البحث الثامن :

" اثـــــر بعض طرق تقدير الدرجات للمفردات على ثبات وصدق درجات اختبار تحصيلي في الرياضيات ذي الاختيار من متعدد لدى طالبات الصف الأول الثانوي بمكة المكرمـة"

إعداد :

أ / عفاف بنت راضي مشخص اللحياني

تخصص مناهج وطرق التدريس

جامعة أم القرى بمكة المكرمة

" اثـــر بعض طرق تقدير الدرجات للمفردات على ثبات وصدق درجات اختبار تحصيلي في الرياضيات ذي الاختيار من متعدد لدى طالبات الصف الأول الثانوي بمكة المكرمة"

أ / عفاف بنت راضي مشخص اللحياني

- **مقدمة**:

إن التقارب الشديد بين التعليم والاختبار جعل من الصعب الفصل بينهما فالمعلم لا يستطيع القيام بتعليم جيد ما لم يكن يحسن الاستفهام، والاستفسار والسؤال قبل وأثناء وبعد القيام بتعليم أي جزء من أجزاء الدرس (جان، ١٩٩٩م: ١٢).

ويعد التقويم إحدى الفعاليات الأساسية في نشاط النظم والمؤسسات التعليمية لضمان التأكد من سيرها في الاتجاه الذي يحقق أهدافها ويزيد من فاعليتها وكفاءتها وانسجام تفاعلها مع البيئة الخارجية على النحو الذي يؤدي إلى تطورها واستمرارها. (الدوسري، ٢٠٠٠م :٣٤)

ويؤدي التقويم دوراً هاماً بارزاً في مجال التربية والتعليم وقد تزايدت أهميته بتعدد وتنوع أساليب التربية وطرقها ومناهجها وأهدافها ووسائلها، وقد عم استخدام أساليب التقويم في المجال التربوي وصولا بالعائد والناتج إلى أعلى المستويات وتزايدت بذلك أهمية التقويم وتنوعت أساليبه وازداد دقة واكتمالا (العبيدي، الجبوري، ١٩٨١م:١٣).

والتطوير التربوي يعتمد اعتماداً كبيراً على التقويم والمتابعة ، فهو يبدأ بالقياس وينتهي بالتقويم لتطوير الواقع وتحديد مشكلاته والغرض الذي يرمي إليه كل قياس وتقويم هو اختبار صحة الفروض التي يقوم عليها التحديث والتجديد التربوي ويرمي القيام والتقويم في التربية إلى معرفة نواحي القوة والضعف في العملية التربوية كلها من أجل تحسينها كما وكيفا. (العبيدي، الجبوري، ١٩٨١م:١٣).

وتتأثر عملية إعطاء الاختبار بعوامل متعددة منها ما يتصل بمن يعطي الاختبار وما يفعله ومنها ما يتصل بمن يقوم بأداء معين على هذا الاختبار وما يحيط به من ظروف ومؤثرات، ومنها مايتصل بالاختبار نفسه وما يتميز به من خصائص وما يقصد إلى أن يقيس ، وإذا كانت هناك عوامل وشروط ينبغي أن تتوافر في الاختبار نفسه حتى يكون أداة قياس يعتمد عليها ويطمأن إليها. (أبو حطب، عثمان، ١٩٧٦م:٥٥).

ولا شك أن قياس تقدم الفرد في المهارات الأساسية والمفاهيم وما إليها من الجوانب المتعلقة بالحياة أمر هام، ولذلك ينبغي أن نعرف كيف يقوم الفرد في هذه النواحي إلى جانب التقويم لفهم الفرد وإدراكه للعلاقات وكيفية تطبيقه

للحقائق والمهارات التي تعلمها في حل مشكلات الحياة. (العبيدي، الجبوري ١٩٨١م:٩) ، ويعتبر الغرض من الاختبارات ترتيب الأفراد وتفسير أدائهم من خلال المقارنة. (الدوسري، ٢٠٠٠م : ١٢١).

إن إدخال مواضيع الاختبارات والقياسات في المناهج أمر ضروري ومهم، حيث يتطلب الأمر تطبيق بعض الاختبارات منها: اختبارات الذكاء، والاختبارات النفسية واختبارات التحصيل على الطلاب لمساعدة المختبر في ملاحظة الفروق الفردية أثناء التدريس وإرشادهم بالنسبة إلى قابليتهم والكشف عن ذكائهم وتوجيههم مهما اختلفت مواهبهم العقلية بحيث يمكن استغلال هذه المواهب استغلالا فعالا ومثمرا. (العبيدي، الجبوري، ١٩٨١م :٢).

ومن الاختبارات التي يمكن أن نقيس بها قدرات المتعلم هي الاختبارات الموضوعية وتظهر هذه الاختبارات الموضوعية في أنواع مختلفة كما أن كل نوع يتفرع إلى أشكال مختلفة وهذه الأنواع تتفاوت في تصحيحها ما بين الموضوعية المطلقة كما في الاختبار المتعدد الاختبارات إلى الأنواع التي يقترب تصحيحها من الذاتية، كما في بعض أنواع اختبارات الإكمال. (جان، ١٩٩٩م:١٢٣).

ويعتبر الاختبار متعدد الاختيارات من أشهر أنواع الاختبارات الموضوعية وأوسعها انتشارا وأكثرها استعمالا لقابلية استخدامها في تقويم أي نوع من أنواع المعرفة أو الخصائص الفكرية المختلفة لدى المتعلمين (حمدان، ١٩٨٠م :٢٨٢).

ويطلق على هذا النوع من الاختبارات أسماء عدة منها الاختبارات ذات الاختيارات المتعددة ومنها الاختبارات ذات الأجوبة المتعددة وغير ذلك من المسميات. (جان، ١٩٩٩م، ص١٢٦).

والميزة الكبرى للاختبار متعدد الاختيارات على غيره من الاختبارات الموضوعية تتجلى في قدرته على قياس تحصيل الطلبة بدقة كبيرة بالإضافة إلى قدرته على قياس مختلف المعارف والمعلومات والموضوعات بكفاءة تامة، كما أنه يمكننا من قياس المستويات المختلفة من الأهداف المعرفية بالإضافة إلى تحرره من النقائص والعيوب التي تتسم بها الأنواع الأخرى. (جان، ١٩٩٩م، ص١٢٦).

ويعد الصدق والثبات من اهم شروط الاختبار الجيد لذلك فإن الثقة في اختبار ما سواء كان جماعي المرجع أو محكي المرجع تعتمد على مقدار صدقه وثباته.

ويستخدم الثبات ليدل على مدى اعتماد الفروق الفردية في درجات الاختبار على أخطاء الصدفة المتضمنة في القياس من ناحية وعلى الاختلافات الحقيقية الأصلية في الصفة أو الخاصية السيكولوجية موضع الاهتمام من ناحية أخرى. (أبو حطب ، عثمان، ١٩٧٦م:٧٨).

ويعتبر التخمين أحد العوامل – إن لم يكن أهمها – التي تقلل من ثبات وصدق الاختبارات ذات الاختيار من متعدد لذلك اتجه أكثر الباحثين مثل سواقد ودودين وشريف عبد الوهاب وغيرهم إلى إتباع الطرق للقضاء على هذه المشكلة وبذلت عدة محاولات لرفع ثبات وصدق هذه الاختبارات بإتباع طرق معينة للتصحيح منها طريقة الدرجة الوزنية للاختبار (طريقة ديفز Davis) والطريقة التقليدية في تقدير درجات مفردات الاختبار وطريقة اختبار الثقة والطريقة اللوغارتيمية، وطرق التصحيح لأثر التخمين.

وقد تمت مقارنة هذه النظم ونظم أخرى بعضها ببعض ومقارنتها أيضاً بالنظام التقليدي القائم على أساس (صفر، ١) وذلك لمعرفة أي النظم تتمتع بثبات وصدق أكبر لدرجات الاختبارات ذات الاختيار من متعدد . ومن هذه الدراسات دراسة ربى (١٩٧٠م)، ودراسة روس وستيفن (١٩٩٤م)، ودراسة ريوز بريمو (١٩٩٨م)، ودراسة عبد الوهاب (٢٠٠١م)، ودراسة النبهان (٢٠٠٢م)، ودراسة الخرشلة (٢٠٠٤م).

وبالنظر إلى بعض الدراسات السابقة المشار إليها نجد أن :

- الدراسات السابقة والتي اتبعت نظام المجموعات المتعددة لم تراع تساوي هذه المجموعات في العدد وهذا تجاوز تجريبي له أثره في صحة النتائج.
- في الدراسات السابقة التي استخدمت أكثر من طريقة لتقدير الدرجات لم تراع اختلاف الزمن من طريقة إلى أخرى بل كان الزمن متساوياً في كل طريقة، وهذا تجاوز له أثره على النتائج.
- عدم إجراء الاختبارات على عينة استطلاعية لتحديد الزمن المناسب للاختبار والذي يؤثر على معامل الثبات والصدق.
- تضارب النتائج الخاصة بالثبات والصدق من دراسة إلى أخرى فمنها ما وجد أثراً للطرق المختلفة لتقدير الدرجات على الثبات والصدق، ومنها ما نفى وجود هذا الأثر.

- **مشكلة البحث**:

من المعروف أن عالمنا المعاصر يتميز بالسرعة والتغير كنتيجة للتطور العلمي وتطبيقاته التكنولوجية التي شملت جميع جوانب حياتنا اليومية التي لا بد أن يكون لها تأثير على عملية التربية والتعليم بعد أن اتسعت المدرسة وزاد عدد طلابها ومدرسيها واتسعت مناهجها وتغيرت أهدافها.

ومن هذا المنطلق فمدرسة اليوم بحاجة إلى التقويم والقياس المستمرين لبعض النواحي الهامة في نظامها وبرنامجها مثل الكشف عن استعدادات الطلبة وقدراتهم.

وفي ضوء هذا يمكن معرفة الواقع بنقاط قوته وضعفه والاستفادة من نتائج التقويم في تحسين وتطوير العملية التعليمية وهذا ما يدعونا إلى القيام بعمل الاختبارات الموضوعية وهذه الاختبارات تستخدم لتقويم أداء الفرد بالنسبة

لإطار سلوكي معين، أي مجموعة من المعارف والمهارات المعرفة بطريقة إجرائية (سلوكية) بصرف النظر عن علاقة أدائه بأداء غيره من الأفراد والذي يطبق عليهم نفس الاختبار (محمود، ١٩٨٦م:٢- ٣).

ولقد بذلت محاولات لرفع ثبات وصدق اختبارات تم بناؤها من قبل، حيث قام بعض المشتغلين بالقياس النفسي بدراسة الثبات والصدق لدرجات الاختبارات ذات الاختيار المتعدد وثبات الاختبارات وأدوات التقويم يقصد بها أن تكون على درجة عالية من الدقة والإتقان والاتساق فيما تزودنا بها من بيانات عن سلوك المفحوص.

ومن خلال نتائج بعض الدراسات حول طرق تقدير الدرجات وأثرها على ثبات وصدق الاختبارات (وايت وسابريس Sabers & Wihte 1969) ، (ثيسين Thissen، 1976)، (برانبوم وتاتسوك Brinbaum & Tatsuoka،1983)، (سليمان محمد سليمان، ١٩٨٦م)، (عادل محمد العدل، ١٩٨٦م) (طه الخرشله ، ٢٠٠٤م)، (النبهان، ٢٠٠٢م) والتي أشارت إلى قصور واضح في تناول هذا الموضوع بالإضافة إلا الاختلاف والتباين بين النتائج لكل الدراسات والبحوث مما دفع الباحثة إلى تناوله بالبحث والدراسة.

