كلية التربية
كلية معتمدة من الهيئة القومية لضمان جودة التعليم
إدارة: البحوث والنشر العلمي (المجلة العلمية)
=======

أثر وجود أداء تفاضلي في الفقرات المشتركة على دقة المعادلة العمودية لاختبار رافن للذكاء في ضوء نموذج فيشر للاستجابة للمفردة

إعـــــــــداد

أ.د / عماد أحمد حسن علي
أستاذ ورئيس قسم علم النفس التربوي
كلية التربية – جامعة أسيوط

د / سميرة محمد أحمد حسن
مدرس علم النفس التربوي
كلية التربية – جامعة أسيوط

م م/ علي صلاح عبد المحسن حسن
مدرس مساعد بقسم علم النفس التربوي
كلية التربية – جامعة أسيوط

﴿ المجلد الثاني والثلاثين – العدد الرابع – جزء ثاني – أكتوبر ٢٠١٦م ﴾
http://www.aun.edu.eg/faculty_education/arabic

مقدمة البحث:

اهتم علماء النفس منذ أن وجدت حركة القياس النفسي، بتحقيق صدق وثبات الاختبارات والمقاييس النفسية، سعياً منهم لتحقيق أعلى درجة من الموضوعية في هذه الأدوات، عند استخدامها في عملية القياس. ووفق نظرية القياس التقليدية Classical Theory يمكن التعبير عن قدرة الفرد من خلال الدرجة الحقيقية والتى تتضح من خلال أدائه على الاختبار، وبناءً عليه فإنه سيتغير وضع قدرة الفرد حسب تغير مستوى الاختبار. إن الاختبار والبنود تتغير خصائصها بتغير خصائص الأفراد، كما أن خصائص الأفراد تتغير بتغير خصائص الاختبار من حيث السهولة والصعوبة.

إن معادلة الاختبارات بواسطة نظرية الاستجابة للمفردة تحل الكثير من المشكلات التي عجزت عنها النظرية التقليدية بشرط أن يكون النموذج المستخدم في النظرية الحديثة مطابقاً للبيانات المعدة للمعادلة(Kolen, M., 1991, 8).

ويعد نموذج فيشر اللوجستي تطبيقاً أوسع لنموذج راش مع وجود بعض القيود الخطية على معلمات البنود، فهذا النموذج يعطي اهتماما خاصا باحتمالية التقدير القصوى، ويوفر هذا النموذج أساساً لاختبار الفرضيات المتعلقة بصعوبة البند، كما تم عمل دراسات مسحية للعديد من الدراسات التجريبية في تحليل البند، تحيز البند، بناء الاختبار، وساعدت هذه الدراسات في تقديم نوع جديد من تطبيقات الاستجابة للمفردة إلى عمليات ديناميكية، وقد اقترح أيضا فيشر نموذجا خاصا من LLTM لقياس التغيير في فضاء متعدد الأبعاد(Fischer, H., 1982, 380).

ولعل وضع نماذج للاختبار عبر مستويات مختلفة من القدرة على مقياس واحد يساعدنا في مقارنة درجات المفحوصين بشكل مباشر ويسهل عملية تطبيق الاختبار، إضافة إلى توفير الجهد والوقت، كما أن وجود فقرات ذات أداء تفاضلي ضمن نماذج الاختبار قد يؤثر في دقة معادلة درجات الاختبارات، لذلك من المهم تحديد هذه الفقرات والكشف عن أثرها في عملية المعادلة (Hidalgo, M. & Lopez, J., 2002, 32).

مشكلة البحث:

إن توافر صور متكافئة أو بديلة للاختبار تسهل من عملية إعادة اختبار المفحوصين عندما يكون ذلك ضرورياً أو مرغوبا فيه للحفاظ على سرية الاختبار، ولمنع أثر الاختلاط بين الأفراد الذين طبق عليهم الاختبار والأفراد الذين سيطبق عليهم الاختبار، وكذلك لمنع أثر تذكر الإجابة على فقرات الاختبار إذا ما أعيد تطبيقه مرة أخرى(على خطاب، ٢٠٠١، ٣٤٣).

وهناك العديد من الدراسات الأجنبية والعربية التي أجريت حول استخدام النظرية الحديثة في القياس (نظرية الاستجابة للمفردة) وعملية معادلة درجات الاختبار، من هذه الدراسات ما أكدت على أثر أداء الفقرات التفاضلي في دقة معادلة الاختبار في ضوء نماذج الاستجابة للمفردة، وأخرى على النقيض من ذلك أكدت على عدم تأثير الفقرات التفاضلية في دقة معادلة الاختبار في ضوء نماذج الاستجابة للمفردة.

ومما تقدم فإن مشكلة البحث تثير الأسئلة التالية:

١- هل تحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون في ظل الاحتفاظ بالفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

٢- هل تحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون عند حذف الفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

أهداف البحث:

يهدف البحث الحالي إلى الكشف عن الفقرات ذات الأداء التفاضلي لمتغير العمر في دقة المعادلة العمودية لاختبار رافن للمصفوفات المتتابعة في ضوء نموذج فيشر للاستجابة للمفردة، وذلك من خلال:

١. الكشف عن الفقرات ذات الأداء التفاضلي لمتغير العمر لصورتي اختبار رافن للمصفوفات المتتابعة الملون ومن ثم إجراء معادلة عمودية لدرجات القدرة للمفحوصين عبر هذين المستويين في ضوء نموذج فيشر، حيث تشير بعض الدراسات إلى أن اختبارات الاستعداد واختبارات القدرة العقلية غالباً ما تبدي فقراتها أداءً تفاضلياً وفقاً للنوع (Harris, A.& Carlton, S., 1995, 140).

٢. التحقق من فعالية المعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون في ظل وجود الفقرات ذات الأداء التفاضلي لمتغير العمر وفقاً لنموذج فيشر.

٣. التحقق من فعالية المعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون عند حذف الفقرات ذات الأداء التفاضلي لمتغير العمر وفقاً لنموذج فيشر.

أهمية البحث:

تتضح أهمية البحث الحالي من خلال ما يلي:

- ✓ تبرز أهمية هذا البحث من فائدة وضع مقياس موحد لمستويات مختلفة للقدرة، والذي بدوره يوفر الوقت والجهد في عملية القياس.
- ✓ مقارنة الأفراد ذوي القدرات المختلفة بشكل مباشر، وملاحظة تطور هذه القدرات واستخدام درجات الاختبارات بشكل متبادل من خلال المعادلة العمودية.
- ✓ السعي للتحقق من أثر الفقرات ذات الأداء التفاضلي على دقة المعادلة للاختبارات باستخدام نموذج فيشر.
- ✓ المساهمة في مجال بناء وتطبيق الاختبارات وتفسيرها وكيفية التعامل مع وجود أداء تفاضلي عند مقارنة قدرات مختلفة خلال المعادلة العمودية للاختبارات.
- ✓ المساهمة في توسيع الفهم حول المعادلة العمودية لدرجات الاختبار باستخدام نموذج فيشر وذلك لمواكبة الاتجاهات الحديثة قي القياس والتقويم.

منهج البحث:

يدور البحث الحالي حول الكشف عن أثر أداء الفقرات التفاضلي العمر في دقة معادلة الاختبار في ضوء نموذج فيشر للاستجابة للمفردة، وفي سبيل ذلك تم استخدام صورتي اختبار رافن الملون للكبار من إعداد (عماد حسن، ٢٠١٦)، ثم استخدام البرامج الإحصائية للكشف عن الفقرات التفاضلية، ثم إجراء عملية المعادلة وفقاً لنموذج فيشر للاستجابة للمفردة، ثم التحقق من دقة المعادلة في حالة حذف الفقرات ذات الأداء التفاضلي للاختبار أو الإبقاء عليها؛ لذلك تم إتباع **المنهج الوصفي**.

حدود البحث:

(أ) الحدود الزمانية:

تم تطبيق أدوات البحث في الفصل الدراسي الثاني من العام الجامعي ٢٠١٥/ ٢٠١٦ بكلية التربية جامعة أسيوط.

(ب) الحدود المكانية:

تم تطبيق أدوات البحث في فصول ومدرجات ومعامل كلية التربية جامعة أسيوط.

(ج) الحدود الموضوعية:

تم استخدام نموذج فيشر كأحد نماذج نظرية الاستجابة للمفردة مع استخدام معيار هانزل مانتل للحكم على أداء الفقرات التفاضلية من خلال برنامج DIFAS 5.0، كما تم استخدام معامل الصدق التقاطعي للحكم على فعالية المعادلة العمودية.

