تقيسها المفردة أو يقيسها الاختبار الذي يحتوي على هذه الفقرة وتسمى هذه الدالة بمنحنى خاصية المفردة والتي تمثل بانحدار غير خطي لاحتمال الإجابة الصحيحة على المفردة بدلالة القدرة التي يتم قياسها بالاختبار.

الانتهاك يحدث من خلال الإخلال بأي افتراض من الافتراضات الرئيسية لنظرية الاستجابة للمفردة ويتم التعرف عليه من خلال مطابقة المفردات للبرنامج المستخدم وكذلك المعايير المستخدمة لتحقيق الافتراضات الرئيسية للنظرية مثل استخدام (χ^2) إذا كانت قيمة كا2 أقل من 0.05 تكون المفردة دالة إحصائية أي أنها غير مطابقة كما تستخدم $P(z)$ للدلالة على مطابقة المفردة للنموذج المستخدم فإذا كانت أقل من 1.96 دل ذلك على عدم مطابقة المفردة أي أنها منتهكة لافتراضات النظرية، دلت النتائج التي أجريت على معرفة أثر معدل انتهاك لافتراضات نظرية الاستجابة للمفردة الاختبارية على قدرات الافراد من ذوي التحصيل المرتفع وذوي التحصيل المنخفض في السمة المقاسة إلى أن مستوى القدرة لدى الطلاب ينخفض بانخفاض نسبة المفردات المنتهكة. (أياد حمادنه، 2011).

9- قضية الأداء التفاضلي (تحيز) مفردات الاختبار Item Bias

تعتبر قضية تحيز الاختبارات Test Bias وتحيز المفردات Item Bias من المشكلات التي تهدد صدق الاختبار؛ وذلك لارتباط الصدق الوثيق ببنية الاختبار وكيفية استخدامه وارتباطه بالاستنتاجات المبنية عليه.

وترجع أهمية دراسة تحيز الاختبارات إلى ضرورة التمييز بين الفروق الحقيقية والفروق غير الحقيقية بين المجموعات، حيث إن الاختبار المتحيز تظهر نتائجه فروقا زائفة بين المجموعات يترتب عليها تقييم خاطئ واتخاذ قرارات غير صحيحة بشأن هذه المجموعات (Montesinos & Benito, 2003).

وتتضمن دراسة التحيز جانبين أولهما إحصائي وهو ما يعرف بالأداء التفاضلي للمفردة / الوظيفية الفارقة للمفردة Differential Item Functioning (DIF) حيث يستخدم الأداء التفاضلي للفقرة لتحديد الفقرات التي يكون احتمال الإجابة الصحيحة مختلفا بين مجموعتين لهما نفس مستوى القدرة، وعند تفسير أسباب ذلك يمكن وصف الفقرات ذات الأداء التفاضلي بأنها فقرات متحيزة لمجموعة ما دون الأخرى لأسباب لا تتعلق بقدرة الأفراد فقط. والجانب الآخر يتلو الشق الإحصائي وهو تحكيمي يعتمد على فحص الخبراء والمتخصصين لمصادر الفروق التي أظهرتها لطرق الإحصائية في الاستجابة على المفردة. وتحديد مدى ارتباطها بالغرض من الاختبار (Camilli. & Shepard, 1994).

وتتعدد الطرق الإحصائية التي تستخدم لدراسة التحيز حيث تعتمد بعض هذه الطرق على النظرية الكلاسيكية ومنها: طريقة الصعوبة المعدلة للمفردة، وطريقة تحليل التباين، والطرق المعتمدة على جداول التوافق (طريقة ماننيل هينزل، طريقة الانحدار اللوغاريتمي). واعتمدت بعض الطرق على النظرية الحديثة ومنها: طريقة مربع كاي للورد، وطريقة المساحة بين المنحنيات المميزة للمفردة لراجو. (إبراهيم محمد علي، 2017)

وتهتم الدراسات المعاصرة في مجال القياس والتقييم بمقارنة الطرق القائمة على النظريتين في علاج مشكلة تحيز المفردات والاختبارات كما تسعى إلى دراسة تأثير التحيز في الاختبارات على دقة تقدير بارامترات الأفراد والمفردات. كما تسعى بعض الدراسات لفحص تأثير التحيز على عدد من المتغيرات باستخدام بيانات واقعية ومقارنتها بالبيانات المولدة عن طريق المحاكاة على عينات كبيرة وتحت شروط مختلفة. (إبراهيم محمد علي، 2017)

10- الوعي بأساليب المعالجة والدلالة الإحصائية ودراسة حجم التأثير:

من الأهمية بمكان أن يكون الباحث على وعى بأسس اختيار الأساليب الإحصائية المناسبة لمعالجة بيانات بحثه أو دراسته في ضوء المنهج المستخدم وهدف البحث ونوعية البيانات وحجم العينة وعدد المجموعات البحثية والشروط والافتراضات الخاصة بكل أسلوب إحصائي حتى يصل إلى نتائج صحيحة ودقيقة وموثوق فيها.

