مجلة ابحاث الذكاء
كلية التربية الاساسية – الجامعة المستنصرية
Journal of Intelligent Research

مستويات العمق المعرفي لمادة الاحصاء التربوي لطلبة الماجستير في اقسام العلوم التربوية والنفسية

هاجر سعد اسود أ.د. صفاء طارق حبيب
جامعة بغداد – كلية التربية – ابن الرشد – قسم العلوم التربوية والنفسية
safa.tairq@ircoedu.uobaghdad.edu.iq spyjojo3@gmail.com

مستخلص البحث:
يهدف البحث الحالي الى بناء اختبار تحصيلي عنقودي في (مادة الاحصاء التربوي) لطلبة الماجستير في اقسام العلوم التربوية والنفسية وفقا لنظرية القياس التقليدية فقد اتبعت الباحثة الخطوات العلمية لتحقيق الهدف الأول في بناء اختبار تحصيلي عنقودي في مادة (الاحصاء التربوي) ابتداءً من تحديد المحتوى الدراسي وصياغة الاهداف السلوكية وتكون الاختبار من (16 عنقود) بواقع (100 فقرة) نوع الاختيار من متعدد ذات (4 بدائل) لكل فقرة وتم تطبيق الاختبار على عينة مكونه من (400) طالب وطالبة تم اختيارهم بأسلوب الطبقي العشوائي من طلبة الجامعات العراقية لأقسام العلوم التربوية والنفسية, استعملت الباحثة نظرية القياس التقليدية لحساب الخصائص السيكومترية المتعلقة بفقرات الاختبار المتمثلة بـ(صعوبة الفقرة, تمييز الفقرة, صدق الفقرة وفعالية البدائل) وكذلك الخصائص السيكومترية المتعلقة بالاختبار وهي (الصدق والثبات) وقد توصلت الى ان جميع فقرات الاختبار وعناقيده متوسطات الصعوبة وان تمييزها جيد جداً وكذلك الاختبار يتميز بثبات عال جدا.

الكلمات المفتاحية: مستويات العمق المعرفي، الإحصاء التربوي.
(بحث مستل من رسالة ماجستير)

مشكلة البحث The problem of the Research:

تعمل مستويات العمق المعرفي على ان تجعل الطلبة يسألون عن الأشياء بـ (لماذا) وليس بـ (كيف) فقط ، وكما تجعلهم يصلون الى أقصى درجات الفهم ويسعوا لإرضاء فضولهم واهتماماتهم المعرفية في جميع الموضوعات والمواد الدراسية ، وكما تجعلهم يستفيدوا من الأدلة والبحث والتقويم وتكسبهم رؤية واسعة لربط أفكارهم ببعضها وسيصبحوا مدفوعين ذاتياً للتعلم وسيمكنهم العمق المعرفي من ربط المفاهيم والمهارات الجديدة بالمواقف والخبرات الحياتية وأخيراً ستجعلهم يميلوا للقراءة ودراسة ما هو أبعد من متطلبات المادة الدراسية (الفيل ،2018: 251). يتطلب تمكن الطلاب من مهارات القرن الحادي والعشرين تنمية مستويات عمق المعرفة من خلال الابداع وتوليد الافكار الجديدة ,وتقديم حلول ذكية للمشكلات وتنمية اساليب جديدة (مندور فرج الله ,2018:20)
ويختلف نموذج عمق المعرفة عن تصنيف بلوم Bloom في النطاق والتطبيق , حيث صنف Bloom المهارات المعرفية المطلوبة من العقل عند تناول مهمة جديدة, وبالتبعية فقد صنف نوع عمليات التفكير اللازمة للإجابة عن سؤال ما ,في حين يرتبط نموذج عمق المعرفة بعمق فهم المحتوى ,ونطاق نشاط التعلم الذي يتضح في فهم المهارات المطلوبة لإكمال المهمة من البداية للنهاية (مثل :التخطيط, البحث, استخلاص النتائج)(Lam & Foong,1996:56).

ان مستويات عمق المعرفة جاءت كردة فعل لبعض مشكلات محتوى الكتب الدراسية ومن ابرزها سطحية المعرفة وتفككها وضعف ترابطها ,حيث انها تفتقر الى أسس المعرفة التي تحقق عمق المادة العلمية وهو ما يؤثر على مخرجات العلمية التعليمية .

أهمية البحث The significance of the Research :

يُعَد القياس من الركائز الاساسية التي تساهم في تحقيق الاهداف العامة للمجتمع، فهو حجر الزاوية لأجراء اي عملية أو قرار بهدف جمع المعلومات لأغراض التغيير والتعديل والتحسين في عمليات وممارسات الانسان في الحياة اليومية، ويوفر قاعدة من المعلومات التي تلزم لمتخذي القرارات حول مدخلات وعمليات ومخرجات اي عملية، وبالتالي المساهمة في التخطيط وصياغة الاهداف واختيار الانشطة والاساليب الملائمة (محاسنة، 2013: 22).

اذ تكمن أهمية القياس النفسي تكمن في مدى قدرته على تشخيص الأفراد بدقة لتوجيههم إلى مهن وأعمال قد تتفق مع ما لديهم من قدرات ومهارات قد تتطلب بها تلك المهن التي تتوفر لدى شخص معين ولا يمتلكها غيره لأن بمثل هذا التشخيص الدقيق بواسطة الاختبارات والمقاييس سيبعد الآخرين أفراداً أو جماعات عن التعرض إلى مخاطر العمل وما سيلحقه بهم من ضرر الأمر الذي يتحتم فيه تكليفهم بأعمال تتلاءَم مع قابليتهم واستعداداتهم فضلاً عن أن القياس النفسي قد يلعب دوراً هاماً في عملية إناطة الأعمال وتوزيعها بين العاملين في المؤسسات التربوية والإدارية من خلال الأخذ بمبدأ وضع الشخص المناسب في العمل المناسب تبعاً للمواصفات التي تتطلبها بعض الأعمال أو المهام كالسرعة والدقة في العمل والتفاعل الجيد مع الآخرين عند الاحتكاك بهم (المياحي ، 2021 : 21).