- **تساؤلات البحث**:

يمكن صياغة المشكلة في التساؤلات التالية:

7 هل يختلف ثبات الاختبار ذي الاختيار من متعدد باختلاف طرق تقدير الدرجات للمفردات؟
7 هل يختلف ثبات الاختبار ذي الاختيار من متعدد باختلاف طرق حساب هذا الثبات؟
7 هل يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق حساب الثبات على ثبات الاختبار ذي الاختيار من متعدد؟
7 هل يختلف صدق الاختبار ذو الاختيار من متعدد باختلاف طرق تقدير الدرجات للمفردات؟
7 هل يختلف صدق الاختبار ذي الاختيار من متعدد باختلاف طرق حساب هذا الصدق؟
7 هل يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق حساب الصدق على الاختبار ذو الاختيار من متعدد؟

- **أهمية البحث**:

- **أولاً: الأهمية النظرية**:

7 يحاول البحث الحالي إيجاد أثر بعض طرق تقدير الدرجات للمفردات على ثبات وصدق الاختبار ذو الاختيار من متعدد، حيث تعتبر هذه البحوث قليلة في البيئة العربية.

7 دراسة الطرق المختلفة لتقدير الدرجات بالإضافة إلى طرق حساب الثبات والصدق وكيفية تقنين الاختبارات.

- **ثانياً: الأهمية العلمية:**

وإلى جانب هذا، تتجلى الأهمية التطبيقية للبحث فيما تسفر عنه من نتائج حيث يمكن الاستفادة منها في:

7 يفيد البحث الحالي في تطوير منهج (الرياضيات) الحالي للصف الأول الثانوي.
7 الوصول إلى أفضل طريقة لتقدير الدرجات تتبع مع طريقة معينة لحساب الثبات والصدق لنوع معين من الاختبارات.
7 التوجيه التربوي للمعلمين وإرشادهم إلى طريقة تقدير الدرجات المثلى والتي تعطي أعلى صدق وثبات للاختبار.
7 تخدم الدراسة الحالية المسئولين بوزارة التربية والتعليم وخاصة المهتمين بأساليب التقويم في إعداد الاختبار ذو الاختيار من متعدد.
7 يسهم البحث الحالي في بناء اختبارات ذات اختيار من متعدد لمقررات دراسية أخرى.
7 دراسة الطرق المختلفة لتقدير الدرجات والطرق المختلفة لحساب ثبات والصدق الاختبارات ذات الاختيار من متعدد.

- **أهداف البحث:**

في ضوء العرض السابق لتحديد المشكلة يمكن صياغة الأهداف كالتالي:

7 بناء اختبار تحصيلي في مادة الرياضيات لطالبات الصف الأول الثانوي.
7 الوصول إلى أفضل طريقة لتقدير الدرجات تتبع مع طريقة معينة لحساب الثبات والصدق لاختبار ذو اختيار متعدد.
7 إمكانية زيادة ثبات وصدق الاختبار ذو الاختيار من متعدد.

- **مصطلحات البحث:**

- **الاختبار التحصيلي:**

هو أداة أو وسيلة أو طريقه تقدم للمفحوص سلسلة من المهمات عليه أن يستجيب لها بحيث أن طريقه استجابته تدل على مقدار ما عنده من تلك الصفة (عدس ، ٢٠٠٢ :٢٩).

- **طرق تقدير الدرجات:**

تحديد القاعدة التي تعطي على أساسها قيم عددية لفقرات الاختبار، أي أنها طريقه لتحديد مقياس معين للمعلومات التي نحصل عليها من إجراءات القياس، يعكس مستوى الأداء عند المفحوصين على الاختبار (سواقد، ١٩٩٢م :٣).

وستستخدم الباحثة في البحث الحالي الطرق التالية:

7 أ- الطريقة التقليدية ب- الطريقة التجريبية

7 ج- طريقة الاختيار الاستنتاجي د- طريقة المكافأة من طرق التصحيح لأثر التخمين.

- **صدق الاختبار**:

يعرف مسك Messick الصدق بأنه "تقييم شامل يوفر من خلاله الدليل المادي والمبرر النظري اللازمين لإثبات كفاية وملاءمة ومعنى أي تأويل أو فعل يبنى على درجة المقياس" (الدوسري، ٢٠٠٠: ٥٣).

- **الاختبار ذو الاختيار من متعدد**:

عرف جان (١٩٩٩م): الاختبار دو الاختيار من متعدد بأنه الاختبار الذي يتكون من اختيارات يكون فيها جواب واحد صحيح فقط يخفي بذكاء بين عدد من الموهات (المحولات أو المشتتات أو المغلطات أو المضللات). (ص١٢٦)

- **ثبات الاختبار**:

ثبات الاختبار هو ما يدل على مدى اعتماد الفروق الفردية في درجات الاختبار على أخطاء الصدفة المتضمنة في القياس من ناحية وعلى الاختلافات الحقيقية الأصيلة في الصفة أو الخاصية السيكولوجية موضع الاهتمام من ناحية أخرى (أبو حطب، عثمان، ١٩٧٣م:٧٨).

- **حدود البحث**:

يتحدد البحث وفقاً لما يلي:

7 الحدود الموضوعية: يتحدد هذا البحث بموضوعه وهو أثر بعض طرق تقدير الدرجات للمفردات على ثبات وصدق الاختبار ذو الاختيار من متعدد، كما يقتصر البحث على مادة الرياضيات للصف الأول الثانوي، الفصل الدراسي الثاني لعام ١٤٢٩هـ على باب حساب المثلثات من هذا المقرر والذي تم تدريسه في هذا الفصل الدراسي.

7 الحدود المكانية : يقتصر هذا البحث على طالبات الصف الأول بنات في مدينة مكة المكرمة في مادة الرياضيات.

7 الحدود الزمانية : تم تطبيق هذه الدراسة خلال الفصل الدراسي الثاني من العام الدراسي ١٤٢٨هـ/ ١٤٢٩هـ.

- **الإطار النظري والدراسات السابقة** :

- **مفهوم القياس** Measurement :

القيام عملية مقارنة شئ بوحدات معينة أو بكمية قياسية أو بمقدار مقنن من نفس الشئ أو الخاصية بهدف معرفة كم من الوحدات التي يتضمنها هذا الشئ. (فرج، ١٩٨٠م: ٤٩). وهو تحديد أرقام حسب قواعد معينة. (ليونا أ . تايلر، ١٩٨٣م: ٢٢). ويعرف جلفورد Guilford القياس بأنه عبارة عن وصف البيانات في صورة رقمية. كما يعرف ستيفنز Stetevs بأنه عبارة عن الأداة التي تستخدم في رصد الظاهرة السلوكية بصورة إجرائية. (عبد الهادي السيد عبده وفاروق السيد عثمان، ٢٠٠٢م : ٢١)

ويعرفه أبو لبده بأنه العملية التي بواسطتها نجد كمية الخاصية أو السمة الموجودة في الشئ (أبو لبده ، ١٩٩٦م : ٤١)

- **معنى التقويم** Evaluation :

يوصف التقويم بأنه عملية تسمح لنا باتخاذ أحكام حول قيمة شئ ما، ويستطيع الفرد أن يقيم المعلومات النوعية أو الكمية على حد سواء. (ويليام أ. ميهرنز وإيرفن لبمان، ٢٠٠٣م: ١٨) .التقويم في التربية يعرف بأنه قياس مدى تحقيق الأهداف عند الفرد. ويعرفه جروتلند Grounland بأن التقويم يعني التعرف إلى مدى ما تحقق من الأهداف عند الطالب واتخاذ قرارات بشأنه. (الهويدي، ٢٠٠٤م: ٢٤). ويعرفه مجدي حبيب بأنه إصدار الأحكام مقترنة بخطط تعديل المسار وتصويب الاتجاه في ضوء ما تسفر عنه البيانات من معلومات. (حبيب، ٢٠٠٠م: ١٥). ويعرفه أحمد زكي صالح بأنه إصدار حكم على مدى وصول العملية التربوية إلى أهدافها ومدى تحقيقها لأغراضها والعمل على كشف نواحي النقص في العملية التربوية أثناء سيرها الطبيعي واقتراح الوسائل لتلافي هذا النقص في المستقبل. (صالح ، ٢٠٠١م: ٤٢٣)

- **الاختبارات** :

نظراً لتعقد وتشابك الظواهر التربوية والنفسية وتعدد التطبيقات الميدانية للقياس والاختبارات فقد تعددت وتنوعت الاختبارات التي يمكن تصنيفها وفقاً للعديد من الأبعاد وفيما يلي أحد التصنيفات الممكنة من حيث:

7 المحتوى: تختلف الاختبارات من حيث محتواها فقد يكون محتواها لغوياً مثل الجمل والمفردات أو غير لغوي مثل المسائل الحسابية أو المواد العيانية مثل المتاهات أو المكعبات وقد يكون المحتوى لقياس الجوانب الوجدانية لدى المفحوص.

7 الغرض (الأداء): وتنقسم إلى اختبارات الأداء المميز واختبارات أقصى الأداء وتقيس اختبارات الأداء الأقصى Maximum Performance درجة الصواب لدى المفحوصين حيث يطلب منهم بذل أقصى جهد للحصول على أعلى درجة ممكن أما اختبارات الأداء المميز Typical Performance فهي تعتمد على الرأي الشخصي للفرد الذي يعكس سلوكه المميز في مواقف معينة.

7 طريقة التصميم والبناء (التقنين) : فقد يكون الاختبار مقنن أو غير مقنن.

7 فالاختبار المقنن هو ذلك الاختبار الذي تكون طريقة تطبيقه ومواده وتعليماته وطريقة تصحيحه موحدة لجميع المختبرين، أما الاختبار غير المقنن مثل ا لذي يعدها المعلمون ويختلف هذان النوعان في طريقة البناء والتصميم.

7 طريقة التطبيق: حيث يمكن أن تطبق على فرد واحد وتسمى اختبار فردي ويمكن أن يطبق الاختبار على مجموعة كبيرة من الأفراد في وقت واحد ويسمى هذا النوع من الاختبارات بالاختبار الجماعي.

- نمط الأداء : يقصد بذلك نوع النشاط الذي يؤديه الفرد في الاختبار فقد يكون الاختبار يتطلب الكتابة بالورقة والقلم (كتابية). وهناك اختبارات تتطلب الإجابة الشفوية وهناك اختبارات تتطلب أداء ما مثل اختبارات التربية الفنية.
- كيفية الاستجابة: حيث تتباين أدوات القياس في كيفية الاستجابة فمنها ما يعتمد على السرعة في الاستجابة Speed Tests ومنها مايعتمد في الاستجابة على القوة Power Tests ففي النوع الأول يكون زمن الاختبار محدد أما النوع الثاني فلا يكون لها زمن محدد للاجابة. وممكن أن تسمى باختبارات موقوته أو غير موقوته عند تصنيفها من حيث الزمن.
- طريقة التصحيح: وهنا يصنف الاختبار إلى موضوعي Objectivity وذاتي Subjectivity فالنوع الأول يعني عدم تأثر تصحيح الاختبار بالأحكام الذاتية للقائمين بها مثل اختبارات الاختيار من متعدد والصواب والخطأ وهذا النوع يصحح عادة بمفتاح تصحيح أما الاختبارات الذاتية فهي تلك التي تتأثر وتختلف درجاتها باختلاف القائمين على التصحيح مثل الاختبارات المقالية.
- تفسير الدرجات : حيث يمكن تفسير درجات الاختبار بطريقتين هما:
  - اختبار مرجعي المعيار Norm-Referenced : وفيه يتم تقدير أداء الفرد بالنسبة لأداء أقرانه وذلك بهدف ترتيبهم بالنسبة لبعضهم البعض في سمة أو قدرة ما. وهذا النوع يمدنا بتفسير درجة الفرد في ضوء معيار الجماعة .
  - اختبار مرجعي المحك Criterion-Referenced : وهو اختبار يهدف إلى تقدير أداء الفرد بالنسبة لمحك أو مستوى أداء معين ، وبالتالي يمدنا هذا النوع بمعلومات تفصيلية محددة عن تحصيل الطالب.

