

البحوث والدراسات

معايير كويتية لاختبار "ريزن" المصفوفات المتدرجة القياسية

أ.د. صلاح أحمد مراد
كلية التربية - جامعة الكويت
smorad44@yahoo.com

أ.د. أحمد محمد عبد الخالق
كلية الآداب - جامعة الإسكندرية
aabdel-khalek@hotmail.com

الملخص:

حظيت المصفوفات المتدرجة القياسية - بوصفها اختباراً للذكاء - باهتمام كبير عبر قارات العالم الخمس، ومنها معظم البلاد العربية وبخاصة دولة الكويت، وذلك نظراً لمزايا عدة في هذا الاختبار. وقد هدفت هذه الدراسة إلى تقديم معايير كويتية لهذا الاختبار، اعتمدت على: (1) استخدام عينة كبيرة الحجم (8472 فرداً) من الجنسين، (2) شمول مدى عمري واسع (من 7 إلى 17 سنة)، (3) اختيار العينة من جميع محافظات دولة الكويت، (4) استخدام النسخة الكاملة للاختبار (60 بنداً)، (5) استخراج المعايير العمرية المئينية المقابلة للدرجات الخام لكل ستة أشهر، (6) حساب معاملات الثبات لكل فئة عمرية، (7) اختبار الفروق بين الجنسين وبين الفئات العمرية. ومن المأمول فيه أن يكون لهذه المعايير تطبيقات نظرية وعملية في دولة الكويت.

Kuwaiti Norms for Raven's Standard Progressive Matrices

Ahmed Mohamed Abdel-Khalek
College of Arts
Alexandria University

Salah Ahmed Mourad
College of Education
Kuwait University

Abstract

The Raven's Standard Progressive Matrices (SPM) has a widespread interest through the five continents of the world, including most of the Arab countries, especially the State of Kuwait because of its several advantages. This study aimed to introduce Kuwaiti normative data on the SPM. This standardization depends on: (1) using a large sample size of both sexes (N= 8472), (2) a wide age range for the sample (from 7 to 17 years.), (3) selecting the sample from all of the Kuwaiti districts, (4) using the complete version of the SPM 60 items, (5) producing percentile norms (each five percentile scores) for each age group every six months, (6) computing reliability for each age group, (7) investigating sex and age group differences. It is hoped that these normative data would have theoretical and practical applications in the Kuwaiti context.

مقدمة:

كتب "ريتشارد لن" (Lynn, 2008) عن المنحنى الاعتدالي عند مناقشته لعوامل السلالة والذكاء وعدم المساواة عبر العالم، وخلص فيه إلى أن الفروق بين الأجناس في الذكاء قادرة على تفسير الفروق في جوانب عدة في الحياة الإنسانية، من بينها التحصيل الدراسي، ومعدل الدخل، والمستوى الاجتماعي الاقتصادي، والجريمة، والعمر المتوقع، ووفيات الأطفال، والخصوبة... وغير ذلك من الظواهر الاجتماعية. كما ذكر "لن، وفانهانن" (Lynn & Vanhanen, 2012) أن الذكاء يمكن أن يكون مفهوماً بنائياً وموحداً ومفسراً في مجال العلوم الاجتماعية، حيث يعد الذكاء محددًا لكثير من الظواهر الاجتماعية المهمة، بما فيها المتغيرات السابق ذكرها أعلاه، فضلاً عن الصحة، والتطور الاقتصادي. وقد أثبت كثير من البحوث، القيمة التفسيرية للذكاء للفروق بين الأفراد، وربما إلى تفسير الفروق بين المجموعات والأمم.

وقد عرض (Lynn & Vanhanen, 2012) معدلات الذكاء في معظم دول العالم (قراءة 197 دولة من مختلف القارات)، ودرسا الارتباطات في هذه الدول بين الذكاء وكل من: التحصيل الدراسي والمعرفي، والدخل القومي، ومعدل الفقر، والمنظمات السياسية، والصحة، والخصوبة، والماء النظيف، ومعدل الجريمة، والدين، والسعادة... وغيرها، وتشير نتائجهما إلى تدعيم الاستنتاج السابق الإشارة إليه. وحيث إن اختبار المصفوفات المتدرجة يستخدم الأشكال والرسوم في بنوده مما يجعله غير متحيز ثقافياً، لذلك يمكن تطبيقه على أي مجموعة في أي دولة دون حاجة إلى ترجمة، وقد أدى هذا إلى استخدامه في كثير من البحوث، وهو الاختبار نفسه الذي تضطلع به الدراسة الحالية.

وتهدف هذه الدراسة إلى تقديم معايير كويتية لاختبار مهم من اختبارات الذكاء، وهو المصفوفات المتدرجة القياسية (SPM) Standard Progressive Matrices، بسبب مضي فترة طويلة على المعايير المتوافرة للاختبار، والتي أُعدت عام 1999. والمعايير هي طرائق لتفسير الدرجات الخام التي نحصل عليها من الاختبارات، لأن الدرجات الخام لا معنى لها، ولا نستطيع تفسيرها إلا إذا توافرت معايير للاختبار لمقارنتها ببيانات عينة التقنين. والمعايير هي معدلات مستنتجة من بيانات عينة معيارية ممثلة للمجتمع الذي سوف يستخدم الاختبار، ومن ثم نستطيع مقارنة الدرجات الخام بالمعدلات (المعايير) الناجمة عن عينة التقنين. وتوجد أنواع مختلفة من المعايير منها المعايير المثنية التي استخدمها ريفن، ومعايير العمر العقلي عند بينيه، ومعايير نسبة الذكاء الانحرافية التي استخدمها وكسلر في مقاييسه الثلاثة لقياس الذكاء... وغيرها من أنواع المعايير (مراد، سليمان، 2012).

تاريخ إعداد اختبار المصفوفات:

توجد ثلاث روايات في هذا الصدد، لكل من "جنسن، وجون ريفن، وكلاين". يسرد الرواية الأولى "جنسن" (Jensen, 1998: 36-37)، فيذكر أن "سبيرمان" Spearman وضع اختباراً مكوناً من مختلف الأشكال الهندسية المألوفة في كل ثقافة، بحيث تتطلب الإجابة عن كل بند استنتاج العلاقات والمتعلقات Induction of relations and

correlates، واتصفت هذه الأشكال بالتجريد Abstractness؛ إذ تكونت من خطوط مستقيمة ومنحنية، ومثلثات ودوائر ومربعات وما شابهها، حيث لا تمثل هذه الأشكال أية موضوعات حقيقية أو ملموسة مثل: الحيوانات والنباتات والأثاث أو المركبات وأدوات النقل. وُسِمى هذا النوع من اختبار العلاقات المكانية "علاقات المصفوفة" Matrix relations. لأن كل بند كان يتكون من ثماني لوحات من الأشكال ولوحة فارغة، وقد نظمت جميعاً على شكل مصفوفة قوامها 3×3.

