برمجة التقويم و بنوك الأسئلة

أ . د/ عبد الرحيم بخيت عبد الرحيم
أستاذ ورئيس قسم الصحة النفسية
كلية التربية – جامعة المنيا

د . هانم حامد ياركندي
أستاذ مشارك علم النفس
كلية البنات – مكة المكرمة

مقدمة :

يـزداد الـتعاون وثوقاً بين البحوث العلمية الميدانية وبين التكنولوجيا العصرية يوماً بعد يوم ، فالبحوث العلمية في سعيها نحو الأصالة والتجديد ترتكز على العديد من المكتشفات التكـنولوجية المعاصرة ، ففي منتصف هذا القرن تحولت العديد من المجتمعات الصناعية إلى " *مجـتمعات معلوماتية* " Information Society امتداد للثورة الصناعية التي بلغت ذروتها فـي بدايـة هـذا القـرن ، وتعـتمد المجـتمعات المعـلوماتية على الصناعات المعلوماتية Information Industry والـتي تتـناول المعـلومات مـن حيث جمعها وتبويبها وتحليلها وتنسيقها ثم تسويقها وبيعها لجمهور المستهلكين في صورة بضائع أو خدمات علمية ، وسوف يشهد العصر القادم سيطرة الدول المتقدمة معلوماتياً ، والدليل الواضح على ذلك ما حدث في قطاعـات العامـلين مـن تحـولات وتغيـرات جذرية أو ما يطلق عليه " التحول المهني " Vocational Transference فـي المجـتمعات الحديـثة وخاصـة فـي الولايات المتحدة الأمـريكية (مكـتب التربية العربي لدول الخليج ،١٥، ١٩٩٤) والتأثير الاقتصادي للمعلومات ليـس ناتجـاً عـن نشوء الصناعات المعلوماتية ، ولكن من مدى تأثير تلك المعلومات على إنـتاجية الأفـراد ، ومـن مـدى تأثيرها في القدرات العقلية وخاصة الذكاء وأنماط التفكير وتطورهـا مـن *التقـليدية المكتبية إلى الابتكارية الإنتاجية لإيجاد فيض من الإبداع الفكري الإنسـاني* ، وفـي وقـت تتـنافى فيـه الـتغيرات الاجتماعية والقيمية وتتهاوى العديد من اسـتراتيجيات الـتربية ، وتتجه أصابع الاتهام نحو المدرسة والجامعة ويشتد النقد إلى المنهج وإلى عملية التقويم Evaluation .

يذكـر " بلوم " Bloom 1987 في كتابة " انغلاق العقل الأمريكي " تلك الجامعات العظيمـة الـتي اسـتطاعت أن تحطم الذرة وتتوصل لشفاء العديد من الأمراض الخطيرة لم تسـتطع أن تنـتج مـنهجاً للتربية العامة ، وتخلف العملية التعليمية خاصة والتربية عامة في

العديد من الدول العربية ، تمثل جزءاً من تخلف غالبية المؤسسات التي مازالت تعيش عصر أهل الثقة وليس أهل الخبرة ، وأكبر دليل على ذلك أن اقتناء الكثير من الأطفال والشباب للحاسبات الشخصية في المنزل وفي خارج المدرسة ذلك قبل أن تفكر مؤسسات التعليم النظامية في إدخاله ، وتلك ظاهرة تستحق التسجيل : *ظاهرة تخلف المدرسة عن تقديم خدمة تعليمية مرغوبة* (وليم عبيد ،٧٥،٩٩،١٩٩٥) ، فالكمبيوتر في المدرسة كمادة تعليمية أمر مرغوب ومطلوب ولا يحتاج إلى جدال أو نقاش كوسيط لتعلم مقررات دراسية أو يستخدم في عملية التقويم الدراسي .

ساير الانتشار التعليمي في الكم هبوط في النوعية والكيفية ، قد يعود ذلك إلى *الانفجار المعرفي* الذي ميز القرن الحالي مع *التسارع في نمو المعارف الإنسانية* والزيادة في إعداد الطلاب زيادة فاقت إمكانيات التعليم ، وقد يرى البعض أن سبب هبوط المستوى الكيفي هو زيادة إعداد الطلاب،ونادى البعض الآخر بتطبيق ما توصلت إليه الدراسات والبحوث التربوية والنفسية وأكدته التجارب في شتى المجالات وهو أن" ***إصلاح التقويم هو المدخل لإصلاح التعليم*** " وهو عنق الزجاجة التي يستطيع من يتحكم فيها أن يتحكم فيما يخرج منها ويشكله وفقاً لما يريد ، خاصة وأن اختبارات التقويم الحالية لا تتناول في الأغلب غير مستوى استدعاء المعرفة Recall مع افتقارها للنظرة الشمولية ، وتستغرق الامتحانات وقتاً وجهداً طويلاً ، ويقتصر التقويم على منتصف أو نهاية العام ، وكذلك اعتماد التقويم على مقارنة الطالب بغيره من الطلاب ، والاقتصار على تقويم المعلم للطالب ، إن التفسير مطلوب، يذكر " فؤاد أبو حطب " (١٩٩٥،٢٣٥) أن هناك أنواع لأسئلة الامتحانات المرتبطة بالجانب المعرفي وهي : أسئلة إنتاج الاستجابة وآلتي تتطلب أن يعطي الطالب إجابته على السؤال أو المشكلة وهذه الفئة لها عدة أنواع وفقاً لمستويات البساطة والتبسيط أسئلة الإنتاج البسيط (الإجابة القصيرة) ، أسئلة الإنتاج المركب (المقال الطويل) ، وهناك أسئلة انتقاء الاستجابة : حيث يقدم للطالب عدة حلول لاختبار الحل السليم وتصنف إلى:أسئلة البديلين،وأسئلة المزاوجة ، وأسئلة الاختبار من متعدد ، ويرى " رشدي فام "(١٩٩٥-١٩٣) أن التقويم موضوع متشعب ومن الضروري أن يعتمد على أكثر من معيار ينسب إليه موقع الطالب وهذه المعايير هي:توقع الفرد بالنسبة للآخرين(المعيار السيكومتري)،موقع الفرد بالنسبة لنفسه من وقت لآخر،موقع الفرد بالنسبة للمحك(المعيار الأديومتري) ، معايير أخرى

كموقعه بالنسبة لمتطلباته وما يتطلع إليه .

تـــري " أميـــنة كاظم " (١٩٩٥ – ٢٤٧) أن هناك ضرورة لتحري الموضوعية في قيـــاس الظواهـــر والمـــتغيرات ، وكـــانت نشـــأة الاختبارات الموضوعية هي بداية الاهتمام بموضوعية التقويم التربوي وأزداد هذا الاهتمام بظهور فكرة " ***مخزن أو مستودع الأسئلة*** " Item pool الـــذي يضم عدداً كبيراً من الأسئلة الموضوعية الخاصة بمقرر دراسي معين ، ومع تطور الحاسبات الإلكترونية وإمكانية حفظ الأسئلة بدأ ظهور ***بنوك الأسئلة*** Item Bank وأصـــبح الإقـــبال متزايد على إنشاء تلك البنوك في الدول الغربية في المرحلة الابتدائية حتى المرحـــلة الجامعية في الوقت الذي تحتاج فيه هذه البنوك إلى عقلية علمية متميزة قادرة على التعامل مع المتغيرات التي تؤثر في دقة وإحكام تلك البنوك .