(د) الحدود البشرية:

اقتصرت هذه الدراسة على عينة من طلاب وطالبات كلية التربية بالفرقتين الأولى والرابعة وجدول ١ يوضح أعداد الذكور والإناث بالفرق الأولى والرابعة والدبلوم العام نظام العام الواحد بكلية التربية.

جدول ١

أعداد الذكور والإناث بالفرق

الأولى والرابعة وطلاب الدبلوم العام بكلية التربية جامعة أسيوط

الفرقة / النوع	ذكور	إناث	المجموع
الأولى	٢٨٠	٣٢٠	٦٠٠
الرابعة	١٩٥	٢٧٠	٤٦٥
دبلوم عام نظام العام الواحد	٤٣٥	٥٠٠	٩٣٥
المجموع	٩١٠	١٠٩٠	٢٠٠٠

أدوات البحث:

اختبار المصفوفات المتتابعة الملون لرافن (تعديل وتقنين عماد أحمد حسن علي)

المفاهيم الإجرائية:

يتناول البحث الحالي المصطلحات الإجرائية التالية:

١- معادلة درجات الاختبار:

وهو تحويل نظام الدرجات الخاص بإحدى نموذجي الاختبار إلى نظام درجات خاص بالنموذج الآخر، بحيث تصبح الدرجات متكافئه (Chong Ho Yu, A., 2005, 3).

٢ – المعادلة العمودية:

وهي تحويل الدرجات على نماذج للاختبار صممت لمجتمعات عند مستويات قدرة مختلفة إلى تدريج مشترك(Stephen, H., 2007, 10; Baker, F., 1984, 261).

٣-الأداء التفاضلي للفقرة:

وهو إبداء الفقرة أداءاً تفاضلياً وفقاً لمعيار مانتل هانزل، بحيث تكون الفقرة ذات أداء تفاضلي إذا كانت قيمة كا تربيع أكبر من ٣.٨٤ عند مستوى دلالة ٠.٠٥، أو أكبر من ٦.٦٣ عند مستوى دلالة ٠.٠١ (Zwick, R.,& Mazzeo, J., 1997, 322).

٤ – اختبار المصفوفات المتتابعة لجون رافن:

تعد المصفوفات المتتابعة الملونة أحد اختبارات المصفوفات الثلاثة التى أعدها عالم النفس الإنجليزي J. Raven ، وهي من اختبارات الذكاء غير اللفظى وهى خالية من تأثير الثقافة إلى حد كبير وتعتمد أساساً على التطبيق الجمعى ويمكن أن تطبق فردياً فى ظروف معينة، ويتوافر منها حالياً ثلاثة مستويات: المصفوفات المتتابعة العادية، المصفوفات المتتابعة الملونة، المصفوفات المتتابعة المتقدمة (عماد حسن، ٢٠١٦، ٣).

٥-نموذج فيشر:

هو أحد النماذج الديناميكية لنظرية الاستجابة للمفردة، وقد اقترحه بوك في كتاباته عام ١٩٧٦، ويستخدم هذا النموذج في قياس منحنيات البارامتر(صلاح الدين عـلام ، ٢٠٠٥ ،٨٤؛ Purya, B.& Klaus, K., 2015, 4).

المفاهيم الأساسية للبحث

– نظرية الاستجابة للمفردة(IRT):

تعد نظرية الاستجابة للمفردة طريقة جديدة لنمذجة البيانات، بمعنى أنها تحاول نمذجة العلاقة بين قدرة الفرد واحتمال استجابته سواء أكان صواباً، أم خطأً، ويعتمد صدق طرق نظرية الاستجابة للمفردة في جزء كبير على البيانات المستمدة من أداء الفرد((Harwell, S., 1997, 266.

وتفترض هذه النظرية أنه يمكن تفسير الأداء الملاحظ للأفراد على اختبار ما بسمة أو قدرة لدى هؤلاء الأفراد. وبطبيعة الحال فإن هذه السمة أو القدرة لا يمكن التنبؤ بها مباشرة، ومن هنا جاءت كلمة سمة كامنة Latent trait. ويذكر صلاح الدين علام (٢٠٠٢، ٢١) أنه عندما يود عالم النفس قياس إحدى السمات العقلية، ولتكن سمة الذكاء، فإنه لا يلاحظ ذكاء الطفل – لأن الذكاء مفهوم مجرد–، ولكنه يلاحظ سلوك الطفل في مواقف مختلفة تتطلب الذكاء، أي أن السمات العقلية تعد بمثابة تكوينات يفترض أنها تعكس مجموعة مترابطة من السلوك الذي يمكن ملاحظته وتسجيله في مواقف اختبارية مقننة.

ولكي يمكن تحديد أي من نماذج الاستجابة للمفردة هو الذي يجب استخدامه، هناك ثلاث محكات لخصها كلا من (Wiberg, M., 2004, 4؛ وصلاح الدين علام، ٢٠٠٥، ٨٤) وهي:

المحك الأول: تحديد افتراضات النموذج.

المحك الثاني: خصائص النموذج المتوقعة.

المحك الثالث: توقعات النموذج لنتائج الاختبار الفعلية.

وتلك المحكات الثلاثة هي ملخص للظاهرة التي يحتمل أن يستخدمها معدي الاختبار، ويمكن أن تنقسم تلك المحكات إلى محكات فرعية أخرى، ولقد اقترح هامبلتون (**في**: ميمي إسماعيل، ٢٠٠٧، ٢٣) أنه يجب على الاختبار أن يلائم أكثر من نموذج واحد للبيانات، ومن ثم يمكن مقارنة النماذج طبقاً للمحكات الثلاثة.

المحك الأول: افتراضات النموذج:

١ – أحادية البعد:

تشير إلى حقيقة أن الاختبار يجب أن يقيس قدرة كامنة واحدة والمفردات التي تكون أحادية البعد ربما تعكس سمات عديدة ويطبق هذا الشرط في معظم نماذج نظرية الاستجابة للمفردة(Yen, M. &Edwardson, S., 1999, 236).

٢ – تساوي التمييز:

يمكن لتساوي التمييز أن يفحص العلاقة بين المفردة (i) والدرجة الكلية للاختبار. فإذا كانت المفردات ذات تمييز متساو فمن الأفضل أن نستخدم النموذج اللوغاريتمي أحادي البارامتر (نموذج راش)، أما إذا لم تكن المفردات ذات تمييز متساو فمن الأفضل أن نستخدم النموذج اللوغاريتمي ثنائي البارامتر وثلاثي البارامتر.

٣ – احتمالية تخمين الإجابة الصحيحة:

هناك طريقة لفحص ما إذا كان يحدث تخمين وذلك عن طريق ملاحظة إجابة الممتحنين ذوي القدرات المنخفضة على المفردات الأكثر صعوبة، أي أن الممتحنين ذوي القدرات المنخفضة الذين يجيبون على المفردات الأكثر صعوبة بشكل صحيح، فإن ذلك يسمح لبارامتر التخمين أن يظهر، وعلى ذلك فإن النموذج اللوغاريتمي ثلاثي البارامتر ملاءمة من النموذج الأحادي والثنائي في إظهار بارامتر التخمين(Wiberg, M., 2004, 5).

المحك الثاني: خصائص النموذج المتوقعة:

ويلخصها (Husen, T. &Postlethwaite, T., 1995, 351) في:

١- إن تقدير بارامتر قدرة الفرد لا يعتمد على تقدير بارامتر صعوبة المفردة.

٢- إن تقدير بارامتر المفردة لا يعتمد على استخدام عينات مختلفة من الأفراد سواء أكانت عينات ذات قدرات مرتفعة، أم منخفضة.

المحك الثالث: توقعات النموذج لنتائج الاختبار الفعلية:

يتم معرفة توقعات النموذج لنتائج الاختبار الفعلية عن طريق:

١- إجراء مقارنة بين المنحنيات المميزة لكل مفردة من المفردات الأخرى.

٢- استخدام رسوم بيانية لتقسيمات الدرجة المتوقعة والملاحظة أو استخدام اختبارات (كا٢).

- نموذج فيشرالديناميكي: (LLTM)

تهتم هذه النماذج بقياس التغير الذي يحدث في السمات الإنسانية عبر الزمن. فالبعض يرى أن هذا التغير يُعد عملية تدريجية، بينما يرى البعض الآخر أنه عملية تحول من حالة إلى أخرى.

