

**تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة
السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار
التحصيلي الجيد من وجهة نظر الطلبة وبعض المتغيرات
الأخرى**

أحمد يعقوب النور*

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي الجيد من وجهة نظر الطلبة وبعض المتغيرات الأخرى

والتقويم التربوي والنفسي، وتدريب أعضاء هيئة التدريس يمثل الصدارة بين مشروعات التطوير التربوي في مؤسسات التعليم العالي في معظم دول العالم، كما أولت المجتمعات الدولية على اختلاف نظمها وفلسفاتها وأهدافها أهمية كبيرة للارتقاء بالأطر التربوية والنفسية ومن بينها القياس والتقويم النفسي والتربوي بهدف بناء أدوات قياسية تربوية لها خصائص سيكومترية جيدة تقدم نتائج ومعلومات تفيد في صنع القرارات التربوية السليمة.

وتشير القرشي [2] إلى أن التقويم التربوي يعد محورياً أساسياً في عملية تجويد التعليم الجامعي، ذلك باعتبار أن التقويم أصبح شرطاً أساسياً تشترطه جميع هيئات الاعتماد الأكاديمي العالمية، حيث أن هيئة (ABET) جعلت في عام 2000م تقويم المخرجات التعليمية المحور الرئيس لعملية الاعتماد الأكاديمي.

كما أن الغرض الأساسي من التقويم هو تصحيح مسار العملية التعليمية وتلافي الأخطاء وليس مجرد إصدار أحكام على الطلبة وعلى مدى تحصيلهم، إذ أن الهدف الأساسي من التقويم هو تطوير العملية التعليمية وتحسينها [3].

وتعد الاختبارات التحصيلية من أكثر أدوات التقويم استخداماً في المؤسسات التربوية التعليمية في العالم، حيث تستخدم من أجل الوقوف على القدرات التحصيلية للطلبة، ونستدل من خلالها على قدراتهم المعرفية والمهارية [4].

يشير الكبيسي [5] إلى أن الاختبارات تعد من الوسائل المهمة المستخدمة في قياس وتقويم قدرات الطلبة باختلاف مستوياتهم الدراسية، حيث تساعد القائمين على أمر التربية والتعليم من كشف عناصر النجاح وتشخيص نقاط القوة

الملخص_ هدفت الدراسة إلى تقويم مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي الجيد من وجهة نظر الطلبة ومتغيرات الجنس والمستوى والتخصص الدراسي والمعدل التراكمي، وتكونت عينة الدراسة من (300) طالبا وطالبة في كلية التربية في الفصل الدراسي الأول من العام 2011م/2012م، تم تطبيق مقياس من إعداد الباحث، وأظهرت النتائج أن مستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد كان متوسطاً من وجهة نظر الطلبة، كما أظهرت النتائج بأنه توجد فروق ذات دلالة إحصائية في وجهة نظر الطلبة تعزى للجنس ولصالح الطالبات الإناث، وأظهرت النتائج عدم وجود فروق تعزى لمتغير المستوى الدراسي، ووجود فروق تعزى لمتغير التخصص الدراسي ولصالح التخصص الأدبي، كما توجد فروق تعزى لمتغير المعدل التراكمي ولصالح المعدل المرتفع.

الكلمات المفتاحية: التقويم، الاختبار التحصيلي، معايير الاختبار الجيد، وجهة نظر.

1. المقدمة

تعد الجامعات معقلاً للفكر الإنساني في أرفع مستوياته، وأن أهم ما يهدف إليه التعليم الجامعي هو خدمة المجتمع والارتقاء به، وتطوير التفكير وتقديم العلم، وتنمية القيم الإنسانية، وتزويد الدول بالمتخصصين والفنيين والخبراء في مجالات شتى [1].

وللتحقق من امتلاك الطلبة للكفايات التي تقدم إليهم بالجامعات لزم استخدام أدوات قياس وتقويم مناسبة تتوافر فيها الخصائص السيكومترية لكي تكون نتائجها موثوقة تمكن القائمين على أمر الكليات والجامعات من اتخاذ القرارات الصائبة. ومن هنا فقد أصبح موضوع تطوير أدوات وأساليب القياس

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي

4- عدم استخدام وسائل التقنية الحديثة في تصحيح وتحليل وتفسير نتائج الاختبارات.

5- الاهتمام بالمستويات الدنيا للأهداف التربوية مثل التذكر والاسترجاع على حساب المستويات العليا مثل التطبيق والتحليل والتركيب والتقييم، والتفكير الاستنباطي والإبداعي والتفكير بأسلوب حل المشكلات.

2. مشكلة الدراسة

إن وجود التقويم التربوي مسألة جوهرية يمكن من خلالها الحكم على مدى النجاح أو الفشل في تحقيق الأهداف التي ترسمها المؤسسات التربوية للعملية التعليمية، وتعد أسئلة الاختبارات التحصيلية جانباً أساسياً في العملية التقييمية، وأن هذه الأسئلة تقيس ما تم تحقيقه من أهداف، فالأسئلة ينبغي أن تغطي جميع الأهداف وتكون شاملة للمحتوى الدراسي المعنى، كما لا بد أن تتوفر في هذه الأسئلة الموضوعية والصدق والثبات والشمول، والقدرة على التمييز وإظهار الفروق الفردية بين الطلبة، كما لا بد لها من أن تنتزع في قياس الجوانب المعرفية والمهارية من التذكر والفهم والتطبيق والتحليل والتركيب والتقييم.

وبمراجعة البحوث والدراسات الجامعية تبين للباحث ندرة الدراسات التي تتناول تقويم مستوى تطبيق أعضاء هيئة التدريس في تخصص اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد من وجهة نظر الطلبة وبعض المتغيرات الأخرى.

ومن هنا يرى الباحث أن الحاجة ملحة لبحث ومعرفة مدى جودة معايير تصميم الاختبارات التحصيلية لمقررات (اللغة العربية، 1، 2، 3) وهي تمثل مقررات اللغة العربية التي تُدرس لطلبة كلية التربية بجامعة السودان للعلوم والتكنولوجيا في السنة الأولى (الفصل الدراسي الأول والثاني)، والسنة الثانية (الفصل الدراسي الأول) ضمن متطلبات الجامعة.

أضف إلى ذلك ما دفع الباحث إلى هذه الدراسة شكاوى طلبة الجامعة من مقررات اللغة العربية ونظرتهم المتمثلة في عدم جدوى هذه المقررات لأنها تُدرس بشكل عام لجميع طلبة الجامعة وفي مختلف الكليات دون مراعاة لخصوصية كل كلية

والضعف في عملية التعليم والتعلم مما يؤدي إلى الارتقاء بالعملية التربوية.

ويؤكد أبوعلام [6] أن سوء تنظيم الورقة الاختبارية يؤثر على أداء الطلبة وكذلك يؤثر على صدق وثبات الاختبار.

وتشير دراسة القحطاني [7] إلى أن أحكام وجودة تصميم الورقة الاختبارية بشمولها لمعايير وخصائص الاختبار الجيد تؤدي إلى نتائج تعد مدخلاً جيداً للتشخيص.

إن المعلم المعاصر يستطيع أن يصمم الاختبارات، ويضع الأسئلة المناسبة لقياس المستويات المختلفة، ويحدد الاختبارات المناسبة لقياس الجوانب المختلفة للتعلم، ويتطلب كل ذلك إلمامه بماهية الاختبارات وشروط تصميمها ونتائجها [8].

ويشير السيد [9] إلى أن بعض الدراسات أشارت إلى نقاط ضعف عديدة بالاختبارات التحصيلية في بنائها وتصميمها، ومن بين هذه النقاط:

1- ضعف قدرات أعضاء هيئة التدريس على صياغة الفقرات التي يمكن أن توصف بالموضوعية.

2- التركيز على أسئلة المقال.

3- عدم شمولية الأسئلة لمحتوى المقرر.

4- عدم اهتمام أعضاء هيئة التدريس بوضع جداول مواصفات عند تصميم وبناء اختباراتهم.

5- الاهتمام بقياس الحفظ والتذكر بدلاً من قياس المهارات العقلية العليا.

6- عدم استخدام الضبط الإحصائي للتعرف على صدق وثبات الاختبارات.

كما يؤكد خضر [10] بأن للاختبارات التحصيلية في البلاد العربية عدداً من السلبيات وهي:

1- افتقار الكثير من المعلمين إلى الأصول العلمية في تصميم وتحليل وتفسير نتائج الاختبارات.

2- عدم استخدام أساليب جديدة من الاختبارات.

3- عدم استثمار نتائج الاختبارات التحصيلية في تطوير المناهج وأساليب التدريس والأهداف التعليمية.

أهم أدوات القياس لجمع المعلومات اللازمة لعلمية التقويم التربوي، كما تتمثل أهميته في الإجابة عن أسئلته، بالإضافة إلى الجوانب الآتية:

1- تسليط الضوء على أهمية تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد عند بناء وتصميم اختباراتهم.

2- تزود نتائج البحث مصممي الاختبارات التحصيلية في اللغة العربية ببيانات عن معايير تصميم الاختبار التحصيلي الجيد.

3- يعد هذا البحث من الدراسات القليلة في العالم العربي والنادرة في السودان، في حدود علم الباحث، والتي تتناول موضوع تقويم مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد.

ج. أهداف الدراسة

تهدف الدراسة الحالية إلى الآتي:

1- الكشف عن وجهة نظر طلبة كلية التربية بجامعة السودان للعلوم والتكنولوجيا حول تقييمهم لمستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد.

2- التعرف على مستوى الفروق في إجابات الطلبة عند تقييمهم لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد وفقاً لمتغيرات الجنس، والمستوى الدراسي، والتخصص الدراسي والمعدل التراكمي.

3- تقديم واقتراح توصيات تسهم في تطوير الاختبارات التحصيلية في اللغة العربية من خلال تطبيق معايير تصميم الاختبار التحصيلي الجيد.

