العلاقة بين مؤشر مطابقة الفقرة ومعالمها في ضوء اختلاف شكل التوزيع وفقًا للإنموذج ثُنائي المَعلَم

م.د. سيف خليل اسماعيل ولي

Dr.saif_wali@ibnsina.edu.iq

جامعة ابن سينا للعلوم الطبية والصيدلانية رئاسة الجامعة

الملخص

يهدف البحث الحالي الى:

١. إعداد اختبار الاستدلال العددي لطلبة المرحلة الاعدادية.

٢. الدلالة الاحصائية للعلاقة الارتباطية بين مؤشر مطابقة الفقرة ومعالمها تبعا لاختلاف شكل التوزيع (ملتوي سالب، اعتدالي، ملتوي موجب).

من أجل تحقيق أهداف البحث اتبع الباحث الخطوات العلمية بدءً من صدق الترجمة مرورا بالصدق الوصفي وصولا لتطبيق الاختبار على عينة التحليل الاحصائي البالغة (550) طالب وطالبة، ثم استعمل البرنامج الإحصائي (Bilog – MG3)، لتحليل البيانات، وقد حققت بيانات جميع الفقرات المطابقة الاحصائية للانموذج المستخدم وتم الاحتفاظ بجميع الفقرات والبالغة (15) فقرة، ثم قام الباحث بتطبيق الاختبار على عينة النتائج البالغة (1556) طالب وطالبة، وبينت نتائج البحث وجود علاقة عكسية دالة احصائيا بين (مؤشر مطابقة الفقرة وصعوبتها) في (الشكل ملتوي سالب)، وعدم دلالة العلاقة في (الشكلين اعتدالي وملتوي موجب)، كما وبينت النتائج وجود علاقة موجبة دالة احصائيا بين (مطابقة الفقرة وتمييزها) في (الشكلين اعتدالي وملتوي موجب) وعدم دلالتها في (الشكل ملتوي سالب). وفي ضوء نتائج البحث توصل الباحث الى مجموعة من التوصيات والاستنتاجات والمقترحات.

الكلمات المفتاحية: مؤشر مطابقة الفقرة ، شكل التوزيع ، الانموذج ثُنائي المعلم.

The relationship between the Item matching index and its parameters In light of the difference in distribution Shape of the according to the two-parameter model

Lect. Dr. Saif Khaleel Ismaeel Wali

Ibn Sina University of Medical and Pharmaceutical Sciences

Abstract:

The research aims to:

1. Preparing a numerical reasoning test for middle school students.

2. Significance of the relationship between the item fit index and its parameters according to the different distribution shapes: negatively skewed, normal, positively skewed.

In order to achieve the research objectives, the researcher followed the scientific steps, starting from: the validity of the translation and descriptive validity, then applied the test to the statistical analysis sample amounting to (550) male and female students. Then he used the statistical program (Bilog – MG3) to analyze the data. The data of all items achieved statistical conformity with the model used, and all paragraphs, amounting to (15) items, were retained. Then the researcher applied the test to the sample of results, amounting to (1556) male and female students. The results of the study showed the existence of a statistically significant inverse relationship between (the index and difficulty) in (the negative skewed shape), and its lack of significance in (the normal and positive skewed shapes). The results also showed the existence of a statistically significant positive relationship between (the index and discrimination) in (the normal and positive skewed shapes), and its lack of significance in (the negative skewed shape). In light of the research results, the researcher reached a set of conclusions, recommendations and proposals.

Key words: item Conformity Index, Distribution Shape, Two-parameter Model.

مُشكلَة البحث:

تَختلَف مؤشرات المُطابقَة الاحصائية من حيثُ معنى المؤشر والأَساس الرياضي الَذي بنَي عَليه، إلا أَن جميعها تعمل عَلى كشف أَنماط اجابة الأَفراد غير المطابقة، وهنالك مجموعة من العوامل يمكن أَن تؤثر عَلى أَداء الأَفراد في الاَختبار كالغش والقَلق والتَخمين والمرض وغيرها،

والتي قد تظهر استجابات غير مطابقة مما يؤدي الى تقدير غير دقيق لقدرات الأفراد ومعالم الفقرات (Yen, 1992: 191).

وأن المطابقة الاحصائية تعد من القضايا المهمة والتي تناولتها العديد من البحوث والدراسات، لكن الى الآن لا توجد اجابات حاسمة عن بعض الاستفسارات والأسئلة حول مطابقة البيانات (Embrestone & Reise, 2000: 52). وقد تباينت وجهات النظر حول العوامل التي تؤثر على دقة التقديرات، فمنهم من يرى أن حجم العينة، وطول الاختبار، وشكل الفقرة وطولها، وعدد بدائل الاجابة، وتوزيع مستويات القدرة، يمكن أن تؤثر على دقة التقديرات ومطابقة البيانات (بني عطا والشريفين، ٢٠١٢: ١٥٣).

كما تباينت نتائج الدراسات التي تناولت العلاقة بين مؤشرات مطابقة الفقرة ومعالمها، فبعضها أشارت بعدم وجود علاقة بين مؤشر المطابقة ومعالمها كدراسة كل من (ريس Reise, 1990 / ماونت وشوماخر Mount & Schumacker, 1998 / سيول Seol, 1999)، في حين أشارت البعض الآخر من الدراسات بوجود علاقة بين مؤشر المطابقة وجزء من المعالم وعدم وجود علاقة مع الجزء الآخر كدراسة كل من (دودين Dodeen, 2004 / سينهاري ولو Sinharay & Lu, 2008 / المومني ٢٠١٥).

وهذا التضارب في نتائج الدراسات السابقة دفع الباحث لإجراء الدراسة الحالية، إذ تتلخص مشكلة البحث الحالي بالإجابة على السؤال الآتي: **هل هنالك علاقة ارتباطية بين مؤشر مطابقة الفقرة ومعالمها (صعوبة وتمييز) في ضوء اختلاف شكل التوزيع (ملتوي سالب ، اعتدالي ، ملتوي موجب) ؟.**

أهمية البحث:

يمكن تلخيص الأهمية النظرية والأهمية التطبيقية للبحث الحالي في النقاط الآتية:

١ . يعد موضوع مطابقة فقرات أداة القياس للانموذج اللوغاريتمي المستخدم من الموضوعات التي نالت اهتمام الباحثين المختصين في القياس النفسي والتربوي منذ حوالي أربعين عام (تقريبا) ولا يزال موضع اهتمام ومدار البحث والتقصي.

٢ . يعد البحث الحالي واحدة من العديد من المحاولات العلمية لمواكبة دراسات المشكلات التي تقلل من دقة القياس وفق النظرية الحديثة (نظرية الاستجابة للفقرة)، كما يعد محاولة لمواكبة ومسايرة ما يحدث من تطورات علمية في مجال القياس والتقويم النفسي والتربوي.

٣ . يوفر البحث الحالي أداة قياس موضوعية ألا وهو (اختبار الاستدلال العددي) يمكن استعماله ليس فقط لعينة البحث الحالي (طلبة المرحلة الاعدادية) بل لشريحة واسعة من المجتمع، كالقبول في بعض الوظائف في الشركات، كما أنه يعد اثراء للمكتبة العلمية إذ يكون الاختبار في متناول الباحثين والدارسين لاستعماله.

٤ . التوصل لنتيجة يمكنها أن تحسم الجدل حول اختلاف وجهات النظر ونتائج الدراسات في موضوع مهم كموضوع العلاقة الارتباطية بين مؤشر مطابقة الفقرة ومعالمها وفق اختلاف شكل توزيع القدرة.

أهداف البحث: يهدف البحث الى:

١. إعداد اختبار الاستدلال العددي (Numerical Reasoning Test) لطلبة المرحلة الاعدادية، وفقا للانموذج ثنائي المعلم.