ويذكـــر " صلاح علام " (١٩٩٤،٨) أن النظرة الضيقة والمحدودة لمفهوم التحصيل الدراسي والفلسفة التربوية التي تهتم بالفروق بين الأفراد والموازنة بينهم وما يترتب على ذلك من أساليب قياس وتقويم للتحصيل تركز على ما اختزنه المعلم في ذهنه من معلومات محددة ومحدودة لم تعد تناسب المتطلبات المستقبلية للتربية ولحتياجاتها المتغيرة ونحن على مشارف القـــرن الحادي والعشرين الذي يتميز بالتفجر المعرفي والتقدم التكنولوجي وثورة المعلومات . فالـــتربية ســـوف تقـــود الـــتغيير وتكـــون في مقدمته ، وسوف تلعب دوراً رئيساً في تنمية المجتمعات المستقبلية التي نطمح إليها . كما أن القضايا والمشكلات التربوية الحيوية المتعلقة بالـــتحديات الـــثلاثة المرتبطة والمتفاعلة التي تواجهها البشرية وهي : التنمية ، والحفاظ على البيئة وحمايتها والسلام، سوف تصبح المحور الرئيسي للسياسات الدولية والقومية .

الأهداف المستقبلية للتربية يصعب تحققها في إطار المفهوم السائد للتحصيل الدراسي والأســـاليب التقـــليدية المتبعة في التقويم . فهي تتطلب مفهوماً للتحصيل أكثر مرونة ورحابة بحيـــث يتـــناول مكتســـبات المتعـــلم وإنجازاته المتعددة التي تمكنه من خلال العمل الجماعي الـــتعاوني من الإسهام الوظيفي الفاعل في رفع مستوى معيشته وتطوير مجتمعه ، كما تتطلب منهجية جديدة للتقويم وأدوات قياس متنوعة.

تعريف بنوك الأسئلة : **Item Banking**

تضم بنوك الأسئلة في صورتها المعاصرة *كماً هائلاً من الأسئلة الموضوعية الجيدة التي تتدرج جميعها بوحدة قياس واحدة تشكل معاً مقياساً واحداً له صفر واحد مشترك وتدرج واحد مشترك يغطي مدى واسعاً من المتغير المراد قياسه* (أحد المتغيرات) ويمكن أن نسحب

مـن هـذه البـنوك مجموعاً من الأسئلة لاستخدامها في عمل الاختبارات التي تحقق أهداف القيـاس ، كمـا يمكـن أن نضـيف إليها أسئلة جديدة تتدرج معها على نفس تدريجها الواحد المشـترك ، ويـبدأ بـنك الأسئلة بدمج بنود اختبارين في تدريج واحد ينتهي بشبكة من عدة اختبارات يمكن أن تغطي المدى الواسع من المتغير المقاس (أمينة كاظم ١٩٩٥ ، ٢٤٧) .

محـتويات بـنوك الأسئلة تتضمن أعداد ضخمة من الأسئلة(المفردات أو البنود Items) والـتي تـم تدريجها على أعداد كبيرة من الأفراد ، حيث من المتعذر على كل فرد الإجابة على كل سؤال من هذه الأسئلة العديدة ، وبالتالي فتكوين بنك الأسئلة يعتمد على دمج مجموعـات مختلفة من الأسئلة في تدريج واحد بصفر واحد مشترك ، حيث يبدأ تدريج عدد كـبير مـن الأسئلة على متغير واحد عن طريق مجموعة من البنود المشتركة بين اختبارين مختـلفين في مستوى صعوبتهما حيث تكون بمثابة رابطة تعمل على ربطهما في تدريج واحد مشـترك وعند تكرار ذلك الرباط بين أكثر من اختبار تتكون أعداد هائلة من البنود في تدريج واحـد مشـترك ، ويـتطور هـذا الـرباط بين الاختبارات المختلفة فتتكون شبكة Net من الارتـباطات ، ثـم تشكل هذه الشبكة من الارتباطات Links نسيجاً Web من الاختبارات ويمكن إنشاء بنوك الأسئلة في أي فرع من فروع المعرفة وكذلك لأي مستوى من مستوياتها (1988 - Jahia) (1996 Stocking) (1995 Hemker) .

تصميم أنظمة بنوك الأسئلة :

يرى خبراء القياس التربوي المعاصر من أمثال (Robert 1990) (Wood1989) (Mengle 1992) أن تكنولوجية بنوك الأسئلة سوف تسهم إسهاماً مؤثراً في تقويم تحصيلات الطـلاب فـي المستقبل . فبناء أنظمة اختباريه يمكن الإفادة منها في قياس التحصيل الدراسي يتطـلب تأسيس بنوك للأسئلة ،ويتضمن بنك الأسئلة مجموعة كبيرة من المفردات أو الأسئلة الاختيارية التي تتميز بخصائص سيكومترية محددة وتقيس جميعها سمة كامنة معينة . ونقصد بالسـمة الكامـنة أي خاصة تميز الفرد مثل قدرته على حل تمارين هندسية أو قدرته اللفظية وهكذا .

لكي يحقق هذا النظام الاختياري أهدافه في مختلف المواقف التعليمية ينبغي أن يتميز بالاتسـاق والفاعلية والمرونة ، ويكون متاحاً للمعلمين والمتهمين ببناء الاختبارات الصفية ، وعـلى الـرغم من أن النظرية الكلاسيكية في القياس قد استخدمت في الماضي لتصميم بنوك

لأسئلة ، إلا أن الإفادة منها كانت محدودة ، ويرجع ذلك إلى اعتماد تحليل مفردات أو أسئلة البنك في إطار هذه النظرية على خصائص عينة الأفراد المستخدمة في هذا التحليل ، لذلك اتجه خبراء القياس إلى مواجهة هذه المشكلة اعتماداً على نماذج السمات الكامنة حيث أن بارامتراتها (صعوبة المفردات وتمييزها بين المستويات المختلفة للقدرة أو السمة المراد قياسها) تتميز بخاصية عدم التباين Invariance بمفهومها الإحصائي غير أن الإفادة من هذه النماذج لا تتأتى إلا إذا روعيت الفروض والأسس التي يستند إليها كل منها في البيانات الاختبارية . وقد اهتمت بعض الدول مثل امريكا وإنجلترا بتنفيذ أنظمة بنوك الأسئلة وفقاً لمفاهيم وأسس هذه النماذج وبخاصة نموذج " راش " ، وأصبحت هذه الأنظمة مكونة رئيسية من مكونات أنظمتها التعليمية ، وتتمثل أهم أهداف واستخدامات بنوك الأسئلة كما تذكر" أمينة كاظم " (١٩٩٥،٢٤٨) في:

١. تحقيق الموضوعية في تقويم التحصيل الدراسي للطلاب وهذا هو الهدف الأساسي لبنوك الأسئلة ، ويعتمد تحقيق هذا الهدف على مجموعة من العوامل التي يمكن أن تؤثر على بناء وتكوين بنك الأسئلة مثل فلسفة القياس ، ومدى الدقة في صياغة الأسئلة ، وتحليلها وتحديد مواصفاتها الإحصائية والتربوية .
٢. إتاحة الفرصة لقياس العمليات العليا التي طالما أغفلت في تقدير تحصيل الطلاب وذلك نظراً لاتساع المدى الذي يمكن أن يغطيه البنك من حيث محتوى المقرر أو من حيث المستويات المعرفية المختلفة .
٣. توفير وقت وجهد المعلم في بناء الاختبارات التحصيلية وذلك بحصوله على اختبارات جيدة من بنوك الأسئلة طبقاً لمواصفات وأهداف المقرر الدراسي .
٤. تخفيض عوامل القلق والرهبة والتوتر المصاحبة لموقف الامتحان ، وذلك باستخدام عينات من بنود البنك للتقويم المستمر أثناء العام الدراسي حتى يألف الطلاب هذا النوع من الأسئلة الموضوعية دون التأثير على سرية الامتحانات .
٥. مرونة القياس حيث يسهل تشكيل أي اختبار في أي وقت بسرعة وسهوله .
٦. سرعة الحصول على صور متكافئة من الاختبارات تبعاً لمواصفات البنود المكونة للبنك .
٧. إمكانية المقارنة الموضوعية بين مستوى تحصيل الطلاب مهما اختلف الاختبارات المستخدمة ومهما أختلفت مستويات المجموعات التي ينتمون إليها ومهما اختلفت الأعوام ،

طالمـا أن البـنود قـد تم تدريجها ومعايرتها ووصفها في بنك واحد وفقاً لعملية التعادل Equating .

