بسم الله الرحمن الرحيم

جامعة اليرموك

كلية التربية

قسم الإرشاد وعلم النفس التربوي

بناء بنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية

" نموذج مقترح "

إعداد

عبد الحكيم علي الحسين مهيدات

إشراف

الأستاذ الدكتور أحمد سليمان عودة

الفصل الصيفي

11/ اب/ 2005 م

6 / رجب / 1426 ه

بناء بنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية

" نموذج مقترح "

إعداد

عبد الحكيم علي الحسين مهيدات

بكالوريوس رياضيات، جامعة اليرموك، 1987 م

ماجستير قياس وتقويم، جامعة اليرموك، 1990 م

قدمت هذه الأطروحة استكمالا لمتطلبات الحصول على درجة الدكتوراة في تخصص القياس والتقويم

في جامعة اليرموك، اربد، الأردن.

وافق عليها

الأستاذ الدكتور أحمد سليمان عودة ..... رئيسا

أستاذ البحث والتقويم التربوي في جامعة اليرموك

الأستاذ الدكتورعبدالرحمن عبد الرحيم عدس ..... عضوا

أستاذ القياس والتقويم التربوي في جامعة عمان العربية للدراسات العليا

الأستاذ الدكتور يوسف محمدالسوالمة ..... عضوا

أستاذ القياس والاحصاء التربوي في جامعة اليرموك

الأستاذ الدكتور ضرار محمد جرادات ..... عضوا

أستاذ القياس والتقويم التربوي في جامعة اليرموك

الدكتورة امل عبدالله خصاونة ..... عضوا

أستاذ مشارك في مناهج الرياضيات واساليب تدريسها في جامعة اليرموك

11/ اب/ 2005 م

6/ رجب / 1426 ه

ISN=662387

مكتبة جامعة اليرموك
رقم التسلسل: 009827
التاريخ: 10-5-2007
رقم التصنيف:

جامعة اليرموك - المكتبة
559826

الإهداء

إلى والديّ الحبيبين متعهما الله بالصحة والعافية.

إلى رفيقة عمري، وفلذات كبدي علي وشيراز وشادن وشجن الذين صبروا معي واحتملوا انشغالي عنهم.

إلى اخوتي وأخواتي الذين شدّوا أزري.

أهدي جهدي هذا.

الباحث

بسم الله الرحمن الرحيم

شكر وتقدير

لا يسعني وقد أعانني الله سبحانه وتعالى على انجاز هذا العمل المتواضع الا أن اتقدم بالشكر العطر والثناء الجميل لأستاذي الدكتور أحمد سليمان عودة المشرف على هذا البحث المواكب لخطواته ، الذي كان لي مرشدا وموجها ومقوما.

كما لا يفوتني أن أتقدم بجزيل شكري وتقديري للأستاذين الكريمين الدكتور يوسف محمد السوالمة ، والدكتور ضرار محمد جرادات على ما أسدوا لي من نصح وتوجيه .

مع خالص شكري لكل من الأستاذ الدكتور عبد الرحمن عدس والدكتورة أمل عبدالله خصاونة لتشريفهم إياي بقبول العضوية في لجنة مناقشة البحث.

أتوجه بجزيل شكري وامتناني إلى والدي واخواني الذين قدموا لي العون المادي والمعنوي، والى زوجتي وابنائي الذين وقفوا إلى جانبي وتحملوا العناء أثناء دراستي.

كما اشكر كل من ساهم بإخراج هذه الرسالة وتنفيذ الإجراءات المتعلقة بها.واتوجه بالشكر الجميل للأستاذ الدكتور عدنان عوض على ما قدمه لي من مساعدة وتوجيه.واتوجه بالشكر للزملاء المبرمجين طارق وعبدالسلام ونشأت ومحمد وامجد.

المحتويات

الموضوع الصفحة

الإهداء.....ث

شكر وتقدير.....ج

المحتوى.....ح

قائمة الجداول.....ذ

قائمة الأشكال.....س

الملاحق.....ط

الملخص.....ظ

الفصل الأول : خلفية إعداد بنك الأسئلة وأهميته.....1

تمهيد.....1

هدف الدراسة.....9

أسئلة الدراسة.....9

تعريف المصطلحات.....10

الفصل الثاني: أدبيات بنوك الأسئلة.....11

الجزء الأول: الإطار النظري لمكونات بنك الأسئلة وعناصره.....11

تعريف بنك الأسئلة.....11

لماذا بنوك الأسئلة؟.....12

أنواع بنوك الأسئلة.....14

أهمية بنوك الأسئلة.....14

مساوئ بنوك الأسئلة.....16

نموذج بنك الأسئلة.....16

المرحلة الأولى.....17
المرحلة الثانية.....18
المرحلة الثالثة.....18
المرحلة الرابعة.....19
المرحلة الخامسة.....20
إدارة وتشغيل بنوك الأسئلة.....21
سحب الاختبارات من بنك الأسئلة.....22
نظريات القياس وبنوك الأسئلة.....25
نظرية القياس التقليدية.....25
تحليل الفقرات الإختبارية حسب النظرية التقليدية.....26
نظرية الاستجابة للفقرة.....28
نماذج نظرية الاستجابة للفقرة.....29
افتراضات نظرية الاستجابة للفقرة.....33
دالة معلومة الفقرة.....35
خصائص اقتران معلومات الاختبار.....37
تدريج الاختبار.....38
دمج فقرات اختبارين.....38
دمج فقرات اختبارين باستخدام مجموعة من الفقرات المشتركة.....40
التحقق من خصائص اختبار مستل من بنك الأسئلة.....39
تخزين الفقرات.....41
النموذج ثنائي المعلمة.....42
مبررات استخدام النموذج ثنائي المعلمة.....43
فوائد النموذج ثنائي المعلمة.....44
محددات النموذج اللوغاريتمي ثنائي المعلمة.....45
الجزء الثاني: الدراسات السابقة.....45
أولا: الدراسات التي تناولت خطوات بناء بنوك الأسئلة.....45

ثانيا: الدراسات التي تناولت المقارنة بين خصائص الفقرات التي تم تحليلها
وفق النظرية التقليدية ونماذج نظرية الاستجابة للفقرة..... 51
الفصل الثالث:الطريقة والإجراءات..... 58
مجتمع الدراسة..... 58
عينة الدراسة..... 58
نماذج الاختبار..... 62
إجراءات الدراسة..... 63
الفصل الرابع: النتائج ومناقشتها..... 69
أولا: الإجابة عن السؤال الأول..... 69
تدريج الفقرات..... 87
الكفاءة النسبية..... 91
وصف البنك..... 92
تقويم البنك وضبط نوعيته..... 103
ثانيا: الإجابة عن السؤال الثاني..... 104
ثالثا: الإجابة عن السؤال الثالث..... 106
رابعا: الإجابة عن السؤال الرابع..... 111
الاستنتاجات والتوصيات..... 119
المراجع العربية..... 122
المراجع الأجنبية..... 123
الملحق(أ) 127
الملحق(ب) 131
الملخص بالإنجليزية..... 135

قائمة الجداول

الجدول الصفحة

جدول1: تصنيف الأهداف التعليمية في بعدين.....7

جدول2: تصنيف الأهداف التعليمية في وزارة التربية والتعليم/ الأردن.....7

جدول3: تصنيف الأهداف التعليمية للرياضيات.....8

جدول4: النموذج التصوري لنموذج بنك الأسئلة.....17

جدول5: توزيع أفراد مجتمع الدراسة وفق متغير الجنس والمسار التعليمي.....58

جدول6: مدارس الذكور وأعداد الطلبة فيها.....59

جدول7: مدارس الإناث وأعداد الطالبات فيها.....60

جدول8: عدد الطلاب المطلوب للعينة.....61

جدول9: توزيع أفراد عينة الدراسة الإناث واسماء المدارس والمسار التعليمي.....61

جدول10: توزيع أفراد عينة الدراسة الذكور واسماء المدارس والمسار التعليمي.....62

جدول11: النماذج الاختبارية وعدد الفقرات في كل منها.....63

جدول12: جدول مواصفات الاختبار.....64

جدول13: أرقام الفقرات الرابطة في النماذج الست.....65

جدول14: عدد أفراد الدراسة موزعين وفق صور الاختبار.....67

جدول15: متوسطات الصعوبة و متوسطات التمييز للفقرات المحسوبة
وفق النظرية الكلاسيكية..... 70

جدول16: لتوزيع التكراري لصعوبة الفقرات وفق النظرية الكلاسيكية للصور
الاختبارية الست.....70

جدول17: التوزيع التكراري لتمييز الفقرات وفق النظرية الكلاسيكية للصور
الاختبارية الست..... 71

جدول18: ملخص متوسطات الصعوبة والتمييز للفقرات المحسوبة وفق
نظرية استجابة الفقرة.....72
جدول19: التوزيع التكراري لصعوبة الفقرات وفق نظرية استجابة الفقرة للصور
الاختبارية الست.....72
جدول20: التوزيع التكراري لتمييز الفقرات وفق نظرية استجابة الفقرة للصور
الاختبارية الست.....73
جدول21: أرقام الفقرات المحذوفة باستخدام نظرية استجابة الفقرة، وكذلك
باستخدام النظرية الكلاسيكية للصور الاختبارية الست.....74
جدول22:معاملات الصعوبة*b ومعاملات التمييز *a والحد الأعلى لدالة معلومة الفقرة
لفقرات البنك بعد التدريج.....88
جدول23: معلومات الاختبار والكفاءة النسبية.....92
جدول24: أرقام الفقرات التي تم اختيارها مع المعلومات الخاصة بها.....105
جدول25: المؤشرات الإحصائية للفقرات المشتركة قيم الصعوبة وقيم التمييز وفق
النظرية التقليدية.....106
جدول26: مصفوفة معاملات الارتباط بين متوسطات صعوبة الفقرات المشتركة وفق
النظرية التقليدية..... 107
جدول27: مصفوفة معاملات الارتباط بين متوسطات معاملات التمييز للفقرات
المشتركة وفق النظرية التقليدية.....107
جدول28: قيم معاملات الصعوبة b والتمييز a للفقرات المشتركة وفق
نظرية الاستجابة للفقرة.....108
جدول29: مصفوفة معاملات الارتباط بين قيم معاملات الصعوبة وفق
نظرية الاستجابة للفقرة.....108
جدول30: مصفوفة معاملات الارتباط بين قيم معاملات التمييز وفق
نظرية الاستجابة للفقرة.....108
جدول31: معالم الفقرات المشتركة وفق نظرية الاستجابة للفقرة ، بالإضافة
إلى إحصائيات الفقرات نفسها وفق النظرية التقليدية.....109
جدول32: مصفوفة معاملات الارتباط لمعاملات صعوبة الفقرات المشتركة

وفق النظرية التقليدية، ونظرية الاستجابة للفقرة..... 110
جدول33: ملخص معاملات الارتباط بين معاملات التمييز للفقرات المشتركة
وفق النظرية التقليدية، ونظرية الاستجابة للفقرة..... 111

جدول34: ملخص لمتوسط الانحراف والمدى المقبول للمقارنة والنسبة
وفق نظرية استجابة الفقرة..... 111
جدول35: ملخص المعلومات حول الاختبارات المسحوبة.....112
جدول36: نتائج تحليل التباين لفريدمان.....112
جدول37:ملخص للفروق بين قدرات المفحوصين ونسبها بين أزواج الاختبارات..... 113
جدول38:ملخص معاملات الارتباط بين قدرات المفحوصين على الاختبارات الخمس..... 115
جدول39 :توزيع قدرات الطلبة على الفئات المختلفة للاختبار السهل والاختبار الصعب.....118
جدول40: عدد حالات التوافق بين قدرات المفحوصين على الاختبارات المختلفة
ونسبها.....119
جدول41:مصفوفة معاملات الارتباط بين الاختبارات الخمس.....119

صفحة قائمة الأشكال

الشكل الصفحة

شكل1: العلاقة بين القدرة θ واحتمال الحصول على الإجابة الصحيحة
لفقرتين مختلفتين في بارا متر التخمين.....31
شكل2: منحنى خصائص الفقرة لفقرتين مختلفتين في التمييز.....32
شكل3: منحنى خصائص الفقرة لفقرتين مختلفتين في الصعوبة
ولهما نفس درجة التمييز..... 33
شكل4: نموذج يوضح منحنى خصائص الفقرة.....35
شكل5: منحنى الفقرة التي تميز بشكل تام عندθ=1.5.....37
شكل6: ثلاث منحنيات لخاصية الفقرة مختلفة في درجة صعوبتها وقوة تمييزها.....43
شكل7: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الأول.....75
شكل8: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الثاني 75
شكل9: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الثالث 76
شكل10: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الرابع 76
شكل11: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الخامس 77

شكل12: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار السادس 77
شكل13: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الأول 78
شكل14: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الثاني 78
شكل15: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الثالث 79
شكل16: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الرابع 79
شكل17: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الخامس 80
شكل18: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار السادس 80
شكل19: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الأول 82
شكل20: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الثاني..... 82
شكل21: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الثالث 83
شكل22: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الرابع..... 83
شكل23: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار الخامس..... 84
شكل24: نسبة الاتفاق بين نموذج ثنائي المعلمة والنظرية التقليدية
في الاحتفاظ بالفقرات للاختبار السادس 84
شكل25: لوحة انتشار بين قيم الصعوبة وفق النموذج الثنائي وقيم

الصعوبة وفق النظرية التقليدية قبل حذف الفقرة.....85
شكل26: لوحة انتشار بين قيم الصعوبة وفق النموذج الثنائي وقيم
الصعوبة وفق النظرية التقليدية بعد حذف الفقرة.....85
شكل27: لوحة انتشار بين قيم معاملات التمييز وفق النموذج الثنائي وقيم
التمييز وفق النظرية التقليدية قبل حذف الفقرة.....86
شكل28: معامل الارتباط بين قيم معاملات التمييز وفق النموذج الثنائي وقيم
معاملات التمييز وفق النظرية التقليدية بعد حذف الفقرة.....86
شكل29: منحنى افضل فقرة.....90
شكل30:منحنى فقرة ضعيفة..... 90
شكل31: منحنى فقرة جيدة..... 91
شكل32: واجهة قاعدة البيانات وموجودات الجداول.....93
شكل33: تصميم حقل ملف الفقرات.....94
شكل34: واجهة قاعدة البيانات ومحتويات الاستعلام.....95
شكل35: جدول المعلومات التي تم الاستعلام عنها.....96
شكل36: واجهة التقارير ومحتوياتها.....96
شكل37: واجهة النماذج ومحتوياتها..... 97
شكل38: الشاشة الرئيسة لبنك الأسئلة..... 97
شكل39: نافذة الفقرات ومحتوياتها..... 98
شكل40: نافذة الفقرة ومحتوياتها.....98
شكل41: نافذة فقرة جديدة..... 99
شكل42: واجهة الاستعلام ومكوناتها.....100
شكل43أ: الحد الأدنى والحد الأعلى لصعوبة الفقرات.....100
شكل43ب: الحد الأدنى والحد الأعلى لتمييز الفقرات.....101
شكل43ج: الحد الأدنى والحد الأعلى لدالة معلومة الفقرات.....101
شكل43د: نوع الفقرة.....101
شكل44: أرقام الفقرات المطلوبة ضمن المواصفات المحددة.....102

شكل45: شاشة إدخال أرقام الفقرات لكل نوع في التقرير المخصص له.....102
شكل46:التمثيل البياني لتوزيع قدرات المفحوصين على الاختبار الوسط.....114
شكل47: التمثيل البياني لتوزيع قدرات المفحوصين على الاختبار الصعب.....114
شكل48: التمثيل البياني لتوزيع قدرات المفحوصين على الاختبار السهل.....114
شكل49: التمثيل البياني لتوزيع قدرات المفحوصين على اختبار الفقرات الفردية.....114
شكل50: التمثيل البياني لتوزيع قدرات المفحوصين على اختبار الفقرات الزوجية.....114
شكل51: ا لتوزيع التكراري للرتب المئينية لقدرات المفحوصين على الاختبار السهل والاختبار الوسط..... 115
شكل52: التوزيع التكراري للرتب المئينية لقدرات المفحوصين على الاختبار السهل والاختبار الصعب..... 115
شكل53: التوزيع التكراري للرتب المئينية لقدرات المفحوصين على الاختبار الصعب والاختبار الوسط 116
شكل54: التوزيع التكراري للرتب المئينية لقدرات المفحوصين على اختبار الفقرات الفردية واختبار الفقرات الزوجية.....116
شكل55: منحنيات جميع أفراد العينة..... 118
شكل56: المنحنيات التوافقية..... 118

فهرس الملاحق

الملحق الصفحة

ملحق أ: واجهات البرامج التي تم استخدامها في التحليل.....127

ملحق ب: النموذج الاختباري الأول من النماذج الست..... 131

الملخص

مهيدات، عبد الحكيم علي. بناء بنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية "نموذج مقترح" أطروحة دكتوراه بجامعة اليرموك.2005 (المشرف:أ. د. أحمد سليمان عودة)

هدفت هذه الدراسة إلى بناء بنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية " نموذج مقترح " والتحقق من فاعلية الاختبارات التي يمكن أن تسحب منه. ولتحقيق أهداف الدراسة تم بناء جدول مواصفات اشتمل على مهارات التفكير الرياضي، وحل المشكلات، والتبرير الرياضي، والنمذجة، والترميز، ومستويات التذ كر، والاستيعاب، والتطبيق، والتحليل، والتقويم، والإبداع، واعتمادا على جدول المواصفات تم بناء تجمع من الفقرات (item pool) من نوع الاختيار من أربعة بدائل.

وقد وزعت هذه الفقرات على ست نماذج اختبارية تراوحت بين(45 ـــ 49) فقرة مع وجود تسع فقرات مشتركة في كل نموذج، ليصبح مجموع الفقرات 237 فقرة في النماذج جميعها. وتكونت عينة الدراسة من (2679) طالبا وطالبة موزعين على 15 مدرسة، من طلبة الصف العاشر في الأردن.

وقد طلب من أفراد عينة الدراسة الإجابة عن فقرات النموذج الاختباري في ظروف صفية عادية، وتم تحليل الإجابات باستخدام برنامجي(bilog-mg, SPSS) حيث استخدمت هذه البرمجيات في تقدير إحصائيات الفقرات باستخدام النظرية التقليدية.

وقد تم تقدير معالم الفقرات باستخدام دالة الأرجحية العظمى والتقريب المتكرر، باستخدام نظرية الاستجابة للفقرة (نموذج المعلمتين) كما استخدمت هذه البرمجيات للتحقق من مدى استقرار معالم الفقرة باختلاف العينات، وذلك من خلال تقدير معالم الفقرات المشتركة. واستخدمت هذه البرمجيات أيضا للتحقق من ثبات قدرات المفحوصين باختلاف الاختبار المسحوب من البنك حيث استخدمت خمس اختبارات متباينة في درجة صعوبتها.

بينت نتائج التحليل أن هناك توافقاً عالياً بين الأسلوبين في تقدير صعوبة وتمييز الفقرات، حيث أشارت النتائج إلى أن الأسلوبين اتفقا في الحكم على 90% من الفقرات عند مدى الصعوبة (0.1- 0.9) و تمييز(اكبر من 0.1) وفق النظرية التقليدية و عند مدى الصعوبة ($^{-}3$ - $^{+}3$) وعند تمييز (اكبر من 0.2) وفق نظرية الاستجابة للفقرة. بينما وصلت هذه النسبة الى85% عند مدى الصعوبة (0.3- 0.7) وتمييز(اكبر من 0.1) وفق النظرية التقليدية ، وعند مدى الصعوبة ($^{-}2$- $^{+}2$) وتمييز (اكبر من 0.2) وفق نظرية الاستجابة للفقرة.

كما بينت نتائج التحقق من مدى استقرار معالم الفقرة في المجموعات المختلفة أن هناك استقرارا عاليا جدا لمعالم الفقرات باختلاف المجموعات.

أما فيما يتعلق بنتائج التحقق من مدى استقرار معلمة القدرة بتعدد الاختبارات المسحوبة من البنك، فقد دلت النتائج إلى أن هناك توافقا عاليا في ترتيب قدرات المفحوصين باختلاف الاختبار المسحوب من البنك.

وبعد بناء قاعدة البيانات تم تخزين 233 فقرة، كما تم تخزين المعلومات التالية مع كل فقرة: نوع المهارة الرياضية، ونص الفقرة، والإجابات الصحيحة، وصعوبة الفقرة، والمجال الذي تقيسه الفقرة، ومعامل التمييز، ودالة معلومة الفقرة.

وقد استخدم البرنامج الحاسوبي الذي تم بناؤه لانتقاء فقرات اختبار ضمن مواصفات محددة، كنموذج لتوضيح آلية السحب وفقا للغرض منه. (الكلمات المفتاحية: بنك الأسئلة، نظرية استجابة الفقرة، المهارات الرياضية، نموذج المعلمتين، الفقرات المشتركة).

الفصل الأول

خلفية اعداد بنك الأسئلة وأهميته.

تمهيد:

يعتبر القياس والتقويم التربوي أحد المكونات الأساسية لأي نظام تعليمي، لما يقدمه من تشخيص وعلاج وتغذية راجعة لتوجيه مسار العملية التعليمية وزيادة فعاليتها وتطويرها، لتحقيق الغايات والأهداف المرجوة منها. ولما كان بهذا القدر من الأهمية فقد اهتم به الباحثون اهتماماً كبيراً، وغاصوا في أهدافه وأدواته ونتائجه، وما زال ميداناً خصباً لإجراء المزيد من الأبحاث والدراسات النظرية والتطبيقية للوصول إلى أفضل الأدوات والأساليب التي تسهل الحصول على أصدق النتائج وأدقها وتحقيق أهدافه المختلفة.

ولقد صاحب التقدم في علم القياس تطورا في أدوات القياس كماً ونوعاً، فقد كان التقويم المدرسي يعتمد بشكل عام والتحصيل الأكاديمي بشكل خاص على الاختبارات الشفوية حتى نهاية النصف الأول من القرن التاسع عشر، ثم بدأ ظهور الاختبارات التحصيلية التحريرية من إعداد المعلمين بشكل أساسي. واستمر الوضع كذلك إلى بداية القرن العشرين حيث ظهرت الاختبارات التحصيلية المقننة ، و تمثلت هذه الاختبارات بما يعرف بالاختبارات معيارية المرجع (NRT) (Norm Referenced Tests) التي تعتمد في أساليب نتائجها وتحليل فقراتها على أسس نظرية القياس التقليدية (Classical Test Theory) التي تركز على الفروق الفردية ، وعلى افتراض التوزيع السوي لدرجات الممتحنين على السمة موضع القياس ، وذلك بمقارنة درجات الممتحنين بمعايير المجموعة التي ينتمون إليها .

وتوالى التطور في أدوات القياس فظهر القياس محكي المرجع (الاختبارات محكية المرجع) (Criterion Referenced Tests) الذي يرتكز على قضية أساسية وهي مقارنة الفرد بمحك معين، حيث يهدف بالدرجة الأولى إلى معرفة مستوى الأداء المطلق للطالب بغض النظر عن أداء زملائه من الطلبة.

ويلاحظ المتتبع لحركة القياس والتقويم أن إضافة جوهرية في الأساس النظري لعملية القياس قد حدثت في الإشارة إلى عيوب النظرية التقليدية ومحاولة التغلب على هذه العيوب بوضع

الأسس لنظرية أخرى التي يتوقع أن يكون لها محددات وجوانب قصور تجعل من البحث في مجال النظريتين عملية مستمرة.

ويرى علماء القياس مثل همبليتون وسواميناثان (Hambleton &Swaminathan , 1985) ورايت وستون (Wright & Stone, 1979) أن بناء نظام موضوعي لتحويل المشاهدات إلى قياسات يتطلب أن تتحرر تداريج الأفراد من معالم الفقرات(Item-Free measurement) وأن تتحرر معالم الفقرات من تداريج الأفراد (Person-Free Measurement) ،وقد أدت هذه الجهود إلى تطوير نظرية معاصرة في القياس أطلق عليها نظرية الاستجابة للفقرة (Item Response Theory) حيث توفر هذه النظرية تقديراً لقدرة الفرد لا يتأثر بعينة الفقرات التي استخدمت في التقدير (Item Free) ، وكما أن تقدير معالم الفقرة لا يتأثر بعينه الأفراد (Sample Free) المستخدمة في التقدير.

وتوالى التقدم في علوم القياس والتقويم التربوي فظهرت فكرة إنشاء بنوك الأسئلة في شتى المجالات المعرفية والتحصيلية وفقاً للنظريتين السابقتين وعلى وجه الخصوص بنوك الأسئلة التي يجري إعدادها وفقاً لنظريات القياس الحديثة والتي تتمثل في نظرية الاستجابة للفقرة الاختبارية (IRT) حيث تتضمن هذه البنوك أنواعاً مختلفة من الفقرات الموضوعية تتمثل في اختبارات مسحوبة منها وتعالج أوجه القصور المختلفة والتي ظلت تعاني منها الاختبارات التقليدية منذ فترة طويلة (Hambleton &Swaminathan, 1985) .

وبالرغم من أهمية الامتحان بشكل عام، وأسئلة الاختيار من متعدد بشكل خاص في المراحل الدراسية المختلفة، إلا أنها تشكل بالنسبة للمدرسين والقائمين على الامتحانات مشكلات وهموما كبيرة تتمثل فيما يلي: كثرة عدد الأسئلة وأعداد النماذج المختلفة لنفس الامتحان، والتغير المستمر لعدد الأسئلة وتزايد أعداد الطلبة في المراحل الدراسية المختلفة، وصعوبة إعداد الصور الاختبارية المتكافئة، والوقت والجهد اللازمين لإعداد وتصحيح الاختبارات، والتكلفة في إعداد الاختبارات لمرات متتالية، وسرية الامتحانات والغش بالامتحانات، وعدم تمكن العديد من المدرسين من مهارات، وأسس بناء اختبارات ذات مواصفات علمية جيدة من جهة، وصعوبة استخراج الخصائص السيكومترية لها (صدق ، ثبات ، تمييز ، ...) من جهة أخرى.

وهناك بعض المعوقات التي تواجه عملية الامتحانات والتقويم؛ كاستخدام الامتحان الواحد لتقويم أداء جميع الطلبة، واختلاف مستوى تحصيل الطلبة تبعا لاختلاف الاختبار المستخدم، وعدم

توافر وحدة قياس مطلقة لقياس معدل نمو المفاهيم والمهارات لدى الطلبة من صف دراسي إلى آخر، وغيرها من المعوقات؛ وعدم وجود قاعدة بيانات دقيقة حول مستوى الطلبة، وصعوبة التعرف على مدى إلمامهم بالمهارات، والمعلومات الأساسية في الرياضيات، و مدى تحقيقهم للأهداف المنشودة من تدريسه.

كل هذه القضايا والصعوبات التي تواجه المدرسين والقائمين على الامتحانات تؤدي إلى عزوفهم عن استخدام الامتحانات بالطريقة المناسبة بالرغم من قناعتهم بها، بالإضافة إلى التنازلات المستمرة في القبول بعدد معين من الأسئلة، و تفضيل نوعية دون أخرى، إضافة إلى الإخراج العشوائي لهذه الاختبارات وعدم إتباع الخطوات العلمية في إعدادها.

لذلك تبدو الحاجة قائمة إلى بلورة خطوات واضحة يمكن تعميمها، و الاحتذاء بها في مواقف مماثلة، ويمكن أن يتحقق هذا الأمر في إنشاء بنوك أسئلة بكل ما تمثله من مزايا، فبنوك الأسئلة التي يمكن إعدادها على أسس منهجية وسيكومترية متطورة بالاستعانة بالتقنيات المناسبة يمكن أن تسهم في تطوير أساليب التقويم وأدواته، وتحسين نوعية الفقرات الاختبارية، وبناء اختبارات ذات خصائص معينة من الرصيد المتجدد لفقرات البنك، فتزويد المدرسين بالفقرات الاختبارية الجيدة يوفر الوقت لديهم ليتمكنوا من تكريسه في تحسين عملية التعليم، ورفع مستوى أداء الطلبة كما أن بنوك الأسئلة تعمل على تقليل كلفة بناء الاختبارات.

ونظراً لأن بنوك الأسئلة تضم كماً هائلاً من الفقرات الموضوعية التي تتدرج جميعها في وحدة قياس واحدة تشكل معاً مقياساً واحداً له صفر واحد مشترك وتدريج واحد مشترك يغطي مدىً واسعاً من الموضوع المراد قياسه، فان هذه الخصائص الهامة تساعد واضعي الاختبارات على بناء نماذج اختباريه متكافئة ومتنوعة، وهذا يسهم في التغلب على مشكلات سوء التصرف في الامتحانات وعملية الغش وزيادة سرية الامتحانات. وقد أشار ملمان وأرتر (Millman &Arter, 1980) إلى أن قدرة بنوك الأسئلة على إنتاج نماذج اختباريه متنوعة يؤدي إلى تقليل عملية الغش، ويعمل على تسوية مشكلات الامتحانات ، وقد أيد فلايجو (Falayajo, 1998) الرأي القائل بأن بنوك الأسئلة تمكن من التغلب على مشكلات سوء التصرف في الامتحانات ، حيث أشار إلى أن بناء وتطوير هذه البنوك، وتدريج الفقرات الاختبارية، وامكانية عمل النماذج الاختبارية المتنوعة، ووضع الفقرات من اختبارات متنوعة على مقياس مشترك ، كل هذه الخصائص تسهم في حل مشكلات الامتحان .

وقد أشار شوبن (Choppin,1981) في دراسته إلى أن بنوك الأسئلة تعمل على زيادة السرية (Security) بطريقتين :

الأولى : كثرة عدد الفقرات الاختبارية في البنك تجعل من المستحيل على أي طالب تذكر جميع هذه الفقرات بدون إتقان المحتوى أولاً .

الثانية: إمكانية عمل النماذج المتعددة للامتحان تمنع من حدوث تسرب الاختبار، لأنه لا أحد يعرف أي نموذج سيطبق.

ونظرا لما تمثله مادة الرياضيات في مرحلة التعليم الأساسي من مجال خصب للمعرفة والمعلومات التي تعد أحد ركائز التطورات العلمية والتكنولوجية التي يتطلبها المجتمع، وتمثل أساساً لعمليات تفكير الطلبة وتساعدهم في تفسير الظواهر والأحداث المحيطة بهم، وفي التعامل مع البيئة والتعرف على مواردها واستثمارها بالإضافة إلى أهمية مادة الرياضيات في تأهيل الطلبة لدراسة كافة الفروع المتخصصة في المرحلتين الثانوية والجامعية، ولما كان تقويم مدى امتلاك الطلبة للمهارات الأساسية في مقرر معين يمثل جزءاً مهماً من الممارسات التعليمية لأي نظام يسعى لتطوير نظامه التعليمي، فقد شهدت الآونة الأخيرة حرص الكثير من الدول على تقويم تحصيل طلبة مرحلة التعليم الأساسي فيها في مادة الرياضيات على فترات زمنية متقاربة ضمن العديد من المشروعات وفيما يلي بعضها:

1. مشروع تيمس العالمي (TIMSS) Trends In International Mathematics And Science Survey

يهدف هذا المشروع إلى تقويم التحصيل الدراسي في مادتي العلوم والرياضيات لطلبة الصفين الرابع والثامن من مرحلة التعليم الأساسي في كثير من دول العالم، وذلك للتعرف على مدى إلمام طلبة الدول المشاركة بالمعارف والمعلومات والمهارات الأساسية في هاتين المادتين، ومقارنة الدول معاً من حيث مستوى التحصيل الدراسي، وترتيبهم على مقياس يمثل خط متصل واحد، حتى يتسنى لكل دولة التعرف على موقع مستوى تحصيل طلبتها بالنسبة لمواقع مستوى تحصيل طلبة دول العالم الأخرى، بحيث تستفيد كل دولة من هذا الترتيب في تطوير مناهجها وممارساتها التعليمية مما يؤهلها لمواكبة التطورات العلمية.

(Wang , 1998; Valverde & Shmid , 2000)

2.مشروع نيب الأمريكي (NAEP) (National Assessment Of Educational Progress) يهدف هذا المشروع إلى تقويم التحصيل الدراسي لطلبة الولايات المتحدة الأمريكية في العديد من المواد المقررة عليهم في مرحلة التعليم الأساسي، ومن بين هذه المواد مادة الرياضيات،

وذلك بغية الحصول على المعلومات التي يستخدمها متخذو القرار في تطوير الممارسات التعليمية المستخدمة في تحقيق الأهداف المرجوة (Rodriguez, 1997)

3.مشروع بيزا العالمي **Program For International Student Assessment (PISA)**
يهدف هذا المشروع إلى تقويم التحصيل الدراسي في مستويي المعرفة والمهارات لدى الطلبة من أعمار (15) سنة في العديد من المواد الدراسية المقررة عليهم مثل (القراءة، و الرياضيات، والعلوم.....) مع التأكيد على استعداد الطلبة للحياة (Aaron R.Douglas & Christian, 2000)

4.المشروع العربي لمنظمة اليونسكو .
يهدف هذا المشروع إلى تقويم التحصيل الدراسي للطلبة في بعض الدول العربية في مواد اللغة العربية، والرياضيات، والعلوم، حتى يتسنى للدول المشاركة التعرف على مستوى تحصيل طلبتها في ضوء محكات محددة، لكي تستفيد من تلك النتائج في تطوير ممارساتها التعليمية (سليمان، وأمينة، 2000).

وحرصاً من الأردن - الذي يتمتع بمجتمع تعليمي طموح - على الاستفادة من التجارب والمشاريع الدولية، فقد شارك في بعض هذه المشاريع واستفاد من النتائج التي حصل عليها في تشخيص الوضع القائم وتحليل قضاياه، حيث توصل إلى مجموعة من القضايا الاستراتيجية، والمشكلات الملحة والعوامل التي يجب التعامل معها في المرحلة القادمة. وانطلاقاً من ذلك فقد عمدت وزارة التربية والتعليم إلى تطوير مشروع الاختبارات الوطنية لضبط نوعية التعليم، والذي يهدف إلى التعرف على مدى امتلاك طلبة مرحلة التعليم الأساسي للمهارات الأساسية اللازمة لتعلمهم المستقبلي، وكذلك للتعرف على القضايا والمشكلات التي تواجههم، ومن ثم تحديد الاحتياجات ليصار إلى تزويد متخذي القرار بمعلومات حول جودة التعليم، تمهيداً لاتخاذ قرارات التطوير المناسبة. ووقعت مسؤولية تنفيذ هذا المشروع على إدارة الامتحانات والاختبارات في وزارة التربية والتعليم من حيث بناء الاختبارات ، وتطبيقها ، وتحليل نتائجها ، وإعداد التقارير اللازمة . وتطبق هذه الاختبارات على طلبة الصف الرابع الأساسي في مبحثي اللغة العربية، والرياضيات، وعلى طلبة الصف الثامن الأساسي في مباحث التربية الإسلامية، واللغة العربية، والعلوم، والرياضيات، واللغة الإنجليزية، والاجتماعيات، والتربية المهنية، وعلى طلبة الصف العاشر الأساسي في مباحث اللغة العربية، والتربية الإسلامية، والعلوم، والرياضيات، والحاسوب،

واللغة الإنجليزية، والاجتماعيات. ولكون النتائج تشير إلى تدني مستوى التحصيل لدى الطلبة ولا تعبر عن مستواهم الحقيقي نتيجة تطبيق اختبارات تفوق أو تقل في مستوى صعوبتها عن مستوى قدرتهم ، لذلك فإن الحاجة إلى وجود فقرات معدة جيداً ومتاحة أمام واضعي الاختبارات من شأنها أن تساعد في قياس موضوعي للتحصيل . ولعل الحل المناسب لمثل هذه الحاجة يكمن في بناء بنوك أسئلة تضم عدداً ضخماً من الفقرات الموضوعية الجيدة المدرجة على مقياس مشترك حيث تمكن هذه البنوك واضعي الأسئلة من سحب الاختبارات اللازمة وإيداع الفقرات المناسبة.

ولما كان معظم الباحثين يتناولون في أبحاثهم بناء بنوك الأسئلة معتمدين على الوحدات الدراسية أو الفصول المقررة في المناهج، والكتب المدرسية، وحيث أن هذه المناهج، والكتب المدرسية متحركة، ودائمة التغيير، والتطوير مما يؤثر على مرونة وسلامة مخزون بنوك الأسئلة من الفقرات الاختبارية، لذا تبدو أهمية هذه الدراسة مرة أخرى لتناولها أسلوبا جديدا في بناء بنوك الأسئلة من منظور آخر بمعزل عن الفصول أو الوحدات الدراسية المتضمنة في المناهج والكتب المدرسية ولتكون نواة لبنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية لحث ذوي الاهتمام بعمل مثل هذه البنوك في مباحث أخرى.

ووفق هذا الأسلوب الجديد تم الاستعانة بمصفوفة كراثوول (Krathwohl, 2002) في تصنيف الأهداف التعليمية لأي مبحث حسب البعدين التاليين:

البعد الأول: بعد العملية المعرفية (The Cognitive Process Dimension).

البعد الثاني: بعد المعرفة نفسها (The Knowledge Dimension).

والجدول (1) يوضح تصنيف الأهداف التعليمية في بعدين وفقا لهذه المصفوفة.

جدول 1: تصنيف الأهداف التعليمية في بعدين

بعد المعرفة The Knowledge Dimension	بعد العملية المعرفية The Cognitive Process Dimension					
	التذكر Recall	الفهم Understand	التطبيق Apply	التحليل Analyze	التقويم Evaluate	الإبداع Create
الحقائق Factual Knowledge						
المفاهيم Conceptual Knowledge						
الإجراءات Procedural Knowledge					*	
ما وراء المعرفة Meta Cognitive Knowledge						

- تعني أن الهدف في مجال معرفة الإجراءات ومستوى التقويم .

بناءً على هذا التصنيف يمكن تصنيف أي فقرة اختباريه في أي مبحث ضمن خلية من خلايا الجدول 1، وبذلك يكون قد تم التخفيف من الالتزام المطلق بالكتاب المقرر أو المناهج المقررة، وفي هذا المجال يمكن الإشارة إلى التجربة الأردنية في مجال تصنيف الأهداف التعليمية، حيث اعتمدت وزارة التربية والتعليم التصنيف الوارد في الجدول 2.

جدول 2: تصنيف الأهداف التعليمية في وزارة التربية والتعليم / الأردن

عناصر المحتوى	مستويات التقويم		
	المعرفة والفهم	حل المشكلات	الاستقصاء
الوحدة الدراسية الأولى			
الوحدة الدراسية الثانية			
الوحدة الدراسية الثالثة			
.....			

ويندرج ضمن مجال المعرفة والفهم المهارات الفرعية التالية:- (التذكر، والتطبيق، وقراءة الرسومات وإنتاجها)، و يندرج ضمن مجال حل المشكلات المهارات الفرعية التالية:-

(التفسير، والإجراءات غير الروتينية، والحكم على معقولية الحل) و يقع ضمن مجال الاستقصاء المهارات الفرعية التالية:- (فهم المسألة، واختيار الاستراتيجية المناسبة، وتطبيق خطوات الاستقصاء، وبرهنة التعميم، وإعطاء مثال).

أما بالنسبة لمبحث الرياضيات المراد بناء بنك أسئلة له وفق التصنيف المشار إليه في جدول 1 فقد تم إجراء بعض التعديلات لتسهيل عملية بناء الفقرات الاختبارية التي تقيس المهارات الرياضية في المرحلة الأساسية، حيث تم دمج البعد الأول لمستويات المعرفة (بعد العملية المعرفية) إلى جزأين:

الجزء الأول: القدرات العقلية الدنيا والتي تشتمل على (التذكر، الاستيعاب، والتطبيق).

الجزء الثاني: القدرات العقلية العليا والتي تشتمل على (التحليل، والتقويم، والإبداع).

وبالنسبة للبعد الثاني (بعد المعرفة نفسها) فقد تم تجزئة عناصر المعرفة إلى مهارات تسهل على الباحث كتابة فقرات اختباريه تقيسها. والجدول 3 يظهر هذه التعديلات، وبهذا الإجراء نوفر مجالا من المرونة لاستخدام جدول تصنيف عمليات المعرفة 1 بتشكيلة وتوليفة وامكانية التعديل عليه حسب خصوصية المبحث الواحد، وتحقق صدقا معقولا للتصنيف سواء فيما يتعلق بمدى التداخل في مستويات العملية المعرفية، والوضوح الإجرائي لمصطلحات بعد المعرفة نفسها.

جدول 3: تصنيف الأهداف التعليمية للرياضيات

بعد المعرفة	بعد العملية المعرفية	
	القدرات العقلية الدنيا (تذكر ، استيعاب ، تطبيق)	القدرات العقلية العليا (تحليل ، تقويم ، إبداع)
مهارة التفكير الرياضي Mathematical Thinking Skill		
مهارة البرهان والتحليل الرياضي Mathematical Reasoning Skill		
مهارة النمذجة Modeling Skill		
مهارة حل المشكلة Problem Solving Skill		
مهارة العرض presentation Skill		
مهارة الترميز Symbolic Technical Skill		
مهارة التواصل Communication Skill		
مهارة الربط Connection Skill		

وتم بناء الفقرات الاختبارية لقياس المهارات الأساسية من الموضوعات المختلفة التي تدرس في مرحلة التعليم الأساسي لمبحث الرياضيات والمتمثلة بـ (الأعداد والعمليات عليها ، والجبر ،والهندسة ، والقياس ، والإحصاء والاحتمالات) .

هدف الدراسة :

هدفت هذه الدراسة إلى بناء نواة لبنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية " نموذج مقترح " واستخراج المؤشرات الإحصائية لفقرات البنك وفق النظرية الكلاسيكية، ونموذج استجابة الفقرة (النموذج ثنائي المعلمة). حيث قامت الدراسة بتحديد معايير للفقرات التي تم انتقاؤها للتخزين في بنك الأسئلة، وتحديد البيانات التي تخزن مع هذه الفقرات.

أسئلة الدراسة:

بعد أن تم إعداد الفقرات التي تقيس المهارات الرياضية في نهاية المرحلة الأساسية، وتطبيق هذه الفقرات على عينات من الطلبة الذين انهوا المرحلة الأساسية، تم تحليل نتائج تطبيق الفقرات للإجابة عن الأسئلة التالية:

س1 : ما خصائص الأسئلة في بنوك الأسئلة ؟

تمت الإجابة على هذا السؤال من خلال ما يمكن أن نقدمه من تقدير لمؤشرات القدرة التمييزية ومؤشرات الصعوبة (بالنظرية الكلاسيكية والنموذج اللوغاريتمي ثنائي المعلمة) للفقرات المتبقية والمحذوفة في ضوء المؤشرات الخاصة بهما (الصعوبة، والتمييز، ودالة معلومة الفقرة: Item Information Function، ومدى مطابقة البيانات للنموذج المستخدم).

س2: ما خصائص الاختبار المستل (المسحوب) من بنك الأسئلة ؟

تمت الإجابة على هذا السؤال من خلال مدى تحقيقه لغرض الاختبار في ضوء الشروط المطلوبة (عدد الفقرات، ومدى شموله للمهارات المقاسة، ومدى القدرة، ومدى الصعوبة، ودالة معلومة الاختبار ((Test Information Function (TIF) والتي تحدد خصائص الاختبار).

س3 : ما مدى استقرار معالم الفقرة في المجموعات المختلفة ؟

تمت الإجابة على هذا السؤال من خلال المقارنة بين أداء المجموعات المختلفة على نفس الفقرة وفق النظريتين التقليدية والحديثة.

س4 : ما مدى استقرار معلمة القدرة للفرد (θ) بتعدد الاختبارات المستلة (المسحوبة) ؟

تمت الإجابة على هذا السؤال من خلال مقارنة قدرة الفرد على الاختبارات المسحوبة من البنك.

تعريف المصطلحات:

* بنك الأسئلة: تجمع من الفقرات التي تم تجريبها وتحليلها للتوصل إلى معالمها (الصعوبة، والتمييز، ودالة معلومة الفقرة)، ومن ثم تدريجها على مقياس مشترك، وتخزينها في قاعدة بيانات وفق برنامج حاسوبي خاص

* معالم الفقرة: هي قيم إحصائية تحسب باستخدام معادلات رياضية وتشمل:

- معلم الصعوبة التي تمثل مستوى القدرة عند الاحتمال(0.5) للاستجابة عندما يكون معامل التخمين مساويا للصفر
- معلم التمييز الذي يمثل تمييز الفقرة بين مستويات القدرة.

* معلم القدرة:هي اللوغاريتم الطبيعي لمرجح نجاح الفرد على الفقرات التي تعبر نقطة صفر التدريج عن صعوبتها.

* دالة معلومات الفقرة: هي دالة رياضية تصف العلاقة بين مستوى القدرة ومقدار المعلومات التي توفرها الفقرة.

* دالة معلومات الاختبار: هي مجموع دوال الفقرات المكونة للاختبار

الفصل الثاني

أدبيات بنوك الأسئلة

الجزء الأول: الإطار النظري لمكونات بنك الأسئلة وعناصره

يتناول الإطار النظري لهذه الدراسة بنوك الأسئلة ونظريات القياس (النظرية الكلاسيكية، ونظرية الاستجابة للفقرة الاختبارية " النموذج اللوغاريتمي ثنائي المعلمة ").

تعريف بنك الأسئلة:

شهد مجال القياس التربوي خلال العقود القليلة الماضية عدداً من المستجدات، بهدف التطوير والتحديث، ومن هذه المستجدات إنشاء ما يعرف ببنوك الأسئلة بشكلها العلمي المتكامل، حيث تتضمن هذه البنوك عدداً كبيراً نسبياً من الفقرات الجيدة لقياس الأهداف والمهارات، وتوفر الجهد والوقت والموضوعية في قياس نتاجات التعلم.

وبالتحديد فقد عرفه شوبن (Choppin, 1979) بأنه حصيلة كبيرة من الفقرات الإختبارية التي يتم تنظيمها وفهرستها كما هو الحال في كتب المكتبة.

أما بيستون (Beeston,2002) فيعرفه: بأنه مخزون من الفقرات الاختبارية تم تصنيفها من حيث موضوع الفقرة وهدف الاختبار ودرجة الصعوبة ، وخزنت في قاعدة بيانات لاستخدامها وقت الحاجة ويمكن استخدام برمجيات حاسوبية لترتيب الفقرات المنتقاة في اختبار ما .

وعرفه بورجوف (Burghof,2001): بأنه مجموعة كبيرة نسبياً من الفقرات التي يسهل الوصول إليها مخزنة في قاعدة بيانات مصنفة ومرمزة من حيث الموضوع والمستوى الدراسي ونوع الفقرة ومهارة التفكير ومستوى الصعوبة .

أما سكويرز(Squires, 2003) فيعرفه بأنه تصنيف وفهرسة الفقرات الاختبارية وفقا للمحتوى الذي تقيسه هذه الفقرات ومستوى صعوبتها (سهلة، متوسطة، صعبة).

وإذا كانت التعريفات مختلفة شكلاً باختلاف الغرض من استخدام بنك الأسئلة إلا أنه توجد أرضية عامة مشتركة تتمثل في أن بنك الأسئلة لا يتضمن إلا نوعية جيدة من الفقرات أو الأسئلة، وبوجه عام فإن السؤال أو الفقرة التي يتضمنها بنك الأسئلة لا بد أن تكون قد تم تكوينها بطريقة صحيحة، ومحتوى السؤال أو الفقرة على درجة عالية من الصدق والثبات والقدرة على التمييز، وقد تحسب الخصائص السيكومترية للأسئلة بالطريقة التقليدية، وهنا تكون بنوك الأسئلة مفيدة للمعلمين. إن هذا المستوى من بنوك الأسئلة كان شائعا في الماضي لدى كثير من مصممي وناشري الاختبارات كمحاولة منهم لتوفير الكم الكافي من الأسئلة لاعداد الصور المتوازية للاختبارات المقننة وهناك مستوى أعلى من ذلك حينما يتم تدريج الأسئلة وفقاً لمؤشرات معينة وبالتالي يتوفر

لدينا بنك أسئلة مدرج أو معاير (Calibrated Item- Bank) بعد التأكد من ملاءمتها (Fit) للنموذج الذي تم اعتماده وفقاً لنظرية الاستجابة للفقرة.

لماذا بنوك الأسئلة ؟

أشار رايت وبل (Wright & Bell,1984) إلى فوائد كثيرة يمكن أن توفرها بنوك الأسئلة منها:

أولاً: تضمينات للمعلم (Implication For The Teacher) حيث يتجلى ذلك في المواقف التالية:

أ. إذا تم إعداد بنك الأسئلة بشكل جيد فهذا يمكن المعلم من إعداد الاختبار المناسب لأي غرض سواء اكان الغرض من الاختبار معرفة مستويات الطلاب أو قياس مستوى طالب منفرد، حيث تتوفر الفقرات الملائمة لكل هدف ولكل مستوى.

ب. بنك الأسئلة الجيد يمكن المعلم من إعداد أنواع مختلفة من الاختبارات حسب الصف أو المدرسة أو المنطقة التعليمية، كما ويمكنه من إضافة فقرات جديدة لقياس أهدافه التعليمية، وهذا يوفر له المجال لإجراء المقارنات العامة التي يتطلبها واقعهم التعليمي.

جـ. يوفر بنك الأسئلة الجيد إمكانية اختيار فقرات لبناء اختبارات مسحية لتقييم الأداء في منطقة تعليمية أو لبناء اختبارات أو مقاييس محكية المرجع تلزم لتخطيط المناهج وأسس القبول.

د. يساعد بنك الأسئلة في تشخيص الأداء وذلك من خلال توفير البنك إمكانية بناء اختبارات ذات قدرة تشخيصية لتشخيص أداء الطلبة.

