مجلة جامعة الملك سعود، م ٥، العلوم التربوية والدراسات الإسلامية (٢)، ص ص٣٧٣-٣٩٠ (١٤١٣هـ/١٩٩٣م).

تحيز بنود اختبار وكسلر لذكاء الأطفال المعدل (الصورة السعودية) حسب الجنس

عبدالله بن علي القاطعي
أستاذ مساعد، قسم علم النفس، كلية التربية،
جامعة الملك سعود، الرياض، المملكة العربية السعودية

ملخص البحث. بدأ الاهتمام بدراسة التحيز في الاختبارات النفسية في النصف الأخير من هذا القرن، نظرًا لاعتبارات إنسانية. إذ نجد أن بواكير دراسة التحيز ركزت على مفهوم العدالة الثقافية للاختبارات. وامتد هذا الاهتمام إلى دراسة التحيز في بنود الاختبار، فاعتبر جزءًا من الأساسيات التي ينبغي التقصي عنها أثناء بناء الاختبار.

وقد درس كثير من الباحثين مفهوم التحيز العرقي لاختبار وكسلر لذكاء الأطفال المعدل، إلا أن أحدًا لم يتناول التحيز الجنسي لهذا الاختبار على الرغم من أهميته. لذا فإن هذه الدراسة تهدف إلى التحري عن تحيز بنود اختبار وكسلر لذكاء الأطفال المعدل (الصورة السعودية) تبعًا لمتغير الجنس.

وتمثل عينة التقنين لاختبار وكسلر لذكاء الأطفال المعدل (الصورة السعودية) عينة البحث. وقد تضمنت عينة التقنين ٥٥٠ طفلاً و٥٥٠ طفلة بواقعَ ٥٠ طفلاً و٥٠ طفلة في كل مرحلة عمرية من المراحل الإحدى عشرة (٦ - ١٦ سنة). أما الأسلوب الإحصائي المستخدم، فهو كميلي كاي تربيع نظرًا لاعتبارات نظرية وعملية.

وأظهرت نتائج البحث أن نسبة البنود المتحيزة تبعًا لمتغير الجنس في اختبار وكسلر لذكاء الأطفال المعدل لم تكن عالية، وأن البنود المتحيزة موزعة بالتساوي تقريبًا بين الجنسين. أي أن عدد البنود المتحيزة لصالح البنين يساوي تقريبًا عدد البنود المتحيزة لصالح البنات. وعليه فإن تأثير التحيز على الاختبار الكلي ليس ذا قيمة. وقد أكدت هذه النظرة نتائج الدراسة الارتباطية لصعوبة البنود لكل من البنين والبنات.

مقدمة

بدأ توجه البحوث نحو الدراسة الجادة لمفهوم تحيز بنود الاختبار بدراسة قام بها Eells وزملاؤه عام ١٩٦١م عن تحيز بنود الاختبار [١، ص ٢١٤]. ومنذ ذلك الحين بدأ الاهتمام بمفهوم التحيز من الجمعيات العلمية الأمريكية وآثاره على سلامة التصنيف أو الانتقاء.

وقد انقسمت دراسة التحيز إلى قسمين رئيسين، ركز أحدهما على عدالة الاختبار في علاقته بمحك معين (أي عدم اختلاف قدرة الاختبار على التنبؤ بأداء مجموعات مختلفة من المجتمع)، وركز ثانيهما على طبيعة الإجابة عن كل سؤال من أسئلة الاختبار (أي عدم اختلاف الإجابة عن السؤال تبعًا لاختلاف المجموعات).

ونظرًا لتعدد مجالات التحيز ونظرة الباحثين لهذا المفهوم، فإنه ليس بالإمكان إعطاء تعريف واحد للتحيز، والمتقصي لدراساتَ التحيز يواجه بعدد من التعريفات التي لا يمكن قبول بعضها ورفض بعضها الآخر. بيد أن صعوبة الحصول على تعريف عام أو شامل للتحيز تعود إلى «ذلك التفاعل المعقد بين كيفية قياس اختبار ما وضع لقياسه وما إذا كان محتوى الاختبار مناسبًا لاستخدام معين» [٢، ص٣١٨].

بيد أنه بالإمكان القول بأن التحيز بصوره المختلفة يعود إلى تفضيل مجموعة دون أخرى ليس بناءً على الخصائص التي يقيسها أو يفترض أن يقيسها الاختبار، وإنما بناء على جوانب ليست ذات علاقة بالسمة المقاسة. وهذه النظرة الشاملة لمفهوم التحيز يمكن أن يدرج تحتها مفهوم التحيز في الاختيار (تحيز الاختبار) ومفهوم التحيز في الاستجابة (تحيز البنود).

التحيز في اختبار وكسلر لذكاء الأطفال المعدل

انقسمت دراسات التحيز لاختبار وكسلر لذكاء الأطفال المعدل إلى قسمين رئيسين، أحدهما تناول التحيز الكلي للاختبار لمجموعات عرقية مختلفة؛ وثانيهما تناول تحيز البنود في بعض الاختبارات الفرعية أو كلها. ونظرًا لكون الدراسة الحالية منصبة على دراسة تحيز بنود

الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل، فسيقتصر العرض هنا للدراسات التي اهتمت بتحيز البنود.

