تصحيح معاملات صعوبة الفقرات لأثر التخمين في اسئلة الإختيار من متعدد: صورة معدّلة لمعادلة جيلفورد

أحمد سليمان عودة*

جامعة جدارا، إربد، الأردن

استُلم بتاريخ: ٢٠١٣/١/٩ عُدل بتاريخ: ٢٠١٣/١/٢٦ قُبل بتاريخ: ٢٠١٣/٢/١٠

بنيت معادلة جيلفورد الأصلية لتصحيح معامل صعوبة الفقرة لأثر التخمين في اختبار الإختيار من متعدد على العقاب مع الإفتراض بأن التخمين يتم عشوائيا، الا أن التقدير لهذا ا لمعامل الصعوبة يقوم على اساس رياضي بحت لايتحقق فيه الإفتراض الذي قام عليه. وأن المؤشرات العملية والمنطقية ترجح اهمية تقديم صورة معدلة لهذه المعادلة لتكون اكثر عدالة ومصداقية في تقدير معاملات الصعوبة لفقرات الإمتحانات العقلية بشكل عام، والتحصيلية بشكل خاص، ترتكز فيه هذه الصورة المعدلة على التوفيق والتكامل بين التخمين القائم على المعرفة الجزئية، والفروق الفردية لدرجة المخاطرة في ضوء التعليمات ذات الصلة بالتخمين وما يترتب عليها من اجراءات في التصحيح، وتحليل الدرجات تحليلا إجماليا وتفصيليا عاى مستوى الفقرات. وقد قدمت الدراسة الحالية هذه الصورة المعدّلة مرفقة بجداول لإستخراج معاملات الصعوبة القائمة على التخمين العشوائي مقابل معاملات الصعوبة المقدرة بالمعادلتين الأصلية والمعدلة وترجح نتائج المقارنة لهذه التقديرات توافقا اعلى بين نتائج التقديرات لمعاملات الصعوبة المشتقة من الصورة المقترحة للمعادلة والأساس النظري الذي ترتكز عليه، وهي معادلة توفيقية تجمع بين نظرية القرار ونظرية القياس.

كلمات مفتاحية: صعوبة الفقرة، التصحيح لأثر التخمين، المخاطرة في الإجابة، فقرات الإختيار من متعدد، المعرفة الجزئية، معادلة التصحيح.

Revision of Guilford Formula to Correct Item Difficulty for Guessing in Multiple Choice Test Items

Ahmad S. Audeh*
Jadara University, Irbid, Jordan

The original Guilford formula for estimation of multiple choice item difficulty was based on a penalty for guessing. This penalty was originally based on completely random or blind guessing, which means that it is purely based on mathematical estimation and on significantly violated assumptions. While authentic and fair estimation is expected to be based on mixed scoring formula which adds another correction factor to integrate measurement theory with decision theory based on partial knowledge and risk- taking behavior. A new formula with two correction factors related to guessing, partial knowledge and risk-taking is presented in this paper. Further studies are suggested for reviewing the validation of the main assumptions of item theory models.

Keywords: item difficulty, correction for guessing, risk-taking, multiple choice items, partial knowledge, scoring formula.

*audeh@yu.edu.jo

يتزايد الاهتمام بصناعة الأسئلة من نوع الاختيار من متعدد لكثرة المزايا التي تتمتع بها، وتزايد استخداماتها في الاختبارات واسعة النطاق (Large scale assessment) التي تطبق على أعداد كبيرة من المفحوصين، وقدرتها على تحقيق درجة عالية من صدق المحتوى التعليمي؛ ولكن يعاب عليها أنها عرضة للتخمين Guessing، الذي يعتبر مصدراً من مصادر الأخطاء (E) في نظرية القياس التقليدية القائمة على العلاقة بين الدرجات الحقيقية (T) والدرجات الظاهرية (X) حيث أن X = T + E. وتأثيره كذلك (اي التخمين c) على احتمال إجابة مفحوص قدرته (θ) على الفقرة االتي تعرّف بصعوبتها (b) وتمييزها (a) وفقا للمعادلة الشاملة لمعالم الفقرة الثلاثة (الصعوبة والتمييز والتخمين $P(\theta) = c + (1-c)/[1+e^{-1.7a(\theta-b)}]$ في نظرية استجابة الفقرة، واشير اليهما بإختصار وعلى التوالي نظرية الاختبار (Test theory) ونظرية الفقرة (Item theory). وبما أن التخمين هو نوع من الخطأ في القياس، فهو عامل مؤثر في دقة تقدير الدرجة الحقيقية للمفحوص من الدرجة الظاهرية ، طالما أن احتمال الإجابة عن الفقرة عند قدرة معينة $P(\theta)$ يعتمد على صعوبة الفقرة المتأثرة بالتخمين.

ووفقا لنظرية القياس القائمة على الاختبار فان معامل صعوبة الفقرة المقدر بنسبة الإجابات الصحيحة R/N (حيث أن N عدد المفحوصين، R عدد الذين أجابوا إجابة صحيحة عن الفقرة)، تفترض بأن كل الإجابات الصحيحة (R) ناجّة عن معرفة حقيقية، اي ان $R = R_T$، ولا يوجد للتخمين أي دور في تحديد الإجابات الصحيحة، بمعنى أن $R_g = 0$. وأكدت أدبيات القياس على عدم صحة هذا الافتراض، وأن مجرد استخدام فقرات (أسئلة) الاختيار من متعدد يعني بالمنطق الرياضي أن هذا الافتراض لا يتحقق (Violated)، وأن درجة عدم التحقق تزداد بنقصان عدد البدائل، وتزداد مقاومة الاختبار (Robustness) لأثر التخمين بزيادة عدد البدائل.

وبما أن عدد البدائل في الحدود القصوى قد يتراوح بين (2-8)، فإن فرصة التخمين تتراوح نظريا بين (.125-.50) ولذلك أكدت أدبيات القياس على أهمية الحد من تأثير التخمين على دقة القياس بأساليب إحصائية منسجمة منطقيا مع تعليمات الإختبار، فقد يطلب من المفحوص في تعليمات الاختبار أن لا يلجأ إلى التخمين إذا شعر بأنه لا يمتلك المعرفة للإجابة عن الفقرة، وأن يتركها دون إجابة، وكأنه يقول للمفحوص بأنك ستخسر إذا لم تلتزم (وهذا هو الشائع في الأوساط التربوية) بمعنى ان على المفحوص ان لايغامر او يخاطر في الإجابة، أو أن يطلب منه الإجابة عن جميع الفقرات، وكأنه يقول للمفحوص عليك الاستفادة من أي فرصة للتخمين طالما أنك لا تخسر شيئاً من درجتك مفترضاً (أي الفاحص) تساوي تأثر جميع المفحوصين إيجابياً بفرصة التخمين، الا أن تباين الخطأ المنتظم الناجم عن التخمين إلى جانب الخطأ العشوائي يستحق الإشارة إليه صراحة في المعادلة الرئيسية للنظرية التقليدية في القياس ليصبح $X = T_t + T_g + e$. بمعنى أن جزءاً من الدرجة الظاهرية (X) حقيقياً فعلاً (T_t) وناجّاً من الاختلاف في قدرات المفحوصين على السمة المقاسة ، وجزءاً آخر ناجّاً عن الاختلاف في القدرة على التخمين الذي يجمع بين التخمين العشوائي (Blind Random) والتخمين الذكي (Intelligent) القائم على المعرفة الجزئية (Partial knowledge)، والمخاطرة (Risk taking)، مع الإفتراض الضمني بأن جميع المفحوصين يستفيدون من فرصة التخمين، وفي نفس الوقت يختلفون في قدرتهم على التخمين، كما يختلفون في درجة التزامهم بالتعليمات عندما يطلب منهم أن لايخمنوا؛ بمعنى أنهم يختلفون في درجة تقديرهم لأثر المخاطرة في الإجابات انطلاقاً من مبدأ الفراق الفردية في السمات الشخصية ذات الصلة بموقف الاختبار وتقديراتهم لمضامين هذه التعليمات، وتوقعاتهم لآثار الالتزام أو عدم الالتزام بها.