ثانياً: تضمينات سيكومترية (Psychometric Implications) .

أ. يزود تدريج الفقرات باني الاختبار بمعلومات هامة: فالفقرات ذات التدريج المنخفض تكون سهلة، بينما الفقرات ذات التدريج المرتفع تكون صعبة. هذا التعاقب خلال الفقرات والذي يرتب تدريجها من السهل إلى الصعب يصف مسار تعلم الطلبة.

ب. الفقرات ذات المحتوى الواحد تدرج على متغير واحد، وكل فقرة تمثل موقعاً على المتغيرلكن نفس المتغير قد يمثل بفقرات أخرى ذات مستوى صعوبة مختلف. وهذا يوفر إمكانية الاستدلال على إتقان الطالب حسب المتغير الأساسي الذي تتشارك فيه، الفقرات بغض النظر عن الفقرات المستخدمة أو الطالب الذي جلس للاختبار.

جـ. مكان الطالب على المتغير يبين وضعه بالنسبة لجميع الذين أخذوا فقرات من البنك وليس فقط بالنسبة للذين أخذوا نفس الفقرات.

د. إذا تيقنا من مصداقية نمط أداء الطالب فأن ذلك يؤكد صحة قياس مستوى الطالب حسب كل فقرات بنك الأسئلة وليس حسب فقرات الاختبار وهذا يعطي معياراً محكي الأداء للطالب (Wright & Bell,1979) .

ثالثاً. تضمينات للمنهاج (Curricular Implications).

أ. يمكن إضافة فقرات جديدة في بنك الأسئلة دون استخدام اختبارات قبلية واسعة المدى (Large Scale Pre-test) شريطة قيام خبير بتقييم ثبات الفقرة الجديدة مع الفقرات الموجودة في البنك، وعندها تدرج الفقرة الجديدة لتستخدم في الاختبار.إن إمكانية إضافة فقرات جديدة للبنك يعني أن بنك الأسئلة يواكب المنهاج وتطوراته كما أنه يشير إلى عدم تقييد المعلمين باستخدام قوائم محددة من الفقرات.

ب. إن فقرات الاختبارات المأخوذة من بنك أسئلة معين تمكن المعلمين من مقارنة نتائج اختباراتهم ومعرفة مدى تعلم موضوع ما من قبل طلبة مختلفين مع معلمين مختلفين، وذلك يوفر إمكانية تقييم طرق التعليم والتعرف إلى طرق بديلة تكون أكثر فاعلية.

رابعاً. تضمينات للطلبة (Implications For The Students).

أ. يجب أن يتضمن بنك الأسئلة فقرات قادرة على قياس أداء الطلاب وتحت كل الظروف، إذ أن بعض الطلاب يلجأون إلى التخمين، أو يخطئون بسبب الإرهاق أو سوء فهم التعليمات، هذا وغيره من العوامل يجعل علامات الطلاب مظللة، لذا فان تحليل ثبات استجابات الطلاب يتطلب كثيراً من الحذر لتجنب تلك العوامل المؤثرة.

ب. من المعروف أن إجابة الطالب على فقرة ما يوضح الفرق بين قدرة الطالب وصعوبة الفقرة، أي ان الاستجابة تمثل النموذج البسيط للتفاعل بين الطالب وسهولة الفقرة باستثناء نواتج أخرى لأي عملية أخرى غير مقصودة. وعليه فإن تحليل باقي فقرات الاختبار يعطي إطار عمل جيد للكشف عن اختلالات الأفراد(Wright & Stone ,1979) وعند تحديد هذه الاختلالات يستطيع المعلم تقصي مصادرها ومعالجتها ليساعد الطالب على مواجهة هذه الاختلالات ، وبالتالي يكون قرار المعلم أكثر دقة ، ويمكن إعادة الاختبار بمجموعة مقارنة من الفقرات ، وتحت ظروف أفضل وإبقاء الفقرات المناسبة ، ويكون القرار أكثر حكمة عند معرفة سبب الاختلال ومعرفة المؤثرات على جلسة الاختبار كما أن التحليل الروتيني لثبات أداء الطالب يساعد المعلمين في عمل اختبار أفضل خاصة عند معرفتهم طبيعة مشاكل الأداء .

أنواع بنوك الأسئلة.

يرجع فشل الجهود المبذولة في أواخر الستينيات وأوائل السبعينيات لاتباع بنوك الأسئلة في مجال الاختبارات إلى عدم وجود برامج كمبيوتر (Computer Software) وغيرها من التسهيلات أو الإمكانيات الأخرى بسبب ضخامة حجم الأوراق والإجراءات والخطوات التي كانت تستخدم في ظل النظام اليدوي قبل الاستعانة بالكمبيوتر. وبناءاً على هذا يمكن القول: أن بنوك الأسئلة قد تصنف طبقاً لنوعية فقراتها التي تتضمنها إلى ثلاثة أنواع هي :

1. بنك أسئلة يدوي بالكامل.
2. بنك أسئلة يدوي باستخدام البطاقة لكل فقرة ولكن يستعان بالكمبيوتر في حساب معاملات صدق وثبات الأسئلة أو الفقرات.
3. بنك أسئلة يستعين بالكمبيوتر كاملاً.

ومن ثم فإن بنوك الأسئلة قد تكون غير مدرجة بالكامل وتدار يدوياً بشكل كلي (Fully Manual) أو تكون نصف آلية أي بنوك غير مدرجة ويستعان بالكمبيوتر في حساب معالم الفقرات أو تكون مدرجة وتدارً بالكمبيوتر بشكل كامل (Calibrated And Fully Computerized).

أهمية بنوك الأسئلة.

أما عن أهمية بنوك الأسئلة فقد ترتبط الفكرة بعمليات تسهيل وتوفير فقرات اختبارات مقننة أمام المعلمين وواضعي الاختبارات، والسرعة في إنجاز هذه العمليات وبدقة كبيرة، كما تفيد بنوك الأسئلة في تنفيذ فكرة تفريد التقويم (Individualized Assessment) بحيث يقدم الاختبار المناسب لظروف كل فرد أو مجموعة من الأفراد، ويتيح بنك الأسئلة مقارنة نتائج تطبيق الاختبار على مجموعات مختلفة من الطلاب في نفس مستوى القياس وبدرجة دقة عالية لإتمام المقارنة. كما ويتيح الكمبيوتر هنا تنفيذ فكرة الاختبار المكيف بالكمبيوتر (Computer Adaptive Testing) على الطلبة. وقد أشار العديد من الباحثين (McCallon, & Schumaker,2002 ,Bello,Wole,&Shola,2000;Hambleton,Swaminathan&Rogers,1991 كاظم، 1988). إلى أهم المزايا التي تحققها بنوك الأسئلة على النحو التالي:

1. تخفيض أعباء وجهود وكلفة إعداد وتصميم الاختبارات (توفير الوقت والجهد والكلفة)
2. إمكانية إتباع لامركزية تطبيق الاختبارات ، وتطبيق اختبارات على المستوى الوطني مع إمكانية إجراء المقارنات الدقيقة بين نتائج تطبيقها .

3. تقلل مشكلة تسرب الأسئلة أو عدم الأمن لها بسبب ضخامة عدد الأسئلة المخزنة في بنك الأسئلة.

4. تحسين مستويات التقويم التربوي والامتحانات حيث يكون اعتماد مصممي الاختبارات والمعلمين على أسئلة جيده يتضمنها البنك.

5. توفير الفرصة أمام المعلمين لوضع أسئلة الاختبارات لطلابهم بأنفسهم بسهولة .

6. يسهل تصميم اختبارات عديدة لأغراض متنوعة .

7. نظراً لأن أساس القياس في بنك الأسئلة هو الفقرة وليس الاختبار ككل، لذلك فإن بنك الأسئلة المدرج يسهل التفسير القائم على أساس محكي (Criterion Referenced).

8. توفير الوقت والجهد الذي يبذله المعلم في تصميم الاختبارات وإتاحة الفرصة أمامه لتكريس جهوده نحو التعليم.

9. تفيد بنوك الأسئلة في حال البلدان ذات النظام التعليمي الذي يضم عدداً كبيراً من المدارس مع قلة الإمكانات، وندرة الخبرة بأسس القياس والتقييم.

10. يساعد مستخدمه على التحكم بطول الاختبار ومحتواه ليلائم الظروف المحلية، دون أن يفقد ثبات النتائج وإمكانية تعميمها.

11. تحقيق الموضوعية في تقويم التحصيل الدراسي للطلاب ، ويعتمد هذا على مجموعة من العوامل التي يمكن أن تؤثر على بناء بنك الأسئلة ، مثل مدى الدقة في صياغة الفقرات ، وتحليلها ، ومدى تغطيتها لمجال القياس من حيث المحتوى والمستوى وكيفية تخزينها .... .

12. إتاحة الفرصة لقياس القدرات العقلية العليا، وذلك نظراً لاتساع المدى الذي يغطيه البنك.

13. مرونة القياس حيث يسهل تشكيل أي اختبار في أي وقت بسرعة وبسهولة.

14. إمكانية المقارنة الموضوعية بين مستوى تحصيل الطلاب مهما اختلفت الاختبارات المستخدمة، ومهما اختلفت مستويات المجموعات التي ينتمون إليها، ومهما اختلفت الأعوام الدراسية، طالما أن الفقرات قد تم تدريجها ومعايرتها ووضعها في بنك واحد.

15. يساعد في تطوير العملية التعليمية من خلال:

أ. التقويم المستمر.

ب. استخدام الصور المختصرة من الاختبارات.

جـ. إمكانية المقارنة.

مساوئ ومحددات بنك الأسئلة: Disadvantages and limitation of Item Banking

نظراً لصعوبة مشروع بنوك الأسئلة وحاجته إلى إمكانات وخبرات كثيرة، فإن تنفيذه يواجه بعض السلبيات والمحددات المرتبطة بجميع جوانب الإعداد والتنفيذ والاستخدام، ومن هذه السلبيات والمحددات ما يلي:

1. كتابة الفقرات. وهي خطوة تواجه صعوبات كثيرة من قبل غير المدربين على ذلك حيث يجدون صعوبة بالغة في التفرقة بين أنواع الأسئلة الموضوعية وفنيات صياغتها والمستويات العقلية التي تقيسها تلك الأسئلة.
2. يلزم تخزين الفقرات في البنوك واستدعائها حاسبات آلية ومبرمجين متدربين.
3. الاستخدام المتزايد لبنوك الأسئلة دون إضافة فقرات جديدة يؤدي إلى الإفلاس. وللمحافظة على البنك وديمومته يلزم إعداد فقرات جديدة بصفة مستمرة وتجريبها وتدريجها ثم تخزينها في البنك. وكذلك تعديل الفقرات التي تم تداولها بين الطلاب لأكثر من مرة (كاظم ،2000).
4. إدارة بنوك الأسئلة. وهي مشكلة تتطلب تحديد المسؤولية حتى لا تتسرب فقرات البنك، ولذلك يمكن وضع الفقرات بعد تدريجها ومعايرتها على (Disks) والاحتفاظ بها في مكان آمن ، بالمقابل يجب أن تتوافر فقرات أخرى في البنك يمكن استخدامها لغايات التقويم المستمر أثناء العام الدراسي.
5. تحتوي بنوك الأسئلة على فقرات مدرجة واختبارات معادلة من حيث الخصائص والموضوعات التي تغطيها، لذا نحتاج إلى عملية مراجعة وتنقيح دورية ودقيقة للفقرات لضمان جودة هذه الفقرات والمحافظة على مخزون البنك (Rudner, 1998).

نموذج بنك الأسئلة.

نظراً لأهمية إنشاء نظام بنك أسئلة يسهم في تطوير أساليب قياس وتقويم للتحصيل الدراسي، تقدم هذه الدراسة أنموذجاً لبنك أسئلة، كيف بتنسيق جيد من عدة نماذج لبنوك أسئلة ذكرت بالأدب الحديث والدراسات البحثية، في خمس مراحل مختلفة مبينة تاليا ومفصلة في الجدول 4 :

1. بناء الفقرات:.Item Development Stage
2. التجريب: Pilot Testing Stage.
3. تعيير (تدريج) الفقرات: Item Calibration Stage .
4. ترميز الفقرات ومعلومات الاختبار: Coding of Item and Test Information .
5. الاستخدام والتقييم للبنك: Using and Assessment of Item Bank .

جدول 4: النموذج التصوري لنموذج بنك الأسئلة

المرحلة/1	المرحلة/2	المرحلة/3	المرحلة/4	المرحلة/5
مرحلة بناء الفقرات	مرحلة التطبيق	مرحلة تدريج الفقرات	مرحلة ترميز الفقرات ومعلومات الاختبار	مرحلة الاستخدام والتقييم للبنك
* تحديد أهداف الاختبار * تصنيف الأهداف التعليمية * تحديد نوع وعدد الفقرات * بناء جدول مواصفات * بناء فقرات مطابقة للمواصفات * غربلة وتحليل كل فقرة * بناء أو تجهيز تعليمات الاختبار * تجميع فقرات الاختبار * بناء مفتاح تصحيح	* تجريب الفقرات على مجموعة أو مجموعات من الطلبة * فحص جاهزية البيانات للتحليل	* تحديد النظرية(CTT) أو (IRT) * تحديد أي النموذج من (IRT) سوف يستخدم * تدريج الفقرة * تحديد الصعوبة والتمييز لكل فقرة من خلال التحليل باستخدام الحاسب * استخدام برمجية Bilog * مراجعة قيم معالم الفقرة a, b	* قيم المعالم للفقرة وإحصائياتها ترمز مقابلا كل فقرة	* تنظم الفقرات وتصنف وتفهرس في البنك باستخدام برمجية حاسبية * بنك أسئلة تام

المرحلة الأولى: مرحلة بناء الفقرات Item development Stage

تعد هذه المرحلة من أهم مراحل بناء مشروع بنك الأسئلة، فالفقرات الاختبارية التي ستخزن في بنك الأسئلة يجب أن تتصف أولاً بالملاءمة (Fit)، حيث أن الفقرات الكثيرة غير الملائمة (Misfit) تجعل البنك عديم الأهمية والفائدة. ومن الضروري أن تتصف هذه الفقرات بالصدق والثبات ليتسنى استخدامها في بناء الاختبارات المتنوعة، مع الأخذ بعين الاعتبار العناية الفائقة باختيار كاتبي الفقرات، ممن تتوافر فيهم المهارة والكفاية في بناء الفقرات(Evans&Masters,1986)، وسوف يتم إتباع نموذج ننتي (Nenty,1991) في بناء الفقرات والذي يمر بالخطوات التالية :

1. تحديد المستوى أو المجال للسلوك المقاس.
2. تحديد أهداف الاختبار للفقرات المقصودة.

3. تصنيف المهارات التعليمية التي سيتم بناء الفقرات من أجلها.
4. تحديد نوع وعدد الفقرات التي سيتم وضعها.
5. بناء جدول مواصفات يحدد الموضوع والسلوك.
6. بناء فقرات اختبارية مطابقة للمواصفات.
7. غربلة(screening) وتحليل كل فقرة بناء على الفقرات التي تم اختيارها.
8. بناء أو إعداد إرشادات الفقرة أو تعليمات الاختبار.
9. تجميع فقرات الاختبار.
10. عمل صور اختبارية أولية.
11. بناء مفتاح تصحيح.

المرحلة الثانية: مرحلة التجريب Pilot Testing Stage.

تمتاز بنوك الأسئلة باحتوائها على رصيد متجدد من الفقرات الاختبارية ذات خصائص قياسية معلومة، فالانتهاء من كتابة الفقرات وغربلتها وإعداد الصور الاختبارية لا يعني أنها أصبحت صالحة للتخزين في البنوك، إذ ينبغي أن تخضع هذه الفقرات لعمليات التجريب الميداني على عينات من الطلبة، واستخدام الأساليب السيكومترية والإحصائية لتحليل البيانات المتعلقة بكل فقرة لتقييم فاعليتها، من حيث مدى ملاءمتها للمجتمع الذي تمثله هذه العينات من الطلبة، وقدرتها على التمييز بين مستوياتهم التحصيلية المتباينة، وذلك باستخدام نموذج سيكومتري مناسب.

المرحلة الثالثة: مرحلة تعيير (تدريج) الفقرات Item Calibration Stage.

بعد الانتهاء من تجريب الصور الاختبارية، وجمع البيانات عن الفقرات الاختبارية، وتفريغها على الحاسب، وتدقيقها، تبدأ عملية التحليل لهذه البيانات حيث تقدر قيم معالم الفقرات في هذه العملية. وفي هذه المرحلة تتحدد مزايا وعيوب كل من النظريتين التقليدية ونظرية الاستجابة للفقرة.

ونظراً لأن البعض يلجأ للنموذج الأحادي لسهولته وقلة الحسابات الرياضية المعقدة فيه، والافتراض أن جميع الفقرات يجب أن تكون قدرتها على التمييز متساوية، ففي حالة اختلاف قدرة الفقرات على التمييز يصبح استخدام هذا النموذج غير صحيح لذا يلزم استخدام النموذج ثنائي

المعلمة لتحديد معالم الفقرات (الصعوبة والتمييز). والتدريج يعني أن يتحدد موقع كل فقرة على نفس المقياس الواحد الذي يأخذ في اعتباره شيئين معاً:

أولاً: مستوى صعوبة / سهولة الفقرة.

ثانياً: قدرة الفرد على الإجابة على الفقرة إجابة صحيحة.

وطبقاً لذلك فإننا نتوقع أن يزودنا الاختبار المدرج بنتيجة إجابة الفرد عنه بمعلومات حول خصائص الفرد بمستوى عال من الدقة. والبنوك المدرجة تسهل عملية إضافة أسئلة جديدة لمخزونها، ويتحدد موقع هذه الأسئلة على المقياس المتدرج الذي تقوم عليه أسئلة هذه البنوك.

وفكرة التدريج هذه تتيح إجراء مقارنات بشكل دقيق بين مجموعات من نتائج الاختبارات الفرعية المسحوبة من البنك، ولإنشاء وتشغيل البنوك المدرجة من الضروري توفر أجهزة الكمبيوتر والبرمجيات اللازمة.

المرحلة الرابعة: الترميز ومعلومات الاختبار Coding of Item and Test Information

يعد تنظيم بنك الأسئلة من الأمور المهمة، إذ من الصعب أن يحقق البنك الفائدة المرجوة منه دون تنظيم دقيق وإدارة محكمة. فأساليب الإيداع والسحب من رصيد الأسئلة والفقرات الاختبارية، والتطوير المستمر لمخزون البنك، وعمليات الإضافة والحذف من رصيده، وإجراءات مراقبة هذه العمليات ومتابعتها والإشراف عليها والتحكم بها وبحركتها تتطلب جميعها إجراءات تنظيمية دقيقة، وضوابط فنية، كما تتطلب أن تكون الهيئة الإدارية المسؤولة عن بنك الأسئلة ذات كفاءة عالية بالإضافة إلى توافر برامج عالية الجودة وقادرة على استيعاب عمليات التخزين والاستعادة.

وعملية إدخال الفقرات تعتمد على مصدر هذه الفقرات وقدرة البرنامج المستخدم على كيفية التعامل مع الأشكال والنصوص ونوعية الفقرات المستخدمة وحجم البنك وسهولة نشر وإضافة فقرات جديدة وسهولة التصنيف.

وقد أشار وارد و موري وارد(Ward&Murray-ward ,1994) إلى أن السبب الرئيس لترميز الفقرات هو تبسيط عملية الاسترجاع لها من أجل بناء الاختبارات وكذلك عمل مراجعة دورية لها . كما أشار بيلو وآخرون (Bello, et.al 2000) إلى أنه يجب أن تكون الفقرات مصنفة بشكل منتظم لتيسير عملية استرجاعها واستخدامها.

ومن خلال بنك الأسئلة فإن من المفيد تصنيف الفقرات بطرق متنوعة تشبه طرق مستخدمي الكمبيوتر، وأمناء المكتبات، وهناك طرق تصنيف متعددة يمكن تأسيسها بحيث ترمز كل فقرة في كل طريقة تصنيف، وان الرموز التصنيفية مفيدة من خلال عملية بناء الاختبار للتأكد من أن الاختبار يسير وفق الخطة الموضوعة.

وهناك طرق مختلفة لتصنيف الفقرات تشمل على:

1. التصنيف حسب المحتوى. مثال، فقرات الرياضيات يمكن تصنيفها في (الجبر، والهندسة، و القياس.....).
2. التصنيف حسب الغرض. مثال ، الفقرات الاختبارية المستخدمة لتحديد الكفاءة في أعمال معينة (الاستيعاب ، الاستقراء ، ...) .
3. التصنيف حسب نوع الفقرة. مثال، إذا كان هناك فقرات تتطلب نوعا معينا من العمليات المعرفية للإجابة، يمكن تصنيفها حسب العمليات المعرفية (حل المشكلات، تذكر، تفسير....) ويمكن الدمج بين هذه التصنيفات المختلفة وتصنيف الفقرة على أنها رياضيات / جبر وغرضها الاستقراء ونوعها حل مشكلات.

ولقد اقترح هاثوي (Hathway, 1985) المشار إليه في (Bello, etal, 2000) بناء ملف لكل فقرة يشتمل على رقم الفقرة، رقم بنك الأسئلة، مفتاح الإجابة، رقم نموذج الاختبار، تاريخ الاختبار، عدد الطلبة الذين أجابوا عن هذه الفقرة، تدريج راش، الخطأ المعياري للتدريج، معامل ارتباط بايسيريال، النسبة المئوية للطلبة الذين أجابوا عنها بشكل صحيح.

المرحلة الخامسة: مرحلة الاستخدام والتقييم للبنك:

Using and Assessment of Item Bank

يتحدد محك فاعلية نظام بنك الأسئلة في ضوء الاستخدام الميداني الفعلي للأسئلة والفقرات الاختبارية المودعة فيه، ومما لا شك فيه أن نتائج الاستخدام وما يترتب عليه تكون بمثابة تغذية راجعة تعين في تقييم فاعليته وإعادة النظر في مكوناته، وأسلوب تنظيمه وتحديثه بصفة مستمرة، إذ يمكن تشكيل اختبارات كاملة بمواصفات محددة من رصيد البنك يمكن على أساسها تقويم تحصيل الطلاب في المجالات الدراسية المختلفة، والإفادة من التطبيق الميداني في المراجعة المستمرة

لأسئلة البنك وفقراته سواء بالإضافة أو الحذف أو إعادة التعيير (التدريج)، أو ربما إعادة النظر في المنهج الدراسي وفقاً لما تسفر عنه نتائج الاستخدام.

كذلك ينبغي تنظيم الخدمات الاختبارية التي يمكن أن يقدمها البنك للمعلمين والطلاب، مثل الاختبارات التدريبية أثناء العام الدراسي، وتصحيح الاختبارات الموضوعية، وبخاصة الاختيار من متعدد، وإعداد دليل متجدد كل عام يوضح للمعلمين والمعنيين بنظام البنك مواصفات ما يشتمل عليه من أسئلة وفقرات اختبارية ومجالاتها، وبذلك يسهم هذا النظام البنكي للأسئلة في تطوير عمليات قياس وتقويم التحصيل الدراسي وجعلها أكثر مرونة وفاعلية ودقة، وجعلها أكثر تقنيناً مما هي عليه في إطار الممارسات الحالية، كما يسهم في تزويد المعنيين ببيانات تتعلق بالتقدم الدراسي للطلاب في وقت قصير، وبذلك يكون البنك مكونا رئيسا من مكونات العملية التعليمية، ويسهم إسهاما فاعلاً في تطويرها ، كما تصبح عمليتا التعليم والتقويم عملية واحدة متكاملة (Wright and Bell, 1984).

إدارة وتشغيل بنوك الأسئلة.

نظراً لضخامة بنوك الأسئلة لما تشتمل عليه من فقرات وما يرتبط بها من معالم الصعوبة، والتمييز، وقدرات الأفراد، فمن الطبيعي استخدام الكمبيوتر في حساب هذه المعالم، وقد أشار شكري (2001) إلى ضرورة توافر ثلاث برمجيات (Computer Software) لتشغيل وإدارة بنوك الأسئلة والبرمجيات هي:

1. برنامج قاعدة بيانات (Data Base Program) حيث يكون مناسباً للتخزين والبحث والاستدعاء وتجهيز الورقة الامتحانية وطباعتها.
2. برنامج (Software) يمكن بواسطته تصحيح استجابات الطلبة على الاختبارات، وعرض النتائج، وإجراء كل الترابطات الممكنة بين فقرات الاختبار طبقاً لاستجابات الطلبة عليها.
3. برنامج تحليل إحصائي يناسب النموذج أو النظرية المستخدمة. يستخدم هذا البرنامج في حساب المعالم الخاصة بكل فقرة طبقاً لنظرية الاستجابة للفقرة، كما ويستخدم في التحليل الإحصائي التقليدي.

وفي هذه الدراسة تم استخدام برنامج (bilog-mg) لتحليل الفقرات باستخدام النموذج ثنائي المعلمة، وكذلك لتحليل الفقرات وفق النظرية التقليدية، لما يتمتع به هذا البرنامج من خصائص تميزه عن بقية البرامج ومن بين هذه الخصائص ما يلي:

1. يتعامل البرنامج مع الفقرات الثنائية (الاختيار من متعدد، الصح أو الخطأ).
2. ينجز تحليل الفقرات ووضع العلامات لأي عدد من الاختبارات الفرعية أو المقاييس الفرعية.
3. يفحص الملاءمة لكل فقرة(Tests of fit for individual items) .
4. يقدر توزيعات السمة(Estimation of latent distributions) .
5. يساعد في إدراج فقرات في الاختبار لتقدير إحصائيات الفقرة دون أن تحسب في علامات الممتحنين (فقرات ثابتة).
6. يتفحص ملائمة الفقرات للنموذج.
7. يقدر الأخطاء المعيارية لكل معالم الفقرة ويدرج العلامات.
8. يربط النماذج الاختبارية من خلال فقرات مشتركة.
9. تعدد الخيارات لمعالجة الإجابات المحذوفة (Omitted Responses).
10. يحلل كل من معلومات الفقرة ومعلومات الاختبار.
11. يرسم اقتران استجابة الفقرة ومنحنيات معلومة الفقرة ومنحنيات معلومة الاختبار.
12. يحفظ كل النتائج الهامة في ملفات منفصلة.

بالإضافة إلى أنواع البرمجيات المذكورة سابقاهناك نظامان فرعيان يحتاج اليهما لإدارة وتشغيل بنوك الأسئلة هما:

1. نظام لإنتاج الفقرات. ويتضمن بالإضافة إلى عملية توليد الأسئلة تدريج الأسئلة وصيانتها.
2. نظام خدمات ودعم. من الضروري أن يكون هناك أشخاص ذوو خبرة ومتفرغون يتولون القيام بإدارة البنك وتشغيله وصيانته.

سحب الاختبارات من بنك الأسئلة.

بعد أن تعد بنوك الأسئلة باستخدام نظرية الاستجابة للفقرة من فقرات ملائمة للنموذج المستخدم، تشكل هذه الفقرات فيما بينها نسيجا من الاختبارات مترابطة مع بعضها بفقرات مشتركة. وتتدرج فقرات هذه الاختبارات جميعها في تدريج واحد مشترك، يتمركز حول نقطة صفر واحدة، وبذا فهو يعرف مستويات مختلفة متدرجة من المتغير المراد قياسه.

ويتحكم في بناء بنك الأسئلة، عدد الفقرات التي نود تدريجها، ومدى الصعوبة التي سيغطيه هذا البنك، والحد الأعلى لعدد الفقرات التي تكون الصورة الواحدة من الاختبارات الفرعية.

ويتميز بنك الأسئلة باستقلالية القياس، وهذا يتيح تحرر تقديرات الأفراد من تأثيرات الفقرات المستخدمة، وهذا يعني تكافؤ تقديرات الأفراد، مهما اختلفت الفقرات المستخدمة، بشرط مناسبتها لمستوى الأفراد.

ونظرا لاحتواء البنك على المدى الواسع من الفقرات المتناظرة والمتكافئة في الصعوبة، فانه يوفر فرصة اختيار مجموعات الفقرات التي تشكل صورا اختبارية مختلفة تناسب مجموعات الأفراد المتشابهة في مستوى القدرة.

وهكذا يشكل بنك الأسئلة مصدرا هاما لعائلة من الاختبارات، التي تعرف مدى واسعا من المتغير، وسواء كانت هذه الاختبارات طويلة أم قصيرة، سهلة أم صعبة، واسعة أم ضيقة، فإنها تتعادل في تقديرها لمستوى قدرة الأفراد. وهذا يساعد في حل مشكلة بناء الصور الاختبارية المختلفة، بما يتيحه بنك الأسئلة من مرونة في اختيار اختبارات مختلفة والتي تتعادل تقديراتها للأفراد بصورة مباشرة. وهكذا يمكن أن نقارن بين مستويات القدرة للأفراد، كما ويمكن قياس التغير الذي يحدث في مستوى الفرد، ويكون قياس هذا التغير باستخدام أي مجموعة من فقرات هذا البنك، طالما أنها مناسبة للفرد، وتكون مناسبة للفرد عندما تقترب المميزات الاحصائية للاختبار من المميزات التي يمكن أن يتصف بها الفرد. وعلى هذا الأساس يمكن اختيار الفقرات التي تكون الاختبار المناسب لقياس قدرته من فقرات البنك.

إذا فرضنا أن المميزات الإحصائية للأفراد المراد تقدير قدراتهم في الصورة الآتية(G(M,S,D حيث تشير كل من:

G: إلى المهارة المراد قياسها.

M: إلى متوسط تقدير قدرة الأفراد.

S: إلى الانحراف المعياري لتقدير قدرة الأفراد.

D: إلى شكل التوزيع.

ويكون تقدير هذه المميزات الخاصة بقدرة الأفراد بصورة تقريبية، إما عن طريق خبرة الباحث، وإما عن طريق تجربته لبعض الصور المختصرة من المقياس. أما المميزات الإحصائية التي ينبغي توفرها في الاختبار المناسب، فيمكن التعبير عنها (T(H, W, L حيث تشير:

T: إلى الاختبار المناسب للقياس.

H: إلى متوسط صعوبة الفقرات المختارة، التي تكون الاختبار.

W: إلى مدى صعوبة فقرات الاختبار.

L: إلى طول الاختبار، وهو عدد فقرات الاختبار.

أما المميزات الإحصائية التي تتوفر فعلا في الاختبار المستخدم، فيرمز لها t(h,w,l) حيث ترمز :

t إلى الاختبار الفعلي المستخدم

h إلى متوسط صعوبة الفقرات الفعلية

w مدى الصعوبة الفعلي

وقد لخص رايت وستون (Wright & Stone, 1979) الخطوات التي يمكن إتباعها لتصميم الاختبار T (H,W,L) ليناسب الأفراد الذين نهدف إلى تقديرهم G (M,S,D) كما يلي :

1. من افتراضنا لتقدير متوسط القدرة (M) نحدد المتوسط المناسب لصعوبة الفقرات M=H.
2. من افتراضنا للانحراف المعياري للقدرة (S) نحدد التشتت المناسب لصعوبة الفقرات (4S=W). (Wright&stone, 1979, p139)
3. تبعاً للخطأ المعياري للقياس (SEM) ، والذي تتطلبه دقة القياس ، نحدد طول الاختبار من العلاقة التالية :

$$L = C/ SEM^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

حيث $C = 6$ أو $C = 6S$.

4. من H,W,L يمكن تحديد مجموعة من الفقرات{ δi }تبعاً للصورة الأتية .

$$\delta i = H - (W/2)((L-2i+1)/ L) , i = 1,L \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

بعد ذلك يحدد الاختبار الذي يستخدم فعلاً t (h,w,l) من هذا التصميم للاختبار المناسب T (H,W,L)

5. نختار مجموعة فقرات di من بنك الأسئلة بحيث تكون أقرب ما يمكن لمجموعة تقديرات الفقرات { δi } وذلك بالتقليل من الفرق (di- δi) قدر المستطاع.
6. عندئذ تحسب متوسط صعوبة الاختبار وانحرافه المعياري.

$$h = \sum_{i}^{L} d_i / L = d \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

$$w = [(d_L + d_{L-1} - d_2 - d_1)/2]\ [L/(L-2)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

حيث d_L, d_{L-1} هما صعوبة أصعب فقرتين وهما L,L-1 و d_1, d_2 هما صعوبة أسهل فقرتين وهما الفقرتان الأولى والثانية .

7. عندئذ تكون مجموعة تقديرات { di} هي الاختبار t (h,w,1) الذي يستخدم في تقدير قدرة الأفراد المراد تقديرها .

8. باستخدام قيم الصعوبة المقدرة في بنك الأسئلة لهذه الفقرات المختارة التي تكون الاختبار، يعوض في المعادلة (1) لتحديد تقديرات الأفراد المقابلة لكل درجة كلية على هذا الاختبار الفرعي، وهذه التقديرات تعادل تقديرات الأفراد أنفسهم على أي اختبار فرعي يسحب من البنك نفسه، وإن اختلفت الدرجة الكلية للفرد على الاختبارين.

نظريات القياس وبنوك الأسئلة.

يعد تحليل الفقرات من الإجراءات الأساسية والهامة في إعداد بنوك الأسئلة، حيث يجري تحليل هذه الفقرات باستخدام عينات كبيرة من الطلبة، من خلال مواقف اختبارية متعددة بهدف تمحيصها وفرزها وانتقاء الصالح منها تمهيداً لاستخدامها في إعداد الاختبارات التي تستخدم كأدوات في عمليات التقويم. لذا فإن بنوك الأسئلة ليست مجرد تجميع فقرات أو أسئلة، وإنما هي أنظمة اختبارية تتطلب تنظيماً وتحديداً للموازين وتعييراً للفقرات الاختبارية لكي تزودنا بأساس علمي لقياس يتصف بالتعميم.

وهذا الوضع لا يتأتى إلا إذا استند تصميم بنوك الأسئلة إلى نظرية أو نموذج إحصائي سيكومتري يفسر كيفية الاستجابة للفقرات الاختبارية ويتم في ضوئه اختيار هذه الفقرات لتضمينها في البنك. ومن هذا المنظور تبدو الحاجة إلى نظريات القياس في مجال بناء وتصميم بنوك الأسئلة.

نظرية القياس التقليدية (Classical Test Theory).

ظهرت هذه النظرية في بداية القرن العشرين، ويعود الفضل إلى العالم البريطاني سبيرمان (Spearman)، وقد عمل العديد من علماء القياس على بلورة أفكار سبيرمان وصياغتها بأن العلامة الملاحظة لاختبار ما تتكون من العلامة الحقيقية والخطأ(Croker & Algina,1986) .

وترتكز نظرية القياس التقليدية على مجموعة من الافتراضات هي:

1. الوسط الحسابي للأخطاء العشوائية لمجتمع الأفراد يساوي صفراً.
2. لا توجد علاقة بين العلامات الحقيقية للممتحنين وأخطاء القياس ($r_{te}=0$).
3. خطأ القياس هو خطأ عشوائي وهو يحدد دقة القياس أو ثبات الاختبار، أما الخطأ المنتظم فهو يعزى لصدق الاختبار.
4. تباينات أخطاء القياس لجميع الأفراد متساوية.

5.يمكن تقدير العلامة الحقيقية من خلال الوسط الحسابي للعلامات الملاحظة الناتجة من تطبيق الاختبار مرات عديدة (Lord, 1980; Croker & Algina, 1986).

وقد أشار هامبلتون وسواميناثان(Hambelton&Swaminathan,1985) إلى أن النظرية التقليدية تعتمد على علامة الاختبار الكلية دون اعتبار لاستجابة الفرد على فقرة معينة بهدف التنبؤ بأدائه .

تحليل الفقرات الاختبارية حسب النظرية التقليدية.

يعنى القائمون على عملية القياس والتقويم التربوي بكتابة فقرات عالية الجودة لقياس السمات المطلوبة قياساً دقيقاً، لذا لابد من مراعاة شروط كثيرة أثناء كتابة هذه الفقرات وصياغتها والتحقق من صدق محتوى كل فقرة وكذلك صدق محتوى الاختبار الذي يتشكل منها، بالأساليب المنطقية وأحكام الخبراء، غير أنه مهما بلغت دقة هذه الأساليب والأحكام فإنها لا تغني عن استخدام الأساليب الإحصائية لتحديد العلاقة بين ما تقيسه الفقرات وبين استجابات الأفراد لها وذلك للتعرف على الخصائص السيكومترية للفقرات (الصعوبة، و التمييز، وفاعلية البدائل).

1. **معامل صعوبة الفقرة** (Item Difficulty (p)).

تعد صعوبة فقرات الاختبار من الخصائص التي تلعب دوراً مهماً في بناء الاختبارات كما تؤثر في إجابات الأفراد على فقراتها، ويتم حساب معامل صعوبة الفقرة عن طريق إيجاد نسبة الذين أجابوا إجابة صحيحة إلى عدد الأفراد الذين حاولوا الإجابة عن الفقرة، وكلما زادت هذه النسبة دل ذلك على سهولة الفقرة وكلما قلت دل ذلك على صعوبتها، وأن أقل قيمة لهذا المعامل (صفر) وأكبر قيمة (1.00) وقد أشار عودة (2002) بأن أي فقرة ضمن مدى صعوبة من 0.2 – 0.8 يمكن أن تكون مقبولة ويحتفظ بها، أما الفقرات الصعبة جداً أو السهلة جداً فيتوقع استبعادها وفقا للغرض من الاختبار، إلا انه من الناحية النظرية يمكن أن نتصور طيفا واسعا من الفقرات موزعة توزيعا اعتداليا مع إمكانية وجود طرفين بعتبة = صفر وسقف = 1.0

2. **معامل تمييز الفقرة** (Item Discrimination (D)).

من الخصائص الأخرى المهمة التي ينبغي أن تتوافر في فقرات الاختبار خاصية التمييز. ونعني بذلك إمكانية قياس الفروق الفردية بواسطة فقرات هذا الاختبار (علام، 2002)، وهو كذلك يمثل الفرق بين نسبة الأفراد الذين أجابوا إجابة صحيحة من الفئة العليا ونسبة الأفراد الذين أجابوا إجابة صحيحة من الفئة الدنيا، ومن بين الأساليب التي يمكن استخدامها في حساب معامل تمييز الفقرة

استخدام أسلوب معامل الارتباط الثنائي المتسلسل النقطي (Point Biserial) الذي يمثل العلاقة بين علامة الفقرة والعلامة الكلية لكل مفحوص. والقيمة التي تحصل عليها تدعى معامل التمييز ، الذي يدل على مدى اتساق علامة كل فقرة مع علامات الاختبار .

والصيغة الرياضية لمعامل الارتباط الثنائي المتسلسل النقطي هي:

$$r_{pbis} = \frac{\bar{x}_{pi} - \bar{x}_{fi}}{S_t} \times \sqrt{p_i q_i} \quad \text{.....(5)}$$

حيث أن:

(r_{pbis}): معامل الارتباط الثنائي المتسلسل النقطي.

($\bar{x}_{pi}$) : متوسط علامات الطلبة الذين أجابوا الفقرة إجابة صحيحة .

($\bar{x}_{fi}$): متوسط علامات الذين لم يجيبوا على الفقرة.

(S_t) : الانحراف المعياري للاختبار .

(p_i) : معامل صعوبة الفقرة .

(q_i) : 1 – P

وقد أشار ايبل (Ebel, 1965) المشار إليه في (Crocker & Algina, 1986) إلى أن قيم معامل التمييز تتراوح بين -1، + 1، وإذا كانت قيمة معامل تمييز الفقرة مساوية (0.4) أو أكثر دل هذا على أن الفقرة تميز بدرجة جيدة بين المجموعتين المتطرفتين ويمكن الاحتفاظ بها، وإذا كانت قيمة معامل التمييز للفقرة محصورة بين 0.3 إلى 0.39 فالفقرة مقبولة وينصح بتحسينها، أما الفقرات ذات معامل التمييز المحصورة بين 0.2 إلى 0.29 تكون هامشية وبحاجة إلى مراجعة وتنقيح، أما الفقرات ذات معامل التمييز الذي تقل قيمته عن 0.19 فتعد ضعيفة التمييز وينبغي حذفها من الاختبار أو إعادة النظر في صياغتها، وربما يتم الاحتفاظ بعدد قليل جدا من هذه الفقرات في ضوء مبررات يحددها باني الاختبار.

3. تقدير فعالية المموهات (البدائل).

المموه هو البديل الخاطئ في فقرة الاختيار من متعدد وقد أشار عوده (2002) إلى أن المموه الجيد هو ذلك المموه الذي يتم اختياره من قبل 5 % من الطلبة أو أكثر وتزداد فاعلية المموهات بزيادة

قدرتها على اجتذاب العدد الأكبر من الطلبة من الفئة الدنيا، كما أشار ألن و ين (Allen & Yen, 1979) إلى اعتبار المموه جيداً إذا كان معامل تمييزه سالباً.

نظرية الاستجابة للفقرة (Item Response Theory).

على الرغم من الاستخدام الواسع النطاق للنظرية التقليدية في تصميم الاختبارات وتحليل وتفسير نتائجها، إلا أن هناك بعض المشكلات التي تقلل من دقة وموضوعية هذا الاستخدام ويمكن تلخيص تلك المشكلات على النحو الاتي:

1. تأثر خصائص فقرات الاختبار بقدرة الأفراد المختبرين: حيث تختلف معاملات الصعوبة والتمييز لفقرات الاختبار باختلاف قدرة أفراد العينة، فالفقرة التي يختبر بها أفراد متمكنون تبدو سهلة لهم بينما تبدو صعبة للأفراد غير المتمكنين. وإذا كانت عينة الأفراد متجانسة القدرة نسبياً، فإن قيم معاملات التمييز للفقرات تكون أقل من القيم التي يتم الحصول عليها من عينة غير متجانسة (Hambleton & Swaminathan, 1985، علام، 1986) لذا فإن نتيجة القياس تختلف باختلاف خصائص العينة موضع الاختبار، أي لا يتوقع ثبات خصائص الفقرات.

2. افتراض تساوي تباين أخطاء القياس لجميع أفراد العينة موضع الاختبار، بالرغم من أنه قد يكون أداء بعض الأفراد على الاختبار أكثر اتساقاً من غيرهم من الأفراد، وأن درجة هذا الاتساق تختلف باختلاف مستوى قدرة الأفراد، أو بمستوى القدرة التي يقيسها الاختبار (Fan,1998) بمعنى أن هذا الافتراض قد لا يتحقق أو من الصعب تحقيقه.

3. لا تؤخذ بعين الاعتبار استجابة الفرد على فقرة معينة ولكنها تعتمد إحصائيا على الدرجة الكلية، ولا توفر قاعدة تساعد على التنبؤ بكيفية أداء الفرد على أي فقرة من فقرات الاختبار، لذا تختلف نتيجة القياس باختلاف الاختبار المستخدم (Hambleton & Swaminathan, 1985) . بمعنى أن قدرة الفرد تعتمد على خصائص عينة الفقرات في الاختبار.

4. صعوبة إثراء الاختبارات المعدة باستخدام النظرية التقليدية بفقرات جديدة: حيث يتعرض المنهاج المدرسي لتغيرات متعددة، قد ينتج عنها إضافة بعض الموضوعات، وإعداد فقرات جديدة لهذه الموضوعات، ودمجها ضمن الفقرات المعدة سلفاً يتطلب إعادة تصنيف جميع الاختبارات مرة أخرى (الشربيني وحجازي، 2002).

5. تأثر ثبات الاختبار بالموقف الاختباري: حيث يعتمد ثبات الاختبار في إطار النظرية الكلاسيكية (التقليدية) إما على إعداد الصور المتكافئة – وإعداد هذه الصور يعد أمراً صعباً – أو على تطبيق الصورة الاختبارية مرتين على أفراد العينة، وبالرغم من أهمية ذلك إلا أنه غير كاف ; حيث يمكن أن يختلف الموقف الاختباري وظروف التطبيق في هاتين المرتين، الأمر الذي يؤثر على دقة ثبات

الاختبار (Hambleton & Swaminathan, 1985,de Gruidter & Vand derkamp, 2003) وقد استثارت تلك المشكلات فكر علماء التربية للبحث عن الدقة والموضوعية في القياس والتقويم ، حتى يقترب القياس السلوكي من القياس في العلوم الطبيعية ، التي تتميز بعدم تأثر نتائج القياس بالأداة المستخدمة- طالما أنها أداة مناسبة لتقدير الظاهرة – كما يكون تدرج الأداة بوحدات قياس متساوية ، لا تعتمد ولا تتأثر بالعناصر التي تقدر عندها الظاهرة .

وقد أسفرت جهود الباحثين في مجال التقويم التربوي عن نظرية تحت المسميات التالية:

* نظرية السمات الكامنةLTT– (Latent Trait Theory) .
* أو نظرية استجابة الفقرة IRT– (Item Response Theory) .
* أو نظرية منحنى خصائص الفقرة ICC– (Item Characteristic Curve) .

وتتميز نظرية الاستجابة للفقرة بما يلي :

أولاً: إمكانية التنبؤ بأداء الفرد وتقدير درجة سمته أو قدرته الكامنة وراء هذا الأداء بغض النظر عن الأداة المستخدمة (Robi & Patricia, 1996) .

ثانياً: تقدم مؤشرات إحصائية ثابتة للاختبار وفقراته لا تختلف باختلاف عينة الأفراد وتوفر هذه النظرية قياسا أكثر ثباتاً (Nitko,2001) .

ثالثاً: يمكن تفسير درجة الفرد على الاختبار في ضوء محكات محددة سلفاً وليس بإرجاعها إلى متوسط العينة التي ينتمي إليها الفرد (Embretson & Reise, 2000).

وانبثقت عن نظرية السمات الكامنة بعض النماذج اللوغارتيمية التي تستهدف تحديد العلاقة المتوقعة بين استجابات الأفراد لاختبار معين، والسمات أو القدرات الكامنة وراء هذه الاستجابات وتفسرها، وذلك من خلال أسس إحصائية تؤدي إلى موضوعية القياس المتمثلة في تحرر خصائص أدوات القياس وفقراتها من خصائص الأفراد المختبرين بها، وكذلك تحرر تقديرات أداء الأفراد من خصائص عينة الفقرات التي يختبرون بها، وفيما يلي نبذة عن بعض تلك النماذج:

نماذج نظرية الاستجابة للفقرة.

نظراً لاعتماد نظرية الاستجابة للفقرة على فرضية أساسية مؤداها أن القيمة الاحتمالية لاستجابة فرد لفقرة اختباريه تكون اقتراناً لكل من القدرة التي يفترض أن الاختبار يقيسها لدى الفرد، وخصائص الفقرة التي يحاول الإجابة عنها، فإن ذالك يتطلب الحصول على معلومات من مصدرين: أحدهما يتعلق بالفرد، والأخر يتعلق بالفقرة الاختبارية، وعادة نحتاج إلى قيمة عددية واحدة تتعلق بالفرد، وهو معلم القدرة المقاسة لدى الفرد Ability Parameter وقيمة عددية أو

أكثر تتعلق بالفقرة الاختبارية (معلم أو معالم الفقرة) Item Parameters وفي ضوء عدد معالم الفقرات، فإنه يندرج تحت نظرية السمات الكامنة مجموعة من النماذج السيكومترية من أهمها: (Baker, 2001;de Gruidter & Vand derkamp, 2003)

1. **النموذج اللوغارتيمي ثلاثي المعلمة (3PL) (Three Parameters Logistic Model)** ويفترض هذا النموذج أن الفقرات تختلف في صعوبتها وتمييزها وكذلك في معلم التخمين الذي يمثل احتمال توصل الأفراد ذوي القدرة المنخفضة إلى الإجابة الصحيحة عن الفقرة وخاصة في الفقرات من نوع الاختيار من متعدد. ومعادلة النموذج الثلاثي المعلمة هي :

$$P_i(\theta) = c_i + \frac{(1-c_i)e^{D_{ai}(\theta-b_i)}}{1+e^{D_{ai}(\theta-b_i)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

حيث:

c_i : معلم التخمين

a_i : معلم التمييز، ويعرف على أنه ميل المماس لمنحنى خاصية الفقرة عند النقطة التي تحدد صعوبتها.

b_i : معلم الصعوبة، وهي القيمة التي يكون عندها احتمال إجابة الفقرة إجابة صحيحة هي $\frac{c+1}{2}$

حيث C احتمال التخمين .