وتعد صياغة فقرات الاختيار من متعدد عملاً فنياً وإبداعياً، وثمة من أعتبر ذلك فناً وعلماً في آن واحد، ولذا يتوجب توزيع قواعد صياغة الفقرات على معدي تلك الفقرات إذا كانوا من غير المتخصصين، وتكثيف البرامج التدريبية والتطبيقية المتعلقة بجودة صياغة فقرات الاختيار من متعدد وفق الإرشادات الخاصة بها من قبل خبراء بناء الأسئلة والمتمرسين عليها، ولا يخلو كتاب في القياس والتقويم دون من قبل متخصصي هذا الفن من إرشادات تتعلق بصياغة فقرات الاختبارات التحصيلية من نوع الاختيار من متعدد (Novick1986 :107&Iord)

وقد نالت حركة بناء الاختبارات اهتماما بالغا لدى المشتغلين في الميدان التربوي من أجل بناء اختبارات جيدة تقدم نتائج ومعلومات موضوعية وواقعية يمكن اعتمادها في صنع القرارات ورسم السياسة التربوية (سواقد ،1983: 127). فهناك مجموعة من القواعد والإرشادات في كتابة فقرات اختبار الاختيار من متعدد منها: التأكد من أن الجذر يطرح مشكلة محددة وواضحة، وجعل البدائل قصيرة ما أمكن، وتجنب صيغ النفي، والتأكد من أن بدائل الإجابة الخطأ تؤلف إجابات معقولة ظاهرياً، وأن تكون جذابة للمفحوصين الذين تنقصهم المعرفة، وأن لا يتضمن الاختبار فقرات تعتمد في إجابتها على فقرات أخرى، وتجنب الخداع والغموض في جذر الفقرة وبدائلها، وجعل بدائل الفقرة متساوية في طولها (باسل خميس أبو فودة ,٢٠١٢ :44). ولابد من الاشارة الى ان الاختبار العنقودي يعد من افضل الطرق المستخدمة في بناء الاختبار التحصيلي ,نظرا لقدرته على تجاوز عيوب ونقاط ضعف الاختبار التحصيلي التقليدي, فضلا عن تحقيقه لمستويات عالية من الصدق

والثبات وفقا للدراسات التي اجريت والتي وظفت الاختبار العنقودي في بناء اختبارات تحصيلية متنوعة في مواد علمية وانسانية (Wang,Bradlow and Wainer,2002:2).

ان اهم ما يميز الاختبار العنقودي هو قدرته على تحقيق ذلك التوزان ما بين المحتوى الدراسي المؤلف من قبل مجموعة من الخبراء والمتخصصين والاهداف التعليمية التي تسعى المؤسسة التعليمية الى الوصول اليها ,حيث يعتمد ذلك على قيام كل مجموعة من الفقرات(حزمة من الفقرات) التي تتعلق بمحتوى معين من المنهج الدراسي بقياس ذلك المحتوى بشكل دقيق ومحدد (Wainer & Stewar 1984:190).

اهداف البحث: Aim of the Research

يهدف البحث الحالي الى:

1.بناء اختبار تحصيلي عنقودي لمادة الإحصاء التربوي لطلبة الدراسات العليا (الماجستير)

2.مستوى طلبة الدراسات العليا (الماجستير) لمحتوى مادة الإحصاء التربوي بناءًا على الاختبار العنقودي المعد بالبحث الحالي

3.التعرف الفروق ذات الدلالة الإحصائية بين متوسط درجات طلبة الدراسات العليا (الماجستير) لمحتوى مادة الإحصاء التربوي بناءًا على الاختبار العنقودي المعد بالبحث الحالي تبعا لمتغيري الجنس والتخصص

حدود البحث: Limitation of the Research

يتحدد البحث الحالي بـ:

طلبة كليات التربيــة – مرحلة الدراسات العليا (ماجستير) – لكلا الجنسين(ذكور –اناث) قسم العلوم التربوية والنفسية – في الجامعات العراقية (بغداد, المستنصرية, الانبار, ديالى, تكريت, البصرة, كربلاء, الكــوفة, الديوانية, واســط, الموصــل, العراقيــة) للعام الدراسي (2021-2022).

تحديد المصطلحات : Definition of terms

اولاً:ـ الاختبار : **Test**

يعرفه كل من :

ـــ كرونباخ **Cronbach** (1984) :

"وسيلة لقياس المعارف والمهارات والمشاعر والذكاء والاستعداد لدى فرد أو جماعة من الافراد". (Cronbach, 1984 : 2)

ـــ انستازي **Anastasi** 1988)):

" طريقة موضوعية ومقننة لتقدير سمة ما من خلال قياس عينة من السلوك.

(Anastasi, 1988 :23)

ثانياً: الاختبار التحصيلي **Achievement test**

عرفه كل من:

-السيد,2000 : "مقياس المستويات المعرفية للطلبة وغيرهم من الافراد في كل ميدان من ميادين المواد الدراسية ,وفي كل مجال من مجالات المعرفة البشرية".(السيد, 2000:327).

- العبادي , 2006: "مجموعة من الفقرات (الاسئلة) التي تعطى للطالب ليجيب عنها شفويا ,او تحريريا, وقد تكون موضوعية او مقاليه ,او رسوما, آو اشكالا تستعمل للمقارنة والقياس". (العبادي,2006 :11)

ثالثاً: مستويات العمق المعرفي (Depth of Knowledge))((DOK:

عرفهُ كل من:

(Webb 1997):- "مستوى التعقيد العقلي الذي يتعلق بكل من المعلومات التي يتوقع ان يعرفها الطلاب وكيفية استفادتهم من تلك المعارف في سياقات مختلفة وكيفية وصولهم الى التعميمات بشكل جيد وكم المعارف التي يجب ان يمتلكوها لفهم الافكار"(Weeb, 1997:15).

(Newton2005): -

أنهُ" فحص متعمق بصورة نقدية للأفكار والمعارف الجديدة ووضعها في البناء القائم وربط الافكار ببعضها لحل مشكلة ما في الحياة الواقعية" (Newton, 2005:43)

الفصل الثاني

اطار نظري ودراسات سابقة

اولا : الاطار النظري

الاختبارات العنقـودية Test lets:

في عام 1987 اقترح wainer and kiely اسم حزمة من فقرات الاختبار التي تدار معًا والتي أطلقوا على الاختبار العنقودي. كانت الاختبارات العنقودية موجودة لفترة طويلة قبل عام 1987 ، وإن كان ذلك بدون هذه التسمية المبهجة. تم استخدامها عادةً لتعزيز كفاءة الاختبار في المواقف التي فحصت قدرة الفرد على فهم نوع من التحفيز ، على سبيل المثال ، مقطع قراءة أو رسم بياني للمعلومات أو مقطع موسيقي أو جدول أرقام. في مثل هذه الحالات يتم إنفاق قدر كبير من وقت الممتحن في معالجة الحافز او المثير وقد وجد أنه يضيع هذا الجهد في طرح سؤال واحد فقط حوله. وبالتالي فقد كانت المحفزات الكبيرة تقترن عادةً بمجموعة من الأسئلة. لقد ساعدت الخبرة في توجيه عدد الأسئلة التي تم استخدامها لتشكيل الاختبار العنقودي, وهكذا تم استخدام الاختبارات العنقودية ببساطة كطريقة لبناء الاختبارات التي تحتوي على محفزات كبيرة بطريقة فعالة.

