

أثر ثلاثة نماذج لترتيب فقرات اختبار تحصيلي في التقويم التربوي على الخصائص السيكومترية للاختبار وفقراته

د.شاهر خالد سليمان

كلية التربية والآداب - جامعة تبوك

المملكة العربية السعودية

الملخص

المقدمة

يعد التقويم التربوي عنصراً أساسياً في منظومة العملية التعليمية ، لما له من دور فاعل في تطويرها وتقدمها ، ولمعرفة مدى التقدم في تحقيق أهداف العملية التعليمية يقوم عضو هيئة التدريس في الجامعة بملاحظة سلوك طلابه باستمرار ، وجمع المعلومات عن هذا السلوك بأدوات وطرق مختلفة ، ولا بد من استخدام أدوات دقيقة لقياس هذا التطور والتقدم ، وفي ضوء هذه النتائج تتخذ القرارات المناسبة من ترفيع الطلبة من مستوى إلى آخر ، وتصنيفهم وما إلى ذلك ، وتعد الاختبارات التحصيلية من أهم وسائل التقويم التي يستخدمها أعضاء هيئة التدريس في الجامعات والمعلمون في المدارس. (الشرفين ، 2006)

كما توفر الاختبارات معلومات يمكن أن تستخدم في قياس مستوى تحصيل الطلاب ومدى نجاح عملية التعليم وتشخيص جوانب ضعف الطلاب ووضع العلاج المناسب وتصنيف الطلبة والتميز بينهم ، وتعد اختبارات الاختيار من متعدد من أكثر أشكال الاختبارات التحصيلية انتشاراً وشيوعاً في الجامعات لسهولة تصحيحها ومرونة استخدامها ، كما تتميز اختبارات الاختيار من متعدد بأنها تقيس مستويات عقلية متنوعة لدى الطلبة كما يمكن استخدامها في مختلف مجالات الدراسة (سليمان ، 2010) .

هدف البحث الحالي إلى التعرف على أثر ترتيب فقرات اختبار تحصيلي على معاملات تمييز وصعوبة الفقرات ، وعلى صدق وثبات الاختبار ، ولتحقيق أهداف البحث تم إعداد اختبار مكون من (30) فقرة في مقرر التقويم التربوي لطلبة البكالوريوس في جامعة تبوك ، ورتبت فقراته على النحو الآتي : اختبار فقراته مرتبة حسب مستوى الهدف الذي تقيسه الفقرة وفق تصنيف بلوم للأهداف (تذكر ، استيعاب ، تطبيق ...) ، واختبار فقراته مرتبة حسب تسلسل المحتوى ، واختبار فقراته مرتبة تصاعدياً حسب قيم معاملات صعوبة الفقرات . طبقت أنماط الاختبار الثلاثة على عينة مكونة من (300) طالب من المسجلين في المقرر الدراسي في الفصل الأول 1432 / 1433 هـ . دلت النتائج على ما يلي : لا توجد فروق دالة إحصائية في المتوسطات الحسابية لمعاملات صعوبة فقرات الاختبار تعزى لاختلاف نمط ترتيب فقراته ، و توجد فروق دالة إحصائية بين المتوسطات الحسابية لمعاملات تمييز الفقرات تعزى لنموذج ترتيب الفقرات لصالح الاختبار الذي رتبت فقراته تصاعدياً حسب قيم الصعوبة ، ولا توجد فروق دالة إحصائية بين معاملات ثبات وصدق الاختبار تعزى لنموذج ترتيب فقراته ، وتوصل البحث إلى مجموعة من التوصيات .

وترتيب فقرات الاختبار حسب درجة صعوبة فقراته تشكل هذه الطريقة الاتجاه العام في ترتيب الفقرات ، فإذا كانت الفقرات من نفس النوع فإن ترتيبها حسب الصعوبة ، حيث يوفر هذا الترتيب للطالب الدافعية للاستمرار في محاولاته للإجابة ، ربما كان الترتيب الأنسب ، حيث يوفر الدافعية للاستمرار في محاولاته للإجابة عندما يشعر بالتعزيز الفوري بسبب قدرته على الإجابة عن الأسئلة الأولى من الاختبار والتي تسمى بالأسئلة التشجيعية أو الأسئلة الماصة للصدمة ، ولكن ما هو مؤشر الصعوبة المتوفر ، أو الذي يمكن توفيره لإجراء هذا الترتيب ؟

ويعتمد عضو هيئة التدريس أكثر من مؤشر لتقدير صعوبة السؤال بعضها مؤشرات منطقية (Logical indices) ويعني أن عضو هيئة التدريس سيعتمد على المنطق في تقديره لصعوبة الفقرة ، مثل تقديره للمستوى العقلي الذي تقيسه الفقرة ، حيث يتوقع زيادة صعوبتها بزيادة المستوى العقلي الذي تقيسه وفق تصنيف بلوم للأهداف . (الحريري، 2007) .

و ترتيب حسب المحتوى : ويقصد بهذا الترتيب تسلسل الفقرات في ورقة الاختبار حسب تسلسل محتوى المادة الدراسية ، أي ترتيب الفقرات حسب تسلسل المادة الدراسية ، بمعنى أن كل موضوع يعتمد على موضوع آخر سابق له . (عودة ، 2005) .

وأجرى سكاير ونورنن (Schuyler & Nornen، 1972) دراسة تناولت أثر ترتيب الفقرات وفقاً لصعوبتها في اختبار تحصيلي من نوع اختيار من متعدد ، وأجرى دانيال (Dinnal، 1988) دراسة استخدم فيها طريقتين في ترتيب الفقرات في اختبار اختيار من متعدد هما الصعوبة الإحصائية للفقرة والصعوبة المنطقية للفقرة ، وأجرى كارستين (Carstens، 1982) دراسة هدفت إلى معرفة أثر ترتيب فقرات اختبار تحصيلي على أداء الطلبة من حيث الترتيب العشوائي لها أو وفقاً لترتيبها في المحتوى

ويعد التقدم والتطور في مجال القياس الركيزة الأساسية لتقدم وتطور العلوم مهما اختلفت أنواعها وتغيرت مشاربها فهي جميعاً تقاس وتقيم بمدى الدقة التي وصل إليها القياس في مجال أبحاثها ودراساتها ، وبالرغم من أن محور القياس في مجال التربية هو الإنسان بكل تغيراته ومرونته وعدم ثباته واستقراره إلا أن علماء القياس اهتموا ببناء أدوات قياس مع الأخذ بالاعتبار كل المتغيرات ووصلوا إلى دقة معقولة حول علمية وموضوعية الاختبارات خاصة فيما يتعلق بخصائصها السيكمومترية من صعوبة وتمييز لفقراتها وصدق وثبات للاختبارات ، ولأن مدى قدرة الاختبار على قياس السمات المختلفة عند الطلاب بقدر عال من الصدق والثبات ويقدر أقل من الأخطاء تعتبر من أهم ركائز عملية القياس والتقويم . (سنان، 2006) .

