

بناء اختبار تحصيلي في مادة منهجية البحث العلمي وفق نظرية الاستجابة للمفردة (نموذج راش)
 constructing an achievement test in the subject of scientific research methodology
 according to the item response theory (Rasch model)

مبروكة بوخالفة*

طالبة دكتوراه، جامعة الجزائر 2

Boukhalfa Mabrouka

Doctorant, Alger2,

boukhalfa.mabrouka@univ-alger2.dz

نصيرة بن نابي

أستاذ، جامعة الجزائر 2

nacera benabi

prof, Alger2, nacer.benabi@univ-alger2.dz

تاريخ النشر: 2024/12/08

تاريخ القبول: 2024/11/28

تاريخ الاستلام: 2024/10/21

الملخص: تهدف الدراسة الحالية إلى توظيف نموذج راش لبناء اختبار تحصيلي في مادة منهجية البحث العلمي لعينة من طلبة السنة الثانية علوم التربية بجامعة الجزائر 2 وفق نظرية الاستجابة للمفردة (نموذج راش)، والذي تكون من (35) سؤالاً في صيغته الأولية مقسمة إلى ثلاثة أنواع من البدائل هي: الاختيار من متعدد، صحيح خطأ، وملاءم الفراغات، وكلها تحتوي على إجابة واحدة فقط صحيحة (1) وأخرى خطأ (0)، ولتحقيق أهداف الدراسة الحالية استخدمنا المنهج الوصفي التحليلي المناسب لطبيعة متغيرات الدراسة وتطبيق الاختبار على عينة ممثلة لمجتمع البحث (الطلبة) تم اختيارها بطريقة قصدية، بلغ عددها (110) طالب وطالبة من طلبة قسم علوم التربية بولاية الجزائر، جامعة الجزائر 2.

وللإجابة على تساؤلات الدراسة تم استخدام العديد من التحليلات الإحصائية منها: الانحراف المعياري المتوسط الحسابي، التباين، التحليل العاملي بطريقة المكونات الأساسية للبواري. لنموذج راش المطور برنامج 3.72.3 WINSTEPS Version.

بعد إجراء التحليل الإحصائي للبيانات كانت نتائج الدراسة كما يلي: حقق الاختبار افتراض أحادية البعد وحقق الاختبار افتراض الاستقلال الموضوعي للقياس. تم حذف 4 عبارات من بين 35 عبارة و6 أفراد من بين (110) طالب وطالبة لعدم مطابقتهم لنموذج راش ماستر، وتمتع الاختبار بثبات عالي، حيث بلغ معامل الثبات للفقرات (0.92) أما مؤشر الفصل فقد بلغ (4.80) في حين معامل الثبات للأفراد (الطلبة) بلغ (0.91) أما مؤشر الفصل فقد بلغ (3.33)، ويمكن اعتبار النتائج المتوصل إليها إضافة علمية تعزز ما توصلت إليه الدراسات السابقة، وهذا الاتفاق لم يكن ناتج عن الصدفة إنما نتيجة لما يتمتع به الاختبار من مواصفات تتجسد في ثلاثة حقائق هامة تتعلق بالاختبار نفسه من حيث اعتماده على إطار نظري متين واحتوائه على 35 عبارة، وهذا العدد في حد ذاته يعتبر في نظر الباحثين مقبول.

الكلمات المفتاحية: نموذج راش؛ نظرية الاستجابة للمفردة؛ الاختبارات التحصيلية.

Abstract: The current study aims to employ the Rasch model to build an achievement test in the subject of scientific research methodology for a sample of second-year students in educational sciences at the University of Algiers 2 according to the item response theory (Rasch model), which

*- المؤلف المرسل

consists of (35) questions in its initial form divided into three types of alternatives: multiple choice, true/false, and fill-in-the-blanks, all of which contain only one correct answer (1) and one incorrect answer (0). To achieve the objectives of the current study, we used the descriptive analytical method appropriate to the nature of the study variables and applied the test to a sample representing the research community (students) that was chosen intentionally, numbering (110) male and female students from the Department of Educational Sciences in the state of Algiers. University of Algiers

To answer the study questions, many statistical analyses were used, including: standard deviation, arithmetic mean, variance, factor analysis using the basic components method of the residuals. For the developed Rasch model, WINSTEPS Version 3.72.3 program. After conducting the statistical analysis of the data, the results of the study were as follows: The test achieved the assumption of one-dimensionality and the test achieved the assumption of positional independence of measurement. 4 statements out of 35 statements and 6 individuals out of (110) male and female students were deleted for not conforming to the Rush Master model. The test has high stability, as the stability coefficient for the paragraphs reached (0.92) and the separation index reached (4.80) while the stability coefficient for individuals (students) reached (0.91) and the separation index reached (3.33). The results reached can be considered a scientific addition that enhances what previous studies have reached. This agreement was not the result of chance, but rather the result of the test's specifications embodied in three important facts related to the test itself in terms of its reliance on a solid theoretical framework and its containing 35 statements, and this number in itself is considered acceptable by researchers.

Keywords: The Rasch model; vocabulary response theory; achievement tests

مقدمة:

يسعى الباحثين في مجال القياس والتقويم التربوي، إلى تطوير العملية التعليمية وأساليب التقويم وبناء الاختبارات، ونتيجة لوجود العديد من المشكلات التي تنتج عن القياس التقليدي والتي تتسبب في عدم دقة النتائج التي تسفر عنها أساليب وأدوات القياس والتقويم، بحيث تعتبر الاختبارات من الوسائل الضرورية التي تعتمد عليها المؤسسات التعليمية العالي بشكل عام والأساتذة بشكل خاص في تقويم الطلبة لمعرفة مواطن القوة وتعزيزها والكشف عن مواطن الضعف لمعالجتها، وكل هذا يمر عبر بناء الاختبارات التحصيلية التي هي موضوع بحثنا، والتي نستعمل فيها نموذج راش الذي سيساعدنا في تحليل البيانات؛ أما فيما يخص المادة التي يبني فيها الاختبار فهي مادة منهجية البحث العلمي، بهدف الوصول إلى اختبار تحصيلي موضوعي.