٢. الدلالة الاحصائية للعلاقة الارتباطية بين مؤشر مطابقة الفقرة ومعالمها تبعا لاختلاف شكل التوزيع (ملتوي سالب، اعتدالي، ملتوي موجب)، وينبثق من هذا الهدف الفرضيتين الآتيتين:

أ. لا توجد علاقة ذات دلالة احصائية بين مؤشر مطابقة الفقرة وصعوبتها تبعا لاختلاف شكل التوزيع (ملتوي سالب / اعتدالي / ملتوي موجب).

ب. لا توجد علاقة ذات دلالة احصائية بين مؤشر مطابقة الفقرة وتمييزها تبعا لاختلاف شكل التوزيع (ملتوي سالب / اعتدالي / ملتوي موجب).

حدود البحث: يتحدد البحث الحالي بـ:

١. طلبة المرحلة الإعدادية، الدراسة الصباحية، للعام الدراسي (2024-2025)، ولكلا الجنسين.

٢. اختبار الاستدلال العددي Numerical Reasoning Test، والذي أعدته مؤسسة الاختبارات العالمية (Job Test Prep).

٣. البرنامج الإحصائي (Bilog – MG3).

٤. ثلاثة أشكال للتوزيع (ملتوي سالب ، اعتدالي ، ملتوي موجب).

تحديد المصطلحات:

أولاً. مطابقة الفقرة: عرفها (دوجلاس Douglas, 1982): تعني ملائمة الفقرة للانموذج المستخدم أي اتساق المشاهدات الامبريقية (المشاهدات التجريبية التي تنتج عن عملية القياس) مع المشاهدات المتوقعة (المشاهدات التي تنتج عن الانموذج) (Douglas, 1982: 32).

- **التعريف النظري لمطابقة الفقرة:** يعرفه الباحث نظريا بأنه مطابقة البيانات المستمدة من الفقرات للانموذج المستخدم أي الاتساق بين البيانات الملاحظة وبين ما هو متوقع من الانموذج.
- **التعريف الاجرائي لمطابقة الفقرة:** يعرفه الباحث اجرائيا بأنه الدلالة الاحصائية لمطابقة البيانات المستمدة من الفقرات للانموذج ثنائي المعلم.

ثانيا. معالم الفقرة: عرفها (علام ٢٠٠٢): تشير للخصائص القياسية للفقرة، تنبثق عن الانموذج ثنائي المعلم وهي كل من (الصعوبة والتي تعني نقطة على متصل القدرة تمثل

احتمالية الاجابة الصحيحة للفرد على الفقرة) و(التمييز وتعني ميل المنحنى المميز للفقرة، ويدل الميل الحاد على التمييز المرتفع والميل البسيط على التمييز المنخفض) (علام، ٢٠٠٢: ٤٥).

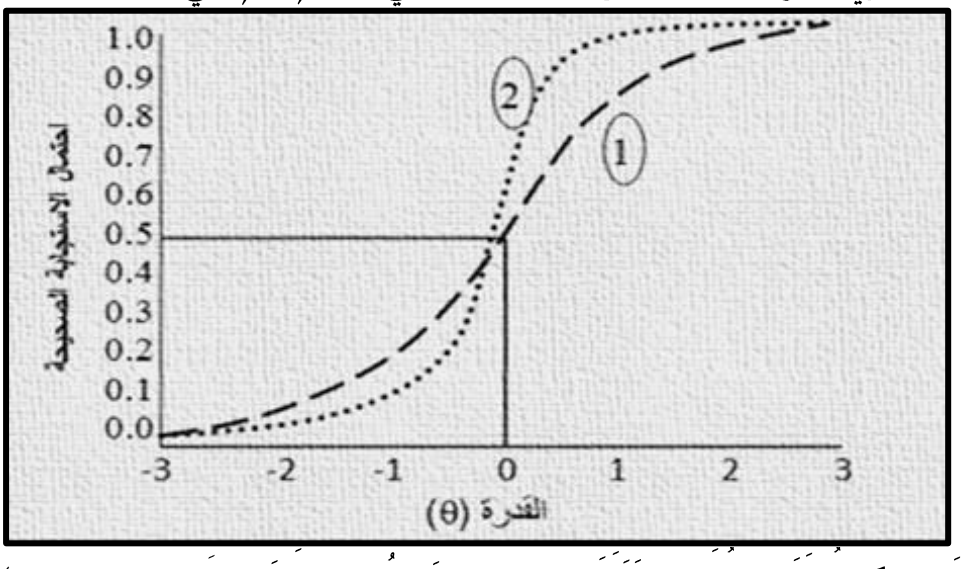
ثالثًا. الانموذج ثُنائي المعلم: عرفها (بيرنبوم Birnbaum, 1968): بأنه أحد النماذج اللوغاريتمية لنظرية الاستجابة للفقرة والذي يمثل الفقرة بواسطة دالة الترجيح اللوغاريتمي ويسمح للفقرات الاختلاف في معلمين (الصعوبة والتمييز)، ويفترض انعدام التخمين وأن الأفراد ذوي القدرة المنخفضة لا يجيبون على الفقرة الصعبة اجابة صحيحة (Hambleton et all, 1991: 12-15).

الخلفية النظرية

الانموذج ثُنائي المعلم **Two Paramtric Model – 2PL:**

هو أحد النماذج الاستاتيكية التراكمية ثُنائية الدرجة أُحادية البُعد في نظرية الاستجابة للفقرة اقترحه العالم (بيرنبوم Birnbaum, 1968)، فقد أضاف بيرنبوم الى هذا الانموذج معلم آخر وهو معلم التمييز فضلاً عن معلم الصعوبة الموجود في إنموذج راش، وأشار بيرنبوم الى أن هذا الانموذج يفترض انعدام التخمين ويسمح لفقرات الاختبار أن تختلف في كل من معلمي (الصعوبة والتمييز) أي أنه يقوم بتقدير قيم صعوبة الفقرات وقيم تمييز الفقرات، ولا يعين أوزانا متساوية في تقدير مستويات القدرة أو السمة (عكس إنموذج راش)، ويترتب على ذلك إمكانية تقاطع المنحنيات المميزة لفقرات الاختبار المعد وفق هذا الانموذج (علام، ٢٠٠٥: ٧١). ويمكن أن نعزو سبب اضافة معلم التمييز الى صعوبة ايجاد مجموعة من الفقرات تُميز بدرجة واحدة بين مستويات القدرة أو السمة التي يقيسها الاختبار (علام، ٢٠٠١: ٢١٢). كما ويفترض الانموذج ثُنائي البارامتر عدم تأثر اجابات الأفراد المختبرين بعامل التخمين، وأن احتمال التوصل الى الاجابة الصحيحة يقل كلما قلت قدرة الفرد (علام، ١٩٨٦: ١١٥). ويعتمد هذا الانموذج على الفرق بين القدرة أو السمة التي يمتلكها الفرد (S) (القدرة أو السمة الكامنة المراد تقديرها والتي تقف خلف إجابات الفرد)، ودرجة صعوبة الفقرة التي يرغب الاجابة عليها والتي يرمز لها (b)، مضروبا في درجة التمييز (a) (التقي، ٢٠٠٩: ٢٢). ويعتمد تأثير هذا الفرق في احتمال الاجابة الصحيحة على قوة تقدير القيمة التمييزية للفقرة، إذ يكبر هذا التأثير عندما تكون تقدير قيمة تمييز الفقرة مرتفعا، كما يمكن الحصول على تقديرين مختلفين لنمطي استجابة لهما نفس الدرجة الكلية، فعندما يجيب الفرد اجابة صحيحة في الفقرات التي تمييزها مرتفع ويخفق في الفقرات التي تمييزها منخفض يؤدي ذلك الى تقدير قدرة مرتفع، لذلك في هذا الانموذج لا يكون للفقرات أوزانا متساوية في تقدير مستويات القدرة أو السمة (Embrestone & Reise, 2000: 51-52). وأن فكرة معلم التمييز تقوم على وصف الطريقة التي تميز فيها الفقرة بين المستجيبين الذين تفوق قدراتهم مستوى صعوبة الفقرة والمستجيبين الذين تكون قدراتهم

أدنى من مستوى صعوبة الفقرة، وينعكس هذا على شدة انحدار المنحنى المميز للفقرة في منتصفه، فالفقرة التي يكون منحناها المميز شديد الانحدار فأنها تميز بشكل جيد بين المستجيبين، أما إذا كان منحناها المميز منبسط فأنها تكون أقل تمييزاً (Baker, 2001: 22). لو افترضنا وجود فقرتين تقدير صعوبة كل منهما يساوي صفر، وتقدير تمييز الاولى يساوي (1) والثانية يساوي (2)، عندها يمكن تمثيل المنحنى الذي يدل على احتمالية الاجابة الصحيحة للمستجيبين ذوي القدرات المختلفة عن كل منهما كما في شكل (1) (التقي، ٢٠١٣: ٢٤).