ويرى بعض علماء القياس أن الأخطاء التي تأتي من بنية المقياس من أخطر الأخطاء ، لذلك جرت محاولات عديدة من علماء القياس والباحثين فيه لاعتماد قواعد وإجراءات دقيقة في بناء أدوات القياس ، غايتها الحد أو التقليل من الأخطاء المتأتية من المقياس ، كما اتفقوا على ضرورة توافر الخصائص السيكمومترية للاختبار وفقراته بدرجة عالية ودقيقة ، كما ينبغي لمصمم الاختبار أن يكون على خبرة ودراية مناسبة بصياغة فقرات الاختبار حتى يخرج الاختبار بصورة صحيحة وتكون نتائجه موثوقاً بها ، ومن الإجراءات التي تتعلق ببنية الاختبار التي يتناولها البحث ترتيب فقرات الاختبار ، ومن الطرق المستخدمة في ترتيب فقرات الاختبار ؛ ترتيب حسب نوع (شكل) الفقرة : إذا احتوى الاختبار على أكثر من نوع من الفقرات ، فمن الأفضل ترتيبها حسب النوع ، بمعنى أن تظهر جميع الفقرات التي من نفس النوع كمجموعة جزئية لعدة مزايا منها : سهولة توضيح التعليمات وسهولة التصحيح والمحافظة على التهيؤ العقلي (mental set) للطالب أثناء الإجابة . (علام ، 2006) .

نتائج الدراسة أنه لا يوجد فروق ذات دلالة إحصائية بين متوسطات درجات أفراد عينة البحث في الأشكال الثلاثة للاختبار، ولا يوجد فروق ذات دلالة إحصائية بين متوسط درجات الطلاب الذكور أو متوسط درجات الطالبات الإناث، وأن أعلى المتوسطات الحسابية بالنسبة للأشكال الثلاثة للاختبار كان الترتيب الذي يبدأ بالأسئلة السهلة (الترتيب التصاعدي) .

وأجرى الليحاني (2010) دراسة بهدف التعرف على أهم طرق ترتيب مفردات الاختبار من حيث تسلسل محتوى المقرر الدراسي للمفردات ومستوى السهولة والصعوبة، والكشف عن نمط الترتيب الأفضل لمفردات الاختبار من حيث الصدق والثبات، ولتحقيق هدف الدراسة تم بناء اختبار تحصيلي متعدد البدائل في اليايين الأول والثاني من مقرر الرياضيات للصف الثاني الثانوي قسم العلوم الطبيعية وهي كما يلي: العمليات الثنائية والرمز K والمصفوفات والمحددات، وبناءً على ذلك تم بناء خمسة نماذج للاختبار نفسه، يتكون كل منها من (25) مفردة على النحو التالي: الأول؛ تم ترتيب المفردات حسب تسلسل عرض المحتوى، والثاني؛ تم ترتيب المفردات من السهل إلى الصعب، والثالث؛ تم ترتيب المفردات من الصعب إلى السهل، والرابع؛ تم ترتيب المفردات ترتيباً متبادلاً (دائرياً)، الخامس؛ تم ترتيب المفردات ترتيباً عشوائياً، وتم تطبيق النماذج الخمسة بشكل عشوائي على (600) طالب من طلاب الصف الثاني الثانوي بمدينة مكة المكرمة. دلت النتائج على أنه توجد فروق ذات دلالة إحصائية بين قيم معاملات الثبات تعزى لاختلاف ترتيب مفردات الاختبار بين النموذجين (الترتيب حسب تسلسل عرض المحتوى والترتيب من السهل إلى الصعب) لصالح نموذج الترتيب حسب تسلسل عرض المحتوى، كما أنه توجد فروق دالة إحصائية بين معاملات الثبات للنموذجين (الترتيب من السهل إلى الصعب والترتيب من الصعب إلى السهل) لصالح نموذج الترتيب من الصعب إلى

الدراسي، كما أجرى ديرك (Dreak, 1984) دراسة تناول فيها أثر ترتيب الفقرات في اختبار تحصيلي وفقاً لصعوبتها على أداء الطلبة.

وأجرى العمري (2011) دراسة هدفت إلى الكشف عن أثر ترتيب الفقرات في المقاييس النفسية على خصائص المقياس وخصائص فقراته، لذا فقد استخدم في هذه الدراسة مقياس الثقة بالنفس، ثم إعادة ترتيب فقراته بأربع طرق، الثلاثة الأولى هي: ترتيب الفقرات تنازلياً حسب مؤثر كل من: تمييز الفقرة، وثبات الفقرة، وصدق الفقرة، إضافة إلى الترتيب الرابع، وهو ترتيب الفقرات عشوائياً .

واستخدم الباحث تصميم القياسات المتكررة في عرض المقياس على عينة من المفحوصين بلغت (264) طالبا، تم تقسيمها إلى أربع مجموعات كل منها (66) مفحوصاً، تلقى كل مفحوص الأنماط الأربعة للمقياس ولكن بترتيب مختلف من مجموعة لأخرى، نتائج الدراسة أشارت إلى تفوق أسلوب ترتيب الفقرات اعتماداً على مؤشرات ثباتها، حيث أعطي هذا الترتيب أعلى معامل ثبات للمقياس بالإضافة إلى أن هذا الترتيب قد تقدم على الطرق الأخرى في تحسين متوسطات الخصائص السيكومترية للفقرات نفسها .

وأجرى رمضان (2010) دراسة بهدف التعرف على مدى تأثير ترتيب فقرات الاختبار حسب درجة صعوبتها، على تحصيل طلاب المرحلة الثانوية لمادة الرياضيات في المنوفية، وقد تكونت عينة الدراسة من (200) طالب وطالبة بالصف الثاني الثانوي (علمي) حيث حل الباحث محتوى الدوال الحقيقية بهدف الحصول على المفاهيم الرياضية في اختبار يحوي على (23) فقرة من نمط الاختبار من متعدد، حيث تم حساب صدق وثبات الاختبار ومعاملات الصعوبة لكل مفردة، وقسم الاختبار إلى ثلاثة أقسام، اختبارات رتبته أسئلتها بالأشكال الثلاثة للاختبار ترتيب تصاعدي، تنازلي، عشوائي). دلت

السهل ولا توجد فروق ذات دلالة إحصائية بين قيم معاملات صدق الاختبار بين النماذج الخمسة. وتختلف الدراسة الحالية عن الدراسة السابقة بأنها جمعت ثلاثة نماذج معا وطبقت على اختبار تحصيلي في مقرر جامعي يُدرس في الجامعة وتناول جميع الخصائص السيكومترية للاختبار التحصيلي من صدق وثبات الاختبار وصعوبة وتمييز فقراته، حيث تناولت معظم الدراسات اختبارات على طلاب المدارس، أو مقاييس نفسية، وقد أضاف البحث الحالي نمط ترتيب حسب مستوى الهدف الذي تقيسه فقرة السؤال وفق تصنيف بلوم للأهداف السلوكية، ليغطي معظم أنماط ترتيب فقرات الاختبار المطروحة وأثرها على جميع الخصائص السيكومترية للاختبار من حيث معاملات الصدق والثبات، والصعوبة، والتمييز.

