# جـامـعة قاصدي مربـــاح ورقلـــة

كليـة العلوم الانسانيـة والاجتماعيــة
قســم علــم النفس وعلــوم التربيـة


العنوان

# اسْتِخْدَام نَمُوذَج رَاش فِي بِنَاءِ اخْتِبَار تحصيلي مَحْكِيّ الْمَرْجِع لمُقرَّر الْإِحْصَاء لِطَلَبَةِ السَنَة أَوْلَى عُلُوم اِجْتِمَاعِيَّة بجامِعَة زيَّان عاشور – الجلفة -

أطروحَةٌ مُقَدَّمةٌ لاستكمال مُتَطَلَّباتِ نَيْل شَهادَةِ الدُّكْتوراه الطُّورِ الثَّالِث
تَخَصُص القِياسُ النَّفْسِيّ وَالتَقْوِيمُ التَرْبَوي

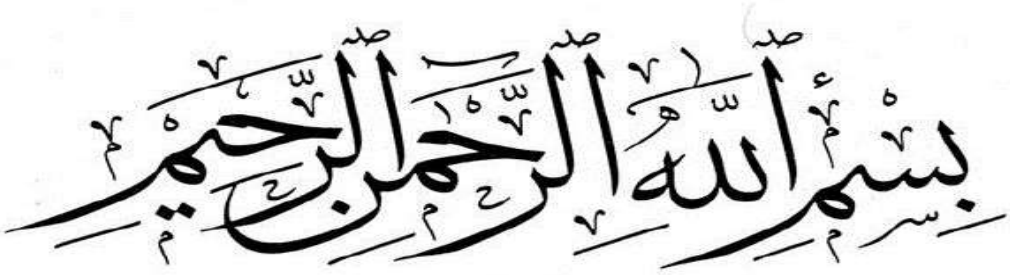

إعداد الطالب
هريمك نبيل

إشراف الدكتورة
باعمر الزهرة


لجنة المناقشة

| الاسم واللقب | الرتبة | الجامعة | الصفة |
|---|---|---|---|
| منصور بن زاهي | أستاذ التعليم العالي | قاصدي مرباح ورقلة | رئيساً |
| الزهرة باعمر | أستاذ محاضر أ | قاصدي مرباح ورقلة | مشرفاً ومقرراً |
| بلخير طبشي | أستاذ التعليم العالي | قاصدي مرباح ورقلة | مناقشاً |
| أمال دربال | أستاذ محاضر أ | جامعة بشار | مناقشاً |
| فارس إسعادي | أستاذ التعليم العالي | جامعة الوادي | مناقشاً |
| رشيد زياد | أستاذ محاضر ب | جامعة تبسة | مدعو |

الموسم الجامعي 2022-2023




**قال الله تعالى**

**﴿وَلَقَدْ آتَيْنَا دَاوُودَ وَسُلَيْمَانَ عِلْمًا وَقَالَا الْحَمْدُ لِلَّهِ الَّذِي فَضَّلَنَا عَلَى كَثِيرٍ مِنْ عِبَادِهِ الْمُؤْمِنِينَ﴾**

سورة النمل الآية (15)

# شكر وتقدير

الحمد لله الذي أيدني بعونه وأصلي وأسلم على أفضل خلقه وبعد

الشكر والحمد لله عز وجل الذي فضلنا على كثير من خلقه ويسر لنا أمورنا وسخر لنا من عباده من كان عوناً لنا

ثم أتوجه بخالص الشكر والعرفان والتقدير للأستاذة المشرفة الدكتورة باعمر الزهرة لما أبدته من جهود علمية وأراء سديدة أغنت الدراسة لتظهر بالصورة الحالية كما أنها كانت خير موجه ومرشد طوال هذا العمل

كما أتوجه بخالص الشكر والثناء لكل الأساتذة الذين ساعدوني من قريب أو بعيد واخص بالذكر الدكتور زياد رشيد وأترحم على الأستاذ غريب مختار طيب الله ثراه الذي كان له فضل كبير في انجاز هذا العمل ولن أنسى أن أتوجه بالشكر إلى الأَساتذة المحكمين على أدوات الدراسة، والشكر موصول للجنة المناقشة التي تفضلت بقبول مناقشة هذا العمل.

والحقيقة لا يسع حروفي إلا أن تمتزج لتكوِّن كلمات شكر

لكل من ساعدني من قريب أو من بعيد

واخيراً أشكر أمي واخوتي وزوجتي وبناتي لدعمهم

وصبرهم على انشغالي عنهم.

هريمك نبيل

**ملخص الدراسة**

هدفت هذه الدراسة إلى بناء اختبار محكي المرجع وفق نموذج راش لمقرر الإحصاء الوصفي لطلاب السنة الأولى علوم اجتماعية جذع مشترك بجامعة زيان عاشور بالجلفة.

وتكوَّن الاختبار بصورته الأولية من (20) مفردة اختبارية من نوع الاختيار من متعدد بواقع أربع بدائل إحداهن تُمثل الاجابة الصحيحة.

وتكوَّنت عينة الدراسة من (502) طالباً وطالبة من طلاب السنة الأولى علوم اجتماعية بجامعة الجلفة، والذين درسوا مقرر الإحصاء الوصفي خلال السداسي الأول من الموسم الجامعي 2021/2020.

وتمَّ التحقق من صدق درجات الاختبار(الصدق الوصفي وصدق الوظيفي وصدق البناء) وكذلك التحقق من معامل ثبات درجات الاختبار، وفق معادلة كيودر ريتشاردسون، الذي بلغ(0.89) وكذا الثبات لمعامل ليفنجستون الذي بلغ(0.95) بالإضافة إلى تحديد درجة القطع المناسبة لتحديد الكفاية إحصائية وفقاً لطريقة أنجوف، والتي بلغت(57.59%) تقريباً الدرجة(12)، والتي من خلالها يتم تصنيف الطلبة إلى متقنين وغير متقنين لهذه الكفاية، والتعرف على مدى تمكن طلاب العلوم الاجتماعية من الكفايات الإحصائية التي تضمَّنها الاختبار الذي تمَّ بناؤه، وقد اعتمد الباحث في هذه الدراسة على المنهج الوصفي التحليلي.

وللإجابة عن أسئلة الدراسة استخدم الباحث البرنامجين(Spss.26) الحزمة الإحصائية للعلوم الاجتماعية وبرنامج(Winsteps) الخاص بتحليل البيانات وفق نموذج راش.

وقد نتج عن تحليل البيانات، حذف مفردة واحدة فقط كانت خارج حدود المطابقة الداخلية والخارجية، كما تمَّ حذف (18) فرداً كانوا غير مطابقين لنموذج راش.

وبينت النتائج أنَّ تدريج الاختبار محكي المرجع لمقرر الاحصاء الوصفي وفق نموذج راش يحقق افتراضات النموذج وهي (أحادية البعد، الاستقلال الموضعي، خاصية المنحنى المميز للمفردة التحرر من عامل السرعة).

كما تمَّ حساب معامل الثبات المتعلق بالأفراد والثبات المتعلق بمفردات الاختبار من خلال عامل الفصل، حيث بلغ معامل الفصل المتعلق بالأفراد (2.53) ومعامل الثبات للأفراد (0.87) في حين وصل

معامل الفصل المتعلق بالمفردات (12.57) ومعامل الثبات للمفردات (0.99) وهما معاملان مرتفعان.

وبينت النتائج أنَّ قيمة دالة المعلومات للاختبار تقدم أكبر كمية معلومات عند مستويات القدرة المتوسطة التي كانت قيمتها (3.32) وخطأ معياري للقياس (0.55) المقابلة لمنتصف الدرجة الحقيقة (09) والتي تقابل صفر تدريج. ويقدم أقل كمية من المعلومات عند مستويات القدرة العالية والمتدنية.

**الكلمات المفتاحية:** اختبار تحصيلي، اختبار محكي المرجع، نظرية الاستجابة للمفردة، نموذج راش.

***Abstract***

*The present study aimed at using the item response theory for constructing a criterion-referenced test in descriptive statistics for 1st year students (social sciences), a common trunk, in the Djelfa University according to Research Model.*

*The test consisted of (20) MCQ with four alternative items. one of which represents the correct answer.*

*The researcher used the descriptive analytical method because it is the most appropriate for this study.*

*The study sample were (502) students in the Zian Achour University in Djelfa.*

*The most significant results were shown in verifying the assumptions of item response theory in this study, in addition to having 19 items that were inconsistent with the Rasch Model while 1 item were excluded as they found to be inconsistent with the Rasch Model. Furthermore, 18 Students were excluded from the sample because they were found inconsistent with the model. The estimates of the difficulty and discrimination coefficients were also acceptable.*

*The analysis also produced a good test that has different forms of validity in addition to (0.99) reliability coefficient of items and (0.87) reliability coefficient of persons.*

*The results also revealed that the maximum value of information was achieved at (e=b) which agrees with Rasch Model and that the test offers the maximum information at medium ability levels while the minimum information was at both lowest and highest ability levels.*

*The results showed that the value of the information function of the test provides the largest amount of information at the average power levels whose value was (3.32) and the standard error of measurement (0.55) corresponding to the middle of the true degree (09), which corresponds to zero scale. It provides the least amount of information at high and low power levels.*

*Keywords: Achievement test, Criterion Referenced Test, Item Response Theory, Rasch model.*

## قائمة المحتويات

**الموضوع** **الصفحة**


شكر وتقدير ..... أ
الملخص بالعربية ..... ب
الملخص بالإنجليزية ..... د
قائمة المحتويات ..... هـ
قائمة الجداول ..... ح
قائمة الأشكال ..... ي

**الباب الأول**
**الفصل الأول**

مقدمة ومشكلة الدراسة ..... 01
**أولاً.** أهداف الدراسة ..... 08
**ثانياً.** أهمية الدراسة ..... 09
**ثالثاً.** حدود الدراسة ..... 09
**رابعاً.** المفاهيم الاجرائية للدارسة ..... 10

**الفصل الثاني: الدراسات السابقة**

**I.** الدراسات السابقة ..... 12
**أولاً:** الدراسات التي استخدمت بناء الاختبارات محكية المرجع ..... 12
**ثانياً:** الدراسات التي استخدمت الاختبارات التحصيلية في مقرر الإحصاء ..... 18
**ثالثاً:** الدراسات التي استخدمت نموذج راش ..... 21
**II.** التعقيب عن الدراسات السابقة ..... 26
**III.** موقع الدراسة الحالية من الدراسات السابقة ..... 28

**الفصل الثالث: الإطار النظري للدراسة**

**تمهيد**

**المحور الأول: الاختبارات التحصيلية محكية المرجع**

**1.** الاختبارات التحصيلية ..... 32
**1.1.** تعريف الاختبار ..... 32
**2.1.** تعريف الاختبار التحصيلي ..... 32
**3.1.** الأسس النظرية للاختبارات التحصيلية ..... 33
**4.1.** مبادئ الاختبارات التحصيلية ..... 34
**5.1.** افتراضات أساسية يرتكز عليها قياس التحصيل ..... 35
**2.** الاختبارات محكية المرجع ..... 35
**1.2.** نشأت الاختبارات محكية المرجع ..... 35
**2.2.** تعريف الاختبارات محكية المرجع ..... 36
**3.2.** أوجه الاختلاف بين الاختبارات محكية المرجع ومعيارية المرجع ..... 37
**4.2.** استخدامات الاختبارات محكية المرجع ..... 38
**5.2.** تصنيف الاختبارات محكية المرجع ..... 38

6.2. أهمية الاختبارات محكية المرجع ..... 39
7.2. عيوب استخدام الاختبارات محكية المرجع..... 40
3. خطوات بناء الاختبارات محكية المرجع ..... 40
4. الخصائص السيكومترية للاختبار محكي المرجع ولفقراته ..... 48
1.4. صدق الاختبارات محكية المرجع ..... 48
2.4. ثبات الاختبارات محكية المرجع..... 51
**المحور الثاني: النظرية الكلاسيكية في القياس**
1.الافتراضات الأساسية للنظرية الكلاسيكية ..... 64
2.المسلمات التي تستند إليها نظرية القياس الكلاسيكية ..... 67
3.أساليب القياس الكلاسيكي ..... 68
4. الخصائص السيكومترية للاختبار وفق النظرية الكلاسيكية..... 69
1.4.مفهوم الصدق ..... 69
1.1.4. معاني الصدق ..... 69
2.1.4. طرق حساب الصدق ..... 70
2.4. مفهوم الثبات وطرق حسابه ..... 75
3.4. العوامل التي تؤثر على ثبات الاختبار..... 81
5. العلاقة بين الصدق والثبات ..... 84
6. النظرية الكلاسيكية والاختبارات ..... 84
7. الانتقادات التي وجهت إلى النظرية الكلاسيكية ..... 84
8. أوجه قصور النظرية الكلاسيكية ..... 86
**المحور الثالث: نظرية القياس الحديثة**
1. الفكرة الأساسية لنظرية القياس الحديثة ..... 90
2.مفهوم نظرية الاستجابة للمفردة ..... 90
3.مسلمات نظرية الاستجابة للمفردة ..... 91
4. افتراضات نظرية الاستجابة للمفردة ..... 92
5.الخصائص السيكومترية وفق النظرية الحديث. ..... 94
6.دالة معلومات الفقرة والاختبار في نماذج IRT ..... 95
7.السمات الكامنة والمنحنيات المميزة للمفردة ..... 97
8.مميزات نماذج نظرية الاستجابة للمفردة ..... 98
9.اجراءات استعمال نظرية الاستجابة للمفردة ..... 99
10. نماذج نظرية الاستجابة للمفردة ..... 102
11.سبب عدم انتشار نظرية الاستجابة للمفردة ..... 109
**المحور الرابع: نموذج راش**
1.الخصائص المميزة لنموذج راش ..... 110
2.الأساس النظري لنموذج راش ..... 110
3.الأساس الرياضي لنموذج راش ..... 113
4.الموضوعية في نموذج راش ..... 114
5.افتراضات نموذج راش ..... 118

6.طرق تقدير نموذج راش ..... 120
7.تدريج فقرات الاختبار ..... 126
8.مفهوم الصدق والثبات وفق نموذج راش ..... 127
9.التقديرات البارامترية لنموذج راش ..... 128
10.وحدة قياس قدرة الفرد وصعوبة المفردة ..... 131
11.استخدامات نموذج راش ..... 135
12.خصائص نموذج راش ..... 136
13.عيوب نموذج راش ..... 136
خلاصة الفصل
**الباب الثاني: الجانب الميداني**
**الفصل الرابع: اجراءات الدراسة وأدواتها**
تمهيد
1.منهج الدراسة ..... 139
2.مجتمع الدراسة ..... 139
3.أداة الدراسة ..... 140
4.مراحل بناء أداة الدراسة ..... 140
5.أساليب المعالجة الإحصائية ..... 159
**الفصل الخامس: تحليل ومناقشة نتائج الدراسة**
تمهيد
**أولاً. تحليل نتائج تساؤلات الدراسة**
1.عرض وتحليل التساؤل الأول ..... 162
2.عرض وتحليل التساؤل الثاني ..... 168
3.عرض وتحليل التساؤل الثالث ..... 178
4.عرض وتحليل التساؤل الرابع ..... 182
5.عرض وتحليل التساؤل الخامس ..... 186
**ثانيا. مناقشة نتائج الدراسة**
1. مناقشة نتائج التساؤل الأول ..... 190
2. مناقشة نتائج التساؤل الثاني ..... 193
3. مناقشة نتائج التساؤل الثالث ..... 195
4. مناقشة نتائج التساؤل الرابع ..... 196
5. نتائج التساؤل الخامس ..... 198
الخلاصة والتوصيات
قائمة المراجع
الملاحق

# قائمة الجداول


| الرقم | عنوان الجدول | الصفحة |
|---|---|---|
| 01 | قيمة معامل كارفر | 57 |
| 02 | الجدول الثنائي لحساب معامل التوافق المرجعي | 61 |
| 03 | الحد الأدنى لمعامل التوافق المرجعي | 62 |
| 04 | عدد الإجابات الصحيحة والخاطئة للمتقنين وغير المتقنين | 63 |
| 05 | عدد الأهداف السلوكية لكل كفاية | 142 |
| 06 | الأهداف السلوكية الإجرائية للكفاية الأولى | 143 |
| 07 | الأهداف السلوكية الإجرائية للكفاية الثانية | 143 |
| 08 | الأهداف السلوكية الإجرائية للكفاية الثالثة | 143 |
| 09 | الأهداف السلوكية الإجرائية للكفاية الرابعة | 144 |
| 10 | الأهداف السلوكية الإجرائية للكفاية الخامسة | 144 |
| 11 | الأهداف السلوكية الإجرائية للكفاية السادسة | 145 |
| 12 | نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الأولى | 145 |
| 13 | نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الثانية | 146 |
| 14 | نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الثالثة | 146 |
| 15 | نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الرابعة | 146 |
| 16 | نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الخامسة | 147 |
| 17 | نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية السادسة | 147 |
| 18 | نتائج التحكيم للأهداف السلوكية | 147 |
| 19 | الوزن النسبي للموضوعات | 148 |
| 20 | الوزن النسبي للأهداف | 149 |
| 21 | مؤشرات الصعوبة لفقرات الاختبار للعينة الاستطلاعية | 152 |
| 22 | معاملات تمييز فقرات الاختبار | 152 |
| 23 | النتائج المتعلقة بمعامل الحساسية | 153 |
| 24 | قيم فاعلية البدائل في الاختبار التحصيلي | 154 |
| 25 | الأوساط الحسابية والانحرافات المعيارية لتقدير المحكمين | 155 |
| 26 | الصدق المحكي بدلالة المجوعة الطرفية | 156 |
| 27 | درجات الاتساق الداخلي للاختبار | 156 |
| 28 | متوسط تقديرات المحكمين لمفردات الاختبار باستخدام طريقة أنجوف | 159 |
| 29 | نتائج التحليل العاملي للمكونات الأساسية للبواقي (PCAR) | 163 |

| | | |
|---|---|---|
| 30 | قيم مؤشرات التمييز لاختبار الإحصاء | 164 |
| 31 | قيم مؤشرات التمييز بمعامل الارتباط الثنائي المتسلسل | 166 |
| 32 | قيم المؤشر الإحصائي(Q3) من خلال قيم ارتباط البواقي المعيارية | 166 |
| 33 | مؤشرات التخمين لمفردات اختبار الإحصاء | 167 |
| 34 | ملخص نتائج التحليل لقدرات الأفراد | 169 |
| 35 | ملخص نتائج التحليل لمعالم الصعوبة لمفردات اختبار الإحصاء | 170 |
| 36 | قيم إحصائيات المطابقة الداخلية والخارجية، ومتوسطات المربعات الموزونة للمفردات في التدريج الأولي للاختبار | 171 |
| 37 | ملخص نتائج التحليل للقيم المتحررة من قدرات الأفراد بعد استبعاد الأفراد والمفردات غير المطابقة | 172 |
| 38 | ملخص نتائج التحليل المتحرر من صعوبة المفردات بعد التدريج | 173 |
| 39 | صعوبة المفردات وأخطاؤها المعيارية مرتبة تنازلياً وفق صعوبة المفردات | 174 |
| 40 | تقديرات قدرة الفرد المقابلة لكل درجة كلية محتملة على الاختبار وانحرافاتها المعيارية وكذا كمية المعلومات التي يقدمها اختبار الإحصاء | 179 |
| 41 | تقدير صعوبة المفردات مقدرة باللوجيت والواط مرتبة تصاعدياً تبعاً للصعوبة | 184 |
| 42 | معايير القدرة لعينة التقنين وفق نموذج راش (وحدة اللوجيت ووحدة الواط) ووفق النظرية الكلاسيكية في القياس (المعايير التائية – الرتب المئينية) | 185 |
| 43 | نتائج التحليل العاملي للمكونات الأساسية للبواقي (PCAR) بعد التدريج | 188 |
| 44 | ملخص لنتائج معاملات الثبات والفصل وأخطاءهما المعيارية للأفراد والمفردات بعد التدريج النهائي لاختبار الإحصاء في صورته النهائية | 189 |

# قائمة الأشكال


| الرقم | عنوان الشكل | الصفحة |
|---|---|---|
| 01 | أنواع الاختبارات محكية المرجع | 38 |
| 02 | خطوات تقدير الاختبار التنبؤي | 72 |
| 03 | خطوات تقدير الصدق التلازمي للاختبار | 74 |
| 04 | تقدير قيمة معامل الاستقرار | 76 |
| 05 | إجراءات تقدير معامل التكافؤ | 77 |
| 06 | تقدير قيمة معامل الاتساق الداخلي | 77 |
| 07 | العلاقة بين القدرة والاستجابة على الفقرة | 97 |
| 08 | دالة الخطوة للمنحنى المميز للفقرة | 98 |
| 09 | تصنيف لنماذج لاستجابة للمفردة الاختبارية | 104 |
| 10 | المنحنيات المميزة للدالة لثلاثة فقرات في نموذج راش | 106 |
| 11 | المنحنيات المميزة للدالة لثلاثة فقرات في نموذج لورد | 107 |
| 12 | المنحنيات المميزة للدالة لثلاثة فقرات في نموذج بيرنبوم | 109 |
| 13 | تأثير الفرق بين مستوى قدرة الفرد ومستوى صعوبة البند في احتمال حدوث الاستجابة الصحيحة | 112 |
| 14 | منحنى الاستجابة | 113 |
| 15 | المنحنى المميز للمفردة في نموذج راش | 119 |
| 16 | كفايات الإحصاء الوصفي | 141 |
| 17 | خصائص مفردات اختبار مقياس الإحصاء 20 مفردة | 165 |
| 18 | مؤشرات المطابقة للمفردات وحدود المطابقة ومواقع للمفردات من هذه الحدود وفقاً الإحصائي متوسط المربعات التباعدية والتقاربية (MNSQ) | 175 |
| 19 | خصائص مفردات اختبار الإحصاء بعد التدريج باستخدام نموذج راش | 176 |
| 20 | خريطة رايت (Wright) لتدريج الأفراد ومفردات اختبار الإحصاء | 177 |
| 21 | العلاقة التقييسية بين الدرجة الكلية المحتملة على الاختبار عند كل مستوى من مستويات القدرة المختلفة للأفراد | 180 |
| 22 | علاقة دالة معلومات الاختبار مع الخطأ المعياري للقدرة الكامنة | 181 |
| 23 | توزيع العتبات الفارقة للمفردات والأفراد على عينة التدريج | 186 |

# الباب الأول

## الفصل الأول

**مقدمة ومشكلة الدراسة**

**أولاً: أهداف الدراسة**

**ثانياً: أهمية الدراسة**

**ثالثاً: حدود الدراسة**

**رابعاً: المفاهيم الاجرائية للدراسة**

**مقدمـــة**

يُعد مجال القياس والتقويم التربوي من المجالات الحيوية الأساسية التي لا غنى عنها عند الباحثين في العلوم السلوكية، والمسؤولين عن اتخاذ القرارات المتعلقة بالأفراد في مختلف الميادين التربوية. ولقد حظيَ هذا المجال باهتمامٍ كبيرٍ من جانب علماء التربية في الآونة الأخيرة.

ويمثل التقويم أحد أهم المداخل الحديثة في هذا المجال، حيث يتم التقويم باستعمال أدوات قياس نوعية وكمية، وتعتبر الاختبارات التحصيلية أهم هذه الأدوات.

كما يتطلب نجاح العملية التعليمية جمع معلومات دقيقة عن الطلبة تُفسر وتبين مستوى الكفايات والمهارات التي يمتلكونها في مجالات مختلفة، ولاتخاذ قرارات ناجحة تساعد الطلبة في المستقبل ووضعهم في أماكنهم المناسبة، فأدوات القياس والتقويم أدوات فائقة تساهم في جمع هذه المعلومات ذات الأهمية في معرفة مستوى الطلبة، وما يحتاجونه وتوجيههم وتشخيصهم.

وقياس وتقويم قدرات الطلبة وتحصيلهم الدراسي جزءاً لا ينفصل عن العملية التربوية عامة والعملية التدريسية خاصة، وللحصول على معلومات مناسبة تتعلق بقدرات الطلبة وتحصيلهم، لا بدَّ أن تصمَم اختبارات تحصيلية جيدة تتسم بالصدق، والثبات، والموضوعية، والشمولية والاستمرارية وأن يكون مصممها قادراً على إدارة هذه الاختبارات وتصحيحها وتبويبها وتفسير نتائجها.

ونظراً لأهمية الاختبارات التحصيلية كونها تُمثل الأداة الأكثر استخداماً للتقويم في الأنظمة التربوية، فقد عرف العصر الحالي اهتماماً متزايداً من طرف الباحثين في علوم التربية بالدراسات في مجال بناء الاختبارات التحصيلية، فمسألة تحسين هذه الاختبارات وتحديثها شكلاً ومضموناً ورفع فاعليتها أصبح ضرورةً في ظروف الثورة العلمية والتقنية، ومع التوسع الهائل في حجم المعرفة العلمية الذي يشهده عصرنا الحديث (عبد الرؤوف وإيهاب، 2017).

وفي أدب القياس النفسي نجد مدخلين متباينين في القياس، أحدهما معياري المرجع Norm-Referenced Test) ، والآخر محكي المرجع( Criterion-Referenced Test) ففي المدخل معياري المرجع، يتم تصنيف الطالب في ضوء بعض الإحصاءات المشتقة من درجات عينته على الاختبار الذي يقيس السمة المراد تصنيفه فيها، ومن أشهر هذه الإحصاءات : مقاييس النزعة المركزية ومقاييس التشتت، وفي هذه الحالة يجب أن يكون توزيع الدرجات توزيعاً اعتدالياً ومن ثمَّ يقوم مُعد الاختبار

بحذف المفردات المتطرفة مهما كانت الأهداف التي تقيسها، مما يؤثر على صدق الاختبار وتُفسر درجة الطالب في الاختبارات بناءً على معايير محددة، ويتم تفسير أداءَه النسبي بناءً على مقارنة درجته مع متوسط تحصيل المجموعة المعيارية(أقرانه في المجموعة الصفية أو العمرية).

وقد تعرض هذا النوع من الاختبارات لمجموعة من الانتقادات وفق ما أشار إليه هامبلتون وسواميناثان(Hambleton & Swaminathan, 1985) من حيث قصورها في قياس فاعلية البرامج التعليمية وتقويم الطلبة تقويماً تشخيصياً يُحَدَّد في ضوئه جوانب التميز في العملية التعليمية والعوامل المؤثرة فيها، إيجاباً أو سلباً بما لا يتيح إمكانية تحسين تلك النواتج، أو تعديل البرامج التعليمية لأجل تحقيق الأهداف التربوية المرجوة(علام، 2001).

كما أشار (Glaser,1994) بأنَّ هذا الاتجاه يعتمد على تفسير درجاته في ضوء معايير محددة أي تُرد درجات الطالب في الاختبار إلى جداول معيارية، ومن ثمَّ تحديد مستوى أداء النسبي بناءً على موقع درجة الطالب مقارنة بمتوسط تحصيل المجموعة المعيارية في الاختبار ويركز هذا الاتجاه على الفروق الفردية بين الطلبة. وبالتالي فأنَّه من المحتمل أن يختلف موقع الطالب باختلاف خصائص مجموعته المعيارية، إلا أنَّه ليس بالضرورة أن يمثل متوسط المجموعة المعيارية الأداء المطلوب للنجاح (Popham & Husk, 1969).

وكان لهذه الانتقادات الدور الأكبر في ظهور نوع أخر من الاختبارات أطلق علية الاختبارات محكية المرجع (CRT)، حيث أثَّرت مقالة(1963) Glaser والتي كان عنوانها (تكنولوجيا التعليم وقياس مخرجات التعلم) تأثيراً عميقاً، حيث قدم من خلالها مدخلاً جديداً في القياس النفسي والتربوي وهو القياس مرجعي المحك.

فقد أشار(1976) Hambleton, et al إلى أنَّ هذا الاتجاه لا يعتمد على تحديد مستوى الطالب النسبي داخل المجموعة، بل يعتمد على تحديد ما يستطيع وما لا يستطيع الطالب انجازه، ويعتمد هذا النوع في بنائه على تحديد نطاق الأهداف السلوكية تحديداً دقيقاً، ومن ثمَّ قياس مدى اتقان الطالب لهذه الأهداف من خلال مقارنة أداء الطالب بمستوى معين من الأداء (درجة قطع)Cut-off Score ومن ثمَّ تصنيف الطلبة إلى مجموعتين الأولى متقنة Mastery والثانية غير متقنة Non-Mastery بالنسبة للمجال الذي يقيسه الاختبار. كما لم يعد هدفها مقتصراً على مجرد التمييز بين الطلبة في القدرة أو

المستوى، وإنَّما تركز على اكتسابهم مهارات معينة وتحقيق أهداف محددة، بل وإتقان تلك المهارات وتلك المعلومات، كما وتهدف الاختبارات محكية المرجع إلى مقارنة أداء الفرد بمستوى أداء مطلق ومحدد في مجال من السلوك أو بمستوى كفاية معين دون الحاجة إلى مقارنة أداء الفرد بأداء الآخرين كما هو الحال في الاختبارات معيارية المرجع (علام، 2001). وبهذا فإنَّ الاختبار المحكي المرجع يعتمد في جوهره على تقدير مستوى أداء الفرد بالنسبة لمحك أو مستوى أداء أو درجة قطع (Popham,1978).

لكن اعتماد بناء جميع خصائص الاختبارات السيكومترية على خصائص عينة الدراسة وعلى صعوبة عينة مفردات الاختبار، يفترض خطية ميزان درجات الطالب في الاختبار، غير أنَّه عادة ما يكون الميزان منحنياً لأن الفرق الثابت بين درجتين من درجات الاختبار يختلف تفسيره بناء على موقع الدرجات على متصل السمة أو القدرة المراد قياسها، كما يفترض أنَّ العلاقة بين زيادة درجة الطالب في الاختبار وبين مقدار السمة أو القدرة علاقة خطية مطردة، وقد لا يكون هذا الافتراض صائباً إذا لم يحصل الطلبة من ذوي القدرات المتميزة على درجات مرتفعة في الاختبار، وفي المقابل يمكن أن يحقق الطلبة من ذوي القدرات المتدنية على درجات مرتفعة في الاختبار، كما تتأثر مفردات الاختبار بالظروف المحيطة به كحذف أو تغيير أي مفردة من مفردات الاختبار الأمر الذي يؤثر في تمثيل المفردات للنطاق السلوكي المراد قياسه .

ويرى علام(2000) أنَّ الاختبارات محكية المرجع(CRT) أكثر أنواع الاختبارات التحصيلية ملاءمة لقياس وتقويم تحصيل الطلبة، وذلك لأنَّها تقوم على تحديد المهارات والكفايات المطلوب إتقانها بدقة فائقة لكي يتمكن المدرس من قياسها وملاحظتها بشكل مباشر، ومن ثمَّ تقدير مدى ما حققه الطالب من تلك الأهداف، بناءً على مستوى أداء محدد، وهذا مما يساعد المدرس على عملية التشخيص، حيث يتم تصنيف الطلبة إلى فئتين، متقنة وغير متقنة للمهارات والكفايات المحددة يلي ذلك تحديد نقاط الضعف والقوة في مستوى تحصيل الطالب ومعرفة المهارات التي أتقنها والتي لم يتقنها، وبذلك يستطيع المدرس وضع البرامج العلاجية المناسبة لذلك.

وبازدياد الحاجة إلى تحديد الكفايات أو المهارات التي يجب أن يمتلكها الطلبة في المستويات التعليمية المختلفة، ازداد التأكيد على أسلوب الاختبارات محكية المرجع.

والمتتبع لتاريخ بناء وإعداد الاختبارات التحصيلية يلاحظ تعدد وتنوع النظريات ونماذج القياس التي تستند إليها في بناء وتصميم الاختبارات وتحليل البيانات المستمدة منها، وعلى الرغم من هذا التباين والتنوع الذي تستند إليه المداخل إلا أنَّه يمكن تصنيفها إلى قسمين رئيسين هما:

قسم يعتمد على النظرية الكلاسيكية في بناء الاختبارات (Classical Test Theory CTT) وما تنطوي عليه من مبادئ ومفاهيم، بعضها يتعلق بخصائص المفردة (الصعوبة والتمييز)، والبعض الآخر يتعلق بخصائص الاختبار (الصدق، الثبات والمعايير)، وقسم يعتمد على النظرية الحديثة أو نظرية الاستجابة للمفردة(Item Response Theory IRT) ، لاهتمامها بالربط بين خصائص المفردة (علام، 2001)

وقد ظلت النظرية الكلاسيكية هي الأساس النظري والعملي للقياس في العلوم السلوكية، وقد استندت هذه النظرية على افتراض بسيط هو أنَّ درجة الفرد الملاحظة في الاختبار تساوي مجموع الدرجة الحقيقة للفرد في السمة المقاسة والدرجة الخطأ في الأداء على الاختبار أي X=T+E حيث X هي درجة الفرد الملاحظة في الاختبار، T هي الدرجة الحقيقية للشخص في السمة المقاسة، E هي درجة الخطأ في الاختبار (إسماعيل، 2007).

وقد عجزت أساليب القياس التقليدية عن تحقيق الدقة والموضوعية المرجوة في القياس السلوكي حيث وجهت لها عدة انتقادات، على أساس أنَّ أداء الأفراد على الاختبار يختلف باختلاف المفردات المستعملة مما ينعكس على الأداء من اختبار لأخر، كما أنَّ هذه النظرية تفترض تساوي خطأ القياس لجميع الأفراد المطبق عليهم الاختبار مما ينتج عنه أنَّ متوسط خطأ القياس لمجموع الأفراد يساوي صفر، فتباين الأخطاء الموجبة يلغي تباين الاخطاء السالبة.

ونتيجة لهذه المشكلات المصاحبة لنظرية القياس الكلاسيكية في قياس الظواهر النفسية والتربوية والتي نتج عنها عدم دقة النتائج وعدم موضوعية تفسيرها، ظهرت الحاجة إلى تطوير أساليب قياس أكثر دقة وموضوعية، فتعددت البحوث والدراسات من قِبل المهتمين بالقياس النفسي والتربوي بحثاً عن أساليب تحقق هذه الدقة وهذه الموضوعية في تفسيرها للنتائج، فأثمرت الجهود إلى ظهور نظرية الاستجابة للمفردة(IRT) .

فقد أشار علام(1995) إلى أنَّ هذه الانتقادات أدت لظهور هذه النظرية وكنتيجة لمحاولات المهتمين بالقياس النفسي والتربوي لتطوير مقاييس أكثر دقة في قياسها للسمات النفسية والتربوية وفي

محاولة لحل المشكلات التي تعاني منها الطرق التقليدية، وتتلخص فكرة هذه النظرة في ربطها بين خصائص الفقرات بمعلم أو أكثر، ومقدار السمة التي يمتلكها الفرد واحتمال إجابته عن مفردة محددة بمستوى معين من الإجابة حسب نوع المفردة، ثنائية الاجابة أو من عدة فئات، ويقوم هذا الربط أساساً على تحديد موقع الفرد على مقياس السمة الكامنة بصرف النظر عن مجموعة المفردات التي يجيب عنها الفرد(إسماعيل،2007. ص.5).

لقد أصبحت نظرية الاستجابة للمفردة وسيلة أساسية وشائعة في بناء وتطوير الاختبارات حيث أنَّها تقدم بديلاً لنظرية القياس الكلاسيكية في تقديرها لمعالم الفرد والمفردة بأقل قدر ممكن من الخطأ وقد بُنيت هذه النظرية والنماذج الرياضية المتعلقة بها– التي حاولت تحديد العلاقة بين أداء الفرد على مفردات الاختبار وبين القدرات التي يفترض أنها تكمن وراء الأداء (علام، 1992) – على افتراضات قوية وصارمة لتحقيق الموضوعية في القياس كأحادية البعد، والاستقلال الموضعي والمنحنيات المميزة للفقرة وعامل السرعة.

وقد اشارت أمينة كاظم(1988) بأنَّ هذه النظرية وفرت نماذج تقدر السمات الكامنة (قدرة الفرد) بصورة مستقلة عن أفراد عينة الدراسة، وتوجد ثلاثة أبعاد توفر أسس الموضوعية في نماذج السمات وهي: بعد السمة المقاسة الذي يعرف بواسطة مجموعة البنود وهي إما أن تكون أحادية (صعوبة المفردة) نموذج راش، أو ثنائية(صعوبة المفردة ومعامل التمييز) نموذج لورد، أو ثلاثية المعلم(صعوبة المفردة ومعامل التمييز وعامل تخمين الاجابة) نموذج بيرنبوم، وبعد استقلالي القياس حيث لا يعتمد تقدير صعوبة المفردة على صعوبة المفردات الأخرى، كما لا يعتمد أيضاً تقدير قدرة الأفراد على قدرة أي مجموعة أخرى، لأنَّ موضوعية نتائج الاختبارات وصدقها تعتمد على دقة الأساليب التي تستخدم في بنائها واختيار فقراتها، وتفسير نتائجها، و وصفها للقدرة التي يقيسها الاختبار لهذا فإنَّه من الضروري استخدام التوجهات الحديثة في القياس، والتي أثبتت البحوث والدراسات التجريبية أنها تحقق الموضوعية والدقة في قياس قدرة الفرد.

ورغم تعدد النماذج في النظرية الحديثة إلاَّ أنَّ نموذج راش أحادي المعلمة يُعد من أهم وأبسط نماذج هذه النظرية، وأكثرها استخداماً في بناء الاختبارات والمقاييس النفسية والتربوية نظراً لبساطته وسهولة تطبيقه وكذا فاعليته في تحقيق موضوعية القياس (علام، 2001).

فنموذج راش يوفر القياس الموضوعي من حيث استقلالية صعوبة البند عن كل من صعوبة بقية بنود الاختبار الاخرى، وقدرات الأفراد المطبق عليهم الاختبار. كما أنَّ جميع بنود الاختبار لها قوة تميزية واحدة(كاظم، 1988 أ)، كما يُعد من أكثر النماذج التي أجريت حولها دراسات متعددة كدراسة كل من لمياء وصبَّاح(2016)، بعنوان (بناء اختبار تحصيلي في الهندسة لطلبة الصف الخامس الأساسي باستخدام نموذج راش، ودراسة العلي(2015) بعنوان (استخدام نموذج راش في تدريج مقياس سوء التكيف الدراسي من اختبار مينسوتا MMPI-2 لدراسة بعض العوامل المؤثرة على دقة القياس) ودراسة كل من السمرائي والخفاجي(2013)، بعنوان (بناء اختبار تحصيلي محكي المرجع في مادة علم النفس الخواص لطلبة أقسام العلوم التربوية والنفسية) ودراسة الخياط(2012) بعنوان (درجة مطابقة اختبار تحصيلي وفق نموذج راش في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن أساسي)، ودراسة أبو جراد(2008) بعنوان (استخدام نموج راش في تطوير اختبار كاتل للذكاء الصورة (أ) ).

كما تطور نموذج راش مستقلاً عن بقية النماذج، ويمكن النظر إليه على أنَّه دالة لوغاريتمية أحادية المعلم للمنحنى المميز للفقرة. أي يمكن اعتباره دالة للفرق بين موقعي المفردة والفرد على بعد السمة التي يقيسها الاختبار، حيث أنَّه يهتم بتحديد موقع المفردة الاختبارية على ميزان صعوبة جميع المفردات التي تشكل الاختبار، كما يهتم بتدريج مستويات قدرات الأفراد باختبار معين على نفس مستويات القدرة المقاسة(Wright, 1977) .

ويستخدم هذا النموذج في تحليل الاختبارات التي تكون الاجابة عنها ثنائية(صح/خطأ) وتكون درجة كل مفردة إما (1 أو0) وعندما يستجيب الفرد لمفردة اختبار ما، فإنَّه يحدث تفاعل بين قدرة هذا الفرد وصعوبة المفردة.

وقد جاءت هذه الدراسة لبناء اختبار تحصيلي محكي المرجع وفق نموذج راش، وذلك تأكيداً على أهمية الاختبارات محكية المرجع وضرورة اعدادها بالشكل الجيد لتحقيق الغرض الذي وضعت من أجله في قياس تحصيل المتعلمين ورفع كفايتهم، وكمؤشر على مستويات أدائهم من جهة، ومن جهةٍ أخرى فأنَّ استخدام نموذج راش في تدريج بنود الاختبارات التحصيلي في مقرر الاحصاء الوصفي يعطينا نتائج موثوقة يمكن الاعتماد عليها في اتخاذ مختلف القرارات.

وعلى الرغم من أهمية الاختبارات التحصيلية في حياة الطلبة والمدارس والمؤسسات التعليمية تشير بعض الدراسات العربية والجزائرية في بناء الاختبارات التحصيلية، أنَّ هناك تفاوت في جودة الاختبارات المستخدمة، وأنَّ أغلب النتائج تشير إلى أنَّ هناك قصوراً في إعداد وبناء اختبارات موضوعية تقيس تحصيل الطلبة، وترتبط بالعديد من المشكلات، منها المتعلقة بأدوات القياس ومنها المرتبط بالمعايير التي يستند إليها في تفسير النتائج المتحصل عليها من الاختبار ومشكلات ترتبط بطبيعة القياس وفلسفته بين النظريتين الكلاسيكية والحديثة، فعلى المستوى العربي في دراسة منار الدسوقي وسهام الشافعي ومنى أبو شنب(2011) على أستاذات جامعة تبوك، أسفر البحث عن انخفاض جودة الاختبارات التحصيلية التي تُعد من قبل عضوات هيئة التدريس، وفي دراسات أخرى لكل من الخضيري(2003) الزيلعي(2001)، والزهراني(1994) والتي تتلخص نتائجهم فيما يلي:
ميل واضعوا الاختبار للأسئلة المقالية أكثر منه للأسئلة الموضوعية، وجود اختلال واضح في توزيع الأسئلة على موضوعات المحتوى الدراسي، عدم الاهتمام بالخطوات العلمية في بناء الاختبارات التحصيلية، عدم مراعاة التسلسل المنطقي في وضع أسئلة الاختبار من السهل إلى الصعب.

وفي الدراسات الجزائرية، دراسة لصباح سعد ووسيلة بن عامر (2017) حول تقييم كفاية بناء الاختبارات التحصيلية لدى اساتذة التعليم الجامعي في الجزائر وفق معاير الاختبار الجيد، فقد توصلت الباحثتان إلى أنَّ معايير بناء الاختبار التحصيلي الجيد وإعداد فقراته جاءت بنسب ضعيفة وأنَّ الأساتذة أكثر استخداماً لأسئلة المقال مقارنةً مع الاسئلة الموضوعية.

كل ذلك وغيره كان له عظيم الأثر في تزايد الحاجة إلى بناء أدوات قياس موثوقة بما يحقق درجة عالية من الموضوعية والدقة في تحديد مختلف ما يمتلكه الطلبة من سمات وقدرات.

وقد اتسمت هذه الدراسة بنظرة شمولية لأهم المهارات الاحصائية التي تكوِّن مقرر الاحصاء الوصفي وما يتوقع من الطالب أن يتقنه في السنة الاولى علوم اجتماعية جذع مشترك، خاصة مع تنامي استخدام الاحصاء في شتى الفروع النظرية والتطبيقية، حيث باتت هذه المادة مقرراً الزامياً على معظم طلاب الجامعات الجزائرية في أغلب التخصصات، وذلك بهدف تمكين الطلبة من استخدام وفهم البحوث المتعلقة بالبيانات والاحصائيات.

غير أنَّ الواقع يشير إلى فتور حماس الطلبة نحو مقرر الاحصاء وموقفهم السلبي منه وكذا تدني درجات الطلبة وتحصيلهم في هذا المقرر.

لذلك تتلخص مشكلة البحث في التساؤلات التالية:

1. **ما درجة تحقيق بيانات اختبار مقرر الإحصاء الوصفي لافتراضات نموذج راش لدى طلاب السنة أولى علوم اجتماعية؟**
2. **ما قيم تقديرات معلم الصعوبة لفقرات الاختبار بعد التدريج وفق نموذج راش؟**
3. **ما كمية المعلومات التي يقدمها اختبار الاحصاء الوصفي عند مستويات القدرة المختلفة؟**
4. **هل يمكن اشتقاق معايير كمية تفسر تقديرات الأفراد على اختبار الإحصاء الوصفي في صورته النهائية بعد تدريجه على الطلبة؟**
5. **ما دلالات الخصائص السيكومترية (الصدق والثبات) لمفردات الاختبار بعد معايرة مفرداته باستخدام نموذج راش؟**

**أولاً: أهداف الدراسة**

- هدفت هذه الدراسة إلى استخدام نظرية الاستجابة للمفردة في إعداد اختبار محكي المرجع في مقياس الإحصاء وفق النموذج اللوجيستي أحادي المعلم (نموذج راش) لقياس تحصيل طلاب السنة الأولى علوم اجتماعية جذع مشترك بجامعة زيان عاشور بالجلفة.
- هدفت هذه الدراسة لتقديم فقرات لبناء اختبار محكي المرجع يمكن الاعتماد عليها في قياس تحصيل الطلبة في الإحصاء بدلاً من الاعتماد على اختبارات غير مقننة.
- هدفت الدراسة لدراسة مؤشرات صدق وثبات الاختبار محكي المرجع المُعد لقياس تحصيل الطلبة في مقياس الإحصاء الوصفي.
- هدفت الدراسة للتحقق من درجة تحقيق بيانات اختبار مقرر الاحصاء الوصفي لافتراضات نموذج راش
- هدفت الدراسة لتقدير قيم معلم الصعوبة لفقرات الاختبار بعد تدريجه وفق نموذج راش.
- هدفت الدراسة للتحقق من مدى اشتقاق معايير كمية تفسر تقديرات الأفراد على اختبار الاحصاء الوصفي.

**ثانياً: أهمية الدراسة**

**أ). الأهمية العلمية:**

يضيف استخدام نموذج (راش) اللوغاريتمي أحادي المعلم لهذه الدراسة أهمية في تحقيق الموضوعية في القياس، وذلك من خلال تدريج معالم فقرات الاختبار التحصيلي التي تعبر عن أداء الطلبة على مقياس السمة الكامنة بشكل مستقل عن خصائص العينة، وذلك بعكس النظرية الكلاسيكية في القياس التي تعطي معالم غير ثابتة لقيم التمييز، وصعوبة الفقرات باختلاف عينة الأفراد التي يطبق عليها الاختبار، إضافة إلى تقديرات غير ثابتة لأداء الأفراد، وباختلاف مجموعة الفقرات التي تتم استجاباتهم عليها، كما أنَّه من الناحية النظرية فإنَّ استقصاء فاعلية نموذج راش والذي يستخدم في تحليل فقرات ثنائية الاستجابة يعتبر ذو فائدة في إضافة دليل على أهمية هذا الأسلوب في تحليل الفقرات، هذا إلى جانب توفير اختبار يتضمن فقرات تعكس قدرات الطلبة مقاسة على نفس المتصل بحيث يمكن إجراء مقارنة مباشرة بين قدرات الطلبة على اختلاف قيمها في متصل القدرة.

**ب). الأهمية التطبيقية:**

تسهم هذه الدراسة في تحقيق الجوانب التالية:

- فحص افتراضات نظرية الاستجابة للمفردة وتحديداً نموذج راش فيما يتعلق بتوضيح أثر خصائص المفحوصين والفقرات، على دقة التقدير لمعالم المفحوصين والفقرات.
- زيادة قدرة معدي اختبارات مقرر الاحصاء على تحديد الطريقة التي بواسطتها يمكن اختيار الفقرات الملائمة، وذلك للحصول على دقة أكبر في تقدير قدرات عينة من المفحوصين في حالة توافر بنك للأسئلة.
- توفير تجمع فقرات لبناء اختبار تحصيلي يمكِّن من قياس تحصيل الطلبة للسنة أولى علوم اجتماعية في مقياس الاحصاء الوصفي بمستوى عالٍ من الدقة والموضوعية.

**ثالثاً: حدود الدراسة**

- **الحدود الزمنية:** تمَّ تطبيق هذه الدراسة في السداسي الأول من السنة الجامعية 2020-2021.
- **الحدود المكانية:** تمَّ تطبيق هذه الدراسة بمعهد العلوم الاجتماعية والانسانية بجامعة زيان عاشور الجلفة.

- **الحدود البشرية:** اقتصرت عينة الدراسة على طلبة السنة أولى علوم اجتماعية جذع مشترك بجامعة زيان عاشور بالجلفة.

- **حدود موضوعية:** تناولت هذه الدراسة بناء اختبار تحصيلي محكي المرجع وفق نموذج راش أحادي المعلم أحد نماذج النظرية الحديثة في القياس، في محتوى مقرر الإحصاء الوصفي المقرر تدريسه في السداسي الأول لطلبة سنة أولى علوم اجتماعية جذع مشترك من الموسم الجامعي2020.2021.

**رابعاً: المفاهيم الاجرائية للدراسة**

**1. اختبار محكي المرجع *Criterion -Referenced Test***

- **اجرائياً:** هو الاختبار في مقرر الاحصاء الذي أعده الباحث والمؤلف من (20) فقرة من نوع الاختيار من متعدد بواقع أربعة بدائل (أحدها صحيح والثلاث الباقية خاطئة) أُعدَّ وفقاً للخطوات العلمية المتعارف والمتفق عليها بين علماء القياس النفسي والتربوي، ويستخدم لقياس تحصيل الطلبة في مقياس الإحصاء الوصفي. وحدد محك الاتقان له باثني عشر (12) درجة وفق طريقة أنجوف، والتي يجب الحصول عليها لوصول الطالب إلى مستوى الاتقان في مقرر الاحصاء الوصفي.

# الفصل الثاني: الدراسات السابقة

I.الدراسات السابقة

أولاً: الدراسات التي استخدمت بناء الاختبارات محكية المرجع

ثانياً: الدراسات التي استخدمت الاختبارات التحصيلية في مقرر الإحصاء

ثالثاً: الدراسات التي استخدمت نموذج راش

II. التعقيب عن الدراسات السابقة

III.موقع الدراسة الحالية من الدراسات السابقة

**I.الدراسات السابقة**

يفيد عرض الدراسات السابقة في البحوث العلمية في التعرف على ما توصلت إليه هذه الدراسات، وليكون نقطة انطلاق أي بحث علمي بناءً عن الأدوات المستعملة، والنتائج والمقترحات التي توصَّل إليها الباحثون للاستفادة منها في الدراسة الحالية، وكون الدراسة الحالية هي بناء اختبار محكي المرجع في مقرر الاحصاء وفق نموذج راش لطلبة السنة الأولى علوم اجتماعية جذع مشترك بجامعة زيان عاشور بالجلفة، فإنَّ الباحث يرى عرض الدراسات السابقة على النحو التالي:

- دراسات تناولت بناء الاختبارات محكية المرجع
- دراسات تناولت الاختبارات التحصيلية في مقرر الاحصاء
- دراسات تناولت استخدام نموذج راش

**أولاً: دراسات استخدمت بناء الاختبارات محكية المرجع**

**1. دراسة قارة (2018)**

**الموضوع: بناء اختبار تشخيصي محكي المرجع لقياس كفايات المعلمين في بناء الاختبارات التحصيلية الموضوعية.**

**أهداف الدراسة:** بناء اختبار تشخيصي مرجعي المحك لقياس كفايات معلمي المرحلة الابتدائية في بناء الاختبارات التحصيلية الموضوعية ببلدية طولقة ولاية بسكرة، والتحقق من خصائصه السيكومترية واستخدام الاختبار المقترح في التعرف على درجة إتقان معلمي المرحلة الابتدائية لكفايات بناء الاختبارات التحصيلية الموضوعية، وتحديد نقاط قوتهم وضعفهم في هذا المجال.

وتمَّ إتباع المنهج الوصفي التحليلي، حيث تكوَّن مجمع الدراسة من معلمي مرحلة التعليم الابتدائي والبالغ عددهم (260) معلم ومعلمة، تمَّ اختيار عينة الدراسة المقدرة بـــ(140) معلماً ومعلمة عن طريق العينة العشوائية متعددة المراحل.

**أداة الدراسة:** قامت الباحثة ببناء اختبار تشخيصي مرجعي المحك وفقاً للخطوات المتعارف عليها لدى علماء القياس والتقويم التربوي وفقا لثلاث مراحل (مرحلة التحليل، مرحلة البناء، مرحلة التجريب) حيث تكون الاختبار من(68) مفردة اختبارية بعد التأكد من خصائصه السيكومترية ومن خصائص الفقرة الاختبارية(معاملات، الصعوبة ومعاملات التمييز)، وكذا التحقق من الصدق بعدة طرق:(الصدق

الوصفي، الصدق التمييزي، الصدق البنائي)، كما تمَّ التحقق من مؤشرات المفردات بطرق متعددة(ثبات التجزئة التصفية، ثبات ألفا كرونباخ، ثبات كيود ريتشاردسون، طريقة هاريس، طريقة ليفنجستون).

**نتائج الدراسة:**

- الاختبار التشخيصي مرجعي المحك الذي تمَّ بناءه يتوفر على خصائص الاختيار الجيد، حيث اتصفت فقرات الاختبار بمعاملات صعوبة جيدة ومعاملات تمييز مقبولة، وكذا توفر دلالات صدق وثبات الاختيار.
- انخفاض درجة إتقان معلمي المرحلة الابتدائية ببلدية طولقة لكفايات بناء الاختبارات التحصيلية الموضوعية على فروع الاختبار التشخيصي مرجعي المحك والاختيار الكلي فيما عدا الاختبار الفرعي الثالث الخاص بكفاية إخراج الاختيار التحصيلي الموضوعي وتصحيحه حيث اظهرت النتائج أنَّ درجة إتقان المعلمين لهذه الكفاية جاءت بدرجة متوسطة.
- نقاط قوة معلمي المرحلة الابتدائية في مجال بناء الاختبارات التحصيلية الموضوعية أقل من نقاط ضعفهم وتمحور نقاط ضعفهم حول:

- تحديد الغرض من الاختبار التحصيلي الموضوعي، صياغة الأهداف السلوكية للمادة الدراسة بناء جدول مواصفات المادة وكذا تحديد الأسس والشروط التي تعتمد في صياغة الفقرات، وكذا تحديد الأخطاء في صياغة بعض الفقرات وتقديم التعديلات المناسبة، وأيضاً الطريقة المعتمدة في ترتيب الاسئلة الموضوعية، وكذا تصحيح الاختبارات الموضوعية والتصحيح من أثر التخمين، وكذا تفسير نتائج المتحصل عليها.

**2. دراسة الشهري (2016)**

**الموضوع: بناء اختبار تشخيصي محكي المرجع لقياس المهارات الإشرافية للمشرف التربوي.**

**أهداف الدراسة:** هدفت الدراسة لبناء اختبار تشخيصي محكي المرجع لقياس المهارات الإشرافية للمشرف التربوي، كما هدفت للكشف عن مدى تمكن المشرفين التربويين من المهارات الأساسية للإشراف التربوي.

**أداة الدراسة:** اختبار التشخيصي محكـــي المرجع لقيـــاس المهارات الإشرافية للمشرف التربـــوي واستخدمت الدراسة المنهج الوصفي المسحي، وتكون مجتمع الدراسة الكلي من المشرفين التربويين

من الجنسين (ذكور وإناث) الذين يتبعون إدارة التربية والتعليم بمنطقة تبوك والمحافظات الست التابعة لها والبالغ عددهم (460) وتكونت عينة الدراسة من (207 مشرفين ومشرفات تربويين) يشكلون ما نسبة 45% من حجم مجتمع الدراسة.

**نتائج الدراسة:** توصلت الدراسة إلى تصميم اختبار تشخيصي محكي المرجع لقياس المهارات الإشرافية للمشرف التربوي، كما أنَّ مجالي توفير الخدمات الأساسية للطلاب وتوثيق العلاقة بين المدرسة والمجتمع المحلي كانت درجة تقديرهما متوسطة، وباقي المجالات كانت درجة تقديرها مرتفعة، وأكبر تقدير كان لمجال تنمية علاقات إيجابية بين العاملين وبلغت قيمة المتوسط الحسابي له (1.29) بانحراف معياري (0.50).

كما أظهرت النتائج وجود فروق في درجة ممارسة المهارات الإشرافية للمشرفين التربويين تعزى للجنس على مجال توفير الخدمات الأساسية للطلبة فقط، وكانت الفروق لصالح الذكور ووجود فروق في درجة ممارسة المهارات الإشرافية للمشرفين التربويين تعزى للمؤهل العلمي على الكلي وعلى جميع المجالات ما عدا مجالي تقديم الخدمات المطلوبة وتوفير الخدمات الأساسية للطلاب وكانت الفروق لصالح حملة درجة البكالوريوس، ووجود فروق في درجة ممارسة المهارات الإشرافية للمشرفين التربويين تعزى للخبرة على الكلي وعلى جميع المجالات ما عدا مجالي تقديم الخدمات المطلوبة وتنظيم العملية التعليمية وإدارته، وكانت الفروق لصالح ذوي الخبرة أكثر من 10 سنوات.

**3. دراسة أفولابي (2015) Afolabi**

**الموضوع: استخدام اختبار الكفاءة الرياضية محكي المرجع للصف الثامن الابتدائي للتنبؤ بسلوك الطالب في المدرسة العليا.**

**هدف الدراسة:** هدفت هذه الدراسة إلى تحديد ما إذا كان بالإمكان استخدام اختبار الكفاءة الرياضية محكي المرجع للصف الثامن الابتدائي للتنبؤ بسلوك الطالب في المدرسة العليا، واستخدمت الدراسة المنهج الكمي ذو التصميم الارتباطي، لتحديد ما إذا كانت هنالك علاقات بين متغيرات مختارة ودرجات الطلبة على الاختبار محكي المرجع والمتغيرات التي تمَّ اختيارها هي: النجاح في مساق الرياضيات لطلبة الصف الثامن، والبقاء في الصف التاسع، والنجاح في اختبار جورجيا للتخرج من المدرسة الثانوية والتخرج من المدرسة العليا.

**أداة الدراسة:** اختبار الكفاءة الرياضية محكي المرجع، وتكون مجتمع الدراسة من مدرسة ريفية صغيرة تقع في مدينة جورجيا، وتمَّ اختيار عينة الدراسة من طلبة الثانوية حديثي التخرج من المدرسة والبالغ عدد طلابها (300) وتمَّ اختيار (120) طالبة للمشاركة في الدراسة. وتمَّ تحليل بيانات الدراسة باستخدام أساليب التحليل الإحصائي.

**نتائج الدراسة:** أظهرت نتائج الدراسة وجود علاقة إيجابية قوية بين علامات الطلبة على اختبار الكفاءة الرياضية محكي المرجع، ونجاح الطلبة في مادة الرياضيات في الصف التاسع، وأظهرت نتائج الدراسة أيضاً وجود علاقة إيجابية بين علامات الطلبة على اختبار الكفاءة الرياضية محكي المرجع ونجاحهم في اختبار جورجيا للتخرج من المدرسة الثانوية، وتعزز نتائج هذه الدراسة الحاجة إلى إيجاد تنبؤات أو مؤشرات ملائمة للتنبؤ مستقبلاً بسلوكيات الطلبة. وأظهرت نتائج الدراسة الحاجة إلى القيام بذلك في أسرع وقت ممكن، للتقليل من خطورة تسرب الطلبة من المدرسة الثانوية.

**4. دراسة الثبيتي (2014)**

**الموضوع: بناء اختبار تشخيصي مرجعي المحك لقياس مهارات المعلمين في بناء الاختبارات التحصيلية لمقرر الحاسب الآلي.**

**أهداف الدراسة:** هدفت هذ الدراسة إلى بناء اختيار تشخيصي مرجعي المحك لقياس مهارات المعلمين في بناء الاختبارات التحصيلية لمقرر الحاسب الآلي، من خلال التعرف على الخصائص السيكومترية التي يتميز بها الاختبار التشخيصي مرجعي المحك.

**أداة الدراسة:** اختبار تشخيصي مرجعي المحك لقياس مهارات المعلمين في بناء الاختبارات التحصيلية لمقر الحاسب الآلي وتحديد درجة القطع بطريقة أنجوف حيث بلغت للاختبار ككل (76%) للتفريق بين المعلمين المتقنين لمهارات بناء الاختبارات التحصيلية وغير المتقنين لذلك، وايجاد أثر كل من نوعية المؤهل وسنوات الخبرة وعدد الدورات التدريبية وهل هناك فروق ذات دلالة إحصائية تعزى لهذه المتغيرات. وتحقيقاً لأهداف الدراسة استخدم الباحث المنهج الوصفي حيث تمَّ اختيار عينة عشوائية عنقودية ممثلة للمجتمع الإحصائي الخاص بالدراسة وهم معلمي الحاسب الآلي في مدارس المرحلة المتوسطة والثانوية التابعة لمكاتب التربية والتعليم بالإدارة العامة للتربية والتعليم بمحافظة الطائف في الفصل الدراسي الأول للعام الدراسي 1434 –1435 هـ.

**نتائج الدراسة:** تمتعت فقرات الاختبار التشخيصي مرجعي المحك بصعوبة جيدة وبقدرة تميزية مقبولة كما أظهرت المؤشرات دلالات كافية على صدق وثبات الاختبار، ودلت النتائج على وجود تدني واضح في مستوى اتقان المعلمين المهارات بناء الاختبارات التحصيلية لمقرر الحاسب الألي على جميع فروع الاختبار التشخيصي مرجعي المحك وعلى الاختبار بشكل كلي، وجود فروق ذات دلالة احصائية عند مستوى الدلالة(0.05) بين متوسطات درجات الاختبارات الفرعية والدرجة الكلية حسب سنوات الخبرة ولصالح سنوات الخبرة الكبيرة، وجود فروق ذات دلالة إحصائية عند مستوى الدلالة(0.05) بين متوسطات درجات الاختبارات الفرعية والدرجة الكلية حسب نوعية المؤهل، ولصالح الماجستير التربوي وجود فروق ذات دلالة إحصائية عند مستوى الدلالة(0.05) بين متوسطات درجات الاختبارات الفرعية والدرجة الكلية حسب عدد الدورات التدريبية، ولصالح الدورات التدريبية المتعددة.

**5.دراسة الزيلعي (2014)**

**الموضوع: بناء اختبار تشخيصي محكي المرجع لقياس مهارات البحث العلمي لدى طلاب الدراسات العليا بكلية التربية بجامعة الباحا.**

**أهداف الدراسة:** هدفت الدراسة لقياس مدى تمكن الطلبة من مهارات البحث العلمي لدى طلاب الدراسات العليا بكلية التربية بجامعة الباحا، ومعرفة الفروق في مدى تمكنهم في ضوء بعض المتغيرات وهي (الجنس، التخصص، دراسة مقرر مناهج البحث، دراسة مقرر التحليل الإحصائي دراسة مقرر الكتابة العلمية).

**عينة الدراسة:** تكونت عينة الدراسة من (210) بواقع (126) طالبة و(84) طالب من طلاب الماجستير بكلية التربية في جامعة الباحة بنسبة (44%) تقريباً من حجم مجتمع الدراسة البالغ عدده (476) طالباً وطالبة. وتمَّ اختيار العينة بالطريقة العشوائية الطبقية.

**أداة الدراسة:** اختبار محكي المرجع مكون من (40) مفردة من نوع الاختيار من متعدد خماسي البدائل. يقيس عشر مهارات فرعية للبحث العلمي، إذ قام الباحث ببناء هذا الاختبار لقياس مهارات البحث العلمي مروراً بثلاث مراحل هي (مرحلة التحليل، مرحلة البناء، مرحلة التجريب). واستخـــدم المنهج الوصفي المسحي.

وقد بلغ متوسط معاملات سهولة فقرات الاختبار (0,33) وبمـــدى تراوح بين (0.07، 0.80).

وبلغ متوسط معاملات صعوبة الاختبار (0.67) بمدى تراوح بين (0.20،0.93) وقد بلغ متوسط معاملات التمييز الاختبار(0.83) بمدى تراوح بين(0.25،1). وحُسب صدق الاختبار وفق عدة طرق من خلال الصدق الوصفي وحساب معامل الاتفاق بين المحكمين والذي بلغ (0.96)، كما حسب صدق المحتوى من خلال حساب معامل الارتباط(بايسيريل) بين درجة كل مفردة مع الدرجة الكلية للاختبار وكانت جميع مفردات الاختبار ذات علاقة ارتباطية دالة إحصائية عن مستوى الدلالة (0.05) وتراوحت قيم معاملات الارتباط بين (0.53،0.42). وحسب صدق الاتساق الداخلي من خلال حساب معامل الارتباط بين درجة كل مهارة مع الدرجة الكلية للاختبار، وقد كانت جميعها دالة عند مستوى الدلالة (0.05) مع الدرجة الكلية للاختبار وتراوحت قيم معاملات الارتباط بين (-0.45،0,52). كما حسب ثبات الاختبار من خلال معامل لفنجستون وتراوحت قيم معاملات لفنجستون بين (0,93،0,87) وجميعها دالة عند مستوى الدلالة (0.05)، وحسب معامل سبكوفياك وتراوحت قيم معاملات سبكوفياك بين (0,92،0.85) وجميعها دالة عند مستوى الدلالة (0.05) ووجدت درجة القطع وفق طريقة أنجوف والتي بلغت (80%) لست مهارات فرعية، و(75%) لأربعة مهارات فرعية و(76,38%) للاختبار الكلي.

**نتائج الدراسة:** بينت النتائج تمتع الاختبار الذي تمَّ بناؤه في هذه الدراسة بخصائص الاختبار الجيد إذ تمتع بدلالات صدق وثبات جيدة، كما أنَّه يتمتع بخصائص مفردات جيدة. وأظهرت نتائج الدراسة انخفاض واضح في مدى التمكن لدى طلبة وطالبات الدراسات العليا بكلية التربية بجامعة الباحة سواءً في نتائج الاختبار الكلي أم نتائج الاختبارات الفرعية إذ كانت نسبة المتمكنين من الاختبار الكلي (11,4%).

**6. دراسة السامرائي والخفاجي (2012)**

**الموضوع: بناء اختبار محكي المرجع في مادة علم النفس الخواص لطلبة اقسام العلوم التربوية.**

**أهداف الدراسة:** هو بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص لطلبة أقسام العلوم التربوية والنفسية في كليات التربية، مبني وفق نظرية السمات الكامنة ونماذج القياس المنبثقة عنها (نموذج راش) تحديداً، لما يتمتع به هذا الاختبار من دقة في قياس خصائص الأفراد التربوية والنفسية، من خلال التحرر من أثر قدرة الفرد على معالم فقرات الاختبار، والتحرر من أثر معالم الفقرات على قدرة الفرد الذي يجيب عنها. وتتلخص أهمية البحث في تحقيق الموضوعية في تقويم

التحصيل الدراسي للطلبة والمساعدة في تطوير العملية التعليمية، لما يحظى به هذا النوع من الاختبارات من مميزات تساعد على تشخيص صعوبات التعلم وسرعة وسهولة عملية التقويم، كما أنَّه يوفر أداة لقياس أقصى أداء يمكن أن يحققه الطالب والتي يمكن بواسطتها معرفة المستويات المختلفة لأداء الطلبة بالنسبة للمستوى المطلوب مما يمكن المدرس من تشخيص حالات التأخر الدراسي.

**أداة الدراسة:** اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص حيث تمَّ تحديد مفردات المادة من خلال كتاب الهيئة القطاعية لوزارة التعليم العالي والبحث العلمي واتفاق مدرسي المادة عليها وعلى الكتاب المنهجي المقرر وتألف الاختبار من (190) فقرة اختبارية من نوع الاختيار من متعدد القياس (190) هدفاً سلوكياً تغطي مفردات المادة وبحسب اتفاق المحكمين، طبق الاختبار على عينة طبقية عشوائية بلغت (399) طالباً وطالبة.

**نتائج البحث:**

- أخضعت بيانات فقرات الاختبار إلى التحليل العاملي بطريقة المكونات الأساسية، الذي أظهر أنَّ فقرات الاختبار تقيس عاملاً عاماً واحداً عدا (22) فقرة تمَّ استبعادها من الاختبار.
- أظهرت النتائج وجود (13) فقرة كانت قيمة كاف تربيع لعشرة فقرات منها دالة إحصائيا لذا تمَّ استبعادها من الاختبار، فيما استبعد برنامج راسكال ثلاث فقرات في الصفحة الأولى من التحليل الإحصائي لعدم مطابقتها لنموذج راش.
- اقتراب معامل التمييز للاختيار من (1) إذ كانت قيمة معامل تمييز الاختبار ضمن المدى المقبول.
- طابقت الفقرات المتبقية والبالغة (125) فقرة متطلبات استقلالية القياس.
- بلغت قيمة ثبات الاختبار (0.957).

**ثانياً: دراسات استخدمت الاختبارات التحصيلية في مقرر الاحصاء.**

**1. دراسة العطوي والمسعودي (2019)**

**الموضوع: بناء اختبار محكي المرجع لقياس كفايات الاحصاء لدى طلبة الدراسات العليا بجامعة تبوك.**

**أهداف الدراسة:** هدفت الدراسة إلى بناء اختبار محكي المرجع لقياس كفايات الإحصاء لدى طلبة الدراسات العليا وفق نموذج راش، ولتحقيق أهداف البحث اعتمدت الباحثة المنهج الوصفي التحليلي

**أداة الدراسة:** اختبار محكي المرجع لقياس كفايات الاحصاء وتمَّ بناء الأداة على أربع مراحل (مرحلة التحليل– مرحلة البناء– مرحلة التجريب– مرحلة الإخراج)؛ حيث تمَّ بناء اختبار محكي المرجع مكون من(47) فقرة، تغطي الكفايات الضرورية للإحصاء وهي:(المفاهيم الأساسية في الإحصاء– مقاييس النزعة المركزية والتشتت – تفسير وتحديد معامل الارتباط – الأساليب الإحصائية البارامترية – الأساليب الإحصائية اللابارامترية – تفسير النتائج المستخرجة من البرنامج الإحصائيSPSS– اختيار الأسلوب الإحصائي المناسب لأهداف وفرضيات البحث، وللتأكد من صدق وثبات الاختبار تمَّ تطبيق الأداة على عينة استطلاعية عددها(30) وتمَّ التحقق من صدق المحتوى وصدق البناء، وبلغ معامل الثبات بطريقة ألفا كرونباخ(0.96) مما يشير إلى أنَّ الاختبار يتمتع بدرجة ثبات عالية، وتمَّ فحص الخصائص السيكومترية للفقرات من خلال برنامج(SPSS) فكانت معاملات الصعوبة متوسطة ومعاملات التمييز مناسبة.

**نتائج الدراسة:** أشارت نتائج الدراسة بعد تطبيق الاختبار على عينة البحث المكونة من (200) فرد باستخدام برنامج (Bilog-mg3) إلى مطابقة (47) فقرة النموذج راش، ومطابقة بيانات (200) فرد للنموذج وتمَّ تحديد درجة القطع بطريقة أنجوف التحكيمية، وبلغت قيمتها (32) درجة، وحاز الاختبار على معامل ثبات مرتفع وفق نظرية القياس الحديثة (0.94) وهذا مؤشر على أنَّ الاختبار يتمتع بدرجة عالية من الدقة والموضوعية، ومن شأنه التمييز بين المتقنين وغير المتقنين لكفايات الإحصاء من طلبة الدراسات العليا.

**2. دراسة الطراونة (2016)**

**الموضوع: بناء اختبار محكي المرجع في الاحصاء التحليلي لطلبة الدراسات العليا في كليات العلوم التربوية في الجامعات الاردنية الحكومية باستخدام نظرية الاستجابة للفقرة.**

**أهداف الدراسة:** هدفت الدراسة إلى بناء اختبار محكي المرجع في الإحصاء التحليلي لطلبة الدراسات العليا في كليات العلوم التربوية في الجامعات الأردنية الحكومية باستخدام نظرية استجابة الفقرة.

**أداة الدراسة:** اختبار محكي المرجع في الاحصاء التحليلي وتكونت عينة الدراسة من (319) طالبـاً وطالبة من الماجستير في كليات العلوم التربوية في الجامعات الأردنية.

**نتائج الدراسة:** تمَّ التحقق من مطابقة عينة الدراسة وفقرات الاختبار لافتراضات نموذج راش، وتمَّ

التحقق من افتراضات نظرية استجابة الفقرة، تمَّ حساب معاملات الصعوبة لفقرات الاختبار باستخدام البرنامج الإحصائي (Bigsteps) وتكون الاختبار بصورته النهائية من (36) فقرة، وأظهرت النتائج تمتَّع الاختبار بدلالات صدق وثبات مقبولة.

**3. دراسة الحباشة (2015)**

**الموضوع: بناء اختبار محكي المرجع لقياس الكفايات البحثية والاحصائية لطلاب الدراسات العليا في الجامعات الأردنية الحكومية.**

**أهداف الدراسة:** هدفت الدراسة إلى بناء اختبار محكي المرجع لقياس الكفايات البحثية والإحصائية لطلبة الدراسات العليا في الجامعات الأردنية الحكومية.

**أداة الدراسة:** اختبار محكي المرجع لقياس الكفايات البحثية والاحصائية لطلاب الدراسات العليا في الجامعة الاردنية الحكومية، وقد تكوَّن الاختبار بصورته الأولية من(50) فقرة من نوع الاختيار من متعدد، وتمَّ تحديد درجة القطع باستخدام طريقة أنجوف، وبلغت(69) ثم طبق الاختبار بصورته النهائية المكون من(45) فقرة على العينة الكلية المؤلفة من(270) طالباً وطالبة من الجامعات الأردنية الحكومية وقدر معامل الثبات باستخدام معامل كيودريتشاردسون20(0.83) لاستخدامه في التحقق من ثبات ليفنجستون عند درجة القطع، حيث بلغت قيمته(0.87).

**نتائج الدراسة:** أظهرت النتائج وجود فروق دالة إحصائية في درجة امتلاك الكفايات تعزى للنوع الاجتماعي، كما أظهرت وجود فروق دالة إحصائية في درجة امتلاك الكفايات تعزى للدرجة العلمية لصالح الدكتوراه، ولم يكن هناك أثر للتفاعل بين المتغيرين (الجنس والدرجة العلمية على امتلاك الكفايات وأظهرت النتائج أنَّ هناك تدنياً واضحاً في إتقان طلبة الدراسات العليا في الجامعات الأردنية للكفايات البحثية والإحصائية.

**4. دراسة البنَّاء (2011)**

**الموضوع: بناء اختبار محكي المرجع لقياس الكفايات الإحصائية لدى طلبة الدراسات العليا بكليات التربية في الجامعات اليمنية.**

**أهداف الدراسة:** بناء اختبار محكي المرجع لقياس الكفايات الإحصائية لدى طلبة الدراسات العليا بكليات التربية في الجامعات اليمنية، والتحقق من خصائصه السيكومترية، وتحديد درجة القطع المناسبة

لكل كفاية إحصائية، والتعرف على مدى تمكن طلبة الدراسات العليا بكليات التربية في الجامعات اليمنية من الكفايات الإحصائية التي تضمنها الاختبار الذي تمَّ بناؤه. وتألفت عينة الدراسة من (136) طالباً وطالبة تمَّ اختيارهم من أربع جامعات بها برنامج للماجستير في كلية التربية.

**أداة الدراسة:** اختبار محكي المرجع من إعداد الباحث، وتمَّ التحقق من صدقه باستخدام الصدق الوصفي وصدق انتقاء المجال السلوكي، وتمَّ التحقق من الثبات باستخدام معادلة ألفا كرونباخ ومعامل ليفنجستون كما قام الباحث بتحديد درجة القطع لكل كفاية باستخدام طريقة ندلسكاي.

**نتائج الدراسة:** بينت نتائج الدراسة أنَّ القياس محكي المرجع أكثر تشخيصاً، وتوضيحاً لمدى تمكن الطلبة من الكفايات الإحصائية، وأنَّ الاختبار محكي المرجع الذي تمَّ بناؤه في هذه الدراسة لقياس الكفايات الإحصائية لدى طلبة الدراسات العليا بكلية التربية في الجامعات اليمنية أكثر شمولاً للكفايات الإحصائية مقارنةً مع الاختبارات المعيارية، وأنَّ هذا يتمتع بدرجة جيدة من الصدق الوصفي وصدق الانتقاء السلوكي وكذلك الثبات.

**ثالثاً: دراسات استخدمت نموذج راش.**

**1. دراسة الزاملي(2019)**

**الموضوع: استخدام نموذج راش في تدريج مقياس التفكير الناقد لدراسة بعض العوامل المؤثرة على دقة القياس وفق نظرية الاستجابة للمفردة.**

**أهداف الدراسة:** تهدف هذه الدراسة إلى الكشف عن بعض العوامل المؤثرة (كحجم العينة المستخدمة في التدريج، وأفضل طريقة للربط بين الصور الاختبارية المختلفة وتدريج ميزان تدرج واحد مفردات مشتركة/ أفراد مشتركين، وأفضل عدد للمفردات المشتركة أو للأفراد المشتركين المستخدمين في ربط الصور الاختبارية المختلفة) بالإضافة إلى أثر مستوى العينة المستخدمة في تدريج المفردات. والتي قد تؤثر على دقة القياس، وذلك للتوصل إلى أفضل تدريج يوفره نموذج راش لمقياس القدرة على التفكير الناقد، كما تهدف الدراسة إلى عمل معايير (الرتب المئينية والدرجات التائية) والتي تفسر على أساسها قدرة الأفراد على المقياس واستخدام برنامج الحاسب الآلي Winsteps في تدريج مقياس القدرة على التفكير الناقد وفقا لنموذج راش.

**أداة الدراسة:** استخدم الباحث تدريج مقياس التفكير الناقد وفق نموذج راش، وأجريت الدراسة على

عينة من الأفراد بلغ عددهم (829) طالباً وطالبةً من طلبة المرحلة الابتدائية (رابع وخامس) والمتوسطة (أول وثاني).

**نتائج الدراسة:** أظهرت النتائج البحث أنَّ أفضل حجم من بين أحجام العينة المستخدمة في تدريج مقياس القدرة على التفكير الناقد باستخدام نموذج راش هو حجم(800) وذلك تبعاً لمحاكاة دقة القياس وأنَّ أفضل عدد للمفردات المشتركة والمستخدمة في ربط الصور الاختبارية المختلفة هو العدد(11) وذلك تبعاً لمحاكاة دقة القياس وأنَّ أفضل عدد للأفراد المشتركين والمستخدمين في ربط الصور الاختبارية هو العدد(50) وذلك تبعاً لمحاكاة دقة القياس وأنَّ أفضل طريقة للربط بين الصور الاختبارية هي طريقة المفردات المشتركة وذلك تبعاً لمحاكاة دقة القياس، وتبعاً لمحاكاة دقة القياس فقد اختلفت دقة القياس باختلاف مستوى العينة وذلك لصالح العينة منخفضة المستوى. وتحققت في مقياس القدرة على التفكير الناقد استقلالية القياس (بمعنى دقة التعادل) وتمَّ تقدير صعوبة المفردات مقياس القدرة على التفكير الناقد وفقاً لنموذج راش وذلك بعد توفر محكات الدقة وكذلك تقدير قدرة الأفراد المقابلة لكل درجة كلية خام محتملة على مقياس القدرة على التفكير الناقد وفقاً لنموذج راش بعد توفر محكات الدقة وتمَّ التوصل إلى معايير الرتب المئينية والدرجات التائية التي تُفسَّر على أساسها مستويات قدرة الأفراد على المقياس.

**2. دراسة بوبو (2019)**

**العنوان: استخدام نموذج راش في تدريج اختبار مهارات معلم الصف في تعرف التلامذة ذوي صعوبات التعلم.**

**أهداف الدراسة:** هدفت الدراسة إلى تدريج اختبار يقيس مهارات معلم الصف في التعرف على صعوبات التعلم عند التلاميذ ذوي صعوبات التعلم.

**أداة الدراسة:** اختبار مهارات معلم الصف في تعرف التلاميذ ذوي صعوبات التعلم وفق نموذج راش تكون الاختبار من (60) مفردة، وقد تمَّ تطبيق البحث على (300) من المعلمين.

**نتائج الدراسة:** بينت الدراسة الإحصائية بأنَّ افتراضات النظرية الحديثة في القياس والتقويم قد تحققت كما أظهرت نتائج تدريج الاختبار بعد استبعاد (5) مفردات بأنَّ الاختبار ثابت، وقد بلغ معامل ثبات المفردات (0.97) ومعامل ثبات الأفراد (0.68) كما تبيَّن أنَّ مفردات الاختبار غطت مدى الصعوبة

بشكل جيد فتراوحت قيمة صعوبة المفردات ما بين (2.48 و2.05) لوجيت، وقيم قدرة الأفراد ما بين (-1.62 و2.13) لوجيت. ووقعت قيم جميع المفردات ضمن قيمة احصائي الملاءمة لنموذج راش.

**3. دراسة زياد وبوقصارة (2018)**

**الموضوع: استخدام نموذج راش اللوغاريتمي أحادي البارامتر في تدريج اختبار لمستويات التفكير الهندسي مبني وفق نموذج "فان هيل".**

**أهداف الدراسة:** هدفت الدراسة إلى استخدام نموذج راش في تدريج اختبار محكي المرجع لقياس مستويات التفكير الهندسي مبني وفق نظرية "فان هيل".

**أداة الدراسة:** اختبار مستويات التفكير الهندسي مبنى وفق نموذج راش، تكون الاختبار من (28) مفردة من نوع الاختيار المتعدد. وطبق الاختبار على عينتين استطلاعية وأساسية مكونتين من (681) تلميذاً وتلميذةً من تلاميذ المرحلة الثانوية.

**نتائج الدراسة:** أظهرت نتائج تدريج الاختبار بعد استبعاد (6) مفردات و(11) فرداً، بأنَّ الاختبار ثابت حيث بلغ معامل الثبات المفردات (0.99) ومعامل الثبات للأفراد (0.65) كما أنَّ مفردات الاختبار غطت مدى الصعوبة بشكل جيد، فتراوحت قيمة الصعوبة للمفردات ما بين (-1.46-1.61) لوجيت وقيم قدرة للأفراد ما بين (3.44-4.63) لوجيت، ووقعت قيم جميع المفردات (22) مفردة ضمن قيمة احصائي الملاءمة لنموذج "راش"(MINSQ) بين (0.7-1.3) وقيمة احصائي الملائمة (ZSTD) بين (0.2-، 0.2).

**4. دراسة بوشي (2018)**

**الموضوع: دراسة مقارنة بين النظرية التقليدية ونموذج راش في اختيار فقرات لدى تلاميذ مرحلة التعليم الثانوي (M.M.I.I.2) اختبار مينسوتا المتعدد الأوجه للشخصية.**

**أهداف الدراسة:** جرت هذه الدراسة على تلاميذ مرحلة التعليم الثانوي في الجزائر، وحاولت أنَّ تقدم تصور عن الخصائص السيكومترية لاختبار مينيسوتا متعدد الأوجه للشخصية وفق النظرية الكلاسيكية ووفق نموذج راش، من أجل تحقيق هذا الهدف استخدمت المنهج الوصفي.

**أداة الدراسة:** اختبار مينسوتا المتعدد الأوجه للشخصية، مع الأساليب الإحصائية، مقاييس النزعة المركــزية، مقاييس التشتت، التحليل العاملي الاستكشافي، معاملات الارتباط، اختبار ت، معامل الفا

كرونباخ، إحصاء الملاءمة تقدير الصعوبة، تقدير القدرة.

وقد طبقت الدراسة على عينة مكونة من (300) تلميذ وتلميذة يتراوح أعمارهم ما بين (16-18) سنة للسنة الدراسية 2015 /2016 حيث تمَّ اختيارها بطريقة قصدية، عنقودية متعددة المراحل.

**نتائج الدراسة:** وفق النظرية الكلاسيكية تمَّ التوصل إلى إنَّ قيم المتوسطات الحسابية للعينة في كل المقاييس أكبر قليلاً من قيَّم الوسيط، وأكبر من قيَّم المنوال، ومؤشر على اقتراب التوزيع الدرجات اعتدالياً كما أظهر التحليل العاملي الاستكشافي وجود ثلاث عوامل تتشبع عليها مقاييس الصدق والمقاييس الإكلينيكية للاختبار، كما تمَّ التوصل إلى وجود فقرات ذات ارتباط دال إحصائياً عند مستوى(0.01) وعند(0.05) وهناك فقرات لا ترتبط بالمقاييس، ووجود فروق ذات دلالة إحصائية بين متوسطي درجات الثلث الأعلى ودرجات الثلث الأدنى في المقاييس، وأنَّ الاختبار له قدرة تشخيصية جيدة.

**وفق نموذج راش:** هناك صدق معتدل لمقاييس اختبار مينيسوتا المتعدد الأوجه من خلال ما يوفره نموذج راش من مؤشرات المطابقة أو الملاءمة التقاربية والتباعدية، وأحادية البعد في القياس وتراوح معامل ثبات الفقرات من (0.95) إلى (0.99) ومعامل ثبات الأفراد من (0.6 فأقل 0.6 فما فوق) وهذا يدل أنَّ بعض مقاييس الاختبار ثابتة بدرجة مقبولة.

**5. دراسة صبَّاح (2016).**

**الموضوع: استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في اللغة الانجليزية لطلبة جامعة القدس المفتوحة وفق نموذج راش.**

**أهداف الدراسة:** كان هدف الدراسة هو استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في اللغة الانجليزية لطلبة جامعة القدس المفتوحة، وفق نموذج راش.

**أداة الدراسة:** لتحقيق هدف الدراسة السابق تمَّ بناء اختبار تحصيلي في اللغة الانجليزية مؤلفاً من (45) فقرة من نوع الاختيار من متعدد من أربعة بدائل، يقيس تحصيل الطالب في اللغة الإنجليزية تمَّ حذف (5) فقرات أجمع المحكمون على أنَّ هناك فقرات تؤدي غرضها، فأصبح الاختبار بصورته الأولية مكوناً من (40) فقرة طبق على أفراد عينة الدراسة البالغ عددهم (956) طالباً وطالبة، من طلبة السنة الأولى في جامعة القدس المفتوحة.

**نتائج الدراسة:** ومن أبرز النتائج التي توصلت إليها الدراسة: تحقق افتراضات نظرية الاستجابة للفقرة في بيانات الدراسة، وكذلك مطابقة الاستجابات عن (38) فقرة من فقرات الاختبار الافتراضات نموذج (راش) وحذف فقرتين، لم تطابق النموذج، وحذف (63) فرداً غير مطابقين للنموذج، كما بينت النتائج المتعلقة بتقديرات معامل الصعوبة أنَّها كانت مقبولة ضمن المحكات التي أوردتها أدبيات القياس التربوي. وأسفر التحليل عن اختبار جيد يتمتع بمظاهر متعددة من الصدق وبثبات (0.99) لفقرات الاختبار (0.80) لقدرات الأفراد، كما بينت النتائج أنَّ قيمة دالة المعلومات كانت أقصى ما يمكن عندما كانت($\theta = b$) وذلك كما هو متوقع من نموذج (راش)، وأنَّ الاختبار يقدم أكبر كمية من المعلومات عند مستويات القدرة المتوسطة، ويقدم أقل كمية من المعلومات عند مستويات القدرة العالية والمتدنية.

**6. دراسة رفقي وصبّاح (2016)**

**الموضوع: بناء اختبار تحصيلي في الهندسة لطلبة الصف الخامس الأساسي باستخدام نموذج راش.**

**أهداف الدراسة:** هدفت هذه الدراسة إلى بناء اختبار تحصيلي في الهندسة لطلبة الصف الخامس باستخدام نموذج راش، وتشخيص فهم الطلبة الصف الخامس للمفاهيم الهندسية، فقد تؤدي الأخطاء المفاهيمية في الهندسة لتكوين صورة غير واضحة للفهم الهندسي لدى الطلبة فيما بعد.

**أداة الدراسة:** اختبار تحصيلي في الهندسة مكون من (30) فقرة من نوع الاختيار من متعدد، وطُبق الاختبار على (216) طالباً وطالبة.

**نتائج الدراسة:** وأظهرت النتائج أنَّ الاختبار ثابت، حيث بلغ معامل الثبات للفقرات (0,98) ومعامل الثبات للأفراد (0,78) وأنَّ فقرات الاختبار غطت مدى الصعوبة بشكل جيد، فتراوحت قيمة الصعوبة للفقرات ما بين (-2,69،3,29) لوجيت، وظهر أنَّ معظم الفقرات تقع ضمن إحصائيات الملاءمة لنموذج راش. وأظهرت النتائج أيضاً وجود أخطاء متنوعة في المفاهيم الهندسية منها: خصائص متوازي الأضلاع، مفهوم المحيط، مفهوم الشبكات، والربط بين الأشكال الهندسية، وفي المقابل هناك مجموعة من المفاهيم الهندسية التي أجاب عليها الطلبة بنسبة أكبر من (88%) منها: مفهوم المضلع مفهوم المضلع المنتظم، ومفهوم الزاوية وقياسها باستخدام المنقلة.

**7. دراسة القرشي (2015)**

**الموضوع: استخدام النموذج الأحادي المعلم (راش) في بناء اختبار تحصيلي محكي المرجع في**

**مقرر طرق البحث العلمي.**

**أهداف الدراسة:** هدفت الدراسة الحالية إلى استخدام النموذج الأحادي المعلم(راش) في تصميم اختبار تحصيلي محكي المرجع في مقرر طرق البحث العلمي لطالبات قسم رياض الأطفال بكلية التربية بجامعة الطائف.

**أداة الدراسة:** لتحقيق هذا الهدف تمَّ بناء اختبار محكي المرجع في مقرر طرق البحث العلمي يقيس تحصيل الطلبة والطالبات ويقدم تقويماً موضوعياً للمهارات المطلوب إتقانها مكون من (40) فقرة من نوع الاختيار من متعدد (أربع بدائل) وطبق الاختبار على عينة بلغت (224) طالبة من طالبات قسم رياض الأطفال.

**نتائج الدراسة:** أشارت النتائج على مطابقة الاستجابات عن (40) فقرة وهي جميع فقرات الاختبار بصورته النهائية لافتراضات نموذج راش، وقد بلغ معامل الثبات للأفراد (0.91) أما معامل الثبات للاختبار فبلغ (0.87).