د. مصطلحات الدراسة

الاختبار Test

يعد الاختبار من أبرز مفاهيم القياس، وهو إجراء منظم لقياس عينة من السلوكيات من خلال عينة من المثيرات [27].

الاختبار أداة أو أسلوب لقياس عينة من السلوك، أي أنه أداة تقويم أداء الفرد في مجال معين، وتتكون هذه الأداة من عدد من

والمهن والأعمال المستقبلية المرتبطة بها، وبالتالي يرون عدم إمكانية حدوث انتقال أثر تعلمهم للغة العربية والاستفادة منها في حياتهم العملية.

ولهذا كله أصبحت الحاجة ملحة لبحث وتقويم جودة معايير تصميم الاختبارات التحصيلية من وجهة نظر الطلبة في جامعة السودان للعلوم والتكنولوجيا، وتقديم التوصيات الكفيلة بتطوير هذه الاختبارات في ضوء المواصفات القياسية والمعيارية الواجب توافرها عند تصميم وبناء الاختبارات التحصيلية، لكي يمكن الاعتماد على نتائجها في اتخاذ القرارات التربوية المناسبة.

أ. أسئلة الدراسة

وتحاول الدراسة الحالية الإجابة عن الأسئلة الآتية:

1- ما تقييم الطلبة لمستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لكل بُعد من أبعاد معايير تصميم الاختبار التحصيلي الجيد في اختبارات اللغة العربية بكلية التربية بجامعة السودان للعلوم والتكنولوجيا؟

2- هل توجد فروق ذات دلالة إحصائية في تقييم الطلبة لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد في اختبارات اللغة العربية بكلية التربية تعزى لمتغير الجنس؟

3- هل توجد فروق ذات دلالة إحصائية بين متوسطات إجابات الطلبة على مستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد في اختبارات اللغة العربية بكلية التربية تعزى لمتغير المستوى الدراسي؟

4- هل توجد فروق ذات دلالة إحصائية بين متوسطات إجابات طلبة كلية التربية بجامعة السودان للعلوم والتكنولوجيا لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد في اختبارات اللغة العربية تعزى للتخصص الدراسي؟

5- هل توجد فروق بين متوسطات إجابات أفراد عينة البحث تعزى للمعدل التراكمي (ضعيف، متوسط، مرتفع)؟

ب. أهمية الدراسة

تتمثل أهمية الدراسة في أهمية الاختبارات التحصيلية باعتبارها

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي

3- أداة الدراسة: هي مقياس تم إعداده لدراسة وجهة نظر طلبة كلية التربية - جامعة السودان للعلوم والتكنولوجيا حول مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد، وتتناسب مع البيئة السودانية وقد بلغ عدد فقراتها (36) فقرة، موزعة علي (6) أبعاد وهي:

1- المعايير المطلوبة لكتابة تعليمات الاختبار التحصيلي الجيد وفقراتها: 1، 2، 3، 4، 5، 6، 7.

2- المعايير المطلوبة لكتابة فقرات الاختبار التحصيلي الجيد وفقراتها هي: 8، 9، 10، 11، 12، 13، 14، 15، 16.

3- معايير موضوعية الاختبار وصلاحيته وفقراتها هي: 17، 18، 19، 20، 21، 22، 23، 24.

4- معايير شمول أسئلة الاختبار للأهداف وللمحتوى الدراسي وفقراتها هي: 25، 26، 27، 28.

5- معايير القدرة التمييزية لأسئلة الاختبار وفقراتها هي: 29، 30، 31، 32.

6- معايير طباعة الاختبار التحصيلي وفقراتها هي، 33، 34، 35، 36.

3. الإطار النظري والدراسات السابقة

إن الهدف الأساسي للعملية التربوية هو إحداث تغييرات مرغوبة في سلوك المتعلمين، ولذلك فإن الدور الرئيسي الذي يجب أن يقوم به التقويم هو تحديد ماهية التغيرات الحادثة في سلوك الطلبة. إن التقويم كعنصر من عناصر العملية التربوية بدأ منذ القدم، ولكن المفهوم الحديث للتقويم بدأ في التركيز على تطوير أساليب التقويم الحديث لمتابعة نمو الطلبة وتقديمهم، حيث أصبحت أساليب التقويم الحديثة تركز على قياس السلوك بأسلوب شمولي في جوانب النمو المتعددة.

إن الاختبارات التحصيلية هي أكثر أنواع التقويم شيوعاً في المدارس والجامعات من أجل الحصول على معلومات عن قدرة الطالب ونموه ومقارنة مستواه بزملائه.

إن إعداد أسئلة الاختبار الجيد يعتمد إلى حد كبير على الخبرة الطويلة والتجربة والممارسة والتدريب، وعلى القدرة اللغوية

البنود أو الفقرات Items أو الأسئلة Questions يجب عنها [20].

الاختبار عبارة عن مجموعة من المثيرات وضعت لتقيس بعض المعطيات العقلية أو الخصائص النفسية بطريقة كمية أو كيفية، حيث تؤدي هذه المثيرات إلى إحداث استجابات يُمنح الفرد على أساسها درجة معينة [28].

ويقصد بالاختبارات في الدراسة الحالية الاختبارات التحصيلية التي أعدها أعضاء هيئة التدريس في اللغة العربية بجامعة السودان للعلوم والتكنولوجيا لقياس التحصيل الأكاديمي لطلبة كلية التربية في الفصل الدراسي الأول من العام الجامعي 2012/2011م.

التقويم EVALUATION

يقصد به إصدار حكم على قيمة الأعمال أو الأفكار أو السلوكيات في ضوء معايير لتقدير مدى كفاية الأشياء وفعاليتها [29].

ويقصد بالتقويم في البحث الحالي بأنه عملية تحليل إجابات الطلبة وتحديد وجهة نظرهم حول مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي الجيد في اختبارات كلية التربية في الفصل الدراسي الأول من العام الجامعي 2012/2011م، بغرض توضيح جوانب القوة والضعف في هذه الاختبارات في ضوء معايير الاختبار التحصيلي الجيد.

هـ. حدود الدراسة

تتحد نتائج هذه الدراسة بالآتي:

1- اقتصر البحث الحالي على دراسة وجهة نظر طلبة كلية التربية بجامعة السودان للعلوم والتكنولوجيا وتقييمهما مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد.

2- تم إجراء البحث خلال الفصل الدراسي الأول للعام الجامعي 2012م/2011م، وبذلك تكون نتائج البحث محددة زمنياً بالفترة التي أجريت فيها.

- والخيال الخصب، والقدرة العلمية المتعمقة في مادة التخصص
 وأساليب تدريسها، إضافة إلى المعرفة التامة بالطلبة الذين يطبق
 عليهم الاختبار [10].
- ويشير النور [4] إلى أن هناك مبادئ أساسية للاختبارات
 التحصيلية تتمثل في الآتي:
- 1- يجب أن تقيس الاختبارات التحصيلية نواتج التعلم مثل
 معرفة الحقائق والمفاهيم والمصطلحات، وتطبيق الحقائق
 والمبادئ، وتحليل الكليات إضافة إلى عمليات التركيب والتقويم
 والجوانب المهارية، لذلك فإن الخطوة الأولى في بناء وتصميم
 الاختبار التحصيلي تتمثل في تحديد نواتج أو مخرجات التعلم
 التي يراد قياسها.
 - 2- يجب أن يقيس الاختبار التحصيلي عينة ممثلة للمقرر
 الدراسي.
 - 3- يجب تنوع الأسئلة لقياس الجوانب السلوكية المختلفة.
 - 4- أن تتوفر الخصائص السيكومترية في الاختبارات
 التحصيلية.
- خطوات إعداد وتصميم الاختبار التحصيلي: من أهم خطوات
 إعداد وتصميم الاختبار التحصيلي الآتي:
- 1- تحديد الأهداف التعليمية (السلوكية).
 - 2- صياغة الأهداف التعليمية بصورة إجرائية قابلة للملاحظة
 والقياس.
 - 3- تحديد محتوى المادة الدراسية موضوع الاختبار.
 - 4- إعداد جدول المواصفات وهو عبارة عن مخطط تفصيلي
 يبين نسبة الأهداف وعدد الأسئلة المخصصة لكل جزء منها.
 - 5- تحديد نوعية الفقرات في الاختبار.
 - 6- كتابة فقرات الاختبار.
 - 7- مراجعة الفقرات وتنقيح صياغتها.
 - 8- ترتيب الفقرات في الاختبار.
 - 9- إعداد تعليمات الاختبار، حيث يساعد وجود تعليمات
 واضحة للاختبار على رفع معامل الصدق والموضوعية والثبات.
 وهذه التعليمات يجب أن تكون بسيطة وواضحة وتتضمن
- المعلومات الآتية:
- هدف الاختبار.
 - الوقت المحدد للاختبار.
 - تاريخ الاختبار.
 - كيفية كتابة الأجوبة.
 - اسم المقرر الدراسي.
- وعندما يتضمن الاختبار الواحد أكثر من فقرة، يتطلب الأمر
 وضع تعليمات عامة للاختبار الكلي، ثم وضع تعليمات خاصة
 بكل جزء، وعادة تُبين التعليمات الخاصة بطريقة تدوين الإجابة
 الخاصة بالجزء المعني
- 10- إخراج الاختبار وطباعته ويجب أن يشتمل على الآتي:
- وضوح الطباعة وخلوها من الأخطاء.
 - السؤال لا بد أن يكون واضحاً ويطرح مشكلة.
 - تُكتب الفقرة على الصفحة نفسها، ولا يجوز تقسيم فقرة السؤال
 إلى صفتين.
 - إعطاء أرقام متسلسلة لفقرات الاختبار.
- شروط الاختبار الجيد:
- إن الاختبار التحصيلي الجيد لا بد أن تتوفر فيه الشروط
 والخصائص الآتية [11,12]:
- أن يقيس الاختبار ما هو مفروض أن يقيسه، وهذا ما يعبر عنه
 بصحة الاختبار وصلاحيته "الصدق".
- 1- أن يكون الاختبار دقيقاً بحيث يعطي نفس النتائج إذا ما
 أعيد تطبيقه في مرات لاحقة، وهذا ما يعبر عنه بثبات الاختبار.
 - 2- أن يكون الاختبار موضوعياً بمعنى أنه يناسب مستوى
 الأفراد الذين يطبق عليهم الاختبار، وألا تختلف درجة الطالب
 باختلاف من يصح ورقته.
 - 3- أن يكون الاختبار صالحاً للتمييز بين الطلبة ذوي المستوى
 العالي والمستوى الضعيف، وهذا ما يعبر عنه بالتمييز.
 - 4- أن يكون الاختبار شاملاً بحيث يغطي أهداف ومفردات
 المقرر، وهذا ما يعبر عنه بالشمول.
 - 5- أن يكون من السهل استعماله وتطبيقه، من حيث الإجراء

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي

ب- جانب يتعلق بطريقة التصحيح وتقدير الدرجة للإجابات. ومن هنا فإن الموضوعية تعني أن واضع الاختبار ينبغي ألا يضعه بناءً على مزاجه وميوله الشخصية لبعض موضوعات المقرر وإهمال موضوعات أخرى، كما تعني ضرورة تحييد رأي المصحح وحكمه الشخصي عن عملية التصحيح [4,14,15].

