تقييم الأسئلة الامتحانية النهائية لقسم العلوم التربوية والنفسية للعام الدراسي 2018-2019

ا.د عبد الرزاق محسن سعود **ا .م .د وليد قحطان محمود**

الجامعة العراقية / كلية التربية / قسم العلوم التربوية و النفسية

المستخلص:

استهدف البحث الحالي تقييم الأسئلة الامتحانية في قسم العلوم التربوية والنفسية للدورين الأول والثاني وللعام الدراسي 2019/2018 . وتحقيقا لأهداف البحث قام الباحثان باستعمال المنهج الوصفي/ منهج تحليل المحتوى/ وحدة الفكرة والتكرار من خلال استعمال بطاق تحليل الأسئلة التي شملت خمس معايير: الصدق، والشمولية، والأهداف التدريسية، صياغة الأسئلة، ومعايير إخراجها. بوصفها أداة للبحث بعد التحقق من صدقها وثباتها على (60) ورقة أسئلة تمثل مجتمع البحث . وقد توصل الباحثان إلى النتائج الاتية:

1. النتائج المتعلقة بالصدق: حيث كانت جميع الأسئلة صادقة عدا نسبة الأسئلة المقالية مقارنة بالموضوعية اذا كانت المقالية أكثر من الموضوعية.
2. النتائج المتعلقة بالشمولية: جاءت الأسئلة منخفضة في هذا المعيار وغير محققة له.
3. النتائج المتعلقة بالأهداف التدريسية: وفيه تم تحقيق الاسئلة لمستويات ثلاثة وفق نمط بلوم، بينما لم تغطي المستويات الثلاثة العليا.
4. النتائج وفق صياغة الأسئلة: واشارت النتائج التحقيق هذا المعيار بوضوح.
5. النتائج وفق الاخراج الجيد للاسئلة: وجاءت النتائج محققة لهذا المعيار بصورة جيدة.

وقد خرج الباحثان بمجموعة من الاستنتاجات والتوصيات والمقترحات.

Assessment the final exam questions for the Department of Educational and Psychological Sciences for the academic year 2018-2019

Abstract:

The present research aimed to evaluate the exam questions in the Department of Educational and Psychological Sciences for the first and second rounds for the academic year 2018/2019. To achieve the objectives of the research, the researchers used the descriptive method / content analysis methodology / unit of idea and repetition through the use of questions analysis card which included five criteria validity, comprehensiveness, teaching objectives, formulation of questions, and criteria for taking them out. As a research tool after verifying its validity and consistency on (60) question papers representing the research community, The researchers reached the following results:

1- Results relating to validity : where all the questions were valid except the proportion of artical questions compared to objectivity if the article was more than objective.
2- Results relating to on inclusiveness: The questions in this standard are low and unfulfilled.
3- Results relating to the teaching objectives: In which the questions were achieved for the three levels according to the Bloom pattern , while not covering the top three levels.
4- Results according to the formulating of the questions : The results indicated that this criterion was clearly achieved.
5- The results according to the good reflecting of the questions: The results indicated that this criterion was achieved well. The researchers came up with a set of conclusions, recommendations and proposals.

مشكلة البحث :-

تعد الاسئلة الامتحانية بأنواعها المقالية والموضوعية والعملية من ضرورات العملية التعليمية ، وهي جزء اساسي منها ، وبدون هذه الاسئلة لا يمكن ان نسير في العملية التربوية برمتها ولا نستطيع ان نقومها وان نجعلها اكثر نجاحاً .

وتعتبر الاسئلة الامتحانية ايضاً بوصفها محكات حقيقية للتعرف والتأكد من قدرات التدريسيين القياسية ، وهي بمثابة اختبار حقيقي يقيس قبولهم ونجاحهم وتفوقهم في مهنة التدريس .

ان الاسئلة الامتحانية السيئة التي لم تصمم على نحو سليم ، فأنها يمكن ان تكون اداه لتثبيط الهمم وتوجيه التعلم توجيهاً خطأً ، واذا تكرر هذا الامر في حياة الطالب فقد تصرفه عن الدراسة وتجعل خبرة الامتحان امراً محفوفاً بالإحباط والكراهية للمدرسة والتعليم عموماً (أبو علام ، 2005 : 132) .بالإضافة الى ان الاسئلة الضعيفة تعيق التفكير وتعطله (سند ، 1985 : 3) .

هذا ما جعل الاسئلة الامتحانية عرضه لكثير من الانتقادات نتيجة قصورها في نواحي عديدة سواء من حيث الاعداد أو اتصافها بالغموض احياناً أو عدم قدرتها على التمييز في احيان اخرى ، والاستعمال العشوائي لأنماط واشكال الاسئلة فيها ، وعدم ملائمتها ومناسبتها للأهداف المنشودة مثلما ذكر في العديد من الدراسات وتقارير المؤتمرات واللجان (كحلوت ، 2000 : 2) .

إذ اظهرت دراسة مارسو وبيج (Marso & Pigge , 2002) التي لخصت نتائج مسح (225) دراسة في مواضيع مختلفة وضعها التدريسيون في مراحل التعليم العالي ، وقد اظهرت هذه الدراسة نتائج عديدة مهمة كان من بينها تركيز الاختبارات على مستوى التذكر من الاهداف المعرفية فقط (Marso & Pigge , 2002 : 9-21) .

وتأسيساً على ما تقدم: يمكن ان تبرز اشكالية حقيقية عندما يضع المدرس اسئلته بطريقة أقل دقة مبتعداً عن المعايير العلمية الواجب توفرها في السؤال الامتحاني . لذلك

جاء البحث الحالي محاولاً الاجابة عن التساؤل الآتي : هل ان الاسئلة الامتحانية في قسم العلوم التربوية والنفسية للعام الدراسي 2019/2018 محققة للشروط القياسية والموضوعية بطريقة علمية ؟

أهمية البحث :-

من اجل ان تتمكن الانظمة التربوية من تحقيق اهدافها لابد من الاستعانة بالتقويم بكل ادواته ، والذي يعتبر المدخل لتطوير المنظومة التعليمية كافة ولا نستطيع ان نصل الى التقويم الا من خلال مرورنا بعمليتي القياس والتقييم ، ومن خلال هذه الاجراءات نستطيع ان نقيم الاهداف التعليمية المنشودة في المحتوى المقرر ، وكل ما يتعلق بالعملية التعليمية.

ومن وسائل التقويم في المواد التعليمية الامتحانات والتي تعد الاكثر انتشاراً اذ تشير الى قياس نتائج تعلم الطالب ، ومقدار ما استوعبه واتقنه من معلومات بعد مدة معينة من الدراسة، وانها وسيلة لنقل الطالب من صف الى آخر وتقدير تحصيله من العملية التعليمية ومدى استحقاقه شهادة المرحلة التي يتجاوزها . وللأسئلة الامتحانية اهميتها وقدرتها على اكتشاف مدى ما تحقق من الاهداف التربوية وبيان اسباب القصور في ذلك ومن ثم قدرتها على توضيح مدى نجاح العملية التعليمية في تحقيق غاياتها، كذلك في تقويم نمو الطالب في جوانب عديدة ابرزها الجانب المعرفي وتقويم طرائق التدريس والمناهج الدراسية ولا يمكن القول بوجود مؤسسة تعليمية يمكن ان تدار دون اجراء الامتحانات بأي نوع من انواعها وذلك لان التغيرات في التعليم التي تسعى التربية الى احداثها لا يمكن تلمسها أو رؤيتها أو اظهارها الا من خلال وسيلة والاسئلة الامتحانية هي الوسيلة لتحقيق هذا الغرض (عيسوي ، 1985 : 397) .