يجب اهتمام الباحثين بدراسة حجم التأثير كأسلوب مكمل للدلالة الإحصائية وعدم الاعتماد المطلق على الدلالة الإحصائية وحدها في اختيار الفروض البحثية وتفسير نتائجها لافتقاد الدلالة الإحصائية إلى الدلالة العملية أو النفسية، وكما أشار (فواد أبو حطب، 1996) فالدلالة الإحصائية لا يتجاوز معناها أن المتغير المستقل له أثر في التغير التابع وهي لا تقيس قوة التأثير أو العلاقة بين المتغيرين. أن استخدام حجم التأثير Effect Size يزود الباحث بمؤشر للدلالة العملية والعلمية بجانب الدلالة الإحصائية، كما يساعد في المقارنة الكمية بين نتائج الدراسات، وتحليل قوة الاختبار الإحصائي، والكشف عن الفروق الحقيقية ومستوى وحجم التأثير أو درجة فعالية البرامج التدريبية.

المرحلة السادسة: رقمنة القياس والتقييم والإحصاء النفسي والتربوي (بنوك الأسئلة)

تهدف هذه المرحلة إلى التعريف ببنوك الأسئلة، مزايا إنشاء بنك الأسئلة باستخدام نظرية الاستجابة للمفردة، مستلزمات تطبيق فكرة بنوك الأسئلة، تحديد المعوقات التي تواجه بناء بنوك الأسئلة، التصور المقترح لتطوير منظومة التقييم الجامعي في ضوء بنوك الأسئلة الإلكترونية باستخدام نظرية الاستجابة للمفردة.

(1) نظرية الاستجابة للمفردة Item Response Theory:

تعد نظرية الاستجابة للمفردة من أهم نظريات القياس الموضوعي التي ظهرت في النصف الثاني من القرن العشرين، حيث سعت تلك النظرية لتحقيق عدالة وموضوعية القياس ويقصد بها تحرره من تأثير كل من الأداة المستخدمة (طالما أنها أداة قياس مناسبة)، ومن العناصر (الأفراد) التي استخدمت هذه الأداة في تقديرهم، بالإضافة إلى تدرج هذه الأداة بوحدة قياس مطلقة ثابتة تتوافق مع تدرج مستويات المتغير موضوع القياس (Brotherton & Preece, 1995). وهذا يتطلب أن يشترك كل من قياس الفرد وتدرج المفردات على متصل خطي واحد، وحداته متساوية؛ ومن ثم يسمح هذا النظام بتقييم القياسات خارج حدود الأداة التي أجرى بها القياس. (صلاح علام، 2002)

كما تفترض تلك النظرية IRT أنه يمكن التنبؤ بأداء الأفراد أو تفسير أدائهم في اختبار نفسي أو تربوي في ضوء خصائص تميز هذا الأداء، تسمى سمات أو قدرات، وهذه السمات غير ملاحظة، ولكن يمكن قياسها من خلال أداء الأفراد على مفردات الاختبار (Fraley et al, 2000)، (أمينة كاظم، 2000).

كما تسمى هذه النظرية أيضا بنظرية المنحنى المميز للمفردة ICC (1990)، Item Characteristic Curve Theory (Suen)، حيث يستخدم المنحنى المميز للمفردة في عمل استقرارات حول السمة الكامنة من خلال الاستجابات الملاحظة، ويمثل هذا المنحنى العلاقة بين احتمال الاستجابة الصحيحة لمفردة من مفردات الاختبار والسمة، أو القدرة موضوع القياس. (Fraley et al, 2000)، (أمينة كاظم، 2000).

ويطلق عليها أيضا نظرية السمات الكامنة (Latent Trait Theory)، ونظرية الاستجابة للمفردة (Item Response Theory) وتوفر نظرية الاستجابة للمفردة مؤشرات إحصائية ثابتة للاختبار ومفرداته لا تختلف باختلاف عينة الأفراد، مما يوفر قياسا أكثر ثباتا مع أخطاء أقل للقياس (صلاح علام، 2005)

ومن أهم مزاياه لتلك النظرية (Hambleton & Swaminathan, 1985):

عندما يتوافر عدد كبير من الفقرات الاختبارية التي تقيس السمة، فإنه يتم تقدير قدرة الفرد بشكل مستقل عن عينة الفقرات التي أجاب عنها، أي أن تقدير قدرة الفرد متحرر من خصائص الفقرات المستخدمة في تقديرها (Free - Item)

عندما يتوافر عدد كبير من الأفراد، فإنه يتم تقدير معالم الفقرات بشكل مستقل عن عينة الأفراد الذين طبقت عليهم هذه الفقرات، أي أن تقدير معالم الفقرات متحرر من خصائص عينة الدراسة. (Sample - Free)

يمكن الحصول على إحصائي خاص بدرجة الدقة عند قياس قدرة كل فرد، وتختلف قيمة هذا الإحصائي من فرد لآخر.

قدمت طريقة في انتقاء الفقرات تتميز بتقديم معالم ثابتة للمفردة هي: معالم الصعوبة والتمييز والتخمين، إضافة إلى توحيد المقياس لكل من الصعوبة وقدرة المفحوص لكي يتمكن مطورو الاختبارات من انتقاء الفقرات الأكثر ملاءمة

لتصنيف الأفراد على مستويات الأداء، التي تتميز في حالة مطابقة البيانات للنموذج بعدم تغير معالم الفقرات بتغير عينة المفحوصين، وعدم تغير تقديرات القدرات بعينة الفقرات، بالإضافة لذلك فإنه يمكن مقارنة معالم الفقرات والقدرة لوقوعهما على المقياس نفسه، فخاصية التغير في نظرية الاستجابة للفقرة تجعل هناك إمكانية لحل مشكلات القياس والاختبار التي من الصعوبة حلها في النظرية الكلاسيكية، وبشكل خاص؛ معادلة الاختبار، وبنوك الأسئلة، وتحيز الفقرات والاختبارات التكيفية (أمانة محمد كاظم 1996 ب).