(Howard Wainer, 2007: 44)

- مفهوم الاختبار العنقودي

ليس هنالك اتجاه واضح في تحديد تعريف محدد لمفهوم الاختبار العنقودي :

فالاتجاه الاول : يرى ان الاختبار العنقودي هو عبارة اختبار يتكون من مجموعة من الفقرات, وكل فقرة يتم التعامل معها على انها وحدة مستقلة بذاتها يتم تصحيحها والتعامل معها كأنها اختبار لوحده

الاتجاه الثاني : يرى ان الاختبار العنقودي ليس بفقرة واحدة مستقلة انما هو عبارة عن مجموعة من (العناقيد) وهذه العناقيد تمثل وحدات او مجالات مستقلة من الاختبار يتم تصحيح هذه الوحدة بشكل مستقلة عن بقية العناقيد (Wang & Wilson,126 :2005) .

- **اهمية الاختبار العنقودي**

ان الاختبار العنقودي مقارنة بالاختبارات الاخرى التقليدية يتسم بمجموعة من السمات التي تميزه عن بقية الاختبارات, ويمكن تلخيص تلك السمات بالاتي:

1.يحقق الاختبار العنقودي التوازن المطلوب ما بين المحتوى الدراسي المؤلف وبين الاهداف التعليمية التي تسعى المؤسسة التعليمية الى الوصول اليها .

2.امكانية تحديد درجة المتعلم الكلية على الاختبار العنقودي ,فضلا عن درجة في كل عنقود او حزمة من حزم الاختبار .

- انواع الاختبارات العنقودية

هنالك نوعين اساسيين من الاختبارات العنقودية يمكن تلخيصها بالاتي:

اولا: الاختبار العنقودي الخطي **Linear Testlets**:

في هذا النوع من الاختبارات يتم تطبيق جميع الحزم او العناقيد على جميع الممتحنين, بمعنى ان هذه الاختبارات مكونة من عناقيد محددة وثابته ,ويتم عرضها بشكل ثابت على المتعلمين, ويتضمن هذا النوع الشكلين التاليين:

1.الاختبارات الورقة وقلم (الورقية) Paper&Pencil tests

2.الاختبارات الالكترونية CBT) Computer-based tests)

ثانيا: الاختبار العنقودي الهرمي Hierarchical Testlets **:**

هي شكل من اشكال اختبارات الموائمة ,يتم من خلالها انتقاء عدد معين من العناقيد او الحزم وفقا للمستوى الذي وصل اليه المتعلم ,فهي تمثل اختبارات مفصلة وفقا لمستوى المتعلم ,وبالتالي يمكن اختيار مجموعة من العناقيد لممتحن ومجموعة اخرى لممتحن اخر وهكذا لبقية الممتحنين (1991:3,Wainer& Kaplan and Lewis).

- **محور مستويات العمق المعرفي**

مقدمة في العمق المعرفي:

يشهد العصر الحالي تغيرات كبيرة في كافة الميادين وابرزها التحول نحو المعرفة، حيث أصبحت المعرفة العامل الأساسي في عملية التنمية وازدادت أهميتها يوماً بعد يوم؛ فأصبحت تمثل ثروة لكافة المجتمعات؛ تمثل المعرفة العميقة المعرفة الغنية بالمعرفة والحقائق والمعلومات, حيث أنها البنية الأساسية لفهم المحتوى المقدم للمتلقي والقدرة على ادراك للأفكار الكامنة واستخراجها والتعامل بالشكل المناسب مع المعلومات وتوظيفها والاستفادة منها، وتعد الأساس لصنع المعرفة الحية والتعلم الفعال, حيث تصل بالمتلقي إلى حالة التأمل المعرفي, ايضاً تساهم في تطوير قدرة التعلم على ممارسة التعلم وبناء معرفته لمواجهة تحديات العالم، وكذلك مساعدته على استرجاع المعلومات التي تم تخزنيها في بنيته العقليه وربطها بالمعلومات الجديدة للوصول إلى معنى اكثر شمولاً، كما تحفزه إلى السعي نحو المزيد من البحث والمعرفة ولا يقتصر فهم المحتوى على الوعي بالمعلومات، بل يمتد إلى الذهاب خلفها وبذلك يتكون الدافع للمعرفة، وعندما يكون هناك دافع لدى الفرد فأنه يسعى لاكتساب اهمية المعلومات والحقائق الهامة, فالفرد الذي يملك دافع للمعرفة ويبحث عنها بقناعة ذاتية يصبح التعلم لديه ذا منفعة ويمكن الاستفادة منه.(دوابه,2020,2733).

في ضوء تلك الاهتمامات، أنشأ Webb أداة تقييم قائمة على المعايير تعتمد بشكل أساسي على تنسيق المعايير والمحتوى والتقييم. من خلال هذه الأداة يتم تصنيف المعرفة العلمية حسب مستويات المعرفة (عمق المعرفة) في ضوء درجة تعقيد التفكير المطلوب لإنجاز المهام العلمية. ويشمل هذا التصنيف أربعة مستويات لعمق المعرفة العلمية: "استدعاء المعرفة (التذكر) ، وتطبيق المفاهيم والمهارات ، والتفكير الاستراتيجي والتفكير الممتد (الموسع)". (عمر، ٢٠١٧، ١٠٣)

- تصنيف نورمان ويب Norman Webb :

هو التصنيف المعتمد في هذه الدراسة وستستعرضه الباحثة بالتفصيل :

مفهوم مستويات العمق المعرفي :

طور في عام 1997 العالم (نورمان لوت ويب) وهو عالم في مركز ويسكونسن لبحوث التعليم ، نموذج العمق المعرفي, للتنسيق بين معايير المناهج وعملية التقييم. يعتمد هذا النموذج على افتراض أن جميع عناصر المنهج يمكن تصنيفها في ضوء المتطلبات المعرفية. وتجدر الإشارة إلى أن مصطلح المعرفة المستخدم هنا يشمل جميع أشكال المعرفة على نطاق واسع (أي: إجرائي ، التصريحية ، إلخ.) (Mississipi Department of Education, 2009: 5)

يعد نموذج العمق المعرفي مدخلاً لتنظيم المعرفة. يختلف تعقيد هذا النموذج وفقًا للصف وما يجب أن يعرفه الطلاب. يعتمد التعقيد في العمق المعرفي أيضًا على البنية المعرفية السابقة للطلاب وقدرتهم على إجراء التعميمات. تؤكد مستويات عمق المعرفة على ما يعرفه الطلاب ويشير ويب إلى أن هناك علاقة بين ما يتم تدريسه وما يتم اختباره ، وشدد على الحاجة إلى المواءمة بين الأهداف ومعايير التقييم عند قياس العمق المعرفي ، ويعتمد ذلك على المستوى التفكير الذي يسمح للطلاب بعمل تعميمات بناء المعرفة. (Viator, 2010: 23)

