تقصي أثر عدد خطوات الأسئلة متعددة التدريج وشكل التوزيع لصعوبتها على تقديرات القدرة للأفراد، والصعوبة للأسئلة، ودالة المعلومات للاختبار وفق نموذج التقدير الجزئي

د. نضال كمال الشريفين د. زايد صالح بني عطا
قسم علم النفس الإرشادي والتربوي - كلية التربية - جامعة اليرموك
الأردن

الملخص

هدفت الدراسة إلى تقصي أثر عدد خطوات أسئلة الاختبار، وشكل التوزيع لصعوبة الأسئلة في دقة التقديرات لقدرات الأفراد، ومعالم الصعوبة للأسئلة، ودالة المعلومات للاختبار، وفق نموذج التقدير الجزئي. ولتحقيق هدف الدراسة، تم توليد بيانات (باستخدام برنامج WINGEN)، خاصة بثلاثة اختبارات مختلفة في عدد خطواتها (ثلاث خطوات، أربع خطوات، خمس خطوات)، وكل منها مكون من ٣٠ سؤالاً بحيث كانت المسافة بين صعوبة الخطوة الأولى والخطوة الأخيرة لها متساوية، وفق ثلاثة أشكال توزيع مختلفة لصعوبة الأسئلة: طبيعي، وملتو لليمين، وملتو لليسار، كما تم توليد استجابات ١٠٠٠ مفحوص، وفق التوزيع الطبيعي، وحللت البيانات وفق نموذج التقدير الجزئي (PCM) باستخدام البرمجية (WINSTEPS)، وبرنامج (SPSS). وبينت النتائج وجود فروق ذات دلالة إحصائية في متوسطات الأخطاء المعيارية لتقديرات القدرات للأفراد تعزى لعدد خطوات الاختبار، وللتفاعل بين عدد الخطوات وشكل التوزيع، في حين لم تظهر فروق ذات دلالة إحصائية تعزى لشكل التوزيع. كما أشارت النتائج إلى وجود فروق ذات دلالة إحصائية في متوسطات الأخطاء المعيارية لتقديرات معالم الصعوبة للأسئلة تعزى لعدد الخطوات، وللتفاعل بين عدد الخطوات وشكل التوزيع، في حين لم تظهر فروق ذات دلالة إحصائية تعزى لشكل التوزيع. كما بيّنت النتائج عدم وجود فروق دالة إحصائياً بين معاملات ثبات التباعد للأفراد المقدرة تعزى لعدد الخطوات أو عند أشكال التوزيع المختلفة. وأشارت النتائج إلى أن الاختبار ذا الخمس خطوات أعطى أكبر كمية من المعلومات سواء كان شكل التوزيع لصعوبة الخطوات طبيعياً أو ملتوياً نحو اليمين، أو ملتوياً نحو اليسار، يليه الاختبار ذو الخطوات الأربع. وأخيراً جاءت كفاءة الاختبار ذي الخطوات الخمس أعلى من كفاءة الاختبار ذي الخطوات الثلاث؛ بمعنى أن الاختبار ذا الخطوات الخمس كان أكثر دقة لتقدير قدرات الأفراد.

المقدمة

تُعد النظرية الحديثة في القياس تطوراً حديثاً ومهماً في القياس النفسي؛ إذ إنها تمكن المتخصص من الإجابة عن أي سؤال عن الفقرة أو الاختبار أو المفحوص، ومن أكثر التطبيقات شيوعاً لهذه النظرية ما يتعلق بفقرات الاختيار من متعدد في اختبارات القدرة. ولها تطبيقات في أنواع أخرى من الأسئلة، وهي تعالج كثيراً من قضايا القياس بشكل أكثر فاعلية من النظرية الكلاسيكية، وتقوم النظرية الحديثة على افتراض أن احتمال إجابة المفحوص إجابة صحيحة عن الفقرة يتحدد بقدرة المفحوص وبصعوبة الفقرة. وتفترض هذه النظرية أنه يمكن التنبؤ بأداء الأفراد، أو يمكن تفسير أدائهم في اختبار نفسي أو تربوي معين، في ضوء خاصية أو خصائص مميزة لهذا الأداء تسمى السمات (Traits)، وتحاول هذه النظرية تقدير درجات الأفراد في هذه السمات. وبالطبع يصعب ملاحظة هذه السمات ملاحظة مباشرة، لذلك يجب تقديرها (أو الاستدلال عليها) من أداء الفرد الذي يمكن ملاحظته - في مجموعة من فقرات الاختبار، ولهذا يطلق عليها السمات الكامنة.

وقد انبثق عن هذه النظرية مجموعة من النماذج تعرف باسم نماذج السمات الكامنة تهدف جميعها إلى تحديد العلاقة بين أداء الأفراد على مجموعة من فقرات الاختبار (وهو ما يمكن ملاحظته مباشرة)، وبين السمات أو القدرات التي تكمن وراء هذا الأداء وتفسيره. وتم تصنيف هذه النماذج إلى ثلاثة أصناف، يرتبط الصنف الأول منها بالاستجابات الثنائية (Dichotomous)، وفيها تكون استجابة الفرد إما إجابة صحيحة فيأخذ العلامة (١)، أو إجابة خطأ فيأخذ العلامة (صفراً). أما الصنف الثاني فيرتبط بالاستجابات المتعددة (Polytomous) وفيها تكون استجابة الفرد متدرجة، وقد تكون غير كاملة، فيحصل على علامة جزئية مثل العلامة ٤ من ٥ أو العلامة ٣ من ٥، والصنف الثالث يكون خليطاً من الصنفين الأول والثاني (Reeve, 2004).

كما تم تطوير نماذج مختلفة في نظرية السمات الكامنة، ويكمن الاختلاف

بين هذه النماذج في اختيار النموذج الرياضي لمنحنى خصائص الفقرة، وأكثر هذه النماذج شيوعاً هي: النماذج اللوجستية (Logistic Models) والنماذج التي تستخدم المنحنى الطبيعي التراكمي (Normal Ogive Models)، ويستخدم كل من هذين النوعين معلمة واحدة تمثل درجة صعوبة الفقرة، أو معلمتين تمثل إحداهما درجة الصعوبة وتمثل الأخرى درجة التميز، أو ثلاث معالم تمثل الأولى درجة الصعوبة للفقرة، وتمثل الثانية درجة تمييزها، أما الثالثة فتمثل خط التقارب الأدنى لمنحنى خصائص الفقرة (The Asymptotic Line to Item Characteristic Curve) الذي يمثل درجة التخمين عليها، ويكون موازياً لمحور القدرة، وغالباً ما تستعمل النظرية الحديثة النماذج اللوجستية أكثر من استعمالها النماذج التي تعتمد على اقتران المنحنى الطبيعي التراكمي وذلك لخصائصها الرياضية، كما أن النموذج اللوجستي ذا المعلمة الواحدة يستخدم بشكل كبير؛ وذلك لسهولة التعامل معه بالمقارنة مع النماذج الأخرى ويسمى بنموذج راش (نسبة إلى العالم Rasch)، أو نموذج المعلمة الواحدة، ومن المميزات التي أعطت هذا النموذج هذه الأهمية هي أنه عندما تتطابق البيانات مع هذا النموذج فإن معالم الأسئلة التي تتمثل في درجة صعوبتها يمكن تقديرها بشكل مستقل عن العينة المستخدمة. كما أن قدرات الأفراد يمكن تقديرها بشكل مستقل عن درجة صعوبة الأسئلة (Masters, 1982; Embretson & Reise, 2000). وفي نموذج راش الذي يتعامل مع الفقرات ذات الاستجابة الثنائية يكون مجموع العلامات التي يحصل عليها الفرد على هذه الفقرات إحصائياً كافياً لتقدير قدرة الفرد، كما أن مجموع الأفراد الذين أجابوا إجابة صحيحة عن الفقرة يكون إحصائياً كافياً لتقدير درجة صعوبتها. ونظراً للأهمية المرتبطة بنموذج راش الذي يتعامل مع الاستجابة الثنائية (Dichotomous) فقد تناول ماسترز (Masters, 1982) نماذج رياضية تتعامل مع الأسئلة ذات الاستجابات المتعددة (Polytomous). ويعد نموذج التقدير الجزئي (Partial Credit) أحد النماذج الذي يمكننا من فصل تقدير معالم الأسئلة ذات الاستجابة المتعددة عن تقدير معالم القدرات لأفراد العينة، كما يمكن فصل تقدير معالم قدرات الأفراد عن تقدير

معالم عينة الأسئلة المستخدمة. وبالتالي يّعد هذا النموذج تعميماً لنموذج راش الذي يستعمل مع الاستجابات الثنائية، لذا فإن هذا النموذج يسمى بنموذج التقدير الجزئي لراش (Rasch Partial Credit) (Eggert & Bogeholz, 2009)، وترتبط الأسئلة متعددة الخطوات بنماذج السمات الكامنة التي تتعامل مع الاستجابات المتعددة (Polytomous).

نموذج التقدير الجزئي

يعد نموذج التقدير الجزئي (Partial Credit) كما قدمه ماسترز (Masters, 1982) أحد النماذج الذي يمكن من خلاله تقدير معالم الأسئلة ذات الخطوات المتعددة من معالم قدرات أفراد العينة، كما يمكن فصل قدرات الأفراد عن معالم عينة الأسئلة المستخدمة. ويعد هذا النموذج تعميماً لنموذج راش الذي يتعامل مع الاستجابات الثنائية؛ لذا يعد هذا النموذج الأمثل للاستخدام في حالة الأسئلة متعددة التدريج (Masters, 1982)، ويتعامل نموذج التقدير الجزئي (Partial Credit Model) مع مواقف اختبارية مختلفة تتضمن المواقف التي تحتوي على فقرات تكون الإجابة عنها متدرجة، تظهر بشكل واضح في الأسئلة التي تتطلب عدداً من الخطوات، أي تكون فيها الإجابة متعددة الخطوات (Polytomous). فالسؤال الذي يتألف من ست مستويات أو ستة مهمات جزئية أو خطوات، يصل المفحوص إلى أي مستوى منها بعد أن يقوم بالخطوة التي تؤدي إلى هذا المستوى، فلا يصل إلى المستوى الأول إلا بعد أن يقوم بالخطوة الأولى، ولا يصل إلى المستوى الثاني إلا بعد أن يقوم بالخطوة الثانية، وهكذا يصل المفحوص إلى المستوى الخامس بعد أن يقوم بالخطوة الخامسة.

وهناك أسلوبان للتعامل مع خطوات السؤال (مستويات التدريج)، فالأسلوب الأول يعدّ امتداداً لأسلوب الفترات المتتابعة (Successive Intervals) لثيرستون في تقدير معالم الصعوبة للفواصل (Thresholds) بين مختلف المستويات، حيث تكون درجة صعوبة أي فاصل أعلى من صعوبة الفاصل الذي يسبقه؛ حيث يستعمل الرمز (dij) للدلالة على صعوبة الفاصل بين المستوى (j) والمستوى (j-1) للسؤال (i). والأسلوب الآخر يتعامل على أساس أن كل مستوى

من مستويات السؤال (كل خطوة من خطوات السؤال) عبارة عن سؤال ثنائي التدريج تكون الإجابة عنها إما (١) عندما يصل إليها المفحوص أو (صفراً) عندما لا يصل إليه المفحوص، ويتعامل مع صعوبة كل خطوة من خطوات السؤال الواحد بشكل منفصل مقارنة بالخطوة السابقة لها؛ حيث يستعمل الرمز (δij) للدلالة على صعوبة الخطوة (j) في السؤال (i) (Masters, 1982)، ويتعامل نموذج التقدير الجزئي (Partial Credit) مع هذا الأسلوب، وهو أسلوب البحث في هذه الدراسة.

فقد قام ماسترز (Masters, 1988) بإعطاء الرمز πkni للدلالة على احتمال أن يقع الفرد (n) ذو القدرة (θn) في المستوى (k) بالنسبة للسؤال (i)، وافترض أن احتمال وصول الفرد إلى مستوى (خطوة) لا يعتمد على الخطوات السابقة له بل يعتمد فقط على درجة صعوبة هذه الخطوة بالإضافة إلى قدرة الفرد، بل افترض أن احتمال تجاوز الفرد (n) للخطوة (k) يعتمد على مستوى صعوبة هذه الخطوة (δik) وعلى قدرة الفرد (θn)، وعبر عن ذلك باستعمال الرمز ϕkni حيث استعمل التعريف الآتي لذلك:

$$\phi_{kni} = \frac{\pi_{kni}}{\pi_{(k-1)ni} + \pi_{kni}} = \frac{e(\alpha_i)\ (-\delta ik)}{1+e(\alpha_i)(\theta n - \delta ik)}$$

حيث: (δij) درجة صعوبة الخطوة k وتفصل المستوى k عن المستوى (k-1): αi درجة التمييز للفقرة i،

θn: قدرة الفرد πkni: احتمال أن يقع الفرد (n) ذو القدرة θn في المستوى (k) بالنسبة للسؤال (i)،

π(k−1)ni: احتمال أن يقع الفرد (n) في المستوى (k-1) بالنسبة للسؤال (i)،

ϕkni: احتمال إنجاز الفرد (n) للخطوة (k) في السؤال (i) ليأخذ العلامة (k) بدلاً من (k-1)، ومن هنا يمكن كتابة العلاقتين الآتيتين لتوضيح الاحتمالات المرتبطة بمثال مكون من خطوتين هي: Φ2ni, Φ1ni، حيث:

$$\phi_{1ni}=\frac{\pi_{1ni}}{\pi_{0ni}+\pi_{1ni}}=\frac{e(\alpha_i)(\theta_n-\delta_{i1})}{(\alpha_i)(\theta_n-\delta_{i1})} \qquad \phi_{2ni}=\frac{\pi_{2ni}}{\pi_{1ni}+\pi_{2ni}}=\frac{e(\alpha_i)(\theta_n-\delta_{i2})}{1+e(\alpha_i)(\theta_n-\delta_{i2})}$$

والمعادلة الخاصة بالخطوة الأولى توضح احتمال نجاح الفرد (n) على السؤال (i) عندما تقتصر نتيجة الاستجابة على بديلين أو قيمتين فقط (ثنائي التدريج)، كما هو الحال في نموذج راش ذي الاستجابات الثنائية التدريج. وحيث $(\pi 0\mathrm{ni}+\pi 1\mathrm{ni})$ تساوي (١) فإن:

$$\pi_{1ni}=\pi_{1ni}=\frac{e(\alpha_i)(\theta_n-\delta_{i1})}{1+e(\alpha_i)(\theta_n-\delta_{i1})}$$

وهذا التعبير في المعادلة يطابق نموذج راش ذا الخطوة الواحدة، وأما إذا كان هناك أكثر من مستويين للاستجابة، أي متعدد المستويات (متعدد الخطوات) كانت قيمة $(\pi 0\mathrm{ni}+\pi 1\mathrm{ni})<1$ وهذا هو الفارق الوحيد بين المعادلة الخاصة بالخطوة الأولى والصيغة التي تمثل الصيغة الرياضية لنموذج راش. ففي حالة السؤال الذي يتكون من ثلاث خطوات يكون $(\pi 0\mathrm{ni}+\pi 1\mathrm{ni}+\pi 2\mathrm{ni}+\pi 3\mathrm{ni}=1)$ وفي حالة السؤال الذي يتكون من خمس خطوات فإن $(\pi 0\mathrm{ni}+\pi 1\mathrm{ni}+\pi 2\mathrm{ni}+\pi 3\mathrm{ni}+\pi 4\mathrm{ni}+\pi 5\mathrm{ni}=1)$ وبشكل عام عندما يكون عدد مستويات الإجابة للسؤال i محدودة بـ $(\mathrm{m}+1)$ من المستويات، ويطلب من المفحوص (n) أن يستجيب للسؤال في أحد هذه المستويات يكون:

$$\sum_{k=0}^{m}\pi_{kni}=1$$

وبإجراء بعض الحلول الجبرية للاحتمالات الشرطية يمكن كتابة ما يساويه احتمال أن يقع الفرد (n) ذو القدرة $(\theta\mathrm{n})$ في المستوى (x) من مستويات الإجابة عن السؤال (i) ذي المستويات الستة الذي يتضمن خمس خطوات درجة صعوبتها هي: $(\delta\mathrm{i}5,\delta\mathrm{i}4,\delta\mathrm{i}3,\delta\mathrm{i}2,\delta\mathrm{i}1)$ الذي يُرمز له بالرمز $(\pi\times\mathrm{ni})$ كما يلي:

$$\pi_{xni} = \frac{\exp . \sum_{j=0}^{x} (\theta_n \delta_{ij})}{\sum_{j=0}^{mi} \exp . \sum_{k=0}^{j} (\theta_n - \delta_{ij})} = \frac{e^{\sum_{j=0}^{x} (\theta n - \delta_{ij})}}{\sum_{j=0}^{mi} e^{\sum_{k=0}^{j} (\theta_n - \delta_{ij})}}$$

حيث: mi: عدد خطوات السؤال، k: عدد المستويات للسؤال.

وتبين المعادلة السابقة النموذج الاحتمالي ($\pi \times$ ni) الذي يدل على وقوع الفرد (n) ذي القدرة (θn) في المستوى (x) من مستويات الإجابة عن السؤال (i) التي عددها يساوي (mi + 1) ويتضمن (mi) خطوة درجة صعوبتها على التوالي هي: δi5, δi4, δi3, δi2, δi1 وهذه المعادلة تعبر عن الصيغة الرياضية لنموذج التقدير الجزئي لماسترز (Masters, 1982: 159-161).