رابعاً: يجب أن يكون الاختبار مميزاً بين الطلاب:

إن الاختبار الجيد ينبغي أن يكون صالحاً لقياس الفروق الدقيقة بين الطلاب أي يميز المجموعتين الدنيا والعليا ويبين بدقة الطالب الضعيف تحصيلياً والمتوسط والمتفوق. فالسؤال الذي تكون درجة تميزه عالية تعني أن نسبة من أجابوا عليه إجابة صحيحة من أفراد المجموعة العليا أكثر من نسبة من أجابوا عليه إجابة صحيحة من أفراد المجموعة الدنيا [16].

ويشير الخولي [17] إلى أن التمييز يعني أن يُبرز الاختبار الفروق بين الطلاب المتفوقين وغير المتفوقين، لذلك ينبغي أن تكون جميع أسئلة الاختبار مميزة، وأن كل سؤال تختلف الإجابة عنه باختلاف المتعلمين مما يتطلب أن يكون هناك مدى واسع بين السهولة والصعوبة في الأسئلة.

ومن الأمور ذات الأهمية العلاقة بين الصعوبة والقدرة التمييزية لل فقرات، فالبنود الصعبة جداً لا تميز بين الطلبة، كما أن البنود السهلة جداً لا تميز بينهم كذلك، فالبنود الصعبة يفشلون في الإجابة عليها كما أن البنود السهلة جداً يتمكن كل الطلاب أو معظمهم من الإجابة عليها [18].

ولتحقيق شرط التمييز ينبغي أن يراعي واضع الاختبار ما يأتي:

أ- يجب أن يكون هناك مدى واسعاً بين الصعوبة والسهولة بحيث نحصل على توزيع معقول للدرجات.

ب- يجب أن يكون هناك أسئلة في كل مستوى من مستويات العضوية بحيث نجد طلاباً يأخذون درجات مختلفة ويكون التوزيع معقولاً ومتماشياً مع تدرج الاختبار من السهل إلى الصعب في مسافات متساوية بقدر الإمكان.

ج- يجب أن يفرق كل سؤال بين الطالب الضعيف والطالب

والتصحيح ومعالجة النتائج، وهذا ما يعبر عنه بالعملية practicality ويمكن توضيح هذه الشروط والخصائص بشيء من التفصيل:

أولاً: صحة الاختبار وصلاحيته "الصدق Validity" يشير صدق الاختبار إلى قدرته على قياس السمة المراد قياسها، والاختبار الصادق والصالح هو الذي يقيس ما هو مفروض أن يقيسه تماماً، وألا يقيس أي أهداف أخرى غير أهداف المقرر، فإذا كان الاختبار يقيس التحصيل مثلاً يجب ألا تكون الأسئلة مشبعة بعامل الذكاء العام بحيث تكون النتائج موضحة لقوة الذكاء لا لقوة التحصيل.

وهناك بعض العوامل التي تؤثر على صدق وصلاحيته الاختبار مثل القدرة على القراءة، ضعف الذاكرة، ضعف المحصول اللغوي، بطء الاستجابة، فإذا لم يراع هذه الجوانب في الاختبار فإن الدرجة المعطاة للطالب لا تعبر عن تحصيله تماماً، ويكون الاختبار معيباً من ناحية صدقه [4,11].

هناك أنواع عديدة للصدق منها صدق المحتوى، وصدق المحك، والصدق الذاتي، وصدق التكوين والصدق العملي [10].

ولتقدير صدق الاختبار وصلاحيته يمكن مقارنة نتائجه بنتائج اختبار آخر سبق أن ثبت صلاحيته وصدقه لقياس نفس السمات وحساب معامل الارتباط بينهما، كما يمكن اللجوء إلى بعض الخبراء في المجال المعنى لإبداء رأيهم حول مدى صلاحية الاختبار وصدقه لقياس ما يجب أن يقيسه.

ثانياً: ثبات الاختبار Reliability

ويقصد به أن يُعطي الاختبار النتائج ذاتها تقريباً إذا ما أعيد تطبيقه على نفس المجموعة من الأفراد. وقد يعطي الاختبار درجات مختلفة إذا ما أعيد تطبيقه ولا يكمن ذلك لعيب فيه، إنما يرجع إلى تغير السمة المقاسة أو يرجع لعيب في تطبيق الاختبار، أو عيب في عملية التصحيح [13].

ثالثاً: يجب أن يكون الاختبار موضوعياً Objectivity، وللموضوعية جانبان هما:

أ- جانب يتعلق بفهم الطالب لكل سؤال وتفسيره.

القوي [11].

الدراسات السابقة:

قام القضاة وآخرون [21]، بدراسة هدفت إلى تعرف مدى تطبيق معايير تصميم الاختبار الجيد من وجهة نظر أعضاء هيئة التدريس والطلبة في كلية المعلمين بأبها، وتكونت عينة الدراسة من (100) عضو من أعضاء هيئة التدريس، كما بلغ عينة الطلبة (60) طالباً، وأظهرت النتائج أن موضوع تصميم الاختبار الجيد ينال أهمية متوسطة لدى أعضاء هيئة التدريس. كما أظهرت النتائج بأن موضوع تصميم الاختبار الجيد من وجهة نظر الطلبة، قد نال أهمية جيدة، وأظهرت النتائج عدم وجود فروق في تقديرات أعضاء هيئة التدريس لمدى تطبيق معايير تصميم الاختبار الجيد تعزى للتخصص، ووجود فروق من وجهة نظرهم تُعزى لسنوات الخبرة.

أما دراسة المصري ومرعي [22] فهذهت إلى التعرف على اتجاهات طلبة جامعة الإسراء الخاصة نحو أساليب التقويم المختلفة والفروق في هذا المتغير تبعاً للجنس والتخصص، والتحصيل الأكاديمي، والتفاعل بينهما، وتكونت عينة الدراسة من (191) طالباً وطالبة، طبقت عليهم استبانة تألفت من (37) فقرة. وأظهرت النتائج أن اتجاهات طلبة جامعة الإسراء الخاصة نحو أساليب التقويم كانت إيجابية، كما أظهرت وجود فروق في الاتجاهات تعزى إلى متغير مستوى التحصيل الأكاديمي (جيد فأعلى)، كما بينت النتائج عدم وجود فروق في الاتجاهات تعزى إلى متغير الجنس والتخصص الأكاديمي والتفاعل بينهما.

وأجرت سنان [23] دراسة هدفت إلى تعرف درجة معرفة وممارسة معلمات الرياضيات بالمرحلة المتوسطة لمهارات بناء الاختبارات التحصيلية في مدينة مكة المكرمة، وتكونت عينة الدراسة من (74) مدرسة تم اختبارها عشوائياً، وأظهرت النتائج بأن مواصفات الاختبار الجيد المتمثل في (الموضوعية، والصدق، والشمول) متدنية في اختبارات أفراد العينة، كما بينت النتائج أن 72% لم يحصلن على دورات تأهيلية على رأس العمل، و21% حصلن على دورة واحدة، وقد كان الأداء متدنياً على درجات المقياس الخاص بإعداد الفقرات، وإخراج الاختبار

خامساً: أن يكون الاختبار شاملاً Comprehensiveness ينبغي أن تكون أسئلة الاختبار عينة ممثلة تقيس الأهداف والمحتوى وفقاً لأولوية تلك الأهداف في مجال العملية التربوية، وفي هذه الحالة لا بد من بناء جدول مواصفات الاختبار التحصيلي الذي يساعد في وضع خطوط عريضة لبناء اختبار يقيس مادة التعلم بشكل أفضل، وتقدم تقديراً لعدد الأسئلة التي يحتاجها كل نوع من الأهداف للمقرر الدراسي [15].

ويشير صبري والرفاعي [19]، والنور [4]. إلى أهمية بناء جدول المواصفات عند تصميم الاختبار التحصيلي الذي يُبنى وفقاً لمجموعة من المواصفات التي تصف المجال الذي يقيسه الاختبار ويكون إطاراً للحصول على عينة ممثلة من مخرجات التعلم، ويمكن اتباع الخطوات الآتية عند بناء جدول مواصفات الاختبار التحصيلي:

أ- تحديد أهداف أو مخرجات التعلم التي يقيسها الاختبار.

ب- تحديد عناصر المحتوى التي يغطيها الاختبار.

ج- بناء جدول مواصفات ذي محورين أحدهما للأهداف والآخر لعناصر المحتوى.

د- تحديد الوزن النسبي للأهداف ولعناصر المحتوى في كل خلية من خلايا الجدول.

هـ- تحديد عدد أسئلة الاختبار لكل خلية في كل عنصر من عناصر المحتوى وفي كل هدف سلوكي.

سادساً: العملية Practicality

يجب أن يكون الاختبار ممكن التطبيق من حيث كل من الآتي:

1- وضوح التعليمات بالنسبة للطلاب والطالبات والقائمين على التطبيق معاً.

