تأثير ترتيب الفقرات على دقة تقدير معالم الفقرات وقدرة الافراد تبعا لمستوى قدرة الافراد وعدد الفقرات

أ.م.د. لقاء شامل خلف
liqaa2019@gamil.com
وزارة التربية/ مديرية تربية الرصافة الاولى

الملخص

دفت الدراسة الحالية الكشف عن تأثير ترتيب الفقرات على دقة تقدير معالم الفقرات وقدرة الافراد تبعا لمستوى قدرة الافراد وعدد الفقرات، وللتحقق من اهداف الدراسة اتبعت الباحثة خطوات علمية في إجراءات تحليل اختبار القدرة العقلية المعد من قبل المؤسسة الهندية (Aglasem Schools) للتكنولوجية التعليمية، المكون من (٥٠) فقرة، بدءاً من ترجمة تعليمات وفقرات الاختبار من اللغة الانكليزية الى اللغة العربية ومن ثم إجراء ترجمة عكسية, وللتثبت من صلاحية فقرات الاختبار للبيئة العراقية تم عرضها على الخبراء والمحكمين والبالغ عددهم (١٠) خبراء لمعرفة اراءهم حول صلاحية الفقرات والتعديلات التي اجريت عليها, واعتمدت الباحثة نسبة مئوية بلغت (100%) وبذلك تم التأكد من صلاحية الفقرات, وقد طبق الاختبار على عينة عشوائية من طلبة المرحلة الاعدادية في مدارس المتفوقين والمتفوقات والبالغ عددهم (٦٠٠) طالب وطالبة ، وتم اجراء التحليل الإحصائي للفقرات وفق أنموذج راش باستخدام البرنامج الاحصائي (Jmetrik4.1.1), ولم تستبعد اي فقرة من فقرات الاختبار وتبين ان جميع فقرات الأداة مطابقة لانموذج راش، ولتحقيق الهدف الاول قامت الباحثة بإعداد نموذجين (ترتيب الفقرات حسب معلم الصعوبة ، وترتيبه بشكل عشوائي وتم الاعتماد على التصميم الاصلي للاختبار)، وبعد ذلك تم تطبيق النموذج (المرتب حسب معلم الصعوبة) على نفس عينة التحليل الاحصائي ذاتها، اما النموذج الثاني تم الاعتماد على نفس بيانات التحليل الاحصائي، وبعد ذلك تم تقسيم العينة لكل من نموذجين الى العينة (المرتفعة والمنخفضة)، وتم تحليل البيانات وفق برنامج (Jmetrik4.1.1 Spss) واستخراج الاخطاء المعيارية لمعلم الصعوبة و القدرة لجميع العينات، واظهرت النتائج عدم وجود فروق ذات دلالة احصائية في دقة تقدير معلم الصعوبة وفق نموذج المرتب حسب معلم الصعوبة ، ووجود فروق ذات دلالة احصائية في دقة تقدير معلم الصعوبة وفق نموذج المرتب بشكل عشوائي لصالح القدرة المنخفضة، ووجود فروق ذات دلالة احصائية في دقة تقدير القدرة لصالح القدرة المنخفضة وفقا للنموذجين، وعليه ان فأن القدرة المرتفعة اكثر دقة في دقة تقدير معلم الصعوبة و القدرة من

القدرة المنخفضة ، ولتحقيق الهدف الثاني قامت الباحثة باعداد اختبار قصير مكون من (١٨) فقرة وفق انموذج راش الذي تم سحبها من الاختبار الاصلي (٥٠) فقرة ، وبعد ذلك تم إعداد اربع نماذج لترتيب الفقرات للاختبارين المكون من (٥٠) فقرة و الاختبار القصير (المختزل) المكون (١٨) فقرة ، وبعد ذلك تم تطبيق النماذج على نفس عينة التحليل ذاتها ، وتم تحليل البيانات باستخدام نفس البرامج الاحصائية، واظهرت النتائج عدم وجود فروق ذات دلالة احصائية في دقة تقدير معلم الصعوبة وفق نموذج المرتب حسب معلم الصعوبة، ووجود فروق ذات دلالة احصائية في دقة تقدير لمعلم الصعوبة وفق نموذج المرتب بشكل عشوائي لصالح الفقرات الاقل (١٨) فقرة، ووجود فروق ذات دلالة احصائية في دقة تقدير القدرة لصالح عدد الفقرات الاقل (١٨) فقرة وفقا للنموذجين، وعليه ان فأن الاختبار عدد فقرات (٥٠) اكثر دقة في دقة تقدير معلم الصعوبة و القدرة من عدد فقرات (١٨) ، وفي ضوء نتائج البحث الحالي توصلت الباحثة الى مجموعة من الاستنتاجات والتوصيات والمقترحات.

الكلمات المفتاحية: ترتيب الفقرات، دقة التقدير معالم الفقرات، قدرة الافراد ، مستوى القدرة ، عدد الفقرات

The effect of the arrangement of the vertebrae on the accuracy of estimating the parameters of the vertebrae and the ability of individuals depending on the ability level of individuals and the number of vertebrae

Dr . liqaa shamil khalaf

Ministry of Education\Rusafa First Education

Abstract

The current study aimed to reveal the effect of the arrangement of the vertebrae on specific discrimination of the features of the vertebrae and the ability of individuals depending on the level of ability of the individuals and the number of patients. To verify the objectives of the study, scientific steps were followed in conducting an analysis of the creativity ability test prepared by the Indian institution (Glassim Schools (for educational technology), component Of (50) items, starting with translating the instructions and test items from English to Arabic and then performing a reverse translation.To verify the validity of the test items for the Iraqi environment, they were presented to the (10) experts

and arbitrators to know their opinions about the validity of the items and the modifications that were made to them. The researcher adopted a percentage of (100%), and thus the validity of the items was confirmed. The test was applied to a random sample. Of the middle school students in schools for outstanding male and female students, numbering (600) male and female students The statistical analysis of the items was conducted according to the Rasch model using the statistical program (Jmetrik 4.1.1). No item was excluded from the test. It was found that all items in the tool were identical to the Rasch model. To achieve the first goal, the researcher prepared two models (arranging the items according to the difficulty parameter, and arranging them randomly). The original design of the test was relied upon, after which the model (arranged according to the difficulty parameter) was applied to the same sample for statistical analysis. As for the second model, the same statistical analysis data was relied upon, after which the sample for each of the two models was divided into the (high and low) sample, and the data was analyzed according to the (Jmetrik4.1.1 Spss) program and the standard errors for the difficulty and ability parameters were extracted for all samples, and the results showed no There are statistically significant differences in the accuracy of estimating the difficulty parameter according to the salary model according to the difficulty parameter. There are statistically significant differences in the accuracy of estimating the difficulty parameter according to the random-order model in favor of low ability, and there are statistically significant differences in the accuracy of estimating ability in favor of low ability according to the two models. Therefore, high ability is more accurate in estimating the difficulty parameter and ability than ability low To achieve the second goal, the researcher prepared a short test consisting of (18) items according to the Rasch model, which was drawn from the original test (50 items). After that, four models were prepared to arrange the

items for the two tests consisting of (50) items and the short (abbreviated) test consisting of (50) items. (18) paragraphs, after which the models were applied to the same analysis sample, and the data were analyzed using the same statistical programs.The results showed that there were no statistically significant differences in the accuracy of estimating the difficulty parameter according to the salary model according to the difficulty parameter, and the presence of statistically significant differences in the accuracy of estimating the difficulty parameter according to the random salary model in favor of the least (18) paragraphs, and the presence of statistically significant differences in accuracy. Estimating ability is in favor of a smaller number of items (18) according to the two models. Accordingly, the test with a number of items (50) is more accurate in accurately estimating the parameter of difficulty and ability than a number of items (18). In light of the results of the current research, the researcher reached a set of conclusions and recommendations. And proposals.

Keywords: arrangement of paragraphs, accuracy of estimation, parameters of paragraphs, ability of individuals, level of ability, number of paragraph.

الفصل الاول

مشكلة البحث:-

يعد الاختبار من أدوات القياس و التقويم التي تستخدم في اتخاذ القرارات التربوية ، فكلما زادت حدة القرار الذي يتوقع أن يتم اتخاذه كلما زادت الحاجة إلى أن تكون المعلومات التي نحصل عليها من الاختبار دقيقة وذات صلة وثيقة بالغرض الذي أعد من أجله ، ومن هنا برزت الحاجة الى التخطيط المحكم لإعداد الاختبار سواء فيما يتعلق بالأهداف والغرض منها، او اعداد الفقرات وتحليلها ، بحيث يقدم الاختبار نتائج دقيقة تعكس المستوى الحقيقي للفرد، وبتتبع أدبيات البحث اتضح أن هناك عوامل اخرى لها تاثير واضحا في نتائج الدراسات منها ترتيب وعدد فقرات الاختبار ومستوى القدرة على الخصائص الاديومترية ودقتها، ولقد اظهر جيو (Chiu, 2012) بان اختلاف ترتيب موقع الفقرة يمكن ان يؤثر على خصائص الفقرة مثل (الصعوبة او التمييز), ونتيجة لذلك له تأثير غير مقصود على الدرجة الكلية للاختبار, وهذه

التأثيرات من الممكن ان تقدم نماذج إختبارية مشكوك فيها, ومن المحتمل انتهاك معايير صناعة الاختبار (Chiu, 2012: 2). وبالرغم من أن الدراسات التي تبدو متفقة بان تأثيرات سياق ترتيب الفقرات يمكن أن تؤدي إلى تقديرات متحيزة، لكن ليس هناك إجماع بخصوص طرق مناسبة لاكتشاف وتقدير هذه التأثيرات, وقد تكون جزئيا بسبب الافتقار الى اتفاق مفاهيمي لسياق ترتيب الفقرات (Albano, 2012: 9). ومن هنا انطلقت مشكلة البحث الحالي في ايجاد تأثير ترتيب الفقرات على دقة تقدير معالم الفقرات وقدرة الافراد تبعا لمستوى قدرة الافراد وعدد الفقرات.

