أثر استخدام طريقتي (التعادل المتزامن/ والتعادل بتثبيت بارامترات المفردات المشتركة) على إحصاءات المفردات والأفراد وفق تصميم المجموعات غير المتكافئة

د. إكرام حمزة السيد صهوان[1]

ملخص الدراسة:

هدفت الدراسة إلى المقارنة بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة في تقديرات بارامترات المفردات والأفراد وفق تصميم المجموعات غير المتكافئة ذو المفردات المشتركة. وقد تمثلت أداة الدراسة في ثلاث صور اختبارية في وحدة المادة وتركيبها بمادة العلوم للصف الأول الإعدادي، وقد بلغ إجمالي مفرداتها (٧٥) مفردة من نوع الاختيار من متعدد. وزعت المفردات على الصور الثلاث بحيث تتكافأ من حيث المحتوى والمستويات المعرفية ومستوى الصعوبة _قدر الإمكان_ وقد بلغ إجمالي المفردات بكل صورة منها (٣٠-٣١ – ٣٤) مفردة على الترتيب يربط بينها (١٠) مفردات مشتركة. وقد تم تطبيق الصور الاختبارية على عينة قوامها (١١٩٨) طالبا وطالبة بالصف الأول الإعدادي. حيث تم تحليل البيانات باستخدام برنامج WINSTEPS وفق طريقة التعادل المتزامن مرة، ووفق طريقة التعادل بتثبيت بارامترات المفردات المشتركة مرة أخرى. وقد تم الحصول على تقديرات بارامترات المفردات والأفراد وفق كل طريقة.

وتمثلت أهم نتائج الدراسة في:

أولا: بالنسبة لتأثير طريقتي التعادل على إحصاءات المفردات:

- **لا توجد** فروق دالة إحصائيا في نسبة المفردات غير الملائمة لأسس القياس الموضوعي وفق طريقتي التعادل؛ حيث حذفت سبع مفردات غير ملائمة عند تدريج الصور الاختبارية الثلاث وفق طريقة تثبيت بارامترات المفردات المشتركة؛ في حين حذفت تسع مفردات غير ملائمة عند التدريج وفق طريقة التدريج المتزامن.
- هناك **اتساق كبير** بين طريقتي التعادل في تصنيف المفردات (ملائمة/ غير ملائمة) حيث جاءت قيمة معامل فاي ومعامل كابا كبيرة ودالة إحصائيا.
- هناك **اتساق كبير** في تقدير صعوبات المفردات بين طريقتي التعادل؛ حيث جاءت قيمة معامل الارتباط كبيرة جدا ودالة إحصائيا، كما قد تساوى معامل الاختلاف بين الطريقتين.
- تكافأت إحصائيا تقديرات الخطأ المعياري لتقديرات صعوبات المفردات الملائمة وفق كلتا الطريقتين؛ حيث أسفرت النتائج عن عدم وجود فروق دالة إحصائيا بين متوسطي الخطأ المعياري

---

[1] مدرس (باحث) بقسم البحوث بالمركز القومي للامتحانات والتقويم التربوي.

لتقديرات صعوبات المفردات الملائمة وفق كل من الطريقتين. كما أشارت النتائج إلى وجود اتساق كبير في توزيع الأخطاء المعيارية لتقديرات الصعوبة للمفردات الملائمة وفق طريقتي التعادل معا.

- جاءت مؤشرات ثبات تقديرات صعوبات المفردات متمثلة في معامل الثبات الهامشي، ومعامل الفصل، ومقدار دالة المعلومات مرتفعة وفق طريقتي التعادل، كما قد جاءت قيم متوسط الأخطاء المعيارية لتقديرات صعوبات المفردات منخفضة بالنسبة لكل من طريقتي التعادل.

- لا توجد فروق دالة في توزيع تقديرات صعوبات المفردات المتطرفة على طرفي متصل التدريج (العليا/ الدنيا) بين طريقتي التعادل.

ثانيا: بالنسبة لتأثير طريقة التعادل على إحصاءات الأفراد:

- **لا توجد** فروق دالة إحصائيا في نسبة الأفراد غير الملائمين لأسس القياس الموضوعي وفق طريقتي التعادل؛ حيث حذف (٢١٠) فردا غير ملائم عند تدريج الصور الاختبارية الثلاث باستخدام طريقة تثبيت بارامترات المفردات المشتركة. في حين حذف (٢٢٦) فردا غير ملائم عند التدريج باستخدام طريقة التدريج المتزامن.

- هناك **اتساق كبير** بين طريقتي التعادل في تصنيف الأفراد (ملائمين/ غير ملائمين) حيث جاءت قيمة معامل فاي، ومعامل كابا كبيرة ودالة إحصائيا.

- هناك **اتساق كبير** في تقديرات قدرات الأفراد بين طريقتي التعادل حيث جاءت قيمة معامل الارتباط كبيرة جدا ودالة إحصائيا، كما تساوى معامل الاختلاف بين طريقتي التعادل.

- **توجد فروق** دالة إحصائيا في متوسط الخطأ المعياري لتقديرات قدرات الأفراد الملائمين بالنسبة لكل طريقة من طرق التعادل حيث كانت الأخطاء المعيارية المناظرة لتقديرات قدرات الأفراد وفق طريقة التدريج المتزامن أكبر منها بالنسبة لطريقة التعادل بتثبيت بارامترات المفردات المشتركة، في حين أشارت نتائج معامل الارتباط بين تقديرات الخطأ المعياري لتقديرات قدرات الأفراد وفق كل من الطريقتين إلى وجود اتساق كبير في توزيع الأخطاء المعيارية لتقديرات القدرة وفق طريقتي التعادل.

- جاءت مؤشرات ثبات تقديرات قدرات الأفراد متمثلة في معامل الثبات الهامشي، ومعامل الفصل مرتفعة وفق طريقتي التعادل كما قد جاءت قيم متوسط الأخطاء المعيارية لتقديرات قدرات الأفراد منخفضة بالنسبة لكل من طريقتي التعادل.

- لا توجد فروق دالة في توزيع تقديرات قدرات الأفراد على طرفي متصل التدريج (العليا/ الدنيا) بين طريقتي التعادل.

أثر استخدام طريقتي (التعادل المتزامن/ والتعادل بتثبيت بارامترات المفردات المشتركة) على إحصاءات المفردات والأفراد وفق تصميم المجموعات غير المتكافئة

د. إكرام حمزة السيد صهوان[٢]

مقدمة:

تعد متابعة النمو الأكاديمي للطالب هدفاً أسمى من أهداف التربية وضمان الجودة والمحاسبية، حيث تبنى معظم القرارات المهمة على درجات الطلاب سواء أكان ذلك على المستوى المؤسسي أو على مستوى السياسة العامة للدولة. وتتكرر المناسبات ومواقف القياس التي تبدو فيها حاجة الأنظمة التعليمية الملحة لإعداد الاختبارات المتكافئة لقياس سمة أو خاصية معينة، مما يتيح لتلك الأنظمة استخدام أكثر من صورة اختبارية لنفس الاختبار كالحاجة إلى تجنب أثر القياسات المتكررة للسمة المقاسة على الطالب عند الرغبة في متابعة تقدمه بعد عمليات التعليم والتعلم والتدريب والمران. أو الحاجة إلى المقارنة بين طلاب المناطق التعليمية المختلفة الذين تعرضوا لاختبارات مختلفة لنفس السمة أو القدرة المقاسة. والحاجة إلى إعادة اختبار بعض الطلاب عندما يكون ذلك ضرورياً أو مرغوب فيه، مع الحفاظ على سرية الاختبار، والمحافظة على عدم تأثر أداء الطلاب _الذين لم يختبروا بعد_ بمعرفة الطلاب الذين اختبروا أولاً، بالإضافة إلى منع انتقال أثر الاختبار المطبق أولاً إذا ما أعيد تطبيق ذات الاختبار مرة أخرى على نفس الطلاب، كتذكر الإجابة أو الألفة بمفردات الاختبار. كما تتطلب الاختبارات واسعة النطاق Large Scale Tests مثل اختبارات القبول بالتعليم العالي، واختبارات الاستعداد المدرسي Scholastic Aptitude Tests(SAT)، واختبارات إتقان اللغة الإنجليزية TOFEL، والاختبارات ذات الدرجة القصوى من الأهمية High Stakes Exams والاختبارات القومية توافر صور اختبارية متعددة للاختبار لضمان سرية الاختبارات، وإتاحة الفرصة لإجراء الاختبار أكثر من مرة على مدار العام. وذلك تحقيقا للمرونة في ملاءمة ظروف واحتياجات المختبرين (Cook, 2007,73:87)؛ حيث تعقد هذه الاختبارات على مدار توقيتات مختلفة؛ ومن ثم فهي لا تتضمن نفس المجموعة من المفردات، كما تسمح فلسفة بعض هذه الاختبارات بدخول الطالب للاختبار أكثر من مرة.

كما تتطلب سياسة الاختبار في بعض هذه الاختبارات تحرير بعض أو كل مفردات الاختبار بعد تطبيقه فعلياً، مما يستلزم الحاجة إلى ضرورة الإثراء المستمر بمفردات جديدة من

---

[٢] مدرس (باحث) بقسم البحوث بالمركز القومي للامتحانات والتقويم التربوي.

خلال بناء عدد كبير من الأسئلة وتجريبيا على الطلاب في آن واحد، حيث يكون من الصعب أن يتعرض طالب واحد لمثل هذا العدد الهائل من الأسئلة الأمر الذي يستلزم خلال عملية جمع البيانات تقسيم هذه الأسئلة إلى عدد من الصور الاختبارية في ضوء المواصفات الأساسية لبناء الاختبار وتجريبها على الطلاب في زمن مناسب.

ومن ثم كانت هناك ضرورة ملحة للبحث عن آلية تضمن تحقيق عدالة وموضوعية القياس وتكافؤ الفرص والمساواة عندما يختبر الطلاب على صور اختبارية مختلفة تقيس نفس السمة، كما تضمن عدم تسرب تلك الاختبارات فيستفيد منها البعض دون الآخرين، أو أن يتقدم الطالب لنفس الاختبار مرة أخرى فيألف الاختبار، ومن ثم تضمن دقة ما يقوم عليها من قرارات.

ولما كان بناء صور اختبارية متعددة على نفس مواصفات الاختبار Test Specification بصفة عامة لا يضمن تساوي صعوبات كل الصور الاختبارية؛ إذ يتعذر تحقيق التكافؤ الكلاسيكي (التكافؤ المنطقي) بدرجة تامة حيث يكون من الصعب الحصول على صورتين اختباريتين أو أكثر متكافئتين تماماً من حيث الصعوبة وغيرها من شروط التكافؤ الكلاسيكي؛ الأمر الذي يثير قضية إمكان مقارنة درجات الاختبار، ويبرز الحاجة إلى البحث عن إمكانية لوضع التدريج المستقل للصور الاختبارية المتعددة على ميزان مشترك واحد يضمن التغلب على مشكلة تفاوت الصعوبة بين تلك الصور الاختبارية من ناحية، كما يتيح إمكان استخدام الدرجات على أي منها بطريقة تبادلية من ناحية أخرى فيما يعرف بعملية معادلة الاختبارات Test Equating(Zhoe & Hambleton, 2017) (Arai & Mayekawa, 2011,1:16).

واستخدم مصطلح التعادل Equating ليصف عملية ضم الصور الاختبارية على ميزان تدرج مشترك وتحديد التقديرات المتناظرة على كل صورة (صلاح الدين محمود علام، ٢٠٠٥، ٢١٣).

ويعتبر تعادل القياس متطلباً مهماً لتحقيق عدالة المقارنة بين الأفراد، كما أن الدرجات بعد تعادلها يمكن أن تتم مقارنتها مباشرة، كما تتغلب على مشكلة اختلاف درجة الفرد باختلاف مستوى صعوبة الاختبار.

وبذلك تحل عملية التعادل مشكلة واجهت القياس التقليدي بقوة وهي عدم إمكان مقارنة درجات الطلاب عبر نتائج تقويم الصفوف الدراسية المختلفة نتيجة لاختلاف مفردات الاختبارات، وكذلك اختلاف مستويات صعوبتها. وذلك من خلال تدريج كافة المفردات على ميزان مشترك واحد يعد بمثابة إطار مرجعي ثابت للمقارنة بين قدرات الطلاب على الاختبارات المختلفة.

ووفق ميسليفي Mislevy (1992) فإن التعادل، وبناء الاختبار عمليتان متلازمتان

وعند تناغمهما فإن تعادل الدرجات على الصور الاختبارية المتكافئة يوفر الأدلة على أداء الطلاب في مجال ما أو مجموعة المهام التي يتم تطبيقها تحت نفس شروط التقنين بمرور الوقت. وفي نفس الوقت عندما تجرى عملية التعادل بدقة فإن ذلك مؤشر على جودة ضبط وتقنين بناء الاختبار، ومن ثم فإن بناء الاختبارات وفق ظروف تقنين صحيحة يضمن إنتاج عملية معادلة جيدة.

وتمر عملية التعادل في إطار نظرية الاستجابة للمفردة بثلاث خطوات تتضمن انتقاء تصميم جمع البيانات الملائم، ثم وضع تقديرات البارامترات على ميزان تدرج واحد، وأخيرا معادلة درجات الاختبارات بحيث يمكن مقارنتها واستخدامها بشكل متبادل. وإذا ما تم تعادل الصور الاختبارية المتماثلة في المحتوى والمتقاربة في مستوى صعوبتها والتي تقيس سمة ما فإن ذلك يعرف بالتعادل الأفقي Horizontal Equating. ويطلق على تقدير الدرجات المتكافئة للصور الاختبارية **المختلفة** في مستوى صعوبتها وتقيس نفس المحتوى في **مستويات مختلفة من القدرة** بالتعادل الرأسي Vertical Equating (أمينة محمد كاظم، ١٩٨٨).

وقد تعددت طرق التعادل حيث صنفها السيكولوجيون وفق أسس متعددة حيث انضوى بعضها تحت نظرية **القياس التقليدية** وتتضمن طريقة المتوسط الحسابي، والطريقة الخطية، وطريقة توكر، وطريقة المئينيات، وانضوى بعضها الآخر تحت **نظرية الاستجابة للمفردة** وصنفت إلى: طرق التعادل المنفصل Separated Methods ومنها طريقة المتوسط/ المتوسط، وطريقة المتوسط / سيجما، وطريقة هايبارا، وطريقة ستوكنج – لورد. وطرق التعادل المتزامن Concurrent Methods، وطريقة التعادل بتثبيت بارامترات المفردات المشتركة Fixed of Common Items Paramters(Pang, et al., 2010).

كما تعددت أيضا تصميمات جمع البيانات Equating Designs المستخدمة لإجراء عملية التعادل في إطار نظرية الاستجابة للمفردة ومن هذه التصميمات: تصميم المجموعة العشوائية، وتصميم المجموعة الواحدة، وتصميم المجموعة الواحدة المتوازن، وتصميم الأفراد المشتركين، وتصميم المجموعات غير المتكافئة ذو المفردات المشتركة.

ويعد تصميم المجموعات غير المتكافئة ذو المفردات المشتركة Common Items Non Equivalent Groups Design(CINEG) أكثرها شيوعا حيث إنه في سياق الاختبارات التربوية غالبا ما تكون المجتمعات غير متكافئة كما قد يختلف مستوى الأفراد من عام إلى آخر وفي هذا التصميم يتم تطبيق صورتين اختباريتين مختلفتين أو أكثر بينهما عدد من المفردات المشتركة على مجموعتين مختلفتين من الأفراد؛ حيث تستخدم استجابات الأفراد على تلك

المفردات المشتركة كأساس لإجراء عملية التعادل (Kolen& Brennan,2004,19). ومن ثم فإنه من الضروري الاهتمام بجودة بناء المفردات المشتركة، وأن يتم انتقاؤها بحيث تكون صورة مصغرة من الاختبار ككل؛ حيث تُقَدَّر بارامترات كافة مفردات الاختبار على أساس تقديرات تلك المفردات (Harris, 2007; Kolen and Bernnan, 2004; Patz & Yao, 2007, 253:272).

وفي إطار تصميم المجموعات غير المتكافئة يمكن إجراء التعادل المنفصل، والتعادل المتزامن، والتعادل بتثبيت بارامترات المفردات المشتركة. ففي طرق التعادل المنفصل يتم تقدير بارامترات الصور الاختبارية المتعددة بشكل منفصل، ثم يتم تقدير معاملات التحويل Transformation Coefficient باستخدام بارامترات المفردات المشتركة Common Items لوضع الصور الاختبارية على ميزان تدرج واحد بصفر مشترك واحد.

وفي طرق التعادل القائمة على تثبيت بارامترات المفردات المشتركة يتم معايرة الاختبار المرجعي أولاً، ثم يتم الحصول على بارامترات المفردات المشتركة وتثبيتها عند تلك التقديرات عند تدريج الصور الاختبارية التي يتم تعادلها مع الاختبار المرجعي، ونتيجة لذلك فإن توزيع تقديرات الاختبار الذي تم معادلته تم وضعها على نفس ميزان تدرج الاختبار المرجعي (Kim, 2006).

وفي التعادل المتزامن يتم دمج استجابات الطلاب على الصور الاختبارية المستخدمة (صورتين أو أكثر) في ملف بيانات واحد (Lord & Winkersky, 1984) مع وضع استجابات المفردات المشتركة في نفس مواضعها وتحلل بيانات الطلاب مرة واحدة حيث يتم وضع بارامترات المفردات والأفراد على ميزان تدرج واحد بصفر مشترك واحد في آن واحد باستخدام بيانات المفردات المشتركة.

مشكلة الدراسة:

إزاء تعدد طرق التعادل وتصميمات جمع البيانات فقد اهتمت دراسات عديدة بمقارنة طرق التعادل المختلفة من حيث ما ينتج عن هذه الطرق من تقديرات للأفراد والمفردات، ومستوى دقة تلك التقديرات، بالإضافة إلى دقة عملية التعادل ذاتها؛ وازدادت أهمية هذه الخطوة خاصة مع ما أيدته نتائج الدراسات من أن لطريقة التعادل المستخدمة تأثير دال على نتائج القياس (Jodoin; Koller and Swaminathan, 2003).

وبمراجعة الدراسات والبحوث السابقة التي اهتمت بالمقارنة بين طرق التعادل المختلفة تبين ما يلي:

- اهتمت بعض الدراسات بالمقارنة بين طرق التعادل القائمة على نظرية القياس الكلاسيكية، وطرق التعادل القائمة على نظرية الاستجابة للمفردة ومنها دراسة: (إيهاب محمد نجيب، ٢٠١٢)،(مايسة فاضل أبومسلم،٢٠١٠)،(يوسف عبد العاطي المحروق،٢٠١٦) &(Arai Maykawa,2011) (Zhao&Hambleton,2017). وأشارت نتائج تلك الدراسات إلى أن طرق التعادل وفق نظرية الاستجابة للمفردة تتميز بكثير من المميزات مقارنة بطرق التعادل التقليدية منها: قدرتها على توفير حلول أفضل للتعادل، والمرونة في تكوين الصور الاختبارية لإجراء عملية التعادل، وسهولة بناء وإثراء بنوك الأسئلة.

- اهتمت بعض الدراسات بالمقارنة بين طرق التعادل المنفصل فقط القائمة على نظرية الاستجابة للمفردة ومنها دراسة: (Kilmen& Demirtasli,2012)(محمد أحمد عبد الوهاب، ٢٠١٥).

- اهتمت بعض الدراسات بالمقارنة بين طرق التعادل المنفصل وطرق التعادل المتزامن ومنها دراسة (Peterson, Cook & Stoking, 1983; Kim & Cohen, 1998, Hanson and Beguin, 2002).

- اهتمت بعض الدراسات بمقارنة طريقة تثبيت بارامترات المفردات المشتركة بطريقة التعادل المتزامن والمنفصل ومنها دراسة (Pang et al., 2010).

- أشار (Kim,2006)(Paek & Young,2005) إلى أن هناك عددا قليلا من الدراسات التي قارنت طريقة تثبيت بارامترات المفردات المشتركة بالنسبة لطريقتي التدريج المتزامن والتدريج المنفصل.

- أوضحت الدراسات السابقة أن هناك تعارضا في نتائج بعض الدراسات التي اهتمت بالمقارنة بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة، حيث أشارت بعض الدراسات إلى تفوق طريقة التعادل المتزامن على طريقة التعادل بتثبيت المفردات المشتركة فيما يتعلق بإنتاج تقديرات أقل تحيزا وذات أخطاء معيارية أقل مصاحبة لهذه التقديرات **خاصة** عند استخدام عدد كاف من المفردات المشتركة، كما أوصت باستخدام طريقة التعادل المتزامن عند إثراء بنوك الأسئلة القائمة ومنها دراسة(Arai&Mayekawa,2011) . كما أشارت نتائج دراسة (Kang&Peterson,2009) إلى تفوق طريقة التعادل المتزامن على طريقة التعادل بتثبيت بارامترات المفردات المشتركة عند استخدام برنامج التحليل الإحصائي BILOG- MG وذلك فيما يتعلق بتوزيع قدرات الأفراد والمنحنى المميز للمفردة والاختبار. وأشارت نتائج دراسة (Keller& Keller,2015) إلى أن طريقة التعادل المتزامن أعطت

تقديرات أكثر ثباتا من طريقة التدريج بتثبيت بارامترات المفردات المشتركة. كما أشارت نتائج دراسة (Keller et al., 2004) إلى أن طريقة التعادل المتزامن كانت أكثر حساسية في الكشف عن التغيرات في قدرات المختبرين مقارنة بطريقة تثبيت بارامترات المفردات المشتركة.

- في حين أشارت نتائج دراسة كل من (Hu.et al.,2008) (Pange.et al.,2010) إلى تفوق طريقة التعادل بتثبيت بارامترات المفردات المشتركة على طريقة التعادل المتزامن؛ حيث صاحبت طريقة التعادل المتزامن تقديرات أكثر تحيزا وأخطاء معيارية أكبر مقارنة بطريقة تثبيت بارامترات المفردات المشتركة. كما أوصت دراسة (Pang et al., 2010) باستخدام طريقة تثبيت بارامترات المفردات المشتركة في تدريج الاختبارات التحصيلية حيث تفوقت على طرق المتوسط/سيجما، والمنحنى المميز، والتدريج المتزامن من حيث انخفاض التحيز والخطأ.

- وأشارت نتائج دراسة (Kang & Petersen,2009)(Kim&Cohen,1998) إلى تأثر كل من طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة عند استخدام عدد قليل من المفردات المشتركة وإن كانت طريقة التعادل باستخدام بارامترات المفردات المشتركة أكثر صمودا في هذه الحالة وتحسن أداء الطريقتين عند زيادة حجم العينة وعدد المفردات المشتركة.

- من جهة أخرى قد أشارت نتائج دراسة (Taherbhai & Sea,2017) إلى أنه لا توجد أفضلية لأي من طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة في تدريج الصور الاختبارية المتكافئة.

- في ضوء تعارض نتائج الدراسات حول فاعلية طرق التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة أوصت دراسة Hanson and Beguin (2002) بضرورة إجراء مزيد من الدراسات للوقوف على أكثر طرق التعادل دقة وفاعلية.

- على الرغم من اهتمام كثير من الدراسات العربية بمقارنة طرق التعادل الكلاسيكية بطرق التعادل التي تندرج تحت نظرية الاستجابة للمفردة خاصة طرق التدريج المنفصل؛ فإن الباحثة لم تعثر على دراسة عربية قارنت طريقة التعادل المتزامن وطريقة التعادل بتثبيت بارامترات المفردات المشتركة.

- ركزت معظم الدراسات على تأثير طرق التعادل على تقديرات المفردات والأخطاء المعيارية المصاحبة لها وكانت هناك ندرة في الدراسات التي تناولت تقديرات الأفراد والأخطاء المعيارية المصاحبة لها، ومقارنة نسبة الأفراد والمفردات الملائمين وغير الملائمين لأسس القياس، ومدى تمدد أطراف التدريج وفق كل طريقة.

وفي ضوء ندرة الدراسات التي اهتمت بمقارنة طريقتي التعادل المتزامن، والتعادل بتثبيت بارامترات المفردات المشتركة من جهة، وتضارب نتائج الدراسات التي اهتمت بالمقارنة بين هاتين الطريقتين، وما أوصت به بعض الدراسات من إجراء مزيد من الدراسات للمقارنة بينهما؛ فقد تمثلت مشكلة الدراسة الحالية في المقارنة بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة من حيث تأثير كل طريقة منها على تقديرات بارامترات المفردات والأفراد، ودقة تلك التقديرات، وذلك كما تم صياغتها في الأسئلة التالية:

أسئلة الدراسة:

١- هل تختلف النسبة المئوية للمفردات غير الملائمة لأسس القياس باختلاف طريقة التعادل؟

٢- إلى أي مدى تتسق تقديرات المفردات على متصل الصعوبة وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة؟

٣- إلى أي مدى تحقق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة دقة التقدير لصعوبات المفردات؟

٤- هل يختلف مدى تقديرات صعوبة المفردات الاختبارية باختلاف طريقة التعادل؟

٥- هل تختلف النسبة المئوية للأفراد غير الملائمين لأسس القياس باختلاف طريقة التعادل؟

٦- إلى أي مدى تتسق تقديرات الأفراد على متصل القدرة وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة؟

٧- إلى أي مدى تحقق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة دقة التقدير لقدرات الأفراد؟

٨- هل يختلف مدى تقديرات قدرات الأفراد عينة التدريج باختلاف طريقة التعادل؟

أهداف الدراسة: هدفت الدراسة الحالية إلى:

- بحث أثر استخدام طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة على التقديرات النهائية لإحصاءات **المفردات** كما تتمثل في: المقارنة بين نسبة المفردات غير الملائمة وغير الملائمة لأسس القياس، واتساق تدرج المفردات على متصل السمة المقاسة، ودقة تقدير صعوبات المفردات كما يعبر عنها بمؤشرات معامل الثبات لتقديرات الصعوبة، ومعامل الفصل للمفردات، والخطأ المعياري للقياس، ودالة المعلومات وفق كلتا الطريقتين.

- بحث أثر استخدام طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة على التقديرات النهائية لإحصاءات **الأفراد** كما تتمثل في: المقارنة بين نسبة الأفراد الملائمين

وغير الملائمين لأسس القياس، واتساق تدرج قدرات الأفراد على متصل القدرة، ودقة تقديرات قدرات الأفراد كما يعبر عنها بمؤشرات معامل الثبات لتقديرات القدرة، ومعامل الفصل للأفراد، والخطأ المعياري للقياس وفق كلتا الطريقتين.

أهمية الدراسة:

- الاستفادة من نتائج الدراسة الحالية في الوقوف على مستوى دقة كل من طريقتي التعادل المتزامن والتدريج بتثبيت بارامترات المفردات المشتركة في تقدير بارامترات الأفراد والمفردات، الأمر الذي من شأنه أن يسهم في تحديد الاستخدامات التطبيقية الأنسب لكل طريقة.
- تقدم الدراسة الخطوات العلمية والعملية المتبعة في معادلة درجات الاختبارات باستخدام طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة؛ الأمر الذي من شأنه أن يساعد الباحثين وطلاب الدراسات العليا في تطبيق هذه الطرق.
- قد تساعد نتائج الدراسة الحالية الجهات والمراكز البحثية المنوطة بالاختبارات وبناء الصور الاختبارية المتعددة في الوقوف على أكثر الطرق ملاءمة لمعادلة درجات الاختبارات، مما يسهم في تحقيق عدالة وموضوعية القياس، وييسر متابعة التقدم الأكاديمي للطلاب وانتقاء الطريقة الأنسب التي توفر أكبر قدر من الدقة وفق ظروف وأهداف القياس المختلفة، كالحفاظ على العينة أو الحفاظ على المفردات أو سهولة وسرعة الإجراءات، وغيرها من الأغراض العملية.
- قد تساعد الدراسة الحالية في سد فجوة في الدراسات العربية في مجال معادلة الاختبارات حيث تعد من الدراسات العربية القليلة التي تناولت المقارنة بين طرق التعادل القائمة على نظرية الاستجابة للمفردة في دقة عملية التعادل باستخدام بيانات واقعية، كما أنها الدراسة العربية الوحيدة – في حدود علم الباحثة- التي تناولت المقارنة بين طريقتي التعادل المتزامن وطريقة التعادل بتثبيت بارامترات المفردات المشتركة.
- قد تقدم الدراسة الحالية حلولا عملية لآلية استخدام كل من طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة في تدريج الصور الاختبارية المستخدمة في مواقف القياس المختلفة.
- تتيح عملية التعادل للمعلمين والقائمين على تقويم أداء الطلاب سحب صور اختبارية من بنك الأسئلة طبقاً للمعايير ومواقف القياس المختلفة مع ضمان عدالة وإمكان المقارنة بين أداء الطلاب في المواقف الاختبارية المختلفة.
- وضع ميزان تدرج واحد للصور الاختبارية الثلاث المستخدمة في الدراسة من الممكن أن يستخدم لعدة أعوام حيث إن المعايير التي توضع في وقت معين يمكن استخدامها في السنوات

التالية مما يساعد في تقليل الوقت والجهد والتكلفة التي تحتاجها عملية وضع المعايير. حيث يتم توفير عدد مناسب من المفردات الاختبارية المدرجة على ميزان تدرج واحد بصفر مشترك واحد ووحدة قياس ثابتة لوحدة المادة وتركيبها يمكن أن يسحب منه عدد من الصور الاختبارية المتعادلة القياس تستخدم لأغراض القياس المختلفة.

حدود الدراسة: اقتصرت الدراسة الحالية على:

- ثلاث صور اختبارية في مادة العلوم للصف الأول الإعدادي وحدة (المادة وتركيبها) بالفصل الدراسي الأول عام ٢٠١٦/ ٢٠١٧م تتضمن (٧٥) مفردة من نوع الاختيار من متعدد موزعة على ثلاثة مستويات معرفية (المعرفة، الفهم، التطبيق).
- استخدام طريقة التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة في تدريج الاختبارات التحصيلية.
- استخدام تصميم المجموعات غير المتكافئة ذى المفردات المشتركة ذات الربط الداخلي لجمع البيانات لإتمام عملية التعادل.
- استخدام برنامج WINSTEPS في تدريج المفردات الاختبارية وإجراء عمليتي التعادل.

مصطلحات الدراسة:

- **التعادل Equating:** العملية الإحصائية التي يمكن من خلالها تحويل نظام وحدات القياس التي تم الحصول عليها من صورة اختبارية أو أكثر إلى نظام وحدات قياس صورة اختبارية أخرى تقيس نفس السمة من خلال تحديد ما تساويه الدرجة التي يحصل عليها فرد ما في أحد الاختبارات من درجة حصل عليها في اختبار آخر يقيس نفس المتغير، وذلك وفقاً لميزان مشترك تتدرج عليه كافة المفردات الاختبارية. و يقصد به إجرائيا في الدراسة الحالية وضع مفردات الصور الاختبارية الثلاث المستخدمة في الدراسة الحالية على ميزان تدرج واحد بصفر مشترك واحد بوحدة قياس واحدة بحيث يمكن مقارنة الدرجات على هذه الصور أو أي صور أخرى فرعية يمكن أن تسحب من هذا البنك.
- **طريقة التعادل المتزامن Concurrent Calibration:** وضع بارامترات كافة المفردات على ميزان تدرج واحد باستخدام استجابات كافة المختبرين على كافة المفردات بكافة الصور الاختبارية في آن واحد بالاعتماد على إحصاءات المفردات المشتركة.
- **طريقة التعادل بتثبيت بارامترات المفردات المشتركة Fixed Common Items Parameters:** وضع بارامترات كافة المفردات على ميزان تدرج واحد من خلال تثبيت تقديرات صعوبات المفردات المشتركة المشتقة من التدريج المرجعي وبناء تدريج الصور

الاختبارية الأخرى على أساس تقديرات هذه المفردات.

