

موصفات الاختبار الجيد

إعداد: نور الهداية

جامعة رادين فتاح الإسلامية الحكومية فالمبانج
nurulhidayah@radenfatah.ac.id

مستخلص

لكل من أراد أن يضع الاختبار أن يفهم الأهداف منه حتى يستطيع أن يعده بشكل مناسب. يعد الاختبار وسيلة من الوسائل المهمة المستخدمة في قياس وتقويم كفاءة الطلبة، ومعرفة ما وصل إليه مستواهم التحصيلي، ومن ناحية أخرى تساعد في معرفة مدى تحقق الأهداف السلوكية، أو النواتج التعليمية المطلوبة. هذه الموصفات الخمسة تعني الموصفات الأساسية التي ينبغي أن تتوفر في كل الاختبار وهي الثبات، الصدق، سهولة التطبيق، التمييز، الموضوعية. أما الموصفات الثانوية تتكون من التمثيل، الوقت، التعليمات، التدريب، الشكل. بعد أن فهم المعلم هذه الموصفات، لزم المعلم أن يقوم به حين إعداد الأسئلة لكل المستوى التعليمي

الكلمات الأساسية: الاختبار، الموصفات، الأسئلة

Abstrak

Setiap yang akan mempersiapkan tes maka dianjurkan untuk memahami tujuan – tujuan yang ada di dalam tes tersebut sehingga guru dapat mempersiapkan dengan bentuk yang sesuai. Tes merupakan media yang penting digunakan untuk mengukur dan mengevaluasi kemampuan siswa, dan mengetahui tingkat prestasi belajar, selain itu tes juga digunakan untuk mengukur tujuan psikomotorik atau hasil pembelajaran yang diharapkan. Lima Karakteristik yang dimaksudkan sebagai karakteristik dasar yang harus terdapat di dalam sebuah tes yaitu reabilitas, validitas, diskriminasi, kepraktisan, dan obyektifitas. Adapun karakteristik sekunder yang selayaknya dimiliki oleh tes adalah penyajian, waktu, instruksi, gradasi, dan bentuk. Setelah guru memahami dengan baik karakteristik ini maka guru bisa membuat soal sesuai dengan kemampuan belajarnya

Kata Kunci: Tes, Karakteristik; soal

مقدمة

كان الاختبار في المدارس وسيلة للقياس و حكما تربويا على انجازات الطالب في المدرسة، وهي أداة يستخدمها المعلم لقياس ما تعلمه الطالب في المدرسة، ولا ينفصل عن عمل المدرس في المدرسة أو الجامعة. ومسؤولية هذا الأمر تقع بالدرجة الأولى على المعلم. إذن المعلم لا يتقن كيف طريقة التدريس فحسب بل لازم أن يعرف كيف يعد الاختبار الجيد أيضا حتى نجح الطلبة في نيل أهداف تعليم اللغة العربية جيدا. كما أن هناك الأخطاء الشائعة في الاختبار

وسوف أكثر من المعلمين لم يفهموا عن وظيفة الاختبار، حتى قامت الباحثة بتحليل مواصفات الاختبار خصوصا في أسئلتها.

كما قال طعيمة أن الاختبار اللغوي يعني مجموعة من الأسئلة التي يطلب من الدارس أن يستجيب لها بهدف قياس مستواه في مهارة لغوية معينة وبيان مدى تقدمه فيها ومقارنته بزملائه (طعيمة، ١٩٨٩، ٢٨٤). والاختبار اللغوي متعلق بمختلف المهارات اللغوية الأربعة وبمختلف عناصر اللغة الأربع. وبعد قيام الاختبار يعرف المعلم عن مستوى المهارات أم كفاءة اللغوية لدى الطلبة. ولا يمكن أن يكون هذه الاختبارات سواء في ظهرها وأيضا في محتواها. بناء على ذلك، حاولت الباحثة أن تبحث الأشياء التي تتعلق بمواصفات الاختبار الجيد حتى نعرف خصائص الاختبار الجيد ونستطيع القيام به في عملية تعليمنا الحاضر.

