مجلة آفاق لعلم الاجتماع

ISSN: 1112-8259 EISSN: 2600-6855

المجلد 11 العدد 2 / ديسمبر 2021 ص ص 44- 58

صلاحية اختبارات الذكاء للتطبيق في بيئات مختلفة في ظل النسبية الثقافية

Assessing the suitability of applying intelligence tests in different environments in the light of cultural relativity

د. بوطالية يمينة*
مخبر القياس والدراسات النفسية
جامعة البليدة2 علي لونيسي

تاريخ الإرسال: 28/ 08/ 2021 تاريخ التقييم: 2021/08/29

تاريخ القبول: 2021/09/10

الملخص

هدفت هذه الدراسة إلى معرفة واقع تطبيق اختبارات الذكاء غير المكيفة في البيئة العربية ومعرفة مدى موضوعية وسلامة الإجراءات التي يقوم بها أصحاب هذه الدراسات من خلال إعادة تحققهم من صدق وثبات الاختبارات.

اعتمدت الدراسة على المنهج الوصفي الذي يهتم بوصف واقع الحال كما هو عن طريق البعد الكيفي أو الكمي معا، وشملت 25 أطروحة دكتوراه و8 دراسات.

توصل هذا البحث إلى أن الدراسات توصلت إلى قبول مستويات صدق وثبات هذه الاختبارات من حيث محتوى بنودها ومعايير تفسير نتائجها، وهذه مفارقة علمية تتضمن خللا نابعا من قصور في طرق إعادة التحقق من صدق وثبات هذه الاختبارات.

الكلمات المفتاحية: اختبارات الذكاء، صلاحية، ثبات، نسبية ثقافية، إعادة التحقق.

Abstract

This study aimed to know the reality of the application of unconditioned intelligence tests in the Arab environment and to know the extent of objectivity and safety of the procedures carried out by the owners of these studies by re-verifying the validity and reliability of the tests.

The study relied on the descriptive approach, which is concerned with describing the reality of the situation as it is through the qualitative or quantitative dimension together, and it included 25 doctoral theses and 8 studies.

The findings revealed that the studies came to a decision to accept the values of validity regarding the items content and of test standards.

Keywords: Intelligence tests, validity, stability, cultural relativity, re-verification.

*بوطالية يمينة، y.boutalia@univ-blida2.dz

1- مقدمة

من خلال تتبع المسار التاريخي لتطور حركة القياس النفسي وبناء الاختبارات يتضح أن الاختبارات المتداولة سواء في الممارسة السيكولوجية الميدانية أو في الدراسات الأكاديمية هي مقاييس غربية المنشأ في أصلها، سواء كفكرة أو كبناء واستعمال، مما جعلها محملة بالأبعاد الثقافية للمجتمعات الغربية ودفع الباحثين في هذه البلدان إلى عدم تطبيقها كما هي بل اللجوء إلى إعادة التحقق من صدقها وثباتها، ومنه جاءت دراستنا بهدف تحليل عدد من أطروحات الدكتوراه والبحوث المنشورة في مجلات علمية متخصصة، نظرا لأن اكتفاء الباحثون بالخصائص السيكومترية الواردة في الاختبارات وفق بيئاتها الأصلية مع العمل على إعادة تقدير مستويات صدق تبين جلها صلاحية هذه الاختبارات دون إجراء التعديلات التي تقتضيها عمليات التقنين لتتناسب والبيئة أو الثقافة الجديدة، ودون الأخذ بعين الاعتبار طبيعة السمات المراد قياسها وظروف تطبيق الاختبار وما يناسبها من دلالات صلاحية. ليتم البحث عن مدى توافق استخدام الدلالات في البيئتين الجديدة والأصلية) الأمر الذي يستلزم التأمل من طرف الباحثين قبل الثقة في نتائج صدق الاختبارات وتوظيفها لمختلف الأغراض البحثية والميدانية).

من هنا تتجلى أكثر إشكالية الدراسة الحالية، فمن جهة ينطلق الباحثون من افتراض إبستيمولوجي نابع من تجليات النسبية الثقافية مؤداه أن الاختبار النفسي الصالح في بيئة معينة لها مميزات وبنية خاصة قد لا يكون صالحا في بيئة مختلفة عنها جذريا، فالاختبار النفسي مشبع بالثقافة التي تعتبر عاملا حاسما في فهم السلوك الفردي وتفسيره، وهذا ما ذهبت إليه المدرسة البنائية الاجتماعية مع بيتر بيرغر(Berger 1996) والتي أشارت إلى أن أنظمة المعرفة التي تقيسها الاختبارات ما هي إلا تركيبات ذهنية تسهم فيها مؤثرات تفرضها القيم الاجتماعية والدينية والايديولوجية السائدة في مجتمع معين، مما يعني أن التفسير الموضوعي للمواقف السلوكية يحتاج إلى فهم عميق للسياق الاجتماعي والعقائدي الذي نشأت فيه والمعاني التي تتشكل في مخيلة الأفراد والضمير الجمعي الذي تشكل على مدار الزمن بفعل مختلف العوامل الثقافية التي تميز المجتمعات عن بعضها البعض، في هذا السياق النسبي الثقافي فمنا بالاطلاع على عينة من نتائج البحوث حول اختبارات الذكاء والمتعلقة بإعادة التحقق من صدق الاختبارات إلى أن الباحثين يعتمدون على تقديرات كمية كأدلة على أن مستويات صدق اختباراتهم على درجات عالية تسمح لهم بقبولها دون تعديلات. وهو الأمر الذي يبطن تنافرا معرفيا عميقا دفعنا إلى محاولة الاقتراب من دلالات الصدق المعتمد عليها في عينة من أطروحات الدكتوراه ودراسات منشورة في مجلات علمية محكمة والتي وظفت للتدليل على النقائص المسجلة في عملية التحقق من صدق الاختبارات، وبشكل أكثر تحديدًا فإن الدراسة تحاول الإجابة عن التساؤلات التالية:

- ماهي الإجراءات العملية التي يتبعها الباحثون لإعادة التحقق من صدق اختبارات الذكاء التي يطبقونها في دراساتهم -أطروحات الدكتوراه ودراسات منشورة في مجلات علمية؟

- هل هناك اختلاف في أنواع الصدق المستخدمة من طرف الباحثين في إعادة التحقق من صلاحية الاختبارات التي يطبقونها في دراساتهم؟

- هل يلتزم الباحثون بتقديم تبريرات موضوعية كدواعي لاستخدامهم أنواع معينة من الصدق لإعادة التحقق من صلاحية الاختبارات النفسية المنقولة من ثقافات أخرى والمطبقة في دراساتهم؟

- هل يحترم الباحثون الإجراءات العلمية الموضوعية لإعادة التحقق من صدق الاختبارات التي يطبقونها، أم أنهم يكتفون بإجراءات إحصائية تتجلى في شكل أرقام لا تعبر عن صلاحية الاختبار؟

وتتحدد الدراسة الحالية بجانبين الأول يتناول صلاحية الاختبارات النفسية وحدود هذه الصلاحية مع تحديد الدواعي التي تدفع الباحثين إلى التعامل مع الاختبارات الجاهزة من زاوية الشك في صلاحيتها واللجوء إلى استخلاص أدلة جديدة نابعة من خصائص العينات المطبق عليها، وندمج في الجانب الثاني نتائج الدراسة المتعلقة بتتبع طريقة تعامل الباحثين مع الاختبارات الجاهزة من خلال عدد من أطروحات الدكتوراه وبحوث منشورة في مجلات عربية متخصصة، بهدف تحديد جوانب التوفيق وجوانب القصور في هذه الإجراءات لنبين الدواعي التي جعلت عمليات إعادة التحقق من الصدق تتوصل إلى نتائج تناقض المنطلق النظري للباحثين وهو تأثر اختبارات الذكاء بالعوامل الثقافية.