ففي الحالة الأولى يكون المتغير الذي ينطوي عليه التغير متصلاً، بينما في الحالة الثانية يكون المتغير منفصلاً، غير أن الحالة الأولى هي التي نالت الاهتمام في النماذج الديناميكية.

فقد اقترح فيشر نموذجاً يمكن استخدامه في قياس منحنيات التحصيل، ويشمل النموذج على عدة مكونات استناداً إلى النموذج ثنائي البارامتر، وتتطلب عملية التعيير باستخدام هذا النموذج توافر عينة متكافئة من المختبرين الذين حصلوا على درجات متفاوتة من التعليم، وقد حاول بوك في نموذجه أن يسمح بتعيين الفرد في نمط معين، ومن ثم تقدير بارامترات هذا النموذج الفردي(صلاح الدين علام، ٢٠٠٥، ٨٤).

ويعد نموذج فيشر اللوجستي تطبيقاً أوسع لنموذج راش مع وجود بعض القيود الخطية على معلمات البنود، فهذا النموذج يعطي اهتماما خاصا باحتمالية التقدير القصوى، ويوفر هذا النموذج أساساً لاختبار الفرضيات المتعلقة بصعوبة البند، كما تم عمل دراسات مسحية للعديد من الدراسات التجريبية في تحليل البند، تحيز البند، بناء الاختبار، وساعدت هذه الدراسات في تقديم نوع جديد من تطبيقات الاستجابة الطارئة إلى عمليات ديناميكية، وقد اقترح أيضا فيشر نموذجا خاصا من LLTM لقياس التغيير في فضاء متعدد الأبعاد(Fischer, H., 1982, 380).

ولنفترض أن لدينا مجموعة من البيانات تتناسب مع افتراضات نموذج راش، فيمكن أن يكون اختبار ملاءمة النموذج المعرفي المحدد بطريقتين: الأولى بطريقة مشابهة تناسب اختبار نموذج راش، والثانية يمكن استخدام نموذج بياني لاختيار معلمات البند المقدرة بنموذج راش ومقارنتها بنموذج فيشر، فنموذج فيشر يساعد على الحد من ضعف ارتباط العناصر ببعضها البعض. ويمكن استخدامه لتعديل أو إعادة صياغة هيكل العمليات الابتدائية، أو للمساعدة في تحديد المشاكل التي تعترض بناء النموذج (Philipp, S., 2008, 347).

إن نموذج LLTM أحدث طفرة عظيمة في علم النفس بصفة عامة والقياس النفسي بصفة خاصة، حيث أنه يسد الفجوة بين العلم والقياس النفسي المعرفي، ويساعد في الكشف عن العمليات العقلية التي تنتج التباين الخاص بالثبات ويساعد في تقديم استراتيجيات معرفية لحل الممتحنين، والتي بدورها تؤدي إلى تنمية العناصر الأكثر كفاءة. وهذا يوفر طريقة منهجية في الحكم على مستوى البند (Baghaei, H. & Kubinger, M., 2015, 9).

وأشار فيشر (Fischer, H, 1973, 360) إلى قوة هذا النموذج في اختبار الفرضيات فهو نموذج اختبار تجريبى يُمكن للباحثين من وضع فرضيات عن البند، ووضع النظريات النفسية الموضوعية، وتحديد العمليات الادراكية التى تحتاجها لحل البنود، وبناء صحة البنود التى يمكن اثباتها لاختبار العمليات الادراكية المحددة.

ويوجد العديد من الطرق التي من خلالها يمكن تقدير المعلمات، من أهمها بروزاً طريقة أقصى احتمال مشروط (Cb4L)، ويفرض نموذج فيشر على معلمة الصعوبة بعض القيود الخطية، ونموذج فيشر يتيح للباحثين تجريبيا اختبار الفرضيات حول عمليات حل البند ووضع النظريات النفسية الموضوعية.

وتلخيصاً لما تقدم عن نموذج فيشر، فإن جميع البنود تعد جزءاً شائعاً من نموذج راش (خاصة في التصميم التجريبي ذي المجموعة الواحدة)، فأينما نجد محتوى نفسي للبنود يربط بينها من خلال بارامترات تزداد إمكانية تطبيق نموذج فيشر في القياس النفسي وذلك عند التعامل مع القياسات القبلية والبعدية، فنموذج فيشر امتداد لنموذج راش (Baghaei, H. & Kubinger, M., 2015, 1; Pono, I, 2002, 24)

ويشير (Baghaei, H. & Kubinger, M., 2015, 3) إلى أنه من الشروط الأساسية قبل تطبيق نموذج فيشر يجب تطبيق نموذج راش أولاً، فقبل إجراء نموذج فيشر يمكن حذف الفقرة أو البند الذي لا يتناسب مع افتراضات نموذج راش، وبهذه الطريقة نتعرف على الاشخاص والفقرات التي يمكن إزالتها مع وجود استراتيجيات بديلة للحل.

وتؤكد كلا من (Purya, B.& ; Susan E., W., & Lisa M. S., 1981, 385 Klaus, K., 2015, 6) على تشابه نموذج راش مع نموذج فيشر في اعتماد النموذجين على بارامتر الصعوبة فقط، إلا أنه في نموذج فيشر تتنوع البارامترات التي تشارك في عمليات تجهيز الأداء.

الصيغة الرياضية لنموذج فيشر:

يعتمد النموذج الرياضي الديناميكي لفيشر على الافتراضات التي أسسها راش في معادلته الرياضية، ويمكن صياغة المعادلة كما سبق في نموذج راش على النحو التالي:

عند وضع صيغة رياضية لاستجابة الفرد (v) للمفردة (i)، يجب البحث عن المتغيرات الأساسية المؤثرة في الاستجابة، وهي قدرة الفرد βv وصعوبة المفردة δi، ويكون النموذج الوسيط الذي يمكن أن يوضح هذين المتغيرين في الاستجابة الملاحظة (Xvi) هو الفرق بين هذين المعلمين، ويعد احتمال حدوث الاستجابة الصواب Xvi= 1 دالة لهذا الفرق، وهذا يتمثل في الدالة الآتية:

$$\mathbf{Pvi = f(\beta v - \delta i)}$$

حيث Pvi احتمال نجاح الفرد (v) على المفردة (i)، أي احتمال حدوث الاستجابة الصواب، ويلاحظ من المعادلة السابقة أن احتمال حدوث الإجابة الصواب Pvi ينحصر بين القيمتين (صفر، ١) في حين أن الفرق (βv – δi) يمكن أن يكون أي عدد حقيقي من −∞ إلى +∞. لذا ينبغي اختيار نموذج احتمال يعتمد على الفرق (βv – δi) ويجعله محصوراً بين القيمتين (صفر، ١)، وليكن احتمال التوزيع الأسي (اللوغاريتمي مثلاً).

ويمكن صياغة المعادلة السابقة في الصورة التالية:

$$Pvi = \frac{e^{\beta v - \delta i}}{1 + e^{\beta v - \delta i}}$$

وهي نفس صيغة نموذج راش، وتشتمل على بارامتر واحد فقط، وهو بارامتر صعوبة المفردات لتمثيل الفروق بين المفردات(صلاح الدين علام، ٢٠٠٥، ٦٩؛ Pono, I, 2002, 23)..

- **معادلة الاختبارات (Test Equating):**

تحتوى أدبيات البحوث على العديد من التعريفات لمعادلة الاختبار، ولا يوجد تعريف محدد لمعادلة الاختبار، وعلميا لا يهتم الباحثون بالتعريفات المتعددة الموجودة (Roos, L., 2002, 10).

فمعادلة الاختبار هي عملية الحصول على درجات متكافئة لأداتين تقيسان السمة نفسها(راشد الدوسري، ٢٠٠١، ١١٢).

وعموما فإن معادلة الاختبار هي إجراء يسمح لمفردات صورتي اختبار أن يتم وضعها على نفس المقياس (same scale) (Roos, L., 2002, 5).

وتعرف معادلة الأختبار بأنها تلك العملية التي يتم فيها اشتقاق (Deriving) دالة رسم الدرجة (Function mapping score) على صورة بديلة للاختبار على مقياس الصورة الرباط (المشتركة) (anchor form) ونتيجة لذلك بعد المعادلة، تكون أيه درجة للمقياس لها نفس المعني بغض النظر عن أية صورة من صور الاختبار تم تطبيقه (Haertel, E., 2004, 1).