## II. التعقيب (التعليقات) على الدارسات السابقة

**أوجه التقارب والاختلاف بين الدراسة الحالية والدراسات السابقة**

قدمت الدراسات السابقة المتعلقـــة بالدراسة الحالية، أرضية خصبة في توسيع ثقافة الباحث واطلاعه على موضوع الدراسة، حيث تشكلت لدى الباحث نظرة شاملة عن الموضوع، وتكونت لديه مجموعة من الأفكار عن مختلف المناهج، وأنسبها للدراسة وأفضل الخطوات لبناء أداة الدراسة.

**1. من حيث المنهج والعينة المستخدمة في الدراسة**

اتفقت هذه الدراسة مع العديد من الدراسات السابقة، من حيث المنهج المستخدم في الدراسة فأغلب هذه الدراسات استعملت المنهج الوصفي، فبعض الدراسات استخدمت المنهج الوصفي التحليلي ومن هذه الدراسات، دراسة كل من العطوى (2019)، وقارة (2018)، والقرشي (2015) وهو المنهج المطبق في هذه الدراسة، واستخدم المنهج الوصفي المسحي في دراسة كل من الشهري (2016) والزيلعي (2014)، وانفرد الصباح (2016)، وافولابي(2015) باستعمال المنهج الكمي.

كما تباينت العينات المستخدمة في الدراسات السابقة التي تمَّ عرضها حيث طبق بعضها على طلبة الدراسات العليا كما في دراسة كل من العطوي والمسعودي (2019)، والطراونة (2016) والحباشة

(2015)، والزيلعي (2014)، والبناء (2011). وطبقت الدراسة على البعض الأخر من الطلبة الجامعيين كما في دراسة، كل من الصباح (2016)، والقرشي (2015)، والسمرائي والخفاجي (2013). وطبقت بعضها على تلاميذ المرحلة الابتدائية والمتوسطة والثانوية، كما في دراسة زياد وبوقصارة(2018) وبوشي (2018)، وفقي والصباح (2016)، وافولابي (2015)، وقد طبق الباحث في هذه الدراسة على عينة من طلبة الجامعة.

أما حجم العينة في الدراسات السابقة فقد تنوعت، فبعض الدراسات اختارت حجم العينة صغيراً والبعض اختار حجم عينة كبير وفق ما تقتضي الدراسة كما في الدراسات المتعلقة بنموذج راش. وتكونت عينة الدراسة الحالية من طلبة السنة الأولى علوم اجتماعية جذع مشترك والبالغ عددهم (502) طالب وطالبة من جامعة زيان عاشور بالجلفة.

كما اتفقت اغلب الدراسات في استخدام العينة العشوائية البسيطة، ففي دراسة قارة (2018) استخدمت العينة العشوائية متعددة المراحل، وفي دراسة الشهري (2016) عينة بسيطة منتظمة، وفي دراسة الثبيتي (2014) عينة عشوائية عنقودية، وفي دراسة السمرائي والخفاجي (2013) عينة عشوائية طبقية، في حين انفردت دراسة بوشي (2018) باستخدام العينة القصدية، أمَّا في الدراسة الحالية فقد حصر الباحث كل طلبة السنة أولى علوم اجتماعية كعينة دراسة.

**2.من حيث الأداة المستخدمة في الدراسة:**

ومن حيث الاداة المستخدمة في الدراسة، فقد اتفقت كل الدراسات السابقة التي استخدمت الاختبارات المحكية مع الدراسة الحالية في هدف واحد وهو بناء اختبار محكي المرجع، حيث ركزت أغلبها في بناء أدوات الدراسة على الخطوات العلمية المتعارف عليها في بناء الاختبارات محكية المرجع (مرحلة التحليل، مرحلة البناء، مرحلة التجريب)، كما في دراسة كل من العطوي والمسعودي(2019)، وقارة(2018)، وطراونة(2016)، والشهري(2016)، والحباشة(2015) والزيلعي (2014)، فيما استخدم مقياس التفكير الناقد في دراسة الزاملي(2019)، واختبار مستويات التفكير الهندسي في دراسة زياد وبوقصارة، واختبار مينسوتا متعدد الاوجه للشخصية في دراسة بوشي (2018). كما استخدمت الاختبارات التحصيلية في دراسة صباح (2016)، ورفقي وصباح (2016).

ومن حيث صدق الأداة فإنَّ الدراسات التي تناولت بناء اختبار محكي المرجع، استخدمت جميعها

للتحقق من الصدق طريقة الصدق الوصفي وصدق المحتوى، واستخدمت دراسة قارة (2018) إضافة إلى الصدق الوصفي، الصدق التمييزي والصدق البنائي، وانفردت دراسة البناء (2011) باستخدام صدق انتقاء المجال السلوكي. وفي هذه الدراسة تمَّ استخدام الصدق الوصفي والصدق الوظيفي، وصدق البناء).

ومن حيث معامل الثبات فإنَّ الدراسات التي تناولت بناء الاختبار محكي المرجع استخدمت في معظمها للتحقق من الثبات معامل الفا كرونباك ومعامل ليفنجستون، باستثناء بعض الدراسات كما في دراسة قارة (2018)، التي استخدمت إضافة إلى هذه المعاملات، معامل هاريس، واستخدم معامل سبكوفياك في دراسة الزيلعي (2014). في حين الدراسات التي استخدمت نموذج راش في دراستها فقد اعتمدت على معامل الفصل لتحديد الثبات المتعلق بالأفراد والثبات المتعلق بالمفردات كما في دراسة كل من بوبو (2019)، زياد وبوقصارة (2018)، وصباح (2016)، والقرشي (2015).
وتميزت هذه الدراسة باستخدام ثبات الاتساق الداخلي، ومعامل كيودر ريتشاردسون، ومعامل ليفنجستون.

ومن حيث طريقة تحديد درجة القطع فإنَّ معظم الدراسات التي كان هدفها بناء اختبار محكي المرجع استخدمت طريقة انجوف لتحديد درجة القطع، ومن هذه الدراسات، دراسة العطوي والمسعودي (2019)، وقارة (2018)، وحباشة (2015)، والزيلعي (2014)، والثبيتي (2014) في حين استعملت دراسة البناء (2011) طريقة ندلسكاي في تحديد درجة القطع.

**III .موقع الدراسة الحالية من الدراسات السابقة**

إن أوجه الاستفادة من الدراسات السابقة، هو أنَّها بمثابة المصدر الأساسي لكثير من المعلومات المهمة والتي توجهنا في إعداد الدراسة الحالية، وتحديد أسئلة الدراسة وإعداد الإجراءات الملائمة لتحقيق أو عدم تحقيق تساؤلات الدراسة، استناداً إلى تلك الدراسات السابقة، وهذا بالإضافة إلى أنَّ تلك الدراسات وجهت الطالب الباحث نحو العديد من البحوث والدراسات والمراجع المناسبة ومكنته من تكوين تصور شامل عن الأطر النظرية التي ينبغي أن تشملها الدراسة الحالية، وقد خلص الباحث إلى ما يلي:

- ✓ اختلاف أدوات الدراسة المستخدمة في تلك الدراسات، والأساليب الإحصائية المستعملة، أدى هذا

إلى استخدام أساليب إحصائية مناسبة لغرض البحث.

- ✓ تباينت العينات في الدراسات السابقة من حيث المرحلة الدراسية والعمرية، فبعضها كانت من المرحلة الثانوية. والبعض الأخر من تلاميذ المرجلة الابتدائية، وكانت عينة هذه الدراسة من طلبة الجامعة.
- ✓ أجمعت معظم الدراسات السابقة على أن نموذج راش يستند إلى افتراضات قوية تستند بدورها إلى أساس إحصائي سليم، مما يثبت فعالية النموذج لتدريج فقرات الاختبارات والمقاييس النفسية والتربوية وعليه يمكن التحقق من الاختبار التحصيلي في مادة الاحصاء ومدى مطابقته لافتراضات نموذج راش.

# الفصل الثالث: الإطار النظري

**تمهيد**

**المحور الأول: الاختبارات التحصيلية محكية المرجع**

**المحور الثاني: النظرية الكلاسيكية في القياس**

**المحور الثالث: النظرية الحديثة في القياس**

**المحور الرابع: نموذج راش**

**خلاصة الفصل**

**تمهيد**

عرف القياس النفسي والتربوي تطوراً، بفضل جهود علماء النفس الذين درسوا المشكلات النفسية والتربوية المختلفة. ونظراً للتطورات المتوالية التي مرَّ بها القياس سوف نركز باختصار في هذا الفصل على مراحل تقدم نظرية القياس وتطبيقاتها من النظرية الكلاسيكية لظهور نظرية القياس الحديثة، أو ما تعرف بنظرية الاستجابة للمفردة، أو نظرية السمات الكامنة. كما واكب هذا التطور لظهور النظرية الحديثة نوع من الاختبارات أطلق عليه اسم الاختبارات محكية المرجع نتيجة لما وجهه علماء القياس والتقويم من نقد إلى فلسفة النظام المعياري المرجع، مما دعا العلماء والمختصين للتوجه نحو الاتجاه الايديومتري الحديث من أجل تحديد قدرة الفرد الفعلية ومستوى اتقانه بعيداً عن جماعة اقرانه.

وفي هذا الفصل المقسم إلى أربع محاور، نتناول في المحور الأول، الاختبارات التحصيلية محكية المرجع، والمحور الثاني، نظرية القياس الكلاسيكية، والمحور الثالث، نظرية القياس الحديثة ولكون بناء الاختبار التحصيلي محكي المرجع في هذه الدراسة وفق نموذج راش، فقد تمَّ تخصيص المحور الرابع لهذا النموذج.

**المحـــور الأول: الاختبـــارات التحصيليـــة محكيـــة المرجـــع**

تعد الاختبارات التحصيلية الأداة الأكثر استخداماً للتقويم في الأنظمة التربوية لكن في الغالب تعاني جل هذه الاختبارات من ضعف التخطيط أو حتى انعدامه، لذلك وجهت لهذه الاختبارات انتقادات شديدة مما دفع المختصون والمربون يتجهون نحو تطوير أشكال جديدة لها بهدف التخفيف من عيوبها (عبد الرؤف وإيهاب، 2017)، كما تعزز الاتجاه الداعي إلى عدم الاقتصار عليها في عملية التقويم وضرورة استخدام طرق وأدوات جديدة للتقويم تحقق الأهداف الأساسية للتربية الحديثة.

وعلى الرغم من ذلك كله فإنَّ اختبارات التحصيل ما زالت تحتل المكانة الأولى في عملية التقويم وتكتسب مسألة تحسين هذه الاختبارات وتحديثها شكلاً ومضموناً، ورفع فاعليتها أهمية خاصة.

ونظراً لتطور وتزايد الدراسات في الحقل التربوي مكَّن من التحكم أكثر في وسائل قياس التحصيل الدراسي ومحاولة جعلها أكثر دقة وقياساً للمستوى التعليمي للمتعلمين لأنَّها تعد الأسلوب

الشائع في مختلف المؤسسات التعليمية، حيث يستخدمها الاساتذة لقياس قدرات الطلبة التحصيلية وكذا في ترتيب المتعلمين وتحديد النجاح والرسوب.

## 1. الاختبارات التحصيلية

### 1.1. تعريف الاختبار

لقد ظهرت عدة تعريفات للاختبارات باختلاف وتباين المختصين في القياس النفسي والتربوي ويمكن تقديم تعريفات مختلفة للاختبار كما يلي:

- يعرفه صفوت فرج (2007): "أي محك أو عملية يمكن استخدامها تهدف إلى الحصول على ملاحظات أو معلومات مقننة، أو تحديد حقائق معينة أو تحديد معايير الصواب أو الدقة أو الصحة سواء في قضية معروضة للدراسة أو المناقشة، أو التحقق من فرض قائم لم يتم التثبت منه بعد"(ص.99).

- يعرفه محمد محاسنة (2013): "هو إجراء لقياس عينة من السلوكيات من خلال عينة من المثيرات"(ص.55). [أي أنَّه أداة للحصول على معلومات حول سلوك المفحوص لأنَّه لا يمكن التنبؤ بجميع سلوكيات الفرد وبالتالي لا يمكن تحديدها واخضاعها للاختبار].

- يعرف بين Bean (1953) الاختبار بأنَّه: "مجموعة من المثيرات أعدت لتقيس بطريقة كمية أو بطريقة كيفية العمليات العقلية والسمات أو الخصائص النفسية، وقد يكون المثير هنا أسئلة شفاهية أو أسئلة كتابية أو تكون سلسلة من الأعداد أو الأشكال الهندسية أو النغمات الموسيقية أو صوراً أو رسومات وهذه كلها مثيرات تؤثر على الفرد وتثير استجابته"(حمدي، 2013، ص.15).

### 2.1. تعريف الاختبار التحصيلي *Achievement Test*

يقصد بالاختبار التحصيلي الأداة التي تستخدم في قياس المعرفة والفهم والمهارة في مادة تدريسية أو تدريبية أو حتى مجموعة من المواد (أبو حطب وأخرون، 2008، ص.467).

ويهدف اختبار التحصيل إلى قياس مدى تحصيل الطالب من حيث الحفظ والفهم والتطبيق والتقييم والمقارنة وهذه الجوانب ترتبط ارتباطاً وثيقاً بموقع الاختبار من التدريس وهنا يشترط أن يكون الاختبار واسع التمثيل للمادة الدراسية ليكون مقياساً صادقاُ، ثابتاً (الخولي، 1998).

- يعرفه الزغلول (2012): "هو عينة من الاسئلة أو المهام التعليمية المصاغة على نحو يمكن معه

معرفة قياس مدى تحقق الأهداف المحددة مسبقاً لمادة تعليمية معينة أو مهارة ما" (ص.331).

- يعرفه عباس (1996): "إنَّ اختبارات التحصيل تشمل على امتحانات في مواد معينة من مواد المنهج الدراسي، أي تعتمد على تحديد الأعمار التحصيلية لكل مادة، أو تحديد المستويات التحصيلية بمستويات الذكاء المختلفة. فهي تهدف إلى تقرير ما حصَّله الفرد من منهج دراسي معين (ص.12-13).

- يعرفه عطوان وأبو شعبان (2019): "هو عينة من السلوك المراد قياسه لمعرفة درجة امتلاك الفرد من هذا السلوك، وذلك من أجل الحكم على مستوى تحصيله من خلال مقارنة أدائه بتحصيل زملائه" (ص.39).

### 3.1. الأسس النظرية للاختبارات التحصيلية

جميع الاختبارات التحصيلية وغيرها تحكمها أسس قياسية وتوجهها في جميع مراحل البناء مروراً بعملية تطبيقها، ومن ثمَّ انتهاء بعملية تفسير نتائجها وتحليلها. فهذا النوع من الاختبارات تؤثر فيه نظريتان أساسيتان هما:

- **النظرية الأولى:** تقوم على أساس أنَّ أسئلة وفقرات الاختبار المتضمنة في الاختبار لا بد أن تكون في مجال كبير وواسع من البنود وهذا يفرض التماثل التام بين الأسئلة وبين البرنامج الدراسي أي أنَّ المعلومات المطلوبة أو أسئلة الفقرات أو البنود لا بدَّ أن تتوفر في المحتوى الدراسي. لكن هذه النظرية من الصعوبة بمكان تطبيقها بصورة عملية لأنَّه يصعب الحصول على أسئلة ممثلة للمجال الكبير الذي تمَّ وضعه للبرنامج الدراسي (الطريري، 1997، ص.283).

- **النظرية الثانية:** تهتم هذه النظرية في بناء الاختبارات التحصيلية بالعلاقة بين الأداء الحقيقي للفرد في سلوكه وتصرفاته وكل ما يصدر عنه من أفعال ذات علاقة بالموضوع، وأدائه في المادة الدراسية التحصيلية، فهذه النظرية تقوم على مفهوم نظرية المنحى المميز للفقرة بمعنى أنَّ أداء الفرد وإجاباته عن الفقرة الاختبارية يعبر عن نتيجة حقيقية لمستوى الفرد في الصفة أو الخاصية المقاسة.

لكن المشكل الذي يمكن توجيهه لهذه النظرية هو صعوبة تقدير العلاقة بين احتمالية الإجابة الصحيحة على الفقرة والتحصيل بشكل عام. حيث تظهر هذه العلاقة في المنحى المميز للفقرة، ومن أجل تقدير هذه العلاقة لا بدَّ من تحقيق فرضية الاستقلال الموضعي (Local Independence) والذي

يعني عدم تأثر أداء الفرد على المفردة في استجاباته بأي مفردة أخرى بالاختبار (الطريري،1997، ص.283-284).

### 4.1. مبادئ الاختبارات التحصيلية

يؤدي العمل بهذه المبادئ والأسس إلى تحسين هذا النوع من الاختبارات كما أكد عليها كثير من علماء القياس أمثال (جرونلاند ومهراز وانستازي) وغيرهم (عبد الرؤوف وإيهاب، 2017).

- يجب أن تأخذ عملية بناء الاختبارات بالحسبان استعمالاته والغرض الخاص منه سواء كان غرض الاختبار هو تقويم التلاميذ أو الغرض الكشف عن صعوبات التعلم أو قياس التمكين أو الاتقان لوحدة دراسية.

- إنَّ طبيعة الأهداف ونواتج التعلم المقاسة، هي من تحدد أشكال البنود حيث أنَّ كل نوع أو شكل من أشكال البنود الاختبارية قد يكون أكثر ملاءمة لقياس نواتج تعليمية معينة وأقل ملاءمة لقياس نواتج أخرى، فالاختبار المقالي مثلاً أكثر فعالية في قياس القدرة على الربط وتنظيم الأفكار والاختبار الموضوعي من نوع التكميل أكثر ملاءمة من سواه لقياس القدرة على استدعاء الحقائق الخاصة، كما أنَّ الاختبار من متعدد يصلح لقياس الكثير من نواتج التعلم ومستوياته الدنيا منها والعليا وهكذا.....

- تؤخذ عينة من بنود الاختبار كإجابة وتعتبر بمثابة دليل على التحصيل في مجال معين، ويفترض عادة أنَّ إجابة التلميذ عن تلك البنود تمثل إجابة محتملة عن كل البنود والأسئلة المحتملة، ويعتبر جدول المواصفات إحدى أهم الوسائل في الحصول على عينة ممثلة للمحتوى الدراسي ونواتج التعلم المراد قياسها.

- يجب أن تكون البنود الاختبارية ملائمة من حيث الصعوبة فمن المعلوم أنَّ المستوى الملائم للصعوبة في اختبارات الإنجاز التي تنتمي عادةً إلى فئة الاختبارات معيارية المرجع هو (50%) والبنود التي يصل معامل سهولتها إلى (0.50) تتمتع بقدرة تميزية عالية وتحقق مستوى عالٍ من الثبات.

- يجب العناية بوضع البنود الاختبارية حتى لا تؤثر العوامل الأخرى في أداء المفحوص وتمنعه من إظهار مستواه الحقيقي في التحصيل، فيجب الابتعاد عن الجمل الطويلة المعقدة، والمفردات الصعبة والكلمات التي تحمل أكثر من معنى، وتجنب الايماءات والتلميحات، وكل ما يعيق الكشف عن الفروق الحقيقية في أداء المفحوصين.

- يجب وضع الاختبار بحيث يسهم في تحسين العملية التعليمية التعلمية لأنَّ الغرض النهائي للقياس هو تحسين تعلم التلاميذ، والاختبار الجيد يجب أن يرفع نوعية التعلم وكميته، ويساعد المعلم في تحسين أدائه التعليمي.

**5.1. افتراضات أساسية يرتكز عليها قياس التحصيل**

يرتكز قياس التحصيل على بعض الافتراضات الأساسية كما أشار إليها علاَّم حتى تكون نتائج قياس التحصيل متسقة، وصادقة وغير متحيزة (علاَّم، 2006، ص.309–310). ومن هذه الافتراضات:

- أن يقيس الاختبار نطاقاً سلوكياً يمكن تحديده بدقة، فالاختبار التحصيلي سواء أعده المعلم لطلبته أو أعدته جهات أخرى ينبغي أن يستند إلى أهداف تربوية يمكن صياغتها صياغة سلوكية محددة بحيث يمكن قياسها.

- أن يقيس الاختبار الأهداف المتعلقة بالمادة الدراسية، أو محتوى البرنامج التدريسي ولا يقيس أهدافاً عارضة أو غير مهمة وهذا يعني أن يكون اختبار التحصيل صادقاً في محتواه.

- أن تكون المعالجة التعليمية مناسبة للأفراد الذين سيطبق عليهم الاختبار بحيث يراعي الفروق الفردية بينهم، وأن تُتاح لهم الفرصة لتعلم المجال الذي يقيسه الاختبار.

**2. الاختبارات محكية المرجع**

**1.2. نشأت الاختبارات محكية المرجع**

بيَّنت العديد من الدراسات أنَّ الاختبارات المستخدمة وفق نظرية القياس التقليدية، تفتقر إلى الموضوعية بسبب تأثرها بخصائص أفراد العينة التي تطبق عليها. ولأنَّ موضوعية نتائج الاختبارات وصدقها تعتمد على دقة الأساليب التي تستخدم في بنائها، واختيار فقراتها، وتفسير نتائجها، ووصفها للقدرة التي يقيسها الاختبار فإنَّه من الضروري استخدام التوجهات الحديثة في القياس، والتي أثبتت البحوث والدراسات التجريبية أنَّها تحقق الموضوعية والدقة في قياس قدرة الفرد.

لذلك حظيت الاختبارات محكية المرجع(CRT) باهتمام كبير من قبل الباحثين والمختصين في مجال القياس النفسي والتربوي.

ويشير العديد من المختصين في القياس التربوي أنَّ من أكثر الاكتشافات فائدة في مجال القياس التربوي خلال عشرين سنة الماضية هو الاختبارات محكية المرجع.

ويعود مصطلح (محكية المرجع) إلى موضوع كتبه العالم الامريكي روبرت جليزر (Robert Glaser) سنة 1963 بعنوان بعض التساؤلات عن تكنولوجيا التعليم وقياس نواتج التعلم(Ebel, 1980) وقد أثار هذا المقال كثيراً من الجدل بين علماء القياس، لكنَّه لم يُحدث نشاطاً ملحوظاً نحو ما نادى به جليزر.

وتعتبر البداية الحقيقية لهذه الاختبارات مع بداية السبعينات على يد كل من بوفام(Popham) وهيوسك(Husek) اللذان حددا استراتيجيات القياس محكي المرجع، الأمر الذي أدى إلى زيادة اهتمام علماء القياس بهذا النوع من الاختبارات، وإلى تحويل القياس محكي المرجع إلى واقع عملي، حيث قام(بوفام) بنشر أول كتاب يجمع فيه كل المقالات والدراسات المتعلقة بالقياس محكي المرجع سنة (1971) مما ترتب عنه حركة واسعة تهتم بدراسة النواحي النظرية والسيكومترية الخاصة بهذا المدخل(إبراهيم، 1991).

### 2.2. تعريف الاختبارات محكية المرجع *Criterion Referenced Test*

يوجد في الأدب التربوي العديد من التعريفات للاختبارات محكية المرجع وهذا يعود لحداثة هذا المفهوم وكذلك للاختلاف حول تحيد مفهوم المحك الذي تنسب إليه درجات الأفراد، ويعد (روبرت جليزر) أول من استخدم مصطلح الاختبار محكي المرجع، ثمَّ أضاف الكثير من الباحثين رؤى حول هذا المفهوم الأمر الذي نتج عنه العديد من الدراسات حول الاستخدامات في النواحي الفنية المرتبطة بمفهوم الاختبارات محكية المرجع ومن بين التعريفات نذكر ما يلي:

- يعرفه بوفام(popham) "هو ذلك الاختبار الذي يستخدم في تقدير أداء الفرد في نطاق سلوكي محدداً تحديداً دقيقاً" (علام،2001، ص.24-25)، أي أنَّ التحديد الدقيق للنطاق السلوكي، هو الركيزة الأساسية لمفهوم المحك من وجهة نظره.

- يعرفه هامبلتون وسواميناثان (Hambleton & Swaminathan,2013) هو الاختبار الذي يستخدم في تقدير أداء الفرد في نطاق سلوكي محدد تحديداً دقيقاً ويمكن أن يكون الاختبار محكي المرجع اختباراً انتقائياً يتألف من عينة من فقرات مأخوذة من مجال سلوكي محدد، ولكن يمكن أن تحدد له نقطة قطع يُميز عندها بين متقنين وغير متقنين(p.61) .

### 3.2. أوجه الاختلاف بين الاختبارات محكية المرجع ومعيارية المرجع

إنَّ الاهتمام بالمقاييس والاختبارات التربوية المعاصرة جعل المهتمين يميزون اختلافات بين الاتجاهين الرئيسين، فالاختبارات معيارية المرجع تختلف عن الاختبارات محكية المرجع في الافتراضات التي يقوم عليها كل اتجاه، وكذلك الاختلاف في الإطار المرجعي الذي يُعتمد عليه في تفسير النتائج.

وقد أشارت سوسن مجيد (2014) إلى أنَّ كثير من علماء القياس التربوي أمثال (جليزر، إيبل، بوفام، وهيوسك) أشاروا إلى وجود الكثير من الاختلافات حول الاختبارات معيارية المرجع ومحكية المرجع في عدة نواحي يمكن اجازها فيما يلي:(ص.178–179)

- **الهدف الرئيس للاختبار:** يهدف الاختبار محكي المرجع إلى مقارنة أداء الطالب بمحك تمَّ تحديده مسبقاً وذلك لمعرفة الأهداف المحققة وغير المحققة لدى الطالب، في حين الاختبار معياري المرجع يهدف إلى مقارنة أداء الطالب بأداء مجموعته المعيارية(أقرانه) من خلال تحديد عدد الأسئلة التي أجاب عليها الطالب بشكل صحيح.
- **استخدام الاختبار:** يستخدم الاختبار محكي المرجع للتعرف على مدى تحقيق المتعلم للأهداف السلوكية المحددة، أما الاختبار معياري المرجع فيعمل على تزويدنا بمؤشرات عن درجة نجاح الطالب، أي إظهار الفروق الفردية بين الطلبة.
- **خصائص أسئلة الاختبار:** في الاختبار محكي المرجع الأسئلة تكون مجمعة حول عدد محدد من الأهداف، بينما تنتشر الأسئلة في الاختبار معياري المرجع بشكل واسع حول نطاق الأهداف.
- **تفسير الأداء:** تفسر الدرجة في الاختبار محكي المرجع بناءً على درجة القطع، بينما تفسر درجة الطالب في الاختبار معياري المرجع بناءً على درجات معيارية تمثل موقع الطالب بين مجموعته الصفية والرتب المئينية.
- **بناء الفقرات:** يعتمد بناء الفقرات في الاختبار محكي المرجع على مدى تحقيق المفردة للهدف السلـــوكي التعليمي، بينما يعتمد بناءها في الاختبـــار معياري المرجع على تباين الدرجـــات ومعامل التمييز، ويُتجنب اختيار المفردات الصعبة جداً والسهلة جداً.
- **التقويم:** يهدف الاختبار محكي المرجع إلى تحديد جوانب الضعف والقوة في أداء الطلبة والبرامج

التعليمية، في حين يسعى الاختبار معياري المرجع إلى تقويم أداء الطالب ومقارنته مع أقرانه.

### 4.2. استخدامات الاختبارات محكية المرجع

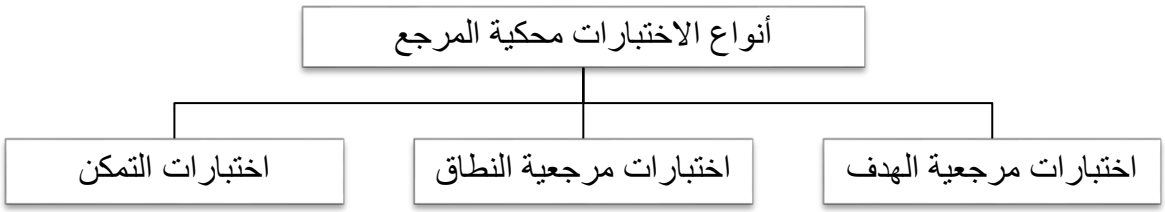
يستخدم هذا النوع من الاختبارات من أجل الكشف عن قدرة الطالب ومدى استعداده للإنجاز في إحدى مجالات الدرس والتحصيل، بحيث يمكن تفسير هذه المعلومات في ضوء محك مطلق دون الرجوع إلى مقارنة الأفراد، ويبنى هذا المحك عادة على الخبرة والمعرفة بخصائص المقرر أو الجال المراد تعلمه (البكري وعجوز، 2011، ص.252).

وقد حدد الخطيب استعمالات الاختبارات محكية المرجع فيما يلي (الخطيب، 1993)

- استخدامها يساعد في اتخاذ القرارات التي بدورها تساعد في تقييم البرامج التعليمية فيما إذا كانت ناجحة أم لا.
- تساعد في تشخيص الصعوبات التعليمية وتخطيط للعملية التعليمية التعلم.
- تساعد في عملية التوجيه والارشاد وفي الأعمال التي تتطلب مهارات خاصة.
- تساعد في تحديد درجة تمكن الطالب من اكتساب كفايات ومهارات معينة.

### 5.2. تصنيف الاختبارات محكية المرجع

إنَّ المتتبع للمراجع والأدبيات المتخصصة في القياس والتقويم التربوي، وخاصة ما تعلق منها بالاختبارات محكية المرجع، يلاحظ أنَّ هناك ثلاث أنواع من الاختبارات المحكية اعتماداً على درجة تحديد النطاق السلوكي الذي يستند إليه بناء كل منها المتمثلة في الشكل الموالي:




**شكل رقم (01) يوضح أنواع الاختبارات محكية المرجع**

**أ) الاختبارات مرجعية الهدف *Objective-Referenced Tests***

هي اختبارات تُبنى وفق مجموعة من الأهداف التعليمية المُصاغة صياغةً سلوكيةً بحيث تكون هناك مزاوجة بين مفردات الاختبار، غير أنَّ النطـــاق السلوكي الذي تمثله هذه الأهداف لا يكـــون محدداً

وبالتالي فإنَّ المفردات التي يشتمل عليها الاختبار تكون قليلة العدد نسبياً، لأنَّها لا تمثل النطاق الشامل للمفردات الممكن، الذي يقيس مجموعة الأهداف، وتطبق هذه الاختبارات عادة عند الانتهاء من وحدة تعليمية، بهدف تصنيف المختبرين في مجموعتين إحداهما حققت الأهداف والأخرى لم تحققها في ضوء نسبة مئوية محددة من المفردات التي ينبغي الإجابة عنها إجابة صحيحة والتعرف على الأهداف التي لم يستطع كل منهم تحقيقها (علام، 1995، ص.24).

**ب)الاختبارات مرجعية النطاق *Domain-Referenced Tests***

في هذا النوع من الاختبارات يتم بناء مفرداته بتحديد نطاق سلوكي شامل من المهام، أو المهارات تحديداً واضحاً دقيقاً، وتُنتقى المفردات التي يشتمل عليها الاختبار من هذا النطاق انتقاءً عشوائياً أو بطرق معاينات عشوائية طبقية، وتستخدم درجات هذه الاختبارات في الحصول على تقديرات إحصائية لاحتمال إجابة فرد أو مجموعة من الأفراد عن مفردات النطاق الشامل الذي تمثله الإجابة الصحيحة على مفردات الاختبار في وقت معين، وهذا يساعد في تعميم نتائج الاختبار على النطاق الشامل (علام، 2001، ص.25).

ويتميز هذا النوع بارتفاع صدق المحتوى حيث يتم تقسيم المقرر إلى مجالات محددة ثمَّ اختبار مفردات تمثل هذه المجالات تمثيلاً منظماً عن طريق تحكيم المفردات المختارة بواسطة خبراء في مجال التخصص للتعرف على مدى اتفاق كل مفردة مع المجال الذي تنتمي إليه، بعد ذلك يتم توزيع المفردات على مجالات الاختبار في صورته النهائية (الزيلعي، 2015، ص.16).

**ج) اختبارات التمكن أو الإتقان *Mastery Tests***

هذه الاختبارات تحدد مدى اكتساب فرد معين للسلوك الذي يهدف البرنامج التعليمي أو التجريبي إلى تنميته لديه، فهذا النوع من الاختبارات يساعد في اتخاذ قرارات تتعلق بإتقان الفرد لهدف تعليمي أو مهارة معينة أو نطاق من المهارات المحددة، ولذلك يمكن أن يكون هذا الاختبار مرجعي النطاق للأداء الذي يقسه الاختبار (علام، 2001، ص.25).

### 6.2. أهمية الاختبارات محكية المرجع

تكمن أهمية الاختبارات محكية المرجع كما أشار إليها كل من (علام، 2010; الشرقاوي وآخرون، 1996) كما يلي:

– توجيه السياسيات التربوية نحو الاهتمام بأداء الطلبة وسلوكياتهم، ومقارنة هذا الأداء بمحكات محددة وثابتة.

- تقارن نتائج الطلبة في هذه الاختبارات بمحكات واضحة ومحددة ومرتبطة بالأداء.
- يمكن الاستفادة من هذه الاختبارات في الحصول على معلومات وبيانات لتقويم المناهج الدراسية وتحديد الكفايات اللازمة والتي نريد من الطالب اتقانها.
- تفيد نتائج هذه الاختبارات في تحديد وتشخيص حالات الضعف والقصور في بعض المهارات.

**7.2. عيوب استخدام الاختبارات محكية المرجع**

أشار عبابنة (2009) إلى عدد من العيوب التي تواجه استخدام الاختبارات محكية المرجع وتطبيقها ومنها:

- مشكلة تحديد المحكات، إذ أنَّ السمات يصعب الاتفاق فيها على مستوى أداء واحد.
- صعوبة استخراج الخصائص السيكومترية لهذه الاختبارات مثل الصدق والثبات.
- حاجة هذه الاختبارات لأشخاص مدربين لإعدادها.
- صعوبة اختيار عينة سلوكية ممثلة من المجال السلوكي الذي يحدد السمة المراد قياسها.

**3. خطوات بناء الاختبارات محكية المرجع**

**1.3. تحديد الكفايات أو النطاق السلوكي للمحتوى التعليمي**

يقصد بالكفاية مجموعة متكاملة من المعارف والمهارات الوظيفية المحددة تحديدا دقيقاً والمرتبطة بمجال تعليمي أو تدريبي معين بحيث يمكن تحقيقها وقياسها من خلال البرامج فردية أو جماعية ويمكن إجراء ذلك عن طريق:

- الاستعانة بمجموعة من خبراء المادة الدراسية بحيث يحددون الكفايات أو المهارات الرئيسية التي يرون أهمية تحققها لدى المتعلمين.
- التحليل المتعمق لمحتوى المنهج الدراسي أو المجال التدريبي.
- إجراء دراسة التقدير احتياجات الفئة المستهدفة في المجال التدريبي يمكن التوصل إلى المهارات والكفايات اللازم إتقانها في مهنة معينة عن طريق تحليل العمل.

### 2.3. تحليل الكفايات أو المهارات الرئيسية إلى مكوناتها

في هذه الخطوة يتم تحليل المهارة الرئيسية تحليلاً إجرائياً أو بنائياً ثمَّ ترتيب مكوناتها ترتيباً منطقياً بإحدى طرق التحليل البنائي للمهارات الرئيسية الأتية:

- **طريقة تحليل الإجراءات *Procedural Approach*:** تتميز هذه الطريقة في تحليل الكفايات بأنَّ إتقان كل خطوة يعتبر مكملاً للخطوة التالية بمثابة سلسلة خطية متتابعة من الخطوات المستقلة والتي تسهم في النهاية إلى تحقيق الكفاية الرئيسية المراد قياسها.
- **طريقة التحليل الهرمي *Hierarchical Approach*:** تستخدم هذه الطريقة عندما يكون من غير الممكن تعلم الكفاية الرئيسية دون تعلم مهارات فرعية تسبقها وهنا يتم تحديد المعارف والمهارات التي ينبغي تعلمها بترتيب هرمي.
- **طريقة تجمع بين تحليل الإجراءات والتحليل الهرمي:** تستخدم هذه الطريقة عندما يكون هناك كفايات معقدة تحتاج إلى تحليل إجرائي ذو خطوات مستقلة متتابعة تتضمن أحكاماً مشددة، وتحليلاً هرمياً يتضمن شبكة من المعارف والمهارات المساعدة والتي تؤدي في النهاية إلى تحقيق تلك الكفايات المعقدة.
- **العلاقات العشوائية بين المهارات:** في هذا النمط تنظم بشكل عشوائي أو غير متسق بين المهارات ويرجع ذلك لعدم ارتباط المهارات بعضها ببعض أو لعدم الاهتمام بتصميم البرنامج التعليمي.

### 3.3. صياغة الأهداف السلوكية للمهارات الرئيسية

هذه النواتج تتضمن المعارف والمهارات المساعدة التي تسهم في تحققها أو اكتسابها إذا استخدمت طريقة التحليل البنائي الهرمي، أين يتم إعادة صياغة هذه النواتج صياغة إجرائية أو سلوكية يمكن ملاحظتها ملاحظةً مباشرةً وقياسها بمفردات اختبارية. فعند بناء الاختبار التشخيصي نعتمد على الأهداف السلوكية التي تحدد نواتج التحليل التعليمي تحديداً واضحاً.

لذلك يجب أن تصاغ الأهداف السلوكية المتعلقة بإجراءات أو مكونات الكفايات أو المهارات الرئيسة التي تمَّ تحليلها بعناية، ولهذه الخطوة عدة أساليب نذكر أهمها:

- **الأهداف المبكرة *(Amplified objectives)*:** اقترح بوفام (Popham) أن يصاغ الهدف السلوكي مبدئياً بالطريقة المعتادة وهي اشتمالها على الفعل السلوكي، ثمَّ المحتوى المرجعي ثمَّ شروط الأداء ثمَّ

مستوى الأداء ويلي ذلك توضيحاً تفصيلياً لكيفية بناء المفردة الاختبارية يتضمن مثال للمفردة وتحديد لنوع المفردة ومكوناته، كذلك محكات الاجابة الصحيحة، وبذلك تصبح الأهداف السلوكية الموجزة أهدافاً مبكرة.

- **المواصفات التفصيلية للاختبار *(Test Specifications)*:** يرى بوفام(Popham) وبيرك(Berk) وميلمان(Millman) أنَّ الاختبار التشخيصي مرجعي النطاق لا تظهر فائدته إلا بوجود مواصفات تفصيلية للاختبار تشتمل على خمسة عناصر تتضمن معلومات محددة ومصاغه بشكل واضح، وهذه العناصر هي:

✓ **وصف عام موجز لما يقيسه الاختبار:** ملخص موجز لفئة السلوك التي سوف تحددها العناصر الباقية بالتفصيل.

✓**مثال لمفردة اختبارية:** يوضح شكل ونوع المفردة الاختبارية التي تبني استرشاداً لمجموعات العناصر التي تليها.

✓**خصائص المثيرات:** التي ينبغي أن تشتمل عليها المفردة الاختبارية التي يستجيب لها المتعلم بطريقة معينة.

✓**خصائص الاستجابات:** يرتبط هذا العنصر بخصائص المثيرات ارتباطاً وثيقاً، نظراً لأنَّ هناك نمطين من أنماط الاستجابات هما الاستجابة المقيدة والاستجابة الحرة أو المفتوحة فإنَّ كلاً منها يتطلب تحديداً لخصائصها التي ينبغي أن تتوفر في كل منها.

✓**ملحق المواصفات:** يستفاد من هذا العنصر عندما تكون خصائص المثير والاستجابة تحتاج تفصيلات عديدة فتكتب هذه التفاصيل في ملحق المواصفات.

### 4.3. بناء المفردات الاختبارية

تتطلب هذه الخطوة دراية تامة من جانب معد الاختبار بطرق إنشاء أنسب أنواع المفردات التي تقيس الأهداف السلوكية المحددة قياساً مباشراً، كما تتطلب التمكن من محتوى البرنامج التعليمي أو التدريبي المعين وفهم خصائص المتعلمين. وذلك لأنَّ هذه المفردات تستخدم في التمييز بين الذين استطاعوا تحقيق الأهداف المحددة والذين واجهتهم صعوبات، وتشخيص أخطاء وفجوات التعلم كذلك يجب على الباحث الأخذ بالاعتبار عدد المفردات الاختبارية التي يمكنها أن تمثل النطاق السلوكي

بشكل تام دون نقص يؤدي إلى خلل في الاختبار التشخيصي أو زيادة تؤدي إلى ملل من قبل المختبرين بحيث تبنى المفردة على أساس مواصفات النطاق السلوكي، وأن يكون مستوى صعوبة كل مفردة مناسباً لصعوبة الهدف، وتكون عينة المفردات ممثلة للنطاق السلوكي، بالإضافة إلى مراعاة الأصول الفنية واللغوية في صياغة المفردات(علام ،1986).

### 5.3. تحليل مفردات الاختبار

توجد عدة طرق لتحليل مفردات الاختبارات محكية المرجع نذكر منها:

#### 1.5.3. طريقة التحقق من تجانس المفردات

تعتمد هذه الطريقة على استخدام مجموعة من الخبراء والمختصين في مجال الاختبارات ليقدِّروا العلاقة بين مجموعة مفردات الاختبار والأهداف التي تقيسها، ويتم تقدير معامل تجانس المفردة بإعداد قائمة بالأهداف والمفردات التي تقيسها، ثمَّ تقدم هذه القائمة للمحكمين، ويطلب من كل واحد أن يقوم بمهمتين هما:

- أن يعطى تقدير (+1) للمفردة إذا كان متأكد من أنَّ المفردة تناسب الهدف الموضوع لقياسه وأن يعطى تقدير(صفر) للمفردة إذا كان غير متأكد من ذلك، وأن يعطى تقدير (-1) للمفردة إذا كان متأكد من أنَّها غير مناسبة لقياس الهدف.
- أن يعطي تقدير (+1) للمفردة إذا كان متأكد من أنَّ المفردة لا تقيس أياً من الأهداف الأخرى وأن يعطي تقدير(صفر) للمفردة إذا كان غير متأكد من ذلك، وأن يعطي تقدير (-1) للمفردة إذا كان متأكد من أنَّ المفردة تقيس أكثر من هدف سلوكي.

ثمَّ يتم تكوين مصفوفة لكل مفردة تمثل صفوفها عدد الخبراء$(n)$، وتمثل أعمدتها مجموعات المواصفات(N)، وتمثل خلايا هذه المصفوفة تقدير$(j)$ من المحكمين للمفردة على أنها تتفق مع مجموعة المواصفات $(i)$، ثمَّ يتم تطبيق المعادلة التالية:

$$L_{ik} = \frac{(N-1)\sum_{j=1}^{n} Xijk - \sum_{k}^{n} N \sum_{j=1}^{n} Xijk + \sum_{j=1}^{n} Xijk}{2[N-1]} \dots\dots\dots (01)$$

حيث:

$L_{ik}$ : يرمز إلى معامل تجانس المفردة(i) مع الهدف(k)

$N$ : عدد مجموعات المواصفات$(\mathrm{i} = 1,2,\dots,\mathrm{N})$

$n$ : عدد المحكمين$(j = 1,2, \dots, n)$

X : ترمز لأحد التقديرات (+1، صفر، −1) للمفردة$(k)$ على أنها تتفق مجموعة المواصفات(i) بواسطة المحكم(j).

ويؤخذ على هذه الطريقة أنَّها تتطلب وقتاً وجهداً كبيرين من المحكمين، وبخاصة عندما يكون عدد المفردات كبيراً، كما أنَّها تناسب الاختبار ذو المحتوى المتجانس(Hambleton, 1984, p209 ، كما ورد في شكري، 2006).

### 2.5.3. طريقة التمايز اللغوي

في هذه الطريقة يتم إعداد قائمتين إحداهما تحتوي على مفردات الاختبار، والأخرى تحتوي على الأهداف، وتقدم القائمتان إلى المحكمين بالإضافة إلى استمارة مدون عليها ميزان يشتمل على درجات متفاوتة لمدى مناسبة المفردة لقياس الهدف، ويتكوَّن هذا الميزان من خمسة بدائل هي :(ممتازة، جيدة جداً، جيدة، ضعيفة، ضعيفة جداً)، ويُطلب من المحكمين اختيار البديل المناسب لكل مفردة ثمَّ يتم حساب متوسط درجات الميزان التي وضعها المحكمون بالنسبة لكل مفردة، ويستخدم هذا المتوسط كمؤشر على مدى اتفاق المحكمين حول مناسبة المفردة لقياس الهدف (Hambleton, 1984, p. 210−226 كما ورد في شكري، 2006).

### 3.5.3. طريقة المزاوجة

في هذه الطريقة يتم إعداد قائمتين إحداهما تحتوي على مفردات الاختبار، والأخرى تحتوي على المواصفات التفصيلية للاختبار أو الأهداف السلوكية، وتقدم القائمتان إلى المحكمين ويطلب منهم أن يقرأون المواصفات التفصيلية للاختبار بعناية، وكذلك المفردات، ثمَّ المزاوجة بينهما بأن يقوم المحكم بكتابة رقم المفردة أمام المواصفات التي تتعلق بها، وبعد إجراء عملية التحكيم يتم تكوين جدول اقتران بين المفردات والمواصفات المتعلقة بها، وتضم خلايا الجدول عدد الخبراء الذين أبدوا موافقتهم على أنَّ المفردة تتعلق بتلك المواصفات، ويستخدم المقياس الإحصائي (كاي تربيع) في التحقق من استقلالية الآراء، كما يمكن فحص جدول الاقتران فحصاً عينياً ، للتعرف على المفردات التي حازت نسبة مرتفعة من الاتفاق، والمفردة التي تتطلب مراجعة أو استبعاد(Hambleton, 1984, p. 211 ، كما ورد في شكري، 2006).

أما تحيز محتوى المفردة فيقصد به أن تتأثر الإجابة على المفردة بالنوع أو المهنة أو العرق

ولذلك يجب مراجعة لغة المفردات بحيث يتم حذف الألفاظ التي تؤدي إلى هذا.

### 6.3. تطبيق الاختبار على العينة الاستطلاعية

نظراً لأنَّ الاختبارات محكية المرجع تستخدم في كثير من الأحيان بغرض التعرف على الطلبة الذين أتقنوا مجموعة من الأهداف والطلبة الذين لم يتقنوا هذه الأهداف، لذلك فإنَّنا نهتم باختيار مجموعة من الطلبة الذين يستطيعون أن يجيبوا على مفردات الاختبار إجابة صحيحة، أي مجموعة متقنة ومجموعة أخرى لا تستطيع الإجابة عن مفردات الاختبار إجابةً صحيحةً، أي مجموعة غير متقنة. وبهذا يمكن استخدام هاتين المجموعين من الطلبة في إجراء التحليل للمفردات.

### 7.3. تحديد درجة القطع للاختبار

يعرف هامبلتون درجة القطع على أنَّها نقطة على متصل درجات الاختبار تستخدم لتصنيف الطلبة إلى فئتين متقنين وغير متقنين والتي تعكس مستويات الأداء المختلفة بالنسبة لهدف معين المراد قياسه في الاختبار، وتفسر نتائج المفحوصين في الاختبارات محكية المرجع وفق درجة القطع(مستوى الأداء) أو المحك الذي يحدد مسبقاً كمرجع لمقارنة أداء المفحوصين فيه(علام،2010; عبابنة،2009) وهذا يتطلب التحقق من صدق القرارات المتعلقة بدرجة القطع وتبرير اختيارها تبريراً علمياً منطقياً، وهناك العديد من الطرق التي يمكن الاستعانة بها في تحديد درجة القطع للاختبارات محكية المرجع، منها:

#### أ). الطرق المتمركزة حول الاختبار *Test Central Methods*

هذه المجموعة من الطرق تستند بصفة رئيسة إلى أحكام الخبراء، سواء كانت فردية أو جماعية وهذه الأحكام تتعلق بمحتوى الاختبار ومفرداته، دون الاسترشاد ببيانات تجريبية مستمدة من التطبيق الفعلي للاختبار. ويرى أصحاب هذه الطرق أنَّ تحديد مستويات الأداء يتطلب بالضرورة أحكاما قيمية تأخذ بعين الاعتبار الهدف من الاختبار ومحتواه وطبيعة السمة التي يقيسها الاختبار. وفيما يأتي أشهر الطرق والأساليب شائعة الاستخدام:

- **طريقة نيدلسكاي *Nedelsky's Method (1954)***

أشار نيدلسكاي إلى أنَّ هذه الطريقة تستخدم مع الاختبار من نوع الاختبار من متعدد وتتطلب من كل محكم فحص كل مفردة من مفردات الاختبار ثمَّ يحدد من البدائل الخاطئة لكل مفردة تلك التي

من الممكن أن يتجنب اختيارها الطالب ذو الحد الأدنى من الكفاية، ويكون الحد الأدنى لاحتمال الإجابة الصحيحة عن المفردة هو مقلوب عدد البدائل الباقية (ذيب، ب.ت).

وتحسب درجة القطع في هذه الطريقة بالنسبة لكل محكم من خلال إيجاد المتوسط الحسابي لتقديراته، وبالتالي تصبح درجة القطع المقدرة من جميع المحكمين هي:(العجمي، 2011، ص.22).

$$MPL = \bar{M}FD + K\sigma FD \dots\dots\dots (02)$$

$$\bar{M}FD = \frac{FD}{N} \dots\dots\dots (03)$$

$$\sigma FD = \sqrt{\sum_{i=1}^{n} \frac{(MFD - \bar{M}FD)^2}{N}} \dots\dots\dots (04)$$

حيث أنَّ:

$\bar{M}PL$: تمثل درجة القطع

$\bar{M}FD$: متوسط الحد الأدنى الذي حدده المحكمون

K: مقدار ثابت يتفق عليه المحكمون

N: عدد المحكمين

ومن عيوب هذه الطريقة ما ذكرته جريجوري سيزك(Gregory, 1996)

- لا تستخدم هذه الطريقة إلا مع الاختبارات من متعدد.
- لا تسمح باحتمالات فيما بين (0.5 –1)
- غالباً ما تعطي هذه الطريقة درجة قطع أقل مقارنة بالطرق الأخرى، لأنَّ المحكمين لا يستخدمون قيمة الاحتمال (واحد صحيح).
- لا يوجد تبرير منطقي لافتراض ندلسكاي بأنَّ الأفراد الذين لا يعرفون الإجابة الصحيحة يلجؤون إلى التخمين، إذ نادراً ما يلجأ الأفراد إلى ذلك كون بدائل فقرات الاختبار من متعدد تُصمم بحيث تجنب غير المتمكنين منهم(Berk, 1986) .

- **طريقة أنجوف *Angoff's Method (1971)***

يطلب من محكمين مدربين على هذه الطريقة ومختصين في المادة، تخيل حداً أدنى مقبول لنجاح الطلبة، بعد ذلك يعرض الاختبار على كل محكم، ثمَّ يطلب منه أن يضع على كل سؤال نسبة احتمال إجابته من قبل الطالب الذي يمثل أداؤه الحد الأدنى من الكفاية في المادة. وتجمع النسب من

قبل كل محكم ليمثل ذلك المجموع الدرجة التي تمثل الحد الأدنى من الكفاية بالنسبة لذلك المحكم. ثمَّ تجمع درجات المحكمين ويؤخذ متوسطها ليمثل درجة القطع لهذا الاختبار (الدوسري، 2001).

- **طريقة إيبل *Ebel's Method (1972)***

تعتمد هذه الطريقة أيضاً على آراء المحكمين في تحديد درجة القطع للاختبار. وتتلخص فكرتها بأن تعرض فقرات الاختبار على مجموعة من المحكمين مع مصفوفة مكونة من أربعة صفوف وثلاثة أعمدة. وتمثل الصفوف مدى أهمية تضمين كل فقرة من فقرات الاختبار في الاختبار (أساسية – مهمة – مقبولة – مشكوك فيها). وتمثل الأعمدة مستوى صعوبة الفقرة (سهلة، متوسطة الصعوبة، صعبة). وبعد ذلك يطلب من كل محكم أن يقدر عدد الفقرات التي يحتمل أن تقع في كل خلية، وكذلك النسبة المئوية لعدد المفحوصين الذين ينبغي أن يجيبوا إجابة صحيحة عن هذه الفقرات، علماً بأنَّه يجب أن يتوافر في هؤلاء المفحوصين الحد الأدنى من الكفاية في المجال الذي يقيسه الاختبار. وتحسب درجة القطع للاختبار من خلال إيجاد متوسط حاصل ضرب عدد الفقرات في النسبة المئوية.

وقد أشار علام (2001) إلى أنَّه على الرغم من أنَّ طريقة إيبل اعتمدت في تحكيم فقرات الاختبار على بعدين يتعلقان بأهمية الفقرة وصعوبتها، إلا أنَّ إيبل لم يقترح مواصفات محددة لكل بعد من هذين البعدين. ولذلك فقد يعتمد كل محكم على مواصفات يحددها لنفسه والتي بالتالي تؤثر في دقة التصنيف واتساقه.

**ب). الطرق المتمركزة حول المفحوص**

تعتمد هذه المجموعة من الطرق اعتماداً أساسياً على البيانات التجريبية المستمدة من تطبيق الاختبار على عينة أو أكثر من الأفراد، وتحليل هذه البيانات باستخدام الأساليب الإحصائية والنماذج الرياضية. وتختلف هذه الطرق اختلافاً ملحوظاً عن الطرق السابقة، من حيث أنَّ عملية التحكيم تتعلق باختيار الطلبة المتمكنين وغير المتمكنين في المقررات الدراسية، وليس بمحتوى الاختبار أو فقراته.

ويغلب على هذه الطرق التوجه، الكمي حيث تسترشد بالنماذج الرياضية والإحصائية. ومن أهم وأشهر هذه الطرق ما يأتي:

- **طريقة المجموعات المحكية *Criterion Group Method***

اقترحها بيرك Berk وتتطلب هذه الطريقة من المحكمين أن ينتقوا مجموعة من الأفراد المتمكنين للسمة التي يقيسها الاختبار، ومجموعة أخرى من غير المتمكنين، بناء على معلومات متاحة لهم.

فمثلاً يمكن اعتبار الطلبة الذين تلقوا تعليماً في محتوى وحدة دراسية أو مقرر دراسي متمكنين والطلبة الذين لم يتلقوا التعليم غير متمكنين. وبعد تحديد المجموعتين يتم تطبيق اختبار محكي المرجع على المجموعتين، ومن ثمَّ يتم الحصول على التوزيع التكراري للمجموعتين والدرجة التي تميز بين المجموعتين تعد بمثابة درجة القطع. ويمكن الحصول على هذه الدرجة بحساب القيم الاحتمالية للتصنيف مع مراعاة إعطاء أوزان مناسبة لأخطاء التصنيف بحسب درجة خطورتها في القرارات التي تتخذ بشأن هؤلاء الأفراد. وتشير شيبارد (Shepard, 1983) إلى أنَّ طريقة بيرك صعبة التطبيق بشكل عملي، وذلك بسبب صعوبة تعريف المجموعة التي لم تتلق التعليم (العجمي، 2011، ص.28).

- **طريقة المجموعات المتضادة *Contrasting Groups Method***

اقترح ليفنجستون وزيكي (Livingston & Zeiky, 1982) هذه الطريقة، وهي مماثلة لطريقة بيرك ولكنها تختلف عنها في كيفية تحديد مستوى الأداء في الاختبار. ففي هذه الطريقة يطلب من المحكمين الذين يكونون على دراية كافية بأفراد المجتمع المستهدف والكفاية التي يقيسها الاختبار تحديد مجموعتين من الأفراد، مجموعة متمكنة ومجموعة غير متمكنة، ثمَّ يطبق الاختبار على المجموعتين ويتم رسم منحنى الدرجات للمجموعتين ونقطة تقاطع المنحنيين تعتبر درجة القطع للاختبار (الجبلي،2005، ص.201).

- **طريقة المجموعة الحدية *Borderline Group Method***

اقترحها (زيكي وليفنجستون) أيضاً، وتتطلب هذه الطريقة من المحكمين تحديد مجموعة الأفراد الذين يرون أنَّ كفايتهم بالسمة التي يقيسها الاختبار تقع عند الحد الفاصل بين التمكن وعدم التمكن وبعد تحديد هذه المجموعة يطبق الاختبار على أفراد هذه المجموعة ويحسب وسيط درجات هذه المجموعة ليكون بمثابة درجة القطع للاختبار(Stephenson et. Al, 2000) .

وبهذا الخصوص، يشير بيرك (1986) Berk إلى أنَّه يصعب على المحكمين تحديد المجموعة الحدية، إذ أنَّ تعريف الفرد الذي يقع عند الحد الفاصل بين التمكن وعدم التمكن أصعب كثيراً من تحديد الفرد المتمكن وغير المتمكن.

## 4. الخصائص السيكومترية للاختبار محكي المرجع ولفقراته

### 1.4. صدق الاختبارات محكية المرجع

لا يوجد اختلاف في المبـــادئ الأساسية في صدق الاختبـــارات، سواء كان الاختبار معياري

المرجع أو محكي المرجع، فمصطلح صدق الاختبار (Validity) يشير إلى ما إذا كان الاختبار يقيس ما صمم أصلاً لقياسه، وبالتالي فإنَّ الصدق عادة ما يرتبط بغرض الاختبار وبهدفه، ولأنَّ الاختبار يستخدم لأغراض مختلفة، فإنَّ نوع الصدق يرتبط بالغرض الخاص من الاختبار وطبيعة القرار الذي يمكن أن يبنى على نتائجه، لذلك فإنَّ دلائل وجود صدق الاختبار تتنوع بتنوع أغراضه.

وعليه، فإنَّ أدلة صدق الاختبار عادة ما تعتمد على أحكام الخبراء، ونتائج التحقق الإمبريقي والتحليل الإحصائي لدرجات الاختبار. ولقد استخدم المختصون بالقياس والتقويم عدة مصطلحات تتعلق بصدق الاختبار منها: الصدق الوصفي، الصدق الوظيفي. صدق انتقاء النطاق السلوكي. وهذه الطرق تناظر صدق المحتوى، والصدق التجريبي، وصدق التكوين الفرضي في الاختبارات معيارية المرجع (علام، 2015).

**1.1.4. الصدق الوصفي**

هو قدرة الاختبار على وصف أداء الفرد للنطاق السلوكي ومحاولة لتأكيد مدى ما يقيسه الاختبار ويعتبر الصدق الوصفي شرطاً أساسياً لأنواع الصدق الأخرى، ويمكن التحقق منه عبر عرض الاختبار على مجموعة من المحكمين و الخبراء وقد أشار (1980) Berk إلى أهمية صدق المحتوى للاختبارات محكية المرجع والذي يتقرر على ثلاثة عوامل، هي: صدق الفقرة، والجودة الفنية للفقرة ودرجة تمثيل الفقرات للمحتوى، حيث يتحدد صدق الفقرة بالمدى الذي يعكس به محتوى هذه الفقرة المجال الذي اشتقت منه، كما أشار(بيرك) إلى وجود منحيان لتحديد صدق الفقرات في الاختبار محكي المرجع: أولهما يتمثل باللجوء إلى خبراء المحتوى المحكمين، والثاني يتمثل بالأساليب التجريبية التي تقوم على تنفيذ دراسة استطلاعية على مجموعة من الطلبة لهم نفس سمات الطلبة الذين سيُعطَون الاختبار وجمع الاستجابات وتحليلها. أما الجودة الفنية للفقرات فإنَّه يتم التوصل إلى مؤشرات عنها عن طريق البحث في ملاءمة الفقرات من حيث المحتوى وتعلقها بالمجال الذي أعدت لقياسه ويتحقق ذلك من خلال إعطائها الخبراء في القياس حيث يطلب منهم مراجعتها من أجل تحديد الأخطاء الفنية في بنيتها، مثل الأخطاء اللغوية أو العلمية أو الخطأ في توزيع الإجابات الصحيحة ووجود مؤشرات للإجابة الصحيحة بين البدائل، وترتيب البدائل وهكذا...

**2.1.4. الصدق الوظيفي**

هو قدرة الاختبار على تحقيق الوظيفة التي صمم من أجلها أو الغرض الذي بني من أجله ولا

يتوقف عن تصنيف المفحوصين وفق أدائهم وينطوي تحت هذا النوع من الصدق نوعان هما:

**أ). صدق القرار *Decision Validity***

يعتبر واحداً من أبرز مؤشرات الصدق للاختبارات محكية المرجع، وذلك لأنَّه في كثير من الأحيان تستخدم هذه الاختبارات لاتخاذ قرارات حول مدى إتقان الطالب المحتوى السلوكي الذي يقيسه الاختبار، وحتى يتم اتخاذ هذه القرارات بدرجة من الدقة واليقين لابدَّ من توفر مؤشرات تدعم دقة هذه القرارات، وعادة يتم اتخاذ القرارات بناءً على درجات الاختبار محكي المرجع، من خلال درجات قطع معينة تقرر أفضل النقاط للتمييز بين الطلبة الذين أتقنوا المجال السلوكي المقاس بالاختبار والذين لم يتقنوه. وهكذا فإنَّ تقييم صدق القرار يتطلب تعيين معيار للإنجاز على درجة القطع، ومن ثمَّ العمل على مقارنة الأداء على الاختبار لمجموعتين أو أكثر من المجموعات المحكية التي حكم عليها بأنَّها متقنة أو غير متقنة بناءً على محك ما، ودراسة مدى الانسجام بين قرار الاختبار بإتقانهم أو عدم إتقانهم والقرارات المبنية على المحك. وبهذا فإنَّ صدق القرار لكل هدف يمكن أن يقيَّم بجمع عدد الطلبة من المجموعة التي خضعت للتدريس وصُنفوا بطريقة صحيحة مع عدد الطلبة من المجموعة التي لم تخضع للتدريس وصنفوا بطريقة صحيحة (Berk, 1980).

**ب). الصدق بدلالة المحك**

يشير هذا النوع من الصدق إلى العلاقة بين نتائج الاختبار المراد التحقق من صدقه بدلالة محك والنتائج التي يتم الحصول عليها من قياس آخر ممثل لمحك محدد، ويمكن أن يكون المحك في هذه الحالة اختباراً آخراً، بحيث يتم حساب معامل الارتباط بين العلامات على الاختبار المطلوب التحقق من صدقه والعلامات على المحك. وفي هذه الحالة يسمى معامل الارتباط بمعامل الصدق بدلالة المحك. وهناك نوعان من الصدق بدلالة محك: الصدق التلازمي، ويُعنى بكشف العلاقة بين الاختبار ومحك تجمع البيانات عليه في وقت ملازم لعملية جمع المعلومات على الاختبار، أي أنَّه يتم مقارنة درجات الأفراد على الاختبار ودرجاتهم على مقياس موضوعي آخر بحسب مركزهم فيما يقيسه الاختبار. والصدق التنبؤي الذي يُعنى بإيجاد العلاقة بين نتائج الاختبار ونتائج محك نحصل عليه في المستقبل(Helton, Workman & Motuszek,1982، كما ورد في أبو عواد، 2006).

**3.1.4. صدق انتقاء النطاق السلوكي**

يقابله الصدق الفرضي(Construct Validity) أو صدق البناء للاختبارات معياريـــة المرجـــع

فبالنسبة لهذا النوع من الصدق يتمثل بالدرجة التي يقيَّم بها الاختبار افتراضات السمة أو المفهوم التي بني الاختبار لقياسها، حيث يفترض أنَّ كل اختبار من الاختبارات يبنى على أساس نظرية معينة يمكن استخدامه في التنبؤ بأداء الأفراد وعندها يكون صادقاً صدق البناء، وإذا لم يكن بالإمكان التنبؤ بالأداء باستخدامه فإما أن يكون الاختبار غير صادق أو النظرية خاطئة، أو أنَّ هنالك خطأ تمَّ ارتكابه أثناء إجراء التجربة. وبعبارة أخرى فإنَّ صدق البناء يعني إلى أية درجة تؤكد نتائج تطبيق الاختبار صحة الافتراضات المستخلصة من النظرية حول مفهوم السمة التي وضع لقياسها، وهناك عدة طرق للتحقق من صدق البناء لاختبار ما منها:(Crocker & Algina, 1986)

- **مقياس التحليل لجوتمان:** إذا كان بالإمكان ترتيب فقرات الاختبار خطياً أو هرمياً فإنَّه يمكن التنبؤ بدرجة الفرد على الاختبار من معرفة هذا الترتيب، وفي الحالات التي لا يمكن التنبؤ فيها بدرجات الأفراد فإنَّ ذلك قد يعود إلى أنَّ التصنيف الهرمي غير دقيق.

- **التحليل العاملي:** حيث يتم استخراج دلالات عن صدق البناء بهذه الطريقة، من خلال مصفوفة الارتباطات بين الدرجات على فترات الاختبار في محاولة لإنقاص عدد العوامل أو المكونات التي تتجمع حولها فقرات الاختبار. ونظراً لأنَّه من المتوقع أن يكون التباين المتوقع في العلامات على الاختبار محكي المرجع منخفضاً فإنَّ هذا من شأنه أن يحد من أهمية هذه الطريقة التي تعتمد على الترابطات الداخلية بين فقرات الاختبار.

- **المقارنات الطرفية:** تستخدم هذه الطريقة في حالة الرغبة في معرفة قدرة الاختبار على التمييز بين المستويات المختلفة للسمة، أو التمييز بين المجموعات التي تمتلك درجات مرتفعة من السمة وتلك التي تمتلك درجات منخفضة منها، فقد تتضمن افتراضات السمة أو المفهوم وجود مجموعات مختلفة في درجة امتلاكها للسمة التي يقيسها الاختبار، فإذا كشفت نتائج الاختبار عن وجود فرق حقيقي بين المجموعات فإنَّ هذا مؤشر لصدق البناء للاختبار.

### 2.4. ثبات الاختبارات محكية المرجع

يعتمد تقدير طرق ثبات الاختبارات محكية المرجع على فحص البيانات أكثر من أن تبنى على الأساس الكمي من النوع الذي تحدد به ثبات المقاييس معيارية المرجع، كما يعتمد معامل الثبات للاختبار محكي المرجع على عدة عوامل، تؤثر على عملية اتخاذ القرار مثل طريقة تحديد المختبرين

لمستويات التفوق واختبار الدرجة الفاصلة(درجة القطع)، وطول الاختبار وتباين المجموعة(عوض الله، 2000)، وقد اشار أبو عواد(2006) أنَّ كل من (هامبلتون، سواميناثان والجينا) أشاروا إلى وجود ثلاثة مفاهيم مختلفة للثبات في سياق الاختبارات محكية المرجع والتي تتمثل في:(أبو عواد، 2006، ص.41-42).

- ثبات قرارات التصنيف حسب محك الإتقان، ويعني اتساق قرارات تصنيف الأفراد من خلال قياسات متكررة.
- ثبات درجات الاختبار محكي المرجع، ويعني اتساق مربع انحرافات درجات الأفراد عن درجة القطع خلال إعادة تطبيق الاختبارات مرات أخرى.
- ثبات تقديرات الدرجات على المجال، أي اتساق درجات الفرد خلال تطبيق الاختبار أكثر من مرة.

وهناك عدة أساليب للتحقق من ثبات الاختبارات محكية المرجع استناداً إلى المفاهيم السابقة ولكن قبل عرض هذه الأساليب ينبغي توضيح مفهوم الاختبارات المتوازية والفروض التي تستند إليه وكذلك مدى تأثر هذه الطرائق بدرجات القطع التي يتم بها تحديد الاختبارات المتوازية Parallel Tests.

تعتمد كثير من طرائق تقدير ثبات الاختبارات محكية المرجع على مفهوم الاختبارات المتوازية وما تستند إليه من فروض، ويتعلق هذا المفهوم بخصائص النطاق السلوكي للمفردات الذي تُستمد منه عينات المفردات، وبكيفية انتقاء عينتي المفردات التي يتكون منها اختباران متوازيان. والاختباران يكونان متوازيين إذا تساوى كل من المتوسط، والانحراف المعياري لدرجاتهما وارتبطا ارتباطاً متساوياً مع مجموعة من الدرجات الحقيقية، وتباين درجات كل منهما الذي لا تفسره الدرجات الحقيقية يعزى إلى الأخطاء العشوائية (علام، 2005، ص.50).

وهنا ينبغي التمييز بين الاختبارات المتوازية التي تبني مفرداتها بالطرائق التقليدية أي استناداً إلى الأهداف السلوكية المتعارف عليها، وبين الاختبارات المتوازية التي تشتمل على عينات عشوائية أو عشوائية طبقية من نطاق سلوكي تمَّ تحديده باستخدام المواصفات التفصيلية أو صيغ المفردات.

ففي الحالة الأولى ينبغي بناء عينتين من المفردات تقيس كل منهما الأهداف السلوكية نفسها وتشكل كل منهما اختبار بصورة يتساوى متوسط وتباين توزيع درجات الاختبارين، وكذلك تتساوى الارتباطات البينية لدرجات مفردات كل منهما، وعندئذ يقال أنَّ الاختبارين متوازيان أو متكافئان

استناداً إلى الأساليب الكلاسيكية (Classically Parallel Tests) ومعظم طرائق تقدير ثبات الاختبارات محكية المرجع تعتمد على هذه الفروض.

أما في الحالة الثانية فإنَّ الاختبارات المتوازية تكون مفرداتها منتقاة عشوائية من النطاق الشامل للمفردات، وبذلك لا تتطلب تحقق أي من الفروض السابقة ويقال عندئذ إنَّ الاختبارات متوازية عشوائية (Parallel Tests Randomly) (علام، 2001، ص.288).

وقد توصل العلماء إلى مجموعة من الطرق لتقدير ثبات الاختبارات محكية المرجع سيتم التطرق إلى أهمها وأكثرها شيوعاً واستخداماً منها:

**1.2.4. طرق تقدير ثبات درجات الأفراد في النطاق السلوكي**

طرق هذه المجموعة تهتم بتقليل خطأ التباين الناتج عن الفروق بين درجات الطلبة في النطاق السلوكي بالإضافة إلى درجات النطاق، ومن هذه الطرق:(مجيد،2014، ص.199).

- **معامل ليفنجستون *Livingston-Index (1972)***

اعتمد ليفنجستون في اشتقاقه لمعامل الثبات الخاص بالاختبارات محكية المرجع على مفهوم درجة القطع والتي تفصل بين المتقين وغير المتقنين(Sax & Newton, 1997) وبهذا فإنَّ هذا المعامل يهتم (بإيجاد تباين درجات الأفراد عن مربع الفرق متوسط تلك الدرجات عن درجة القطع) بدلاً من انحرافها عن المتوسط، كما يهتم هذا المعامل بمفهوم متوسط مجموع مربعات انحرافات الدرجات الملاحظة المتوقعة وذلك بحساب قيمة انحراف درجات كل طالب عن درجة القطع في الاختبار (Crocker & Allgina, 1986) ويتم ذلك وفق الصيغة الرياضية التالية:

$$K^2(X,T) = \frac{\delta^2 T + (M_T - n_i C)^2}{\delta^2 X + (M_x - n_i C)^2} \dots\dots\dots (05)$$

حيث أنَّ:

$K^2(X,T)$: تشير إلى معامل ليفنجستون.

$\delta^2 T$: تشير إلى مجموع مربعات انحرافات الدرجات الملاحظة حول درجة القطع (C).

$\delta^2 X$: تشير إلى مجموع مربعات انحرافات الدرجات الحقيقية حول درجة القطع (C).

M: تشير إلى متوسط درجات الطلبة في النطاق السلوكي الذي يقيسه الاختبار.

$n_i$ : تشير إلى عدد الأسئلة.

C : درجة القطع.

ولتقدير معامل ليفنجستون في حالة استخدام اختبارين لكل طالب نستخدم الصيغة الرياضية التالية:(الجبلي، 2005، ص.182).

$$\widehat{K}^2(X,T) = \frac{\hat{P}xx\ \hat{\sigma}x\ \hat{\sigma}x + (\widehat{M}x - niC)(\widehat{M}x - niC)}{\sqrt{\left\{\hat{\sigma}X^2 + (\widehat{M}x - niC)^2\right\}\left\{\hat{\sigma}X^2 + (\widehat{M}x - niC)^2\right\}}} \dots\dots\dots (06)$$

حيث أنَّ:

$\hat{P}xx$ : معامل الارتباط بين درجات كلا الاختبارين.

$\hat{\sigma}X^2$ : متوسط مربعات انحرافات الدرجات الحقيقية عن درجة القطع.

$\widehat{M}$ : متوسط الدرجات الخام.

$ni$: عدد الأسئلة.

$C$: درجة القطع.

ولتقدير معامل ليفنجستون في حالة استخدام اختبار واحد نستخدم الصيغة الرياضية التالية:

$$\widehat{K}^2(X,T) = \frac{\sigma^2 x(KR20) + (Mx - niC)^2}{\sigma^2 x + (M - niC)^2} \dots\dots\dots (07)$$

حيث: ترمز(KR20) إلى معامل ارتباط كيودر ريتشاردسون.

**2.2.4. طرق تتعلق بتصنيف الافراد في مجموعات بحسب تمكنهم من نطاق السلوكي**

تهتم طرق هذه المجوعة بتحديد الأخطاء الناتجة عن عدم اتساق التصنيف عن تطبيق اختبارين متوازيين أو في حالة إعادة الاختبار على أفراد العينة ومن هذه الطرق ما يلي:

**أ). طرق تتطلب تطبيق الاختبار مرة واحدة:**

- **طريقة هاريس *Harris Method***

لا تعتمد هذه الطريقة في تحديدها على متغير طول الاختبار كما أشار هاريس1974 (Harris) إنَّما تقوم على الارتباط بين متغير يمثل مجموع درجة الطالب في الاختبار ومتغير ثنائي آخر يمثل تصنيف الطالب إلى متقن أو غير متقن على ضوء درجة القطع، وعليه يتم تصنيف الطلبة إلى أربع مجموعات في (جدول رباعي) وهي طلبة متقنين فعلاً، وطلبة غير متقنين فعلاً، ومجموعة طلبة الخطأ الأول (α) ومجموعة طلبة الخطأ الثاني (β)، ويتم تحديدهم بناء على درجة قطع الاختبار. وقد أطلق على هذه الطريقة بمعامل(هاريس) ويرمز له بالرمز (Mc) ويتم تقديره بالصيغة التالية: (الجبلي،2005،

ص.182).

$$M_c = \frac{SSb}{SS_b + SS_w} \dots\dots\dots (08)$$

حيث أنَّ:

$M_c$: معامل ثبات هاريس.

$SS_b, SS_w$: عبارة عن مجموع المربعات (داخل، بين) المجموعات.

وتمتد قيمة المعامل ما بين (صفر وواحد).

- ***معامل كابا (هاينا) Huynh Kappa Coefficient***

يتم تقدير معامل(هاينا) في حال تطبيق اختباراً واحداً أو اختيارين متوازيين، وفي حالة تقدير معامل الثبات(هاينا) لاختبارٍ واحدٍ فقط، فقد أشار سابكوفياك(Subkoviak) بأنَّ تقدير هذا المعامل يستند إلى عدة فروض هي:

- أن يتمثل توزيع درجات الطلبة الحقيقية في الاختبار (Beta, Distribution) شكل بيتا إذا طبق عدد من الاختبارات على طالب ما، فإنَّه يفترض أن يتخذ توزيع درجاته في الاختبار شكل التوزيع ذي الحدين، وفي هذه الحالة يجب أن يكون تقدير درجات أسئلة الاختبار إما صفراً أو واحداً.
- أن تكون الإجابة على الأسئلة مستقلة إحصائياً بحيث لا تؤثر الإجابة على إحدى المفردات على إجابة المفردة الأخرى.
- أن تتساوى تقريباً معاملات صعوبة جميع أسئلة الاختبار.

ويتم تقدير معامل ثبات(هاينا) في إحدى الصيغتين التاليتين:(مجيد،2014، ص.201–202)

$$\widehat{K} = \frac{\hat{P}_{zz} - \hat{P}_z^2}{\hat{P}_z - \hat{P}_{z^2}} \dots\dots\dots (09)$$

$$\hat{P}o = 1 - 2\left(\hat{P}_z - \hat{P}_{zz}\right) \dots\dots\dots (10)$$

وتشير سوسن مجيد (2014) أنَّه يتطلب تحقيق الفروض السابقة إجراء الخطوات التالية:

- تقدير متوسط درجات الطلبة.
- تقدير متوسط درجات الطلبة ($\widehat{M}$) والانحراف المعياري ($\sigma^2$) ومعامل ثبات كيودر ريتشاردسون21 وتحديد درجة القطع(C) وبارامتر توزيع ($\alpha$).
- تحويل القيم السابقة إلى بيانات معيارية وفق الصيغ التالية:

$$\widehat{M} = sin - 1\sqrt{\frac{\widehat{M}}{n}} \dots\dots\dots (11)$$

حيث أنَّ:

$\widehat{M}$: ترمز إلى المتوسط

n: ترمز إلى عدد المفردات

$$\sigma = \sqrt{(\hat{\alpha}21 + 1) + (\hat{\alpha} + N)}$$

$\sigma$ : ترمز إلى الانحراف المعياري.

$$\hat{P} = \hat{\alpha}21\sqrt{(n-1)/(n + \hat{\alpha}21)}$$

$\hat{P}$: ترمز إلى معامل الثبات.

$$\hat{C} = sin^{-1}\sqrt{(C - 0.5)n}$$

$\hat{C}$: درجة القطع.

تقدير قيمة الاحتمال($\widehat{\mathrm{PZ}}$) من الجداول الخاصة بذلك بحيث تكون هذه القيمة أقل من قيمة(z) ثمَّ تحدد قيمة الاحتمال($\widehat{\mathrm{PZZ}}$) من جداول دالة التوزيع الثنائي الاعتدالي بحيث تكون قيمة الاحتمال أيضاً أقل من قيمة (z).

- ***معامل الاتفاق لسابكوفياك Subkoviak Coefficient of Agreement***

اقترح سابكوفياك(1976) طريقة تعتمد على مدى اتساق تصنيف الأفراد إلى مجموعتين إحداهما متمكنة والأخرى غير متمكنة باستخدام نتائج تطبيق الاختبار مرة واحدة(علام، 1995) وتتضمن الطريقة تقدير مقدار احتمال تصنيف كل فرد تصنيفاً صحيحاً، وجمع هذه المقادير الاحتمالية وإيجاد متوسطها، وبذلك نحصل على معامل اتفاق خاص بالمجموعة التي يطبق عليها الاختبار والصيغة الرياضية لهذا المعامل هي:

$$P_o = \frac{\sum P_x}{N} \dots\dots\dots (12)$$

حيث:

N: عدد أفراد المجموعة.

$P_x$: معامل الاتفاق التقديري لكل فرد من أفراد المجموعة.

ويحسب $P_x$ من العلاقة التالية:

$$Px = KR20\left(\frac{x}{n}\right) + (1 - KR20)\left(\frac{M}{N}\right) \dots\dots\dots (13)$$

حيث أنَّ:

KR20: معامل كيودر ريتشاردسون.

x: درجة الطالب في الاختبار.

n: عدد الفقرات.

**ب). طرق تتطلب تطبيق الاختبار مرتين**

- **معامل كارفر *Carver Method***

تقوم هذه الطريقة على تطبيق اختبارين متوازيين على مجموعة واحدة من الأفراد، ويتم مقارنة نسبة عدد الطلبة الذين تمَّ تصنيفهم ضمن المتقنين في كلا الاختبارين، فإذا كانت النسبتين متماثلتين أو قريبتين من بعضهما يعتبر الاختبارين ثابتين، لكن ما يعاب على هذه الطريقة أنَّها غير حساسة لاتساق قرارات كل طالب في الاختبارين.(Subkoviak, 1984, p269 كما ورد في شكري، 2006) ويمكن إيجاد قيمة معامل (كارفر) وفق الجدول التالي:

**جدول رقم (01) ايجاد قيمة معامل كارفر**

| الاختبار (أ) | الاختبار (ب) | | |
|---|---|---|---|
| | | غير متقن | متقن |
| متقن | B | A | |
| غير متقن | C | D | |

ويحسب معامل كارفر باستخدام الصيغة التالية:

$$\frac{A + C}{N} = \cdots \dots\dots (14)$$

حيث أنَّ:

A: مجموع الطلبة المتقنين بالفعل.

C: مجموع الطلبة غير المتقنين بالفعل.

N= A + B + C + D (مجيد، ص.203)

- ***معامل كابا لـــ Swaminathan, Hambleton, Algina***

أشارت سوسن مجيد (2014) إلى أنَّ كل من هامبلتون، وآخرون أشاروا إلى أنَّ هذه الطريقة تعتبر امتداداً لطريقة(كارفر) حيث تعتمد على فرق النسبة بين الطلبة المتمكنين وغير المتمكنين في صورتي الاختبار المتوازيين، أو في مرتي تطبيق الاختبار، في ضوء درجة قطع محددة وتمتد قيمة معامل الثبات بين (+1، -1)، ويتم تقدير معامل وفق الصيغة التالية:

$$\hat{P}o = \sum_{k=1}^{m} (\hat{P}\,KK) \dots\dots\dots (15)$$

حيث أنَّ:

$\hat{P}o$: ترمز إلى معامل الثبات أي اتساق التصنيف.

$\hat{P}\,KK$: ترمز إلى نسبة الطلبة الذين تمَّ تصنيفهم ضمن المتقنين في المستوى(K) في مرتي تطبيق الاختبار.

M: ترمز إلى عدد مستويات الاتقان.

ونظراً لأنَّ هذه الطريقة لا تأخذ بعين الاعتبار اتساق التصنيف الناتج عن عامل الصدفة، لذا اقترح كل من (سواميناثان، هامبلتون والجينا) استخدام معامل كابا الذي ينسب إلى كوهن.

$$K = \frac{\hat{P}o - \hat{P}c}{1 - \hat{P}c} \dots\dots\dots (16)$$

حيث أن:

K: معامل كابا.

$\hat{P}o$: نسبة الاتفاق الملاحظ في التصنيفات وتحسب بالصيغة السابقة $\hat{P}o = (\hat{P}\,KK)$

$\hat{P}c$: ترمز إلى نسبة الاتفاق المتوقع في التصنيفات وتحسب بالصيغة التالية:

$$\hat{P}o = \sum_{k=1}^{m} PK.P.K \dots\dots\dots (17)$$

حيث أنَّ: (PK)(P.K) نسبة الطلبة المصنفين ضمن فئة المتقنين في المجموعة(K) في مرتي تطبيق الاختبار على التوالي.

- **معامل الاعتمادية لبرينان وكين**

يسمى أيضاً "معامل إمكانية التعميم" إذ يوضح كرونباك أنَّ هناك نوعين من تباين الخطأ أحدهما

يتعلق بتفسير البيانات المستمدة من الاختبارات معيارية المرجع، والأخر يتعلق بتفسير البيانات المستمدة من الاختبارات محكية المرجع، ومعامل التعميم هنا هو النسبة بين تباين الدرجة الشاملة والتباين المتوقع للدرجة الملاحظة، وأطلق عليه برينان معامل الاعتمادية (علام،2001، ص.292-293).

وقيم معامل الاعتمادية تختلف باختلاف درجات القطع، وأنَّ أصغر قيمة له تساوي الصفر وأكبر قيمة له تساوي الواحد الصحيح. وقد وضَّح بيرك (Berk,1986) أنَّه إذا كانت:

درجة القطع = متوسط درجات عينة الأفراد، فإنَّ معامل ليفنجستون يصبح مطابقاً لمعامل كيودر ريتشاردسون (الصورة 20) في حين معامل الاعتمادية يصبح مطابقاً لمعامل كيودر ريتشاردسون (الصورة 21). وهناك تشابه كبير بين معامل الاعتمادية ومعامل ليفنجستون، لكن المعامل الثاني يعتمد على المفهوم الكلاسيكي للاختبارات المتوازية، بينما يعتمد المعامل الأول على مفهوم الاختبارات المتوازية العشوائية (عوض الله، 2000، ص.67).

### 3.4. صعوبة وتمييز فقرات الاختبار محكي المرجع

يختلف مفهوم الصعوبة في نظرية الاستجابة للفقرة عنه في النظرية الكلاسيكية وذلك أنَّ النظرية الحديثة تفترض وجود دالة مميزة لكل فقرة من فقرات الاختبار تتخذ شكل المنحنى اللوغاريتمي التراكمي وتتميز هذه الدالة بخصائص يمكن الاستفادة منها في جعل مقياس علامات الاختبار خطياً وفي تقدير الخصائص السيكومترية للاختبارات تقديراً مستقلاً عن خصائص عينة الأفراد المختبرين (Hambleton & Swaminathan, 1985).

#### أ). معامل صعوبة الفقرة

إنَّ حساب معامل صعوبة مفردات الاختبارات محكية المرجع لا يختلف عنه في حالة مفردات الاختبار معيارية المرجع. ولكن التركيز على صعوبة المفردة في هذا النوع من الاختبارات لا يعتمد على مقدار ما تضيفه من تباين إلى الدرجات الكلية للاختبار، كما أنَّها ليست مؤشراً أساسياً يستخدم لتقويم الفقرة، إذ يتوقع أن تكون المفردات محكية المرجع أسهل عموماً من المفردات معيارية المرجع فالفقرة التي معامل صعوبتها 50% ربما تصنف على أنَّها متوسطة الصعوبة في الاختبار معياري المرجع، وتكون صعبة جداً في الاختبار محكي المرجع شريطة تطبيقها على الأفراد أنفسهم، كما ينصح بتحديد متوسط أو وسيط معاملات صعوبة المفردات محكية المرجع التي تقيس هدفاً ما وتستخدم

هذه القيمة في تقييم فاعلية تدريس ذلك الهدف (النبهان، 2004، ص.205).

**ب). معامل تمييز الفقرة**

يشير معامل تمييز الفقرة إلى قدرة الفقرة على التمييز بين مجموعة الطلبة التي وصلت إلى درجة الإتقان والمجموعة التي لم تصل إلى تلك الدرجة، ويطلق أحياناً على معامل التمييز المحسوب بمعامل الحساسية، وهو مؤشر إحصائي يبين درجة الزيادة في نسبة الطلبة الذين يجيبون عن الفقرة إجابة صحيحة بعد انتهاء عملية التدريس. وهذا يتطلب تطبيق الاختبار مرتين، قبل التعرض لعملية التدريس وبعد التعرض لها. والواقع أنَّ هنالك أربع طرق تستخدم لإيجاد معامل التمييز تلخيصها فيما يلي:

- **معامل حساسية المفردة (معامل برينان)**

معامل حساسية المفردة هو قدرة الفقرة على التمييز بين الممتحنين الذين تلقوا التدريس أو التدريب والذين لم يتلقوا التدريب، وقد اقترح كل من كوكس وفرقاس(Cox & Vargas) طريقة لحساب حساسية المفردة(S) عندما يتم عرضها على مجموعة من الممتحنين قبل التدريب وبعده وفق المعادلة التالية: (الحسن، 2012، ص.56).

$$S = P\ post - P\ pre \dots\dots\dots (18)$$

حيث أنَّ:

P post: نسبة الممتحنين الذين أجابوا عن الفقرة اجابة صحيحة بعد التدريب.

$P\ pre$: نسبة الممتحنين الذين أجابوا عن الفقرة اجابة صحيحة قبل التدريب.

وتتراوح قيمة(S) بين (−1، +1) وتفضل القيم الموجبة الكبيرة، وتكون الفقرة عديمة الحساسية عندما تكون قيمة(S) صفراً.

وقد اقترح برينان أسلوباً آخراً لحساب معامل حساسية الفقرة شبيها بمقياس كوكس وفرقاس ويختلف عنه فقط في طريقة اختيار مجموعة المحك، فبدلاً من تصنيف الأفراد في مجموعتين إحداهما تلقت التعليم والأخرى لم تتلق، فإنَّهم يصنفون إلى متمكنين وغير متمكنين في ضوء مستوى الأداء المحدد، أي أنَّ معامل برينان كالآتي:

$$S = \frac{U}{n1} - \frac{L}{n2} \dots\dots\dots (19)$$

حيث تشير:

n1: عدد الذين درجاتهم فوق درجة القطع.

$n2$: عدد الذين درجاتهم تحت درجة القطع.

U: عدد الذين أجابوا اجابة صحيحة عن الفقرة ممن درجاتهم الكلية فوق درجة القطع.

L: عدد الذين أجابوا اجابة صحيحة عن الفقرة ممن درجاتهم الكلية تحت درجة القطع.

وتتراوح قيمة هذا المعامل بين (−1،+1) وتعتبر المفردة جيدة وفق معامل برينان إذا كانت قيمتها تساوي أو أكبر من (0.20) (الحسن، 2012، ص.57).

- **معامل التوافق**

يشير هذا المعامل إلى درجة توافق فقرتين في قياس هدف تعليمي أو مهاري واحد، أو ينتمي لمحتوى دراسي واحد، ويحسب وفق القانون الإحصائي كاي تربيع (Chi-square) لاختبار درجة استقلال الفقرتين عن بعضهما في قياس ذلك الهدف (النبهان، 2004، ص.207−208)، كما يمكن التعرف على نسبة الممتحنين الذين اجتازوا أو أخفقوا في الفقرتين معاً وفق المعادلة الآتية:

$$X^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(b+d)(a+c)} \dots\dots\dots (20)$$

إذ تمثل الرموز a,b,c,d تكرارات خلايا جدول التوافق الآتي:

**جدول رقم (02) الجدول الثنائي لحساب معامل التوافق المرجعي**

| | الفقرة 1 | | |
|---|---|---|---|
| | | + | − |
| الفقر2 | + | A | B |
| | − | C | D |

ويمكن حساب درجة الأداء المتسق على فقرتي الاختبار اللتين تقيسان هدفاً واحداً بإيجاد نسبة الممتحنين الذين أجابوا عن الفقرتين إجابة صحيحة، أو خاطئة وفق المعادلة التالية:(عبابنة، 2009، ص.99-101).

$$P = \frac{a+d}{n} \dots\dots\dots (21)$$

- **معامل التوافق المرجعي**

يعد هذا المعامل من أحدث معاملات التمييز الخاصة بتحليل بنود اختبارات الإتقان، والذي يهدف إلى معرفة احتمالية التوافق بين نواتج سؤال معين ونواتج الاختبار، وهو من الطرائق التي تعتمد على تطبيق الاختبار مرة واحدة على مجموعة واحدة من الأفراد، ومن ثمَّ يتم تصنيف أفراد

هذه المجموعة إلى متقنين وغير متقنين، بناءً على تحقيقهم للمستوى المطلوب للإتقان. وقد اقترح كل من هاريس (Harris)، سابكوفياك(Subkoviak) المعادلة الآتية الحساب معامل التوافق المرجعي:

$$\frac{A+D}{N} \dots\dots\dots (22)$$

حيث تشير:

A: عدد الأشخاص المتقنين الذين أجابوا عن السؤال إجابة صحيحة.