وتعتمد المصفوفات هنا على قدرة الفرد على التوصل إلى القاعدة التي يمكن أن تحدد الخصائص المميزة للشكل الذي يجب أن يلائم الفراغ الناقص، حتى يكمل النمط المنطقي للمصفوفة كلها. ويتم اختيار الشكل الصحيح من بين مجموعة من ستة اختيارات أو ثمانية تقدم أسفل المصفوفة، وهي مجموعة من البدائل غير الصحيحة (مشتتات) وبديل واحد صحيح مثل بدائل أسئلة الاختيار من متعدد.

وقد تطور اختبار "سبيرمان" أكثر بواسطة أحد تلاميذه وهو "جون ريفن" J. Raven وعالم الوراثة "ليونيل بينروز" Penrose، وهو المسمى الآن بمصفوفات "ريفن" المتدرجة (PM) Raven's Progressive Matrices، وتسمى متدرجة لأن البنود تزداد في الصعوبة بشكل متنسق لدى إجابة المبحوث عنها، وذلك اعتماداً على عدد الأسس أو المبادئ، التي تدخل بشكل تلقائي في العلاقات والمتعلقات التي يجب أن تستنتج وصولاً إلى الإجابة الصحيحة (Jensen, 1998: 36).

ويسرد الرواية الثانية عن تاريخ تأليف هذا الاختبار "جون جون ريفن" (ابن مؤلف الاختبار) (Raven, 2008b) فيذكر أن (والده) "جون ريفن" قد وضع اختبار المصفوفات المتدرجة، لأنه كان يعمل مع عالم من علماء الوراثة هو "ليونيل بينروز" في دراسة عن الأصول الوراثية والبيئية للتأخر العقلي، وكان ذلك يتطلب أن كلاً من الراشدين والأطفال يجب أن تجرى عليهم الاختبارات، وكان معظم هؤلاء الأفراد من الأميين، ومن ثم فقد كانوا غير قادرين على تتبع التعليمات المكتوبة، وكان من الضروري أن يتم اختبارهم في المنازل، والمدارس، وأماكن العمل، التي كانت غاصة بالضوضاء، ومن ثم جعل سؤالهم صعباً، هذا فضلاً عن بعض المشكلات في اختبارات الذكاء المتاحة وقتها. ومن ثم، بدأ "ريفن" في تأليف اختبار يمكن أن يكون سهلاً في تطبيقه، ويعتمد على أساس نظري قوي، ويمكن تفسيره دون الحاجة إلى حسابات معقدة.

وكان "ريفن" من تلاميذ "سبيرمان"، ومن المعروف أن "سبيرمان" كان أول من لاحظ اتجاه الاختبارات التي يفترض أنها قدرات مستقلة إلى أن ترتبط معاً ارتباطاً مرتفعاً، فسره بعض الباحثين بافتراض عامل واحد كامن، سماه كثير من الناس "القدرة المعرفية العامة"، وهو ما أعطاه "سبيرمان" اسم "g" أو "ع"؛ إذ تجنب عامداً استخدام كلمة "ذكاء" لوصف هذا العامل، نظراً لأن المصطلح الأخير يشير إلى معانٍ مختلفة جداً عند مختلف الناس. وقد اعتقد "سبيرمان" أن "ع" أو "g" تتكون من قدرتين مختلفتين تعملان معاً عن قرب، وهما "القدرة على تكوين معنى" (Induction (formation meaning)) و"القدرة على الإنتاج" Reproductive ability، وهي القدرة على تكوين معلومات صريحة عن المهارات المتعلمة (Raven, 2008b).

وبالإضافة إلى ما كتبه "جون ريفن" الابن في الفقرة السابقة، فقد ذكر في اتصال بالبريد الإلكتروني (Raven, Personal Communication, 21 March, 2013) أن "بينروز" كان يبحث عن باحث مساعد، فخطب "سبيرمان"، وكان "ريفن" يعمل مع الأخير، فأرسل "سبيرمان" معه خطاباً، وأصبح "ريفن" هذا المساعد، وفي هذه الأثناء، أدرك "ريفن" حدود الاختبارات من نوع "بينيه / تيرمان"، وصمم "ريفن" على تطوير اختبار يعتمد على مبادئ "سبيرمان"، ومع ذلك فلم يستخدم هذا الاختبار في بحوث "بينروز" مطلقاً.

والرواية الثالثة لعالم النفس البريطاني الراحل "بول كلاين" المتخصص في القياس النفسي، إذ يؤكد أن بنود المصفوفات قد تم تأليفها بوساطة كل من "بينروز، وريفن" (Kline, 1988: 204). وبالنظر إلى هذه الروايات الثلاث، تجب التفريق بين فكرة المصفوفات من الناحية النظرية (وقد تعزى هذه في المقام الأول إلى "سبيرمان")، وتأليف أشكال المصفوفات من الناحية العملية، ويبدو أن "سبيرمان" وضع نماذج لهذه الأشكال، ثم طورها "بينروز، وريفن". لكن الشائع الآن في مراجع القياس النفسي أن المصفوفات المتدرجة من تأليف "جون ريفن".

وللمصفوفات المتدرجة صيغ ثلاث، هي القياسية Standard، والموتة، والمتقدمة Advanced. والصيغة القياسية هي المستخدمة في هذه الدراسة، وتطبق فردياً أو جمعياً، وتصلح لمدى عمري واسع: من 6 إلى 80 سنة. وتشتمل المصفوفات المتدرجة على 60 بنداً، مصنفة إلى خمس مجموعات، في كل منها 12 بنداً تتدرج في الصعوبة، وكل مصفوفة بها جزء مفقود، ويطلب من المبحوث أن يختار الجزء الناقص من بين مجموعة من البدائل (6 أو 8). ولهذا الاختبار معاملات ثبات وصدق مرتفعة في صيغته الأصلية (انظر: Court & Raven, 1995, Gregory, 1992: 230-231; Kline, 2000: 463; Murphy & Davidshofer, 1998: 318).

المصفوفات المتدرجة مقياس للعامل العام:

تقيس المصفوفات المتدرجة القدرة على الاستنتاج Inductive ability أو العامل العام (g) General factor، الذي أسماه "سبيرمان" القدرة على إدراك الترابطات (Richardson, 2000: 33). وتؤكد كثير من البحوث أن المصفوفات المتدرجة أفضل اختبار متاح لقياس العامل العام (Kaplan & Saccuzzo, 1997: 359). وتتشعب بالعامل العام قرابة 0.88، ومن المحتمل أن يكون أهم ما يميز المصفوفات المتدرجة تشعبها المنخفض جداً بأي عامل آخر خلاف العامل العام (Jensen, 1998: 38). كما تعد المصفوفات أفضل مقياس للقدرة على الاستنتاج المجرد أو غير اللفظي، وهذا هو جوهر الذكاء المرن (السائل) Fluid intelligence الذي أشار إليه كل من: هورن (Horn, 1985)، ولين (lynn, 2009).

