دراسة تقويمية لنماذج من اختبارات قسم اللغة الإنجليزية (بكليات التربية/جامعة صنعاء)

في ضوء المعايير العلمية لإعداد الاختبارات التحصيلية

د. عايض شريان *

ملخص

تهتم الدراسة الحالية بتشخيص واقع الاختبارات التحصيلية (في قسم اللغة الإنجليزية في كليات التربية) بجامعة صنعاء؛ حيث لم يسبق أن أجريت دراسة (منهجية) من هذا النوع. ولذات السبب قام رئيس جامعة صنعاء بتشكيل فريق علمي لتقييم الاختبارات بالجامعة في مختلف الكليات و التخصصات. وتأتى هذه الدراسة ضمن هذا السياق لتحاول الكشف عن مدى مطابقة الأسئلة في قسم اللغة الإنجليزية بجامعة صنعاء للفصل الأول من العام الجامعي 2004- 2005 لبعض المعايير المتفق عليها أكاديمياً ويتبناها فريق الدراسة بوصفها معايير يجب توافرها في الأسئلة بصورة عامة. وقد حاولت الدراسة الإجابة عن سؤال: مدى مراعاة أعضاء هيئة التدريس في قسم اللغة الإنجليزية في كليات التربية جامعة صنعاء لـ:

1. ما مواصفات الورقة الاختبارية الجيدة؟
2. ما معيار التنوع في الأسئلة؟
3. ما معيار شمول الأسئلة لموضوعات المقرر؟
4. ما معيار التنوع في المستويات المعرفية التي تقيسها الأسئلة؟
5. ما معايير صياغة الأسئلة بحسب أنواعها المتضمنة (المقال، إكمال الفراغ، الصواب والخطأ، الاختيار المتعدد، المزاوجة)؟

للإجابة عن هذه الأسئلة قام الباحث ضمن الفريق بتطوير أداة خاصة بهذه الدراسة لغرض تحليل الاختبارات التي وصلت من بعض كليات التربية (صنعاء: 13 أنموذج، المحويت: 15 أنموذج، أرحب: 4 نماذج (المجموع 32)؛ و 50 ورقة اختبارية تحوي 1149 سؤالاً و تمثل 40 مقرراً). رتبت المعايير بشكل قوائم حول

---

* د. عايض شريان، أستاذ مشارك، قسم اللغة الإنجليزية بكلية التربية، جامعة صنعاء. مؤلف ومحرر لسلسلة الكتب الأدبية التي تضم حوالي 32 كتاباً حتى ألان وتدرس في معظم الجامعات اليمنية. بريد إلكتروني: ayids@yahoo.com.

تنظيم وإخراج ورقة الأسئلة والتنوع بشكل عام ثم أعطيت لكل نوع من الأسئلة قوائم مستقلة: أسئلة المقال، إكمال فراغ، صواب وخطأ، اختيار من متعدد، مزاوجة.

و بعد اتفاق الفريق على القوائم، تم تصنيف الاختبارات إلى أنواعها الرئيسية واختيار النماذج التي يمكن أن تدخل ضمن الدراسة، ثم قورنت أوراق الاختبارات بهذه القوائم التي تحوي المعايير. وأخيراً تم تفريغ البيانات في القوائم المتفق عليها ثم حساب التكرار والنسب المئوية لوصف البيانات، وكشف الجوانب السلبية أو الإيجابيّة في اختبارات اللغة الإنجليزية. فمثلا 79٪ من الأوراق لا تحتوي على تحديد لدرجة الاختبار وهذا يترك الطلبة في حيرة من ما سيحدث في النتائج النهائية. تخلص الدراسة إلى توصيات لتحسين الاختبارات في جامعة صنعاء.

1. 0 المقدمة:

لم تلق الاختبارات التحصيلية الاهتمام المناسب في بحوث التقويم على الرغم من أنها تؤدي دوراً رئيساً في النظام التربوي لمختلف البلدان. فالاختبارات تساعد في معرفة مدى تحقق الأهداف التعليمية، إلى جانب تحسين مستوى تعلم الطلبة من خلال التغذية الراجعة من ناحية، و تحسين مستوى التدريس من خلال التغذية الراجعة التي يحصل عليها عضو هيئة التدريس من نتائج اختباراته للطلبة من ناحية أخرى. كما أنّها تسهم في تشخيص جوانب الخلل في العملية التعليمية- التعلمية، ومحاولة إصلاحها. كما إنّ لها دوراً حاسماً في اتخاذ القرارات التعليمية المختلفة، مثل: نجاح الطالب أو فشله، ومن ثمّ انتقاله إلى مستوى أعلى، ومنحه درجة علمية، وحصوله على مقعد في الدراسات العليا، أو منحة دراسية (العجيلي، 2005: 13- 14؛ أبو علام، 1987: 112- 114؛ عودة، 1985: 28).

وبالنظر إلى أثر تطبيق تلك المعايير من عدمه نجد عند الطلبة عبارات دالّة من مثل: هذا اختبار جيد، ممثل للمادة، عادل، فيه مصداقية، وأخر متحيز، لا يمثل المادة، أو يعتمد بشكل كبير على آخر محاضرة إلخ....

إذن الاختبارات في الجامعات تلعب دوراً كبيراً؛ حيث تحدد نوع العملية التربوية أو العلمية ومن ثم المخرجات التي قد تتوافق مع متطلبات سوق العمل. وفي هذا السياق تأتي هذه الدراسة ضمن التطوير المستمر الذي تسعى إليه جامعة صنعاء مثل جامعات العالم التي تحرص على مراجعة برامجها ومناهجها وأدواتها وأساليبها من أجل تحقيق وظائفها وأهدافها بكفاءة. وفي هذا الصدد قام رئيس الجامعة بتشكيل فريق علمي متخصص (قرار 57 لسنة 2005) لدراسة وتقويم اختبارات الفصل الأول 2004/ 2005- كان الباحث واحداً من أعضائه- لمراجعة الاختبارات في كليات جامعة صنعاء المختلفة؛ لتقويم اختبارات التحصيل التي طبقت على الطلبة في الفصل الدراسي الأول من العام الجامعي 2004/ 2005، والذي آل إلى الدراسة الحالية كواحدة من

دراسات متعددة: العلوم التربوية والنفسية، العلوم (فيزياء، كيمياء، أحياء، رياضيات)، اللغة العربية ، اللغة الإنجليزية، العلوم الاجتماعية (تاريخ، جغرافية، اقتصاد، فلسفة)، العلوم الهندسية، والعلوم الزراعية. واستغرقت هذه الدراسة ما يقارب السنة وتم عرضها ضمن هذه الدراسات في ورشة عمل وطنية (في 1/ 6/ 2006) برعاية رئيس الجامعة؛ حضرها أكاديميون من الجامعات اليمنية.

ويأتي هذا الجهد اتساقاً مع الرؤية التي تؤمن بأنّ الاختبارات هي من أهم حلقات التعليم الجامعي التي تشكل منظومةً متكاملةً، تشمل العديد من المكونات المترابطة: الأهداف، المحتوى، التدريس، الأنشطة، التقنيات والتقويم؛ وذلك لربط مدخلات العملية التعليمية ومخرجاتها؛ بغية معرفة مدى كفاءة الأنشطة التعليمية في تحقيق الأهداف التربوية للبرامج الجامعية. فـ(الاختبار) يعني اتخاذ قرارات تعليمية تؤثر على حياة فئة غير عادية من المجتمع، من حيث النجاح أو الرسوب، الانتقال إلى مستوى أعلى، منح درجة علمية، الحصول على مقعد في الدراسات العليا أو منحة دراسية، كل هذا يضفي على موضوع هذه الدراسة أهميّة خاصّة. علاوةً على أنّ الاختبارات تسهم في معرفة مدى تحقق الأهداف التعليمية، و تشخيص جوانب الخلل في العملية التعليمية؛ للعمل على تطويرها.

والاختبارات قد تكون تحصيلية (**Achievement Tests**) أو مشاريع عمل، أو ملاحظة الأداء، أو بحوث قصيرة، أو ملف إنجاز (**Portfolio**)... الخ. لكنّ الاختبارات التحصيليّة تعدّ من أهم وسائل التقويم؛ لأنها تحدد مصير أجيال. وكشاهد على ذلك نورد حادثةً توضح خطورة هذا الموضوع فقد قيل إن إحدى الطالبات في هذه الكلية قتلت نفسها بعد أن رأت نتيجة إحدى المواد، ولعلّها شعرت أنّ المدرس لم يعطها ما تستحق (وبذا يغدو موضوع التصحيح من المواضيع التي تستحق الدراسة). ونسمع الكثير عمن ترك الدراسة، أو من عانى إلى حد بعيد، بسبب الاختبارات التحصيلية. وفي المقابل، يوجد من الطلبة من ترك معظم الدراسة وطور استراتيجيات اخذ الاختبار: التركيز على الملاحظات التي أخذها طلاب سابقون من المدرس، بالإضافة إلى الاختبارات السابقة. وهذا يعطي فرصةً للنجاح- وربما بمستوى عال- دون جهد وبدون تعلم. من هنا كانت فكرة التقييم لهذه الاختبارات ضرورة ملحة. فالتقييم الصادق والمتثبت هو الذي يتم تخطيطه بعناية ودقة، والتفكير فيه قبل القيام به بفترة. ولكي تؤدي الاختبارات دورها بشكل فعال، لابد أن تتوفر فيها خصائص أساسية منها: الصدق، والثبات، والموضوعية، وسهولة الاستخدام (العجيلي، 2005: 110).