ويقصد بعملية معادلة درجات اختبارين (test Score Equationg) تلك الإجراءات والعمليات الإحصائية التي تتم بغرض تحويل نظام وحدات القياس الخاص بإحدى الصورتين إلى نظام وحدات القياس الخاص بالصورة الأخرى، بحيث تصبح القياسات المستمدة من درجات كل من الصورتين متكافئة بعد إجراء هذا التحويل(إسماعيل الوليلي، ٢٠٠٥، ١٠٨).

أهمية معادلة الاختبارات:

هناك الكثير من المشكلات التي تعانيها نظم التعليم في الكثير من دول العالم، ومن أهم هذه المشكلات عملية بناء الامتحانات الفصلية، والامتحانات الملحقة للشهادات العامة بالمرحلتين الإعدادية والثانوية، إذ يقوم المتخصصون في إدارة المناهج ببناء الكثير من الامتحانات لنهاية كل فصل دراسي، ويتطلب ذلك وقتا كبيراً، وجهداً أكبر، ويضطرون إلى بناء اختبارات ملحقة للطلبة الذين لم يتمكنوا من اجتياز الامتحانات الفصلية بنجاح، ولا يوجد دليل علمي على قدرة هذه الامتحانات على تمثيل المحتوى المطلوب لتأكيد جانب الصدق فيها، كما لا يمكن اعتبارها متكافئة فيما بينها بسبب عدم تحليلها للحصول على الخصائص السيكومترية الضرورية لعملية التكافؤ، والمعادلة، كالوسط الحسابي، والانحراف المعياري، ومعامل الصعوبة، ومعامل التمييز، وكثيراً ما يشتكي الطلبة في كل عام دراسي من تفاوت الامتحانات من امتحان إلى أخر، وعدم العدل في قياس سمة التحصيل قياساً علمياً سليماً ومما يزيد من خطورة المشكلة أن نتائج الطلبة في الامتحانات العامة تترتب عليها قرارات متعلقة بالانتقال إلى مراحل دراسية أعلى، أو الالتحاق بالجامعات، أو بسوق العمل(راشد الدوسري، ٢٠٠١، ١٠).

طرق معادلة الاختبارات وفقا لنظرية الاستجابة للمفردة:

تعتمد هذه الطريقة على تحويل التدريج المستقل لمفردات اختبارين مختلفين حسب درجة صعوبة مفردات كل منهما، وكذلك قدرات التلاميذ الذين طبق عليهم الاختباران إلى تدريج واحد مشترك له صفر مشترك، وذلك بعد التحق من توافر افتراضات نظرية الاستجابة للمفردة في البيانات الناتجة عن تطبيق الاختبارين(إسماعيل الوليلي، ٢٠٠٥، ١٠٧).

إن وضع تقديرات المعلمات للمفردة المشتقة من الصور المختلفة للاختبار على مقياس مشترك، يعد إحدى خطوات المعادلة وفق نظرية الاستجابة للمفردة، وتوجد ثلاثة أنواع رئيسية للمعايرة وإجراءات الربط وذلك لوضع تقديرات معلمة المفردة المشتقة من الصور المختلفة للاختبار على مقياس مشترك في حالة المعادلة وفق تصميم المفردات المشتركة وهي :

- **وصل المعايرة(التدريج)المنفصلة**.The linking separate calibration(LSC)

- **المعايرة (التدريج) المتلازمة**. The concurrent calibration

- **معايرة (تدريج) المعلمة المثبتة**. The fixed parameter calibration

ومن أشهر الطرق المستخدمة في حساب معاملات التحويل (Transformation Coefficients) (A, B) في حالة وصل المعايرة (التدريج) المنفصلة (Zhonghua, Z., 2010, 18) الطرق التالية:

- **الطرق الآنية / اللحظية / التامة (the moment methods) وهي:**

- طريقة المتوسط / المتوسط (Mean / Mean method)

- طريقة المتوسط / الانحراف المعياري (Mean / sigma method)

أسئلة البحث:

من خلال ما تم عرضه من الإطار النظري حول متغيرات البحث ونتائج الدراسات السابقة يمكن عرض أسئلة البحث فيما يلي:

١. هل تتحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون في ظل الاحتفاظ بالفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

٢. هل تتحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون عند حذف الفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

إجراءات البحث:

أولاً– عينة البحث: اقتصر هذا البحث على عينة من طلاب وطالبات الفرقة الأولى والرابعة وطلاب وطالبات الدبلوم العامة بكلية التربية، للعام الجامعي ٢٠١٥/ ٢٠١٦م.

ثانياً– الخطوات الإجرائية للبحث:

تتلخص خُطوات إجراء البحث في الخُطوات الآتية:

١– استعان الباحث باختبار جون رافن المصور للذكاء من إعداد جون رافن وتعديل وتقنين عماد حسن(٢٠١٦).

٢– تم تقسيم الاختبار إلى صورتين كل صورة تتكون من ٢٤ فقرة، بينهما ١٢ فقرة مشتركة.

٣– تم تطبيق الصورتين على طلاب وطالبات كلية التربية، وذلك بهدف تقنين الصورتين على العينة الاستطلاعية.

٤– تم تطبيق صورتي الاختبار النهائيتين على العينة الأساسية، وأصبح لكل طالب وطالبة درجة كلية على إحدى الصورتين.

٥– تم تطبيق برنامج **DIFAS 5.0** على صورتي الاختبارين للكشف عن الفقرات التفاضلية.

٦– تم استخدام الأساليب الاحصائية المناسبة لتحليل البيانات، بعد أن تم التأكد من توافر شروط نظرية الاستجابة للمفردة (IRT).

نتائج البحث وتفسيرها:

وينص السؤال الأول على " هل تحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون في ظل الاحتفاظ بالفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

وللإجابة عن هذا السؤال: تم أولاً الكشف عن الفقرات ذات الأداء التفاضلي في صورتي اختبار رافن، وذلك من خلال استخدام برنامج DIFAS4.0، بأن يكون طلاب الدراسات العليا مجموعة مرجعية وطلاب الفرقتين الأولى والرابعة مجموعة مستهدفة، أو العكس – **حيث أظهر برنامج DIFAS4.0 نفس الفقرات بالترتيب عند وضع طلاب الدراسات العليا مجموعة مستهدفة وطلاب الفرقتين الأولى والرابعة مجموعة مرجعية**– وأظهرت نتائج البرنامج، وجود ١٨ فقرة من فقرات الاختبارين ذات أداءاً تفاضلياً، ثماني فقرات بالصورة الاختبارية الأولى، وعشر فقرات بالصورة الاختبارية الثانية.

ثم قام الباحث بالتحقق من تساوي متوسطي درجات طلاب الدراسات العليا وطلاب الفرقتين الأولى والرابعة على صورتي الاختبار كمعيار لتكافؤ الصورتين، ولذلك تم استخدام اختبار (ت) لعينتين مستقلتين من خلال البرنامج الاحصائي Spss، وجدول ٢ يوضح ذلك.

جدول ٢

المتوسط الحسابي والانحراف المعياري وقيمة ت للفروق بين درجات الطلاب على صورتي اختبار رافن

المتغير	الصورة الاختبارية الأولى (ن١= ١٠٦٥)		الصورة الاختبارية الثانية (ن٢= ٩٣٥)		قيمة ت	الدلالة
	م	ع	م	ع		
الدرجة الكلية للاختبار	١٩.٠٥٤٧	٣.٦٣٣١٨	١٨.٢٨٠٢	٤.٦٣٠٨٣	١٠.١٢	دال عند ٠.٠١

يتضح من جدول ٢ وجود فروق ذات دلالة إحصائية بين متوسطي درجات طلاب الفرقتين الأولى والرابعة على صورتي الاختبار، وذلك لصالح متوسط درجات الطلاب في الصورة الأولى، وبذلك فإن صورتي اختبار رافن غير متكافئتين وفقاً لمعيار التكافؤ المستخدم وهو تساوي متوسطي درجات الطلاب على الصورتين الاختباريتين، وعدم التكافؤ ينتج عنه عدم القدرة على المقارنة بين مستوى ذكاء الطلاب الذين أجابوا على صور مختلفة من الاختبار، وبذلك تكون صورتا الاختبار في حاجة إلى إجراء عملية معادلة للدرجات لإتخاذ قرارات صحيحة بشأن مستوى ذكاء الطلاب على الصورتين.