توفر نظرية الاستجابة للمفردة عددا كبيرا من قيم الثبات، وذلك عن طريق الدالة المعلوماتية للاختبار (Test Information Function)، والدالة المعلوماتية للفقرة (Item Information Function)، فقيمة هاتين الدالتين تتوقف على مستوى القدرة، لذا فإن مقدار ما تقدمه الفقرة من معلومات يختلف باختلاف مستويات الأفراد، فالفقرة قد توفر معلومات هامة بالنسبة للأفراد ذوي القدرة المتوسطة، ومعلومات أقل أهمية بالنسبة للأفراد ذوي القدرة العالية أو الضعيفة (صلاح علام، 2005).

وتستند نظرية الاستجابة للمفردة إلى نماذج رياضية احتمالية تعمل على ضبط العوامل المركبة المختلفة التي تؤثر في المقارنات بين الدرجات وذلك بإجراء نوع من النمذجة الرياضية لبارامترات العملية الاختبارية. ومنها النماذج الأحادية المعلم وثنائية وثلاثية المعلم. (صلاح علام، 2005).

وتعد بنوك الأسئلة (Item Banking) ومن أهم التطبيقات لنظرية الاستجابة للمفردة.

(2) بنوك الأسئلة Item Banking

تعد بنوك الأسئلة من أهم التطورات المعاصرة في مجال القياس والتقويم والإحصاء النفسي والتربوي، فتهدف بنوك الأسئلة إلى تحقيق مجمل معايير التقويم من موضوعية وعدالة القياس والحكم والشفافية والشمول وغيرها، فضلا عن كونها أحد التطبيقات العملية للنماذج السيكمترية التي توفر عدد كبير من المفردات في محتوى دراسي معين، هذه المفردات لها خصائص محددة، وتستخدم في بناء اختبارات تحصيلية متنوعة طبقاً لما تستهدفه عملية التقويم.

كما يعد التقويم الإلكتروني Electronic assessment، المعروف أيضاً باسم التقويم الرقمي digital assessment أو التقويم عبر الإنترنت online assessment أو التقويم الحاسوبي computer-based assessment، هو استخدام تكنولوجيا المعلومات في التقويم مثل التقويم التربوي والتقويم الصحي والتقويم النفسي والقياس النفسي. يغطي هذا مجموعة واسعة من الأنشطة التي تتراوح من استخدام معالج النصوص للتعينات إلى الاختبار على الشاشة. تشمل الأنواع المحددة للتقويم الإلكتروني الاختيار من متعدد، والتسليم عبر الإنترنت / إلكترونياً، والاختبار التكيفي المحوسب computerized adaptive testing،

واختبار التصنيف المحوسب computerized classification testing .
(Ponton & Baker, 2008, Rovai).

(3) بنوك الأسئلة المحوسبة:

يعرف بنك الأسئلة بأنه: مكان آمن توضع فيه مجموعات متنوعة من الأسئلة ذات مستويات مختلفة في تقدير قدرة معينة ويسهل عن طريقه سحب أو إضافة مجموعة أو عدد من الأسئلة المختلفة المقننة أي التي لها خصائص سيكومترية مميزة ومعلومة مثل: معامل السهولة والصعوبة، ومعامل التمييز، وفعالية المشتتات وكذلك صدق وثبات المفردات، والمصنفة وفق وحدات المقرر الدراسي وحسب المستويات العقلية المعرفية المطلوب أداؤها أثناء الإجابة عليها، وذلك بطريقة تشبه إلى حد ما تنظيم وفهرسة الكتب. وهذا التصنيف يمكننا من معرفة كل سؤال والهدف الذي يقيسه والمجال الذي ينتمي إليه كل سؤال. وتخزن الأسئلة في ذاكرة حاسب آلي وفق برنامج سابق التجهيز، ويتم معايرتها عن طريق نماذج معينة وباستخدام برامج خاصة بها. (Havens, 2002)

وتتم الاستعانة بالكمبيوتر في مجال التقويم النفسي والتربوي من خلال تصميم ما يعرف ببنوك الأسئلة، والفكرة التي تقوم عليها تتمثل في تخزين كم هائل من الأسئلة في مختلف مجالات المعرفة، مع تصنيف هذه الأسئلة في ذاكرة الحاسب طبقاً لخصائص عديدة، بحيث يكون لكل سؤال خصائصه الإحصائية الناتجة عن تطبيقه في الميدان لحساب هذه الخصائص، ويتم عمل تدريج لهذه الأسئلة طبقاً لمستويات السهولة والصعوبة متحررة من طبيعة خصائص أفراد العينة، وكذلك معادلة هذه الأسئلة لإمكانية عمل المقارنات بين نتائجها من حيث المستوى، ويعتمد ذلك على برامج إحصائية وبرامج تخزين واستدعاء باستخدام أحد نماذج الاستجابة للمفردة.

(4) مزايا إنشاء بنك الأسئلة باستخدام نظرية الاستجابة للمفردة:

- أ- موضوعية الاختبارات المسحوبة منه وعدالتها
- ب- تحرر صعوبة المفردة من توزيع أداء عينة الأفراد.
- ج- تحرر قدرة الفرد من تأثير صعوبة المفردات:
- د- إمكانية سحب صور اختبارية متعادلة Equated Test Forms
- هـ- إضافة إلى اقتصادية التكلفة، والأهداف، والموضوعية، العدالة، الكفاءة، الموثوقية، الاستقرار، المتانة.