- أنواع المعرفة من حيث العمق:

صنف كلاً من (Alex Bennet & David Bennet,2008), أنواع المعرفة من حيث عمقها الى ثلاثة أنواع رئيسة وهي:

1- المعرفة السطحية Surface knowledge**:**

تتمثل في المعلومات الأساسية التي يمتلكها الفرد حول موضوع ما ، ويتم استخدام الحد الأدنى من مستوى الفهم والذاكرة قصيرة المدى ، والتي من خلالها يجيب الفرد على الأسئلة: (ماذا ، من ، متى ، أين). يتضمن هذا المستوى من المعرفة الحقائق والبيانات والمعلومات والمفاهيم والتعامل معها على مستوى الحفظ والتذكر لاستخدامها من أجل النجاح في الامتحان. يصعب تذكر هذا النوع من المعرفة ويسهل نسيانه ؛ لأنها ذات معنى بسيط وروابط قليلة فيما بينها. (Bennet.A & Bennet.D,2008,4)

2-المعرفة الضحلة Shallow knowledge**:**

عند هذا المستوى يمتلك الفرد مقياسًا للمعرفة وفهمًا جزئيًا لهذه المعرفة ، وبعض المعاني لها من خلال السياق ، ويحدث هذا من خلال ربط المعرفة الجديدة بالخبرات السابقة الموجودة في البنية المعرفية للفرد. . قد يتكون هذا المعنى أيضًا من المنطق والتحليل والملاحظة والتأمل والتنبؤ أحيانًا . (Bennet.A & Bennet.D,2008,4)

3-المعرفة العميقة Deep knowledge**:**

عند هذا المستوى من المعرفة ، يكون الفرد قادرًا على فهم المعرفة ومعناها السياقي بطريقة دقيقة ومتكاملة ، وكذلك القدرة على تغيير إطاره المرجعي مع تغيير السياق. الشخص الخبير الذي يمتلك معرفة عميقة حول موضوع ما هو الشخص الذي يمكنه استنتاج الأنماط وتقييم أهميتها في الموقف واستخدام النظريات التي توجه معرفتهم وإبداعهم وابتكارهم. كلما اكتسب الإنسان معرفة عميقة، تعمق فهمه وزاد إبداعه ، تعمقت معرفته ، وهكذا. يتطلب اكتساب المعرفة العميقة وتطويرها تحفيز الطالب على دراستها وتعلمها ، والتركيز على تفاصيل السياق لكل تجربة يمر بها، وطرح الأسئلة، وتحليل الأحداث ، ودمج كل هذا في هيكل المعرفة الخاص به الذي يتعمق بمرور الوقت، استخدامه لفهم العالم من حوله وتفسير ظواهره وحل مشاكله.

(2008,5,Bennet.A & Bennet.D)

مستويات العمق المعرفي: Depth Of Knowledge Levels

وفيما يأتي توضيح لهذه المستويات الأربعة:

المستوى الاول: التذكر وإعادة الإنتاج: **Recall and Reproduction**

يتضمن هذا المستوى المهام الأساسية التي تطلب من المتعلمين تذكر أو إعادة إنتاج المعرفة أو المهارات ، وهذا المستوى يتضمن عادة الطلاب الذين يتعاملون مع حقائق وخصائص الأشياء. قد يشمل أيضًا استخدام إجراءات وصيغ بسيطة. فيما يتعلق بأسئلة هذا المستوى ، فإن الطالب الذي يجيب على سؤال إما يعرف الإجابة أم لا ؛ أي أن الإجابة لا تحتاج إلى أن تكون "احسب" أو "تحل". (7Mississippi State University, 2009, p)

المستوى الثاني: تطبيق المفاهيم والمهارات: **Basic Application of Skills/Concepts**

يصف هذا المستوى أداء بعض عمليات وأنشطة المعالجة العقلية التي تتجاوز مستوى التذكر ، وتتطلب القدرة على إبراز الاختلافات أو المقارنات والتصنيف والفرز والوصف وتوضيح العلاقات العمق المعرفي في تطبيق المفاهيم والمهارات اعتمادًا على نوع من التفكير المرن والاختيار بين البدائل والحلول مقارنة بالمستوى الأول ، وتتطلب العناصر المدرجة في هذه الفئة تكامل بعض العمليات الذهنية وليس مجرد الاستجابات العادية ، حيث يتطلب الطلاب تقرر ما يجب القيام به باستخدام أساليب غير مألوفة للعمليات المنطقية والاستراتيجيات. يتم تحديد دور المعلم في العرض والملاحظة والتنظيم والتسهيل والتقييم.(شاهين,423,2020)

المستوى الثالث: التفكير الاستراتيجي: Strategic Thinking:

يتطلب هذا المستوى استخدامًا قصير المدى لعمليات التفكير العليا، مثل التحليل والتقييم، لحل مشاكل العالم الحقيقي بنتائج يمكن التنبؤ بها. يعتبر ذكر المنطق علامة رئيسية للمهام التي تقع في هذا المستوى. تميل التوقعات المحددة للمهام على هذا المستوى أيضًا إلى تنسيق المعرفة والمهارة من مجالات متنوعة ومتعددة لتنفيذ العمليات والتوصل إلى حلول. تشمل العمليات الرئيسية التي تشير غالبًا إلى هذا المستوى: التحليل والتفسير والدعم بالأدلة والتعميم والإبداع. في هذا المستوى من التعقيد ، يجب على الطلاب استخدام التخطيط والأدلة والتفكير بشكل أكثر تجريدًا. تحتوي المهمة

على إجابات صحيحة متعددة ، حيث يجب على الطلاب تبرير اختياراتهم ، مثل حل المشكلات غير الروتينية ، أو تصميم تجربة ، أو تحليل خصائص الأنواع.. (2014,11,Gerald Aungst)

المستوى الرابع: التفكير الممتد: Extended Thinking

يتميز بمتطلبات معرفية عالية ومعقدة ، وعادة ما يتطلب بناء روابط منطقية بين الأفكار، وربطها ببقية الموضوعات الأخرى من منظور متعدد التخصصات ، واختيار أو اقتراح استراتيجية مناسبة لتوظيفها في حل المشكلة العلمية المطلوبة. من بين العديد من الخيارات المتاحة ، وتجاهل الاختبارات التقليدية في العلوم باستخدام أي أنشطة تقييم ضمن هذا المستوى تتعلق بصقل وتطوير مهارات التفكير العليا الأكثر تقدمًا بين الطلاب في القرن الحادي والعشرين ، ويتطلب هذا المستوى فترات طويلة نسبيًا من الوقت سواء لإجراء تجارب علمية أو لتحقيق أهداف معينة أو لتنفيذ مجموعة من الخطوات المتعددة لحل أسئلة تقييم معينة. (حسين,2019,18).