وحتى يتصف نموذج التقدير الجزئي بالقياس الموضوعي فلا بد من فصل معالم الخطوات للأسئلة عن قدرات الأفراد الذين يجيبون عنها (Person-Free)، وكذلك فصل قدرات الأفراد عن معالم الخطوات للأسئلة (Item-Free) أيضاً. فإذا كانت مستويات الإجابة مرتبة وأن المفحوصين جميعهم يجب أن يبدؤوا بالخطوة الأولى ثم ينتقلوا إلى الخطوة الثانية... وهكذا، فإن ذلك يفضي إلى أن عدد المفحوصين الذين يصلون المستوى (k) لا يمكن بأي حال من الأحوال أن يزيد على عدد المفحوصين الذين يصلون المستوى (k-1). أي أن: δi1 $\geq$ δi2 $\geq$ δi3.... $\geq$ δim؛ حيث: δim تدل على عدد المفحوصين الذين يصلون إلى المستوى (m) في السؤال (i)، إلّا أن قيم الصعوبة لخطوات السؤال لا تأخذ بالضرورة ترتيباً معيناً (Masters, 1982). وهذا لا يعني أيضاً عدم وجود أسئلة تكون صعوبتها مرتبة أي: δi1 $\geq$ δi2 $\geq$ δi3 $\geq$ δi4 $\geq$ δi5θ واعتماداً على فرضية استقلال إجابة كل فرد في مختلف أسئلة الاختبار، وتطبيقاً لنموذج التقدير الجزئي فإن احتمال حصول الفرد (n) ذي القدرة (θn أو βn) على المتجه (xni) شرط معرفة مصفوفة درجات صعوبة مختلف خطوات الأسئلة جميعها وهي (δij) يعطى بالعلاقة الآتية:

حيث

$$P\{(x_m)|\beta n,((\delta_{ij}))\} = \prod_{i=1}^{L}\left[\frac{\exp\sum_{j=0}^{xni}(\beta n-\delta_{ij})}{\sum_{k=0}^{mi}\exp\sum_{j=0}^{k}(\beta n-\delta_{ij})}\right]$$

$$\exp\sum_{i=1}^{L}\sum_{j=0}^{xni}(\beta n-\delta_{ij})$$

$$\psi n=\prod_{i=1}^{L}=\left[\sum_{k=0}^{mi}\exp\sum_{j=0}^{k}\psi n(\beta n-\delta ij)\right]$$

حيث (L): هو عدد الأسئلة في أي اختبار، والرمز (j): يدل على عدد خطوات السؤال، فإذا دلت القيمة (mi) على عدد خطوات السؤال (i) فإن قيم المتغير (j) تأخذ أي قيمة بين القيمتين (١) و (n.(mi: تدل على أرقام الأفراد الذين تقدموا للاختبار، والرمز δij يدل على درجة صعوبة الخطوة (j) في السؤال (i) والرمز ((δij)) للدلالة على المصفوفة التي تكون مدخلاتها هي درجات صعوبة الخطوات المختلفة لمختلف الأسئلة، والرمز ((xni)) للدلالة على المصفوفة التي تكون مدخلاتها العلامات التي حصل عليها مختلف الأفراد في مختلف الأسئلة، والرمز (xni) للدلالة على المتجه الذي يدل على علامات الفرد (n) في مختلف الأسئلة، والرمز $\prod_{i=1}^{L}$: يدل على حاصل ضرب المقادير المرتبطة بالأسئلة جميعها من السؤال الأول (i = 1) إلى السؤال L. وقيمة صعوبة الخطوة δij يمكن فصلها بشكل مستقل عن قدرات الأفراد (n□) الذين يستجيبون للسؤال (Person- Free)، وكذلك يمكن فصل قدرات الأفراد بشكل مستقل عن معالم الخطوات للأسئلة (Item- Free)، وعند فصل هذه المعالم باستخدام طريقة الأرجحية العظمى غير المشروطة (Unconditional Maximum

(Likelihood Procedure وهو ما يسمى بالموضوعية النوعية (Specific Objectivity) (Masters, 1982; Masters, 1984). تصبح المعادلة لتقدير معالم صعوبة الخطوات بالشكل الآتي:

$$P\{((x_m))|(r_n),((\delta_{ij}))\} = \prod\left[\frac{\exp\left(-\sum_{j=0}^{L}\sum^{m_{ni}}\delta_{ij}\right)}{\sum\exp-\left(-\sum_{j=0}^{L}\sum^{x_{ni}}\delta_{ij}\right)}\right]$$

حيث P: هي الاحتمال الشرطي لمتجه الاستجابة (Response Vector) للحصول على مصفوفة علامات الأفراد ((xni)) على الأسئلة جميعها بمعرفة متجه علامات الأفراد (rn). ويلاحظ من المعادلة إمكانية فصل متجه القدرة (θn أو βn) لجميع الأفراد عن معالم الخطوات (δij) عن طريق إيجاد الاحتمال الشرطي للحصول على مصفوفة علامات الأفراد عن الأسئلة ((xin)) بمعرفة متجه علامات الأفراد ومصفوفة خطوات مختلف الأسئلة. ويمكن الحصول على علامة كل فرد بإيجاد مجموع العلامات التي حصل عليها كل منهم على الأسئلة في الاختبار، أو مجموع عدد الخطوات التي استطاع الفرد اجتيازها بنجاح. ويلاحظ من المعادلة أن قدرة الفرد (θn) قد اختفت من المعادلة. أي أن متجه علامات الأفراد (rn) هو الشرط الكافي لتقدير متجه قدراتهم (θn) بمعرفة مصفوفة درجات صعوبة خطوات مختلف الأسئلة. والمعادلة تبرر بموجب نموذج التقدير الجزئي تقدير درجة صعوبة كل خطوة من خطوات الأسئلة بغض النظر عن عينة الأفراد الذين تقدموا للاختبار تحت شرط معرفة متجه العلامات التي حصلوا عليها، وذلك بتطبيق مبدأ الأرجحية العظمى.

وكذلك يمكن البرهنة على أن احتمال الحصول على متجه العلامات للسؤال (i)، وهو (xni) بمعلومية المتجه الذي يدل على مجموع الأفراد في كل مستوى من مستويات السؤال (i) وهو Si الذي يعرف بالمتتالية الآتية:

(Si)= Si1, Si2, ..., Sim)، لا يعتمد على مستوى صعوبة خطوات السؤال، بل يعتمد فقط على قدرات الأفراد. والمعادلة الآتية توضح احتمال الحصول على المتجه (xni) لعلامات الأفراد على السؤال (i) بمعرفة المتجه (Si) الذي يدل على أعداد الأفراد الذين يقعون في كل من مستويات السؤال ومتجه قدرات الأفراد (θn).

$$P\{((x_m))|(S).(\beta_n)\} = \frac{P\{(x_{ni})|(\beta_n),(\delta_{ij})\}}{P\{(S)|(\beta_n),(\delta_{ij})\}} = \frac{\exp\left(\sum_{n=1}^{N} X_{ni}\beta_n\right)}{\sum_{\frac{Xni}{5}}^{(S)} \exp\left(\sum_{n=1}^{N} X_{ni}\beta_n\right)}$$

ويلاحظ من المعادلة أن درجة صعوبة الخطوات المختلفة للسؤال (i) الممثلة بالمتجه (δij) لم تظهر في الاحتمال المشروط، مما يدل على أن المتجه الذي يشير إلى أعداد الأفراد في كل مستوى من مستويات السؤال (i) وهو المتجه ((Sij)) هو الشرط الكافي لتقدير درجة صعوبة خطوات السؤال (i)، ومن هنا فإن المصفوفة ((Sij)) التي تحتوي على عدد الأفراد في كل مستوى من مستويات الأسئلة تحتوي على جميع المعلومات التي تحتويها مصفوفة درجات صعوبة الخطوات في جميع الأسئلة (δij).

ومن هذه المعادلة يمكن إيجاد الاحتمال الشرطي للحصول على مصفوفة العلامات ((xni)) بمعرفة مصفوفة ((Sij)) التي تدل على عدد الأفراد الذين وقعت إجاباتهم في المستوى (j) من السؤال (i)، وتبيان أن هذا الاحتمال لا يعتمد على درجات صعوبة خطوات الأسئلة الممثل بالمصفوفة (δij) كما في المعادلة الآتية:

$$P\{((X_m))|((S_{ij})),(\beta n)\} = \prod_{i=1}^{L} \frac{\exp.(\sum_{n-1}^{N}(X_{ni}\beta n)}{\sum_{(Xni)}^{(S)} \exp.(\sum_{n-1}^{N}(X_{ni}\beta n)}$$

حيث: N عدد الأفراد الذين تقدموا للاختبار.

ويلاحظ من المعادلة بأنها لا تعتمد على مصفوفة درجات صعوبة مختلف خطوات الأسئلة ((δij))، لذلك يمكن استعمال هذه العلاقة في تقدير معالم (قدرات) الأفراد المختلفة الممثلة بالمتجه (βn) بمعرفة مصفوفة ((δij)) دون الرجوع إلى مصفوفة درجات صعوبة مختلف خطوات الأسئلة التي يتقدم لها الأفراد، سواء كانت سهلة أم صعبة أم متوسطة. وبصورة أخرى فإن المصفوفة ((Sij)) هي الشرط الكافي لتقدير المصفوفة ((δij)). لذلك عندما تطابق البيانات التي تلائم نموذج التقدير الجزئي يصبح بالإمكان فصل معالم (قدرات) الأفراد عن معالم (صعوبة) الخطوات في الأسئلة التي تم استخدامها، وهذا ما جعل نموذج التقدير الجزئي حالة خاصة من نموذج راش.

الدراسات السابقة

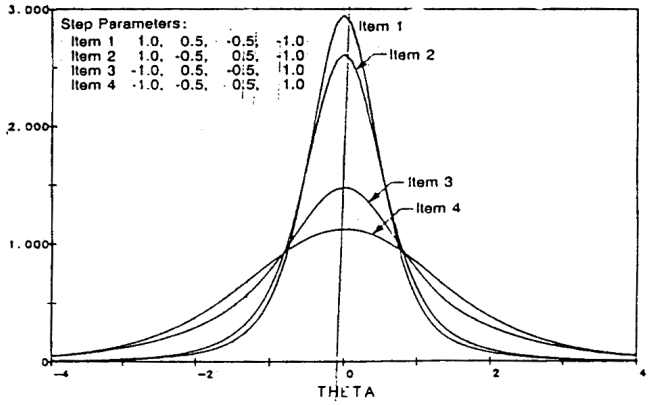
ومن الدراسات التي تناولت نموذج التقدير الجزئي الخاص بالأسئلة متعددة التدريج، الدراسة التي قام بها رايت وماسترز (Wright and Masters, 1982) باستخدام نموذج التقدير الجزئي لبيانات حول المعرفة الفيزيائية تم جمعها عن طريق نموذج "أجب حتى تصل إلى الإجابة الصحيحة" (Answer Until Correct)، تكون الاختبار من (٣٠) فقرة من نوع الاختيار من متعدد، وتكونت عينة الدراسة من (٣٢١) مفحوصاً، حيث قام الباحثان بعمل ثلاثة نماذج للتصحيح: النموذج الأول: يحصل فيه المفحوص على العلامة (١) إذا كانت الإجابة صحيحة من المحاولة الأولى، ويحصل على العلامة صفر لغير ذلك، والنموذج الثاني: يحصل فيه المفحوص على العلامة (١) إذا كانت الإجابة صحيحة من المحاولة الأولى أو الثانية، ويحصل على العلامة صفر لغير ذلك، بينما النموذج الثالث: يحصل فيه المفحوص على العلامة (٢) إذا كانت الإجابة الصحيحة من المحاولة الأولى والعلامة (١) إذا كانت الإجابة الصحيحة من المحاولة الثانية ويحصل المفحوص على العلامة صفر لغير ذلك، واستخدما في تحليل البيانات نموذج راش الخاص بالفقرات ثنائية التدريج للنموذجين الأول

والثاني، ونموذج التقدير الجزئي الخاص بالأسئلة متعددة التدريج للنموذج الثالث. وأشارت نتائج الدراسة إلى زيادة متوسط تقديرات قدرات المفحوصين بمقدار (١,٤٦) وحدة لوغاريتمية في النموذج الثاني على النموذج الأول حيث كانت معظم الأسئلة في النموذج الثاني أسهل منها في النموذج الأول، وإن ترتيب درجة الصعوبة للسؤال الواحد في النماذج الثلاثة متطابق تقريباً، حيث لم تتأثر قيمة صعوبة الفقرة في النماذج الثلاثة، وكذلك الى عدم اختلاف معامل التباعد للأسئلة (Item Separation Index) في النموذج الثالث عنه في النموذج الأول، وسبب ذلك وجود عدد من الأفراد ذوي القدرة المتدنية الذين يمكن أن يستفيدوا من المحاولة الثانية للإجابة، كما أشارت نتائج الدراسة إلى أن أخطاء القياس للأفراد ذوي القدرة المنخفضة في النموذج الثالث أقل من أخطاء القياس في النموذج الثاني، وأن أخطاء القياس للأفراد ذوي القدرة العالية كانت متشابهة تقريباً، أي أن النموذج الثالث (نموذج التقدير الجزئي) أكثر دقة وفاعلية في قياس قدرات الأفراد ذوي القدرة المنخفضة، بالاضافة الى ذلك فإن الكفاءة النسبية للنموذج الثالث بالنسبة للنموذج الأول تساوي (١,٥) وذلك عند مستوى القدرة ($\theta = -2$)، بمعنى أن الأفراد ذوي القدرة ($\theta = -2$) يحتاجون إلى ١٥ سؤالاً في النموذج الأول لإعطاء كمية المعلومات نفسها التي تنتج عن (١٠) أسئلة في النموذج الأول.

وفي دراسة أخرى قام بها دود وكوخ (Dodd and Koch, 1985) تناولت التغيرات التي تحدث في قيم الخطوات للفقرات ودالة المعلومات للاختبار باستخدام نموذج التقدير الجزئي (Partial Credit)؛ حيث تم استخدام أسئلة متعددة الخطوات/ التدريج. خلصت الدراسة إلى أن دالة المعلومات للفقرات تختلف باختلاف الفاصل بين صعوبة الخطوة الأولى والخطوة الأخيرة. فإذا كان الفاصل صغيراً، فإن الأسئلة تعطي الحد الأقصى من المعلومات ولكن في نطاق ضيق على متصل السمة الكامنة، وإذا كان الفاصل تقريباً (Proximity) كبيراً فالأسئلة تعطي معلومات أقل ولكن في نطاق أوسع على متصل السمة الكامنة،

وإذا كانت الأسئلة واقعة ضمن تقديرات صعوبة الخطوات، فإن الأسئلة تعطي الحد الأقصى من المعلومات لمستويات السمة الكامنة.

وحول النقطة نفسها، قام الباحثان دود وكوخ (Dodd and Koch, 1987) بتقصي العلاقة بين توزيع قيم دالة المعلومات للسؤال وبين ترتيب قيم صعوبة خطوات السؤال في نموذج التقدير الجزئي. وتحقيقاً لغرض الدراسة فقد استخدم الباحثان بيانات مولّدة (Simulated) لنوعين من الأسئلة: أسئلة بأربع خطوات قيم صعوبتها (-١، -٠,٥، ٠,٥، +١) وأسئلة بثلاث خطوات تكون قيم صعوبة خطواتها (-١، صفر، +١)، حيث تم إيجاد (٣٠) سؤالاً مختلفاً منها (٢٤) مرتبة قيم الصعوبة لها بأربع خطوات، و(٦) أسئلة بثلاث خطوات. وباستخدام نموذج التقدير الجزئي، تم توليد (٢٤٠) سؤالاً لتلائم النموذج، نصفها يعطي دالة معلومات حادة (Peaked)، والنصف الآخر يعطي دالة معلومات مفلطحة (Flatted)، وكذلك تم توليد استجابات لعينة من (١٠٠٠) مفحوص. أشارت نتائج الدراسة إلى أن الأسئلة المتماثلة في عدد الخطوات تعطي المساحة ذاتها تحت منحنى دالة المعلومات بالنسبة لنوعي الأسئلة، (ذات الأربع خطوات والثلاث خطوات)، في حين أن الأسئلة ذات الخطوات الأربع أعطت كمية من المعلومات أكبر عبر متصل السمة الكامنة من الأسئلة ذات الثلاث خطوات. وكذلك فيما يتعلق بدالة المعلومات الخاصة بالفاصل بين الخطوة الأولى والخطوة الأخيرة في الأسئلة وكذلك ترتيب قيم صعوبة الخطوات، فقد أشارت النتائج إلى أن الأسئلة التي تكون فيها الفواصل متساوية، فإن دالة المعلومات تكون أكثر تفلطحاً إذا كانت خطوات الأسئلة مرتبة من الأسهل للأصعب أي من (-١، -٠,٥، ٠,٥، +١)، وتكون دالة المعلومات أكثر حدة إذا كانت خطواتها مرتبة من الأصعب للأسهل (+١، ٠,٥، -٠,٥، -١) كما هو مبين في الشكل رقم (١) الذي يبين العلاقة بين دالة المعلومات وترتيب صعوبة الخطوات للفقرات.



الشكل رقم (١) - دالة المعلومات لأربعة أسئلة لها الفواصل نفسها بين الخطوة الأولى والخطوة الأخيرة ولكنها في عدد الخطوات التي تخرج عن الترتيب المطرد لصعوبة خطوات السؤال

حيث يلاحظ من الشكل رقم (١) أن الفقرة (١) كانت الأكثر حدّة؛ لأن خطواتها مرتبة من الأصعب إلى الأسهل (+١، ٠,٥، ٠,٥-، -١)، في حين أن الفقرة (٤) كانت الأكثر تفلطحاً؛ لأن خطواتها مرتبة من الأسهل إلى الأصعب (-١، -٠,٥، ٠,٥، +١). وأما الأسئلة التي يكون فيها الفاصل بين صعوبة الخطوة الأولى وصعوبة الخطوة الأخيرة غير متساو، ولكن عدد خطواتها متساو، فإن الأسئلة التي يكون فيها الفاصل صغيراً تعطي دالة معلومات أكثر حدّة من الأسئلة التي يكون فيها الفاصل كبيراً. وكلما كانت هذه الفواصل كبيرة كانت دالة المعلومات أكثر حدة. أما دراسة ولسون وايفينتوش (Wilson & Iventosch, 1988) فقد هدفت إلى دراسة الاستجابات لاختبارات جزئية متعددة الخطوات ومرتبة خطواتها من الأسهل إلى الأصعب باستخدام نموذج التقدير الجزئي. تكونت عينة الدراسة من (٢٥٧) طالباً من صفوف الرابع والسادس والثامن، جمعت البيانات اللازمة للدراسة بوساطة سبعة أسئلة من اختبار لقياس القدرة العقلية. أشارت نتائج هذه الدراسة إلى أن نموذج التقدير الجزئي يقدم كمية من المعلومات أكبر، وأكثر تفصيلاً من نموذجي راش ومقياس جتمان وذلك في مجالي عدم الملاءمة وفي الاستقلالية الموضعية للاستجابة.

وأجرى هاريس ولان وموسينسون (Harris, Laan & Mossenson, 1988) دراسة هدفت إلى بناء اختبارات الكتابة القصصية باستخدام نموذج التقدير الجزئي. تم تطوير خمسة اختبارات لصفوف مختلفة من الصف الرابع وحتى الصف الثاني عشر، وكل منها يتطلب بناء قصة، وتم إعطاء التقديرات لهذه الكتابات القصصية باستخدام ثمانية مقاييس وكل منها مدرج بتسع درجات/ خطوات. حيث تمت معاملة كل مقياس من هذه المقاييس سؤالاً مستقلاً. وتم تحليل البيانات باستخدام نموذج التقدير الجزئي، وأشارت النتائج إلى فاعلية نموذج التقدير الجزئي في تدريج الأفراد والأسئلة بحيث يمكن التنبؤ بعلامات الأفراد على المقياس وبدرجة عالية من الدقة من خلال الدرجة الكلية لقدرته الكتابية.

كما أجرى آدمز (Adams, 1988) دراسة بعنوان تطبيق نموذج التقدير الجزئي للتشخيص التربوي، ولتحقيق هدف الدراسة فقد تم تطبيق اختبار تشخيصي في الرياضيات مكون من فقرات اختيار من متعدد بحيث أعطيت البدائل علامات تتناسب وقربها من الإجابة الصحيحة، وتم معاملة هذه البدائل وكأنها خطوات. تكونت عينة الدراسة من (٥٦٠) فرداً تراوحت أعمارهم ما بين (٩-١٠) سنوات، وتم استخدام نموذج التقدير الجزئي لتحليل البيانات. أشارت نتائج الدراسة إلى فاعلية نموذج التقدير الجزئي في تفسير درجة الطالب على الاختبار وذلك بدلالة المهمات (الخطوات) التي تعبر عن مستويات المهارة والفهم من خلال استجابة الطالب.

وفي دراسة قام بها التقي (١٩٩٢) لفحص مدى تحقيق افتراضين أساسيين من افتراضات نظرية السمات الكامنة في أسئلة المقال (متعددة الخطوات)، وهذان الافتراضان هما: وجود لا تغير في تقدير درجات صعوبة الأسئلة من خلال استعمال عينات مختلفة من الأفراد، ووجود لا تغير في تقدير قدرات الأفراد من خلال التقدم لاختبارات تتضمن أسئلة متفاوتة في درجات الصعوبة، وتم فحص الافتراضين أعلاه من خلال استعمال نموذج التقدير

الجزئي ونموذج سلم التقدير. أما الأداة المستخدمة في عملية الفحص، فهي اختبار في الرياضيات مكون من (٢٠) سؤالاً يقيس قدرة طلاب الصف الأول الثانوي العلمي على حل المعادلات الجبرية والمثلثية في مجهول واحد وفي مجهولين، وتم استخدام برمجيتين حاسوبيتين هما (MSTEPS) و(MSCALE). أشارت نتائج الدراسة إلى تحقق الافتراضين عند استعمال نموذج التقدير الجزئي، بينما لم يتحقق هذان الافتراضان عند استعمال نموذج سلم التقدير، وأوصى الباحثان باستخدام نموذج التقدير الجزئي لبناء أسئلة الاختبارات التحصيلية المقالية.