1. 1 مشكلة الدراسة:

قابل الباحث العديد من الطلبة الذين غالباً ما يشكون من أن الاختبارات غير عادلة، تعتمد على التركيز

على نسبة بسيطة من المنهج، أو أن الاختبار كان من المحاضرة الأخيرة... الخ. هذه الشكاوى المستمرة التي غالباً ما تكون بين الطلبة، والتي لا تولى من الاهتمام القدر الكافي- كانت وراء اختيار موضوع هذه الدراسة تأسيساً على الأهمية المركزيّة للاختبارات في حياة الطلبة؛ من حيث كونها العامل الأبرز في تحديد مصائرهم. هذا بالإضافة إلى ملاحظات حول ضعف العديد من الاختبارات التحصيلية، الأمر الذي آثار لدى القائمين على الجامعة التساؤل الآتي:

- ما مدى مراعاة أعضاء هيئة التدريس في كليات جامعة صنعاء للمعايير العلمية لإعداد الاختبارات التحصيلية؟ وفي نفس السياق تتساءل الدراسة الحالية:
- ما مدى مراعاة أعضاء هيئة التدريس في أقسام اللغة الإنجليزية بكليات التربية بجامعة صنعاء للمعايير العلمية لإعداد الاختبارات التحصيلية؟

1. 2 أهمية الدراسة:

من خلال كشوفات الكنترول، يلاحظ أنّ هناك مواد تصل نسبة الرسوب فيها إلى 70٪ ونسمع الكثير عن تأثير نتائج الاختبارات على الطلبة وتغيير مسارات حياتهم، الآمر الذي يعطي أهمية للبحث ومشكلته. وتتميز هذه الدراسة بأنها من الدراسات النادرة التي تحلل الاختبارات في قسم اللغة الإنجليزية بجامعة صنعاء. فهي تساعد في الكشف عن نقاط القوة والضعف في الاختبارات لتعطي صورة واضحة لمتخذي القرار في جامعة صنعاء عن مدى تطابق المنهج الذي ترجوه الجامعة مع المنهج الذي يقدمه المدرسون فعلا و المنهج المكتسب لدى الطلبة كما تبينه الإختبارات كمؤشر للمحصلة النهائية. وتعطي هذه الدراسة فكرة عن المعايير التي يمكن الاستفادة منها عند تصميم الاختبارات. فهي تلفت انتباه الأساتذة في قسم اللغة الإنجليزية بجامعة صنعاء والجامعات الأخرى إلى بعض الجوانب الفنية والأكاديمية عند تصميم الاختبار. وهي بهذا تهدف إلى تقديم مقترحات لتطوير الاختبارات في الجامعة لتحقيق الأهداف المرسومة والمطلوبة من هذه الاختبارات، بوصفها وسيلةً ناجحةً لفحص الطلبة، وإخراج نوعيّات جيدة من الخريجين.

و أمّا من حيث العنصر الزمني لهذه الدراسة، فإنّها تأتي في وقت يتزامن مع قدوم الفريق الدولي لتقييم برنامج القسم والذي يأتي ضمن مشروع تطوير الأداء النوعي وكفاءة التخطيط المؤسسي- في الجامعات العربية، ضمن برنامج الأمم المتحدة الإنمائي[1] من 27/ 4/ إلى 3/ 5/ 2006. [2] حيث تخضع عينة من الطلبة للاختبار ضمن المشروع[3]؛ وذلك لغرض فحص حقل التخصص الجامعي (Major Field Test) في 30/ 5/ 2006. وهو عبارة عن اختبار موضوعي و مقنن من قبل مؤسسة خدمات الاختبارات التعليمية (ETS or

(Educational Testing Services).

و ا لحاجة إلى الدراسة تتلخص- بناءً على رأى الفريق- في النقاط الآتية:

1. تلبية حاجة جامعة صنعاء لتقويم اختبارات التحصيل التي يعدها أعضاء هيئة التدريس بقسم اللغة الإنجليزية لطلبتهم.
2. تعريف أعضاء هيئة التدريس بمعايير إعداد الاختبارات وكيفية الاستفادة منها في تطوير اختباراتهم.
3. قد تفيد نتائج الدراسة في اقتراح برامج تدريبية لتطوير كفاءات أعضاء هيئة التدريس المتعلقة بإعداد الاختبارات وصياغة الأسئلة المتنوعة.
4. إعطاء اقتراحات لتحسين الاختبارات في الجامعة.

1. 3 أهداف الدراسة:

في ضوء التساؤل المطروح في مشكلة الدراسة آنفاً، فإن أهداف الدراسة تتركز في الإجابة عن التساؤلات الآتية:

1. ما مدى مراعاة أعضاء هيئة التدريس لمواصفات الورقة الاختبارية الجيدة؟
2. ما مدى مراعاة أعضاء هيئة التدريس لمعيار التنوع في الأسئلة؟
3. ما مدى مراعاة أعضاء هيئة التدريس لمعيار شمول الأسئلة لموضوعات المقرر؟
4. ما مدى مراعاة أعضاء هيئة التدريس لمعيار التنوع في المستويات المعرفية التي تقيسها الأسئلة؟
5. ما مدى مراعاة أعضاء هيئة التدريس لمعايير صياغة الأسئلة بحسب أنواعها المتضمنة (المقال، إكمال الفراغ، الصواب والخطأ، الاختيار المتعدد، المزاوجة)؟

1. 4 حدود الدراسة : حددت الدراسة عينتها بآخذ نماذج من أسئلة مقررات قسم اللغة الإنجليزية من ثلاث كليات تربويّة في جامعة صنعاء: صنعاء، أرحب، المحويت للفصل الدراسي الأول من العام الجامعي 2004/ 2005.

1. 5 تحديد المصطلحات:

الاختبارات Tests: يقصد بها في عرف الدراسة الحالية الاختبارات التي أعدها أعضاء هيئة التدريس في كليات التربية، جامعة صنعاء للفصل الدراسي الأول من العام الجامعي 2004/ 2005م كاختبارات تحصيلية.

التقويم Evaluation : : التقويم في عرف الدراسة الحالية هو إصدار أحكام على الاختبارات التحصيلية في قسم اللغة الإنجليزية في كليات التربية بجامعة صنعاء وذلك من خلال عملية تحليل الاختبارات المشار إليها

أعلاه، ومقارنتها بمعايير إعداد الاختبارات، لغرض تحسينها مستقبلا.

2. 0 الدراسات السابقة:

هناك الكثير من الكتب والدراسات عن الاختبارات في عدة مجالات، وخصوصا في التعليم العام بالعربية والإنجليزية (احمد شريم و يوسف سوالمة، 2006؛ العجيلي، 2005؛ ملحم، 2005 **Al Baidani, Anastasi, 1997; Gregory, 1997; McIntyre 2000, 2003;** أبو علام، 1987؛ علام، 1986؛ الخطيب وآخرون، 1985؛ عودة، 1985؛ حمدان، 1980؛ زكريا 1975، بركات وموفق الحمداني 1972، **Ten Brink 1974, Ebel 1972**). تناولت هذه الدراسات الصدق، الموضوعية، التنوع للمستويات المعرفية، الشمول لموضوعات المقرر، تنظيم ورقة الاختبار، وضوح الكتابة أو الطباعة، توزيع الدرجة، مراعاة الزمن... إلخ.

وهناك نماذج نظرية متعددة للمستويات المعرفية، تشكل خلفيّةً نظريةً لهذه الدراسة. وهذه المستويات المعرفية تتحقق في البرامج والمناهج الدراسية، وبصورة خاصّة في الجامعة. وقد آلت تلك النماذج إلى فرز أنواع الأسئلة، معتمدةً على مستويات معرفية مرتبطة ببعض النماذج النظرية الشائعة مثل:

- تصنيف بلوم و آخرين (Bloom,1958)، والذي يأخذ مستويات متعددة من مجال المعرفة، متدرجةً وفق تراتب هرمي: المعرفة Knowledge، الاستيعاب Comprehension، التطبيق Application، التحليل Analysis، التركيب Synthesis، التقويم. Evaluation
- جيلفورد (Guilford 1956) يتناول أسئلة الإدراك Reasoning، التذكر Recall، التفكير Convergent، التفكير المتشعّب Divergent، التقويم Evaluation.
- آما تصنيف ويفر وسنسي- (Weaver & Cenci 1960)فيتناول أسئلة الذاكرة Memory Questions والأسئلة المثيرة للتفكير Thought Provoking Questions. تتعدى هذه الأسئلة الذاكرة إلى الفهم الحقيقي للمادة الدراسية وتتطلب إجابة منطقية ومركبة كماً ونوعاً، فيها توضيح وتفسير بعيدين عن التخمين أو التذكر العشوائي، كما أنّها تنمي القدرة على الحكم، والتحليل، والتنظيم والمقارنة، والتنبؤ، والمنطق. (حمدان، 1980: 242- 243).
- تصنيف كارنر (Carner 1963) يركز على الأسئلة الحسية Concrete Questions، و المجردة Abstrac، والإبداعية Creative. وهذه الأسئلة الأخيرة تمثل قمة التعقيد الفكري، ويدخل فيها خصائص من الأسئلة الحسية والمجردة (عودة، 1985: 156)

- تصنيف ساندرز (Sanders 1966) يتشابه مع تصنيف بلوم من حيث أنه يركز على: أسئلة التذكر Memory Questions، التطبيق Application، التحليل Analysis، التركيب Synthesis، التقويم Evaluation إلا أنه أضاف الترجمة Translation والتفسير Interpretation، الذي يقتضي إدراك العلاقة بين الظواهر، والقيم والحقائق.
- آما تصنيف بريمر وبات (Bremer & Bat 1967) فيقسم الأسئلة إلى الأسئلة التجميعية Convergent Questions و التشعبية Divergent (عودة، 1985: 155)
- تصنيف ستال وانزالون (Stahl & Anzlone 1970) يركز على أسئلة التذكر Recall Question، والاستدلال Reasoning والتبرير Rationalization، ثمّ يخلص هذا التصنيف إلى أن غالبية الأسئلة في التعليم هي في غالبيتها تمييزية وتذكرية (الخطيب وآخرون، 1985: 245- 246).

ويمكن سرد العديد من الدراسات في هذا المجال والتي تشير إليها المصادر السابقة- وتم بالفعل مراجعة العديد من الدراسات ضمن الفريق حسب الاختصاصات المختلفة. وكمثال على الدراسات لمادة اللغة الإنجليزية في التعليم العام، قام (البيضاني) Al-Baidhani (2003) بتقييم بعض أسئلة امتحانات الثانوية العامة لمادة اللغة الإنجليزية بالجمهورية اليمنية وحاولت الدراسة وصف تصميم الاختبار، والمدى الذي تغطيه فقرات الاختبارات من 1995 إلى 2001. عمل الباحث على تطبيق الاختبارات على طلاب من المرحلة الثانوية العامة، وحلل الفقرات. ثم أعد إستبانة لموجهي ومدرسي اللغة الإنجليزية؛ لفحص مدى وعيهم بمبادئ اختبار اللغة وتقييم مناقشة النتائج. بعد ذلك أعد الباحث اختباراً موضوعياً وطبقه على مجموعة من طلاب الثانوية العامة. وخلص إلى اقتراحات حول تحسين اختبارات الثانوية العامة لمادة اللغة الإنجليزية بالجمهورية اليمنية.