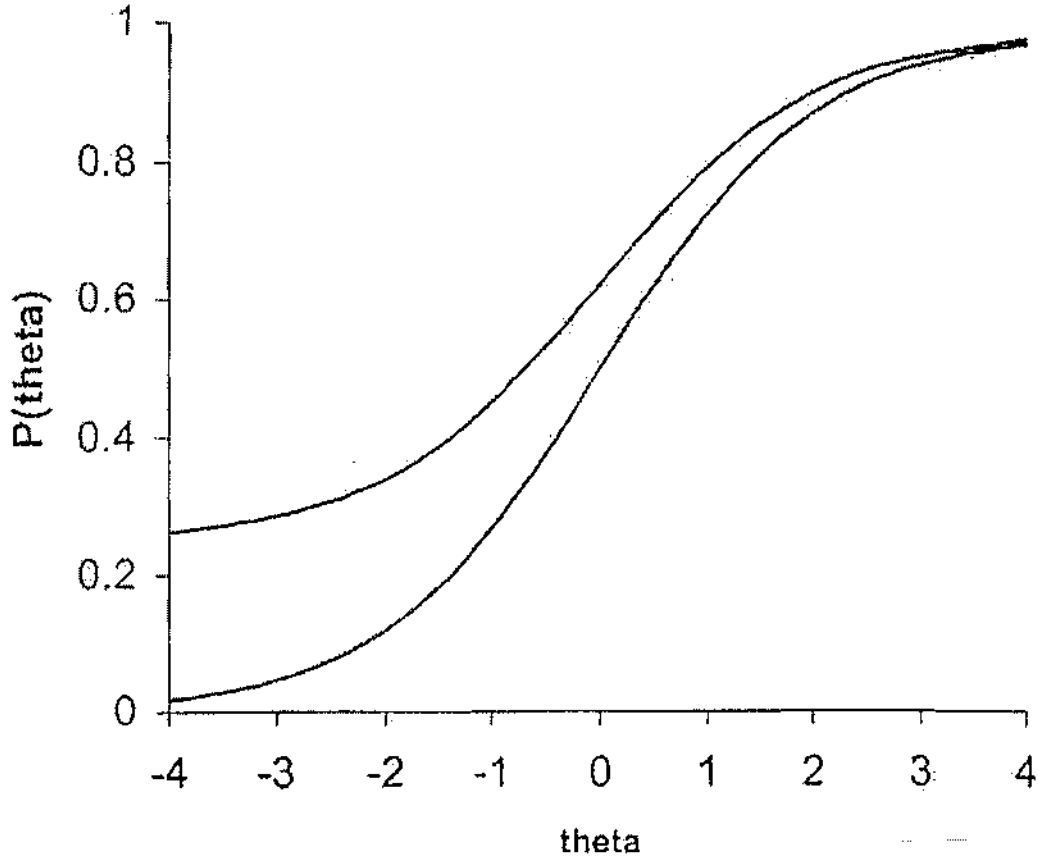
D : ثابت وقيمته تساوي 1.7 .

θ : مستوى قدرة الفرد.

e : الأساس اللوغارتيمي ويساوي (2.7183) .

$P_i(\theta)$: احتمال إجابة الفقرة (i) من قبل الفرد الذي قدرته(θ) إجابة صحيحة.

والشكل 1 يحدد العلاقة بين قدرة الفرد (θ) واحتمال الحصول على الإجابة الصحيحة لفقرتين مختلفتين في معلم التخمين.



قدرة المفحوص (θ)

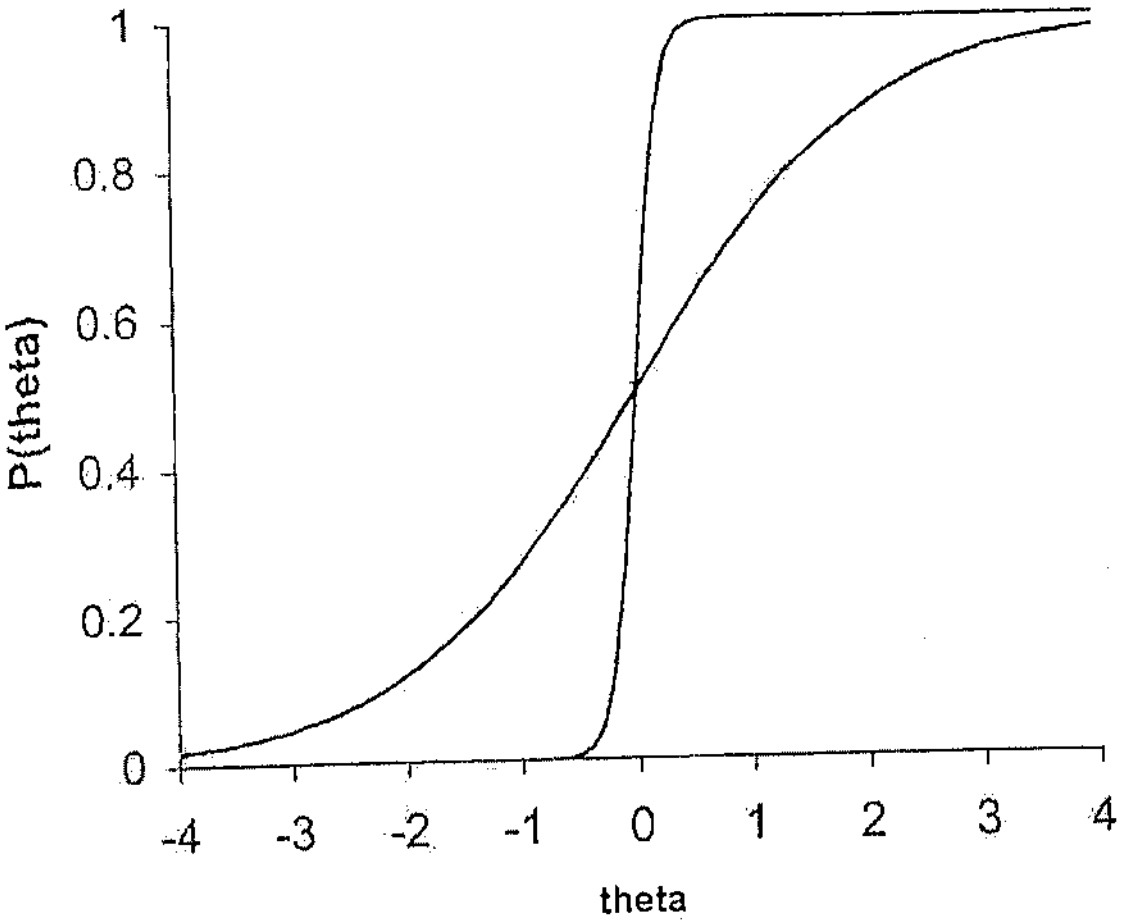
شكل 1: العلاقة بين قدرة الفرد θ واحتمال الحصول على الإجابة الصحيحة لفقرتين مختلفتين في معلم التخمين

2.النموذج اللوغارتيمي ثنائي المعلمة (Two Parameters Logistic Model) (2PL)

يفترض هذا النموذج أن الفقرات تختلف في صعوبتها وتمييزها بين المستويات مختلفة القدرة، كما يفترض أن معلم التخمين يساوي صفرا. ومعادلة النموذج الثنائي المعلمة هي :

$$P_i(\theta) = \frac{e^{D a_i(\theta - b_i)}}{1 + e^{D a_i(\theta - b_i)}} \quad \text{.....(7)}$$

والشكل 2 يوضح منحنى خصائص الفقرة لفقرتين مختلفتين في معلم التمييز.



قدرة المفحوص (θ)

شكل 2 : منحنيات خصائص الفقرة لفقرتين مختلفتين في التمييز.

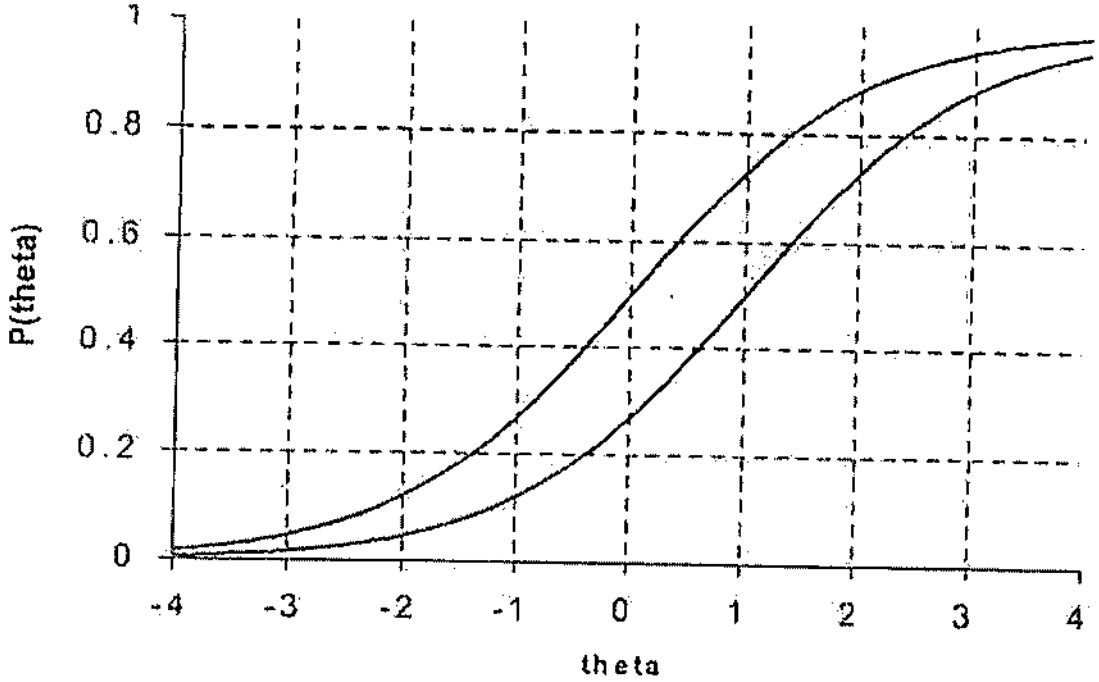
3.نموذج راش: النموذج اللوغارتيمي أحادي المعلمة (1PL)

(One- Parameters Logistic Model)

معادلة نموذج راش تعطى بالعلاقة:

$$P(\theta)=1/(1+e^{-1(0-b)}).....(8)$$

ويفترض هذا النموذج أن فقرات الاختبار تختلف فقط في صعوبتها وتتساوي في تمييزها ولا يوجد أثر للتخمين والشكل 3 يوضح منحنى خصائص الفقرة لفقرتين مختلفتين في الصعوبة ولكن لهما نفس القدرة التمييزية.



قدرة المفحوص (θ)

شكل 3: منحنى خصائص الفقرة لفقرتين مختلفتين في الصعوبة ولهما نفس التمييز.

افتراضات نظرية الاستجابة للفقرة Assumptions of Item Response Theory

تؤكد نظرية الاستجابة للفقرة على أن صعوبة الفقرة وقدرة الفرد يمكن أن توضع على مقياس مشترك للسمة الكامنة وفق الافتراضات التالية: (Hambleton & Swaminathan, 1985):

1. افتراض أحادية البعد Unidimensionality أي أن الفقرات تقيس سمه أحادية، وبمعنى آخر تفترض نماذج السمات الكامنة وجود قدرة واحدة تفسر أداء الفرد في الاختبار. أما النماذج التي تفترض وجود أكثر من قدرة واحدة تكمن وراء هذا الأداء فإنها تسمى نماذج متعددة الأبعاد.

2. الاستقلال المحلي أو الموضعي Local Independence وتظهر في :

أ. تحرر القياس من توزيع العينة المستخدمة Sample Free وهذا يعني ثبات تقدير كل من قدرة الفرد وصعوبة الفقرة واستقرارهما بالرغم من اختلاف عينة الأفراد المستخدمة في تدريج المقياس طالما أنها عينة ملائمة.

ب. تحرر القياس من مجموعة الفقرات المستخدمة Item Free وهذا يعني ثبات تقدير كل من قدرة الفرد وصعوبة الفقرة واستقرارهما بالرغم من اختلاف مجموعة الفقرات المستخدمة في القياس، طالما أنها فقرات ملائمة، وطالما أن هذه المجموعات المختلفة من الفقرات تقع على ميزان تدرج واحد، أي أنها تعرف متغيراً واحداً.

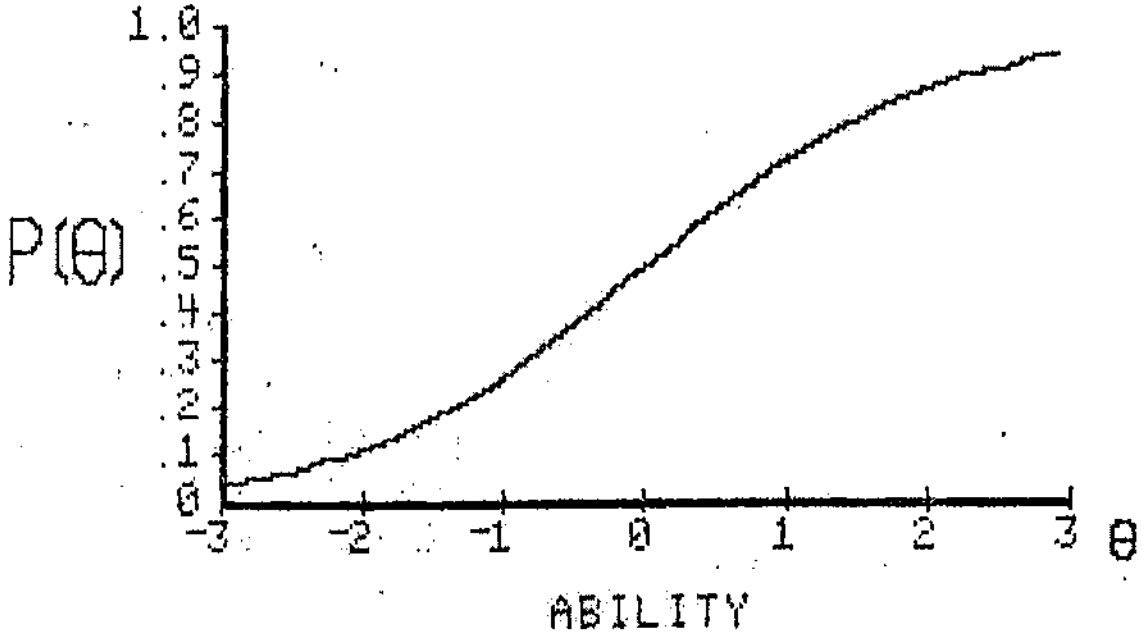
فإذا تحقق فرض الاستقلال المحلي يكون احتمال الحصول على نمط معين لدرجات فرد مساوياً حاصل ضرب احتمالات حدوث هذا النمط لكل فقرة من فقرات الاختبار التي أجاب عنها.

فعلى سبيل المثال فإن احتمال حدوث نمط استجابة فرد على خمس فقرات (10110) = U حيث (0) تعني إجابة خاطئة، (1) تعني إجابة صحية للفقرة يعبر عنه بالشكل P1(1-P2)(P3)(P4)(1-P5)

حيث Pi: احتمال إجابة المفحوص إجابة صحيحة عن الفقرة i.

1-Pi: احتمال إجابة المفحوص إجابة خطأ عن الفقرة i.

3. افتراض منحنى خصائص الفقرة Item Characteristic Curve (ICC). منحنى خصائص الفقرة هو عبارة عن اقتران تراكمي صاعد، من خلاله يمكن تحديد العلاقة بين تحصيل الفرد على الفقرة وقدرته المقيسة بالاختبار، ويوفر هذا المنحنى احتمال إجابة الأفراد عن الفقرة في مستويات القدرة المختلفة إجابة صحيحة وكونه تراكمياً صاعداً، فإنه يشير إلى أن احتمال إجابة فقرة ما إجابة صحيحة يزداد بازدياد قدرة الفرد، ويعزى سبب الاختلاف فيما بين نماذج السمات الكامنة إلى اختلاف شكل منحنى خصائص الفقرة (ICC) والشكل 4 يوضح هذا الافتراض



قدرة المفحوص (θ)

شكل 4: نموذج يوضح منحنى خصائص الفقرة.

4. فرض التحرر من عامل السرعة في الإجابة:Speediness . تفترض نماذج نظرية السمات الكامنة أن عامل السرعة لا يلعب دوراً في الإجابة عن فقرات الاختبار، بمعنى أن إخفاق الأفراد في الإجابة على فقرة ما يرجع إلى انخفاض قدرتهم وليس إلى تأثير عامل السرعة على إجابتهم، ويمكن تقدير ما إذا كان عامل السرعة قد لعب دوراً في الإجابة، وذلك عن طريق معرفة عدد الأفراد الذين لم يتمكنوا من الانتهاء من إجابة جميع فقرات الاختبار الذي تقدموا له. وهذا الافتراض ليس له تأثير مباشر على شكل المنحنى المميز للفقرة وإنما تأثيره غير مباشر. ويعتبر عامل السرعة من العوامل التي تجعل الأفراد يلجأون إلى التخمين في الإجابة وأحياناً تصبح سرعة الأداء إحدى القدرات التي يقيسها الاختبار.

دالة معلومات الفقرة Item Information Function

دالة معلومات الفقرة هي اقتران يمثل العلاقة بين متغيرين هما قدرة الفرد والمعلومات المستخلصة من خلال الفقرات، ويعبر هذا الاقتران عن كمية المعلومات التي تقدمها الفقرة عن مستوى القدرة التي تقيسها، وذلك بتحديد أقصى ارتفاع لمنحنى اقتران معلومات الفقرة عند مستوى معين من القدرة، وبواسطة هذا يمكن تحديد أي الفقرات الاختبارية تقيس المتغير المراد قياسه

بدرجة أفضل عند مستويات محددة للقدرة (Hambleton & Swaminathan, 1985,Baker,2001) . فالمعلومات عند نقطة ما على متصل القدرة تتناسب مع مربع القوة التمييزية للفقرة (Lord,1980).

ويمكن اعتبار إسهام الفقرة في دقة الاختبار كجزء من إسهامات بقية الفقرات وتعطى معلومات الفقرة (i) بالعلاقة:

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)[1 - P_i(\theta)]} \quad \text{.......(9)}$$

حيث $P_i'(\theta)$ هي مشتقة $P_i(\theta)$ بالنسبة لـ θ.

من هذا يتبين أن معلومات الفقرة هي عبارة عن مربع ميل منحني خصائص الفقرة (θ) مقسوما على تباين الخطأ، الذي يكون تباين الفقرة المعطى.

أما اقتران معلومات الاختبار فهو المجموع الحسابي لاقتران معلومات الفقرات المكونة له ويعطى بالعلاقة:

$$I(\theta) = \sum_{i=1}^{N} I_i(\theta) \quad \text{.......(10)}$$

وبالنسبة للنموذج ثنائي المعلمة (2pl) فإن اقتران معلومات الفقرة يعطى بالعلاقة:

$$I_i(\theta) = a_i^2 P_i(\theta)[1 - P_i(\theta)] \quad \text{.......(11)}$$

ومع زيادة القدرة التمييزية (a) تزداد المعلومات عند $\theta = b$ بينما المعلومات عند القدرات البعيدة تتناقص. أما الفقرة التامة (Perfect Guttman Item) فإنها تميز عند نقطة واحدة، بمعنى أنها تميز بين الأفراد ذوي قدرة (θ) أقل من (b) و بين الأفراد ذوي القدرة (θ) أكبر من (b)،والشكل 5 يوضح ذلك. (de Gruijter & Vander Kamp, 2003) .

الشكل 5: فقرة تمييز بشكل تام عند θ =1.5

خصائص اقتران معلومات الاختبار.

يتصف اقتران معلومات الاختبار بالخصائص الآتية:

1- يعرف اقتران المعلومات لمجموعة من فقرات اختبار عند كل نقطة من مقياس القدرة

2- يتأثر كم المعلومات من هذا الاقتران بنوعية وعدد فقرات الاختبار

3- لا يعتمد اقتران معلومات الفقرة على أي تجميع لاقترانات معلومات الفقرات، حيث أن اقتران معلومات الفقرة لا يتأثر باقتران المعلومات للفقرات الأخرى.واقتران معلومات الاختبار هو مجموع حسابي لاقترانات معلومات الفقرات المكونة له .

4- يتناسب كم المعلومات المعطى عند كل مستوى للقدرة (θ) عكسيا مع خطا التقدير عند مستوى القدرة.

$$SE(\theta) = 1 / \sqrt{I(\theta)} \ldots\ldots\ldots\ldots\ldots 12$$

Or $I(\theta) = 1/SE^2(\theta)$

حيث $SE(\theta)$ هو الخطا المعياري لتقدير القدرة عند مستوى معين من القدرة (Hambleton & Swaminathan, 1985;Baker, 2001) ولذلك فان دقة القياس في نظرية الاستجابة للفقرة يمكن أن يحدد عند أي مستوى من مستويات القدرة θ، بالإضافة إلى أنها تمكن باني الاختبار من تحديد الدقة والمعلومات التي تضيفها كل فقرة إلى الاختبار.

تدريج الاختبار

تفسر القيم العددية لمعلمات الفقرة وقدرات الأفراد على مقياس واحد فيه نقطة المنتصف= صفر ووحدة القياس له تساوي (1)، بينما يعرض هذا المقياس المفاهيم الأساسية لنظرية استجابة الفقرة، في حين لا يظهر الوضع الحقيقي للاختبار. فعندما يضع باني الاختبار الفقرة فانه يكون على علم مسبق- بالسمة التي تقيسها هذه الفقرة، وما هو مستوى القدرة اللازمة لها(منخفض، متوسط، عال)- ولكن من غير الممكن تحديد قيم معالم الفقرة مسبقا. فعند تطبيق اختبار على مجموعة من الأفراد لا نستطيع معرفة عدد السمات التي يتجاوزها كل مفحوص. فعملية تحديد معالم الفقرات وقدرات الأفراد باستخدام نظرية الاستجابة للفقرة تدعى التدريج (calibration)، وتزودنا هذه العملية بإطار مرجعي لتفسير نتائج الاختبار.

تتم عملية تدريج فقرات الاختبار بواسطة تطبيق اختبار مكون من(N) من الفقرات الثنائية الاستجابة على مجموعة (M) من المفحوصين. بعد ذلك تجرى عمليات رياضية على البيانات المكونة من استجابات المفحوصين لتكوين مقياس القدرة الذي يربط بين معالم الفقرات ومعالم المفحوصين، ومن ثم تعرض معالم الفقرة وقدرات المفحوصين على هذا المقياس، أي يتحدد موقع كل فقرة على نفس المقياس الواحد الذي يأخذ في حسابه:

1) مستوى صعوبة/سهولة الفقرة

2) قدرة الفرد على إجابة السؤال إجابة صحيحة

وطبقا لذلك فان الاختبار المدرج يتوقع أن يزودنا بنتيجة إجابة الفرد عن فقراته بمعلومات حول خصائص الفرد بمستوى عال من الدقة.

دمج فقرات اختبارين

تهدف هذه العملية إلى تحويل التدريج المستقل لكل من الاختبارين إلى تدريج واحد مشترك، ويتطلب هذا التحويل موازنة أو معادلة لتدريج الفقرات المكونة لكل من الاختبارين، مما يساعد في بناء صور متكافئة من الاختبارات المتنوعة والتي تحافظ على سرية الامتحانات والحد من إمكانية تطبيقها اكثر من مرة على المفحوص.وهناك نوعان من المعادلة :

1)المعادلة الرأسية: Vertical Equating وهي التي تكون فيها الصور الاختبارية مختلفة في مستوى صعوبتها، وكذلك تكون عينات الأفراد التي تطبق عليها هذه الصور مختلفة في مستوى قدراتها.

2)المعادلة الأفقية: Horizontal equating وهي التي تكون فيها الصور الاختبارية متقاربة في مستوى صعوبتها، والهدف من إجراء المعادلة هو تعديل الفروق الناتجة عن الاختلاف في الصعوبة والتمييز.

دمج فقرات اختبارين باستخدام مجموعة من الفقرات المشتركة(الرابطة).

تهدف هذه العملية إلى وضع مقياس عام واحد لفقرات الاختبارين، بغض النظر عن الصورة التي وجدت فيها هذه الفقرة، ولإنشاء هذا المقياس يجب تحويل صعوبة الفقرات ومقياس القدرة إلى مقياس معياري بمتوسط(صفر) وانحراف معياري (1) حتى يمكن إجراء تحويل خطي بين الصور المختلفة. وبوضع المتوسط صفر، تلغى الاختلافات عند نقطة البدء في معايرة الفقرات من الصور المختلفة، وبذلك يمكن تحويل مقاييس جميع صور الاختبار إلى مقياس واحد بوحدات تسمى logits تخزن الفقرات في بنك الأسئلة على أساسه.

وتتم هذه العملية باستخدام النموذج اللوغاريتمي ثنائي المعلمة (2pl) كما يلي:

1- يطبق كل نموذج اختباري على عينة كبيرة من المفحوصين من مجتمعات مختلفة.

2- يدرج كل اختبار على حدة.

3- تستخدم الفقرات الرابطة لتحويل التقديرات إلى مقياس عام مشترك.

4- بافتراض أن توزيعات السمة لها وسط=صفر، وانحراف معياري =1.

5- تحسب الأوساط الحسابية للصعوبات المقدرة للفقرات المشتركة وكذلك الأوساط الهندسية المقدرة لميول الفقرات المشتركة.

ولإجراء ذلك نفترض وجود نموذجين A, B وان الأوساط الحسابية لصعوبات m من الفقرات المشتركة في التدريجات المنفصلة للنماذج هي:

$$\bar{b}_A = 1/m \sum_{j=1}^{m} b_{Aj} \quad \text{.....(13)}$$

$$\bar{b}_B = 1/m \sum_{j=1}^{m} b_{Bj} \quad \text{.....(14)}$$

ولنفترض كذلك أن الأوساط الهندسية لميول هذه الفقرات في تدريج النموذجين هي:

$$\bar{a}_A = \exp\left\{1/m\sum_{j=1}^{m}\log_e a_{Aj}\right\} \qquad (15)$$

$$\bar{a}_B = \exp\left\{1/m\sum_{j=1}^{m}\log_e a_{Bj}\right\} \qquad (16)$$

ولنفرض أن: $r = \sqrt{\bar{a}_A / \bar{a}_B}$

عندها تكون الميول المعدلة لكل الفقرات في النموذج A والنموذج B هو

$$a^*_{Aj} = a_{Aj} / r \qquad (17)$$

$$a^*_{Bj} = r a_{Bj} \qquad (18)$$

ولنفرض أن: $d = (r\bar{b}_A - \bar{b}_B / r)/2$

عندها تكون الصعوبة المعدلة لكل الفقرات هي:

$$b^*_{Aj} = r b_{Aj} - d \qquad (19)$$

$$b^*_{Bj} = b_{Bj} / r + d \qquad (20)$$

التحقق من خصائص اختبار مستل من بنك الأسئلة

عند سحب الاختبارات من بنك الأسئلة ضمن مواصفات محددة، تحدد من خلالها صعوبة وتمييز الفقرات المنتقاة بالإضافة إلى اقتران معلومات الفقرة. هذا كله يساعد باني الاختبار على التحقق من خصائص الاختبار المسحوب من البنك قبل تطبيقه.

ثبات الاختبار.

لقد بينا سابقا أن اقتران معلومات الاختبار يمكن حسابه من خلال جمع اقترانات معلومات الفقرات المكونة له. وبما انه يمكن استخدام اقتران معلومات الاختبار لتحديد مستوى الخطا المعياري في التقدير من خلال المعادلة:

$$SE(\theta) = 1/\sqrt{I(\theta)} \quad ...(21)$$

حيث SE هو الخطا المعياري في التقدير و $I(\theta)$ هو اقتران معلومات الاختبار لذلك فان اقتران معلومات الاختبار يزود باني الاختبار بمستوى الخطا المعياري في التقدير للاختبار. وعند معرفة الخطأ المعياري للاختبار يستطيع باني الاختبار حساب ثبات الاختبار من العلاقة التالية:

$$SE^2=1- r...(22)$$

صدق الاختبار

تخزن الفقرات في بنك الأسئلة بعد أن يتم تحقيق الشروط اللازمة للتخزين، مثل أحادية البعد، والاستقلال الموضعي، وان كلا من صعوبة الفقرات وقدرات الأفراد تتدرج على متصل المتغير المقيس، بالإضافة إلى الإجراءات الأولية من تحليل للمحتوى، واحكام لجنة بناء الفقرات. ولذلك، فان صدق الاختبار المستل من بنك الأسئلة سيكون بالمواصفات التي تمثل فيه الفقرات المستلة منه المجال السلوكي لمجموع فقرات البنك.

تخزين الفقرات

بعد الانتهاء من تحليل استجابات الطلبة على الفقرات، وبعد أن تتم عملية حساب معالم الفقرات وتدريج معالمها على مقياس واحد، تخزن الفقرات في قاعدة البيانات المعدة لذلك وفق المعلومات الأساسية عن كل فقرة وهي:

1- رمز الفقرة ونصها والصورة أو الشكل المرافق

2- البدائل

3- رمز الإجابة الصحيحة

4-المحتوى(المهارة)الذي/التي تعبر عنه الفقرة

5- المستوى المعرفي الذي تعبر عنه الفقرة

6- صعوبة الفقرة

7- تمييز الفقرة

8- اقتران معلومات الفقرة

9- زمن الإجابة

10- عدد مرات الاستخدام

ويجب أن تتسم قاعدة البيانات بالمرونة ليتسنى لباني الاختبار الحصول على فقرات بمعايير محددة ويسهل عليه أيضا عملية حذف فقرات موجودة أو إضافة فقرات جديدة أو تعديل على الفقرات الموجودة تمشيا مع التغيرات التي قد تطرأ على موجودات البنك.

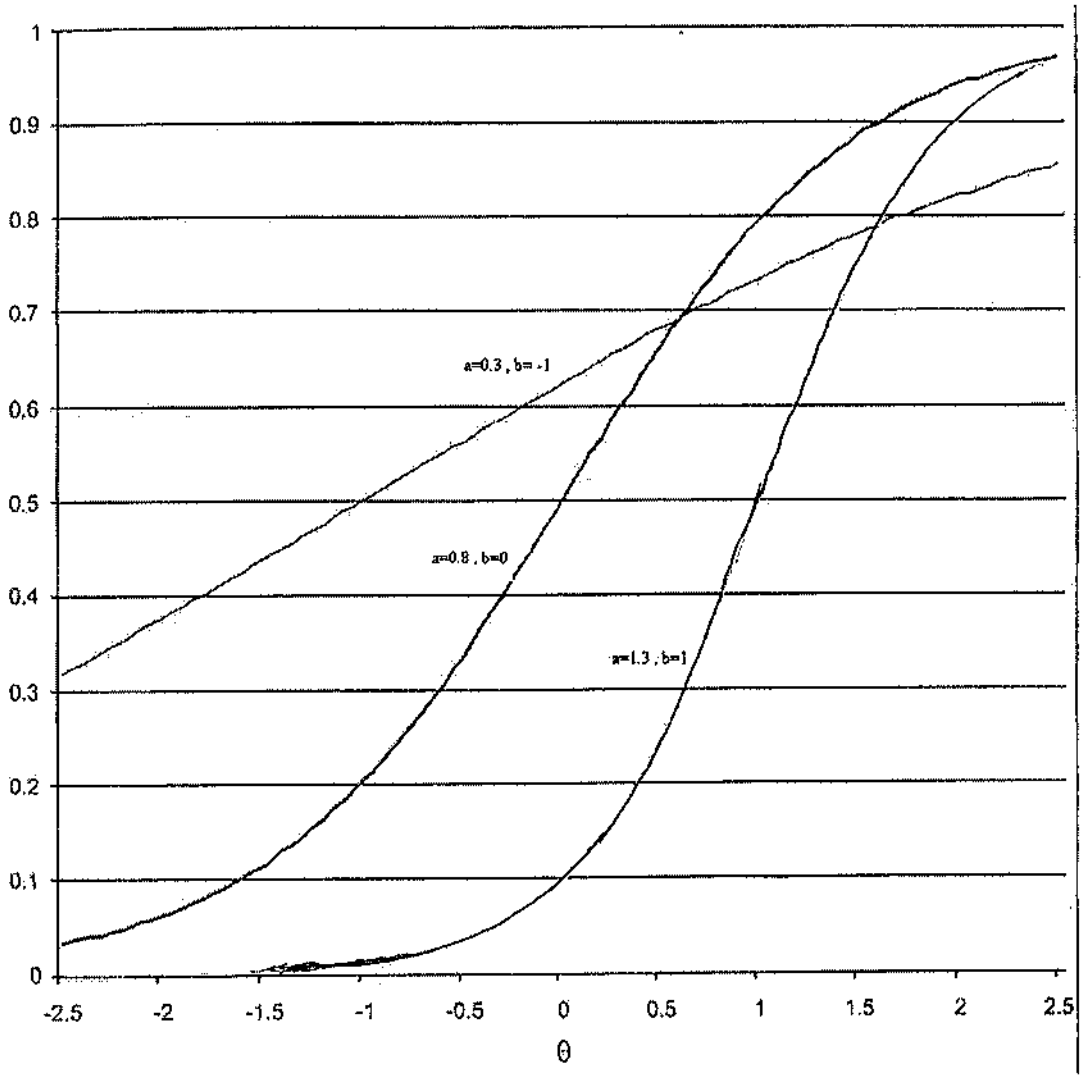
النموذج ثنائي المعلمة (Two-Parameter Logistic Model: 2PL).

ينسب النموذج ثنائي المعلمة إلى لورد (1952). والذي استخدم المنحنى الطبيعي التراكمي (Normal Ogive model)، لكن بيرنبوم (1968) عدل على النسخة الرياضية للنموذج مستخدماً المنحنى اللوجستي (Logistic Ogive) ، وذلك بإضافة معلم التمييز لكل فقرة j (aj) ، ومعامل قياس (D = 1.7 لتقريب المنحنى الطبيعي التراكمي 1 %)، وذلك لحساب احتمال إجابة الفرد (i) على الفقرة j بالمعادلة التالية :

$$P_i(X_{ij}=1|\theta_i a_j, b_j)=\frac{e^{D_{at}(\theta-b_i)}}{1+e^{D_{at}(\theta-b_i)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (23)$$

ويعزى معلم التمييز (a) إلى ميل المنحنى اللوجستي عند نقطة الانعطاف لمنحنى خصائص الفقرة. يتحددالميل بواسطة التباين لتوزيع الخطأ للفقرة، كما يعكس توافق(combination) تفرد الفقرة والخطأ،(Lord,1980) .

والشكل 6 يظهر ثلاث منحنيات لخصائص الفقرة وفق النموذج ثنائي المعلمة متباينة في معالم الصعوبة ومعالم التمييز. يتبين من الشكل 6 أن هذه المنحنيات متقاطعة وهذا يسبب تغيراً في ترتيب الاحتمالية وفقاً للقدرة، بينما صعوبة الفقرات تبقى مستقرة نسبة للقدرات على المقياس. وهذا النموذج يؤيد الفكرة القائلة بأن بعض الفقرات تميز بشكل كبير مع أن لها توزيعات خطأ ضيقة، بينما فقرات أخرى ربما تكون عرضة لمصادر أخرى من التفرد أو الخطأ نسبة للسمة الكامنة مثل (التعقيد، قدرات ثانوية، حلول متعددة) والتي لها تأثيراً على توسيع توزيع الخطأ (Lord,1980).



شكل 6: ثلاث منحنيات لخاصية الفقرة مختلفة في درجة صعوبتها وقوة تمييزها.

مبررات استخدام النموذج ثنائي المعلمة:

1. يسمح بأن تختلف فقرات الاختبار في صعوبتها وتمييزها وهو ما يلاحظ في بناء الاختبارات. إذ من الصعب إيجاد مجموعة من الفقرات تميز بدرجة واحدة بين مستويات السمة أو القدرة التي يقيسها الاختبار، ولذلك غالبا مايتم اعتماد مدى من التمييز يعتبر الفقرات التي تقع ضمن هذا المدى متساوية إحصائيا.

2. يعطي تقديرات أكثر دقة لصعوبة الفقرة حتى عندما تكون تباينات التمييز قليلة، والتخمين اقل ما يمكن، لانه يعمل على تسوية التمييزات المختلفة.(pelton,2002)

3. معالجة منحنيات خصائص الفقرات غير المتوازية في أغلب مقاييس الاختبارات الواسعة النطاق (Large – Scale Tests). لانه يعمل على تسوية استقلالية الأبعـاد المتعددة (independent multidimensionality). (pelton,2002)

4. يعطي نتائج متميزة عندما يكون حجم العينة كبيرا.

5. يستخدم طريقة الترجيح العظمى في تقدير المعالم Maximum Likelihood Estimation

فوائد النموذج ثنائي المعلمة:

يستخدم النموذج اللوغارتيمي ثنائي المعلمة معلم تمييز الفقرة لشرح الفروقات في تباينات الخطأ للفقرة. ويقترح مكدونالد (mcDonald,1981) :

1. إن التباين في الميل عند نقطة الانعطاف يربط نظرية الاستجابة للفقرة بالانحدار الخطي.
2. يمكن أن يعامل النموذج اللوغاريتمي ثنائي المعلمة على انه نموذج عامل عام غير خطي.
3. تمثيل ن من المتغيرات الثنائية على سمة كامنة منفردة بشكل عام، يتطلب ن+1 من الأبعاد dimensions أو العوامل factors ، واحد للسمة الكامنة وواحد لكل متغير ثنائي.

إن التضمين في نظرية الاستجابة للفقرة، للفقرات الثنائية يكون منقسما إلى الجزء الفريد لتباين الفقرة، والذي يكون ثابتا بشكل عام على مدار الوقت، ويمكن أن يختلف من فقرة إلى أخرى. وجزء الخطأ الذي يعزى للعشوائية، والذي يفترض أن يكون متساويا من فقرة لأخرى. وهذا يؤيد الفكرة بأنه لا يوجد هناك سبب نظري يؤكد بأن ميول منحنيات خصائص الفقرات، والتي هي اقترانات تفرد الفقرة يجب أن تكون متساوية (mcDonald,1981).

ويدعي (رايت wright)" أن المنحنيات المتقاطعة جعلت هرمية صعوبة الفقرة النسبي تتغير عند مستوى كل قدرة " (Wright, 1999 P.95) . وكمثال لتوضيح هذا قام رايت بإعادة ترتيب مواضع الفقرات بناء على تدريجات المقياس الفئوي الجديد وفقا لاحتمالاتها النسبية وعلاقتها بمنحنى خصائص فقرة ما، وأشار لهذه المقاييس الجديدة على أنها دليل داعم لادعائه. بينما التباينات في التمييز والتي سببها تفرد الفقرة تسمح للاحتمالات النسبية للفقرات بان تتغير مع القدرة ، وهذا ببساطة منسجم مع الفكرة بان بعض الفقرات اكثر تمييز من الأخرى .ووفق النموذج ثنائي المعالم تبقى الصعوبات ثابتة على مقياس البناء حتى مع تنوع التمييز، وان صعوبات الفقرة المرتبطة البناء والتي تتشارك في القياس مع قدرة الفرد المقدرة هي الغرض الأساسي للقياس.

وعندما تحتوي الفقرات على درجات تمييز مختلفة، أومكونات فريدة والتي تسبب تقاطع لمنحنيات خصائص الفقرة، فان التخمين(intuition) يشير إلى أن تقدير صعوبة الفقرة للفقرات البعيدة عن المركز (اكثر من انحراف معياري عن مركز قدرة المجتمع) سوف تقدر بشكل

ضعيف عند استخدام نموذج راش. وسبب ذلك أن ملائمة الميل الثابت(لمنحنى خصائص الفقرة وفق نموذج راش) للبيانات سوف يكون مسيطرا عليه عند طرف أو أن الطرف الأخر يجبر تقديرات الصعوبة بالابتعاد عن موقعها الصحيح.

محددات (عيوب) النموذج اللوغاريتمي ثنائي المعلمة (2PL):

1.نظرا لاحتواء النموذج اللوغاريتمي ثنائي المعلمة على (معلم التمييز) زيادة على النموذج الاحادي المعلمة، من المتوقع أن يجعله أكثر حساسية للنتائج المشكوك بها(Spurious Results) من النموذج الاحادي المعلمة، والتي تعزى إلى العشوائية في مجموعة البيانات الصغيرة.

2.ربما يعطي هذا النموذج تقديرات أقل (underestimate) لصعوبة الفقرات عندما تحتوي البيانات أثراً لعامل التخمين .

3.تتطلب عملية المعايرة استخدام قيود اصطناعية(artificial constraints) على بعض المعالم(الصعوبة، أو القدرة)، لاجل ضمان التقارب (mislevy&bock,1990) . بالإضافة إلى افتراض أن توزيعات معالم التمييز تزيد معدل التقارب. ولم يعرف كيف ستؤثر هذه القيود والتوزيعات الدقة النهائية للتقديرات.(pelton,2002)

الجزء الثاني : الدراسات السابقة

لقد ركزت بعض الدراسات المحلية، والعربية، والأجنبية، التي تسنى للباحث الرجوع إليها، على خطوات بناء بنوك الأسئلة، والتي تبدأ من إعداد الفقرات، وتجريبها، ومعايرتها، ثم تخزينها ومعالجتها باستخدام برنامج حاسوبي خاص، بينما ركزت دراسات أخرى على المقارنة بين نظرية الاستجابة للفقرة، والنظرية التقليدية، وسيتناول هذا الجزء عرضا للدراسات المحلية، والعربية، والأجنبية التي تناولت بناء بنوك الأسئلة، و أخرى التي تناولت المقارنة بين خصائص الفقرات الناتجة من التحليل وفق النظريتين.

وقد قام الباحث بتصنيف الدراسات إلى:

أولا: الدراسات التي تناولت خطوات بناء بنوك الأسئلة.

ثانيا: الدراسات التي تناولت المقارنة بين خصائص الفقرات التي تم تحليلها وفق النظرية التقليدية، ونماذج نظرية الاستجابة للفقرة.

أولا: الدراسات التي تناولت خطوات بناء بنوك الأسئلة.

1- **الدراسات الأجنبية:** يتضمن هذا البند وصفا موجزا لأربع دراسات حول بناء بنوك الأسئلة:

ففي دراسة اوبرين و هامبيلوس(O, brien & Hampilos, 1988) والتي هدفت إلى بناء بنك أسئلة في مساق مدخل إلى القياس بكلية المعلمين في جامعة كاليفورنيا

معتمدين على نموذج راش ذي المعلم الواحد (صعوبة الفقرة).وقد قامت الدراسة على تطبيق 67 فقرة من إعداد معلم المساق، هي أسئلة امتحان نصف الفصل لطلاب المساق في فصل الخريف عام 1984 ، وتم إعادة تطبيقه في فصل الربيع عام 1985 على مجموعة أخرى من الطلبة.تكونت كل مجموعة من60 طالبا وطالبة، وكانت فقرات الاختبار من نوع الاختيار من متعدد.

بعد تطبيق فقرات الاختبار، استخدم نموذج راش في تحليلها، حيث تبين ثبات مؤشرات صعوبة الفقرات في مستوى مقبول بمرور الوقت، من الاختبار الأول في الخريف، إلى إعادته في الربيع بما يسمح بتخزين الفقرات في بنك الأسئلة، إذ بلغ معامل الارتباط بين صعوبة الفقرات في كلا التطبيقين 0.78 وكان معامل ثبات الاختبار المحسوب بمعادلة كودر-ريتشاردسون 20 من تطبيق الخريف 0.87 بينما كان في الربيع 0.86.

وكذلك تم حساب معامل الصعوبة ومعامل فاي وذلك لفحص العلاقة بين الفقرات المقبولة والمرفوضة على أساس صدق الفقرة وثباتها، وبينت نتائج الدراسة أن الجزء الأكبر من الفقرات يمكن تخزينه بناء على صدقها وثباتها، وتقدير صعوبتها باستخدام نموذج راش، وهذا يقود إلى إمكانية بناء بنك أسئلة من إعداد المعلم وتخزينها معتمدين على مؤشرات صدقها وثباتها، وصعوبتها المحسوبة باستخدام نموذج راش.

واما روبيتيالي و اوشيا(Robitaille & O,shea,1983) فقد اجريا دراسة هدفت إلى بناء ثلاثة بنوك للأسئلة في الرياضيات في كولومبيا البريطانية باستخدام نموذج راش بهدف استخدامها في قياس تحصيل طلبة الصفوف (الثالث،والسابع،والعاشر) واعداد اختبارات تشخيصية لطلبة الصفوف (الرابع،والثامن،والحادي عشر). تم تحليل المحتوى و تحديد نتاجات التعلم لهذه الصفوف ثم حسبت معاملات صعوبة الفقرات وتم تقسيمها إلى ثلاثة مستويات:

- مستوى منخفض (0 – 29)%
- مستوى متوسط (30 – 79)%
- مستوى مرتفع (80 – 100)%

ثم تم تقسيم البناء المعرفي إلى ثلاثة مستويات:

- المعرفة 50%
- الاستيعاب 25%
- التحليل 25%

بعد ذلك تم بناء خمسة نماذج اختباريه لكل بنك من بنوك الأسئلة، بحيث يشتمل كل نموذج على 40 فقرة من نوع الاختيار من متعدد، كما تم تحديد عشر فقرات رابطة بين كل نموذجين.بهدف إجراء عملية التعادل بين الاختبارات وتدريجهما على مقياس عام مشترك.

تم تطبيق الاختبارات على عينة مؤلفة من 1500 طالب وطالبة، ومن ثم أجريت عملية التحليل لاستجابات الطلبة باستخدام النظرية الكلاسيكية حيث استبعدت بعض الفقرات غير المطابقة،ثم تم تحليل الاستجابات مرة أخرى باستخدام نموذج راش بواسطة البرنامج الحاسوبي(Bical).

وقد اعتمد الباحث وحدة تدريج لهذه الفقرات إذ حول وحدة اللوجيت إلى 100 وحدة معتمدا القيمة 1000 كمتوسط لدرجات كل بنك. خلصت هذه الدراسة إلى أن نموذج راش كان أكثر ملاءمة لتدريج الفقرات،ومعادلتها. واستخدام هذا البنك في إنتاج صور اختبارية متكافئة بهدف قياس التحصيل في الرياضيات.

كما أجرى كوننجهام (Cunningham, 1979) دراسة بعنوان الأهداف الرياضية والفقرات الاختبارية للصفوف من الأول ولغاية الثامن (Mathematical Objectives and Test Item for Grades One Through Eight)

حيث هدفت هذه الدراسة إلى بناء بنك أسئلة في الرياضيات باستخدام نموذج راش لتدريج فقرات البنك ومعادلتها. وقد حدد الباحث الموضوعات التي سيغطيها عند كتابة الفقرات على النحو التالي: الأعداد والعمليات عليها، ونظام العد، والكسور العادية والعشرية، والهندسة، وحل المشكلات. ثم قام الباحث بصياغة أهداف المحتوى لهذه الموضوعات حيث بلغت 100 هدف ، وبعد ذلك اعد الفقرات بواقع 5- 8 فقرات اختبارية لكل هدف ، ثم عمل على بناء نماذج اختبارية متنوعة وفقا للمستوى الصفي ، واشتمل كل اختبارين متتاليين من حيث المستوى الصفي على 8 فقرات مشتركة لاستخدامها في عملية التدريج والمعادلة وفق نموذج راش ، حيث استخدم الباحث طريقة المعادلة الراسية (vertical equating) في معادلة الاختبارات .خلصت الدراسة إلى انه يمكن الاستفادة من بنك الأسئلة في بناء اختبارات تشخيصية، واختبارات التمكن، واختبارات الكفاءة.

ومن الدراسات ذات الصلة بالموضوع وجد الباحث دراسة بيللو و وولي و شولا (Bello,Wole,Shola,2000) بعنوان بناء أنموذج بنك أسئلة لمركز الانتساب والقبول في نيجيريا

(The Dvelopment of a Prototype Item Bank for the Nigerian JAMB Test Items), حيث يعنى هذا المركز ببناء وتطبيق اختبارات (placement examinations) للجامعات النيجيرية.

ولبناء هذا البنك فقد تحققت الدراسة من المراحل التالية:

- بناء الفقرات
- تجريب الفقرات
- تدريج الفقرات
- ترميز الفقرات وتخزينها
- جاهزية البنك

وتكون مجتمع الدراسة من 350, 000 طالب وطالبة، أما عينة الدراسة فتألفت من 626 طالب و 374 طالبة، تراوحت أعمارهم بين 15 - 37 عاما. أما أداة الدراسة التي استخدمت فقد تألفت مما يلي:

اختبار اللغة الإنجليزية وهو إجباري ومن نوع الاختيار من أربع بدائل، الوقت المخصص لهذا الاختبار ساعة وربع. والأسئلة موزعة كما يلي:

- الأسئلة من 1 - 15 لكل سؤال 3 علامات.
- الأسئلة من 16 - 30 لكل سؤال علامتان.
- الأسئلة من 31 - 100 لكل سؤال علامة واحدة.

تم استخدام برنامج(Bilog-mg) لتدريج الفقرات ، وكذلك استخدمت طريقة الترجيح العظمى الهامشية (marginal maximum liklehood estimation)، والطريقة البايزية في التقدير (Bayes estimation method) ، حيث تم تقدير معالم الصعوبة ، والتمييز ، والتخمين وفق نظرية استجابة الفقرة . بالإضافة إلى إحصائيات الفقرة وفق النظرية التقليدية.

ولتصميم البنك تم استخدام برنامج حاسوبي(FastTEST) لتخزين الفقرات .

أشارت نتائج هذه الدراسة إلى أن بنوك الأسئلة تساعد باني الاختبار في بناء اختبارات متوازية (Parallel Tests)، بالإضافة إلى سحب اختبارات مكيفة(Adaptive Tests)، أوصت الدراسة انه في حالة القيام ببناء بنك أسئلة كبير لا بد من عمل دراسة تحليل التكلفة (cost analysis) ، من حيث الأشخاص ، والبرمجيات ، والمواد، والتدريب للعاملين في البنك والمستخدمين له.

2- الدراسات المحلية:

يتضمن هذا البند دراستين لكل من فرجات(2004) وحرز الله(2004).

أجرى فرجات (2004) دراسة هدفت إلى بناء نواة لبنك أسئلة في مبحث الكيمياء للصف الثاني ثانوي العلمي وفق نموذج راش والنظرية الكلاسيكية في القياس، وتدريج الفقرات لتكون جاهزة للتخزين في البنك، والسحب منه عند الحاجة. تكونت مجموعة أفراد الدراسة من 2168 طالبا وطالبة من طلبة الصف الثاني الثانوي العلمي موزعين على 12 مديرية تربية وتعليم في الأردن. ولتحقيق هذا الهدف تم تطبيق 120 فقرة من نوع الاختيار من متعدد، حيث وزعت الفقرات على

ثلاثة نماذج اختبارية، اشتمل كل اختبار على 40 فقرة وتم ربطها بفقرات مشتركة عددها ست فقرات.

