

مواصفات الاختبار الجيد

أ/ لبني زعرور

أ/ عبد الرحمان محمدي

الملخص:

نستعرض في هذا المقال مواصفات الاختبار الجيد، على اعتبار أن صياغة الاختبار أو بنائه تتطلب الدقة مهما كان نوع هذا الاختبار، إذ يجب أن يتّصف الاختبار بالموضوعية، الثبات الصدق، معامل التمييز، معامل السهولة والصعوبة، والشمولية والقابلية للاستعمال
مقدمة:

إنّ أي قصور في عملية إعداد الاختبارات التحصيلية سيؤدي حتما إلى خلل في عملية التقويم التربوي فيؤثر بدوره على العملية التعليمية التعلمية، وعليه كان من الضروري توفر هذه الأداة على خصائص ومواصفات محددة تجعل منها وسيلة صالحة للتقويم وهذه المواصفات كما يراها أحمد يعقوب النور تتمثل فيما يلي: الموضوعية، الثبات، الصدق، معامل التمييز، معامل السهولة والصعوبة، والشمولية والقابلية للاستعمال و سنوضح هذه المواصفات فيما يأتي من عناصر بالتفصيل.

1- الموضوعية:

نقصد بموضوعية الاختبار أي تحرر معده من العوامل الذاتية الشخصية عند إعداده للاختبار أو تصحيحه له. أي عدم تأثر درجات الاختبار بذاتية المصحح أو انطباعه.

لكي يصبح اختبار ما موضوعيا ينبغي التركيز على جعل فقراته ممثلة تمثيلا سليما للموضوعات المقررة ويعني ذلك توزيع فقراته على مختلف جوانب المنهاج المقرر.

(عقل، 2001، ص 410).

يعني أيضا الالتزام بالحياد أثناء عملية التصحيح أن لا يتأثر المصحح بعوامل أخرى خارجية لا تعكس الأداء الحقيقي للمفحوص، كأن يتأثر برداءة الخط أو شخصية المفحوص الذي يقيمه، فموضوعية التصحيح تعني إعطاء نفس الدرجة لنفس ورقة الاختبار مهما تعدد مصححو هذه الورقة وهذا خصوصا في الاختبارات المقالية، حيث يذكر سبع أبولبدة (1987) أن دراسة أجريت في بريطانيا تم فيها إعطاء أوراق إجابات مادة التاريخ إلى خمسة عشرة مصححا لتصحيحها، فاختلّفوا في تقدير الدرجات اختلافا كبيرا، ثم أعطيت نفس الإجابات لنفس المصححين بعد سنة ونصف فانعكس التقدير في (92) حالة من راسب إلى ناجح. (النور، 2007، ص 172)

لتوخي الموضوعية في الاختبارات التحصيلية هناك بعض الإجراءات التي تساعد على ذلك والتي تتعلق بتصميم الامتحان وتصحيحه وتتمثل هذه الإجراءات فيما يلي:
(أبولبدة، 2002، ص ص 209-210)

- أن تكون أسئلة الامتحان عينة ممثلة لمختلف أجزاء المادة، وعادة ما يكون ذلك بالالتزام بجدول المواصفات.
- أن تكون هذه الأسئلة واضحة محددة لا لبس فيها ولا غموض، وان تكون لغتها في مستوى التلاميذ كي لا تحد اللغة المعقدة من قدرتهم على الإجابة.
- محاولة تصميم الأسئلة على الطريقة الموضوعية قدر الإمكان.
- عندما تكون الأسئلة مقالية يجب أن توضع إجابة نموذجية يتم في ضوءها تصحيح إجابات التلاميذ ووضع سلم تصحيح.
- يفضل قراءة عينة من إجابات التلاميذ ثم تقارن بسلم التصحيح ليتم تعديله قبل الاستعمال.

2- الثبات:

يتصف الاختبار بالثبات عند حصول الطلاب المختبرين على نفس الدرجة تقريبا بعد تطبيق نفس الاختبار عليهم لأكثر من مرة وخلال فترات زمنية متقاربة ومحددة. والمقصود من ذلك أن يعطي الاختبار نفس النتائج إذا ما كرر تطبيقه على نفس المجموعة وفي نفس الظروف. (عقل، 2001، ص 404)

هذا ما ذهبت له أنستازي في تعريفها للثبات على أنه "اتساق القياسات التي يتم الحصول عليها من نفس الأفراد عندما يتم إعادة اختبارهم باستعمال نفس الأداة في مناسبات مختلفة وتحت نفس الظروف" (Anne Anastasi & Susana Urbina, 1997, p 84)

أي أن يحافظ أفراد العينة التي طبق عليها الاختبار على نفس ترتيبهم إذا تم تطبيقه عليهم مرة أخرى بعد فترة زمنية معينة، ويتراوح معامل الثبات للاختبار بين (0، 1) وكلما اقتربت قيمته من الواحد الصحيح كانت درجات الاختبار أكثر ثباتا، لكن عادة لا يمكن الحصول على معامل ثبات يساوي الواحد تماما لأن هناك بعض الظروف التي تؤثر في ذلك ولا يمكن التحكم فيها، وفي هذا العنصر سنفكر الطرق التي يمكن بها حساب معامل ثبات الاختبارات التحصيلية وتحديد القيمة المقبولة له والعوامل التي يمكن أن تؤثر على الثبات.