ورغم أنّ الباحث كان مقيّماً وناقداً لهذه الاختبارات؛ عبر وضع استبيانات للموجهين والمدرسين حول هذه الاختبارات، وإعادة تطبيق هذه الأسئلة على عينة مشابهة للمجموعة التي خضعت لها، ثم تحليل النتائج لهذه الأسئلة وكذا الاستبيانات- إلا أنّ الدراسة لم تركز على المستويات المعرفية المشار إليها في الإطار النظري (التذكر، الفهم، التطبيق، التحليل، التركيب، التقويم... الخ) أو على معايير مثل الشمول، الموضوعية، مباشرة أو غير مباشرة، استقلالية الأسئلة، نسبة مطابقتها مع أسئلة الكتاب المدرسي . وكذلك لم تركز على النواحي الفنية للورقة الاختبارية: الزمن، توزيع الدرجات، الأخطاء الإملائية والعلمية والمطبعية.

3. 0 إجراءات الدراسة :اعتمدت الدراسة وصف المقررات الموجودة في قسم اللغة الإنجليزية بكلية التربية[4] (صنعاء)؛ على أساس أنّ هذا القسم هو الأقدم زمنيّاً، والأكثر تنظيماً ووضوحاً في خطته الدراسية ويعدّ المرجع

لباقي الأقسام.

3. 1 منهج الدراسة : اقتفت الدراسة الحالية المنهج الوصفي التحليلي؛ حيث استخدم أسلوب تحليل المحتوى، الذي يعدّ أحد أساليب منهج الدراسة الوصفية التحليلية.

3. 2 عينة الدراسة : تكون مجتمع الدراسة من الأوراق الاختبارية، التي أعدها أعضاء هيئة التدريس في أقسام اللغة الإنجليزية في كليات ا لتربية (صنعاء، أرحب، المحويت)؛ وجميعها للمقررات التي درسها طلاب تلك الأقسام في الفصل الدراسي الأول من العام الجامعي **2004 / 2005** م. (وقد تم إرسال الاختبارات من الكليات إلى فريق العمل في كلية التربية – صنعاء). بعد فحص الاختبارات التي وصلت إلى (**50**) اختباراً، في تخصص اللغة الإنجليزية، ثم تم اختيار **32** نموذجاً (**64**٪) كعينة ممثلة لمواد قسم اللغة الإنجليزية في كليات التربية- جامعة صنعاء. تشمل هذه النماذج مواد مختلفة في التخصص. أمّا المقررات التي استبعدت فهي مواد متطلبة من أقسام غير قسم اللغة الإنجليزية، وكذلك مواد الباحث؛ لغرض تحقيق عنصر الموضوعيّة في الدراسة.

أمّا النماذج (**32**) الخاضعة للتحليل فهي من كلية التربية صنعاء (**13**)، المحويت (**15**) ارحب (**4**)، وبعض النماذج لم تكتب عليها الكلية التي وردت منها.[5]

وبناء على اتفاق فريق الدراسة فقد تم اختيار العينة بطريقة مرحلية عشوائية؛ بحيث تمثل المستويات الدراسية، من ناحية، وتأخذ بعين الاعتبار تنوع الأوراق الاختبارية، من حيث الطباعة، وشكل الورقة، وتنوع الأسئلة فيها، قدر المستطاع. والجدول التالي يبين توزيع أعداد الكليات، والمقررات، و الأوراق الاختبارية ، والأسئلة التي خضعت للتحليل في إطار تخصص اللغة الإنجليزية:

جدول (**1**): أعداد الكليات، و المقررات، والأوراق الاختبارية ، والأسئلة التي خضعت للتحليل ضمن تخصص اللغة الإنجليزية:

م	الفئة	العدد
1	الكليات	3
2	المقررات	40
3	الأوراق الاختبارية	50
4	الأسئلة	1149

3. 3 أداة الدراسة

لغرض تحقيق أهداف الدراسة، من حيث وصف الحالة الراهنة للاختبارات، وتشخيص جوانب القوة و جوانب الضعف، قام الفريق بتطوير استمارات لتحليل المحتوى كأداة للتحليل تحتوي على معايير مبنية على خبرة

الباحثين، ومستقاة من الدراسات السابقة التي حاولت تقييم الاختبارات، خصوصاً في المدارس مثل دراسة شريم و سوالمة (2006: 4)، التي اهتمت بالاختبارات الموضوعية، وبمعايير من مثل مدى ارتباط الفقرة بالهدف الذي تقيسه؛ من حيث المحتوى والمستوى المعرفي، و صياغة الفقرة بلغة واضحة، ومدى تناسب لغة الفقرة مع مستوى الطلبة، وكذلك ما إذا كانت الفقرة تخلو من أو تحتوي على آية إشارات لفظية للإجابة الصحيحة، وما إذا كانت إجابة الفقرة لا تؤثر في إجابة غيرها من فقرات الاختبار و أخيراً ما إذا كانت ممهوات الفقرة جذابة للطلاب ومناسبة لهم.

وتجدر الإشارة إلى أنّ المعايير قد جمعت من مجالات مختلفة، وأخذت بعين الاعتبار التصنيفات المختلفة لمجالات المعرفية، وكذلك مفردات المنهج، وشمول الأسئلة وموضوعيتها. كان هذا هو السبيل المتيسّر للبحث؛ حيث أنّه لا توجد استمارات جاهزة في الدراسات السابقة التي تم مراجعتها. ولهذا فقد أعد فريق الدراسة أداة لتحليل الاختبارات وفق الخطوات الآتية:

أولاً: اشتقاق معايير إعداد الاختبارات التحصيلية الجامعية من خلال مراجعة الأدب المتخصص في القياس والتقويم التربوي، وكذلك الدراسات السابقة التي تناولت تحليل وتقويم اختبارات التحصيل، وفي هذا الإطار تم إعداد قائمة بمعايير الاختبارات التحصيلية الجامعية، حيثُ روعي أن تكون المعايير عملية وقابلة للتحقيق في الدراسة الحالية وقد اتفق على المعايير الآتية:

1. التنظيم والإخراج: تنظيم الورقة الاختبارية، وخلوُّها من الأخطاء الإملائية أو الطباعية ، واحتوائها على البيانات الأساسية.
2. الشمول: شمول أسئلة الورقة الاختبارية لموضوعات المنهج؛ على أساس أن الشمول يعد إحدى المؤشرات الرئيسة لصدق محتوى الاختبار التحصيلي.
3. التنوّع: تنوع الأسئلة في الورقة الاختبارية، وعدم اقتصارها على نوع واحد. فلما كانت أهداف المقررات الجامعية متنوعة فإنها تتطلب تنوع الأسئلة الاختبارية التي تقيسها؛ لأن هناك أهدافاً يسهل قياسها باختبارات معينة، وهناك أهداف يسهل قياسها باختبارات أخرى، مع الأخذ بعين الاعتبار "معيار" الموضوعية الذي يصعب تناوله بمفرده.
4. المستوى المعرفي: ويقصد به المستوى المعرفي الذي يقيسه كل سؤال، وذلك بحسب النموذج الذي تبناه فريق الدراسة. ولغرض تسهيل عملية التحليل ودقتها، خاصةً وان المحلل ليس واضع الأسئلة ولا يمتلك

الصورة المتكاملة للأهداف المعرفية (المخرجات التعليمية المتوقعة) للمقرر ومحتوياته التفصيلية، من كتب أو ملازم- فقد اتفق أعضاء فريق الدراسة على تبني النموذج الثلاثي الآتي:

أ. المستوى الأول: التذكُّر ويتضمن تذكر الطالب للمعلومات المتعلقة بالمقرر، وهذا المستوى مطابق لمستوى التذكر في تصنيف بلوم لأهداف المجال المعرفي.

ب. المستوى الثاني: الفهم البسيط: ويتضمن قدرة الطالب على توضيح المعلومات التي وردت في المقرر بأسلوب مختلف عما ورد في محتويات المقرر. ومن أمثلة هذا المستوى أن يأتي الطالب بمثال من عنده لمصطلح أو مفهوم ورد في المقرر، أو تعريف مصطلح أو مفهوم بأسلوبه الخاص، أو أن يشرح الطالب مخططا أو رسماً أو أن يبيّن مخططاً لمفهوم ورد بشكل مكتوب. وهذا المستوى مطابق لمستوى الاستيعاب (الفهم) في تصنيف بلوم لأهداف المجال المعرفي.

ج. المستوى الثالث: الفهم العميق: ويقصد به قدرة الطالب على تطبيق القواعد والمبادئ في مواقف جديدة؛ للوصول إلى حل مشكلة مطروحة في موقف جديد. وكذلك تحليل المعلومات المعطاة إلى مكوناتها الجزئية، وإعادة تنظيم أجزاء المعلومات وجعلها في وحدة مترابطة ذات فكرة جديدة، وإصدار الأحكام القيمية على المواقف الجديدة المطروحة في الأسئلة. وبهذا المعنى يشمل هذا المستوى الفئات الأربع الأخيرة من تصنيف بلوم، وهي التطبيق والتحليل والتركيب والتقويم.

5. الصياغة: أي مراعاة قواعد صياغة الأسئلة بحسب أنواعها المختلفة، ودقتها من الناحية العلمية. وهذا يتضمن مراعاة قواعد صياغة (إعداد) كل نوع من أنواع الأسئلة التي ترد في الأوراق الاختبارية الخاضعة للتحليل.

