

تقويم بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN)
على مستوى المدرسة الثانوية الإسلامية للسنة الدراسية 2014-2015 م
(في المدرسة الثانوية الإسلامية الحكومية 3 مالانق)

Rizka Anugerah

rizka.irtadhoni@gmail.com

ABSTRACT

UAMBN is one of exam which held by the ministry of religion for all students in grades 3 MTS and MA. This exam is one of the requirements that must be met students before they graduate. This test also aims to determine the quality of a madrasah, madrasah whether it has a good quality or bad quality. Therefore UAMBN exam has a very important role to improve the quality of learning and education in madrasah. Because this test as a benchmark to determine the success in learning Arabic after knowing the value of the students in the exam, then the exam should have a matter of good quality. Therefore in this study the researchers wanted to analyze further the quality level items owned by UAMBN. And The result of this study is UAMBN 2015 have quality content validity was good, and the quality of the items that are less good, and the quality of reliability items were good, and the quality of power difficulties were medium, and quality of distinguishing power were good, and the quality of deception power less well.

Keywords: *Item Evaluation, language test, UAMBN*

أ. مقدمة

القياس هو عملية نظامية لتعيين القيمة على موضوع معين، القياس في التعليم يتعلّق وثيقاً بالاختبار. لأن الاختبار هو أداة المستخدمة كثيرة لقياس كفاءة الطلبة المحصولة. ماعدا الاختبار، تستخدم غير الاختبار لقياس كفاءة الطلبة في المجال العاطفي (عبد الحميد، 2013: 7).
و الاختبار هو مجموعة الأسئلة التي يطلب من الدارسين أن يستجيبوا لها بهدف قياس مستواهم في مهارة لغوية معينة وبيان مدى تقدمه فيها ومقارنته بزملائه (رشدي أحمد طعيمة، 1989: 248). إذا، كان الاختبار هو أمر مهم و واجب أداءه في عملية التعليم والتعلم.
تعتبر أيضاً بأن الاختبارات هي وسيلة من وسائل التقويم التربوية الشائعة و المهمة لتقويم عملية التعليم و العناصر المؤثرة في نجاح التعليم، إما من مجال طريقة التعليم المستخدمة أو مادة

التعليم وغيرهما. ويستخدم اختبار أيضا لقياس ولتقويم قدرات الطلبة، ومعرفة مدى مستواهم التحصيلي بعد اشتراكهم عملية التعليم والتعلم. فلذلك كان الاختبار جزء مهم في عملية التعليم، لا يوجد التعليم الناجح بدون التقويم، فكيف نعرف نجاح التعليم إذا كان لانعرف نتائج التعليم، ولن نعرف نتائج التعليم إلا بالتقويم يعني بوسيلة الإختبار أم بالتقويم بغير الاختبار.

من أدوات التقويم في تعليم اللغة العربية هي الاختبار. وضع اختبارات اللغة يوافق أهداف التعليم ومحتوياته وتركيز التعليم ذاته، وهذا الاختبار هو الذي يستخدمه المعلمون أو المحاضرون غالبا. لأنه مرتبط بالواجبات التعليمية، وهو إلقاء التقويم والقيمة لمعرفة نتائج التعلم من الطلاب.

الاختبار النهائي الوطني للمدرسة الإسلامية نسّميه ب (UAMBN) هو إحدى أشكال الاختبار الذي عقده وزارة الشؤون الدينية لجميع الطلبة في المراحل الأخيرة من المدرسة المتوسطة أو المدرسة الثانوية الإسلامية.

منذ سنة 2010/2009 عقدت وزارة الشؤون الدينية الاختبار النهائي الوطني ليكون شرط التخرج للطلبة، لذلك كانت أسئلة اختبار العلوم الدينية واللغة العربية صممتها فرقة خاصة من وزارة الشؤون الدينية.

وبوجود الاختبار النهائي الوطني للمدرسة كان واجبة الطلبة المدرسة الإسلامية تكون أثقل من طلبة للمدرسة العامية. لأن لقياس الكفاءة المحصولية للطلبة مطابقة بمعايير كفاءات المتخرجين وعلى جميع الطلبة أن يشتركوا الاختبار النهائي الوطني للعلوم العامية (كعلم الطبيعة و علم الفيزيا و علم الرياضية) وأن يشتركوا الاختبار النهائي الوطني للمدرسة الإسلامية للعلوم الدينية (كعلم القرآن والحديث و علم الفقه و علم تاريخ الإسلام و علم اللغة العربية).

وكانت أهداف تطبيق الاختبار النهائي الوطني للمدرسة هي تقوية المميّز تربية وتعليم في المدرسة الإسلامية، لأجل ذلك الشيء فيكون أمر مهمّا لعقد تقوية تعليم اللغة العربية وتعليم العلوم الدينية في جميع المستوى الدراسي في المدرسة الإسلامية.

يستخدم هذا الاختبار لتعيين نوعية المدرسة، هل هذه المدرسة لها نوعية جيدة أو لها نوعية غير جيدة. هذه نوعية نعرفها بالاختبار الشامل النهائي. لذلك هذا الاختبار الشامل النهائي له دور مهم لترقية نوعية التعليم والتعلم في المدرسة.

من مواصفات الاختبار الجيد هو الشمول، والمراد بالشمول هو كان للاختبار أسئلة مناسبة بالمادة والمهارات اللغوية المدروسة في المدرسة. بعد أن لحظت الباحثة قليلا كان مضمون بنود الأسئلة لهذا الاختبار غير متساو بين عناصر اللغة والمهارات في اللغة العربية، كان بعض الأسئلة يدل كثيرا إلى عدّة مهارات اللغة أو عناصر اللغة فقط. يتكون هذا الاختبار بخمسين سؤالاً، وكان بنود الأسئلة تتكلم كثيرا عن القواعد تعني خمسين في المئة. وأما بقيّة من الأسئلة تتكلم عن مهارة القراءة ومهارة الكتابة والمفردات والترجمة. تعني: مهارة القراءة 3 أسئلة، مهارة الكتابة 5 أسئلة، مفردات 11 سؤالاً، ترجمة 3 أسئلة، قواعد 26 سؤالاً، والثقافة سؤالين.

من ذلك البيانات القليلة عرفت الباحثة بأنّ هذا الاختبار له نقصان، لذلك أردت الباحثة أن تحلل بنود الأسئلة من ناحية الأخرى بأعماق لمعرفة النقائص الأخرى في بنود أسئلة الاختبار، من النقائص المحلّلة فنعرف جودة بنود الأسئلة الموجودة. عسى يكون هذا التقويم مداخلات واقتراحات لتصميم بنود الأسئلة لاختبار الآتي حتّى يكون هذا الاختبار اختباراً معيارياً جيّداً في هذا الوطن.

ومن البيان المذكور، أرادت الباحثة أن تحلل جودة اختبار اللغة العربية الشامل الوطني في السنة الدراسية 2014-2015م. لأن هذا الاختبار يكون مقياساً لمعرفة نجاح تعليم اللغة العربية في المدرسة الثانوية ولتحسين عوامل التعليم في المدرسة بعد معرفة نتائج الطلبة من هذا الاختبار. إذا لبدّ على هذا الاختبار أن يستحق مواصفات الاختبار الجيّد. والاختبار يكون جيّداً إذا كان له درجة الصدق و الصعوبة والتمييز والتشّت جيّدة لكي يستطيع أن يقيس كفاءة الطلبة بالصدق.

تلك الأشياء تكون مهمّاً لتصميم بنود الأسئلة الجيّدة لأن بدرجة الصدق الجيّدة يستطيع الاختبار أن يقيس كفاءة الطلبة بالصدق، وبدرجة الصعوبة الجيّدة يستطيع الاختبار أن يدافع الطلبة على إجابة الأسئلة كلها، وبدرجة التمييز الجيّدة يستطيع الاختبار أن يفرق بين الطلبة الماهرون والطلبة غير الماهرين، وبدرجة التشّت الجيّدة يستطيع الاختبار أن يعطي التحديات إلى الطلبة عند إجابة الأسئلة.

ولقصر الطاقة والوقت، قامت الباحثة بالتحليل في هذا البحث لإحدى المدرسة الثانوية الإسلامية فحسب وهي المدرسة الثانوية الإسلامية الحكومية الثالثة ، اختارت الباحثة هذه

المدرسة لأن هذه المدرسة من مدارس النموذجية في جاوى الشرقية، لها هممة عالية في تنمية اللغة العربية والإنجازات الكثيرة في مسابقات اللغة العربية. ولها معهد للسكن الطلبة لتأثير الطلبة على التكلم باللغة العربية بوجود البيئة اللغوية هناك.