2- الوقت اللازم للتطبيق.

3- توافر الكفايات المطلوبة لدى القائمين على التطبيق.

4- توافر الأمان اللازم للتطبيق.

5- سهولة تقدير الدرجات وفي الوقت المخصص لذلك.

6- ذات قيمة تعليمية أي تزويد الطلاب بتغذية راجعة [20].

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي

وتطبيقه وتصحيحه.

وطالبة من مختلف كليات الجامعة، أظهرت نتائج الدراسة أن اتجاهات الطلبة تمثلت في العدالة في تقويم الطلبة، وتتنوع أساليب التقويم، واستثارة الاختبارات لذهن الطالب، وعدم تركيزها على الحفظ، وتركيزها على التفكير الناقد.

4. الطريقة والإجراءات

أ. منهج الدراسة

إن طبيعة الدراسة الحالية تتطلب استخدام الوصف والتحليل والمقارنة والتفسير، ولذلك استخدم الباحث المنهج الوصفي التحليلي ومهج البحث ملائم لطبيعة الدراسة.

ب. مجتمع الدراسة

يتكون مجتمع الدراسة من جميع طلبة كلية التربية بجامعة السودان للعلوم والتكنولوجيا الذين درسوا مقرر اللغة العربية 1، واللغة العربية 2، واللغة العربية 3 للفصل الدراسي الأول من العام 2012م/2013م والبالغ عددهم (1820) طالباً وطالبة.

ج. عينة الدراسة

تكونت عينة الدراسة من (300) طالب وطالبة تم اختيارهم بالطريقة العشوائية والجدول (1) يوضح العينة حسب الجنس.

أما دراسة الناجم [24] فهدفت إلى تعرف المشكلات التي يشعر بوجودها طلبة كلية التربية - جامعة الملك فيصل تبعاً لمتغير الجنس، والتخصص، والمستوى الدراسي، وتكونت عينة الدراسة من (319) طالباً وطالبة. وأظهرت النتائج أكثر المشكلات التي يراها الطلبة هي سوء أسئلة الاختبارات وافتقارها إلى تنمية مستويات التفكير العليا، وعدم موضوعية نتائج الاختبارات، وعدم مراعاة الطلبة في وضع جداول الاختبارات، ووجود فروق بين طلبة القسم العلمي وطلبة القسم الأدبي في وجود هذه المشكلات وكذلك في المستويات الدراسية المختلفة.

وأوضحت دراسة جرينولد [25] إلى أن الطلبة يعطون تقديرات عالية للأساتذة المتساهلين في وضع العلامات والعكس صحيح بالنسبة للأساتذة الذين يتشددون في إعطاء الدرجات، أي أن تقييم الطلبة للأساتذة يعتمد على مدى تساهل أو تشدد الأساتذة في إعطاء الدرجات للطلبة.

كما أجرى نبروي ويحيى [26] دراسة هدفت إلى كشف اتجاهات طلبة جامعة الإمارات العربية المتحدة نحو الممارسات التربوية في الجامعة، وتكونت عينة الدراسة من (246) طالباً

جدول 1

يوضح عينة الدراسة

المتغير	العدد	المجموع
الجنس	ذكر	300
	أنثى	165
التخصص	علمي	300
	أدبي	127
		173

تم التحقق من صدق محتوى أداة الدراسة من حيث الوضوح والشمول، ومناسبة البند للمجال الذي ينتمي إليه من خلال عرض الأداة بصورتها الأولية على عدد من المتخصصين في مجال التربية وعلم النفس والقياس والتقويم حيث طُلب منهم إبداء رأيهم حول أبعاد المقياس وتحديد مدى الارتباط مع مشكلة الدراسة، وفي ضوء ملاحظات المحكمين واقتراحاتهم قام الباحث بإجراء التعديلات المناسبة على الأداة لتأخذ صورتها النهائية كما في ملحق البحث.

الخصائص السيكومترية لأداة الدراسة:

لتحقيق أهداف البحث والإجابة عن تساؤلاته، قام الباحث بتصميم أداة وكتابة بنوده في ضوء أدبيات البحث والدراسات السابقة في هذا المجال، حيث تكونت الأداة من (36) فقرة وفقاً لنموذج (ليكرت) ذات البدائل الخمسة وهي بدرجة كبيرة جداً، بدرجة كبيرة، بدرجة متوسطة، بدرجة ضعيفة، بدرجة ضعيفة جداً.

د. صدق الأداة Validity

هـ. ثبات الأداة Reliability

للتحقق من ثبات الأداة قام الباحث بحساب ثبات الاتساق الداخلي باستخدام إحصائيات الفقرة للمقياس، حيث تم تطبيق أداة الدراسة على عينة تكونت من (90) طالباً وطالبة تم اختيارهم عشوائياً، وبلغ معامل ثبات الاتساق الداخلي للمقياس ككل (0.78) وللمجالات بلغ معاملات الاتساق الداخلي 0.79، و0.76، و0.81، و0.79 و0.75، و0.82، على التوالي، وهي قيم مقبولة لأغراض هذا البحث.

طريقة تصحيح استجابات أفراد العينة على فقرات المقياس:

استخدم الباحث طريقة مقياس ليكرت الخماسي لقياس وجهة نظر الطلبة في تقييمهم لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد، وقد أعطيت الدرجات (5-1) لاحتمالات الإجابة عن كل فقرة من الفقرات، حيث وزعت الدرجات تنازلياً على التقديرات بدرجة كبيرة جداً، بدرجة كبيرة، بدرجة متوسطة، بدرجة ضعيفة، بدرجة ضعيفة جداً، كما أعطيت الدرجات تصاعدياً (5-1) للفقرات السلبية وهي الفقرة 21، 22، 23، 24، 26، 28.

لتحديد استجابة العينة وبعد الاطلاع على الدراسات السابقة وفي ضوء آراء المحكمين والأدب النظري تم تحديد درجة الالتزام التي يمكن من خلالها الحكم على وجهة نظر الطلبة وتقييمهم لمستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار الجيد بمستويات ثلاثة هي: كبيرة، متوسطة، ضعيفة، وفقاً للمعادلة التالية:

القيمة العليا لبدائل الإجابة في أداة البحث - القيمة الدنيا لبدائل الإجابة في أداة البحث مقسومة على عدد المستويات الثلاثة (كبيرة، ومتوسطة، وضعيفة) حيث إن:

$1.33 = (5-1) \div 3$ ، وهذه القيمة تساوي طول الفئة بين المستويات الثلاثة (كبيرة، متوسطة، ضعيفة).

$$1 + 1.33 = 2.33.$$

$$2.33 + 1.33 = 3.66.$$

$$3.66 + 1.33 = 5.00$$

وبذلك تعد قيم المتوسطات الحسابية لتقييم الطلبة لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد تتراوح ما بين:

(من 1.00 - أقل من 2.33) قيمة ضعيفة.

(من 2.33 - أقل من 3.66) قيمة متوسطة.

(من 3.66 - 5.00) قيمة كبيرة.

متغيرات الدراسة:

تتضمن الدراسة متغيراً أساسياً وهو تقييم الطلبة لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد، كما تتضمن المتغيرات المستقلة الآتية:

1- الجنس: وله مستويان (طلاب، طالبات).

2- التخصص الدراسي: وله مستويان (علمي، أدبي).

3- المستوى الدراسي: وله مستويان (السنة الأولى، السنة الثانية).

4- المعدل التراكمي وله ثلاثة مستويات (ضعيف، متوسط، مرتفع).

المعالجة الإحصائية:

- استخدم المتوسطات الحسابية والانحرافات المعيارية للإجابة عن السؤال الأول وتحديد وجهة نظر الطلبة على كل بعد من أبعاد أداة البحث.

- استخدم نتائج اختبار (ت) لعينتين مستقلتين لحساب دلالة الفروق بين المتوسطات الإجابة عن السؤال الثاني والثالث والرابع.

- للإجابة عن السؤال الخامس استخدم نتائج تحليل التباين الأحادي (ONE WAY ANOVA)، بالإضافة إلى اختبار شيفيه للمقارنات البعدية.

5. النتائج

النتائج المتعلقة بالسؤال الأول: ما تقييم الطلبة لمستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لكل بعد من أبعاد معايير تصميم الاختبار التحصيلي الجيد في اختبارات اللغة العربية بكلية التربية جامعة السودان للعلوم والتكنولوجيا؟.

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي

للإجابة عن هذا السؤال استخدم المتوسط الحسابي والانحراف المعياري لكل فقرة من فقرات المقياس، ولكل بُعد من أبعاده، والدرجة الكلية للمقياس، وتوضح الجداول من (2-7) هذه النتائج تبعاً لكل بُعد.

البُعد الأول: المعايير المطلوبة لكتابة تعليمات الاختبار التحصيلي الجيد.

لتحليل نتائج البعد تم استخراج المتوسطات الحسابية والانحرافات المعيارية والرتبة لإجابات وحساب المتوسط الكلي للبعد والجدول رقم (2) يوضح ذلك.

جدول 2

المتوسطات الحسابية والانحرافات المعيارية والرتبة لإجابات العينة عن بعد المعايير المطلوبة لكتابة تعليمات الاختبار التحصيلي الجيد

م	العبرة	المتوسط	الانحراف المعياري	الرتبة
7	تكتب التعليمات بصيغة بسيطة وسهلة يفهما الجميع	4.34	1.49	1
2	توضح التعليمات الزمن المخصص للإجابة عن الاختبار	4.26	1.37	2
1	يتضمن ورقة الاختبار تعليمات عامة (اليوم، التاريخ، اسم المقرر)	4.08	1.49	3
4	تحدد التعليمات الدرجة الخاصة بكل سؤال	3.19	1.62	4
3	تحدد التعليمات عدد أسئلة الاختبار	3.12	1.64	5
6	هناك تعليمات واضحة خاصة بكل سؤال	3.03	1.48	6
5	تحد التعليمات عدد صفحات الاختبار	2.92	1.67	7
	الكلي للبعد	3.56	1.53	

يوضح الجدول (2) أن مستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد كان متوسطاً من وجهة نظر الطلبة حيث بلغ المتوسط (3.56) بانحراف معياري (1.53).

