الفصل الثانى

الفصل الثانى

الدراسات السابقة

الفصل الثانى
الدراسات السابقة

يهدف هذا الفصل إلى تحليل مجموعة من الدراسات السابقة، للاستفادة منها فى بيان منهج البحث الذى ستتبعه الباحثة ، وفى تحديد الأدوات التى تستخدم فى جمع البيانـــات وتصنيفـــها وتحليلها ، وفى كيفية بناء الاختبار ، وإجراءات تطبيقه . ومن أجل تحقيق هذا الهدف ســـتقوم الباحثة بعرض أهم الدراسات التى أجريت فى مجال الاختبارات محكية المرجع.

وقد تنوعت أهداف الدراسات التى أتيح للباحثة الاطلاع عليها ، فمنها ما اقتصـــر هدفـــه على بناء الاختبار ، ومنها ما تعدى ذلك إلى الوصول لأفضل الطرائق لتقنينه .

ولذلك تم تقسيم هذه الدراسات إلى المحورين التاليين :

أولاً : دراســـات تناولت بنـــاء الاختبار .
ثانياً : دراسات تنـاولـــت تقنين الاختبار .

وسوف يتم عرض الدراسات من القديم إلى الحديث، مع توضيح أوجه الاختلاف والاتفاق فيما بينها وبين الدراسة الحالية ، وما تم الإفادة منه فى البحث الحالى .

أولاً :- دراسات تناولت بناء الاختبار

واشتملت هذه الدراسات على :-

١- دراسة إسماعيل الفقى (١٩٨٢)
٢- دراسة عماد عبد المسيح (١٩٨٢)
٣- دراسة محمد الدسوقى (١٩٨٥)
٤- دراسة تاج الســـــــر (١٩٨٦)
٥- دراسة ماسترز و إيفانز (١٩٨٦) Masters and Evans
٦- دراسة محمود إبراهيم (١٩٩٠)
٧- دراسة صلاح علام (١٩٩١)
٨- دراسة ريفيز وليفن (١٩٩٢) Reves and levin
٩- دراسة محمد السيد على (١٩٩٢)
١٠- دراسة نادية عبد السلام (١٩٩٢)
١١- دراسة آمال أبو حديد (١٩٩٤)
١٢- دراسة محمد فتح الله (١٩٩٥)
١٣- دراسة نادية عبد السلام (١٩٩٦)
١٤- دراسة إسماعيل حسن فهيم (١٩٩٦)

١ – دراسة إسماعيل الفقى (١٩٨٢)

هدفت الدراسة إلى :-

١- التحقق من فعالية نظام حديث للتقويم التكوينى بمقارنته بالطرائق التقليدية (باستخدام الطرائق المعتادة للتقويم – الامتحانات الحالية)

٢- التحقق من فعالية نظام التقويم التكوينى بمقارنته باستخدام اختبار للتقويم التجميعى يعتمد على الأسس العلمية لبناء الاختبارات التحصيلية فى هذا الصدد .

منهج الدراسة :- استخدم الباحث المنهج التجريبى ، وذلك باشتمال عينة البحث على مجموعتين ، إحداهما تجريبية والآخرى ضابطة من طلاب الصف الثانى من المرحلة الإعدادية بمدرسة الطبرى بمصر الجديدة ، وكان عدد تلاميذ المجموعة الأولى (٤٨) ثمانية وأربعين طالباً (٤٨) ، والثانية (٤٩) تسعة وأربعين طالباً، تتراوح أعمارهم ما بين ١١- ١٤ سنة .

وتكونت أدوات البحث من :-

١- اختبار قبلى لقياس مدى التمكن فى العلوم للصف الثانى الإعدادى (وهو من إعداد الباحث) واشتمل على عدة أنواع من الأسئلة الموضوعية (الصواب والخطأ – الاختيار من متعدد – المزاوجة – التكملة) .

٢- اختبارات للتقويم التجميعى فى مادة العلوم للصف الثانى الإعدادى (وهو من إعداد الباحث) واحتوت على ثمانية اختبارات تكوينية قصيرة، قدمها لطلاب المجموعة التجريبية. وحدد مستوى الإتقان ٩٠٪ لوصول التلاميذ للإجابة .

٣- الامتحانات المدرسية العادية فى مادة العلوم للصف الثانى الإعدادى التى يؤديها التلاميذ فى نهاية العام الدراسى .

وقد اتبع الباحث **الإجراءات التالية فى بناء الاختبار التكوينى** (وهو الاختبار الذى يعتمد على الاختبارات المحكية المرجع)

١- التحليل المناسب للمحتوى والسلوك لوحدة التعلم .

---

[١] إسماعيل محمد عبد الرؤف الفقى : دراسة تجريبية لفاعلية برنامج للتقويم التكوينى فى مادة العلوم بالمرحلة الاعدادية . رسالة ماجستير غير منشورة ، كليه التربية – جامعة عين شمس ، ١٩٨٢ .

٢- إعداد جدول المواصفات .
٣- اختيار مفردات الاختبار بحيث تشكل تدرجاً فى المستويات من الأقل صعوبة إلى الأكثر صعوبة .
٤- تحليل مفردات الاختبار وتحديد مستوى الإتقان ، وقد حدد بـ ٩٠٪ .

ومن أهم النتائج التى توصلت إليها الدراسة :-

١- أكدت النتائج أثر استخدام الاختبارات التكوينية (المحكية المرجع) فى استمرار التحسن فى سلوك التلاميذ التحصيلى .
٢- وجود اتجاه للتحسن بتكرار هذه الاختبارات .
٣- أن استخدام البرنامج التكوينى فى مقرر البيولوجى للصف الثانى من المرحلة الإعدادية حقق نتائج إيجابية، حيث استخدمت نتائج التقويم فى أغراض التشخيص وتقديم التغذية الراجعة لكل من المدرس والتلميذ ، ويكشف عن الصعوبات والنقائص فى تعلم التلاميذ أولاً بأول ،ويقدم البدائل العلاجية المناسبة لطبيعة الفروق الفردية للتلاميذ لعلاج نقائص التعلم .

وقد تميزت هذه الدراسة أنها الأولى من نوعها- فى حدود علم الباحثة- التى قامت بالتقويم التكوينى ومقارنته بالطرائق التقليدية المعتادة للتقويم .

وتتفق هذه الدراسة مع الدراسة الحالية فى نوعية الاختبار مع اختلاف مادة الدراسة، وفى بعض خطوات بناء الاختبار المحكى المرجع، وتدرج مفرداته من الأسهل إلى الأصعب ، وأنواع الأسئلة الموضوعية، ***وستستفيد*** الباحثة من النقاط السابقة عند إعدادها للاختبار .
وتختلف هذه الدراسة مع الدراسة الحالية فى نوع العينة ، والأدوات .

٢ - دراسة عماد عبد المسيح (١٩٨٢)

هدفت الدراسة إلى :

١- كيفية بناء مقياس مرجعى - الميزان فى مادة الطبيعة و(حدة المادة) المقررة على طلاب الصف الأول الثانوى العام .

---

[٢] عماد عبد المسيح : **دراسة لبناء مقياس مرجعى المحك فى مادة الطبيعة بالصف الأول الثانوى** . رسالة ماجستير غير منشورة ، كلية التربية - جامعة المنيا ، ١٩٨٢ .

٢- قياس مستوى تحصيل الطلاب بالصف الأول الثانوى العام فى وحدة خواص المادة بمقرر الطبيعة لهذا الصف فى المستويات الثلاثة الأولى لتقسيم بلوم للمجال المعرفى .
٣- تقديم بعض المقترحات والتوصيات التى تسهم فى تحصيل أفضل لوحدة خواص المادة .

منهج الدراسة :- استخدم الباحث المنهج التجريبى على عينة مكونة من (٤٩٢) طالب و (٤٥١) طالبة، من طلبة وطالبات الصف الأول الثانوى بالمدارس العامة بإدارة المنيا التعليمية، وتم اختيارهم بطريقة عشوائية . طبق عليهم المقياس المرجعى الميزان فى وحدة خواص المادة بمنهج الفيزياء للصف الأول الثانوى، والذى قام بإعداده الباحث، وهى الأداة الوحيدة التى استخدمها الباحث .

وقد اتبع الباحث الخطوات التالية فى إعداد المقياس :-

١- تحديد المجال أو محتوى المادة الدراسية المراد قياسها حسب مستويات بلوم (تذكر - فهم - تطبيق) .
٢- تحديد الأهداف التعليمية لمجال المحتوى المراد قياسه فى صورة سلوكية .
٣- إعداد مواصفات القياس مرجعى الميزان .
٤- بناء مفردات المقياس . وقد استخدم (الاختيار من متعدد) .
٥- تحديد مستوى الأداء (الدرجة الفاصلة) ٨٠ ٪ .
٦- حساب صدق محتوى المقياس (عن طريق المحكمين وتطابق المفردة بالهدف)
٧- حساب ثبات المقياس (بإعادة تطبيق المقياس على عينة من الطلاب)

ومن أهم النتائج التى توصلت إليها الدراسة :-

١- أن مستويات تحصيل الطلاب فى المقياس يتدرج تنازليا من المعرفة إلى الفهم، ثم إلى التطبيق . مما يدل على اهتمام الطلاب بالحفظ ، والاسترجاع، واستظهار المعلومات، والحقائق أكثر من الاهتمام بالفهم والتطبيق لهذه المعلومات والحقائق ، والاستفادة مما يدرس فى مواقف جديدة غير مألوفة لهم فى الحياة .

وقد تميزت هذه الدارسة بأنها الأولى من نوعها- فى حدود علم الباحثة- التى قامت ببناء مقياس مرجعى الميزان؛ لقياس مستوى تحصيل الطلاب بالصف الأول الثانوى العام فى مادة الطبيعة .

وتتفق هذه الدراسة مع الدراسة الحالية فى خطوات بناء الاختبار المحكى المرجع ، والطريقة التى استخدمت لتقدير صدق الاختبار وثباته ، والمستويات المعرفية لبلوم (تذكر - فهم - تطبيق)

وستستفيد الباحثة من النقاط السابقة عند إعدادها للاختبار. **وتختلف** هذه الدراسة مع الدراسة الحالية من حيث الموضوع ، ونوع العينة وعدم اقتصارها عند بناء المفردات (على الاختيار من المتعدد) . وكذلك عدم اقتصارها على المستويات الثلاثة الأولى لتقسيم بلوم للمجال المعرفى ، وترى الباحثة أن هذه المستويات تعتبر مستويات دنيا للمرحلة الثانوية ، وستحاول الباحثة تطبيق مستويات أعلى للمجال المعرفى بما يتناسب مع المرحلة الابتدائية وأهداف المادة والمرحلة . واختلفت دراسة (عبد المسيح) مع دراسة (إسماعيل الفقى) فى تحديد درجة القطع . **واتفقت** مع نتائج دراسة إسماعيل الفقى (١٩٨٢) على أهمية الاختبارات المحكية المرجع فى أغراض التشخيص ، مما يدل على أهميتها فى عملية التقويم .

٣ - دراسة محمد الدسوقى الشافعى (١٩٨٥)

هدفت الدراسة إلى :-

١- بناء اختبار مرجع المحك فى مادة الهندسة الفراغية لطلاب الصف الثانى الثانوى العلمى.

٢- قياس اتقان الطلاب فى الصف الثانى الثانوى العلمى فى مادة الهندسة الفراغية المقررة باستخدام الاختبار .

منهج الدراسة :- استخدم الباحث المنهج التجريبى على عينة اشتملت على (٤٦٠) طالب وطالبة ، طبق عليهم الصورة (أ) من الاختبار ، و(٣٢٦) طالباً وطالبة ، طبق عليهم الصورة (ب) وتم استبعاد (١٥٣) ورقة من أوراق الإجابة . وتم اختيار العينة بطريقة عشوائية من المدارس الثانوية بمحافظة الدقهلية والغربية . وطبق عليهم أداة البحث وهى الاختبار المرجعى المحك فى مادة الهندسة الفراغية لطلاب الصف الثانى الثانوى العلمى بصورتيه (أ) ، (ب) .

وقد اتبع الباحث الخطوات التالية فى بناء الاختبار :-

١- تحديد مجال ومحتوى المادة الدراسية المراد قياسها .

٢- تحديد أهداف المحتوى إجرائياً .

