

الفصل الأول

" المشكلة وأهميتها "

- أولا : المجال العام لمشكلة البحث .
- ثانيا : تحديد المشكلة .
- ثالثا : تحديد المصطلحات .
- رابعا : أهمية البحث .
 - الأهمية النظرية .
 - الأهمية التطبيقية .
- خامسا : حدود البحث .

أولا : المجال العام لمشكلة البحث :

تعد مشكلة قياس تقويم تحصيل الطلاب من المشكلات الرئيسية التي شغلت ولا تزال تشغل اهتمام المربين وخبراء القياس النفسي والتربوي حتى وقتنا الحالي ، والسبب في ذلك أن أساليب القياس والتقويم تلعب دورا هاما في تطوير العملية التعليمية .
" فالعملية التعليمية تعتبر منظومة تتألف من حلقات مترابطة متفاعلة تشمل أهداف التعليم أو محتواه — أى أساليب التدريس والتعليم — والقياس والتقويم ، وبدأ التفاعل بين هذه المكونات يوءد أن أى تطوير في احدها لا بد أن يكون له صدى وأثره المباشر على المكونات الأخرى . (١٣ : ١١٢) ،

" وهناك مداخل متعددة لهذا التطوير الشامل للتعليم من بينها ذلك المدخل الذى يلقي قبولا متزايدا من جانب المربين ألا وهو اتخاذ تطوير نظم التقويم وأساليبه نقطة بدء لاجداث الاصلاح فى الجوانب الأخرى التعليمية . (٦ : ٥) ،

ويستدعى ذلك إعادة النظر بصورة مستمرة فى أدوات القياس الحالية وتطويرها لى تصبح أكثر دقة وموضوعية وتقرب من المقاييس المستخدمة فى العلوم الطبيعية ومن بين أساليب التقويم المعروفة الاختبارات المقالية والاختبارات الشفهية والاختبارات الموضوعية . وكما وجه نقد شديد فى كل من الاختبارات المقالية والشفهية من حيث عدم موضوعيتها فى اعطاء الدرجة الدقيقة التى تعبر عن المستوى الحقيقى للتلميذ فنجد أنه وجه كذلك نقدا شديدا للاختبارات الموضوعية فى العقد الماضى — يرجع ذلك الى الاعتقاد بأن هذه الاختبارات ليست أدوات مصبوبة بسبب التخمين الذى يحدث من جانب المفحوصين عندما يكونوا غير متأكدين من الاجابة الصحيحة أو غير عارفين لها . (٢ : ٢) ،

ومن بين الاختبارات الموضوعية ، الاختبارات مرجعية المعيار والاختبارات مرجعية المحك ، ولما كانت الاختبارات التقليدية المرجعية المعيار غير مفيدة بالدرجة الكافية فى تحقيق أغراض التقويم التربوى مثل : تطوير المناهج ، معرفة مدى تقدم

التلاميذ ، معرفة مدى فاعلية أساليب التدريس المختلفة ، معرفة مدى سلامة المناخ المدرسى ، فقد ظهرت طريقة حديثة وحديثة مختلفة القياس وهى طريقة أكثر اتساقا مع متطلبات التقييم التربوى وهذه الطريقة الجديدة يطلق عليها القياس المرجعى المحك Criterion-Referenced Measurement (٣٧٦ : ٩) ، والاختبارات المرجعية المحك Criterion-Referenced tests تعتبر من المقاييس التربوية المعاصرة التى بدأت تنتشر فى الولايات المتحدة الأمريكية منذ أوائل السبعينات حتى الآن ، وأصبحت تطبق فى دول كثيرة ضمن برامج التقييم التربوى .

ويستخدم هذا النوع من الاختبارات لتقدير أداء الفرد بالنسبة الى اطار سلوكى معين ، أى مجموعة من المعارف والمهارات المعرفة بطريقة اجرائية (سلوكية) بصرف النظر عن علاقة أدائه بأداء غيره من الافراد الذين يطبق عليهم نفس الاختبار ، حيث أن هذا النوع من الاختبارات يناسب المواقف التعليمية التى تتطلب قياس اتقان الطالب لمجموعة من المهارات الاساسية التى ترتبط عادة دراسية أو محتوى دراسى معين (١٢ : ٣) وعند بناء الاختبار لابد من مراعاة شروطه : وهى الثبات والصدق والموضوعية والمعايير ، حيث يعتبر الصدق والثبات من أهم شروط الاختبار الجيد ، لذلك فإن الثقة فى اختبار ما تتوقف الى حد كبير على مقدار صدقه وثباته .