فقد درس كل من Oakland and Feigenbaum تحيز بنود اختبار وكسلر لذكاء الأطفال المعدل (WISC-R) على عينة من الأطفال الأنجلوساكسون والزنوج وذوي الأصول المكسيكية. وقد استخدما معاملات الصعوبة والتمييز كأساس للمقارنة (متغيرات تابعة). ووجدا ـ من خلال استخدام تحليل التباين ـ أن المجموعات العرقية الثلاث لا تختلف سوى في معاملات الصعوبة. كما أنه لا يوجد نمط معين من الاختلاف يمكن إرجاعه إلى تحيز الاختبار. وقد خلصا إلى القول «بأن الاختلاف في الصعوبة، قد يكون راجعًا إلى الاختلاف في القدرة العامة» [٣، ص٩٧٢].

وكذلك درس Sandoval تحيز بنود WISC-R على عينة من الأصول العرقية الثلاثة التي ورد ذكرها في الدراسة السابقة. وقد عرَّف التحيز على أنه التفاعل بين البند والمجموعة، واستخدم بذلك تحليل التباين المتعدد (MANOVA). وقد أظهرت النتائج أن ٨٠ بندًا من الاختبار الكلي تختلف فيها الصعوبة اختلافًا دالًّا إحصائيًّا لصالح الأنجلوساكسونيين، أي أنها أقل صعوبة للأطفال الأنجلوساكسونيين منها للأطفال الزنوج. كذلك فإن ٧٦ بندًا من الاختبار الكلي كانت صعوبتها أقل للأنجلوساكسونيين منها لذوي الأصول المكسيكية [٤، ص٩٢٣].

أما Meile، فقد درس تحيز بنود اختبار WISC-R على عينة من الأطفال الأنجلوساكسونيين والأطفال الزنوج. وعُرِّف التحيز على أنه التفاعل بين البند والعرق. وقد أظهرت النتائج أن التفاعلات دالة إحصائيًّا لكل الاختبارات الفرعية. بيد أن التفاعلات لم تفسر سوى ٥٪ من التباين العام. وقد لاحظ Miele في بعض التحليلات الإضافية أن البنود الأكثر تمييزًا داخل كل مجموعة عرقية، هي البنود نفسها الأكثر تمييزًا بين المجموعات العرقية. وهذا بالتالي جعل Miele يقول إن الاختلاف قد يكون ناتجًا عن الفروق في النضج العقلي وليس عن تحيز بنود الاختبار [٥، ص١٦٢].

وتـزيد مصـداقية التفسـير البـديل الـذي قال به Miele على طبيعـة الفـروق بين المجموعتين حين يؤخذ في الحسبان ما أثير من جدل حول مصداقية استخدام تحليل التباين في دراسة تحيز البنود. إذ نجد أن Hunter [٦] يعتبر شواهد التحيز الناتجة عن التفاعل بين البند والاختبار غير معبرة عن التحيز بقدر ما هي معبرة عن اختلاف مستويات القدرة بين المجموعات المقارنة. كذلك فإن Cotter and Berk [٧] أشارا إلى صعوبة القول بأن الفروق الناتجة عن التفاعل بين البند والمجموعة ناتجة إما عن تحيز البند أو اختلاف مستويات القدرة.

أمـا حين درس Camilli and Shepard مصـداقية استخـدام تحليل التباين لدراسة التحيز ـ من خلال دراسة تمثلية simulated study ـ فقد وجدا أن تحليل التباين قد لا يظهر التحيز عندما يكون موجودًا أصلاً، كما أنه في المقابل يظهر بعض البنود على أنها متحيزة في الوقت الذي لا يكون هذا صحيحًا، نظرًا لكون التحيز مرتبطا بأثر المتغيرات main effects وليس بالتفاعل بينها. وعليه فإنهما خلصا إلى القول بأن «دراسات التحيز الموجودة والمبنية على تحليل التباين ينبغي إلغاؤها. كما أن تحليل التباين ينبغي عدم التوصية باعتباره طريقة من طرق دراسة التحيز حتى ولو كان هذا في مراحل الغربلة الأولية للبنود» [٨، ص٩٨].

وقد يكون من الأفضل ـ في ضوء النقد الآنف لطريقة تحليل التباين ـ النظر إلى نتائج الدراسات السابقة نظرة الشك لاعتمادها على طريقة تحليل التباين لدراسة تحيز البنود.

أمـا Mishra ، فقد درست تحيز بنود بعض الاختبارات الفرعية من اختبار وكسلر لذكاء الأطفال المعدل على عينة من الهنود الحمر والأنجلوساكسونيين. بيد أن الباحثة عرفت التحيز من خلال مفهوم التفاعل، واستخدمت الإحصائي «ل» تربيع likelihood chi-square (L^2) للتعـرف على البنـود المتحيزة (العـرق × القـدرة × الدرجة) في كل من المعلومات والمتشابهات والمفردات. وأظهرت النتائج أن خمسة بنود من اختبارات المعلومات، وأربعة بنود من المتشابهات، وستة بنود من المفردات كانت متحيزة بناء على نتائج «ل» تربيع. لكن الباحثة تحفظت على مصداقية النتائج نظرًا لكون مجموعة الهنود أخذت من عائلات

ذات مستوى اجتماعي منخفض. وعليه فإن النتائج ينبغي ألا تفسر على أنها دعوة «لحذف هذه البنود من الاختبار» [٩، ص٤٦٣].