ثانيا: إعداد أداة التحليل: في ضوء المعايير المشار إليها أعلاه، وفي هذا السياق أعدت أداة التحليل التي تألفت من ست استمارات. وقد تضمنت الاستمارة الأولى بنوداً (عناصر) تمثل المعايير الثلاثة الأولى، وهي: التنظيم والإخراج، والشمول، والتنوع، من حيثُ كونها تتعلق بمجمل (عناصر) الورقة الاختبارية. أما الاستمارات الخمس الأخرى فتتضمن بنود معياري المستوى المعرفي، والصياغة والتي ترتبط بكل نوع من أنواع الأسئلة الخمسة الآتية: المقال، إكمال الفراغ، الصواب والخطأ، الاختيار المتعدد، والمزاوجة.

أمّا من حيث التحقق من وضوح بنود استمارات التحليل لضمان دقة التحليل من أعضاء فريق الدراسة، فقد قام كل عضو من أعضاء الفريق بتحليل عينة من الأوراق الاختبارية والأسئلة في مجال تخصصه، ثم نوقشت الملاحظات التي ظهرت من خلال التحليل الأولي في اجتماع حدد للفريق، وفي ضوء ذلك أجريت بعض التعديلات على الاستمارات لتصبح بصيغتها التي اتفق عليها.

3. 4 خطوات التحليل: تم تحليل الأوراق الاختبارية وفق الخطوات الآتية: -

أ.تم تصنيف الأوراق الاختبارية كما هو موضح بالجدول الأتي:

الجدول (2): يوضح الجدول الأتي النماذج المختارة.

رقم المقرر	اسم المقرر	المستوى	نسبة الأسئلة إلى الموضوعات الرئيسة
	قواعد I-Grammar (1)	1	٪70
	قواعد I-Grammar (1)	1	٪60
	قواعد I-Grammar (1)	1	٪80
	لغة الحديث I-Spoken English (1)	1	٪50
	لغة الحديث I- Spoken English(1)	1	٪50
	كتابة I-Writing (1)	1	٪60
	كتابة I-Writing (1)	1	٪70
	قراءة I-Reading (1)	1	٪70
	قراءة I-Reading (1)	1	٪70
	لغة الحديث III-Spoken English (3)	2	٪40
	لغة الحديث III-Spoken English (3)	2	٪60
	قراءة III-Reading (3)	2	٪50
	قراءة III-Reading (3)	2	٪70
	كتابة III-Writing (3)	2	٪70
	كتابة III-Writing (3)	2	٪70
	مقدمة في اللغة Introduction to Language (1)	2	٪30
	قواعد Grammar-Grammar (3) III - I)	2	٪70
	مقدمة في اللغة Introduction to Language (1)	2	٪70
	صوتيات Phonology and Phonetics	3	٪90
	صوتيات Phonology and Phonetics	3	٪60
	شعر القرن الثامن عشر 18th Cent Poetry	3	٪60
	طرق تدريس Methods (1)	3	لا يوجد وصف
	تربية عملية (نظري) Theoretical Practicum	3	لا يوجد وصف
	كتابة متقدمة Advanced Writing Skills	3	٪60
	مسرح القرن العشرين 20th cent Drama	3	٪80
	مهارات الخطاب English Discourse Skills	3	٪60
	مهارات الخطاب English Discourse Skills	3	٪70
	مناهج Curriculum 1	3	٪50
	كتابة متقدمة Advanced Writing Skills	4	٪60
	مسرح القرن العشرين 20th cent Drama	4	٪50
	قواعد تحويلية Generative Grammar-Transformational	4	٪70
	قواعد تحويلية Generative Grammar-Transformational	4	٪40
	المجموع - 32 مقرر		٪60

يلاحظ من الجدول أنّ معظم الأسئلة في المستوى الأول، وتقل في المستويات الأخرى. و غالبية الأوراق الاختبارية في المستوى الثالث. و بعد استبعاد بعض النماذج التي لا توجد ضمن مفردات قسم اللغة الإنجليزية-

كلية التربية صنعاء (من كليات الآداب، الأعلام آو كليات التربية، المواد المختلفة)، تم اختيار نماذج كما هو موضح في الجدول. بيّن الجدول السابق أنّ هناك مواداً بلا وصف، و بعض أوراق الأسئلة لا تغطى المفردات بالشكل المطلوب. يلي هذا تصنيف الأسئلة إلى فروعها الرئيسية.

ب. حساب نسبة الشمول من خلال حساب عدد الموضوعات التي تغطيها الورقة الاختبارية (ككل) وقسمتها على العدد الكلي للموضوعات الرئيسة بالمقرر وضرب الناتج في (100).

ج. تحليل كل ورقة إختبارية بحسب بنود التنظيم والإخراج ورصدها في المواضع المخصصة لها بالاستمارة (1).

د. رصد عدد الأسئلة بحسب أنواعها في كل ورقة إختبارية. ثم حساب عدد الأوراق الاختبارية التي تحوي نوعا واحدا من الأسئلة، وعدد الأوراق التي تحوي نوعين، و كذلك عدد الأوراق التي تحوي ثلاثة أنواع، وتلك التي تحتوي على أربعة أنواع، والتي تحتوي على خمسة أنواع. ثم تحويل هذه الأعداد إلى نسب مئوية؛ لتحقيق معيار التنوع.

ه. ولتحديد المستوى المعرفي الذي يقيسه كل سؤال، فانّه يتم قراءة كل سؤال من الأسئلة الواردة في الورقة الاختبارية بدقة وتحديد الأفكار الواردة فيه، ثم يرصد في المستوى الذي يقيسه وإذا وردت أكثر من فكرة في السؤال؛ إحداها ضمن مستوى وأخرى ضمن مستوى آخر، فإن السؤال يوضع في المستوى الأعلى. وترصد البيانات المتعلقة بالمستوى المعرفي في الاستمارات (2، 3، 4، 5، 6) بحسب نوع السؤال المتضمن في الورقة.

و. ثم تعاد قراءة كل سؤال؛ لتحديد مراعاته لقواعد الصياغة بحسب البنود الواردة في الاستمارة المتعلقة بنوع ذلك السؤال، ورصد ذلك في المواضع المحددة من الاستمارة.

ز. حساب مجموع التكرارات لكل بند من بنود كل استمارة وتحويلها إلى نسب مئوية ورصدها في جداول تسهل عملية عرض النتائج وتفسيرها.

ح. وأخيراً تصنيف الأوراق على حسب المستوى.

و الجدول التالي يبين عدد المقررات، و الأوراق الاختبارية ، والأسئلة التي خضعت للتحليل، في إطار تخصص اللغة الإنجليزية موزعة بحسب المستوى.

جدول (3): عدد المقررات، و الأوراق الاختبارية ، والأسئلة التي خضعت للتحليل موزعة بحسب المستوى

المستوى	عدد المقررات	عدد الأوراق الاختبارية	نسبة كل مستوى في الدراسة	عدد الأسئلة
1	4	9	٪28	313
2	5	9	٪28	90
3	8	10	٪31	87
4	4	4	٪13	37
المجموع	21	32	٪100	539

3. 5 ثبات التحليل:

لغرض التحقق من ثبات تحليل الأوراق الاختبارية والأسئلة المتضمنة لحساب ثبات التحليل قام الباحث بالتحليل مرتين متتاليتين يفصل بينهما فترة زمنية، ما يقارب الشهرين، وتم حساب معامل الثبات باستخدام معادلة كوبر

$$\text{نسبة الاتفاق} = \frac{\text{عدد مرات الاتفاق}}{\text{عدد مرات الاتفاق + عدد مرات}} \times 100$$

وكانت النتيجة 87 ٪ وهي نتيجة مرتفعة.

3. 6 تطبيق الأداة

بعد الاتفاق على قوائم تحتوي على معايير محددة في إطار الفريق، تم تصنيف الاختبارات إلى أنواعها الرئيسية، واختيار النماذج التي يمكن أن تدخل ضمن الدراسة، ثم قورنت أوراق الاختبارات بهذه القوائم التي تحوي المعايير. بعد ذلك استخدمت التكرارات والنسب المئوية، ثم تم تفريغ البيانات في القوائم كما هو موضح في الجداول السابقة. الورقة الأولى (الجدول 4 & 5) تحتوي على معايير تبين تنظيم و إخراج ورقة الأسئلة، و كذلك معيار التنوع.

4. 0 عرض النتائج:

تم استخدام التكرارات والنسب المئوية؛ للإجابة عن تساؤلات الدراسة الستة التي سبق إيرادها. وفيما يلي عرض لتساؤلات الدراسة والإجابة عنها:

4. 1 السؤال الأول: ما مدى مراعاة أعضاء هيئة التدريس لمعايير الورقة الاختبارية الجيدة؟

الجدول (4): التكرار والنسبة المئوية لبنود معيار احتواء الورقة الاختبارية على البيانات الأساسية.

م	البند	تكرار	نسبة ٪
1	تحتوي الورقة على اسم المقرر	32	٪100
2	تحتوي الورقة على اسم المستوى	28	٪87
3	تحتوي الورقة على العام الجامعي	26	٪78
4	تحتوي الورقة على الزمن المحدد للإجابة	26	٪78
	العدد الكلي للأوراق الاختبارية	32	٪100

يبين جدول (4) أن بعض الأوراق الاختبارية لا تحتوي على المعلومات الأساسية التي تعرف بالورقة الاختبارية (مثل العام الجامعي والمستوى)، وقدتم استبعاد بعض الأوراق؛ لأنها لا تحتوي على بيانات للتعرف على مكانها وتاريخها.

الجدول (5): التكرار والنسبة المئوية لبنود معيار تنظيم الورقة الاختبارية وخلوها من الأخطاء.

م	البند	تكرار	نسبة ٪
1	الورقة خالية من الأخطاء	27	٪84
2	توجد فواصل واضحة بين الأسئلة	30	٪94
3	الدرجة الكلية موزعة على الأسئلة	7	٪21
	العدد الكلي للأوراق الاختبارية	32	٪100

تنظيم الورقة الاختيارية وخلوها من الأخطاء يأخذ بعداً مهماً في تقييم الاختبارات. يلاحظ في جدول (5) أن التكرار في المعيار "الدرجة الكلية موزعة على الأسئلة بشكل متوازن" هو 7 من 32 نموذج اختبار (21٪). نسبة 21٪ من الأسئلة توزع الدرجات على الأسئلة هي نسبة بسيطة؛ لأنّ 79٪ من الاختبارات لا تحتوي على توزيع للدرجات، وهذا يدل على عدم الإعداد الجيد والمرتب لهذه الأسئلة. كما أنّ هذا يشير إلى أن المدرس يمكن أن يغير رأيه في التصحيح حسب إجابات الطلبة. وهوا بهذا لا يلزم نفسه بتوزيع؛ يستطيع الطالب من خلاله تقدير درجته النهائية ومعرفة مصداقية الاختبار المصحح. الجدول (6) يبين أعداد الأسئلة المحللة موزعة بحسب أنواعها.