D: عدد الأشخاص غير المتقنين الذين أجابوا عن السؤال إجابة صحيحة.

N: العدد الكلي للأشخاص.

وينحصر مدى معامل التوافق المرجعي بين (الصفر، +1) ويمكن حساب الحد الأدنى لمعامل التوافق حينما لا يكون هناك علاقة بين مستوى الاتقان والاستجابة على السؤال، ويحسب الحد الأدنى لمعامل التوافق المرجعي من خلال الجدول الثنائي كما يلي:

**جدول رقم (03) لحساب الحد الادنى لمعامل التوافق المرجعي**

| | | |
|---|---|---|
| صح | A | B |
| خطأ | C | D |

وعن طريق المعادلة الأتية:

الحد الأدنى لمعامل التوافق المرجعي= $\frac{(A+B)(A+C)(C+D)(B+D)}{N^2}$

حيث أنَّ:

A: عدد الأشخاص المتقنين الذين أجابوا عن السؤال إجابة صحيحة.

B: عدد الأشخاص غير المتقنين الذين أجابوا عن السؤال إجابة صحيحة.

C: عدد الأشخاص المتقنين الذين أجابوا عن السؤال إجابة خاطئة.

D: عدد الأشخاص غير المتقنين الذين أجابوا عن السؤال إجابة خاطئة.

N: العدد الكلي للأشخاص.

ويمكن اعتبار الفقرة جيدة وفق معامل التوافق المرجعي إذا كان الفرق بين الحد الأدنى لمعامل التوافق المرجعي ومعامل التوافق المرجعي يساوي أو أكبر من (0.05) (الحسن، 2012، ص.59).

- **معامل فاي *Coefficient- Phi***

يقوم هذا المعامل بحساب درجة التوافق في التصنيف بين الفقرة والاختبار للمفحوصين، وهو

من الطرائق التي يتم فيها تطبيق الاختبار مرة واحدة على مجموعة واحدة من الأفراد، ويتم اختيار درجة فاصلة تمثل مستوى الإتقان، وتتحدد فاعلية الفقرة بقدرتها على التمييز بين المفحوصين عند درجة فاصلة محددة على العلامة الكلية على الاختبار، ويتم إيجاد معامل فاي وفق الجدول الثنائي(2x2) إذ يبين هذا الجدول عدد الإجابات الصحيحة والخاطئة للمتقنين وعدد الإجابات الصحيحة والخاطئة لغير المتقنين.

**جدول رقم (04) يبين عدد الإجابات الصحيحة والخاطئة للمتقنين وعدد الإجابات الصحيحة والخاطئة لغير المتقنين.**

| | ناجح (متقن) | راسب (غير متقن) |
|---|---|---|
| صح | A | B |
| خطأ | C | D |

ولتحقيق قيمة معامل فاي يتم استخدام المعادلة التالية:

$$\emptyset = \frac{AD - BC}{\sqrt{(A+b)(C+D)(A+C)(B+D)}}$$

وتعتبر الفقرة جيدة وفق معامل فاي إذا كانت قيمته تساوي أو أكبر (0.30) (الحسن، 2012، ص.60).

## المحور الثاني: النظرية الكلاسيكية في القياس *Classical True Score Theory*

سيطرت مبادئ النظرية الكلاسيكية على القياس طوال القرن الماضي، على عملية بناء الاختبارات وتحليل نتائجها، وتفسيرها، ولايزال استعمالها في العديد من المجالات رغم وجود قصور في العديد من جوانبها كما أشار لذلك هامبلتون وسوامينثان(Hambleton & Swaminthan, 1985) منها أنَّ خصائص الاختبار وفقراته تتحدد بالنسبة لمجموعة معينة من المفحوصين، كما أنَّ خصائص المفحوصين أنفسهم تتحدد أيضاً بالنسبة لاختبار معين، أي أنَّ تقدير قدرة المفحوصين يعتمد على عينة فقرات الاختبار، وتقدير احصائيات الفقرة يعتمد على عينة المفحوصين الذين طبق عليهم الاختبار.

وتعرف أيضاً النظرية الكلاسيكية بنظرية الدرجة الحقيقية ودرجة الخطأ Theory of True Error Score وتقوم على مبدأين الدرجة الحقيقية وهي تقدير حقيقي لما لدى الفرد من قدرة معينة أو صفة محدد، ودرجة الخطأ التي تكون عشوائية وغير مرتبطة ببعضها البعض (Hambleton & Zaal, 1991, p. 69) وأنَّ درجات الخطأ والدرجات الحقيقية والدرجات الملاحظة تكون مرتبطة خطياً. ويعبر عنها بنموذج الدرجة الحقيقية الذي يأخذ شكل المعادلة التالية:

$$X = T + E \dots\dots\dots\dots (1)$$

حيث:

X: الدرجة الملاحظة (المشاهدة).

T: الدرجة الحقيقية.

E: الدرجة الخطأ.

وتعكس الدرجة الملاحظة تأثير كل من عامل الوقت، نوع الأسئلة والظروف النفسية مما يفترض أنَّه لو أمكن إجراء الاختبار عدة مرات على الفرد بعناصر جديدة وتحت ظروف مختلفة فإننا نحصل على درجات مختلفة متوسطها هو أقرب تقدير غير متحيز لقدرة الفرد.

مثال ذلك سمة القلق من السمات المجردة التي تأثر على سلوك ولكن يمكن الاستدلال عليها من خلال الأعراض التي يظهرها الفرد أو من خلال حساب عدد الإجابات على المقياس الذي يعد الدرجة الملاحظة، أمَّا الدرجة الحقيقية هي التقدير الحقيقي لهذه السمة، أمَّا الدرجة الخطأ مصدرها ما يسمى بالخطأ العشوائي، وهي راجعة للظروف المحيط به أثناء الإجابة مثل الوقت كافي أم لا، نوع الأسئلة ووضوحها، الظروف النفسية للمفحوص أثناء إجراء الاختبار...الخ.

وتعكس النتيجة الملحوظة تأثير عامل الوقت على نوع الأسئلة والظروف النفسية، حيث أنَّه من المفترض إذا كان يمكن إجراء الاختبار عدة مرات على الفرد مع عناصر جديده وتحت ظروف مختلفة، نحصل على درجات مختلفة.

**1. الافتراضات الأساسية للنظرية الكلاسيكية**

وفقاً لما ذكرته(Gregory) فانَّ نظرية القياس الكلاسيكية تتبلور من الفكرة الأساسية أنَّ درجة الفرد على اختبار معين ينتج عن تأثير مجموعتين من العوامل، العوامل التي تؤدي إلى الاتساق والتي تتشكل من السمات الثابتة للفرد التي يتم قياسها في الاختبار، والعوامل التي تؤدي إلى عدم الاتساق وتتكون من مجموعة من العوامل غير المرغوب فيها، وتؤثر على درجة الفرد في الاختبار، وتسمى عوامل الخطأ، وهذا يعني أنَّه إذا كنا نستطيع أن نقوم بالاختبار عدة مرات على نفس الفرد مع عناصر جديدة وتحت ظروف مختلفة، نحصل على درجات مختلفة ومتوسط ذلك هو أقرب تقدير غير متحيز من قدرة الفرد أو درجته الحقيقية، وبقصد تفسير الدرجة من حيث صحتها أو عدمه(Gregory, 2004).

وتشكل النظرية الكلاسيكية أسلوباً بسيطاً بين العوامل المؤثرة في الدرجة الظاهرية للمفحوص والتي تُسبب ما يسمى بأخطاء القياس ويظهر ذلك من خلال الافتراضات الخاصة به، بمعنى أنَّ دقة وصحة النتائج التي يمكن الحصول عليها تبقى مرهونة بمدى تحقق تلك الافتراضات، وأنَّ الإخلال بهذه الافتراضات تجعل النتائج موضع شك (محاسنة، 2013، ص.100).

وتفترض هذه النظرية مجموعة من الافتراضات هي:

- افتراض يتعلق بمكونات الدرجة الحقيقة للمفحوص والتي تمثل قدرته الحقيقية، وينص هذا الافتراض على أنَّ الدرجة الحقيقية للمفحوص تتكون من جزأين، الأول يمثل الدرجة الظاهرية أو الملاحظة، والثاني يمثل درجة الخطأ أو خطأ القياس للدرجة الحقيقية ويمكن صياغة هذا الافتراض على النحو التالي:

د ح = د ظ + د خ  حيث:

د ح: الدرجة الحقيقية.

د ظ: الدرجة الظاهرية.

د خ: الدرجة الخطأ.

- عند خضوع مفحوص لاختبار ما، يقيس سمة أو قدرة معينة، فإنَّ الدرجة التي سيتحصل عليها من إجاباته على فقرات الاختبار تعتبر درجة ملاحظة أو ظاهرية، لأنَّ الدرجة الحقيقة وفق افتراض النظرية الكلاسيكية هي "المتوسط الحسابي للتوزيع النظري للدرجات الظاهرية التي يمكن أن يحصل عليها المفحوص إذا خضع لاختبار لعدة مرات، أو لعدة نماذج اختبارية متوازية منه وفقراتها مستقلة وتقيس نفس السمة، وهذا من الناحية العملية أمراً صعباً، لأنَّه من الصعب أن يحصل المفحوص على نفس الدرجة في كل مرة، إضافة إلى أنَّ الاختبارات المتوازية لها شروطها من حيث توزيع الدرجات وانحرافها المعياري، والوسط...(محاسنة، 2013، ص.103) ويمكن التعبير عن الدرجة الحقيقية للمفحوص بـــ:

د ح = $\underline{س}_1$ + س$_2$ + س$_3$ +.... س$_ن$ + د خ  حيث:

د ح = س + د خ

د ح: الدرجة الحقيقة.

س : الوسط الحسابي لدرجات المفحوص في مرات التطبيق.

دخ: درجة الخطأ.

- ينص على أنَّ قيمة معامل الارتباط بين الدرجة الحقيقية والدرجة الخطأ تساوي صفراً، وهذا يعني أنَّ أخطاء القياس المنتظمة لدرجات المفحوصين ذوي الدرجات العالية سواءً أكان إيجابياً أو سلبياً تكون أكبر منها بالنسبة للمفحوصين ذوي الدرجات المتدنية.

- إنَّ درجات الخطأ لمفحوص ما، خضع لاختبار ما، لا ترتبط سلباً أو ايجاباً بدرجات الخطأ لنفس المفحوص أذا خضع لاختبار أخر، وهذا الافتراض منطقي في حالة توفر الظروف الجيدة لتطبيق الاختبار، أما في حالة تأثر الدرجات بعوامل مباشرة فيزيقية أو غيرها فإنَّ هذا الافتراض يصبح غير منطقي.

- الدرجة الخطأ على اختبار ما لا ترتبط بالدرجة الحقيقية على اختبار آخر، بمعنى عدم ارتباط درجة الخطأ في اختبار ما بالدرجة الحقيقية على اختبار آخر يقيس نفس السمة.

- إنَّ تصميم نماذج متعددة لاختبارات تقيس سمة واحدة ليس بالضرورة أن تتحقق خاصية التوازي ومضمون هذا الافتراض يبين شروط الاختبارات أو النماذج المتوازية في توفر الشروط التالية:

- تشابه المحتوى الذي يقيسه كل نموذج.

- تطابق توزيع الدرجات الظاهرية (الملاحظة) لكل نموذج.

- يتعلق بنماذج الاختبارات المتكافئة (Equivalent) من حيث شروط التكافؤ، حيث أنَّ اعتبار تحقق خاصية التكافؤ قد لا تعكس الواقع الحقيقي لقدرة المفحوص، وشروط التكافؤ للاختبارات هي:

- تشابه المحتوى الذي يقيسه كل من النموذجين.
- بناء الاختبارات بنفس المواصفات (العدد، الإجراءات، النوع...).
- ليس بالضرورة تساوي الدرجات الحقيقية في كلا النموذجين.
- ليس بالضرورة تساوي الوسط الحسابي للدرجات الظاهرية لكل من النموذجين.
- تساوي تباين الدرجات الظاهرية لكل من النموذجين.
- تساوي التباينات المشتركة لهذه الاختبارات.

**2. القواعد (المسلمات) التي تستند إليها نظرية القياس الكلاسيكية.**

حتى تكون للنظرية الكلاسيكية في القياس القدرة على تفسير النتائج والتحليل كان لا بدَّ وأن تكون صالحة للاستخدام باعتمادها على مجموعة من الفروض والمسلمات، التي تقوم عليها من أجل تفسير الظواهر التي ترتبط بها، وتستند نظرية القياس الكلاسيكية على أربع مسلمات رئيسية هي: (عبد الرحمن، 1998، ص.69)

**أ) يمكن قياس أداء الفرد وتقديره**

أي تحويل النمط السلوكي للفرد من الصيغة النوعية الوصفية إلى صيغة كمية باستخدام الأرقام وفقاً لقواعد معينة، ويمكن قياس الأداء المقدر من التنبؤ بالأداء في وقت لاحق (ردود الفعل)، وهذا القياس يعتمد على إخضاع الفرد إلى التجريب المرتبط بموضوع القياس وينبغي أن يراعي موضوع القياس جميع الظروف والعوامل التي قد تؤثر على أداء الفرد.

**ب) أداء الفرد هو دالة خصائصه**

بمعني أنَّ كل أداء يتم إصداره من قبل سمة أو مجموعة من الخصائص التي تميز الفرد عند قياس الأداء الذي يرتبط بخاصية من القدرة الرياضية يجب أنَّ نعرف أنَّ هذا الأداء هو نتيجة لخاصية القدرة الحسابية، إلى جانب خصائص أخرى مثل الذكاء، وهذه الخاصية تعطي أكثر من نوع واحد من الأداء، على سبيل المثال، الخاصية الرياضية هي القدرة الحسابية وحل المشاكل الرياضية لذا يصبح من الضروري أن يكون المقياس على درجة عالية من الحساسية لشدة العلاقة بين القدرة والأداء من أجل أن تكون قادرة على قياس الأداء والعودة إلى خاصية واحدة أو مجموعة الصفات، مما يعني ضرورة الانتباه إلى فحص دلالات القياس وتوضيح النتائج وتفسيرها.

**ج) الخاصية والأداء والعلاقة بينهما تختلف من فرد إلى آخر**

ويعني "الفروق الفردية" إذ أنَّ أدوات القياس النفسي تعمل على الضبط في إطار محدد ومختلف نسبياً، والاختلاف الموجود بين خصائص الأفراد ينعكس في أدائهم وهنا نركز على أداة القياس نفسها واختيار وحداتها، والتأكد من أنَّ هذه الوحدات فعالة، وتحليل وتفسير الدرجات في الإطار المرجعي الذي تنسب إليه هذه القياسات.

**د) القياس الظاهري الكلي يتكون من قياس حقيقي وآخر يرجع إلى الخطأ**

مما يشير إلى أنَّ النتيجة الإجمالية تساوي النتيجة الحقيقية + الدرجة التي تعود إلى الخطأ والتي

تصنف على النحو التالي:

- الخطأ الثابت يعود إلى نفس المقياس ويتكرر بنفس الصفة ويؤثر على جميع درجات المقياس.
- خطأ القياس يرجع إلى استخدام الدرجة الظاهرية في القياس بدلاً من الدرجة الحقيقية.
- الخطأ العرضي أو العشوائي لا يمكن السيطرة عليه أو ضبطه.
- متوسط الدرجات التي تعود إلى الخطأ العشوائي يساوي الصفر.
- معامل الارتباط بين النتائج الحقيقية والدرجات بسبب الخطأ العشوائي = صفر أي لا وجود لعلاقة بينهما.
- علامات الخطأ العشوائي إذا تمَّ تطبيق مقياس ما على الأفراد ليس لها علاقة مع درجات الخطأ العشوائي عندما يتم تطبيق مقياس آخر لنفس الأفراد.

**3. أساليب القياس الكلاسيكي**

**1.3. القياس الجماعي المرجع**

تستند المقاييس جماعية المرجع إلى فكرة تقدير الاختلافات الفردية بين الأفراد والتمييز بينهم حيث أنَّ هذه المقاييس تقرر، أنَّه لا يكون لدرجة الفرد معنى إذا لم تقارن بمعيار يستند إلى مستوى مجموعة الأقران التي ينتمي اليها الفرد، حيث أنَّ معدل أعضاء المجموعة يصبح مستوى مماثل يقاس عليه مدى انحراف درجة الفرد على المتوسط، والمعايير المحسوبة لقفرات من مجموعة الاختبار وهذه المعايير التي تتغير مع تغير المجموعة. وقد أجرى بينيه اختباراً لعزل الأطفال المتخلفين عن الطلبة العاديين من خلال اختيار البنود الأكثر تمييزاً (كاظم، 1988 أ، ص.18).

**2.3. القياس المحكي المرجع**

يركز هذا الأخير على مستوى أداء الفرد، والحجة هي أنَّ النشاط التعليمي هو نشـــاط متعمد ينبغي القيام به من أجل إتقان الطلبة لما تعلموه، وينبغي ألا يخضع توزيع الأداء للمتغيرات الطبيعية مثل الوزن والطول.

وهذا يعني أنَّ التركيز ليس على مقارنة أداء الفرد فيما يتعلق بالأقران، بل على الأحرى وفقاً للاختبار المحدد ووفقاً للأهداف الموضوعة للقياس (كاظم، 1988 أ، ص.19).

## 4.الخصائص السيكومترية للاختبار وفق النظرية الكلاسيكية:

### 1.4.مفهوم الصدق

تقول ليونا تايلر (1988) إنَّ أهم ما يؤخذ في الاعتبار هو الصدق الذي يتصل مباشرة بالسؤال الذي يقول: (ماذا يقيس الاختبار؟) وإذا لم يكن لدينا إجابة كافية بدرجة معقولة لهذا السؤال، فإنَّ أي اختبار سوف يكون عديم الفائدة (ص.50).

كما يشير علام (1995) إلى أنَّ مفهوم الصدق يشير إلى الاستدلالات الخاصة، التي تنتج عن درجات المقياس من حيث الفائدة، ومدى مناسبتها، وتحقيق الصدق في المقياس معناه تجمع الأدلة التي تؤيد مثل هذه الاستدلالات (ص.274).

بعبارة أخرى، هل يقيس الاختبار ما وضع لقياسه، أو بشكل صحيح أكثر، هل يعطي الاختبار درجة تعكس السمة أو القدرة المراد قياسها بالإضافة إلى قدرة الاختبار على التمييز بين الأفراد.

ويرتبط صدق الاختبار النفسي بالنسبية، على سبيل المثال، اختبار فهم الموضوع (TAT) هو صادق بالنسبة للمراهقين والبالغين، ولكنَّه ليس كذلك بالنسبة للأطفال، وكذلك الاختبار الذي أعد في البلدان الغربية ليس صحيحاً تماماً بالنسبة للدول العربية إلا بعد التكيف والتقنين، دون نسيان الوقت فاختبار(مينيسوتا) متعدد الأوجه، الطبعة الأولى التي أعدت في(1942)، بعد مرور خمسين سنة أصبح غير صالح، مما أُضطر إلى تجديده، أي أنَّ الصدق نسبي، استناداً إلى الفئة العمرية في المجتمع الذي تمَّ تصميم المقياس له والوقت الذي أنجزت فيه.

#### 1.1.4. معاني الصدق

يرى عبد المحسن (2019) أنَّ الصدق مفهوم واسع، وله عدة معاني تختلف بحسب استخدام الاختبار والصدق محدود بمجموعة من الخصائص والشروط التي يجب أن يضعها مصمم الاختبار في اعتباره أهما:(ص.35)

**أ.** الصدق صفة تتعلق بنتائج الاختبار، أو نتائج أداة القياس وليس بالاختبار، أو أداة القياس نفسها ولكنها ترتبط بالاختبار أو أداة القياس من قبيل الاختصار فقط.

**ب.** الصدق صفة نسبية متدرجة وليس مطلقة، لذا يجب تجنب التفكير في نتائج القياس على أنَّها صادقة أو لا، وبالتالي يمكن القول بأنَّ الصدق مرتفع، أو متوسط أو منخفض.

**ج.** الصدق صفة نوعية ترتبط دائماً باستعمال خاص، حيث لا يوجد اختبار صادق وصالح لجميع الأغراض.

**د.** الصدق مفهوم واحد فلا يوجد أنواعاً مختلفة للصدق ولكن طرقاً وأساليباً تستخدم لجمع الأدلة في هذا المفهوم هناك ثلاث طرق رئيسية تستخدم في ذلك هي:

الصدق بدلالة المحتوى، الصدق بدلالة المحك، والصدق بدلالة التكوين. (إسماعيل، 2004، ص.85)

### 2.1.4. طرق حساب الصدق

وفقا لمعايير القياس التربوي والنفسي للجمعية الأمريكية النفسية American Psychological Association، هناك ثلاث طرق لحساب الصدق حددتها معايير القياس التربوي والسيكولوجي الصادرة عن الرابطة الامريكية لعلم النفس (1985)، وتذكرها(أناستازي) وهي:(عبد الخالق، ص.52- 55).

**أ). الصدق المرتبط بالمحتوى *Content Related Validity***

ويسمى أيضاً بصدق المضمون ويعني مدى قياس الاختبار لتمثيل مجال السلوك الذي يجب قياسه عن طريق تحليل مكونات الاختبار والفقرات، ثمَّ تحديد نطاق السلوك الذي سيتم قياسه، ثمَّ التطابق بين الاختبار والسمة التي تمَّ قياسها، لتحديد إلى أي مدى يمثل الاختبار الوظيفة المطلوبة ولضمان صدق المحتوى يعرض الاختبار على متخصصين في الموضوع الذي يقيسه الاختبار للتأكد من مدى قياس المفردات للأهداف التدريسية، وتغطيتها للمحتوى التعليمي المدروس، فكلما زادت نسبة الاتفاق بين المحكمين دلَّ ذلك على ارتفاع صدق المحتوى (الكيلاني وعدس، 2011، ص.221).

ويشير علام (2015) إلى أنَّ صدق محتوى الاختبار يتعلق بالحكم على مدى كفاية مفرداته لعينة ممثلة لنطاق المحتوى، أو أهداف يفترض أنَّ الاختبار يقيسها، وهناك نوعان من الصدق يتصلان بصدق المحتوى هما:

- **الصدق الظاهري *Face Validity***

يعني مدى ملاءمة الاختبار للغرض الذي وضع من أجله، وذلك بإعطاء الاختبار لأكثر من محكم – لتفادي الذاتية- لإبداء الرأي حول درجة قياس الاختبار للسمة المقاسة، ويمكن تقييم درجة الصدق الظاهري للاختبار من خلال التوافق بين تقديرات المحكمين، فكلما كان هناك توافق كبير دل ذلك على قوة صدق الاختبار، وكلما كانت تقديرات المحكمين متدنية فهذا مؤشر على ضعف الصدق (أبو دينار، 2012).

- **الصدق العيني (المنطقي) *Sampling Validity***

يرتبط الصدق المنطقي بتصنيف الفقرات وفقاً لمجالات أداة القياس، ويتطلب هذا النوع تحديداً أدق للموضوعات الدراسية التي يغطيها الاختبار، وكلما كانت هذه الموضوعات أكثر تحديداً فإنَّه يُتَوقع أن يكون الصدق المنطقي أعلى، ويُعد جدول المواصفات خطوة أساسية في بناء الاختبارات التحصيلية لأنَّه يحصر الموضوعات ويحدد أهمية كل منها. ومن الجدير بالملاحظة أنَّ الصدق المنطقي يركز على عدد الأسئلة بينما يركز الصدق الظاهري على محتوى الأسئلة بصرف النظر على عددها أو مدى تغطيتها للموضوعات أو الأهداف التدريسية (عودة، 2002، كما ورد في أبو الديار، 2012).

**ب). الصدق المرتبط بالمحك *Criterion related validity***

تركز هذه الطريقة على مقارنة نتائج الاختبار بمحكات، وهذه المحكات قد تكون مستوى أداء للأفراد في نشاطات أخرى مثل التحصيل، أو الأداء على اختبارات أخرى.. وهذا يتم عادة بواسطة حجم الترابط بين الاختبار والمحك (عباس، 1996، ص.25). وتعتمد الطريقة العامة في حساب معامل الارتباط على الدرجات الخام مباشرة ومربعات هذه الدرجات. والمعادلة الخاصة بها:(أمطانيوس، 2015، ص.91).

$$r = \frac{n\sum xy - \sum x.\sum y}{\sqrt{[(n\sum x^2) - (\sum x)^2][(n\sum y^2) - (\sum y)^2]}} \dots\dots\dots\dots (02)$$

$(\sum x^2)$: مجموع مربعات الاختبار الأول $x$

$(\sum x)^2$: مربع مجموع درجات الاختبار الأول x

$(\sum y^2)$: مجموع مربعات الاختبار الثاني y

$(\sum y)^2$: مربع مجموع درجات الاختبار الثاني y

ويميز بعض الباحثين نوعين لهذا الصدق هما:

- **الصدق التنبؤي *Predictive Validity***

يقوم الصدق التنبؤي على أساس حساب القيمة التنبؤية للاختبار، بمعنى معرفة مدى صحة التنبؤات المتوقعة معتمدين على درجات الاختبار. ويتم حساب معامل الصدق وفق هذه الطريقة بمعامل الارتباط بين الدرجات على الاختبار ودرجات الأداء الفعلي للأفراد، ويرى (الزوبعي، وآخرون، 1981) أنَّه عند تعيين الصدق التنبؤي يجب تطبيق المقياس على عينة ملائمة من الأفراد والانتظار

لفترة من الزمن بحسب طبيعة السمة المقاسة أو الدراسة وهدف القياس، ثمَّ إجراء قياس ثانٍ للحصول على درجات المحك المعينة نفسها، وحساب معامل الارتباط بين درجات المقياس ودرجات المحك، إذ يمثل معامل الارتباط هنا معامل صدق التنبؤي (حيدر، 2017، ص.100). ويمكن توضيح خطوات هذه الطريقة بالشكل التالي:

**شكل رقم (02) يوضح خطوات تقدير الاختبار التنبؤي**

ويشير الانصاري (2000) أنَّه بعد جمع البيانات حساب الدرجات على الاختبار، نقوم بإيجاد العلاقة بينهما بإحدى الطرق الثلاث التالية:(ص.100).

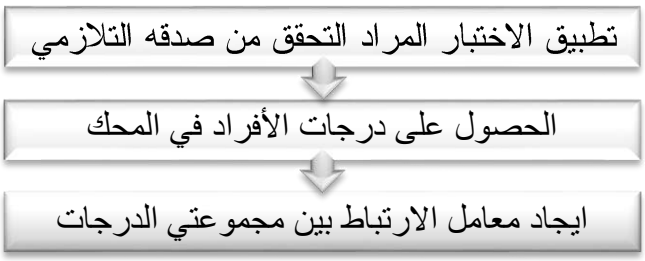
- **طريقة النسب المئوية:** حيث يتم تقسيم الأفراد وفق رتبهم في مقياس المحك إلى قسمين متقابلين كالناجحين في مقابل الفاشلين مثلاً، ومن ثمَّ حساب النسب المئويـــة من الأفراد الذين حصلوا على درجات مرتفعة أو متوسطة أو منخفضة في الاختبار في كل من المجموعتين، ليتضح ما إذا كان ارتفاع الدرجات على الاختبار يصاحبه زيادة في نسبة النجاح أو الفشل، فإذا كان الفرق جوهرياً دل هذا على أنَّ الاختبار صادق.
- **طريقة المتوسطات:** فيها يتم حساب دلالة الفروق بين درجات مجموعتين من الأفراد في الاختبار إحداهما أخذت تقديراً مرتفعاً في مقياس المحك والأخرى أخذت تقديراً منخفضاُ في مقيـــاس المحك أيضاً فإذا ثبت أنَّ هناك فرقاً دالاً بين درجات هاتين المجموعتين في الاختبار فإنَّ الاختبار صادق.
- **طريقة معامل الارتباط:** يستعمل هنا معامل الارتباط الثنائي Biserial Correlation Coefficient بين درجات الأفراد على المحك ودرجاتهم على الاختبـــار، وهذه الطريقـــة أدق من الطريقتين السابقتين لأنَّها لا تعتمد على المتوسطات أو النسب.

- **الصدق التلازمي *Concurrent Validity***

يكشف عن قدرة الاختبار لقياس أداء الفرد أو السمة، ليتم قياسها عن طريق مقارنة النتائج التي تمَّ الحصول عليها مع نتائج الاختبار الآخر التي تختبر نفس الأداء أو السمة ليتم قياسها بمعنى أنَّه يمكن حساب معامل الصدق بمعامل الارتباط بين درجات الأفراد على الاختبار ودرجاتهم في الأداء الفعلي في جوانب التي يقيسها الاختبار، بشرط أن تكون درجات الأداء الفعلية للأفراد قد تمَّ جمعها وقت إجراء الاختبار أو قبلها.

وتشير سوسن مجيد (2014) إلى أنَّ معاملات الصدق المرتبطة بمحك خارجي تتأثر بعدد من العوامل التي يجب أن تؤخذ بعين الاعتبار عند تفسير الدرجات (ص.112-113).

- يؤدي زيادة تجانس العينة إلى انخفاض معامل الصدق، وكلما كانت العينة غير متجانسة كلما زاد التباين، وبالتالي زاد معامل الصدق، ذلك أنَّ أحد المفاهيم الهامة لصدق المقياس هو قدرته على تمييز واظهار الفروق لفردية لدى الأفراد في مجال سمة أو قدرة معينة، هذا من جهة ومن جهة أخرى فإنَّ معامل الصدق هو في جوهره معامل ارتباط يتأثر كثيراً بمدى زيادة أو نقصان الفروق الفردية يجب أن يكون معامل درجة ثبات المحك وثبات المقياس عالياً، لأنَّ القيمة الحقيقية لمعامل الصدق لا يمكن أن تتجاوز مؤشر الثبات (الجدر التربيعي لمعامل الثبات).
- ذاتية واطلاع المُقِّيمين لدرجات المقياس قد تؤثر في تقديراتهم للأفراد وبالتالي يتأثر معامل الصدق.
- طول المقياس له دور كبير في زيادة معامل صدق الاختبار فنظراً لأنَّ القيمة القصوى للصدق المرتبطة بالمحك تعتمد على مؤشرات الثبات، فإنَّ القيمة تزداد بزيادة قيمة معامل الثبات، وزيادة قيمة معامل الثبات تعتمد على طول الاختبار، أي على عدد فقراته، فكلما زاد العدد زادت قيمة معامل الثبات أذا طبق الاختبار على مجموعة من الأفراد نفسها، غير أنَّ هذا التأثير يكون أقل فيما يتعلق بقيم الصدق.
- المدة الزمنية الفاصلة بين تطبيق المقياس التنبؤي ومقياس المحك لها تأثيرات، نظراً لأنَّ الصدق ينخفض بازدياد المدة الزمنية الفاصلة، وذلك لتأثير الأخطاء العشوائية في درجات كل من المقياسين.
- يفضل أن تبدأ بعينة كبيرة من الأفراد تحسباً لفقدان بعظهم عند تطبيق مقياس المحك مما يؤدي إلى انخفاض معامل الصدق. ويمكن توضح خطوات هذه الطريقة بالشكل التالي:(علام، 2006)




**شكل رقم (03) يوضح خطوات تقدير الصدق التلازمي للاختبار**

**ج). صدق التكوين الفرضي (صدق المفهوم) *Construct Validity***

يعرف صدق التكوين الفرضي أو صدق المفهوم على أنَّه قدرة الاختبار أو أداة القياس على قياس سمة أو خاصية معينة تعتبر أساساً لتحديد السلوك البشري في مواقف معينة، ويعتمد صدق المفهوم على نتائج مجموعة من التحريات هدفها إثبات أنَّ الأشخاص الذين حصلوا على درجات عالية في الاختبار تتوافق تصرفاتهم في عدد من المواقف مع ما هو متوقع من الأفراد الذين يمتلكون كمية كبيرة من هذا المفهوم، وفي الوقت ذاته تتوافق تصرفات أولئك الذين حصلوا على معدلات درجات متدنية في الاختبار مع ما هو متوقع من أولئك الذين يمتلكون كمية ضئيلة من المفهوم، ويشير صدق المفهوم إلى أنَّ الارتباط بين مقاييس المفهوم ذاته التي تمَّ الحصول عليها بطرق مختلفة(الصدق التقاربي) أعلى من الارتباط بين عدة مفاهيم تم قياسها بالطريقة ذاتها(الصدق التميزي)(أيكن، 2007، ص.79- 80).

ويشير صدق التكوين الفرضي إلى مدى الارتباط بين الجوانب التي يقيسها الاختبار وبين مفهوم هذه الجوانب، أو إلى أي مدى يكون الاختبار مقياساً للتكوين النظري؟، ويمكن حساب معامل صدق الاختبار بهذه الطريقة بتحديد معامل الارتباط بين درجات الأفراد على الاختبار وبين مفهوم هذه الجوانب، وصدق التكوين الفرضي مرتبط بما يلي:

- التحليل العاملي: طريقه إحصائيـــة تهدف إلى تفسير معاملات الارتباط الإيجابي التي لها دلالـــة إحصائية هامة بين المتغيرات المختلفة، وبمجرد أن يتم استخراج العوامل وتقسيمها، يمكن استخدامها لوصف التكوين العام للاختبار.
- الاتساق الداخلي: المحك هنا يصبح النتيجة الإجمالية للاختبار نفسه.
- الصدق التميزي: استناداً إلى الدليل على أنَّ الاختبار لا يرتبط ارتباطاً جوهرياً بالمتغيرات التي

يجب أن يختلف عنها. مثال لقياس القدرة على القراءة دون ذكر الحساب.

**2.4.مفهوم ثبات وطرق حسابه**

إنَّ مفهوم الثبات يعني حصول الفرد على نفس الدرجات في حالة ما إذا أعيد تطبيق عليه الاختبار مرة أخرى، بنفس الأداة وفي نفس الظروف تقريباً، والحقيقة أنَّ هذا الأمر صعب التحقيق بحكم أنَّنا لا نستطيع التحكم في ظروف الإجراء ولا في تغير مزاج المفحوصين ناهيك عن الخبرة التي قد يكتسبها المفحوصين، ولذلك فإنَّ الثبات في هذه الحالة يعني مقدار التباين أو التقارب بين درجات الأفراد إذا ما أعيد تطبيقه على نفس الأفراد في ظروف متشابهة. (صلاح وسليمان، 2005، ص.359).

فمفهوم ثبات الاختبار يشير إلى درجة التوافق أو الاتساق في علامات مجموعة من الأفراد عند تكرار تطبيق الاختبار أو صورة أخرى منه مكافئة له على نفس المجموعة، لذلك فإنَّ ثبات الاختبار يرتبط بـــ:

الموثوقية: أي الاعتماد عليه في الحصول على النتائج.

الاستقرار في النتائج: بمعنى أنَّه إذا قمنا بإعادة تطبيق الاختبار على نفس الفرد سنحصل على نفس النتائج أو تقريباً نفس النتيجة.

الموضوعية: أي يحصل الفرد على نفس النتائج بتغير الأخصائي الذي يطبق الاختبار ويصححه.

وتشير أدبيات القياس النفسي والتربوي لوجود عدة طرق لتقدير معامل ثبات هي:

**1.2.4. إعادة تطبيق الاختبار (معامل الاستقرار) *Test-Retest Reliability***

الفكرة الأساسية لهذه الطريقة هي تطبيق الاختبار على مجموعة من الأفراد، ثمَّ إعادة تطبيق نفس الاختبار على نفس المجموعة في نفس الظروف، بعد فترة من الوقت بين التطبيق الأول والتطبيق الثاني، ولا ينبغي أن تكون الفترة قصيرة وهذا لتجنب عوامل التذكر والاستفادة مـــن تجربـــة الأداء الأولي وتجنب عدم اهتمام المفحوص بالاختبار لأنَّه قد أجريَّ بالفعل، وتجنب طول الفترة مما يسمح بتغيرات النضج أو التاريخ ليظهر أثرها في درجات المفحوصين الحقيقية(Crocker & Algina,1986) ثم يحسب معامل الارتباط بين نتائج الإجراء الأول ونتائج الإجراء الثاني، فيكون معامل الارتبـــاط الناتج هو معامل ثبات الاختبار بطريق الإعادة، والذي تتراوح قيمته بين الصفر والواحد الصحيح

وكلما ارتفعت قيمة معامل الارتباط كلما كان الاختبار أكثر ثباتاً.

**شكل رقم (04) يوضح تقدير قيمة معامل الاستقرار**

- **عيوب طريقة إعادة الاختبار**

حدد كل من فرج وعامر (2016) عدة عيوب لطريقة إعادة الاختبار منها:(ص. 2016).

- استفادة الطلبة من تجربة الامتحان السابق، فتكون الدرجات أعلى في المرة الثانية أحياناً.
- نمو الطالب يؤثر في استجابته للاختبار في كل مرة، ويغير من موقفه وخاصة عندما تعطى فترة طويلة بين الاختبار الأول والثاني.
- عوامل أخرى مثل التعب والحالة الصحية والنفسية للمفحوص وغيرها تؤثر في كل مرة على مستوى المفحوص عند اختباره.

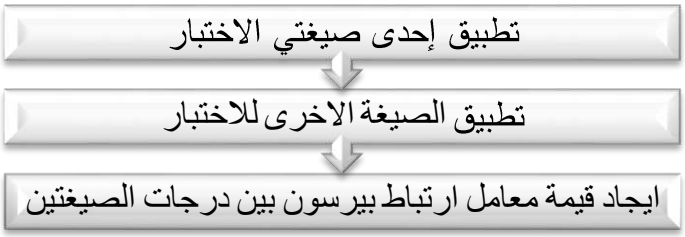
ولذلك اقترح الاخصائيون في القياس النفسي والتربوي أن تكون المدة الفاصلة بين الاختبارين ليست بالقصيرة ولا بالطويلة حتى يقل أثر النمو والتعلم.

### 2.2.4. الصور المتكافئة (معامل التكافؤ) *Equivalent forms Reliability*

في حالة صعوبة تطبيق الاختبار على نفس الأفراد مرتين، أو أنَّ السمة المقاسة تتميز بعدم الاستقرار تُعتمد طريقة الصور المتكافئة، وتَعتمد على إعداد صورتين مكافئتين للاختبار، ويشار إليه بمعامل التكافؤ، وكلما ارتفعت قيمة معامل التكافؤ كان الاختبار أكثر ثباتاً. والمقصود بالصورة المكافئة أن تكون صور الاختبار متساوية في جميع النقاط التالية:(شحاتة، 2011، ص.85).

- عدد الأسئلة لكل اختبار.
- مستوي صعوبة أسئلة الاختبار.
- تعليمات الإجراء واختيار طريقة التصحيح.

ومعامل الثبات في هذه الحالة هو معامل الارتباط بين مجموعة من الأفراد على هذه الصور المتكافئة. والشكل التالي يوضح طريقة حساب معامل التكافؤ.




**شكل رقم (05) يوضح إجراءات تقدير معامل التكافؤ**

- **عيوب طريقة الصور المتكافئة**

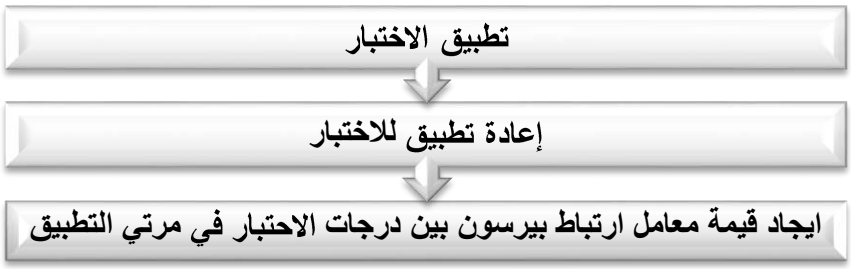
• تُجهد الباحثين بعمل اختبارين بدلاً من اختبارٍ واحدٍ.

• صعوبة وضع أسئلة متكافئة في الصورتين.

• تأثر الفرد بصياغة المفردات، فقد يجيب عن مفردة أو سؤال في الصورة الأولى ولا يجيب عنها في الصورة الثانية لعدم فهمه لصياغة السؤال.

### 3.2.4. التجزئة النصفية (معامل الاتساق الداخلي) *Split-half Reliability*

التجزئة النصفية أو التقسيم النصفي من الأساليب المستخدمة لحساب معامل الثبات، وتقسيم الاختبار يأخذ الأساليب الثلاث التالية: التربيب، أو التقسيم، (فردي– زوجي)، أو المزاوجة العشوائية والشرط لاستخدام هذا الأسلوب هو إجراء الاختبار بأكمله في وقت واحد، ثمَّ يقسم الاختبار إلى جزئيين متساويين، ثمَّ يحسب معامل الارتباط بين درجات الأفراد في الجزء الأول من الاختبار ودرجاتهم في الجزء الثاني من الاختبار. إلا أنَّ معامل الثبات المحسوب هو لنصف الاختبار فقط، لذا يمكن إجراء تعديل للقيمة المحسوبة باستخدام عدة معادلات لتصحيح معامل ثبات نصفي الاختبار (Crocker& Algina,1986)




**شكل رقم (06) يوضح تقدير قيمة معامل الاتساق الداخلي**

وهناك عدة طرق تستخدم لتصحيح معامل ثبات نصفي الاختبار نذكر منها:

**أ). معادلة سبيرمان وبراون *Spearman & Brown***

حيث يتم التعويض بمعامل الارتباط بين نصفي الاختبار لنحصل على معامل ثبات الاختبار

ككل وهذه المعادلة في صورتها المختصرة هي:

$$r_{SB} = \frac{2r_{\frac{1}{2}\frac{1}{2}}}{1 + r_{\frac{1}{2}\frac{1}{2}}} \dots\dots\dots\dots\dots\dots (03)$$

$r_{SB}$ : معامل ثبات درجات الاختبار.

$r_{\frac{1}{2}\frac{1}{2}}$ : معامل الارتباط بين نصفي الاختبار.

يتم حساب $r$ لبيرسون من العلاقة..... (01)

**ب). معادلة *Rulon***

لأنَّ معادلة سبيرمان براون تفترض أنَّ التغيرات في درجات نصفي الاختبار متساوية تماماً أي تفترض تكافؤ ثباتها، وهو افتراض يصعب توافره حتى ولو بدأ نصف الاختبار على درجة كبيرة من التكافؤ بالفعل، وعلى ذلك اقترح رولون (Rulon) تبسيط معادلة سبيرمان براون مع وضع تباين نصفي الاختبار في الاعتبار وتصاغ لذلك المعادلة التالية:(إسماعيل، 2004، ص.76)

$$r_{tt} = 1 - \frac{\sigma^2 d}{\sigma^2 t} \dots\dots\dots (04)$$

$r_{tt}$: معامل ثبات الاختبار.

$\sigma^2 d$: متوسط مربعات الفروق لنصفي الاختبار.

$\sigma^2 t$: التباين الكلي للاختبار.

**ج). معادلة جتمان *Guttman***

تستخدم معادلة جتمان للتجزئة النصفية في حالة عدم تساوي نصفي الاختبار، بمعنى إذا كان عدد فقرات النصف الأول يختلف عن عدد فقرات النصف الثاني، فإنَّه يمكن حساب معامل ثبات الاختبار بطريقة جثمان التي تصلح لحساب معامل الثبات مباشرة دون الحاجة إلى حساب معامل الارتباط، ويمكن صياغة هذه المعادلة على النحو التالي:(إسماعيل،2004، ص.77).

$$r_{tt} = 2\left(1 - \frac{\sigma_1^2 + \sigma_2^2}{\sigma_t^2}\right) \dots\dots\dots (05)$$

$r_{tt}$: معامل ثبات الاختبار.

$\sigma_1^2$ : تباين درجات النصف الأول.

$\sigma_2^2$ : تباين درجات النصف الثاني.

$\sigma_t^2$ : التباين الكلي للاختبار.

**د). معادلة هوست *Horst formula***

في بعض الحالات يصعب تقسيم الاختبار إلى نصفين متساويين، وعندها يكون استعمال المعادلات السابقة غير مجدٍ، وفي هذه الحالة تكون معادلة هوست هي المناسبة وهي كالتالي: (الطريري، 2014، ص.208)

$$H = \frac{r\sqrt{r^2 + 4p_1p_2(1-r^2)} - r}{2p_1p_2(1-r^2)} \dots\dots\dots (06)$$

r: الارتباط بين جزئي الاختبار.

$r^2$: مربع الارتباط بين جزئي الاختبار.

$p_1$: جزء الاختبار الأقل في عدد البنود.

$p_2$: جزء الاختبار الأكثر في عدد البنود.

**ه). معادلة موزير *Mosier Formula***

تقوم هذه المعادلة على أساس الاستفادة من النواتج الإحصائية للاختبار خاصة الانحراف المعياري لدرجات الأفراد الذين طبق عليهم الاختبار، وكذا تباين الدرجات، إضافة إلى معامل الارتباط بين أحد نصفي الاختبار والدرجة الكلية والمعادلة على النحو التالي:(الطريري، 1997، ص.209).

$$r_{oe} = \frac{r_{ot}\sigma_t\sigma_o}{\sqrt{\sigma_t^2 + \sigma_o^2 - 2r_{ot}\sigma_o\sigma_t}} \dots\dots\dots (07)$$

$r_{ot}$: معامل الارتباط بين البنود الفردية والدرجة الكلية.

$\sigma_t$: الانحراف المعياري لدرجات الاختبار بصورته الكاملة.

$\sigma_o$: الانحراف المعياري لدرجات البنود الفردية.

$\sigma_t^2$: تباين درجات الاختبار بصورته الكاملة.

$\sigma_o^2$: تباين درجات البنود الفردية.

### 4.2.4. طريقة الاتساق الداخلي

وتستند هذه الطريقة إلى أهمية اتساق وارتباط الفقرات ببعضها في الاختبار، وكذلك الارتباط بين جميع الفقرات مع الاختبـــار ككل، والمعادلات المستخدمـــة لحساب الاتساق الداخلي هي معادلتي

كيودر وريتشاردسون20 ومعادله الفا كرونباخ.

**أ). معادلة كيودر– ريتشاردسون20 *Kuder & Richardson***

تعرف اختصاراً(KR20) وتعتمد على توفر بيانات عن تباين كل مفردة من مفردات الاختبار وفى حالة عدم توفر هذه البيانات يمكن استخدام الصيغة رقم (21) بنفس المعادلة وهي تتميز بالسهولة والسرعة في حسابها حيث أنَّها لا تحتاج إلى معرفة تباين البنود، ولكن يعيبها أنَّها أقل دقة من الصيغ السابقة، وقد وضع كيودر ريتشاردسون شروطاً لاستخدام هذه المعادلة وهي:

- أن تكون درجة أسئلة الاختبار (صفر أو واحد).
- ألا يكون عدد الأسئلة المتروكة كبير.
- تقارب مستوى صعوبة الأسئلة.
- تساوى معاملات الارتباط بين درجات الأسئلة.

صيغة معادلة كيودر– ريتشاردسون كما يلي:

$$\rho_{KR20} = \frac{n}{n-1}\left(1 - \frac{\sum_{j=1}^{k} p_j q_j}{\sigma^2}\right) \dots\dots\dots (08)$$

$n$: عدد فقرات الاختبار.

$p_j$: عدد الإجابات الصحيحة على الاختبار.

$q_j$: عدد الإجابات الخاطئة على الاختبار.

$\sigma^2$: التباين الكلي لعدد الدرجات.

**ب). معادلة كيودر– ريتشاردسون21**

تعتبر شروط معادلة كيودر ريتشاردسون الصيغة (21) امتداد لشروط الصيغة (20) إضافة إلى شرط تساوي جميع مفردات المقياس في درجة الصعوبة، ونظراً لصعوبة تحقيق هذا الشرط كان استخدام الصيغة (20) الأكثر انتشاراً حيث يتطلب استخدامها إيجاد المتوسط والانحراف المعياري للدرجات الكلية فقط، وهو ما جعلها أكثر ملاءمة للقياس التربوي واختبارات التحصيل الدراسي (علام، 2006، ص.164).

الصيغة العاملة لمعادلة معادلة كيودر– ريتشاردسون21 هي:

$$\rho_{KB21} = \frac{n}{n-1}\left(1 - \frac{\overline{x}(n-\overline{x})}{n(\sigma^2)}\right) \dots\dots\dots (09)$$

n : عدد مفردات الاختبار.

$\overline{x}$ : متوسط الدرجات الكلية للاختبار.

$\sigma^2$ : تباين الدرجات الكلية للاختبار.

**ج). معامل الفا كرونباخ *Cronbach's alpha***

اشتق كرونباخ صورة عامة لمعادلة معامل الثبات على أساس معادلة كيودر ريتشاردسون وسماه معامل الفا(α) (الأنصاري، 2000، ص.129). ويتم فيها حساب معامل الثبات عن طريق حساب معامل الفا، وهو معامل التجانس الداخلي للمقياس، ويشير إلى الخاصية الداخلية التي يتمتع بها المقياس التي تنشأ من العلاقة الإحصائية بين الفقرات (العفون وجليل، 2013، ص.228).

والصيغة العاملة لمعادلة الفا كرونباخ هي:

$$r_{\propto} = \frac{n}{n-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma^2}\right) \dots\dots\dots (10)$$

$n$ : عدد مفردات المقياس

$\sigma_i^2$: تباين درجات كل مفردة من مفردات المقياس

$\sigma^2$: مجموع تباين درجات جميع المفردات

### 3.4. العوامل التي تؤثر على ثبات الاختبار

هناك جملة من العوامل التي تؤثر في درجة ثبات الاختبار وهي: (أبو هاشم، 2009).

- **مفردات الاختبار:** إنَّ درجة صعوبة أو سهولة مفردات الاختبار من شأنها أن تؤثر في قيم معامل ثبات الاختبار، فمفردات الاختبار الغاية في السهولة أو الصعبة جداً لا نستطيع من خلالها قياس الفروق الفردية، ففي الحالتين يكون توزيع الدرجات منتظماً، مما ينعكس على قيَّم ثبات الاختبار(علام،2006)، لذلك يفضل أن يكون الاختبار متوسط الصعوبة، ويشير المختصين في القياس أنَّ الاختبار الجيد هو ما تتراوح بنود أسئلته بين (0.25 و0.75) إلا أنَّ الافضل ما كان مستوى صعوبتها (0.50).
- **طول الاختبار:** يزيد معامل الثبات مع زيادة عدد مفردات الاختبار، وذلك لأنَّ زيادة عدد الأسئلة

يؤدي إلى محتوى أكثر شمولاً، وهذا يعني أنَّ العلاقة بين عدد فقرات الاختبار ومعامل استقراره هي علاقة مباشرة وطردية، وحيث أنَّ ثبات الاختبار يتأثر بطوله، فهذا يعني أنَّ الاختبارات القصيرة المكونة من عدد قليل من الفقرات تعطي معاملات ثبات منخفضة، في حين أنَّ الاختبارات الطويلة والمكونة من عدد أكثر من الأسئلة تعطي معاملات ثبات مرتفعة، وقد اشتق سبيرمان براون علاقة بين طول الاختبار وتقدير الثبات والتي عبر عنها بالمعادلة التالية: (إسماعيل، 2004، ص.82).

$$n = \frac{\rho_{xx'}(1-\rho_{jj'})}{\rho_{jj'}(1-\rho_{xx'})} \dots\dots\dots (11)$$

$n$: مضاعفات الطول المطلوب.

$\rho_{xx'}$: معامل الثبات المطلوب للاختبار.

$\rho_{jj'}$: معامل الثبات الحالي للاختبار قبل رفعه.

فإذا كان معامل الثبات الحالي لاختبار ما هو $(0.5)$ مثلاً وأردنا أن نرفعه $(0.8)$، فبالتعويض في المعادلة رقم $(10)$ نستطيع معرفة الطول المطلوب لفقرات الاختبار كما يلي:

$$n = \frac{0.8(1-0.5)}{0.5(1.0.8)} = \frac{0.4}{0.1} = 4$$

أي أنَّ المطلوب مضاعفة عدد بنود الاختبار إلى أربع أضعاف عن عددها الحالي.

- **زمن أداء الاختبار:** يتأثر ثبات الاختبار بدرجة كبيرة بالزمن المستغرق في الحل، فترتفع قيَّم معامل الثبات بارتفاع الزمن لأنَّه يتيح الوقت الكاف للإجابة والتأكد منها، وينخفض معامل الثبات بسبب تدخل عوامل أخرى كالتعب والارهاق والملل أي أنَّ ارتفاع معامل الثبات يتناسب طردياً مع زيادة وقت الاختبار ولكن إذا كانت المدة طويلة بشكل كبير ينخفض الثبات.

- **تباين درجات الأفراد:** ويقصد به أنَّ معامل الثبات يقل بزيادة تجانس المجموعة، وكلما ارتفع تباين الأفراد كلما ارتفع معامل الثبات، ولذلك يجب أن يصاحب معامل الثبات وصف دقيق لنوع المجموعة المطبق عليها الاختبار. وهذا يعني أنَّ معامل الثبات ينخفض عن طريق زيادة تجانس المجموعة المطبق عليها.

وبناءً على ذلك يمكن استخدام المعادلة التالية لتحديد العلاقة بين معامل الثبات للاختبار وتباين درجاته. (إسماعيل، 2004، ص.83)

$$\rho_{jj'} = 1 - \frac{\sigma_x^2(1-\rho_{xx'})}{\sigma_u^2} \dots\dots\dots (12)$$

$\rho_{jj'}$: الثبات المتنبئ به للعينة الجديدة.

$\sigma_x^2$: تباين العينة الأصلية.

$\sigma_u^2$: تباين العينة الجديدة.

$\rho_{xx'}$: تقدير ثبات العينة الأصلية.

من المعادلة يتبين أنَّ معامل ثبات الاختبار يزداد بزيادة تباين درجات المجموعة، ويقل معامل الثبات عندما يقل التباين في المجموعة، بمعنى أنَّ هناك علاقة طردية بين التباين ومعامل التباين.

- **التعليمات الخاصة بالاختبار:** تلعب تعليمات الاختبار وفهمه من طرف المفحوص أهمية بالغة فالمفحوص الذي يستوعب ويفهم هذه التعليمات قد يتحصل على درجات مرتفعة في حين يضعف أداء المفحوص الذي لم يفهم هذه التعليمات نتيجة القلق أو الاضطراب الذي أصابه أثناء موقف إجراء الاختبار.

- **مستوى القدرة في مجموعة الاختبار:** عندما يكون اختبار لمجموعة من الأفراد صعب جداً فإنَّها تلجأ إلى التخمين، كما أنَّ احتمال الدقة في الاختبار تكون منخفضة، ومن ناحية أخرى، إذا كان اختبار الفئة سهلاً جداً، فمن المتوقع أن يكون الاختبار غير فعال في التمييز بين أعضاء تلك الفئة.

- **التخمين:** ينخفض معامل ثبات الاختبار تبعاً لزيادة أثر التخمين، لأنَّ الإجابة التي تعتمد على التخمين في المرة الأولى لا تعتمد على التخمين في المرة الثانية على نفس الاختبار ونفس العينة كما يؤثر الغش أيضاً على ثبات الاختبار، ويشير اليعقوبي (2013) أنَّ التخمين في درجة الاختبار يعالج باستخدام معادلة تصحيح أثر التخمين التالية:

$$\text{س} = \text{ص} - \frac{\text{خ}}{\mathbf{1} - \text{ن}}$$

حيث أنَّ:

س: الدرجة المصححة

ص: عدد المفردات التي أجاب عنها المفحوص إجابة صحيحة.

خ: عدد المفردات التي أخطأ المفحوص في الإجابة عليها.

ن: عدد البدائل للإجابة عن المفردة.

**5. العلاقة بين الصدق والثبات**

يعتبر الصدق والثبات وجهان لعملة واحدة خاصة أنَّ كلا المفهومين يبحث في مدى كفاءة الاختبار ومناسبته للمسلمات الرئيسية لنظرية القياس. فمفهوم الثبات يبحث في مدى استقرار درجات الاختبار عندما تتغير الظروف الخارجية، ومفهوم الصدق يتعلق بتشبعه بالعوامل التي يقيسها بمعنى مدى صلاحية الاختبار في قياس ما وضع لقياسه (عبد الرحمن، 2008، ص.213).

ولأنَّ الاختبار لا يمكن أن يكون صادقاً إذا لم تتسم مفرداته بالاتساق والثبات، حيث أنَّ قيمة مؤشر الثبات تعد الحد الأدنى لقيمة معامل الصدق (علام، 2000)، كما أنَّ المقياس الصادق لا بدَّ وأن يكون ثابتاً، بينما قد يكون الاختبار ثابتاً ولكنَّه غير صادق، فاختبار ذكاء الأطفال رغم تمتعه بدرجة عالية من الثبات يكون صادقاً في قياسه لذكاء الأطفال، لكن في حالة استخدامه لقياس ذكاء الكبار فلن يكون صادقاً في النتائج المستخرجة ولا يعطي تفسيرات ودلالات صحيحة.

**6. النظرية الكلاسيكية والاختبارات**

تهدف هذه النظرية إلى تحديد العوامل المؤثرة على درجة الفرد والتي تسبب ما يسمى خطأ المقياس، ولتحقيق ذلك استندت النظرية على مجموعة من الفروض منها:

- درجة الفرد التي يحصل عليها ليس من الضروري أن تمثل درجته الحقيقية فهي قابلة للتغيير حسب الظروف والمواقف.
- زيادة خطأ القياس يقلل من الدرجة الحقيقية والعكس صحيح.
- الدرجة الحقيقية يمكن معرفتها من خلال تكرار تطبيق الاختبار عدة مرات وهي عبارة عن متوسط مرات التطبيق.
- درجة الفرد هي نتاج نوعين من الدرجات (الدرجة الحقيقية والدرجة الخطأ).
- لا يوجد اقتران بين الدرجات التي يحققها الأفراد وبين الدرجات الخاطئة.

**7.الانتقادات التي وجهت إلى النظرية الكلاسيكية.**

في كل الأحوال، فإنَّ النظرية الكلاسيكية للقياس تزودنا بمعلومات محدودة للغاية عن الثبات حيث لا يمكن استخدام المعلومات المتوفرة عن نوع ما من أنواع الثبات مثلاً (ثبات إعادة التطبيق)

للوصول إلى استنتاجات خاصة بنوع آخر مثلاً (ثبات الاتساق الداخلي)، وحتى في حالة حساب أكثر من نوع من أنواع الثبات فإنَّه يظل من الصعب استخدام هذه المعلومات مجتمعة لتقرير عدد مرات تطبيق الاختبار أو صور الاختبار، أو عبارات الاختبار التي نحتاجها للحصول على قياسات ثابتة.

ومن الملاحظات المهمة في النظرية الكلاسيكية للقياس أنَّ للمفحوص درجة حقيقية True Score واحدة فقط تمثل متوسط الدرجات التي يحصل عليها إذا طبق عليه عدد لا نهائي من الاختبارات المتوازية أو المتكافئة Parallel Measures وهي اختبارات يتماثل محتواها ويتساوى متوسطات وتباين درجاتها، وكذلك تتساوى الارتباطات البينية لدرجات مفرداتها المتناظرة والطريقة والبعد، وهذه الأوجه لها نظائرها أو ما يقابلها في النظرية الكلاسيكية للقياس.

ورغم الاستخدام الكبير لهذه النظرية من العديد من الباحثين في بناء الاختبارات وتحليل نتائجها وتفسيرها، إلا أنَّ هناك العديد من المشاكل التي تقلل من دقة موضوعية هذا الاستخدام منها ما ذكره كل من:( علام،2000 ; كاظم، 1996 ;Nitko,2001; Randall,1998 ، كما ورد في بركات، 2018).

- **تأثر الدرجة الكلية للفرد في اختبار ما بمفرداته:** حيث تكون درجة الفرد عندما يختبر بمفردات سهلة أعلى من المفردات الصعبة، فلا يمكن تقدير قدرته فيما تقيسه هذه المفردات تقديراً دقيقاً، لذا تختلف نتيجة القياس باختلاف الاختبار المستخدم.

- **تأثر خصائص مفردات الاختبار بقدرة الأفراد:** حيث تختلف معاملات الصعوبة أو السهولة والتمييز لمفردات الاختبار باختلاف قدرة أفراد العينة، فالمفردة التي يختبر بها أفراد ذوي قدرات مرتفعة تبدو سهلة بينما تبدو نفس المفردة صعبة لذوي القدرات المنخفضة، وإذا كانت العينة متجانسة نسبياً، فإنَّ قيم معاملات التمييز تكون أقل من القيم التي يمكن الحصول عليها.

- تقتصر الموازنة بين الأفراد في السمة أو القدرة التي يقيسها الاختبار على تطبيق نفس مفردات الاختبار أو مجموعة مفردات مكافئة أو موازية لها على كل فرد من الأفراد، وبالتالي لا يمكن الموازنة بين مستويات القدرة أذا أجاب الأفراد على مفردات مختلفة ومتباينة في صعوبتها.

-**عدم وجود وحدة قياس ثابتة:** حيث لا تحدد مواضيع القياس على متصل المتغير بصورة خطية فاعتماد درجات الأفراد على مفردات الاختبار قد يؤدي إلى اختلاف المسافة بين كل درجتين متتاليتين ويؤدي إلى اختلاف المعنى الكمي لأي فرق محدد عبر مدى درجات الاختبار.

-**تأثر ثبات الاختبار بالموقف الاختباري:** حيث يعتمد ثبات الاختبار على تطبيق الصورة الاختبارية مرتين على أفراد العينة، أو على أعداد صور متكافئة من الاختبار ويعد هذا في الواقع أمراً صعباً وبالرغم من أهمية ذلك، إلا أنَّه غير كافٍ، حيث يمكن أن يختلف الموقف الاختباري وظروف التطبيق في هاتين المرتين، الأمر الذي يؤثر على دقة الاختبار وثباته.

- تساوي تباين أخطاء القياس لجميع أفراد العينة موضع الاختبار وهذا بالرغم من أنَّه قد يكون أداء بعض الأفراد على الاختبار أكثر اتساقاً من غيرهم من الأفراد، التي يقيسها الاختبار.

- لا تقدم هذه النظرية تفسيراً سيكولوجياً يوضح كيف يحاول الفرد إجابة إحدى مفردات الاختبار على الرغم من أنَّ هذا التفسير يعد ضرورياً ولازماً للتنبؤ بخصائص الدرجات المستمدة من مجتمع معين أو مجتمعات مختلفة من الأفراد، أو من أجل تصميم اختبارات تتميز بخصائص سيكومترية معينة تناسب مجتمعاً من الأفراد.

- إنَّ تكوين مفردات الاختبار ومعناها تتغير بتغير عامل الزمن، أي بمضي الزمن بالنسبة لعينة الأفراد الذين أُعد لهم الاختبار، فالظروف البيئية تتغير، والظروف الاختبارية ليست دائماً مقننة كما أنَّ حذف أو تغيير أي مفردة من مفردات الاختبار يؤدي إلى تغيير في درجات الأفراد، وهذا التغيير يصعب التنبؤ به.

- جميع خصائص الاختبارات التي تستند في بنائها على أسس النظرية الكلاسيكية، مثل معامل الصعوبة والتمييز والثبات، تعتمد على خصائص عينة الأفراد التي يجري عليها الاختبار، وعلى عينة مفردات الاختبار التي يشتمل عليها الاختبار، وهذا بدوره له أثار سلبية ومشكلاته متعددة في الوصول إلى مستوى من الدقة والموضوعية في القياس.

**8. أوجه قصور النظرية الكلاسيكية**

رغم استخدام العديد من المختصين في القياس النفسي، والتربوي، للنظرية الكلاسيكية في بناء الاختبارات، وتحليل البياناتّ، المستمدة منها، الا أنها واجهت العديد من جوانب القصور والانتقادات منها ما ذكره كل من علام، 2001 ;Allen & Yen, 1979 ,Hambleton & Swaminathan, 2009) كما ورد في بركات ،2018).

- أنَّ معالم الفقرات الشائع استخدامها (الصعوبة والتمييز) تعتمد على عينة المفحوصين المستخدمة

للحصول على هذه المعالم، وأنَّ طرق تحليل الاختبارات المبنية على النظرية الكلاسيكية والمفاهيم السيكومترية المرتبطة بها، مثل معاملات الصعوبة والتمييز، تختلف باختلاف خصائص العينة المستخدمة في حساب هذه المعاملات، إذاً فإنَّ معلم الصعوبة ليس من المعالم المستقرة لأنَّه مرتبط ارتباطاً مباشراً بطبيعة العينة، كما أنَّ معامل التمييز(ارتباط البند بالدرجة الكلية) يتأثر بتباين استجابة العينة، لذا فإنَّ عملية الاعتماد على معامل الصعوبة ومعامل التمييز لا تعتبر من الجوانب الجيدة نظراً لأنَّ هذين العاملين مرتبطان إلى حد كبير بطبيعة وخصائص العينة المستجيبة، وبذلك فإنَّه من الممكن الحصول على معاملات تمييز وصعوبة مختلفة من عينة لأخرى تبعاً لاختلاف خصائص هذه العينة.

- عند تفسير درجات الاختبار فإنَّ درجات الأفراد في اختبار ما تعتمد على عينة الفقرات التي يحتويها الاختبار، فدرجاتهم تختلف باختلاف عينة فقرات الاختبار.

- إنَّ النظرية الكلاسيكية تفترض تساوي تباين أخطاء القياس لجميع الأفراد الذين يطبق عليهم الاختبار.

ولكن من الملاحظ في بعض الأحيان أنَّ بعض الأفراد يكون أداؤهم في الاختبار أكثر اتساقاً من غيرهم من الأفراد، وأنَّ هذا الاتساق يختلف باختلاف مستوى قدرتهم. ولهذا فإنَّه ربما يتوقع أنَّ الأفراد من ذوي القدرات المرتفعة يكون أداؤهم في صورة موازية من صور الاختبار أكثر اتساقاً وأكثر تجانساً من أداء الأفراد ذوي القدرات المتوسطة. فمن الطبيعي أن يزيد الخطأ في الاختبار الصعب المطبق على مجموعة من أصحاب القدرات المنخفضة، ويقلل الخطأ أذا طبق على أصحاب القدرات المرتفعة.

- إنَّ أداء الأفراد على الاختبار يختلف باختلاف مفردات الاختبار التي تمَّ سحبها من مجتمع المفردات الكبير، فهذه المفردات قد تختلف في صعوبتها وهذا الاختلاف في صعوبة المفردات سينعكس في نهاية المطاف على الأداء المختلف للأفراد من اختبار لآخر، وقد ينتفي وجود الاختلاف في أداء الأفراد إذا تضمنت الاختبارات نفس الفقرات أو بنوداً مكافئة لها من حيث مستوى الصعوبة.

- يعرف ثبات الاختبار وفــق هذه النظرية إما باستخــدام الاختبارات المتوازيــة التي من الصعب الحصول عليها عملياً، أو إعادة الاختبار والتي تتأثر باستجابات الأفراد بالتطبيق الثاني بعوامل متعددة مثل النسيان وتعلم مهارة جديدة بالإضافة لتغير مستوى القلق والدافعية لديهم، ونتيجة ذلك ظهرت النظرية الحديثة في القياس، والتي جاءت لمعالجة القصور في لنظرية الكلاسيكية.

- تتطلب النظرية الكلاسيكية نماذج متكافئة تماماً للاختبار الواحد، وهذا متطلب صعب في التطبيق العملي مما يقلل من قيمة النتائج المبنية وفق هذه النظرية.

**المحور الثالث: نظرية القياس الحديثة *Item Response Theory***

تُعد نظرية القياس الحديثة أو ما تعرف باسم نظرية استجابة المفردة أو نظرية السمات الكامنة من أهم التطورات الحديثة في مجال القياس النفسي والتربوي بسبب ما قدمته من طرق سيكومترية ذات فعالية كبيرة في بناء المقاييس النفسية والتربوية، وطريقة تفسير درجات الأفراد على هذه المقاييس مقارنة بنظرية القياس الكلاسيكية.

وقد ساعد ظهور نظرية الاستجابة للمفردة على تطوير أساليب القياس النفسي والتربوي في كافة جوانبه بالإضافة أنَّها قدمت أساساً نظرياً جديداً للقياس النفسي والتربوي من خلال تقويم دقة وكفاءة القياس.

فنظرية القياس الحديثة، هي الاتجاه المعاصر في القياس النفسي والتربوي لأنَّها تفترض أنَّ أداء الفرد في اختبار ما يمكن تفسيره أو التنبؤ به من خاصية معينه من هذا الأداء يسمي السمات Traits فعلى سبيل المثال، عند قياس سمة عقلية مثل الذكاء، لا يمكن رؤية الذكاء أو قياسه مباشرة لأنَّه مفهوم تجريدي، ومع ذلك، فانَّه يستدل من رد الفرد في اختبار يقيس سمه الذكاء. هذه النظرية تحاول تقدير درجات الأفراد في هذه السمات، نظراً لصعوبة مراقبة هذه السمات مباشرة أو قياسها مباشرة، ولا بدَّ من التقدير، أو يستدل عليها من ردود الأفراد الملاحظة على مفردات الاختبار التي يقيس السمة أو القدرة المطلوبة، هذا هو السبب في أنَّها تسمى "نظرية السمات الكامنة" أو نظرية الاستجابة للمفردة. (عماد، 1991، ص.447).

وترى النظرية الحديثة أنَّه يمكن تحديد سمة البعد الكمي بمواقف الأفراد الذين لا يكون موقف الفرد منهم مرهوناً ببعد السمة على الصفات، أي من العينات التي ينتمي اليها الفرد، وتهدف إلى تقدير القدرة أو السمة لتكون مستقلة عن العينة، وتركز هذه النظرية أيضاً على فقرة الاختبار، وهذا يسمح بإضافة الفقرات أو حذفها أو تعديلها دون أن يتأثر الاختبار ككل بما يخالف النظرية الكلاسيكية.

ويمكن تحديد القواعد التي تعتمد عليها نظرية الاستجابة للمفردة كما يلي: (إسماعيل، 2007)

- الدرجات المختلفة أو أنماط الاستجابة تؤدي إلى اختلاف الخطأ المعياري للمقياس.
- يمكن أن يكون الاختبار الأقصر أكثر استقراراً من الاختبار الأطول.
- عندما تختلف مستويات صعوبة الاختبار بين الأفراد تكون المقارنة بدرجات الاختبار في أشكـــال

متعددة أفضل.

- يمكن الحصول على تقديرات غير متحيزة لخصائص المفردة من العينات غير الممثلة للأفراد.
- درجات الاختبار لها أهمية عند مقارنتها بمسافتها عن المفردات.

**1.الفكرة الأساسية لنظرية القياس الحديثة**

تتمحور الفكرة الأساسية لنظرية الاستجابة لفقرة الاختبار في محاولتها لاشتقاق القيَّم المقدرة للسمة أو السمات المتضمَّنة في مجموعة من الردود لاختبار الفقرات من خلال النماذج التي تعمل على تفسيرها، ولذلك فإنَّ هذه النماذج هي نماذج احتمالية، وكلها تستند إلى دالة الترجيح اللوغاريتمي بدلاً من دالة الكثافة الاحتمالية الاعتدالية التي تعتمد عليها النظرية الكلاسيكية، وتشرح النماذج رد الفرد على اختبار الفقرات التي تتضمن سمة تقاس بالفقرات تجاوز حدود مجموع درجات الاختبار المرجحة، ويتم اختيار نموذج البيانات المناسب وفقا لمدى الوفاء بهذه الافتراضات في البيانات (علام، 2001، ص.686).

**2.مفهوم نظرية الاستجابة للمفردة**

نظرية استجابة للفقرة أو المفردة هي وسيلة لنمذجة البيانات، أي أنَّها تحاول نمذجة العلاقة بين المتغير غير الملاحظ(قدرة الفرد)، واحتمال استجابة الفاحص بشكل صحيح لفقرة ما(المتغير الملاحظ) وتتوقف صحة طرق نظرية استجابة الفقرات إلى حدٍ كبيرٍ على البيانات المستمدة من الأداء الفردي وتستند نظرية استجابة الفقرة إلى وجود اثنين من المتغيرات في علاقة رياضية وتمثل نموذجاً احتمالياً وذلك لأنَّ طبيعة البيانات تسمح بالتنبؤ باحتمال النجاح في أي من الفقرات المدرجة وتوفر تقديرات القدرات لوحدات فردية من النموذج المستخدم(إسماعيل، 2007، ص.13).

وعندما يُراد قياس السمة الذهنية على سبيل المثال، عندما يلاحظ الطفل لا يُرى ذكاءه، لأنَّ الذكاء هو مفهوم مجرد، لكنَّه يلاحظ أسلوب الطفل في الحالات التي تتطلب الذكاء، أي أنَّ السمات العقلية هي التكوينات التي من المفترض أن تشكل مجموعة متماسكة من السلوك والتي يمكن ملاحظتها وتسجيلها في حالات الاختبار المقنن (علام، 2000، ص.21-22).

والفكرة الأساسية لنماذج نظرية الاستجابة للمفردة هي تتلخص في ربطها خصائص الفقرات بالبارامتر المحدد لمقدار السمة التي يمتلكها الفرد، واحتمال أنَّه سيجيب على فقرة معينة بمستوي معين

من الإجابة تبعاً لنوع الفقرة، ويستند هذا الارتباط على تحديد موقع الفرد على مقياس السمة بغض الاعتبار عن مجموعة إجاباته الفردية على الفقرات المختلفة (عودة، 1992).

**3. مسلمات نظرية الاستجابة للفقرة**

تستند نظرية استجابة الفقرة على فرضية أساسية هي أنَّ القيمة الاحتمالية لاستجابة الفرد لفقرة اختبارية ما، هي وظيفة دالة لكل سمة أو قدرة التي من المفترض أن الاختبار يقيسها في فرد وخصائص الفقرة التي يحاول الإجابة عليها، أي إنَّها تفترض أنَّ هناك دالة احتمال أن يربط بارامترين واحد يتعلق بالفرد والآخر يتعلق بالفقرة التي يتم اختباره فيها، وتهدف النظرية إلى الوصول إلى القيم المقدرة لكل من هذين البارامترين (Linden & Hambleton, 1997, p 5).

ومن الجدير بالذكر أنَّ نظرية الاستجابة للمفردة تستند إلى بعض المسلمات التي يجب على كل مستخدم للنظرية أن يكون على بيِّنة بها، أو أحد النماذج المتعلقة بها، وأهم هذه المسلمات هي: (الوليلي، 2002، ص.47-48).

- يمكن التنبؤ بأداء المختبِر في أي اختبار من خلال مجموعة من العوامل التي تسمي سمات أو القدرات الكامنة.

- يمكن وصف العلاقة بين أداء المختبِر على أي فقرة اختبارية ومجموعة من السمات أو القدرات المحتملة التي يفترض أن تؤثر على أدائه على هذه الفقرة بدالة تزايدية (تزايد طردي)، وتسمي دالة خصائص الفقرة، لأنَّ هذه الدالة تحدد المختبرين الذين حققوا درجات عالية في السمات التي لديها توقعات احتمال عالية للإجابة الصحيحة على الفقرة من المختبرين الذين حققوا درجات منخفضه على السمات.

- يعرف بأنَّ المنحنى المميز للفقرة، ولكل فقرة في الاختبار منحنى مميز لها، واحتمال الإجابة الصحيحة يقترب من الصفر عند مستويات منخفضة من القدرة.

- يزيد احتمال الإجابة الصحيحة عند مستويات عالية من القدرة حتى يصل احتمال الإجابة الصحيحة إلى الحل الصحيح (Baker, 2001, p. 7).

ومفهوم المنحني المميز للفقرة مفهوم هام في نماذج استجابة الفقرة التي تعالج سمة كامنة ومستمرة يتم قياسها باستخـــدام عناصر اختبار النوع الثنائي أي يتطلب إجـــابة صحيحة أو إجابة خاطئة، هذا

المنحنى هو الدالة الرياضية التي تربط احتمال نجاح الفرد في الإجابة على الفقرة، وبين السمة أو القدرة المقاسة بمجموعة من الفقرات أو باختبار معين. أي أنَّها دالة انحدار درجة الفقرة على السمة الكامنة التي تقاس بالاختبار، وهذه الدالة غير خطية (علام، 2000، ص.690).

**4. افتراضات نظرية الاستجابة للمفردة**

تتميز نظرية استجابة المفردة بالعديد من الافتراضات التي يجب أن تكون متوفرة في بيانات الاختبار حيث يتم فيها اختيار نموذج البيانات المناسب وفقاً لمدى الوفاء بهذه الافتراضات في البيانات وهذا يؤدي إلى دقة القياس والاحكام والقرارات اللاَّحقة. وفيما يلي شرح للافتراضات الأساسية التي تستند إليها نماذج نظرية استجابة الفقرة:

**1.4. أحادية البعد *Unidimensionality***

إنَّ أحادية البعد تشير إلى الافتراض بأنَّ الاختبار يقيس قدرة كامنة واحدة، وهي تعني أنَّ التغاير بين الفقرات يمكن أن يفسر بواسطة بعد واحد يقف خلف ذلك التغاير، بمعنى أن يقيس الاختبار قدرة كامنة واحدة، وهذا يتفق مع أنَّ المفحوص ذو القدرة العالية يفترض أن تكون احتمالية إجابته على الفقرة بشكل صحيح عالية أيضاً.

ويشير(دودين) إلى نقطة تطبيقية في غاية الأهمية، وهي أنَّ العديد من الاختبارات عادة ما تحتوي على عدة مجالات، والسؤال كما يثيره دودين: كيف ندعي عندئذٍ في مثل هذا الوضع أنَّ الاختبار أحادي البعد في حين أنَّه يقيس عدة مجالات؟، وأنَّ أحادية البعد يشير إلى الافتراض بأنَّ الاختبار يقيس قدرة محتملة واحدة، ويعني ذلك أنَّ التباين بين الفقرات يمكن تفسيره ببعد واحد وراء هذا التباين (دودين، 2004).

وخلاصة القول هي أنَّه وفقاً للنظرية الحديثة التي تفترض أحادية البعد، أنَّ الاختبار يقيس امكانية واحدة، وهذا يتسق مع حقيقة أنَّه من المفترض أن يكون لدى المفحوص ذو القدرة العالية احتمالاً كبيراً بالإجابة على الفقرة بشكل صحيح أيضاً. والعكس صحيح (كاظم، 1988 أ).

كما أنَّ هناك نقطة مهمة جداً عند التطبيق، وهي أنَّ العديد من الاختبارات عادة ما تحتوي على عدة مجالات، والسؤال كما يثيره دودين أيضاً: كيف إذاً ندعي في مثل هذه الحالة أنَّ الاختبار هو أحادي البعد في حين أنَّه يقيس العديد من المجالات؟

ويُجيب دودين (2004) على هذا التساؤل فيقول: "والحقيقة هي أنَّه لا يوجد تناقض بالنظر إلى أنَّ هناك عاملاً أساسياً مشتركاً بين هذه المجالات، وهي قياس السمة التي سيتم قياسها".

ثم يضيف دودين: 'من المهم التفريق بين أحادية البعد النظرية والعملية، نظرياً، تتكون المسائل الحسابية من أربع مكونات أو أبعاد، على سبيل المثال: الجمع والطرح والضرب والقسمة. يمكن للتحليل الشامل للاختبار الحسابي إظهار هذه الهياكل أو الأبعاد الأربعة. ولكن في الممارسة العملية نفترض أنَّ هذه الهياكل الأربعة تقيس بالفعل موضوعاً واحداً (القدرة الحسابية). وفي الواقع، من الصعب العثور على الاختبار الحسابي الذي لا يجمع بينهم. (ص.115).

### 2.4. الاستقلال المحلي *Local Independence*

يري الدوسري (2001) أنَّ السمة المميزة للاستقلال المحلي تعني أنَّ الإجابة على فقرة ما تكون مستقلة إحصائياً عن الإجابة على أي فقرة أخرى في الاختيار (الدوسري، 2001، ص.5).

والمفهوم الأكثر دقة كما هو محدد من قبل هامبلتون وآخرون: أنَّه "عندما تقوم بتثبيت القدرة المسؤولة عن أداء المفحوصين للاختبار، فإنَّ ردود الاستجابات على أي زوج من الفقرات تكون مستقلة إحصائياً " (Hambleton, Swaminathan, & Rogers, 1991, p. 10).

ويعتقد(اورلاندو) أنَّ هذا الافتراض، من الناحية الفنية، هو امتداد للافتراض الأوسع وهو أحادي البعد، وهذا يعني أنَّه إذا تمت إزالة تأثير العامل أو البعد أو السمة الكامنة التي خلف المقياس، فلن يكون هناك اختلاف منتظم إضافي بين الفقرات، وبالتالي يمكن أنَّ يظهر الاستقلال الذاتي المحلي بين المجموعات الفرعية من الفقرات التي تسير في اتساق واحد (Orlando, n.d. كما ورد في زكري، 2007)

### 3.4. عامل السرعة في الإجابة *assumption of speediness*

هناك فرضية أساسية عامة لكل نماذج من نظرية الاستجابة للفقرة المستخدمة، وهي أنَّه لا يتم تطبيق الاختبار الذي يسعى النموذج لتطبيقه تحت ظرف السرعة، بمعنى أنَّ الأشخاص الذين فشلوا في الإجابة على عناصر الاختبار لم يكن إخفاقهم بسبب الوقت(السرعة) بما يكفي لإكمال الاختبار ولكن ذلك بسبب قدراتهم المحدودة (علام، 1986).

ويرى هامبلتون وسوامينثان، إلى أنَّه نادراً ما يكـــون هذا الافتراض مرجعاً، لأنَّه قد يكـــون

متضمناً في افتراض أحادي البعد، فعندما تؤثر السرعة في الأداء على الاختبار، فإنَّ هذا يعني أنَّ هناك سمتان على الأقل تؤثران على الأداء، وهما: السرعة في الأداء، والسمات الأخرى التي يسعى الاختبار لقياسها، هي الافتراضات السابقة التي تمت مناقشتها، هي الافتراضات العامة لجميع نماذج نظرية الاستجابة للفقرة(Hambleton & Swaminathan, 1985) .

### 4.4. تساوي مؤشرات التمييز *Equal Discrimination Indices*

يرى الدوسري (2004) أنَّ المشاكل المرتبطة بالتحقق من فرضيات النظرية مسألة اختيار النموذج الصحيح لاختبار تحليل البيانات، بمعنى، هل يتم استخدام نموذج ثلاثي أو ثنائي أو أحادي المعلمة؟ (الدوسري، 2004).

فالخطوة الأولى في تحديد أنواع المعالم التي يمكن تضمينها في النموذج تعتمد على التحليل النظري الكلاسيكي، على سبيل المثال: معلمه التمييز (الميل Slope) في نظرية الاستجابة للفقرة والتي يرمز لها بالرمز(ai) توازي ارتباط الفقرة بالدرجة الإجمالية في النظرية الكلاسيكية.

والمقصود بتساوي مؤشرات التمييز، أن تكون كل المنحنيات المميزة لفقرات الاختبار متوازية (متساوية الميل) وهي تختلف عن بعضها البعض في نقاط تقاربها مع المحور الأفقي الذي يمثل استمرارية القدرة أو السمة الكامنة.

### 5.4.الحد الأدنى من أثر التخمين *Minimal Guessing*

أما معلمة التخمين فيمكن التحقق مما إذا كان من المناسب ادراجها في النموذج المقترح وذلك يمكِّن استخدام النموذج (3PLM) حيث إنَّ نموذج ثلاثي المعلمة له مواصفات خاصة بالتخمين وهو مقدار اقتراب المنحني من نقطة الصفر.

وهذا يعني أنَّ معلمة التخمين يجب أن تكون جزءاً من النموذج، خاصة إذا كانت الفقرات هي من نوع الاختيار من متعدد، فإنَّ إمكانية التخمين تزداد بشكل عام(Wiberg, 2004) .

## 5.الخصائص السيكومترية وفق النظرية الحديثة

### 1.5. معلمة الصعوبة *Difficulty Parameter*

مفهوم الصعوبة في نظرية الاستجابة للمفردة يعبر عن نقطة على متصل السمة الكامنة المتوقع عنها حدوث احتمال إجابة الفرد إجابة صحيحة عن المفردة بدون تخمين (0.50) (علام، 2007،

ص.224) و يرى بيكر أنَّ الصعوبة تتراوح دراجاتها من(+∞) إلى(−∞) إلا أنَّه لاعتبارات علمية عادة ما تحدد مدى القدرة من (−3) إلى (+3) (Baker , 2001, p. 5) .

### 2.5. معلمة التمييز *Discrimination Parameter*

تعبر معلمة التمييز عن مدى قدرة الفقرة على التمييز بين المختبرين ذوي القدرات التي تقع دون موقع الفقرة أو البند على متصل القدرة، وأولئك ذوي القدرات التي تقع فوق موقع هذه الفقرة وترتبط قيَّم معلم التمييز بميل منحنى خصائص الفقرة، وتتراوح قيمته من(صفر) إلى (+ ∞) حيث تقع قيَّم التمييز المقبولة ضمن الفقرة بين (صفر إلى 2)، وبزيادة ميل المنحنى تزداد قيمة معلم التمييز (حامد، 2008، ص.19).

### 3.5. معلمة التخمين *Guessing Parameter*

هو احتمال الوصول للإجابة الصحيحة عن طريق التخمين فقط، ويتراوح قيمة التخمين من (الصفر) إلى (+ 1) ولكن في الممارسة العملية فإن قيمته تتراوح من(الصفر) إلى (0.35) ( Baker, 2001,p.28.)

### 4.5. معلمة القدرة *Ability Parameter*

هي قدرة الفرد المقدرة باللوجيت وهي اللوغاريتمي الطبيعي لمرجع نجاح الفرد على الفقرات التي تعتبر النقطة(صفر) التدريج عن صعوبتها(Wright & Stone,1979,p.17) ويرى بيكر أنَّ الصعوبة تتراوح درجاتها من (+∞) إلى (− ∞)، ولكن عملياً لا يتجاوز المدى (−3) إلى (+3).

## 6.دالة معلومات الفقرة والاختبار في نماذج نظرية الاستجابة للمفردة

إنَّ استخدام دالة المعلومات في نماذج نظرية الاستجابة للمفردة تفظي إلى اختبارات أفضل كونها توفر العديد من المعلومات حول السمات الكامنة التي يتم تقديرها، إضافة إلى التأكد من دقة تقدير معالم الفقرة، ومعلمة القدرة، وبالتالي تبرز أهمية الفقرة من خلال مساهمة كل فقرة بدالة معلومات الاختبار ويقدم (بيكر) دالة معلومات الفقرة لكل نموذج كالتالي(Baker,2001) .

1. دالة المعلومات للفقرة وفق نموذج اللوجيستي أحادي المعلمة:

$$Ii(\theta) = [Pi(\theta)Qi(\theta)] \dots\dots\dots (01)$$

2. دالة المعلومات للفقرة وفق النموذج اللوجيستي ثنائي المعلمة

$$Ii(\theta) = a^2[Pi(\theta)Qi(\theta)] \dots\dots\dots (02)$$

3. دالة المعلومات للفقرة وفق النموذج اللوجيستي ثلاثي المعلمة

$$Ii(\theta) = a^2\left[\frac{Q(\theta)}{Pi(\theta)}\right]\left[\frac{Pi(\theta - c)^2}{(1 - c^2)}\right] \dots\dots\dots (03)$$

4. دالة المعلومات للفقرة وفق النموذج اللوجيستي رباعي المعلمة

$$Ii(\theta) = \frac{a^2[Pi(\theta) - c]^2[d - Pi(\theta)]^2}{[(d - c)^2 Pi(\theta)][1 - Pi(\theta)]} \dots\dots\dots (04)$$

حيث أنَّ:

$Ii(\theta)$ : دالة معلومات الفقرة $(i)$

$Pi(\theta)$: احتمال أنَّ يجيب المفحوص ذو القدرة $(\theta)$ عن الفقرة $(i)$ إجابة صحيحة وتختلف باختلاف النموذج المستخدم(Loken & Rulison, 2010) .

أما دالة معلومات الاختبار تستخرج بالاعتماد على دوال معلومات الاختبار كاملاً حيث تعطى بالمعادلة التالية:

$$Ii = \sum_{i=1}^{N} Ii(\theta) \dots\dots\dots (05)$$

حيث:

$Ii(\theta)$: دالة معلومات الفقرة $(i)$

$I(\theta)$: دالة معلومات الاختبار.

