المجلة الأردنية في العلوم التربوية، مجلد 15، عدد 1، 2019، 99- 111

أثر خصائص الفقرة ومستوى قدرة الأفراد في افتراض الاستقلال الموضعي في نظرية الاستجابة للفقرة

زايد بني عطا*

تاريخ تسلم البحث 2018/7/30 تاريخ قبوله 2018/12/27

ملخص: هدفت الدراسة الحالية إلى تقصي أثر خصائص الفقرة ومستوى قدرة الأفراد على افتراض الاستقلال الموضعي. ولتحقيق الهدف من الدراسة تم توليد استجابات (1000) فرد على (100) فقرة، وتحليلها باستخدام النموذج اللوجستي ثنائي المعلمة لإيجاد معالم الفقرات وقدرة الأفراد، حيث تم تشكيل أربعة نماذج من الاختبارات حسب مستوى الصعوبة والتمييز للفقرة، وتقسيم الأفراد إلى ثلاثة مستويات حسب القدرة. أظهرت نتائج الدراسة أن نسبة أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي كانت الأعلى في حالة النموذج الأول(فقرات صعبة وتمييز مرتفع) عند مستوى القدرة العالية والأقل عند مستوى القدرة المنخفضة، وكذلك أشارت النتائج إلى أن نسبة أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي كانت تزداد بزيادة مستوى القدرة في النموذج الأول، في حين كانت تقل في النموذج الرابع (فقرات سهلة وتمييز منخفض). وأظهرت النتائج بشكل عام ان لخصائص الفقرة ومستوى القدرة للأفراد أثراً على افتراض الاستقلال الموضعي، وأن تمييز الفقرة كان له دور أكبر في التأثير عليه من صعوبة الفقرة، وذلك من خلال اختبار دلالة الفروق الإحصائية بين نسب أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي حسب مستوى القدرة للأفراد ونموذج الاختبار.

(**الكلمات المفتاحية:** الاستقلال الموضعي، النموذج اللوجستي ثنائي المعلمة، Q_3، القدرة، خصائص الفقرة)

مقدمة: هيمنت نظرية الاستجابة للفقرة (Item Response Theory) منذ ظهورها في منتصف القرن العشرين على أدبيات القياس النفسي والتربوي؛ بسبب ما قدمته من طرق سيكومترية ذات فعالية عالية متعلقة ببناء الاختبارات والمقاييس النفسية، وطرق تحليل بياناتها مقارنة بنظرية القياس التقليدية، التي واجهت بعض القصور بدراسة الخصائص السيكومترية لفقرات المقاييس والاختبارات المتمثلة بالصعوبة والتمييز، بالإضافة إلى الخصائص السيكومترية للمقاييس والاختبارات المتعلقة بالثبات والصدق، حيث كانت تتأثر تلك الخصائص بخصائص الأفراد الذين تطبق عليهم (Van der Linden, 2010). ولعل من أهم المزايا التي ارتبطت بنظرية الاستجابة للفقرة وجعلتها تسيطر على أدبيات القياس النفسي هي استقلالية القياس، ويعني أن تقدير معالم الفقرات يكون مستقلاً عن خصائص الأفراد التي استخدمت في تقدير هذه المعالم (Sample Free)، وأن تقدير قدرة الأفراد يكون مستقلا عن عينة الفقرات التي تطبق عليهم(Item Free)، حيث اعتبرت هذه الميزة بمثابة النقطة المفصلية بين النظرية التقليدية للقياس ونظرية الاستجابة للفقرة (Hambleton, 1994).

The Effect of Item Characteristics and Persons Ability Levels on the Assumption of Local Independence in Item Response Theory

Zaid Bani Ata, Yarmouk University, Jordan.

Abstract: This study aimed at investigating the effect of item characteristics and persons, ability level on item local independence. To achieve the aim of the study, responses of (1000) persons were generated on (100) items and were analyzed using the two-parameter logistic model to find the parameters of the items and the ability of persons. Four forms of tests were formed according to the level of difficulty and discrimination of the item, and persons were divided into three levels of ability. The results of the study showed that the percentage of item pairs that violate the assumption of local independence was the highest in the case of the first form (high difficult, high discriminate) at the high ability and the lowest at low ability level. The results also indicated that the percentage of item pairs that violate the assumption of local independence increased by increasing the ability level in the first form, whereas it decreased in the fourth form (low difficult, low discriminate). In general, the results showed the item characteristics and the persons ability level influenced the assumption of local independence. The discrimination of the item had a greater role in influencing it than the difficulty of the item by testing the significance of statistical differences between the percentages of item pairs to the assumption of local independence according to the persons ability level and the test form.

(**Keywords:** Local Independence, Two-Parameter Logistic Model, Q_3, Ability, Item Characteristics)

وتقوم هذه النظرية على مجموعة من النماذج الاحتمالية القائمة على الاقتران اللوغارتيمي الذي يحدد العلاقة بين أداء الفرد على الفقرة والقدرة أو السمة التي تكمن وراء هذا الأداء، وأن العلاقة بين أداء الفرد على الفقرة وقدرته يمكن أن تحدد من خلال ما يسمى بمنحنى خاصية الفقرة (Item Characteristic Curve)، وتفترض كذلك أن مقدار الاحتمال يكون دالة متزايدة وتيريا لموقع الفرد على متصل السمة، مما يعني أن احتمال الإجابة الصحيحة يزداد بزيادة قدرة الفرد (Hambleton, 1994; De Gruijter & Van der Kamp, 2008; Henard, 2000). وتصنف نماذج نظرية الاستجابة للفقرة إلى نماذج ثنائية التدريج (Dichotomous Models) ونماذج متعددة التدريج (Polytomous Models) (De Aylaa, 2009; Hambleton & Swaminthan. (1985).

* جامعة اليرموك، الأردن.

© حقوق الطبع محفوظة لجامعة اليرموك، إربد، الأردن.

وتعد النماذج ثنائية التدريج من أشهر النماذج استخداما في بناء وتطوير المقاييس النفسية والاختبارات، حيث تختلف هذه النماذج في عدد معالم الفقرة التي تقدرها، ويوجد منها ثلاثة نماذج، أولها: النموذج اللوغارتيمي ذو المعلمة الواحدة (One-Parameter Logistic Model) أو المشهور باسم نموذج راش (Rasch Model)، وهو من أكثر النماذج الثنائية استخداماً، حيث لا يتطلب عمليات حسابية معقدة، على الرغم من تشدد افتراضاته، وهي أن الفقرات متساوية في تمييزها، ولا يوجد تخمين، وأن الفقرات تتباين في الصعوبة فقط (Harris, 1989). أما النموذج الثاني فهو النموذج اللوغارتيمي ثنائي المعلمة (Two- Parameter Logistic Model)، وهو يختلف عن النموذج الأحادي بأنه يتيح تقدير معلمة التمييز للفقرة إضافة إلى معلمة الصعوبة، ويعد النموذج ثنائي المعلمة أقل عمومية من النموذج الأخير والثالث من بين هذه النماذج، وهو النموذج اللوجستي ثلاثي المعلمة (Three –Parameter Logistic Model) الذي يقدر معلمة الصعوبة، والتمييز والتخمين (De Ayala, 2009; Embretson & Reise, 2000).

وحتى يؤدي استخدام نظرية الاستجابة للفقرة في تحليل البيانات الناتجة من تطبيق الاختبارات والمقاييس النفسية للوصول إلى نتائج يمكن الوثوق بها، لا بد من التحقق من الافتراضات التي استندت إليها هذه النظرية، وهي: افتراض أحادية البعد (Unidimensionality)، ويعني أن هناك قدرة واحدة أو سمة واحدة تفسر أداء الفرد على الاختبار أو المقياس، ويرى ورم (Warm, 1978) أن افتراض أحادية البعد أكثر الافتراضات تعقيداً في نظرية الاستجابة للفقرة؛ بسبب تدخل عوامل أخرى لها علاقة بظروف تطبيق الاختبار، والتي تؤثر في الأداء على الاختبار مثل: الدافعية، والقلق، والعوامل الشخصية، ولذلك فإن مثل هذا الافتراض لا يتحقق بشكل مؤكد دائما، وعلى الرغم من وجود هذه العوامل فقد رأى هامبلتون وسوامينيثان وروجرز (Hambleton, Swaminathan, & Rogers, 1991) بأنه يتطلب وجود عامل أساسي في الأداء، وهو ما يشار إليه بالقدرة التي يقيسها الاختبار.

أما الافتراض الثاني فيتمثل بالاستقلال الموضعي (Local Independence)، ويقصد به كما عرفه لورد ونوفيك (Lord & Novick, 1968) بأنه لأي مستوى من مستويات القدرة للأفراد تكون التوزيعات الشرطية لدرجات الاختبار جميعها مستقلة عن بعضها بعضا. في حين أشار هامبلتون واسوامنيثان (Hambleton & Swaminthan, 1985) إلى أن الاستقلال الموضعي يعني أن استجابات الأفراد على فقرات الاختبار من نفس القدرة تكون مستقلة إحصائيا، ويعني هذا أن استجابة الفرد عن فقرة ما يجب أن لا تؤثر سلبا أو إيجابا في استجابته عن أي فقرة أخرى.