حللت إجابات 1770 طالبا وطالبة باستخدام برنامج تحليل:SPSS للتحليل وفق النظرية الكلاسيكية، وبرنامج Winstep للتحليل وفق نموذج راش، وذلك للحصول على الإحصائيات الخاصة بالفقرات والأفراد، كما استخدم برنامج قاعدة البيانات Access لتخزين الفقرات ومؤشراتها الإحصائية وفق النظريتين بهدف استدعاء أي من الفقرات وفقا لمؤشراتها الإحصائية، والمواصفات التي بني عليها الاختبار.تراوح معامل الصعوبة للفقرات وفق النظرية التقليدية ما بين 0.3 – 0.8 كذلك من 2.24—3.1 لوجيت وفق نموذج راش.

بينت النتائج أن عدد الفقرات التي تم اختيارها وفق المؤشرات التقليدية كان 109 فقرات، مقابل 88 فقرة وفق نموذج راش. كذلك دلت النتائج على وجود اتفاق عالي بين الأسلوبين في تقدير صعوبة الفقرات.

كما أوضحت الدراسة أن استخدام نموذج راش في بناء اختبار مكون من مجموعة من الفقرات الملائمة يعني أن صعوبات الفقرات تقيس ما تقيسه قدرات الأفراد وتعبر عنه على نفس المقياس، كما أن لها نفس وحدة القياس اللوجيت.

وقد أوصت الدراسة على بناء بنوك أسئلة لمختلف المباحث كذلك إعادة تدريج وتقدير صعوبة الفقرات وفقا لنموذجي ثنائي وثلاثي المعلمة، بالإضافة إلى استخدام الأسئلة المقالية في بناء بنوك الأسئلة.

وفي دراسة حرز الله (2004)، والتي هدفت إلى بناء بنك أسئلة في مبحث الرياضيات، والتحقق من فاعليته في انتقاء فقرات اختبار محكي المرجع في مستوى امتحان الثانوية العامة. ولتحقيق أهداف الدراسة، تم بناء تجمع فقرات من خلال تحليل المحتوى لكتاب الرياضيات للصف الثاني ثانوي العلمي، وبناء جدول مواصفات لتمثيل الأهداف بفقرات. وزعت الفقرات على ست صور اختبارية متكافئة من حيث المحتوى، تكون كل منها من 50 فقرة، منها 10 فقرات مشتركة في كل صورة للفصل الأول، و 10 فقرات أخرى مكررة في كل صورة للفصل الثاني ليصبح مجموع الفقرات الكلي 260 فقرة.

طبقت الاختبارات الستة على عينة مؤلفة من 1500 من طلبة الصف الثاني الثانوي العلمي للعام الدراسي 2001\2002. حيث تم تطبيق الصورة الأولى للفصلين الأول والثاني على 764 مفحوص، والصورة الثانية على 375 مفحوص،والصورة الثالثة على 361مفحوص.

استخدمت الباحثة برمجية المايكروكات (Microcat)، وبرمجية (Bilog) لتحليل البيانات باستخدام النموذج ثلاثي المعلم في نظرية استجابة الفقرة. وكذلك تقدير إحصائيات الفقرة باستخدام النظرية الكلاسيكية.

بينت نتائج التحليل أن 22 فقرة من 260 لم تطابق النموذج من خلال تحليل البواقي، وأن 32 فقرة لم تطابق النموذج من خلال معامل الارتباط الثنائي النقطي. كما بينت نتائج التحقق من تحرر تقدير القدرة من معالم الفقرات أن الفروق في تقدير القدرة من خلال الاختبار الأول(مدى صعوبة فقراته ضمن التوزيع الطبيعي) والاختبار الثاني(صعوبة فقراته أدنى من المتوسط) كانت في 77, 6 % من الحالات لم تتجاوز القيمة الحرجة المتوقعة ضمن احتمالات الصدفة عند مستوى (α=0.05). أما الفروق في تقدير القدرة بين الاختبار الأول والاختبار الثالث(صعوبة فقراته أعلى من المتوسط) كانت في 55.7 % من الحالات لم تتجاوز القيمة الحرجة المتوقعة ضمن احتمالات الصدفة عند مستوى (α=0.05). والفروق في تقدير القدرة من خلال الاختبار الثاني(صعوبة فقراته أدنى من المتوسط) والاختبار الثالث(صعوبة فقراته أعلى من المتوسط) كانت في 62.7 % من الحالات لم تتجاوز القيمة الحرجة المتوقعة ضمن احتمالات الصدفة عند مستوى (α=0.05).

كما أشارت نتائج هذه الدراسة إلى أن بناء بنك أسئلة يعتمد على بناء تجمع من الفقرات، وحساب معالمها، ثم تخزينها في برنامج حاسوبي خاص، يستجيب للمستخدم في التزود بفقرات ذات خصائص معينة تتلاءم مع غرض الاختبار، ومجموعة المفحوصين الذين سيطبق عليهم.

3- الدراسات العربية:

أجرت كاظم(1988) دراسة تجريبية هدفت إلى بناء بنك أسئلة لمقرر علم النفس 101 الذي يدرس في جامعة الكويت مستخدمة نموذج راش في تفسير أداء الطلاب، كما هدفت الدراسة لمعرفة اثر(تقارب- تباعد) مستوى عينتي التحليل مع مستوى كل من اختبارين أحدهما سهل والأخر صعب على دقة عملية التعادل الراسي لتدريجهما، ومن ثم تدريجهما في تدريج واحد مشترك.

وقد افترضت الباحثة انه كلما ازداد التقارب في المستوى بين كل اختبار من الاختبارين والعينة المستخدمة في تحليله، كلما ازدادت الدقة في عملية التعادل الراسي لتدريجي الاختبارين وأمكن دمجهما بدقة في تدريج مشترك واحد له صفر مشترك واحد.

واستخدمت الباحثة في هذه الدراسة اختبارا موضوعيا في علم النفس قامت ببنائه عام(1985) وتكون الاختبار في صورته النهائية من 95 فقرة صالحة للقياس. وتمثل هذه الفقرات ثمانية أبواب من مقرر (101) مدخل إلى علم النفس الذي يدرس بقسم علم النفس-جامعة الكويت.حيث قسم هذا الاختبار إلى اختبارين فرعيين يشتركان في 14 فقرة تستخدم كفقرات رابطة بينهما وذلك على الصورة الآتية:

- **الاختبار السهل**: ويتضمن 40 فقرة سهلة + 14 فقرة رابطة

- **الاختبار الصعب** : ويتضمن 41 فقرة صعبة + 14 فقرة رابطة

وتم تطبيق الاختبارين على عينة حجمها 418 طالبا وطالبة من طلاب مساق علم النفس (101) في جامعة الكويت، وقد كان متوسط قدرة العينة الكلية على الاختبار الكلي0.7 وانحرافها المعياري 0.49 مقدرة بوحدة اللوجت. وباستخدام وسيط الدرجات قسمت العينة الكلية إلى:

- عينة مرتفعة المستوى:وتكونت من 222 طالبا وطالبة درجاتهم من وسيط الدرجات فاكثر.

- عينة منخفضة المستوى: وتكونت من 213 طالبا وطالبة درجاتهم من وسيط الدرجات فاقل.

وبهذا الوضع تشترك العينتان في 17 طالب وطالبة .

وبعد إجراء المعالجات الإحصائية اللازمة تبين أن عدد الفقرات التي لها فروق دالة إحصائيا بين تقديري الصعوبة المشتق من التقدير المرجعي والتدريج المشترك هو (24) فقرة. أي أن حوالي 25% من عدد الفقرات الكلي (95 فقرة) تختلف صعوبتها بصورة جوهرية ، في التدريج المرجعي الناتج من تحليل استجابات العينة الكلية على جميع الفقرات باعتبارها اختبارا واحدا ، عنها في التدريج المشترك الناتج من عملية التعادل الرأسي للاختبارين السهل والصعب في حالة تباعد كل اختبار مع عينة تحليله . وان 75% من هذه الفقرات تتساوى تقديرات صعوبتها إحصائيا في كل من التدريجين. وعند مقارنة نتائج عملية التعادل الرأسي للاختبارين السهل والصعب لوضعهما على تدريج واحد مشترك وصفر واحد مشترك وذلك في كل حالة تقارب كل من الاختبارين لعينتي التحليل وفي حالة تباعدهما تبين انه:

- في حالة التقارب يتفق التدريج المشترك للاختبارين مع التدريج المرجعي في 85% من الفقرات ويختلفان في 15% منها ، بينما في حالة التباعد يتفق التدريجان في 75% من الفقرات ويختلفان في 25% منها .

وبحساب النسبة الحرجة لدلالة تفوق نسبة الاتفاق بين التدريجين 85% في حالة التقارب، عن نسبة الاتفاق بينهما 75% في حالة التباعد، نجد أنها تساوي 1.72 وهي نسبة دالة إحصائيا عند مستوى 0.05 .

ثانيا: الدراسات التي تناولت المقارنة بين النظرية التقليدية ونماذج النظرية الحديثة:

يتضمن هذا البند ست دراسات تناولت المقارنة بين النظرية التقليدية والنظرية الحديثة في القياس.

فقد أجرى عودة (1992) دراسة هدفت إلى الكشف عن مدى التوافق بين نموذج راش ذو المعلم الواحد والمؤشرات الإحصائية التقليدية في اختيار فقرات مقياس اتجاهات المعلمين نحو الامتحانات المدرسية من حيث عدد ونوع الفقرات في صورتي المقياس ومدى تغطية الفقرات لمجال الاتجاه ومعاملات الثبات. طبقت الدراسة على عينة مؤلفة من 458 معلما ومعلمة في

مدارس منطقة اربد ، تكون المقياس من 42 فقرة من نوع العبارات اللفظية سباعية التدريج . تم تحليل استجابات أفراد العينة وفق نموذج راش باستخدام برنامج (Mscale)، بالإضافة إلى ثلاثة مؤشرات إحصائية تقليدية. بلغ عدد الفقرات في التحليل الأولي وفق النموذج (19) فقرة. وبعد أن تم حذف الأفراد غير الملائمين في مرحلة التحليل الثانية ارتفع عدد الفقرات الملائمة إلى (28) فقرة، وقد أشارت النتائج إلى درجة عالية من التوافق في عدد الفقرات ومعامل الثبات للمقياس على عينة الأفراد الذين انسجمت إجاباتهم مع النموذج، إلا أن اختيار الفقرات على أساس معامل الارتباط المعدل لا يعني بالضرورة انسجامها مع النموذج في مقياس يفترض بأنه أحادي البعد.

وفي دراسة فان (Fan,1998) بعنوان المقارنة بين النظرية التقليدية ونظرية الاستجابة للفقرة ، بحث في إمكانية تقدير الفروق بين إحصائيات الفرد والفقرة لكل نموذج.استخدم الباحث بيانات من مركز التقويم في تكساس لاختبارين من نوع الاختيار من متعدد، وبلغ حجم العينة 193,000 مشارك ،اختار الباحث عينة الدراسة كما يلي : 40 عينة عشوائية ، 80 عينة مختلفة بالجنس ، 80 عينة عالية ومتدنية في القدرة ، و اشتملت كل عينة من هذه العينات على 1000 مشارك.

أظهرت نتائج التحليل للبيانات أن درجة الارتباط بين تقديرات القدرة المقدرة باستخدام نماذج نظرية الاستجابة للفقرة، وتقديرات القدرة باستخدام النظرية التقليدية كانت جميعها اكبر من 0.96، كما أن درجة الارتباط كانت عالية لتقديرات الصعوبة باستخدام النظرية التقليدية، والتقديرات المحسوبة باستخدام نموذج راش، حيث كانت جميعها اكبر من 0.998 ، ومتوسطة مع النموذج الثنائي والثلاثي ، حيث كانت $0.83 < r < 0.957$ ، أما بالنسبة لدرجة الارتباط بين معاملات التمييز المحسوبة باستخدام النظرية التقليدية ، ودرجة الارتباط المحسوبة باستخدام النموذجين الثنائي والثلاثي فقد كانت غير قوية وليست ثابتة. كما تفحص الباحث ثبات تقديرات صعوبة الفقرات عبر العينات، فوجد أن درجة الارتباط متشابهة بين تقديرات النظرية التقليدية ونموذج راش، بينما كانت متباينة عند المقارنة مع النموذجين الثنائي والثلاثي. أشار الباحث إلى انه لا يوجد مبرر لتفضيل نظرية الاستجابة للفقرة على النظرية التقليدية.

كما أجرى الحربي (2003) دراسة هدفت إلى التحقق التجريبي من المقارنة والثبات لتقديرات معالم الفقرة ومعلمة الفرد باستخدام النظرية التقليدية ونظرية الاستجابة للفقرة (النموذج أحادى المعلمة، والنموذج ثنائي المعلمة، والنموذج ثلاثي المعلمة)، وذلك من خلال عينات عشوائية مختلفة في أحجامها. استخدم الباحث ثلاث عينات لتقييم المقارنة لتقديرات معالم الفقرة ومعلمة الفرد، وبالمثل لدرجة ثبات تقديرات معالم الفقرة. واستخدم أيضا عينتان لتفحص درجة ثبات تقدير معلمة الفرد.

وتم إنجاز هذه الدراسة على مرحلتين:

- اشتملت مرحلة التحليل الأولى على نتائج تقييم المقارنة والثبات لتقديرات معالم الفقرة والفرد، عندما يتم اختيار العينات عشوائيا من المشاركين دون شروط.
- اشتملت مرحلة التحليل الثانية على نتائج تقييم المقارنة والثبات لتقديرات معالم الفقرة والفرد، عندما يتم اختيار العينات عشوائيا من المشاركين الذين ثبتت ملاءمتهم لنماذج نظرية الاستجابة للفقرة.

استخدم الباحث نوعين من الأدلة لضمان ملاءمة مجموعة البيانات لنماذج نظرية الاستجابة للفقرة:

الأول: استخدم تقنيات مختلفة لتفحص افتراضات نمـاذج نظرية الاستجابة للفقرة مثل (PAFA, MPA, and DIMTEST)

الثاني: تفحص دقة تنبؤات النموذج عند كل من الفرد، الفقرة، مستويات الاختبار.

أشارت نتائج هذه الدراسة إلى أن المقارنة بين تقديرات النظرية التقليدية ونظرية الاستجابة للفقرة لكل من معالم الفقرة ومعلمة الفرد جديرة بالمقارنة. كذلك تقديرات الفقرة والفرد باستخدام النظرية التقليدية، ونظرية الاستجابة للفقرة أعطت نتائج متشابهة فيما يتعلق بالثبات. لأجل ذلك فشلت نتائج هذه الدراسة في إثبات تفوق نظرية الاستجابة للفقرة على النظرية التقليدية فيما يتعلق بثبات تقديرات المعالم. بالإضافة إلى أن حذف الفقرات غير الملائمة، والأفراد غير الملائمين لم يحسن الثبات لتقديرات المعالم وفق نظرية الاستجابة للفقرة.

وأجرت جمحاوي (2000) دراسة هدفت إلى المقارنة بين النظرية التقليدية والنظرية الحديثة في القياس، في مقياس للقدرة الرياضية والمطور وفق المؤشرات التقليدية ليوائم طلاب البيئة الأردنية، وبالتحديد الصفوف الثامن والتاسع والعاشر. حيث قورنت خصائص فقرات هذا المقياس والمكون من 39 فقرة وفق الأسلوبين، وتركزت المقارنة في مدى اتفاق الأسلوبين في انتقاء الفقرات من حيث المحتوى والعدد، وفي مدى اتفاقهما في تقدير صعوبة الفقرات وتمييزها أيضا، وكذلك في مدى اتفاقهما في توزيع القدرة للمفحوصين، وفي ثبات المقاييس الناتجة من كل أسلوب.

تكون مجتمع الدراسة من طلبة الصف التاسع في منطقة اربد، أما عينة الدراسة فقد تكونت من 1061 طالبا وطالبة من طلبة الصف التاسع في مديريتي اربد1، واربد2. وقد طلب من أفراد عينة الدراسة الإجابة عن فقرات المقياس في ظروف صفية عادية، وحللت إجابات الطلبة باستخدام برامج التحليل (Rascal, Ascal) للتحليل وفق نماذج النظرية الحديثة ، كما تم استخدام برامج التحليل (Iteman , SPSS) للتحليل وفق النظرية التقليدية .

أشارت النتائج إلى وجود اتفاق عال بين الأسلوبين في تقدير صعوبة وتمييز الفقرات، وكذلك أظهر الأسلوبان اتفاقا عاليا في توزيع القدرة للمفحوصين، ودلت النتائج أيضا إلى أن عدد الفقرات التي تم اختيارها وفق المؤشرات التقليدية كان 33 فقرة مقابل 20 فقرة انسجمت ونموذج راش، و35 فقرة انسجمت ونموذج ثنائي المعلمتين، و38 فقرة انسجمت ونموذج ثلاثي المعلمة من أصل 39 فقرة، وهذا يشير إلى أن عدد الفقرات المنسجمة يزداد بزيادة عدد معالمه.

كما ويجب ملاحظة أن قيم معاملات التمييز، والصعوبة التقليدية العالية لا تحقق أو لا تضمن تطابق الفقرات المختارة مع أحد نماذج السمات الكامنة، كما أن انخفاض معاملات التمييز والصعوبة وفق النظرية التقليدية لبعض الفقرات لا يعني عدم اتفاقها مع أحد نماذج السمات الكامنة، كما وأن الزيادة في عدد الفقرات المنسجمة مع النموذج لا يوفر نسبة اتفاق عالية مع المؤشرات التقليدية في انتقاء الفقرات.

أما بالنسبة إلى مؤشرات الثبات للمقاييس المطورة بنماذج السمات الكامنة وبالمؤشرات التقليدية، فقد أشارت النتائج إلى عدم وجود فرق ذي دلالة إحصائية بين معاملي ثبات المقياسين الناتجين من المؤشرات التقليدية ونموذج المعلمتين، ووجود فرق ذي دلالة إحصائية بين معامل ثبات المقياس الناتج من المؤشرات التقليدية ونموذج ثلاثي المعالم وكذلك نموذج راش. وقد تبين أن النموذج ثنائي المعالم كان الأكثر اتفاقا والمؤشرات التقليدية من النماذج الأخرى.

كما أجرى الشريفين (2003) دراسة هدفت للكشف عن مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والأخر متعدد التدريج وفق نماذج النظرية الحديثة والنظرية التقليدية في القياس. ولتحقيق هذا الهدف أعد الباحث اختبارين تحصيليين في مبحث الفيزياء، أحدهما ثنائي التدريج تألف من 75 فقرة من نوع الاختيار من أربعة بدائل، والأخر متعدد التدريج تألف من 20 سؤالا.

تكون مجتمع الدراسة من طلبة الصف الثاني عشر العلمي في مدارس مديرية اربد 1، وكان عددهم 2590 طالبا وطالبة. أما عينة الدراسة فتكونت من 1003 طالب وطالبة، حيث تم اختيار المدارس بطريقة عشوائية، أما عينة الدراسة فقد تم اختيارها بالطريقة العشوائية العنقودية.

طبق الاختباران على أفراد عينة الدراسة، وصححت الأوراق وجمعت البيانات، ثم استخدم الباحث البرامج الإحصائية التالية (SPSS ,Bigstep , Ministep) لتحليل البيانات ، والحصول على تقديرات لقدرات الأفراد ، وقيم الصعوبة ، والأخطاء المعيارية ، وإحصائيات المطابقة ، ومتوسطات المربعات ، وقيم معاملات الارتباط الثنائي النقطي ، ومعاملات الثبات للأفراد والفقرات .

أظهرت الدراسة النتائج الآتية:

- تم التحقق من مطابقة الاستجابات عن فقرات الاختبار ثنائي التدريج مع نموذج راش، وتكون الاختبار بصورته النهائية من 50 فقرة.
- تم التحقق من مطابقة الاستجابات عن أسئلة الاختبار متعدد التدريج لنموذج التقدير الجزئي، وتكون الاختبار بصورته النهائية من 20 سؤالا.
- كانت جميع قيم معالم الفقرات للاختبار الثنائي التدريج ضمن حدود المطابقة بالنسبة لمؤشري متوسطات المربعات الداخلية والخارجية وهي (0.7- 1.3) ، وبلغ معامل الثبات للأفراد 0.9 وللاختبار 0.96
- كانت قيم معالم الأسئلة للاختبار متعدد التدريج واقعة جميعها ضمن حدود المطابقة بالنسبة لمؤشري متوسطات المربعات.
- تطابقت إحصائيات الفقرات ثنائية التدريج المحسوبة وفق النظرية التقليدية مع نتائج التحليل باستخدام نموذج راش. أما فيما يتعلق بالاختبار متعدد التدريج فقد تطابقت نتائج التحليل مع نموذج التقدير الجزئي.
- أشارت النتائج إلى أن الاختبار متعدد التدريج يزودنا بمعلومات اكثر من الاختبار ثنائي التدريج وعند كل مستويات القدرة.
- أشارت نتائج المعادلة الأفقية لتصميم المجموعات المتكافئة إلى أن النموذج الأحادي المعلمة كان الأكثر فاعلية من طريقتي المعادلة الخطية والمئينية.

كما أجرت ستيج(Stage,1998) دراسة بعنوان مقارنة تحليل الفقرات وفق النظرية التقليدية والنظرية الحديثة في القياس (Acomparison Between Item Analysis Based on Item Response Theory and Classical Test Theory . A Study of the SweSAT Subtest WORD) .

هدفت الدراسة إلى الإجابة عن الأسئلة التالية:

1- كيف يمكن مقارنة مؤشرات الصعوبة للفقرة المحسوبة وفق النظرية التقليدية مع معالم الصعوبة للفقرة المقدرة وفق نظرية الاستجابة للفقرة.

أ) لبيانات الاختبار القبلي(pretest)

ب) لبيانات الاختبار العادي (natural test)

2- كيف يمكن مقارنة مؤشرات التمييز للفقرة المحسوبة وفق النظرية التقليدية مع معالم التمييز للفقرة المقدرة وفق نظرية الاستجابة للفقرة.

أ) لبيانات الاختبار القبلي

ب) لبيانات الاختبار العادي

3) ما درجة الثبات لمؤشرات الفقرة من بيانات الاختبار القبلي إلى الاختبار العادي وفق النظرية التقليدية.

4) ما درجة الثبات لمعالم الفقرة من بيانات الاختبار القبلي إلى بيانات الاختبار العادي وفق نظرية الاستجابة للفقرة

استخدمت الباحثة 20 فقرة من اختبار المفردات العادي الذي طبق في ربيع 1997، ونفس الفقرات طبقت في ربيع 1996 كاختبار قبلي. حسبت الباحثة المؤشرات الإحصائية(معاملات الصعوبة، ومعامل الارتباط الثنائي النقطي (معامل التمييز)) لكلا الاختبارين وفق النظرية التقليدية. بلغ معامل ارتباط سبيرمان بين قيم معاملا ت الصعوبة على الاختبارين 0.92 ، في حين كان معامل الارتباط بين قيم معاملات التمييز للاختبارين مساويا 0.81 .

كما حسبت معالم الفقرة (الصعوبة ، والتمييز ، والتخمين) لكلا الاختبارين وفق نظرية الاستجابة للفقرة باستخدام برمجية (Bilog-mg) ، كما تم رسم منحنيات خصائص الفقرة لكلا الاختبارين .

وتم حساب معامل الارتباط بين قيم معالم الصعوبة للاختبارين حيث بلغ 0.92 ، في حين بلغ معامل الارتباط بين قيم معاملات التمييز 0.74 ، و أظهرت الرسومات تطابقا عاليا بين منحنيات خصائص الفقرات لجميع فقرات الاختبارين .

خلصت الدراسة إلى النتائج التالية :

- كانت درجة التوافق بين نتائج التحليل للفقرات عالية جدا في كل من النظريتين ، حيث أظهرت نتائج التحليل أن 13 فقرة من 20 فقرة كان لها معاملات الصعوبة والتمييز نفسها على الاختبارين . أما بقية الفقرات فقد كانت الفرو قات بين معاملات الصعوبة أو التمييز قليلة جدا .

- التنبؤ من الاختبار القبلي ببيانات الاختبار العادي مرض، و القسم الأكبر من الفرق في التنبؤ يمكن تفسيره بالتغيرات للفقرات. وهذه النتيجة صحيحة لكلا التحليلين، بغض النظر عن النظرية المستخدمة.

- بلغت درجة الارتباط بين معالم الصعوبة للفقرات المحسوبة وفق النظرية الحديثة ، وقيم الصعوبة للفقرات المحسوبة وفق النظرية التقليدية 0.93 ، للاختبارين ، كما بلغت درجة الارتباط لمعاملات التمييز لفقرات الاختبار القبلي 0.63 و 0.64 للاختبار العادي وفق النظريتين.

وفي ضوء ما سبق من عرض لأهم الدراسات السابقة فقد استخلص الباحث منها مايلي:

1- أن معظم الدراسات السابقة في بناء بنوك الأسئلة اعتمدت على تحليل محتوى الكتب المقررة لبناء الفقرات الاختبارية.

2- استخدمت اغلب الدراسات السابقة نظرية استجابة الفقرة في بناء بنوك الأسئلة وتدريجها وذلك وفق نموذج راش .

3- أكدت معظم الدراسات السابقة على وجود درجة تطابق عالية بين إحصائيات الفقرة المحسوبة وفق النظرية التقليدية ونظرية استجابة الفقرة.

4- عدم وجود دراسات محلية (بحدود علم الباحث) لبناء بنك أسئلة باستخدام نموذج ثنائي المعلمة.

لذا تكمن أهمية هذه الدراسة في بناء بنك أسئلة متحرر من الكتب المدرسية المقررة، والتي هي دائمة التغيير والتعديل، ومحاولة الكشف عن مؤشرات تفوق النظرية الحديثة في القياس على النظرية التقليدية في تطوير المقاييس والاختبارات النفسية والتربوية.

ويمكن الاستفادة من هذه الدراسة في عمل بنوك أسئلة متحررة من الكتب المدرسية لكافة المباحث الدراسية، وعمل مقارنات بين نماذج نظرية الاستجابة للفقرة من جهة، وكذلك المقارنة بين هذه النماذج والنظرية التقليدية.

الفصل الثالث

الطريقة والإجراءات

يتناول هذا الفصل وصفاً لمجتمع الدراسة ، وطريقة اختيار عينة الدراسة ، وبناء أداة الدراسة وإجراءات تطبيقها ، ومراحل بناء بنك الأسئلة ، بالإضافة إلى المعالجات الإحصائية التي أجريت

مجتمع الدراسة:

يتألف مجتمع الدراسة من طلاب وطالبات الصف الأول الثانوي للعام الدراسي 2005/2004 التابعين لمدارس مديرية التربية والتعليم لمنطقة اربد الأولى، ويبلغ عددهم (7399) طالباً وطالبة، منهم (3614) طالباً و(3785) طالبة موزعين على (30) مدرسة للذكور و(38) مدرسة للإناث، وذلك من خلال الرجوع إلى التقرير الإحصائي السنوي الذي يعده قسم التخطيط في المديرية للعام الدراسي 2005/2004. ويبين الجدول 5 توزيع أفراد مجتمع الدراسة وفق متغير الجنس والمسار التعليمي.

جدول 5: توزيع أفراد مجتمع الدراسة وفق متغير الجنس والمسار التعليمي.

الجنس / المسار التعليمي	الذكور	الإناث	المجموع
العلمي	1320	1100	2420
الأدبي	1199	1682	2881
الإدارة المعلوماتية	497	400	897
المهني	598	603	1201
المجموع	3614	3785	7399

عينة الدراسة:

بعد قيام الباحث بإجراء مسح لمدارس مديرية التربية والتعليم لمنطقة اربد الأولى، وذلك من خلال الرجوع إلى التقرير الإحصائي السنوي الذي يعده قسم التخطيط في المديرية للعام الدراسي 2004/ 2005، تم تحديد المدارس التي يوجد فيها الصف الأول الثانوي (العلمي، الأدبي، الإدارة المعلوماتية، المهني) وكان عدد هذه المدارس (68) مدرسة منها (30) مدرسة للذكور و(38) مدرسة للإناث كما هو موضح في الجدولين 6 و 7.

جدول 6: مدارس الذكور وأعداد الطلبة فيها .

الرقم	المدرسة	علمي	إدارة معلوماتية	ادبي	مهني
.1	ثانوية اربد للبنين	349	*	*	*
.2	ثانوية خالد بن الوليد	97	*	80	*
.3	ثانوية شفيق ارشيدات	44	*	59	*
.4	ثانوية الأمير حسن	175	*	161	*
.5	ثانوية علي خلقي الشراري	82	143	63	29
.6	ثانوية الوليد بن عبد الملك	134	*	125	*
.7	ثانوية حسن كامل الصباح	*	*	118	*
.8	ثانوية سعد بن أبي وقاص	*	*	133	*
.9	ثانوية حوارة للبنين	49	*	42	*
.10	ثانوية بشرى للبنين	73	*	25	*
.11	ثانوية وصفي التل	*	*	*	180
.12	ثانوية حطين	76	36	59	*
.13	ثانوية التطوير الحضري للبنين	6	*	18	*
.14	ثانوية المغير للبنين	*	*	27	*
.15	ثانوية بيت راس للبنين	21	*	25	*
.16	ثانوية الزرنوجي للبنين	79	51	46	*
.17	ثانوية بيت يافا للبنين	28	*	*	*
.18	ثانوية فوعرا للبنين	13	*	17	*
.19	اليرموك النموذجية للبنيين	73	*	*	*
.20	دار العلوم	21	*	*	*
.21	ثانوية ججين للبنين	*	*	13	*
.22	ثانوية زحر للبنين	*	*	12	*
.23	ثانوية سوم للبنين	*	*	18	*
.24	ثانوية دوقرا للبنين	*	*	27	*
.25	ثانوية الخراج للبنين	*	*	10	*
.26	ثانوية حكما للبنين	*	*	24	*
.27	ثانوية علعال للبنين	*	*	15	*
.28	ثانوية كفر جايز للبنين	*	*	11	*
.29	ثانوية حور للبنين	*	*	15	*
.30	ثانوية سال للبنين	*	*	24	*

جدول 7: مدارس الإناث وأعداد الطالبات فيها .

الرقم	المدرسة	علمي	ادبي	إدارة معلوماتية	مهني
.1	ثانوية اربد للبنات	165	*	*	*
.2	صفية بنت عبد المطلب	87	150	*	49
.3	جمانه الشاملة بنات	*	323	123	25
.4	ثانوية عين جالوت الشاملة	94	91	*	25
.5	نور الحسين الثانوية الشاملة	39	29	108	22
.6	الشاملة للبنات	108	155	62	180
.7	ثانوية بنات طبريا الشاملة	191	267	*	*
.8	عائشة بنت أبي بكر الشاملة	*	*	*	140
.9	حواره الثانوية للبنات	33	56	*	*
.10	بشرى الثانوية للبنات	25	45	*	*
.11	سال الثانوية للبنات	22	27	32	16
.12	ثانوية بنات المغير	12	36	*	*
.13	ثانوية بنات بيت راس	38	67	*	62
.14	ثانوية بنات كفر يوبا	39	54	36	31
.15	ثانوية بنات بيت يافا	24	15	*	21
.16	ثانوية بنات زحر	22	73	*	*
.17	ثانوية بنات سوم	11	24	*	*
.18	ثانوية بنات دوقرا	9	26	*	*
.19	ثانوية بنات قميم	33	36	20	42
.20	ثانوية بنات كفر أسد	26	32	20	*
.21	ثانوية بنات اليرموك النموذجية	47	*	*	*
.22	ثانوية بنات النموذجية	11	*	*	*
.23	ثانوية راهبات الوردية	8	*	*	*
.24	ثانوية الروم الارثوذوكس	16	*	*	*
.25	ثانوية بنات النهضة	4	*	*	*
.26	ثانوية بنات دار العلوم	7	*	*	*
.27	ثانوية بنات نهاوند	*	21	*	*
.28	ثانوية بنات الأندلس	*	37	*	*
.29	ثانوية بنات حكما	*	43	*	*
.30	ثانوية بنات علعال	*	43	*	*
.31	ثانوية بنات التطوير الحضري	*	17	*	16
.32	ثانوية بنات كفر جايز	*	18	*	*
.33	ثانوية بنات فوعرا	*	36	*	*
.34	بنات حور	*	15	*	*
.35	بنات ناطفه	*	6	*	*
.36	بنات كفر عان	*	11	*	*
.37	بنات حوفا	*	14	*	*
.38	بنات الخراج	*	10	*	*

وبعد ذلك تم حصر أعداد الطلبة في كل مدرسة، كما تم تحديد حجم العينة المطلوب والذي يساوي (10) أضعاف عدد الفقرات الاختبارية مضافاً إليها (10%) من حجم العينة كما هو في الجدول 8

جدول 8: عدد الطلاب المطلوب للعينة.

المسار التعليمي	عدد الطلاب	عدد الطالبات	العدد المطلوب للعينة
العلمي	1320	1100	850
الأدبي	1199	1682	1100
إدارة معلوماتية	497	400	300
مهني	598	603	429
المجموع	3614	3785	2679

ثم بعد ذلك تم اختيار عينة من المدارس بالطريقة العشوائية البسيطة كما يلي :

1. تم استبعاد كافة المدارس التي يقل عدد الطلاب فيها عن (20) .
2. تمت كتابة أسماء المدارس المتبقية على أوراق صغيرة ووضعت في صندوق، و تم السحب عشوائيا من الصندوق وتسجيل اسم المدرسة وعدد الطلاب فيها إلى أن اكتمل العدد المطلوب، ويبين الجدولان 9 و 10 أسماء المدارس والمسار التعليمي وأعداد الطلبة الذين اختيروا للعينة من الجنسين.

جدول 9: توزيع أفراد عينة الدراسة الإناث وأسماء المدارس والمسار التعليمي

الرقم	اسم المدرسة	العلمي	الأدبي	إدارة معلوماتية	مهني	المجموع
1.	ثانوية بنات كفر أسد	26	32	20	42	120
2.	ثانوية بنات قميم	33	36	20	*	89
3.	عائشة أم المؤمنين	*	*	*	140	140
4.	صفية بنت عبد المطلب	87	150	*	49	286
5.	ثانوية بنات اربد	165	*	*	*	165
6.	جمانه الثانوية للبنات	*	232	123	25	380
7.	ثانوية بنات طبريا	158	231	*	*	389
	المجموع	469	681	163	256	1569

*تشير الى عدم وجود المسار التعليمي في المدرسة

جدول 10: توزيع أفراد عينة الدراسة الذكور وأسماء المدارس والمسار التعليمي

الرقم	اسم المدرسة	العلمي	الأدبي	إدارة معلوماتية	مهني	المجموع
.1	حطين الثانوية للذكور	*	59	36	*	95
.2	ثانوية الزرنوجي الشاملة	75	43	50	80	248
.3	خالد بن الوليد	*	140	*	*	140
.4	وصفي التل	*	*	*	110	110
.5	ثانوية الأمير حسن	100	100	*	*	200
.6	علي خلقي الشراري الشاملة	82	*	70	29	181
.7	الوليد بن عبد الملك	50	50	*	*	100
.8	ثانـــوية ججـــين للبنين	*	36	*	*	36
	المجموع	307	428	156	219	1110

*تشير إلى عدم وجود المسار التعليمي في المدرسة

وقد روعي في اختيار عينة الدراسة أن تكون من مستويات تعليمية مختلفة وذلك للحصول على تباين في استجابات الطلبة . ومما يجدر ذكره هنا أن غرض الدراسة لم يكن قياس تحصيل الطلبة وإنما هو الحصول على استجابات يمكن من خلالها بناء بنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية .

نماذج الاختبار:

هدفت هذه الدراسة إلى بناء بنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية وفق نظرية الاستجابة للفقرة(النموذج ثنائي المعلمة) والنظرية الكلاسيكية في القياس ولتحقيق هذا الهدف تم تطبيق اختبار مكون من (237) فقرة من نوع الاختيار من متعدد ولكل فقرة أربعة بدائل واحد منها فقط صحيح، وتم تقسيم الفقرات على ستة نماذج اختبارية، والجدول 11 يظهر عدد الفقرات الاختبارية في كل نموذج.

جدول 11: النماذج الاختبارية وعدد الفقرات في كل منها.

النموذج	1	2	3	4	5	6
عدد الفقرات	49	46	46	47	45	49

ولكي نتمكن من وضع جميع الفقرات على تدريج واحد مشترك تم ربط النماذج الستة من خلال فقرات مشتركة وعددها (9) فقرات من ضمن الفقرات الكلية لكل نموذج.

إجراءات الدراسة:

تضمنت عملية بناء الفقرات الاختبارية ووضعها في صورتها النهائية الاسترشاد بالأسس العلمية المتبعة في بناء الاختبارات ، وفي ما يلي الإجراءات التي تم إتباعها :

أولا : تحديد الغرض من الاختبار وهو بناء بنك أسئلة للمهارات الرياضية في نهاية المرحلة الأساسية .

ثانياً :تم تحديد المهارات الرياضية المراد قياسها في الاختبار وهي :

مهارة التفكير الرياضي، مهارة التبرير الرياضي، مهارة النمذجة، مهارة حل المشكلة، مهارة الرموز.

ثالثاً : تم تحديد مستويات المعرفة للعملية المعرفية كما يلي :

1. القدرات العقلية الدنيا والتي تشمل (التذكر ، الاستيعاب ، التطبيق) .
2. القدرات العقلية العليا والتي تشمل (التحليل ، التقويم ، الإبداع) .

رابعاً : تم بناء جدول مواصفات جرى فيه ربط مستويات المعرفة بالمهارات الرياضية المراد قياسها في الاختبار والجدول 12 يبين ذلك .

جدول 12: جدول مواصفات الاختبار

المهارة	الوزن	مجالات التقويم: تذكر 10%	استيعاب 20%	تطبيق 20%	تحليل 15%	تقويم 15%	ابداع 20%	
التفكير الرياضي	%40	10	20	20	15	15	20	
حل المشكلات	%30	6	15	15	12	12	15	
التبرير الرياضي	%10	2	5	5	4	4	5	
النمذجة	%10	2	5	5	4	4	5	
الترميز	%10	2	5	5	4	4	4	
المجموع	100%	22	50	50	39	39	49	249

خامساً: بناء الفقرات. قام الباحث بتشكيل لجنة من المختصين وأصحاب الخبرة والكفاية، حيث تألفت هذه اللجنة من (6) مشرفين تربويين يحملون درجة البكالوريوس في الرياضيات ودرجة الماجستير في التربية. وبعد ذلك عقدت ورشات عمل تم خلالها مناقشة المهارات الرياضية ومستويات المعرفة وصياغة الفقرات الاختبارية حيث تمت صياغة فقرات البنك في صورة الاختيار من متعدد، وقد حرصت اللجنة على مراعاة شروط صياغة فقرة جيدة ذات بدائل ملائمة. كما راعت اللجنة تحديد المستوى المعرفي ومدى تمثيل الفقرات للمهارات المقابلة لها والصحة العلمية لهذه الفقرات . وقد بلغ عدد فقرات البنك في صورته الأولية(249) فقرة.

سادساً: التجريب الأولي للاختبار. بعد أن تمت صياغة الفقرات من قبل اللجنة المشكلة، تمت طباعة هذه الفقرات بصورتها الأولية، ومن ثم تم تشكيل (6) نماذج اختبارية بطريقة عشوائية حيث تألف كل اختبار من (40) فقرة بالإضافة إلى (9) فقرات مشتركة بين الاختبارات الستة، ومن ثم طبق الباحث هذه الاختبارات على عينه مؤلفة من (120) طالب من طلبة مدرسة كفر أسد الأساسية للذكور (الصف العاشر الأساسي)، وذلك بالتنسيق مع معلمي الرياضيات، إذ كان هؤلاء الطلبة متفاوتين في المقدرة التحصيلية في الرياضيات وتتوفر لديهم الرغبة والجدية في الأداء، وتم إعلامهم بأن الغرض من هذه الاختبارات القيام بدراسة علمية. و قد روعي أن لا يكون أفراد العينة الاستطلاعية من ضمن أفراد الدراسة الرئيسة. وكان الغرض من التجريب هو تحديد الزمن الذي يستغرقه الطلبة في الإجابة على النماذج الاختبارية وجمع أي ملحوظات حول الفقرات. وبالتالي

الكشف عن الفقرات التي تحتاج إلى تعديل أو حذف في ضوء ذلك . هذا و قد تم إبلاغ معلمي الرياضيات في مدرسة كفر أسد الأساسية مسبقاً عن موعد تطبيق الاختبار، وعن طبيعة المادة التي سيطبق عليها الاختبار. وبعد دراسة إجابات الطلبة على النماذج الاختبارية. والملاحظات التي تم تسجيلها من قبل المشرفين على الاختبار تم حذف (12) فقرة، ليصبح عدد الفقرات بصورتها النهائية (237) فقرة.

سابعاً: الصورة النهائية للاختبار: تألف الاختبار في صورته النهائية من (237) فقرة اختبارية مرتبطة بالمهارات الرياضية للمرحلة الأساسية. وقد تم توزيع هذه الفقرات على ستة نماذج من (1 –6) بحيث توزعت الفقرات على هذه النماذج (49 ،46 ، 46، 47 ،45 ،49) على التوالي بحيث يختبر الطالب في نموذج واحد فقط . ولغرض وضع الفقرات جميعها على تدريج واحد مشترك لتكوين بنك الأسئلة تم ربط هذه النماذج الستة بـ(9) فقرات مشتركة. وأشارت دراسة راجو وإدوارد وأوسبرج (Raju, Edwards & Osberg, 1983) إلى أن العدد القليل من الفقرات المشتركة 6 أو 8 فقرات تنجز على الأغلب نفس ما ينجزه العدد الأكبر من الفقرات الرابطة، وفي دراسة ونقرسكاي ولورد (Wingersky & Lord 1984) وجدا أن فقرتين جيدتين تعملان عمل (25) فقرة. وفي دراسة كلاين و كولن (Klein & Kolen,1985) وجدا أن طول اختبار المرساة (Anchor Test) له تأثير قليل على نوعية المعادلة في حالة المجموعات المتشابهة القدرة . أما بارنا رد (Barnard, 1996) فقد أشار إلى أن الفقرات المشتركة يجب أن لا يقل عددها عن (20%) من طول الاختبار. وتوزعت الفقرات الرابطة بين النماذج الستة بشكل عشوائي . والجدول 13 (يمثل أرقامها في كل نموذج).

جدول 13: أرقام الفقرات الرابطة في النماذج الستة

النموذج/1	النموذج/2	النموذج/3	النموذج/4	النموذج/5	النموذج/6
9	26	29	10	30	48
15	46	35	16	45	12
6	39	26	7	27	45
7	40	27	8	28	46
10	41	30	11	40	7
11	42	31	12	41	8
12	43	32	13	42	9
13	44	33	14	43	10
14	45	34	15	44	11

ثامناً: تعليمات تطبيق الاختبار: تم إرفاق تعليمات مع نموذج الاختبار توضح للطالب كيفية الإجابة على الاختبار. وقد تضمن تعريفا بالاختبار، وحث الطلبة على بذل أقصى جهد ممكن وتوخي الدقة عند الإجابة، ووضع إشارة (×) على رمز البديل الذي يمثل الإجابة الصحيحة وذلك على نموذج تفريع الإجابة مع إعطاء مثال على ذلك. وكذلك تم تزويد الطلبة بقائمة من القوانين يمكن الرجوع إليها.

تاسعاً : مرحلة التحضير لتطبيق الاختبار : تم التحضير لتطبيق الاختبار من خلال التنسيق المسبق مع مشرفي الرياضيات في مديرية التربية والتعليم لمنطقة أربد الأولى للإشراف على تطبيق الاختبار ومتابعته . وكذلك تم الاجتماع مع معلمي المبحث الذين سيطبقون الاختبار لتوضيح تعليمات الاختبار وآلية تطبيقه .

وقد تم تحديد يومي الأربعاء 8 / 9 / 2004 لتطبيق الاختبار في مدارس الذكور والخميس 2004/9/9 لتطبيق الاختبار في مدارس الإناث.

عاشراً: مرحلة تطبيق الاختبار: طبقت الدراسة في بداية الفصل الأول للعام الدراسي 2005/2004 وقد تضمنت عملية التطبيق قيام المعلمين بتطبيق الاختبار تحت إشراف ومتابعة المشرف المختص وذلك بعد قراءة التعليمات لضمان أكبر قدر ممكن من الجدية في الإجابة على الاختبار. وقد روعي ما يلي:

1. توزيع الاختبارات على الطلبة بطريقة الزق زاق، فالطالب الذي يجيب على الاختبار الأول يجيب الطالب المجاور له على الاختبار الثاني، في حين يجيب الذي يجلس أمامه أو خلفه على الاختبار الثالث.... وهكذا، ولما كان عدد الفقرات في الاختبارات الفرعية متقارباً فصارت المدة الزمنية اللازمة للإجابة عن هذه الاختبارات أيضاً متقاربة، وبتنوع اختبارات الطلبة وقيام المشرفين بمتابعة التطبيق بأنفسهم داخل الصفوف تم ضبط الموقف الاختباري والتحكم فيه.
2. الإجابة على النموذج الخاص بالإجابة.
3. عدم البدء في الإجابة لحين إشعارهم بالبدء.
4. الطلب من الطلبة بذل قصارى جهدهم وتوخي الدقة في الإجابة.

حادي عشر: مرحلة جمع استجابات الطلبة. بعد أن قام الباحث بجمع إجابات الطلبة على الاختبارات الستة والبالغ عددهم (2679) طالباً وطالبه تم مراجعة أوراق الاختبارات جميعها للكشف عن إجابات الطلبة، وقد تم استبعاد الأوراق التي لم يجب عنها، وكذلك الأوراق التي لم تحمل الجدية في الإجابة وغيرها من الأشكال المختلفة للإجابة(أوراق أجاب عنها الطالب باختيار البديل (ج) لكل الفقرات، وأوراق اشتملت على كتابات ساخرة ليس لها علاقة بالاختبار...). ويبين

الجدول 14 توزيع الطلبة الذين تم تحليل إجاباتهم فعلاً وفق صور الاختبار المستخدمة باستخدام البرنامج الإحصائي (Bilog- MG).

جدول 14: عدد أفراد الدراسة موزعين وفق صور الاختبار.

الاختبار	عدد الأفراد	النسبة المئوية
الأول	400	%16.40
الثاني	404	%16.50
الثالث	422	%17.30
الرابع	412	%16.90
الخامس	403	%16.45
السادس	402	%16.45
المجموع	2443	%100

ثاني عشر: مرحلة تخزين الإجابات على البرنامج الحاسوبي

(Statistical Package For Social Science – SPSS).

1. تعبئة البيانات : تم تعبئة البيانات بإعطاء العمود الأول للرقم المتسلسل للطالب (Identification Number) يليها عمود لرقم المدرسة والعمود الثالث لجنس الطالب يليها أعمدة تمثل إجابات الفقرات (حسب عدد الفقرات في كل اختبار) .
2. طريقة تفريغ البيانات: يوضع للطالب الرقم (1) إن اختار البديل (أ)،و يوضع للطالب الرقم (2) إن اختار البديل (ب) و يوضع للطالب الرقم (3) إن اختار البديل (جـ) و يوضع للطالب الرقم (4) إن اختار البديل (د) .
3. مفتاح التصحيح: في بداية البيانات (السطر الأول) تم تخزين مفتاح التصحيح لكل اختبار من اليسار إلى اليمين لفقرات الاختبار المختلفة. كما تم إعطاء الرقم (5) ليدل على الفقرات التي تركها الطالب دون إجابة أي (Missing Values)، وعند إجراء التحليل اعتبرت أنها إجابة خاطئة. لتحليل البيانات استناداً إلى النظرية الكلاسيكية. فقد تم إنشاء ملف أوامر في البرنامج الإحصائي (bilog- mg) وهو (Phase 1) لتحويل الرموز (1، 2، 3، 4، 5) إلى (0، 1) بحيث يأخذ الطالب العلامة (1) على الإجابة الصحيحة و(0) على الإجابة الخاطئة، وبالتالي فإن كل نموذج يصحح ثنائيا.

تم إيجاد معامل صعوبة الفقرة (p) وهو نسبة الاستجابات الصحيحة أو نسبة الأفراد الذين أجابوا إجابة صحيحة على الفقرة. وتم الاحتفاظ بالفقرات ذات معاملات صعوبة تراوحت ما بين

(0.1– 0.9) وهذا ما أشار إليه عودة (2002)، كذلك تم إيجاد معامل التميز (d) باستخدام العلاقة الارتباطية بين علامة الفقرة والعلامة الكلية وهو ما يشار إليه بمعامل ارتباط (بوينت بايسيريال)، و تم الاحتفاظ بالفقرات ذات القدرة التمييزية المقبولة التي أخذت القيم أكبر من 0.1 وذلك حسب ما أشار إليه كل من ألين و ين (Allin & Yen,1979) وكروكر و ألجينا (Crocker & Algina,1986) وبهذا الإجراء بلغ عدد الفقرات المحذوفة (39 فقرة) من أصل (237) فقرة .

وقد تضمنت مرحلة التحليل أيضاً تحليل البيانات وفق نظرية استجابة الفقرة (النموذج ثنائي المعلمة). وقد عوملت الفقرة المتروكة (Missing Value) على أنها خطأ، ثم جهز كل نموذج بإنشاء ملف أوامر في البرنامج الإحصائي (bilog-mg) تكون من(phase2&phase3)، حيث تم حساب معالم الفقرات (الصعوبة، والتمييز)، و معلمة الأفراد (القدرة) مع الأخطاء المعيارية لتقدير كل معلم، كما قدم البرنامج بالإضافة للمعالم، دالة معلومات الاختبار ،مع رسم بياني لها ،بالإضافة إلى الرسم البياني لمنحنى خصائص الاختبار ، ورسم بياني لمنحنيات خصائص الفقرات . وقد بلغ عدد الفقرات المحذوفة (44 فقرة) من اصل (237).

