أسس إعداد تعليمات الاختبار، وتحليل بنود الاختبار

*بحث في القياس والتقويم التربوي*

*إعداد/ شادية بيومي حامد عطية*

*قسم التربية*

*كلية التربية– جامعة المدينة العالمية*

شاه علم - ماليزيا

*shadia@mediu.ws*

*خلاصة*—هذا البحث يبحث في أسس إعداد تعليمات الاختبار، وتحليل بنود الاختبار.

*الكلمات المفتاحية: المراجعة، الاختبار، المقارنة، البنود*.

# ***المقدمة***

معرفة أسس أسس إعداد تعليمات الاختبار، وتحليل بنود الاختبار والمقصود بها هي مراجعة الاختبار للكشف عن نواحي القوة ونواحي الضعف فيه، يتم ذلك من خلال تحليل بنود أو فقرات الاختبار، من حيث مستوى صعوبة كل بند، وقوة تمييز البند. المقصود بمستوى الصعوبة هنا: النسبة المئوية لعدد المتعلمين الذين تمكنوا من الإجابة عن ذلك البند بصورة صحيحة، أو قوة التميز لهذا البند، فهي قدرة كل بند من البنود على التمييز بين المتعلمين الذين حصلوا على درجات عالية في الاختبار، والمتعلمين الذين حصلوا على اختبارات منخفضة.

1. *المقالة*

الأسس التي يجب أن نضعها في الاعتبار عند إعداد تعليمات الاختبار:

1- أن تكون تلك التعليمات مرتبطة بنوع الأسئلة: يقصد بذلك: أن يكون لكل نوع من أنواع الأسئلة تعليمات خاصة به، إلى جانب التعليمات العامة في أول الورقة، مثلًا: الورقة متضمنة خمسة أسئلة، المفروض أن يجاوب المتعلم عن أربعة فقط من تلك الأسئلة. أيضًا هناك تعليمات خاصة بالورقة، تلك التعليمات تشير إلى زمن الاختبار الكلي، تشير إلى عدد الأسئلة التي يتم اختبارها، تشير للصف الذي يتم فيه تطبيق ذلك الاختبار، تشير إلى الزمن الكلي للاختبار، فتلك تعليمات عامة.

2- أيضًا تعليمات خاصة بكل سؤال على حِدة، توضح أسلوب الإجابة عن ذلك السؤال، وزمن كل سؤال، ودرجة كل سؤال. أيضًا يمكن إعطاء أمثلة بطريقة الإجابة على الأسئلة الموضوعية، وذلك بالنسبة للمرحلة الابتدائية بالذات.

3- تقدير كيفي لدرجات كل سؤال، وما هي الدرجة التي سوف يحصل عليها الطالب؟ يجب أن تكون التعليمات في ورقة الأسئلة بالنسبة لجميع الصفوف لو كانت في بدايتها، لو كانت تلك التعليمات عامة، ولو كانت تلك التعليمات خاصة كل التعليمات أعلى السؤال المرتبطة به.

4- التعليمات أيضًا توضح كيفية الإجابة عن الاختبار، وهل يتم تسجيل الإجابة على ورقة أسئلة بصورة مباشرة، وفي هذه الحالة ما هو المكان المحدد للإجابة؟ لو كانت ورقة الإجابة هي ورقة الأسئلة، إذًا محدد فيها مكان الإجابة.

5-هناك مرحلة عمرية تتطلب أن تكون ورقة الأسئلة في ورقة الإجابة: وهي المرحلة المتقدمة لدى المتعلمين الصغار، ولكن يتم فصلُها في المراحل المتأخرة الدراسية، في المرحلة المتوسطة، والمرحلة الثانوية، والجامعية طبعًا. إذًا، فاصل الورقة يتوقف عند عمر المتعلم.

6- طبيعة الاختبار نفسه: فلو كان هذا الاختبار الإكمال يفضل أن تكون الإجابة على نفس الورقة، إذا كان الاختبار اختيارًا من متعدد تُفضل الإجابة على نفس الورقة، إذا كان الاختبار صوابًا وخطأً أيضًا تفضل الإجابة على نفس الورقة.

فإنه يمكن القول بشكل عام: إن ورقة الإجابة المنفصلة تساعد على دقة أكبر في تقدير الدرجات، والحصول على ثبات أكبر للدرجات، وتساعد على استخدام ورقة الأسئلة أكثر من مرة، ولكن عند استخدام ورقة منفصلة يجب أن تنظم بشكل خاص، وأن تكتب تعليمات عامة أعلى الكراسة، ومعلومات وبيانات خاصة بكل سؤال داخل كراسة الاختبار.