- إحصاءات المفردات **Items Statistics** : مجموعة الإحصاءات التي يوفرها برنامج التحليل الإحصائي القائم على نظرية الاستجابة للمفردة المستخدم في الدراسة الحالية المرتبطة بالمفردات، كما تتمثل إجرائيا في النسب المئوية للمفردات الملائمة وغير الملائمة لأسس القياس، تقديرات صعوبات المفردات، معامل الثبات الهامشي، ومعامل الفصل، والخطأ المعياري للقياس، ودالة المعلومات.
- إحصاءات الأفراد **Persons Statistics**: مجموعة الإحصاءات التي يوفرها برنامج التحليل الإحصائي القائم على نظرية الاستجابة للمفردة المستخدم في الدراسة الحالية المرتبطة بالأفراد، كما تتمثل إجرائيا في النسب المئوية للأفراد الملائمين وغير الملائمين لأسس القياس، وتقديرات قدرة الأفراد، ومعامل الثبات الهامشي، ومعامل الفصل، والخطأ المعياري للقياس.

الإطار النظري:

تعددت المصطلحات التي ارتبطت بمفهوم التعادل في سياق نظرية الاستجابة للمفردة، ومن أهم هذه المصطلحات التدريج، والربط. ووفقا لكيلر (Keller,2007,18:19) فإن **التدريج** Scaling يعني وضع مفردات الصور الاختبارية على ميزان تدرج واحد بصفر مشترك واحد. بينما يتجاوز **التعادل** Equating عملية التدريج بالحصول على جداول تحويل الدرجة الخام إلى ما يناظرها من قدرة على الصور الاختبارية المختلفة. وعرف دورانس (Dorans et.al,2010) **الربط** Linking: بأنه التحويل من درجات اختبار ما إلى اختبار آخر، واعتبر معادلة الدرجات نوع خاص من الربط. انطلاقا من أن الربط يتضمن وضع تقديرات البارامترات على ميزان تدرج مشترك. وتناوله كيم وكوهين (Kim & Kohen, 1996) على أنه بناء مقياس مشترك باستخدام نظرية الاستجابة للمفردة بتحويل تقديرات بارامترات مجموعة من المفردات من تدريج لآخر.

ووفق تصنيف هولاند ودورانس (2006) Holland & Dorans فإن الربط يشتمل على ثلاث فئات تتمثل في: التنبؤ Predicting، بناء التدريج ووضع المفردات على تدريج واحد Scale Aligning، التعادل Equating.

وتتبنى الدراسة الحالية مفهوم التعادل بأنه المفهوم الأشمل الذي يتضمن عمليات انتقاء تصميم مناسب لجمع البيانات يمكن من خلاله ربط الصور الاختبارية بإحدى طرق الربط المناسبة ومن ثم تدريجها على ميزان تدرج واحد بصفر مشترك واحد يمكن أن تسحب من خلاله سحب صور

اختبارية متعادلة القياس بحيث يمكن مقارنة أداء الطلاب عليها مباشرة.

مفهوم التعادل:

يتمثل التعادل بعائلة من النماذج والطرق الإحصائية التي تستخدم لمواءمة الدرجات على اثنتين أو أكثر من الصور الاختبارية بحيث يمكن استخدامها بشكل متبادل Interchangeably (Gonzalez,2014).

ويتمثل الهدف من عملية التعادل في الحصول على تحويل مناسب للدرجات من صورة اختبارية إلى تدريج صورة اختبارية أخرى سواء أكان ذلك خلال نفس الصف أو عبر الصفوف الدراسية (Holland, et al.,2006).

وعرفها راشد حماد الدوسري (٢٠٠١، ١١٢) بأنها عملية الحصول على درجات متكافئة لأداتين تقيسان السمة نفسها. وعرفها روس (Roos, 2002, 5) بأنها ذلك الإجراء الذى يتيح إمكان وضع مفردات صورتى اختبار يقيس سمة ما على ميزان تدرج واحد.

ووفقاً لكل من جريجتر وكام (Gruijter & Kam, 2008و201) فإن معادلة الاختبار عملية إمبريقية تقوم على بناء علاقة بين الدرجات على صورتي اختبار أو أكثر. وعرفها هامبلتون وثوامنثان (Hambleton & Swaminathan, 1985و197) بأنها مجموعة من الإجراءات لإيجاد علاقة بين الدرجات على مجموعة من الصور الاختبارية. ويرى كل من (Hills al el., 1988) أن المعادلة أسلوب لعزل التأثير على درجات الاختبار من الفروق الناجمة عن الاختلافات في متوسط صعوبة المفردات، ومتوسط مستويات تمييزها، بين صورتين أو أكثر من صور الاختبار التي تبنى بغرض معادلتها.

ويعرفها (Lord & Stocking, 1990) بأنها عملية إيجاد الدرجات المتناظرة على الصور المختلفة للاختبار. في حين يعرفها (Dorans, 1990, 145) بأنها مجموعة من الإجراءات الإحصائية التي تحول بها درجات الأفراد على صورة ما للاختبار بحيث تكافئ الدرجات على صورة أخرى وذلك حتى يمكن المقارنة بين هذه الدرجات.

وتناولها صلاح الدين محمود علام (٢٠٠٥) على أنها تحويل نظام وحدات القياس الخاص بإحدى الصور الاختبارية إلى نظام وحدات القياس الخاص بالصورة الاختبارية الأخرى، بحيث تصبح القياسات المستمدة من درجات كل من الصورتين متعادلة بعد إجراء هذا التحويل.

ووفقا لكل من (Andersson & Wiberg, 2017) فإن التعادل عملية إحصائية تطبق غالبا على الاختبارات المقننة للتحقق من أن نتائج التطبيقات المختلفة للاختبارات المقننة يمكن مقارنتها.

وفى ضوء ما تقدم فإن عملية التعادل تتضمن:

- إجراء تعديل إحصائي على درجات صورة واحدة للاختبار لجعلها معادلة لدرجات صورة أخرى للاختبار ذاته.
- عملية إحصائية تستخدم لضبط الدرجات على صور الاختبار بحيث يمكن استخدام الدرجات على هذه الصور **بطريقة تبادلية**.
- تحويل التدريج المستقل لاختبارين أو أكثر إلى تدريج واحد مشترك.
- يكمن جوهر عملية التعادل في الآليات التي يتم فيها تحويل الدرجات من صورة لأخرى بما يضمن عدالة وموضوعية المقارنة بين الصور الاختبارية المختلفة.
- تهدف عملية معادلة الاختبارات إلى بناء علاقة بين صورتين اختباريتين، حتى يمكن استخدام أي صورة منها في تقييم الأفراد واتخاذ القرارات أو إصدار الأحكام المختلفة في ضوئها. حيث إن الدرجات المستخلصة من الصور الاختبارية حينئذ يمكن معالجتها كما لو كانت مشتقة من اختبار واحد.

متطلبات عملية التعادل:

حددت أدبيات القياس النفسي عدة شروط لإتمام عملية التعادل تتمثل فيما يلي:

- أن تقيس الصور الاختبارية نفس السمة / البنية، بمعني أن تقيس الصور الاختبارية المتغير ذاته. فلكي يتم تحقيق التعادل بين اختبارين يقيسان المعرفة العلمية ينبغي أن يشتمل الاختباران على مفردات تتناول نفس المزيج من المحتوى، وهو ما يعرف بتوازي المحتوى أو توازنه (صلاح الدين محمود علام، ٢٠٠٥، ٢١٥) (Dorans & Holland, 2000, 283) ويتجسد أكثر الطرق شيوعاً في التحقق من هذا الشرط في فحص محتوى وصياغة مفردات الصور الاختبارية من خلال عرضها على الخبراء في المجال مما يسهم في تحقيق صدق البناء.
- أن تحقق عملية المعادلة شرط التساوي Equity Properties بمعنى " أن يكون التوزيع التكراري المشروط للدرجات على أي صورة اختبارية للأفراد في مستوى معين من القدرة متساوياً $x(y):f_{x10}=f_{x(y)10}$، حيث x(y) دالة تناظر أحادي تستخدم في تحويل درجات الاختبار y إلى ميزان الاختبار X (صلاح الدين محمود علام، ٢٠٠٥، ٢١٦)، وبمعنى آخر، عدم وجود فروق بين الأفراد في أي مستوى من مستويات القدرة على كل من الصورتين الاختباريتين. ويؤكد (Lord, 1980) أن تساوى متوسط كل من f_{x10} . $f_{x(y)10}$ لا يعتبر كافياً لتحقيق شرط التساوي، ولكن يجب أن يتساوى كذلك تباين التوزيعين والأخطاء المعيارية ويضيف كولن (Kolen, 1981, 2) ضرورة أن تكون الصورتان متوازيتين/ متكافئتين تماماً، وهو ما إن حدث

في الحقيقة فلن يكون هناك داع لإجراء عملية المعادلة بين الصور الاختبارية. ويعتبر من المستحيل عملياً الوصول إلى صورتين متوازيتين تماماً، لذلك تظل الحاجة إلى المعادلة من الأمور الضرورية والمهمة. وعلى الرغم من صعوبة تحقيق شرط المساواة عملياً بالصورة الدقيقة، فإنه ينبغي العمل على تحقيقه قدر الإمكان (Spence, 1996, 7).

- أن تحقق عملية المعادلة **شرط التماثل** Symmetry Properties بمعنى أن تكون معادلة الصورة X إلى ميزان الصورة Y هي معكوس معادلة الصورة Y إلى ميزان الصورة X، وهنا يظل ناتج عملية التعادل كما هو بغض النظر أي الاختبارين أجريت معادلته على ميزان الآخر، أي أن تكون عملية التحويل قابلة للانعكاس (Cruijter & Kamp, 2008, 200).

- أن تتسم التحويلة المستخدمة في معادلة الصور الاختبارية بالثبات وعدم التغير عبر المجموعات الفرعية المأخوذة من المجتمع موضوع القياس، وهذا ما يتم التحقق منه بالتأكد من تشابه دوال تعادل الصور الاختبارية في تلك المجموعات أو على الأقل تقاربها من الدالة الخاصة بالمجتمع ككل(Dorans & Holland, 2000, 286). ويعنى تحقق هذه الخاصية بالضرورة أن التحويلة المستخدمة لإجراء عملية التعادل بين الصورتين X، Y لا تختلف من مجموعة لأخرى مثل (الذكور/ الاناث)، (الريف/ الحضر) وغيرها ومن ثم يؤكد (Kolen & Brennan, 2004, 13) على ضرورة مراعاة أن تكون العينات التي تستخدم لإجراء عملية التعادل ممثلة جيداً للمجتمع المستهدف بتطبيق الصور الاختبارية بعد معادلتها فيما يعرف بخاصية **اللاتباين في المجتمع الأصلي** بمعنى أن التحويل من الصورتين X، Y هو ذاته بصرف النظر عن عينة الأفراد التي أجريت عليها عملية التعادل، أي أن عملية التحويل من صورة لأخرى لا تختلف باختلاف المجتمعات الفرعية كالبنين والبنات بمعنى أن نموذج القياس يكون ثابتاً على كل المجتمعات الفرعية الممثلة للمجتمع الأصلي (صلاح الدين محمود علام، ٢٠٠٥، ٢١٧)، (Maris, 2010, 11-12).

وبالإضافة إلى الشروط السابقة أضاف بعض السيكولوجيين عددا من الشروط منها تساوي ثبات الصور الاختبارية؛ حيث يسهم في تحقيق هذه الشروط عمليا بناء الصور الاختبارية على نفس الإطار المرجعي من حيث المحتوى والمؤشرات الإحصائية وتطبيقها تحت نفس الشروط (Kolen, 2007).

أنواع التعادل: تتضمن معادلة الاختبارات نوعين للتعادل هما: التعادل الأفقي والتعادل الرأسي، وفيما يلي نبذة مختصرة عن كلا النوعين.

التعادل الأفقي Horizontal Equating: وفيه تكون الاختبارات التي يتم معادلتها متكافئة

إحصائيا _بقدر الإمكان_ حيث تصمم الاختبارات لتكون متكافئة من حيث المحتوى ومستوى الصعوبة (Kolen, 2007). ويكون هدف عملية التعادل هو تعديل الفروق الناشئة عن الاختلاف في مستوى الصعوبة حيث ترتكز على أن الصور الاختبارية المتعددة للاختبار الواحد تعد لمجال سلوكي محدد أو نفس توزيع القدرة أو السمة المقاسة، ومن ثم لمجتمع محدد من الأفراد يتمثل فيه توزيع القدرة نفسها؛ بمعنى أن الصور الاختبارية المتعددة للاختبار تكون متماثلة في المحتوى ومتقاربة في مستوى صعوبتها وتقيس السمة ذاتها في المجتمع نفسه، ويعود الاختلاف المحتمل في قدرة المفحوص التي يسفر عنها الاختبار إلى الاختبار وليس إلى قدرة المفحوص العقلية؛ حيث يشكل الأفراد الذين تطبق عليهم الصور الاختبارية المتعددة مجتمعاً واحداً من حيث المستوى التعليمي، ويتمثل هدف عملية التعادل في هذه الحالة في تعديل الفروق الناتجة عن اختلاف مستوى صعوبة المفردات في الصور الاختبارية المختلفة وتبدو الحاجة إلى هذا النوع من المعادلة عندما تكون هناك حاجة لتوافر صور اختبارية متعددة من اختبار معين. ويطبق هذا النوع من التعادل في معظم القياسات الكبرى مثل مقاييس الاستعداد المدرسي SAT. (أمينة محمد كاظم، ١٩٨٨) (أروى الحورانى وأحمد عودة، ٢٠٠٩) (نضال كمال الشريفين، ٢٠٠٩، ٢٣-٢٩).

التعادل الرأسي Vertical Equating: في هذا النوع من التعادل تتباين الصور الاختبارية المختلفة في مستوى صعوبتها كما تتباين قدرات الأفراد المستهدفين بتطبيق هذه الصور الاختبارية إلا أن هذه الصور جميعا تقيس نفس السمة؛ حيث ترتكز على أن القدرة أو السمة على متصل واسع المدى، أو متعدد المستويات كما هو الحال في السمات النمائية أو مقارنة التحصيل الدراسي بين المختبرين من ذوي المستويات المتباينة، أو عند الحاجة لإجراء المعادلة بين مجموعة من الاختبارات المكونة لإحدى بطاريات الاختبارات المصممة لصفوف متعددة. فقد لا يسمح المدى الواسع لمتصل السمة المقاسة بأن يتم القياس بصورة اختبارية واحدة؛ ومن ثم تكون هناك ضرورة إلى استخدام عدة صور اختبارية تقيس السمة ذاتها بحيث تكون كل صورة اختبارية جزءا أو مستوى من المدى للمتصل يمكن قياسه به.

ورغم افتراض أن هذه الاختبارات تقيس السمة ذاتها لكنها فى مستويات متباينة منها فإن هناك ضرورة ماسة لمعادلة هذه الصور الاختبارية بحيث تكون الدرجة على أي منها تناظر موقعاً محدداً على متصل السمة المقاسة بنفس وحدة القياس (نضال كمال الشريفين، ٢٠٠٩، ١٦-١٧) (محمد منصور الشافعي، ١٩٩٦، ١١٥). ويكون الهدف من إجراء عملية المعادلة الرأسية هو تحويل درجات الصور الاختبارية المعدة لمجتمعات من الأفراد في مستويات تعليمية متباينة إلى ميزان واحد بصفر مشترك واحد ووحدة قياس ثابتة. ويطلق على الميزان الذي يبنى من خلال

التعادل الرأسي بالميزان الرأسي أو الميزان النمائي (Briggs & Weaks, 2009). وعادة ما يتم ربط هذه الاختبارات من خلال مجموعة من المفردات المشتركة (الرابطة) والتي تتوزع عبر الصفوف/ المستويات المستهدفة. واتباع الإجراءات الإحصائية في إطار نظرية الاستجابة للمفردة لوضع كافة المفردات الاختبارية على ميزان تدرج واحد بصفر مشترك واحد. حيث يكون لكل مفردة بارامتراتها المحددة التي تحدد موضعها على ميزان تدرج المفردات.

وعموماً تشترك كل من طريقتي المعادلة الأفقية والرأسية في أن ناتج كل منها هو وضع درجات الصور الاختبارية المختلفة على ميزان تدرج واحد بصفر مشترك واحد ووحدة قياس ثابتة.

تصميمات معادلة الاختبارات Test Equating Designs

يتطلب إجراء معادلة الاختبارات، تحديد طريقة جمع البيانات المستخدمة لهذا الغرض، وتتعدد تصميمات التعادل، ويمكن تصنيفها على أساس استخدام المجموعات المتكافئة في مقابل المجموعات غير المتكافئة، وكذلك يمكن تصنيفها من حيث طريقة الربط ما إذا كان داخليا Internal Anchoring أم خارجيا External Anchoring. ومهما اختلفت طرق تصنيف تصميمات التعادل فإنه من الضروري مراعاة أن تمثل مجموعات المختبرين المجتمع المستهدف بتطبيق الصور الاختبارية الأساسية المستخدمة في مواقف القياس الفعلية، ويتوقف الاختيار من بين هذه التصميمات على بعض الاعتبارات العملية والإحصائية (Kolen & Brennan, 2004, 13). وفيما يلي عرض مختصر لهذه التصميمات.

١- **تصميم المجموعات العشوائية** Random Group Design: ويطلق عليه تصميم المجموعات المتكافئة Equivalent Groups، وفيه يتم توزيع الصورتين الاختباريتين على مجموعتين مستقلتين عشوائيتين ممثلتين للمجتمع المستهدف بحيث تطبق صورة واحدة فقط على كل مجموعة، ويتم ترتيب الصور عند تجهيزها بحيث يتم توزيعها على الممتحنين بشكل متتال؛ حيث يأخذ الممتحن الأول الصورة الاختبارية الأولى، ويأخذ المختبر الثاني الصورة الاختبارية الثانية ثم يأخذ الثالث الصورة الأولى... إلخ فيما يعرف بطريقة الترتيب الحلزوني/ الزجزاج spiraling (Kim et al., 2008, 84) (Kolen & Brennan, 2004, 15). وتأتي سهولة هذا التصميم من تعيين الأفراد عشوائياً لتوزع عليهم صوراً اختبارية مختلفة؛ الأمر الذي يجعل تنفيذ أو تطبيق هذا التصميم أكثر سهولة وأقل في الوقت والجهد مقارنة بالتصميمات التي يطبق فيها أكثر من صورة على نفس الأفراد. إلا أن استخدام هذا التصميم يتطلب وجود أكثر من صورة اختبارية ليتم تطبيقها في نفس الوقت، وهو ما يجعل تأمين عملية التطبيق مصدر تهديد

يواجه استخدام هذا التصميم، هذا بالإضافة إلى أن توفر أكثر من صورة اختبارية في نفس الوقت قد يكون من الصعب في بعض المواقف، كما يتطلب هذا التصميم أيضاً عدداً كبيراً من المختبرين بسبب توزيع صور اختبارية مختلفة على أفراد مختلفين (Baldwin, 2008, 19). ويؤكد (Kolen & Brennan, 2004, 15) على أهمية التحقق من وضع وترتيب الأفراد المختبرين في موقف التطبيق، خاصة عند استخدام طريقة الترتيب الحلزوني، لتجنب مشكلات مثل ترتيب الأفراد على سبيل المثال (ولد- بنت- ولد- بنت....) فتتوزع الصورة الأولى على الأولاد مثلا، وتوزع الصورة الثانية على البنات فيختل بذلك شرط العشوائية والتكافؤ بين المجموعتين.

٢- **تصميم المجموعة الواحدة** (Single Group Design (SG:

يتم في هذا التصميم تطبيق صورتين اختباريتين X وY على **نفس المجموعة** من المختبرين؛ إلا أنه قد يترتب على تطبيق الصورتين على نفس المجموعة، تحيز نتائج عملية التعادل نتيجة ثبات ترتيب تطبيق الصور الاختبارية على أفراد نفس المجموعة، فعند توزيع الصورة الأولى على أفراد المجموعة وبعد الانتهاء منها، تم توزيع الصورة الثانية، فإنه من الممكن أن يؤثر عامل التعب والإرهاق على المختبرين فتبدو لهم الصورة الثانية صعبة نسبياً مقارنة بالصورة الأولى، مما قد يؤثر على دقة عملية التعادل، كما أنه من الممكن أن تؤثر ألفة المختبرين بمحتوى الصورة الأولى فتبدو لهم الصورة الثانية سهلة نسبياً حتى ولو كان هناك فاصل زمني بين التطبيقين (Von Davier et al., 2004, 22- 23)، (Kolen & Brennan, 2004, 15).

٣- **تصميم المجموعة الواحدة المتوازن** Single Group Design with Counterbalancing

وفيه يتم ترتيب تطبيق الصور الاختبارية للتغلب على تأثير ثبات الترتيب الذي يواجه تصميم المجموعة الواحدة. حيث يتم في هذا التصميم إعداد الأوراق الاختبارية لموقف التطبيق بحيث يتضمن نصف هذه الأوراق الصورة الاختبارية الأولى يليها الصورة الاختبارية الثانية، والنصف الآخر تترتب فيه الصور الاختبارية عكس الترتيب السابق بحيث يأخذ الطالب الأول الصورة الاختبارية X ثم الصورة Y في حين يأخذ الطالب الثاني الصورة الاختبارية Y يليها الصورة X في نفس ظروف التطبيق، مع مراعاة الفصل بين تطبيق كل صورة من الصور الاختبارية بفاصل زمني (Kolen & Brennan, 2004, 16)، (von Davier et al., 2004, 27).

ويساعد التوزيع الحلزوني للصور الاختبارية في إمكان المقارنة بين المجموعة التي

تأخذ الصورة X أولاً والمجموعة التي تأخذ الصورة Y أولاً، مع ملاحظة أن هاتين المجموعتين تتماثل مع المجموعات في التصميم ذى المجموعات العشوائية. ولذلك فمعادلة الدرجات على الصورة الاختبارية X مع الدرجات على الصورة Y تتم بالاعتماد على بيانات الصورة التي يتم تطبيقها أولاً، وعندما يتشابه تأثير تطبيق الصورة X قبل الصورة Y مع تأثير تطبيق الصورة Y قبل الصورة X فإن علاقة التعادل لا تختلف في الحالتين.

وبالإضافة إلى مشكلة ترتيب تطبيق الصور الاختبارية التي يحاول هذا التصميم ضبطها ومعالجتها، فإنه يواجه مشكلة أخرى تتمثل في ضرورة تطبيق الصورتين على نفس المختبرين، الأمر الذي يحتاج إلى وقت مضاعف للتطبيق. وإذا ما تم ضبط عامل الإرهاق ولم يكن هناك اختلاف في تأثير ترتيب تطبيق الصور الاختبارية فإن استخدام هذا التصميم يصبح متميزاً عن تصميم المجموعات العشوائية لحاجته إلى حجم أقل من الأفراد المختبرين. ووفقا لكولن وبرنان (Kolen & Brennan, 2004, 17) فإنه من الناحية العملية يمكن استخدام تصميم المجموعة الواحدة المتوازن بدلاً من تصميم المجموعات العشوائية عندما يتحقق تطبيق الصورتين الاختباريتين على المختبرين عملياً. وكذلك عندما لا يكون من المتوقع وجود تأثير لترتيب تطبيق الصور الاختبارية على الممتحنين. وفي الحالات التي لا يتوافر فيها عدد كاف من المختبرين يمكن الاعتماد عليه في إجراء عملية التعادل بين الصور الاختبارية.

٤- **تصميم الأفراد المشتركين** Common Person Design

في هذا التصميم يتم تقديم الصور الاختبارية المطلوب معادلتها لمجموعتين من الأفراد (صورة واحدة لكل مجموعة)، على أن **تأخذ مجموعة مشتركة من الأفراد كلتا الصورتين**. فعندما تطبق الصورتان على مجموعة مشتركة من الأفراد يكون لكل فرد تقدير للقدرة على الصورة الأولى وتقدير آخر للقدرة على الصورة الثانية.، ووفقاً للنموذج المستخدم في تدريج مفردات الصورتين يجب أن تتكافأ التقديرات المتناظرة للقدرة لجميع الأفراد، وحيث إن متوسط الصعوبة للمفردات يختلف في الصورتين، فإن ذلك يؤدى إلى اختلاف موضع نقطة صفر التدريج لكل من الصورتين على متصل المتغير موضع القياس، وبالتالي يؤدى هذا إلى وجود اختلاف ثابت في تقديرات القدرة المشتقة من الصورتين، نتيجة لوجود إزاحة حادثة بين صفري التدريجين اللذين ينسب إليهما كل من تقديري القدرة ومن ثم يكون من الضروري تعديل تدريج مفردات الصورتين لتصبح على تدريج واحد وصفر مشترك واحد، وذلك باستخدام تقديرات الأفراد المشتركين بين الصورتين(أمينة محمد كاظم، ١٩٨٨، ١٠٩). ويعاني هذا التصميم من نفس العيوب التي يواجهها تصميم المجموعة الواحدة، حيث تطبق الصورتان على مجموعة واحدة من الأفراد (Hambleton

.(et al., 1991, 128

٥-**تصميم المجموعات غير المتكافئة ذو المفردات المشتركة** Common Item Non-Equivalent Groups Design (CINEG): يعتبر أكثر تصميمات التعادل شيوعا، ويطلق عليه تصميم المجموعات غير المتكافئة ذو الاختبار الرابط Non-Equivalent Group with Anchor Test (NEAT). وفيه يتم اختيار مجموعة من المفردات لتكون مشتركة بين الصورتين الاختباريتين، وقد تكون المفردات الرابطة داخلية Internal Anchor بمعنى أن تكون درجة المختبر عليها جزءًا من درجته الكلية، أو تكون خارجية External Anchor بمعنى أن الدرجة عليها ليست جزءًا من الدرجة الكلية على الاختبار، وفى كلتا الحالتين **تستخدم استجابات الأفراد على تلك المفردات المشتركة كأساس** لإجراء عملية التعادل (Kolen & Brennan, 2004, 19). ويستخدم هذا التصميم عندما لا يكون متاحا غير تطبيق صورة اختبارية واحدة في نفس يوم التطبيق بسبب بعض المخاوف أو المحاذير العملية كضرورة تطبيق صورة واحدة على المختبرين في المراحل الدراسية؛ لضمان تحقيق العدالة بينهم أو لضمان السرية. ونتيجة لعدم تكافؤ المجموعتين فإنه يفترض أن تكونا مختلفتين في مستوى القدرة، الأمر الذي يجعل الفروق بين الدرجات على الصورتين يعكس كلاً من الفروق بين المختبرين والفروق بين الصور الاختبارية نفسها. ولضمان أن تعكس الفروق بين الدرجات على الصور الاختبارية الفروق بين المجموعات يجب أن تكون مجموعة مفردات الربط وتسمى المفردات المشتركة بمثابة جزء مصغر من الاختبار، بمعنى أن تكون ممثلة نسبياً للصور الاختبارية من حيث المحتوى والخصائص الإحصائية ونوع المفردات. كما يراعى أن توزع في مواقع متشابهة/ متقاربة بما يضمن لها نفس آلية الظهور عبر الصور الاختبارية بحيث يراعى أن تظهر المفردات المشتركة في نفس مواضعها في الصور الاختبارية المختلفة أو في مواقع متشابهة حتى يتم تجنب تأثير موقع المفردة على مستوى صعوبتها(Kolen and Brennan, 2004; Patz and Yoa, 2007).

وتزيد شهرة وشيوع هذا التصميم عن غيره بسبب تطبيق صورة اختبارية واحدة فقط في الموقف الاختباري، وهو ما يتفق مع ضوابط وسياسات المواقف الاختبارية بشكل عام. كما أشارت نتائج الدراسات السابقة إلى فعالية هذا التصميم مقارنة بتصميمات التعادل الأخرى حيث أشارت دراسة (Kesamang, 2017) إلى أن تصميم التعادل القائم على المفردات المشتركة أكثر ثباتاً من تلك القائمة على الأفراد المشتركين.

ومن مشكلات هذا التصميم حاجته لافتراضات إحصائية صارمة من أجل فصل الفروق بين الصور الاختبارية عن الفروق بين مجموعات المختبرين الذين يتعرضون لهذه الصور

.(Kolen & Brennan, 2004: 21- 22)

طرق التعادل Equating Methods:

تتعدد طرق التعادل كما تتعدد تصميماته، وتجرى عملية التعادل في إطار كل من نظريتي القياس الكلاسيكية والحديثة، ونظرا لأن الدراسة الحالية تهتم بمقارنة اثنتين من طرق التعادل وفق النظرية الحديثة لذا يتناول الجزء التالي عرضا مختصرا لطرق التعادل في إطار هذه النظرية.

طرق التعادل وفق نظرية الاستجابة للمفردة

تصنف طرق التعادل المنبثقة من نظرية الاستجابة للمفردة إلى ثلاث مجموعات تتضمن: طرق التعادل المنفصل، وطرق التعادل المتزامن، وطرق التعادل بتثبيت بارامترات المفردات المشتركة. في حين ضمت تصنيفات أخرى طرق التعادل بتثبيت بارامترات المفردات المشتركة إلى طرق التدريج المنفصل على اعتبار أن إجراءات التعادل وفق كلتا الطريقتين تتم على أكثر من مرحلة (Hanson& Begun,2002; Hu,Rogers& Vrukmiyovic,2008). وفيما يلي عرض مختصر لكل طريقة منها.

طرق التعادل المنفصل Separate Equating Methods

في عملية التعادل المنفصل يتم تقدير بارامترات الصور الاختبارية المتعددة بشكل منفصل ثم يتم تقدير **معاملات التحويل** Transformation Coefficients باستخدام تقدير بارامتر المفردات المشتركة. حيث يتم معايرة كل صورة اختبارية على حدة، ثم تستخدم المفردات المشتركة لوضع الصورتين الاختباريتين أو أكثر على ميزان تدرج واحد ويطلق على الصور الاختبارية المستخدمة في وضع التدريج المشترك **بالاختبار المرجعي** Reference Test، ويطلق على الصورة الاختبارية التي تضم إلى هذا الميزان بالاختبار المتعادل The Equated Test ويستخدم التحويل الخطى Linear Transformation لوضع بارامترات المفردات من الاختبار المتعادل على ميزان الاختبار المرجعي بناء على بارامترات المفردات المشتركة بين الاختبارين ويمكن إجراء التحويل الخطى من خلال عدة طرق تصنف في فئتين هما:

١- **طرق العزوم** Moment Methods وتتضمن:

- **طريقة المتوسط / المتوسط** Mean/Mean Approach قدمها (Loyd&Hoover,1980) وتقوم على تحويل بارامترات الصعوبة والتمييز لمفردات الاختبار، ويستخدم متوسط **بارامترات التمييز** في تحديد منحنى المعادلة، ومتوسط **بارامترات الصعوبة** في تحديد ثابت المعادلة.

