

**الكشف عن انتهاك افتراض الاستقلال الموضوعي في الاختبارات
الوطنية لضبط نوعية التعليم في الأردن**

بحث منشور في مجلة بحوث جامعة حلب/سلسلة الآداب والعلوم الانسانية
والتربوية. العدد ٦٤ لعام ٢٠٠٨

إعداد

الدكتور عز الدين عبدالله عواد النعيمي

جامعة الإمام محمد بن سعود الإسلامية/كلية العلوم الاجتماعية/قسم علم النفس

الكشف عن انتهاك افتراض الاستقلال الموضوعي في الاختبارات الوطنية لضبط
نوعية التعليم في الأردن

د. عزالدين عبدالله عواد النعيمي

أستاذ مساعد

قسم تربية الطفل

كلية العلوم التربوية

جامعة الإسراء الخاصة

القياس والتقويم التربوي

عمان - الأردن

الكشف عن انتهاك افتراض الاستقلال الموضوعي في الاختبارات الوطنية لضبط نوعية التعليم في الأردن

د. عزالدين النعيمي

أستاذ مساعد في قسم تربية الطفل

كلية العلوم التربوية - جامعة الإسراء الخاصة

الملخص

هدفت الدراسة إلى الكشف عن انتهاك افتراض الاستقلال الموضوعي في الاختبارات الوطنية لضبط نوعية التعليم في الأردن، ولتحقيق هذا الهدف، تم اختيار عينة مكونة من ١٠٠٠ من طلبة الصف الثامن الأساسي في جميع مباحث الاختبار الوطني، وتكونت عينة الصف العاشر الأساسي من ٢٩٨٢ في مبثي اللغة العربية واللغة الإنجليزية (الورقة الأولى)، ٣٠٥٥ في مبثي الرياضيات والعلوم (الورقة الثانية)، ٣١١٩ في مبثي التربية الإسلامية والتربية الاجتماعية والوطنية (الورقة الثالثة)، وقد استخدم المؤشر Q_3 للكشف عن درجة انتهاك افتراض الاستقلال، وتم الاستعانة بالبرامج الإحصائية (Bilog-mg, LDID) لاستخراج النتائج المتعلقة بالكشف عن انتهاك هذا الافتراض، حيث استخدم النموذج اللوجستي ثلاثي المعلم. وأظهرت النتائج الخاصة بالصف الثامن الأساسي، أن جميع المباحث قد انتهكت افتراض الاستقلال الموضوعي، حيث تجاوزت نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي ٩٨%، فيما عدا مبثي التربية الاجتماعية والوطنية، حيث كانت النسبة تساوي ٩١%. وبالنسبة للصف العاشر الأساسي، أظهرت النتائج أن مبثياً فقط قد انتهك افتراض الاستقلال الموضوعي، وهو مبثي اللغة العربية، حيث كانت نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي تساوي ٩٤%، في حين أن المباحث الأخرى لم تنتهك هذا الافتراض، حيث تراوحت النسبة بين ١٣% في مبثي اللغة الإنجليزية، وبين ٢١% في مبثي التربية الإسلامية.

وأوصت الدراسة بضرورة إتباع العديد من الطرق للتخفيف من مقدار الاعتماد الموضوعي في فقرات الاختبار الوطني، منها: كتابة فقرات مستقلة إحصائياً عن بعضها البعض، واستخدام أسلوب العناقيد

الكلمات المفتاحية: نظرية استجابة الفقرة، افتراض الاستقلال الموضوعي، المؤشر Q_3

Exploring Violation of the Local Independence Assumption in the National Education Quality Control Tests in Jordan

Dr. Izzeldin Al-Naimi

Assistant Professor, Child Education Dept.

Faculty of Educational Sciences-Al Isra'a Private University

ABSTRACT

The purpose of this study was to explore violation of local independence Assumption in the national education quality control tests in Jordan. To achieve this goal, a sample of 10,000 primary 8th graders was drawn out and administered the all-subjects national test. The National Test of Arabic & English (booklet1), Mathematics & Science (Booklet2), Islamic Education & Social and National Education (Booklet3) versions were administered to a sample of 2982, 3055 and 3119 to 10th graders respectively. To find out violation degree of the local independence, Q3 index was used. Similarly, to find out results of whether violation of the local independence Assumption exists, Bilog-Mg, LDID statistical programs and the 3-parameters logistic model were used. Results from the primary 8th graders showed that almost all subjects violated the local independence Assumption, where 98% of items were in violation of the local independence Assumption except for the Social & National Education (91%). As for the 10th graders, Results showed that only one subject (Arabic Language) was in violation of the local independence Assumption, where 94% of items were in violation of the local independence Assumption. The remaining subjects were not violating that Assumption, where only 13% and 21% of items were in violation of the local independence Assumption for the English and Islamic Education subjects respectively.

The study recommended using various methods to mitigate the local correlation among national test items including developing statistically independence items and using testlets method.

Keywords: Item Response Theory, Local independence Assumption, Q3 Index

تستخدم نظرية استجابة الفقرة (IRT) Item Response Theory بشكل واسع من قبل المتخصصين في القياس النفسي والتربوي^(١) [١٤]. ونماذج استجابة الفقرة أحادية البعد نماذج رياضية تعرّف العلاقة بين أداء المفحوص الملاحظ على الاختبار وبين السمة أو القدرة الغير الملاحظة للمفحوص والتي تكون مسؤولة عن أداء المفحوص على الاختبار. وترتكز نماذج نظرية استجابة الفقرة IRT Models، مثل كثير من النماذج الرياضية على افتراضات عدة يجب تحققها في البيانات التي تخضع للتحليل باستخدام هذه النماذج، وهي أحادية البعد للسمة التي يقيسها الاختبار Unidimensionality، والاستقلال الموضوعي لل فقرات Local Item Independence (LII)، والعلاقة الوتيرية بين القدرة والأداء على الفقرة Item Characteristic Curve (ICC)، وعدم السرعة Non-Speededness^(٢) [٧].

ومن المعروف أن حجر الزاوية في الاختبار هي الفقرة، فالمؤسسات المسؤولة عن تطوير الاختبارات تقضي جلّ وقتها واهتمامها في اختيار وتطوير الفقرات ليتم تضمينها فيه. إذ ينبغي أن تغطي هذه الفقرات المحتوى أو المجال المراد قياسه وتعمل على إيجاد تقديرات دقيقة لقدرات المفحوصين^(٣) [١٧]. ومن الطرق المعروفة لزيادة ثبات علامات الاختبار زيادة عدد الفقرات لذلك الاختبار، ولكن تضمين فقرات تكرر بعضها البعض لن يحقق الهدف بالنسبة لقياس ثابت وصادق بنفس الوقت، ولهذا فإن مطوري الاختبار يحاولون اختيار فقرات تزودنا بمعلومات فريدة، بمعزل عن أي فقرة أخرى في الاختبار، حول قدرات ومهارات الأفراد الذين يستجيبوا لفقرات الاختبار.