---

[٣] محمد الدسوقى عبد العزيز الشافعى : **بناء اختبار مرجع لميزان فى مادة الهندسة الفراغية لطلاب الصف الثانى الثانوى العلمى** . رسالة ماجستير غير منشورة ، كلية التربية - جامعة طنطا ، ١٩٨٥ .

٣- إعداد مواصفات الاختبار .

٤- إعداد مفردات الاختبار. وقد أعدها من أنواع (الاختبار من متعدد ، أسئلة تحتاج إجابة قصيرة لسؤال أو تكملة لعبارة ناقصة ، مقال) .

٥- تحديد الدرجة الفاصلة . وحددت بـ (٨٠٪) كمستوى أداء جيد ومقبول لكل جزء من أجزاء الاختبار (المعرفة - الفهم - التطبيق - حل المشكلات)، وأيضاً للاختبار ككل .

٦- تقدير صدق الاختبار. وتم تقدير صدق محتوى الاختبار بتقدير صدق مفرداته باستخدام معامل تطابق المفردة بالهدف .

٧- تقدير ثبات الاختبار. وتم باستخدام الطرائق غير الارتباطية التى اقترحها . سواميتزان وزملاؤه (١٩٧٤) . والطرائق الارتباطية التى يصح استخدامها فى هذا القياس (كيودر ريتشارد سون ، معامل كروبناخ)

ومن أهم النتائج التى توصلت إليها الدراسة :

١- أن نسبة الذين حققوا مستوى الاتقان (٨٠٪) كانت منخفضة جداً، إذ بلغت (٢،٩٣٪)، وأن نسبة المتقنين فى مستوى التذكر أكبر منها فى مستوى الفهم ، أكبر منها فى مستوى حل المشكلات ، وهذا يؤكد على اهتمام التلاميذ بالحفظ والاستظهار فى تعلمهم هذه المادة . مما يباعد بينهم وبين تحقيق الأهداف المرغوبة .

وقد تميزت هذه الدراسة بأنها الأولى من نوعها - فى حدود علم الباحثة - التى قامت ببناء اختبار مرجع لميزان فى مادة الهندسة الفراغية لطلاب الصف الثانى الثانوى ، وكذلك تطبيقها . . فى محافظتى الدقهلية والغربية .

وتتفق هذه الدراسة مع الدراسة الحالية فى خطوات بناء الاختبار ، وأنواع المفردات المستخدمة ، وطرائق تقنين الاختبار ، وعدم اقتصار الاختبار على المستويات المعرفية الدنيا (تذكر - فهم - تطبيق)، وهذا ما ستستفيد منه الباحثة .

وتختلف هذه الدراسة مع الدراسة الحالية فى الموضوع ، ونوع العينة ، ومستوى (حل المشكلات) فى بناء المفردات ؛ لأن هذا المستوى يناسب الرياضيات أكثر من اللغة العربية ، مع الاقتصار على صورة واحدة للاختبار .

وقد اتفقت هذه الدراسة مع دراسة عبد المسيح (١٩٨٢) فى نـــوع العينـــة ، واســـتخدام عينات كبيرة ، وفى تقدير صدق الاختبار بتطابق المفردة بالهدف ، وتقدير الدرجـــة الفاصلـــة بـــ ٨٠٪ ، كذلك اتفقت مع نتائج الدراسة فى أن نسبة المتقنين فى مستوى التذكر أو المعرفـــة أكبر منها فى المستويات الأعلى ، واهتمام الطلاب بالحفظ و الاستظهار . واتفقت مع **دراســـة الفقى** (١٩٨٢) ، وعبد المسيح (١٩٨٢) أيضا فى أهمية الاختبارات المحكيـــة المرجـــع فـــى أغراض التشخيص ؛ مما يؤكد على أهميتها فى عملية التقويم .

٤ – دراسة تاج السر (١٩٨٦)

هدفت الدراسة :–

إلى اختبار فعالية استراتيجيتين من استراتيجيات التقويم التكوينى ، هما استراتيجية (بلوم) واستراتيجية (كيلر) فى تدريس علم البيئة فى المدارس الثانوية بالسودان ، وتسعى كل منهما إلى تحقيق درجة عالية من التحصيل عن طريق الاختبارات المنسوبة إلى المحك .

منهج الدراسة :– استخدم الباحث المنهج التجريبى وذلك بتقسيم العينة الكلية التى اشتملت على (٢٣٥) تلميذاً وتلميذة من تلاميذ الصف الأول الثانوى إلى أربع مجموعات . مجموعتان مـــن البنين (٥٩ ، ٦١) ، ومجموعتان من البنات (٥٩ ، ٥٦) وتراوحت أعمارهم بين ١٦، ١٧ سنة ، وينتمون إلى مدرستين من مدارس أم درمان بالسودان . واســـتخدمت مـــع المجموعـــة الأولى استراتيجية بلوم ، والآخرى استراتيجية كيلر .

وقد تضمنت أدوات البحث :

١– أدوات وصف العينة واشتملت على :–

أ– اختبار الذكاء المصور .

ب– السجلات المدرسية .

٢– **الاختبارات وتشمل :–**

أ – الاختبار التحصيلى القبلى / البعدى .

ب– الاختبارات التكوينية لاستراتيجية (كيلر) .

جـ– الاختبارات التكوينية لاستراتيجية (بلوم) .

د – اختبارات الإتقان لكل من المجموعتين الدارستين.

---

[٤] – تاج السر عبد الله الشيخ : **التقويم التكوينى كاستراتيجية تعلم للإتقان فى تدريس علم البيئة فى المدارس الثانوية بالسودان** . رسالة ماجستير غير منشورة ، كلية التربية – جامعة عين شمس ، ١٩٨٦م .

٣- بطاقات الملاحظة الفردية .

وباعتبار أن الاختبارات التكوينية اختبارات (مرجعية المحك) ؛ فقد اتبع الباحث المبـــادئ العامة التى وضعها جرونلاند (١٩٧٦) ، لبناء الاختبارات المرجعية المحك وهى :-

أ - تحديد الهدف من الاختبارات .

ب- تحليل محتوى المادة وإيجاد الأوزان النسبية .

جـ- بناء جدول المواصفات .

د - اختيار أنواع المفردات .

هـ- إعداد مفردات الاختبار المناسبة .

و - تقنين الاختبار .

وقد قام الباحث باستخدام صدق المحتوى للتحقق من صدق الاختبار بعرضه على مجموعة من المحكمين . وقد حدد الباحث محك التحصيل بـــ ٨٠ ٪ .

ومن أهم النتائج التى توصلت إليها الدراسة :-

١- أن الاختبارات التى تستخدمها كل من الاستراتيجيتين من نـــوع الاختبـــارات المرجعيـــة المحك؛ للتأكد من وصول التلاميذ إلى مستوى محدد من إتقان المادة موضع التعلم ، وأنه مـــن السهل نسبياً لمعظم الطلاب إنجاز هذه الاختبارات ، وأن يحققوا ٨٠٪ وما فوق، وهذا يرجـــع إلى أن الاختبار يتصل بوحدات صغيرة ، يتم تعلمها بالمعدل المناسب لكل تلميذ .

٢- أن الاختبارات التكوينية (المرجعية المحك) هى اختبارات تشخيصية للكشف عن مواطن الضعف فى تعليم التلاميذ ، وتقدم البرامج العلاجية المناسبة لعلاج نقائص التعلم لدى التلاميذ .

٣- أن التحسن فى التحصيل يرجع إلى تتابع واستمرارية التقويم .

وقد تميزت هذه الدراسة بأنها اختبرت فعالية استراتيجيتين (بلـــوم) و(كيلـــر) باســـتخدام اختبارات محكية المرجع ، مع تعدد الأدوات الآخرى .

وتتفق هذه الدراسة مع الدراسة الحالية فى بعض خطوات بناء الاختبار ، وطرائق تقنين الاختبار.

وتختلف هذه الدراسة مع الدراسة الحالية فى الموضوع ، ونوع العينة ، وتقســـيمها إلـــى بنين وبنات ، والأدوات .

وقد اتفقت هذه الدراسة مع دراسة عبد المسيح (١٩٨٢) ، والدســـوقى (١٩٨٥) ، فـــى المرحلة التعليمية (طلاب المرحلة الثانوية) ، وتحديد درجة القطع بـــ (٨٠٪) واتفقا أيضـــا

مع الفقى (١٩٨٢) ، فى أن الاختبارات المحكية المرجع اختبارات تشخيصية للكشف عن مواطن الضعف فى تعليم التلاميذ ، وأنها تقدم البرامج والبدائل العلاجية المناسبة لعلاج نقائص التعلم لدى التلاميذ .

٥ - دراسة ماسترز و ايفانز (١٩٨٦) Masters and Evans

هدفت الدراسة :

إلى تحليل أداء التلاميذ فى الاختبارات المحكية المرجع ؛ لمساعدة المعلمين على استخدام هذه التقارير المولدة بواسطة الكمبيوتر؛ لمعرفة مدى تقدم التلاميذ ، وتشخيص مناطق القوة والضعف لديهم .

عينة البحث :- اشتملت العينة على جميع التلاميذ فى بعض المدارس الابتدائية فى ولاية فيكتوريا باستراليا . وطبق اختبارين على كل تلميذ فى كل صف دراسى من أفراد العينة .

أدوات الدراسة :-

١- تم بناء سبعة اختبارات فى مادة الرياضيات، بعد تقسيمها إلى عشرة موضوعات . وبنيت بطريقة منطقية ؛ أى من السهل إلى الصعب ، ورمزت من (B) إلى (H) . ويحتوى كل اختبار على سؤالين من الموضوعات العشر للمنهج .

- وطبق اختبارين على كل تلميذ فى كل صف دراسى من أفراد العينة . فقد طبق على التلاميذ فى الصف الأول اختبار (B) و (C) والتلاميذ فى الصف الثانى (C) و(D) وهكذا.

٢- القيام بعمل أداة لقياس مدى استفادة التلاميذ فى كل صف لموضوعات الرياضة المقسمة إلى عشرة موضوعات .

ومن أهم ما توصلت إليه الدراسة :

تم التوصل إلى ترجمة دقيقة لمستوى تحصيل التلاميذ باستخدام الاختبارات المحكية المرجع فى أى وقت أثناء العام الدراسى . والتفسير فى ضوء أنماط السلوك وفى ضوء فهم خصائص مستوى التحصيل الذى وصل إليه التلميذ .

---

٥ Masters, G. N and Evans, J. A sense of Direction in Criterion Referenced Assessment . In: Studies in Educational Evaluation, V. ١٢, PP ٢٥٧-٢٦٥, Printed in Great Britain, ١٩٨٦.

وقد تميزت هذه الدراسة بأن العينة شملت كل التلاميذ فى جميع الصفوف الدراسية لبعض المدارس الابتدائية ، وهذا يدل على مدى أهمية هذه الاختبارات .

وتتفق هذه الدراسة مع الدراسة الحالية فى بناء الاختبار بطريقة منطقية ، أى من السهل إلى الصعب ، ونوع العينة (تلاميذ المرحلة الابتدائية) . وأهمية الاختبارات المحكية فى الكشف عن مواطن الضعف فى تعليم التلاميذ .

واختلفت الدراسة مع الدراسة الحالية فى المادة الدراسية ، وحجم العينة. وتطبيق اختبارين على كل تلميذ ، واستخدام أداة لقياس مدى الاستفادة للتلاميذ فى كل صف للموضوعات ، وهذا أيضا ما اختلفت فيه مع ما سبق من دراسات الفقى (١٩٨٢) ، عبد المسيح (١٩٨٢) ، الشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) .

واتفقت مع نتائج الفقى (١٩٨٢) ، وتاج السر (١٩٨٦) ، فى أهمية استخدام الاختبارات المحكية المرجع للوصول إلى ترجمة حقيقية لتحصيل التلاميذ .

٦ – دراسة محمود إبراهيم (١٩٩٠)

هدفت الدراسة إلى :

التوصل إلى أفضل الطرائق الاحصائية التى يمكن استخدامها فى حساب معاملات الثبات للاختبارات المرجعة إلى المحك فى ضوء الشروط التالية : طول الاختبار ، درجة القطع ، حجم العينة .

منهج الدراسة :– استخدم الباحث المنهج التجريبى على العينة التى تم اختيارها عشوائيا من طلاب وطالبات الصف السابع بمرحلة التعليم الأساسى بمحافظة القاهرة ، واشتملت على (٨١٨) طالبا وطالبة من مناطق : الوايلى والزيتون ومصر الجديدة . وطبق عليهم اختبار مرجع إلى المحك من إعداد الباحث فى مادة الرياضيات فى باب المجموعات كتاب (الجبر)

٦– محمود محمد ابراهيم عبد الله : **دراسة سيكومترية مقارنة لطرق حساب ثبات الاختبارات المرجعية إلى المحك.** رسالة ماجستير غير منشورة ، كلية التربية – جامعة عين شمس ، ١٩٩٠م .