ولتدريج الفقرات جميعها على تدريج واحد مشترك، بوحدة قياس معرفة (اللوجيت)، فقد تم استخدام الفقرات المشتركة بين الاختبارات، حيث تم بناء معادلات خطية لتحويل الصعوبة إلى تدريج مشترك، وذلك بالاستفادة من معالم الفقرات المشتركة. وقد تم تثبيت تدريج النموذج الأول وتحويل تدريج النماذج الأخرى إلى تدريج النموذج الأول ، وبذلك تم تدريج الفقرات جميعها .

الفصل الرابع

النتائج ومناقشتها

يتناول هذا الفصل نتائج التحليلات المختلفة ذات الصلة بإجراءات بناء بنوك الأسئلة بشكل عام وللمهارات الرياضية في نهاية المرحلة الأساسية وفق النظرية الكلاسيكية والنظرية الحديثة في القياس بشكل خاص والتي تتضمن توفير معلومات عن خصائص كل فقرة من الفقرات التي يحتفظ بها البنك أهمها صعوبة الفقرة وتمييزها المتحررين من خصائص العينة والتي ستوفر بالتالي اختبارات تقدر قدرة المفحوص متحررة من الفقرات الداخلة فيها. ولتحقيق ذلك تم تطبيق ستة نماذج اختبارية ثنائية التدريج، وبعد أن جمعت البيانات من هذا التطبيق أدخلت في ذاكرة الحاسوب حيث استخدمت البرمجيات SPSS،(bilog-mg) لمعالجة وتحليل البيانات واستخراج النتائج وفيما يلي عرض لهذه النتائج:

أولا:

الإجابة عن السؤال الأول: ما خصائص الأسئلة في بنوك الأسئلة؟

لقد تمت الإجابة عن هذا السؤال وفقا للإجراءات التالية:

1) إيجاد الخصائص السيكومترية لفقرات البنك وفق النظرية الكلاسيكية باستخدام برنامج التحليل (bilog-mg). (phase 1)
2) إيجاد الخصائص السيكومترية لفقرات البنك وفق نظرية استجابة الفقرة باستخدام برنامج التحليل (bilog-mg). (phase 2&3)

لمعرفة درجة صعوبة الفقرات الاختبارية ثنائية التدريج ودرجة تمييزها، تم إدخال إجابات الطلبة في ذاكرة الحاسوب باستخدام البرمجية (SPSS)، ثم أخضعت البيانات للتحليل الإحصائي باستخدام البرنامج(bilog-mg)المصمم لتحليل البيانات وفقا للنظريتين السابقتين، حيث تشتمل مخرجات هذا البرنامج على المكونات التالية:

- **المرحلة الأولى** phase one وتحتوي (معامل ارتباط بيرسون، و معامل الارتباط الثنائي النقطي، وصعوبة الفقرة).
- **المرحلة الثانية** phase two وتحتوي (صعوبة الفقرة، وتمييز الفقرة (ميل الفقرة)، والتخمين، والخطأ المعياري للصعوبة، والخطأ المعياري للتمييز).
- **المرحلة الثالثة** phase three وتحتوي (علامة المفحوص، وقدرة المفحوص، والخطأ المعياري للقدرة).

وفيما يلي عرض للنتائج التي تم التوصل إليها وفقا للإجراءات السابقة:

ا) النتائج التي أسفر عنها الإجراء الأول وهي الخصائص السيكومترية للفقرات الاختبارية وفق النظرية الكلاسيكية باستخدام برنامج التحليل (bilog-mg). حيث حسب البرنامج معامل الصعوبة للفقرة على انه نسبة الأفراد الذين أجابوا عن الفقرة بشكل صحيح إلى مجموع الأفراد الذين حاولوا الإجابة عن هذه الفقرة. كما حسب البرنامج معامل التمييز للفقرة على انه معامل الارتباط بين علامة الفقرة والعلامة الكلية لكل فرد (معامل ارتباط بوينت بايسيريال)، و يتبين من الجدول 15 متوسطات الصعوبة ومتوسطات التمييز للفقرات الاختبارية في الصور الاختبارية الست الناتجة من التحليل وفق النظرية الكلاسيكية باستخدام برنامج التحليل (bilog-mg).

جدول 15: متوسطات الصعوبة ومتوسطات التمييز للفقرات المحسوبة وفق النظرية الكلاسيكية

الصورة الاختبارية	متوسط صعوبة الفقرات	متوسط التمييز للفقرات
الأولى	0.39	0.32
الثانية	0.37	0.27
الثالثة	0.43	0.32
الرابعة	0.47	0.36
الخامسة	0.36	0.24
السادسة	0.46	0.40

يتضح من الجدول 15 أن متوسطات صعوبة الفقرات للصور الاختبارية الست تراوحت بين 0.36 - 0.47. أما متوسطات تمييز الفقرات فقد تراوحت بين0.24 – 0.4 .كما يتبين من الجدولين 16 و17 التوزيعات التكرارية لصعوبة الفقرات الاختبارية ودرجة التمييز وفق النظرية الكلاسيكية.

جدول 16:التوزيع التكراري لصعوبة الفقرات وفق النظرية الكلاسيكية للصور الاختبارية الست

الفئات	التكرار للصورة الأولى	التكرار للصورة الثانية	التكرار للصورة الثالثة	التكرار للصورة الرابعة	التكرار للصورة الخامسة	التكرار للصورة السادسة
0.19 – 0.0	5	3	1	1	2	0
0.39– 0.2	21	25	20	19	28	23
0.59– 0.4	14	14	19	13	14	18
0.79– 0.6	9	4	6	11	1	7
0.99– 0.8	0	0	0	3	0	1
المجموع	49	46	46	47	45	49

وقد أظهرت النتائج أن قيم معاملات الصعوبة للصورة الاختبارية الأولى تراوحت بين 0.16 – 0.75 ، بينما تراوحت قيم معاملات الصعوبة للصورة الثانية بين 0.14- 0.75 ، وتراوحت قيم الصعوبة للصورة الثالثة بين 0.20 – 0.73 ، وللصورة الرابعة تراوحت بين 0.20 –0.86 وللصورة الخامسة تراوحت بين 0.13 – 0.66 وللصورة السادسة تراوحت بين 0.23 – 0.80 ، وهذا يشير إلى أن هناك فقرات متطرفة في الصعوبة.

جدول 17:التوزيع التكراري لتمييز الفقرات وفق النظرية الكلاسيكية للصور الاختبارية الستة

الفئات	التكرار للصورة الأولى	التكرار للصورة الثانية	التكرار للصورة الثالثة	التكرار للصورة الرابعة	التكرار للصورة الخامسة	التكرار للصورة السادسة
0.11- - 0.3-	0	2	1	2	1	1
0.09 – 0.10-	6	3	6	0	9	0
0.29 – 0.1	15	17	12	14	19	9
0.49 – 0.3	16	24	21	19	13	26
0.69 – 0.50	12	0	6	12	3	13
المجموع	49	46	46	47	45	49

وتراوحت قيم معاملات التمييز للصورة الأولى بين 0.05- – 0.63 ،كما تراوحت قيم التمييز للصورة الثانية بين0.17- – 0.48 ، وتراوحت قيم التمييز للصورة الثالثة بين0.11- 0.60 ،وللصورة الرابعة تراوحت قيم التمييز بين 0.29- – 0.62 ، وللصورة الخامسة تراوحت قيم التمييز بين 0.24- – 0.52 ، وللصورة السادسة تراوحت قيم التمييز بين0.12- – 0.67 .يتضح من الجدول 17 أن مدى معاملات التمييز تراوح بين 0.29- – 0.67 ، وهذا يشير إلى وجود فقرات عديمة التمييز.

وإذا أردنا الاحتفاظ بالفقرات الاختبارية في بنوك الأسئلة، فان أدبيات القياس والتقويم في ضوء تحليل البيانات وفق النظرية التقليدية توصي بان يتم الاحتفاظ بالفقرات التي تتمتع بخصائص جيدة، وأهم هذه الخصائص صدقها التمييزي، ومستوى صعوبتها، كما و أشارت هذه الأدبيات إلى أنه يمكن الاحتفاظ بالفقرات التي قيمة متوسط صعوبتها يساوي 0.8، بينما معامل تمييزها يساوي صفر، وذلك عندما يكون الهدف من الاختبار قياس التمكن من مهارة معينة .وكذلك الفقرات التي تنحصر قيم معامل صعوبتها بين 30. – 0.7 ، ومعامل تمييزها أكبر من أو يساوي 0.3، بينما في حالة الاختبارات محكية المرجع يوصى بان يحتفظ بالفقرات التي تنحصر قيم معاملات صعوبتها بين 0.1 – 0.9 ، ومعامل التمييز اكبر من 0.1، وقد يتسع هذا المدى لأكثر من ذلك في حالة الاختبارات المفصلة(عودة، 2002، علام، 2002).

2) النتائج التي أسفر عنها الإجراء الثاني وهي الخصائص السيكومترية للفقرات الاختبارية وفق نظرية الاستجابة للفقرة باستخدام برنامج (bilog-mg):

تم تحليل الفقرات الاختبارية وفق نظرية الاستجابة للفقرة، نموذج ثنائي المعلمة(two-parameter logistic model) مستخدما برمجية (bilog-mg).

يبين الجدول 18 ملخصا لمتوسطات معاملات الصعوبة والتمييز للصور الاختبارية الست وفق نظرية الاستجابة للفقرة.

جدول 18: ملخص متوسطات الصعوبة والتمييز للفقرات المحسوبة وفق نظرية الاستجابة للفقرة

الصورة الاختبارية	متوسط الصعوبة	متوسط التمييز
الاولى	1.84	0.78
الثانية	1.39	0.67
الثالثة	0.87	0.75
الرابعة	0.48	0.89
الخامسة	1.61	0.64
السادسة	0.52	0.90

يتضح من الجدول 18 أن مدى متوسطات الصعوبة لفقرات الصور الاختبارية الست تتراوح بين 0.48 – 1.80 ، في حين تراوح مدى قيم متوسطات معاملات التمييز بين (0.6 – 0.9). يتبين من الجدولين 19، 20 التوزيعات التكرارية لصعوبة الفقرات الاختبارية، ودرجة التمييز لهذه الفقرات وفق نظرية استجابة الفقرة.

جدول19:التوزيع التكراري لصعوبة الفقرات وفق نظرية استجابة الفقرة للصور الاختبارية الستة

فئات الصعوبة	التكرار للصورة الأولى	التكرار للصورة الثانية	التكرار للصورة الثالثة	التكرار للصورة الرابعة	التكرار للصورة الخامسة	التكرار للصورة السادسة
-3 -2	0	0	1	0	0	0
-2 -1	2	3	1	6	0	3
-1 0	14	6	13	16	9	19
0 -1	9	13	14	8	8	13
1 -0	8	11	7	8	11	9
2 -1	7	3	4	3	5	4
3 -2	4	5	4	3	7	0
4 -3	3	2	2	1	2	0
5 -4	2	2	0	0	2	1
6 -5	49	45	46	45	44	49
المجموع						

يتضح من الجدول 19 أن مدى معاملات الصعوبة للفقرات تراوح بين -2.06 – 5.31 ، وهذا يشير إلى وجود فقرات متطرفة في الصعوبة.

جدول 20:التوزيع التكراري لتمييز الفقرات وفق نظرية استجابة الفقرة للصور الاختبارية الستة

فئات التمييز	التكرار للصورة الأولى	التكرار للصورة الثانية	التكرار للصورة الثالثة	التكرار للصورة الرابعة	التكرار للصورة الخامسة	التكرار للصورة السادسة
0.2– 0.0	0	0	0	0	0	1
0.4 – 0.2	10	12	11	4	12	3
0.6 – 0.4	12	8	6	10	16	10
0.8 – 0.6	7	9	10	5	3	10
1.0 – 0.8	4	13	5	11	5	7
1.2 – 1.0	6	1	10	4	4	9
1.4 – 1.2	7	2	2	6	2	4
1.6 – 1.4	1	0	1	4	2	1
1.8 – 1.6	2	0	1	0	0	2
2.0 – 1.8	0	0	0	1	0	1
2.2 – 2.0	0	0	0	0	0	0
2.4 – 2.2	0	0	0	0	0	1
المجموع	49	45	46	45	44	49

يتضح من الجدول 20 أن مدى معاملات التمييز للفقرات الاختبارية تراوح بين (0.20 – 2.35) وهذا يشير إلى أن هناك فقرات ذات تمييز عال.

وإذا ما أريد الاحتفاظ بالفقرات الاختبارية في بنك الأسئلة بناء على تحليل البيانات وفق نظرية الاستجابة للفقرة، فقد أشارت أدبيات القياس والتقويم إلى انه يمكن الاحتفاظ بالفقرات التي تنحصر معاملات صعوبتها بين $^{-}2 - ^{+}2$ أو بين $^{-}2.5 - ^{+}2.5$ أو بين $^{-}3 - ^{+}3$ ، ومعامل التمييز اكبر من 0.2، ويرتبط ذلك كله بالغرض الذي انشأ البنك من أجله، وكذلك الهدف من الاختبارات التي ستسحب من البنك. ولإجراء المقارنة بين الفقرات التي يتم حذفها وفق النظرية التقليدية ، والفقرات التي يتم حذفها وفق نظرية الاستجابة للفقرة . فقد تم اعتماد المحكات التي أوصت بها أدبيات القياس والتقويم، والجدول 21 يبين ذلك.

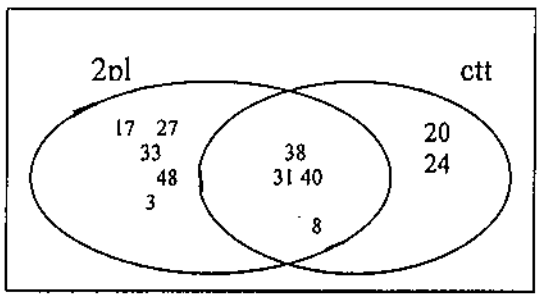
الجدول (21):أرقام الفقرات المحذوفة باستخدام نظرية الاستجابة للفقرة ،وكذلك باستخدام النظرية التقليدية للصور الاختبارية الست

الصورة الاختبارية	النظرية التقليدية			النظرية الحديثة		
	مدى الصعوبة	التمييز	أرقام الفقرات المحذوفة	مدى الصعوبة	التمييز	أرقام الفقرات المحذوفة
الأولى	0.9 – 0.1	rbis> .1	8، 20، 24، 31، 38، 40	3^{+} - 3^{-}	a>.2	48، 40، 38، 33، 31، 27، 17، 8، 3
الثانية			29، 25، 22، 12، 8			38، 35، 29، 28، 25، 22، 12، 11، 8، 2
الثالثة			46، 45، 44، 28، 16، 7، 3			45، 44، 28، 19، 16، 7
الرابعة			9، 6			43، 42، 41، 21، 9، 6
الخامسة			38، 29، 23، 17، 16، 14، 8، 6، 3، 2			26، 23، 21، 17، 16، 14، 8، 3، 38، 3، 29
السادسة			47			47
	0.8 – 0.2	rbis> .1		2.5^{+} – 2.5^{-}	a>.2	
الأولى			48، 40، 38، 33، 31، 24، 20، 8			31، 27، 25، 24، 20، 17، 8، 3، 48، 47، 42، 40، 38، 33
الثانية			38، 35، 29، 25، 22، 12، 8، 6			38، 35، 29، 28، 25، 22، 12، 11، 8، 2
الثالثة			46، 45، 44، 28، 16، 7، 6، 3			19،28،44،45،7،16،18
الرابعة			21، 20، 9، 6، 4			43، 42، 41، 40، 21، 9، 6
الخامسة			38، 29، 26، 23، 17، 16، 14، 8، 6، 3، 2			38، 30، 29، 26، 23، 21، 17، 16، 15، 14، 11، 8، 6، 3
السادسة			47			47، 36، 15
	0.7 – 0.3	rbis> .1		2^{+} -2^{-}	a>.2	
الأولى			27، 25، 24، 20، 19، 17، 9، 8، 5، 3، 41، 40، 39، 38، 35, 33، 32, 31 ، 12، 1، 48، 47، 45, 42			48، 47، 42، 40، 38، 35، 33، 31، 27، 25، 24، 20، 17، 9، 8، 3
الثانية			23، 22، 19، 15، 12، 11، 8، 6، 2، 37،36، 35 ، 34، 29، 28، 26، 25، 21، 38			38، 35، 29، 28، 34، 25، 22، 19، 12، 11، 8، 6، 2
الثالثة			46، 45، 44، 29، 28، 19، 16، 13، 7، 6، 4، 3			46، 45، 44، 29، 28، 19، 18، 16، 13، 7، 3
الرابعة			45، 43، 42، 41، 40، 32، 30، 29، 26، 21، 20، 13، 9، 7، 6، 5، 4			45، 43، 42، 41، 40، 30، 21، 9، 6
الخامسة			38، 37، 36، 30، 29، 26، 23، 22، 21، 17، 16، 15، 14، 11، 8، 6، 3، 2			38، 37، 31، 30، 29، 26، 23، 22، 21، 17، 16، 15، 14، 11، 8، 6، 3
السادسة			48، 47، 38، 31، 26، 19، 15، 6، 3			47، 38، 37، 36، 15

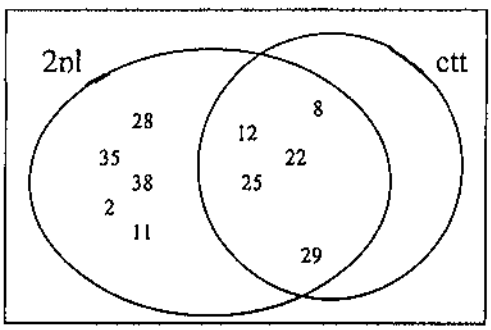
يتضح من الجدول 21 أن 81% من الفقرات المحذوفة عند استخدام النظرية التقليدية بمدى الصعوبة 0.1 – 0.90 والتمييز أكبر من 0.10 تكون متوافقة مع الفقرات المحذوفة عند استخدام نظرية الاستجابة للفقرة بمدى الصعوبة ($^{-}3$ – $^{+}3$) والتمييز أكبر من 0.2، و قد بلغ عدد الفقرات المحذوفة التي اختلف الأسلوبان في الحكم عليها 24 فقرة، حيث تم حذف 6 فقرات بناء على المؤشرات التقليدية، و 18 فقرة حذفت بناء على نظرية الاستجابة للفقرة، وبذلك تكون نسبة الفقرات التي اتفق الأسلوبان عليها 90%. بينما تصل النسبة إلى 83% عند مدى الصعوبة 0.20 – 0.80 والتمييز أكبر من 0.10 وفق النظرية التقليدية مقابل($^{-}2.5$ - $^{+}2.5$) ، والتمييز اكبر من 0.20 وفق نظرية الاستجابة للفقرة، و قد بلغ عدد الفقرات المحذوفة التي اختلف الأسلوبان في الحكم عليها 28 فقرة، حيث تم حذف 7 فقرات بناء على المؤشرات التقليدية، و 21 فقرة حذفت بناء على نظرية الاستجابة للفقرة، وبذلك تكون نسبة الفقرات التي اتفق الأسلوبان عليها 88%. بينما تقل هذه النسبة إلى 68.7% عند مدى الصعوبة 0.30 – 0.70 والتمييز اكبر من 0.10، وفق النظرية التقليدية، ومدى الصعوبة $^{-}2$ - $^{+}2$، والتمييز اكبر من 0.20 وفق نظرية الاستجابة للفقرة، و قد بلغ عدد الفقرات المحذوفة التي اختلف الأسلوبان في الحكم عليها 36 فقرة، حيث تم حذف 32 فقرة بناء على المؤشرات التقليدية، و 4 فقرات حذفت بناء على نظرية الاستجابة للفقرة، وبذلك تكون نسبة الفقرات التي اتفق الأسلوبان عليها 85%.

ويمكن توضيح نسبة التوافق بين النظرية التقليدية والنموذج ثنائي المعلمة بأشكال فن كما يلي:

1- عند اعتماد مدى الصعوبة (0.10 – 0.90) والتمييز اكبر من 0.10 وفق النظرية التقليدية، وعند استخدام نظرية الاستجابة للفقرة(نموذج ثنائي المعلمة) بمدى الصعوبة ($^{-}3$ - $^{+}3$) والتمييز اكبر من 0.20 للصور الاختبارية الستة والأشكال12,11,10,9,8,7 تبين ذلك.



شكل 7: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الأول



شكل 8: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الثاني

من الشكل 7 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الأول) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{38 + 4}{49} = 0.86$$

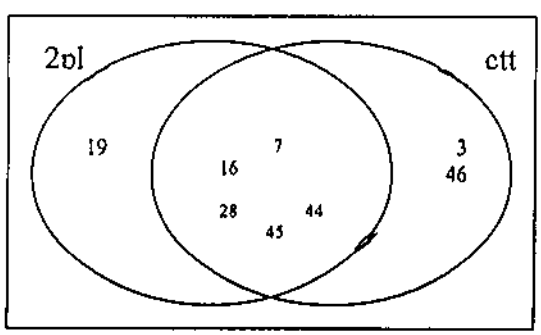
أي أن 86% من الفقرات (42 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 38 فقرة، وكذلك اتفقا على استبعاد (4 فقرات)، أما النسبة 14% من فقرات الاختبار(7 فقرات) فقد اختلفا في الحكم عليها، فقرتان استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وخمس فقرات استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

يتبين من الشكل 8 أن:

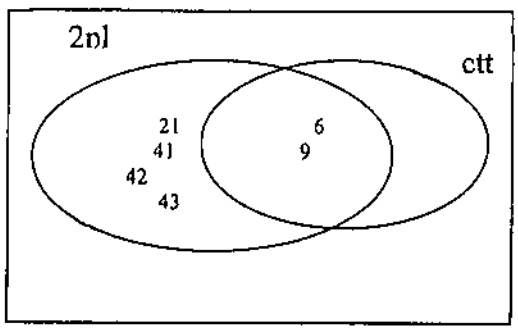
نسبة الاتفاق بين الأسلوبين (للاختبار الثاني) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{36 + 5}{46} = 0.89$$

أما بالنسبة للاختبار الثاني نجد أن 89% من الفقرات (41 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 36 فقرة، وكذلك اتفقا على استبعاد (5 فقرات)، أما النسبة 11% من فقرات الاختبار(5 فقرات) فقد اختلفا في الحكم عليها، خمس فقرات استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.



شكل 9: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الثالث



شكل 10: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الرابع

من الشكل 9 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الثالث) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{38 + 5}{46} = 0.94$$

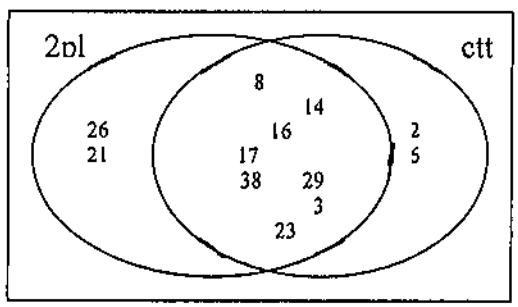
يتبين أن 94% من الفقرات (43 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 38 فقرة ،وكذلك اتفقا على استبعاد (5 فقرات)،أما النسبة 6% من فقرات الاختبار(3 فقرات) فقد اختلفا في الحكم عليها،فقرتان استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وفقرة واحدة استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

من الشكل 10 يتبين أن:

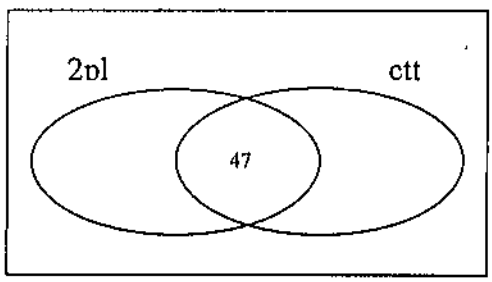
نسبة الاتفاق بين الأسلوبين (للاختبار الرابع) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{41 + 2}{47} = 0.92$$

أي أن 92% من الفقرات (43 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 41 فقرة ،وكذلك اتفقا على استبعاد (2 فقرة)،أما النسبة 8% من فقرات الاختبار(4 فقرات) فقد اختلفا في الحكم عليها،أربع فقرات استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.



شكل 11: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الخامس



شكل 12: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار السادس

من الشكل 11 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الخامس) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{33 + 8}{45} = 0.91$$

يتبين أن 91% من الفقرات (41 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 33 فقرة ،وكذلك اتفقا على استبعاد (8 فقرات)،أما النسبة 9% من فقرات الاختبار(4 فقرات) فقد اختلفا في الحكم عليها،فقرتان استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وفقرتان استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

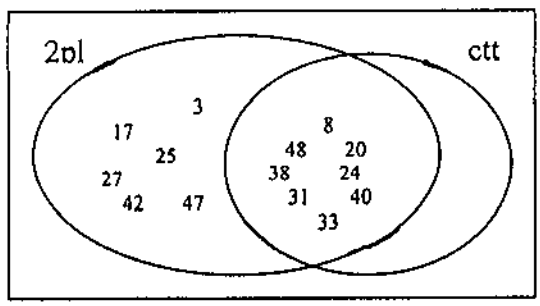
من الشكل 12 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار السادس) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

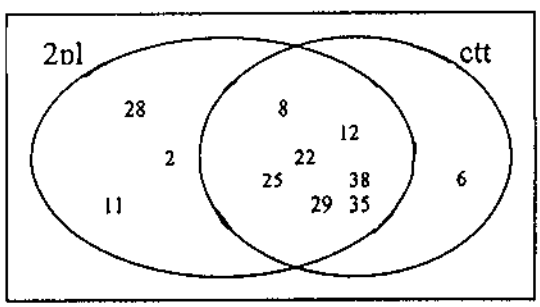
$$= \frac{48 + 1}{49} = 1$$

وهذا يعني أن 100% من الفقرات اتفق الأسلوبان في الحكم عليها، حيث اتفقا على استبعاد فقرة واحدة، وابقاء 48 فقرة.

2- عند اعتماد مدى الصعوبة (0.2 – 0.8) والتمييز أكبر من 0.10 وفق النظرية التقليدية. وعند استخدام نظرية الاستجابة للفقرة بمدى الصعوبة (-2.5 – $+2.5$) والتمييز أكبر من 0.20 للصور الاختبارية الستة والأشكال 13,14,15,16,17,18 تبين ذلك.



شكل 13: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الأول



شكل 14: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الثاني

من الشكل 13 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الأول) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{35 + 8}{49} = 0.88$$

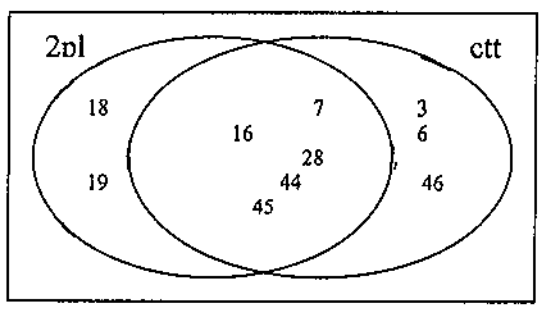
يتبين أن 88% من الفقرات (43 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 35 فقرة ،وكذلك اتفقا على استبعاد (8 فقرات)،أما النسبة 12% من فقرات الاختبار(6 فقرات) فقد اختلفا في الحكم عليها،ست فقرات استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

من الشكل 14 يتبين أن:

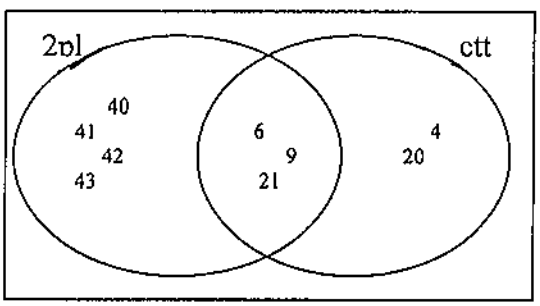
نسبة الاتفاق بين الأسلوبين (للاختبار الثاني) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{35 + 7}{46} = 0.91$$

يتبين أن 91% من الفقرات (42 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 35 فقرة ،وكذلك اتفقا على استبعاد (7 فقرات)،أما النسبة 9% من فقرات الاختبار(4 فقرات) فقد اختلفا في الحكم عليها،فقرة استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وثلاث فقرات استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.



شكل 15: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الثالث



شكل 16: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الرابع

من الشكل 15 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الثالث) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{36 + 5}{46} = 0.89$$

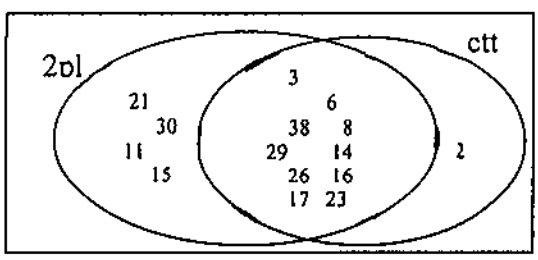
يتبين أن 89% من الفقرات (41 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 36 فقرة ،وكذلك اتفقا على استبعاد (5 فقرات)،أما النسبة 11% من فقرات الاختبار(5 فقرات) فقد اختلفا في الحكم عليها،ثلاث فقرات استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وفقرتان استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

من الشكل 16 يتبين أن:

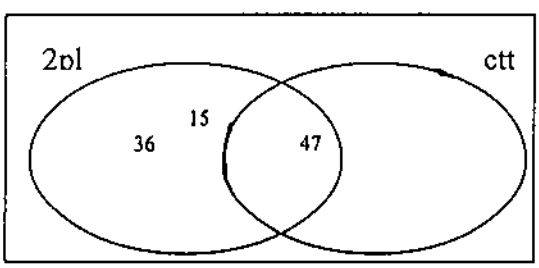
نسبة الاتفاق بين الأسلوبين (للاختبار الرابع) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{38 + 3}{47} = 0.87$$

يتبين أن 87% من الفقرات (41 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 38 فقرة، وكذلك اتفقا على استبعاد (3 فقرات)، أما النسبة 13% من فقرات الاختبار(6 فقرات) فقد اختلفا في الحكم عليها، فقرتان استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة أربع فقرات استبعدها النموذج ولم تستبعدها المؤشرات التقليدية



شكل 17: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الخامس



شكل 18: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار السادس

من الشكل 17 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الخامس) = (عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين) / العدد الكلي للفقرات

$$= \frac{30 + 10}{45} = 0.89$$

يتبين أن 89% من الفقرات (40 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 30 فقرة ،وكذلك اتفقا على استبعاد (10 فقرات)،أما النسبة 11% من فقرات الاختبار(5 فقرات) فقد اختلفا في الحكم عليها،فقرة واحدة استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وأربع فقرات استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

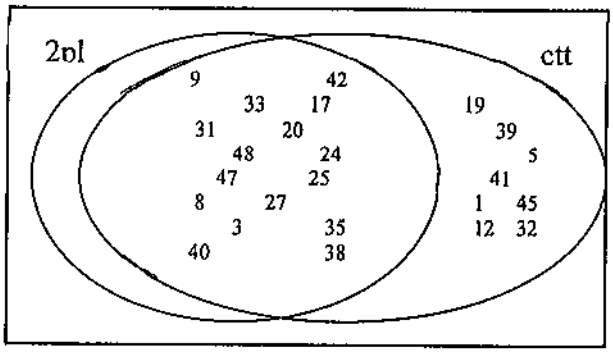
من الشكل 18 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار السادس) = (عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين) / العدد الكلي للفقرات

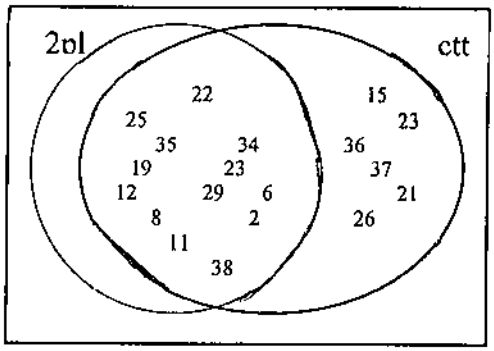
$$= \frac{46 + 1}{49} = 0.96$$

يتبين أن 96% من الفقرات (47 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 46 فقرة، وكذلك اتفقا على استبعاد (فقرة واحدة)، أما النسبة 4% من فقرات الاختبار(فقرتان) فقد اختلفا في الحكم عليها، فقرتان استبعدها النموذج ولم تستبعدها المؤشرات التقليدية

3- عند اعتماد مدى الصعوبة (0.30– 0.70) والتمييز أكبر من 0.10 وفق النظرية التقليدية. وعند استخدام نظرية الاستجابة للفقرة بمدى الصعوبة (2^- - 2^+) والتمييز أكبر من 0.20 للصور الاختبارية الستة والأشكال 19,20,21,22,23,24 تبين ذلك.



شكل 19: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الأول



شكل :20 نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الثاني

من الشكل 19 يتبين أن:

$$\text{نسبة الاتفاق بين الأسلوبين (للاختبار الاول)} = \frac{\text{عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين}}{\text{العدد الكلي للفقرات}}$$

$$= \frac{25 + 16}{49} = 0.84$$

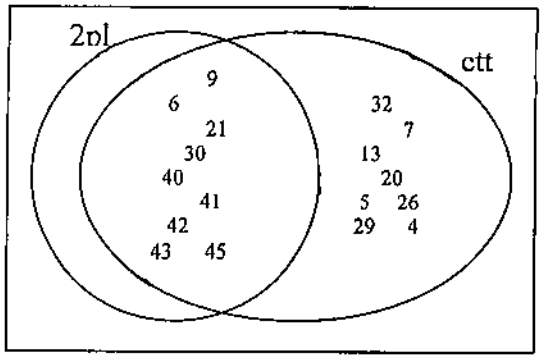
يتبين أن 84% من الفقرات (41 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 25 فقرة ،وكذلك اتفقا على استبعاد (16 فقرة)،أما النسبة 16% من فقرات الاختبار(8 فقرات) فقد اختلفا في الحكم عليها،ثماني فقرات استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة.

من الشكل 20 يتبين أن:

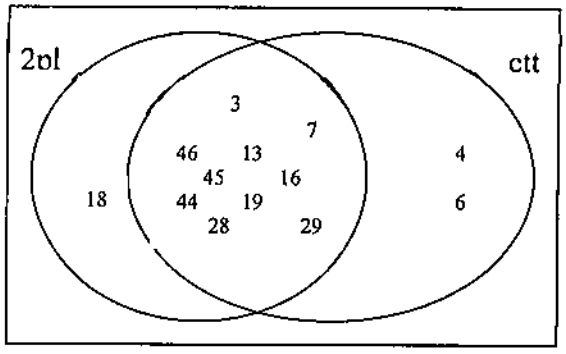
$$\text{نسبة الاتفاق بين الأسلوبين (للاختبار الثاني)} = \frac{\text{عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين}}{\text{العدد الكلي للفقرات}}$$

$$= \frac{27 + 13}{46} = 0.87$$

يتبين أن 87% من الفقرات (40 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 27 فقرة ،وكذلك اتفقا على استبعاد (13 فقرة)،أما النسبة 13% من فقرات الاختبار(6 فقرات) فقد اختلفا في الحكم عليها،ست فقرات استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة.



شكل 21: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الثالث



شكل 22: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الرابع

من الشكل 21 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الثالث) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{33 + 10}{46} = 0.94$$

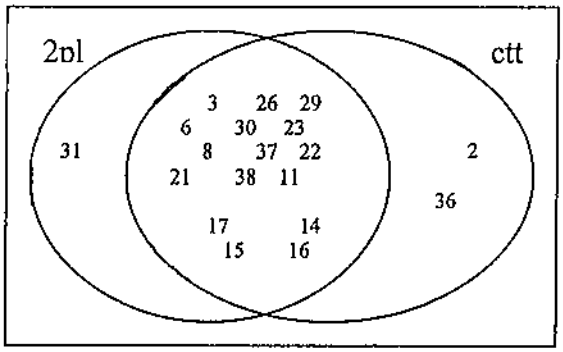
يتبين أن 94% من الفقرات (43 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 33 فقرة، وكذلك اتفقا على استبعاد (10 فقرات)، أما النسبة 6% من فقرات الاختبار(3 فقرات) فقد اختلفا في الحكم عليها، فقرتان استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وفقرة واحدة استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

من الشكل 22 يتبين أن:

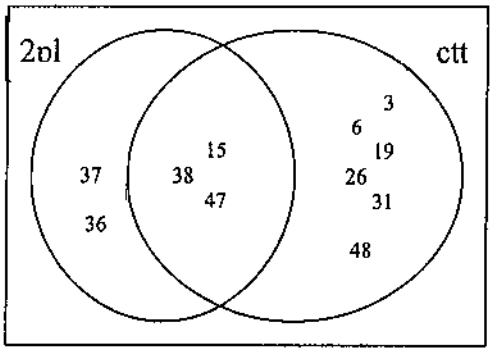
نسبة الاتفاق بين الأسلوبين (للاختبار الرابع) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{30 + 9}{47} = 0.83$$

يتبين أن 83% من الفقرات (39 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 30 فقرة ،وكذلك اتفقا على استبعاد (9 فقرات)،أما النسبة 17% من فقرات الاختبار(8 فقرات) فقد اختلفا في الحكم عليها، ثماني فقرات استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة.



شكل 23: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار الخامس



شكل 24: نسبة الاتفاق بين النظرية التقليدية والنموذج ثنائي المعلمة للاختبار السادس

من الشكل 23 يتبين أن:

نسبة الاتفاق بين الأسلوبين (للاختبار الخامس) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{26 + 16}{45} = 0.93$$

يتبين أن 93% من الفقرات (42 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 26 فقرة ،وكذلك اتفقا على استبعاد (16 فقرة)،أما النسبة 7% من فقرات الاختبار(3 فقرات) فقد اختلفا في الحكم عليها،فقرتان استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وفقرة استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

من الشكل 24 يتبين أن:

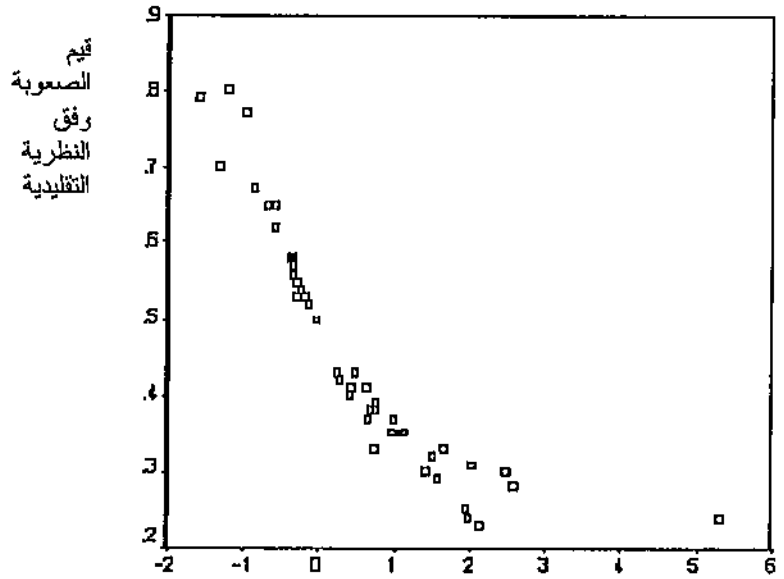
نسبة الاتفاق بين الأسلوبين (للاختبار السادس) = عدد الفقرات المتفق على استبعادها وفق الأسلوبين + عدد الفقرات المتفق على بقائها وفق الأسلوبين / العدد الكلي للفقرات

$$= \frac{38 + 3}{49} = 0.84$$

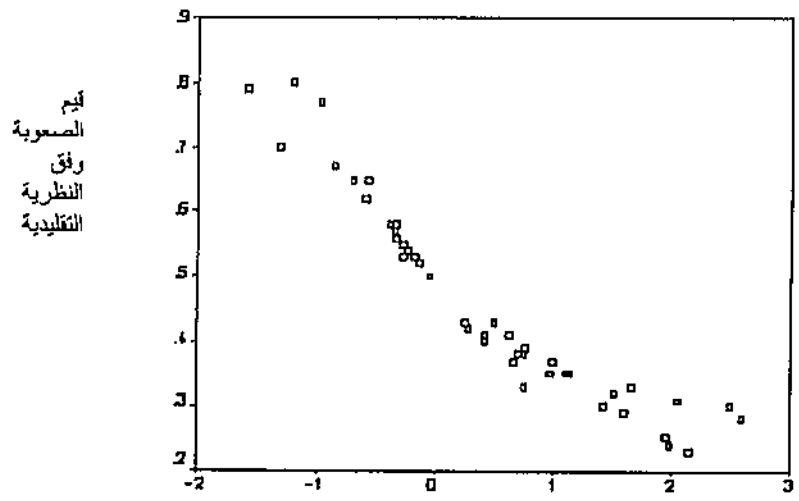
يتبين أن 84% من الفقرات (41 فقرة) اتفق الأسلوبان في الحكم عليها، حيث اتفقا على إبقاء 38 فقرة ،وكذلك اتفقا على استبعاد (3 فقرات)،أما النسبة 16% من فقرات الاختبار(8 فقرات) فقد اختلفا في الحكم عليها،ست فقرات استبعدتها المؤشرات التقليدية ولم يستبعدها النموذج ثنائي المعلمة وفقرتان استبعدها النموذج ولم تستبعدها المؤشرات التقليدية.

وعند حساب معامل ارتباط بيرسون بين معاملات صعوبة الفقرات وفق النظرية التقليدية للاختبار الأول، ومعاملات الصعوبة وفق نظرية الاستجابة للفقرة لنفس الاختبار وجد أنه يساوي

-0.88. وهو ارتباط عكسي عال، وإذا حذفت الفقرة 47 التي معاملي صعوبتها 0.24 و 5.32 نجد أن معامل الارتباط قد ارتفع إلى -0.95 والشكلان 25، 26 يوضحان ذلك

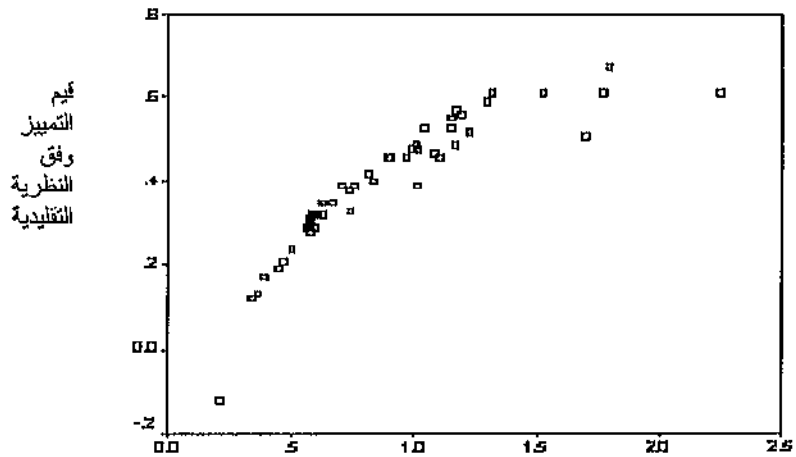


شكل 25: لوحة انتشار معاملات صعوبة الفقرات وفق النظرية التقليدية للاختبار الأول ومعاملات الصعوبة وفق نظرية الاستجابة للفقرة(النموذج الثنائي) لنفس الاختبار



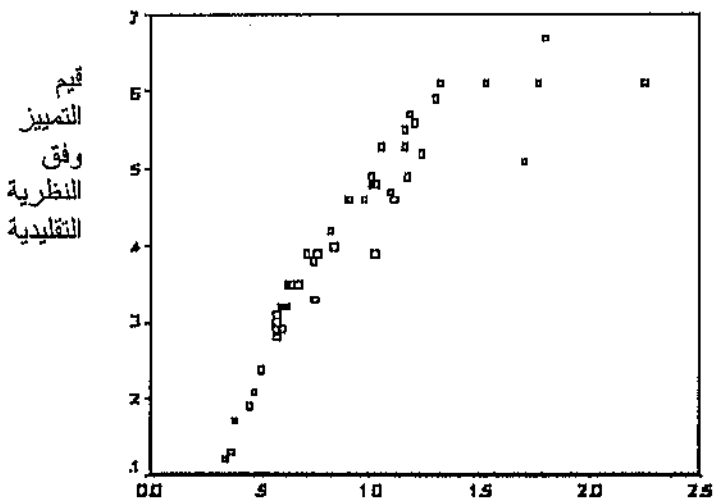
شكل 26: لوحة انتشار معاملات صعوبة الفقرات وفق النظرية التقليدية للاختبار الاول ومعاملات الصعوبة وفق نظرية الاستجابة للفقرة(النموذج الثنائي) لنفس الاختبار

وعند حساب معامل ارتباط بيرسون بين معاملات تمييز الفقرات وفق النظرية التقليدية للاختبار الأول، ومعاملات التمييز وفق نظرية الاستجابة للفقرة لنفس الاختبار وجد أنه يساوي 0.88 .وهو ارتباط عال، وإذا حذفت الفقرة 47 التي معاملي تمييزها –0.12 و 0.21 نجد أن معامل الارتباط قد ارتفع الى 0.90 والشكلان27 ،28 يوضحان ذلك



قيم التمييز وفق النموذج الثنائي

شكل27: لوحة انتشار معاملات تمييز الفقرات وفق النظرية التقليدية للاختبار الاول ومعاملات التمييز وفق نظرية الاستجابة للفقرة(النموذج الثنائي) لنفس الاختبار



قيم التمييز وفق النموذج الثنائي

شكل28: لوحة انتشار معاملات تمييز الفقرات وفق النظرية التقليدية للاختبار الاول ومعاملات التمييز وفق نظرية الاستجابة للفقرة(للنموذج الثنائي) لنفس الاختبار

وقد جرى مطابقة البيانات للنموذج ثنائي المعلمة من خلال اعتماد معامل الارتباط الثنائي النقطي بين علامة الفقرة والعلامة الكلية على الاختبار كمعيار لحذف الفقرات غير الملائمة للنموذج، وقد تم حذف جميع الفقرات التي كان معامل ارتباطها أقل من –0.15 وقد بلغ عدد هذه الفقرات 4 فقرات.

تدريج الفقرات

بعد مطابقة البيانات للنموذج ثنائي المعلمة والحصول على معالم الفقرات باستخدام برمجية (bilog-mg) تم استخدام الفقرات الرابطة (المشتركة) لتحويل التقديرات إلى مقياس عام مشترك. بافتراض أن توزيعات السمة لها وسط =صفر وانحراف معياري =1 .ثم حسبت الأوساط الحسابية للصعوبات المقدرة للفقرات المشتركة وكذلك الأوساط الهندسية المقدرة لميول الفقرات المشتركة. ولتحويل التقديرات إلى مقياس عام مشترك تم تثبيت النموذج الأول كمرجع و إجراء التعديلات للصعوبة والتمييز لكافة النماذج حسب المعادلات الخطية الآتية:

بالنسبة للنموذجين الأول والثاني لنفرض أن الوسطين الحسابيين لصعوبة الـ(9) فقرات المشتركة في التدريجين المنفصلين للنموذجين هما:

$$\bar{b}_1 = \frac{1}{9}\sum_{j=1}^{9} b_{1j} \quad \text{.....24}$$

$$\bar{b}_2 = \frac{1}{9}\sum_{j=1}^{9} b_{2j} \quad \text{.....} \quad \text{.....25}$$

ولنفرض ان الوسطين الهندسيين لميول هذه الفقرات في التدريجين هما :

$$\bar{a}_1 = \exp\left\{\frac{1}{9}\sum_{j=1}^{9} \log_e a_{1j}\right\} \quad \text{..... 26}$$

$$\bar{a}_2 = \exp\left\{\frac{1}{9}\sum_{j=1}^{9} \log_e a_{2j}\right\} \quad \text{..... 27}$$

لنفرض ان $r = \sqrt{\bar{a}_1 / \bar{a}_2}$

عندها تكون الميول المعدلة لكل الفقرات في النموذجين 1، 2 هي :

$$a^*_{2j} = r.a_{2j} \quad \text{..... 28}$$

$$a^*_{1j} = {a_{1j}}/{r} \quad \text{..... 29}$$

ولنفرض أن :

$$d = (r.\bar{b}_1 - \bar{b}_2 / r)/2$$

عندها تكون الصعوبة المعدلة لكل الفقرات هي :

$$b^*_{1j} = rb_{1j} - d \quad \text{----------30}$$

$$b^*_{2j} = b_{2j}/(r+ d) \quad \text{----------31}$$

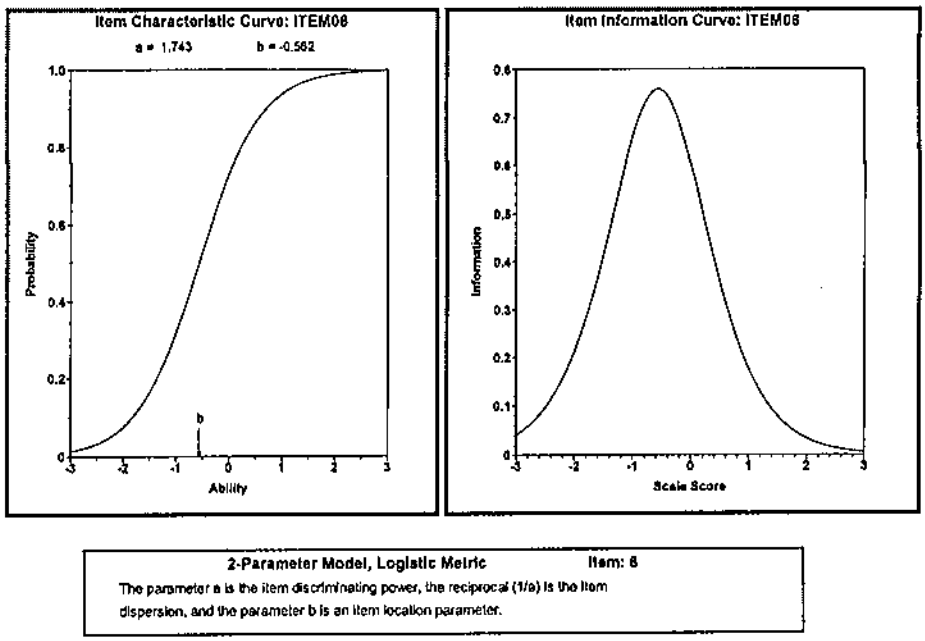
وبنفس الطريقة السابقة يمكن إيجاد المعادلات اللازمة بين الأول والثالث، والأول والرابع، والأول والخامس، والأول والسادس.وبهذه الطريقة تم الحصول على معالم الفقرات (الصعوبة b والتمييز a) والحد الأعلى (MaxIIF) من النماذج الاختبارية الست على تدريج واحد كما يظهر في الجدول 22 .وهذا الإجراء يوفر إمكانية أن تكون صعوبة الفقرات المنتقاة من أي صورة اختبارية على نفس المقياس الذي يمثل بنك الأسئلة، وقد تمت معالجة هذه العمليات جميعا باستخدام برنامج الماتلاب (MATLAB v6.5).