ان كل من عمل في مجال التعليم يعلم تماماً اثر الاسئلة الامتحانية في الطلبة فهي على اقل تقدير تدفعهم الى الاستذكار والاستعداد للامتحان ، اي ان لها وظيفة دافعية (أبو علام ، 2005 : 132) .

ومن العوامل المهمة التي تعين الطلبة على ممارسة التفكير النشيط هو نوع الاسئلة التي تقدم لهم وان تؤدي الاسئلة الجيدة دوراً في تحفيز التفكير واستمراريته وانتاجيته (سند، 1985 : 3) .

وصياغة الاسئلة الجيدة يعد من الوسائل المهمة في انجاح العملية التعليمية وعلى نوعية هذه الاسئلة يتوقف هذا النجاح ، والمدرس الجيد هو الذي يجيد صياغة الاسئلة واثارتها ، لان الاسئلة ماهي الا مثيرات تتضمن عمليات عقلية تلقي اهتماماً لما تؤديه من وظائف مهمة في العملية التعليمية (التل ، 1988 : 19) .

وكلما قام المعلم في رسم الاهداف التعليمية من تذكر وفهم وتطبيق وتحليل وتركيب وتقويم ، من شأنه ان يساعد الطالب في التعليم بشكل افضل مما لو اقتصر على التذكر والحفظ فقط وهذه المعرفة لمخرجات عملية التعلم تمكن المعلم من توجيه جهوده وتركيز انتباهه على مكونات هذا الاداء ليكون قادر على انجازه ، ومعرفة المعلم بهدف التعلم تعمل معززاً وكلما تقدم الطالب في سنوات دراسته اكتسبت الاسئلة الامتحانية اهمية اكبر، وكثيراً ما يصبح الاختبار الوسيلة الرئيسية ، بل الوحيدة لتحديد ما اكتسبه الطالب اذ يؤيد الاختيار التحصيلي هذا الدور البارز في جميع مراحل التعليم، بدءاً من المرحلة الابتدائية الى مرحلة التعليم العالي (ابو علام ، 2005 : 131) .

ويشعره بالتحصيل والانجاز ، وتصبح بذلك احدى المدعمات العقلية وتعطيه تغذية راجعه عن ادائه (بركات وصباح ، 2007 : 7) .

وفق ما تقدم نستطيع ان نبين اهمية دراسة تقييم الاسئلة الامتحان كونها :-

1- من ابرز الوسائل القياسية والنفسية الاكثر انتشاراً في المؤسسات التعليمية .

2- لها القدرة على قياس نتائج الطالب ومقدار استيعابه واتقانه للمعلومة .

3- انها وسيلة لنقل الطالب من صف الى آخر وتقدير تحصيله في العملية التعليمية ، ومدى استحقاقه في نيل الشهادة.

4- لها الاهمية والقدرة على اكتشاف مدى ما تحقق من الاهداف التدريسية وتبيان اسباب القصور في ذلك ومن ثم تحديد مدى نجاح العملية التعليمية في تحقيق مقاصدها .

5- تساهم في تقويم نمو الطالب في عده جوانب ابرزها الجانب المعرفي .

6- المساهمة في تقويم طرائق التدريس والمناهج الدراسية .

7- لها الدور البارز في رؤية وتلمس التغيرات الحاصلة في المتعلم وهي ما تسعى له التربية .

8- تدفع الطلبة للاستذكار والاستعداد لأداء الواجبات الدراسية .

9- تساعد الطلبة على ممارسة التفكير النشط وديمومته واعادة انتاجه .

10- تعد محكات حقيقية للحكم على جودة واداء المعلمين والمدرسين من خلال اعدادهم الجيد للأسئلة الامتحانية .

11- تعتبر مثيرات عقلية تؤدي دورها في العملية التعليمية .

12- تعمل بوصفها معززات للمتعلم تشعره بالإنجاز والتحصيل .

أهداف البحث :-

يستهدف البحث الحالي التعرف على مدى تحقيق اسئلة الامتحانات النهائية في قسم العلوم التربوية والنفسية للأهداف التدريسية تبعاً لنمط بلوم وتبعاً لمعايير تقييم الاسئلة (الصدق ، الشمولية ، الموضوعية ، الوضوح ، العدالة، السهولة، ونوع الاسئلة : مقاليه، موضوعية) من خلال الاجابة عن التساؤلات أو الاستفسارات الآتية :-

1- هل الاسئلة الامتحانية ملتزمة بمعيار بلوم في المجال المعرفي .

2- هل الاسئلة الامتحانية ملتزمة بمعايير تقييم الاسئلة وتأخذ بها .

حدود البحث :-

يتحدد البحث الحالي بالأسئلة الامتحانية النهائية لجميع المواد الدراسية ولأربع مراحل والبالغ عددها (60) مادة للعام الدراسي 2019/2018 وللدور الاول والثاني في قسم العلوم التربوية والنفسية/ كلية التربية/ الجامعة العراقية.

تحديد المصطلحات :- سيتم تعريف المفاهيم البارزة التي تناولها البحث الحالي وعلى النحو الآتي :

أولاً:- التقييم (Assessment)

لغةً: منْ قيَمَ يُقيمَ ، أي قدَر القيمة (بو حُمصي ، 1996 : 21) : هو عملية اضافة قيمة للشيء وهذا ما اجازة مجمع اللغة العربية فوضع لفظة التقييم في المعجم الوسيط بمعنى بيان قيمة الشيء ، وهي لفظة مختلفة عن التقويم (وديع ،2018: 2).

اصطلاحاً :- وعرفه كل من :

- (بو حُمصي ، 1996):

اعطاء المُقيم قيمته وحقه ، وهو تقدير كيفي ووصفي (حسن ، جيد ، ناقص) ، يروم من خلاله تشخيص واصدار حُكم (بو حُمصي ، 1996 : 21) .

- أنجيلو (Angelo , 1993):

هو مدخل لمساعدة المعلمين على اكتشاف ما يتعلمه الطلاب ودرجة جودة التعلم. وهذا المدخل متمركز حول الطالب ، موجه من المعلم ، مفيد للاثنين ، مرتبط بمحتوى ، مستمر ، معلم بمجريات الامور وله جذور في الممارسات الصفية الجيدة (Angelo , 1993 : 42) .

- (الشواورة ، 1995):

عملية جمع ووصف وتكميم البيانات على مستوى اداء معين تثمينها لغرض استعمالها في اتخاذ قرارات معينة وهو اكثر خصوصية من التقويم واكثر اتساعاً وشمولاً من القياس (الشواورة ، 1995 : 6) .

وقد تبنى الباحثان تعريف الشواورة تعريفاً نظرياً لأنه يتناسب مع اهداف البحث الحالي.