ومن زاوية أخرى، يمكن ان ينظر الى التخمين على انه عامل (Factor) من العوامل المحددة للعلامة X، ويضيف تباينا حقيقياً ($S^2_{T_g}$) يفترض بأنه لا يرتبط بالتباين في القدرة الحقيقية الخاصة بما وضع الاختبار لقياسه ($S^2_{T_t}$)، وقد يترتب عليه بشكل جوهري تخفيض صدق الاختبار الذي يعتمد إحصائياً على نسبة التباين الحقيقي المرتبط بالسمة المقاسة من التباين الكلي ($S^2_{T_t}/S^2_x$)، بمعنى أن نسبة التباين الناجّة من عامل التفاعل بين التخمين والمخاطرة ($S^2_{T_g}/S^2_x$) تساهم في نسبة التباين المحددة للصدق (S^2_T/S^2_x)، وعندها قد ينعكس ذلك جوهرياً أيضاً على تشبعات (Loadings) فقرات الاختبار بهذا العامل؛ وبالتالي عدم تحقق افتراض أحادية البعد (Unidimentionality) التي تقوم عليه النماذج الرئيسية في النظرية الحديثة في القياس (نظرية الفقرة) الأحادية والثنائية والثلاثية المعلمة، التي تفترض جميعها (اي النماذج) أن دقة كل المعالم التي تقدنها البرامج الإحصائية القائمة على هذه النماذج مبنية على مدى تحقق افتراض أحادية البعد.

الدراسات السابقة

اختلفت الدراسات السابقة في منهجيات التعامل مع اساليب التصحيح لأثر التخمين بشكل عام. فقد اشارت دراسة مارتن واروين وجورج (Martin, Erwin & George, 1972) الى ضعف الأساس المنطقي لهذا التصحيح القائم على التباين في سرعة الطلبة التي تدفعهم الى الاجابة العشوائية عن الاسئلة الواقعة في نهاية الامتحان. واشارت لاحقا دراسة روبرت (Robert, 1988) الى أن العقاب لأثر التخمين يمكن أن يكون مناسبا لاختبارات السرعة المتطرفة او اختبارات القوة المتطرفة، ولكنها لا تناسب معظم الامتحانات الصفّية التي تجمع بين القوة والسرعة؛ ولذلك فان المعادلة التقليدية في التصحيح لأثر الخمين لا تراعي المعرفة الجزئية. كما أن معرفة الطلبة بأن هناك عقاب لأثر التخمين سيزيد من وقت الاجابة، ويؤثر بالتالي على صدق الاختباروفقراته. أما فلدت (Feldt, 1993) فقد اشار الى ان أدبيات القياس تقدم ارشادات ووصفات لمطوري الاختبارات ومنها اختيار معاملات صعوبة الفقرات بمتوسط = ٠.٥٠ للحصول على ثبات أعلى.وقد اشار في نتائج دراسته عن أثر توزيع صعوبة الفقرات على ثبات الاختيار الى ان معاملات الصعوبة في حالة الاسئلة التي تصحح (١،٠) لا بد وأن تكون أعلى من ٠.٥٠، ووجد ان الثبات كان أعلى عندما تراوحت معاملات الصعوبة بين ٠.٥٧ ، ٠.٦٧ المقابلة لفرصة تخمين تراوحت بين ٠.٥٠ ، ٠.١٥ وأشار الى التأثير العكسي على ثبات الاختيار عندما تزداد فرصة التخمين الى ٠.٢٥ مع معاملات صعوبة تصل الى ٠.٧٤ . كما أشار لو ولو وهنج ويسب (Lau, Lau, Hong & Usop, 2011) الى أن استخدام الطريقة التقليدية للتصحيح لأثر التخمين تؤدي بشكل عام الى التحفظ في الاجابة الى الحد الذي يحرم الطلبة من توظيف المعرفة الجزئية، وقد أشاروا في دراستهم ووفق تصميم معين استخدموا فيه الحذف الجزئي للبدائل أن مراعاة المعرفة الجزئية في حساب الدرجة يؤدي الى تقليل خطأ القياس.

وفي هذا السياق ايضا قارن ثوماس ونيل والكس وآن (Thomas, Neal, Alex & Anne, 2006) الدرجات على اختبار فقراته على شكل اجابة قصيرة (بدون تخمين) مع الدرجات على الإختبار نفسه بوجود بدائل مع اوزان احتمالية ٢٠% للتخمين الصحيح ويعطى (١) واحتمال ٨٠% للتخمين غير الصحيح ويعطى (1/4) بحيث يكون مجموع ما يحصل المفحوص بالتخمين يساوي صفر [-1/4*0.80+1*0.20]. حيث اشارت النتائج الى ان التصحيح لأثر التخمين يزيد صدق الاختبار عندما تكون الفقرات وبدائل الاجابة غير قابلة لتوظيف المعرفة الجزئية، الذي اعتبراه غير واقعي. وبهذا الصدد فقد اشار لوسي وتريسي وجيمز ومارك Lucy, Tracey, James & Mark (2009) الى الإستماع الى الطلبة وردود افعالهم عند التعامل مع التصحيح لأثر التخمين في الوقت الذي يتزايد فيه استخدام اختبارات الاختيار من متعدد، حيث اشارت نتائج الدراسة الى أن الطلبة لا يفضلون استخدام التصحيح لأثر التخمين إلا في حالة الامتحان الذي يسمح فيه بفتح الكتاب (open-book) وأنهم يعاقبون على أدائهم عندما تصحح درجاتهم لأثر التخمين، وان الدرجات كانت أقل عندما طلب منهم ان لا يخمنوا، وفسر ذلك بأن هذا العقاب لأثر التخمين بالأسلوب التقليدي يؤدي الى أن تكون قيم معاملات صعوبة الاسئلة أقل مما هي عليه فعلاً..

قامت دراسات اخرى في تعاملها مع التخمين في الإطار نفسه على نظرية الفقرة (item theory) (او ما يطلق عليها نظرية استجابة الفقرة) مقابل دراسات اخرى قائمة على نظرية الإختبار (test theory). فقد قارنت دراسة كلوز وكريستيان (Klaus,. & Christian, 2007) نتائج التحليل لإجابات الطلبة على نماذج اختبار تختلف فقط في عدد بدائل فقراته بنتائج التحليل لنموذج مرجعي لايعطى فيه بدائل (اجابة قصيرة) . وأشارت النتائج وفق نموذج راش ان صعوبة الفقرة بخمسة بدائل أواكثر تعادل صعوبتها الفقرة بدون بدائل (أي لا يكون هناك فرصة للتخمين)، بمعنى ان فرصة التخمين تصبح ضعيفة جداً بخمسة بدائل أو أكثر، وأن التغير في صعوبة الفقرة لا يكون جوهرياً. وبحثت دراسة شانج ولن ولن (Chang, Lin, & Lin, 2007) في اثر مراعاة المعرفة الجزئية القائمة على اسلوب الحذف المتدرج (subset elimination) والاختيار المتدرج (subset selection) على خصائص الاختبار والفقرات وفق النموذج المتدرج القائم على المعرفة الجزئية (partial credit step function) لنظرية الفقرة مقارنة بالأسلوب الذي يعتمد عدد الاجابات التي يتم اختيارها دون اي مراعاة للمعرفة الجزئية. وأكدت نتائج الدراسة على اهمية حوسبة الاختبارات واستخدام برامج تسهل عملية التصحيح بالأساليب التي تراعي المعرفة الجزئية. كما اكدت دراسة اسبنوسا وجارديزابال (Espinosa & Gardeazabal, 2010) على اهمية التعمق في النظر الى تصرف الطلبة عند اجاباتهم عن اسئلة الاختيار من متعدد والتي تتفاعل فيها المعرفة مع المخاطرة، والى ضرورة التمييز بين نظرية القرار (Decision theory) مقابل نظرية القياس،عند التعامل مع الجوانب النظرية والتطبيقية للإختبارات. فقد اشارت الى ان سلوك الطالب يميل الى التحفظ وعدم المخاطرة ، كلما توقع عقاباً عالياً بسبب التخمين. وبالتالي زيادة عدد الفقرات المتروكة دون اجابة، كما اشارا الى ان الإفتراضات التي تقوم عليها نظرية القياس القائمة على الفقرة لا تتحقق طالما أن هناك عامل مخاطرة، ولذلك تحدث عن نموذج توفيقي يجمع بين نظرية