٨. المسـاعدة فـي تطوير العملية التعليمية وذلك لما تحظى به بنوك الأسئلة من مميزات من أهمهـا (١) السـرعة والسـهولة الـتي يمكن بها – باستخدام الحاسب الآلي – تكوين اختـبارات موضوعية جيدة مسحوبة من بنك الأسئلة ، يتيح فرصة التقويم المستمر. (٢) تدرج صعوبة البنود ومعايرتها لقياس قدرة الطلاب على التحصيل الدراسي تتيح الفرصة لاسـتخدام صـور مختصـرة مـن الاختبارات تضم عددا قليلاً من البنود تحقق أهداف المقيـاس، (٣) سـهولة تكويـن الاختبارات المسحوبة من بنك الأسئلة وسهولة طباعتها وتصحيحها باستخدام الحاسب الآلي،واستخدامها للأعداد الكبيرة والأماكن البعيدة .

والأهـداف والاستخدامات السابقة تحقق متطلبات القياس الموضوعي والتي أوردها " رايـت " Wright (١٩٧٩-٩) والتي تتضمن البنود الصادقة والتي تعرف المتغير موضوع القيـاس تعـريفاً إجـرائياً ، وكذلك صدق التدرج لهذه البنود بحيث يمكن تمثيل هذا المتغير بواسـطة مستقيم ، وأيضاً صدق أنماط الاستجابات لتحديد مواقع الأفراد ، وكذلك التوافق بين تـدرج الأفراد على الاختبار ومميزات البنود بحيث تؤدي إلى تقدير مستوى الفرد دون التقيد باختـبار معين ، وأيضاً القياسات الخطية التي يمكن استخدامها لدراسة النمو أو للمقارنة بين المجموعات .

البنك الدولي للأسئلة :

يعتبر مشروع البنك الدولي للأسئلة من التطورات المساهمة في مجال التقويم ، وقد بدأته الجمعية الدولية لقياس التحصيل IEA بالتعاون مع مركز دراسات التقويم التابع لجامعة كاليفورنيا CSE ، وقد اشترك في هذا المشروع الذي بدأ تنفيذه عام ١٩٨٠ عشرين دولة من الـدول الأعضـاء فـي هذه الجمعية ، واهتم المشروع ببناء مركز دولي منبع وشبكة بنوك الأسـئلة **Item Bank Network** تسـاعد في تحسين أساليب وعمليات تقويم التحصيل في الـدول الأعضاء ، ويرى المسئولين عن هذا المركز أن بناء بنك دولي للأسئلة متصل ببنوك الأسـئلة الخاصـة بكـل دولـة سـوف يؤدي إلى ابتكار نظم تصنيفية للأسئلة وخصائصها الإحصـائية على موازين دولية موحدة **International Common Scale** بحيث يسهل على جميع الدول استخدامها ، ويتميز بالمرونة الكافية بحيث يتفق والأنظمة التعليمية في هذه

الـدول ،كما يساعد في الحصول على معلومات عن جوانب تعليمية معينة في الدول المختلفة (صـلاح عـلام ١٩٩٤، Nissan 1996)، وقـد تمكن مركز الاختبارات التربوية الامريكية بولاية واشنطن بالولايات المتحدة الأمريكية من إعداد (١٨٠٠٠) اختبار لبنوك الأسئلة تغطى العديد من المناهج وتبدأ من المرحلة الابتدائية حتى المرحلة الجامعة ، مما يدعو للعمل الجاد بمصر للاعداد المماثل .

الأساس النظري لاتجاهات القياس و بنوك الأسئلة :

١. لا خـلاف بين علماء التربية وعلم النفس من أن عملية التقويم حالياً فقدت الكثير من موضـوعيتها ، ومع تزايد الخطورة تضافرت الجهود التي أسفرت عن قيام علماء القياس والـتقويم ببناء وصناعة أحدث وأدق الاختبارات وظهرت نظرية السمات الكامنة Item Response Theory (Latent Trait Theory) IRT- LTT ومن خلال هذه النظرية بدأت فكرة معادلة درجات الاختبارات ومنها فكرة بنوك الأسئلة Item Banks ، وبذلك تم التغلب على معظم مشكلات عمليات القياس والتقويم التقليدية (Hambleton, 1987: 13) ، ويذكـر " رايت " (Wright 1979,1982) أن بنوك الأسئلة تمثل تجمع ضخم من البنود الممثلة لمتغير ما والتي لا تختلف فيما بينها إلا من حيث مستوى صعوبتها ولها تدريج واحد وصفر مشترك وانتظام مواقع المفردات حول صفر التدريج يعبر عن مستوى صعوبتها فقط (Vecrkamp 1994 & Bergerstrom 1992a) .

أولاً : الاتجاهات التقليدية للقياس : Classical Measurement Theory

١ . (القياس جماعي المرجع Norm Referenced Tests (NRT's

هـو القياس النفسي (السيكومتري)ويهتم **بتباين الفروق الفردية للأداء بين الأفراد** ، والذي أسس على ما جاءت به النظرية السيكومترية التقليدية (الكلاسيكية) وهو ما يطلق عليه: القيـاس جماعي-المرجع Group - Referenced أو القياس مرجعي-المعيار - Normal Referenced Measurement ، وهـذا الـنوع مـن القياس يهتم في جوهره على تصنيف الأفـراد بحسـب مركـزهم النسبي بين أقرانهم في القدرات المختلفة وتقدير الفروق الفردية للأداء،وإخضـاع هـذه الفروق للمنحنى الاعتدالي،وبذلك فالقياس جماعي المرجع يسعى إلى تقديـر مسـتوى أداء الفـرد بالنسبة لأداء أقرانه أو جماعته المعيارية في القدرة التي يقيسها الاختبار .

٢. (القياس محكي المرجع Criterion - Referenced Tests (CRT's

هو القياس التربوي (الأديومتري) ويهتم بقياس التحصيل لدى الأفراد **لتحديد مدى التوصل إلى مستوى أو محك الإتقان** لتحصيل أحد المقررات الدراسية ، وهو الاهتمام المألوف للاختبارات التربوية ، وقد أطلق عليه : القياس محكي – المرجع Criterion Referenced وهذا النوع من القياس يهدف إلى موازنة أداء الفرد بمحك،أو بمستوى أداء مطلق ومحدد جيداً (نطاق من السلوك Behavioral Domain)أو مستوى كفاية معين دون الحاجة إلى موازنة أداء الفرد بأداء الأفراد الآخرين،وبهذا فالقياس محكي المرجع يعتمد في جوهره على تقدير مستوى أداء الفرد بالنسبة لمحك أو مستوى الإتقان المطلوب لأحد المقررات التي يدرسها.

لا يقتصر الاختلاف بين هذين الأسلوبين(CRT's , NRT's)على هدف القياس فقط وإنما يتعدى هذا ويمتد إلى الاختلاف في بناء أو تصميم الاختبار نفسه ، فهما يختلفان من حيث اختيار المفردات ومستويات صعوبتها وقدرتها على التمييز بين الأفراد،وكذا من حيث شروط صدقها وثباتها ومعايرتها،فأفضل المفردات من وجهة القياس جماعي المرجع هي الأقدر على التمييز،والتي يتساوى فيها كل من معاملي صعوبة المفردة وسهولتها،أي تبلغ (٠,٥)، وأسوأ المفردات تلك التي لا تستطيع التمييز بين الأفراد،وأفضل المفردات من وجهة نظر القياس محكي المرجع تلك المفردة التي يكون مستوى سهولتها مساوياً للصفر قبل البدء في البرنامج التعليمي ثم يصبح مساوياً للواحد الصحيح بعد تعلم البرنامج،حيث يستطيع جميع الأفراد الإجابة على هذه المفردات على عكس ما كانوا عليه قبل دراستهم للبرنامج (Bergstrom 1992 b' Koch 1995)