7- إعداد مفتاح الإجابة: لهذا يجب إعداد مفتاح الإجابة عقب إعداد ورقة الأسئلة مباشرة، وقبل موعد الاختبار. إذًا على المعلم أن يعد مِفتاحًا للإجابة عقب إعداد ورقة الأسئلة مباشرة، وذلك قبل موعد الاختبار بفترة كافية؛ ولهذا يجب على المعلم اتباع خطوات محددة؛ حتى يتأكد من سلامة مفتاح الإجابة. مفتاح الإجابة ييسر على المعلم عملية التصحيح، ومراجعة المفتاح للاطمئنان على صحته.

وهناك عدة طرق متنوعة للمراجعة، منها: أن يقوم معلم آخر بمراجعة ذلك المفتاح، أو عمل مفتاح آخر، ويتم المقارنة بين المفتاحين؛ للتأكد من صحة المفتاح. إذًا مفتاح الإجابة مهم جدًّا، يتم إعداده عقب إعداد الاختبار. عند إعداد المفتاح يجب مراعاة تساوي أوزان الأسئلة، وهذا يصدق بوجه عام أو خاص على الأسئلة الموضوعية، وأسئلة الإكمال، وأسئلة التعرف بأنواعها المختلفة، وإعطاء الوزن أوزانًا مختلفة الأسئلة معناه في الواقع أن المعلم يعتقد أن معرفة بعض المفاهيم أهم من المفاهيم الأخرى، أو إدراك تلك المفاهيم لها أثر مباشر في عملية التعلم، وتوجيه سلوك المتعلم.

المقصود بتحليل بنود الاختبار: هي مراجعة الاختبار للكشف عن نواحي القوة ونواحي الضعف فيه، يتم ذلك من خلال تحليل بنود أو فقرات الاختبار، من حيث مستوى صعوبة كل بند، وقوة تمييز البند. المقصود بمستوى الصعوبة هنا: النسبة المئوية لعدد المتعلمين الذين تمكنوا من الإجابة عن ذلك البند بصورة صحيحة، أو قوة التميز لهذا البند، فهي قدرة كل بند من البنود على التمييز بين المتعلمين الذين حصلوا على درجات عالية في الاختبار، والمتعلمين الذين حصلوا على اختبارات منخفضة.

مدى الصعوبة، ومدى السهولة هذه من خلال مجموعة صاد على نون في مائة، يعني: تدل هنا على نسبة عدد الناجحين في ذلك البند، نوع الطلبة الذين أجابوا إجابة صواب أو صحيحة، مثلًا: لدينا عدد خمسين طالبًا، ولدينا سؤال، الخمسون طالبًا إجاباتهم عن هذا السؤال صحيحة، فالمجموع هنا خمسون حصلوا على الإجابة صحيحة خمسون، يتم قسمتهم على عدد الطلاب، خمسون ضرب مائة نحصل على ناتجة مائة في المائة، معنى ذلك: أن مؤشر الصعوبة هنا يعتبر منخفضًا؛ لأن الإجابة تامة لكل المتعلمين، فتقدير درجات جميع الأوراق يجب أن تكون درجة كل متعلم مساوية لعدد الإجابات الصحيحة أيضًا في الاختبار، يجب أن نضع الجزئية هذه في الاعتبار، هي الخطوات الخاصة لتحديد مستوى الدرجات.

أولًا: كل درجة حصل عليها الطالب في الورقة تمثل عدد الإجابات الصحيحة، ثم يتم ترتيب أوراق الإجابة ترتيبًا تنازليًّا من أعلى درجة إلى أقل درجة في الاختبار، معنى ذلك: أننا حسبنا معامل الصعوبة والسهولة لذلك البند فقط، أو ذلك السؤال، ولكن لو أردنا حساب ذلك بالنسبة للاختبارات ككل يتم ترتيب أوراق الإجابة ترتيبًا تنازليًّا من أعلى إلى درجة أقل، ثم نحدد المجموعة الحاصلة على أعلى درجات، والمجموعة الحاصلة على أدنى درجات.

تحديد أعلى الدرجات:

يتم تحديد أعلى درجات الحاصلين على ربع الدرجة العليا، والحاصلين على أدنى درجات الذين حصلوا على ربع الدرجة الدنيا، فلو كانت -مثلًا- الدرجة من خمسين نجد أن الطلبة الذين حصلوا من أربعين إلى خمسين هم يمثلون الدرجة العليا، الطلبة الذين حصلوا من صفر إلى عشر هم الحاصلون على الدرجة الدنيا؛ فيتم حساب معامل الصعوبة للاختبار من خلال -أولًا- نسبة عدد الناجحين في البند الكلي، حددنا بالنسبة للبند الأول، بعد ذلك يتم تحديده بالنسبة للاختبار ككل.