ويعد افتراض الاستقلال الموضعي غاية في الأهمية عند استخدام نماذج نظرية الاستجابة للفقرة (Item Response Theory) في تحليل بيانات الاختبارات النفسية والتربوية، وغياب التحقق من هذا الافتراض ينعكس سلباً على دقة تقديرات معالم الفقرة وقدرة الأفراد والذي بدوره ينعكس على أخطاء القياس ودالة معلومات الاختبار (Emberston & Rise, 2000; Killer, Swamiathan & Sireci, 2003; Reese, 1999; Thomposon & Pommerrich, 1996; Zenisty, Hambleton & Sireci, 2003).

ومما تجدر الإشارة إليه، أن الاستقلال الموضعي لا يتعارض مع الارتباطات الموجودة بين الفقرات التي شكل منها الاختبار أو المقياس عندما يتم تطبيقه على أفراد العينة كاملاً، وهذا في الحقيقة يعد مطلباً أساسيّاً للحصول على اختبار يتمتع بدقة عالية في قياس السمة، مما يعني الحاجة لفقرات متسقة فيما بينها في قياس السمة، أو بمعنى آخر الارتباط بينها عال. ولذا، فإن مفهوم الاستقلال الموضعي يعني استقلالية الاستجابة عن الفقرة عندما يتم ضبط مستوى القدرة للأفراد، ويكافئ ذلك بعدم وجود علاقة ارتباطيه بين أي زوج من الفقرات لنفس المستوى من القدرة للأفراد (Emberston & Rise, 2000).

وبين البحث السيكومتري الذي اهتم بالتحقق من افتراض الاستقلال الموضعي أن الأسباب المسئولة عن انتهاك هذا الافتراض والتي تظهر الاعتمادية بين الفقرات (Local Dependence) تأتي من مصادر مختلفة، من أهمها تعددية أبعاد الاختبار(Chen & Thisson, 1997; Tuerlinckx & De Boeck, 2001). وأشارت ين (Yen, 1993) أن عوامل أخرى قد تكون خارجية تظهر الاعتمادية بين الفقرات، أهمها: الغش، والسرعة، وصياغة الفقرات، والتبعية لنص معين، بحيث إذا تم معرفة الإجابة على فقرة ما في الاختبار يؤدي ذلك لزيادة احتمالية الاستجابة الصحيحة على الفقرة التالية.

وللكشف عن الاستقلال الموضعي للفقرات فقد تم إيجاد العديد من المؤشرات الإحصائية، من أشهر هذه المؤشرات مؤشر Q_3 الذي اقترحته ين (Yen,1984) كمؤشر إحصائي للكشف عن الاستقلال الموضعي بين فقرات الاختبار، وهو عبارة عن معامل الارتباط بين البواقي لزوج من الفقرات بعد ضبط قدرة الفرد (θ). ويكمن الأساس الرياضي في حساب هذا المؤشر في تقدير قدرة الأفراد على فقرات الاختبار من أجل استخدام هذه القدرة في تقدير احتمالية الاستجابة الصحيحة عن فقرات الاختبار لكل قدرة من قدرات الأفراد، وبعد ذلك يتم إيجاد الباقي ويرمز له بالرمز (d_{ik}) من خلال الفرق بين الأداء الملاحظ للفرد (k) والأداء المتوقع له على الفقرة (i) كما في المعادلة رقم 1

$$d_{ik} = u_{ik} - \hat{p}_i(\theta_k) \quad \ldots\ldots\ldots\ldots\ldots 1$$

حيث يشير الرمز(u_{ik}) الى علامة الفرد (k) الملاحظة على الفقرة (i)، وتكون في هذه الحالة إما صفراً، إذا كانت الاستجابة خاطئة، وواحداً إذا كانت الاستجابة صحيحة، أما الرمز $\hat{p}_i(\theta_k)$ فيشير الى احتمال الاجابة الصحيحة عن الفقرة المقدر من النموذج المستخدم. وبعد أن يتم إيجاد البواقي لفقرات الاختبار يحسب

معامل الارتباط بين كل زوج من الفقرات (i,j) للبواقي وفق المعادلة رقم 2، ويكون هذا المعامل هو Q_3 (Yen, 1993).

$$Q_{3ij} = r(d_i, d_j) \quad \cdots\cdots\cdots\cdots \quad 2$$

وعندما يتحقق افتراض الاستقلال الموضعي فإن القيمة المتوقعة لمؤشر Q_3 تساوي (-1/n-1)، حيث يشير الرمز n إلى عدد فقرات الاختبار ((Yen, 1993). وعلى الرغم من شيوع استخدام هذا المؤشر في الكشف عن أزواج الفقرات المستقلة من غير المستقلة (المعتمدة)، إلا أن المتصفح للأدب النظري يجد تباينا في تحديد القيم الحرجة لهذا ألمؤشر، فقد أشار كريستنس وماكرونس وهورتن (Christensen, Makrons, & Horton, 2017) أنه لا يوجد معيار محدد يمكن من خلاله تحديد درجة الاعتمادية لكل المواقف، بل يجب النظر إلى القيمة المتوقعة للمؤشر Q_3، وهذا ما أكده شن (Shen, 1997) بأن أحد المشاكل الرئيسة لهذا المؤشر هو عدم وجود معيار محدد يمكن من خلاله الاستناد إليه في تحديد درجة انتهاك الاستقلال الموضعي، حيث يبقى تحديد درجة الاعتمادية اعتباطية. وقد اعتبر شن وثيسن (Chen & Thissein, 1997) أن أي قيمة مطلقة لمؤشر Q_3 أعلى من (0.2) تعد بمثابة مؤشر على وجود الاعتمادية بين أي زوج من الفقرات، إلا أن كيم وكوهين ولين (Kim, Cohen, & Lin, 2005) قد استخدموا القيمة (0.05) كقيمة حرجة لمؤشر Q_3 في برنامجهم (LDID) لحساب مؤشرات الكشف عن الاستقلال الموضعي، والتي تعد بمثابة تباين البواقي المشترك في زوج من الفقرات، وهذه القيمة استخدمت في الدراسة الحالية كحد فاصل بين أزواج الفقرات التي بينها استقلال موضعي.

وعلى الرغم من الاستخدام المتزايد لنماذج نظرية الاستجابة للفقرة وتوظيفها في التطبيقات العملية المتعلقة بالاختبارات والمقاييس النفسية والتربوية، نجد أن البحث السيكومتري ما زال قائماً حول تلك النماذج وفي معرفة مدى تأثير انتهاك افتراضات النموذج المستخدم على دقة تقديرات معالم النموذج، ومن هذه الافتراضات التي نالت اهتماماً في البحث السيكومتري افتراض الاستقلال الموضعي.

وبالنظر إلى الدراسات والأبحاث التي اهتمت بتناول افتراض الاستقلال الموضعي للفقرات، نجد أن غالبيتها قد جاءت في اتجاهين؛ الاتجاه الأول ركز على دراسة أثر انتهاك هذا الافتراض على دقة التقديرات لمعالم النماذج المستخدمة، أما الاتجاه الثاني فقد ركز على المقارنة بين مؤشرات الكشف عن الاستقلال الموضعي. فقد هدفت دراسة ثومبسون وبومرج (Thompson & Pommerich, 1996) الى دراسة أثر انتهاك افتراض الاستقلال الموضعي على ثبات الاختبار ودالة المعلومات من خلال استخدام مؤشر G^2، ومؤشر البواقي المستخدم في برنامج NOHRAM في الكشف عن الاعتماد الموضعي بين الفقرات. أظهرت نتائج الدراسة بان هناك تأثيراً واضحاً لانتهاك افتراض الاستقلال الموضعي على كل من ثبات الاختبار، ودالة معلومات الاختبار، حيث جعل تقديراتهما متحيزة للأعلى، وكذلك فعالية المؤشرين في الكشف عن الاعتماد الموضعي بين الفقرات، واعتبرت القيمة (10) بمثابة الحد الفاصل لمؤشر الاستقلال الموضعي G^2 لأي زوج من الفقرات، وهي القيمة الحرجة التي استخدمت في برنامج LDID.

وفي دراسة شين (Shen, 1997) التي استخدم فيها (770) فقرة من نوع الاختيار من متعدد من الاختبارات الوطنية للرخصة الطبية (National Medical Licensing Examination) للمقارنة بين مؤشرين للكشف عن الفقرات التي بينها اعتماد موضعي (غير مستقلة)، وهما: مؤشر Q_3 ومؤشر Fisher Z. توصلت الدراسة الى أن المؤشر Fisher Z كان يمتلك خصائص إحصائية أفضل من مؤشر Q_3، لكونه يعتمد على البواقي المعيارية، وفي الوقت نفسه كان حساساً أكثر من مؤشر Q_3 للاعتماد الموضعي.