واختارت الباحثة الطلبة الفصل الثاني عشر في قسم علوم الطبيعة ولا قسم اللغة أو قسم علوم الدينية لأن من هذه ثلاثة أقسام الموجود في هذه المدرسة، من الذي يستحق أضعف الكفاءة في تعلم اللغة العربية هو قسم علوم الطبيعة. إذا بهذا الاختبار أردت الباحثة أن تعرف ما مدى كفاءة الطلبة في قسم علوم الطبيعة في اللغة العربية.

ب. التقويم

1. مفهوم التقويم

التقويم هو بيان قيمة الشيء و هو يعني الإصلاح بعد التشخيص. و قد قال رسول الله (صلى الله عليه وسلم): رحم الله امرءً قوم لسانه (أي أصلح اعوجاجه) (رائد خليل العبادي، 2005: 9).

التقويم هو عملية إصدار أحكام و الوصول إلى قرارات، وذلك من خلال التعرف على نواحي القوة والضعف فيها، وعلى ضوء الأهداف التربوية المقبولة بقصد تحسين عملية التعلم والتعليم (أنسى محمد أحمد قاسم، 2003: 262).

وبهذا المعنى فإن التقويم يصبح عملية مستمرة وشاملة ولا تقف عند مجرد إعطاء درجة أو تقدير، وإنما ترتبط بإصدار أحكام على ضوء أهداف معايير محددة.

لذلك كان تقويم مهمّ جدًّا لترقية جودة التعليم والتعلم. بالاختبار نستطيع أن نرى الأهداف المحسولة لنجاح الطلاب بعد اشتراكهم في عملية التعليم والتعلم. وبتنتائج الاختبار نستطيع أن نقوم بالتشخيص والعلاج لجوانب الضعف من ناحية الطريقة المستخدمة أو المواد المدروسة والأشياء الأخرى التي تأثر كثيرا لنجاح التعليم والتعلم.

2. أنواع التقويم

لقد صار التقويم عملية تخصصية في التربية، يتم تدريب أفراد متخصصين لكي يصبحوا مقومين (قائمين بالتقويم). ولقد أصبح التقويم جزءا هاما وركنا، أساسيا للعملية التعليمية وتطويرها وتحسينها.

وعموما، كان موضوع التقويم نجد هناك أغراض مختلفة للتقويم، وهنا يختلف التقويم تبعا لهذه الأغراض. ويمكننا هنا أن نحدد أنواع التقويم على النحو التالي:

الأول التقويم القبلي: مثلما يحدث في الإجراءات التي تسبق دخول الطالبة كلية رياض الأطفال من عملية الاختبار لتحديد استعدادها وقدراتها. الثاني التقويم التشخيصي: ويقصد به التعرف على نواحي القوة والضعف أو الخبرات السابقة لدى المتعلمين. أي تحديد مالمدى الفرد بالفعل من مهارات وإمكانيات. الثالث التقويم البنائي أو مستمر: ويستخدم بهدف متابعة مدى تقدم المتعلمين أثناء دراستهم، ويرتبط هذا النوع من التقويم بمبادئ التدعيم والتغذية الراجعة. الرابع التقويم النهائي: وذلك للتعرف على مدى تحقيق الأهداف المرجوة مثل الوصول إلى المستوى المطلوب لمخرجات التعلم.

اختبار اللغة العربية الشامل الوطني هو من أدوات التقويم النهائي لأن عقد هذا الاختبار في نهاية المستوى الدراسي. تستخدم نتائج هذا الاختبار لمعرفة مدى تحقيق الأهداف المرجوة مثل الوصول إلى المستوى المطلوب لمخرجات التعلم، ولمعرفة جود.

3. مفهوم اختبارات اللغة

مفهوم الاختبار هو: مجموعة من المفردات " الأسئلة" التي تعطي للطالب للإجابة عنها شفويا أو تحريريا وقد تكون موضوعية أو مقالية أو رسوما أو أشكالا تستخدم للمقارنة والمقياس (أنسي محمد أحمد قاسم، 2003: 11).

وقال طعيمة ، أنّ الاختبار هو مجموعة الأسئلة التي يطلب من الدارسين أن يستجيبوا لها بهدف قياس مستواهم في مهارة لغوية معينة وبيان مدى تقدمه فيها ومقارنته بزملائه (رشدي أحمد طعيمة، 1989: 247).

ومن التعريف الموجود في الاختبار قامت الباحثة على أخذ الإستنباط بأنّ اختبار اللغة العربية هو أداة أو آلة لتقييم كفاءة الطلبة في اللغة العربية بعد اشتراكهم في عملية التعليم والتعلم،

أو نستطيع أن نقول بأن الاختبار هو إحدى من وسائل التقويم لمعرفة نجاح أو فشل عملية التعليم والتعلم.

إنّ عملية التعلّم للدارس والتعليم للمدرّس لم تكن خالية من التقييم لأنه من إحدى العوامل في نشاطات عملية التعلم والتعليم، وكذلك لمعرفة الوصول إلى الهدف المعلوم من عدمه في سير التعليم والتعلّم (محمد على) لأن في الاختبار هو ينمو كفاءتهم في عملية التعليم.

4. أنواع اختبارات اللغة

قسم يوسماني الاختبارات على سبعة أنواع (Joesmani, 1998: 26)

الأول الاختبار التكويني (formative test): وهذا الاختبار لفحص تنمية الطلاب في تعليم موضوع كل الدرس أو مواضع كل فصل. الثاني الاختبار الختامي (summative test): اختبار الذي له هدف لمعرفة حصول الطلاب بحسب كفاءتهم في التعليم الذي قدمه المدرس من قبل. ويعقد المدرس هذا الاختبار بعد نهاية التعليم، وهذا الاختبار سمي باختبار النصفى والاختبار النهائي.

الثالث الاختبار القبولي (entrance test): اختبار الذي يعقد لإعادة الاختيار الطلاب الجديد قبل دخولهم إلى المدارس المعينة. والهدف لوجود هذا الاختبار هو عدم توفر الإمكانيات الضرورية المتسعة لعدد الطلبة بالمدرسة. الرابع الاختبار التصنيفي (classification test): كان الاختبار التصنيفي يصمم بهدف توزيع الدارسين الجدد كل حسب مستواه في مجموعة من المجموعات التي تناسبه حتى يتسنى له البدء في دورة اللغة، وحتى لا يجلس مع مجموعة أعلى من مستواه فيضيع بينهم، أو مع مجموعة أدنى من مستواه فيفقد الدافعية والحماس. (عبد الرحمن بن إبراهيم الفوزان، 1423: 90). الخامس الاختبار التشخيصي (diagnostic test): كان الاختبار التشخيصي يصمم بهدف مساعدة كل من المدرس والدارس على معرفة نقاط الضعف والقوة لدى الدارس ومدى تقدمه في تعلم عناصر بعينها في دورة اللغة. ويعقد مثل هذا الاختبار في العادة بعد نهاية كل الوحدة في الكتاب المقرر أو بعد كل درس في الوحدة (عبد الرحمن بن إبراهيم الفوزان، 1423: 91). السادس الاختبار التحصيلي (achievement test): الاختبار التحصيلي يصمم لقياس ما يكون قد درسه الدارس خلال فترة قد تطول أو تقصر، فقد تكون عاما أو أقل، أو

لقياس مدرسه في دورة دراسية بأكملها، ويقصد به اكتشاف المستوى الذي توصل إليه الدارس مقارنة بزملائه الآخرين في المستوى نفسه (عبد الرحمن بن إبراهيم الفوزان، 1424: 90).

5. تحليل بنود الأسئلة في الاختبار

لكي يخرج الاختبار بصورة جيدة ويمكن الاعتماد عليه يجب أن يتسم بعدد من المعايير والشروط التي تحدد إمكانية استخدامه من عدمه فمتى ما اتصف الاختبار بتلك المعايير والشروط دل على أنه اختبار جيد ومتى ما فقدت منه أو لم يتصف بها كان غير ذلك. ومن هذه المعايير والشروط هي: الموضوعية و الصدق و الثبات و الواقعية و الاقتصادية (, suharsimi ari kunto, 2012: 72).