جدول 22: معاملات الصعوبة*b ومعاملات التمييز*a والحد الأعلى لدالة معلومة الفقرة لفقرات البنك بعد التدريج

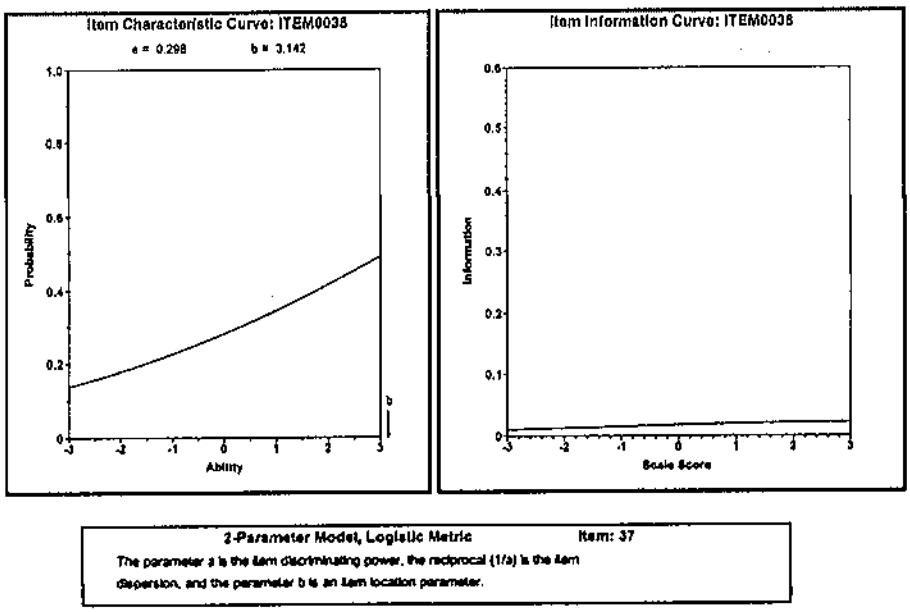
رقم الفقرة	الصعوبة b*	التمييز a*	Max iif	رقم الفقرة	الصعوبة b*	التمييز a*	Max iif	رقم الفقرة	الصعوبة b*	التمييز a*	Max iif
3/88	2.18-	.34	0.03	1/46	.11	1.01	.32	1/44	1.49	.55	.1
4/126	1.7-	1.04	0.27	2/78	.17	.94	.24	6/216	1.52	.51	.06
1/12	1.63-	.58	0.12	4/149	.18	.76	.06	1/30	1.54	.63	.12
1/1	1.62-	.66	0.14	3/97	.19	.41	.06	2/64	1.56	.42	.21
3/89	1.61-	.77	0.15	1/16	.2	.49	.1	6/199	1.6	.6	.15
6/196	1.59-	1.01	0.24	2/79	.21	1.07	.04	4/123	1.6	.44	.04
4/131	1.48-	1.58	0.14	1/43	.22	1.21	.46	4/155	1.6	.54	.08
6/206	1.21-	1.67	0.74	1/13	.22	.95	.3	2/84	1.61	.7	.04
2/56	1.13-	.86	0.15	6/229	.25	1.2	.38	4/139	1.62	.58	.08
4/127	1.12-	.99	0.24	1/37	.26	.97	.28	3/96	1.66	.42	.04
2/69	1.11-	1.53	0.02	6/214	.27	1.26	.4	6/198	1.67	.45	.13
4/137	1.05-	1.92	0.08	4/136	.28	1.35	.2	5/191	1.68	.5	.02
3/105	1.02-	1.05	0.27	5/186	.29	.55	.2	2/80	1.7	.42	.07
6/213	.97-	2.25	1.38	1/11	.35	1.41	.65	1/5	1.76	1.01	.76
5/159	.91-	1.41	.42	3/116	.37	1.07	.28	5/185	1.77	.42	.07
6/207	.85-	1.1	.31	3/102	.4	.76	.17	5/179	1.8	.52	.02
3/119	.72-	1.23	.36	5/181	.41	.65	.34	2/81	1.85	.56	.03
5/170	.69-	1.22	.02	6/233	.42	1.05	.28	3/98	1.97	.63	.1
3/107	.68-	1.33	.42	6/215	.42	.91	.22	3/122	1.98	.28	.02
2/50	.67-	1.01	.21	5/188	.42	1.21	.07	6/218	1.99	.63	.14
5/177	.66-	1.13	.05	2/53	.49	1.05	.24	2/55	2.01	.7	.11
2/86	.65-	.95	.22	6/220	.5	.63	.14	5/172	2.02	.46	.03
4/125	.65-	1.36	.43	2/68	.51	.76	.46	6/224	2.05	.39	.04

.12	.44	2.06	2/67	.03	.97	.53	2/58	.21	1.41	.65-	4/148
.08	.51	2.11	4/156	.06	.5	.54	3/91	.07	1.27	.62-	5/182
.08	.59	2.15	6/225	.22	.78	.55	1/22	.3	1.09	.59-	6/205
.1	.54	2.23	1/32	.19	.73	.58	2/57	.96	1.33	.59-	4/135
.08	.45	2.32	1/45	.27	.78	.61	5/176	.8	1.77	.58-	6/211
.2	.58	2.34	4/141	.11	.6	.63	6/221	.32	.88	.58-	5/169
.1	.57	2.35	1/41	.04	.94	.63	4/150	.39	1.28	.56-	3/117
.17	.46	2.37	5/168	.31	1.14	.65	3/109	.04	.50	.51-	2/66
.04	.34	2.38	3/103	.08	.53	.65	3/110	.34	1.03	.51-	1/36
.07	.43	2.44	1/39	.2	.9	.66	6/203	.19	1.54	.49-	4/128
.04	.34	2.49	6/223	.06	.52	.66	2/52	.36	1.55	.45-	1/6
.06	.41	2.55	1/9	.11	.64	.67	3/113	.28	.93	.44-	4/138
.03	.54	2.57	4/151	.08	.97	.74	2/62	.42	1.54	.38-	4/134
.04	.37	2.59	6/202	.38	1.17	.75	6/201	.46	1.19	.38-	1/10
.06	.33	2.6	5/192	.14	.67	.76	6/210	.4	1.23	.37-	6/226
.08	.49	2.61	1/19	.16	.64	.77	6//197	.05	1.11	.35-	1/7
.04	.35	2.62	5/161	.22	.53	.78	5/193	.44	1.17	.35-	1/26
.06	.4	2.65	5/178	.05	.58	.8	5/167	.19	.84	.33-	5/194
.09	.32	2.66	5/180	0.22	.55	.82	4/158	.36	1.22	.33-	3/115
.05	.37	2.73	2/60	.13	.59	.84	5/175	.42	1.14	.32-	1/15
.03	.36	2.92	3/104	.58	.86	.85	4/129	.27	1.05	.32-	2/94
.08	.46	2.99	1/35	.12	.69	.86	3/95	.26	1.01	.29-	6/219
.03	.39	3.01	4/152	.1	.62	.86	3/112	.68	1.46	.29-	1/21
.04	.44	3.02	2/51	.28	.87	.87	4/144	.06	.95	.28-	2/65
.07	.38	3.03	4/154	.04	.34	.92	3/99	.06	.47	.28-	6/209
.06	.45	3.03	1/47	.21	.77	.95	2/77	.24	1.01	.28-	3/118
.52	.58	3.03	5/183	.11	.57	.97	1/18	.04	1.11	.27-	2/74
.03	.36	3.04	2/59	.17	.71	.97	6/200	.2	.77	.25-	1/2
.06	.37	3.06	4/153	.09	.57	.99	6/231	.26	1.05	.25-	3/100
.02	.38	3.08	2/75	.24	.87	1.01	1/34	.18	.81	.24-	6/228
.02	.27	3.18	1/24	.07	.63	1.04	5/166	.24	.97	.17-	6/212
.04	.33	3.29	3/101	.23	1.1	1.16	4/140	.8	1.8	.14-	6/208
.02	.32	3.31	5/173	.02	.29	1.09	5/160	.26	1.02	.13-	6/204
.02	.24	3.31	3/121	.05	.46	1.1	5/162	.02	.98	.12-	2/72
.14	.34	3.34	2/76	.08	.47	1.1	1/14	.35	1.22	.1-	3/108
.06	.41	3.38	1/25	.11	.58	1.12	6/230	.26	1.04	.09-	4/124
.02	.27	3.38	1/20	.08	.58	1.13	6/217	.31	.98	.08-	5/187
.04	.38	3.41	1/42	.1	.57	1.14	6/222	.41	.75	.07-	4/133
.09	.37	3.46	5/165	.18	.73	1.15	1/49	.31	.95	.06-	1/29
.02	.29	3.49	3/120	.09	.59	1.16	3/114	.18	.74	.04-	6/195
.35	.41	3.5	2/85	.05	.65	1.17	5/190	.08	.48	.04-	1/4
.04	.33	3.73	1/27	0.28	.57	1.18	4/157	.2	.99	.03-	2/54
.08	.3	3.77	5/174	.09	.57	1.23	5/189	.03	.94	.02-	5/164
.05	.35	3.87	1/17	.06	.52	1.4	5/163	.42	1.12	.01-	1/23
.02	.33	4.06	3/92	.19	.45	1.42	2/61	.18	.98	.01-	4/142
.12	.29	4.17	2/70	.11	.62	1.42	6/227	.14	.98	.01-	4/147
.56	.43	4.27	4/132	.04	.57	1.44	2/63	.47	1.1	0	4/146
.04	.27	4.39	5/184	.2	.75	1.44	2/71	.13	.74	.01	2/87
.01	.21	4.42	3/111	.05	.44	1.44	3/106	.04	.9	.03	4/130
.04	.28	4.44	5/171	.04	.71	1.44	4/143	.44	1.19	.04	1/28
.08	.39	4.46	2/82	.22	.39	1.45	4/145	.16	.79	.05	3/90
.05	.39	4.5	1/48	.1	.61	1.46	2/83	.1	.63	.06	3/93
.05	.24	4.62	2/73	0.02	0.26	5.94	1/40	0.02	0.25	6.01	1/31
.04	.32	4.64	1/3	0.03	0.23	5.71	1/8	0.02	0.29	5.81	1/38
-	-	-	-	0.04	0.33	5.27	1/33	0.01	0.21	5.32	6/232

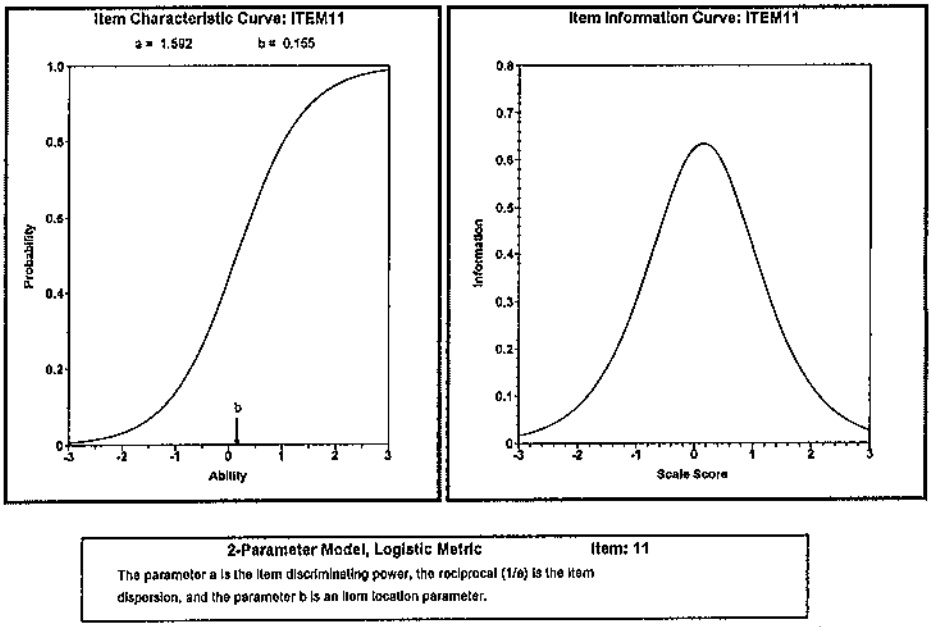
وعند رسم منحنيات خصائص الفقرات ICC، ودالة معلومة الفقرة IIF بعد تدريج كل اختبارعلى انفراد، تبين أن أفضل الفقرات هي الفقرة رقم6 /1 والشكل 29 يوضح ذلك . بينما كانت الفقرة رقم 37/ 5 من أسوا الفقرات والتي تظهر في الشكل 30 ، في حين كانت الفقرة رقم 1/11 جيدة والشكل 31 يوضح ذلك .



شكل 29: منحنى خاصية الفقرة على اليسار ومنحنى معلومات الفقرة على اليمين لأفضل فقرة



شكل 30: منحنى خاصية الفقرة على اليسار ومنحنى معلومات الفقرة على اليمين لأسوا فقرة



شكل 31: منحنى خاصية الفقرة على اليسار ومنحنى معلومات الفقرة على اليمين لفقرة جيدة

أشارت نتائج التدريج إلى أن جميع الفقرات في الصور الاختبارية تقع على تدريج واحد مقدرة بوحدة اللوجيت، وهذا يؤدي إلى أن كل الفقرات المسحوبة من البنك لها نفس التدريج، وبالتالي تعطي تقديرا لقدرة المفحوصين على نفس التدريج.

الكفاءة النسبية(Relative Efficiency)

لمقارنة الكفاءة النسبية لاختبارين يقيسان نفس البناء (Construct) فإننا نستخدم الصيغة التالية :

$$RE(\theta) = \frac{I_A(\theta)}{I_B(\theta)} \quad \text{..... 32}$$

حيث RE(θ): الكفاءة النسبية

$I_A(\theta)$: معلومات الاختبار A

$I_B(\theta)$: معلومات الاختبار B

وهذه القيم تعتمد على :

- عدد الفقرات
- نوعية الفقرات
- التوافق بين صعوبة الفقرة وقدرة المفحوصين.

اما عن كيفية ارتباط المعالم بمقدار المعلومات في الاختبار:

- تزداد المعلومات عندما تقترب قيمة معلمة الصعوبة (b) من (θ)
- تزداد المعلومات عندما تزيد قيمة معلم التمييز (a)
- تزداد المعلومات عندما تقترب قيمة معلم التخمين (c) من الصفر.

وهذا يساعد في تقويم الاختبار وانتقاء فقراته. والجدول 23 يظهر معلومات الاختبارات TIF الستة بالإضافة إلى الكفاءة النسبية عند مقارنة الاختبارات جميعها مع الاختبار الأول.

جدول 23:معلومات الاختبارات الستة والكفاءة النسبية

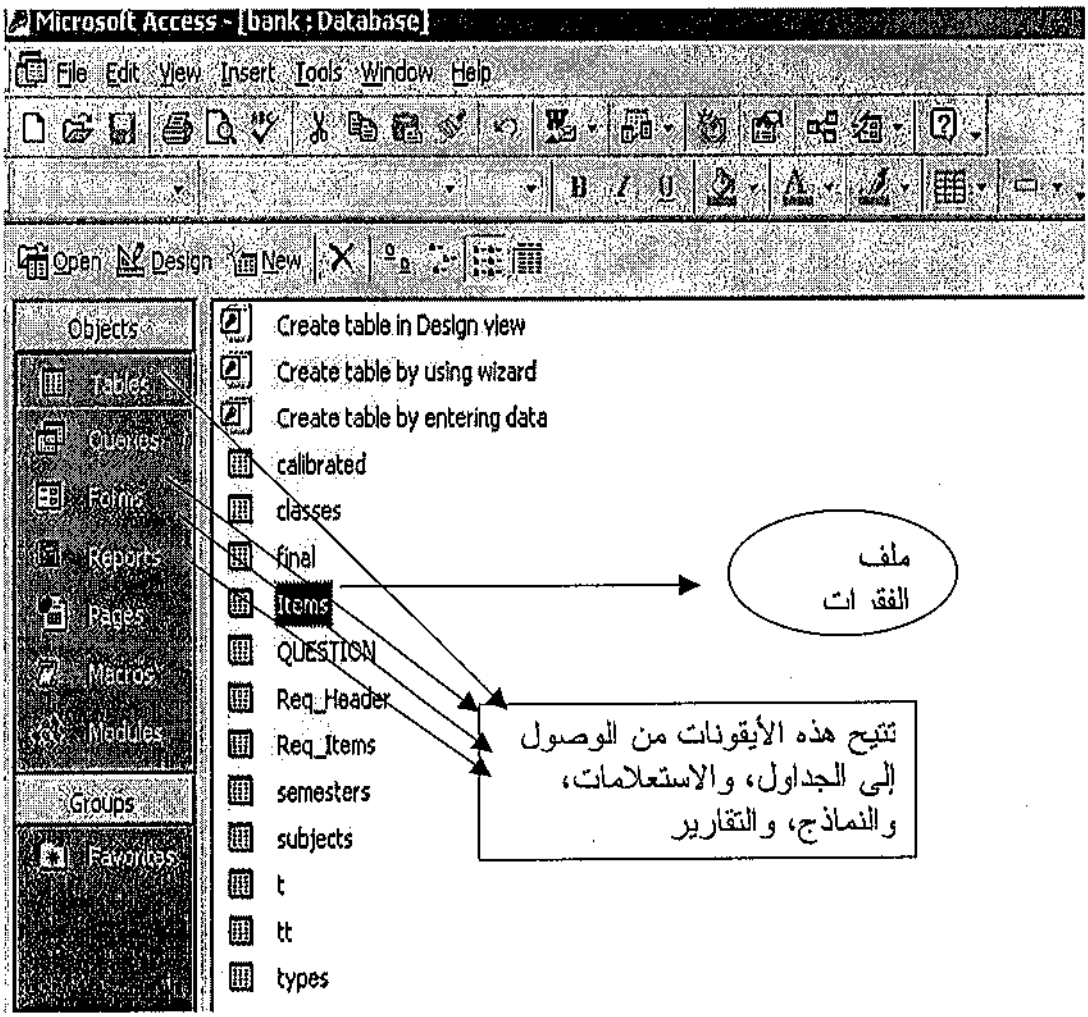
$RE(\theta)$	TIF	عدد الفقرات	رقم الاختبار
1	8.6	49	1
0.67	5.2	46	2
0.88	7.6	46	3
1.07	9.2	47	4
0.61	5.2	45	5
1.20	10.3	49	6

عند مقارنة الاختبارات (2 ، 3 ، 4 ، 5 ، 6) مع الاختبار الأول في الجدول 23 نلاحظ أن الكفاءة النسبية للاختبار الثاني تساوي 0.67 عند مقارنته بالاختبار الأول ، وهذا يدل على أن الاختبار الثاني يؤدي وظيفته كما لو كان أقصر من الاختبار الأول بـ 33% ، وكذلك يؤدي الاختبار الثالث وظيفته كما لو كان أقصر من الاختبار الأول بـ 12% ، و يؤدي الاختبار الرابع وظيفته كما لو كان أطول من الاختبار الأول بـ 7% ، أما بالنسبة للاختبار الخامس فانه يؤدي وظيفته عندما يكون اقصر من الأول بنسبة 39% ، في حين أن الاختبار السادس يؤدي وظيفته كما لو كان أطول من الاختبار الأول بنسبة 20% .

وصف البنك:

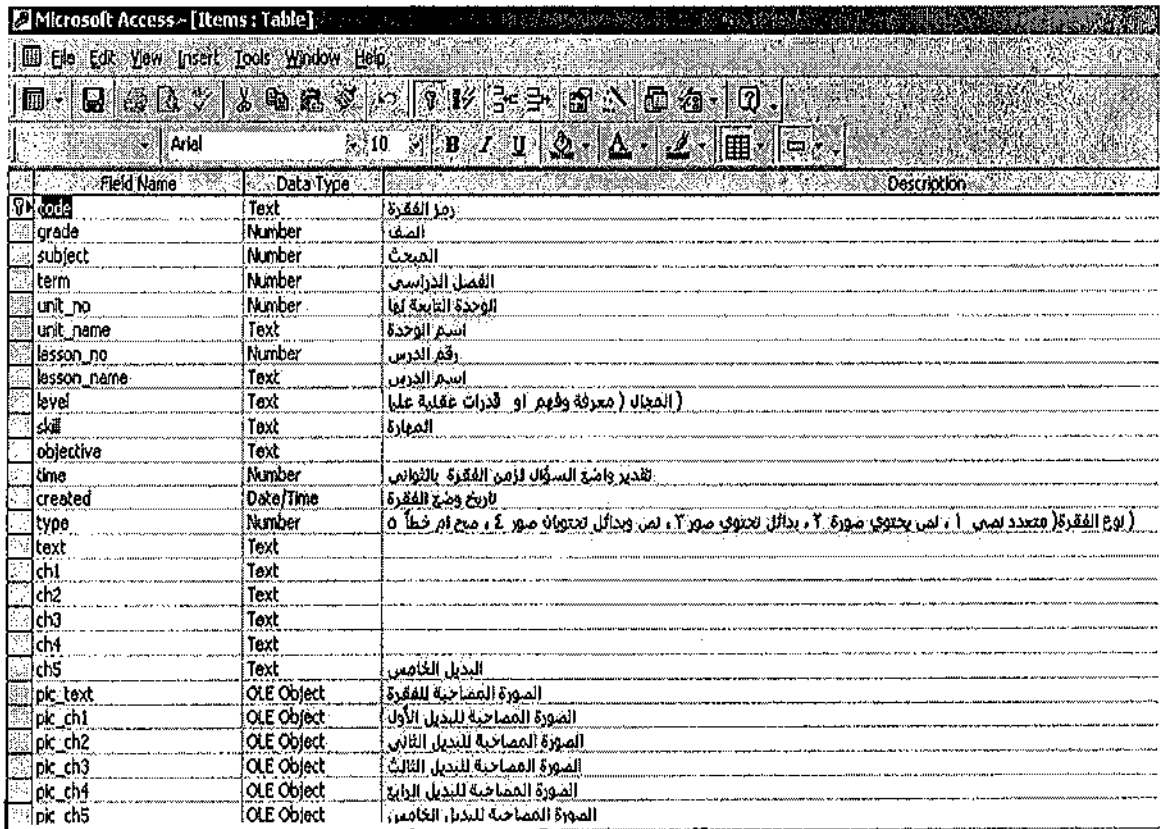
بعد تدريج الفقرات وتحديد خصائصها تم تخزينها في بنك الأسئلة، بحيث يمكن استخدامها لاستخراج اختبار حسب المواصفات المطلوبة للاختبار، ولتصميم نظام بنك الأسئلة، تم إنشاء قاعدة البيانات لبناء البنك باستخدام برنامج قواعد البيانات (Microsoft Access) حيث أن هذا البرنامج يتميز بالسهولة في التعامل والمرونة في التطوير،وسهولة الحصول عليه ، وتخزين قدر هائل من البيانات ، والسماح بفرز البيانات بسرعة وسهولة ، والبحث عن البيانات وتحديدها. واشتملت قاعدة البيانات على المكونات الأساسية التالية:

1) الجداول Tables: وتختص الجداول بالتعامل مع البيانات في صورة ملفات.وهناك مجموعة من الملفات مثل : ملف الفقرات ، و ملف الاختبارات ، والشكل 32 يمثل واجهة قاعدة البيانات التي تحوي المكونات(objects) وما هو موجود داخل الجداول.



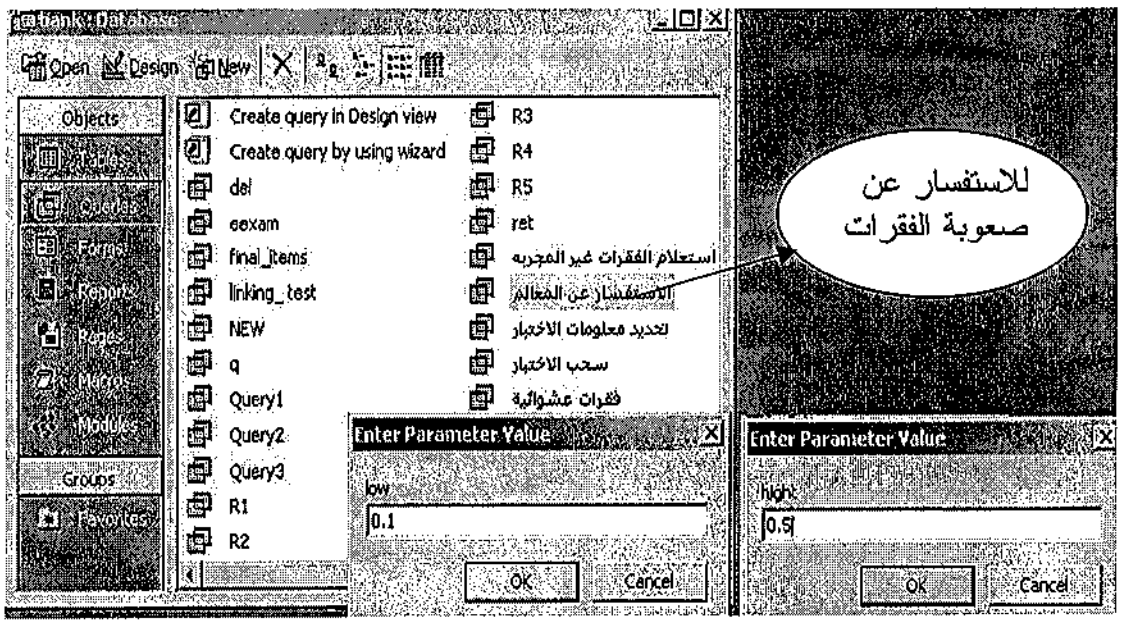
شكل32: واجهة قاعدة البيانات ومكوناتها ومحتويات الجداول

عند النقر على أيقونة الجداول (tables) تظهر محتويات الجداول وهي عبارة عن ملفات ، اختر ملف الفقرات (items)، ثم اضغط كبسة يسار يظهر الشكل 33، والذي يمثل تصميم حقل ملف الفقرات.



شكل 33: تصميم حقل ملف الفقرات

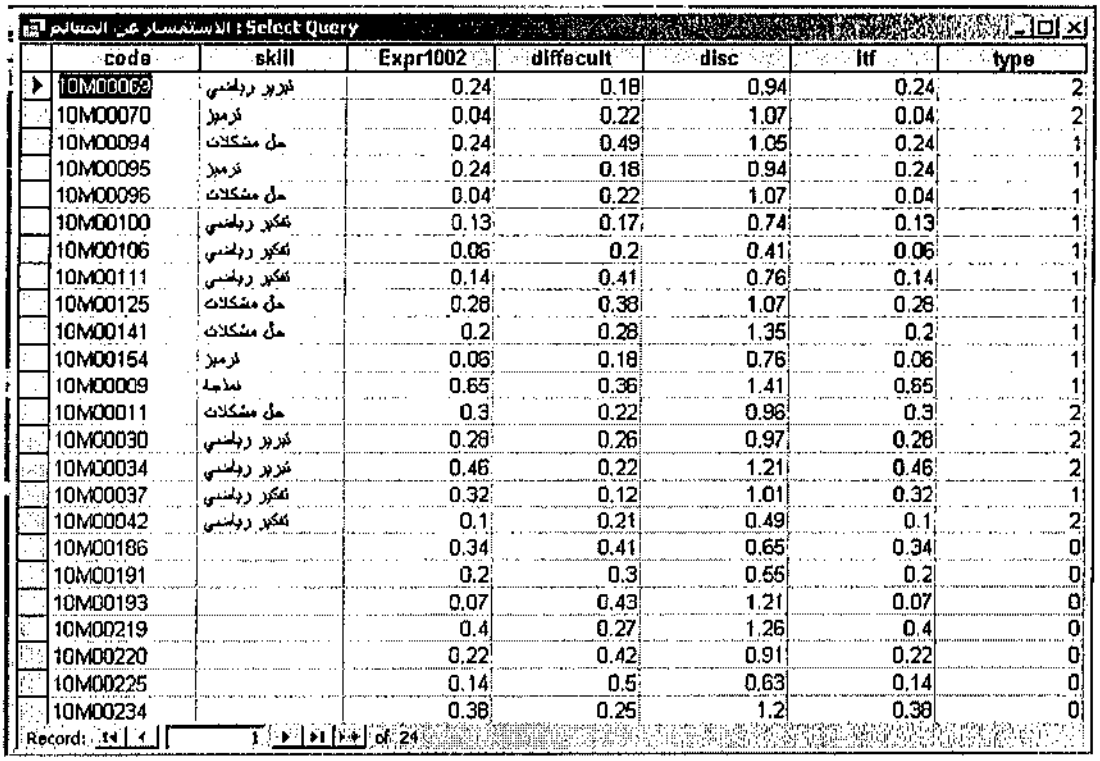
2) الاستعلامات query: عبارة عن سؤال خاص بالبيانات، مثل الاستفسار عن الفقرات التي معامل صعوبتها 0.5-0.1. الشكل 34 يظهر واجهة قاعدة البيانات ومحتويات الاستعلام.



شكل 34: واجهة قاعدة البيانات ومحتويات الاستعلام.

مثال : الاستفسار عن جميع الفقرات التي تنحصر قيم معاملات الصعوبة لها بين (0.5 - 0.1)

- اختار الكائن queries
- اختار ايقونة الاستفسار عن المعالم
- تظهر الواجهة enter parameter value
- ادخل الحد الادنى ثم اضغط ok
- ادخل الحد الاعلى ثم اضغط ok
- يظهر الشكل 35 والذي يمثل المعلومات التي تم الاستعلام عنها



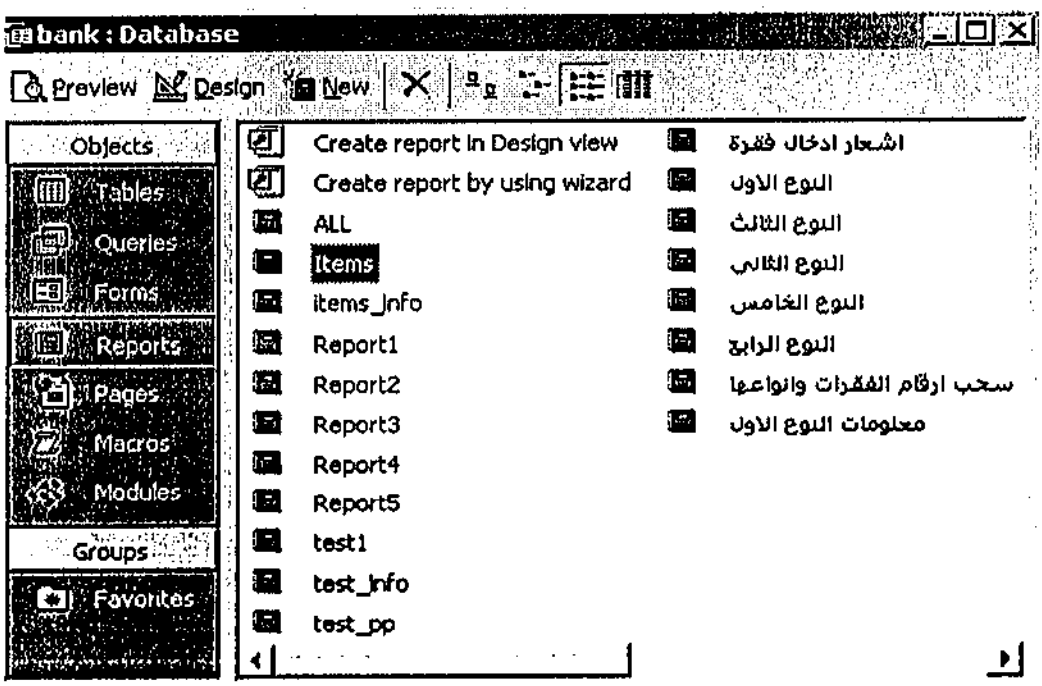
الاستفسار على الصفائح : Select Query

code	skill	Expr1002	diffecult	disc	itf	type
10M00069	تفريق رياضي	0.24	0.18	0.94	0.24	2
10M00070	ترميز	0.04	0.22	1.07	0.04	2
10M00094	حل مشكلات	0.24	0.49	1.05	0.24	1
10M00095	ترميز	0.24	0.18	0.94	0.24	1
10M00096	حل مشكلات	0.04	0.22	1.07	0.04	1
10M00100	تفكير رياضي	0.13	0.17	0.74	0.13	1
10M00106	تفكير رياضي	0.06	0.2	0.41	0.06	1
10M00111	تفكير رياضي	0.14	0.41	0.76	0.14	1
10M00125	حل مشكلات	0.28	0.38	1.07	0.28	1
10M00141	حل مشكلات	0.2	0.28	1.35	0.2	1
10M00154	ترميز	0.06	0.18	0.76	0.06	1
10M00009	نمذجة	0.65	0.36	1.41	0.65	1
10M00011	حل مشكلات	0.3	0.22	0.96	0.3	2
10M00030	تفريق رياضي	0.28	0.26	0.97	0.28	2
10M00034	تفريق رياضي	0.46	0.22	1.21	0.46	2
10M00037	تفكير رياضي	0.32	0.12	1.01	0.32	1
10M00042	تفكير رياضي	0.1	0.21	0.49	0.1	2
10M00186		0.34	0.41	0.65	0.34	0
10M00191		0.2	0.3	0.55	0.2	0
10M00193		0.07	0.43	1.21	0.07	0
10M00219		0.4	0.27	1.26	0.4	0
10M00220		0.22	0.42	0.91	0.22	0
10M00225		0.14	0.5	0.63	0.14	0
10M00234		0.38	0.25	1.2	0.38	0

Record: 1 of 24

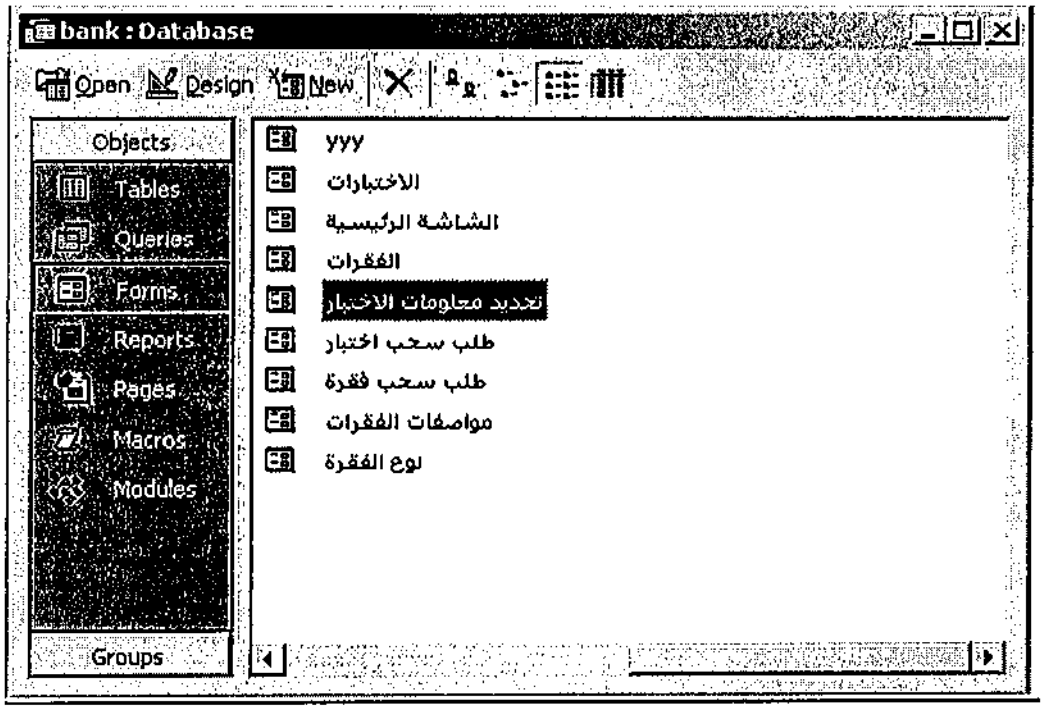
شكل 35: جدول المعلومات التي تم الاستعلام عنها

3) التقارير report: أما التقرير فيستخدم لاستخراج نماذج مطبوعة من الملفات والاستفسارات التي تم تصميمها، وعند تصميم التقرير يتم وضع الصورة التي يكون عليها المخرجات الورقية أو المرئية على الشاشة. والشكل 36 يبين واجهة التقارير ومحتوياتها.



الشكل 36: واجهة التقارير ومحتوياته

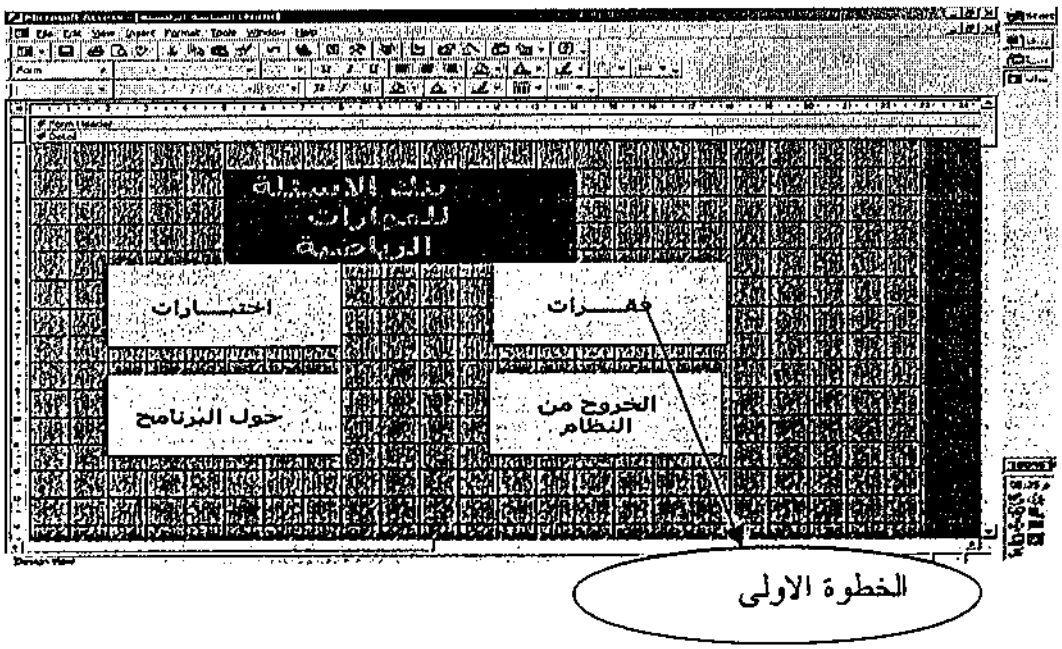
4) النماذج forms: يمكنك أن تستخدم النموذج في إدخال أو تحرير أو عرض المعلومات في سجلات قاعدة البيانات.والشكل 37 يبين واجهة النماذج ومحتوياتها



شكل 37: واجهة النماذج ومحتوياتها

وعندما نريد تخزين فقرات جديدة في البنك نتبع الخطوات التالية:

1- نختار من الشاشة الرئيسة زر الفقرات فتظهر نافذة المكونات (objects) كما في الشكل 38



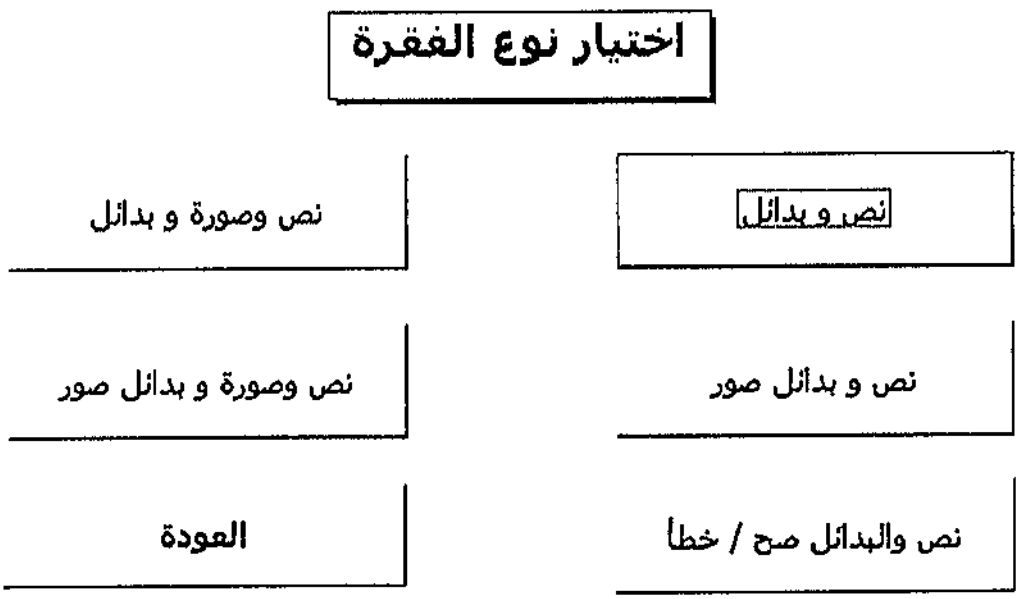
شكل38: الشاشة الرئيسة لبنك الأسئلة

2- من نافذة المكونات (objects) نختار النموذج form ومن محتوياته نختار زر الفقرات (items)

3- تظهر نافذة الفقرات ومن محتوياتها(إضافة فقرة، إرجاع فقرة، حذف فقرة، . . .) نختار زر إضافة فقرة كما في الشكل 39

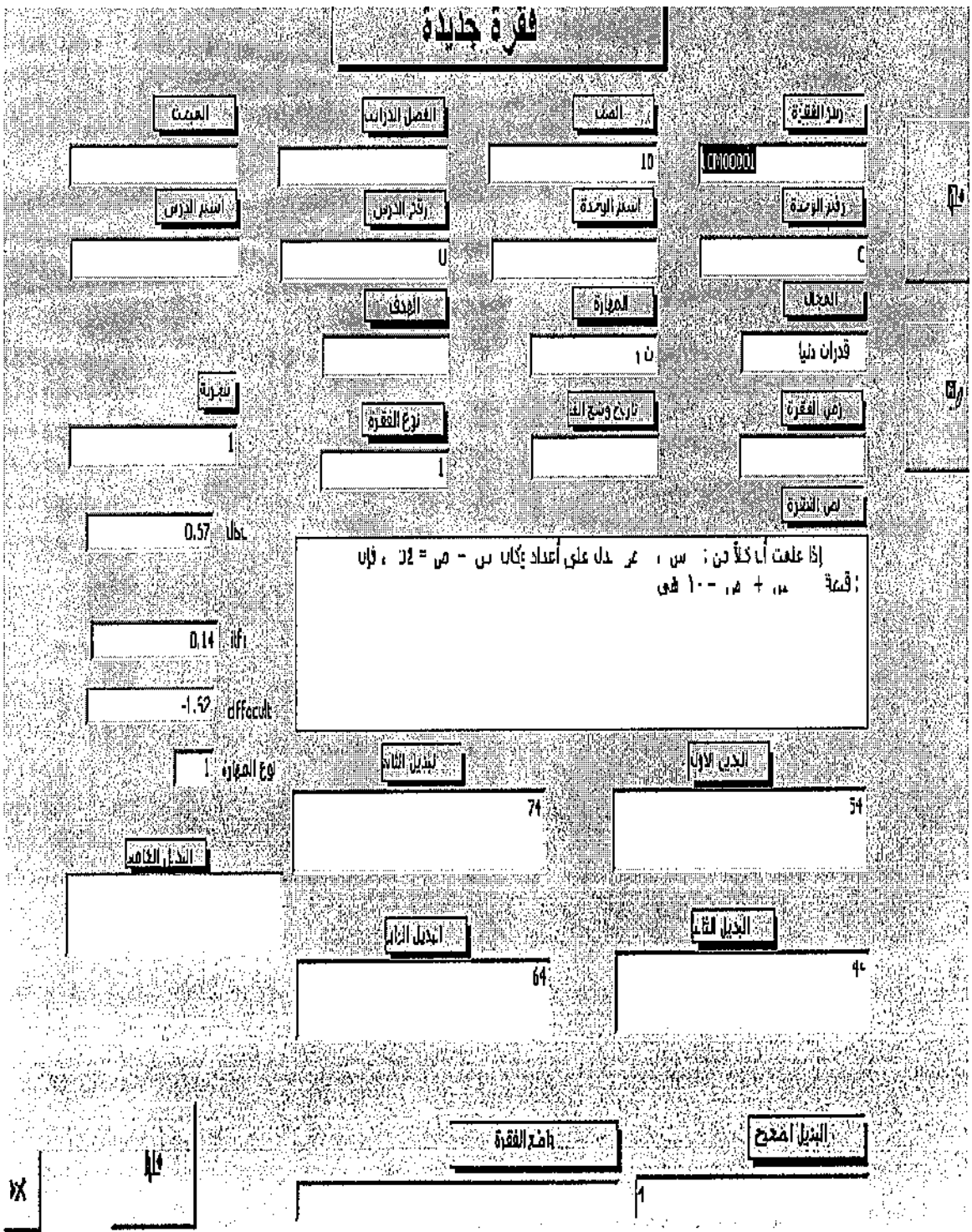
شكل39: نافذة الفقرات ومحتوياتها

4- تظهر نافذة اختيار نوع الفقرة ومن محتوياتها (نص وبدائل، نص وبدائل صور. . .) كما في الشكل 40



شكل 40: نافذة نوع الفقرة ومحتوياتها

5- نختار ما يتناسب ونوع الفقرة فمثلا عند اختيار نص وبدائل تظهر نافذة بعنوان فقرة جديدة كما في الشكل 41



* هناك بعض الخلايا غير مفعلة في هذه النافذة، قد تستخدم عند الحاجة، مثل عند مرات استخدام الفقرة...