اهمية البحث:-

وإن الاختبار بأنواعه ونماذجه المختلفة هو الأداة المستخدم لقياس السمات النفسية للافراد، وبالتالي يتم خلالها ان نحدد كمية القدرة التي يمتلكها في المجال الذي يتم قياسها لذلك سعى الباحثون والدارسون في مجال القياس والتقويم إلى وضع استراتيجية لقياس الخصائص المتعلقه بالافراد ، سواء كانت طرق القياس للجوانب المتعلقة بشخصية الفرد وسلوكه، أو ذكائه وقدراته، أو جوانب معينة من تحصيله الأكاديمي، والتي تسيطر عليها ظروف وعوامل ومتغيرات كثيرة ومتنوعة. وقد تتفاعل هذه العوامل مع بعضها البعض بشكل واضح. (مساعدة، ٢٠١٣: ١)

وأن التطورات الحاصله في الاختبارات بانواعها تبعا للتطورات الحاصلة لنظريات القياس ، مما انعكس ايجابيا في التوجه نحو البحوث التي تستخدم النظريات والطرق الحديثة في القياس (Embrettson & Reise , 2000: 13).ومن اهم هذه النظريات هي نظرية (IRT) والتي استطاعت بنماذجها اللوغاريتمية المتعدده من بناء ادوات تشمل (اختبارات والمقاييس) وتتمتع بخصائص قياسية جيدة من الصدق والثبات والمرونه مما انعكس بشى من الايجاب حول قدرتها على حل الكثير من المشكلات الحاصلة بالاختبارات والمقاييس مما ساعد بانشاء نظام قياسي موضوعي يركز على الفقره نفسها وليس على الاداة ككل نفسها (Hambleton & Swaminathan, 1985 : 8) فقد انتشرت دراسات عده وفق نظرية الاستجابة للفقرة واحد المواضيع المهمة التي تم تناولها في الاونة الاخيرة تأثير عدد الفقرات وترتيبها في المجالات التربوية والنفسية المختلفة، وفي هذا السياق يشير جين Chen, 2010 الى انه أجريت بحوث لتأثيرات ترتيب الفقرة في المجالات المختلفة مثل: (التسويق، والتعليم، والعلوم الطبية). حيث تم بحث تأثيرات ترتيب الفقرات في دراسات اختبارات التحصيل (Achievement Tests), واختبارات الاستعدادات (Aptitude Tests), والاختبارات المكيفة (Adaptive Testing) ومقاييس الاتجاهات(Attitude Measures), من خلال بعض المفاهيم او التأثيرات: كمعيار عدم التحيز (Norm of Evenhandedness) إذ, يميل الأفراد إلى تعديل إجاباتهم على أساس قيمة جوابهم السابق, وتأثير ترتيب الأسئلة الخاصة والعامة أو تأثير الإضافة (Addition

(Effect, وتأثير الطرح (Subtraction Effect) وهو بالنقيض من تأثير الإضافة. إذ, قد يقوم الأفراد بطرح او استبعاد الأسباب المستعملة للإجابة على الفقرة الأولى لضبط استجابتهم للفقرة الثانية. وبحوث الترسيخ والتعديل (Anchoring and Adjustment), وذلك للإشارة الى ان بعضها تبقى ثابتة, وأخرى تتأثر بالسياق او ترتيب الفقرات المكونة لأداة القياس (Chen, 2010: 5- 25)

وتأتي أهمية هذه الدارسة الحالية حول موضوع عدد الفقرات وترتيبها في الاختبار القدرة العقلية حيث إن النتائج التي يتوصل إليها البحث الحالي سيرشد واضعي الاختبارات إلى الطريقة المناسبة لعدد الفقرات وترتيبها للحصول على اختبار يتمتع بخصائص جيدة و مما يقلل من الخطأ المعياري في القياس مما يزيد من دقة القياس.

الأهمية النظرية:

- أهتم البحث الحالي بموضوع (تأثير عدد الفقرات وترتيبها على معالم الفقرات والافراد) وهو من الموضوعات المهمة في مجال القياس التي شغلت اهتمام الدارسين ، ويحتاج إلى مزيد من البحث والتقصي في هذا المجال.
- سيقدم البحث الحالي طرق قياسيه جديداً لكي نحصل الى نتائج اكثر دقة في الخصائص القياسية، معتمداً على نمط ترتيب فقرات الاختبار ومستوى القدرة وعدد الفقرات.

الأهمية التطبيقية:

- يقدم البحث الحالي اختبار يقيس القدرة العقلية, له خصائص قياسية وموضوعية. وهو ايضا يتمتع بأهمية واهتمام الباحثين والمختصين في المجالات التربوية والنفسية.
- عينة البحث متمثلة بشريحة مهمة في المجتمع, وهم طلبة المرحلة الاعدادية في المدارس المتفوقين والمتفوقات، بوصفها مرحلة دراسية مهمة, إذ تقوم هذه المرحلة بإعداد الطلبة معرفياً, ونفسياً, واجتماعياً لمواصلة دراستهم الجامعية.

حدود البحث :-

- طلبة المرحلة الاعدادية في مدارس المتفوقين والمتفوقات للمديريات التابعة لمحافظة بغداد الرصافة (١ ، ٢ ، ٣) للعام الدراسي (٢٠٢٣- ٢٠٢٤).
- اختبار القدرة العقلية المعد من قبل المؤسسة الهندية (Aglasem Schools) للتكنولوجية التعليمية.

أهداف البحث :-

- **الهدف الاول** : التعرف على تأثير ترتيب الفقرات على دقة تقدير معالم الفقرات وقدرة الافراد تبعا لمستوى القدرة، وينبثق من هذا الهدف الفرضيتين الصفريتين الاتيتين:-

١- لايوجد فروق ذات دلالة احصائية في دقة تقدير معلم الصعوبة تبعا لمستوى القدرة (المرتفعة والمنخفضة) للنموذجين.

٢- لايوجد فروق ذات دلالة احصائية في دقة تقدير معلمة القدرة تبعا لمستوى القدرة (المرتفعة والمنخفضة) للنموذجين.

- **الهدف الثاني** : التعرف على تأثير ترتيب الفقرات على دقة تقدير معالم الفقرات وقدرة الافراد تبعا لعدد الفقرات وينبثق من هذا الهدف الفرضيتين الصفريتين الاتيتين:

١- لايوجد فروق ذات دلالة احصائية في دقة تقدير معلم الصعوبة تبعا لعدد الفقرات (١٨، ٥٠) للنموذجين.

٢- لايوجد فروق ذات دلالة احصائية في دقة تقدير معلمة القدرة تبعا لعدد الفقرات (١٨، ٥٠) للنموذجين.

تحديد المصطلحات :-

اولا:- تأثير ترتيب الفقرات عرفه ميللر Miller (2004): "تأثير الاستجابة لأحد الفقرات على الاستجابة أو الاستجابات لفقرة سابقة أو قادمة أو مجموعة فقرات. وتسمى تأثيرات ترتيب الفقرة أيضا بتأثيرات ترتيب السؤال, وتأثيرات السياق والتي بدورها قد توثر على الخصائص القياسية للأداة وفقراتها". (Miller, 2004: 1)

- التعريف الإجرائي لتأثير ترتيب الفقرات:

"الفروق ذات الدلالة الإحصائية المحتملة في الإحصائيات الخاصة دقة تقدير معالم قدرة الأفراد والفقرات, والمستمدة من نتائج البيانات لنماذج ترتيب فقرات الاختبار حسب مستوى القدرة وعدد الفقرات أداة البحث الحالي".

ثانيا :- دقة تقدير:- عرفه الكيلاني (٢٠٠٤):- "هو تعبير يشير الى جودة التقدير التي يميزها الاحتمالية الكبيرة في ان التقدير قريب من القيمة الحقيقية حيث يمكن الوصول الى ذلك باختيار التقدير غير المتحيز الذي يتصف تباينه بأنه اقل تباين من اي تقدير اخر غير متحيز، وذلك باستخدام معدل مربعات الاخطاء او الخطأ المعياري في التقدير (الكيلاني ، ٢٠٠٤:٣٩)".

ثالثا : معالم الفقرات :- هنالك عدة معالم للفقرات منها (الصعوبة ، التمييز ، التخمين) سيتم تعريف المعلم الذي يتم استخرجه وفق انموذج راش **معلم الصعوبة** (Threshold) :- "هي نقطة على متصل السمة الكامنة التي نتوقع ان يكون احتمال الإجابة الصحيحة عن الفقرة (٠.٥٠), وبدون تخمين" (علام, ٢٠٠١: ٢١٨).

- التعريف الاجرائي لتقدير معالم الفقرات:- وهي قيمة التدريجية التي نحصل عليها من تطبيق الدالة الرياضية لنموذج راش احد نماذج استجابة للفقرة عند استجابة المفحوص على مجموعة

من مفردات الاختبار، والتي تقيس الخصائص الاديومترية المتعلقة بالفقرة (كالصعوبة والتمييز لمنحنى خصائص الفقرة).