1- الإشكالية

شهد تاريخ القياس النفسي تطورات عدة تحولت بمرور الزمن من القياسات التقليدية إلى الاتجاهات الحديثة، حيث كان القياس التقليدي يعتمد على مفاهيم بسيطة تعود إلى بدايات القرن العشرين، ولكنه تطور ليشمل نظريات أكثر تعقيداً حيث شُرع في البحث عن أساليب جديدة للتقويم والقياس، مما يسهل فهم السمات النفسية والتربوية بصورة أشمل وأدق ويحقق نتائج أكثر توازناً فوق نظرية القياس التقليدية، يمكن التعبير عن الدرجة الحقيقية من خلال قدرة الفرد التي تتضح من خلال أدائه على الاختبار، تعتمد النظرية التقليدية على الاعتقاد بأن الأداء الفردي يتغير استناداً إلى مستوى الاختبار الذي يُقدم له، وبالتالي يُعتمد على أن كل اختبار له سماته الخاصة التي تؤثر في النتائج المستخرجة من الأفراد، مما يحتم على العلماء فهم هذه الديناميكيات بعناية لضمان دقة القياس، حيث يتأثر أداء الأفراد بشكل كبير بمستوى الاختبار، سواء من حيث السهولة أو الصعوبة، فالاختبارات بمستويات منخفضة قد تؤدي إلى نتائج تعكس قدرات الأفراد بشكل غير عادل، في حين أن الاختبارات ذات مستوى عالٍ قد تبرز فروقات الأداء بين الأفراد بشكل أوضح لذلك، من الضروري اختيار الاختبارات التي تتماشى مع مستوى القدرات الفردية لضمان دقة النتائج، مما يساعد في تعزيز فعالية استراتيجية القياس (عودة، 1992، ص. 155).

وأُسفرت جهود العلماء عن ظهور بعض الاتجاهات الحديثة في مجال القياس والتقويم، والتي جاءت نتيجة كثرة الانتقادات الموجهة لنظرية التقليدية في القياس ومنها نظرية الاستجابة للمفردة (Item Response Theory)، تعمل هذه النظرية على تحليل استجابات الأفراد بالمفردات المختلفة، مما يسمح بفهم أعمق لكيفية تأثير الخصائص النفسية للأفراد على أدائهم وتعتبر هذه النظرية بديلاً حديثاً للقياس التقليدي، حيث تتغلب على العديد من المشكلات المرتبطة بالاختبارات التقليدية، حيث تقوم الاختبارات النفسية والتربوية على فرضية وجود سمات أو خصائص معينة يشترك فيها جميع الأفراد، رغم اختلافهم في مقدارها، يُعبر عنها بالسمات الكامنة من خلال سلوك الفرد واستجاباته لمفردات الاختبار، وتُعد هذه الأخيرة مهمة في تقديم فهم شامل للأداء الفردي، مما يدعم اتخاذ قرارات أكثر دقة بشأن احتياجات الأفراد وأهدافهم التعليمية، تُشير كذلك إلى الخصائص النفسية غير المرئية، وتمثل نماذج السمات الكامنة أداة حيوية في مجالات القياس النفسي والتقويم، هذه النماذج تهدف إلى فهم العلاقة بين الاستجابات الملاحظة في الاختبارات وما يمكن أن يعرف بالسمات غير الملاحظة، وتعتبر هذه النظرية ضرورية لتحديد كيفية تأثير السمات الفردية على استجابات الأفراد، مما يساهم في تطوير أدوات قياس أكثر دقة وموضوعية (علام، 1987، ص. 22).

تشير نماذج السمات الكامنة إلى إطار إحصائي يُستخدم لتعزيز الفهم حول كيف تؤدي السمات غير المرئية إلى استجابات معينة على الاختبارات، هذه النماذج تمكن الباحثين من التمييز بين الصفات الفطرية والبيئية، وبالتالي تقديم تفسيرات موضوعية للنتائج الملاحظة، تعتمد أهميتها على قدرتها على توفير رؤى عميقة حول الأفراد وخصائصهم، إن العلاقة التي تربط الاستجابات الملاحظة بالسمات غير الملاحظة هي محور الدراسات النفسية. كما تشير "أمانة كاظم" (1996)، فإن هذه النماذج تحدد شكل ودرجة هذه العلاقة، حيث توضح كيف يمكن أن تُعزى الاستجابات إلى السمات الأساسية التي تُعتبر غير مرئية، مما يتيح للباحثين تحديد وقياس هذه السمات بدقة. تعتبر هذه العلاقة جوهرية للحصول على قياسات صحيحة وموثوقة (كاظم، 1996).

وأحد الانتقادات الرئيسية في هذا الإطار هو تأثير العينات على مواضع الأفراد. إذ يؤكد "صلاح الدين علام" (2000) أن التقديرات الكمية التي تم الحصول عليها لمواقع الأفراد يجب أن تكون مستقلة عن طبيعة العينات التي ينتمون إليها. هذا الاعتراض يشير إلى ضرورة تطوير نماذج تسمح بالتقييم العادل لكل فرد دون التأثير السلبي للعينات المتنوعة (علام، 2000، ص. 203).

وتهدف نماذج الاستجابة للمفردة إلى اشتقاق قيم تقديرية تقريبية للسمات بناءً على مجموعة من الاستجابات لمجموعة محددة من المفردات، يشير هذا إلى أنه عند قياس القدرة أو الخاصية يجب أن تكون العلاقة بين مستويات الأداء واحتمالات الاستجابة صحيحة غير منتظمة، هذه الفرضية تدعم الفهم الأعمق لطبيعة القياس النفسي وتساعد في تطوير أدوات أكثر دقة. وتعتبر غياب العلاقة المنتظمة بين السمات والاستجابات الصحيحة أحد أعمق الأفكار ضمن هذا الإطار، يمثل هذا الغموض تحديًا كبيرًا للباحثين، إذ أن العديد من الفرضيات حول السمات والمقاييس تحتاج إلى إعادة تقييم، وتعتبر هذه المسألة نقطة مهمة للنقاش، حيث تسلط الضوء على الصعوبات في القياس وتتطلب إعادة التفكير في كيفية تصميم الاختبارات السلوكية والنفسية. (عبد المسيح، 1991، ص. 443)

ونموذج راش من نماذج السمات الكامنة وأبسطها وأكثرها استخداماً في بناء الاختبارات والمقاييس النفسية والتربوية. يقوم هذا النموذج على الفرضية الأساسية التي تقول إن كل فقرة تحمل شحنة انفعالية، هذه الشحنة تتضاف مع شحنات الفقرات الأخرى، مما يساهم في تكوين شحنة انفعالية إجمالية تعبر عن اتجاه الفرد، ويتم تقييم هذه الشحنات بناءً على تقديرات الفقرات، مما يتيح للمقياس تقييم استجابات الأفراد بطريقة كمية، يتأسس نموذج راش على فكرة أن الأداء في اختبارات متعددة الفقرات يمكن أن يُعبر عنه من خلال نموذج رياضي، يعتمد النموذج على عدم الارتباط بين استجابات المشاركين وأي متغيرات خارجية، ويركز على الشحنات الانفعالية

لكل فقرة وكيف تساهم هذه الشحنات في تحديد سلوك الأفراد، ويمكن اعتبار أن كل فقرة تنقل مستوى مختلف من الاتجاهات النفسية بناءً على تقديرات المشاركين، هذه الأساسيات تجعل من نموذج راش أداة قوية للفهم والتحليل النفسي (علام، 1987، ص. 22).