2- أهمية الدراسة والحاجة إليها

تنبع أهمية الدراسة الحالية من أهمية وخطورة القرارات التي تتخذ من طرف الباحثون في مختلف المستويات والتي تعتمد على اختبارات ومقاييس نفسية مبنية في ثقافة غربية أكدت البحوث عبر الثقافية ذات الطابع المقارن في علم النفس وجود عدة مشكلات تنشأ من جراء التباينات الثقافية.

فكل ثقافة تتميز بمعايير وقيم ومفاهيم تختلف عن الثقافات الأخرى، أولى هذه المعايير اللغة من حيث قدرتها عن التعبير عن مختلف المعاني الواردة في اللغة الأصل للاختبار، فنقل عبارة من لغة إلى أخرى عادة ما يطرح مشكل المقابل في اللغة المقصودة والذي ينقل بأمانة المعنى المقصود، بمقابل ذلك نرى الباحثين في العالم العربي بصفة عامة وفي الجزائر بصفة خاصة يتعمدون على هذه الاختبارات وهم يدركون هذه الحقيقة العلمية فيقومون بإعادة التحقق من صلاحيتها من دون منظور حيوي نظري واضح المعالم فتصبح الإجراءات شكلية بحتة، إلى الحد الذي جعل بعض من الباحثين يلجأ إلى المبرمجين والمتمكنين من تطبيق بعض البرامج الإحصائية الالكترونية لكي يجري لهم تحليلا إحصائيا للبيانات المستمدة من تطبيقهم للاختبارات وهم يجهلون تماما المقاربات والخلفيات النظرية الكامنة وراء الأرقام التي يتوصلون إليها والهدف منها.

كما أن الكثير من الباحثين يفتقرون إلى تصورات نظرية واضحة ومحددة الأبعاد يسعون إلى قياس السمة في إطارها(رجاء محمود، 2001، ص147) ويتزامن ذلك مع غياب فروض واضحة يتم التحقق منها، فاختيار نوع من أنواع الصدق كثيرا ما تؤدي إلى نتائج لا تحمل معنى نظري وليست لها قيمة علمية، فالوصول إلى نتائج ذات قيمة علمية يتطلب التعمق في دراسة المفاهيم نظريا أولا والاطلاع على أكبر عدد من الدراسات والاختبارات المشابهة ثانيا، قبل الخوض في غمار إعادة التحقق من الصلاحية الذي يعتبر عملا متخصصا نظريا وتقنيا، فيجد الباحث نفسه أمام إطار نظري مستمد من النماذج الغربية وبعدا تطبيقيا متحايل عليه من خلال عينة جديدة من الأفراد تختلف كلية عن العينة التي بني الاختبار وقنن عليها (بوسالم، 2014، ص89).

أمام هذه المشكلات المنهجية والمنطقية جاءت هذه الدراسة التحليلية بهدف توضيح الاختلال الحاصل من طرف الباحثين في إعادة التحقق من صدق اختباراتهم، نظرا لإتباعهم إجراءات تؤدي في غالب الحالات إلى اختبارات غير صالحة وسيئة البناء وتصبح نتائجها ليست ذات مصداقية سواء في البحث العلمي الأكاديمي أو في الخدمة النفسية الاجتماعية المراد تقديمها للأفراد في مختلف مجالات الحياة حيث يتواجد المختص النفسي.

3- أهداف الدراسة

تهدف هذه الدراسة إلى توجيه التركيز البحث في مجال الاختبارات النفسية نحو الاعتناء بالخصائص الكيفية للاختبارات كقرينة ملازمة لدلالة الأرقام، بدل التركيز على البعد الرياضي الكمي كمؤشر وحيد فقط لصدق وثبات الاختبارات في البيئة الجديدة ومنه تطبيق الاختبار وتفسير نتائجه. كما هدفت الدراسة الى تقصي السبب الكامن وراء توصل جل الباحثين إلى أن هذه الاختبارات والمقاييس صالحة في البيئة الجزائرية والعربية، بعد نقلها من البيئة الغربية وبعد تطبيقها تجريبيا وإعادة استخراج مؤشرات صدقها وثباتها. ومنه حاولنا تقديم تصور شامل موضوعي يرتكز على المفهوم الحديث لصلاحية الاختبارات قصد تقديم منهجية لإعادة التحقق من صدق وثبات الاختبارات المراد تطبيقها والمنقولة من بيئات غربية.

4- ماهية العلاقة بين الاختبار النفسي والنظرية النفسية

متعارف عليه بأن العلوم الاجتماعية (علم النفس وعلوم التربية جزء منها) يمكن اعتبارها نتاج خمسة عناصر أساسية، العنصر الأول يتمثل في وجود فلسفة اجتماعية ومفاهيم أنطولوجية (Ontologie) أما الثاني فيتمثل في الأدوات المنهجية التي توضع من أجل التقرب من الواقع لمحاولة دراسته وفهمه وتفسيره بطريقة علمية، أما الثالث فيشير إلى المعلومات البنيوية أو البنى الاجتماعية الموضوعية الموجودة في الواقع. والرابع يشير إلى معطيات إدراكية وتجريدية للفرد والمجتمع الذي يعيش فيه (Perception) وخامسا مجموع الخدمات الاجتماعية والنفسية والتربوية التي يمكن تقديمها عن طريق توظيف النظريات والإفادة منها واقعيا، وبدون توفر هذه العناصر الخمس لا يمكن لأي علم من العلوم الاجتماعية أن يكون له جدوى وقيمة أكاديمية أو مجتمعية. (عبد السلام، 2002، ص 148).