مشكلة الدراسة وأسئلتها

تعد الاختبارات من أدوات التقييم المهمة، حيث تستخدم نتائجها في اتخاذ القرارات التربوية، كلما زادت أهمية القرار الذي يتوقع أن يتم اتخاذه كلما زادت الحاجة إلى أن تكون المعلومات التي نحصل عليها من الاختبار دقيقة وذات صلة وثيقة بالفرض الذي أعد من أجله، ومن هنا برز الاهتمام بالتخطيط المحكم لإعداد الاختبار سواء فيما يتعلق بتحديد الأهداف والفرض من الاختبار، وتحضير الفقرات وتحليلها أو فيما يتعلق بإخراج الاختبار وتطبيقه، وجميع العوامل التي يمكن أن تؤثر على نتائجه.

وكان للثورة العلمية الحديثة في مجال البحث العلمي والتي تبنت الطرق الاستقرائية، الأثر الكبير في دفع حركة القياس، وتطوير العديد من أدوات القياس المختلفة وفق أساليب وطرق علمية منظمة تبدأ بتحديد الهدف وكتابة الفقرات التي تعكس الخصائص والسمات المراد قياسها وانتهاء بتعيين الخصائص السيكومترية للأداة، وتقدير معايير الدرجات.

ويكثر أعضاء هيئة التدريس في الجامعات من استخدام الاختبارات الاختيارية كأداة في قياسهم لتحصيل الطلبة، حيث تتطلب الأسئلة الاختيارية مهارات معينة خصوصا اختبارات الاختيار من متعدد والاهتمام بإخراج الاختبار يوفر خصائص سيكومترية تجعل الفقرات أفضل ما يمكن من حيث الصعوبة والتمييز، وتجعل من الاختبار أفضل ما يمكن من حيث الصدق والثبات، وبالتالي يقدم الاختبار نتائج دقيقة تعكس المستوى الحقيقي للطلاب، مما يجعل عضو هيئة التدريس في الجامعة على وعي بأهمية تحديد الأهداف وترتيب فقرات الاختبار، وتتبع أدبيات البحث اتضح أن هناك اختلافاً واضحاً في نتائج الدراسات التي اهتمت بتأثير نمط ترتيب فقرات الاختبار على الخصائص السيكومترية، وقد يعود ذلك إلى اختلاف العينات وأدوات البحث، وتقصي الحقيقة في هذا المجال يتطلب مزيداً من الدراسات في بيئتنا العربية حتى يمكن تقديم التوصيات العلمية التي ينبغي أخذها في الاعتبار عند اختيار أو بناء أداة البحث، وبناء على ما تقدم ونتيجة لندرة الأبحاث في هذا المجال على المستوى المحلي والعالمي، فإن الدراسة الحالية محاولة متواضعة للإسهام في التحري عن علاقة نمط ترتيب الفقرات في الاختبار بخصائصه السيكومترية ومدى اختلاف هذه العلاقة وعلى وجه التحديد سيركز البحث الحالي على محاولة الإجابة عن التساؤلات الآتية:

1. هل توجد فروق دالة إحصائية في المتوسطات الحسابية لمعاملات صعوبة فقرات اختبار تحصيلي في التقويم التربوي تعزى لنموذج ترتيب فقراته؟
2. هل توجد فروق ذات دلالة إحصائية في المتوسطات الحسابية لمعاملات تمييز فقرات اختبار تحصيلي في التقويم التربوي تعزى لنموذج ترتيب فقراته؟
3. هل توجد فروق ذات دلالة إحصائية في معاملات ثبات اختبار تحصيلي في التقويم التربوي تعزى لنموذج ترتيب فقراته؟

1

بحوث ودراسات

على فرص العمل، وما يتخذ بناء عليها من قرارات تعليمية تتعلق بالطلبة، كما يمكن للبحث الحالي أن يساعد القائمين على تطوير الاختبارات في المؤسسات التعليمية، على الاهتمام ببنية الاختبار وإعداده وإخراجه، كما تتمثل أهمية الدراسة في بيان آلية بناء اختبار تحصيلي في أحد المقررات الدراسية الجامعية الذي يدرس في جامعة تبوك، وفي جامعات أخرى.

مصطلحات البحث :

الخصائص السيكومترية وتتكون من :

معامل صعوبة الفقرة : هونسبة الطلاب الذين أجابوا عن الفقرة إجابة صحيحة إلى مجموع الطلاب الذين حاولوا الإجابة عن الفقرة .

معامل تمييز الفقرة : هو قيمة معامل الارتباط بين الأداء على الفقرة والدرجة الكلية على الاختبار .

معامل ثبات الاختبار : هو معامل ثبات الاتساق الداخلي المحسوب باستخدام معادلة ألفا كرونباخ والطريقة النصفية.

معامل صدق الاختبار : يتكون من نوعين :

الصدق التنبؤي : هو قيمة معامل ارتباط بيرسون المحسوب بين درجات الطلاب على الاختبار الذي أعده الباحث ودرجاتهم على الاختبار النهائي في ذلك المقرر .

الصدق التلازمي : هو قيمة معامل ارتباط بيرسون المحسوب بين درجات الطلاب على الاختبار الذي أعده الباحث ودرجاتهم على الاختبار النصفية في نفس المقرر .

اختبار تحصيلي : عينة من فقرات من نوع اختيار من متعدد أربعة بدائل أعدها الباحث في الموضوعات التالية؛ خصائص ومستويات القياس ، أنواع الاختبارات وخصائصها، تصميم الاختبارات واستخداماتها ، تطبيق الاختبارات ، والموضوعات

4. هل توجد فروق ذات دلالة إحصائية في معاملات صدق اختبار تحصيلي في التقويم التربوي تعزى لنموذج ترتيب فقراته ؟

أهداف البحث

يهدف البحث الحالي إلى ما يأتي :

- الكشف عن أثر اختلاف ثلاثة نماذج من ترتيب فقرات اختبار تحصيلي في التقويم على معاملات صعوبة وتمييز فقراته .
- الكشف عن أثر ثلاثة نماذج من ترتيب فقرات اختبار تحصيلي في التقويم التربوي على معاملات صدق وثبات الاختبار .
- تحديد نمط الترتيب الأمثل لترتيب فقرات اختبار تحصيلي في التقويم التربوي الذي يتميز بمعاملات صدق وثبات مرتفعة .