استخدامات المصفوفات ومزاياها:

تستخدم المصفوفات المتدرجة - بوصفها اختباراً للذكاء - استخدامات واسعة: مع

الأطفال، والمراهقين، والراشدين، والمحرومين ثقافياً، والمعاقين لغوياً، وتعد أداة صالحة للبحوث في مجال القدرة العقلية وتدهورها، وتعد المصفوفات علامة هادية أو مؤشراً للعامل العام في البحوث التحليلية العملية، وهي أقل اختبارات الذكاء تأثراً بالثقافة، كما أنها تتسم بالمرونة في تطبيقها فردياً أو جمعياً، وتعليماتها بسيطة، ويمكن أن تطبق من دون استخدام اللغة، ويمكن أن تكون موقوتة Timed أو غير موقوتة.

واختبار المصفوفات المتتابعة مهم في الاستخدام مع مختلف الجماعات الثقافية، ولكن من المعروف الآن أن تأليف اختبار لقياس القدرات المعرفية مستقلاً عن الخبرة أمر مستحيل تقريباً، وقد بينت البحوث التي أجريت في الدول النامية أن الفروق في الأمية وعدد سنوات التعليم أكثر أهمية من اللغة، والدولة، والسلالة أو العنصر بوصفها محددات للفروق الثقافية في درجات اختبارات الذكاء (Frisby, 1999).

الدراسات السابقة:

حظيت المصفوفات المتدرجة بالاهتمام الكبير على مستوى غالبية الدول العربية، ونعرض فيما يلي ما أتيج من هذه الدراسات، ففي دراسة مصرية على طلاب الجامعة وجد عبد الخالق أن اختبار المصفوفات المتدرجة يتسم بثبات إعادة تطبيق مرتفع، واستخرج من المجموعات الخمس للاختبار عاملاً عاماً مرتفع التشبعات يشير إلى اتساق داخلي مرتفع، وأسفر التحليل العاملي للدرجة الكلية في المصفوفات وأربعة من المقاييس الفرعية لمقياس "ثرستون" القدرات العقلية الأولية PMA عن عامل عام قوي، تشبعت به المصفوفات بمقدار 0.77، وهو ما يشير إلى صدق تلازمي مرتفع (Abdel-Khalek, 1988). وقن عبد الحليم، والشريف (2001) اختبار المصفوفات المتدرجة على التلاميذ الصم من 6-8 سنوات في مصر.

وهدفت دراسة نور الدين (2002) إلى بيان مدى فاعلية الاختبار المحوسب (Computerized adaptive testing (CAT في تقدير قدرة الأفراد، وتحديد الخصائص السيكومترية للقياس مقارنة بأساليب أخرى للاختبار، وذلك من خلال أداء المبحوثين (أكثر من 5000 طالب وطالبة في الجامعة) في اختباري "ريفن" للمصفوفات المتدرجة. وأسفرت النتائج عن أن الاختبار التواؤمي المحوسب أكثر فاعلية من الاختبار الخطي المحوسب، حيث تمكن من الوصول إلى دقة أعلى من خلال تطبيق (25) مفردة في المتوسط، وحتى عند تطبيق خمس مفردات فقط، مع زيادة دقة القياس.

كما أجرى أبو حطب وزملاؤه (1977) تقنياً للاختبار في المملكة العربية السعودية، وأسفرت دراسة الجميسان (1998) على عينات سعودية من طلاب الجامعة عن ارتباط بين الدرجات على هذا الاختبار والمعدل التراكمي. كما قننت آل ثاني (2002) الاختبار نفسه على طلاب في المرحلة الابتدائية بمدينة الدوحة في دولة قطر. وفي البيئة العُمانية قن كاظم، وزملاؤه (2008) المصفوفات الملونة، في حين قن يحيى، وإبراهيم، وجلال (2003) المصفوفات القياسية.

وقورنت بيانات التقنين المستخرجة من تطبيق اختبار المصفوفات على عينات كبيرة الحجم من كل من: المملكة العربية السعودية، وعمان، والأردن؛ بالمعايير البريطانية،

فحصلت هذه العينات الثلاث على نسب ذكاء تساوي 78، 85، 86 على التوالي (Abdel-Khalek & Lynn, 2008b, 2009, Lynn & Abdel-Khalek, 2009).

وفيما يختص بدولة الكويت فقد نشر القرشي (1987) اختبار المصفوفات المتتابعة الملونة ودليل الاستخدام، ونشر دراسة أخرى عن ثبات الاختبار وصدقه على الأطفال الكويتيين (القرشي، 1988). أما اختبار المصفوفات المتدرجة القياسية SPM فقد قننه في الكويت أبو علام (1981)، ثم عوض (1999)، ولكن المؤلفة الأخيرة حذفت 12 بنداً لأسباب عدة.

وفي وقت أحدث، طبق عبد الخالق (Abdel-Khalek, 2005) المصفوفات المتدرجة القياسية على 6529 طفلاً ومراهقاً كويتياً، تراوحت أعمارهم بين 8 و15 سنة، واستخرج معاملات ثبات إعادة التطبيق (ن=968)؛ إذ تراوحت بين 0.69 و0.85، في حين تراوحت معاملات ثبات ألفا "كرونباخ" بين 0.88 و0.93، إشارة إلى الثبات عبر الزمن والاتساق الداخلي بين المقبول والمرتفع، وتراوحت تشبعات مجموعات المصفوفات الخمس بالعامل العام بين 0.73 و0.89. كما نشرت على العينة السابقة نفسها معايير كويتية (عبد الخالق، 2006)، وقورنت المعايير الكويتية بالبريطانية (Abdel-Khalek & Raven, 2006, 2008).

وباستخدام عينات كويتية ظهرت فروق صغيرة بين الجنسين في اختبار المصفوفات (البنات أعلى بمقدار 1.2 نقطة في نسبة الذكاء)، كما حصلت البنات على متوسطات أعلى من الأولاد في الاختبارات التحصيلية الخاصة بالفهم اللفظي، واللغة الإنجليزية والرياضيات، ولا تؤكد هذه النتائج الاستنتاج القائل بأن الإناث في المجتمعات التقليدية يحصلن على درجات أقل في الذكاء والاختبارات التحصيلية. وفي دراسة أخرى على عينة كويتية، لم تتأكد نظرية "زاجونك" Zajonc من أن هناك علاقة بين الذكاء وكل من حجم الأسرة ورتبة المولد (Abdel-Khalek & Lynn, 2006, 2008a).

واستخدم الاختبار في دولة الكويت استخدامات كثيرة في المجال العملي. ويلاحظ من هذه الدراسات السابقة ما يلي:

- أ - قننت المصفوفات المتدرجة في كثير من البلاد العربية.
- ب - استخدمت معظم الدراسات عدداً كبيراً من المبحوثين وممثلاً لمجتمع الدراسة.
- ج - استخدمت بعض الدراسات جزءاً من الاختبار وألغت الآخر، وفي ذلك إخلال بشروط المقارنة بين النتائج المحلية وغيرها من النتائج في دول أخرى.