إن دالة معلومات الاختبار تسهم في الوصول إلى الخطأ المعياري في التقدير وبالتالي فهي تشير إلى مقدار يتناسب عكسياً مع الخطأ المعياري في التقدير، وهذا يعد مؤشر على ثبات الاختبار لأنَّه بزيادة الخطأ المعياري في التقدير يقل الثبات والعكس، ويصبح ذلك من خلال الصيغة الآتية:

$$S.E(\theta) = \frac{1}{\sqrt{I(\theta)}} \dots\dots\dots (06)$$

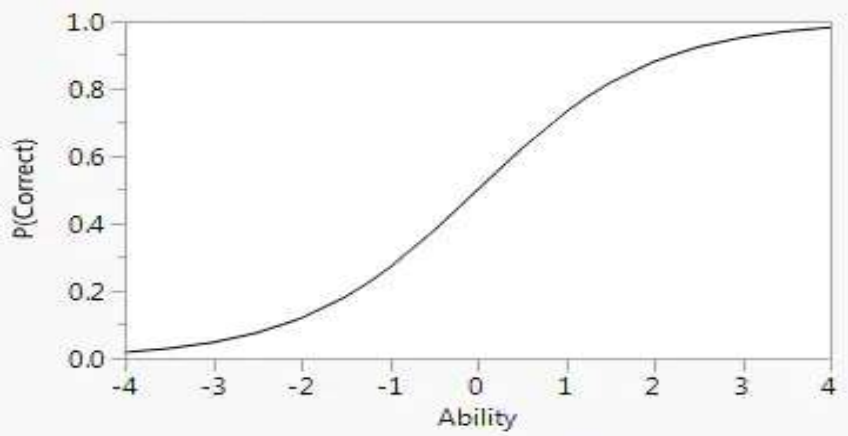
$$I(\theta) = \sum_{i=1}^{N} Ii(\theta)$$

$$Ii(\theta) = 1/(S.E(\theta))^2$$

$S.E(\theta)$: الخطأ المعياري في تقديرات القدرة (Hambleton & Swaminathan,1985)

**7.السمات الكامنة والمنحنيات المميزة للمفردة**

يعد المنحنى المميز للفقرة أحد المفاهيم المركزية لنظرية الاستجابة للفقرة Item characteristic Curve (ICC) وهو احتمالية الاستجابة الصحيحة على الفقرة كدالة للسمة الكامنة ويرمز لها بالرمز ($\theta$) ففي معظم تطبيقات نظرية الاستجابة للفقرة نفترض أنَّ المنحني المميز للفقرة يأخذ الشكل(S) ويبين المنحنى أنَّه بزيادة قيمة السمة الكامنة تزداد احتمالية الاستجابة الصحيحة على الفقرة، وتكمن أهمية المنحنى المميز للفقرة مقارنة بإحصائيات الصعوبة والتمييز أنها تسمح برؤية أنَّ احتمالية الاستجابة الصحيحة على الفقرة تعتمد على درجة السمة درجة الكامنة(كروكر والجينا،2009، ص.451).

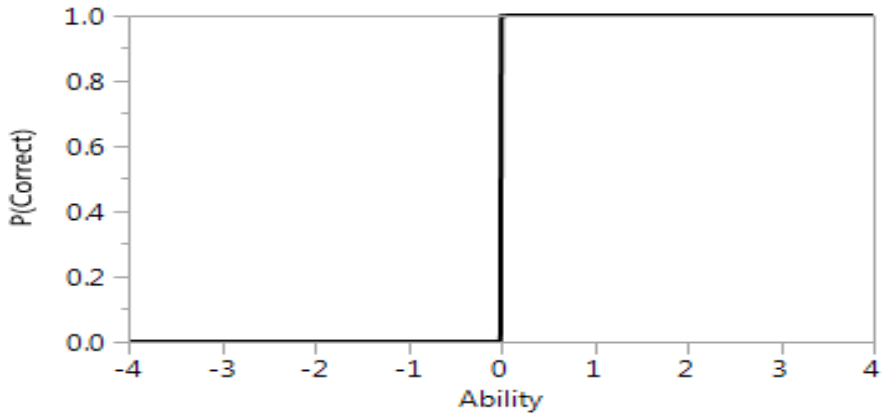



**شكل رقم (07) يبين العلاقة بين القدرة والاستجابة على الفقرة**

وتشير كل من كروكر والجينا(2009) أنَّه من الأهمية التفسير الصحيح لاحتمالية الاستجابة الصحيحة على الفقرة، وقد قدَّم لورد(Lord, 1980) تفسيرين مقبولين لهذه الاحتمالية، فالتفسير الأول يجب تسمية مجموعة فرعية من المفحوصين عند كل نقطة على تدريج السمة الكامنة والخصائص المحددة لكل مجموعة فرعية هي الأفراد الذين لديهم درجة السمة الكامنة نفسها، وتفسر احتمالية الإجابة الصحيحة عندئذٍ على أنَّها احتمالية الإجابة الصحيحة لمفحوص اختير عشوائياً من المجموعة الفرعية المتجانسة ويفترض أنَّ المنحنى المميز للفقرة يشير إلى أنَّه عندما تكون(θ)=2 فإنَّ احتمالية الإجابة الصحيحة على الفقرة تكون(0.87)، ويفسر هذا على أنَّ احتمالية الإجابة الصحيحة لمفحوص اختير عشوائياً من المجموعة الفرعية المتجانسة يساوي(0.87)، ويشير التفسير الثاني إلى المجموعة

الفرعية من الفقرات التي لها المنحنى المميز للفقرة نفسه، وتفسر احتمالية الإجابة الصحيحة عندئذ على أنَّها احتمالية استجابة مفحوص معين على فقرة اختيرت عشوائياً من مجموعة الفقرات الفرعية وأوصى لورد(1980) بالتحديد تجنب تفسير احتمالية الإجابة الصحيحة باحتمالية إجابة مفحوص معين على فقرة محددة إجابة صحيحة.

ومع ذلك فإنَّ الشكل(S) للمنحنى المميز للفقرة يستخدم بشكل واسع في بناء الاختبارات، لكنَّه ليس الشكل الوحيد المحتمل للمنحنى المميز للفقرة، بل هناك منحنيات اخرى يطلق عليها اسم دالة الخطوة(Step Function) حيث تتضمن وجود درجة سمة كامنة دنيا يرمز لها بالرمز $(\theta^2)$ وعند درجة أقل منها لا يستطيع المفحوص الإجابة بشكل صحيح على الفقرة، وتفيد دوال الخطوة في إنتاج مفاهيم مهمة عديدة لنظرية الاستجابة للمفردة، ومع ذلك تعد هذه الدوال أقل شيوعاً في بناء الاختبارات من منحنيات الشكل(S) كون بيانات الاختبارات الحقيقية تكون أكثر اتساقاً في هذا النوع من المنحنيات (كروكر والجينا، 2009، ص.452–453)




**شكل رقم (08) دالة الخطوة للمنحنى المميز للفقرة**

### 8.مميزات نماذج نظرية الاستجابة للمفردة

أشارت ميمي إسماعيل (2007) أنَّ (Henson,1999) ذكر ثلاث مميزات رئيسية لنماذج نظرية الاستجابة للمفردة وهي:

- إمكانية الحصول على عدد كبير من المفردات تقيس نفس السمة الكامنة ولكن بشـــرط أنَّ تقديـــر قدرة الممتحن يكون مستقلاً عن مجموعة مفردات الاختبار التي أجاب عنها.
- إمكانية وجود عدد كبير من الممتحنين ولكن بشرط أنَّ تقدير إحصائيات المفردة (صعوبة. تمييز)

يكون مستقلاً عن عينة الأفراد التي استخدمت في حساب هذه الإحصائيات.

- إمكانية تقدير معالم صعوبة المفردة، تمييز المفردة، التخمين حيث أنَّها تعطي معلومات مهمة مع الأخذ في الاعتبار صعوبة كل من الاختبار والمفردة على مدى قدرة معينة.

ويضيف هوسن وبوستل ثوايت (Postlethwaite & Husen ,1995 المشار إليه في إسماعيل، 2007) ميزة رابعة لنماذج نظرية الاستجابة للمفردة وهي:

- إمكانية التنبؤ بخصائص الاختبار مقدماً قبل تنفيذه.

وتضيف ميمي إسماعيل أنَّ أنستازي وأوربينا(1997) يرَيا أنَّ استخدام نظرية الاستجابة للمفردة يوفر العديد من الفوائد، والميزة المهمة لهذه الطريقة في القياس هي معالجتها لموضوع الثبات، والخطأ في القياس من خلال دوال معلومات المفردة حيث يمكن حساب تلك الدوال لكل مفردة، وتوفر قواعد دقيقة لاختبار المفردات عند بناء الاختبار، نظراً لأنَّ دالة معلومات المفردة تأخذ في الاعتبار جميع القدرة (إسماعيل، 2007، ص.32).

**9. إجراءات استعمال نظرية الاستجابة للمفردة**

**أ. اختيار النموذج المناسب**

إنَّ اختيار النموذج المناسب في نظرية الاستجابة للمفردة يعتبر دالة لكل من القيم المتوقعة لدرجة المفردة، وعدد السمات الكامنة المؤثرة في الاستجابة، وعدد المفردات، وحجم العينة ومجال المحتوى (Panter & Reeve,2002, p25 ، كما ورد في محمود،2012).

فنماذج النظرية الحديثة تنقسم تبعاً لمستوى الإجابة (ثنائية أو متعددة) أو تبعاً لعدد السمات الكامنة المؤثرة في الاستجابة (أحادية البعد أو متعددة الأبعاد) وتنقسم تبعاً لعدد بارامترات المفردة التي يتم تقديرها والافتراضات الإحصائية التي تحكم شكل المنحنى المميز للمفردة (أحادية البارامتر ثنائية البارامتر، أو ثلاثية البارامتر) (Bejar, 1983, pp. 31-32).

ويعد تسمية النماذج حسب عدد بارامترات المفردة الذي يأخذه النموذج في الاعتبار هو أفضل التصنيفات وأكثرها شيوعاً (Szabo, 2008, p. 47).

**ب. تقدير بارامترات النموذج**

تعتبر إجراءات تقدير البارامترات مكوناً أساسياً في عملية التحليل الإحصائي، ويختلف مدى

تعقدها تبعاً لعدد بارامترات المفردة التي يتضمنها النموذج، وفي نموذج راش على الرغم من وجود بارامترين فقط، هما: (قدرة الفرد، وصعوبة المفردة)، فإنَّه لا يكون أي منهما معلوماً في أية مجموعة مثالية من البيانات.

لذلك فقد ظهر وتطورت العديد من الإجراءات الرياضية لتقدير البارامترات في نظرية الاستجابة للمفردة، التي كانت مجالاً خصباً لعلماء الإحصاء، كما ظهرت الأبحاث التي تقارن بينها ومن أهم هذه الطرق: طريقة الإمكان الأعظم المرتبط (غير الشرطي)، وطريقة الإمكان الأعظم الهامشي وطريقة الإمكان الأعظم الشرطي، وطريقة بييز، وطريقة التحليل العاملي غير الخطي وطريقة التقدير (إسماعيل، 2004).

**ج. التحقق من ملاءمة البيانات للنموذج**

باعتبار نماذج نظرية الاستجابة للمفردة نماذج افتراضية، فإنَّ استعمال النماذج يقتضي التحقق من أنَّه سيشرح البيانات الفعلية (Wright & Stone, 1979, p. 60)، ومن ثمَّ يعد تحليل البيانات للنموذج جزءاً أساسياً عند استعمال نظرية الاستجابة للمفردة(Szabo,2008, p 44) وتكمن أهمية هذه الخطوة في ارتباطها بالصدق إذ يعتمد هذا الأخير في النظرية الحديثة على درجة تحقق افتراضات النموذج المستعمل ولذلك يعد الصدق هنا هو الاعتبار الأكثر أهمية في تقييم الاختبار.

ويقصد بملاءمة النموذج للبيانات مدى مناسبة المكون النظري للنموذج لتحليل البيانات ويشير ثورندايك إلى أنَّه بمجرد الحصول على تقديرات صعوبة المفردات وتقديرات القدرة المقابلة لكل درجة على الاختبار، يكون السؤال: إلى أي درجة يمكن أن نستنتج البيانات الأصلية من مقياس التقديرات الناتج؟ (ثورندايك وهيجن، 1986).

وتتضمن اختبارات الملاءمة التحقق من افتراض أحادية البعد، وعدم تغاير البارامترات، أي أنَّ تقديرات البارامترات داخل مجتمع ما ترتبط خطياً مع مثيلتها داخل مجتمع آخر (Bejar,1983, p.37) وهنا تقع المسؤولية الكبيرة على مطور الاختبار في التأكد من أنَّ افتراضات نظرية الاستجابة للمفردة محققة، وأنَّ النموذج المستعمل يقدم ملاءمة جيدة للبيانات.(Bachman, 1991, p. 209).

**د. الاعتبارات المتعلقة بالنموذج**

هناك عدة اعتبارات وعوامل تؤثر في تحديد واختيار نوع النموذج المناسب منها مدى ملاءمة

مفردات الاختبار للنموذج وتباين تمييز المفردات، وعدد البارامترات ...إذ أن استعمال النموذج في غير موضعه قد يؤدي الى نتائج أقل دقة.

ولكي يمكن تحديد أي من نماذج الاستجابة للمفردة هو الذي يجب استخدامه، هناك ثلاث محكات لخَّصها ويبرج كما يلي: (Wiberg, 2004, p.4).

- المحك الأول: تحديد افتراضات النموذج.
- المحك الثاني: خصائص النموذج المتوقعة.
- المحك الثالث: توقعات النموذج لنتائج الاختبار الفعلية.

ويمكن أن تنقسم تلك المحكات إلى محكات فرعية أخرى. ولقد اقترح همبلتون أنَّه يجب على الاختبار أن يلائم أكثر من نموذج واحد للبيانات، ومن ثمَّ يمكن مقارنة النماذج طبقا للمحكات الثلاثة (Hambleton, et al., 1991).

**المحك الأول: افتراضات النموذج *Assumptions of the Model***

**أ. أحادية البعد *Unidimensionality***

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

**ب. تساوى التمييز *Equal discrimination***

يمكن لتساوي التمييز أن يفحص العلاقة بين المفردة(i) والدرجة الكلية لدرجة الاختبار فإذا كانت المفردات ذات تمييز متساوٍ فمن الأفضل أن نستخدم النموذج اللوغاريتمي أحادي المعلم، أما إذا لم تكن المفردات ذات تمييز متساوٍ فمن الأفضل أن يستخدم النموذج اللوغاريتمي ثنائي المعلم والنموذج اللوغاريتمي ثلاثي المعلم أكثر من النموذج أحادي المعلم.

**ج. احتمالية تخمين الإجابة الصحيحة *Possibility of guessing the correct answer***

هناك طريقة لفحص ما إذا كان يحدث تخمين وذلك عن طريق ملاحظة إجابة الممتحنين ذوي القدرات المنخفضة على المفردات الأكثر صعوبة، أي أنَّ الممتحنين ذوي القدرات المنخفضة الذين يجيبون على المفردات الأكثر صعوبة بشكل صحيح، فإنَّ ذلك يسمح لمعلم التخمين أن يظهر، وعلى ذلك فإنَّ النموذج اللوغاريتمي ثلاثي المعلم أكثر ملاءمة من النموذج اللوغاريتمي ثنائي المعلم والنموذج

اللوغاريتمي أحادي المعلم في إظهار معلم التخمين(Wiberg, 2004, p.5) .

**المحك الثاني: خصائص النموذج المتوقعة *Expected Model Features***

- أنَّ تقدير مَعلَم قدرة الفرد لا يعتمد على تقدير مَعلَم صعوبة المفردة.
- أنَّ تقدير معلم المفردة لا يعتمد على استخدام عينات مختلفة من الأفراد سواء كانت عينات ذات قدرات مرتفعة أم منخفضة.

**المحك الثالث: توقعات النموذج لنتائج الاختبار الفعلية *Model predictions of actual test results***

يتم معرفة توقعات النموذج لنتائج الاختبار الفعلية عن طريق:

- إجراء مقارنة بين المنحنيات المميزة لكل مفردة مع المفردات الأخرى.
- استخدام رسوم بيانية لتقسيمات الدرجة المتوقعة والملاحظة أو استخدام اختبار كاف تربيع(كا$^2$).

**10.نماذج نظرية الاستجابة للمفردة**

لقد أدت الجهود الدؤوبة التي بذلها المهتمون بنماذج الاستجابة للفقرة للتوصل الى مجموعة من النماذج السيكومترية الجديدة المعروفة باسم نماذج الاستجابة للمفردة(IRM) وتتخلص الفكرة الأساسية لهذه النماذج في أنَّها محاولة اشتقاق القيم المقدرة للسمة أو السمات المضمَّنة في الاستجابة لمجموعة من فقرات الاختبار، بمعنى أنَّ هذه النماذج تقدم شرحاً لاستجابة الفرد لفقرة الاختبار التي تتضمن السمة المقاسة بهذه الفقرة، والسمة المقاسة قد تكون قدرة معينة أو سمة من سمات الفرد بحيث يوجد هناك علاقة منتظمة بين مستويات السمة المقاسة للأفراد المختلفين وبين احتمالية الاستجابة الصحيحة لفقرات الاختبار(علام، 2000، ص.686).

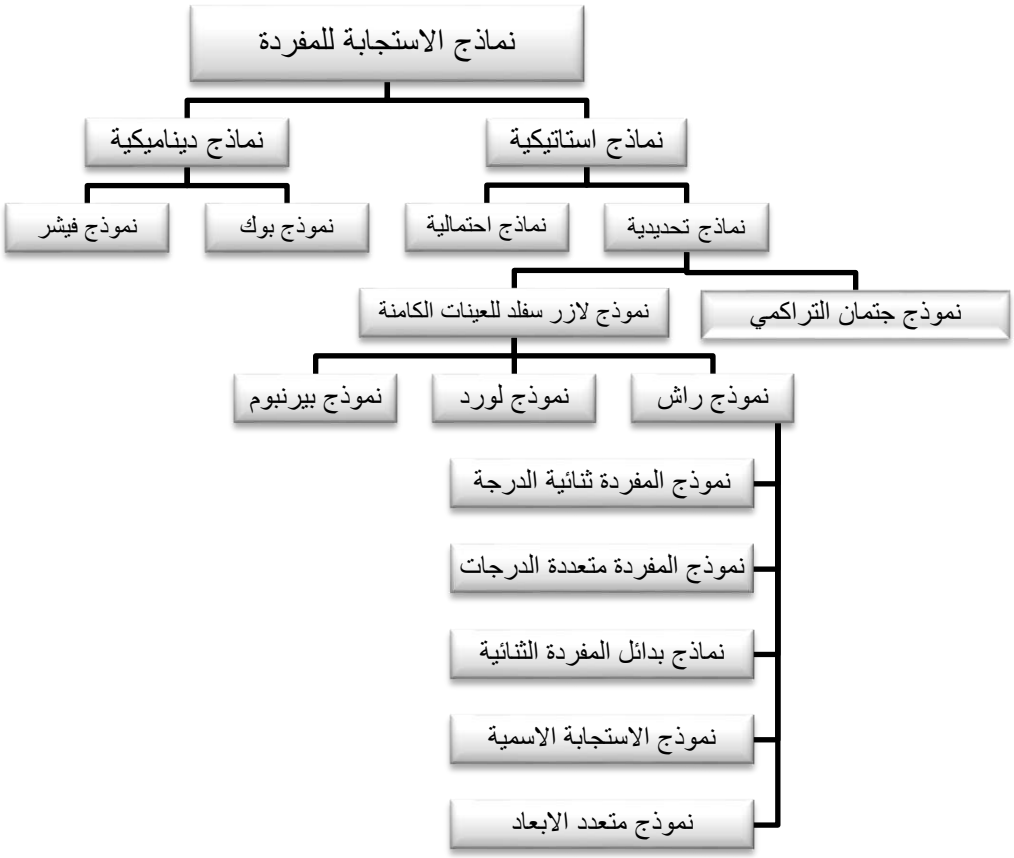
والهدف الأساسي لأي نموذج ناتج من أي نظرية تجريبية بشكل عام هو استخدامه في التدريج Scaling، وهذا يتطلب إعطاء القيم للأفراد على التدريج Scale values، سيحدث هذا فقط إذا قام النموذج بتحديد العلاقة بين الأداء الفردي، والسمة التي تقف وراء هذا الأداء، وبالتالي يحدد نموذج السمة الكامنة العلاقة بين البيانات (الأداء الفردي في الاختبار) تدريج وتقدير القيم أو التقديرات على تدريب السمة الكامنة.

وكلما كان هناك ملاءمة بين النموذج المستخدم ومجموعة البيانات، ينتج عن ذلك تقديرات دقيقة للسمة المقاسة، حيث يجب وضع الأفراد والفقرات على مقياس تدريج السمة، ويتم ذلك من خلال

التقديرات طالما هناك تقارب محتمل بين احتمالات الفاحصين (الاحتمالات الواقعية الحقيقية) لأدائهم في كل مستوى من مستويات السمة المقاسة، وتتميز هذه النماذج بأنَّها دوال احتمالية Probabilistic وليست دوال حتمية Deterministic، وهذا يعني أنَّ العلاقة بين أداء الفرد والسمة التي يجري قياسها تسير وفقاً لنظرية الاحتمالات(علام ،1982، ص.40).

ويرجع الاختلاف الرئيسي بين هذه النماذج إلى اختلاف صورة الدالة الرياضية الممثلة لها بسبب اختلاف عدد البارامترات المطلوبة لوصف هذه الدالة، مما يؤدي إلى اختلاف شكل المنحنى المميز للفقرة حيث تشمل النماذج كلاً من النماذج الاستاتيكية والنماذج الديناميكية، وقد قدَّم علام تصنيفاً شاملاً لنماذج الاستجابة للمفردة في الشكل رقم (09) (علام، 2000، ص.616) وعلى الباحث أن يختار إحدى هذه الدوال الرياضية، ولكن عليه أن يتحقق من أنَّ الدالة التي اختارها تطابق البيانات المستمدة من الاختبار بدرجة أفضل من غيرها.

توجد نماذج متعددة في إطار نظرية السمات الكامنة، وذلك لاختلاف الافتراضات المتعلقة بالبيانات الاختبارية، إذ ربما يفترض أنَّ الأداء في الاختبار ينطوي على سمة أحادية البعد، أو على سمات متعددة أو ربما يفترض أنَّ المفردات تتباين في قدرتها على التمييز بين المستويات المختلفة للقدرة، أو الإجابات الصحيحة عن فقرات الاختبار من متعدد ثنائية أم متعددة الأقسام (علام،2000، ص.67).




**الشكل رقم (09) تصنيف لنماذج لاستجابة للمفردة الاختبارية**

وسنعرض هذه النماذج مع التركيز على نموذج راش كونه المعتمد في إجراءات الدراسة الحالية أما بقية النماذج فسنعرض بعضها بشيء من الإيجاز.

### 1.10. نموذج راش (*Rasch*) (النموذج اللوجستي ذو المعلمة الواحدة).

يعد نموذج راش الأحادي المعلم One parameter أبسط نماذج الاستجابة للمفردة وأكثرها استخداماً في بناء الاختبارات وتحليل مفرداتها، ويرجع الفضل في بناء هذا النموذج إلى عالم الرياضيات الدنماركي جورج راش(Rasch) وطوره للتطبيق العملي العالم الأمريكي بنجامين رايت (Benjamin Wright) وكان هدف راش تحقيق الموضوعية، بمعنى أنَّ درجة المفحوص في الاختبار يجب ألا تكون دالة لعينة المفحوصين في التدريج الأصلي للفقرات المشمولة في الاختبار، كما أنَّه يجب أن يحصل المفحوص على الدرجة نفسها في أي اختبارين يقيسان القدرة أو السمة نفسها(علام، 1986، ص.118).

ويقــوم نموذج راش على التقديــر الثنائي للاستجابة، فإما أن يصيب الفرد(الممتحن) الهدف

ويجيب على الفقرة إجابة صحيحة، وعندئذ يحصل على الدرجة(واحد)، وإما أن يخطئ الفرد الهدف ويجيب على الفقرة إجابة خاطئة وعندئذ يحصل على الدرجة(صفر) (كاظم، 1988أ ;علام، 1986).

ويقوم نموذج راش في نظرية القياس على نتائج تفاعل قدرة الفرد مع صعوبة الفقرة، وتمثل نتائج هذا التفاعل بالاستجابات الملاحظة على فقرات الاختبار، التي يمكن من خلاله التوصل للتدريجات المتحررة للفقرات، وتقديرات الأفراد التي تتحقق بها مطالب الموضوعية في القياس (الشرقاوي، وآخرون، 1996، ص.311).

ويصف نموذج راش الفقرة بدلالة بارامتر واحد وهو الصعوبة، إذ أنَّه يفترض أنَّ بارامتر التخمين يساوي(صفر) وبارامتر التمييز يساوي(واحد)، لذا فعلى معد الاختبار العناية بكتابة فقرات الاختبار بحيث يراعي الافتراضات الأساسية التي يتطلبها النموذج مثل تقليل نسبة التخمين إلى أقصى حد.

ممكن من خلال العناية باختيار بدائل الإجابة عند استعمال فقرات اختيار من متعدد والاهتمام بأن تكون الفقرات ذات قوة تمييزية واحدة (Crislip & Chance, 2001, p. 2).

وأنَّ احتمالية الإجابة الصحيحة هي المتغير التابع في نموذج راش وغيره من نماذج السمات الكامنة، أما المتغيرات الكامنة المستقلة في هذا النموذج، فهي درجة قدرة الفرد وصعوبة الفقرة ويشترط في البيانات التي تعمد إلى تحليلها باستعمال نموذج راش أن تستوفي فروض النموذج. وتأخذ معادلة نموذج راش الشكل التالي:

$$p_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} \dots\dots\dots (07)$$

حيث:

$p_i(\theta)$: احتمال الإجابة الصحيحة لفرد قدرته $\theta$ عن الفقرة i

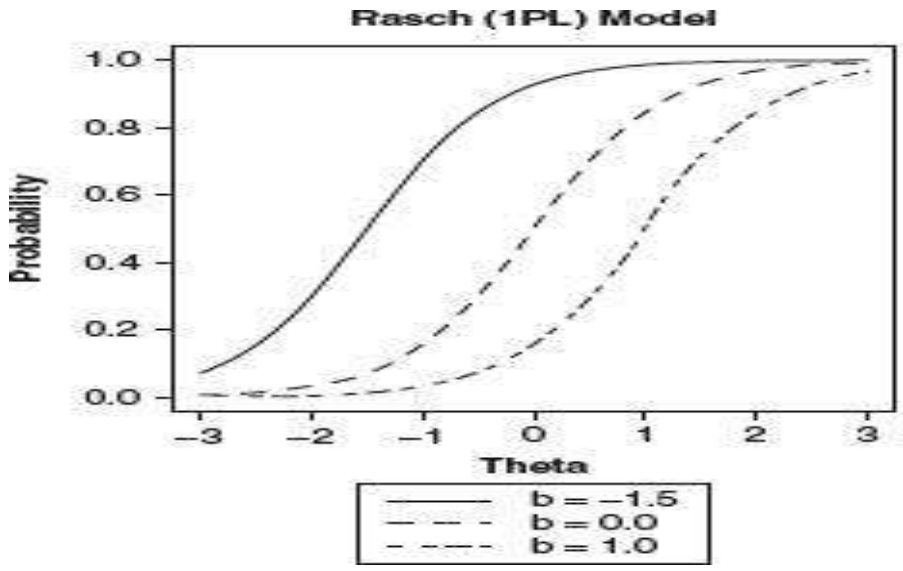
$D$: عامل التدريج ويساوي 1.7

$\theta$ : قدرة الفرد.

$b_i$ : معلم صعوبة الفقرة $i$

$e$ : تمثل الاساس الطبيعي اللوغاريتمي

والشكل التالي يبين المنحنيات المميزة للدالة لثلاثة فقرات في نموذج راش




**شكل رقم (10) المنحنيات المميزة للدالة لثلاثة فقرات في نموذج راش**

## 2.10. نموذج لورد (Lord) (النموذج اللوجستي ثنائي المعلمة)

يتم في هذا النموذج إضافة معلم جديد إلى نموذج راش وهو معلم التميز(ai) لكل فقرة والذي هو عبارة عن ميل منحنى خاصية المفردة عن نقطة انقلاب المنحنى، وهذا يؤدي إلى تقاطع المنحنيات المميزة لفقرات الاختبار الذي يصمم وفق هذا النموذج، أي أنَّ هذا النموذج يفترض انعدام التخمين(ci)، ويقوم بتقدير صعوبة الفقرات(bi) ومعاملات التميز(ai) (عبد الوهاب، 2010).

ويتم إضافة المعلم الجديد (معلم التمييز) لكل فقرة لأنَّه من الصعب إيجاد مجموعة من الفقرات تميز بدرجة واحدة بين مستويات السمة أو القدرة التي يقيسها اختبار معين (زكري، 2007، ص.51) مما يؤدي إلى تقاطع المنحنيات المميزة لفقرات الاختبار، إذ يفترض هذا النموذج انعدام التخمين ويعمل على تقدير صعوبة الفقرات ومعاملات التمييز وتتمثل الصيغة الرياضية لهذا النموذج في:

$$p_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \dots\dots\dots (08)$$

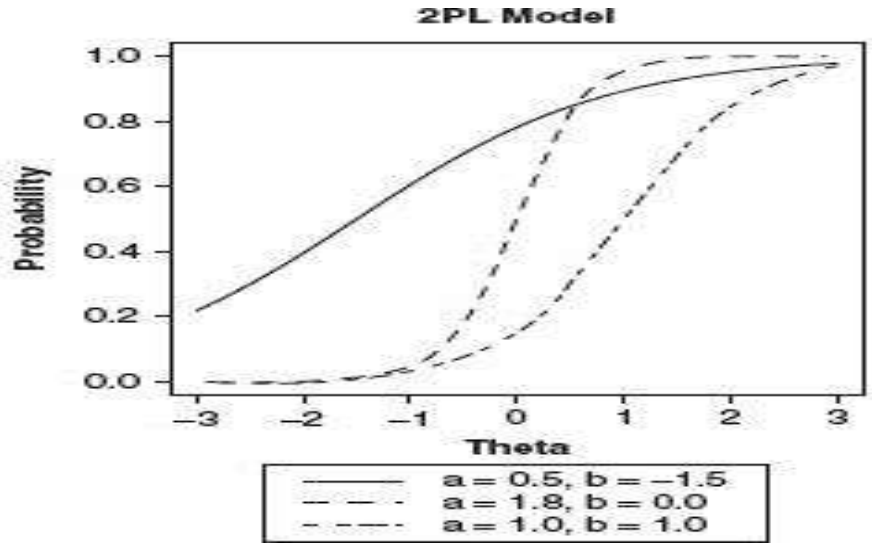
حيث:

$a_i$: معلمة تمييز الفقرة.

ويعتمد هذا النموذج على الفرق بين القدرة(θ) التي يمتلكها الطالب(S) في السمة التي يراد تقديرها (القدرة الكامنة وراء إجابات الطلبة)، ودرجة صعوبة الفقرة التي يرغب الطالب الإجابة عنها والتي يمثلها الرمز(b) مضروباً في درجة التميز التي تتصف فيها الفقرة(i) والتي يرمز لها بالرمز(ai)

وسيتم كذلك افتراض أنَّ هناك بعداً واحداً وراء الفروق الفردية في إجابات الطلبة (التقي، 2009، ص.22).

إنَّ إضافة معلم التمييز أتت بسبب صعوبة إيجاد مجموعة من الفقرات تميز بدرجة واحدة بين مستويات السمة أو القدرة التي يقيسها الاختبار وهو الافتراض الذي اعتمد عليه نموذج راش ويعرف بارامتر الصعوبة وفق هذا النموذج بأنَّه نقطة على ميزان القدرة عندما تكون احتمالية الاستجابة الصحيحة(0.5) ويتراوح مدى الصعوبة بين(-∞، +∞) وبارامتر التميز فانَّ فكرته تقوم بصورة عامة في نظرية السمات الكامنة على أنَّ هذا البارامتر يصف الطريقة التي تفرق فيها الفقرة بين المفحوصين الذين تكون قدراتهم أعلى من مستوى صعوبة الفقرة والذين تكون قدراتهم أدنى من صعوبة الفقرة وهذا ينعكس على شدة انحدار المنحنى المميز للفقرة في منتصفه، فإذا كان المنحنى شديد الانحدار فانَّ الفقرة تميز بشكل جيد بين المفحوصين، وإذا كان المنحنى منبسط كانت الفقرة أقل تميزاً ( Baker, 2001,p.22). ويشير ويبرج(Wiberg, 2004) أنَّه في أغلب الأحيان يكون بارامتر التميز ذو قيمة موجبة أي أنَّ احتمالية الاستجابة الصحيحة تتزايد كلما تزايدت القدرة، ولكن قد نحصل على تمييز سالب فيها واحتمالية الاستجابة الصحيحة تتناقص كلما تزايدت القدرة (عصفور، 2007، ص.56). والشكل التالي بين المنحنيات المميزة للدالة لثلاثة فقرات في نموذج لورد.




**شكل رقم (11) المنحنيات المميزة للدالة لثلاثة فقرات في نموذج لورد**

### 3.10. نموذج بيرنبوم *(Birnbaum)*

ويعرف النموذج اللوجستي ذو المعلمات الثلاثة (Three-parameter logistic model) بنموذج بيرنبوم الذي أضاف له معلما ثالثاً أطلق عليه بارامتر الخط ألتقاربي الأدنى أو بارامتر التخمين الذي يتعلق بالفقرات الاختبارية التي تتطلب الاختيار من بدائل متعددة، وهذا البارامتر يحدد احتمال أن يجيب فرد يفترض أنَّ مستوى قدرته منخفض انخفاضاً لانهائياً، ومع هذا يجيب إجابة صحيحة عن غالبية مفردات الاختبار عن طريق التخمين (علام، 1986، ص.695).

يصف هذا النموذج العلاقة بين احتمالية الاستجابة الصحيحة عن الفقرة والقدرة في ضوء ثلاث بارامترات وهي: التخمين، والصعوبة، والتميز، حيث إنَّ بعض المفحوصين قد يعمدون التخمين في التوصل إلى الإجابة الصحيحة وينجحون بذلك (Hullin et al., 1983,p.29).

إنَّ بارامتر التخمين هو احتمالية الاجابة الصحيحة عن الفقرة بالتخمين لوحده، وقيمة هذا البارامتر لا تكون متغيرة وكأنَّها دالة لمستوى القدرة، لذلك فأنَّ المفحوصين من ذوي القدرة العالية أو الدنيا، لهم نفس الاحتمالية بالتوصل إلى الإجابة الصحيحة عن طريق بارامتر التخمين، ولهذا فأنَّ القيمة النظرية لهذا البارامتر تتراوح بين (1،0)، وبارامتر الصعوبة في ضوء هذا النموذج هو نقطة على ميزان القدرة عندما تكون احتمالية الاستجابة الصحيحة في منتصف المسافة بين بارامتر التخمين و(1) (Baker, 2001,pp. 28-29).

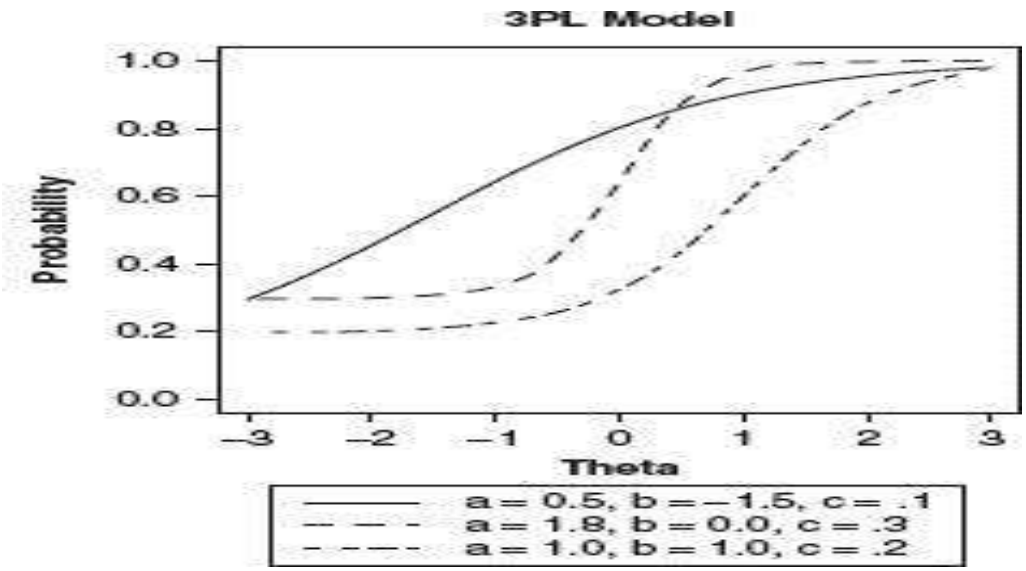
أي أنَّ هذا المعلم يحدد أثر التخمين في احتمال الإجابة الصحيحة على فقرات الاختبار وتأخذ الصيغة الرياضية في هذا النموذج الشكل التالي:(Hambleton & Swaminthan,1985).

$$p_i(\theta) = c_i + (1 + c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \dots\dots\dots (09)$$

حيث:

$c_i$ : معلمة تخمين الفقرة والذي يمثل الخط التقاربي الأدنى لمنحنى خصائص الفقرة.

والشكل التالي يبين المنحنيات المميزة للدالة لثلاثة فقرات في نموذج بيرنبوم




**شكل رقم (12) المنحنيات المميزة للدالة لثلاثة فقرات في نموذج بيرنبوم**

**11.سبب عدم انتشار نظرية الاستجابة للمفردة**

يرجع سبب عدم انتشار مبادئ وأسس هذه النظرية المعاصرة انتشاراً واسعاً إلى عدة أسباب منها:

**1.** تمثل هذه النظرية فرعاً معقداً من فروع نظرية الاختبارات، واستيعاب البحوث السيكومترية التي تجري في هذا المجال وتنشر في الدوريات العلمية المتخصصة يحتاج إلى قدر كبير من الفهم الرياضي العالي، والإحصاء متعدد المتغيرات.

**2.** معظم خبراء القياس الذين تناولوا هذا الموضوع وجهوا دراساتهم وبحوثهم لمنظري القياس وليس للذين سيقومون بتطبيقه في بناء الاختبارات وتحليل وتفسير نتائجها، وفى غير ذلك من التطبيقات السيكولوجية والتربوية (زكري، 2008، ص.60).

**3.** صعوبة تحقق بعض فروض النماذج المتعلقة بهذه النظرية في البيانات الفعلية المستمدة من الاختبارات العقلية والتحصيلية (علام، 2000).

**4.** لم يكن هناك برامج حاسوبية فعالة ومناسبة لتقدير بارامترات نماذج السمات الكامنة التي تتضمنها هذه النظرية (علام، 2000، ص.83-84).

**المحـــور الرابـــع: نمـــوذج راش *Rasch Model***

يعد نموذج راش أحد نماذج نظرية الاستجابة للمفردة ثنائية التدرج، ويطلق عليه اسم نموذج البارامتر الواحد، ويعد هذا النموذج من أكثر النماذج شيوعاً في تصميم وبناء الاختبارات والمقاييس النفسية والتربوية، وسنتناوله بشيء من التفصيل في هذا المحور.

**1. الخصائص المميزة لنموذج راش**

حاز نموذج راش على اهتمام معظم علماء القياس النفسي والتربوي المعاصر، وذلك للمزايا التي يمتاز بها عن غيره من النماذج ومن هذه المزايا:

- يعد نموذج راش أقل النماذج في عدد الفروض اللازم توفرها أو تحقيقها في البيانات المستمدة من الاختبارات، والذي يستخدم للحصول على تقديرات دقيقة لكل من صعوبة الفقرة وقدرة الفرد وأبسطها كونه يشتمل على بارامترٍ واحدٍ يتعلق بصعوبة الفقرة (علام، 2001، ص.696).
- توازي المنحنيات المميزة للفقرات، أي أنَّ منحنياتها اللوغاريتمية متساوية الانحدار، إذ أنَّ النموذج يفترض تساوي القوة التمييزية للفقرات جميعها، ونتيجة لتساوي انحدار الفقرات فأنَّه يمكن التعبير عن معلم الفرد ومعلم الفقرة على متصل واحد وبوحدة قياس واحدة (يعقوب، 1990، ص.44).
- إنَّ استخدام نموذج راش يتيح الفرصة لحذف كل من الأفراد غير الملائمين وكذلك المفردات غير الملائمة، ويستبقي عند تدريج المقاييس استجابات الأفراد الصادقة، وكذا المفردات الصادقة في تدرجها وفي قياس القدرة موضوع القياس، وهذا يوفر صدق القياس لهذه المقاييس.
- إنَّ استقلالية القياس التي يوفرها نموذج راش يتيح الفرصة لثبات المقياس، حيث لا يختلف القياس (سواء لقدرة الفرد، أو صعوبة المفردة) باختلاف عينة التدريج أو باختلاف الاختبار الفرعي المستخدم لقياس القدرة (كاظم وأخرون، 1996، ص.123).
- إيجاد الحلول الإحصائية المناسبة لمشكلة تقدير بارامترات النموذج، إذ تعطي القيم (0) و(1) في حالة الإجابات الثنائية، على عكس النماذج الأخرى حيث تواجه فيها مشكلات إحصائية ويتطلب استخدامها درجات موزونة (Wright & Masters, 1982, p. 2).

**2. الأساس النظري لنموذج راش**

ظل نموذج راش فكراً رياضياً نظرياً بحتاً حتى عام 1965، حين بدأ العالم الأمريكي بنجامين

رايت (Benjamin Wright) وزملائه في تبني الفكرة الرئيسية لهذا النموذج، المتمثلة في موضوعية القياس فقد اهتم(رايت) بالجوانب التطبيقية لهذا النموذج بهدف تسهيل استخدامه في مجال القياس والتقويم التربوي (الوليلي، 2001، ص.96).

وقد أشار هامبلتون وكوك (1977) إلى أنَّ الأنشطة البحثية الخاصة بنموذج راش قد ازدادت وشملت عدة أمور أهما:(Hambleton, & cook, 1977, p.90 ).

- استخدام النموذج في تحليل وبناء مفردات الاختبار وبنوك الأسئلة خاصة في الاختبارات التحصيلية.
- استخدام النموذج في تقدير المتغيرات واختبارات حسن المطابقة.
- تصميم البرامج الحاسوبية.

ويؤكد كل من (علام، 2001; Boeck & Wilson, 2006, p. 26 ; Fischer, 1997 ) على أنَّ نموذج راش يُعد من أكثر نماذج النظرية الاستجابة للمفردة استخداماً في بناء الاختبارات وانشاء بنوك الأسئلة لما تتضمنه من مزايا منها:

- يعد نموذج راش من أبسط النماذج الاحتمالية، لاحتوائه على بارامتر واحد (صعوبة المفردة) مع افتراضه تساوي المفردات في قدرتها التمييزية

- يتوفر نموذج راش على خاصية استقلال المعلمة، أي عدم تباين تقديرات قدرة الفرد بتباين عينة المفردات التي يختبرون بها، وعدم تباين الخصائص السيكومترية للمفردات بتباين عينة الأفراد التي أجري عليها الاختبار.

- يمثَّل نموذج راش الاستجابة على الاختبار بمنحنيات متوازية دون تقاطع من خلال منحنيات مميزة للمفردة، ومنحنيات مميزة للأفراد، مما يسهل عملية التقدير الاحصائي للنموذج بعكس النماذج الأخرى التي أضافت بارامترات أخرى على نموذج راش.

- يقوم نموذج راش على تحليل البيانات التي تمَّ استخراجها من فقرات الاختبار التي تعتمد إجابتها على طريقة صح بدرجة (1) وخطأ بدرجة (0) وعندما يستجيب الفرد لفقرة اختبارية ما فإنَّه يحدث تفاعل بين قدرة الفرد وصعوبة هذه الفقرة، ووفق هذا التفاعل يتضح الأساس الذي يعتمده نموذج راش.

ويشير علام إلى أنَّ نموذج راش يعتمد على قضيتين أساسيتين هما:(علام، 2001، ص.227)

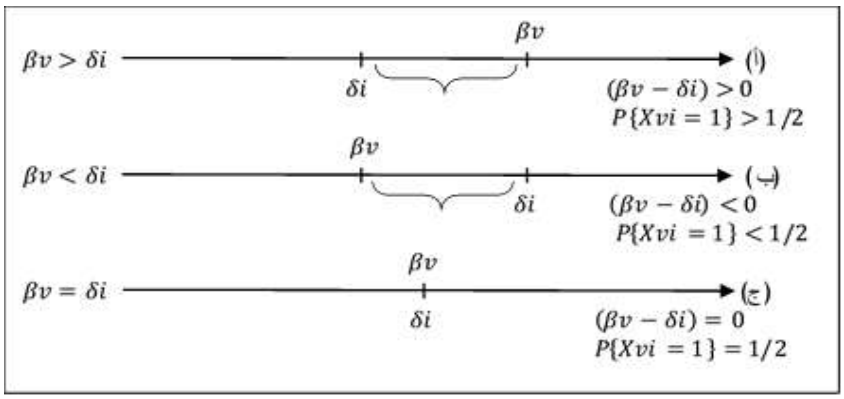
أ. أنَّ احتمال الإجابة الصحيحة عن مفردة اختبارية يزيد بزيادة قدرة الفرد.

ب. أنَّ احتمال أن يجيب فرد إجابة صحيحة عن مفردة سهلة أكبر من احتمال أن يجيب إجابة صحيحة عن مفردة صعبة.

فعندما يشرع الفرد($V$) في الاستجابة لفقرة معينة($i$) فإنَّ كلاً من مستوى القدرة لدى هذا الفرد($\beta v$) وصعوبة المفردة(δi) تعبران عن وضع هذا الفرد على متصل المتغير المقاس وتتحكمان في احتمال حدوث الاستجابة المناسبة على الفقرات المتدرجة على متصل هذا المتغير (كاظم، 1988أ، ص.43) وتكون هذه الاحتمالية محكومة بالفرق بين معلم قدرة الفرد وصعوبة الفقرة.

لذلك فإنَّه من المنطقي أنَّه كلما زاد مستوى قدرة الفرد على مستوى صعوبة المفردة ($\delta i$) > ($\beta v$) فإنَّ الفرق يكون أكبر من الصفر، وعندئذ يكون احتمال حدوث الاستجابة الصحيحة أكبر من(0.5)، أمَّا إذا كان مستوى قدرة الفرد($\beta v$) أقل من مستوى صعوبة المفردة($\delta i$)، فإنَّ الفرق ($\beta v - \delta i$) يكون أقل من الصفر وعندئذ يكون احتمال حدوث الاستجابة الصحيحة أقل من(0.5) أمَّا في حالة تساوي قدرة الفرد($\beta v$) مع صعوبة المفردة($\delta i$) فإنَّ الفرق($\beta v - \delta i$) يساوي الصفر وعندئذ يكون احتمال حدوث الاستجابة الصحيحة يساوي(0.5) (كاظم، ص.44).

وتضع كاظم تصور لثلاثة حالات (أ، ب، ج) توضح منطقية تأثير الفرق($\beta v - \delta i$) على احتمالية الاستجابة الصحيحة.



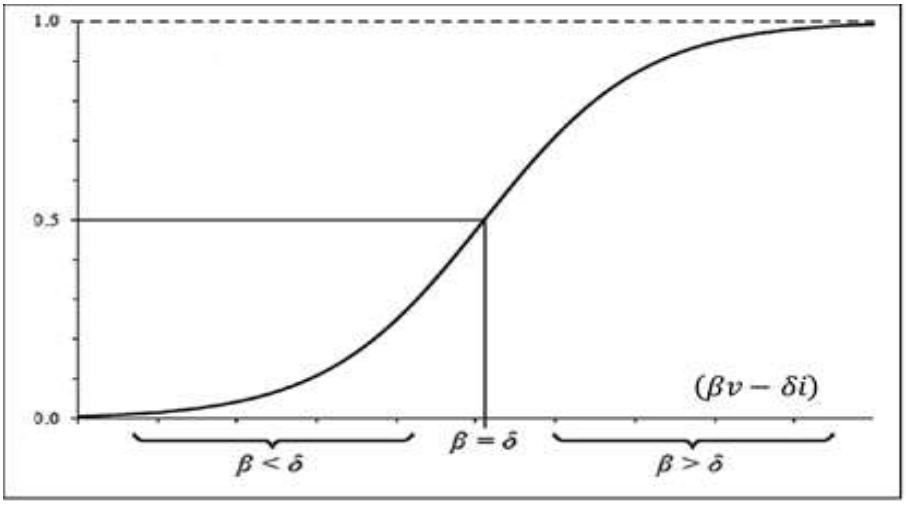

**الشكل رقم (13) تأثير الفرق بين مستوى قدرة الفرد ومستوى صعوبة البند في احتمال حدوث الاستجابة الصحيحة**

أي أنَّه إذا كانت:

$P[Xvi = 1] > 0.5 \longrightarrow (\beta v - \delta i) > 0$
$P[Xvi = 1] < 0.5 \longrightarrow (\beta v - \delta i) < 0$
$P[Xvi = 1] = 0.5 \longrightarrow (\beta v = \delta i)$

$P$: ترمز إلى الاحتمال

والمنحنى الموضح في الشكل رقم (14) يتضمن تلك العلاقات المنطقية بين الفرق$(\beta v - \delta i)$ واحتمالات الإجابة الصحيحة.




**شكل رقم (14) منحنى الاستجابة**

### 3. الأساس الرياضي لنموذج راش

تعتمد الصيغة الرياضية لاستجابة الفرد$(v)$ على المفردة(i) على تحديد متغيرات أساسية مؤثرة في الاستجابة، وهما قدرة الفرد$(\beta v)$ لحل جميع البنود، وصعوبة البند$(\delta i)$، ويظهر تأثير هذين المتغيرين في الاستجابة الملاحظة$(Xvi)$ من خلال النموذج الوسيط المتمثل في الفرق بين المتغيرين(Bv – δi) ويعد احتمال حدوث الاستجابة الصواب$(Xvi = 1)$ دالة لهذا الفرق، وهذا يتمثل في الدالة التالية: (كاظم، 1988ب، ص.49).

$$Pvi = f(\beta v - \delta i) * \dots\dots\dots (01)$$

حيث $Pvi$ احتمال نجاح الفرد$(v)$ على المفردة$(i)$، أي احتمال حدوث الاستجابة الصواب وقد تمكنت الباحثة أمينة الكاظم تصور كيفية التوصل إلى النموذج في صورته النهائية، إذ أنَّ احتمال الإجابة الصواب $(Pvi)$ تنحصر بين القيمتين(صفر) و(واحد) في حين أنَّ الفرق$(\beta v - \delta i)$ يمكن أنَّ يكون أي عدد حقيقي، وقد يصل إلى $-\infty$ وحتى $+\infty$ أي$(-\infty < (\boldsymbol{\beta v - \delta i}) < +\infty)$ لذا يتم تحويل $(\beta v - \delta i)$ إلى الصيغة الأسية للأساس الطبيعي $(e)$ فتكون الصيغة:

---

*$(\beta v)$: بارامتر قدرة الفرد، $(\delta i)$: بارامتر صعوبة المفردة أو البند

$$e^{(\beta v-\delta i)} = exp(\beta v - \delta i) ** \dots\dots\dots (02)$$

وتتراوح هذه الصيغة بين صفر ومالا نهاية+ $0 \leq e^{(\beta v-\delta i)} \leq \infty$ ولتحويلها إلى المدى من (صفر) إلى (واحد) نصل إلى النسبة:

$$0 \leq \frac{e^{(\beta v-\delta i)}}{1 + exp\ (vp - \delta i)} \leq 1$$

ويمكن مساواتها بالطرف الأيسر من الدالة (01) فتصبح المعادلة

$$Pvi - \frac{\exp(\beta v - \delta i)}{1 + \exp(\beta v - \delta i)} \dots\dots\dots (03)$$

عندما تكون الاستجابة $Xvi = 1$ تعبر عن احتمال النجاح يكون:

$$P\left(Xvi = \frac{1}{\beta v}, \delta i\right) = \frac{exp\ (\beta v - \delta i)}{1 + \exp(\beta v - \delta i)} \dots\dots\dots (04)$$

الاستجابة $Xvi = 0$ تعبر عن احتمال الخطأ***

$$P(Xvi = 0/\beta v, \delta i) = 1 - \frac{exp\ (\beta v - \delta i)}{1 + exp\ (\beta v - \delta i)}$$

وبالتبسيط تصبح المعادلة

$$P(Xvi = 0/\beta v, \delta i) = \frac{1}{1 + exp\ (\beta v - \delta i)} \dots\dots\dots (5)$$

ومن المعادلتين (4) و(5) تصبح المعادلة العامة للنموذج:

$$P(Xvi = X/(\beta v, \delta i) = 1 - \frac{exp[X(\beta v - \delta i)]}{1 + exp(\beta v - \delta i)} \quad ,X = 0,1 \dots\dots\dots (06)$$

(Willmot, S. & Fowles, D.,1974, Murray, 87;.p D.,1974, Wright, Mead & 424;.p 1988أ). ، كما ورد في كاظم Bell,1980; Wright & Stone, 1979

وتعد هذه الصورة العامة لنموذج راش من بين مجموعة من نماذج القياس التي ترجع لجورج راش، والتي لها خواص فريدة هي الأساس للموضوعية في القياس.

### 4. الموضوعية في نموذج راش

يقصد بالموضوعية في نموذج راش "موضوعية المقارنة بين نتيجة تفاعل قدرتي فردين مع صعوبة مفردة مناسبة أي موضوعية المقارنة بين استجابتي فردين لمفـــردة مناسبة، كما تعني أيضاً

** $exp$ : تعني المقابل اللوغاريتمي
*** احتمال الخطأ = (1- احتمال النجاح)

موضوعية المقارنة بين صعوبة فقرتين أو مفردتين استجابة لهما فرد مناسب (كاظم، 1988أ، ص.53).

وتمثل الموضوعية حجر الأساس في القياسات جميعها فهي تتطلب استقلالية العامل المقاس مثل (صعوبة الفقرة) عن العينات المقاسة (مثل المفحوصين) المستخدمة في عملية التدريج، كذلك يجب أن يستقل قياس الأشياء عن أدوات قياسها، ويتحقق هذا في القياس النفسي بدرجة تقريبية لذلك يرى راش بأنَّ هذا التحقق التقريبي يمكِّننا من القياس الموضوعي الذي يرشد عملية اختيار النموذج المناسب في القياس وصولاً إلى قياس كمي للخصائص النفسية (علام، 1986).

وترى أمينة كاظم (1996) أنَّ الموضوعية تعني المقارنة بين استجابتي فردين لفقرة مناسبة فضلاً عن ذلك تعني موضوعية المقارنة بين صعوبة فقرتين استجاب لهما فرد مناسب (الشرقاوي، وآخرون، 1996، ص.221).

وعلى هذا الأساس اعتمدت النظرية الحديثة في القياس على النماذج الرياضية الاحتمالية التي تقوم على نظرية الاحتمالات للوصول إلى الموضوعية في القياس السلوكي، وبذلك فأنَّ هناك إجراءات مهمة ينبغي أن يقوم بها الباحث لتحقيق متطلبات الموضوعية في القياس، وهذه الإجراءات هي تحرير صعوبة الفقرة من عينة المفحوصين الذين طبق عليهم الاختبار وتحرير تقديرات القدرة من عينة الفقرات التي يشتمل عليها الاختبار (Baker, 2001, p. 55).

بمعنى مقارنة بين استجابة الفرد $(V)$ واستجابة الفرد$(U)$ للفقرة$(i)$ والمقارنة بين الفقرة$(i)$ والفقرة $(c)$ استجابة لهما الفرد(v). ويتحقق ذلك من خلال:(كاظم، 1988أ، ص.43– 49).

**أ. استقلال معلم قدرة الفرد عن الفقرة المستخدمة *Item-Free***

إنَّ معامل صعوبة الفقرة المتعارف عليه في القياس الجماعي المرجع هو نسبة الإجابات الصحيحة للعينة المستعملة إلا أنَّ هذا المعامل تتغير قيمته تبعاً لمستوى أو توزيع قدرة تلك العينة أي يكون عالياً للفقرة نفسها إذا كانت العينة المستعملة ذات قدرة عالية في تلك السمة، وينخفض في الحالة المعاكسة لذلك فهدف نموذج راش إيجاد تدريج للفقرات متحررة من هذا التأثير (قدرة الأفراد) (Wright & Panchapakesan, 1969, p.23).

وفي هذا النموذج يتم تحديد احتمالية الاستجابة الصحيحة عن أية فقرة في ضوء الفرق بين قدرة الفرد وصعوبة الفقرة، فإذا ما تمت المقارنة بين أي فقرتين باستعمال فقرة واحدة، فإنَّ قدرة الفرد

ستكون المتغير الوحيد المؤثر في هذه المقارنة، وبذلك يمكن استعمال أية فقرة للمقارنة بين قدرة هذين الفردين ولن تتأثر هذه المقارنة باستعمال أي فقرة من هذه الفقرات، وهذا يدل على أنَّ المقارنة بين قدرات الأفراد متحررة من تأثيرها بخصائص الفقرات المستعملة في المقارنة(الشرقاوي وآخرون، 1996، ص.321)، ويتم ذلك من خلال تحديد احتمال نجاح الفرد$(v)$ على الفقرة$(i)$ وفق المعادلة الآتية:

$$Pvi = \frac{exp\ (\beta v - \delta i)}{1 + exp\ (\beta v - \delta i)} \dots\dots\dots (07)$$

$Pvi$ تمثل احتمال نجاح فرد قدرته $(\beta i)$على فقرة صعوبتها $(\delta i)$.

وحيث $exp\ (\beta v - \delta i)$ مميز النجاح، أو مرجع النجاح:

$$exp(\beta v - \delta i) = \frac{Pvi*}{1 - Pvi}$$

وبأخذ لوغاريتم الطرفين تصبح:

$$(\beta v - \delta i) = ln\frac{Pvi}{1 - Pvi} \dots\dots\dots (08)$$

$(Pvi)$ تمثل احتمال النجاح.

أما $(1 - Pvi)$ فهو احتمال الخطأ.

أما $(l)$ ترمز للوغاريتم الطبيعي الذي أساسه $(e)$.

وبالمثل في حالة استجابة فرد آخر $(u)$ على الفقرة نفسها $(i)$ فإن:

$$(\beta u - \delta i) = ln\frac{Pui}{1 - Pui} \dots\dots\dots (09)$$

وبطرح المعادلة (8) من المعادلة (9) يحذف معلم صعوبة الفقرة$(\delta i)$ وتبدو معادلة المقارنة بين معلمي قدرة كل من الفرد $(v)$ والفرد $(u)$ من المعادلة الآتية:

$$(\beta v - \beta u) = ln\frac{Pvi}{1 - Pvi} - ln\frac{Pui}{1 - Pui} \dots\dots\dots (10)$$

* إذا كان $(Pvi)$ احتمال النجاح فإن $(1 - Pvi)$ احتمال الخطأ.
ويكون مرجح النجاح = احتمال النجاح/احتمال الخطأ

وعليه فإنَّ ما نصل إليه ليست القدرة المطلقة للفرد$(v)$ وإنَّما بعده عن قدرة فرد آخر هو$(u)$ وهذا الفرق يجعل الفرد$(u)$ نقطة أصل تقاس منها قدرة الفرد$(v)$.

**ب. استقلال معلم صعوبة الفقرة عن الفرد الذي يجيب عليها *Person-free***

يشير هذا المبدأ إلى أنَّ هدف نظرية السمات الكامنة المتمثل في إيجاد تدريج للفقرات يكون متحرراً من تأثير قدرات الأفراد، فمثلاً إذا اختبرنا فرداً واحداً بمجموعتين من الفقرات تقيسان متغيراً واحداً تمَّ تدريجهما، واستخراج خصائصهما وفق نموذج راش، ولكن إحدى هاتين المجموعتين متوسط صعوبتها سالب أي أنَّها سهلة، والأخر متوسط صعوبتها موجب أي أنها صعبة وأعدنا تدريج فقرات كل من المجموعتين، ويكون هذان التدريجان متكافئين إحصائياً، مع مراعاة الخطأ المعياري للتقدير بالنسبة للمفحوص (Baker, 2001, p. 91).

وهذا يعني إمكانية مقارنة قدرة فرد أجاب عن أية فقرة اختبارية بشرط تحقيقها لافتراضات نموذج راش، إلا أنَّ القياس المتحرر لا يكون مطلقاً في قياسه لقدرة المفحوصين، ولكن ما يقاس هو بعد قدرة أحد المفحوصين عن قدرة مفحوص أخر، أي أنَّ قدرة أحد المفحوصين يكون مرجعاً أو نقطة الأصل لقياس قدرة الآخر (الشرقاوي وآخرون، 1996، ص.321).

ويعني ذلك أنَّه إذا أجاب الفرد$(v)$ على فقرتين من الفقرات المناسبة $(c)$ و$(i)$ فإننا نصل إلى المعادلتين الآتيتين:

$$(\beta v - \delta i) = ln\frac{(P_{vi})}{1 - P_{vi}} \dots\dots\dots (11)$$

$$(\beta v - \delta i) = ln\frac{(P_{vc})}{1 - P_{vc}} \dots\dots\dots (12)$$

وبطرح المعادلة (11) من المعادلة (12) يحذف معلم قدرة الفرد$(\beta v)$ وتبدو المقارنة بين الفقرتين$(c)$ و$(i)$ من المعادلة:

$$(\delta_c - \delta_i) = ln\frac{(P_{vi})}{1 - P_{vi}} - \frac{(P_{vc})}{1 - P_{vc}} \dots\dots\dots (13)$$

يمكن التوصل إلى نفس المقارنة بين معلمي صعوبة الفقرتين، عن طريق أي فرد آخر يكون مناسب. وعليه فإنَّ ما يتم التوصل إليه لا يمثل الصعوبة المطلقة للفقرة$(i)$ وإنَّما بُعدُها عن صعوبة فقرة أخرى هي$(c)$ وهذا الفرق يجعل من صعوبة الفقرة$(c)$ نقطة أصل تقاس منها صعوبة الفقرة$(i)$.

**5. افتراضات نموذج راش**

يوفر نموذج راش متطلبات القياس الموضوعي عند تحقق الافتراضات الخاصة بالنموذج وهي:

**أ). أحادية البعد *Unidimensionality***

يفترض نموذج راش أنَّ أداء الفرد في الاختبار يمكن تفسيره بواسطة مفردة واحدة تشكل الأساس لجميع المفردات الأخرى، وهذا الافتراض هو أحادية البعد، حيث يجب أن تقيس المفردات بعداً واحداً فقط، ويرى هانسون (1999) أنَّ هناك صعوبات في تحقيق هذا الافتراض على الوجه الأكمل لكون هناك العديد من الصعوبات المعرفية وغيرها، التي قد تؤثر في استجابة الفرد على المفردة الاختبارية مما يجعل هذا الافتراض صعب التحقيق بالشكل المطلق.

وتشير أمينة كاظم (1988) إلى أنَّ أحادية البعد في نموذج راش تتحقق عندما:

- تتدرج المفردات بحيث تعرف متغيراً واحداً فقط، بحيث تختلف فيما بينها في صعوبتها فقط.
- يتدرج الأفراد على متصل هذا المتغير، تبعاً لمستوى أدائهم على المفردات، حيث يكون هؤلاء الأفراد ذوي قدرة أحادية فتحدد وحدها هذا المستوى للأداء على المفردات.

**ب). استقلالية القياس:*Local independence***

وهو يعني أن تكون استجابات الفرد عن المفردة الاختبارية مستقلة إحصائياً، بمعنى ألا تؤثر استجابة الفرد لإحدى مفردات الاختبار على استجاباته للمفردات الأخرى، وتبدو استقلالية القياس في

- استقلال القياس عن قدرة العينة التي تؤدي الاختبار بمعنى:
- تحرر قدرة الفرد المقدرة لهذا الاختبار عن قدرة باقي الأفراد الذين يجيبون عن الاختبار.
- تحرر تقدير صعوبة المفردة عن قدرة الأفراد الذين يجيبون عليها.
- استقلال القياس عن المفردات التي يجيب عليها الأفراد بمعنى:

▪ تحرر تقدير المفردة عن باقي مفردات الاختبار.

▪ تحرر تقدير قدرة الفرد عن صعوبة المفردات المعينة التي يجيب عليها.(Hong,2004, Kim & p,18; Avery et al., 2003, p,698) ، كما ورد في إسماعيل، 2007، ص.39)

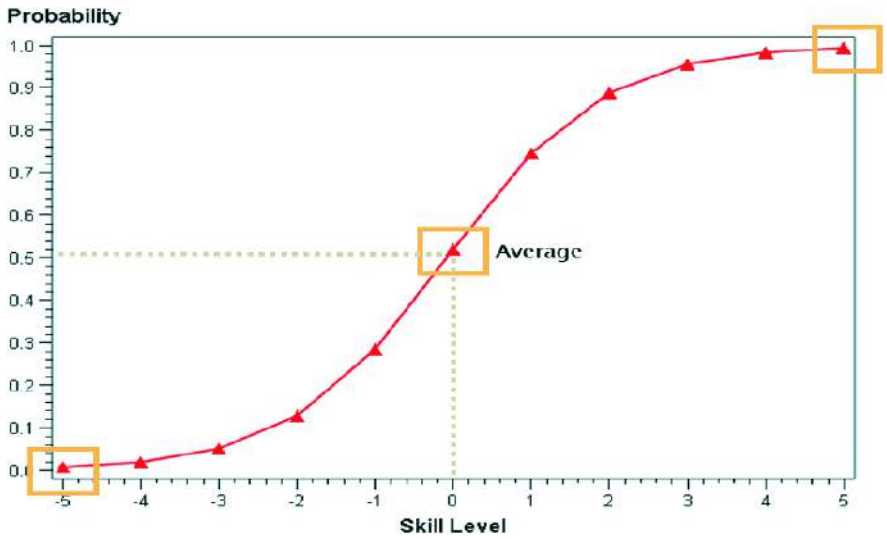
**ج). خطية القياس**

تعني أنَّ هناك معدلاً ثابتاً لتدرج القياس، وذلك على المدى الواسع من متصل المتغير موضع

القياس، والذي يتمثل بواسطة وحدة قياس واحدة، عندئذٍ يكون تقدير الفرق بين أي قياسين متتاليين على هذا التدريج ثابتاً عند أي مستوى من مستويات المتغير، ولا يختلف المعنى الكمي لأي فرق بين أي قياسين على هذا التدريج بتغير أداة القياس طالما أنَّها أداة مناسبة، وعندما تتوافر الخطية في القياس يمكننا تقدير التغير الحادث في المتغير موضع القياس (إسماعيل، 2007، ص.40).

**د). المنحنيات المميزة للفقرة *Item characteristic Curves***

يعرف شكل المنحى المميز للمفردة بأنَّه تعبير رياضي بين احتمال الاجابة الصحيحة على المفردة بالنسبة لمستوى القدرة وخصائص المفردة نفسها (بركات، ص.64) ولذلك فإنَّ احتمال إجابة الفرد الصحيحة عن مفردة ما تعتمد على شكل المنحنى.




**شكل رقم(15) المنحنى المميز للمفردة في نموذج راش**

**هـ).انعدام أثر التخمين**

يفترض هذا النموذج أنَّ الفرد لا يلجأ إلى التخمين العشوائي في إجاباته عن المفردات الاختبارية لأنَّ هذا التخمين يقلل من ثبات الاختبار(Hambleton & Swaminathan,1985,p.38–39) ويرى أنَّ التقليل من لجوء الفرد إلى التخمين يعتمد على قدرة مُعد الاختبار ومهاراته في وضع البدائل غير الصحيحة بصورة تجذب المفحوصين، ويمكن تحقيق هذا أيضاً في حالة ما إذا كان الاختبار مناسباً لقدرة المفحوصين. ويضيف (Eliis & Van Der & Wollenberg, 1993, p:417 المشار اليهم في إسماعيل،2007، ص41) افتراضاً سادساً لنموذج راش وهو:

- **التجانس الموضعي (المركزي)**

ويعني أنَّ الموضوعات التي يتم تمثيلها بنفس القيمة على السمة الكامنة تكون لها احتمالات استجابة مترابطة و متطابقة.

إضافة الى بعض الافتراضات الآخرى الواجب مراعاتها عند استخدام نموذج راش وأهمها:

- دراسة بعض المعلومات والمعرفة الرياضية لإجراء العمليات الحسابية.
- التحكم في البرمجيات والحاسوب خاصة البرامج الإحصائية.
- تطبيق واستخدام عدد كافٍ وكبير لتقدير معالم النموذج.
- أن يكون التمييز متماثلاً من خلال التفاعل بين الأفراد والمفردات داخل اطار محدد من المعرفة.

## 6.طرق تقدير نموذج راش

### 1.6.مؤشرات مطابقة الفرد المستندة إلى الأرجحية *Likelihood-based Person-fit Indices*

تعتبر مؤشرات مطبقة الفرد المستندة إلى الأرجحية الأكثر مؤشرات مطابقة الفرد شيوعاً في الكشف عن أنماط الاستجابة غير المطابقة لأنَّها تمتاز بسهولة الحساب، والفهم ومن الطرق الشائعة لتقدير وحساب مؤشرات مطابقة الفرد:

#### 1.1.6.طريقة الارجحية العظمى *(Maximum Likelihood Estimation) (MLE)*

والتي يتم فيها تحديد دالة صعوبة الفقرة ودالة قدرة الفرد ثمَّ تبسيط المعادلتين وتصنيف الأفراد إلى مجموعات تبعاً لدرجاتهم الكلية على الاختبار، ثمَّ تعيين عدد الأفراد في كل مجموعة من مجموعات الدرجات الكلية، ثمَّ حل المعادلتين بواسطة الإعادة المتعاقبة حتى التوصل إلى قيَّم لمعالم الصعوبة والقدرة التي تحقق المعادلتين، ثمَّ يتم حساب الأخطاء المعيارية لصعوبة البند وقدرة الفرد من خلال الاشتقاق الثاني من دالة لوغاريتم الترجيح الأعظم (كاظم، 1988 أ، ص.61).

ويعتمد تقدير قيمة القدرة($\theta$) في هذه الطريقة على نمط استجابة المفحوص على فقرات الاختبار ويرمز للدالة الأرجحية لنمط الاستجابة على مجموعة من الفقرات بالرمز $L(u\backslash\theta)$ وتعطى دالة الأرجحية في الاختبارات مختلطة الشكل بالمعادلة التالية: (Tao, Shi & Chang, 2012 ، المشار إليهم في طوالبة، 2018، ص.20).

$$L(u\backslash\theta) = \left(\prod_{i=1}^{n} p_i(\theta)^{U_i}\, Q_i(\theta)^{1-U_i}\right)\left(\prod_{i=1}^{n}\prod_{k=0}^{m}(p_{jk(\theta)})^{u_{jk}}\right) \dots\dots\dots (13)$$

ففي الجزء الأول من المعادلة$(\prod_{i=1}^{n} p_i(\theta)^{U_i}\, Q_i(\theta)^{1-U_i})$ يعبر عن دالة الأرجحية في الفقرات ثنائية التدريج وتمثل$(U_i)$ نمط استجابة المفحوص $(1،0)$ وتعبر$(p_i)$ عن احتمال الإجابة الخاطئة عن الفقرة$(i)$ ، وتعبر$Q_i(\theta)$ عن احتمال الإجابة الخاطئة، وتعبر$(n)$ عن عدد الفقرات.

ويعبر الجزء الثاني من المعادلة$(\prod_{i=1}^{n}\prod_{k=0}^{m}(p_{jk(\theta)})^{u_{jk}})$ عن دالة الأرجحية في الفقرات متعددة التدريج، حيث تمثل$(m)$ فئة الاستجابة، وتعبر$(u_{jk})$ عن نمط استجابة المفحوص $u = (u_{1k}, u_{2k}, \dots, u_{jk}, u_j + 1k, \dots, u_{nk})$ وتكون $u_{jk} = 1$ إذا وقعت استجابة المفحوص على الفقرة$(i)$ في الفئة$(k)$ وتكون $u_{jk} = 0$ إذا لم تقع استجابة المفحوص في الفئة$(k)$ تمثل$p_{jk(\theta)}$ احتمالية الفئة عند مستوى القدرة$(\theta)$ (طوالبة، 2018، ص.19– 20).

ويتم الحصول على تقديرات القدرة$(\theta)$ عن طريق إيجاد المشتقة الأولى للوغاريتم دالة الأرجحية ومساواتها بالصفر حسب المعادلة التالية:

$$\frac{\partial}{\partial\theta} lnL(u\backslash\theta) = 0 \dots\dots\dots\dots (14)$$

ويمكن استخدام عمليات التقريب للحصول على أكثر القيم دقة لتقدير القدرة$(\theta)$ كطريقة نيوتن رافسون(Newton Raphson) للحصول على تقديرات ثابتة لمعلم القدرة حسب العلاقة التالية: (Hambleton, & Swaminathan,1985)

$$\theta_{m+1} = \theta_{m-}\frac{\left[\frac{\partial}{\partial\theta} ln\,(u\backslash\theta)\right]}{\left[\frac{\partial^2}{\partial\theta^2} ln\,(u\backslash\theta)\right]} \dots\dots\dots\dots (15)$$

حيث:

$\theta_m$ : تمثل القيمة $m$ للقدرة المقدرة

$\theta_{m+1}$: التقدير الأكثر دقة للقدرة

ويتم تكرار العملية حتى يصبح الفرق بين$(\theta_m)$ و$(\theta_{m+1})$ أقل من مستوى الفرق المحدد $(0.005)$ عندها يتم الحصول على القدرة بطريقة الأرجحية العظمى(Ho,2010 ، المشار إليه في طوالبة، 2018 ، ص.21).

- **عيوب استخدام طريقة الارجحية العظمى**

أشار كل من (Ho, 2010 ;Kim & Nicewander,1993; Bock & Mislevy,1982) إلى أنَّ هناك عدة عيوب في استخدام هذه الطريقة منها: (طوالبة، 2018، ص.21).

- أنَّ التحيز في تقدير القدرة يرتبط بعلاقة إيجابية مع مستوى القدرة($\theta$) وبالتالي فإنَّ التقديرات تكون أعلى من المتوقع عند مستويات القدرة العالية وأقل من المتوقع عند مستويات القدرة المتدنية.
- يصعب الحصول على تقديرات للقدرة عندما تكون إجابة المفحوص على جميع الفقرات صحيحة ($\theta = +\infty$) أو تكون جميع إجابات المفحوص خاطئة ($\theta = -\infty$).
- تولَّد طريقة الأرجحية العظمى قيَّماً غير مقبولة عند وجود أنماط استجابة شاذة على الفقرات.

### 2.1.6. طريقة الأرجحية الموزونة *Weighted Likelihood Estimation (WLE)*

قدم وارم (Warm,1989) هذه طريقة كبديل للتقليل من التحيز في تقدير القدرة الذي عانت منه طريقة الأرجحية العظمة(MLE) كمشكلة من مشكلات استخدامها، حيث تعتمد طريقة الأرجحية الموزونة(WLE) في تقدير القدرة على موازنة دالة الأرجحية ببعض الدوال الأخرى المناسبة، ويتم الحصول على تقدير الأرجحية الموزونة للقدرة من خلال المعادلة.

$$\frac{\partial lnL(u\backslash\theta)}{\partial\theta} + \frac{\partial\, ln\, w(\theta)}{\partial\theta} = 0 \dots\dots\dots (16)$$

حيث تمثل:

($\frac{\partial lnL(u\backslash\theta)}{\partial\theta}$): المشتقة الأولى لدالة وزن الأرجحية $w(\theta)$ ، ويعبر عنها بالدالة التالية:

$$\frac{\partial\, ln\, w(\theta)}{\partial\theta} = -Bias\big(MLE(\theta)\big) \times I(\theta) \dots\dots\dots (17)$$

يمثل: $Bias\big(MLE(\theta)\big)$ التحيز في تقدير القدرة بطريقة(MLE) وتمثل$I(\theta)$ دالة المعلومات عند مستوى القدرة($\theta$)(طوالبة، 2018، ص.22).

### 3.1.6. طريقة كوهين التقريبية

هي من أكثر الطرق اقتصاداً في تقدير معالم نموذج راش، وتقترب التقديرات الناتجة منها من طريقة الأرجحية القصوى غير المشروطة، ويزيد هذا التقارب في التقدير بالنسبة للاختبارات الطويلة نوعاً ما أو تلك التي تتوزع درجاتها بصورة مماثلة على وجه العموم، حتى يصل الاختلاف بينها إلى

مجرد الكسر من الخطأ المعياري لهذه التقديرات، ولتطبيق هذه الطريقة التقريبية هناك ثلاث خطوات. (كاظم، 1988أ، ص.65– 66).

**أ. تعيين تقديرات أولية لمعالم كل من صعوبة المفردات وقدرات الأفراد وتباينها**

حيث يكون التقدير الأولي لمعلم الصعوبة للبند$(i)$ هو$(d_i^\circ)$ حيث:

$$d_i^\circ = ln\left[\frac{N - s_i}{s_i}\right] - \sum_i^L ln\left[\frac{N - S_i}{S_i}\right] \quad /l \quad i = 1, L \dots\dots\dots (18)$$

ومنها تحسب القيمة (D) حيث:

$$\mathrm{D} = \sum_{\mathrm{i}}^{\mathrm{L}} \left(\mathrm{d}_{\mathrm{i}}^{\circ}\right)^2 / (\mathrm{L} - 1) \quad (2.89) \dots\dots\dots (19)$$

كما يكون التقدير الأولي لمعلم قدرة الفرد الحاصل على الدرجة الكلية هو $(r)$ هو$(b_r^\circ)$ حيث:

$$b_r^\circ = \ln\left[\frac{r}{L - r}\right] \qquad r = i, L - i \dots\dots\dots (20)$$

ومنها يكون

$$b^\circ = \frac{\sum_{r=1}^{L-1} n_r \; b_r^\circ}{N} \quad \dots\dots\dots (21)$$

من المعادلتين (20،21) تحسب القيمة (B) (كاظم، 1988، ص.65– 66)

$$B = \sum_{r=1}^{L=1} b_r (b_r^n - b^\circ)^2 / (N - 1) \; (2.89) *$$

**ب.حساب معاملي الامتداد *Expansion Coefficient***

- **معامل الامتداد لصعوبة البند** $(X)$: يختص بتصحيح التقدير الأولي لمعلم صعوبة المفردة من تأثير اتساع مدى قدرة أفراد العينة، ويعرف هذا المعامل بالمعادلة التالية: (كاظم، 1988أ)

$$X = [(1 + B)/(1 - BD)]^{1/2} \dots\dots\dots (22)$$

- **معامل الامتداد لقدرة الفرد (Y):** تختص بتصحيح التقدير الأولي لمعلم قدرة الفرد من تأثير مدى صعوبة مفردات الاختبار، ويعرف هذا المعامل بالمعادلة التالية:

$$Y = [(1 + D)/(1 - BD)]^{1/2} \dots\dots\dots (23)$$

* المقدار 2.89= $(1.7)^2$ وحيث أن 1.7 هو معامل التدريج الذي يحول المنحنى اللوغاريتمي إلى تطابق تقريبي مع المنحنى الاعتدالي.

**ج. حساب التقديرات النهائية للمعالم وأخطائها المعيارية**

يحسب التقدير النهائي لمعلم صعوبة المفردة $(d_i)$ وذلك بضرب معامل الامتداد$(X)$ في التقدير الأولي لمعلم صعوبة المفردة أي:

$$di = Xd_i^{\circ} \quad i = 1, L \dots \dots \dots (24)$$

أما الخطأ المعياري لمعلم الصعوبة فهو:

$$SE(d_i) = X[N/S_i(N - S_i)]^{1/2} \dots \dots \dots (25)$$

يحسب التقدير النهائي لمعلم قدرة الفرد$(b_r)$ وذلك بضرب معامل الامتداد$(Y)$ في التقدير الأولي لمعلم قدرة الفرد أي:

$$b_r = Yb_r^{\circ} \qquad r** = i, L - i \dots \dots \dots (26)$$

والخطأ المعياري لمعلم القدرة هو:

$$SE(b_r) = Y[L/r(L - r)]^{1/2} \dots \dots \dots (27)$$

وقد أشارت أمينة كاظم (1988) أنَّه لوحظ بالنسبة للاختبارات الطويلة نوعاً ما، أو التي لها توزيع درجات متماثل، أنَّ هذه الخطوات التقريبية تعطي تقديرات تختلف بمقدار كسر من الخطأ المعياري من القيم التي نحصل عليها من طريقة الارجحية العظمى (كاظم، 1988أ، ص.68).

### 4.1.6. طريقة بييز في تقدير القدرة *(Bayesian Ability Estimation)*

من أكثر الطرق استخداما في تقدير معلمي نموذج راش وتعتمد هذه الطريقة على افتراض توزيعات معينة مسبقة للمعالم يمكن اشتراطياً من خلالها الحصول على توزيعات بعدية تمكن من الوصول إلى تقديرات أكثر دقة، ويتم تقدير قدرات الأفراد بطريقتين هما:

**أ. طريقة التقدير البعدي المتوقع *(Expected A Posteriori) (EAP)***

يشير(Bock & Aitkin,1981) بأنَّ تستخدم هذه الطريقة قيمة الوسط الحسابي للتوزيع البعدي (posterior) كتقدير للقدرة(θ) ويتم التعبير عن الوسط الحسابي والتباين والتوزيع البعدي$(p(\theta/u)$ بطريقة التقدير البعدي المتوقع حسب المعادلتين (طوالبة، 2018، ص.23).

$$E(\theta \backslash u) = \int_{-\infty}^{\infty} \theta P(\theta \backslash u) d\theta \dots \dots \dots (28)$$

** أي من الدرجة الكلية (r=1) إلى الدرجة الكلية (r=L-1)

$$Var(\theta\backslash u) = \int_{-\infty}^{\theta} \theta^2 \, P(\theta\backslash u) d\theta - E(\theta\backslash u)^2 \dots\dots\dots (29)$$

وباستخدام نقاط التربيع (quadrature points) يتم إجراء عمليات تقريب للمعادلتين السابقتين للحصول على التقديرات الأكثر دقة للقدرة (EAP) حسب المعادلتين(Wang & Vispoel,1998 كما ورد في طوالبة ،2018).

$$\hat{\theta} = EAP(\theta) = E(\theta\backslash u) = \frac{\sum_{k=1}^{q} X_k L_j(X_k) W(X_k)}{\sum_{k=1}^{q} L_j(X_k) W(X_k)} \dots\dots\dots (30)$$

$$\hat{\sigma}^2(\hat{\theta}) = Var(\theta\backslash u) = \frac{\sum_{k=1}^{q} (X_k - \hat{\theta})^2 L_j(X_k) W(X_k)}{\sum_{k=1}^{q} L_j(X_k) W(X_k)} \dots\dots\dots (31)$$

حيث تمثل:

$X_k$: أحد نقاط التربيع.

$W(X_k)$: الوزن المرتبط بنقطة التربيع المقابلة.

$L_j$: دالة الأرجحية بعد $j$ من المفردات.

تمتاز هذه الطريقة بأنَّ قيم(θ) المقدرة تقترب من القيم الحقيقية بزيادة عدد الفقرات، وتوزيع تقديرات القدرة له تباين أقل من توزيع القدرة الحقيقية، كما تولد هذه الطريقة قيماً غير مقبولة للقدرة عند وجود أنماط استجابة غير مطابقة لنموذج الاستجابة للفقرة المستخدمة (Bock & Mislevy ; lord, 1980 ; Ho, 2010 كما ورد في طوالبة).

ومن عيون هذه الطريقة:

- التحيز في تقدير القدرة قد يزداد عند دمج المعلومات السابقة مع أرجحية الاستجابة على الفقرة
- كما يمكن أن يتأثر التوزيع البعدي بخصائص التوزيع القبلي.

**ب. طريقة التقدير البعدي الأعظم *(Maximum A Posteriori) MAP***

قامت Samejima(1969) بتقديم هذه الطريقة كبديل لطريقة الأرجحية العظمى، وفيها يدمج التوزيع القبلي للقدرة($\theta$) في عملية التقدير، وتستخدم هذه الطريقة القيمة المنوالية(mode) للتوزيع البعدي كتقدير للقدرة($\theta$) كما يتم تقدير القدرة بطريقة التقدير البعدي الأعظم لنمط الاستجابة($u$) عن طريق حساب المشتقة الأولى للوغاريتم التوزيع البعدي للقدرة ومساواتها بالصفر حسب المعادلة التالية

(طوالبة، 2018).

$$\frac{\partial log[P(\theta\backslash u)]}{\partial \theta} = \frac{\partial log[L(\theta\backslash u)]}{\partial \theta} + \frac{\partial log[P(\theta)]}{\partial \theta} = 0 \dots \dots \dots (32)$$

حيث تمثل:

$P(\theta\backslash u)$ : التوزيع البعدي للقدرة $(\theta)$ لنمط الاستجابة $(u)$.

$P(\theta)$ : التوزيع القبلي للقدرة $(\theta)$.

$L(\theta\backslash u)$: دالة الأرجحية لنمط الاستجابة $(u)$ .

ويتم بعد ذلك استخدام طرق التدوير المتعاقب (iterative method) كطريقة (نيوتن رافسون) للحصول على التقدير الأكثر دقة للقدرة، وتؤثر العديد من العوامل على فاعلية مؤشرات مطابقة الفرد مثل شكل توزيع معالم الفقرات (الصعوبة والتمييز) وكذا نموذج الاستجابة للفقرة المستخدم لتقدير معالم الفقرات والأفراد.

**7.تدريج فقرات الاختبار**

يقصد بتدريج الفقرات بأنها إنشاء ميزان لتدرج الصفة السلوكية الذي يحدد مستوى الأفراد على هذه الصفة مقدراً بوحدة معروفة، ويتمثل هذا في انتظام مواقع الفقرات على متصل صعوبتها حول صفر تدريج بحيث يعبر موقع كل فقرة عن صعوبتها (الشرقاوي وآخرون، 1996).

ويذهب القريشي إلى تعريف تدريج الفقرات بأنَّها انتظام مواقع الفقرات على متصل صعوبتها حول صفر تدريج متغير مستوى الصعوبة بحيث يعبر موقع كل فقرة (بعدها عن صفر التدرج) عن مستوى صعوبتها أي أنَّ تدريج الفقرات هو وضع ميزان يحدد بانتظام مواقع الفقرات على متصل صعوبتها حول صفر إذ يعبر موقع كل فقرة مستوى صعوبتها ولتدريج الفقرات يجب:(زكري، 2007، ص.51)

- تحديد معايير اختيار الفقرات التي يشمل عليها الاختبار.
- إحصاء الملائمة بمعنى أن يكون حساساً لأنماط الاستجابة غير المتوقعة للمفحوصين الذين يقرب مستوى قدرتهم من مستوى صعوبة الفقرة.
- الأخطاء المعيارية لتدرجات الفقرات تمثل دلالات لثبات الفقرة، وكلما كانت الأخطاء المعيارية صغيرة كانت المعلومات التي يشتمل عليها الاختبار دقيقة.

- توزيع تدرجات الفقرة لا بدَّ أن يكون بشكل مستمر ومتغير.
- محتوى فقرات الاختبار لا بدَّ أن يكون لفظاً أو عددياً ويرتبط بالقدرة المراد فيها القياس.

**8.مفهوم الصدق والثبات وفق نموذج راش:**

**أ. مفهوم الصدق**

يشير مفهوم صدق القياس وفق نموذج راش "بأن تعرف الفقرات فيما بينها متغيراً واحداً، أي أنَّ فقرات الاختبار تتدرج من حيث صعوبتها بحيث تعرف متغيراً واحداً، كما يعني تدرج قدرات الأفراد على المتغير محددة تقديرات أدائهم على هذا الاختبار (زكري، 2007، ص.106).

وهذا يوضح أنَّ كلاً من صعوبة الفقرات، وقدرات الأفراد تتدرج على متصل واحدٍ يمثل متغيراً واحداً. ويعد صدق القياس من الخصائص المهمة في تحقيق متطلبات هذا النموذج، فيمثل أول مطالب الموضوعية في أداة القياس وتتمثل في أن تعرف الفقرات فيما بينها متغيراً واحداً أي أنَّ فقرات الاختبار تتدرج من حيث صعوبتها بحيث تعرف متغيراً واحداً، وكذلك تدرج قدرات الأفراد على المتغير محددةٍ تقديرات أدائهم على هذا الاختبار، وهذا يوضح أنَّ كل من صعوبة الفقرات وقدرات الأفراد تتدرج على متصل واحد يمثل متغيراً واحداً.

ولذلك تتيح اختبارات الملاءمة المختلفة الخاصة بنموذج راش حذف كل من الأفراد غير المناسبين وكذلك الفقرات غير المناسبة ونستبقي فقط استجابات الأفراد الصادقة، وكذا الفقرات الصادقة في تدرجها على متصل قياس هذه السمة.

**ب.مفهوم الثبات**

أما مفهوم ثبات القياس في ضوء نموذج" راش فهو "بتحقق ما بقي من مطالب الموضوعية في القياس، أي عندما تستخدم أداة القياس التي أنشئت باستخدام هذا النموذج، حيث يتحقق:

- استقلال القياس عن الاختبار المستخدم.
- استقلال القياس عن مجموعة الأفراد المؤدية للاختبار.

وهذا يعني أنَّ استقلالية القياس وتحرره الذي يوفره نموذج راش تتيح الفرصة لثبات القياس بحيث لا يختلف القياس سواءً لقدرة الفرد أو صعوبة الفقرة باختلاف عينة التدريج، أو باختلاف الاختبار المستخدم لقياس قدرة ما.

وتعتبر عملية القياس في ميدان التربية أو علم النفس من العمليات المعقدة نظراً لطبيعة الظواهر الإنسانية التي تمتاز بالتعقيد، والتداخل والنسبية، حيث يصعب معها فصل مؤشرات ظاهرة عن مؤشرات ظاهرة أخرى، كما تُعد الوسائل المعتمدة في القياس عاملاً مؤثراً من حيث موضعيتها وقدرتها على قياس الخصائص بدقة والاقتراب قدر الإمكان من حقيقة المتغيرات النفسية أو التربوية المراد قياسها، وبناء على ذلك فإنَّ وجود أدوات أي اختبارات نفسية تتميز بالمصداقية والموضوعية أمر ضروري للوصول إلى نتائج حقيقية يمكن الاعتماد عليها في عملية التشخيص ونتيجة التكامل بين مناهج القياس النفسي وبعض النظريات الرياضية والإحصائية من جهة، ونظريات علم النفس من جهة أخرى، قد أدى إلى ميلاد نظريات القياس النفسي(Psychological measurement) التي أصبحت الخلفية الضرورية لأي عملية قياس تتوخى الموضوعية.