هذه المكونات تحلينا إلى حقيقة علمية أخرى وهي أن دراسة الوقائع الاجتماعية والنفسية لابد أن يتم في إطار تصور نظري متفق حوله نسبيا يسمح بتحويل الوقائع الامبريقية إلى استدلالات علمية. وعلم النفس كبقية العلوم يرتبط بعمليات تاريخية ويتأثر بالمجتمع الذي أنتجه، فالظاهرة السيكولوجية هي عملية تعبير للنسبية الثقافية والاجتماعية والذي يختص بسياق دون أخر، كما أن النظرية النفسية هي نتاج إطار اجتماعي يعبر عن مستوى النمو الصناعي والتجاري وتطور المعارف ونسبة التعليم والديمقراطية والحريات والحقوق الفردية والمعتقدات، هذه الخصائص تختلف من مجتمع إلى أخر وفي ظلها أنتج علم النفس الغربي في الولايات المتحدة بالدرجة الأولى (علم الرغم من أن بوادره الأولى كانت في أوروبا) فهو علم أنجلو- ساكسوني تم التفكير فيه وإنتاجه باللغة الانجليزية وهو يأخذ في حسبانه مشكلات هذه المجتمعات مركزا على تصوراتها للإنسان وللقيم وللخير والشر ومختلف المعايير الاجتماعية(Ardila, 1995,p34)، وبنظرة تاريخية نجد أنفسنا في الجزائر والعالم العربي نستورد المعرفة السيكولوجية الغربية من أجل تطبيقها في بلادنا لدراسة وفهم سلوك الفرد الجزائري والعربي.

إذن النظريات النفسية مستمدة من المجتمعات الغربية وتعبر عن واقعها وعن نظرتها للإنسان. والاختبارات النفسية كأدوات منهجية للاقتراب من الواقع ودراسته بطريقة موضوعية تعتبر مشبعة بروح النظريات وتجسيد لها، وبالتالي فعند بناء اختبار نفسي فإن السمة التي يقيسها هذا الاختبار تتموقع في سياق نظري يختاره الباحث من ضمن السياقات النظرية المتداولة والتي يرى أنه مناسب لدراسة وتفسير السمة المراد قياسها، فيختار المؤشرات وفقها ويقدم التفسيرات الموضوعية بناء عليها.

هل يمكن نقل هذه الاختبارات من بيئة ثقافية ذات خصوصية معينة وتطبيقها في بيئة أخرى واعتبارها صالحة أم أن الأمر يحتاج إلى إعادة فحص هذه الاختبارات للتحقق من صلاحيتها، فإذا كان الجواب المنطقي هو الحاجة إلى إعادة فحص، فماهي الإجراءات المنهجية والحدود الابستيمولوجية التي يتحرك في إطارها الباحث وهو يحاول الاقتراب من اختبار بني في الغرب ويريد تطبيقه في الجزائر أو في إحدى بلدان العالم العربي.

وهل يمكننا الحديث عن توطين النظريات النفسية والتربوية قبل توطين الأدوات التي تعكس هذه النظريات، أم يمكننا لدواعي نفعية محضة أن نوطن الأدوات ونبقى على النظريات العلمية غربية، في هذه الحالة هل يمكننا نحت مصطلح "التبيئة" لنصف من خلاله العمليات والإجراءات المنهجية والقياسية التي تسمح لنا بجعل الاختبارات الغربية صالحة للتطبيق في البيئة الجزائرية أو العربية بصفة عامة.

5- من مفهوم صدق الاختبارات النفسية إلى مفهوم صلاحيتها

يعتبر الصدق الصفة الأهم في أي أداة من أدوات القياس في كل العلوم السلوكية عامة وبدونه لا يمكن الاعتماد على نتائجها في اتخاذ أي قرار من القرارات الفردية أو الجماعية.

لقد شاع تعريفا في مختلف المراجع والدراسات النفسية والتربوية يشير إلى أن معنى الصدق هو أن يقيس الاختبار فعلا القدرة أو السمة أو الاستعداد الذي وضع لقياسه، أي أن يقيس فعلا ما وضع لقياسه. فالاختبار الذي أعد لقياس سمة نفسية (الدافعية مثلا) يكون مقياسا صادقا بمدى قدرته على قياس سمة الدافعية فعلا(عبد السلام، 1996، ص 211). ونتيجة لذلك سيطر هذا المعنى على أدبيات القياس وبناء الاختبارات النفسية من كتب ودراسات أكاديمية (بحوث الدكتوراه والماجستير الدراسات المنشورة في المجلات العلمية المحكمة) هذه السيطرة استمرت في الكثير من البلدان وخاصة العربية منها على الرغم من التطور الذي عرفته نظريات القياس.

فالتعمق في دراسة هذا التعريف وربطه بمفهوم الصدق وفق النظرية الحديثة في القياس النفسي يبين أن التعريف ينطوي على نقص لأنه يشير الى بعد واحد من أبعاد الصدق، بمعنى أن الاختبار يمكن أن يقيس السمة المراد قياسها ولكنه لا يتوفر على درجات مقبولة من القدرة على التنبؤ بسلوك الأفراد في المستقبل، وعليه اقترحت الجمعية الأمريكية لعلم النفس (APA) إدخال بعد أخر للتعريف كمرحلة تاريخية أولى وهوقدرة الاختبار على تحقيق الغرض أو الهدف الذي وضع من أجله، بمعنى إدخال بعد التنبؤ كواحد من أهم أغراض عملية القياس في العلوم السلوكية. أي قدرة الاختبار على التنبؤ بأداء الأفراد في المستقبل، كأن نطبق اختبارا لقياس القدرة الميكانيكية ولكن الغرض من تطبيقه هو التنبؤ بالفرد الذي يمكن أن ينجح مستقبلا في تخصص جامعي معين. فقد يكون الاختبار صادقا في قياس القدرة الميكانيكية ولكنه غير قادرا على التنبؤ بأداء الفرد مستقبلا.

6-عينة الدراسة

نشير في البداية إلى أن اختيار دراسة معينة لتكون من ضمن العينة خضع لمعيار أن تكون إما أطروحة دكتوراه أو بحثا منشورا في مجلة علمية محكمة في إحدى التخصصات الثلاثة (علم النفس، علوم التربية، الارطوفونيا)، وأن يكون الاختبار المطبق في الدراسة هو اختبار ذكاء، وهذا أهم سبب جعل العينة تبدو قليلة على اعتبار أن دراسات الذكاء قليلة بالمقارنة مع الدراسات التي تناولت أبعاد مختلفة من الشخصية.

كما أن هذه الدراسات تنتمي إلى جامعات جزائرية مختلفة، بمعنى أن متغير الجامعة ليس له قيمة في هذه الدراسة بل أكتفينا بالبحث في مكتبات والمواقع الالكترونية لبعض الجامعات التي تضمن مناقشة الدكتوراه في الشعب الثلاث المعنية بالدراسة، وعليه لم نوزع الدراسات على الجامعات لأن بعض الجامعات أخذنا منها أطروحة واحدة والبعض 06 أطروحات وهكذا حتى تمكنا من الحصول على 25 أطروحة طبق فيها اختبار ذكاء ليس من بناء الباحث صاحب الدراسة.