وقبل إجراء عملية المعادلة بين درجات الطلاب على صورتي الاختبار، تم التحقق من توافر شروط نظرية الاستجابة للمفردة(أحادية البعد– الاستقلال الموضعي– التحرر من السرعة– توافر خاصية المنحنى المميز للمفردة ICC) وذلك لكل صورة من صورتي الاختبار، كما ورد عند الاجابه عن السؤال الأول.

وبعد أن تم التأكد من توافر شروط نظرية الاستجابة للمفردة، وكذلك جودة المفردات المشتركة في صورتي الاختبار، سيتم الإجابة على السؤال الذي ينص على " ما مدى تحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون في ظل الاحتفاظ بالفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

وللإجابة عن هذا السؤال، تم تحديد طرق المعادلة المستخدمة في ضوء نموذج فيشر لنظرية الاستجابة للمفردة، وهي كالتالي:

أ- معادلة المتوسط/ المتوسط.

ب- معادلة المتوسط/ الانحراف المعياري.

ولحساب المعادلة بطريقة المتوسط/ المتوسط وكذلك طريقة المتوسط/ الانحراف المعياري، يستلزم ذلك حساب بارامتر المفردة(الصعوبة) لأن النموذج المستخدم هو نموذج فيشر، حيث يكون فيه بارامتر التمييز للمفردات قيمة ثابتة وتساوي الواحد الصحيح أو قيمة قريبة منها ويتم تثبيتها في البرنامج المستخدم لجميع المفردات، وكذلك بارامتر التخمين مساوياً للصفر، ولتقدير بارامترات الصعوبة الخاصة بمفردات صورتي الاختبار فقد تم استخدام برنامج IRTEQ وقد تم ضبط البرنامج بحيث يكون معامل التدريج D = 1.7 أي في حالة القياس الطبيعي، وقد تم الحصول على بارامترات صعوبة المفردات لكل صورة من صورتي الاختبار، وجدول ٣ يوضح ذلك.

جدول ٣
بارامتر الصعوبة لكل مفردة من مفردات صورتي الاختبار (ن= ١٠٦٥)للصورة الأولى، (ن= ٩٦٥) للصورة الثانية عدد المفردات لكل صورة اختبارية(٢٢) مفردة

رقم المفردة بالصورة الاختبارية الأولى	بارامتر الصعوبة المقدر (b- parameter)	رقم المفردة بالصورة الاختبارية الثانية	بارامتر الصعوبة المقدر (b- parameter)
١	-٢.٥٧٦	١	-٢.٩٥٤
٢	-١.٣٢١	٢	٠.١٨٩
٣	-٠.٩٩٥	٣	-١.٢٧٨
٤	-١.٣٢١	٤	٠.٠٤١٥
٥	-١.٢٨٤	٥	-٠.٤٢٦٣
٦	-٠.٢٠٤٨	٦	-٠.٢١١٧
٧	-٠.٩٢٠	٧	-١.٢٥٤
٨	-١.١٠١	٨	٠.٠٠٤٨
٩	٠.٢٢١٤	٩	-٠.٥٢٨٨
١٠	-١.٢٣٣	١٠	١.٨٤٧
١١	-١.٥٣٤	١١	-٠.٣٣٧٥
١٢	٠.٢٤٢٩	١٢	-٠.٥٦٧٨
١٣	-١.٣٢	١٣	١.٢٢٦
١٤	٠.١٩٧٨	١٤	١.٤٢٥
١٥	٠.٠٨٥٥	١٥	-١.٧٨٥
١٦	٠.٠١٧٨	١٦	-٠.٧١٥
١٧	-٠.٨٩٥٥	١٧	٠.٩٨٥٦
١٨	١.١٧٥	١٨	٠.٦٥٨٠
١٩	-٠.٨٥٩	١٩	٠.٥٨٩٠
٢٠	٠.٢٤٤	٢٠	٠.٢٨٩٥
٢١	-٠.٤١٨٧	٢١	٠.٤٣٥٢
٢٢	٠.٩٨١	٢٢	-٠.٥٩٤٧

يلاحظ من جدول ٣ اختلاف بارامترات الفقرات من رقم (١٣- ٢٢) بالرغم من أنها مفردات مشاركة في الصورتين الاختباريتين، ويرجع ذلك لاختلاف الأفراد الذين طبقت عليهم الصورتين.

ويمكن الحصول على معاملات صعوبة وتمييز المفردات المشتركة كما يوضحها جدول ٤.

جدول ٤
متوسط تمييز وصعوبة المفردات المشتركة والانحراف المعياري لها على الصورتين الاختباريتين الأولى والثانية، عدد المفردات المشتركة (١٠) مفردات

المؤشر الاحصائي / الصورة الاختبارية	الصورة الأولى	الصورة الثانية
متوسط تمييز المفردات المشتركة	١.٠٠	١.٠٠
متوسط صعوبة المفردات المشتركة	-٠.٤٨٦	-٠.١٥٢
الانحراف المعياري للمفردات المشتركة	٠.٨٥٠	٠.٧٧٨

يتضح من جدول ٤ أن متوسط تمييز المفردات المشتركة لكل من صورتي الاختبار ثابت ويساوي الواحد الصحيح، وذلك لأن نموذج فيشر (النموذج المعمم لنموذج راش) يقوم بتثبيت معلم التمييز وجعله مساوياً للواحد الصحيح، وكذلك فإن متوسطي صعوبة المفردات المشتركة لكل من الصورة الأولى والثانية (-٠.٤٨٦، -٠.١٥٢) على الترتيب، بينما كان الانحراف المعياري للمفردات المشتركة على صورتي الاختبار(٠.٨٥٠، ٠.٧٧٨) على الترتيب.

أ- معادلة المتوسط/ المتوسط:

ولإيجاد معاملات المعادلة (A, B) عن طريق المتوسط/ المتوسط، يمكن استخدام تلك المعادلات:

$$A = \frac{\mu(\alpha_j)}{\mu(\alpha_i)}$$

$$B = \mu(b_J) - A\,\mu(b_I)$$

حيث أن

$\mu(\alpha_I)$ هو متوسط تمييز المفردات المشتركة بالصورة الاختبارية الأولى (الهدف).

$\mu(\alpha_J)$ هو متوسط تمييز المفردات المشتركة بالصورة الاختبارية الثانية (القاعدة).

$\mu(b_I)$ هو متوسط صعوبة المفردات المشتركة بالصورة الاختبارية الأولى (الهدف).

$\mu(b_J)$ هو متوسط صعوبة المفردات المشتركة بالصورة الاختبارية الثانية (القاعدة).

ويوضح جدول ٥ ثوابت التحويل الخاصة بطريقة المتوسط/ المتوسط.

جدول ٥

ثوابت التحويل الخاصة بطريقة المتوسط/ المتوسط

A	B
١.٠٠	٠.٣٣٤

وبعد الحصول على ثوابت التحويل (A, B) فإن ميزان الصورة الأولى سيتحول إلى ميزان الصورة الثانية، حيث أن الصورة الأولى عبارة عن المقياس (الهدف)، وقد تم اختيار الصورة الثانية كمرجع (قاعدة) لها وذلك لأن معامل ثبات الصورة الثانية أعلى من الصورة الأولى، وجدول ٤٤ يوضح ذلك التحويل في ميزان الصورة الأولى.

جدول ٦

بارامتر الصعوبة لكل مفردة من مفردات الصورة

الاختبارية الأولى قبل عملية التحويل وبعدها وفقاً لطريقة المتوسط/ المتوسط

رقم المفردة بالصورة الاختبارية الأولى	بارامتر الصعوبة المقدر قبل عملية التحويل	بارامتر الصعوبة المقدر بعد عملية التحويل
١	٢.٥٧٦-	٢.٢٤٢-
٢	١.٣٢١-	٠.٩٨٧-
٣	٠.٩٩٥-	٠.٦٦١-
٤	١.٣٢١-	٠.٩٨٧-
٥	١.٢٨٤-	٠.٩٥-
٦	٠.٢٠٤٨-	٠.١٢٩٢
٧	٠.٩٢٠-	٠.٥٨٦-
٨	١.١٠١-	٠.٧٦٧-
٩	٠.٢٢١٤	٠.٥٥٥٤
١٠	١.٢٣٣-	٠.٨٩٩-
١١	١.٥٣٤-	١.٢-
١٢	٠.٢٤٢٩	٠.٥٧٦٩
١٣	١.٣٢-	٠.٩٨٦-
١٤	٠.١٩٧٨	٠.٥٣١٨
١٥	٠.٠٨٥٥	٠.٤١٩٥
١٦	٠.٠١٧٨	٠.٣٥١٨
١٧	٠.٨٩٥٥-	٠.٥٦١٥-
١٨	١.١٧٥	١.٥٠٩
١٩	٠.٨٥٩-	٠.٥٢٥-
٢٠	٠.٢٤٤	٠.٥٧٨
٢١	٠.٤١٨٧-	٠.٠٨٤٧-
٢٢	٠.٩٨١	١.٣١٥

يلاحظ من جدول ٦ أن الصورة الاختبارية الأولى (الهدف) تحولت إلى ميزان الصورة الاختبارية الثانية (القاعدة).