4. 2 السؤال الثاني: ما مدى مراعاة أعضاء هيئة التدريس لمعيار التنوع في الأسئلة؟

وللإجابة عن هذا السؤال تم حساب عدد الأسئلة المحللة بجميع أنواعها، والتي بلغت بمجملها 539 سؤالاً توزعت بحسب النوع إلى خمسة أنواع، كما هو موضح في الجدول الأتي.

الجدول (6): أعداد الأسئلة المحللة موزعةً بحسب أنواعها.

م	نوع السؤال	العدد
1	مقال	141
2	إكمال الفراغ	101
3	صواب وخطأ	36
4	اختيار من متعدد	240
5	مزاوجة	21
	المجموع	539

الجدول (6) يشير إلى أن هناك أنواعاً متعددة من الأسئلة في قسم اللغة الإنجليزية. الاختيار من متعدد، آو الموضوعية يأتي في المرتبة الأولى من حيث التكرار (45٪) ثم الأسئلة المقالية (27٪). وربما يعود استخدام الاختيار من متعدد بشكل واسع إلى انتشار الاختبارات المقننة التي تستخدم هذه النوع بشكل كبير، وتأثر المدرسين بهذا النوع من الأسئلة؛ كونها سهلة التصحيح للمجموعات الكبيرة ويمكن أن تغطي المنهج، وتحقق الشمول. لكن من عيوبها أنها تخضع للمصادفة ، والحظ في نتائجها ، وهذا يقلل من دقتها وثبات نتائجها، حيث أن الطالب يمكن أن يحقق درجات حتى لو انه لا يعرف شيئاً ذا بال في المادة.

لكنّ الأسئلة الموضوعيّة، من جهة أخرى، لها سلبياتها من حيث محدوديتها في قياس قدرات التفكير العالي، التي تتميز بها الاختبارات المقالية؛ وذلك مثل القدرة علي عرض وتنظيم الأفكار، آو حل المشكلات و التفكير الابتكاري، فالمتعلم يقوم بالوصف، والمناقشة، والمقارنة، و التفسير، والتلخيص، والنقد.

الجدول (7): التكرار والنسبة المئوية للأوراق الاختبارية بحسب أنواع الأسئلة التي تحويها.

م	البند	تكرار	نسبة ٪
	تحتوي الورقة على:		
1	أسئلة مقال فقط	141	٪26
2	أسئلة صواب وخطأ فقط	36	٪7
3	أسئلة اختيار متعدد فقط	240	٪45
4	نوعين من الأسئلة	5	٪16
5	ثلاثة أنواع من الأسئلة	2	٪6
6	أربعة أنواع من الأسئلة	1	٪3
7	خمسة أنواع من الأسئلة	0	٪0
	العدد الكلي للأوراق الاختبارية الداخلة في التحليل	32	٪100

يبين الجدول (7) أنّ نسبة أسئلة الصواب والخطأ بسيطة (7٪) لعدد الأسئلة ككل؛ وهذا ناتج عن أن الطلبة في قسم اللغة الإنجليزية يتوقع منهم أن يتعلموا مهارات لغوية يدخل فيها نوع من الإبداع والمشاركة الذاتية، أكثر من التعرف على حقائق وتصحيحها. وتتميز أوراق الاختبار في قسم اللغة الإنجليزية- كلية التربية جامعة صنعاء- بأن جزءاً لا بأس به- (41 ٪)- من الأسئلة تتيح للطلاب التعبير عن آرائهم عند الإجابة، وهذا

يتيح للطالب استخدام عمليات عقلية عليا حسب التصنيفات المذكورة في الدراسات السابقة. و الفقرات الثنائية الإجابة (الصواب والخطأ) تستخدم في قياس قدرة الطلبة على التمييز بين متغيرين للإجابة عنها وتركز على التذكر واستدعاء الحقائق والمفاهيم وهي سهلة البناء وسريعة الإنجاز، تغطي المنهج، وهي كذلك سهلة التصحيح. لكن من عيوبها ارتفاع نسبة التخمين وتشجع على الحفظ والاستذكار، بالإضافة إلى أنها سهلة الغش.

وهي تستخدم لتمييز العبارات الصحيحة من الخاطئة، وتعريف المصطلحات والنصوص والقوانين والحقائق العلمية من الآراء أو النظريات، وتحديد الأسباب وتمييزها، وكذلك تحديد الاستنتاج السليم من غير السليم. ويفضل عند بنائها استخدام عبارات واضحة، وذلك عن طريق تجنب استخدام نفي، وأن تتضمن الفقرة الواحدة حقيقة واحدة والبعد عن الأفكار الجدليّة، وتجنب الفقرة الطويلة، بالإضافة إلى تجنب التضليل والأفكار الهامشية.

4. 3 السؤال الثالث: ما مدى مراعاة أعضاء هيئة التدريس لمعيار شمول الأسئلة لموضوعات المقرر الرئيسية؟

الجدول (8): التكرار والنسبة المئوية لبنود معيار الشمول.

م	البند	تكرار	نسبة ٪
1	تحتوي الورقة على أسئلة اختيارية	21	٪67
2	متوسط (وسيط) نسبة الشمول	1920	٪60
3	الأوراق التي كانت نسبة الشمول فيها أقل من 50 ٪	6	٪19
4	الأوراق التي كانت نسبة الشمول فيها أكثر من 75 ٪	3	٪9
5	المدى الذي وقعت فيه نسبة شمول الأسئلة لموضوعات المقرر الرئيسة	61=1+90-30	
	العدد الكلي للأوراق الاختبارية	50	
	العدد الكلي للأوراق الاختبارية التي حسبت منها نسبة الشمول	32	

الجدول (8) يبين أن نسبة (9٪) من الأوراق كانت نسبة الشمول فيها أكثر من 75 ٪. ولوحظ أن بعض المواد لا تحتوي على وصف، و بعض الأساتذة يخلطون في وصف المقررات.

يلاحظ أن التكرار في المعيار "الدرجة الكلية موزعة على الأسئلة بشكل متوازن" هو 7 من 32 نموذج اختبار (21٪). نسبة 21٪ من الأسئلة توزع الدرجات على الأسئلة هي نسبة بسيطة لان 79٪ من الاختبارات لا تحتوي على توزيع الدرجات؛ وهذا يدل على عدم توفر الإعداد الجيد والمرتب لهذه الأسئلة. وهذا يشير إلى أن المدرس يمكن أن يغير رأيه في التصحيح، حسب إجابات الطلبة. وهو بهذا لا يلزم نفسه بتوزيع؛ يستطيع الطالب من خلاله تقدير درجته النهائية ومعرفة مصداقية الاختبار.

المصحح.

و يلاحظ أن التكرار في المعيار "الدرجة الكلية موزعة على الأسئلة بشكل متوازن" هو 7 من 32 نموذج اختبار (21٪).

4. 4 السؤال الرابع: ما مدى مراعاة أعضاء هيئة التدريس لمعيار تنوع المستويات المعرفية في الأسئلة؟

بعد تحليل الأوراق الداخلة في الاختبار من حيث إعداد الورقة، يتم تحليل الأوراق حسب الأنواع الرئيسية؛ الاستمارة الثانية تحلل الاختبارات المقالية والجدول التالي يبين التكرار والنسبة المئوية لمطابقة معيار المستويات المعرفية التي تقيسها أسئلة المقال.

الجدول (9): التكرار والنسبة المئوية لمطابقة معيار المستويات المعرفية التي تقيسها أسئلة المقال.

م	البند	تكرار	نسبة ٪
1	يقيس السؤال مستوى التذكر	31	٪21
2	يقيس السؤال مستوى الفهم البسيط	85	٪60
3	يقيس السؤال مستوى الفهم العميق	32	٪23
4	يقيس السؤال هدفا يصعب قياسه بالاختبارات الأخرى	112	٪79
	العدد الكلي لأسئلة المقال المحللة	141	٪100

تتميز اختبارات قسم اللغة الإنجليزية بجامعة صنعاء بأنها تتيحُ مجالاً كبير للفهم البسيط (60%) والعميق (79٪) أكثر من التذكر (21%) الذي يتطلب عمليات عقلية بسيطة، علاوة على أنّ أثرهُ محدود في تشكيل شخصية الفرد. فالأسئلة المقالية تفيد في إعطاء الطالب فرصةً للتعبير عن نفسه، وعن الحياة والبيئة من حوله. وتنقسم إلى أسئلة ذات إجابة قصيرة، و أسئلة ذات إجابة طويلة.

ورغم الانتقادات الموجهة إلى الاختبارات المقالية إلا أنها لازال لها استخدام واسع، بالإضافة إلى وجود نواتج مهمة لا يصلح لقياسها اختبارات أخرى؛ وذلك مثل القدرة علي عرض وتنظيم وتكامل الأفكار، والتفسير، و القدرة علي التعبير الكتابي، والتلخيص والإيجاز، والنقد و إبداء الرأي وإصدار الأحكام، والمناقشة والحوار، وكذلك القدرة على حل المشكلات؛ لأن المتعلم يصف، يناقش، يقارن، يعبر وهو بهذا يستخدم مهارات عليا، ومستويات عليا من التفكير؛ من خلال التخطيط للإجابة وعرض الأفكار، وتنظيمها، وربط بعضها ببعض، ومدى تكاملها وانسجامها.