وقد اتبع الباحث **الخطوات التالية فى بناء الاختبار** فى ضوء طريقة بابام :-

١- تحديد النطاق السلوكى الذى يقيسه الاختبار .

٢- تحديد الأهداف الإجرائية .

٣- إعداد مواصفات الاختبار .

٤- إعداد مفردات الاختبار . وقد تم اختيارها من أسئلة الاختيار من متعدد .

٥- تقدير صدق محتوى الاختبار عن طريق المحكمين ، واستخدام معادلة (هاملتون و رفينلى).

٦- تحديد طريقة التصحيح ، وتقدير الدرجات .

وقد حدد الباحث تسع درجات للقطع من ١٠٪ إلى ٩٠٪

ومن أهم النتائج التى توصلت إليها الدراسة :-

١- يفضل استخدام طريقة (هينا) للاختبارات التى تتكون من عدد كبير من المفردات والمتوسط ولأى حجم من العينات عند الأخذ فى الاعتبار تأثير طول الاختبار.

٢- وتفضل طريقة (سيكوفياك) للاختبارات القصيرة ولحجم العينة الكبير .

٣- وتفضل طريقة (ليفنجستون) لحجم العينة الصغير والمتوسط .

٤- تفضل طريقة (ليفنجستون) عند اتخاذ درجات القطع بين (٧٠: ٩٠٪) ولأى حجم للعينة ولأى طول للاختبار . كما يمكن استخدام طريقة (هينا) لنفس درجات القطع .

وقد **تميزت** هذه الدراسة بمحاولة إلقاء الضوء على أفضل الطرائق الإحصائية التى يمكن استخدامها فى حساب معاملات الثبات للاختبارات المرجعة إلى المحك فى ضوء طول الاختبارات ، درجه القطع ، حجم العينة .

وتتفق هذه الدراسة مع الدراسة الحالية فى خطوات بناء الاختبار ، واعتماد صدق المحتوى عن طريق المحكمين، واستخدام طريقة ليفنجستون لحساب ثبات الاختبار؛ لأنها أنسب الطرائق، وهذا ما ستستفيد منه الباحثة عند إعداد الاختبار . واتفقت مع دراسة كل من الفقى (١٩٨٢) ، وعبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) ، فى خطوات البناء ومع عبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) ، فى تقدير صدق المحتوى.واتفقت فى المرحلة التعليمية مع الفقى (١٩٨٢)، وفى الموضوع مع كل من الشافعى (١٩٨٥) ، وماسترز و ايفانز (١٩٨٦) ، وتختلف هذه الدراسة مع الدراسة الحالية ودراسة كل من عبد المسيح (١٩٨٢) ، وتاج السر (١٩٨٦) ، فى الموضوع ونوع العينة .

٧ – دراسة صلاح علام (١٩٩١)

هدفت الدراسة إلى :

المقارنة بين أربع طرائق – اثنتان تعدان من الطرائق المطلقة ؛ أى التى تعتمد على تحكيم محتوى الاختبار و مفرداته، وهما طريقتان Nedelesky , Angoff ، والطريقتان الأخريان تعدان من الطرائق النسبية؛ أى التى تعتمد على تحكيم أداء المختبرين، وجمع بيانات تجريبية فعلية تتعلق بهذا الأداء (وهما طريقتا المجموعة الحدية والمجموعة المتناقضة) – لتحديد مستويات الأداء فى اختبار محكى المرجع ، لقياس كفايات طلاب كليات التربية والمعلمين قبل الخدمة فى مجال القياس والتقويم التربوى والنفسى .

منهج الدراسة استخدام الباحث المنهج التجريبى على عينة من طلاب كليات التربية ممن يدرسون مقرراً سيكلوجياً فى الفروق الفردية ، بلغت (٣٨٢) طالب، كذلك استعان بـ (٢٦) محكما قسمها الباحث إلى ثلاث مجموعات متباينة فى مستوياتها .

وقد اتبع الباحث الخطوات التالية فى بناء الاختبار المحكى المرجع فى أساسيات القياس والتقويم التربوى والنفسى :-

١- تحديد موضوعات الاختبار .

٢- حدد الباحث مجموعة متكاملة من الأهداف السلوكية ، التى تتضمن المعارف والمهارات المتعلقة بالموضوعات المحددة مسبقاً .

٣- عرض الأهداف السلوكية على بعض المحكمين لتقدير أولويات هذه الأهداف على ميزان من ثلاث درجات (٣=مهم جداً ، ٢=مهم ،١= قليل الأهمية)

٤- حدد الباحث (٥٧) هدفاً وقام بكتابة المفردات (الاختيار من متعدد)

٥- عرض الاختبار على نفس مجموعة الخبراء الذين قاموا بتحكيم الأهداف ، وذلك للتحقق من أن كل مفردة مقياس صادق للهدف الذى تقيسه.

---

[٧] – صلاح الدين محمود علام : **دراسة مقارنة لبعض طرق تحديد مستويات الأداء فى اختبار مرجعى المحك** . المجلة المصرية للدراسات النفسية ، العدد الأول ، سبتمبر ١٩٩١ . مكتبة الانجلو المصرية ، القاهرة ، ص٧٧-٩٦.

ومن **أهم النتائج** التى توصلت إليها الدراسة :-

١- ضرورة استخدام طرائق مختلفة عند تقدير الحد الأدنى لمستويات الأداء فى الاختبارات .

٢- تزويد المحكمين ببيانات كافية فعلية عن أداء الطلاب ؛ حتى يتم تقدير مستويات الأداء بصورة أكثر ثباتاً واستقراراً .

وقد **تميزت** هذه الدراسة بأنها قارنت بين أربع طرائق لتحديد مستويات الأداء فى اختبار محكى المرجع .

وتتفق هذه الدراسة مع الدراسة الحالية ودراسة كل من الفقى (١٩٨٢) ، وعبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) ، ومحمود إبراهيم (١٩٩٠) فى بعض خطوات بناء الاختبار ، وتقدير صدق محتوى الاختبار عن طريق مطابقة المفردة للهدف وصدق المحكمين . وكذلك اقتصار المفردات على (الاختيار من متعدد). ولم تذكر الدراسة فى خطوات البناء إعداد جدول مواصفات ، وهذا ما اختلفت فيه الدراسة مع جميع الدراسات السابقة والدراسة الحالية ، كما اختلفت معهم فى نوع العينة والموضوع .

٨ - **دراسة ريفيز و ليفين (١٩٩٢) Reves and Levin**

هدفت الدراسة إلى :

تطوير نموذج اختبار محكى المرجع لإتقان مهارة القراءة للفهم لدى الدارسين بدورات تعلم اللغة الإنجليزية كلغة ثانية .

عينة الدراسة :- تكونت العينة من (٦٨) طالباً من طلاب السنة الأولى المسجلين لدراسة دورة القراءة للفهم فى اللغة الإنجليزية كلغة ثانية فى جامعة إسرائيل .

أدوات الدراسة :-

١ - استبانة وزع على عينة من :-

أ- معلمى القراءة فى مقرر اللغة الإنجليزية كلغة ثانية فى جامعات إسرائيل .

ب- الأقسام الجامعية المستفيدة من تدريس مقرر القراءة فى اللغة الإنجليزية كلغة ثانية .

جـ- الطلاب المتقدمين لدراسة مقرر القراءة كلغة ثانية .

---

٨ Reves, T and Levine, A . From Needs Analysis To Criterion - Referenced Testing." In: System , V. ٢٠ , N. ٢ , PP ٢٠١-٢١٠, printed in Great Britain , ١٩٩٢.

د – الطلاب الدارسين لمقررات القراءة فى اللغة الانجليزية كلغة ثانية فـــى مختلـف السنوات الجامعية .

وقد قسم محتوى الاستبانة إلى عشرة محاور مستقلة بواسطة المحكمين ، وتوصل إلـــى التحديد النهائى لعشرة محاور للاستبانة بناءً على موافقة الأغلبية من المحكمين .

٢– اختبار قبلى وبعدى فى إتقان القراءة ، حيث طبق الاختبار القبلى فى بداية العام الدراسى، والاختبار البعدى بعد سبعة شهور من التطبيق القبلى .

٣– اختبار التقدم فى الأداء لطلاب الفرقة الأولى – المكونة من (٦٨) طالباً – فى توظيف مهارات القراءة الموجودة فى المقرر .

٤– كتابة تقرير عن ملاحظات الطلاب لدرجة تقدمهم الذاتى فى الاستفادة من المقرر .

ومن أهم ما توصلت إليه الدراسة :

١– وجود علاقة بين حاجات المتعلمين ووظيفة القراءة ومحتوى المقرر والإسهامات التراكمية فى صياغة المحكات لاختبارات إتقان فهم القراءة .

٢– أن أهداف اختبارات إتقان مهارات القراءة هى :–

– معرفة أساليب البلاغـــــــة فـــى القراءة .

– تحديد الفكرة الرئيسة فـــى الموضـــــــوع .

– تحديد الأفكار المساعدة للفكـــــرة الرئيسة .

– اكتشاف غرض المؤلف من كتابة الموضوع .

– فهم ما استخلصـــــه المـــــؤلـــــــف .

– تقييـــــــم الموضـــــــــــــــوع .

٣– عدم تحديد المهارات الخاصة بفهم القراءة قبل إجراء الاختبار وتحليل النتائج ، ولكن بعد إجرائه وتحليل نتائجه .

وتميزت هذه الدراسة بأنها حاولت أن تطور فى نموذج الاختبار محكى المرجـــع لإتقـــان مهارة القراءة للفهم والأدوات المستخدمة لهذا التطوير .

وستستفيد الباحثة من هذه الدراسة بما توصلت إليه من النقاط الخاصة بأهداف اختبارات إتقان مهارات القراءة عند بناء الاختبار الخاص بها ،بحيث تكون هذه المهارات مناسبة لمستوى تلاميذ المرحلة الابتدائية وأهدافها .

وتختلف هذه الدراسة مع الدراسة الحالية ودراسة كل من الفقى (١٩٨٢) ، وعبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) ، وماسترز و ايفانز (١٩٨٦) ، ومحمود

إبراهيم (١٩٩٠) ، فى نوع العينة ، وحجمها ، والموضوع ، ولكنها اتفقت مع صلاح علام (١٩٩١) ، فى نوع العينة (المرحلة الجامعية)

كذلك **اختلفت** مع جميع الدراسات السابقة عليها والدراسة الحالية فى أن تحديد المهارات يتم بعد إجراء الاختبار وتحليل النتائج .

٩ – دراسة محمد السيد على (١٩٩٢)

هدفت الدراسة إلى دراسة :-

١- أثر التدريس باستخدام التقويم البنائى على تحصيل طلاب القسم الأدبى الثانوى فى مادة العلوم .

٢- دراسة أثر التدريس باستخدام التقويم البنائى على اتجاهات طلاب القسم الأدبى الثانوى نحو العلوم .

٣- دراسة أثر التدريس باستخدام التقويم على قلق طلاب القسم الأدبى الثانوى نحو الامتحان.

منهج الدراسة :- اعتمد الباحث على المنهج التجريبى من خلال مجموعتين تجريبتين ومجموعة ضابطة ، اشتملت على (٩٨) طالباً من طلاب الصف الثالث الثانوى بمدرسة سيف بن سلطان الثانوية للبنين بولاية عبرى فى سلطنة عمان . وقد تم اختيار العينة بصورة عشوائية . وقد طبق عليهم الباحث **أدوات البحث** وهى:

١- اختبارات بنائية Quizzes (من إعداد الباحث) .

٢- اختبار تحصيلى (من إعداد الباحث) .

٣- مقياس الاتجاه نحو مادة العلوم (من إعداد الباحث) .

٤- قائمة قلق الامتحان (من إعداد سبيلبرجر وآخرين)

وقد اتبع الباحث الخطوات التالية فى تصميمه للاختبارات البنائية (**المحكية المرجع**)

١- تحديد الأهداف التعليمية لكل درس من دروس الوحدة .

٢- تحليل محتوى كل درس من الدروس السابقة .