- **الاختبارات التحصيلية** Achievement Test

لقد بحث العديد من العلماء المختصين مفهوم التحصيل الدراسي بطرق مختلفة ولعل أبرز الاتجاهات في تحديد هذا ا لمفهوم هو ربطه بمفهوم التعلم الدراسي. فقد استخدمت اختبارات التحصيل لتحديد ما تعلمه الفرد بعد أن تعرض لنوع معين من التعليم، لا يستطيع أحد أن ينكر ضرورة الاختبارات التحصيلية بالنسبة للتربية الرسمية الفعالة، حيث أن عملية التربية معقدة ومكلفة جهدا ومالا لذا لابد أن توجه الجهود نحو تحقيق أهداف معينة وحتى يتم ذلك لابد لجميع المهتمين بالتربية من معلمين وطلاب ومسئولين وأباء لابد لهم من المعرفة الدورية لمدى نجاح جهودهم ومواطن القوة والضعف لديهم ليتم تصحيحها وعلاجها.

- **ثانياً : دراسات وبحوث سابقة:**

- **أولاً : دراسات تناولت ثبات وصدق الاختبار مع متغيرات أخرى:**

دراسة القضيب (٢٠٠١م): أساليب التحقق من ثبات وصدق الاختبارات المستخدمة في رسائل الماجستير في أقسام علم النفس في بعض جامعات المملكة

العربية السعودية. وكانت نتائج الدراسة في جانب الثبات: هناك مجموعة من الاختبارات لم يتم التحقق من ثباتها على الإطلاق، أيضا هناك بعض الممارسات الخاطئة تمثلت في استخدام أساليب ثبات غير مناسبة لطبيعة الاختبار، في جانب الصدق: هناك مجموعة من الاختبارات لم يتم التحقق من صدقها على الإطلاق، كما هناك بعض الممارسات الخاطئة تمثلت في استخدام أساليب صدق غير سليمة علميا، وكان ابرز الأخطاء هي عدم الدقة في اختيار المحكات المناسبة، وعدم إتباع الشروط المنهجية في التحقق من صدق البناء.

وأضاف الزهراني (٢٠٠٠م) حيث قدم في دراسته مقارنه بين ثمان طرق لتقدير الثبات في القياس محكي المرجع وهي معامل اتفاق هاينا ومعامل اتفاق سبكوفياك ومعامل كابا لهاينا ومعامل كابا لسبكوفياك ومعامل بيتا ومعامل كودر ريتشاردسون ٢٠ KR20 ومعامل كودر ريتشاردسون٢١ KR21 ومعامل لفنجستون، حيث كشفت دراسته عن الفروق العملية بينهما من حيث اتجاهات التغير وشدته في قيم معاملات الثبات في ظل تغير قيم التباين وعدد الأفراد وعدد البنود والدرجة الفاصلة .

كما وردت دراسة عليان (٢٠٠٤) بعنوان تقدير الثبات للعلامات المدرسية في مدينة اربد الكبرى في الفصل الدراسي الأول ٢٠٠٣/٢٠٢٢م. وأشارت النتائج إلى عدم وجود فروق ذات دلالة إحصائية بين توزيعات مستويات الثبات حسب متغيرات نوع المدرسة والمبحث ومتغير دراسة المعلم لمساقات القياس والتقويم أو عدم دراستها. بينما أشارت النتائج إلى وجود فروق ذات دلالة إحصائية في التوزيع التكراري لقيم ثبات العلامات المدرسية حسب الصف.

- **ثانياً : دراسات تناولت طرق تقدير الدرجات**:-

قدم النبهان (Alnabhan،2002) تناول فيها اختبار ثلاث طرق لتصحيح أثر التخمين على الخصائص السيكومترية للاختبار حيث تمت المقارنة بين طريقة عدد الإجابات الصحيحة وطريقة استخدام معادلة التصحيح لأثر التخمين (العقاب) وطريقة الاختيار الجزئي، وأظهرت نتائج هذه الدراسة إلى أن تأثير معادلة التصحيح لأثر التخمين (العقاب) وطريقة الاختيار الجزئي على معامل ثبات الاختبار تأثير متساوي، وتعطيان معامل ثبات أعلى من الطريقة الأولى وهي طريقة عدد الإجابات الصحيحة. وفي دراسة لـ جرادات وسواقد (Jaradat & Sawaged،1986) تناولت اختبار ثلاث طرق لتصحيح اختبارات الاختيار من متعدد لأثر التخمين. وأثرها، على الخصائص السيكومترية للاختبار عند المفحوصين من مستويات تحصيل، ودرجات مخاطرة مختلفة. والطرق الثلاث لتصحيح الاختبار هي: ١ - الطريقة التقليدية. ٢ – معادلة التصحيح لأثر التخمين (العقاب). ٣ - طريقة الاختيار الجزئي من بين مجموعة البدائل الموضوعة للسؤال. وقد أظهرت النتائج فيما يتعلق بأثر طريقة تصحيح الاختبار على معامل ثباته، ان تصحيح الاختبار باستخدام معادلة التصحيح لأثر التخمين (العقاب) أو بطريقة الاختيار الجزئي يعطي معامل ثبات أعلى مقارنة بتصحيح الاختبار بطريقة التقليدية (تصحيح الاختبار على أساس عدد الإجابات الصحيحة).

دراسة روبرت م . ربي (١٩٧٠) Robert M. Rippy وتهدف إلى دراسة ومقارنة أثر خمس طرق من طرق تقدير الدرجات والطرق التي استخدمها الباحث هي: – الاحتمال المقترح للإجابة الصحيحة. Probability assigned to correct respinse . – الطريقة اللوغاريتمية Logarithmic Method – الطريقة الكروية Spherical Method – الطريقة الاقليدية Euclidean Method – طريقة الاختيار الاستنتاجي Inferred choice Method ، وكانت النتائج كما يلي: معامل الثبات للطريقة الأولى بلغ ٠,٦٩ ، وبلغ للطريقة الثانية ٠,٥٠ ، وبلغ في الطريقة الثالثة ٠,٤٩ ، وبالنسبة للطريقة الرابعة بلغ ٠,٥٨ والطريقة الخامسة ٠,٤٧. ومن هذه النتائج نلاحظ أن الطريقة الأولى تعطي أعلى ثبات يليها الطريقة الرابعة، ثم الطريقة الثانية، ثم الطريقة الثالثة، ثم تعطى الطريقة الخامسة أقل معدل للثبات ويؤخذ على هذه الدراسة عدم تساوي إعداد الطلاب في المجموعات المختلفة.

دراسة سواقد (١٩٨٣): هدفت الدراسة إلى الإجابة على سؤالها الرئيسي هل لطرق تصحيح الاختبار أثر على معامل ثباته. تم إعداد ثلاثة أشكال متكافئة لاختبار يقيس تحصيل التلاميذ في مبحث العلوم العامة للصف الثالث الإعدادي، وأظهرت النتائج إن استخدام طريقة الاختبار الجزئي لتصحيح اختبارات الاختيار من متعدد، أفضل من استخدام الطريقة التقليدية. واستخدام معادلة التصحيح لأثر التخمين. ولذلك لقدرتها على ضبط اثر التخمين، أنها تسمح للمفحوص باستغلال المعرفة الجزئية المتوفرة لديه.

دراسة سواقد (١٩٩٢) : هدفها الأول الإجابة على السؤال التالي: ما صحة الافتراضات النظرية التي تقوم عليها طرق التصحيح لأثر التخمين. وكان من نتائج الدراسة: عدم صحة الافتراض النظري الذي تقوم عليه معادلة التصحيح لأثر التخمين. صحة الافتراض النظري الذي تقوم عليه كل من طريقة: الاختيار الجزئي. واستبعاد البدائل غير الصحيحة. كما اظهر التحليل الإحصائي للنتائج المتعلقة بين أثر استخدام الطرق الأربع موضوع الدراسة على الخصائص السيكومترية التي يتألف منها الاختبار ما يلي: كان معامل صعوبة الفقرات عند استخدام الطريقة التقليدية. أعلى منه عند استخدام الطرق الثلاث الأخرى في حين أدى استخدام معادلة التصحيح لأثر التخمين، إلى جعل معامل صعوبة الفقرات، أقل منه عند استخدام الطرق الثلاث الأخرى. كان كل من: معامل التمييز، ومؤشر الثبات، ومؤشر الصدق للفقرات عند استخدام كل من طريقة؛ الاختيار الجزئي، واستبعاد البدائل مقارنة، بالطريقتين الأخريين؛ التقليدية ومعادلة التصحيح، كما أن استخدام معادلة التصحيح أدى إلى جعل كل من، معامل التمييز، ومؤشر الثبات، ومؤشر الصدق للفقرات أقل منه عند استخدام الطرق الثلاث الأخرى.

قدم عبد الوهاب عام (٢٠٠١) دراسة عنوانها أثر بعض الطرق الوزنية لتقدير الدرجات على صدق الاختبارات مرجعية المحك ذات الاختيار من متعدد . وخرجت الدراسة بأنه يوجد فروق دالة إحصائيا عند مستوى دلالة ٠,٠٥ ، ٠,٠١ بين قيم معاملات صدق الإختبار مرجعي بين مجموعتي الطريقة التقليدية وطريقة الدرجات التجريبية لصالح الثانية التي بلغ معامل صدقها ٠,٩١٧

والثانية بلغ فيها ٠,٨٣٨ كما أن معامل اصدق في طريقة الاحتمال المقترح كانت أعلى (٠,٨٩٥) من الطريقة التقليدية كذلك (٠,٨٣٨).

دارسة داريل ل. سابربس وجوردن و. وايت (١٩٦٩م) Darrell L. Sabera and Gordon W. White: وتهدف إلى دراسة أثر اعطاء أوزان مختلفة للاستجابات أو طرق تقدير الدرجات على ثبات وصدق اختبارات الاستعداد. وكانت النتائج كما يلي: المجموعة الأولى كانت الزيادة في الثبات ٠,٠٠٤ والمجموعة الثانية بلغ التحسن في معامل الثبات ٠,٠١٥ وبلغ في المجموعة الثالثة ٠,٠٠٦ ، وفي المجموعة الرابعة ٠,٠١٣، (وذلك لصالح طريقة ديفز الوزنية) وهذه الزيادة غير دالة إحصائيا، أي أن طريقة ديفز الوزنية لا تؤدي إلى زيادة في ثبات الاختبار عن الطريقة التقليدية. بالنسبة للصدق فلم يتأثر بتغير طريقة التصحيح، وبالنسبة للصدق فلم يتأثر يتغير طريقة التصحيح ويؤخذ على هذه الدراسة عدم إجراء الاختبارات على عينة استطلاعية لتحديد الزمن المناسب للاختبار.

دراسة مينشا برانبوم ، كيوم ك . تاتسوك (١٩٨٣): وتهدف إلى تأكيد قلة الفائدة من تصحيح الاختبارات بالطريقة التي تعتمد على الإجابة الصحيحة فقط في قياس التحصيل. وبحساب الثبات بمعامل ألفا وجد أن : ثبات الدرجات التقليدية ٠,٨٤ ثبات الدرجات التجريبية ٠,٩٣ أي أن هناك زيادة في ثبات الاختبار لصالح الطريقة التجريبية. لحساب الصدق تم استخدام اختبار آخر لمحك ووجد أن هناك فروقا دالة إحصائيا لصالح طريقة تقدير الدرجات التحويلية.

دراسة جرادات (١٩٨٨) Garadat هدفت إلى المقارنة بين عدد من طرق تقدير درجات الاختبارات مرجعية المحك ذات الاختيار من متعدد في تأثيرها على صدق وثبات تلك الاختبارات وكذا للتغلب على عامل التخمين عند الإجابة على مثل هذا النوع من المفردات، حيث بلغت حجم العينة ٥٤ طالبا جامعيا وبلغت مفردات الاختبار ٢٤ مفردة من نوع الاختيار من متعدد ، ويعد تطبيق الاختبار على أفراد العينة المقسمة إلى مجموعتين إحداهم ضابطة والأخرى تجريبية وباستخدام الطريقة التقليدية والطريقة الإقليدية تقدير الدرجات تم حساب معاملات الصدق والثبات لمفردات الاختيار المستخدم حيث تبين أن الفروق في معامل الثبات غير دالة بين الطريقتين بينما عند حساب معامل الصدق وجد أن هناك فروقا واضحة وارتفاع ملحوظ في معامل الصدق عند استخدام الطريقة التجريبية المذكورة.