اما التعريف الاجرائي: فهو جميع الاجراءات القياسية والاحصائية المتخذة بشأن الاسئلة الامتحانية وصولاً الى تشخيصها كونها جيدة أم أنها اقل جودة .

ثانياً :- الاسئلة الامتحانية (Examinations Questions)

عرفها كل من :

- جود (Good , 1973) :- بأنها بعض العمليات لاختبار قدرة أو تحصيل الطلبة في بعض المجالات (Good , 1973 : 222) .
- (عبيدات ، 1989): هي المقياس الذي بين يدي المعلم يستطيع من خلاله معرفه مستوى طلبته ودرجة التحصيل التي وصل اليها كل طالب منهم (عبيدات ، 1989 : 233) .
- (العفون ، 1991) : هي المقياس الذي يستعمله لتقدير مدى تحصيل الطلبة للمعلومات وقدراتهم على التعبير الصحيح وفهم العلاقات بين الحقائق وتفكيرهم في الاسباب والنتائج وتطبيق ما تعلموه من مواقف جديدة وتعد الاسئلة على وفق اسس ومعايير خاصة لتكون شاملة وصادقة وموضوعية ومحققة للأهداف التعليمية (العفون، 1991 : 34) .

وقد تبنى الباحثان تعريف العفون تعريفاً نظرياً لأنه اكثر شمولية ونستطيع من خلاله تحقيق اهداف البحث الحالي .

أما التعريف الاجرائي: فهو اداءات الطلبة الجامعيين في قسم العلوم التربوية والنفسية على الاسئلة الامتحانية النهائية التي تم اختبارهم فيها .

الاطار النظري ودراسات سابقة:

يتضمن الفصل الحالي الموضوعات الرئيسية ذات العلاقة بموضوع البحث متمثلة بمفهوم التقييم والاسئلة الامتحانية ونمط بلوم التدريسي والخصائص القياسية التي تم وفقها التقييم فضلاً عن بعض الدراسات السابقة ذات العلاقة وعلى النحو الآتي :

أولاً:- التقييم (Assessment) :

يعد التقييم حلقة الوصل بين تكميم مستوى الطلبة قياسياً وبين اصدار حكم عليهم تقويمياً من اجل الوصول الى القرار المناسب واتخاذه ، لذا فهو يعد عملية وسيطة ضرورية تعد بمثابة تثمين لما تم قياسه وصولاً الى عملية تقويمه ، ولا تحدث هذه العملية الا وفق معايير معتمد عليها من اجل تحقيق الغايات التربوية المنشودة والتي تتحدد بتمثيل الاهداف التدريسية وتحقيقها تماشياً مع الخصائص القياسية المعتمدة عند صياغة الفقرات الاختيارية ويحصل هذا من خلال جمع وتحليل واستعمال المعلومات من النواتج التي تم قياسها بطريقة منتظمة ومستمرة لتحسين تعلم الطلبة من حيث المعرفة المطلوبة والفهم والمهارات والكفايات التي تم اكتسابها .

ويخدم التقييم الصفي اغراضاً مختلفة تشمل تقييم التعلم ، والتقييم من اجل التعلم ، والتقييم بوصفة تعلماً ، وقد كان التركيز تقليدياً على تقييم التعليم، أما التقييم من اجل التعلم فكان يستعمل احياناً لعمليات التشخيص والتقييم البنائي ولكن بشكل غير رسمي وضمني، وكان التقييم المنتظم بوصفه عملية تعلم نادرٍ ، إذ استعمل قليل من المعلمين التقييم الذاتي أو التفكير في التعلم . أما في ضوء الاتجاهات الحديثة للتعلم فهناك تركيز على جميع هذه الاغراض (درندري ، 2010 : 4) .

ويمكن ان نبين اغراض التقييم وعلى النحو الآتي :-

1- تقييم التعلم (Assessment of Learning):

وهو قياس التعلم بعد حدوثة ، باستعمال المعلومات عن اداء وانجاز الطلبة، وهو ذو بعد واحد، وتجميعي، ويشكل جزءً من المنهج ، لكنه يقود عملية التدريس (التدريس من اجل الاختبار) ويكون غير حقيقي ، ومستقلٍ عن السياق وغير مرن ، ويستعمل للتأكيد على ما يعرفه الطلبة وما يستطيعون عمله ، وما اذا كانوا قد حققوا نواتج المنهج. كما يساعد على تحديد مكانتهم مقارنة بغيرهم وهو واسع الاستعمال(Harlan , 2007) .

2- التقييم من اجل التعلم (Assessment for Learning):

ويكون متعدد الابعاد ، وثنائياً ، ومدمجاً في المنهج وحقيقياً وموجوداً في سياق ومتصفاً بالمرونة . ويصمم لإعطاء المعلم معلومات لتعديل أنشطة التعلم و التعليم، ويقوم على اعتبار ان هناك انماطاً يستعملها الطلبة للتعلم يمكن التنبؤ بها . ويتطلب من المعلم ان يجمع معلومات لا لتحديد ما يعرفه الطلبة فقط ، بل ليعرف متى وكيف يطبق الطلبة ما تعلموه . وفي هذا السياق يصبح الطلبة مستهلكين لمعلومات التقييم ايضاً .

3- التقييم بوصفة تعلماً (Assessment as Learning):

يركز هذا التقييم على دور الطلبة في ربط التعلم بالتقييم، ويكون فيه الطلبة نشطين ، ومشاركين ، وناقدين يستطيعون تكوين معاني من المعلومات وربطها بما لديهم والتعلم الجديد، وليكون الطلبة متفاعلين بشكل نشط في تكوين فهمهم ، يجب ان يتعلموا ان يكونوا مقيّمين ناقدين ، يستطيعون استعمال المعلومات ويربطونها بالمعرفة السابقة ويستعملونها في التعلم الجديد . وهذه العملية التنظيمية في عمليات ما وراء المعرفة ، بحيث يتعود الطلبة على مراقبة ما يتعلموه بشكل ذاتي ، ويستعملون ما

اكتشفوه من مراقبة تعلمهم لعمل التعديلات ، وتغيير عمليات تفكيرهم (Stiggins , 2007) .

وأياً كان التقييم معتمداً على المعلم أو على المتعلم ذو بعد واحد أو متعدد الابعاد يعبر عن جزء من المنهج أو مدمجاً فيه، متلقي للتقييم أو ناقداً نشطاً له فلابد من وجود اسئلة أو فقرات اختيارية تكون محور هذه العملية التقييمية وجوهرها وتتنوع الاسئلة بتنوع الاهداف ومعايير تصنيفها وهنالك وجهات نظر عديدة لتصنيف الاسئلة منها :

1- الاسئلة وفق ويفر وسنسي وهي نوعان:

أ- اسئلة تذكر: وتتطلب اجابة قصيرة واحدة وتحتاج قليل من التفكير ومقدار عالٍ من التخمين .

ب-أسئلة تفكير: وتتطلب مستويات عالية من التفكير كالتحليل والتركيب والتقويم والتنظيم والمقارنة والتنبؤ ، وهي بحاجة الى قدرة للإجابة عليها .

2- الاسئلة وفق ستال وانزالون :- وهي اربعة اصناف .

أ- اسئلة التصنيف: وتميز تفريق الطلبة وتمييزهم للمادة وهي ابسط انواع الاسئلة تأخذ نسبة (20%) .