شكل (1) منحنيان مميزان لفقرتين في الانموذج ثنائي المعلم (التقي، ٢٠١٣: ٢٤).

ويتبين من الشكل أعلاه أنه كلما زادت قيمة صعوبة الفقرة زاد انحدار المنحنى، ولأن معلم التمييز في هذا الانموذج متغير لكل فقرة فأن ميلان المنحنيات لا تتساوى لذلك يمكن أن تتقاطع (Lord, 1980: 84).

الاستدلال العددي **Numerical Reasoning:**

يمكن القول أن الاستدلال العددي هو احدى مهارات التفكير الاستدلالي، وتتمثل هذه المهارة في جوهرها على كل ما يتعلق باستخدام الأرقام، أي سهولة ودقة إجراء العمليات العددية والحسابية والسرعة في حل المسائل المختلفة ومرونة في استخدام عمليات الجمع والطرح والضرب وغيرها (الأشقر، ٢٠١١: ١١٢). ويشير الاستدلال العددي الى قدرة الفرد على ادراك العلاقات العددية والتعرف عليها وعلى أنماطها، وإجراء العمليات الحسابية، وفهم وتحليل المعلومات الرقمية والبيانات الاحصائية والرسوم البيانية والجداول والرموز والنسب المئوية، واستخلاص استنتاجات منطقية واستخدامها في حل المشكلات ثم إتخاذ قرارات مستنيرة مبنية على حقائق رقمية (العثامين، ٢٠١٥: ١٥). وفيما يخص قياس الاستدلال العددي فانه من الممكن قياسه بعدة اختبارات، منها: اختبار الجمع (القدرة على الاضافة العددية والرقمية)، واختبار العلاقات

المحذوفة (القُدرة على ادراك العلاقات العددية)، واختبار الضرب الناقص والقسمة الناقصة (القُدرة على ادراك المتعلقات العددية) (الأشقر، ٢٠١١: ١١٣).

منهجية البحث واجراءاته

أولاً. منهج البحث:

تم اعتماد المنهج الوصفي المقارن للأسباب، والذي يُعد أنسب المناهج ملائمة لأهداف البحث.

ثانياً. مجتمع البحث:

تألف مجتمع البحث من طلبة المرحلة الاعدادية في محافظة بغداد العام الدراسي (2025-2024) في المدارس الحكومية الصباحية فقط، والذين بلغ عددهم (215320) طالب وطالبة يتوزعون حسب المديرية والتخصص والصف والجنس، كما موضح في جدول (1).

جدول (1) مجتمع البحث موزع بحسب المديرية والصف والجنس

الصفوف / المديريات	الصف الرابع		الصف الخامس		الصف السادس		المجموع			النسبة المئوية
	ذكور	إناث	ذكور	إناث	ذكور	إناث	ذكور	إناث	مجموع	
الرصافة الاولى	7290	6210	7069	6022	7733	6587	22092	18819	40911	%19
الرصافة الثانية	6523	5557	6325	5388	6918	5893	19766	16838	36604	%17
الرصافة الثالثة	4988	4249	4837	4120	5290	4507	15115	12876	27991	%13
الكرخ الاولى	5372	4576	5209	4437	5697	4854	16278	13867	30145	%14
الكرخ الثانية	8825	7518	8558	7290	9360	7973	26743	22781	49524	%23
الكرخ الثالثة	5372	4576	5209	4438	5697	4853	16278	13867	30145	%14
المجموع	38370	32686	37207	31695	40695	34667	116273	99047	215320	%100
	71056		68902		75362					

ثالثاً. عينة البحث:

يتطلب البحث واجراءاته عدة عينات، سيتم توضيح كُل منها بحسب الإجراءات المتبعة في البحث.

رابعاً- أداة البحث:

تبنى الباحث (اختبار الاستدلال العددي Numerical Reasoning Test) والذي أعدته مؤسسة الاختبارات العالمية (Job Test Prep). يحتوي الاختبار على نموذج رقمي يعتمد على تفسير الرسوم البيانية والجداول الرقمية، يتكون من (15 فقرة) لكل منها خمسة بدائل أحدها الاجابة الصحيحة، ويكون تصحيح الاختبار (0) للإجابة الخاطئة (1) للإجاب الصحيحة، وبذلك فأن أعلا درجة يمكن ان يحصل عليها المستجيب هي (15) وادنى درجة هي (0) (www.jobtestprep.com).

خامساً- إجراءات إعداد الاختبار:

أ. صدق الترجمة:

لأجل التحقق من سلامة الترجمة عرض الباحث الاختبار بصيغته الأصلية على مختص باللغة الانكليزية من أجل ترجمته الى اللغة العربية، ثم عرضه على مختص آخر من أجل

أعادته الى اللغة الإنكليزية، ثم عرض النسختين (الأصلية والمترجمة من العربي الى الانكليزي) على مختص ثالث للمقارنة بينهما، وقد تبين ان النسختين متطابقتين.

ب. التحليل المنطقي لفقرات الاختبار:

قام الباحث بعرض الاختبار وتعليماته على (11) محكم مختص في المقاييس والاختبارات والعلوم التربوية والنفسية، لأجل التأكد من الفقرات (صياغة وجودة ومضمونا)، ووفقا لآرائهم تم التعديل على بعضها، وقبول جميع الفقرات استنادا لقيمة مربع كاي كمحك للرفض أو القبول، وكذلك بمعيار (80%) كنسبة لاتفاق المختصين.

ج. تجربة وضوح الفقرات والتعليمات:

قام الباحث بتطبيق الاختبار على عينة عددها (50) من طلبة المرحلة الاعدادية، لمعرفة مدى وضوح تعليمات الاختبار وفقراته، وكذلك الوقت المستغرق في الاجابة، فتراوح وقت الإجابة بين (12 – 19) دقيقة بمتوسط (16 دقيقة) أما تعليمات الاختبار وفقراته فكانت واضحة ومفهومة لدى أفراد العينة.

د. التحليل الاحصائي:

عينة التحليل الإحصائي: بلغت عينة البحث (550) طالب وطالبة، اختيرت بطريقة عشوائية متناسبة من المديريات الست التابعة لمحافظة بغداد، مع مراعاة نسب كل مديرية والجنس والتخصص والصف في المجتمع الاصلي، وجدول (2) يوضح ذلك.