وقدم روس ستيفن (1994) Ross-Steven إلى الكشف عن أثر إختلاف طرق تقدير الدرجات طرق تقدير الدرجات على كل من الصدق والثبات قبل وبعد تقديم المحتوى للاختبارات التحصيلية مرجعية المحك ذات الإختيار من متعدد في اللغة، وإستخدام تلك النتائج في تطوير دراسة اللغة لدى المتعلمين، حيث بلغت العينة ١٢٧ طالبا بالمدارس الثانوية مما يدرسون اللغة وتم تطبيق الإختبار قبل تقديم المحتوى التعليمي وأيضا بعد تقديم المحتوى وباستخدام طريقة تقليدية وأخرى تجريبية لتقدير الدرجات حيث أتضح من النتائج تباين في قيم الثبات والصدق بالنسبة لدرجات الإختبارات التي تم تطبيقها قبل وبعد تقديم المحتوى وذلك لصالح الإختبارات التي طبقت بعد تقديم المحتوى أما

بالنسبة لقيم الثبات والصدق مع اختلاف طرق تقدير الدرجات أتضح أن هناك فروق دالة في قيم الثبات لصالح الطريقة التجريبية مقارنة بالطريقة التقليدية.

- **ثالثاً : دراسات تناولت الاختبار ذو الاختيار من متعدد :**

دراسة الدويري (١٩٨٤) بعنوان مقارنة أثر اختبارات الاختيار من متعدد واختبارات الإجابة القصيرة في احتفاظ طلاب الصف الأول الثانوي بمادة قواعد اللغة العربية. وأظهرت نتائج هذه الدراسة تفوق مجموعة الإجابة القصيرة على مجموعة الاختيار من متعدد، عندما قيس الاحتفاظ باختبار مكون من مزيج من مزيج من النوعين معاً من الاختبارات، وعلى المستويات الثلاثة، كما تفوقت عندما قيس الاحتفاظ باختبار الإجابة القصيرة على مستوى المعرفة، وعلى مستوى الأداء الكلي، وتعادل أداؤهما على اختبار نوع الاختيار من متعدد، وعلى المستويات الثلاثة. كما تساوي أداؤهما على مستوى الفهم. من اختبار الإجابة القصيرة وهذا يدل على قوة الأثر التعليمي لاختبارات الإجابة القصيرة. كما تفوق مجموعتي الإجابة القصيرة والاختيار من متعدد على المجموعة الضابطة على الاختبارات الثلاثة بشكل عام.

دراسة العلي (١٩٨٩): المقارنة بين أثر اختبارات الاختيار من متعدد واختبارات الإجابة القصيرة واختبارات المزيج من النوعين معا على بعض الخصائص السيكومترية للاختبار.. وأظهرت نتائج الدراسة، وجود فروق بين معاملات ثبات الاختبار ذات دلالة إحصائية في مستوى الأداء الكلي، ومستوى التذكر وكانت هذه الفروق لصالح الإجابة القصيرة في مستوى الأداء الكلي، ولصالح الاختبار من نوع المزيج بين النوعين معا، في مستوى التذكر، كما أظهرت النتائج وجود فروق ذات دلالة إحصائية عن بين مستوى الأداء الكلي، ومستوى الاستيعاب، ولصالح الإجابة القصيرة في مستوى التذكر، في حين لم تكن هنالك فروق ذات دلالة إحصائية لباقي المعاملات.

- **تعليق عام على الدراسات والبحوث السابقة:**

بالنظر إلى الدراسات والبحوث السابقة نجد ما يلي:

7 بالنسبة للدراسات التي تناولت ثبات وصدق الاختبار نجد في بعضها عدم تنوع المعادلات التي تستخدم في حساب الثبات والصدق.

7 الدراسات التي تناولت طرق تقدير الدرجات وأثرها على الثبات والصدق يلاحظ عليها ما يلي:-

- لم تراعي دراسة روبرت م. رابي Robert M. Rippy ودراسة رونالد ك. هامبلتون Ronald K. Hambleton تساوى المجموعات في العدد مما يؤثر في النتائج. بينما راعت ذلك كل من دراسة سليمان محمد سليمان محمود. ودراسة عادل محمد محمود العدل، ودراسة الخرشة.
- بعض الدراسات السابقة لم تراعي صغر عدد أفراد المجموعة مما كان له أثر على النتائج.

- معظم الدراسات السابقة لم يتم إجراء الاختبار على عينة استطلاعية لتحديد الزمن المناسب للاختبار، وكما نعلم أن للزمن أثر كبير في حساب الثبات.
- تضارب النتائج الخاصة بالثبات من دراسة لأخرى فدراسة سواقد ١٩٨٣ ودراسة محمود ودراسة العدل ١٩٨٦ ودراسة مينشيا برانبوم، تاتسوك Brinbaum & Tatsuoka ١٩٨٣ ودراسة ربي Rippy 1970 ودراسة الخرشة (٢٠٠٤م) ودراستي سواقد والنبهان (٢٠٠٢م) وجدت أثراً لطرق تقدير الدرجات على ثبات الاختبار أما دراسات كل من سابريس وايتٍ Sabers&White 1970 ودراسة ثيسين Thissen 1976 فلم تجد أثراً للطرق المختلفة لتقدير الدرجات على ثبات الاختبار.

7 الدراسات التي تناولت الاختبار ذو الاختيار من متعدد يلاحظ عليها ما يلي:

7 استخدمت الدراسات إختبارات تحصيليه.

7 أكدت الدراسات على أهمية أسئلة الاختبار ذو الاختيار من متعدد في قياس مستوى الفهم كما في دراسة الدويري ١٩٨٤م.

7 شجعت الدراسات على اتخاذ عدد أربع بدائل للاختبار عدداً مناسباً.

- **فروض البحث**:

7 يختلف ثبات الاختبار ذو الاختيار من متعدد باختلاف طرق تقدير الدرجات للمفردات.

7 يختلف ثبات الاختبار ذو الاختيار من متعدد باختلاف طرق حساب هذا الثبات.

7 لا يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق حساب الثبات على ثبات الاختبار ذو الاختيار من متعدد.

7 يختلف صدق الاختبار ذو الاختيار من متعدد بإختلاف طرق تقدير الدرجات للمفردات.

7 يختلف صدق الاختيار من متعدد بإختلاف طرق حساب هذا الصدق.

7 لا يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق حساب الصدق على صدق الاختبار ذو الاختيار من متعدد.

- **إجراءات البحث**:

- **منهج البحث**:

استخدمت الباحثة المنهج التجريبي طبقاً للمتغيرات التالية:

7 المتغير المستقل: هو طرق تقدير الدرجات متمثلة في : ١- الطريقة التقليدية. ٢- الطريقة التجريبية.٣ - طريقة الاختيار الاستنتاجي. ٤ – طريقة المكافأة من طرق التصحيح لأثر التخمين.

المتغيرات التابعة هو كل من صدق الاختبار وثبات الاختبار.

- **ثانياً : مجتمع البحث** :

يتكون مجتمع البحث من طالبات الصف الأول ثانوي بمدارس مدينة مكة المكرمة للعام الدراسي ١٤٢٨/١٤٢٩هـ.

- **ثالثاً : عينة البحث :**
- **أ - اختيار العينة :**

يتم سحب العينة العشوائية من تلميذات الصف الأول ثانوي في كل من إدارة تعليم شمال مكة المكرمة وإدارة تعليم جنوب مكة المكرمة في الفصل الدراسي الثاني لعام ١٤٢٨/١٤٢٩هـ.

- **ب- حجم العينة :**

تكونت عينة البحث الاستطلاعية من ٧٥ طالبة من طالبات الصف الأول ثانوي بمدينة مكة المكرمة. والجدول التالي يوضح التوصيف العددي العينة الاستطلاعية

جدول رقم (٢) التوصيف العددي للعينة الاستطلاعية

البيان		عدد الطالبات
الإدارة	المدرسة	
إدارة تعليم شمال مكة	المدرسة (١٧) الثانوية	٤٥
إدارة تعليم جنوب مكة	المدرسة (٢٣) الثانوية	٣٠
المجموع الكلي	-	٧٥

- **٢ - العينة النهائية :**

تكونت عينة البحث النهائية من (٤٠٠) طالبه من طالبات الصف الأول ثاوي وقد استخدمت بيانات هذه العينة في اختبار فروض البحث الحالي. والجدول (٣) يوضح بيانات العينة النهائية.

جدول رقم (٣) وصف العينة النهائية

البيان		الطريقة	عدد الطالبات
الإدارة	المدرسة		
إدارة تعليم شمال مكة	المدرسة (١٨) الثانوية	التقليدية	٥٠
		المكافأة	٥٠
	المدرسة (١٧) الثانوية	التجريبية	٥٠
		الاختيار الاستنتاجي	٥٠
إدارة تعليم جنوب مكة	المدرسة (٢٣) الثانوية	التقليدية	٥٠
		الاختيار الاستنتاجي	٥٠
	المدرسة (٢٩) الثانوية	التجريبية	٥٠
		المكافأة	٥٠

- **رابعاً : أدوات البحث :**
- **بناء اختبار تحصيلي معياري المرجع:**

يتطلب البحث الحالي بناء اختبار تحصيلي معياري المرجع في باب حساب المثلثات للصف الأول ثانوي الفصل الدراسي الثاني للعام الدراسي ١٤٢٨/١٤٢٩هـ، ولقد تم بناء هذا الاختبار بإتباع الخطوات التالية:

7 إعداد خطة الاختبار بفحص المادة الدراسة جيداً واستشارة المعلمات ذوات الخبرة لاختيار باب من المقرر الدراسي نظرا لأهميته ووقع الاختيار على باب حساب المثلثات.

- الاستعانة بكتاب المعلمة لمادة الرياضيات للفصل الدراسي الثاني طبعة عام ١٤٢٨هـ- ١٤٢٩هـ. للحصول على الأهداف الخاصة بباب حساب المثلثات.
- أخذ آراء معلمات ذوات خبره لا تقل عن ٥ سنوات حول عدد الحصص الفعلية التي يتم تدريسها لكل موضوع والتعبير عن ذلك بنسب مئوية.
- تعيين عدد الحصص المقررة من وزارة التربية والتعليم لتدريس كل موضوع من موضوعات باب حساب المثلثات والتعبير عن ذلك كنسب مئوية.
- وذلك بالقانون :

$$\text{نسبة التركيز} = \frac{\text{عدد الحصص للوحدة الدراسية}}{\text{عدد الحصص}} \times ١٠٠$$

- حساب متوسط النسب المئوية لأهمية موضوعات باب حساب المثلثات من خلال تعيين المتوسط الحسابي لأهمية موضوعات الباب المستخرجة من عدد الصفحات التي يشغلها كل موضوع –عدد الحصص الفعلية اللازمة لتدريس كل موضوع حسب رأي معلمات المادة- عدد الحصص المقررة من وزارة التربية والتعليم لتدريس كل موضوع من موضوعات الباب.
- بناء جدول المواصفات: قامت الباحثة بتصنيف الأهداف المعرفية الخاصة بالباب حسب تصنيف بلوم إلى التذكر والفهم والتطبيق والتحليل والتركيب والتقويم ولم تجد الباحثة لمستوى التقويم أثراً في هذا الباب وبعد ذلك قامت الباحثة بتحديد الأهداف المعرفية في كل موضوع من موضوعات باب حساب المثلثات ونسبة تواجدها في هذا الموضوع وذلك بالاستعانة بالنسب المئوية والتي تبين النسب المئوية لأهمية الموضوعات في باب حساب المثلثات ثم تحديد عدد الأسئلة للموضوع بالقانون التالي: عدد الأسئلة لكل موضوع = نسبة مستوى الهدف × نسبة التركيز × عدد الأسئلة الكلي وبإتباع الطريقة السابقة تم بناء جدول المواصفات ،ولما كان من الصعب وضع مائةٍ سؤال ليتكون منه الاختبار حيث أنهٍ يجب أن يكون طول الاختبار مناسباً ، وجعل الاختبار يتكون من ٢٥ سؤالاً قامت الباحثة بقسمة كل النسب المئوية الموجودة في جدول المواصفات على رقم (٤) وتم الحصول على الجدول رقم (٤) الذي يبين عدد الأسئلة في كل موضوع.

