*ما المقصود بثبات الاختبار؟*

*بحث في القياس والتقويم التربوي*

 *إعداد/ أحمد كمال فوزي*

*قسم التربية*

*كلية التربية– جامعة المدينة العالمية*

شاه علم - ماليزيا

*ahmed.fawzy@mediu.ws*

*خلاصة*—هذا البحث يبحث في ما المقصود بثبات الاختبار؟.

*الكلمات المفتاحية: المقياس، الصدق، الثبات*.

# ***المقدمة***

معرفة أسس ما المقصود بثبات الاختبار؟، من المفيد أن نطبق الثبات على الاختبارات بهدف تحديد الأخطاء في القياس، ولكن يجب أن نضع في الاعتبار أنواع الأخطاء التي يمكن أن نقع فيها بالقياس. عندنا نوعان من الأخطاء: أخطاء عشوائية غير مقصودة، أخطاء منتظمة مقصودة. الأخطاء العشوائية ترجع إلى الصدفة،والخطأ العشوائي قد يرفع درجات الثبات، وقد تنكمش على أثره درجات الثبات. فأخطاء الصدفة أو الأخطاء العشوائية غير الأخطاء المنتظمة.

1. *المقالة*

تعرضنا إلى إحدى الوسائل التي نتأكد بها من صحة الاختبار، وهو قياس صدق الاختبار. قياس الصدق وأساليب قياس الصدق. أيضًا يُضاف إلى القياس درجة الثبات، التأكد من ثبات الاختبار.

المقصود بثبات الاختبار:

يقصد بثبات الاختبار الدقة في المقياس واتساقه، معنى ذلك أننا لو طبقنا اختبار على مجموعة ما أو عينة ما حصلنا على مجموعة درجات، بعد فترة زمنية محددة طبقنا نفس الاختبار على نفس المجموعة في نفس الظروف، سوف نجد أن الدرجة التي نحصل عليها هي نفس الدرجة التي حصلنا عليها من قبل، أو درجة قريبة جدًّا منها.

فعند تطبيقه أكثر من مرة فإننا بنصف الاختبار أو المقياس. في هذه الحالة إن الدرجة يجب أن تكون على مستوى عال من الثبات، ومعنى هذا أن مفهوم الثبات مرتبط بما يسمى إحصائيًّا أخطاء القياس، أخطاء القياس هي التي تعبر عن كل درجة من درجات القياس أثناء تطبيق ذلك القياس.

العامل الذي يرتبط بالثبات:

عندنا ما يسمى معامل الارتباط بين درجات الأفراد في المرات التي تم إجراؤه فيها، إذن معامل الارتباط بين عدد التكررات لتلك الدرجات هي معامل ثبات المقياس أو ثبات الاختبار. لو حصل تداول شديد بين تلك الدرجات في المرات المختلفة من خلال النتائج التي يتوصل إليها المقياس، سواء كان المقياس على مجموعة واحدة من الأفراد، فنجد هذا التباين يوضح لنا عدم وجود ثبات في ذلك الاختبار.

هناك ارتباط بين درجات الأفراد في المقياس، أحيانًا يحدث تباين، هذا التباين بين الدرجات يرجع إلى عدة أسباب، قد يرجع إلى اختلاف من طبق الاختبار على الأفراد، يعني: الذين قاموا بالإجراءات. إذن معامل الثبات هو معامل الارتباط بين المقياس ونفسه هو نفس المقياس، فنحن نحسب معامل الثبات بحساب معامل الارتباط بين الدرجات الخاصة بالمقياس ونفس الدرجات لنفس المقياس، ولكن بعد فترة زمنية.

ويمكن الاستعانة بمقياس آخر يتضمن نفس مستوى الدرجات، نفس عددها، ونفس حجمها، نفس المفردات اللغوية، ولكن مقياس آخر ولكن يتساوى مع نفس المقياس الأول في الحجم وفي عدد المفردات وفي قوة المفردات أو ضعفها. ولذلك عندما نتكلم عن الثبات نحن نتكلم عن مستويين أو طريقتين مختلفتين، هما استقرار درجات الاختبار، أو التناسق الداخلي لمفردات الاختبار.