وبينت نتائج دراسة بيل وباتسون وذرس (Bell, Pathison, & Withers, 1988) التي هدفت إلى الكشف عن الارتباط الموضعي بين فقرات الاختبار الدراسي الأسترالي (The Australian Scholastic Test) أن الارتباط الموضعي كان داخل العنقود ملاحظاً أكثر مما هو بين العناقيد، وأن هذا الارتباط للفقرات المرتكزة على الرياضيات أكثر من الفقرات المرتكزة على اللغة، وكان يزداد بزيادة قدرة الأفراد،

وأجرى ريس (Reese, 1999) دراسة هدفت إلى معرفة تأثير انتهاك الاستقلال الموضعي على خصائص فقرات الاختبار وثباته. ولتحقيق الغرض من الدراسة قام بتطبيق اختبار(LSAT)(Law School Admission Test) على عينة الدراسة حيث استخدم مؤشر Q_3 للكشف عن الاستقلال الموضعي للفقرات. أشارت نتائج الدراسة أنه عندما كان هناك ارتباط موضعي عالٍ بين الفقرات كانت معاملات التمييز للفقرات وثبات الاختبار متحيزة للأعلى، وفي الحالات التي كان الارتباط الموضعي قليلاً بين الفقرات، لم يكن هناك أثر على قيم معاملات التمييز وثبات الاختبار، وأشارت النتائج كذلك أن الاعتماد الموضعي بين الفقرات كان له أثر واضح في تقدير معلمة القدرة للأفراد، حيث أدى ذلك إلى تحيز سلبي لذوي القدرة المنخفضة، وتحيز إيجابي لذوي القدرة المرتفعة.

وفي دراسة زينسكي وهامبلتون وسيريس (Zenisky, Hambleton, & Sireci, 2003) التي هدفت إلى الكشف عن مدى تأثير انتهاك افتراض الاستقلال الموضعي على تقدير معلمة القدرة للأفراد وثبات الاختبار. ومن أجل تحقيق هدف الدراسة تم تطبيق اختبار القبول في كلية الطب، واستخدموا مؤشرين للكشف عن أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي، مؤشر Q_3، ومؤشر مقارنة تقديرين للثبات. أظهرت نتائج الدراسة أنه عندما ينتهك افتراض الاستقلال الموضعي كان لذلك أثر واضح في تقدير معلمة القدرة للأفراد، وكذلك كان له أثر في تقدير ثبات الاختبار، حيث كانت قيمة الثبات للاختبار متحيزة بشكل إيجابي. وبينت نتائج الدراسة كذلك فعالية مؤشرQ_3 في الكشف عن أزواج الفقرات التي بينها اعتماد موضعي.

وقام شن ووانج (Chen & Wang, 2007) بدراسة هدفت إلى معرفة التأثيرات المحتملة لتجاهل التفاعلات بين فقرات الاختبار، من خلال استخدام بيانات مولدة وفق النموذج اللوجستي ثلاثي المعلمة، واستخدام مؤشر Q_3 للكشف عن الفقرات التي بينها اعتماد موضعي. أشارت نتائج الدراسة أن تجاهل التفاعلات الموجبة بين الفقرات أدى إلى جعل معلمة التمييز للفقرة تتحيز إلى الأعلى، ومعلمة الصعوبة للفقرة تتحيز إلى الأسفل، بينما تجاهل التفاعلات السالبة بين الفقرات جعل معلمة التمييز تتحيز إلى الأسفل، ومعلمة الصعوبة تتحيز للأعلى. وأشارت النتائج كذلك بأنه عندما تم تجاهل التفاعلات الموجبة بين الفقرات جعلت قيم مؤشر Q_3 سالبة، والعكس عندما تم تجاهل التفاعلات السالبة.

وهدفت دراسة أبوشندي (2011) إلى تفحص تأثير قوة العلاقة الارتباطية بين الأبعاد (0.0 ,0.5, 0.9)، ومستوى قدرة المفحوصين (منخفضة، متوسطة، عالية) على قدرة أربعة من مؤشرات للاستقلال الموضعي، وهي: G^2, Z, Q_3,Fisher Z. أظهرت نتائج الدراسة عدم تأثر المؤشر G^2بتغير قوة الارتباط، أما المؤشران Q_3 ,Z لم يتأثرا فقط عند قوة الارتباط (0.0)، وأظهرت النتائج كذلك أن نسبة الأزواج للفقرات المنتهكة للاستقلال الموضعي لم تتغير عند استخدام المؤشر G^2، وكانت تقل بنقصان مستوى القدرة عند استخدام المؤشرين Q_3, Fisher Z وقوة العلاقة بين الابعاد (0.9).

وأجرى حمادنه (2011) دراسة هدفت إلى الكشف عن الارتباط الموضعي بين أزواج الفقرات باستخدام مؤشرQ_3 ، ومعرفة نسبة أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي باختلاف مستوى القدرة للأفراد، وكذلك العلاقة بين افتراض أحادية البعد والاستقلال الموضعي. ولتحقيق الهدف من الدراسة استخدم الباحث بيانات اختبار مستوى الحاسوب الذي تعده جامعة آل البيت. أظهرت نتائج الدراسة أن نسبة عدد أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي قد بلغت (0.135)، وكذلك أظهرت النتائج أن هذه النسبة تنخفض بانخفاض مستوى القدرة للأفراد، وبينت كذلك النتائج أن افتراض الاستقلال الموضعي وافتراض أحادية البعد لنظرية الاستجابة للفقرة متكافئان.

يتضح بشكل عام من الدراسات والأبحاث بأن هناك اهتماما ملحوظا بدراسة الآثار المترتبة على انتهاك افتراض الاستقلال الموضعي على التقديرات المختلفة لمعالم نماذج نظرية الاستجابة للفقرة، خصوصاً النماذج الأحادية، وفي الوقت نفسه البحث عن فعالية الأساليب والإجراءات المستخدمة في الكشف عن الاستقلال الموضعي، حيث اتفقت غالبية الدراسات على أن انتهاك افتراض الاستقلال الموضعي كان له أثر واضح في تقديرات القدرة، وعلى معاملات الثبات للاختبارات، وكذلك شيوع استخدام مؤشر Q_3 في الكشف عن الفقرات التي بينها اعتماد موضعي. ولأن افتراض الاستقلال الموضعي يعد أحد الافتراضات الرئيسة في نظرية الاستجابة للفقرة، ولما له من آثار سلبية على التقديرات المختلفة لنظرية الاستجابة الفقرة، جاءت هذه الدراسة بالاتجاه المعاكس للدراسات السابقة من حيث تناولها دراسة أثر خصائص الفقرات ومستوى القدرة على افتراض الاستقلال الموضعي لندرة الدراسات –في حدود اطلاع الباحث – التي تطرقت لمثل ذلك.

مشكلة الدراسة وأسئلتها

تعد الفقرة الاختبارية حجر الزاوية الأساسية في بناء وتطوير الاختبارات والمقاييس النفسية والتربوية، حيث يهتم مطورو الاختبارات والمقاييس بانتقاء الفقرات التي تسهم إسهاماً عالياً في قياس السمة المستهدفة من الاختبار، وكما هو معروف في الأدب السيكومتري (علام، 2002؛ *Anstasi & Urbina, 2005; Allen & Yen, 1979; Crocker & Algina, 1986*) إن من أهم المؤشرات التي يعتمد عليها مطورو الاختبارات في انتقاء الفقرات هو معامل الصعوبة والتمييز للفقرة، وفي سياق استخدام نظرية الاستجابة للفقرة فإن الخصائص السيكومترية للفقرة هي التي تحدد العلاقة بين السمة الكامنة التي يهدف الاختبار إلى قياسها وقدرة الأفراد في ذلك الاختبار (*Hambelton, 1994*)، ولأن هناك اطراداً متزايداً في استخدام نماذج نظرية الاستجابة للفقرة في التطبيقات العملية المختلفة التي تستند على افتراضات يجب التحقق منها، ومن هذه الافتراضات افتراض الاستقلال الموضعي، حيث أشارت الدراسات (*Chen & Wang, 2007; Shen, 1997; Zenisky, Hambleton, & Sireci, 2003*) أن انتهاكه يؤثر على التقديرات المحتملة لنظرية الاستجابة للفقرة من مثل خصائص الفقرة، وقدرة الأفراد، ودالة المعلومات.