البحث

لكي يوصف الاختبار بأنه جيد وخال من الثغرات اللغوية والفنية يجب أن يتميز بعدة صفات عرفها وقتن لها خبراء الاختبارات من علماء اللغة ، ووجدت الباحثة اختلاف آراء الخبراء عن أنواع المواصفات الجيدة،

١- المواصفات الأساسية عند آراء الخبراء

قال طعيمة هذه المواصفات الخمسة تعني المواصفات الأساسية التي ينبغي أن تتوفر في كل الاختبار (طعيمة، ١٩٨٩، ٧٢٠) أولئك هم:

أ. الثبات (*reliability*)

يقصد بالثبات يعني 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 يتغيرا. وعلى هذا فإن ثبات الاختبار يرتبط إلى حد كبير بثبات التقدير العام أو حتى الدرجات التي يحرزها الدارس نفسه، فإذا ما تذبذبت درجاته فإن هذا يعني أن المقياس أو الاختبار لا يتصف بالثبات (محمد عبد الخالق محمد، ١٩٨٩، ٤٣)

وإذا كان الاختبار صادقاً وقياس حقا ما صمم من أجله فسوف تتميز نتائجه بالثبات، ولكن لا يمكن الجزم بأن يكون العكس هو الصحيح، إذ ليس كل اختبار ثابت صادقاً. لذا فمن الأفضل دائماً أن ننظر في أمر صدق الاختبار قبل أن نشرح في الحصول على ثباته حتى لا تضيع الجهود السدى. يمكننا الحصول على معامل ثبات الاختبار بعدة طرق منها (محمد عبد الخالق محمد، ١٩٨٩، ٤٤-٤٦) :

١- إعادة تطبيق الاختبار، وذلك بأن يطبق الاختبار نفسه للمرة الثانية لمجموعة الدارسين أنفسهم. وبعد الحصول على النتائج نحسب الارتباط بين النتائج التي حصل عليها الدارسون في المرتين. فإذا كانت درجة الارتباط عالية يمكننا القول بأن الاختبار يتمتع بدرجة ثبات مؤقتة ممتازة.

٢- إعادة اختبارين في وقت واحد، ويكون ذلك بتقديم اختبارين متطابقين أو شبه متطابقين من حيث المادة ومستوى السهولة والصعوبة وصياغة التعليمات، ومتساويين في الطول والزمن ومتساويين حتى في شكل الاختبار وبنائه. ويجرى الاختباران على المجموعة نفسها من الدارسين، ويصححان بالطريقة نفسها.

٣- تجزئة الاختبار، يجرى اختبار واحد ولمرة واحدة فقط ثم نقوم بتجزئة الاختبار إلى نصفين. فإن كان لدينا اختبار يحتوي على مئة بند فيكون للنصف الأولى الخمسون بنداً الأولى، وللنصف الثاني الخمسون بنداً الثانية. ولكن الأفضل أن يجرأ إلى نصفين على أساس الأسئلة ذات الأرقام الفردية والزوجية.

٤- طريقة تحليل التباين، وفيها يعطى الاختبار مرة واحدة، ثم تحسب نسبة الدارسين الذين وفقوا في كل بند على حدة، ونسبة الدارسين الذين أخفقوا فيه ونحصل على معامل الثبات باستخدام صيغة كودر-ريتشاردسون كالآتي:

$$\text{ث} = \frac{\text{ب} \cdot \sum \text{ح}^2}{\text{ح} \cdot (\text{ب} - 1)}$$

$$\begin{aligned} \text{حيث} \quad \text{ث} &= \text{معامل ثبات الاختبار} \\ \text{ب} &= \text{عدد بنود الاختبار} \\ \text{نج} &= \text{نسبة الطلاب الذين نجحوا في البند} \\ \text{نخ} &= \text{نسبة الطلاب الذين أخفقوا في البند} \\ \text{ح}^2 &= \text{الانحراف المعياري للاختبار} \\ \Sigma &= \text{مجموع} \end{aligned}$$

ب. الصدق (*validity*)