أهمية البحث

تأتي أهمية البحث بمقدار ما يضيفه إلى المعرفة من الناحيتين النظرية والتطبيقية ، وعليه فإن الأهمية للبحث الحالي يمكن أن تتحدد في كونه سيقدم مؤشراً قياسياً جديداً لزيادة دقة أدوات القياس (الاختبار)، حيث تعد الاختبارات التحصيلية من أكثر أدوات القياس انتشاراً وشيوعاً في الجامعات ومؤسسات التعليم الأخرى، إذ إن ما توصل إليه البحث من تفضيل لأحد أنماط ترتيب فقرات الاختبار سيرشد الباحثين إلى أي نمط من ترتيب فقرات الاختبار ستعطي موثوقية وصدق لنتائج الاختبار التي تبني عليها قرارات منها تتعلق بانتقال الطلبة من مستوى دراسي إلى آخر ، وبمستوى إتقان الطلبة للمهارات والمعارف التي اكتسبوها أثناء الدراسة في الجامعة، مما يعطي صورة دقيقة وصادقة عن مستوى الخريجين من الجامعات ، وينعكس إيجاباً على قدرة الطلبة على المنافسة في سوق العمل، ولأهمية دقة تقدير مستويات الأداء في الجامعات وانعكاسها

هي مفردات لمقرر التقويم التربوي للفصل الدراسي الأول من العام الدراسي 1432 هـ .

محددات البحث

- اقتصار البحث على الطلاب الذكور المسجلين في مقرر التقويم التربوي في جامعة تبوك في المملكة العربية السعودية في الفصل الدراسي الأول 1432 / 1433 هـ .
- اقتصار البحث على اختبار تحصيلي في مقرر التقويم التربوي في الموضوعات التالية: خصائص ومستويات القياس، أنواع الاختبارات وخصائصها، تصميم الاختبارات واستخداماتها ، تطبيق الاختبارات .

الطريقة والإجراءات

عينة البحث: تكونت عينة البحث من (300) طالب وهم جميع الطلاب المسجلين في مقرر التقويم التربوي في الفصل الدراسي الأول 1432 / 1433 هـ، وهم جميع أفراد مجتمع البحث .

أداة البحث

تم بناء الاختبار في موضوعات القياس والتقويم وفق الخطوات الآتية :

- تحديد الغرض من الاختبار ، وهو قياس تحصيل الطلاب في موضوعات التقويم التربوي.
- تحديد محتوى الاختبار بناء على قائمة توصيف مقرر التقويم التربوي والأهداف السلوكية التي ينبغي تحقيقها لدى الطلبة .
- اشتملت موضوعات الاختبار على ما يلي : خصائص ومستويات القياس ، الاختبارات أنواعها وخصائصها ، تصميم الاختبارات واستخداماتها ، تطبيق الاختبارات .
- تم صياغة الأهداف السلوكية التي تضمنتها الموضوعات والتي ينبغي تحقيقها لدى الطلاب.

• إعداد جدول مواصفات لبناء الاختبار حيث تم ربط مستويات الأهداف السلوكية بمحتوى الموضوعات الدراسية التي هي موضوع الاختبار.

• تم صياغة (40) فقرة من نوع اختيار من متعدد بأربعة بدائل أحدها بديل صحيح ، لقياس كل هدف من الأهداف السلوكية ، وقد تطلب قياس بعضها أكثر من فقرة واحدة ، وقد روعيت الأسس الفنية في كتابة هذا النوع من الفقرات ، ومطابقتها للهدف السلوكي الذي تقيسه من حيث المحتوى ، ومستوى الهدف .

• للتأكد من صدق المحتوى ، تم عرض فقرات الاختبار على ستة محكمين من أعضاء هيئة التدريس في الجامعات والذين يدرسون القياس والتقويم ، وذلك لأخذ وجهات نظرهم حول الفقرات والاختبار واقترح ما يروونه مناسباً ، وطلب من كل منهم أن يبدي رأيه من حيث : درجة الاتفاق بين الأهداف السلوكية والهدف الأساسي للاختبار وتمثيلها لها ، ومدى قياس الفقرات للهدف السلوكي المتعلق بها ، ومدى اتساق الاختبار مع الهدف الأساسي منه ، وبناء على ملاحظاتهم تم حذف بعض الفقرات التي أجمع المحكمون على حذفها وبلغ عددها (7) ، وبعد ذلك عدل الباحث الفقرات وراجعها وبلغ عددها في صورتها الأولية (38) فقرة ، حيث أجمع المحكمون على أنها عينة كافية وممثلة للمجال السلوكي الذي يقيسه الاختبار

• **التجريب الأولي للاختبار** : قام الباحث بتطبيق الاختبار المكون من (38) فقرة على عينة استطلاعية تكونت من (30) طالباً ، تم اختيارهم عشوائياً من مجتمع البحث وخارج عينة الدراسة ، بغرض التعرف على مستوى صعوبة الفقرات وتحديد الزمن المناسب للاختبار والتأكد من وضوح الصياغة اللغوية للفقرات والكشف عن الصعوبات المتوقعة التي

وذلك بإيجاد نسبة الطلاب الذين أجابوا عن الفقرة إجابة صحيحة من بين المفحوصين الذين حاولوا الإجابة عنها وإيجاد معامل تمييز والجدول رقم (1) يوضح معاملات الصعوبة والتمييز لكل فقرة من الاختبار .

تواجه الطلاب ، وكذلك إيجاد معامل الثبات للاختبار ، ولم تواجه الطلاب أية صعوبات في الإجابة عن الفقرات .

لقد أجريت عملية التحليل للبيانات التجريبية الأولية حيث تم حساب معامل الصعوبة لكل فقرة ،

جدول رقم (1)

معاملات الصعوبة والتمييز لفقرات الاختبار على العينة الاستطلاعية

رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
١	٠,٤٢	٠,٦٧	١٤	٠,٢٢	٠,١٨	٢٧	٠,٤٠	٠,٤٥
٢	٠,٧٤	٠,٤٥	١٥	٠,١٩	٠,٠٩	٢٨	٠,٧٠	٠,٣٩
٣	٠,٣٩	٠,٤١	١٦	٠,٥٦	٠,٥٠	٢٩	٠,٢٦	٠,١٧
٤	٠,٤٠	٠,٥٨	١٧	٠,٧٧	٠,٤٣	٣٠	٠,٢٦	٠,٤٢
٥	٠,١٢	٠,٠٨	١٨	٠,٦٥	٠,٦٧	٣١	٠,٣٥	٠,٤٢
٦	٠,٣٥	٠,٤٢	١٩	٠,٩٣	٠,٢٨	٣٢	٠,١٤	٠,١٠
٧	٠,٣٣	٠,٥٨	٢٠	٠,٣٧	٠,٤٣	٣٣	٠,٦٥	٠,٧٥
٨	٠,٥٨	٠,٤٧	٢١	٠,٥١	٠,٥٠	٣٤	٠,٤٩	٠,٤١
٩	٠,٢٣	٠,٤٢	٢٢	٠,٢٦	٠,٤٧	٣٥	٠,٦٥	٠,٤٥
١٠	٠,٦٣	٠,٥٠	٢٣	٠,٧٠	٠,٥٨	٣٦	٠,٤٤	٠,٤٥
١١	٠,٦٥	٠,١٨	٢٤	٠,٢٥	٠,٤٧	٣٧	٠,٣٣	٠,٤٨
١٢	٠,٣٥	٠,٣٩	٢٥	٠,٤٩	٠,٥٠	٣٨	٠,٣٧	٠,٤٧
١٣	٠,٤٩	٠,٥٠	٢٦	٠,٢٠	٠,٠٨			

حسب الصعوبة المنطقية وهي هنا حسب مستوى الهدف الذي تقيسه الفقرة وفق تصنيف بلوم (تذكر، استيعاب، تطبيق، تحليل، تركيب، تقويم) ، والنموذج الثاني: اختبار فقراته مرتبة حسب تسلسل المحتوى الذي هو موضوع الاختبار ، النموذج الثالث اختبار فقراته مرتبة حسب قيم معاملات الصعوبة وكان الترتيب من السهل إلى الصعب .