وعلى الرغم من تزايد الاهتمام البحثي في دراسة الآثار المترتبة على انتهاك افتراض الاستقلال الموضعي في التقديرات المحتملة لنظرية الاستجابة للفقرة أو الأسباب المؤدية إلى انتهاكه، إلا أنه من الملاحظ أثناء استخدام الاختبارات في التطبيقات العملية قد تظهر بعض المواقف الاختبارية التي تتطلب تطبيق عينات جزئية من الفقرات على عينات من الأفراد من مثل: القياس التكيفي، بنوك الأسئلة، والاختبارات الوطنية، مما يجعل مثل هذه العينات من الفقرات قد تتباين في خصائصها السيكومترية، وكذلك عينات الأفراد التي قد تختلف أيضاً في مستوى القدرة، وهذا بدوره يثير استفسارا حول مدى تأثير مثل هذا التباين في خصائص الفقرات ومستوى قدرة الأفراد على افتراض الاستقلال الموضعي، مما جعل الحاجة ماسة لإجراء هذه الدراسة خصوصاً ندرة الدراسات – في حدود علم الباحث – التي جاءت في هذا الاتجاه. وبالتحديد حاولت هذه الدراسة الإجابة عن السؤالين الآتيين:

1. ما نسبة الفقرات التي بينها ارتباط موضعي باستخدام مؤشر Q_3 حسب مستوى القدرة للفرد ونموذج الاختبار(خصائص الفقرات)؟ وهل تزيد هذه النسبة بزيادة مستوى القدرة للأفراد؟

2. هل تختلف نسبة الفقرات التي بينها ارتباط موضعي باختلاف مستوى القدرة للأفراد ونموذج الاختبار (خصائص الفقرات)؟

أهمية الدراسة

تكمن أهمية الدراسة في محاولتها الكشف عن أثر خصائص فقرات الاختبار ومستوى قدرة الأفراد على الاستقلال الموضعي كأحد الافتراضات الرئيسة لنماذج نظرية الاستجابة، لما لهذا الافتراض من انعكاسات سلبية على دقة تقدير معالم الفقرات وقدرة الأفراد، إضافة إلى ذلك ندرة الدراسات بشكل عام – في حدود اطلاع الباحث- التي اهتمت بمثل هذه المشكلة. نقطة أخرى جديرة بالذكر بأنه إذا جاءت النتائج لتؤيد زيادة نسبة الفقرات المنتهكة للاستقلال الموضعي بزيادة قدرة الأفراد، مثل ذلك يفتح الآفاق أمام مستخدمي نماذج نظرية الاستجابة للفقرة أحادية البعد بإعادة تقييم التقديرات الناتجة من النموذج، خصوصاً في سياق استخدام الاختبارات أقصى أداء (*Maximal Performance Testing*)؛ لأن مثل هذا النوع يتضمن بالتأكيد قدرات عالية ومتوسطة ومنخفضة وفي الوقت نفسه يحتوي على فقرات صعبة وفقرات عالية التمييز، والذي بدوره قد يتطلب إعادة تدريج الفقرات عند كل مستوى من مستويات القدرة.

مصطلحات الدراسة

خصائص الفقرات: ويقصد بها في هذه الدراسة معلمة الصعوبة ومعلمة التمييز المنبثقة عن النموذج اللوجستي ثنائي المعلمة.

الاستقلال الموضعي: مصطلح يشير إلى أن إجابة الفرد عن فقرة من فقرات الاختبار لا تعتمد على إجابة فقرة ثانية من الاختبار، ويتم الكشف عنه باستخدام مؤشر Q_3 في هذه الدراسة.

تصميم الدراسة

لتحقيق أهداف الدراسة الحالية في الكشف عن أثر خصائص فقرات الاختبار ومستوى القدرة للأفراد على افتراض الاستقلال الموضعي، فقد تم استخدام التصميم التجريبي العاملي لبعدين، العامل الأول، وله ثلاث مستويات، وهو مستوى القدرة للفرد (قدرة عالية، قدرة متوسطة، قدرة منخفضة). أما العامل الثاني فهو مستوى خصائص فقرات الاختبار المشكل، وله أربعة مستويات (النموذج الأول، النموذج الثاني، النموذج الثالث، النموذج الرابع) بالاعتماد على أسلوب المحاكاة.

توليد البيانات

تم توليد البيانات باستخدام برنامج *WINGEN*، وهو من تصميم وإنتاج هان وهامبيلتون (*Han & Hambleton, 2007*). وتمت عملية توليد البيانات وفق المراحل الآتية:

المرحلة الأولى: تم توليد اختبار مكون من (100) فقرة باستخدام النموذج اللوجستي ثنائي المعلمة وفق الشروط التالية:

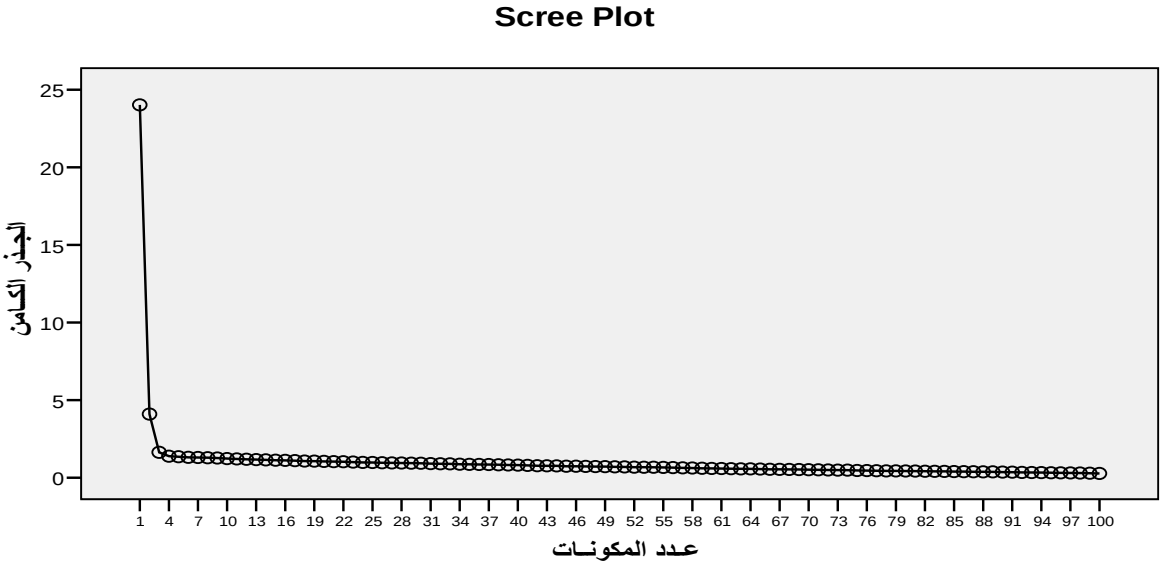
- توليد معلمة التمييز للفقرات وفق التوزيع المنتظم *uniform (2 , 0) ~*، حيث يعطي هذا التوزيع قيما متباينة لمعلمة التمييز تماثل القيم الحقيقية التي استخدمها هامبلتون وسواميثان (*Hambleton & Swaminthan, 1985*) حيث أكدا على أن تكون قيم معلمة التمييز تتراوح ما بين 0 و 2 لوجيت.
- توليد معلمة الصعوبة للفقرات وفق التوزيع الطبيعي *Normal~ (0 , 1)*، وهذا ينتج فقرات متباينة في الصعوبة تتراوح ما بين (-3 و3) كما أشار (*Baker, 2001*).

المرحلة الثانية: توليد استجابات (*1000*) فرد وفق التوزيع الطبيعي *Normal ~ (0 , 1)* باستخدام المعالم الحقيقية للفقرات التي تم توليدها في المرحلة الأولى.

تحليل البيانات

من أجل تحقيق الهدف من الدراسة والإجابة عن سؤالي الدراسة، قام الباحث بتحليل البيانات التي تم توليدها في المرحلة الثانية وفق الخطوات التالية:

أولا: تم التحقق من افتراض أحادية البعد لاستجابات الأفراد المولدة على فقرات الاختبار ككل المكون من (*100*) فقرة باستخدام المكونات الرئيسة (*Principle Component Analysis*) كافتراض يستند إليه النموذج اللوجستي ثنائي المعلمة، فتبين من نتائج التحليل بأن قيمة الجذر الكامن للعامل الأول قد بلغت (*24.023*)، في حين كانت قيمة الجذر الكامن للعامل الثاني (*4.101*)، مما يعني أن حاصل قسمة الجذر الكامن للعامل الأول على الجذر الكامن للعامل الثاني أكبر من (2)، والذي يعد مؤشراً على تحقق أحادية البعد (*Hatti, 1985*)، ويتضح هذا كذلك من خلال التمثيل البياني للجذور الكامنة *Scree Plot* كما في الشكل (1).



الشكل (1): الرسم البياني للجذور الكامنة للبيانات المولدة

وكمؤشر آخر على تحقق افتراض أحادية البعد للبيانات، تم استخدام برنامج *NOHARM* وإجراء تحليل البواقي *(Residual Analysis)*، من أجل فحص المواءمة بين عدد الأبعاد للاختبار والبيانات من خلال حساب مصفوفة البواقي *(Residual Matrix)* لإيجاد قيمة الجذر التربيعي لمتوسطات مربعات البواقي *Root Mean Square of Residuals: RMSR)*، حيث بلغت قيمته *(0.005)* وإيجاد قيمة مؤشر تناكا *(Tanaka Index)* وكانت قيمته *(0.991)*، وقد بين ميلر *(Miller, 1991)* أنه إذا كانت قيمة مؤشر *RMSR* قريبة من الصفر، وقيمة مؤشر تناكا *(Tanaka Index)* قريبة من الواحد، فيعد مثل هذا دليلاً على تحقق أحادية البعد. وبالنظر إلى قيمة *RMSR* نجد بأنها قريبة من الصفر، وكذلك اقتراب قيمة مؤشر تناكا من الواحد، مما يعني تحقق افتراض أحادية البعد في الاختبار.

ثانيا: تم التحقق من مطابقة البيانات للنموذج اللوجستي ثنائي المعلمة، حيث تم استخدام برنامج بايلوج *(Bilog-Mg 3)* لتقدير معالم الفقرات (الصعوبة والتمييز) وقدرة الأفراد **(θ)، وقد تبين من** نتائج التحليل بالاعتماد على المؤشر الإحصائي كاي تربيع ($\chi 2$) مطابقة جميع فقرات الاختبار للنموذج، وكذلك مطابقة جميع استجابات الأفراد للنموذج، وتم الاحتفاظ بمعالم الفقرات في ملف خاص وكذلك قدرات الأفراد.