ولقد أجريت العديد من الدراسات حول فعالية نموذج راش في اختيار فقرات الاختبارات. منها دراسات "رامسدين وانتوستل" و"انتوستل"، والتي وفرت إشارات قوية على اتساق نموذج راش مع القياسات النفسية. توضح نتائج هذه الأبحاث كيف أن نموذج راش يساهم في تحسين جودة البيانات بالإضافة إلى ذلك، تعتبر هذه الدراسات دليلاً على قوة النموذج في مجالات متعددة من العلم العوامل المستخدمة في هذه الدراسات تؤكد على فعالية الشحنات الانفعالية في تحكم الاختبارات النفسية (تركي، 1990، ص. 85).

إن الدقة في اختيار الأساليب المستخدمة في بناء الاختبارات تعني أنه يجب أن تكون هذه الأساليب مدعومة بأدلة وأبحاث تجريبية قوية. تكمن أهمية هذه الأساليب في كونها تعكس القدرة الحقيقية التي يقيسها الاختبار. مهاجمة العيوب المرتبطة بالنماذج التقليدية تتطلب دراسة مدى نجاعة الأساليب المتبعة. من الأهمية بمكان تحليل مدى فعالية هذه الأساليب في مختلف السياقات التعليمية والنفسية لتحديد مدى قدرتها على تقديم نتائج قوية وموثوقة (العيسوي، 1998، ص. 75).

وتعتبر الاختبارات التحصيلية أداة مهمة من أدوات القياس. تبحث في مدى فهم الطالب واستيعابه للمواد الدراسية. وفقاً لعودة (2002) يُعرف الاختبار التحصيلي بأنه: مجموع الأسئلة المقدمة بطريقة منظمة لتحديد مستوى تحصيل الطالب لمعلومات ومهارات محددة. يتم هذا التحديد من خلال إجابات الطلبة على الأسئلة المصممة لقياس قدراتهم ومدى استيعابهم للمادة المحددة. باستخدام هذه الطريقة، يمكن إعطاء نظرة شاملة حول الأداء الأكاديمي للطلاب، وتتضمن الاختبارات التحصيلية عدة مكونات، تشمل نوعية الأسئلة، والتنسيق، وطرق التقييم ويجب أن تكون الأسئلة مصممة بشكل دقيق لتقيس المهارات والمعلومات المستهدفة. بالإضافة إلى ذلك، ينبغي أن تعكس الأسئلة المنهج الدراسي وأن تكون مناسبة لمستوى الطلاب. فهم هذه العناصر يساعد في بناء اختبارات فعالة تعكس الأداء الحقيقي للطلاب (محاسنة، 2013، ص. 288).

يمكن استخدام الاختبارات التحصيلية في العديد من السياقات التعليمية، مثل التقييمات الفصلية أو النهائية. تساعد هذه الاختبارات في تحديد مدى فهم الطلاب لمحتوى الدروس. كما يمكن أن توفر استبيانات تفصيلية تساعد المعلمين في تحسين أساليب التدريس، وبهذا فإن الاختبارات التحصيلية تعزز من دور التعليم في تطوير مهارات الطلاب

تتطلب المناهج التعليمية في التعليم العالي نمطاً مختلفاً مقارنة بالمرحلة التعليمية السابقة يتمثل هذا الاختلاف في تحويل دور الطالب من مجرد متلقي للمعلومات إلى باحث نشط يسعى للحصول على المعرفة. وهذا يتطلب فهماً عميقاً للمنهجيات الخاصة بكتابة وتوثيق المعلومات وفي هذا الإطار، تم تناول الأدبيات المتعلقة ببناء الاختبارات التحصيلية بشكل مكثف، ومن المهم أن نفهم أن بناء اختبار تحصيلي ناجح يتطلب مراعاة الخصائص السيكومترية وتطبيق منهجيات الكتابة الصحيحة. وإن تعزيز الفهم لدى الطلاب حول كيفية بحثهم عن المعلومات سيساهم في تحسين جودة التعليم ومن المؤكد أن هذه التطورات سترفع من مستوى المعرفة والمهارة لدى الأجيال المستقبلية.

وعلي حسب حدود اطلاع الباحثة في أدبيات الدراسة التي تناولت بناء اختبار تحصيلي علي طلبة الجامعة وفق النظرية الحديثة للقياس في منهجية البحث العلمي تعتبر هذه الدراسة من الدراسات القليلة التي تناولت هذه المتغيرات وفق نموذج راش لذلك قامت الباحثة ببناء اختبار تحصيلي باتباع جميع خطوات البناء والتأكد من الخصائص السيكومترية لهذه الأداة لذلك طرحنا تساؤلات الدراسة التالية:

- هل فقرات الاختبار التحصيلي في مادة منهجية البحث العلمي لدي طلبة السنة الثانية علوم التربية متطابقة مع نموذج راش وذلك باستعمال برنامج الحاسب الآلي؟
 - هل تحقق فقرات الاختبار التحصيلي في مادة منهجية البحث العلمي لدي طلبة السنة الثانية علوم التربية ميزة أحادية البعد كما يتطلبه نموذج راش -ماسترز؟
 - ماهي خطوات بناء اختبار تحصيلي في مادة منهجية البحث العلمي لطلبة السنة الثانية علوم تربية بجامعة الجزائر 2 وفق نموذج راش؟
- 2-فرضيات الدراسة:

- إن هذه الدراسة تهدف بشكل أساسي لاي تحقيق الافتراضات التالية:
- فقرات الاختبار التحصيلي في مادة منهجية البحث العلمي لدي طلبة السنة الثانية علوم التربية مطابقة لنموذج راش وذلك-ماسترز باستعمال برنامج الحاسب الآلي winsteps.
 - تحقق فقرات الاختبار التحصيلي في مادة منهجية البحث العلمي لدي طلبة السنة الثانية علوم التربية ميزة أحادية البعد كما يتطلبه نموذج راش -ماسترز.
 - تحقق فقرات الاختبار التحصيلي في مادة منهجية البحث العلمي لدي طلبة السنة الثانية علوم التربية ميزة الاستقلال الموضوعي كما يتطلبه نموذج راش -ماسترز.