٣. (الاختبارات هدفيه المرجع Objective Referenced Tests (ORT's

تستخدم في قياس مجموعة من **الأهداف التعليمية التي تم صياغتها إجرائياً وسلوكياً**،ومن أهم مميزات احتوائها على عدد قليل نسبياً من المفردات التي تمثل كل المفردات التي تقيس نفس المجموعة من الأهداف،وتصنف تحت الاختبارات محكية المرجع لأنها لا تعتمد في تفسيرها للدرجات على معايير المجموعة التي ينتمي إليها الفرد (Thorndike *et al* , 1991, 214)

٤. (الاختبارات المرجعية النطاق Domain Referenced Tests (DRT's

تختلف هذه النوعية من الاختبارات عما سواها من الاختبارات في أنها تشتمل على

عينة عشوائية أو عشوائية طبقية من **المفردات التي يتم سحبها من النطاق الشامل Domain من المفردات التي أعدت لقياس مهام Tasks محددة** ، ويصنف هذا النوع تحت الإختبارات محكية المرجع (Thorndike , 1991:185)

ويرجع الفضل الأول في هذا المجال إلى العالم الأمريكي روبرت جليزر R . Glaser ١٩٦٣ حيث أثار اهتمام علماء القياس بالقياس محكي – المرجع نظراً لأهميته البالغة في قياس نواتج التعلم ، وفي عام ١٩٦٩ تبنى جيمس بوبام Popham فكرة هذا النوع من القياس وتزعم حركته بهدف تحقيق المثالية في القياس السلوكي والموضوعية في تفسير النتائج وتلبية المطالب والاحتياجات الاجتماعية المتزايدة والتي كان من أهمها *تغيير الهدف الأساسي للعملية التعليمية من الانتقاء والتصنيف والغربلة إلى التعميمية* ، ومن ثم لا يكون موقع الفرد المتدني من الجماعة سبباً في حرمانه من حقه في التعليم واستمراره ، مع تدارك الهبوط المستمر في مستويات الأفراد نتيجة الاقتصار على اتخاذ الجماعة المعيارية مرجعاً عند تفسير الدرجات ، وشيوع مفهوم التعليم حتى *الإتقان أو التمكن* Mastery Learning بحيث ينسحب هذا على الفرد والجماعة على السواء ، والاستجابة للصيحات المتزايدة والنداءات المستمرة للتعليم المستمر ، والتعليم الذاتي ومدارس بلا رسوب ، ولا مزيد من الفشل ولا مزيد من المعايير ، واستخدام تكنولوجيا التعليم في التعلم ، مما يدعم تقدم الفرد بالنسبة لأهداف سلوكية معينة بغض النظر عن موقعه ومكانته النسبية من جماعته المعيارية ، حتى أصبح الشعار السائد : *القياس والتقويم كمدخل لإصلاح التعليم ورفع مستوى كفاية كل متعلم أولاً بأول بديلاً عن القياس والتقويم من أجل التصنيف والفرز والانتقاء والغربلة* (Ward 1994)(Hambilton ,Jones 1994)

ومن ثم يبدو الخطر عندما تستخدم المقاييس جماعية – المرجع لتقدير مدى إتقان التحصيل ، وعندما تستخدم المقاييس محكية – المرجع CRT للتمييز بين الأفراد ، ولهذا برز اتجاه جديد معاصر لا يهتم بهذا التصنيف بل يهدف إلى التوصل إلى مفهوم جديد يحل مشكلة الموضوعية في القياس السلوكي ويحققها بصورة تختلف عما هو سائد ومألوف طبقاً للنظرية السيكومترية الكلاسيكية ، ويعتمد هذا الاتجاه الجديد على النماذج الاحتمالية ، هذه النماذج ما هي إلا نماذج سيكومترية جديدة غير تقليدية تفي وتحقق الموضوعية في القياس النفسي والتربوي ، يطلق عليها اسم " **نظرية أو نماذج السمات الكامنة** " Latent Trait-Models أو " نماذج الاستجابة للمفردة أو البند " Item Response ، وهذه النماذج لها مزاياها الهامة

فـــي مواجهة الكثير من المشكلات التي عجزت عن حلها النماذج التقليدية ، ومعالجة القصور والعيـــوب الـــتي لحقـــت بتلك النماذج الكلاسيكية التي سادت وسيطرت على الفكر التربوي والســـيكولوجي فـــترة طويـــلة تـــزيد عـــن نصـــف قرن من الزمن ، ومن أهم هذه النماذج " **نمـــوذج راش** " Rasch Model وهـــذا *النموذج يقترب في تحقيقه للموضوعية في القياس السلوكي من المقاييس الفيزيائية* .

ثانياً : الاتجاهات الحديثة للقياس : Modern Measurement Theory

ســـعياً وراء تحقيـــق الموضوعية في القياس النفسي والتربوي ، والتغلب على أوجه القصـــور والعيـــوب ســـالفة الذكـــر ،فقد تكاتفت جهود علماء القياس بقيادة" بنجامين رايت " B.Wright وتوصـــلوا إلى نظـــرية أو نماذج السمات الكامنة (LTM) هذه النماذج تفترض إمكانيـــة التنـــبؤ بأداء الأفراد ولإمكانية تفسير هذا الأداء على الاختبارات في ضوء خاصية معينة تسمى السمات Traits .

نظرية السمات الكامنة (الاستجابة المفردة) : Item Response Theory

فالاختـــبارات النفســـية والتربوية بعامة تفترض أن **هناك سمات أو خصائص معينة يشـــترك فيها جميع الأفراد ولكنهم يختلفون في مقدارها** ، وبالرغم من أن هذه السمات غير مـــنظورة ولا يمكـــن قياسها بطريقة مباشرة إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحـــظ للفرد المتمثل في استجاباته على مفردات الاختبار ، وهذا ما يبرر تسميتها بالسمات الكامـــنة ، فالســـمة الـــتي تكمن وراء استجابة الفرد على مفردات اختبار لفظي ، تختلف عن السمة التي تكمن وراء استجابته على مفردات اختبار مكاني أو عددي . ولكن يمكن أن تكمن سمة واحدة وراء استجابته على مفردات اختبارين مختلفين متعلقين بنفس المحتوى . لذلك فإن الهـــدف الأساســـي لكل من النماذج الكلاسيكية ونماذج السمات الكامنة هو تحديد العلاقة بين اســـتجابات الأفـــراد على اختبار معين ، والسمة الكامنة وراء هذه الاستجابات والأمر الأكثر أهميـــة في القياس النفسي والتربوي بعامة هو تحديد مقدار السمات الكامنة وراء أداء الأفراد على الاختبارات المختلفة والاستفادة منها في تفسير النتائج والتنبؤ بسلوكهم في مواقف مماثلة ، واتخاذ قرارات معينة بشأنهم في ضوء هذا التقدير الكمي للسمات (صلاح علام ، ١٩٨٧ ، ٢)

وتعتـــبر هـــذه النماذج في جوهرها دوالاً رياضياً تحدد العلاقة بين أداء الفرد على مفـــردات الاختـــبار (الاستجابات الملاحظة المباشرة) وبين السمات أو القدرات التي تكمن

وراء هـذا الأداء ، وتـتميز هـذه الـدوال بأنهـا احـتمالية Probabilistic وليست حتمية Deterministic ، وهـذا يعـني أن العلاقـة التي تحددها هذه الدوال الرياضية تتحدد وفقاً لـنظرية الاحـتمالات ،ويشير لورد (Lord 1993)إلى أن نماذج السمات الكامنة تقوم على افتراضيين أساسيين هما :

١. أن تكون جميع مفردات الاختبار متجانسة،أي أن جميعها تقيس سمة واحدة أو قدرة واحدة فقط

٢. أن تكـون إجابة الفرد على إحدى مفردات الاختبار مستقلة إحصائياً عن إجابته على بقية المفـردات فـي الاختـبار ، ولكن يجب ملاحظة أن هذا الافتراض لا يعني عدم ارتباط مفـردات الاختـبار بالنسبة للعينة الكلية للأفراد ، ومن أهم هذه النماذج التي حازت على اهتمام علماء القياس النفسي والتربوي :

(Hambleton & Cook 1977,Bergstorm 1992b,a & Lord 1980)

١ - " **نموذج بيرنبوم** " Birnbaum's Model : يعرف باسم"النموذج اللوغاريتمي ثنائي البارامتر " Two-Parameter Logistic Model

٢- **النموذج اللوغاريتمي ثلاثي البارامتر**Three Parameter Logistic Model وهذا النموذج يعد امتداداً للنموذج اللوغاريتمي ثنائي البارامتر .