رابعا : تقدير القدرة : عرفه علام (٢٠٠٥) : "عملية تقدير مقدار السمة التي تكمن وراء اداء الفرد على فقرات الاختبار، ويتميز هذا التقدير بخصائص اللاتغاير اي عدم تغير هذه التقديرات بتغير عينة الفقرات التي يحتويها الاختبار، ويتم تقدير هذه القدرة بطرق متعددة منها طرق تعتمد على الارجحية العظمى وطرق تعتمد على نظرية بييز" (علام، ٢٠٠٥: ٣٤).

- **التعريف الاجرائي لتقدير القدرة** :- وهي قيمة تدريجية التي تنتج من تطبيق الدالة الرياضية لانموذج راش عند استجابة المفحوص على مجموعة من الفقرات الاختبار، والتي تقيس مدى امتلاكه للسمة المراد قياسها، حيث يكون لكل فرد معلم يسمى معلم الفرد أو القدرة. .

الفصل الثاني

الاطار النظري والدراسات السابقة

أثر ترتيب الفقرات :

بدء الاهتمام بالبحث عن تأثير ترتيب المفردة على الاداء في فترة الخمسينات. وبعد انتهاء الحرب العالمية الثانية وبسنوات، حصلت تغييرات وتطورات كثير ومهمة في الاختبار التربوي والنفسي، منها تطوير الحاسوب، والتحليل الإحصائية التي تحولت من النظريات البسيطة إلى الاساليب الحسابية الاكثر فعالة وكفاءة. (Leary, and Dorans, 1985: 387-389). وان الاهتمام في موضوعات تأثير سياق المفردة ليس بجديد على جماعة القياس التربوي والنفسي (Leary & Dorans, 1985: 388). ، وهناك تسميات عديد لترتيب الفقرات ومنها مصطلح (الأسبقية والحداثة والسياق) (Chen, 2010: 2). لقد تم التقصي عن تأثير ترتيب المفردة, فعلى سبيل المثال, وجد كل من وايتلي وديوس Whitely and Dawis, 1976 صعوبات متباينة للمفردات المعروضة على الافراد ضمن صيغ مختلفة وترتيبات مختلفه. ووجد مايرز، وزميليه Meyers, Miller, & Way, 2009 وان التغير الحاصل في ترتيب المفردة بين أشكال الاداة يرتبط بشكل إيجابي بالتغيير الحاصل في صعوبة المفردة (Hartig, & Buchholz, 2012: 419). حيث ظهرت عدة دراسات في القياس التربوي والنفسي بان تباين موقع المفردة اثناء التطبيق على الافراد احتمال تاثيرها على معالم المفردة مثل (الصعوبة), وجد لها تأثير على الدرجة الكلية بشكل غير مباشر (Pommerich & Harris, 2003: 12). تقصى كل من هامبلتون وتروب Hambleton and Traub, 1974 عن مستوى الطلبة في الصف الخامس الثانوي على اختبار لمادة الرياضيات, وتوصلوا الى نتيجة بان متوسط معلم الصعوبة ذات الترتيب من الصعب إلى السهل كان أقل بكثير من الإجابة على المفردات ذات الترتيب العكسي (Chiu, 2012: 3). وفي حال تقديم الاداة للافراد مع ترتيب ثابت للمفردة، لا يمكن العزل بين

تأثير ترتيب وموقع المفردة وصعوبتها, لكون كل مفردة تقدم للافراد بنفس الترتيب، اما اذا قدم الاختبار الافراد بعدة ترتيبات فان هذا الترتيب من المحتمل ابن يؤثر على صعوبة المفردات. وهذا ما توصلت اليه البحوث التي تناولت تأثيرات ترتيب المفردة. (Hartig, & Buchholz, 2012: 419)

معالم الفقرات:

هنالك عدة معالم للفقرات منها (الصعوبة ، التمييز ، التخمين) حيث ان البحث الحالي سيتناول معلم الصعوبة بشي من التفصيل لانه المعتمد في الدراسة الحالية:-

معلم الصعوبة (Threshold) :-

يشير بيكر,(Baker, 2001) من الخصائص التقنية لمنحنى المميز للفقره التي تستعمل لوصف المنحنى نفسه، و صعوبة الفقره الذي بدوره يصف موقع دوال الفقره على طول اداة القدره ومن امثلتها هي دالة الفقره الصعبة وهي متوسط الافراد من ذوي القدره المرتفعة، ودالة الفقره السهله والتي تعبر عن المتوسط للافراد ذوي القدره المنحفضة(Baker, 2001: 7). "وهذا ما يتم التعبير عنه باحتمالية الإجابة الصحيحة على الفقرة مقرونة بمستوى القدرة التي يمتلكها الفرد، ويرمز لها بالحرف اللاتيني (β) (محاسنة، ٢٠١٣: ١٧٠) ". ، عندما يتم تحويل درجات القدرة مقياس محدد لمجموعة من الافراد فان متوسط قدرتهم سيكون (θ) = صفر والانحراف المعياري = ١. قيم معالم الصعوبة هو من (- ٢ إلى + ٢). وهذا يعني أن القيم القريبة من (-٢) تمثل فقرات سهلة للغاية، والقيم القريبة من (+٢) تمثل فقرات صعبة للغاية (ابو حجر، ٣٨-٣٩) .

تقدير قدرة الافراد:-

يعد تقدير القدرة هو التوزيع النظري لمتغير السمة المستمر والذي يعد احد المفاهيم الاساسية لنظرية القياس ويتم فيها القياس لعدد لانهائي من المجتمع، وجاء اهتمام علماء مثل كوتليت وجالتون بالتوزيعات التكراريه وايجاد عدد كبير من الملاحظات على متغير السمات التي تكون قريبة من شكل الجرس. "إن العديد من السمات، مثل الطول والوزن ومستوى الذكاء وما إلى ذلك، إذا تم قياسها على عدد كبير من الملاحظات فإن توزيعها يقترب من التوزيع الطبيعي. ولم يأخذ شكل التوزيع الطبيعي، له مسميات عديدة المنحنى الاعتدالي او التوزيع الجرسي لان يشبه الجرس (كروكر, والجينا,٢٠٠٩: ٤٥) ". ان النظريه الحديثة للقياس تتفق اغلب نماذجها على افتراض وجود سلسلة متصلة من السمات وهذه النمادج تفترض انها تاخذ شكل منحنى لوغارتمبي وهذا مايطلق عليه دالة المنحنى المميز للفقره ومن خلال افتراضات النظريه يمكن تقدير احتمالية اجابة الفرد على مفردة الاختبار بشكلها الصحيح اذا عرفنا موقع الفرد وفق سلسلة متصله من السمات إلا أنهم يختلفون في تصورهم لشكل هذه الدالة، إذ تفترض هذه النماذج أنها تأخذ شكل منحنى لوغاريتمي. وهذا ما يسمى دالة المنحنى المميز للفقرة.

(علام, ٢٠٠١: ٢٠٨)

عدد فقرات الاختبار :

ان لعدد فقرات الاختبار له تأثير على قيمة معامل الثبات يزداد بزيادة طول الاختبار وعدد مفرداته. كما ان قيمة الصدق المرتبطة بمحك يعتمد على مؤشر الثبات، فإن القيمة المتعلقة بالصدق تزداد مع زيادة قيمة الثبات، وتعتمد قيمة معامل الثبات على طول الاداة ، اي عدد فقراته . وكلما زاد عدد الفقرات زاد من قيمة معامل الثبات (ماجد، ٢٠٠٧: ١٠٢-١١٣) أن احد الاسباب التي لها تاثير على دقة تقدير معالم المفردات والافراد هو عدد المفردات في الاختبار وهذا ماشار اليه كل من هامبلتون وكوك ، ومن الممكن أن تؤثر ايضا على (الثبات والصدق)، وفي هذا الموضوع وضح لورد بشكل اوضح أن لطول الاختبار له تاثير على جودة تقدير المعالم .(Hambeleton& Cook, 1983:133) إي بزيادة عدد المفردات ارتفعت قيمة الثبات، وذلك لأن العدد الأكثر من المفردات نحصل على عينة اكبر من السلوك، بينما المفردات القليلة في الاداة تقلل من احتمال استقرار الصفة المراد قياسها، وبالتالي ينخفض قيمة معامل الثبات، وكلما زاد من عدد المفردات ادى زيادة الفرصة أن الخطأ المعياري يلغي بعضه بعضاً، ومن ثم يزداد قيمة التباين، وبالتالي يزداد معامل الثبات (مجيد، ٢٠٠٧ :١٢٠،١٢١) .

نظرية الاستجابة للفقرة

عتبر نظرية الاستجابة للفقرة احدى نظريات القياس الحديثة ولها عدة تسميات ومنها نظرية المنحنى المميز ونظرية الاستجابة للفقرة او نظرية السمات الكامنة، ومن مهام هذه النظرية الاحصائية تحليل مفردات الاختبار وأداء الافراد في الاختبار وكيفية ارتباط الأداء بالقدرات المقاسة بمفردات الاختبار. يمكن أن تكون الاستجابات لمفردة ما منفصلة أو مستمرة ويمكن أن تكون القيم ثنائية أو متعددة. يمكن ان تكون مرتبة او غير مرتبة على شكل فئات أو أقسام. يمكن أن تكون هناك قدرة واحدة أو أكثر من القدرات للأداء في الاختبار، هناك العديد من النماذج التي يمكن من خلالها تحديد العلاقة بين الاستجابة على الفقرة والسمة الكامنة أو القدرات (Hambleton, & Jones, 1993: 40). وان نماذج هذه النظرية (IRT) هي تعتبر من النماذج الإحصائية التي يتم استخدامها من قبل الباحثين لكي تصف سلوكيات استجابة المفحصوين تجاه عدد من الفقرات على نحو مطلق (Johnson, et. al, 2007: 586).