وفي دراسة شاهين (١٩٩٣) التي هدفت إلى استقصاء أثر عدد خطوات تصحيح الأسئلة المقالية في تقدير قدرات الأفراد وقيم صعوبة الأسئلة باستخدام نموذج التقدير الجزئي، تم إعداد اختبار تحصيلي في العمليات على الكسور مكون من (٤٠ سؤلاً)، (٤) منها تصحح بخطوة واحدة، و(١٦) منها تصحح بخطوة أو خطوتين و(٢٠) منها تصحح بخطوة واحدة أو خطوتين أو ثلاث خطوات أو أربع خطوات. طبق الاختبار على عينة من (٥٢٩) طالباً وطالبة من طلبة الصف السادس الابتدائي. صححت النتائج بطرق أربع: الطريقة الأولى: أن يعطى الطالب العلامة (١) لكل إجابة صحيحة والعلامة (٠) لكل إجابة خاطئة - والطريقة الثانية: أن يعطى الطالب العلامة (١) إذا اجتاز الخطوة الأولى، وأن يعطى العلامة (٢) إذا اجتاز الخطوة الثانية، ويعطى العلامة (٠) لغير ذلك عن الأسئلة التي تصحح بخطوة أو خطوتين أما الطريقة الثالثة: فهي أن يعطى الطالب العلامة (١) إذا اجتاز الخطوة الأولى، والعلامة (٢) إذا اجتاز الخطوة الثانية، والعلامة (٣) إذا كانت الإجابة صحيحة، ويعطى العلامة (٠) لغير ذلك عن الأسئلة التي يمكن أن تصحح بخطوة أو اثنتين أو ثلاث خطوات - والطريقة الرابعة هي أن يعطي الطالب العلامة (١) إذا اجتاز الخطوة الأولى، والعلامة (٢) إذا اجتاز الخطوة الثانية، والعلامة (٣) إذا اجتاز الخطوة الثالثة، والعلامة (٤) إذا اجتاز الخطوة الرابعة، والعلامة (٠) لغير ذلك عن الأسئلة التي يمكن أن تصحح بخطوة أو اثنتين أو ثلاث خطوات أو أربع

خطوات. وتم استخدام البرمجية (MSTEPS) التي أعدت لتحليل هذا النوع من الاختبارات، وتم الحصول على تقديرات لقدرات الأفراد وقيم صعوبة الأسئلة والأخطاء المعيارية في قياسها وإحصائيات الملاءمة في كل من الطرق الأربع. واستخرجت دلالات الفروق بين طرق التصحيح الأربع وتم التوصل إلى النتائج الآتية: عدم وجود فروق ذات دلالة إحصائية في تقديرات القدرة للأفراد وعند مستويات القدرة جميعها بين طرق التصحيح الأربع، وأن الخطأ المعياري في قياس قدرات الأفراد وفي تقديرات قيم الصعوبة للأسئلة يتناقص بزيادة عدد خطوات تصحيح الأسئلة، وكذلك الى وجود فروق ذات دلالة إحصائية في تقديرات قيم صعوبة الأسئلة، حيث تصبح الأسئلة أكثر سهولة بزيادة عدد خطوات تصحيحها، كما أن الكفاءة النسبية للاختبار تزداد بزيادة عدد خطوات التصحيح وعند مستويات جميعها القدرة، بمعنى أن قيمة دالة المعلومات تزداد بزيادة عدد الخطوات.

وأجرت سوزان بيريتفاز (Beretvas, 2001) دراسة لتقصي فاعلية ثلاثة نماذج في نظرية الاستجابة للفقرة في تقدير قدرات الأفراد ودرجة الصعوبة للفقرات باستخدام نموذج راش ذي المعلمة الواحدة، ونموذج التقدير الجزئي، ونموذج الاستجابة المتدرجة. وقامت الباحثة ببناء اختبارات محكية المرجع تتضمن مزيجا من الفقرات من نوع الاختيار من متعدد، وأسئلة متعددة التدريج مكونة من أربع خطوات، وتم تدريج الفقرات باستخدام النماذج اللوجستية: الأحادي المعلمة، والثنائي المعلمة، والثلاثي المعلمة، وكذلك تدريج الفقرات متعددة التدريج باستخدام نموذج التقدير الجزئي. أشارت نتائج الدراسة إلى فاعلية نموذج التقدير الجزئي في معايرة الأسئلة من النماذج اللوجستية، وذلك عند مستويات القدرة العالية، في حين أن النماذج اللوجستية كانت أكثر فاعلية في معايرة الفقرات عند مستويات القدرة المتدنية.

وفي دراسة أخرى أجراها الزعبي (٢٠٠٣) حول فاعلية نموذج التقدير الجزئي في بناء فقرات بنك أسئلة متعددة الخطوات في مادة الكيمياء لطلبة

الصف الثاني الثانوي العلمي، تم استقصاء فاعلية نموذج التقدير الجزئي في تقدير قدرات الأفراد، ومعايرة أسئلة متعددة الخطوات، استخدم الباحثان برنامج (BIGSTEPS) في تحليل النتائج. ولتحقيق هدف الدراسة تم بناء اختبار كيمياء مكون من ٢٠ سؤالاً مختلفة في عدد خطواتها، وتم تطبيقها على عينة مكونة من ٧٧١ طالبا وطالبة، وتوصلت الدراسة إلى أن الخطأ المعياري يزداد بزيادة عدد خطوات الاستجابة، وعدم وجود فروق ذات دلالة إحصائية في متوسطات الفروق المعايرة لتقديرات معالم الصعوبة للأسئلة، ووجود فروق دالة إحصائيا بين تقديرات القدرة المعايرة، وأن الكفاءة النسبية للاختبار بشكل عام تزداد بزيادة عدد خطوات الاستجابة.

كما أجرى الشريفين (٢٠٠٣) دراسة هدفت إلى الكشف عن مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس. ولتحقيق هدف الدراسة تم بناء اختبارين تحصيليين في الفيزياء: أحدهما ثنائي التدريج مؤلف من ٧٥ فقرة من نوع الاختيار من أربع بدائل، والآخر متعدد التدريج مؤلف من ٢٠ سؤالا، كل سؤال مكون من أربع خطوات. أما عينة الدراسة فتكونت من ١٠٠٣ طالباً وطالبة، تم اختيار المدارس بصورة عشوائية، وتم اختيار عينة الدراسة بالطريقة العشوائية العنقودية، وتم استخدام البرامج الإحصائية (SPSS MINSTEPS,, BIGSTEPS) للحصول على تقديرات لقدرات الأفراد، وقيم الصعوبة لأسئلة الاختبارين، والأخطاء المعيارية في قياسها، وإحصائيات المطابقة الداخلية والخارجية، ومتوسطات المربعات الداخلية والخارجية، وقيم معاملات الارتباط ثنائي النقطة (Point biserial coefficient of correlation)، ومعاملات الثبات للأفراد والفقرات. أشارت النتائج إلى أن متوسط الكفاءة النسبية للاختبار متعدد التدريج مقارنة بالاختبار ثنائي التدريج بلغ ٣,٨، وهذا يعني أن قيمة دالة المعلومات للاختبار متعدد التدريج تساوي ٣,٨ قيمة دالة المعلومات للاختبار ثنائي التدريج، أي أن الاختبار متعدد التدريج يزودنا بمعلومات أكثر من الاختبار ثنائي التدريج وعند كل مستويات القدرة،

وفيما يتعلق بفاعلية المعادلة للاختبارين الثنائي التدريج والمتعدد التدريج فقد استخدم معياران للحكم على فاعلية طرق المعادلة المستخدمة هما: معامل الصدق التقاطعي، والخطأ المعياري للمعادلة. أشارت نتائج المعادلة الأفقية التي أجريت وفق تصميم المجموعات المتكافئة أن النموذج أحادي المعلمة كان الأكثر فاعلية من طريقتي المعادلة الخطية والمئينية.

وفي دراسة أجرتها الشريف (٢٠٠٦)، هدفت الى بحث أثر عدد فئات تدريج ليكرت على خصائص المقياس وخصائص فقراته السيكومترية وفقا للنظرية الحديثة في القياس، وذلك باستخدام نموذج مقياس التقدير لموراكي (Muraki)، تكونت أداة الدراسة من ٢٦ فقرة تقيس قلق الرياضيات لدى عينة مكونة من ٣٠٢٠ طالباً وطالبة، وتم تدريج المقياس بأسلوب ليكرت الثنائي، والثلاثي، والرباعي، والخماسي، والسداسي، والسباعي، وأظهرت نتائج الدراسة أنه لا يوجد أثر لعدد فئات التدريج على: تقديرات معالم الصعوبة للفقرات، ومعالم التمييز للفقرات، باستثناء أن متوسط الأخطاء المعيارية في تقدير معالم الصعوبة ومعالم التمييز وتقدير السمة لدى الأفراد كانت تقل بزيادة عدد فئات التدريج، كما أظهرت النتائج أن دالة المعلومات للفقرة تزداد بزيادة عدد فئات التدريج، كما أن دالة المعلومات للمقياس تزداد بزيادة عدد فئات التدريج، وتقل قيمة الخطأ المعياري للمقياس بزيادة عدد فئات التدريج، وكشفت الدراسة أن التدريج السباعي يعطي معلومات أكثر عند المستويات المختلفة من السمة باستثناء المستويات العالية جداً من السمة، وأظهر التدريج الثنائي أنه غير دقيق في اعطاء معلومات كافية عن خصائص المقياس، وخصائص فقراته مقارنة بباقي التدريجات، وأظهرت النتائج أن أنسب عدد من الفئات لتحقيق اللاتغير في تقديرات معالم الفقرة وتقديرات السمة باختلاف الفقرات والأفراد هو التدريج السداسي.

كما قام هاريس ولان وموسينسون (Harris, Laan, & Mossenson, 2009) بدراسة هدفت إلى تطبيق أسلوب التقدير الجزئي في بناء اختبارات الكتابة الروائية، حيث تم كتابة مجموعة من المهام الروائية التي تستخدم مع الأطفال من سن الثالثة وحتى العاشرة استخدم نموذج التقدير الجزئي في معايرتها، وتم تقييم

القصص الروائية التي قام الطلبة بكتابتها باستخدام ثمانية مقاييس (مثل: الجلسة، استخدام التماسك، بناء القصة)، كما تم التعامل مع هذه المقاييس كثماني فقرات. بينت النتائج دقة التقدير باستخدام نموذج التقدير الجزئي، وأن مقياس الكتابة المعاير يتوجب استخدامه كمقياس عالمي لقياس كفاءة الكتابة الروائية؛ لأن بإمكانها تفسير القدرة على الكتابة الروائية بشكل دقيق وحسب الخصائص الأكثر احتمالا لكتابة كل طالب نظراً لدقة أسلوب التحليل المستخدم.

ومن خلال هذا العرض لنتائج الدراسات السابقة، التي تناولت تحليل البيانات باستخدام نموذج التقدير الجزئي فقد تبين أن معظم الدراسات أشارت إلى أن نموذج التقدير الجزئي هو الأكثر دقة وفاعلية في قياس قدرات الأفراد ذوي القدرة المنخفضة، وأن الكفاءة النسبية للاختبار المبني بخطوات متعددة أكبر من الاختبار المبني بخطوتين. وأن نموذج التقدير الجزئي يقدم معلومات أكثر وأكثر تفصيلاً من نموذجي راش ومقياس جتمان، وذلك في مجالي عدم الملاءمة وفي الاستقلالية الموضعية للاستجابة. وكذلك فاعلية نموذج التقدير الجزئي في تدريج الأفراد والأسئلة بحيث يمكن التنبؤ بعلامات الأفراد على المقياس وبدرجة عالية من الدقة من خلال الأسئلة ذات الخطوات المتعددة. كما أن شروط موضوعية القياس وهي: (اختبار متحرر من الفقرات، اختبار متحرر من الأفراد) يتحققان بشكل أكبر عند استعمال نموذج التقدير الجزئي. لذا أوصى الباحثون باستخدام نموذج التقدير الجزئي، عند تقنين الاختبارات التحصيلية المتعددة الخطوات، وكذلك فإن المعلومات التي تعطيها الأسئلة تختلف باختلاف الفاصل بين الخطوة الأولى والأخيرة، فتكون المعلومات في حدها الأقصى عندما يكون الفاصل صغيراً وتغطي حيزاً محدوداً على متصل السمة (القدرة). وتكون في حدها الأدنى عندما يكون الفاصل كبيراً وتغطي حيزاً أوسع على متصل السمة. وأن الأسئلة التي لها العدد نفسه من الخطوات تعطي الكمية نفسها من المعلومات تقريباً، وأنه بزيادة عدد الخطوات يزداد حجم المعلومات. كما أن دالة المعلومات تكون أكثر تفلطحاً إذا كانت الخطوات مرتبة من الأسهل للأصعب، وتكون أكثر حدة إذا كانت الخطوات

مرتبة من الأصعب للأسهل. وذلك عندما تكون الفواصل بين الأسئلة متساوية، وإذا كانت غير ذلك (غير متساوية) فإن الأسئلة التي يكون فيها الفاصل صغيراً تعطي دالة معلومات أكثر حدّة من الأسئلة التي يكون فيها الفاصل كبيراً. أي كلما كانت الفواصل كبيرة كانت دالة المعلومات أكثر حدة. لذا جاءت هذه الدراسة استكمالاً لجهود الباحثين السابقين لتقصي العلاقة بين عدد خطوات الأسئلة متعددة التدريج وشكل التوزيع لصعوبة الفقرات والتفاعل بينهما على تقديرات القدرة للأفراد، والصعوبة للأسئلة، ودالة المعلومات للاختبار وفق نموذج التقدير الجزئي.

مشكلة الدراسة

برزت الدعوات المتعددة لبناء اختبارات تتضمن أسئلة ذات إجابة مصوغة تراعي المعرفة الجزئية بجانب الفقرات الموضوعية، حيث إن الفقرات ثنائية التدريج تضع قيوداً غير مقبولة في الواقع التربوي الميداني في كثير من الأحيان، وتحد من صيغ التقويم التي ينادي فيها خبراء القياس في الوقت الحاضر. وكذلك فإن تقويم مهارات حل المشكلات، وتنفيذ إجراءات، وكتابة مقال، وتجميع جهاز، يتطلب فقرات اختبارية متعددة التدريج، وكذلك ارتفعت الدعوات لبناء بنوك الأسئلة. لذا فقد ظهرت في الدراسات الحديثة عمليات معايرة الأسئلة الإنشائية، واستخراج معالم وإحصائيات لها تظهر على السطح، وتتحدى علماء القياس والتربويين في إيجاد النماذج التي تتعامل مع هذا النوع من الأسئلة. فاعتماداً على معالم الأسئلة، والإحصائيات المستخرجة لها، يمكن بناء اختبارات تتلاءم وقدرات الأفراد؛ إذ يمكن وضع اختبارات سهلة تتلاءم وقدرات الطلبة المتدنية، واختبارات صعبة تتلاءم وقدرات الطلبة العالية، وباستخدام الأساليب المناسبة يمكن تقدير قدرات الأفراد، وحتى تكون تلك التقديرات دقيقة لا بد أن يكون مستوى صعوبة الأسئلة ملائما لمستوى قدرة الفرد المتقدم لها، بحيث يكون الخطأ في التقدير أقل ما يمكن، وبالتالي عملية القياس أدق ما يمكن.

ولنموذج التقدير الجزئي أهمية كبيرة في الأسئلة التي تراعي المعرفة

الجزئية، حيث يلجأ واضع السؤال إلى تجزئة الإجابة إلى خطوات (أجزاء) تعطى درجات جزئية على هذه الخطوات، تشكل في مجموعها العلامة الكلية على السؤال. ولكن السؤال المطروح هو كم عدد الخطوات لهذا النوع من الأسئلة، التي تعطي تقديراً دقيقاً لقدرة المفحوص في السمة المقيسة؟ فهل الزيادة في عدد الخطوات تؤدي إلى دقة أكبر في تقدير قدرة المفحوص في السمة المقيسة؟ لا سيما أن إعطاء الدرجة صفر أو واحد لا يعطي المفحوص الفرصة بالتعبير عما يسمى بالمعرفة الجزئية. فهل تجزئة السؤال إلى عدد من الخطوات تؤدي إلى اختلاف في مستوى الصعوبة للسؤال؟ حيث إن احتمال الإجابة عن السؤال هو حصيلة التفاعل بين قدرة المفحوص ومستوى صعوبة السؤال، وتعد هذه التساؤلات هي المشكلة الرئيسة لهذه الدراسة، وبالتحديد فإن هذه الدراسة هدفت للإجابة عن الأسئلة الآتية:

١ - ما أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، في دقة تقدير معالم القدرة للأفراد؟

٢ - ما أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، في دقة تقدير معالم الصعوبة للأسئلة؟

٣ - ما أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، في ثبات تباعد الأفراد؟

٤ - ما أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، في دالة المعلومات للاختبار؟

٥ - هل تختلف الكفاءة النسبية باختلاف عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)؟

أهمية الدراسة

تتمثل أهمية هذه الدراسة في جانبين الأول نظري: يتوقع من خلال هذه الدراسة استقصاء أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) في دقة التقدير لمعالم الأسئلة، ودقة التقدير لمعالم القدرة للأفراد، وكذلك استقصاء أثر شكل توزيع الصعوبة للخطوات(طبيعي، ملتو نحو اليمين، ملتو نحو اليسار) في دقة التقدير لمعالم الأسئلة، ودقة التقدير لقدرات الأفراد، والكشف عن العلاقة بين دالة المعلومات للسؤال من جهة، وكل من: عدد خطوات السؤال، وشكل توزيع الصعوبة للخطوات من جهة أخرى. كما يتوقع أن تكشف النتائج عما إذا كانت الكفاءة النسبية للاختبار تختلف باختلاف كل من: عدد خطوات السؤال، وباختلاف شكل توزيع الصعوبة للخطوات، وذلك من خلال بيانات محاكاة نظراً لصعوبة الحصول على فقرات بهذه المواصفات. وأما الجانب الثاني فهو تطبيقي: إذ يتوقع أن توفر هذه الدراسة أدلة إمبريقية من خلال أساليب الإحصاء الاستدلالي؛ بمعنى محاولة التوصل إلى تبريرات عملية تقدم إلى الباحثين يمكن اعتمادها بناءاً على أساس منطقي حول الطريقة الفضلى التي يمكن استخدامها من قبل مطوري الاختبارات لمواصفات الأسئلة التي تراعي التقدير الجزئي، كما يمكن أن يستفيد مطورو الاختبارات الذين يلجؤون إلى نظرية الاستجابة للفقرة لحل المشكلات المتعلقة ببناء الاختبارات لاختيار الطريقة الأنسب عند تقدير قدرات الأفراد في كيفية بناء هذا النوع من الاختبارات، ومواصفات تلك الأسئلة بحيث تتحقق فيها شروط القياس الموضوعي.

التعريفات الاصطلاحية والإجرائية

- **نموذج التقدير الجزئي** (Partial Credit Model (PCM)): أحد نماذج النظرية الحديثة في القياس، ويعد توسيعاً لنموذج راش (نموذج المعلمة الواحدة)؛ أي

أنه نموذج سمة كامنة يناسب الأسئلة التي تتطلب خطوات متعددة (Polytomous) مثل حل المسائل الحسابية، حيث يكون من الأفضل تعيين درجات جزئية عندما تستكمل خطوات متعددة في عملية الحل قام بتطويره ماسترز (١٩٨٢) في أستراليا.

- **دالة المعلومات للفقرة:** دالة رياضية تبين مدى إسهام الفقرة في دالة المعلومات للاختبار(للمقياس) بشكل مستقل عن الفقرات الأخرى.
- **دالة المعلومات للاختبار:** دالة رياضية تعبر عن مجموع دوال المعلومات لجميع فقرات الاختبار عند مستوى معين من القدرة.
- **معالم الأفراد:** وهي معالم قدرات الأفراد، وهي مقدار ما يمتلكه الفرد من السمة التي يتم قياسها تبعاً لنظرية السمات الكامنة.
- **تقدير المعالم للفقرات والأفراد:** هي عملية التعبير الكمي عن المعالم.
- **الاختبار متعدد التدريج** (Polytomous): اختبار تكون الاستجابة عن فقراته متدرجة، إذ قد تكون إجابة الفرد غير كاملة، فقد يحصل الفرد على علامة جزئية مثل العلامة ٣ من ٤ أو العلامة ١ من ٤، حيث يلجأ واضع السؤال عادة إلى تجزئة الإجابة إلى خطوات وتعطى علامات جزئية على هذه الخطوات لتشكل في مجموعها العلامة الكلية للسؤال، وفي هذه الحالة تكون الاستجابة من النوع المتعدد المستويات (Polytomous) أي أن هذا النوع من الأسئلة يراعي المعرفة الجزئية.
- **صعوبة الخطوة:** النقطة الواقعة على محور القدرة (السمة الكامنة) التي يتقاطع فيها منحنيا استجابة لخطوتين متتاليتين؛ لذلك يشار أحياناً إليها ببارامترات تقاطع الفئات (الخطوات) (Category Intersection Parameters).
- **اللوجيت:** وحدة قياس كل من قدرة الفرد، وتعرف بأنها قدرة الفرد على النجاح على الفقرات التي تعبر نقطة صفر التدريج عن صعوبتها، عندما يساوي هذا المرجح ثابتاً هو الأساس الطبيعي e؛ أي (٢,٧٢)، ويكون عنده احتمال النجاح يساوي (٠,٧٣=٣,٧٢/٢,٧٢)، أو هو اللوغاريتم الطبيعي

لمرجح نجاح الفرد على الفقرات التي تعبر نقطة صفر التدريج عن صعوبتها، عندما يساوي هذا المرجح ثابتاً هو الأساس الطبيعي e، (٧,٧٢) ويكون عندئذ احتمال نجاحه (٠,٧٣).