وبتطبيق الاختبار على العينة الاستطلاعية الذي سبق وصفها تم الحصول على درجات الطالبات المبينة في الجدول (٤):

حيث م٢ هو المتوسط المرتقب وهو يساوي النهاية العظمى للاختبار مقسومة على م١ هو المتوسط التجريبي وهو عبارة عن متوسط درجات التلاميذ في الاختبار ن١ هو متوسط الزمن الذي استغرقه مجموعة التلاميذ في الاختبار (البهي، ٢٠٠٦م:٤٦٧) ولتحديد الزمن التجريبي للاختبار قامت الباحثة بتسجيل الزمن الذي استغرقته كل طالبه للإجابة على الاختبار ووجدت الباحثة أن : م٢ = ١٢,٥ ، م١= ١٥,١٩، ن١ = ٢٠ دقيقه. وبتطبيق المعادلة السابقة نجد أن : الزمن المناسب لإجراء الاختبار هو ١٦,٥دقيقه.

جدول رقم (٤) التكرارات لعينة الدراسة

الدرجة	التكرار	الدرجة × التكرار	الدرجة	التكرار	الدرجة × التكرار
٨	٢	١٦	١٧	١٠	١٧٠
٩	٢	١٨	١٨	٧	١٢٦
١٠	٨	٨٠	١٩	٨	١٥٢
١١	٦	٦٦	٢٠	٣	٦٠
١٢	٢	٢٤	٢١	٢	٤٢
١٣	١٠	١٣٠	٢٢	١	٢٢
١٤	٥	٧٠	٢٣	١	٢٣
١٥	٤	٦٠	٢٤	٢	٤٨
١٦	٢	٣٢	المجموع	٧٥	١١٣٩

المتوسط الحسابي = ١٥,١٩

حساب معاملات السهولة والصعوبة لمفردات الاختبار :

تقاس سهولة مفردات الاختبار بحساب المتوسط الحسابي للإجابات الصحيحة وذلك بالنسبة للذين أجابوا فعلاً على السؤال إجابة صحيحة مع استبعاد المفردات المحذوفة والمتروكة، كما أن العلاقة بين السهولة والصعوبة علاقة عكسية مباشرة أي : معامل الصعوبة = ١ – معاملة السهولة

وقد حسبت الباحثة معاملات السهولة والصعوبة لكل سؤال من أسئلة الاختبار وذلك بهدف إعادة صياغة الأسئلة شديدة السهولة والأسئلة شديدة الصعوبة والجدول (٥) يوضح معاملات سهولة وصعوبة كل مفرده.

ووجدت الباحثة أن المفردات ذات الأرقام (١٢- ١٥- ١٧- ٢١- ٢٤) تحذف والمفردات (٥- ١٤- ٢٣- ٢٥) تعدل.

- **ثبات المفردات** :

يعتمد ثبات الاختبار اعتماداً مباشراً على ثبات مفرداته، ولقد استخدمت الباحثة لحساب ثبات مفردات الاختبار طريقه الاحتمال المنوالي وذلك لأنها الطريقة التي تصلح لحساب المفردات التي تعتمد إجابتها على اختيار اجابه واحدة من إجابتين أو إجابة واحدة من عدة إجابات محتمله ويبن الجدول (٦) ثبات مفردات الاختبار التحصيلي.

كما قامت الباحثة بحساب الثبات الكلي للاختبار من خلال حساب قيمة معامل الثبات الفا كرونباخ Cronbach's التي بلغت (٠,٨٥٥) وهذا يدل على ثبات عالي.

- **صدق المفردات** :

استخدمت الباحثة طريقة الفروق الطرفية لأنها لا تستغرق وقتاً وجهداً كبيراً في طريقة حسابها وكانت النتائج كما بالجدول (٧) :

يلاحظ من الجدول (٧) أن جميع المفردات تحتوي على ثبات وصدق دال إحصائياً أي أن المقياس صالح للاستخدام.

جدول رقم (٥): معاملات سهولة وصعوبة الاختبار التحصيلي

رقم السؤال	معامل السهولة	معامل الصعوبة	معامل التمييز	التباين
١	٠,٧٤	٠,٢٦	٠,٣	٠,١٩
٢	٠,٧١	٠,٢٩	٠,٥٥	٠,٢١
٣	٠,٨٣	٠,١٧	٠,٤	٠,١٤
٤	٠,٧٧	٠,٢٣	٠,٤	٠,١٨
٥	٠,٩٥	٠,٠٥	٠,١٥	٠,٠٤٧
٦	٠,٥٣	٠,٤٧	٠,٥	٠,٢٥
٧	٠,٦	٠,٤	٠,٦٥	٠,٢٤
٨	٠,٦٩	٠,٣	٠,٧	٠,٢١
٩	٠,٥٩	٠,٤١	٠,٨٥	٠,٢٤
١٠	٠,٨	٠,٢	٠,٤٥	٠,١٦
١١	٠,٧٩	٠,٢١	٠,٥٥	٠,١٧
١٢	٠,٢٩	٠,٧١	٠,٠٥	٠,٢١
١٣	٠,٦	٠,٤	٠,٧	٠,٢٤
١٤	٠,٨٤	٠,١٦	٠,١٥	٠,١٣
١٥	٠,٢٣	٠,٧٧	٠,٠٥	٠,١٨
١٦	٠,٦٤	٠,٣٦	٠,٦٥	٠,٢٣
١٧	٠,٢٤	٠,٧٦	٠	٠,١٨
١٨	٠,٧٤	٠,٢٦	٠,٣٥	٠,١٩
١٩	٠,٦	٠,٤	٠,٦	٠,٢٤
٢٠	٠,٨١	٠,١٩	٠,٤	٠,١٥
٢١	٠,٢١	٠,٧٩	٠,٠٥	٠,١٧
٢٢	٠,٤٥	٠,٥٥	٠,٧	٠,٢٥
٢٣	٠,٢٤	٠,٧٦	٠,٢	٠,١٨
٢٤	٠,٢٨	٠,٧٢	٠,٠٥	٠,٢٠
٢٥	٠,٩١	٠,٠٩	٠,١	٠,٠٨

جدول رقم (٦) يبين ثبات مفردات الاختبار التحصيلي

رقم السؤال	التكرار الأكثر للاستجابة	التكرار النسبي	معامل الثبات
١	٥٦	٠,٧٤	٠,٦٤
٢	٥٣	٠,٧١	٠,٦
٣	٦٢	٠,٨٣	٠,٧٤
٤	٥٨	٠,٧٧	٠,٥٢
٥	٧١	٠,٩٥	٠,٩١
٦	٤٠	٠,٥٣	٠,٣٦
٧	٤٥	٠,٦	٠,٤٦
٨	٥٢	٠,٦٩	٠,٤٤
٩	٤٤	٠,٥٩	٠,٤٤
١٠	٦٠	٠,٨	٠,٧٢
١١	٥٩	٠,٧٩	٠,٧٠
١٢	٣٩	٠,٥٢	٠,٣٥
١٣	٤٥	٠,٦	٠,٤٦
١٤	٦٣	٠,٨٤	٠,٧٧
١٥	٢٩	٠,٣٩	٠,١٨
١٦	٤٨	٠,٦٤	٠,٥١
١٧	٢٨	٠,٣٧	٠,١٦
١٨	٥٦	٠,٧٥	٠,٧
١٩	٤٥	٠,٦	٠,٤٦
٢٠	٦١	٠,٨١	٠,٧٣
٢١	٢٦	٠,٣٥	٠,١٣
٢٢	٣٤	٠,٤٥	٠,٢٦
٢٣	٢٧	٠,٣٦	٠,١٤
٢٤	٣٠	٠,٤	٠,١٩
٢٥	٦٨	٠,٩٠	٠,٨٥

جدول رقم (٧) معاملات صدق المفردات للاختبار التحصيلي

رقم السؤال	ص ع	ص س	معامل صدق المفردة
١	١٨	١٢	٠,٣
٢	٢٠	٩	٠,٥٥
٣	٢٠	١٢	٠,٤
٤	٢٠	١٢	٠,٤
٥	٢٠	١٧	٠,١٥
٦	١٧	٧	٠,٥
٧	٢٠	٧	٠,٦٥
٨	١٩	٥	٠,٧
٩	٢٠	٣	٠,٨٥
١٠	١٩	١٠	٠,٤٥
١١	٢٠	٩	٠,٥٥
١٢	٦	٥	٠,٠٥
١٣	١٩	٥	٠,٧
١٤	١٩	١٦	٠,١٥
١٥	٥	٤	٠,٠٥
١٦	٢٠	٧	٠,٦٥
١٧	٥	٥	٠
١٨	١٩	١٢	٠,٣٥
١٩	١٧	٥	٠,٦
٢٠	٢٠	١٢	٠,٤
٢١	٤	٣	٠,٠٥
٢٢	١٨	٤	٠,٧
٢٣	٧	٣	٠,٢
٢٤	٦	٥	٠,٠٥
٢٥	١٩	١٧	٠,١

- **تحليل نتائج الدراسة ومناقشتها** :

تناول هذا الفصل عرض النتائج التي توصلت إليها هذه الدراسة من خلال الإجابة على أسئلة الدراسة التالية :-

- **للإجابة على السؤال الأول : والذي ينص على** :-

هل يختلف ثبات الاختبار ذي الاختيار من متعدد باختلاف طرق تقدير الدرجات للمفردات؟ ولحساب معاملات ثبات الاختبارات بطرق التصحيح المختلفة يمكن استخدام برنامج الحزم الإحصائية المعروف ب SPSS ، وكانت نتائج تطبيق معاملات الثبات (معادلة ألفا كرونباخ ، والتجزئة النصفية ومعادلة جتمان) كما في الجدول (٨) :-

جدول رقم (٨) يوضح معامل ثبات الاختبار بطرق التصحيح المختلفة باستخدام معادلة ألفا كرونباخ

طريقة تقدير الدرجات	الطريقة التقليدية	الطريقة الاحتمالية	الطريقة المكافأة	الطريقة التجريبية
معامل الثبات	0.8158	٠,٧٨٧٧	٠,٧٦٧٩	٠,٩٣٣٩

يتضح من جدول (٨) وجود اختلاف في معاملات ثبات الاختبار التحصيلي ذي الاختيار من متعدد باختلاف طريقة تقدير الدرجات حيث بلغ أعلى معامل ثبات للطريقة التجريبية بقيمة (٠,٩٣٣٩) وبلغت أقل قيمة لمعامل ثبات الاختبار بطريقة المكافأة وبلغت قيمته (٠,٧٦٧٩) ، في حين احتلت الطريقة التقليدية المرتبة الثانية بعد الطريقة التجريبية وبلغ معامل ثبات الاختبار بطريقة تقدير الدرجات بالطريقة التقليدية (٠,٨١٥٨) ، واحتلت الطريقة الاحتمالية المرتبة الثالثة وقبل الأخيرة بمعامل ثبات بلغ (٠,٧٨٧٧) .