ضرورة تحديث معايير الاختبارات:

المعايير Norms قيم أو أرقام أو مستويات ممثلة لمجموعة ما، وتستخدم لمقارنة الحالات الفردية أو لمجموعات بها. ومن نافلة القول أن نذكر أن المعايير يجب أن تكون محلية، مستخرجة من عينة مماثلة للعينات أو الأفراد الذين سيتم مقارنتهم بها في المجتمع نفسه، فلا يصح مثلاً أن نقارن أداء طفل كويتي بمعايير بريطانية.

ومن الضروري كذلك أن تستخدم معايير حديثة، فينبه "ريفن" (Raven, 2008a) إلى خطورة الخطأ الناتج عن استخدام معايير قديمة لاختبارات الذكاء، ويضرب

لذلك مثلاً واقعياً، بمبجوث عمره خمسون عاماً، اختبر عام 1942، وصنفته درجته على أنه يقع في المئين 95 قياساً إلى معايير عام 1942، ولكن هذه الدرجة تصنفه في المئين 25 إذا حكم عليه بمعايير اليوم. ومن ثم يجب التنبيه إلى الزمن الذي تحددت فيه المعايير، حيث تتغير من زمن إلى آخر. هذا فضلاً عن التغير تبعاً للثقافة، بالإضافة إلى ما يعرف "بأثر فلن" Flynn effect.

فقد بين "جيمس فلين" (Flynn, 1987, 2000) أن متوسط نسبة الذكاء لشخص متوسط في العشرين من عمره في عام 1980، أعلى في نسبة ذكائه بمقدار 15 نقطة عن نظيره في عام 1940، وأن الذكاء يستمر في الزيادة بمقدار (0.33) نقطة كل عام، وأن هذه الفروق عبر الأجيال في متوسط نسبة الذكاء تظهر بصورة أكبر في اختبارات مثل المصفوفات المتدرجة، أكثر من اختبارات "وكسلر، وبينيه"، وخلص "فلين" إلى أن هذه الزيادة المشاهدة عبر الأجيال في متوسط درجات اختبارات الذكاء ترجع إلى البيئة أكثر من العوامل الوراثية، ولكنها لا تعزى فقط إلى التحسن في الدراسة الشكلية في المدارس، وهناك عوامل أخرى يحتمل تأثيرها، ومن بينها التحصيل الدراسي للوالدين، والاهتمام الزائد من الآباء بأطفالهم، وتحسن المستوى الاجتماعي الاقتصادي، والتغذية الجيدة، وتناقص أمراض الأطفال، وزيادة تعقد المجتمع التكنولوجي (Aiken & Groth-Marnat, 2006: 162).

ويرى "جرينفيلد" (Greenfield, 1998) أن غالبية الزيادة في نسبة الذكاء التي بينها "فلين" ترجع إلى التأثيرات البصرية، نتيجة التليفزيون، والحاسبات، وألعاب الفيديو، ومختلف الأجهزة التكنولوجية. واتضح كذلك أن سوء التغذية الشديد، ونقص اليود Iodine والحديد، وغيرهما من المواد الغذائية ترتبط بانخفاض نسب الذكاء، فضلاً عن قصر القامة، التي تناقصت بوضوح خلال هذا القرن.

أهداف الدراسة:

تتاح لاختبار المصفوفات المتدرجة القياسية بيانات معيارية كويتية استخرجت في دراسات سابقة، وقد استخدم غالبية هذه الدراسات عينات كبيرة الحجم. وتهدف الدراسة الحالية إلى إعداد معايير كويتية لاختبار المصفوفات المتدرجة القياسية (SPM)، وذلك لعدم تحديث معايير هذا الاختبار في الكويت منذ فترة طويلة حيث قام بتقنيه في الكويت أبو علام (1981)، وأعيد التقنين مرة أخرى (عوض، 1999) وفي هذا التقنين الأخير تم فيها حذف 12 بنداً من الاختبار. ونظراً للاستخدام المستمر لهذا الاختبار ومضي خمسة عشر عاماً على المعايير المتاحة أصبح من الضروري توفير معايير حديثة للاختبار.

كما تهدف الدراسة إلى فحص الفروق بين الأعمار الزمنية، والفروق بين الجنسين لكل مجموعة من مجموعات الأعمار الزمنية.

وتستخدم الدراسة الحالية عينة كبيرة الحجم (ن=8493 فرداً) تتقارب فيها أعداد الأولاد والبنات (4245 و 4248 على التوالي)، وتغطي مدى عمرياً كبيراً (من 6.5 إلى 17 سنة فأكثر). وطبق عليهم النسخة الكاملة من اختبار المصفوفات (60 بنداً).

أهمية الدراسة:

من الملاحظ أن اختبار "ريفن" للمصفوفات المتدرجة القياسية قد حظي بالاهتمام في مختلف الدول العربية، ويتضح ذلك أيضاً في دولة الكويت، حيث قنن المقياس عدة مرات، واستخدم استخدامات عدة في البحوث الأساسية والتطبيقية، ومن بينها استكشاف حالات التأخر العقلي والتفوق العقلي وغير ذلك، ولكن هذه الدراسة تتسم بعدة مزايا كما أسلفنا، والمتوقع أن تكون هذه المعايير ذات فائدة كبيرة في البحوث، وفي تشخيص التأخر والتفوق العقلي، فمن السهل لمستخدم الاختبار أن يستخرج الدرجة المئينية المقابلة للدرجة الخام لفرد أو مجموعة ما.

إجراءات الدراسة:

عينة التقنين:

تم تطبيق اختبار "ريفن" المصفوفات المتدرجة القياسية (60 بنداً). على عينة حجمها 8493 (4245 ذكور، 4248 إناث) من أعمار زمنية مختلفة تراوحت بين 6 سنوات ونصف وحتى 19 سنة، وتمثل محافظات الكويت الست. وقد تم تصنيف العينة إلى (22) مجموعة عمرية كل فئة تمثل ستة أشهر، ويوضح الجدولان (1، 2) توصيف العينة وفقاً للنوع والمحافظات والأعمار الزمنية.