(5) معوقات تواجه بناء بنوك الأسئلة:

- ندرة الخبرة في بعض مجالات بنوك الأسئلة خاصة تدريج الاختبارات ومعادلتها.
- أعمال التطبيق الاختبارية وضبط الموقف الاختباري تقدير الدرجات لأسئلة إنتاج الإجابة.

المرحلة الأولى: التخطيط وتحديد المهام	المرحلة الثامنة: إعداد مجموعة من البرامج
المرحلة الثانية: التدريب وإعداد الكوادر	المرحلة التاسعة: مراجعة البنك
المرحلة الثالثة: إعداد مواصفات بنك الأسئلة	المرحلة العاشرة: تصميم نظام كودي لتخزين الأسئلة
المرحلة الرابعة: كتابة مفردات بنك الأسئلة	المرحلة الحادي عشر: وضع خطة تنظيمية لاستخدام البنك
المرحلة الخامسة: تحديد الخصائص السيكومترية لمفردات الاختبار (التجريب والتدريج):	المرحلة الثاني عشر: متابعة عملية تطوير البنك
المرحلة السادسة: التنظيم والتخزين	المرحلة الثالث عشر: إصدار دليل البنك
المرحلة السابعة: الاستخدام الميداني	المرحلة الرابعة عشر: تحرير البنك على المنصة الإلكترونية

جدول (10) تصور (مقترح) لنموذج هيكل الخطة الزمنية الإجمالية لإدارة بنوك الأسئلة

المخرجات النهائية	الإجراءات المطلوبة	المدة الزمنية 2021-2022					تكلفة الوحدة (المادة)	الاحتياجات والموارد
		2021/1	2021/2	2021/3	2021/4	2021/5		
البنك الإلكتروني أسئلة الصفوف الثلاث للمرحلة الثانية لجميع المقررات	1- تشكيل الفرق لكل مادة							
	2- الإطار المفاهيمي للاختبار في كل مادة ويتضمن:							
	1-2- دراسة وفحص المعايير القومية لنواتج التعلم 2007 ومعايير هيئة الجودة والاعتماد التربوي للمواد الدراسية المختارة (لكل مادة دراسية).							
	2-2- تحليل محتوى للمناهج الدراسية للمواد المختارة.							
	3-2- تحديد مواصفات البنك ويتضمن المفاهيم، والأهداف، والمستويات المعرفية المحددة ونسبة الأسئلة التي تقيسها وصياغة نواتج التعلم في صورة أهداف سلوكية..							

الاحتياجات والموارد	تكلفة الوحدة (المادة)	المدة الزمنية 2021 - 2022							الإجراءات المطلوبة	المخرجات النهائية
		2021/12	2021/11	2021/10	2021/9	2021/8	2021/7	2021/6		
6- منسقي وموافقي بيانات. ثانياً: الموارد المادية									3- بناء المفردات الاختبارية.	
1- Clustered Server (With special specification UPS 2- optical 3- scanners (to be estimated) 4- برنامج لإدارة بنوك الأسئلة.									3- 1-item Card ومراجعتها.	
									3- 2- تحكم المفردات الاختبارية.	
									4- عمل خرائط الاختبارات وأكواد الاختبار ومفتاح التصحيح.	
									5- سحب صور اختبارية استطلاعية.	
									6- التطبيق الاستطلاعي للمفردات لتحديد خصائصها السيكومترية. (الترم الأول) (والترم الثاني)	
									6- 1- عمل دليل تطبيق مقنن.	

الاحتياجات والموارد	تكلفة المادة	المدة الزمنية 2021 - 2022							الإجراءات المطلوبة	المخرجات النهائية
		2022/7	2022/6	2022/5	2022/4	2022/3	2022/2	2022/1		
5- برنامج لإدارة التحليل الإحصائي.									13- إدخال البيانات	
6- برنامج لإدارة البيانات.									14- تدقيق ومراجعة البيانات.	
7- برنامج لإدارة العينة.									15- تحليل نتائج التطبيق الاستطلاعي في ضوء النظريات الحديثة للقياس IRT والكلاسيكية.	
8- برنامج لإدارة الكنترول.										
9- برنامج										

إدارة التصحيح.								الترم الأول والثاني).	
								16- تخزين المفردات وإحصاءاتها.	
								17- تكوين الصور الاختيارية واختبار تعادل الصور الاختيارية.	
								18- تخزين الصور الاختيارية النهائية.	
								19- تكوين صور اختباريه متعددة لنفس الاختبار.	
								20- استدعاء صور متكافئة للاختبارات في ضوء ما سبق إعداد من مواصفات اختباريه.	

الاحتياجات والموارد	تكلفة المادة	المدة الزمنية 2021 - 2022							الإجراءات المطلوبة	المخرجات النهائية
		2022/4	2022/3	2022/2	2022/1	2021/12	2021/11	2021/10		
									التكلفة الإجمالية للمادة الواحدة	
									21- توحيد فكر الأسئلة من خلال التدريب على إعداد الأسئلة بمستوياتها المعرفية المختلفة وعمل دليل التصحيح تكوين الصور الاختيارية	

									لجميع المواد.
									22- نشر ثقافة بنوك الأسئلة قبل وأثناء وبعد تنفيذ المشروع (إعداد مطويات عن البنوك).
	10% من الإجمالي								23- خدمات إدارية ومستلزمات وأدوات مكتبية. الإجمالي النهائي لجميع المواد.