- **طريقة المتوسط/ سيجما** Mean/ Sigma Approach: وضعها ماركو (Marco,1977) وتقوم على استخدام **الانحراف المعياري** في تحديد منحنى المعادلة، ويستخدم متوسط بارامترات **الصعوبة** في تحديد ثابت المعادلة.

٢- **طرق المنحنى المميز** Characteristic Curve Methods وتتضمن:

- **طريقة هايبارا (Haebara,1980):** وتعتمد على الحصول على منحنى المعادلة وثابت المعادلة من خلال الفرق بين المنحنيات المميزة للمفردات، حيث يمثل هذا الفرق بين المنحنيات مجموع مربعات الفروق بين المنحنيات المميزة لكل مفردة. ثم يحسب ثابت المعادلة ومنحنى المعادلة المصغر لهذا الفرق.
- **طريقة ستوكنج- لورد** (Stocking-Lord 1983): ويتم فيها الحصول على منحنى المعادلة وثابت المعادلة من خلال الفرق بين المنحنيات المميزة للمفردات، ويمثل الفرق بين المنحنيات المميزة في هذه الحالة مربع مجموع الفروق بين المنحنيات المميزة لكل مفردة، ثم يتم حساب ثابت المعادلة ومنحنى المعادلة لأقل فرق.

طريقة التعادل بتثبيت بارامترات المفردات المشتركة Fixed Common Items Parameters: تتضمن هذه الطريقة خطوتين للتدريج والمعايرة حيث يتم فيها معايرة الاختبار المرجعي أولاً. وعند تدريج الاختبار المعادل Equated Test فإن بارامترات المفردات المشتركة يتم تثبيتها عند تلك التقديرات التي تم اشتقاقها من معايرة الاختبار المرجعي. وكنتيجة لذلك فإن توزيع تقديرات الاختبار الذي تم معادلته تم وضعها على نفس ميزان تدرج الاختبار المرجعي (Kim, 2006).

حيث تقدر بارامترات المفردات المشتركة من المجموعات المرجعية وتثبت وتقدر بارامترات المفردات غير المشتركة باستخدام بيانات المجموعة المستهدفة/ المرجعية (Kim & Cohen, 1998).

ووفق (Keller and Hambleton, 2013) فإن طريقة تثبيت بارامترات المفردات المشتركة تعتبر من الطرق الأكثر استخداماً لكفاءتها العملية ومرونتها. كما تستخدم بشكل كبير في برامج التقييم والاختبارات العالمية.

وفى التدريج بتثبيت تقديرات بارامترات المفردات المشتركة، فإن تقديرات المفردات لكل صور اختبارية تقدر بشكل منفصل/ مستقل **ولكن لا تقدر معاملات التحويل**، حيث تقدر بارامترات المفردات من خلال تثبيت بارامترات المفردات المشتركة بالقيم التي تم الحصول عليها من تقديرات سابقة، وتكون النتيجة أن تقديرات بارامترات المفردات يتم وضعها جميعاً على ميزان

تدرج واحد.

فعند استخدام طريقة تثبيت بارامترات المفردات المشتركة في التحليلات الإحصائية المستخدمة لإثراء البنوك القائمة بالفعل تستخدم بارامترات المفردات المشتركة المقدرة من قبل بالبنك، ويتم تدريج المفردات الجديدة فقط باستخدام هذه التقديرات (Kim, 2006).

ووجدت طريقة تثبيت بارامترات المفردات أكثر مناسبة لتدريج المفردات الجديدة والكشف عن تحيز المفردات في الاختبارات المواءمة بالكمبيوتر (CAT) (Ban et al., 2001) (Conzalz & et al., 2014, 334).

وعلى ذلك فإن طريقة التعادل بتثبيت بارامترات المفردات المشتركة تشترك مع طرق التعادل المنفصل في أن عملية التدريج تتم على أكثر من مرحلة، حيث يتم تدريج كل صورة اختبارية بشكل مستقل أولا، ثم يتم معادلة الصور الاختبارية؛ إلا إنه في حالة التدريج المنفصل يتم استخدام معاملات التحويل لإتمام عملية التعادل في حين تستخدم بارامترات المفردات المشتركة ذاتها في عملية التعادل وفق طريقة تثبيت بارامترات المفردات المشتركة.

طريقة التعادل المتزامن Concurrent Equating Method

في هذه الطريقة يتم دمج استجابات الطلاب على الصور الاختبارية المستخدمة (صورتين أو أكثر) في ملف بيانات واحد باستخدام عدد مناسب من المفردات المشتركة، مع وضع استجابات المفردات المشتركة في نفس مواضعها وتحلل الاستجابات مرة واحدة؛ حيث يتم وضع بارامترات المفردات والأفراد على ميزان تدرج واحد بصفر مشترك واحد في آن واحد باستخدام استجابات الطلاب على المفردات المشتركة. ويوضح شكل (١) التالي نموذجا لتكوين ملف البيانات في حالة التعادل المتزامن.

	المفردات	مفردات تظهر بالصورة X فقط			مفردات مشتركة تظهر بالصورتين X & Y			مفردات تظهر بالصورة Y فقط		
		مفردة 1	مفردة 2	مفردة 3	مفردة مشتركة 1	مفردة مشتركة 2	مفردة مشتركة 3	مفردة 1	مفردة 2	مفردة 3
العينة الأولى	طالب 1	√	√	√	√	√	√	.	.	.
	طالب 2	√	√	√	√	√	√	.	.	.
	طالب 3	√	√	√	√	√	√	.	.	.
العينة الثانية	طالب 1	.	.	.	√	√	√	√	√	√
	طالب 2	.	.	.	√	√	√	√	√	√
	طالب 3	.	.	.	√	√	√	√	√	√

شكل (١): رسم تخطيطي لتكوين ملف البيانات في حالة التعادل المتزامن.

يتضح من شكل (١) أن هناك مجموعتين من الأفراد: العينة الأولى وتتعرض للصورة Xوالتي تحتوي مفردات خاصة بالصورة X فقط بالإضافة إلى مفردات مشاركة تظهر بالصورتين (X & Y) معا. وتتعرض العينة الثانية لمفردات تظهر بالصورة Y فقط بالإضافة إلى مفردات تظهر بالصورتين (X & Y) معا.

وعلى هذا تحتسب تقديرات أفراد كل مجموعة على المفردات الاختبارية التي تعرضوا لها فقط، وتعامل المفردات التي لم يختبر عليها الطالب على أنها لم يتعرض لها الطالب Not-Reached وذلك باستخدام برامج التحليل الإحصائي وفق نظرية الاستجابة للمفردة المناسبة.

ويمكن أن تستخدم مع طريقة التعادل المتزامن طريقة تقدير الأرجحية العظمى المشتركة Joint Maximum Likelihood Estimation (JMLE) وطريقة تقدير الأرجحية العظمى الهامشية Marginal Maximum Likelihood Estimation (MMLE). كما يمكن استخدام تقدير بيز Bayes Estimation.

حيث تختلف طريقة التحليل المتزامن عن التدريج المنفصل حيث لا تكون هناك حاجة لإجراء عملية التحويل الخطي، واستخلاص معاملات التحويل لإتمام عملية التعادل كما هو الحال في طريقة التعادل المنفصل وبدلاً من ذلك فإن البارامترات يتم وضعها على مقياس تدرج واحد بصفر مشترك واحد خلال إجراءات التقدير الأساسية.

ويقدر عدد كبير من برامج الحاسب الآلي الخاصة بنظرية الاستجابة للمفردة مثل BILOG-MG, PARSCALE، IRT Command language ICL توزيع القدرة لكل مجموعة عملياً (تجريبياً) عند إجراء التدريج المتزامن، في حين تضع بعض برامج الحاسب الآلي الأخرى مثل FlexMIRT (Cai, 2017) وبرنامج MulTILOG توزيع القدرة لكل مجموعة في وضعها الافتراضي على المنحنى الاعتدالي ثم يقدر المتوسط والانحراف المعياري للتوزيع.

أخطاء التعادل Equating Errors: يوجد مصدران للخطأ عند إجراء عملية التعادل هما الخطأ العشوائي، والخطأ المنتظم، وفيما يلي عرض مختصر لكل منهما.

- **الخطأ العشوائي** Random Error : ينتج عند استخدام عينة من الأفراد في تقدير بارامترات المجتمع بدلاً من استخدام **المجتمع ككل** (Kolen, 1988, 121). وتحسب الأخطاء العشوائية من خلال الأخطاء المعيارية للتعادل، والتي تمثل الانحراف المعياري للدرجات المعادلة عندما يتكرر استخدام طريقة التعادل عددا كبيرا من المرات باستخدام عينات مختلفة من المختبرين (Livingston, 2004, 42). ويسهم استخدام عينات كبيرة بقدر مناسب واختيار التصميم الملائم لجمع البيانات في اختزال هذا النوع من الخطأ (Kolen & Brennan, 2004, 23).

- **الخطأ المنتظم** Systematic Error: من أسباب الخطأ المنتظم اختراق الافتراضات الإحصائية لطريقة التعادل أو التصميم المستخدم في جمع البيانات. كما يمكن أن تحدث طريقة التقدير المستخدمة تحيزاً في تقدير علاقة التعادل. بالإضافة إلى الاستخدام غير المناسب لتصميم جمع البيانات، وكذلك الاعتماد على عينة أو عينات من الأفراد غير ممثلة للمجتمع المستهدف بتطبيق الصور الاختبارية بعد معادلتها.

ويمكن ضبط الأخطاء المنتظمة من خلال الاهتمام ببناء الصور الاختبارية بحيث تقيس نفس المحتوى ويكون لها نفس الخصائص السيكومترية، بالإضافة إلى استخدام التصميم المناسب في جمع البيانات والطريقة المناسبة في إجراء عملية المعادلة (Kolen & Brennan, 2004, 23).

البرامج الإحصائية المستخدمة في إجراء التعادل بطريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة: تستخدم كثير من البرامج الإحصائية لإجراء التدريج المتزامن، وكذلك التدريج عن طريق تثبيت بارامترات المفردات المشتركة ومنها: WINSTEPS(Linecar,2003)، MULTILOG (Thissen, 1991) or BILOG-MG (Zimowski et al., 1996).

الدراسات السابقة:

تتناول الباحثة في هذ الجزء عرض الدراسات السابقة المرتبطة بموضوع الدراسة.

هدفت دراسة بترسن، وكوك، وستوك (Petersen, Cook & Stocking, 1983) إلى مقارنة طريقتي التعادل المتزامن والتعادل المنفصل متمثلا في طريقة ستوكنج – لورد. واستخدمت في التحليل الإحصائي للبيانات وفق طريقتي التعادل برنامج LOGIST وذلك لتدريج ست صور اختبارية للاستعداد اللغوي، وست صور للاستعداد الرياضي؛ حيث تم تدريج كل صورتين مثنى مثني بشكل متزامن مرة، وتدريجها وفق طريقة ستوكنج – لورد مرة أخرى. وقد أسفرت نتائج الدراسة عن تفوق طريقة التدريج المتزامن على طريقة ستوكنج – لورد فيما يتعلق بانخفاض الخطأ المعياري المصاحبة للتقديرات.

وتناولت دراسة كيم وكوهين (Kim & Cohen, 1998) مقارنة ثلاث طرق لبناء تدريج مشترك وفق نظرية الاستجابة للمفردة وتمثلت هذه الطرق في: طريقة الربط المنفصل باستخدام معاملات التعادل المشتقة من طريقة المنحنى المميز، وطريقة التدريج المتزامن القائمة على طريقة التقدير الهامشية العظمى البعدية، وطريقة التحليل المتزامن القائمة على طريقة الأرجحية العظمى الهامشية. وقد أسفرت نتائج الدراسة عن تفوق طريقة التحليل المنفصل على الطريقتين الآخريين في حالة استخدام عدد قليل من المفردات المشتركة، واتضح ذلك في انخفاض

أخطاء القياس المصاحبة لتقديرات صعوبات المفردات وتمييزها. وقد تساوت طرق التعادل الثلاث في حالة استخدام عدد كبير من المفردات المشتركة.

واستخدمت دراسة (Li, Tam, & Tompkins,2004) المحاكاة للمقارنة بين طريقة تثبيت بارامترات المفردات المشتركة وطريقة المنحنى المميز للمفردة باستخدام بيانات محاكاة وبيانات واقعية؛ حيث استخدمت ثلاث صور اختبارية تتضمن ١٥ مفردة مشتركة، وتراوح عدد المفردات غير المشتركة بين ٣٥: ٣٨ مفردة. وقد أسفرت نتائج الدراسة عن اتساق كبير في نتائج التعادل بين الطريقتين.

وهدفت دراسة (Taherbhai& Seo,2007) إلى مقارنة طريقة تثبيت بارامترات المفردات بطريقة التدريج المتزامن باستخدام نموذج راش أحادي المعلم للمفردات ثنائية الاستجابة ونموذج راش للتقدير الجزئي Rasch Partial Credit . واستخدم برنامج WINSTEPS لتدريج المفردات وفق طريقتي التحليل، وذلك باستخدام بيانات مولدة. وأوضحت نتائج الدراسة أن أداء الطريقتين يتأثر بحجم العينة وعدد المفردات المشتركة وغيرها ولا توجد أفضلية لاستخدام أي من الطريقتين مقارنة بالطريقة الأخرى في تدريج الصور الاختبارية المتكافئة.

وفحصت دراسة هيو وآخرون Hu et al, (2008) تأثير عدة عوامل على أربع طرق للمعادلة هي: طريقة تثبيت بارامترات المفردات المشتركة، وطريقة المنحنى المميز، وطريقة المتوسط/ سيجما، وطريقة التعادل المتزامن في وجود وعدم وجود مفردات مشتركة متطرفة وذلك باستخدام المحاكاة.

وقد أسفرت النتائج عن أن تحويلات المنحنى المميز، وطريقة المتوسط/ سيجما كانتا الأفضل في حين كان التدريج وفق طريقتي التعادل المتزامن، وطريقة تثبيت بارامترات المفردات ذا تفاعلات متداخلة مع متغيري تكافؤ المجموعات، وعدد النقاط المحددة للمفردات المتطرفة. أما في حالة عدم وجود مفردات متطرفة في مجموعة المفردات المشتركة فإن الطرق الأربع كانت ذات حساسية مناسبة ولكنها غير متساوية في حالة وجود مجموعات غير متكافئة.
وفى حالة عدم وجود فروق بين متوسطات القدرة للمجموعتين كانت طرق المعادلة الأربع متكافئة من حيث (التحيز والأخطاء العشوائية) ولكن عندما اختلفت متوسطات القدرة للمجموعتين بمقدار انحراف معياري واحد أنتجت الطريقتين المتوسط/ سيجما، وطريقة المنحنى المميز تقديرات أقل تحيزاً وذات خطأ عشوائي صغير. وكانت طريقة تثبيت بارامترات المفردات المشتركة متعادلة من حيث التحيز والخطأ المعياري. في حين كانت طريقة التعادل المتزامن أكثر الطرق من حيث

التقديرات المتحيزة والأخطاء المعيارية الكبيرة.

واهتمت دراسة (Ito& al., 2008) بمقارنة طريقتي التعادل المتزامن والتعادل المنفصل باستخدام بيانات واقعية لأحد الاختبارات القومية المقننة للأطفال من الحضانة حتى الصف التاسع في القراءة والرياضيات وتوصل إلى أن طريقتي التعادل المتزامن والمنفصل أدت إلى نتائج متشابهة فيما يتعلق بصعوبات المفردات وتمييزها وتقديرات قدرات الأفراد. في حين كانت نتائج طريقة التعادل المنفصل أفضل فيما يتعلق بالمحافظة على تمدد أطراف ميزان التدرج مقارنة بطريقة التدريج المتزامن.

وهدفت دراسة (Kang& Peterson,2009) إلى مقارنة ثلاث طرق لمعادلة الاختبارات تمثلت في طريقة التدريج المنفصل متمثلة في طريقة ستوكنج – لورد، وطريقة التدريج المتزامن، وطريقة التدريج بتثبيت بارامترات المفردات المشتركة. تم تطبيق طريقتي التدريج المتزامن والمنفصل باستخدام برنامج BILOG-MG، واستخدم كل من برنامجي BILOG-MG، PARSCALE لإجراء التدريج باستخدام طريقة تثبيت بارامترات المفردات المشتركة وذلك لأن كلا البرنامجين يقدرها بطريقة مختلفة. تم مقارنة الطرق الثلاث باستخدام بيانات محاكاة بناء على بيانات اختبار فعلي، وتم تقييم النتائج في ضوء توزيع القدرة، المنحنى المميز للمفردة، المنحنى المميز للاختبار. وتمثلت متغيرات الدراسة في حجم العينة، توزيعات القدرة، عدد المفردات المشتركة. وقد أدت طريقتي التدريج المتزامن والتدريج المنفصل إلى نتائج جيدة متماثلة بالنسبة للمتغيرات موضوع الدراسة، كما أسفر التدريج بتثبيت بارامترات المفردات المشتركة لنتائج مماثلة عند استخدام برنامج PARASCALE فقط.

واهتمت دراسة باتج وآخرون (Pang, et al., 2010) بتقييم فاعلية أربع طرق للتعادل/ للتدريج في تقدير تحصيل الطلاب وتمثلت طرق التعادل المستخدمة في: طرق التدريج المتزامن، وطريقة تثبيت بارامترات المفردات وطريقة ستوكنح – لورد باستخدام المنحنى المميز للاختبار وطريقة المتوسط/ سيجما وتمت المقارنة بين الطرق الأربع باستخدام **بيانات مولدة** لتقييم الاختبار لمفردات متعددة الأنواع (مفردات انتقاء إجابة، ومفردات إنتاج إجابة).

وأسفرت نتائج الدراسة عن تماثل أداء ثلاث من طرق التعادل وهي طريقة تثبيت بارامترات المفردات المشتركة، وطريقة المنحنى المميز للاختبار وطريقة المتوسط/ سيجما وتفوقها جميعاً على طريقة التحليل المتزامن.

وهدفت دراسة (Arai & Mayekawa, 2011) إلى مقارنة أداء عدد من طرق التدريج المستخدمة لإضافة مفردات جديدة إلى بنك أسئلة قائم موجود بالفعل. باستخدام تصميم المجموعات

غير المتكافئة ذي المفردات المشتركة باستخدام بيانات محاكاة؛ حيث تم مقارنة أربعة طرق لمعايرة الصور الاختبارية وتعادل القياس ، وقد تمثلت في طريقة المتوسط/ سيجما، طريقة تحليل العامل، طريقة تثبيت بارامترات المفردات المشتركة وطريقة التدريج المتزامن. حيث تمت محاكاة البيانات في ضوء بارامترات بنك أسئلة مكون من (٥٠) مفردة قائم بالفعل بصورتين اختباريتين تتكون كل منها من ٤٠ مفرده منها (١٠) مفردات مشتركة من بنك أسئلة قائم بالفعل. وأسفرت النتائج عن تفوق طريقة الربط المتزامن على طريقة تثبيت بارامترات المفردات المشتركة خاصة عند استخدام عدد كاف من المفردات المشتركة في حين كانت طريقة تثبيت بارامترات المفردات المشتركة أكثر دقة في حالة استخدام عدد قليل من المفردات المشتركة.

وقارنت دراسة تيان (Tian, 2011) طريقة الربط المنفصل متمثلا في طريقة ستوكنج – لورد، وطريقة الربط المتزامن مستخدم بيانات مولدة.

وأشارت النتائج إلى أن طريقة الربط المتزامن تنتج تقديرات أكثر دقة مقارنة بطريقة الربط المنفصل.

وذلك على اختبار مكون من مفردات اختيار من متعدد وأسئلة إنتاج إجابة باستخدام أعداد مختلفة من المفردات المشتركة (٥، ١٠، ١٥) مفردة وصور مختلفة للمفردات المشتركة (اختيار من متعدد فقط–اختيار من متعدد، وأسئلة إنتاج إجابة–أسئلة إنتاج إجابة فقط).

وهدفت دراسة كيلر و كيلر (Keller & Keller, 2015) إلى مقارنة ثلاث طرق للتدريج وفق نظرية الاستجابة للمفردة متمثلة في طريقة تثبيت بارامترات المفردات المشتركة، طريقة ستوكنج– لورد وطريقة التدريج المتزامن مع اعتبار تصنيف المختبرين في مستويات الأداء المتعددة وتقدير بارامتر القدرة عندما تتغير محتوى الاختبار قليلاً من عام إلى آخر ويتغير توزيع تقديرات قدرة الطلاب من عام لآخر وأشارت النتائج إلى أن طريقة التدريج المتزامن أنتجت تقديرات أكثر ثباتاً من طرق التدريج المنفصل.

وهدفت دراسة زهايو وهامبلتون (Zhao & Hambleton, 2017) إلى دراسة مدى حساسية عدد من طرق التعادل متمثلة في طريقة المتوسط/ سيجما، وطريقة المنحنى المميز لستوكنج– لورد وطريقة تثبيت بارامترات المفردات المشتركة لعدد من المفردات تتضمن: ملاءمة البيانات للنموذج، والتغير في قدرات الطلاب من اختبار لآخر بالإضافة إلى تصنيف الطلاب عبر مستويات الأداء المختلفة. وذلك باستخدام **بيانات محاكاة**. وجاءت طريقة تثبيت بارامترات المفردات المشتركة الأكثر حساسية لعدم ملاءمة البيانات، وكذلك بالنسبة للتغير في قدرات الطلاب عبر مرات التطبيق المختلفة. وتفوقت طريقة تثبيت بارامترات المفردات المشتركة على طريقتي

التعادل المنفصل (المتوسط/سيجما، ستوكنج-لورد).

تعليق على نتائج الدراسات السابقة: بمراجعة الدراسات السابقة يتضح ما يلي:

- تباينت الدراسات السابقة التي اهتمت بمقارنة طرق التعادل من حيث الطرق المستخدمة في كل منها حيث قارنت بعض الدراسات طرق التعادل المتزامن ببعض طرق التعادل المنفصل مثل دراسة (Tian,2011)(Kim&cohen,1998)، في حين قارنت بعض الدراسات بين طريقة التعادل بتثبيت بارامترات المفردات بطرق التعادل المنفصل مثل دراسة(Zhao&Hhambleton,2017)، وقارنت بعض الدراسات بين طرق التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة وبعض طرق التعادل المنفصل مثل دراسة (Hu et al.,2008)(Kang&Peterson,2009)(Keller&Keller,2015)، وقارنت بعض الدراسات بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة فقط مثل دراسة (Taherbhai&SSeo,2007)(Li&Tan&Tompkins,2004).
- استخدمت معظم الدراسات بيانات، ولم تتضمن الدراسات دراسات عربية حيث إنه _ في حدود علم الباحثة-اقتصرت الدراسات العربية على المقارنة بين طرق التعادل القائمة على النظرية الكلاسيكية وبعض طرق التعادل المنفصل، أو المقارنة بين بعض طرق التعادل المنفصل فقط.
- تباينت البرامج الإحصائية المستخدمة في التدريج وعملية التعادل، حيث استخدم في بعضها برنامج WINSTEPSمثل دراسة (Taherbhai&Seo,2007)، استخدمت دراسات أخرى برامج BILOGE-MG, PARASCAL مثل دراسة(Li&Tan&Tompkins,2004) واستخدم في بعضها برنامج LOGIST مثل دراسة (Petersen, Cook & Stocking, 1983).
- تباينت نتائج الدراسات السابقة؛ حيث اتضح أن هناك تعارضا في نتائج الدراسات السابقة التي اهتمت بالمقارنة بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة؛ حيث أشارت بعض الدراسات إلى تفوق طريقة التعادل المتزامن على طريقة التعادل بتثبيت المفردات المشتركة فيما يتعلق بإنتاج تقديرات أقل تحيزا وأخطاء معيارية مصاحبة لهذه التقديرات أقل **خاصة** عند استخدام عدد كاف من المفردات المشتركة، ومنها دراسة (Arai&Mayekawa,2011) . كما أشارت دراسة (Kang&Peterson,2009) إلى تفوق طريقة التدريج المتزامن على طريقة التدريج بتثبيت بارامترات المفردات المشتركة عند استخدام برنامج التحليل الإحصائي BILOG- MG وذلك فيما يتعلق بتوزيع قدرات الأفراد

والمنحنى المميز للمفردة والاختبار. وأشارت دراسة (Keller& Keller,2015) إلى أن طريقة التدريج المتزامن أعطت تقديرات أكثر ثباتا من طريقة التدريج بتثبيت بارامترات المفردات المشتركة. في حين أشارت نتائج دراسة كل من (Pange.et al.,2010) (Hu.et al.,2008) إلى تفوق طريقة التعادل بتثبيت بارامترات المفردات المشتركة على طريقة التعادل المتزامن؛ حيث صاحبت طريقة التعادل المتزامن تقديرات أكثر تحيزا وأخطاء معيارية أكبر مقارنة بطريقة تثبيت بارامترات المفردات المشتركة. وقد أشارت دراسة (Kim&Cohen,1998) إلى تأثر كلتا الطريقتين عند استخدام عدد قليل من المفردات المشتركة وإن كانت طريقة التعادل باستخدام بارامترات المفردات المشتركة أكثر صمودا في هذه الحالة. ومن جهة أخرى أشارت دراسة (Taherbhai & Sea,2017) إلى أنه لا توجد أفضلية لأي من طريقتي التعادل (المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة في تدريج الصور الاختبارية المتكافئة.

منهج الدراسة وإجراءاتها:

عينة الدراسة: تكونت عينة الدراسة من (١١٩٨) طالبا وطالبة بالصف الأول الإعدادي بالعام الدراسي ٢٠١٦/ ٢٠١٧ بالفصل الدراسي الأول بمحافظات القاهرة والجيزة والقليوبية. ويتضمن جدول (١) التالي أعداد الطلاب موزعين على الصور الاختبارية الثلاث.

جدول (١): عينة الدراسة موزعة على الصور الاختبارية الثلاث.

الصورة الاختبارية	المحافظة						المجموع	
	القاهرة		الجيزة		القليوبية			
	العدد	النسبة المئوية	العدد	النسبة المئوية	العدد	النسبة المئوية	العدد	النسبة المئوية
الأولى	١٧٠	١٤.١٩	١٠٨	٩.٠٢	١٢٢	١٠.١٨	٤٠٠	٣٣.٣٩
الثانية	١٥١	١٢.٦٠	١٣٥	١١.٢٧	١١٤	٩.٥٢	٤٠٠	٣٣.٣٩
الثالثة	١٤٧	١٢.٢٧	١٢٠	١٠.٠٢	١٣١	١٠.٩٣	٣٩٨	٣٣.٢٢
المجموع	٤٦٨	٣٩.٠٧	٣٦٣	٣٠.٣٠	٣٦٧	٣٠.٦٣	١١٩٨	١٠٠

أداة الدراسة: تمثلت أداة الدراسة في مستودع من المفردات مكون من (٧٥) مفردة تم صياغتها لتقيس نواتج التعلم الرئيسة بوحدة المادة وتركيبها بمادة العلوم للفصل الدراسي الأول بالصف الأول الإعدادي للعام الدراسي ٢٠١٦/ ٢٠١٧، نظرا لأهمية هذه الوحدة بالنسبة لمقرر العلوم بالمرحلة الإعدادية؛ حيث تعد الأساس للمفاهيم الكيميائية المهمة، بالإضافة إلى أن تخصص الباحثة بالدرجة الجامعية الأولى هو الطبيعة والكيمياء. وكانت الباحثة قد قامت ببناء صورة اختبارية على هذه

الوحدة مكونة من ٣٠ مفردة في دراسة سابقة لها؛ لذا فقد آثرت الباحثة إثراء هذه الصورة بمفردات جديدة تسمح بتكوين ثلاث صور اختبارية متكافئة تسمح باختبار طريقتي التعادل المستخدمة في الدراسة الحالية. وقد اتبعت الباحثة في بناء الصور الاختبارية الخطوات التالية:

- **مراجعة محتوى المقرر الدراسي المراد قياسه وتحليله إلى أوجه التعلم الرئيسة المتضمنة فيه:** وتناولت الدراسة الحالية محتوى وحدة المادة وتركيبها من محتوى مادة العلوم للصف الأول الإعدادي للعام الدراسي (٢٠١٦/ ٢٠١٧). مراجعة تحليل محتوى هذه الوحدة في ضوء البناء المعرفي للعلم متمثلا في: **المعرفة العلمية، المفاهيم، والتعميمات والمبادئ والقوانين والنظريات**، بغرض تحديد التوزيع النسبي لكل منها (ملحق ١).

- **مراجعة المعايير القومية للتعليم في جمهورية مصر العربية الخاصة بمادة العلوم:** حيث قامت الباحثة بالاطلاع على المعايير القومية للتعليم لمحتوى مادة العلوم للوقوف على أهم المعايير التي ينبغي أن تتحقق من خلال الوحدة محل الدراسة، وكذا مؤشرات تحقيق كل منها؛ ومن ثم إمكان قياسها إجرائيا. وبعد تحليل المحتوى إلى أوجه التعلم المتضمنة، وتحديد المعايير المراد تحقيقها ومؤشراتها من خلال **الوحدة محل الدراسة** كجزء من معايير العلوم في المرحلة الاعدادية، قامت الباحثة بترجمة هذه المؤشرات إجرائيا من خلال صياغتها إلى أهداف إجرائية تقيس جوانب التعلم المتضمنة في الوحدة محل الدراسة. وتغطي ثلاثة مستويات من مستويات النشاط العقلي وهي: المعرفة، والفهم، والتطبيق (ملحق ٢).

- **بناء المفردات الاختبارية:** وتضمن بناء المفردات الاختبارية عدة خطوات تمثلت في: تصنيف الأهداف الإجرائية التي تم التوصل إليها في شكلها النهائي، وبناء جدول مواصفات الاختبار تبعا لبعدي المحتوى والمستوى المعرفي. ثم تحديد عدد المفردات التي ترتبط بالأهداف في موضوعات الوحدة محل الدراسة. ويوضح جدول (٢) التالي أعداد المفردات التي تقيس تلك الأهداف موزعة على الصور الاختبارية.

جدول (٢): أعداد المفردات الاختبارية والمستويات المعرفية التي تقيسها موزعة على الصور الاختبارية.

الموضوع		المستوى المعرفي للهدف										المجموع
		المعرفة				الفهم				التطبيق		
		صورة١	صورة٢	صورة٣	مشترك	المجموع	صورة١	صورة٢	صورة٣	مشترك	المجموع	
تركيب الذرة	عدد الأسئلة	٤	٧	٥	٢	١٨	٤	٢	٢	١	٩	٤٤
خواص المادة	عدد الأسئلة	٣	٤	٦	٣	١٦	٣	٣	٤	١	١١	٣١
المجموع	عدد الأسئلة	٧	١١	١١	٥	٣٤	٧	٥	٦	٢	٢٠	٧٥
	النسبة	٩.٣	١٤.٧	١٤.٧	٦.٧	٤٥.٣	٩.٣	٦.٧	٨	٢.٧	٢٦.٧	١٠٠

		التطبيق				
		صورة١	صورة٢	صورة٣	مشترك	المجموع
تركيب الذرة	عدد الأسئلة	٤	٤	٦	٣	١٧
خواص المادة	عدد الأسئلة	٢	١	١	-	٤
المجموع	عدد الأسئلة	٦	٥	٧	٣	٢١
	النسبة	٨	٦.٧	٩.٣	٤	٢٨

وبالوصول إلى هذه المرحلة يكون قد توفر قدر من الصدق الملائم للمفردات الاختبارية أو ما يسمى بصدق المحتوى ويعني أن محتوى المفردات الاختبارية يعكس الأهمية النسبية للأهداف التعليمية ويعبر عنها تعبيراً صحيحاً.