إن صدق الاختبار يعني إلى أي مدى يقيس الاختبار الشيء الذي وضع من أجله. لتوفير درجة صدق عالية للاختبار يجب أن يكون المحتوى ذا علاقة بالشيء الذي يراد قياسه، ويجب استبعاد أية مشكلات ثانوية لاعلاقة لها به، والتي قد تكون في ذاتها أكثر صعوبة من الصعوبات ذات الصلة بالاختبار. وكلما كان الاختبار قائما على تحليل دقيق للمهارة أو العناصر التي نحن بصدد اختبارها، وكلما كانت درجاته ترتبط ارتباطا عاليا بالقدرات التي نود قياسها، كان ذلك مؤشرا على أن الاختبار صادق وأنه يؤدي إلى الهدف المطلوب

وفي تعليم اللغة، إذ أراد مدرس المحادثة معرفة مهارة الكلام من الطلاب ثم عرض إليهم أسئلة في الورقة ليجيبوها تحريريا فلن يصل إلى غرضه ويكون الاختبار الذي يقوم به طلابه غير صادق لأن الإجابة التحريرية لا تقيس مهارة كلامهم أصلا، بل الصحيح يقوم المدرس بالحوار مع طلابه حول موضوع ما حتى يعرف ويقيس كفاءتهم اللغوية الشفهية (أسراري، ١٢، ٢٠١٣)

والصدق نفسه ينقسم إلى ثلاثة أنواع (Soenardi, 2008, p. 165)، أكثرها أهمية

لمدرس اللغة الذي يضع اختبارته بنفسه وهي:

١- الصدق الظاهري (*Face Validity*)

هو من الطرق الشائعة لمعرفة مدى مناسبة وصلة الاختبار بالهدف الذي يريده واضع الاختبار. ويعني أن الاختبار يبدو مناسباً للهدف الذي وُضع من أجله، وهو ليس صادقاً بالمعنى العلمي للكلمة، لأنه يدل على ما يبدو أن الاختبار يقيسه من الظاهر، لا على ما يقيسه الاختبار بالفعل ((كاظم مرشد ذرب الزبيدي).

نظر الصدق الظاهري بنظرة سريعة إلى أسئلة الاختبار لأن أسئلة الاختبار واضحة ولها علاقة بالأهداف التي أرادها المعلم أو واضع الاختبار، فإذا كان هدف الاختبار قياس التعبير الشفهي من الطالب فكل سؤال في الاختبار يتعلق به. إذا حدث ذلك فيكون الاختبار ذا صدق ظاهر. المثال الآتي لاختبار مهارة الكلام تظهر فيه الأسئلة الآتية: تكلم عن أسرتك!، ماذا تعرف عن شهر رمضان، تحدث عنه!.

٢- صدق المحتوى (Content Validity)

يعنى في المقام الأول بما يجب أن يتضمنه الاختبار، وما يتضمنه الاختبار يعتمد على تحليل جيد اللغة المراد اختبارها وعلى تحليل أدق للمهارة قيد الاختبار ولأهداف الدورة. لذا يجب أن تغطي بنود الاختبار بالتقريب كل الجوانب الرئيسة في المقرر المدروس، وتكون بحسب نسب ثقل كل منها فيه، وبحيث تكون العلاقة بين بنود الاختبار وأهداف الدورة دائماً واضحة. وعلى سبيل المثال نؤتى بعض السؤوال في الاختبار عن العنوان للطلاب في الفصل السابع في المرحلة الواحدة، ولكن في المنهج، هذه المادة تعطى بالمعلم في المرحلة الثانية. إذا حدث ذلك فيكون الاختبار غير صادق محتوي (محمد عبد الخالق محمد، ١٩٨٩، ص ٥٠).

٣- الصدق التجريبي أو الصدق الإحصائي (Empirical Validity)

حصل على هذا النوع من الصدق عن طريق مقارنة نتائج الاختبار بنتائج معيارية أخرى يعتقد أنه صادقة، مثل اختبار آخر ثبت أنه صادق أو في ضوء

تقديرات مدرس الصف التي تعطى في نهاية المقرر الدراسي، أو أي مقياس آخر مستقل. فإذا دلت التجربة على أنه يوجد ارتباط عال بين علامات الاختبار وأي معيار من المعايير المذكورة سلفا نكون قد حصلنا على ما يسمى بالصدق التجريبي. وهذا الصدق نوعان: صدق تنبئي (*predictive validity*) وصدق تلازمي (*concurrent validity*).