ثالثا: لغاية التحقق من افتراض الاستقلال الموضعي لفقرات الاختبار المكون من *(100)* فقرة، تم استخدام برنامج *(LDID)* (*A Computer program for Local Dependence Indices for Dichotomous Items*)، حيث تم تخزين ملف قدرة الأفراد باسم *(ldid sco)*، وملف معالم الفقرات باسم *(ldid par)*، ويعود السبب وراء هذه التسمية هو من أجل عرض هذين الملفين على برنامج *LDID*؛ لأنه لا يتعرف على ملف القدرة وملف معالم الفقرات إلا إذا كانت التسمية بهذا الشكل. وقد تم حساب مؤشر Q_3 للكشف عن أزواج الفقرات التي بينها ارتباط موضعي، حيث أظهرت قيم هذا المؤشر بتحقق افتراض الاستقلال الموضعي بين الفقرات، وكانت قيمه ضمن المدى المعتمد في البرنامج، وهي القيمة المطلقة للعدد *(0.05)* كحد فاصل بين أزواج الفقرات التي بينها انتهاك لافتراض الاستقلال الموضعي من غيرها *(Kim, Cohen, & Lin, 2005)*.

رابعا: اعتبرت فقرات الاختبار الذي تم توليده الحالة المرجعية لكونها محققة لافتراض الاستقلال الموضعي استنادا إلى النقطة الثالثة. وقد تم تشكيل أربعة مستويات من خصائص الفقرات اعتماد على معلمة **التمييز** والصعوبة للفقرة، عبر عنها بنماذج الاختبار بهذه الدراسة، وقد صنفت الفقرة على أن تمييزها مرتفع إذا كانت قيمة معلمة التمييز لها أعلى من *(1.35)* ومنخفض إذا كان تمييزها اقل من ذلك، وبالمثل تم تصنيف الفقرة على أنها صعبة إذا كانت قيمة معلمة الصعوبة لها أكبر من الصفر، وسهلة إذا كانت أقل من ذلك *(Baker, 2001)*، وفي ضوء ذلك تم تشكيل أربعة مستويات من خصائص الفقرات عبر كل مستوى بنموذج مكون من *(20)* فقرة كما يلي:

- النموذج الأول: (فقرات صعبة وتمييز مرتفع).
- النموذج الثاني: (فقرات صعبة وتمييز منخفض).
- النموذج الثالث: (فقرات سهلة وتمييز مرتفع).
- النموذج الرابع: (فقرات سهلة وتمييز منخفض).

خامسا: تم تقسيم الأفراد والبالغ عددهم *(1000)* حسب مستوى القدرة إلى ثلاثة مستويات، قدرات عالية ومتوسطة ومنخفضة بالاعتماد على نتائج التحليل الواردة في النقطة الثانية، حيث تم اختيار أعلى *(25%)* من الأفراد ليشكلوا المستوى الأول

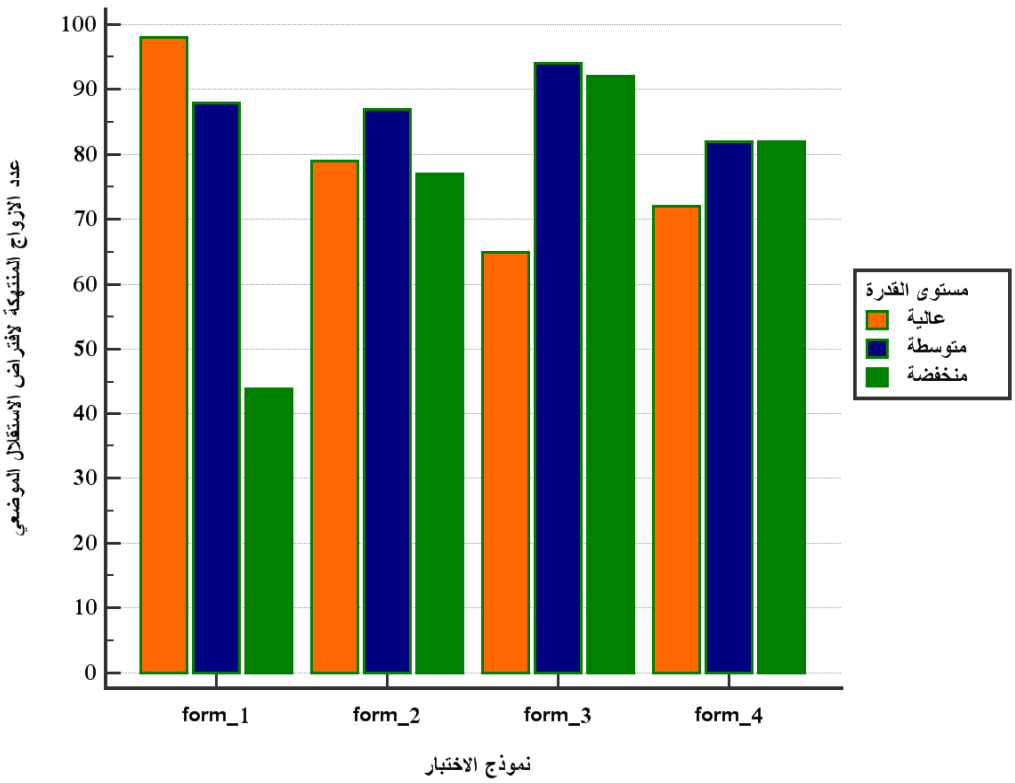
من القدرة (عالية)، وأدنى (25%) ليشكلوا المستوى الثالث من القدرة (منخفضة)، والمتبقي (50%) شكلوا المستوى الثاني من القدرة (متوسطة) وكان عددهم (500) فردٍ، تم اختيار (250) فردا منهم بالطريقة العشوائية، وذلك حتى يصبح عدد أفراد كل مستوى من مستويات القدرة متساويا بواقع (250) فردا لكل مستوى.

سادسا: تم الاعتماد على برمجيات إحصائية متنوعة لإجراء العمليات الحسابية والتحليلات الإحصائية الخاصة بالإجابة عن أسئلة الدراسة، وتلخصت الأساليب الإحصائية بما يلي:

- تقدير معالم الفقرات وقدرة الأفراد وفق النموذج اللوجستي ثنائي المعلمة باستخدام برنامج بايلوج *Bilog-Mg 3*.
- إيجاد قيم مؤشر الاستقلال الموضعي Q_3 باستخدام برمجية *LDID*.
- استخدام برمجية *SPSS* في تشكيل نماذج الاختبار حسب خصائص الفقرات ومستويات القدرة للأفراد، بالإضافة إلى إيجاد مؤشرات التحقق من أحادية البعد لاستجابات الأفراد المولدة على الاختبار ككل.
- استخدام برمجية *NOHARM* لإيجاد قيمة (*RMSR*) ومؤشر تناكا (*Tanaka Index*) كمؤشرات للتحقق من أحادية البعد.
- إيجاد نسبة أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي اعتمادا على قيم المؤشر الإحصائي Q_3 باستخدام برمجية *SPSS*.
- مقارنة الفروق بين نسب أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي باستخدام الإحصائي وولد زد (*Wald Z*) باستخدام برمجية (*NCSS 11 Statistical Software, 2016*).

نتائج الدراسة ومناقشتها

للإجابة عن السؤال الأول للدراسة، تم إيجاد قيم المؤشر الإحصائي Q_3 لكل مستوى من مستويات القدرة للأفراد، ولكل نموذج من نماذج الاختبار حسب خصائص الفقرات باستخدام برنامج *LDID*، حيث تم إيجاد عدد الأزواج المنتهكة لافتراض الاستقلال الموضعي، والشكل (2) يوضح ذلك.



الشكل (2): عدد الأزواج المنتهكة لافتراض الاستقلال الموضعي تبعا لمستوى القدرة وخصائص الفقرات (النموذج)

يتضح من الشكل (2) التباين في عدد أزواج الفقرات التي بينها اعتماد موضعي وذلك وفقا لمتغيري الدراسة: مستوى قدرة الأفراد وخصائص فقرات الاختبار(النموذج)، حيث كان الأثر الذي تركه كل من مستوى القدرة وخصائص فقرات النموذج على انتهاك افتراض الاستقلال الموضعي للفقرات واضحا من خلال الشكل. وبناء على عدد الأزواج المنتهكة لافتراض الاستقلال الموضعي تم إيجاد نسبة أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي وذلك من خلال قسمة عدد الأزواج المنتهكة للاستقلال الموضعي

على عدد الأزواج الكلية وهو (190) زوجاً*، والجدول (1) يوضح نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي حسب مستوى القدرة وخصائص الفقرات (النموذج).