وجاءت بعض العبارات بدرجة تحقق كبيرة وهي: (تكتب التعليمات بصيغة بسيطة وسهلة يفهما الجميع) في المرتبة الأولى، بمتوسط حسابي (4.34) وانحراف معياري (1.49)، والعبرة (توضح التعليمات الزمن المخصص للإجابة عن الاختبار) في المرتبة الثانية، بمتوسط (4.26) وانحراف معياري (1.37)، والعبرة (يتضمن ورقة الاختبار تعليمات عامة (اليوم،

وجاءت بقية عبارات البعد بدرجة متوسطة من حيث درجة التحقق، وتراوح متوسطاتها بين (2.92 - 3.19).

البُعد الثاني: المعايير المطلوبة لكتابة فقرات الاختبار التحصيلي الجيد:

لتحليل نتائج هذا البعد استخدم المتوسطات الحسابية والانحرافات المعيارية والرتب لإجابات عينة البحث، إضافة إلى حساب المتوسط الكلي للبعد كما هو موضح بالجدول (3).

جدول 3

المتوسطات الحسابية والانحرافات المعيارية والرتب لعبارات بعد المعايير المطلوبة لكتابة فقرات الاختبار التحصيلي الجيد

م	العبرة	المتوسط	الانحراف المعياري	الرتبة
16	تتناول الأسئلة جوانب ذات قيمة في المقرر	3.94	1.36	1
9	لغة أسئلة الاختبار واضحة وسهلة الفهم	3.87	1.55	2
8	تحدد المشكلة بشكل واضح في السؤال	3.75	1.49	3
15	تستخدم أسئلة محددة ذات إجابات متفق عليها	3.11	1.58	4
11	أشعر بأن الزمن المخصص للاختبار لا يسمح بالإجابة عن جميع الأسئلة	2.89	1.64	5
10	مستويات أسئلة الاختبار مناسبة لمستوى الأهداف	2.62	1.76	6
13	يشتمل أسئلة الاختبار على أسئلة موضوعية	2.17	1.58	7
14	توجد بدائل للإجابة في أسئلة الاختبار من متعدد تتراوح بين (4-5).	2.16	1.62	8
12	تنظم أسئلة الاختبار وتسلسل من السهل إلى الصعب	2.13	1.56	9

الكلبي للبعد	2.96	1.59
يوضح جدول (3) أن مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية للمعايير المطلوبة لكتابة فقرات الاختبار التحصيلي الجيد كان متوسط من وجهة نظر الطلبة، حيث بلغ المتوسط الحسابي (2.96) بانحراف معياري (1.57)، مما يشير إلى أن تقييم الطلبة لهذا البعد جاء متوسطاً.	حيث التحقق بمتوسط (3.87) وانحراف معياري (1.55). أما العبارات (يشمل أسئلة الاختبار على أسئلة موضوعية)، وتوجد بدائل للإجابة تتراوح بين 4-5)، وتنظيم أسئلة الاختبار وتسلسل من السهل إلى الصعب) فجاءت جميعها ضعيفة من حيث درجة التحقق.	
وجاءت العبارة (تتناول الأسئلة جوانب ذات قيمة في المقرر) في المرتبة الأولى من حيث درجة التحقق بمتوسط (3.94)، وانحراف معياري (1.37)، كما جاءت العبارة (لغة أسئلة الاختبارات واضحة وسهلة الفهم) في المرتبة الثانية من	البُعد الثالث: معايير موضوعية الاختبار وصلاحيته لتحليل نتائج هذا البعد استخدم المتوسطات الحسابية والانحرافات المعيارية والترتيب الإجابات عينة البحث، إضافة إلى حساب المتوسط الكلبي للبعد والجدول (4) يبين ذلك:	

جدول 4

المتوسطات الحسابية والانحرافات المعيارية والترتيب لعبارات بعد معايير موضوعية الاختبار وصلاحيته

م	العبارة	المتوسط	ع	الرتبة
19	يقيس الاختبار موضوعات جميعها داخل المقرر	4.21	1.34	1
20	الأسئلة واضحة وجميع موضوعاتها درست لي	3.84	1.54	2
22	الاختبار بوصفه الحالي لا يعكس قدراتي على الإجابة	3.52	1.64	3
18	موضوعية التصحيح هي الأساس في تصحيح الاختبارات	3.32	1.59	4
23	أشعر بأن الاختبار يقيس ما وضع لقياسه	3.24	1.67	5
17	أرى أن الدرجة التي تمنح لي في الاختبار تعبر عن أدائي	3.16	1.77	6
24	أرى أن درجاتي تكون أفضل إذا صححت ورقتي من قبل عضو آخر.	3.00	1.74	7
21	أرى أن خطي له دور في انخفاض درجاتي	2.82	1.79	8
	الكلبي للبعد	3.38	1.64	

يوضح جدول (4) أن مستوى تطبيق أعضاء هيئة التدريس لمعايير موضوعية الاختبار وصلاحيته في اختبارات اللغة العربية بكلية التربية كان متوسطاً من وجهة نظر الطلبة حيث بلغ المتوسط (3.38) بانحراف معياري (1.064)، حيث جاءت العبارة (يقيس الاختبار موضوعات جميعها بالمقرر الدراسي) في المرتبة الأولى بدرجة تحقق كبيرة بمتوسط (4.21)، وانحراف معياري (1.34)، وجاءت العبارة (الأسئلة واضحة وجميع موضوعاتها درست لي) في المرتبة الثانية بدرجة تحقق كبيرة

حيث بلغ المتوسط الحسابي (3.84) بانحراف معياري (1.54)، وجاءت بقية فقرات البعد في مستوى تحقق متوسط حيث بلغ متوسط العبارات بين (2.82-3.52).

البُعد الرابع: معايير شمول الاختبار للأهداف والمحتوى الدراسي لتحليل نتائج هذه البعد استخدام المتوسطات الحسابية والانحرافات المعيارية والترتيب لإجابات عينة البحث، إضافة إلى حساب المتوسط الكلبي للبعد كما هو مبين بالجدول (5).

جدول 5

المتوسطات الحسابية والانحرافات المعيارية والترتيب لبعده شمول الاختبار للأهداف والمحتوى الدراسي

م	العبارة	المتوسط	الانحراف المعياري	الرتبة
26	أرى أن الأسئلة مصممة لقياس قدراتي على الحفظ والتذكر	3.94	1.42	1
28	أرى أن الاختبار يركز على الأسئلة المقالية ولا تغطي ما استذكره من دروس	3.26	1.53	2
25	أرى أن أسئلة الاختبار تغطي معظم الموضوعات بالمقرر الدراسي	2.96	1.62	3

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي

27	أرى أن أسئلة الاختبار مصممة لقياس مهاراتي العقلية العليا (التحليل، التركيب، التقويم).	2.67	1.64	4
	الكلّي للبعد	3.21	1.55	

يوضح جدول (5) أن مستوى تطبيق أعضاء هيئة التدريس لمعايير شمول الاختبار للأهداف والمحتوى في اختبارات اللغة العربية من وجهة نظر الطلبة كان متوسطاً، حيث بلغ المتوسط (3.21)، وانحراف معياري (1.55)، وجاءت العبارة (أرى أن الأسئلة مصممة لقياس قدراتي على الحفظ والتذكر) في المرتبة الأولى بدرجة تحقق كبيرة، وبمتوسط (3.94)، وانحراف معياري (1.42)، وجاءت العبارة (أرى أن الاختبار يركز على الأسئلة المقالية ولا تغطي ما استذكره من دروس) في المرتبة الثانية وبدرجة تحقق متوسطة، وبمتوسط (3.26)، وانحراف معياري (1.53). كما جاءت العبارة (أرى أن أسئلة الاختبار تغطي

معظم الموضوعات بالمقرر الدراسي) في المرتبة الثالثة وبدرجة تحقق متوسطة، وبمتوسط (2.96)، وانحراف معياري (1.62)، وجاءت العبارة (أرى أن أسئلة الاختبار مصممة لقياس مهاراتي العقلية العليا (التحليل، التركيب، التقويم) في المرتبة الأخيرة وبدرجة تحقق متوسطة، وبمتوسط (2.67)، وانحراف معياري (1.64).

البُعد الخامس: معايير القدرة التمييزية لأسئلة الاختبار لتحليل نتائج هذا البعد استخدام المتوسطات الحسابية والانحرافات المعيارية والرتب لإجابات أفراد العينة، إضافة إلى حساب المتوسط الكلّي للبعد كما هو مبين بالجدول (6).

جدول 6

المتوسطات الحسابية والانحرافات المعيارية والرتب لبعدها معايير القدرة التمييزية لأسئلة الاختبار

م	العبارة	المتوسط	الانحراف المعياري	الرتبة
30	إن الاختبار صعب بدرجة معقولة	2.91	1.58	1
29	أرى أن جميع الطلبة الذين يدرسون المقرر لا يحصلون على درجات متقاربة.	2.78	1.64	2
31	عندما اجتهد في الاستدكار أجد أن درجاتي مرتفعة مقارنة بزملائي	2.73	1.67	3
32	السهولة هي السمة المميزة للاختبار	2.58	1.58	4
	الكلّي للبعد	2.75	1.62	

يوضح الجدول (6) أن مستوى تطبيق أعضاء هيئة التدريس لمعايير القدرة التمييزية لأسئلة اختبارات اللغة العربية كان متوسطاً من وجهة نظر الطلبة، حيث بلغ المتوسط الكلّي (2.75) بانحراف معياري (1.62).

البُعد السادس: معايير طباعة الاختبار التحصيلي: لتحليل نتائج هذا البعد استخدم المتوسطات الحسابية والانحرافات المعيارية والرتب لإجابات أفراد عينة البحث بالإضافة إلى المتوسط الكلّي كما هو مبين بالجدول (7).