وتفترض كل من النظرية التقليدية في القياس Classical Test Theory (CTT)، وكذلك نظرية استجابة الفقرة Item Response Theory (IRT) بأن فقرات الاختبار مستقلة موضعياً locally Independent، ويوضح لي (Lee, 2004)^(٤) [١٠] أن نظرية القياس التقليدية تفترض بأن أخطاء القياس مستقلة إحصائياً بين الفقرات المختلفة، وبالتالي فهي مستقلة عبر الاختبار ككل، أما في نظرية استجابة الفقرة فإن افتراض الاستقلال الموضوعي للفقرات يعني أن الاستجابات لفقرتين مختلفتين ينبغي أن تكون مستقلة إحصائياً عن بعضها البعض للمفحوصين عند مستوى قدرة معين، أي أن استجابة المفحوص

^(١) XIE, Y., 2001-Dimensionality, Dependence, or Both? An Application of the Item Bundle Model to Multidimensional Data. Policy, Organization, Measurement & Evaluation Graduate School of Education: University of California at Erkeley

^(٢) HAMBLETON, R., 1989-Principles and selected applications of item response theory in R.L. Linn (Ed.), Educational measurement (3rd ed. PP. 147-200) New York: American Council on Education and Macmillan.

^(٣) ZENISKY, A., HAMBLETON, R., & SIRECI, S., 2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/mcat/research/monograph5.df.

^(٤) LEE, Y., 2004-Examining passage – related local item dependence (LID) and measurement construct using Q₃ statistics in an EFL reading comprehension test. Language testing, vol. 21, No. 1, pp: 74-100.

فقرة ما لا تتأثر سواء إيجاباً أو سلباً باستجابته على فقرة أخرى. وإذا ما تم انتهاك افتراض الاستقلال الموضوعي بين الفقرات سيظهر ما يسمى بالارتباط الموضوعي بين الفقرات (LID) Local Item Dependence ، وأشار زينسكي وهامبلتون وسيريبي (Zenisky, Hambleton & Sireci, 2002) ⁽¹⁾ إلى أن وجود مثل هذا الارتباط بين الفقرات قد يؤدي إلى تقديرات غير دقيقة لمعالم الفقرة وإحصائيات الاختبار وقدرة المفحوص، وأن مثل هذا الارتباط يضيف بُعداً إضافياً (وهو غالباً غير مقصود) للاختبار ويكون على حساب البناء الذي نهتم به. وبالتالي فإذا كانت هناك سلسلة من الفقرات تتبع نص معين، فإنه من الممكن أن افتراض الاستقلال الموضوعي للفقرة (LII) Local Item Independence سيتم انتهاكه بين هذه الفقرات، ونتيجة لذلك فإنه سيظهر تقديرات غير دقيقة لقدرات المفحوصين ودالة معلومات الفقرة وثبات الاختبار إذا ما تم تطبيق نماذج نظرية استجابة الفقرة التي تفترض الاستقلال الموضوعي للفقرات على الاختبارات التي تشتمل على مجموعة فقرات مرتبطة موضعياً Locally dependent Items. وهناك العديد من الدراسات وضحت المشاكل التي يمكن أن تنجم عند عدم أخذ الارتباط الموضوعي للفقرات (LID) بعين الاعتبار عند إجراء التحليلات المختلفة. ففي دراسة سيريبي وآخرون وثيسن وآخرون (Sirecie, et al., 1991; Thisson, et al., 1989; Cited in Zenisky et al., 2002) ⁽²⁾ تم تحليل الفقرات المتعلقة بأسئلة استيعاب المقروء ووجدوا بأنه عندما عوملت الفقرات وبشكل خاطئ بأنها منفصلة وبأنها فقرات مستقلة موضعياً كانت تقديرات دالة معلومات الاختبار والثبات متحيزة وبشكل خطير لأعلى، وهذه الخطورة لها خصوصية في اختبارات التكيف المحوسبة (CAT) Computerized Adaptive Test حيث يستخدم الخطأ المعياري في التقدير (SEE) Standard Error of Estimate غالباً كمييار للانتهاء من الاختبار بالنسبة للمفحوص، ونحن نعرف بأن الخطأ المعياري في التقدير هو معكوس دالة معلومات الاختبار $(SEE = \frac{1}{\sqrt{I(\Phi)}}$ ، وبالتالي تحيز دالة معلومات أعلى سيؤدي إلى أن يجعل الخطأ المعياري في التقدير قليل والذي يؤدي بدوره إلى الانتهاء مبكراً للاختبار لذلك المفحوص (Fennessy, 1995; Cited in Zenisky et al., 2002) ⁽³⁾ [17].

⁽¹⁾ZENISKY, A., HAMBLETON, R., & SIRECI, S.,2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/mcat/research/monograph5.df.

⁽²⁾ZENISKY, A., HAMBLETON, R., & SIRECI, S.,2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/mcat/research/monograph5.df.

⁽³⁾ZENISKY, A., HAMBLETON, R., & SIRECI, S.,2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/mcat/research/monograph5.df.

وتعاني الطرق التقليدية في القياس أيضاً من عدم الدقة في التقديرات المختلفة الناتجة من عدم أخذ الارتباطات البينية للفقرات بالحسبان، فكما يشير زينسكي وآخرون (Zenisky et al., 2002) ⁽¹⁾ [١٧] إلى أن هناك عدة باحثين (Anastasi, 1961; Guilford, 1936; Kelly, 1927; Thorndike, 1951) وغيرهم نبهوا إلى ضرورة أن الفقرات التي تنتمي إلى سيناريو أو مثير مشترك مثل مجموعة الفقرات المرتبطة بفقرة استيعاب المقروء أو المرتبطة بجدول معين أو شكل أو خريطة معينة وغيرها ينبغي وضعها في نفس نصف الاختبار عند تقدير الثبات باستخدام طريقة التجزئة النصفية، وعدم فعل ذلك سيؤدي إلى تضخيم الثبات المقدّر لأن ذلك سيؤدي إلى تضخيم معامل الارتباط بين نصفي الاختبار. ومن المعروف أن معامل كرونباخ ألفا أيضاً يمثل كل الأجزاء النصفية المحتملة. وبالتالي سيكون هذا المعامل أيضاً متحيزاً لأعلى من خلال عدم مراعاة هذه الارتباطات البينية للفقرات، وبالتالي فإن المشاكل المتعلقة بالارتباط الموضوعي للفقرات لا تقتصر على نظرية استجابة الفقرة لوحدها.

وإذا ما ظهرت مثل هذه الارتباطات الموضوعية بين الفقرات، فإن إحدى الطرق المستخدمة في تصحيح مثل هذه الفقرات هو استخدام عنقود Testlet أي تجميع الفقرات التي بينها ارتباط موضوعي ومعاملتها كأنها فقرة عديدة الاستجابات Polytomous item ، وبعد ذلك نستخدم أحد نماذج نظرية استجابة الفقرة عديدة الاستجابات Polytomous IRT Models، فعلى سبيل المثال فإن فقرة استيعاب المقروء في امتحان اللغة الإنجليزية وما تحتويه من فقرات لقياس استيعاب المفحوصين لهذه الفقرة يمكن أن تشكل عنقود Testlet، وإذا ما استخدمنا نموذج لنظرية استجابة الفقرة عديدة الاستجابات Polytomous IRT Models لتصحيح هذه العناقيد Testlets، فإنه بإمكاننا تحليل البيانات مع المحافظة على الاستقلال الموضوعي بين العناقيد المختلفة. وبالتالي نستطيع ومن خلال استخدام ما يسمى Testlet (عنقود) حل مشكلة الاستقلال الموضوعي بين الفقرات وتقدير المعامل المختلفة سواء للفقرة أو للمفحوص أو للاختبار ككل.