توليد البيانات

تم توليد البيانات باستخدام برنامج (WINGEN)، وهو من تصميم هان وهامبيلتون (Han & Hambelton, 2007)، ويمكن من خلال هذا البرنامج توليد استجابات لاختبارات أحادية البعد (ثنائية التدريج، متعددة التدريج)، بالإضافة إلى استجابات متعددة الأبعاد. كما أنه باستخدام البرنامج يمكن توليد استجابات لأكثر من مجموعة من المفحوصين، وبتوزيعات مختلفة، وتستخدم ثلاثة توزيعات لتوليد قدرات المفحوصين وهي: التوزيع الطبيعي (Normal Distribution)، التوزيع المنتظم (Uniform Distribution)، وأخيرا توزيع بيتا (Beta Distribution)، الذي من خلاله يمكن توليد توزيعات ملتوية سواء كانت ملتوية التواءً موجباً أو التواءً سالباً. وتمت عملية توليد البيانات وفق المراحل التالية:

المرحلة الأولى: تم توليد بيانات خاصة باختبار مكون من ٣٠ سؤالاً وفق نموذج التقدير الجزئي ((Partial Credit Model (PCM) باستخدام برنامج WINGEN، وتم اختيار هذا الطول للاختبار حتى يكون مناسباً للوصول إلى دقة في التقديرات وفق نموذج التقدير الجزئي، بالإضافة إلى مناسبة هذا العدد من الأسئلة للظروف التربوية القائمة في المدارس والجامعات والمؤسسات التربوية حول الاختبارات، وتم توليد معالم فقرات الاختبار وفق الشروط التالية: تم توليد بيانات خاصة بثلاثة اختبارات مختلفة في عدد خطواتها (ثلاث خطوات، أربع خطوات، خمس خطوات)، وكل منها مكون من ٣٠ سؤالاً، بحيث كانت المسافة بين صعوبة الخطوة الأولى والخطوة الأخيرة لها متساوية، حيث تراوحت بين -٣ و+٣ لوجيت، وفق التوزيع الطبيعي Normal~(0.1)، والجدول رقم (١) يبين القيم الحقيقية لمعالم الصعوبة لفواصل الخطوات للاختبارات التي تم توليدها تبعاً لعدد الخطوات.

الجدول رقم (١)

القيم الحقيقية لمعالم الصعوبة لفواصل الخطوات للاختبارات التي تم توليدها تبعاً لعدد الخطوات

رقم السؤال	عدد خطوات السؤال													
	ثلاث خطوات			أربع خطوات				خمس خطوات						
	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الرابعة	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الرابعة	الخطوة الخامسة
١	٠,٣٦٢	٠,٥٣٨	١,٤٧٩	-٠,٠٤٩	٠,٣٥٨	٠,٤٢٤	١,٠٠٣	-١,٤٥٥	-٠,٢١٤	٠,٤٣	٠,٨٨	١,٠٦٢
٢	-٠,١٤	٠,٣٦٣	١,٠٩٦	٠,٣٢٩	٠,٤١٤	٠,٦١٤	٠,٧٢٤	-١,٨٩٢	-٠,٣٧٣	-٠,١٦٩	٠,٢٩٢	٠,٤٤٨
٣	-٠,٢٨٣	٠,٨٩	٠,٩٩٦	-٠,٩٣	-٠,٦٠١	٠,٢٣	٠,٧٠١	-١,٠٠٤	-٠,٠٩٢	٠,٧٥٣	١,٦٧٦	٣,٠٠٩
٤	-٢,٧٥٥	٠,٢٧٤	٠,٣١٦	-٠,٥٨٦	-٠,٣٩٨	-٠,٣٠٧	-٠,١٥٧	-٢,٣٤١	-١,٠٦٥	-٠,٦٣٨	-٠,٣٠٩	٠,٢٢٧
٥	٠,١٢٣	١,٥٢٥	١,٦٤٥	-٠,٦٦٣	٠,١٨	٠,٣٣٧	٠,٩١٥	-١,٣٩٦	-٠,٦١١	٠,٦٤٥	١,٣٨٩	١,٥٢
٦	-١,٩٤١	-٠,٢٦	-٠,١٣٩	-١,١٩٢	-٠,٣٣	-٠,١٥٥	١,٠٨٧	-٠,٧١٦	-٠,١٩١	١,٣٧٢	١,٩١٧	٢,١٦٣
٧	-٢,٣٢	-٢,٠٥٣	٠,٧٧٣	-٢,٠٥٢	-٠,٦٥٨	٠,٤٨٥	١,٦٣٥	-٢,٠٠٤	-١,٤٢٢	-١,٣٣٦	-٠,٩٣٨	-٠,٤١٢
٨	-٠,٠٩٥	٠,٣٣٣	٢,٥٤	-١,٢٦	-٠,٣٣٦	٠,١٠٦	٠,٧٨١	-٠,٨٤٦	-٠,٨٢٩	٠,٢٠٣	٠,٥٥٧	١,٢٠٢
٩	-٠,٢٢٩	٠,٠٠٣	٠,٧٩٤	-٢,٢٣٨	-٠,٧٣٥	-٠,٣	٠,٥٩٣	-٠,٧٢٦	٠,٠٩٩	٠,٤٥٥	٠,٨١٧	١,٥٥١
١٠	٠,٣٨٦	٠,٤٤٥	٠,٧٣٦	-١,١٠٢	٠,١١٧	٠,٤٣٤	١,٣٢٣	-١,١٥٦	-٠,٨٦٧	-٠,٥٩٢	-٠,٥٥٥	٣,٥١٨
١١	-١,٢٩٦	-٠,٦٦٢	٠,٥١	-٢,٢٦٧	-١,٧٩٢	-٠,٣٤	٠,٢٩١	-٢,٠٥٤	-٠,١٣٤	٠,١١٩	٠,١٨٥	٠,٧٣
١٢	-٠,٢٤٣	-٠,٠٩٧	٠,٥٣٥	-١,٨٤١	-٠,٥١٤	٠,٣٧	١,٤٧	-٠,٦٥٩	٠,٠٤٧	٠,٠٦٨	٠,٤١٧	٠,٧٧١

تابع/ الجدول رقم (١)

القيم الحقيقية لمعالم الصعوبة لفواصل الخطوات للاختبارات التي تم توليدها تبعاً لعدد الخطوات

رقم السؤال	عدد خطوات السؤال														
	ثلاث خطوات			أربع خطوات				خمس خطوات							
	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الرابعة	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الرابعة	الخطوة الخامسة
١٣	١,٣٣٨-	٠,٨٦٦	٢,٣٨٧	٢,٢٢٨-	١,٣٧٣-	٠,٤٢٨-	٠,٥٠٢	٠,٨٣٥-	٠,٥٧٦-	٠,٢٢٥	٠,٧٠٥	١,١٣٣
١٤	١,٢٢٣-	٠,٣	١,٢٥٣	١,١١١-	٠,٣٦٦-	٠,٠٧٤	١,٩٦٢	٠,٩٠٤-	٠,٤٧٤	٠,٥٣٥	٠,٨١٢	١,١٨٢
١٥	٠,٠٣٨	٠,٦٧٧	٢,٢٦٣	١,٠٥٥-	٠,٩٥٣	١,٨٩١	٢,٦٤١	١,٤٦٢-	١,٠٢٦-	٠,٠٦٤	٠,٧٠٤	٠,٩٢٣
١٦	١,٦٤٣-	٠,٢٦٩-	٠,٢٥٥-	١,٠١٥-	٠,٠٠٣	٠,٣٦٤	٠,٣٩٣	١,٦٧٤-	٠,٢٨٩-	٠,٠٧٤-	٠,٤٣٣	١,١٨١
١٧	١,١٤٩-	٠,٢٨٤-	٠,٨٤٥	٠,١٩٥	٠,٣٥٧	٠,٨٤	٠,٨٨٢	١,٣٩٤-	١,١٣١-	٠,٢١٨-	٠,٢٥٥	٠,٤٠٣
١٨	٠,٢٤٤-	١,١٥٤	١,٤٧٤	٢,٢٠١-	٠,٥٦١-	٠,٤١٧	١,١٢٥	٠,٨٣٥-	٠,٢٨٥-	٠,٢١٣-	٠,١٦٨-	٠,٤١٨
١٩	١,١٨٥-	٠,٦٥-	٠,١٠٧	٠,٩١٤-	٠,١٥٢-	٠,١٣٢-	٠,٢٦٢	١,٢٧١-	٠,٠٩٨-	٠,٦٤	١,٦٩٣	٢,٢١
٢٠	٠,٣٧٥-	٠,٠٩٤	٢,٣٨٩	٠,٣٠٣	٠,٣٥٢	٠,٤٣٥	٠,٩٦٩	٠,٦٨٣-	٠,٤٢٦-	٠,٠٨٩-	٠,٢٠٩	١,٢٠٧
٢١	١,٠٥-	٠,٣٥-	٠,٢٢٩	٠,٨٩٣-	٠,٣٢٥-	٠,٠٨٢-	٠,٢٦٣	٠,٢٣١-	٠,١٧١	٠,٣٧٣	٠,٦٠٨	٠,٩٥٩
٢٢	٠,٨٠٢-	٠,٥٣٨	٠,٧٤	١,٣٣٥-	٠,٧٨٨-	٠,١٣٦	٢,٠٧٥	٢,٧٣٣-	٠,٢٣٧-	٠,٢١٨-	٠,١٦٧	٠,٨٣٧
٢٣	٢,١٧٣-	٠,٥٦٥-	٠,٦١٥	١,٤٥٧-	١,١١٦-	٠,٨٦١-	٢,٥٤٨	١,٣١٢-	٠,١٣٣-	٠,٠٥٦	٠,٥٦٦	١,١٦٧
٢٤	١,١٢٨-	٠,١٢٦	٠,٥٧٢	٠,٩٨٢-	٠,٦٢-	٠,٠٢٦-	٠,٠٦٤	٣,٣٢٧-	١,١٤٤-	٠,٥٧٦-	٠,٠٥٢-	١,٥٢٧

تابع/ الجدول رقم (١)

القيم الحقيقية لعالم الصعوبة لفواصل الخطوات للاختبارات التي تم توليدها تبعاً لعدد الخطوات

رقم السؤال	عدد خطوات السؤال									
	ثلاث خطوات			أربع خطوات			خمس خطوات			
	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الرابعة	الخطوة الأولى	الخطوة الثانية	الخطوة الثالثة	الخطوة الرابعة	الخطوة الخامسة
٢٥	١,١٢٥-	٠,٠٥٣-	٠,٠٣٤-	١,٥٦٩-	١,٥١٧-	٠,٩٨٣-	٠,٨٣٨-	٠,٧٨٩-	٠,٢٧٨	٠,٦٩٥	٠,٧١٤	٢,١٤١
٢٦	٠,٣٠٤-	٠,٠٦١-	٢,٣٣٢	١,٨٣٢-	١,٤١١-	٠,٠٧٤	٢,١٤٦	١,٧١٥-	١,٣٢٥-	٠,٠٢٨	٠,١٥٩	١,٥٣
٢٧	٠,٠٣-	٠,٠٨٣	٠,٤٠٨	٢,٢٩٤-	١,٠٤٩-	٠,٤٣٥-	٠,٩٨٤	٢,٤٢٨-	١,٦٧٧-	١,٣٣٧-	٠,٨٠٧-	٠,٤٠٢-
٢٨	١,٥٤٦-	١,٠٠٣-	١,٢٧١	١,٨٢٣-	١,٥٨٥-	٠,١١-	٠,٥١٩	١,١٣١-	٠,٦٤٨-	٠,٥٥٤	١,١٤٥	١,٤٥٧
٢٩	١,٠٧٨-	٠,٧٠٥-	٠,٣٦٨-	١,٦٢٩-	١,٥٦٤-	٠,١٥٧-	١,٧٦٢	٢,٨٤٨-	١,٩٦-	١,١٨٦-	٠,٢٥٥-	٠,٠٣٣-
٣٠	٠,٥٩-	٠,٥٧٦-	٠,٧٨٩	٠,٧٧٣-	٠,١٥٣	١,٣٤١	١,٩٩٧	٠,١٩٨-	٠,٠٠١	٠,٣	٠,٨٧٥	١,٠٢٦

كما تم توليد بيانات خاصة بستة اختبارات، اثنان منها كل سؤال مكون من ثلاث خطوات، واثنان منها كل سؤال مكون من أربع خطوات، واثنان منها كل سؤال مكون من خمس خطوات، وجميعها وفق توزيع بيتا (٤،٢) .~ Beta كل اختبار مكون من ٣٠ سؤالاً بحيث كانت المسافة بين صعوبة الخطوة الأولى والخطوة الأخيرة لها متساوية، وتراوحت بين -٣ و+٣، حيث كان توزيع الصعوبة للاختبار الأول ملتويا نحو اليمين بوسط حسابي وانحراف معياري ٠,٦٠-، ١,٢٠، وكان توزيع الصعوبة للاختبار الثاني ملتوياً نحو اليسار بوسط حسابي وانحراف معياري ٠,٦١، ١,١٩ على الترتيب، ويلخص الجدول رقم (٢) القيم الحقيقة لمعالم الصعوبة للفقرات التي تم توليدها تبعاً لعدد الخطوات وشكل التوزيع وفق نموذج التقدير الجزئي.

الجدول رقم (٢)

القيم الحقيقة لمعالم الصعوبة للفقرات التي تم توليدها تبعاً لعدد الخطوات وشكل التوزيع وفق نموذج التقدير الجزئي

رقم السؤال	عدد خطوات السؤال								
	ثلاث خطوات			أربع خطوات			ست خطوات		
	طبيعي	ملتو لليمين	ملتو لليسار	طبيعي	ملتو لليمين	ملتو لليسار	طبيعي	ملتو لليمين	ملتو لليسار
١	٠,٢٧٩	٠,٨٧	٠,٣٢٣	٠,٣٥٩	٠,٨٧٨	٠,٣٩٣	٠,٤٣٤	٠,٦٩٧	٠,١٨٥
٢	٠,٣٦٦	٠,٧٦	٠,٣٤٩	٠,٣٣١	٠,٨٢٧	٠,٤٢٣	٠,٥٨١	٠,٦٩٥	٠,١١٤
٣	٠,٣٤٨	٠,٩٠٥	٠,٣٣	٠,٥٤٣	٠,٤	٠,٢٨٣	٠,٣٣٦	٠,٦٠٥	٠,٢١٤
٤	٠,٥٩١	٠,٧٤٩	٠,١٤٧	٠,٦١٣	٠,٧٣٣	٠,٣٢٦	٠,٧	٠,٨٧	٠,٣٦٥
٥	٠,٢٢	٠,٧٣٧	٠,٠٧٣	٠,٤٤٤	٠,٨٦٦	٠,١٥٥	٠,٤٠٤	٠,٦٣٧	٠,٢٣٥
٦	٠,٦٧٢	٠,٥٢٤	٠,٠٩٤	٠,٥٣٧	٠,٦٢٢	٠,٢٣٧	٠,٢٩٩	٠,٦٠٣	٠,٤٥
٧	٠,٧٣٣	٠,٤٠٢	٠,٢٨٣	٠,٥٢٤	٠,٧٢٥	٠,٦٣٢	٠,٨٢٤	٠,٨٦٣	٠,٢٨١
٨	٠,٣٢٢	٠,٧٦٩	٠,٢٣٩	٠,٥٤٨	٠,٧٣٢	٠,٢٧	٠,٤٨٥	٠,٦٥٣	٠,٢٠٢
٩	٠,٤٥٣	٠,٧٣٨	٠,٣٦٦	٠,٦٤٤	٠,٧٤٨	٠,٢١٩	٠,٣٨٧	٠,٧٠٧	٠,٢٠٩
١٠	٠,٣٣٣	٠,٨١٣	٠,٢٧٣	٠,٤٤٩	٠,٧٥	٠,١٨٥	٠,٥٩٤	٠,٦٣٧	٠,٣٣١

تابع/ الجدول رقم (٢)

القيم الحقيقة لمعالم الصعوبة للفقرات التي تم توليدها تبعاً لعدد الخطوات وشكل التوزيع وفق نموذج التقدير الجزئي

رقم السؤال	عدد خطوات السؤال								
	ثلاث خطوات			أربع خطوات			ست خطوات		
	طبيعي	ملتو لليمين	ملتو لليسار	طبيعي	ملتو لليمين	ملتو لليسار	طبيعي	ملتو لليمين	ملتو لليسار
١١	٠,٦٢٣	٠,٨٥٧	٠,٠٩٤	٠,٧١٧	٠,٨١١	٠,٢٢٢	٠,٥٢٥	٠,٩٢٤	٠,٣٥٨
١٢	٠,٤٦٨	٠,٩١٦	٠,١٩٣	٠,٥١٦	٠,٧٨١	٠,٣٠٢	٠,٤٤	٠,٦٨١	٠,٣٥٦
١٣	٠,٣٧٩	٠,٨٠١	٠,٣٠٢	٠,٦٩	٠,٨٨	٠,١٩٩	٠,٤٦٣	٠,٤٠٧	٠,٣٧
١٤	٠,٤٦٤	٠,٦٨١	٠,٢٦٦	٠,٥	٠,٨٩١	٠,٢٨٥	٠,٣٦٤	٠,٨٣٨	٠,٢٤٨
١٥	٠,٢٨٥	٠,٨١١	٠,٢٨٢	٠,٢٨٢	٠,٤٥٨	٠,٣٠٤	٠,٥٣٢	٠,٨٠٩	٠,١٤١
١٦	٠,٦٨٣	٠,٩٥٥	٠,٣٧٢	٠,٥١٤	٠,٧١	٠,٣٤٩	٠,٥١٥	٠,٧٠٧	٠,١٤٥
١٧	٠,٥٥٨	٠,٥٨٩	٠,٣٧٥	٠,٣٢٤	٠,٧٤٩	٠,٥٠٩	٠,٦٠١	٠,٤٧١	٠,١٩٥
١٨	٠,٢٩٨	٠,٧٢٣	٠,١٨٨	٠,٥٥	٠,٧٥٢	٠,٤٦١	٠,٥٦٦	٠,٧٨	٠,٣٧٥
١٩	٠,٦٦٨	٠,٩٠٩	٠,٠٦٢	٠,٥٦٦	٠,٨٣٢	٠,١٨٢	٠,٣٥٨	٠,٨٧٩	٠,٢٦٥
٢٠	٠,٣٧٥	٠,٥١٤	٠,٢٠٧	٠,٣٤١	٠,٨٠٤	٠,٢٧٩	٠,٤٩١	٠,٨	٠,٤١٩
٢١	٠,٦٠٦	٠,٦٣٢	٠,٠٨	٠,٥٨٤	٠,٥٣٧	٠,٠٨٢	٠,٣٨٤	٠,٧٥٣	٠,١١٤
٢٢	٠,٤٣٦	٠,٨٠٣	٠,٠٦٢	٠,٥٢٥	٠,٦٦٦	٠,٤٠٤	٠,٥٥٧	٠,٦٩٧	٠,٢٦
٢٣	٠,٦٢٥	٠,٧٨٢	٠,٢٤٢	٠,٦٢٤	٠,٧٧٧	٠,٤	٠,٤٧٧	٠,٧٧١	٠,٣١
٢٤	٠,٥٢١	٠,٧٣٨	٠,٤٢٧	٠,٦٢٣	٠,٧٧٣	٠,٤٠٥	٠,٦٣١	٠,٦٢٥	٠,٣٥٥
٢٥	٠,٦١٢	٠,٨٢٨	٠,١٧	٠,٨٢٧	٠,٧٠٢	٠,٢٦٤	٠,٣٤٧	٠,٨٧٦	٠,١١٧
٢٦	٠,٣٩٣	٠,٥٨٦	٠,٥٢٩	٠,٥٧٧	٠,٤٨٣	٠,٤٥٣	٠,٥٦٨	٠,٧١٣	٠,١٠٦
٢٧	٠,٤٤٤	٠,٧٤٢	٠,٠٣٥	٠,٦٥٢	٠,٧٢١	٠,١٦	٠,٨١٩	٠,٨٣٢	٠,١٨٩
٢٨	٠,٦٠٨	٠,٨٨٢	٠,١٩٢	٠,٦٨٢	٠,٧٣٦	٠,١١١	٠,٤٢١	٠,٧٩٤	٠,١٤٩
٢٩	٠,٧٢٣	٠,٤٦٧	٠,٣٠٦	٠,٦٠٨	٠,٨٥٥	٠,١٥٩	٠,٧٦١	٠,٨٩٥	٠,١٦٩
٣٠	٠,٥٣٨	٠,٤٨٥	٠,١٢٢	٠,٣٤٩	٠,٥٧٥	٠,١٢٧	٠,٣٧٥	٠,٥٠٧	٠,١٩٩

المرحلة الثانية: تم توليد استجابات ١٠٠٠ مفحوص، وفق التوزيع

الطبيعي (0, 1) ~ Normal، حيث كان الوسط الحسابي والانحراف المعياري لهذا التوزيع ٠,٠٠٠٥، ٠,٩٨٨ على الترتيب، وكانت درجة الالتواء والتفلطح له ٠,٠٢ -، ٠,٠٧ على الترتيب، بالاستناد على دراسة دراسجو (Drasgow, 1982) للوصول إلى دقة في التقديرات عند استخدام نموذج التقدير الجزئي، وكذلك وفق رأي لورد (Lord, 1980) بأن يكون عدد المفحوصين ١٠٠٠ للحصول على أفضل التقديرات.