جدول 7

المتوسطات الحسابية والانحرافات المعيارية والرتب لإجابات العينة في بعد معايير طباعة الاختبار التحصيلي

م	العبارة	المتوسط	الانحراف المعياري	الرتبة
35	نسخ الاختبار واضحة بعد التصوير	4.52	1.33	1
36	أرقام الأسئلة مناسبة من البداية إلى النهاية ولا يوجد تداخل في الأرقام	4.48	1.37	2
34	المسافات بين الأسئلة مناسبة	4.36	1.39	3
33	لا توجد أخطاء طباعية أو إملائية بالاختبار	4.12	1.46	4
	الكلّي للبعد	4.37	1.39	

يوضح الجدول (7) أن مستوى تطبيق أعضاء هيئة التدريس لمعايير طباعة الاختبار التحصيلي لاختبارات اللغة العربية كان كبيراً من وجهة نظر الطلبة، حيث بلغ المتوسط (4.37) بانحراف معياري (1.39) وجاءت جميع عبارات البعد بدرجة تحقق كبيرة.

تقويم مستوى تطبيق معايير تصميم الاختبار التحصيلي الجيد لاختبارات اللغة العربية (بشكل كلي).

لتقويم مستوى تطبيق معايير تصميم الاختبار الجيد لاختبارات

اللغة العربية من وجهة نظر الطلبة استخدم المتوسطات الحسابية والانحرافات المعيارية والرتب لكل بعد إضافة إلى المتوسط الكلي

جدول 8

المتوسطات الحسابية والانحرافات المعيارية والترتيب لكل بعد من أبعاد المقياس، والمتوسط الكلي للأبعاد

الرتبة	الانحراف المعياري	المتوسط	العبرة
1	1.39	4.37	معايير طباعة الاختبار
2	1.53	3.56	المعايير المطلوبة لكتابة تعليمات الاختبار
3	1.64	3.38	معايير موضوعية الاختبار وصلاحيته
4	1.55	3.21	معايير شمول أسئلة الاختبار للأهداف والمحتوى
5	1.59	2.96	المعايير المطلوبة لكتابة فقرات الاختبار
6	1.62	2.75	معايير القدرة التمييزية لأسئلة الاختبار
	1.55	3.37	متوسطة الدرجة الكلية

الاختبار، وجميع هذه الأبعاد أنتت بدرجة تحقق متوسطة عدا بُعد معايير طباعة الاختبار التحصيلي فتتحقق بدرجة كبيرة. السؤال الثاني: هل توجد فروق ذات دلالة إحصائية في تقييم الطلبة لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد في اختبارات اللغة العربية بكلية التربية تعزى لمتغير الجنس؟ وللإجابة عن هذا السؤال استخدم المتوسطات الحسابية والانحرافات المعيارية ونتائج اختبار (ت) لمتغير الجنس كما هو موضح بجدول (9):

يوضح الجدول (8) أن المتوسط الكلي لتقويم الطلبة لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد في اختبارات اللغة العربية كان متوسطاً حيث بلغ المتوسط (3.37)، وانحراف معياري (1.55)، وجاء بعد معايير طباعة الاختبار في المرتبة الأولى ويليه بعد المعايير المطلوبة لكتابة تعليمات الاختبار في المرتبة الثانية، كما جاء بعد معايير موضوعية الاختبار وصلاحيته في المرتبة الثالثة، ومعايير شمول أسئلة الاختبار للأهداف والمحتوى الدراسي في المرتبة الرابعة، وبعد المعايير المطلوبة لكتابة فقرات الاختبار في المرتبة الخامسة، وأخيراً بعد معايير القدرة التمييزية لأسئلة

جدول 9

نتائج اختبار (ت) لدلالة الفروق تبعاً لمتغير الجنس

مستوى الدلالة	قيمة ت	ع	م	المتغير	مجالات الأداة
.29	1.13	1.46	4.09	ذكور	المعايير المطلوبة لكتابة تعليمات الاختبار
		1.37	4.27	إناث	
.38	.28	1.48	2.93	ذكور	المعايير المطلوبة لكتابة فقرات الاختبار
		1.57	2.98	إناث	
.05	*3.06	1.41	3.17	ذكور	معايير موضوعية الاختبار وصلاحيته
		1.38	3.66	إناث	
.05	*2.68	1.54	3.05	ذكور	معايير شمول أسئلة الاختبار للأهداف وللحموى
		1.46	3.48	إناث	
.05	*2.76	1.51	3.16	ذكور	معايير القدرة التمييزية لأسئلة الاختبار
		1.42	2.69	إناث	
.36	.29	1.48	4.74	ذكور	معايير طباعة الاختبار
		1.48	4.69	إناث	
.07	1.7				الكلي

متوسطات إجابات الطلبة على مستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد في اختبارات اللغة العربية تعزى للمستوى الدراسي؟
للإجابة عن هذا السؤال استخدام المتوسطات الحسابية والانحرافات المعيارية واختبار (ت) لدلالة الفروق بين متوسطات إجابات الطلبة على مستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد في اختبارات اللغة العربية تعزى للمستوى الدراسي كما هو موضح بجدول (10).

يوضح جدول (9) أنه توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha = 0.05$) بين متوسطات عينة البحث تعزى لمتغير الجنس، ولصالح الطالبات الإناث في بعض الأبعاد وهي: بعد معايير موضوعية الاختبار وصلاحيته حيث بلغ قيمة ت (3.06)، وبعد معايير شمول أسئلة الاختبار للأهداف وللمحتوى حيث بلغ قيمة ت (2.68)، وبعد معايير القدرة التمييزية لأسئلة الاختبار حيث بلغ قيمة ت (2.76).
السؤال الثالث: هل توجد فروق ذات دلالة إحصائية بين

جدول 10

نتائج اختبار (ت) لدلالة الفروق بين متوسطات إجابات الطلبة تبعاً للمستوى الدراسي

الدلالة	قيمة ت	ع	م	المتغير	مجالات الأداة
.34	.56	1.37	4.41	السنة الأولى	المعايير المطلوبة لكتابة تعليمات الاختبار
		1.41	4.32	السنة الثانية	
.29	.89	1.57	3.11	السنة الأولى	المعايير المطلوبة لكتابة فقرات الاختبار
		1.49	2.95	السنة الثانية	
.34	.47	1.38	3.49	السنة الأولى	معايير موضوعية الاختبار وصلاحيته
		1.47	3.41	السنة الثانية	
.36	.31	1.42	3.32	السنة الأولى	معايير شمول أسئلة الاختبار للأهداف وللمحتوى
		1.39	3.37	السنة الثانية	
.17	.94	1.56	2.99	السنة الأولى	معايير القدرة التمييزية لأسئلة الاختبار
		1.58	2.82	السنة الثانية	
.07	1.12	1.47	4.29	السنة الأولى	معايير طباعة الاختبار
		1.43	4.48	السنة الثانية	
.08	.72				الكلية

والتكنولوجيا لمستوى تطبيق أعضاء هيئة التدريس معايير تصميم الاختبار التحصيلي الجيد في اختبارات اللغة العربية تعزى للتحقق الدراسي؟
للإجابة عن هذا السؤال تم استخدام المتوسطات والانحراف المعياري واختبار (ت) لدلالة الفروق كما هو موضح بالجدول (11).

يوضح الجدول (10) أنه لا توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha = 0.05$) بين متوسطات عينة البحث تعزى لمتغير المستوى الدراسي (السنة الأولى، والسنة الثانية)، حيث بلغ قيمة ت الكلية (0.72).
السؤال الرابع: هل توجد فروق ذات دلالة إحصائية بين متوسطات إجابات طلبة كلية التربية بجامعة السودان للعلوم

جدول 11

اختبار (ت) لدلالة الفروق بين متوسطات إجابات طلبة التخصصات العلمية والتخصصات الأدبية في تقييمهم لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد في اختبار اللغة العربية

مستوى الدلالة	قيمة ت	ع	م	المتغير	مجالات الأداة
.07	1.3	1.51	4.26	أدبي	المعايير المطلوبة لكتابة تعليمات الاختبار
		1.48	4.02	علمي	
.05	*2.25	1.39	3.17	أدبي	المعايير المطلوبة لكتابة فقرات الاختبار
		1.48	2.81	علمي	

.23	.71	1.53	3.41	أدبي	معايير موضوعية الاختبار وصلاحيته
		1.44	3.29	علمي	
.03	*4.25	1.45	3.69	أدبي	معايير شمول أسئلة الاختبار للأهداف والمحتوى
		1.49	3.01	علمي	
.05	*3.44	1.37	3.24	أدبي	معايير القدرة التمييزية لأسئلة الاختبار
		1.43	2.69	علمي	
.26	.35	1.52	4.48	أدبي	معايير طباعة الاختبار
		1.41	4.42	علمي	
.05	*2.05				الكلي

* دال عند مستوى دالة $\alpha = 0.05$

الجيد في اختبارات اللغة العربية أعلى من تقييم طلبة التخصصات العلمية، وهذه النتيجة تبدو منطقية بسبب تفضيل طلبة التخصصات الأدبية للغة العربية كونها الأقرب إلى تخصصاتهم الأدبية.

النتائج المتعلقة بالسؤال الخامس:

هل توجد فروق بين متوسطات إجابات أفراد عينة البحث تعزى للمعدل التراكمي (ضعيف، متوسط، مرتفع).

للإجابة عن هذا السؤال استخدم نتائج تحليل التباين الأحادي لدلالة الفروق بين متوسطات أداء أفراد العينة وفقاً للمعدل التراكمي؟ كما هو موضح بالجدول (11).

يوضح الجدول (11) أن هناك فروقاً ذات دلالة إحصائية عند مستوى (0.05α) لمتغير التخصص الدراسي (أدبي، علمي) ولصالح التخصص الأدبي، حيث بلغت قيمة ت الكلية (2.05) وهي دالة إحصائياً عند مستوى (0.05α) ، وكانت الفروق دالة في بعض الأبعاد وهي: المعايير المطلوبة لكتابة فقرات الاختبار، ومعايير شمول أسئلة الاختبار للأهداف والمحتوى، ومعايير القدرة التمييزية لأسئلة الاختبار لصالح التخصص الأدبي بينما لم تظهر فروقاً دالة في الأبعاد الثلاثة الأخرى.