ثانياً – دراســات ســابقة Anterior Studies

- **دراسة محمد 2010:**

هدفت دراسة (محمد,2010) على دراسة اختلاف في تحقيق كل من افتراض (احادية البعد ,الاستقلال الموضعي, منحني خصائص المفردة وقيم معاملات ثبات الافراد والمفردات) لاختبار عنقودي مرجعي المحك باختلاف (انموذج راش, وانموذج التقدير الجزئي)اذ تبين من النتائج تحقق افتراضات نظرية الاستجابة للفقرة (احادية البعد _الاستقلال الموضعي_ توازن منحنيات خصائص الفقرة) باستخدام نموذج التقدير الجزئي مقارنةً بنموذج راش, والدقة في قياس دالة معلومات العناقيد ودالة الاختبار اضافتاً ارتفاع معدل الثبات لكل من الافراد والعناقيد باستخدام الانموذج التقدير الجزئي عنهُ عند استخدام انموذج راش.

- **ولسون و ادمس Wilson& Adams 1995**

توصلت نتائج دراسة ولسون و ادمس Wilson& Adams 1995 الى امكانية الوصول الى حسن مطابقة ممتازة للبيانات المستمدة من الاختبارات العنقودية لفروض نماذج الاستجابة للفقرة عند تحليل العناقيد باستخدام نموذج التقدير الجزئي ,إذ هدفت الدراسة الى استخدام الانموذج الاسمي في تحليل الاختبارات العنقودية حيث تم المقارنة بين اربع عناقيد تحتوي على (22) فقرة من خلال استخدام طريقتين في تحليل كل عنقود احدها اعتمدت على تحليل كل عنقود على انه وحدة مستقلة والاخرى اعتمدت على تحليل كل فقرة على انها وحدة مستقلة.

الفصل الثالث

اجراءات البحث

اولاً: منهجية البحث Method Of Research:

ونظراً لتعدد المناهج في إجراء البحوث في العلوم الإنسانية فأن طبيعة الدراسة وموضوعها والهدف منها هو الذي يحدد نوع البيانات المراد جمعها وطبيعة المنهج المستخدم في إجراء البحث، وبما ان البحث الحالي يهدف الى بناء اختبار تحصيلي عنقودي في مادة الاحصاء التربوي لدى طلبة الدراسات العليا (الماجستير)، فإن المنهج الملائم لموضوع البحث الحالي هو المنهج الوصفي المقارن للأسباب اي (الدراسات السببية المقارنة) وهو نوع من البحوث الوصفية التي تحاول إيجاد الحلول للمشاكل من خلال التحليل المترابط فيما بينها.

ثانياً : مجتمع البحث population of Research:

يتكون مجتمع البحث الحالي من جميع طلبة الدراسات العليا (الماجستير) في كليات التربية بالجامعات العراقية الذين درسوا مادة الاحصاء التربوي في(المرحلة التحضيرية) للعام الدراسي (2021-2022م) والبالغ عددهم (724) طالباً وطالبة، ومن كلا الجنسين بواقع (332) طالباً من الذكور وبنسبة (45.86%) ,و (392)طالبة من الاناث وبنسبة (54.14%), ومن التخصصات في أقسام العلوم التربوية والنفسية (المجموعة النفسية بواقع (23) طالباً وطالبة بنسبة (3.18%), و(244) طالباً وطالبة للمجموعة التربوية وبنسبة (33.70%), و(457) طالباً وطالبة لمجموعة الطرائق التدريس وبنسبة (63.12%).

ثالثاً: عينة البحث Sample of Research :

ولأجل الوصول الى نتائج دقيقة يمكن الوثوق بها , يتطلب ذلك عدداً من العينات وسيتم عرض عينات البحث بشيء من التفصيل كل عينة بحسب الاجراء المتبع لتحقيق أهداف البحث الحالي.

رابعاً : أداة البحث Instrument of the Research:

ولتحقيق هدف البحث الحالي تم بناء اختبار تحصيلي عنقودي في مادة الاحصاء التربوي يتكون من 100 فقرة موزعة على 16 عنقود وفقا للخطوات المتبعة والمعتمدة لبناء هذا النوع من الاختبارات.

التحليل الاحصائي لفقرات الاختبار التحصيلي:

تكونت عينة التحليل الاحصائي من (400) طالباً وطالبة من طلبة الدراسات العليا(ماجستير) في مختلف الجامعات العراقية وهي تمثل نسبة (55.2%) من مجتمع البحث الحالي والذي تم اختيارهم بأسلوب العينة الطبقية العشوائية ، فقد اختير أعداد من الطلبة من كل جامعة بصورة عشوائية وبما يتناسب مع حجمهم في مجتمع البحث حيث اختارت الباحثة عينة طبقية عشوائية من (13) جامعة من جامعات العراقية والمتمثلة بـ (بغداد ، المستنصرية، الانبار، ديالى، الكوفة، بابل، البصرة، كربلاء، تكريت، الموصل، الديوانية، واسط ، العراقية) وأختارت الباحثة عينة البحث بشكل عشوائي تناسبي تبعاً لمتغيرات (الجامعة، الجنس ، التخصص).

التحليل الاحصائي للفقرات:

طبق الاختبار التحصيلي قياس التحصيل في مادة الاحصاء التربوي لطلاب الدراسات العليا / الماجستير على عينة مكونة من (400) طالباً وطالبة أُختيروا بالأسلوب الطبقي العشوائي ،وبعد تطبيق الاختبار على أفراد عينة التحليل الإحصائي البالغ عددهم (400) طالباً وطالبة ، إذ تم تصحيح إجابات عينة التحليل الإحصائي باستخدام مفتاح تصحيح مثقب (الملحق) على أساس إعطاء الدرجة (1) للإجابة الصحيحة عن كل فقرة اختبارية والدرجة (صفر) للإجابة الخاطئة، وبذلك تكون درجة الطالب الكلية في الاختبار مساوية لعدد الفقرات التي أجاب عنها بصورة صحيحة ، وقد تراوحت درجات أفراد عينة التحليل الإحصائي على الاختبار بأكمله ما بين (37 - 79) درجة بمتوسط حسابي قدره (63.915) وانحراف معياري (7.454) درجة ، والجدول التالي يوضح المؤشرات الإحصائية لعينة تحليل الفقرات .