فإذا استعملنا على سبيل المثال اختبار اللغة العربية بوصفها لغة أجنبية لاختيار المتقدمين للكليات الجامعية، ثم بعد التصحيح قمنا باستخراج معامل الارتباط بين دراستهم بالكلية نكون بذلك قد قمنا بتحديد الصدق التنبئي للاختبار، أي أن هذا النوع من الصدق ينبئنا بالنجاح المستقبلي. أما إذا قمنا بعد استعمال الاختبار وتصحيحه مباشرة بتقويم كفاية كل دارس على حده مقارنة بمعيار آخر في الفترة نفسها ثم أوجدنا معامل الارتباط نكون بذلك قد قمنا بتحديد الصدق التلازمي للاختبار (محمد عبد الخالق محمد، ١٩٨٩، ص ٥١).

ث. سهولة التطبيق

قد يتمتع الاختبار بدرجة ثبات وصدق عاليتين إلا أنه لا يمكن تطبيقه لسبب من الأسباب التي تتصل بالتصحيح، والإمكانات المادية، وعدم توافر الظروف التي يتطلبها إجراء الاختبار من حيث الوقت، وعدم توافر أجهزة بعينها يستلزمها تطبيق الاختبار وتخرج عن إمكانات وقدرات الجهة المنفذة له. لذا كان من اللازم أن يضع الشخص المكلف بوضع الاختبار بعض الاعتبارات العلمية نصب عينيه ومن هذه الاعتبارات: التصحيح، الاقتصاد، إجراء الاختبار وتطبيقه (محمد عبد الخالق محمد، ١٩٨٩، ص ٥٢)

ث. التمييز (*discrimination*)

الاختبار الجيد يميز بين مستويات الطلاب المختلفة. فالاختبار الذي يأخذ فيه جميع الطلاب صفرا أو مئة أو درجة موحدة ليس اختبارا مميزا، وليس هو بالاختبار

- المطلوب (محمد علي الخولي، ٢٠٠٠، ص ١٦٨) وهناك طرق كثيرة لحساب معامل التمييز وسيكتفى هنا بذكر أحدها وتتلخص في الخطوات الآتية (Thanwiya):
- ١- ترتب أوراق الطلاب تصاعدياً حسب الدرجات ، ونفترض أن عددها ١٠٠ ورقة
 - ٢- قسم الأوراق إلى مجموعتين عليا ودنيا وتمثل أعلى ٢٧% من الأوراق ذات الدرجات العليا ، وأدنى ٢٧% منها ذات الدرجات الدنيا .
 - ٣- وبذلك يكون عدد أفراد المجموعة العليا = عدد أفراد المجموعة الدنيا = ٢٧ (وإذا كان عدد الطلاب قليلاً فيمكن تقسيمهم إلى مجموعتين أعلى ٥٠% وأدنى ٥٠%). يتم حصر عدد الطلاب الذين أجابوا عن السؤال الأول إجابة صحيحة من بين أولئك الذين حصلوا على الدرجات العليا . ونفترض أنهم كانوا ٢٠ طالباً.
 - ٤- يتم حصر عدد الذين أجابوا عن السؤال الأول إجابة صحيحة من بين أولئك الذين حصلوا على الدرجات الدنيا . ونفترض أنهم كانوا ١٢ طالباً.
 - ٥- تطرح الخطوة (٤) من الخطوة (٣). (الفرق بين المجموعتين ٢٠ - ١٢ = ٨).
 - ٦- يقسم الناتج من الخطوة (٥) على عدد أفراد إحدى المجموعتين.

س - ص
ن

يكون معامل التمييز

- حيث س : عدد طلاب الفئة العليا في التحصيل الذين أجابوا على السؤال إجابة صحيحة
- ص : عدد طلاب الفئة الدنيا في التحصيل ممن أجابوا على السؤال إجابة صحيحة
- ن : عدد افراد إحدى المجموعتين.