لترقية نوعية أدوات القياس كالاختبار، فعلى كل المدرّس أن يحلّل بنود الأسئلة لكشف صفات الاختبار، لمعرفة نوعية بنود الأسئلة، لتصميم الاختبار و لتطوير الاختبار، لمعرفة استخدام الاختبار الجيد ويمكن الدفاع عنه و يحسن النقصان لعقد الاختبار الآتي (sri wahyuni dan , 2012: 128).

وأهداف تحليل بنود الأسئلة في الاختبار هي: لمعرفة بنود الأسئلة القبيحة، لتحسين بنود الأسئلة في عقد الاختبار الآتي، لمعرفة صورة الاختبار المصمم، لمعرفة قوة و ضعف بنود الأسئلة لتفقيح تلك بنود الأسئلة، لمعرفة المشكلة الموجودة في بنود الأسئلة كالخطأ في وضع مفتاح الإجابة و وجود الأسئلة السهلة أو الصعبة جدًا و أسئلة التي لاتستطيع أن تفرّق بين الطالب المستعد والطالب غير المستعد في وجه الاختبار وغير ذلك (sri wahyuni dan abd syukur , 2012: 128).

والتقنيات المستخدمة لتحليل بنود الأسئلة هي تحليل درجة صدق بنود الأسئلة، وتحليل درجة الصعوبة و تحليل درجة التمييز وتحليل درجة التشبّت. ويمكن البيان بالتفصيل كما الآتي:

1. درجة الصدق

إن صدق الاختبار يعني إلى أي مدى يقيس الاختبار الشيء الذي وضع من أجله. بمعنى أن يكون الاختبار الصادق هو الذي يقيس ما أعد من أجل قياسه فعلا أي يقيس الهدف الذي أعد لقياسه ولا يقيس شيئاً آخر مختلفاً عنه. فمثلاً، الاختبار الذي أعد لقياس التحصيل في مادة

معينة لا يجب أن يكون بين أسئلته أسئلة متعلقة بقياس مادة أخرى. أي أن الاختبار مصمم لقياس سلوك معين لذلك يجب أن تكون فقرات هذا الاختبار كلها مرتبطة بهذا السلوك المراد قياسه. ولتوفير درجة الصدق العالية للاختبار، يجب أن يكون المحتوى ذا علاقة بالشئ الذي يراد قياسه. ويجب استبعاد أية مشكلة ثانوية لاعلاقة لها به، والتي قد تكون في ذاتها أكثر صعوبة من الصعوبات ذات الصلة بالاختبار.

والاختبار الصادق يتم ب: تحديد أهداف الاختبار، الابتعاد عن تأثير العوامل الخارجية مثل الصعوبات اللغوية وصعوبات القراءة، تحديد زمن الاختبار المناسب حتى يكون ذلك ملائماً للتلاميذ. وأنواع صدق الاختبار: (sri wahyuni dan abd syukur ibrahim, 2012:86-99).

صدق المحتوى (content validity)

صدق المحتوى هو دقة أداة القياس من حيث المحتوى ذلك أداة القياس. ويقال أداة القياس لديه صدق المحتوى إذا كان المحتوى أو المواد هذه أداة القياس مناسبة بالمواد المدروسة. بمعنى كان محتوى من أداة القياس مناسباً بالمواد المدروسة تستند إلى المنهج. طريقة للتحقيق صدق المحتوى هي غير يقارن مواد أداة القياس مع المكونات في إعداد أداة القياس مع التحليل العقلاني. إذا كانت المادة أداة القياس مناسبة مع إعداد المواد أداة القياس، إذا تلك أداة القياس لديها صدق المحتوى.

صدق البناء (construct validity)

صدق البناء يشير إلى التوافق بين نتائج أداة القياس مع القدرة على قياس. يقال أداة القياس لديها صدق البناء عالياً إذا كانت نتائج القياس وفقاً للخصائص السلوك الذي يقيسه.

ليحصل إلى صدق البناء يمكن أن ينظر إلى بنود الأسئلة والاقتران تلك البنود مع بعض الأهداف التي تهدف إلى كشف المستوى الجوانب المعرفية. لتحديد مدى صدق البناء، ويمكن أن يتم إعداد بنود الأسئلة مستندة على شعرية من أدوات القياس.

الأول صدق المقياس (measurement validity): أظهر صدق المقياس إلى فهم أعلى قدرة الطلاب لمن يتم تدريسه من الطلاب الذين لمن لم يتم تدريسه. الثاني صدق التنبؤي

(predictive validity): صدق التنبؤي هو أن دقة أداة القياس يظهر من قدرة الأدوات للتنبؤ الإنجاز الذي سيحقق في الوقت الآتي. الثالث صدق بنود الأسئلة (item validity): اختبار الصدق الذي سبق ذكره هو اختبار الصدق بالإجمال، دون الإهتمام على كل بند الاختبار. إذا، الاختبار لا يكون صادقاً إذا كان أدوات الاختبار وبنود الأسئلة ليس لهما درجة الصدق. صدق بنود الأسئلة يؤثر كثيراً إلى صدق الاختبار، إذا كان بنود الأسئلة لها درجة الصدق المرتفعة وكان درجة الصدق ذلك الاختبار مرتفعة أيضاً. (sudaryono, 2012: 149).

2. درجة الثبات

الثبات هو معيار القياس هل أداة القياس تستطيع أن تقيس بالنتائج المتناسقة من وقت إلى وقت. الثبات بمعنى انتظام عندما يتم ذلك الاختبار يختبر مرارا والنتيجة هي متساوية، وهذا هو القول بعد ارتباط نتائج الاختبار لأول مع الاختبار التالي، وهنا ارتباط كبير بينهما. الثبات بمعنى استقرار الثبات عندما تم اختبار هو تحليل نتائج الثبات باستخدام المعايير الداخلية في تلك الاختبارات. طريقة لمعرفة معامل الاستقرار مع بعض الصيغ باستخدام اختبار واحد وكان أداء الاختبار يكفي بمرّة واحدة. ويقصد بالثبات عدم التذبذب في الاختبار إذا ما قصد به أن يكون بمثابة المقياس. فالمقياس المتري مثلا يمكن أن تقيس به الطول والعرض لعدة أشياء، ويمكن بعد فترة أن تقيس الأشياء نفسها بالمقياس المتري نفسه وتحصل على النتائج نفسها دون تذبذب مادام الطول والعرض كما هما لم يتغيران. وعلى هذا فإن ثبات الاختبار يرتبط إلى حد كبير بثبات التقدير العام أو حتى الدرجات التي يحرزها الدارس نفسه، فإذا ما تذبذبت درجته فإن هذا يعني أن المقياس أو الاختبار لا يتصف بالثبات. فمثلا إذا قدمنا اختبارا معيناً لخبطة من الدارسين اليوم، وبعد تحصيل نوعي وكمي معلوم لدينا، ثم قدمنا هذا الاختبار نفسه للخبطة نفسها تحت ظروف مطابقة أو مشابهة دون إضافة جديدة ثم جاءت النتائج مختلفة، فيمكن القول في هذا الحال بأن هذا الاختبار يفتقر إلى الثبات فهو اختبار غير جيد (عبد الرحمن بن إبراهيم الفوزان، 1424: 92).

وطرق لبحث درجة ثبات الاختبار هي (sri wahyuni dan abd syukur ibrahim, 2012: 106-115).