ب-اسئلة التذكر: وهي معروفة وفق النماذج الاخرى وايضاً من المستويات البسيطة وتأخذ نسبة (40%) .

ت-اسئلة الاستدلال: وهي متوسطة الصعوبة ونسبتها (25%) تعبر عن قدرة الطالب في توظيف التفكير الاستدلالي .

ث-اسئلة التبرير: وهي صعبة وتحتاج الى مستويات عالية من التفكير للإجابة عنها وهي بنسبة (15%) .

3- الاسئلة وفق جون ولسون :- وهي اربعة اصناف تقع بنفس النسب وهي :

أ- الاسئلة التحليلية: وتتطلب هذه الاسئلة قدرة تحليلية تصاغ بعبارة أو جملة .

ب- الاسئلة التجريبية: وتتطلب هذه الاسئلة اضفاء التجارب الحياتية الواقعية للمستجيب يستدل بها عند اجابته .

ت- الاسئلة التقييمية: وتتطلب تثمين المادة العلمية بقيم تحمل في ثناياها المدح والثناء واللوم والنقد .

ث- اسئلة ما وراء الطبيعة: وتتناول هذه الاسئلة القضايا الروحية مثل الدين والوجود كذلك القضايا الميتافيزيقية والظواهر غير المحسوسة .

4- الاسئلة وفق جيلفورد : وصنفها الى خمسة انواع وهي :

أ- اسئلة التعرف : وهي ابسط انواع الاسئلة التي يمكن ان تطرح حول ظاهرة معرفيه محددة وتمثل (10%) .

ب-الاسئلة التفكيرية : وهي اسئلة الاساس المعرفي الذي يعتمد عليه النظام الفكري وتمثل (30%) من مجموع الاسئلة.

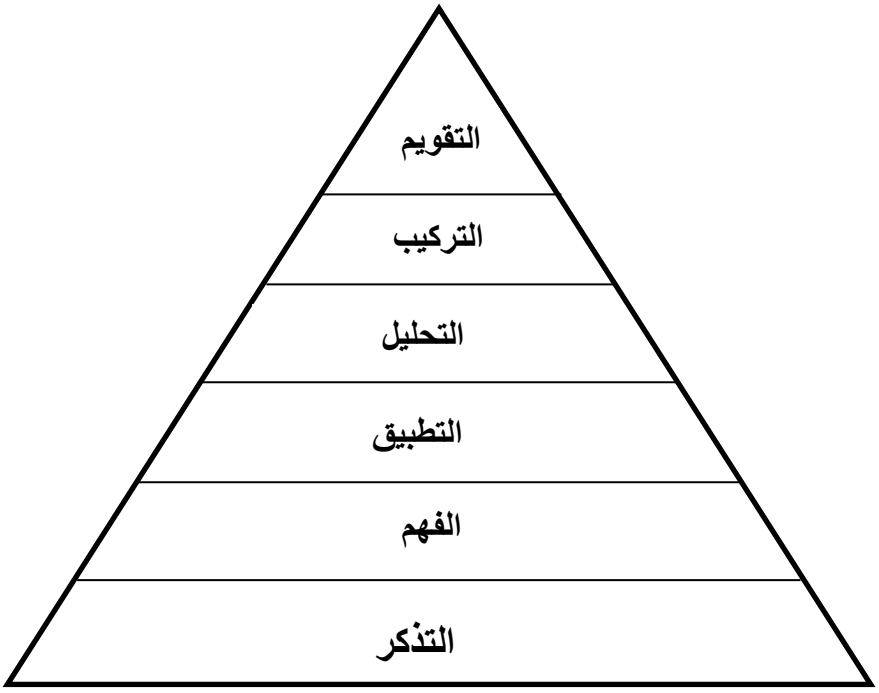
ت-الاسئلة التجميعية : وتعتمد هذه الاسئلة على اثارة قدرة الطالب على تجميع ودمج المعلومات واستخلاص الافكار ونسبتها (30%).

ث-الاسئلة التفريقية :- وهي عكس التجميعية إذ تؤدي الى تحليل وتصنيف العناصر المختلفة للموضوع ونسبتها (15%).

ج- الاسئلة التقييمية: وتشتمل هذه الاسئلة اصدار حكم تثميني بخصوص الموضوعات والظواهر والاشياء للكشف عن صحة المعلومات واهميتها وتشكل نسبة (15%) من مجموع الاسئلة الكلية (زيتون ، 54:1997) (قطامي ، 76:1998) .

5- **الاسئلة وفق تصنيف بلوم:** إذ وضع بلوم تصنيفه للأسئلة على اساس هرمي الذي وضعه لتصنيف الاهداف التعليمية في المجال المعرفي ، وهو بذلك يرى توزيع الاسئلة في أي امتحان بالشكل الذي ينسجم مع تصنيفه للأهداف المعرفية في ست مستويات متفاوتة في سهولتها وصعوبتها ومرتبة هرمياً في ضوء ذلك من قاعدة الهرم

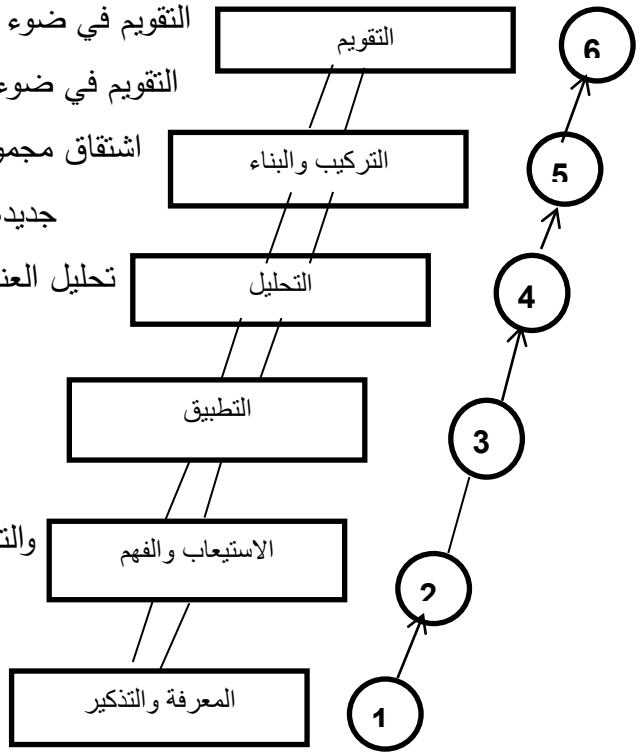
التي تمثل المستويات السهلة وتزداد الصعوبة كلما ارتفعنا نحو قمة الهرم كما موضحة في الشكل الآتي :



تصنيف بلوم لمستويات الأهداف المعرفية

ويمكن ربط التقسيمات الفرعية المعرفية وفق انموذج بلوم بالمهمات التي يتوقع ان يضمها

كل مستوى فرعي كما هي مبينه في الشكل الآتي :



التقويم في ضوء معايير داخل الموضوع (داخلية) .
التقويم في ضوء معايير خارج الموضوع (داخلية).
اشتقاق مجموعة العلاقات المجردة انتاج افكار جديدة ، انتاج خطط معينة .
تحليل العناصر والعلاقات والمبادئ التنظيمية
تطبيق المفاهيم والحقائق والمبادئ والقواعد والطرائق والنظريات .
التحويل والترجمة والشرح والتفسير، الاستنتاج والاستقراء والتنبؤ.
معرفة خصوصيات الاشياء ، معرفة طرائق معالجتها واساليبها ، معرفة عموميات الاشياء.