ونظرًا لأن دراسة Mishra اقتصرت على ثلاثة اختبارات فرعية من الجانب اللفظي لاختبار وكسلر لذكاء الأطفال المعدل، فإنه لا يمكن اعتبارها دراسة متكاملة لجوانب التحيز لبنود اختبار وكسلر لذكاء الأطفال المعدل.

ولم تقتصر دراسة التحيز على الجوانب الموضوعية، بل تناولت أيضًا الطرق التقديرية وفاعليتها في التعرف على التحيز في بنود الاختبار. فنجد أن كثيرًا من نقاد اختبار وكسلر ينظرون إلى البند السادس من اختبار الفهم على أنه خير مثال للبنود المتحيزة ضد الزنوج (البند السادس: ماذا تفعل لو أن طفلاً «طفلة» أصغر منك ضربتك)؟ إذ إن الإجابة المنطقية للطفل الزنجي هي «أضربه» بينما تعد هذه الإجابة خاطئة. أما الإجابة الصحيحة والمتمثلة في «عدم مضاربته،» فإنها تعد كارثة في الوسط الاجتماعي الزنجي [١٠، ص٥٠٢]. ويشير كل من Jensen [١١] و Miele [٥] و Reynold and Kaiser [١٠] إلى أن هذا الادعاء لم تدعمه النتائج الإمبريقية. إذ أظهرت النتائج أن البند المشار إليه أقل صعوبة للزنوج منه للأنجلوساكسن. فقد كانت رتبته من حيث السهولة للزنوج ٤٢ وللأنجلوساكسن ٤٧ [١١، ص ٥٥٤].

وتؤكد دراسة Sandoval and Miille التوجه السابق حين درست دقة الطريقة الحكمية في التعرف على البنود الصعبة في اختبار وكسلر لذكاء الأطفال المعدل للأصول العرقية الثلاثة (الأنجلوساكسن والزنجي والمكسيكي). فقد اختار الباحثان عددًا من البنود التي صنفت بالطرق الإحصائية على أنها صعبة إما للزنوج أو لذوي الأصول المكسيكية، كما اختارا عددًا آخر من البنود التي كانت صعوبتها متساوية لكل الأصول العرقية الثلاثة. وقد طلب من مجموعة المحكمين أن يقرأوا كل بند ويقرروا ما إذا كان البند أسهل لأطفال الأقليات مقارنة بأطفال الأكثرية؛ أو أنه أسهل لأطفال الأكثرية مقارنة بأطفال الأقليات؛ أو أن سهولته متساوية لكل الفئات. وقد أظهرت النتائج أن المحكمين لم يستطيعوا التمييز

بين البنـود الصعبة لفئة دون أخرى وبين البنود التي تعد صعوبتها متساوية لكل الفئات [١٢، ص٢٤٢].

هدف الدراسة وأهميتها

ينظر بعض الباحثين إلى مفهوم التحيز على أنه نوع من عدم صدق الاختبار [١٣؛ ١٤]، كما أن البعض الآخر ينظر إليه على أنه جزء من الاختبار [١٥]. فالتحيز كما يراه Shepard خاصية متـأصلة في الاختبار، بينما الصدق خاصية مرتبطة باستخدام الاختبار [١٥، ص١٠]. وهذا الفصل الدقيق بين مفهومي التحيز والصدق يهدف إلى إحداث نوع من الأولويات. إذ إنه حين ينظر للتحيز على أنه مرتبط بجوهر الاختبار فإن ذلك يعني حتمية دراسـة تحيز البنـود في الـوقت نفسـه الـذي يدرس فيه خصـائصها الإحصائية الأخرى (كالصعوبة والتمييز مثلاً).

لذا فإن دراسـة تحيز بنـود الاختبار تعد ضرورية لكل الاختبارات النفسية [١٦؛ ١٧]. فهي مهمة لمعدي الاختبارات لما يترتب عليها من أمور مرتبطة بصدق البنود وعدلها الثقافي [١، ص٢١٤]. ولذلك نجد أن الاهتمام بمفهوم التحيز لم يقتصر على الجوانب العرقية في الوقت الحاضر، بل تعداه ليشمل الجوانب المرتبطة بالجنس، أي الاختلافات بين البنين والبنات [١٧]. وعلى الرغم من أهمية دراسة التحيز الناتج عن اختلاف الجنس، إلا أن هذا المجال لم يحظ بالعناية الكافية [١٨]، ولهذا نجد أن Wilson et al. يقولون إن «هناك ندرة في البحوث التي أجريت لدراسة التحيز الناتج عن الجنس مقارنة بالتحيز الناتج عن العرق» [١٧، ص٢٨٩].