٣- وضع اختبار موجز (Quiz) لكل درس من دروس الوحدة فى ضوء أهدافه السلوكية وتحليل محتواه .

٤- صياغة مفردات الاختبارات الموجزة (Quizzes) من نوع (الاختيار من متعدد ، والتكملة).

---

[٥] – محمد السيد على : **فعالية التقويم البنائى على تحصيل طلاب القسم الأدبى بالمرحلة الثانوية واتجاهاتهم نحو العلوم ، وقلقهم للامتحان فى سلطنة عمان** . التربية المعاصرة ، العدد (٢٣) السنة التاسعة، سبتمبر ١٩٩٢ ، ص ٢٣٧ : ص ٢٦٧ .

٥- عرض الاختبارات الموجزة فى صورتها الأولية على مجموعة المحكمين ؛ للوقوف على صدقها وشمولها . وبعد التعديلات أصبح الاختبار لكل درس يتكون من (٣) مفردات إكمال ، (٤) مفردات اختيار من متعدد .

ومن **أهم النتائج** التى توصلت إليها الدراسة :-

١- أن استخدام التقويم البنائى فى تدريس مادة العلوم العامة لطلاب القسم الأدبى بالمرحلة الثانوية يسهم فى تحسين مستوى تحصيلهم الدراسى واتجاهاتهم نحو المادة الدراسية، ويخفف من قلقهم من الامتحان . مما يؤكد على أهمية التقويم البنائى فى العملية التدريسية.

وقد **تميزت** هذه الدراسة بأن الباحث قام بإعداد ثلاث أدوات من أربع ــ التى سبق الاشارة إليهم ــ استخدمها فى هذه الدراسة ، ليبين أثر التدريس باستخدام التقويم البنائى فى تحصيل القسم الأدبى الثانوى فى مادة العلوم .

وتتفق هذه الدراسة مع الدراسة الحالية فى بعض من خطوات بناء الاختبار البنائى (المحكى المرجع) ، مع استخدام (الإكمال والاختيار من متعدد) فى بناء المفردات . وهذا ما ستستفيد منه الباحثة من هذه الدراسة .و **اتفقت** مع دراسة عبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) ، فى نوعية العينة (طلاب المرحلة الثانوية) .

واختلفت هذه الدراسة مع الدراسة الحالية والدراسات السابقة عليها فى عدم إدراج جدول المواصفات من ضمن خطوات بناء الاختبار ، واتفقت مع دراسة صلاح علام (١٩٩١) ، فى هذه النقطة . واختلفت مع الدراسة الحالية فى الأدوات الآخرى المستخدمة .

واتفقت نتائج الدراسة مع دراسة الفقى (١٩٨٢) ، وتاج السر (١٩٨٦) ، و ماسترز وايفانز (١٩٨٦) على أن الاختبارات البنائية (المحكية المرجع) تسهم فى تحسين مستوى تحصيل التلاميذ . **واتفقت** مع دراسة كل من عبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) ، و تاج السر (١٩٨٦) ، و محمود إبراهيم (١٩٩٠) ، و صلاح علام (١٩٩١) ، فى تقدير صدق الاختبار .

١٠ – دراسة نادية عبد السلام (١٩٩٢)

هدفت الدراسة إلى :-

تحديد الدرجات الفاصلة لاختبار محكى المرجع والحكم على مناسبتها لتقدير إتقان الطلاب باستخدام ثلاثة محكات هى :-

١- متوسط الفرق المطلق . ٢- قرار الإتقان . ٣- أنموذج الاحتمالية .

منهج الدراسة :- اعتمدت الباحثة على المنهج التجريبى بتطبيق الاختبار المحكى المرجع لقياس منهج الإحصاء الوصفى على عينة تتكون من (٤٩) طالباً وطالبة من طلبة الدراسات العليا (دبلوم خاص و تمهيدى ماجستير) فى كلية البنات – جامعة عين شمس .

وقد اتبعت الباحثة الخطوات التالية فى بناء الاختبار المحكى المرجع :-

أ – تحديد النطاق السلوكى الذى يقيسه الاختبار (أى إعداد مواصفات المجال) .

ب- يتكون اختبار البحث الحالى من ٦٠ مفردة تقيس (٦) مجالات .

جـ- بناء مفردات الاختبار ، (الاختيار من متعدد) .

د – تقدير صدق المحتوى عن طريق مضاهاته بالمجالات السابقة .

هـ- تحديد الدرجة الفاصلة .

ومن أهم النتائج التى توصلت إليها الدراسة :-

١- أن الحد الأدنى للنجاح الذى حدده المحكمون مناسب لأقسام الاختبار عموماً .

٢- الاختيار من متعدد هو الأسلوب الأفضل إذا ما قورن بأسلوب الصواب والخطأ .

تتفق هذه الدراسة مع الدراسة الحالية فى بعض خطوات بناء الاختبار المحكى المرجع، و نوعية المفردات، وتقدير صدق المحتوى . وهذا ما اتفقت فيه أيضاً مع دراسة عماد عبد

---

[١٠] نادية محمد عبد السلام : **ثلاث محكات رئيسية لتحديد مستوى الإتقان فى الاختبارات محكية المرجع (دراسة امبريقية) . المجلة المصرية للدراسات النفسية** ، العدد الثالث ، ديسمبر ١٩٩٢ ، مكتبة الأنجلو المصرية ، القاهرة . ص ٨٨ : ص ١٠٦ .

المسيح (١٩٨٢)، محمود إبراهيم (١٩٩٠)، وصلاح علام (١٩٩١)، فـــى اقتصـــار نوعيـــة الاختبار على (الاختيار من متعدد) ، ومع دراسة كل من عبد المســـيح (١٩٨٢)، الشـــافعى (١٩٨٥)، وتاج السر (١٩٨٦)، ومحمود إبراهيم (١٩٩٠)، وصلاح علام (١٩٩١)، ومحمـــد السيد (١٩٩٢) .

وتختلف هذه الدراسة مع الدراسة الحالية فى نوع العينة والموضوع ، ولكنها اتفقت مـــع دراسة كل من صلاح علام (١٩٩١) وريفيزوليفن (١٩٩٢) فى نوع العينة .

وتعد هذه الدراسة هى الدراسة الأولى التى توصلت إلى أن أسئلة (الاختيار من متعدد) أفضل فى الاختبارات المحكية عن الصواب والخطأ . وهذا ما ستستفيد منه الباحثة أيضا عنـــد بناء مفردات الاختبار .

١١ – **دراسة آمال أبو حديد (١٩٩٤)**

هدفت الدراسة إلى :-

علاج أخطاء تلاميذ الصف الأول الثانوى فى استخدام علامات الترقيم فى مـــادة اللغـــة الإنجليزية .

منهج البحث :- استخدمت الباحثة المنهج التحليلى التجريبى وقد اشتملت عينـــة البحـــث على (٦٠) تلميذة، التى قسمتهن إلى مجموعتين ، تجريبية وضابطة و قامت الباحثة بـــإعداد اختبار المحكى المرجع .

وقد اتبعت الباحثة الخطوات التالية عند بناء الاختبار :-

١- تحليل محتوى الكتاب المدرسى .

٢- تحديد الأهداف ، لمستوى (التطبيق) .

٣- تحديد عدد بنود الامتحان ، حسب الوزن النسبى لكل هدف .

٤- بناء جدول الموصفات .

٥- تقدير صدق وثبات الاختبار ، واعتمدت الباحثة على صدق المحتوى .

وقد اعتبرت الباحثة محك الإتقان المطلوب فى هذه الدراسة هو ٨٠٪ .

---

[١١] Amal Abou Hadid, Analytical and Experimental study of Secondary School Students, Punctuation Errors, M.A. Thesis . Women's College , Ain Shams University , ١٩٩٤ .

ومن أهم النتائج التى توصلت إليها الدراسة :- أن تلاميذ الصف الأول الثانوى وصلوا إلى مرحلة الإتقان عند إجاباتهم عن الاختبار المحكى المرجع ، والـــذى يقيـس اسـتخدامهم لعلامات الترقيم المختلفة ،وذلك فى المهارات الأربع الأساسية فى اللغة (الاستماع ، القـراءة، الكتابة ، التحدث) .

وقد تميزت هذه الدراسة أنها الأولى من نوعها – فى حدود علم الباحثة – التـى قـامت ببناء اختبار محكى المرجع لقياس إتقان طالبات المرحلة الثانوية بعد علاجـهن مـن أخطـاء استخدام علامات الترقيم فى مادة اللغة الإنجليزية .

وتتفق هذه الدراسة مع الدراسة الحالية فى خطوات بناء الاختبار المحكى المرجع . وهذا ما اتفقت فيه مع دراسة كل من عبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) وتاج السـر (١٩٨٦) ، ومحمود إبراهيم (١٩٩٠) ، وصلاح علام (١٩٩١) ، ومحمد السـيد (١٩٩٢) ، نادية عبد السلام (١٩٩٢) . كذلك تتفق مع الدراسة الحالية فى استخدام مستوى (التطبيـق) عند إعداد مفردات الاختبار ، والموضوع .

واتفقت هذه الدراسة مع دراسة عماد عبد المسيح (١٩٨٢) ، الشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) فى اتخاذ ٨٠٪ درجه قطع .

واتفقت مع دراسة ريفيز وليفين فى نوعية الموضوع (اللغة) ، ومع دراسة مـن عبـد المسيح (١٩٨٢) ، الشافعى (١٩٨٥) ، وتاج السر (١٩٨٦) ، ومحمد السيد (١٩٩٢)، فى نوع العينة (المرحلة الثانوية) و مع كل من دراسة محمد السيد (١٩٩٢) ، ونادية عبـد السلام (١٩٩٢) فى حجم العينة الصغير (أقل من مائه) .

١٢ – دراسة محمد فتح الله (١٩٩٥)

هدفت الدراسة إلى :-

بناء اختبار محكى المرجع فى مادة العلوم لتلاميذ نهاية مرحلة التعليم الأساسى فى ضـوء الأهداف التعليمية الوظيفية لهذه المادة فى هذه المرحلة الأساسية .

---

[١٢] محمد محمد فتح الله سيد أحمد : **بناء اختبار محكى المرجع فى العلوم لتلاميذ نهاية مرحلة التعليم الأساسى** ، رسالة ماجستير ، غير منشورة ، كلية التربية – جامعة الأزهر ، ١٩٩٥ م .

منهج الدراسة :- استخدم الباحث المنهج التجريبى ، على عينة اشتملت (٢٠٠) تلميذ ، قسمهم الباحث إلى مجموعتين ؛ كل مجموعة (١٠٠) تلميذ .

وقد قام الباحث ببناء اختبار محكى المرجع ، استخدم الخطوات التالية فى بنائه :-

١- تحديد النطاق السلوكى للمحتوى التعليمى .

٢- تحليل النطاق إلى مكونات سلوكية .

٣- إعداد مواصفات الاختبار .

٤- بناء مفردات الاختبار . (الاختيار من متعدد)

٥- تحديد الخصائص السيكومترية للاختبار .

وقد حدد الباحث درجة القطع بـ ٨٠٪

وقد قام الباحث للتحقق من صدق محتوى الاختبار باستخدام طريقتين ، إحداهما تعتمد على التحكيم ، والآخرى على البيانات المستمدة من تطبيق الاختبار على العينة الاستطلاعية .

وقد قام بالتأكد من صدق الاختبار ككل عن طريق صدق المحكمين واستخدام معادلة روفينيلى وهامبلتون .

ولتحديد ثبات الاختبار استخدم الباحث أربع طرائق : طريقتان من الطرائق المستخدمة فى الاختبارات الجماعية المرجع (كيودر - ريتشاردسون ، معامل ارتباط بيرسون)، وطريقتان من الطرائق الخاصة بالاختبارات محكية المرجع (طريقة ليفنجستون، ومعامل كابالسواميناثان، هامبلتون، ألجينا) .

ومن أهم النتائج التى توصلت إليها الدراسة :

١- أن الاختبار (محكى المرجع) موضوع هذه الدراسة يساعد فى تحسين عملية التقويم لتلاميذ نهاية مرحلة التعليم الأساسى فى مادة العلوم .

٢- أن المعلم يستطيع أن يحدد مكانة التلميذ من التسلسل التعليمى للكفايات المحددة .

٣- أن الاختبار له فائدة تشخيصية فى توضيح جوانب القوة والضعف فى تحصيل التلاميذ أثناء تعلمهم الكفايات الأساسية لمادة العلوم .