جدول (1) توصيف العينة وفقاً للنوع والمنطقة

المجموع	النوع		المنطقة
	إناث	ذكور	
1765	829	936	العاصمة
1155	622	533	حولي
940	467	473	الفروانية
949	437	512	مبارك الكبير
934	497	437	الأحمدي
845	422	423	الجهراء
1905	974	931	غير محدد المنطقة
8493	4248	4245	المجموع الكلي

جدول (2) توصيف العينة وفقاً للنوع والمجموعات العمرية

المجموع	النوع		فئة العمر	المجموع	النوع		فئة العمر
	إناث	ذكور			إناث	ذكور	
403	194	209	12 سنة	21	5	16	6½ سنة
393	210	183	12½ سنة	358	203	155	7 سنوات
476	236	240	13 سنة	504	242	262	7½ سنة
411	177	234	13½ سنة	259	139	120	8 سنوات
389	195	194	14 سنة	415	211	204	8½ سنة
444	228	216	14½ سنة	388	204	184	9 سنوات
420	202	218	15 سنة	445	237	208	9½ سنة
385	196	189	15½ سنة	328	149	179	10 سنوات

492	197	295	16 سنة	418	194	224	10½ سنة
421	237	184	16½ سنة	438	210	228	11 سنة
266	174	92	17 سنة فأكثر	419	208	211	11½ سنة
8493	4248	4245	المجموع الكلي				

وقد تم تصنيف العينة إلى فئات عمرية كل ستة أشهر من 7 سنوات وحتى 17 سنة فأكثر (وهي 21 فئة عمرية) عند تحليل البيانات وحساب معاملات الثبات والفروق بين الجنسين والفروق بين الأعمار الزمنية، وكذلك عند حساب المعايير المئينية للدرجات الخام، وقد تم حذف مجموعة الأعمار 6.5 لصغر حجم هذه المجموعة.

الطريقة المستخدمة في تصنيف العينة إلى فئات عمرية هو تبويب الأعمار الزمنية بالشهور إلى فئات طول كل منها ستة أشهر، فعلى سبيل المثال الفئة العمرية 7 سنوات تشمل الأعمار من ست سنوات وعشرة أشهر إلى سبع سنوات وثلاثة أشهر، وكذلك الفئة العمرية 7.5 سنوات تشمل الأعمار الزمنية من سبع سنوات وأربعة أشهر إلى سبع سنوات وتسعة أشهر، وهكذا في بقية الفئات العمرية.

اختبار المصفوفات المتدرجة القياسية:

استخدمت الصيغة الأصلية للاختبار بعد الحصول على تصريح كتابي من صاحب حقوق الطبع، وأعدت صيغة عربية تختلف عن الإنجليزية في أن النسخة العربية للاختبار قد عكست فيها المصفوفة الرئيسية والبدائل الستة (أو الثمانية) من اليسار إلى اليمين في الصفحة نفسها (كصورة مرآة)، مع احتفاظ كل البدائل بأرقامها الأصلية، ومن ثم فإن كتيب الاختبار العربي تتابع فيه بنود الاختبار من اليمين إلى اليسار لتتبع الاستخدام المؤلف في اللغة العربية.

وتشتمل المصفوفات المتدرجة القياسية على خمس مجموعات، يضم كل منها 12 بنداً، فيكون المجموع 60 بنداً. ويتكون كل بند من نمط أو شكل قطع منه جزء، يطلب من المبحوث تكملته من أحد البدائل المرسومة أسفله (ستة بدائل في المجموعتين أ، ب، وثمانية في المجموعات ج، د، هـ). وقد استخدمت في هذه الدراسة الصيغة الكاملة (60 بنداً). (انظر: (Court & Raven, 1995; Raven, 1991; Raven, Raven & Court, 1998).

إجراءات التطبيق:

طبقت المصفوفات المتدرجة القياسية على الطلاب عن طريق مجموعة من الاختصاصيين النفسيين والاجتماعيين الأكفاء (وهم مساعدا الباحثين) الذين تلقوا تدريباً جيداً على تطبيق الاختبار. وقد طبق الاختبار على عينة الذكور الباحثين الذكور، في حين طبق الاختبار على عينة الإناث مجموعة من الأخصائيات (مساعداً الباحثين). وكان يقوم بتطبيق الاختبار - في كل فصل دراسي - باحث ومساعد له، مع وجود مدرس الفصل الذي كان دوره يقتصر على المحافظة على الانضباط. وكان التطبيق جمعياً على الفصل الدراسي بأكمله (من 25 - 30 تلميذاً)، وأعطى الاختبار للتلاميذ دون حدود زمنية. بعد ذلك استبعدت أوراق غير الكويتيين، كما استبعدت أوراق الإجابة غير المكتملة، وألقيت تعليمات

الاختبار على التلاميذ شفويًا.

نتائج الدراسة:

تم حساب معاملات الثبات لدرجات اختبار ريفن لكل فئة عمرية وللعيينة الكلية، كما تم اختبار الفروق بين الجنسين، وبين الفئات العمرية تمهيداً لحساب المعايير المئينية لدرجات الاختبار.

أولاً - نتائج معاملات ثبات اختبار ريفن:

تم حساب معاملات ثبات اختبار ريفن وفقاً للأعمار الزمنية (21 فئة عمرية) طول كل منها ستة أشهر (من 7-17 سنة فأكثر) وقد تم استخدام معادلة ألفا كرونباخ في حساب معاملات الثبات. ويتضمن جدول رقم (3) المتوسطات الحسابية والانحرافات المعيارية ومعاملات الثبات (ألفا كرونباخ) لكل فئة عمرية. حيث يتضح من الجدول أن معاملات الثبات تتراوح بين 0.78 إلى 0.92 للفئات العمرية، 0.94 للعيينة الكلية وهي معاملات ثبات مرتفعة.

جدول (3) المتوسطات والانحرافات المعيارية ومعاملات الثبات لكل فئة عمرية

معامل الثبات	الانحراف المعياري	المتوسط	الفئة العمرية	معامل الثبات	الانحراف المعياري	المتوسط	الفئة العمرية
0.90	8.22	32.38	12½ سنة	0.78	5.67	14.63	7 سنوات
0.88	7.21	33.97	13 سنة	0.82	6.23	15.64	7½ سنة
0.88	7.24	34.80	13½ سنة	0.89	8.30	18.32	8 سنوات
0.88	7.09	35.37	14 سنة	0.90	8.56	19.19	8½ سنة
0.87	6.46	37.38	14½ سنة	0.91	9.04	21.86	9 سنوات
0.85	6.03	37.44	15 سنة	0.92	9.72	23.58	9½ سنة
0.89	6.90	37.58	15½ سنة	0.92	9.36	24.11	10 سنوات
0.87	6.48	37.63	16 سنة	0.91	8.90	26.87	10½ سنة
0.84	5.57	38.56	16½ سنة	0.91	9.10	28.24	11 سنة
0.83	5.65	38.34	17 سنة فأكثر	0.90	8.17	30.85	11½ سنة
0.94			العيينة الكلية	0.89	7.75	31.79	12 سنة

الفروق بين الجنسين:

تم حساب الفروق بين متوسطي الذكور والإناث لكل فئة عمرية، ويوضح الجدول رقم (4) نتائج الفروق بين الجنسين.