نماذج نظرية الاستجابة للفقرة

هنالك عدة نماذج لنظرية الاستجابة للفقرة (احادية المعلم ، ثنائية المعلم ، ثلاثي المعلم، رباعي المعلم ، خماسي المعلم) ستتناول الباحثة انموذج راش احادي المعلم بشي من التفصيل لانه الانموذج المعتمد في الدراسة الحالية :-

أنموذج راش :

قدم هذا الانموذج العالم الدنماركي جورج راش (1960,Rasch) وتناول راش النموذج الأحادي المعلم تحليل البيانات للاختبارات وذلك من منظور نظرية الاحتمالات. على الرغم من كونه انطلق من إطار مرجعي مختلف إلا أنه في النهاية توصل إلى المنحنى المميز للمفردة والذي يعد نموذجا لوجستيا (بيكر، ٢٠١٠). ويعتبر النموذج أحادي المعلم أبسط نماذج نظرية (IRT) وذلك لأنه يصف كل فقرة بدلالة معلم واحد بعد تحليل بيانات الفقرة التي تم الاستجابة عليها من قبل الأفراد، وهذا المعلم هو موقع الفقرة على متصل السمة الذي يمثل المفهوم والذي يمثل معلم الصعوبة للفقرة، ويفترض هذا النموذج أن جميع الفقرات لها قيمة تمييزية واحدة وأن التخمين يكون في حده الأدنى، أي يفترض أن التخمين يساوي صفرا تقريبا وأن التمييز ثابت لكل فقرات الاختبار وأن صعوبة الفقرة تأخذ قيمة متغيرة. وطبقاً لنظرية الاستجابة للمفردة فإن النموذج الرياضي المعياري للمنحنى المميز للفقرة هو الصيغة التراكمية للدالة اللوجستية التالية:-

$$P_i(\theta) = \frac{e^{D(\theta-bi)}}{1+e^{D(\theta-bi)}}$$

حيث أن:

Pi(θ): احتمالية الاستجابة الصحيحة على الفقرة i لشخص قدرته θ

bi: معلم الصعوبة للفقرة i

D: عامل التدريج =1.7

e: أساس اللوغاريتم الطبيعي = 2.718

(القرني : ٢٠٢٢: ٨-٩)

الدراسات السابقة

دراسة دامبروت (1980 Dambrot) التي هدفت إلى معرفة أثر ترتيب الفقرات والقدرة الاكاديمية على أداء الطالب في الاختبار. ولتحقيق هدف الدراسة، تم تطبيق نموذجين لاختبار تحصيلي على عينة مكونة من (٨٤٧) طالبا وطالبة مسجلين في مساق علم النفس في جامعة أكروفاهيت، حيث تم ترتيب الفقرات في النموذج الأول تصاعديا حسب مؤشر صعوبة الفقرة، في حين رتبت الفقرات عشوائيا في النموذج الثاني. وكشفت نتائج الدراسة عدم وجود فروق ذات دلالة إحصائية في تحصيل الطلبة على أي من النموذجين تعزى لطريقة ترتيب الفقرات. (1980 Dambrot)

دراسة (حجازي ٢٠٢٢) : هدفت هذه الدراسة إلى تقصي أثر ترتيب فقرات اختبار اختيار من متعدد وفق مؤشرات إحصائية في دقة تقدير معالم الفقرات وقدرة الأفراد. ولتحقيق الهدف من الدراسة، تم توليد استجابات ثنائية لِ ١٠٠٠ مفحوص، بواقع ١٠٠ مرة لسبعة نماذج من الاختبار حسب طريقة ترتيب الفقرا ت. وتم تحليل البيانات المولدة؛ وذلك لتقدير معالم الفقرات وقدرة الأفرا د وإيجاد قيم الجذر التربيع لوسط مربعات الإنحرافات، والتحيز لهذه التقديرا ت. وكشفت نتائج الدراسة عن وجود أثر ذي دلالة إحصائية لطريقة ترتيب الفقرات على معالم

الفقرات وقدرة الأفراد، حيث اشارت النتائج إلى أن الوسط الحسابي لقيم الجذر التربيعي لوسط مربعات الإنحرافات لمعالم الفقرات وقدرة الأفراد كان الاقل عندما رتبت الفقرات عشوائيا (النموذج الرابع)، وعندما رتبت تنازليا حسب مؤشر صدق الفقرة (النموذج الخامس)، كما أشارت النتائج إلى أن قيم الوسط الحسابي لمؤشر التحيز لهذين النموذجين كانت قريبة من الصفر (حجازي، ٢٠٢٢).

الفصل الثالث

اولاً – مجتمع البحث:

يتكون مجتمع البحث الحالي من طلبة المرحلة الاعدادية لمدارس المتفوقين والمتفوقات للدراسة الصباحية في محافظة بغداد في مديريات التابعة لتربية الرصافة (١, ٢, ٣) في المدارس الحكومية، للعام الدراسي (٢٠٢٣-٢٠٢٤)، إذ بلغ المجموع الكلي لمجتمع الدراسة (4351) طالبا وطالبة، ويتوزع مجتمع الدراسة بحسب متغير الجنس، بواقع (2530) ذكور ، و(1821) اناثا، وبلغ عدد الطلبة في الرصافة الأولى (1193) بنسبة (٢٨%)، بواقع (698) ذكور بنسبة (58%)، و(495) إناثا بنسبة (42%)، والرصافة الثانية (2016) بنسبة (46%)، بواقع (1179) ذكور بنسبة(59%)، و(837) إناثا بنسبة(41%)، الرصافة الثالثة (1142) بنسبة (26%)، بواقع (653) ذكور بنسبة(57%)، و(489) إناثا بنسبة(43%) .

ثانياً:عينة البحث:

من أجل تحقق من أهداف البحث يتطلب اختيار عينات عده منها:- العينة الاستطلاعية تمثل (٥٠) طالب وطالبة، عينة التحليل الإحصائية (600) طالب وطالبة.

ثالثاً : اداة البحث

استخدمت الباحثة اختبار القدرة العقلية لطلبة للمرحلة الاعدادية في المدارس المتفوقين والمتفوقات المعد من قبل المؤسسة الهندية (Aglasem Schools) للتكنولوجية التعليمية التي تعد اختبارات تحصيلية واختبارات القدرات تتناسب مع كافة الاعمار الدراسية في المدارس.

وصف الاختبار :-

يتكون اختبار من (٥٠) مفردة ذات (الاربعة بدائل) لكل مفردة إجابة صحيحة واحدة.

ثالثا :- إجراءات إعداد اختبار القدرة العقلية :-

بترجمة اختبار القدرة العقلية من لغة الام (الاجنبية) الى اللغة العربية عن طريق متخصص بالترجمة[١] ، ومن ثم أعيد ترجمة الاختبار بشكل عكسي عن طريق مترجم ثاني[٢]. ثم دقق عن طريق متخصص في اللغة العربية [١].

[١] أ. د ضياء مزهر كلية تربية ابن الرشد .

[٢] أ. د فاتن خليل الرفاعي كلية تربية ابن الرشد.

التحليل المنطقي لفقرات اختبار القدرة العقلية :

من أجل التأكد من صلاحية المفردات بشكلها الظاهري تم عرضها على عدد من الخبراء في مجال العلوم النفسية والتربوية لإبداء آرائهم. وكانت نسبة الاتفاق المعتمدة لقبول المفردات وهي (١٠٠%) ، وبهذا الاجراء تم التحقق من الصدق الظاهري لاختبار القدرة العقلية

تجربة وضوح الفقرات والتعليمات:

وللتأكد من وضوح التعليمات والمفردات الزمن اللازم للإجابة على الاداة ، طبق الاختبار على عينه وعددها (٥٠) طالباً وطالبة ، ومن خلال هذه التجربة تبين ان التعليمات وجميع المفردات واضحة ، وان متوسط الوقت الحقيقي اللازم للإجابة هو (٥٥) دقيقة.

التحليل الإحصائي للاختبار

وقد شمل(التحليل الإحصائي) إجراءات عده ومنها:(تحديد العينة, تطبيق الاداة، التحليل وفق انموذج راش باستخدام برنامج (jmetrik) والتحقق من فرضيات البحث من خلال برنامج(spss).

عينة التحليل الاحصائي :-

عينة التحليل الإحصائية مكونه (٦٠٠) طالب وطالبة من طلبة المرحلة الاعدادية للمدارس المتفوقين والمتفوقات وقد تم اختيارها بأسلوب (العشوائية الطبقية المتناسبة).والجدول (١) يوضح ذلك :

جدول (١) عينة البحث بحسب المديريات والجنس

مديرية	الجنس		العدد	النسبة المئوية
	ذ	أ		
مديرية تربية الرصافة الأولى	98	70	168	%28
مديرية تربية الرصافة الثانية	162	114	276	%46
مديرية تربية الرصافة الثالثة	89	67	156	%26
المجموع الكلي	349	251	600	%100

التحقق من افتراضات انموذج راش

الافتراض الاول:- التحقق من افتراض أحادية البعد:

هنالك عدة اساليب ومؤشرات للتأكد من افتراض الاحادية، وقد اعتمد البحث الحالي على (التحليل العاملي) :

- **مؤشرات تعتمد على التحليل ألعاملي:**

[١]أ . د سعد زاير كلية تربية ابن الرشد.