وهذه النتيجة تشير إلى أن تقييم طلبة التخصصات الأدبية لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار

جدول 11

نتائج تحليل التباين الأحادي لدلالة الفروق بين متوسطات أداء أفراد العينة وفقاً للمعدل التراكمي

الدلالة	قيمة ف	متوسط المربعات	د.ح	مجموع المربعات	مصدر التباين
.001	7.033*	506.545	2	1013.09	بين المجموعات
		72.019	277	19949.34	داخل المجموعات
			299	20962.43	المجموع

* دال عند مستوى الدلالة $\alpha = 0.05$.

لمعايير تصميم الاختبار الجيد.

وللتعرف على مصادر هذه الفروق تم إجراء مقارنات بعدية بطريقة (شيفيه) كما هو موضح بالجدول (12).

يوضح الجدول (11) وجود فروق ذات دلالة إحصائية بين متوسطات أداء الطلبة عينة البحث تعزى لمتغير المعدل التراكمي، مما تعني بأن للمعدل التراكمي تأثيراً دالاً على وجهة نظر الطلبة عند تقييمهم لمستوى تطبيق أعضاء هيئة التدريس

جدول 12

نتائج اختبار (شيفيه) للمقارنات البعدية بين متوسطات الأداء تبعاً لمتغير المعدل التراكمي

مرتفع (3-4)	متوسط (2.51-2.99)	ضعيف (2-2.5)	المعدل التراكمي
8.49**	068.	-	ضعيف (2-2.5)
3.42*	-	-	متوسط (2.51-2.99)

دال عند مستوى الدلالة $\infty = 0.05$ ، *دال عند مستوى الدلالة $\infty = 0.01$.

يوضح الجدول (12) أن هناك فروقاً ذات دلالة إحصائية على اختبار (شيفيه) بين المعدل التراكمي (ضعيف، مرتفع) ولصالح المعدل المرتفع، كما يبين أن هناك فروقاً بين المعدل التراكمي (متوسط، مرتفع) لصالح المعدل المرتفع، بينما لم تظهر فروق بين المعدل التراكمي الضعيف والمتوسط. أشارت النتائج إلى أن هناك فروقاً ذات دلالة إحصائية في تقييم الطلبة ووجهة نظرهم حول مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار الجيد تبعاً للمعدل التراكمي، لصالح المعدل التراكمي المرتفع.

6. مناقشة النتائج

هدف البحث إلى تقويم مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد من وجهة نظر طلبة كلية التربية بجامعة السودان للعلوم والتكنولوجيا وعلاقة ذلك بكل من متغيرات الجنس والمستوى الدراسي، والتخصص الدراسي، والمعدل التراكمي.

أشارت نتائج السؤال الأول لهذا البحث إلى أن مستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد كان متوسطاً من وجهة نظر الطلبة، حيث جاء بُعد معايير طباعة الاختبار في المرتبة الأولى ويليها بُعد المعايير المطلوبة لكتابة تعليمات الاختبار في المرتبة الثانية، كما جاء بُعد معايير موضوعية الاختبار وصلاحيته في المرتبة الثالثة، ومعايير شمول أسئلة الاختبار للأهداف وللمحتوى الدراسي في المرتبة الرابعة، وبُعد المعايير المطلوبة لكتابة فقرات الاختبار في المرتبة الخامسة، وأخيراً بُعد معايير القدرة التمييزية لأسئلة الاختبار، وجميع هذه الأبعاد أنت بدرجة تحقق متوسطة عدا بُعد معايير طباعة الاختبار التحصيلي فتتحقق بدرجة كبيرة.

وتشير هذه النتيجة بأن مقررات اللغة العربية مصممة لتلائم أسلوب التلقين والإلقاء والإصغاء ولا تسترعي لدى الطلبة التفكير والتحليل والتركيب والتقويم وبالتالي جاءت الاختبارات لتتماشى

مع تصميم المنهج.

وقد تفسر ضعف تقييم الطلبة للفقرات المتعلقة بالأسئلة الموضوعية بقلة إمام أعضاء هيئة التدريس الذين يدرسون اللغة العربية بكلية التربية بأساسيات تصميم الاختبارات الموضوعية وخاصة أسئلة الاختيار من متعدد ومدى معرفتهم وخبرتهم ببناء جدول مواصفات الاختبار الذي يُعتمد عليه في كتابة أسئلة الاختبارات التحصيلية، حيث يشير صبري والرفاعي [19]، والنور [4]. إلى أهمية بناء جدول المواصفات عند تصميم الاختبار التحصيلي الذي يُبنى وفقاً لمجموعة من المواصفات التي تصف المجال الذي يقيسه الاختبار ويكون إطاراً للحصول على عينة ممثلة من مخرجات التعلم، وبالتالي فإن ضعف معرفة وخبرة عضو هيئة بهذه الجوانب قد يؤدي إلى عدم جودة تصميم الورقة الاختبارية وعدم شمولها لمعايير وخصائص الاختبار الجيد، وبالتالي لا تعد نتائجها مدخلاً جيداً للتشخيص [7]، الأمر الذي يؤدي إلى ضعف في تقييم ووجهة نظر الطلبة لمستوى تطبيق أعضاء هيئة التدريس لكتابة فقرات الاختبارات الموضوعية.

وتتفق هذه النتائج مع ما أشار إليه السيد [9] بضعف قدرات أعضاء هيئة التدريس على صياغة فقرات الأسئلة الموضوعية والتركيز على أسئلة المقال.

كما يمكن تفسير هذه النتيجة بأن أعضاء هيئة التدريس في اللغة العربية لديهم إمام كبير بالجانب الشكلي للاختبار المتمثل في طباعته وإخراجه أكثر من إمامهم ومعرفتهم بالمعايير الأخرى لتصميم الاختبار التحصيلي والمتعلقة بجودة فقرات الاختبار وموضوعيته وصلاحيته وشموله وتمييزه لمستويات الطلبة، وقد يعزى ذلك إلى ندرة التأهيل التربوي لأعضاء هيئة التدريس الذين يدرسون اللغة العربية، وخاصة فيما يتعلق بجوانب القياس والتقويم التربوي علماً بأن معظم هؤلاء الأعضاء تخرج في كليات غير تربوية. وتتفق هذه النتائج مع ما أشارت إليها دراسة سنان [23] بأن مواصفات الاختبار الجيد المتمثل في

مستوى تقييمهم لمعايير تصميم الاختبار التحصيلي الجيد. أشارت نتائج السؤال الرابع إلى أن هناك فروقاً ذات دلالة إحصائية عند مستوى (0.05 α) لمتغير التخصص الدراسي (أدبي، علمي) ولصالح التخصص الأدبي، وكانت الفروق دالة في بعض الأبعاد وهي: المعايير المطلوبة لكتابة فقرات الاختبار، ومعايير شمول أسئلة الاختبار للأهداف وللمحتوى، ومعايير القدرة التمييزية لأسئلة الاختبار لصالح التخصص الأدبي بينما لم تظهر فروقاً دالة في الأبعاد الثلاثة الأخرى.

وهذه النتيجة تشير إلى أن تقييم طلبة التخصصات الأدبية لمستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار الجيد في اختبارات اللغة العربية أعلى من تقييم طلبة التخصصات العلمية، وهذه النتيجة تبدو منطقية بسبب تفضيل طلبة التخصصات الأدبية للغة العربية كونها الأقرب إلى تخصصاتهم الأدبية.

ويعزي الباحث سبب اختلاف تقييم وجهة نظر طلبة التخصصات الأدبية عن طلبة التخصصات العلمية لمستوى تطبيق أعضاء هيئة التدريس في اللغة العربية لمعايير تصميم الاختبار التحصيلي الجيد بالدرجة المتحصلة للطلاب في اختبارات اللغة العربية، إذ أن التحصيل الأكاديمي في اللغة العربية لطلبة التخصصات الأدبية أعلى منه لدى طلبة التخصصات العلمية، ذلك لأن استيعاب الطالب لمقررات التخصصات الأدبية ترتبط بدرجة كبيرة بفهمه وإلمامه باللغة العربية، وبالتالي فإن تحصيله الأكاديمي المرتفع في اللغة العربية قد يؤثر على تقييمه ووجهة نظره نحو اختبارات اللغة العربية، حيث أشارت دراسة المصري ومرعي [22] إلى أنه كلما كان تحصيل الطالب أعلى كان اتجاهه نحو أساليب التقويم أفضل. وتتفق هذه النتيجة مع أشارت إليها دراسة الناجم [24] في وجود فروق في وجهة نظر طلبة القسم العلمي وطلبة القسم الأدبي حول المشكلات المتعلقة بجودة ومعايير تصميم الاختبار.

أشارت نتائج السؤال الخامس إلى أن هناك فروقاً ذات دلالة

(الموضوعية، والصدق، والشمول) متدنية في اختبارات أفراد العينة، كما تختلف هذه النتيجة مع ما أشارت إليها دراسة القضاة وآخرون [21] بأن موضوع تصميم الاختبار الجيد من وجهة نظر الطلبة، قد نال أهمية جيدة، ويفسر الباحث سبب هذا الاختلاف إلى اختلاف البيئة التعليمية للمتعلمين بجامعة السودان للعلوم والتكنولوجيا وكلية المعلمين بابها بالمملكة العربية السعودية.