جدول (4) المتوسطات والانحرافات المعيارية وقيم (ت) للفروق بين الجنسين لكل فئة عمرية

حجم التأثير	مستوى الدلالة	قيمة (ت)	إناث			ذكور			الفئة العمرية
			ع	م	ن	ع	م	ن	
-	غير دالة	1.89	5.54	15.09	203	5.80	14.02	155	7 سنوات
-	غير دالة	1.70	6.37	16.15	242	6.09	15.19	262	7½ سنة
-	غير دالة	0.15	8.30	18.25	139	8.33	18.40	120	8 سنوات
-	غير دالة	1.59	8.56	19.84	211	8.52	18.51	204	8½ سنة

0.232	0.023	2.28	8.97	22.86	204	9.01	20.76	184	9 سنوات
0.204	0.012	2.15	9.55	24.66	237	9.79	22.36	208	9½ سنة
0.222	0.047	2.00	8.99	25.26	149	9.45	23.18	179	10 سنوات
0.267	0.007	2.72	8.53	28.10	194	9.09	25.80	224	10½ سنة
0.232	0.016	2.43	8.65	29.35	210	9.41	27.21	228	11 سنة
-	غير دالة	1.60	8.06	31.48	208	8.25	30.22	211	11½ سنة
-	غير دالة	1.85	7.02	32.54	194	8.31	31.11	209	12 سنة
-	غير دالة	0.17	7.71	32.83	210	8.79	31.87	183	12½ سنة
-	غير دالة	1.02	6.87	33.66	236	7.36	34.33	240	13 سنة
-	غير دالة	0.15	7.73	34.52	177	6.84	35.01	234	13½ سنة
-	غير دالة	0.17	7.22	35.32	195	6.99	35.43	194	14 سنة
-	غير دالة	1.74	6.19	37.91	228	6.71	36.85	216	14½ سنة
0.208	0.034	2.13	5.68	38.09	202	6.26	36.83	218	15 سنة
-	غير دالة	0.48	6.23	37.75	196	7.55	37.41	189	15½ سنة
-	غير دالة	1.93	6.17	38.31	197	6.53	37.17	295	16 سنة
0.217	0.045	2.21	5.57	39.09	237	5.35	37.90	184	16½ سنة
0.342	0.01	2.65	5.60	38.71	174	5.53	37.64	92	17 سنة فأكثر

يتضح من الجدول رقم (4) وجود فروق دالة بين الجنسين في ثمان فئات عمرية (من 21 فئة) وهي الأعمار: 9، 9½، 10، 10½، 11 سنة وفي الفئات 15، 16½، 17 سنة، حيث كانت متوسطات الإناث أعلى من الذكور فيها جميعاً (لكنها لم تتعدى درجتين)، ولم توجد فروق في 13 فئة عمرية. وباستخدام معادلة كوهن Cohen لحساب حجم التأثير (مراد، 2011) وجد أن حجم التأثير للفروق الدالة يتراوح بين 0.20 إلى 0.34 وهو حجم تأثير ضعيف وفقاً لما حدده كوهن (0.5 هي حجم تأثير متوسط، 0.8 فأكثر لحجم تأثير مرتفع)؛ ولذلك فإن هذه الفروق قد ترجع إلى كبر حجم العينة، ونظراً لأن الفروق الدالة في 9 فئات فقط بنسبة 36%، ولضعف حجم التأثير فقد تم دمج الذكور مع الإناث في حساب المعايير المئينية.

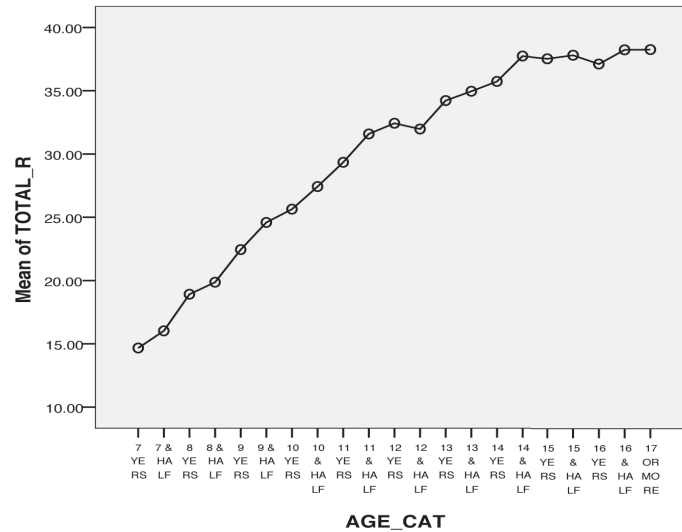
الفروق بين الفئات العمرية:

تم إجراء تحليل تباين أحادي لتعرف الفروق بين الفئات العمرية (21 فئة)، وقد تم جمع الذكور والإناث في هذا التحليل، ويوضح الجدول (5) وجود فروق دالة بين 13 فئة عمرية من الـ (22).

جدول (5) تحليل التباين الأحادي (الفئات العمرية) لدرجات اختبار ريزن للذكاء

مصدر التباين	مج المربعات	د.ح.	متوسط المربعات	قيمة (ف)	مستوى الدلالة
بين الفئات العمرية	475985.9	20	23799.29	476.21	0.001
داخل الفئات	410409.5	8452	49.98		
الكلية	886395.4	8472			

تدل نتائج جدول (5) على وجود فروق دالة بين الفئات العمرية، وبإجراء المقارنات المتعددة بين المتوسطات تبين وجود فروق دالة بين 13 مجموعة عمرية من أصل 21 مجموعة، ويوضح الشكل الآتي لمتوسطات الفئات العمرية هذه الفروق:



المعايير المئينية:

تم حساب الرتب المئينية للدرجات الخام لاختبار "ريفن" للذكاء (المصفوفات المتدرجة القياسية)، لكل فئة عمرية من الفئات الـ (21) بعد دمج الذكور والإناث معاً بسبب ضعف الفروق بين المتوسطات وضعف حجم التأثير في حال وجود فروق دالة، ويوضح الجدول (6) المئينيات (على يمين ويسار الجدول) وما يقابلها من الدرجات الخام (داخل الجدول وأسفل كل فئة عمرية).

وقد تم استخراج الرتب المئينية (كل خمس رتب مئينية) المقابلة للدرجات الخام وذلك لكل فئة عمرية (سنة أشهر).