الاحتياجات والموارد

- 2. Clustered Server
- 3UPS (at least)
- 3optical scanners (to be estimated)
- برنامج لإدارة بنوك الأسئلة (بجميع مراحلها).
- برنامج لإدارة التحليل الإحصائي.
- برنامج لإدارة البيانات.
- برنامج لإدارة العينة.
- برنامج لإدارة الكونترول.
- برنامج لإدارة التصحيح.

خلاصة الورقة:

هدفت هذه الورقة إلى وضع رؤية مستقبلية لتطوير تخصص القياس والتقويم والإحصاء النفسي والتربوي بالجامعات المصرية وقد اتبع الباحث منهج التخطيط الإستراتيجي في تقديم هذه الرؤية المقترحة، فقد تم وصف واقع القياس والتقويم والإحصاء النفسي والتربوي بالجامعات المصرية، وكذلك تحليل هذا الواقع وبيان جوانب القوة والضعف والفرص والتحديات وبناء على نتائج هذا التحليل، ثم صياغة الرؤية المستقبلية لتطوير تخصص القياس والتقويم والإحصاء النفسي والتربوي من خلال تحديد الرؤية المقترحة، والرسالة التي يجب أن يحملها، والقيم التي يستوجب عليه استهدافها، والأهداف الإستراتيجية

المطلوب تحقيقها، وآليات لتنفيذ هذه الرؤية المستقبلية ووضع مؤشرات أداء رئيسية لقياس مدى التقدم في تحقيقها.

التوصيات:

- في ضوء الرؤية المستقبلية التي طرحت في هذه الورقة، يقترح الباحث ما يلي:
- 1- إنشاء قسم خاص بمجال القياس والتقويم والإحصاء النفسي والتربوي يتبع مرحليا قسم علم النفس التربوي لحين إصدار التشريعات اللازمة لذلك.
 - 2- تبني مشروع لتطوير مقررات القياس والتقويم والإحصاء النفسي والتربوي على مستوى الجامعات المصرية (الحكومية - الأهلية - الخاصة) والمراكز البحثية.
 1. توحيد توصيف مقررات القياس والتقويم والإحصاء النفسي والتربوي على مستوى الجامعات المصرية، وكذا طرق وأساليب قياس وتقويم تحصيل المتعلمين.
 - 3- تبني دراسة مسحية على مستوى التخصص للتعرف على اتجاهات البحوث والفجوات في المعرفة بالمجال.
 2. 5تبني مشروع لربط الدراسات والبحوث باحتياجات المجتمع.
 3. عقد مؤتمر عام في التخصص يتناول واقع القياس والتقويم والإحصاء النفسي والتربوي والتحديات التي يواجهها وكيفية مواجهتها.
 4. 7تبني مشروع لترجمة البحوث العربية إلى اللغة الانجليزية وبخاصة أدوات القياس لتعزيز الوجود العربي على شبكة المعلومات الدولية لإبراز هوية علم النفس العربي.

المراجع:

- إبراهيم محمد على أحمد، (2017). الكشف عن الأداء التفاضلي لمفردات اختبار تحصيلي في الرياضيات لدى تلاميذ المرحلة الإعدادية وفقا لنظريتي القياس التقليدية والاستجابة للمفردة، رسالة دكتوراه غير منشورة، كلية التربية جامعة الزقازيق.
- إخلاص العطيان (2018). أثر معالجة القيم المتطرفة في اختبارات تيمس الدولية (TIMSS,2011) على تقديرات المعالم للفقرات ودقتها وفقاً لنظرية الاستجابة للفقرة. مجلة جامعة النجاح للأبحاث (العلوم الانسانية) المجلد 32 (8).
- أمطانيوس ميخائيل (2015). القياس والتقويم النفسي والتربوي للأسوياء وذوي الحاجات الخاصة. عمان : دار الإعصار العلمي للنشر والتوزيع.
- أمينة محمد كاظم (1988). دراسة نظرية نقدية حول القياس الموضوعي للسلوك (نموذج راش). الكويت: سلسلة الكتب المختصة، مؤسسة الكويت للتقدم العلمي.

أمينة محمد كاظم (1994). تدريج ومعايرة المقاييس: معايير نمو طفل ما قبل المدرسة. في: محمد عماد الدين إسماعيل وناهد رمزي وليلى كرم وأمينة كاظم وهدي ناشف، المجلد الثاني، الدراسة النفسية، القاهرة: رئاسة مجلس الوزراء، المجلس القومي للطفولة والأمومة، 114-232.

أمينة محمد كاظم (1996 أ). دراسة نظرية نقدية حول القياس الموضوعي للسلوك - نموذج راش". في أنور الشرقاوي وآخرون، اتجاهات معاصرة في القياس والتقويم النفسي والتربوي، القاهرة: الأنجلو المصرية، ص ص 281-430.

أمينة محمد كاظم (1996 ب). مستوى العينة وتدرج بنك الأسئلة باستخدام نموذج "راش" دراسة تجريبية في أنور الشرقاوي وآخرون، اتجاهات معاصرة في القياس والتقويم النفسي والتربوي، القاهرة: الأنجلو المصرية، 547-583.

أمينة محمد كاظم (2000). اتجاهات معاصرة في بناء بنوك الأسئلة. في: الأسس التربوية لإعداد المعلم الجامعي. القاهرة: جامعة عين شمس، ط3، 321 - 342.