-صياغة المفردات الاختبارية التي تقيس الأهداف السلوكية: في ضوء جدول المواصفات قامت الباحثة ببناء مجموعة من المفردات الاختبارية، وقد تم اختيار شكل المفردات من نوع **الاختيار من متعدد** ذات البدائل الأربعة يمثل أحدها الإجابة الصحيحة للمفردة.

-التحكيم على المفردات: بعد الانتهاء من صياغة المفردات تم عرضها على عدد من المحكمين، بغرض الوقوف على مدى صحتها العلمية وأسلوب صياغتها ومدى تغطيتها وشمولها للأهداف الإجرائية وملاءمتها للمستويات المعرفية المحددة. وقد تضمنت توجيهات المحكمون إجراء بعض التعديلات في صياغة سبع مفردات. وبعد إجراء التعديلات تكونت الصورة الأولية لمستودع الأسئلة الذي تسعى الدراسة الحالية إلى تدريجه من (٧٥) مفردة.

-صياغة التعليمات: تم صياغة تعليمات التطبيق بحيث تتضمن: الهدف من التطبيق، وتعليمات الإجابة على المفردات.

-**التجريب الاستطلاعي للمفردات:** قامت الباحثة بتجريب المفردات الاختبارية استطلاعياً بتطبيقها على مجموعة من التلاميذ بلغ عددها(٩٠) تلميذاً وتلميذة من تلاميذ الصف الأول الإعدادي بإدارة المقطم التعليمية للوقوف على مدى وضوح التعليمات، ووضوح الصياغة، وملاءمة البدائل، وكذلك حساب الزمن اللازم للإجابة على الاختبار وقد استغرق تطبيق الصورة الاختبارية الواحدة حوالي (٤٥) دقيقة تقريباً[٣].

تدريج الصور الاختبارية وفق طريقتي التعادل:

لتدريج الصور الاختبارية وفق طريقتي التعادل اتبعت الباحثة الخطوات التالية:

-**تطبيق الصور الاختبارية** على عينة التدريج.

-**ترميز البيانات:** حيث تم تحديد رمز لكل مفردة من مفردات الاختبار.

-**تجهيز ملفات الإدخال** الخاصة بكل صورة اختبارية وإدخال البيانات الخاصة بأداء أفراد عينة التدريج على المفردات المختلفة على الحاسب الآلي باستخدام الحزمة الإحصائية للعلوم الاجتماعية Statistical Package for Social Sciences (SPSS).

-**تصحيح المفردات** باستخدام برنامج الحاسب الآلي SPSS.

-**تحليل بيانات الصور الاختبارية باستخدام برنامج WINSTEPS:** وفقاً لنموذج راش البارامتري أحادي المعلم؛ بهدف تحديد واستبعاد البيانات غير الملائمة لأسس القياس الموضوعي وتدريج المفردات الملائمة على ميزان تدرج واحد بصفر مشترك واحد، وكذلك الحصول على تقديرات القدرة للأفراد وفق كل من طريقتي التدريج المتزامن والتدريج بتثبيت بارامترات المفردات المشتركة.

تدريج الصور الاختبارية الثلاث باستخدام طريقة تثبيت بارامترات المفردات المشتركة.

استخدم برنامج WINSTEPS لتحليل الصور الاختبارية وفق طريقة التعادل بتثبيت بارامترات المفردات المشتركة، وذلك وفق الخطوات التالية:

١-تحليل الصور الاختبارية الثلاث كل صورة على حدة؛ وذلك لحذف الأفراد غير الملائمين لأسس القياس الموضوعي، وكذلك تنقية الصور الاختبارية من المفردات غير الملائمة.

---

[٣] نظرا لأن الدراسة الحالية تتبنى تدريج الصور الاختبارية وفق نظرية الاستجابة للمفردة، فقد تم إرجاء عرض دلالات الصدق والثبات للصور الاختبارية بعد عرض نتائج التدريج.

٢- تحديد صورة من الصور الاختبارية الثلاث لتكون الصورة المرجعية للتدريج (الصفر المرجعي للتدريج)، وقد تم اعتبار الصورة الاختبارية الأولى **الصورة المرجعية** للتدريج وفق طريقة تثبيت بارامترات المفردات المشتركة.

٣-تحديد صعوبات المفردات المشتركة كما اشتقت من الصورة المرجعية واستخدامها في الحصول على التدريج النهائي لمفردات الصورتين الثانية والثالثة، ومن ثم ضمها جميعا على ميزان مشترك واحد بصفر مشترك واحد.

وفيما يلي عرض نتائج كل خطوة من هذه الخطوات:

أولا: تدريج الصورة الاختبارية الأولى (الإطار المرجعي): مر تدريج الصورة الاختبارية الأولى بالخطوات التالية:

(أ) حذف البيانات التامة والصفرية: يستبعد برنامج WINSTEPS آليا أثناء إجراء التحليل البيانات الخاصة بالأفراد الحاصلين على الدرجة النهائية أو الدرجة صفر على الاختبار، كما يستبعد البيانات الخاصة بكل مفردة التى أجمع كافة أفراد العينة على إجابتها سواء أكانت صواباً أو خطأ. ووفق هذا المحك لم يسفر التحليل الأولى للبيانات عن استبعاد أي من الأفراد أو مفردات هذه الصورة.

(ب) حذف الأفراد غير الملائمين لأسس القياس: بعد إجراء التحليل الأولى للبيانات باستخدام برنامج WINSTEPS تم تحديد وحذف الأفراد الذين تجاوزوا محكات الملاءمة الإحصائية التقاربية والتباعدية (±٢) وفق مؤشر (ZSTD) التي يوفرها البرنامج؛ حيث إن تجاوز الحد الفاصل (+٢) يعني أن نمط استجابات الفرد يختلف عما يتوقع منه. وقد يرجع ذلك إلى التسرع، أو الإهمال أو اللجوء إلى التخمين، أما تجاوز الحد الفاصل(-٢) فيعني أن نمط استجابات الفرد متسق بدرجة غير واقعية. وقد يرجع ذلك إلى البطء أو الحرص الشديد، أو اللجوء إلى الغش.

(ج) حذف المفردات غير الملائمة لأسس القياس: أعيد التحليل بعد حذف الأفراد غير الملائمين لأسس القياس الموضوعي، بهدف تحديد وحذف المفردات التي تجاوزت محكات الملاءمة الإحصائية التقاربية والتباعدية (± ٢) وفق مؤشر (ZSTD) التي يوفرها برنامج WINSTEPS. ويعني تجاوز الحد الفاصل (+٢) من تدريج الملاءمة عيباً في صياغة المفردة، أو أنها تعرف متغيراً آخراً يختلف عن المتغير الذي تعرفه باقي المفردات، أما تجاوز الحد الفاصل (-٢) يعني أن المفردة قد تعتمد على غيرها من المفردات أو أنها تقيس متغيراً آخر شديد الارتباط بالمتغير محل القياس.

ويتضمن جدول (٣) التالي ملخص نتائج تدريج الصورة الاختبارية الأولى (الإطار المرجعي).

جدول (٣): ملخص نتائج تدريج الصورة الاختبارية الأولى (الإطار المرجعي)

التحليل	عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
			الصعوبة	القدرة	الصعوبة	القدرة
الأول	٣٠	٤٠٠	٠.٠٠	٠.١٤-	٠.٩٧	٠.٧٨
الثاني	٣٠	٣٤٦	٠.٠٠	٠.٠٤-	٠.٩٨	٠.٧٧
الثالث	٢٨	٣٤٦	٠.٠٠	٠.٠١	٠.٩٨	٠.٧٨

يتضح من جدول (٣) ما يلي:

- حذفت مفردتان غير ملائمتين لأسس القياس الموضوعي. كما حذف ٥٤ فردا غير ملائم.
- هناك تقارب شديد بين متوسط تقديرات الأفراد ومتوسط تقديرات القدرة؛ حيث بلغ متوسط تقدير القدرة في التحليل النهائي ٠.٠١ لوجيت.
- تمتعت تقديرات صعوبات المفردات بدرجة ثبات عالية، ازدادت بعد حذف المفردات غير الملائمة مما يشير إلى دقة التدريج.
- تمتعت تقديرات قدرات الأفراد بدرجة ثبات عالية، ازدادت أيضا بعد حذف المفردات غير الملائمة.

التدريج النهائي لمفردات الصورة الاختبارية الأولى (الإطار المرجعي للتدريج): يتضمن جدول (٤) التالي التدريج النهائي لمفردات الصورة الاختبارية الأولى (الإطار المرجعي) مقدرة بوحدتي اللوجيت والمنف.

جدول (٤): التدريج النهائي لمفردات الصورة الاختبارية الأولى مقدرة بوحدتي اللوجيت والمنف.

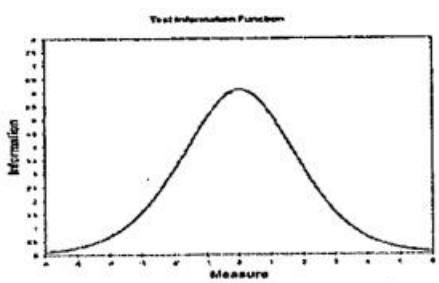
كود المفردة	رقم المفردة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف
COM2	٦	٢.٢٤-	٣٨.٨٠	٠.١٧	٠.٨٥
COM4	٨	١.٤٩-	٤٢.٥٥	٠.١٤	٠.٧٠
COM45	٢١	١.٠١-	٤٤.٩٥	٠.١٣	٠.٦٥
COM10	١٤	٠.٨٦-	٧٠.٤٥	٠.١٢	٠.٦٠
COM44	٢٠	٠.٨٤-	٤٥.٨٠	٠.١٢	٠.٦٠
COM37	٤	٠.٦٣-	٤٦.٨٥	٠.١٢	٠.٦٠
COM5	٩	٠.٥٥-	٤٧.٢٥	٠.١٢	٠.٦٠
COM6	١٠	٠.٣٩-	٤٨.٠٥	٠.١٢	٠.٦٠
COM49	٢٥	٠.١٧-	٤٩.١٥	٠.١٢	٠.٦٠
COM38	٥	٠.١٦-	٤٩.٢٠	٠.١٢	٠.٦٠
COM3	٧	٠.١١-	٤٩.٤٥	٠.١٢	٠.٦٠
COM48	٢٤	٠.١٠-	٤٩.٥٠	٠.١٢	٠.٦٠
COM50	٢٦	٠.٠٧-	٤٩.٦٥	٠.١٢	٠.٦٠

كود المفردة	رقم المفردة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف
COM1	٣	-٠.٠٦	٤٩.٧٠	٠.١٢	٠.٦٠
COM40	١٦	٠.٠٠	٥٠.٠٠	٠.١٢	٠.٦٠
COM41	١٧	٠.٠٠	٥٠.٠٠	٠.١٢	٠.٦٠
COM53	٢٩	-٠.٠٨	٥٠.٤٠	٠.١٢	٠.٦٠
COM47	٢٣	٠.١١	٥٠.٥٥	٠.١٢	٠.٦٠
COM8	١٢	٠.٢٦	٥١.٣٠	٠.١٢	٠.٦٠
COM42	١٨	٠.٣٢	٥١.٦٠	٠.١٢	٠.٦٠
COM9	١٣	٠.٥٢	٥٢.٦٠	٠.١٢	٠.٦٠
COM51	٢٧	٠.٦٨	٥٣.٤٠	٠.١٢	٠.٦٠
COM54	٣٠	٠.٧٠	٥٣.٥٠	٠.١٢	٠.٦٠
COM39	١٥	٠.٨٧	٥٤.٣٥	٠.١٣	٠.٦٥
COM36	٢	٠.٩٥	٥٤.٧٥	٠.١٣	٠.٦٥
COM7	١١	١.١٥	٥٥.٧٥	٠.١٣	٠.٦٥
COM52	٢٨	١.٣٥	٥٦.٧٥	٠.١٤	٠.٧٠
COM46	٢٢	١.٦٩	٥٨.٤٥	٠.١٥	٠.٧٥

يتضح من جدول (٤) ما يلي:

- تراوح تقدير صعوبة مفردات الصورة الاختبارية الأولى (الإطار المرجعي) بين ٣٨.٨: ٥٨.٤٥ منف، كما تراوحت قيم الخطأ المعياري لتقديرات الصعوبة بين ٠.٦٠: ٠.٨٥ منف.
- يوضح شكل (٢) التالي دالة معلومات الصورة الاختبارية الأولى وكذلك خريطة الاختبار لهذه الصورة.

خريطة مفردات الصورة الاختبارية الأولى.



دالة معلومات الصورة الاختبارية الأولى.

شكل (٢): دالة المعلومات وخريطة الاختبار للصورة الاختبارية الأولى.

يتضح من شكل (٢) السابق ما يلي:

- تغطي مفردات الصورة الأولى مدى القياس بشكل جيد حيث جاءت تقديرات قدرة الأفراد قريبة جدا من صعوبات الأسئلة.
- تعطى الصورة الاختبارية الأولى كما مناسبا من المعلومات خاصة عند المستوى المتوسط للقدرة؛ حيث بلغت أقصى قيمة لدالة المعلومات (٦.٠٩) عند مستوى قدرة (٠.١٢) لوجيت.

تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الأولى (الإطار المرجعي):

جدول (٥): تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الأولى.

الدرجة الخام	تقدير القدرة باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري لتقدير القدرة باللوجيت	الخطأ المعياري لتقدير القدرة بالمنف
0	٤.٩١-	٢٥.٤٥	٨٥.١	٩.٢٥
1	٣.٦٤-	٣١.٨٠	١.٠٤	٥.٢٠
2	٢.٨٧-	٣٥.٦٥	٠.٧٦	٣.٨٠
3	٢.٣٩-	٣٨.٠٥	٠.٦٤	٣.٢٠
4	٢.٠٣-	٣٩.٨٥	٠.٥٧	٢.٨٥
5	١.٧٣-	٤١.٣٥	٠.٥٢	٢.٦٠
6	١.٤٧-	٤٢.٦٥	٠.٤٩	٢.٤٥
7	١.٢٥-	٤٣.٧٥	٠.٤٧	٢.٣٥
8	١.٠٤-	٤٤.٨٠	٠.٤٥	٢.٢٥
9	٠.٨٥-	٤٥.٧٥	٠.٤٣	٢.١٥
10	٠.٦٦-	٤٦.٧٠	٠.٤٢	١٠.٢
11	٠.٤٩-	٤٧.٥٥	٠.٤١	٢.٠٥
12	٠.٣٢-	٤٨.٤٠	٠.٤١	٢.٠٥
13	٠.١٥-	٤٩.٢٥	٠.٤١	٢.٠٥
14	٠.٠١	٥٠.٠٥	٠.٤٠	٢.٠٠
15	٠.١٨	٥٠.٩٠	٠.٤١	٢.٠٥
16	٠.٣٤	٥١.٧٠	٠.٤١	٢.٠٥
17	٠.٥١	٥٢.٥٥	٠.٤١	٢.٠٥
18	٠.٦٨	٥٣.٤٠	٠.٤٢	٢.١٠
19	٠.٨٦	٥٤.٣٠	٠.٤٣	٢.١٥
20	١.٠٥	٥٥.٢٥	٠.٤٤	٢.٢٠

الدرجة الخام	تقدير القدرة باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري لتقدير القدرة باللوجيت	الخطأ المعياري لتقدير القدرة بالمنف
21	١.٢٦	٥٦.٣٠	٠.٤٦	٢.٣٠
22	١.٤٨	٥٧.٤٠	٠.٤٨	٢.٤٠
23	١.٧٣	٥٨.٦٥	٠.٥٢	٢.٦٠
24	٢.٠٢	٦٠.١٠	٠.٥٦	٢.٨٠
25	٢.٣٧	٦١.٨٥	٠.٦٣	٣.١٥
26	٢.٨٤	٦٤.٢٠	٠.٧٥	٣.٧٥
27	٣.٥٩	٦٧.٩٥	١.٠٣	٥.١٥
28	٤.٨٤	٧٤.٢٠	١.٨٤	٩.٢٠

يتضح من جدول (٥) أن مدى تقديرات القدرة المقابل لكل درجة خام على الصورة الاختبارية الأولى امتد بين (4.91-) لوجيت إلى (٤.٨٤) لوجيت.

ثانيا: **تدريج الصورة الاختبارية الثانية:** مر تدريج الصورة الاختبارية الثانية وفق طريقة التدريج بتثبيت بارامترات المفردات المشتركة بأربع خطوات، شملت الخطوات الثلاث الأولى منها نفس خطوات تدريج الصورة الاختبارية الأولى/ الإطار المرجعي من حيث حذف الدرجات التامة والصفرية، وحذف الأفراد والمفردات غير الملائمة لأسس القياس. ثم تأتي **الخطوة الرابعة** وتتمثل في ربط مفردات الصورة الاختبارية الثانية بتدريج مفردات الصورة الاختبارية الأولى/ الإطار المرجعي باستخدام صعوبات المفردات المشتركة المستخلصة من التدريج النهائي للصورة الأولى. وفيما يلي عرض نتائج كل خطوة من هذه الخطوات على حدة.

- **حذف البيانات التامة والصفرية:** لم يسفر التحليل وفق هذه الخطوة عن حذف أي من المفردات أو الأفراد.
- **حذف الأفراد غير الملائمين لأسس القياس:** حذف ٦٤ فردا غير ملائم لأسس القياس الموضوعي.
- **حذف المفردات غير الملائمة لأسس القياس:** حذفت (٥) مفردات غير ملائمة لأسس القياس الموضوعي.

ويتضمن جدول (٦) التالي ملخص نتائج المراحل الثلاث لتدريج الصورة الاختبارية الثانية.

جدول (٦): ملخص نتائج تدريج الصورة الاختبارية الثانية.

التحليل	عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
			الصعوبة	القدرة	الصعوبة	القدرة
الأول	٣١	٤٠٠	٠.٠٠	٠.٥٢	٠.٩٧	٠.٨٨
الثاني	٣١	٣٣٦	٠.٠٠	٠.٥٦	٠.٩٧	٠.٨٧
الثالث	٢٦	٣٣٦	٠.٠٠	٠.٦٤	٠.٩٧	٠.٨٥

- **تعادل الصورة الاختبارية الثانية مع الصورة الاختبارية الأولى (التدريج المرجعي):** وذلك بهدف ضم مفردات الصورة الاختبارية الثانية الملائمة إلى تدريج مفردات الإطار المرجعي ووضعها على نفس الصفر؛ حيث تم إعادة تحليل الصورة الثانية بعد الحصول على تقديرات المفردات المشتركة من الصورة الاختبارية الأولي (الإطار المرجعي) وإعادة التحليل. ويتضمن جدول (٧) التالي ملخص نتائج التحليل النهائي للصورة الاختبارية الثانية بعد ضمها لتدريج الصورة الاختبارية الأولى.

جدول (٧): ملخص نتائج تحليل الصورة الاختبارية الثانية بعد ضمها إلى تدريج الصورة المرجعية.

التحليل	عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
			الصعوبة	القدرة	الصعوبة	القدرة
الرابع للصورة الثانية	٢٦	٣٣٦	-٠.٠١	٠.٦٣	٠.٩٧	٠.٨٥

يلاحظ من جدول (٦)، وجدول (٧) ما يلي:

- تمتعت تقديرات المفردات الاختبارية، وتقديرات قدرة الأفراد على الصورة الاختبارية الثانية بدرجة ثبات مرتفعة حيث جاءت قيمة معامل ثبات تقديرات صعوبة المفردات (٠.٩٧)، وبلغت (٠.٨٥) بالنسبة لتقديرات قدرات الأفراد.
- **التدريج النهائي لمفردات الصورة الاختبارية الثانية بعد ضمها للإطار المرجعي للتدريج:** يتضمن جدول (٨) التالي التدريج النهائي لمفردات الصورة الاختبارية الثانية بعد وضعها على نفس تدريج الإطار المرجعي مقدرة بوحدتي اللوجيت والمنف.

جدول (٨): التدريج النهائي لمفردات الصورة الاختبارية الثانية مقدرة بوحدتي اللوجيت والمنف.

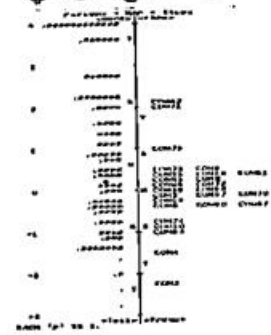
كود المفردة	رقم المفردة بالصورة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري لتقدير الصعوبة باللوجيت	الخطأ المعياري لتقدير القدرة بالمنف
COM2	٦	٢.٢٤-	٣٨.٨٠	٠.٢٠	١.٠٠
COM4	٨	١.٤٩-	٤٢.٥٥	٠.١٧	٠.٨٥

كود المفردة	رقم المفردة بالصورة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري لتقدير الصعوبة باللوجيت	الخطأ المعياري لتقدير القدرة بالمنف
COM63	١٩	-٠.٩٧	٤٥.١٥	٠.١٥	٠.٧٥
COM10	١٤	-٠.٨٦	٤٥.٧٠	٠.١٥	٠.٧٥
COM74	٣٠	-٠.٧٥	٤٦.٢٥	٠.١٥	٠.٧٥
COM67	٢٣	-٠.٤٣	٤٧.٨٥	٠.١٤	٠.٧٠
COM6	١٠	-٠.٣٩	٤٨.٠٥	٠.١٤	٠.٧٠
COM60	١٦	-٠.٣٩	٤٨.٠٥	٠.١٤	٠.٧٠
COM64	٢٠	-٠.٢٠	٤٩.٠٠	٠.١٤	٠.٧٠
COM70	٢٦	-٠.١٨	٤٩.١٠	٠.١٤	٠.٧٠
COM57	٤	-٠.١٢	٤٩.٤٠	٠.١٤	٠.٧٠
COM3	٧	-٠.١١	٤٩.٤٥	٠.١٤	٠.٧٠
COM68	٢٤	-٠.٠١	٤٩.٩٥	٠.١٤	٠.٧٠
COM59	١٥	٠.٠٥	٥٠.٢٥	٠.١٤	٠.٧٠
COM66	٢٢	٠.٠٧	٥٠.٣٥	٠.١٤	٠.٧٠
COM72	٢٨	٠.١٧	٥٠.٨٥	٠.١٤	٠.٧٠
COM65	٢١	٠.٢١	٥١.٠٥	٠.١٤	٠.٧٠
COM8	١٢	٠.٢٦	٥١.٣٠	٠.١٤	٠.٧٠
COM55	١	٠.٣٦	٥١.٨٠	٠.١٤	٠.٧٠
COM58	٥	٠.٣٨	٥١.٩٠	٠.١٤	٠.٧٠
COM61	١٧	٠.٣٨	٥١.٩٠	٠.١٤	٠.٧٠
COM73	٢٩	٠.٤٤	٥٢.٢٠	٠.١٤	٠.٧٠
COM9	١٣	٠.٥٢	٥٢.٦٠	٠.١٤	٠.٧٠
COM75	٣١	٠.٩٥	٥٤.٧٥	٠.١٤	٠.٧٠
COM71	٢٧	١.٩٧	٥٩.٨٥	٠.١٥	٠.٧٥
COM62	١٨	٢.١٦	٦٠.٨٠	٠.١٦	٠.٨٠

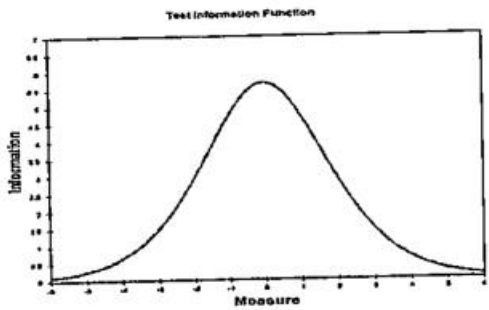
يتضح من جدول (٨) ما يلي:

- تراوح تقدير الصعوبة بين٣٨.٨: ٦٠.٨ منف، وتراوحت قيم الخطأ المعياري لتقديرات الصعوبة بين٠.٧: ١ منف.
- يوضح شكل (٣) التالي دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثانية بعد

ضمها للتدريج المرجعي.



خريطة مفردات الصورة الاختبابة الثانية



دالة معلومات الصورة الاختبابة الثانية

شكل (٣): دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثانية بعد ضمها للتدريج المرجعي.

يتضح من شكل (٣) السابق ما يلي:

- تغطي مفردات الصورة الثانية مدى القياس بشكل جيد حيث جاءت تقديرات قدرة الأفراد قريبة جدا من صعوبات الأسئلة. مع وجود بعض الفجوات خاصة عند أطراف التدريج وهذه الفجوات يقل وجودها بعد ضم مفردات الصور الاختبارية الثلاث على ميزان تدرج واحد بصفر مشترك واحد.
- توفر الصورة الاختبارية الثانية بعد ضمها للتدريج المرجعي كما مناسبا من المعلومات خاصة عند المستوى المتوسط للقدرة؛ حيث بلغت أقصى قيمة لدالة المعلومات (٥.٧٦) عند مستوى قدرة (-٠.٠٦) لوجيت.

تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثانية بعد ضمها للإطار المرجعي:

جدول (٩): تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثانية بعد ضمها للتدريج المرجعي.

الدرجة	تقدير القدرة باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف
0	٤.٨٥-	٢٥.٧٥	١.٨٥	٩.٢٥
1	٣.٥٨-	٣٢.١٠	١.٠٤	٥.٢٠
2	٢.٨٠-	٣٦.٠٠	٠.٧٦	٣.٨٠

الدرجة	تقدير القدرة باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف
3	٢.٣٢-	٣٨.٤٠	٠.٦٤	٣.٢٠
4	١.٩٥-	٤٠.٢٥	٠.٥٧	٢.٨٥
5	١.٦٥-	٤١.٧٥	٠.٥٣	٢.٦٥
6	١.٣٩-	٤٣.٠٥	٠.٤٩	٢.٤٥
7	١.١٦-	٤٤.٢٠	٠.٤٧	٢.٣٥
8	٠.٩٤-	٤٥.٣٠	٠.٤٥	٢.٢٥
9	٠.٧٤-	٤٦.٣٠	٠.٤٤	٢.٢٠
10	٠.٥٥-	٤٧.٢٥	٠.٤٣	٢.١٥
11	٠.٣٧-	٤٨.١٥	٠.٤٢	٢.١٠
12	٠.١٩-	٤٩.٠٥	٠.٤٢	٢.١٠
13	٠.٠١-	٤٩.٩٥	٠.٤٢	٢.١٠
14	٠.١٦	٥٠.٨٠	٠.٤٢	٢.١٠
15	٠.٣٤	٥١.٧٠	٠.٤٣	٢.١٥
16	٠.٥٣	٥٢.٦٥	٠.٤٣	٢.١٥
17	٠.٧٢	٥٣.٦٠	٠.٤٤	٢.٢٠
18	٠.٩٢	٥٤.٦٠	٠.٤٦	٢.٣٠
19	١.١٣	٥٥.٦٥	٠.٤٧	٢.٣٥
20	١.٣٧	٥٦.٨٥	٠.٥٠	٢.٥٠
21	١.٦٣	٥٨.١٥	٠.٥٣	٢.٦٥
22	١.٩٤	٥٩.٧٠	٠.٥٨	٢.٩٠
23	٢.٣١	٦١.٥٥	٠.٦٥	٣.٢٥
24	٢.٨٠	٦٤.٠٠	٠.٧٧	٣.٨٥
25	٣.٥٩	٦٧.٩٥	١.٠٥	٥.٢٥
26	٤.٨٦	٧٤.٣٠	١.٨٥	٩.٢٥

يتضح من جدول (٩) أن مدى تقديرات القدرة المقابل لكل درجة خام على الصورة الاختبارية الثانية امتد بين (-٤.٨٥) لوجيت إلى (٤.٨٦) لوجيت.

ثالثا: **تدريج الصورة الاختبارية الثالثة:** مر تدريج الصورة الاختبارية الثالثة وفق طريقة التدريج بتثبيت بارامترات المفردات المشتركة بأربع خطوات، شملت الخطوات الثلاث الأولى منها نفس خطوات تدريج الصورة الأولى والثانية؛ من حيث حذف الدرجات التامة والصفرية، وحذف الأفراد والمفردات غير الملائمة لأسس القياس. ثم تأتي **الخطوة الرابعة** وتتمثل في ربط مفردات

الصورة الاختبارية الثالثة بتدريج مفردات الصورة الاختبارية الأولى (التدريج المرجعي) باستخدام صعوبات المفردات المشتركة المستخلصة من التدريج النهائي للصورة الأولى. وفيما يلي عرض نتائج كل خطوة من هذه الخطوات على حدة.

- **حذف البيانات التامة والصفرية:** لم يسفر التحليل وفق هذه الخطوة عن حذف أي من المفردات أو الأفراد.

- **حذف الأفراد غير الملائمين لأسس القياس:** حذف (٩٢) فردا غير ملائم لأسس القياس الموضوعي.

- **حذف المفردات غير الملائمة لأسس القياس:** حذفت (٤) مفردات غير ملائمة لأسس القياس الموضوعي.

ويتضمن جدول (١٠) التالي ملخص نتائج المراحل الثلاث لتدريج الصورة الثالثة.

جدول (١٠): ملخص نتائج تدريج الصورة الثالثة.

التحليل	عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
			الصعوبة	القدرة	الصعوبة	القدرة
الأول	٣٤	٣٩٨	٠.٠٠٠	٠.٤٦	٠.٩٩	٠.٨٣
الثاني	٣٤	٣٠٦	٠.٠٠٠	٠.٧٢	٠.٩٩	٠.٨٣
الثالث	٣٠	٣٠٦	٠.٠٠٠	٠.٩٥	٠.٩٨	٠.٨١

- **تعادل الصورة الاختبارية الثالثة مع الصورة الاختبارية الأولى (التدريج المرجعي):** وذلك بهدف ضم مفردات الصورة الاختبارية الثالثة الملائمة إلى تدريج مفردات الإطار المرجعي ووضعها على نفس الصفر؛ حيث تم إعادة تحليل الصورة الاختبارية الثالثة بعد الحصول على تقديرات المفردات المشتركة من الصورة الاختبارية الأولى (الإطار المرجعي) وإعادة التحليل مرة ثانية. ويتضمن جدول (١١) التالي ملخص نتائج التحليل النهائي للصورة الاختبارية الثالثة بعد ضمها للصورة الاختبارية الأولى على نفس التدريج.

جدول (١١)

ملخص نتائج تحليل الصورة الاختبارية الثالثة بعد ضمها إلى التدريج المرجعي.

التحليل	عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
			الصعوبة	القدرة	الصعوبة	القدرة
الرابع للصورة الثالثة	٣٠	٣٠٦	-٠.٢٢	٠.٧	٠.٩٨	٠.٨٠

يلاحظ من جدول (١٠)، وجدول (١١) ما يلي:

- تمتعت تقديرات المفردات الاختبارية، وتقديرات قدرة الأفراد على الصورة الاختبارية الثالثة بدرجة ثبات مرتفعة؛ حيث جاءت قيمة معامل ثبات تقديرات صعوبة المفردات

(٠.٩٨)، وبلغت (٠.٨٠) بالنسبة لتقديرات قدرات الأفراد.