إذن المقصود بالثبات هنا ثبات درجات الاختبار، الاختبار طبقناه على نفس المجموعة بنفس الظروف أكثر من مرة، يجب أن يكون ذلك الاختبار درجاته واحدة. أيضًا لو تم إعداد صورتين للاختبار فيجب أن يكون هناك تناسق بينهما، وأيضًا هناك تناسق بين مفردات الاختبار الواحد، مفردات الأسئلة الخاصة بأسئلة الاختبار نفسه، أو المقياس نفسه أن يكون هناك بينها درجة من الوضوح، درجة من المناسبة لمستوى التلاميذ، إذن صفات خاصة بإعداد الاختبار الداخلي.

ثبات الخط يشير إلى درجة من الالتصاق الخاصة بالدرجات: عامة لا يوجد مقياس ثابت بشكل مطلق؛ لأن الدرجات لا تكون ثابتة عبر الفترة الزمنية، يعني: لا تكون الدرجة مستقرة؛ لأنه تمر فترة زمنية، الفترة الزمنية قد تكون بعيدة، فالطالب ينسى الاختبار تمامًا، فبالتالي مستوى الثبات يكون منخفضًا. قد تكون الفترة الزمنية قريبة جدًّا وبالتالي يكون مستوى الثبات عالٍ، فمن ثم الفترة الزمنية تؤثر على على استقرار الدرجات، أو عبر عينة مماثلة من البنود، هذه خاصة بالتماثل بين الأسئلة، التناسق في أسئلة الاختبار بينها وبين بعضها، في درجة السهولة والصعوبة، وهذه خاصة بالتناسق الداخلي لمفردات الاختبار.

 نجد أن ثبات الاختبار سيتأثر بتلك الأمور، فقد ترتفع نسبة الثبات أو تنخفض نسبة الثبات بسبب الظروف التي تؤثر فيه.

 الأسلوب الذي سوف يتم من خلاله استخدام ذلك المقياس أو فيه، هو الذي سيؤثر في نوع الثبات المناسب لو أردنا نعرف حالة الأفراد مثلًا، لو نتكلم عن الأفراد في فترة زمنية، سنتكلم عن حالة في المستقبل، معنى هذا: أن المقياس الذي سنطبقه في الوقت الحالي سيطبق بعد فترة زمنية بعيدة، فمعنى هذا أن درجة الثبات ستتأثر عندما نطبق المقياس بعد فترة زمنية تزيد عن ستة شهور، أو أربعة شهور.

والجدير بالذكر أن الثبات ضروري للقياس، ولكن الثبات ليس بديلًا للصدق، يعني: تطبيق الصدق أساسي لا يمكن أن نستغنى عن قياس الصدق بقياس الثبات، فالمقياس ممكن أن يعطينا درجة غير مستقرة أو غير متسقة، لا يمكن أن يكون صادقًا، يعني: حينما يكون درجة الثبات الخاصة بالمقياس غير مستقرة، فبالتالي ذلك المقياس لا يكون صادقًا.

ومن ناحية أخرى إذا كانت درجة الثبات عالية ومرتفعة، فذلك ليس دليلًا على أن هذا الاختبار صادق، يعني: لو طبقنا مقياس الثبات على الاختبار، وأعطانا نسبة ثبات أو درجة ثبات عالية جدًّا للاختبار، هل معنى ذلك أنه يمكن الاستغناء بقياس الثبات عن قياس الصدق؟ القوانين كلها قالت والعلماء أجمعوا على أنه لا يمكن الاستغناء عن الصدق بالثبات. ولكن حين تكون الدرجة غير مستقرة وغير متسقة في قياس الثبات، فبالتالي الاختبار لا يمكن أن يكون صادقًا، إذن عدم استقرار الدرجة في الثبات تشير إلى عدم صدق الاختبار، ولكن ثبات الاختبار لا يشير إلى صدقه.