- وزعت نماذج الاختبار على طلاب العينة على أن يجيب جميع الطلاب على النماذج ، بحيث يجيب الطالب الأول على النموذج الأول ، والطالب الثاني وفق النموذج الثاني ، والطالب الثالث وفق النموذج الثالث ، وهكذا لباقي الطلاب ، ثم جمع الباحث نماذج الاختبارات وقد وزع الباحث أعداداً متساوية من النماذج .

يتضح من الجدول رقم (1) أن معاملات الصعوبة لفقرات الاختبار تراوحت بين (0,93-0,12) بوسط مقداره (0,53)، أما معاملات تمييز الفقرات فقد تراوحت بين (0,750,08) بوسط مقداره (0,42) وتم حذف الفقرات الآتية (5, 11, 14, 15, 19, 26, 29, 32) نظراً لتدني معاملات تمييزها فقد كانت (0,08 ، 0,18 ، 0,18 ، 0,09 ، 0,28 ، 0,08 ، 0,17 ، 0,10 ، 0,12 ، 0,65 ، 0,23 ، 0,19 ، 0,93 ، 0,20 ، 0,26 ، 0,14) على التوالي . وقد أصبح عدد فقرات الاختبار بصورته النهائية (30) فقرة .

- قام الباحث ببناء ثلاثة نماذج للاختبار لها نفس المحتوى ؛ النموذج الأول اختبار فقراته مرتبة

النتائج وتفسيرها

السؤال الأول : هل توجد فروق دالة إحصائية في المتوسطات الحسابية لمعاملات صعوبة فقرات اختبار

تحصيلي في التقويم التربوي تعزى لنمط ترتيب فقراته ؟ ولإجابة عن سؤال الدراسة فقد حسبت معاملات صعوبة فقرات الاختبار كما في الجدول رقم (2)

جدول رقم (2)

المتوسطات الحسابية لمعاملات صعوبة الفقرات حسب نمط ترتيب فقرات الاختبار

م	نموذج ترتيب فقرات الاختبار	العدد	المتوسط الحسابي العام لمعامل صعوبة الفقرات	الانحراف المعياري
١	نموذج الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر، استيعاب،)	٣٥	٧٣٧٧.	٠,٣١
٢	نموذج الترتيب حسب تسلسل المحتوى	٣٥	٦٨٥.	٠,٢٨
٣	نموذج الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب)	٣٥	٦٦٨٧.	٠,٢٤

يتضح من الجدول رقم (2) أن قيم معاملات الصعوبة ولجميع الفقرات ، قد جاءت ضمن المدى المقبول ، حيث أشار عودة (2005) أن الفقرات التي تتراوح معاملات صعوبتها بين (0,800, 20 -) تعتبر مقبولة إحصائياً ، وكلما اقتربت معاملات صعوبة فقرات الاختبار من قيمة (0,50) ، وهي قيمة معامل الصعوبة الأمثل بمراعاة أثر التخمين ، الذي يستطيع فيه الاختبار أن يميز بين الطلبة بدرجة عالية ، كلما كان الاختبار أفضل ، ولمعرفة دلالة الفروق بين متوسط معاملات الصعوبة لكل نموذج من نماذج الاختبار باستخدام اختبار (t) كما في الجدول رقم (3)

جدول رقم (3)

دلالات الفروق بين متوسطات صعوبات الفقرات والمتوسط الأمثل لصعوبة الفقرة في نماذج الاختبار

النموذج	متوسط صعوبة الفقرة	متوسط الصعوبة الأمثل	درجة الحرية	قيمة (t)	الدلالة الاحصائية
الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر، استيعاب،)	٠,٧٣٧٧	٠,٥	٢٩	٤,٩٦	٠٠٠
الترتيب حسب تسلسل المحتوى	٠,٦٨٥	٠,٥	٢٩	٣,٧٤٨	٠,٠٠١
الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب)	٠,٦٦٨٧	٠,٥	٢٩	٣,٤٨٠	٠,٠٠١

الترتيب حسب قيم الصعوبة تصاعدياً يكون محسوباً بطريقة دقيقة بما يحقق التدرج في الأسئلة من السهل إلى الصعب ، مما يوفر للطالب الدافعية للاستمرار في محاولة للإجابة عندما يشعر بالتعزير الفوري بسبب قدرته على الإجابة عن الأسئلة الأولى من الاختبار والتي تسمى بالأسئلة التشجيعية أو الأسئلة الماصة للصدمة .

ولمعرفة دلالات الفروق بين المتوسطات الحسابية لمعاملات صعوبة الفقرات بين نماذج ترتيب الفقرات، استخدم اختبار تحليل التباين الأحادي كما في الجدول رقم (4) .

يتضح من الجدول رقم (3) وجود فروق دالة إحصائية بين متوسط صعوبة الفقرة في كل نموذج من نماذج الاختبار الثلاثة ، ومتوسط صعوبة الفقرة الأمثل ، والفروق جوهرية إلا أنه يمكن ترتيب الاختبارات وفق نماذج ترتيب الفقرات فيها حسب قربها من معامل الصعوبة الأمثل - لاختبار اختيار من متعدد أربعة بدائل - على النحو الآتي : اختبار فقراته مرتبة حسب قيم معامل الصعوبة تصاعدياً ، يليه الاختبار الذي فقراته مرتبة حسب تسلسل المحتوى، يليه الاختبار الذي فقراته مرتبة حسب مستوى الهدف الذي تقيسه الفقرة . ربما قد يعود ذلك إلى أن

الجدول رقم (4)

تحليل التباين الأحادي لمعاملات صعوبة الفقرات حسب أنماط ترتيب الفقرات

البيان	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	الدلالة الإحصائية
بين المجموعات	٠,٨٧	٢	٠,٠٣٩	٠,٥٦٢	٠,٥٧٢
داخل المجموعات	٦,٠٢٧	٨٧	٠,٠٦٩		
الكلية	٦,١٠٥	٨٩			

جميع الأنماط تحمل نفس المحتوى فلا يتأثر أداء الطلاب بأنماط ترتيب الفقرات ، حيث إن المستويات العقلية التي تقيسها الفقرات هي واحدة أيضاً بغض النظر عن ترتيب، وكل فقرة تحتاج إلى نفس القدرات العقلية للإجابة عنها بغض النظر عن موقعها في الاختبار.