تحليل البيانات

قبل معالجة البيانات للإجابة عن أسئلة الدراسة تم التحقق من افتراض أحادية البعد للسمة المقيسة، باعتبار أن أحادية البعد أحد الافتراضات الأساسية التي يجب توافرها في البيانات عند تطبيق نموذج التقدير الجزئي لتحقيق الهدف من الدراسة، قام الباحثان باختيار استجابات ١٠٠٠ مفحوص على كل توزيع من التوزيعات (معتدل، ملتو نحو اليمين، ملتو نحو اليسار)، وكل اختبار من الاختبارات (ثلاث خطوات، أربع خطوات، ست خطوات)، وقد تم الحفاظ على طول الاختبار(٣٠)، وحجم العينات (١٠٠٠) لما لها من أثر في دقة التقديرات، واستخدم الباحثان برنامج SPSS لتحليل البيانات المولدة، وتمت عملية التحليل وفق المراحل التالية:

أولاً - التحقق من افتراضات النموذج: تم التحقق من افتراضات نموذج التقدير الجزئي وهو أحادية البعد، لما له من أثر في دقة التقديرات، وذلك باستخدام التحليل العاملي لكل توزيع من التوزيعات الثلاثة باستخدام طريقة المكونات الرئيسة، ولكل نوع من أنواع الاختبارات. ويظهر الجدول رقم (٣) قيم الجذر الكامن ونسب التباين المفسر للعامل الأول والثاني وناتج قسمة الجذر الأول على العامل الثاني في التوزيعات المختلفة.

الجدول رقم (٣)

قيم الجذر الكامن ونسب التباين المفسر للعامل الأول والثاني وناتج قسمة قيمة الجذر الأول على العامل الثاني في التوزيعات المختلفة تبعاً لشكل التوزيع وعدد خطوات الاختبار

عدد خطوات أسئلة الاختبار	شكل التوزيع	الإحصائي	العامل الأول	العامل الثاني	ناتج القسمة
ثلاث خطوات	طبيعي	الجذر الكامن نسبة التباين المفسر	١٤,٢٥ ٤٧,٥٪	٢,٢٣ ٧,٤٣٪	٦,٣٩
	ملتو نحو اليمين	الجذر الكامن نسبة التباين المفسر	١٤,٦٣ ٤٨,٧٧٪	٢,٤١ ٨,٠٣٪	٦,٠٧
	ملتو نحو اليسار	الجذر الكامن نسبة التباين المفسر	١٤,٥٨ ٤٨,٦٪	٢,٥٦ ٨,٥٣٪	٥,٧٠
أربع خطوات	طبيعي	الجذر الكامن نسبة التباين المفسر	١٥,١٧ ٥٠,٥٦٪	٢,٣١ ٧,٧٪	٦,٥٧
	ملتو نحو اليمين	الجذر الكامن نسبة التباين المفسر	١٥,٠٤ ٥٠,١٣٪	٢,٤٦ ٨,٢٪	٦,١١
	ملتو نحو اليسار	الجذر الكامن نسبة التباين المفسر	١٤,٩٨ ٤٩,٩٣٪	٢,٥٠ ٨,٣٣٪	٥,٩٩
ست خطوات	طبيعي	الجذر الكامن نسبة التباين المفسر	١٥,٧٣ ٥٢,٤٣٪	٢,٢٨ ٧,٦٪	٦,٩٠
	ملتو نحو اليمين	الجذر الكامن نسبة التباين المفسر	١٥,٣٦ ٥١,٢٪	٢,٣٦ ٧,٨٧٪	٦,٥٠
	ملتو نحو اليسار	الجذر الكامن نسبة التباين المفسر	١٥,٢٢ ٥٠,٧٣٪	٢,٤١ ٨,٠٣٪	٦,٣٢

يتضح من النتائج الواردة في الجدول رقم (٣) بأن نسب التباين المفسر للعامل الأول لجميع أشكال التوزيع ولكل الاختبارات كانت أعلى من ٢٠٪، وأن ناتج قسمة الجذر الكامن للعامل الأول على العامل الثاني لجميع أشكال التوزيع ولكل الاختبارات أكبر من ٢، مما يشير إلى وجود بعد سائد لجميع التوزيعات الذي يشير إلى تحقق افتراض أحادية البعد في البيانات تبعاً إلى ما أشار إليه هيتي (Hatti, 1985).

أما فيما يتعلق بافتراض الاستقلال الموضعي (Local Independence)؛

فقد أشار هامبلتون وسوامنيثان (Hambelton & Swaminathan, 1985) إلى أن افتراض الاستقلال الموضعي يكافئ افتراض أحادية البعد، وهذا يعني أنه إذا تحقق افتراض أحادية البعد في المقياس، فإن المقياس يحقق افتراض الاستقلال الموضعي، وكذلك لا بد من الإشارة إلى مؤشرات أخرى على تحقق افتراضات أخرى، من أبرزها تكافؤ الأسئلة في القدرة التمييزية، وما يؤكد هذه النتيجة التقارب في قيم معاملات الارتباط الثنائية (rpbis)، التي تعكس تقارباً في معاملات التمييز للفقرات، وبالتالي، فإن هذا مؤشر على تحقق افتراض من افتراضات النموذج، وهو أن تكون للفقرات قدرة تمييزية متكافئة تقريبا. وقد أشار هامبلتون وسوامنيثان (Hambelton & Swaminathan,1985) إلى أنه لكي يتحقق تكافؤ مؤشرات التمييز ومطابقتها للنموذج، يجب أن تكون قيمها واقعة ضمن حدود المدى (متوسط معاملات التمييز ±0.15)؛ حيث إن متوسط معاملات التمييز للأسئلة جميعها ولكل الاختبارات بأشكال التوزيع المختلفة كانت جميعها ضمن المدى، وكذلك كانت قيمة الانحراف المعياري لهذه المعاملات صغيرة، ويعد ذلك مؤشراً على تحقق هذا الافتراض.

ثانياً - مطابقة البيانات للنموذج: بهدف التحقق من درجة مطابقة الأفراد والفقرات لنموذج التقدير الجزئي، تم استخدام برنامج WINSTEPS لتحليل البيانات لكل اختبار ولكل توزيع من التوزيعات الثلاثة، وبناءاً على نتائج التحليل، بالنسبة للأفراد فقد ارتأى الباحثان بإبقاء حجم العينات ١٠٠٠ مفحوص نظرا لقلة عدد الأفراد الذين لم تتطابق استجاباتهم مع توقعات النموذج لإجراء عملية المقارنة بين معالم الفقرات الناتجة من كل توزيع من التوزيعات المختلفة للقدرة، كم بينت نتائج التحليل أن الأسئلة جميعها كانت ضمن مدى المطابقة (٠,٧-١,٣)، وبالتالي تم الإبقاء على الفقرات جميعها (٣٠)، والأفراد جميعهم (١٠٠٠).

نتائج الدراسة ومناقشتها

للإجابة عن أسئلة الدراسة تم استخدام برنامج (WINSTEPS)؛ وذلك للحصول على تقديرات القدرة للأفراد ومعلم الصعوبة للفقرات تبعا لعدد خطوات أسئلة الاختبار (ثلاث خطوات، أربع خطوات، خمس خطوات)، وشكل التوزيع لصعوبة الخطوات وكانت النتائج على النحو الآتي:

فيما يتعلق بالإجابة عن السؤال الأول: ما أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، في دقة تقدير معالم القدرة للأفراد؟ تم الحصول على تقدير قدرة واحدة لكل فرد في كل اختبار من الاختبارات الثلاثة المختلفة في عدد خطواتها (ثلاث خطوات، أربع خطوات، ست خطوات)، وكل منها مكون من ٣٠ سؤالاً بحيث كانت المسافة بين صعوبة الخطوة الأولى والخطوة الأخيرة لها متساوية، حيث تراوحت بين -٣ و+٣ لوجيت، تبعاً لشكل التوزيع لصعوبة الخطوات(طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، بالإضافة إلى ذلك تم الحصول على الخطأ المعياري في تقدير القدرة، وإحصائي الملاءمة لكل تقدير من هذه التقديرات. والجدول رقم (٤) يبين الأوساط الحسابية والانحرافات المعيارية لكل من: قدرات الأفراد، والأخطاء المعيارية في تقديرها، وإحصائي المطابقة لكل اختبار من الاختبارات الثلاثة تبعاً لعدد خطواتها.

الجدول رقم (٤)

الأوساط الحسابية والانحرافات المعيارية لكل من تقديرات القدرة للأفراد والأخطاء المعيارية وإحصائيات الملاءمة الداخلية والخارجية لكل اختبار تبعاً لعدد الخطوات الأسئلة

شكل التوزيع	عدد خطوات الاختبار	القدرة		الخطأ المعياري		إحصائي الملاءمة الداخلية (Z INFIT)		إحصائي الملاءمة الخارجية (Z OUTFIT)	
		الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	الانحراف المعياري
طبيعي	ثلاث	-٠,٠٧	١,٨٢	٠,٣٠	٠,١٠	-٠,١	٠,٩	-٠,١	٠,٩
	أربع	-٠,٠٢	١,٧٩	٠,٢٩	٠,١١	-٠,١	١	-٠,١	١
	خمس	٠,٢٥	١,٧١	٠,١٧	٠,٠٨	-٠,١	٠,٩	-٠,١	٠,٩
ملتو نحو اليمين	ثلاث	١,٦٦	١,٧٤	٠,٣٦	٠,١٣	-٠,١	٠,٩	-٠,١	٠,٩
	أربع	١,٧٣	١,٦٨	٠,٣١	٠,١٤	-٠,١	٠,٩	-٠,١	٠,٩
	خمس	١,٧٤	١,٦٢	٠,٢٠	٠,١١	-٠,١	١	-٠,١	١
ملتو نحو اليسار	ثلاث	-٢,٠٧	١,٨٢	٠,٤٧	٠,١٦	-٠,١	٠,٩	-٠,١	٠,٨
	أربع	-١,٩٤	١,٧٦	٠,٣١	٠,١٣	-٠,١	٠,٩	-٠,١	٠,٩
	خمس	-١,٥٩	١,٧٠	٠,٢٨	٠,١٠	-٠,١	٠,٩	-٠,١	٠,٩

يلاحظ من الجدول رقم (٤) أن الوسط الحسابي لتقديرات القدرة يزداد وبشكل مطرد بزيادة عدد خطوات أسئلة الاختبار، حيث أشارت النتائج إلى زيادة كبيرة في متوسط تقديرات قدرات الأفراد بازدياد عدد خطوات الاختبار إذ بلغت (٠,٣٢) وحدة لوجيت بين الاختبار ذي الخطوات الخمس والاختبار ذي الخطوات الثلاث عندما كان شكل التوزيع لصعوبة الفقرات طبيعياً، وبلغت (٠,٠٨) عندما كان شكل التوزيع ملتوياً لليمين، و(٠,٤٧)، في حالة التوزيع الملتوي لليسار. كما أشارت النتائج إلى نقص كبير في الوسط الحسابي للخطأ المعياري في القياس لهذه القدرات بزيادة عدد خطوات الاختبار. ويبدو أن الوسط الحسابي للخطأ المعياري يتناقص بزيادة عدد خطوات أسئلة الاختبار، كما أن الوسط الحسابي والانحراف المعياري لإحصائيات المطابقة كانت قريبة من القيم المثالية المتوقعة (صفر، ١)، وفي الاختبارات الثلاثة المختلفة في عدد خطواتها، وفي أشكال التوزيع المختلفة. ويتبين من الجدول رقم (٤) أن الانحراف المعياري لتقديرات القدرة للأفراد يقل بزيادة عدد خطوات أسئلة الاختبار؛ أي أن التشتت في تقديرات القدرة للأفراد يقل بزيادة عدد خطوات أسئلة الاختبار؛ وهذا يعني أن تجمع قدرات الأفراد حول الوسط الحسابي قد ازداد بزيادة عدد خطوات أسئلة الاختبار. ويلاحظ من الجدول رقم (٤) أن الوسط الحسابي للأخطاء المعيارية لتقديرات القدرة للأفراد الملاحظ كان أعلى وذلك عند التقدير باستخدام الاختبار ذي الثلاث خطوات، وذلك عند أشكال التوزيع المختلفة، يليه الاختبار ذو الأربع خطوات ثم الاختبار ذو الخمس خطوات؛ أي أن الاختبار ذا الخمس خطوات كان الأكثر دقة في تقديرات القدرة للأفراد. كما أن أعلى قيمة للأخطاء المعيارية والانحراف المعياري لها للاختبار ذي الثلاث خطوات؛ وهذا يعني أن نقصان عدد الخطوات يعمل على المغالاة في تقدير قدرات الأفراد، وبالتالي زيادة الأخطاء المعيارية. كما يلاحظ من الجدول رقم (٤) أن الانحراف المعياري للأخطاء المعيارية للاختبار ذي الخمس خطوات كان الأقل قيمة وذلك عند أشكال التوزيع المختلفة؛ وهذا يعني أن هذا الاختبار هو المفضل من حيث التجانس في قيم الأخطاء المعيارية، وعدم تقديم أخطاء معيارية شاذة، يليه الاختبار ذو الأربع خطوات، ثم الاختبار ذو الثلاث خطوات.

ويبدو للوهلة الأولى أن هذه النتيجة تناقض إحدى الخصائص الرئيسة لنموذج التقدير الجزئي وهي إمكانية فصل معالم الأفراد بشكل مستقل عن عينة الأسئلة التي يستجيبون لها. إذ لم يكن الغرض من البحث دحض افتراض رئيس في النظرية الحديثة للقياس، وإنما هو تقييم لأثر عدد خطوات الاختبار في دقة تقديرات القدرة للأفراد من خلال الأخطاء المعيارية في القياس. لذا وللتعرف إلى دلالة الفروق بين متوسطات التقديرات للأخطاء المعيارية لتقدير قدرات الأفراد عند الاختبارات الثلاثة المختلفة في عدد خطواتها، وشكل التوزيع، والتفاعل بينهما، تم استخدام التصميم العاملي للقياسات المتكررة على أحد العاملين فقط؛ لأن التحليل تم على العينة نفسها بأشكال مختلفة لتوزيعات الصعوبة للخطوات (متغير غير مكرر عليه القياس)، ولكن اختلف عدد خطوات الاختبار (متغير مكرر عليه القياس). لذا تم استخدام تحليل التباين الثنائي ٢×٢ في حالة القياسات المتكررة على أحد العاملين فقط. كما هو موضح في الجدول رقم (٥).

الجدول رقم (٥)

نتائج تحليل التباين للقياسات المتكررة لأثر متغيري عدد خطوات الاختبار، وشكل التوزيع في تقدير معالم القدرة للأفراد

مصدر التباين	مجموع المربعات	درجات الحرية	وسط المربعات	قيمة F	الدلالة الإحصائية
بين الأفراد (شكل التوزيع)	٠,٨٩٤	٢	٠,٤٤٧	٢,١٨٠	٠,٠٩٨
الخطأ	٦١٤,٢٥٨	٢٩٩٧	٠,٢٠٥		
داخل الأفراد (عدد خطوات الاختبار)	١٨,٤٥٧		٩,٢٢٨٥	٧٨,٢٠٧	٠,٠٠٠❖
التفاعل	٤,٦٨٩	٤	١,١٧٢	٩,٩٣٢	٠,٠٠٣
الخطأ	٧١٠,٢٤٥	٥٩٩٤	٠,١١٨		

❖ ذات دلالة إحصائية عند مستوى الدلالة (α=٠,٠٥)

يلاحظ من الجدول رقم (٥) وجود فروق ذات دلالة إحصائية تعزى لأثر متغير عدد خطوات الاختبار، وعدم وجود فروق دالة إحصائياً تعزى لأثر متغير

شكل التوزيع، وكذلك أثر التفاعل بين عدد خطوات الاختبار وشكل التوزيع لصعوبة الخطوات. وللكشف عن مواقع الفروق بين متوسطات الأخطاء المعيارية لتقديرات القدرة، تم استخدام اختبار شفيه للمقارنات البعدية، كما هو موضح في الجدول رقم (٦).