القرار ونظرية القياس يأخذ بالاعتبار الفقرات المتروكة بسبب ضيق وقت الامتحان بالنسبة للمفحوص مقابل الفقرات التي قامت اجابتها على معرفة جزئية أو المخاطرة التامة في الاجابة، وربط هذه العوامل المؤثرة في الإجابة على توزيع وقت الاجابة (Response time)

تعكس نتائج الدراسات السابقة بصورة مباشرة اوغير مباشرة جدلية التعامل مع التخمين في اسئلة الإختيار من متعدد في ضوء عدة عوامل وممارسات في الإطارين النظري والتطبيقي، وكيفية التعامل مع التخمين والمخاطرة والمعرفة الجزئية ، واختلاف تأثيرها على خصائص الإختبار وفقراته باختلاف التعليمات وعدد بدائل الإجابة ، وقد بينت هذه الدراسات بصورة غير مباشرة انه لايوجد وصفة سحرية لضبط أثر التخمين ، ولا الى معرفة حجم التأثيرعلى هذه الخصائص ودقة القرارات ، ولكنها عكست اهمية التصحيح لأثر التخمين بأساليب أكثر عدالة من الأسلوب التقليدي القائم على العقاب، ومن الطبيعي ان تكون صعوبة الفقرات في اختبارات الإختيار من متعدد محور هذا التصحيح،وهو الغرض الأساسي في هذه الدراسة.

مشكلة الدراسة

أدرك المنظّرون في القياس مبكراً أهمية الحد من التلوث الإحصائي (Contamination) في التقديرات لدرجات المفحوصين وخصائص الفقرات بمعادلات إحصائية وفقاً لتعليمات الإختبار تحت عناوين مختلفة أبرزها التصحيح لأثر التخمين، إلا أن مساهمات الباحثين للحد من هذا التلوث في إطار النظرية الحديثة القياس في القياس مازالت متواضعة، وقد كانت أبرز هذه المساهمات من شو (Chiu, 2010)، وهو على اي حال ليس موضوع البحث هنا، لأن النظرية التقليدية ما زالت مسيطرة ، وأنه ينظر لنظرية الفقرة بانها مكملة وليست منافسة لها (Osterlind, 2006, p263)

وبالمقابل ، يأتي هذا البحث في إطار الجهود العلمية للباحثين المهتمين في القياس النفسي والتربوي بشكل عام والاختبارات من نوع الاختيار من متعدد بشكل خاص المرتكزة على النظرية التقليدية، وتحديدا تعديل معادلة تصحيح معامل صعوبة الفقرة لأثر التخمين التي بدأها جيلفورد (Guilford, 1936)، في ضوء تعدد معادلات تصحيح درجات المفحوصين لأثر التخمين التي تأخذ بالإعتبار الإختلاف بين التخمين العشوائي والتخمين الذكي والتخمين القائم على المعرفة الجزئية وتفاعل التخمين مع المخاطرة، وهو تعديل مقترح لمعادلة جيلفورد لتصحيح صعوبة الفقرة لآثر التخمين، انطلاقا من تعظيم عدالة هذه المعادلة في ضوء مدى مقاومتها للافتراضات الضمنية التي قامت عليها اوانطلقت منها ، ومدى انسجامها مع جدلية التصحيح لأثر التخمين في ضوء تعليمات الاختبار من جهة، والأسس المنطقية التي ارتكز عليها التعدد في المعادلات المقترحة للتصحيح لأثر التخمين من جهة اخرى.

تعليمات الاختبار ومعادلات التخمين

تؤكد ادبيات القياس التربوي على اهمية الإلتزام بقواعد وشروط تطبيق الاختبارات العقلية بشكل عام والتحصيلية بشكل خاص وتصحيحها وتفسير درجاتها وفقا لهذه الشروط، ومن ضمنها التعليمات المتعلقة بالتخمين التي تعتبر عقداً بين الفاحص والمفحوص يترتب عليها اجراءات متعلقة بالتصحيح لأثر التخمين، كأن يطلب من المفحوص:

١. الإجابة عن جميع الفقرات، وهي في هذه الحالة لا تخول الفاحص اتخاذ أي إجراء لمعاقبة المفحوصين لأثر التخمين، وأن صعوبة الفقرة الثنائية التدريج (0,1) تقدر نسبة المفحوصين الذين أجابوا إجابة صحيحة عن الفقرة (R/N) على افتراض أن N عدد المفحوصين.
٢. أن لا يخمن: وهي تعني ضمناً أن على المفحوص أن لا يخاطر في الإجابة، ويتوقع منه ان يحد من درجة مخاطرته، مع الاعتراف بالفروق بين المفحوصين في درجة المخاطرة أو التجاوب مع التعليمات، أو على المفحوص أن يكون مهيئاً لتعديل درجته الظاهرية (X)، علماً بأن التعديل قد يحتمل أكثر من خيار يمكن تلخيصها على النحو الآتي:

(١) العقاب التخمين وفقاً للمعادلة التقليدية، وذلك بتخفيض العلامة الظاهرية (X) بمقدار ما يمكن أن يحققه المفحوص من علامة إضافية عن طريق التخمين (عدد الفقرات تتناسب مع الإجابات التي أجاب عنها إجابة خاطئة (W))، الذي يفترض بأنه يتم بصورة عشوائية تامة، ولا تعترف بالتخمين الذكي أو استخدام المعرفة الجزئية، ولكنها تأخذ بالاعتبار درجة المخاطرة التي تتناسب مع عدد الفقرات المتروكة دون إجابة (O)، وتحسب الدرجة المصححة (X) بتخفيض (X) العدد المتوقع رياضياً للإجابات الصحيحة حسب الفرصة المهيأة للإجابة الصحيحة من البدائل الخاطئة (L-1)، على الافتراض أن للفقرة أربعة بدائل واحدة منها صحيحة، أي أن $X^* = x - \frac{W}{L-1}$)، وقد اعتبر لتل (Little,1962) أن هذه المعادلة متشددة أو قدم معادلة أقل تشدداً، وذلك بتقليل التخفيض في

الدرجة إلى النصف لتصبح $X^* = x - \frac{W}{2(L-1)}$). مفترضاً أن ما يقوم به المفحوص في الإجابة لا يمكن أن يكون عشوائياً تاماً أو تخميناً ذكياً تاماً، وافتراض أن التخمين يتم مناصفة بين النوعين.

(٣) المكافأة لعدم المخاطرة، والتي تتناسب مع عدد الفقرات المتروكة، وهي تعطي المفحوص درجة إضافية تتناسب مع الاحتمال الرياضي للإجابة الصحيحة على الفقرة فيما لو أجاب عنها إجابة عشوائية لتصبح

$X^* = X + \frac{O}{L}$

). والتي يعاب عيها بأنها تبالغ (Overestimate) في تقدير قدرة المفحوص، بعكس طريقة العقاب التي يعاب عليها بأنها تبخس (Underestimate) قدرة المفحوص؛ ولذلك جاء التفكير بموازنة التحيز (Balancing bias) بمعادلة تجميع بين الاستدلال المنطقي والاستدلال الإحصائي

$X^* = X - \frac{W}{L-1} + \frac{O}{L}$

). والسؤال المطروح هنا يتمثل بانعكاس هذه الخيارات على تصحيح معاملات صعوبة الفقرات وليس فقط على درجات المفحوصين.