- **طريقة إعادة الثبات:** هذه الطريقة لاختبار ثبات أداة لقياس مع استخدام تختبر تلك أداة مرتين أو أكثر، ثم يرتبط النتائج من ذلك الاختبار. الخطوات المستخدمة لمعرفة درجة الثبات بالطريقة إعادة الثبات هي: الأول تصميم أداة القياس المختبر درجة ثباتها، ثم تختبر تلك أداة القياس، ثم حساب نتائج أداة القياس الأولى، ثم إعادة الاختبار لتلك أداة القياس، ثم حساب نتائج إعادة الاختبار، ثم حساب درجة الثبات تلك أداة القياس بارتباط نتائج أداة القياس الأولى بالنتائج إعادة الاختبار باستخدام الرمز (korelasi product moment)
 - **طريقة المنحرفة:** تتم هذه الطريقة بالتصميم أداتين، لهما صفة متساوية، بعد أن يتم اختبار كل الأداةين ثم ترتبط النتائج كلاهما. ولو كان تتكون من أداتين للقياس وبالتحقيق كانت هذين أداتين تقاس شيئاً متساوياً. الخطوات المستخدمة في هذه الطريقة هي: الأول تصميم أداتان متساويتان للقياس، ثم اختبار تلك أداتان في وقت واحد، ثم تعيين النتائج لكل أداة القياس، ثم يرتبط تلك النتائج باستخدام الرمز (korelasi product moment)
 - **طريقة التجزئة النصفية:** تتم هذه الطريقة بتقسيم أداة القياس، على سبيل المثال، تقسيم الاختبار إلى قسمين ثم ترتبط نتيجتهما باستخدام رمز معين. وطريقة تقسيم أداة القياس هي تقسيم بين النتيجة الفردية والنتيجة المتزوجة أو تقسيم بين النمرة العلوي والنمرة السفلي.
 - الخطوات المستخدمة في هذه الطريقة هي: الأول تصميم الاختبار وله النمرة الفردية، ثم أداء الاختبار، ثم حساب نتائج الاختبار ثم تقسمها إلى النتيجة الفردية والنتيجة المتزوجة أو النتيجة العلوي والنتيجة السفلي، ثم ترتبط تلك نتيجتين بالرمز (korelasi product moment) لبحث درجة الثبات نصف الاختبار، ثم استخدام الرمز (spearman brown) لبحث درجة ثبات الاختبار بالكامل
- تقويم درجة الثبات في هذا البحث لمعرفة استقرار بنود الأسئلة في الاختبار، الفرضية الأساسية لتقويم درجة الثبات بأن هذا الاختبار مصمم بأنماط نظامية، حتى إذا كان نقسم نتائج الاختبار بين النتيجة العلوي والنتيجة السفلي فلا يغير موقف النتيجة لكل الطالب.

3. درجة الصعوبة

الاختبار الجيد له بنود الأسئلة الجيدة. وبنود الأسئلة الجيدة هي لا يصعب الطلبة كثيرا و لا يسهل الطلبة كثيرا. الأسئلة السهلة جدًا لا يحفز الطلبة لإجابة تلك الأسئلة، وبالعكس، الأسئلة الصعبة جدًا يجعل الطلبة يشعر باليأس عند إجابة تلك الأسئلة لأنهم يشعرون غير قادرين لإجابة تلك الأسئلة (sri wahyuni dan abd syukur ibrahim, 2012 : 129).

لذلك يحتاج إلى تحليل درجة الصعوبة لكل بنود الأسئلة لمعرفة درجة الصعوبة ودرجة السهولة من الاختبار. بهذا التحليل يغتاب درجة بنود الأسئلة، هل هذا الاختبار السهل، أو الاختبار الصعب وغير ذلك.

تقويم درجة الصعوبة في اختبار اللغة العربية الشامل الوطني لمعرفة جودة هذا الاختبار، لأن أسئلة الاختبار الجيدة لها موصفة متعددة بين درجة صعوبتها ودرجة سهولتها. حتى هذا الاختبار يستطيع أن ينمي حماسة الطلبة في إجابة أسئلة اختبار اللغة العربية.

4. درجة التمييز

من صفات الاختبار الجيد أن تكمن فيه القدرة على التمييز بين مختلف الدارسين من حيث الأداء، ففي كل صف من الصفوف نجد تباينا بين الدارسين، فهناك المتفوق والضعيف ومستويات بين هؤلاء وهؤلاء. ولكي يفرق الاختبار بين هذه الفئات، فإن على واضعي الاختبارات أن يتوخوا الدقة قدرة الإمكان في مدى سهولة الأسئلة وصعوبتها بحيث لا تكون كلها صعبة يبرز فيها المتفوقون فقط، أو متوسطة يجيب عنها التفوقون والتوسطون دون الضعاف، أو سهلة كلها بحيث لا تفرق بين الجميع. فعلى سبيل المثال، فإن الحصول على 80% لا يعد شيئا كبيرا إذا ما حصل كل الطلاب على درجات في المدى ما بين 67% و 82% ولا شك في أن اختبارا يحصل فيه معظم الدارسين على المدى المذكور هو اختبار يفتقر إلى عنصر التمييز لأن أسئلته غير جيدة ولأن السؤال الجيد ينبغي أن يكون غير غامض، أي أنه يجب أن يكون للسؤال تفسير واحد. وكذلك فإن السؤال الجيد هو المميز الذي تتفق نتيجته مع النتيجة العامة للاختبار كل (عبد الرحمن بن إبراهيم الفوزان، 1424: 93).

ويمكن تحديد مستوى سهولة وصعوبة كل سؤال في الاختبار بناء على نتائج أو استجابة المتعلمين على كل سؤال أو فقرة الاختبار وبحساب المعادلة التالية:

- عدد الإجابات الصحيحة أو عدد الإجابات الخاطئة

- المجموع الكلي للمتعلمين أو (عدد الإجابات الصحيحة + عدد الإجابات الخاطئة).

بمعرفة درجة تمييز في بنود الأسئلة، نستطيع أن نعرف قدرة ذلك الاختبار في تفریق بين الطلاب الماهر و الطلاب الجاهل. ينبغي على الطلاب الماهر أن يجيب الأسئلة بالأجوبة الصحيحة أكثر إجابة من الطلاب الجاهلين. ارتفع مستوى تمييز الاختبار فيرتفع قدرته في تمييز الطلاب الماهر والطلاب الجاهل (عبد الرحمن بن إبراهيم الفوزان، 1424: 133).

يرتفع مستوى تمييز الاختبار إذا كان نجد الفرق في عدد الإجابة الصحيحة المنتجة من المجموعتين الموجودة. تزايد عدد الإجابة الصحيحة بين المجموعتين فتزايد مستوى التمييز من بنود الأسئلة.

تقويم درجة التمييز في اختبار اللغة العربية الشامل الوطني لمعرفة قوّة هذا الاختبار لتمييز الطلاب الماهرون والطلاب غير الماهرين. كلما ارتفعت درجة التمييز لبنود الأسئلة فارتفعت قوتها لتمييز الطلاب الماهرون والطلاب غير الماهرين.

5. درجة التشتت

لمعرفة جودة الاختبار الجيد ليس إلا بالنظر إلى مستوى الصعوبة و مستوى التمييز فقط. بل علينا أن نهتم إلى درجة التشتت من بنود الأسئلة الموجودة. تحليل درجة التشتت في بنود الأسئلة يكون أمرا مهماً لوجود الفرق في تردد الإجابة بين الطلاب الماهر والطلاب الجاهل. لبدائل الإجابة الصحيحة، على كل الطلاب الماهر أن يختار هذه البدائل أكثر من الطلاب الجاهل. و هذا الشيء يدلّ على فهرس مستوى التمييز في بنود الأسئلة. على العكس، لبدائل الإجابة المشتتة على كل الطلاب الجاهل أن يختار هذه البدائل أكثر من الطلاب الماهر (عبد الرحمن بن إبراهيم الفوزان، 1424: 41).

فلذلك، التشتت الجيد هو التشتت الذي يمكن تجنبه الطلبة المرتفعون و يختار كثيرا الطلبة المنخفضون. إلى ذلك، كلّ بدائل الأجوبة الموجودة كانت مختارة كلها عند الطلبة. لمعرفة

فعالية بدائل الإجابة في الاختبار، لا بدّ منها القيام بتحليل درجة التشتت في بنود الأسئلة. بهذا التحليل نستطيع أن نعرف توزيعية تردد الإجابة.

ت. منهجية البحث

1- مدخل البحث ومنهجه

تستخدم الباحثة في إعداد بحثها المنهج الوصفي بالمدخل الكمي. ويقصد بالمنهج الوصفي هو وصف الظاهرة التي يريد الباحث دراستها وجمع أوصاف ومعلومات دقيقة عنها مع إجراء عملية التقويم المناسب لها.

وهذا البحث من نوع البحث تحليل المحتوى (content analysis). لأن قامت الباحثة بتحليل و تقويم بنود الأسئلة لاختبار اللغة العربية الشامل بعد حصلت على ورقة أجوبة الطلبة في الاختبار. والبيانات تتكون من البيانات الكميّة والبيانات الوصفيّة.