2-1-1- طرق حساب الثبات:

يمكن تقدير معامل الثبات للاختبارات التحصيلية من خلال الطرق التالية:

2-1-1-1: طريقة إعادة الاختبار:

تعتمد هذه الطريقة على إعطاء الاختبار مرتين للمجموعة نفسها مع وجود فاصل زمني معين بين المرتين ومراعاة إجراء الاختبار في نفس الظروف لكلا المرتين، وعندما يتم حساب معامل ارتباط درجات الأفراد في التطبيقين، ويقبل معامل ثبات الاختبار إذا وصل معامل الارتباط (معامل الثبات) إلى (0,7) أو (0,8) ويمكن اعتباره معامل ثبات مرتفع أما إذا انخفض عن (0,7) فيعتبر ثباته منخفضا، وهذا يعني عدم توافر الثبات. (النور، 2007، ص 176)

لكن مما يعاب على هذه الطريقة أنها تطبق نفس الاختبار في أكثر من مرة مما يجعل التلاميذ يألفونه ويتذكرون إجابات التطبيق الأول ويستعملونها في المرة الثانية خصوصا إذا كانت المدة الفاصلة بين التطبيقين قصيرة، أما إذا طالت المدة فإن التلاميذ يتعرضون لعوامل النسيان، أو التذكر، أو التعلم. (أبوليدة، 2002، ص 229)

لذلك عند حساب معامل الثبات لهذه الطريقة يجب توخي الحذر وأخذ الحيطة في تقدير زمن المدة الفاصلة بين التطبيقين بحيث لا تؤثر على النتائج بشكل كبير.

2-1-2 - التجزئة النصفية:

في هذه الطريقة بدلا من إعادة التطبيق فإننا نحسب معامل الثبات من نتائج نفس التطبيق وذلك بقسمة الاختبار إلى جزأين متساويين ولحساب معامل الثبات يحسب معامل الارتباط بين هذين الجزأين، وقد يكون التقسيم إلى جزأين متساويين النصف الأول والنصف الثاني، وقد يكون بتقسيم الأسئلة إلى أسئلة فردية وأسئلة زوجية، بحيث يكون الجزأين المتحصل عليهما متكافئين في المحتوى ومستوى صعوبة المفردات وتباين الدرجات وعدم اعتماد إجابة المفردات على السرعة.

(علام، 2007، ص 235)

بعد حساب معامل الارتباط بين درجات الأداء المتحصل عليهما في الجزأين فإننا نتحصل على معامل ثبات نصف الاختبار لذلك لتصحيح الطول والحصول على ثبات الاختبار ككل يمكن استخدام بعض المعادلات والتي من بينها نذكر:

(أبولبدة، 2002، ص ص 231 - 232)

- معادلة سبيرمان . براون: في هذه الطريقة يقسم الاختبار إلى نصفين متكافئين ، ويقصد بالتكافؤ هنا تساوي متوسطي النصفين وتساوي تباينهما وتمائل معاملات الصعوبة ، وقانونها كما يلي:

$$\text{رث ت} = \frac{\text{ن ر}}{1 + (\text{ن} - \text{ر})}$$

رث ت : معامل الثبات التقديري أو معامل ثبات الفحص ككل .

ن : عدد المرات التي يجب أن يزداد بها طول الفحص .

ر : معامل الثبات النصفي للفحص الذي التوصل إليه بطرق تجريبية إحصائية .

- معادلة رولون: يفترض رولون أن تباين درجات الأفراد على الاختبار ككل يرجع إلى تباين حقيقي في مستويات القدرة المقاسة، كما يرجع إلى أخطاء تجريبية في القياس، وهذه المعادلة كما يلي: (بوسالم، 2014، ص94)

$$\text{رث ت} = 1 - \left(\frac{\text{ع}^2 \text{ف}}{\text{ع}^2} \right)$$

رث ت : معامل ثبات الاختبار ككل .

ع² ف : تباين الفرق بين علامات الاختبارين المكون كل منهما من نصف الاختبار الكلي.

ع²: تباين علامات الفحص الكلي .

- معادلة جتمان العامة : وهذه المعادلة تصلح عندما لا تتساوى الانحرافات المعيارية لجزئي الاختبار، وتصلح كذلك في حال تساوي هذه الانحرافات، وهذه المعادلة كما يلي: (النور، 2007، ص181)

$$\text{رث ت} = 2 \left(\frac{-1 \text{ع}^2_1 + \text{ع}^2_2}{\text{ع}^2} \right)$$

ع²₁: تباين درجات الأسئلة الفردية .

ع²₂: تباين درجات الأسئلة الزوجية.

2-1-3 - طريقة الاتساق الداخلي:

ترى بشرى إسماعيل (2004) أن هذه الطريقة لحساب الثبات تعتمد على الاتساق في أداء الأفراد على الاختبار من فقرة لأخرى، ففيها يقسم الاختبار إلى عدد كبير من الأجزاء بحيث يتكون كل جزء من فقرة واحدة من فقرات الاختبار، وكلما زاد الاتساق بين هذه الفقرات زاد ثبات الاختبار ككل، ومن بين الطرق المستخدمة لذلك ما يلي:

(مراد وسليمان، 2005، ص ص 365، 366)

- معادلة كيو دريتشاردسون: توصل كيو دريتشاردسون Kuder Richardson إلى معادلة لحساب معامل ثبات الاختبار عن طريق تحليل الإجابات على البنود وحساب تبايناتها وسميت KR. 20 وهي:

$$r = \frac{N \cdot \sigma^2 - M^2}{N(N-1) \cdot \sigma^2}$$

ن : تشير إلى عدد أسئلة الاختبار .

م : تشير إلى متوسط درجات الاختبار .

ع² : تشير إلى تباين درجات الاختبار ككل .