أياد محمد حمادنة (2009). استخدام نظرية الاستجابة للمفردة في بناء اخبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلم، مجلة العلوم النفسية والتربوية، الأردن. م (10)، ع (2). ص 215 - 249.

راشد حماد الدوسري (2004). القياس والتقويم التربوي الحديث. عمان: دار الفكر للنشر والتوزيع.

زايد صالح ابراهيم بني عطا (2018). أثر القيم المتطرفة على الأداء التفاضلي لفقرات اختبار الرياضيات في الدراسة الدولية تيمس تبعًا لمتغير الجنس، دراسات العلوم التربوية، المجلد 45 ، العدد 4 ، ملحق 2. الجامعة الاردنية، عمادة البحث العلمي.

شحته عبد المولى (1999): تقويم بناء الاختبارات المرجعة إلى المحك / المعيار في ضوء نظرية الاستجابة للمفردة والنظرية التقليدية. رسالة دكتوراه، كلية التربية، جامعة عين شمس.

صالح محمد علي أبو جادو، (2005). علم النفس التربوي، دار المسيرة، عمان، الأردن.

صفوت فرج (2002). القياس النفسي. (ط6). القاهرة : مكتبة الأنجلو المصرية .

صلاح أحمد مراد (1997). اتجاهات معاصرة في القياس النفسي. المؤتمر التربوي الأول بسلطنة عمان، اتجاهات التربية وتحديات المستقبل (7- 10 ديسمبر)، ص ص: 1-19.

صلاح الدين محمود علام (1991). : دراسة مقارنة لبعض طرق تحديد مستويات الأداء في اختبار مرجعي المحك " المجلة المصرية للدراسات النفسية،

- تصدرها الجمعية المصرية للدراسات النفسية، القاهرة، العدد1، سبتمبر ص 77-96.
- صلاح الدين محمود علام (2002). القياس والتقويم التربوي والنفسي، القاهرة، دار الفكر العربي.
- صلاح الدين محمود علام (2004). التقويم التربوي البديل أسسه النظرية والمنهجية وتطبيقاته الميدانية، القاهرة، دار الفكر العربي.
- صلاح الدين محمود علام (٢٠٠٥). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي، الطبعة الأولى، دار الفكر العربي، القاهرة.
- صلاح الدين محمود علام (2007). الاختبارات التشخيصية محكية المرجع في المجالات التربوية والنفسية والتدريبية. القاهرة: دار الفكر العربي.
- عادل كمال خضر (1999). وضع البحث النفسي في مصر. الهيئة المصرية للكتاب: مجلة علم النفس، 52، 20-38.
- عبد المجيد نشواتي (2003)، علم النفس التربوي (الطبعة الرابعة)، عمان: الفرقان للنشر والتوزيع، صفحة 17-20.
- عبد الناصر السيد عامر (2004). أداء مؤشرات حسن المطابقة لتقويم نموذج المعادلة البنائية، المجلة المصرية للدراسات النفسية، 14(45)، 105-157.
- عزت عبد الحميد حسن (2016). الإحصاء المتقدم للعلوم التربوية والنفسية والاجتماعية تطبيقات باستخدام برنامج Listel 8.8. القاهرة: دار الفكر العربي.
- فؤاد أبو حطب وآمال صادق (1996). علم النفس التربوي. القاهرة: الأنجلو المصرية.
- محمد المري محمد إسماعيل (2011). تقييم بعض البحوث والدراسات العربية والأجنبية في مجال العلوم النفسية. دراسات تربوية ونفسية، مجلة كلية التربية جامعة الزقازيق، 70، ص 273-298
- محمد بشير زينة (2017). لغة البرمجة الإحصائية R (الإصدار الأول). حلب.
- محمد حبشي (2011). الخصائص السيكمترية لاستبانة إدارة الوقت لدى عينة من طلاب الجامعة في مصر والسعودية: دراسة تقويمية لنظرية القياس التقليدية ونظرية الاستجابة للمفردة. المجلة التربوية. 25(99) الجزء الأول. 353-410.
- محمد محمد فتح الله (2016). بنوك الأسئلة الإجراءات التنفيذية والمتطلبات ونماذج تطبيقية ورقة عمل قدمت لمؤتمر العلمي الدولي الثاني للقياس والتقويم في مصر جامعة الزقازيق بعنوان " القياس والتقويم والمؤسسات التعليمية – الواقع والرؤى المستقبلية، يومي السبت والأحد 30-31 يونيو 2016 م