من اجل التحقق من افتراض أحادية البعد (اختبار القدره العقلية), بتحليل العينة المكونة من (٦٠٠) طالب وطالبة، قامت الباحثة بتحليل من خلال برنامج الاحصائي(SPSS) ، وقد تم التوصل الى عام واحد مسيطر ذي جذر كامن تقديره (٢٦.١٩٨) وتباين تفسيره (٥٢.٣٩٦) من التباين الكلي ، والجدول (٢) يوضح تشبع المفردات بالعامل العام.

الجدول (٢) تشبع فقرات الاختبار بالعامل العام

فقرة	تشبع	فقرة	تشبع	فقرة	تشبع	فقرة	تشبع	فقرة	تشبع
١	٠.٩٤٢	١١	٠.٤٧٩	٢١	٠.٨٥٥	٣١	٠.٧٠٤	٤١	٠,٧٣٨
٢	٠.٧٢٦	١٢	٠.٦٤٣	٢٢	٠.٩٥٤	٣٢	٠.٧٠٩	٤٢	٠,٦٣٠
٣	٠.٦٦٧	١٣	٠.٧٨٣	٢٣	٠.٧٥٣	٣٣	٠.٨٠٩	٤٣	٠,٦٧٣
٤	٠.٨٥٤	١٤	٠.٧٢٦	٢٤	٠.٦٠٨	٣٤	٠.٦٧٨	٤٤	٠,٨٦١
٥	٠.٧٢٨	١٥	٠.٦٤٠	٢٥	٠.٦٣٠	٣٥	٠.٧٨٨	٤٥	٠,٧٨١
٦	٠.٨٨٧	١٦	٠.٨٤٠	٢٦	٠.٥٣٢	٣٦	٠.٧٩٤	٤٦	٠,٧٣٦
٧	٠.٦٥٧	١٧	٠.٧٥٢	٢٧	٠.٧٨٤	٣٧	٠.٨٩٤	٤٧	٠,٦٩٩
٨	٠.٩٣٩	١٨	٠.٦٧٩	٢٨	٠.٥٨١	٣٨	٠.٥١٦	٤٨	٠,٧٨٦
٩	٠.٥١٦	١٩	٠.٧٢٤	٢٩	٠.٦٦٢	٣٩	٠.٦٥٨	٤٩	٠,٦٥٩
١٠	٠.٦٩٢	٢٠	٠.٥٩٠	٣٠	٠.٥٧٩	٤٠	٠.٤٤٥	٥٠	٠,٧٥٢.

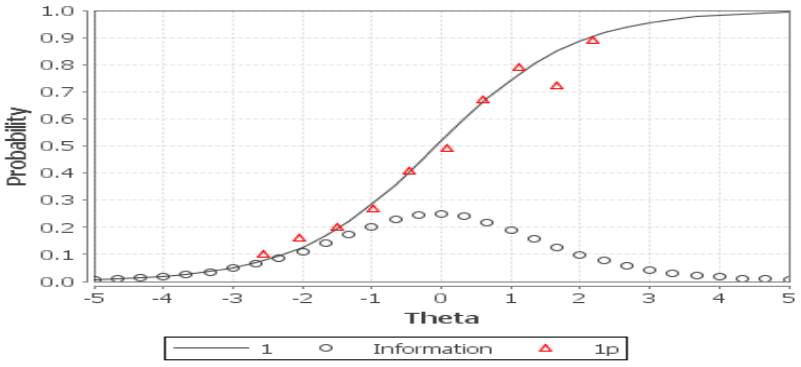
ويلاحظ من الجدول أعلاه أن جميع المفردات حصلت على تشبعات أكبر من (٠.٣٠) حسب معيار جيلفورد، ولم يتم استبعاد اي مفردة.

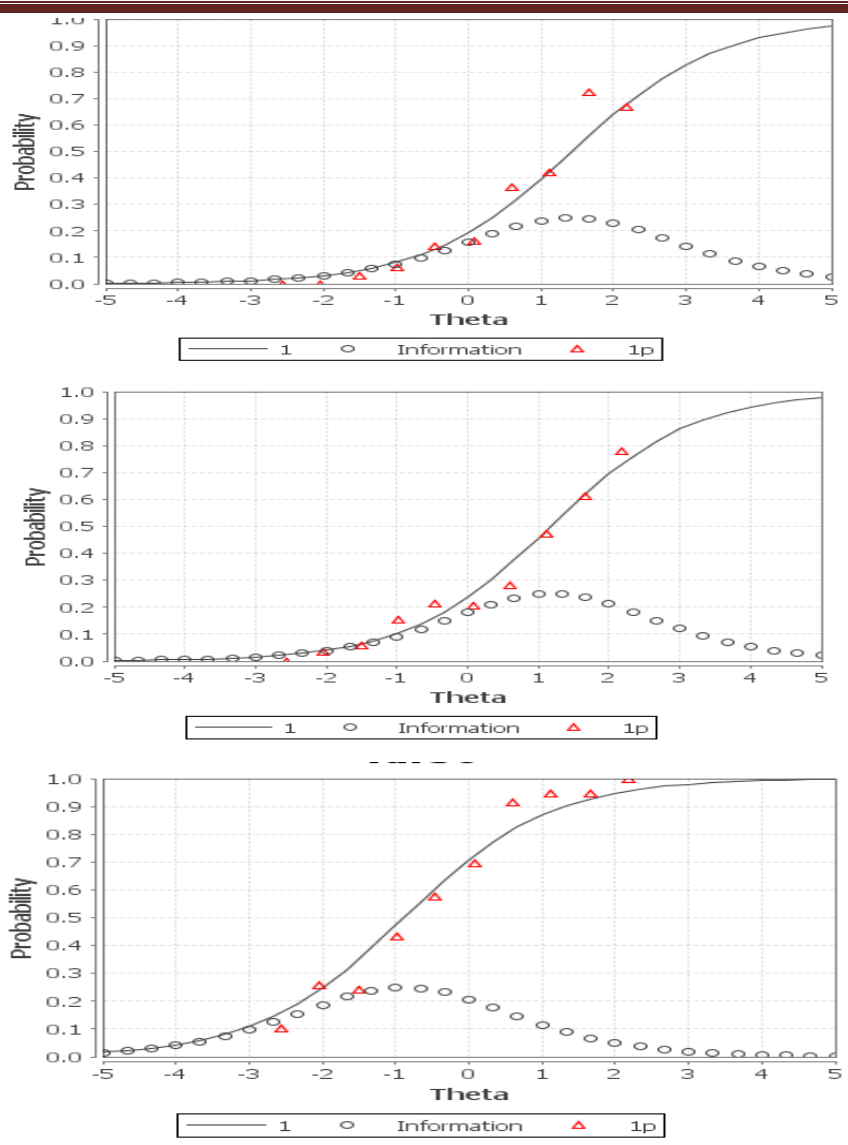
الافتراض الثاني:- التحقق من افتراض الاستقلال المحلي

ويتم التحقق من هذا الافتراض على النحو التالي: بما أن الافتراض الأحادي يعادل افتراض الاستقلال المحلي، ولكن ليسا متساويين مفاهيمياً، فاذا تحقق الأول ويتحقق الثاني.

الافتراض الثالث:- طبيعة المنحنى المميز للفقرة Nature of the Item Characteristic Curve:

المنحنى المميز للفقرة (ICC). ويتم تحديد ميل المنحنى وفق معلم الصعوبة (β), و قدرة الأفراد (θ). والشكل (١) يوضح ذلك





شكل (١) المنحنى المميز لبعض فقرات الاختبار

مطابقة الفقرات لانموذج راش باستخدام برنامج (jmetrik) :-

بعد تحليل البيانات يقوم البرنامج بعرض عدة مخرجات على النحو الأتي:

أولاً: المخرجات النصية (Text Output) تظهر قيمة مربع كاي ودرجة الحرية والاحتمالية ومعالم الفقرات وقدرة الافراد حسب أنموذج راش كما موضح في الجدول (٣) :

جدول (٣) احصائية مربع كاي للمطابقة لفقرات الاختبار

ت	مربع كاي	df	قيمة الاحتمالية P<	ت	مربع كاي	df	قيمة الاحتمالية P<
١	٢٩.٨٥٩٥	٤١.	٠.٩٠١١	26	51.8577	٤١.	0.0991
٢	٣٦.٦	٤١	٠.٦٦٦١	27	34.0663	٤١	0.7336

						٠٨٦	
0.5940	٤١	38.2385	28	٠.٥٩٥٦	٤١	٣٨.٢٠٣٦	٣
0.4523	٤١	41.4196	29	٠.٢٤٢١	٤١	٤٦.٩٣٨٥	٤
0.6475	٤٢	37.0339	30	٠.٨١٧٦	٤٢	٣٣.٦٤٤٥	٥
0.8038	٤١	33.1435	31	٠.٩٦٣٩	٤١	٢٦.٢٨٨١	٦
0.3821	٤١	43.0864	32	0.8642	٤١	٣٨.٠٠٩١	٧
0.1612	40	30.9828	33	0.8750	40	30.878 5	٨
0.9984	40	19.3745	34	0.9245	40	28.788 5	٩
0.8494	٤١.	30.8747	35	0.9245	٤٢	25.756	١٠
0.0004	٤١	8.3262	36	0.9605	٤١	43.023 1	١١
0.5204	٤١	39.8784	37	0.3431	٤١.	48.847 1	١٢
0.9359	٤١	28.1859	38	0.2167	٤١	47.775 7	١٣
0.7275	٤٢	34.2150	39	0.4429	٤١	41.638 3	١٤