كما أن كودر ريتشاردسون توصل إلى معادلة أخرى وهي " KP-21 ":

$$r_{21} = \frac{N}{N-1} \left[\frac{M(N-M)}{N \cdot \sigma^2} - 1 \right]$$

- معامل α كرونباخ: ويعد معامل α كرونباخ حالة خاصة من قانون كودر ريتشاردسون وقد اقترحه كرونباخ 1951 ونوفاك ولويس 1967، وتعتمد هذه المعادلة على تباينات أسئلة الاختبار وتستعمل عندما لا تكون أسئلة الاختبار إجابتها ثنائية مثل نعم أو لا (0، 1)، على عكس معادلة كودر ريتشاردسون التي تستعمل عندما تكون إجابة الأسئلة ثنائية، وهذه المعادلة كما يلي:

ن	مج تباينات الأسئلة
$r_{1-1} = [1 -]$	
ن-1	تباين الدرجات الكلية

القيمة المقبولة لمعامل الثبات في الاختبارات التحصيلية:

وفي هذا الشأن يذكر سبع محمد أبو لبدة (2008) أن معاملات الثبات لاختبارات التحصيل المقننة تبلغ (0,50).

أما الاختبارات التي يعدها المعلم أقل ثباتا إذا ما قورنت بالاختبارات التحصيلية المقننة، وهذا خصوصا إذا كانت مقالية فإن معاملات ثباتها تكون منخفضة ما لم يهتم المعلم بعملية بنائها بشكل جيد وفق معايير محددة وسلم تصحيح واضح.

2-3- العوامل المؤثرة في ثبات الاختبارات التحصيلية:

هناك عدة عوامل يمكن أن تؤثر على ثبات الاختبارات التحصيلية من بينها ما يلي:

2-3-1 المصحح:

لا تتوقف درجة الطالب على معلوماته فقط بل على من يصحح ورقته أيضا خاصة في الاختبارات المقالية، وهذا ما أكدته بعض الدراسات مثل ما قام الباحثان ستارك وإليوت Elliot et Starch فقد أرسلوا ورقة إجابة في امتحان هندسة مستوية إلى 138 مدرس

هندسة طالبين تصحيحها فتراوحت درجاتها بين (28-95) (أبوليدة، 2002 ص236)، وهذا يدل على أن تصحيح الورقة لا يعكس إجابة التلميذ فقط وإنما يتدخل في ذلك رأي المصحح وحالته النفسية وتأثره بخط التلميذ وغير ذلك من الأمور التي تؤثر على ذلك .

2-3-2 بناء الاختبار:

- تتأثر الاختبارات التحصيلية بدرجة كبيرة بعامل الصدفة، فقط تكون أسئلة الاختبار تعكس جزءا يتقنه التلميذ فيتحصل على درجة مرتفعة، وقد تكون هذه الأسئلة من جزء لا يتقنه التلميذ فيرسب، وهذا راجع إلى عدم تمثيل الأسئلة للمادة الدراسية تمثيلا صحيحا، كما أن غموض الأسئلة وعدم فهمها فهما صحيحا يؤثر في ثبات الاختبارات .

- احتمال الاختبار على عدد قليل من المفردات أو الأسئلة يحتمل أن يكون ثبات درجاته منخفضا، وهذا يتضح عند تطبيق معادلة سبيرمان . براون في التجزئة النصفية للاختبار، حيث تزداد القيمة التقديرية لثبات درجات الاختبار ككل عن قيمته باستخدام نصف عدد مفرداته فقط، لذا كلما زاد عدد المفردات في الاختبار كلم ازداد ثبات الاختبار، لكن الزيادة في عدد المفردات تتطلب وقتا أطول في الإجابة مما يشعر الطلاب بالإجهاد فيقلل ذلك من ميزات إطالة الاختبار إذن من المرغوب فيه عادة كتابة عدد من الأسئلة يمكن الإجابة عليها في وقت قصير، لكن هنا يعتمد على طبيعة المادة الدراسية والأهداف التعليمية والصف الدراسي. إذ أنه في بعض الأحيان ربما يفضل أن يشتمل الاختبار على عدد أقل من الأسئلة التي تتطلب إجابة مطولة. (علام، 2007، ص243)

- تباين قدرات المفحوصين يؤثر على ثبات الاختبار، فإذا كانت مجموعة الطلاب الذي أجري عليهم الاختبار متجانسة يؤدي ذلك إلى انخفاض ثبات الاختبار، ويفسر ذلك بأنه عندما تكون علاماتهم أقل انتشارا أي تكون متقاربة فإن احتمال تغير مركز الفرد بين جماعته إذا أعيد تطبيق الاختبار عليهم احتمال كبير، مما يؤدي إلى انخفاض معامل الثبات الذي سبق وأن أسلفنا أنه محافظة الفرد على مركزه وترتيبه بين أفراد

مجموعته، أما إذا كانت هذه الدرجات مختلفة اختلافا كبيرا فإن ذلك يزيد من معامل ثبات الاختبار.

- مدى صعوبة الأسئلة يؤثر بدوره على معامل ثبات الاختبار، فإذا كانت معظم أسئلة الاختبار صعبة جدا فإن الصدفة تلعب دورا كبيرا في إجابات الطلاب، مما يؤثر على معامل ثبات هذا الاختبار .

- الإرهاق أو السأم على الممتحنين أو أية عوامل أخرى مما يجعلهم يستجيبون بصورة سيئة للاختبار وكل ذلك يؤدي إلى أن تكون قيمة معامل الثبات خاطئة.

(Edel, .Frisbie, 1995, p 89)

- صدق الاختبار يؤثر على ثباته هذا لأن الاختبار الصادق ثابت و ليس بالضرورة أن الاختبار الثابت صادق. (منسي وآخرون، 2003، ص150)

إذن من خلال ما سبق يتضح أن هناك عديد العوامل التي تتدخل و تؤثر على معامل ثبات الإختبار، لذلك للحفاظ على ثبات الاختبار يفضل أخذ التدابير اللازمة في ضوء العوامل المؤثرة المذكورة أعلاه .