ومن خلال البرامج الاحصائية يمكننا الحصول على الدرجات الخام للصورة الأولى والقدرات المقابلة لها والدرجات المعادلة للدرجات الخام على الصورة الثانية وفقاً لطريقة (المتوسط/ المتوسط)، وجدول ٧ يوضح ذلك.

جدول ٧

الدرجات الخام على الصورة الأولى والقدرات المقابلة لها والدرجات المعادلة لها على الصورة الثانية وفقاً لطريقة المتوسط/ المتوسط

الدرجة الخام على الصورة الأولى	القدرة المقابلة لكل درجة	الدرجة المعادلة لها على الصورة الثانية
١	-٢.١٣٦	٠.٦٤
٢	١.٠٠٧٦	١.٣٢
٣	٠.٥٦٨	٢.٠٤
٤	-١.٠٠٧	٢.٨٤
٥	٠.٩٦٢	٣.٥١
٦	-١.٢٣٦٨٢	٤.٢٦
٧	٠.٤٥٢٦٧٧	٥.٠٥
٨	٠.٧٢٢٣٣	٥.٩٠
٩	-١.١٧٦١٣	٦.٦٣
١٠	٠.٨٩٨٣٢٦	٧.٤٥
١١	١.٢٤٩٧٣٧	٨.٢٩
١٢	-١.١٠٠١٧	٩.١٨
١٣	١.٠٠٦٣٩٣	١٠.٠٢
١٤	-١.٢٦٢٩٨	١٠.٩٨
١٥	-١.٧٣٧٣٨	١١.٨٨
١٦	-٢.٠٨٩٣٨	١٢.٨٧
١٧	٠.٤١٣	١٣.٦٦
١٨	٠.٨٢٢٨٩٤	١٤.٦٥
١٩	٠.٣٥٢٩٤٢	١٦.٦٥
٢٠	-١.٠٩٦٣٦	١٧.٥٦
٢١	-٠.٥٦٨١٨	١٨.٦٥
٢٢	٠.٥٤٧٦٧٣	١٩.٥٥

ويوضح جدول ٧ الدرجات الخام على الصورة الأولى(الهدف)، والقدرات المقابلة لها، وكذلك الدرجة المعادلة لها على الصورة الثانية(القاعدة)، فالطالب الذي تكون قدرته مساوية (-٢.١٣٦) تكون درجته على الصورة الأولى(١) والدرجة المعادلة لها على الصورة الثانية (٠.٦٤)، والطالب الذي تكون قدرته مساوية (٠.٤١٣) تكون درجته على الصورة الأولى(١٧) والدرجة المعادلة لها على الصورة الثانية (١٣.٦٦).

ثانياً: نتائج السؤال الثاني ومناقشتها:

وينص السؤال الثاني على " هل تحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون عند حذف الفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

وللإجابة عن هذا السؤال: تم أولاً الكشف عن الفقرات ذات الأداء التفاضلي في صورتي اختبار رافن، وذلك من خلال استخدام برنامج DIFAS4.0، بأن يكون طلاب الدراسات العليا مجموعة مرجعية وطلاب الفرقتين الأولى والرابعة مجموعة مستهدفة، أو العكس –**حيث أظهر برنامج DIFAS4.0 نفس الفقرات بالترتيب عند وضع طلاب الدراسات العليا مجموعة مستهدفة وطلاب الفرقتين الأولى والرابعة مجموعة مرجعية**– وأظهرت نتائج البرنامج وجود ١٨ فقرة من فقرات الاختبارين ذات أداءاً تفاضلياً، ثماني فقرات بالصورة الاختبارية الأولى هم الفقرة رقم (١، ٧، ٩، ١١، ١٢، ١٣، ١٨، ٢١)، وعشر فقرات بالصورة الاختبارية الثانية هم الفقرة رقم (١، ٦، ٩، ١٢، ١٥، ١٦، ١٧، ١٩، ٢٠، ٢١).

وقبل إجراء عملية المعادلة بين درجات الطلاب على صورتي الاختبار، كان من الضروري التحقق من توافر شروط نظرية الاستجابة للمفردة(أحادية البعد– الاستقلال الموضعي– التحرر من السرعة– توافر خاصية المنحنى المميز للمفردة ICC) وذلك لكل صورة من صورتي الاختبار خاصة بعد حذف الفقرات ذات الأداء التفاضلي.

وبعد أن تم التأكد من توافر شروط نظرية الاستجابة للمفردة، وكذلك جودة المفردات المشتركة في صورتي الاختبار، سيتم الإجابة على السؤال الذي ينص على " ما مدى تحقق معايير الفعالية للمعادلة العمودية لصورتي اختبار رافن للمصفوفات المتتابعة الملون عند حذف الفقرات ذات الأداء التفاضلي للعمر في ضوء نموذج فيشر؟

وللإجابة عن هذا السؤال، تم تحديد طرق المعادلة المستخدمة في ضوء نظرية الاستجابة للمفردة، وهي كالتالي:

أ- معادلة المتوسط/ المتوسط.

ب- معادلة المتوسط/ الانحراف المعياري.

ولحساب المعادلة بطريقة المتوسط/ المتوسط وكذلك طريقة المتوسط/ الانحراف المعياري، يستلزم ذلك حساب بارامتر المفردة(الصعوبة) لأن النموذج المستخدم هو نموذج فيشر وكما أشرنا إلى أن نموذج فيشر هو نموذج ديناميكي معمم لنموذج راش، حيث يكون فيه بارامتر التمييز للمفردات قيمة ثابتة وتساوي الواحد الصحيح أو قيمة قريبة منها ويتم تثبيتها في البرنامج المستخدم لجميع المفردات، وكذلك بارامتر التخمين مساوياً للصفر، ولتقدير بارامترات الصعوبة الخاصة بمفردات صورتي الاختبار فقد تم استخدام برنامج IRTEQ وقد تم ضبط البرنامج بحيث يكون معامل التدريج D = 1.7 أي في حالة القياس الطبيعي، وقد تم الحصول على بارامترات صعوبة المفردات لكل صورة من صورتي الاختبار، وجدول ٨ يوضح ذلك.

جدول ٨
بارامتر الصعوبة لكل مفردة من مفردات صورتي الاختبار (ن= ١٠٦٥)للصورة الأولى، (ن= ٩٣٥) للصورة الثانية عدد المفردات للصورة الأولى(١٤) مفردة، عدد المفردات للصورة الثانية(١٢) مفردة

رقم المفردة بالصورة الاختبارية الأولى	بارامتر الصعوبة المقدر (b- parameter)	رقم المفردة بالصورة الاختبارية الثانية	بارامتر الصعوبة المقدر (b- parameter)
١	-١.٣٤١	١	٠.١٩٩
٢	-٠.٩٦٥	٢	-١.٢٧٨
٣	-١.٣٧١	٣	٠.١٥١٥
٤	-١.٢٩٤	٤	-٠.٤٥٦٣
٥	-٠.٢٢٤٨	٥	-١.٣٥٤
٦	-١.١٢١		
٧	-١.٢٥٣		
٨	٠.١٩٠٨	٦	٠.٠٢٤٨
٩	٠.١٨٥٥	٧	١.٨٥٧
١٠	٠.٠١٩٨	٨	-٠.٣٣٦٥
١١	-٠.٨٦٥٥	٩	١.٢٣٦
١٢	-٠.٨٨٩	١٠	١.٤٤٥
١٣	٠.٢٩٤	١١	٠.٦٧٨٠
١٤	٠.٩٨٧	١٢	-٠.٥٩٠٧

يلاحظ من جدول ٨ اختلاف بارامترات الفقرات من رقم (٨- ١٤) بالرغم من أنها مفردات مشاركة في الصورتين الاختباريتين، ويرجع ذلك لاختلاف الأفراد الذين طبقت عليهم الصورتين.