شكل (2)

التقسيمات الفرعية لمكونات الاهداف المعرفية

ويلاحظ من الانموذج ان كل مستوى من مستوياته يعد متطلباً سابقاً للمستوى الذي يليه فلكي يحدث الفهم لابد وان يسبقه المعرفة والتذكر وهكذا حتى يتم الوصول الى مستوى التقويم (الزغول والمحامير ، 2007 : 55 – 57) .

وبغض النظر عن نوع الاسئلة ومستواها ومن وضعها ينبغي ان تتمتع هذه الاسئلة بمجموعة من الخصائص القياسية (السايكومترية) التي تعد من مواصفات الاختبار الجيد ومن دون هذه الخصائص تفقد الاسئلة الكثير من قدرتها في قياس المستوى التحصيلي للطلبة وابرز هذه الخصائص هي :-

1- **الصدق (Validity)** : ويعد الاختبار صادقاً عندما يقيس ما اعد لقياسه، أو الذي يحقق ما أعُد لأجله . فمثلاً يعد اختبار الاستعداد المدرسي صادقاً إذا كان قادراً على قياس الاستعداد الدراسي بشكل صادق . ويقال ان الاختبار صادق اذا توافرت الادلة الكافية على انه يقيس فعلاً العامل الذي صمم الاختبار لقياسه، وانه لا يقيس عوامل اخرى ، فالاختبار التحصيلي الصادق هو اذا تمكن من قياس مدى تحقيق الاهداف التربوية المعرفية للمادة التي وضع من اجلها بنجاح (العزاوي ، 2008 : 93) . وللصدق عدة انواع منها :

أ- الصدق الظاهري : وهو الاشارة الى مدى قياس الاختبار للغرض الذي وضع من اجله ظاهرياً ، وهو المظهر العام للاختبار من حيث المفردات وكيفية صياغتها ومدى وضوحها، ويتم التوصل اليه من خلال توافق تقديرات المحكمين على درجة قياس الاختبار للسمة (العزاوي ، 2008 : 94).

ب- صدق المحتوى :- ويسمى بصدق المضمون ، أو الصدق الشامل ، والاختبار الذي يتصف بصدق المحتوى ، هو الذي تعتبر فقراته عينه ممثلة لمجال السلوك المراد قياسها ، ومن ثم اختيار عدد من الاسئلة أو الفقرات يفترض بها ان تمثل هذا المجال تمثيلاً صادقاً (الدليمي والمهداوي ، 2005 : 118) .

ويتحقق هذا النوع من الصدق في الاختبار التحصيلي من خلال الخارطة الاختبارية أو جدول المواصفات الذي يحدد فيه الاهداف السلوكية وتكون تلك الاهداف ممثلة ومن ثم الاعتماد على المحكمين لتحديد صلاحية فقرات الاختبار . وتعد النسبة المئوية من اكثر الوسائل شيوعاً في حساب هذا النوع من الصدق (مجيد، عيال ، 2012 : 96) .

ت- الصدق المرتبط بمحك (Related criterion): يتركز الاهتمام في هذا النوع من الصدق على مدى صلاحية الاختبار من الناحية الوظيفية أو العملية ،إذ ان اجراءاته

تركز على مدى تقديم الاختبار أو تحقيقه لغرض معين وهو على نوعين (العزاوي ، 2008 : 95) . هما :

- الصدق التلازمي : ويعني كشف العلاقة بين الاختبار ومحك ، تجمع البيانات عليه في وقت قريب أو قبل اجراء الاختبار . اي اننا نقارن بين درجات الافراد على الاختبار ودرجاتهم على مقياس موضوعي آخر ، على ان يعطي الافراد درجات على المحك في نفس الوقت الذي نطبق فيه الاختبار.

فعندما نريد ان نعرف هل ان الاختبار الذي صمم لقياس القدرة الرياضية مثلاً يتمتع بالصدق التلازمي . وهنا نأتي باختبار معروف بصدقه يقيس نفس القدرة الرياضية ، ويتم تطبيق الاختباران على نفس المجموعة من الطلبة ، ومن ثم يحسب معامل الارتباط بين درجات الطلبة على الاختبارين (الدليمي والمهداوي ، 2005 : 122 – 123) .

- الصدق التنبؤي : وفيه يجري التركيز على درجات الفرد في موضوعاً معيناً الآن ، وبعد فترة زمنية ربما شهور أو سنوات يتم الحصول على درجات تقديرات نفس الفرد في المحك المتصل بموضوع الاختبار مثل درجة طالب في مادة الرياضيات (90) في الثانوية بعد سنة تفحص درجته في نفس المادة في كلية الهندسة . فإذا كان الارتباط عالياً يدل على ان الاختبار الذي وضع في الثانوية صادق تنبؤياً (العزاوي ، 2008 : 98) .

ث- صدق البناء : يتحقق هذا النوع من الصدق بالدرجة التي يقيس فيها المقياس بناءاً نظرياً أو خاصية معينة . ويعد هذا النوع من افضل انواع الصدق بالنسبة للمقاييس الجديدة وتقوم دراسات هذا النوع من الصدق بفحص النظرية القائمة خلف الاختبار (مجيد وعيال ، 2012 : 98) . ويشير صدق البناء الى :

- اي حد يقيس الاختبار خاصية أو سمة لها وجود فعلي ؟

- اي حد يكون هناك تناظر بين التفسير المقترح للسمة أو الخاصية وما يقيسه الاختبار فعلاً ؟ (الدليمي والمهداوي ، 2005 : 125) .

2- **الثبات (Reliability)** : وهو دقة في القياس أو الملاحظة وعدم تناقض المقياس مع نفسه واتساقه واطراداه فيما يزودنا به من معلومات عن سلوك المستجيب ، ويعد الاختبار ثابتاً اذا كان يؤدي الى النتائج نفسها في حالة تكراره في حالة تماثل الظروف المحيطة بالاختبار والمستجيب في كلا التطبيق ويمكن ايجاد ثبات الاختبار بعده طرق ابرزها:

أ- طريقة الاختبار – اعادة الاختبار (ثبات السكون أو الاستقرار) :- وتتلخص هذه الطريقة بإعطاء المقياس أو الاختبار نفسه مرتين لمجموعة من الافراد لحساب درجة اتساق الاختبار عبر الزمن ، اي ان التقدير الذي يتحصل عليه المستجيب في اختبارها في وقت محدد فعند اعادة الاختبار عليه بعد مرور فترة زمنية يحصل على التقدير نفسه أو قريب جداً منه .

ب- طريقة الصور المتكافئة (ثبات التكافؤ) :- وفيها يتم اعداد اختبارين مختلفين لكنهما متماثلان أو متكافئان أو متوازيان في كل شيء ثم يجري تطبيقهما في وقت واحد على مجموعة من الطلبة اذ يطبق في المرة الاولى الاختبار الاول وبعد الانتهاء يقدم لهم الاختبار الثاني الذي يمثل صورة متكافئة للصورة الاولى (مجيد وعيال ، 2012 : 84) .