3- الصدق :

يعتبر الصدق ضروري لأداة القياس لذا فهو شرط أساسي من شروط الاختبار التحصيلي الجيد، ويعني الصدق أن يقيس الاختبار ما وضع لقياسه، حيث يرى جيه Gay (1990) أن الاختبار الصادق اختبار يقيس الوظيفة التي يزعم أنه يقيسها ولا يقيس شيئا آخر بدلا منها أو بالإضافة إليها (ملحم، 2007، ص334)، أي أن يحقق الاختبار الأهداف التي وضع من أجلها فمثلا إذا وضع معلم مادة العلوم اختبار تحصيلي في مادة العلوم الطبيعية، فيجب أن يقيس هذا الاختبار تحصيل الطلاب في مادة العلوم الطبيعية وألا يقيس شيئا آخر معها كحسن الخط، أو إجابة التعبير، لأن هذا يقاس باختبارات تحصيلية أخرى في اللغة. وللصدق عدة أنواع وعدة طرق يقاس بها كما أنه يتأثر بعدة عوامل وسنحاول في هذا العنصر توضيح ذلك على النحو التالي :

1-3 - أنواع الصديق

للصديق عدة أنواع نذكرها كما يلي:

1-1-3 - صديق المحتوى

يشير صديق المحتوى إلى مدى تمثيل الاختبار التحصيلي تمثيلا صادقا لمختلف أهداف وأجزاء المادة المدروسة، حيث كلما مثل الاختبار محتوى المقرر تمثيلا شاملا كلما كانت درجة صدقه مرتفعة، حيث يذكر صلاح مراد وآخر (2005) أن صديق المحتوى دليل على شمول الأداة ودرجة تمثيلها للمحتوى، ولتحقيق هذا النوع من أنواع الصديق يجب الاعتناء ببناء جدول المواصفات للمحتوى المراد إجراء اختبار عليه لأن ذلك يسهم بشكل كبير المواصفات للمحتوى المراد إجراء اختبار عليه لأن ذلك يسهم بشكل كبير في التوزيع العادل للأسئلة الموضوعية حول الأهداف التعليمية وذلك بحسب الأهمية النسبية لكل منها، ويطلق على صديق المحتوى أيضا صديق المضمون المقصود بالمضمون هو المحتوى، وللتأكد من صديق المحتوى للاختبار لا بد من عرضه على عدد من المتخصصين في المجال المعني وألا يقل عددهم عن خمسة لتحكيمه، حيث يرى أحمد يعقوب النور أنه إذا كانت درجة الاتفاق بين المحكمين عالية بحيث لا تقل عن 75% فإن معامل صديق المحتوى يعد مقبولا، أما إذا قلت نسبة الاتفاق عن 50% فينبغي تعديل فقرات أداة القياس، إلا أنه كلما زاد اتفاق المحكمين ارتفع صديق المحتوى

2-1-3 - الصديق المرتبط بمحك:

ويعتمد هذا النوع من أنواع الصديق على مدة ارتباط علامات الاختبار المراد الكشف عن صدقه وعلامات فحص آخر أعده المعلم، أو معدل الطالب الفصلي أو السنوي ويسمى في هذه الحالة فحص المعلم أو معدل الطالب بالمحك، حيث يشترط في هذا الأخير أن يكون موضوعيا صادقا، وثابتا (أبو لبد، 2002، ص218)، والصديق المرتبط بمحك ينقسم إلى نوعين:

3-1-3: الصدق التنبؤي:

يعتمد الصدق التنبؤي لأداة القياس على حساب القيمة التنبؤية لهذه الأداة على أساس أن السلوك يتضح بدرجة عالية من الثبات في المستقبل. (دروزة، 2005، ص175) وهذا يعني أن الفرد الذي نجح في أداء اختبار تحصيلي مادة معينة فإنه يتوقع له أن ينجح في دراسته الجامعية إذا تخصص هذه المادة الدراسية، ويحسب معامل صدق هذا الاختبار ودرجات محك تجمع عنه المعلومات لاحقا كأن يكون هذا المحك هو معدل الفصل الأول أو العام الأول من الدراسة الجامعية.

3-1-4 - الصدق التلازمي:

يُحسب هذا النوع من الصدق بحساب معامل الارتباط بين درجات الاختبار المراد الكشف عن صدقه، ودرجات اختبار آخر، على أن يطبقان في فترات متقاربة، كذلك يمكن استخدام معدلات التلاميذ كمحك لحساب صدق هذا الاختبار (أبولبدة، 2002، ص219) ويكون هذا عندما يريد المعلم استبدال اختبار جديد مكان اختبار كأن يستغله ويستغرق له الكثير من الوقت عند عملية التطبيق أو التصحيح، وبالتالي الاختبار الجديد يوفر له الوقت، ويكون هذا الاختبار صادقا إذا اقترب معامل الارتباط بين نتائجه ونتائج المحك إلى الواحد الصحيح.

3-1-5 - الصدق العاملي :

يهتم الصدق العاملي بمعرفة مدى تشعب بنود الاختبار بعوامل معينة (مكونات) سواء كانت عاملا عاما أو عوامل طائفية. (مراد وسليمان، 2005، ص355)، ويعتمد هذا النوع من الصدق على أسلوب التحليل العاملي الذي يقوم على تحليل معاملات الارتباط بين الاختبار والمحكات المختلفة من أجل الوصول إلى العوامل التي أدت إلى إياد هذه المعاملات.