ولقد أشار كل من كرونباك cornbach وميهل Meehl (١٩٥٥) الى أن الصدق التكويني Construct validity يمكن أن يتحسن بواسطة الاجراءات المختلفة التى تتبع للتغلب على مشكلة التخميس ، كما أشار كل من مور Moor وبلود Blood (١٩٥١) الى أنه من الممكن احداث زيادة فى ثبات الاختبار دون احداث نقص فى صدقه .

ولقد بذلت محاولات لرفع صدق وثبات اختبارات تم بناؤها من قبل حيث قام بعض المشتغلين بالقياس النفسى بدراسة الثبات والصدق لدرجات الاختبارات الموضوعية ذات الاختيار من متعدد . وكذلك أجريت مجموعة من الدراسات فى هذا المجال كانت تقوم على أساس الافتراض القائل ، بأنه اذا كانت هناك رغبة فى زيادة ثبات مفردات

الاختبار وثبات الاختبار نفسه وصدقه فلا بد من اتباع نظام معين من الدرجات أو بعبارة أخرى اتباع طريقة معينة للتصحيح ، وعندما يتضح أن معاملات الثبات ضعيفة فيجب تغيير نظام الدرجات الحالي بنظام جديد واجراء فحص آخر لصدق وثبات تلك المفردات والاختبار ككل ، ومع أساس هذا الافتراض قدمت نظم عديدة لتصحيح الاختبارات ذات الاختيار من متعدد وتقدير درجات المفردات على أساس هذه النظم ومن هذه النظم على سبيل المثال نظام كومبس Combs ، ومور Moor (١٩٥٦) ونظام ديفز Davis ، ربي Pippay (١٩٦٨) ونظام دافيد Davaid (١٩٧٢) ونظام دايموند Diamond وايفانس Evans (١٩٧٣) ، ولقد تمت مقارنة هذه النظم بعضها ببعض ومقارنتها أيضا بالنظام التقليدي لتقدير درجات مفردات الاختبار وذلك لمعرفة أى من هذه النظم تتمتع بثبات وصدق أكبر للاختبار المتعدد الاختيارات (٢ : ٢ — ٤)

ومن هذه الدراسات : دراسة داريل . ل . سابرز ، جوردن وايت Sabers, D L. and White, G.W.

(١٩٦٩) ، والتي توصلت الى أن هناك زيادة لقيمة لها في الصدق والثبات باستخدام طريقة ديفز عند مقارنتها بالطريقة التقليدية ، وقد أثبتت كذلك دراسة رونالد ك ، هامبلتون ، دينسم ، برورتنس ، بروث ، ترايوب (١٩٧٠)

Ronald K., Hambleton, Dennis M., Roberts, Rosse, Trouh

أن معامل الصدق التنبؤى الناتج من درجات المجموعة التقليدية بلغ ٠٦٢١ وأن معامل الصدق الناتج من درجات الثقة بلغ ٠٧٢٠ ، أما بالنسبة للثبات فلم تدل النتائج الى وجود فروق دالة بين الطريقتين التقليدية ودرجات وزن الثقة ، (٢ : ٢٦١ — ٢٧٢) .

وقد توصلت دراسة روبرت م . ربي (١٩٧٠) Rippey, R. M.

(١٩٧٠) الى أن طريقة الاحتمال المقترح للإجابة الصحيحة تعطى أفضل ثبات ممكن ، يليها الطريقة الاقلية ، طريقة الاختيار الاستنتاجي التي تعطى أقل ثباتا في الطرق الخمسة التي استخدمت في تلك الدراسة (٤٧ — ١٦٥ : ١٧٠)

(١٩٧١) Koehler, R.A.

وقد قارنت دراسة روجر أ. كوهلر

بين ثلاث طرق في تقدير الدرجات:

الاولى : يجيب فيها التلاميذ بطريقة وضع نسب احتمالية تعبر عن مقدار صواب كسل
اختيار احتمالي .

والثانية : يجيب فيها التلاميذ بالطريقة التقليدية .