نظرية الثبات:

من المفيد أن نطبق الثبات على الاختبارات بهدف تحديد الأخطاء في القياس، ولكن يجب أن نضع في الاعتبار أنواع الأخطاء التي يمكن أن نقع فيها بالقياس. عندنا نوعان من الأخطاء: أخطاء عشوائية غير مقصودة، أخطاء منتظمة مقصودة. الأخطاء العشوائية ترجع إلى الصدفة،والخطأ العشوائي قد يرفع درجات الثبات، وقد تنكمش على أثره درجات الثبات. فأخطاء الصدفة أو الأخطاء العشوائية غير الأخطاء المنتظمة.

الأخطاء العشوائية تؤثر على الثبات، والأخطاء المنتظمة على صدق الاختبار، يعني: الخطأ المقصود يؤثر على صدق الاختبار. مثال على ذلك، مثلًا: عندنا اختبار ركل الكرة، قياس المسافة التي يركل بها التلميذ الكرة. الهدف من الاختبار تحديد قدرة الطالب على ركل الكرة. نكلف التلميذ بركل الكرة، ويتم قياس تلك المسافة، لو حدث أن التلميذ دخل الملعب واختار كرة بمحض الصدفة، تلك الكرة كانت جامدة أو منكمشة، فتلك ترجع إلى الصدفة البحتة، وبالتالي أثر الركلة سيختلف بين الكرة التي هي جامدة والكرة الأخرى غير متضخمة أو غير جافة أو يابسة. هنا سيؤثر على معامل الثبات. ولكن لو حدث وأمطرت السماء فإن عدد التكرارات الضربة للطالب برجله ستتأثر أيضًا.

وبذلك لو قِسنا قدرة المتعلم على ركل الكرة المختلفة الحجم أو المختلفة كثافة الكرة من قوة الضغط، أو من انكماش الكرة، أو قسنا ثبات اختبار الطالب أثناء المطر، ففي كلا الحالتين هنا درجة الثبات منخفضة، ولكن ترجع إلى الصدفة المحضة، وكذلك لو قلنا نحسب الصدق فهي غير صادقة. لكن لو طلب أو كلف المدرب ذلك الطالب أن يضرب الكرة، أو يقيس مسافات الكرة على خلال يومين أو ثلاثة أيام، أو على مدار الأسبوع، هل المقياس في اليوم الواحد مثل قياس المسافات على أيام مختلفة؟ طبعًا الخطأ هنا خطأ منظم ومقصود، وبالتالي لا يصلح لقياسه؛ لأننا غيرنا في عامل هنا دخل وغير في أحداث التلميذ، أنه من اليوم الأول إلى اليوم التالي التلميذ هنا اختلف وضعه لعدد من الأسباب.

ممكن أن يكون في اليوم الثاني مريضًا، ممكن أن يكون في اليوم الثاني مكتئبًا، أو عنده أرق لم ينم جيدًا، فتلك كلها عوامل ممكن أن تؤثر في ثبات قوة ضربة أو ركلة الكرة لدى التلميذ.

إذن يجب أن نضع في الاعتبار أن الأخطاء وليدة الصدفة وهي أخطاء عشوائية تؤدي إلى عدم الاتساق بين الضربات بخلاف الأخطاء المقصودة. العوامل التي تؤثر على الثبات بين الضربات هنا، سنجد التلميذ نفسه ممكن أن يتغير، العمل نفسه قد يكون تغير، قد تكون الكرة المستخدمة في اليوم الأول غير الكرة المستخدمة في اليوم الثاني، أيضًا عينة السلوك المستخدمة ممكن أن تتأثر، عينة السلوك المحددة ممكن أن تؤدي إلى عدم الاستقرار. بمعنى: أن الفرد معرض للصدفة، الركلة معرضة للرياح، معرضة للأمطار، معرضة لعوامل كثيرة متعددة ممكن تتأثر بثبات ذلك الاختبار وقياسه.