سؤال الدراسة الثاني : هل توجد فروق ذات دلالة إحصائية في المتوسطات الحسابية لمعاملات تمييز فقرات اختبار محكي المرجع تعزى لنموذج ترتيب فقراته ؟ وللإجابة عن سؤال الدراسة تم حساب المتوسطات الحسابية والانحرافات المعيارية لمعاملات صعوبة الفقرات لكل نمط من أنماط ترتيب الفقرات، كما في الجدول رقم (5)

يتضح من الجدول رقم (4) عدم وجود فروق ذات دلالة إحصائية عند مستوى ($0,05 \leq \alpha$) بين المتوسطات الحسابية لتقدير معاملات صعوبة فقرات الاختبار تعزى لنموذج ترتيب فقراته ، أي أنه لا يوجد أثر لنموذج ترتيب فقرات الاختبار في تقديرات المتوسطات الحسابية لمعاملات صعوبة فقرات الاختبار لكل نموذج من نماذج ترتيبها ، حيث إن الفروق موجودة بين متوسطات معاملات صعوبة الفقرات حسب أنماط ترتيب الفقرات لكن الفروق بسيطة وليست جوهرية ، تتفق نتائج البحث الحالي مع دراسة رمضان (2010) حيث أشارت إلى أنه لا يوجد أثر لنموذج ترتيب الفقرات على أداء الطلبة على الاختبار . ربما يعود ذلك إلى أن الفقرات في

جدول رقم (5)

المتوسطات الحسابية لمعاملات تمييز الفقرات حسب نمط ترتيب فقرات الاختبار

م	نموذج ترتيب الفقرات في الاختبار	العدد	المتوسط الحسابي العام لمعامل تمييز الفقرات	الانحراف المعياري
١	نموذج الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر، استيعاب،)	٢٠	٠,٤٥	٠,٣٠٨
٢	نموذج الترتيب حسب تسلسل المحتوى	٢٠	٠,٢٧	٠,٢٣٤
٣	نموذج الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب)	٢٠	٠,٥٧	٠,٣٢٢

التمييز بين الطلاب وإفراز الطلاب إلى عدة مستويات تحصيلية . كما يظهر وجود فروق بين المتوسطات الحسابية لمعاملات تمييز فقرات الاختبارات حسب أنماط ترتيب الفقرات ، ربما ارتقاء معاملات التمييز للاختبار الذي فقراته مرتبة تصاعدياً محسوبة بطريقة إحصائية حسب قيم معامل الصعوبة يعود إلى عدم وجود فقرات ذات معاملات صعوبة متطرفة في صعوبتها وسهولتها ، مما يجعل عدداً كبيراً من الفقرات ذات معاملات تمييز مرتفعة وموجبة . ولاختبار دلالة الفروق بين المتوسطات الحسابية لمعاملات تمييز الفقرات ، استخدم اختبار تحليل التباين الأحادي كما في الجدول رقم (6)

يتضح من الجدول رقم (5) أن قيم معاملات التمييز المقبولة في اختبار تحصيلي متوفرة في نماذجين من ترتيب الفقرات ، الاختبار الذي فقراته مرتبة حسب قيم معامل الصعوبة من السهل إلى الصعب (حسب الطريقة الإحصائية) ، حيث المتوسط الحسابي لمعاملات التمييز للفقرات (0,57) ، ثم يليه اختبار فقراته مرتبة حسب تسلسل المحتوى حيث كان متوسط معاملات تمييز فقرات الاختبار (0,45) ، حيث أشار عودة (2005) أن معاملات تمييز الفقرات $\leq 0,39$ تعتبر ضمن المدى المقبول والفعل في التمييز بين الطلاب ، حيث إن أحد الأهداف الرئيسة للاختبارات التحصيلية هو

جدول رقم (6)

تحليل التباين الأحادي لمعاملات تمييز الفقرات حسب أنماط ترتيب الفقرات

البيان	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	الدلالة الإحصائية
بين المجموعات	٠,٦٢	٢	٠,٣١	٢,٥٧	٠,٠٢٢
داخل المجموعات	٧,٥٥	٨٧	٠,٠٨٧		
الكلية	٨,١٧	٨٩			

لمعاملات تمييز فقرات الاختبار لكل نمط من أنماط ترتيبها ، ولتحديد أي المتوسطات ذات الفروق الدالة إحصائياً أُجري اختبار توكي للمقارنات البعدية كما في الجدول رقم (7) .

يتضح من الجدول رقم (6) وجود فروق ذات دلالة إحصائية عند مستوى ($\alpha \leq 0,05$) بين تقديرات معاملات تمييز فقرات الاختبار تعزى لنماذج ترتيب فقراته ، أي أنه يوجد أثر لنمط ترتيب فقرات الاختبار في تقديرات المتوسطات الحسابية

جدول رقم (7)

اختبار شافيه للمقارنات البعدية للمتوسطات الحسابية لمعاملات تمييز فقرات حسب نمط الترتيب

نمط ترتيب الفقرات في الاختبار	نمط الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر ، استيعاب)	نمط الترتيب حسب تسلسل المحتوى	نمط الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب)
المتوسط الحسابي	٠,٤٥	٠,٣٧	٠,٥٧
نمط الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر ، استيعاب)	-	٠,٠٨	٠,١٢
نمط الترتيب حسب تسلسل المحتوى	٠,٣٧	-	× ٠,٢٠
نمط الترتيب حسب قيم معامل الصعوبة للفقرات (من السهل إلى الصعب)	٠,٥٧	-	-

× القيمة دالة عند مستوى ($\alpha \leq 0,05$)

الصعب، حيث بلغت قيمته 0,20 ، والقيمة دالة إحصائياً عند مستوى $\alpha \leq 0,05$ حيث نمط الاختبار الأخير يميز بين الطلبة بدرجة عالية .

سؤال الدراسة الثالث: هل توجد فروق ذات

دلالة إحصائية في معاملات ثبات اختبار تحصيلي في مقرر التقويم التربوي تعزى لنموذج ترتيب فقراته؟
ويجب عن سؤال الدراسة الجدول رقم (8)

يتضح من الجدول رقم (7) وجود فرق دال إحصائياً بين المتوسط الحسابي لمعاملات تمييز نموذج اختبار مرتبة فقراته حسب تسلسل المحتوى والمتوسط الحسابي لمعاملات تمييز نموذج اختبار مرتبة فقراته حسب قيم معاملات الصعوبة من السهل إلى الصعب لصالح الاختبار الذي فقراته مرتبة حسب قيم معاملات الصعوبة من السهل إلى

جدول رقم (8)

معاملات ثبات الاختبار التحصيلي بطريقة الفا كرونباخ لنماذج الاختبار الثلاثة

م	نموذج ترتيب الفقرات في الاختبار.	معامل الثبات بطريقة الفا كرونباخ
١	نموذج الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر ، استيعاب)	٠,٧١٨٧
٢	نموذج الترتيب حسب تسلسل المحتوى.	٠,٧٢٨٦
٣	نموذج الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب) .	٠,٧٧٤٢

فقراته مرتبة حسب قيم معاملات الصعوبة من السهل إلى الصعب وقد جاء في المرتبة الأولى ، وقد كان المتوسط الحسابي لمعامل الثبات (0,7742)

يتضح من الجدول رقم (8) أنه يمكن ترتيب نماذج الاختبار التحصيلي من نوع اختيار من متعدد حسب قيم معاملات الثبات على النحو الآتي : اختبار

كما يساعد ترتيب الفقرات على التهيؤ العقلي للطلاب مما يجعل الاختبار يقيس قدرة الطالب على التحصيل بدرجة مقبولة من الدقة أو بأقل خطأ، بمعنى أن يكون أكثر ثباتا .