جدول (1): نسب أزواج الفقرات المنتهكة للاستقلال الموضعي حسب مستوى القدرة للأفراد وخصائص الفقرات (النموذج)

مستوى القدرة	خصائص الفقرات (النموذج)			
	الأول (صعبة-تمييز عالي)	الثاني (صعبة-تمييز منخفض)	الثالث (سهلة -تمييز عالي)	الرابع (سهلة-تمييز منخفض)
عالية	0.5158	0.4158	0.3421	0.3789
متوسطة	0.4632	0.4579	0.4947	0.4316
منخفضة	0.2316	0.4053	0.4842	0.4316

تشير النتائج الواردة في الجدول (1) بأن هناك تبايناً ملحوظاً في نسبة أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي حسب قدرة المفحوصين وخصائص فقرات الاختبار، حيث يتضح بأن أعلى نسبة للفقرات المنتهكة لافتراض الاستقلال الموضعي كانت قيمتها تساوي (*0.5158*) عندما كان مستوى القدرة للأفراد عالياً، وخصائص فقرات النموذج صعبة وعالية التمييز، في حين كانت الأقل لنفس خصائص الفقرات (النموذج الأول) عندما كان مستوى القدرة للأفراد منخفضاً، حيث بلغت قيمتها (*0.2316*).

وبينت النتائج كذلك بأن نسبة الفقرات المنتهكة لافتراض الاستقلال الموضعي كانت تزداد بزيادة مستوى القدرة للأفراد عندما كانت خصائص الفقرات صعبة وذات تمييز عالٍ (النموذج الأول)، في حين أخذت هذه النسبة بالتناقص عندما كانت خصائص الفقرات سهلة وذات تمييز منخفض (النموذج الرابع). وكذلك أشارت النتائج المتعلقة بالنموذجين (الثاني والثالث) بأنه لم يكن هناك اتجاه واضح لسلوك نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي حسب مستوى القدرة للأفراد، حيث كانت هذه النسبة تتباين من مستوى قدرة إلى آخر عندما كانت خصائص الفقرات صعبة ومنخفضة التمييز (النموذج الثاني)، وكذلك عندما كانت خصائصها سهلة ومرتفعة التمييز (النموذج الثالث).

وبالنظر إلى النتائج التي أظهرها الشكل (2) والجدول (1) يتضح الأثر الذي أحدثته خصائص الفقرات ومستوى قدرة الأفراد على افتراض الاستقلال الموضعي، على الرغم من أن الفقرات جاءت من مجمع فقرات(الحالة المرجعية *100* فقرة) لم يكن بينها اعتماد موضعي، وكانت مستقلة فيما بينها، ولكن عندما تم تصنيفها في نماذج حسب مستوى الصعوبة والتمييز لها وإعادة فحص الاستقلال الموضعي لها، لوحظ الأثر الواضح الذي أحدثته على افتراض الاستقلال الموضعي، والتغير الكبير نسبيا في نسبة الأزواج التي بينها اعتماد موضعي ضمن كل مستوى من مستويات خصائص الفقرات ومستوى القدرة للأفراد. فقد أبرزت الدراسة نتيجة مهمة وهي: عندما كانت خصائص الفقرات (فقرات صعبة وعالية التمييز) أو (فقرات سهلة ومنخفضة التمييز) كانت هناك نزعة واضحة لنسبة الفقرات المنتهكة لافتراض الاستقلال الموضعي، حيث كانت تزداد هذه النسبة بزيادة مستوى القدرة عندما كانت خصائص الفقرات (صعبة وعالية التمييز)، وتقل عندما كانت خصائص الفقرات (سهلة ومنخفضة التمييز)، وهذه النتيجة جاءت متعارضة مع نتيجة دراسة بيل وباتسون ووذرس (*Bell, Pathison, & Withers, 1988*)، التي أشارت إلى أن الاعتماد الموضعي بين الفقرات قد يزداد بزيادة مستوى القدرة للأفراد، وكذلك تعارضت مع نتيجة دراسة أبوشندي (*2011*) التي تناولت أثر قوة العلاقة الارتباطية بين الأبعاد ومستوى قدرة الأفراد على الاستقلال الموضعي، وأشارت في نتائجها بأن نسبة الفقرات المنتهكة لافتراض الاستقلال الموضعي تزداد بزيادة مستوى القدرة للأفراد، والسبب في ذلك أن هذه الدراسات لم تأخذ بعين الاعتبار مستوى خصائص الفقرات كمتغير بل تعاملت مع نفس الفقرات.

إن السبب الذي جعل نسبة أزواج الفقرات في حالة الفقرات الصعبة والتمييز العالي تزيد بزيادة مستوى القدرة، يكمن في الأساس الرياضي الذي يستند إليه المؤشر Q_3 كمعامل ارتباط بين البواقي لاحتمال الإجابة الملاحظة والإجابة المتوقعة، باستخدام معالم الفقرات وقدرة الأفراد لكل زوج من أزواج الفقرات، حيث إن التمييز العالي للفقرات والصعوبة العالية جعلت أنماط الاستجابة المتوقعة بعيدة عن أنماط الاستجابة الملاحظة عند مستوى القدرة العالية والذي أدى بدوره إلى ارتفاع قيم المؤشر الإحصائي Q_3 لكل زوج من أزواج الفقرات، في حين أخذت الأنماط المتوقعة تتقارب مع الأنماط الملاحظة عندما تغير مستوى القدرة من مستوى عالٍ إلى مستوى متوسط ثم إلى مستوى منخفض باضطراد؛ مما جعل نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي الأقل عندما كان مستوى القدرة للأفراد منخفضة. أما عندما كانت الفقرات (سهلة ومنخفضة التمييز) فقد أخذت نسبة الأزواج المنتهكة لافتراض

* عدد الأزواج هو $k(k-1)\backslash 2$ حيث k تمثل عدد فقرات كل نموذج وهو 20 فقرة

الاستقلال الموضعي تقل بزيادة مستوى القدرة، والسبب في ذلك يعزى كما أشير سابقا إلى الأساس الرياضي الذي يستند إليه المؤشر Q_3، حيث أخذت أنماط الاستجابة الملاحظة تتقارب مع أنماط الاستجابة المتوقعة بزيادة مستوى القدرة للأفراد، وبالتالي انعكس ذلك على قيم المؤشر Q_3 مما جعل نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي تتناقص بزيادة مستوى القدرة.

وفي السياق نفسه، لوحظ أنه عندما كانت خصائص الفقرات (صعبة وتمييز منخفض) بأن نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي كانت متقاربة ضمن مستويات القدرة (عالية، ومتوسطة ومنخفضة)، والسبب في ذلك أنه عندما تكون قيم التمييز لفقرات الاختبار منخفضة، هذا يعني عدم قدرة الفقرات على التمييز بين مستويات القدرة المختلفة، فاحتمال الإجابة الصحيحة يكون متقاربا لجميع المستويات، الذي بدوره جعل قيم المؤشر Q_3 متقاربا لكل المستويات، في حين عندما كانت خصائص الفقرات (سهلة وعالية التمييز)، فقد ميزت الفقرات بشكل جيد بين مستويات القدرة العالية والمتوسطة والمنخفضة، حيث كانت نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي الأقل عند مستوى القدرة العالية، وكانت متماثلة عند المستويين المتوسطة والمنخفضة، ومثل هذه النتيجة تؤكد أهمية التمييز كعامل يؤثر على الاستقلال الموضعي أكثر من صعوبة الفقرة التي يتحدد في ضوئها موقع الفرد على متصل القدرة.

وللإجابة عن السؤال الثاني من أجل الوقوف على الفروق بين نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي حسب مستوى القدرة وخصائص الفقرات (نموذج الاختبار)؛ فقد تم استخدام اختبار Z (*Wald Z*) لمقارنة نسبتين مستقلتين من خلال برنامج (*NCSS 11 Statistical Software, 2016*)، والجدول 2 يبين قيم الإحصائي Z والدلالة الإحصائية (*sig*) للفرق بين نسبتين مستقلتين حسب مستوى القدرة للأفراد.

جدول (2): قيم الإحصائي Z والدلالة الإحصائية للفرق بين نسبتين مستقلتين حسب مستوى القدرة للأفراد

المقارنة / مستوى القدرة	الأول مع الثاني		الأول مع الثالث		الأول مع الرابع		الثاني مع الثالث		الثاني مع الرابع		الثالث مع الرابع	
	Z	sig	Z	sig	Z	sig	Z	sig	Z	sig	Z	sig
عالية	1.954	0.052*	3.420	0.001*	2.682	0.007*	1.480	0.139	0.734	0.463	-0.748	0.455
متوسطة	.103	0.918	-0.616	0.538	0.619	0.536	-0.719	0.472	0.516	0.606	1.235	0.217
منخفضة	-3.634	0.0003*	-5.137	0.000*	-4.141	0.000*	-1.548	0.122	-0.520	0.603	1.030	0.303

* دال إحصائيا عند مستوى الدلالة $\alpha = 0.05$

يتضح من النتائج الواردة في الجدول (*2*) بأن الفروق في نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي كانت دالة إحصائيا عند مقارنة النموذج الأول (صعبة وعالية التمييز) مع النموذج الثاني ((صعبة وتمييز منخفض)، وعند مقارنة النموذج الأول مع النموذج الثالث (سهلة وعالية التمييز)، وكذلك عند مقارنة النموذج الأول مع النموذج الرابع (سهلة ومنخفضة التمييز) عندما كان مستوى القدرة للأفراد عالياً ومنخفضاً. وأشارت النتائج كذلك بأنه لم تظهر فروق دالة إحصائيا عند مقارنة نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي لجميع النماذج عند مستوى القدرة المتوسطة.