٣ - **نموذج " راش "** Rasch Model : يعرف بإسم " النموذج اللوغاريتمي أحادي البارامـتر " One-Parameter Logistic Model وهذا النموذج أيضاً يعد حالة خاصة من نموذج بيرنبوم ثنائي البارامتر .

وعـلى الـرغم مـن اخـتلاف الـنماذج الثلاثة في تركيبها البارامتري الذي يفسر اسـتجابات الأفراد على مفردات الاختبارات إلا أنها تتفق في مجموعة الافتراضات التي يجب أن تـتحقق في البيانات قبل استخدام أي منها في بناء هذه الاختبارات أو المقاييس ، ومن بين هذه الافتراضات:

(١) أن تكون مفردات الاختبار من النوع الذي يتطلب الإجابة بصح أو خطأ ، ولا يعتمد إجابة أي منها على إجابة المفردات الأخرى .

(٢) يفضل أن يكون الاختبار من نوع اختبارات القوة التي لا تلعب السرعة دوراً في الإجابة عن مفرداته .

(٣) لا تتأثر الإجابة بعامل التخمين .

(٤) أن تقيس جميع مفردات الاختبار سمة أو قدرة أحادية البعد

وفي حالة عدم تحقيق أي من هذه الافتراضات فإن هذا يؤدي إلى عدم مطابقة البيانات المستمدة من الاختبار للنموذج المستخدم بدرجة أو بأخرى لذلك يجب على الباحثين والمشتغلين بالقياس ، العناية باختبار النموذج الذي يتلاءم وطبيعة السمة المراد قياسها ، ونوع البيانات المستمدة من الاختبار الذي يبني لقياس هذه السمة بحيث تحقق هذه البيانات فروض النموذج المستخدم ، والجدير بالذكر أن نماذج السمات الكامنة تختلف في عدد الفروض اللازم توافرها في البيانات المستمدة من الاختبارات ، ويعتبر " نموذج راش " أقل هذه النماذج في عدد الفروض اللازم توافرها في البيانات ، حيث يعتمد على افتراض أن جميع *مفردات الاختبار تميز بنفس القدر بين الأفراد ، ولا تختلف عن بعضها البعض إلا في معامل الصعوبة*.ولهذا سمى بنموذج راش أحادي البعد أو أحادي البارامتر .

وتتميز أو نظرية السمات الكامنة بثلاث مميزات أساسية تتعلق بمفهوم الموضوعية Objectivity في القياس هي : (Wright ,1977,1979,1985)، (صلاح علام، ١٩٨٦:١٠٣)،(Roznowski 1991)

١. وجود عينة كبيرة من مفردات الاختبار التي تقيس نفس السمة أو القدرة،فإنه يمكن الحصول على قيمة تقديرية Estimate لقدرة الفرد ، تكون مستقلة إحصائياً عن عينة المفردات،أي متحررة من صعوبة المفردات التي اختير بها Item Free Person Measurement .

٢. وجود عينة كبيرة من الأفراد ، فإنه يمكن الحصول على قيمة تقديرية لمعاملات الصعوبة ومعاملات التمييز للمفردات التي اختبروا بها ،وهذه القيمة تكون مستقلة إحصائياً عن عينة الأفراد التي طبق عليها الاختبار Person-Free Test Calibration .

٣. تمكننا نماذج أو نظرية السمات الكامنة من الحصول على معامل إحصائي يبين مدى دقة تقدير قدرة كل فرد من أفراد العينة بواسطة مفردات الاختبار ، وربما يختلف هذا المعامل من فرد إلى آخر بحسب مستوى قدرة كل منهم .

نموذج راش : Rasch Model

يرتبط نموذج " راش " باسم عالم الرياضيات الدانمركي جورج راش G.Rasch

الـذي أرسـى قواعـده عام ١٩٦٠ ، منادياً بأهمية القياس الموضوعي وضرورة بناء الأداة الموضـوعية فـي قياس السلوك ، ويرجع الفضل الأول إلى العالم الأمريكي بنجامين رايت B.Wright الذي طوعه للتطبيق العملي بهدف تحقيق الموضوعية التي يعنى بها : أن درجة الفـرد فـي الاختبار لا يجب أن تكون دالة لعينة الأفراد التي استخدمت في التدريج الأصلي لـلمفردات Item Calibration الـتي يشتمل عليها الاختبار ، كما أنه يجب أن يحصل الفرد عـلى نفس الدرجة في كل اختبارين يقيسان نفس السمة أو القدرة ، ولتوضيح ذلك : نفترض أنـنا قمـنا ببـناء وتطبيق اختبارين يشتملان على نفس العدد من المفردات التي تقيس نطاقاً سـلوكياً معيـناً طبقاً للنظرية السيكومترية الكلاسيكية Classical Psychometric Theory عـلى مجموعة معينة من الأفراد ، وبفرض أن أحد الاختبارين يشتمل على المفردات الصعبة ، والآخـر يشـتمل عـلى المفردات السهلة ، فسوف نجد أن درجة فرد معين من أفراد هذه المجموعـة فـي كلا الاختبارين تختلف باختلاف صعوبة المفردات ، حيث نتوقع أن يحصل الفـرد على درجة أعلى في الاختبار الأقل صعوبة بينما يحصل على درجة أقل في الاختبار الأكـثر صـعوبة . وهـذا يعـني أن القياسات الناتجة عن مثل هذه الاختبارات تعتمد على المفـردات التي يشتمل عليها كل اختبار منها ، أي أن هذه القياسات غير متحررة أو مستقلة عن المفردات المستخدمة في القياس .

وكذلـك إذا استخدمنا في تقنين اختبار معين تلاميذ المرحلة الإعدادية فإنه لا يجوز تطـبيق هذا الاختبار على تلاميذ مرحلة مختلفة كالمرحلة الابتدائية طبقاً للنظرية السيكومترية الكلاسـيكية ، إذ يجـب أن تكـون خصائص الأفراد في عينة التقنين مماثلة لخصائص أفراد المجـتمع المسـتهدف ،وبذلك لا تكون أدوات القياس التي تبنى على هذا الأساس مستقلة عن خصـائص العيـنة Not Sample-Free، ولهـذا يذكر " راش " أن مفهوم الموضوعية في القيـاس غيـر مـتوفر في القياس النفسي والتربوي كما هو متوفر في العلوم الفيزيائية*،لكي يـتحقق شـرط الموضـوعية في القياس يجب أن تتحرر تداريج أدوات القياس من خصائص الأشـياء المـراد قياسها ، وأن يتحرر قياس الأشياء من خصائص أدوات القياس* ، حيث أن توافـر الشرطان يمكننا من تعميم القياس بحيث لا يقتصر على أداة قياس بعينها ، كما يمكننا مـن موازنـة الأشـياء المـراد قياسـها باستخدام أدوات قياس متشابهة وليست متطابقة .
(Mengel & Schorr 1992) (Willmot & Fowles, 1974:15)

بنوك الأسئلة في ضوء الاتجاهات الحديثة للقياس :

تعتـبر بنوك الأسئلة من أهم التطبيقات العملية لنموذج (راش) حيث يتيح استخدام الـنموذج تكوين بنك للأسئلة (أمينة كاظم ١٩٩٥) ، (Robert 1990) (Waimer 1990) ، بحيث يضم البنك عدة اختبارات تتدرج جميعها في تدريج واحد مشترك وصفر واحد مشترك ، بحيث تعرف هذه البنود مدى واسعاً من مستويات المتغير موضوع القياس ، وعندما يستخدم نمـوذج راش فـي بناء بنوك الأسئلة ، عندئذ نتغلب على مشكلتين وهما : موضوعية تقدير المسـتوى الأكـاديمي للطالب بالرغم من أختلاف الاختبارات وتباين مستوياتها ، موضوعية تقديـر المسـتوى الأكـاديمي للطالب مع اختلاف مستوى مجموعته ، يبدأ بنك الأسئلة بدمج اختبارين في تدريج واحد ، وينتهي بشبكة من الاختبارات التي تغطي مدى واسع، وتقوم فكرة تكوين بنك الأسئلة على ما يتمتع به النموذج من خاصية استقلال القياس عن كل من تأثيرات عيـنة الأفـراد ومجموعة البنود المستخدمة في النقاط الآتية : دمج بنود اختبارين في تدريج واحد ، تكوين بنك الأسئلة ، سحب الاختبارات الفرعية من بنك الأسئلة ، حبك الاختبار .