معادلة جيلفورد لتعديل صعوبة الفقرة:

لم يكن الهدف من هذا العرض التعريف بتصحيح الدرجات لأثر التخمين أو المخاطرة لأنها متوفرة في أدبيات القياس التربوي، ولكن الهدف هو الربط بين تصحيح الدرجات وتصحيح صعوبة الفقرات التي بدأها جيلفورد (Guilford, 1936)، حيث تعرّف صعوبة الفقرة إحصائياً بأنها نسبة المفحوصين الذين أجابوا عنها إجابة صحيحة (R/N)، وهو تعريف تقليدي يقدّر صعوبة الفقرة بأعلى مما هي عليه حقيقة رقيمياً حسب التعريف الإجرائي للصعوبة، ولذلك قدم جيلفورد تصحيحاً لمعامل الصعوبة (d_c) قائما على أنه نسبة المفحوصين الذين أجابوا إجابة صحيحة بدون تخمين $S = R - \frac{W}{L-1}$) (S/N) معتبراً أن). وانتهت عمليات الاشتقاق بالتعويض للتوصل إلى أن معامل الصعوبة المصحح لأثر التخمين

$d_c = \frac{S}{N} = \frac{Ld-1}{L-1}$

). وهي معادلة تصحيح قائمة على العقاب لأثر التخمين دون التمييز بين الفقرات المحذوفة (O)، والفقرات التي أجاب عنها إجابة خاطئة (W)، مفترضاً أن لدى جميع المفحوصين نفس الدرجة من المخاطرة، وكأنه لا علاقة بين التصحيح وتعليمات الاختبار، ولذلك فهي ممارسة غير منطقية من ناحية، والافتراضات القائمة عليها لا تتحقق من ناحية اخرى، ولا تتجاوب مع التحذيرات الخاصة بانتهاك (Violation) الافتراضات التي تقوم معادلة تصحيح الدرجات بالعقاب، ولذلك فهي(اي معادلة التصحيح بطريقة جيلفورد) تتجاوز افتراض عدم وجود تباين في المخاطرة بين المفحوصين، وتتجاوز حقيقة التخمين الذكي والمعرفة الجزئية، حيث أن (W) في معادلة جيلفورد غير تلك التي ترد عادة في معادلة التصحيح لأثر التخمين بالعقاب انسجاما مع التعليمات التي تنص على أن لا يخمن أو لا يخاطر، فهو يفترض (اي جيلفورد) أن جميع المفحوصين لا يتقيدون بالتعليمات كما لو أنه طلب منهم الإجابة عن جميع الفقرات.

التعديل المقترح لمعادلة جيلفورد

(١) ينطلق التعديل المقترح من أن: التعليمات تنص على أنه يطلب من المفحوص أن لا يخمن أو لا يخاطر في الإجابة العشوائية، ويتحمل بالتالي مسؤولية عدم الالتزام بهذه التعليمات، كما أن عدد الإجابات الخاطئة (w) في التعديل المقترح تختلف عنها عند جيلفورد، فهي عند جيلفورد تساوي (N-R)، إلا أنها في التعديل المقترح (N-R-O).

(٢) معامل الصعوبة المصحح d_c^* قائم على نسبة الإجابات الصحيحة المعدّلة لأثر التخمين $\frac{1-d}{L}$ التي توازن بين المكافأة والعقاب، $\frac{Ld-1}{L-1}$ لأن التصحيح الذي يجمع بين المكافأة والعقاب يتوقع ان يحقق التوازن النسبي في التحيز بتقدير معامل صعوبة الفقرة وفقا للمعادلة المقترحة .

$$d_c^* = \frac{Ld-1}{L-1} + \frac{1-d}{L}$$

ويبين جدول١ مقارنة ثنائية لقيم معاملات الصعوبة المعدلة بالمعادلتين d_c ، d_c^* المقابلة لمعاملات الصعوبة على متصل الصعوبة المحسوب من نسبة الإجابات الصحيحة (d).

جدول ١

معاملات الصعوبة المقدرة من نسبة الإجابات الصحيحة (d) المشار اليها بالحرف G مقابل المعاملات المعدلة بمعادلة جيلفورد d_c^* ، والمعدلة بالمعادلة المقترحة d_c^* المشار اليها بالحرف A محسوبة عند عدد من البدائل (٢-٨)