جدول (2) عينة التحليل الاحصائي موزعة بحسب المديرية والصف والجنس

الصفوف / المديريات	الصف الرابع		الصف الخامس		الصف السادس		المجموع			النسبة المئوية
	ذكور	إناث	ذكور	إناث	ذكور	إناث	ذكور	إناث	مجموع	
الرصافة الاولى	18	16	18	15	19	18	55	49	104	%19
الرصافة الثانية	17	14	15	15	18	15	50	44	94	%17
الرصافة الثالثة	13	10	12	10	14	12	39	32	71	%13
الكرخ الاولى	14	12	14	10	15	12	43	34	77	%14
الكرخ الثانية	22	20	22	20	24	19	68	59	127	%23
الكرخ الثالثة	14	12	14	10	15	12	43	34	77	%14
المجموع	98	84	95	80	105	88	298	252		%100
	182		175		193		550			

تحقق افتراضات الانموذج ثنائي المعلم: هنالك مجموعة افتراضات رئيسة في نظرية الاستجابة للفقرة، ينبغي التحقق منها عند استعمال أي من نماذجها، لذلك تحقق الباحث منها كما يأتي:

1- افتراض أُحادية البعد:

للتحقق من هذا الافتراض اعتمد الباحث على التحليل العاملي والذي يكشف عن العوامل التي يقيسها الاختبار، ومن خلال (برنامج SPSS) استعمل التحليل العاملي الاستكشافي بطريقة المكونات الأساسية اسلوب التدوير المتعامد الفاريماكس، وبين التحليل عامل واحد بجذر كامن

بلغ (4.335) ويفسر (28.9%) من التباين الكلي، ثم بعدها ثلاثة عوامل ثانوية كل منها له جذر كامن يفوق الـ (1) ويفسرن جميعا (59.62%) من التباين، وجدول (3) يوضح ذلك.

جدول (3) الجذور الكامنة والتباين المفسر والتباين الكلي للعامل السائد والعوامل الثانوية

رقم العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين التراكمي
الأول	4.335	28.9%	28.9
الثاني	1.899	12.66%	41.56
الثالث	1.517	10.11%	51.67
الرابع	1.193	7.95%	59.62

ووفقًا لما اقترحه جوتمان بحسب (محك كايزر) ورأي رياكس فقد تحقق أُحادية البعد، لأن التحليل العاملي أظهر عاملاً واحدا يتميز على ثلاث عوامل صغيرة، إذ بلغ الجذر الكامن للعامل الأول أكثر من (20%) من التباين الكلي، وهذا مؤشر جيد لأُحادية البعد (Wiberg, 2004: 5). فضلاً عن مؤشر التشبع على العامل العام، فهو يشير إلى العلاق بين المتغير والعامل (أبو حطب وآخرون، ٢٠٠٨: ١٩٦). ويتبين من جدول (4) إن كل الفقرات يعد تشبعها بالعامل العام مقبول بحسب محك جيلفورد، إذ كانت جميعها أكبر من (0.30) (فرج، ١٩٩١: ١٠٤). وبذلك تم التحقق من مؤشرين لأحادية البعد.

جدول (4) قيم تشبع فقرات اختبار الاستدلال العددي بالعامل العام

الفقرة	قيمة التشبع بالعامل العام	الفقرة	قيمة التشبع بالعامل العام	الفقرة	قيمة التشبع بالعامل العام	الفقرة	قيمة التشبع بالعامل العام
1	0.471	5	0.637	9	0.541	13	0.618
2	0.385	6	0.487	10	0.494	14	0.458
3	0.582	7	0.523	11	0.566	15	0.397
4	0.574	8	0.739	12	0.463		

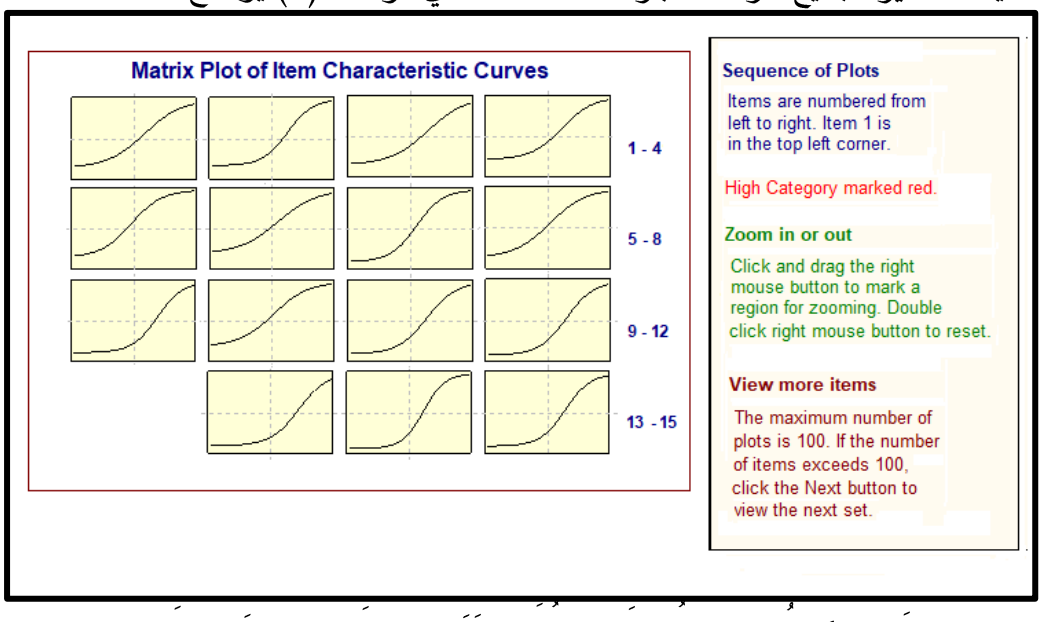
٢- افتراض الاستقلال الموضعي:

تم التحقق من هذا الافتراض من طريق التحقق من افتراض أُحادية البعد إذ أن الافتراضان متشابهان ومرتبطان ارتباطًا وثيقًا، وإن تحقق أُحادية البعد يعد مؤشرا للاستقلال الموضعي (علام، ٢٠٠٥: ٦٤).

٣- افتراض المنحنى المميز للفقرة:

يصف المنحنى المميز لكل فقرة طبيعة العلاقة بين قدرة الفرد وأداءه على الفقرة ويعتمد شكلها على معالم الفقرة (صعوبتها و تمييزها) وعلى قدرة الفرد، ويمثل المنحنى المميز انحدار درجة

الفقرة على قدرة الفرد وهو يأخذ شكل حرف (S)، وإن التغيير في انحدار المنحنى يعبر عن تمييز الفقرة والتغيير في الجزء المحصور بين المنحنى والمحور الأفقي يعبر عن صعوبة الفقرة (Hambleton, 1990: 97). وباستخدام برنامج (Bilog – MG3) تم الحصول على المنحنيات المميزة لجميع فقرات اختبار الاستدلال العددي، وشكل (1) يوضح ذلك.



شكل (1) يوضح المنحنيات المميزة لفقرات اختبار الاستدلال العددي

٤- افتراض التحرر من عامل السرعة:

تم التحقق من هذا الافتراض من خلال حساب نسبة الأفراد الذين أجابوا عن جميع الفقرات فبلغت نسبتهم (97%)، إذ أشار (هامبلتون Hambleton) الى ان هذا الافتراض يمكن ان يتحقق إذا تمكنوا (75%) من الأفراد المفحوصين من اجابة جميع الفقرات، أو تتم الاجابة عن (80%) من الفقرات من قبل جميع الأفراد (Hambleton et al, 1991: 13).