إنَّ التقدير المتحرر لكل من قدرة الفرد، وصعوبة الفقرة وتدريجهما على ميزان خطي يعطينا تقديرات مباشرة لنمذجة تباين الخطأ لكل من تقديرات القدرة، وصعوبة الفقرة، لذلك يكون الثبات هو مدى الدقة في تقدير موقع كل من الفقرات، والأفراد على متصل السمة التي نهدف إلى قياسها (علام، 2001، ص.500).

ويحسب نوعان من معاملات الثبات الأول يطلق عليه اسم معامل الفصل بين الفقرات ( Item separation Index) ويمثل النسبة بين الانحراف المعياري للقيم التقديرية المحررة لصعوبة الفقرات ومتوسط الخطأ المعياري لهذه القيم، وبذلك فإنَّ هذا المعامل سيمثل نسبة التباين الحقيقي إلى التباين الملاحظ، وأنَّ قيمة معامل الفصل ستمتد من (0،∞−) ويجب أن تكون أكبر من (1) لكي يكون مقبولاً.

أما الثاني فيسمى معامل الفصل بين الأفراد (Person separation Index) ويمثل نسبة بين الانحراف المعياري للقيم التقديرية المحررة لقدرة الأفراد ومتوسط الخطأ المعياري لهذه القيم وبذلك فإنَّ معامل الفصل سيمثل نسبة التباين الحقيقي إلى التباين الملاحظ وأنَّ قيمة معامل الفصل ستمتد من (صفر، − ∞) ويجب أن تكون أكبر من (1) لكي يكون مقبولاً(Schumacher, 2006, p.5) .

## 9. التقديرات البارامترية لنموذج راش

### 1.9. تقدير قيم صعوبة المفردات وقدرة الأفراد *Difficulty and ability estimates*

عند تطبيق نموذج الاستجابة للفقرة على اختبار تحصيلي، أو اختبار للقدرة العقلية، يتم تقدير

بارامتري صعوبة الفقرة وقدرة الفرد بخطوات قبل التحليل، وهي ثلاث خطوات قبل التحليل ثمَّ تتم عملية التحليل التي تتضمن الخطوتين الأخيرتين كما يلي:(حمَّاد، 2010).

**أ- إنشاء مصفوفة الاستجابات**

تتبنى مصفوفة لنتائج استجابات مجموعة من الأفراد ($N$)على مجموعة من فقرات اختبار ما (L) محورها الأفقي يمثل الأفراد، ومحورها الرأسي الفقرات، وخلاياها تمثل استجابات كل فرد على كل فقرة من فقرات الاختبار، تعطي استجابة الفرد($i$) على الفقرة(i) القيمة (1) في حالة النجاح و(0) في حالة الإخفاق، فتكون قيمة كل خلية من خلايا هذه المصفوفة إمَّا (1) أو (0) ومجموع قيَّم خلايا كل عمود يعطي الدرجة الكلية لكل فرد، ومجموع خلايا الصفوف يعطي عدد الأفراد الذين أجابوا إجابة صحيحة على الفقرة (حمَّاد، ص.50-53).

**ب- حذف الأفراد غير الملائمين**

يحذف كل فرد أخفق في حل فقرات الاختبار (أي حصل على درجة صفر)، حيث يعد عندئذ أقل من مدى مستوى الاختبار، كما يحذف كل فرد نجح في حل جميع فقرات الاختبار (أي حصل على الدرجة النهائية) حيث يعد أعلى من مدى مستوى الاختبار، ويكون هؤلاء الأفراد غير ملائمين أي غير مناسبين للإجابة على هذا الاختبار.

**ج- حذف الفقرات غير الملائمة**

تحذف كل فقرة يخفق في الإجابة عليها جميع الأفراد، حيث يعد أعلى من مستوى الأفراد وكذلك الحال بالنسبة لكل فقرة يجيب عليها جميع الأفراد إجابة صحيحة، حيث يعد تحت مستوى الأفراد وتكون هذه الفقرات غير ملائمة، أي غير مناسبة لاستجابة الأفراد عليها (حمَّاد، 2010، ص.50-53).

ويعد تقدير قيَّم بارامتري النموذج (صعوبة المفردات، قدرة الأفراد) من المشكلات الرئيسة عند استخدام نماذج الاستجابة للفقرة في تحليل البيانات التي من المفترض أن تكون هذه التقديرات دقيقة وغير متحيزة، لذا وجدت طرق عديدة لتقدير بارامترات النموذج المستخدم، حيث تعتمد هذه الطرق على أساليب التحليل العددي (Numerical Analysis) والتي تتطلب برامج خاصة للحاسوب ومن البرامج التي تعتمد على الحاسوب: مثل

- برنامج (BICAL) الذي صممه رايت وميد (Wright and Mead, 1976) الذي يعتمد على البواقي المعيارية.

- وكذلك نموذج (CALFT) الذي اقترحه رايت وستون (Wright and Stone) والذي يعتمد على البواقي المعيارية أيضاً.

- برنامج (Pmi) الذي صممه جاستافسون(Gustaffson) ويعتمد على دالة الأرجحية القصوى المشروطة(CMI).

- برنامج مايكرسكيل (MICR, SCALE) الذي صممه رايت (Wright, 1984) وهو برنامج شامل يضم نماذج ثيرستون، وليكرت، وجثمان، وراش في تحليل شامل.

- برنامج (BILOG-MG3) ويعد من البرامج القوية الفعالة التي تتسم بالمرونة والسرعة في تقدير بارامترات نموذج راش وبيرنيوم ولورد باستخدام الارجحية القصوى (علام، 2001، ص.239–240).

- برنامج راسكال (RASCAL) هو برنامج حاسوب مصمم بلغة(dBase) لتدريج الفقرات الاختبارية بأسلوب نموذج راش من شركة (Micro At) عام 1992(عوض، 2000، ص.140).

- برنامج Winsteps من تصميم (Mike Linacre) ويعد من البرامج القوية السهلة الاستخدام ويعمل على بيئة النوافذ ويمكن استخدامه في إدخال مفردات اختبارية يبلغ عددها (60) ألف مفردة و(10) ملايين لعدد الأفراد الذين يختبرون بهذه المفردات (موقع برنامج,Winsteps, 2021) .

### 2.9. اختبار حسن المطابقة لنموذج راش Test of in the Rasch models

تعد مطابقة أو ملاءمة الفقرات لنموذج راش خطوة أساسية في إجراءات تحليل الفقرات باستخدام نموذج راش، وينبغي الإبقاء على الفقرات التي تتطابق مع افتراضات النموذج في الصيغة النهائية للاختبار وحذف الفقرات غير المطابقة للنموذج.

ويستهدف اختبار حسن المطابقة تحديد ما إذا كانت قيم معلم النموذج التي تقدر من نتائج الاختبار تحقق الفروض التي يرتكز عليها النموذج، فإذا تحققت الفروض يمكن القول أنَّ النموذج يفسر أو يصف ما يحدث من تفاعل بين الأفراد وفقرات الاختبار (علام، 1986، ص.141). ومطابقة البيانات تبدأ من تقدير بارامترات النموذج من النتائج الفعلية للاختبار، هذه البارامترات تستخدم في تحديد مقدار احتمال توصل الفرد إلى الإجابة الصحيحة عن كل فقرة، وذلك لكل مجموعة من

مجموعات الأفراد الذين حصلوا على نفس الدرجة الكلية في الاختبار، ثمَّ يتم حساب العدد المتوقع للأفراد الذين ينبغي أن يجيبوا عن الفقرة إجابة صحيحة، يليه تطبيق مقارنة إحصائية بين العدد المتوقع والعدد الفعلي في كل هذه المجموعات (حمَّاد، 2010، ص.53).

ويعد تحليل أسباب عدم مطابقة مفردة ما لنموذج راش في الحقيقة أمراً أساسياً لاختبار حسن المطابقة، فعن طريق هذا التحليل نستطيع معرفة المفردات المتحيزة، والمفردات التي أعتمد الطلبة في إجاباتها على التخمين، والمفردات التي لم يبذل الطلبة ذوي القدرة المرتفعة الجهد اللازم للإجابة عنها وقد قام كل من ميد و والر(1976) بدراسة مشكلة التخمين في الإجابة عن مفردات الاختيار من متعدد واستطاعا باستخدام تحليل أسباب عدم مطابقة المفردة لنموذج راش، معرفة الأفراد الذين اعتمدوا على التخمين العشوائي في إجاباتهم، فقد بني هذا التحليل على دراسة(احتمال) توصل هؤلاء الأفراد إلى الإجابة الصحيحة عن المفردة طبقاً للنموذج، فإذا كان احتمال توصل فرد إلى الإجابة الصحيحة عن مفردة ما أقل من قيمة معينة ولتكن(0.05) ولكنَّه أجاب بالفعل عنها إجابة صحيحة يمكن اعتبار أنَّ هذا الفرد اعتمد التخمين في الإجابة، وبذلك لا تستطيع قياس قدرته باستخدام هذه المفردة(علام، 1986، ص.151).

**10. وحدة قياس (قدرة الفرد وصعوبة المفردة)**

إن الهدف من وضع اختبار حسب نظرية الاستجابة للمفردة هو تحديد قدرة المفحوصين والقيام بمقارنتهم ببعضهم البعض، وذلك من خلال تقدير قدرة المفحوصين وتقدير معلم صعوبة الفقرة حيث يمكننا الوصول إلى ذلك من خلال إجابة عدد من المفحوصين على عينة من الفقرات ويتم تقدير استجابة الفرد على الفقرة بواحد في حالة الإجابة الصحيحة على الفقرة وصفر في حالة الإجابة الخاطئة على الفقرة وتشير كاظم أنَّه قبل البدء في عملية التحليل يجب:(كاظم، 1988أ، ص.60-61).

- حذف كل فرد فشل في كل بند من بنود الاختبار لانَّ مستوى هذا الفرد أقل من مستوى الاختبار.
- حذف كل فرد نجح في حل كل بند من بنود الاختبار لأنَّ مستوى هذا الفرد أعلى من مستوى الاختبار.
- حذف كل بند أخفق جميع الأفراد في الإجابة عليه لأنَّ هذا البند أعلى من مستوى الأفراد.
- حذف كل بند نجح جميع الأفراد في الإجابة عليه لأنَّ هذا البند أقل من مستوى الأفراد.

وتندرج صعوبة الفقرة وقدرة الأفراد على ميزان مقياس واحد، وأنَّ وحدة قياس كل من قدرة الفرد وصعوبة الفقرة تقدر بوحدة قياس واحدة مشتقة مباشرة من نموذج راش(كاظم، 1988أ)، وتعد وحدة القياس هذه وحدة عامة تستخدم في قياس قدرات الأفراد بواسطة الاختبارات التي تبنى على أساس هذا النموذج، وهذه الوحدة تعد من الوحدات المناسبة التي يمكن التعامل بها في القياس النفسي والتربوي ويسهل تحويلها إلى وحدات قياس أخرى تناسب التطبيقات المختلفة للاختبارات وهذه الوحدة الرياضية العامة تسمى الترجيح اللوغاريتمي(علام، 1986، ص.131).

وتضيف أمينة كاظم (1988) أنَّه عندما يجيب الفرد $(v)$على الفقرة$(i)$ فإن أرجحية حدوث أي من الاستجابتين (نجاح، فشل) يعتمد فقط على قدرة الفرد وصعوبة الفقرة، ويحددها المقابل اللوغاريتمي للفرق بين هذين المعلمين مما يعني أنَّ الاستجابة الصحيحة تتوقف في حالة قدرة الفرد أكبر من صعوبة الفقرة والاستجابة الخاطئة في حالة صعوبة الفقرة أكبر من قدرة الفرد. (ص.55).

واحتمال أن يجيب فرد إجابة صحيحة عن فقرة صعوبتها تساوي 0 هو:

$$\frac{e-\delta}{1+e-\delta} \dots\dots\dots (33)$$

$$\boldsymbol{e\beta = \frac{P}{1-P}} = \frac{\text{احتمال حدوث الاجابة الصحيحة}}{\text{احتمال حدوث الاجابة الخاطئة}}$$ ترجيح الاجابة الصحيحة هو

وبتحويل الصورة الأخيرة إلى الصورة اللوغاريتمية فإنβ = اللوغاريتم الطبيعي لمرجح الإجابة الصحيحة وكذلك فانَّ قياس صعوبة الفقرة بنفس الوحدة (Logit) هو عبارة عن اللوغاريتم الطبيعي لترجيح الإجابة الخطأ عن هذه الفقرة للأفراد الذين تقع قدراتهم عند نقطة ميزان القدرة وعليه فإنَّ احتمال أن يجيب الفرد الذي قدرته (β =0) إجابة صحيحة عن الفقرة صعوبتها δ هو:

$$\frac{e\beta}{1+e\beta} \dots\dots\dots (34)$$

وحيث أنَّ ترجيح الإجابة الخطأ هو:

$$\boldsymbol{e\delta = \frac{1-P}{P}} = \frac{\mathbf{1} - \text{احتمال حدوث الاجابة الصحيحة}}{\text{احتمال حدوث الاجابة الخاطئة}}$$

وبتحويل المعادلة للصورة اللوغاريتمية فإن δ = اللوغاريتم الطبيعي لمرجح الإجابة الخاطئة إذن تعرف قدرة الفرد أنَّها اللوغاريتم الطبيعي لمميز نجاح الفرد على الفقرات التي تعتبر نقطة صفر التدريج عن صعوبتها.

أما صعوبة البند، هي اللوغاريتم الطبيعي لمميز الفشل لدى الأفراد التي تعتبر نقطة صفر التدريج عن صعوبتها وتقدر باللوجيت (كاظم، 1988أ، ص.56).

**أ. تعريف وحدة القياس (اللوجيت)**

هي وحدة قياس كل من قدرة الفرد وصعوبة الفقرة، وتعرف باللوغاريتم الطبيعي لمميز نجاح الفرد على الفقرات التي تعتبر نقطة صفر التدرج عن صعوبتها عندما يساوي هذا المميز ثابتا هو الأساس الطبيعي($e$) أي (2.72) ويكون عندئذ احتمال نجاح هذا الفرد (0.73).

فالوحدة المستخدمة هنا في قياس كل من القدرة الكامنة ودرجة صعوبة الفقرة هي اللوجيت ويمكن تعريفها من خلال إيجاد اللوغاريتم الطبيعي لناتج قسمة احتمال الإجابة عن الفقرة على احتمال عدم الإجابة عنها، لكل درجة كلية لكل فقرة (التقي، 2009، ص.19).

وتحقق وحدة القياس اللوجيت مميزات التدريج ذو الوحدات المتساوية، ولكنَّها قد تتضمن قيماً سالبة وصفرية، كما قد تكون عدداً صحيحاً أو كسرياً، مما قد يؤدي إلى صعوبة التفسير لدى الباحثين أو المعلمين(Gregory, 2004) .

ويمكن تلافي التقديرات السالبة لكل من القدرة والصعوبة وذلك بتعديل نقطة أصل موضع الصفر في هذا التدرج حيث يعد أمراً اعتبارياً، ويمكن أيضاً تغير حجم وحدة القياس بحيث تتلافى التقديرات الكسرية لكل من القدرة والصعوبة، وعليه فأنَّ إضافة أو طرح قيمة ثابتة إلى قيم كل من القدرة والصعوبة لا يغير قيمة هذا الاحتمال، ومن ثمَّ يمكن تحديد نقطة أصل، وبهذا يمكن التوصل إلى تناسب صدق المقياس، تدرجات جديدة بوحدات جديدة(Wright & Stone, 1979, p.201) مثل النيت(Nit) وحدة التدريج الجماعية Normative Scaling Units، والسيت(Sit) وحدات التدريج المعتمد على محك مستقل Substantive scaling units، و وحدة(الواط) حيث يتم تحويل تدريج اللوجيت إلى تدريج يمتد من القيمة(0،100) ومتوسط صعوبة المفردات يساوى(50). ويتميز هذا التدريج المئوي بسهولة التفسير لتقديرات قدرة الأفراد، ووحدة القياس البريطانية للقدرات باس (Bas) The British

Ability Scales، وكذلك وحدة(المنف) حيث تحول تدريج اللوجيت إلى تدريج له نفس خصائص تدريج الواط تدريج مئوي.

ويتم تحويل وحدة القياس باستخدام معادلة التحويل الخطي الآتية:

$$Y = a + Bx \dots\dots\dots (35)$$

X: وهو تدريج اللوجيت.

Y: وهو التدرج الجديد.

a: هي الإزاحة المطلوبة لنقطة أصل التدرج وتحديد نقطة أصل في التدريج الجديد.

B: هو عامل المسافة الذي يحدد سعة وحدة القياس الجديدة.

تحافظ المعادلة على خصائص الوحدات المتساوية لوحدة اللوجيت، وتساعد على التوصل إلى تدرجات(وحدات) جديدة متعددة تتغلب على مشكلات وحدة اللوجيت (الشرقاوي وآخرون، 1996، ص.102).

وتتنوع تدرجات الفقرات الجديدة وتختلف تبعاً لأهداف القياس المختلفة، وفيما يأتي بعض الأنواع المهمة من التدريجات.

- **وحدة التدريج الجماعية (نيت) *Normative scaling Nits***

تحويل اللوجيت إلى تدريج مبني على معيار الجماعة ويمكن حسابه بالمعادلة الآتية:

(B = 50 + 10 (b-m)/ S
(D = 50 + 10 (d-m)/ S

B: تقدير قدرة الأفراد

D: تقدير صعوبة الفقرة

M: متوسط درجات العينة

S: الانحراف المعياري (صلاح وسليمان، 2002، ص.432).

- **وحدة التدريج المعتمدة على محك مستقل *(site) (Substantive scaling units)***

ويمكن حسابه بالمعادلة الآتية:

(B = 38 + 6d)
(D = 38 + 6d)

- **وحدة التدريج الخاصة باحتمال الاستجابة الصواب*(Response probability scaling unite) (Chips)***

تعين الفرق بين قدرة الفرد وصعوبة الفقرة عبر احتمالات الاستجابة الصواب.

- **وحدة القياس المستخدمة في المقاييس البريطانية (*BAS*).**

لتناسب مقياس القدرات الفعلية ويمكن حسابه بالمعادلة الآتية:

(DBas = 10 (d1 + (1-a1))

(BBas = 10 (ar + (1-a1))

DBas = صعوبة الفقرة $(i)$ مقدرة بوحدة اللوجيت

BBas = قدرة الفرد الحاصل على درجة كلية $(r)$ بعد التحويل.

- **وحدة الواط**

حيث يتم تحويل تدريج اللوجيت إلى تدريج يمتد من صفر إلى 100، ومتوسط صعوبة الفقرات يساوي (50)، ويمتاز هذا التدريج المئوي بسهولة التفسير لتقديرات قدرة الأفراد ومن الممكن تحويل وحدة اللوجيت إلى وحدة الواط في تقدير كل من الصعوبة والقدرة باستخدام المعادلة الآتية:

$B = 50 + (15) \times b$

$D = 50 + (15) \times d$

(B,D) هما تقديرات كل من القدرة والصعوبة مقدرة بالواط.

(d،b) هما تقديرات باللوجيت (Masters, 1984, p. 140).

- **وحدة المنف**

حيث تحول تدريج اللوجيت إلى تدريج له خصائص تدريج الواط نفسها (تدريج مئوي، نقطة أصل تساوي 50) إلا أنَّ طول أو سعة وحدة القياس(منف) يختلف عن سعة وحدة الواط، ويتم الحصول على تدريج المنف(a = 50، B = 5) في معادلة التحويل الخطي السابقة (الشرقاوي، وآخرون، 1996، ص.126).

- **وحدة أكا**

هي وحدة تحقق ميزات وحدتي ألواط والمنف، التي يمكن الحصول عليها بوضع (a=50 ،B=7.41) في معادلة التحويل الخطي(35).

**11. استخدامات نموذج راش:**( إسماعيل، 2007، ص.41-42).

- يستخدم نموذج راش في القياس النفسي أي تحديـــد موضع الفرد على متصل السمة الكامنـــة أو

التحليل الاستكشافي للبيانات مثل فهم تركيب المفردة.

- يسمح النموذج بدمج المفردات التي تمَّ تنظيمها.
- يستخدم النموذج كنظرية مفسرة في العديد من المجالات المعرفية، بسبب خاصيته المرغوب فيها وهي الموضوعية.
- يستخدم النموذج في بناء وتحليل الاختبارات النفسية والتربوية، وكذلك بناء بنوك الأسئلة التي تحقق الموضوعية.
- يستخدم النموذج في حل بعض المشكلات الموضوعية في القياس سواء في مجال التحصيل أو قياس الذكاء أو غيرها.
- يستخدم النموذج لتقييم أحادية البعد، عن طريق تحليل المكونات الرئيسية للعامل باستخدام الفروق المعيارية واحصائيات الملاءمة للنموذج.

## 12. خصائص نموذج راش

تشير بشرى إسماعيل (2004) أن نموذج راش يتميز بالخصائص التالية:

- لدى جميع المفردات القدرة التمييزية ذاتها فيمل يتعلق بالسمة الكامنة.
- لا يحدث تفاعل بين الأفراد والمفردات.
- المجموع النهائي لكمية الاستجابة الموجهة لمفردات الاختبار كافياً لتقييم مقياس السمة المرتبط بهذه الكمية الموجهة بمعنى أنَّ تقييم القدرة يعتمد فقط على عدد المفردات التي تمَّ حلها وليس نماذج المفردة.
- سهولة عملية تقدير المعالم المفردة حيث يمكن تقييمها باستخدام طريقة الترجيح الأكبر المشروط ويكون التقييم مستقلاً عن مقاييس القدرة والعكس.
- وصف العلاقة بين الاستجابة والسمة عن طريق دالة متزايدة مطردة وهي المنحنى المميز للمفردة

## 13. عيوب نموذج راش

تتلخص أهم عيوب نموذج راش فيما يلي:

- يتطلب توفر خاصية الاستقلال لبارامتر نموذج راش التحقق بصورة دورية من عدم حدوث تغيرات قد تطرأ على القيم التقديرية لبارامتراته نتيجة لتغير خصائص أفراد مجتمع المختبرين.

- يفترض نموذج راش أنَّ جميع المفردات تتساوى في قدرتها على التمييز وهذا القيد يضيف قيداً آخر على انتقاء المفردات التي يشتمل عليها الاختبار.

**خلاصة الفصل**

في ضوء ما سبق في هذا الفصل، نستخلص أنَّ الاختبارات محكية المرجع ترتبط بمحتوى تعليمي محدد في ضوء محكات أداء محددة، وتسعى لتقديم نظام يتسم بالعدالة أكثر من الاختبارات معيارية المرجع، ولها شروط وقواعد عند بنائها يجب الالتزام بها.

كما تناولنا كل من النظرية الكلاسيكية ونظرية القياس الحديثة بشيء من التفصيل ومدى مساهمة كل منهما، في بناء الاختبارات خاصة مساهمة النظرية الحديثة وما ارتبط بها من نماذج مختلفة، في وضع إطارٍ علميٍ جديدٍ ووثيقٍ في بناء واختيار فقرات الاختبار، لما انطوت عليه هذه النظرية للعديد من النماذج ولعل أشهرها وأكثرها استخداماً نموذج راش الذي تطرقنا لاهم افتراضاته وشروط تطبيقه.

# البــاب الثــاني

## الفصل الرابع: اجراءات الدراسة وأدواتها

**تمـهيد**

**أولاً: منـهج الدراسة**

**ثانياً: مجتمع الدراسة وعينته**

**ثالثاً: أداة الدراسة**

**رابعاً: إجراءات البحث**

**خامساً: المعالجات الإحصائية**

**تمهيد**

يتضمن هذا الفصل تعريفاً بمنهج الدراسة وعينتها، وطريقة اختيارها ووصفها، ويتضمن أيضاً وصفاً لأداة الدراسة المتمثلة في الاختبار التحصيلي محكي المرجع في مقرر الإحصاء الوصفي لطلبة أقسام السنة الأولى علوم اجتماعية جذع مشترك، كما يتضمن الخطوات المنهجية لبناء الأداة وإجراءات تطبيقها، والتحقق من صدقها وثباتها، واستعراض الوسائل الإحصائية التي استخدمها الباحث لتحقيق أهداف الدراسة.

**1. منهج الدراسة**

تقتضي الدراسة التي تتناول بناء اختبار تحصيلي محكي المرجع في مقرر الاحصاء الاعتماد على المنهج الوصفي التحليلي، لقدرته على الوصف الدقيق والتفصيلي للظاهرة بصورة نوعية أو كمية فالتعبير الكيفي يصف الظاهرة ويوضح خصائصها، أما التعبير الكمي فيعطينا وصفاً رقمياً يوضح مقدار هذه الظاهرة أو حجمها ودرجة ارتباطها مع الظواهر الأخرى المختلفة، وذلك من أجل الحصول على نتائج علمية، ثمَّ تفسيرها بطريقة موضوعية، بما ينسجم مع المعطيات الفعلية للظاهرة (دويدري، 2000، ص.183)

**2. مجتمع الدراسة وعينته**

**1.2. عينة الدراسة الأساسية**

يشتمل مجتمع الدراسة الحالي على طلاب السنة الأولى علوم اجتماعية جذع مشترك الذين درسوا مقرر الإحصاء الوصفي للموسم الجامعي (2021–2020) بجامعة زيان عاشور بالجلفة والبالغ عددهم (502) طالباً وطالبة، وبمأنَّ هذه الدراسة تحتاج إلى عينة كبيرة نسبياً للحصول على نتائج للبيانات يمكن الوثوق بها، فقد أعتمد الباحث المجتمع الأصلي كله للدراسة.

**2.2. عينة الدراسة الاستطلاعية**

من أجل استطلاع الظروف المحيطة بالدراسة، وتحديد أوجه القصور في إجراءاتها، تمَّ تطبيق الاختبار التحصيلي على عينة استطلاعية من خارج عينة الدراسة (خارج الجامعة) ومن نفس مجتمع الدراسة وعددها (67) طالب وطالبة.

**3. أداة الدراسة**

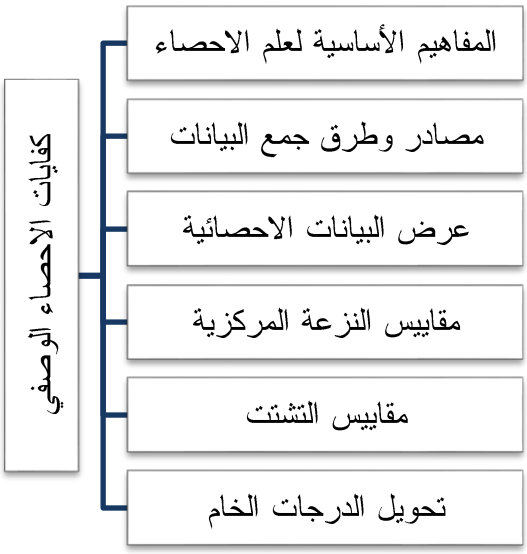
هذه الدراسة قائمة على فكرة بناء اختبار تحصيلي محكي المرجع في مقرر الإحصاء الوصفي لطلبة السنة الأولى علوم اجتماعية جذع مشترك، فقد اعتمد الباحث في بناء فقراته على الأهداف التعليمية، والسلوكية التي تتضمنها الموضوعات في مقرر الإحصاء الوصفي، وعليه فالأداة هي من نوع الاختيار من متعدد، وعند بناء هذا النوع من الاختبارات محكية المرجع فإنَّ هناك خطوات هامة ينبغي مراعاتها، لذلك تطلَّب من الباحث الرجوع إلى المصادر والمراجع المتخصصة في بناء الاختبارات محكية المرجع مثل علام(2012)، واليعقوبي(2013)، ومجيد( 2014) والدراسات السابقة مثل دراسة كل من طراونة(2016)، ربيع(2014)، (شكري،2006) (الجهني،2004) (العنزي،2003). وهذا لتحديد خطوات بناء أداة الدراسة.

**4. مراحل بناء أداة الدراسة**

**المرحلة الأولى: مرحلة التحليل وتشتمل على ثلاث خطوات**

**أ). تحديد النطاق السلوكي للمحتوى التعليمي**

يعتمد تحديد النطاق السلوكي للمحتوى المراد قياسه على طبيعة وحدود ذلك المحتوى، ويشير (علام، 1985) إلى أنَّه يُقصد بالنطاق السلوكي "مجموعة المعارف والمهارات التي يجب أن يتمكن منها الطالب والتي يقيسها الاختبار". وكون الدراسة الحالية تهدف إلى بناء اختبار محكي المرجع لقياس الكفايات في مقرر الإحصاء الوصفي، فقد قام الباحث بالاطلاع على محتوى البرنامج المقرر من الوزارة الوصية والخاص بطلبة السنة الأولى علوم اجتماعية جذع مشترك ملحق رقم (01) كما تمَّ التواصل بعدد من الأساتذة ممن يدِّرسون هذا المقرر للوقوف على مدى تطابق المقرر بما يدَّرس في الجامعات، وقد وجد الباحث أنَّ هناك اختلاف في بعض الجامعات من حيث الدروس المقررة للسنة الأولى علوم اجتماعية جذع مشترك. ومن خلال المقرر الوزاري لهذه الفئة من الطلبة يتضح أنَّ هناك ستة (06) كفايات يتكون منها النطاق السلوكي للاختبار كما في الشكل التالي:




**شكل رقم (16) كفايات الإحصاء الوصفي**

**ب). تحليل النطاق السلوكي (تحديد الكفايات الإحصائية)**

قام الباحث في هذه الخطوة بتحليل النطاق السلوكي تحليلاً اجرائياً إلى مكوناته من أهداف سلوكية مستخدماً أسلوب تحليل الإجراءات، حيث يتم ترتيب الأهداف في سلسلة خطية متتابعة من الأهداف المستقلة والمتكاملة تؤدي إلى تحقيق الكفاية المطلوبة(علام، 1995، ص.69) بعد ذلك تمَّ عرض الكفايات التي تمَّ التوصل إليها على عدد من المحكمين (أساتذة اختصاص قياس نفسي والتقويم التربوي وكذا أساتذة يدرسون مقياس الإحصاء) وهذا لإبداء الرأي في دقة الصياغة الإجرائية للأهداف السلوكية، ثمَّ قام الباحث بتحليل الكفايات الإحصائية إلى مكوناتها الفرعية على النحو التالي:

- كفاية المفاهيم الأساسية لعلم الإحصاء (التعريف بالإحصاء).
- كفاية مصادر وطرق جمع البيانات (مصادر وأساليب جمع البيانات).
- كفاية عرض البيانات الإحصائية (تصنيف البيانات وتبويبها، العرض البياني للبيانات).
- كفاية استعمال مقاييس النزعة المركزية (مفهوم النزعة، المنوال، الوسط الحسابي، الوسيط).
- كفاية استعمال مقاييس التشتت (المدى، المدى الربيعي، الانحراف المتوسط، الانحراف المعياري).
- كفاية تحويل الدرجات الخام (أنواع الدرجات المعيارية).

**ج). تحديد الأهداف الرئيسية لكل كفاية احصائية**

في هذه الخطوة وبعد أن تمَّ صياغة الأهداف الرئيسية للكفايات الإحصائية قام الباحث بصياغة

وترجمة هذه الأهداف العامة إلى أهداف سلوكية إجرائية، بحيث تمثل هذه الأهداف الموضوعات التي تتكون منها كل كفاية إحصائية، وقد تمَّ إعداد قائمة من الأهداف المصاغة صياغةً سلوكيةً للموضوعات التي يتضمنها مقرر الإحصاء الوصفي للسنة الأولى علوم اجتماعية جذع مشترك، وتمَّ عرضها على مجموعة من المحكمين للتأكد من مدى صحتها ووضوحها وشمولها وكذا تصنيف كل هدف سلوكي ضمن المستويات المعرفية لتصنيف بلوم، وتمَّ عرض استمارة التحكيم الملحق رقم(02) على مجموعة من المحكمين المختصين في القياس النفسي و التقويم التربوي و كذا الأساتذة الذين يدرسون مقرر الإحصاء الملحق رقم(09).

حيث تمَّ الأخذ بملاحظات المحكمين والتعديل على بعض الأهداف السلوكية التي بلغ عددها (54) هدفاً سلوكياً، موزعة وفق مستويات بلوم المعرفية كما يلي:

(17) هدفاً في مستوى التذكر (04) اهدافً في مستوى الفهم (16) هدفاً في مستوى التطبيق (12) هدفاً في مستوى التحليل (03) أهداف في مستوى التركيب (02) هدفان في مستوى التقويم.

والجدول التالي يبين عدد الأهداف السلوكية لكل كفاية إحصائية.

**جدول رقم (05) عدد الأهداف السلوكية لكل كفاية**

| الكفايات | عدد الأهداف السلوكية |
|---|---|
| المفاهيم الأساسية لعلم الإحصاء | 05 |
| مصادر وطرق جمع البيانات | 07 |
| عرض البيانات الإحصائية | 07 |
| استعمال مقاييس النزعة المركزية | 15 |
| استعمال مقاييس التشتت | 14 |
| تحويل الدرجات الخام | 06 |
| **المجموع** | 54 |

ويمكن تفصيل كل الأهداف السلوكية لكل كفاية إحصائية كما يا يلي:

**اولاً: كفاية المفاهيم الأساسية لعلم الإحصاء**

**جدول رقم (06) الأهداف السلوكية الإجرائية للكفاية الأولى**

| الكفاية | الدرس | الأهداف السلوكية |
|---|---|---|
| المفاهيم الأساسية لعلم الاحصاء | التعريف بالإحصاء | يعدد أهمية الاحصاء |
| | | يعطي تعريفاً للإحصاء |
| | | يحدد وظائف علم الاحصاء |
| | | يلخص علاقة الإحصاء بالعلوم الاجتماعية والانسانية |
| | | يفرق بين أنواع المتغيرات وطرق قياسها |

**ثانياً: كفاية مصادر وطرق جمع البيانات**

**جدول رقم (07) الأهداف السلوكية الإجرائية للكفاية الثانية**

| الكفاية | الدرس | الأهداف السلوكية |
|---|---|---|
| مصادر وطرق جمع البيانات | مصادر وأساليب جمع البيانات | يلم بمصادر جمع البيانات الإحصائية |
| | | يختار أسلوب جمع البيانات المناسب |
| | | يميز بين مصادر جمع البيانات |
| | | يتعرف على خصائص العينة |
| | | يميز بين أنواع العينات |
| | | يعدد أنواع العينات |
| | | يبين أخطاء المعاينة |

**ثالثاً: كفاية عرض البيانات الإحصائية**

**جدول رقم (08) الأهداف السلوكية الإجرائية للكفاية الثالثة**

| الكفاية | الدرس | الأهداف السلوكية |
|---|---|---|
| عرض البيانات الاحصائية | تصنيف البيانات وتبويبها | يشرح مفهوم تبويب البيانات |
| | | يعرف آليات التعامل مع البيانات المجمعة |
| | | يصمم جداول التوزيع التكراري |
| | | ينظم البيانات في جداول تكراري |
| | | يمثل بيانات الجداول التكرارية المختلفة بيانياً (أعمدة، دوائر، مدرج...) |
| | | يفرق بين أنواع الجداول التكرارية (البسيطة والمركبة) |
| | | يستخرج طول الفئة |

## رابعاً: كفاية استعمال مقاييس النزعة المركزية

**جدول رقم (09) الأهداف السلوكية الإجرائية للكفاية الرابعة**

| الكفاية | الدرس | الأهداف السلوكية |
|---|---|---|
| مقاييس النزعة المركزية | مفهوم النزعة | يعرِّف مفهوم النزعة المركزية |
| | | يقارن بين حالات استخدام مقاييس النزعة المركزية |
| | المنوال | يعرِّف المنوال |
| | | يحدد خصائص المنوال |
| | | يحسب المنوال للبيانات المبوبة وغير المبوبة |
| | | يحسب المنوال بالطريقة البيانية |
| | الوسط الحسابي | يعرِّف الوسط الحسابي |
| | | يستخرج خصائص الوسط الحسابي |
| | | يحسب الوسط الحسابي للبيانات المبوبة وغير المبوبة |
| | | يوظِّف المتوسط في الحالات المناسبة |
| | الوسيط | يعرِّف الوسيط |
| | | يستخرج خصائص الوسيط |
| | | يحسب الوسيط للبيانات المبوبة وغير المبوبة |
| | | يحسب الوسيط بالطريقة البيانية |
| | | يوظف العلاقة بين مقاييس النزعة المركزية |

## خامساً: كفاية استعمال مقاييس التشتت

**جدول رقم (10) الأهداف السلوكية الإجرائية للكفاية الخامسة**

| الكفاية | الدرس | الأهداف السلوكية |
|---|---|---|
| مقاييس التشتت | المدى | يعرِّف المدى |
| | | يحسب المدى المطلق للبيانات المبوبة وغير المبوبة |
| | | أن يحسب المدى النسبي |
| | | يعدد ايجابيات ومساوئ المدى |
| | المدى الربيعي | يعرِّف مفهوم المدى الربيعي ونصف المدى الربيعي |
| | | يحسب المدى الربيعي للبيانات المبوبة وغير المبوبة |
| | | يحسب نصف المدى الربيعي |
| | الانحراف المتوسط | يعرِّف الانحراف المتوسط |
| | | يحسب الانحراف المتوسط للبيانات المبوبة وغير المبوبة |
| | الانحراف المعياري | يعرِّف الانحراف المعياري |
| | | يحسب الانحراف المعياري للبيانات المبوبة وغير المبوبة |
| | | يعدد مزايا وعيوب الانحراف المعياري |
| | | يوضح ضرورة مقاييس التشتت |
| | | يفرق بين استخدامات مقاييس النزعة المركزية ومقاييس التشتت |

**سادساً: كفاية تحويل الدرجات الخام**

**جدول رقم (11) الأهداف السلوكية الإجرائية للكفاية السادسة**

| الكفاية | الدرس | الأهداف السلوكية |
|---|---|---|
| تحويل الدرجات المعيارية | أنواع الدرجات المعيارية | يعرِّف مفهوم الدرجة المعيارية |
| | | يحسب الدرجة المعيارية |
| | | يعرف مفهوم الدرجة التائية |
| | | يحسب الدرجة التائية |
| | | يعرف مفهوم المئينيات |
| | | يحسب المئينيات |

وقد قام الباحث بعرض هذه الأهداف الرئيسية ومضامينها من الأهداف السلوكية الإجرائية على مجموعة من المحكمين المتخصصين في القياس والتقويم وكذا من الأساتذة الذين يدرسون مقرر الإحصاء، وهذا لأخذ آرائهم في الأهداف السلوكية الإجرائية التي تمَّ اشتقاقها من الأهداف العامة الرئيسية وقد بلغ عدد المحكمين(17) محكماً من بعض الجامعات الجزائرية، لكن الباحث اعتمد على (13) محكماً فقط ممن أبدو الرغبة والاهتمام بالاختبار وقاموا بالتحكيم، ونتيجة للتوجيهات المقدمة من المحكمين تمَّ إجراء بعض التعديلات من حيث وضوح الهدف السلوكي الاجرائي.

ولمعرفة مدى ارتباط الهدف السلوكي الاجرائي بالهدف الرئيسي العام قام الباحث بإيجاد نسبة الاتفاق بين المحكمين باعتبار أن (مرتبط=1 وغير مرتبط=0) الملحق رقم (03). وقد اعتمد الباحث الحصول على النسبة (0.80) للإبقاء على الهدف السلوكي، كون أنَّ هذه النسبة هي الأمثل والتي اعتمدتها أغلب الدراسات وأنَّ المختصين في القياس يرون أنَّها النسبة المناسبة لاعتماد الأهداف السلوكية (البنَّاء،2011، ص.111). وقد توصَّل الباحث الى النسب الموضحة في الجداول التالية:

- **الكفاية الأولى: المفاهيم الأساسية لعلم الإحصاء**

**جدول رقم (12) نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الأولى**

| عدد الأهداف السلوكية | 01 | 02 | 03 | 04 | 05 |
|---|---|---|---|---|---|
| الاتفاق | 12 | 13 | 12 | 11 | 12 |
| نسبة الاتفاق | %92.30 | %100 | %92.30 | %84.61 | %92.30 |

من خلال الجدول السابق يتضح أنَّ أغلب الأهداف السلوكية للكفاية الأولى حصلت على نسب اتفاق عالية فاقت نسبة (84%) وهو مؤشـــر إيجابي يعكس ارتباط الأهداف السلوكيـــة بالهدف العام الذي

يقيس مدى تمكن الطلبة من المفاهيم الأساسية لعلم الإحصاء.

- **الكفاية الثانية: مصادر وطرق جمع البيانات**

**جدول رقم (13) نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الثانية**

| عدد الأهداف السلوكية | 01 | 02 | 03 | 04 | 05 | 06 | 07 |
|---|---|---|---|---|---|---|---|
| الاتفاق | 12 | 13 | 12 | 11 | 12 | 13 | 13 |
| نسبة الاتفاق | %92.30 | %100 | %92.30 | %84.61 | %92.30 | %100 | %100 |

من الجدول السابق يتضح أنَّ معظم الأهداف السلوكية حصلت على نسبة اتفاق فاقت 80% حول مدى ارتباطها بالهدف العام وهو طرق ومصادر جمع البيانات.

- **الكفاية الثالثة: عرض البيانات الإحصائية**

**جدول رقم (14) نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الثالثة**

| عدد الأهداف السلوكية | 01 | 02 | 03 | 04 | 05 | 06 | 07 |
|---|---|---|---|---|---|---|---|
| الاتفاق | 12 | 13 | 12 | 11 | 12 | 13 | 11 |
| نسبة الاتفاق | % 92.30 | % 100 | % 92.30 | % 84.61 | % 92.30 | % 100 | % 84.61 |

الملاحظ من خلال الجدول أنَّ نسبة الاتفاق في الكفاية الثالثة قد تجاوزت النسبة المقبولة وهي 80% وبالتالي فإنَّ الأهداف السلوكية مرتبطة بالهدف العام الذي يقيس طرق عرض البيانات الإحصائية.

- **الكفاية الرابعة: استعمال مقاييس النزعة المركزية**

**جدول رقم (15) نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الرابعة**

| الأهداف السلوكية | 01 | 02 | 03 | 04 | 05 | 06 | 07 |
|---|---|---|---|---|---|---|---|
| الاتفاق | 12 | 13 | 12 | 11 | 12 | 13 | 11 |
| نسبة الاتفاق | % 92.30 | % 100 | % 92.30 | % 84.61 | % 92.30 | % 100 | % 84.61 |
| الأهداف السلوكية | 08 | 09 | 10 | 11 | 12 | 13 | 14 |
| الاتفاق | 13 | 13 | 12 | 11 | 13 | 11 | 12 |
| نسبة الاتفاق | % 100 | % 100 | % 92.30 | % 84.61 | % 100 | % 84.61 | % 92.30 |
| الأهداف السلوكية | 15 | | | | | | |
| الاتفاق | 13 | | | | | | |
| نسبة الاتفاق | % 100 | | | | | | |

من خلال الجدول السابق يتضح أنَّ كل الأهداف السلوكية حصلت على نسبة تفوق 80% في مدى ارتباطها بالهدف العام وهو استعمال مقاييس النزعة المركزية.

- **الكفاية الخامسة: استعمال مقاييس التشتت**

**جدول رقم (16) نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية الخامسة**

| الأهداف السلوكية | 01 | 02 | 03 | 04 | 05 | 06 | 07 |
|---|---|---|---|---|---|---|---|
| الاتفاق | 12 | 13 | 12 | 11 | 12 | 13 | 13 |
| نسبة الاتفاق | %92.30 | % 100 | %92.30 | % 84.61 | %92.30 | % 100 | % 100 |
| الأهداف السلوكية | 08 | 09 | 10 | 11 | 12 | 13 | 14 |
| الاتفاق | 13 | 13 | 12 | 11 | 13 | 13 | 12 |
| نسبة الاتفاق | % 100 | % 100 | %92.30 | % 84.61 | % 100 | % 100 | %92.30 |

من خلال الجدول السابق يتضح أنَّ كل الأهداف السلوكية حصلت على نسبة تفوق 80% في مدى ارتباطها بالهدف العام وهو استعمال مقاييس التشتت.

- **الكفاية السادسة: تحويل الدرجات الخام**

**جدول رقم (17) نتائج نسبة الاتفاق بين المحكمين في الأهداف السلوكية للكفاية السادسة**

| الأهداف السلوكية | 01 | 02 | 03 | 04 | 05 | 06 |
|---|---|---|---|---|---|---|
| الاتفاق | 13 | 12 | 13 | 13 | 12 | 13 |
| نسبة الاتفاق | %100 | %92.30 | %100 | %100 | %92.30 | %100 |

من الجدول السابق يتضح أنَّ هناك بعض الأهداف السلوكية حصلت على الاتفاق التام بين المحكمين 100% والبعض الأخر فاق نسبة 80% مما يعكس ارتباطها بالهدف العام وهو تحويل الدرجات الخام. ويمكن تلخيص أهم التعديلات والاضافات التي تمت على الأهداف السلوكية من خلال أراء المحكمين كما يلي:

**جدول رقم (18) نتائج التحكيم للأهداف السلوكية**

| الكفايات | عدد الأهداف قبل التحكيم | عدد الأهداف بعد التحكيم | عدد الأهداف المعدلة |
|---|---|---|---|
| الأولى | 05 | 05 | 2 |
| الثانية | 07 | 07 | / |
| الثالثة | 09 | 07 | 1 |
| الرابعة | 14 | 15 | 2 |
| الخامسة | 12 | 14 | 1 |
| السادسة | 06 | 06 | / |
| المجموع | 53 | 54 | |

### 5.بناء جدول المواصفات

بعد تحديد الأهداف العامة لمقرر الإحصاء الوصفي للسنة الأولى جذع مشترك علوم اجتماعية والمؤشرات السلوكية الدالة عليها، قام الباحث بالاستعانة ببرنامج Excel ببناء جدول المواصفات ملحق رقم (18)، حيث تمَّ إعداد جدول ثنائي البعد يتضمن البعد الأفقي(الأهداف) ويتضمن البعد الرأسي (موضوعات المحتوي) ثمَّ تحديد الأوزان النسبية (الأهمية النسبية) لبعدي الجدول وفق الطريقة التالية:

- تحديد الوزن النسبي للأهداف وتمَّ وفق أراء المختصين والمحكمين.
- تحديد الوزن النسبي لموضوعات المحتوى وتمَّ في ضوء المحكات الآتية:

- رأي المختصين والمحكمين.
- عدد الحصص المقررة لتدريس مقرر الاحصاء الوصفي.

تمَّ حساب نسبة الاهمية للموضوع وفق الطريقة التالية:

$$\text{نسبة الأهمية للموضوع} = \frac{\text{عدد حصص الموضوع}}{\text{مجموع الحصص لكامل المقرر}} \times 100$$

تمَّ حساب نسبة الأهمية للأهداف وفق الطريقة التالية

$$\text{نسبة الأهمية للأهداف} = \frac{\text{عدد أهداف المستوى كاملة}}{\text{مجموع الأهداف الكامل المقرر}} \times 100$$

تمَّ حساب عدد أسئلة كل مستوى وفق الطريقة التالية

عدد أسئلة كل مستوى = عدد الأسئلة الكلي x نسبة الأهمية للموضوع x نسبة الأهمية للأهداف

**جدول رقم (19) يبين الوزن النسبي للموضوعات**

| الموضوعات | عدد الحصص | الوزن النسبي لكل موضوع |
|---|---|---|
| المفاهيم الأساسية لعلم الاحصاء | 2 | % 13 |
| مصادر وطرق جمع البيانات | 2 | % 13 |
| عرض البيانات الاحصائية | 3 | %18 |
| مقاييس النزعة المركزية | 4 | % 25 |
| مقاييس التشتت | 4 | % 25 |
| تحويل الدرجات الخام | 1 | % 06 |
| المجموع | 16 | % 100 |

**الجدول رقم (20) يبين الوزن النسبي للأهداف**

| مستوى الأهداف | عدد الأهداف | الوزن النسبي للأهداف |
|---|---|---|
| تذكر | 17 | %31 |
| فهم | 04 | %07 |
| تطبيق | 16 | %30 |
| تركيب | 03 | %06 |
| تحليل | 12 | %22 |
| تقويم | 02 | %04 |
| المجموع | 54 | % 100 |

وفي ضوء هذه المعطيات تمَّ تحديد عدد الأسئلة، الملحق رقم (18) يبين عدد الاسئلة لكل موضوع.

**المرحلة الثانية: وتشتمل على:**

**أ). بناء مفردات الاختبار**

يخضع بناء مفردات الاختبارات محكية لمرجع إلى قواعد وشروط محددة ينبغي إتباعها لذلك قام الباحث بالاطلاع على مجموعة من الدراسات السابقة ذات الصلة وكذا الأدبيات المتعلقة ببناء الاختبارات محكية المرجع، حتى يتم صياغة وبناء مفردات الاختبار وقد روعيَّ ما يلي:

- أن تمثل مفردات الاختبار محتوى ومضمون المقرر الإحصائي المقدم لطلبة السنة أولى علوم اجتماعية جذع مشترك.
- أن يكون مستوى صعوبة كل مفردة مناسباً لمستوى صعوبة الهدف الذي يقيسه.
- أن تبنى مفردات الاختبار على أساس مواصفات النطاق السلوكي الذي سبق إعداده.
- أن يحدد النطاق السلوكي الذي يقيسه الاختبار محكي المرجع حسب طريقة المواصفات السلوكية التي اقترحها بوفام (مجيد،2014، ص.180–183). كما أنَّها الطريقة الأكثر استخداماً في الدراسات السابقة والتي اهتمت ببناء هذا النوع من الاختبارات.
- أن يمثل الاختبار جميع المستويات المعرفية المتمثلة بمستويات التعلم كما حددها بلوم (التذكر، الفهم، التطبيق، التحليل، التركيب، التقويم).
- مراعاة التوازن في عدد مفردات الاختبار بين كافة محاور المقرر (المحاور الستة) وبما يتفق وحجم موضوعات الكفاية كما هي مقررة لطلبة السنة الأولى علوم اجتماعية جذع مشترك.

**ب). صياغة مفردات الاختبار.**

في بداية هذه الخطوة قام الباحث بتحديد العدد المناسب من المفردات التي تقيس الأهداف الممثلة في النطاق السلوكي، لذلك تمَّ اختيار (30) فقرة من نوع الاختيار من متعدد، تغطي الأهداف السلوكية للكفايات الستة للاختبار وبعد مراجعتها تمَّ اعتماد عشرين (20) فقرة لهذا الاختبار وهذا للأسباب التالية:

- الزمن المحدد لتطبيق الاختبار والمقدر بساعتين.
- أنَّ النقص في قيمة أخطاء التصنيف عند الانتقال من اختبار طوله (20) فقرة إلى اختبار طوله (30–40) فقرة يكون صغيراً نسبياَ لدرجة لا تبرر استخدام فقرات تزيد عن (20) فقرة (مجيد،2014، ص.191).
- الاختبار الطويل يؤدي إلى الملل ويقلل من دافعية المفحوص لإكمال الاختبار مما يفقد الاختبار أهم خاصية وهي الاختبار تحت ظروف أقصى أداء.
- الظروف التي شابت الموسم الدراسي 2020–2021 من انتشار لفيروس كوفيد19 حيث اعتمدت الوزارة الوصية التعليم عن بعد، وهو ما أثر على تحصيل الطلبة وعلى السير الحسن للدروس مما دعا بالباحث إلى مراعاة هذه الظروف والعمل بالحد الأدنى المقبول من الأسئلة.

وقد تمَّ تحديد الاختبار من نوع الأسئلة ذات الاختيار من متعدد بواقع أربع بدائل لكل سؤال ويبرر الباحث سبب اختياره لهذا النوع من الأسئلة للأسباب التالية:

- أغلب الدراسات التي تناولت هذا النوع من الأسئلة استخدمت الاختبارات رباعية البدائل.
- تمثل المحتوى المراد قياسه تمثيلاً جيداً.
- صادقة وثابتة بدرجة أكبر.
- يمكن استخدامها في قياس أهداف معرفية مختلفة مثل (التذكر، الفهم، التطبيق).
- يمكن تحليل نتائجه احصائياً بسهولة.
- تقلل من نسبة تخمين الجواب الصحيح حيث نسبة التخمين فيها يتوقف على عدد البدائل (25%) إذا كان عدد البدائل أربعة.

**ج). تحكيم مفردات الاختبار**

بعد صياغة فقرات الاختبار، تمَّ عرض الاختبار على مجموعة مكونة من(13) محكماً من ذوي الاختصاص والخبرة بالقياس والتقويم، واصحاب الكفاية في مجال الاحصاء، وذلك من أجل الحكم على مدى ارتباط كل فقرة بالهدف المحدد لها، وإبداء ملاحظاتهم حول وضوح الفقرة من حيث الصياغة اللغوية وملاءمة البدائل، حيث طلب من كل محكم أن يعطي رائه بوضع علامة(x) أمام الاجابة(نعم أولا) الملحق(05)، وبناءً على تقديرات المحكمين تمَّ إعادة النظر في بعض بدائل الفقرات، وتمَّ التعديل على صيغة البعض الأخر، وتبين أنَّ الاختبار ممثل للمحتوى المراد قياسه.

**المرحلة الثالثة: مرحلة التجريب:** وتتضمن عدة خطوات وهي:

**الخطوة الأولى: تطبيق الاختبار على عينة استطلاعية**

جُرب الاختبار المكون من عشرين (20) فقرة من أربع بدائل على عينة مكون من (67) طالباً وطالبة، لتحديد الزمن الذي يحتاجه المفحوصين للإجابة على بنود الاختبار وللتأكد من وضوح وسلامة الصياغة اللغوية للفقرات إضافة الى التأكد من الخصائص السيكومترية للاختبار وكذا للفقرات، وقد تمَّ تحديد الزمن اللازم للاختبار بساعتين (120 د) كما هو معمول به في الامتحانات الرسمية كما أنَّها كانت مدة كافية لينهي جميع المفحوصين الاختبار.

**الخطوة الثانية :تحليل مفردات الاختبار.**

على ضوء تطبيق الاختبار وتصحيحه يتم تقييم فعالية مفرداته، بهدف التعرف على:

- معامل صعوبة المفردة.
- معامل التمييز.
- معامل حساسية المفردات.
- فعالية كل بديل من بدائل المفردة.

**1.حساب مؤشر صعوبة الفقرة**

قام الباحث بهذه الخطوة للتأكد من مدى صعوبة مفردات الاختبار، حيث تمَّ تحليل كل فقرة من فقرات الاختبار لكل مفحوص من العينة الاستطلاعية، حيث يتم حساب عدد الطلبة الذين أجابوا إجابة صحيحة على كل فقرة من فقرات الاختبار مقسوماً على عددهم ليتم الحصول على مؤشر الصعوبة لكل فقرة من فقرات الاختبار. الملحق رقم (12) لذلك تمَّ ادخال البيانات في البرنامج (Spss).

والجدول التالي يبين درجة صعوبة كل فقرة من فقرات الاختبار.

**جدول رقم (21) مؤشرات الصعوبة لفقرات الاختبار للعينة الاستطلاعية**

| رقم الفقرة | معامل الصعوبة | رقم الفقرة | معامل الصعوبة |
|---|---|---|---|
| 01 | 0.66 | 11 | 0.64 |
| 02 | 0.67 | 12 | 0.58 |
| 03 | 0.73 | 13 | 0.61 |
| 04 | 0.52 | 14 | 0.75 |
| 05 | 0.60 | 15 | 0.55 |
| 06 | 0.72 | 16 | 0.63 |
| 07 | 0.60 | 17 | 0.55 |
| 08 | 0.60 | 18 | 0.69 |
| 09 | 0.63 | 19 | 0.69 |
| 10 | 0.58 | 20 | 0.52 |

يتبين من الجدول أنَّ متوسط معاملات الصعوبة لفقرات اختبار الإحصاء الوصفي هو (0.62) وأنَّ أقل معامل هو (0.52) للفقرتين 04 و20 وأعلى معامل هو (0.75) للفقرة 14، ويرى علماء القياس والتقويم انَّ المدى المقبول لمعامل الصعوبة هو الذي يتراوح بين (0.20 و0.80) (عودة، 2005). وبالتالي يرى الباحث أنَّ معاملات الصعوبة لجميع فقرات الاختبار مقبولة للدراسة.

**2.حساب معامل تمييز المفردة.**

تمَّ حساب معامل تميز المفردة من خلال ادخال البيانات لبرنامج (Spss) ملحق رقم (13). والجدول التالي يبين معامل تمييز فقرات الاختبار

**جدول رقم (22) يبين معاملات تمييز فقرات الاختبار**

| رقم الفقرة | معامل التمييز | رقم الفقرة | معامل التمييز |
|---|---|---|---|
| 01 | 0.74 | 11 | 0.44 |
| 02 | 0.56 | 12 | 0.46 |
| 03 | 0.44 | 13 | 0.58 |
| 04 | 0.40 | 14 | 0.75 |
| 05 | 0.41 | 15 | 0.68 |
| 06 | 0.54 | 16 | 0.45 |
| 07 | 0.44 | 17 | 0.57 |
| 08 | 0.47 | 18 | 0.64 |
| 09 | 0.48 | 19 | 0.43 |
| 10 | 0.63 | 20 | 0.44 |

يتضح من خلال الجدول أنَّ أغلب مفردات الاختبار ذات تمييز عالٍ فاقت (0.40) وهو الحد الذي يُعتبر فيه تمييز الفقرة ذات تمييز عالٍ وممتاز، باستثناء فقرة واحدة رقم 04 حيث تمييزها (0.40) وهو أضعف تمييز، وبالتالي فأنَّ المفردة تحتاج لتعديل وعليه يرى الباحث أنَّ جميع مفردات الاختبار ذات تمييز عالٍ ومقبول لإجراء الدراسة.

**3.حساب معامل حساسية البنود (معامل برينان) *Brennan***

تم حساب معامل حساسية البنود للاختبار محكي المرجع وفق القانون التالي:(علام،2001)

$$S = \frac{U}{n1} - \frac{L}{n2} \dots\dots\dots (01)$$

حيث أن:

$n1$: عدد الذين درجاتهم الكلية فوق درجات القطع.

$n2$: عدد الذين درجاتهم الكلية تحت درجات القطع.

$U$: عدد الذين اجابوا إجابة صحيحة عن الفقرة ممن درجاتهم الكلية فوق درجة القطع.

$L$: عدد الذين اجابوا إجابة صحيحة عن الفقرة ممن درجاتهم الكلية تحت درجة القطع.

**الجدول رقم (23) يبين النتائج المتعلقة بمعامل الحساسية**

| S | L/n2 | U/n1 | الفقرة | S | L/n2 | U/n1 | الفقرة |
|---|---|---|---|---|---|---|---|
| 0.52 | 0.29 | 0.80 | 11 | 0.75 | 0.14 | 0.89 | 01 |
| 0.50 | 0.24 | 0.74 | 12 | 0.49 | 0.33 | 0.83 | 02 |
| 0.61 | 0.19 | 0.80 | 13 | 0.44 | 0.43 | 0.87 | 03 |
| 0.60 | 0.33 | 0.93 | 14 | 0.48 | 0.19 | 0.67 | 04 |
| 0.73 | 0.05 | 0.78 | 15 | 0.52 | 0.24 | 0.76 | 05 |
| 0.50 | 0.29 | 0.78 | 16 | 0.56 | 0.33 | 0.89 | 06 |
| 0.67 | 0.10 | 0.76 | 17 | 0.52 | 0.24 | 0.76 | 07 |
| 0.58 | 0.29 | 0.87 | 18 | 0.52 | 0.24 | 0.76 | 08 |
| 0.45 | 0.38 | 0.83 | 19 | 0.57 | 0.24 | 0.80 | 09 |
| 0.48 | 0.19 | 0.67 | 20 | 0.64 | 0.14 | 0.78 | 10 |

نلاحظ من خلال الجدول رقم (23) أنَّ قيم معامل الحساسية تراوحت بين (0.44 إلى 0.75) وهي قيم مقبولة وتميز بين المفحوصين.

**4.فعالية كل بديل من بدائل المفردة**

يحسب معامل جاذبية المشتت (المموه) من خلال المعادلة التالية:

$$DA = \frac{Pu - Pi}{n} \dots\dots\dots (02)$$

حيث أن:

$DA$: جاذبية المشتت.

$Pu$: عدد الطلبة الذين اختاروا المشتت من الفئة العليا.

$Pi$ : عدد الطلبة الذين اختاروا المشتت من الفئة الدنيا.

n: عدد الطلبة في أحد الفئتين.

**والجدول رقم (24) يبين قيم فاعلية البدائل في الاختبار التحصيلي محكي المرجع**

| رقم المفردة | فاعلية البدائل | | | | رقم المفردة | فاعلية البدائل | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | | 1 | 2 | 3 | 4 |
| 01 | -0.03 | -0.13 | --- | -0.13 | 11 | -0.17 | -0.03 | --- | 0.03 |
| 02 | --- | -0.27 | -0.13 | -0.07 | 12 | 0.03 | -0.07 | -0.03 | --- |
| 03 | -0.07 | -0.20 | -0.17 | --- | 13 | -0.03 | --- | -0.07 | -0.03 |
| 04 | 0.10 | -0.17 | --- | -0.17 | 14 | --- | 0.00 | -0.13 | 0.10 |
| 05 | -0.10 | --- | -0.03 | 0.17 | 15 | -0.13 | 0.13 | --- | 0.07 |
| 06 | --- | -0.10 | -0.30 | 0.17 | 16 | --- | 0.03 | -0.10 | -0.20 |
| 07 | -0.20 | -0.07 | -0.03 | --- | 17 | 0.03 | 0.00 | -0.13 | --- |
| 08 | --- | 0.00 | -0.07 | -0.13 | 18 | -0.03 | --- | -0.13 | -0.07 |
| 09 | -0.13 | --- | -0.03 | 0.10 | 19 | --- | 0.03 | 0.03 | -0.10 |
| 10 | -0.10 | -0.27 | -0.13 | --- | 20 | -0.20 | -0.03 | --- | -0.07 |

من خلال الجدول يتضح أنَّ أغلبية المشتتات(المموهات) سالبة باستثناء بعض البدائل التي تحتاج إلى تعديل أو تبديل في الترتيب كما أنَّ أغلبية المشتتات حازت على اختيار الطلبة بنسبة (5%) فأكثر وعلية فإنَّ فاعلية المشتتات كانت اغلبها جيدة.

**الخطوة الثالثة :التحقق من صدق وثبات درجات الاختبار**

**اولاً: صدق الاختبار**

يشير (علام، 1995) إلى أنَّ إتباع الخطوات المناسبة في بناء الاختبارات محكية المرجع سيؤدي بالضرورة إلى اختبارات صادقة دون الحاجة إلى مزيد من التحقق. ولإضفاء نوع من المصداقية للاختبار، وقد تمَّ التحقق من صدق الاختبار بالطرق التالية:

**أ). الصدق الوصفي (المحتوى)**

قام الباحث بالتحقق من الصدق الوصفي، لتحديد مدى مطابقة مفردات النطاق السلوكي للأهداف، وذلك بقيام الباحث بعرض الفقرات الاختبارية التي تمَّ بناؤها وصياغتها على مجموعة من الخبراء والمحكمين عددهم (13) للتحقق من صلاحية الفقرات التي تمَّ إعدادها لبناء الاختبار التحصيلي

وكذلك الوضوح وسلامة الصياغة اللغوية الخاصة بفقرات الاختبار والتأكد من عدم وجود أخطاء في صياغة الفقرات.

وفيما يلي الخطوات المتبعة في تنفيذ الاجراءات:

- قام الباحث بإعداد استمارة تشتمل على ثلاث عناصر رئيسية هي (الأهداف، مفردات الاختبار ومقياس تدريجي به ملاحظات (متفق تماما=5) (متفق=4) (لا أدري=3) (غير متفق=2) (غير متفق تماماً=1) ملحق رقم (06).

- تم توزيع الاستمارة على المحكمين من أساتذة الجامعات المتخصصين في القياس النفسي والتقويم التربوي وممن يدرسون مقرر الاحصاء ثمَّ طلب منهم تحديد من وجهة نظرهم الاختيار المناسب لارتباط الهدف بالمفردة المعدة لقياسه.

- اعتمد الباحث الوسط الحسابي (3) كمعيار أساسي لتقييم الفقرة وذلك كنموذج خاص كما قام الباحث بإجراء بعض التعديلات على الفقرات الاختبارية لتكون أكثر ملاءمة من الناحية الفنية والعلمية وذلك بناءً على ملاحظات الخبراء والمحكمين، حيث أنَّ التغيرات شملت فقط بعض التعديلات والتغيرات على بعض الفقرات وبدائلها، بحيث تمَّ صياغتها بشكل أوضح وأنسب. مع الإبقاء على جميع الفقرات نظراً لاتفاق المحكمين على أهميتها ومن ثمَّ كتابتها بالصورة النهائية. ويبين الجدول رقم (25) متوسط تقدير آراء المحكمين حول فقرات الاختبار وانحرافاتها المعيارية وفقاً لمدى ملاءمة مواصفات الاختبار

- **جدول رقم (25) الأوساط الحسابية والانحرافات المعيارية لتقديرات المحكمين**

| رقم الفقرة | الوسط الحسابي | الانحراف المعياري | رقم الفقرة | الوسط الحسابي | الانحراف المعياري |
|---|---|---|---|---|---|
| 01 | 4.15 | 0.987 | 11 | 4.31 | 0.751 |
| 02 | 4.15 | 0.801 | 12 | 4.38 | 0.768 |
| 03 | 4.62 | 0.506 | 13 | 4.31 | 0.855 |
| 04 | 4.54 | 0.776 | 14 | 4.62 | 0.650 |
| 05 | 4.69 | 0.480 | 15 | 4.54 | 0.660 |
| 06 | 4.62 | 0.650 | 16 | 4.54 | 0.519 |
| 07 | 4.62 | 0.650 | 17 | 4.54 | 0.776 |
| 08 | 4.38 | 0.961 | 18 | 4.62 | 0.506 |
| 09 | 4.69 | 0.630 | 19 | 4.46 | 0.660 |
| 10 | 4.54 | 0.877 | 20 | 4.54 | 0.519 |

نلاحظ من خلال الجدول أنَّ متوسط تقديرات المحكمين تراوحت بين (4.15 و4.69) أي فاقت المتوسط الحسابي المقدر بـــ(3) وبالتالي فأنَّ المفردات تقيس النطاق السلوكي للأهداف حسب أراء المحكمين وقد تمَّ على أساس هذه الأوساط افتراض أنَّ فقرات لاختبار مصاغة بشكل جيد وأنَّ الاختبار صادق من حيث المحتوى.

**ب). الصدق الوظيفي**

تمَّ التحقق من الصدق الوظيفي للاختبار من خلال الصدق بدلالة محك المجموعات الطرفية في عينة الدراسة، حيث قسم الطلبة إلى مجموعتين مجموعة عليا فوق درجة القطع (12) ومجموعة دنيا تحت درجة القطع ثمَّ طبق عليهم اختبار(ت) لعينتين مستقلتين والجدول التالي يوضح النتائج.

**الجدول رقم (26) الصدق المحكي بدلالة المجموعة الطرفية**

| الفئة | العدد | المتوسط | الخطأ المعياري | درجة الحرية | قيمة (ت) | Sig | القرار |
|---|---|---|---|---|---|---|---|
| العليا | 21 | 17.57 | 0.926 | 23.80 | 18.625 | 0.00 | دال |
| الدنيا | 21 | 4.86 | 2.988 | | | | |

يلاحظ من خلال الجدول (26) وحسب Leven's test فإنَّ قيمة sig أقل من القيمة 5% وهي دالة وبالتالي شرط التباين غير محقق أي هناك فروقاً واضحة بين متوسطي علامات المجموعتين العليا والدنيا على الاختبار لصالح المجموعة العليا، حيث بلغت قيمة(T) المحسوبة (18.62) وهي قيمة دالة احصائياً عند مستوى الدلالة($\alpha \leq 0.01$) لأنَّ قيمة مستوى الدلالة المحسوب من العينة قد بلغ(0.00) وهذا دليل على أنَّ الاختبار يتمتع بصدق تمييزي جيد بين الطلبة المتمكنين وغير المتمكنين.

**ج). صدق البناء**

استخدم البحاث للدلالة على صدق البناء طريقة الاتساق الداخلي كمؤشر على هذا النوع من الصدق حيث قام بحساب درجة الترابط بين مفردات الاختبار التحصيلي والجدول التالي يبين درجة ترابط فقرات الاختبار مع بعضها البعض:

**جدول رقم (27) يبين درجات الاتساق الداخلي للاختبار**

| متوسط الارتباط بين الفقرات | الانحراف المعياري | القيمة الصغرى | القيمة العظمى |
|---|---|---|---|
| 0.310 | 5.66 | 0.051 | 0.826 |

من خلال الجدول يتضح أنَّ متوسط ارتباط الفقرات ببعضها البعض يساوي (0.31) بانحراف معياري قدره (5.66) في حين كانت أقل قيمة للارتبـــاط بين الفقرات هي (0.05) وأعلى قيمة (0.82) وهي قيم

مقبولة حسب (محمد حسين حبشي) وعليه يرى الباحث أنَّ قيمة متوسط ارتباط المفردات ببعضها البعض جيدة تعكس مدى ارتباط فقرات الاختبار ببعضها البعض، مما يدل على وجود اتساق داخلي بين فقرات الاختبار.

**ثانياً: ثبات الاختبار**

لحساب معامل ثبات الاختبار محكي المرجع طبق الباحث طريقتين هما:

أ). **معامل ثبات كيودر ريتشاردسون 20:** حيث تعتمد هذه المعادلة على حساب نسبة الأفراد الذين نجحوا في كل فقرة أو جزء من أجزاء الاختبار وعلى مدى تباين درجات هذه الفقرات والاجزاء ويحسب في حالة الفقرات ثنائية التصحيح (1،0) وفق المعادلة التالية:(Crocker & Algina,1986).

$$\mathrm{KR_{20}} = \frac{\mathrm{k}}{\mathrm{k}-1}\left(1 - \frac{\sum \mathrm{pq}}{\sigma_{\mathrm{x}}^2}\right) \dots\dots\dots (03)$$

حيث أنَّ:

$k$: عدد فقرات الاختبار.

$p$: نسبة الاجابات الصحيحة على الفقرة.

q: نسبة الاجابات الخاطئة على الفقرة.

$\sigma_x^2$: تباين الدرجة الكلية على الاختبار.

وبالتعويض نجد:

$$KR_{20} = \frac{20}{19}\left(1 - \frac{4.60}{28.80}\right) = 0.89$$

إذن قيمة معامل كيودر ريتشاردسون (0.89).

**ب). معادلة ليفنجستون *(Livingston Method)*:** بما أنَّ ثبات الاختبارات محكية المرجع التي تطبق مرة واحدة على الطلبة تعتمد بشكل كبير على الفرق بين متوسط درجات الطلبة المفحوصين ودرجة قطع اختبار التحصيل محكي المرجع، قام الباحث بتقدير قيمة معامل الثبات الاختبار من خلال معادلة ليفنجستون الخاصة بالاختبارات محكية المرجع التالية:

$$\mathrm{K^2(X,T)} = \frac{\delta_{\mathrm{x}}^2(\mathrm{KR_{20}}) + (\mu_{\mathrm{x}} - \mathrm{n_i c})^2}{\delta_{\mathrm{x}}^2 + (\mu_{\mathrm{x}} - \mathrm{n_i c})^2} \dots\dots\dots (04)$$

$\mathrm{K^2(X,T)}$: معامل ثبات ليفنجستون.

$KR_{20}$: معامل ثبات كيودر ريتشاردسون 20.

$\delta_x^2$: تباين درجات الاختبار الكلية.

$\mu_x$: متوسط درجات الاختبار الكلية.

c: درجة القطع.

$n_i$: عدد فقرا الاختبار.

نعوض في المعادلة رقم (01)

$$K^2(X,T) = \frac{32.01(0.89) + (12.51 - 20.12)^2}{32.01 + (12.51 - 20.12)^2} = 0.95$$

ومنه فإن قيمة ثبات الاختبار بطريقة ليفنجستون هي (0.95) وهي قيمة عالية جداً ومقبولة في الاختبارات التحصيلية محكية المرجع.

**الخطوة الرابعة: تحديد درجة القطع**

لتفسير أداء الطلبة على الاختبار محكي المرجع استخدم الباحث في هذه الدراسة طريقة أنجوف (Angoff's method) لتحديد درجة القطع ويبرر الباحث استخدام هذه الطريقة للأسباب التالية:

- عدم جدوى تطبيق الطرائق الاحصائية لعدم اخضاع عينة الدراسة لبرامج تعليمية.
- صعوبة تطبيق الطرائق التوافقية وكلفتها من ناحية الوقت والجهد.

إضافة إلى كون هذه الطريقة مستعملة في أغلب الدراسات، كما أنَّها أثبتت ارتفاع درجة القطع الناتجة مقارنة مع الطرق الأخرى، إضافة إلى انخفاض معدل الثبات الداخلي بين المحكمين كما أشارت لذلك كل من دراسة(علام،1991)،( Fehrmann et al., 1991)، (Chang, 1996) (Shen,2001) (Wang, 2003)، وقد ذكر(Hurtz & Hurtz,1999,p887 كما ورد في شكري، 2006) أنَّ من (9−11) مُحَكَّماً يكون كافياً للوصول إلى تقديرات موثوق بها. لذلك اعتمد الباحث على استشارة (13) محكماً من ذوي الخبرة في تدريس الإحصاء واختصاص القياس والتقويم التربوي.

وفيما يلي الاجراءات التي اتبعها الباحث في تحديد درجة القطع بطريقة أنجوف.

- تم أعداد استمارة لتقدير أنجوف ملحق رقم (15).
- اختيار عينة من المحكمين من ذوي الخبرة في مجال تدريس الإحصاء ومن تخصص القياس النفسي

والتقويم التربوي وممن يدرس مقرر الإحصاء والبلغ عددهم (11) محكماً فقط ممن أبدوا الرغبة من (13) محكماً وقاموا بالتحكيم.

- تقديم كل من المفردات الاختبارية واستمارة تقدير أنجوف إلى المحكمين لإعطاء التقدير المناسب لكل مفردة من وجهة نظرهم.
- حساب متوسط تقديرات المحكمين لكل مفردة، ثمَّ حساب متوسط تقديرات مفردات الاختبار ككل ملحق رقم (16).

والجدول رقم (28) يوضح النتائج المتحصل عليها من متوسط تقديرات المحكمين للمفردات بطريقة أنجوف.

**جدول رقم (28) يبين متوسط تقديرات المحكمين لمفردات الاختبار باستخدام طريقة أنجوف**

| رقم المفردة | متوسط تقدير المحكمين | رقم المفردة | متوسط تقدير المحكمين |
|---|---|---|---|
| 01 | 69.09 | 11 | 67.27 |
| 02 | 65.91 | 12 | 45.00 |
| 03 | 73.64 | 13 | 45.45 |
| 04 | 59.09 | 14 | 57.27 |
| 05 | 55.45 | 15 | 40.91 |
| 06 | 62.27 | 16 | 73.18 |
| 07 | 55.45 | 17 | 41.36 |
| 08 | 54.55 | 18 | 66.82 |
| 09 | 56.36 | 19 | 46.82 |
| 10 | 68.64 | 20 | 47.27 |
| المجموع | | | 1151.82 |

وبقسمة مجموع متوسطات تقديرات المحكمين لمفردات الاختبار على العدد الكلي لهذه الفقرات نتجت درجة القطع كما يلي:

$$\frac{1151.82}{20} = 57.59\ \%$$

وبتحويل النسبة المئوية ينتج عدد المفردات 11.52 وبالتقريب يصبح (12) وهي درجة القطع.

- **أساليب المعالجة الاحصائية**

للإجابة على أسئلة الدراسة استعمل الباحث المعالجات الاحصائية التالية:

- حساب معاملات الصعوبة والتمييز لكل فقرة من فقرات الاختبار من خلال برنامج(Spss26) .
- حساب قيمة معامل ثبات كيودر ريتشاردسون20(KR-20).
- حساب قيمة معامل الثبات ليفنجستون(Livingston Index) .

- تحديد درجة القطع باستخدام طريقة أنجوف(Angoff's Method) .
- أجراء اختبار (T-test) للكشف عن الصدق الوظيفي.
- استخدام برنامج(Winsteps) للإجابة على أسئلة الدراسة.

# الفصل الخامس: تحليل ومناقشة نتائج الدراسة

**تمهيد**

**أولاً: تحليل نتائج تساؤلات الدراسة**

**ثانياً: مناقشة نتائج تساؤلات الدراسة**

**الخاتمة**

**التوصيات**

**تمهيـــد**

يتضمن هذا الفصل عرضاً مفصلاً للنتائج التي توصلت إليها الدراسة باستخدام الأساليب الإحصائية المناسبة بهدف الاجابة على تساؤلات الدراسة وتفسيرها ومناقشتها.

**I). عـــرض وتحليـــل نتائـــج التســـاؤل الأول والـــذي ينص:**

- **ما درجة تحقيق بيانات اختبار مقرر الإحصاء لافتراضات نموذج راش لدى طلبة السنة أولى علوم اجتماعية جذع مشترك؟**

من أجل اختبار صحة هذه التساؤلات، قام الباحث بتفحص الافتراضات الرئيسة التي يتطلبها نموذج راش، وهي:

- أحادية البعد
- الدالة المميزة لكل مفردة والتي تصف العلاقة الوتيرية بين القدرة والأداء على المفردة، والتي تسمى اختصاراً " دالة خصائص المفردة "
- الاستقلال الموضعي
- عامل التخمين
- التحرر من عامل السرعة.

**أولاً: التحقق من افتراضات نموذج راش أحادي البارامتر.**

**1. افتراض أحادية البعد**

قام الباحث باختبار هذا الافتراض بطريقة تحليل نموذج(راش) للمكونات الأساسية المعتمدة على البواقي (PCAR).

**1-2. طريقة تحليل نموذج (راش) للمكونات الأساسية المعتمد على البواقي**

تهدف نظرية الاستجابة للمفردة إلى التعرف على هل أنَّ الانحرافات عن السمة المقاسة ترقى إلى أن تكون عامل مستقل أم لا؟ لذلك يوفر برنامج(Winsteps) تحليل نموذج (راش) للمكونات الأساسية المعتمدة على البواقي(PCAR) والذي يُظهر الاختلافات بين الأبعاد والجدول رقم (29) يلخص أهم مؤشرات نتائج هذا التحليل:

**الجدول (29) نتائج التحليل العاملي للمكونات الأساسية للبواقي (PCAR) باستخدام نموذج (راش) قبل التدريج.**

| حجم تباين البواقي المعيارية مقدرا بوحدات القيم Eigen. | الجذر الكامن | الملاحظ | المتوقع |
|---|---|---|---|
| التباين الكلي في الاستجابات | 37.7151 | %100 | %100 |
| التباين الذي فسره العامل الرئيسي، تقديرات نموذج (راش) | 17.7151 | %47 | %47 |
| التباين المفسر بواسطة الأفراد | 9.0763 | %24.1 | %24.1 |
| التباين المفسر بواسطة المفردات | 8.6389 | %22.9 | %22.9 |
| مجموع التباين غير المفسر | 20.0000 | %53 | %53 |
| التباين الذي فسره العامل الثاني (الأول في البواقي) | 2.1270 | %5.6 | %10.6 |

يُظهر الجدول رقم (29) مؤشرات التحليل العاملي بالمكونات الأساسية للبواقي باستخدام نموذج (راش) وللحكم على أحادية البعد تمَّ التحقق من بعض المحكات، وهي:(Linacre, 2008, p.272).

أ. **المحك الأول:** كقاعدة عامة إذا كانت قيمة التباين المفسر بواسطة التقديرات ( Raw variance explained by measures) أكبر من أو تساوي (50%)، في حين أكد لينـاكر مطور (Winsteps) في موضع أخر حيث حدد المجال من (20%) إلى (80%) بأنَّه مجال مقبول ويعد مؤشراً قوياً لأحادية البعد.

وهو ما تحقق في نتائج الدراسة الحالية حيث أشار مؤشر التباين الذي فسره العامل الرئيسي لتقديرات نموذج(راش) للقيمة الملاحظة (47%) وهي قريبة جداً من القيمة (50%) وفي المجال الذي اقترحه لينـاكر، كما أنَّ القيمة الملاحظة هي نفس القيمة المتوقعة للنموذج.

**ب. المحك الثاني:** والذي يُحدَّد بقيمة الجذر الكامن لنسبة التباين الذي يفسره العامل الثاني (Unexplained Variance in 1st contrast)، بأن لا تتعدى أو تكون أقل من القيمة (3)، وهذا ما تحقق في الدراسة الحالية بتسجيلنا للقيمة (2.1270) وهي أقل من القيمة (3) المحددة.

**ج. المحك الثالث:** والذي يحدد قيمة نسبة التباين الذي يفسره العامل الثاني (الأول في البواقي) (Unexplained Variance in 1st contrast) بأقل أو قريب من القيمة (5%) وهذا ما تحقق في الدراسة الحالية بالقيمة (5.6%) وهذا دليل آخر على أحادية البعد.

يتضح من خلال المؤشرات الإحصائية في الجدول رقم (29)، بأنَّ محكات التحليل المستخدمة تحققت كلياً، وهو ما يؤكد على أنَّ اختبار مقياس الإحصاء، له تكوين فرضي ذو فضاء كامن أحادي البعد.

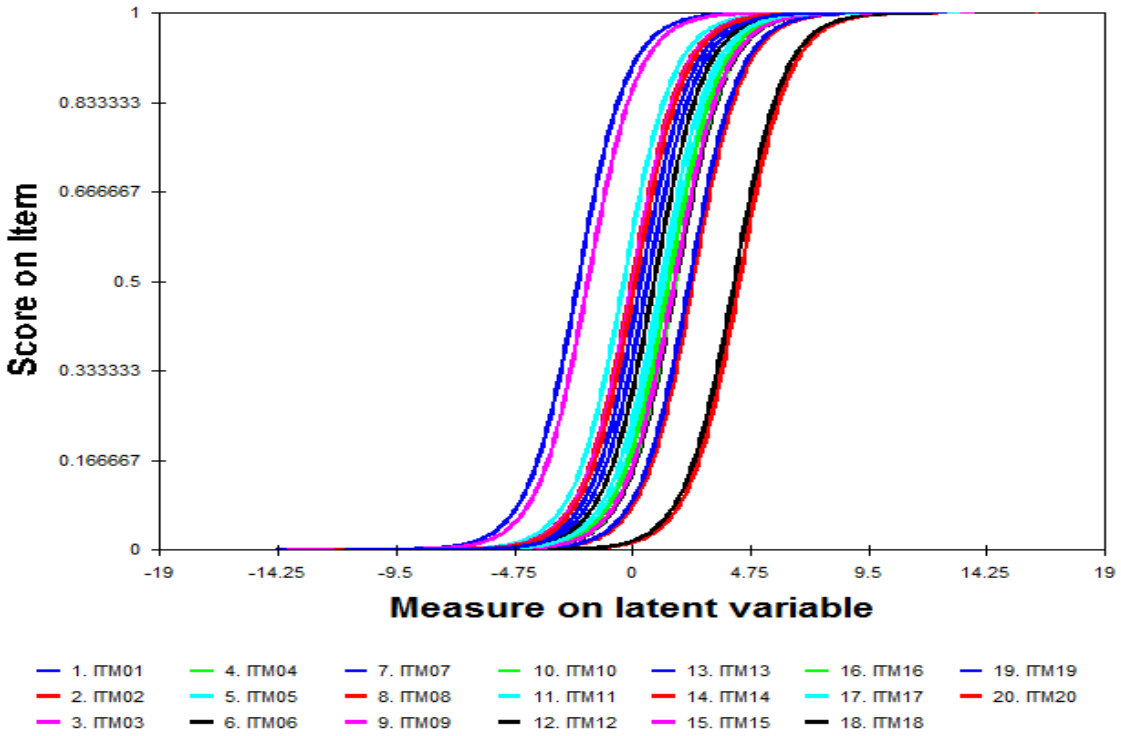
**ثانيا: التحقق من افتراض دالة استجابة المفردة**

استخدم الباحث البرنامج الإحصائي (Winsteps) الذي يوفر تقدير معلم التمييز، والجدول رقم (30) يوضح نتائج مؤشرات تمييز مفردات اختبار الاحصاء.

**الجدول (30) يظهر قيم مؤشرات التمييز لاختبار الإحصاء.**

| رقم المفردة | مؤشر التمييز | رقم المفردة | مؤشر التمييز |
|---|---|---|---|
| 1 | 0.98 | 11 | 1.02 |
| 2 | 1.10 | 12 | 0.97 |
| 3 | 0.99 | 13 | 0.97 |
| 4 | 1.00 | 14 | 1.04 |
| 5 | 0.96 | 15 | 1.02 |
| 6 | 1.08 | 16 | 1.12 |
| 7 | 1.00 | 17 | 0.96 |
| 8 | 0.95 | 18 | 1.12 |
| 9 | 1.06 | 19 | 1.02 |
| 10 | 0.79 | 20 | 0.90 |

يتضح من الجدول رقم (30) أنَ مؤشرات تمييز مفردات اختبار الإحصاء، كلها قيم أعلى من القيمة (0.50) والتي تراوحت ما بين (0.79-1.12) وحسب كلٍ من (لينكر، ماسترز) فإنَّ قيم التمييز المقبولة يجب أن تكون ما بين (0.5، 2) ويمكن أن تكون من (0.5) إلى ما لانهاية ( Linacre, 2012, p.596; Masters, 1988 ,pp.15–29) ويرى لينكر أنَّ متوسط قيم تمييز الميل (Slope) المفردات في نموذج (راش) يجب أن يكون قريب من(1.0) تقريباً. (Linacre, 2012, p.597) وهذا ما أشارت إليه القيم في الدراسة الحالية حيث أنَّ متوسط قيم مؤشرات التمييز كان (1.0025) وهي قيمة قريبة جداً من الواحد، وتظهر مؤشرات تمييز المفردات على شكل منحنيات بيانية توضح احتمال التوصل إلى الاستجابة الصحيحة على المفردة كدالة للسمة الكامنة في ضوء الأداء على مفردات الاختبار وأنَّه في تطبيقات نظرية الاستجابة للمفردة يفترض أنَّ منحنيات خصائص المفردة لها شكل حرف(S)، وهو ما يوضحه الشكل رقم (17)، لخصائص مفردات الاختبار المستخرجة بواسطة برنامج (Winsteps).




**الشكل (17) منحنى خصائص مفردات اختبار مقياس الإحصاء20 مفردة.**

يتضح من الشكل رقم(17) أنَّ الدرجة تتزايد بزيادة السمة الكامنة، وبالتالي تزداد احتمالية التوصل إلى الاستجابة الصحيحة، ومن المؤشرات التي تدل كذلك على تحقق تكافؤ مؤشرات تمييز المفردات ومطابقتها للنموذج، حسب ما أشار إليه هامبلتون وسوامينثان(Hambleton & Swaminathan, 1985) هو أن تكون قيمها واقعة ضمن حدود المدى لمتوسط معاملات التمييز يساوي(0.15) وبما أنَّنا وجدنا قيمة متوسط معاملات التمييز(0.57) وبالتالي فهي في حدود المدى لقيم معامل الارتباط الثنائي المتسلسل ($\rho_{pbis}$) ما بين القيمة(0.41) للمفردة رقم(20) والقيمة(0.64) للمفردة رقم(16).

كما أسفرت نتائج التحليل الموضحة في الجدول رقم (31) لقيم معاملات الارتباط الثنائي المتسلسل ($\rho_{pbis}$)، والتي كانت جميعها ضمن المدى المقبول، وتوزيعها إلى حدٍ ما متجانس، مما يدل على تساوي مؤشرات تمييز مفردات الاختبار، مع تسجيل قيم صغيرة للانحراف المعياري لهذه المؤشرات وهو مؤشر آخر على تحقق هذا الافتراض.

**الجدول (31) يظهر قيم مؤشرات التمييز بمعامل الارتباط الثنائي المتسلسل ($\rho_{pbis}$) لاختبار الإحصاء.**

| رقم المفردة | مؤشر التمييز | رقم المفردة | مؤشر التمييز |
|---|---|---|---|
| 1 | 0.60 | 11 | 0.61 |
| 2 | 0.61 | 12 | 0.45 |
| 3 | 0.58 | 13 | 0.58 |
| 4 | 0.60 | 14 | 0.59 |
| 5 | 0.60 | 15 | 0.47 |
| 6 | 0.63 | 16 | 0.64 |
| 7 | 0.95 | 17 | 0.55 |
| 8 | 0.57 | 18 | 0.63 |
| 9 | 0.62 | 19 | 0.57 |
| 10 | 0.56 | 20 | 0.41 |

**ثالثاً: التحقق من افتراض الاستقلال الموضعي**

للتحقق من هذا الافتراض، استخدم البرنامج الإحصائي (Winsteps) في استخراج المؤشر الاحصائي(Q3)، والنتائج يوضحها الجدول رقم (32):

**الجدول (32) يوضح قيم المؤشر الإحصائي(Q3) من خلال قيم ارتباط البواقي المعيارية.**

| أزواج المفردات | الارتباط | أزواج المفردات | الارتباط |
|---|---|---|---|
| المفردة 03 مع المفردة 16 | 0.16- | المفردة 09 مع المفردة 11 | 0.13- |
| المفردة 14 مع المفردة 16 | 0.15- | المفردة 07 مع المفردة 08 | 0.13- |
| المفردة 08 مع المفردة 18 | 0.15- | المفردة 09 مع المفردة 15 | 0.13- |
| المفردة 05 مع المفردة 07 | 0.14- | المفردة 09 مع المفردة 17 | 0.12- |
| المفردة 09 مع المفردة 14 | 0.14- | المفردة 03 مع المفردة 06 | 0.12- |
| المفردة 16 مع المفردة 11 | 0.14- | المفردة 07 مع المفردة 09 | 0.12- |
| المفردة 16 مع المفردة 1 | 0.14- | المفردة 16 مع المفردة 18 | 0.12- |
| المفردة 03 مع المفردة 09 | 0.13- | المفردة 09 مع المفردة 18 | 0.12- |
| المفردة 07 مع المفردة 16 | 0.13- | المفردة 06 مع المفردة 16 | 0.12- |
| المفردة 09 مع المفردة 06 | 0.13- | …. | .... |

يتضح من الجدول رقم(32) أنَّ قيم مؤشر الاحصائي(Q3) تراوحت ما بين(-0.12 و-0.16) بمتوسط حسابي يساوي(-0.13) وبالتالي فإنَّ هذه القيم لم تتجاوز مستوى الانتهاك(0.30) حسب زينسكي وهامبلتون، وسيريسي (Zenisky , Hambleton & Sireci, 2002) وكذلك المجال الذي اقترحه كوهن وتيسن(Cohen & Thissen, 1997) وهو أن تكون القيمة ما بين(0.20 و-0.20) ومنه يمكننا الحكم على أنَّ مفردات اختبار الإحصاء لا توجد فيها مفردات تتأثر الإجابة عنها بمفردات أخرى من مفردات نفس الاختبار، مما يعني عدم تداخل منحنيات الخصائص المميزة للمفردات كذلك، وهو ما يشار إليه بالاستقلال الموضعي(Linacre,2012, p.405).