١٥	29.109 7	٤١	0.8986	40	44.0935	٤١	0.3026
١٦	30.774 4	٤٢	0.8525	41	47.7044	٤١	0.3026
١٧	45.341 2	٤١	0.8986	42	45.7123	40	0.2827
١٨	49.207 3	٤١	0.1508	43	34.7334	40	0.7441
١٩	37.933 0	40	0.6077	44	32.2170	٤١	0.8046
٢٠	36.353 8	40	0.6277	45	38.3588	٤١	0.5887
٢١	35.687 3	٤١	0.7053	46	12.8336	٤٢	0.1019
٢٢	44.190 2	٤٢	0.3384	47	31.1690	٤١	0.8401
٢٣	5.0376	٤١	0.0098	48	39.9943	٤١	0.4705
٢٤	20.701	٤١	0.0020	49	12.7102	40	0.1040
٢٥	20.723 7	40	0.0623	50	31.6738	42	0.8520

قيمة مربع كاي الجدولية تساوي (٥٥.٧٦) عند مستوى دلالة (٠.٠٥)، ودرجة حرية (٤٠ ، ٤١ ، ٤٢)

ويتضح من الجدول اعلاه ان جميع مفردات الاداة جاءت ملائمة لانموذج راش، اذ ان قيم كاي (كا) انها لم تكن ذو دلاله احصائيا، حيث ان القيمة المحسوبة اصغر من قيمة كاي الجدولية (٥٥.٧٦) وعند درجة حرية (٤٠ ، ٤١ ، ٤٢).

الجدول (4) معلم الصعوبة والاخطاء المعيارية

ت	صعوبة	الخطا المعياري	ت	صعوبة	الخطا المعياري
١	١.٠٠	٠.١٠	26	−1.78	٠.١٠
٢	1.23	٠.١١	27	−0.91	٠.٠٩
٣	0.22	٠.٠٩	28	1.79	٠.١٠
٤	0.35	0.09	29	0.94	٠.١٠
٥	−0.40	0.09	30	−0.68	٠.٠٩
٦	1.42	0.10	31	−0.26	0.09
٧	0.41	0.09	32	0.98	0.10
٨	1.82	0.11	33	−0.54	٠.٠٩
٩	1.64	0.12	34	1.50	٠.١٠
١٠	1.55	0.11	35	−1.48	٠.١١
١١	0.51	0.09	36	0.76	٠.١٠
١٢	−0.30	0.09	37	1.90	٠.١٠
١٣	1.38	٠.١١	38	1.90	٠.١٠
١٤	0.88	٠.١٢	39	−1.64	٠.٠٩

١٥	−1.13	٠.١٠	40	−0.69	٠.١٠
١٦	0.83	٠.٠٩	41	0.98	٠.٠٩
١٧	−1.37	٠.١١	42	−0.62	٠.١٠
١٨	1.81	٠.١١	43	2.00	٠.٠٩
١٩	0.40	٠.٠٩	44	0.11	٠.١٠
٢٠	−0.63	٠.٠٩	45	1.55	٠.١٠
٢١	1.21	٠.١٠	46	0.93	٠.٠٩
٢٢	0.47	٠.٠٩	47	−0.72	٠.٠٩
٢٣	0.35	٠.٠٩	48	−0.61	٠.١٩
٢٤	−0.75	٠.١٣	49	0.54	٠.٠٩
٢٥	2.00	٠.١٠	50	−0.40	٠.٠٩
معلم التمييز = ٠.٩٣٧ لجميع الفقرات					

يتضح من جدول (٤) اعلاه ان جميع الفقرات ضمن المعيار المقبول لمعلم صعوبة الفقرة (٢.٥) – (٢.٥ −)

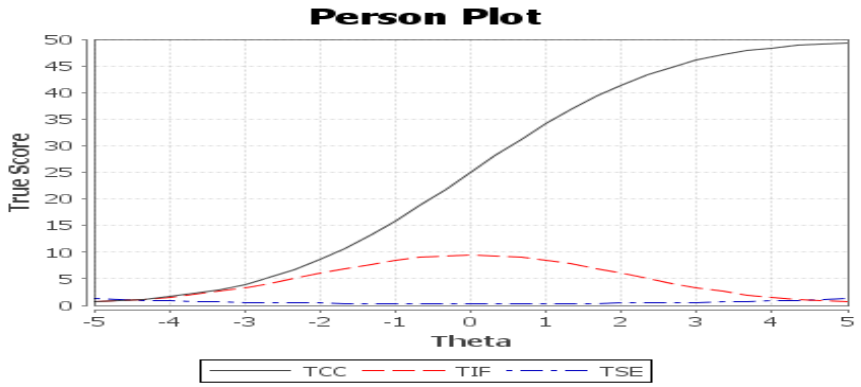
• ملف قدرة الأفراد Persons Parameters File : يعرض البرنامج في هذا الملف معالم تقديرات قدرة أفراد العينة البالغ عددهم (٦٠٠). والأخطاء المعيارية لهذه التقديرات والمتوسط الحسابي لتقديرات القدرة والاخطاء المعيارية ، وكما يقدمها البرنامج، كما موضح في جدول (٥).

الجدول (٥) يبين وصفاً مختصراً لمعالم العينة

عدد الافراد	القيم	اعلى قيمة	ادنى قيمة	الوسط الحسابي	الانحراف المعياري
٦٠٠	القدرة	2.789	−3.339	0.006	0.876
	الخطا المعياري للقدرة	0.555	0.300	٠.٣٣٢	٠.٠٥٠

يتبين من جدول (٥) اعلاه أن قيم القدرة تتراوح بين (2.789) – (3.339−) بوسط حسابي (0.006) وانحراف معياري (0.876).

ثانيا :- المخرجات البيانية.(Graph Output) :- تتضمن الرسوم البيانية لمنحنى المميز ودالة معلومات الفقرة الاختبار كما موضوح في شكل (1) اعلا، ومنحنى مميز ودالة المعلومات للاختبار ككل كما موضح في الشكل (2) التالي.



شكل (2) المنحنى المميز ودالة المعلومات لاختبار القدرة العقلية

ثبات الاختبار :

تم حساب الثبات وفق مؤشر دالة المعلومات، وذلك من خلال الرجوع للمخرجات البيانية للبرنامج (jmetrik) واعتماد على الرسم البياني لمنحنى معلومات الاختبار ككل كما موضح في الشكل (٢) ان أعلى ارتفاع منحنى المعلومات عند مستوى القدرة بقيمة (٩.٠) على محور (المعلومات) الأفقي، والخطا المعياري (٠.١١)، فان قيمة ثبات الاختبار (0.89) وهو يعد مناسب للاداة من حيث دقتها.

الوسائل الاحصائية :-

- البرنامج الاحصائي (spss).
- البرنامج الاحصائي (jmtrik).

الفصل الرابع

يتضمن هذا الفصل عرضاً لنتائج البحث الحالي, وتحليلها, فضلا عن ابرز الاستنتاجات والتوصيات والمقترحات التي توصل إليها البحث الحالي:-

- **التعرف على تأثير ترتيب الفقرات على دقة تقدير معالم الفقرات وقدرة الافراد تبعا لمستوى القدرة:**

للتحقيق من الهدف الاول للبحث الحالي, وبعد مطابقة اختبار القدرة العقلية مع أنموذج راش وتحقق من ملائمته كما موضح في (الفصل الثالث)، وبعد ذلك قامت الباحثة باطلاع على عدة دراسات سابقة لترتيب الفقرات، وقد تم اختيار مؤشرين احصائيا لاعداد نموذجين لترتيب فقرات

لاختبار القدرة العقلية ومنها: (النموذج المرتب تصاعديا حسب معلم الصعوبة, اضافة الى النموذج المرتب بشكل عشوائيا للمصمم الاصلي لاختبار في ترتيب الفقرات), وقد تم ترتيب الفقرات وفقاً لمؤشر صعوبة الفقرات بشكل تصاعديا, تم تحديدها وفقاً لمخرجات البرنامج الإحصائي (jmetrik). وذلك بترتيب الفقرات بدءً بالفقرة التي تتمتع بأقل قيمة لمعلم الصعوبة الى اعلى قيمة لها ، وبعد ذلك تم تطبيق النموذج (المرتب حسب معلم الصعوبة) على عينة التحليل الإحصائي ذاتها، أما بالنسبة (النموذج المرتب بشكل العشوائي) اعتمادا على نفس بيانات التحليل الاحصائي الاول ، وبعد ذلك تم تقسيم العينة لكل من نموذجين الى عينة مرتفعة ومنخفضة ومعتمده على قيمة الوسيط ومن خلال برنامج (اكسل)، وبعد ذلك تم تحليل البيانات المستمدة من هذين النموذجين لكل عينة على حدة, وباستعمال البرنامج الإحصائي (jmetrik) لاستخراج الاخطاء المعيارية لمعلم الصعوبة وقدرة الافراد ، وللتحقق من الهدف الاول قامت الباحثة بصياغة الفرضيتين الاتيتين:-

الفرضية الاولى : لايوجد فروق ذات دلالة احصائية في دقة تقدير معلم الصعوبة تبعا للمستوى القدرة (المرتفعة ، المنخفضة) للنموذجين:

ولغرض للتحقق من صحة هذا الفرض اذ تم حساب المتوسطات الحسابية والانحرافات المعيارية لقيم الخطأ المعياري لتقدير معلم الصعوبة باختلاف القدرتين (المرتفعة ، المنخفضة) للنموذجين، باستعمال الاختبار "ت" لعينتين مستقلتين للتعرف على الفروق بين المتوسطات الحسابية لقيم الخطأ المعياري والجدول (٦) يوضح ذلك.