- تحقق فقرات الاختبار التحصيلي في مادة منهجية البحث العلمي لدي طلبة السنة الثانية علوم التربية درجة عالية من الصدق والثبات باستخدام نموذج راش -ماسترز.
- معرفة كيف يمكن استخدام نموذج راش في بناء اختبار تحصيلي في منهجية البحث العلمي
- الكشف عن كيفية استخدام برنامج الحاسب الآلي "winsteps" في تحقيق القياس الموضوعي وتفسير للنتائج.
- معرفة كيف تفسر نتائج هذا الاختبار باستخدام نموذج راش.
- الكشف عن قيم تقديرات معلم الصعوبة لفقرات الاختبار وفق نموذج راش.

3. أهداف الدراسة:

- إن هذه الدراسة تهدف بشكل أساسي ومباشر إلى بناء اختبار تحصيلي في مادة منهجية البحث العلمي:
- * التعرف على مدى اتساق الخصائص السيكومترية من صدق وثبات لاختبار تحصيلي في مادة منهجية البحث العلمي بعد تطبيقه على عينة من الطلبة.
- * توفر اختبار مقنن على البيئة المحلية.
- * إضافة جديدة لأدوات القياس يراعى فيها خصائص الاختبار الجيد وفق نظرية الاستجابة للمفردة.

4-أهمية الدراسة:

- تكمن أهمية الدراسة الحالية فيما يلي:
- * أنها تتطرق إلى موضوع على درجة كبيرة من الأهمية على المستوى البحثي في القياس النفسي وهو بناء اختبار تحصيلي في مادة منهجية البحث العلمي وفق نموذج راش.
- * الاهتمام باختيار هذه العينة لأن معظم الدراسات التي تناولت بناء الاختبارات ركزت على كل المستويات ما عدا طلبة السنة الثانية.
- * الشعور بأهمية دراسة عملية بناء الاختبار التحصيلي وفق نموذج راش والسعي لمعرفة المراحل التي يجب مراعاتها للحصول على اختبار موضوعي.
- * كما تبرز أهمية الدراسة في أنها تلفت نظر المختصين في القياس النفسي إلى القيام بدراسات في مجال بناء الاختبارات وتقنياتها لإثراء البحث العلمي وفق النظرية الحديثة للقياس.

5-تحديد المصطلحات:

- الاختبار التحصيلي: هو أداة تحدد مستوي الإنجاز المعرفي لطالب في مقرر دراسي كان قد تعلمها مسبقا في المدرسة من خلال إجاباته على عينة من المثيرات التي تمثل السمة أو القدرة المراد قياسها. (عوذه، 2002).

- إجرائيا: هو الدرجة التي يحصل عليها الطالب في الاختبار التحصيلي لمادة منهجية البحث العلمي من خلال إجابته على مجموع الأسئلة المقدمة.

- نموذج راش: يستخدم في بناء الاختبارات وتحليل الفقرات من نماذج نظرية السمات الكامنة وأبسطها، جاء هذا النموذج لتصحيح وأوجه القصور المقدمة لنظرية التقليدية (شحتة، 1999).
5. الدراسات السابقة: سنقوم بعرض بعض الدراسات العربية والأجنبية والتي لا تلم بموضوع الدراسة مباشرة، ولكنها قد تمس جانبا من جوانبه، قصد الاستعانة بها في مناقشة نتائج الدراسة ولقد جاءت على النحو الآتي:

- دراسة مرنيز: (2019): تهدف هذه الدراسة إلى البحث في بناء الاختبارات التحصيلية وفق نظرية الاستجابة للمفردة (نموذج راش)، لتحقيق هذا الهدف تم بناء اختبار تحصيلي في مادة المعلوماتية لسنة أولى ثانوي وفق نموذج راش يتكون من (51) بند ذات أشكال متعددة على عينة مكونة من (405) تلميذ بالنسبة إلى أولى ثانوي من مدينة غليزان للعام الدراسي 2019/2018.

المنهج المستخدم في الدراسة المنهج الوصفي، ومن أجل تحليل البيانات استخدمت البرمجية الإحصائية (winsteps)، وبرنامج الرزمة الإحصائية للعلوم الاجتماعية (spss) النسخة (25).

أظهرت نتائج الدراسة عدم تطابق (14) فقرة لافتراضات نموذج راش، وأيضا عدم ملائمة (19) فردا مع (14) فقرة بياناتهم للنموذج، أما بخصوص صدق وثبات الاختبار فقد تم التحقق من صدق التكوين الفرضي وهذا من خلال مختلف الأدلة حسب الصدق الحديث، أما بخصوص الثبات فكان (0.92) بالنسبة للفقرات مع معامل الفصل (6.88) (فقرات الاختبار مرتبة بدقة واتساق وفق صعوبة الفقرات) و(0.66) بالنسبة للأفراد مع معامل فصل (1.32) (الاختبار استطاع التمييز بين فئتين من التلاميذ وفق كفاءتهم). وتوصلت الدراسة أيضا إلى عدم وجود اختلاف بين أشكال الفقرة وصعوبتها وفق نموذج راش.

- دراسة عبد الوهاب: (2017): هدف الباحث إلى بناء اختبار تحصيلي محكي المرجع في مادة القياس النفسي مكون من 33 بندا، كلها من نوع الاختيار من المتعدد بأربعة بدائل، وقد تم التحليل السيكميومي وفق النظريتين، النظرية الكلاسيكية للقياس والنظرية الحديثة للقياس (نموذج راش)، تم تطبيقه على (146) طالبا وطالبة، وقد تم قبول كل المفردات في هذه الدراسة بالنسبة للنظرية الكلاسيكية للقياس في حين تم حذف (04) مفردات لعدم مطابقتها للنموذج، أي أن الاختبار أصبح مكونا من (29) مفردة، وقد أشارت النتائج إلى ارتفاع تقدير درجة الثبات وفق نموذج راش منه في النظرية الكلاسيكية مما يعطى أكثر دقة وموضوعية في القياس بالنسبة للنظرية الحديثة للقياس.