أما بخصوص الدراسات المنشورة في مجلات علمية محكمة فقد اعتمدنا على الأرضية الوطنية للمجلات العلمية (https://www.asjp.cerist.dz) وهي متوفرة على شبكة الانترنيت ويمكن الوصول إليها بسهولة والبحث عن الدراسات التي تناولت الذكاء ومنه اختيار الدراسات التي اعتمدت اختبارات جاهزة، وعليه لم نتحكم في حجم العينة بل توفر هذا النوع من الدراسات هو الذي أوصلنا الى 33 دراسة نراها كافية لتحقيق أهداف الدراسة.

تتمثل عينة الدراسة الحالية في بحوث علمية في علم النفس، علوم التربية والارطوفونيا بمختلف تخصصاتها، وقد أعتمد أصحاب هذه الدراسات لقياس الذكاء كأحد متغيرات الدراسة على اختبارات أو مقاييس لباحثين آخرين بنيت في بيئات أجنبية (بيئات غير عربية). وقد بلغ عدد الدراسات 33 دراسة موزعة على 25 أطروحة دكتوراه بغض النظر عن كونها من نمط دكتوراه الطور الثالث (ل.م.د) أو دكتوراه علوم، وثمانية (8) دراسات منشورة في مجلات علمية محكمة. والجدول التالي يبين توزيع الدراسات عينة البحث حسب التخصص وطبيعة البحث من حيث كونها دكتوراه أو بحث منشور.

الجدول رقم 1: تكرار الدراسات عينة البحث حسب التخصص

طبيعة البحث	أطروحة دكتوراه		بحث منشور		المجموع	
التخصص	التكرار	٪	التكرار	٪	التكرار	٪
علم النفس	08	32	03	37.50	11	33.34
علوم التربية	13	52	03	37.50	16	48.48
الأرطوفونيا	04	16	02	25	06	18.18
المجموع	25	100	08	100	33	100

من خلال الجدول رقم 01 يتضح أن العدد الكلي للدراسات هو 33 دراسة موزعة على الشعب الثلاث، بحيث علم النفس بـ 11 بحثا منها 08 بحوث لأطروحة دكتوراه وثلاثة لبحوث منشورة في مجلات علمية محكمة، بالمقابل هناك 16 بحثا في شعبة علوم التربية منها 13 لأطروحة دكتوراه و03 لبحوث منشورة في مجلات، وأخيرا شعبة الأرطوفونيا بـ 06 دراسات منها 04 لأطروحة دكتوراه و02 لبحوث منشورة في مجلات علمية محكمة.

6- منهج الدراسة

اعتمدت الدراسة على المنهج الوصفي الذي يهتم بوصف واقع الحال كما هو عن طريق البعد الكيفي أو الكمي أو البعدين معا. والدراسة الحالية اعتمدت البعدين معا، فهي حاولت وصف النتائج الكمية التي تتعلق بالخصائص السيكومترية للاختبارات والمقاييس التي طبقت من طرف أصحاب الدراسات المعنية بالبحث بهدف معرفة مدى صلاحية هذه الاختبارات وفق البعد الكمي الإحصائي. كما اهتمت الدراسة أيضا بالبعد الكيفي في المنهج الوصفي من خلال وصف مختلف الإجراءات المنهجية والخطوات التي اتبعها أصحاب الدراسات من أجل إعادة التحقق من مستويات صدق وثبات الاختبارات التي طبقت وهي اختبارات بنيت في بيئات غير عربية مما تطلب من الباحثين إعادة التحقق من مدى صلاحيتها للتطبيق في البيئة الجزائرية.

7- نتائج الدراسة

قبل عرض النتائج التي توصلت اليها الدراسة والخاصة بإعادة التحقق من صلاحية اختبارات الذكاء من طرف مجموعة من الباحثين، نشير إلى أن تعاملنا مع مؤشري صلاحية الاختبارات النفسية وهما الصدق والثبات، كان بوصف الصدق على أنه أنواع كما ورد في تصنيف الجمعية الأمريكية لعلم النفس، في حين تم وصف الثبات من حيث أنه طرق، فالثبات واحد ويتم التحقق منه بعدة طرق، أي قولنا أنواع الصدق مقابل طرق الثبات، وكلاهما مؤشرا أو دليلا من أدلة الصلاحية.

7-1- نتائج إعادة التحقق من صدق الاختبارات

كان لظهور نظرية الاستجابة لمفردات الاختبار دور بارزا في تطور النظرة إلى مختلف الخصائص السيكومترية للاختبارات النفسية وبالدرجة الأولى صفتي الثبات والصدق ويشير مفهوم الصدق وفق النظرية الحديثة للقياس إلى الاستدلالات الخاصة التي تخرج بها درجات الاختبار من حيث مناسبتها ومعناها وفائدتها. وتحقيق صدق الاختبار معناه تجميع عدد من الأدلة التي تؤيد هذه الاستدلالات، ولذلك يشير الصدق إلى مدى صلاحية استخدام درجات الاختبار في القيام بتفسيرات معينة، فإذا كان الاختبار يستخدم لوصف الذكاء فيجب أن تفسر الدرجات أو النتائج على أساس أنها تمثل سمة الذكاء الذي يقيسه الاختبار.

وقصد التحقق من مدى التزام الباحثين بالمفاهيم الجديدة للصدق أثناء إعادة التحقق من مستويات صدق اختبارات الذكاء المطبقة في دراساتهم ويجدونها جاهزة ومبينة في بيئات غربية إما ترجمت إلى اللغة العربية وأعيد التحقق من صلاحيتها من طرف باحثين آخرين، وإما قيام البحث في حد ذاته بترجمة الاختبار من اللغة الأم (لغة انجليزية في حالة اختبارين في دراستنا)، وبالتالي يتبع ترجمة الاختبار قيام الباحث بإجراءات التحقق من صدق وثبات الاختبار، وفي الجدول الموالي نعرض نتائج تحليل هذه الدراسة بخصوص أنواع الصدق المستعملة من طرف الباحثين في الدراسة.

الجدول رقم 2: أنواع الصدق المستعملة من طرف الباحثين في عينة الدراسة وطبيعة القرارات بقبول أو رفض مستوى صدقها

القرار بخصوص الصدق	مستوى صدق مقبول		مستوى صدق غير مقبول		المجموع	
نوع الصدق	التكرار	%	التكرار	%	التكرار	%
صدق المحكمين	18	29.50	00	00	18	29.50
صدق المحتوى	15	24.59	01	01.69	16	26.22
الصدق التمييزي	12	19.67	00	00	12	19.67
صدق المحكات	08	13.11	01	01.69	09	14.75
صدق المفهوم	06	09.83	00	00	06	09.83
المجموع	59	96.17	02	03.83	61	100

في البداية نشير إلى أن أنواع الصدق الموضحة في الجدول أعلاه أكبر من عدد الدراسات التي تم تقييمها، لكون الباحثين عادة ما يعتمدوا على أكثر من طريقة واحدة لإعادة التحقق من الصدق وأكثر من طريقة لتقدير ثبات الاختبار الواحد، كما أن بعض التسميات المذكورة في الجدول قد تثير انتباه بعض الباحثين على أنها تسميات غير علمية ولم تذكرها التصنيفات المختلفة للجمعية الأمريكية لعلم النفس والتي تعتبر مرجعية تعريف أنواع الصدق بين الباحثين عالميا.