الجدول المؤشرات الإحصائية لعينة تحليل الفقرات

عدد فراد العينة	الوسط الحسابي	الوسيط	المنوال	الانحراف المعياري	الالتواء	التفلطح	اقل قيمة	أعلى قيمة
400	**63.915**	**54.000**	**54.124**	**7.454**	**- 0.083**	**- 0.544**	**37**	**79**

وفيما يأتي توضيح للخصائص الإحصائية للفقرات:

اولاً : صعوبة الفقرة : Item Difficulty

ولحساب معامل صعوبة الفقرات الى الأختبار التحصيلي العنقودي في مادة الاحصاء التربوي لدى طلبة الدراسات العليا (الماجستير) بعد تطبيقها على عينة البحث اعتمدت الباحثة على المجموعتين المتطرفتين العليا والدنيا، حيث تم حساب معاملات الصعوبة لكل فقرة من فقرات الأختبارات من خلال تطبيق معادلة الصعوبة والتي تشير الى تحديد عدد الافراد ذوي الاجابات الصحيحة في المجموعة العليا وكذلك جمع عدد افرادها مع افراد ذوي الاجابات الصحيحة في المجموعة الدنيا ، ومن ثم تقسم على عدد افراد المجموعتين العليا والدنيا وبنسبة (27%). وأن فقرات الأختبار العنقوي لمادة الأحصاء التربوي لطلبة الدراسات العليا / الماجستير المكون من (100) فقرة لم تتضمن أي فقرة واحدة صعبة جداً أو فقرة سهلة جدا ، وتراوحت مديات صعوبة فقرات لأختبارات فيليب كارتر للرشاقة العقلية من (0.200 – 0.790) وبمتوسط صعوبة قدره (0.348) وهذا يقع ضمن مستوى الفقرات متوسطة الصعوبة . كما حسبت الباحثة معاملات صعوبة عناقيد الأختبار التحصيلي الستة عشر لمادة الإحصاء التربوي لطلبة الدراسات العليا (الماجستير) ، وذلك بحساب عدد الطلبة الذين أجابوا عن العنقود وضربها بوزن الدرجات المحددة له وللمجموعتين العليا والدنيا ، وبعد حساب حاصل الضرب تجمع للمجموعتين وتقسم على عدد أفراد المجموعتين مضروبا بأعلى وزن للدرجة الخاصة بالعنقود ، فكانت نتيجة حساب معاملات الصعوبة لعناقيد الأختبار التحصيلي الستة عشر.

ثانيا :- القوة التمييزية للفقرات Discrimination power of Items **:**

رتبت الباحثة درجات العينة البالغ عددها (400) طالب وطالبة من أعلى درجة إلى أقل درجة. وتم تحديد المجموعتان المتطرفتان في الدرجة الكلية من خلال اعتماد نسبة (27%) في كل مجموعة، بحيث اصبحت كل مجموعة من المجموعتان العليا والدنيا تتضمن (216) طالبا وطالبة. وتم حساب التمييز من خلال تطبيق معادلة التمييز لحساب معامل التمييز لكل فقرة من فقرات الأختبار العنقودي لمادة الإحصاء التربوي لطلبة الدراسات العليا (الماجستير).

ويتضح من خلال النتائج أن فقرات الأختبار العنقودي لمادة الإحصاء التربوي لطلبة الدراسات العليا / الماجستير المكون من (100) فقرة لم تتضمن أي فقرة ذات تمييز ضعيف أو فقرة ذات تمييز حدي ، وتراوحت مديات صعوبة فقرات الأختبار العنقودي لمادة الإحصاء التربوي من (0.311 – 0.633) وبمتوسط صعوبة قدره (0.420) وهذا يقع ضمن مستوى الفقرات تمييز جيد وتمييز جيد جدا .

كما حسبت الباحثة معاملات تمييز عناقيد الأختبار التحصيلي الستة عشر لمادة الإحصاء التربوي لطلبة الدراسات العليا / الماجستير، وذلك بحساب عدد الطلبة الذين أجابوا عن العنقود وضربها بوزن الدرجات المحددة له وللمجموعتين العليا والدنيا، وبعد حساب حاصل الضرب تطرح للمجموعتين وتقسم على عدد أفراد أحدى المجموعتين مضروبا بأعلى وزن للدرجة الخاصة بالعنقود.

فعالية البدائل الخاطئة Effectiveness of Distractors**:**

المفروض ان تكون البدائل فعالة بما فيه الكفاية لكي يخطئ البعض بها وليس الكل ، فلا فائدة من بديل خاطئ يخطئ به الكل أو يعرفه الكل . وينبغي ان يكون عدد الاختيارات الخاطئة أكثر لدى المجموعة الدنيا من المجموعة العليا. يتبين من النتائج أن قيم معاملات فعالية البدائل الخاطئة لجميع فقرات الأختبار العنقودي لمادة الإحصاء التربوي لطلبة الدراسات العليا / الماجستير كانت سالبة ، وهذا يدل على أن جميع البدائل فعالة فهي بدائل قد جذبت أفراد المجموعة الدنيا أكثر مما جذبت أفراد المجموعة العليا ، وهذا ما ينبغي أن تكون عليه البدائل الخاطئة للحكم على أنها فعالة .

الخصائص السيكومترية للأختبار التحصيلي العنقودي لمادة الإحصاء التربوي :

صدق الاختبار (Validity of test)**:**

يمثل الصدق العملية التي يجمع مطور الاختبار أو مستخدمه من خلالها الأدلة التي تدعم الاستدلالات التي استخلصها من درجات الاختبار (كروكر والجينا ، 2009 : 291).

تحققت الباحثة صدق الأختبار التحصيلي العنقودي لمادة الإحصاء التربوي على النحو الأتي:

أ- الصدق الظاهري (Face Validity) **:**

قامت الباحثة بعرض اختبار التحصيلي العنقودي لمادة الاحصاء التربوي على مجموعة من المحكمين الخبراء، حيث طلبت الباحثة من لجنة الخبراء ابداء حكمهم على مدى صلاحية كل فقرة من فقرات الاختبارات من حيث التعديل أو الحذف ، واعتمدت الباحثة على نسبة (100%) الأتفاق التام محكاً لقبول الفقرة من عدمه وقد حازت جميع الفقرات على نسب عالية مقارنة بنسبة المحك (٨٣%)، ولم تحذف اي فقرة من فقرات الأختبار التحصيلي العنقودي لمادة الإحصاء التربوي.

- صدق البناء(التكوين) (Construct Validity):

وتحقق هذا الأفتراض بأستخراج معاملات الصعوبة ومعاملات التمييز ومعاملات فعالية البدائل الخاطئة لفقرات للأختبار التحصيلي العنقودي لمادة الإحصاء التربوي لطلبة الدراسات العليا (الماجستير) حيث تعد مؤشرات جيدة على صدق البناء.

ثبات الاختبارات (Reliability Tests) :

حسبت الباحثة الثبات للأختبار التحصيلي العنقودي لمادة الإحصاء التربوي لطلبة الدراسات العليا (الماجستير) بعد اختيار عينة بصورة عشوائية تتألف من (100) طالب وطالبة تم اختيارهم من عينة البحث . وتم اعتماد طريقة تحليل التباين بأستعمال معادلة هويت Hoyt .