وتتميز– الاختبارات المقاليّة– بأنها تقيس عدة أهداف في وقت واحد مثل التعرف علي معلومات الطلبة، وخبراتهم، ومدى فهمهم واستيعابهم وقدرتهم على الربط بين المعلومات والخبرات وتنظيمها، بالإضافة إلى قدرتهم علي التعبير عنها. ومن عيوبها قلة الأسئلة، مما يجعلها –بصورة غالبة– تخضع للمصادفة ، والحظ في نتائجها

، وهذا يقلل من دقتها وثبات نتائجها، ويسمُّها بالانطباعية في التصحيح؛ فقد تتفاوت الدرجات إلى حد كبير من مصحح إلى آخر، وتأخذ وقتاً كبيراً في التصحيح. و كونها لا تغطي المحتوى مثل الأسئلة الموضوعية فقد يحقق الطالب درجات عالية لأنه ركز جهده على جزء محدد.

4. 5 السؤال الخامس: ما مدى مراعاة أعضاء هيئة التدريس لمعايير صياغة الأسئلة بحسب أنواعها المتضمنة (المقال، إكمال الفراغ، الصواب والخطأ، الاختيار من متعدد، المزاوجة)؟

4. 5. 1 أسئلة المقال

الجدول (10): يوضح التكرار والنسبة المئوية لمطابقة معيار دقة صياغة أسئلة المقال وارتباطها بقضايا جوهرية.

م	البند	تكرار	نسبة ٪
1	السؤال مصاغ بشكل واضح ودقيق وخال من الأخطاء العلمية	136	٪96
2	يحدد السؤال عناصر الإجابة المطلوبة بدقة ووضوح	118	٪84
3	يتناول السؤال قضية جوهرية في محتوى المقرر	112	٪79
4	السؤال مستقل ولا تعتمد إجابته على إجابة سؤال آخر	141	٪100
5	يحتوي السؤال على الدرجة المخصصة له	49	٪34
	العدد الكلي لأسئلة المقال المحللة	141	٪100

يتبين من الجدول (10) أن بعض الأوراق لا تحتوي على تعليمات خاصة توضح طريقة الإجابة. وهذا يسبب إرباك في وقت الاختبار بكثرة الأسئلة من الطلبة ويسبب مشكلة في طريقة التصحيح.

4. 5. 2 إكمال الفراغ

الاستمارة الثالثة خصصت لأسئلـــة إكمال الفراغ. الجدول الأتي يبين نتائج هذه الاستمارة.

الجدول (11): يوضح التكرار والنسبة المئوية لمطابقة معيار المستويات المعرفية التي تقيسها أسئلة إكمال الفراغات.

م	المعيار	التكرار	نسبة ٪
1	يقيس السؤال مستوى التذكر	42	٪14
2	يقيس السؤال مستوى الفهم البسيط	57	٪56
3	يقيس السؤال مستوى الفهم العميق	2	٪09.
	العدد الكلي لأسئلة إكمال الفراغ المحللة	101	٪100

أسئلـــة إكمال الفراغ تمثل نسبة محدودة- 19٪- من حيث عدد الأسئلة ككل وهي أقل من الأسئلة المقالية. و هذا يعطي ميزة لأسئلة قسم اللغة الإنجليزية بجامعة صنعاء؛ لأن هذا النوع غالبا ما يستدعي التذكر، لكن النتائج في الجدول (7) تبين أن الأسئلة أولت اهتماماً كبيراً للفهم أكثر من التذكر. والأسئلة تقيس معلومات

جوهرية في المقرر (90٪).

وأسئلـــة إكمال الفراغ تستخدم لقياس قدرة الطلبة على تذكر الكلمات أو المصطلحات المحذوفة من فقرات السؤال، والتعويض عنها بجمل أو أشباه جمل مفيدة وهي أسئلة ذات إجابات محددة. و تتميّز بسهولة الإعداد والبناء والقدرة على تغطية المنهج، وسهولة التصحيح وإمكانية استخدامها في كافة المراحل الدراسية، بالإضافة إلى تيسر توفير الصور المتعددة للفقرة الواحدة. لكنها بالمقابل، تهتم بالحفظ والاستذكار ويسهل فيها الغش. ويفضل في بنائها أن تحتوي على هدف ذي قيمة وأهمية، وأن تكون الفقرة محددة بدقة، كما يجب استخدام لغة سليمة وسهلة الفهم، و ألا تترك فراغات كثيرة ومتعددة في الجملة الواحدة، وأن يترك فراغ مناسب للإجابة، وان تتجنب التعقيدات أو الكلمات الغامضة.

الجدول (12): يوضح التكرار والنسبة المئوية لمطابقة معيار دقة صياغة أسئلة إكمال الفراغات وارتباطها بقضايا جوهرية.

م	المعيار	التكرار	نسبة ٪
1	السؤال مصاغ بشكل واضح ودقيق وخال من الأخطاء العلمية	101	٪100
2	مكان الفراغ في بداية العبارة	32	٪32
3	مكان الفراغ في وسط العبارة	62	٪61
4	مكان الفراغ في نهاية العبارة	37	٪37
5	كل فراغ يقيس هدفا بذاته	93	٪92
6	السؤال محدد بحيث لا يحتمل أكثر من إجابة واحدة لكل فراغ	93	٪92
	العدد الكلي لأسئلة إكمال الفراغ المحللة		

3. 5. 4 الصواب والخطأ

الاستمارة الرابعة خصصت لأسئلـــة الصواب و الخطأ. والجدول الأتي يبين نتائج هذه الاستمارة.

الجدول (13): التكرار والنسبة المئوية لمطابقة معيار المستويات المعرفية التي تقيسها أسئلة الصواب والخطأ.

م	المعيار	التكرار	نسبة ٪
1	يقيس السؤال مستوى التذكر	10	28٪
2	يقيس السؤال مستوى الفهم البسيط	11	٪31
3	يقيس السؤال مستوى الفهم العميق	25	٪69
	العدد الكلي لأسئلة الصواب والخطأ المحللة	36	٪100

الجدول (14): يوضح التكرار والنسبة المئوية لمطابقة معيار دقة صياغة أسئلة الصواب والخطأ وارتباطها بمعرفة جوهرية.

م	المعيار	تكرار	نسبة ٪
1	السؤال مصاغ بشكل واضح ودقيق وخال من الأخطاء العلمية	36	٪100
2	السؤال يتضمن فكرة رئيسة واحدة	36	٪100
3	يقيس السؤال معرفة جوهرية في المقرر	18	٪50
4	المعلومة التي يقيسها السؤال إما صحيحة تماما أو خاطئة تماما	22	٪61
5	تجنب السؤال استخدام كلمات الإطلاق (مثل دائما، أبدا، ....)	36	٪100
6	تجنب السؤال استخدام كلمات التحفظ (مثل قد، ربما، ....)	36	٪100
7	السؤال مختصر ولا يحتوي على معلومات زائدة	36	٪100
8	طول العبارة مقارب لأطوال العبارات الأخرى	36	٪100
9	تجنب السؤال أسلوب النفي المزدوج (نفي النفي)	36	٪100
	العدد الكلي لأسئلة الصواب والخطأ المحللة	36	٪100

أسئلة الصواب و الخطأ تمثل نسبة بسيطة- (07 %)- من مجموع الأسئلة وتكرارها محدود (36) ضمن الأسئلة الداخلة في الاختبار. الملاحظ أن نسبة (39٪) لا تحتوى على إجابة صحيحة تماماً أو خاطئة تماما.

4. 5. 4 الاختيار من متعدد

الاستمارة الخامسة خصصت لأسئلـــة الاختيار المتعدد. أسئلة الاختيـار المتعـدد هـي الأكثـر تكـرارا (240) وتحتوي على نسبة عالية- 45٪- بالنسبة لعدد الأسئلة ككل. يبين الجدول الأتي التكرار والنسب المئوية.

الجدول (15): يوضح التكرار والنسبة المئوية لمطابقة معيار المستويات المعرفية التي تقيسها أسئلة الاختيار المتعدد.

م	المعيار	تكرار	نسبة ٪
1	يقيس السؤال مستوى التذكر	114	٪48
2	يقيس السؤال مستوى الفهم البسيط	97	٪40
3	يقيس السؤال مستوى الفهم العميق	27	٪11
	العدد الكلي لأسئلة الاختيار المتعدد المحللة	240	٪100

أسئلـــة الاختيار المتعدد مصاغة بشكل جيد، وتكاد تخلو من العيوب كما هو مبين في الجدول (15). يلاحظ على هذه الأسئلة أنها لا تعطي الفهم العميق اهتماماً كبيراً حيث أن نسبة الفهم والتذكُّر تكاد تكون متساوية في هذا النوع من الأسئلة.

الجدول (16): يوضح التكرار والنسبة المئوية لمطابقة معيار دقة صياغة أسئلة الاختيار المتعدد وارتباطها بمعرفة جوهرية.

م	المعـــــيار	تكرار	نسبة ٪
1	السؤال مصاغ بشكل واضح ودقيق وخال من الأخطاء العلمية	229	٪95
2	يحدد السؤال كل ما يحتاجه من معلومات لفهم المشكلة قبل قراءة البدائل	230	٪96
3	أصل السؤال تضمن الجزء الأكبر منه، والبدائل مختصرة	240	٪100
4	يقيس السؤال معرفة جوهرية في المقرر	225	٪94
5	تجنب السؤال أسلوب النفي المزدوج (نفي النفي)	238	٪99
6	السؤال مختصر ولا يحتوي على معلومات زائدة	239	٪99
7	يوجد تناسق لغوي بين أصل السؤال والبدائل	223	٪93
8	أصل السؤال لا يحتوي على كلمات توحي بالإجابة الصحيحة	238	٪99
9	طول الإجابة الصحيحة مقارب لأطوال البدائل الخاطئة	237	٪99
10	في كل فقرة يوجد بديل واحد فقط هو الأصح	237	٪99
11	موقع الإجابة الصحيحة موزع بشكل غير منتظم بين البدائل الأخرى	237	٪99
	العدد الكلي لأسئلة الاختيار المتعدد المحللة	240	٪100

4. 5. 5 أسئلة المزاوجة

الاستمارة السادسة خصصت لأسئلـــة المزاوجـــــة. فهذه الأسئلة هي الأقل تكرارا (21) وتشكل نسبة بسيطة (03. ٪) من الأسئلة. يوضح الجدول الآتي التكرارات والنسب المئوية.