2- مجتمع البحث وعينته

إن مجتمع البحث في هذا البحث هو طلبة في الفصل الثاني عشر بمدرسة الثانوية الإسلامية الحكومية الثالثة مالانق في السنة الدراسية 2015-2016م. ينقسم أقسام العلوم في هذه المدرسة إلى أربعة أقسام تعني قسم علوم الدينية و قسم علوم اللغة وقسم علوم الطبيعة وقسم علوم الإجتماعية.

وعينة البحث في هذا البحث هو طلبة في الفصل الثاني عشر في قسم علوم الطبيعة وكان عددهم 35 طالبا. اخترت الباحثة قسم علوم الطبيعة لأن الطلبة في هذا القسم له أضعف الكفاءة من الطلبة في قسم اللغة أو قسم علوم الدين، وأمرت مدرسة اللغة العربية لاختيار هذا القسم لمعرفة كفاءتهم في اللغة العربية بالضبط.

3- البيانات ومصادرها

يستعمل مصادر البيانات في هذا البحث كما يالي:

مصادر البيانات الأساسية أخذت الباحثة من تحليل ورقة إجابة الطلبة وبنود أسئلة

الاختبار. و مصادر البيانات الثانوية هي دليل المقابلة بمدرسة اللغة العربية.

4- أدوات جمع البيانات

الأدوات المستخدمة لجمع البيانات هي :

الوثائق: قامت الباحثة بتحليل ورقة أجوبة الطلبة لمعرفة جودة بنود الأسئلة في اختبار اللغة العربية الشامل (UAMBN) للسنة الدراسية 2014-2015 م من ناحية درجة الصدق والثبات والصعوبة والتمييز والتشتت، ثم المقابلة، قامت الباحثة بالمقابلة مع مدرسة اللغة العربية لنيل البيانات عن الأحوال الطلبة والمدرسة.

5 - أسلوب تحليل البيانات

تحليل البيانات عند فاتون (في مولونج) هو عملية تنظيم ترتيب البيانات، وتنظيمها إلى التصميم، وفصيلة و وحدة فكك اساسي (moleong lexy I, 2005: 103). وتحليل البيانات في هذا البحث هو :

- قامت الباحثة بإعادة الاختبار، لحصول على ورقة أجوبة الطلبة في اختبار اللغة العربية الشامل (UAMBN) للسنة الدراسية 2014-2015م.

- قامت الباحثة بتحليل ورقة أجوبة الطلبة في اختبار اللغة العربية الشامل (UAMBN) للسنة الدراسية 2014-2015م. لمعرفة جودة ذلك الاختبار من ناحية درجة الصدق و الثبات و الصعوبة و التمييز والتشتت. و الرموز المستخدم لمعرفة مستوى جودة الاختبار هو:

(1) طريقة تحليل درجة الصدق تعني صدق بنود الأسئلة هي بتحليل إجابة الطلبة في بنود الأسئلة الممتحنة درجة صدقه، ثم يختبر لكل بنود الأسئلة بارتباط بين بنود الأسئلة بنتيجة الطلبة المحسولة. بعد ذلك ادخال تلك البيانات إلى الرمز “product moment” باستخدام برامج (excel) كما يلي: (sri wahyuni dan abd syukur halim, 2012: 100)

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

(1) تحليل درجة الثبات باستخدام طريقة التجزئة النصفية، والخطوات منها كما يلي: (sri wahyuni dan abd syukur halim, 2012: 136).

- تصنيف نتائج الاختبار تكون النتيجة الفردية و النتيجة المتزوجة أو تصنيف نتائج الاختبار تكون النتيجة العلوي والنتيجة السفلي.
- بحث ثبات نصف اختبار بارتباط نتيجتين باستخدام الرمز "product moment" و باستخدام برامج (excel).
- بحث ثبات الاختبار الشامل باستخدام الرمز "spearman - Brown"
- إذا كانت نتيجة (r statistik Spearman Brown) أكبر من نتيجة (r tabel). وهذا دليل على وجود الإتصال على شكل ملحوظ بين X و Y معنى ذلك الاختبار له ثبات جيد
- الرمز "spearman - Brown" كما يلي:

$$r = \frac{2 \times r}{1 + r}$$

ثبات بنود الأسئلة الإجمالية: x^2 ثبات نصف الاختبار

+1 ثبات نصف الاختبار

(2) تحليل درجة الصعوبة باستخدام الرمز كما يلي:

$$IF = \frac{FH + FL}{N}$$

IF: *Item Facility*, (درجة الصعوبة بنود الأسئلة)

FH: *Frequency High*, (عدد الإجابة الصحيحة للمجموعة المرتفعة)

FL: *Frequency Low*, (عدد الإجابة الصحيحة للمجموعة المنخفضة)

N: عدد الطلاب من مجموعتين

قائمة درجة الصعوبة هي سهلة تعني: $0.70 \leq IF < 1.00$. صعبة تعني: $IF < 0.30$

متوسطة تعني: $0.30 \leq IF < 0.70$

(2) تحليل درجة التمييز باستخدام الرمز كما يلي: (*sri wahyuni dan abd*)

. (*syukur halim, 2012: 136*)

$$ID = \frac{FH - FL}{N}$$

ID: Item Discriminability, (درجة التمييز بنود الأسئلة)

FH: Frequency High, (عدد الإجابة الصحيحة للمجموعة المرتفعة)

FL: Frequency Low, (عدد الإجابة الصحيحة للمجموعة المنخفضة)

N: عدد الطلاب من المجموعة الواحدة فقط

قائمة درجة التمييز هي جيّد جدًا تعني: $40,0 < ID$. جيّد تعني: $30,0 < ID \leq 39,0$.

متوسطة تعني: $20,0 < ID \leq 29,0$

(3) تحليل درجة التشتت بهذه الخطوات: (sri wahyuni dan abd syukur)

. (halim, 2012: 142-143)

- تنظيم نتيجة الطلبة من أرفع النتيجة حتّى أسفل النتيجة.
- تقسيم 27,5% من عدد الطلبة يدخل إلى مجموعة النتيجة المرتفعة، 27,5% من عدد الطلبة يدخل إلى مجموعة النتيجة المتوسطة، 27,5% من عدد الطلبة يدخل إلى مجموعة النتيجة المنخفضة.
- حساب اختيار الطلبة على الإجابات الموجودة حتّى نعرف البيانات المنتشرة عن تردد الإجابة بين ثلاثة مجموعات.
- تحليل تشتت بنود الأسئلة بالنظرها ويقارنها عدد تشتت الإجابة المختارة بين المجموعة المرتفعة والمجموعة المنخفضة.

ث. نتيجة البحث ومناقشتها

1- درجة صدق المحتوى اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة الدراسية

2015 في المدرسة الثانوية الإسلامية الحكومية الثالثة مالانق.

طريقة التي لقياس دراجة صدق المحتوى هي باستخدام آراء الخبير في تعليم اللغة العربية و في مجال تقويم التربية والتعليم. أو نستطيع أيضا بمقارنة مواد الاختبار بمواد تأليف أسئلة الاختبار.

في هذا البحث قارنت الباحثة أسئلة الاختبار بالمواد المدروسة لدى الطلاب. ووصفت وزارة الشؤون الدينية تلك المواد إلى المعايير الكفائية للمتخرجين. تلك المعايير هي المعايير الكفائية المقررة من لجنة اختبار اللغة العربية الشامل الوطني (UAMBN). وكانت هذه المعايير

قد تناسب بالمواد الدراسية للطلاب. بعد ذلك حللت الباحثة ورقة أسئلة اختبار اللغة العربية الوطني وتناسبها بالمعايير الموجودة.

من ذلك التحليل هذا اختبار اللغة العربية الوطني تكون صادق من ناحية المحتوى لتوافق أو لتناسب بعض الأسئلة بالمعايير المقررة من وزارة الشؤون الدينية للمتخرجين وتلك المعايير قد تناسب بالمواد اللغة العربية المدروسة لدى الطلاب. ولكن رأيت الباحثة نقصان من ناحية عدد الأسئلة لكل عناصر اللغة العربية. كان مضمون هذه الأسئلة لم يكن بالتساوي بين عناصر اللغة ومهارات اللغة العربية. كان بعض الأسئلة تدل كثيرا إلى عدّة عناصر اللغة أو مهارات اللغة فقط. يتكون هذا الاختبار بخمسين سؤالاً، وكانت بنود الأسئلة هذا الاختبار تتكلم كثيرا عن القوائد اللغوية فقط تعني خمسين في المئة. وأمّا بقية من الأسئلة تتكلم عن مهارة القراءة ومهارة الكتابة والمفردات والترجمة. تعني: مهارة القراءة (3) أسئلة، مهارة الكتابة (5) أسئلة، مفردات (11) سؤالاً، ترجمة (3) أسئلة، قواعد (26) سؤالاً، والثقافة سؤالين.