وباستخدام نظرية استجابة الفقرة (IRT) في تحليل بيانات الاختبارات النفسية والتربوية، تم إيجاد حلول لمعظم المشكلات التي تعاني منها النظرية التقليدية في القياس (CTT) Classical Test Theory. ومن أهم هذه الحلول الحصول على تقديرات لمعامل الفقرة متحررة من قدرات المفحوصين، وكذلك الحصول على تقديرات لمعامل القدرة متحررة من خصائص الفقرات ⁽²⁾ [٨].

ويعتبر افتراض الاستقلال الموضوعي للفقرة من الافتراضات الرئيسة لكل من النظرية التقليدية في القياس (CTT) ونظرية استجابة الفقرة (IRT) ⁽³⁾ [١٠]، حيث أشار هامبلتون وسوامينثان (Hambleton)

⁽¹⁾ZENISKY, A., HAMBLETON, R., & SIRECI, S.,2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/mcat/research/monograph5.df.

⁽²⁾HAMBLETON, R., & SWAMINATHAN, H.,1985-Item response theory: principles and applications. Boston, MA: Kluwer - Nijhoff.

⁽³⁾LEE, Y.,2004-Examining passage - related local item dependence (LID) and measurement construct using Q₃ statistics in an EFL reading comprehension test. Language testing.vol. 21,No.1, pp:74-100.

Swaminathan, 1985) & (1) [8] إلى أن افتراض الاستقلال الموضوعي يعني أن استجابات المفحوصين على فقرات الاختبار تكون مستقلة إحصائياً عند مستوى قدرة معينة، أي أن استجابة المفحوص لفقرة ما يجب أن لا تؤثر سلباً أو إيجاباً على استجابته لفقرة أخرى؛ بمعنى أن إجابة الفقرة لا تعطي أية تلميحات أو معيقات لإجابة فقرة أخرى ، وهذا الافتراض يوضح أن قدرة المفحوص وخصائص الفقرة هي وحدها التي تؤثر في استجابته على فقرات الاختبار. ويتحقق افتراض الاستقلال الموضوعي للفقرة (LII) إذا كانت مجموعة الفقرات أحادية البعد، أي إذا كانت تقيس قدرة أو سمة واحدة، فإذا لم يتحقق افتراض أحادية البعد فإن درجات فقرات الاختبار ستكون مرتبطة عند مستوى قدرة معين لأن فقرات الاختبار تقيس سمة ثنائية أو أكثر (1) [8].

وذكرت ين (Yen, 1993) (2) [16] أن من أهم الأسباب لظهور الاعتماد الموضوعي (LD) بين

فقرات الاختبار، هي المساعدة الخارجية والسرعة والتبعية لنص معين .

وقد ظهرت العديد من المؤشرات التي تستخدم للكشف عن انتهاك افتراض الاستقلال الموضوعي، ومن أكثر هذه المؤشرات استخداماً المؤشر Q_3 ، وهو عبارة عن معامل الارتباط للبواقي لزوج من الفقرات بعد ضبط السمة المقدره وقد استخدم في عدد من الدراسات للكشف عن انتهاك الاستقلال الموضوعي (4) [4,11,17]، والمؤشر الإحصائي G^2 ، وهو يبين الفروق بين ما هو متوقع وبين ما هو ملاحظ لتكرار الاستجابات لزوج من الفقرات، وقد استخدم هذا المؤشر في عدد من الدراسات (5) [4,13]، والمؤشر فيشر (ز) Fisher Z الذي استخدمه شن (Shen, 1997) (1) [12]، وهو تحويل للإحصائي Q_3 ، ويتم ذلك بتحويل الأخطاء الملاحظة إلى أخطاء معيارية.

(1) HAMBLETON, R., & SWAMINATHAN, H., 1985-Item response theory: principles and applications. Boston, MA: Kluwer - Nijhoff.

(2) HAMBLETON, R., & SWAMINATHAN, H., 1985-Item response theory: principles and applications. Boston, MA: Kluwer - Nijhoff.

(3) YEN, W., 1993-Scaling Performance Assessments: Strategies for Managing Local item Dependence. Journal of Educational Measurements, vol.30, No.3, pp 187-213.

(4) CHEN, W.H. & THISEN, D., 1997-Local dependence indexes for item pairs using item response theory. Journal of Educational and Behavioral Statistics, vol. 22, pp: 265-289.

REESE, L., 1999-A classical Test Theory perspective on LSAT local Item Dependence. Report No. LSAC-R-96-01. Law School Admission council, Newtown, PA (Document Reproduction Service No. ED 469172).

ZENISKY, A., HAMBLETON, R., & SIRECI, S., 2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/ mcat/ research/ monograph5. df.

(5) CHEN, W.H. & THISEN, D., 1997-Local dependence indexes for item pairs using item response theory. Journal of Educational and Behavioral Statistics, vol. 22, pp: 265-289.

THOMPSON, T. & POMMERICH, M., 1996-Examining the sources and effects of local item dependence. paper presented at the Annual meeting of the American Educational Research Association. New York, NY, April 8-12, 1996 (ERIC Document Reproduction Service No. ED 400311).

(6) SHEN, L., 1997-Quantifying Item Dependency by Fisher's Z. paper presented At the annual meeting of the American Educational Research Association, Chicago, IL, March 24-28, 1997 (ERIC Document Reproduction Service No. ED 410241).

ولابد من التأكيد هنا على أهمية افتراض الاستقلال الموضوعي، حيث إن انتهاك هذا الافتراض سيؤثر على النتائج المتعلقة بتقدير معالم الفقرات والمفحوصين، وذلك لأن المعادلة المحورية التي تقوم عليها نماذج نظرية استجابة الفقرة أحادية البعد هي المعادلة التي يتم من خلالها حساب أرجحية نمط استجابة المفحوص، وانتهاك الاستقلال الموضوعي سيؤدي إلى التأثير على المعالم المقدره والمعادلة هي :

$$P_i^{u_i} Q_i^{(1-u_i)} L(u_1, u_2, u_3, \dots, u_n / \theta) = \prod_{i=1}^n \dots (1)$$

حيث u_i : استجابة المفحوص على الفقرة، وتحمل قيمة الواحد الصحيح عند استجابة المفحوص على الفقرة استجابة صحيحة ، والقيمة صفر عند الاستجابة الخاطئة للمفحوص على الفقرة، و $L(u_1, u_2, u_3, \dots, u_n / \theta)$ تمثل الاحتمال المشروط لنمط استجابة المفحوص ، ومن خلال هذه المعادلة سيكون احتمال إجابة مفحوص إجابة صحيحة على أي فقرتين يساوي حاصل ضرب احتمال إجابة المفحوص على الفقرة الأولى إجابة صحيحة في احتمال إجابة المفحوص على الفقرة الثانية إجابة صحيحة، و يتحقق هذا الشرط في حالة استجابة الفقرات لافتراض الاستقلال الموضوعي للفقرات، أما إذا لم يتحقق هذا الشرط فان هنالك خطورة تقع على دقة تقدير كل من معالم الفقرات والمفحوصين⁽¹⁾ [6,8].