ت- طريقة ثبات التجانس :- لا تتطلب هذه الطريقة سوى الاختبار مرة واحدة في ضوء هذه الاستجابة التي تعتمد على الارتباطات الداخلية على بنود الاختبار يتم تطبيق المعادلات الخاصة بكل نوع وهذه الانواع هي :

- الثبات بطريقة التجزئة التصفية .

– معامل جتمان – معامل رولون – معامل الفاكروبناخ – معامل كودر – ريتشاردستون 20– معامل كودر – ريتشاردستون 21– معامل هويت – معامل بطريقة التحليل العاملي (مجيد وعيال ، 2012 : 85 – 91) .

3– **الشمولية (comprehensiveness)**: تتضمن الشمولية جانبين هما تغطية الاهداف وتغطية المحتوى ولا يحصل هذا بدقة في الاختبار التحصيلي الا بفضل وضع جدول المواصفات لتتوزع الاسئلة بحسب الاهداف التي يسعى المدرس الى تحقيقها في الاختبار ومحتوى المادة الدراسية المطلوبة اجراء اختبار فيها . وبمعنى اخر يجب ان تكون الاسئلة متنوعة من حيث السهولة والصعوبة ومتنوعة من حيث محتوى المادة الدراسية، ان درجة الشمولية تؤثر عادةً في مدى ثبات الاختبار وصدقه (العزاوي ، 2008 : 2– 1 – 103) .

4– **الموضوعية (Objectivity)** : ان الاختبار الجيد هو الاختبار الذي لا يسمح ان تتأثر نتائجه باعتقادات من يصححه وآرائه ، وان تكون درجات الاختبار التحصيلي معبرة عن الأداء الفعلي للطالب وفي حالة انتهاك هذا المفهوم فأن الدرجة تختلف باختلاف المصحح وطريقة التصحيح ، لذا ينبغي ان تكون هناك قواعد وخطوات محددة في تصحيح الاختبار لا يختلف عليها اثنان وبهذا يمكن الحكم على الاختبار انه موضوعي .

5– **الوضوح (Clarity)** : وهي ان تكون الاسئلة واضحة وغير خادعة ولا مموهة على الطالب وان تكون لغتها تتناسب مع عمر الطالب وفي حالة وجود بدائل للاستجابة ينبغي ان تتصف بالوضوح ولا يحمل السؤال اكثر من اجابة محددة ، وقد جرى تجريبه على عينة غير مشمولة بالتطبيق .

6– **العدالة (Justice)** : وهي ان تتناسب اسئلة الاختبار مع المستوى العام للطلبة وان تؤخذ ظروف الطلاب جميعاً في الاعتبار مع مراعاة الفروق الفردية .

7- **السهولة** :- وتقسم الى قسمين :

أ- سهولة التطبيق : تعد صعوبة التطبيق عائقاً امام تحقيق موضوعية الاختبار وصدقه وثباته ، فعدم فهم تعليمات الاختبار قد يؤدي الى انخفاض درجة الطالب . وتتأثر عملية تطبيق الاختبار بعده عوامل : منها ما يتصل بمقدم الاختبار ومنها ما يتصل بالطلبة الذي يطبق عليهم الاختبار ، لذلك على المدرس ان يراعي الجو النفسي والاجتماعي المناسب للطلبة بحيث يمكنهم من اداء الاختبار بدقة أو يحدد الهدف الذي وضع من اجله .

ب-سهولة التصحيح :- ينبغي ان تعد سلفاً خطوات التصحيح عند اعداد الاختبار وخصوصاً في الاختبارات المقالية ، لان الغاية المرجوة منه لا تتحقق حينما تكون طريقة التصحيح معقدة أو غير دقيقة، وحينما تحدد اجابات اسئلة الاختبار للمصحح فأن ذلك يسهل من طريقة التصحيح ويساعد على تحقيق الدقة فيه (مجيد وعيال، 2012 : 81 - 82) .

وتأسيساً على ما تقدم سيتم تقييم الاسئلة الامتحانية وفق غرض تقييم التعلم بوصفة احد اساليب التقييم الصفي المعتمد في المؤسسات التعليمية اعتماداً على تصنيف بلوم للأسئلة المعد وفق تصنيفه للأهداف التدريسية في المجال المعرفي للاختبارات الموضوعية والمقالية ، متحرين الخصائص القياسية (الصدق الظاهري، صدق المحتوى ، الموضوعية، الوضوح ، السهولة ، الشمولية) التي اعتمدت عند وضع وصياغة الاسئلة الامتحانية .

دراسات سابقة : سيتم ذكر بعض الدراسات السابقة التي قيمت الاسئلة ولها علاقة بموضوع البحث وعلى النحو الآتي

أولاً: دراسة مولتون (Moulton , 2003) :

(Effects of level review and sequence of inserted questions of test processing) (اثر مستوى الاسئلة في مدى جاهزية الاختبار).

استهدفت الدراسة الحالة التعرف على أثر مستوى وقيمة الاسئلة في مدى تجهيز الاختبار، لذلك حاولت التعرف على مدى توزيع اسئلة الاختيارات على الاهداف المعرفية، حيث تم تحليل (78) اختبار في مواد مختلفة ، وبعد استعمال التحليل المناسب ومعالجة البيانات بمجموعة من الوسائل الاحصائية المناسبة ، توصلت الدراسة الى ان الاسئلة تركزت على الاهداف الدنيا ، كما أظهرت عدم وجود فروق حقيقية في نسبة توزيع هذه الاسئلة وفق متغيرات (الجنس ، والتخصص ، والخبرة) . (Moulton , 2003 : 427 – 436) .

ثانياً: دراسة (بركات وصباح ، 2007) :

مدى تحقيق اسئلة الامتحانات النهائية في جامعة القدس المفتوحة للأهداف التعليمية تبعاً لهرم بلوم .

استهدفت الدراسة الحالية التعرف على مدى تحقيق اسئلة الامتحانات النهائية للأهداف التعليمية المعرفية في المستويات المختلفة تبعاً لهرم بلوم ،إذ تم تحليل جميع الاسئلة الامتحانية النهائية التي اعدها المدرسون للفصل الدراسي الثاني للعام (2004/2003) والبالغ عددها (133) أنموذجا باستعمال المنهج الوصفي بأسلوب تحليل المحتوى من خلال استمارة خاصة استعملت معياراً لتحليل اسئلة الاختيارات تضمنت هذه الاستمارة مستويات المجال المعرفي لبلوم ويقابل كل واحد منها الافعال التي تناسبها ولتحقيق اهداف الدراسة استعملت مجموعة من الوسائل الاحصائية مثل المتوسطات الحسابية والانحرافات المعيارية ، والنسب المئوية ، والاختيار التائي ، وتحليل التباين الاحادي ، واختبار (LDS) للمقارنات البعدية وقد توصلت الدراسة الى النتائج الآتية :

1- مستوى الاهداف المعرفية السائد في الاسئلة الامتحانية النهائية هو مستوى المعرفة (التذكر) إذ بلغت نسبته (81%) من مجمل الاسئلة .