ولاختبار دلالة الفروق بين قيم معاملات الثبات لنماذج الاختبارات الثلاثة ، استخدم الباحث الإحصائي (M) المقترح من قبل هاكستين ولين (Hakstain&Whalen, 1979) والجدول رقم (9) يظهر ذلك :

جدول رقم (9)

دلالة الفروق باستخدام الإحصائي (M) بين معاملات ثبات لثلاثة نماذج من اختبار تحصيلي

نموذج الاختبار	نموذج الترتيب حسب مستوى الهدف الذي تقيسه	نموذج الترتيب حسب تسلسل المحتوى	نموذج الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب)	درجات الحرية	قيمة (M)	مستوى الدلالة
معامل الثبات بطريقة الفا كرونباخ	٠,٧١٨٧	٠,٧٢٨٦	٠,٧٧٤٢	٢	١,٥٤	٠,٤٦٦

عن سؤال الدراسة الرابع ، حسب معاملات الصدق التنبؤي التي هي معامل ارتباط بيرسون بين درجات الطلاب على كل نموذج من نماذج الاختبارات الثلاثة ودرجاتهم على الاختبار النهائي في نفس المقرر ، كما حسب الصدق التلازمي لكل نمط من أنماط ترتيب فقرات الاختبار الذي هو معامل ارتباط بيرسون بين درجات الطلاب على كل نموذج من نماذج الاختبارات الثلاثة ودرجاتهم على اختبار منتصف الفصل الدراسي في نفس المقرر ، والنتائج كما في الجدول رقم (10)

جدول رقم (10)

معاملات ثبات الاختبار التحصيلي بطريقة الفا كرونباخ لنماذج الاختبار الثلاثة

م	نموذج ترتيب فقرات الاختبار	معامل الصدق التلازمي	معامل الصدق التنبؤي
١	نموذج الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر ، استيعاب ،.....)	٠,٦٠٥	٠,٥٠٤
٢	نموذج الترتيب حسب تسلسل المحتوى	٠,٦٢٢	٠,٥١٢
٣	نموذج الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب)	٠,٥٥٤	٠,٤١٧

ثم يليه الاختبار الذي فقراته مرتبة حسب تسلسل المحتوى ، وكان معامل الثبات (0 ,7286) ، يليه الاختبار الذي فقراته مرتبة حسب مستوى الهدف الذي تقيسه الفقرة ، وكان معامل ثباته (0 ,7187) ربما يعود ارتفاع معامل الثبات للاختبار الذي فقراته مرتبة تصاعديا حسب قيم الصعوبة إلى زيادة دافعية الطالب للإجابة لإعطاء أفضل إجابة ، مما يؤدي إلى زيادة قدرته على مواصلة الإجابة حتى نهاية فترة الامتحان ، وتساؤه على زيادة قدرته على الضبط النفسي عند مواجهة أي مشكلة في الامتحان ،

يتضح من الجدول رقم (9) عدم وجود فروق ذات دلالة إحصائية بين معاملات الثبات بطريقة ألفا كرونباخ لنماذج الاختبار الثلاثة تعزى لنمط ترتيب الفقرات ، بمعنى أن اختلاف نموذج ترتيب فقرات الاختبار لا يؤثر على ثباته بشكل جوهري .

سؤال الدراسة الرابع : هل توجد فروق ذات دلالة إحصائية في معاملات الصدق المرتبط بمحك (التلازمي و التنبؤي) لاختبار تحصيلي في مقرر التقويم التربوي تعزى لنمط ترتيب فقراته ؟ وللإجابة

1

بحوث ودراسات

تسلسل المحتوى وكان متوسط معامل صدقه التنبؤي (0,512) ، يليه نموذج اختبار فقراته مرتبة حسب مستوى الهدف الذي تقيسه الفقرة وفق تصنيف بلوم ، وكان معامل صدقه التنبؤي (0,504) ، يليه اختبار فقراته مرتبة حسب قيم معامل الصعوبة من السهل إلى الصعب وكان متوسط معامل صدقه التنبؤي (0,417) .

وللكشف عن دلالة الفروق بين معاملات الصدق التلازمي لنماذج الاختبار الثلاثة ، ودلالة الفروق بين معاملات الصدق التنبؤي لنماذج الاختبار الثلاثة استخدم الاختبار الاحصائي (V) الذي وضعه هيز (Hays, 1981) ويظهر كما في الجدول رقم (11)

جدول رقم (11)

دلالة الفروق باستخدام الإحصائي (V) بين معاملات الصدق المرتبط بمحك (التلازمي - التنبؤي) لثلاثة نماذج من اختبار تحصيلي

نموذج الاختبار	نموذج الترتيب حسب مستوى الهدف الذي تقيسه الفقرة (تذكر ، استيعاب)	نموذج الترتيب حسب تسلسل المحتوى	نموذج الترتيب حسب قيم معامل الصعوبة (من السهل إلى الصعب)	درجات الحرية	قيمة (V)	مستوى الدلالة
معامل الصدق التلازمي	٠,٦٠٥	٠,٦٢٢	٠,٥٥٤	٢	١,٠٨	٠,٥٨٦
معامل الصدق التنبؤي	٠,٥٠٤	٠,٥٠٤	٠,٤١٧	٢	٠,٩٨٣	٠,٦١١

المحتوى لذلك جاء معامل صدق نموذج الاختبار فقرات مرتبة حسب تسلسل المحتوى ويمكن ترتيب فقرات بعض الاختبار حسب مستوى الهدف الذي تقيسه ولكن نادرا ما يستخدم ترتيب الفقرات حسب قيم صعوبتها .