٢ - الاختبار الوطني لضبط نوعية التعليم في الأردن

يعد مستوى أداء الطلبة على اختبارات التحصيل معياراً أساسياً للحكم على فعالية النظم التربوية القائمة، ووسيلة تساعد في التخطيط الفعال لمستقبل المسيرة التربوية في أي بلد، إذ يمكن في ضوء نتائج ذلك الأداء رسم سياسة تربوية شاملة تهدف إلى تحديد مدخلات العملية التربوية الفعالة، ومن ثم تحسينها، وكذلك مراقبة وتطوير عملياتها وتقويم مخرجاتها. ومن هنا جاء اهتمام القائمين على العملية التعليمية في الأردن على إعداد اختبار وطني لقياس مدى امتلاك طلبة الفئة المستهدفة المهارات الأساسية التراكمية لعدد من المباحث الدراسية، حيث تقوم عدة لجان مختارة بعناية بعملية إعداد هذا الاختبار، ويتم انقاء فقرات الاختبار النهائية بالاعتماد على التحليلات الإحصائية للفقرات، ويتم الحكم على جودة الفقرات بالاعتماد على بعض الإحصائيات الخاصة بالفترة كعاملات الصعوبة والتمييز، والتي يتم تحليلها وفق النماذج الكلاسيكية والحديثة. وللوقوف على دقة الاختيار لهذه الفقرات لابد من القائمين على مثل هذه الاختبارات أن يكونوا واعين لمخاطر عدم تحقق الافتراضات التي تقوم عليها هذه النماذج، ومن أهم هذه الافتراضات افتراض الاستقلال الموضوعي للفقرات، والذي إذا ما تم انتهاكه أدى إلى تقديرات غير دقيقة للإحصائيات الخاصة بالاختبار ومعالم الفقرات. ويعد قياس مستوى أداء الطلبة على الاختبار الوطني لضبط نوعية التعليم معياراً

(1) EMBRETSON, S. & REIASE, S., 2000-Item response theory for psychologists. New Jersey: Lawrence Erlbaum Associates, Inc.

- HAMBLETON, R., & SWAMINATHAN, H., 1985-Item response theory: principles and applications. Boston, MA: Kluwer - Nijhoff.

يمكن من خلاله تحديد مدى تحقيق الطلبة للنتائج المتوقع منهم تحقيقها، والكشف عن فاعلية العملية التعليمية التعليمية، وتحديد نقاط القوة لتعزيزها، ونقاط الضعف لمعالجتها.

٣- مؤشر Q3

اقترحت ين (Yen,1984) ^(١) [١٥] المؤشر الإحصائي Q3 كمؤشر للكشف عن الارتباط الموضوعي بين فقرات الاختبار. ومؤشر Q3 هو معامل الارتباط للبواقي لزوج من الفقرات بعد ضبط السمة المقدره، ولحساب المؤشر Q3 ينبغي تقدير القدرة لكل مفحوص وذلك لاستخدام هذه القدرة لتقدير احتمال الإجابة الصحيحة للمفحوص الذي يمتلك هذه القدرة لكل الفقرات، ويتم حساب الباقي ويرمز له بالرمز (d_{ja}) ، وذلك بأخذ الفرق بين الأداء الملاحظ والأداء المتوقع للمفحوص على الفقرة، ويمكن التعبير عن ذلك رياضياً كما يلي :

$$d_{ja} = U_{ja} - P_j(\theta_a) \dots \dots \dots (2)$$

حيث (U_{ja}) تمثل العلامة الملاحظة للمفحوص (a) على الفقرة (j) وهي إما واحد عند الاستجابة الصحيحة أو صفر عند الاستجابة الخاطئة، و $P_j(\theta_a)$ تمثل احتمال إجابة الفقرة (j) إجابة صحيحة من قبل المفحوص (a) وهي تقع بين الصفر والواحد الصحيح. وبالتالي لو تم أخذ الفقرتين (i, j) ، فإن $Q3$ يكون عبارة عن معامل الارتباط بين البواقي لهاتين الفقرتين لجميع المفحوصين، ويمكن التعبير عنه كما يلي:

$$Q3_{ji} = r(d_j, d_i) \dots \dots \dots (3)$$

وقد تم استخدام هذا المؤشر الإحصائي وأثبتت فاعليته في الكشف عن الارتباط الموضوعي للفقرات في عدة

دراسات ^(٢) [٤,15,16,17].

٤- مشكلة الدراسة وأسئلتها :

بعد افتراض الاستقلال الموضوعي من الافتراضات الأساسية التي تقوم عليها نماذج نظرية استجابة الفقرة ، وذلك لان ظهور أي انتهاك لافتراض الاستقلال الموضوعي سيؤثر على التقديرات المختلفة والمتعلقة بمعالم الفقرات والاختبار ومعالم المفحوصين ، ومن هنا لابد من استخدام مؤشرات فعالة في الكشف عن الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي، حتى نتقاضي أي تأثير على تقديرات نماذج استجابة الفقرة.

مجلة بحوث حلب

^(١)YEN, W.,1984-Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. Applied psychological Measurement,vol. 8,pp: 125-145.

^(٢) CHEN,W.H. & THISSEN,D.,1997-Local dependence indexes for item pairs using item response theory. Journal of Educational and Behavioral Statistics,vol. 22,pp: 265-289.

YEN, W.,1984-Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. Applied psychological Measurement,vol. 8,pp: 125-145.

YEN, W.,1993-Scaling Performance Assessments:Strategies for Managing Local item Dependence. Journal of Educational Measurements,vol.30,No.3,pp 187-213.

ZENISKY, A., HAMBLETON, R., & SIRECI, S.,2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/ mcat/ research/monograph5. df.

وهدفت هذه الدراسة إلى الكشف عن درجة انتهاك افتراض الاستقلال الموضوعي لل فقرات في مباحث الاختبار الوطني لضبط نوعية التعليم المستخدم في وزارة التربية والتعليم الأردنية، وبشكل أكثر تحديد حاولت الدراسة الإجابة عن السؤال البحثي الآتي:

١- هل يوجد هناك انتهاك لافتراض الاستقلال الموضوعي في مباحث الاختبار الوطني لضبط نوعية التعليم في الأردن؟ وهل تختلف نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي باختلاف المباحث الدراسية؟

٥- التعريفات الإجرائية:

الاختبار الوطني لضبط نوعية التعليم: اختبار تعدد وزارة التربية والتعليم (مديرية الاختبارات) لقياس مدى امتلاك طلبة الفئة المستهدفة المهارات الأساسية التراكمية لعدد من المباحث الدراسية (إدارة الامتحانات والاختبارات العامة، وزارة التربية والتعليم الأردنية).