التحليل الاحصائي وفق الانموذج ثنائي المعلم Two Paramtric Model – 2PL:

تم استعمال برنامج (بايلوج Bilog – MG3) الذي يقوم التحقق من ملائمة فقرات الاختبار للانموذج المستخدم، إذ قام الباحث بإدخال اجابات (546) فرد على (15) فقرة بعد أن تم استبعاد استمارات الأفراد الذين حصلوا على درجة كاملة والذين حصلوا على درجة صفر، ولم يجد الباحث أي فقرة أجابوا عليها جميع الأفراد بشكل صحيح أو بشكل خاطئ. وقد أظهر التحليل (مخرجات البرنامج) المخرجات الآتية:

١- **معالم الفقرات:** يتضمن هذا الملف احصاءات قيم مربع كاي (χ^2) وكذلك معالم الفقرات (الصعوبة والتمييز) والأخطاء المعيارية في تقدير كل منها، كما يوضحها جدول (5) الآتي:

جدول (5) قيم معالم الفقرات وأخطاءها المعيارية ومربع كاي وفق الانموذج ثُنائي المعلم

الفقرة	معالم الفقرات				مربع كاي		درجة الحرية	المطابقة
	الصعوبة	خ.المعياري	التمييز	خ.المعياري	المحسوبة	الجدولية		
1	-0.876	0.0544	1.311	0.1340	11.563	15.51	8	ملائمة
2	0.964	0.0823	1.040	0.1123	10.245	12.59	6	ملائمة
3	1.032	0.0487	0.927	0.1031	9.893	15.51	8	ملائمة
4	-0.735	0.0396	0.887	0.0976	10.166	15.51	8	ملائمة
5	1.438	0.0641	1.200	0.1265	11.388	12.59	6	ملائمة
6	-1.191	0.0950	0.768	0.0896	8.521	12.59	6	ملائمة
7	0.922	0.0772	0.901	0.0984	9.758	14.07	7	ملائمة
8	2.121	0.0562	1.430	0.1311	10.594	14.07	7	ملائمة
9	-0.681	0.0618	1.187	0.1417	11.937	12.59	6	ملائمة
10	-1.378	0.0855	1.658	0.1528	9.916	12.59	6	ملائمة
11	0.744	0.0717	0.941	0.0994	12.052	15.51	8	ملائمة
12	-2.059	0.0491	1.532	0.1335	8.743	14.07	7	ملائمة
13	0.793	0.0367	0.871	0.0902	10.446	12.59	6	ملائمة
14	-0.993	0.0931	0.952	0.9364	13.271	15.51	8	ملائمة
15	0.947	0.0748	1.606	0.1483	10.335	14.07	7	ملائمة

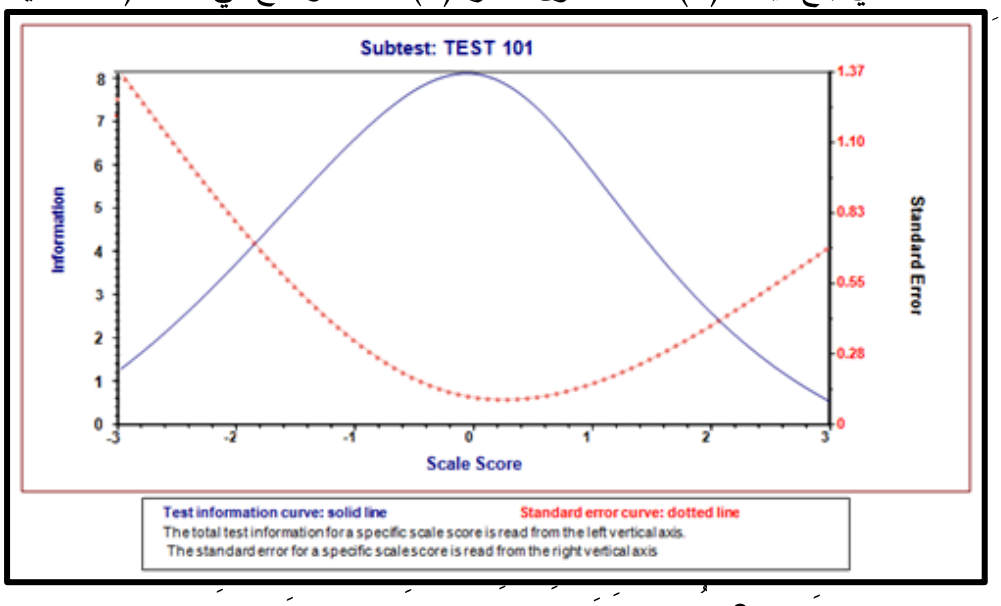
ويتضح من الجدول أعلاه ان قيم مربع كاي المحسوبة لجميع الفقرات غير دالة احصائيا إذ أنها أصغر من القيم الحرجة، وعليه تم الاحتفاظ بجميع الفقرات إذ أن جميع فقرات اختبار الاستدلال العددي مطابقة للانموذج ثُنائي المعلم.

٢- **قُدرات الأفراد:** يتضمن هذا الملف القُدرة المحسوبة لـ(546) درجة كُلية خام، وكذلك الخطأ المعياري في تقدير القُدرة لكُل الدرجات، وتم حساب المتوسط الحسابي والانحراف المعياري وتحديد أعلى وأدنى قيمة لقيم كُل من القُدرة (THETA) والخطأ المعياري (TSTD)، موضحة في جدول (6) التالي:

جدول (6) الاحصاءات الوصفية لقيم تقديرات القُدرة وأخطاءها المعيارية

عدد الأفراد	القيم	أعلى قيمة MAX	أدنى قيمة MIN	المتوسط الحسابي	الانحراف المعياري
546	القدرة (THETA)	3.368	-3.149	0.984	1.807
	الخطأ المعياري (TSTD)	0.577	0.115	0.216	0.422

٣- **دالة معلومات الاختبار:** يقدم برنامج (بايلوج Bilog – MG3) دالة معلومات الاختبار (TIF) وتمثل هذه الدالة مقدار المعلومات التي يقدمها الاختبار عند قيمة معينة للقدرة (θ)، وأن معلومات الاختبار هو معكوس مربع الخطأ المعياري في تقدير القدرة، وهذا يعني كلما ارتفعت معلومات الاختبار تؤدي لدقة أكثر في التقديرات، وأن أقصى درجة من المعلومات قدمها اختبار الاستدلال العددي بلغ قيمته (8) عند مستوى القدرة (0)، كما موضح في شكل (2) التالي:



شكل (2) يوضح دالة معلومات اختبار الاستدلال العددي

الخصائص القياسية لإختبار الاستدلال العددي:

أولًا. صدق الاختبار **Validity Test:**

هنالك بعض المؤشرات للدلالة على خاصية الصدق تم الاستناد اليها في صدق اختبار الاستدلال العددي ، منها (الصدق الوصفي، الصدق الوظيفي، صدق انتقاء السلوكي) وفيما يلي توضيح لكيفية التحقق من تلك المؤشرات لأنواع الصدق الثلاثة:

١- الصدق الوصفي: تم التحقق من الصدق الوصفي من طريق عرض فقرات اختبار الاستدلال العددي وتعليماته على مجموعة من المحكمين، وقد قام الباحث ببعض التعديلات الطفيفة إستنادا الى ملاحظاتهم، وفي المجمل العام اتفق جميع المحكمين على صلاحية جميع الفقرات، وبذلك تحقق مؤشر الصدق الوصفي لاختبار الاستدلال العددي.

٢- الصدق الوظيفي: تم التحقق من الصدق الوظيفي من طريق مؤشرات ملائمة الفقرات للإنموذج المستخدم، إذ ان برنامج (بايلوج Bilog-MG) يقدم احصاءات مربع كاي (كا^2)، وقد تم اعتمادها كمؤشر للصدق الوظيفي، وتبين ان جميع فقرات اختبار الاستدلال العددي ملائمة للإنموذج المستخدم (الانموذج ثنائي المعلم Two Paramtric Model – 2PL) كما مبينة في جدول (5).

٣- صدق انتقاء النطاق السلوكي: تحقق الباحث من هذا النوع من الصدق من طريق مؤشر التحليل العاملي، اذ ان هذا الاجراء أظهر وجود عامل واحد لاختبار الاستدلال العددي وأن جميع الفقرات كانت تشبعها على العامل العام أكبر من (0.30) وهو قيمة تشبع جيدة ومتفق عليها، وجدولين (4 - 3) يبينان ذلك، وهذا يعني فقرات الاختبار جميعها تشترك في بنية الاختبار لقياس سمة كامنة واحدة وتنتمي لنطاق سلوكي واحد، وهذا مؤشر جيد بأن الاختبار صادق في تعريف وقياس السمة التي صمم من أجلها.