وتتفق هذه الدراسة مع الدراسة الحالية فى بعض خطوات بناء الاختبار وهذا ما اتفقت فيه مع دراسة كل من عبد المسيح (٨٢) ، الشافعى (٨٥) ، وتاج السر (٨٦) ، ومحمود إبراهيم (٩٠)، وصلاح علام (٩١) ، ومحمد السيد (٩٢) ، نادية عبد السلام (٩٢) . وفى نوعية مفردات الاختبار (الاختيار من متعدد) والاعتماد على صدق المحكمين للتحقق من صدق الاختبار ، واستخدام طريقة ليفنجستون فى التحقق من ثبات الاختبار .

واتفقت مع دراسة عماد عبد المسيح (٨٢) ، الشافعى (٨٥) ، وتاج السر وآمال أبو حديد (٩٤) فى درجة القطع ٨٠٪ . ومع دراسة الفقى (٨٢) ، وتاج السر (٨٦) ، وماسترو ايفانز (٨٦) ، ومحمد السيد على (٩٢) ، فى أن الاختبارات المحكية لها فوائد تشخيصية توضح جوانب القوة والضعف فى تحصيل التلاميذ . **واختلفت** مع الدراسة الحالية فى المرحلة التعليمية للعينة .

١٣ – دراسة نادية عبد السلام (١٩٩٦)

هدفت الدراسة إلى :-

بناء اختبار تشخيصى هدفى المرجع لقياس إتقان الطلاب لمجموعة من المهارات الأساسية التى ترتبط بمادة الإحصاء الوصفى .

منهج الدراسة : استخدمت الباحثة المنهج التجريبى على عينة مكونة من (٢٢) طالباً وطالبة من دارسى الدبلوم الخاص والتمهيدى ماجستير عام ١٩٨٥/٨٤م ، وطبق عليهم صورتين متوازيتين (أ،ب) لاختبار محكى المرجع ، يشتمل على مجموعة من المهارات الأساسية اللازمة لتحصيل مادة الإحصاء الوصفى .

وقد اتبعت الباحثة الخطوات التالية عند بناء الاختبار المحكى المرجع وهى (خلاصة المراجعة المعدلة لفريمر ١٩٧٤) :-

١-تحديد الأهداف .

٢- بناء مفردات الاختبار .

٣- تحليل المفردات .

- نادية محمد عبد السلام : **بناء اختبار هدفى المرجع وجماعى المرجع (دراسة مقارنة بين النموذجين)** . كتاب اتجاهات معاصرة فى القياس والتقويم النفسى ، القاهرة ، مكتبة الأنجلو المصرية ، ط١ ، ١٩٩٦ ، من ص٨٦ إلى ص١١٩ .

٤- انتقاء المفردات / طول الاختبار / الدرجات الفاصلة .
٥- دراسات الثبات والصدق .

وحددت الباحثة معيار الأداء فى هذا الاختبار واحد بالنسبة لجميع الأهداف وهو ١٠٠٪ . أما المختبر الذى لم تصل إجابته لهذه النسبة المئوية يعتبر غير متقن حتى ولو كانت النسبة المئوية لإجابته الصحيحة ٩٠٪ أو ٩٥٪ .

ومن أهم **النتائج** التى توصلت إليها الدراسة :-
١- أن الوصول للتعلم من أجل الإتقان لم يتحقق بالقياس جماعى المرجع .
٢- أن تشخيص نقاط القوة والضعف لا يتحقق إلا بالاتجاه الأديومترى الذى يعنى بتفسير الدرجات الخام ، لأن لها معنى فى علاقتها بالمحك أو الهدف الذى هو مستقل عن الفروق الفردية .
٣- لا يمكن تفسير الدرجة الخام فى حالة الاتجاه السيكومترى ، لأنها تؤسس بغرض : إعطاء مقارنات بين الأفراد .

تميزت هذه الدراسة عن الدراسات السابقة فى أنها حددت معيار الأداء فى هذا الاختبار ١٠٠٪ ، واعتبار المختبر الذى لم تصل إجابته لهذه النسبة المئوية (غير متقن) . وهذا ما اختلفت فيه مع الدراسات السابقة عليها بل وعلى الدراسة الحالية أيضاً . وكذلك استخدامها عند بناء الاختبار الخلاصة المعدلة لفريمر (١٩٧٤) .

كذلك لاحظت الباحثة أن هذه الدراسة استخدمت نفس المحتوى (مادة الإحصاء الوصفى) ونفس العينة (طلاب الدبلوم الخاص وتمهيدى الماجستير) فى دراستها السابقة (١٩٩٢) .

وتختلف هذه الدراسة عن الدراسة الحالية فى تحديد معيار الأداء ١٠٠٪ ، وفى نوعية العينة ، والموضوع ، واستخدام صورتين متكافئتين للاختبار .

واتفقت هذه الدراسة مع دراسة كل من محمد السيد (١٩٩٢) ، ونادية عبد السلام (١٩٩٢) ، وآمال أبو حديد (١٩٩٤) ، فى حجم العينة الصغير .

وكذلك مع دراسة الشافعى (١٩٨٥) ، ونادية عبد السلام (١٩٩٢) ، فى إعداد صورتين متوازيتين (أ،ب) للاختبار .

واتفقت مع دراسة كل من الفقى (٨٢) ، وتاج السر (٨٦) ، وماسترز و ايفانز (١٩٨٦) ، ومحمد السيد (١٩٩٢) ، وآمال أبو حديد (٩٤) ، محمد فتح الله (٩٥) ، بأن الاختبارات المحكية المرجع فى تشخيص نقاط القوة والضعف ووصول المتعلم إلى مستوى الإتقان من خلال هذه الاختبارات .

١٤ – دراسة إسماعيل حسن فهيم (١٩٩٦)

هدفت الدراسة إلى :-

التعرف على أثر ثلاث طرائق للتغذية المرتدة الاختبارية على ثبات درجات اختبار هدفى المرجع فى مادة حساب المثلثات لطلاب الصف الأول الثانوى ، والتعرف على أثر هذه الطرائق من التغذية المرتدة فى استبقاء المعلومات لدى هؤلاء الطلاب .

منهج الدراسة :- استخدم الباحث المنهج التجريبى ، وتكونت عينة البحث من البنين فقط بالصف الأول الثانوى ، بمدرسة بلبيس الثانوية ، بمحافظة الشرقية ، وقد قام باختيار ثمانية فصول بطريقة عشوائية ، وزعهم الباحث على أربع مجموعات ، وبلغ العدد الكلى (٢٨٥) طالب . واقتصر الاختبار على وحدة دراسية فقط من وحدات مادة حساب المثلثات ، وهى (الدوال المثلثة) .

وقد اتبع الباحث الخطوات التالية فى بناء اختبار (**هدفى المرجع**) :-

١- تحديد المحتوى المراد قياسه .

٢- تحديد الأهداف العامة التى يقيسها الاختبار .

٣- تحليل الأهداف العامة إلى أهداف سلوكية .

٤- صياغة الأهداف السلوكية .

٥- إعداد مواصفات الاختبار .

٦- كتابة مفردات الاختبار .

٧- تحديد صدق محتوى الاختبار وثبات درجاته .

---

[١٤] إسماعيل حسن فهيم إسماعيل الوليلى : **أثر استخدام بعض طرق التغذية المرتدة على ثبات درجات اختبار هدفى المرجع وعلى استيفاء المعلومات** ، رسالة ماجستير غير منشورة ، كلية التربية – جامعة الأزهر ، ١٩٩٦ .

- وقد تحقق الباحث من صدق محتوى الاختبار باستخدام صدق المحكمين .
- واستخدم الباحث لتحديد ثبات الاختبار معامل (كابا لسوامينا ثان، وهامبلتون، و ألجينا).

ومن أهم النتائج التى توصلت إليها الدراسة :-

- أن التغذية الراجعة الكلية وكذلك المؤجلة ٢٤ ساعة تزيد من ثبات درجات الاختبارات هدفية المرجع ، وتزيد من درجة احتفاظ الطلاب بالمعلومات .
- أن استخدام طرائق التغذية الراجعة وسيلة مهمة تسهم فى سير عملية التعلم .
- أن الاختبار (هدفى المرجع) له فائدة تشخيصية حيث إنها توضح جوانب القوة والضعف فى تحصيل الطلاب أثناء تعلمهم المهارات الأساسية لمادة حساب المثلثات .

- يمكن الاستفادة من الاختبار (هدفى المرجع) فى تحسين عملية تقويم طلاب الصف الأول الثانوى فى مادة حساب المثلثات،وإثراء تعلمهم لهذه المادة الهامة.

تميزت هذه الدراسة باستخدام ثلاث طرائق للتغذية المرتدة ؛ للتعرف على أثرها فى ثبات درجات الاختبار (هدفى المرجع) .

وقد اتبع الباحث فى خطوات بنائه للاختبار **(هدفى المرجع)** نفس خطوات بناء الاختبار محكى المرجع ، وهو ما استفادت منه الباحثة عند بناء الاختبار .

واتفقت مع دراسة (عبد المسيح) ، (٨٢) ، الشافعى (٨٥) ، تاج السر (٨٦) ، محمود إبراهيم (٩٠) ، صلاح علام (٩١) ، محمد السيد على (٩٢) ، محمد فتح الله (٩٥) ، بالتأكد من صدق الاختبار باستخدام صدق المحكمين ، و مع دراسة عبد المسيح (٨٢) ، محمد الشافعى (٨٥) ، محمود إبراهيم (٩٠) ، محمد فتح الله (٩٥) ، فى طريقة تحديد ثبات الاختبار، واتفقت فى استخدام طريقة معامل (كابا لسوامينا ثان - هامبلتون - ألجينا) مع كل من محمود إبراهيم (٩٠) ، ومحمد فتح الله (٩٥) ، لتحديد ثبات الاختبار . ومع كل من الفقى (٨٢) ، تاج السر (٨٦) ، محمد السيد على (٩٢) ، محمد فتح الله (٩٥) ، ونادية عبد السلام (٩٦) ، فى أهمية الاختبارات المحكية المرجع ، هدفيه المرجع فى تشخيص جوانب القوة والضعف فى تحصيل الطلاب أثناء تعلمهم .

واختلفت مع الدراسة الحالية فى نوع العينة (المرحلة الثانوية) واقتصارها على البنين .

ثانيا : دراسات تناولت تقنين الاختبار

واشتملت هذه الدراسات على :-

١- دراسة سليمان محمد سليمان (١٩٨٦)

٢- دراسة رشيل Russell (١٩٨٦)

٣- دراسة آمال محروس (١٩٨٨)

٤- دراسة زيكى Zieky (١٩٨٩)

٥- دراسة براون Brown (١٩٩٠)

٦- دراسة ناديه عبد السلام (١٩٩٠)

٧- دراسة كينان Kunnon (١٩٩٢)

٨- دراسة احمد مختار شبارة (١٩٩٣)

٩- دراسة هدسون Hudson (١٩٩٣)

١ - دراسة سليمان محمد سليمان (١٩٨٦)

هدفت الدراسة إلى :

دراسة الطرائق المختلفة لتقدير الدرجات والطرائق المختلفة لحساب ثبـــات الاختبـــارات المرجعية المحك .

منهج الدراسة استخدم الباحث لطرائق تقدير الدرجات المنهج التجريبـــى بالاســـتعانة بعينـــة استطلاعية ، بلغت (١٨٠) طالب من الصف التاسع الأساسى ، والاستعانة بعينة نهائية بلغـــت (١١٦٠) طالب ، قسمها الباحث إلى قسمين متساويين ، طبق عليهم الصورتين المتكافئتين مـــن الاختبار التحصيلى المرجعى المحك فى الجبر للصف التاسع الأساسى ، وهو من إعداد الباحث . ثم قام الباحث بتقسيم العينة إلى أربع مجموعات ، وقام بتصحيح كل مجموعة بطريقة مـــن الأربع الطرائق التالية :-

١- الطرائق التقليدية.

٢- الدرجات الذاتية .

٣- وزن الثقة .

٤- توزيع الثقة.

وقام بحساب ثبات درجات الاختبار بثلاث طرائق (ليفنجســـتون - هـــامبلتون - هـــارس و ألجينا).

ومن أهم النتائج التى توصلت إليها الدراسة :-

١- أن الثبات المحسوب بالطرائق المختلفة يختلف باختلاف الطريقة المستخدمة لتقدير الدرجات للاختبارات المحكية المرجع .