ج. الموضوعية (objectivity)

ويقصد بها عدم تأثير شخصية المصحح على وضع أو تقدير علامات الطلاب في الاختبار. ومما ساعد على تحقيق الموضوعية أن يفهم الطلاب تعليمات الاختبار بدقة وأن يكون هناك تفسير واحد للأسئلة والإجابة المطلوبة منه فضلا عن توفير الظروف المادية والنفسية للطلاب لأداء الاختبار (طعيمة، ١٩٨٩، ص ٧٢٠).

ومن أجل ذلك، لا بد على المعلم أن يعد دليلا أو معيارا خاصا ويعتمد عنه في تصحيح إجابات الأسئلة، لأن خوفا من أثار شخصياته. فهذا الدليل معد من موافقة المعلمين، وأرادت الباحثة أن تعطي المثال في تصحيح الإنشاء الموجه لدى الطلبة في الصف الثالث بالمدرسة الثانوية:

رقم	المهات المقيسة	١	٢	٣	٤	٥
١	مناسبة محتوى الإنشاء بالموضوعات الموجودة					
٢	صحة في كتابة الأحرف الهجائية					
٣	صحة في استخدام القواعد اللغوية					
٤	عدد الكلمات المكتوبة					

وبعد يحاسب المعلم النتيجة بهذا الدليل ثم يدخلها بالرموز:

$$\text{النتيجة الأخرى} = ٤ \times \frac{\text{النتيجة المحسولة}}{\text{النتيجة الأعلى}}$$

٢- المواصفات الثانوية

وللاختبار الجيد خصائص أخرى كما ذكرها الخولي، وهي كما يلي (محمد علي الخولي، ٢٠٠٠، ص ١٩) :

أ- التمثيل

لاختبار الجيد يمثل المادة الدراسية موضع الاختيار تمثيلاً متوازناً. الاختبار الجيد يستثني الأجزاء المعادة من المادة أو الأجزاء التي درسها الطلاب في سنوات سابقة ويركز على الأجزاء الجديدة من المادة كلما كان ذلك ممكناً.

ولتحقيق التمثيل هناك عدة طرق: أولاً، طريقة المستجدات يعني تتطلب هذه الطريقة حصر الجوانب الجديد في المادة وتركيز الاختبار على هذه الجوانب. ثانياً، الطريقة النسبية يعني عندما نريد أن نضع اختباراً يتكون من مئة بند موضوعي من كتاب يتكون من ثلاث مئة صفحة، فإذا نضع بنداً واحداً من كل ثلاث صفحات.

ب- الوقت

الاختبار الجيد يراعي عامل الوقت. يجب أن يكون الوقت المحدد للاختبار كافياً للطالب المتوسط كي يجيب عن أسئلة الاختبار بكيفية مريحة. قد يفشل الاختبار إذا كان وقته أقصر كثيراً أو أطول كثيراً من الوقت الكافي. الوقت الأقصر يحرم الطلاب من فرصة الإجابة عن جزء من أسئلة الاختبار، والوقت الأطول قد يؤدي إلى بعض الفوضى في إدارة الاختبار وإجرائه وقد يؤدي إلى حالات من الغش.

وباستطاعة المعلم أو الفاحص أن يقرر الوقت الكافي للاختبار عن طريق تقدير الوقت الكافي للإجابة عن كل سؤال أو بند على حدة مع مراعاة احتساب الوقت بالنسبة للطالب المتوسط، وليس بالنسبة للطالب المتفوق (محمد علي الخولي، ٢٠٠٠، ص ٢٠)

ج- التعليمات

يجب أن تكون تعليمات الاختبار واضحة محددة مكتوبة في ورقة الاختبار ذاتها. والتعليمات نوعان (محمد علي الخولي، ٢٠٠٠، ص ٢١)

١- تعليمات عامة تخص الاختبار كله بوجه عام، من أمثلة التعليمات العامة:

أ- أجب عن جميع الأسئلة التالية.

ب- أجب عن خمسة أسئلة مما يلي.

ج- وقت الاختبار ساعتان.

د- يجوز استخدام المعجم.

٢- تعليمات خاصة بكل سؤال على حدة. ومن أمثلة التعليمات الخاصة :

أ- إقرأ القطعة التالية وأجب عن الأسئلة التي بعدها في الفراغ المحدد لكل جواب.