عند رأي الباحثة كان من موصفات الاختبار الجيد هو التناسب و الشمول. والمراد بالتناسب هو كانت مواد الاختبار أن تكون مناسبة بالمواد المدروسة لدى الطلاب. فطبعا أن تكون تلك المواد مناسبة بالمنهج المستخدم.

والمراد بالشمول هنا يعني أن مضمون الاختبار أن يشمل جميع مكونات مواد الدراسة بالتساوي. ولكن كانت بعض الأسئلة في الاختبار الوطني تتكلم كثيرا إلى عدة بعض المادة فقط، وهي القواعد اللغوية. أمّا هذا الاختبار لا يقيس مادة القواعد فقط، لكن كثير المواد و المهارات اللغوية التي ستقيس بهذا الاختبار كالمفردات و الترجمة و مهارة القراءة وغير ذلك.

2- درجة صدق بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة

الدراسية 2015 في المدرسة الثانوية الإسلامية الحكومية الثالثة مالانق.

كما عرفت الباحثة هناك الطرق الكثيرة لقياس درجة صدق الاختبار هو صدق المحتوى، صدق التوقع، صدق المبني، صدق القياس وغير ذلك. تلك الطرق الكثيرة نستخدمها لقياس درجة صدق الاختبار على شكل عام ولا يهتم إلى أحوال بنود الأسئلة واحدا فواحدا. كما عرفت الباحثة كان صدق بنود الأسئلة تأثر كثيرا إلى صدق الاختبار على شكل عام. إذا كان بنود

الأسئلة لها درجة الصدق العالية وكان ذلك الاختبار له درجة الصدق العالية أيضا. وقامت الباحثة بتحليل ورقة أجوبة الطلاب لحصول إلى البيانات المقصودة.

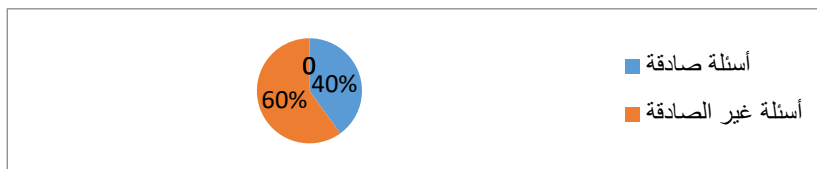
بعد أن فتشت الباحثة نتائج الطلاب، ثم حسبت الباحثة عدد النتائج الصحيحة لكل طالب. كان عدد مشترك هذا الاختبار هو 25 مشتركا. وكانت عدد أسئلة الاختبار هي 50 سؤالا. و مجموع الإجابات الصحيحة من الطلاب هي 820 إجابة. وجدول الذي يبين تلك البيانات ستعرض الباحثة في ملاحق البحث.

بعد أن عرفت الباحثة نتائج الطلاب ثم اختبرت الباحثة كل بنود الأسئلة بارتباط تلك بنود الأسئلة بالنتائج المحسولة عند الطلاب. بعد ذلك أدخلت الباحثة تلك البيانات إلى الرمز "product moment" باستخدام برامج (excel) كما يلي:

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

بعد أن حسبت الباحثة نتائج الطلاب بذلك الرمز سوف نعرف نتيجة (r hitung) لكل بنود الأسئلة ثم قارنت تلك النتيجة بنتيجة (r tabel). إذا كانت نتيجة (r hitung) أكبر من نتيجة (r tabel) وكان ذلك السؤال هو سؤال صادق. وإذا كانت نتيجة (r hitung) أصغر من نتيجة (r tabel) وكان ذلك السؤال هو سؤال غير صادق.

من الجدول السابق رأت الباحثة بأن أسئلة التي صادقة هي (40%) من 50 سؤالا تعني (20 سؤالا). وأما أسئلة التي غير الصادقة هي (60%) تعني (30 سؤالا). والبيانات بالتفصيل كما في الصورة الآتية:



من البيانات الموجودة حللت الباحثة بأن هذا الاختبار ليست لها درجة الصدق العالية نظرا لكثرة أسئلة التي غير الصادقة في هذا الاختبار. وهذا النقصان يكون ملاحظة لترقية جودة اختبار اللغة العربية الشامل الوطني للسنة الآتية لتكون اختبارا معياريا جيدا في هذا الوطن.

عند رأي الباحثة من مواصفات الاختبار الجيد هي الصدق. الصدق على شكل العام أو الصدق على شكل الخاص وهو صدق بنود الأسئلة. على مؤسس الأسئلة أن يهتم بصدق الأسئلة اجمالية أو صدق الأسئلة واحدا فواحدا وهو لكل بنود الأسئلة. لكي يكون ذلك الاختبار هو الاختبار الجيد.

3- درجة ثبات بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة الدراسية 2015 في المدرسة الثانوية الإسلامية الحكومية الثالثة مالانق.

معرفة درجة الثبات في الاختبار هي أمر مهم. نستطيع أن نعرف جودة ذلك الاختبار من هذه الصفة تعني ثبات بنود الأسئلة في الاختبار. لمعرفة ثبات بنود الأسئلة لابد لنا أن نوزع الأسئلة على قسمين: الأسئلة الفردية والأسئلة المتزوجة. بعد تحليل الباحثة ورقة أجوبة الطلاب . حصلت النتائج كما تلي:

بعد أن حللت الباحثة ورقة أجوبة الطلاب نعرف أن عدد الممتحنين هو 25 ممتحنا ($\Sigma 25$) . وثانيا نعرف أن مجموع الإجابات الصحيحة من الأسئلة الفردية هو (415 إجابة) . وثالثا نعرف أن مجموع الإجابات الصحيحة من الأسئلة المتزوجة هو (405 إجابة). ورابعا نعرف أن مجموع تركيب الأسئلة الفردية هو (7237). وخامسا نعرف أن مجموع تركيب الأسئلة المتزوجة هو (6911). وسادسا نعرف أن مجموع ضرب بين الأسئلة الفردية والأسئلة المتزوجة هو (6936). وسابعا نأتي بقيام تحليل ثبات الاختبار بالرمزين : product moment و sperman .brown

أما تحليل برمز product moment هو لمعرفة ثبات نصف السؤال. والرمز كما يلي:

$$r_{xy} = \frac{N \Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{\{N \Sigma X^2 - (\Sigma X)^2\} \{N \Sigma Y^2 - (\Sigma Y)^2\}}}$$

البيان:

N: عدد الطلاب المشتركين في الاختبار

Y: عدد النتيجة المتزوجة

X: عدد النتيجة الفردية

ومن ذلك الرمز عرفت الباحثة أن ثبات نصف اختبار اللغة العربية الوطني هو (610,0).
ثم استخدمت الباحثة الرمز sperman brown لمعرفة ثبات جميع الاختبار. والرمز كما

يلي:

$$r: \frac{2 \times r}{1 + r}$$

ثبات جميع الاختبار: ثبات نصف الاختبار X 2

ثبات نصف الاختبار +

1

من ذلك الرمز عرفت الباحثة أن ثبات جميع الاختبار هو (758,0) بعد ذلك قارنت
الباحثة تلك النتيجة بنتيجة r tabel (396,0). ثم رأت الباحثة بأن النتيجة r statistik
(Spearman Brown) أكبر من نتيجة (r tabel). وهذا دليل على وجود الإتصال على شكل
ملحوظ بين X و Y. إذا، ذلك التحليل يدل على هذا الاختبار له درجة الثبات الداخلية
المرتفعة.

4- درجة صعوبة بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة
الدراسية 2015 في المدرسة الثانوية الإسلامية الحكومية الثالثة مالانق.

بنود الأسئلة الجيدة هي بنود التي لاتصعب طلبة كثيرا و لا تسهل طلبة كثيرا، وكانت درجة
صعبتها في المستوى المتوسط. في الخطوة الأولى نظمت الباحثة نتائج الطلبة من أكبر النتائج
إلى أصغر النتائج ثم قسّمت الباحثة تلك النتائج إلى مجموعتين تعني المجموعة الأولى للطلاب
المرتفعين والمجموعة الثانية لطلاب المنخفضين .