٢- أن طريقة ليفنجستون تعد أفضل الطرائق المستخدمة عند مقارنة طرائق حساب الثبات بالطرائق التقليدية، وأنه لا توجد فروق بين طريقة هاريس وطريقة هامبلتون وألجينا .

٣- لا توجد فروق بين ليفنجستون وطريقة هارس وطريقة هامبلتون وألجينا عند استخدام أسلوب الدرجات الذاتية.

٤- لا توجد فروق بين طريقة ليفنجستون وطريقة هارس وطريقة هاملبتون وألجينا عند

---

[١] سليمان محمد سليمان محمود : **أثر بعض طرق تقدير الدرجات على ثبات الاختبارات المرجعية المحك ذات الاختيار من متعدد** . رسالة ماجستير غير منشورة ، كلية التربية - جامعة الزقازيق ، ١٩٨٦م .

استخدام طريقة وزن الثقة.

٥- لا توجد فروق بين طريقة لينفنجستون وطريقة هارس . وأن طريقة ليفنجستون أفضل من طريقة هامبلتون وألجينا . وطريقة هارس أفضل من طريقة هامبلتون وألجينا عند استخدام طريقة توزيع الثقة.

٦- أن هناك تأثيرا مشتركا بين طرائق تقدير الدرجات وطرائق حساب الثبات .

تميزت هذه الدراسة بأنها قامت بمقارنة أربع طرائق لحساب ثبات الاختبار المحكى المرجع .

وتعتبر هذه الدراسة من الدراسات التى استخدمت العينة الكبيرة، حيث بلغت (١١٦٠) طالبا . وقسمت إلى قسمين متساويين، طبقت عليهما صورتان متكافئتان للاختبار .

٢ - **دراسة رشيل** (١٩٨٦) Russell

هدفت الدراسة إلى :

وضع أساليب لتقديرات النجاح من الاختبارات المحكية المرجع .

وقد حددت الدراسة خطوات القياس المحكى المرجع كالآتى :-

أ - المواصفات :- ١- تحديد الأهداف

٢- تحديد استراتيجية صناعة القرار الخاص بتقديرات الاختبار .

ب- التطبيق :- ٣- جمع بيانات عن أداء الطالب وسجل بمستواهم التحصيلى .

٤- تطبيق القرارات والأحكام وكتابة تقرير عن ذلك .

جـ- المراجعة :- ٥- إعادة تحليل بيانات الطالب لتقييم :-

أ - شمولية ، ووضوح ، ومناسبة الأهداف .

ب- صدق أداة التقييم .

جـ- آثار وتأثيرات صناعة قرارات الأحكام .

وركزت الدراسة على النقطة الثانية وهى تحديد استراتيجية صناعة القرار الخاص بتقديرات الاختبار وحددها الباحث عن طريق خمس استراتيجيات :-

---

٢ **Russell A . Docking . " Criterion - Referenced Grading Techniques " In : Studies in Educational Evaluation , V. 2 , PP 281 - 294 , printed in Great Britain , 1986 .**

١- النسبة المئوية :- وهى الطريقة السهلة والبسيطة للتعامل مع الأعداد الكبيرة لبيان نسبة ما تحقق من أهداف، ثم ترجمة هذه النسبة إلى تقدير رمزى (A-B-C-D-E) ويسمى هذا التقدير محكى المرجع ؛ لأن التقدير أعطى بناء على محك معين أو معيار هو تحقيق الهدف .

٢- الاعتماد على ترتيب الأهداف من السهل إلى الصعب، وذلك فى شكل تدرج فى القدرات فى موضوع معين ، وبذلك يصبح للنسبة المئوية معنى طيب .

٣- استخدام التقدير ذى المحك أو التقدير المحكى ، وهو يزيل الحواجز بين الأهداف ، وكل مجموعة من الأهداف حددت فى صورة مستويات أدائية .

٤- تقسيم الأهداف إلى موضوعات ، وفى كل موضوع حددت المستويات الدنيا للأداء ، وذلك بالاعتماد على أهميتها النسبية

٥- تحديد المجموعة المهمة من الأهداف بواسطة مصمم المقرر . ويعنى الرسوب فى أى هدف رسوباً فى المقرر كله . ويعنى بالأهداف المهمة الأهداف الأساسية التى تمثل أهمية شديدة للتعليم فى المستقبل ، أو تمثل مهارات خاصة .

ومن أهم النتائج التى توصلت إليها الدراسة :- أن الاستراتيجية الخامسة يمكن تطبيقها فى المقاييس المحكية المرجع حيث تنفرد بمميزات كثيرة منها :-

- الموضوعية ، المرونة ، الفائدة ، الوضوح ، البساطة .

وقد تميزت هذه الدراسة بأنها حاولت إلقاء الضوء على أهم الاستراتيجيات ، والتى يمكن استخدامها فى تقدير النجاح فى الاختبارات المحكية المرجع .

والتى تتلخص فى :-

١- النسبة المئوية.

٢- الاعتماد على ترتيب الأهداف من السهل إلى الصعب.

٣- استخدام التقدير المحكى .

٤- تقسم الأهداف إلى موضوعات .

٥- تحديد المجموعة المهمة من الأهداف .

وستستفيد الباحثة من هذه الدراسة فى أنها ستحدد مجموعة الأهداف المهمة التى ستمثل أهمية للتعليم والتلميذ فى المرحلة الابتدائية ، والتى هى ترجمة عن أهم ما ستتفق فيه الباحثة

مع هذه الدراسة عند تقدير النجاح فى الاختبار ، وهو اختبار التلميذ لمجموعة الأهداف الموضوعة كمحك أداء ، وهى الاستراتيجية الخامسة التى أوصت بها الدراسة .

ولكن **ستختلف** الدراسة الحالية مع هذه الدراسة فى أن الرسوب فى أى هدف لن تعتبره الباحثة رسوباً فى المقرر كله ، ولكن ستعتبره الباحثة قصوراً لدى التلميذ، يمكن أن يساعد فى تشخيص مواطن الضعف لديه والاهتمام بها ، حتى نصل به إلى درجة الإتقان و التمكن .

٣ – آمال محروس (١٩٨٨)

هدفت الدراسة إلى :

مقارنة أربع طرائق لحساب ثبات الاختبارات المحكية المرجع بحيث تعطى أقل خطأ ممكن بأقل مجهود .

منهج الدراسة : اتبعت الباحثة المنهج التجريبى على عينة من الممتحنات قسمتهن إلى ثلاث مجموعات متجانسة من طالبات الكليات المتوسطة للبنات فى المملكة العربية السعودية ، وقامت الباحثة بتطبيق ثلاث صور متكافئة من إعدادها لاختبارات تحصيلية موضوعية مرجعية المحك لأفرع مادة علم النفس للصف الثانى بكليات البنات المتوسطة – بالمملكة العربية السعودية . وكانت الأفرع :-

١– التقويم والقياس النفسى ،وبلغت العينة (٦٦) طالبة .

٢– اختبار علم النفس العام ،وبلغت عينته (١٥٢) طالبة .

٣– اختبار علم نفس التعلم ،وبلغت عينته (١١١) طالبة ، وقد جعلت الباحثة درجة القطع ٦٠٪ .

ومن أهم النتائج التى توصلت إليها الدراسة :-

١– إذا تمت المقارنة بين الطرائق الثلاث التى تعتمد على صورة واحدة للاختبار فى إيجاد معامل الثبات، وهى طريقة (لفنجستون – هارس – جيكوفياك) نجد أن طريقة (لفنجستون) تعطى ثباتاً أعلى من ثبات الطريقتين المذكورتين .

٢– إذا تمت المقارنة بين طريقة هامبلتون التى تعتمد على صورتى الاختبار المتكافئين والطرائق الثلاث (لفنجستون – هارس – جيكوفياك) نجد تفوق طريقة هامبلتون .

---

[٣] – آمال أحمد محروس : **مقارنة عدة طرق لحساب ثبات الاختبارات المرجعية المحك** . رسالة دكتوراه غير منشورة، كلية التربية – جامعة الزقازيق ، ١٩٨٨ .

تميزت هذه الدراسة بأنها قامت بمقارنة أربع طرائق (هامبلتون - لفنجستون - هارس - جيكوفياك) لحساب ثبات الاختبارات المحكية المرجع لإعطاء أقل خطأ ممكن بأقل مجهود .

وتتفق هذه الدراسة من الدراسة الحالية فى نوعية الاختبار (الموضوعى) . **وستستفيد** الباحثة من هذه الدراسة عند بناء الاختبار وتقنينه ، ومع دراسة سليمان (١٩٨٦) فى أنها قامت باختبار ثلاث طرائق من أربع استخدمها أيضاً ، وهى (لفنجستون - هاريس - وهامبلتون) .

واختلفت هذه الدراسة عن الدراسة الحالية ودراسة سليمان (١٩٨٦) فى العينة ، والموضوع ، والأدوات ، وإعداد ثلاث صور متكافئة لثلاث مقررات فى علم النفس ، و فى تحديد درجة القطع بـــ ٦٠٪ .

٤ - دراسة زيكى (١٩٨٩) Zieky

هدفت الدراسة إلى :

تحديد مستويات الأداء فى الاختبارات المحكية المرجع ، وذلك من خلال التركيز على النقاط الثلاث الآتية :-

١ - **ذاتية الدرجة المعيارية** :- بينت الدراسة على أنه لا يوجد طريق موضوعى دقيق لتحديد الدرجة المعيارية ، وذلك لأنها تعتمد على الذاتية ، كذلك لا توجد طريقة لإثبـــات أى درجـــة معيارية تعتبر صحيحة ؛ لأن الأساس الذى بنيت عليه هذه الطريقة تعتمـــد علـــى التقديـــرات الشخصية ، أو التقييم الشخصى . وحيث إن تحديد المستوى يعتمد على تقديـــرات شخصيـــة، والطلاب مختلفين، فبالتالى تتباين تقديراتهم وقيمهم المختلفة ؛ فمن المستحيل تجنب المتناقضات عند تحديد الدرجة المعيارية . فالدرجة المعيارية عملية بنائية وليست استكشافية .

٢ - **الطرائق التقليدية** :- مثل تقدير الدرجة المعيارية (تقدير افتراضى ٧٠٪ أو ٨٠٪) بناءً على تشكيل لجنة من الخبراء لتحديدها ، وعادة ما تكون مهمتهم المواءمة بين رأيـــن مؤيـــد للدرجة المرتفعة ٨٠٪ للحفاظ على المستوى العالى للتربية، والرأى الآخـــر يدعـــى أن هـــذه

---

[4] Zieky , M . Methods of Setting Standards of Performance on Criterion Referenced Tests . In : Studies in Educational Evaluation , v. 15 , PP335-338. Printed in Great Britain, 1989 .

٢- **الطرائق التقليدية** :- مثل تقدير الدرجة المعيارية (تقدير افتراضى ٧٠٪ أو ٨٠٪) بناءً على تشكيل لجنة من الخبراء لتحديدها ، وعادة ما تكون مهمتهم المواءمة بين رأيين مؤيد للدرجة المرتفعة ٨٠٪ للحفاظ على المستوى العالى للتربية، والرأى الآخر يدعى أن هذه الدرجة غير معقولة . وهذه الطريقة فشلت فى الوضع فى الاعتبار صعوبة بعض الأسئلة المستخدمة فى أغراض متعددة .

٣- **الطرائق المؤسسة على التحكيم على الأسئلة** :- وتعتمد هذه الطريقة على مفهوم الخط الفاصل **Borderline** للمختبر ، وهو تحرى الدقة والعدالة التامة ، فى تقدير الدرجات التى تضع الممتحنين قرباً من حدود النجاح أو الرسوب . وهذه الدرجة الفاصلة تحدد من خلال نتائجها مجموعة من الأسئلة المحددة فى الاختبار ذاته .

وقد توصلت الدراسة إلى :- أن الطرائق الثلاث السابقة مناسبة لتحديد مستويات الأداء فى الاختبارات **CRT** ، ويمكن تطبيقها إما قبل أو بعد الاختبار . كما أن عملية التحكيم على الأسئلة تجعل المحكم يركز عن قرب على محتوى الاختبار ، وأن هذه الطرائق توفر البيانات التى يؤسس عليها التحكيم .

اعتماد كل طريقة من الطرائق الثلاث السابقة على خمس خطوات هى :-

١- اختيار المحكمين

٢- تحديد الـ Borderline للمعلومات والمهارات .

٣- تدريب المحكمين على استخدام هذه الطرائق .

٤- تجميع الأحكام .