وللوقوف على دلالات الفروق بين نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي حسب نموذج الاختبار (خصائص الفقرات)، تم إيجاد قيمة الإحصائي Z والدلالة الإحصائية (*sig*)، والجدول *3* يوضح هذه القيم.

جدول (3): قيم الإحصائي Z والدلالة الإحصائية للفرق بين نسبتين مستقلتين حسب مستوى نموذج الاختبار (خصائص الفقرات)

المقارنة / النموذج	قدرة عالية مع متوسطة		قدرة عالية مع منخفضة		قدرة متوسطة مع منخفضة	
	Z	***sig***	***Z***	***sig***	***Z***	***sig***
الاول	1.026	0.305	-5.726	0.000*	4.711	0.000*
الثاني	-0.827	0.408	0.209	0.835	1.036	0.300
الثالث	-3.016	0.003*	-2.813	0.005*	0.205	0.837
الرابع	-1.045	0.296	-1.045	0.296	0.000	1.000

* دال إحصائيا عند مستوى الدلالة $\alpha = 0.05$

تشير النتائج الواردة في الجدول (*3*) بأن الفروق بين نسب أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي كانت دالة إحصائياً عند مستوى الدلالة ($\alpha = 0.05$) عند مقارنة مستوى القدرة العالية مع المنخفضة، وكذلك عند مقارنة مستوى القدرة المتوسطة مع المنخفضة على النموذج الأول للاختبار (فقرات صعبة وعالية التمييز). وبينت النتائج كذلك بأن الفروق بين نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي كانت دالة إحصائياً أيضاً عند مقارنة مستوى القدرة العالية مع القدرة المتوسطة، وكذلك عند مقارنة مستوى القدرة العالية مع القدرة المنخفضة على النموذج الثالث للاختبار(فقرات سهلة وعالية التمييز)، في حين لم تكن الفروق في نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي دالة إحصائياً عند مقارنة مستويات القدرة المختلفة لكلا النموذجين الثاني (فقرات صعبة وتمييز منخفض) والنموذج الرابع (فقرات سهلة وتمييز منخفض).

وبالنظر إلى النتائج الواردة في الجدول (*2*) يتضح بأن الفروق بين نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي جاءت دالة إحصائياً عند مستوى الدلالة ($\alpha = 0.05$) لدى مقارنة النسبتين بين النموذج الأول والنموذج الثاني، وكذلك ما بين النموذج الأول مع الثالث والنموذج الأول مع الرابع عندما كان مستوى القدرة للأفراد عاليا، وكانت النسبة الأعلى لأزواج الفقرات التي يوجد بينها اعتماد موضعي في حالة النموذج الأول (فقرات صعبة وعالية التمييز). إن مثل هذه النتيجة ربما جاءت في الاتجاه الذي تناولته دراسة ريس (*Reese, 1999*) التي أشارت إلى أنه عندما يكون افتراض الاستقلال الموضعي منتهكا بدرجة عالية يؤدي إلى المغالاة في قيم تمييز الفقرات، حيث كان واضحاً من نتيجة الدراسة الحالية بأنه عندما كانت الفقرات ذات قدرة تمييزية عالية وصعوبة عالية أدت إلى وجود انتهاك عالٍ لافتراض الاستقلال الموضعي بين أزواج الفقرات(النموذج الأول) وكانت الأعلى في جميع الحالات، على الرغم من أنها جاءت من مجمع فقرات مستقلة إحصائيا، وهذا يمكن تفسيره على أن مثل هذا النوع من خصائص الفقرات (النموذج الأول) تكون أكثر حساسية لذوي القدرة العالية، مما يترتب عليه أن يجعل احتمالات الإجابة الصحيحة عن الفقرات(الإجابات الملاحظة) بعيدة عن الاحتمالات المتوقعة للإجابة الصحيحة، وهذا بدوره يؤدي إلى زيادة الفرق بين البواقي الملاحظة والمتوقعة لكل زوج من أزواج الفقرات، مما يجعل قيم المؤشر الإحصائي Q_3 لأن ترتفع لكل زوج من أزواج الفقرات، خاصة وأن هذا المؤشر يعتمد على دالة الأرجحية القصوى وتوزيع القدرة للأفراد.

وفي السياق نفسه بينت نتائج الجدول (*2*) بأن الفروق بين النسب لأزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي بين نماذج الاختبار عندما كان مستوى القدرة للأفراد متوسطاً لم تكن دالة إحصائياً عند مستوى الدلالة ($\alpha = 0.05$)، وفي الوقت نفسه جاءت النسب عالية لجميع نماذج الاختبار (انظر جدول 1)، وقد يعزى السبب في ذلك إلى أن أداء هذه الفئة على جميع النماذج بغض النظر عن خصائص الفقرات كان متماثلا؛ مما جعل أنماط الاستجابة الملاحظة لهم على جميع النماذج متقاربة، في حين كانت الأنماط المتوقعة من النموذج اللوجستي ثنائي المعلمة لاحتمالات الإجابة الصحيحة متباينة حسب نموذج الاختبار، وهذا جعل الفروق بين أزواج الفقرات بين ما هو ملاحظ ومتوقع تقريبا متماثلة ولكل النماذج، حيث جعل قيم المؤشر Q_3 متماثلة لأزواج الفقرات، والذي بدوره جعل هذه النسب متقاربة ولكل النماذج، مما ترتب عليه بأن جاءت الفروق بينها غير دالة إحصائياً.

أما عند مقارنة النسب بين نماذج الاختبار عند مستوى القدرة المنخفضة، يلاحظ من نتائج الجدول (*2*) بان الفروق الدالة إحصائيا بين نسب أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي جاءت دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) ما بين النموذج الأول مع الثاني، والأول مع الثالث، والأول مع الرابع على غرار المقارنة بينها عندما كان مستوى القدرة للأفراد عاليا، ولكن بشكل معاكس، حيث يتضح بأن النموذج الأول كان الأقل في نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي، وهذه النتيجة منطقية حيث يتضح بأن استجابات أفراد القدرة المنخفضة على هذا النموذج (فقرات صعبة وعالية التمييز) متماثلة لدرجة كبيرة، وربما

تعتمد على التخمين كون فقرات الاختبار أعلى بكثير من قدرة المفحوصين، وجاءت متوافقة مع الاستجابات المتوقعة، وهذا جعل الفروق بين ما هو متوقع وملاحظ لكل زوج من أزواج الفقرات منخفضة، مما جعل نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي الأقل لهذا النموذج ولكل الحالات، ولكن عندما تحولت خصائص الفقرات لتصبح فقرات سهلة وعالية التمييز (النموذج الثالث) أصبحت نسبة الفقرات المنتهكة لافتراض الاستقلال الموضعي هي الأعلى عند مستوى القدرة المنخفضة، والسبب في ذلك يعود إلى قيم تمييز الفقرات لكونها عالية، حيث تجعل منحنيات خصائص الفقرات حادة؛ مما يدل على أن التغير البسيط نسبياً في قدرة الفرد يجعل احتمال الإجابة الصحيحة عن الفقرة يتباين تبايناً ملحوظاً، وهذا التباين في الإجابات الصحيحة له أثره على قيم المؤشر Q_3 على الرغم من أن الفقرات سهلة.

وعند مقارنة النسب حسب نموذج الاختبار(خصائص الفقرات) بينت نتائج الجدول (3) بأن الفروق بين نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي جاءت دالة إحصائياً عند مستوى الدلالة ($\alpha = 0.05$) عند مقارنة القدرة العالية مع القدرة المنخفضة، والقدرة المتوسطة مع القدرة المنخفضة على النموذج الأول، حيث يتضح زيادة نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي على هذا النموذج بزيادة مستوى القدرة للأفراد، فكانت أنماط استجابة الأفراد ذوي القدرة المنخفضة على هذا النموذج متماثلة، وربما يلعب عامل التخمين هنا دوراً في الوصول إلى الإجابة الصحيحة، لكون مستوى صعوبة وتمييز فقرات الاختبار أعلى بكثير من مستوى قدرتهم، بينما عندما تم مقارنة نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي ما بين أفراد القدرة العالية والمتوسطة على هذا النموذج لم تكن دالة إحصائيا، وذلك بسبب ابتعاد الاستجابات الملاحظة عن المتوقعة لكل زوج من أزواج الفقرات كما أشير سابقاً، مما جعل نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي متقاربة عند مستوى القدرتين.