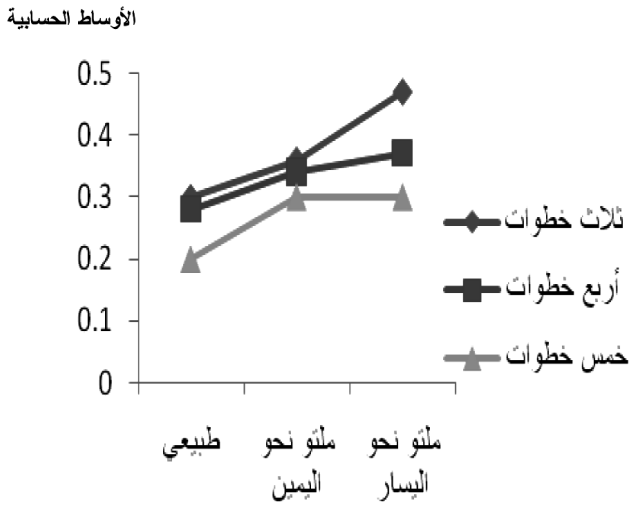
الجدول رقم (٦)

نتائج المقارنات البعدية للكشف عن مواقع الفروق بين متوسطات الأخطاء المعيارية لتقديرات القدرة تبعا لشكل التوزيع باستخدام اختبار شفيه

شكل التوزيع	عدد خطوات الاختبار	ثلاث (الوسط الحسابي = ٣٠,)	أربع (الوسط الحسابي = ٢٩,)	خمس (الوسط الحسابي = ٢٧,)
طبيعي	ثلاث (الوسط الحسابي = ٣٠,)	---	٠,٠١	٠,١٣*
	أربع (الوسط الحسابي = ٢٩,)		---	٠,١٢*
	خمس (الوسط الحسابي = ١٧,)		---	
	الفئات	**ثلاث (الوسط الحسابي = ٣٦,)**	**أربع (الوسط الحسابي = ٣١,)**	**خمس (الوسط الحسابي = ٢٠,)**
ملتو نحو اليمين	ثلاث (الوسط الحسابي = ٣٦,)	---	٠,٠٥	٠,١٦*
	أربع (الوسط الحسابي = ٣١,)		---	٠,١١*
	خمس (الوسط الحسابي = ٢٠,)		-----	
	الفئات	**ثلاث (الوسط الحسابي = ٤٧,)**	**أربع (الوسط الحسابي = ٠٣١,)**	**خمس (الوسط الحسابي = ٢٨,)**
ملتو نحو اليسار	ثلاث (الوسط الحسابي = ٤٧,)	-----	٠,١٦*	٠,١٩*
	أربع (الوسط الحسابي = ٣١,)		------	٠,٠٣
	خمس (الوسط الحسابي = ٢٨,)			-----

يلاحظ من الجدول رقم (٦) وجود فرق ذي دلالة بين الاختبار ذي الخمس خطوات والاختبار ذي الثلاث خطوات ولصالح الاختبار ذي الخمس خطوات لكل أشكال التوزيع. وكذلك تبين وجود فرق ذي دلالة إحصائية بين الاختبار ذي الخمس خطوات والاختبار ذي الأربع خطوات ولصالح الاختبار ذي الخمس خطوات عندما يكون شكل التوزيع طبيعياً، وملتوياً نحو اليمين، وكذلك وجود فرق

بين الاختبار ذي الأربع خطوات والاختبار ذي الثلاث خطوات ولصالح الاختبار ذي الأربع خطوات عندما يكون شكل التوزيع ملتوياً نحو اليسار، وبالتالي يمكن القول بأن معظم تقديرات القدرات للأفراد من خلال الأخطاء المعيارية في القياس قد أصبحت أكثر دقة بازدياد عدد خطوات الاختبار من خلال النقص في الأخطاء المعيارية في قياسها. وهذه النتيجة تتفق مع النتيجة التي توصل إليها كل من التقي (١٩٩٢)، ودود وكوخ (Dodd and Koch, 1987)، ورايت وماسترز (Wright and Masters, 1982)، ولتوضيح أثر التفاعل بين شكل التوزيع وعدد خطوات الاختبار في دقة تقدير معلمة القدرة للأفراد؛ فقد تم تمثيل هذا التفاعل بيانيا كما في الشكل رقم (٣) وذلك بالاستعانة بالأوساط الحسابية في الجدول رقم (٥)، ويوضح الشكل رقم (٢) التفاعل بين شكل التوزيع وعدد فقرات الاختبار في دقة تقدير قدرات الأفراد. ويلاحظ أن الأخطاء المعيارية كانت أعلى للاختبار ذي الثلاث خطوات، يليه الاختبار ذو الأربع خطوات، ثم الاختبار ذو الخمس خطوات وذلك عند أشكال التوزيع جميعها؛ أي أن الاختبار ذا الخمس خطوات كان الأكثر دقة في تقديرات القدرة للأفراد.



الشكل رقم (٢) - التمثيل البياني للتفاعل بين عدد خطوات الاختبار وشكل التوزيع لصعوبة السؤال

يلاحظ من الشكل رقم (٢) أن الأخطاء المعيارية الملاحظة كانت أعلى وفقاً للاختبار ذي الثلاث خطوات، يليه الاختبار ذو الأربع خطوات، ثم الاختبار ذو الخمس خطوات، وذلك عند أشكال التوزيع المختلفة؛ أي أن الاختبار ذا الخمس خطوات كان الأكثر دقة في تقديرات القدرة للأفراد. ويلاحظ من الشكل نفسه كذلك أن التقديرات كانت أكثر دقة عندما يكون الاختبار ذا خمس خطوات، في حين أن التقديرات كانت أقل دقة عندما يكون الاختبار ذا ثلاث خطوات، وعند أشكال التوزيع جميعها؛ وهذا يعني أنه بزيادة عدد خطوات الأسئلة، تزداد دقة التقديرات للقدرة، وما يلاحظ أيضاً أن التقديرات تكون أكثر دقة عندما يكون شكل التوزيع طبيعياً، يليه شكل التوزيع الملتوي نحو اليمين، وأخيراً عندما يكون شكل التوزيع ملتوياً نحو اليسار؛ وهذا يشير إلى قيمة قليلة لمتوسطات الأخطاء المعيارية، تشير إلى دقة أعلى في التقدير.

وفيما يتعلق بالإجابة عن السؤال الثاني: ما أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، في دقة تقدير معالم الصعوبة للأسئلة؟ وكما أشرنا فيبدو أيضاً وللوهلة الأولى أن هذا السؤال يناقض إحدى الخصائص الرئيسة لنموذج التقدير الجزئي، وهي إمكانية فصل معالم الأسئلة بشكل مستقل عن عينة الأفراد الذين يستجيبون لها، إذ لم يكن الغرض من البحث دحض افتراض رئيس في النظرية الحديثة للقياس. وإنما هو تقييم لأثر عدد خطوات الاختبار في دقة تقديرات معالم الصعوبة من خلال الأخطاء المعيارية في القياس. لذا وللإجابة عن السؤال، تم تقدير قيم معالم الصعوبة لكل سؤال من أسئلة الاختبارات المختلفة في عدد خطواتها تبعاً لشكل التوزيع لصعوبة الخطوات؛ علماً بأن كل اختبار من الاختبارات مكون بصورته النهائية من (٣٠) فقرة. والجدول رقم (٧) يبين الإحصائيات الوصفية الخاصة بالأخطاء المعيارية لمعلمة الصعوبة للأسئلة، تبعاً لعدد خطوات السؤال، وشكل التوزيع لصعوبة الخطوات.

جدول رقم (٧)

الإحصائيات الوصفية الخاصة بالأخطاء المعيارية للتقدير لمعلمة الصعوبة، تبعا لعدد خطوات الاختبار وشكل التوزيع

شكل التوزيع	الإحصائيات الوصفية	عدد خطوات الاختبار		
		ثلاث خطوات	أربع خطوات	خمس خطوات
		الخطأ المعياري للصعوبة	الخطأ المعياري للصعوبة	الخطأ المعياري للصعوبة
طبيعي	الوسط الحسابي	٠,٠٥	٠,٠٣	٠,٠٢
	أعلى قيمة	٠,٠٧	٠,٠٦	٠,٠٤
	أدنى قيمة	٠,٠٤	٠,٠٤	٠,٠٣
	الانحراف المعياري	٠,٠١	٠,٠٠	٠,٠٠
	RMS	٠,٠٥	٠,٠٤	٠,٠٤
ملتو نحو اليمين	الوسط الحسابي	٠,٠٥	٠,٠٤	٠,٠٣
	أعلى قيمة	٠,٠٩	٠,٠٦	٠,٠٥
	أدنى قيمة	٠,٠٤	٠,٠٤	٠,٠٤
	الانحراف المعياري	٠,٠١	٠,٠١	٠,٠١
	RMS	٠,٠٦	٠,٠٥	٠,٠٥
ملتو نحو اليسار	الوسط الحسابي	٠,٠٦	٠,٠٤	٠,٠٤
	أعلى قيمة	٠,٠٨	٠,٠٧	٠,٠٦
	أدنى قيمة	٠,٠٥	٠,٠٤	٠,٠٣
	الانحراف المعياري	٠,٠١	٠,٠١	٠,٠١
	RMS	٠,٠٠	٠,٠٥	٠,٠٥

ويلاحظ من الجدول رقم (٧) أن الوسط الحسابي للأخطاء المعيارية لتقديرات معالم الفقرات الملاحظ كان أعلى، وذلك عند استخدام الاختبار ذي الثلاث خطوات، ولكل أشكال التوزيع الثلاثة، يليه الاختبار ذو الأربع خطوات، ثم الاختبار ذو الخمس خطوات؛ أي أن طريقة التقدير باستخدام الاختبار ذي الخطوات الخمسة كان الأكثر دقة في تقديرات معالم الصعوبة. كما أن أعلى

قيمة للأخطاء المعيارية والانحراف المعياري لها كانت في حالة الاختبار ذي الثلاثة خطوات؛ وهذا يعني أن استخدام الاختبار ذي الثلاثة خطوات يعمل على المغالاة في تقدير معالم الصعوبة للأسئلة، وبشكل أقل عندما يكون شكل التوزيع طبيعياً، أما عندما يكون شكل التوزيع ملتوياً سواء كان الالتواء لليمين او لليسار فقد كان هناك تطابق في تقدير معالم الصعوبة للأسئلة، وبالتالي زيادة الأخطاء المعيارية.

وللتعرف إلى دلالة الفروق بين متوسطات الأخطاء المعيارية لتقدير معلمة الصعوبة للفقرات عند أشكال التوزيع المختلفة، وعدد خطوات الاختبار والتفاعل بينهما، فقد تم استخدام التصميم العاملي للقياسات المتكررة على أحد العاملين فقط؛ لأن التحليل تم على العينة نفسها بثلاثة أشكال توزيع مختلفة (متغير غير مكرر عليه القياس)، ولكن اختلفت عدد خطوات الاختبار (متغير مكرر عليه القياس). لذا تم استخدام تحليل التباين الثنائي ٢×٢ في حالة القياسات المتكررة على أحد العاملين فقط. كما هو موضح في الجدول رقم (٨).

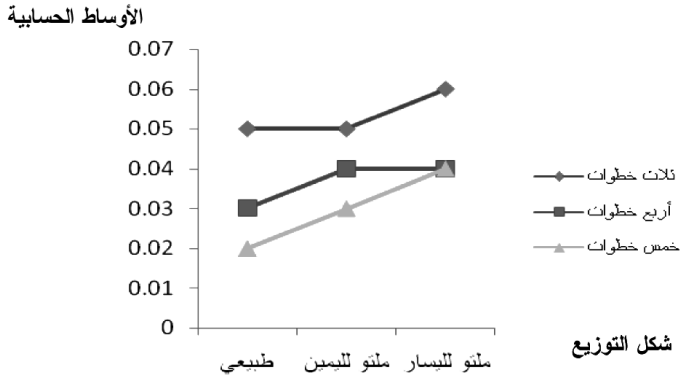
الجدول رقم (٨)

نتائج تحليل التباين للقياسات المتكررة لأثر متغيري عدد خطوات الاختبار، وشكل التوزيع في تقدير معالم الصعوبة للأسئلة

مصدر التباين	مجموع المربعات	درجات الحرية	وسط المربعات	قيمة F	الدلالة الاحصائية
بين الأفراد (شكل التوزيع)	٠,٤٨٨	٢	٠,٢٤٤	١,٧١٦	٠,١٥٦٢
الخطأ	٤٢٦,١٨٠	٢٩٩٧	٠,١٤٢٢		
داخل الأفراد (عدد خطوات الاختبار)	١,٢١٤	٢	٠,٦٠٧	١٠,٢٠٢	٠,٠١٩❖
التفاعل	٢,٥٦٨	٤	٠,٦٤٢	١٦,٨٠٧	٠,٠٠٨
الخطأ	٣٥٦,٦٥٤	٥٩٩٤	٠,٠٥٩٥		

❖ ذات دلالة إحصائية عند مستوى الدلالة (α = ٠,٠٥)

يلاحظ من الجدول رقم (٨) عدم وجود فروق ذات دلالة إحصائية تعزى لأثر متغير شكل التوزيع، كما يلاحظ وجود فروق ذات دلالة إحصائية في الأوساط الحسابية للأخطاء المعيارية تعزى لعدد خطوات الاختبار، ووجود أثر التفاعل بين شكل التوزيع وعدد خطوات الاختبار على تقديرات معلمة الصعوبة للفقرات؛ أي أن هناك تأثيراً مشتركا لكل من العاملين. وللكشف عن مواقع الفروق بين متوسطات الأخطاء المعيارية لتقديرات معالم الصعوبة للأسئلة، تم استخدام اختبار شفيه للمقارنات البعدية، حيث تبين وجود فرق ذي دلالة إحصائية بين الاختبار ذي الخمس خطوات والاختبار ذي الثلاث خطوات ولصالح الاختبار ذي الخمس خطوات، وكذلك تبين وجود فرق ذي دلالة إحصائية بين الاختبار ذي الخمس خطوات والاختبار ذي الأربع خطوات ولصالح الاختبار ذي الخمس خطوات، وعدم وجود فرق بين الاختبار ذي الأربعة خطوات والاختبار ذي الثلاث خطوات، وبالتالي يمكن القول بأن تقديرات معالم الصعوبة للفقرات من خلال الأخطاء المعيارية في القياس قد أصبح أكثر دقة بازدياد عدد خطوات الاختبار من خلال النقص في الأخطاء المعيارية في قياسها. وهذه النتيجة تتفق مع النتيجة التي توصل اليها كل من التقي (١٩٩٢)، ودود وكوخ (Dodd and Koch, 1987)، ورايت وماسترز (Wright and Masters, 1982). ولتوضيح أثر التفاعل بين شكل التوزيع وعدد خطوات الاختبار في دقة تقدير لمعالم الصعوبة للأسئلة؛ فقد تم تمثيل هذا التفاعل بيانيا كما في الشكل رقم (٣) وذلك بالاستعانة بالأوساط الحسابية في الجدول رقم (٥)، ويوضح الشكل رقم (٣) التفاعل بين شكل التوزيع وعدد فقرات الاختبار في دقة معالم الصعوبة للفقرات. ويمكن تمثيل هذا التفاعل بيانيا كما في الشكل رقم (٣) وذلك بالاستعانة بالأوساط الحسابية في الجدول رقم (٤).



الشكل رقم (٣) - التمثيل البياني للتفاعل بين شكل التوزيع وعدد خطوات الاختبار في تقدير معلم الصعوبة للفقرات

يلاحظ من الشكل رقم (٣) أن الأخطاء المعيارية الملاحظة كانت أعلى للاختبار ذي الخطوات الثلاثة، يليه الاختبار ذو الأربع خطوات، ثم الاختبار ذو الخمس خطوات، وذلك عند أشكال التوزيع المختلفة؛ أي أن الاختبار ذا الخطوات الخمس كان الأكثر دقة في تقديرات معالم الصعوبة للسؤال. كما يتبين من الشكل رقم (٣) أن هناك تفاوتا واضحا بين قيم الأخطاء المعيارية في حالة شكل التوزيع طبيعي، بينما تقلص هذا الفرق عندما كان شكل التوزيع ملتوياً لليمين، وملتوياً نحو اليسار، وتلاشى تماما عندما كان شكل التوزيع ملتوياً نحو اليسار بين الاختبارين المكونين من ثلاث وأربع خطوات؛ أي أن الاختبار ذا الثلاث خطوات كان أقل الاختبارات دقة في تقدير معالم الصعوبة، وعند أشكال التوزيع جميعها. وتتفق هذه النتيجة مع النتيجة التي حصل عليها رايت وماسترز (Wright and Masters, 1982)À وكذلك مع النتيجة التي حصل عليها شاهين (١٩٩٣)؛ ويمكن تفسير هذه النتيجة أيضا أن زيادة عدد خطوات السؤال ربما تجعل السؤال أكثر سهولة، حيث إن إعطاء فرصتين للإجابة عن السؤال تجعل الأسئلة أكثر سهولة من إعطاء فرصة واحدة.

وفيما يتعلق بالإجابة عن السؤال الثالث؛ حول تقصي أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة

الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)، في ثبات تباعد الأفراد؟ فقد تم حساب معاملات التباعد (Person Separation Index)، ومعامل الثبات لتباعد الأفراد (Test Reliability of Person Separation)، وعدد الطبقات الإحصائية المتمايزة للأفراد وعدد الطبقات الإحصائية المتمايزة للأفراد (Ability Strata)، تبعاً لاختلاف الاختبارات في عدد خطواتها وشكل التوزيع، والجدول رقم (٩) يبين معاملات التباعد بين الأفراد وثبات هذه المعاملات وعدد الطبقات الإحصائية المتمايزة للأفراد في كل اختبار من هذه الاختبارات تبعاً لعدد خطواتها.

الجدول رقم (٩)

معاملات التباعد بين الأفراد وثبات هذه المعاملات وعدد الطبقات الإحصائية المتمايزة للأفراد في كل اختبار من هذه الاختبارات تبعاً لعدد خطواتها، وشكل التوزيع

	عدد خطوات الاختبار	ثبات تباعد الأفراد	معامل التباعد	عدد الطبقات الإحصائية
ثلاث خطوات	طبيعي	٠,٩٦	٥,٠٨	٧,١١
	ملتو نحو اليمين	٠,٩٤	٤,٠٦	٥,٧٥
	ملتو نحو اليسار	٠,٩٢	٣,٢٩	٤,٧٢
أربع خطوات	طبيعي	٠,٩٧	٥,٩٨	٨,٣١
	ملتو نحو اليمين	٠,٩٦	٤,٧٨	٦,٧١
	ملتو نحو اليسار	٠,٩٦	٥,١٦	٧,٢١
خمس خطوات	طبيعي	٠,٩٧	٥,٣٤	٧,٤٥
	ملتو نحو اليمين	٠,٩٧	٥,٤٦	٧,٦١
	ملتو نحو اليسار	٠,٩٦	٥,١٦	٧,٢١

يتضح من الجدول رقم (٩) أن قيم معاملات التباعد للأفراد وقيم معاملات الثبات لهذا المعامل كانت جميعها عالية، (وهي تناظر معامل الثبات لكودر- ريتشاردسون-٢٠) (KR-20)، وهي بلا شك قيم مرتفعة جداً وتشير إلى كفاية أسئلة الاختبار في الفصل بين الأفراد والتمييز بين مستويات القدرة

المختلفة للأفراد، وكفاية عينة الأفراد في الفصل بين الأسئلة وتعريف متصل القدرة الذي تقيسه هذه الأسئلة وكانت هذه معاملات متقاربة في الاختبارات الثلاثة المختلفة في عدد خطواتها سواء كان شكل التوزيع طبيعياً أو ملتوياً نحو اليمين أو ملتوياً نحو اليسار. ويلاحظ من الجدول رقم (٩)، أن قيم معاملات الثبات الملاحظة، والمقدرة باستخدام الاختبار ذي الخمس خطوات كانت الأعلى، وذلك عند أشكال التوزيع المختلفة، يليها الاختبار ذو الأربع خطوات، ثم الاختبار ذو الثلاث خطوات. كما يلاحظ بأن قيم معاملات الثبات كانت الأعلى عندما يكون شكل التوزيع طبيعياً يليه شكل التوزيع الملتوي نحو اليمين، ثم وأخيرا شكل التوزيع الملتوي نحو اليسار.

بعد التحقق من جودة مطابقة أسئلة الاختبارات المختلفة في عدد الخطوات بصورتها النهائية وكذلك تحقيقها لافتراضات نموذج التقدير الجزئي، تم الحصول على نوعين من معاملات الثبات اللذين يطلق عليهما الثبات المتعلق بالأفراد (Person- Reliability) أو (Sample- Reliability)، والثبات المتعلق بفقرات بالاختبار (Test- Reliability) أو (Item- Reliability)؛ حيث يشير مفهوم الثبات في إطار نموذج التقدير الجزئي أيضاً إلى مدى الدقة في تحديد موقع كل من الأسئلة والأفراد على متصل القدرة الذي يهدف إلى قياسه. ويمكن تحديد مدى دقة الأسئلة في تعريف هذا المتصل بما يسمى بمعامل الفصل/ التباعد بين الفقرات (Item- Separation Index) (Gi) الذي يجب أن تكون قيمته أكبر ما يمكن. فإذا لم تزد قيمته على (٢) يصعب عندئذ قياس السمة بوساطة هذه الأسئلة. وقد وجد الباحثان أن قيم معاملات الفصل بين الفقرات في الاختبار المختلفة في عدد خطواتها بصورتها النهائية، ولكل أشكال التوزيع كبيرة وأكبر من ٢؛ وهذا يؤكد قدرة أسئلة الاختبار وكفايتها في تعريف متصل السمة الذي تقيسه، وكذلك ما يسمى معامل الفصل بين الأفراد (Gp). وقد وجد الباحثان أن قيم هذا المعامل جميعها تزيد على (٢)، لذا فإن عينة الأفراد كافية كذلك للفصل بين الفقرات وتعريف متصل السمة أيضاً.

كما وجد الباحثان أن قيم معاملات الثبات المتعلقة بالأفراد والاختبار

باستخدام الصيغة الرياضية (R = G2/1 + G2) (0.99) كانت جميعها مرتفعة، إذ تدل الأولى على كفاية عينة الأفراد في الفصل بين الأسئلة؛ وبالتالي في تعريف متصل القدرة الذي تقيسه هذه الأسئلة، وتدل الثانية على كفاية عينة الأسئلة في الفصل بين الأفراد؛ وبالتالي في التمييز بين مستويات القدرة لدى هؤلاء الأفراد.