٥- تجميع الأحكام لتقرير درجة النجاح .

كما أكدت الدراسة على أهمية جلوس المحكمين مع بعضهم، لتقدير ما يعنون بـ **Borderline** ، وعلى ضرورة إلمامهم التام بما يقيسه الاختبار، وكيف توظف درجاته، وعلى أهمية أن يصف المحكمون بدقة تامة للشخص الذى تمثل مهاراته الـ **Borderline** .

تميزت هذه الدراسة بأنها حاولت إلقاء الضوء على أهم النقاط اللازمة لتحديد مستويات الأداء فى الاختبارات المحكية المرجع وهى :-

١- ذاتية الدرجة المعيارية

٢- الطرائق التقليدية

٣- الطرائق المؤسسة على التحكيم على الأسئلة .

وقد أفادت الباحثة من هذه الدراسة فى الإطار النظرى للبحث ، واستخدام أى الطرائق السابقة فى تحديد مستوى التلاميذ التى يمكن تطبيقه فى الاختبارات، ومدى تحقيق النقاط منها مع

الاهتمام بالخطوات الخمس اللازمة لتحرى الدقة فى اختيار المحكمين ، أو اختيارهم من الأساتذة المتخصصين فى المجال؛ حتى يكون تحديد مستوى الأداء دقيق .

وقد اتفقت هذه الدراسة مع الدراسات السابقة فى تحديد الدرجة الفاصلة باتخاذها الطريقة التقليدية المشار إليها فى هذه الدراسة عند تحديدهم للدرجة الفاصلة المعتمدة على آراء الخبراء.

٥ - دراسة براون Brown (١٩٩٠)

هدفت الدراسة إلى :

تقديم تقديرات الدرجة الفاصلة لاتساق الاختبارات المحكية المرجع باستخدام بيانات مصدرها تطبيق ثلاث إستراتيجيات على الطلاب الدارسين لمقرر القراءة فى مادة اللغة الإنجليزية كلغة ثانية .

عينة البحث :- تكونت العينة من (١٠٠) طالب من طلاب جامعة هاواى فى (مانوا)، وقسموا إلى قسمين : متقنين وغير متقنين .

أدوات الدراسة :-

- اختبار محكى المرجع طبق مرتان ، بشرط وجود فترة زمنية بين تطبيق الاختبارين لتجنب تأثير التدريب على حل الاختبار ، وقد فرضت مسبقا الدرجة الفاصلة لتصنيف الطلاب إلى متقنين وغير متقنين على أساس مدى تحقيق الطلاب لهدف الاختبار .

- استخدام ثلاث استراتيجيات فى تحليل بيانات الاختبارات المحكية المرجع هى :-

١- استخدام مدخل الفروق حول العتبة الفارقة، والذى يتم حسابه باستخدام معامل الإتقان ومعامل الارتباط كابا، وهذه المعاملات تستخدم فقط فى حالة حساب الثبات بطريقة إعادة الاختبار .

٢- استخدام مدخل الفرق حول مربع الخطأ، والذى يحسب باستخدام معامل فاى، ومعامل الارتباط الرتبى، ويسمى معامل لامدا .

Brown, D. J. Short - Cut Estimators Of Criterion Referenced Test Consistency . In : Language Testing , 7, PP 77- 97 , Printed in Great Britain, 1990.

٣- مدخل استقلالية الدرجة ويستخدم طريقة التجزئة القصيرة؛ لتقدير القيمة العامة لمعامل فاى. وهذا أسلوب اقترحه بيرنون (١٩٨٣) .

ومن أهم ما توصلت إليه الدراسة :-

١- أن الطرائق الثلاثة المستخدمة لحساب الاتساق الداخلى للاختبار المحكى المرجع تؤدى إلى نفس النتيجة عند حساب الاتساق الداخلى للاختبار .

٢- كان من المتوقع أن الطريقة الثالثة (مدخل استقلالية الدرجة ويستخدم طريقة التجزئة القصيرة لتقدير القيمة العامة لمعامل فاى) هى الأنسب فى حساب معامل الاتساق الداخلى فى اختبارات اللغة المحكية المرجع ، ولكن ثبت أن استخدام الطريقة الأولى (مدخل الفروق حول العينة الفارقة، والذى يتم حسابه باستخدام معامل الإتقان ومعامل الارتباط كابا)، هو أكثر الطرائق انتشاراً على مستوى العالم فى حساب معامل الاتساق الداخلى فى اختبارات اللغة المحكية المرجع .

وقد تميزت هذه الدراسة بأنها ألقت الضوء على تقديرات الدرجة الفاصلة لاتساق الاختبارات المحكية باستخدام ثلاث استراتيجيات خاصة باللغة . هى :-

١- استخدام الفروق حول العتبة الفارقة .

٢- استخدام مدخل الفرق حول مربع الخطأ .

٣- مدخل استقلالية الدرجة (التجزئة القصيرة) .

وتتفق هذه الدراسة مع الدراسة الحالية فى تقسيم التلاميذ إلى متقنين وغير متقنين على أساس مدى تحقيقهم لأهداف الاختبار، ولكن بعد الانتهاء من أداء الاختبار .

واختلفت هذه الدراسة مع الدراسة الحالية فى العينة ، وموضوع الدراسة، كما أنها اختلفت مع دراسة كل من سليمان (١٩٨٦) ، وآمال محروس (١٩٨٨) فى النقاط ذاتها . وكذلك فى تحديدها مسبقاً للطلاب كمتقنين وغير متقنين قبل إجراء الاختبار . ولكن ستحدد الباحثة هذه النقطة بعد تطبيق الاختبار، وهذا ما اختلفت فيه أيضا مع دراسة آمال محروس (١٩٨٨) .

٦ – دراسة نادية عبد السلام (١٩٩٠)

هدفت الدراسة إلى :-

استخدام الاختبارات المحكية المرجع فى مجالات تقدير الفرد وتقويم أدائه .

منهج الدراسة : استخدمت الباحثة المنهج التجريبى على عينة اشتملت على (٤٩) طالباً وطالبة من طلاب الدراسات العليا (تمهيدى ماجستير و دبلوم خاص) فى كلية البنات – جامعة عين شمس (٨٨-٨٩) الدارسين لمادة الإحصاء الوصفى . وقد قسمت الباحثة المادة إلى ستة موضوعات ، قامت بترتيبهم من (القمة إلى القاعدة) ، وقد قسمت كل موضوع إلى أربعة مستويات من حيث تحقيق الهدف؛ لتحدد أى مستوى وصل إليه الطالب أو الطالبة فجعلت :-

المستوى الأول : يمثل عدد الأفراد الذين حصلوا على درجات من صفر إلى ما قبل ٥٠٪

المستوى الثانى : يمثل عدد الأفراد الذين حصلوا على درجات من ٥٠٪ إلى ما قبل ٧٠٪

المستوى الثالث : يمثل عدد الأفراد الذين حصلوا على درجات من ٧٠٪ إلى ما قبل ٨٠٪

المستوى الرابع : يمثل عدد الأفراد الذين حصلوا على درجات من ٨٠٪ إلى ١٠٠٪

ومن أهم النتائج التى توصلت إليها الدراسة :-

١- أن التفسيرات المحكية المرجع تساعدنا على شرح ماذا يستطيع الفرد عمله دون الرجوع إلى أداءات الآخرين .

٢- شرح الأعمال التعليمية المحددة التى يستطيع أداءها الطلاب .

٣- مقارنة أداء الاختبار لمعيار معين واتخاذ قرار إتقان أو عدم إتقان .

٤- نستدل على النسبة المئوية للأعمال التى أنجزها الطلاب بدقة .

تميزت هذه الدراسة فى أنها ساهمت فى إلقاء الضوء على أهمية الاختبارات المحكية فى اتخاذ قرارات الإتقان وعدم الإتقان لدى الطلاب والتلاميذ ، وأهميته فى تحديد النسب المئوية لإنجاز الطلاب بدقة .

٦- نادية محمد عبد السلام : **القياس محكى المرجع وتطوير التعليم** "دراسة تجريبية" .مجلة علم النفس ، العدد الرابع عشر ، السنة الرابعة ، ١٩٩٠ ، ص٤٤ : ص٦٤ .

واتفقت هذه الدراسة مع الدراسة الحالية فى تقسيم المنهج إلى موضوعات، ومن ثم إلى أهداف عند بناء الاختبار المحكى المرجع .

واختلفت هذه الدراسة مع الدراسة الحالية فى نوع العينة ، والموضوع ، وعدد أفراد العينة ، وكذلك فى تقسيم كل موضوع من الموضوعات (الستة) إلى أربع مستويات وفى ترتيب الموضوعات من القمة إلى القاعدة . وهذا ما اختلفت فيه أيضا مع كل من سليمان (١٩٨٦) ، ورشيل (١٩٨٦) ، وآمال محروس (١٩٨٨) ، وزيكى (١٩٨٩) ، بروان (١٩٩٠) .

٧ – **دراسة كينان** (١٩٩٢) Kunnan

هدفت الدراسة إلى :

استخدام ثلاث طرائق لتحليل إجراءات اختبار محكى المرجع فى اللغة الإنجليزية باستخدام نظرية التعميم ، والتحليل العاملى ، والعنقودى ، للتحقق من اعتمادية وصدق الاختبار وتطوره فى جامعة كاليفورنيا بلوس أنجلوس عام ١٩٨٩ .

عينة البحث :- اشتملت عينة البحث على (٣٩٠) دارسا من السكان غير الأصليين المتحدثين للغة الإنجليزية والدارسين للغة الإنجليزية كمتطلب لقبولهم بالجامعة .

أدوات الدراسة :- اختبار محكى المرجع استخدموا فيه (الاختيار من متعدد ، الصواب والخطأ) وتكون من (١٠٠) مفردة قسمت كالآتى :-

١) ٣٠ مفردة للاستماع .

٢) ٤٠ مفردة للقراءة والتعرف على الكلمات الصعبة .

٣) ٣٠ مفردة للقواعد والتعبير .

وحدد وقت الاختبار بـــ (٩٠) دقيقة لكل (١٠٠) مفردة ، بالإضافة إلى (٣٠) دقيقة لموضوع التعبير .

– وتم التأكد من صدق الاختبار عن طريق تحليل المحتوى والصدق البنائى .

---

[٧] Kunnan, A. J. An investigation of a criterion - referenced test using G . theory , and factor and cluster analyses , In : Language Testing V. q , n1, , PP30 49 , Printed in Great Britain , June , 1992.

ومن أهم ما توصلت إليه الدراسة : أن اعتمادية وصدق الاختبار المحكى المرجع لاختبار الدارسين للغة الانجليزية كمطلب لقبولهم بالجامعة (NESLPE) يمكن التوصل إليها باستخدام نظرية التعميم والتحليل العاملى والعنقودى . كما أكدت على ضرورة تحرى الدقة فى تطبيق الاختبار محكى المرجع (NESLPE) كضرورة لتطوير اعتمادية وصدق الاختبار ولتحديد الطول المناسب ومستويات الأداء للاختبار .

تميزت هذه الدراسة بمحاولة تطوير الاختبار المحكى المرجع فى جامعة كاليفورنيا ، وباستخدام ثلاث طرائق لتحليل إجراءات الاختبار .

واتفقت هذه الدراسة مع الدراسة الحالية فى التأكد من صدق الاختبار عن طريق تحليل المحتوى ، والصدق البنائى ، واختلفت فى نوع العينة .

وستستفيد الباحثة من هذه الدراسة تحرى الدقة فى تطبيق الاختبار ؛ حتى يكون اختباراً صادقاً محققا للأهداف التى وضعت من أجله .

واتفقت مع دراسة كل من آمال محروس (١٩٨٦) ، براون (١٩٩٠)، نادية عبد السلام (١٩٩٠) ، فى المرحلة التعليمية ، ومع براون (١٩٩٠) ، أيضا فى اتخاذ مقرر القراءة كمادة لبناء الاختبار .

واختلفت هذه الدراسة مع الدراسة الحالية فى المرحلة التعليمية. وكذلك فى نوعية أسئلة الاختبار .

٨ – دراسة أحمد مختار شبارة (١٩٩٣)

هدفت الدراسة إلى :

١- التعرف على مستوى إتقان الطالبات المعلمات للكفايات التعليمية المنشودة لبرنامج التربية العملية بالكلية المتوسطة للمعلمات بمسقط .