والثالثة : تعالج فيها الدرجات بالدالة اللوغاريتمية .

وقد أظهرت نتائج هذه الدراسة أن هناك زيادة طفيفة في معامل الثبات بين
المجموعة الاولى والمجموعة الثانية . كما أن هناك زيادة في معامل الثبات قدرها ٠٣ ر
بين المجموعة الثانية والمجموعة الثالثة ، ولقد أثبتت النتائج وجود فروق طفيفة في الصدق
لصالح الطريقة التقليدية عن الطريقتين الاخرتين (٢٧٧ : ٢ — ٢٧٩)

، دراسة وانلوب كسوب Kansup, W. (١٩٧٥) والتي قد

توصلت الى أنه توجد فروق في معاملات الثبات بين الدرجات الوزنية والدرجات التقليدية
لصالح الدرجات الوزنية ، وذلك عند حساب الثبات بمعامل ألفا (α) وذلك
في المجموعة الضابطة ، أما بالنسبة للثبات في المجموعة التجريبية فقد وجدت بعض
الفروق في معاملات الثبات بين الإجراءات الخمسة التي أتبعته في معالجة
الدرجات وذلك عند استخدام معامل ألفا ، وهذه الفروق وجدت بصفة خاصة بين
الطريقة اللوغاريتمية (الطريقة الثالثة) والطريقة التربيعية (الطريقة الرابعة)
لصالح الطريقة التربيعية ، أما بالنسبة لاعادة الاختبار فقد أظهرت النتائج أنه لا توجد
فروق بين معاملات الثبات المحسوبة بالطرق الخمس المذكورة ، وبصفة عامة أظهرت
النتائج أن معاملات الثبات (المحسوبة) باستخدام معامل ألفا واعادة الاختبار
في المجموعة التجريبية أعلى منها في المجموعة الضابطة . (٢٢٩ : ٤٠ — ٢٣٩)

، دراسة جيمس ج. ديموند Damond, James, J. (١٩٧٥) والتي

وصلت الى أن الدرجات التجريبية تعطى ثباتا أكبر من الثبات التي تعطيه الدرجات

التقليدية (١٢٩ : ٢٧ — ١٣٣) ، (٢٩٦ : ٢) ، دراسة ريتشارد م. ايفانيس

، كاسي مسفلدت Evans R.M. & Misfeldt, K. (١٩٧٤) والتي

وصلت الى أن اسلوب الدرجات الذاتية Self-Scoring Method (المتمثل في اعطاء المفحوصين تغذية رجعية كلية) تزيد الثبات عن الاسلوب التقليدي فسي جميع الحالات • (٢٨ : ١٢٤٨) ، (٢ : ٣٠٥ — ٣٠٦)

ويتضح من عرض الدراسات السابقة أن هناك بعض نواحي القصور يمكن ايجازها فيما يلي :-

أولا : الدراسات السابقة كانت تركز على الاختبارات المرجعية المعيار Norme-Referenced-Tests (NRT)

ولم تتعرض الى الاختبارات المرجعية المحك Criterion-Referenced-Tests (CRT)

ثانيا : في الدراسات السابقة تم تطبيق الاختبارات مرة واحدة مباشرة دون عمل تجارب استطلاعية ، وفي ذلك قصور لتحديد الزمن المناسب للاختبار مع كل طريقة من طرق التطبيق •

لذلك فان الامر يحتاج الى اجراء تجارب استطلاعية لتحديد الزمن المناسب للاختبارات المستخدمة مع كل طريقة من طرق التطبيق ثم اجراء التطبيق ثانيا بالزمن الجديد الذي تم تجديده من التجارب الاستطلاعية •

ثالثا : لم تتناول الدراسات السابقة ثبات مفردات الاختبارات المستخدمة وانما اكتفت بالثبات الكلي للاختبار ، علما بأن ثبات الاختبار الكلي يتأثر بثبات مفرداته •

ثانيا : تحديد المشكلة :

ومن خلال نواحي القصور الموجودة في الدراسات السابقة ، وعدم تعرضها للاختبارات المرجعية المحك وطرق حساب ثباتها ، نبعت مشكلة الدراسة الحالية التي تتمثل في التساؤلات الآتية :