- معادلة هويت **Hoyt**:

واستعملت الباحثة تحليل التباين الثنائي (بدون تفاعل) لدرجات عينة الإفراد ، ومن ثم أستعملت معادلة هويت (Hoyt) من نتائج تحليل التباين والتي يمكن ان تعتمد على تباين الخطأ التباين بين الإفراد. وبهذا فقد بلغت قيمة معامل ثبات اختبار التحصيلي العنقودي (0.857) ، التي تم حسابها بطريقة تحليل التباين، كانت كلها معاملات جيدة عند مقارنتها بالدراسات السابقة التي تناولت الثبات.

عرض نتائج البحث وتفسيرها ومناقشتها :

●لتحقيق الهدف الأول الذي ينص على (بناء أختبار تحصيلي عنقودي لمادة الإحصاء التربوي لطلبة الدراسات العليا / الماجستير، ولتحقيق الهدف الأول قامت الباحثة بجميع الإجراءات والخطوات العملية اللازمة لبناء الأختبار، بدءً من تحديد المحتوى الدراسي وتحديد الأهداف السلوكية، ومن ثم إجراءات التحليلات المنطقية والإحصائية للتحقق من الخصائص السيكومترية للفقرات والأختبار الكلي، وقد تبين ان جميع الفقرات والأختبار الكلي يتمتع بخصائص قياسية جيدة ، وتم الاستبقاء على جميع الفقرات، وبذلك تم تحقيق هذا الهدف.

لتحقيق الهدف الثاني الذي ينص على قياس الخصائص السيكومترية المتمثلة بـ(صعوبة الفقرات ,تمييز الفقرات, صدق الفقرات , الصدق ,الثبات) للاختبار العنقودي وحصلت على التالي:

1.جميع عناقيد الاختبار على معاملات صعوبة تراوحت من (0.312 – 0.745) وهي معاملات متوسطة الصعـوبة.

2.حصلت جميع عناقيد الاختبار على معاملات تمييز تراوحت بين (0.310- 0.612) وهي معاملات تمييزها جيد جداً.

3.كانت جميع قيم فعالية البدائل الخاطئة بالسالب وتراوحت بين (-0.008 الى -0.124) وهي بدائل فعالة.

4.جميع فقرات الاختبار ترتبط ارتباطاً دالاً احصائياً مع الدرجة الكلية له اذ كانت قيم معامل ارتباط الفقرة بالدرجة الكلية اكبر من القيمة الحرجة لمعامل الارتباط وذلك يعني انها تتسق فيما بينها في قياس قدرة واحدة .

5.تبين أن جميع عناقيد الاختبار ترتبط ارتباطا دالاً احصائياً مع الدرجة الكلية للاختبار إذ كانت قيم معامل الارتباط أكبر من القيمة الحرجة لمعامل الارتباط ويشير ذلك على أن هناك اتساق داخلي بين العناقيد جميعها .

التوصيـــات Recommendations **:**

في ضوء النتائج التي توصلت إليها الدراسة الحالية ، توصي الباحثة بما يأتي:

1. على الباحثين في مجال العلوم التربوية والنفسية عموما والقياس والتقويم خصوصا الاستمرار في استخدام الأختبارات العنقودية في بناء الاختبارات التحصيلية، وذلك لما يتميز به من الدقة في القياس فضلاً عن الخصائص النظرية الأخرى التي يتميز بها.

2. على المؤسسات التعليمية (وزارة التربية ووزارة التعليم العالي) توافر البرامج الإحصائية الحديثة التي تستخدم في تحليل بيانات الاختبارات النفسية والتربوية التي تعتمد على الأسلوب المعاصر في القياس وتدريب ممارسي الاختبارات عليها مثل (المدرسين وطلبة الدراسات العليا).

المقترحات Suggestions

من خلال ما أسفرت عنه الدراسة الحالية من نتائج ، تقترح الباحثة بعض الدراسات والبحوث الآتية:

1. إجراء دراسة مقارنة بين بناء الاختبار العنقودي والاختبار التقليدي في بناء الاختبار التحصيلي وذلك لمعرفة أي النماذج أكثر دقة في بناء الاختبار التحصيلي .

2. دراسة أثر طول الاختبار وحجم العينة في الخصائص السيكومترية للاختبار التحصيلي الذي بُني بالبحث الحالي.

المصادر العربية:

- محاسنة، ابراهيم محمد .(2013).القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة . ط1 . دار جرير للنشر والتوزيع عمان – الأردن .
- سواقد، ساري سليم سلامة.(1983). المقارنة بين أثر ثلاث طرق لتصحيح اختبارات الاختيار من متعدد على الخصائص السيكومترية للاختبار وعلى أداء المفحوصين من مستويات تحصيل ودرجات مخاطرة مختلفة على الاختبار، رسالة ماجستير غير منشورة ، كلية التربية ، جامعة اليرموك ، الاردن.
- الفيل ، حلمي محمد حلمي. (2018). برنامج مقترح لتوظيف التعلم القائم على السيناريو في التدريس والتأثير في تنمية مستويات عمق المعرفة وخفض التجول العقلي لدى طلاب كلية التربية النوعية جامعة الإسكندرية، مجلة كلية التربية ،جامعة المنوفية :القاهرة – العدد2(من 1 الى6).
- الساعدي ، حسن حيال ، والمياحي، مقداد سطر جراد . (2021) . المنهج التكاملي (مفهومه-نظرياته -طرائق تدريسه – تحليله – طرق بنائه) (ط 1) .بغداد :اليمامة للطباعة والنشر .
- ابو فود ، باسل خميس وبني يونس، نجاتي احمد (2012) . الاختبارات التحصيلية، ط 1، عمان، الاردن، دار المسيرة لمنشر والتوزيع والطباعة.
- أبن منظور، 630(711 ه) لسان العرب ، ط 2 ، دار إحياء التراث العربي، مؤسسة التاريخ العربي ، بيروت لبنان(1992).