يمكن إرجاع هذه الفروق بين معاملات ثبات الاختبار إلى الاختلاف بين درجات الطالبات على الاختبار كما يتضح من الجداول (٩) ، (١٠) ، (١١) ، (١٢) :

جدول (٩) يبين معالم درجات طالبات المجموعة التجريبية (قبلياً و بعدياً)

المجموعة التجريبية	العدد	القيمة الدنيا	القيمة القصوى	الوسط	الانحراف المعياري	التباين	الالتواء
قبليا	١٠٠	١٠	٣٠	٢٤,٦	٤,٥٧	٢٠,٨٤	- ١,٠٦
بعديا	١٠٠	٤١	٦٠	٥٣,٢٧	٤,٩٨	٢٤,٨٠	- ٠,٢٠

جدول (١٠) يبين معالم درجات طالبات المجموعة الاحتمالية (قبلياً و بعدياً)

المجموعة الاحتمالية	العدد	القيمة الدنيا	القيمة القصوى	الوسط	الانحراف المعياري	التباين	الالتواء
قبليا	١٠٠	١٠	٣٠	٢٢,٠٧	٦,٧٩	٤٦,١٥	- ٠,٤٦
بعديا	١٠٠	٤١	٢٠	١٦,٧١	٢,٦٢	٦,٨٧	- ٠,٤٤

جدول (١١) يبين معالم درجات طالبات المجموعة المكافأة (قبلياً و بعدياً)

المجموعة الاحتمالية	العدد	القيمة الدنيا	القيمة القصوى	الوسط	الانحراف المعياري	التباين	الالتواء
قبليا	١٠٠	١٠	٣٠	٢٢,٠٧	٦,٧٩	٤٦,١٥	- ٠,٤٦
بعديا	١٠٠	٤١	٢٠	١٦,٧١	٢,٦٢	٦,٨٧	- ٠,٤٤

جدول (١٢) يبين معالم درجات طالبات المجموعة الضابطة (قبلياً و بعدياً)

المجموعة الاحتمالية	العدد	القيمة الدنيا	القيمة القصوى	الوسط	الانحراف المعياري	التباين	الالتواء
قبليا	١٠٠	٨	٣٠	١٩,٤٤	٧,٦٣	٥٨,٢١	٠,٦٤
بعدياً	١٠٠	٨	٢٠	١٥,٥٦	٣,٢٣	١٠,٤٥	- ٠,٣١

وبتحليل التباين أحدي الاتجاه (ANOVA) بين درجات الطالبات على الاختبار التحصيلي ذي الاختيار من متعدد يمكن توضيح الفروق بين الطرق المختلفة لتقدير الدرجات أثرها على معامل ثبات الاختبار كما يتضح في الجدول (١٣) : ومن الجدول يتضح وجود فروق ذات دلالة بين درجات الطالبات على الاختبار ذي الاختيار من متعدد حيث بلغت قيم معامل F (٦,٧٣) لطريقة المكافأة بالمقارنة بالطريقة التقليدية ، و(1.475) لطريقة الاحتمالية و(1.287) لطريقة التجريب وجميعها ذات دلالة إحصائية ، مما يؤكد وجود فروق بين معاملات الثبات للاختبار يرجع لطريقة لتقدير الدرجات .

جدول رقم (١٣) : يبين معاملات تحليل التباين أحادي الاتجاه (ANOVA) بين الطرق الثلاثة لتقدير الدرجات

مصادر التباين		مجموع المربعات	متوسط المربعات	درجات الحرية	قيمة F	القيمة الإحصائية	الدلالة
طريقة المكافأة	بين المجموعات	٣١٨.١٦٤	١١.٣٦٣	٢٨	٦.٧٣	٠.٠٠	دالة
	داخل المجموعات	١١٩.٠٤	١.٦٨٨	٧١			
طريقة الاحتمالية	بين المجموعات	٢٥٠.٢٩٨	٨.٩٣٩	٢٨	١.٤٧٥	٠.٩٦	دالة
	داخل المجموعات	٤٣٠.٢٩٢	٦.٠٦	٧١			
الطريقة التجريبية	بين المجموعات	٨٢٦.٩٥٢	٢٩.٥٣٤	٢٨	١.٢٨٧	٠.١٩٦	دالة
	داخل المجموعات	١٦٢٨.٧٥٨	٢٢.٩٤	٩١			

وتتفق هذه النتائج مع دراسة كل من (الزهراني، ٢٠٠٠م) والتي كشفت عن الفروق العملية بين ثمان طرق لتقدير ثبات الاختبار من حيث اتجاهات التغير وشدته رغم تشابه بعض مسلكيات بعض المعاملات ،ودراسة (النبهان Alnabhan, 2002) والتي تتفق هي أيضا مع دراسة (جرادات وسواقد Jaradat & Sawaged,1986) ، ودراسة (روبرت م . ربي (١٩٧٠) Robert M. Rippy) ودراسة (مينشا برانبوم ، كيوم ك . تاتسوك ،١٩٨٣)،(دراسة سواقد ،١٩٨٣)، ودراسة محمود ،١٩٨٦م) والتي أظهرت نتائجها أن معامل الثبات المحسوب يختلف باختلاف طرق تقدير الدرجات للاختبار مما يؤكد صحة الفرض الأول للدراسة والقائل بأن معامل ثبات الاختبار يختلف باختلاف طرق تقدير درجات الاختبار.

وتختلف مع دراسة كل من (داريل ل. سابربس وجوردن و. وايت (١٩٦٩م) Darrell L. Sabera and Gordon W. White) وكل من دراسة (سابريس وايت 1970 Sabers&White) ودراسة ثيسين Thissen 1976التي لم تجد أثرا للطرق المختلفة لتقدير الدرجات على ثبات الاختبار وقد ترجع هذه النتائج لاختلاف المرحلة العمرية لعينة الدراسة وحجم هذه العينات وكذلك عدد فقرات الاختبار ونوع الاختبار ذاته.

ويمكن تفسير هذه النتائج بأن الكثير من الدراسات لم يتم التحقق من ثبات وصدق الاختبارات الموجودة بها كما أوضحت دراسة القضيب (٢٠٠١م): حيث توصلت إلى أنٍ هناك مجموعة من الاختبارات لم يتم التحقق من ثباتها على الإطلاق، أيضا هناك بعض الممارسات الخاطئة تمثلت في استخدام أساليب ثبات غير مناسبة لطبيعة الاختبار، في جانب الصدق: هناك مجموعة من الاختبارات لم يتم التحقق من صدقها على الإطلاق، كما هناكٍ بعض الممارسات الخاطئة تمثلت في استخدام أساليب صدق غير سليمة علميا، وكان ابرز الأخطاء هي عدم الدقة في اختيار المحكات المناسبة، وعدم إتباع الشروط المنهجية في التحقق من صدق البناء.

- **للإجابة على السؤال الثاني والذي ينص على: هل يختلف ثبات الاختبار ذي الاختيار من متعدد باختلاف طرق حساب هذا الثبات؟**

تم حساب بمعاملات الثبات بطرق (معادلة ألفا كرونباخ ،والتجزئة النصفية ومعادلة جتمان) وكانت النتائج كما في الجدول (١٤) :

جدول رقم (١٤) يبين معاملات الثبات للاختبار ذي الاختيار من متعدد لطريقة تقدير الدرجات بطريقة المكافأة

معامل الثبات	ألفا كرونباخ	التجزئة النصفية	معادلة جتمان
قيمة المعامل	٠,٧٦٧٩	٠,٧٦٧٩	٠,٧٦٧٩

من الجدول (١٤) يتضح وجود اختلافات بين معامل ثبات الاختبار ذي الاختيار من متعدد باختلاف طريقة حساب معاملات الثبات حيث بلغت قيمة معامل الثبات بمعادلة ألفا كرونباخ (٠,٧٦٧٩) وتتساوى مع قيمة معامل ثبات الاختبار بطريقة التجزئة النصفية و معامل جتمان أيضا، وتم حساب معاملات الثبات للاختبار التحصيلي والذي تم تقدير الدرجات بطريقة الاحتمالية

جدول رقم (١٥) يبين معاملات الثبات للاختبار ذي الاختيار من متعدد لطريقة تقدير الدرجات بطريقة الاحتمالية

معامل الثبات	ألفا كرونباخ	التجزئة النصفية	معادلة جتمان
قيمة المعامل	٠,٧٨٧٧	٠,٧٨٧٧	٠,٧٨٧٧

من الجدول (١٥) يتضح وجود اختلافات بين معامل ثبات الاختبار ذي الاختيار من متعدد باختلاف طريقة حساب معاملات الثبات حيث بلغت قيمة معامل الثبات بمعادلة ألفا كرونباخ (٠,٧٨٧٧) وتتساوى بقيمة معامل الثبات للاختبار بواسطة التجزئة النصفية وقيمة معامل الثبات للاختبار .

وتم حساب معاملات الثبات للاختبار التحصيلي والذي تم تقدير الدرجات بطريقة التجريبية وكانت هذه المعاملات كما في الجدول التالي :-

جدول رقم (١٦) يبين معاملات الثبات للاختبار ذي الاختيار من متعدد لطريقة تقدير الدرجات بطريقة التجريبية

معامل الثبات	ألفا كرونباخ	التجزئة النصفية	معادلة جتمان
قيمة المعامل	٠,٩٣٣٩	٠,٩٣٣٩	٠,٩٣٣٩

من الجدول (١٦) يتضح وجود اختلافات بين معامل ثبات الاختبار ذي الاختيار من متعدد باختلاف طريقة حساب معاملات الثبات حيث بلغت قيمة معامل الثبات بمعادلة ألفا كرونباخ (٠,٩٣٣٩) وتتساوى بقيمة معامل الثبات للاختبار بواسطة التجزئة النصفية وقيمة معامل الثبات للاختبار .

وتم حساب معاملات الثبات للاختبار التحصيلي والذي تم تقدير الدرجات بطريقة التقليدية وكانت هذه المعاملات كما في الجدول التالي :-

جدول رقم (١٧) يبين معاملات الثبات للاختبار ذي الاختيار من متعدد لطريقة تقدير الدرجات بطريقة التقليدية

معامل الثبات	ألفا كرونباخ	التجزئة النصفية	معادلة جتمان
قيمة المعامل	٠,٨١٥٨	٠,٨١٥٨	٠,٨١٥٨

من الجدول (١٧) يتضح عدم وجود اختلافات بين معامل ثبات الاختبار ذي الاختيار من متعدد باختلاف طريقة حساب معاملات الثبات حيث بلغت قيمة

معامل الثبات بمعادلة ألفا كرونباخ (٠,٨١٥٨) وتتساوى بقيمة معامل الثبات للاختبار بواسطة التجزئة النصفية وقيمة معامل الثبات للاختبار .

من خلال عرض معاملات ثبات الاختبار بمعادلات مختلفة؛ يتضح أن معاملات الثبات متفقة تماما رغم اختلاف طرق حساب معاملات الثبات للطريقة الواحدة في تقدير الدرجات.

وتتفق هذه النتيجة مع نتائج دراسة كل من دراسة الزهراني (٢٠٠٠م) ودراسة (العدل :١٩٨٦) حيث توصلت إلى أنه لا توجد فروق ذات دلالة إحصائية بين معاملات الثبات المحسوبة بمعادلة سبيرمان- براون ومعادلة كرونباخ وكذلك لا توجد فروق ذات دلالة إحصائية بين معاملات الثبات المحسوبة بمعادلة كرونباخ وطريقة إعادة الاختبار وذلك لطرق تقدير الدرجات الأربع.

وتختلف نتائج هذه مع نتائج بعض الدراسات ؛ومنها دراسة (روبرت م . ربي Robert M. Rippy (١٩٧٠) ، ودراسة (دراسة مينشا برانبوم ، كيوم ك. تاتسوك (١٩٨٣) ، ودراسة (محمود، ١٩٨٦) والتي تُظهر نتائجها المجملة أن معاملات ثبات الاختبار الواحد تختلف باختلاف طرق حساب معاملات ثبات الاختبار مما يؤدي لرفض الفرض الثاني والذي ينص على (يختلف ثبات الاختبار ذو الاختيار من متعدد باختلاف طرق حساب هذا الثبات)

- **للإجابة على السؤال الثالث للدراسة والذي ينص على : هل يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق حساب الثبات على ثبات الاختبار ذي الاختيار من متعدد؟**

يمكن مقارنة معاملات ثبات الاختبار بطرق تقدير درجات الاختبار و طرق حساب معاملات ثبات الاختبار كما في الجدول التالي :-

جدول (١٨) يبين العلاقة بين معاملات ثبات الاختبار وطرق تقدير درجات الاختبار و طرق حساب معاملات ثبات الاختبار

طريقة تقدير درجات الاختبار	طرق حساب معاملات الثبات		
	معامل ألفا كرونباخ	التجزئة النصفية	معامل جتمان
طريقة المكافأة	٠,٧٦٧٩	٠,٧٦٧٩	٠,٧٦٧٩
طريقة الاحتمالية	٠,٧٨٧٧	٠,٧٨٧٧	٠,٧٨٧٧
الطريقة التجريبية	٠,٩٣٣٩	٠,٩٣٣٩	٠,٩٣٣٩
الطريقة التقليدية	٠,٨١٥٨	٠,٨١٥٨	٠,٨١٥٨

من جدول (١٨) يتضح عدم وجود اختلاف بين معاملات ثبات الاختبار بالطريقة الواحدة لتقدير الدرجات رغم اختلاف طريقة حساب هذه المعاملات؛ فمعاملات ثبات الاختبار بطريقة المكافأة هي قيمة واحدة بلغت (٠,٧٦٧٩) وتم حسابها بمعادلة ألفا كرونباخ والتجزئة النصفية ومعامل جتمان ، وكذلك معاملات ثبات الاختبار بطريقة الاحتمالية بلغت (٠,٧٨٧٧) بحسابها بمعادلة ألفا كرونباخ والتجزئة النصفية و معامل جتمان، وكذلك معامل ثبات الاختبار بالطريقة التجريبية بلغ (٠,٩٣٣٩) عند حسابه بطريقة بمعادلة ألفا

كرونباخ والتجزئة النصفية و معامل جتمان ، وبلغ معامل ثبات الاختبار بالطريقة التقليدية لتقدير الدرجات (٠,٨١٥٨) بمعادلة ألفا كرونباخ والتجزئة النصفية و معامل جتمان .