وثمة مؤشر إحصائي آخر يمكن الإشارة إليه وهو ما يسمى عدد الطبقات الإحصائية المتمايزة لتقديرات قدرات الأفراد والتي تعرّف أنها طبقات لقدرات الأفراد (Ability Strata)، حيث تبتعد مراكزها بعضها عن بعض بمقدار ثلاث وحدات من وحدات الأخطاء المعيارية في القياس، ويلاحظ أن قيمة معامل الثبات المتعلق بالأفراد كانت أقل للاختبار ذي الخطوات الثلاث، ولكنها كانت كبيرة وقادرة على التمييز بين سبع طبقات إحصائية متمايزة للأفراد؛ أي بين سبعة مستويات للقدرة لدى هؤلاء الأفراد؛ في حين أن معامل الثبات للأفراد في الاختبار ذي الخطوات الخمس كانت قيمته كبيرة وقادرة على التمييز بين سبعة طبقات إحصائية متمايزة للقدرات لدى أفراد عينة الدراسة.

ولمعرفة دلالة الفروق بين قيم معاملات الثبات للأفراد تبعاً لعدد خطوات أسئلة الاختبار الثلاثة ولكل شكل من أشكال التوزيع الثلاثة، فقد استخدم الباحثان الإحصائي (M) المقترح من هاكستين وولن (Hakstin and Whalen, 1976) الذي يتبع توزيع كاي تربيع بدرجات حرية تساوي (عدد المعاملات -١) وتحسب قيمة M من الصيغة الرياضية الآتية:

$$M = \frac{(J-1)(9n-11)^2}{18J(n-1)} \left[k - \frac{\sum_{k=1}^{3}(1-r_k)^{-\frac{1}{3}}]^2}{\sum_{k=1}^{3}(1-r_k) - \frac{2}{3}} \right]$$

حيث تمثل الرموز: K عدد خطوات الاختبار، :n عدد أفراد المجموعة في

الاختبار المستخدم، r_k: معامل ثبات للأفراد من خلال الاختبار المستخدم، j: عدد فقرات الاختبار المستخدم.

وقد كشفت تلك القيم عن عدم وجود فروق بين معاملات الثبات للأفراد تعزى لعدد خطوات الاختبار عندما كان شكل التوزيع لصعوبة الفقرات طبيعياً، حيث كانت قيمة الإحصائي المحسوبة (٢,٤١)، وهي أقل من القيمة الحرجة لكاي تربيع بدرجات حرية (٢) عند مستوى الدلالة الإحصائية (α=٠,٠٥)، كما كشفت عن عدم وجود فروق بين معاملات الثبات للأفراد تعزى لعدد خطوات الاختبار عندما كان شكل التوزيع لصعوبة الأسئلة ملتوياً نحو اليمين، حيث كانت قيمة الإحصائي المحسوبة (٢,٣٣)، وهي أقل من القيمة الحرجة لكاي تربيع بدرجات حرية (٢) عند مستوى الدلالة الإحصائية (α=٠,٠٥). وكذلك كشفت عن عدم وجود فروق بين معاملات الثبات للأفراد، تعزى لعدد خطوات الاختبار عندما كان شكل التوزيع لصعوبة الأسئلة ملتوياً نحو اليسار، حيث كانت قيمة الإحصائي المحسوبة (١,٩٨)، وهي أقل من القيمة الحرجة لكاي تربيع بدرجات حرية (٢) عند مستوى الدلالة الإحصائية (α=٠,٠٥). ومن المؤشرات الأخرى التي يستدل منها على معامل الثبات للاختبار في النظرية الحديثة في القياس هي دالة المعلومات للاختبار (Information Function) (Ozaki & Toyoda, 2009) ومنحنى دالة المعلومات للاختبار ينتج من تكوين منحنيات الفقرات فوق بعضها بعضا وفق العلاقة:

$$I(\theta) = \sum_{j} Ig(\theta)$$

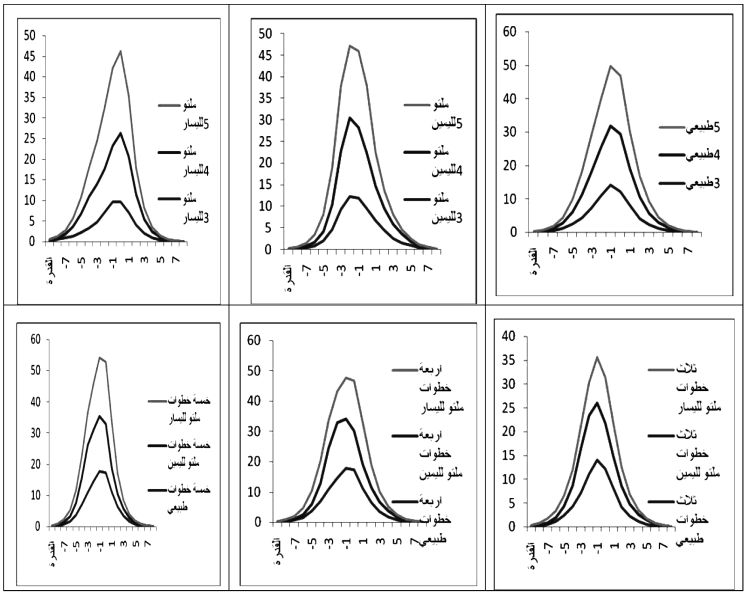
حيث ($I(\theta)$): هي كمية المعلومات للاختبار عند مستوى القدرة، $\sum_{g} Ig(\theta)$ مجموع دوال المعلومات لفقرات ذلك الاختبار عند مستوى القدرة نفسه. ولذلك فإن زيادة عدد الفقرات يعطي خطأ معيارياً صغيراً (θ) S.E، ونقصان قيمة الخطأ المعياري في تقدير القدرة يؤدي إلى زيادة كمية المعلومات للاختبار وفق العلاقة الآتية:

$$I(\theta) = \frac{1}{\sqrt{S.E(\theta)}}$$

ونقصان قيمة الخطأ المعياري في تقدير القدرة يؤدي إلى زيادة قيمة معامل الثبات شأنه في ذلك شأن النظرية الكلاسيكية. وما يميز هذه الطريقة عن الطرق الكلاسيكية أن تقدير الثبات في النظرية الكلاسيكية مرتبط بالعينة، وهذه خاصية سيئة لتقدير الثبات، كما أنها تؤدي لتقدير جماعي للأخطاء الفردية في علامات الاختبار، وهو ما يسمى بالخطأ المعياري في القياس، بينما في نظرية الاستجابة للفقرة تزودنا بتقدير للخطأ المعياري للقياس عند كل مستوى من مستويات القدرة، ويمكن باستخدامها تحديد مدى إسهام كل فقرة في تحديد دقة الاختبار (Hambelton & Swaminathan, 1985).

ويشير كروكر وألجينا (Crocker & Algina, 1987) أن دالة المعلومات لها فوائد متعددة مثل مقارنة اختبارات مختلفة تقيس السمة نفسها، أو مقارنة طريقتين مختلفتين لتصحيح الاختبار نفسه، بالإضافة إلى مقارنة استراتيجيات اختبار مختلفة مثل المقارنة بين الاختبار التقليدي والاختبار التكيفي وذلك من خلال كمية المعلومات التي يعطيها كل اختبار، لذا جاء السؤال الرابع.

وفيما يتعلق بالإجابة عن السؤال الرابع؛ حول تقصي أثر عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار) في دالة المعلومات للاختبار، فقد تم إيجاد منحنيات دالة المعلومات للاختبارات المختلفة في عدد خطواتها، ولكل شكل من أشكال التوزيع لصعوبة الخطوات، ويبين الشكل رقم (٤) هذه المنحنيات للاختبارات الثلاثة المختلفة في عدد خطواتها، ولكل شكل من أشكال التوزيع.



❖ المحور الأفقي: القدرة؛ المحور الرأسي: كمية المعلومات

الشكل رقم (٤). منحنيات دالة المعلومات للاختبارات تبعا لعدد خطوات الأسئلة وشكل التوزيع لصعوبة الفقرة

يلاحظ من الشكل رقم (٤) أن الاختبار ذا الخمس خطوات أعطى أكبر كمية من المعلومات سواء كان شكل التوزيع لصعوبة الخطوات طبيعياً أو ملتوياً نحو اليمين، أو ملتوياً نحو اليسار، يليه الاختبار ذو الخطوات الأربع، وأخيراً الاختبار ذو الخطوات الثلاث؛ وهذا يعني أنه بزيادة عدد الخطوات للاختبار فإن المساحة تحت المنحنى تزداد. وأن كمية المعلومات التي يعطيها الاختبار ذو الخمس خطوات كانت أكبر من كمية المعلومات التي يعطيها الاختبار ذو الخطوات الأربع، وكذلك ذو الخطوات الثلاثة، وعند جميع مستويات القدرة، ولكل أشكال التوزيع لصعوبة الأسئلة. ويلاحظ من الشكل رقم (٥) أيضاً أن دالة المعلومات الأكثر تسطحا كانت للأسئلة المكونة من ثلاث خطوات، ودالة المعلومات الأكثر تحدياً كانت للأسئلة التي عدد خطواتها خمسة. كما يلاحظ أن الاختبار ذا شكل التوزيع الملتوي نحو اليسار يقدم أكبر كمية من المعلومات، ويقدم أقصى كمية من المعلومات، وعند مستويات القدرة جميعها، يليه الاختبار ذو شكل

التوزيع الملتوي نحو اليمين، وأخيراً الاختبار ذو شكل التوزيع الطبيعي. واتفقت هذه النتيجة مع ما توصل إليه رايت وماسترز (Wright and Masters, 1982).

ويلاحظ من الشكل رقم (٤) أن قيم دالة المعلومات التي تقدمها الاختبارات المختلفة في عدد خطواتها، وشكل التوزيع لصعوبة خطواتها تكون أكبر ما يمكن عند مستوى القدرة صفر لوجيت؛ بمعنى أن الاختبارات تزودنا بمعلومات أكثر فاعلية عند الأفراد ذوي القدرة المتوسطة بينما تكون قيم دالة المعلومات التي تقدمها الاختبارات أقل ما يمكن عن مستويات القدرة العالية والمتدنية رغم اختلاف شكل التوزيع، بمعنى أن الاختبارات تزودنا بمعلومات قليلة عند الأفراد ذوي القدرات العالية والقدرات المتدنية. كما يلاحظ من الشكل رقم (٤) أن قيمة دالة المعلومات للاختبارات المختلفة في عدد خطواتها وشكل توزيع صعوبتها تتزايد تدريجياً بتزايد القدرة (θ)، حتى تصل إلى أقصى قيمة ممكنة لها عندما تكون قيمة القدرة (صفراً) لوجيت، وهي تقابل متوسط الصعوبة للأسئلة صفر لوجيت، وهذا ما يتطابق مع ما هو متوقع من نموذج التقدير الجزئي حيث إن دالة المعلومات تكون أقصى ما يمكن عندما تكون قيمة ($\theta = b$) وذلك لكل من النموذجين الأحادي المعلمة والثنائي المعلمة. وهذا ما تؤكده قيمة الخطأ المعياري للتقدير التي كانت أقل ما يمكن عند متوسط الصعوبة لفقرات لكل الاختبارات، ويلاحظ أن كل الاختبارات تقدم أكبر كمية من المعلومات عند مستويات القدرة المتوسطة، وتقدم أقل كمية من المعلومات عند مستويات القدرة العالية والمتدنية، بمعنى أن كل الاختبارات تزودنا بمعلومات قليلة عن الأفراد ذوي القدرات العالية والمتدنية. وقد اختلفت نتيجة هذه الدراسة مع دراسة دود وكوخ (Dodd and Koch, 1987)، التي أبانت أن الأسئلة التي لها العدد نفسه من الخطوات تعطي القدر نفسه من المعلومات، حيث لوحظ بأن الأسئلة التي لها العدد نفسه من الخطوات قد أعطت كميات مختلفة من المعلومات، وعند جميع مستويات القدرة. وكذلك تم تقدير قيم دالة المعلومات للاختبارات المختلفة في عدد خطواتها عند كل مستوى من مستويات القدرة وتبعاً لشكل التوزيع والجدول رقم (١٠) يبين تقديرات قدرات الأفراد

بوحدة اللوجيت وقيم دالة المعلومات عند مستويات مختارة من القدرة للاختبارات المختلفة في عدد خطواتها، وشكل التوزيع لصعوبة الأسئلة.

الجدول رقم (١٠)

تقديرات قدرات الأفراد بوحدة اللوجيت وقيم دالة المعلومات عند مستويات مختارة من القدرة للاختبارات المختلفة في عدد خطواتها، وشكل التوزيع لصعوبة الأسئلة

	عدد خطوات الاختبار وشكل التوزيع								
	ثلاث خطوات			أربعة خطوات			خمسة خطوات		
القدرة	طبيعي	ملتو لليمين	ملتو لليسار	طبيعي	ملتو لليمين	ملتو لليسار	طبيعي	ملتو لليمين	ملتو لليسار
-٨	٠,٠٩	٠,٠٤	٠,٢١	٠,٠٩	٠,٠٤	٠,٢١	٠,٠٩	٠,١	٠,١٤
-٧	٠,٢٢	٠,١٢	٠,٤٨	٠,٢٦	٠,١٢	٠,٥١	٠,٢٦	٠,٢٧	٠,٣٦
-٦	٠,٥٤	٠,٣٢	٠,٨٢	٠,٦٨	٠,٣٢	١,٠٨	٠,٦٨	٠,٧١	٠,٨٥
-٥	١,١٦	٠,٧٨	١,٣٤	١,٦	٠,٨١	٢,٢٧	١,٦	١,٦٥	١,٩٨
-٤	٢,٣٩	١,٩٣	٢,١٥	٣,٧١	٢,١٧	٤,٥٣	٣,٧١	٣,٨٥	٤,٠٦
-٣.	٤,٢١	٤,٤٨	٣,٢٩	٧	٥,٨١	٧,٥٦	٧	٨,٥٦	٧,١٦
-٢	٧,١	٩,٢٧	٤,٧٣	١٠,٩٢	١٣,٥٧	٩,٢٢	١٠,٩٢	١٥,١٣	١٠,٥
-١	١١,٠٢	١٢,٢٧	٧,٠١	١٤,٥٧	١٨,٢٧	١٠,٦	١٤,٥٧	١٦,٥٦	١٤,٨٤
٠	١٤,٠٩	١١,٩٢	٩,٦٢	١٧,٧٩	١٦,٣١	١٣,٦٩	١٧,٧٩	١٧,٦	١٨,٧٢
١	١٢,١٨	٩,٥٤	٩,٥٩	١٧,٣٤	١٢,٧٨	١٦,٧٧	١٧,٣٤	١٥,٥٣	١٩,٨٦
٢	٨,٠٢	٦,٦٥	٧	١١,١٣	٧,٩٩	١٣,٧	١١,١٣	٧,٧١	١٤,٨٢
٣	٤,١١	٤,٣٦	٤,٠٦	٦,٣٧	٥,٠٦	٧,٣١	٦,٣٧	٤,١٤	٦,٤١
٤	٢,١٩	٢,٥٤	٢,٠٧	٣,٤٨	٣,٢٨	٣,٣٩	٣,٤٨	٢,١٨	٢,٧٢
٥.	١,١٢	١,٤٦	٠,٩٢	١,٧	١,٩٥	١,٤٥	١,٧	١,١٢	١,١٢
٦	٠,٥٥	٠,٨	٠,٣٩	٠,٧٥	٠,٩٩	٠,٦٢	٠,٧٥	٠,٥٨	٠,٤٧
٧	٠,٢٢	٠,٣٨	٠,١٤	٠,٣٣	٠,٤٨	٠,٢٣	٠,٣٣	٠,٢٦	٠,١٨
٨	٠,٠٩	٠,١٧	٠,٠٥	٠,١٢	٠,٢	٠,٠٩	٠,١٢		٠,٠٦

يبين الجدول رقم (١٠) أن القيمة القصوى لدالة المعلومات بوحدة اللوجيت للاختبارات المختلفة في عدد خطواتها، وعند أشكال التوزيع المختلفة، كانت أقصى ما يمكن عند مستوى القدرة صفر لوجيت. ويلاحظ من هذه القيم أن الاختبار ذا الخطوات الخمس يقدم أكبر قدر من المعلومات، وأن أدنى قيم قصوى للمعلومات تقدمها أسئلة الاختبار ذي الخطوات الثلاث، ويمكن استنتاج أن دالة المعلومات للفقرة تزداد بزيادة عدد خطوات الاختبار بصورة ملحوظة، وبالتالي يزداد المقياس ثباتاً ودقةً. وهذه النتيجة تساعد مطوري المقاييس في اختيار الفقرات الأنسب بناءً على منحنى دالة المعلومات بزيادة عدد الخطوات، كما تساعد في تحسين الاختبار وزيادة دقته وثباته إذا كانت دالة معلوماته متدنية وذلك بزيادة عدد خطواته، كما بينت النتائج أن الاختبار ذا الخطوات الثلاث أقل دقة في إعطاء معلومات عن تقدير القدرة لتذبذب قيم دالة معلومات أسئلته مقارنةً بقيم دالة معلومات لباقي الاختبارات. ويمكن الاستفادة من دالة المعلومات للاختبار في حساب الكفاءة النسبية للاختبار مقارنة باختبار آخر. وقد عرّف لورد (Lord, 1980) الكفاءة النسبية للاختبار Y بالنسبة للاختبار X بأنها النسبة بين دالة المعلومات للاختبارين، ويعبر عنها كالآتي:

$$RE\{Y, X\} = \frac{I\{\theta, Y\}}{I\{\theta, X\}}$$

حيث $I\{\theta, Y\}$، $I\{\theta, X\}$: هي دالة المعلومات للاختبارين X,Y على الترتيب.

ودرجات الاختبارين X,Y يمكن أن تكون الدرجات على اختبارين مختلفين عند مستوى القدرة نفسه، أو يتم الحصول على هذه الدرجات من الاختبار نفسه بطريقتين. وتعرف الكفاءة النسبية فقط عندما نأخذ النسبة بين دالة المعلومات للاختبارين X,Y عند مستوى القدرة نفسه، وتختلف الكفاءة النسبية لاختبارين باختلاف مستوى القدرة. وفي ضوء ذلك تم حساب الفاعلية النسبية (Relative Efficiency) للاختبارات الثلاثة المختلفة في عدد خطواتها، بالنسبة إلى الاختبار ذي الخطوات الثلاث وذلك للإجابة عن السؤال الخامس.

وللإجابة عن السؤال الخامس للدراسة: هل تختلف الكفاءة النسبية باختلاف عدد خطوات السؤال (ثلاث خطوات، أربع خطوات، خمس خطوات) وشكل التوزيع لصعوبة الخطوات (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار)؟ تم إيجاد قيمة دالة المعلومات عند مستويات مختارة من القدرة بحيث تغطي متصل القدرة، وذلك عند أشكال التوزيع المختلفة، وتم الإشارة لها في الجدول رقم (١٠). وللإجابة عن السؤال تم إيجاد الكفاءة النسبية بالنسبة إلى الاختبار ذي الخطوات الثلاثة، وعند أشكال التوزيع المختلفة. ويبين الجدول رقم (١١) يبين قيم الكفاءة النسبية للاختبارات المختلفة في عدد الخطوات بالنسبة إلى الاختبار ذي الخطوات الثلاث عند قيم مختلفة للسمة (?) تبعا لشكل التوزيع.