وبما أن أغلب دراسات التحيز لبنود اختبار وكسلر لذكاء الأطفال المعدل استخدمت طرقًا إحصائية مشكوكًا في صحتها ـ في أفضل الأحوال ـ ونظرًا لقلة دراسات التحيز المبنية على متغير الجنس، ولكون دراسة التحيز تعد مطلبًا أساسيًا لأي اختبار جديد، فإن الدراسة الحالية تحاول التعرف على تحيز بنود اختبار وكسلر لذكاء الأطفال المعدل (الصورة السعودية) للبنين والبنات وذلك عن طريق الإجابة عن السؤال التالي: هل تختلف احتمالية الإجابة

الصحيحة عن كل بند من بنود الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل تبعًا لاختلاف الجنس؟

الطريقة الإحصائية المستخدمة

رُبط مفهوم التحيز باختلاف خصائص البند تبعًا لاختلاف المجموعات المقارنة [١٥]. أما مفهوم عدم التحيز، فإنه عرف على النحو التالي:

يُعد البند غير متحيز إذا كانت نسبة الإجابة الصحيحة للفئة الأولى (بنات مثلاً) مساوية لنسبة الإجابة الصحيحة للفئة الثانية (بنين) عن البند، شريطة تساوي درجاتهم في الاختبار الذي يحتوي البند [١٩، ص٢].

ويعتبر Camilli X^2 مناسبًا للدراسة الحالية نظرًا لتمشيه مع التعريف السابق، ولكونه سهل الاستخدام وذا قاعدة نظرية جيدة. ففي دراسة مقارنة لطرق تحيز البنود، أوضح القاطعي [٢٠] أن كميلي كاي تربيع يعد أفضل بديل لطريقة السمات الكامنة، إذ إن نتائج الدراسات الإمبريقية والخلفية النظرية لكميلي كاي تربيع تجعله بديلاً عمليًا لطريقة السمات الكامنة differential item function(DIF).

الإجـــراءات

العينة

تمثل عينة التقنين للصورة السعودية من اختبار وكسلر لذكاء الأطفال المعدل عينة البحث الحالي، أي أن عينة البحث الحالي هي عينة التقنين نفسها. وقد تضمنت عينة التقنين ٥٥٠ طفلاً و٥٥٠ طفلة بواقع ٥٠ طفلاً و٥٠ طفلة في كل فئة عمرية من الفئات العمرية الإحدى عشرة (من ٦ سنوات إلى ١٦ سنة). وقد تم سحب عينة التقنين عشوائيًا عن طريق الجداول العشوائية بعد أن تم حصر الفئات العمرية المختلفة في صفوف مدارس التعليم العام والخاص. وقد كانت النسب المختارة من كل فئة عمرية في كل منطقة من مناطق الرياض متمشية مع نسبتها للمجموع الكلي لتلك الفئة العمرية.

الأداة

أجريت العديد من الدراسات عن ثبات الصورة السعودية وصدقها لاختبار وكسلر لذكاء الأطفال المعدل [١] ودلت جميعها على أن الاختبار يتمتع بقدر عال من الثبات والصدق (تراوحت قيم الثبات بين ٠,٧٢ - ٠,٩٤ للاختبارات الفرعية، وبين ٠,٩٥ - ٠,٩٦ للأجزاء اللفظية والعملية، و٠,٩٦ للاختبار الكلي). وقد تمت دراسة الصدق بعدة طرق، منها صدق التعلق بمحك وصدق التكوين الفرضي من خلال المقارنة بين المجموعات والمقارنة بين الأعمار المختلفة، وكذلك التحليل العاملي. وجميع تلك الطرق كانت مؤيدة لصدق الاختبار [٢١].

التصميم والإجراء الإحصائي

أوضحت Scheuneman [١٩، ص١٤٥] أن مستويات القدرة المناسبة لمربع كاي تتراوح بين «٣ و ٥» لضمان عدم الإخلال بشرط مربع كاي (مثل عدم الحصول على تكرار متوقع أقل من ٥٪ في أي خلية). وعليه فإنه أمكن التوصل إلى ثلاثة مستويات من القدرة لكل سؤال دون الإخلال بشرط كاي تربيع في الدراسة الحالية. بيد أن مستويات القدرة تختلف باختلاف صعوبة السؤال. فمثلاً كانت مستويات القدرة المناسبة للسؤال الخامس من اختبار المفردات على النحو التالي:

ا ـ مستوى القدرة المنخفض ويتضمن جميع الأطفال الذين حصلوا على نسبة ذكاء أقل من ٨٠.

---

١ تعتبر هذه الأداة ضمن مجموعة من الأدوات كيفت لتلائم الثقافة السعودية، وتم حساب ثباتها وصدقها بعدة طرق، وكذلك استخرجت معاييرها على عينة ممثلة ليتم استخدامها في مشروع برنامج الكشف عن الموهوبين ورعايتهم المدعوم من قبل مدينة الملك عبدالعزيز للعلوم والتقنية ووزارة المعارف والرئاسة العامة لتعليم البنات. ويتكون فريق البحث لهذا الجزء من:

(ا) عبدالله النافع	باحثًا رئيسًا
(ب) عبدالله القاطعي	باحثًا مشاركًا
(جـ) الجوهرة السليم	باحثة مشاركة

وقد تم استئذان مدينة الملك عبدالعزيز للعلوم والتقنية لاستخدام بيانات عينة التقنين.