2- وجود فروق دالة احصائياً بين المتوسطات النسبية للأسئلة تبعاً لنوع الاسئلة (موضوعية أو مقاليه) في خمسة مستويات وهي (التذكر ، والفهم والتحليل ،

والتركيب، والتقويم) وذلك لصالح الاسئلة الموضوعية في مستوى التذكر ، ولصالح الاسئلة المقالية في المستويات الاخرى . بينما لا توجد فروق جوهرية في مستوى التطبيق .

3- وجود فروق ذات دلالة احصائية بين المتوسطات النسبية للأسئلة في مستويات (التذكر ، والفهم ، والتطبيق ، والتركيب تبعاً للبرامج الاكاديمية ، بينما لا توجد فروق جوهرية في مستويين (التحليل ، والتقويم) .

4- وجود فروق دالة احصائياً بين المتوسطات النسبية للأسئلة في مستويات (التذكر، والتطبيق ، والتركيب) تبعاً للتخصصات الاكاديمية المختلفة ، بينما لا توجد فروق جوهرية في مستويات (الفهم ، والتحليل ، التقويم) (بركات و صباح ، 2007).

ثالثا: دراسة (حسن، 2017):

تقويم أسئلة الكتاب لمادة القران الكريم والتربية الإسلامية للصف السادس الإعدادي في ضوء معايير بناء الأسئلة.

استهدف الدراسة تقويم أسئلة كتاب القران الكريم والتربية الإسلامية للصف السادس الإعدادي في ضوء معايير بناء الأسئلة.واستعمل تحليل محتوى أسئلة كتاب القران الكريم والتربية الإسلامية بحسب معايير بناء الأسئلة (أداة البحث)، وللتأكد من ثبات التحليل استعملت الباحثة طريقتين الأولى تحليل محتوى الأسئلة من محللين خارجين والأخرى بإعادة التحليل بين الباحثة ونفسها بفاصل زمني، وباستعمال معادلة (هولستي).

وتوصلت الباحثة إلى عدة نتائج تتمثل بضعف توافر معايير بناء الاسئلة التي استرشد بها في تحليل اسئلة كتاب القران الكريم والتربية الاسلامية للصف السادس الاعدادي، وإلى ضعف المام المعنيين والمختصين بإعداد اسئلة كتاب القران الكريم والتربية الاسلامية بمعايير بناء اسئلة كتاب القران الكريم والتربية الاسلامية، كما توصلت إلى أن واضعي أسئلة الكتاب قد ركزوا على المستويات الدنيا وأهملوا المستويات العليا،

اعتقاداً منهم أن أغلب تفكير الطلبة تناسب المستويات الدنيا من التصنيف، وتوصلت أيضا إلى عدم تحقق شروط الاخراج الجيد لأسئلة كتاب القران الكريم والتربية الاسلامية. (حسن،2017).

رابعا : دراسة (حبيب وعلوة ، 2017):

تحليل وتقويم اسئلة الاختبارات النهائية للمواد الدراسية في الاقسام العلمية كلية التربية للبنات/ جامعة الكوفة وفق معايير معتمدة.

هدفت الدراسة الحالية الى تحليل وتقويم اسئلة الاختبارات النهائية للعام الدراسي (2014 – 2015) ، وهدفت ايضاً الى معرفة درجة شمولية تلك الاسئلة على مستويات الاهداف المعرفية وفقاً لتصنيف بلوم ودرجة مطابقتها لمعايير اعداد الاسئلة المعرفي الباحثين ، وتألف مجتمع البحث من (301) انموذج اسئلة يمثل اسئلة الاختبارات النهائية لجميع المواد الدراسية ولأربعة مراحل في اقسام كلية التربية وللإجابة على اسئلة البحث قام الباحثان بحساب تكرارات الاسئلة ونسبها المئوية والمتوسطات الحسابية بحسب الاقسام الاكاديمية ، كما استعملا الانحرافات المعيارية والاختيار التائي لمعرفة دلالة الفروق وفق التخصص الاكاديمي ، وقد توصلت الدراسة الى النتائج الآتية :-

ان اسئلة الاختبارات النهائية قد شملت جميع مستويات التفكير بحسب تصنيف بلوم، إلا انها ركزت بشكل ملحوظ على المستويات المتدنية للتفكير (التذكر، والاستيعاب ، والتطبيق) في حين كان الاهتمام بالمستويات العليا للتفكير (التحليل، والتركيب ، والتقويم) متواضعاً . (حبيب وعلوة ، 2017 : 87) .

منهج البحث وإجراءاته:

اعتمد المنهج الوصفي لتحقيق أهداف البحث، ذلك أنه "يُعنى بجمع البيانات وتحديدها ، وتحليلها وتفسيرها ، ومتابعتها بدقة، والمقارنة بينها، ثم الوصول إلى تعميمات مفسّرة ومكمّمة يمكن الإفادة" (عودة وفتحي،1987:

100)؛ واستعمل أسلوب تحليل المحتوى، الذي يعد من أنواع الدراسات الوصفية، كما جاء في (فان دالين،1985: 318)؛ ومما يستعمل بهذا الأسلوب دراسة المناهج والكتب المدرسية والأسئلة والأجوبة .. (داود وأنور،1991: 175)، وإن أسلوب تحليل المحتوى هو أسلوب منظم يقوم على ضوء خطة علمية، تتحد على أساسها خطوات التحليل ونتائجه، وهو أسلوب يعتمد التقدير الكمي من قبيل استعمال التكرارات والنسب المئوية (الرشيدي،2000: 123)، وعلى ضوء ذلك فقد أعتمد أسلوب تحليل المحتوى باستعمال وحدة الفكرة والتكرار.

مجتمع البحث:

تضمن مجتمع البحث الأسئلة الامتحانية النهائية للدورين الأول والثاني في قسم العلوم التربوية والنفسية للمواد الاختصاص للعام الدراسي 2018/2017 والبالغ عددها (30) ورقة أسئلة للدور الأول موزعة على المستوى الأول (5) وعلى المستوى الثاني (7) والثالث (10) والرابع (8)؛ ومثلها للدور الثاني. وعلى ذلك يكون عدد نماذج الأسئلة الكلي (60) نموذجا.

عينة البحث:

تم اعتماد أسلوب الحصر الشامل لمجتمع البحث، إذ تم إجراء التحليل على جميع الأسئلة الامتحانية لمواد الاختصاص في القسم البالغة (60) نسخة للدورين الأول والثاني، وتحتوي هذه الأسئلة (285) سؤالا بمعدل (4.75) لكل نسخة، بواقع (250) سؤال مقالي، و(35) سؤال موضوعي.

أداة البحث:

بعد الاطلاع على الخلفية النظرية والدراسات السابقة ولتحقيق أهداف البحث تم اعتماد بطاقة تحليل الأسئلة أداة للبحث، والتي تم إعدادها على وفق الخطوات الآتية:

أ- تحديد المعايير: تم اعتماد خمسة معايير لتحليل الأسئلة هي: (الصدق، والشمولية، وشروط صياغة الأسئلة، وتوزيع الأسئلة على المستويات المعرفية لبلوم، والإخراج الجيد للأسئلة).