3-1-6- صدق التكوين الفرضي:

ويشير هذا النوع من أنواع الصدق إلى مدى قياس الاختبار لتكوين فرضي أو مفهوم نفسي أو سمة، وهذا مثل الذكاء، الرياضيات... وقد وضع ثرستون ترجمة أحمد زكي اختبار

للقدرات العقلية الأولية ويقاس هذا الاختبار أربع قدرات هي: معاني الكلمات، والإدراك المكاني والتفكير، والقدرة العددية، بحيث تستخدم كل هذه القدرات في حساب نسبة الذكاء وواضع الاختبار هنا يقوم بإعداد أسئلة تقيس تلك المكونات ويتأكد من صدقها بعد تجربتها على عينة ويتم إجراء تحليل عاملي لتحديد مدى قياسه للمكونات المفترضة. (النور، 2007، ص201)

2-3- طرق حساب معامل الصدق:

من الطرق المستخدمة لحساب معامل الصدق يذكر أحمد يعقوب النور (2007)

ما يلي:

3-2-1- استخدام معامل الارتباط للدلالة على صدق الاختبار:

وهنا مثلا إذا تم استخدام طريقة المحك الخارجي، فإنه لا بد من إيجاد معامل الارتباط بين درجات الاختبار موضع التطبيق ودرجات المحك الخارجي الذي ثبت صدقه سابقا ويبدل هذا المعامل على صدق الاختبار، ويمكن استخدام معامل ارتباط بيرسون لإيجاد درجة الصدق، والتي نصها كما يلي :

$$r = \frac{\text{مج ف}_1 \text{ ف}_2}{\sqrt{\text{مج ف}_1^2 \text{ ف}_2^2}}$$

حيث :

ر: معامل ارتباط بيرسون .

ف₁ : فرق درجات الاختبار الأول عن متوسطها .

ف₂ : فرق درجات الاختبار الثاني عن متوسطها .

3-2-2- طريقة استطلاع آراء المحكمين:

في هذه الطريقة يقدر المحكم المتخصص مدى علاقة كل بند من بنود الاختبار بالسمة أو القدرة المقاسة، وذلك بعد توضيح المعنى الإجرائي لهذه السمة، ويمكن استخدام المعادلة التالية لحساب صدق المحتوى عن طريق التحكيم :

$$S_1^{-1} S_2$$

$$S_m = -$$

$$K$$

حيث :

S_m : مؤشر صدق المحتوى .

S₁ : عدد الأسئلة التي اتفق المحكمون على أنها تقيس الهدف.

S₂ : عدد الأسئلة التي اتفق المحكمون على أنها لا تقيس الهدف.

K : العدد الكلي لأسئلة الاختبار.

3-2-3- طريقة المقارنة الطرفية :

وفي هذه الطريقة يتم التعرف على مدى قدرة الاختبار على إبراز الفروق الفردية لدى أفراد العينة، حيث يرى رجاء أبو علام 1987 أن الاختبارات محكية المرجع يجب أن تزيد الفروق بين المجموعات التي أتقنت المقرر، ولمعرفة ذلك يجب إتباع الخطوات التي يذكرها بشير معمريه 2007 كما يلي : (معمريه، 2007، ص158)

- ترتب نتائج الاختبار التي حصل عليها أفراد العينة ترتيبا تنازليا أو تصاعديا .
- يسحب 27% من طرفي التوزيع، فنحصل على مجموعة عليا من حيث ارتفاع درجاتها مجموعة دنيا من حيث انخفاض درجاتها في الاختبار .

- يقارن بين أداء المجموعتين باستعمال أسلوب إحصائي ملائم وهو اختبار "ت" لدلالة الفرق بين متوسطين حسابيين

وفي الأخير وبعد الحصول على (ت) المحسوبة يتم مقارنتها مع (ت) الجدولة لنستطيع الحكم على صدق الاختبار، بحيث إذا كانت قيمة (ت) المحسوبة دالة يمكن القول أن الاختبار صادق.

3-3- تفسير معامل الصدق:

كلما اقترب معامل الصدق من الواحد الصحيح، كلما كان صدق الاختبار عاليا وكلما اقترب من الصفر كان الصدق منخفضا، ولا يوجد خط فاصل إذا تعداه معامل الصدق المقدر نعتبر الاختبار صادقا، وإذا انخفض عنه نعتبره غير صادق، ويذكر سبع محمد أبو لبددة أن معامل صدق اختبار يؤلفه المدرس ونتائج اختبار ذكاء يتراوح بين (0.30) إلى (0.50) وهذا الانخفاض راجع إلى أن الاختبار الذي يعده المدرس يتأثر بعوامل أخرى كاجتهاد التلميذ، وذاتية المصحح وغير ذلك من الأمور التي يمكن أن تؤثر فيه، أما إذا كان معامل الصدق المحسوب بين نتائج اختبار تحصيلي وضعه المدرس ونتائج اختبار تحصيلي آخر مقنن، يكون أعلى مما سبق كذلك معامل الصدق بين بطارية اختبارات تحصيل ومعدل التلميذ أو رتبته في الصف فإنه يتراوح على الأغلب بين (0.60) إلى (0.70)، أما إذا حسب معامل الصدق بين اختبارين تحصيليين مقننين فإنه يتراوح بين (0.60) إلى (0.80).