وعليه فإن الموجود في الجدول هو ذكر نوع الصدق كما أوردها صاحب الدراسة وبدون أي تأويل من طرفنا. بالرغم من اعتراضنا على بعض التسميات الواردة في الجدول أعلاه.

وبالعودة إلى النتائج نلاحظ من خلال الجدول الذي يبين أنواع الصدق التي أعتمد عليها الباحثون في دراساتهم -عينة البحث- البالغة 25 أطروحة دكتوراه و8 دراسات في مجلات علمية محكمة، نلاحظ أن ما سماه الباحثون بصدق المحكمين هو النوع الأكثر -استعمالا بـ 18 مرة أي بنسبة 29.5% من مجموع أنواع الصدق المستعملة في الدراسات وقد أسفر على قبول مستوى صدق الاختبارات من طرف الخبراء المحكمين بنسبة 100%، ويأتي في الرتبة الثانية صدق المحتوى والذي هو في الحقيقة إشارة إلى النوع الأول، والذي يعبر معناه عن مدى تمثيل بنود الاختبار لمؤشرات السمة المراد قياسها، وجاء هذا النوع من الصدق بتكرار 16 مرة ونسبة مئوية تقدر بـ26.22%، منها 15 حالة كان القرار فيها بقبول مستواه وحالة واحدة فقط كان القرار فيها بعدم قبول مستواه، أي أن الباحث توصل من خلال الإجراءات التي قام بها إلى أن مستوى صدق الاختبار دون المستوى المطلوب، أي أن إعادة التحقق من اختبار سبق وأن تحقق من صلاحيته باحثون آخرون أكدت ان درجة صدقه لا ترقى الحد المقبول.

ويأتي في الرتبة الثالثة ما سمي من طرف الباحثون بالصدق التمييزي بتكرار 12 مرة ونسبة 19.67% وقبل مستوى صدق الاختبارات في جميع الحالات التي وظف فيها الباحثون هذا النوع من الصدق، وفي الرتبة الرابعة نجد صدق المحكات بتكرار 09 مرات ونسبة 14.75% منها 08 حالات قبل فيها مستوى الصدق وحالة واحدة لم يقبل مستواه. ويأتي في الرتبة الأخيرة من حيث الاستعمال صدق المفهوم بتكرار 06 حالات ونسبة 09.87%.

2-7- نتائج إعادة التحقق من ثبات الاختبارات

كما هو الشأن بالنسبة لأنواع الصدق قمنا في هذه الدراسة برصد طرق الثبات التي أستعملها الباحثون من أجل إعادة التحقق من مستويات ثبات الاختبارات المطبقة في دراساتهم، مع تحديد طبيعة القرارات التي توصل إليها الباحثون بقبول مستوى الثبات أو رفضه، والجدول الموالي يلخص هذه النتائج.

الجدول رقم 3: طرق التحقق من الثبات المستعملة من طرف الباحثين في عينة الدراسة وطبيعة القرارات بقبول أو رفض مستواه

القرار بخصوص الثبات	مستوى ثبات مقبول		مستوى ثبات غير مقبول		المجموع	
طريقة الحساب	التكرار	٪	التكرار	٪	التكرار	٪
تحليل التباين	28	44.44	01	00.74	29	46.03
التجزئة النصفية	17	26.98	03	04.76	20	31.74
إعادة التطبيق	12	19.06	00	00	12	19.06
الصور المتكافئة	02	03.17	00	00	02	03.17
المجموع	59	93.63	04	05.50	63	100

كما حصل حين التحقق من أنواع صدق الاختبارات حيث كان تكرار الأنواع أكبر من عدد الدراسات بسبب أن الباحث يعتمد على أكثر من نوع مع الاختبار الواحد، الشيء نفسه حصل مع طرق التحقق من الثبات حيث نجد مجموع الطرق أكبر من عدد الدراسات والبالغ 33 دراسة، لكون الباحث عادة ما يعتمد على أكثر من طريقة لإعادة تقدير ثبات الاختبار الواحد.

من خلال النتائج المبينة في الجدول نلاحظ والخاصة بطرق إعادة التحقق من ثبات اختبارات الذكاء قبل تطبيقها في الدراسات عينة البحث والبالغ عددها 33 دراسة، نلاحظ أن الطريقة الأكثر استعمالا هي تحليل التباين بمجموع 29 مرة، منها 28 أي ما نسبته 44.44% توصل فيها الباحثون إلى قبول قيم الثبات، في حين مرة واحدة لم يقبل فيها قيم الثبات، ونشير إلى أن طريقة تحليل التباين تعتمد على قيمة إحصائية واحدة ناتجة عن تطبيق أحد المعادلتين، إما ألفا (α) في حالة الاستجابات المتعددة أو معادلة كيودر وريتشار دسون (Kuder-Richardson) في حالة الاستجابة الثنائية (0-1) في حالة اختبارات التحصيل والذكاء، وهو الحال مع اختبارات الذكاء عينة هذه الدراسة.

أما الطريقة الثانية التي اعتمدها الباحثون فهي طريقة التجزئة النصفية بـ 20 مرة أي بنسبة 31.74%، منها 17 مرة تم قبول معامل ثبات الاختبارات أي بنسبة 26.98% في حين توصلت ثلاث دراسات إلى قيم مستويات ثبات منخفضة وهو ما يمثل 04.76 %، وهنا نشير إلى أن طريقة التجزئة النصفية في التحقق من الثبات تعتبر صالحة مع الاختبارات التي تكون فيها السمة المقاسة أحادية البعد وليست متعددة الأبعاد.

أما الطريقة الثالثة التي أعتمدها الباحثون من أجل إعادة التحقق من الثبات فهي طريقة التطبيق وإعادة التطبيق، بمعنى التحقق من ثبات الاستقرار والذي يشترط ألا تكون السمة المقاسة حساسة للزمن، بمعنى أنها سريعة التغير مع الزمن، والذكاء سمة من هذا النمط وبالتالي تطبيق الاختبار على عينة من الأفراد وإعادة تطبيقه بعد فترة زمنية محددة (من 15 يوما إلى شهر) لن يؤثر على مستويات السمة عند الأفراد، وعليه فقد اعتمدت 12 مرة بنسبة 19.06% وقد أسفرت النتائج عن قبول قيم الثبات في جميع الحالات، وهذا يوافق المنطق القياس الذي يشير الى أن الاختبار اذا صادقا وطبق وأعيد تطبيقه فإنه يعطينا نتائج متقاربة تشير إلى ارتفاع مستويات ثباته بشرط احترام الباحث لظروف التطبيق في المرتين.