ب ـ مستوى القدرة المتوسط ويتضمن جميع الأطفال الذين حصلوا على نسبة ذكاء بين ٨٠ و٩٠.

جـ ـ مستوى القدرة المرتفع ويتضمن جميع الأطفال الذين حصلوا على نسبة ذكاء أكثر من ٩٠.

أما مستويات القدرة المناسبة للسؤال رقم ٢٩ من الاختبار نفسه، وهو سؤال عالي الصعوبة، فقد كانت على النحو التالي:

ا ـ مستوى القدرة المنخفض ويتضمن جميع الأطفال الذين حصلوا على نسبة ذكاء أقل من ١٠٥.

ب ـ مستوى القدرة المتوسط ويتضمن جميع الأطفال الذين حصلوا على نسبة ذكاء بين ١٠٥ و ١١٩.

جـ ـ مستوى القدرة المرتفع ويتضمن جميع الأطفال الذين حصلوا على نسبة ذكاء أكثر من ١١٩.

وعليه فإن الدرجات الفاصلة لمستويات القدرة وعدد الأطفال في كل مستوى اختلف باختلاف صعوبة البند، إلا أن التصميم الإحصائي أصبح على النحو الموضح في جدول رقم ١.

جدول رقم ١. التصميم الإحصائي لكميلي كاي تربيع.

المستوى	القدرة	×	الجنس	×	الدرجة
	٣		٢		٢
١	مرتفع		ذكور		واحد
			إناث		صفر
٢	متوسط		ذكور		واحد
			إناث		صفر
٣	منخفض		ذكور		واحد
			إناث		صفر

ولحساب Camilli X^2 فإنه تم اتباع الخطوات التالية:

١ ـ حددت مستويات القدرة تبعًا لشروط مربع كاي.

٢ ـ حسب كاي تربيع لكل مستوى من مستويات القدرة لمصفوفة مكونة من ٢×٢.

٣ ـ جمعت قيم كاي تربيع لمستويات القدرة الثلاثة.

٤ ـ حسبت درجة الحرية على النحو التالي: ا (ب ـ ١) حيث:

ا = مستويات القدرة

ب = مستويات الجنس

وتعتبر الخطوات السابقة خطوات عملية لحساب كميلي كاي تربيع [٢٢، ص ١٢٩].

النتائج

يوضح جدول رقم ٢ قيمة كميلي كاي تربيع لكل بند من بنود الاختبارات اللفظية والعملية من الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل. وحين تقارن القيم المدرجة في الجدول بالقيمة الجدولية ٧,٨١، نجد أن عدد البنود الدالة إحصائيا عند مستوى الدلالة ٠,٠٥ من بين بنود الجانب اللفظي هي ٢٧ بندًا من أصل ١٢٩ بندًا في الجانب اللفظي من الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل.

أما عدد البنود الدالة إحصائيًا عند مستوى الدلالة ٠,٠٥ من بين بنود الجانب العملي فهي ٨ بنود من أصل ٦١ بندًا. لذا فإن عدد البنود المتحيزة من الاختبار الكلي لاختبار وكسلر لذكاء الأطفال المعدل (الصورة السعودية) بلغ ٣٥ بندًا من أصل ١٩٠ بندًا، أي أن نسبة البنود المتحيزة من الصورة السعودية لاختبار وكسلر بلغت ١٨٪.

كذلك يوضح جدول رقم ٣ أرقام البنود المتحيزة في كل اختبار من اختبارات الجزء اللفظي والجزء العملي. إلا أن أكثرية البنود المتحيزة كانت من اختبار المفردات يليه المعلومات، ثم الفهم، فالحساب وترتيب الصور، ثم المتشابهات والمتاهات. ويأتي في المرتبة الأخيرة اختبار إعادة الأرقام وتجميع الأشياء. أما اختبار المكعبات فلم تظهر بنوده أي نوع من التحيز.

لذا فإن الإجابة عن سؤال البحث الخاص «باختلاف احتمالية الإجابة الصحيحة عن كل بنـد من بنود الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل تبعًا لاختلاف الجنس» هي أن جميع بنود الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل لا تتحيز لجنس دون آخر عدا البنود الموضحة في جدول رقم ٣. بيد أن بعض هذه البنود متحيزة لصالح البنين وبعضها الآخر متحيز لصالح البنات. لذا فقد وضعت إشارة ناقص أمام رقم البنـد حين يكون متحيزا لصالح البنين، وإشارة زائد أمام رقم البند حين يكون متحيزًا لصالح البنات لكي تكون الإشارة مماثلة لتلك التي أمام قيمة مربع كاي في جدول رقم ٢، وهذا الإجراء يدعى signed test. وقد استخدم كثيرًا في مجال التحيز بيد أن الإشارة توضع قبل جمع قيمة مربع كاي لمستويات القدرة المختلفة [٢٢، ص١٢٧].