d	G٢	A٢	G٣	A٣	G٤	A٤	G٥	A٥	G٦	A٦	G٧	A٧	G٨	A٨
,٩٩	,٩٨٠	,٩٨٥	,٩٨٥	,٩٨٨	,٩٨٧	,٩٨٩	,٩٨٨	,٩٩٠	,٩٨٨	,٩٩٠	,٨٨	,٩٩٠	,٩٨٩	,٩٩٠
,٩٨	,٩٦٠	,٩٧٠	,٩٧٠	,٩٧٧	,٩٧٣	,٩٧٨	,٩٧٥	,٩٧٩	,٩٧٦	,٩٧٩	,٩٧٧	,٩٨٠	,٩٧٧	,٩٨٠
,٩٧	,٩٤٠	,٩٥٥	,٩٥٥	,٩٦٥	,٩٦٠	,٩٦٨	,٩٦٣	,٩٦٩	,٩٦٤	,٩٦٩	,٩٦٥	,٩٦٩	,٩٦٦	,٩٦٩
,٩٦	,٩٢٠	,٩٤٠	,٩٤٠	,٩٥٣	,٩٤٧	,٩٥٧	,٩٥٠	,٩٥٨	,٩٥٢	,٩٥٩	,٩٥٣	,٩٥٩	,٩٥٤	,٩٥٩
,٩٥	,٩٠٠	,٩٢٥	,٩٢٥	,٩٤٢	,٩٣٣	,٩٤٦	,٩٣٨	,٩٤٨	,٩٤٠	,٩٤٨	,٩٤٢	,٩٤٩	,٩٤٣	,٩٤٩
,٩٤	,٨٨٠	,٩١٠	,٩١٠	,٩٣٠	,٩٢٠	,٩٣٥	,٩٢٥	,٩٣٧	,٩٢٨	,٩٣٨	,٩٣٠	,٩٣٩	,٩٣١	,٩٣٩
,٩٣	,٨٦٠	,٨٩٥	,٨٩٥	,٩١٨	,٩٠٧	,٩٢٤	,٩١٣	,٩٢٧	,٩١٦	,٩٢٨	,٩١٨	,٩٢٨	,٩٢٠	,٩٢٩
,٩٢	,٨٤٠	,٨٨٠	,٨٨٠	,٩٠٧	,٨٩٣	,٩١٣	,٩٠٠	,٩١٦	,٩٠٤	,٩١٧	,٩٠٧	,٩١٨	,٩٠٩	,٩١٩
,٩١	,٨٢٠	,٨٦٥	,٨٦٥	,٨٩٥	,٨٨٠	,٩٠٣	,٨٨٨	,٩٠٦	,٨٩٢	,٩٠٧	,٨٩٥	,٩٠٨	,٨٩٧	,٩٠٨
,٩٠	,٨٠٠	,٨٥٠	,٨٥٠	,٨٨٣	,٨٦٧	,٨٩٢	,٨٧٥	,٨٩٥	,٨٨٠	,٨٩٧	,٨٨٣	,٨٩٨	,٨٨٦	,٨٩٨
,٨٩	,٧٨٠	,٨٣٥	,٨٣٥	,٨٧٢	,٨٥٣	,٨٨١	,٨٦٣	,٨٨٥	,٨٦٨	,٨٨٦	,٨٧٢	,٨٨٧	,٨٧٤	,٨٨٨
,٨٨	,٧٦٠	,٨٢٠	,٨٢٠	,٨٦٠	,٨٤٠	,٨٧٠	,٨٥٠	,٨٧٤	,٨٥٦	,٨٧٦	,٨٦٠	,٨٧٧	,٨٦٣	,٨٧٨
,٨٧	,٧٤٠	,٨٠٥	,٨٠٥	,٨٤٨	,٨٢٧	,٨٥٩	,٨٣٨	,٨٦٤	,٨٤٤	,٨٦٦	,٨٤٨	,٨٦٧	,٨٥١	,٨٦٨
,٨٦	,٧٢٠	,٧٩٠	,٧٩٠	,٨٣٧	,٨١٣	,٨٤٨	,٨٢٥	,٨٥٣	,٨٣٢	,٨٥٥	,٨٣٧	,٨٥٧	,٨٤٠	,٨٥٨
,٨٥	,٧٠٠	,٧٧٥	,٧٧٥	,٨٢٥	,٨٠٠	,٨٣٨	,٨١٣	,٨٤٣	,٨٢٠	,٨٤٥	,٨٢٥	,٨٤٦	,٨٢٩	,٨٤٧
,٨٤	,٦٨٠	,٧٦٠	,٧٦٠	,٨١٣	,٧٨٧	,٨٢٧	,٨٠٠	,٨٣٢	,٨٠٨	,٨٣٥	,٨١٣	,٨٣٦	,٨١٧	,٨٣٧
,٨٣	,٦٦٠	,٧٤٥	,٧٤٥	,٨٠٢	,٧٧٣	,٨١٦	,٧٨٨	,٨٢٢	,٧٩٦	,٨٢٤	,٨٠٢	,٨٢٦	,٨٠٦	,٨٢٧
,٨٢	,٦٤٠	,٧٣٠	,٧٣٠	,٧٩٠	,٧٦٠	,٨٠٥	,٧٧٥	,٨١١	,٧٨٤	,٨١٤	,٧٩٠	,٨١٦	,٧٩٤	,٨١٧
,٨١	,٦٢٠	,٧١٥	,٧١٥	,٧٧٨	,٧٤٧	,٧٩٤	,٧٦٣	,٨٠١	,٧٧٢	,٨٠٤	,٧٧٨	,٨٠٥	,٧٨٣	,٨٠٧
,٨٠	,٦٠٠	,٧٠٠	,٧٠٠	,٧٦٧	,٧٣٣	,٧٨٣	,٧٥٠	,٧٩٠	,٧٦٠	,٧٩٣	,٧٦٧	,٧٩٥	,٧٧١	,٧٩٦
,٧٩	,٥٨٠	,٦٨٥	,٦٨٥	,٧٥٥	,٧٢٠	,٧٧٣	,٧٣٨	,٧٨٠	,٧٤٨	,٧٨٣	,٧٥٥	,٧٨٥	,٧٦٠	,٧٨٦
,٧٨	,٥٦٠	,٦٧٠	,٦٧٠	,٧٤٣	,٧٠٧	,٧٦٢	,٧٢٥	,٧٦٩	,٧٣٦	,٧٧٣	,٧٤٣	,٧٧٥	,٧٤٩	,٧٧٦
,٧٧	,٥٤٠	,٦٥٥	,٦٥٥	,٧٣٢	,٦٩٣	,٧٥١	,٧١٣	,٧٥٩	,٧٢٤	,٧٦٢	,٧٣٢	,٧٦٥	,٧٣٧	,٧٦٦
,٧٦	,٥٢٠	,٦٤٠	,٦٤٠	,٧٢٠	,٦٨٠	,٧٤٠	,٧٠٠	,٧٤٨	,٧١٢	,٧٥٢	,٧٢٠	,٧٥٤	,٧٢٦	,٧٥٦
,٧٥	,٥٠٠	,٦٢٥	,٦٢٥	,٧٠٨	,٦٦٧	,٧٢٩	,٦٨٨	,٧٣٨	,٧٠٠	,٧٤٢	,٧٠٨	,٧٤٤	,٧١٤	,٧٤٦
,٧٤	,٤٨٠	,٦١٠	,٦١٠	,٦٩٧	,٦٥٣	,٧١٨	,٦٧٥	,٧٢٧	,٦٨٨	,٧٣١	,٦٩٧	,٧٣٤	,٧٠٣	,٧٣٥
,٧٣	,٤١٠	,٥٩٥	,٥٩٥	,٦٨٥	,٦٤٠	,٧٠٨	,٦٦٣	,٧١٧	,٦٧١	,٧٢١	,٦٨٥	,٧٢٤	,٦٩١	,٧٢٥
,٧٢	,٤٤٠	,٥٨٠	,٥٨٠	,٦٧٣	,٦٢٧	,٦٩٧	,٦٥٠	,٧٠٦	,٦٦٤	,٧١١	,٦٧٣	,٧١٣	,٦٨٠	,٧١٥
,٧١	,٤٢٠	,٥٦٥	,٥٦٥	,٦٦٢	,٦١٣	,٦٨٦	,٦٣٨	,٦٩٦	,٦٥٢	,٧٠٠	,٦٦٢	,٧٠٣	,٦٦٩	,٧٠٥
,٧٠	,٤٠٠	,٥٥٠	,٥٥٠	,٦٥٠	,٦٠٠	,٦٧٥	,٦٢٥	,٦٨٥	,٦٤٠	,٦٩٠	,٦٥٠	,٦٩٣	,٦٥٧	,٦٩٥
,٦٩	,٣٨٠	,٥٣٥	,٥٣٥	,٦٣٨	,٥٨٧	,٦٦٤	,٦١٣	,٦٧٥	,٦٢٨	,٦٨٠	,٦٣٨	,٦٨٣	,٦٤٦	,٦٨٤
,٦٨	,٣٦٠	,٥٢٠	,٥٢٠	,٦٢٧	,٥٧٣	,٦٥٣	,١٠٠	,٦٦٤	,٦١٦	,٦٦٩	,٦٢٧	,٦٧٢	,٦٣٤	,٦٧٤
,٦٧	,٣٤٠	,٥٠٥	,٥٠٥	,٦١٥	,٥٦٠	,٦٤٣	,٥٨٨	,٦٥٤	,١٠٤	,٦٥٩	,٦١٥	,٦٦٢	,٦٢٣	,٦٦٤
,٦٦	,٣٢٠	,٤٩٠	,٤٩٠	,١٠٣	,٥٤٧	,٦٣٢	,٥٧٥	,٦٤٣	,٥٩٢	,٦٤٩	,٦٠٣	,٦٥٢	,٦١١	,٦٥٤
,٦٥	,٣٠٠	,٤٧٥	,٤٧٥	,٥٩٢	,٥٣٣	,٦٢١	,٥٦٣	,٦٣٣	,٥٨٠	,٦٣٨	,٥٩٢	,٦٤٢	,١٠٠	,٦٤٤
,٦٤	,٢٨٠	,٤٦٠	,٤٦٠	,٥٨٠	,٥٢٠	,٦١٠	,٥٥٠	,٦٢٢	,٥٦٨	,٦٢٨	,٥٨٠	,٦٣١	,٥٨٩	,٦٣٤
,٦٣	,٢٦٠	,٤٤٥	,٤٤٥	,٥٦٨	,٥٠٧	,٥٩٩	,٥٣٨	,٦١٢	,٥٥٦	,٦١٨	,٥٦٨	,٦٢١	,٥٧٧	,٦٢٣
,٦٢	,٢٤٠	,٤٣٠	,٤٣٠	,٥٥٧	,٤٩٣	,٥٨٨	,٥٢٥	,١٠١	,٥٤٤	,١٠٧	,٥٥٧	,٦١١	,٥٦٦	,٦١٣
,٦١	,٢٢٠	,٤١٥	,٤١٥	,٥٤٥	,٤٨٠	,٥٧٨	,٥١٣	,٥٩١	,٥٣٢	,٥٩٧	,٥٤٥	,١٠١	,٥٥٤	,١٠٣
,١٠	,٢٠٠	,٤٠٠	,٤٠٠	,٥٣٣	,٤٦٧	,٥٦٧	,٥٠٠	,٥٨٠	,٥٢٠	,٥٨٧	,٥٣٣	,٥٩٠	,٥٤٣	,٥٩٣
,٥٩	,١٨٠	,٣٨٥	,٣٨٥	,٥٢٢	,٤٥٣	,٥٥٦	,٤٨٨	,٥٧٠	,٥٠٨	,٥٧٦	,٥٢٢	,٥٨٠	,٥٣١	,٥٨٣
,٥٨	,١٦٠	,٣٧٠	,٣٧٠	,٥١٠	,٤٤٠	,٥٤٥	,٤٧٥	,٥٥٩	,٤٩٦	,٥٦١	,٥١٠	,٥٧٠	,٥٢٠	,٥٧٣
,٥٧	,١٤٠	,٣٥٥	,٣٥٥	,٤٩٨	,٤٢٧	,٥٣٤	,٤٦٣	,٥٤٩	,٤٨٤	,٥٥٦	,٤٩٨	,٥٦٠	,٥٠٩	,٥٦٢
,٥٦	,١٢٠	,٣٤٠	,٣٤٠	,٤٨٧	,٤١٣	,٥٢٣	,٤٥٠	,٥٣٨	,٤٧٢	,٥٤٥	,٤٨٧	,٥٥٠	,٤٩٧	,٥٥٢
,٥٥	,١٠٠	,٣٢٥	,٣٢٥	,٤٧٥	,٤٠٠	,٥١٣	,٤٣٨	,٥٢٨	,٤٦٠	,٥٣٥	,٤٧٥	,٥٣٩	,٤٨٦	,٥٤٢
,٥٤	,٠٨٠	,٣١٠	,٣١٠	,٤٦٣	,٣٨٧	,٥٠٢	,٤٢٥	,٥١٧	,٤٤٨	,٥٢٥	,٤٦٣	,٥٢٩	,٤٧٤	,٥٣٢
,٥٣	,٠٦٠	,٢٩٥	,٢٩٥	,٤٥٢	,٣٧٣	,٤٩١	,٤١٣	,٥٠٧	,٤٣٦	,٥١٤	,٤٥٢	,٥١٩	,٤٦٣	,٥٢٢
,٥٢	,٠٤٠	,٢٨٠	,٢٨٠	,٤٤٠	,٣١٠	,٤٨٠	,٤٠٠	,٤٩٦	,٤٢٤	,٥٠٤	,٤٤٠	,٥٠٩	,٤٥١	,٥١١
,٥١	,٠٢٠	,٢٦٥	,٢٦٥	,٤٢٨	,٣٤٧	,٤٦٩	,٣٨٨	,٤٨٦	,٤١٢	,٤٩٤	,٤٢٨	,٤٩٨	,٤٤٠	,٥٠١
,٥٠	,٠٠٠	,٢٥٠	,٢٥٠	,٤١٧	,٣٣٣	,٤٥٨	,٣٧٥	,٤٧٥	,٤٠٠	,٤٨٣	,٤١٧	,٤٨٨	,٤٢٩	,٤٩١
,٤٩		,٢٣٥	,٢٣٥	,٤٠٥	,٣٢٠	,٤٤٨	,٣٦٣	,٤٦٥	,٣٨٨	,٤٧٣	,٤٠٥	,٤٧٨	,٤١٧	,٤٨١
,٤٨		,٢٢٠	,٢٢٠	,٣٩٣	,٣٠٧	,٤٣٧	,٣٥٠	,٤٥٤	,٣٧٦	,٤٦٣	,٣٩٣	,٤٦٨	,٤٠٦	,٤٧١
,٤٧		,٢٠٥	,٢٠٥	,٣٨٢	,٢٩٣	,٤٢٦	,٣٣٨	,٤٤٤	,٣٦٤	,٤٥٢	,٣٨٢	,٤٥٧	,٣٩٤	,٤٦١
,٤٦		,١٩٠	,١٩٠	,٣٧٠	,٢٨٠	,٤١٥	,٣٢٥	,٤٣٣	,٣٥٢	,٤٤٢	,٣٧٠	,٤٤٧	,٣٨٣	,٤٥٠
,٤٥		,١٧٥	,١٧٥	,٣٥٨	,٢٦٧	,٤٠٤	,٣١٣	,٤٢٣	,٣٤٠	,٤٣٢	,٣٥٨	,٤٣٧	,٣٧١	,٤٤٠
,٤٤		,١٦٠	,١٦٠	,٣٤٧	,٢٥٣	,٣٩٣	,٣٠٠	,٤١٢	,٣٢٨	,٤٢١	,٣٤٧	,٤٢٧	,٣٦٠	,٤٣٠
,٤٣		,١٤٥	,١٤٥	,٣٣٥	,٢٤٠	,٣٨٣	,٢٨٨	,٤٠٢	,٣١٦	,٤١١	,٣٣٥	,٤١٦	,٣٤٩	,٤٢٠
,٤٢		,١٣٠	,١٣٠	,٣٢٣	,٢٢٧	,٣٧٢	,٢٧٥	,٣٩١	,٣٠٤	,٤٠١	,٣٢٣	,٤٠٦	,٣٣٧	,٤١٠
,٤١		,١١٥	,١١٥	,٣١٢	,٢١٣	,٣٦١	,٢٦٢	,٣٨٠	,٢٩٢	,٣٩٠	,٣١٢	,٣٩٦	,٣٢٦	,٣٩٩
,٤٠		,١٠٠	,١٠٠	,٣٠٠	,٢٠٠	,٣٥٠	,٢٥٠	,٣٧٠	,٢٨٠	,٣٨٠	,٣٠٠	,٣٨٦	,٣١٤	,٣٨٩