- محمد محمد فتح الله (2018). بنوك الأسئلة والتصحيح الإلكتروني التطوير الحقيقي لمنظومة تقويم تحصيل المتعلمين المتطلبات والإجراءات التنفيذية مجلة كلية التربية، جامعة العريش، العدد السادس عشر أكتوبر 2018.
- محمد محمد فتح الله (2020 أ). جودة إعداد مفردات إنتاج الاستجابة الاختبارية (المقال) وموضوعية تقدير درجاتها في ضوء معايير جودة التقويم التربوي، مجلة القياس والتقويم، يصدرها مركز القياس بجامعة المنيا، العدد الأول يناير 2020. ص 1-28
- محمد محمد فتح الله (2020 ب). الدليل التنفيذي لمنظومة بنوك الأسئلة، مجلة القياس والتقويم، يصدرها مركز القياس بجامعة المنيا، العدد الأول يناير 2020. ص 125-146
- محمد منصور الشافعي (2014). تأثير انتهاك بعض افتراضات نظرية الاستجابة للمفردة على تدريج بنك الأسئلة ودقة معادلة الاختبارات البنكية المسحوبة، مجلة كلية التربية بنها، العدد (98) ص 431-498.
- محمد يحيى ناصف (2014). تصور مقترح لتطوير منظومة تقويم طلاب المرحلة الثانوية في ضوء خبرات بعض الدول، المركز القومي للبحوث التربوية والتنمية، القاهرة.
- المركز القومي للامتحانات والتقويم التربوي (2015). تطوير بنك أسئلة لقياس الاستعداد للدراسة بالمرحلة الثانوية، قسم البحوث، القاهرة.
- مروان عبد العزيز دبدوب (2007). المتغيرات الوهمية بديلة عن تقدير القيم المفقودة في تجارب التصميم الأساسية. مجلة علوم الرفادين، 18، 15-33.
- هار وهالت وغينكل وماركو زارستد (2019). الاساس في نمذجة المعادلات الهيكلية بالمربعات الصغرى الجزئية، ترجمة زكريا بلخامسة، عمان: مركز الكتاب الأكاديمي.
- وليد محمد أبو المعاطي 2017: رؤية مستقبلية مقترحة لتطوير مجال القياس والتقويم النفسي والتربوي أحد مجالات علم النفس التربوي في الجامعات المصرية، المجلة الدولية للدراسات التربوية والنفسية، المجلد (1)، العدد 2، ص ص 139-176.

Banks, K. (2015) "An Introduction to Missing Data in the Context of Differential Item Functioning," Practical Assessment, Research, and Evaluation: Vol. 20, Article 12. DOI: <https://doi.org/10.7275/fpg0-5079>

Brotherton, P.N.& Preece, F.W. (1995). Science Process Skills: The nature and interrelation ships. Research in

- Science & Technology Education, 13, 5-12.
- Camilli, G.& Shepard, L. (1994). Methods for Identifying Biased test items. California: Sage Publications.
- Davey, T. (2011). A guide to computer adaptive testing systems. Council of Chief State School Officers
- De Ayala. R. J., Plake, B., & Impira, J. (2001). The Impact of Omitted Responses On the Accuracy of Ability Estimation in Item Response Theory. Journal of Educational Measurement, 38, 213-234
- Fraley, R.C. & Others (2000). An Item Response Theory Analysis of Self- Report Measures of Adult Attachment. Journal of Personality, and Social Psychology, 78(2), 350: 365.
- Hair, J.F., Sarstedt, M., Ringle, C.M. and Gudergan, S.P. (2018). Advanced Issues in Partial Least Squares Structural Equation Modeling (PLS-SEM), Sage, Thousand Oaks, CA.
- Havens, A. (2002). Examinations and Learning: An Activity – Theoretical Analysis of the Relationship between Assessment and Learning.
- Retrieved October 20, 2016 from <http://www.leeds.ac.uk/educol/documents/00002238.html>
- King, L. (2008). A Comparison of the Performance of Testlet – based Computer Adaptive Tests and Multistage Tests. PhD Dissertation, Texas University.
- Kline, R. B. (2016). Principles and practice of structural equation modeling (4th ed.). New York, NY: Guilford Publications.
- Lin, S and Huang, Y. (2016). Psychometric properties and factorial structure of the Chinese version of the Gratitude Resentment and Appreciation Test. British Journal of Guidance & Counselling, 44(3), 347–361.

- Montesunos, M., & Benito, J. (2003). Test Purification and the Evaluation of Differential Item Functioning with Multinomial Logistic Regression. *European Journal of Psychological Assessment*, 19 (1), 1-11.
- Petter, S. (2018). "Haters gonna hate": PLS and information systems research. *The Data Base for Advances in Information Systems*, 49, 10–13.
- Peugh, J. L., & Enders, C. K. (2004). Missing Data in Educational Research A review of Reporting Practices and Suggestion for Improvement. *Review of Educational Research*, 74, 525-566.
- Popham, W. J. (1978). *Criterion-Referenced Measurement*. Englewood Cliffs N.J.: prentice Hall Inc.
- Ranger, J., & Kuhn, J.T. (2011). A flexible latent trait model for response times in tests. *Psychometrika*, 77, 31–47.
- Rovai A, Ponton M & Baker J (2008). *Distance Learning in Higher Education: A programmatic Approach to Planning, Design, Instruction, Evaluation, and Accreditation*. Teachers College Press.
- Schumacker, R.E., & Lomax, R.G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Erlbaum.
- Suen, H.K. (1990). *Principles of Test Theories*. New Jersey : Laurence Erlbaum Associates Publishers.
- TIAN Jian-quan, MIAO Dan-min, ZHU Xia, GONG Jing-jing (2007). An Introduction to the Computerized Adaptive Testing., *US-China Education Review*. 4(1): ,72- 81
- Wainer, H., Kaplan, B., Lewis, C. (1991). Some empirical guideline for building testlets, *Educational Testing Service*, available at : www.eric.ed.gov.
- Zahng, B., & Cindy, M., W. (2008). Impact of Missing Data on Person – Model Fit and Person Trait Estimation. *Applied Psychological Measurement*. 30(6). 466-479.