ثانيا. ثبات الاختبار Reliability Test:

أكدت نظرية الاستجابة للفقرة ان الاختبارات والمقاييس ذات الفقرات الجيدة من الممكن أن تكون أكثر ثباتا (علام، ٢٠٠٥: ٥٦). ووفقًا لهذه النظرية هنالك بعض المؤشرات للدلالة على ثبات الاختبار تم الاستناد اليها في ثبات اختبار الاستدلال العددي، منها (مؤشر نسبة التباين، ومؤشر دالة المعلومات) وفيما يلي توضيح لكيفية التحقق من تلك المؤشرين:

١- مؤشر نسبة التباين: يعتمد الثبات على المجتمع المدروس وكذلك على أداة القياس، وان ثبات مجموعة من القياسات هو نسبة التباين لهذه القياسات، وهذا هو التباين الحقيقي إذ هنالك مصدرين للتباين (التباين الحقيقي وتباين الخطأ) وأن مجموعهما يساوي التباين الملاحظ، فالثبات يساوي نسبة التباين الحقيقي الى التباين الملاحظ (Wright, 1996: 472). وقام الباحث بحساب الثبات وفقا لنسبة التباين الحقيقي (التباين تقديرات قدرة الأفراد) الى التباين الملاحظ (الحقيقي + الخطأ)، كما مبين في جدول (7) التالي:

جدول (7) قيمة معامل الثبات لاختبار الاستدلال العددي (وفق مؤشر نسبة التباين)

عدد الأفراد	تقديرات القدرة (THETA)		تقديرات الخطأ المعياري (TSTD)		الثبات (r)
	الانحراف المعياري	التباين σ^2_T	الانحراف المعياري	التباين σ^2_E	
546	1.807	3.265	0.422	0.178	0.948

٢- مؤشر دالة المعلومات: تقدم نظرية الاستجابة للفقرة عدة مزايا أهمها معالجتها للثبات وخطأ القياس، من طريق دالة معلومات لكل فقرة (أنستازي وأوربينا، ٢٠١٥: ١١٥) يمكن بتجميعها أن تشكل دالة معلومات الاختبار الكلي على طول متصل السمة (Martin, 2010: 163). وأن دالة معلومات الاختبار والمنحنى الدال عليها تعد مؤشرا لثبات تقدير القدرة على اختبار ما (Hambleton et al, 1991: 93). واعتمد الباحث على دالة معلومات اختبار الاستدلال العددي في تحديد ثبات القدرة، كما مبين في جدول (8) التالي:

جدول (8) قيم المعلومات والخطأ المعياري للقياس والثبات لاختبار الاستدلال العددي

معلومات الاختبار الكلي	مقدار المعلومات عند القدرة (0)	الخطأ المعياري (SEM)	تباين الخطأ المعياري $(SEM)^2$	الثبات (r)
	8	0.353	0.125	0.875

عينة النتائج:

بعد اعداد الإختبار والتحقق من مطابقة فقراته لافتراضات الانموذج واستخراج خصائصه القياسية، ونظرا لمطابقة جميع الفقرات للانموذج وعدم استبعاد أي منها، اعتمد الباحث اجابات عينة التحليل الاحصائي البالغة (546 فرد) فضلا عن تطبيق الاختبار على عينة بلغ عددها (1010 فرد) اختيروا بطريقة عشوائية متناسبة من المديريات الست التابعة لمحافظة بغداد، مع مراعاة نسب كل مديرية والجنس والتخصص والصف في المجتمع الاصلي، ليصبح عدد عينة النتائج النهائية (1556) طالب وطالبة، وجدول (9) يوضح ذلك.

جدول (9) عينة البحث النهائية موزعة بحسب المديرية والصف والجنس

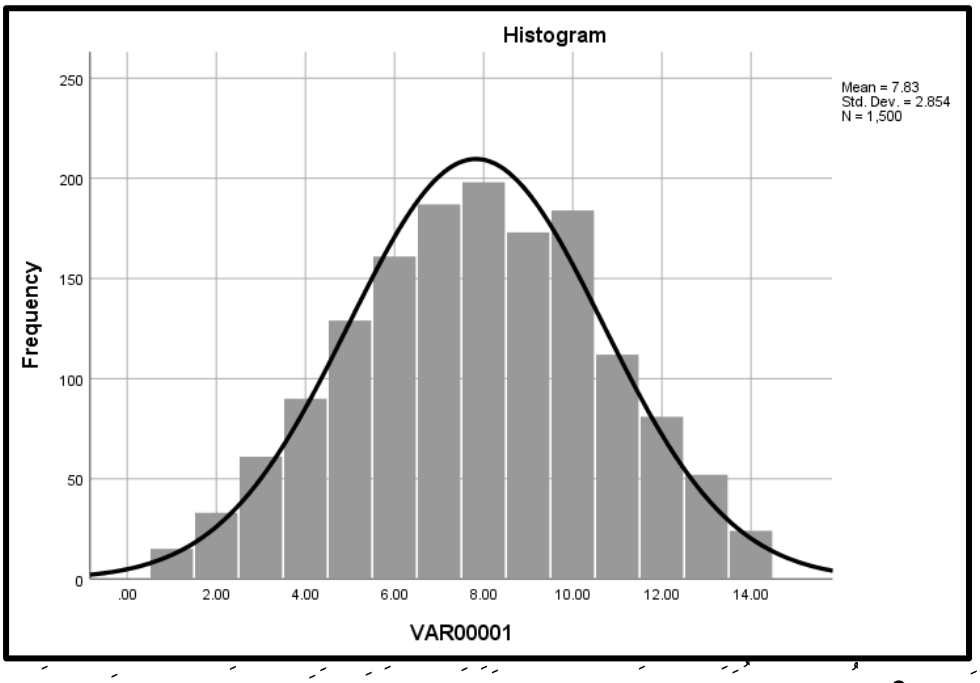
الصفوف / المديريات	الصف الرابع		الصف الخامس		الصف السادس		المجموع			النسبة المئوية
	ذكور	إناث	ذكور	إناث	ذكور	إناث	ذكور	إناث	مجموع	
الرصافة الاولى	51	47	51	42	55	51	157	140	297	%19
الرصافة الثانية	47	38	42	42	51	42	140	122	262	%17
الرصافة الثالثة	38	30	34	30	38	34	110	94	204	%13
الكرخ الاولى	38	34	38	30	42	34	118	98	216	%14
الكرخ الثانية	64	55	64	55	68	55	196	165	361	%23
الكرخ الثالثة	38	34	38	30	42	34	118	98	216	%14
المجموع	276	238	267	229	296	250	839	717		%100
	514		496		546		1556			

أخذ الباحث (١٥٠٠) اجابة وقام بحساب المؤشرات الاحصائية لدرجاتهم من أجل توفر معلومات عن شكل توزيع الدرجات فكانت قيم المؤشرات كما مبينة في جدول (10) الآتي:

جدول (10) المؤشرات الاحصائية لدرجات عينة البحث النهائية

المؤشر الاحصائي	الدرجة	المؤشر الاحصائي	الدرجة	المؤشر الاحصائي	الدرجة
عدد العينة	1500	الانحراف . م	2.854	المدى	13.000
المتوسط	7.827	التباين	8.147	أعلى درجة	14.000
الوسيط	8.000	الالتواء	-0.079	أدنى درجة	1.000
المنوال	8.000	التفرطح	-0.528	عدد الفقرات	15

وبالرغم من أن المؤشرات الاحصائية دلت على أن الدرجات تتوزع توزيعا اعتداليا، إذ أشار (فيركسون ١٩٩١) أن التوزيع التكراري يكون متماثلا عندما تتساوى او تقترب قيم كل من (الوسط والوسيط والمنوال) (فيركسون، ١٩٩١: ٧٨). إلا أن الباحث تحقق من اعتدالية توزيع القدرة للعينة، من طريق اجراء اختبار (كولمجروف–سميرنوف) باستخدام برنامج (Spss)، وأظهرت النتائج عدم الدلالة الاحصائية عند مستوى دلالة (0.01)، إذ كانت مستوى دلالة الاختبار (0.022) وهي أكبر من مستوى دلالة (0.01)، ويشير ذلك الى أن الدرجات تتوزع توزيعا اعتداليا، وشكل (3) يوضح التوزيع التكراري لدرجات العينة.