3-4- العوامل المؤثرة على صدق الاختبار:

هناك عديد العوامل التي قد تقلل من صدق الاختبار وتجعله غير مناسب للاستعمال الذي خطط من أجله، ويمكن تقسيم هذه العوامل كما يلي: (ملحم، 2007، ص336)

3-4-1- عوامل متعلقة بالتلميذ:

وتتمثل فيما يلي :

- اضطراب التلميذ أو خوفه عند أدائه للاختبار، قد يشل قدرته على الإجابة وبالتالي يحصل على درجة لا يمثل قدراته الحقيقية .
 - لجوء التلميذ للتخمين أو الغش، أو يخادع المصحح بأسلوبه في التعبير ويؤثر عليه
- 3-4-2- عوامل متعلقة بالاختبار:

وتتمثل فيما يلي :

- صعوبة لغة أسئلة الاختبار تجعل التلميذ يعجز عن الإجابة، وكذلك غموضها يجعل التلاميذ يفسرونها تفسيرات مختلفة وبالتالي تكون إجاباتهم مختلفة.
- صياغة الأسئلة التي تحمل في طياتها الإجابة تمكن التلاميذ من الحصول على علامات مرتفعة.
- سهولة الأسئلة أو صعوبتها تجعلنا نحكم حكما خاطئا على التلميذ، فإذا كانت سهلة دون مستواهم العقلي تجعلهم يحصلون على درجات مرتفعة فنحكم عليهم بالتفوق والعكس صحيح، لذا في الحالتين لن نتصف درجة التلميذ بالصدق، لأنها لا تمثل قدرته الحقيقية .
- علاقة الأسئلة بما تعلمه التلميذ يؤثر على صدق الاختبار، فإذا كانت هذه الأسئلة الجيدة وتقيس المعرفة والفهم والتحليل والتطبيق لكن المعلم علم التلاميذ الإجابة عنها وطريقة حلها فتصبح هذه الأسئلة تقيس تذكر الحقائق فقط، كذلك إذا وضع المعلم أسئلته من موضوع لم يدرسه التلاميذ فإن هذا يؤدي على تدني درجاتهم وهذا التدني لا يمثل قدراتهم الحقيقية.(النور، 2007، ص196).

3-4-3- عوامل متعلقة بإدارة الفحص :

وتتمثل فيما يلي :

- تؤثر العوامل البيئية تأثيرا سلبيا على إجابة التلميذ، وهذا مثل شدة الحرارة أو البرودة أو الضوضاء المحيطة .
- عدم وضوح طباعة الأسئلة ووجود أخطاء مطبعية أو سوء ترتيب الأسئلة .

عدم وضوح تعليمات الاختبار .

- استعمال الاختبار في غير ما وضع له، كأن يوضع اختبار قواعد لقياس المقدرة اللغوية لدى التلميذ.

- عدم استعمال الاختبار مع الفئة التي وضع من أجلها، كأن يتم استخدام اختبار وضع للموهوبين لفئة من الضعفاء .

4- معامل السهولة والصعوبة:

يعد معامل السهولة والصعوبة من مؤشرات الصلاحية لفقرات الاختبار التحصيلي وتدلنا صعوبة الفقرات للاختبار النسبة المئوية للطلاب الذين أجابوا عن السؤال إجابة خاطئة، وتختلف صيغة معامل الصعوبة حسب نوع الفقرة فإذا كانت الأسئلة موضوعية فإن حساب معامل صعوبة السؤال حسب (جلال، 2008، ص 209) كالتالي:

$$\text{معامل الصعوبة للسؤال الموضوعي} = \frac{\text{عدد الذين أجابوا إجابة خاطئة}}{\text{العدد الكلي للطلاب}} \times 100$$

كما تعني سهولة فقرات الاختبار النسبة المئوية للطلبة الذين تمكنوا من الإجابة على فقرات الاختبار إجابة صحيحة ومعادلتها كالتالي:

$$\text{معامل السهولة للسؤال الموضوعي} = \frac{\text{عدد الذين أجابوا إجابة صحيحة}}{\text{العدد الكلي للطلاب}} \times 100$$

أما إذا كانت الفقرات من النوع المقالي فإن معامل سهولة السؤال (العبيسي، 2010، ص 206)

كالتالي:

$$\text{معامل صعوبة سؤال المقال} = \frac{\text{مجموع علامات الطلبة في السؤال}}{\text{عدد الطلبة} \times \text{علامة السؤال}}$$

ويميل بعض الباحثين إلى حساب معاملات صعوبة المفردات عن طريق حساب سهولتها، بحيث يتراوح معامل الصعوبة بين 0 و1 ؛ و تأخذ العلاقة بين السهولة والصعوبة بحيث: علاقة عكسية مباشرة (السيد، 1978، ص 447)

$$\text{معامل الصعوبة} + \text{معامل السهولة} = 1$$

ولطبيعة النوع المقالي لأسئلة الاختبارات التحصيلية - عينة الدراسة- فإنه ولتقدير معامل السهولة-الصعوبة تم استخدام معامل السهولة للسؤال المقالي.

و يشير عودة (1993) إلى أن معامل الصعوبة المقبول (الجيد) يتفاوت تبعا لنوع السؤال كما يلي: (كاظم، 2001، ص 102)

- معامل الصعوبة المناسب لأسئلة: الصواب والخطأ هو: 0.75 .
- الاختيار من 4 بدائل هو : 0.63 .
- الاختيار من 5 بدائل هو : 0.60 .
- المقال هو : 0.50 .

يضيف أحمد يعقوب النور النوع التالي من أنواع الأسئلة :

- الاختيار من 3 بدائل، ومعامل الصعوبة المناسب له هو : (0.67)، ووضع معيار عام لمعامل الصعوبة لجميع المفردات وهو من (0.40) إلى (0.60).