ملاحق الورقة البحثية

ملحق (1) قائمة المحكمين

الاسم	الوظيفة
أ.د/ محمود احمد عمر	أستاذ علم النفس التربوي كلية التربية جامعة عين شمس
أ.د/ امانى سعيدة	أستاذ علم النفس التربوي كلية الدراسات العليا للتربية جامعة القاهرة
أ.د/ كريمان عويضة منشار	أستاذ علم النفس التربوي كلية التربية جامعة بنها
أ.د/ احمد عاشور	أستاذ علم النفس التربوي كلية التربية جامعة بنها
أ.د/ إسماعيل الفقي	أستاذ علم النفس التربوي كلية التربية جامعة عين شمس
أ.د/ سيد صقر	أستاذ علم النفس التربوي كلية التربية جامعة كفر الشيخ
أ.د/ عادل البنا	أستاذ علم النفس التربوي كلية التربية جامعة دمنهور
أ.د/ نجيب خزام	أستاذ علم النفس التربوي كلية التربية جامعة عين شمس
أ.د/ كمال اسماعيل	أستاذ علم النفس التربوي كلية التربية جامعة بنها
أ.د/ محمد سعد محمد علي	أستاذ علم النفس التربوي كلية التربية جامعة الازهر
أ.د/ محمود عكاشة	أستاذ علم النفس التربوي كلية التربية جامعة دمنهور
أ.د/ أميمة كامل	أستاذ علم النفس التربوي كلية التربية النوعية جامعة القاهرة
أ.د/ محمود عبد الحليم منسي	أستاذ علم النفس التربوي كلية التربية جامعة الاسكندرية
أ.د/ أسماء عبد الحميد	أستاذ علم النفس التربوي كلية التربية جامعة المنيا
أ.د/ محمود احمد شبيب	أستاذ علم النفس التربوي كلية التربية جامعة فنا
أ.د/ أبو اليزيد الشويقي	أستاذ علم النفس التربوي كلية التربية جامعة طنطا
أ.د/ رمضان محمد رمضان	أستاذ علم النفس التربوي مدير المركز القومي للامتحانات والتقويم التربوي
أ.د/ هانم أبو الخير	أستاذ علم النفس التربوي كلية التربية جامعة المنصورة
أ.د/ احمد ثابت	أستاذ مساعد ورئيس قسم علم النفس التربوي كلية التربية جامعة مدينة السادات
أ.د/ نادية عبده عواض	أستاذ علم النفس التربوي كلية التربية جامعة الازهر
أ.د/ محمد غازي	أستاذ علم النفس التربوي المركز القومي للبحوث التربوية والتنمية

الاسم	الوظيفة
أ.د/ محمد يحيى ناصف	أستاذ علم النفس التربوي المركز القومي للبحوث التربوية والتنمية
أ.د/ محمد ابراهيم محمد	أستاذ مساعد علم النفس التربوي كلية التربية جامعة المنيا
أ.د/ اكرام حمزة	أستاذ مساعد علم النفس التربوي المركز القومي للامتحانات والتقويم التربوي
أ.د/ حلمي الفيل	أستاذ مساعد علم النفس التربوي كلية التربية النوعية جامعة الاسكندرية
د/ دعاء هريدي	مدرس علم النفس التربوي كلية الدراسات العليا للتربية جامعة القاهرة

ملحق رقم (2)

مقياس تقييم المفردة الاختبارية إنتاج الاستجابة
اسم المحكم: ... المادة: ... العام: ...
نص المفردة:

م	العبارة.	4	3	2	1	ادلة وشواهد
1-	ذات صياغة لغوية صحيحة ومناسبة					
2-	ذات صياغة علمية دقيقة وواضحة					
3-	تميز بين المتعلمين بحسب تحقيقهم لنواتج التعلم المستهدف قياسه					
4-	تراعي الأصول والقواعد الفنية لكتابة المفردات الاختبارية					
5-	تقيس إتقان المتعلم لنواتج التعلم المحدد					
6-	دقة وموضوعية تقدير الدرجات للإجابة عليها					
7-	تقيس موضوع مهم مرتبط بالمستويات المعيارية للمحتوى الدراسي					
8-	تنوافر في المفردة عنصر الجودة والحدثة (عدم تكرارها في اختبارات سابقة)					
9-	مناسبة الدرجة المقررة للمفردة لما هو					

				مطلوب من إنتاج المتعلم للاستجابة المحددة	
				10- مناسبة الزمن المخصص للإجابة لما تتطلبه المفردة من مهام محددة	
				11- تقيس قدرة المتعلم في التفكير الناقد وحل المشكلات	
				12- تعكس قدرة المتعلم على توظيف المعلومات المتعلمة في مواقف مختلفة جديدة	
				13- تتضمن المفردة معطيات محددة وواضحة للمتعلم	
				14- تتضمن المفردة مطلوب محدد واضح للمتعلم	
				15- يوجد مقياس متدرج للاستجابات المتوقعة للمتعلم على المفردة المقالية (إنتاج الاستجابة) لتقدير الدرجات	
				16- تتناسب درجة صعوبة المفردة ومستواها المعرفي مع ما وضعت لقياسه من المستويات المعيارية للمحتوى الدراسي	
				17- وضوح المعطيات والمطلوب الإجابة عنه في السؤال.	
				18- مدى استخدام السؤال لكلمات عامة غير محددة مثل (ناقش - اشرح - اكتب نبذة مختصرة).	
				19- مدى ملائمة نوع السؤال مع ناتج التعلم المستهدف قياسه.	
				20- مدى تناول السؤال لمشكلة واضحة ومحددة	