- **التدريج النهائي لمفردات الصورة الاختبارية الثالثة بعد ضمها للإطار المرجعي للتدريج:**

يتضمن جدول (١٢) التالي التدريج النهائي لمفردات الصورة الاختبارية الثالثة بعد وضعها على نفس تدريج الإطار المرجعي مقدرة بوحدتي اللوجيت والمنف.

جدول (١٢)

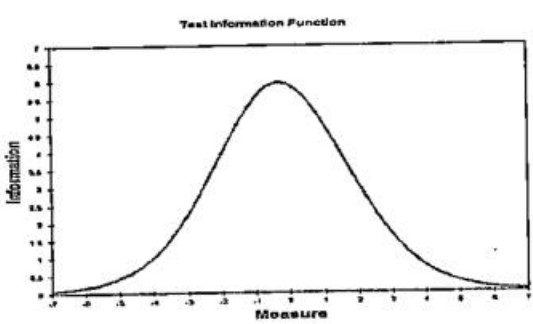
التدريج النهائي لمفردات الصورة الاختبارية الثالثة مقدرة بوحدتي اللوجيت والمنف.

كود المفردة	رقم المفردة في الصورة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف
COM2	٦	٢.٢٤-	٣٨.٨٠	٠.٢٢	١.١٠
COM22	٣٢	٢.١٧-	٣٩.١٥	٠.٢٢	١.١٠
COM28	١٨	٢.٠٠-	٤٠.٠٠	٠.٢٠	١.٠٠
COM4	٨	١.٤٩-	٤٢.٥٥	٠.١٧	٠.٨٥
COM17	٢٧	١.٣٥-	٤٣.٢٥	٠.١٧	٠.٨٥
COM25	١٥	١.١٩-	٤٤.٠٥	٠.١٦	٠.٨٠
COM27	١٧	١.٠٤-	٤٤.٨٠	٠.١٦	٠.٨٠
C0M19	٢٩	٠.٩٩-	٤٥.٠٥	٠.١٥	٠.٧٥
COM10	١٤	٠.٨٦-	٤٥.٧٠	٠.١٥	٠.٧٥
COM18	٢٨	٠.٨٣-	٤٥.٨٥	٠.١٥	٠.٧٥
COM12	٢	٠.٨٠-	٤٦.٠٠	٠.١٥	٠.٧٥
COM30	٢٠	٠.٦٥-	٤٦.٧٥	٠.١٤	٠.٧٠
COM21	٣١	٠.٤٥-	٤٧.٧٥	٠.١٤	٠.٧٠
COM6	١٠	٠.٣٩-	٤٨.٠٥	٠.١٤	٠.٧٠
COM16	٢٦	٠.٣٤-	٤٨.٣٠	٠.١٤	٠.٧٠
COM26	١٦	٠.١٢-	٤٩.٤٠	٠.١٣	٠.٦٥
COM3	٧	٠.١١-	٤٩.٤٥	٠.١٣	٠.٦٥
COM33	٢٣	٠.١٠-	٤٩.٥٠	٠.١٣	٠.٦٥
COM1	٣	٠.٠٦-	٤٩.٧٠	٠.١٣	٠.٦٥
COM31	٢١	٠.٢٤	٥١.٢٠	٠.١٣	٠.٦٥
COM8	١٢	٠.٢٦	٥١.٣٠	٠.١٣	٠.٦٥
COM23	٣٣	٠.٢٩	٥١.٤٥	٠.١٣	٠.٦٥
COM29	١٩	٠.٣٦	٥١.٨٠	٠.١٣	٠.٦٥

كود المفردة	رقم المفردة في الصورة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف
COM9	١٣	٠.٥٢	٥٢.٦٠	٠.١٣	٠.٦٥
COM11	١	٠.٦١	٥٣.٠٥	٠.١٣	٠.٦٥
COM7	١١	١.١٥	٥٥.٧٥	٠.١٣	٠.٦٥
COM14	٥	١.٣٩	٥٦.٩٥	٠.١٤	٠.٧٠
COM34	٢٤	١.٣٩	٥٦.٩٥	٠.١٤	٠.٧٠
COM20	٣٠	١.٨٤	٥٩.٢٠	٠.١٤	٠.٧٠
COM32	٢٢	٢.٤٨	٦٢.٤٠	٠.١٦	٨٠.٠

يتضح من جدول (١٢) ما يلي:

- تراوحت تقديرات صعوبات مفردات الصورة الثالثة بعد ضمها لتدريج الصورة الأولى بين ٣٨.٨: ٦٢.٤ منف، وتراوحت الأخطاء المعيارية لتقديرات الصعوبة بين ٠.٦٥: ١.١ منف.
- يوضح شكل (٤) التالي دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثالثة بعد ضمها للتدريج المرجعي.



دالة معلومات الصورة الاختبارية الثالثة.

خريطة مفردات الصورة الاختبارية الثالثة.

شكل (٤): دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثالثة بعد ضمها للتدريج المرجعي.

يتضح من شكل (٤) ما يلي:

- تغطي مفردات الصورة الاختبارية الثالثة مدى القياس بشكل جيد حيث جاءت تقديرات قدرة الأفراد قريبة من صعوبات الأسئلة. مع وجود بعض الفجوات خاصة عند أطراف التدريج وهذه الفجوات يقل وجودها بعد ضم مفردات الصور الاختبارية الثلاث على

ميزان تدرج واحد بصفر مشترك واحد.

- تعطى الصورة الاختبارية الثالثة بعد ضمها للتدريج المرجعي كما مناسبا من المعلومات خاصة عند المستوى المتوسط للقدرة حيث بلغت أقصى قيمة لدالة المعلومات (٥.٩٦) عند مستوى قدرة (-٠.٢٨) لوجيت.

تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثالثة بعد ضمها للإطار المرجعي

جدول (١٣)

تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثالثة بعد ضمها للتدريج المرجعي.

الدرجة	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
0	٥.٣٦-	١.٨٥	٢٣.٢٠	٩.٢٥
1	٤.١١-	١.٠٤	٢٩.٤٥	٢٠.٥
2	٣.٣٤-	٠.٧٦	٣٣.٣٠	٣.٨٠
3	٢.٨٦-	٠.٦٤	٣٥.٧٠	٣.٢٠
4	٢.٥٠-	٠.٥٧	٣٧.٥٠	٢.٨٥
5	٢.٢٠-	٠.٥٣	٣٩.٠٠	٢.٦٥
6	١.٩٤-	٠.٤٩	٤٠.٣٠	٢.٤٥
7	١.٧١-	٠.٤٧	٤١.٤٥	٢.٣٥
8	١.٥٠-	٠.٤٥	٤٢.٥٠	٢.٢٥
9	١.٣٠-	٠.٤٤	٤٣.٥٠	٢.٢٠
10	١.١١-	٠.٤٣	٤٤.٤٥	٢.١٥
11	٠.٩٣-	٠.٤٢	٤٥.٣٥	٢.١٠
12	٠.٧٥-	٠.٤٢	٤٦.٢٥	٢.١٠
13	٠.٥٨-	٠.٤١	٤٧.١٠	٢.٠٥
14	٠.٤١-	٠.٤١	٤٧.٩٥	٢.٠٥
15	٠.٢٥-	٠.٤١	٤٨.٧٥	٢.٠٥
16	٠.٠٨-	٠.٤١	٤٩.٦٠	٢.٠٥
17	٠.٠٩	٠.٤١	٥٠.٤٥	٢.٠٥
18	٠.٢٧	٠.٤٢	٥١.٣٥	٢.١٠
19	٠.٤٤	٠.٤٣	٥٢.٢٠	٢.١٥
20	٠.٦٣	٠.٤٤	٥٣.١٥	٢.٢٠
21	٠.٨٢	٠.٤٥	٥٤.١٠	٢.٢٥
22	١.٠٣	٠.٤٦	٥٥.١٥	٢.٣٠
23	١.٢٥	٠.٤٨	٥٦.٢٥	٢.٤٠
24	١.٤٩	٠.٥١	٥٧.٤٥	٢.٥٥
25	١.٧٧	٠.٥٤	٥٨.٨٥	٢.٧٠
26	٢.٠٨	٠.٥٩	٦٠.٤٠	٢.٩٥
27	٢.٤٦	٠.٦٥	٦٢.٣٠	٣.٢٥
28	٢.٩٦	٠.٧٧	٦٤.٨٠	٣.٨٥

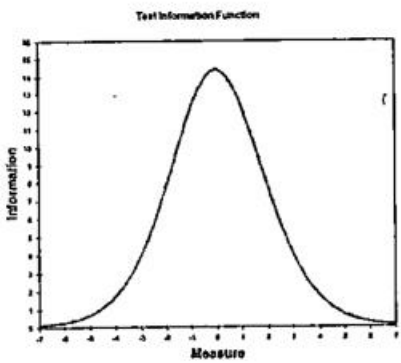
الدرجة	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
29	٣.٧٦	١.٠٥	٦٨.٨٠	٥.٢٥
30	٥.٠٤	١.٨٥	٧٥.٢٠	٩.٢٥

يتضح من جدول (١٣) أن تقديرات القدرة المقابلة لكل درجة خام كلية محتملة على الصورة الاختبارية الثالثة بعد ضمها للتدريج المرجعي تراوحت بين (٢٣.٢) إلى (٧٥.٢) منف.

<u>تعليق على تدريج الصور الثلاث بطريقة تثبيت بارامترات المفردات المشتركة:</u>

<u>بالنسبة لتوزيع المفردات على متصل الصعوبة:</u> يوضح شكل (٥) التالي خريطة الاختبار، ودالة معلومات الصور الاختبارية الثلاث معا.

خريطة مفردات الاختبار للصور الاختبارية الثلاثة معا.



دالة معلومات للصور الاختبارية الثلاثة معا.

شكل (٥): خريطة الاختبار ودالة معلومات الصور الاختبارية الثلاث معا.

يتضح من شكل (٥) ما يلي:

- عند استخلاص خريطة الاختبار ودالة المعلومات للصور الاختبارية الثلاث امتد توزيع تقديرات الصعوبة على طرفي متصل القياس. كما أن قدرات الأفراد تتناسب بشكل كبير مع صعوبات المفردات وتغطي صعوبات المفردات متصل الصعوبة بشكل جيد. حيث لا توجد فجوات حقيقية عدا عند أطراف متصل الصعوبة.
- تقدم الصور الاختبارية الثلاث معا كما مناسبا من المعلومات خاصة عند المستويات المتوسطة للقدرة، حيث بلغت قيمة دالة المعلومات (١٤.٣٨-) عند مستوى قدرة (٠.٠٧-).

<u>بالنسبة لتعادل القياس على الصور الاختبارية الثلاث:</u>

- بانتهاء تدريج الصورة الاختبارية الثالثة أصبحت كافة المفردات الملائمة لأسس القياس الموضوعي على ميزان تدرج واحد بصفر مشترك واحد، وأصبحت الصور الثلاث متعادلة

القياس بحيث لا يختلف تقدير قدرة الفرد باختلاف الصورة الاختبارية التي يتعرض لها. كما أن هذه الصور تكون متعادلة مع أي صورة فرعية أخرى يمكن سحبها من البنك الكلي؛ حيث يمكن سحب صور اختبارية أخرى متعادلة القياس من البنك الكلي مكونة من مفردات اختبارية من كل من الصور الثلاث ومع ذلك تكون الصور الجديدة متعادلة القياس مع الصورة الاختبارية الأصلية ومع أي صورة اختبارية أخرى تسحب من التدريج الكلي.

- وفيما يلي مثال لتعادل التقديرات على الصور الاختبارية الثلاث بعد تدريجها على ميزان تدرج واحد بصفر مشترك واحد.

جدول (١٤)

نماذج لتعادل القياس على الصور الاختبارية الثلاث بعد معادلتها.

الدرجة الخام	الصورة الاختبارية الأولى		الصورة الاختبارية الثانية		الصورة الاختبارية الثالثة	
	تقدير القدرة بالمنف	الخطأ المعياري لتقدير القدرة بالمنف	تقدير القدرة بالمنف	الخطأ المعياري بالمنف	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
3	٣٨.٠٥	٣.٢٠	٣٨.٤٠	٣.٢٠	٣٥.٧٠	٣.٢٠
4	٣٩.٨٥	٢.٨٥	٤٠.٢٥	٢.٨٥	٣٧.٥٠	٢.٨٥
5	٤١.٣٥	٢.٦٠	٤١.٧٥	٢.٦٥	٣٩.٠٠	٦٥.٢
10	٤٦.٧٠	٢.١٠	٤٧.٢٥	٢.١٥	٤٤.٤٥	٢.١٥
15	٥٠.٩٠	٢.٠٥	٥١.٧٠	٢.١٥	٤٨.٧٥	٢.٠٥
20	٥٥.٢٥	٢.٢٠	٥٦.٨٥	٢.٥٠	٥٣.١٥	٢.٢٠

يتضح من جدول (١٤) ما يلي:

- الصور الاختبارية الثلاث متقاربة في مستوى صعوبتها ويتجلى ذلك من تقارب قيم تقديرات القدرة المقابلة لكل درجة خام على كل صورة اختبارية منها. جاءت الصورة الثالثة أسهل الصور. على سبيل المثال الدرجة الخام ١٥ يناظرها على الصورة الأولى تقدير للقدرة (٥٠.٩) منف، في حين يقابلها على الصورة الاختبارية الثانية (٥١.٧)، ويناظرها على الصورة الثالثة (٤٨.٧٥) منف.

تدريج مفردات الصور الاختبارية الثلاث باستخدام طريقة التعادل المتزامن:

اتبع في تدريج مفردات الصور الاختبارية الثلاث الخطوات التالية:

- ضم بيانات الصور الاختبارية الثلاث في ملف واحد مع إعطاء المفردات المشتركة **نفس الكود.**
- تصحيح استجابات الطلاب على المفردات الاختبارية وفق مفتاح الإجابة.

- تحليل البيانات باستخدام برنامج WINSTEPS وفق نفس الخطوات المتبعة في تدريج الإطار المرجعي السابقة حيث أن تدريج مفردات الصور الاختبارية الثلاث يتم في حالة التدريج المتزامن مرة واحدة حيث إنه بانتهاء مراحل التحليل الثلاث من حيث حذف البيانات التامة والصفرية وحذف الأفراد والمفردات غير الملائمة _ وذلك وفق نفس محكات وقيم الملاءمة السابقة_ يتم الحصول على نتائج التدريج النهائي للمفردات الاختبارية وفق طريقة التدريج المتزامن.

- وفيما يلي عرض لهذه الخطوات كل على حدة.

- **حذف البيانات التامة والصفرية:** لم يسفر التحليل وفق هذه الخطوة عن حذف أي من المفردات أو الأفراد.

- **حذف الأفراد غير الملائمين لأسس القياس:** حذف (٢٢٦) فردا غير ملائم لأسس القياس الموضوعي.

- **حذف المفردات غير الملائمة لأسس القياس:** حذفت (٩) مفردات غير ملائمة لأسس القياس الموضوعي.

ويتضمن جدول (١٥) التالي ملخص نتائج تحليل الصور الاختبارية الثلاث وفق طريقة التدريج المتزامن

جدول(١٥)

ملخص نتائج تحليل الصور الاختبارية الثلاث وفق طريقة التدريج المتزامن.

التحليل	عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
			الصعوبة	القدرة	الصعوبة	القدرة
الأول	٧٥	١١٩٨	٠.٠٠٠	٠.٢	٠.٩٨	٠.٨٥
الثاني	٧٥	٩٧٢	٠.٠٠٠	٠.٣٣	٠.٩٨	٠.٨٤
الثالث	٦٦	٩٧٢	٠.٠٠٠	٠.٤١	٠.٩٨	٠.٨٣

يتضح من جدول (١٥) ما يلي:

-تمتعت تقديرات صعوبات المفردات بدرجة ثبات عالية حيث بلغت (٠.٩٨) في مراحل التحليل الثلاث مما يشير إلى دقة نتائج التدريج.

-كما تمتعت تقديرات قدرات الأفراد بدرجة ثبات عالية حيث بلغت (٠.٨٣) في التحليل النهائي.

التدريج النهائي لمفردات الصور الاختبارية الثلاث معا وفق طريقة التدريج المتزامن:

يتضمن جدول (١٦) التالي التدريج النهائي لمفردات الصور الاختبارية الثلاث مقدرة بوحدتي اللوجيت والمنف.

جدول (١٦):

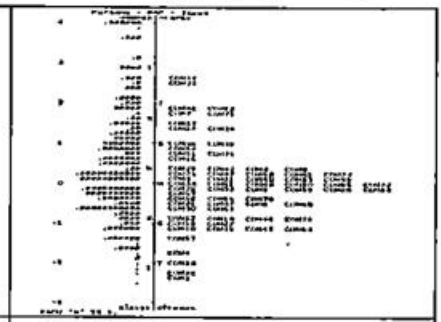
تقديرات الصعوبات النهائية لمفردات الصور الاختبارية الثلاث وفق طريقة التدريج المتزامن.

كود المفردة	رقم المفردة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف	رقم الصورة الاختبارية التي تقع بها المفردة
COM2	٦	٢.٣٧-	٣٨.١٥	٠.١٢	٠.٦٠	١-٢-٣
COM22	٧٣	٢.٢٤-	٣٨.٨٠	٠.٢٢	١.١٠	٣
COM28	٥٩	١.٩٥-	٤٠.٢٥	٠.٢٠	١.٠٠	٣
COM4	٨	١.٧٩-	٤١.٠٥	٠.١٠	٠.٥٠	١-٢-٣
COM17	٦٨	١.٣٢-	٤٣.٤٠	٠.١٦	٠.٨٠	٣
COM25	٥٦	١.١٦-	٤٤.٢٠	٠.١٦	٠.٨٠	٣
COM10	١٤	١.١٣-	٤٤.٣٥	٠.٠٨	٠.٤٠	١-٢-٣
COM63	٣٩	١.١٢-	٤٤.٤٠	٠.١٥	٠.٧٥	٢
COM45	٢١	١.١٠-	٤٤.٥٠	٠.١٤	٠.٧٠	١
COM27	٥٨	١.٠٢-	٤٤.٩٠	٠.١٥	٠.٧٥	٣
COM19	٧٠	١.٠٢-	٤٤.٩٠	٠.١٥	٠.٧٥	٣
COM74	٥٠	٠.٩٠-	٤٥.٥٠	٠.١٥	٠.٧٥	٢
COM12	٥٣	٠.٨٨-	٤٥.٦٠	٠.١٥	٠.٧٥	٣
COM44	٢٠	٠.٨٥-	٤٥.٧٥	٠.١٣	٠.٦٥	١
COM18	٦٩	٠.٨٤-	٤٥.٨٠	٠.١٥	٠.٧٥	٣
COM30	٦١	٠.٦٥-	٤٦.٧٥	٠.١٤	٠.٧٠	٣
COM67	٤٣	٠.٥٨-	٤٧.١٠	٠.١٤	٠.٧٠	٢
COM37	٤	٠.٥٤-	٤٧.٣٠	٠.١٣	٠.٦٥	١
COM60	٣٦	٠.٥٤-	٤٧.٣٠	٠.١٤	٠.٧٠	٢
COM6	١٠	٠.٥٠-	٤٧.٥٠	٠.٠٨	٠.٤٠	١-٢-٣
COM21	٧٢	٠.٤٤-	٤٧.٨٠	٠.١٤	٠.٧٠	٣
COM64	٤٠	٠.٣٦-	٤٨.٢٠	٠.١٤	٠.٧٠	٢
COM70	٤٦	٠.٣٤-	٤٨.٣٠	٠.١٤	٧٠.٠	٢
COM16	٦٧	٠.٣٣-	٤٨.٣٥	٠.١٤	٠.٧٠	٣
COM57	٣٣	٠.٣١-	٤٨.٤٥	٠.١٤	٠.٧٠	٢
COM68	٤٤	٠.١٧-	٤٩.١٥	٠.١٤	٠.٧٠	٢
COM5	٩	٠.١٦-	٤٩.٢٠	٠.٠٨	٠.٤٠	١-٢-٣
COM26	٥٧	٠.١٥-	٤٩.٢٥	٠.١٣	٠.٦٥	٣
COM59	٣٥	٠.١٤-	٤٩.٣٠	٠.١٤	٠.٧٠	٢
COM33	٦٤	٠.٠٩-	٤٩.٥٥	٠.١٣	٠.٦٥	٣
COM66	٤٢	٠.٠٨-	٤٩.٦٠	٠.١٤	٠.٧٠	٢
COM38	٥	٠.٠٦-	٤٩.٧٠	٠.١٢	٠.٦٠	١
COM50	٢٦	٠.٠١-	٤٩.٩٥	٠.١٢	٠.٦٠	١
COM72	٤٨	٠.٠١-	٤٩.٩٥	٠.١٤	٠.٧٠	٢
COM41	١٧	٠.٠٣	٥٠.١٥	٠.١٢	٠.٦٠	١
COM65	٤١	٠.٠٣	٥٠.١٥	٠.١٤	٠.٧٠	٢
COM48	٢٤	٠.٠٥	٥٠.٢٥	٠.١٢	٠.٦٠	١
COM3	٧	٠.١١	٥٠.٥٥	٠.٠٧	٠.٣٥	١-٢-٣
COM40	١٦	٠.١٢	٥٠.٦٠	٠.١٢	٠.٦٠	١
COM53	٢٩	٠.١٤	٥٠.٧٠	٠.١٢	٠.٦٠	١
COM55	٣١	٠.١٨	٥٠.٩٠	٠.١٤	٠.٧٠	2

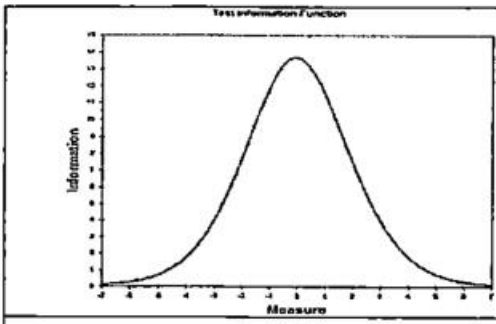
كود المفردة	رقم المفردة	الصعوبة باللوجيت	الصعوبة بالمنف	الخطأ المعياري باللوجيت	الخطأ المعياري بالمنف	رقم الصورة الاختبارية التي تقع بها المفردة
COM31	٦٢	٠.١٨	٥٠.٩٠	٠.١٣	٠.٦٥	٣
COM61	٣٧	٠.٢٠	٥١.٠٠	٠.١٤	٠.٧٠	٢
COM23	٧٤	٠.٢١	٥١.٠٥	٠.١٣	٠.٦٥	٣
COM47	٢٣	٠.٢٣	٥١.١٥	٠.١٢	٠.٦٠	١
COM58	٣٤	٠.٢٣	١٥.٥١	٠.١٤	٠.٧٠	٢
COM73	٤٩	٠.٣١	٥١.٥٥	٠.١٤	٠.٧٠	٢
COM8	١٢	٠.٣٣	٥١.٦٥	٠.٠٧	٠.٣٥	١ - ٢ - ٣
COM42	١٨	٠.٣٧	٥١.٨٥	٠.١٢	٠.٦٠	١
COM29	٦٠	٠.٣٧	٥١.٨٥	٠.١٣	٠.٦٥	٣
COM9	١٣	٠.٤٣	٥٢.١٥	٠.٠٨	٠.٤٠	١ - ٢ - ٣
COM11	٥٢	٠.٦٣	٥٣.١٥	٠.١٣	٠.٦٥	٣
COM75	٥١	٠.٧٢	٥٣.٦٠	٠.١٤	٠.٧٠	2
COM51	٢٧	٠.٨٠	٥٤.٠٠	٠.١٣	٠.٦٥	١
COM54	٣٠	٠.٨٤	٥٤.٢٠	٠.١٣	٠.٦٥	١
COM39	١٥	١.٠١	٥٥.٠٥	٠.١٣	٠.٦٥	١
COM36	٢	١.٠٦	٥٥.٣٠	٠.١٣	٠.٦٥	١
COM34	٦٥	١.٣٥	٥٦.٧٥	٠.١٣	٠.٦٥	٣
COM14	٥٥	١.٣٧	٥٦.٨٥	٠.١٣	٠.٦٥	٣
COM52	٢٨	١.٤٦	٥٧.٣٠	٠.١٤	٠.٧٠	١
COM7	١١	١.٧٣	٥٨.٦٥	٠.٠٨	٠.٤٠	١ - ٢ - ٣
COM71	٤٧	١.٧٥	٥٨.٧٥	٠.١٥	٠.٧٥	٢
COM46	٢٢	١.٨٥	٥٩.٢٥	٠.١٦	٠.٨٠	١
COM62	٣٨	١.٩١	٥٩.٥٥	٠.١٥	٠.٧٥	٢
COM32	٦٣	٢.٤٩	٦٢.٤٥	٠.١٦	٠.٨٠	٣
COM24	٧٥	٢.٦٨	٦٣.٤٠	٠.١٧	٠.٨٥	٣

يتضح من جدول (١٦) ما يلي:

- تراوحت التقديرات النهائية لصعوبات المفردات بين ٣٨.١٥: ٦٣.٤ منف، وتراوحت الأخطاء المعيارية لتلك التقديرات الصعوبة بين ٠.٣٥: ١.١ منف.
- يوضح شكل (٦) التالي خريطة الاختبار، ودالة معلومات الصور الاختبارية الثلاث وفق طريقة التعادل المتزامن.



خريطة الاختبار الكلي وفق طريقة التعادل المتزامن.



دالة معلومات الاختبار الكلي وفق طريقة التدريج المتزامن.

شكل (٦): خريطة الاختبار ودالة معلومات الصور الاختبارية الثلاث وفق طريقة التعادل المتزامن.

يتضح من شكل (٦) ما يلي:

- تغطي مفردات الاختبار الكلي (الصور الاختبارية الثلاث معا) متصل الصعوبة بشكل جيد، كما توجد أكثر من مفردة لها نفس مستوى الصعوبة.
- تغطي دالة المعلومات للاختبار الكلي (الصور الاختبارية الثلاث معا) متصل القياس بشكل مناسب خاصة عند المستوى المتوسط للقدرة؛ حيث بلغت أقصى قيمة لدالة المعلومات (١٣.٦٨) عند مستوى قدرة (-٠.٠٧) منف.

تقديرات القدرة المقابلة لكل درجة خام محتملة على الاختبار الكلي وفق طريقة التعادل المتزامن:

جدول (١٧): تقديرات القدرة المقابلة لكل درجة خام محتملة على الاختبار الكلي وفق طريقة التعادل المتزامن.

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
0	٥.٨٨-	١.٨٤	٢٠.٦٠	٩.٢٠
1	٤.٦٥-	١.٠٢	٢٦.٧٥	٥.١٠
2	٣.٩٢-	٠.٧٣	٣٠.٤٠	٣.٦٥
3	٣.٤٨-	٠.٦١	٣٢.٦٠	٣.٠٥
4	٣.١٦-	٠.٥٣	٣٤.٢٠	٢.٦٥
5	٢.٩٠-	٠.٤٨	٣٥.٥٠	٢.٤٠
6	٢.٦٨-	٠.٤٥	٣٦.٦٠	٢.٢٥
7	٢.٥٠-	٠.٤٢	٣٧.٥٠	٢.١٠
8	٢.٣٣-	٠.٤٠	٣٨.٣٥	٢.٠٠
9	٢.١٨-	٠.٣٨	٣٩.١٠	١.٩٠
10	٠٤.٢-	٠.٣٦	٣٩.٨٠	١.٨٠
11	١.٩١-	٠.٣٥	٤٠.٤٥	١.٧٥

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
12	١.٧٩-	٠.٣٤	٤١.٠٥	١.٧٠
13	١.٦٨-	٠.٣٣	٤١.٦٠	١.٦٥
14	١.٥٧-	٠.٣٢	٤٢.١٥	١.٦٠
15	١.٤٧-	٠.٣٢	٤٢.٦٥	١.٦٠
16	١.٣٧-	٠.٣١	٤٣.١٥	١.٥٥
17	١.٢٨-	٠.٣٠	٤٣.٦٠	١.٥٠
18	١.١٨-	٠.٣٠	٤٤.١٠	١.٥٠
19	١.١٠-	٠.٢٩	٤٤.٥٠	١.٤٥
20	١.٠١-	٠.٢٩	٤٤.٩٥	١.٤٥
21	٠.٩٣-	٠.٢٩	٤٥.٣٥	١.٤٥
22	٠.٨٥-	٠.٢٨	٤٥.٧٥	١.٤٠
23	٠.٧٧-	٠.٢٨	٤٦.١٥	١.٤٠
24	٠.٦٩-	٠.٢٨	٤٦.٥٥	١.٤٠
25	٠.٦١-	٠.٢٨	٤٦.٩٥	١.٤٠
26	٠.٥٣-	٠.٢٨	٤٧.٣٥	١.٤٠
27	٠.٤٦-	٠.٢٧	٤٧.٧٠	١.٣٥
28	٠.٣٨-	٠.٢٧	٤٨.١٠	١.٣٥
29	٠.٣١-	٠.٢٧	٤٨.٤٥	١.٣٥
30	٠.٢٤-	٠.٢٧	٤٨.٨٠	١.٣٥
31	٠.١٦-	٠.٢٧	٤٩.٢٠	١.٣٥
32	٠.٠٩-	٠.٢٧	٤٩.٥٥	١.٣٥
33	٠.٠٢-	٠.٢٧	٤٩.٩٠	١.٣٥
34	٠.٠٦	٠.٢٧	٥٠.٣٠	١.٣٥
35	٠.١٣	٠.٢٧	٥٠.٦٥	١.٣٥
36	٠.٢٠	٠.٢٧	٥١.٠٠	١.٣٥
37	٠.٢٨	٠.٢٧	٥١.٤٠	١.٣٥
38	٠.٣٥	٠.٢٧	٥١.٧٥	١.٣٥
39	٠.٤٣	٠.٢٨	٥٢.١٥	١.٤٠
40	٠.٥٠	٠.٢٨	٥٢.٥٠	١.٤٠
41	٠.٥٨	٠.٢٨	٥٢.٩٠	١.٤٠
42	٠.٦٦	٠.٢٨	٥٣.٣٠	١.٤٠
43	٠.٧٤	٠.٢٨	٥٣.٧٠	١.٤٠
44	٠.٨٢	٠.٢٩	٥٤.١٠	١.٤٥
45	٠.٩١	٠.٢٩	٥٤.٥٥	١.٤٥
46	٠.٩٩	٠.٢٩	٥٤.٩٥	١.٤٥
47	١.٠٨	٠.٣٠	٥٥.٤٠	١.٥٠
48	١.١٧	٠.٣٠	٥٥.٨٥	١.٥٠
49	١.٢٦	٠.٣١	٥٦.٣٠	١.٥٥
50	١.٣٦	٠.٣١	٥٦.٨٠	١.٥٥
51	١.٤٦	٠.٣٢	٥٧.٣٠	١.٦٠
52	١.٥٧	٠.٣٣	٥٧.٨٥	١.٦٥
53	١.٦٨	٠.٣٤	٥٨.٤٠	١.٧٠
54	١.٨٠	٠.٣٥	٥٩.٠٠	١.٧٥
55	١.٩٢	٠.٣٦	٥٩.٦٠	١.٨٠
56	٢.٠٥	٠.٣٧	٦٠.٢٥	١.٨٥

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
57	٢,٢٠	٠,٣٩	٦١,٠٠	١,٩٥
58	٢,٣٥	٠,٤٠	٦١,٧٥	٢,٠٠
59	٢,٥٢	٠,٤٣	٦٢,٦٠	٢,١٥
60	٢,٧٢	٠,٤٥	٦٣,٦٠	٢,٢٥
61	٢,٩٤	٠,٤٩	٦٤,٧٠	٢,٤٥
62	٣,٢٠	٠,٥٤	٦٦,٠٠	٢,٧٠
63	٣,٥٣	٠,٦١	٦٧,٦٥	٣,٠٥
64	٣,٩٨	٠,٧٤	٦٩,٩٠	٣,٧٠
65	٤,٧٢	١,٠٢	٧٣,٦٠	٥,١٠
66	٥,٩٦	١,٨٤	٧٩,٨٠	٩,٢٠

يتضح من جدول (١٧) ما يلي:

- تغطي تقديرات القدرة المقابلة لكل درجة خام محتملة على الاختبار الكلي (الصور الثلاث بعد ضمها في ملف بيانات واحد) مدى واسعا من تقديرات القدرة. امتد بين (٢٠.٦) إلى (٧٩.٨) منف. ويعتبر هذا المدى لتقديرات القدرة أكبر مدى للاختبار ككل أو أي صورة فرعية متعادلة القياس مسحوبة منه.