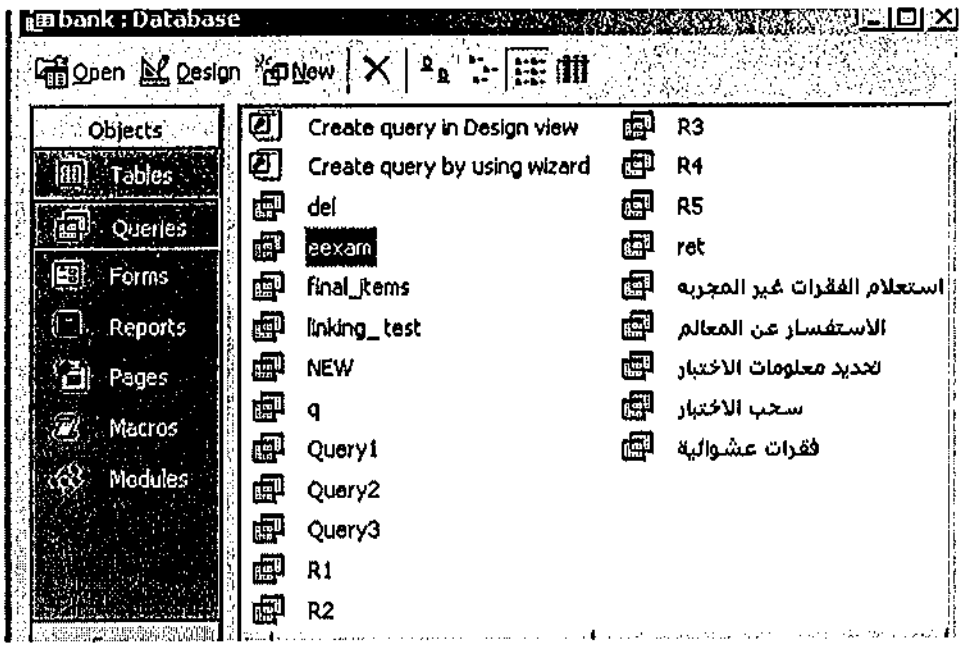
شكل 41: فقرة جديدة نص وبدائل

6)تملأ البيانات المتوفرة عن الفقرة كما في الشكل 41 ثم يتم حفظها

خطوات سحب اختبار من البنك:

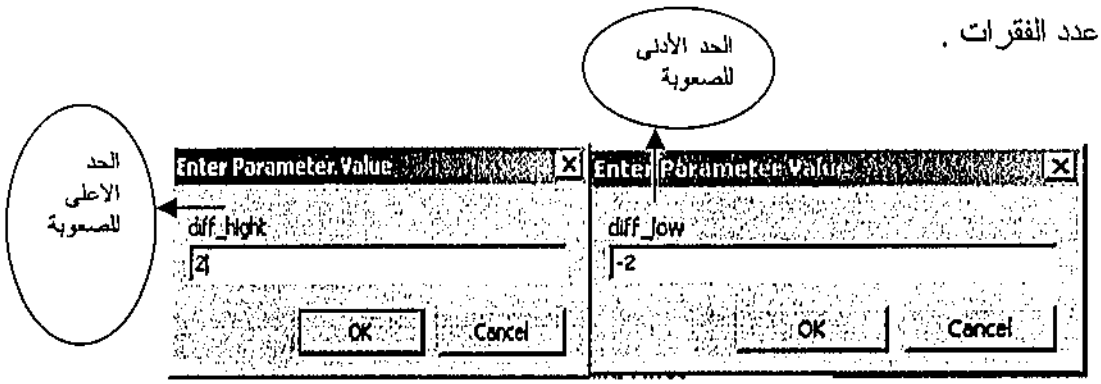
1) من الشاشة الرئيسة اختار زر الاختبارات فتظهر نافذة المكونات

2) من نافذة المكونات اختار الاستعلامات (queries) كما في الشكل 42 ومن محتويات هذه النافذة اختار (eexam)

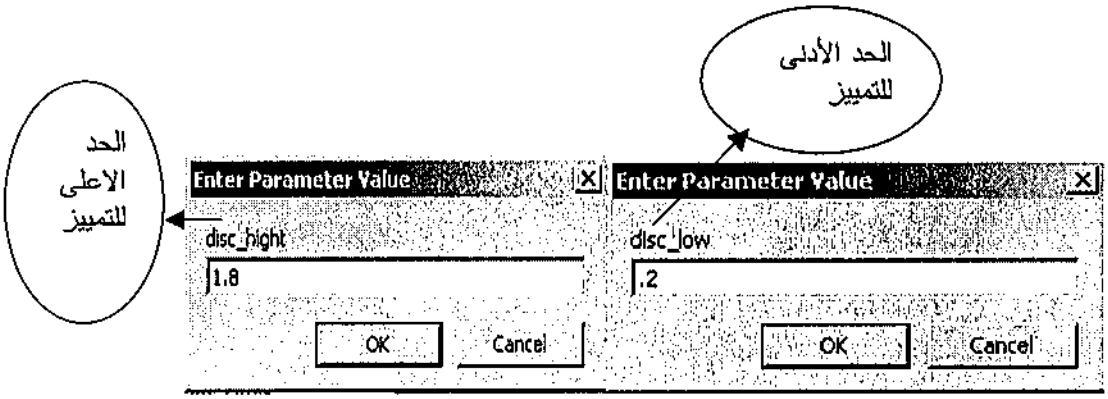


شكل 42 : واجهة الاستعلام ومكوناتها

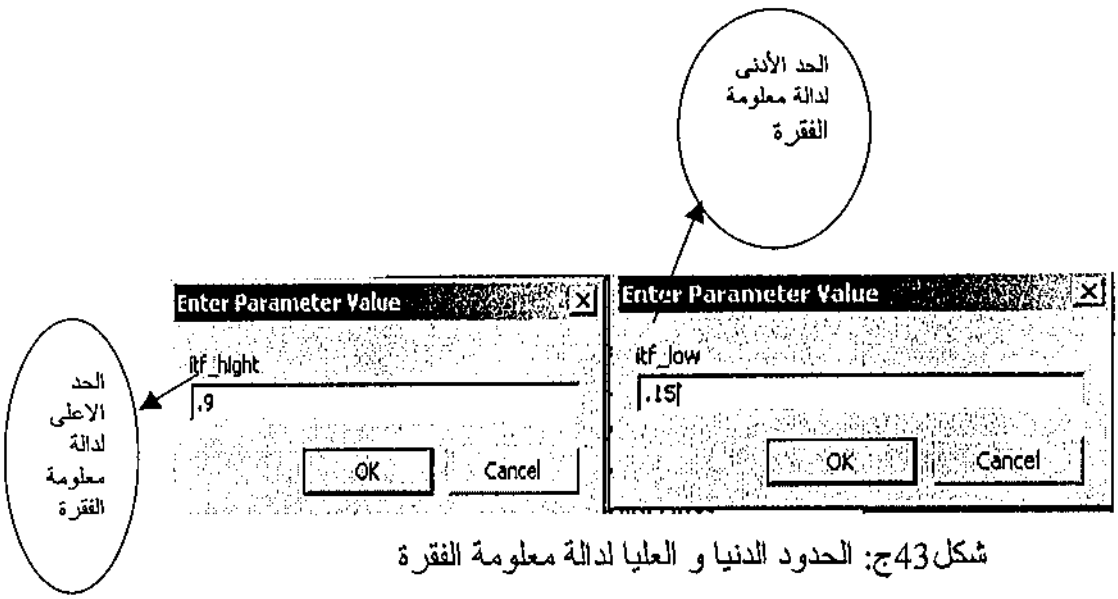
3)تظهر نافذة بعنوان(enter parameters values) كما في الاشكال 43أ و 43ب و 43ج و 43د ادخل الحدود الدنيا والعليا(للصعوبة ، والتمييز ، ودالة معلومة الفقرة) ، ونوع المهارة ، وعدد الفقرات .



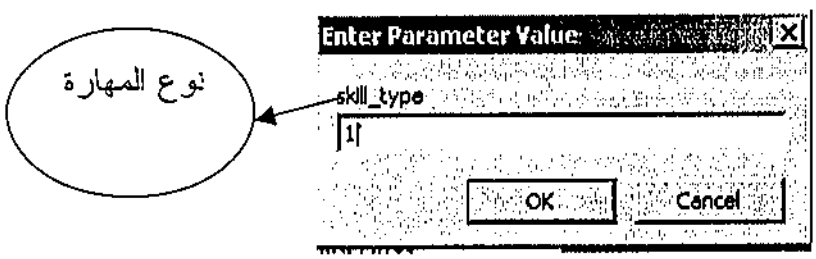
شكل 43أ: الحدود الدنيا و العليا لصعوبة الفقرة



شكل 43ب : الحدود الدنيا و العليا لتمييز الفقرة



شكل43ج: الحدود الدنيا و العليا لدالة معلومة الفقرة



شكل 43د : نوع المهارة

3) بعد ملء المعلومات التي تم تحديدها اضغط على زر ok تظهر الفقرات المطلوبة ضمن المواصفات كما في الشكل44.

bank : Database

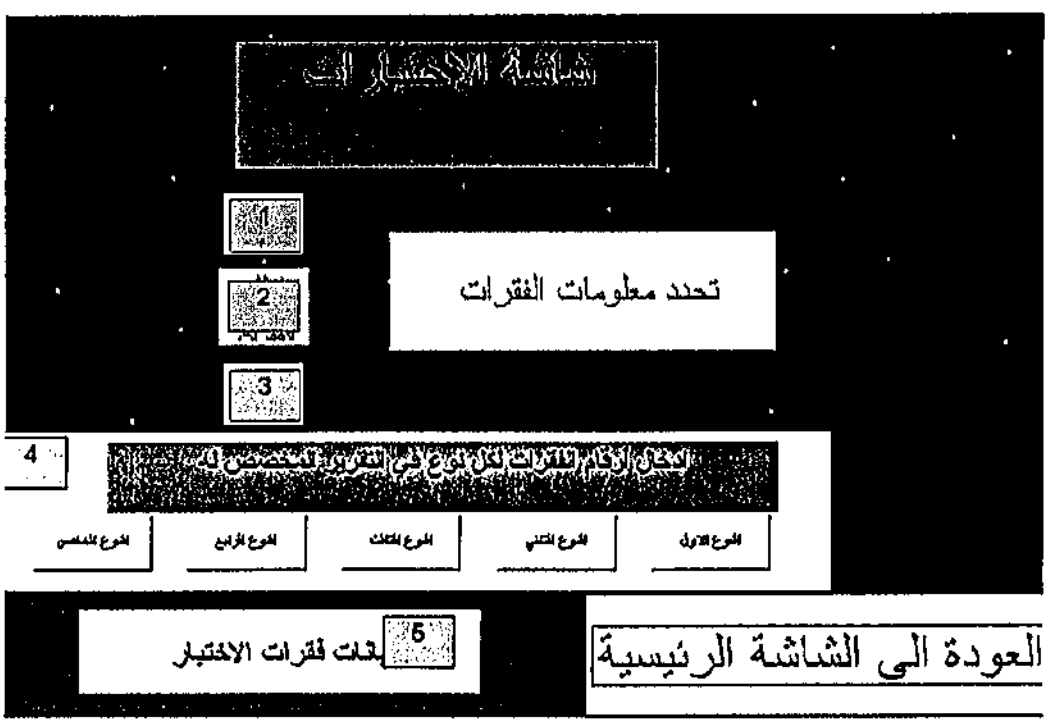
eexam : Select Query

code	type	diffecult	disc	itf	skill type	SumOfitf
10M00005	1	1.76	1.01	0.76	1	0.76
10M00037	1	0.12	1.01	0.32	1	0.32
10M00075	4	-1.61	0.77	0.15	1	0.15
10M00117	1	-0.1	1.22	0.35	1	0.35
10M00118	1	0.65	1.14	0.31	1	0.31
10M00126	1	-0.56	1.28	0.39	1	0.39
10M00135	1	-0.49	1.54	0.19	1	0.19
10M00143	1	-0.44	0.93	0.28	1	0.28
10M00147	1	-0.01	0.91	0.18	1	0.18
10M00149	1	0.87	0.87	0.28	1	0.28
10m00164	0	-0.91	1.41	0.42	1	0.42
10m00167	0	1.09	0.46	0.5	1	0.5
10m00169	0	-0.02	0.94	0.3	1	0.3

Record: 10 of 13

شكل 44: أرقام الفقرات المطلوبة ضمن مواصفات محددة

4) من الشاشة الرئيسة اختار زر الاختبارات تظهر النافذة كما في الشكل 45، املأ أرقام الفقرات من الشكل 44 حسب أنواعها، وبعد ذلك تحصل على الاختبار المطلوب.



شكل 45:شاشة إدخال أرقام الفقرات لكل نوع في التقرير المخصص له

القرار بالمعلومات المطلوبة. ولاثراء هذا البنك يمكن تطبيق العديد من الاختبارات التي تحتوي على فقرات جديدة وفقرات مشتركة من البنك.

ويمكن الإفادة من مميزات البنك وهي بناء الاختبارات المتكافئة حيث يتم استخدام أحدها للتطبيق القبلي والأخر للتطبيق البعدي، وامكانية المقارنة بين نتائج التطبيقين، وتمكن هذه الميزة المستخدم من تحديد موضع أداء الفرد على متصل القدرة ومتابعة مدى تقدمه، كما ويتيح الفرصة أمام مستخدمه للمقارنة بين أداء طالبين أو اكثر باستخدام فقرات مختلفة.

ثانيا: للإجابة عن السؤال الثاني: ما خصائص الاختبار المسحوب من بنك الأسئلة ؟

و لأجل سحب اختبار ضمن مواصفات معينة مثل: عدد فقرات الاختبار، النسبة المئوية للفقرات من كل مهارة، مدى صعوبة الفقرات، مدى تمييز الفقرات، دالة معلومة الاختبار. تم تجريب البرنامج الحاسوبي لاختيار فقرات الاختبار، حيث قدمت للبرنامج المدخلات الاتية:

عدد الفقرات 25 فقرة	
النسبة المئوية لمهارة التفكير الرياضي	40%
النسبة المئوية لمهارة حل المشكلة	40%
النسبة المئوية لمهارة التعبير بالرموز	10%
النسبة المئوية لمهارة التبرير	10%
مدى صعوبة الفقرات	3^{+} - 3^{-}
مدى تمييز الفقرات	2 - 0.35
دالة معلومة الاختبار	12 - 10

بعد تخزين هذه المواصفات في قاعدة البيانات حصلنا على المخرجات آلاتية:

عدد الفقرات من مهارة التفكير الرياضي 10، وعدد الفقرات من مهارة حل المشكلة 10، وعدد الفقرات من مهارة التعبير بالرموز 3، وعدد الفقرات من مهارة التبرير 2

والجدول 24 يوضح أرقام الفقرات التي تم اختيارها، والمعلومات الخاصة بكل فقرة.

جدول 24 :أرقام الفقرات التي تم اختيارها مع المعلومات الخاصة بها

رقم الفقرة	المهارة	الصعوبة	التمييز	دالة معلومة الفقرة
1\2	تفكير رياضي	0.25-	0.77	0.20
1\6	حل مشكلات	0.45-	1.55	0.76
1\7	حل مشكلات	0.35-	1.11	0.39
1\10	تبرير	0.38-	1.19	0.44
1\11	تعبير بالرموز	0.36	1.41	0.64
1\21	تفكير رياضي	0.29-	1.46	0.68
6\19	حل مشكلات	1.21-	1.68	0.73
6\21	حل مشكلات	0.14-	1.8	0.81
6\24	تفكير رياضي	0.58-	1.8	0.80
5\1	تفكير رياضي	0.91-	1.41	0.44
5\12	حل مشكلات	0.58-	0.88	0.32
5\19	حل مشكلات	0.61	0.78	0.26
4\23	تفكير رياضي	0.38-	1.54	0.42
4\35	حل مشكلات	0.001	1.1	0.57
3\24	تبرير	0.65	1.14	0.30
3\23	تفكير رياضي	0.10-	1.22	0.35
3\22	تفكير رياضي	0.68-	1.33	0.42
3\15	تفكير رياضي	0.25-	1.05	0.27
3\9	تعبير بالرموز	0.32-	1.05	0.26
2\31	تفكير رياضي	0.18	0.94	0.24
2\23	حل مشكلات	1.44	0.75	0.19
2\16	تعبير بالرموز	1.56	0.42	0.20
2\5	حل مشكلات	0.03-	0.99	0.20
2\4	تفكير رياضي	0.49	1.05	0.26
2\30	حل مشكلات	0.96	0.77	0.22

بلغت دالة معلومة الاختبار 10.27، ولحساب الخطأ المعياري في القياس من خلال دالة معلومة الاختبار يمكن استخدام المعادلة:

$$SE(\theta) = 1/\sqrt{I(\theta)} \quad \text{OR} \quad SE^2(\theta) = 1/I(\theta)$$

$$= 1/\sqrt{10.27}$$

$$= 0.31$$

حيث بلغ 0.31. يتضح من الجدول 24 أن معاملات الصعوبة للفقرات الاختبارية تباينت من الفقرات السهلة والمتوسطة والصعبة حيث تراوحت هذه المعاملات بين (-1.21 - 1.56)، وكذلك تمييز هذه الفقرات تراوح بين (0.42 - 1.8) وهو تمييز يشير إلى معيار الفقرة الفاعلة في التمييز بين مستويات القدرة المختلفة. أما معامل ثبات الاختبار فقد بلغ 0.904

ونظرا لاحتواء البنك على عدد كبير من الفقرات الاختبارية ، فان اختيار الفقرات الاختبارية لبناء اختبار عملية سهلة ، ولوجود قاعدة البيانات فإنها توفر المرونة في استدعاء وإضافة وحذف الفقرات ،وتكون هذه المواصفات تحت سيطرة المستخدم .

ثالثا: للإجابة عن السؤال الثالث: ما مدى استقرار معالم الفقرة في المجموعات المختلفة ؟ تم استخدام المؤشرات الإحصائية للفقرات المشتركة بين الصور الاختبارية الست وعددها 9 فقرات كما يلي:

أ) المؤشرات الإحصائية للفقرات المشتركة وفق النظرية التقليدية. يبين الجدول 25 المؤشرات الإحصائية للفقرات المشتركة وفق النظرية التقليدية.

جدول 25: المؤشرات الإحصائية للفقرات المشتركة قيم الصعوبة (P) وقيم التمييز(rbis) وفق النظرية التقليدية

الرقم	الأول		الثاني		الثالث		الرابع		الخامس		السادس	
	P	rbis	P	rbis	P	rbis	P	rbis	P	rbis	P	rbis
1	0.64	0.58	0.58	0.44	0.59	0.57	0.74	0.52	0.59	0.51	0.57	0.53
2	0.60	0.51	0.47	0.48	0.56	0.47	0.66	0.62	0.51	0.49	0.53	0.48
3	0.28	0.25	0.29	0.22	0.28	0.22	0.30	0.18	0.24	0.13	0.25	0.29
4	0.60	0.50	0.49	0.39	0.54	0.47	0.65	0.52	0.50	0.36	0.65	0.49
5	0.44	0.63	0.40	0.41	0.39	0.47	0.50	0.48	0.38	0.51	0.52	0.55
6	0.73	0.30	0.66	0.26	0.66	0.39	0.74	0.30	0.59	0.33	0.70	0.33
7	0.48	0.51	0.40	0.36	0.42	0.49	0.47	0.44	0.36	0.26	0.50	0.60
8	0.40	0.26	0.37	0.18	0.34	0.22	0.42	0.36	0.32	0.28	0.38	0.40
9	0.60	0.53	0.50	0.39	0.48	0.60	0.57	0.45	0.49	0.44	0.58	0.61

بلغ متوسط الصعوبة للفقرات المشتركة في الاختبار الأول 0.54 ، وللاختبار الثاني 0.46 ، وللاختبار الثالث0.47 ، وللاختبار الرابع 0.56 وللاختبار الخامس 0.44 ، وللاختبار السادس 0.52 . وهذا الوضع ربما يفسر على انه تكافؤ في التحصيل لمجموعات المفحوصين الست، أو/ وتكافؤ مستويات الصعوبة للمجموعات الست. ويبين الجدول 26 مصفوفة معاملات الارتباط بين قيم معاملات الصعوبة للمجموعات الست (بيرسون) .

جدول 26: مصفوفة معاملات الارتباط بين متوسطات صعوبة الفقرات المشتركة وفق النظرية التقليدية

الاختبار	الثاني	الثالث	الرابع	الخامس	السادس
الاول	0.96	0.98	0.97	0.98	0.93
الثاني		0.96	0.94	0.96	0.88
الثالث			0.98	0.98	0.89
الرابع				0.99	0.89
الخامس					0.89

يتضح من الجدول 26 أن درجة الارتباط قوية بين معاملات الصعوبة للفقرات المشتركة رغم اختلاف المجموعات.

يبين الجدول 27 مصفوفة معاملات الارتباط (بيرسون) بين معاملات التمييز للفقرات المشتركة بين الاختبارات الست وفق النظرية التقليدية.

جدول 27 : مصفوفة معاملات الارتباط بين معاملات التمييز للفقرات المشتركة وفق النظرية التقليدية.

الاختبار	الثاني	الثالث	الرابع	الخامس	السادس
الأول	0.92	0.87	0.82	0.81	0.86
الثاني		0.86	0.88	0.81	0.71
الثالث			0.71	0.74	0.83
الرابع				0.82	0.71
الخامس					0.6

يظهر من الجدول 27 أن معاملات الارتباط كانت عالية بين معاملات التمييز للفقرات المشتركة لكافة الاختبارات باستثناء الارتباط بين معاملات التمييز للاختبارين الخامس والسادس حيث كان مقبولا.

ب) المعالم الإحصائية للفقرات المشتركة وفق نظرية الاستجابة للفقرة. يبين الجدول 28 المعالم الإحصائية للفقرات المشتركة وفق نظرية الاستجابة للفقرة.

جدول 28: المعالم الإحصائية الصعوبة (b) والتمييز (a) للفقرات المشتركة وفق نظرية الاستجابة للفقرة

رقم الفقرة	الاختبار الأول		الاختبار الثاني		الاختبار الثالث		الاختبار الرابع		الاختبار الخامس		الاختبار السادس	
	b	a	b	a	b	a	b	a	b	a	b	a
1	-.55	1.75	-.41	1.16	-.37	1.64	-1.03	1.39	-.39	1.44	-.34	1.16
2	-.47	1.25	.09	1.21	-.28	1.2	-.65	1.61	-.07	1.45	-.17	1.0
3	2.11	.46	2.03	.46	2.17	.46	1.99	.44	3.15	.38	1.96	.6
4	-.50	1.34	.05	.93	-.23	1.21	-.67	1.27	-.02	.94	-.69	1.17
5	.16	1.59	.51	.90	.53	.98	-.03	1.02	.5	1.2	-.13	1.16
6	-1.61	.66	-1.04	.71	-.89	.9	-1.72	.66	-.55	.71	-1.3	.75
7	.04	1.08	.57	.77	.39	.95	.13	.89	1.11	.56	-.04	1.32
8	.82	.53	1.19	.46	1.43	.47	.52	.69	1.34	.62	.7	.76
9	-.44	1.28	.04	.90	.02	1.46	-.36	.99	.05	.99	-.33	1.53

بلغ متوسط قيم معاملات الصعوبة (b) للاختبارات الستة على الترتيب كمايلي: -0.05 ،0.30 ، 0.31 ،-0.20 ، 0.57 ، -0.04

ويمكن تفسير هذا كتكافؤ في مستوى صعوبة الاختبارات.

بينما بلغت قيم معاملات الارتباط (بيرسون) بين قيم معاملات الصعوبة (b) على الاختبارات الست على الترتيب كما في الجدول 29 .

جدول 29: مصفوفة معاملات الارتباط بين قيم معاملات الصعوبة وفق نظرية الاستجابة للفقرة

الاختبار	2	3	4	5	6
1	0.99	0.98	0.99	0.96	0.98
2		0.98	0.99	0.95	0.96
3			0.97	0.96	0.96
4				0.97	0.96
5					0.95

يتضح من الجدول 29 أن معاملات الارتباط عالية جدا وهذا مؤشر على استقرار معالم الفقرات رغم اختلاف المجموعات.

يبين الجدول 30 معاملات الارتباط بين قيم معاملات التمييز للفقرات المشتركة وفق نظرية الاستجابة للفقرة.

جدول 30: مصفوفة معاملات الارتباط بين قيم معاملات التمييز وفق نظرية الاستجابة للفقرة

الاختبار	2	3	4	5	6
1	0.87	0.85	0.8	0.84	0.73
2		0.89	0.95	0.92	0.57
3			0.77	0.78	0.75
4				0.51	0.9
5					0.43

يتضح من الجدول 30 أن معاملات الارتباط بين قيم معاملات التمييز كانت عالية، وكانت مقبولة بين معاملات التمييز للاختبارين الرابع والخامس، لكنها كانت منخفضة قليلا بين الخامس والسادس.

يتبين من الجدول 31 معالم الفقرات المشتركة وفق نظرية الاستجابة للفقرة، بالإضافة إلى إحصائيات الفقرات نفسها وفق النظرية التقليدية.

جدول 31: معالم الفقرات المشتركة وفق نظرية الاستجابة للفقرة، بالإضافة إلى إحصائيات الفقرات نفسها وفق النظرية التقليدية.

الفقرة	النظرية	الرمز	الأول	الثاني	الثالث	الرابع	الخامس	السادس
1	Irt	b	0.55-	-.41	-.37	-1.03	-.39	-.34
	ctt	p	0.64	0.58	0.59	0.74	0.59	0.57
	Irt	a	1.75	1.18	1.64	1.39	1.44	1.16
	ctt	rbis	0.58	0.44	0.57	0.52	0.51	0.53
2	Irt	b	0.47-	0.09	-.28	-.65	-.07	-.17
	ctt	p	0.60	0.47	0.56	0.66	0.51	0.48
	Irt	a	1.25	1.21	1.2	1.61	1.45	1.0
	ctt	rbis	0.52	0.48	0.47	0.62	0.49	0.25
3	Irt	b	2.11	2.03	2.17	1.99	3.15	1.96
	ctt	p	0.28	0.29	0.28	0.3	0.24	0.25
	Irt	a	0.46	0.46	0.46	0.44	0.38	0.6
	ctt	rbis	0.25	0.22	0.22	0.18	0.13	0.29
4	Irt	b	-.50	0.05	-.23	-.67	-.02	-.69
	ctt	p	0.61	0.49	0.54	0.65	0.5	0.65
	Irt	a	1.34	0.96	1.21	1.27	0.94	1.17
	ctt	rbis	0.5	0.39	0.47	0.52	0.36	0.49
5	Irt	b	0.16	0.51	0.53	-.03	0.5	-.13
	ctt	p	0.44	0.4	0.39	0.5	0.38	0.52
	Irt	a	1.59	0.90	0.98	1.02	1.2	1.16
	ctt	rbis	0.63	0.41	0.47	0.49	0.51	0.55
6	Irt	b	-1.61	-1.04	-.89	-1.72	-.55	-1.3
	ctt	p	0.73	0.66	0.66	0.74	0.59	0.7
	Irt	a	0.67	0.71	0.9	0.66	0.71	0.75
	ctt	rbis	0.3	0.26	0.39	0.3	0.33	0.33
7	Irt	b	0.04	0.57	0.39	0.13	1.12	-.04
	ctt	p	0.48	0.4	0.42	0.47	0.36	0.5
	Irt	a	1.08	0.77	0.95	0.90	0.56	1.32
	ctt	rbis	0.51	0.36	0.49	0.44	0.26	0.6
8	Irt	b	0.82	1.19	1.43	0.52	1.34	0.7
	ctt	p	0.4	0.37	0.34	0.42	0.32	0.38
	Irt	a	0.53	0.46	0.47	0.69	0.62	0.76
	ctt	rbis	0.26	0.18	0.22	0.36	0.28	0.4
9	Irt	b	-.44	0.04	0.02	-.36	0.05	-.33
	ctt	p	0.6	0.5	0.48	0.57	0.49	0.58
	Irt	a	1.28	0.90	1.46	0.99	0.99	1.53
	ctt	rbis	0.53	0.39	0.6	0.45	0.44	0.61

يتضح من الجدول 31 أن قيم معاملات ا لصعوبة (b) وفق نظرية الاستجابة للفقرة، للفقرة الأولى تزايدت من –0.55 في الاختبار الأول إلى –0.41 في الاختبار الثاني، والى –0.37 في الاختبار الثالث، والى –0.39 في الاختبار الخامس، والى -0.34 في الاختبار السادس. بينما تناقصت إلى –1.03 في الاختبار الرابع. وكانت القيم المقدرة لمعامل التمييز(a) للفقرة الأولى 1.75 في الاختبار الأول تناقصت إلى 1.18 في الاختبار الثاني، والى 1.64 في الاختبار الثالث ، والى 1.39 في الاختبار الرابع ، والى 1.44 في الاختبار الخامس ، والى 1.16 في الاختبار السادس. وهذا مؤشر على أن الفقرة كانت صعبة قليلا مع انخفاض قليل في التمييز. ولنفس الفقرة نلاحظ ان قيم معاملات الصعوبة وفق النظرية التقليدية قد تناقصت في جميع الاختبارات، بينما زادت في الاختبار الرابع. وان قيم معاملات التمييز تناقصت في جميع الاختبارات ، وهذا يشير إلى التوافق العالي بين خصائص الفقرة في كل من النظرية الحديثة والتقليدية، كما انه مؤشر على الثبات النسبي لمعالم الفقرة باختلاف العينات.

يتبين من الجدول 32 معاملات الارتباط بين قيم معاملات الصعوبة للفقرات المشتركة وفق النظرية التقليدية، ونظرية الاستجابة للفقرة.

جدول 32: مصفوفة معاملات الارتباط لمعاملات صعوبة الفقرات المشتركة وفق النظرية التقليدية، ونظرية الاستجابة للفقرة.

الاختبار	الحديثة\الاول	الثاني	الثالث	الرابع	الخامس	السادس
التقليدية\الاول	-.96	-.97	-.98	-.96	-.93	-.92
الثاني	-.93	-.97	-.91	-.94	-.88	-.87
الثالث	-.94	-.96	-.96	-.95	0.90	-.87
الرابع	-.93	-.95	-.96	-.96	-.94	-.88
الخامس	-.92	-.96	-.95	-.95	-.94	-.87
السادس	-.97	-.95	-.96	-.95	-.93	-.99

يتضح من الجدول 32 أن معاملات الارتباط عالية جدا وهذا مؤشر على استقرار معالم الفقرة.

يتبين من الجدول 33 معاملات الارتباط بين معاملات التمييز للفقرات المشتركة وفق النظرية التقليدية، ونظرية الاستجابة للفقرة.

جدول 33: ملخص معاملات الارتباط بين معاملات التمييز للفقرات المشتركة وفق النظرية التقليدية ،ونظرية الاستجابة للفقرة.

الاختبار	الحديثة\الاول	الثاني	الثالث	الرابع	الخامس	السادس
التقليدية\الاول	0.96	0.82	0.78	0.75	0.76	0.81
الثاني	0.96	0.84	0.9	0.85	0.69	0.71
الثالث	0.86	0.83	0.94	0.68	0.67	0.89
الرابع	0.82	0.88	0.72	0.95	0.85	0.63
الخامس	0.85	0.85	0.78	0.8	0.95	0.55
السادس	0.77	0.59	0.68	0.65	0.48	0.96

ونظرا لعدم تجانس المجموعات، ولكي نتمكن من قياس الاستقرار النسبي لمعالم الفقرات باختلاف المجموعات، سنأخذ متوسط الصعوبة للفقرات المشتركة في المجموعات الست. ثم نحسب أكبر انحراف عن متوسط الصعوبة واصغر انحرافا عنه. ثم نأخذ الوسط الحسابي للانحرافين. فيكون المدى المقبول لمقارنة الاستقرار النسبي لمعالم الفقرات بين المجموعات ضمن متوسط الصعوبة (±) الوسط الحسابي للانحرافين، والجدول 34 يظهر الاستقرار النسبي لمعالم الفقرات باختلاف المجموعات وفق نظرية الاستجابة للفقرة.

جدول 34:ملخص لمتوسط الانحراف والمدى المقبول للمقارنة والنسبة وفق نظرية الاستجابة للفقرة

رقم الفقرة	متوسط صعوبة الفقرة	متوسط الانحراف عن الوسط	المدى المقبول للمقارنة	النسبة
1	-.52	0.35	0.87- - 0.17-	%83
2	-.26	0.26	0.52- - 0.0	%67
3	2.34	0.6	2.94 – 1.74	%83
4	-.34	0.32	0.02 – 0.66-	%100
5	0.26	0.33	0.63 – 0.07-	%83
6	-1.19	0.64	1.83- - 0.55-	%100
7	0.37	0.54	0.81 – 0.17-	%83
8	-1	0.46	1.46 – 0.54	%83
9	-.17	0.2	0.03 – 0.37-	%67

يتضح من الجدول 34 أن معالم الفقرات 5،3،1،8،7 مستقرة بنسبة 83% باختلاف المجموعات ، ومعالم الفقرتين 9،2 مستقرة بنسبة 67% باختلاف المجموعات ، بينما معالم الفقرتين 6،4 مستقرة تماما.

رابعا: و للإجابة عن السؤال الرابع: ما مدى استقرار معلمة القدرة للفرد (θ) بتعدد الاختبارات المسحوبة من البنك ؟ فقد تم التحقق من ثبات تقدير قدرة الفرد من خلال إجابته على الاختبارات المسحوبة من هذا البنك، ولتحقيق ذلك فقد تم ما يلي:

سحب 5 اختبارات من النموذج الأول، الاختبار الأول (السهل) تكون من 20 فقرة مستوى صعوبتها امتد من -1.63 – 0.36 ، والاختبار الثاني(الصعب) تكون من 20 فقرة مستوى صعوبتها امتد من 0.55 – 3.41 ، والاختبار الثالث (المتوسط) تكون من 20 فقرة مستوى صعوبتها امتد من -1.63 – 3.41 ، أما الاختبار الرابع فقد تكون من (20) فقرة من الفقرات الفردية للنموذج الأول امتد مستوى صعوبتها من -1.62 – 3.38 ،الاختبار الخامس تكون من 20 فقرة من الفقرات الزوجية للنموذج الأول امتد مستوى صعوبتها من -1.63 – 3.41 .يتبين من الجدول 35 المعلومات حول هذه الاختبارات المسحوبة.

جدول35: ملخص المعلومات حول الاختبارات المسحوبة.

مدى صعوبة الفقرات	الاختبار	عدد الفقرات	معامل الثبات(كرونباخ)	عدد أفراد العينة
0.36 – 1.63-	السهل	20	0.86	200
3.41– 0.55	الصعب	20	0.66	200
3.41 – 1.63-	المتوسط	20	0.79	200
3.38– 1.62-	اختبار الفقرات الفردية	20	0.78	200
3.41 – 1.63-	اختبار الفقرات الزوجية	20	0.77	200

طبقت الاختبارات الخمسة على نفس العينة، والتي بلغ حجمها (200) طالب وطالبة، ثم حللت النتائج باستخدام برنامج (bilog-mg)، وللتعرف على قدرات الطلبة على الاختبارات المختلفة، ولمعرفة ما إذا كان هناك فروق ذات دلالة إحصائية بين قدرات الطلبة على هذه الاختبارات، أي أن قدرات الطلبة ليست معتمدة على أي اختبار معين من هذه الاختبارات، تم استخدام تحليل التباين لفريدمان (The Friedman Two-way Analysis of Variance) . والجدول 36 يظهر نتائج التحليل التي تشير إلى عدم وجود فرق ذو دلالة إحصائية. مما يعني توافقا عاليا في ترتيب قدرات الطلبة أو المفحوصين باختلاف الاختبار المسحوب من البنك.

جدول 36: نتائج تحليل التباين لفريدمان

الاختبار	مجموع الرتب	متوسط الرتب	احصائي فريدمان	الاحتمال	درجات الحرية
1	596	3.12			
2	559	2.93	2.718	0.606	4
3	554	2.9			
4	569	2.98			
5	587	3.07			

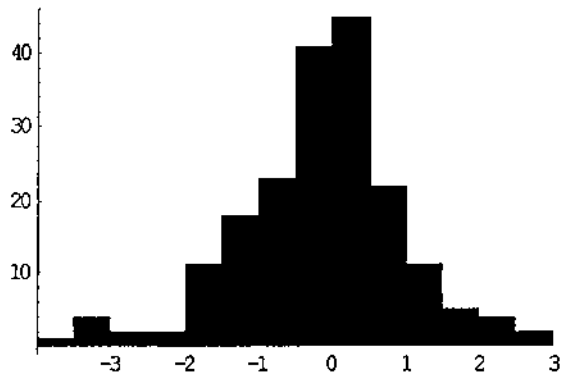
وكإجراء آخر للتحقق من التوافق في ترتيب قدرات الطلبة أو المفحوصين باختلاف الاختبار المسحوب من البنك، تم عمل تدريج خماسي لقدرات الطلبة على كل اختبار من الاختبارات الخمس، ثم حسب الفرق المطلق لهذه القدرات على كل زوج من الاختبارات، وسجلت النسب التي تمثل هذه الفروق، حيث أن الفرق (0) يعني تطابق قدرة المفحوص على الاختبارين، والفرق (1) يعني اختلاف قدرة المفحوص بوحدة واحدة، والفرق (2) يعني اختلاف قدرة المفحوص بوحدتين، والفرق (3) يعني اختلاف قدرة المفحوص بثلاث وحدات. والجدول 37 يبين توزيع هذه الفرو قات والنسب المقابلة لها لكل زوج من الاختبارات

جدول 37: ملخص للفروق بين قدرات المفحوصين ونسبها بين ازواج الاختبارات

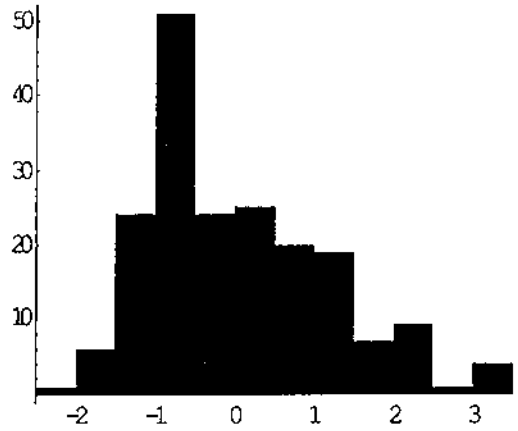
	الفروقات والنسب							
ازواج الاختبارات	0	النسبة	1	النسبة	2	النسبة	3	النسبة
السهل – الصعب	46	0.24	91	0.48	48	0.25	3	0.03
السهل – الوسط	100	0.52	88	0.46	3	0.02	0	0.0
الصعب – الوسط	43	0.23	121	0.63	27	0.14	0	0.0
الفردي – الزوجي	90	0.47	96	0.5	5	0.03	0	0.0

يتبين من الجدول 37 أن الفرق 4 لم يحدث مطلقا، وأن الفرق 3 حصل بنسبة 3%في حالة واحدة بين الاختبار الاول والثاني(الاختبار السهل، الاختبار الصعب)، ولم يحدث لبقية الأزواج الأخرى من الاختبارات. بينما الفرق 2 قد حدث بأعلى نسبة مقدارها 25% بين الاختبار الأول والاختبار الثاني(السهل،والصعب) وبنسبة 14% بين الاختبار الثاني والثالث(الصعب،والوسط). أما بقية الفروق فقد كانت صغيرة ويمكن إهمالها.

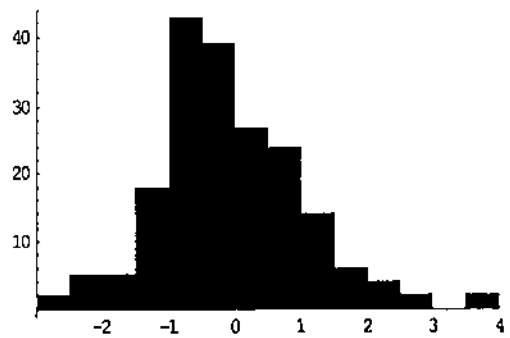
أما في حالة كون الفروق لا تزيد على درجة واحدة فقد تراوحت هذه النسب بين 72% إلى 98%، وهذا مؤشر عال على التوافق في ترتيب قدرات المفحوصين باختلاف الاختبار المستخدم . تمثل الرسومات البيانية 46،47،48،49،50 توزيع قدرات المفحوصين على الاختبارات الخمس المختلفة.



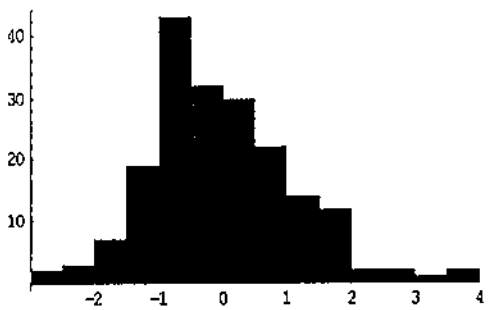
شكل46: التمثيل البياني لقدرات المفحوصين على الاختبار الوسط



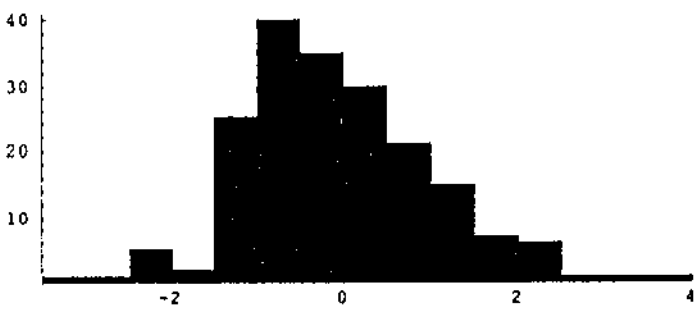
شكل47: التمثيل البياني لقدرات المفحوصين على الاختبار الصعب



شكل48: التمثيل البياني لقدرات المفحوصين على الاختبار السهل



شكل49: التمثيل البياني لقدرات المفحوصين على اختبار الفقرات الفردية



شكل50: التمثيل البياني لقدرات المفحوصين على اختبار الفقرات الزوجية

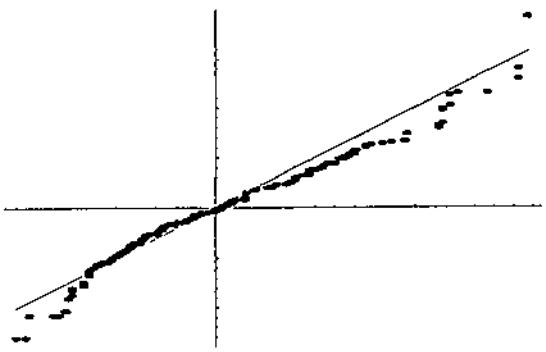
الجدول 38 يبين معاملات الارتباط بين قدرات المفحوصين على الاختبارات الخمس.

جدول 38: ملخص معاملات الارتباط بين قدرات المفحوصين على الاختبارات الخمس

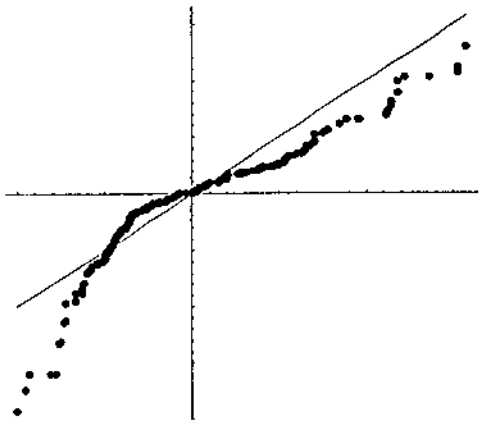
	الصعب	الوسط	الزوجي
السهل	0.54	0.85	_
الصعب		0.68	_
الفردي			0.68

يتبين من الجدول 38 أن معاملات الارتباط كانت عالية بين جميع الاختبارات باستثناء الارتباط بين الاختبار السهل والاختبار الصعب كان متوسطا نسبيا.

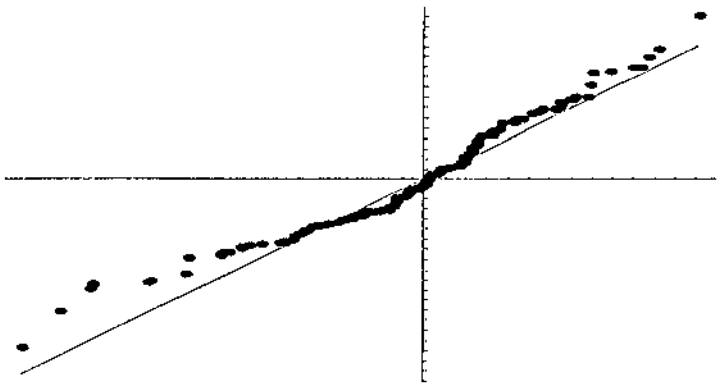
ولفحص مدى تطابق توزيع القدرات على كل زوج من الاختبارات، فانه تم رسم الأشكال البيانية 54,53,52,51, والتي تمثل التوزيعات التكرارية للرتب المئينية لقدرات المفحوصين على كل زوج من الاختبارات.



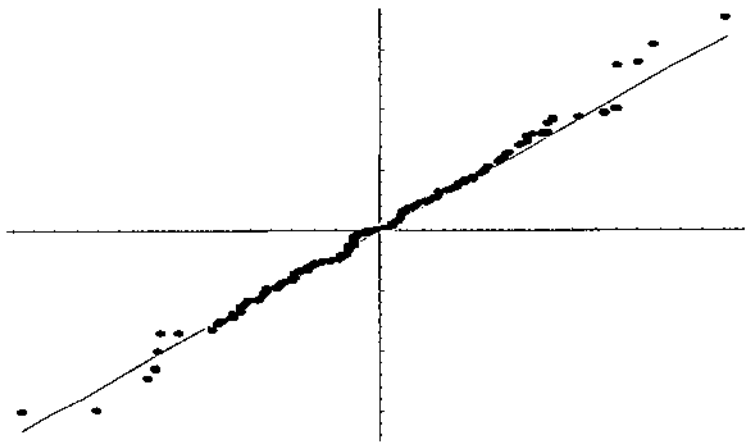
شكل51: التوزيع التكراري للرتب المئينية لقدرات المفحوصين على الاختبار السهل والاختبار الوسط



شكل52: التوزيع التكراري للرتب المئينية لقدرات المفحوصين على الاختبار السهل والاختبار الصعب



شكل53: التوزيع التكراري للرتب المئينية لقدرات
المفحوصين على الاختبار الصعب والاختبار الوسط



شكل54: التوزيع التكراري للرتب المئينية لقدرات
المفحوصين على الاختبار الفردي والاختبار الزوجي

الأشكال 54,53,52,51 تبين التوزيعات التكرارية للرتب المئينية لقدرات المفحوصين على كل زوج من الاختبارات، وذلك بهدف فحص مدى تطابق توزيع القدرات على كل زوج من الاختبارات. يتبين من الشكل (52) الذي يمثل توزيعي الاختبارين السهل والصعب عدم تطابق التوزيعين مع الخط الذي معادلته ص= س بدرجة عالية، وهذا الوضع ربما يكون طبيعيا، أما الأشكال الأخرى التي لم يظهر فيها التطابق بشكل تام حيث يتبين من الأشكال أن هناك عدد قليل من الطلبة ابتعدوا عن الخط في الجهة العليا وهم يمثلون الطلبة المتميزين ذوي القدرات المرتفعة،

ونفس العدد تقريبا من الطلبة ابتعدوا عن الخط في الجهة السفلى وهؤلاء يمثلون الطلبة ذوي القدرات المنخفضة، ومن الأسباب الأخرى التي يتوقع أن تكون قد ساهمت في ابتعاد الطلبة وعدم التطابق هي عدم جديتهم في الأداء، ولجوئهم إلى التخمين، ولجوئهم أحيانا إلى الغش سواء كان فرديا أو جماعيا.

وإذا ما حاولنا التعرف على درجة المطابقة بين قدرات المفحوصين باستخدام المنحنيات التوافقية (harmonic curves)* لمجتمع الدراسة، ومنحنيات جميع أفراد العينة والتي يمكن توضيحها من خلال المعادلات الآتية:

Consider the r^{th} data point $x_r = (x_{r1}, x_{r2},\ldots,x_{rp})$

$r = 1, \ldots, n$

Plot the curve

$F_{xr}(t) = x_{r1}/\sqrt{2} + x_{r2}\sin t + x_{r3}\cos t + x_{r4}\sin 2t + x_{r5}\cos 2t + \ldots$

For each data point x_r, $r = 1,\ldots, n$, other the interval $-\pi \le t \le \pi$

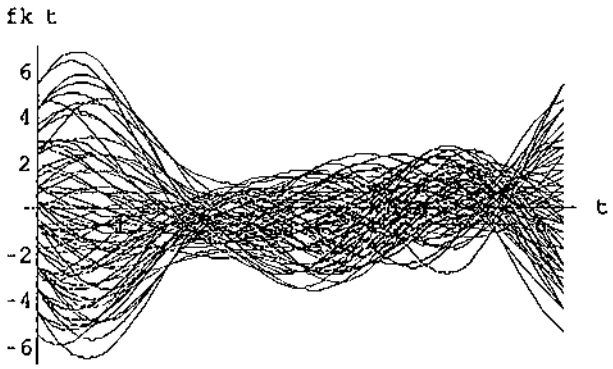
Note : $$\int_{-\pi}^{\pi} [\, f_x(t) - f_y(t)\,]^2 \, dt = \pi \, \| \, x - y \, \|^2$$

Which is proportional to the square of the Euclidean distance between x & y. (Mardia, Kent &Bibby.1975, Andrews, 1972).

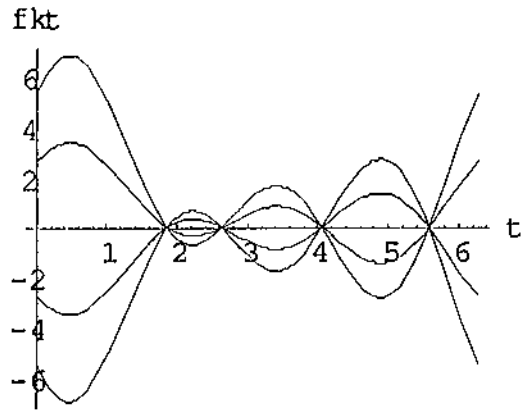
يتبين من الشكل 55 أن منحنيات أفراد العينة ضمن مدى التوزيع للمنحنيات التوافقية شكل 56 في الفترة 0-2، بينما يبدوا أن هناك تضخما طفيفا في الفترة 2- 3 ، وبعض العشوائية البسيطة في الفترة 4-5 ، بينما كانت المنحنيات لأفراد العينة مشابهة للمنحنيات التوافقية في الفترة 5-6 ، وربما تعزى أسباب التضخم هذا والعشوائية إلى عدم جديتهم في الأداء ،والتخمين، ولجوئهم أحيانا إلى الغش سواء كان فرديا أو جماعيا.

---

* من د. عدنان عوض الأستاذ الدكتور بقسم الرياضيات- الجامعة الأردنية



شكل55: منحنيات جميع أفراد العينة



شكل56: المنحنيات التوافقية

وكإجراء ثالث للتحقق من مدى التوافق في ترتيب قدرات المفحوصين باختلاف الاختبار المسحوب من البنك، تم عمل تدريج سباعي لقدرات الطلبة على كل اختبار من الاختبارات الخمس، ثم عملت جداول تقاطعية(cross tabulation) كما في جدول (39) لحالات التوافق لهذه القدرات على كل زوج من الاختبارات، وسجلت النسب التي تمثل هذه الحالات. والجدول 39 يبين توزيع هذه الحالات والنسب المقابلة لها لكل زوج من الاختبارات.