جدول رقم ٢. قيم كميلي كاي تربيع لبنود الاختبارات الفرعية للجوانب اللفظية والعملية من الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل.

	كميلي كاي تربيـــع								
رقم	الجانب اللفظي						الجانب العملي		
البند	المعلومات	المتشابهات	الحساب	المفردات	الفهم	إعادة الأرقام	تكملة الصور	ترتيب الصور	رسوم المكعبات	تجميع الأشياء	المتاهات
١	٠,٠٠	٣,٣٢	٢,٥٦	٦,٨٩	٠,٠٠	٣,٥١	٠,٩٧	+٢٦,٥٥	,٠٠	,٠٠	,٠٠
٢	,٠٠	+١٠,١٠	١,٢٠	٧,٥٥	,٠٠	,٢٣٧	,٧٨	٢,٦٢	٢,٢٨	٣,٠٨	٧,٢٨
٣	+٢٧,٥٥	٠,٠٣٧	٤,٨٩	١,١٠	٠,٨٩	٦,٦١	٢,٥٤	-٢٠,٣٢	٠,٣٦	٤,١٨	٧,٢٤
٤	٣,٦٢	٣,١١	٥,٤٣	٣,٦١	١,٤٤	٤,٣٠	+٢٣,٠٢	٠,٧٩	٣,١١	+٢٧,٠٩	٥,٧١
٥	١,٦٥	٠,٦٠	٦,٣٧	+١٥,٩٧	+٢٢,٠٠	٠,٠٠	١,٨٤	٠,٠٣٦	,٠٢٠		٤,٨٨
٦	٣,٣٩	٤,١٢	٠,٦٤	٤,٣٢	+٣٢,١٣	+٧٢,٤٧	٣,٣٦	١,١٠	٢,٣٣		+٢٣,٥٦
٧	٣,٣٩	٢,٠٧	٣,١١	-٢٦,٨٢	-٣٩,٥٢	١,٠٦	٦,٨٢	٥,٤٨	١,٠٦		-١٥,٧٤

تابع جدول رقم ٢ .

كمبلي كاي تربيـــع

رقم	الجانب اللفظي						الجانب العملي		
البند	المعلومات	المتشابهات	الحساب	المفردات	الفهم	إعادة الأرقام	تكملة الصور	ترتيب الصور	رسوم المكعبات	تجميع المتاهات الأشياء
٨	٠,٨٣٤	٢,٣٠	٥,٩٥	٢٣,٧٠-	٧,٣٢	٠,٢٧٢	٠,٩٤	١,٨٤	٢,٣٥	٦,٠٦
٩	٤,٣٩	٢,٩٦	٥,١٦	١,٣٧	٦,٦٧	٠,٤٠٨	١,١٩	٥,٢٩	٤,٧٧	١,٩٣
١٠	٥,١٠	٢,٩٢	+٢٩,٣٢	٤٦,٤٢-	٥٤,٧٥-	٠,٠٠	٤,٧٤	٤,٥٢	٢,١١	
١١	٢,٠٥	٢,٢٨	٦,٧٨	٠,٠٧	١,٨١٠	٠,٠٠	٢,٧٦	٣,٦٤	٢,٣٦	
١٢	٧,٧١	+١٦,٣٤	٧,٥٨	+١٤,٥٧	٤,٩٠		٣,٧٦	١٨,٧٠-	١,٥١	
١٣	٣,٠٩	٠,٠٠	+١٢,٢١	٢٢,٦٩-	٠,٣٨		٥,٣٨		١,٦٢	
١٤	٢,١٨	٠,٠٠	٤٣,١٣-	١٤,٩٤-	٣,٤٣		١,٣٩			
١٥	+٢١,٤٢	٠,٠٠	٤,٤٨	١,٨٢	٧,٠٥		٢٠,٧٣-			
١٦	٧,٣٩	٠,٠٠	٠,٠٠	١٦,٩٨-	٣,٣٨		١,٥٠			
١٧	٠,٠٥	٠,٠٠	٠,٠٠	٢,٦٧	١٩,١٨-		٠,٣٢			
١٨	٥٧,٧-	٠,٠٠	٠,٠٠	+٢٦,٩٦	٠,٠٠		٠,٣٢			
١٩	٦,٦٦			٧,٧٤			٤,٣٩			
٢٠	٧,١٢			٦,١٢			١,٠٦			
٢١	٦,٨٨			٣,٢٩			١,٣٨			
٢٢	٤,٢٢			١,١٠			٠,٠٠			
٢٣	٢٨,٧٦-			٣,٠٧			٠,٨٠			
٢٤	١,٨٨			٣,٩٤						

تابع جدول رقم ٢ .