ومن الجدول السابق يتضح تأثر معامل ثبات الاختبار بطريقة حساب معاملات ثبات الاختبار ، بينما لا يتأثر بطرق تقدير الدرجات مما يؤكد عدم وجود تأثير مشترك بين طرق تقدير درجات الاختبار وطرق حساب معاملات الثبات للاختبار .

- **للإجابة على السؤال الرابع والذي ينص على: هل يختلف صدق الاختبار ذو الاختيار من متعدد باختلاف طرق تقدير الدرجات للمفردات؟**

للإجابة على هذا السؤال السابق تم حساب معامل الارتباط (بيرسون) بين حالتي تطبيق الاختبار قبليا وبعديا وكانت النتائج كما يلي :-

جدول رقم (١٩) يبين معاملات ارتباط (بيرسون) درجات الطالبات على الاختبار التحصيلي بطرق تقدير الدرجات المختلفة .

طريقة تقدير الدرجات	الطريقة التقليدية	الطريقة الاحتمالية	الطريقة المكافأة	الطريقة التجريبية
معامل الارتباط (بيرسون)	٠,٩٥٩	٠,٩٦٧	٠,٨٩٢	٠,٨٧٩

من جدول (١٩) يوضح اختلاف معاملات ارتباط درجات الطالبات على الاختبار التحصيلي الذي تم تطبيقه قبليا بعديا . يشير معامل الارتباط بين درجات الطالبات على الاختبار بتطبيقه قبليا وبعديا على معامل الصدق التلازمي . وبلغت قيمة معامل صدق الاختبار عند تقدير الدرجات بالطريقة التقليدية (٠,٩٥٩) وبالطريقة الاحتمالية (٠,٩٦٧) وهي أعلى قيمة لمعاملات صدق الاختبار، وبطريقة المكافأة (٠,٨٩٢) وأقل قيمة لمعامل الصدق لطريقة التقدير بالطريقة التجريبية وبلغت (٠,٨٧٩) .

ويتضح من الجدول السابق اختلاف معامل الصدق باختلاف طرق تقدير الدرجات وتتفق هذه النتيجة مع نتائج دراسة كل من دراسة جرادات (١٩٨٨م) ، ودراسة عبد الوهاب (٢٠٠١م) والتي تؤكد وجود اختلاف معاملات صدق الاختبار باختلاف طرق تقدير درجات الاختبار ، مما يؤكد أيضا قبول الفرض الرابع من فروض الدراسة الذي ينص على أنه : يختلف صدق الاختبار ذو الاختيار من متعدد باختلاف طرق تقدير الدرجات للمفردات

- **للإجابة على السؤال الخامس والذي ينص على : هل يختلف صدق الاختبار ذي الاختيار من متعدد باختلاف طرق حساب هذا الصدق ؟**

قامت الباحثة بحساب معاملات الصدق للاختبار بطرق مختلفة لحساب معاملات الارتباط كما في الجدول (٢٠) :

من الجدول (٢٠) يتضح أن معامل صدق الاختبار عند حسابه بمعامل ارتباط بيرسون وتقدير الدرجات بالطريقة التقليدية قد بلغت قيمته (٠,٩٥٩) وبلغت قيمته بمعامل كاندال (٠,٩٢١) وبمعامل سبيرمان (٠,٩٧٦) وبالجذر

التربيعي لمعامل الثبات بلغت قيمته (٠,٩٠٢) ، وبتقدير الدرجات بالطريقة المكافأة قد بلغت قيمة معامل ارتباط بيرسون (٠,٨٩٢) وبلغت قيمته بمعامل كاندال (٠,٨١٥) وبمعامل سبيرمان (٠,٨٩٨) وبالجذر التربيعي لمعامل الثبات بلغت قيمته (٠,٨٧٦)،وبتقدير الدرجات بالطريقة الاحتمالية قد بلغت قيمة معامل ارتباط بيرسون (٠,٩٦٧) وبلغت قيمته بمعامل كاندال (٠,٩١٨) وبمعامل سبيرمان (٠,٩٧١) وبالجذر التربيعي لمعامل الثبات بلغت قيمته (٠,٨٨٧) وبتقدير الدرجات بالطريقة التجريبية قد بلغت قيمة معامل ارتباط بيرسون (٠,٨٧٩) وبلغت قيمته بمعامل كاندال (٠,٨٣٦) وبمعامل سبيرمان (٠,٩٣٤) وبالجذر التربيعي لمعامل الثبات بلغت قيمته (٠,٩٦٦). ومن خلال الجدول السابق يتضح وجود اختلاف بين معاملات الصدق باختلاف طرق تقدير درجات الاختبار وكذلك باختلاف طرق تحديد معامل الصدق للطريقة الواحدة لتقدير درجات الاختبار ، مما يؤكد قبول الفرض الخامس الذي ينص على أنه يختلف معامل صدق الاختبار باختلاف المعادلات المتبعة لحساب هذا المعاملات .

جدول رقم (٢٠) يبين معاملات صدق الاختبار المختلفة وعلاقاتها بطرق تقدير درجات الاختبار

طريقة تقدير الدرجات	معامل بيرسون	معامل كاندال	معامل سبيرمان	الجذر التربيعي لمعامل الثبات
الطريقة التقليدية	٠,٩٥٩	٠,٩٢١	٠,٩٧٦	٠,٩٠٢
طريقة المكافأة	٠,٨٩٢	٠,٨١٥	٠,٨٩٨	٠,٨٧٦
الطريقة الاحتمالية	٠,٩٦٧	٠,٩١٨	٠,٩٧١	٠,٨٨٧
الطريقة التجريبية	٠,٨٧٩	٠,٨٣٦	٠,٩٣٤	٠,٩٦٦

- **للإجابة على السؤال السادس والذي ينص على : هل يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق حساب الصدق على الاختبار ذو الاختيار من متعدد ؟**

من الجدول (٢٠) يتضح أن معامل صدق الاختبار عند حسابه بمعامل ارتباط بيرسون وتقدير الدرجات بالطريقة التقليدية قد بلغت قيمته (٠,٩٥٩) وبلغت قيمته بمعامل كاندال (٠,٩٢١) وبمعامل سبيرمان (٠,٩٧٦) وبالجذر التربيعي لمعامل الثبات بلغت قيمته (٠,٩٠٢) ، وبتقدير الدرجات بالطريقة المكافأة قد بلغت قيمة معامل ارتباط بيرسون (٠,٨٩٢) وبلغت قيمته بمعامل كاندال (٠,٨١٥) وبمعامل سبيرمان (٠,٨٩٨) وبالجذر التربيعي لمعامل الثبات بلغت قيمته (٠,٨٧٦)،وبتقدير الدرجات بالطريقة الاحتمالية قد بلغت قيمة معامل ارتباط بيرسون (٠,٩٦٧) وبلغت قيمته بمعامل كاندال (٠,٩١٨) وبمعامل سبيرمان (٠,٩٧١) وبالجذر التربيعي لمعامل الثبات بلغت قيمته (٠,٨٨٧) وبتقدير الدرجات بالطريقة التجريبية قد بلغت قيمة معامل ارتباط بيرسون (٠,٨٧٩) وبلغت قيمته بمعامل كاندال (٠,٨٣٦) وبمعامل سبيرمان (٠,٩٣٤) وبالجذر التربيعي لمعامل الثبات بلغت قيمته (٠,٩٦٦). مما يؤكد اختلاف معامل الصدق باختلاف كل من طريقة تقدير درجات الاختبار ومعادلة حساب معامل الصدق وبالتالي يؤدي لرفض الفرض السادس من فروض الدراسة والذي ينص على: لا يوجد تأثير مشترك لكل من طرق تقدير الدرجات وطرق

حساب الصدق على صدق الاختبار ذو الاختيار من متعدد. ولا توجد دراسة تؤكد مثل هذه النتيجة وقد يرجع السبب في عدم وجود دراسات حول تأثر معامل صدق الاختبار أو معاملات بطرق تقدير الدرجات لأحد من السببين التاليين وهما :

- ارتباط معامل الصدق للاختبار بمعامل الثبات (أحد معادلات حساب صدق الاختبار هي الجذر التربيعي لمعامل الثبات) وبالتالي فتركيز معظم الدراسات على حساب معامل ثبات الاختبار والذي تتعدد معادلات حسابه .
- وجود أنواع متعددة من الصدق منها صدق المحكمين الذي من الصعب قياسه أو تقديره ، ومنها أيضا الصدق التلازمي والذي يعتمد على معامل الارتباط بين اختبارين تزداد بذلك فرصة تأثر معاملات الصدق بطرق تقدير الدرجات و طرق حساب معاملات الصدق .

- **توصيات الدراسة** :

في ضوء مما سبق من عرض النتائج التي توصلت إليها الباحثة يمكن الخروج بمجموعة من التوصيات ومنها :

- ضرورة وأهمية تحديد الطرق الأكثر ملاءمة لحساب معاملات صدق وثبات الاختبارات التحصيلية وذلك للحصول على نتائج تصلح للحكم على ضوئها على تحصيل الطلاب .
- ضرورة و أهمية تحديد أنسب الطرق لتقدير درجات الطلاب في الاختبارات التحصيلية ومن ثم القدرة على الحكم على مستوى تحصيلهم بموضوعية .
- ضرورة وأهمية تنوع طرق تقدير تصحيح الاختبارات بصفة عامة واختبارات الاختيار من متعدد بصفة خاصة وعدم الاقتصار على الطريقة التقليدية .
- ضرورة تقنين الاختبارات الموضوعية بصفة عامة والاختيار من متعدد بصفة خاصة ، تحديد معاملات الثبات والصدق الخاص بها ، ومن ثم اتخاذ مثل هذه الاختبارات لمحك للحكم على مستوى تحصيل الطلاب .
- ضرورة تدريب المعلمين على طرق جديدة لتقدير درجات الطلاب في الاختبارات المختلفة ومن ثم ضرورة اختيار الأنسب منها للحكم على مستوى الطلاب .

- **مقترحات الدراسة**

من خلال نتائج الدراسة يمكن التوصل لعدة مقترحات للدراسات المقترحة ومنها :

- إجراء دراسة توضح أثر اختلاف طرق تقدير درجات اختبار اختيار من متعدد على معاملات ثباته وصدقه وعلاقة ذلك بعوامل مختلفة مثل السن المادة الدراسية ، عدد أسئلة الاختبار
- إجراء دراسة توضح إمكانية الاتفاق على معادلة لتحديد معاملات ثبات وصدق الاختبارات سهلة الاستخدام .
- إجراء دراسة توضح أثر اختلاف طرق تقدير درجات الاختبار على معاملي ثباته وصدقه لدى طالبات المرحلة الإعدادية ، والابتدائية .