- دراسة هاتن: (1980) Hatton: هدف الباحث إلى المقارنة بين الاختبارات التي تبني مفرداتها بالطريقة الكلاسيكية والاختبارات التي تبني مفرداتها باستخدام نموذج راش من حيث الدقة والفعالية وطبق الباحث اختبارا في الرياضيات يتكون من (40) مفردة على عينة بلغت (6987) تلميذا من الصف التاسع، ثم تم سحب عينتين عشوائيتين منها، حجم كل منهما 200 تلميذ، وقسمت كل من العينتين إلى عينتين فرعيتين متساويتين، تم استخدام العينة الفرعية الأولى في تدريج المفردات بالطرق التقليدية، وقد تمت عملية تحليل البيانات من خلال ثلاث مراحل هي:

- المرحلة الأولى: اختيار المفردات التي تحقق أفضل ملائمة للنموذج باستخدام إحصاء متوسط المربعات للملاءمة التقاربية (إحصاء الملاءمة التقاربي).

- المرحلة الثانية Validation Cross: بحث الصديق المستعرض حيث خضعت مفردات الاختبارين لبرنامج موحد لتحلل المفردات باستخدام أربع متغيرات تابعة لقياس دقة الاختبار هي: الثبات والخطأ المعياري للقياس، وصعوبة المفردة، ومعامل الارتباط الثنائي للمفردة

- المرحلة الثالثة: أظهر التحليل وجود فرق واحد ذي دلالة إحصائية في الدقة، حيث ميز الاختبار الذي استخدم الأسلوب الكلاسيكي في اختيار مفرداته بطريقة أفضل من الاختبار الذي استخدم نموذج راش في اختيار مفرداته، كما أظهر التحليل الإحصائي عدم وجود دليل ثابت يدعم فعالية أي من الطريقتين في عملية اختيار المفردات. وقد أشار في النهاية (Hation (1980) أن المقارنة بين طريقتي اختيار المفردات ربما تميل لصالح نموذج راش لأنه أكثر جاذبية نظرا لما يتمتع به من مميزات ثبات القياس

- دراسة (1999) Afrassa & Keeves: هدف الباحث إلى إحداث تغييرات في التحصيل الدراسي لمادة الرياضيات لطلاب المدارس الإعدادية بأستراليا على مدار فترة زمنية باستخدام برنامج الحاسب الآلي الكويست Quest. وتم استخدام نموذج راش أحادي البارامتر في بناء مقياس مشترك لتحصيل التلاميذ المادة الرياضيات، ويكون هذا المقياس المشترك مستقلا عن كل من عينات التلاميذ الذين يطبق عليهم الاختبار وعينات المفردات المستخدمة ويستخدم المقياس لفحص التغيرات في تحصيلات التلاميذ في مادة الرياضيات بأستراليا خلال (30) سنة ابتداء من عام 1964 إلى عام 1994، وانقسمت الدراسة إلى سبعة أجزاء الجزء الأول يستعرض إجراءات العينة الجزء الثاني يفحص إجراءات القياس المستخدمة في الدراسة، الجزء الثالث يهتم بالإجراءات الإحصائية المطبقة في التدريج وإعطاء الدرجات في اختبارات الرياضيات، الجزء الرابع ما إذا كانت مفردات الرياضيات المستخدمة في الدراسة تلائم نموذج راش من عدمه، الجزء الخامس يناقش إجراءات التكافؤ المستخدمة في الدراسة الجزء السادس يفحص النتائج والخلاصة، واشتملت عينة الدراسة على

(4320) تلميذا في المرحلة الإعدادية للمراحل الدراسية السابعة والثامنة، والتاسعة، وبلغت أعمارهم (13) سنة، وكانوا موزعين عبر المراحل الدراسية السابعة والثامنة والتاسعة. وأوضحت النتائج قوة المقياس العام والتغيرات في التحصيل الدراسي لمادة الرياضيات لدى التلاميذ خلال الفترة الزمنية ابتداء من 1964 إلى 1984 في أستراليا وأن مؤشرات إحصاء متوسط المربعات للملاءمة التباينية توضح أن مفردات الاختبار لاءمت نموذج راش وأن جميع استجابات التلاميذ لاسـت نموذج راش، ما عدا التلاميذ الذين حصلوا على درجات صفـرية أو درجات تامة.

• الجانب الميداني

فيما يلي عرض لما جاء في الدراسة الأساسية من منهج ومجالات ومجتمع وعينة الدراسة وأدواتها وإجراءاتها التطبيقية، وخصائصها السيـكومترية، وأخيرا أساليبها الإحصائية.

1-منهج الدراسة: اعتمد في هذه الدراسة على المنهج الوصفي لأنه المنهج المناسب لمشكلة الدراسة، بهدف وصف الظاهرة والأحداث وجمع الحقائق.

2. حدود الدراسة:

تحدد الدراسة الحالية بثلاثة مجالات:

- المجال البشري: طلبة السنة الثانية علوم التربية وتشمل الذكور والإناث.
- المجال المكاني: تم تنفيذ مجريات الدراسة الحالية في جامعة الجزائر 2.
- المجال الزمني: تم تنفيذ إجراءات الدراسة الحالية خلال الفترة الممتدة بين 2023/04/05 و2023/06/11 من العام الدراسي 2023-2024.

3. مجتمع الدراسة:

يعرف محمود أبو علام مجتمع الدراسة بأنه جميع الأفراد أو الأشياء أو العناصر الذين لهم خصائص واحدة يمكن ملاحظتها (علام، 2006، ص. 149). ويتكون مجتمع الدراسة من طلبة السنة ثانية علوم تربية بكلا الجنسين، ولقد بلغ عدد الطلبة 110 طالب وطالبة.

4. عينة الدراسة:

تكونت عينة الدراسة من (110) طالب وطالبة، تم اختيارهم بطريقة القصدية، وقد تميز أفراد هذه العينة بالخصائص التالية المبينة في الجدول (1).