5- معامل التمييز:

يشير معامل تمييز السؤال إلى مدى قدرة السؤال على إبراز الفروق الفردية بين مرتفعي ومنخفضي التحصيل، فالغرض من تحديد معامل التمييز لكل سؤال هو معرفة قدرة السؤال على التمييز بين المجموعتين العليا والدنيا (مراد وسليمان، 2005، ص 218) ولتحديد المجموعة العليا والمجموعة الدنيا يتم عمل ما يلي :

ترتب أوراق الطلاب ترتيبا تنازليا، أي من أعلى درجة إلى أقل درجة، ثم تقسم إلى مجموعتين متساويتين، الأولى مجموعة الأوراق ذات الدرجات العليا، والثانية مجموعة الأوراق ذات الدرجات الدنيا، وإذا كان عدد الطلاب فرديا فيمكن استبعاد ورقة إجابة الطالب الذي يقع ترتيبه في الوسط، فمثلا إذا كان عدد الطلاب 41 فيمكن استبعاد الطالب رقم 21 في الترتيب، أما إذا كان عدد الطلاب كبيرا كأن يكون 100 فيمكن الاكتفاء برقع الأوراق من كل طرف.

وعليه يقترح سبع محمد أبو لبدة أن يأخذ المعلم أية نسبة لتمثيل المجموعة العليا والمجموعة الدنيا شريطة أن يكون عدد الطلاب كافيا أو معقولا، فإذا كان عدد طلاب الصف من (30-40) طالبا يستحسن إدخالهم جميعا في التحليل أما إذا كان الامتحان لعدد من الشعب يبلغ طلابها 100 فإمكاننا أخذ أعلى وأدنى 30 % أو 25 % أو 20 % بحيث يدخل في التحليل ما بين 40-60 % من الطلاب، وكلما زاد عدد المفحوصين كلما انخفضت هذه النسبة. (أبولبدة، 2002، ص309)

إذن بعد الحصول على المجموعتين العليا والدنيا، يتم حساب معامل التمييز بتطبيق إحدى الصيغتين التاليتين:
- فإذا كانت الأسئلة موضوعية:

معامل التمييز =	$\frac{\text{عدد الإجابات الصحيحة في المجموعة العليا} - \text{عدد الإجابات الصحيحة في المجموعة الدنيا}}{\text{عدد الطلاب في إحدى المجموعتين}}$
-----------------	--

- أما إذا كانت المفردة من نوع المقال فالصيغة التي يمكن استخدامها كما يلي: (علام، 2007، ص254)

معامل تمييز سؤال المقال =	$\frac{\text{ق-ع-ق د}}{\text{س العظمى} \times (\text{ن})}$
---------------------------	--

ق ع: مجموع عدد النقاط التي حصل عليها طلاب المجموعة العليا .

ق د : مجموع عدد النقاط التي حصل عليها طلاب المجموعة الدنيا .

س العظمى: أكبر من النقاط التي يمكن أن يحصل عليها طالب في السؤال .

ن: عدد طلاب أي من المجموعتين العليا أو الدنيا.

ويمكن تفسير قيم معامل التمييز كما يراها ايبل EBEL (1965) على النحو التالي:

معامل التمييز ≤ 0.40 الفقرة تلي الغرض أو الهدف .

$0.30 \geq$ معامل التمييز ≥ 0.39 الفقرة تتطلب مراجعة قليلة .

$0.20 \geq$ معامل التمييز ≥ 0.29 الفقرة تقع على الحد الفاصل و تحتاج إلى مراجعة

معامل التمييز ≥ 0.19 يجب حذف هذه الفقرة أو إجراء مراجعة تامة لها.

(كروكروالجينا، 2009، ص418)

6- الشمول:

يقتضي الشمول أن تكون مفردات الاختبار ممثلة لعينة السلوك المراد قياسه تمثيلا صادقا (منسي وآخرون، 2003، ص135)، فالمقصود بالشمول في الاختبارات التحصيلية أن لا يضع المعلم أسئلة الاختبار الذي يعده من جزء محدد من المقرر ويتجاهل الجوانب الأخرى فالاختبار ينبغي أن يشمل محتوى المقرر ويشمل الأهداف التعليمية ويمثل ما تم تدريسه للمتعلمين تمثيلا شاملا متوازنا، لأن هذا يسهم بشكل كبير في تحديد وتشخيص مواطن الخلل وبالتالي يسهل وضع برنامج علاجي لها في ضوء النتائج المتحصل عليها.

7- القابلية للاستعمال :

يصبح الاختبار قابلا للاستعمال إذا ما كان عمليا من حيث :

1-7- سهولة تطبيقه: فنجد أنه كلما كان تطبيق الاختبار سهلا كلما كان ذلك أفضل

لأن صعوبة التطبيق قد تكون عائق يمنع من تحقيق صدق الاختبار وثباته وموضوعيته وتشمل سهولة التطبيق وضوح الأسئلة وتعليمات الاختبار، فالتعليمات توجه الطالب وترشده في أداء الاختبار كونها يخبره بما ينبغي أن يفعله بالضبط، وتوضح الوقت المخصص لحل أسئلة الاختبار، وهذه التعليمات يجب صياغتها بلغة بسيطة وواضحة توضح الهدف من الاختبار

وطريقة تسجيل الإجابات، وتوضح كذلك للطالب أن يتجنب التخمين في حال ارتيابه في تحديد الإجابة الصائبة حيث يرى ألن وين (1979) Allen et Yen أنه لتحقيق تصميم تعليمات واضحة للاختبار فإنه يمكن إتباع ما يلي: (ملحم، 2007، ص444)