شكل (3) يوضح المدرج التكراري لتوزيع درجات العينة على اختبار الاستدلال العددي

وبعد أن تأكد الباحث من أن درجات العينة البالغة (1500) تتوزع توزيعا اعتداليا رتب الدرجات ترتيبا تنازليا من أعلى درجة الى أدنى درجة، ثم قسمها الى ثلاث أقسام باعتبار أول (500) درجة بأنهم الأفراد ذوي القدرة المرتفعة (ملتوي سالب) وثاني (500) درجة بأنهم الأفراد ذوي القدرة المعتدلة (اعتدالي) وثالث (500) درجة بأنهم الأفراد ذوي القدرة المنخفضة (ملتوي موجب).

الوسائل الإحصائية:

١. **برنامج (Spss)**، تم استعماله لاستخراج كل من: (التحليل العاملي للتحقق من أحادية البعد)، (اختبار (كولمجروف–سميرنوف) للتحقق من اعتدالية التوزيع)، (المؤشرات الاحصائية لعينة النتائج النهائية)، (الوسط الحسابي والانحراف المعياري لقيم تقدير القدرة وأخطاءها المعيارية)، (معامل ارتباط بيرسون للتعرف على العلاقة بين المؤشر مطابقة الفقرات ومعالمها).

٢. **البرنامج الاحصائي (بايلوج Bilog–MG3)**، تم استعماله للتحقق من (ملائمة فقرات الاختبار للانموذج المستخدم)، ولاستخراج (تقديرات معالم الفقرات، وتقديرات قدرات الأفراد، ودالة المعلومات).

٣. **وسائل احصائية استعملت يدويا**، وهي:(الخطأ المعياري للقياس بحسب المعادلة $SEM=1/\sqrt{I(\theta)}$)، (معامل الثبات المستمد من نسبة التباين بحسب المعادلة $R=\sigma^2_T / \sigma^2_O$)، (معامل الثبات المستمد من الخطأ المعياري للقياس ودالة المعلومات بحسب المعادلة $R=1-(SEE)^2$... $or...R=1-(1/I(\theta))$)، (اختبار مربع كاي لحساب الصدق الوصفي بحسب المعادلة

$\chi 2 = \sum \frac{(O-E)2}{E}$). (الاختبار التائي لمعرفة دلالة العلاقة الارتباطية بين مؤشر مطابقة الفقرات ومعالمها تبعا لشكل التوزيع).

نتائج البحث

الهدف الأول: **إعداد اختبار الاستدلال العددي لطلبة المرحلة الاعدادية وفقًا للانوذج ثُنائي المعلم.**

من أجل تحقيق هذا الهدف اتبع الباحث الخطوات العلمية بدءً من ترجمة الاختبار والتحقق من صدق ترجمته ثم عرضه على مجموعة من المحكمين للتحقق من الصدق الوصفي لفقرات الاختبار، ثم طبق الاختبار على عينة التحليل الاحصائي البالغة (550) طالب وطالبة، ثم تحقق من افتراضات الانموذج (انموذج ثُنائي البارامتر) ثم استعمل البرنامج الإحصائي بايلوج (Bilog-MG3) لتحليل البيانات، وقد حققت بيانات جميع الفقرات المطابقة الاحصائية للانموذج المستخدم وفقا لقيم مربع كاي للحكم على ملائمتها كما يوضحها جدول (5)، وتم الاحتفاظ بجميع الفقرات والبالغة (15) فقرة، وبذلك تحقق الهدف الأول.

٢. الهدف الثاني: **الدلالة الاحصائية للعلاقة الارتباطية بين مؤشر مطابقة الفقرة ومعالمها تبعا لاختلاف شكل التوزيع (ملتوي سالب / اعتدالي / ملتوي موجب).**

من أجل تحقيق هذا الهدف استعمل الباحث برنامج (بايلوج Bilog – MG3) وقام بإدخال (500) إجابة على (15 فقرة) لكل من (عينة الأفراد ذوي القدرة المرتفعة، وعينة الأفراد ذوي القدرة المعتدلة، وعينة الأفراد ذوي القدرة المنخفضة)، لتحليل بياناتهم والحصول على قيم مؤشر المطابقة وتقديرات صعوبة الفقرات وتقديرات تمييز الفقرات لكل من العينات الثلاث. **وقد وضع الباحث لهذا الهدف الفرضيتين الآتيتين:**

أ. الفرضية الأولى: لا توجد علاقة ذات دلالة احصائية بين مؤشر مطابقة الفقرة وصعوبتها تبعا لاختلاف شكل التوزيع (ملتوي سالب / اعتدالي / ملتوي موجب).

من أجل التحقق من هذه الفرضية أخذ الباحث قيم مؤشر مطابقة الفقرة وقيم تقديرات الصعوبة لكل من العينات الثلاث، واستعمل معامل ارتباط بيرسون من طريق برنامج (Spss) لإيجاد العلاقة الارتباطية بينهما، ثم استعمل الاختبار التائي لمعرفة دلالة العلاقة، وظهرت النتائج كما

جدول (11) معاملات الارتباط بين قيم مؤشر المطابقة وقيم تقديرات الصعوبة

ت	شكل التوزيع	العدد	قيمة معامل الارتباط بين مؤشر المطابقة والصعوبة	القيمة التائية		مستوى الدلالة
				المحسوبة	الجدولية	
١	ملتوي سالب	15	-0.524	2.218	2.160	دالة
٢	اعتدالي	15	-0.197	0.724	2.160	غير دالة
٣	ملتوي موجب	15	0.106	0.384	2.160	غير دالة

بينت نتائج هذه الفرضية انه عندما يكون شكل التوزيع (اعتدالي) أو (ملتوي موجب) فان العلاقة بين مؤشر مطابقة الفقرة و(معلم الصعوبة) غير دالة احصائيا، وتتفق هذه النتيجة مع دراسة كل من (ريس Reise, 1990، ماونت وشوماخر Mount & Schumacker, 1998 / سيول Seol, 1999 / دودين Dodeen, 2004 / سينهاري ولو Sinharay & Lu, 2008). أما عندما يكون شكل التوزيع (ملتوي سالب) فان العلاقة بين مؤشر مطابقة الفقرة و(معلم الصعوبة) هي علاقة عكسية دالة احصائيا، ولا تتفق هذه النتيجة مع اي من الدراسات السابقة باستثناء (دراسة المومني ٢٠١٥)، التي أشارت وجود علاقة ضعيفة بين مؤشر المطابقة ومعلم الصعوبة عندما تكون فترات القدرة ١٢ فترة ولكنها غير دالة. ويرى الباحث هذه النتيجة هي متوافقة مع منطق نظرية الاستجابة للفقرة فكلما إرتفعت قدرة الفرد كانت الفقرات أقل صعوبة وبالتالي تحقق الفقرات المطابقة الاحصائية للانموذج المستخدم.

ب. الفرضية الثانية: لا توجد علاقة ذات دلالة احصائية بين مؤشر مطابقة الفقرة وتمييزها تبعا لاختلاف شكل التوزيع (ملتوي سالب / اعتدالي / ملتوي موجب).

من أجل التحقق من هذه الفرضية أخذ الباحث قيم مؤشر مطابقة الفقرة وقيم تقديرات التمييز لكل من العينات الثلاث، واستعمل معامل ارتباط بيرسون من طريق برنامج (Spss) لإيجاد العلاقة الارتباطية بينهما، ثم استعمل الاختبار التائي لمعرفة دلالة العلاقة، وظهرت النتائج كما مبينة في جدول (12) الآتي.