جدول (1) يمثل خصائص أفراد عينة الدراسة

المتغير	التكرار	النسبة %
إناث	73	66%
ذكور	37	34%
المجموع	110	100

4- الدراسة الاستطلاعية:

1.4. أهداف الدراسة الاستطلاعية: تسعى الباحثة من خلال الدراسة الاستطلاعية لتحقيق الأهداف التالية:

- التعرف على مجتمع الدراسة وتحديد عينته.
- التعرف على مدى إمكانية تطبيق الاختبار في البيئة الجامعية وفق الإمكانيات المتاحة.
- التعرف على مدى استيعاب المفحوصين لتعليمات الاختبار.
- التعرف على مدى تفاعل الطلبة من خلال استجاباتهم على الاختبار.

2.4. إجراءات الدراسة الاستطلاعية:

قمنا بتطبيق الاختبار على عينة مكونة من (38) طالب وطالبة من طلبة الجامعة، قسم علوم التربية بجامعة الجزائر 2، وتم اختيارها قصديا، وذلك بتاريخ 2023/11/11 وأثناء الدراسة تم معاينة مكان إجراء التطبيق واللقاء ببعض الأساتذة والطلبة وإجراء حوار ومناقشات حول موضوع الدراسة، وتقديم شروحات، والغرض العام من تطبيق الاختبار.

3.4. نتائج الدراسة الاستطلاعية: كانت نتائج الدراسة الاستطلاعية التي قامت بها الباحثة كالتالي:

- قدمت تسهيلات من قبل السيد عميد الكلية والأساتذة.
- تفاعل وتجاوب من طرف أفراد العينة
- تعليمات الاختبار كانت واضحة بدرجة كبيرة.
- تذمر واستغراب الطلبة من طول الاختبار
- تراوح زمن الإجابة على الاختبار ما بين (30-60) دقيقة، بمتوسط قدره (45) دقيقة.

5- أداة الدراسة: اعتمدت الباحثة في دراستها لجمع البيانات على بناء اختبار تحصيلي في مادة منهجية البحث العلمي وفق نموذج راش:

5.5-1- بناء الاختبار التحصيلي: تم اتباع الخطوات العملية في بناء الاختبار التحصيلي في وحدة منهجية البحث العلمي وفق نموذج راش لطلبة السنة الثانية علوم التربية والاعتماد على البرنامج السنوي المعد من طرف وزارة التعليم العالي هذه الخطوات هي:

- تحديد الهدف من الاختبار: تهدف الدراسة الحالية إلى قياس تحصيل طلبة السنة الثانية علوم تربية في مادة منهجية البحث العلمي وفق نموذج راش وذلك من خلال بناء اختبار تحصيلي في هذه المادة.

- تحليل محتوى مادة منهجية البحث العلمي: من البرنامج السنوي المقدم من الوزارة تم تحليل محتوى مادة منهجية البحث العلمي إلى أربع محاور رئيسية هي:

أولاً: موضوع الدراسة.

ثانياً: إشكالية الدراسة.

ثالثاً: فرضيات الدراسة.

رابعاً: مناهج البحث.

بعدها تم عرض هذا التحليل على مجموعة تتكون من 15 خمسة عشرة من المحكمين، قام المحكمون بإعطاء ملاحظات حول تحليل المحتوى وتحديد عناصر الاتفاق وعناصر الاختلاف في هذا التحليل. ولقد أخذت الباحثة بعين الاعتبار المقترحات المقدمة

- صياغة الأهداف السلوكية: تم صياغة 35 هدفا سلوكيا متعلقا بمادة منهجية البحث العلمي، وهي متمثلة في أربعة مستويات معرفية مختلفة حسب تصنيف بلوم هي كما في الجدول التالي:

الجدول رقم (02) يمثل مستويات الأهداف وعددها حسب تصنيف بلوم للأهداف

عدد الأهداف	مستويات الأهداف
15	التذكر
12	الفهم
5	التطبيق
3	التحليل

- إعداد جدول المواصفات: هو مخطط يربط العناصر الأساسية لمحتوى مادة منهجية البحث العلمي بمجالات التقييم ومهاراته الفرعية، يقيس مدى صدق المحتوى، والصدق يحتوي على عنصرين مهمين هما، الشمولية والتمثيل تعني يجب أن تكون فقرات الاختبار تمثل عينة المحتوى تمثيلاً صادقا لجميع جوانب التحصيل.

- صياغة الأهداف التعليمية للمادة بعبارات سلوكية تعكس نواتج تعليمية يمكن ملاحظتها وقياسها وهذا يأتي تأكيداً على أن العملية التعليمية -ينبغي أن يخطط لها وتنفذ وتقوم في ضوء أهداف تعليمية محددة.

- تحديد نوع فقرات الاختبار وتقييمها: تم إعداد الصيغة الأولية لفقرات الاختبار التحصيلي، وهي من نوع أسئلة الاختيار من متعدد وأسئلة صح والخطأ وأسئلة ملئ الفراغ لقياس الأهداف السلوكية واشتملت كل فقرة من فقرات الاختيار من متعدد على أربع اختيارات واحد منها فقط صحيحة.

وقد اقترح المحكمون بعد المراجعة تعديل الفقرات المكررة أو المتشابهة والتي تقيس نفس الهدف السلوكي مثل الفقرات (1، 7، 10، 21، 23)، كذلك تعديل البديل رقم (2) الخاص بالفقرة 13 والبديل رقم (3) الخاص بالفقرة 23

- تجريب المفردات: قام الباحثة بتجريب الاختبار على عينة تكونت من (110) طالب وطالبة من طلبة السنة الثانية علوم تربية تم اختيارهم بالطريقة القصدية وذلك لتحقيق من:
 - سلامة صياغة الفقرات وملاءمة البدائل ولتعرف على الفقرات الغامضة والتي تحتاج إلى تعديل.
 - تحديد الزمن اللازم للإجابة على الاختبار وقد تم حسابه بإيجاد متوسط الزمن بين تسليم أول ورقة إجابة وبين تسليم آخر ورقة، لذلك فقد حدد الزمن اللازم للاختبار بـ 45 دقيقة.
 - طريقة التصحيح وتقدير الدرجة: طبق هذا الاختبار بطريقة جماعية، أي يجب عليه مجموعة طلبة في نفس الجلسة، تتم الإجابة على الاختبار بتحديد رمز الإجابة الصحيحة.
6. الأساليب الإحصائية: وتمثل في:

- الانحراف المعياري.
- التحليل العاملي بطريقة المكونات الأساسية (Masters) والذي يسمى بنموذج التقدير.
- برنامج WINSTEPS Version 3.72.3

7. عرض ومناقشة وتفسير النتائج:

هل فقرات الاختبار التحصيلي في مادة منهجية البحث العلمي لدى طلبة السنة الثانية علوم التربية متطابقة مع نموذج راش وذلك باستعمال برنامج الحاسب الآلي "winsteps"؟

تم اعتماد طريقة نموذج راش للمكونات الأساسية المعتمدة على البواقي: نظرا لأن نموذج راش بكل امتداداته يفترض أحادية البعد إلا أن أحادية البعد ليست مطلقة، ويجب ألا ينظر إلى مفهوم أحادية البعد كما يعنهما التحليل العاملي حيث يختلف الهدف بينهما، فالتحليل العاملي يهدف إلى تحديد العوامل التي يتكون منها الاختبار ولكن في نظرية الاستجابة للمفردة تهدف إلى التعرف على هل الانحرافات عن السمة المقاسة ترقى إلى أن تكون عاملا مستقلا أم لا؟ ويقصد بذلك مدى قدرة الفقرة على قياس السمة موضع القياس، واتفاقها مع باقي الفقرات في ذلك، ويستخدم في هذا الشأن اختبار (T) للمطابقة الكلية (Total (T) Fit Statistics) ويعتمد هذا الاختبار على الافتراض القائل بأن الفقرات جميعها تقيس السمة نفسها موضع القياس، أي أن الفقرات تتصف بأنها أحادية البعد، حيث توفر دليلا على مدى اتساق الاستجابات على فقرة معينة مع الدرجة الكلية على الاختبار (كتفي، 2020، ص. 253)

وفقاً لـ Hattie، أن بعض الأساليب الإحصائية، مثل تحليل العوامل، والتي هي جزء من فئة نماذج النظرية الكلاسيكية، ستكون أقل أهمية من تلك النماذج في نظرية السمات الكامنة لكشف بُعدي البيانات المرتبطة بالاستجابة على البند. (Linacre, John.M, 1998).

في تحليل العوامل المشتركة، نحاول تحسين القواسم المشتركة، والتكبير، والتدوير، والانحراف لإعطاء أقوى بنية عاملية ممكنة، حيث يتم تفسير عمليات تحميل العوامل على أنها ارتباطات مع العوامل الكامنة المرجوة. وفي نموذج التقدير الجزئي من خلال تحليل البواقي، نحاول تنفيذ الفرضية القائلة بأن البواقي مبعثرة عشوائية (random noise) من خلال البحث عن المكون الذي يفسر أكبر درجة ممكنة من التباين في البواقي. هذا هو "التباين الأول" (أو مكون PCA الأول في مصفوفة ارتباط البواقي). إذا كانت قيمة التباين الأول صغيرة (عادة أقل من 2.0)، فإن التباين الأول يكون عند مستوى العشوائية ولم يتم تنفيذ افتراض البعثرة العشوائية بشكل عام، ولابد من النظر إلى نسبة هذه القيمة، فإن كانت أقل من 10 من المئة فهذا يؤكد أحادية البعد ولا ينفيها (كتفي، 2020، ص. 254)

الجدول رقم (02): يمثل التحليل العاملي للمكونات الأساسية للبواقي باستخدام نموذج راش – ماسترزود تم

استخراجها باستخدام (Winsteps)

المتوقع Expected	الملاحظ Observed	الجزر الكامن Empirical	حجم تباين البواقي المعيارية Table of STANDARDIZED RESIDUAL Variance (in Eigenvalue units)
100	100	71.1	التباين الكلي في الاستجابات Total raw variance in observations
22.2	22.5	23.1	التباين الذي فسره العامل الرئيسي (تقديرات نموذج راش) Raw variance explained by measures
11.0	11.1	07.9	التباين المفسر بواسطة الأفراد Raw variance explained by persons
21.2	21.4	15.2	التباين المفسر بواسطة المفردات Raw Variance explained by items
55.8	55.5	36.0	مجموع التباين غير المفسر Raw unexplained variance (total)

نلاحظ من خلال البيانات الواردة في الجدول رقم (02) أن نسبة التباين المفسر بتقديرات النموذج المستخدم (راش) بلغت 22.5%، والذي يمثل العامل الرئيسي والذي تشبعت عليه عدد من المفردات بلغ عددها (23) من أصل (35) مفردة، أن القيمتين الملاحظة (22.5) والمتوقعة (22.2) متقاربتين وهذا يعني أن النموذج جيد ومطابق (Fit) لأن النسبة الحقيقية قريبة من المتوقعة حيث إذا كان النموذج جيد فإن التباين سيكون (22.5) وهذه النسبة مقبولة حسب ما ورد في دليل برنامج winsteps. أن قيمة التباين الكلي غير المفسر بواسطة Raw unexplained variance أكبر أو

يساوي 36.0 وهذا مؤشر يعد مقبولا لأحادية البعد بعد التحقق من التباين غير المفسر الأول والثاني ونجد قيمته قد بلغت 55.5% ويسمى ببعد راش الأول، وهذا يعني عدم وجود أبعاد حقيقية أخرى تفسر التباين المتبقي.

1- حذف أفراد عينة البحث **Persona Fit** : لأخذ قرار حول حذف أو إبقاء أفراد عينة البحث، وتم حساب الخطأ المعياري في قياس قدرة أفراد العينة، الجدول رقم (3).

الجدول رقم (3) الانحراف المعياري والمتوسط الحسابي خاص بتقديرات قدرات الأفراد

الخطأ المعياري	القدرة	
0.30	0.99	المتوسط الحسابي
0.09	1.00	الانحراف المعياري

يلاحظ، من خلال الجدول رقم (3) المتوسط الحسابي، حيث كانت قيمته كما يلي: (0.99) وهي قريبة من الواحد الصحيح، في حين كانت قيم الانحراف المعياري لقدرة الأفراد قد بلغت (1.00) وهي قيمة مثالية يفترضها نموذج راش وهي بين (صفر، 1) على الترتيب، حيث تعتبر علامة قوية على تطابق بين أفراد العينة ونموذج راش، وقد تم استبعاد الأفراد الغير المقبولين في نموذج راش الذين تتراوح قدراتهم بين (2+) أو (2-) والجدول التالي يبين ذلك:

الجدول رقم (4): الأفراد غير المقبولين في نموذج راش

رقم الطالب	القدرة	الخطأ المعياري
30	-2.11	0.41
110	-1.20	0.34
16	-0.50	0.33
10	0.60	0.30
57	2.66	0.45
25	1.40	0.32

من الجدول رقم (4)، يتبين أن قيم قدرات الأفراد غير المقبولين في نموذج راش، أن الخطأ المعياري تراوح بين 0.30 و 0.41 وهي قيم غير متطابقة مع نموذج راش لذلك تم حذفها.