الجدول (17): التكرار والنسبة المئوية لمطابقة معيار المستويات المعرفية التي تقيسها أسئلة المزاوجة.

م	المعـــيار	تكرار	نسبة ٪
1	يقيس السؤال مستوى التذكر	10	٪48
2	يقيس السؤال مستوى الفهم البسيط	10	٪48
3	يقيس السؤال مستوى الفهم العميق	1	٪0.48
	العدد الكلي لأسئلة المزاوجة المحللة	21	

الجدول (18): التكرار والنسبة المئوية لمطابقة معيار دقة صياغة أسئلة المزاوجة وارتباطها بمعرفة جوهرية.

م	المعيار	تكرار	نسبة ٪
1	السؤال مصاغ بشكل واضح ودقيق وخال من الأخطاء العلمية	21	٪100
2	فقرات القائمتين متجانسة (تدور حول محور واحد)	21	٪100
3	فقرات القائمتين مختصرة (لا تحتوي على كلمات زائدة)	21	٪100
4	فقرات الاستجابة اقصر من فقرات المقدمات	21	٪100
5	فقرات الاستجابة اختيرت بحيث تكون إجابة محتملة للمقدمات	21	٪100
6	توجد تعليمات توضح طريقة الإجابة	11	٪52
7	تم تجنب التطابق التام في فقرات القائمتين	21	٪100
	العدد الكلي لأسئلة المزاوجة المحللة	21	٪100

5. 0 مناقشة نتائج الدراسة

اشتملت الدراسة على ست استمارات تدور حول تنظيم وإخراج ورقة الأسئلة والتنوع، الأسئلة المقالية، و أسئلة إكمال الفراغ، و أسئلة الصواب و الخطأ، و أسئلة الاختيار المتعدد، و أسئلة المزاوجة. و يلاحظ من الجدول (3) الذي يوضح نسبة الأسئلة إلى الموضوعات الرئيسية بأن بعض الأسئلة كانت متميزة وتمثل المنهج بشكل كبير. فالأوراق التي تزيد نسبة الشمول فيها عن 50٪ هي كبيرة إذ تصل إلى أكثر من 90٪. البقية (0. 09 ٪) تمثل جزءاً بسيطاً. والمدى (61) يعطي مؤشراً لهذا. يلاحظ على الأوراق التي لم تكن ممثلة للمنهج بشكل كبير أنها ركيكة الإعداد؛ و ربما يرجع السبب إلى أنّ الذي أعدها لم يضمن كل الأجزاء، ونورد هنا مثالاً على هذا النوع من الاختبارات التي لا تغطي المفردات، وهو عبارة عن ورقة اختبار في(طرق التدريس) مكتوبة بخط يد تطرح الأسئلة آلاتية: عرف Method، Techniques، Approach وESL. ثم قارن بين Lingual and direct method -Audio اشرح Grammar Translation Method. هذه المادة ليس لها وصف، ولكن واضح من الأسئلة أنها لا تغطي المفردات التي تؤهل الطالب لأن يكون مدرساً متمكناً، خصوصا أن مادةً كهذه هي أساسية لطالب التربية الذي يؤهل بالأصل لأن يكون مدرساً. أتضح من تحليل الاستمارة الأولى التي خصصت لتنظيم و إخراج أوراق الأسئلة والتنوع أن بعض الأوراق لا يمكن تصنيفها لأية كلية؛ لأنه لا يوجد ما يشير إلى مصدرها. فبعض المدرسين لا يذكر اسمه ولا اسم الكلية. بعض الأوراق لا يوجد عليها أرقام صفحات ولا علامة تشير إلى أن الورقة قد انتهت، وتحتوى مثل هذه الورقة الاختبارية على أسئلة من نحو: Make notes on the following paragraph.

وهذا السؤال يفترض أن يقيس مهارات الكتابة كونه يأتي في ورقة اختباريه تقيس الكتابة. لكنة يخلو من تحديد طريقة الإجابة ويعطى هذا لطلاب السنة الأولى- تخصص إنجليزي. وورقة في الكتابة المتقدمة؛ تطلب من الطالب كتابة خطية وهذا غير موجود في وصف المقرر، وقليل ما يتصل بمستقبل الطالب العملي كمدرس لغة إنجليزية. ويمكن أن تدخل هذه الفقرة في لغة الحديث أكثر منها في مادة الكتابة. وتطلب الورقة من الطلبة إكمال الفراغ؛ كون المادة كتابة متقدمة يتوقع أن تقيس مهارات عليا تعتمد على الإبداع والخلق ولا يتوقع أن يحفظ الطالب شيئاً ويأتي ليكتبه ويملئ الفراغ وهو في المستوى الرابع. وقل مثل ذلك عن مادة أخرى هي (تحليل أسلوب الخطاب)؛ حيث يعطى الطالب مطلع رواية ويطلب منه كتابة خمس جمل في نفس الموضوع، فالقطعة غير معروفة المصدر بالنسبة للطالب والتجربة ليست قريبة من الحياة التي يمكن أن يواجهها الطلبة في الميدان ويخوض غمارها في قابل الأيام، وتوزيع الدرجات على الأسئلة يبين أن بعض المدرسين لا يولي اهتماماً للاختبارات أو التقويم التكويني ويعطي اختباراً نهائياً من 150 (الدرجة الكاملة). وبدلاً عن تدريس مسرحيتين أو ثلاث يدرس مسرحية ويعطي عليها كل الأسئلة. هناك أوراق اختبارية بها أرقام في الصفحة الأولى على قطعة للقراءة، و الورقة الثانية بها قطعة وأرقام بنفس الطريقة. هذا يعني أن المصحح لن يستطيع التمييز بين الأرقام في الورقة الأولى والثانية لأنها تخلو من التسلسل (نموذج 9). وعدد الأسئلة محدود (13) فقرة موضوعية لا تحتوي على ورقة تغطية تساعد المصحح ولا على مكان لوضع الإجابة. وتطلب من الطالب الإشارة إلى معلومة بسطر محدد برقم دون ترقيم أسطر الفقرة. ومما يتصل بهذا الموضوع بسبب أنّه يتم اختيار قطع للقراءة لا تنم عن صلة بثقافة الطالب الذاتيّة التي تتصل بمشربه الأصلي. فالنموذج (8) يحتوى على قطعة للقراءة تتحدث عن الهندوسية. ولأنّ القطعة تحاول أن تقيس مهارات في القراءة فإنّها يمكن أن تؤخذ من واقع الطالب فتحتوى مثلا على معلومات يمكن أن يواجهها الطالب في الحياة اليومية؛ وذلك بطبيعة الحال أفضل من قطعة قد لا يواجهها إلا لماماً. وهذا يجعل الدراسة ذات صلة بحياة الطالب وبيئته ويعزز حب التعلم عند الطالب عندما يجد انه يدرس شيئاً ذا صلة بواقعه وبمشربه الثقافيّ. بعض المدرسين لازال يستخدم خط اليد في إعداد الأسئلة وهذه ظاهرة وأن كانت بسيطة (19٪) إلاّ أنها ملفتة للانتباه؛ كونها موجودة في عصر المعلومات والثورة الإلكترونية إذ تنم عن أن بعض الأساتذة لم يندمجوا في سياق هذه التغيرات مما يتسبب في انفصام نفسي بينهم وبين طلابهم الذين لا يفتئون يتابعون لكل جديد في إطار هذه الموجة الثقافيّة الجديدة. و ربما احتاجوا إلى إعادة تأهيل ليتكيفوا مع العصر وإفرازاته المستمرّة. أما آثرها على الاختبار فلاشك بأنها مربكة خصوصاً إذا كانت تحوي اختباراً موضوعياً يحتاج إلى الكثير من التفاصيل (مثال على هذا النموذج رقم3).

6. 0 الخلاصة

يلاحظ آن بعض أوراق الاختبار لم تراعِ شروط الورقة الجيدة. مواد المهارات اللغوية مثلا بها تداخل كبير في الأسئلة التي يمكن آن تقيس مهارات أخرى؛ فمثلا النموذج 4 يفحص مهارات الحديث (Spoken English)، لكن هناك أسئلة تطلب تصحيح جمل ووضع علامات الترقيم المناسبة (Punctuation Marks)؛ لتصحيح الجمل. وواضح بأنّ هذا من مهارات الكتابة. وفي ذات السياق تكون بعض الأسئلة عبارة عن مقالة مأخوذة من صحيفة والمطلوب الإجابة(بصح آو خطأ) على الأسئلة لفحص الفهم ومطابقة هذا مع القطعة. و واضح أن هذا النوع من الأسئلة يقيس مهارة القراءة (Reading). وأسئلة أخرى تطلب من الطالب رسم خريطة وكتابة رسالة وتلك هي الأسئلة التي تتعلق بمادة تقيس مهارات الحديث (Spoken English)، و تضيف اكتب خمس جمل: Write five sentences about shops and facilities in your area.

ربما تعود طبيعة هذه الأسئلة إلى آنّه لا يوجد وصف واضح للمقرر في خطة القسم، واستتباعاً لذلك فإنّ هذا يستدعي مراجعة وصف المقرر ووضع خطوط عريضة للفصل بين المهارات اللغوية، ووضع تدرج لها؛ حيثُ يلاحظ أنه لا يوجد وصف يحدد الفرق بين مهارات الحديث Spoken 1, 2, 3 and 4 أو مهارة القراءة Reading 1, 2, 3, 4 مثلا. فما يدخل في محتوى Spoken 1 أو Reading 1 يمكن أن يدخل في Spoken 2 أو Reading 2. وهذا واضح من النماذج الاختبارية المتضمنة في التحليل.

بعض الأوراق الاختبارية طويلة و بدون ورقة تغطية، مع العلم بأنّ ورقة التغطية ضرورية للأسئلة الطّويلة (نموذج 4) و متنوعة. ووجود أكثر من نموذج A and B يقلل من إمكانية الغش لكن هذا الخيار غير موجود في كل الأوراق المضمنة في هذه الدراسة.