- نبه طلابك إلى قراءة التعليمات قبل البدء في الإجابة على فقرات الاختبار .
 - نبه طلابك إلى مكونات الاختبار، عرفهم بعد فقرات الاختبار، ضع قيمة العلامة لكل فقرة من فقرات أسئلة الاختبار .
 - نبه طلابك إلى عدد الأوراق الاختبار .
 - حدد زمن الاختبار وقدر لتلاميذك الوقت الذي يستغرق في الإجابة على كل فقرة من فقرات أسئلة الاختبار .
 - دون في صفحة التعليمات طريقة الإجابة على أسئلة الاختبار إذا كانت الأسئلة من نفس النوع، أما إذا كانت الأسئلة تشمل على عدة أنواع نحدد لطلبتك كل نوع من أنواع فقرات الاختبار.
 - نبه طلابك إلى عدم إمكانية الاستفسار عن أي شي يتعلق بأسئلة الاختبار من المراقب، ما عدا الأخطاء المطبعية التي قد ترد في ورقة أسئلة الاختبار أو تعليماته .
 - نبه طلابك إلى ضرورة الكتابة بخط واضح ومنظم.
 - أكد على الطلبة ضرورة كتابة أسمائهم على ورقة الإجابة بمجرد استلامهم الأوراق الخاصة بالإجابة .
 - أكد على طلابك ضرورة عدم التخمين في إجاباتهم لفقرات أسئلة الاختبار، خاصة إذا كنت تنوي التصحيح من أثر التخمين.
- 2-7- سهولة تصحيحه: عند وضع الاختبار، على المعلم التفكير في طريقة تصحيحه لأن عملية التصحيح كلما تعقدت يؤدي ذلك إلى السماح بتدخل ذاتية المصحح، كما أن ذلك يؤدي إلى حدوث الأخطاء و يتطلب وقتا وجهدا كبيرين، وهذا خصوصا في اختبارات المقال فتصحيحها بالغ التعقيد لان الجواب يختلف من مفحوص لآخر وفيها يصعب التقيد بدليل للتصحيح مما يسمح بتدخل ذاتية المصحح، لذا من الأفضل أن يضع المعلم تصحيحا نموذجيا مرفقا بسلم تنقيط واضح ليتم التصحيح في ضوء ذلك بأقل أخطاء ممكنة.

خاتمة:

إذا ما توفر الاختبار التحصيلي على شروط ومواصفات الاختبار التحصيلي الجيد فإن القرارات الناجمة عنه تحقق الغرض المحدد له سواء كانت لجمع المعلومات والتعرف على نقاط القوة لدى التلاميذ لترسيخها أو الضعف لمعالجتها أو كانت بغرض اتخاذ القرارات المتعلقة بالنجاح أو الرسوب. و أهم شروط مواصفات الاختبار التحصيلي الجيد توافره على الموضوعية، وابتعاده قدر الإمكان عن الذاتية، كما أنه ينبغي أن يغطي ما تم تدريسه تغطيه شاملة متوازية وأن يتوفر على صدق وثبات مناسبين، لكن نستطيع الوثوق بنتائجه كما أن مفرداته يجب أن تتمتع بمستوى صعوبة وقدرة تمييزية مناسبة لنوع المفردة والغرض من الاختبار، وأن يكون هذا الاختبار سهل التطبيق والتصحيح، فيسهل استعماله، وبالتالي يؤدي الغرض المنشود منه .

قائمة المراجع:

- أبو لبد، سبع محمد (2002): مبادئ القياس النفسي والتقويم التربوي، ط1، دار الفكر للنشر والتوزيع، الأردن.
- بوسالم، عبد العزيز (2014): القياس في علم النفس والتربية، الأسس النظرية والمبادئ التطبيقية، ط1، دار قرطبة للنشر والتوزيع، الجزائر.
- جلال، أحمد سعد (2008): تطبيقات وتدريبات عملية علي برنامج SPSS ط1، دار الدولية للاستثمارات الثقافية، القاهرة.
- دروزة، أفنان نظير، (2005): الأسئلة التعليمية والتقويم المدرسي، ط1، دار الشروق للنشر والتوزيع، الأردن.
- السيد، فؤاد البهي (1978): علم النفس الإحصائي وقياس العقل البشري، ط3، دار الفكر العربي، مصر.
- العبسي، محمد مصطفى (2010): التقويم الواقعي في العملية التدريسية ، ط1، دار المسيرة للنشر والتوزيع والطباعة، الأردن.
- عقل، أنور (2001): نحو تقويم أفضل، ط1، دار النهضة العربية، بيروت.
- علام، صلاح الدين محمود (2007): الاختبارات التشخيصية، مرجعية المحك في المجالات التربوية والنفسية، د. ط، القاهرة.
- كاظم، علي مهدي (2001): القياس والتقويم في التعلم والتعليم، ط1، دار الكندي للنشر والتوزيع، الأردن.
- مراد، صلاح أحمد وسليمان، أمين علي (2005): الاختبارات والمقاييس في العلوم النفسية والتربوية، خطوات إعدادها وخصائصها، ط3، دار الكتاب الحديث، دبي.
- ملحم، سامي محمد (2007): مناهج البحث في التربية وعلم النفس، ط5، دار للنشر والتوزيع الأردن.

- منسي، محمود عبد الحليم والصالح، أحمد وناجي، محمد قاسم (2003): التقويم التربوي ومبادئ الإحصاء، شركة الجمهورية الحديثة للتحويل وطباعة الورق، الإسكندرية، مصر.
- النور، أحمد يعقوب (2007): القياس والتقويم في التربية وعلم النفس، الجنادرية للنشر والتوزيع الأردن.
- Anastase, A, Urbina, S,(1997),Psychological Testing, New Jersey .
- Edel, R-Frisbir, (1995),Essentials Of Educational Measurement, New Jersey.