**رابعاً: التحقق من افتراض تدني عامل التخمين**

تمَّ استخدام البرنامج الإحصائي (Winsteps)، والذي يوفر تقدير معلمة التخمين للمفردات وتظهر قيم معلمة التخمين تحت العمود "التخمين الأدنى"، والجدول رقم (33) يوضح نتائج تقدير مؤشرات التخمين لمفردات اختبار الإحصاء.

**الجدول (33) يوضح مؤشرات التخمين لمفردات اختبار الإحصاء.**

| رقم المفردة | مؤشر التخمين | | رقم المفردة | مؤشر التخمين | |
|---|---|---|---|---|---|
| | الأعلى | الأدنى | | الأعلى | الأدنى |
| 1 | 1.00 | 0.01 | 11 | 1.00 | 0.00 |
| 2 | 1.00 | 0.00 | 12 | 1.00 | 0.00 |
| 3 | 1.00 | 0.00 | 13 | 1.00 | 0.01 |
| 4 | 1.00 | 0.00 | 14 | 1.00 | 0.00 |
| 5 | 1.00 | 0.03 | 15 | 1.00 | 0.00 |
| 6 | 1.00 | 0.00 | 16 | 1.00 | 0.00 |
| 7 | 1.00 | 0.03 | 17 | 0.99 | 0.00 |
| 8 | 0.99 | 0.00 | 18 | 1.00 | 0.00 |
| 9 | 1.00 | 0.00 | 19 | 1.00 | 0.00 |
| 10 | 0.96 | 0.03 | 20 | 0.97 | 0.01 |

يتضح من الجدول رقم (33)، أنَّ جميع قيم مؤشرات التخمين الأدنى لمفردات اختبار الإحصاء أقل من القيمة(0.10) وقريبة جداً من الصفر وفي المجال المقبول لقيم التخمين الذي حدده كل من ( Linacre, 2012, p.112; Steven, 2003, pp.164-184; Baker, 2001, p.28).

**خامساً: التحقق من التحرر من عامل السرعة**

تمَّ التأكد عملياً من هذا الافتراض من خلال تطبيق الباحث للاختبار، وأنَّ فشل الطلبة في الوصول إلى الإجابة الصحيحة عن مفردات الاختبار يعود فقط لانخفاض قدراتهم، ولا يعود إلى ضيق وقت الاختبار أو لعوامل السرعة في الاستجابة، وذلك بملاحظة الإجابات المتروكة والتي توزعت عشوائياً على طول مفردات الاختبار ولم تتمركز في منطقة معينة، مما يشير إلى أنَّ زمن الاختبار لم يؤثر على استجابات الطلبة، وأنَّ مدة واحد ساعة كانت كافية للانتهاء من الإجابة على جميع مفردات الاختبار بالنسبة لأغلبية الطلبة، كما أنَّه لم يُظهر أي طالب أي شكوى أثناء التطبيق من حجم الوقت وعدم كفايته، فإنَّ هذا يعني أنَّ الاختبار موضع البحث يقيس القوة ولا يقيس السرعة.

**2- عـــرض وتحليـــل نتائـــج التســـاؤل الثانـــي والـــذي ينـــص علـــى:**

**- ما قيم تقديرات معلم الصعوبة لفقرات الاختبار بعد التدريج وفق نموذج راش؟**

لتقدير معلم الصعوبة لفقرات اختبار مقياس الإحصاء، قام الباحث بتحليل استجابات الطلبة اعتماداً على برنامج(Winsteps) الذي يستخدم طريقة تقدير الأرجحية القصوى المشتركة(JMLE) في تقدير بارامترات النموذج، حيث يقوم البرنامج بعملية " التحرير" لتدريج أو معايرة الاختبار بعد حذف الاستجابات التامة، وهي حذف استجابات الأفراد الذين أجابوا إجابة صحيحة على جميع مفردات الاختبار، وكذلك حذف الأفراد الذين لم يُجيبوا على أي مفردة إجابةً صحيحةً، حيث حذف خمس(5) إجابات صحيحة تامة، ونفس العدد(5) إجابتهم كانت كلها خاطئة على فقرات الاختبار أي حذف عشرة(10) طلبة، في حين لم يَحذف أي من فقرات الاختبار والتي تباينت إجابات الطلبة عليها ولم نسجل أي فقرة أجاب عليها الطلبة إجابة صحيحة تامة، أو لم يجب عليها أي طالب لتدخل فقرات اختبار مقياس الإحصاء جميعها في تحليل استجابات أفراد العينة(492) طالباً وطالبة المتبقية، حيث تمَّ تحليل استجابات الأفراد عن جميع مفردات اختبار مقياس الإحصاء الوصفي بهدف الكشف عن الأفراد غير المطابقين للنموذج، ولتحقيق ذلك تمَّ تقدير قدرة كل فرد والخطأ المعياري في مقياس القدرة، وقيم إحصائي المطابقة الداخلية(Infit ZSTD)، أو إحصائي متوسط المربعات التقاربي(Infit MNSQ)، وهو مؤشر إحصائي للسلوكيات غير المتوقعة التي تؤثر في الاستجابات عن الفقرات التي تكون قريبة من مستوى قدرة الفرد. كما تمَّ تقدير قيم إحصائي المطابقة الخارجية (Outfit ZSTD) أو إحصائي متوسط

المربعات التباعدي (Outfit MNSQ) وهو مؤشر إحصائي أكثر حساسية للسلوكيات غير المتوقعة التي تؤثر في الاستجابات عن الفقرات التي تكون بعيدة عن مستوى قدرة الفرد، وذلك لكل تقدير من تقديرات القدرة، والجدول رقم (34) يوضح ملخص لنتائج تحليل قدرات الأفراد.

**الجدول (34) ملخص نتائج التحليل لقدرات الأفراد.**

| الإحصاءات | الدرجة الخام Raw Score | القدرة Measure | الخطأ المعياري للنموذج Model S.E | إحصائي المطابقة الداخلية Infit | | إحصائي المطابقة الخارجية Outfit | |
|---|---|---|---|---|---|---|---|
| | | | | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD |
| المتوسط الحسابي | 10.7 | 0.26 | 0.66 | 1.00 | 0.00 | 0.95 | 0.1 |
| الانحراف المعياري | 5.2 | 1.86 | 0.15 | 0.30 | 0.8 | 0.78 | 0.6 |
| أعلى علامة | 19 | 4.04 | 1.13 | 2.06 | 2.2 | 9.90 | 3.8 |
| أدنى علامة | 1 | 3.98- | 0.53 | 0.38 | 1.9- | 0.13 | 1.0- |

يتضح من الجدول رقم (34) أنَّ قيم المتوسط الحسابي لمتوسطات المربعات الداخلية والخارجية (MNSQ) يساوي الواحد بالنسبة لمتوسطات المربعات الداخلية، وقريب جداً من الواحد بالنسبة لمتوسطات المربعات الخارجية، وأنَّ المتوسط الحسابي لمتوسط إحصائي المطابقة الداخلية والخارجية (ZSTD) تساوي الصفر، وهو الوضع المثالي والذي يفترضه النموذج، كما أنَّ الانحراف المعياري قريب من الواحد.

وعند تفحص قيَّم إحصائي المطابقة الخارجية والداخلية للأفراد، التي تشير إلى مطابقة قدرة الفرد مع قدرات مجموعة الأفراد التي ينتمي إليها في قياس السمة التي يقيسها الاختبار، فإذا كانت قيمة هذان الإحصائيان تزيد على(+2) تعتبر قدرة الفرد غير مطابقة لقدرات مجموعة الأفراد (Alastair and Hutchinson,1987 ; julian,1988) حيث تبين أنَّ(08) أفراد غير مطابقين للنموذج تزيد قيَّم إحصائي المطابقة الخارجية والداخلية لدرجاتهم على القيمة(+2) أو قيَّم متوسطات المربعات المقابلة لدرجاتهم تزيد على(1) وهي القيم التي يتوقعها النموذج(Wright and Stone, 1979).

بعد استبعاد الأفراد غير المطابقين للنموذج، تمت إعادة التحليل للكشف عن المفردات غير المطابقة للنموذج، حيث تمَّ تقدير معلم الصعوبة لكل مفردة، والخطأ المعياري في قياس معلم الصعوبة

وقيَّم إحصائي المطابقة الداخلية للمفردات(Infit ZSTD)، والذي يعبر عنه بإحصائي متوسطات المربعات للمفردات(Infit MNSQ)، كما تمَّ تقدير قيَّم إحصائي المطابقة الخارجية للمفردات( Outfit ZSTD) والذي يعبر عنه بإحصائي متوسطات المربعات للمفردات(Outfit MNSQ) لكل معلم من معالم الصعوبة، والجدول رقم(35) يوضح ملخص نتائج التحليل لمعالم صعوبة مفردات اختبار الإحصاء.

**الجدول (35) ملخص نتائج التحليل لمعالم الصعوبة لمفردات اختبار الإحصاء**

| الإحصاءات | الدرجة الخام Raw Score | الصعوبة Measure | الخطأ المعياري للنموذج Model S.E | إحصائى المطابقة الداخلية Infit | | إحصائى المطابقة الخارجية Outfit | |
|---|---|---|---|---|---|---|---|
| | | | | متوسط لمربعات MNSQ | قيمة الإحصائي ZSTD | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD |
| المتوسط الحسابى | 259.4 | 0.00 | 0.13 | 1.00 | 0.1 | 0.95 | 0.2- |
| الانحراف المعياري | 102.8 | 1.66 | 0.02 | 0.05 | 0.8 | 0.12 | 0.8 |
| أعلى علامة | 443 | 3.43 | 0.19 | 1.14 | 2.4 | 1.21 | 1.7 |
| أدنى علامة | 53 | 3.29- | 0.12 | 0.93 | 1.2- | 0.65 | 1.8- |

يتضح من الجدول رقم (35) أنَّ المتوسط الحسابي لمتوسطات المربعات الداخلية والخارجية (MNSQ) يساوي الواحد بالنسبة لمتوسطات المربعات الداخلية، وقريب جداً من الواحد بالنسبة متوسطات المربعات الخارجية، كما أنَّ المتوسط الحسابي لمتوسط إحصائي المطابقة الداخلية والخارجية (ZSTD) قريب من الصفر، فالوضع المثالي الذي يفترضه النموذج هو (0،1) كما أنَّ الانحراف المعياري قريب من الواحد. وفيما يتعلق بالمفردات، تبين عدم مطابقة مفردة واحدة لنموذج راش، حيث زادت قيم متوسطات المربعات الموزونة لها الواحد الصحيح، والجدول رقم (36) يوضح قيم احصائيات المطابقة الداخلية والخارجية ومتوسطات المربعات الموزونة.

**الجدول (36) يوضح المفردات غير المطابقة، وقيم إحصائيات المطابقة الداخلية والخارجية، ومتوسطات المربعات الموزونة لهذه المفردات في التدريج الأولي لاختبار الإحصاء.**

| المفردة Item | الدرجة الخام Raw Score | إرتباط بونت بيسريال PTBIS-CORR | الصعوبة Measure | الخطأ المعياري للنموذج Model S.E. | إحصائى المطابقة الداخلية Infit | | إحصائى المطابقة الخارجية Outfit | |
|---|---|---|---|---|---|---|---|---|
| | | | | | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD |
| 01 | 294 | 0.59 | 0.44- | 0.12 | 1.03 | 0.5 | 1.03 | 0.3 |
| 02 | 328 | 0.61 | 0.94- | 0.12 | 0.94 | 0.9- | 0.84 | 0.9- |
| 03 | 332 | 0.57 | 1.00- | 0.12 | 1.02 | 0.3 | 0.91 | 0.5- |
| 04 | 291 | 0.60 | 0.40- | 0.12 | 0.99 | 0.1- | 1.01 | 0.1 |
| 05 | 250 | 0.60 | 0.16 | 0.12 | 1.02 | 0.3 | 1.00 | 0.1 |
| 06 | 273 | 0.62 | 0.15- | 0.12 | 0.95 | 0.8- | 0.62 | 0.61 |
| 07 | 310 | 0.59 | 0.67- | 0.12 | 1.01 | 0.2 | 1.00 | 0.1 |
| 08 | 336 | 0.56 | 1.06- | 0.12 | 1.02 | 0.4 | 1.12 | 0.7 |
| 09 | 212 | 0.62 | 0.68 | 0.12 | 0.97 | 0.5- | 0.88 | 0.9- |
| 10 | 227 | 0.55 | 0.47 | 0.12 | 1.14 | 2.4 | 1.21 | 1.7 |
| 11 | 237 | 0.62 | 0.34 | 0.12 | 0.99 | 0.2- | 1.03 | 0.3 |
| 12 | 66 | 0.42 | 3.09 | 0.15 | 1.08 | 0.9 | 0.92 | 0.2- |
| 13 | 170 | 0.57 | 1.26 | 0.12 | 1.04 | 0.7 | 1.04 | 0.3 |
| 14 | 159 | 0.58 | 1.42 | 0.12 | 0.97 | 0.5- | 1.00 | 0.00 |
| 15 | 432 | 0.43 | 2.95- | 0.17 | 0.97 | 0.2- | 0.93 | 0.1- |
| 16 | 211 | 0.64 | 0.69 | 0.12 | 0.93 | 1.2- | 0.78 | 1.8- |
| 17 | 357 | 0.54 | 1.39- | 0.13 | 1.06 | 0.9 | 1.02 | 0.2 |
| 18 | 207 | 0.63 | 0.75 | 0.12 | 0.93 | 1.2- | 0.84 | 1.3- |
| 19 | 443 | 0.39 | 3.29- | 0.19 | 0.99 | 0.00 | 0.94 | 0.00 |
| 20 | 53 | 0.41 | 3.43 | 0.17 | 1.02 | 0.3 | 0.65 | 1.0- |
| Mean | 259.4 | | 0.00 | 0.13 | 1.00 | 0.1 | 0.95 | 0.2- |
| S.D | 100.2 | | 1.61 | 0.02 | 0.05 | 0.8 | 0.12 | 0.7 |

بعد استبعاد الأفراد غير المطابقين، والمفردات غير المطابقة، ولغرض التأكد من مطابقة البيانات الناتجة عن استجابات أفراد العينة للاختبار والتحقق من موضوعية الاختبار بصورته النهائية (19) مفردة، تمت إعادة التحليل للحصول على تقديرات متحررة من صعوبة المفردات وقدرات الأفراد والجدول رقم (37) يبين ملخص نتائج التحليل للقيم المتحررة من قدرات الأفراد بعد استبعاد المفردات والأفراد غير المطابقين لنموذج راش.

**الجدول (37) ملخص نتائج التحليل للقيم المتحررة من قدرات الأفراد بعد استبعاد الأفراد والمفردات غير المطابقة.**

| الإحصاءات | الدرجة الخام Raw Score | القدرة Measure | الخطأ المعياري للنموذج Model S.E | إحصائي المطابقة الداخلية Infit | | إحصائي المطابقة الخارجية Outfit | |
|---|---|---|---|---|---|---|---|
| | | | | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD |
| المتوسط الحسابي | 10.2 | 0.28 | 0.68 | 1.00 | 0.00 | 0.95 | 0.1 |
| الانحراف المعياري | 4.9 | 1.90 | 0.15 | 0.33 | 0.9 | 0.87 | 0.7 |
| أعلى علامة | 18 | 4.06 | 1.14 | 2.17 | 2.4 | 9.90 | 3.7 |
| أدنى علامة | 1 | 3.96- | 0.55 | 0.37 | 2.1- | 0.13 | 1.1- |

يتضح من الجدول رقم(37) أنَّ التقديرات النهائية المتحررة من قدرات الأفراد تراوحت ما بين الدرجتين(1 و18) وبمتوسط توزيع القدرة بلغ(0.28) لوجيت، وبانحراف معياري بلغ(1.9) لوجيت أمَّا الخطأ المعياري للمتوسط الحسابي لتقديرات القدرة فقد بلغ(0.68) وهي قيمة تقترب مما يفترضه النموذج، كما يتضح كذلك أنَّ متوسط توزيع القدرة كانت قيمته(0.26) لوجيت قبل حذف المفردة رقم(10) القيمة(0.17) لوجيت، وبانحراف معياري(1.86) أما الخطأ المعياري للمتوسط الحسابي لتقديرات القدرة بلغ(0.66) لمَّا كان متوسط صعوبة المفردات يساوي صفر، كما هو مبين في الجدولين رقم(34) ورقم(35) وبالتالي فإنَّ متوسط قدرة الأفراد ارتفع من القيمة(0.26) قبل الحذف إلى القيمة(0.28) بعد حذف المفردات والأفراد غير المطابقين لنموذج راش، وبالتالي فهذه القيمة أعلى من صعوبة المفردات أي أنَّ مستوى الاختبار بشكل عام في متناول مستوى أفراد العينة مما يقلل من عوامل التخمين كما تؤكد ذلك(كاظم، 1988ب). مما يشير إلى دقة تحديد مــواقع الأفراد على متصل

القدرة، وهو ما يوضحه الجدول رقم (38).

**الجدول (38) يوضح ملخص نتائج التحليل المتحرر من صعوبة المفردات بعد التدريج.**

| الإحصاءات | الدرجة الخام Raw Score | الصعوبة Measure | الخطأ المعياري للنموذج Model S.E | إحصائى المطابقة الداخلية Infit | | إحصائى المطابقة الخارجية Outfit | |
|---|---|---|---|---|---|---|---|
| | | | | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD |
| المتوسط الحسابي | 261.1 | 0.00 | 0.13 | 1.01 | 0.1 | 0.96 | 0.2- |
| الانحراف المعياري | 105.3 | 1.72 | 0.02 | 0.05 | 0.7 | 0.12 | 0.7 |
| أعلى علامة | 443 | 3.50 | 0.19 | 1.09 | 1.0 | 1.13 | 0.7 |
| أدنى علامة | 53 | 3.30- | 0.12 | 0.93 | 1.1- | 0.64 | 1.7- |

يشير الجدول رقم(38) إلى أنَّ متوسط توزيع القدرة بلغ(صفراً) لوجيت، والانحراف المعياري بلغ(1.72) لوجيت، وقيم التقديرات المتحررة من صعوبة المفردات تراوحت ما بين(0.64 و1.13) أمَّا الخطأ المعياري للمتوسط الحسابي لتقديرات الصعوبة فقد بلغ(0.13) وهي قيمة متدنية نوعاً ما، مما يشير إلى دقة تقديرات الصعوبة، فالمتوسط الحسابي لمتوسطات المربعات الداخلية والخارجية (MNSQ) يساوي الواحد بالنسبة لمتوسطات المربعات الداخلية، وقريب جداً من الواحد بالنسبة متوسطات المربعات الخارجية(0.96) بينما المتوسط الحسابي لمتوسط إحصائي المطابقة الداخلية والخارجية(ZSTD) تقريباً يساوي الصفر، فالوضع المثالي الذي يفترضه النموذج هو(0،1) كما أنَّ الانحراف المعياري قريب من الواحد، وقد تمَّ تقدير قيَّم معالم مفردات اختبار مقياس الإحصاء بصورته النهائية(19) مفردة، باستخدام طريقة الأرجحية العظمى غير المشروطة (UCON) لتقدير أدق القدرات والصعوبة، ولتقليل الخطأ في التقدير لصعوبة المفردات. والجدول رقم (39) يبين صعوبة المفردات والخطأ المعياري للقياس في تقدير الصعوبة مرتبة تنازلياً بحسب صعوبتها بعد التدريج النهائي.

**الجدول (39) يوضح صعوبة المفردات وأخطاؤها المعيارية مرتبة تنازلياً وفق صعوبة المفردات.**

| إحصائي المطابقة الخارجية Outfit | | إحصائي المطابقة الداخلية Infit | | الخطأ المعياري للنموذج Model S.E | الصعوبة Measure | ارتباط بونت بيسريال PTBIS-CORR | الدرجة الخام Raw Score | المفردة Item |
|---|---|---|---|---|---|---|---|---|
| قيمة الإحصائي ZSTD | متوسط المربعات MNSQ | قيمة الإحصائي ZSTD | متوسط المربعات MNSQ | | | | | |
| 1.0- | 0.64 | 0.5 | 1.05 | 0.17 | 3.50 | 0.41 | 53 | 20 |
| 0.2 | 1.04 | 1.0 | 1.09 | 0.16 | 3.16 | 0.42 | 66 | 12 |
| 0.00 | 1.00 | 0.4- | 0.97 | 0.12 | 1.47 | 0.59 | 159 | 14 |
| 0.4 | 1.06 | 0.8 | 1.05 | 0.12 | 1.31 | 0.58 | 170 | 13 |
| 0.8- | 0.89 | 0.4- | 0.98 | 0.12 | 0.71 | 0.63 | 212 | 09 |
| 1.7- | 0.78 | 1.0- | 0.94 | 0.12 | 0.72 | 0.64 | 211 | 16 |
| 1.2- | 0.85 | 1.1- | 0.93 | 0.12 | 0.78 | 0.64 | 207 | 18 |
| 0.4 | 1.04 | 0.1- | 0.99 | 0.12 | 0.37 | 0.62 | 237 | 11 |
| 0.00 | 0.99 | 0.3 | 1.02 | 0.12 | 0.19 | 0.61 | 250 | 05 |
| 0.4- | 0.95 | 0.7- | 0.96 | 0.12 | 0.13- | 0.63 | 273 | 06 |
| 0.3 | 1.03 | 0.5 | 1.03 | 0.12 | 0.43- | 0.59 | 294 | 01 |
| 0.1 | 1.01 | 0.00 | 1.00 | 0.12 | 0.38- | 0.61 | 291 | 04 |
| 0.2 | 1.03 | 0.4 | 1.02 | 0.12 | 0.66- | 0.59 | 310 | 07 |
| 1.0- | 0.83 | 0.9- | 0.95 | 0.12 | 0.93- | 0.61 | 328 | 02 |
| 0.2- | 0.95 | 0.8 | 1.05 | 0.12 | 0.99- | 0.57 | 332 | 03 |
| 0.7 | 1.13 | 0.6 | 1.04 | 0.12 | 1.05- | 0.56 | 336 | 08 |
| 0.3 | 1.04 | 0.9 | 1.06 | 0.13 | 1.39- | 0.54 | 357 | 17 |
| 0.1 | 0.98 | 0.1- | 0.99 | 0.17 | 2.96- | 0.42 | 432 | 15 |
| 0.00 | 0.95 | 0.1- | 0.98 | 0.19 | 3.30- | 0.40 | 443 | 19 |
| 0.2- | 0.96 | 0.1 | 1.01 | 0.13 | 0.00 | | 261.1 | Mean |
| 0.6 | 0.11 | 0.7 | 0.04 | 0.02 | 1.67 | | 102.5 | S.D |

يتضح من الجدول رقم (39) أنَّ مؤشرات المطابقة لكل مفردة من مفردات اختبار الإحصاء الوصفي كانت ضمن حدود المطابقة الداخلية(التقاربية) والخارجية(التباعدية) والشكل رقم (18) يوضح قيَّم كل

من صعوبة المفردات ومجال قيَّم الإحصائي لمتوسط المربعات التباعدي للمفردات (Outfit MNSQ) وقيَّم الإحصائي متوسط المربعات التقاربي للمفردات (Infit MNSQ).

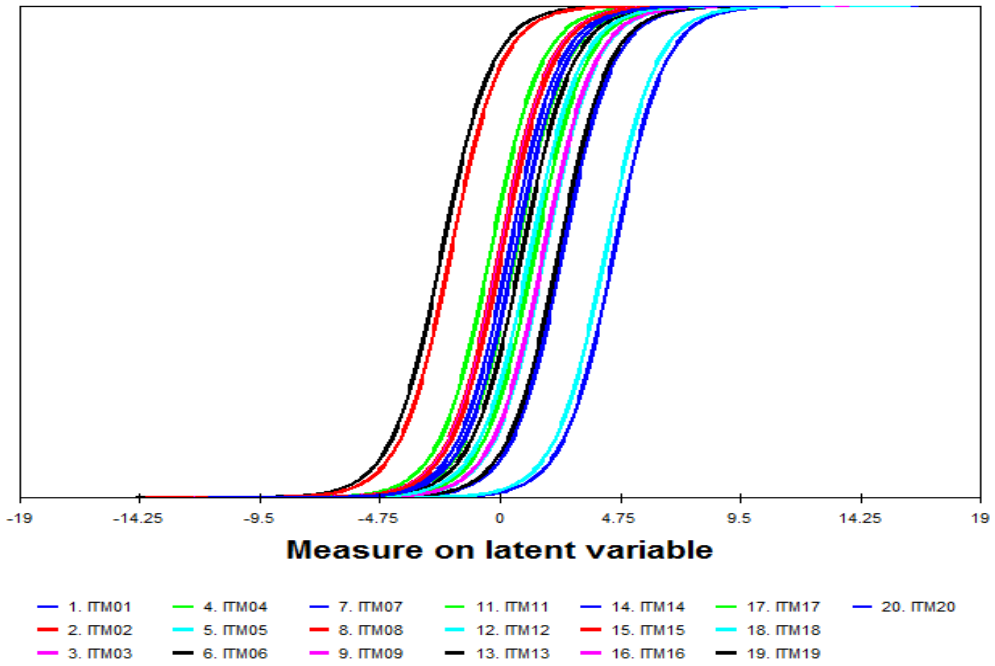
ITEM FIT GRAPH: MISFIT ORDER

| ENTRY NUMBER | MEASURE - + | INFIT MEAN-SQUARE 0.0 1 2 | OUTFIT MEAN-SQUARE 0.0 1 2 | ITEM |
|---|---|---|---|---|
| 8 | * | : * | : .* | ITM08 |
| 12 | *| | : * | : * | ITM12 |
| 13 | * | : * | : * | ITM13 |
| 17 | * | : * | : * | ITM17 |
| 3 | * | : * | : *. | ITM03 |
| 20 | *| | : * | : * . | ITM20 |
| 11 | * | : *. | : * | ITM11 |
| 1 | * | : * | : * | ITM01 |
| 7 | * | : * | : * | ITM07 |
| 5 | * | : * | : *. | ITM05 |
| 4 | * | : *. | : * | ITM04 |
| 14 | * | : *. | : *. | ITM14 |
| 15 |* | : *. | : *. | ITM15 |
| 9 | * | : *. | : * . | ITM09 |
| 19 |* | : *. | : *. | ITM19 |
| 6 | * | : *. | : *. | ITM06 |
| 2 | * | : *. | : * . | ITM02 |
| 16 | * | : *. | : * . | ITM16 |
| 18 | * | : *. | : * . | ITM18 |

**الشكل (18) يوضح مؤشرات المطابقة للمفردات وحدود المطابقة ومواقع للمفردات من هذه الحدود وفقاً لإحصائي متوسط المربعات التباعــدي والتقاربي (MNSQ).**

يتضح من الشكل رقم(18) أنَّ مؤشرات المطابقة للمفردات وحدود المطابقة ومواقع المفردات من هذه الحدود وفقاً لإحصائي متوسط المربعات التباعدي و التقاربي(MNSQ) تشير إلى أنَّ صعوبة مفردات اختبار مقياس الإحصاء في صورته النهائية(19) مفردة، والمعبر عنها في العمود(Measure) بأنَّها في المجال المقبول للصعوبة، كما تشير قيَّم مؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية والتقاربية لمفردات الاختبار إلى أنَّها قريبة من الوضع المثالي الذي يفترضه النموذج وهو القيمة(1) والذي تراوحت قيمه ما بين(0.64 و1.13) بمتوسط حسابي(0.96) وانحراف معياري(0.12) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية (الخارجية)، وتراوحت ما بين(0.93 و1.09) وبمتوسط حسابي(1.01) وانحراف معياري(0.05) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية(الداخلية) وهي قيَّم في مجال القياس المنتج الذي حدده كلٍ من( Wright and Linacre, 1994) ما بين(0.70 و1.30) أو ما بين(0.50 و1.5) عموماً والقريبة من الواحد، وهو

مؤشر إحصائي أكثر حساسية للسلوكيات غير المتوقعة من النموذج(المتطرفة مثل التخمين و الأخطاء غير المتوقعة) التي تؤثر في الاستجابات عن المفردات التي تكون بعيدة عن مستوى قدرة الفرد. لذلك فالقيمة الأكبر من (1.0) لمؤشر (Outfit–Underfit) أو مؤشر (2–/ 0.7> MNSQ) تعني وجود تباين كبير أو تداخل في البيانات غير مبرر. والقيمة أقل من (1.0) لمؤشر (Infit-Overfit) أو مؤشر (2/ 1.3> MNSQ) يعني أنَّ النموذج مفرط في البيانات وقد يتسبب في تضخم إحصاءات الثبات (William, et al, 2014, p.166). ويدعم هذه النتائج الشكل رقم (19) الذي يظهر منحنيات خصائص مفردات اختبار الإحصاء بعد التدريج النهائي.




**الشكل (19) منحنى خصائص مفردات اختبار الإحصاء بعد التدريج باستخدام نموذج راش.**

يلاحظ من الشكل رقم (19) بأنَّ المنحنيات المميزة لمفردات اختبار الإحصاء بعد تدريجه باستخدام نموذج(راش) أحادي البارامتر لا تتقاطع، ولكن تختلف عن بعضها في الصعوبة فقط وتتلاقى فقط عند المحور الأفقي الذي يمثل متصل القدرة أو السمة، وهو ما تؤكد عليه خريطة (فرد– بند) لرايت في الشكل رقم (20).

```
INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  19 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------

MEASURE    PERSON - MAP - ITEM
            <more>||<freq>
    5             .  ++
                     ||
                     ||
                     ||
                     ||
                     ||
    4         .#### T++
                     ||T ITM19
                     ||
                     ||  ITM15
                     ||
                     ||
    3     #########  ++
                     ||
                     ||
                     ||
          .########  ||
                    S||S
    2                ++  ITM17
           .#######  ||
                     ||  ITM08
           .#######  ||  ITM02  ITM03
                     ||
        .##########  ||  ITM07
    1                ++  ITM01  ITM04
                     ||
           .#######  ||  ITM06
            .######  ||M
                    M||  ITM05
             .#####  ||  ITM11
    0                ++
             ######  ||  ITM09  ITM16  ITM18
                     ||
             #######  ||
                     ||  ITM13
             #######  ||  ITM14
   -1                ++
             .######  ||S
                     ||
              .#####  ||
                    S||
          #########  ||
   -2                ++
                     ||
            .#####  ||
                     ||
                     ||  ITM12
                     ||T
   -3          .###  ++  ITM20
                     ||
                     ||
                    T||
                     ||
                     ||
   -4           ###  ++
               <less>||<rare
```

**الشكل (20) يوضح خريطة رايت (Wright) لتدريج الأفراد ومفردات اختبار الإحصاء.**

تشير خريطة رايت في الشكل رقم (20) وبصورة جلية إلى وجود تناسب تام بين توزيع تقدير مفردات اختبار الإحصاء من الجانب الأيسر للخريطة، وتدريج المفردات من الجانب الأيمن.

وبالرغم من انتظام التدريج لصعوبة المفردات بصورة واضحة من المدى (-3.30 و3.50) لوجيت فقد تبدو بعض الفراغات الضيقة عند بعض المستويات الأخرى من المتغير والتي تظهرها خريطة (رايت) والتي يمكن التحقق منها من خلال العلاقة بين مسافة الفراغ بين صعوبة المفردتين ومجموع الخطأ المعياري لهما، والتي يجب أن يكون الفرق بين صعوبتي المفردتين المحددتين للفراغ أقل من مجموع الخطأ المعياري للقياس لكلاهما والفراغات هي:

الفراغ الأول: الفرق بين صعوبتي المفردتين(14و12) يساوي (1.69) ومجموع الخطأ المعياري للقياس (0.28) أي أنَّ الخطأ المعياري للقياس لا يغطي الفراغ.

الفراغ الثاني: الفرق بين صعوبتي المفردتين(13و18) يساوي (0.53) ومجموع الخطأ المعياري للقياس (0.24) أي أنَّ الخطأ المعياري للقياس لا يغطي الفراغ.

الفراغ الثالث: الفرق بين صعوبتي المفردتين(15و17) يساوي (1.57) ومجموع الخطأ المعياري للقياس (0.30) أي أنَّ الخطأ المعياري للقياس لا يغطي الفراغ.

الفراغ الرابع: الفرق بين صعوبتي المفردتين(15و19) يساوي (0.34) ومجموع الخطأ المعياري للقياس (0.36) أي أنَّ الخطأ المعياري للقياس يغطي الفراغ.

وهكذا فإنَّ الخطأ المعياري للقياس لصعوبة كل مفردتين مكونتين للفراغ، لا يغطي مسافة الفراغات الثلاثة الأول بينهما، ولكنَّه يغطي مسافة الفراغ الرابع، وبالرغم من أنَّ هناك انتظام في تدرج المفردات على مدى ميزان القياس، بما يعني تعريف مفردات الاختبار للمتغير الذي يُمثَّل بمدى هذا المتصل وهو القدرة على الإحصاء، وهو ممَّا يدل على جودة تدريج مفردات الاختبار.

إلا أنَّ هذا يدعونا إلى ضرورة إضافة مفردات أخرى للاختبار المطور مستقبلاً لتغطية الفراغات المسجلة على مستوى متصل القدرة.

**3. عـرض وتحليـل نتائـج التسـاؤل الثـالث والـذي ينـص علـى:**

**- ما كمية المعلومات التي يقدمها اختبار الاحصاء الوصفي عند مستويات القدرة المختلفة؟**

لتقدير كمية المعلومات التي يقدمها اختبار الإحصاء عند مستويات القدرة المختلفة، والعلاقة بين الدرجة الخام والقدرة بصورته النهائية، باستخدام نموذج راش أحادي البارامتر، وتقدير دالة المعلومات المقابلة لكل درجة خام وكل قدرة، وتقدير الخطـأ المعياري والقدرة، والجدول رقم (40)

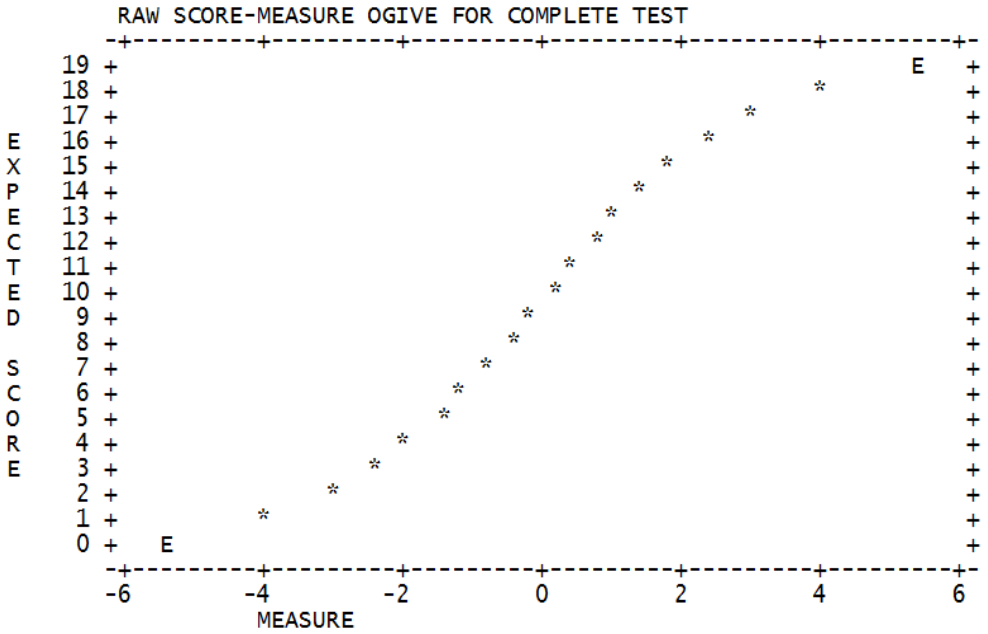
يوضح هذه المؤشرات.

**الجدول (40) يتضمن تقديرات قدرة الفرد المقابلة لكل درجة كلية محتملة على الاختبار، وانحرافاتها المعيارية وكذا كمية المعلومات التي يقدمها اختبار الإحصاء عند كل مستويات القدرة المختلفة، والعلاقة بين الدرجة الخام والقدرة بوحدة اللوجيت.**

| الدرجة الخام | تكرار الأفراد | التكرار التراكمي للأفراد | القيم المتحررة لقدرات الأفراد بوحدة اللوجيت | الخطأ المعياري | دالة المعلومات |
|---|---|---|---|---|---|
| 1 | 12 | 12 | 3.97- | 1.12 | 0.80 |
| 2 | 14 | 26 | 3.03- | 0.86 | 1.35 |
| 3 | 23 | 49 | 2.39- | 0.74 | 1.82 |
| 4 | 36 | 85 | 1.90- | 0.67 | 2.24 |
| 5 | 23 | 108 | 1.49- | 0.62 | 2.61 |
| 6 | 25 | 133 | 1.13- | 0.59 | 2.90 |
| 7 | 28 | 161 | 0.80- | 0.57 | 3.12 |
| 8 | 28 | 189 | 0.48- | 0.55 | 3.26 |
| 9 | 24 | 213 | 0.18- | 0.55 | 3.32 |
| 10 | 23 | 236 | 0.12 | 0.55 | 3.31 |
| 11 | 26 | 262 | 0.43 | 0.56 | 3.22 |
| 12 | 30 | 292 | 0.75 | 0.57 | 3.05 |
| 13 | 41 | 333 | 1.09 | 0.60 | 2.82 |
| 14 | 31 | 364 | 1.46 | 0.63 | 2.51 |
| 15 | 31 | 395 | 1.89 | 0.68 | 2.14 |
| 16 | 34 | 429 | 2.41 | 0.76 | 1.72 |
| 17 | 36 | 465 | 3.08 | 0.88 | 1.28 |
| 18 | 18 | 483 | 4.07 | 1.14 | 0.77 |
| 19 | 1 | 484 | 5.48 | 1.90 | 0.28 |

يوضح الجدول رقم (40) تقديرات قدرات الأفراد، وكذا الأخطاء المعيارية المقابلة لكل تقدير للقدرة المقابل للدرجات الكلية التي حصل عليها أفراد عينة التدريج، وكذلك المحتملة على الاختبار حيث امتد مدى الدرجات المحتملة(الحقيقية) من الدرجة الكلية واحد، إلى الدرجة الكلية (18) حيث يحذف من التدريج الطلبة الحاصلون على الدرجة(صفر) والدرجة النهائية (19) وتراوحت تقديرات قدرات الطلبة من (3.97- إلى 4.07) لوجيت. ويمكننا إيجاد الدرجة الحقيقية المقابلة لأي مستوى محدد من القدرة من خلال المنحنى المميز للاختبار، والذي يعتمد شكله على عدد من العوامل منها عدد مفردات الاختبار ونموذج المنحنى المميز للمفردات المستخدمة، وقيَّم معالم المفردات فالمنحنى المميز للاختبار

هو العلاقة الوظيفية(التقييسية) بين الدرجة الحقيقية ومقياس القدرة، كما يوضحه الشكل رقم (20) للعلاقة المتجمعة بين الدرجة الكلية الخام المحتملة على الاختبار، والقدرة المقابلة لها باللوجيت.

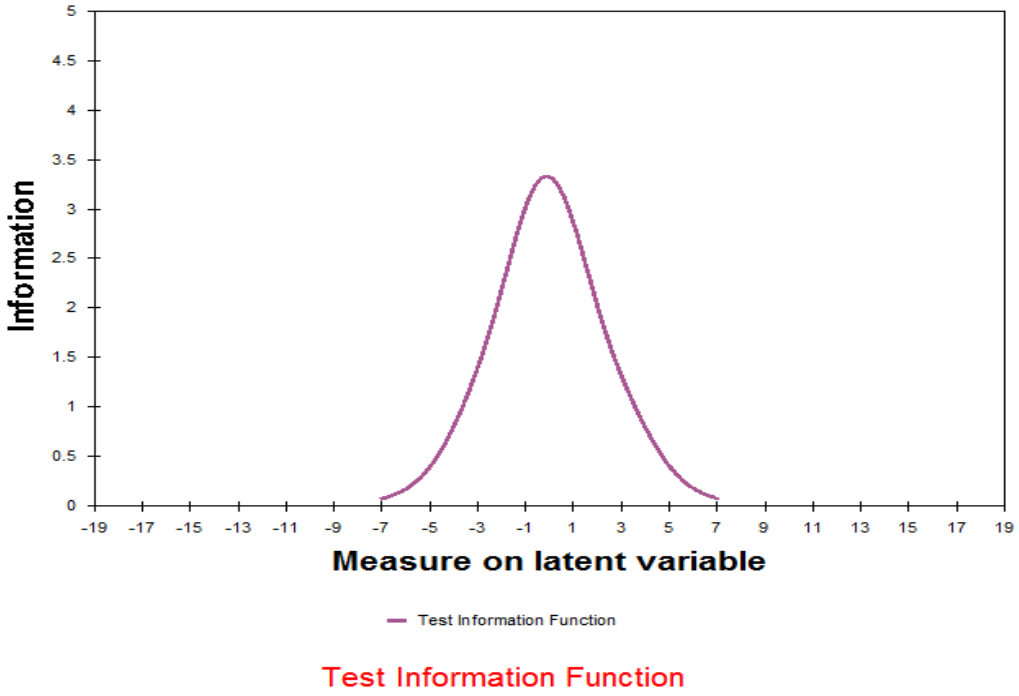



**الشكل (21) العلاقة التقييسية بين الدرجة الكلية المحتملة على الاختبار عند كل مستوى من مستويات القدرة المختلفة للأفراد.**

يوضح الشكل رقم (21) رسم بياني للدرجات الحقيقية المقابلة لمستويات القدرة، كدالة للقدرة على المحور الأفقي، في حين يمثل المحور العمودي الدرجات الحقيقية المتوقعة، والتي تراوحت ما بين(1و18) وهو عدد المفردات المكونة للاختبار وهو مقياس القدرة، بعد حذف الدرجة التامة والصفرية، كما يتضح أنَّ العلاقة بين الدرجة الكلية الخام والمحتملة على الاختبار، والقدرة المقابلة لها علاقة طردية موجبة في شكل دالة منتظمة متزايدة، حيث تزيد القدرة كلما ازدادت الدرجة الكلية الخام المقابلة لها.

فالدور الأساسي للمنحنى المميز للاختبار في نظرية الاستجابة للمفردة، هو إمدادنا بوسائل لتحويل درجات القدرة إلى درجات حقيقية، كما يتضح من الجدول رقم(40)، أنَّ قيَّم دالة معلومات اختبار الإحصاء تراوحت ما بين(0.80 و3.26) في النصف الأول من مدى الاختبار، وبعد ذلك بدأت في الانحدار تدريجياً، وهذا يدل على أنَّ الدالة تعبر عن الوضع الحقيقي والواقعي لمتطلبات أفراد العينة، كما تتفق تماماً مع القدرات المتوقعة لأفراد العينة، مما يؤكد دقة الاختبار في عملية القياس

والثقة في نتائجه، وقدرته على التمييز، وتحديد القدرات المختلفة لأفراد العينة. فمتوسط القدرة يساوي تقريباً صفراً، أي أنَّها عند أقصى دالة للمعلومات، فالاختبار يعطي أقصى دالة للمعلومات عند مستوى القدرة المتوسط حيث بلغت قيمتها (3.32) وخطأ معياري للقياس (0.55) المقابلة لمنتصف الدرجة الحقيقية (9) والتي تقابل صفر، أي صفر التدريج، أو القيمة المتوسطة لمعالم صعوبة المفردات وهو مؤشر على المكان الذي سيقوم الاختبار فيه بوظيفته على مقياس القدرة، كما يتضح من الشكل رقم (21).




**الشكل (22) يوضح علاقة دالة معلومات الاختبار مع الخطأ المعياري للقدرة الكامنة.**

يتضح من الشكل رقم (22) أنَّ منحنى دالة معلومات الاختبار متماثلاً بطريقة مقبولة، وذو ارتفاع واضح جداً في الوسط، وبالتالي فهو يأخذ الشكل الاعتدالي المتوقع لمثل هذه المنحنيات ويشير شكل المنحنى إلى أنَّه يتم تقدير القدرة بأعظم دقة بمصاحبة متوسط مقياس القدرة، وبما أنَّ القدرة متغير متصل تكون المعلومات متغيراً مستمراً أيضاً، لهذا تحدث ذروة دالة معلومات الاختبار في نقطة أعلى بدرجة طفيفة من النقطة المتوسط لمقياس القدرة.

كما يلاحظ انخفاض في الخطأ المعياري أين كانت أقصى قيَّم للمعلومات عن مستوى قدرات الطلبة التي يمدنا بها اختبار الإحصاء بدقة مرتفعة إلى حدٍ ما داخل مدى القدرة بين(-3.97 و4.07) لوجيت وهو ما كان متوقعاً، لأنَّ معظم مفردات الاختبار غطت هذا المدى، وخارج هذا المدى ينخفض مقدار المعلومات بسرعة ولا يتم تقدير مستويات القدرة بدرجة جيدة، مما يعني أنَّ الخطأ المعياري في التقدير يكون أقل ما يمكن عند مستويات القدرة التي تناظر أقصى المعلومات فأقل قيمة للخطأ المعياري، كما نلاحظ في الجدول أعلاه يساوي(0.55) والذي يناظر أقصى قيمة للمعلومات التي يمنحها الاختبار، وهي قيمة للقدرة تساوي(3.32) لوجيت عند الدرجة الخام(9) ومستوى القدرة(-0.18). وتنخفض قيمة المعلومات التي يمكن الحصول عليها من الاختبار بشكل واضح مع زيادة القدرة عن(4.07) أو تنخفض عن مستوى القدرة(-3.97) وهو ما يؤكد حاجة الاختبار إلى إضافة مفردات تغطي تلك المستويات من القدرة، فمقدار المعلومات التي تسهم بها مجموعة من المفردات عند مستوى قدرة معين يتناسب تناسباً عكسياً مع مربع سعة فترة الثقة حول تقديرات القدرة عند مستوى القدرة ($\theta$) معين، أي أنَّ الخطأ المعياري للقياس يساوي $SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$ فبالتعويض في هذه المعادلة بالقيم السابقة، فإنَّ القيمة للخطأ المعياري للقياس تساوي تقريباً القيمة(0.55) والتي تناظر أقصى قيمة لدالة معلومات اختبار الإحصاء.

**4. عــرض وتحليــل نتائــج التســاؤل الرابــع والــذي ينــص علــى:**

- **هل يمكن اشتقاق معايير كمية تفسر تقديرات الأفراد على اختبار الإحصاء في صورته النهائية بعد تدريجه على الطلبة؟**

تمَّ تحويل تدريج وحدات اللوجيت لصعوبة مفردات الاختبار في صورته النهائية(بعد التدريج) إلى تدريج وحدات مئوية بوحدة الواط، حيث تشتق وحدة اللوجيت من نموذج راش اللوغاريتمي والذي يتناول التقدير الاحتمالي للاستجابة الصحيحة للفرد على المفردة ضمن التدريج الذي نقطة صفره متوسط صعوبات المفردات، ونظراً لوجود مشكلة الكسور، وعدم ألفة الباحثين والمدرسين لهذا النوع من التدريج، ولمعالجة هذه العيوب، وبما أنَّ نظام التدريج المئوي هو أكثر التدريجات ألفة في أغلب مجالات القياس، أوجد بعض المتخصصين في القياس عدداً من الوحدات الجديدة مثل وحدة التدريج المعتمد على محك مستقل(سيت) ووحدات التدريج الجماعية(نيت) ووحدات التدريج الخاصة باحتمال

الاستجابة الصواب(شيبس) ووحدة التدريج المستخدمة في المقاييس البريطانية للقدرات(باص) ولهذا وقع اختيار الباحث من تلك الوحدات على وحدة قياس الواط التي قدمها(Masters,1984) لتحويل تقديرات كل من صعوبات المفردات وقدرات الأفراد من وحدة القياس

باللوجيت إلى هذه الوحدة الجديدة وهي وحدة الواط، وذلك باستخدام المعادلتين التاليتين:

$$B= 50 + (15 / \log 4)b$$

$$D= 50 + (15 / \log 4)d$$

حيث أنَّ B، D هما تقدير كل من القدرة والصعوبة على الترتيب مقدرتين بالواط، وb، d هما تقدير كل من القدرة والصعوبة على الترتيب مقدرتين باللوجيت، وحيث log هو اللوغاريتم الطبيعي ولذا يكون متوسط صعوبة المفردات 50، كما تتدرج كل من B، D من الصفر إلى المائة.

إضافة إلى ذلك استخرج الباحث معايير الدرجات التائية، وكذلك معايير الرتب المئينية المقابلة لكل درجة من الدرجات الخام التي حصل عليها الأفراد في عينة التقنين، وهذه المعايير هي الطريقة الكلاسيكية المألوفة، وقد قدمها الباحث لإجراء مقارنة بين المعايير المستخرجة وفق النظريتين الكلاسيكية في القياس، ونظرية الاستجابة للمفردة (نموذج راش). وفيما يلي عرض لهذه المعايير التي على أساسها يمكن تحديد وتفسير قدرة الطلبة على اختبار الإحصاء سواء بالتفسير المحكي أو المعياري المرجع. والجدول رقم (41) يوضح تقدير صعوبة مفردات اختبار الإحصاء في صورته النهائية مقدرة بوحدة اللوجيت، ووحدة الواط.

**الجدول (41) تقدير صعوبة المفردات مقدرة باللوجيت والواط مرتبة تصاعدياً تبعاً للصعوبة (19) مفردة.**

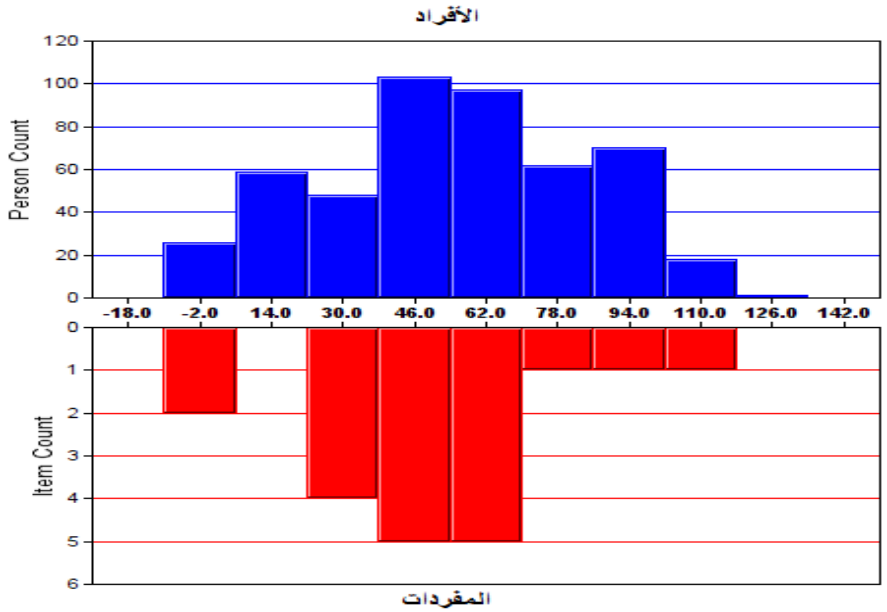
| المفردة | الدرجة الخام | تقديرات الصعوبة | | أخطائها المعيارية | |
|---|---|---|---|---|---|
| | | باللوجيت | بالواط | باللوجيت | بالواط |
| 20 | 53 | 3.50 | 103 | 0.17 | 2.53 |
| 12 | 66 | 3.16 | 97 | 0.16 | 2.34 |
| 14 | 159 | 1.47 | 72 | 0.12 | 1.84 |
| 13 | 170 | 1.31 | 70 | 0.12 | 1.82 |
| 18 | 207 | 0.71 | 62 | 0.12 | 1.77 |
| 16 | 211 | 0.72 | 61 | 0.12 | 1.77 |
| 9 | 212 | 0.78 | 61 | 0.12 | 1.77 |
| 11 | 237 | 0.37 | 55 | 0.12 | 1.76 |
| 5 | 250 | 0.19 | 53 | 0.12 | 1.76 |
| 6 | 273 | 0.13- | 48 | 0.12 | 1.77 |
| 4 | 291 | 0.43- | 44 | 0.12 | 1.79 |
| 1 | 294 | 0.38- | 44 | 0.12 | 1.79 |
| 7 | 310 | 0.66- | 40 | 0.12 | 1.81 |
| 2 | 328 | 0.93- | 36 | 0.12 | 1.85 |
| 3 | 332 | 0.99- | 35 | 0.12 | 1.86 |
| 8 | 336 | 1.05- | 34 | 0.12 | 1.87 |
| 17 | 357 | 1.39- | 29 | 0.13 | 1.93 |
| 15 | 432 | 2.96- | 6 | 0.17 | 2.55 |
| 19 | 443 | 3.30- | 1 | 0.19 | 2.79 |

يتضح من الجدول رقم(41) تقدير الصعوبة لمفردات الاختبار مقدرة بوحدة اللوجيت، وكذا بوحدة الواط مرتبة تصاعدياً تبعاً لصعوبتها، والتي انحصرت فيها الدرجات الخام ما بين(53 و443) وكانت أقل قيمة من نصيب المفردة(20) والذي تقابله قيمة تقديرية للصعوبة بوحدة اللوجيت(3.50) وخطأ معياري قدره(0.17) والذي يقترب من الصفر، والدرجة(103) بوحدة الواط وبخطأ معياري قدره(2.53) في حين كانت أكبر قيمة من نصيب المفردة(19) والذي تقابله قيمة تقديرية للصعوبة بوحدة اللوجيت (3.30-) وخطأ معياري قدره(0.19) والذي يقترب من الصفر هو كذلك، والدرجة(1) بوحدة الواط وبخطأ معياري قدره(2.79) فالطريقة الملائمة لتفسير أي قيمة رقمية لمعلم الصعوبة لمفردات اختبار الإحصاء، تقوم على تحديد أين تعمل المفردة على مقياس القدرة. والجدول رقم (43) يوضح معايير القدرة لأفراد عينة البحث، باستخدام نموذج راش (وحدة اللوجيت – وحدة الواط) والنظرية الكلاسيكية في القياس (المعايير التائية – الرتب المئينية).

**الجدول (42) يوضح معايير القدرة لعينة التقنين (484) طالباً وطالبة، وفق نموذج راش (وحدة اللوجيت ووحدة الواط) ووفق النظرية الكلاسيكية في القياس (المعايير التائية – الرتب المئينية).**

| الدرجة الخام | معايير القدرة المقابلة | | | |
|---|---|---|---|---|
| | وحدة اللوجيت | وحدة الواط | الدرجة التائية | الرتب المئينية |
| 1 | 3.97- | 10- | 10 | 1 |
| 2 | 3.03- | 5 | 20 | 4 |
| 3 | 2.39- | 14 | 26 | 8 |
| 4 | 1.90- | 21 | 31 | 14 |
| 5 | 1.49- | 28 | 35 | 20 |
| 6 | 1.13- | 33 | 39 | 25 |
| 7 | 0.80- | 38 | 42 | 30 |
| 8 | 0.48- | 43 | 45 | 36 |
| 9 | 0.18- | 47 | 48 | 42 |
| 10 | 0.12 | 52 | 51 | 46 |
| 11 | 0.43 | 56 | 54 | 51 |
| 12 | 0.75 | 61 | 57 | 57 |
| 13 | 1.09 | 66 | 61 | 65 |
| 14 | 1.46 | 72 | 65 | 72 |
| 15 | 1.89 | 78 | 69 | 78 |
| 16 | 2.41 | 86 | 74 | 85 |
| 17 | 3.08 | 96 | 81 | 92 |
| 18 | 4.07 | 111 | 91 | 98 |
| 19 | 5.48 | 132 | 105 | 99 |

يتضح من الجدول رقم (42) تقدير القدرة المقابل لكل درجة من درجات عينة التقنين مقدرة بوحدة اللوجيت، وتقابلها كذلك نفس هذه التقديرات مقدرة بوحدة الواط. وفي نفس هذا الجدول تقدم الرتب المئينية والدرجات التائية المقابلة لكل درجة من هذه الدرجات الخام التي سجلها أفراد عينة التقنين.




**الشكل (23) توزيع العتبات الفارقة للمفردات والأفراد على عينة التدريج (484) طالباً وطالبة.**

يتبين من الشكل رقم (23) أنَّ توزيع القدرات لمجموعة(التدريج) المعايرة، بأنَّه متماثلاً بطريقة مقبولة وذو ارتفاع واضح جداً في الوسط، والذي يشير إلى أنَّه يتم تقدير الصعوبة والقدرة بأعظم دقة بمصاحبة متوسط مقياس القدرة، والذي يستطيع أنَّ يأخذ أياً من الأشكال، حيث لا يفترض فيه أنَّ يأخذ شكل توزيع المنحنى الطبيعي، والذي يبين العتبات الفارقة للمفردات بدلاً من الصعوبة، والذي يظهر كذلك توزع عتبات مفردات اختبار الإحصاء، من أسهل المفردات (-3.30) لوجيت إلى الأصعب نوعاً ما (3.50) لوجيت. وتتدرج قدرات الأفراد على الاختبار من أقل قدرة حوالي (-3.97) لوجيت إلى أكبر قدرة حوالي (5.48) لوجيت.

**5. عـرض وتحليـل نتائـج التسـاؤل الخـامس والـذي ينـص علـى:**

**- ما دلالات الخصائص السيكومترية (الصدق والثبات) لمفردات الاختبار بعد معايرة مفرداته باستخدام نموذج راش؟**

اعتمد الباحث على برنامج(Winsteps) لتقدير مؤشرات صدق وثبات درجات اختبار الإحصاء بصورته النهائية (بعد التدريج).

**أولاً: تقويم الصدق**

تبين للباحث من خلال تحليل النتائج الموضحة في الجدول رقم (39) والشكل رقم (18) إلى أنَّ

قيَّم مؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية والتقاربية لمفردات اختبار الإحصاء كانت قريبة من الوضع المثالي الذي يفترضه النموذج وهو القيمة واحد، والذي تراوحت قيمه ما بين (0.64 و1.13) بمتوسط حسابي (0.96) وانحراف معياري (0.12) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية(الخارجية)، وتراوحت ما بين (0.93 و1.09) وبمتوسط حسابي (1.01) وانحراف معياري (0.05) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية (الداخلية). وهي قيم في مجال القياس المنتج الذي حدده كلٍ من (Wright and Linacre, 1994) وما بين (0.70 و1.30) أو ما بين (0.50 و1.5) عموماً والقريبة من الواحد. وتتعلق هذه المؤشرات بصدق تدرج مفردات الاختبار في تعريفها للمتغير موضوع القياس. كما تتعلق أيضاً بصدق تدرج قدرات الأفراد على متصل هذا المتغير، الذي يقوم على صدق استجابات الأفراد على الاختبار.

ولإضفاء المزيد من المصداقية على الاختبار الحالي، قام الباحث باستخدام الأساليب التالية للتحقق من الصدق:

**1. التحقق من قطبية المفردات (*Polarity*)**

أظهرت نتائج التحليل الموضحة في الجدول رقم (39) أنَّ نقطة قياس معاملات الارتباط لكل مفردات اختبار مقياس الإحصاء كانت قيمها موجبة والتي تراوحت ما بين (0.40) للمفردة (19) والقيمة (0.64) للمفردتين (16 و18) مما يدل على أنَّ مفردات اختبار الإحصاء تقيس التكوين الفرضي الذي ينوي الباحث قياسه.

**2. التحقق من أحادية البعد**

استخدم الباحث طريقة التقدير وفقاً لنموذج (راش) للمكونات الأساسية المعتمدة على البواقي والتي تُظهر الاختلافات بين الأبعاد، والنتائج يوضحها الجدول رقم (43).

**الجدول (43) نتائج التحليل العاملي للمكونات الأساسية للبواقي (PCAR)**
**باستخدام نموذج (راش) بعد التدريج.**

| حجم تباين البواقي المعيارية مقدرا بوحدات القيم Eigen. | الجذر الكامن | الملاحظ | المتوقع |
|---|---|---|---|
| التباين الكلي في الاستجابات | 36.8909 | %100 | %100 |
| التباين الذي فسره العامل الرئيسي، تقديرات نموذج (راش) | 17.8909 | %48.5 | %48.6 |
| التباين المفسر بواسطة الأفراد | 9.1130 | %24.7 | %24.8 |
| التباين المفسر بواسطة المفردات | 8.7779 | %23.8 | %23.9 |
| مجموع التباين غير المفسر | 19.0000 | %51.5 | %51.4 |
| التباين الذي فسره العامل الثاني (الأول في البواقي) | 2.1284 | %5.8 | %11.2 |

يتضح من نتائج الجدول رقم(43) تحقق المحكات الثلاثة المعتمدة في تحديد أحادية البعد، حيث يتحقق المحك الأول وكقاعدة عامة إذا كانت قيمة التباين المفسر بواسطة التقديرات( Raw variance explained by measures) أكبر من أو يساوي(50%) واسترشاداً بما أشار إليه لينكر(Linacre) مطور برنامج(Winsteps) الذي حدد المجال من(20% إلى 80%) بأنَّه مجال مقبول و يعد مؤشراً قوياً لأحادية البعد، وهو ما تحقق في بحثنا الحالي، حيث ارتفع مؤشر القيمة الملاحظة للتباين الذي فسره العامل الرئيسي إلى(48.5%) بعد حذف مفردة واحدة من مفردات الاختبار عن قيمته السابقة والتي كانت(47%) قبل الحذف، كما أنَّ الفرق بين القيمة الملاحظة والقيمة المتوقعة للنموذج هي(0.1%) فقط، أما المحك الثاني هو أنَّ نسبة التباين الذي يفسره العامل الثاني(Unexplained Variance in 1st contrast) يجب أن تكون قيمته أقل أو قريب من(5%) وهذا ما تحقق في البحث الحالي تقريباً بالقيمة(5.8%) وهذا دليل آخر على أحادية البعد. في حين يعتمد المحك الثالث، على قيمة الجذر الكامن لنسبة التباين الذي يفسره العامل الثاني (Unexplained Variance in 1stcontrast) والذي يجب ألا يتعدى أو يكون أقل قيمة له (3) وهذا ما تحقق في البحث الحالي بتسجيل القيمة (2.1284) وهي أقل من القيمة (3) المنصوح بها من ليناكر (Linacre, 2008, p.272).

**ثانياً: ثبات الاختبار**

يبدو ثبات القياس بتحقق مطالب الموضوعية في القياس، عندما تستخدم أداة القياس التي تبنى باستخدام نموذج(راش) ولإضافة المزيد من المعلومات حول ثبات الاختبار استخدم الباحث الأساليب التالية:

**1. معامل ثبات الأفراد ومعامل ثبات المفردات**

باستخدام نموذج(راش) أمكن الحصول على القيم المتحررة لكل من صعوبة المفردات وقدرات الأفراد، ثمَّ الحصول على نوعين من المعاملات، معامل الثبات الخاص بالأفراد، ومعامل الثبات الخاص بالمفردات، فالثبات وفق نظرية الاستجابة للمفردة، يعني الدقة في تقدير موقع كل من الأفراد والمفردات على متصل السمة، ويمكن تحديد دقة المفردات في تعريف هذا المتصل بحساب معامل الفصل للمفردات(GI) والذي يعرف بأنَّه النسبة بين الانحراف المعياري للقيم التدريجية المتحررة من المفردات، ومتوسط الخطأ المعياري لهذه القيم، والتي يوضحها الجدول رقم(44).

**الجدول (44) ملخص لنتائج معاملات الثبات والفصل وأخطاءهما المعيارية للأفراد والمفردات وفق نموذج راش بعد التدريج النهائي لاختبار الإحصاء في صورته النهائية 19 مفردة.**

| معامل الثبات | | معامل الفصل | | الانحراف المعياري SD | | الجذر التربيعي لمتوسط مربعات الأخطاء المعيارية للقياس RMSE | |
|---|---|---|---|---|---|---|---|
| الأفراد | المفردات | الأفراد | المفردات | الأفراد | المفردات | الحقيقي للأفراد REAL RMSE | الحقيقي للمفردات REAL RMSE |
| 0.85 | 0.99 | 2.36 | 12.42 | 1.75 | 1.67 | 0.74 | 0.13 |
| | | | | | | لنموذج الأفراد MODEL RMSE | لنموذج المفردات MODEL RMSE |
| 0.87 | 0.99 | 2.53 | 12.57 | 1.77 | 1.67 | 0.70 | 0.13 |
| معامل الثبات ألفا كرونباخ أو (KR-20) يساوي 0.89. والخطأ المعياري للقياس يساوي SEM=1.66 | | | | | | | |

يتضح من الجدول رقم (44) أنَّ قيمة معامل ثبات المفردات بلغت (0.99) ومعامل الفصل لمفردات الاختبار بلغ (12.57) وهي قيمة أكبر من (2) كما بلغت قيمة معامل الثبات الأفراد (0.87) والقيمة (2.53) لمعامل الفصل للأفراد (GP) وهي قيمة كذلك أكبر من (2).

كما حدد الباحث عدد الطبقات الإحصائية المتمايزة لكل من المفردات والأفراد باستخدام الصيغة الرياضية التالية:(H=(4*G+1) /3 ) حيث يرمز H إلى عدد الطبقات الإحصائية، ويرمز G لمعامل الفصل وبالتعويض في المعادلة السابقة بالنسبة للأفراد(H=(4*2.53+1) /3) أي تساوي(4.71) تقريباً خمس طبقات إحصائية، وكذلك بالنسبة للمفردات(H=(4*12.57+1) /3) وبالتالي فإن عدد الطبقات الإحصائية للمفردات بلغ(18) طبقة .

كما يتضح من الجدول رقم (44)، أنَّ قيم جذر مربع متوسط الخطأ المعيـــاري للثبات الحقيقي

للأفراد والمفردات لمعامل الفصل للأفراد والمفردات (REAL RMSE) وهو تباين الخطأ الحقيقي وهو يمثل أسوأ حالة للثبات أو أدنى قيمة ثبات للاختبار، حيث سجل الباحث القيم (0.13، 0.74) على التوالي، فقيمة جذر مربع متوسط الخطأ المعياري للنموذج بالنسبة للأفراد، تكافئ قيمة الخطأ المعياري للقياس في النظرية الكلاسيكية للقياس والتي بلغت (1.66) (Linacre, 202, p.36).

**II). منـــاقشة نتائـــج تســـاؤلات الـــدراسة.**

**1.مناقشة نتائج التساؤل الأول الذي ينص على:**

- **ما درجة تحقيق بيانات اختبار مقرر الإحصاء لافتراضات نموذج راش لدى طلبة السنة أولى علوم اجتماعية جذع مشترك؟**

بينت نتائج الدراسة ملاءمة بيانات اختبار الإحصاء الوصفي لنموذج راش وذلك بتحقيقها لافتراضات النموذج حيث أظهرت النتائج تحقق كل من(أحادية البعد، دالة استجابة المفردة، افتراض الاستقلال الموضعي، افتراض تدني عامل التخمين، التحرر من عامل السرعة) أي أنَّ التساؤل الأول قد تحقق بتحقق جميع افتراضات النموذج، وبالتالي تمَّ التحقق من افتراضات نموذج راش بتحليل استجابات الطلبة على فقرات الاختبار باستخدام برنامج(Winsteps) على ثلاثة مراحل، في المرحلة الأولى تمَّ التحقق من مطابقة الأفراد للنموذج وحذف الأفراد غير المطابقين، وفي المرحلة الثانية تمَّ التحقق من مطابقة الفقرات للنموذج وحذف الفقرات غير المطابقة، وذلك بعد حذف الأفراد غير المطابقين للنموذج، وفي المرحلة الثالثة وبعد حذف الأفراد غير المطابقين للنموذج والفقرات غير المطابقة تمَّ تدريج فقرات الاختبار بصورته النهائية وفق نموذج راش وهذا يعتبر دليلاً على موضوعية تفسير نتائج الاختبار وموضوعية القياس وقد تمَّ التحليل باستخدام نموذج راش كما يلي:

قام البرنامج بحذف استجابات الأفراد الذين يحصلون على العلامة الكاملة على الاختبار وعددهم(05) وكذلك بحذف استجابات الأفراد الذين يحصلون على العلامة(صفر) على الاختبار وكان عددهم(05) وبالتالي قام البرنامج بحذف استجابات الأفراد، وأصبح عدد أفراد عينة الدراسة التي دخلت في التحليل الأولي(492) فرداً كما يقوم البرنامج بحذف الفقرات التي يجيب عنها جميع الأفراد إجابة صحيحة، وكذلك جميع الفقرات التي يجيب عنها جميع الأفراد إجابة خاطئة، ولم يكن هناك أي فقرة أجاب عنها جميع الأفراد إجابة صحيحة، وكذلك لم يكن هناك أي فقرة أجاب عنها جميع الأفراد

إجابة خاطئة، وبالتالي لم يحذف البرنامج أي فقرة ودخلت جميع فقرات الاختبار(20) في التحليل الأولي وقد يُعزى ذلك لحرص الباحث على أن لا يحتوي الاختبار على فقرات صعبة جداً يخفق جميع الأفراد في الإجابة عنها إجابة صحيحة، ولا فقرات سهلة جداً يجيب عنها كل الأفراد إجابة صحيحة.

ثم تمَّ التحقق من مطابقة الأفراد للنموذج بتقدير قدرة كل فرد بالإضافة إلى الخطأ المعياري في قياس هذه القدرة، وكذلك قيم إحصائي المطابقة الكلى وله مؤشران، مؤشر المطابقة الداخلية التقاربي(ZSTD) ويعبر عنه كذلك بإحصائي متوسط المربعات التقاربي، وكذلك مؤشر المطابقة الخارجية التباعدي(ZSTD) ويعبر عنه كذلك بإحصائي متوسط المربعات التباعدي(MNSQ) وهو مؤشر إحصائي يعد بديلاً عن السابق، وله صفات مقاربة ومشابهة ولكنَّه أكثر حساسية للسلوكيات غير المتوقعة في استجابات الأفراد عن الفقرات التي تكون بعيدة عن مستوى قدراتهم، وذلك لكل تقدير من تقديرات القدرة، كما في الجدول(34) الذي يبين المتوسط الحسابي والانحراف المعياري لكل تقدير من تقديرات القدرة والخطأ المعياري في قياس هذه القدرة، وقيم إحصاءات المطابقة الداخلية والخارجية ومتوسطات المربعات للمطابقة الداخلية والخارجية.

وعند تفحص قيم إحصائي المطابقة الخارجية والداخلية للأفراد، التي تشير إلى مطابقة قدرة الفرد مع قدرات مجموعة الأفراد التي ينتمي إليها في قياس السمة التي يقيسها الاختبار، فإذا كانت قيمة هذان الإحصائيان تزيد على (+2) تعتبر قدرة الفرد غير مطابقة لقدرات مجموعة حيث تبين أن (08) أفراد غير مطابقين للنموذج، تزيد قيم إحصائي المطابقة الخارجية والداخلية لدرجاتهم على القيمة (+2) أو قيم متوسطات المربعات المقابلة لدرجاتهم تزيد على (1) وهي القيم التي يتوقعها النموذج.

وقد يُعزى هذا إلى أنَّ الأفراد يجيبون صواباً عن الفقرات التي تزيد درجة صعوبتها عن مستوى قدراتهم أو يفشلون في الإجابة عن فقرات درجة صعوبتها أدنى من مستوى قدراتهم وقد يعزى ذلك أيضاً إلى ميل بعض المفحوصين من ذوي القدرات الدنيا إلى التخمين عند الإجابة على الفقرات التي قيَم معامل صعوبتها أكبر من قدراتهم، أما بالنسبة للأفراد من ذوي القدرات العليا الذين تمَّ حذفهم فقد يعزى إخفاقهم في الإجابة عن الفقرات التي تقل درجة عن مستوى قدراتهم إلى أنَّهم يميلون إلى تفسير ما تقيسه الفقرة بأكثر مما تحتمل، وقد يعزى أيضاً إلى الإهمال أو التسرع.

بعد استبعـــاد المفردات والأفـــراد غير المطابقين لنموذج راش تمَّ اعادة التحليل للحصول على

تقديرات متحررة من صعوبة المفردات وقدرات الأفراد، ويتضح من الشكل رقم(39) أن مؤشرات المطابقة للمفردات وحدود المطابقة ومواقع المفردات من هذه الحدود وفقاً لإحصائي متوسط المربعات التباعدي والتقاربي، تشير إلى أنَّ صعوبة مفردات اختبار مقياس الإحصاء في صورته النهائية(19) مفردة، بأنَّها في المجال المقبول للصعوبة، كما تشير قيم مؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية والتقاربية لمفردات الاختبار إلى أنَّها قريبة من الوضع المثالي الذي يفترضه النموذج وهو القيمة(1)، والذي تراوحت قيمه ما بين(0.64 و1.13) بمتوسط حسابي(0.96) وانحراف معياري(0.12) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية (الخارجية) وتراوحت ما بين(0.93 و1.09) وبمتوسط حسابي(1.01) وانحراف معياري(0.05) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية(الداخلية)، وهي قيم في مجال القياس الذي حدده كلٍ من(Wright & Linacre, 1994) ما بين(0.70 و1.30) أو ما بين(0.50 و1.5) عموماً والقريبة من الواحد.

إنَّ حذف الأفراد(18) الذين لم تتطابق استجاباتهم مع افتراضات النموذج، يعني الإبقاء على الأفراد الذين تطابقت استجاباتهم مع افتراضات النموذج، أي الذين يكون هناك اتساقاً بين استجاباتهم على الفقرة واحتمال نجاحهم فيها، وهذا يعني وجود اتساق بين استجاباتهم على الفقرة ودرجاتهم الكلية على الاختبار، مما يدل على الاتفاق بين السمة التي تقيسها الفقرة والسمة التي تقيسها باقي فقرات الاختبار، ويعد هذا مؤشرا لصدق القياس، كما تمَّ التحقق من مطابقة الفقرات للنموذج بعد استبعاد الأفراد الذين لم تتطابق استجاباتهم الملاحظة مع توقعات النموذج، تمَّ إعادة التحليل الحذف الفقرات غير المطابقة لنموذج راش، وفي ضوء محكات المطابقة للفقرات المستخدمة، تبين وجود فقرة واحدة رقم(10) حيث زادت قيم متوسطات المربعات الموزونة لها الواحد الصحيح، وهذا مؤشر إلى إنَّها فقرة غير مطابقة وبعيدة عن توقعات النموذج، وقد يعزى ذلك إلى عيب في صياغتها أو عدم صدقها في قياس ما تقيسه بقية الفقرات، أو يعني أن الفقرة شديدة التشابه أو الاعتماد على بعض فقرات المقياس، أي لا تكون مستقلة عنها.

كما أنَّ حذف الفقرات غير المطابقة الافتراضات النموذج وعددها (1) يدل على أنَّ كل فقرة من الفقرات المتبقية في الاختبار وعددها (19) فقرة مطابقة الافتراضات النموذج وتتفق في تعريفها

للمتغير مع ذلك الذي تعرفه بقية فقرات الاختبار، وهذا يعني أنَّ هناك اتفاقاً بين السمة التي تقيسها الفقرة والسمة التي تقيسها باقي الفقرات، وهذا يعد مؤشرا لصدق القياس.

تمَّ التحقق من جودة مطابقة البيانات الناتجة عن استجابات أفراد العينة (484) على الاختبار والتحقق من موضوعية تفسير نتائج الاختبار بصورته النهائية (19) فقرة أي بعد حذف الأفراد غير المطابقين والفقرات غير المطابقة للنموذج، وكذلك التحقق من الافتراضات الخاصة بنموذج راش إذ يعد ذلك تحقيقاً لمدى صدق النموذج في تحقيقه لموضوعية القياس.

وقد اتفقت هذه النتائج إلى حد كبير مع دراسات كل من العطوي ومسعودي (2019) ودراسة كل من بوبو وشروق وسمان (2019) ودراسة رشيد وبوقصارة (2018) ودراسة الطراونة (2016) الصبَّاح (2016)، المصري (2015)، السامرائي والخفاجي (2012)، الربيعي (2012)، الجبوري (2012) ويلسون وماك(Wilson and Mac) (2007). إلا انَّها اختلفت مع بعض هذه الدراسات من حيث حجم العينة المستخدمة في الدراسة، وكذا من حيث البرامج التي تستعمل للتحقق من افتراضات نموذج راش فقد وافقت هذه الدراسة استعمال برنامج (Winsteps) كما في دراسة كل من الطراونة (2016) وبوبو وشروق وسمان (2019) وزياد وبوقصارة (2018) واختلفت مع دراسة السامرائي والخفاجي (2013) التي استعملت برنامج راسكال (Rascal) ودراسة العطوي والمسعودي (2019) التي استخدمت برنامج (Bilog–mg3) ودراسة الطراونة (2016) التي استعملت برنامج (Bigsteps).

**2. منـــاقشة نتائـــج التســـاؤل الثانـــي الـــذي ينـــص علـــى:**

**- ما قيم تقديرات معلم الصعوبة لفقرات الاختبار التحصيلي بعد التدريج وفق نموذج راش؟**

بعد تحليل استجابات الطلبة في برنامج(Winsteps) لتقدير الارجحية القصوى المشتركة (JMLE) لتحرير ومعايرة الاختبار، وبعد استبعاد الأفراد غير المطابقين للنموذج تمت اعادة التحليل للكشف عن المفردات غير المطابقة للنموذج حيث تمَّ تقدير معلم الصعوبة لكل مفردة والخطأ المعياري وقد توصل اباحث الي أنَّ المتوسط الحسابي لمتوسط المربعات الداخلية والخارجية(MNSQ) يساوي واحد بالنسبة لمتوسطات الداخلية، وقريب جداً من الواحد بالنسبة لمتوسطات المربعات الخارجية، كما أنَّ المتوسط الحسابي لمتوسط احصائي المطابقة الداخلية والخارجية(ZSTD) قريب من الصفر الوضع المثالي الذي يفترضه النموذج هو(0،1) إضافة إلى أنَّ الانحراف المعياري قريب من الواحد، ليتبين

عدم مطابقة مفردة واحدة لنموذج راش حيث زادت المربعات الموزونة لها الواحد الصحيح وهي المفردة رقم(10) وبعد استبعاد الأفراد والمفردات غير المطابقين للنموذج تمت إعادة التحليل للحصول على تقديرات متحررة من صعوبة المفردات وقدرات الافراد ليتضح أنَّ التقديرات النهائية المتحررة من قدرات الأفراد تراوحت بين الدرجتين(18-1) وبمتوسط توزيع القدرة بلغ(0.28) لوجيت وبانحراف معياري بلغ(1.9) لوجيت أما الخطأ المعياري للمتوسط الحسابي لتقديرات القدرة فقد بلغ(0.68) وهي قيمة قريبة مما يفترضه النموذج وبالتالي فإنَّ مستوى الاختبار بشكل عام في متناول مستوى أفراد العينة مما يقلل عوامل التخمين كما تؤكد ذلك (كاظم، 1988ب).

وعليه فإنَّ نتائج التحليل تشير إلى أنَّ صعوبة المفردات الاختبار في صورته النهائية(19) مفردة بأنَّها في المجال المقبول للصعوبة كما تشير قيم ملاءمة الإحصائية لمتوسط المربعات التباعدية والتقاربية لمفردات الاختبار إلى أنَّها قريبة من الوضع المثالي الذي يفرضه النموذج وهو(1) والذي تراوحت قيمته ما بين(0.64 و1.13) بمتوسط حسابي(0.96) وانحراف معياري(0.12) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية وتراوحت بين(0.93 و1.09) ومتوسط حسابي(1.01) وانحراف معياري(0.05) بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية وهي قيم في مجال القياس الذي حدده كل من(Wright & Linacre,1994) ما بين(0.70 و1.30) و هو مؤشر احصائي أكثر حساسية للسلوكيات غير المتوقعة.

كما يلاحظ بأنَّ المنحنيات المميزة لمفردات اختبار الإحصاء بعد تدريجه باستخدام نموذج (راش)لا تتقاطع، وإنَّما تختلف عن بعضها في الصعوبة فقط وتتلاقى فقط عند المحور الأفقي الذي يمثل متصل القدرة أو السمة.

كما تمَّ حساب قيَّم معلم الصعوبة لفقرات الاختبار بعد تدريجها باستخدام نموذج راش والخطأ المعياري لكل معلم كما في الجدول رقم (39) حيث يتضح أن قيم معلم الصعوبة للفقرات تراوح بين (-3.30) للفقرة(19) وهي فقرة سهلة إلى(3.50) للفقرة(20) وهي فقرة صعبة وحيث أن معظم الطلبة تقع قدراتهم بين المدى(-2) إلى(2) فإنَّ فقرات الاختبار مقبولة.

وللتعرف على جودة تدريج فقرات الاختبار تمَّ حساب الفروق بين مستوى صعوبة كل فقرتين متتاليتين في التدريج ومقارنته بمجموع الخطأ المعياري لهما لمعرفة فيما إذا كانت هناك فراغات فإذا

كان الفرق بين الصعوبتين أكبر من مجموع الخطأ المعياري لهما فإنَّ هذا يعني وجود فراغ بين الفقرتين على متصل القدرة المقاسة مما يستدعي إضافة فقرات جديدة لتغطية الفراغ وقياس القدرة ويتضح أنَّ الفرق بين الصعوبات أقل من مجموع الخطأ المعياري ما عدى في الحالات التالية: الفرق بين صعوبتي المفردتين(14 و12) يساوي(1.69) ومجموع الخطأ المعياري للقياس(0.28) الفرق بين صعوبتي المفردتين(13 و18) يساوي(0.53) ومجموع الخطأ المعياري للقياس (0.24) الفرق بين صعوبتي المفردتين(15 و17) يساوي (1.57) ومجموع الخطأ المعياري للقياس (0.30) الفرق بين صعوبتي المفردتين(15 و19) يساوي (0.34) ومجموع الخطأ المعياري للقياس (0.36) أي أنَّ الخطأ المعياري للقياس لهذه المفردات لا يغطي الفراغ، حيث يمكن اضافة فقرات إلى الاختبار لتغطية هذه الفراغات مما يساعد في توفير صور متكافئة من الاختبار عند الاستخدام.

**3. مناقشة نتائج التساؤل الثالث الذي ينص على:**

- **ما كمية المعلومات التي يقدمها الاختبار عند مستويات القدرة المختلفة؟**

تمَّ تقدير كمية المعلومات التي يقدمها اختبار الإحصاء عند مستويات القدرة المختلفة، والعلاقة بين الدرجة الخام والقدرة بصورته النهائية، باستخدام نموذج راش، وتقدير دالة المعلومات المقابلة لكل درجة خام وكل قدرة، وتقدير الخطأ المعياري للقدرة، حيث امتد مدى الدرجات المحتملة(الحقيقية) من الدرجة الكلية واحد، إلى الدرجة الكلية(18) حيث يحذف من التدريج الطلبة الحاصلون على الدرجة(صفر) والدرجة النهائية(19) وتراوحت تقديرات قدرات الطلبة من (-4.07،3.97) لوجيت و يمكننا إيجاد الدرجة الحقيقية المقابلة لأي مستوى محدد من القدرة من خلال المنحنى المميز للاختبار ومن خلال الرسم بياني للدرجات الحقيقية المقابلة لمستويات القدرة يتضح أنَّ العلاقة بين الدرجة الكلية الخام والمحتملة على الاختبار، والقدرة المقابلة لها علاقة طردية موجبة في شكل دالة منتظمة متزايدة حيث تزيد القدرة كلما ازدادت الدرجة الكلية الخام المقابلة لها. كما يتضح من خلال تحليل الاختبار أنَّ قيَّم دالة معلومات اختبار الإحصاء تراوحت ما بين (0.80 و2.26) في النصف الأول من مدى الاختبار وبعد ذلك بدأت في الانحدار تدريجياً، وهذا يدل على أنَّ الدالة تعبر عن الوضع الحقيقي والواقعي لمتطلبات أفراد العينة، كما تتفق تماماً مع القدرات المتوقعة لأفراد العينة، مما يؤكد دقة الاختبار في عملية القياس والثقة في نتائجه، وقدرته على التمييز، وتحديد القدرات المختلفة لأفراد

العينة.

ويتضح أيضاً أنَّ منحنى دالة معلومات الاختبار متماثلاً بطريقة مقبولة، وذو ارتفاع واضح جداً في الوسط، وبالتالي فهو يأخذ الشكل الاعتدالي المتوقع لمثل هذه المنحنيات، ويشير شكل المنحنى إلى أنَّه يتم تقدير القدرة بأعظم دقة بمصاحبة متوسط مقياس القدرة، وبما أنَّ القدرة متغير متصل تكون المعلومات متغيراً مستمراً أيضاً، لهذا تحدث ذروة دالة معلومات الاختبار في نقطة أعلى بدرجة طفيفة من النقطة المتوسط لمقياس القدرة.

كما يلاحظ انخفاض في الخطأ المعياري أين كانت أقصى قيم للمعلومات عن مستوى قدرات الطلبة التي يمدنا بها اختبار الإحصاء، بدقة مرتفعة إلى حدٍ ما داخل مدى القدرة بين(-3.97 و4.07) لوجيت وهو ما كان متوقعاً، لأنَّ معظم مفردات الاختبار غطت هذا المدى، وخارج هذا المدى ينخفض مقدار المعلومات بسرعة ولا يتم تقدير مستويات القدرة بدرجة جيدة، مما يعني أنَّ الخطأ المعياري في التقدير يكون أقل ما يمكن عند مستويات القدرة التي تناظر أقصى المعلومات فأقل قيمة للخطأ المعياري كما نلاحظ في الجدول أعلاه يساوي(0.55) والذي يناظر أقصى قيمة للمعلومات التي يمنحها الاختبار وهي قيمة للقدرة تساوي(3.32) لوجيت عند الدرجة الخام(9) ومستوى القدرة(-0.18) وتنخفض قيمة المعلومات التي يمكن الحصول عليها من الاختبار بشكل واضح مع زيادة القدرة عن(4.07) أو تنخفض عن مستوى القدرة(-3.97) وهو ما يؤكد حاجة الاختبار إلى إضافة مفردات تغطي تلك المستويات من القدرة، فمقدار المعلومات التي تسهم بها مجموعة من المفردات عند مستوى قدرة معين يتناسب تناسباً عكسياً مع مربع سعة فترة الثقة حول تقديرات القدرة عند مستوى القدرة ($\theta$) أي أنَّ الخطأ المعياري للقياس يساوي $SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$ فبالتعويض في هذه المعادلة بالقيم السابقة، فإنَّ القيمة للخطأ المعياري للقياس تساوي تقريباً القيمة(0.55) والتي تناظر أقصى قيمة لدالة معلومات اختبار الإحصاء الوصفي.

**4. مناقشة نتائج التساؤل الرابع الذي ينص على:**

**- هل يمكن اشتقاق معايير تفسر تقديرات الأفراد على اختبار الإحصاء في صورته النهائية بعد تدريجه على الطلبة؟**

تمَّ تحويل تدريج وحدات اللوجيت لصعوبة مفردات الاختبار في صورته النهائية إلى تدرج وحدات مئوية بوحدة الواط، حيث تشتق وحدة اللوجيت من نموذج راش اللوغاريتمي، والذي يتناول التقدير الاحتمالي للاستجابة الصحيحة للفرد على المفردة ضمن التدريج الذي نقطة صفره متوسط صعوبات المفردات، ونظراً لوجود مشكلة الكسور، وقع اختيار الباحث على وحدة قياس الواط التي قدمها (Masters,1984) لتحويل تقديرات كل من صعوبات المفردات وقدرات الأفراد من وحدة القياس باللوجيت إلى هذه الوحدة الجديدة وهي وحدة الواط.

إضافة إلى ذلك استخرج الباحث معايير الدرجات التائية، وكذلك معايير الرتب المئينية المقابلة لكل درجة من الدرجات الخام التي حصل عليها الأفراد في عينة التقنين، وهذه المعايير هي الطريقة الكلاسيكية المألوفة، وقد قدمها الباحث لإجراء مقارنة بين المعايير المستخرجة وفق النظريتين الكلاسيكية في القياس، ونظرية الاستجابة للمفردة (نموذج راش).

وبعد تحليل النتائج يتضح أنَّ تقدير الصعوبة لمفردات الاختبار مقدرة بوحدة اللوجيت، وكذا بوحدة الواط، مرتبة تصاعدياً تبعاً لصعوبتها، والتي انحصرت فيها الدرجات الخام ما بين(53 و443) وكانت أقل قيمة من نصيب المفردة(20) والذي تقابله قيمة تقديرية للصعوبة بوحدة اللوجيت(3.50) وخطأ معياري قدره(0.17) والذي يقترب من الصفر، والدرجة(103) بوحدة الواط وبخطأ معياري قدره(2.53) في حين كانت أكبر قيمة من نصيب المفردة(19) والذي تقابله قيمة تقديرية للصعوبة بوحدة اللوجيت(-3.30) وخطأ معياري قدره(0.19) والذي يقترب من الصفر هو كذلك، والدرجة(1) بوحدة الواط وبخطأ معياري قدره(2.79).