7 إجراء دراسة توضح العلاقة بين معاملات صدق وثبات الاختبارات واختلاف البيئة والعادات والمستويات الاقتصادية (دراسة مقارنة)

- **المراجـع** :
- **أولا : المراجع العربية** :

١- أبو لبدة، سبع محمد (١٩٨٥م) . مبادئ القياس النفسي والتقييم التربوي، ط٣، عمان.

٢- جان، محمد صالح بن علي (١٩٩٩م) . إعداد الاختبارات الموضوعية بين الواقع والمأمول، ط١، الطائف، دار الطرفين.

٣- حبيب ، مجدي عبد الكريم (٢٠٠٠م) . التقويم والقياس في التربية وعلم النفس، المجلد الثاني، ط١، القاهرة، مكتبة النهضة المصرية.

٤- الدوسري، إبراهيم بن مبارك (٢٠٠٠م). الإطار المرجعي للتقويم التربوي، ط٢، الرياض، مكتب التربية العربي لدول الخليج.

٥- الدويري، محمد عايد (١٩٨٤م). مقارنة أثر اختبارات الاختيار من متعدد واختبارات الإجابة القصيرة، في احتفاظ طلاب الصف الأول الثانوي بمادة قواعد اللغة العربية، رسالة ماجستير، جامعة اليرموك ، أربد، الأردن.

٦- الزهراني، صالح عبد العزيز (١٤٢١هـ). مقارنة طرق تقدير الثبات في القياس محكي المرجع، رسالة ماجستير، كلية التربية، جامعة أم القرى، مكة المكرمة.

٧- سواقد، ساري سليم سلامة (١٩٩٢م). بعنوان اختبار صحة الافتراضات النظرية لطرق التصحيح لأثر التخمين، ومقارنة أثر استخدام هذه الطرق على الخصائص السيكومترية للفقرة، رسالة ماجستير غير منشورة، كلية الدراسات العليا، الجامعة الأردنية.

٨- صالح، أحمد زكي (٢٠٠١م). الأسس النفسية للتعليم الثانوي، ط٢، القاهرة، مكتبة النهضة المصرية.

٩- عبد الوهاب، صلاح شريف (٢٠٠١م). أثر بعض الطرق الوزنية لتقدير الدرجات على صدق الاختبارات مرجعية المحك ذات الاختيار من متعدد. مجلة كلية التربية ببنها، عدد أكتوبر ٢٠٠١م، المجلد ١٢ ، العدد ٤٩.

١٠- عبده، عبد الهادي السيد؛ عثمان، وفاروق السيد (٢٠٠٢م) . القياس والاختبارات النفسية، ط١، القاهرة، دار الفكر العربي.

١١- العبيدي، غانم سعيد ؛ الجبوري، حنان عيسى سلطان (١٩٨٤م): أساسيات البحث العلمي بين النظرية والتطبيق، ط١، دار العلوم للطباعة والنشر. الرياض. المملكة العربية السعودية.

١٢- العبيدي، غانم سعيد شريف؛ و الجبوري، حنان عيسى سلطان (١٩٨١م): أساسيات القياس والتقويم في التربية والتعليم، ب ط.

١٣- عدس، عبد الرحمن (٢٠٠٢م): دليل المعلم في بناء الاختبارات التحصيلية، ط٣، عمان، دار الفكر.

١٤- العدل، عادل محمد محمود(١٩٨٦م): أثر بعض طرق تقدير الدرجات للمفردات على ثبات الاختبارات ذات الاختيار من متعدد، رسالة ماجستير غير منشورة، مودعة بكلية التربية، جامعة الزقازيق.

١٥- العساف ، صالح محمد (٢٠٠٣م). المدخل إلى البحث في العلوم السلوكية، ط الثالثة، مكتبة العبيكان ، الرياض، السعودية.

١٦- العلي، محمد محمود البشير (١٩٨٩م). المقارنة بين أثر اختبارات الاختيار من متعدد واختبارات الإجابة القصيرة واختبارات المزج من النوعين معا على بعض الخصائص السيكومترية للاختبار، رسالة ماجستير، جامعة اليرموك، أربد ، الأردن.

١٧- عليان، ريم محمد مصطفى (٢٠٠٢م) . تقدير الثبات للعلامات المدرسية في مدينة إربد الكبرى في الفصل الدراسي الأول ٢٠٠٢/٢٠٠٣م، رسالة ماجستير، جامعة اليرموك، اربد، الأردن.

١٨- العيسوي ، عبد الرحمن محمد (٢٠٠٣م) . الاختبارات والمقاييس النفسية والعقلية، ب ط، الإسكندرية، منشأة المعارف.

١٩- أبو حطب، فؤاد؛ وسيد أحمد عثمان (١٩٩٧م): التقويم النفسي، ط٤، القاهرة، مكتبة الأنجلو المصرية.

٢٠- فرج ، صفوت (٢٠٠٠م). القياس النفسي، ط٤، القاهرة، الانجلو المصرية.

٢١- القضيب، نورة عبد الرحمن (٢٠٠١م) . أساليب التحقق من ثبات وصدق الاختبارات المستخدمة في رسائل الماجستير في أقسام علم النفس في بعض جامعات المملكة العربية السعودية. رسالة ماجستير غير منشورة، كلية التربية، قسم علم النفس، جامعة الملك سعود.

٢٢- ليونا أ. تايلر (١٩٨٣م) . الاختبارات والمقاييس، ترجمة سعد عبد الرحمن، مراجعة محمد عثمان نجاتي، ط١، بيروت، دار الشروق.

٢٣- محمود، سليمان محمد سليمان (١٩٨٦م). أثر بعض طرق تقدير الدرجات على ثبات الاختبارات المرجعية ذات الاختيار من متعدد، كلية التربية، قسم علم النفس، جامعة الزقازيق.

٢٤- النبهان ، موسى (٢٠٠٤م). اساسيات القياس في العلوم السلوكية، ط١، عمان، دار الشروق للنشر والتوزيع.

٢٥- نبيل عبد الهادي (٢٢٠١م).القياس والتقويم التربوي واستخدامه في مجال التدريس الصفي، ط٢، عمان، دار وائل.

٢٦- الهويدي، زيد (٢٠٠٤م). أساسيات القياس والتقويم التربوي، ط١، الإمارات العربية المتحدة، دار الكتاب الجامعي، .

٢٧- ويليام أ. مهرنز وايرفن ج. ليمان (٢٠٠٣م). القياس والتقويم في التربية وعلم النفس، ترجمة هيثم كامل الزبيدي، مراجعة ماهر أبو هلاله، ط١، العين، الإمارات العربية المتحدة، دار الكتاب الجامعي.

- **ثانياً : الدراسات الأجنبية:**

54) Aiken، l. (2000)، Psychological Testing and Assessment. Boston. Boton.

55) American Educational Research Association، American Psychological Association، National Council on Measurement in Education (1995). Standards for Education and psychological Testing. Washington، Dc: Author.

56) Anastasi، A. and Urbina، S. (1997). Psychological Testing. New

Jersey 07458: Upper Saddle River.

57) **Baldanf، R.B. (1982): The effect of guessing and item dependence on the reliability and validity of recognition based cloze tests. J. of E. and psychol. M. ، Vol . 42**

58) **Bush، m. (2001). A Muliple choice Test that Rewards Partial knowledge، Journal of futher and higher education، Vol. 25، No (2).**

59) **Edel، R. Frisbie، D. (1995). Essentials of Educational Mesurement. New Jersey.**

60) **Hambleton، R.K. ، Dennis M. Roberts and Ross E. Traube (1970): A comparison of the Reliability and Validity of Two Method for Asspssing Partial Knowledge on a Multiple choice tests. J. of E.M.، Vol. 7، No. 1**

61) **Hammersley، M. (1987). Some Notes on the terms Validity and Reliability. British Educational Research Journal، 13، 1**

62) **Hanna، G.S. and Long، C.A. (1974): Effect of Answer Until Correct Testing on Reliability، Perceptual and Motor Skills، 1974.**

63) **Jaradat – Derar، Tollefson، Nona (1988): The Impact of Alternative Scoring Procedures for Multiple Measurement' Vol، 48، No. 3، pp. 627 – 635.**

64) **Kansup، W. & Hakstian، A.R. (1975): A Comparison of Several Methods of Assessing Partial Knowledge in Multiple Choice Tests: Scoring procedures، Journal of Educational Measurement، Vol. 12، No. 4، pp. 219- 229.**

65) **Kline، P (1993) The Handbook of Psychological Testing،**

66) **Koehler4، R.A. (1971): A Comparison of the Validities of Conventional Choice Testing and Various Confidence Marking Procedures، Journal of Educational Measurement، Vol. 8، No. 4.**

67) **Michael، J.J. (1968): The Reliability of A multiple choice Examination Under Various test Taking Instruction، Journal of Educational Measurement، Vol. 5**

68) **Rippey، R.M.، (1970): A comparison of Five Different Scoring Function for confidence tests. J.of E، M. Vol. 7. No. 3**

69) **Rudner، Lawrence. M. (1994) . Questions to ask when Evaluating Tests (ERIC/ AE Digest April، ED 385607)**

70) **Sabers، D.L.، and Gordon W.White (1969): The effect of differential weighting of indidiual item responses on the predictive validity and reliability of on aptitude test. J. of E.M.،Vol.6، No.2.**

******

((نموذج اشتراك في مجلة دراسات عربية في التربية وعلم النفس))

سعادة / الأستاذ الدكتور : رئيس تحرير مجلة دراسات عربية في التربية وعلم النفس السلام عليكم ورحمة الله وبركاته وبعد ،،

أرغب الاشتراك في المجلة لمدة : (سنة واحدة ☐) (سنتان ☐) (ثلاث سنوات ☐)

على أن تصلني نسخ أعداد المجلة على عنواني البريدي الموضح بهذا النموذج.

الاســـــــم	.....
الوظيفـــــــة	.....
جهـــة العمـــل	.....
الجنسيـــــــة	.....
عنوان المراسلـــة	.....
البريد الإلكتروني	.....
الهاتف/ الفاكس	.....

اسم المشترك : .....

التوقيـــــع : .....

- قيمة الاشتراك للأفراد بالدول العربية : (سنة ٢٠٠ريالا ، سنتان ٣٥٠ريالا ، ثلاث سنوات ٥٠٠ ريالا).
- قيمة الاشتراك للأفراد بباقي دول العالم: (سنة ٥٠ دولار ، سنتان ١٠٠دولا، ثلاث سنوات ١٥٠دولار).
- قيمة الاشتراك للمؤسسات بالدول العربية: (سنة ٢٥٠ريالا، سنتان ٤٥٠ ريالا، ثلاث سنوات ٦٥٠ريالا).
- قيمة الاشتراك للمؤسسات بباقي دول العالم: (سنة ١٠٠دولار، سنتان ٢٠٠دولار، ثلاث سنوات ٢٥٠دولار).
- قيمة الاشتراكات هذه شاملة تكاليف البريد العادي ، ومن يرغب في البريد الممتاز يتحمل الفرق.
- تسدد قيمة الاشتراكات مباشرة أو بحوالة بنكية بالريال السعودي باسم رئيس تحرير المجلة (أ. د / ماهر إسماعيل صبري) على بنك الرياض فرع حي أحد بالمدينة المنورة ، رقم آي بان : SA1620000001070160599940
- يمكن سداد قيمة الاشتراكات بالجنيه المصري مباشرة لمكتب المجلة بجمهورية مصر العربية، أو بحوالة بنكية باسم رئيس التحرير (أ. د / ماهر إسماعيل صبري)على بنك فيصل الإسلامي المصري فرع بنها رقم الحساب ١٨٥٠٦
- ترسل صورة من قسيمة تحويل الاشتراكات على البريد الإلكتروني لرئيس التحرير mahersabry21@yahoo.com
- يرسل هذا النموذج بعد تعبأة بياناته عبر البريد الإلكتروني لرئيس تحرير المجلة، أو عبر البريد العادي على عنوان رئيس التحرير الحالي : المدينة المنورة ، جامعة طيبة ، كلية التربية ، قسم المناهج وطرق التدريس. أو على عنوان مكتبنا بمصر : ١ش أحمد ماهر متفرع من ش الشعراوي، أتريب ، بنها .