- الامام , مصطفى محمود واخرون(2005).التقويم والقياس. بغداد-العراق, دار الحكمة
- السيد(2000). الذكاء. ط6، القاهرة، دار الفكر العربي.
- دوابه, ايمان عز الدين محمد (2020): دور القنوات الثقافية باليوتيوب في تنمية المعرفة العميقة والدافع المعرفي لدى الطلاب, العدد (55), ج (5), كلية التربية النوعية, جامعة المنصورة.
- عمر, عاصم محمد إبراهيم (2017): اثر تدريس العلوم باستخدام وحدات التعلم الرقمي في تنمية مستويات عمق المعرفة العلمية والثقة بالقدرة على تعلم العلوم لدى طلاب الصف الثاني المتوسط, المجلة التربوية, جامعة الكويت, مجلد(32), عدد(4).
- شاهين ،عبدالرحمن بن يوسف (2020): مدى توفر مستويات العمق المعرفي في كتب الاحياء للمرحلة الثانوية نظام المقررات في المملكة العربية السعودية ، المجلة العلمية بكلية التربية :جامعة أسيوط.
- حسين, اشرف عبدالمنعم محمد (2019): اثر تدريس العلوم باستخدام مدخل حل المشكلات مفتوحة النهاية على التحصيل وتنمية عمق المعرفة العلمية لدى طلاب الصف الأول المتوسط, المجلة المصرية للتربية العلمية, العدد(7), المجلد(22).
- كروكر، ليندا والجينا ، جيمز (2009) : مدخل إلى نظرية القياس التقليدية والمعاصرة ، ترجمة: دعنا زينات يوسف، الطبعة الاولى ، دار الفكر ناشرون وموزعون.

- ترجمة المصادر العربية الى الاجنبية:

- Mahasneh, Ibrahim Mohamed (2013). Psychological measurement in light of traditional theory and modern theory. I 1. Jarir House for Publishing and Distribution Amman – Jordan

- Sawaqed, Sari Salim Salameh. (1983). Comparison between the effect of three methods of correcting multiple choice tests on the psychometric characteristics of the test and the performance of the respondents of different levels of achievement and degrees of risk on the test, unpublished master's thesis, Faculty of Education, Yarmouk University, Jordan.

-Al-fil, Helmy Mohamed Helmy. (2018). A proposed program to employ scenario-based learning in teaching and influence the development of levels of depth of knowledge and reduce mental wandering among students of the Faculty of Specific Education, Alexandria University, Journal of the Faculty of Education, Menoufia University: Cairo - Issue 2 (from 1 to6)

- Al-Saadi, Hassan Hayal, and Al-Mayahi, Miqdad, Jarad line. (2021). The integrative curriculum (its concept - theories - teaching methods - analysis construction methods) (1st edition). Baghdad: Al-Yamamah for printing and publishing.
- Ibn Manzoor, 630 (711 AH) Lisan al-Arab, 2nd Edition, Dar Revival of Arab Heritage, Arab History Foundation, Beirut, Lebanon (1992).
- Imam, Mustafa Mahmoud and others (2005). Evaluation and - measurement. Baghdad-Iraq, House of Wisdom
- Dawabah, Iman Ezz El-Din Mohamed (2020): The Role of Cultural Channels on YouTube in Developing Deep Knowledge and Cognitive Motivation among Students, Issue (55), Part (5), Faculty of Specific Education, Mansoura University.
- Omar, Asim Muhammad Ibrahim (2017): The effect of teaching science using digital learning units in developing the levels of depth of scientific knowledge and confidence in the ability to learn science among second grade intermediate students, Educational Journal, Kuwait University, Vol. (32), No. (4).
- Shaheen, Abdul Rahman Bin Youssef (2020): The extent to which levels of knowledge depth are available in biology books for the secondary stage, the course system in the Kingdom of Saudi Arabia, Scientific Journal of the Faculty of Education: Assiut University.
- Hussein, Ashraf Abdel Moneim Mohamed (2019): The effect of teaching science using an open-ended problem-solving approach on achievement and developing the depth of scientific knowledge among intermediate first-grade students, The Egyptian Journal of Scientific Education, Issue (7), Volume (22).
- Crocker, Linda and Algina, James (2009): An Introduction to the Traditional and Contemporary Measurement Theory, translated by: Da'na Zeenat Youssef, first edition, Dar Al-Fikr Publishers and Distributors

- **المصادر الاجنبية**

*Lam, R., & Foong, Y. (1996): **Rasch analysis of math SOLO taxonomy levels using hierarchical items in testlets**., available online : www.eric.ed.gov, ED 398 271.

Iord, F . M & Novick, M. B. (1986) . **statistical theories of menta.***

*Wang ,Bradlow and Wainer ,H.(2002): **A general Bayesian for Testlets : Theory and Applications,** Educational Testing Service , Princeton, available online : www.ets.com ,NJ 08541.

*Weiner,E.A.&Stewar,B.J.(1984): **Assessing Individuals and Educational test and measurement**, Now York, Litlle Brown & co.

*Wainer, H., & Kiely, G. (1987): **Item clusters and computerized adaptive testing: A case for testlets.** Journal of Educational Measurement, 24, 185-201.

*Ferrara, S., Huynh, H. and Baghi, H. (1997): **Contextual characteristics of locally dependent open-ended item clusters in a large-scale assessment**. Applied Measurement in Education, 10, 123-144.

*Wang & Wilson,(2005): **The Rasch Testlet Model**, Applied Psychological Measurement , 29,126-149.

* Mississippi State University , **Researchers group** .(2009) . Webb's Depth of Knowledge Guide Career and Technical Education Definitions.

*Viator , E . (2010) **. A Critical Analysis of the Implementation of Depth of Knowledge and Preliminary Findings Regarding Its Effectiveness in Language Arts Achievement,** PhD Dissertation, University of Southern Mississippi.

* Bennet,A & Bennet, D.(2008). **The Depth of Knowledge** :Surface ,Shallow ,ordeep ,United States The Mountain Quest Institute.

* Gerald Aungst. (2014). **Using Webb's Depth of Knowledge to Increase Rigor https**://www.edutopia.org/blog/webbs-depth-knowledge increase-rigor-gerald-aungst.

مجلة ابحاث الذكاء
كلية التربية الاساسية – الجامعة المستنصرية
Journal of Intelligent Research

Levels of cognitive depth of educational statistics for master's students in educational and psychological sciences departments

Hajir Saad Aswad spyjojo3@gmail.com
Professor Safa Tariq Habeeb(Ph.D.) safa.tairq@ircoedu.uobaghdad.edu.iq

Abstract: The current research aims to:

Constructing a cluster achievement test in (educational statistics) for master's students in educational and psychological sciences departments according to the traditional measurement theory The researcher followed the scientific steps to achieve the first goal in building a cluster achievement test in the subject (educational statistics), starting from defining the academic content and formulating behavioral goals. The test was applied to a sample of (400) male and female students They were chosen in a random stratified manner from Iraqi university students for the departments of educational and psychological sciences. The researcher used the traditional measurement theory to calculate the psychometric characteristics related to the test items represented by (the difficulty of the items , the distinction of the items , the validity of the items and the effectiveness of alternatives), as well as the psychometric properties related to the test, which are (Validity and reliability). I concluded that all the items and clusters of the test are of medium difficulty, and that their discrimination is very good, as well as the test is characterized by very high stability.

Keywords: levels of cognitive depth, educational statistics
Research for an M.A. student.