رقم	كميلي كاي تربيـــع										
	الجانب اللفظي						الجانب العملي				
البند	المعلومات	المتشابهات	الحساب	المفردات	الفهم	إعادة الأرقام	تكملة الصور	ترتيب الصور	رسوم المكعبات	تجميع الأشياء	المتاهات
٢٥	٤٤,٠١-				٤,٩٨						
٢٦	٣,١٢				٦,٥٧						
٢٧	٣٨,٨-				٦,٠١						
٢٨	٠.٨٥٧				١,٠٥						
٢٩	٠,٠٠٠				٢٦,٢٨+						
٣٠	٠,٠٠				٧,٣٦						
٣١					٥,٨٨						
٣٢					٢,٢٨						
٣٣					٥,١٦						
٣٤					١,٥٢						
٣٥					١,٦٨						

مناقشة النتائج

أظهرت نتائج الدراسة الحالية أن نسبة البنود المتحيزة في الصورة السعودية من اختبار وكسلر لذكاء الأطفال المعدل ليست عالية، إذ بلغت ١٨٪ من البنود البالغ عددها ١٩١ بندًا. ونظرًا لعدم وجود دراسات عن تحيز بنود اختبار وكسلر لذكاء الأطفال المعدل تبعًا لمتغير الجنس في المجتمع الأمريكي، فإن إمكان مقارنة الصورة السعودية من الاختبار بالصورة

جدول رقم ٣. أرقام البنود المتحيزة في الاختبارات الفرعية في الجوانب اللفظية والعملية من الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل.

	رقم البنود المتحيزة في						المجموع
	المعلومات	المتشابهات	الحساب	المفردات	الفهم	إعادة الأرقام	الكلي
	+٣	+٢	+١٠	+٥	+٥	+٦	
				-٧			
الاختبارات	+١٥	+١٢	+١٣	-٨	+٦		
	-١٨		-١٤	-١٠	-٧		
اللفظية	-٢٣			+١٢	-١٠		
	-٢٥			-١٣	-١٧		
	-٢٧			-١٤			
				-١٦			
				+١٨			
				+٢٩			
المجموع	٦	٢	٣	١٠	٥	١	٢٧

	رقم البند المتحيز في					
	تكملة الصور	ترتيب الصور	المكعبات	تجميع الأشياء	المتاهات	
الاختبارات	+٤	+١	بدون	+٤	+٦	٠
العملية	-١٥	-٣			-٧	
		-١٢				
المجموع	٢	٣	-	١	٢	٨
مجموع البنود المتحيزة في الاختبار الكلي						٣٥

الأصلية في هذا الجانب من جوانب التحيز غير ممكن في الوقت الحاضر. بيد أنه من خلال استعراض دراسات التحيز العرقي لهذا الاختبار في المجتمع الأمريكي، نجد أن نسبة البنود المتحيزة ضد الزنوج ٤٥٪، وضد ذوي الأصول المكسيكية ٤٣٪ في دراسة Sandoval [٤]. وبالرغم من أن نسبة البنود المتحيزة لعرق دون آخر في الدراسات الأمريكية أعلى من نسبة البنود المتحيزة في الدراسة الحالية، إلا أن تلك الدراسات لم تنظر إلى الاختبار نظرة تحفظ.

كذلك أظهرت النتائج الحالية أن اتجاه التحيز لم يكن لصالح فئة دون أخرى. فقد كان عدد البنود المتحيزة لصالح البنين ١٨ بندًا، أي ٥١٪ من البنود المتحيزة؛ أما عدد البنود المتحيزة لصالح البنات فقد كان ١٧ بندًا، أي ٤٩٪ من البنود المتحيزة. لذا فإن تأثير التحيز سوف لا ينعكس على الدرجة الكلية للاختبار نظرًا لتقارب عدد البنود المتحيزة لصالح البنين أو البنات. وللتأكد من مصداقية هذا التفسير فقد تم حساب معامل الارتباط الرتبي (سبيرمان) لمستويات صعوبة البنود للبنين والبنات على درجات الاختبار الكلي، وكان معامل الارتباط ٠,٩١. وتعني قيمة الارتباط هذه أن هناك تماثلاً كبيرًا في صعوبة البنود لكل من البنين والبنات. إذ إن Reynolds أشار إلى أن الاختبار يعتبر غير متحيز عندما تكون قيمة معامل الارتباط الرتبي لمستويات الصعوبة ٠,٩٠ أو أكثر [١٣، ص١٢٠].

بناء على نتائج الدراسة، فإنه يمكن القول بأن الصورة السعودية لاختبار وكسلر لذكاء الأطفال المعدل تعتبر غير متحيزة لأي من البنين أو البنات حين ينظر للدرجة الكلية على الاختبار. كما أن نسبة البنود المتحيزة في الصورة السعودية من اختبار وكسلر لذكاء الأطفال المعدل تعد أقل من نسبة البنود المتحيزة في الصورة الأمريكية. بيد أنه ينبغي النظر إلى المقارنة الحالية بكثير من التحفظ نظرا لكون دراسات التحيز للصورة الأمريكية اعتمدت على طرق مشكوك في سلامتها [٨]. ولكون دراسات التحيز لم تكن ضمن إطار متغير الجنس [١٨]، بل كانت لمتغير العرق.

وتعد الدراسة الحالية من أول الدراسات التي حاولت أن تدرس تحيز البنود لاختبار وكسلر بناء على متغير الجنس، ويمكن القول بأن الأداة غير متحيزة بشكل عام.