الجدول (٦) اختبار (ت) لدلالة الفروق بين متوسطي الاخطاء المعيارية والانحراف المعياري لمعلم صعوبة الفقرات تبعا لمستوى القدرة

نوع النموذج	مستوى القدرة	عدد الفقرات	متوسط الحسابي لخطا المعياري لمعلم الصعوبة	الانحراف المعياري	المحسوبة	الجدولية	الدلالة عند مستوى (٠,٠٥)
ترتيب حسب الصعوبة	القدرة المرتفعة	٥٠	0.1432	0.0282	1.019	1.96	غير دالة
	القدرة المنخفضة	٥٠	0.1486	0.0248			
ترتيب بشكل عشوائي	القدرة المرتفعه	٥٠	0.1454	0.0269	4.482	1.96	دالة
	القدرة المنخفضة	٥٠	0.1802	0.0478			

يتبين من الجدول (٦) ان القيمة التائية المحسوبة الاولى بلغت (1.019) عند مستوى دلالة (٠,٠٥) وهي اصغر من القيمة الجدولية البالغة (١,٩٦) ، وهذا يعني عدم وجود فرق دال احصائيا بين المتوسطين وفق النموذج المرتب حسب الصعوبة ، ان القيمة التائية المحسوبة الثانية بلغت(٤.٤٨٢) عند مستوى دلالة (٠,٠٥) وهي اكبر من القيمة الجدولية البالغة (١,٩٦)، وهذا يعني وجود فرق دال احصائيا بين المتوسطين لصالح القدرة المنخفضة، وعليه

فان القدرة المرتفعة اكثر دقة في دقة تقدير معلم الصعوبة لانخفاض قيمة الخطأ المعياري في معلم الصعوبة لنموذج المرتب بشكل عشوائي.

الفرضية الثانية :- لايوجد فروق ذات دلالة احصائية في دقة تقدير القدرة الافراد تبعا لمستوى القدرة (المرتفعة ، المنخفضة) للنموذجين

ولغرض للتحقق من صحة هذا الفرض اذ تم حساب المتوسطات الحسابية والانحرافات المعيارية لقيم الخطأ المعياري لتقدير القدرة باختلاف القدرتين (المرتفعة ، المنخفضة) للنموذجين، باستعمال الاختبار "ت" لعينتين مستقلتين للتعرف على الفروق بين المتوسطات الحسابية لقيم الخطأ المعياري (٧) يوضح ذلك:-

الجدول (٧)

اختبار (ت) لدلالة الفروق بين متوسطي الاخطاء المعيارية والانحراف المعياري لقدرة الافراد تبعا لمستوى القدرة

نوع النموذج	مستوى القدرة	عدد الافراد	متوسط الحسابي لاخطاء المعياري لتقدير القدرة	الانحراف المعياري	المحسوبة	الجدولية	الدلالة الاحصائية عندى مستوى ٠.٠٥
ترتيب حسب الصعوبة	القدرة المرتفعة	٣٠٠	0.3376	0.0379	11.412	1.96	دالة
	القدرة المنخفضة	٣٠٠	0.3726	0.0371			
ترتيب بشكل عشوائيا	القدرة المرتفعة	٣٠٠	0.3357	0.0383	10.193	1.96	دالة
	القدرة المنخفضة	٣٠٠	0.3684	0.0401			

يتبين من الجدول (٧) ان القيمة التائية المحسوبة الاولى بلغت (11.412) عند مستوى دلالة (٠,٠٥) وهي اكبر من القيمة الجدولية البالغة (١,٩٦) ، وهذا يعني وجود فرق دال احصائيا بين المتوسطين لصالح القدرة المنخفضة، اما القيمة التائية المحسوبة الثانية بلغت(10.193) عند مستوى دلالة (٠,٠٥) وهي اكبر من القيمة الجدولية البالغة (١,٩٦) ، وهذا يعني وجود فرق دال احصائيا بين متوسطين لصالح القدرة المنخفضة وعليه فان القدرة المرتفعة وفق النموذجين اكثر دقة في دقة تقدير قدرة الافراد لانخفاض قيمة الاخطاء المعيارية في قدرة الافراد للنموذجين .

- **التعرف على تأثير ترتيب الفقرات على دقة تقدير معالم الفقرات وقدرة الافراد تبعا لعدد الفقرات**

للتحقق من الهدف الثاني قامت الباحثة بأخذ آراء الخبراء حول النسبة المئوية المناسبة لتحديد طول ثاني للاختبار باعتباره اختبار مختزل من الاختبار الاصلي وقد أشاروا الى نسبة (٣٥%) من فقرات الاختبار ، لذا تم سحب (١٨) فقرة من الاختبار الاصلي وبصورة قصدية من خلال فقرات ذات الملائمة الجيدة على وفق انموذج راش، بحيث يكون الخطأ المعياري بتقدير معامل

صعوبة الفقرة أقل ما يمكن في الاختبار ككل للفقرات الملائمة لأنموذج راش، وبعد ذلك تم تطبيق الاختبار القصير على نفس عينة التحليل الاحصائي لاستخراج مدى ملائمته مع انموذج راش، وقد تبين ان جميع فقرات الاختبار القصير يتمتع بخصائص قياسية جيدة, وبعد ذلك قامت الباحثة بإعداد اربعة نماذج لترتيب فقرات الاختبارين , متكون من نموذجين لكل اختبار المكون من (٥٠) فقرة و الاختبار القصير (المختزل) المكون من (١٨) فقرة ومنها النموذج المرتب حسب معلم الصعوبة وبشكل تصاعدي, وذلك بترتيب الفقرات بدءً بالفقرة التي تتمتع بأقل قيمة لمعلم الصعوبة ، اضافة الى النموذج المرتب بشكل عشوائي, والذي معد مسبقاً من قبل المصمم. وبعد تطبيق هذه النماذج على عينة التحليل الإحصائي ذاتها, قامت الباحثة بتحليل البيانات المستمدة من هذه النماذج, وباستعمال البرنامج الإحصائي (jmetrik). وذلك لتحديد اخطاءهم المعيارية لمعلم الصعوبة وقدرة الافراد وللتحقق من الهدف الثاني قامت الباحثة بصياغة الفرضيتين الاتيتين:

- **الفرضية الاولى** :لايوجد فروق ذات دلالة احصائية في دقة تقدير معلم الصعوبة تبعا لعدد الفقرات (١٨، ٥٠) للنموذجين .

ولغرض للتحقق من صحة هذا الفرض اذ تم حساب المتوسطات الحسابية والانحرافات المعيارية لقيم الخطأ المعياري لتقدير معلم الصعوبة باختلاف عدد الفقرات (١٨، ٥٠) للنموذجين، باستعمال الاختبار "ت" لعينتين مستقلتين للتعرف على الفروق بين المتوسطات الحسابية لقيم الخطأ المعياري والجدول (٨) يوضح ذلك.

الجدول (٨)

اختبار (ت) لدلالة الفروق بين متوسطي الخطا المعياري والانحراف المعياري لمعلم صعوبة الفقرات تبعا لعدد الفقرات

نوع النموذج	عدد الفقرات	متوسط الحسابي لخطا المعياري لمعلم الصعوبة	الانحراف المعياري	المحسوبة	الجدولية	الدلالة الاحصائية عندى مستوى ٠.٠٥
ترتيب حسب الصعوبة	١٨	0.1094	0.0262	1.045	1.96	غير دالة
	٥٠	0.1028	0.0219			
ترتيب بشكل عشوائي	١٨	0.1461	0.0545	5.392	1.96	دالة
	٥٠	٠.١٠٠٢	٠.٠١٦١			

يتبين من الجدول (٨) ان القيمة التائية المحسوبة الاولى بلغت (1.045) عند مستوى دلالة (٠,٠٥) وهي اصغر من القيمة الجدولية البالغة (١,٩٦) ، وهذا يعني عدم وجود فرق دال احصائيا بين المتوسطين، اما القيمة التائية المحسوبة الثانية بلغت(5.392) عند مستوى دلالة (٠,٠٥) وهي اكبر من القيمة الجدولية البالغة (١,٩٦) ، وهذا يعني وجود فرق دالة احصائيا بين المتوسطين لصالح الاختبار القصير، وعليه فان الاختبار الطويل اكثر دقة في دقة تقدير

معلم الصعوبة لانخفاض قيمة الاخطاء المعيارية في معلم الصعوبة وفق النموذج المرتب بشكل عشوائي.

●**الفرضية الثانية** : لايوجد فروق ذات دلالة احصائية في دقة تقدير معلمة القدرة تبعا لعدد الفقرات (١٨، ٥٠) للنموذجين :

ولغرض للتحقق من صحة هذا الفرض اذ تم حساب المتوسطات الحسابية والانحرافات المعيارية لقيم الخطأ المعياري لتقدير القدرة باختلاف عدد الفقرات (١٨، ٥٠) للنموذجين، باستعمال الاختبار "ت" لعينتين مستقلتين للتعرف على الفروق بين المتوسطات الحسابية لقيم الخطأ المعياري والجدول (٩) يوضح ذلك.