أما الطريقة الرابعة والأخيرة فهي طريقة الصور المتكافئة والتي تعتبر أقل الطرق استعمالا من طرف الباحثين في التحقق من ثبات الاختبارات والمقاييس النفسية سواء أثناء البناء أو أثناء إعادة التحقق من الثبات للاختبارات الجاهزة، وعليه فقد لجأ إليها الباحثون في عينة الدراسة مرتين (02) مما يمثل نسبة 3.17% وهي اقل نسبة لطرق التحقق من الثبات في البحوث عينة الدراسة.

7-3- تحليل وتفسير النتائج في ضوء المفهوم الحديث لصلاحية الاختبارات النفسية

بالرغم من أن منطلق الباحثين عند إعادة التحقق من صدق وثبات الاختبارات هو افتراض أن هذه الاختبارات على اختلاف طبيعتها فهي بنيت في بيئات غربية في بعض الحالات في بيئة عربية ونقلت الى بلد عربي أخرفي حالات أخرى، وعليه لجأ الباحثون إلى محاولة الإجابة عن التساؤل الذي يطرح عادة عند الشروع في تطبيق اختبار أو مقياس نفسي جاهز بني في بيئة غير بيئة العينة والتساؤل هو: هل هذه الاختبارات تتمتع بمستويات مقبولة من الصلاحية عندما تطبق على عينات جزائرية والأخذ بالنتائج وتفسيرها علميا لفهم السلوك بطريقة موضوعية.

من خلال هذه النتائج التي تشير إلى اعتماد الباحثين على أنواع متعددة من الصدق وطرق للثبات للتحقق من صلاحية اختباراتهم، نقول أن صدق الاختبار النفسي أصبح ينظر إليه حسب (علام صلاح الدين، 2002، ص104) على أنه مفهوما واحدا وخاصية واحدة للاختبار، وما أصطلح على تسميته بأنواع الصدق في بعض المراجع والاستعمالات المختلفة هي في الحقيقة إشارة إلى مؤشرات وطرائق لجمع الأدلة عن مدى صدق الاختبار وليست نوع من أنواعه (Jenkins, 1978,p 93). لذلك كلما كان الاختبار يحمل أكثر من مؤشر للصدق زادت ثقة الباحث فيه لقياس ما أعد لقياسه وتحقيقه للغرض الذي وضع من أجله، كما أصبح البعد القيمي الاجتماعي لصلاحية الاختبارات والمقاييس النفسية يشكل جزءا لا يتجزأ من مفهوم الصدق. ولذلك أصبح الحديث عن بعد جديد للصدق (دليل أو مؤشر جديد) يتمثل في المآل أو الغرض الاجتماعي لعملية القياس (علام صلاح الدين، 1995، ص220)، كما أن السعي الى تحقيق صدق الاختبار هو في الحقيقة عملية تجميع للأدلة التي نستدل بها على قدرة الاختبار على قياس ما اعد لقياسه، وقد حددت ثلاثة مؤشرات رئيسة لصدق المقاييس والاختبارات النفسية، هي صدق المحتوى، صدق المحك (التلازمي والتنبؤي) وصدق التكوين الفرضي أو ما يسمى بصدق المفهوم.

فالدراسة توصلت إلى قبول الباحثين لمستويات صلاحية الاختبارات التي أعادوا التحقق من صدقها وثباتها على الرغم من كونها اختبارات مبنية في بيئات غربية وعلى أفراد يفترض أنهم يختلفون في قيمهم وعاداتهم وتقاليدهم، وهذا ينطبق على الرأي الذي يشير إلى أن صلاحية الاختبارات النفسية صفة تتعلق بنتائج الاختبارات وقدرتها على تقديم تفسير موضوعي لأداء الفرد على الاختبار، بمعنى أن الاهتمام يركز على التفسيرات التي تأخذها النتائج بالنسبة لعينة من

الأفراد، وللوصل إلى درجة مقبولة من الصدق والثبات على الباحث أن يقدم مجموعة من الأدلة البعض منها نظري والأخر تجريبي إحصائي للقول بأن اختباره صالحا لقياس سمة الذكاء عند مجموعة من الأفراد خصائصهم معروفة ومحددة مسبقا (بوسالم، 2014، ص198).

وهنا نشير إلى أن ثبات الاختبار أصبح ينظر إليه وفق المفهوم الحديث للصدق على أنه دليل من أدلة صلاحية الاختبار وليس خاصية معزولة تحسب بصفة انفرادية، فمن خصائص الاختبار الصالح توفره على درجة من الثبات، فلا يمكن للباحث أن يجزم بصدق اختباره إذا لم يدلل على أن اختباره على درجة مقبولة من الثبات. فالتوجه الحديث ينظر إلى الصدق على أنه يرتبك بمعنى الاختبار وتفسير النتائج، أي التركيز على صدق التكوين الفرضي للاختبار واعتبار الأنواع الأخرى جوانب متضمنة في صدق البناء. وقد تطور هذا التوجه من خلال كتابات كرونباك وميسيك (Messick et Cronbach ,1990, p158) حيث تم وصف الصدقعلى أنه نوع واحد وهو صدق التكوين الفرضي وهناك عدة إجراءات للتحقق منه.

ويري منسيك (أحد منظري نظرية الصدق الحديثة) أن صدق المحتوى يقدم شواهد حكيمة عن انتماء الاختبار لمحتوى معين، فهو يركز على انتماء الاختبار لمحتوى ومدى تمثيله له، إلا أنه لا يقدم لنا ذلك دليلا ولا معلومات عن قدرة درجات الاختبار على تفسير أداء الأفراد، وهذا أهم عامل حسب رأينا الذي جعل الباحثين في هذه الدراسة يقبلون مستويات صدق اختبارات الذكاء انطلاقا من إجراءات إحصائية بعيدة عن البعد التفسيري لنتائج الاختبار.

فالاكتفاء بما يسمى بصدق محتوى الاختبار وخاصة مع اختبارات الذكاء، يجعله يقتصر على الحكم على صياغة بنود الاختبار وليس على السلوكات التي تستثيرها هذه البنود، ويؤكد منسك أن صدق محتوى الاختبار ينبغي ألا ينظر إليه على أنه خاصية في الاختبار ولكن يمكن اعتباره خاصية ترتبط بعلاقة الاختبار بالسمة المراد قياسها، فصدق المحتوى هو مرحلة من مراحل بناء الاختبار وليس صفة يتم التحقق منها عند الانتهاء من البناء أو من طرف باحث أخر.

نستخلص من هذا أن تقدير المحكمين لصدق محتوى اختبارات الذكاء كما هو الحال مع الباحثين في دراستنا هذه يركز على خواص مرتبطة بالاختبار وليس باستخدامه أو تفسير نتائجه، وعلى اعتبار الصدق صفة ترتبط باستخدام درجات الاختبار من أجل اتخاذ قرارات معينة، فإن الاعتماد على تقرير المحكمين لا يعد دراسة كافية للتحقق من صدق الاختبار النفسي، ويبدو الأمر أكثر غموضا عندما نعتمد على صدق المحكمين لاختبارات تقيس الذكاء كسمة لا ترتبط بمحتوى معين، أي أن محتواها افتراضي، حسب (تيغزى، 2018، ص ص 22-38)، حيث لا يمكن الحديث عن محتوى محدد للذكاء وتكون بنود الاختبار عبارة عن مؤشرات للذكاء وليست عينة ممثلة لها.