ومن خلال توزيع العتبات الفارقة للمفردات والأفراد على عينة التدريج مفحوص اتضح أنَّ توزيع القدرات لمجموعة(التدريج) المعايرة، بأنَّه متماثلاً بطريقة مقبولة وذو ارتفاع واضح جداً في الوسط والذي يشير إلى أنَّه يتم تقدير الصعوبة والقدرة بأعظم دقة بمصاحبة متوسط مقياس القدرة والذي يستطيع أنَّ يأخذ أياً من الأشكال، حيث لا يفترض فيه أن يأخذ شكل توزيع المنحنى الطبيعي والذي يبين العتبات الفارقة للمفردات بدلاً من الصعوبة، والذي يظهر كذلك توزع عتبات مفردات اختبار الإحصاء، من أسهل المفردات(-3.30) لوجيت إلى الأصعب نوعاً ما(3.50) لوجيت. وتتدرج قدرات الأفـــراد على الاختبـــار من أقل قدرة حوالي (-3.97) لوجيت إلى أكبر قدرة حوالي (5.48)

لوجيت.

**5. منـــاقشة التســـاؤل الخـــامس الـــذي ينـــص علـــى:**

**- ما دلالات الخصائص السيكومترية (الصدق والثبات) لمفردات الاختبار بعد معايرة مفرداته باستخدام نموذج راش؟**

يتحقق صدق الاختبار بتحقق موضوعية تفسير نتائج الاختبار وموضوعية القياس بتحقق افتراضات نموذج راش على بيانات الاختبار بصيغته النهائية وذلك بعد حذف الأفراد غير المطابقين وحذف الفقرات غير المطابقة لافتراضات النموذج ويرى(Stenner,1990) أن التحقق من افتراضات النموذج يعتبر تحقيق لموضوعية القياس، كما تمَّ التحقق من صدق الاختبار من خلال صدق المحكمين من خلال الاجراءات التي اتبعت في تحديد مجال التحصيل في موضوعات الاحصاء الوصفي وتحليل المحتوى وكتابة الأهداف السلوكية وتمَّ التحقق من خلال الأحكام التي قدمها المحكمون المختصون عن مدى التوافق بين مضمون الفقرات والاهداف التي تحقها تلك الفقرات إضافة إلى صدق البناء الداخلي من خلال حساب معامل الارتباط بين فقرات الاختبار الذي كانت تساوي(0.310) بانحراف معياري قدره(5.66) وهي قيمة مقبولة.

ولإضفاء المزيد من المصداقية على الاختبار، قام الباحث باستخدام الأساليب التالية للتحقق من الصدق وهو التحقق من قطبية المفردات حيث أظهرت نتائج التحليل الموضحة في الجدول رقم(39) أنَّ نقطة قياس معاملات الارتباط لكل مفردات اختبار مقياس الإحصاء الوصفي كانت قيمها موجبة والتي تراوحت ما بين(0.40) للمفردة (19) والقيمة(0.64) للمفردتين(16 و18) مما يدل على أنَّ مفردات اختبار الإحصاء تقيس التكوين الفرضي الذي ينوي الباحث قياسه، أيضاً تمَّ التحقق من أحادية البعد حيث استخدم الباحث طريقة التقدير وفقاً لنموذج(راش) للمكونات الأساسية المعتمدة على البواقي، والذي تظهر الاختلافات بين الأبعاد، كما يوضحها الجدول رقم(29) حيث تحققت المحكات الثلاثة المعتمدة وهي المحك الأول وكقاعدة عامة إذا كانت قيمة التباين المفسر بواسطة التقديرات أكبر من أو يساوي(50%) واسترشاداً بما أشار إليه لينكر(Linacre) مطور برنامج (Winsteps)، الذي حدد المجال من(20%) إلى (80%) بأنَّه مجال مقبول ويعد مؤشراً قوياً لأحادية البعد، وهو ما تحقق في بحثنا الحالي بقيمة(48.5%) أما المحك الثاني، هو أنَّ نسبة التباين الذي يفسره العامل الثاني يجب أن

تكون قيمته أقل أو قريب من(5%) وهذا ما تحقق في البحث الحالي تقريباً بالقيمة(5.8%) وهذا دليل آخر على أحادية البعد. في حين يعتمد المحك الثالث، على قيمة الجذر الكامن لنسبة التباين الذي يفسره العامل الثاني والذي يجب ألا يتعدى أو يكون أقل قيمة له (3) وهذا ما تحقق في البحث الحالي بتسجيل القيمة (2.1284) وهي أقل من القيمة (3) المنصوح بها من لينكر(Linacre, 2008,p.272)

أما ثبات الاختبار فبعد التحقق من جودة مطابقة فقرات الاختبار النموذج (راش)، واستبعاد الفقرات التي لم تحقق هذه المطابقة، وكذلك الأفراد غير المطابقين للنموذج، وتحديد القيم التقديرية لكل من صعوبة الفقرات وقدرات الأفراد، تمَّ الحصول على نوعين من المعاملات، هما الثبات المتعلق بالأفراد والثبات المتعلق بفقرات الاختبار. فمفهوم الثبات في إطار نموذج (راش)، يشير إلى مدى الدقة في تقدير موقع كل من الفقرات والأفراد على متصل السمة، الذي تهدف إلى قياسه.

ويمكن تحديد مدى دقة الفقرات في تعريف هذا المتصل، بإيجاد النسبة بين الانحراف المعياري للقيم التقديرية المتحررة لصعوبة الفقرات، ومتوسط الخطأ المعياري لهذه القيم. ويطلق على هذه النسبة معامل الفصل بين الفقرات(G) والذي يجب أن تكون قيمته أكبر ما يمكن. فإذا لم تزد هذه القيمة عن (2) فإنَّه يصعب قياس المتغير بوساطة هذه الفقرات، وقد وجد الباحث آنَّ قيمة معامل الفصل بين الفقرات في الاختبار بصورته النهائية (12.57). وبما أنَّ هذه القيمة تزيد عن (2) فإنَّ فقرات الاختبار تعتبر كافية لتعريف متصل السمة الذي تقيسه. وكذلك وجد الباحث أنَّ قيمة معامل الفصل بين الأفراد (Gp) تساوي(2.53) وترمز إلى النسبة بين الانحراف المعياري للقيم التقديرية المتحررة لقدرات الأفراد ومتوسط الخطأ المعياري لهذه القيم، ونظراً لأنَّ هذه القيمة أيضاً تزيد عن(2) فإنَّ عينة الأفراد تعد كافية للفصل بين الفقرات، وتمَّ حساب معامل الثبات المتعلق بالأفراد فكان يساوي(0.87) وهي قيمة مقبولة وتدل على أنَّ عينة الأفراد كافية للفصل بين الفقرات والتمييز بين مستويات الفقرات أي أنَّ الأفراد يتوزعون بشكل مناسب على متصل السمة الممثل بفقرات الاختبار التي بُنيت وفق نموذج راش وتمَّ حساب معامل الثبات للفقرات فكان يساوي(0.99) عند المستويات المختلفة للقدرة، وهي قيمة مرتفعة جداً وتدل على أنَّ عينة الفقرات كافية للفصل بين الأفراد وبالتالي التمييز بين مستويات التحصيل لدى هؤلاء الأفراد.

كما حدد الباحث عدد الطبقات الإحصائية المتمايزة لكل من المفردات والأفراد باستخدام الصيغة

الرياضية المشروحة سابقاً وبالتعويض في المعادلة بالنسبة للأفراد وُجِد أنَّ النتيجة تساوي(4.71) تقريباً خمس طبقات إحصائية متمايزة للأفراد أي خمسة مستويات للتحصيل لدى هؤلاء وقد يعزى ذلك إلى قصر الاختبار بصورته النهائية، في حين أنَّ قيمة معامل الثبات المتعلق بالفقرات كانت كبيرة جداً وبتطبيق المعادلة فقد تبين أنَّ الفقرات كانت قادرة على إظهار ثمانية عشر طبقة إحصائية متمايزة تقريباً لصعوبة الفقرات لدى أفراد عينة الدراسة أي إظهار الفروق الفردية بين الأفراد بشكل كبير، وقد يعزى ذلك إلى جودة الفقرات المكونة للاختيار بصورته النهائية.

وبعد أن تمَّ بناء الاختبار عبر هذه المراحل من استبعاد الفقرات غير المطابقة لنموذج راش واستبعاد الأفراد غير المطابقين، وبعد تحديد الخصائص السيكومترية للاختبار وفقراته، فإنَّ هذا الاختبار المؤلف من (19) فقرة يمثل المجال بشكل متوازن مع الأهداف.

وتتفق هذه النتيجة مع دراسات كل من الطراونة (2016)، العنزي (2014)، هادي ومراد (2013) السامرائي والخفاجي (2012)، على (2012)، الخياط (2012)، الجبوري (2012).

**الخلاصة**

سعت هذه الدراسة إلى بناء اختبار تحصيلي مرجعي المحك لقياس كفايات الطلبة في مقياس الاحصاء الوصفي لطلبة السنة الأولى علوم اجتماعية جذع مشترك، والتحقق من خصائصه السيكومترية، وكذا التحقق من درجة مطابقة الاختبار لافتراضات نموذج راش احادي المعلمة.

وتوصلت الدراسة الحالية إلى أنَّ الاختبار التحصيلي مرجعي المحك الذي تمَّ بناءه يتوفر على خصائص الاختبار الجيد، حيث اتصفت فقرات الاختبار بمعاملات صعوبة جيدة ومعاملات تمييز مقبولة، وكانت فقراته مطابقة لنماذج الاستجابة للمفردة، وكذا توفر دلالات صدق وثبات جيدتين.

وكان الهدف من هذه الدراسة هو التغلب على مشكلات النظرية الكلاسيكية إذ أنَّ الاختبارات التحصيلية التي تُصمم في إطار النظرية الحديثة تكون على درجة عالية من الموضوعية ويمكن تعميم نتائجها بعكس النظرية الكلاسيكية التي لاقت العديد من الانتقادات لضعف افتراضاتها، ويعد نموذج راش أحادي المعلمة من أفضل النتائج استخداماً في بناء الاختبارات والمقاييس النفسية والتربوية وتطورها وهو ما تمَّ تبنيه في هذه الدراسة حيث تمَّ بناء اختبار تحصيلي وفق نموذج راش.

واعتماداً على النتائج التي توصلت اليها الدراسة فإنَّ الباحث يوصي بما يلي:

- ✓ إجراء دراسات مماثلة في مقررات دراسية مختلفة.
- ✓ استخدام نماذج الاستجابة للمفردة المختلفة في بناء وتطوير الاختبارات والمقاييس المختلفة.
- ✓ برمجة دورات تدريبية في كيفية استخدام نماذج النظرية الحديثة والبرامج الخاصة بها.
- ✓ إعداد اختبارات محكية المرجع لجميع مجالات الاحصاء.
- ✓ إجراء دراسة لإيجاد استراتيجية لبناء الاختبارات محكية المرجع قائمة على طرق احصائية تحليلية عوضاً عن الطرق السائدة والمعتمدة على تحكيم الخبراء.
- ✓ تدريب الأساتذة لإكسابهم مهارة إعداد الاختبارات التحصيلية محكية المرجع واستخدام نتائجها لتحديد نقاط الضعف ومعالجتها.
- ✓ إعداد بنوك اسئلة ذات مواصفات جيدة يمكن استخدامها عند الحاجة.

# المـــــراجـــع

**المراجع باللغة العربية**


- إبراهيم، محمود محمد. (1991). اتجاهات معاصرة في القياس النفسي والتربوي. *رسالة التربية، دائرة البحوث التربوية، 1(18)*، 149-133.
- أبو حطب، فؤاد، السيد أحمد عثمان، صادق، أمال. (2008). *التقويم النفسي* (ط.4). المكتبة الأنجلو المصرية.
- أبو دينار، مسعد.(2012). *القياس والتقويم لذوي صعوبات التعلم*. مركز تقويم وتعليم الطفل.
- أبو عواد، محمد عثمان فريال.(2006). *تطوير اختبار تشخيصي محكي المرجع للكشف عن الأخطاء التي يقع فيها طلبة الصفوف الخامس والسادس والسابع في مادة الرياضيات في الأردن* [رسالة دكتوراه غير منشورة]. الجامعة الأردنية.
- أبو هاشم، السيد.(2009). *الدليل الاحصائي في تحليل البيانات باستخدام SPSS*. مكتبة الرشد للنشر والتوزيع.
- إسماعيل، السيد أحمد ميمي.(2007). *الخصائص السيكو مترية لاختبار القدرة العقلية باستخدام نموذج راش لدى طلبة المرحلة الثانوية العامة* [أطروحة ماجستير غير منشورة]. جامعة الزقازيق.
- إسماعيل، بشرى.(2004). *المرجع في القياس النفسي*. القاهرة: مكتبة الأنجلو المصرية.
- أمطانيوس، نايف مخائيل.(2015). *القياس والتقويم النفسي والتربوي للأسوياء ودوي الحاجات الخاصة*. دار الإعصار العلمي للنشر والتوزيع.
- الأنصاري، بدر محمد.(2000). *قياس الشخصية*. دار الكتاب الحديث.
- أيكن، لويس.(2007). *الاختبارات والامتحانات قياس القدرات والأداء*. (فرج السراج، ترجمة) مكتبة العبيكان.(1998).
- بركات، زياد.(2018). *القياس والتقويم النفسي والتربوي بين النظرية والتطبيق*. جامعة القدس المفتوحة.
- البكري، أمل وعجوز، ناديا.(2011). *علم النفس المدرسي*. المعتز للنشر والتوزيع.
- البناء، مأمون علي ناجي.(2011). *بناء اختبار محكي المرجع لقياس الكفايات الإحصائية لدى طلبة الدراسات العليا بكليات التربية في الجامعات اليمنية* [أطروحة ماجستير غير منشورة]. جامعة الملك سعود.
- تايلر، ليونا.(1988). *الاختبارات والمقاييس* (ط.2). (عبد الرحمن سعد، مترجم). دار الشروق للنشر والتوزيع.(1971).
- التقي، أحمد محمد.(2009). *النظرية الحديثة في القياس*. دار المسيرة.

- ثورندايك، روبرت وهيجن، اليزابيت.(1986). *القياس والتقويم في علم النفس والتربية* (الكيلاني عبد الله زيد، عدس عبد الرحمن، مترجم)( ط.4). مركز الكتب الأردني.
- الجبلي، شاكر سوسن.(2005). *أساسيات بناء الاختبارات والمقاييس النفسية والتربوية*. مؤسسة علاء الدين للطباعة والنشر.
- حامد، صبحي سيرين.(2008). *أثر نموذج الاستجابة للفقرة وتعدد الأبعاد وطريقة المطابقة في تقدير معالم الأفراد والفقرات* [رسالة دكتوراه غير منشورة]. جامعة عمان.
- الحسن، عدنان أسماء.(2012). *بناء اختبار محكي المرجع لقياس درجة إتقان المعلمين لقواعد بناء الاختبارات التحصيلية* [رسالة ماجستير غير منشورة]. جامعة دمشق.
- حمَّاد، فهمي علي ديانا.(2010). *علاقة الحكمة الاختبارية بالأداء الناتج من اختبار تحصيلي ذو اختيار من متعدد مبني وفق نموذج راش لدى طالبات كلية التربية للبنات الأقسام الأدبية بجامعة أم القرى* [رسالة دكتوراه غير منشورة]. جامعة أم القرى.
- حمدي، عبد الله عبد العظيم.(2013). *موسوعة الاختبارات والمقاييس*. مكتبة أولاد الشيخ للتراث.
- حيدر، عبد الكريم الزهيري.(2017). *مناهج البحث التربوي*. مركز ديبونو لتعليم التفكير.
- الخطيب، مصطفى فريد.(1993). *مبادئ القياس والتقويم في التربية*. دار الفكر للطباعة والنشر.
- الخولي، محمد علي.(1998). *الاختبارات التحصيلية اعدادها وإجراؤها وتحليلها*. دار الفلاح للنشر والتوزيع.
- دودين، محمد حمزة.(2004). *استخدام نظرية الاستجابة على الفقرة في تحليل الاستجابات على اختبار المهارات الكمية لدى طلبة الجامعة*، المجلة التربوية، 18(72)، 125-105.
- الدوسري، إبراهيم.(2001). الإطار المرجعي للتقويم التربوي (ط.3). مكتب الدول العربية لدول الخليج.
- الدوسري، راشد، حماد.(2004). *القياس والتقويم التربوي الحديث مبادئ وتطبيقات وقضايا معاصرة*. دار الفكر للنشر والتوزيع.
- دويدري، رجاء وحيد.(2000). *البحث العلمي أساسياته النظرية وممارساته العلمية*. دار الفكر المعاصر
- ذيب، عبد الكريم إيمان.(ب. ت). أثر طول الاختبار على تحديد درجة القطع لاختبار تحصيلي محكي المرجع في مادة علم نفس الشخصية. *مجلة مداد الأدب، 6(1)*، 408-330.
- الزغلول، عبد الرحيم عماد.(2012). *علم النفس التربوي* (ط.2). دار الكتاب الجامعي.
- زكري، علي بن محمد.(2007). *الخصائص السيكو مترية لاختبار (أوتيس-لينون) للقدرة العقلية مقدرة*

*وفق القياس الكلاسيكي ونموذج راش لدى طلبة المرحلة المتوسطة بمحافظة صبيا التعلمية* [رسالة دكتوراه غير منشورة]. جامعة أم القرى.

- الزوبعي، عبد الجليل إبراهيم وبكر، محمد الياس والكناني، عبد المحسن.(1981). *الاختبارات والمقاييس النفسية*. الجامعة الموصلية.
- الزيلعي، محمد بن علي بن عمر.(2015). *بناء اختبار تشخيصي محكي المرجع لقياس مهارات البحث العلمي لدى طلاب الدراسات العليا*. [أطروحة ماجستير غير منشورة]. جامعة أم القرى.
- شحاتة، محمد ربيع.(2011). *قياس الشخصية* (ط.3). دار المسيرة للنشر والتوزيع.
- الشرقاوي، أنور وكاظم، أمينة والشيخ، سليمان وعبد السلام، نادية.(1996). *اتجاهات معاصرة في القياس والتقويم النفسي والتربوي*. مكتبة الأنجلو مصرية.
- شكري، محمد سومية.(2006). *فاعلية بعض نرق تقدير درجة القطع في التنبؤ بالتحصيل اللاحق في الهندسة للمرحلة الإعدادية* [أطروحة ماجستير غير منشورة]. جامعة المنيا.
- صلاح الدين، محمود علام. (2010). *القياس والتقويم التربوي في العملية التدريسية* (ط.3). دار المسيرة للنشر والتوزيع.
- صلاح الدين، محمود علام.(1982). *استخدام نموذج السمات الكامنة في بناء بنوك أسئلة*. صحيفة التربية، (5)، 344-381.
- صلاح الدين، محمود علام.(1986). *تطورات معاصرة في القياس النفسي والتربوي*. مطابع القبس التجارية.
- صلاح الدين، محمود علام.(1995). *الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريسية*. دار الفكر العربي.
- صلاح الدين، محمود علام.(2000). *القياس والتقويم التربوي والنفسي اساسياته وتطبيقاته وتوجهاته المعاصرة*. دار الفكر العربي.
- صلاح الدين، محمود علام.(2001). *الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية*. دار الفكر العربي.
- صلاح الدين، محمود علام.(2005). نماذج الاستجابة للمفردة الاختبارية. دار الفكر العربي.
- صلاح الدين، محمود علام.(2006). *القياس والتقويم التربوي والنفسي أساسياته وتوجهاته المعاصرة*. دار الفكر العربي.

- صلاح الدين، محمود علام.(2007). *الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية* (ط.2). دار الفكر العربي.
- صلاح الدين، محمود علام.(2015). *القياس والتقويم النفسي والتربوي، أساسياته وتطبيقاته وتوجيهاته المعاصرة*. دار الفكر العربي.
- صلاح، أحمد مراد وسليمان، علي أمين.(2005). *الاختبارات والمقاييس في العلوم النفسية والتربوية* (ط.2). دار الكتاب الحديث.
- صلاح، أحمد مراد وسليمان، علي، أمين. (2002). *الاختبارات والمقاييس في العلوم النفسية والتربوية، خطوات إعدادها وخصائصها*. دار الكتاب الحديث.
- الطريري، عبد الرحمن بن سلمان.(1997). *القياس النفسي والتربوي*. مكتبة الرشد للنشر والتوزيع.
- الطريري، عبد الرحمن بن سليمان.(2014). *القياس النفسي والتربوي نظريته، أسسه، تطبيقاته* (ط.2). مكتبة الرشد للنشر والتوزيع.
- طوالبة، علي عبد الرحيم.(2018). *الكشف عن أنماط استجابة الفرد غير المطابقة لنموذج الاستجابة للفقرة في الاختبارات مختلطة الشكل في ضوء بعض المتغيرات* [رسالة دكتوراه غير منشورة]. جامعة اليرموك.
- عبابنة، عماد.(2009). *الاختبارات محكية المرجع فليفتها وأسس تكويرها*. دار المسيرة للنشر والتوزيع.
- عبد الخالق، أحمد محمد.(1996). *قياس الشخصية*. جامعة الكويت.
- عبد الرحمن، سعد.(1998). *القياس النفسي، النظرية والتطبيق* (ط.3). دار الفكر العربي.
- عبد الرحمن، سعد.(2008). *القياس النفسي النظرية والتطبيق*. (ط.5). هبة النيل العربية للنشر والتوزيع.
- عبد الرؤوف، طارق وإيهاب، عيسى.(2017). *المقاييس والاختبارات، التصميم، الإعداد، التنظيم*. المجموعة العربية للتدريب والنشر.
- عبد المحسن، علي صلاح.(2019). *تعلم الاحصاء*. كاستر النشر والتوزيع.
- العجمي، عبد الهادي شاجع عبد الله.(2011). *مدى التباين في تقديرات المحكمين لدرجة القطع لاختبار محكي المرجع في اللغة العربية بناء على شكل الفقرة* [رسالة دكتوراه غير منشورة]. جامعة عمَّان العربية للدراسات العليا.
- عصفور، خلود رحيم.(2007). *تشخيص كفايات التعلم لطلبة كلية التربية باستخدام نموذج راش والطريقة التقليدية* [رسالة دكتوراه غير منشورة]. جامعة بغداد.
- عطوان، أسعد حسين وأبو شعبان، صبحي شيماء.(2019). *القياس والتقويم التربوي*. دار الكتب العلمية.

- العفون، نادية حسن وجليل، ماهر وسن.(2013). *التعلم المعرفي واستراتيجية معالجة المعلومات*. دار المناهج النشر والتوزيع.
- عماد، عبد المسيح يوسف.(1991). استخدام نموذج راش اللوغاريتمي أحادي البارامتر في تحليل مفردات الاختبارات المعرفية مرجعية المعيار ثنائية القطب (دراسة تجريبية). *مجلة البحث في التربية وعلم النفس، 4(4)* 443-473.
- عودة، أحمد.(1992). مدى توافق بين نموذج راش والمؤشرات التقليدية في اختبار فقرات مقياس اتجاه سباعي التدريج. *مجلة كلية التربية، 7(8)*، 179-152.
- عودة، أحمد.(2005). *القياس والتقويم في العملية التدريسية* (ط.3). دار الامل للنشر والتوزيع.
- عوض الله، محمد عبد الرحيم.(2000). *مقارنة بين أسلوبي راش والطريقة التقليدية في بناء اختبارات الذكاء باستخدام محك التنبؤ بالتحصيل الدراسي* [اطروحة دكتوراه غير منشورة]. جامعة بغداد.
- فرج، المبروك وعامر، عمر.(2016). *التقويم والقياس التربوي الحديث*. دار حميثرا للنشر والترجمة.
- فرج، صفوت.(2007). *القياس النفسي* (ط.6). مكتبة الأنجلو المصرية.
- فيصل، عباس.(1996). *الاختبارات النفسية تقنياتها وإجراءاتها*. دار الفكر العربي.
- كاظم، أمينة محمد.(1988أ). *استخدام نموذج راش في بناء اختبار تحصيلي في علم النفس وتحقيق التفسير الموضوعي للنتائج*. مطبوعات جامعة الكويت.
- كاظم، أمينة محمد.(1988ب). دراسة نظرية نقدية جول القياس الموضوعي للسلوك نموذج راش. مؤسسة الكويت للتقدم العلمي.
- كروكر، ليندا والجينا، جيمز.(2009). *مدخل إلى نظرية القياس التقليدية والمعاصرة*.(الحموري هند، زينات دعنا، ترجمة). دار الفكر ناشرون وموزعون.
- الكيلاني، عبد الله وعدس، عبد الرحمن. (2011). *القياس والتقويم في التربية وعلم النفس*. مركز الكتاب الأردني.
- مجيد، شاكر سوسن.(2014). *أسس بناء الاختبارات والمقاييس النفسية والتربوية*. (ط.3).مركز ديبونو لتعليم التفكير.
- محاسنة، إبراهيم محمد.(2013). *القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة*. دار جرير للنشر والتوزيع.
- محمود، شكري، محمد سومية.(2012). *استخدام نموذج راش في بناء بنك أسئلة لقياس التحصيل في مقرر*

*سيكولوجية التعلم لدى طلاب كلية التربية* [رسالة دكتوراه غير منشورة]. جامعة المنية.

- النبهان، موسى.(2004). *أساسيات القياس في العلوم السلوكية*. دار الشروق للنشر والتوزيع.
- الوليلي، إسماعيل.(2002). *دراسة سيكو مترية مقارنة لبعض نماذج الاستجابة للمفردة في انتقاء مفردات الاختبارات مرجعية المحك* [أطروحة دكتوراه غير منشورة]. جامعة الأزهر.
- يعقوب، عيسى ابراهيم محمد.(1990). *دراسة مقارنة للخصائص السيكومترية لمقياس مفهوم الذات المبني بالطريقة التقليدية وطريقة نموذج راش* [أطروحة دكتوراه غير منشورة]. الجامعة الأردنية.
- اليعقوبي، حيدر.(2013). *القياس والتقويم في العلوم التربوية والنفسية رؤيا تطبيقية*. مركز المرتضى للتنمية الاجتماعية.

**المراجع باللغة الأجنبية**

- Alastair, P., Hutchinson, C. (1987). Calibrating graded assessments: Rasch partial credit analysis of performance in writing. *SAGE Journals, 4(1),* 72-92. https://doi.org/10.1177/026553228700400107
- Bachman, L. (1991). *Fundamental considerations in language testing*. Oxford univ. Press.
- Baker, F. (2001). *The Basics of item response theory,2nd Ed*. Eric clearing House on Assessment and Evaluation. university of Mary Land. college Park, MD.
- Bejar, I. (1983). *Achievement testing, Recent Advances*. Sage Publication Inc.
- Berk, R. (1980). *Criterion-Referenced Measurement. The state of the Art*. The John Hopkins University Press.
- Berk, R. (1986). A Consumer's guide to Setting Performance Standards on Criterion-Referenced tests. *Review of educational research, 56(1),* 137-172. https://doi.org/10.2307/1170289
- Boeck, P., & Wilson, M. (2006). *Explanatory Item Response Models a Generalized Linear and Nonlinear Approach*. statistics for Social Science and Behavioral. Publisher.
- Crislip, M. A., & Chin-Chance, S. (2001). *Using traditional psychometric methodologies and the Rasch model in designing a test*, paper presented at the annual meeting of the American Educational Research Association. Seattle, WA.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Cengage Learning, Inc.
- Ebel, R. (1980). *Practical Problems in Education Measurement*. D.C Health &company.
- Fischer, G. (1997). *Handbook of Modern Item Response Theory*. Van Der Linden Etal: Springer.
- Gregory, J, .(1996). An NCME Instructional Module on Setting Passing Scores. *EducationalMeasurement,15(2)*20-31.https://doi.org/10.1111/j.1745-3992.1996.tb00809.x

- Gregory, J. (2004). *Psychological testing: History, Principal*. 4thed. Pearson Education Group Inc.
- Hambleton, R, K., & cook, L. (1977). Latent trait models and their use in the analysis of Education Test data. *Journal of Education Measurement, 14(02)*, 75-96.
- Hambleton, R, K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Kluwer-Nijhoff Publishing Co.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Response Theory*. SAGE Publications, Inc.
- Hambleton, R.K., Zaal, J. (1991). *Advances in Educational and Psychological Testing*. Springer Netherlands.
- Hambleton. R, K., & Swaminathan, H. (2013). Item Response Theory, Principles and Application. *Springer & Business Media*. https://doi.org/10.1111/j.1745-3984.1977.tb00030.x
- Hullin, C. & Drasgow, F., & Parsons, C. (1983). *Item Response Theory: Application to Psychological Measurement*. Dorsey edition.
- Linacre, J. M. (2008). Winsteps (version 3.92.1). [*computer software*] Retrieved from http://www.Winsteps.com.
- Linacre, J. M. (2012). *A User's Guide to Winsteps Ministep Rasch- Model Computer Programs*. https://www.winsteps.com/manuals.htm.
- Linden Van der, W.J., & Hambleton, R.K. (1997). *Handbook of modern Item Response Theory*. Springer-Verlag.
- Linden, W.J., & Hambleton, R.K. (1997). *Handbook of modern Item Response Theory*. Springer Science.
- Livingston, S,. Zieky,M.(1983). A Comparative study of standard-setting methods. *ETS. 1983(2)*, i-48.https://doi.org/10.1002/j.2330-8516.1983.tb00038.x
- Loken, E., & Rulision, K.l. (2010). Estimation of a 4 Parameter Item Response Theory model. *The Britisg Journal of Mathematical and Statistical Psycholody, 63(3),* 509-525. https://doi.org/10.1348/000711009X474502
- Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Lawrence Erlbaum associates.
- Masters G.N. (1988). Item discrimination: when more is worse. *Journal of Educational Measurement 25(1)*, 15-29. https://doi.org/10.1111/j.1745-3984.1988.tb00288.x
- Masters, G. (1984). DICOT Analyzing classroom test with the Rasch model. *SAGE journal,44(01),* 145-150. https://doi.org/10.1177/0013164484441015
- Sax, G., & Newton, J.W. (1997). *Principe of Educational and Psychological Measurement and Evaluation*. (4thEd). Belmont, CAM: Wadsworth Publishing.
- Schumacher, R. (2006). *Latent trait reliability using Rasch measurement models*. paper presented at the Annual meeting of the south west. education research Association.
- Stephenson, A. S, Patricia, B., & John. A. E. (2000). Standard setting Technique: *An application for counselling programs. Measurement and Evaluation in Counseling and Development, 32(4)*,229-244. doi:10.1080/07481756.2000.12068989

- Steven, P., R, Niels G. W. (2003). How Many IRT Parameters Does It Take to Model, *Psychological Methods,8(2),*164-184. https://doi.org/10.1037/1082-989X.8.2.164
- Szabo, G. (2008). *Applying item response theory in language test item bank building*. Peter Lang GmbH.
- Wiberg, M. (2004). *Classical Test Theory vs. Item Response Theory An evaluation of the theory test in the Swedish driving-license test*. Umea University.
- William j,. John,R. & Melissa,.S. (2014). *Rasch Analysis in the Human Sciences*. Springer Dordrecht.
- Wright, B. (1977). Misunderstanding the Rasch model, *JEM, 14(03),* 219-225. https://doi.org/10.1111/j.1745-3984.1977.tb00039.x
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8(3),* 370-371
- Wright, B. D., & Masters, G. N. (1982). *Rating scale. Analysis*. MESA Press.
- Wright, B., & Panchapakesan, N. (1969). Procedures for sample-free Item analysis, *SAGE Journal, 29(1),* 23-48. https://doi.org/10.1177/001316446902900102
- Wright, B., & Stone, M. (1979). *Best Test Design. Rasch Measurement*. MESA Press.
- Yen, M., & Edwardson, S.R. (1999). Item Response Theory Approach in Scale Development *Nursing Research. 48(4),* 234-238. DOI: 10.1097/00006199-199907000-00008.
- Zenisky, A. L., Hambleton, R. K., & Sireci, S. G. (2002). Identification and evaluation of local item dependencies in the Medical College Admissions Test. *Journal of Educational Measurement, 39(4),* 291–309. https://doi.org/10.1111/j.1745-3984.2002.tb01144.x

# الملاحق

الملحق رقم (01)

# البرنامج المفصل لكل مادة

(تقديم *بطاقة مفصلة لكل مادة*)

**الوحدة: وحدات التعليم المنهجية**

**المادة: الإحصاء الوصفي**

**أهداف التعليم:**

(ذكر *ما يفترض على الطالب اكتسابه من مؤهلات بعد نجاحه في هذه المادة، في ثلاثة أسطر على الأكثر*) **تأتي هده المادة كدليل ومرشد للطلبة في الكليات والجامعات في التعريف بالأسس العلمية التي ينبغي إتباعها أو مراعاتها خلال إجراء البحث العلمي**

**المعارف المسبقة المطلوبة:**

(وصف *تفصيلي للمعرفة المطلوبة والتي تمكن الطالب من مواصلة هذا التعليم، سطرين على الأكثر)*

**محتوى المادة:**

**_علم الإحصاء وعلاقته بالعلوم الأخرى**

**_دور الإحصاء في العلوم الاجتماعية**

**_تنظيم البيانات**

**_مقاييس النزعة المركزية (المتوسط الحسابي ــ الوسيط ــ المنوال)**

**_ مقاييس التشتت (الانحراف عن المتوسط ــ المدى ــ التباين ــ الانحراف المعياري)**

**_ تحويل درجات الخام (الدرجة المعيارية ــ الدرجة الثنائية _ المئينيات ).**

الملحق رقم (02)

**نموذج استمارة تحكيم الأهداف السلوكية والبنود الاختبارية المقابلة لها والمعدة كاختبار من متعدد لقرر الاحصاء الوصفي**

السيد (ة) الدكتور(ة):

تحية طيبة وبعد

يقوم الباحث ببناء اختبار محكي المرجع في مقرر الاحصاء الوصفي للسنة الاولى علوم اجتماعية كجزء من موضوع بحث يعده للحصول على درجة الدكتوراه في القياس النفسي والتقويم التربوي وهو بعنوان (استخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع في مقرر الاحصاء للسنة الأولى علوم اجتماعية).

وقد قام الباحث بإعداد بنود اختبارية من نوع الاختيار من متعدد وبأربعة بدائل لقياس كل هدف من الأهداف السلوكية المحددة في المجال المعرفي وفق تصنيف بلوم مقابلاً للبنود الاختيارية التي أعدت لقياس هذا الهدف في الاختبار محكي المرجع.

لذا يرجو الباحث من سيادتكم التفضل بتحكيم العمل في بناء البنود الاختبارية والتكرم بالإجابة عن الاسئلة التالية:

1- هل الهدف السلوكي سليم وواضح لغوياُ؟

2- هل البند الاختباري سليم وواضح لغوياً؟

3- هل البند الاختياري يقيس الهدف الذي وضع لقياسه؟

يرجى التكرم بتقديم أية ملاحظة حول بناء الاختبار شاكراً تعاونكم ومساعدتكم في انجاز هذا البحث العلمي

| الكفاية | الدروس | الأهداف السلوكية | التذكر | الفهم | التطبيق | التحليل | التركيب | التقويم |
|---|---|---|---|---|---|---|---|---|
| المفاهيم الأساسية لعلم الاحصاء | التعريف بالإحصاء | يعدد أهمية الإحصاء | x | | | | | |
| | | يعطي تعريفاً للإحصاء | | x | | | | |
| | | يحدد وظائف علم الاحصاء | | | | x | | |
| | | يلخص علاقة الإحصاء بالعلوم الاجتماعية والانسانية | | x | | | | |
| | | يفرق بين أنواع المتغيرات وطرق قياسها | | | | x | | |
| مصادر وطرق جمع البيانات | مصادر وأساليب جمع البيانات | يلم بمصادر جمع البيانات الاحصائية | x | | | | | |
| | | يختار أسلوب جمع البيانات المناسب | x | | | | | |
| | | يميز بين مصادر جمع المعلومات | | | | x | | |
| | | يحدد طرق اختيار كل عينة | | | | x | | |
| | | يميز بين أنواع العينات | | | | x | | |
| | | يعدد أنواع العينات | x | | | | | |
| | | يبين أخطاء المعاينة | | x | | | | |
| عرض البيانات الاحصائية | تصنيف البيانات وتبويبها | يشرح مفهوم تبويب البيانات | x | | | | | |
| | | يعرف آليات التعامل مع البيانات المجمعة | x | | | | | |
| | | يصمم جداول التوزيع التكراري | | | | | x | |
| | | ينظم البيانات في جدول تكراري | | | | | x | |
| | العرض البياني للبيانات | يمثل بيانات الجداول التكرارية المختلفة بيانياً (أعمدة- دوائر- مدرج... | | | | | x | |
| | | يفرق بين أنواع الجداول التكرارية (البسيطة،المركبة) | | | | x | | |
| | | يستخرج طول الفئة | | | | x | | |
| مقاييس النزعة المركزية | مفهوم النزعة | يعرف مفهوم النزعة المركزية | x | | | | | |
| | | يقارن بين حالات استخدام مقاييس النزعة المركزية | | | | x | | |
| | المنوال | يعرف المنوال | x | | | | | |
| | | يحدد خصائص المنوال | | | | x | | |
| | | يحسب المنوال للبيانات المبوبة وغير المبوبة | | | x | | | |
| | | يحسب المنوال بالطريقة البيانية | | | x | | | |
| | الوسط الحسابي الوسيط | يعرف الوسط الحسابي | | | | x | | |
| | | يستخرج خصائص الوسط الحسابي | | | | x | | |
| | | يحسب الوسط الحسابي للبيانات المبوبة وغير المبوبة | | | x | | | |
| | | يوظف المتوسط الحسابي في الحالات المناسبة | | | x | | | |
| | | يعرف الوسيط | x | | | | | |
| | | يستخرج خصائص الوسيط | | | | x | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | يحسب الوسيط للبيانات المبوبة وغير المبوبة | | | | x | | |
| | | يحسب الوسيط بالطريقة البيانية | | | | x | | |
| | | يوظف العلاقة بين مقاييس النزعة المركزية | | | | x | | |
| مقاييس التشتت | المدى | يعرف المدى | x | | | | | |
| | | يحسب المدى المطلق للبيانات المبوبة وغير المبوبة | | | | x | | |
| | | يحسب المدى النسبي | | | | x | | |
| | | يعدد ايجابيات ومساوئ المدى | | | | | | x |
| | المدى الربيعي | يعرف مفهوم المدى ونصف المدى الربيعي | x | | | | | |
| | | يحسب المدى الربيعي للبيانات المبوبة وغير المبوبة | | | | x | | |
| | | يحسب نصف المدى الربيعي | | | | x | | |
| | الانحراف المتوسط | يعرف الانحراف المتوسط | x | | | | | |
| | | يحسب الانحراف المتوسط للبيانات المبوبة وغير المبوبة | | | | x | | |
| | الانحراف المعياري | يعرف الانحراف المعياري | x | | | | | |
| | | يحسب الانحراف المعياري للبيانات المبوبة وغير المبوبة | | | | x | | |
| | | يحدد مزايا وعيوب الانحراف المعياري | | | | | | x |
| | | يوضح ضرورة وأهمية مقاييس التشتت | | x | | | | |
| | | يفرق بين استخدامات مقاييس النزعة المركزية ومقاييس التشتت | | | x | | | |
| تحويل الدرجات الخام | أنواع الدرجات المعيارية | يعرف مفهوم الدرجة المعيارية | x | | | | | |
| | | يحسب الدرجة المعيارية | | | | x | | |
| | | يعرف مفهوم الدرجة التائية | x | | | | | |
| | | يحسب الدرجة التائية | | | | x | | |
| | | يعرف مفهوم المئينيات | x | | | | | |
| | | يحسب المئينيات | | | | x | | |

ملحق رقم (03)

استمارة تقدير المحكمين لمدى ارتباط الهدف السلوكي الاجرائي بالهدف الرئيسي

| الكفاية | الدروس | الأهداف السلوكية | مرتبط | غير مرتبط |
|---|---|---|---|---|
| المفاهيم الأساسية لعلم الاحصاء | التعريف بالإحصاء | يعدد أهمية الإحصاء | | |
| | | يعطي تعريفاً للإحصاء | | |
| | | يحدد وظائف علم الاحصاء | | |
| | | يلخص علاقة الإحصاء بالعلوم الاجتماعية والانسانية | | |
| | | يفرق بين أنواع المتغيرات وطرق قياسها | | |
| مصادر وطرق جمع البيانات | مصادر وأساليب جمع البيانات | يلم بمصادر جمع البيانات الاحصائية | | |
| | | يختار أسلوب جمع البيانات المناسب | | |
| | | يميز بين مصادر جمع المعلومات | | |
| | | يحدد طرق اختيار كل عينة | | |
| | | يميز بين أنواع العينات | | |
| | | يعدد أنواع العينات | | |
| | | يبين أخطاء المعاينة | | |
| عرض البيانات الاحصائية | تصنيف البيانات وتبويبها | يشرح مفهوم تبويب البيانات | | |
| | | يعرف آليات التعامل مع البيانات المجمعة | | |
| | | يصمم جداول التوزيع التكراري | | |
| | | ينظم البيانات في جدول تكراري | | |
| | العرض البياني للبيانات | يمثل بيانات الجداول التكرارية المختلفة بيانياً (أعمدة، دوائر، مدرج..) | | |
| | | يفرق بين أنواع الجداول التكرارية (البسيطة والمركبة) | | |
| | | يستخرج طول الفئة | | |
| مقاييس النزعة المركزية | مفهوم النزعة | يعرف مفهوم النزعة المركزية | | |
| | | يقارن بين حالات استخدام مقاييس النزعة المركزية | | |
| | المنوال | يعرف المنوال | | |
| | | يحدد خصائص المنوال | | |
| | | يحسب المنوال للبيانات المبوبة وغير المبوبة | | |
| | | يحسب المنوال بالطريقة البيانية | | |

| | | | | |
|---|---|---|---|---|
| | الوسط الحسابي الوسيط | يعرف الوسط الحسابي | | |
| | | يعرف الوسط الحسابي | | |
| | | يستخرج خصائص الوسط الحسابي | | |
| | | يحسب الوسط الحسابي للبيانات المبوبة وغير المبوبة | | |
| | | يوظف المتوسط الحسابي في الحالات المناسبة | | |
| | | يعرف الوسيط | | |
| | | يستخرج خصائص الوسيط | | |
| | | يحسب الوسيط للبيانات المبوبة وغير المبوبة | | |
| | | يحسب الوسيط بالطريقة البيانية | | |
| | | يوظف العلاقة بين مقاييس النزعة المركزية | | |
| مقاييس التشتت | المدى | يعرف المدى | | |
| | | يحسب المدى المطلق للبيانات المبوبة وغير المبوبة | | |
| | | يحسب المدى النسبي | | |
| | | يعدد ايجابيات ومساوئ المدى | | |
| | المدى الربيعي | يعرف مفهوم المدى الربيعي ونصف المدى الربيعي | | |
| | | يحسب المدى الربيعي للبيانات المبوبة وغير المبوبة | | |
| | | يحسب نصف المدى الربيعي | | |
| | الانحراف المتوسط | يعرف الانحراف المتوسط | | |
| | | يحسب الانحراف المتوسط للبيانات المبوبة وغير المبوبة | | |
| | الانحراف المعياري | يعرف الانحراف المعياري | | |
| | | يحسب الانحراف المعياري للبيانات المبوبة وغير المبوبة | | |
| | | يحدد مزايا وعيوب الانحراف المعياري | | |
| | | يوضح ضرورة وأهمية مقاييس التشتت | | |
| | | يفرق بين استخدامات مقاييس النزعة المركزية ومقاييس التشتت | | |
| تحويل الدرجات الخام | أنواع الدرجات المعيارية | يعرف مفهوم الدرجة المعيارية | | |
| | | يحسب الدرجة المعيارية | | |
| | | يعرف مفهوم الدرجة التائية | | |
| | | يحسب الدرجة التائية | | |
| | | يعرف مفهوم المئينيات | | |
| | | يحسب المئينيات | | |

## ملحق رقم (04)

### استبانة تحكيم الاختبار وتكرارات تقدير المحكمين في الصدق الوفي

| الفقرات | أتفق تماماً | أتفق | لا اعرف | لا اتفق | لا اتفق تماماً |
|---|---|---|---|---|---|
| 01 | 6 | 2 | | 1 | |
| 02 | 5 | 3 | 1 | | |
| 03 | 7 | 2 | | | |
| 04 | 6 | 2 | | 1 | |
| 05 | 7 | 1 | | 1 | |
| 06 | 8 | 1 | | | |
| 07 | 9 | | | | |
| 08 | 7 | 1 | 1 | | |
| 09 | 8 | | 1 | | |
| 10 | 4 | 5 | | | |
| 11 | 8 | 1 | | | |
| 12 | 6 | 2 | 1 | | |
| 13 | 6 | 3 | | | |
| 14 | 7 | 2 | | | |
| 15 | 8 | 1 | | | |
| 16 | 6 | 2 | 1 | | |
| 17 | 6 | 3 | | | |
| 18 | 8 | 1 | | | |
| 19 | 4 | 3 | 1 | 1 | |
| 20 | 8 | 1 | | | |

ملحق رقم (05)

جدول خاص بتحكيم بنود الاختبار

| الأسئلة الاختبارية | إجابات المحكم | | | | | |
|---|---|---|---|---|---|---|
| | الهدف السلوكى سليم وواضح لغوياً | | البند الاختباري سليم وواضح لغوياً | | البند الاختباري يقيس الهدف الذي وضع لقياسه | |
| | نعم | لا | نعم | لا | نعم | لا |
| 01 | | | | | | |
| 02 | | | | | | |
| 03 | | | | | | |
| 04 | | | | | | |
| 05 | | | | | | |
| 06 | | | | | | |
| 07 | | | | | | |
| 08 | | | | | | |
| 09 | | | | | | |
| 10 | | | | | | |
| 11 | | | | | | |
| 12 | | | | | | |
| 13 | | | | | | |
| 14 | | | | | | |
| 15 | | | | | | |
| 16 | | | | | | |
| 17 | | | | | | |
| 18 | | | | | | |
| 19 | | | | | | |
| 20 | | | | | | |

**ملاحظات أخرى:**

.....
.....
.....
.....

الملحق رقم (06)

نموذج تقدير المحكمين لفقرات الاختبار في الصدق الوصفي

| الفقرة | متفق تماماً | متفق | لا أدري | غير متفق | غير متفق تماماً |
|---|---|---|---|---|---|
| 01 | | | | | |
| 02 | | | | | |
| 03 | | | | | |
| 04 | | | | | |
| 05 | | | | | |
| 06 | | | | | |
| 07 | | | | | |
| 08 | | | | | |
| 09 | | | | | |
| 10 | | | | | |
| . | | | | | |
| . | | | | | |
| . | | | | | |
| 20 | | | | | |

ملحق رقم (07)

نتائج متوسط تقديرات المحكمين لفقرات الاختبار

| المحكمين | الفقرات | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 01 | 4 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 |
| 02 | 5 | 5 | 4 | 4 | 5 | 5 | 5 | 3 | 5 | 3 | 5 | 3 | 4 | 4 | 5 | 4 | 5 | 4 | 5 | 5 |
| 03 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 3 | 5 | 5 | 4 | 5 | 5 | 3 | 5 | 5 | 5 | 5 | 4 | 4 |
| 04 | 4 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 3 | 3 | 4 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| 05 | 5 | 4 | 4 | 5 | 4 | 4 | 5 | 3 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 06 | 4 | 3 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 4 |
| 07 | 3 | 5 | 5 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 5 | 5 | 5 | 4 | 5 | 4 | 5 | 4 | 5 | 4 |
| 08 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 5 | 3 | 4 | 3 | 4 | 4 | 4 |
| 09 | 4 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 4 | 5 | 3 | 5 | 4 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| متوسط التقدير | 4.33 | 4.22 | 4.56 | 4.67 | 4.67 | 4.78 | 4.89 | 4.11 | 4.78 | 4.56 | 4.22 | 4.22 | 4.44 | 4.56 | 4.44 | 4.44 | 4.67 | 4.56 | 4.56 | 4.56 |

الملحق رقم (08)

الأوساط الحسابية والانحرافات المعيارية لتقديرات المحكمين لصياغة فقرات الاختبار

**Descriptive Statistics**

| | N | Mean | Std. Deviation |
|---|---|---|---|
| فقرة_01 | 13 | 4.15 | .987 |
| فقرة_02 | 13 | 4.15 | .801 |
| فقرة_03 | 13 | 4.62 | .506 |
| فقرة_04 | 13 | 4.54 | .776 |
| فقرة_05 | 13 | 4.69 | .480 |
| فقرة_06 | 13 | 4.62 | .650 |
| فقرة_07 | 13 | 4.62 | .650 |
| فقرة_08 | 13 | 4.38 | .961 |
| فقرة_09 | 13 | 4.69 | .630 |
| فقرة_10 | 13 | 4.54 | .877 |
| فقرة_11 | 13 | 4.31 | .751 |
| فقرة_12 | 13 | 4.38 | .768 |
| فقرة_13 | 13 | 4.31 | .855 |
| فقرة_14 | 13 | 4.62 | .650 |
| فقرة_15 | 13 | 4.54 | .660 |
| فقرة_16 | 13 | 4.54 | .519 |
| فقرة_17 | 13 | 4.54 | .776 |
| فقرة_18 | 13 | 4.62 | .506 |
| فقرة_19 | 13 | 4.46 | .660 |
| فقرة_20 | 13 | 4.54 | .519 |
| Valid N (listwise) | 13 | | |

ملحق رقم (09)

قائمة بأسماء المحكمين

| الرقم | لقب واسم الأستاذ | التخصص | جهة العمل |
|---|---|---|---|
| 01 | بن عطاالله عائشة | اقتصاد كمي | عمار ثليجي الأغواط |
| 02 | بوسالم عبد العزيز | القياس النفسي والتقويم التربوي | القطب الجامعي العفرون (البليدة) |
| 03 | بوطي محمد نور الدين | علوم التربية | المدرسة العليا للأساتذة القبة |
| 04 | تواتي عيسى إبراهيم | علم النفس العيادي | 8 ماي 1945 قالمة |
| 05 | جدي العربي | اقتصاد كمي | عمار ثليجي الأغواط |
| 06 | جعفور ربيعة | علم النفس وعلوم التربية | قاصدي مرباح ورقلة |
| 07 | حمي سليم | علم النفس التربوي | محمد البشير الابراهيمي برج بو عريريج |
| 08 | حماد محمد | علوم التربية | حمة لخضر الواد |
| 09 | داودي خيرة | علوم التربية | زيان عاشور الجلفة |
| 10 | رماضنية أحمد | علوم التربية | المدرسة العليا للأساتذة الأغواط |
| 11 | سويسي دحمان | علم النفس تنظيم وعمل | الحاج لخضر باتنة |
| 12 | شلوفي عمير | اقتصاد كمي | عمار ثليجي الأغواط |
| 13 | شلوفي عمير | اقتصاد كمي | عمار ثليجي الأغواط |
| 14 | عداد وسام | تربية خاصة | العربي بن المهيدي أم البواقي |
| 15 | غريب مختار | علم النفس التربوي | زيان عاشور الجلفة |
| 16 | مرباح أحمد تقي الدين | علوم التربية | زيان عاشور الجلفة |
| 17 | مزوز عبد الحليم | علم النفس المدرسي | محمد لمين دباغين |

الملحق رقم (10)

تعليمات الاختبار التحصيلي محكي المرجع

واسئلة الاختبار ونموذج لورقة الاجابة

عزيزي الطالب

- ✓ يهدف الاختبار إلى قياس درجة اتقانكم لقواعد مقرر الإحصاء الوصفي.
- ✓ يتكون الاختبار من (03) أوراق.
- ✓ عدد أسئلة الاختبار (20) سؤالاً من نوع الاختيار من متعدد.
- ✓ مجموع درجات الاختبار (20) درجة. علماً أن لكل سؤال درجة واحدة.
- ✓ اقرأ أسئلة الاختبار جيداً ومن ثمَّ أجب عنها وذلك بوضع علامة صح في الخانة المناسبة على ورقة الاجابة. علماً أنَّه يوجد لكل سؤال، بديل واحد فقط صحيح من البدائل الأربعة.

ملاحظة01: يرجى الاجابة عن جميع الأسئلة وفق ما تمتلكونه من معلومات متحرين الدقة والسرعة قدر الامكان.

ملاحظة02: يرجى التأكد من وضع العلامة المحددة أمام الاجابة الصحيحة.

| | | | |
|---|---|---|---|
| 1. علم الإحصاء الوصفي هو العلم الذي: | | | |
| Ⓐ | يبحث في استقراء النتائج واتخاذ القرارات | Ⓑ | يقوم بوضع تصور لكيفية حل مشكلة ما دون محاولة حلها |
| Ⓒ | يهتم بجمع وتبويب وعرض ووصف البيانات دون الوصول إلى نتائج أو استدلالات خاصة | Ⓓ | يضع فرضيات لمشكلة ما ثم يقوم بدراسة صحة هذه الفرضيات |
| 2. عملية الحصول على القياسات والبيانات الخاصة بظاهرة معينة تسمى: | | | |
| Ⓐ | جمع البيانات | Ⓑ | تنظيم وعرض البيانات |
| Ⓒ | تحليل البيانات | Ⓓ | استقراء النتائج واتخاذ القرارات |
| 3. المجموعة التي تتكون من كل الأفراد محل الدراسة تسمى: | | | |
| Ⓐ | العينة | Ⓑ | البيانات |
| Ⓒ | المتغيرات | Ⓓ | المجتمع |
| 4. تقدير الطلبة في مقرر الاحصاء يعتبر متغير: | | | |
| Ⓐ | الاسمي | Ⓑ | كمي |
| Ⓒ | الترتيبي | Ⓓ | النسبي |
| 5. عدد الطلبة المتقدمين لامتحان الإحصاء الوصفي هو متغير هو: | | | |
| Ⓐ | متصل | Ⓑ | منفصل |
| Ⓒ | مستقل | Ⓓ | تابع |
| 6. أحد هذه المقاييس ليس مقياساً للنزعة المركزية: | | | |
| Ⓐ | الانحراف المعياري | Ⓑ | الوسيط |
| Ⓒ | الوسط الحسابي | Ⓓ | المنوال |
| 7. في المدرج التكراري لبيانات متصلة تكون الأعمدة الممثلة للفئات: | | | |
| Ⓐ | متداخلة فيما بينها | Ⓑ | فوق بعضها |
| Ⓒ | منفصلة عن بعضها بمسافات مناسبة | Ⓓ | متلاصقة تماماً |
| 8. عملية إيجاد قيم لمقاييس تتحدد قيمها من البيانات الخاصة بظاهرة معينة وتعطي بعض الدلالات عن تلك الظاهرة تسمى: | | | |
| Ⓐ | تحليل البيانات | Ⓑ | تنظيم وعرض البيانات |
| Ⓒ | استقراء النتائج واتخاذ القرارات | Ⓓ | جمع البيانات |
| 9. في طريقة عرض البيانات المنفصلة في منحنى تكراري تمثل كل قيمة من قيم المتغير x بـــ: | | | |
| Ⓐ | خط رأسي طوله يعبر عن تكرار تلك القيمة | Ⓑ | نقطة احداثياتها هي قيمة المتغير وتكرارها ويقوم بتوصيل هذه النقاط يدوياً |
| Ⓒ | خط افقي طوله يعبر عن تكرار تلك القيمة | Ⓓ | بقطاع من دائرة طبقاً لتكرارها |

| | | | |
|---|---|---|---|
| 10.تسمى القيم الأكثر تكراراً لمجموعة من القيمً بـــ: | | | |
| Ⓐ | المدى | Ⓑ | الوسط الحسابي |
| Ⓒ | الوسيط | Ⓓ | المنوال |
| 11.المدى لمجموعة من القيم هو : | | | |
| Ⓐ | العدد الأوسط | Ⓑ | العدد الأكثر تكراراَ |
| Ⓒ | الفرق بين أكبر عدد وأصغر عدد | Ⓓ | مجموع الأعداد مقسوماً على 2 |
| 12.الانحراف المعياري للبيانات غير المبوبة هو: | | | |
| Ⓐ | تباين هذه القيم | Ⓑ | مربع تباين هذه القيم |
| Ⓒ | نصف التباين للقيم | Ⓓ | الجذر التربيعي لتباين هذه القيم |
| 13.الدرجات المعيارية درجات محولة تحسب من خلال المتوسط والانحراف المعياري للعينة ويكون: | | | |
| Ⓐ | متوسطها (50) وانحرافها المعياري (10) | Ⓑ | متوسطها (صفر) وانحرافها المعياري (1) |
| Ⓒ | متوسطها (100) وانحرافها المعياري 50) | Ⓓ | متوسطها (واحد)وانحرافها المعياري صفر) |
| 14.الانحراف المعياري لمجموعة من القيم هو أحد مقاييس: | | | |
| Ⓐ | التشتت | Ⓑ | النزعة المركزية |
| Ⓒ | الارتباط | Ⓓ | التفرطح والالتواء |
| 15. المدى الربيعي يساوي | | | |
| Ⓐ | نصف الانحراف الربيعي | Ⓑ | الانحراف الربيعي |
| Ⓒ | ضعف الانحراف الربيعي | Ⓓ | المدى المئين |
| 16.المنوال للقيم التالية: 17 – 13 – 9 – 13 – 21 – 9 – 7 هو: | | | |
| Ⓐ | 13 و 9 | Ⓑ | 9 |
| Ⓒ | 13 | Ⓓ | 13 و 7 |
| 17.إذا كان المتوسط الحسابي لعلامات 5 طلاب في الإحصاء هو 12 فما هو مجموع علاماتهم: | | | |
| Ⓐ | 70 | Ⓑ | 50 |
| Ⓒ | 40 | Ⓓ | 60 |
| 18.لمجموعة من القيم: 9 – 4 – 2 – 4 – 12 – 6 فإن الوسيط لهذه القيم هو: | | | |
| Ⓐ | 2 | Ⓑ | 6 |
| Ⓒ | 5 | Ⓓ | 4 |
| 19.الانحراف المعياري لمجموعة القيم 9 – 8 – 5 – 6 – 2 هو: | | | |
| Ⓐ | 2.73 | Ⓑ | 3.45 |
| Ⓒ | 1.15 | Ⓓ | 0.74 |
| 20.الانحراف المتوسط لمجموعة القيم 17 – 7 – 3 هو: | | | |
| Ⓐ | 3.32 | Ⓑ | 1.25 |
| Ⓒ | 4.66 | Ⓓ | 2.42 |

## ورقة الاجابة عن بنود الاختبار التحصيلي محكي المرجع

اللقب:..... الاسم:.....

الفوج:..... رقم الطالب:.....

ضع اشارة في المربع المناسب المقابل للاحتمال الذي تراه صحيحاً على ورقة الاجابة علماً أنَّه للإجابة الصحيحة احتمال واحد فقط.

| رقم البند | Ⓐ | Ⓑ | Ⓒ | Ⓓ |
|---|---|---|---|---|
| 01 | | | | |
| 02 | | | | |
| 03 | | | | |
| 04 | | | | |
| 05 | | | | |
| 06 | | | | |
| 07 | | | | |
| 08 | | | | |
| 09 | | | | |
| 10 | | | | |
| 11 | | | | |
| 12 | | | | |
| 13 | | | | |
| 14 | | | | |
| 15 | | | | |
| 16 | | | | |
| 17 | | | | |
| 18 | | | | |
| 19 | | | | |
| 20 | | | | |

ملحق رقم (11)

مفتاح الاجابة عن فقرات الاختبار محكي المرجع لمقرر الإحصاء الوصفي

| رقم البند | Ⓐ | Ⓑ | Ⓒ | Ⓓ |
|---|---|---|---|---|
| 01 | | | ■ | |
| 02 | ■ | | | |
| 03 | | | | ■ |
| 04 | | | ■ | |
| 05 | | ■ | | |
| 06 | ■ | | | |
| 07 | | | | ■ |
| 08 | ■ | | | |
| 09 | | ■ | | |
| 10 | | | | ■ |
| 11 | | | ■ | |
| 12 | | | | ■ |
| 13 | | ■ | | |
| 14 | ■ | | | |
| 15 | | | ■ | |
| 16 | ■ | | | |
| 17 | | | | ■ |
| 18 | | ■ | | |
| 19 | ■ | | | |
| 20 | | | ■ | |

## الملحق رقم (12)
## مؤشرات الصعوبة لفقرات الاختبار للعينة الاستطلاعية

**Item Statistics**

| | Mean | Std. Deviation | N |
|---|---|---|---|
| ITEM01 | .66 | .478 | 67 |
| ITEM02 | .67 | .473 | 67 |
| ITEM03 | .73 | .447 | 67 |
| ITEM04 | .52 | .503 | 67 |
| ITEM05 | .60 | .494 | 67 |
| ITEM06 | .72 | .454 | 67 |
| ITEM07 | .60 | .494 | 67 |
| ITEM08 | .60 | .494 | 67 |
| ITEM09 | .63 | .487 | 67 |
| ITEM10 | .58 | .497 | 67 |
| ITEM11 | .64 | .483 | 67 |
| ITEM12 | .58 | .497 | 67 |
| ITEM13 | .61 | .491 | 67 |
| ITEM14 | .75 | .438 | 67 |
| ITEM15 | .55 | .501 | 67 |
| ITEM16 | .63 | .487 | 67 |
| ITEM17 | .55 | .501 | 67 |
| ITEM18 | .69 | .467 | 67 |
| ITEM19 | .69 | .467 | 67 |
| ITEM20 | .52 | .503 | 67 |

## الملحق رقم (13)
## معاملات تمييز فقرات الاختبار للعينة الاستطلاعية

**Item-Total Statistics**

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|---|
| ITEM01 | 11.85 | 28.038 | .739 | . | .888 |
| ITEM02 | 11.84 | 28.957 | .556 | . | .893 |
| ITEM03 | 11.78 | 29.692 | .436 | . | .897 |
| ITEM04 | 11.99 | 29.591 | .396 | . | .898 |
| ITEM05 | 11.91 | 29.537 | .415 | . | .897 |
| ITEM06 | 11.79 | 29.137 | .544 | . | .894 |
| ITEM07 | 11.91 | 29.416 | .439 | . | .897 |
| ITEM08 | 11.91 | 29.234 | .474 | . | .896 |
| ITEM09 | 11.88 | 29.228 | .483 | . | .895 |
| ITEM10 | 11.93 | 28.403 | .635 | . | .891 |
| ITEM11 | 11.87 | 29.482 | .438 | . | .897 |
| ITEM12 | 11.93 | 29.282 | .462 | . | .896 |
| ITEM13 | 11.90 | 28.731 | .577 | . | .893 |
| ITEM14 | 11.76 | 28.336 | .746 | . | .889 |
| ITEM15 | 11.96 | 28.165 | .676 | . | .890 |
| ITEM16 | 11.88 | 29.410 | .447 | . | .896 |
| ITEM17 | 11.96 | 28.680 | .574 | . | .893 |
| ITEM18 | 11.82 | 28.573 | .644 | . | .891 |
| ITEM19 | 11.82 | 29.604 | .430 | . | .897 |
| ITEM20 | 11.99 | 29.379 | .436 | . | .897 |

# ملحق رقم (14)
## قيم فاعلية البدائل في الاختبار التحصيلي

| الفقرة01 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 7 | 4 | 16 | 3 | 30 |
| الفئة الدنيا | 8 | 8 | 7 | 7 | 30 |
| جاذبية المشتت | 0.03- | 0.13- | 0.30 | 0.13- | |

| الفقرة02 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 21 | 4 | 4 | 1 | 30 |
| الفئة الدنيا | 7 | 12 | 8 | 3 | 30 |
| جاذبية المشتت | 0.47 | 0.27- | 0.13- | 0.07- | |

| الفقرة03 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 9 | 2 | 2 | 17 | 30 |
| الفئة الدنيا | 11 | 8 | 7 | 4 | 30 |
| جاذبية المشتت | 0.07- | 0.20- | 0.17- | 0.43 | |

| الفقرة04 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 14 | 0 | 14 | 2 | 30 |
| الفئة الدنيا | 11 | 5 | 7 | 7 | 30 |
| جاذبية المشتت | 0.10 | 0.17- | 0.23 | 0.17- | |

| الفقرة05 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 11 | 7 | 3 | 9 | 30 |
| الفئة الدنيا | 14 | 8 | 4 | 4 | 30 |
| جاذبية المشتت | 0.10- | 0.03- | 0.03- | 0.17 | |

| الفقرة06 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 13 | 8 | 0 | 9 | 30 |
| الفئة الدنيا | 11 | 8 | 7 | 4 | 30 |
| جاذبية المشتت | 0.07- | 0.20- | 0.17- | 0.43 | |

| الفقرة07 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 0 | 7 | 8 | 15 | 30 |
| الفئة الدنيا | 6 | 9 | 9 | 6 | 30 |
| جاذبية المشتت | 0.20- | 0.07- | 0.03- | 0.30 | |

| الفقرة08 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 16 | 8 | 3 | 3 | 30 |
| الفئة الدنيا | 10 | 8 | 5 | 7 | 30 |
| جاذبية المشتت | 0.20 | 0.00 | 0.07- | 0.13- | |

| الفقرة09 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 7 | 11 | 6 | 6 | 30 |
| الفئة الدنيا | 11 | 9 | 7 | 3 | 30 |
| جاذبية المشتت | 0.13- | 0.10- | 0.30- | 0.10 | |

| الفقرة10 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 6 | 4 | 0 | 20 | 30 |
| الفئة الدنيا | 9 | 12 | 4 | 5 | 30 |
| جاذبية المشتت | 0.10- | 0.27- | 0.13- | 0.50 | |

| الفقرة11 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 3 | 8 | 12 | 7 | 30 |
| الفئة الدنيا | 8 | 9 | 7 | 6 | 30 |
| جاذبية المشتت | 0.17- | 0.03- | 0.17 | 0.03 | |

| الفقرة12 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 9 | 3 | 7 | 11 | 30 |
| الفئة الدنيا | 8 | 5 | 8 | 9 | 30 |
| جاذبية المشتت | 0.03 | 0.07- | 0.03- | 0.07 | |

| الفقرة13 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 2 | 13 | 9 | 6 | 30 |
| الفئة الدنيا | 3 | 9 | 11 | 7 | 30 |
| جاذبية المشتت | 0.03- | 0.13 | 0.07- | 0.03- | |

| الفقرة14 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 15 | 9 | 0 | 6 | 30 |
| الفئة الدنيا | 11 | 9 | 4 | 6 | 30 |
| جاذبية المشتت | 0.13 | 0.00 | 0.13- | 0.00 | |

| الفقرة15 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 2 | 7 | 11 | 10 | 30 |
| الفئة الدنيا | 6 | 3 | 13 | 8 | 30 |
| جاذبية المشتت | -0.13 | 0.13 | -0.07 | 0.07 | |

| الفقرة16 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 17 | 4 | 4 | 5 | 30 |
| الفئة الدنيا | 9 | 3 | 7 | 11 | 30 |
| جاذبية المشتت | 0.27 | 0.03 | -0.10 | -0.20 | |

| الفقرة17 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 14 | 8 | 0 | 8 | 30 |
| الفئة الدنيا | 13 | 8 | 4 | 5 | 30 |
| جاذبية المشتت | 0.03 | 0.00 | -0.13 | 0.10 | |

| الفقرة18 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 3 | 18 | 4 | 5 | 30 |
| الفئة الدنيا | 4 | 11 | 8 | 7 | 30 |
| جاذبية المشتت | -0.03 | 0.23 | -0.13 | -0.07 | |

| الفقرة19 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 11 | 4 | 6 | 9 | 30 |
| الفئة الدنيا | 10 | 3 | 5 | 12 | 30 |
| جاذبية المشتت | 0.03 | 0.03 | 0.03 | -0.10 | |

| الفقرة20 | A | B | C | D | المجموع |
|---|---|---|---|---|---|
| الفئة العليا | 1 | 8 | 17 | 4 | 30 |
| الفئة الدنيا | 7 | 9 | 8 | 6 | 30 |
| جاذبية المشتت | -0.20 | -0.03 | 0.30 | -0.07 | |

ملحق رقم (15)

قائمة تقدير درجة القطع حسب طريقة أنجوف

الأستاذ الفاضل.

يقوم الباحث بإعداد رسالة للحصول على درجة الدكتوراه بعنوان (استخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر الاحصاء لطلبة سنة أولى علوم اجتماعية).

وفي هذا الإطار يقوم الباحث بتحديد درجة القطع للاختبار التحصيلي محكي المرجع لتحديد الطلبة المتمكنين وغير المتمكنين. وقد أعد الباحث لهذا الغرض قائمة موضح بها رقم كل مفردة وتقابلها مجموعة من فئات التقديرات والتي تمثل نسبة الطلبة الذين يستطيعون الإجابة على المفردة إجابة صحيحة من العدد الكلي للطلاب من أصحاب أقل مستوى مقبول من الأداء.

والمطلوب من سيادتكم الاطلاع على كل مفردة من مفردات الاختبار واختيار الفئة المناسبة من وجهة نظركم ووضع دائرة حولها لتحديد عدد الطلبة الممكن أن يجيبوا عن السؤال من ذوي أقل مستوى مقبول من الاداء.

مثال:

إذا كان عدد الطلبة من ذوي أقل مستوى مقبول من الأداء 100 طالب فكم عدد الطلبة من هؤلاء يستطيعون الاجابة عن المفردة إجابة صحيحة من بينهم، حسب وجهة نظرك.

- إذا كان العدد ما بين 0 و10 ضع دائرة حول الفئة (0 – 10)
- إذا كان العدد ما بين 11 و20 ضع دائرة حول الفئة (11 – 20) ...وهكذا

ولكم فائق التقدير والاحترام

## قائمة تقدير أنجوف لمفردات الاختبار

| رقم الفقرة | الفئات | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 01 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 02 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 03 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 04 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 05 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 06 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 07 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 08 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 09 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 10 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 11 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 12 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 13 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 14 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 15 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 16 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 17 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 18 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 19 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |
| 20 | 10 - 0 | 20 - 11 | 30 - 21 | 40 - 31 | 50 - 41 | 60 - 51 | 70 - 61 | 80 - 71 | 90 - 81 | 100 - 91 |

## ملحق رقم (16)

## نتائج تقدير أنجوف لمفردات الاختبار

| المحكمون | الفقرات | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 01 | 85 | 80 | 65 | 75 | 45 | 75 | 70 | 75 | 50 | 80 | 75 | 45 | 55 | 45 | 50 | 80 | 40 | 85 | 40 | 45 |
| 02 | 65 | 95 | 65 | 40 | 45 | 60 | 30 | 45 | 50 | 60 | 65 | 40 | 40 | 45 | 25 | 70 | 50 | 80 | 45 | 45 |
| 03 | 90 | 60 | 75 | 45 | 55 | 50 | 35 | 55 | 40 | 65 | 60 | 45 | 40 | 55 | 40 | 75 | 45 | 75 | 35 | 40 |
| 04 | 60 | 55 | 75 | 60 | 60 | 55 | 55 | 50 | 35 | 60 | 55 | 35 | 45 | 50 | 40 | 70 | 40 | 60 | 45 | 50 |
| 05 | 65 | 50 | 70 | 70 | 40 | 60 | 50 | 50 | 75 | 55 | 60 | 40 | 40 | 65 | 50 | 60 | 50 | 60 | 55 | 45 |
| 06 | 75 | 55 | 80 | 65 | 60 | 55 | 55 | 55 | 55 | 60 | 65 | 45 | 45 | 50 | 45 | 75 | 45 | 65 | 50 | 50 |
| 07 | 80 | 75 | 80 | 45 | 75 | 60 | 75 | 50 | 65 | 70 | 65 | 45 | 60 | 55 | 30 | 65 | 40 | 75 | 45 | 45 |
| 08 | 60 | 60 | 75 | 55 | 50 | 65 | 50 | 45 | 70 | 65 | 70 | 45 | 55 | 70 | 40 | 80 | 40 | 50 | 40 | 50 |
| 09 | 65 | 85 | 70 | 60 | 55 | 70 | 55 | 60 | 60 | 85 | 85 | 55 | 45 | 65 | 45 | 75 | 35 | 70 | 55 | 45 |
| 10 | 55 | 45 | 80 | 70 | 45 | 55 | 60 | 40 | 65 | 75 | 75 | 50 | 35 | 60 | 45 | 85 | 40 | 60 | 50 | 50 |
| 11 | 60 | 65 | 75 | 65 | 80 | 80 | 75 | 75 | 55 | 80 | 65 | 50 | 40 | 70 | 40 | 70 | 30 | 75 | 55 | 55 |
| **المجموع** | 69,09 | 65,91 | 73,64 | 59,09 | 55,45 | 62,27 | 55,45 | 54,55 | 56,36 | 68,64 | 67,27 | 45,00 | 45,45 | 57,27 | 40,91 | 73,18 | 41,36 | 66,82 | 46,82 | 47,27 |
| **1151.82** | | | | | | | | | | | | | | | | | | | | |

# ملحق رقم (17)
# نتائج تحليل اختبار (ت)

**T test**

**Group Statistics**

| | الطلبة | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| الدرجات | العليا | 21 | 17.57 | .926 | .202 |
| | الدنيا | 21 | 4.86 | 2.988 | .652 |

**Independent Samples Test**

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 95% Confidence Interval of the Difference | |
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| الدرجات | Equal variances assumed | 13.372 | .001 | 18.625 | 40 | .000 | 12.714 | .683 | 11.335 | 14.094 |
| | Equal variances not assumed | | | 18.625 | 23.805 | .000 | 12.714 | .683 | 11.305 | 14.124 |

ملحق رقم (18)

جدول المواصفات

| المحتوى | المجموع | مجاميع مستويات الاهداف ووزنها النسبي | | | | | | المجموع | المجموع | المجموع | عدد أسئلة الاختبار | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 16 | 17 | 4 | 16 | 3 | 12 | 2 | 54 | 100% | 100% | 20 | | | | | | |
| | | 31% | 7% | 30% | 6% | 22% | 4% | | | | 07 | 02 | 05 | 01 | 05 | 00 | |
| | عدد الحصص | تذكر | فهم | تطبيق | تركيب | تحليل | تقويم | أهداف الموضوع | الوزن النسبي للموضوع | الوزن النسبي للأهداف | اسئلة تذكر | اسئلة فهم | اسئلة تطبيق | اسئلة تركيب | اسئلة تحليل | اسئلة تقويم | |
| المفاهيم الأساسية لعلم الإحصاء | 2 | 1 | 2 | 0 | 0 | 2 | 0 | 5 | 13% | 9% | 1 | 0 | 1 | 0 | 1 | 0 | |
| مصادر وطرق جمع البيانات | 2 | 3 | 1 | 0 | 0 | 3 | 0 | 7 | 13% | 13% | 1 | 0 | 1 | 0 | 1 | 0 | |
| عرض البيانات الإحصائية | 3 | 2 | 0 | 0 | 3 | 2 | 0 | 7 | 19% | 13% | 0 | 0 | 1 | 0 | 1 | 0 | |
| مقاييس النزعة المركزية | 4 | 4 | 0 | 7 | 0 | 4 | 0 | 15 | 25% | 28% | 2 | 1 | 1 | 1 | 1 | 0 | |
| مقاييس التشتت | 4 | 4 | 1 | 6 | 0 | 1 | 2 | 14 | 25% | 26% | 2 | 1 | 1 | 0 | 1 | 0 | |
| تحويل الدرجات الخام | 1 | 3 | 0 | 3 | 0 | 0 | 0 | 6 | 6% | 11% | 1 | 0 | 0 | 0 | 0 | 0 | |

# ملحق رقم (19)

## مصفوفة الارتباطات بين فقرات الاختبار

Inter-Item Correlation Matrix

| | ITEM01 | ITEM02 | ITEM03 | ITEM04 | ITEM05 | ITEM06 | ITEM07 | ITEM08 | ITEM09 | ITEM10 | ITEM11 | ITEM12 | ITEM13 | ITEM14 | ITEM15 | ITEM16 | ITEM17 | ITEM18 | ITEM19 | ITEM20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ITEM01 | 1.000 | .499 | .271 | .316 | .431 | .521 | .367 | .303 | .547 | .471 | .443 | .343 | .521 | .590 | .487 | .482 | .360 | .596 | .257 | .316 |
| ITEM02 | .499 | 1.000 | .078 | .159 | .268 | .406 | .203 | .462 | .315 | .374 | .140 | .245 | .421 | .761 | .393 | .118 | .393 | .487 | .281 | .222 |
| ITEM03 | .271 | .078 | 1.000 | .095 | .051 | .291 | .257 | .189 | .089 | .306 | .320 | .237 | .277 | .266 | .335 | .159 | .335 | .316 | .824 | .229 |
| ITEM04 | .316 | .159 | .095 | 1.000 | .067 | .194 | .250 | .250 | .313 | .220 | .283 | .280 | .342 | .335 | .161 | .189 | .341 | .320 | .062 | .282 |
| ITEM05 | .431 | .268 | .051 | .067 | 1.000 | .091 | .131 | .380 | .436 | .291 | .211 | .168 | .345 | .220 | .423 | .310 | .239 | .232 | .232 | .128 |
| ITEM06 | .521 | .406 | .291 | .194 | .091 | 1.000 | .158 | .158 | .199 | .340 | .359 | .340 | .314 | .470 | .432 | .405 | .366 | .503 | .217 | .327 |
| ITEM07 | .367 | .203 | .257 | .250 | .131 | .158 | 1.000 | .131 | .058 | .599 | .148 | .229 | .407 | .360 | .423 | .310 | .178 | .298 | .232 | .189 |
| ITEM08 | .303 | .462 | .189 | .250 | .380 | .158 | .131 | 1.000 | .310 | .168 | .211 | .106 | .345 | .500 | .484 | .121 | .484 | .363 | .232 | .128 |
| ITEM09 | .547 | .315 | .089 | .313 | .436 | .199 | .058 | .310 | 1.000 | .222 | .325 | .222 | .336 | .401 | .298 | .426 | .298 | .277 | .077 | .251 |
| ITEM10 | .471 | .374 | .306 | .220 | .291 | .340 | .599 | .168 | .222 | 1.000 | .124 | .448 | .443 | .480 | .637 | .410 | .393 | .406 | .276 | .401 |
| ITEM11 | .443 | .140 | .320 | .283 | .211 | .359 | .148 | .211 | .325 | .124 | 1.000 | .124 | .172 | .280 | .329 | .325 | .266 | .368 | .300 | .220 |
| ITEM12 | .343 | .245 | .237 | .280 | .168 | .340 | .229 | .106 | .222 | .448 | .124 | 1.000 | .133 | .410 | .211 | .160 | .272 | .276 | .145 | .826 |
| ITEM13 | .521 | .421 | .277 | .342 | .345 | .314 | .407 | .345 | .336 | .443 | .172 | .133 | 1.000 | .451 | .515 | .272 | .392 | .320 | .320 | .097 |
| ITEM14 | .590 | .761 | .266 | .335 | .220 | .470 | .360 | .500 | .401 | .480 | .280 | .410 | .451 | 1.000 | .510 | .259 | .510 | .715 | .271 | .404 |
| ITEM15 | .487 | .393 | .335 | .161 | .423 | .432 | .423 | .484 | .298 | .637 | .329 | .211 | .515 | .510 | 1.000 | .298 | .577 | .492 | .297 | .161 |
| ITEM16 | .482 | .118 | .159 | .189 | .310 | .405 | .310 | .121 | .426 | .410 | .325 | .160 | .272 | .259 | .298 | 1.000 | .174 | .277 | .144 | .189 |
| ITEM17 | .360 | .393 | .335 | .341 | .239 | .366 | .178 | .484 | .298 | .393 | .266 | .272 | .392 | .510 | .577 | .174 | 1.000 | .362 | .297 | .161 |
| ITEM18 | .596 | .487 | .316 | .320 | .232 | .503 | .298 | .363 | .277 | .406 | .368 | .276 | .320 | .715 | .492 | .277 | .362 | 1.000 | .306 | .256 |
| ITEM19 | .257 | .281 | .824 | .062 | .232 | .217 | .232 | .232 | .077 | .276 | .300 | .145 | .320 | .271 | .297 | .144 | .297 | .306 | 1.000 | .127 |
| ITEM20 | .316 | .222 | .229 | .282 | .128 | .327 | .189 | .128 | .251 | .401 | .220 | .826 | .097 | .404 | .161 | .189 | .161 | .256 | .127 | 1.000 |

## ملحق رقم (20)

## نتائج الارتباط بين المفردات

**Summary Item Statistics**

| | Mean | Min | Max | Range | Max / Min | Variance | N of Items |
|---|---|---|---|---|---|---|---|
| Item Means | 0.625 | 0.522 | 0.746 | 0.224 | 1.429 | 0.004 | 20 |
| Item Variances | 0.234 | 0.192 | 0.253 | 0.061 | 1.318 | 0.000 | 20 |
| Inter-Item Covariances | 0.072 | 0.011 | 0.206 | 0.195 | 18.260 | 0.001 | 20 |
| Inter-Item Correlations | 0.310 | 0.051 | 0.826 | 0.774 | 16.115 | 0.021 | 20 |

## ملحق رقم (21)

## ثبات الفا كرونباك

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| **0.899** | 0.900 | 20 |

# ملحق رقم (22)

## واجهة برنامج WINSTEPS

```
INPUT: 502 PERSON  20 ITEM  REPORTED: 502 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
|                                                                             |
|                     * * * *  W I N S T E P S  * * * *                       |
|                     ---------------------------------                       |
|                                                                             |
|                 -  RASCH ANALYSIS FOR TWO-FACET MODELS  -                   |
|                                                                             |
|       PERSON, ITEM & RESPONSE STRUCTURE MEASUREMENT AND FIT ANALYSIS        |
|                                                                             |
|                      WINSTEPS®(R)     www.winsteps.com                      |
|                                                                             |
|               COPYRIGHT (C)©JOHN MICHAEL LINACRE, 1991-2017                |
|                      Permission to copy is granted.                         |
|                    July 24, 2017      VERSION 4.0.1                         |
-------------------------------------------------------------------------------
TITLE= Napil.sav
CONTROL FILE: C:\Users\user\Desktop\nabil.txt
OUTPUT FILE: C:\Users\user\Desktop\ZOU814WS.TXT
502 PERSON Records Input
INPUT: 502 PERSON  20 ITEM  REPORTED: 502 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------

                            CONVERGENCE TABLE

-------------------------------------------------------------------------------
|    PROX          ACTIVE COUNT       EXTREME 5 RANGE      MAX LOGIT CHANGE   |
| ITERATION  PERSON    ITEM  CATS     PERSON    ITEM      MEASURES  STRUCTURE|
|-----------------------------------------------------------------------------|
|         1      502      20     2       5.89    2.42        2.9444          |
|         2      492      20     2       6.90    3.23        -.7281          |
|         3      492      20     2       7.61    3.46         .3533          |
|         4      492      20     2       7.83    3.64        -.1517          |
|-----------------------------------------------------------------------------|
|    JMLE      MAX SCORE    MAX LOGIT     LEAST CONVERGED     CATEGORY STRUCTURE|
| ITERATION   RESIDUAL*      CHANGE  PERSON    ITEM    CAT   RESIDUAL   CHANGE|
|-----------------------------------------------------------------------------|
|       1        -6.84       .3046      70     14*                           |
|       2        -2.96      -.0893       7     14*                           |
|       3        -1.54       .0396       7     14*                           |
|       4         -.92       .0227      14     12*                           |
|       5         -.61       .0143      14     12*                           |
|       6         -.40       .0094      14     12*                           |
|       7         -.26       .0062       3     12*                           |
|       8         -.17       .0041      10     12*                           |
-------------------------------------------------------------------------------
Time for estimation: 0:0:0.402
INPUT: 502 PERSON  20 ITEM  REPORTED: 502 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------
&INST
Title= "Napil.sav"
; SPSS file created or last modified: 17/06/2021 15:44:33
; SPSS Cases processed = 502
; SPSS Variables processed = 21
ITEM1 = 1 ; Starting column of item responses
NI = 20 ; Number of items
NAME1 = 22 ; Starting column for person label in data record
NAMLEN = 4 ; Length of person label
XWIDE = 1 ; Matches the widest data value observed
CODES = 01 ; matches the data
TOTALSCORE = Yes ; Include extreme responses in reported scores
```

## ملحق رقم(23)

## أحادية البعد قبل الحذف

INPUT: 502 PERSON 20 ITEM REPORTED: 502 PERSON 20 ITEM 2 CATS WINSTEPS 4.0.1

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

| | Eigenvalue | Observed | | Expected |
|---|---|---|---|---|
| Total raw variance in observations = | 37.7151 | 100.0% | | 100.0% |
| Raw variance explained by measures = | 17.7151 | 47.0% | | 47.0% |
| Raw variance explained by persons = | 9.0763 | 24.1% | | 24.1% |
| Raw Variance explained by items = | 8.6389 | 22.9% | | 22.9% |
| Raw unexplained variance (total) = | 20.0000 | 53.0% | 100.0% | 53.0% |
| Unexplned variance in 1st contrast = | 2.1270 | 5.6% | 10.6% | |
| Unexplned variance in 2nd contrast = | 1.3734 | 3.6% | 6.9% | |
| Unexplned variance in 3rd contrast = | 1.2966 | 3.4% | 6.5% | |
| Unexplned variance in 4th contrast = | 1.2503 | 3.3% | 6.3% | |
| Unexplned variance in 5th contrast = | 1.2085 | 3.2% | 6.0% | |

## ملحق رقم(24)

## الاستقلال الموضعي

INPUT: 502 PERSON 20 ITEM REPORTED: 502 PERSON 20 ITEM 2 CATS WINSTEPS 4.0.1

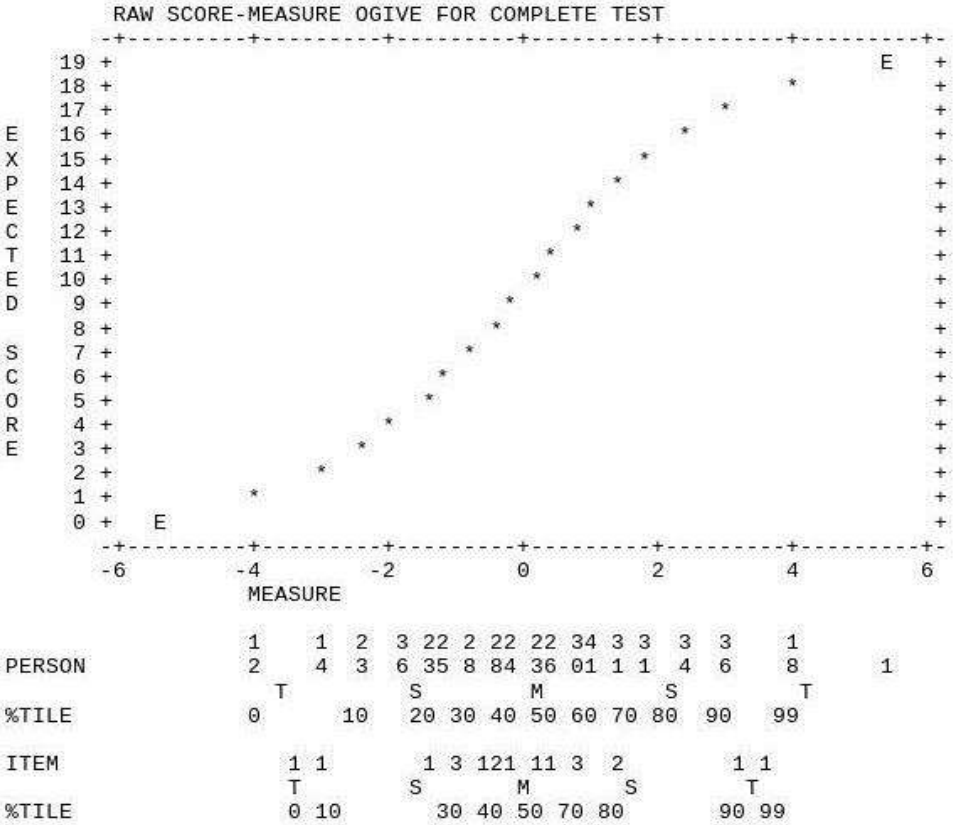
LARGEST STANDARDIZED RESIDUAL CORRELATIONS
USED TO IDENTIFY DEPENDENT ITEM

| CORREL-ATION | ENTRY NUMBER | ITEM | ENTRY NUMBER | ITEM |
|---|---|---|---|---|
| .87 | 9 | ITM09 | 16 | ITM16 |
| -.16 | 3 | ITM03 | 16 | ITM16 |
| -.15 | 14 | ITM14 | 16 | ITM16 |
| -.15 | 8 | ITM08 | 18 | ITM18 |
| -.14 | 5 | ITM05 | 7 | ITM07 |
| -.14 | 9 | ITM09 | 14 | ITM14 |
| -.14 | 11 | ITM11 | 16 | ITM16 |
| -.14 | 1 | ITM01 | 16 | ITM16 |
| -.13 | 3 | ITM03 | 9 | ITM09 |
| -.13 | 7 | ITM07 | 16 | ITM16 |
| -.13 | 6 | ITM06 | 9 | ITM09 |
| -.13 | 9 | ITM09 | 11 | ITM11 |
| -.13 | 7 | ITM07 | 8 | ITM08 |
| -.13 | 9 | ITM09 | 15 | ITM15 |
| -.12 | 9 | ITM09 | 17 | ITM17 |
| -.12 | 3 | ITM03 | 6 | ITM06 |
| -.12 | 7 | ITM07 | 9 | ITM09 |
| -.12 | 16 | ITM16 | 18 | ITM18 |
| -.12 | 9 | ITM09 | 18 | ITM18 |
| -.12 | 6 | ITM06 | 16 | ITM16 |