ويمكن الحصول على معاملات صعوبة وتمييز المفردات المشتركة كما يوضحها جدول ٩.

جدول ٩
متوسط تمييز وصعوبة المفردات المشتركة والانحراف المعياري لها على الصورتين الاختباريتين الأولي والثانية، عدد المفردات المشتركة (٧) مفردات

المؤشر الاحصائي / الصورة الاختبارية	الصورة الأولى	الصورة الثانية
متوسط تمييز المفردات المشتركة	١.٠٠	١.٠٠
متوسط صعوبة المفردات المشتركة	-٠.٥٧٦	-٠.٢٧٢
الانحراف المعياري للمفردات المشتركة	٠.٨٣٠	٠.٧٥٠

يتضح من جدول ٩ أن متوسط تمييز المفردات المشتركة لكل من صورتي الاختبار ثابت ويساوي الواحد الصحيح، وذلك لأن نموذج فيشر (المعمم لنموذج راش) يقوم بتثبيت معلم التمييز وجعله مساوياً للواحد الصحيح، وكذلك فإن متوسطي صعوبة المفردات المشتركة لكل من الصورة الأولى والثانية (-٠.٥٧٦، -٠.٢٧٢) على الترتيب، بينما كان الانحراف المعياري للمفردات المشتركة على صورتي الاختبار(٠.٨٣٠، ٠.٧٥٠) على الترتيب.

أ- معادلة المتوسط/ المتوسط:

ولإيجاد معاملات المعادلة (A, B) عن طريق المتوسط/ المتوسط، يمكن استخدام تلك المعادلات:

$$A = \frac{\mu(\alpha_J)}{\mu(\alpha_I)}$$

$$B = \mu(b_J) - A\,\mu(b_I)$$

حيث أن

$\mu(\alpha_I)$ هو متوسط تمييز المفردات المشتركة بالصورة الاختبارية الأولى (الهدف).

$\mu(\alpha_J)$ هو متوسط تمييز المفردات المشتركة بالصورة الاختبارية الثانية (القاعدة).

$\mu(b_I)$ هو متوسط صعوبة المفردات المشتركة بالصورة الاختبارية الأولى (الهدف).

$\mu(b_J)$ هو متوسط صعوبة المفردات المشتركة بالصورة الاختبارية الثانية (القاعدة).

ويوضح جدول ١٠ ثوابت التحويل الخاصة بطريقة المتوسط/ المتوسط.

جدول ١٠

ثوابت التحويل الخاصة بطريقة المتوسط/ المتوسط

A	B
١.٠٠٠	٠.٣٠٤

وبعد الحصول على ثوابت التحويل (A, B) فإن ميزان الصورة الأولى سيتحول إلى ميزان الصورة الثانية، حيث أن الصورة الأولى عبارة عن المقياس (الهدف)، وقد تم اختيار الصورة الثانية كمرجع (قاعدة) لها وذلك لأن معامل ثبات الصورة الثانية أعلى من الصورة الأولى، وجدول ١١ يوضح ذلك التحويل في ميزان الصورة الأولى.

جدول ١١

بارامتر الصعوبة لكل مفردة من مفردات الصورة

الاختبارية الأولى قبل عملية التحويل وبعدها وفقاً لطريقة المتوسط/ المتوسط

رقم المفردة بالصورة الاختبارية الأولى	بارامتر الصعوبة المقدر قبل عملية التحويل	بارامتر الصعوبة المقدر بعد عملية التحويل
١	١.٣٤١-	١.٠٣٧-
٢	٠.٩٦٥-	٠.٦٦١-
٣	١.٣٧١-	١.٠٦٧-
٤	١.٢٩٤-	٠.٩٩-
٥	٠.٢٢٤٨-	٠.٠٧٩٢
٦	١.١٢١-	٠.٨١٧-
٧	١.٢٥٣-	٠.٩٤٩-
٨	٠.١٩٠٨	٠.٤٩٤٨
٩	٠.١٨٥٥	٠.٤٨٩٥
١٠	٠.٠١٩٨	٠.٣٢٣٨
١١	٠.٨٦٥٥-	٠.٥٦١٥-
١٢	٠.٨٨٩-	٠.٥٨٥-
١٣	٠.٢٩٤	٠.٥٩٨
١٤	٠.٩٨٧	١.٢٩١

يلاحظ من جدول ١١ أن الصورة الاختبارية الأولى (الهدف) تحولت إلى ميزان الصورة الاختبارية الثانية (القاعدة).

ومن خلال البرامج الاحصائية يمكننا الحصول على الدرجات الخام للصورة الأولى والقدرات المقابلة لها والدرجات المعادلة للدرجات الخام على الصورة الثانية وفقاً لطريقة (المتوسط/ المتوسط)، وجدول ١٢ يوضح ذلك.

جدول ١٢
الدرجات الخام على الصورة الأولى والقدرات المقابلة لها
والدرجات المعادلة لها على الصورة الثانية وفقاً لطريقة المتوسط/ المتوسط

الدرجة الخام على الصورة الأولى	القدرة المقابلة لكل درجة	الدرجة المعادلة لها على الصورة الثانية
١	-٢.١٦٦	٠.٦٩
٢	١.٠١٧٠	١.٣٨
٣	٠.٥٨٩	٢.١٦
٤	-١.٠٢٧	٢.٩٨
٥	٠.٩٨٢	٣.٦٩
٦	-١.٢٥٢	٤.٥٦
٧	٠.٤٦٧٧	٥.٤٥
٨	٠.٧٤٣	٦.٢٣
٩	-١.١٧٦	٦.٩٣
١٠	٠.٨٩٥٦	٧.٨٥
١١	١.٢٤٤٧	٨.٦٩
١٢	-١.١٠٢٧	٩.٦٨
١٣	١.٠٣٦	١٠.٥٢
١٤	-١.٤٩٨	١١.٠٢

ويوضح جدول ١٢ الدرجات الخام على الصورة الأولى(الهدف)، والقدرات المقابلة لها، وكذلك الدرجة المعادلة لها على الصورة الثانية(القاعدة)، فالطالب الذي تكون قدرته مساوية (-٢.١٦٦) تكون درجته على الصورة الأولى(١) والدرجة المعادلة لها على الصورة الثانية (٠.٦٩)، والطالب الذي تكون قدرته مساوية (٠.٧٤٣) تكون درجته على الصورة الأولى(٨) والدرجة المعادلة لها على الصورة الثانية (٦.٢٣).

وللتحقق من فاعلية المعادلة عند حذف الفقرات ذات الأداء التفاضلي أو في ظل الابقاء عليها للعمر في ضوء نموذج فيشر، تم حساب معامل الصدق التقاطعي للمعادلة العمودية، وقد بلغ معامل الصدق التقاطعي في ظل الاحتفاظ بالفقرات ذات الأداء التفاضلي (٠.٠٠٠٨)، في حين بلغ معامل الصدق التقاطعي عند حذف الفقرات ذات الأداء التفاضلي (٠.٠٠٠٥)، وكلا المعاملين يشيرا إلى توافر الفاعلية للمعادلة العمودية بدرجة مقبولة، وبمقارنة قيم معامل الصدق التقاطعي الناتجة عن إجراء المعادلة العمودية لمستويي الاختبار (صورتي الاختبار) في ظل الاحتفاظ بالفقرات ذات الأداء التفاضلي للعمر، نلاحظ أن قيمة هذا المعامل انخفضت من (٠.٠٠٠٨) إلى (٠.٠٠٠٥) عند حذف الفقرات ذات الأداء التفاضلي للعمر مما يدل على زيادة دقة المعادلة العمودية لصورتي الاختبار عند استبعاد الفقرات ذات الأداء التفاضلي، وأكد ذلك كل من (حسين عبد القادر، ١٩٩٤؛ نضال الشريفين، ٢٠٠٣؛ أروى الحواري وأحمد عودة، ٢٠٠٨) في دراساتهم حيث توصلوا إلى أنه كلما قل معامل الصدق التقاطعي دل ذلك على دقة المعادلة، ويتفق مع هذه النتيجة دراسة (Hidalog, M.& Lopez, J., 2002) والتي أكدت أن الفقرات ذات الأداء التفاضلي قد تأثر بدرجة كبيرة على دقة المعادلة العمودية، ومن المهم تحديد هذه الفقرات وتقليل تأثيرها، فقد تحسنت درجة دقة المعادلة عند حذف الفقرات ذات الأداء التفاضلي، لذلك يجب أخذ الأداء التفاضلي للفقرات بعين الاعتبار عند إجراء المعادلة لدرجات الاختبارات وذلك لتقليل أثر الأداء التفاضلي على دقة فاعلية المعادلة.