- الحصول على تقديرات القدرة المقابلة لكل درجة خام محتملة على كل صورة اختبارية من الصور الثلاث على حدة.

للحصول على تقديرات القدرة المقابلة لكل درجة خام محتملة على كل صورة من الصور الثلاث **كل على حدة** تم استخدام تقديرات صعوبات كافة المفردات المتضمنة بكل صورة اختبارية والتي تم تقديرها في الاختبار الكلي وفق طريقة التعادل المتزامن كما سبق، ويتضمن الجزء التالي هذه النتائج.

- الحصول على تقديرات القدرة المقابلة لكل درجة خام على الصورة الأولى من الاختبار الكلي.

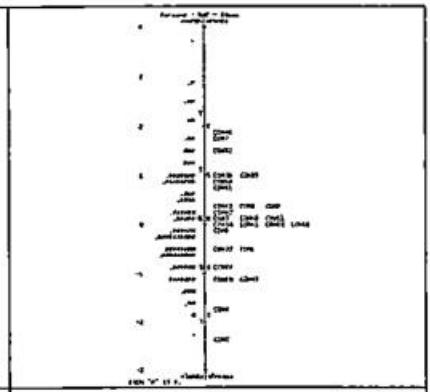
باستخدام تقديرات كافة مفردات الصورة الأولى الملائمة من الاختبار الكلي والتي تم الحصول عليها في التحليل النهائي وفق طريقة التدريج المتزامن بالخطوة السابقة تم إعادة التحليل للحصول على تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الأولى. ويتضمن جدول (١٨) التالي ملخص نتائج إعادة التحليل.

جدول(١٨): ملخص نتائج إعادة التحليل للصورة الاختبارية الأولى.

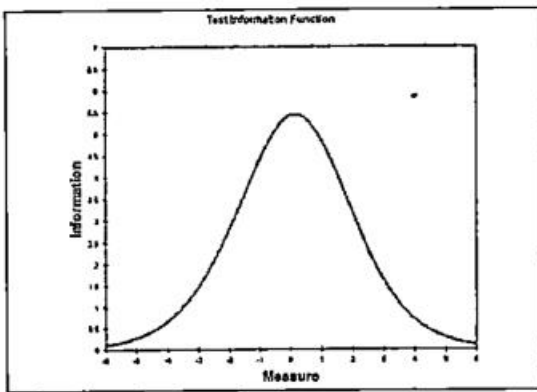
عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
		الصعوبة	القدرة	الصعوبة	القدرة
٢٦	٣٢٢	٠,٠٨	٠,١٢	٠,٩٨	٠,٧٨

ويتضمن شكل (٧) التالي دالة المعلومات وخريطة الاختبار للصورة الاختبارية الأولى وفق.

طريقة التعادل المتزامن.



خريطة مفردات الصورة الأولى وفق طريقة التعادل المتزامن.



دالة معلومات الصورة الأولى وفق طريقة التعادل المتزامن.

شكل (٧): دالة المعلومات وخريطة الاختبار للصورة الاختبارية الأولى وفق طريقة التعادل المتزامن.

يتضح من شكل (٧) أن دالة المعلومات تغطي مدى مناسبا من القدرة؛ خاصة عند المستوى المتوسط؛ حيث بلغت أقصى قيمة لدالة المعلومات (٥.٤٤) عند مستوى القدرة (٠.٠٦).

- تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الأولى وفق طريقة التدريج المتزامن:

يتضمن جدول(١٩) التالي تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الأولى وفق طريقة التعادل المتزامن.

جدول(١٩): تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الأولى وفق طريقة التعادل المتزامن.

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
0	٤.٨٩-	١.٨٦	٢٥.٥٥	٩.٣٠
1	٣.٦١-	١.٠٥	٣١.٩٥	٥.٢٥
2	٢.٨٢-	٠.٧٨	٣٥.٩٠	٣.٩٠
3	٢.٣١-	٠.٦٦	٣٨.٤٥	٣.٣٠
4	١.٩٣-	٠.٥٩	٤٠.٣٥	٢.٩٥
5	١.٦١-	٠.٥٤	٤١.٩٥	٢.٧٠
6	١.٣٤-	٠.٥١	٤٣.٣٠	٢.٥٥
7	١.١٠-	٠.٤٨	٤٤.٥٠	٢.٤٠
8	٠.٨٧-	٠.٤٧	٤٥.٦٥	٢.٣٥
9	٠.٦٦-	٠.٤٥	٤٦.٧٠	٢.٢٥

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
10	-٠,٤٦	٠,٤٤	٤٧,٧٠	٢,٢٠
11	-٢٧,٠	٠,٤٣	٤٨,٦٥	٢,١٥
12	-٠,٠٨	٠,٤٣	٤٨,٦٠	٢,١٥
13	٠,١٠	٠,٤٣	٥٠,٥٠	٢,١٥
14	٠,٢٨	٠,٤٣	٥١,٤٠	٢,١٥
15	٠,٤٧	٠,٤٣	٥٢,٣٥	٢,١٥
16	٠,٦٦	٠,٤٤	٥٣,٣٠	٢,٢٠
17	٠,٨٥	٠,٤٥	٥٤,٢٥	٢,٢٥
18	١,٠٦	٠,٤٦	٥٥,٣٠	٢,٣٠
19	١,٢٨	٠,٤٧	٥٦,٤٠	٢,٣٥
20	١,٥١	٠,٥٠	٥٧,٥٥	٢,٥٠
21	١,٧٧	٥٣,٠	٥٨,٨٥	٢,٦٥
22	٢,٠٧	٠,٥٧	٦٠,٣٥	٢,٨٥
23	٢,٤٤	٠,٦٤	٦٢,٢٠	٣,٢٠
24	٢,٩٢	٠,٧٦	٦٤,٦٠	٣,٨٠
25	٣,٦٩	١,٠٤	٦٨,٤٥	٥,٢٠
26	٤,٩٤	١,٨٥	٧٤,٧٠	٩,٢٥

يتضح من جدول (١٩) أن مدى تقدير القدرة على الصورة الاختبارية الأولى وفق طريقة التعادل المتزامن تراوح بين (25.55) إلى (٧٤.٧) منف.

- الحصول على تقديرات القدرة المقابلة لكل درجة خام على الصورة الاختبارية الثانية من الاختبار الكلى.

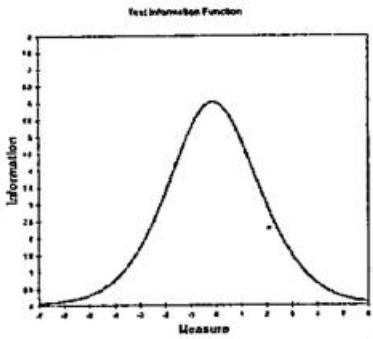
باستخدام تقديرات كافة مفردات الصورة الاختبارية الثانية الملائمة من الاختبار الكلي والتي تم الحصول عليها في التحليل النهائي وفق طريقة التعادل المتزامن بالخطوة السابقة تم إعادة التحليل للحصول على تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثانية. ويتضمن جدول (٢٠) التالي ملخص نتائج إعادة التحليل.

جدول(٢٠)

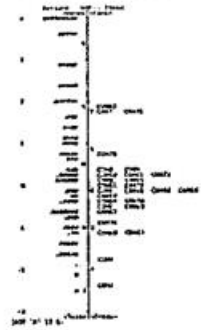
ملخص نتائج إعادة التحليل للصورة الاختبارية الثانية.

عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
		الصعوبة	القدرة	الصعوبة	القدرة
٢٨	٣٣٦	-٠,٠٩	٠,٤٦	٠,٩٨	٠,٨٦

ويتضمن شكل (٨) التالي دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثانية وفق طريقة التدريج المتزامن.



دالة معلومات الصورة الثانية وفق طريقة التعادل المتزامن.



خريطة مفردات الصورة الثانية وفق طريقة التعادل المتزامن.

شكل (٨): دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثانية وفق طريقة التدريج المتزامن.

يتضح من شكل (٨) أن دالة المعلومات تغطي مدى مناسبا من القدرة؛ خاصة عند المستوى المتوسط للقدرة؛ حيث بلغت قيمة دالة المعلومات (٦.٠٢) عند مستوى القدرة (-٠.١١).

- **تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثانية وفق طريقة التدريج المتزامن:** يتضمن جدول(٢١) التالي تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثانية وفق طريقة التدريج المتزامن.

جدول(٢١)

تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثانية وفق طريقة التعادل المتزامن.

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
0	-٤.٥٠	١.٨٥	٢٤.٨٠	٩.٢٥
1	-٣.٧٨	١.٠٤	٣١.١٠	٥.٢٠
2	-٣.٠٠	٠.٧٦	٣٥.٠٠	٣.٨٠
3	-٢.٥٢	٠.٦٤	٣٧.٤٠	٣.٢٠
4	-٢.١٥	٠.٥٧	٣٩.٢٥	٢.٨٥
5	-١.٨٥	٠.٥٢	٤٠.٧٥	٢.٦٠
6	-١.٦٠	٠.٤٩	٤٢.٠٠	٢.٤٥
7	-١.٣٧	٠.٤٧	٤٣.١٥	٢.٣٥
8	-١.١٦	٠.٤٥	٤٤.٢٠	٢.٢٥
9	-٠.٩٦	٠.٤٣	٤٥.٢٠	٢.١٥
10	-٠.٧٨	٤٢.٠	٤٦.١٠	٢.١٠
11	-٠.٦٠	٠.٤٢	٤٧.٠٠	٢.١٠
12	-٠.٤٣	٠.٤١	٤٧.٨٥	٢.٠٥

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
13	-٠.٢٦	٠.٤١	٤٨.٧٠	٢.٠٥
14	-٠.١٠	٠.٤١	٤٩.٥٠	٢.٠٥
15	٠.٠٧	٠.٤١	٥٠.٣٥	٢.٠٥
16	٠.٢٤	٠.٤١	٥١.٢٠	٢.٠٥
17	٠.٤١	٠.٤٢	٥٢.٠٥	٢.١٠
18	٠.٥٩	٠.٤٣	٥٢.٩٥	٢.١٥
19	٠.٧٧	٠.٤٤	٥٣.٨٥	٢.٢٠
20	٠.٩٧	٠.٤٥	٥٤.٨٥	٢.٢٥
21	١.١٨	٠.٤٧	٥٥.٩٠	٢.٣٥
22	١.٤١	٠.٤٩	٥٧.٠٥	٢.٤٥
23	١.٦٧	٠.٥٣	٥٨.٣٥	٢.٦٥
24	١.٩٨	٠.٥٧	٥٩.٩٠	٢.٨٥
25	٢.٣٤	٠.٦٤	٦١.٧٠	٣.٢٠
26	٢.٨٣	٠.٧٦	٦٤.١٥	٣.٨٠
27	٣.٦١	١.٠٤	٦٨.٠٥	٥.٢٠
28	٤.٨٧	١.٨٥	٧٤.٣٥	٩.٢٥

يتضح من جدول (٢١) السابق أن مدى تقدير القدرة على الصورة الاختبارية الثانية وفق طريقة التعادل المتزامن تراوح بين (24.8) إلى (٧٤.٣٥) منف.

- الحصول على تقديرات القدرة المقابلة لكل درجة خام على الصورة الاختبارية الثالثة من الاختبار الكلى.

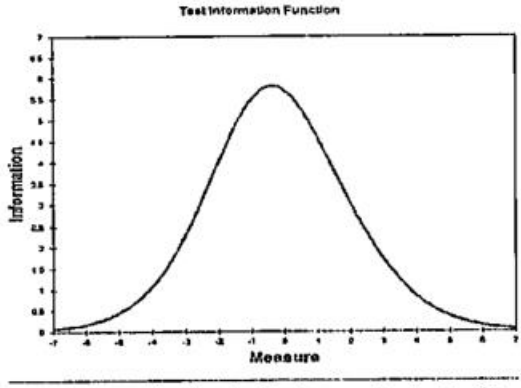
باستخدام تقديرات صعوبات كافة مفردات الصورة الثالثة الملائمة من الاختبار الكلي والتي تم الحصول عليها في التحليل النهائي وفق طريقة التعادل المتزامن بالخطوة السابقة تم إعادة التحليل للحصول على تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثالثة. ويتضمن جدول (٢٢) التالي ملخص نتائج إعادة التحليل.

جدول(٢٢)

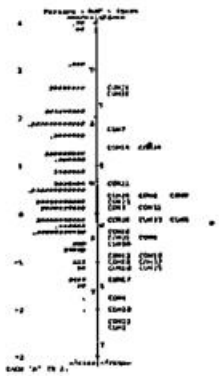
ملخص نتائج إعادة التحليل للصورة الاختبارية الثالثة.

عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
		الصعوبة	القدرة	الصعوبة	القدرة
٣٠	٣١٤	-٠.٢١	٠.٦٨	٠.٩٨	٠.٨١

ويتضمن شكل (٨) التالي دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثالثة وفق طريقة التعادل المتزامن.



دالة معلومات الصورة الثالثة وفق طريقة التعادل المتزامن.



خريطة مفردات الصورة الثالثة وفق طريقة التعادل المتزامن.

شكل (٨): دالة المعلومات وخريطة الاختبار للصورة الاختبارية الثالثة وفق طريقة التعادل المتزامن.

يتضح من شكل (٨) أن دالة المعلومات تغطي مدى مناسبا من القدرة؛ خاصة عند المستوى المتوسط للقدرة؛ حيث بلغت أقصى قيمة لدالة المعلومات (5.8) عند مستوى القدرة (-0.42).

- **تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثالثة وفق طريقة التعادل المتزامن:** يتضمن جدول (٢٢) التالي تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثالثة وفق طريقة التدريج المتزامن.

جدول (٢٢): تقديرات القدرة المقابلة لكل درجة خام محتملة على الصورة الاختبارية الثالثة وفق طريقة التعادل المتزامن.

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
0	-٥.٤٢	١.٨٥	٢٢.٩٠	٩.٢٥
1	-٤.١٦	١.٠٤	٢٩.٢٠	٥.٢٠
2	-٣.٣٩	٠.٧٦	٣٣.٠٥	٣.٨٠
3	-٢.٩١	٠.٦٤	٣٥.٤٥	٣.٢٠
4	-٢.٥٤	٠.٥٧	٣٧.٣٠	٢.٨٥
5	-٢.٢٤	٠.٥٣	٣٨.٨٠	٢.٦٥
6	-١.٩٨	٠.٥٠	٤٠.١٠	٢.٥٠
7	-١.٧٤	٠.٤٧	٤١.٣٠	٢.٣٥
8	-١.٥٣	٠.٤٦	٤٢.٣٥	٢.٣٠
9	-١.٣٣	٠.٤٤	٤٣.٣٥	٢.٢٠
10	-١.١٣	٠.٤٣	٤٤.٣٥	٢.١٥
11	-٠.٩٥	٠.٤٣	٤٥.٢٥	٢.١٥

الدرجة الخام	تقدير القدرة باللوجيت	الخطأ المعياري باللوجيت	تقدير القدرة بالملف	الخطأ المعياري بالملف
12	-٠.٧٧	٠.٤٢	٤٦.١٥	٢.١٠
13	-٠.٦٠	٠.٤٢	٤٧.٠٠	٢.١٠
14	-٠.٤٢	٠.٤٢	٤٧.٩٠	٢.١٠
15	-٠.٢٥	٠.٤٢	٤٨.٧٥	٢.١٠
16	-٠.٠٨	٠.٤٢	٤٩.٦٠	٢.١٠
17	٠.١٠	٠.٤٢	٥٠.٥٠	٢.١٠
18	٠.٢٨	٠.٤٣	٥١.٤٠	٢.١٥
19	٠.٤٦	٠.٤٣	٥٢.٣٠	٢.١٥
20	٠.٦٥	٠.٤٤	٥٣.٢٥	٢.٢٠
21	٠.٨٦	٠.٤٦	٥٤.٣٠	٢.٣٠
22	١.٠٧	٠.٤٧	٥٥.٣٥	٢.٣٥
23	١.٣١	٠.٤٩	٥٦.٥٥	٢.٤٥
24	١.٥٦	٠.٥٢	٥٧.٨٠	٢.٦٠
25	١.٨٥	٠.٥٥	٥٩.٢٥	٢.٧٥
26	٢.١٨	٠.٦٠	٦٠.٩٠	٣.٠٠
27	٢.٥٩	٠.٦٧	٦٢.٩٥	٣.٣٥
28	٣.١١	٠.٧٩	٦٥.٥٥	٣.٩٥
29	٣.٩٣	١.٠٦	٦٩.٦٥	٥.٣٠
30	٥.٢٢	١.٨٦	٧٦.١٠	٩.٣٠

يتضح من جدول (٢٢) السابق أن مدى تقدير القدرة على الصورة الاختبارية الثالثة وفق طريقة التعادل المتزامن تراوح بين (٢٢.٩) إلى (٧٦.١) منف.

بالنسبة لتعادل القياس على الصور الاختبارية الثلاث وفق طريقة التعادل المتزامن:

- بانتهاء مراحل تدريج ملف البيانات الكلي الذي يضم استجابات كافة الأفراد على مفردات الصور الاختبارية الثلاث أصبحت كافة المفردات الملائمة لأسس القياس الموضوعي على ميزان تدرج واحد بصفر مشترك واحد، وأصبحت الصور الثلاث متعادلة القياس بحيث لا يختلف تقدير قدرة الفرد باختلاف الصورة الاختبارية التي يتعرض لها. كما أن هذه الصور تكون متعادلة مع أي صورة فرعية أخرى يمكن سحبها من البنك الكلي حيث يمكن سحب صور اختبارية أخرى متعادلة القياس من البنك الكلي مكونة من مفردات اختبارية من كل من الصور الثلاث ومع ذلك تكون الصور الجديدة المسحوبة من الاختبار الكلي متعادلة القياس مع الصورة الاختبارية الأصلية ومع أي صورة اختبارية تسحب من التدريج الكلي.

- وفيما يلي مثال لتعادل التقديرات على الصور الاختبارية الثلاثة بعد تدريجها على ميزان تدرج واحد بصفر مشترك واحد.

جدول (٢٣): نماذج لتعادل القياس على الصور الاختبارية الثلاث وفق طريقة التدريج المتزامن.

الدرجة الخام	الصورة الاختبارية الأولى		الصورة الاختبارية الثانية		الصورة الاختبارية الثالثة	
	تقدير القدرة بالمنف	الخطأ المعياري لتقدير القدرة بالمنف	تقدير القدرة بالمنف	الخطأ المعياري بالمنف	تقدير القدرة بالمنف	الخطأ المعياري بالمنف
3	٣٨.٤٥	٣.٣٠	٣٧.٤٠	٣.٢٠	٣٥.٤٥	٣.٢٠
4	٤٠.٣٥	٢.٩٥	٣٩.٢٥	٢.٨٥	٣٧.٣٠	٢.٨٥
5	٤١.٩٥	٢.٧٠	٤٠.٧٥	٢.٦٠	٣٨.٨٠	٢.٦٥
10	٤٧.٧٠	٢.٢٠	٤٦.١٠	٢.١٠	٤٤.٣٥	٢.١٥
15	٥٢.٣٥	٢.١٥	٥٠.٣٥	٢.٠٥	٤٨.٧٥	٢.١٠
20	٥٧.٥٥	٢.٥٠	٥٤.٨٥	٢.٢٥	٥٣.٢٥	٢.٢٠

يتضح من جدول (٢٣) ما يلي:

- الصور الاختبارية الثلاث متقاربة في مستوى صعوبتها ويتجلى ذلك فى تقارب قيم تقديرات القدرة المقابلة لكل درجة خام على كل صورة اختبارية منها. وقد جاءت الصورة الثالثة أسهل الصور، وتقاربت بدرجة كبيرة صعوبة الصورتين الأولى والثانية. على سبيل المثال الدرجة الخام ١٥ يناظرها على الصورة الأولى تقدير للقدرة (٥٢.٣٥) منف، في حين يقابلها على الصورة الاختبارية الثانية (٥٠.٣٥)، ويناظرها على الصورة الثالثة (٤٨.٧٥) منف.

حساب صدق وثبات القياس:

أولا: حساب الصدق: للتحقق من صدق القياس استخدمت الباحثة كل من: صدق المحتوى، وصدق التدرج كما يوفره " نموذج راش".

(١) **صدق المحتوى:** يشير صدق المحتوى إلى درجة التوافق بين مفردات الاختبار ومحتوى جدول المواصفات وللتأكد من صدق المحتوى لمستودع الأسئلة اعتمدت الباحثة على صدق البناء، وصدق المحكمين. ويرتبط **صدق المحتوى** بخطة بناء الاختبار؛ حيث تم بناء الاختبار وفق خطة محددة تشمل تحديد المجال السلوكي الذي يغطيه الاختبار وما يتضمنه من تحليل للمحتوى وتحديد أوجه التعلم المتضمنة، والأهداف السلوكية وما يترتب على ذلك من إعداد جدول المواصفات والذي في ضوئه يتم صياغة المفردات.

وقد سارت الباحثة بهذه الخطوات في بنائها للصورة الأولية لمستودع الأسئلة. وبهذا يكون قد تحقق صدق المحتوى وهو **صدق البناء** كما يتمثل في إعداد المفردات بحيث تقيس متغيراً واحداً هو التحصيل الدراسي في مادة العلوم، وأن تكون الإجابة على كل مفردة من مفردات الاختبار مستقلة عن الإجابة عن أي من مفرداته الأخرى.

أما **بالنسبة لصدق المحكمين** فقد قامت الباحثة بعرض مجموعة المفردات الاختبارية على مجموعة من المحكمين وذلك لمراجعتها، والحكم على مدى تحقيقها للأهداف التي وضعت لقياسها ومدى ملاءمتها للمستوى المعرفي الذي وضعت لقياسه. وفى ضوء آراء المحكمين تم تعديل المفردات التي تحتاج إلى تعديل. وبتحقق التطابق بين المفردات الاختبارية والأهداف التي تقيسها يتحقق صدق البناء للمفردات وبالتالي يتحقق صدق المحتوى.

(٢) صدق التدرج: ويعتمد على أن تدرج مفردات مجموعة من الصور الاختبارية تقيس نفس السمة على ميزان واحد مشترك باستخدام نموذج "راش" يعنى أن هذه المفردات تحقق شرط أحادية القياس أي أنها تعرف متغيراً واحداً.

وتتحقق أحادية البعد بتحقق صدق المحتوى، وبملاءمة كل من الأفراد والمفردات لأسس القياس الموضوعي تبعاً لمحكات الملاءمة الخاصة ببرنامج WINSTEPS المستخدم في التحليل الإحصائي. حيث توضح هذه المحكات مدى قياس المفردة لما تقيسه باقي المفردات على متصل المتغير موضوع القياس، كما تبين مدى اتساق تدرج قدرة الفرد مع تدرج قدرات باقي الأفراد على هذا المتصل. وتم حذف المفردات غير الملائمة وبلغ عددها سبع مفردات عند استخدام طريقة التعادل المتزامن، وتسع مفردات في حالة التعادل بتثبيت بارامترات المفردات المشتركة وفقا لتلك المحكات الإحصائية. وذلك لأسباب مختلفة كأن تعتمد المفردة على غيرها من المفردات، أو أن تكون المفردة ضعيفة الصياغة أو غير ذلك. كما تم حذف الأفراد غير الملائمين وبلغ عددهم (٢٢٦) فردا عند استخدام طريقة التعادل المتزامن، و(٢١٠) فردا في حالة التعادل بتثبيت بارامترات المفردات المشتركة وفقا لتلك المحكات الإحصائية، وذلك لأسباب مختلفة أيضاً كأن يلجأ الطالب إلى التخمين أو الغش أو يتسم أداؤه بالبطء الشديد أو السرعة أو يستجيب بعكس ما هو متوقع منه.

وبحذف الحالات غير الملائمة من الأفراد والمفردات يستبقى الأفراد الصادقون وتكون المفردات المتبقية هي المفردات الصادقة في تعريف متغير **التحصيل الدراسي في مادة العلوم**؛ مما يوفر الموضوعية في تقدير الأفراد على متصل التحصيل الدراسي، كما يوفرها نموذج راش من خلال شروطه. وبالإضافة إلى المفردات المستبقاة يستبقى أيضاً الأفراد الصادقون في استجاباتهم على تلك المفردات ويعنى هذا توفر شرطي الصدق والثبات لتقديرات كل من صعوبات مفردات الاختبار، وقدرات الأفراد أي يعنى تحقق صدق وثبات القياس.

ثانياً ثبات القياس: اعتمد حساب ثبات القياس على دراسة كل من: ثبات درجات الصور الاختبارية باستخدام معامل ألفا كرونباك، وثبات التدرج وحساب الخطأ المعياري لتقديرات كل من

صعوبات المفردات وقدرات الأفراد.

ثبات درجات الصور الاختبارية:

حسب معامل الثبات ألفا كرونباخ لكل صورة اختبارية على حدة وبلغت قيمته ٠.٧٩٢ بالنسبة للصورة الاختبارية الأولى، وبلغت قيمته ٠.٩٢٤ للصورة الاختبارية الثانية، كما بلغت قيمته ٠.٨٣٩ للصورة الاختبارية الثالثة. وتعتبر هذه القيم مرتفعة في إطار القيم المقبولة لمعامل الثبات مما يشير إلى أن الصور الاختبارية الثلاث تتمتع بدرجة ثبات مناسبة.

▪ **ثبات التدرج:** إن تدرج مفردات اختبار التحصيل الدراسي الكلية معاً على ميزان تدرج واحد مشترك وفقا لنموذج "راش" – بعد حذف الحالات غير الملائمة من الأفراد والمفردات – يعني تحقق شروط النموذج ومنها شرط استقلال القياس، ويعنى ذلك ثبات تقديرات كل من الصعوبة والقدرة وعدم تأثرهما باختلاف الصورة الاختبارية المستخدمة والمسحوبة من الاختبار الكلي أو باختلاف عينة الأفراد المستخدمة في تدريجه (أمينة محمد كاظم،٢٠٠٠، ٢٣١).

وقد حسب معامل ثبات تقدير كل من صعوبة المفردات وقدرة الأفراد باستخدام معامل ثبات **كيودريتشاردسون** الذي يوفره برنامج WINSTEPS. وتشير قيم معاملات الثبات الموضحة بجدول ( ) إلى أن الصور الاختبارية الثلاث تتمتع بتقديرات تتميز بتقديرات ثبات مرتفعة لتقديرات صعوبات المفردات، كما توفر تقديرات ثبات مرتفعة لتقديرات قدرة الأفراد حيث بلغ معامل ثبات تقدير القدرة ( ) تقريباً.

▪ **حساب الخطأ المعياري لتقديرات كل من صعوبات المفردات وقدرات الأفراد:** يوفر برنامج الحاسب الآلي WINSTEPS تقديرات للخطأ المعياري لتقدير قدرة كل فرد، وكذا صعوبة كل مفردة، والتي تعد مؤشرا لدقة القياس. ويتضح من جدول () أن قيم الخطأ المعياري لصعوبات المفردات تعتبر منخفضة بوجه عام، حيث تتراوح هذه القيم بين (٠.١٣: ٠.١٥) لوجيت. كما يوضح جدول(٢٤) انخفاض قيم الخطأ المعياري لتقديرات قدرات الأفراد على الصور الاختبارية الثلاث حيث تراوحت بين(٠.٤٥: ٠.٥٥) لوجيت.

جدول(٢٤)

معامل الثبات والأخطاء المعيارية لتقديرات صعوبات المفردات وقدرات الأفراد وفق طريقتي التعادل.

الإحصاءة	الصورة الأولى		الصورة الثانية		الصورة الثالثة	
	تثبيت بارامترات المفردات	التدريج المتزامن	تثبيت بارامترات المفردات	التدريج المتزامن	تثبيت بارامترات المفردات	التدريج المتزامن
معامل الثبات لتقديرات الصعوبة	٠.٩٨	٠.٩٨	٠.٩٧	٠.٩٨	٠.٩٨	٠.٩٨
متوسط الخطأ المعياري لتقدير صعوبة المفردات.	٠.١٣	٠.١٤	٠.١٥	٠.١٥	٠.١٥	٠.١٥
معامل الثبات لتقديرات القدرة.	٠.٧٨	٠.٧٨	٠.٨٥	٠.٨٦	٠.٨٠	٠.٨١
متوسط الخطأ المعياري لتقدير القدرة.	٠.٤٥	٠.٤٩	٠.٥٥	٠.٥٢	٠.٥٠	٠.٥١

- الاتساق الداخلي:

تم حساب معامل ارتباط كل مفردة بالدرجة الكلية للصورة الاختبارية التي تندرج تحتها. ويتضمن جدول (٢٥) التالي قيم معاملات الارتباط بين درجة كل مفردة والدرجة الكلية للصورة الاختبارية.

جدول (٢٥)

معامل ارتباط كل مفردة بالدرجة الكلية للصورة الاختبارية.