جدول ١

معاملات الصعوبة المقدرة من نسبة الإجابات الصحيحة (d) المشار اليها بالحرف G مقابل المعاملات المعدلة بمعادلة جيلفوردd_c ، والمعدلة بالمعادلة المقترحة d_c^* المشار اليها بالحرف A محسوبة عند عدد من البدائل (٢-٨)

d	G٢	A٢	G٣	A٣	G٤	A٤	G٥	A٥	G٦	A٦	G٧	A٧	G٨	A٨
,٣٩		,٠٨٥	,٠٨٥	,٢٨٨	,١٨٧	,٣٣٩	,٢٣٧	,٣٥٩	,٢٦٨	,٣٧٠	,٢٨٨	,٣٧٥	,٣٠٣	,٣٧٩
,٣٨		,٠٧٠	,٠٧٠	,٢٧٧	,١٧٣	,٣٢٨	,٢٢٥	,٣٤٩	,٢٥٦	,٣٥٩	,٢٧٧	,٣٦٥	,٢٩١	,٣٦٩
,٣٧		,٠٥٥	,٠٥٥	,٢٦٥	,١٦٠	,٣١٧	,٢١٢	,٣٣٨	,٢٤٤	,٣٤٩	,٢٦٥	,٣٥٥	,٢٨٠	,٣٥٩
,٣٦		,٠٤٠	,٠٤٠	,٢٥٣	,١٤٧	,٣٠٧	,٢٠٠	,٣٢٨	,٢٣٢	,٣٣٩	,٢٥٣	,٣٤٥	,٢٦٩	,٣٤٩
,٣٥		,٠٢٥	,٠٢٥	,٢٤٢	,١٣٣	,٢٩٦	,١٨٧	,٣١٧	,٢٢٠	,٣٢٨	,٢٤٢	,٣٣٥	,٢٥٧	,٣٣٨
,٣٤		,٠١٠	,٠١٠	,٢٣٠	,١٢٠	,٢٨٥	,١٧٥	,٣٠٧	,٢٠٨	,٣١٨	,٢٣٠	,٣٢٤	,٢٤١	,٣٢٨
,٣٣		,٠٠٠	,٠٠٠	,٢١٨	,١٠٧	,٢٧٤	,١٦٢	,٢٩١	,١٩١	,٣٠٨	,٢١٨	,٣١٤	,٢٣٤	,٣١٨
,٣٢				,٢٠٧	,٠٩٣	,٢٦٣	,١٥٠	,٢٨١	,١٨٤	,٢٩٧	,٢٠٧	,٣٠٤	,٢٢٣	,٣٠٨
,٣١				,١٩٥	,٠٨٠	,٢٥٢	,١٣٧	,٢٧٥	,١٧٢	,٢٨٧	,١٩٥	,٢٩٤	,٢١١	,٢٩٨
,٣٠				,١٨٣	,٠٦٧	,٢٤٢	,١٢٥	,٢٦٥	,١٦٠	,٢٧٧	,١٨٣	,٢٨٣	,٢٠٠	,٢٨٧
,٢٩				,١٧٢	,٠٥٣	,٢٣١	,١١٢	,٢٥٤	,١٤٨	,٢٦٦	,١٧٢	,٢٧٣	,١٨٩	,٢٧٧
,٢٨				,١٦٠	,٠٤٠	,٢٢٠	,١٠٠	,٢٤٤	,١٣٦	,٢٥٦	,١٦٠	,٢٦٣	,١٧٧	,٢٦٧
,٢٧				,١٤٨	,٠٢٧	,٢٠٩	,٠٨٧	,٢٣٣	,١٢٤	,٢٤٦	,١٤٨	,٢٥٣	,١٦٦	,٢٥٧
,٢٦				,١٣٧	,٠١٣	,١٩٨	,٠٧٥	,٢٢٣	,١١٢	,٢٣٥	,١٣٧	,٢٤٢	,١٥٤	,٢٤٧
,٢٥				,١٢٥	,٠٠٠	,١٨٧	,٠٦٢	,٢١٢	,١٠٠	,٢٢٥	,١٢٥	,٢٣٢	,١٤٣	,٢٣٧
,٢٤				,١١٣		,١٧٧	,٠٥٠	,٢٠٢	,٠٨٨	,٢١٥	,١١٣	,٢٢٢	,١٣١	,٢٢٦
,٢٣				,١٠٢		,١٦٦	,٠٣٧	,١٩١	,٠٧١	,٢٠٤	,١٠٢	,٢١٢	,١٢٠	,٢١٦
,٢٢				,٠٩٠		,١٥٥	,٠٢٥	,١٨١	,٠٦٤	,١٩٤	,٠٩٠	,٢٠١	,١٠٩	,٢٠٦
,٢١				,٠٧٨		,١٤٤	,٠١٢	,١٧٠	,٠٥٢	,١٨٤	,٠٧٨	,١٩١	,٠٩٧	,١٩٦
,٢٠				,٠٦٧		,١٣٣	,٠٠٠	,١٦٠	,٠٤٠	,١٧٣	,٠٦٧	,١٨١	,٠٨٦	,١٨٦
,١٩				,٠٥٥		,١٢٢		,١٤٩	,٠٢٨	,١٦٣	,٠٥٥	,١٧١	,٠٧٤	,١٧٦
,١٨				,٠٤٣		,١١٢		,١٣٩	,٠١٦	,١٥٣	,٠٤٣	,١٦٠	,٠٦٣	,١٦٥
,١٧				,٠٣٢		,١٠١		,١٢٨	,٠٠٤	,١٤٢	,٠٣٢	,١٥٠	,٠٥١	,١٥٥
,١٦				,٠٢٠		,٠٩٠		,١١٨	,٠٠٠	,١٣٢	,٠٢٠	,١٤٠	,٠٤٠	,١٤٥
,١٥				,٠٠٨		,٠٧٩		,١٠٧		,١٢٢	,٠٠٨	,١٣٠	,٠٢٩	,١٣٥
,١٤				,٠٠٠		,٠٦٨		,٠٩٧		,١١١	,٠٠٠	,١٢٠	,٠١٧	,١٢٥
,١٣						,٠٥٧		,٠٨١		,١٠١		,١٠٩	,٠٠٦	,١١٤
,١٢						,٠٤٧		,٠٧٦		,٠٩١		,٠٩٩	,٠٠٠	,١٠٤
,١١						,٠٣٦		,٠٦٥		,٠٨٠		,٠٨٩		,٠٩٤
,١٠						,٠٢٥		,٠٥٥		,٠٧٠		,٠٧٩		,٠٨٤
,٠٩						,٠١٤		,٠٤٤		,٠٦٠		,٠٦٨		,٠٧٤
,٠٨						,٠٠٣		,٠٣٤		,٠٤٩		,٠٥٨		,٠٦٤
,٠٧						,٠٠٠		,٠٢٣		,٠٣٩		,٠٤٨		,٠٥٣
,٠٦								,٠١٣		,٠٢٩		,٠٣٨		,٠٤٣
,٠٥								,٠٠٢		,٠١٨		,٠٢٧		,٠٣٣
,٠٤								,٠٠٠		,٠٠٨		,٠١٧		,٠٢٣
,٠٣										,٠٠٠		,٠٠٧		,٠١٣
,٠٢												,٠٠٠		,٠٠٢
,٠١														,٠٠٠