٦- محددات الدراسة

١- اقتصرت الدراسة على استخدام النموذج اللوجستي ثلاثي المعلمة للفقرات التي تتطلب الاستجابة الثنائية.

٢- اقتصرت الدراسة على استخدام المؤشر Q3 للكشف عن افتراض الاستقلال الموضوعي.

٣- اهتمت الدراسة بافتراض الاستقلال الموضوعي للنماذج أحادية البعد فقط.

٧- الطريقة الإجرائية

يحتوي هذا الجزء وصفاً للبرامج المستخدمة في الدراسة، ووصفاً لعينة الدراسة، والطريقة التي تمّ بها استخدام البيانات لحساب المؤشر Q3 والذي يكشف عن أزواج الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي، وذلك للإجابة عن أسئلة الدراسة حول درجة انتهاك مباحث الاختبار الوطني لضبط نوعية التعليم في الأردن لافتراض الاستقلال الموضوعي لكل من الصفين الثامن والعاشر الأساسيين.

٧-١ البرامج المستخدمة في الدراسة

- BILOG : وهو برنامج يستخدم لتحليل بيانات تحتوي على فقرات ثنائية Binary Items للاختبارات ذات التصحيح الثنائي ويستخدم لتقدير كل من معالم الفقرات (الصعوبة، والتمييز، والتخمين) ، ومعالم المفحوصين، حيث تتضمن عملية التحليل ثلاث مراحل، هي:

١. المرحلة الأولى (Phase 1): في هذه المرحلة يتم تقدير إحصائيات الفقرات حسب النظرية التقليدية في القياس مثل معامل صعوبة الفقرة وهو نسبة الاستجابات الصحيحة على الفقرة، ومعامل تمييز الفقرة وهو معامل ارتباط بوينت بيسيرال بين علامة الفقرة والعلامة الكلية على الاختبار.

٢. المرحلة الثانية (Phase 2): في هذه المرحلة يتم تقدير معالم الفقرة ، كما يتم تحديد الفقرات الملائمة للنموذج المستخدم .

٣. المرحلة الثالثة (Phase 3): في هذه المرحلة يتم تقدير علامات القدرة و معالم الفقرة حسب النموذج المستخدم في التحليل.

واستخدم في هذه الدراسة النموذج اللوجستي الثلاثي المعلم، والذي يتم من خلاله تقدير معالم الفقرة الثلاثة: الصعوبة، والتمييز، والتخمين، وتقدير معلمة القدرة للمفحوصين.

- برنامج مؤشرات الاعتماد الموضوعي للفقرات الثنائية:

A Computer program for Local Dependence Indices For Dichotomous Items (LDID) وهو برنامج مكتوب بلغة فورتران ٩٠، يتم من خلاله حساب المؤشر Q3، ويحتاج تشغيل هذا البرنامج إلى ثلاثة ملفات هي: ملف قدرات المفحوصين المقدر، و ملف البيانات الخام المولدة، و ملف معالم الفقرات المقدر (الصعوبة، والتمييز، والتخمين)، والتي يتم استغلالها من مخرجات التحليل لبرنامج BILOG ، وبالاعتماد على برنامج LDID تم اعتماد القيمة ٠.٠٥ كحد فاصل بين أزواج الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي من غيرها من الفقرات عندما تم التعامل مع المؤشر Q3 ؛ حيث إذا كانت القيمة المطلقة لأزواج الفقرات عند استخدام المؤشر Q3 أعلى من ٠.٠٥ اعتبرت هذه الأزواج منتهكة لافتراض الاستقلال الموضوعي، وإذا كانت القيمة المطلقة للمؤشر Q3 أقل من ٠.٠٥ كانت أزواج الفقرات مستقلة موضعياً، وتعني القيمة الموجبة للمؤشر Q3 أن إجابة فقرة ما تؤثر إيجاباً على إجابة فقرة أخرى بينما القيمة السالبة تعني أن إجابة فقرة ما تؤثر سلباً على إجابة فقرة أخرى (١) [٩].

٧-٢ عينة الدراسة

تم تطبيق الاختبار الوطني لضبط نوعية التعليم للعام الدراسي ٢٠٠٧/٢٠٠٨ على جميع طلبة الصف الثامن الأساسي في الأردن، وتم تطبيقه على عينة عشوائية لطلبة الصف العاشر الأساسي، والجدول الآتي يوضح عينة الدراسة التي تم تطبيق هذه الدراسة عليهم، حيث اعتمد الباحث على نفس عينة طلبة الصف العاشر الأساسي، وفي حين تم اختيار عينة عشوائية من ١٠٠٠٠ طالب وطالبة للصف الثامن الأساسي.

(1) KIM, S; COHEN, A. & LIN, Y., 2005-LDID: A Computer Program for Local Dependence Indices for Dichotomous Items. Version 1.0.

الجدول ١

عينة الدراسة لطلبة الصف الثامن والعاشر الأساسي حسب مباحث الاختبار الوطني لضبط نوعية التعليم

الصف	المباحث				
	اللغة العربية	اللغة الإنجليزية	الرياضيات	العلوم	التربية الإسلامية والتربية الاجتماعية والوطنية
الثامن	١٠٠٠٠	١٠٠٠٠	١٠٠٠٠	١٠٠٠٠	١٠٠٠٠
العاشر	٢٩٨٢	٢٩٨٢	٣٠٥٥	٣٠٥٥	٣١١٩

٣-٧ أداة الدراسة

تم استخدام بيانات الاختبار الوطني لضبط نوعية التعليم في الأردن، والذي يتم تطبيقه سنوياً على عينات مختارة من طلبة الصف الثامن والعاشر الأساسي، حيث تم تقسيم الاختبار الوطني إلى ثلاث أوراق اختبارية هي: الورقة الأولى (ورقة اللغات): وتشمل اللغة العربية واللغة الإنجليزية، والورقة الثانية (الورقة العلمية): وتشمل الرياضيات والعلوم، والورقة الثالثة (الورقة الثقافية) وتشمل التربية الإسلامية والتربية الاجتماعية والوطنية، ويخضع كل طالب إلى ورقة اختبارية واحدة فقط وبشكل عشوائي. ويضم كل اختبار ٣٠ فقرة من نوع اختيار من أربعة بدائل. وقد تم الحصول على هذه البيانات من وزارة التربية والتعليم الأردنية.

٨- النتائج

يتضمن هذا الجزء عرضاً لنتائج الدراسة المتعلقة بنسبة الفقرات التي كشف المؤشر Q_3 عن وجود انتهاك لافتراض الاستقلال الموضوعي فيما بينها لمباحث الاختبار الوطني لضبط نوعية التعليم لكل من الصف الثامن والعاشر الأساسيين، والمتعلقة بالتعرف على أكثر المباحث انتهاكاً لهذا الافتراض.