قوة تمييز البند:

البند هو السؤال، هل السؤال له قوة للتمييز بين المستويات العليا، والمستويات الدنيا؟ ذلك هو المقصود بقوة التمييز، معنى هذا: إن نسبة عدد المتعلمين في المجموعة العليا ونسبة المتعلمين في المجموعة الدنيا بينهم هنا نسبة واضحة ومحددة، فهنا ذلك البند ميز بين المستويات العليا والمستويات الدنيا من الطلاب. يمكن حساب مؤشر قوة تمييز البند من خلال مجموعة المستويات العليا من الطلاب الذين حصلوا على درجة عليا -الربع الأعلى- يُطرح منها مجموع الطلاب الذين حصلوا على الدرجة الدنيا، الكل على نصف العدد.

إذًا، مؤشر قوة التمييز يرمز له بالرمز تاء، تاء تساوي مجموع الحاصلين على الدرجة العليا، يطرح منها مجموعة حاصلين على الدرجة الدنيا، الكل على نصف عدد الطلاب، فلو كان عدد المتعلمين في كل مجموعة عشرة، وكان عدد الناجحين في ذلك البند من المجموعة العليا تسعة متعلمين، وعدد الناجحين من المجموعة الدنيا يمثل في نفس البند أربعة متعلمين فقط، وكان المجموع الكلي للطلاب عشرين، إذًا يتم طرح أربعة من تسعة على عشرة يساوي خمسين من مائة، ويتضح من هذه المعادلة أن الحد الأقصى لمؤشر التمييز هو واحد صحيح، والحد الأدنى هو واحد سالب واحد صحيح.

وفي هذه الحالة هنا بالموجب، وهنا بالسالب، وفي هذه الحالة الأولى يكون جميع المتعلمين في المجموعة العليا ناجحين في ذلك البند، ويكون جميع المتعلمين في المجموعة الدنيا فاشلين في هذا البند، وهذا هو الوضع الذي نريد تحقيقه قدر الإمكان من التمييز، وفي هذه الحالة يكون جميع المتعلمين في المجموعة العليا.

إذًا، يتوفر لدينا بند سالب، وتمييز سالب، وتمييز موجب، خمسون من مائة ذلك بند موجب، معنى هذا: نجاح المجموعة العليا في التمييز، أما لو كان بالسالب، معنى هذا: أن المجموعة العليا فشلت أيضًا في تمييز ذلك البند.

ما هي مستويات الصعوبة المطلوبة للبند حتى يصبح بندًا من بنود التمييز؟

من المعروف أن قوة التمييز تتأثر بمستوى صعوبة البند، قد تكون مؤشرات صعوبتها قريبة من خمسين، يكون مؤشر التمييز فيها أعلى ما يمكن، كذلك يتأثر مستوى الصعوبة بنوع الأسئلة في الاختبار؛ إذ يزيد مؤشر الصعوبة في أسئلة الصواب والخطأ، ويكون أقل ما يمكن في أسئلة الإكمال.

وهنا لدينا متوسط مؤشرات الصعوبة في كل نوع، أسئلة الإجابة بالإجابة القصيرة والإكمال، مؤشر متوسط الصعوبة في الاختبار عالي التمييز يجب أن يصل إلى خمسين، الاختبار المتعدد من خمسة بدائل مؤشر التمييز سبعون، الاختيار متعدد من أربع مؤشر التمييز أربعة وسبعون، مؤشر متعدد من ثلاثة بدائل سبعة وسبعون، الصواب والخطأ من بديلين، مؤشر التمييز خمسة وثمانون.

وبهذا يتضح لنا أن البدائل المتوفرة لها أثر في مستوى صعوبة الأسئلة؛ ولهذا نجد ما يسمى: "فعالية المشتتات"، ما هي المشتتات؟ المشتتات هي البدائل الخاطئة في الأسئلة، وهنا يتضح لنا أن أثر البدائل غير فعالة، ولا قيمة لها.

1- نسبة اختيارها في المجموعة الدنيا قليلة.

2- نسبة اختيارها في المجموعة العليا متساوية.

3- نسبة اختيارها في المجموعة العليا أكبر من نسبة اختيارها في المجموعة الدنيا.