يتضح من الجدول رقم (11) عدم وجود فروق ذات دلالة إحصائية في معاملات الصدق المرتبط بمحك (التلازمي - التنبؤي) بين نماذج الاختبار الثلاثة ، بمعنى أنه لا يوجد أثر جوهري لنموذج ترتيب فقرات الاختبار على الصدق المرتبط بمحك (التلازمي - التنبؤي) للاختبار التحصيلي . ويمكن تفسير ذلك بأن معظم الاختبارات التحصيلية التي تجرى في الجامعات تعتمد في ترتيبها على تسلسل

بناء على نتائج البحث يمكن الخروج بالتوصيات الآتية :

- استخدام نموذج ترتيب الفقرات تصاعدياً حسب قيم معامل الصعوبة عند بناء الاختبارات التحصيلية متعددة البدائل لارتفاع متوسطات معاملات تمييز فقراته ، و ثباته وصدقه . إذا تعذر ذلك ينبغي استخدام ترتيب الفقرات حسب المحتوى مع الأخذ بعين الاعتبار مستوى الهدف الذي تقيسه الفقرة على افتراض أن الفقرة التي تقيس مستوى هدف سلوكي أعلى حسب تصنيف بلوم يفترض أن تكون أكثر صعوبة .
- يمكن أن يعتبر عضو هيئة التدريس تطبيقه للامتحان أول مرة كتجريب أولي ثم يستفيد من تحليل إجابات الطلبة في توفير معلومات إحصائية عن كل فقرة .
- لا بد من تدريج أسئلة الاختبارات من الأسهل إلى الأكثر صعوبة وتدريب أعضاء هيئة التدريس في الجامعة على كيفية صياغة أسئلة الاختبارات، بحيث تراعي ترتيب الفقرات .
- إجراء دراسات مماثلة تتعلق بمقررات ونماذج أخرى من ترتيب الفقرات .

- الحريري، رافدة عمر (2007) . التقويم التربوي الشامل للمؤسسة المدرسية، دار الفكر: عمان .
- سنان ، إيناس محمد (2006) . درجة معرفة ممارسة معلمات الرياضيات في المرحلة المتوسطة لمهارات بناء الاختبار بمدينة مكة المكرمة ، رسالة ماجستير غير منشورة بكلية التربية ، جامعة أم القرى .
- سليمان ، شاهر خالد (2008) أثر مخالفة قواعد صياغة فقرات اختبار اختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته ، المؤتمر السنوي لقسم علم النفس ، تطبيقات العلوم النفسية ومشكلات المجتمع في الفترة من 4-6 ، جامعة طنطا ، جمهورية مصر العربية، ص315
- سليمان ، شاهر خالد (2010) ممارسات أعضاء هيئة التدريس في جامعة تبوك في تقييم تحصيل طلابهم ، مجلة جامعة أم القرى للعلوم التربوية والنفسية بمكة المكرمة، 2(2) ، ص 50-80
- الشريفين ، نضال (2006) الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق النظرية الحديثة في القياس التربوي والنفسى ، مجلة العلوم التربوية والنفسية ، المجلد (7) ، العدد (4) ص80 .
- الراداي ، عبد العالي (2006) أثر اختلاف نمط الاختبار على الخصائص السيكومترية للاختبار وفقراته ، رسالة ماجستير غير منشورة بكلية التربية ، جامعة أم القرى بمكة المكرمة .
- رمضان ، رمضان (2010) أثر ترتيب فقرات الاختبار حسب درجة صعوبتها على تحصيل

- (1979). Introduction to Measurement Theory. Wadsworth inc.
- Anastasi .A(1988). Psychological Testing . New York :Macmillan. co
 - Carstens .w, MCkeaq .R .A .(1982). The effect of change item sequence order on performances in a test re-test experiments. ERIK DOCUMENT ED241548.
 - Crocker, L. & Allgina, J. (1986). Introduction to Classical and Modern Test Theory. California : Cole Publishing Company.
 - Dahlstrom, W. G., Brooks, J. D., & Peterson, C. D. (1990). The Beck depression inventory: Item order and the impact of response sets. Journal of Personality Assessment, 55, 224-233.
 - Derek Hodson. (1984) The effect of changes in item sequence on student performance in a multiple-choice chemistry test. Journal of Research in Science Teaching 21:5, 489
 - Diannal_Newman, Deborah .k. kudert .(1988) Effect of varying item order on multiple- choice test scores: importance of statistical and cognitive difficulty. Applied measurement in education vol 1 , pp 89-97
 - Hambelton, R.K and Swaminathan ,H.(1985). Item response theory :principals And applications Kluwer, Nijhoff publishing. Boston
 - Hambelton ,R.K.and Traub, R.R .(1971).information curves and efficiency of three logistic models. British journal of mathematical and statistical psychology.24,273-281
 - Jodi .B, Claudio. N, Ilse De. B, Christine. M, Jay .M.(2007). The Effect of Item Order on Physical Activity Estimates Using the IPAQ. Californian
 - طلاب المرحلة الثانوية لمادة الرياضيات ، مجلة الأبحاث التربوية ، كلية التربية ، جامعة الأزهر، العدد(8) .ص45.
 - العريفي ، هدى محمد (1427) تباين الخصائص السيكمترية لاختبار اختيار من متعدد بتباين عدد مفرداته وبعض الخصائص الدافعية للمفحوصين ، رسالة ماجستير غير منشورة بكلية التربية ، جامعة أم القرى.
 - العمري ، حسان (2011) أثر ترتيب الفقرات على المقاييس النفسية على التجانس الداخلي للمقياس وخصائص فقراته ، مجلة جامعة دمشق مجلد 27 (3،4) . ص111
 - عودة ، أحمد (2005) القياس والتقويم في العملية التدريسية ، الطبعة الثالثة : دار الأمل للنشر والتوزيع ، اربد ، الأردن .
 - علام ، صلاح الدين (2007) القياس والتقويم التربوي ، الطبعة الأولى ، دار المسيرة، عمان .
 - الليحاني ، عدنان (2010) أثر ترتيب فقرات الاختبار على صدق وثبات اختبار اختيار من متعدد البدائل في مادة رياضيات لطلاب المرحلة الثانوية بمكة المكرمة ، رسالة ماجستير غير منشورة ، جامعة أم القرى بمكة المكرمة.
 - Aiken ,L.R(1988).psychological testing Assessment Boston: allynabacon , inc.
 - Anderson, D.C &Brokewski ,J,G.(1978)Experimental Psychology Glenview ,Illinois :Scott, foresmen and company
 - Al banes, M ,A, and Forsyth, A.(1984). The one-two and modified two parameter latent trait models. An empirical study of relative fit. Education and psychological measurement .44,229-246
 - Allen, Mary J. and yen, Wendy M.

Journal of Health Promotion V 5, Issue 1, 23-29

- Maloney. P.M &Ward, P.M.(1980). Psychological Assessment Conceptual Approach . New York Oxford University Press.
- Neely, D. L., Springston, F. J., & McCann, S. J. (1994). Does item order affect performance on multiple-choice exams? Teaching of Psychology, 21, 44-45.
- Pedhasur.E.J.(1982).Multiple regression in behavioral research. New York: Holt .Rinehart and Winston.
- Reckase ,M.D.(1979). Unifactor latent trait models applied to multifactor test. Journal of educational statistics.4.207-230
- Schuyler W. Huck, Norman D. Bowers.(1972).Item difficulty level and sequence effects in Multiple choice Achievement Tests. Journal of Educational Measurement, Vol. 9, No. 2 (Summer, 1972), pp. 105-111
- Stanley C.J, Hopkins.K.D.(1972). Educational and Psychological Measurement Evaluation, New Jersey Prentice-hall.