(١) دمج بنود اختبارين في تدريج واحد :

تهـدف هذه العملية إلى تحويل التدرج المستقل لكل من الاختبارين إلى تدريج واحد مشترك ، ويتطلب هذه التحويل موازنة أو معادلة لتدرج البنود المكونة لكل من الاختبارين . وتـتم هـذه العملية بأسلوبين يقوم أولهما (وهو الأكثر استخداماً) على استخدام بعض البنود المشـتركة كـرابطة بيـن الاختـبارين . ويقوم الأسلوب الثاني على استخدام بعض الأفراد المشتركين في أداء كل من الاختبارين .

(٢) تكوين بنك الأسئلة :

تتضـمن أغـلب بـنوك الأسئلة مئات من الأسئلة أو البنود المدرجة على آلاف من الأفـراد ولما كان من المستحيل على كل فرد من الأفراد أن يؤدي كل بند من هذه البنود فإن بناء البنك يقوم على دمج المجموعات المختلفة من البنود (أي الاختبارات المختلفة التي تعرف متغيراً واحداً) وذلك في تدريج واحد مشترك.ويوفر نموذج (راش) طريقة خاصة لبناء بنك الأسـئلة . ويبدأ الأسلوب الأكثر شيوعاً بتدريج عدد كبير من البنود على متغير واحد ، وذلك باسـتخدام مجموعـة من البنود المشتركة بين اختبارين مختلفين تعمل كرابطة ، تضمها في تدريـج مشترك واحد وبتكرار الرباط بين اختبارين أو أكثر تتجمع أعداد كبيرة من البنود في تدريـج مشترك واحد تشكل بنكاً للبنود أو الأسئلة . ويتطور الرباط بين الاختبارات المختلفة

حتى يصل إلى شبكة من الارتباطات العديدة وكل رباط منها يربط بين اختبارين وتشكل هذه الشبكة من الارتباطات نسيجاً من الاختبارات المدرجة جميعها على تدريج واحد . ويغطي بنك الأسئلة في هذه الحالة مدى واسعاً من المتغير ، كما يتضمن صوراً متكافئة من الاختبارات التي تغطي المستويات المختلفة من المتغير موضوع القياس .

(٣) سحب الاختبارات المطلوبة من بنك الأسئلة :

كما سبق فإن بنك الأسئلة يتضمن عدداً من المفردات جميعها في تدريج واحد وصفر واحد والتي يتسع مدى صعوبتها ويزيد عمل يستطيع أداؤه فرد واحد من الأفراد . ويتحكم في بناء البنك عدد البنود التي نود تدريجها في البنك ، والحد الأعلى لعدد البنود المكونة لكل اختبار من الاختبارات المراد سحبها ، ومدى الصعوبة التي نود أن يغطيه هذا البنك .

ويتميز بنك الأسئلة هذا بما يتميز به نموذج (راش) من استقلالية القياس،ويتحرر بذلك تقدير الأفراد من تأثير البنود المستخدمة،وهذا يعني تكافؤ تقدير الأفراد(باعتبار الخطأ المعياري) مهما اختلفت البنود المستخدمة ، المسحوبة من البنك بشرط مناسبتها لمستوى الأفراد ، ويتيح مدى القدرة الذي يغطيه بنك الأسئلة فرصة اختبار المجموعات المختلفة من البنود تشكل الاختبارات المناسبة لمستويات الأفراد المتباينة ، كما يتيح ما يتضمنه البنك من ذلك المدى العريض من البنود المتناظرة والمتكافئة الصعوبة ، والفرصة لاختيار مجموعات البنود التي تشكل الصور المتكافئة من الاختبارات التي تناسب مجموعات الأفراد المتشابهة في مستوى القدرة .

وبذلك يشكل بنك الأسئلة مصدراً لمجموعة من الاختبارات ، التي تقيس مدى واسعاً من مستويات المتغير . وسواء أكانت هذه الاختبارات طويلة أو قصيرة ، سهلة أم صعبة ، واسعة من حيث مدى الصعوبة أم ضيقة ، فإنها تتعادل في تقديرها لمستوى قدرة الأفراد . وبهذا يمكن اعتبار مشكلة بناء الصور المختلفة من الاختبارات الموضوعية في طريقها للحل ، وذلك بما يتيحه بنك الأسئلة من مرونة في اختيار الاختبارات المختلفة التي تتعادل تقديراتها للأفراد بصورة مباشرة .وهكذا يمكن أن نقارن بين مستويات القدرة للأفراد أو المجموعات المختلفة . كما يمكن أيضاً قياس التغير الذي يحدث في مستوى الفرد أو مستوى الأفراد . وتكون هذه المقارنة ، أو قياس هذا التغير باستخدام أي من مجموعات البنود (الاختبارات الفرعية) طالما أنها مسحوبة من بنك واحد للأسئلة ، وطالما أنها مناسبة للأفراد الذين يؤدون

الاختبار وكلما كان الاختبار مناسباً للفرد كلما كان تقدير القدرة أقرب للدقة ؛ وعلى هذا فعندما يسحب اختبار فرعي ما من بنك الأسئلة ينبغي أن يكون مناسباً للفرد أو مجموعة الأفراد الذين يقومون بأدائه . ويكون الاختبار مناسباً عندما تقترب المميزات الإحصائية للاختبار من المميزات التي يمكن أن يتصف بها الفرد أو الأفراد الذين تهدف إلى تقديرهم . وعلى هذا الأساس يمكن اختيار البنود التي تكون الاختيار المناسب لقياس قدراتهم من بين البنود التي يضمها بنك الأسئلة

ويكون تقدير المميزات الخاصة بالفرد أو الأفراد بصورة تقريبية ، أما عن طريق خبرة الباحث وتوقعه ، وإما عن طريق تجربته لبعض الصور من المقياس ،أما المميزات الإحصائية التي ينبغي توافرها في الاختبار المناسب فتتضمن صعوبته ومدى هذه الصعوبة وعدد بنوده ، وتحتاج بنوك الأسئلة إلى مبرمجين أو مستخدمي البرامج ولا يستلزم ذلك دراية بمحتوى المواد الدراسية ، وتكون مهمة المبرمج أو مستخدم البرنامج هي تخزين البنود بطريقة معينة واستدعاء البنود وقت الحاجة إليها ، ويتم التخزين للبنود الجيدة التي تم تدريجها ومعايرتها عن طريق إعطاء رقم رمزي للمادة الدراسية والمحتوى والهدف المرتبط به البنود ثم يخزن مع كل سؤال البيانات الخاصة به وعند الاستدعاء تستخدم نفس الأرقام الكودية والبيانات الخاصة بالأسئلة،وقد وضعت لذلك العديد من البرامج .