- (١) هل يختلف ثبات الاختبارات المرجعية المحك ذات الاختيار من متعدد باختلاف طرق تقدير الدرجات للمفردات ؟
- (٢) هل يختلف ثبات الاختبارات المرجعية المحك ذات الاختيار من متعدد باختلاف طرق حساب هذا الثبات ؟
- (٣) هل يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق حساب الثبات على ثبات الاختبارات المرجعية المحك ؟

ثالثا : تحديد المصطلحات :

طرق تقدير الدرجات : هي نظم معينة لاعطاء الدرجات في تصحيح العينات أو هي طرق معينة للتصحيح يختلف فيها درجة كل مفردة على حسب صعوبتها بحيث لا يشترط أن تكون جميع درجات الاسئلة متساوية .

الثبات : Reliability

يقصد بمصطلح الثبات دقة الاختبار (أو أى أداة قياس أخرى) في القياس أو عدم تناقضه مع نفسه ومعنى ذلك أن الثبات هو عبارة عن الاتساق بين قياسات الاختبار حيثما كانت هذه القياسات . أو بعبارة أخرى فإن الثبات يعرف على أنه عبارة عن درجة التناسق أو الاتساق بين مقياس الشيء الواحد (١٩ : ٦٠) ، فالمقياس الثابت يعطى نفس النتائج إذا قاس نفس الشيء مرات متتالية (٢٠ : ٥١٣) .
وتعرف الدراسة الحالية الثبات بأنه استقرار النتائج مقدرة بالدرجات للاختبار مرجعى المحك ذات الاختيار من متعدد إذا قاس نفس الشيء مرات متتالية .

المحك : Criterion

يستخدم لفظ Criterion (مفرد) بمعنى محك Criteria (جمع) بمعنى محكات " ويقصد بالمحك مستوى أداء معين ذي مواصفات محددة يمكن استخدامه في مقارنة أداء الفرد بهدف تقويم ذلك الأداء (٥ : ٣٠) .

وتعرف الدراسة الحالية المحك بأنه الوصول الى مستوى أداء ٨٠٪ من درجات الاختبار مرجعي المحك ذات الاختيار من متعدد ويهدف مقارنة أداء الافراد بهذا المحك المحدد مسبقا من قبل الباحث

رابعاً : أهمية البحث :

(أ) الأهمية النظرية :

(١) تتضح الأهمية النظرية للبحث الحالي في أنها محاولة لدراسة الطرق المختلفة لتقدير الدرجات والطرق المختلفة لحساب ثبات الاختبارات المرجعية المحك .

(٢) قد يسهم البحث الحالي في بناء اختبارات مرجعية المحك فسي المواد الدراسية الأخرى .

(ب) الأهمية التطبيقية للبحث الحالي في أنها تساعد على :

(١) إمكانية زيادة ثبات الاختبارات المرجعية المحك ذات الاختيار من متعدد .

(٢) الوصول الى أفضل طريقة عن طرق تقدير الدرجات للمفردات مع أفضل طريقة من طرق حساب ثبات الاختبارات المرجعية المحك .

(٣) قد يسهم الاختبار المرجعي المحك الذي تناولته الدراسة الحالية في تقويم تلاميذ الصف التاسع من مرحلة التعليم الاساسى بطريقة أفضل .

(٤) قد تفيد الدراسة الحالية في تطوير منهج الجبر الحالى للصف التاسع من مرحلة التعليم الاساسى .

(٥) تخدم الدراسة الحالية المسؤولين بوزارة التربية والتعليم وخاصة المهتمين بأساليب التقويم فى اعداد اختبارات مرجعية المحك من نوع الاختيار من متعدد .

خامسا : حدود البحث :

(١) اقتصر البحث الحالى على وحدتين فى منهج الجبر المقرر على طلاب الصف التاسع الاساسى فى العام الدراسى ١٩٨٥ / ١٩٨٦ وذلك فى وحدتى الاعداد الحقيقية والتحليل . وذلك لان هاتين الوحدتين تشتملا على مجموعة من المفاهيم والقوانين التى تعتبر كأساس لمواصلة الدراسة فى الرياضيات بصفة خاصة فى الاعوام المقبلة .

(٢) اشتملت عينة البحث الحالى على طلاب (بنين وبنات) الصف التاسع من مرحلة التعليم الاساسى بمحافظة الشرقية .