2 -حذف مفردات الاختبار: تم تطبيق الاختبار في صورته الأولية البالغة (35) فقرة، تبين أن (d4) فقرات لم تتحقق فيها شرط المطابقة للنموذج وهي الفقرات التالية (11، 18، 21، 34) وأما بعض الفقرات التي تجاوزت قيم متوسط المربعات لها الواحد صحيح، أو قل عن الصفر وهذا ليس مؤشرا على أن هذه الفقرات غير مطابقة لتوقعات النموذج (Linacre and Wright.1993) بما أن الفقرات لم تتجاوز قيم إحصائي المطابقة عن (2+) و (2) الجدول رقم (06): يوضح بيانات المفردات غير المطابقين للنموذج:

الجدول رقم (06): للمفردات غير المطابقة للنموذج راش

رقم المفردة	الصعوبة	الخطأ المعياري
11	-8.00	1.90
21	7.63	1.88
34	-7.98	1.85
18	7.99	1.86

من الجدول رقم (6) يتبين أن هذه المفردات خارج مجال القبول للنموذج لذلك تم حذفها من التحليل الإحصائي،
الخاتمة:

تعد النظرية الحديثة في القياس من النظريات المهمة في ميدان القياس النفسي، لا غني عنها في جميع العلوم، فهي ضرورية في أغلب المعارف والعلوم لأن القياس يؤدي إلى معرفة خصائص الشيء وما يميزه عن غيره، وبالتالي تأكد من صلاحيته ومناسبته لاستخدام ما، حيث تتمتع النظرية الحديثة بدرجة عالية من الموضوعية التي تحتاجها أغلب العلوم والعلوم النفسية والتربوية واحدة من تلك العلوم التي تحتاج إلى الموضوعية، وقد جاءت النظرية الحديثة في القياس النفسي، نتيجة للانتقادات التي وجهها عديد العلماء للنظرية التقليدية وغياب اختبارات ومقاييس موضوعية يمكن استخدامها في اختبارات القدرات والميول والاستعدادات والتي وفرت للباحثين النفسيين والتربويين إمكانية تحقيق الموضوعية في المقاييس السيكمومترية المختلفة. لذلك فقد جاء هذا المقال ليكون مدخلاً للباحثين المختصين بالقياس النفسي والتربوي.

- الاقتراحات: رغبتنا منا في إثراء البحث العلمي قمنا باقتراح ما:
- إقامة دورات تدريبية وإرشادية لتوضيح أهمية بناء الاختبارات التحصيلية وفق النظرية الحديثة للقياس.
- تطوير أساليب ووسائل التعليم بحيث تواكب التطورات الحاصلة في مجال البحث العلمي ومستوى كل طالب حتى يتسنى للجميع التعلم.
- تدريب الأساتذة على كيفية بناء اختبارات موضوعية.
- ضرورة تأهيل الأساتذة مهنيًا وبصورة مستمرة لزيادة فاعليتهم في العمل من جهة وتماشياً مع التطورات الحاصلة في المجال التربوي.
- عقد دورات وإعداد برامج خاصة وتكوين الأساتذة في مجال بناء الاختبارات خاصة وفق النظرية الحديثة

قائمة المراجع:

- أيلالا دي (2017): النظرية والتطبيق في نظرية استجابة الفقرة. ترجمة الكيلاني، عبد الله؛ البرصان، إسماعيل، الرياض: جامعة الملك سعود.
- بني عطاء زايد والشريفين، نضال، كمال. (2012). أثر اختلاف شكل توزيع القدرة على معلم الفقرة ودالة المعلومات للاختبار. المجلة الأردنية في العلوم التربوية، 8 (2)
- الشمراني محمد (2014): استخدام نظرية استجابة الفقرة والنظرية التقليدية في تقدير الخصائص الإحصائية السيكمومترية لاختبار مهارات التفكير والتعلم لطلبة السنة التحضيرية الجامعية. التربية جامعة الأزهر، مصر. م 2. ع 158.
- شحنة عبد المولى (1999): تقويم بناء الاختبارات المرجعة إلى المحك / المعيار في ضوء نظرية الاستجابة للمفردة والنظرية التقليدية، رسالة دكتوراه، جامعة عين شمس، كلية التربية
- العيسوي، عبد الرحمان محمد (2003م)، الاختبارات والمقاييس النفسية والعقلية، منشأة المعارف بالإسكندرية، مصر.
- علام، صلاح الدين (1986): أساسياته وتطبيقاته وتطورات معاصرة في القياس النفسي والتقويم التربوي. الكويت: مطابع التجارية
- علام، صلاح الدين (2006): القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة، ط2. القاهرة: دار الفكر العربي.
- عماد عبد المسيح (1991): استخدام نموذج راش اللوغاريتمي أحادي البارامتر في تحليل مفردات الاختبارات المعرفية مرجعية المعيار ثنائية القطب (دراسة تجريبية)، جامعة المنيا، كلية التربية، مجلة البحث في التربية وعلم النفس، العدد (4)، أبريل 1991.
- محاسنة إبراهيم محمد. (2013). القياس النفسي في ضل النظرية التقليدية والنظرية الحديثة، الأردن: دار جرير للنشر والتوزيع.
- مقدم عبد الحفيظ (2003): الإحصاء والقياس النفسي والتربوي، بن عكنون، الجزائر: ديوان المطبوعات الجامعية، ط2.
- كتفي عبد الحق (2020): بناء اختبار تحصيلي في مادة علوم الطبيعة والحياة مستوى السنة الأولى من التعليم المتوسط وفق نموذج راش، أطروحة دكتوراه، كلية العلوم الإنسانية والاجتماعية، جامعة الجزائر 2، الجزائر.
- Linacre, J. M. (1998). Detecting multidimensionality: Which residual data-type works best? Journal of Outcome Measurement (2(3)), 266-283.