وعلى هذا الأساس -على أساس تحليل المشتتات- يتم الكشف عن البدائل غير الفعالة؛ لتعديلها إذا ما أريد استخدام البنود القوية، نجد أن من تحليل الجدول يتضح أن الصواب والخطأ من بديلين نسبة التمييز فيه خمسة وثمانون، الاختيار من متعدد من ثلاثة بدائل سبعة وسبعون، الاختيار من متعدد من خمسة بدائل سبعون. إذًا، كلما قلت البدائل أو المشتتات كانت عملية التمييز فعالة أكثر.

وبهذا يتضح لنا أهمية صياغة الاختبار الجيد، وأثر إعداد البدائل في تحديد صعوبة وسهولة البند، فصعوبة البند من العوامل الأساسية، فالبنود يجب أن تكون ذات فعالية في الاختبار، ولكن صعوبة الفقرة بتميز قدرة المتعلمين وسهولتها. بعد الانتهاء من المقرر نجد أن تدريس المقرر أو جزء منه يتضح لنا طبيعة ذلك البند، هل يتناسب مع الاختبار، أو لا يتناسب مع الاختبار؟

وقوة التمييز أيضًا لها علاقة مهمة جدًّا في اختيار البنود، بهذا يتضح لنا أن إخراج الاختبار وإجراءه وتحليل بنوده عناصر أساسية جدًّا؛ حتى يكون ذلك الاختبار يقوم بدوره في قياس مستوى التلاميذ بموضوعية؛ بحيث تكون أسئلة جيدة ومتعددة، وأسئلة قادرة على التمييز بين الطلاب.

وبهذا يتضح لنا أن الاختبار الجيد يجب أن يتصف بعدد من المواصفات المهمة:

1- أن يكون موضوعيًّا، فلا يتأثر بعوامل ذاتية عند صياغة الأسئلة، والحكم في ذلك الاختبار يصبح حكمًا على مستوى أداء المتعلمين.

2- أيضًا زمن الاختبار، يجب أن يراعَى في الزمن جيدًا ملائمة الأسئلة للزمن الموجود، وذلك من خلال تحديد زمن أول طالب أنهَى الاختبار -التجربة- أسرع طالب في الاختبار، والزمن الذي استغرقه أبطأ طالب، إذًا الزمن الذي استغرقه أسرع طالب زائد "+" الزمن الذي استغرقه أبطأ طالب، الكل على اثنين، يحدد لنا الزمن اللازم لإجراء ذلك الاختبار بدقة وموضوعية، إذًا الموضوعية في الاختبار. ثانيًا تحديد زمن يتناسب مع حجم الاختبار.

3- الصدق، يقصد بصدق الاختبار: أن الاختبار يقيس فعلًا ما وضع لقياسه، أي: ذلك الاختبار اختبار صادق؛ لأنه مرتبط بأهداف المحتوى، وأهداف التعلم، ومرتبط يقيس في الفعل مخرجات التعلم المنشودة.

4- الثبات، فثبات الاختبار هو الذي يعطي نفس النتائج تقريبًا إذا تم إعادة تطبيقه على نفس الأفراد في نفس الظروف، ومعنى هذا: أنه إذا أعطي الاختبار لمجموعة من التلاميذ تم حساب درجة كل تلميذ في هذا الاختبار، ثم أُعيد تطبيق نفس الاختبار على نفس مجموعة التلاميذ، ولكن بعد فترة زمنية كافية، يتم الحصول على نفس النتائج، وهناك طرق لحساب معامل الثبات، ولعل أسهل الطرق هي طريقة معامل التجانس الداخلي التي يتم فيها حساب الثبات باستخدام معادلة "ريتشارد شون" التي تحدد معامل الارتباط بالاختبار، نتائج الاختبار في الاختبار الأول، والاختبار الثاني.

# المراجع والمصادر

1. اللقاني، اللقاني احمد حسين (المناهج بين النظرية والتطبيق)، عالم الكتب، القاهرة، 1981م
2. محمد حسين، آل ياسين. محمد حسين (مبادئ في طرق التدريس العامة)، بيروت، الطبعة الرابعة، 1991م
3. القصيري، القصيري. موفق عبد الله (الدليل العملي في تعليم اللغة العربية وآدابها)، ماليزيا، دار التجديد، 2006م
4. حسيني، حسيني. محمد سمير (التربية أصول وأساسيات)، القاهرة، مطبعة سعيد، 1978م
5. حامد، منصور أحمد حامد (تكنولوجيا التعليم وتنمية القدرة على التفكير والابتكار)، الكويت، دار السلاسل، 1986م