جدول (6) المعايير المئينية المقابلة للدرجات الخام لكل فئة عمرية

المئيني	10	9½	9	8½	8	7½	7	المئيني
5	12	11	10	9	9	8	7	5
10	13	12	11	10	10	9	8	10
15	14	13	12			10	9	15
20	16	14	13	11	11		10	20
25	18	15	14	12		11		25
30	20	17	15	13	12	12	11	30
35	23	19	17	14	13			35
40	24	21	18	15	14	13	12	40
45	25	23	20	16	15			45
50	26	25	22	17	16	14	13	50
55	27	26	23	19	17	15		55
60	29	28	25	21	18	16	14	60
65	30	29	26	23	20	17	15	65
70	31	31	28	25	22	18	16	70
75	32	32	30	27	25	19	17	75
80	33	33	31	28	27	21	18	80
85	34	34	32	30	29	23	20	85

90	35	36	34	31	31	25	23	90
92	36	37	35	32	32	26	24	92
95	37	38	36	34	34	28	27	95
97	39	39	37	35	34	30	28	97
99	40	41	39	37	36	32	32	99
99.5	41	43	40	38	37	34	34	99.5
99.9	42	45	42	41	41	39	36	99.9

تابع جدول (6) المعايير المثبتة المقابلة للدرجات الخام لكل فئة عمرية

المثني	14	13½	13	12½	12	11½	11	10½	المثني
5	23	20	19	17	18	16	13	12	5
10	26	24	23	18	18	16	16	14	10
15	28	26	26	21	21	21	19	16	15
20	30	28	28	25	25	24	21	18	20
25	32	31	30	28	28	27	23	20	25
30	33	32	31	29	29	28	25	22	30
35	34	33	32	30	30	29	26	24	35
40	35	34	33	31	31	30	28	26	40
45		35	34	32	32	31	29	27	45
50	36	36	35	33	33	32	30	28	50
55	37	37	36	34	34	33	31	29	55
60	38			35	35	34	32	31	60
65		38	37			35	33	32	65
70	39	39	38	36	36	36	34	33	70
75	40			37	37	37	35	34	75
80	41	40	39	38	38		36	35	80
85	42	41	40	39	39	38	37	36	85
90		42	41	40	40	39	38	37	90
92	43	43	42	41	41	40	39	38	92
95	44	44		42	42	41	40	39	95
97	45	45	43	43	43	42	41	40	97
99	47	47	45	44	44	43	42	42	99
99.5	48	48	46	46	46	44	43	43	99.5
99.9	49	49	48	47	47	45	45	44	99.9

تابع جدول (6) المعايير المثبتة المقابلة للدرجات الخام لكل فئة عمرية

المثني	17 فأكثر	16½	16	15½	15	14½	المثني
5	28	28	26	27	27	26	5
10	31	31	29	29	29	30	10
15	33	32	31	31	31	31	15
20	34	34	33	33	33	33	20
25	35	35	34	34	34	34	25
30	36	36	35	35	35	35	30
35	37	37	36	36	36	36	35
40			37	37	37	37	40
45	38	38					45

50			38	38	38	38	50
55	39	39	39	39			55
60					39	39	60
65	40	40	40	40	40	40	65
70	41	41	41	41	41	41	70
75							75
80	42	42	42	42	42	42	80
85	43	43	43	43	43	43	85
90	44	44	44	44	44	44	90
92	45	45	45	45	45		92
95	46	46	46	46	46	45	95
97	47	47	47	47	47	46	97
99	48	48	48	49	48	47	99
99.5	50	50	50	50	49	48	99.5
99.9	51	51	51	51	50	49	99.9

مناقشة عامة:

تم تقسيم عينة الدراسة إلى 22 فئة عمرية بمعدل ستة أشهر لكل فئة، وتم اختبار الفروق بين الجنسين في كل فئة عمرية، وأسفرت النتائج عن وجود فروق دالة بين الجنسين في ثمان فئات من أصل 21 فئة (بنسبة 38%)، وبحساب حجم التأثير (الدلالة العملية) في حال الفروق الدالة وُجد أنه حجم تأثير ضعيف، وقد توصل عبد الخالق (2006) إلى فروق بين الجنسين للأعمار من 8 - 14 سنة لكنها ضعيفة أيضاً، وعليه تم دمج مجموعتي الذكور والإناث معا في كل فئة عمرية لحساب المعايير المثبتة لدرجات اختبار ريفن.

ويتضح من جداول المعايير اختلاف الدرجات باختلاف الفئة العمرية مما يدل على تطورها مع العمر، كما يتضح أن المعايير الحالية مختلفة بعض الشيء مع المعايير السابق نشرها لاختبار ريفن، ربما يرجع ذلك إلى النمو في بعض الأعمار الزمنية، والدافعية لدى الإناث في الإجابة عن أسئلة الاختبار، بينما نقصت الدرجات في بعض الفئات العمرية ويرجع هذا إلى وجود نسبة (من 3-5%) من ذوي الاحتياجات الخاصة في كل فئة عمرية بعينة التقنين حتى تكون ممثلة للمجتمع.

وبصفة عامة فقد اعتمد حساب المعايير على عينة كبيرة الحجم وممثلة للمجتمع باستثناء الفئة العمرية 6.5 سنوات التي بلغ حجمها (21) طفلاً وطفلة، وقد تم حذفها من التحليلات الإحصائية ومن جداول المعايير.

فعلى سبيل المثال إذا حصل طفل عمره 9 سنوات وثلاثة أشهر على درجة خام (27) فإن المثني المقابل لهذه الدرجة الخام يقع بين 65، 70 (من جدول 6، عمر 9 سنوات) يساوي 68. وإذا حصل آخر عمره الزمني 12 سنة وثمانية أشهر على درجة خام (34) فإن المثني المقابل لها هو 55 من معايير عمر 12.5 سنة.

المراجع

المراجع العربية:

- أبو حطب، فؤاد (1977). تقنين اختبار المصفوفات على البيئة السعودية. في: فؤاد أبو حطب (محرر) بحوث في تقنين الاختبارات النفسية. القاهرة: مكتبة الأنجلو المصرية، المجلد 1، 191-246.
- أبو علام، رجاء (1981). اختبار المصفوفات المتتابعة لرافن. الكويت: إدارة الخدمات الاجتماعية والنفسية.
- آل ثاني، العنود مبارك (2002). تقنين اختبار المصفوفات المتتابعة لرافن على طلاب وطالبات المرحلة الابتدائية بمدينة الدوحة بدولة قطر. ملخص رسالة ماجستير، مجلة مركز البحوث التربوية، جامعة قطر، 11، 229-234.
- الجميسان، محمد إبراهيم (1998). تطبيق اختبار المصفوفات المتتابعة لرافن على عينة من طلاب كلية المعلمين بالرياض وكلية التربية بجامعة الملك سعود. مجلة البحوث النفسية والتربوية، جامعة المنوفية، 13-97.
- عبد الحليم، رياض، والشريف، صلاح الدين (2001). تقنين المصفوفات المتدرجة لرافن على التلاميذ الصم للأعمار 6-8 سنة بمحافظات أسيوط وسوهاج وأسوان. مجلة كلية التربية، جامعة أسيوط، 17، 252-281.
- عبد الخالق، أحمد (2006). تقنين مصفوفات "ريفن" المتدرجة القياسية على عينات كويتية من 8-15 سنة. مجلة الطفولة العربية، 7، 8-18.
- عوض، فتحية عبد الرؤوف (1999). اختبار المصفوفات المتتابعة: كراسة التعليمات. الكويت، وزارة التربية، إدارة الخدمات الاجتماعية والنفسية، مراقبة الخدمة النفسية.
- القرشي، عبد الفتاح (1987). اختبار المصفوفات المتتابعة الملونة لرافن: دليل الاستخدام. الكويت: دار القلم.
- القرشي، عبد الفتاح (1988). تقدير ثبات اختبار مصفوفات رافن الملونة وصدقه على الأطفال الكويتيين. المجلة التربوية، جامعة الكويت، المجلد 5 (17)، 211-226.
- كاظم، على مهدي؛ والزبيدي، عبد القوي سالم؛ والصارمي، عبد الله محمد؛ ويوسف، يوسف حسن؛ والجمالي، فوزية عبد الباقي؛ والمشهداني، سكرين إبراهيم؛ والبلوشي، سناء سبيل؛ والخروصي، حسين علي؛ والبوسعيدي، أميمة بدر؛ والبحراني، وداود عبد الله؛ والفوري، سعاد مبارك (2008). تقنين اختبار رافن للمصفوفات المتتابعة الملونة على الأطفال العمانيين في المرحلة العمرية بين 5-11 سنة. دراسات نفسية، 18 (3)، 391-429.
- مراد، صلاح (2011). الأساليب الإحصائية في العلوم النفسية والتربوية والاجتماعية ط2. القاهرة: مكتبة الأنجلو المصرية.
- مراد، صلاح؛ سليمان، أمين (2012). الاختبارات والمقاييس في العلوم النفسية والتربوية ط3. الكويت: دار الكتاب الحديث.
- نور الدين، أمين محمد صبري (2002). فاعلية استخدام الاختبار الموائم باستخدام الحاسب في تقدير قدرة الأفراد وتحديد الخصائص السيكومترية للمقياس. رسالة دكتوراه (غير منشورة)، كلية التربية، جامعة عين شمس، مصر.
- يحيى، على محمد؛ وإبراهيم، على محمد؛ وجمال، سعد (2003). تقنين اختبار المصفوفات المتتابعة لريفن في البيئة العمانية (مسقط). سلسلة الدراسات النفسية والتربوية - جامعة السلطان قابوس، 6، 35-58.

المراجع الأجنبية:

- Abdel-Khalek, A. M. (1988). Egyptian results on the Standard Progressive Matrices. *Personality and Individual Differences*, 9, 193-195.
- Abdel-Khalek, A. M. (2005). Reliability and factorial validity of the Standard Progressive Matrices among Kuwaiti children ages 8 to 15 years. *Perceptual and Motor Skills*, 101, 409-412.
- Abdel-Khalek, A. M., & Lynn, R. (2006). Sex differences on the Standard Progressive Matrices and in educational attainment in Kuwait. *Personality and Individual Differences*, 40, 175-182.

- Abdel-Khalek, A. M., & Lynn, R. (2008a). Intelligence, family size and family position: Some data from Kuwait. *Personality and Individual Differences*, 44, 1032-1038.
- Abdel-Khalek, A. M., & Lynn, R. (2008b). Norms for intelligence assessed by the Standard Progressive Matrices in Oman. *The Mankind Quarterly*, 69, 183-188.
- Abdel-Khalek, A. M., & Lynn, R. (2009). Norms and sex differences for intelligence in Saudi Arabia assessed by the Standard Progressive Matrices. *Mankind Quarterly*, 50, 106-113.
- Abdel-Khalek, A. M., & Raven, J. (2006). Normative data from the standardization of Raven's Standard Progressive Matrices in Kuwait in an international context. *Social Behavior and Personality*, 34, 169-179.
- Abdel-Khalek, A. M., & Raven J. C. (2008). Kuwaiti norms for the classic SPM in an international context. In J. & J. Raven (Eds.), *Uses and abuses of intelligence: Studies advancing Spearman and Raven's quest for non-arbitrary metrics* (353-362). New York, Unionville: Royal Fireworks Press.
- Aiken, L. R., & Groth-Marnat, G. (2006). *Psychological testing and assessment* (12th ed.). Boston: Allyn & Bacon.
- Court, J. H., & Raven, J. (1995). *Normative, reliability and validity studies: References*. Raven Manual: Section 7. Oxford: Oxford Psychologists Press.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure? *Psychological Bulletin*, 101, 171-191.
- Flynn, J. R. (2000). The hidden history of IQ and special education: Can the problems be solved? *Psychology, Public Policy, and Law*, 6, 191-198.
- Frisby, C. L. (1999). Culture and test session behavior: Part II. *School Psychology Quarterly*, 14, 281-303.
- Greenfield, P. M. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *Intelligence on the rise?* Washington DC: American Psychological Association.
- Gregory, R. J. (1992). *Psychological testing: History, principles, and applications*. Boston: Allyn & Bacon.
- Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence*. New York: Wiley, 267-300.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger.
- Kaplan, R. M., & Saccuzzo, D. P. (1997). *Psychological testing: Principles, applications, and issues* (4th ed.). Pacific Grove: Brooks/ Cole.
- Kline, P. (1988). The British 'cultural influence' on ability testing. In S. H. Irvine & J. W. Berry (Eds.), *Human abilities in cultural context* (187-207). Cambridge: Cambridge University Press.
- Kline, P. (2000). *Handbook of psychological testing* (2nd ed.). London: Routledge.
- Lynn, R. (2008). *The global bell curve: Race, I. Q., and inequality worldwide*. Augusta, Georgia: Washington Summit Publishers.
- Lynn, R. (2009). Fluid intelligence but not vocabulary has increased in Britain, 1979-2008. *Intelligence*, 37, 249-255.
- Lynn, R., & Abdel-Khalek, A. (2009). Intelligence in Jordan: Norms for the Advanced Progressive Matrices. *Mankind Quarterly*, 50, 114-119.
- Lynn, R. & Irwing, P. (2004). Sex differences on the Progressive Matrices: A meta-analysis. *Intelligence*, 32, 481-498.
- Lynn, R., & Vanhanen, T. (2012). *Intelligence: A unifying construct for the social sciences*. London: Ulster Institute for Social Research.
- Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing: Principles and applications* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
- Raven, J. (1991). *Irish and British standardization. Raven Manual: Research supplement 1*. Oxford: Oxford Psychologists Press.
- Raven, J. (2008a). Change and stability in RPM scores over culture and time: The story at the turn of the century. In J. Raven & J. Raven (Eds.), *Uses and abuses of intelligence: Studies*

advancing Spearman and Raven's quest for non-arbitrary metrics (213-257). New York: Royal Fireworks Press.

Raven, J. (2008b). General introduction and overview: The Raven Progressive Matrices Tests: Their theoretical basis and measurement model. In J. Raven & J. Raven (Eds.), *Uses and abuses of intelligence: Studies advancing Spearman and Raven's quest for non-arbitrary metrics* (17-68). New York: Royal Firework Press.

Raven, J. (2013, March, 21). Personal Communication.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual: Section 1; General overview*. Oxford: Oxford Psychologists Press.

Richardson, K. (2000). *The making of intelligence*. New York: Columbia University Press.