في مزج الأسئلة المقالية والموضوعية آو مزج أكثر من نوع من الاختبارات (مقال، إكمال فراغ، صواب وخطأ، اختيار من متعدد... الخ) ما يجعل الاختبار أكثر قوة؛ بالقدرة على التميز بين المفحوصين وهذا قليل في هذه الدراسة. بعض الأسئلة تفحص الهدف أكثر من مرة بأشكال مختلفة والقدر الكافي في هذا الصدد هو الفحص لمرة واحدة، لكنّ بعض الأوراق تعيد نفس الهدف. وبناءً على المعايير فالاختبار يحتاج إلى تحديد الغرض من الاختبار؛ أي الأهداف التي سوف يقيسها الاختبار، وكذلك تحديد الزمن المناسب، وإعداد جدول مواصفات الاختبار، ووضع تعليمات واضحة... إلخ.

وكتابة أسئلة الاختبار تحتاج إلى التقيد بجدول المواصفات، كما تحتاج إلى ربط السؤال بشكل مباشر مع النتائج التعليمية المراد قياسها، ويجب أنّ تتحاشى بعض المحاذير وذلك من خلال؛ تجنب المصطلحات أو الكلمات

غير المحددة، تجنب الكلمات التي تقترن عادة بالتعميمات مثل: دائماً، أبداً، جميعا، ومن الأمور المهمة التي يجب توفرها، صياغة الأسئلة بشكل واضح يتناسب مع مستوى الطالب، وان يكون طول السؤال مناسباً، وان يكون السؤل مستقلاً؛ بحيث لا تعتمد إجابته على سؤال آخر، كما يجب إسناد العبارات التي تتعلق بالرأي إلى مصادرها؛ لأن عبارات الرأي ليست صحيحة أو خاطئة تماما، ومما يجب على واضع الأسئلة توخيه عند وضع الأسئلة – حسب رأي الفريق– الابتعاد عن إتاحة فرصة الاختيار من بين مجموعة لأن البعض قد يستبعد بعض أجزاء المنهج. ومن الأمور المهمة في هذا المنحى، مراجعة الأسئلة بعد كتابتها وتدقيقها بصورة مستمرة. على أنّ هذه المعايير تختلف طبقاً لنوع الاختبار.

6. 1 التوصيات والمقترحات

تخلص الدراسة إلى أنّ الاختبار الجيد يأخذ في الاعتبار الأهداف المتوخاة منهُ، كما يراعي المفردات اللغوية لدى المتعلم، بالإضافة إلى الموازنة بين الأسئلة والمهارات اللغوية المطلوب اختبارها، مع وضوح الصياغة والتعليمات، وتوفر خاصيّة التنوّع؛ بحيث يعطي فرصةً للطالب أن يبين قدراته ويقلل نسبة الصدفة، ويسهم في توفر عنصر الموضوعية في التصحيح، و أن يكون سهل التطبيق بصورة تخلو من التكلُّف، ومجرب على عينة من الطلبة قبل التطبيق.

كون الاختبارات تقيس نواتج متنوعة من التعلم فينبغي أن تكون منسجمة مع الأهداف والمحتوى. وينصح بصياغة مخرجات التعلم بعد صياغة الأهداف، وتحديد المحتوى، و طرق التدريس، وكذلك طرق التقييم. ولذلك فان الاختبار الجيد يأتي من تحديد نواتج التعلم (ILOs or Intended Learning Outcomes)، بحيث يقيس عينة ممثلة للمادة التي تحويها موضوعات المقرر من خلال أسئلة واضحة، متنوعة، متدرجة، دقيقة، قادرة على التمييز بين مستويات الطلبة، مناسبة من حيث الزمن، وموضوعية في التفسير والتصحيح، بعيدة عن التعقيد اللغوي وتسمح بإبداء رأي المتعلم، محددة ولا يوجد بها خيارات كثيرة. وأسئلة الاختبار الجيد لا توحي بالإجابة آو أن تعتمد إجابة سؤال على إجابة سؤال آخر. ومن خلال العرض السابق توصي الدراسة بـــ:

- عقد دورات تدريبية لأعضاء هيئة التدريس لاقتسام اللغة الإنجليزية في مختلف الكليات؛ لمناقشة مختلف أساليب التقويم المختلفة وان لا تقتصر على قياس مستوى التذكر والفهم بل تشمل باقي المستويات من تطبيق وتحليل وتركيب وتقويم حسب التصنيفات المختلفة.

- عرض الاختبارات على لجنة اختباريه في القسم تراعي أن يشتمل الاختبار على جميع مفردات المقرر، وأن تشتمل الورقة على تعليمات واضحة ودقيقة، تراعي معايير صياغة الأسئلة من تنظيم وإخراج وتنوع ووقت مناسب، بالإضافة إلى مراعاة الفروق الفردية بين الطلبة.
- الأخذ في الحسبان التنويع في أساليب التقويم، بحيث لا يقتصر على الاختبارات بل تستخدم الملاحظات والمقابلات والاستبيانات و تنفيذ المشاريع.
- إجراء بحوث ذات علاقة بتصحيح الاختبارات.
- مناقشة الاختبارات في مجالس الأقسام لتبادل الآراء حول كيفية بناء الاختبارات (وبالذات عند التحاق مدرسين جدد بالقسم) في كل مادة لتحديد مواصفات عامة للورقة الاختبارية وأن توزع قائمة بمعايير صياغة الاختبار على الأساتذة قبل بناء الاختبارات بوقت كاف، و تكون كبداية لإلزام كل أستاذ بتقويم مفردات مادته وإعادة بناء هذه المفردات في صورة مقرر، له أهدافه ومحتواه وأساليب عرضه وأنشطته وأساليب تقويمه.

المراجـــع

أبـو عـلام، رجـاء محمود(1987). *قياس وتقويم التحصيل الدراسي*، الكويت دار القلم للنشر والتوزيع.

. بركات، محمد خليفة وموفق الحمداني (1972). دليل مبسط عن تحسين أسئلة الامتحانات، بغداد مركـز البحـوث التربويـة والنفسية.

حمدان، (1980). *تقييم الطالب، أسسه وتطبيقاته*، بيروت، دار العلم للملايين.

الخطيب، أحمـد وآخـرون (1985). *دليل البحث والتقويم التربوي*، عمان، دار المستقبل.

زكريا، جيرولد، (1975). *"الاختبارات المدرسية عون أم عقبة"* . *مستقبل التربية*، العدد الأول، مكتب اليونسكو الاقليمي، القاهرة، مطابع الهيئة المصرية العامة للكتاب.

شريم، احمد و سوالمة، يوسف (2006). *"تحيد درجة لاختبار محكي المرجع في الرياضيات باستخدام نموذجي "انجوف" و "بدلكسي": دراسة مقارنة بمعرفة صعوبات الفقرات وعدم معرفته"* . مستل من رسالة ماجستير. المجلة الأردنيـة في العلوم التربوية (جامعة اليرموك). مجلد (2)، عدد (1) 2006، ص ص 1- 10.

العجيلي، صباح، (2005). *القـياس والتقويم*. اليمن. مركز التربية للتصوير والطباعة والنشر.

علام، صلاح، (1986). *تطورات معاصرة في القياس النفسي والتربوي*. الكويت، القبس التجارية.

عودة، احمد (1985). *القـياس والتقويم فـي العملية التدريسية*، اربد دار الأمل.

ملحم، محمد (2005). القياس والتقويم في التربية و علم النفس. اربد. دار المسيرة.

Al- Baidhani, Abdussalam (2003). *Critique of Secondary School English Language Examinations from 1994- 1995 to 2000- 2001*. Unpublished MA thesis, Faculty of Arts, Sana'a University.

Anastasi, A. & Urbina, S. (1997). *Psychological Testing*. New Jersey: Prentice- Hall.

Bloom, B. *etl.* (1958). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Longman. New York. (This is the 25th printing 1982 by Longman Inc.)

Brown, J. and Hudson, T. (2003). *Criterion-referenced Language Testing*. Cambridge: Cambridge University Press. (Reviewed in TESL_EJ, Vol. 7. No. 2, September 2003)

Ebel, R. (1972). *Essentials of Education Measurement.* Englewood cliffs, N. J.

Gregory, R. (1997). *Psychological Testing*. Boston: Allyn & Bacon.

McIntyre, S. & Miller. A. (2000). *Foundations of Psychological Testing*. Boston: Mc Graw Hill.

Ten Brink, T. (1974). *Evaluation: A Practical Guide for Teachers*. New York, Mograw-Hill.

الهوامش

1) كان الباحث هو ممثل القسم والمنسق لهذا المشروع وقد تم تدريبه- على التقييم وضبط الجودة والإشراف على الاختبار- في تونس (12-16/ 12/ 2006) والأردن (18-20/ 2006) ضمن فريق من كل الجامعات العربية قام بتدريبهم هذا المشروع؛ ليقوموا بالتنسيق بين القسم وإدارة المشروع.

2) The project entitled "Enhancement of Quality Assurance and Institutional Planning at Arab Universities", is sponsored, and supervised by the Regional Bureau of Arab States of the United Nations Development Programme (UNDP/ RBAS). It was launched in January 2002 and it completed until now two phases that are executed by The United Nations Office for Project Services (UNOPS). The researcher was acting as a UNOPS consultant and as the Subject Representative for a project to organize external review mission to his department.

3) أعلى 50 طالب في دفع سنة رابعة .تم ترتيب طلاب السنة الرابعة من الأعلى إلي الأسفل حسب معدلاتهم التراكمية وتم اختيار أعلى خمسين طالب- كان منهم 49 طالبة وطالب واحد؛ وهذا مؤشر يستحق الدراسة.

4) كلية التربية-صنعاء هي الكلية الأولى للتربية في الجامعة (بها 12 قسماً و 110 أستاذة) و7390 طالبا وطالبة حسب إحصائية 2004/ 2005م_ الأقسام الأدبية 4379 (59٪)، الأقسام العلمية 3011 (41٪) ونسبة الطالبات في الأقسام الأدبية (60٪) وفي قسم اللغة الإنجليزية (85٪) ويتوقع آن يتخرج منها (2540) لعام 2005.

5) من الصعوبات أن بعض الاختبارات لم تصل من الكليات. بالإضافة إلى مشكلة أخرى و هي أن بعض الزملاء لا يسمح بان يحمل الأسئلة متوفرة حتى لا تتسرب.