الصورة الاختبارية الأولى		الصورة الاختبارية الثانية		الصورة الاختبارية الثالثة	
كود المفرد	معامل ارتباط المفردة بالدرجة الكلية المصحح	كود المفرد	معامل ارتباط المفردة بالدرجة الكلية المصحح	كود المفرد	معامل ارتباط المفردة بالدرجة الكلية المصحح
COM35	٠.٠٦	COM55	٠.٥١	COM11	٠.٣٤
COM36	٠.١٧	COM56	٠.٥٥	COM12	٠.٤٤
COM1	٠.٣٠	COM1	٠.٣١	COM1	٠.٣٠
COM37	٠.٣٤	COM57	٠.٦٠	COM13	-٠.٠٣
COM38	٠.٣٤	COM58	٠.٦٠	COM14	٠.٢٠
COM2	٠.٣٠	COM2	٠.٤٠	COM2	٠.٣٠
COM3	٠.٣٣	COM3	٠.٤٥	COM3	٠.٥١
COM4	٠.٤٠	COM4	٠.٤٥	COM4	٠.٣١
COM5	٠.٢٣	COM5	٠.٤٢	COM5	٠.٢٣
COM6	٠.٣٨	COM6	٠.٤٣	COM6	٠.٤١
COM7	٠.١٤	COM7	٠.٣٢	COM7	٠.٣٠

الصورة الاختبارية الأولى		الصورة الاختبارية الثانية		الصورة الاختبارية الثالثة	
كود المفرد	معامل ارتباط المفردة بالدرجة الكلية المصحح	كود المفرد	معامل ارتباط المفردة بالدرجة الكلية المصحح	كود المفرد	معامل ارتباط المفردة بالدرجة الكلية المصحح
COM8	٠.٤٠	COM8	٠.٦٠	COM8	٠.٤١
COM9	٠.٣٥	COM9	٠.٥٢	COM9	٠.٣١
COM10	٠.٥٠	COM10	٠.٥٣	COM10	٠.٣٠
COM39	٠.٢٤	COM59	٠.٤٦	COM25	٠.٤٢
COM40	٠.٣٦	COM60	٠.٥٩	COM26	٠.٣٩
COM41	٠.٣٦	COM61	٠.٦٠	COM27	٠.٤٠
COM42	٠.٣٤	COM62	٠.٥٥	COM28	٠.٣٣
COM43	٠.١٤	COM63	٠.٤٦	COM29	٠.٥٣
COM44	٠.٣٩	COM64	٠.٥٠	COM30	٠.٤٠
COM45	٠.٤٢	COM65	٠.٤٩	COM31	٠.٣٥
COM46	٠.١٢	COM66	٠.٥٦	COM32	٠.٣٠
COM47	٠.٤٤	COM67	٠.٥٩	COM33	٠.٤١
COM48	٠.٢٥	COM68	٠.٥٥	COM34	٠.٣٤
COM49	٠.١٦	COM69	٠.٣٣	COM15	٠.١٣
COM50	٠.٢٧	COM70	٠.٦٠	COM16	٠.٣٢
COM51	٠.٢٢	COM71	٠.٥١	COM17	٠.٤٠
COM52	٠.٣٥	COM72	٠.٦٧	COM18	٠.٥١
COM53	٠.٣٨	COM73	٠.٥٧	C0M19	٠.٤٦
COM54	٠.٣٢	COM74	٠.٤٨	COM20	٠.١٩
		COM75	٠.٦٣	COM21	٠.٣٩
				COM22	٠.٣٧
				COM23	٠.٣١
				COM24	٠.١٧

يتضح من جدول (٢٥) ما يلي:

- إن عددا كبيرا من المفردات يتمتع بدرجة اتساق مناسبة في قياس ما يقيسه الاختبار، حيث تجاوزت قيم معامل ارتباط المفردة بالدرجة الكلية للصورة الاختبارية التي تنتمي إليها القيمة (٠.٣٠)، في حين انخفضت قيمة معامل ارتباط (١٥) مفردة عن هذه القيمة.

ونظرا لأن الدراسة الحالية تستهدف المقارنة بين اثنتين من طرق التعادل القائمة على

نظرية الاستجابة للمفردة فلم تقم الباحثة بحذف أي من هذه المفردات وفق هذه الخطوة؛ حيث إن طرق التحليل القائم على نظرية الاستجابة للمفردة تتضمن إجراءات تحديد وحذف المفردات غير الملائمة لأسس القياس الموضوعي، وتستهدف الدراسة الحالية المقارنة بين الطريقتين من حيث نسبة المفردات غير الملائمة وفق كل طريقة منها.

- **منهج الدراسة**: للإجابة عن أسئلة الدراسة استخدم المنهج الوصفي.

- **أساليب المعالجة الإحصائية**:

- الإحصاءات الوصفية مشتملة على (المتوسط، الوسيط، الانحراف المعياري، معامل الاختلاف، المدى الربيعي).

- معامل الثبات الهامشي، ومعامل الفصل، والخطأ المعياري لتقديرات القدرة والصعوبة، وحجم التأثير.

- اختبار شابيرو – ويلك للتحقق من اعتدالية توزيع متغيرات الدراسة المتصلة (تقديرات صعوبة المفردات-تقديرات قدرة الأفراد-تقديرات الخطأ المعياري لتقديرات الصعوبة والقدرة).

- التحليل الإحصائي للبيانات وفق نموذج راش باستخدام برنامج التحليل الإحصائي WINSTEPS للحصول على تقديرات الأفراد والمفردات وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة.

- اختبار دلالة الفرق بين نسبتين مرتبطتين لدراسة فروق النسب بين نسبة المفردات المحذوفة وفق طريقتي التعادل، وكذلك دراسة فروق النسب بين نسبة الأفراد المحذوفين وفق طريقتي التعادل.

- معاملات الارتباط بين تقديرات الأفراد وفق طريقتي التعادل، وكذلك بين تقديرات المفردات وفق طريقتي التعادل.

- معامل فاي ومعامل كابا لدراسة الاتساق بين طريقتي التعادل في تصنيف الأفراد والمفردات وفق محكات الملائمة.

- اختبار مان وتني للمقارنة بين قيم الخطأ المعياري لتقديرات صعوبة المفردات بالنسبة لكل من طريقتي التعادل. وكذلك المقارنة بين قيم الخطأ المعياري لتقديرات قدرات الأفراد بالنسبة لكل من طريقتي التعادل.

- اختبار Moses Test لدلالة تساوي المدى لتقديرات صعوبات المفردات وقدرات الأفراد وفق طريقتي التعادل.

نتائج الدراسة:

تتناول الدراسة في هذا الجزء عرض نتائج الدراسة، وذلك على محورين: النتائج المرتبطة بإحصاءات المفردات وفق طريقتي التعادل، ثم النتائج المرتبطة بإحصاءات الأفراد، ثم يختتم هذا الجزء بالتوصيات والبحوث المقترحة في ضوء ما أسفرت عنه الدراسة من نتائج.

نتائج الإجابة على السؤال الأول وينص على:

"هل تختلف النسبة المئوية للمفردات الاختبارية غير الملائمة لأسس القياس باختلاف طريقة التعادل؟"

للإجابة عن هذا السؤال تم إجراء التحليل الإحصائي للبيانات الكلية لمفردات الصور الاختبارية الثلاث (٧٥ مفردة، ١١٩٨ فردا) باستخدام برنامج WINSTEPS مرة باستخدام طريقة التعادل المتزامن، ومرة ثانية باستخدام طريقة التعادل بتثبيت بارامترات المفردات المشتركة، ووفقاً لإحصاءات الملاءمة التي يوفرها برنامج WINSTEPS اعتبر تجاوز إحصاءات الملاءمة المدى (±٢) وفقا لإحصائي الملاءمة التقاربي والتباعدي للمؤشر (ZSTD) محكاً لعدم ملاءمة المفردة.

- ولاختبار الدلالة الإحصائية للفروق بين نسبتي المفردات غير الملائمة وفق طريقتي التعادل استخدم اختبار النسبة الحرجة لدلالة الفرق بين نسبتين مرتبطتين كما هو موضح بجدول (٢٦) التالي.

جدول (٢٦)

اختبار النسبة الحرجة للفرق بين نسبتي المفردات غير الملائمة وفق طريقتي التعادل.

طريقة التعادل	عدد المفردات	عدد المفردات المحذوفة	النسبة المئوية للمفردات المحذوفة%	Z	مستوى الدلالة
المتزامن	٧٥	٩	١٢	٠.٧٩٤	غير دالة
تثبيت بارامترات المفردات المشتركة	٧٥	٧	٩.٣٣		

يتضح من جدول (٢٦) ما يلي:

- عدم اختلاف نسبة المفردات غير الملائمة لأسس القياس الموضوعي في حالتي التعادل حيث بلغت قيمة (Z) (٠.٧٩٤) وهي غير دالة إحصائيا عند مستوى دلالة (٠.٠٥).
- ولاختبار دلالة الاتفاق بين طريقتي التعادل في الحكم على ملاءمة المفردات استخدمت الباحثة معامل فاي ويتضمن جدول (٢٧) التالي نتائج هذا الاختبار.

جدول (٢٧)

حجم ودلالة الاتساق بين طريقتي التعادل في تقييم ملاءمة مفردات الاختبار

		التحليل المتزامن		المجموع	قيمة معامل فاي	الدلالة	معامل كابا	الدلالة
		غير ملائم	ملائم					
التحليل تثبيت بارامترات المفردات المشتركة	غير ملائم	٦	١	٧	**٠.٧٢٨	٠.٠٠٠	**٠.٧٢١	٠.٠٠٠
	ملائم	٣	٦٥	٦٨				
المجموع		٩	٩	٦٦	٧٥			

- يتضح من جدول (٢٧) ما يلي:
- أسفر التحليل الإحصائي للبيانات باستخدام طريقة التعادل المتزامن عن حذف (٩) مفردات بنسبة بلغت (١٢%) من إجمالي عدد المفردات، وقد أسفر التحليل الإحصائي للبيانات باستخدام طريقة التعادل بتثبيت بارامترات المفردات المشتركة عن حذف (٧) مفردات بنسبة بلغت (٩.٣٣%) من إجمالي عدد المفردات.
- اتفقت طريقتي التعادل على ملاءمة (٦٥) مفردة كما اتفقت على عدم ملاءمة (٦) مفردات.
- صنفت مفردة واحدة على إنها ملائمة وفقا لطريقة التعادل المتزامن؛ في حين صنفت على أنها غير ملائمة وفق طريقة التعادل بتثبيت بارامترات المفردات المشتركة.
- صنفت (٣) مفردات على أنها ملائمة وفق طريقة التحليل باستخدام التعادل بتثبيت بارامترات المفردات المشتركة؛ في حين صنفت غير ملائمة وفق طريقة التعادل المتزامن.
- اتساق طريقتي التعادل في الحكم على ملاءمة المفردات الاختبارية؛ حيث إن معامل الارتباط قوى ودال إحصائياً عند مستوى دلالة أقل من (٠.٠١) وقد أكدت هذه النتائج معامل الاتفاق (كابا) حيث جاء الاتساق كبيرا (٠.٧٢١ و ٠.٠٠٠**) ودالاً إحصائياً.

وقد يرجع انخفاض عدد المفردات المحذوفة وفق طريقة التعادل بتثبيت بارامترات المفردات المشتركة إلى أنه وفق هذه الطريقة يتم تحليل كل صورة اختبارية بشكل مستقل أولا وتحديد الأفراد والمفردات غير الملائمين قبل ربط الصور الاختبارية بالصورة المرجعية؛ حيث يكون الحكم على ملاءمة الأفراد والمفردات وفق الصورة الاختبارية التي تعرض لها الفرد، والتي وردت فيها المفردات. ثم يأتي بعد ذلك ربط الصور الاختبارية بالإطار المرجعي للبنك، وعند الربط مع الإطار المرجعي فإن بعض المفردات المشتركة قد تكون ملائمة في إحدى الصور وغير ملائمة في صورة أخرى؛ لذا يستبقى عليها في البنك ولكن لا تستخدم كمفردة مشتركة وإنما تستخدم فقط كمفردة ملائمة من المفردات المكون للبنك. على سبيل المثال فإن المفردة المشتركة الأولى (كود المفردة COM1) جاءت ملائمة وفق تدريج الصورة المرجعية ولكنها كانت غير

ملائمة بالنسبة للصورة الاختبارية الثانية؛ لذلك تم الإبقاء على هذه المفردة في البنك كمفردة ملائمة ولكن تغيرت وظيفتها حيث لم تعد مفردة مشتركة عند ربط الصورة الثانية بالصورة الاختبارية المرجعية، وإنما اعتبرت مفردة من مكونات البنك.

أما بالنسبة للتعادل المتزامن فإن عدم ملاءمة مفردة مشتركة يعني حذفها من البنك تماما؛ حيث يتم تحليل كافة الصور مرة واحدة. على سبيل المثال المفردة المشتركة الأولى (كود المفردة COM1) جاءت غير ملائمة وفق التدريج المتزامن لذلك حذفت تماما من البنك عند التحليل وفق هذه الطريقة. وتتسق هذه النتائج مع ما أشار إليه (Hanson&Begin,2002) من أن طريقة التدريج المتزامن تضع قيودا كثيرة على تقديرات بارامترات المفردات أكثر من طريقة تثبيت بارامترات المفردات المشتركة خاصة عندما يتم تدريج كثير من الصور الاختبارية في آن واحد.

نتائج الإجابة عن السؤال الثاني وينص على:

"إلى أي مدى تتسق تقديرات المفردات على متصل الصعوبة وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة؟"

نظراً لإجراء تحليل البيانات وفق كل من طريقتي التعادل بشكل مستقل عن الأخرى فإن لكل منهما **صفر تدريج مختلف**؛ لذا فقد تم استخدام المدى ومعامل الاختلاف ومعامل الارتباط بين تقديرات الصعوبة وفق طريقتي التعادل، واقتصرت المقارنة على **المفردات الملائمة فقط** وفق طريقتي التعادل، ويتضمن جدول (٢٨) التالي الوصف الإحصائي لتقديرات صعوبات المفردات الملائمة وفق طريقتي التعادل ومعامل الارتباط بينهما.

جدول (٢٨)

الوصف الإحصائي لتقديرات صعوبات المفردات الملائمة وفق طريقتي التعادل ومعامل الارتباط بينهما.

طريقة التعادل	تقديرات الصعوبة بالمنف			معامل الاختلاف	معامل الارتباط	مستوى الدلالة
	المدى	المتوسط	الانحراف المعياري			
التعادل المتزامن	٢٤.٣٠	٤٩.٧٩٥٤	٤.٩٣٤٥٨	٠.٠٩٩	**٠.٩٨٨	٠.٠٠٠
التعادل تثبيت بارامترات المفردات المشتركة	٢٣.٦٠	٤٩.٩٢٦٩	٤.٧٨٠١٢	٠.٠٩٦		

ويتضح من جدول (٢٨) ما يلي:

- لم يختلف تباين تقديرات صعوبات المفردات وفق طريقتي التعادل المتزامن والتعادل عن طريق تثبيت بارامترات المفردات المشتركة حيث تساوت قيم معاملي الاختلاف تقريبا

مما يشير إلى عدم تغير تشتت تقديرات هذه الصعوبات وفق طريقتي التعادل.

- لم يختلف مدى تقديرات صعوبة المفردات الملائمة وفق طريقتي التعادل حيث لم يتجاوز الفرق بين قيمتي المدى متوسط الخطأ المعياري لتقديرات صعوبة تلك المفردات في أي من طريقتي التعادل.
- هناك اتساق في توزيع المفردات على متصل الصعوبة وفق طريقتي التعادل؛ حيث جاء معامل ارتباط بيرسون[٤] بين قيم تقديرات الصعوبة وفق طريقتي التعادل المتزامن وتثبيت بارامترات المفردات المشتركة طردي قوى -يكاد يكون تام- دال إحصائياً عند مستوى دلالة أقل من (٠.٠١). وتتسق هذه النتائج مع دراسة (Taherbhai&Sea,2017) حيث أشارت إلى عدم وجود فروق بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة في تدريج الصور الاختبارية المتكافئة. وقد يرجع ذلك إلى أن الباحثة قد راعت تكافؤ الصور الاختبارية في مرحلة بنائها وتكوينها من حيث تمثيل المحتوى المعرفي للوحدة محل الدراسة حيث تقيس كل صورة اختبارية من الصور الثلاث نواتج التعلم الرئيسة بوحدة المادة وتركيبها، كما حرصت على تكافؤ الصورتين من حيث مستوى الصعوبة من وجهة نظر المحكمين خلال مرحلة تكوين الصور الاختبارية مما ساعد في الحد من تأثير العوامل التي يمكن أن تؤثر على عملية التعادل خاصة فيما يتعلق بالتفاوت بين الصور الاختبارية من حيث مستوى صعوباتها أو ما تقيسه من نواتج تعلم. بالإضافة إلى أنه وفق طريقتي التعادل قد تم حذف الأفراد غير الملائمين والمفردات غير الملائمة مما يساعد على الإبقاء على الاستجابات الصادقة على المفردات الملائمة فيقلل من أخطاء القياس والدقة في التقدير.

الإجابة عن السؤال الثالث وينص على:

"إلى أي مدى تحقق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة دقة تقديرات صعوبات المفردات؟"

يوفر برنامج التحليل الاحصائي WINSTEPS عددا من المؤشرات الإحصائية كمؤشرات لدقة تدرج صعوبات مفردات الاختبار ككل منها معامل الثبات الهامشي، ومتوسط الخطأ المعياري

---

[٤] استخدم معامل ارتباط بيرسون نظرا لاعتدالية توزيع تقديرات الصعوبة وفق طريقتي التدريج المتزامن والتدريج بتثبيت بارامترات المفردات المشتركة حيث جاء اختبار اعتدالية التوزيع لكل من شابيرو ويلك Shapiro-Wilk غير دالة إحصائيا.

لتقديرات الصعوبة، ومعامل الفصل Separation Index، وكذلك قيمة دالة المعلومات للاختبار ككل، كما يوفر قيمة للخطأ المعياري المصاحبة لتقدير صعوبة كل مفردة على حدة؛ لذلك اعتمدت الباحثة في الإجابة على هذا السؤال على مقارنة مؤشرات دقة الاختبار الكلية لكل صورة اختبارية، ثم إجراء المقارنات بين متوسطات الأخطاء المعيارية المصاحبة لكل مفردة.

ويتضمن جدول (٢٩) التالي ملخصاً لمؤشرات دقة تقديرات صعوبات المفردات المستخلصة من التدريج وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة كل على حدة.

جدول (٢٩)

مؤشرات دقة التدريج النهائي لتقديرات المفردات وفق طريقتي التعادل.

الإحصاءة	الصورة الأولى		الصورة الثانية		الصورة الثالثة	
	تثبيت بارامترات المفردات	التدريج المتزامن	تثبيت بارامترات المفردات	التدريج المتزامن	تثبيت بارامترات المفردات	التدريج المتزامن
العدد النهائي للمفردات	٢٨	٢٦	٢٦	٢٨	٣٠	٣٠
معامل الثبات لتقديرات الصعوبة	٠.٩٨	٠.٩٨	٠.٩٧	٠.٩٨	٠.٩٨	٠.٩٨
متوسط الخطأ المعياري لتقدير صعوبة المفردات.	٠.١٣	٠.١٤	٠.١٥	٠.١٥	٠.١٥	٠.١٥
معامل الفصل لتقدير صعوبة المفردات.	٦.٥٣	٧.٠٩	٥.٨٦	٦.٢٣	٧.٢٨	٧.٩٢
قيمة دالة المعلومات.	٦.١	٥.٤٤	٥.٦٧	٦.٠٢	٥.٩٦	٥.٨

يتضح من جدول (٢٩) ما يلي:

- ارتفاع قيمة معامل الثبات الهامشي لتقديرات صعوبة المفردات حيث بلغت ٠.٩٨ لكافة الصور الاختبارية وفق طريقتي التعادل المتزامن، والتعادل بتثبيت بارامترات المفردات المشتركة؛ فيما عدا معامل الثبات لتقديرات صعوبة مفردات الصورة الثانية وفق طريقة تثبيت بارامترات المفردات المشتركة حيث بلغ ٠.٩٧.

- ارتفاع قيمة معامل الفصل لتقديرات صعوبة المفردات؛ حيث تراوحت قيمته بين (٥.٨٦: ٧.٢٨) بالنسبة لطريقة التعادل بتثبيت بارامترات المفردات المشتركة؛ وتراوحت بين (٦.٢٣: ٧.٩٢) بالنسبة لطريقة التعادل المتزامن، وجمعها قيم مرتفعة جدا. وحيث إن معامل الفصل للمفردات يعرف بأنه النسبة بين الانحراف المعياري لتقديرات المفردات، ومتوسط الخطأ المعياري لهذه القيم. ويستخدم للتعبير عن مدى كفاية مجموعة المفردات لتعريف متصل السمة الذي تقيسه، ووفق دليل البرنامج فإنه كلما زادت قيمة معامل الفصل عن (٢) فإن ذلك يعد مؤشرًا على أن المفردات المتضمنة بالاختبار كافية لتعريف متصل السمة المقاسة. وقد تجاوزت

قيم معامل الفصل لكافة الصور الاختبارية وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة هذه القيمة؛ مما يشير إلى كفاية هذه الصور وفق طريقتي التدريج لتغطية متصل التحصيل الدراسي في وحدة المادة وتركيبها محل الدراسة. إلا إنه يلاحظ ارتفاع قيم معاملات الفصل لتقديرات صعوبة المفردات وفق طريقة التدريج المتزامن قليلا بالنسبة للصور الثلاث مقارنة بمثيلاتها وفق طريقة التعادل بتثبيت بارامترات المفردات المشتركة.

- انخفاض قيمة متوسط الخطأ المعياري لتقديرات صعوبة المفردات حيث تراوحت قيمته بين (٠.١٣: ٠.١٥) بالنسبة لطريقة التعادل بتثبيت بارامترات المفردات المشتركة؛ وتراوحت بين (٠.١٤: ٠.١٥) بالنسبة لطريقة التعادل المتزامن، وجميعها قيم **منخفضة**. حيث تساوت متوسط قيم الخطأ المعياري لتقديرات الصعوبة بالنسبة لكافة الصور الاختبارية وفق طريقتي التعادل وانخفضت فقط بالنسبة للصورة الأولى وفق طريقة تثبيت بارامترات المفردات المشتركة عن مثيلتها بالنسبة لطريقة التعادل المتزامن بمقدار (٠.٠١). مما يشير إلى دقة القياس وفق طريقتي التعادل.

- ارتفاع قيمة دالة المعلومات للصور الاختبارية الثلاث وفق طريقتي التعادل حيث تراوحت قيمتها بين (٥.٦٧: ٦.١) بالنسبة لطريقة التعادل بتثبيت بارامترات المفردات المشتركة؛ وتراوحت بين (٥.٤٤: ٦.٠٢) بالنسبة لطريقة التعادل المتزامن. أي أن كلتا الطريقتين تقدمان كما مناسباً من المعلومات. خاصة عند مستوى القدرة المتوسط ويشير ارتفاع قيمة دالة المعلومات إلى دقة القياس باستخدام كل من طريقتي التعادل، ويتفق ذلك مع ما اتضح من انخفاض لقيمة الخطأ المعياري، وارتفاع قيم الثبات ومعاملات الفصل للصور الاختبارية الثلاث وفق طريقتي التعادل المتزامن، وتثبيت بارامترات المفردات المشتركة. وحيث إن قيمة دالة المعلومات تتأثر بعدد مفردات الاختبار؛ فقد كانت قيمة دالة المعلومات للصور الاختبارية وفق طريقة تثبيت بارامترات المفردات المشتركة أكبر بالنسبة للصورتين الاختباريتين الأولى والثانية؛ حيث كان عدد المفردات الملائمة أكبر، كما زادت قيمة دالة المعلومات قليلا بالنسبة وفق هذه الطريقة عند تساوي عدد مفردات الصورة الاختبارية الملائمة وفق طريقتي التعادل، وذلك كما هو الحال في الصورة الثالثة.

<u>بالنسبة للخطأ المعياري لتقديرات صعوبة المفردات</u>

تمت المقارنة بين الخطأ المعياري لتقديرات صعوبة مفردات الاختبار ككل الملائمة وفق كل طريقة من طرق التعادل، وكذا الارتباط بين القيم المتناظرة للخطأ المعياري لتقديرات صعوبة المفردات الملائمة وفق طريقتي التعادل معا وفيما يلي نتائج كل منها.

-**نتائج المقارنة بين قيم الخطأ المعياري لتقديرات صعوبة مفردات الاختبار الملائمة وفق كل طريقة من طرق التعادل.** استخدم اختبار مان ويتني نظرا لالتواء توزيع قيم الخطأ المعياري، ويتضمن جدول (٣٠) التالي نتائج هذا الاختبار.

جدول (٣٠)

نتائج اختبار مان ويتني للفروق بين قيم الخطأ المعياري لتقديرات الصعوبة وفق طريقتي التعادل.

طريقة التعادل	عدد المفردات	مجموع الرتب	Z	مستوى الدلالة
المتزامن	٦٦	٤٣٨٠	٠.٣٤٣	٠.٧٣٢
تثبيت بارامترات المفردات المشتركة	٦٨	٤٦٦٥		

يتضح من جدول (٣٠) ما يلي:

- لا يوجد فرق دال إحصائيا بين قيم الخطأ المعياري لتقديرات الصعوبة لمفردات الصور الاختبارية الثلاثة الملائمة في تدريجها النهائي وفق طريقتي التعادل.
- **الارتباط بين قيم الخطأ المعياري لتقديرات صعوبة مفردات الاختبار الملائمة وفق طريقتي التعادل معا.** استخدم معامل ارتباط "سبيرمان" نظرا لالتواء التوزيع للمقارنة بين قيم الخطأ المعياري لتقديرات صعوبات المفردات الملائمة وفق كل من طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة.

جدول (٣١)

معامل الارتباط بين قيم الخطأ المعياري لتقديرات الصعوبة للمفردات الملائمة وفق طريقتي التعادل.

طريقة التعادل	المدى	الوسيط	الانحراف الربيعي	معامل الارتباط	مستوى الدلالة
التعادل المتزامن	٠.٧٥	٠.٧٠	٠.١٠	**٠.٨٦٢	٠.٠٠٠
التعادل بتثبيت بارامترات المفردات المشتركة	٠.٥٠	٠.٧٠	٠.١٠		

يتضح من جدول (٣١) ما يلي:

- اتساق توزيع قيم الخطأ المعياري للمفردات **الملائمة** في حالتي التعادل، حيث جاء معامل ارتباط سبيرمان طردي قوى دال إحصائياً عند مستوى دلالة أقل من (٠.٠١) بين تقديرات الخطأ المعياري لتقديرات صعوبات المفردات الملائمة وفق كل من طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة.

كما أن تشتت قيم الخطأ المعياري للمفردات الملائمة وفق طريقتي التعادل جاء متساوياً في

الحالتين حيث بلغت قيمة الانحراف الربيعي (٠.١٠) منف، مع تساوي قيمتي الوسيط (٠.٧) منف.

وتتسق هذه النتائج مع دراسة (Taherbhai&Sea,2017) حيث أشارت إلى عدم وجود فروق بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة في تدريج الصور الاختبارية المتكافئة.

نتائج الإجابة عن السؤال الرابع وينص على:

"هل يختلف مدى تقديرات صعوبة المفردات الاختبارية باختلاف طريقة التعادل؟"

لدراسة تساوي مدى تقديرات صعوبة المفردات الاختبارية باختلاف طريقة التعادل استخدم اختبار Moses لدلالة الفروق بين مدى توزيعين من خلال تحديد القيم المتطرفة لتوزيع التقديرات وفق كل طريقة من طرق التعادل.

جدول(٣٢):

نتائج اختبار Moses لتقديرات صعوبات المفردات وفق طريقتي التعادل.

	تقدير الصعوبة بالمنف
Observed Control Group Span	١٣٤
Sig. (1-tailed)	١,٠٠٠
Trimmed Control Group Span	١٢٢
Sig. (1-tailed)	٠,٧٣٩
Outliers Trimmed from each End	٣

يتضح من جدول (٣٢) أنه لا توجد فروق دالة إحصائيا بين توزيع تقديرات صعوبات المفردات بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات في توزيع القيم المتطرفة؛ مما يشير إلى تساوي مدى توزيع تقديرات الصعوبة وفق طريقتي التعادل.

النتائج المتعلقة بإحصاءات الأفراد:

نتائج الإجابة على السؤال الخامس وينص على:

هل تختلف النسبة المئوية للأفراد غير الملائمين لأسس القياس باختلاف طريقة التعادل؟

للإجابة عن هذا السؤال تم إجراء التحليل الإحصائي للبيانات الكلية لمفردات الصور الاختبارية الثلاثة (٧٥ مفردة، ١١٩٨ فردا) باستخدام برنامج WINSTEPS مرة باستخدام طريقة التعادل المتزامن، ومرة ثانية باستخدام طريقة التعادل بتثبيت بارامترات المفردات المشتركة، ووفقاً لإحصاءات الملاءمة التي يوفرها برنامج WINSTEPS اعتبر تجاوز إحصاءات الملاءمة المدى (±٢) وفقا لإحصائي الملاءمة التقاربي والتباعدي للمؤشر (ZSTD) محكاً لعدم ملاءمة الفرد.

- ولاختبار الدلالة الإحصائية للفروق بين نسبتي الأفراد غير الملائمين وفق طريقتي التعادل استخدم اختبار النسبة الحرجة لدلالة الفرق بين نسبتين مرتبطتين كما هو موضح بجدول (٣٣) التالي.

جدول (٣٣)

اختبار النسبة الحرجة للفروق بين نسبتي الأفراد غير الملائمين وفق طريقتي التعادل.

طريقة التعادل	عدد الأفراد	عدد الأفراد المحذوفين	النسبة المئوية للأفراد المحذوفين%	Z	مستوى الدلالة
المتزامن	١١٩٨	٢٢٦	٠.١٨٨٦	١.٢١٦	غير دالة
تثبيت بارامترات المفردات المشتركة	١١٩٨	٢١٠	٠.١٧٥٣		

يتضح من جدول (٣٣) ما يلي:

- عدم اختلاف نسبة الأفراد غير الملائمين لأسس القياس الموضوعي في حالتي التعادل؛ حيث بلغت قيمة (Z) (١.٢١٦) وهي غير دالة إحصائيا عند مستوى دلالة (٠.٠٥).

ولاختبار دلالة الاتفاق بين طريقتي التعادل في الحكم على ملاءمة الأفراد (ملائم/ غير ملائم) لأسس القياس الموضوعي استخدمت الباحثة معامل فاي ويتضمن جدول (٣٤) التالي نتائج هذا الاختبار.

جدول (٣٤):

حجم ودلالة الاتساق بين طريقتي التعادل في تقييم ملاءمة الأفراد.

		التعادل بتثبيت بارامترات المفردات المشتركة		المجموع	قيمة معامل فاي	الدلالة	معامل كابا	الدلالة
		غير ملائم	ملائم					
التعادل المتزامن	غير ملائم	١٩١	٣٥	٢٢٦	٠.٨٤٩	٠.٠٠٠	٠.٨٤٩	٠.٠٠٠
	ملائم	١٩	٩٥٣	٩٧٢				
المجموع		٢١٠ .	٩٨٨	١١٩٨				

يتضح من جدول (٣٤) ما يلي:

- أسفر التحليل الإحصائي للبيانات باستخدام طريقة التعادل المتزامن عن حذف (٢٢٦) فردا بنسبة بلغت (١٨.٨٦%) من إجمالي عدد الأفراد، في حين أسفر التحليل الإحصائي للبيانات باستخدام طريقة التعادل بتثبيت بارامترات المفردات المشتركة عن حذف (٢١٠) فردا بنسبة بلغت (١٧.٥٣%) من إجمالي عدد الأفراد.
- جاء (٩٥١) فرداً ملائماً **وفقاً لطريقتي التعادل معًا**، وجاء (١٩) فرداً غير ملائم وفق طريقة التعادل بتثبيت بارامترات المفردات المشتركة **فقط**، و(٣٥) فردا غير ملائم وفقا لطريقة التعادل المتزامن **فقط**، في حين جاء (١٩١) فردا غير ملائم **وفقا لطريقتي التعادل معا.**

- هناك اتساق بين طريقتي التعادل في الحكم على ملاءمة الأفراد (ملائم/ غير ملائم)؛ حيث جاء معامل الارتباط قوى ودال إحصائياً عند مستوى دلالة أقل من (٠.٠١). وأكد هذه النتائج معامل الاتفاق (كابا) حيث جاء الاتساق كبيرا ودالاً إحصائياً (٠.٨٤٩ و ٠.٠٠٠***).