جدول(39):توزيع قدرات الطلبة على الفئات المختلفة للاختبار السهل والاختبار الصعب

الاختبار الصعب	الاختبار السهل $3^{-}-4^{-}$	$2^{-}-3^{-}$	$1^{-}-2^{-}$	$1-1^{-}$	2 - 1	3 - 2	4 - 3	المجموع
$3^{-}-4^{-}$	**1**							1
$2^{-}-3^{-}$		**0**		1				1
$1^{-}-2^{-}$	2	2	**7**	21	1			33
$1-1^{-}$	3	2	20	**86**	7			118
2 - 1			2	18	**4**	1		25
3 - 2				4	3	**3**		10
4 - 3				1	1	2	**0**	4
المجموع	6	4	29	131	16	6	0	192

جدول 40:عدد حالات التوافق بين قدرات المفحوصين على الاختبارات المختلفة ونسبتها

ازواج الاختبارات	حالات التوافق	النسبة
السهل مع الصعب	101	%53
السهل مع الوسط	128	%67
الصعب مع الوسط	124	%65
الفردي مع الزوجي	107	%56

يتبين من الجدول 39 أن النسبة المئوية لحالات التوافق بين قدرات المفحوصين تراوحت بين 53% إلى 67%، وهذا مؤشر جيد على ثبات ترتيب قدرات المفحوصين باختلاف الاختبار المسحوب من البنك.

أما بالنسبة لمعاملات الارتباط بين الاختبارات الخمس فيبينها الجدول 40.

جدول 41: مصفوفة معاملات الارتباط بين الاختبارات الخمس

	الصعب	الوسط	الزوجي
السهل	0.53	0.85	-
الصعب		0.67	-
الوسط			-
الفردي			0.68

يتبين من الجدول 40 ان معاملات الارتباط كانت عالية بين جميع ازواج الاختبارات باستثناء معاملات الارتباط بين الاختبار الصعب وبقية الاختبارات حيث كان متوسطا.

الاستنتاجات والتوصيات

يلاحظ من هذه النتائج والتي أجابت عن أسئلة الدراسة أن هناك درجة عالية من التوافق بين أسلوب النظرية التقليدية وأسلوب نظرية استجابة الفقرة (النموذج الثنائي) في الحكم على الفقرات من حيث الإبقاء عليها أو استبعادها بناء على مؤشرات الصعوبة والتمييز. فقد بلغت نسبة الاتفاق بين الأسلوبين في الحكم على الفقرات (الفقرات المتبقية والفقرات المحذوفة) 90% عند تصنيف صعوبة الفقرات وفق النظرية التقليدية من 0.1 – 0.9 ومعامل التمييز اكبر من 0.1، ووفق نظرية استجابة الفقرة من $^{-}3$ - $^{+}3$ والتمييز أكبر من 0.2 .في حين بلغت نسبة الاتفاق بين الأسلوبين في الحكم على الفقرات 88% عند تصنيف صعوبة الفقرات وفق النظرية التقليدية من 0.2 – 0.8 ومعامل التمييز اكبر من 0.1، ووفق نظرية استجابة الفقرة من $^{-}2.5$ - $^{+}2.5$ والتمييز أكبر من 0.2بينما وصلت نسبة الاتفاق بينهما 85% عند تصنيف صعوبة الفقرات وفق النظرية التقليدية من 0.3 – 0.7 ومعامل التمييز أكبر من 0.1، ووفق نظرية استجابة الفقرة من $^{-}2$ - $^{+}2$ والتمييز اكبر من 0.2 .

كما بينت النتائج أن هناك اتفاقا عاليا بين الأسلوبين في تقدير صعوبة الفقرات وتمييزها، فقد بلغت قيمة معامل ارتباط بيرسون بين معاملات الصعوبة لفقرات الاختبار الأول وفق النظرية التقليدية، ومعاملات الصعوبة لنفس فقرات الاختبار وفق نظرية الاستجابة للفقرة –0.88 وهو ارتباط عال جدا، بينما بلغت قيمة معامل الارتباط بين معاملات التمييز وفق الأسلوبين 0.88 وهو ارتباط عال أيضا.

من هذه النتائج يمكن القول بأن الأسلوبين اتفقا اتفاقا كبيرا في الحكم على الفقرات (الإبقاء عليها أو حذفها)، وكذلك في تقدير معالم الفقرات (الصعوبة والتمييز)، وهذه النتيجة تنسجم ونتيجة دراسة جمحاوي (2000).

وفيما يتعلق بخصائص الاختبار المسحوب من بنك الأسئلة فقد دلت النتائج إلى أن استخدام البرنامج الحاسوبي يسهل وييسر عملية الحصول على اختبار ضمن المواصفات المطلوبة، حيث كانت معالم الفقرات المنتقاة ضمن محكات الصعوبة، والتمييز، ودالة معلومة الاختبار التي حددت في المواصفات المطلوبة، وغطت هذه الفقرات المهارات الرياضية المستهدفة، بالإضافة إلى سهولة معرفة الخطأ المعياري في التقدير، وعلاوة على ذلك فان البرنامج الحاسوبي يوفر إمكانية إضافة أي فقرة جديدة إلى البنك تشترك معه في نفس التدريج، وكذلك يسهل عملية حذف أي فقرة من البنك. وكلما زاد عدد الفقرات المخزنة في البنك كلما سهل عملية سحب الاختبارات المختلفة والمتكافئة. وهذه النتيجة تنسجم ونتيجة دراسة حرزالله (2004)، ودراسة فرجات (2004).

وكذلك دلت نتائج هذه الدراسة إلى درجة عالية من الاستقرار النسبي لمعالم الفقرات باختلاف المجموعات وفق النظرية التقليدية ونظرية الاستجابة للفقرة. حيث بلغت نسبة الاستقرار النسبي لمعالم الفقرتين(2،9) 67%، في حين وصلت نسبة الاستقرار لمعالم الفقرات (1، 3، 5، 8، 7) إلى 83%، بينما كانت معالم الفقرتين (4، 6) مستقرة تماما.

ودعم هذه النتيجة قيم معاملات الارتباط بين متوسطات صعوبة هذه الفقرات وفق النظرية التقليدية ووفق نظرية الاستجابة للفقرة، حيث كانت هذه القيم عالية جدا وشبه تامة. كما أن قيم معاملات الارتباط بين معاملات التمييز لهذه الفقرات وفق الأسلوبين كانت عالية جدا. وهذا مؤشر قوي على اتفاق النموذج الثنائي مع النظرية التقليدية في الكشف عن مدى استقرار معالم الفقرات باختلاف المجموعات.

أما بالنسبة للنتائج المتعلقة بمدى استقرار معلمة القدرة بتعدد الاختبارات المسحوبة من البنك، فقد دلت هذه النتائج إلى أن هناك توافقا عاليا في ترتيب قدرات المفحوصين باختلاف الاختبار المسحوب من البنك، حيث تراوحت النسب بين 72% الى 100% في حالة كون الفروق لا تزيد

عن درجة واحدة بين ترتيب قدرات المفحوصين على الاختبارات المختلفة، وهذا مؤشر عال على ثبات ترتيب قدراتهم. ودعم هذه النتيجة قيم معاملات الارتباط بين قدرات المفحوصين على الاختبارات المختلفة حيث كانت عالية جدا بين جميع الاختبارات.

من هذه النتائج نلمس أن هناك اتفاقا كبيرا بين سلوك النموذج ثنائي المعلمة والمؤشرات التقليدية وربما يفسر هذا لانه الأكثر شبها بالمؤشرات التقليدية في اختيار الفقرات.

وهكذا فقد تم بناء بنك أسئلة للمهارات الرياضية، تحققت فيه الموضوعية، وتدرجت فقراته، ويمكن هذا البنك مستخدمه من سحب الاختبارات المتنوعة، ويسهل عليه إضافة أي فقرة جديدة للبنك أو حذف أي فقرة منه.

وبناء على نتائج هذه الدراسة فان الباحث يوصي بإجراء المزيد من الدراسات:

1) في مجال المقارنة بين نماذج نظرية الاستجابة للفقرة من حيث استقرار معالم الفقرة باختلاف العينات وكذلك استقرار معلمة القدرة باختلاف الاختبار.
2) دراسة اثر حجم العينة وعدد الفقرات المشتركة على دقة وكفاءة ضم اختبارين في تدريج مشترك.
3) إعادة تدريج فقرات البنك وفقا للنموذجين راش والنموذج ثلاثي المعلمة.
4) بناء بنوك أسئلة في مباحث أخرى.
5) أثر إجابات حالات الغش وعدم الجدية على معالم فقرات بنك الأسئلة.

المراجع العربية

الشربيني، أحلام و حجازي، رضا (2002). تقويم تحصيل تلاميذ مرحلة التعليم الأساسي في العلوم باستخدام نظرية السمات الكامنة. ورقة بحث مقدمة إلى المؤتمر العلمي السادس للجمعية المصرية للتربية العلمية " التربية العلمية وثقافة المجتمع ". المركز القومي للامتحانات والتقويم التربوي. القاهرة.

الشريفين، نضال (2003). مدى تحقق معايير الفاعلية في معادلة اختبارين احدهما ثنائي التدريج و الآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس. اطروحة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا :عمان

الفرجات، هشام عقيلة.(2004). بناء بنك أسئلة لمبحث الكيمياء للصف الثاني ثانوي العلمي . رسالة ماجستير غير منشورة، جامعة مؤتة:الكرك.

جمحاوي، إيناس (2000). مقارنة خصائص الفقرات وفق النظرية التقليدية ونظرية استجابة الفقرة في مقياس للقدرة الرياضية. رسالة ماجستير غير منشورة ، جامعة اليرموك : اربد.

حرزالله ، علية محمد.(2004). بناء بنك أسئلة في الرياضيات والتحقق من فاعليته في انتقاء فقرات اختبار محكي المرجع في مستوى امتحان شهادة الدراسة الثانوية العامة في الأردن . أطروحة دكتوراة غير منشورة، جامعة عمان العربية للدراسات العليا:عمان

سليمان الخضري، أمينة كاظم (2000) المشروع العربي للتقويم المقارن لمستويات التحصيل العام، اليونسكو.

عودة، احمد سليمان.(1992). مدى التوافق بين نموذج راش والمؤشرات التقليدية في اختيار فقرات مقياس اتجاه سباعي التدريج.حولية كلية التربية، جامعة الإمارات، العدد(8)، 153-179.

عوده، احمد سليمان (2002). القياس والتقويم في العملية التدريسية. اربد: دار الأمل للنشر والتوزيع.

علام، صلاح الدين محمود. (2002). القياس والتقويم التربوي والنفسي أساسيا ته وتطبيقاته وتوجهاته المعاصرة. القاهرة: دار الفكر العربي.

علام، صلاح الدين محمود. (1986). تطورات معاصرة في القياس النفسي والتربوي، الكويت، مطابع القبس التجارية.

كاظم، أمينة.(1988). مستوى العينة وتدريج بنك الأسئلة باستخدام نموذج راش ، مؤسسة الكويت للتقدم العلمي ، الكويت.

كاظم، أمينة.(2000). اتجاهات معاصرة في بناء بنوك الأسئلة. جامعة عين شمس:القاهرة.

شكري، سيد احمد(2001، كانون أول). العلاقة بين ثلاثة اتجاهات معاصرة في مجال القياس التربوي النفسي ومشكلات تعميمها والإفادة منها. ورقة بحث مقدمة في المؤتمر العربي الأول، المركز القومي للامتحانات والتقويم التربوي، الامتحانات والتقويم التربوي: رؤية مستقبلية. القاهرة.

المراجع الاجنبية

Aaron, R. Douglas & Christian, M. (2000). Issues in the design of the student Assessment Instrument for PISA. Australian Council for Educational research.

Allen, M. J. & Yen, W. N. (1979). Introduction to Measurement Theory. Californai: Cole Publishing Company.

Al-harbi, khaleel abdulrahman, 2003. An empirical investigation of comparability and invariance of classical test theory and item response theory person and item parameter estimates. DAI-A64/02, P.477, Aug2003.Ohio University.

Andrews, D. F. (1972). Plots of high dimensional data. Biometrics, 28, 125-136.

Baker,F.B.(2001). The basics of item response theory.ERIC Clearinghouse on assessment and Evaluation. Retrieved May15,2004 from: www.ericae.net/irt.

Barnard, J. J. (1996,Septemper). In search for equity in educational measurement:raditional versus modern equating methods. paper presented at ASEESA's national conference,prretoria South Africa. Retrieved August 28,2004,from http://www.aseesa-edu.co.za/jeec .htm

Beeston, S. (2002). The ucles EFL item Banking System (1-3). Retrieved January 29,2004 From http://www.Cambridge-efl.org/rs-notes/0002/rs-notes2-4.cfm

Bello, S. Wole, F. and Shola, A. (2000). The development of a prototype item Bank for the Nigerian JAMB Test items. Retrieved October 15,2004, from .www.aqa.org.uk/support /iaea/papers/salim-falayajo;a kindele.

Burghof, K. L. (2001). Assembling an item-bank for computerized linear and adaptive testing in Geography. International Education.2, (4). Retrieved September 10,2004, from http://www.flinders.edu.au/education/iej

Choppin, B. H. (1981). Is education getting better? British Educational Research Journal 7:3-16

Choppin, B. (1979). Testing the questions. The Rasch model and item banking.

Crocker & Algina, J. 1986. Introduction to Classical and modern test theory. Holt Rinehart and Winston, New York.

Cunninghams, S. E. (1979). Mathematical Objectives and test items for Grades one through Eight. ERIC:ED237528.

De Gruijter, D. N. M. & Leo J.Th.van der Kamp(2003). Statistical Test Theory for Education and Psychology.

Embreston, S.E &REISE, S. P. (2000). Item Response Theory for Psychologist, London, Lawrence erlbaum associates, Publishers.

Evans, .J & Masters. G. N. (1986). Banking non-dichotomously scored items. Applied Psychological Measurement 10(4), 355-367.

Falayajo, W. (1998). From a gate keeper to a gateway: University of ibadan, Nigeria Inaugural lectures. September 1998.

Fan, x. (1998). Item Response theory and classical test theory:An empirical comparison of their item/person statistics.Educational and psychological measurement, 58(3), 357-381.

Hambleton,R.K.& Swaminathan,H.(1985).Item Response Theory Principles and Applications, Boston, Klwer Nijhoof Publishing.

Hambleton, R. K. Swaminathan,H.and Rogers,H.J.(1991). Fundamentals of item response theory . Newbury Park CA: Sage Publications. London.

Hathaway, W. (1985). A Regional and local item response theory based test item bank system, Portland Public Schools Dept of Research and Evaluation, Portland, Oregon.

Klein, L. W. & Kolen, M. J. (1985) . Effect of number of common items in common- item equating with non-random groups . Paper presented at the annual meeting of American Educational Research Association, Chicago.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: an overview. Theory into Practice, 41, (4), 212-218. Retrieved January 20,2004,from {EBSCO}.

Lord, F. M. (1980). Applications of Item Response Theory to Practical Testing Problems. Hillsdale NJ:Lawrence Erlbaum Associates.

Mardia, K. Kent, J. and Bibby, J. (1975). Multivariate Analysis, Academic Press, London, Page 20.

McDonald, R. P. (1981). The dimensionality of tests and items ,British Journal of Mathe-ematical and Statistical Psychology , 34, 100-117.

McCallon, E. L. & Schumacker, R. E. (2002). Development and maintaining an item bank. ELM Metrics, Inc.

Millman, J. and Arter, J. A. (1984). Issues in Item Banking . Journal of Educational Meas-urement. 21, 315-330.

Mislevy, R. J.& Bock, R. D. (1990) . Item Analysis and Test Scoring with Binary Logistic Models, Second Edition .

Nenty, H. J. (1991) . Item Banking for continous assessment- Conference of the National Association of Educational Psychologist-ABU,Zaria.March 24-28,1991.

Nitko, A. J. (2001). Educational Assessment 0f Students, third edition, Newjersy, merrill prontice hall.

O`brien, M. L. and Hampilos, J. P. (1988).The Feasibility of Creating an Item Bank from a teacher- made test, using the rash model . Educational and Psychological Measure-ment , 48.

Pelton, T. W. (2002) . The Accuracy of Unidimensionality Measurement Models in the presence of deviations from the underlying Assumptions. Unpublished doctoral disser-tation, Brigham Young University.

Raju, N. S. Edwards, J. E. & Osberg, D. W. (1983). The effect of anchor test size inverti-cal equating with rasch and the three-parameter models. paper presented at the annual meeting of the National Council on Measurement in Education, Montreal.

Robi, L.& Patricia, E. (1996). Predicating academic : An application of youths universal scale for evades , online Ericir syr.Eduy.

Robitaille, F. and O, shea, T. (1983). The development of an item bank in mathematics using the rasch model, Canadian Journal of Education.8 (1), 57-70.

Rodriguez, A. J. (1997)."The dangerous discourse of Invisibility: A critique of the Natio-nal Science Education Standards" Journal of research in science Teaching, 34, no.1, 19-37.

Rudner, L. M. (1998). Item Banking. Practical Assessment, Research & Evaluation , *6(4)* . Retrieved May 2,2004, from http://ericae.net/pare/getvn.asp?v=6&n=4

Squires, P. (2003). Concept paper on an item bank approach to testing. Retrieved April1, 2004,from: www. applied skills.com.

Stage, C. (1998). A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory. A Study of the SweSAT Subtest WORD. (Educational Measurement). Umea University, Department of Educational Measurement .Retrieved April10,2004 from: www.cepchile.cl

Valverde, G. A. and Shmidt, W. H. (2000). Grater expectations learning from other nations in the quest for world-class standards in USA school mathematics and science, Journal of curriculum studies, 32(5), 651-687.

Wang, J. (1998)."Comparative study of student science Achievement betweenUnited States and China ", Journals of research in science teaching, 35(3), 329-336 .

Ward, A.W. & Murray-ward, M.(1994). Guidelines for the development of item banks. Educational Measurement Issues and Practice, 13(1), 34-39

Wingersky, M. S. & Lord, F. M. (1984).An investigation of methods for reducing sampling error in certain IRT procedures. Applied Psychological Measurement. 8, pp.347-364.

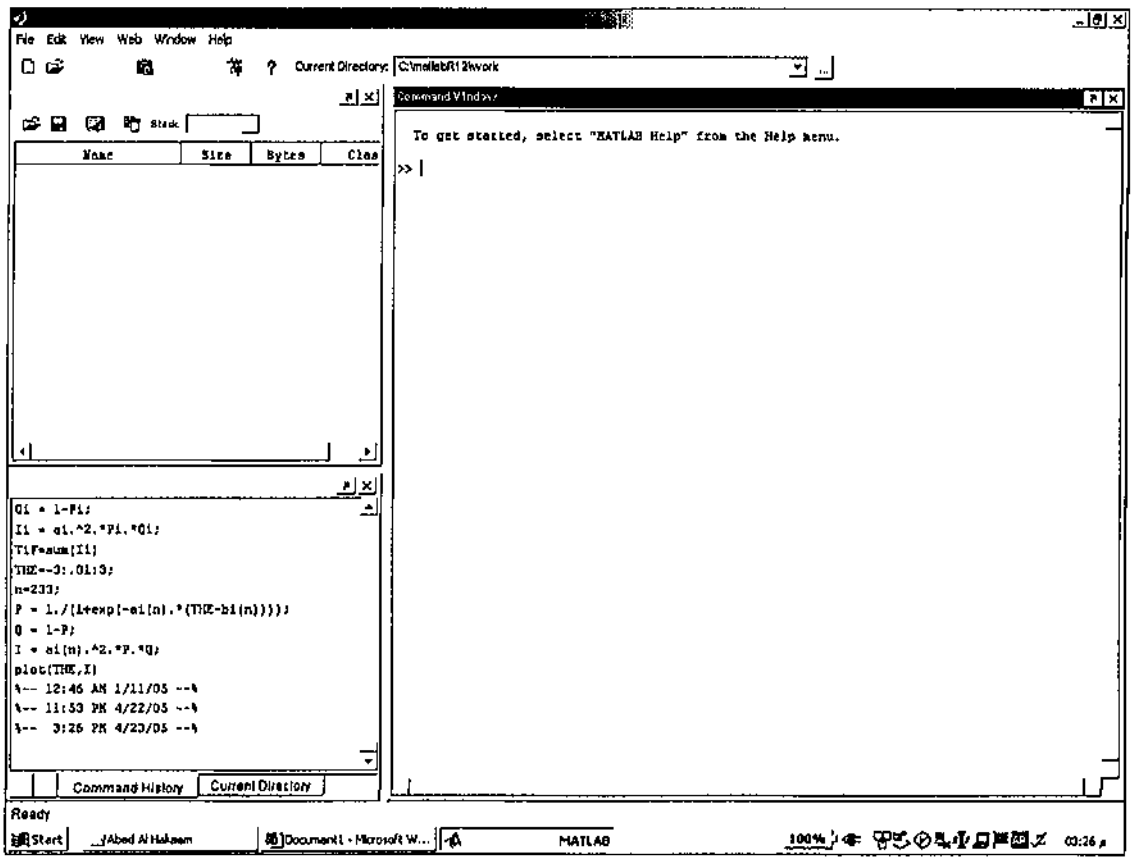
Wright, B. D. & Stone, M. H. (1979). Best Test Design. Chicago: MESA Press.

Wright, B. D. and Bell, S. R. (1984). Item Banks: what, why, how. Journal of Educational Measurement 21(4), 331-345. Retrieved November 20,2003, from http://209.41.153/ Mem043.htm

Wright, B. D. (1999). Fundamental measurement for psychology. The new rules of measurement's. Albertson and S.L. Hershberger.Mahwah NJ, Lawrence Erlbaum Associates.

الملحق (أ)

1) واجهة التطبيق لبرنامج الماتلاب والبرنامج المستخدم لتدريج فقرات الاختبارات الست. ويحتوي هذا البرنامج على مجموعة كبيرة من الأوامر الخاصة به والمبنية على لغة (C / C++) ولغة (Fortran) ويستخدم لإجراء جميع العمليات الرياضية والهندسية.



البرنامج المستخدم لتدريج الاختبارين الأول والثاني باعتبار الاختبار الأول مرجعا. ولباقي الاختبارات يتم استبدال Test2 برقم الاختبار المراد تدريجه علما بأن البيانات للاختبارات جميعها محفوظة في ملف – M (M-Files) في الماتلاب.

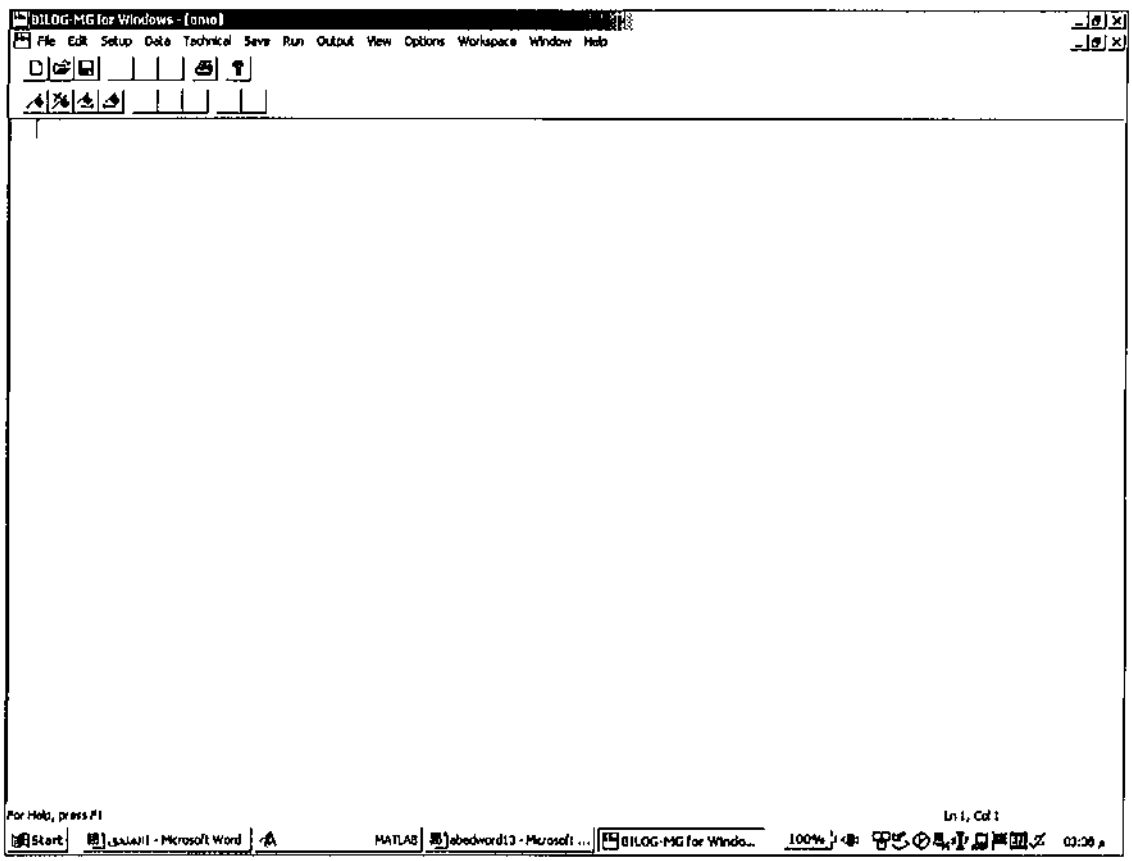
```
load Test1
N1= Test1(:,1);
a1j=Test1(:,3);
b1j=Test1(:,4);
%Common ITEMS of Test 1 are in the same order : 08  09  15  06  07  10  11  12  13  14
%Common ITEMS of Test 3 are in the same order : 28  29  35  26  27  30  31  32  33  34
%Common ITEMS of Test 4 are in the same order : 09  10  16  07  08  11  12  13  14  15
%Common ITEMS of Test 5 are in the same order : 29  30  45  27  28  40  41  42  43  44
%Common ITEMS of Test 6 are in the same order : 47  48  12  45  46  07  08  09  10  11
```

```
% number of common Items m=10 with omitted items m =9.
m=9;
b1cj =[ b1j(9) ; b1j(15) ; b1j(6) ; b1j(7) ; b1j(10) ; b1j(11) ; b1j(12) ; b1j(13) ; b1j(14) ];
a1cj =[ a1j(9) ; a1j(15) ; a1j(6) ; a1j(7) ; a1j(10) ; a1j(11) ; a1j(12) ; a1j(13) ; a1j(14) ];
b1cav=sum(b1cj)/m;
a1cav=exp(sum(log(a1cj))/m);
% ITEM 8 is Missing
load Test2
N2= Test2(:,1);
a2j=Test2(:,3);
b2j=Test2(:,4);
%Common ITEMS of Test 2 are in the same order : 25 26 46 39 40 41 42 43 44 45
b2cj =[ b2j(25) ; b2j(45) ; b2j(38) ; b2j(39) ; b2j(40) ; b2j(41) ; b2j(42) ; b2j(43) ;
b2j(44) ];
a2cj =[ a2j(25) ; a2j(45) ; a2j(38) ; a2j(39) ; a2j(40) ; a2j(41) ; a2j(42) ; a2j(43) ; a2j(44)
];
b2cav=sum(b2cj)/m;
a2cav=exp(sum(log(a2cj))/m);
r=sqrt(a1cav/a2cav);
a1star=a1j/r;
a2star=r*a2j;
d=0.5*(r*b1cav - b2cav/r);
b1star = r*b1j - d;
b2star = b2j/r + d;
R1=[ N1 a1star b1star];
R2=[ N2 a2star b2star];
% Item Number =R (:,1)
% Discrimination a= R (:,2)
% Difficulty b = R (:,3)
R1(:,1)
R1(:,2)
R1(:,3)
R2(:,1)
R2(:,2)
R2(:,3)
```

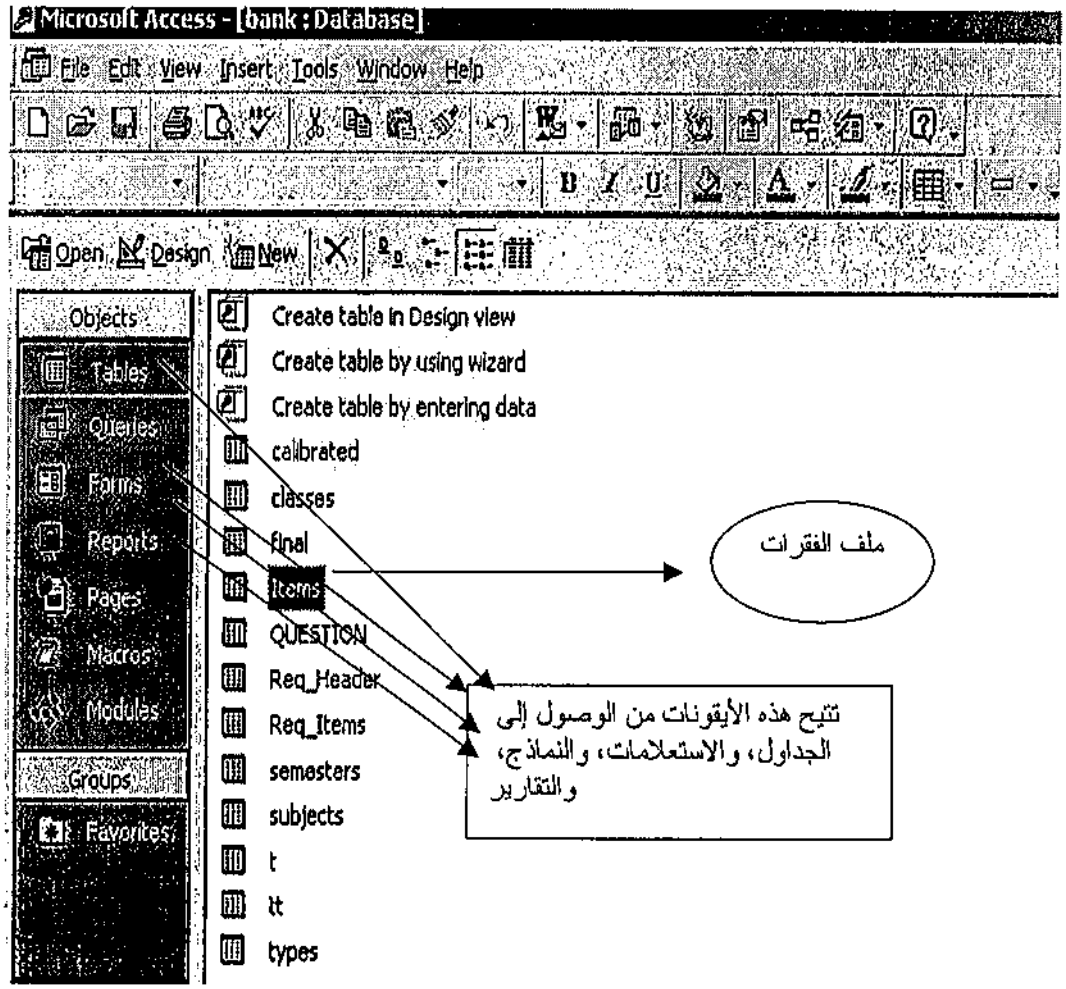
2) واجهة التطبيق لبرنامج (Bilog-Mg)

برنامج جاهز يحتوي على مجموعة من الأوامر التي تقوم بإجراء عمليات المعالجة الإحصائية المختلفة بالطرق التقليدية والحديثة.

بعد إدخال البيانات على برنامج (SPSS) وحفظها على نظام أل (ASCI) وإرسالها إلى برنامج أل (Bilog-Mg) وإجراء الخطوات اللازمة لمعالجة هذه البيانات إحصائيا باستخدام النظريتين: التقليدية والحديثة. ونتائج المعالجات طويلة بالنسبة للبحث، ومتوفرة لدى الباحث.



3) واجهة قاعدة البيانات bank:database حيث تم تصميمها باستخدام برنامج(Microsoft Acssess) لتخزين الفقرات واستدعائها. و تم تخزين جميع الفقرات التي تم تدريجها مع معالمها الاحصائية، وهي محفوظة ومتوفرة لدى الباحث.



الملحق ب: النموذج الاختباري الأول

1) إذا علمت أن كلاً من : □ ، △ تدل على أعداد وكان □ + △ = 54

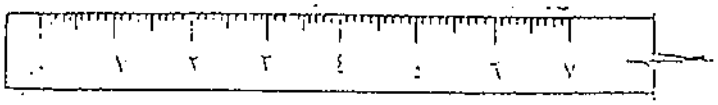
فان قيمة □ + △ + 10 هي

أ) 54 ب) 74 ج) 44 د) 64

2) قدّر ناتج 415 - 193 لأقرب مئة:

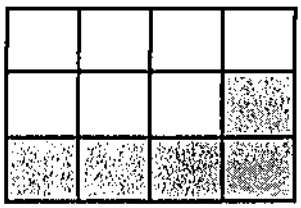
أ) 230 ب) 200 ج) 300 د) 322

3) باستخدام مسطرة مدرجة بالسنتمترات كما في الشكل يمكنك القياس لأقرب :



أ) مليمتر ب) نصف مليمتر ج) سنتمتر د) نصف سنتمتر

4) الكسر الذي تمثله المنطقة المظللة في المستطيل المجاور يساوي تقريباً الكسر الذي تمثله المنطقة المظللة في :



أ) ب) ج) د)

5)أي الأعداد (0,525 ، 0,25 ، 0,5 ، 0,125) هو الأصغر ؟

أ) 0.525 ب) 0.25 ج) 0.5 د) 0.125

6) حل المعادلة 2س + 5 = 17 هو :

أ) 2 ب) 10 ج) 6 د) 12

7) اذا كان س= 3 ، ص= -1 ، فما قيمة المقدار 2س - 3ص ؟

أ) 9 ب) 45 ج) 3 د) 18

8)انفق احمد1 / 3 ما معه من نقود ثمن قميص و2 /3 ما تبقى معه ثمن بنطا ل. وتبقى معه 6دنانير. كم مجموع ما أنفقه؟

أ) 27 دينار ب)24 دينار ج)21 دينار د)18 دينار

9) خزان به3 /5 سعته ماء.إذا علمت أن الخزان بحاجة إلى30م3 ليمتلىء كليا.أوجد سعة الخزان الكلية؟

أ)25م3 ب)45م3 ج)64م3 د)75م3

10) أي الأعداد (621 ، 622 ، 612 ، 226) يقبل القسمة على العددين (2) ، (3) معاً ؟

أ) 621 ب) 622 ج) 612 د) 226

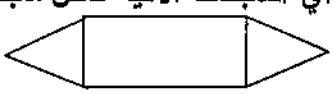
11) أي مما يلي يعتبر إجراءً صحيحاً لإيجاد ناتج $\frac{1}{7} - \frac{1}{5}$ ؟

أ) $\frac{1}{7} - \frac{1}{5} = \frac{1}{5-7}$ ب) $\frac{1}{5} - \frac{1}{7} = \frac{1}{5-7}$

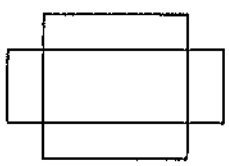
ج) $\frac{1}{7} - \frac{1}{5} = \frac{7-5}{5 \times 7}$ د) $\frac{1}{5} - \frac{1}{7} = \frac{1-1}{5 \times 7}$

12) أي الشبكات الآتية تمثل شبكة هرم رباعي قائم ؟

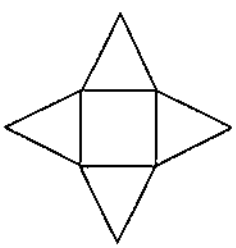
أ)



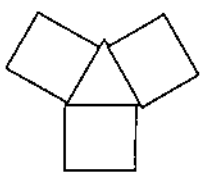
ب)



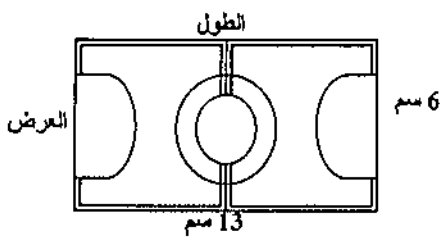
ج)



د)



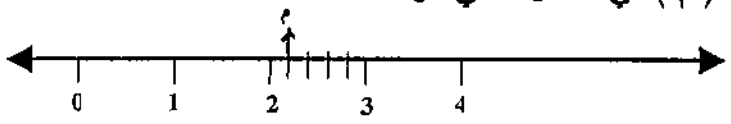
13) رسم ملعب كرة سلة باستخدام مقياس الرسم 1سم : 3م كما في الشكل ، ما الطول الحقيقي للملعب بالأمتار ؟



أ) 39م ب) 18م

ج) 57م د) 13م

14) العدد الذي تمثله النقطة (م) في الشكل الآتي هو :



أ) $\frac{1}{5}$ ب) $2\frac{1}{4}$ ج) $2\frac{1}{10}$ د) $2\frac{1}{5}$

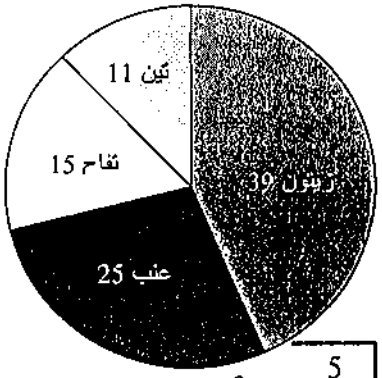
15)العدد الذي إذا قسم على (– 8) كان الناتج (16) هو :

أ) –128 ب) –2 ج) 2 د) 128

16) اشترى عامل طلاء (الدهان) (8) علب طلاء بسعر 30 دينار للعلبة. لم يستخدم 25% من علب الطلاء التي اشتراها لأنها جفت قبل استعمالها ،ما هو المبلغ الذي خسره العامل في هذه الحالة؟

أ) 60 دينار ب) 180 دينار ج) 240 دينار د) 360 دينار

17) قطعة أرض مساحتها 90 دونماً مزروعة بالأشجار كما هو ممثل بالرسم المجاور ، قياس الزاوية المركزية لقطاع أشجار التفاح هو :



أ) 75 ب) 30

ج) 50 د) 60

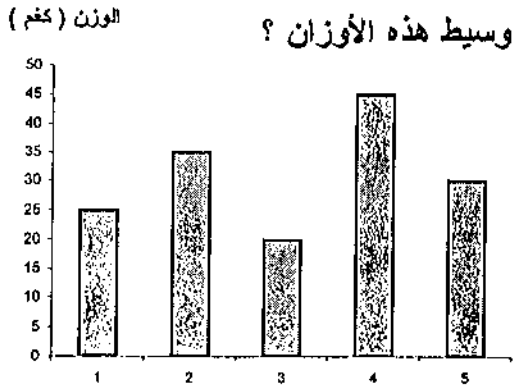
18) إذا كان $\sqrt{15} = 3.87$ تقريباً فما القيمة التقريبية لـ $\sqrt{\frac{5}{3}}$ ؟

أ) 0.2 ب) 0.4 ج) 1.29 د) 3.66

19) ما عدد علب الكبريت التي ابعادها 1سم، 3سم ، 5سم والتي يمكن وضعها في صندوق مكعب طول ضلعه 30سم؟

أ) 60 علبة ب) 100 علبة ج) 1800 علبة د) 2400 علبة

20) مثلت بالأعمدة أوزان خمسة أطفال كما في الشكل، ما وسيط هذه الأوزان ؟



أ) 31 ب) 30

ج) 45 د) 20

21) إذا كان $\frac{3}{8} = \frac{س}{32}$ فان قيمة س =

أ) 27 ب) 4 ج) 88 د) 12

22) $جا^2 30 + جتا^2 30$ =

أ)4 ب)3 ج)2 د)1

23) ينص قانون هوك على ان تمدد الزنبرك يتناسب مع القوة المؤثرة .اذا اثرت قوة مقدارها 150 نيوتن تمدد الزنبرك 8 سم.ما مقدار التمدد في الزنبرك اذا بلغت القوة 300 نيوتن ؟

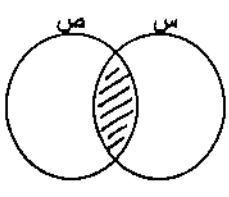
أ) 24سم ب) 21 سم ج) 16 سم د) 61 سم

24) إذا كان $س^2 + ص^2 = 15$ ، س ص = 5 ، فما قيمة $(س + ص)^2$

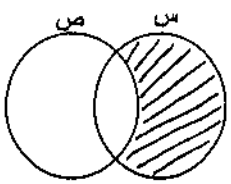
أ) 20 ب) 15 ج) 10 د) 25

25) أي الأشكال الآتية ظللت فيها المجموعة س- ص ؟

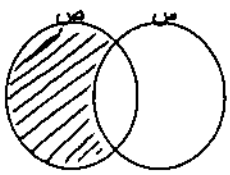
أ)



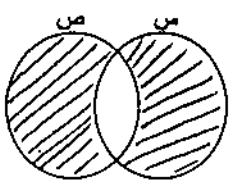
ب)



ج)



د)



26) 2.935 - 5,72 =

أ) 2.785 ب) 2.163 ج) 2.363 د) 2.637

27) في الشكل المجاور ، كم يصبح سعر القطعة بعد الخصم ؟

أ) 6 دنانير ب) 5.5 دينار

ج) 1.5 دينار د) 5 دنانير

28) قيمة (س) في الجدول المجاور هي

$(-3)^0$	$(-3)^1$	$(-3)^2$	$(-3)^3$	$(-3)^4$	$(-3)^5$
1	3-	9	27-	س	243-

أ) 81- ب) 81

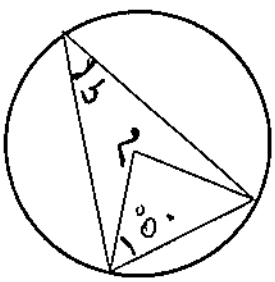
ج) 12 د) 12-

29) تُرتب الأعداد 3 ، 4- ، 2- تنازلياً كما يلي :

أ) 4- ، 2- ، 3 ب) 2- ، 4- ، 3

ج) 3 ، 4- ، 2- د) 3 ، 2- ، 4-

30) في الشكل المجاور م مركز الدائرة ، إن قياس الزاوية س هو :



أ) 40° ب) 80°

ج) 50° د) 25°

31) إذا كانت س ، س+2 ، س+4 تمثل أطوال أضلاع مثلث فما قيمة أو (قيم) س

أ) س < 2 ب) س > 2 ج) س = 2 د) 0 < س < 4

32) الجدول الآتي يمثل توزيع طلاب الصف العاشر في إحدى المدارس وفق هواياتهم ما الكسر الدال على الطلبة الذين لا يفضلون استخدام الحاسوب ؟

الهواية	مطالعة	سباحة	استخدام الحاسوب
عدد الطلاب	8	10	16

أ) $\frac{16}{18}$ ب) $\frac{18}{16}$ ج) $\frac{16}{34}$ د) $\frac{18}{34}$

33) ما قيمة $1 - 5 \times (-2)$ ؟

أ) -11 ب) 8 ج) -8 د) 11

34) ما قيمة س التي تحقق المعادلة $7^{\text{س}-2} = 49$ ؟

أ) 4 ب) س^{-2} ج) -2 د) صفر

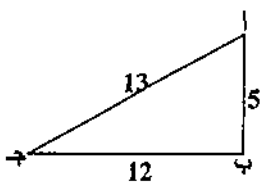
35) سلك رفيع طوله 20سم ، عمل منه مستطيل عرضه 4سم، ما طول هذا المستطيل ؟

أ) 5سم ب)6سم ج) 12سم د)16سم

36) إذا علمت أن كتلة كيس الإسمنت (50 كيلوغراما) فما عدد أكياس الإسمنت التي كتلتها (900 كيلوغرام)

أ)10 ب) 15 ج)18 د) 21

37) قيمة جا جـ من الشكل المجاور هي :



أ) $\frac{5}{13}$ ب) $\frac{12}{13}$

ج) $\frac{5}{12}$ د) $\frac{13}{5}$

38) إذا كان | س - 15 | = 4 س، حيث | | هي القيمة المطلقة فإن س =

أ) 3 ب) -5 ج) 3 ، 5 د) -5 ، 3

39) إذا كانت س هي مجموعة الأعداد الزوجية المحصورة بين العددين 8 ، 24 فأي العبارات الآتية صحيحة ؟

أ) 9 ∉ س ب) 6 ∉ س ج) 10 ∉ س د) 4 ∈ س

40) ما هو عدد المكعبات بارتفاع 1سم التي نحتاجها لتشكيل مكعب واحد بارتفاع 5 سم؟

أ) 625 ب) 125 ج) 25 د) 5

41) إذا كان حاصل ضرب 5 أعداد صحيحة أقل من أصغرهم، ما هو الشرط الواجب توفره في هذه الأعداد لتحقيق ذلك؟

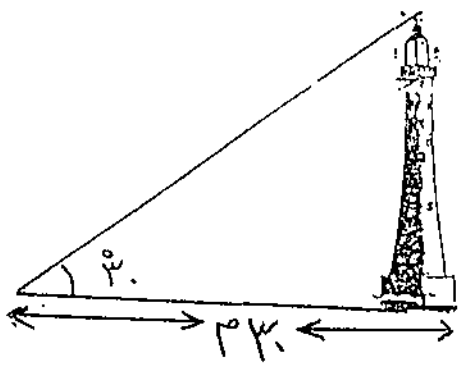
أ) أن يكون أحد الأعداد سالبا أو تكون جميعها سالب

ب) أن يكون اثنين من الأعداد سالبين والثالث موجبا

ج) أن لا تكون إشارة الثلاث أعداد متشابهة.

د) لا يوجد أعداد يمكن أن تحقق ذلك

42) رُصدت زاوية ارتفاع قمة برج من نقطة تبعد 30م عن قاعدة البرج فكانت 30° ، جد ارتفاع البرج ؟



أ) $10\sqrt{3}$ م ب) $30\sqrt{3}$ م

ج) $100\sqrt{3}$ م د) $5\sqrt{3}$ م

43)يبين الجدول المجاور علامات 3 طلاب في أربعة اختبارات في الرياضيات ما متوسط علامات حسن ؟

الطالب	الاختبار الأول	الاختبار الثاني	الاختبار الثالث	الاختبار الرابع
سعيد	7	8	10	6
حسن	6	7	9	10
أحمد	8	10	10	9

أ) 6 ب) 7

ج) 8 د) 9

44) **12.5 % في صورة كسر عشري تساوي :**

أ) 1.25 ب) 0.125 ج) 12.5 د) 0.0125

45) **أي مما يلي يساوي (470 × 998) + (470 × 2)**

أ) 470 × 1000 ب) 472 × 998

ج) 940 × 998 د) 470 × 998 × 2

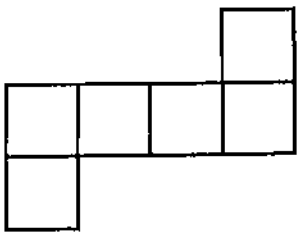
46) **0.1 × 0.1 =**

أ) 0.01 ب) 1 ج) 0.1 د) 0.2

47) **أبسط صورة للمقدار ($\sqrt{75}$ - $\sqrt{12}$) هي :**

أ) 3 $\sqrt{3}$ ب) 3 ج) $\sqrt{63}$ د) 21 $\sqrt{3}$

48) **في الشكل المجاور إذا كانت مساحة كل مربع 25 سم^2 فما محيط الشكل ؟**



أ) 120 سم ب) 150 سم

ج) 70 سم د) 60 سم

49) **توفي رجل عن زوجة وولد وبنت وترك 24000 ديناراً ، فما نصيب البنت من التركة**

(إرشاد : الزوجة تأخذ $\frac{1}{8}$ التركة وللذكر مثل حظ الانثيين) .

أ) 3000 دينار ب) 6000 دينار ج) 7000 دينار د) 1400 دينار

ملاحظة :

تطلب المعلومات من الباحث على العنوان التالي:

amuhaidat@yahoo.com

أو الاتصال على الارقام التالية:

منزل:02/7300489

جوال:0777746346

Muhaidat, Abdel-hakim Ali. Establishing a bank of items for Mathematical skills at the end of the Basic Stage. (A suggested paradigm), Yarmouk University, 2005 (Supervisor: prof. Ahmad Suleiman Audeh).

This study aimed to establish a mathematical skills Item Bank at the end of the basic stage " A suggested Paradigm" and to test the efficacy of the tests that could be withdrawn from it. To achieve the purpose of the study, a table of specification was constructed; it included skills such as, mathematical thinking, problem solving, mathematcal reasoning, modeling, and symboling, and the levels of recalling, understanding, application, analysis, evaluation, and creativity. According to the table of specification an item pool of the multiple choice type was produced.

The items were divided into six test models, each of about (45 – 49) items, and 9 items of each were common among all. The sum total of all items was 237 items. The study sample consisted of 2679 10^{th} graders (males and females) from 15 schools.

The study subjects were asked to answer the items of one of six models in normal schooling conditions. The responses were analyzed by using (bilog-mg, SPSS) to estimate item statistics following the classical test theory.

Item parameters were estimated using maximum likelihood function and reiteration, following the item response theory (two-parameter model). Also, these programmes, which were used to verify the stability of examinees abilities on different tests, were withdrawn from the Item Bank whereas 5 different tests in terms of difficulty degree were used.

The statistical analysis indicated high agreement between the two methods in estimating the difficulty degree and item discrimination. The results pointed out a 90% agreement for the difficulty degree between (0.1 – 0.9) and a discrimination of more than (0.1) according to the classical test theory and at a range of difficulty ($^{-}3$ - $^{+}3$) and a discrimination of more than (0.2) according to the item response theory. This percentage came to 85% at a range of difficulty degree between (0.3 – 0.7) and a discrimination of more than (0.2) according to the item response theory.

The results of verifying the extent of stability of item parameters from different groups indicated a high degree of stability of item parameters. Also, the results of verifying

the extent of stability of ability parameters using different tests withdrawn from the item bank, indicated high agreement in ordering ranking of the testees abilities using different forms.

After the construction of a database, 233 items were stored, in addition to the following information: type of mathematical skill, item structure, correct response, difficulty, the domain that could be tested, the item discrimination, and the item information function.

The computerized programme, developed to select test items, was used to illustrate the test process and the withdrawing mechanism according to the purpose of the test.

(**Key words**: Item Banking, Item response theory, mathematical skills, two-parameter logistic model, Common items).