وأشارت النتائج الواردة في الجدول (3) أيضا إلى أنه عند مقارنة نسب أزواج الفقرات المنتهكة لافتراض الاستقلال الموضعي على النموذج الثالث (فقرات سهلة وعالية التمييز) وعند مقارنة القدرة العالية مع القدرة المتوسطة، والقدرة العالية مع القدرة المنخفضة قد جاءت دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$)، بينما لم تكن المقارنات دالة إحصائياً عند مقارنة مستويات القدرة المختلفة على النموذجين: الثاني (فقرات صعبة ومنخفضة التمييز) والنموذج الرابع (فقرات سهلة ومنخفضة التمييز). إن هذه النتيجة تؤكد بأن تمييز الفقرات يلعب دوراً أكبر من صعوبة الفقرات في التأثير على الاستقلال الموضعي، حيث إنه عندما تكون الفقرات ذات قدرة تمييزية منخفضة فإنها تفشل في التمييز بين مستويات القدرة المختلفة.

الاستنتاجات والتوصيات

إن النتائج التي توصلت إليها الدراسة الحالية أشارت عموماً إلى وجود أثر لخصائص الفقرات ومستوى القدرة للأفراد على افتراض الاستقلال الموضعي، حيث أظهرت النتائج نتيجة مهمة بأنه عندما تكون خصائص الفقرات صعبة وعالية التمييز لوحظ بأن هناك نزعة لأن تزيد نسبة الفقرات المنتهكة لافتراض الاستقلال الموضعي بين الفقرات بزيادة مستوى القدرة للأفراد، في حين أخذت هذه النسبة تتناقص بزيادة مستوى القدرة للأفراد عندما تكون خصائص فقرات الاختبار سهلة ومنخفضة التمييز، بينما لم تظهر نزعة معينة لسلوك نسبة الفقرات المنتهكة لافتراض الاستقلال الموضعي عندما تتباين خصائص الفقرات (صعبة ومنخفضة التمييز أو سهلة وعالية التمييز) عبر مستويات القدرة المختلفة، حيث كانت تتباين من مستوى قدرة إلى مستوى قدرة أخرى، وهذا قد يبدو كمشكلة عندما يتم استخدام نماذج نظرية الاستجابة للفقرة أحادية البعد عند تطبيق مجموعة من الفقرات تتصف بخصائص معينة على مستويات مختلفة لقدرة الأفراد، والذي بدوره يطرح استفسارا بحثيا وهو: هل هناك حاجة لإعادة معايرة الفقرات مرة أخرى عند تطبيقها على مستوى معين من قدرة الأفراد؟ ومثل هذا الاستفسار بحاجة لمزيد من الأبحاث والدراسات.

ومن النتائج التي توصلت إليها هذه الدراسة كذلك بأن معلمة التمييز للفقرات كانت تلعب دورا أكثر من معلمة الصعوبة في التأثير على افتراض الاستقلال الموضعي، حيث كانت الفروق بين نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي دالة إحصائيًا حسب مستوى القدرة على النموذج الأول (فقرات صعبة وعالية التمييز) وعلى النموذج الثالث (فقرات سهلة وعالية التمييز)، في حين لم تكن هناك دلالة إحصائية بين نسب الأزواج المنتهكة لافتراض الاستقلال الموضعي حسب مستوى القدرة على النموذج الثاني (فقرات صعبة ومنخفضة التمييز) والنموذج الرابع (فقرات سهلة ومنخفضة التمييز).

وعلى الرغم من أهمية النتائج التي توصلت إليها الدراسة الحالية، إلا أن بعض القيود على هذه الدراسة تحد من تعميم النتائج بشكل عام، وذلك لوجود عوامل أخرى قد تلعب دوراً في التأثير على افتراض الاستقلال الموضعي بين الفقرات، منها قيم الحد الفاصل للمؤشر Q_3، والتي قد يكون لها الأثر في زيادة أو تقليل نسبة الأزواج المنتهكة لافتراض الاستقلال الموضعي لكونه مؤشراً وليس إحصائياً يتبع توزيع إحصائي معين، كذلك قيم معلمة التخمين للفقرات والتي لها الأثر على دقة تقديرات معالم الفقرات وقدرة الأفراد، مما يجعل الحاجة ماسة لإجراء دراسة مماثلة، تأخذ بالحسبان الحد الفاصل (القيمة الحرجة) لمؤشر Q_3 كمتغير ومستوى التخمين للفقرات، بالإضافة إلى متغيرات الدراسة ودراسة أثرها على افتراض الاستقلال الموضعي.

المراجع

أبو شندي، يوسف (2011). كواشف الاستقلال الموضعي في نظرية الاستجابة للفقرة ومدى تأثرها بقوة العلاقة الارتباطية بين الأبعاد ومستوى القدرة للمفحوصين. *مجلة جامعة تشرين للبحوث والدراسات* – سلسلة الآداب والعلوم الإنسانية، *33* (*3*)، *49* – *62*.

حمادنه، إياد (2011). الكشف عن الارتباط الموضعي بين أزواج فقرات الاختبار باستخدام مؤشر Q_3. *مجلة العلوم التربوية والنفسية*، 12 (1)، 40 – 68.

علام، صلاح الدين (2002). *القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة*. القاهرة: دار الفكر العربي

Anstasi, A. & Urbina, S. (2005). *Psychological Testing (7 th ed*.). New Jersey: Prentic-Hall, Inc.

Allen, M. & Yen, W. (1979). *Introduction to measurement theory*. California: Wasdsworth, Inc.

Baker, F. (2001). *The basics of item response theory (2nd ed)*. College Park, MD: ERIC Clearing House on Assessment and Evaluation.

Bell, R., Pattison, P. & Withers, G.(1988). Conditional independence in a clustered item test. *Applied Psychological Measurement, 12, 15*-26.

Chen, W. & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics, 22(3),* 265-289.

Chen, C. & Wang, W.(2007).Effects of ignoring item interaction on item parameter estimation and detection of interacting items. *Applied Psychological Measurement, 31, 388*-411

Christensen K., Makransky G.,& Horton M.(2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch Model using residual correlations. *Applied Psychological Measurement, 41(3),* 178-194

Crocker, L. & Algina, J.(1986). *Introduction to classical and modern test theory*. N. Y.: Harcourt Jovanovich.

De Ayala, R. (2009).*The theory and practice of item response theory*. NY: Guliford.

De Gruijter, D. & Van der Kamp, L. 2008).*Statistical test theory for the behavioral sciences.* New York: Chapman and Hall.

Embretson , S. & Reise, S. (2000). *Item response theory for psychologists.* New Jersey: Lawrence Erlbaum Associates, Publishers.

Hambleton, R.(1994). Item response theory: A broad psychometric frame work for measurement advances. *Psicothema, 6*(3), 535-556.

Hambleton, R. & Swaminathan, H.(1985). *Item response theory: Principles and applications*. Boston: Kluwer-Nijhoff.

Hambleton, R., Swaminathan, H., & Rogers, H.(1991). *Fundamentals of item response theory*. Newbury Park CA: Sage Publication.

Han, K. & Hambleton , R.(2007). *User's manual for WinGen:Windows Software that Generated IRT Model Parameter and Item Response.* Center for Educational Assessment Research Report No 642, Amherst, MA: University of Massachusetts Center for Educational Assessment.

Harris, D. (1989).Comparison of 1-, 2-, and 3-parameter IRT Models. *Educational Measurement: Issues and Practice, 8,* 35-41.

Hattie, J.(1985).Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement. 9*, 139-164

Henard, D. (2000).Item response theory. In L. Grimm & Yarnold (Ed), *Reading and understanding more multivariate statistics*.(pp 67-97).Washington DC: American Psychological Association.

Killer, L. Swaminathan, H. & Sireci, S. (2003). Evaluating scoring Procedures for context–dependent item sets. *Applied Measurement in Education, 16*(3), 207-222.

Kim. S., Cohen. A., & Lin.Y. (2005).*LDID: A Computer program for local dependence indices for dichotomous Items.* Version 1.0.

Lord, F., & Novick, M. (1968). *Statistical Theories of mental test scores*. London, Addison, Wesley: Publishing Company.

Miller, T. (1991). *Empirical estimation of standard errors of compensatory MI model parameters obtained from the NOHARM estimation program.*(ACT Research Report No. onr91-2). Iowa City IA: ACT Inc.

NCSS 11 Statistical Software (2016). *NCSS, LLC*. Kaysville, USA, css.com/software/ncss.

Van der Linden, W. (2010).Item response theory. *International Encyclopedia of Education, 4*, 81-88.

Warm, T. (1978). *A Primer of Item Response Theory*. Technical Report No, 940279, Oklahoma City: U.S, Coast Guard Institute.

Yen, W. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125-145.

Yen, W. (1993).Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement, 30* (3), 187-213.

Zenisky, A., Hambleton, R., & Sireci, S. (2003). Identification and evaluation of local item dependencies in the medical college admissions test. *Journal of educational Measurement*, 39, 1-16.

Reese, L. (1999). *A classical test theory perspective on LSAT local item dependence (Report No. LSAC-R-96-01).*Newtown, PA: Law School Admission Council.(ERIC Document Reproduction Service No. ED469172).

Shen, L. (1997).*Quantifying item dependency by Fisher's Z.* paper presented A the annual meeting of the American Educational Research Association Chicago, IL,March 24-28,1997(ERIC Document Reproduction.

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. NY, April8-12, 1996 (ERIC Document Reproduction Service No. ED 400311).

Tuerlinckx, F., & De Boeck, P. (1998). Modeling local item dependencies in item response theory. *Psychologica Belgica, 38(2),* 61-82.