المشكلات المتعلقة ببنوك الأسئلة وكيفية التغلب عليها :

نظراً لتعدد مشروع بنوك الأسئلة وتعدد خبرات المشاركين فيه وتنفيذه يواجه صعوبات متعددة مرتبطة بجميع جوانب الإعداد والتنفيذ والاستخدام ومن هذه الصعوبات ما يلي : (أمينة كاظم ١٩٩٥ - ٢٦٢)

١ - تحديد الأهداف الخاصة بالمواد الدراسية حيث تصاغ الأهداف في صورة إجرائية سلوكية قابلة للتنفيذ والقياس .

٢ - وضوح الأهداف الخاصة بالمادة الدراسية لكاتبي الأسئلة وعدم الخلط بين الأهداف والمحتوى وطرق التدريس ، ويمكن التغلب على ذلك من خلال الدورات التدريبية ووجود متخصص في التقويم مع كل مجموعة للمواد .

٣ - كتابة الأسئلة أو البنود وهي خطوة تواجه صعوبات كثيرة من قبل غير المتدربين حيث يجدون صعوبات بالغة في التفرقة بين أنواع الأسئلة ، وفنيات صياغتها ، والمستويات

العقـلية التي تقيسها تلك الأسئلة وغير ذلك من مشكلات . والتغلب على تلك المشكلات يسـتلزم عمل دورات تدريبية لكتابة بنود الأسئلة في كل مادة دراسية وكل مرحلة ،كما تستلزم وجود متخصص في التقويم والمادة الدراسية على رأس كل مجموعة . ويجب أن يشـارك في كل مجموعة عدد كبير من المتخصصين في المادة الدراسية والمتفوقين في الدورات التدريبية لكتابة الأسئلة .

٤ – وضـع خطـة لتجريب بنود الأسئلة التي تمت كتابتها حتى يمكن تحليل الإجابات عليها واختبار الملائم ورفض البنود الأخرى أو تعديلها . وهنا ليس من الضروري أن تجرب البنود على عينات كبيرة كما أنها لا تستلزم أن يجيب كل طالب على جميع البنود ويجب أن تحـدد الأسـاليب الـتي تستخدم في تحليل إجابات البنود للحكم على جودتها . ومن المناسب استخدام أحد نماذج السمات الكامنة مثل راش أو نموذج برنيوم ، وما يلزم ذلك من برامج خاصة بها

٥ – تخـزين بـنود الأسـئلة فـي البنوك واستدعائها ويستلزم ذلك حاسبات آلية ومبرمجين مـتدربين وبرامج موضوعة خصيصاً لذلك . ويجب أن يخصص بنك مستقل لكل مادة دراسية .

٦ – إدارة بـنوك الأسئلة ، وهي مشكلة تستلزم تحديد المسئولية حتى لا تتسرب بنود الأسئلة خـارج البـنك ، ولذلك يمكن وضع البنود بعد تدريجها ومعايرتها على شرائط Disks يحتفظ بها شخص واحد ، على أن توجد بنود أخرى في البنوك يمكن تداولها واستخدامها فـي الـتقويم المستمر أثناء العام الدراسي . بينما تستخدم البنود المحتفظ بها في التقويم الـنهائي . ويمكن ترك البنود التي استخدمت في الاختبارات جانباً لفترة ثم يعاد وضعها ثانيـة أو وضع بنود جديدة غيرها ،ومن الممكن إعادة استخدام بنود الاختبار النهائية في التقويم المستمر .

٧ – اسـتخدام بنك الأسئلة وكثرة السحب منه دون إضافة قد تؤدي إلى الإفلاس . ولذلك فإن المحافظـة عـلى تواجـد البنك تستلزم أعداد بنود أسئلة جديدة بصفة مستمرة وتجريبها ومعايـرتها ثم وضعها في البنك ، وكذلك تعديل البنود التي تم تداولها أكثر من مرة بين الطـلاب عـن طـريق الاختـبارات أو تـركها لفترة قبل إعادة استخدامها ولذلك فإن مجموعـات كتابة بنود الأسئلة تستمر في عملها وكذلك نظم التجريب والتحليل وما الى ذلك من سحب وإضافة .

٨ – وتوجد مشكلة أخرى هامة تواجه مشروع بنوك الأسئلة في الدول النامية وهي متعلقة بالرأي العام المناهض للفكرة وللقياس الموضوعي بصفة عامة . فمازالت فكرة استخدام أسئلة المقال مسيطرة على تفكير العديد من القائمين بالتقويم التربوي . ولذلك فإن الحاجة إلى تعديل الرأي العام لتقبل الفكرة أمر ضروري .

الحاسبات الآلية و القياس و برامج بنوك الأسئلة :

لم يقتصر استخدام الحاسب على معالجة البيانات للتحليلات الإحصائية في عملية القياس ، فقد تطورت أساليب الاستخدام من خلال ٤ أجيال لخصها وأوردها " بندرسن Bunderson عام ١٩٨٩ (عبد الله الكيلاني ١٩٩٤) وهي :

١– الجيل الأول : الاختبار بالحاسوب Computerized Testing :

يتم فيه تطوير الاختبارات العادية لتخزن في الحاسوب ، ويتم تشغيلها وتطبيقها عن طريقه بواسطة بعض البرمجيات Software تنظم عمليات التشغيل والتطبيق والتصحيح وتحليل النتائج ، ومن مميزاته:دقة عمليات التقنين دون تدخل العنصر البشرى،والحفاظ على السرية،إمكانية الحذف والتعديل لفقرات الاختبارات وطرق التصحيح،أمكن البداية في عمل بنوك الأسئلة Item Bank .

٢ – الجيل الثاني : الاختبارات التكيفية Adaptive (Tailored) Testing :

تلك الاختبارات يصعب تطبيقها آلا بالحاسوب،فالفقرة الاختبارية التي تعرض على المفحوص ترتبط باستجابته لما قبلها،أو أنها تتحدد بنتيجة أدائه على الاختبارات السابقة.

٣ – الجيل الثالث : القياس المتواصل Continuous Testing :

لازال في مرحلة الدراسة والتجريب ، حيث تستخدم فيه مقاييس مدرجة متضمنة في المنهج المدرسي لقياس مدى التغير أو التقدم في كل مرحلة لدى الطالب بشكل منتظم ومتواصل ، حيث يفترض أن خبرات التعلم مصممة جيدا للوصول لنتائج محددة عند كل مرحلة،وهناك معايير لتقويم الأداء عند كل مرحلة من مراحل التقدم نحو الهدف النهائي.

٤ – الجيل الرابع " القياس الذكي " Intelligent Testing :

محوره قاعدة معلومات حاسبية مع أساليب استدلالية في عمليات تقدير الدرجات ، وتفسير الأداء ، والتوجيه للمعلمين ، حيث يتم إيداع ما لدى خبراء القياس والتقويم من معرفة في ذاكرة الحاسوب في نظام رمزي يشكل " قاعدة المعلومات "،وهذا ذكاء مصطنع يستفيد

منه المعلمين ممن يفتقرون خبرة القياس والتقويم ،ويتميز الجيل الرابع عن الثالث بقاعدة المعلومات التفسيرية والاستدلالية .

من أهم برامج الحاسبات الآلية Computers الشهيرة و المستخدمة في إعدادات بنوك الأسئلة في الدول المتقدمة وبعض الدول العربية :

١ - برنامج يسمى Micro Cat وضعته مؤسسة تقويم الأنظمة في ولاية مينوسوتا بالولايات المتحدة الأمريكية Assessment Systems Cooperation وهو من أحدث البرامج ويحتوي على ١٦ برنامج مرتبطة بتخزين الأسئلة واستدعائها وإعداد الاختبارات وطباعتها وتصحيحها وتحليل البيانات بعدة طرق منها الطريقة العادية باستخدام برنامج تحليل بنود يسمى Iteman ، وبطريقة نموذج راش الأحادي المعلم RASCAL .