نتائج الإجابة على السؤال السادس وينص على:

إلى أي مدى تتسق تقديرات الأفراد على متصل القدرة وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة؟

نظراً لإجراء تحليل البيانات وفق كل من طريقتي التعادل بشكل مستقل عن الأخرى فإن لكل منهما **صفر تدريج مختلف**؛ لذا فقد تم استخدام المدى ومعامل الاختلاف ومعامل الارتباط بين تقديرات القدرة وفق طريقتي التعادل، واقتصرت المقارنة على **الأفراد الملائمين فقط** وفق الطريقتين معا. ويتضمن جدول (٣٥) التالي الوصف الإحصائي لتقديرات قدرات الأفراد الملائمين وفق طريقتي التعادل ومعامل الارتباط بينهما.

جدول (٣٥)

الوصف الإحصائي لتقديرات قدرات الأفراد الملائمين وفق طريقتي التعادل ومعامل الارتباط بينهما.

طريقة التعادل	تقديرات القدرة بالمنف			معامل الاختلاف	معامل الارتباط	مستوى الدلالة
	المتوسط	المدى	الانحراف المعياري			
المتزامن	٥٢,٨٤	٣٨,٧٠	٧.٤٣	٠.١٤	٠.٩٨٩**	٠.٠٠٠
تثبيت بارامترات المفردات المشتركة	٥٣,٠٨	٣٩,٢٠	٧.٦٢	٠.١٤		

يتضح من جدول (٣٥) ما يلي:

- اتساق توزيع الأفراد على متصل القدرة وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة؛ حيث كان معامل الارتباط طرديا قويا (يكاد يكون تاما) دالا إحصائياً عند مستوى دلالة أقل من (٠.٠١).
- لا يختلف تشتت تقديرات القدرة وفق طريقتي التعادل؛ حيث تساوت قيم معاملي الاختلاف مما يشير إلى عدم اختلاف تباين تلك التقديرات في الحالتين.

الإجابة عن السؤال السابع وينص على:

"إلى أي مدى تحقق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة دقة التقدير لقدرات الأفراد؟"

يوفر برنامج التحليل الإحصائي WINSTEPS عددا من المؤشرات الإحصائية كمؤشرات لدقة تدرج قدرات الأفراد ككل منها معامل الثبات الهامشي، ومتوسط الخطأ المعياري لتقديرات القدرة، ومعامل الفصل Separation Index، كما يوفر قيمة للخطأ المعياري المصاحب لتقدير قدرة كل فرد على حدة لذلك اعتمدت الباحثة في الإجابة عن هذا السؤال على مقارنة مؤشرات دقة تقديرات القدرة ككل، ثم إجراء المقارنات بين متوسطات الأخطاء المعيارية المصاحبة لكل فرد.

ويتضمن جدول (٣٦) التالي ملخصاً لمؤشرات دقة تقديرات قدرات الأفراد المستخلصة من التدريج وفق طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة كل على حدة.

جدول (٣٦)

مؤشرات دقة التدريج النهائي لتقديرات القدرة وفق طريقتي التعادل.

الإحصاءة	الصورة الأولى		الصورة الثانية		الصورة الثالثة	
	تثبيت بارامترات المفردات	التدريج المتزامن	تثبيت بارامترات المفردات	التدريج المتزامن	تثبيت بارامترات المفردات	التدريج المتزامن
معامل الثبات لتقديرات القدرة.	٠.٧٨	٠.٧٨	٠.٨٥	٠.٨٦	٠.٨٠	٠.٨١
متوسط الخطأ المعياري لتقدير القدرة.	٠.٤٥	٠.٤٩	٠.٥٥	٠.٥٢	٠.٥٠	٠.٥١
معامل الفصل لتقدير القدرة.	١.٨٩	١.٩٠	٢.٣٨	٢.٥	٢.٠٣	٢.٠٤

يتضح من جدول (٣٦) ما يلي:

- ارتفاع قيمة معامل الثبات الهامشي لتقديرات قدرات الأفراد حيث تراوحت بين (٠.٧٨: ٠.٨٥) بالنسبة لطريقة التعادل بتثبيت بارامترات المفردات المشتركة وتراوحت بين (٠.٧٨: ٠.٨٦) بالنسبة لطريقة التدريج المتزامن. وتساوت قيم معامل الثبات تماما بالنسبة للصورة الأولى، وكان الفرق بين معاملات الثبات صغيرا جدا لم يتجاوز ٠.٠١ بالنسبة للصورتين الثانية والثالثة.

- ارتفاع قيمة معامل الفصل لتقديرات قدرة الأفراد؛ حيث تراوحت قيمته بين (١.٨٩: ٢.٣٨) بالنسبة لطريقة التعادل بتثبيت بارامترات المفردات المشتركة؛ وتراوحت بين (١.٩: ٢.٥) بالنسبة لطريقة التعادل المتزامن، وجميعها قيم مرتفعة، مما يشير إلى أن عينة الأفراد كافية للفصل بين المفردات وفق كل من طريقتي التعادل بالنسبة للصور الاختبارية الثلاث. إلا إنه يلاحظ ارتفاع قيم معاملات الفصل لتقديرات قدرة الأفراد وفق طريقة التدريج المتزامن قليلا بالنسبة للصور الثلاث مقارنة بمثيلاتها وفق طريقة التعادل بتثبيت بارامترات المفردات المشتركة حيث لم يتجاوز هذا الفرق ٠.١٢.

- انخفاض قيمة متوسط الخطأ المعياري لتقديرات قدرة الأفراد حيث تراوحت قيمته بين

(٠.٤٥: ٠.٥٥) بالنسبة لطريقة التعادل بتثبيت بارامترات المفردات المشتركة؛ وتراوحت بين (٠.٤٩: ٠.٥٢) بالنسبة لطريقة التعادل المتزامن، وجميعها قيم منخفضة.

- ب-بالنسبة للخطأ المعياري لتقديرات قدرات الأفراد عينة التدريج

تمت المقارنة بين الخطأ المعياري لتقديرات قدرات الأفراد عينة التدريج ككل الملائمة وفق كل طريقة من طرق التعادل، وكذا الارتباط بين القيم المتناظرة للخطأ المعياري لتقديرات قدرات الأفراد الملائمة وفق طريقتي التعادل معا وفيما يلي نتائج كل منها.

- **نتائج المقارنة بين قيم الخطأ المعياري لتقديرات قدرات الأفراد عينة التدريج الملائمة وفق كل طريقة من طرق التعادل.** استخدم اختبار مان ويتني نظرا لالتواء توزيع قيم الخطأ المعياري، ويتضمن جدول (٣٧) التالي نتائج هذا الاختبار.

جدول (٣٧)

نتائج اختبار مان ويتني للفروق بين قيم الخطأ المعياري لتقديرات قدرات الأفراد وفق طريقتي التعادل.

طريقة التعادل	عدد الأفراد	مجموع الرتب	Z	مستوى الدلالة	حجم التأثير
المتزامن	٩٧٢	١٠١٤٢٤٩	٤.٩٠٩	٠.٠٠٠	٠.١٢٧ صغير
تثبيت بارامترات المفردات المشتركة	٩٨٨	٩٠٧٥٣١			

يتضح من جدول (٣٧) ما يلي:

- توجد فروق دالة إحصائيا عند مستوى (٠.٠١) بين قيم الخطأ المعياري لتقديرات قدرات الأفراد قدرة الأفراد عينة التدريج وقد جاء حجم التأثير صغيرا. حيث كانت قيم الأخطاء المعيارية المصاحبة لتقديرات قدرة الأفراد وفق طريقة التدريج المتزامن أكبر منها بالنسبة لطريقة تثبيت بارامترات المفردات المشتركة. ويعني ذلك أن طريقة التعادل بتثبيت بارامترات المفردات المشتركة ينتج تقديرات لقدرات الأفراد أكثر دقة مقارنة بطريقة التدريج المتزامن، وإن كان حجم التأثير صغيرا. ويتفق ذلك مع نتيجة دراسة (Hu et al.,2008)(Pang et al.,2010) وتتعارض مع نتيجة دراسة(Arai&Mayekawa,2011)(Keller&Keller,2015) وقد يرجع ذلك إلى أن طريقة التعادل المتزامن تفرض قيودا كثيرة خلال مرحلة التحليل مما قد يؤدي إلى ارتفاع الأخطاء المعيارية لتقديرات قدرات الأفراد. وفي هذا الصدد يشير Kolen and Brennan (2004) إلى أنه رغم استخدام التحليل المتزامن كافة المعلومات المتاحة حيث يتم تحليل كافة الاستجابات في آن واحد إلا أن ذلك قد يؤدى إلى انتهاك افتراضية أحادية البعد

وفى هذه الحالة قد تكون طريقة التدريج تثبيت بارامترات المفردات المشتركة أفضل.

- الارتباط بين قيم الخطأ المعياري لتقديرات قدرة عينة التدريج الملائمة وفق طريقتي التعادل معا.

استخدم معامل ارتباط "سبيرمان" نظرا لالتواء التوزيع للمقارنة بين قيم الخطأ المعياري - لتقديرات قدرة عينة التدريج الملائمة وفق كل من طريقتي التعادل.

جدول (٣٨)

معامل الارتباط بين قيم الخطأ المعياري لتقديرات قدرة الأفراد الملائمة وفق طريقتي التعادل.

طريقة التعادل	المدى	الوسيط	الانحراف الربيعي	معامل الارتباط	مستوى الدلالة
التعادل المتزامن	٧,٢٥	٢,٢٥٠٠	٠,٤	٠,٨٣٧**	٠,٠٠٠
التعادل بتثبيت بارامترات المفردات المشتركة	٧,٣٠	٢,٢٠٠٠	٠,٤٥		

يتضح من جدول (٣٨) ما يلي:

- اتساق توزيع قيم الخطأ المعياري لقدرات الأفراد الملائمين في حالتي التعادل، حيث جاء معامل ارتباط سبيرمان طرديا قويا دالا إحصائياً عند مستوى دلالة أقل من (٠.٠١) بين تقديرات الخطأ المعياري لتقديرات قدرات الأفراد الملائمين وفق كل من طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات المشتركة.

كما أن تشتت قيم الخطأ المعياري للمفردات الملائمة وفق طريقتي التعادل قد جاء متساوياً في الحالتين؛ حيث بلغت قيمة الانحراف الربيعي (٠.١٠) منف، مع تساوي قيمتي الوسيط (٠.٧) منف.

نتائج الإجابة عن السؤال الثامن وينص على:

"هل يختلف مدى تقديرات قدرات الأفراد عينة التدريج باختلاف طريقة التعادل؟"

لدراسة تساوي مدى تقديرات قدرة عينة التدريج باختلاف طريقة التعادل استخدم اختبار Moses لدلالة الفروق بين مدى توزيعين من خلال تحديد القيم المتطرفة لتوزيع التقديرات وفق كل طريقة من طرق التعادل.

جدول(٣٩)

نتائج اختبار **Moses** لتقديرات قدرات الأفراد وفق طريقتي التعادل.

	قدرات الأفراد بالمنف
Observed Control Group Span	١٩٥٦
Sig. (1-tailed)	٠,١٩٢
Trimmed Control Group Span	١٧٥٨
Sig. (1-tailed)	٠,٣١٩
Outliers Trimmed from each End	٤٨

يتضح من جدول (٣٩) السابق أنه لا توجد فروق دالة إحصائيا بين توزيع تقديرات قدرات الأفراد بين طريقتي التعادل المتزامن والتعادل بتثبيت بارامترات المفردات في توزيع القيم المتطرفة؛ مما يشير إلى تساوي مدى توزيع تقديرات القدرة وفق طريقتي التعادل.

خلاصة وتوصيات:

- أشارت النتائج إلى تكافؤ طريقتي التعادل من الناحية الإحصائية في تقدير إحصاءات الأفراد والمفردات، وقد جاءت طريقة التعادل بتثبيت بارامترات الأفراد والمفردات أكثر حفاظا على الأفراد والمفردات من حيث انخفاض أعداد المفردات والأفراد المحذوفين، كما تفوقت على طريقة التدريج المتزامن في انخفاض قيم الأخطاء المعيارية المصاحبة لتقديرات قدرات الأفراد. لذلك توصي الدراسة من الناحية العملية عندما يكون هدف القياس المحافظة على أكبر قدر من المفردات خاصة في الاختبارات ذات الدرجة العالية من الأهمية؛ حيث تكون هناك درجة عالية من السرية، فضلا عن كون بنائها مكلف جدا فيكون من الأنسب استخدام طريقة التعادل بتثبيت بارامترات المفردات المشتركة، وكذلك في الحالات التي يتم فيها ربط مفردات جديدة ببنك قائم بالفعل. أما التدريج المتزامن فيكون من الناحية العملية أنسب عندما يكون هناك عدد كبير من الصور الاختبارية التي يتم تدريجها للمرة الأولى لإنشاء بنك أسئلة خاصة عندما يكون عدد الصور الاختبارية كبيراً؛ حيث يمكن ضم مفردات الصور الاختبارية واستجابات الأفراد عليها في ملف واحد، وإجراء عملية التدريج المتزامن مرة واحدة بدلا من أن يتم تدريج الصور بشكل مستقل، ثم يتم ضمها بعد ذلك على ميزان تدرج واحد. وبعد انتهاء تدريجها فإنه في المراحل التالية التي يتم فيها إثراء البنك يمكن استخدام طريقة تثبيت بارامترات المفردات مع مراعاة استخدام عدد كاف من المفردات المشتركة؛ حيث يقوم عليها

بصفة أساسية إنشاء التدريج. حيث يساعد ذلك في توفير الجهد كما يساعد في تجنب مشكلة أي الصور الاختبارية تستخدم كإطار مرجعي للبنك.

- نظرا لأن تصميم المجموعات غير المتكافئة ذا المفردات المشتركة يقوم بصفة أساسية على تقديرات المفردات المشتركة في بناء التدريج المشترك؛ لذلك كان من المهم العناية بانتقاء المفردات المشتركة من حيث جودة خصائصها السيكومترية وخصائصها الفنية، مع كفاية عددها بحيث تكون صورة مصغرة من الاختبار الكلي.
- للتغلب على مشكلة كشف عدد كبير من المفردات الاختبارية للمختبرين عند تطبيق الصور الاختبارية فإنه يمكن استخدام تصميم المجموعات غير المتكافئة ذى المفردات المشتركة ذي الرابط الداخلي.

دراسات مقترحة:

في ضوء ما أسفرت عنه الدراسة من نتائج وتوصيات تقترح الباحثة ما يلي:

- دراسة دقة التعادل باستخدام طريقة التعادل المتزامن وطريقة التعادل بتثبيت البارامترات في إطار تصميم الأفراد المشتركين.
- دراسة أثر اختلاف ثبات الصور الاختبارية وطول الاختبار وعدد المفردات المشتركة على دقة معادلة درجات الاختبارات باستخدام طريقتي التعادل المتزامن والتعادل باستخدام طريقة تثبيت بارمترات المفردات المشتركة.
- دراسة أثر انتهاك بعض افتراضات نظرية الاستجابة للمفردة على دقة التعادل باستخدام طريقتي التعادل المتزامن وطريقة تثبيت بارامترات المفردات المشتركة.
- دراسة تأثير طريقتي التعادل (المتزامن / تثبيت بارامترات المفردات المشتركة،) على دقة تدريج الاختبارات المعرفية الأخرى مثل اختبارات الاستعداد واختبارات الذكاء.

المراجع:

أروى الحواري وأحمد عودة. (٢٠٠٨). الخصائص السيكومترية لصور مختارة من اختبارات الرخصة الدولية لقيادة الحاسوب فى الأردن، ومعادلة درجاتها. ***المجلة الأردنية فى العلوم التربوية***،٤(٤)، ٢٩٧-٣١٩.

إسماعيل حسن الوليلى. (٢٠٠٥). تكافؤ درجات الاختبارات فى ضوء نظريتى القياس الكلاسيكية والحديثة (دراسة سيكومترية مقارنة). ***مجلة كلية التربية جامعة بنها***، ١٥(٦٣)، ٩٩-١٤٩.

أمينة محمد كاظم. (١٩٨٨). دراسة نظرية نقدية حول القياس الموضوعي للسلوك (نموذج راش). الكويت: سلسلة الكتب المختصة، مؤسسة الكويت للتقدم العلمي.

أمينة محمد كاظم. (٢٠٠٠). اتجاهات معاصرة في بناء بنوك الأسئلة. في: الأسس التربوية لإعداد المعلم الجامعي. القاهرة: جامعة عين شمس، ط٣، ٣٢١ - ٣٤٢.

إيهاب محمد نجيب. (٢٠١٢). معادلة الاختبارات وعلاقتها ببعض المتغيرات السيكومترية دراسة محاكاة وتطبيقية. *رسالة دكتوراة*، غير منشورة، كلية التربية جامعة عين شمس.

راشد حماد الدوسرى. (٢٠٠١). معادلة الاختبارات مفهومها، وطرقها، ومشكلات تطبيقها. ***مجلة العلوم التربوية والنفسية***، ٤، ١٠٧-١٤١.

زياد عبد الحسيب الخولى. (٢٠٠٦). أثر طرق معادلة درجات الاختبار فى تدريج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئى. رسالة دكتوراة، غير منشورة، معهد الدراسات التربوية، جامعة القاهرة.

شادية عبد العزيز مهتدى. (٢٠١٧). التكامل بين التحليل الإحصائي والبياني في انتقاء وتدريج مفردات الاختيار من متعدد باستخدام نظرية الاستجابة للمفردة. ***المجلة المصرية للدراسات النفسية***. ٢٧(٩٧)، ٣٢٤ - ٣٨٧.

صلاح الدين محمود علام. (٢٠٠٥). ***نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها فى القياس النفسى والتربوى***. القاهرة، دار الفكر العربى.

مايسة فاضل أبو مسلم. (٢٠١٠). معادلة صورتى اختبار "تونى" للذكاء غير اللفظى باستخدام طرق مختلفة للمعادلة فى ضوء بعض المتغيرات المؤثرة على نتائجها. ***المجلة المصرية للدراسات النفسية***، ٢٠(٦٦)، ٣٧١-٤١١.

محمد أحمد عبد الوهاب.(٢٠١٥). أثر اختلاف معامل ثبات صورتي اختبار في دقة معادلة درجاتها باستخدام الطرق القائمة على نظرية الاستجابة للمفردة. ***مجلة كلية التربية***، جامعة الإسكندرية. ٢٥(٣)، ٢٣-٥٤.

محمد منصور الشافعى. (١٩٩٦). أثر طرق معادلة درجات الاختبار وضوابط اختبار العينة على تدريج بنود بنك الأسئلة باستخدام (نموذج راش). رسالة دكتوراه، غير منشورة، كلية التربية، جامعة المنصورة.

نضال كمال الشريفين. (٢٠٠٣). مدى تحقق معايير الفاعلية فى معادلة اختبارين أحدهما ثنائى التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة فى القياس. رسالة ماجستير غير منشورة، كلية العلوم التربوية، جامعة عمان العربية للدراسات العليا.

يوسف عبد العاطي المحروق.(٢٠١٦). أثر حجم العينة وطول الاختبار في دقة معادلة درجات الاختبارات متعددة الحدود باستخدام طريقة كيرنيل. ***مجلة العلوم التربوية والنفسية***. ١٧(٣)، ٢٠١-٢٢٨.

Andersson,B& Wiberg.M.(2017). Item Response Theory Observed-Scor Kerenl Equating. ***Psychometrika***, 82(1), 48–66. March 2017.do 10.1007/s11336-016-9528-7.

Arai,S,& Mayekawa,S.(2011).A Comparison of Equating Methods an Linking Designs for Developing an Item Pool Under Item Respons Theory. ***Behaviormetrika***.38(1),1- 16.

Ban, J, HanSon, B, Wang, T, Yi, Q, & Harris, D. (2001). A Comparativ Study of on- line pretest item Calibration/Scaling Methods i Computerized Adaptive Testing. ***Journal of Educationa Measurement,*** 38, 191-212.

Briggs, D. C., & Weeks, J. P. (2009). The impact of vertical scaling decision on growth interpretations. ***Educational Measurement: Issues an Practice*** Winter, 28(4), 3–14.

Cohen, A. S., & Kim, S. H. (1998). An investigation of linking method under the graded response model. ***Applied Psychologica Measurement***, 22(2), 116-130.

Cook, L. L. (2007). ***Practical problems in equating test scores: practitioner's perspective.*** In N. J. Dorans, M. Pommerich, & P. W Holland (Eds.), Linking and aligning scores and scales (2nd ed pp.73-87). New York: Springer.

Dorans, N. & Moses, T. and Eignor, D. (2010). ***Principles and practices o test score Equating.*** ETS, Princeton, New Jersey.

Gonz´alez J (2013). "Book Review: Statistical Models for Test Equating Scaling, and Linking." ***Applied Psychological Measurement***, 37(4 336–339.

Gruijter, D. & Kamp, L. (2008). ***Statistical Test Theory for the Behaviora Sciences***. London, Taylor.

Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). ***Fundamental of item response theory.*** USA: Sage.

Hanson, B. A., & Béguin, A. A. (2002). Obtaining a common scale for iter response theory item parameters using separate versus concurrer estimation in the common-item equating design. ***Applie Psychological Measurement***, 26(1), 3-24.

Harris, D. J. (2007). ***Practical issues in vertical scaling.*** In N. J. Dorans, M Pommerich, & P. W. Holland (Eds.), Linking and aligning scores an scales (pp. 233–251). New York: Springer.

Hills, J. R., Subhiyah, R. G. & Hirsch, T. M. (1988). Equating minimum competency tests: Comparisons of methods. ***Journal of Educationa Measurement***, 25(3), 221-231. Retrieve from:http://www.jstor.org/stable/1434501 on November, 12, 2013.

Hills, J. R., Subhiyah, R. G., & Hirsch, T. M. (1988). Equating minimur competency tests: Comparison of methods. ***Journal of Educationa Measurement***, 25, 221-231.

Ito, K., Sykes, R. C., & Yao, L. (2008). Concurrent and separate grade groups linking procedures for vertical scaling. ***Applied Measuremen in Education***, 21(3), 187–206.

Jodoin, M. G. (2003). Measurement efficiency of innovative item formats i computer-based testing. ***Journal of Educational Measuremen.*** 40(1), 1-15. doi: 10.1111/j.1745-3984.2003.tb01093.x

Jodoin, M. G., Keller, L. A., & Swaminathan, H. (2003). A comparison c linear, fixed common item, and concurrent parameter estimatio equating procedures in capturing academic growth. ***The Journal o Experimental Education***, 71, 229–250.

Keller, L, & Keller, R. (2015). The Effect of Changing Content on IR Scaling Methods. ***Applied Measurement in Education***. Vol. 2 Issu2.

Keller.R.(2007).A Comparisons of item response theory true score equatin and item response theory – based local equating. Submitted to th Graduate School of the University of Massachusetts Amherst i partial fulfillment of the requirements for the degree of Doctor c Education.

Kesamang M. (2017) The comparison of item and person item respons theory (IRT) parameter estimates for the anchoritems and common persons designs .***Herald J. Edu. Gen. Stud.*** 4(1), pp. 001 – 012

Kilmen,S.& Dermirtasli,N.(2012).Comparison of Test Equating Method Based on Item Response Theory According to the Sample Size an Ability Distribution. ***Procedia –Social and Behavioral Science:*** 46,130-134.

Kim, S & Cohen, A. (1996). A comparison of Linking and Concurrer Calibration under Item Response Theory. Paper Presented at th Annual ***Meeting of the American Educational Researc Association***. (New York, N4, April 8-12.

Kim, S. (2006). A comparative study of IRT fixed parameter calibratio methods.
Journal of Educational Measurement, 43(4), 355-381.

Kolen MJ, Brennan RL (2004). ***Test Equating, Scaling, and Linking Methods and Practices***. Springer-Verlag.

Li, Y. H., Griffith, W. D., & Tam, H. P. (1997). Equating multiple tests via a IRT linking design: Utilizing a single set of anchor items with fixe common item parameters during the calibration process. ***Pape presented at the meeting of the Psychometric Society,*** Knoxville TN.

Linacre, J. M. (2003). WINSTEPS [Computer Program]. Chicago: MES/ Press. Loyd, B. H., & Hoover, H. D. (1980). Vertical equating usin the Rasch model. ***Journal of Educational Measurement***, 17, 179 193.

Linacre, J. M. (2011). Winsteps (Version 3.72.3) [Computer software] Chicago: Winsteps.com.

Livingston, S. A. (2004). ***Equating test scores (without IRT).*** Princetor Educational Testing Service.

Lord, F. M., & Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score "equatings." ***Applied Psychologica Measurement***, 8, 453-461.

Loyd, B. H. & Hoover, H. D. (1980). Vertical equating using the Rasc model. ***Journal of Educational Measurement***, 17(3), 179-19: Retrieved from http://www.jstor.org/stable/1434833 on Novembe 12, 2013.

Marco, G. L. (1977). Item characteristic curve solutions to three intractabl testing problems. ***Journal of Educational Measurement***, 14, 139 160.

Mislevy, R. J. (1992). ***Linking educational assessments: Concepts, issue methods, and prospects***. Princeton, NJ: Policy Information Center.

Paek, I., & Young, M. J. (2005). Investigation of student growth recovery in fixed-item linking procedure with a fixed-person prior distributio for mixed-format test data. ***Applied Measurement in Education***, 1٤ 199–215.

Pang, X; Madera, E; Radwan, N; Zhang, S. (2010). A Comparison of Fou Test Equating Methods. Report Prepared For the Education Qualit and Accountalility. Office (EQAO). April. 2010

Patz, R. J., & Yao, L. (2007). Methods and models for vertical scaling. In N J. Dorans, M. Pommerich, & P. W. Holland (Eds.), Linking an aligning scores and scales (pp. 253–272). New York: Springer.

Petersen, N. S. , Cook, L. L., & Stocking, M. L. (1983). IRT versu conventional equating methods: A comparative study of scal stability. ***Journal of Educational Statistics***, 8, 137-156.

Spence, P. D. (1996). The effect of multidimensionality on unidimensiona equating with item response theory. Unpublished doctora dissertation, University of Florida America.

Stocking, M. L. & Lord, F. M. (1983). Developing a common metric in iter response theory. ***Applied Psychological Measurement***, 7(2), 201 210. doi: 10.1177/014662168300700208

Taherbhai,H.& Seo.D.(2007). Comparing concurrent versus fixed paramete equating with common items: using the dichotomous and partia credit models in a mixed-item format test. ***Journal of Applie Measurement***. 8(1).84: 96.

Thissen, D. (1991). Multilog user's guide: Multiple categorical item analysi and test scoring using item response theory [Computer program] Chicago: Scientific Software International.

Tian, F. (2011). A comparison of equating/linking using the Stocking-Lor method and concurrent calibration with mixed-format tests in th non-equivalent groups common-item design under IRT. Unpublishe doctoral dissertation, Boston College.

von Davier AA, Holland PW, Thayer DT (2004). ***The Kernel Method of Tes Equating***. Springer-Verlag.

Yen, W. M. (2007). ***Vertical scaling and no child left behind***. In N. J Dorans, M. Pommerich, & P. W. Holland (Eds.), Linking an aligning scores and scales (pp. 273–282). New York: Springer.

Zhao,J & Hambleton, R. (2017). Practical Consequences of Item Respons Theory Model Shift in the Context of Test Equating with Mixed Format Test Data. ***Frontiers in psychology.*** Vol. 8, Article, 484. do 10338911 psyg.2017.00484.

Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (1996). ***BILOG MG: Multiple-group analysis and test maintenance for binary item*** [Computer software]. Chicago: Scientific Software International.

The Effect of Using (Concurrent / and Fixed Common Items Parameters Equating Methods).in Estimating Items and Persons Statistics According to Non Equivalent Groups Design

Dr. Ekram Hamza El Sayed Sahwan[5]

Abstract

The study aimed to compare between the concurrent calibration and fixing item parameters equating methods in estimating item and person statistics According to nonequivalent groups design. The tool of the study consists of 3 equivalent test forms developed on a unit of the science book (material and its composition) from grade 7. The total number of items was 75 MCQs that were distributed among three test forms (30, 31, and 34 items) to be equivalent in terms of content, cognitive levels and difficulty level with 10 common items. The test forms were administered on 1198 students from grade 7. The data was analyzed twice using WINSTEPS - once by each equating methods, and item and person statistics were extracted from each analysis. Comparisons were made between the results of each method to answer the study questions.

The most important results were:

First: Regarding the effect of the equating methods on item statistics:

- There are no statistically significant differences in the percentages of the unfit items to the basis of objective measurement between the two equating methods. In the fixed common items parameters equating method, seven items were deleted from the three test forms. While in the concurrent equating method, nine items were deleted.
- There is large significant consistency between the two methods in classification of items as the coefficients of Phi and Kaba are large and statistically significant.
- There is large consistency in estimating the items difficulties between the two methods where the correlation coefficient is very large and statistically significant, and the coefficient of difference between the two methods is equal.

[5] Researcher in the Research Department in the National Center for Examinations and Educational Evaluation (NCEEE).

- The standard errors (SE) estimates of item estimation are statistically equivalent between the two methods where the results shows that there is no statistically significant difference of the mean SE of item parameters between the two methods. Moreover, the results of the correlation between the SE estimates of items parameters in both methods indicate a significant consistency between the distribution of SE of items difficulty according to the two equating methods.
- The items reliability indices, as represented in the Marginal Reliability, the Separation Coefficient, and the Information Function, are high according to the two equating methods. Moreover, the mean of the items SE estimates are low for both equating methods.
- There are no statistically significant differences in the distribution of items difficulties estimates on both sides of the difficulties calibration scale.

Secondly: Regarding the effect of the equating methods on person statistics:

- There are no statistically significant differences in the percentage of the unfit persons to the basis of objective measurement between the two equating methods. In the fixed common items parameters equating method, 210 persons were deleted from the three test forms. While in the concurrent equating method, 226 persons were deleted.
- There is large significant consistency between the two methods in classification of persons as the coefficients of Phi and Kaba are large and statistically significant.
- There is large consistency in estimating the persons' abilities between the two methods where the correlation coefficient is very large and statistically significant, and the coefficient of difference between the two methods is equal.
- There are statistically significant differences in the SE estimates of person abilities SE estimates are larger in the concurrent calibration than the fixing item parameters methods. Yet, the results of the correlation between the SE estimates of person abilities in both methods indicated a large consistency between the distribution of SE of person abilities according to the two calibration methods.
- The person reliability indices, as represented in the Marginal Reliability, the Separation Coefficient, and the Information Function, is high according to the two equating methods. Moreover, the mean of the persons' SE estimates are low for both equating methods.

- There are no statistically significant differences in the distribution of person's abilities estimates on both sides of the abilities calibration scale.