جدول رقم (١١)

قيم الكفاءة النسبية للاختبارات المختلفة في عدد الخطوات بالنسبة إلى الاختبار ذي الخطوات الثلاث عند قيم مختلفة للسمة (?) تبعا لشكل التوزيع

القدرة	شكل التوزيع								
	طبيعي			ملتو لليمين			ملتو لليسار		
	RE4/3	RE5/3	RE5/4	RE4/3	RE5/3	RE5/4	RE4/3	RE5/3	RE5/4
-٨	١,٠٠	١,٠٠	١	١,٠٠	٢,٥٠	٢,٥٠	١,٠٠	٠,٦٧	٠,٦٧
-٧	١,١٨	١,١٨	١	١,٠٠	٢,٢٥	٢,٢٥	١,٠٦	٠,٧٥	٠,٧١
-٦	١,٢٦	١,٢٦	١	١,٠٠	٢,٢٢	٢,٢٢	١,٣٢	١,٠٤	٠,٧٩
-٥	١,٣٨	١,٣٨	١	١,٠٤	٢,١٢	٢,٠٤	١,٦٩	١,٤٨	٠,٨٧
-٤	١,٥٥	١,٥٥	١	١,١٢	١,٩٩	١,٧٧	٢,١١	١,٨٩	٠,٩٠
-٣.	١,٦٦	١,٦٦	١	١,٣٠	١,٩١	١,٤٧	٢,٣٠	٢,١٨	٠,٩٥
-٢	١,٥٤	١,٥٤	١	١,٤٦	١,٦٣	١,١١	١,٩٥	٢,٢٢	١,١٤
-١	١,٣٢	١,٣٢	١	١,٤٩	١,٣٥	٠,٩١	١,٥١	٢,١٢	١,٤٠
٠	١,٢٦	١,٢٦	١	١,٣٧	١,٤٨	١,٠٨	١,٤٢	١,٩٥	١,٣٧
١	١,٤٢	١,٤٢	١	١,٣٤	١,٦٣	١,٢٢	١,٧٥	٢,٠٧	١,١٨
٢	١,٣٩	١,٣٩	١	١,٢٠	١,١٦	٠,٩٦	١,٩٦	٢,١٢	١,٠٨
٣	١,٥٥	١,٥٥	١	١,١٦	٠,٩٥	٠,٨٢	١,٨٠	١,٥٨	٠,٨٨

تابع/ جدول رقم (١١)

قيم الكفاءة النسبية للاختبارات المختلفة في عدد الخطوات بالنسبة إلى الاختبار ذي الخطوات الثلاث عند قيم مختلفة للسمة (?) تبعا لشكل التوزيع

القدرة	شكل التوزيع								
	طبيعي			ملتو لليمين			ملتو لليسار		
	RE4/3	RE5/3	RE5/4	RE4/3	RE5/3	RE5/4	RE4/3	RE5/3	RE5/4
٤	١,٥٩	١,٥٩	١	١,٢٩	٠,٨٦	٠,٦٦	١,٦٤	١,٣١	٠,٨٠
٥.	١,٥٢	١,٥٢	١	١,٣٤	٠,٧٧	٠,٥٧	١,٥٨	١,٢٢	٠,٧٧
٦	١,٣٦	١,٣٦	١	١,٢٤	٠,٧٣	٠,٥٩	١,٥٩	١,٢١	٠,٧٦
٧	١,٥٠	١,٥٠	١	١,٢٦	٠,٦٨	٠,٥٤	١,٦٤	١,٢٩	٠,٧٨
٨	١,٣٣	١,٣٣	١	١,١٨	١,٥١	١,٢٩	١,٨٠	١,٢٠	٠,٦٧
الوسط الحسابي	١,٤٠	١,٤٠	١	١,٢٢	١,٥١	١,٢٩	١,٦٥	١,٥٥	٠,٩٢

يلاحظ من الجدول رقم (١١) أن الأوساط الحسابية لقيم الكفاءة النسبية للاختبار ذي الخطوات الخمس بالنسبة للاختبار ذي الخطوات الثلاث عندما تكون أشكال التوزيع لصعوبة الفقرات: (طبيعي، ملتو نحو اليمين، ملتو نحو اليسار) تساوي (١,٤٠، ١,٥١، ١,٥٥) على التوالي، وهذا يعني أن قيم دالة المعلومات للاختبار في الاختبار ذي الخطوات الخمس تساوي بالتقريب ١,٤٩ قيم دالة المعلومات للاختبار ذي الثلاث خطوات لأشكال التوزيع الثلاثة على التوالي. وهذا يعني أن الأفراد وعند مستويات القدرة كلها يحتاجون من (١٥) سؤالاً في الاختبار ذي الخطوات الثلاث لإعطاء القدر نفسه من المعلومات الذي تعطيه (١٠) أسئلة من أسئلة الاختبار ذي الخمس خطوات. كما يلاحظ من الجدول رقم (١١) أيضاً أن الأوساط الحسابية لقيم الكفاءة النسبية للاختبار ذي الخطوات الأربعة بالنسبة للاختبار ذي الخطوات الثلاثة عندما تكون أشكال التوزيع لصعوبة الفقرات: طبيعي، ملتو نحو اليمين، ملتو نحو اليسار تساوي (١,٤٠، ١,٦٥، ١,٢٢) على التوالي؛ وهذا يعني أن قيم دالة

المعلومات للاختبار في الاختبار ذي الخطوات الأربعة تساوي بالتقريب ١,٤٢ قيم دالة المعلومات للاختبار ذي الثلاث خطوات لأشكال التوزيع الثلاثة على التوالي، وهذا يعني أن الأفراد وعند مستويات القدرة كلها يحتاجون إلى (١٤) سؤالاً في الاختبار ذي الخطوات الثلاث لإعطاء القدر نفسه من المعلومات الذي تعطيه (١٠) أسئلة من أسئلة الاختبار ذي الأربع خطوات. وهذه النتيجة تتفق مع ما توصلت إليه دراسة رايت وماسترز (Wright and Masters, 1982)، التي بينت أن الأسئلة المتعددة الخطوات تعطي كمية من المعلومات أكبر من الأسئلة ثنائية الخطوة، وأيدت ذلك دراسة شاهين (١٩٩٣) التي أشارت إلى أن الأسئلة التي لها العدد نفسه من الخطوات تعطي تقريباً الكمية نفسها من المعلومات، وأنه بزيادة عدد الخطوات (عدد التدريجات) يزداد حجم المعلومات، وأن الخطأ المعياري في قياس قدرات الأفراد وفي تقديرات قيم الصعوبة للأسئلة يتناقص بزيادة عدد خطوات التدريج.

ولعل السبب في أن كفاءة الاختبار ذي الخطوات الخمس أفضل منه للاختبار ذي الخطوات الثلاث، أي أن الاختبار ذي الخطوات الخمس كان مناسباً أكثر لقدرات الأفراد، وكان في متناول الأفراد بينما ظهر مستوى الاختبار ذي الخطوات الثلاث أكثر صعوبة من مستوى أفراد العينة، إذ يكون الخطأ المعياري صغيرا كلما اقترب مستوى القدرة للأفراد من مستوى الصعوبة للأسئلة، حيث تزداد كمية المعلومات بنقصان الخطأ المعياري؛ إذ إن زيادة عدد خطوات السؤال ربما تجعل السؤال أكثر سهولة، حيث إن إعطاء فرصتين للإجابة عن السؤال تجعل الأسئلة أكثر سهولة من إعطاء فرصة واحدة، وإعطاء خمس فرص تجعل الأسئلة أكثر سهولة من إعطاء ثلاث فرص. وما يلاحظ أن قيم الكفاءة النسبية للاختبار ذي الخطوات الخمس بالنسبة للاختبار ذي الخطوات الثلاث، وكذلك قيم الكفاءة النسبية للاختبار ذي الخطوات الخمس بالنسبة للاختبار ذي الخطوات الثلاث كانت أعلى ما يمكن في حالة الاختبار شكل التوزيع لصعوبة الأسئلة ملتو نحو اليسار.

الخلاصة والتوصيات

أشارت نتائج الدراسة أن الوسط الحسابي لتقديرات القدرة يزداد وبشكل مطرد بزيادة عدد خطوات أسئلة الاختبار، كما تبين أن الوسط الحسابي للخطأ المعياري يتناقص بزيادة عدد خطوات أسئلة الاختبار؛ أي أن الاختبار ذا الخمس خطوات كان الأكثر دقة في تقدير القدرات للأفراد، حيث كشفت النتائج عن وجود فروق ذات دلالة إحصائية في دقة التقدير لقدرات الأفراد تعزى لعدد خطوات الاختبار، وعدم فروق دالة إحصائياً تعزى لأثر متغير شكل التوزيع، وكذلك وجود أثر للتفاعل بين عدد خطوات الاختبار وشكل التوزيع لصعوبة الخطوات، كما يلاحظ وجود فروق ذات دلالة إحصائية في الأوساط الحسابية للأخطاء المعيارية لتقدير معالم الصعوبة تعزى لعدد خطوات الاختبار، ووجود أثر التفاعل بين شكل التوزيع وعدد خطوات الاختبار على دقة التقديرات لمعالم الصعوبة للفقرات؛ أي أن هناك تأثيراً مشتركا لكل من العاملين؛ أي أن الاختبار ذا الخطوات الخمس كان الأكثر دقة في تقدير معالم الصعوبة للسؤال، إذ إن زيادة عدد خطوات السؤال ربما تجعل السؤال أكثر سهولة، حيث إن إعطاء فرصتين للإجابة عن السؤال تجعل الأسئلة أكثر سهولة من إعطاء فرصة واحدة. كما بينت النتائج أن قيم معاملات التباعد للأفراد وقيم معاملات الثبات لهذا المعامل كانت جميعها عالية، ومتقاربة في الاختبارات الثلاثة المختلفة في عدد خطواتها سواء كان شكل التوزيع طبيعياً أم ملتوياً نحو اليمين أم ملتوياً نحو اليسار. وقد كشفت النتائج عن عدم وجود فروق بين معاملات الثبات للأفراد، تعزى لعدد خطوات الاختبار عندما كان شكل التوزيع لصعوبة الفقرات طبيعياً، وكذلك عدم وجود فروق بين معاملات الثبات للأفراد، تعزى لعدد خطوات الاختبار عندما كان شكل التوزيع لصعوبة الأسئلة ملتوياً نحو اليمين، وكذلك كشفت عن عدم وجود فروق بين معاملات الثبات للأفراد، تعزى لعدد خطوات الاختبار عندما كان شكل التوزيع لصعوبة الأسئلة ملتوياً نحو اليسار.

وخلصت نتائج الدراسة أيضاً إلى أن الاختبار ذا الخمس خطوات أعطى أكبر كمية من المعلومات سواء كان شكل التوزيع لصعوبة الخطوات طبيعياً أو ملتوياً نحو اليمين، أو ملتوياً نحو اليسار، يليه الاختبار ذو الخطوات الأربع، وأخيراً الاختبار ذو الخطوات الثلاث؛ وهذا يعني أنه بزيادة عدد الخطوات للاختبار فإن المساحة تحت المنحنى تزداد، وعند جميع مستويات القدرة، ولكل أشكال التوزيع لصعوبة الأسئلة، كما تبين أن قيم دالة المعلومات للاختبار في الاختبار ذي الخطوات الخمس تساوي بالتقريب ١,٤٩ قيم دالة المعلومات للاختبار ذي الثلاث خطوات لأشكال التوزيع الثلاثة على التوالي؛ وهذا يعني أن الأفراد وعند مستويات القدرة جميعها يحتاجون إلى (١٥) سؤالاً من الاختبار ذي الخطوات الثلاث لإعطاء القدر نفسه من المعلومات الذي تعطيه (١٠) أسئلة من أسئلة الاختبار ذي الخمس خطوات، في حين يحتاجون إلى (١٤) سؤالاً من الاختبار ذي الخطوات الثلاث لإعطاء القدر نفسه من المعلومات الذي تعطيه (١٠) أسئلة من أسئلة الاختبار ذي الأربع خطوات.

وقد أثبت نموذج التقدير الجزئي فعالية في الحصول على نتائج تعد بداية لاستخدام هذا النوع من الأسئلة في إعطاء تقديرات أكثر دقة لقدرات الأفراد وقيم الصعوبة للأسئلة، مما قد يفيد في عمليات التشخيص الدقيقة للأفراد. واعتماداً على ذلك، فإن الباحثيْن يوصيان مطوري الاختبارات، والجهات الحريصة على دقة التقدير لقدرات المفحوصين بإمكانية استخدام اختبارات متعددة التدريج، يتم بناؤها باستخدام نموذج التقدير الجزئي، ومراعاة الزيادة في عدد خطواتها لزيادة الدقة في تقدير قدرات الأفراد، وكذلك دقة التقدير لمعالم الصعوبة، مما ينعكس إيجابا على جودة التقديرات، وربما تسهم هذه التوصية في تمكين بانئي الاختبارات التكيفية (المفصلة) (Adaptive Testing) بشكل أكثر، لا سيما وأن هناك توجهاً إلى استخدام الاختبارات المتعددة التدريج في بناء الاختبارات المقننة، حيث اتصفت النتائج بنوع من الاستقرار، وكانت دقة التقدير أكثر وضوحاً في حالة شكل التوزيع الطبيعي لصعوبة الفقرات. كما يمكن أن تقدم هذه الدراسة دليلاً امبريقياً بأن تقدير

معاملات الثبات للأفراد تبقى قيمته إلى حد بعيد مستقرة، حتى وإن اختلف شكل التوزيع لصعوبة الأسئلة. فعلى سبيل المثال، لا أهمية لأسئلة متعددة الخطوات أو مختلفة في شكل التوزيع لصعوبة خطواتها عند تقدير الثبات للأفراد، كما يوصي الباحثان بإجراء دراسات أخرى مماثلة لكن باستخدام بيانات فعلية ناتجة عن تطبيق الاختبارات على مواد تعليمية مختلفة لا سيما وأن طبيعة الأسئلة قد تختلف من مادة تعليمية إلى مادة تعليمية أخرى، كما يمكن إعادة الدراسة نفسها باستخدام نماذج أخرى من نماذج نظرية الاستجابة للفقرة كالنموذجين الأحادي المعلمة والثنائي المعلمة، وكذلك استخدام مقاييس التقدير والاختبارات بأشكالها المختلفة للوصول إلى نتائج أكثر قابلية للتعميم.

An Investigation of the Effect of Number of Steps and Difficulty Distribution of Polytomous Test Questions on Persons' Ability, Questions Difficulty Estimations and Information Function of the Test According to Partial Credit Model

Dr. Nedal k. Alshraifin **Dr. Zaid s. Bani - Ata**

College of Education
Yarmouk Univesity
H. K. J.

Abstract

The study aims to investigate the effect of number of steps and difficulty distribution of test questions on the accuracy of persons' ability, questions difficulty parameters estimation, and Information function of the test according to Partial Credit Model (PCM). To achieve the objectives of the study, data from three 30-question tests with varying number of steps for every question, (three steps, four steps, and five steps) were developed using WINGEN program. The degree of difficulty between first and last steps in those tests was equal for all tests according to three different distributions related to difficulty of questions: normal, positively skewed, and negatively skewed. Also, data related to responses from 1000 persons were generated. Data were analyzed according to Partial Credit Model and SPSS program. The results showed statistically significant differences in the standard errors of means for persons' ability estimations attributed to the number of test steps and the interaction between the number of test steps and the shape of distribution, and the results also showed no statistically significant differences attributed to the shape of distribution. Moreover, the results showed statistically significant differences in the standard errors of means for difficulty parameters estimation attributed to the number of test steps and the interaction between the number of test steps and the shape of distribution. No statistically significant differences, attributed to the shape of distribution, was shown. In addition, no statistically significant differences were found among estimated persons reliability coefficients, attributed to the number

of test steps or at different shapes of distributions. The results indicated that the five-step test provided the largest amount of information under different shapes of distributions (normal, positively skewed, and negatively skewed), than the fourÙ-step test. Finally, the relative efficiency of the five-step test was higher than that of the three-step test. In other words, the five-step test was more accurate in estimating persons abilities.

المراجع

١ - التقي، أحمد (١٩٩٢). اللاتغير في تقدير معالم قدرات الأفراد ودرجات صعوبة أسئلة المقال من خلال نموذجي التقدير الجزئي وسلم التقدير كحالتين خاصتين من نماذج راش. رسالة دكتوراه غير منشورة، عمان: الجامعة الأردنية.

٢ - الزعبي، محمد(٢٠٠٣). فاعلية نموذج التقدير الجزئي في بناء فقرات بنك أسئلة متعددة الخطوات في مادة الكيمياء للصف الثاني الثانوي العلمي. رسالة دكتوراه غير منشورة، عمان: جامعة عمان العربية للدراسات العليا.

٣ - شاهين، محمد (١٩٩٣). أثر عدد خطوات تصحيح الأسئلة المقالية في تقدير قدرات الأفراد وقيم صعوبة الأسئلة باستخدام نموذج التقدير الجزئي. رسالة دكتوراه غير منشورة، عمان: الجامعة الأردنية.

٤ - الشريف، فاتنة (٢٠٠٦).أثر عدد فئات تدريج ليكرت على خصائص المقياس وخصائص فقراته السكومترية وفقا للنظرية الحديثة في القياس. رسالة دكتوراه غير منشورة، إربد: جامعة اليرموك.

٥ - الشريفين، نضال (٢٠٠٣). مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج و الآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس. رسالة دكتوراه غير منشورة، عمان: جامعة عمان العربية للدراسات العليا.

6 - Adams, R. J. (1988). Applying the partial credit model to educational diagnosis. **Applied Measurement in Education**, 1 (4), 347-378.

7 - Beretvas, Susan N. (2000). **To meet or not to meet standards:** Proficiency estimation using different polytomous IRT models. University of Washington. Degree: PHD.

8 - Crocker, L.& Algina, J. (1987). **Introduction to classical and modern test theory**. New York: Harcourt Jovanovich.

9 - Dodd, B., & Koch, W. (1985, April). Item and scale information functions for the partial credit model. Paper presented at the meeting of the American Educational Research Association, Chicago.

10 - Dodd, B. G., & Kock, W. R. (1987). Effect of variations in steps values on item and test information in the partial credit model. **Applied Psychological Measurement**, 11(4), 371-384.

11 - Drasgow, H. (1982). What constitutes an adequate sample size for calibrating the graded response model (GRM) and the three-parameter logistic (3PL) model. Retrieved January 15, 2012 from http://www.measured progress.org/documents/.../AdequateSampleSize.pdf.

12 - Eggert, S. & Bogehols, S. (2009). Students use of decision-making strategies with regard to socioscientific issues: an application of the Rasch Partial Credit Model. Retrieved December 21, 2011 from www.interscience.wiley.com..

13 - Embretson, S. & Reise, S. (2000). **Item response theory for psychologists**. Mahawah, NJ: Erlbaum

14 - Han, K.T. and Hambleton, R.K (2007). User's Manual for WinGen: Windows Software that Generated IRT Model Parameter and Item Response. Center for Educational Assessment Research Report No 642, Amherst, MA: University of Massachusetts Center for Educational Assessment.

15 - Hakstin, A. R. & Whalen, T. E. (1976). A k- sample significance test for independent alpha coefficients. **Psychometrica**, 41(2), 219- 231.

16 - Hambleton, R. K & Swaminthan, H. (1985). **Item response theory principle and application**. Boston: Kluwer: Nijhoff publishing.

17 - Harris, J., Laan, S. & Mossenson, L. (2009). Applying partial credit analysis to the construction of narrative writing. Retrieved January 15, 2012 from http://www.informaworld.com/smpp/title~content=t775653631, Online publication date: 14 December 2009.

18 - Harris, J., Laan, S. & Mossenson, L. (1988). Applying partial credit analysis to construction of narrative writing tests. **Applied Measurement in Education**, 1 (4), 335-345.

19 - Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. **Applied Psychological Measurement**, 9(2), 139-164.

20 - Lord, F. M. (1980). **Application of item response theory to practical testing problems**. Hillsadle, New Jersey: Lawrence Erlbaum Associates.

21 - Masters, G. N. (1982). A Rasch model for partial credit scoring. **Psychometrika**, 2, 149-174.

22 - Masters, G. N. (1988). The analysis of partial credit scoring. **Applied Measurement in Education**, 1(4), 149-174

23 - Ozaki, K. & Toyoda, H. (2009). Item difficulty parameter estimation using the idea of the graded response model and computerized adaptive testing. **Japanese Psychological Research**, 51(1), 1-12.

24 - Reeve, B. (2004). Applications of item response theory (IRT) modeling for building and evaluating questionnaires measuring patient: Reported outcomes. Unpublished Paper.{On line} http: //outcomes.Cancer.gov/ conference/irt/reeve.pdf.

25 - Wilson, M., & Iventosch, I. (1988). Using partial credit model to investigate responses to structured subtests. **Applied Measurement in Education**, 1(4), 319-333.

26 - Wright, B. D. & Masters, G. N., (1982). **Rating Scale Analysis**. Chicago: MESA Press.