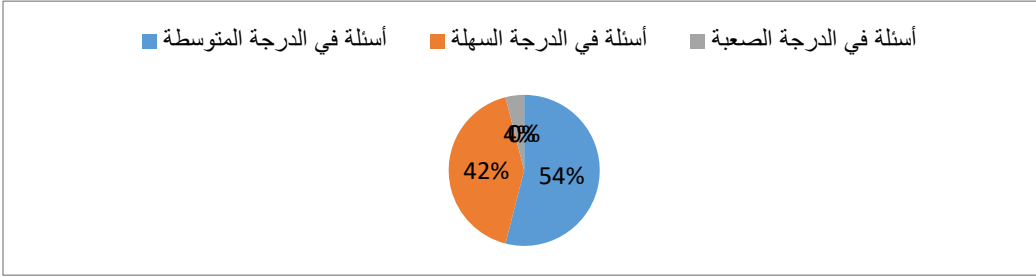
بعد ذلك، حلّلت الباحثة درجة الصعوبة لكل رقم من الأسئلة. بإضافة نتائج الطلاب
المرتفعون و نتائج الطلاب المنخفضون ثم تقسّم تلك النتائج بعدد الطلاب من مجموعتين، كما
في الرمز الآتي:

$$IF: \frac{FH + FL}{N}$$

(درجة الصعوبة: النتائج الصحيحة للطلاب المرتفعين + النتائج الصحيحة للطلاب المنخفضين)

عدد الطلاب من مجموعتين

من التحليل الموجود رأَت الباحثة بأنّ درجة صعوبة بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة الدراسية 2015 تتكون من (54%) أسئلة في درجة متوسطة، و (42%) أسئلة في درجة سهولة، و (4%) في درجة صعوبة. والبيانات بالتفصيل كما في الصورة الآتية:



إذا أستطيع أن ألخص بأنّ 50% أسئلة في هذا الاختبار هي أسئلة جيّدة لأن تلك الأسئلة لها درجة الصعوبة على المستوى المتوسط، وهذا المستوى إحدى من موصّفات أسئلة الاختبار الجيّدة. وأمّا الباقي من الأسئلة الموجودة في هذا الاختبار هي أسئلة غير جيّدة لأن تلك الأسئلة لها درجة الصعوبة على المستوى السهولة جدًّا وبعضها على المستوى الصعوبة جدًّا. وهذا المستوى هو إحدى من موصّفات أسئلة الاختبار غير جيّدة، ومن المستحسن لاتستخدم تلك الأسئلة في الاختبار أو كانت تلك الأسئلة تحتاج إلى الإصلاح.

5- درجة تمييز بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة الدراسية 2015 في المدرسة الثانوية الإسلامية الحكومية الثالثة مالانق.

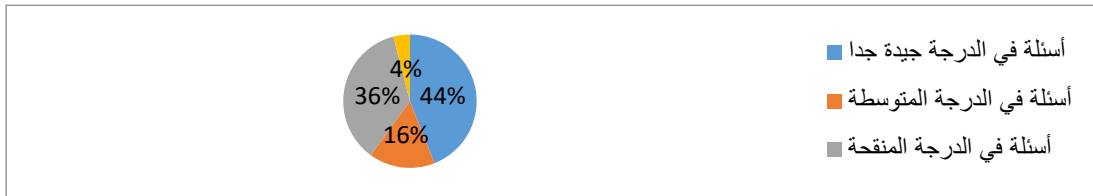
بنود الأسئلة الجيّدة لها درجة التمييز في كل سؤال وهي أسئلة التي تستطيع أن تميّز بين الطلاب المرتفعين والطلاب المنخفضين. في الخطوة الأولى نظمت الباحثة نتائج الطلبة من أكبر النتائج إلى أصغر النتائج ثمّ قسّمت الباحثة تلك النتائج إلى مجموعتين تعني المجموعة الأولى للطلاب المرتفعين والمجموعة الثانية لطلاب المنخفضين. بعد ذلك حلّلت الباحثة درجة التمييز لكلّ رقم من الأسئلة. بنقص نتائج الطلاب المرتفعون و نتائج الطلاب المنخفضون ثمّ تقسّم تلك النتائج بعدد الطلاب من المجموعتين أو 27,5% منه، كما في الرمز الآتي:

$$ID = \frac{FH - FL}{N}$$

(درجة الصعوبة: النتائج الصحيحة للطلاب المرتفعين – النتائج الصحيحة للطلاب المنخفضين)

عدد الطلاب من مجموعتين أو 27,5% منه

من التحليل السابق رأت الباحثة بأنّ درجة تمييز بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة الدراسية 2015 تتكون من (44%) أسئلة في درجة جيّدة جدًا وهي (22 سؤالاً) ، و (0%) أسئلة في درجة جيّدة، و (16%) في درجة متوسطة وهي (8 أسئلة)، و (36%) أسئلة في درجة منقحة وهي (18 أسئلة)، و 4% أسئلة في درجة مرفوضة وهي (سؤالين). والبيانات بالتفصيل كما في الصورة الآتية:



إذا أستطيع أن ألخص بأن بعض الأسئلة في هذا الاختبار تعني (50%) أسئلة هي أسئلة جيّدة لأن تلك الأسئلة لها درجة التمييز على مستوى جيّد جدًا و مستوى متوسط، وهذا المستوى إحدى من موصّفات أسئلة الاختبار الجيّدة من ناحية دراجة التمييز. وكانت هذه الأسئلة تستطيع أن تميّز بين الطلاب المرتفعين والطلاب المنخفضين. وكانت بعض الأسئلة التي لها الدرجة المنقحة فتحتاج إلى الإصلاحات لتكون الأسئلة الجيدة وقادرة على تمييز كفاءة الطلبة في اللغة العربية. وأسئلة التي في الدرجة المرفوضة لانستطيع أن نستخدم تلك الأسئلة مرة أخرى ولا بد علينا كالمواضع الأسئلة أن نبدلها بأحسن الأسئلة.

6- درجة تشتت بنود الأسئلة في اختبار اللغة العربية الشامل الوطني (UAMBN) للسنة الدراسية 2015 في المدرسة الثانوية الإسلامية الحكومية الثالثة مالانق.

تحليل درجة تشتت بنود الأسئلة في اختبار اللغة العربية الشامل الوطني للسنة الدراسية

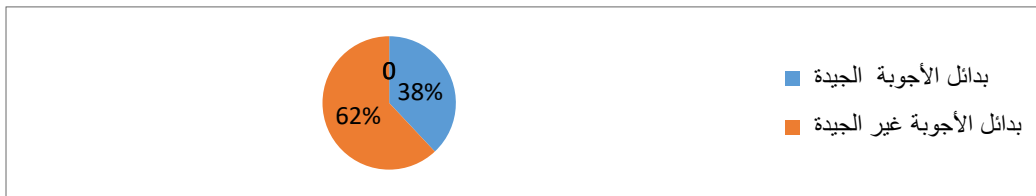
2015 في المدرسة الثانوية الإسلامية الحكومية الثالثة مالانق هو:

أولاً، نظمت الباحثة نتيجة الطلبة من أرفع النتيجة إلى أسفل النتيجة. ثانياً، قسّمت الباحثة عدد الطلاب إلى 27,5% من عدد الطلبة يدخل إلى مجموعة النتيجة المرتفعة، و27,5% من عدد الطلبة يدخل إلى مجموعة النتيجة المنخفضة. في هذا الاختبار كان عدد الطلبة لكل مجموعة هو 7 طلاب. وجدول الذي يبين تلك البيانات ستعرض الباحثة في الملاحق.

ثالثاً، حسبت الباحثة اختيار الطلبة على الإجابات الموجودة حتى نعرف عدد الانتشار إجابة الطلبة لكل مجموعة. رابعاً، حللت الباحثة تشتت بنود الأسئلة بالنظرها ويقارنها عدد تشتت الإجابة المختارة بين المجموعة المرتفعة والمجموعة المنخفضة.

ومن ذلك التحليل رأت الباحثة بأن بعض بدائل الأجوبة لها درجة التشتت القبيحة وتحتاج إلى الإصلاح لتكون بدائل الأجوبة الجيدة. كما عرفت الباحثة بأن التشتت الجيد له الترددات في الإجابة بين الطلاب المرتفعين و الطلاب المنخفضين. التشتت الجيد له أكثر الإختيار من الطلاب المنخفضين. ولجميع بدائل الأجوبة أن تكون مختارة جميعاً.