الإستنتاجات والتوصيات

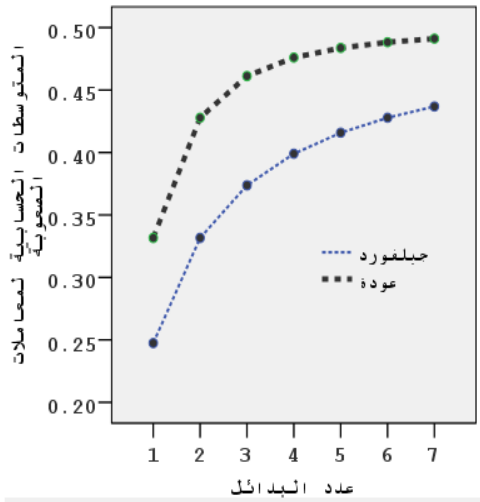
الملاحظ على النتائج من خلال البيانات في جدول ١ مايلي: اختلاف النهايات لمعاملات صعوبة الفقرات المعدلة بالمعادلتين (جيلفورد وعودة)، اي قيمة معامل الصعوبة قبل التعديل (d) المقابل للمعامل الذي يؤول الى الصفر في كل منهما(d_{cG})، (d_{cA}^*) كما هو مبين في جدول ٢.

جدول ٢

معاملات الصعوبة قبل التعديل (d) المقابلة للمعاملات التي تؤول الى الصفر في كل من معادلتي التصحيح (d_{cG})، (d_{cA}^*)

	عدد البدائل (L)						
	٢	٣	٤	٥	٦	٧	٨
(d_{cG})	٠,٥٠	٠,٣٣	٠,٢٥	٠,٢٠	٠,١٦	٠,١٤	٠,١٢
(d_{cA}^*)	٠,٣٣	٠,١٤	٠,٠٧	٠,٠٤	٠,٠٣	٠,٠٢	٠,٠١

كما بين الشكل ١ الفروق التفصيلية بين متوسطات القيم المقدرة بالمعادلتين القائمتين على التصحيح لأثر التخمين مقارنة بقيم معاملات الصعوبة الأصلية عندعدد مختلف من بدائل الإجابة (٢-٨). مع ملاحظة ان المتوسط بدون اجراء اي تصحيح = ٠,٠٥ متبوعا بجدول ٣ الذي يبين هذه المتوسطات والإنحرافات المعيارية للقيم المقدرة على متصل الصعوبة (٠,٠- ١,٠)، ثم نتائج فحص الفروق بين هذه المتوسطات باستخدام الإختبار الإحصائي ت (t) للبيانات المتجانسة في التباين حسب اختبار ليفين (Levene) المبينة في جدول ٤.



الشكل ١: الفروق الفعلية بين القيم المقدرة بالمعادلتين الناتجين عن التصحيح لأثر التخمين عندعدد مختلف من البدائل

ويلاحظ من الشكل ١ أن العلاقة بين عدد البدائل ومعاملات الصعوبة المقدرة بالمعادلتين علاقة وتيرية (monotonic)، وأن الفروق الرقمية واضحة على مستوى المتوسطات، وعلى مستوى الفقرات عند كل بديل من البدائل المبينة في جدول ١. ويلاحظ من جدول ٥ الموضح للفروق بين متوسطات التقديرات بالمعادلتين الانخفاض الجوهري مقارنة بالمتوسط المرجعي عند معامل الصعوبة المتوسط في غياب التصحيح لأثر التخمين (d= 0.50) وخاصة عندما يكون عدد البدائل (٢). ويتناقص هذا الفرق مع تزايد عدد البدائل، ويكاد يتلاشى عند العدد (٨) من البدائل وفق معادلة جيلفورد، كما يصبح الفرق بين التقديرات بالمعادلتين غير داال احصائياً عند هذا العدد.

جدول ٣

المتوسطات الحسابية والإخرافات المعيارية لمعاملات الصعوبة المقدرة بمعادلتي جيلفورد وعوده عند عدد مختلف من البدائل

عدد البدائل	المعادلة	الوسط الحسابي	الإخراف المعياري
	جيلفورد	,٠٠٥	٢٨٧
بدون تصحيح (d)	عودة	,٠٠٥	,٢٨٧
١٢	جيلفورد	,٢٤٧	,٣٢١
	عودة	,٣٣٢	,٣٢٢
١٣	جيلفورد	,٣٣٢	,٣٢٢
	عودة	,٤٢٨	,٣١٩
١٤	جيلفورد	,٣٧٤	,٣٣٠
	عودة	,٤٦١	,٣٠٧
١٥	جيلفورد	,٣٩٩	,٣٢٦
	عودة	,٤٧٦	,٣٠٠
١٦	جيلفورد	,٤١٦	,٣٢٢
	عودة	,٤٨٤	,٢٩٦
١٧	جيلفورد	,٤٢٨	,٣١٩
	عودة	,٤٨٨	,٢٩٤
١٨	جيلفورد	,٤٣٧	,٣١٦
	عودة	,٤٩١	,٢٩٢

أما دلالة هذه الفروق عن (d= 0.50) كما هي مبينة في جدول ٥ فتشير الى ان معادلة جيلفورد المعدّلة أكثرعدالة منها قبل التعديل في مراعاة المعرفة الجزئية . فمن المنطقي ان يصبح الفرق غير دال احصائيا بزيادة عدد البدائل.