وللإجابة عن السؤال البحثي الرئيس: هل يوجد هناك انتهاك لافتراض الاستقلال الموضوعي في مباحث الاختبار الوطني لضبط نوعية التعليم في الأردن؟ وهل تختلف نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي باختلاف المباحث الدراسية؟

تمّ حساب نسبة أزواج الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي من خلال قسمة عدد أزواج الفقرات و التي بينها انتهاك لافتراض الاستقلال الموضوعي على عدد أزواج الفقرات الكلية والعلاقة الآتية توضح كيفية حساب نسبة أزواج الفقرات:

نسبة أزواج الفقرات = عدد أزواج الفقرات / ٤٣٥ (٤)

وعدد أزواج الفقرات في كل اختبار من الاختبارات الوطنية لضبط نوعية التعليم قد بلغ ٤٣٥، وذلك لان عدد الفقرات في كل حالة من الحالات بلغ ٣٠ فقرة، ومن خواص التوافق الرياضية فان عدد أزواج الفقرات لاختبار عدد فقراته ٣٠ فقرة يساوي ٤٣٥ وهذه قيمة ثابتة في كل الحالات.

د. عزالدين النعيمي

ويبين الجدول ٢ نسبة أزواج الفقرات التي يوجد بينها انتهاك لافتراض الاستقلال الموضوعي ولجميع المباحث الدراسية لكل من الصف الثامن والعاشر الأساسيين ، وذلك باستخدام المؤشر Q_3 .

الجدول ٢

نسبة أزواج الفقرات للمؤشر Q_3 التي بينها انتهاك لافتراض الاستقلال الموضوعي لمباحث الاختبار الوطني لكل من الصف الثامن والعاشر الأساسيين

الصف	المباحث					
	اللغة العربية	اللغة الإنجليزية	الرياضيات	العلوم	التربية الإسلامية	التربية الاجتماعية والوطنية
الثامن	٠.٩٨٦	٠.٩٩٨	١	٠.٩٨٦	١	٠.٩٠٦
العاشر	٠.٩٣٨	٠.١٢٩	٠.١٩٣	٠.١٨٩	٠.٢١٤	٠.١٩٨

بالنسبة للصف الثامن الأساسي، يتبين من الجدول ٢ أن جميع المباحث كانت فقراتها منتهكة لافتراض الاستقلال الموضوعي، حيث كانت نسبة الفقرات التي بينها اعتماد موضوعي في هذه المباحث تتجاوز ٩٠%، وكانت تتراوح هذه النسب من ٠.٩٠٦ إلى ١، مما يعني أن هذه الافتراض قد تم انتهاكه بصورة كبيرة. وقد كانت نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال تتجاوز ٩٨% لجميع المباحث، فيما عدا التربية الوطنية والاجتماعية، حيث كانت النسبة تساوي حوالي ٩١%. ويلاحظ أن جميع فقرات الرياضيات والتربية الإسلامية قد انتهكت افتراض الاستقلال الموضوعي باستخدام هذا المؤشر.

وبالنسبة للصف العاشر الأساسي، يتبين من الجدول ٢ أن جميع المباحث كانت نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي أقل مما هو عليه في الصف الثامن الأساسي، حيث تراوحت نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي من ٠.١٢٩ إلى ٠.٩٣٨. وكانت نسب الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي متدنية فيما عدا مبحث اللغة العربية، حيث كانت نسب الفقرات التي بينها انتهاك مرتفعة جداً لهذا المبحث مقارنة مع باقي المباحث، ويلاحظ من الجدول السابق أن أقل المباحث انتهاكاً لهذا الافتراض مبحث اللغة الإنجليزية، حيث لم تتجاوز نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي ١٣%، يليه مبحث العلوم، حيث لم تتجاوز نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي ١٩%، ونلاحظ أن باقي المباحث لم تتجاوز نسبة الفقرات التي بينها انتهاك لهذا الافتراض ٢٠%، مما يعني أن هذه المباحث باستثناء مبحث اللغة العربية كان انتهاكها لافتراض قليلاً.

ولأخذ صورة أوضح حول انتهاك هذه المباحث لافتراض الاستقلال الموضوعي، تم إيجاد الوسيط للمؤشر Q₃ ولكل مبحث من مباحث الاختبار الوطني في الصفين الثامن والعاشر الأساسيين، والجدول ٣ يوضح هذه الوسيطات للمباحث المختلفة.

الجدول ٣

الوسيط للمؤشر Q₃ للمباحث المختلفة في كل من الصف الثامن والعاشر الأساسيين

الصف	المباحث					
	اللغة العربية	اللغة الإنجليزية	الرياضيات	العلوم	التربية الإسلامية	التربية الاجتماعية والوطنية
الثامن	٠.٢٤٣	٠.٣٠٨	٠.٢٩٧	٠.٢٠٩	٠.٢٩٠	٠.١٦٨
العاشر	٠.٢٦١	٠.٠١٣-	٠.٠٢٥-	٠.٠٢١-	٠.٠٢٩-	٠.٠٢٨-

يتبين من الجدول ٣ أنه بالنسبة إلى الصف الثامن الأساسي كانت الوسيطات للمؤشر Q₃ أعلى من ٠.٠٥ وهذا يؤكد بأن جميع المباحث قد انتهك فيها افتراض الاستقلال الموضوعي، حيث تراوحت هذه الوسيطات بين ٠.١٦٨ إلى ٠.٣٠٨. وكانت جميع المباحث يتجاوز الوسيط فيها النسبة ٠.٢٠، فيما عدا مبحث التربية الاجتماعية والوطنية، وانفقت هذه النتائج مع ما تم عرضه سابقاً فيما يتعلق بنسبة الأزواج التي بينتها انتهاك لافتراض الاستقلال الموضوعي.

ويتبين من الجدول ٣، أنه بالنسبة للصف العاشر الأساسي، كانت الوسيطات للمؤشر Q₃ أقل من ٠.٠٥، فيما عدا مبحث اللغة العربية، وهذا يؤكد بأن هذه المباحث قد استجابت لافتراض الاستقلال الموضوعي، وكانت أكثر المباحث استجابةً لهذا الافتراض مبحث اللغة الإنجليزية، وكان الوسيط يساوي ٠.٠١٣-، في حين أن مبحث اللغة العربية قد انتهك هذا الافتراض، حيث كان الوسيط يساوي ٠.٢٦١، وانفقت هذه النتائج أيضاً مع ما تم عرضه سابقاً فيما يتعلق بنسبة الأزواج التي بينتها انتهاك لافتراض الاستقلال الموضوعي.