جدول (12) معاملات الارتباط بين قيم مؤشر مطابقة الفقرة وقيم تقديرات التمييز

ت	شكل التوزيع	العدد	قيمة معامل الارتباط بين مؤشر المطابقة والتمييز	القيمة التائية		مستوى الدلالة
				المحسوبة	الجدولية	
١	ملتوي سالب	15	-0. 202	0.743	2.160	دالة
٢	اعتدالي	15	0.613	2.797	2.160	دالة
٣	ملتوي موجب	15	0.568	2.488	2.160	غير دالة

بينت نتائج هذه الفرضية انه عندما يكون شكل التوزيع (ملتوي سالب) فان العلاقة بين مؤشر مطابقة الفقرة و(معلم التمييز) غير دالة احصائيا، أما عندما يكون شكل التوزيع (اعتدالي) أو (ملتوي موجب) فان العلاقة الارتباطية بين مؤشر مطابقة الفقرة و(معلم التمييز) علاقة موجبة دالة احصائيا، وتتفق هذه النتيجة مع دراسة (دوديـن Dodeen, 2004) ومع دراسة (سينهاري ولو Sinharay & Lu, 2008) ومع دراسة (المومني ٢٠١٥)، ويرى الباحث ان هذه النتيجة هي تتفق مع ما جاءت به مضامين نظرية الاستجابة للفقرة ونماذجها إذ أشار (لورد Lord, 1980) الى أنه كلما زادت قيمة صعوبة الفقرة زاد انحدار المنحنى المميز للفقرة، وهذا يشير لارتفاع تمييز الفقرة (Lord, 1980: 84).

الاستنتاجات:

١. مناسبة الانموذج ثنائي المعلم لاختبار الاستدلال العددي، إذ أن بيانات جميع الفقرات مطابقة للانموذج.

٢. اتساق التقديرات الملاحظة المتمثلة بتقديرات قدرة الأفراد، مع التقديرات المتوقعة المتمثلة بالتقديرات الناتجة عن الانموذج ثنائي المعلم.

٣. تزداد مطابقة الفقرات عندما تكون قدرات الأفراد أعلى من صعوبة الفقرات (الى حد ما).

٤. تزداد مطابقة الفقرات عندما ترتفع قيم تقديرات تمييز الفقرات.

التوصيات:

١. استعمال الانموذج ثنائي المعلم في بناء وتطوير اختبارات تقيس القدرات العقلية والجوانب المعرفية.

٢. استعمال اختبار الاستعداد العددي في الدراسات والبحوث بهدف قياسه لدى طلبة المرحلة الاعداد، وكذلك استعماله للقبول في الوظائف في بعض الشركات.

٣. استعمال اختبار الاستعداد العددي من قبل المرشدين التربويين في المدارس الاعدادية للكشف عن قدراتهم العددية من اجل ارشادهم في تحديد مستقبلهم التعليمي.

المقترحات:

١. اجراء دراسة لمعرفة العلاقة بين مؤشر مطابقة الفقرة ومعالمها وفقاً لاختلاف نسب البيانات المفقودة.

٢. اجراء دراسة لمعرفة دقة تقدير معالم الفقرات وفقاً لاختلاف شكل التوزيع.

المصادر العربية والأجنبية:

١.ابو حطب، فؤاد وعثمان، سيد أحمد وصادق، آمال (٢٠٠٨): **التقويم النفسي**، ط٤، القاهرة، مكتبة الأنجلو المصرية.

٢.الأشقر، فارس راتب (٢٠١١): **فلسفة التفكير ونظريات في التعليم والتعلُم**، دار زهران للنشر والتوزيع، المملكة الأردنية الهاشمية.

٣.أنستازي، آنا وأوربينا، سوزانا (٢٠١٥): **القياس النفسي**، ترجمة: صلاح الدين محمود علام، ط١، عمان، دار الفكر ناشرون وموزعون.

٤.بني عطا، زايد والشريفين، نضال (٢٠١٢): أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات، **المجلة الاردنية في العلوم التربوية**، مجلد (٨)، العدد (٢).

٥.التقي، احمد محمد، (٢٠٠٩): **النظرية الحديثة في القياس**، ط١، دار المسيرة، عمان- الأردن.

٦. ..... (٢٠١٣): **النظرية الحديثة في القياس**، ط٢، دار المسيرة للنشر والتوزيع والطباعة، عمان- الاردن.

٧. العثامين، إنعام أحمد (٢٠١٥): **تطوير اختبار للقُدرة الاستدلالية لطَلَبة المرحلة الثانوية في المملكة الأُردنية الهاشمية**، رسالة ماجستير، عمادة الدراسات العليا، جامعة مؤتة.

٨. علام، صلاح الدين محمود (١٩٨٦): **تطورات معاصرة في القياس النفسي**، جامعة الكويت.

٩. ..... (٢٠٠١): **الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية**، القاهرة، دار الفكر للنشر والتوزيع.

١٠. ..... (٢٠٠٢): **القياس والتقويم التربوي والنفسي، أساسياته وتطبيقاته وتوجهاته المعاصرة**, القاهرة, دار الفكر العربي.

١١. ..... (٢٠٠٥): **نماذج الاستجابة للمفردة الاختيارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**، القاهرة، دار الفكر العربي.

١٢. فرج، صفوت (١٩٩١): **التحليل العاملي في العلوم السلوكية**، القاهرة: مكتبة الأنجلو المصرية.

١٣. فيركسون، جورج (١٩٩١): **التحليل الإحصائي في التربية وعلم النفس**، ترجمة هناء محسن العكيلي، بغداد- دار الحكمة للطباعة.

14. Baker, Frank, B. (2001): **The Basics of item responses Theory**, 2ed, ERIC Clearinghouse on Assessment and Evaluation.

15. Dodeen, H. (2004). The relationship between item parameter and item fit. **Journal of Educational Measurement,** 14 (3), 261-270.

16. Douglas, Graham . (1982). **Issues in the Fit of Data to Psychometric Models** .Educational Research and Perspectives. 9 (1), 32-43.

17. Embretson, S.E, Reise, S.P.(2000). **Item response theory for psychologists**. New Jersey :Lawrence Erlbaum Associates inc. publishers.

18. Hambleton,R.K. (1990): **Item Response theory: Introduction And Bibliography**, *Psicothema*, vol,2,No1, pp. 97-107.

19. Hambleton, R.K. ; Swaminthan, H.; Rogers, h.j. (1991). **Fundamentals Of Item Response Theory,** Newbury Park,Ca:Sage. The International Professional Publishers.

20. Lord, M. F. (1980): **Application of Item Response Theory to Practical Testing Problems**. Hillsdale, NJ: Erlbaum.

21. Mount, R. E., & Schumacker, R. E. (1998). Identifying measurement disturbance effects using Rasch item fit statistic and the logit residual index, **Journal of Outcomes Measurement,** 2(4), 338-350.

22. Reise, S. (1990). A comparison of item and person fit method of assessing model data fit. **Applied Psychological Meeasureement,** 14 (2), 127-137.

23. Seol, H. (1999). Detecting differential item functioning with five standardized item-fit indices in the Rasch model, **Journal of Outcomes Measurement,** 3 (3), 233-247.

24. Sinharay, S. & Lu, Y. (2008). A further look at the correlation between item parameters and item fit statistics, **Journal of Educational Measurement,** 45 (1), 1-15.

25. Wiberg, M. (2004). **Classical Test Theory vs. Item Response Theory An Evaluation of The Theory Test in the Swedish Driving-License Test.** Umea University, Department of Educational measurement.

26. Wright, B. D. (1996). **Reliability and separation.** ***Rasch Measurement Transactions*****,** 9, 472.

27. www.jobtestprep.com

28. Yen, W. (1992). Scaling Performance Assessments: Strategies for Managing Local Item Dependence, **Journal of Educational Measurements,** 30(3), 187-213.