ب- ما مرادف كل كلمة مما يلي؟ اكتب الجواب في الفراغ المحدد!

ج- اكتب عن موضوع كذا فقرة تتكون من عشرة جمل تتراوح كلماتها بين ١٠٠-١٥٠ كلمة.

د- اختر الجواب الصحيح وضع دائرة حول حرفه، أي حول أ، ب، ج،

د

٥- التدرج

الاختبار الجيد سهل التدرج، بعيد عن التعقيد. وهذا يستدعي عدة أمور منها (محمد علي الخولي، ٢٠٠٠، ص ٢٢) :

١- أن يصمم المعلم مفتاحا للإجابات عند بناء الاختبار ذاته يلتزم به عند التدرج، وهذا يجعل التدرج سهلا ثابتا موضوعيا.

- ٢- أن يحدد المعلم عند بناء الاختبار درجة كل سؤال ودرجة كل بند فيه، أي الدرجة النسبية مقارنة بالأسئلة الأخرى في الاختبار ذاته.
- ٣- أن تتساوى أوزان الأسئلة من حيث درجاتها إذا تساوت في وقت الإجابة. لا تختلف الأوزان النسبية للأسئلة إلا لسبب يمكن الدفاع عنه.

و- الشكل

- الاختبار الجيد ذو شكل منسق تراعى فيه الأمور التالية (محمد علي الخولي، ٢٠٠٠، ص ٢٣) :
- ١- يبدأ بمعلومات المقدمة التي تبين اسم المادة الدراسية واسم القسم ومدة الاختبار واسم أستاذ المادة واسم المدرسة أو الجامعة أو الكلية وتصنيف الاختبار.
 - ٢- ينقسم الاختبار إلى أجزاء وأسئلة وبنود واضحة التقسيم يتسلسل ترقيمها على نحو جيد. وعلى سبيل المثال، يستعان بالكلمات لترقيم الأجزاء: الجزء الأول، الجزء الثاني،... إلخ. أما الأسئلة فترقم باستخدام الحرف: أ، ب، ج،... إلخ.
 - ٣- تظهر على ورقة الإجابة (التي قد تكون ورقة الأسئلة ذاتها) الفراغات المحددة للإجابات.
 - ٤- يدقق الاختبار طباعياً تدقيقاً وافياً حتى يخلو من الأخطاء الطباعية أو سواها كيلا تنشأ مشكلات في أثناء إجراء الاختبار.
 - ٥- يفضل أن يكون الاختبار مطبوعاً على أن يكون مكتوباً بخط اليد لتحقيق قدر أوفر من الوضوح والترتيب والاتساق

الخلاصة

وجدنا عشرة مواصفات الاختبار الجيد وتقسم إلى قسمين المواصفات الأساسية و المواصفات الثانوية. المواصفات الأساسية هي الثبات، والصدق، وسهولة التطبيق، والتمييز،

والموضوعية، والمواصفات الثانوية هي التمثيل، والوقت، والتعليمات، والتدريب، والشكل. سوف المعلم أن يطبق هذه المواصفات عندما يقوم بالاختبار كي نجد الانجازات الجيدة في التعليم.

المراجع

Ainin,. muhammad T. dan M. (2013). *evaluasi pembelajaran bahasa arab*. malang.

Soenardi, D. (2008). *Tes Bahasa, Pegangan bagi Pengajar bahasa Arab*. Jakarta: Indeks.

Thanwiya. (n.d.). No Title.

طعيمة, ر. أ. (1989a). *تعليم العربية لغير الناطقين بها مناهجه وأساليبه*.

طعيمة, ر. أ. (1989b). *المرجع في تعليم اللغة العربية لغير الناطقين بها*. (2, Ed.). مكة.

كاظم مرشد ذرب الزبيدي. (n.d.). *الصدق الظاهري*.

محمد عبد الخالق محمد. (١٩٨٩). *اختبار اللغة. الرياض: جامعة الملك السعود*.

محمد علي الخولي. (2000a). *أساليب تدريس اللغة العربية. الأردن: دار الفلاح للنشر و التوزيع*.

محمد علي الخولي. (2000b). *الاختبارات اللغويات. الأردن: دار الفلاح للنشر و التوزيع*.