فالتحكيم الذي أعتمده الباحثون بأكبر نسبة في عينة دراستنا ينحصر دوره في الحكم على الجانب اللغوي ومحتوى الاختبار، أي الحكم على صياغة البنود ومدى قياسها لأهداف معينة. إلا أن هذا الحكم لا يرتبط بصدق البند نظرا لعدم قدرة التحكيم (الخبراء) على الغوض في جوانب مرتبطة بصدق البنود، لكون الصدق يرتبط بتفسير الدرجات على البنود وليس سلامة البند لغويا أو ملائمته لمستوى المفحوص أو رأي الخبير على أنه يقيس فعلا ما أعد لقياسه.

كما نشير الى أن الافتراضات النظرية التي تقوم عليها أنواع الصدق قد همشت من طرف الباحثين في عينة الدراسة، فكل نوع من الصدق ينطوي على مجموعة من الافتراضات النظرية تتعلق بطبيعة المعطيات من جهة وخصائص العينة وظروف التطبيق من جهة أخرى، فعلى الباحث

ألا يهمل هذه الافتراضات وألا يكتفي فقط كما حدث في الدراسات عينة البحث بذكر نوع الصدق وقيمته العددية. فصدق المحتوى ومعه ما سماه الباحثون بصدق المحكمين له إفترضات لم ينتبه لها الباحثين ولم يتم التحقق منها عمليا، وهي أن يكون المفهوم النظري وهو للذكاء في دراستنا معروفا عند من يحكم الاختبار وسبق له التعامل مع مؤشراته، وفي حالة عدم تحقق هذا الشرط فإن التحكيم يصبح مجرد انطباعات لغوية وربما منطقية لن تضيف لمفهوم الصدق أي قيمة علمية.

كما أن الاعتماد على صدق المحكات إجراء يتطلب توفر محك على درجة عالية من الصلاحية (الصدق والثبات) وفي البيئة الجديدة للباحث، فلا يكفي أن نأتي باختبار أخر لم يتم التحقق من صلاحيته ونعتمده كمحك للحكم على مدى صلاحية اختبار أو مقياس أخر، وهذا هو الاجراء الذي قام به الباحثون في عينة الدراسة، فهي إما أنها تحققت من صلاحية المحكات ولم توضح ذلك، أو أنها فعلا أعتمدت محكات من دون التحقق من صلاحيتها. وهذا الشرط تغاضت عنه كل الدراسات عينة بحثنا هذا.

ويمكننا أن نستخلص مجموعة من الملاحظات تتعلق بقيام الباحثين باعادة التحقق من ثبات الاختبارات بمختلف الطرق،يمكننا تحديدابعاد همشتها الدراسات عند إعادة التحقق من ثبات الاختبارات التي تعاملوا معها.

- تجاهل الافتراضات النظرية

لما كانت كل طريقة من طرق معامل الثبات تنطوي على مجموعة من الافتراضات النظرية تتعلق بطبيعة المعطيات من جهة وخصائص العينة وظروف التطبيق من جهة أخرى، فعلى الباحث ألا يهمل هذه الافتراضات وألا يكتفي فقط كما هو معمول به في الغالبية العظمى من الدراسات والبحوث بذكر نوع الثبات والقيمة العددية له، فمعامل ثبات التجزئة النصفية تتطلب إفترضات صارمة قلما أنتبه الباحثين لها ونادرا ما حاولوا التحقق منها عمليا، وهي التوازي ما بين درجات النصفين، والتساوي التام في قيم تباين خطأ القياس التي تنطوي عليها الدرجات وفي حالة عدم تحقق التوازي فإن الاعتماد على معدالة سبرمان- براون لتصحيح معامل ارتباط درجات النصفين يؤدي إلى تقديرمتحيز وغير دقيق لمعامل الثبات. وبالتالي قيمة الاختبار تصبح محل شك حسب كل من(Spearman- Brown) صاحبا المعادلة، وهو الافتراض الذي تغاضى عنه كل الباحثين فيعينة الدراسة وتم الاكتفاء بتقديم قيمة واحدة تشير إلى معامل الثبات.

- طبيعة السمة أو الخاصية المقاسة

إن طبيعة السمة أو الخاصية التي يتعامل معها الباحث من حيث كونها أحادية البعد أو متعددة الأبعاد، ومستوى قياسها (تصنيفي، رتبي، مسافات متساوية) وما ينتج عن ذلك من القيم العددية المتحصل عليها، كل هذا له علاقة مباشرة بنوع معامل الثبات الذي يختاره الباحث، فطريقة التجزئة النصفية تقدم قيما مزيفة وغير دقيقة مع الاختبارات التي تقيس سمات متعددة الأبعاد، كما أن ثبات التجانس الداخلي وفق معادلة ألفا أو كيودر رتشردسون يتطلب كل منها معطيات مختلفة يتطلب من الباحث معرفتها والإشارة إليها عند التحقق من ثبات اختباره وفق إحدى هاتين الطريقتين.

- عدم ذكر مصدر الخطأ

من المعروف أن مختلف طرق التحقق من الثبات هي في الأصل طرق للتقليل من أخطاء القياس التي تعود إلى مصادر مختلفة، وعليه فإن اختيار طريقة دون أخرى ينطوي على تقدير من

الباحث لمصادر الخطأ التي يتوقع أنها تؤثر على درجات إختباره في ظروف وخصوصية تختلف من دراسة إلى أخرى ومن عينة على أخرى، وهنا يمكن للباحث حسب (علام صلاح الدين، 2000، ص421) أن يعتمد على أكثر من طريقة لتقدير ثبات الاختبار، ولكن بشرط أن يقدم التبريرات العلمية والعملية التي جعلته يختار التحكم في مصدر معين للخطأ دون الأخر. وهو الشيء الذي لم يقم به جميع الباحثين في عينة دراستنا.

فلكي يساهم كل بند بجانب جديد من المعلومات، فإن إثراء المفهوم أو السمة موضوع القياس عملية متسقة تتخذ منحنى منتظما وغير عشوائي ولذلك تساهم في الرفع من التباين المنتظم أو المتسق الذي يفترض فيه أنه يعكس السمة المقاسة أو جوانب منها، والخطأ العشوائي يختلف عن الخطأ المنتظم لأنه لاترتبط بالمفهوم النظري للسمة أو المحتوى الدلالي للمفهوم وإنما يرتبط بمصادر خارجية تتعلق بظروف تطبيق الاختبار وهو مصدرالخطأ.