المراجـــع

[١] Rudner, L., P. Getson, and D. Kinight."Biased Item Detection Techniques." *Journal of Educational Statistics*, 5, No. 2 (1980), 213-33.

[٢] Shepard, L., G. Camilli, and M. Averill. "Comparison of Procedures for Detecting Test-Item Bias with Both Internal and External Ability Criteria." *Journal of Educational Statistics*, 6 (1981), 317-75.

[٣] Oakland, T., and D. Feigenbaum. "Multiple Sources of Test Bias on the WISC-R and the Bender-Gestalt Test." *Journal of Consulting and Clinical Psychology*, 47, No. 5 (1979), 968-74.

[٤] Sandoval, J. "The WISC-R and Internal Evidence of Test Bias with Minority Groups." *Journal of Consulting and Clinical Psychology*, 47, No. 5 (1979), 919-27.

[٥] Meile, F. "Cultural Bias in the WISC." *Intelligence*, 3, No. 2 (1979), 149-64.

[٦] Hunter, J. "A Critical Analysis of the Use of Item Means and Item-Test Correlations to Determine the Presence or Absence of Content Bias in Achievement Test Items." Paper presented at the National Institute of Education Conference on Test Bias, Annapolis, MD, 1975.

[٧] Cotter, D., and R. Berk. "Item Bias in the WISC-R Using Black, White, and Hispanic Learning Disabled Children." Paper Presented at the annual meeting of the American Educational Research Association, Los Angeles, 1981.

[٨] Camilli, G., and L. Shepard. "The Inadequacy of ANOVA for Detecting Test Bias." *Journal of Educational Statistics*, 12, No. 1 (1987), 87-99.

[٩] Mishra, S. "The WISC-R and Evidence of Item Bias for Native-American Navajos." *Psychology in the Schools*, 19, No. 4 (1982), 458-64.

[١٠] Reynolds, C., and S. Kaiser. "Test Bias in Psychological Assessment." In T. B. Gutkin and C.R. Reynolds, eds. *The Handbook of School Psychology*. New York: Wiley, 1990, pp. 487-525.

[١١] Jensen, A. *Bias in Mental Testing*. New York: Free Press, 1980.

[١٢] Sandoval, J. and M. Miille. "Accuracy of Judgements of WISC-R Item Difficulty for Minority Groups." *Journal of Consulting and Clinical Psychology*, 48, No. 2 (1979), 249-53.

[١٣] Reynolds, C. "Methods for Detecting Construct and Predictive Bias." In R. Berk, ed. *Handbook of Methods for Detecting Test Bias*. Baltimore, MD: The Johns Hopkins University Press, 1982, pp. 199-227.

[١٤] Green, D. "What Does It Mean to Say a Test Is Biased?" *Educational and Urban Society*, 8 (1975), 33-52.

[١٥] Shepard, L. "Definitions of Bias." In R.A. Berk, ed. *Handbook of Methods for Detecting Test Bias*. Baltimore, MD: Johns Hopkins University Press, 1982, pp. 9-30.

[١٦] Berk, R. A. *Handbook of Methods for Detecting Test Bias*. Baltimore, MD: Johns Hopkins University Press, 1982.

[١٧] Wilson, V., R. Nolan, and C. Reynolds. "Race and Gender Effects on Item Functioning on the Kaufman Assessment Battery for Children." *Journal of School Psychology*, 27 (1989), 289-97.

[١٨] Hale, R., and A. Potok. "Sexual Bias in the WISC-R." *Journal of Consulting and Clinical Psychology*, 48, No. 6 (1980), 776.

[١٩] Scheuneman, J. "A New Methods of Assessing Bias in Test Items." Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., 1975.

[٢٠] القاطعي، عبدالله. «دراسة مقارنة لبعض طرق تحيز بنود الاختبار.» قيد النشر.

[٢١] النافع، عبدالله، وعبدالله القاطعي، والجوهرة السليم. «إعداد اختبارات ومقاييس للتعرف على الموهوبين والكشف عنهم: (القسم ا) صدق وثبات اختبار وكسلر لذكاء الأطفال المعدل.» الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية، ١٤١١هـ، ص ص ٥ ـ ٢٥.

[٢٢] Ironson, G. H. "Use of Chi-Square and Latent Trait Approaches for Detecting Item Bias." In R.A. Berk, ed. *Handbook of Methods for Detecing Test Bias*. Baltimore, MD: Johns Hopkins University Press, 1982.

Item Bias of the Wechsler Intelligence Scale for Children-Revised (Saudi Version) According to Respondents' Sex

Abdullah A. Qataee
Assistant Professor, Department of Psychology,
College of Education, King Saud University,
Riyadh, Saudi Arabia

Abstract. Investigation of test bias is becoming more important in recent years due to social pressure and ethical concerns about the soundness of selection. This study was intended to investigate the extent of sex bias in the items of the Wechsler Intelligence Scale for Children-Revised WISC-R (Saudi Version). The records of the normative sample were evaluated. The results showed minimal sex bias in the items of the WISC-R (Saudi Version). The implications of the results were discussed.