٩- مناقشة النتائج والتوصيات

أشارت النتائج المتعلقة بالصف الثامن الأساسي، أن جميع مباحث الاختبار الوطني قد انتهكت افتراض الاستقلال الموضوعي، حيث كانت نسبة الفقرات التي بينها اعتماد موضوعي في هذه المباحث تتجاوز ٩٠%، وأن أعلى المباحث انتهاكاً لافتراض الاستقلال الموضوعي كانت في مباحث الرياضيات والتربية الإسلامية واللغة الإنجليزية واللغة العربية والعلوم على الترتيب، حيث تجاوزت نسبة الفقرات التي بينها انتهاك لهذا الافتراض ٩٨%، وفي المقابل نجد أن أقل المباحث انتهاكاً لافتراض الاستقلال الموضوعي كانت في مبحث التربية الوطنية والاجتماعية، حيث كانت نسبة أزواج الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي تساوي ٠٠.٩٠٦، وبالرجوع إلى الوسيطات، نجد أن أكثر المباحث انتهاكاً لهذا الافتراض كانت في مباحث اللغة الإنجليزية والرياضيات والتربية الإسلامية واللغة العربية والعلوم، حيث تجاوزت القيم للمؤشر Q_3 ٠٠.٢٠، وفي المقابل نجد أن أقل المباحث انتهاكاً لافتراض الاستقلال الموضوعي كانت في مبحث التربية الاجتماعية والوطنية، حيث لم تتجاوز قيمة المؤشر Q_3 ٠٠.١٧.

وأشارت النتائج أنه بالنسبة للصف العاشر الأساسي، كانت أكثر المباحث انتهاكاً لافتراض الاستقلال الموضوعي باستخدام المؤشر Q_3 مبحث اللغة العربية، حيث كانت نسبة الفقرات التي بينها انتهاك لافتراض الاستقلال الموضوعي تساوي ٠٠.٩٣٨، في حين أن المباحث الأخرى كان انتهاكها لهذا الافتراض قليلاً، حيث لم تتجاوز نسبة الفقرات التي بينها انتهاك ٠٠.٢٠. وبالرجوع إلى الوسيطات، أشارت النتائج إلى أن أكثر المباحث انتهاكاً لهذا الافتراض كان في مبحث اللغة العربية، وفي المقابل نجد أن المباحث الأخرى قد استجابت لافتراض الاستقلال الموضوعي.

يبدو واضحاً أن جميع مباحث الاختبار الوطني للصف الثامن قد انتهكت وبشكل كبير افتراض الاستقلال الموضوعي، ومبحث واحد من مباحث الصف العاشر الأساسي وهو مبحث اللغة العربية، قد انتهك هذا الافتراض، وبالنسبة إلى الأسباب التي يمكن أن تؤدي إلى الارتباط الموضوعي، فكانت متفقة مع بعض الأسباب التي ذكرتها ين (Yen, 1993) ^(١) [١٦]، منها: الأسباب المتعلقة بالتبعية لنص معين والفقرة المتسلسلة والمساعدة الخارجية، حيث كانت بعض المباحث وخاصة مبحثي اللغة العربية واللغة الإنجليزية تحتوي على فقرات تتبع نص معين، ومن المعروف أن إجابة أي فقرة من هذه الفقرات تعتمد على الفهم العام للنص المعطى، ولا يستبعد الباحث لجوء بعض الطلبة إلى أسلوب الغش، وهو يعتبر نوع من المساعدة الخارجية، وذكرت آلن وين (Allen & Yen, 1979) ^(٢) [١١] بأنه لا يقتصر الاستقلال الموضوعي على أن كون استجابات المفحوصين على الفقرات المختلفة مستقلة إحصائياً، بل من المفروض أن تكون استجابات المفحوصين الذين لديهم نفس القدرة على أي فقرة مستقلة إحصائياً أيضاً وهو ما يعرف بـ Local

(1) YEN, W., 1993-Scaling Performance Assessments: Strategies for Managing Local item Dependence. Journal of Educational Measurements, vol. 30, No. 3, pp 187-213.

(2) ALLEN, M. & YEN, W., 1979-Introduction to measurement theory. California: Brooks/Cole publishing company.

independence of examinees، فاستجابة مفحوص على فقرة معينة ينبغي أن لا تتأثر باستجابة مفحوص آخر (له نفس القدرة) على نفس الفقرة سواء بالمناقشة بين المفحوصين أو التعاون فيما بينهم، وبالتالي من المحتمل لجوء بعض الطلبة إلى الغش، وخصوصاً أن جميع فقرات هذه الاختبارات هي من نوع اختيار من متعدد، ومن المعروف سهولة الغش في هذا النوع من الاختبارات. ويمكن أن تكون عدد الفقرات المتضمنة في العنقود عاملاً مساهماً لوجود الارتباط الموضوعي، فمن الواضح أنه في بعض الأحيان يكون هناك احتمالية لأن يكون مقدار الارتباط الموضوعي للعناقيد التي تحتوي على فقرات قليلة أعلى من العناقيد التي تحتوي على فقرات كثيرة.

وهناك العديد من الطرق التي يمكن استخدامها من أجل التخفيف من مقدار الارتباط الموضوعي في فقرات الاختبار، وكذلك التقليل من الأثر الذي يمكن أن ينجم عن هذا الارتباط على الإحصائيات المتعلقة بفقرات الاختبار. وفي هذا السياق ذكر الن و سدويكس (Allen & Sudweeks, 2001) ⁽¹⁾ [٢] بأن الحل المباشر هو كتابة فقرات مستقلة إحصائياً عن بعضها البعض. ومن الطرق المتبعة لتحليل العلامات لاختبار يحتوي على فقرات مرتبطة بسياق معين هو استخدام ما يسمى *testlet scoring*. وهذا الإجراء يتطلب تصحيح الفقرات داخل كل مجموعة من الفقرات (تشكل عنقود) باعتبارها فقرة واحدة. واستجابة المفحوص لهذه الفقرات تأخذ علامة واحدة تتراوح ما بين (zero إلى n) ، حيث (n) عدد الفقرات داخل العنقود. وهذه الطريقة في التصحيح تفترض بأن يكون الاستقلال محقق بين العناقيد المختلفة وكذلك لا بد أن يكون هناك استقلال بين الفقرات الموجودة داخل العناقيد وفقرات الاختبار الأخرى. ولا تتطلب هذه الطريقة إزالة الارتباطات بين الفقرات داخل العنقود نفسه.

ويشير درشر (Dresher, 2003) ⁽²⁾ [٥] إلى أن معظم الباحثين يلجأ إلى حلول وصفها بأنها الأسوأ عندما يتم الكشف عن وجود الارتباط الموضوعي بين الفقرات، ومن هذه الحلول هو استخدام أسلوبين عند التعامل مع الفقرات التي بينها ارتباط أحدهما يتمثل في تجاهل هذا الارتباط (التصحيح على أساس أنها فقرات مستقلة)، والآخر يعتمد على حذف الفقرات التي بينها ارتباط موضوعي (أي إزالتها من الاختبار)، وهذه الأساليب تؤدي إلى تعقيدات تم الإشارة إليها سابقاً من مثل: تحيز كل من الثبات و دالة المعلومات ومعامل التمييز إلى أعلى، ويؤدي حذف بعض الفقرات إلى اختلال في مواصفات الاختبار.

(1) ALLEN, S & SUDWEEKS, R., 2001-Identifying and managing local item dependence in context-dependent item sets. (ERIC Document Reproduction Service No.ED 453236).