أشارت نتائج السؤال الثاني بأنه توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha = 0.05$) بين متوسطات عينة البحث تعزى لمتغير الجنس، ولصالح الطالبات الإناث في بعض الأبعاد وهي: بُعد معايير موضوعية الاختبار وصلاحيته، وبُعد معايير شمول أسئلة الاختبار للأهداف وللمحتوى حيث بلغت قيمة ت (2.68)، وبُعد معايير القدرة التمييزية لأسئلة الاختبار، ويفسر الباحث وجود فروق في بعض فقرات البُعد تعزى لمتغير الجنس ولصالح الطالبات الإناث بأن الطالبات أكثر اهتماماً بالدراسة وخاصة في مجال اللغات مقارنة بالطلاب، لذلك فهن يبذلن قصارى الجهد لحضور محاضرات اللغة العربية وتحصيل المعلومات في المقررات المرتبطة بها، وقد يكون ذلك نابع من محاولتهن إثبات ذواتهن في المجتمع الجامعي من خلال أدائهن في اللغة العربية، الأمر الذي ينعكس أثره في تقييمهن ووجهة نظرهن حول اختبارات اللغة العربية.

أشارت نتائج السؤال الثالث بأنه لا توجد فروقاً ذات دلالة إحصائية عند مستوى الدلالة (0.05) بين متوسطات عينة البحث تعزى لمتغير المستوى الدراسي (السنة الأولى، والسنة الثانية).

ويمكن أن يعزى عدم وجود فروق ذات دلالة إحصائية بين متوسطات عينة البحث تعزى لمتغير المستوى الدراسي (السنة الأولى، والسنة الثانية) إلى أنه ليس هناك فارق كبير بين طلبة السنة الأولى وطلبة السنة الثانية من حيث معرفتهم وإلمامهم بجودة الاختبارات التحصيلية نظراً لعدم وجود فروق بينة من حيث الفترة التي قضاها بالجامعة، وبالتالي لا توجد تباين في

تقويم مستوى تطبيق أعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي

- الإحصائية على اختبار شيفيه بين المعدل التراكمي (ضعيف، مرتفع) ولصالح المعدل المرتفع، كما أن هناك فروقاً بين المعدل التراكمي (متوسط، مرتفع) لصالح المعدل المرتفع، بينما لم تظهر فروق بين المعدل التراكمي الضعيف والمتوسط.
- ويعزي الباحث سبب هذه الفروق في تقييم الطلبة ووجهة نظرهم بأن المعدل التراكمي المرتفع يعني ارتفاع التحصيل الأكاديمي للطلاب الذي بدوره يؤدي إلى ارتفاع درجة تقييم الطالب ووجهة نظره حول الاختبارات التحصيلية [22].
- ### 7. التوصيات
- في ضوء نتائج الدراسة يوصي الباحث بالآتي:
- 1- تطوير كفايات أعضاء هيئة التدريس في مجالات التقويم الحديث وخاصة فيما يتعلق ببناء وتصميم الاختبارات التحصيلية في ضوء معايير تصميم الاختبار التحصيلي الجيد، ذلك من خلال إقامة دورات تدريبية وورش عمل متخصصة.
 - 2- مراجعة وتقويم الاختبارات التحصيلية بالجامعة بشكل مستمر للوقوف على مستوى جودتها من حيث شمولها على معايير الاختبار الجيد وبنائها في ضوء جدول مواصفات الاختبار التحصيلي.
 - 3- التفكير في آليات تعزز دافعية الطلبة ورضاهم من المقررات التي يدرسونها وأساليب تقويمها.
 - 4- استخدام طرائق التعليم القائم على الأنشطة والمهارات لإثراء جوانب التفكير العليا والتفكير التحليلي للطلبة.
 - 5- إجراء بحوث أخرى واستخدام مناهج وأدوات بحثية ومتغيرات أخرى لدراسة مستوى تطبيق أعضاء هيئة التدريس لمعايير تصميم الاختبار التحصيلي الجيد.
- ### المراجع
- أ. المراجع العربية
- [1] شحاته، حسن (2001) التعليم الجامعي والتقويم الجامعي بين النظرية والتطبيق، الدار العربية للكتاب.
 - [2] القرشي، خديجة ضيف الله (2010) تقويم الورقة الاختبارية، المؤتمر الإقليمي الثاني لعلم النفس، رابطة
- الإحصائيين النفسية المصرية، 9-49.
- [3] عبيد، وليم وآخرون (2000)، تربويات الرياضيات، مكتبة الانجلو المصرية، القاهرة.
- [4] النور، أحمد يعقوب، (2008)، القياس والتقويم في التربية وعلم النفس، الجنادرية للطباعة، عمان، الأردن.
- [5] الكبيسي، عبد الواحد حميد (2007)، القياس والتقويم تجديبات ومناقشات، دار جرير، عمان.
- [6] أبو علام، رجاء (2005)، تقويم التعلم، دار الميسرة، الأردن.
- [7] القحطاني، سالم (1996)، تقويم أسئلة الاختبارات النهائية، والأسئلة الواردة في كتب الدراسات الاجتماعية للمرحلة المتوسطة في منطقة أبها التعليمية.
- [8] فتح الله، مندور عبد السلام (2000)، التقويم التربوي، دار النشر الدولي، الرياض.
- [9] السيد، يسري (2006)، بناء الاختبارات التحصيلية وتقنيها، جامعة الإمارات العربية كلية التربية.
- [10] خضر، فخري رشيد (2003) الاختبارات والمقاييس في التربية وعلم النفس، دار القلم، دبي الإمارات العربية المتحدة.
- [11] أبوحويج، مروان والخطيب، إبراهيم وأبو مغلي، سمير (2002)، القياس والتقويم في التربية وعلم النفس، دار العلمية الدولية للنشر والتوزيع.
- [12] بعيص، نادية وزيدان، سحر (2012)، مبادئ القياس والتشخيص، دار النشر الدولي، الرياض.
- [13] عوض، عباس محمود (1999)، القياس النفسي بين النظرية والتطبيق، دار المعرفة الجامعية الإسكندرية.

- [14] أبو لبدة سعد (1987)، مبادئ القياس النفسي والتقييم التربوي، ط4، جمعية عمال المطابع التعاونية، عمان، الأردن.
- [15] ملح، سامي محمد (2002)، القياس والتقويم في التربية وعلم النفس، دار الميسرة للنشر والتوزيع والطباعة، عمان، الأردن.
- [16] مراد، صلاح أحمد وسليمان، أمين علي (2005)، الاختبارات والمقاييس في العلوم التربوية والنفسية خطوات إعدادها وخصائصها، دار الكتاب الحديث، القاهرة.
- [17] الخولي، محمد علي (2004)، الاختبارات التحصيلية، دار الفلاح للنشر والتوزيع، عمان، الأردن.
- [18] الطيريري، عبد الرحمن سليمان (1418هـ)، القياس النفسي والتربوي نظريته، أسسه، تطبيقاته، مكتبة الرشد، الرياض.
- [19] صبري، ماهر إسماعيل والرفاعي، محب محمود كامل، (2003)، التقويم التربوي، أسسه وإجراءاته، مكتبة الرشد، الرياض.
- [20] زيتون، حسن حسين (1428هـ)، أصول التقويم والقياس التربوي، المفهومات والتطبيقات، الدار الصولتية، الرياض.
- [21] القضاة، محمد فرحان وآخرون (1429هـ)، مدى تطبيق معايير تصميم الاختبار الجيد من وجهة نظر أعضاء هيئة التدريس والطلبة في كلية المعلمين بأبها، العدد (13)، 302-275.
- [22] المصري، محمد ومرعي، توفيق (2007)، اتجاهات طلبة جامعة الإسراء الخاصة نحو أساليب التقويم، مجلة العلوم التربوية والنفسية. مجلد 8، العدد (1)، 110-92.
- [23] سنان، إيناس محمد (2006)، درجة معرفة ممارسة معلمات الرياضيات في المرحلة المتوسطة لمهارات بناء الاختبارات التحصيلية في مكة المكرمة، رسالة ماجستير غير منشورة، جامعة أم القرى.
- [24] الناجم، سعيد عبد الرحمن (2002) المشكلات التي تواجه طلاب وطالبات كلية التربية بجامعة الملك فيصل في ضوء متغيرات الجنس والتخصص والمستوى الدراسي، المجلة العلمية، جامعة الملك فيصل، مجلد 3، عدد (1).
- [26] نبراوي، يوسف وإبراهيم، يحيى (1984)، اتجاهات طلبة جامعة الإمارات العربية المتحدة نحو الممارسات التربوية في الجامعة، المجلة العربية للبحوث التربوية، مجلد 4، عدد (2)، 63-79.
- [27] محاسنة، إبراهيم محمد (2013)، القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة، دار جدير للنشر والتوزيع، الأردن.
- [28] تايلر، ليونا (1998)، الاختبارات والمقاييس، دار الاسراء للنشر والتوزيع، القاهرة.
- ب. المراجع الاجنبية**
- [25] Greenwald, A. G. (1995): Applying a social Psychology to revel a major flow in student evaluation of teaching. Paper Presented on the annual meeting of the American psychological association, New York City.
- [29] Thorndike, R. M. (2005). Measurement & Evaluation in psychology saddle River. NJ: Pearson Education pinc.

THE EVALUATION OF CRITERIA OF THE GOOD ACHIEVEMENT TEST APPLIED BY THE ARABIC LANGUAGE STAFF AT SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY ACCORDING TO STUDENTS' POINT OF VIEW AND OTHER VARIABLES

AHMED YAGOUB ELNOUR
SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
UNIVERSITY OF JAZAN

Abstract This study aims to identify and evaluate the criteria of good achievement test construction used by staff members of the Department of Arabic language, Sudan University of Science and Technology. It investigated students' point of view taking in consideration the variables of gender, level of education, field of study and GPA "Grade Point Average". The study population constituted 300 students, male and female, in the Faculty of Education during the first term of the year 2011-2012.

Having finished testing the reliability and the validity of the scale designed by the researcher, the study came out with the findings that: the criteria of good achievement test construction used by the staff members is the average. The study also showed that there are some statistically significant differences in students view related to the variable of gender in favor of females. The results show no differences related to the level of education but there are some significant differences related to the field of study in favor of arts students. There are also some significant differences related to the GPA in favor of high GPA.

Key words: Evaluation, achievement test, criteria of good achievement test, point of view.