– أوجه الاستفادة من البحث :

في ضوء نتائج البحث الحالي يمكن تقديم هذه التوصيات:

- ✓ يقوم المتخصصون في الحقل التربوي ببناء الكثير من الاختبارات لنهاية كل فصل دراسي، ويتطلب ذلك وقتاً كبيراً وجهداً مضاعفاً، ومن خلال إجراء معادلة للاختبارات يمكننا التغلب على هذه المشكلات بصورة حقيقية، وعلى العاملين بالحقل التربوي دراسة هذا الموضوع بشيء من التفصيل، وإعطائه المزيد من الاهتمام والعناية.
- ✓ إن توافر صور متكافئة أو بديلة للاختبار تسهل من عملية إعادة اختبار المفحوصين عندما يكون ذلك ضرورياً أو مرغوباً فيه للحفاظ على سرية الاختبار ولمنع أثر الاختلاط بين الأفراد الذين طبق عليهم الاختبار والذين سيطبق عليهم فيما بعد.
- ✓ إن الاعتماد بشكل كبير على اختبارات الذكاء عند استخدامها في الدراسات أو البحوث، كأداة رئيسة أو محك أو غير ذلك يجعل من الأهمية بمكان إعداد صوراً متكافئة، لذا كان من الضروري عمل معادلة لصورتي اختبار رافن للذكاء.

قائمة المراجع

- أروى الحواري وأحمد عوده (٢٠٠٨). الخصائص السيكومترية لصور مختارة من اختبارات الرخصة الدولية لقيادة الحاسوب في الأردن، ومعادلة درجاتها، المجلة الأردنية في العلوم التربوية، ٤(٤)، ص ص ٢٩٧ – ٣١٩.
- إسماعيل حسن فهيم الوليلي (٢٠٠٥). تكافؤ درجات الاختبارات في ضوء نظريتي القياس الكلاسيكية والحديثة (دراسة سيكومترية مقارنة)، مجلة كلية التربية جامعة بنها، ١٥(٦٣)، ص ص٩٩ – ١٤٩.
- حسين محمد عبد القادر أيوب(١٩٩٤). المقارنة بين أربع طرق للمعادلة عندما يكون التصميم من مجموعات متكافئة وغير متكافئة، رسالة دكتوراه، كلية الدراسات العليا، الجامعة الأردنية.
- راشد حماد الدوسري (٢٠٠١). معادلة الاختبارات مفهومها، وطرقها، ومشكلات تطبيقها، مجلة العلوم التربوية والنفسية ، ١(٤)، ص ص ١٠٧ – ١٤١.
- صلاح الدين محمود علام(٢٠٠٢). القياس والتقويم التربوي والنفسي أساسياته وتوجهاته المعاصرة، القاهرة: دار الفكر العربي.
- صلاح الدين محمود علام(٢٠٠٥). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي، القاهرة: دار الفكر العربي.
- على ماهر خطاب (٢٠٠١). علم النفس الفارق ط ٣، القاهرة ، مكتبة الأنجلو المصرية.
- عماد احمد حسن على (٢٠١٤). المصفوفات المتتابعة الملونة للأطفال (تعديل وتقنين)، ط ١، القاهرة: الانجلو المصرية.
- عماد أحمد حسن على(١٩٩٦). العمر والخبرة وعلاقتهما بالبناء العاملى للقدرة المكانية، رسالة دكتوراه، كلية التربية، جامعة أسيوط.

- عماد أحمد حسن علي(٢٠١٦). تقنين اختبار رافن الملون للمصفوفات المتتابعة للكبـــار. القاهرة: الأنجلو المصرية.
- ميمي السيد إسماعيل(٢٠٠٧). الخصائص السيكومترية لاختبار القدرة العقلية باســـتخدام نموذج راش لدى طلبة المرحلة الثانوية العامة، رسالة ماجستير، كليـــة التربية – جامعة الزقازيق.
- نضال كمال محمد الشريفين (٢٠٠٣). مدى تحقق معايير الفاعلية في معادلة اختبـــارين أحدهما ثنائي التدريج والآخر متعدد التـــدريج وفـــق نمـــاذج النظريـــة الكلاسيكية والنظرية الحديثة في القياس، رسالة دكتوراه، كليـــة العلـــوم التربوية، جامعة عمان العربية للدراسات العليا.

Baghaei, H. & Kubinger, M.(2015). Linear Logistic Test Modeling with R, Practical Assessment, Research & Evaluation, 20 (1), pp., 1-11.

Chong Ho Yu, A. (2005). Test Equating by Common Items and Common Subjects: Concepts and Applications, A peer-reviewed electronic journal, 10(4). PP., 1-5.

Fischer, H. (1973). The linear logistic test model as an instrument in educational research Acta Psychologica, 37(1), pp., 359-374 .

Fischer, H., (1982). Some Applications of Logistic Latent Trait Models with Linear Constraints on the Parameters, Applied Psychological Measurement, 6(4), pp., 397- 416.

Haertel, E. (2004). The Behavior of zn Linking Items in Test Equating. CSE, 1-8.

Hambleton, R.& Swaminathan ,H.(1989).Item response theory: Principals and application. Boston, MA: Kluwer, NIJ publishing.

Harris, A.& Carlton, S.(1995). Pattern of gender differences on mathematics items on the scholastic Applied Measurement in Education, 6(2), pp137-150.

Harwell, S. (1997). Cities and Camps of the Confederate States, Univresity of Illinois, United States of America.

Hidalgo –Montesinos, M.& Lopez-Pina, J.(2002).Two Stage Equating in Differential Item Functioning Detection Under the Graded Response and the Lord Statistic. Educational and Psychological Measurement, 62(1),32.

Husen, T. &Postlethwaite, T., (1995). Item Response Theory, the Intenational Encyclopedia of Education, 1(5), pp., 351- 355.

Korawan, S.& Sirichai, K.& Nonglak, W. (2016). Student factors affecting latent transition of mathematics achievement measuring from latent transition analysis with a mixture item response theory measurement model, Procedia - Social and Behavioral Sciences, 1(217), pp., 729- 739.

Kolen , M. & Brennan , R. (2004) . Test equating, scaling, and linking: methods and practices. 2nd Ed. The United States of America, Springer.

Kolen, M. (1991). Comparison of traditional and item response theory methods for equating tests. Journal of Educational Measurement, 18(1), pp ,. 1- 11.

Philipp, S.(2008). Using the LLTM to evaluate an item-generating system for reading comprehension, sychology Science Quarterly, 50(3), pp., 345-362.

Pono, I,.(2002). On the Applicability of some IRT Models For Repeated Measurement Designs: Conditions, Consequences, and Goodness- of- Fit Tests, Methods of Psychological Research on line, 7(1), pp., 21- 40.

Purya, B.& Klaus, K.(2015). Linear Logistic Test Modeling with R, Practical Assessment, Research & Evaluation, 20(1), pp., 2-10.

Roos, L.(2002). Testing the Robustness of Item Response Theory Equating to Violation of the Condition of Equivalent Item Difficulty parameters in both the Extant pool Infused Items: A Comparison of the Effects on True-Score and Observed-Score Equating. Ph. D.. University of Nebraska.

Stephen, H.(2007). The impact of differential discrimination on vertical equating, Murdoch University, Western Australia.

Susan E., W., & Lisa M. S., (1981). Information Structure for Geometric Analogies: A Test Theory Approach, Applied psychological Measurement. 5 (3), pp, 383-397

Wiberg, M.(2004). Classical test theory vs Item response theory, an evaluation of the theory test in the Swedish driving-license test, Universities-Umea, 1(50), pp., 1- 25.

Yen, M. & Edwardson, S.(1999). Item response theory Approach in scale development, Official Journal of the Eastern Nursing Research Society and the Institute of Nursing, 1(48), pp., 234- 238.

Zhonghua, Z. (2010). Comparison of Different Equating Methods and An Application to link Testlet- Based Tests. Ph. D., the Chinese university of Hong Kong.

Zwick, R.,& Mazzeo, J.(1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items, Applied Measurement in Education, 1(10), pp. , 321- 334.