(2) DRESHER, A., 2003-An empirical investigation of local item dependency in NAEP data. paper presented at the annual meeting of the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME) held between April 21 to 25, 2003, In Chicago, IL. (on line). available: www.ets.org/legal/copyright.html

ويعتبر درشر (Dresher,2003) ⁽¹⁾ [٥] بأن أسلوب تشكيل العناقيد هو الأسلوب الأفضل لاحتواء الارتباط الموضوعي في الوقت الحاضر. واتفق معه في ذلك ثيسن وآخرون (Thissen et al., 1989, cited in Dresher, 2003) ⁽²⁾ [٥] حيث بينوا بأن استخدام أسلوب التصحيح متعدد الاستجابات نظرياً يعتبر مقنعاً أكثر من القيام بتطبيق نماذج استجابة الفقرة أحادية البعد على الفقرات التي تكون مرتبطة بسياق معين مثل فقرات استيعاب المقروء، في كونها تجعل افتراض الاستقلال الموضوعي محققاً في البيانات التي بين أيدينا. وبين لنا بيشوب وعمر (Bishop & Omar, 2002) ⁽³⁾ [٣] بأنه على الرغم بأن افتراض الاستقلال الموضوعي لا يمكن تحقيقه على مستوى الفقرة، إلا أنه يمكن تحقيق هذا الافتراض إذا ما تم تشكيل عناقيد للفقرات التي بينها ارتباط موضوعي، وتطبيق نماذج نظرية استجابة الفقرة متعددة الاستجابات لمجموعة العناقيد التي تم تشكيلها يمكن أن يكون هو البديل الأمثل لتجنب المشاكل التي يمكن أن تنجم عن عدم تحقق هذا الافتراض، مثل المشكلات الخاصة بتقدير الثبات ودالة معلومات الاختبار، والمعادلة الأفقية والتحليلات الخاصة بالأداء التفاضلي للفقرة.

وذكرت ين (Yen,1993) ⁽⁴⁾ [١٦] ستة أساليب يمكن أن تساعد في التخفيف من اثر الارتباط الموضوعي، هي:

- بناء فقرات مستقلة Construct Independent Items.
- تطبيق الاختبار تحت ظروف ملائمة Administer Test Under Appropriate Conditions
- تجميع العلامات للفقرات التي تحتوي على ارتباط موضوعي Combine The Grading of LID Items.
- مراجعة الاختبار لتحديد الارتباط الموضوعي للفقرات Review Tests to Identify LID Items
- بناء مقاييس مستقلة Construct Separate Scales.
- استخدام أسلوب العناقيد Use Testlets.

⁽¹⁾DRESHER,A.,2003-An empirical investigation of local item dependency in NAEP data. paper presented at the annual meeting of the American Educational Research Association(AERA)and the National Council on Measurement in Education(NCME) held between April 21 to 25,2003,In Chicago,IL. (on line).available:www.ets.org/legal/copyright.html

⁽²⁾DRESHER,A.,2003-An empirical investigation of local item dependency in NAEP data. paper presented at the annual meeting of the American Educational Research Association(AERA)and the National Council on Measurement in Education(NCME) held between April 21 to 25,2003,In Chicago,IL. (on line).available:www.ets.org/legal/copyright.html

⁽³⁾BISHOP,N,&OMAR,M.,2002- comparing vertical scales derived from dichotomous and Polytomous IRT models for a test composed of testlets. (ERIC Document Reproduction Service No.ED 466639).

⁽⁴⁾YEN, W.,1993-Scaling Performance Assessments:Strategies for Managing Local item Dependence.Journal of Educational Measurements,vol.30,No.3,pp 187-213.

- [١] ALLEN, M. & YEN, W., 1979-**Introduction to measurement theory**. California: Brooks/Cole publishing company.
- [٢] ALLEN, S. & SUDWEEKS, R., 2001-**Identifying and managing local item dependence in context-dependent item sets**. (ERIC Document Reproduction Service No. ED 453236).
- [٣] BISHOP, N. & OMAR, M., 2002- **comparing vertical scales derived from dichotomous and Polytomous IRT models for a test composed of testlets**. (ERIC Document Reproduction Service No. ED 466639).
- [٤] CHEN, W. H. & THISSEN, D., 1997-**Local dependence indexes for item pairs using item response theory**. Journal of Educational and Behavioral Statistics, vol. 22, pp: 265-289.
- [٥] DRESHER, A., 2003-**An empirical investigation of local item dependency in NAEP data**. paper presented at the annual meeting of the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME) held between April 21 to 25, 2003, In Chicago, IL. (online). available: www.ets.org/legal/copyright.html
- [٦] EMBRETSON, S. & REIASE, S., 2000-**Item response theory for psychologists**. New Jersey: Lawrence Erlbaum Associates, Inc.
- [٧] HAMBLETON, R., 1989-**Principles and selected applications of item response theory** in R. L. Linn (Ed.), Educational measurement (3rd ed. PP. 147-200) New York: American Council on Education and Macmillan.
- [٨] HAMBLETON, R., & SWAMINATHAN, H., 1985-**Item response theory: principles and applications**. Boston, MA: Kluwer - Nijhoff.

- [٩] KIM, S; COHEN, A. & LIN, Y., 2005-LDID: A Computer Program for Local Dependence Indices for Dichotomous Items. Version 1.0.
- [١٠] LEE, Y., 2004-Examining passage - related local item dependence (LID) and measurement construct using Q_3 statistics in an EFL reading comprehension test. Language testing, vol. 21, No.1, pp:74-100.
- [١١] REESE, L., 1999-A classical Test Theory perspective on LSAT local Item Dependence. Report No. LSAC-R-96-01. Law School Admission council, Newtown, PA (Document Reproduction Service No. ED 469172).
- [١٢] SHEN, L., 1997-Quantifying Item Dependency by Fisher's Z, paper presented At the annual meeting of the American Educational Research Association, Chicago, IL, March 24-28, 1997 (ERIC Document Reproduction Service No. ED 410241).
- [١٣] THOMPSON, T & POMMERICH, M., 1996-Examining the sources and effects of local item dependence, paper presented at the Annual meeting of the American Educational Research Association, New York, NY, April 8-12, 1996 (ERIC Document Reproduction Service No. ED 400311).
- [١٤] XIE, Y., 2001-Dimensionality, Dependence, or Both? An Application of the Item Bundle Model to Multidimensional Data. Policy, Organization, Measurement & Evaluation Graduate School of Education: University of California at Berkeley
- [١٥] YEN, W., 1984-Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. Applied psychological Measurement, vol. 8, pp: 125-145.
- [١٦] YEN, W., 1993-Scaling Performance Assessments: Strategies for Managing Local item Dependence. Journal of Educational Measurements, vol. 30, No. 3, pp 187-213.
- [١٧] ZENISKY, A., HAMBLETON, R., & SIRECI, S., 2002-Effects of local item Dependent on the Validity of IRT item, test and ability statistics. On line, available: www.aamc.org/students/mcat/research/monograph5.df.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100