وبذلك نضع في الاعتبار الفرقَ بين الأخطاء العشوائية وليدة الصدفة، والأخطاء المنظمة التي يمكن أن نتنبأ بها. فالمدرس مثلًا يتحدث بالعامية، والطالب اعتاد على فهم المعاني باللغة العامية، فيأتي في مرة من المرات يتحدث معه باللغة العربية الفصحى، فبالتالي أصبح هنا الخطأ مقصودًا، فقد لا يتمكن الطالب من فهم اللغة، قد يثير قلق الطالب، قد يثير توتر الطالب، ومن ثم تنخفض قدرته على ركل الكرة بطريقة صحيحة.

الأخطاء المنتظمة تؤدي إلى عدم صدق الاختبار،، وبالتالي ستؤثر أيضًا على ثباته، وهنا يحدث تغيير منتظم في الدرجات، التغيير هنا لا يرجع لعملية القياس نفسها، ولكن يرجع إلى سلوك الذي يركل الكرة نفسه، قد تأثر، ومن ثم أثر في الدرجات.

إذن يجب أن نميز بين الأخطاء العشوائية والأخطاء المقصودة.

ممكن أن تكون صادقة إلا إذا كانت ثابتة، إذن المقاييس لا يمكن أن تكون صادقة إلا إذا كانت ثابتة، ومن ثم يتم قياس الثبات قبل قياس الصدق. ولقد سبق أن ذكرنا مقياسًا مثل مقياس الذكاء، كثبات المقياس لا دليل على صحته، ولكنه ليس دليلًا على صدقه، ولكن عدم ثباته يدل على عدم صدقه.

نعم يتأثر، ولكن ليس بنفس الدرجة التي يتأثر بها الأخطاء المنتظمة والمقصودة؛ لأن الأخطاء المنظمة أو المقصودة تؤثر بصورة مباشرة أيضًا على الصدق.

ما العوامل التي تؤثر بصورة مباشرة في المقياس وتتصف بدرجة عالية من السهولة؟ يعني: إلى جانب ذلك أن الأسئلة سهلة جدًّا وبسيطة ومباشرة، تكون عاملًا رئيسًا في أن الدرجة تكون عالية، ولذلك يستجيب لها عدد كبير من الطلبة.

أيضًا قد تكون أسئلة الاختبار كلها غامضة وصعبة فمعظم الطلبة لا تستطيع حلها، ويعطينا الكلام ثبات عال للاختبار.

إذن عندنا الأخطاء العشوائية ممكن تكون عشوائية ترجع إلى طبيعة الأداء، متأصلة في الأداة التي تم إعدادها: أداة القياس أو الاختبار، وممكن تكون الأسباب ترجع إلى الأسلوب أو الطريقة التي يتم من خلالها إجراء الاختبار. مثلًا: الفاحص الذي يطبق تلك الطريقة أو يطبق المقياس، قدرة الفاحص على تطبيق المقياس واستخدامه للأداة، وتعامله مع الطلبة هو مؤثر جدًّا، الظروف الخاصة بإجراء الاختبار ممكن أن يكون في يوم شديد الحرارة، فلا يوجد تهوية جيدة، فتؤثر على قدرة الطالب على التركيز، وحل الاختبار. أيضًا ممكن أن يكون الاختبار التعليمات الموجودة فيه غامضة، أو التعليمات التي يوضحها لنا الفاحص لا تكون جيدة، ومن ثم تؤثر على درجة الاختبار.

أيضًا هناك أخطاء ترجع للتلميذ. إذن عندنا الأخطاء العشوائية مصادرها كم مصدر؟ قلنا: المصدر الأول هنا الأخطاء المتمثلة في الأداء، المصدر الثاني طريقة إجراء ذلك الاختبار، المصدر الثالث يرجع إلى طبيعة التلميذ نفسه، التقلبات التي ممكن يقابلها التلميذ ما الدوافع، الميول، الاتجاهات، الحالة الصحية، القلق، والعوامل الانفعالية، وكل تلك الأمور ممكن تؤثر بصورة مباشرة على ثبات المقياس.

# 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