---

[٨] أحمد مختار شبارة : تقويم اكتساب الكفايات التعليمية فى ضوء مبادئ التعلم الإتقانى لدى طالبات الكلية المتوسطة للمعلمات بمسقط (دراسة مرجعية المحك لبرنامج التربية العملية) ، **التربية المعاصرة** ، العدد ٢٧ ، السنه العاشرة ، يونيو ١٩٩٣ ، ص ١٩٧ : ٢٤٤ .

٢- محاولة التعرف على مدى تأثر اكتساب هذه الكفايات بنوع التخصص الأكاديمى علمى - أدبى .

٣- محاولة التعرف على مدى فاعلية البرامج التدريبية العلاجية كأحد فعاليات برنامج التربية العملية فى اكتساب الطالبات للكفايات التعليمية ، التى يعانين ضعفاً فيها فى ضوء محك الإتقان المطلوب .

عينة الدراسة :- تكونت العينة الميدانية من (٩٣) طالبة معلمة من الطلبات المعلمات من جميع التخصصات العلمية والأدبية بالكلية المتوسطة للمعلمات بمسقط فى أثناء قيامهن بالتدريس الفعلى فى المدارس المتعاونة بمنطقة مسقط التعليمية .

أدوات الدراسة :- بناء أداة الدراسة الميدانية (بطاقة الملاحظة)

منهج الدراسة :- استخدم الباحث فى دراسته أحد أنواع المنهج الوصفى، وهو المنهج المسحى .

- وقد اعتبر الباحث محك الإتقان المطلوب فى هذه الدراسة هو ٨٥٪

ومن أهم النتائج التى توصلت إليها الدراسة :- أن البرنامج الحالى - للتربية العملية الميدانية بالكلية المتوسطة للمعلمات بمسقط - لا يؤدى دوره فى اكتساب الطالبات المعلمات للكفايات التعليمية المنشودة بالمستوى المطلوب، مقارنة بمحك الإتقان المطلوب لعملية إعداد المعلم، والتى نفترض أن يكون الحد الأدنى لامتلاك هذه الكفايات لدى المعلم (٨٥٪) على الأقل؛ حتى يكون أداؤه ناجحاً فى عملية التعليم والتعلم .

اعتمدت هذه الدراسة على تحديد مستوى الإتقان لدى الطالبات المعلمات بمحك الإتقان المطلوب وهو ٨٥٪ .

وقد لاحظت الباحثة أن الدراسات السابقة اعتمدت على الدرجات الفاصلة من ٨٠٪ فى الإعدادى والثانوى، ماعدا دراسة الفقى التى حدد مستوى الإتقان ٩٠٪ فى المرحلة الإعدادية، وأمال محروس (٦٠٪) فى الكليات المتوسطة فى السعودية . ولم تتطرق أية دراسة إلى مستوى الإتقان فى المرحلة الابتدائية . ولهذا ستستفيد الباحثة من هذه الدراسة عند تحديد درجة الإتقان لدى تلاميذ الصف الخامس الابتدائى .

واتفقت هذه الدراسة مع دراسة أمال محروس (١٩٨٨) فى المرحلة التعليمية.

٩ – **دراسة هدسون** (١٩٩٣) Hudson

هدفت الدراسة إلى :

تطوير تطبيق الاختبارات المحكية المرجع عملياً، وذلك بإيضاح العلاقة بين وظائف المعلومات الموجودة فى السؤال وثلاثة أنواع من اختبارات اللغة .

عينة البحث :

١- اشتملت العينة على (٤٦٣) ممتحن لاختبار التحصيل، ولهم خلفية لغوية متجانسة ، وهم ممن تجاوزوا اختبارات التويفول (TOEFL)، وتراوحت درجاتهم بين (٥٠٠ إلى ٥٩٩) درجة .

٢- (٢١٦) ممتحن من معهد اللغة الانجليزية بجامعة هاواى .

٣- (٢٠٦) دارس للهندسة الكيماوية فى الصف الثالث والرابع فى جامعة (Guadalajara)

أدوات الدراسة :- ثلاثة أنواع من اختبارات اللغة المحكية المرجع .

الاختبار الأول :- اختبار تحديد المستوى فى القراءة ، صمم لوضع الطلاب أو استثنائهم من المقررات الأكاديمية فى القراءة. وقد اشتمل على القراءة ست قطع كل منها مكونة من (٢٥٠) كلمة تقريباً. تبعت بأسئلة شاملة من اختيار من متعدد عددها من (٣:٥) أسئلة، بالإضافة إلى سؤال عن معانى (٢٥) كلمة .

الاختبار الثانى :- يحتوى على (٤٣) عبارة فى مقرر ذى مستوى عالٍ فى التحصيل فى القراءة . اشتمل على: القراءة السريعة ، اختيار الأفكار الرئيسة ، والقراءة للبحث عن التفصيلات ، قراءة الأفكار الأساسية ، تحديد المصطلحات وإشارات الكلام ، القراءة الناقدة .

الاختبار الثالث :- يتكون من (٢٥) عبارة – اختيار من متعدد – للطلاب المتحدثين الأسبانية ولأغراض أخرى ، واشتمل الاختبار على قطعة أساسية بها (٢٥) فراغاً تحت كل فراغ أربعة اختيارات .

---

٩ Hudson , T. Surrogate Indices for item Information Function in Criterion Referenced Language Testing . In : **Language Testing** , V.١٠ ,No ٢, . PP ١٧١- ١٩١ . Printed in Great Britain , ١٩٩٣ .

من أهم ما توصلت إليه الدراسة :- أن معاملى الصعوبة فاى ومعامل B (Btndex) يمكن أن يفيدا فى عملية اختيار الأسئلة ، وأنهما أظهرا علاقة ارتباطية بالمعلومات الموجودة فى السؤال (TTF) عند درجة قاطعة ٦٥٪ ، ٧٠٪ ، ٧٥٪ .

تميزت الدراسة بأنها حاولت تطوير تطبيق الاختبارات المحكية المرجع على أنواع من اختبارات اللغة .

وستستفيد الباحثة من هذه الدراسة فيما أشارت إليه من مهارات رئيسة للقراءة، والتى تناسب تلاميذ الصف الخامس من اختبار الأفكار الرئيسة ، والتعرف على معانى بعض الكلمات .

وقد **اتفقت** هذه الدراسة مع دراسة كل من آمال محروس (١٩٨٨) ، وبراون (١٩٩٠) ، ونادية عبد السلام (١٩٩٠) ، وصلاح علام (١٩٩١) ، وكينان (١٩٩٢) ، وأحمد شبارة (١٩٩٣) ، فى نوعية العينة (طلاب جامعات) ، ودراسة براون (١٩٩٠) ، وكينان (١٩٩٢)، ومع الدراسة الحالية فى اتخاذ مقرر القراءة كمادة لبناء الاختبار المحكى المرجع .
واختلفت هذه الدراسة مع الدراسة الحالية فى نوعية أسئلة الاختبار .

تعليق عام على الدراسات السابقة :-

١- تشابهت الدراسات التى هدفت إلى بناء اختبارات محكية المرجع أو هدفية المرجع ، فى استخدام معظم الخطوات الخاصة ببناء الاختبار .
٢- اعتمدت معظم الدراسات السابقة فى تصميم الاختبار على الاختبارات الموضوعية . واقتصار بعضها على (الاختيار من متعدد) منها دراسة عبد المسيح (١٩٨٢) ، محمود إبراهيم (١٩٩٠) ، صلاح علام (١٩٩١) ، هدسون (١٩٩٣) ، فتح الله (١٩٩٥) ، إسماعيل فهيم (١٩٩٦) .
٣- عدم اشتمال أية دراسة من الدراسات السابقة فى تصميم الاختبارات (لأسئلة المقال)
٤- اعتماد جميع الاختبارات على صدق المحتوى عند تحديد صدق الاختبار.
٥- تعدد طرائق تحديد ثبات الاختبار مثل استخدام طريقة (كيودر ريتشارسون) طريقة اتساق القرار ، ومعامل كابا لسواميناثان وهامبلتون و ألجينا ، ليفنجستون .
٦- استخدام طرق تقليدية تستخدم فى الاختبارات معيارية المرجع، للتأكد من ثبات الاختبار محكى المرجع .

٧- ركزت معظم الدراسات على اختيار العينة من طلاب وطالبات الجامعة ، والثانوى، وندرة الدراسات التى اختارت العينة من تلاميذ المرحلة الابتدائية ، بالرغم من أهمية هذه المرحلة. فيماعدا دراسه ماسترز وإيفانز (١٩٨٦) .

٨- قلة الاختبارات التى استخدمت اللغة عامة والعربية خاصة ، كمادة لبناء الاختبار مثل دراسات ريفنزوليفن (١٩٩٢) ، آمال أبو حديد (١٩٩٤) ، وبراون (١٩٩٠) ، وكينان (١٩٩٢)

٩- اتفقت معظم الدراسات على استخدام الدرجة الفاصلة ٨٠٪ ، وتفاوتت هذه الدرجة بين المرحلة الإعدادية والثانوية ، ماعدا دراسة الفقى (٨٢) ، فقد حددها بـــ ٩٠٪ للمرحلة الإعدادية ، وآمال محروس (٦٠٪) ، وهذا لطبيعة العينة (كليات متوسطة فى المملكة العربية السعودية) وتحديد نادية عبد السلام (١٩٩٦) ، بـــ ١٠٠٪ .

١٠- استخدمت بعض الدراسات العينات كبيرة العدد ١٠٠ طالب وطالبة فأكثر منها دراسة عبد المسيح (١٩٨٢) ، الدسوقى (١٩٨٥) ، تاج السر (١٩٨٦) ، فتح الله (١٩٩٥) ، واسماعيل فهيم (١٩٩٦) .

١١- استخدمت بعض الدراسات صورتين أو ثلاث صور لاختبارات متكافئة منها دراسة الدسوقى (١٩٨٥) ، ماسترزوايفانز (١٩٨٦) ، نادية عبد السلام (١٩٩٦) ، سليمان (١٩٨٦) ، آمال محروس (٨٨) .

١٢- خلصت أغلب الدراسات السابقة إلى أهمية الاختبارات المحكية المرجع ، هدفية المرجع فى :-

أ - استمرار تحسن سلوك التلاميذ التحصيلى .

ب- تشخيص جوانب القوة والضعف لدى التلاميذ فى مادة الاختبار نحو الامتحان .

جـ- الكشف عن الصعوبات والنقائص فى تعلم التلاميذ أولا بأول .

د - تقديم البدائل العلاجية المناسبة لطبيعة الفروق الفردية للتلاميذ لعلاج نقائص التعلم .

و- تخفيف قلق الدارس نحو الامتحان .

هـ- تحسن التحصيل بتكرار هذه الاختبارات، وهذا ما توصلت إليه دراسات كل من الفقى (١٩٨٢) ، عماد عبد المسيح (١٩٨٢) ، الشافعى (١٩٨٥) ، تاج السر (١٩٨٦) ، ماسترز و ايفانز (١٩٨٦) ، أمال أبو حديد (١٩٩٤) ، نادية عبد السلام (١٩٩٦) ، إسماعيل فهيم (١٩٩٦) .

١٣- اعتمدت معظم الدراسات على تقسيم بلوم فى تحديد مستويات الأهداف الإجرائيـــة مثـــل عماد عبد المسيح (١٩٨٢) ، الشافعى (١٩٨٥) ، آمال أبو حديد (١٩٩٤) ، محمود إبراهيـــم (١٩٩٠) .

١٤- وتميزت بعض الدراسات على مستوى البلاد العربية بأنها من أوائل الدراسات التى قامت ببناء هذه الاختبارات مثل عبد المسيح (١٩٨٢) ، والشافعى (١٩٨٥) .

وتميزت دراسة ريفيز وليفن (١٩٩٢) ، بأنها حاولت أن تطور فى أنموذج الاختبار المحكى وتميزت آمال أبو حديد (١٩٩٤) ، بأنها الأولى التى استخدمت الاختبار المحكـــى لتشخيـــص أخطاء استخدام علامات الترقيم فى اللغة الإ نجليزية.

١٥- توصلت معظم الدراسات إلى أن التلاميذ والطلاب لم يحققوا درجة الإتقان المحددة فى كل اختبار ، وأن نسبة الإتقان فى المستويات الدنيا (التذكر) أعلى من المستويات المعرفية الأعلى (الفهم - التطبيق - حل المشكلات.... الخ) منها دراسة عبد المسيح (١٩٨٢) ، الشافعى (١٩٨٥) .