وترجح نتائج المقارنة اهمية اعتماد الصورة المعدلة لمعادلة جيلفورد في تقدير معاملات صعوبة الفقرات التي توفر توازنا نسبيا بين عوامل التخمين العشوائي والمخاطرة والمعرفة الجزئية، ومن السهل على الباحثين ومطوري الاختبارات وبنوك الاسئلة استخدام الجدول المعدّ في هذه الدراسة لهذا الغرض. والتوصية أيضا باعتماد النهايات الطرفية المعدلة في تقدير معامل الصعوبة الذي اعتاد المنظرون في مجال القياس التربوي والإختبارات التحصلية بشكل خاص والقائم على الصيغة ($\frac{1+c}{2}$) حيث تقابل فرصة التخمين كما وردت في جدول رقم(٢)، وللتوضيح بمثال، فالمتوسط المقترح لمعاملات الصعوبة في حال الفقرات ذات البديلين ($\frac{1+0.33}{2}$) وليس ($\frac{1+0.50}{2}$) وهكذا بالنسبة للبدائل الأخرى الواردة في جدول(٢). وترجح الأرقام في الجدول (وتحديدا عند مقارنة قيم الدلالات الإحصائية لقيمة t بالمعادلتين مع تزايد عدد البدائل) التوصية بعدم ضرورة التصحيح لأثر التخمين عندما يكون عدد بدائل فقرات الإختبار (٤ فأكثر)، وهذا يتفق مع ماأشارت اليه دراسة كلوز وكريستيان (Klaus & Christian, 2007) بأن فرصة التخمين تصبح ضعيفة جداً بخمسة بدائل أو أكثر، وأن التغير في صعوبة الفقرة لا يكون جوهرياً.

.جدول ٤

نتائج فحص الفروق بين المتوسطات الحسابية والإخرافات المعيارية لمعاملات الصعوبة المقدرة بمعادلتي جيلفورد وعوده عند عدد مختلف من البدائل

اختبار (t) لفحص الفروق بين المتوسطات		اختباراليفين لفحص تجانس التباين Levene	
الدلالة الإحصائية	قيمة الإختبار t	الدلالة الإحصائية	قيمة الإختبار ف (F)
٠,٠٧١	١,٨١٢-	٠,٤٧٩	٠,٥٠٣
٠,٠٣٩	٢,٠٧٨-	٠,٤٥١	٠,٥٧١
٠,٠٥٥	١,٩٣١-	٠,٢٤١	١,٣٨١
٠,٠٨٦	١,٧٢٧-	٠,٢١٤	١,٥٥٥
٠,١٢٤	١,٥٤٤-	٠,٢٢٧	١,٤٧١
٠,١٦٧	١,٣٨٧-	٠,٢٥٤	١,٣٠٧
٠,٢١١	١,٢٥٥-	٠,٢٨٦	١,١٤٣

جدول ٥

الفروق بين متوسطات التقديرات بالمعادلتين والمتوسط العام لمعاملات صعوبة الفقرات دون اي تعديل المشار اليه بالمتوسط المرجعي عند (d= ٠,٠٥)

	Original Guilford Correction Formula				Modified Guilford Correction Formula			
	نتائج التحليل بمعادلة جيلفورد الأصلية				نتائج التحليل بمعادلة عودة (جيلفورد المعدّلة)			
	الوسط الحسابي	الإنحراف المعياري	قيمة ت (t)	الدلالة	الوسط الحسابي	الإنحراف المعياري	قيمة ت (t)	الدلالة
القيمة المرجعية (بدون تخمين)	٠,٥٠٠	٠,٢٨٧			٠,٥٠٠	٠,٢٨٧		
١٢	٠,٢٤٧	٠,٣٢١	٧,٨٢-	٠,٠٠٠	٠,٣٣٢	٠,٣٣٢	٥,٠٤-	٠,٠٠٠
١٣	٠,٣٣٢	٠,٣٣٢	٥,٠٤-	٠,٠٠٠	,٤٢٨	,٣١٩	٢,٢٥-	٠,٠٢٧
١٤	٠,٣٤٧	٠,٣٣٠	٣,٨١-	٠,٠٠٠	,٤٦١	,٣٠٧	١,٢٦-	٠,٢١١
١٥	٠,٣٩٩	٠,٣٢٦	٣,٠٨-	٠,٠٠٣	,٤٧٦	,٣٠٠	٠,٨٠-	٠,٤٢٨
١٦	٠,٤١٦	٠,٣٢٢	٢,٦٠-	٠,٠١١	,٤٨٤	,٢٩٦	٠,٥٥-	٠,٥٨٦
١٧	٠,٤٢٨	٠,٣١٩	٢,٢٥-	٠,٠٢٧	,٤٨٨	,٢٩٤	٠,٤٠-	٠,٦٩٢
١٨	٠,٤٣٧	٠,٣١٦	١,٩٩-	٠,٠٥٠	,٤٩١	,٢٩٢	٠,٣٠-	٠,٧١٤

وهذا يرجح التعامل مع جداول التصحيح المقدمة في هذه الدراسة، وأن التصحيح لمعاملات الصعوبة بالمعادلة المقترحة أكثر مصداقية وواقعية (more authentic). ومن خلال الربط بين الأسس التي انطلقت منها الدراسة الحالية ونتائجها، والدراسات السابقة بشكل عام، ودراسة اسبنوسا وجارديزابال (Espinosa & Gardeazabal, 2010) بشكل خاص، يقترح الباحث على السيكومتريين اعادة النظر في التعريف الإجرائي لإختبار السرعة من جهة وافتراض التحرر من السرعة speediness من جهة اخرى ، وعلاقته بالإفتراضات الأخرى التي تقوم عليها نظرية الفقرة ونماذجها بشكل عام وافتراض أحادية البعد (unidimensionality) بشكل خاص.

المراجع

References

Chang, S.-H., Lin, P. -C., & Lin, Z. C. (2007). Measures of partial knowledge and unexpected responses in multiple-choice tests. *Educational Technology & Society*, *10* (4), 95-109.

Chiu, T. (2010). *Correction for guessing in the framework of the 3pl item response theory*. (Unpublished Dissertation) Rutgers, The State University of New Jersey.

Espinosa, M. P. & Gardeazabal, J. (2010). Optimal correction for guessing in multiple-choice tests. *Journal of Mathematical Psychology*, *54*, 415 - 425.

Feldt, L. S. (1993).The relationship between the distribution of item difficulties and test reliability. *Applied Measurement In Education*, *6*(1), 37-48.

Guilford, J. (1936). The determination of item difficulty when chance success is a factor. *Psychometrika*, 1(4), 259-264.

Klaus, D. K., & Christian, H. G. (2007). Item difficulty of multiple choice tests dependant on different item response formats – An experiment in fundamental research on psychological assessment. *Psychology Science*, *49* (4), 361-374.

Lau, P. N., Lau, S. H., Hong, K. S. & Usop, H. (2011). Guessing, partial knowledge, and misconceptions in multiple-choice tests. *Educational Technology & Society*, *14* (4), 99–110.

Little, E. (1962). Overestimation for Guessing in multiple choice test scoring. *Journal of Educational Research*, *2*, 245-252.

Lucy, R. B., Tracey J. E., James, H. & Mark, T. (2009). Does correction for guessing reduce students' performance on multiple-choice examinations? Yes? No? Sometimes? *Assessment and Evaluation in Higher Education*, *34*(1), 1-15.

Martin, G. W., Erwin, K. T., & George W. H. (1972). The formula score -a correction for chance or a chance correction? *Personnel Psychology*, 25, 75-93.

Osterlind, S. (2006). Modern measurement: theory, principles, and applications of mental appraisal. Columbus: Pearson.

Robert, B. F. (1988). Formula Scoring of multiple choice tests (correction for guessing). *Educational Measurement: Issues and Practice*, 7(2), 33-38.

Thomas, J., Neal, R., Alex, C., & Anne, C. (2006). Correcting for guessing increases validity in multiple-choice examinations in an oral and maxillofacial pathology course. *Journal of Dental Education*, 70 (4), 378-386.