الجدول (٩)

اختبار (ت) لدلالة الفروق بين متوسطي الخطا المعياري والانحراف المعياري لقدرة الافراد تبعا لعدد الفقرات

نوع النموذج	عدد الفقرات	عدد الافراد	متوسط الحسابي لخطا المعياري لتقدير القدرة	الانحراف المعياري	المحسوبة	الجدولية	الدلالة الاحصائية عندى مستوى ٠.٠٠١
ترتيب حسب الصعوبة	١٨	٦٠٠	0.4221	0.0713	25.867	٣.٢٩١	دالة
	٥٠	٦٠٠	0.3304	0.0494			
ترتيب بشكل عشوائي	١٨	٦٠٠	0.4336	0.0643	30.382	٣.٢٩١	دالة
	٥٠	٦٠٠	0.3323	0.0503			

يتبين من الجدول (٩) ان القيمة التائية المحسوبة الاولى بلغت (25.867) عند مستوى دلالة (٠,٠٠١) وهي اكبر من القيمة الجدولية البالغة (٣.٢٩١) ، وهذا يعني وجود فرق دال احصائيا بين المتوسطين لصالح الاختبار القصير، واما القيمة التائية المحسوبة الثانية بلغت (30.382) عند مستوى دلالة (٠,٠٠١) وهي اكبر من القيمة الجدولية البالغة (٣.٢٩١)، وهذا يعني وجود فرق دال احصائيا بين متوسطين لصالح الاختبار القصير، وعليه فان الاختبار الطويل وفق النموذجين اكثر دقة في دقة تقدير قدرة الافراد لانخفاض قيم الاخطاء المعيارية لقدرة الافراد .

الاستنتاجات:

في ضوء وما توصل إليه البحث الحالي من نتائج, يمكن أن نستنتج ما يأتي:

١. فاعلية برنامج (jmetirk) الإحصائي في إجراء التحليلات الإحصائية لبيانات الاختبار.

٢. مناسبة اختبار القدرة العقلية لأفراد العينة (طلبة المرحلة الاعدادية) للمدارس المتفوقين والمتفوقات, وذلك لاتساق قدرات أفراد العينة وترتيب الفقرات.

٣. افضلية النموذج المرتب حسب معلم الصعوبة لانه يقلل من الاخطاء المعيارية لتقديرات معلم الصعوبة والقدرة.

٤. كلما ارتفعت قدرة الافراد نحصل على نتائج اكثر دقة وبالتالي يقلل من الاخطاء المعياري لتقدير معلم الصعوبة والقدرة وفق النموذجين.

٥. الاختبار الطويل افضل من الاختبار القصير من حيث دقة تقدير معلم الصعوبة والقدرة، وهذا يشير الى انه كلما زاد عدد الفقرات قلت قيمة الاخطاء المعيارية وبالتالي يزيد من الثبات وفق النموذجين.

التوصيات:

في ضوء نتائج يمكن للباحثة ان توصي بالآتي:

١. استخدام اختبار القدرة العقلية في مجالات عدة (الدراسات التربوية والنفسية وميدانية) كأداة لقياس (القدرة العقلية).

٢. استخدام البرنامج الإحصائي (jmetrik) في بناء وتطوير الاختبارات والمقاييس .

٣. عند بناء او اعداد فقرات الاختبار سواء كانت الاختبارات(تحصيليه او استعدادات او قدرات) يفضل الاعتماد على ترتيب الفقرات حسب معلم الصعوبة.

المقترحات:

في ضوء النتائج التي توصل اليها البحث الحالي, تقترح الباحثة إجراء الآتية:

١.إجراء دراسة تتناول تأثير ترتيب المفردات سواء (للمقياس أو للاختبار) على الخصائص القياسية أو المؤشرات الأخرى مثل: دالة المعلومات (IIF) الاختبار ، أو دالة استجابة المفردة (IRF)، أو المنحنى المميز للمفردة (ICC).

٢. إجراء نفس الدراسة بإضافة متغير آخر يتعلق بخصائص المفردة مثل (ترتيب أو عدد او نوع البدائل الاستجابة)، وذلك باستخدام أي تصميم عاملي، وحسب مستويات المتغيرين. .

٣.إجراء دراسات مشابه باستخدام النماذج البارامترية، وغير البارامترية، والاستجابة الثنائية، والمتعددة الاستجابات، بالإضافة إلى مقارنة النماذج.

المصادر :

• Al-Qarni, Ahmed Salem Saeed (2022). The effect of methods for estimating the ability of individuals on the accuracy of the test equation in the item response theory, Arab Journal of Educational and Psychological Sciences, Arab Foundation for Education, Science and Arts, Egypt, Volume (6), Issue (30), October, 1-22

• Abu Hajar, Amna Ibrahim Khalil (2011): Building a multi-level scale for thinking in a problem-solving method and evaluating psychometric properties according to paragraph response theory among basic stage

students, unpublished doctoral thesis, College of Arab Educational Studies, Amman University.

- Albano, A.D. (2013). Multilevel modeling of item position effects. Journal of Educational Measurement, 50(4), 408-426.
- Albano, D. A. (2012). Multilevel Modeling of Item Position Effects. Thesis of Doctoral. Faculty of the Graduate school of the University of Minnesota.
- Allam, Salah El-Din Mahmoud (2001). Diagnostic tests are a benchmark in the educational and psychological fields, Cairo, Dar Al-Fikr for Publishing and Distribution.
- Allam, Salah El-Din Mahmoud (2005). Response models for one-dimensional and multi-dimensional optional items and their applications in psychological and educational measurement, Cairo. Dar Al-Fikr Al-Arabi.
- Al-Taqi, Ahmed Muhammad (2013). Modern Theory of Measurement, 2nd edition, Amman, Dar Al-Masirah for Publishing and Distribution.
- Masuid, Muhammad Riyad (2013): The effect of the shape of the test items on matching the items to the non-parametric Mokin model and the two-parametric model, unpublished doctoral thesis, Yarmouk University, College of Education, Jordan.
- Atta, Zaid Bani, Hijazi, Taghreed (2022), The effect of arranging the items of a multiple choice test on the accuracy of estimating the parameters of the items and the ability of individuals, The Jordanian Journal of Educational Sciences, Volume 18, Number 1.
- Chen, Pei-Hua (2010). Item Order Effects on Attitude Measures. Theses of Doctoral, The Morgridge College of Education, University of Denver. ProQuest Dissertations & Theses Database.
- Chiu, P. C. (2012). The effect of item position on state mathematics assessment. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, Canada.

• Crocker, Linda, and Algina, James (2009). An Introduction to Traditional and Contemporary Measurement Theory, translated by Da'na, Zeinat Yousef, 1st edition, Dar Al-Fikr, publishers and distributors.

• Dambrot, F. (1980). Test item order and academic ability, or should you shuffle the test item deck? *Teaching of Psychology,* 7(2), 94–96 .

• Embretson, S. E. & Reise, S. P. (2000). Item response theory for psychologists. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

• Hambleton& Cook, L.L.(1977) :Latent Trait Models And Their Use In The Analysis of Educational Test Data ,Journal of Educational Measurement,Vol,14No .2,pp.75-95.

• Hambleton, R. K. &Jonse, R. W. (1993). Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development.Educational Measurement.

• Hartig, J. & Buchholz, J. (2012). A multilevel item response model for item position effects and individual persistence. *Psychological Test and Assessment Modeling*, Volume 54, 2012 (4), 418-431.

• Johnson, M. S., Sinharay, S., & Bradlow, E. T. (2007). Hierarchical Item Response Theory Models. Ch, 17, In Rao, C.R., & Sinharay, S. (Edt). Handbook of Statistics, 1st Ed- Psychometrics, Amsterdam; Boston: Elsevier North-Holland.

• Kim, S. (2001). An Evaluation of A Markov Chain Monte Carlo Method for the Rasch Model. Applied Psychological Measurement, Vol. 25, No. 2, 163-176.

• Leary, L. F., & Dorans, N. J. (1985). Implications for altering the context in which test items appear: A historical perspective on an immediate concern. Review of Educational Research, 55, 387–413.

- Majeed, Sawsan. (2007). Foundations for constructing psychological and educational tests and standards. First edition. Amman. Debono Publishing and Distribution House.
- Miller, Jeffrey M. (2004). Susceptibility to Survey Item Order Fffects as a Function of Perceived. Thesis of Master, The Graduate School of the University of Florida. UF Graduate School.
- Pommerich, M., & Harris, D. J. (2003). Context Effects in Pretesting: Impact on Item Statistics and Examinee Scores. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Sass, D. A., Schmitt, T. A., & Walker, C. M.(2004). An evaluation of BILOG-MG with skewed theta distribution using various estimation procedures : A simulation study. Poster presented at the National Council on Measurement in Education. San Diego, California.
- Stage, Christina. (1997). Applicability of Item Response Models to the SweSAT, (Education Measurement, No. 25), Umeå University, Department of Education Measurement.
- Uhan, S., & Hafner Fink, M. (2013). Context Effects in Social Surveys: Between Instrument and Respondent. TEORIJA IN PRAKSA let. 50, Vol. L, No. 1, pg. 233– 248.