- تفسير معامل الثبات نظريا

مما سبق تبين أن التحقق من ثبات الاختبار عملية معقدة ومركبة تتطلب من الباحث أن يقدر أبعاد مختلفة، بداية من الافترضات التي تقوم عليها كل طريقة، مرورا بطبيعة السمة المقاسة ووصولا إلى مصادر الخطأ التي قدر الباحث ضرورة التحكم فيها، هذه المراحل التنظيرية بالدرجة الأولى تحتاج إلى تقديم الباحث لقيمة عددية تشير إلى مقدار ثبات الاختبار من خلال مختلف المعادلات الاحصائية يحتاج الى تفسير وهو الشيء الغائب في البحوث التي شملتها دراستنا وهي النتيجة نفسها التي توصل اليها (بوسالم، 2017، ص ص 85-196) الذي توصل الى أن أغلب الباحثين يتجنبون تفسير قيمة معامل الثبات مما يجعله مجرد قيمة عددية خالية من البعد النظري.

- الخاتمة

من خلال هذه الدراسة توصلنا الى أن الباحثين قاموا بإعادة التحقق من الصدق والثبات بطرق غير مناسبة لطبيعة سمة الذكاء التي تقيسها الاختبارات، كما أنهم لم يقدموا التبرير العلمي لاختيارهم أنواع معينة من الصدق أو الثبات دون الأنواع الأخرى، هذا الإخلال بالمنطق العلمي وتهميش الأبعاد والمقاربات الحديثة لمفهوم الصلاحية نتج عنه قبول مستويات صلاحية هذه الاختبارات كما هي، والعلة في ذلك أن هذا الإجراء تحول إلى عمل روتيني كمي محض الهدف منه أن يثبت الباحث أنه أعاد التحقق من الصدق والثبات ووجدهما على درجات مقبولة وبالتالي يمكن الثقة في نتائج الاختبار الذي طبقه، وهذه مفارقة علمية غير منطقية، فلما كانت مستويات صدق وثبات الاختبارات مقبولة فهذا يعني أنها اختبارات عابرة للثقافات، بمعنى أنها لا تتأثر بالعامل الثقافي، ويمكن تطبيقها في الثقافات المختلفة.

ومنه نقول أن النظرة القاصرة لمفهومي الصدق والثبات والمبنية على البعد الكمي المحض هي التي جعلت الباحثين يصلون إلى قبول مستويات صدق وثبات اختبارات الذكاء التي نقلوها من بيئات مختلفة، مما يدفعنا إلى القول بضرورة إعادة النظر في طريقة التحقق من الصلاحية واعتبارها مفهوما مشبعا بالقيم الاجتماعية وله مألات وتفسيرات ترتبط بالخصائص الثقافية التي طبق فيها الاختبار مما يجعل الاختبار الصالح في بيئة معينة غير صالح إذا نقل إلى بيئة أخرى مختلفة عنها في عدد من المؤشرات الثقافية والعقائدية والاجتماعية.

- قائمة المراجع

- رجاء محمود أبو علام. (2001). النظريات الحديثة في القياس والتقويم وتطوير نظام الامتحانات، المؤتمر العربي الأول للامتحانات والتقويم التربوي، المركز القومي للامتحانات والتقويم التربوي، القاهرة.
- بوسالم عبد العزيز. (2014). *القياس في علم النفس وعلوم التربية*، الجزائر: جار قرطبة للنشر والتوزيع.
- عبد السلام نادية محمد. (2002). تعدد معطيات تحديد الدرجة الفاصلة في القياس محكي المرجع باستخدام نموذج راش أحادي البرامتر، دراسة تجريبية، حولية كلية البنات المجلد الأول، (16)، القاهرة: جامعة عين شمس، ص ص 112-125.
- عبد السلا نادية محمد. (1996). *ثبات وصدق القياس محكي المرجع*، في أنور الشرقاوي وآخرون، اتجاهات معاصرة في القياس والتقويم النفسي والتربوي، القاهرة: مطبعة الأنجلو المصرية.
- الشرقاوي أنور محمد. (1996). *الاختبارات المرجعية إلى محك، وسائل جديدة في القياس النفسي والتربوي*، في أنور الشرقاوي وآخرون، اتجاهات معاصرة في القياس والتقويم النفسي والتربوي، ط1، القاهرة: المكتبة الأنجلـومصرية.
- كاظم أمينة محمد. (1995). *اتجاهات معاصرة في بناء بنوك الأسئلة، في الأسس التربوية لإعداد المعلم الجامعي*، ط2، القاهرة: جامعة عين شمس.
- كاظم أمينة محمد. (1997). *دراسة نظرية نقدية حول القياس الموضوعي للسلوك "نموذج راش"* في أنور الشرقاوي وآخرون، اتجاهات معاصرة في القياس والتقويم النفسي والتربوي، القاهرة: مكتبة الأنجلومصرية.
- الأنصاري بدر محمد. (2000). *قياس الشخصية*، بدون طبعة، الكويت: دار الكتاب الحديث.
- ملحم سامي محمد، (2005). *القياس والتقويم في التربية وعلم النفس*، ط3، عمان، الأردن: دار المسيرة للنشر والتوزيع والطباعة.
- روبرت ثورندايك.، إليزابيت هاجن. (1989). *القياس والتقويم في علم النفس والتربية*، ترجمة: عبد الله زيد الكيلاني وعبد الرحمان عدس، ط1، عمان، الأردن: مركز الكتاب الأردني.
- صلاح أحمد مراد وآخرون. (2005). *الاختبارات والمقاييس في العلوم النفسية والتربوية*، ط2، القاهرة، مصر: دار الكتاب الحديث.
- علام صلاح الدين محمود. (1995). *الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية*، ط1، القاهرة: دار الفكر العربي.
- علام صلاح الدين محمود. (2002). *القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة*، بدون طبعة، القاهرة: دار الفكر العربي.

- Cronbach LJ. (1990). Educational Psychology, 5th Edition, Harpex and Row, New York.
- Messick S. (1990). Validity in R.linn (Ed), Educational Measurement New York, , 3rd ed, American Council on education and Macmillan publishing company.
- Wright B. (1980). Solving Measurement Problems With The Rasch Model Journal Of Educational Measurement, New York.
- Hambleton R K. (1998). Toward An Integration of Theory and Method for Criterion Referenced Tests, Review of Educational Research, New York.
- Hashway R M. (1998). Assessment and evaluation of developmental learning, Qualitative individual assessment and evaluation models. London West. port, Connecticuts, Praiger Publishers.
- Anastasi A.(1976). Psychological testing, New York, Macmillan ,4th.

- تغزى محمد (1998). في: قراءات في التقويم التربوي، كتاب الرواسي، تأليف محمد نقادي وآخرين جمعية الاصلاح الاجتماعي والتربوي، باتنة، الجزائر.
- علام صلاح الدين محمود (2000). *القياس والتقويم النفسي والتربوي*، القاهرة: دار الفكر العربي.
- بوسالم عبد العزيز (2017). الصدق الثقافي للاختبارات النفسية المطبقة في الجزائر ومشكلة التكييف من أجل الصلاحية، مجلة العلوم النفسية والتربوية، المجلد3، (2)، كلية العلوم الإنسانية والاجتماعية، جامعة الوادي، ص ص 185-196.