٢ - برنامج آخر من نفس المؤسسة السابق الإشارة إليها يسمى ASTEC يهتم بتخزين واسترجاع البنود وإعداد الاختبارات طبقاً لمواصفات معينة

٣ - برنامج تستخدمه جامعة كامبردج يسمى MICRO SCALE وهو يستخدم لتخزين واستدعاء الأسئلة وتحليل البنود باستخدام نموذج راش والنموذج ثنائي المعلم . ويوجد بالمركز القومي للامتحانات والتقويم التربوي بالقاهرة .

٤ - برنامج BICAL وهو برنامج من إعداد رايت Wright ويستخدم نموذج راش في تحليل البنود،ويوجد بالجامعة الأمريكية بالقاهرة وبكلية البنات عين شمس .

٥ - برنامج يسمىLOGIST من إعداد لورد Lord في مؤسسة الاختبارات الأمريكية ETS وهو يطبق نموذج راش في تحليل البنود .

٦ - برنامج يسمى DICOT أعده ماسترز Masters وهو يطبق نموذج راش .

٧ - برنامج Form يمثل هذا البرنامج أولى خطوات إنشاء بنك الأسئلة وهو تجميع وكتابة أعداد وفيرة من المفردات Pool of Item والتي تقيس متغيراً واحداً .

٨ - برنامج Forcal يشكل هذا البرنامج المرحلة الثانية في تكوين وبناء بنك الأسئلة وهو مرحلة تطبيق الصور المختلفة للاختبار على عينات ملائمة من الطلاب ، حيث تسجل كل الدرجات التي يحصل عليها الطلاب في ملف File

٩ - برنامج Shift يعمل على إيجاد صعوبة كل صورة من صور الاختبار على حده ، ويضيف القيمة التي تنتج إلى تقدير صعوبة كل مفردة تبعاً للصورة التي ظهر فيها حتى تتم تدريج صعوبة المفردات على تدريج مشترك .

١٠ – برنامج Item list :يعمل هذا البرنامج على إعداد قائمة تحتوي على جميع مفردات بنك الأسئلة ، والأرقام الخاصة بهذه المفردات وكذلك الإجابة الصحيحة لكل منها، وصور الاختبار التي تظهر فيها كل مفردة ، وصعوبة هذه المفردة لكل صورة من الصور وكذلك حسن مطابقة المفردة لكل صورة .

١١ – برنامج Item map : يهدف البرنامج إلى التحقق من أن مفردات البنك تقيس المتغير المطلوب قياسه يقوم برسم الأشكال البيانية التي توضح الصعوبة .

١٢ – برنامج Kid list يقوم بإعداد قوائم الطلاب والمعلومات الخاصة عنهم ، ويقوم بإجراء حسن مطابقة هذه التقديرات لنموذج " راش " ويعد أشكال وصور بيانية توضح الدرجات التي حصل عليه الطلاب في مفردات الاختبار .

١٣ – برنامج Kid map يختص البرنامج بتمثيل أداء كل طالب في الاختيار بيانياً ، وإعداد خريطة لكل طالب توضح مستوى .

مراجع :

١ – أمينة محمد كاظم (١٩٩٥) : اتجاهات معاصرة في بناء بنوك الأسئلة . الأسس التربوية لإعداد المعلم الجامعي . ط٢ . ٢٤٧ .

٢ – صلاح الدين محمود علام (١٩٨٦) : تطورات معاصرة في القياس النفسي والتربوي ، الكويت ، جامعة الكويت .

٣ – صلاح الدين محمود علام (١٩٨٧) : دراسة موازنة ناقدة لنماذج السمات الكامنة والنماذج الكلاسيكية في القياس النفسي والتربوي ، الكويت : المجلة العربية للعلوم الإنسانية ، العدد السابع والعشرون ، المجلد السابع .

٤ – صلاح الدين محمود علام (١٩٩٤):التوجهات المستقبلية لتقويم تحصيل الطلاب في ضوء متطلبات القرن الحادي والعشرين،جامعة البحرين ،كلية التربية،المؤتمر العلمي الثالث .

٥ – فؤاد أبو حطب (١٩٩٥) : أنواع الأسئلة . الأسس التربوية لإعداد المعلم الجامعي . ط٢ . جامعة عين شمس . ٢٣٥ .

٦ – عبد الله زيد الكيلاني (١٩٩٤) : المجالات التطبيقية للتقويم التربوي . جامعة البحرين ،كلية التربية،المؤتمر العلمي الثالث . دولة البحرين .

٧ – مكتب التربية العربي لدول الخليج (١٩٩٤) : التعليم والحاسوب في دول الخليج العربية " الواقع وافاق التطوير " . ندوة الحاسوب في جامعات دول الخليج العربية . المنامة . دولة البحرين ، ص ١٥.

٨ – وليم عبيد (١٩٩٥) : المنهج والعملية التعليمية . الأسس التربوية لإعداد المعلم الجامعي . ط٢ . جامعة عين شمس – ٧٥ .

٩ – وليم عبيد (١٩٩٥) : استخدام الكمبيوتر في التعليم . الأسس التربوية لإعداد المعلم الجامعي . ط٢ . جامعة عين شمس – ٩٩ .

10 – Bergstrom, BA (1992 a) Assessing existing item bank depth for computer adaptive testing, National Council on measurement in education, San Francisco, C.A., USA

11 – Bergstrom, B.A. (1992 b) Alternative the level of difficulty in adaptive testing, Journal of Applied Measurement in Education, 5(2), 137, San Francisco, C.A., USA

12 – Hambleton.R.K. &Cook, L.L. (1977). Latent trait models and their use in the analysis of education test data, Journal of Education Measurement, 14(2), and 52.

13 – Hambleton.R.K. (1987). Item response theory principles and application, NY.Kluwer.Nijhoff publication.

14 - Hambleton.R.K.Jones, R.W. (1994). Item parameter estimation errors and their influence on test infornation function, Journal of Applied Measurement inEducation.7 (3), 171.

15 – Hemker B. T. et al. (1995). Selection of unidimensional scale from a multidimensional item bank in the polygamous Mokken Model, Applied psychological Measurment, 19(4), 337-52.

16 – Jahja, v. (1988). Robustness of the Rach model. DAI, 1749.

17 – Lord F. (1980). Application of item response theory to practical tests problems. Hill Dale, New Jersey.

18 – Nissan, S.et al. (1996) An analysis of factor affecting the difficulty of dialogue item in TOEFL Listening comprehension. TOEFL Research Reports. 51,EducationTesting Service, Princeton, and NJ.

19 – Robert, J &Darrell, B .V. (1990). Bilog 3: Item analysis and test scoring and logistic models. Scientific Software, Inc., 2nd.Ed.35.
20 – Roznowsky, M. and othérs (1991). Three approaches to delivering the dimensionality of binary items, Jornal of applied psychology measurment.12 (2). 109.
21 – Stocking, M.L. (1994). Three practical issues from modern adaptive testing item pools. Educational testing service, Princton, N J.
22 – Thorndike, R. et al. (1991) Measurement and evaluation in psychology and education. N J: Macmillan publishing company.
23 – Vercamp.W.Berger M. (1994). A simple and fast item selection procedure for adaptive testing research report, Faculty of Education, Netherlands.
24 – Wainer, H.et al. (1990). Computerized adaptive testing, Erl Baum Associates Publisher.New Jersey.
25 – Ward A. & Murray W. 91994). Guideline for the development of item bank, Journal of Education and Measurement.
26 – Willmot, S. & Fowles, D. (1974). The objective interpretation of test performance: The Rach model applied, National foundation for educational research, England.
27 – Wood, R. &Skurnik D. (1989). Item banking, National foundation for Education Research, England and Wales.
28 – Wright, B. &Douglass.G. (1977). Conditional versus unconditional procedures for sampler-free-item analysis, Journal of Educational and psychological Measurment, 37(3), 573-586.
29 – Wright, B. &Stone, M. (1979). Best test design. Chicago Press.
30 – Wright, b&Masters G. (1982). Rating scale analysis, Chicago Mesa Press.