# ملحق رقم(25)

## منحنى دالة المعلومات

```
INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  19 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------
                        TABLE OF MEASURES ON TEST OF 19 ITEM
-------------------------------------------------------------------------------
| SCORE  MEASURE    S.E. | SCORE  MEASURE    S.E. | SCORE  MEASURE    S.E. |
|------------------------+------------------------+------------------------|
|     0    -5.35E   1.89 |     7     -.80     .57 |    14     1.46     .63 |
|     1    -3.97    1.12 |     8     -.48     .55 |    15     1.89     .68 |
|     2    -3.03     .86 |     9     -.18     .55 |    16     2.41     .76 |
|     3    -2.39     .74 |    10      .12     .55 |    17     3.08     .88 |
|     4    -1.90     .67 |    11      .43     .56 |    18     4.07    1.14 |
|     5    -1.49     .62 |    12      .75     .57 |    19     5.48E   1.90 |
|     6    -1.13     .59 |    13     1.09     .60 |                        |
-------------------------------------------------------------------------------
CURRENT VALUES, UIMEAN=.0000 USCALE=1.0000
TO SET MEASURE RANGE AS 0-100, UIMEAN=49.4165 USCALE=9.2379
TO SET MEASURE RANGE TO MATCH RAW SCORE RANGE, UIMEAN=9.3891 USCALE=1.7552
Predicting Score from Measure: Score = Measure * 2.1608 + 9.4934
Predicting Measure from Score: Measure = Score * .4442 + -4.2173
Maximum statistically different levels of performance (strata) = 3.6
Wright's Sample-independent Person (Test) Reliability based on maximum strata
= .93
```




```
                1    1  2  3 22 2 22 22 34 3 3  3  3    1
PERSON          2    4  3  6 35 8 84 36 01 1 1  4  6    8      1
                  T         S        M         S         T
%TILE             0      10   20 30 40 50 60 70 80  90   99

ITEM                 1 1       1 3 121 11 3  2        1 1
                     T        S       M       S        T
%TILE                0 10       30 40 50 70 80       90 99
```

## الملحق رقم (26)

## الاجابات التامة الصحيحة والخاطئة

```
INPUT: 502 PERSON  20 ITEM  REPORTED: 502 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------
PERSON: REAL SEP.: 2.43  REL.: .86 ... ITEM: REAL SEP.: 12.03  REL.: .99

         PERSON STATISTICS:  MEASURE ORDER
 ------------------------------------------------------------------------------------
|ENTRY  TOTAL TOTAL       MODEL|  INFIT  | OUTFIT  |PTMEASUR-AL|EXACT MATCH|      |
|NUMBER SCORE COUNT MEASURE S.E.|MNSQ ZSTD|MNSQ ZSTD|CORR.  EXP.| OBS%  EXP%|PERSON|
|-------------------------------+---------+---------+-----------+-----------+------|
|    70    20   20   5.37   1.89| MAXIMUM MEASURE   |  .00   .00|100.0 100.0|  70  |
|    78    20   20   5.37   1.89| MAXIMUM MEASURE   |  .00   .00|100.0 100.0|  78  |
|   175    20   20   5.37   1.89| MAXIMUM MEASURE   |  .00   .00|100.0 100.0| 175  |
|   361    20   20   5.37   1.89| MAXIMUM MEASURE   |  .00   .00|100.0 100.0| 361  |
|   375    20   20   5.37   1.89| MAXIMUM MEASURE   |  .00   .00|100.0 100.0| 64
     .      .    .     .      .                        .     .   .     .     .
     .      .    .     .      .                        .     .   .     .     .
     .      .    .     .      .                        .     .   .     .     .
|    97     0   20  -5.29   1.89| MINIMUM MEASURE   |  .00   .00|100.0 100.0|  97  |
|   173     0   20  -5.29   1.89| MINIMUM MEASURE   |  .00   .00|100.0 100.0| 173  |
|   315     0   20  -5.29   1.89| MINIMUM MEASURE   |  .00   .00|100.0 100.0| 315  |
|   379     0   20  -5.29   1.89| MINIMUM MEASURE   |  .00   .00|100.0 100.0| 379  |
|   499     0   20  -5.29   1.89| MINIMUM MEASURE   |  .00   .00|100.0 100.0| 499  |
|-------------------------------+---------+---------+-----------+-----------+------|
| MEAN  10.6  20.0     .23   .68|1.00   .0| .98    .1|           | 81.0  81.0|
| P.SD   5.3    .0    1.97   .23| .30   .9| .92    .7|           | 10.3   7.0|
```

## ملحق رقم (27)

## صعوبة المفردات قبل الحذف

INPUT:492 PERSON 20 ITEM REPORTED:484 PERSON 20 ITEM 2 CATS WINSTEPS 4.0.1

PERSON: REAL SEP.: 2.41 REL.: .85 ... ITEM: REAL SEP.: 12.13 REL.: .99

ITEM STATISTICS: ENTRY ORDER

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MODEL MEASUR | MODEL S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEASUR-AL CORR. | PTMEASUR-AL EXP. | EXACT MATCH OBS% | EXACT MATCH EXP% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 294 | 484 | -.44 | .12 | 1.03 | .5 | 1.03 | .3 | .59 | .60 | 78.7 | 78.8 |
| 2 | 328 | 484 | -.94 | .12 | .94 | -.9 | .84 | -.9 | .61 | .58 | 81.4 | 80.2 |
| 3 | 332 | 484 | -1.00 | .12 | 1.02 | .3 | .91 | -.5 | .57 | .58 | 79.3 | 80.4 |
| 4 | 291 | 484 | -.40 | .12 | .99 | -.1 | 1.01 | .1 | .60 | .60 | 79.3 | 78.8 |
| 5 | 250 | 484 | .16 | .12 | 1.02 | .3 | 1.00 | .1 | .60 | .61 | 77.9 | 78.0 |
| 6 | 273 | 484 | -.15 | .12 | .95 | -.8 | .94 | -.5 | .62 | .61 | 81.2 | 78.4 |
| 7 | 310 | 484 | -.67 | 12 | 1.01 | .2 | 1.00 | .1 | .59 | .59 | 76.9 | 79.3 |
| 8 | 336 | 484 | -1.06 | .12 | 1.02 | .4 | 1.12 | .7 | .56 | .58 | 82.2 | 80.7 |
| 9 | 212 | 484 | .68 | .12 | .97 | -.5 | .88 | -.9 | .62 | .61 | 79.1 | 77.9 |
| 10 | 227 | 484 | .47 | .12 | 1.14 | 2.4 | 1.21 | 1.7 | .55 | .61 | 72.9 | 77.8 |
| 11 | 237 | 484 | .34 | .12 | .99 | -.2 | 1.03 | .3 | .62 | .61 | 76.7 | 77.8 |
| 12 | 66 | 484 | 3.09 | .15 | 1.08 | .9 | .92 | -.2 | .42 | .44 | 87.4 | 87.9 |
| 13 | 170 | 484 | 1.26 | 12 | 1.04 | .7 | 1.04 | .3 | .57 | .59 | 77.3 | 79.0 |
| 14 | 159 | 484 | 1.42 | .12 | .97 | -.5 | 1.00 | .0 | .58 | .58 | 81.6 | 79.7 |
| 15 | 432 | 484 | -2.95 | .17 | .97 | -.2 | .93 | -.1 | .43 | .42 | 90.1 | 90.5 |
| 16 | 211 | 484 | .69 | 12 | .93 | -1.2 | .78 | -1.8 | .64 | .61 | 79.8 | 77.9 |
| 17 | 357 | 484 | -1.39 | 13 | 1.06 | .9 | 1.02 | .2 | .54 | .56 | 80.8 | 82.0 |
| 18 | 207 | 484 | .75 | .12 | .93 | -1.2 | .84 | -1.3 | .63 | .60 | 78.9 | 77.9 |
| 19 | 443 | 484 | -3.29 | .19 | .99 | .0 | .94 | .0 | .39 | .38 | 92.1 | 92.3 |
| 20 | 53 | 484 | 3.43 | .17 | 1.02 | .3 | .65 | -1.0 | .41 | .41 | 89.3 | 90.1 |
| MEAN | 259.4 | 484.0 | .00 | .13 | 1.00 | .1 | .95 | -.2 | | | 81.1 | 81.3 |
| P.SD | 100.2 | .0 | 1.61 | .02 | .05 | .8 | .12 | .7 | | | 4.8 | 4.6 |

ITEM CATEGORY/OPTION/DISTRACTOR FREQUENCIES: ENTRY ORDER

| ENTRY NUMBER | DATA CODE | SCORE VALUE | DATA COUNT | DATA % | ABILITY MEAN | ABILITY P.SD | S.E. MEAN | INFT MNSQ | OUTF MNSQ | PTMA CORR. | ITEM |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 190 | 39 | -1.11 | 1.46 | .11 | 1.0 | 1.0 | -.59 | ITM01 |
| | 1 | 1 | 294 | 61 | 1.14 | 1.54 | .09 | 1.0 | 1.1 | .59 | |
| 2 | 0 | 0 | 156 | 32 | -1.38 | 1.38 | .11 | 1.0 | .8 | -.61 | ITM02 |
| | 1 | 1 | 328 | 68 | 1.04 | 1.53 | .08 | .9 | .9 | .61 | |
| 3 | 0 | 0 | 152 | 31 | -1.32 | 1.41 | .12 | 1.0 | .9 | -.57 | ITM03 |
| | 1 | 1 | 332 | 69 | .98 | 1.57 | .09 | 1.0 | 1.0 | .57 | |
| 4 | 0 | 0 | 193 | 40 | -1.12 | 1.45 | .10 | 1.0 | 1.0 | -.60 | ITM04 |
| | 1 | 1 | 291 | 60 | 1.17 | 1.51 | .09 | 1.0 | 1.1 | .60 | |
| 5 | 0 | 0 | 234 | 48 | -.90 | 1.44 | .09 | 1.0 | .8 | -.60 | ITM05 |
| | 1 | 1 | 250 | 52 | 1.35 | 1.53 | .10 | 1.1 | 1.2 | .60 | |
| 6 | 0 | 0 | 211 | 44 | -1.06 | 1.45 | .10 | 1.0 | 1.0 | -.62 | ITM06 |
| | 1 | 1 | 273 | 56 | 1.28 | 1.46 | .09 | .9 | .9 | .62 | |
| 7 | 0 | 0 | 174 | 36 | -1.21 | 1.39 | .11 | 1.0 | .9 | -.59 | ITM07 |
| | 1 | 1 | 310 | 64 | 1.09 | 1.56 | .09 | 1.0 | 1.1 | .59 | |
| 8 | 0 | 0 | 148 | 31 | -1.31 | 1.48 | .12 | 1.1 | 1.2 | -.56 | ITM08 |
| | 1 | 1 | 336 | 69 | .95 | 1.57 | .09 | 1.0 | 1.0 | .56 | |
| 9 | 0 | 0 | 272 | 56 | -.77 | 1.51 | .09 | 1.0 | .9 | -.62 | ITM09 |
| | 1 | 1 | 212 | 44 | 1.58 | 1.38 | .09 | 1.0 | .9 | .62 | |
| 10 | 0 | 0 | 257 | 53 | -.71 | 1.59 | .10 | 1.1 | 1.2 | -.55 | ITM10 |
| | 1 | 1 | 227 | 47 | 1.35 | 1.51 | .10 | 1.1 | 1.2 | .55 | |
| 11 | 0 | 0 | 247 | 51 | -.87 | 1.47 | .09 | 1.0 | .9 | -.62 | ITM11 |
| | 1 | 1 | 237 | 49 | 1.43 | 1.46 | .10 | 1.0 | 1.2 | .62 | |
| 12 | 0 | 0 | 418 | 86 | -.05 | 1.75 | .09 | 1.1 | 1.1 | -.42 | ITM12 |
| | 1 | 1 | 66 | 14 | 2.24 | 1.21 | .15 | 1.1 | .9 | .42 | |
| 13 | 0 | 0 | 314 | 65 | -.52 | 1.59 | .09 | 1.0 | 1.0 | -.57 | ITM13 |
| | 1 | 1 | 170 | 35 | 1.70 | 1.42 | .11 | 1.1 | 1.1 | .57 | |
| 14 | 0 | 0 | 325 | 67 | -.50 | 1.58 | .09 | 1.0 | 1.0 | -.58 | ITM14 |
| | 1 | 1 | 159 | 33 | 1.82 | 1.36 | .11 | 1.0 | 1.0 | .58 | |
| 15 | 0 | 0 | 52 | 11 | -2.03 | 1.37 | .19 | 1.0 | .9 | -.43 | ITM15 |
| | 1 | 1 | 432 | 89 | .53 | 1.72 | .08 | 1.0 | .9 | .43 | |
| 16 | 0 | 0 | 273 | 56 | -.79 | 1.51 | .09 | 1.0 | .9 | -.64 | ITM16 |
| | 1 | 1 | 211 | 44 | 1.62 | 1.33 | .09 | .9 | .7 | .64 | |
| 17 | 0 | 0 | 127 | 26 | -1.41 | 1.43 | .13 | 1.1 | 1.0 | -.54 | ITM17 |
| | 1 | 1 | 357 | 74 | .85 | 1.62 | .09 | 1.0 | 1.0 | .54 | |
| 18 | 0 | 0 | 277 | 57 | -.76 | 1.49 | .09 | .9 | .9 | -.63 | ITM18 |
| | 1 | 1 | 207 | 43 | 1.63 | 1.37 | .10 | .9 | .8 | .63 | |
| 19 | 0 | 0 | 41 | 8 | -2.14 | 1.28 | .20 | 1.0 | .9 | -.39 | ITM19 |
| | 1 | 1 | 443 | 92 | .48 | 1.75 | .08 | 1.0 | 1.0 | .39 | |
| 20 | 0 | 0 | 431 | 89 | -.01 | 1.76 | .08 | 1.1 | 1.0 | -.41 | ITM20 |
| | 1 | 1 | 53 | 11 | 2.45 | 1.08 | .15 | 1.0 | .6 | .41 | |

# ملحق رقم (28)

## معامل الثبات والفصل بعد التدريج

```
INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  19 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------
      SUMMARY OF 483 MEASURED (NON-EXTREME) PERSON
-------------------------------------------------------------------------------
|          TOTAL                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      10.2      19.0         .28     .68      1.00     .0    .95     .1 |
| P.SD       4.9         .0        1.89     .15       .33     .9    .87     .7 |
| S.SD       4.9         .0        1.90     .15       .33     .9    .87     .7 |
| MAX.      18.0      19.0        4.06    1.14      2.17    2.4   9.90    3.7 |
| MIN.       1.0      19.0       -3.96     .55       .37   -2.1    .13   -1.1 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .74 TRUE SD    1.74  SEPARATION  2.35  PERSON RELIABILITY  .85 |
|MODEL RMSE    .70 TRUE SD    1.76  SEPARATION  2.53  PERSON RELIABILITY  .87 |
| S.E. OF PERSON MEAN = .09                                                   |
-------------------------------------------------------------------------------
  MAXIMUM EXTREME SCORE:      1 PERSON .2%
              DELETED:      8 PERSON

     SUMMARY OF 484 MEASURED (EXTREME AND NON-EXTREME) PERSON
-------------------------------------------------------------------------------
|          TOTAL                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      10.2      19.0         .29     .68                                 |
| P.SD       4.9         .0        1.91     .16                                 |
| S.SD       4.9         .0        1.91     .16                                 |
| MAX.      19.0      19.0        5.47    1.90                                 |
| MIN.       1.0      19.0       -3.96     .55                                 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .74 TRUE SD    1.75  SEPARATION  2.36  PERSON RELIABILITY  .85 |
|MODEL RMSE    .70 TRUE SD    1.77  SEPARATION  2.53  PERSON RELIABILITY  .87 |
| S.E. OF PERSON MEAN = .09                                                   |
-------------------------------------------------------------------------------
PERSON RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .89  SEM = 1.66

     SUMMARY OF 19 MEASURED (NON-EXTREME) ITEM
-------------------------------------------------------------------------------
|          TOTAL                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN     261.1     484.0         .00     .13      1.01     .1    .96    -.2 |
| P.SD     102.5         .0        1.67     .02       .04     .7    .11     .6 |
| S.SD     105.3         .0        1.72     .02       .05     .7    .12     .7 |
| MAX.     443.0     484.0        3.50     .19      1.09    1.0   1.13     .7 |
| MIN.      53.0     484.0       -3.30     .12       .93   -1.1    .64   -1.7 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .13 TRUE SD    1.67  SEPARATION 12.42  ITEM   RELIABILITY  .99 |
|MODEL RMSE    .13 TRUE SD    1.67  SEPARATION 12.57  ITEM   RELIABILITY  .99 |
| S.E. OF ITEM MEAN = .39                                                     |
-------------------------------------------------------------------------------
DELETED:      1 ITEM
ITEM RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000
```

## الملحق رقم (29)

## صعوبة المفردات النهائية مرتبة تنازلياً

INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  19 ITEM  2 CATS WINSTEPS 4.0.1

PERSON: REAL SEP.: 2.36  REL.: .85 ... ITEM: REAL SEP.: 12.42  REL.: .99

ITEM STATISTICS:  DISPLACEMENT ORDER

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEASUR-AL CORR. | PTMEASUR-AL EXP. | EXACT MATCH OBS% | EXACT MATCH EXP% | DISP-LACE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 53 | 484 | 3.50 | .17 | 1.05 | .5 | .64 | -1.0 | .41 | .42 | 89.2 | 90.2 | .00 |
| 12 | 66 | 484 | 3.16 | .16 | 1.09 | 1.0 | 1.04 | .2 | .42 | .45 | 87.4 | 88.1 | .00 |
| 14 | 159 | 484 | 1.47 | .12 | .97 | -.4 | 1.00 | .0 | .59 | .58 | 80.3 | 79.9 | .00 |
| 13 | 170 | 484 | 1.31 | .12 | 1.05 | .8 | 1.06 | .4 | .58 | .59 | 76.6 | 79.4 | .00 |
| 9 | 212 | 484 | .71 | .12 | .98 | -.4 | .89 | -.8 | .63 | .61 | 78.5 | 78.2 | .00 |
| 16 | 211 | 484 | .72 | .12 | .94 | 1.0 | .78 | -1.7 | .64 | .61 | 79.1 | 78.2 | .00 |
| 18 | 207 | 484 | .78 | .12 | .93 | -1.1 | .85 | -1.2 | .64 | .61 | 80.7 | 78.3 | .00 |
| 11 | 237 | 484 | .37 | .12 | .99 | -.1 | 1.04 | .4 | .62 | .62 | 77.4 | 78.2 | .00 |
| 5 | 250 | 484 | .19 | .12 | 1.02 | .3 | .99 | .0 | .61 | .62 | 77.2 | 78.3 | .00 |
| 6 | 273 | 484 | -.13 | .12 | .96 | -.7 | .95 | -.4 | .63 | .61 | 81.4 | 78.6 | .00 |
| 1 | 294 | 484 | -.43 | .12 | 1.03 | .5 | 1.03 | .3 | .59 | .61 | 78.3 | 79.0 | .00 |
| 4 | 291 | 484 | -.38 | .12 | 1.00 | .0 | 1.01 | .1 | .61 | .61 | 79.7 | 79.0 | .00 |
| 7 | 310 | 484 | -.66 | .12 | 1.02 | .4 | 1.03 | .2 | .59 | .60 | 76.6 | 79.6 | .00 |
| 2 | 328 | 484 | -.93 | .12 | .95 | -.9 | .83 | -1.0 | .61 | .59 | 82.0 | 80.4 | .00 |
| 3 | 332 | 484 | -.99 | .12 | 1.05 | .8 | .95 | -.2 | .57 | .58 | 79.5 | 80.6 | .00 |
| 8 | 336 | 484 | -1.05 | .12 | 1.04 | .6 | 1.13 | .7 | .56 | .58 | 81.6 | 80.8 | .00 |
| 17 | 357 | 484 | -1.39 | .13 | 1.06 | .9 | 1.04 | .3 | .54 | .56 | 81.2 | 82.1 | .00 |
| 15 | 432 | 484 | -2.96 | .17 | .99 | -.1 | .98 | .1 | .42 | .42 | 90.1 | 90.5 | .00 |
| 19 | 443 | 484 | -3.30 | .19 | .98 | -.1 | .95 | .0 | .40 | .38 | 92.3 | 92.3 | .00 |
| MEAN | 261.1 | 484.0 | .00 | .13 | 1.01 | .1 | .96 | -.2 | | | 81.5 | 81.7 | .00 |
| P.SD | 102.5 | .0 | 1.67 | .02 | .04 | .7 | .11 | .6 | | | 4.6 | 4.6 | .00 |

INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  19 ITEM  2 CATS WINSTEPS 4.0.1

ITEM CATEGORY/OPTION/DISTRACTOR FREQUENCIES:  DISPLACEMENT ORDER

| ENTRY NUMBER | DATA CODE | SCORE VALUE | DATA COUNT | DATA % | ABILITY MEAN | ABILITY P.SD | S.E. MEAN | INFT MNSQ | OUTF MNSQ | PTMA CORR. | ITEM |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 0 | 0 | 431 | 89 | .01 | 1.80 | .09 | 1.1 | 1.1 | -.41 | ITM20 |
| | 1 | 1 | 53 | 11 | 2.54 | 1.12 | .15 | 1.0 | .6 | .41 | |
| 12 | 0 | 0 | 418 | 86 | -.03 | 1.79 | .09 | 1.1 | 1.1 | -.42 | ITM12 |
| | 1 | 1 | 66 | 14 | 2.33 | 1.27 | .16 | 1.1 | 1.0 | .42 | |
| 14 | 0 | 0 | 325 | 67 | -.50 | 1.60 | .09 | 1.0 | 1.0 | -.59 | ITM14 |
| | 1 | 1 | 159 | 33 | 1.90 | 1.40 | .11 | 1.0 | 1.0 | .59 | |
| 13 | 0 | 0 | 314 | 65 | -.52 | 1.60 | .09 | 1.0 | 1.0 | -.58 | ITM13 |
| | 1 | 1 | 170 | 35 | 1.78 | 1.47 | .11 | 1.1 | 1.1 | .58 | |
| 9 | 0 | 0 | 272 | 56 | -.77 | 1.54 | .09 | 1.0 | .9 | -.63 | ITM09 |
| | 1 | 1 | 212 | 44 | 1.64 | 1.41 | .10 | 1.0 | .9 | .63 | |
| 16 | 0 | 0 | 273 | 56 | -.79 | 1.53 | .09 | 1.0 | .9 | -.64 | ITM16 |
| | 1 | 1 | 211 | 44 | 1.68 | 1.36 | .09 | .9 | .7 | .64 | |
| 18 | 0 | 0 | 277 | 57 | -.76 | 1.50 | .09 | .9 | .9 | -.64 | ITM18 |
| | 1 | 1 | 207 | 43 | 1.70 | 1.42 | .10 | 1.0 | .8 | .64 | |
| 11 | 0 | 0 | 247 | 51 | -.87 | 1.49 | .10 | 1.0 | .9 | -.62 | ITM11 |
| | 1 | 1 | 237 | 49 | 1.50 | 1.50 | .10 | 1.0 | 1.2 | .62 | |
| 5 | 0 | 0 | 234 | 48 | -.91 | 1.45 | .09 | 1.0 | .8 | -.61 | ITM05 |
| | 1 | 1 | 250 | 52 | 1.42 | 1.56 | .10 | 1.1 | 1.2 | .61 | |
| 6 | 0 | 0 | 211 | 44 | -1.07 | 1.47 | .10 | 1.0 | 1.0 | -.63 | ITM06 |
| | 1 | 1 | 273 | 56 | 1.34 | 1.50 | .09 | 1.0 | .9 | .63 | |
| 1 | 0 | 0 | 190 | 39 | -1.12 | 1.47 | .11 | 1.0 | 1.0 | -.59 | ITM01 |
| | 1 | 1 | 294 | 61 | 1.20 | 1.58 | .09 | 1.0 | 1.1 | .59 | |
| 4 | 0 | 0 | 193 | 40 | -1.13 | 1.47 | .11 | 1.0 | 1.0 | -.61 | ITM04 |
| | 1 | 1 | 291 | 60 | 1.23 | 1.55 | .09 | 1.0 | 1.1 | .61 | |
| 7 | 0 | 0 | 174 | 36 | -1.21 | 1.41 | .11 | 1.0 | 1.0 | -.59 | ITM07 |
| | 1 | 1 | 310 | 64 | 1.13 | 1.60 | .09 | 1.1 | 1.1 | .59 | |
| 2 | 0 | 0 | 156 | 32 | -1.40 | 1.39 | .11 | 1.0 | .8 | -.61 | ITM02 |
| | 1 | 1 | 328 | 68 | 1.09 | 1.56 | .09 | .9 | .9 | .61 | |
| 3 | 0 | 0 | 152 | 31 | -1.31 | 1.45 | .12 | 1.1 | .9 | -.57 | ITM03 |
| | 1 | 1 | 332 | 69 | 1.02 | 1.62 | .09 | 1.0 | 1.0 | .57 | |
| 8 | 0 | 0 | 148 | 31 | -1.31 | 1.50 | .12 | 1.1 | 1.2 | -.56 | ITM08 |
| | 1 | 1 | 336 | 69 | 1.00 | 1.62 | .09 | 1.0 | 1.0 | .56 | |
| 17 | 0 | 0 | 127 | 26 | -1.42 | 1.45 | .13 | 1.1 | 1.0 | -.54 | ITM17 |
| | 1 | 1 | 357 | 74 | .90 | 1.66 | .09 | 1.0 | 1.0 | .54 | |
| 15 | 0 | 0 | 52 | 11 | -2.03 | 1.38 | .19 | 1.0 | 1.0 | -.42 | ITM15 |
| | 1 | 1 | 432 | 89 | .57 | 1.76 | .09 | 1.0 | 1.0 | .42 | |
| 19 | 0 | 0 | 41 | 8 | -2.19 | 1.30 | .21 | .9 | .9 | -.40 | ITM19 |
| | 1 | 1 | 443 | 92 | .52 | 1.79 | .08 | 1.0 | 1.0 | .40 | |

## الملحق رقم (30)

### دالة المعلومات والمئينيات

| SCORE | MEASURE | S.E. | INFO | NORMED | S.E. | FREQUENCY | % | CUM.FREQ. | % | PERCENTILE |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | -5.35 | 1.89 | .28 | 204 | 99 | 0 | .0 | 0 | .0 | 0 |
| 1 | -3.97 | 1.12 | .80 | 277 | 59 | 12 | 2.5 | 12 | 2.5 | 1 |
| 2 | -3.03 | .86 | 1.35 | 326 | 45 | 14 | 2.9 | 26 | 5.4 | 4 |
| 3 | -2.39 | .74 | 1.82 | 359 | 39 | 23 | 4.8 | 49 | 10.1 | 8 |
| 4 | -1.90 | .67 | 2.24 | 385 | 35 | 36 | 7.4 | 85 | 17.6 | 14 |
| 5 | -1.49 | .62 | 2.61 | 407 | 32 | 23 | 4.8 | 108 | 22.3 | 20 |
| 6 | -1.13 | .59 | 2.90 | 426 | 31 | 25 | 5.2 | 133 | 27.5 | 25 |
| 7 | -.80 | .57 | 3.12 | 443 | 30 | 28 | 5.8 | 161 | 33.3 | 30 |
| 8 | -.48 | .55 | 3.26 | 459 | 29 | 28 | 5.8 | 189 | 39.0 | 36 |
| 9 | -.18 | .55 | 3.32 | 475 | 29 | 24 | 5.0 | 213 | 44.0 | 42 |
| 10 | .12 | .55 | 3.31 | 491 | 29 | 23 | 4.8 | 236 | 48.8 | 46 |
| 11 | .43 | .56 | 3.22 | 507 | 29 | 26 | 5.4 | 262 | 54.1 | 51 |
| 12 | .75 | .57 | 3.05 | 524 | 30 | 30 | 6.2 | 292 | 60.3 | 57 |
| 13 | 1.09 | .60 | 2.82 | 542 | 31 | 41 | 8.5 | 333 | 68.8 | 65 |
| 14 | 1.46 | .63 | 2.51 | 561 | 33 | 31 | 6.4 | 364 | 75.2 | 72 |
| 15 | 1.89 | .68 | 2.14 | 584 | 36 | 31 | 6.4 | 395 | 81.6 | 78 |
| 16 | 2.41 | .76 | 1.72 | 611 | 40 | 34 | 7.0 | 429 | 88.6 | 85 |
| 17 | 3.08 | .88 | 1.28 | 646 | 46 | 36 | 7.4 | 465 | 96.1 | 92 |
| 18 | 4.07 | 1.14 | .77 | 698 | 60 | 18 | 3.7 | 483 | 99.8 | 98 |
| 19 | 5.48 | 1.90 | .28 | 772 | 100 | 1 | .2 | 484 | 100.0 | 99 |

# الملحق رقم (31)

## خريطة ملاءمة المفردات

INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  19 ITEM  2 CATS WINSTEPS 4.0.1

PERSON: REAL SEP.: 2.36  REL.: .85 ... ITEM: REAL SEP.: 12.42  REL.: .99
ITEM STATISTICS:  MISFIT ORDER

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEASUR-AL CORR. | PTMEASUR-AL EXP. | EXACT MATCH OBS% | EXACT MATCH EXP% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 336 | 484 | -1.05 | .12 | 1.04 | .6 | 1.13 | .7 | A .56 | .58 | 81.6 | 80.8 |
| 12 | 66 | 484 | 3.16 | .16 | 1.09 | 1.0 | 1.04 | .2 | B .42 | .45 | 87.4 | 88.1 |
| 13 | 170 | 484 | 1.31 | .12 | 1.05 | .8 | 1.06 | .4 | C .58 | .59 | 76.6 | 79.4 |
| 17 | 357 | 484 | -1.39 | .13 | 1.06 | .9 | 1.04 | .3 | D .54 | .56 | 81.2 | 82.1 |
| 3 | 332 | 484 | -.99 | .12 | 1.05 | .8 | .95 | -.2 | E .57 | .58 | 79.5 | 80.6 |
| 20 | 53 | 484 | 3.50 | .17 | 1.05 | .5 | .64 | -1.0 | F .41 | .42 | 89.2 | 90.2 |
| 11 | 237 | 484 | .37 | .12 | .99 | -.1 | 1.04 | .4 | G .62 | .62 | 77.4 | 78.2 |
| 1 | 294 | 484 | -.43 | .12 | 1.03 | .5 | 1.03 | .3 | H .59 | .61 | 78.3 | 79.0 |
| 7 | 310 | 484 | -.66 | .12 | 1.02 | .4 | 1.03 | .2 | I .59 | .60 | 76.6 | 79.6 |
| 5 | 250 | 484 | .19 | .12 | 1.02 | .3 | .99 | .0 | J .61 | .62 | 77.2 | 78.3 |
| 4 | 291 | 484 | -.38 | .12 | 1.00 | .0 | 1.01 | .1 | i .61 | .61 | 79.7 | 79.0 |
| 14 | 159 | 484 | 1.47 | .12 | .97 | -.4 | 1.00 | .0 | h .59 | .58 | 80.3 | 79.9 |
| 15 | 432 | 484 | -2.96 | .17 | .99 | -.1 | .98 | .1 | g .42 | .42 | 90.1 | 90.5 |
| 9 | 212 | 484 | .71 | .12 | .98 | -.4 | .89 | -.8 | f .63 | .61 | 78.5 | 78.2 |
| 19 | 443 | 484 | -3.30 | .19 | .98 | -.1 | .95 | .0 | e .40 | .38 | 92.3 | 92.3 |
| 6 | 273 | 484 | -.13 | .12 | .96 | -.7 | .95 | -.4 | d .63 | .61 | 81.4 | 78.6 |
| 2 | 328 | 484 | -.93 | .12 | .95 | -.9 | .83 | -1.0 | c .61 | .59 | 82.0 | 80.4 |
| 16 | 211 | 484 | .72 | .12 | .94 | -1.0 | .78 | -1.7 | b .64 | .61 | 79.1 | 78.2 |
| 18 | 207 | 484 | .78 | .12 | .93 | -1.1 | .85 | -1.2 | a .64 | .61 | 80.7 | 78.3 |
| MEAN | 261.1 | 484.0 | .00 | .13 | 1.01 | .1 | .96 | -.2 | | | 81.5 | 81.7 |
| P.SD | 102.5 | .0 | 1.67 | .02 | .04 | .7 | .11 | .6 | | | 4.6 | 4.6 |

INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  19 ITEM  2 CATS WINSTEPS 4.0.1

ITEM FIT GRAPH:  MISFIT ORDER

| ENTRY NUMBER | MEASURE - + | INFIT MEAN-SQUARE 0.0 1 2 | OUTFIT MEAN-SQUARE 0.0 1 2 | ITEM |
|---|---|---|---|---|
| 8 | * | : * | : .* | ITM08 |
| 12 | * | : * | : * | ITM12 |
| 13 | * | : * | : * | ITM13 |
| 17 | * | : * | : * | ITM17 |
| 3 | * | : * | : *. | ITM03 |
| 20 | * | : * | : * . | ITM20 |
| 11 | * | : *. | : * | ITM11 |
| 1 | * | : * | : * | ITM01 |
| 7 | * | : * | : * | ITM07 |
| 5 | * | : * | : *. | ITM05 |
| 4 | * | : *. | : * | ITM04 |
| 14 | * | : *. | : *. | ITM14 |
| 15 | * | : *. | : *. | ITM15 |
| 9 | * | : *. | : * . | ITM09 |
| 19 | * | : *. | : *. | ITM19 |
| 6 | * | : *. | : *. | ITM06 |
| 2 | * | : *. | : * . | ITM02 |
| 16 | * | : *. | : * . | ITM16 |
| 18 | * | : *. | : * . | ITM18 |

INPUT:492 PERSON 20 ITEM  REPORTED:484 PERSON 19 ITEM 2 CATS WINSTEPS4.0.1

ITEM CATEGORY/OPTION/DISTRACTOR FREQUENCIES:  MISFIT ORDER

| ENTRY NUMBER | DATA CODE | SCORE VALUE | DATA COUNT | % | ABILITY MEAN | P.SD | S.E. MEAN | INFT MNSQ | OUTF MNSQ | PTMA CORR. | ITEM |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | A | 0 | 148 | 31 | 1.31 | 1.50 | .12 | 1.1 | 1.2 | -.56 | ITM08 |
| | | 1 | 336 | 69 | 1.00 | 1.62 | .09 | 1.0 | 1.0 | .56 | |
| 12 | B | 0 | 418 | 86 | -.03 | 1.79 | .09 | 1.1 | 1.1 | -.42 | ITM12 |
| | | 1 | 66 | 14 | 2.33 | 1.27 | .16 | 1.1 | 1.0 | .42 | |
| 13 | C | 0 | 314 | 65 | -.52 | 1.60 | .09 | 1.0 | 1.0 | -.58 | ITM13 |
| | | 1 | 170 | 35 | 1.78 | 1.47 | .11 | 1.1 | 1.1 | .58 | |
| 17 | D | 0 | 127 | 26 | 1.42 | 1.45 | .13 | 1.1 | 1.0 | -.54 | ITM17 |
| | | 1 | 357 | 74 | .90 | 1.66 | .09 | 1.0 | 1.0 | .54 | |
| 3 | E | 0 | 152 | 31 | 1.31 | 1.45 | .12 | 1.1 | .9 | -.57 | ITM03 |
| | | 1 | 332 | 69 | 1.02 | 1.62 | .09 | 1.0 | 1.0 | .57 | |
| 20 | F | 0 | 431 | 89 | .01 | 1.80 | .09 | 1.1 | 1.1 | -.41 | ITM20 |
| | | 1 | 53 | 11 | 2.54 | 1.12 | .15 | 1.0 | .6 | .41 | |
| 11 | G | 0 | 247 | 51 | -.87 | 1.49 | .10 | 1.0 | .9 | -.62 | ITM11 |
| | | 1 | 237 | 49 | 1.50 | 1.50 | .10 | 1.0 | 1.2 | .62 | |
| 1 | H | 0 | 190 | 39 | 1.12 | 1.47 | .11 | 1.0 | 1.0 | -.59 | ITM01 |
| | | 1 | 294 | 61 | 1.20 | 1.58 | .09 | 1.0 | 1.1 | .59 | |
| 7 | I | 0 | 174 | 36 | .21 | 1.41 | .11 | 1.0 | 1.0 | -.59 | ITM07 |
| | | 1 | 310 | 64 | .13 | 1.60 | .09 | 1.1 | 1.1 | .59 | |
| 5 | J | 0 | 234 | 48 | -.91 | 1.45 | .09 | 1.0 | .8 | -.61 | ITM05 |
| | | 1 | 250 | 52 | 1.42 | 1.56 | .10 | 1.1 | 1.2 | .61 | |
| 4 | i | 0 | 193 | 40 | -1.13 | 1.47 | .11 | 1.0 | 1.0 | -.61 | ITM04 |
| | | 1 | 291 | 60 | 1.23 | 1.55 | .09 | 1.0 | 1.1 | .61 | |
| 14 | h | 0 | 325 | 67 | -.50 | 1.60 | .09 | 1.0 | 1.0 | -.59 | ITM14 |
| | | 1 | 159 | 33 | 1.90 | 1.40 | .11 | 1.0 | 1.0 | .59 | |
| 15 | g | 0 | 52 | 11 | -2.03 | 1.38 | .19 | 1.0 | 1.0 | -.42 | ITM15 |
| | | 1 | 432 | 89 | .57 | 1.76 | .09 | 1.0 | 1.0 | .42 | |
| 9 | f | 0 | 272 | 56 | -.77 | 1.54 | .09 | 1.0 | .9 | -.63 | ITM09 |
| | | 1 | 212 | 44 | 1.64 | 1.41 | .10 | 1.0 | .9 | .63 | |
| 19 | e | 0 | 41 | 8 | -2.19 | 1.30 | .21 | .9 | .9 | -.40 | ITM19 |
| | | 1 | 443 | 92 | .52 | 1.79 | .08 | 1.0 | 1.0 | .40 | |
| 6 | d | 0 | 211 | 44 | -1.07 | 1.47 | .10 | 1.0 | 1.0 | -.63 | ITM06 |
| | | 1 | 273 | 56 | 1.34 | 1.50 | .09 | 1.0 | .9 | .63 | |
| 2 | c | 0 | 156 | 32 | -1.40 | 1.39 | .11 | 1.0 | .8 | -.61 | ITM02 |
| | | 1 | 328 | 68 | 1.09 | 1.56 | .09 | .9 | .9 | .61 | |
| 16 | b | 0 | 273 | 56 | -.79 | 1.53 | .09 | 1.0 | .9 | -.64 | ITM16 |
| | | 1 | 211 | 44 | 1.68 | 1.36 | .09 | .9 | .7 | .64 | |
| 18 | a | 0 | 277 | 57 | -.76 | 1.50 | .09 | .9 | .9 | -.64 | ITM18 |
| | | 1 | 207 | 43 | 1.70 | 1.42 | .10 | 1.0 | .8 | .64 | |

```
INPUT:492 PERSON 20 ITEM REPORTED: 484 PERSON 19 ITEM 2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------
MOST MISFITTING RESPONSE STRINGS
ITEM     OUTMNSQ  |PERSON
                  |4331144442211  44332221  21  43 33 23322431324322
                  |3106182102185622190220O2 44 763851 968638993419317
                  |7151441568984553158545579966165184872389191862387G876
               high-----------------------------------------------------
 8 ITM08    1.13 A|0.............0000.............................1
12 ITM12    1.04 B|.........................1.1...1.11..1............
13 ITM13    1.06 C|.0..............................1....1..11.......
17 ITM17    1.04 D|.........0...0.....0.....0.0...0..................
 3 ITM03     .95 E|...................0..0.00.................1.1......
20 ITM20     .64 F|.....................1.........1.1..................
11 ITM11    1.04 G|.....................................1..1......1...
 1 ITM01    1.03 H|.....00.0.0.........0.0.......................1....
 7 ITM07    1.03 I|....0..0...0.................................1..
 5 ITM05     .99 J|..........0...0.....0...................1...1.....
 4 ITM04    1.01 i|.0....0.....0...0.......0.0.....................1.
14 ITM14    1.00 h|.....................................1..11.........
15 ITM15     .98 g|.......................0...0..0.....................
 9 ITM09     .89 f|......0.....................................1......
19 ITM19     .95 e|.....................0.......0......0..............
 6 ITM06     .95 d|..00.......................0.........................
 2 ITM02     .83 c|............0.............................1.......
                  |-----------------------------------------------------low
                  |4331144442211624433222129216743833823322431324322 7
                  |3106182102185552190220O7 44 163151 968638993419316
                  |7151441568984  31585455  96  65 84 72389191862387
```

```
INPUT:492 PERSON 20 ITEM REPORTED:484 PERSON 19 ITEM 2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------
MOST UNEXPECTED RESPONSES
ITEM     MEASURE  |PERSON
                  |4331144442211  44332221  21  43 33 23322431324322
                  |3106182102185622190220O2 44 763851 968638993419317
                  |7151441568984553158545579966165184872389191862387G876
               high-----------------------------------------------------
19 ITM19   -3.30 e|.....................0.......0......0..............
15 ITM15   -2.96 g|.......................0...0..0.....................
17 ITM17   -1.39 D|.........0...0.....0.....0.0...0...................
 8 ITM08   -1.05 A|0.............0000.............................1
 3 ITM03    -.99 E|...................0..0.00.................1.1......
 2 ITM02    -.93 c|............0.............................1.......
 7 ITM07    -.66 I|....0..0...0.................................1..
 1 ITM01    -.43 H|.....00.0.0.........0.0.......................1....
 4 ITM04    -.38 i|.0....0.....0...0.......0.0.....................1.
 6 ITM06    -.13 d|..00.......................0.........................
 5 ITM05     .19 J|..........0...0.....0...................1...1.....
11 ITM11     .37 G|.....................................1..1......1...
 9 ITM09     .71 f|......0.....................................1......
13 ITM13    1.31 C|.0..............................1....1..11.......
14 ITM14    1.47 h|.....................................1..11.........
12 ITM12    3.16 B|.........................1.1...1.11..1............
20 ITM20    3.50 F|.....................1.........1.1..................
                  |-----------------------------------------------------low
                  |4331144442211624433222129216743833823322431324322 7
                  |3106182102185552190220O7 44 163151 968638993419316
                  |7151441568984  31585455  96  65 84 72389191862387
```

INPUT:492 PERSON 20 ITEM REPORTED:484 PERSON 19 ITEM  2 CATS WINSTEPS 4.0.1

MOST UNEXPECTED RESPONSES

| DATA | OBSERVED | EXPECTED | RESIDUAL | ST.RES. | MEASDIFF | ITEM | PERSON | ITEM | PERSON |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | .99 | -.99 | -13.43 | 5.19 | 19 | 224 | ITM19 | 229 |
| 1 | 1 | .01 | .99 | 12.58 | -5.06 | 12 | 383 | ITM12 | 392 |
| 0 | 0 | .99 | -.99 | -11.29 | 4.85 | 15 | 105 | ITM15 | 108 |
| 0 | 0 | .99 | -.99 | -9.10 | 4.42 | 15 | 146 | ITM15 | 149 |
| 0 | 0 | .99 | -.99 | -8.98 | 4.39 | 19 | 71 | ITM19 | 72 |
| 1 | 1 | .01 | .99 | 8.71 | -4.33 | 11 | 393 | ITM11 | 402 |
| 0 | 0 | .98 | -.98 | -7.88 | 4.13 | 8 | 437 | ITM08 | 446 |
| 1 | 1 | .02 | .98 | 6.89 | -3.86 | 14 | 481 | ITM14 | 490 |
| 0 | 0 | .98 | -.98 | -6.67 | 3.80 | 17 | 228 | ITM17 | 233 |
| 0 | 0 | .98 | -.98 | -6.67 | 3.80 | 17 | 65 | ITM17 | 65 |
| 0 | 0 | .98 | -.98 | -6.48 | 3.74 | 7 | 114 | ITM07 | 117 |
| 1 | 1 | .02 | .98 | 6.47 | -3.73 | 9 | 338 | ITM09 | 344 |
| 1 | 1 | .02 | .98 | 6.35 | -3.70 | 13 | 399 | ITM13 | 408 |
| 1 | 1 | .02 | .98 | 6.35 | -3.70 | 13 | 191 | ITM13 | 196 |
| 1 | 1 | .02 | .98 | 6.31 | -3.68 | 20 | 358 | ITM20 | 365 |
| 1 | 1 | .03 | .97 | 5.99 | -3.58 | 4 | 217 | ITM04 | 222 |
| 1 | 1 | .03 | .97 | 5.86 | -3.54 | 1 | 412 | ITM01 | 421 |
| 0 | 0 | .97 | -.97 | -5.65 | 3.46 | 4 | 311 | ITM04 | 317 |
| 0 | 0 | .97 | -.97 | -5.64 | 3.46 | 8 | 25 | ITM08 | 25 |
| 0 | 0 | .97 | -.97 | -5.43 | 3.38 | 15 | 466 | ITM15 | 475 |
| 1 | 1 | .03 | .97 | 5.39 | -3.37 | 14 | 239 | ITM14 | 244 |
| 1 | 1 | .03 | .97 | 5.32 | -3.34 | 12 | 314 | ITM12 | 320 |
| 1 | 1 | .03 | .97 | 5.32 | -3.34 | 12 | 8 | ITM12 | 8 |
| 0 | 0 | .97 | -.97 | -5.30 | 3.34 | 2 | 154 | ITM02 | 157 |
| 1 | 1 | .04 | .96 | 5.22 | -3.31 | 7 | 238 | ITM07 | 243 |
| 0 | 0 | .96 | -.96 | -5.15 | 3.28 | 17 | 225 | ITM17 | 230 |
| 1 | 1 | .04 | .96 | 4.98 | -3.21 | 5 | 246 | ITM05 | 251 |
| 0 | 0 | .96 | -.96 | -4.98 | 3.21 | 6 | 305 | ITM06 | 310 |
| 0 | 0 | .96 | -.96 | -4.98 | 3.21 | 6 | 161 | ITM06 | 164 |
| 1 | 1 | .04 | .96 | 4.97 | -3.21 | 13 | 268 | ITM13 | 273 |
| 1 | 1 | .04 | .96 | 4.66 | -3.08 | 20 | 335 | ITM20 | 341 |
| 0 | 0 | .96 | -.96 | -4.64 | 3.07 | 7 | 415 | ITM07 | 424 |
| 0 | 0 | .96 | -.96 | -4.64 | 3.07 | 7 | 188 | ITM07 | 193 |
| 1 | 1 | .05 | .95 | 4.58 | -3.04 | 12 | 81 | ITM12 | 83 |
| 1 | 1 | .05 | .95 | 4.39 | -2.96 | 14 | 362 | ITM14 | 369 |
| 0 | 0 | .95 | -.95 | -4.35 | 2.94 | 8 | 423 | ITM08 | 432 |
| 0 | 0 | .95 | -.95 | -4.35 | 2.94 | 8 | 411 | ITM08 | 420 |
| 0 | 0 | .95 | -.95 | -4.35 | 2.94 | 8 | 395 | ITM08 | 404 |
| 1 | 1 | .05 | .95 | 4.30 | -2.92 | 8 | 76 | ITM08 | 77 |
| 0 | 0 | .95 | -.95 | -4.22 | 2.88 | 3 | 308 | ITM03 | 313 |
| 0 | 0 | .95 | -.95 | -4.22 | 2.88 | 3 | 205 | ITM03 | 210 |
| 0 | 0 | .95 | -.95 | -4.22 | 2.88 | 3 | 27 | ITM03 | 27 |
| 0 | 0 | .95 | -.95 | -4.22 | 2.88 | 3 | 9 | ITM03 | 9 |
| 0 | 0 | .95 | -.95 | -4.15 | 2.85 | 17 | 249 | ITM17 | 254 |
| 0 | 0 | .95 | -.95 | -4.15 | 2.85 | 17 | 6 | ITM17 | 6 |
| 0 | 0 | .94 | -.94 | -4.13 | 2.84 | 1 | 484 | ITM01 | 493 |
| 0 | 0 | .94 | -.94 | -4.13 | 2.84 | 1 | 421 | ITM01 | 430 |
| 0 | 0 | .94 | -.94 | -4.13 | 2.84 | 1 | 406 | ITM01 | 415 |
| 0 | 0 | .94 | -.94 | -4.13 | 2.84 | 1 | 219 | ITM01 | 224 |
| 0 | 0 | .94 | -.94 | -4.10 | 2.82 | 19 | 297 | ITM19 | 302 |

# ملحق رقم (32)

## الملخص بعد حذف الأفراد

```
INPUT: 492 PERSON  20 ITEM  REPORTED: 484 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------

     SUMMARY OF 484 MEASURED PERSON
-------------------------------------------------------------------------------
|          TOTAL                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      10.7      20.0         .26     .66      1.00     .0    .95     .1 |
| P.SD       5.2         .0        1.86     .15       .30     .8    .78     .6 |
| S.SD       5.2         .0        1.86     .15       .30     .8    .78     .6 |
| MAX.      19.0      20.0        4.04    1.13      2.06    2.2   9.90    3.8 |
| MIN.       1.0      20.0       -3.98     .53       .38   -1.9    .13   -1.0 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .71 TRUE SD    1.72  SEPARATION  2.41  PERSON RELIABILITY  .85 |
|MODEL RMSE    .67 TRUE SD    1.74  SEPARATION  2.58  PERSON RELIABILITY  .87 |
| S.E. OF PERSON MEAN = .08                                                   |
-------------------------------------------------------------------------------
                DELETED:      8 PERSON
PERSON RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .89  SEM = 1.71

     SUMMARY OF 20 MEASURED ITEM
-------------------------------------------------------------------------------
|          TOTAL                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN     259.4     484.0         .00     .13      1.00     .1    .95    -.2 |
| P.SD     100.2         .0        1.61     .02       .05     .8    .12     .7 |
| S.SD     102.8         .0        1.66     .02       .05     .8    .12     .8 |
| MAX.     443.0     484.0        3.43     .19      1.14    2.4   1.21    1.7 |
| MIN.      53.0     484.0       -3.29     .12       .93   -1.2    .65   -1.8 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .13 TRUE SD    1.61  SEPARATION 12.13  ITEM   RELIABILITY  .99 |
|MODEL RMSE    .13 TRUE SD    1.61  SEPARATION 12.26  ITEM   RELIABILITY  .99 |
| S.E. OF ITEM MEAN = .37                                                     |
-------------------------------------------------------------------------------
ITEM RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000
```

## الملحق رقم (33)

## الدرجات المحولة بالواط

INPUT: 492 Person 20 Item REPORTED: 484 Person 19 Item 2 CATS WINSTEPS 4.0.1

Person: REAL SEP.: 2.36 REL.: .85 ... Item: REAL SEP.: 12.42 REL.: .99

Item STATISTICS: DISPLACEMENT ORDER

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEASUR-AL CORR. | PTMEASUR-AL EXP. | EXACT MATCH OBS% | EXACT MATCH EXP% | DISP LACE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 53 | 484 | 102.56 | 2.53 | 1.05 | .5 | .64 | -1.0 | .41 | .42 | 89.2 | 90.2 | .04 |
| 12 | 66 | 484 | 97.45 | 2.34 | 1.09 | 1.0 | 1.04 | .2 | .42 | .45 | 87.4 | 88.1 | .03 |
| 14 | 159 | 484 | 72.04 | 1.84 | .97 | -.4 | 1.00 | .0 | .59 | .58 | 80.3 | 79.9 | .02 |
| 13 | 170 | 484 | 69.59 | 1.82 | 1.05 | .8 | 1.06 | .4 | .58 | .59 | 76.6 | 79.4 | .01 |
| 18 | 207 | 484 | 61.69 | 1.77 | .93 | -1.1 | .85 | -1.2 | .64 | .61 | 80.7 | 78.3 | .01 |
| 16 | 211 | 484 | 60.86 | 1.77 | .94 | -1.0 | .78 | -1.7 | .64 | .61 | 79.1 | 78.2 | .01 |
| 9 | 212 | 484 | 60.65 | 1.77 | .98 | -.4 | .89 | -.8 | .63 | .61 | 78.5 | 78.2 | .01 |
| 11 | 237 | 484 | 55.48 | 1.76 | .99 | -.1 | 1.04 | .4 | .62 | .62 | 77.4 | 78.2 | .00 |
| 5 | 250 | 484 | 52.80 | 1.76 | 1.02 | .3 | .99 | .0 | .61 | .62 | 77.2 | 78.3 | .00 |
| 6 | 273 | 484 | 48.03 | 1.77 | .96 | -.7 | .95 | -.4 | .63 | .61 | 81.4 | 78.6 | .00 |
| 4 | 291 | 484 | 44.23 | 1.79 | 1.00 | .0 | 1.01 | .1 | .61 | .61 | 79.7 | 79.0 | -.01 |
| 1 | 294 | 484 | 43.59 | 1.79 | 1.03 | .5 | 1.03 | .3 | .59 | .61 | 78.3 | 79.0 | -.01 |
| 7 | 310 | 484 | 40.13 | 1.81 | 1.02 | .4 | 1.03 | .2 | .59 | .60 | 76.6 | 79.6 | -.01 |
| 2 | 328 | 484 | 36.11 | 1.85 | .95 | -.9 | .83 | -1.0 | .61 | .59 | 82.0 | 80.4 | -.01 |
| 3 | 332 | 484 | 35.19 | 1.86 | 1.05 | .8 | .95 | -.2 | .57 | .58 | 79.5 | 80.6 | -.01 |
| 8 | 336 | 484 | 34.27 | 1.87 | 1.04 | .6 | 1.13 | .7 | .56 | .58 | 81.6 | 80.8 | -.01 |
| 17 | 357 | 484 | 29.22 | 1.93 | 1.06 | .9 | 1.04 | .3 | .54 | .56 | 81.2 | 82.1 | -.02 |
| 15 | 432 | 484 | 5.66 | 2.55 | .99 | -.1 | .98 | .1 | .42 | .42 | 90.1 | 90.5 | -.03 |
| 19 | 443 | 484 | .45 | 2.79 | .98 | -.1 | .95 | .0 | .40 | .38 | 92.3 | 92.3 | -.03 |
| MEAN | 261.1 | 484.0 | 50.00 | 1.97 | 1.01 | .1 | .96 | -.2 | | | 81.5 | 81.7 | .00 |
| P.SD | 102.5 | .0 | 25.11 | .31 | .04 | .7 | .11 | .6 | | | 4.6 | 4.6 | .02 |

INPUT:492 Person 20 Item REPORTED:484 Person 19 Item 2 CATS WINSTEPS 4.0.1

Item FIT GRAPH: DISPLACEMENT ORDER

| ENTRY NUMBER | MEASURE - + | INFIT MEAN-SQUARE 0.0 1 2 | OUTFIT MEAN-SQUARE 0.0 1 2 | Item |
|---|---|---|---|---|
| 20 | \| * \| | : * : | : * . : | ITM20 |
| 12 | \| * \| | : * : | : * : | ITM12 |
| 14 | \| * \| | : *. : | : *. : | ITM14 |
| 13 | \| * \| | : * : | : * : | ITM13 |
| 18 | \| * \| | : *. : | : * . : | ITM18 |
| 16 | \| * \| | : *. : | : * . : | ITM16 |
| 9 | \| * \| | : *. : | : * . : | ITM09 |
| 11 | \| * \| | : *. : | : * : | ITM11 |
| 5 | \| * \| | : * : | : *. : | ITM05 |
| 6 | \| * \| | : *. : | : *. : | ITM06 |
| 4 | \| * \| | : *. : | : * : | ITM04 |
| 1 | \| * \| | : * : | : * : | ITM01 |
| 7 | \| * \| | : * : | : * : | ITM07 |
| 2 | \| * \| | : *. : | : * . : | ITM02 |
| 3 | \| * \| | : * : | : *. : | ITM03 |
| 8 | \| * \| | : * : | : .* : | ITM08 |
| 17 | \| * \| | : * : | : * : | ITM17 |
| 15 | \|* \| | : *. : | : *. : | ITM15 |
| 19 | \|* \| | : *. : | : *. : | ITM19 |

INPUT: 492 Person 20 Item REPORTED: 484 Person 19 Item 2 CATS WINSTEPS 4.0.1

Item CATEGORY/OPTION/DISTRACTOR FREQUENCIES: DISPLACEMENT ORDER

| ENTRY NUMBER | DATA CODE | SCORE VALUE | DATA COUNT | % | ABILITY MEAN | P.SD | S.E. MEAN | INFT MNSQ | OUTF MNSQ | PTMA CORR. | Item |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 0 | 0 | 431 | 89 | 50.20 | 26.95 | 1.30 | 1.1 | 1.1 | -.41 | ITM20 |
| | 1 | 1 | 53 | 11 | 88.07 | 16.73 | 2.32 | 1.0 | .6 | .41 | |
| 12 | 0 | 0 | 418 | 86 | 49.53 | 26.80 | 1.31 | 1.1 | 1.1 | -.42 | ITM12 |
| | 1 | 1 | 66 | 14 | 84.91 | 19.12 | 2.37 | 1.1 | 1.0 | .42 | |
| 14 | 0 | 0 | 325 | 67 | 42.54 | 24.02 | 1.33 | 1.0 | 1.0 | -.59 | ITM14 |
| | 1 | 1 | 159 | 33 | 78.50 | 20.99 | 1.67 | 1.0 | 1.0 | .59 | |
| 13 | 0 | 0 | 314 | 65 | 42.25 | 24.06 | 1.36 | 1.0 | 1.0 | -.58 | ITM13 |
| | 1 | 1 | 170 | 35 | 76.70 | 22.08 | 1.70 | 1.1 | 1.1 | .58 | |
| 18 | 0 | 0 | 277 | 57 | 38.57 | 22.56 | 1.36 | .9 | .9 | -.64 | ITM18 |
| | 1 | 1 | 207 | 43 | 75.47 | 21.23 | 1.48 | 1.0 | .8 | .64 | |
| 16 | 0 | 0 | 273 | 56 | 38.19 | 22.96 | 1.39 | 1.0 | .9 | -.64 | ITM16 |
| | 1 | 1 | 211 | 44 | 75.27 | 20.43 | 1.41 | .9 | .7 | .64 | |
| 9 | 0 | 0 | 272 | 56 | 38.52 | 23.05 | 1.40 | 1.0 | .9 | -.63 | ITM09 |
| | 1 | 1 | 212 | 44 | 74.66 | 21.22 | 1.46 | 1.0 | .9 | .63 | |
| 11 | 0 | 0 | 247 | 51 | 36.99 | 22.37 | 1.43 | 1.0 | .9 | -.62 | ITM11 |
| | 1 | 1 | 237 | 49 | 72.44 | 22.50 | 1.46 | 1.0 | 1.2 | .62 | |
| 5 | 0 | 0 | 234 | 48 | 36.30 | 21.74 | 1.42 | 1.0 | .8 | -.61 | ITM05 |
| | 1 | 1 | 250 | 52 | 71.25 | 23.43 | 1.49 | 1.1 | 1.2 | .61 | |
| 6 | 0 | 0 | 211 | 44 | 33.98 | 22.09 | 1.52 | 1.0 | 1.0 | -.63 | ITM06 |
| | 1 | 1 | 273 | 56 | 70.10 | 22.43 | 1.36 | 1.0 | .9 | .63 | |
| 4 | 0 | 0 | 193 | 40 | 33.05 | 21.99 | 1.59 | 1.0 | 1.0 | -.61 | ITM04 |
| | 1 | 1 | 291 | 60 | 68.48 | 23.20 | 1.36 | 1.0 | 1.1 | .61 | |
| 1 | 0 | 0 | 190 | 39 | 33.27 | 22.07 | 1.61 | 1.0 | 1.0 | -.59 | ITM01 |
| | 1 | 1 | 294 | 61 | 67.97 | 23.63 | 1.38 | 1.0 | 1.1 | .59 | |
| 7 | 0 | 0 | 174 | 36 | 31.82 | 21.21 | 1.61 | 1.0 | 1.0 | -.59 | ITM07 |
| | 1 | 1 | 310 | 64 | 67.00 | 24.05 | 1.37 | 1.1 | 1.1 | .59 | |
| 2 | 0 | 0 | 156 | 32 | 29.05 | 20.87 | 1.68 | 1.0 | .8 | -.61 | ITM02 |
| | 1 | 1 | 328 | 68 | 66.39 | 23.44 | 1.30 | .9 | .9 | .61 | |
| 3 | 0 | 0 | 152 | 31 | 30.39 | 21.77 | 1.77 | 1.1 | .9 | -.57 | ITM03 |
| | 1 | 1 | 332 | 69 | 65.32 | 24.32 | 1.34 | 1.0 | 1.0 | .57 | |
| 8 | 0 | 0 | 148 | 31 | 30.29 | 22.47 | 1.85 | 1.1 | 1.2 | -.56 | ITM08 |
| | 1 | 1 | 336 | 69 | 64.95 | 24.24 | 1.32 | 1.0 | 1.0 | .56 | |
| 17 | 0 | 0 | 127 | 26 | 28.69 | 21.78 | 1.94 | 1.1 | 1.0 | -.54 | ITM17 |
| | 1 | 1 | 357 | 74 | 63.48 | 24.93 | 1.32 | 1.0 | 1.0 | .54 | |
| 15 | 0 | 0 | 52 | 11 | 19.57 | 20.75 | 2.91 | 1.0 | 1.0 | -.42 | ITM15 |
| | 1 | 1 | 432 | 89 | 58.54 | 26.47 | 1.28 | 1.0 | 1.0 | .42 | |
| 19 | 0 | 0 | 41 | 8 | 17.22 | 19.49 | 3.08 | .9 | .9 | -.40 | ITM19 |
| | 1 | 1 | 443 | 92 | 57.79 | 26.80 | 1.27 | 1.0 | 1.0 | .40 | |

# ملحق رقم (34)

## معامل التمييز والتخمين

INPUT: 502 PERSON  20 ITEM  REPORTED: 502 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1

PERSON: REAL SEP.: 2.43  REL.: .86 ... ITEM: REAL SEP.: 12.03  REL.: .99

ITEM STATISTICS:  MISFIT ORDER

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEASUR-AL CORR. | PTMEASUR-AL EXP. | EXACT MATCH OBS% | EXACT MATCH EXP% | ESTIM DISCR | ASYMPTOTE LOWER | ASYMPTOTE UPPER | ITEM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 63 | 502 | 3.27 | .16 | 1.08 | .8 | 2.21 | 2.8 | A .41 | .45 | 88.4 | 89.4 | .90 | .01 | .97 | ITM20 |
| 10 | 235 | 502 | .45 | .12 | 1.14 | 2.3 | 1.19 | 1.6 | B .56 | .61 | 72.8 | 77.7 | .79 | .03 | .96 | ITM10 |
| 8 | 345 | 502 | -1.05 | .12 | 1.02 | .3 | 1.10 | .6 | C .57 | .58 | 82.1 | 80.3 | .95 | .00 | .99 | ITM08 |
| 12 | 71 | 502 | 3.08 | .15 | 1.07 | .8 | .89 | -.3 | D .45 | .47 | 87.6 | 88.0 | .97 | .00 | 1.00 | ITM12 |
| 17 | 366 | 502 | -1.37 | .13 | 1.04 | .6 | 1.00 | .1 | E .55 | .56 | 80.9 | 81.6 | .96 | .00 | .99 | ITM17 |
| 5 | 260 | 502 | .12 | .12 | 1.03 | .5 | 1.00 | .0 | F .60 | .61 | 77.2 | 77.8 | .96 | .03 | 1.00 | ITM05 |
| 13 | 175 | 502 | 1.27 | .12 | 1.03 | .5 | 1.02 | .2 | G .58 | .59 | 77.6 | 79.1 | .97 | .01 | 1.00 | ITM13 |
| 11 | 245 | 502 | .32 | .12 | .99 | -.1 | 1.02 | .2 | H .61 | .61 | 76.4 | 77.7 | 1.02 | .00 | 1.00 | ITM11 |
| 1 | 300 | 502 | -.41 | .12 | 1.01 | .2 | 1.01 | .1 | I .60 | .60 | 78.7 | 78.5 | .98 | .01 | 1.00 | ITM01 |
| 3 | 342 | 502 | -1.01 | .12 | 1.01 | .2 | .91 | -.5 | J .58 | .58 | 79.1 | 80.2 | .99 | .00 | 1.00 | ITM03 |
| 4 | 300 | 502 | -.41 | .12 | .99 | -.1 | 1.00 | .0 | j .60 | .60 | 79.1 | 78.5 | 1.00 | .00 | 1.00 | ITM04 |
| 7 | 319 | 502 | -.68 | .12 | 1.00 | .1 | 1.00 | .1 | i .59 | .60 | 77.0 | 79.1 | 1.00 | .03 | 1.00 | ITM07 |
| 15 | 439 | 502 | -2.81 | .16 | .99 | -.1 | .89 | -.2 | h .47 | .46 | 89.2 | 89.6 | 1.02 | .00 | 1.00 | ITM15 |
| 19 | 453 | 502 | -3.21 | .18 | .99 | .0 | .93 | -.1 | g .44 | .43 | 91.7 | 91.8 | 1.02 | .00 | 1.00 | ITM19 |
| 14 | 165 | 502 | 1.41 | .12 | .96 | -.6 | .98 | -.1 | f .59 | .58 | 81.7 | 79.7 | 1.04 | .00 | 1.00 | ITM14 |
| 9 | 219 | 502 | .66 | .12 | .97 | -.5 | .89 | -.9 | e .62 | .61 | 79.1 | 77.8 | 1.06 | .00 | 1.00 | ITM09 |
| 6 | 280 | 502 | -.14 | .12 | .94 | -1.0 | .92 | -.6 | d .63 | .61 | 81.3 | 78.1 | 1.08 | .00 | 1.00 | ITM06 |
| 16 | 218 | 502 | .68 | .12 | .94 | -1.1 | .79 | -1.7 | c .64 | .61 | 79.7 | 77.8 | 1.12 | .00 | 1.00 | ITM16 |
| 2 | 335 | 502 | -.91 | .12 | .93 | -1.1 | .83 | -1.1 | b .61 | .59 | 81.3 | 79.7 | 1.10 | .00 | 1.00 | ITM02 |
| 18 | 213 | 502 | .75 | .12 | .93 | -1.3 | .83 | -1.3 | a .63 | .60 | 79.1 | 77.9 | 1.12 | .00 | 1.00 | ITM18 |
| MEAN | 267.1 | 502.0 | .00 | .13 | 1.00 | .0 | 1.02 | -.1 | | | 81.0 | 81.0 | | | | |
| P.SD | 100.9 | .0 | 1.57 | .02 | .05 | .8 | .29 | 1.0 | | | 4.7 | 4.5 | | | | |

INPUT:502 PERSON 20 ITEM REPORTED:502 PERSON 20 ITEM 2 CATS WINSTEPS4.0.1

ITEM CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER

| ENTRY NUMBER | DATA CODE | SCORE VALUE | DATA COUNT | % | ABILITY MEAN | P.SD | S.E. MEAN | INFT MNSQ | OUTF MNSQ | PTMA CORR. | ITEM |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | A | 0 | 439 | 87 | -.08 | 1.81 | .09 | 1.1 | 1.1 | -.41 | ITM20 |
| | | 1 | 63 | 13 | 2.37 | 1.66 | .21 | 1.1 | 2.4 | .41 | |
| 10 | B | 0 | 267 | 53 | -.80 | 1.67 | .10 | 1.1 | 1.2 | -.56 | ITM10 |
| | | 1 | 235 | 47 | 1.40 | 1.60 | .10 | 1.1 | 1.2 | .56 | |
| 8 | C | 0 | 157 | 31 | -1.43 | 1.59 | .13 | 1.1 | 1.1 | -.57 | ITM08 |
| | | 1 | 345 | 69 | .99 | 1.63 | .09 | 1.0 | 1.0 | .57 | |
| 12 | D | 0 | 431 | 86 | -.13 | 1.81 | .09 | 1.1 | 1.1 | -.45 | ITM12 |
| | | 1 | 71 | 14 | 2.44 | 1.40 | .17 | 1.1 | .9 | .45 | |
| 17 | E | 0 | 136 | 27 | -1.55 | 1.56 | .13 | 1.1 | 1.0 | -.55 | ITM17 |
| | | 1 | 366 | 73 | .89 | 1.67 | .09 | 1.0 | 1.0 | .55 | |
| 5 | F | 0 | 242 | 48 | -.99 | 1.54 | .10 | 1.0 | .9 | -.60 | ITM05 |
| | | 1 | 260 | 52 | 1.37 | 1.61 | .10 | 1.1 | 1.2 | .60 | |
| 13 | G | 0 | 327 | 65 | -.60 | 1.65 | .09 | 1.0 | 1.0 | -.58 | ITM13 |
| | | 1 | 175 | 35 | 1.79 | 1.52 | .12 | 1.0 | 1.0 | .58 | |
| 11 | H | 0 | 257 | 51 | -.95 | 1.56 | .10 | 1.0 | .9 | -.61 | ITM11 |
| | | 1 | 245 | 49 | 1.47 | 1.55 | .10 | 1.0 | 1.2 | .61 | |
| 1 | I | 0 | 202 | 40 | -1.20 | 1.55 | .11 | 1.0 | 1.0 | -.60 | ITM01 |
| | | 1 | 300 | 60 | 1.20 | 1.60 | .09 | 1.0 | 1.1 | .60 | |
| 3 | J | 0 | 160 | 32 | -1.44 | 1.53 | .12 | 1.0 | .9 | -.58 | ITM03 |
| | | 1 | 342 | 68 | 1.01 | 1.64 | .09 | 1.0 | 1.0 | .58 | |
| 4 | j | 0 | 202 | 40 | -1.22 | 1.55 | .11 | 1.0 | 1.0 | -.60 | ITM04 |
| | | 1 | 300 | 60 | 1.21 | 1.58 | .09 | 1.0 | 1.0 | .60 | |
| 7 | i | 0 | 183 | 36 | -1.31 | 1.50 | .11 | 1.0 | .9 | -.59 | ITM07 |
| | | 1 | 319 | 64 | 1.12 | 1.63 | .09 | 1.0 | 1.1 | .59 | |
| 15 | h | 0 | 63 | 13 | -2.21 | 1.56 | .20 | 1.0 | .9 | -.47 | ITM15 |
| | | 1 | 439 | 87 | .58 | 1.76 | .08 | 1.0 | 1.0 | .47 | |
| 19 | g | 0 | 49 | 10 | -2.40 | 1.55 | .22 | 1.0 | .9 | -.44 | ITM19 |
| | | 1 | 453 | 90 | .52 | 1.79 | .08 | 1.0 | 1.0 | .44 | |
| 14 | f | 0 | 337 | 67 | -.58 | 1.65 | .09 | 1.0 | 1.0 | -.59 | ITM14 |
| | | 1 | 165 | 33 | 1.89 | 1.46 | .11 | 1.0 | 1.0 | .59 | |
| 9 | e | 0 | 283 | 56 | -.84 | 1.59 | .09 | 1.0 | .9 | -.62 | ITM09 |
| | | 1 | 219 | 44 | 1.62 | 1.48 | .10 | 1.0 | .9 | .62 | |
| 6 | d | 0 | 222 | 44 | -1.16 | 1.54 | .10 | .9 | .9 | -.63 | ITM06 |
| | | 1 | 280 | 56 | 1.33 | 1.53 | .09 | .9 | .9 | .63 | |
| 16 | c | 0 | 284 | 57 | -.87 | 1.58 | .09 | .9 | .9 | -.64 | ITM16 |
| | | 1 | 218 | 43 | 1.66 | 1.44 | .10 | .9 | .7 | .64 | |
| 2 | b | 0 | 167 | 33 | -1.48 | 1.49 | .12 | .9 | .8 | -.61 | ITM02 |
| | | 1 | 335 | 67 | 1.09 | 1.59 | .09 | .9 | .9 | .61 | |
| 18 | a | 0 | 289 | 58 | -.84 | 1.57 | .09 | .9 | .8 | -.63 | ITM18 |
| | | 1 | 213 | 42 | 1.68 | 1.47 | .10 | .9 | .8 | .63 | |

```
INPUT: 502 PERSON  20 ITEM  REPORTED: 502 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------

MOST MISFITTING RESPONSE STRINGS
ITEM     OUTMNSQ  |PERSON
                  |243314421  44221   2 1 4432 3 433 422421 43224422
                  |14111923562202102  3 47792286 39680749993545520427
                  |467073437550490879206925506358429083408607431123 27
               high----------------------------------------------------
20 ITM20    2.21 A|.......................1.1.1.1..1.......1........
10 ITM10    1.19 B|0.............................1..1...1.1..1......
 8 ITM08    1.10 C|.0........000.....................................1
12 ITM12     .89 D|.....................1....1.1.1.1.................
17 ITM17    1.00 E|.......0.0.........00.............................
 5 ITM05    1.00 F|..........0..0.....................1......1.....
13 ITM13    1.02 G|..0..............................11...1..........
11 ITM11    1.02 H|..............................11..1.........1...
 1 ITM01    1.01 I|.....0.......0.0.............................1....
 3 ITM03     .91 J|.............0.00.......................1.......
 4 ITM04    1.00 j|..0.....0..0.....0...0.........................1.
 7 ITM07    1.00 i|....0.0.................................11.....1..
15 ITM15     .89 h|...............0.....0.0..0...0.................
19 ITM19     .93 g|............0........0.0.......................
14 ITM14     .98 f|.......................1.......1...11..........
 9 ITM09     .89 e|.........................................1.......
 6 ITM06     .92 d|...0.............................................
 2 ITM02     .83 b|........0.........0....0..............1.........
18 ITM18     .83 a|.............................1.................
                  |-----------------------------------------------low
                  |24331442162442212922617443283843384224213432244227
                  |14111923555202107  3 42792236 39600749990545520427
                  |467073437  04908   0 9 5506 5 429 834086 74311232


INPUT: 502 PERSON  20 ITEM  REPORTED: 502 PERSON  20 ITEM  2 CATS WINSTEPS 4.0.1
-------------------------------------------------------------------------------

MOST UNEXPECTED RESPONSES
ITEM     MEASURE  |PERSON
                  |243314421  44221   2 1 4432 3 433 422421 43224422
                  |14111923562202102  3 47792286 39680749993545520427
                  |467073437550490879206925506358429083408607431123 27
               high----------------------------------------------------
19 ITM19   -3.21 g|............0........0.0.......................
15 ITM15   -2.81 h|..............0.....0.0..0...0.................
17 ITM17   -1.37 E|.......0.0.........00.............................
 8 ITM08   -1.05 C|.0........000.....................................1
 3 ITM03   -1.01 J|.............0.00.......................1.......
 2 ITM02    -.91 b|........0.........0....0..............1.........
 7 ITM07    -.68 i|....0.0.................................11.....1..
 1 ITM01    -.41 I|.....0.......0.0.............................1....
 4 ITM04    -.41 j|..0.....0..0.....0...0.........................1.
 6 ITM06    -.14 d|...0.............................................
 5 ITM05     .12 F|..........0..0.....................1......1.....
11 ITM11     .32 H|..............................11..1.........1...
10 ITM10     .45 B|0.............................1..1...1.1..1......
 9 ITM09     .66 e|.........................................1.......
18 ITM18     .75 a|.............................1.................
13 ITM13    1.27 G|..0..............................11...1..........
14 ITM14    1.41 f|.......................1.......1...11..........
12 ITM12    3.08 D|.....................1....1.1.1.1.................
20 ITM20    3.27 A|.......................1.1.1.1..1.......1........
                  |-----------------------------------------------low
                  |24331442162442212922617443283843384224213432244227
```

INPUT:502 PERSON 20 ITEM REPORTED:502 PERSON 20 ITEM 2 CATS WINSTEPS 4.0.1

MOST UNEXPECTED RESPONSES

| DATA | OBSERVED | EXPECTED | RESIDUAL | ST.RES. | MEASDIFF | ITEM | PERSON | ITEM | PERSON |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | .00 | 1.00 | 23.15 | -6.28 | 20 | 457 | ITM20 | 457 |
| 0 | 0 | .99 | -.99 | -13.05 | 5.14 | 19 | 229 | ITM19 | 229 |
| 1 | 1 | .01 | .99 | 11.08 | -4.81 | 20 | 80 | ITM20 | 80 |
| 0 | 0 | .99 | -.99 | -10.66 | 4.73 | 15 | 108 | ITM15 | 108 |
| 1 | 1 | .01 | .99 | 10.03 | -4.61 | 12 | 392 | ITM12 | 392 |
| 1 | 1 | .01 | .99 | 9.32 | -4.46 | 20 | 434 | ITM20 | 434 |
| 1 | 1 | .01 | .99 | 8.35 | -4.24 | 11 | 402 | ITM11 | 402 |
| 0 | 0 | .98 | -.98 | -7.77 | 4.10 | 8 | 446 | ITM08 | 446 |
| 0 | 0 | .98 | -.98 | -7.60 | 4.06 | 19 | 72 | ITM19 | 72 |
| 0 | 0 | .98 | -.98 | -7.29 | 3.97 | 15 | 149 | ITM15 | 149 |
| 1 | 1 | .02 | .98 | 6.73 | -3.81 | 14 | 490 | ITM14 | 490 |
| 0 | 0 | .98 | -.98 | -6.66 | 3.79 | 17 | 233 | ITM17 | 233 |
| 0 | 0 | .98 | -.98 | -6.66 | 3.79 | 17 | 65 | ITM17 | 65 |
| 0 | 0 | .98 | -.98 | -6.45 | 3.73 | 7 | 117 | ITM07 | 117 |
| 1 | 1 | .02 | .98 | 6.28 | -3.68 | 9 | 344 | ITM09 | 344 |
| 1 | 1 | .02 | .98 | 6.27 | -3.67 | 13 | 196 | ITM13 | 196 |
| 1 | 1 | .03 | .97 | 5.96 | -3.57 | 20 | 365 | TM20 | 365 |
| 0 | 0 | .97 | -.97 | -5.87 | 3.54 | 10 | 214 | ITM10 | 214 |
| 1 | 1 | .03 | .97 | 5.78 | -3.51 | 1 | 421 | ITM01 | 421 |
| 1 | 1 | .03 | .97 | 5.78 | -3.51 | 4 | 222 | ITM04 | 222 |
| 0 | 0 | .97 | -.97 | -5.67 | 3.47 | 8 | 25 | ITM08 | 25 |
| 0 | 0 | .97 | -.97 | -5.65 | 3.46 | 4 | 317 | ITM04 | 317 |
| 1 | 1 | .03 | .97 | 5.65 | -3.46 | 10 | 253 | ITM10 | 253 |
| 1 | 1 | .03 | .97 | 5.40 | -3.37 | 12 | 8 | ITM12 | 8 |
| 0 | 0 | .97 | -.97 | -5.35 | 3.35 | 15 | 475 | ITM15 | 475 |
| 1 | 1 | .03 | .97 | 5.32 | -3.34 | 14 | 244 | ITM14 | 244 |
| 0 | 0 | .97 | -.97 | -5.26 | 3.32 | 2 | 157 | ITM02 | 157 |
| 1 | 1 | .04 | .96 | 5.19 | -3.29 | 20 | 495 | ITM20 | 495 |
| 1 | 1 | .04 | .96 | 5.19 | -3.29 | 20 | 226 | ITM20 | 226 |
| 1 | 1 | .04 | .96 | 5.07 | -3.25 | 7 | 243 | ITM07 | 243 |
| 1 | 1 | .04 | .96 | 4.95 | -3.20 | 13 | 408 | ITM13 | 408 |
| 1 | 1 | .04 | .96 | 4.95 | -3.20 | 13 | 273 | ITM13 | 273 |
| 0 | 0 | .96 | -.96 | -4.94 | 3.19 | 6 | 310 | ITM06 | 310 |
| 0 | 0 | .96 | -.96 | -4.94 | 3.19 | 19 | 495 | ITM19 | 495 |
| 1 | 1 | .04 | .96 | 4.79 | -3.13 | 5 | 251 | ITM05 | 251 |
| 1 | 1 | .04 | .96 | 4.70 | -3.09 | 12 | 320 | ITM12 | 320 |
| 1 | 1 | .04 | .96 | 4.70 | -3.09 | 12 | 83 | ITM12 | 83 |
| 0 | 0 | .96 | -.96 | -4.70 | 3.09 | 7 | 424 | ITM07 | 424 |
| 0 | 0 | .95 | -.95 | -4.43 | 2.98 | 8 | 420 | ITM08 | 420 |
| 0 | 0 | .95 | -.95 | -4.43 | 2.98 | 8 | 404 | ITM08 | 404 |
| 1 | 1 | .05 | .95 | 4.36 | -2.95 | 14 | 369 | ITM14 | 369 |
| 0 | 0 | .95 | -.95 | -4.34 | 2.93 | 3 | 210 | ITM03 | 210 |
| 0 | 0 | .95 | -.95 | -4.34 | 2.93 | 3 | 27 | ITM03 | 27 |
| 0 | 0 | .95 | -.95 | -4.34 | 2.93 | 3 | 9 | ITM03 | 9 |
| 0 | 0 | .95 | -.95 | -4.25 | 2.89 | 17 | 230 | ITM17 | 230 |
| 0 | 0 | .95 | -.95 | -4.25 | 2.89 | 17 | 6 | ITM17 | 6 |
| 1 | 1 | .05 | .95 | 4.20 | -2.87 | 8 | 77 | ITM08 | 77 |
| 1 | 1 | .05 | .95 | 4.17 | -2.85 | 10 | 298 | ITM10 | 298 |
| 1 | 1 | .05 | .95 | 4.17 | -2.85 | 10 | 30 | ITM10 | 30 |
| 0 | 0 | .94 | -.94 | -4.12 | 2.83 | 2 | 2 | ITM02 | 2 |

# الملحق رقم (35)

## التحويل للدرجة التائية

| SCORE | MEASURE | S.E. | INFO | NORMED | S.E.NORMED | FREQ | FREQ% | CUM.FREQ | CUMFREQ% | PERCENTILE |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | -3.49 | 18.94 | 0.28 | 204 | 99 | 0 | 0 | 0 | 0 | 0 |
| 1 | 10.33 | 11.2 | 0.8 | 277 | 59 | 12 | 2.5 | 12 | 2.5 | 1 |
| 2 | 19.74 | 8.59 | 1.35 | 326 | 45 | 14 | 2.9 | 26 | 5.4 | 4 |
| 3 | 26.06 | 7.41 | 1.82 | 359 | 39 | 23 | 4.8 | 49 | 10.1 | 8 |
| 4 | 30.99 | 6.68 | 2.24 | 385 | 35 | 36 | 7.4 | 85 | 17.6 | 14 |
| 5 | 35.11 | 6.19 | 2.61 | 407 | 32 | 23 | 4.8 | 108 | 22.3 | 20 |
| 6 | 38.73 | 5.87 | 2.9 | 426 | 31 | 25 | 5.2 | 133 | 27.5 | 25 |
| 7 | 42.04 | 5.66 | 3.12 | 443 | 30 | 28 | 5.8 | 161 | 33.3 | 30 |
| 8 | 45.17 | 5.54 | 3.26 | 459 | 29 | 28 | 5.8 | 189 | 39 | 36 |
| 9 | 48.21 | 5.49 | 3.32 | 475 | 29 | 24 | 5 | 213 | 44 | 42 |
| 10 | 51.22 | 5.5 | 3.31 | 491 | 29 | 23 | 4.8 | 236 | 48.8 | 46 |
| 11 | 54.28 | 5.58 | 3.22 | 507 | 29 | 26 | 5.4 | 262 | 54.1 | 51 |
| 12 | 57.46 | 5.72 | 3.05 | 524 | 30 | 30 | 6.2 | 292 | 60.3 | 57 |
| 13 | 60.87 | 5.96 | 2.82 | 542 | 31 | 41 | 8.5 | 333 | 68.8 | 65 |
| 14 | 64.62 | 6.31 | 2.51 | 561 | 33 | 31 | 6.4 | 364 | 75.2 | 72 |
| 15 | 68.92 | 6.84 | 2.14 | 584 | 36 | 31 | 6.4 | 395 | 81.6 | 78 |
| 16 | 74.12 | 7.62 | 1.72 | 611 | 40 | 34 | 7 | 429 | 88.6 | 85 |
| 17 | 80.82 | 8.83 | 1.28 | 646 | 46 | 36 | 7.4 | 465 | 96.1 | 92 |
| 18 | 90.66 | 11.39 | 0.77 | 698 | 60 | 18 | 3.7 | 483 | 99.8 | 98 |
| 19 | 104.76 | 19.03 | 0.28 | 772 | 100 | 1 | 0.2 | 484 | 100 | 99 |

# الملحق رقم (36)

## القدرة بالواط

| SCORE | MEASURE | S.E. | INFO | NORMED | S.E.NORMED | FREQ | FREQ% | CUM.FREQ | CUMFREQ% | PERCENTILE |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | -30.24 | 28.41 | 0.28 | 204 | 99 | 0 | 0 | 0 | 0 | 0 |
| 1 | -9.51 | 16.8 | 0.8 | 277 | 59 | 12 | 2.5 | 12 | 2.5 | 1 |
| 2 | 4.61 | 12.89 | 1.35 | 326 | 45 | 14 | 2.9 | 26 | 5.4 | 4 |
| 3 | 14.09 | 11.11 | 1.82 | 359 | 39 | 23 | 4.8 | 49 | 10.1 | 8 |
| 4 | 21.48 | 10.01 | 2.24 | 385 | 35 | 36 | 7.4 | 85 | 17.6 | 14 |
| 5 | 27.66 | 9.29 | 2.61 | 407 | 32 | 23 | 4.8 | 108 | 22.3 | 20 |
| 6 | 33.1 | 8.8 | 2.9 | 426 | 31 | 25 | 5.2 | 133 | 27.5 | 25 |
| 7 | 38.07 | 8.49 | 3.12 | 443 | 30 | 28 | 5.8 | 161 | 33.3 | 30 |
| 8 | 42.76 | 8.31 | 3.26 | 459 | 29 | 28 | 5.8 | 189 | 39 | 36 |
| 9 | 47.31 | 8.23 | 3.32 | 475 | 29 | 24 | 5 | 213 | 44 | 42 |
| 10 | 51.83 | 8.25 | 3.31 | 491 | 29 | 23 | 4.8 | 236 | 48.8 | 46 |
| 11 | 56.42 | 8.36 | 3.22 | 507 | 29 | 26 | 5.4 | 262 | 54.1 | 51 |
| 12 | 61.2 | 8.58 | 3.05 | 524 | 30 | 30 | 6.2 | 292 | 60.3 | 57 |
| 13 | 66.3 | 8.94 | 2.82 | 542 | 31 | 41 | 8.5 | 333 | 68.8 | 65 |
| 14 | 71.93 | 9.47 | 2.51 | 561 | 33 | 31 | 6.4 | 364 | 75.2 | 72 |
| 15 | 78.39 | 10.26 | 2.14 | 584 | 36 | 31 | 6.4 | 395 | 81.6 | 78 |
| 16 | 86.18 | 11.43 | 1.72 | 611 | 40 | 34 | 7 | 429 | 88.6 | 85 |
| 17 | 96.22 | 13.25 | 1.28 | 646 | 46 | 36 | 7.4 | 465 | 96.1 | 92 |
| 18 | 110.99 | 17.08 | 0.77 | 698 | 60 | 18 | 3.7 | 483 | 99.8 | 98 |
| 19 | 132.13 | 28.55 | 0.28 | 772 | 100 | 1 | 0.2 | 484 | 100 | 99 |

# الملحق رقم (37)

## دالة المعلومات

| SCORE | MEASURE | S.E. | INFO | NORMED | S.E.NORMED | FREQ | FREQ% | CUM.FREQ | CUMFREQ% | PERCENTILE |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | -5.35 | 1.89 | 0.28 | 204 | 99 | 0 | 0 | 0 | 0 | 0 |
| 1 | -3.97 | 1.12 | 0.8 | 277 | 59 | 12 | 2.5 | 12 | 2.5 | 1 |
| 2 | -3.03 | 0.86 | 1.35 | 326 | 45 | 14 | 2.9 | 26 | 5.4 | 4 |
| 3 | -2.39 | 0.74 | 1.82 | 359 | 39 | 23 | 4.8 | 49 | 10.1 | 8 |
| 4 | -1.9 | 0.67 | 2.24 | 385 | 35 | 36 | 7.4 | 85 | 17.6 | 14 |
| 5 | -1.49 | 0.62 | 2.61 | 407 | 32 | 23 | 4.8 | 108 | 22.3 | 20 |
| 6 | -1.13 | 0.59 | 2.9 | 426 | 31 | 25 | 5.2 | 133 | 27.5 | 25 |
| 7 | -0.8 | 0.57 | 3.12 | 443 | 30 | 28 | 5.8 | 161 | 33.3 | 30 |
| 8 | -0.48 | 0.55 | 3.26 | 459 | 29 | 28 | 5.8 | 189 | 39 | 36 |
| **9** | **-0.18** | **0.55** | **3.32** | **475** | **29** | **24** | **5** | **213** | **44** | **42** |
| 10 | 0.12 | 0.55 | 3.31 | 491 | 29 | 23 | 4.8 | 236 | 48.8 | 46 |
| 11 | 0.43 | 0.56 | 3.22 | 507 | 29 | 26 | 5.4 | 262 | 54.1 | 51 |
| 12 | 0.75 | 0.57 | 3.05 | 524 | 30 | 30 | 6.2 | 292 | 60.3 | 57 |
| 13 | 1.09 | 0.6 | 2.82 | 542 | 31 | 41 | 8.5 | 333 | 68.8 | 65 |
| 14 | 1.46 | 0.63 | 2.51 | 561 | 33 | 31 | 6.4 | 364 | 75.2 | 72 |
| 15 | 1.89 | 0.68 | 2.14 | 584 | 36 | 31 | 6.4 | 395 | 81.6 | 78 |
| 16 | 2.41 | 0.76 | 1.72 | 611 | 40 | 34 | 7 | 429 | 88.6 | 85 |
| 17 | 3.08 | 0.88 | 1.28 | 646 | 46 | 36 | 7.4 | 465 | 96.1 | 92 |
| 18 | 4.07 | 1.14 | 0.77 | 698 | 60 | 18 | 3.7 | 483 | 99.8 | 98 |
| 19 | 5.48 | 1.9 | 0.28 | 772 | 100 | 1 | 0.2 | 484 | 100 | 99 |