من الجدول السابق حللت الباحثة بأن هذا الإختبار تتكون من (62%) بدائل الأجوبة غير الجيدة من ناحية الدرجة التشتت تعني (124 بدائل الأجوبة). وأما الباقي (38%) لها درجة التشتت الجيدة تعني (76 بدائل الإجابة). كما حسبت الباحثة كان هذا الاختبار يتكون من 200 بدائل الأجوبة. والبيانات بالتفصيل كما في الصورة الآتية:



وأمر التي تسبب إلى قبيح بدائل الأجوبة من ناحية درجة التشتت هي:
 لاالطلاب الذي يختار هذه بدائل الأجوبة إقما من الطلاب المرتفعين أو الطلاب المنخفضين.
 ثم يختار الطلاب المرتفعون بدائل الأجوبة أكثر من الطلاب المنخفضين. ثم كانت بدائل الأجوبة لها ترددات متساوية في إجابة الطلاب المرتفعين والطلاب المنخفضين. و يختار الطلاب المرتفعون بدائل الإجابة ولا الطلاب المنخفضون.

ج. خاتمة

أ- نتائج البحث

النتائج من هذا البحث هي:

- 1- درجة صدق المحتوى اختبار اللغة العربية الشامل الوطني للسنة 2015 هي جيدة من ناحية المواد، لتوافق أو لتناسب بعض الأسئلة بالمعايير المقررة من وزارة الشؤون الدينية للمتخرجين. ولكن مضمون هذه الأسئلة لم يكن بالتساوي بين عناصر اللغة ومهارات اللغة العربية وكانت بنود الأسئلة هذ الاختبار تتكلم كثيرا عن القوائد اللغوية فقط تعني خمسين في المئة. وأما بقيّة من الأسئلة تتكلم عن مهارة القراءة ومهارة الكتابة والمفردات والترجمة. تعني: مهارة القراءة و أسئلة مهارة الكتابة و أسئلة، مفردات و سؤالاً، ترجمة وأسئلة، قواعد اللغة و سؤالاً، والثقافة سؤالين.
- 2- درجة صدق بنود الأسئلة في اختبار اللغة العربية الشامل الوطني للسنة 2015 هي: كان هذ الاختبار غير الصادق نظرا لكثرة الأسئلة غير الصادقة في هذا الاختبار. 40% من 50 سؤالاً تعني 20 سؤالاً هي أسئلة صادقة. و أما الباقي هو 60% تعني 30 سؤالاً هي أسئلة غير صادقة.
- 3- درجة ثبات بنود الأسئلة في اختبار اللغة العربية الشامل الوطني للسنة 2015 هي: كانت النتيجة (r statistik /Spearman Brown) أكبر من نتيجة (r tabel). وتلك النتيجة تدل على وجود الإتصال على شكل ملحوظ بين X و Y. إذا، ذلك التحليل يدل على هذ الاختبار له درجة الثبات الداخلية العالية.
- 4- درجة صعوبة بنود الأسئلة في اختبار اللغة العربية الشامل الوطني للسنة 2015 هي: تتكون من 50% أسئلة في درجة متوسطة، و 42% أسئلة في درجة سهولة، و 4% في درجة صعوبة. إذا أستطيع أن ألخص بأنّ 50% أسئلة في هذا الاختبار هي أسئلة جيّدة. وأما الباقي من الأسئلة الموجودة في هذا الاختبار هي أسئلة غير جيّدة.
- 5- درجة تمييز بنود الأسئلة في اختبار اللغة العربية الشامل الوطني للسنة الدراسية 2015 هي: 44% أسئلة في درجة جيّدة جدّاً ، و 0% أسئلة في درجة جيّدة ، و 16% في درجة متوسطة، و 18% أسئلة في درجة منقحة، و 4% أسئلة في درجة

مرفوضة. إذا أستطيع أن ألخص بأن بعض الأسئلة في هذا الاختبار تعني 50% أسئلة هي أسئلة جيّدة.

6- درجة تشتت بنود الأسئلة في اختبار اللغة العربية الشامل الوطني للسنة الدراسية 2015 هي: 62% هي بدائل الأجوبة القبيحة من ناحية الدرجة التشتت تعني: 124 بدائل الأجوبة. وأمّا الباقي تعني 76 بدائل الإجابة لها درجة التشتت الجيدة. كما حسبت الباحثة كان هذا الاختبار يتكون من 200 بدائل الأجوبة.

ب- الاقتراحات

بناء على نتائج البحث تقدم الباحثة الاقتراحات لجميع مدرسين ومدرسات اللغة العربية أن يعمل تقويم بنود الأسئلة بتحليل ورقة أجوبة الطلاب لمعرفة جودة ذلك الاختبار. والحاصل من ذلك التحليل سوف يكون ملاحظة لمواضع الأسئلة في اختبار اللغة العربية لتصميم أحسن الأسئلة في الاختبار الآتي.

المراجع

- أنسي محمد أحمد قاسم، الفروق الفردية والتقويم، عمان: دار الفكر، 2003.
- رائد خليل العبادي، الاختبارات المدرسية، عمان: مكتبة المجتمع العربي، 2005.
- رجاء وحيد ويدر، البحث العلمي أساسياتها لنظرية وممارستها العملية، دمشق: دارالفكر، 2008.
- رشدي أحمد طعيمة، تعليم اللغة العربية لغير الناطقين بها: مناهجه وأساليبه، مصرى: جامعة المنصورة، 1989.
- عبد الخالق محمد، اختبار اللغة، الرياض: جامعة الملك سعود، 1089.
- عبد الرحمن بن إبراهيم الفوزان، الرياض: مؤسسة الوقف الإسلامي، 1424هـ
- محمد علي الخولي، الاختبارات اللغوية، صولنج: دار الفلاح للنشر والتوزيع، 2000.
- محمود كامل الناقه ورشدي أحمد طعيمة، طرق التدريس اللغة العربية لغير الناطقين بها، الرياض: مطبع المعارف الجديدة، 2003.
- Ainin, Moh., *Evaluasi dalam Pembelajaran Bahasa Arab*, 2006. Malang: Misykat.

- Ali, Muhammad, *Penelitian Pendidikan: Prosedur dan Strategi*, 1987. Bandung: Angkasa.
- Arikunto, Suharsimi, *Dasar-Dasar Evaluasi Pendidikan Edisi 2*, 2012. Jakarta: PT Bumi Aksara.
- Arikunto, Suharsimi, *Prosedur Penelitian Suatu Pendekatan Praktek Edisi Revisi*, 2010. Jakarta: PT. Rineka Cipta.
- Djalal, fahrudin, M., *Penelitian dalam Pengajaran Bahasa Asing/Arab*, 1989. Malang: IKIP Malang
- Djiwantoro, Soenardi, M., *Tes Dalam Pengajaran*, 1996. Bandung: ITB.
- Hamid, Abdul, *Mengukur Kemampuan Bahasa Arab Untuk Studi Islam*, 2013. Malang: UIN Press.
- Jakarta: .Joesmani, *Pengukuran dan Evaluasi dalam pengajaran*, 1998 Depdikbud.
- Keputusan Direktur Jendral Pendidikan Islam Nomor 32: 2015.
- Moleong, Lexy, L, *Metode Penelitian Kualitatif*, 2005, Bandung: Rosda Karya,
- Nanar, Muhammad, *Metode penelitian*, 2009, Ghalia Indonesia.
- Nuruddin, dkk, *Ujian Nasional di Madrasah Persepsi dan Aspirasi Masyarakat*, 2007, Jakarta: Gaung Persada Press.
- Prawiradilaga, Dewi Salma. *Prinsip Disain Pembelajaran*. 2007. Jakarta: Prenada Media Group.
- Sugiono, *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*, 2010. Bandung: Alfabeta.
- Sudjana, Nana, *Hasil Proses Belajar Mengajar*. 2001. Bandung: Remaja Rosdakarya.
- Thoha, Chabib, M., *Teknik Evaluasi Pendidikan*, 2003. Jakarta: PT. Raja Grafindo Persada.
- Abd Ibrahim Syukur, *Asesmen Pembelajaran Bahasa*, 2012, Wahyuni Sri dan Bandung: PT. Refika Aditama