ب- تحديد وصياغة الشروط: بعد الاطلاع على الدراسات السابقة تم صياغة الشرط المندرجة تحت كل معيار من المعايير الخمسة، إذ تضمن المعيار الأول (7) شروط، والمعيار الثاني (3) شروط، والمعيار الثالث (5) شروط، والرابع (6) شروط، وأخيرا المعيار الخامس (6) شروط.

ت- صياغة التعليمات: تم صياغة التعليمات التي تتضمن كيفية استدخال البيانات في بطاقة التحليل من قبل مستخدميها بما يضمن موضوعية التحليل.

ث- الصدق الظاهري: تم عرض بطاقة التحليل على لجنة مكونة من خمسة تدريسيين مختصين في العلوم التربوية والنفسية بغرض النظر في المعايير المعتمدة في بطاقة التحليل وشروط التأكد منها، وقد اتفقت اللجنة بعد التداول على صلاحية فقرات بطاقة التحليل بتعليماتها ومعاييرها وشروط التحقق منها، وبنسبة (80%) فما فوق.

ج- ثبات التحليل: تم اعتماد طرقة الاتفاق بين الباحث ونفسه وبين الباحث وزميله للتحقق من ثبات التحليل، وباستعمال معادلة هولستي (بداري،2004: 50)، والتي تبين من نتائجها:

- إن معاملات ثبات التحليل بين الباحث ونفسه بعد مرور اسبوعين لعشرين نموذج من نماذج الأسئلة أختيرت عشوائيا من نماذج مجتمع البحث، كانت على وفق ما مبين في جدول (1) الآتي:

جدول (1)

قيم معاملات الثبات لتحليل الباحث مع نفسه

ت	المعايير	معامل الثبات
1	الصدق	0.912
2	الشمولية	0.932
3	صياغة الأسئلة	0.924
4	توزيع الأسئلة الامتحانية على المستويات المعرفية لبلوم	0.936
5	الإخراج الجيد لأسئلة الامتحانية	0.953
6	الكلي	0.938

- إن معاملات ثبات التحليل بين الباحث وزميل له بعد تدريبه على كيفية التحليل وللعشرين نموذج من نماذج الأسئلة التي اختيرت سابقا، كانت على وفق ما مبين في جدول (2) الآتي:

جدول (2)

قيم معاملات الثبات لتحليل الباحث مع نفسه

ت	المعايير	معامل الثبات
1	الصدق	0.845
2	الشمولية	0.873
3	صياغة الأسئلة	0.854
4	توزيع الأسئلة الامتحانية على المستويات المعرفية لبلوم	0.853
5	الإخراج الجيد لأسئلة الامتحانية	0.862
6	الكلي	0.856

إن معاملات الثبات تشير إلى تحقق ثبات بطاقة التحليل، وهذا ما جعلها جاهزة للتطبيق.

عرض نتائج البحث وتفسيرها:

تتضمن عرض نتائج البحث في ضوء معايير بناء الاسئلة المتضمنة في بطاقة التحليل، وعلى النحو الآتي:

أولا: النتائج المتعلقة بمعيار الصدق: من خلال تحليل الأسئلة الامتحانية النهائية على وفق معيار الصدق تبين الآتي:

1. حصلت أربعة من الشروط على نسبة مئوية (100%).
2. حصل شرط (إجابة كل سؤال مستقلة عن إجابة أي سؤال آخر) على نسبة مئوية (98.245%).
3. حصل شرط (الأسئلة ذات مستوى مناسب من الصعوبة) فقد حصل على نسبة مئوية (75.439%)
4. حصل شرط (الأسئلة متنوعة بين المقالي والموضوعي) على نسبة مئوية مقدارها (12.281%). والجدول (3) يوضح هذه النتائج.

جدول (3)

نتائج التقييم بحسب معيار الصدق

ت	معيار الصدق	عدد الأسئلة	التكرار	النسبة المئوية %
1	الأسئلة تقيس ما وضعت لقياسه من المقرر.	285	285	100
2	الأسئلة موزعة على مفردات المقرر.		285	100
3	الأسئلة ذات مستوى مناسب من الصعوبة.		215	75.439
4	الأسئلة متنوعة بين المقالي والموضوعي.		35	12.281
5	وضوح المطلوب من كل سؤال.		285	100
6	خلو الأسئلة من الأخطاء العلمية.		285	100
7	إجابة كل سؤال مستقلة عن إجابة أي سؤال آخر.		280	98.245

يتضح من الجدول أعلاه تحقق ستة شروط من الشروط السبعة المتمثلة بمعيار الصدق، وقد يرجع سبب ذلك إلى وعي أعضاء الهيئة التدريسية بهذه الشروط وذلك من خلال تخصصهم (العلوم التربوية والنفسية)، إذ أن إعداد الاختبارات التحصيلية يدخل ضمن صميم هذا التخصص، أما ما يتعلق بعدم تحقق شرط بنسبة ضعيفة جدا فقد يرجع سببه إلى عزوف التدريسيين عن استعمال الاختبارات الموضوعية بسبب عدد الطلبة الكبير مما يزيد من احتمال الغش عند استعمال هذا النوع من الاختبارات.

ثانيا: النتائج المتعلقة بمعيار الشمولية: من خلال تحليل الأسئلة الامتحانية النهائية على وفق معيار الشمولية، تبين الآتي:

1. حصل شرط (الأسئلة متنوعة من حيث الأهداف التي تقيسها) على نسبة مئوية (23.509%).
2. أما شرط (توازن عدد الأسئلة قياسا بعدد مفردات المقرر) فقد حصل على نسبة مئوية (19.298%).
3. أما شرط (الأسئلة تغطي جميع مفردات المقرر) فقد حصل على نسبة مئوية (11.93%). والجدول (4) يوضح هذه النتائج.

جدول (4)

نتائج التقييم بحسب معيار الشمولية

ت	شروط معيار (الشمولية)	عدد الأسئلة	التكرارات	النسبة المئوية %
1	الأسئلة تغطي جميع مفردات المقرر.	285	34	11.93
2	الأسئلة متنوعة من حيث الأهداف التي تقيسها.		67	23.509
3	توازن عدد الأسئلة قياسا بعدد مفردات المقرر.		55	19.298

إن النتائج تبين تحقق معيار الشمولية بنسبة منخفضة وربما يعود سبب ذلك إلى عدم استعمال التدريسيين إلى جدول المواصفات في إعداد الأسئلة واعتمادهم على الاجتهادات الشخصية في إعدادها.

ثالثا: النتائج المتعلقة بمعيار الصياغة الجيدة للأسئلة:

من خلال تحليل الأسئلة الامتحانية النهائية على وفق معيار الشمولية، تبين الآتي:

1. حصل شرط (صياغة السؤال بطريقة واضحة، بعيدة عن غريب اللغة، وتجنب الصيغ المفتوحة أو الناقصة) على نسبة مئوية (83.509%).
2. وحصل شرط (صياغة السؤال بطريقة تدل على حدوث الناتج التعليمي المرغوب فيه) فقد حصل على نسبة مئوية (71.93%).
3. في حين حصل شرط (صياغة السؤال بطريقة تثير اهتمام المتعلمين) فقد حصل على نسبة مئوية (64.561%).
4. بينما حصل شرط (صياغة السؤال بما يحقق فهما واحدا لا يقبل التأويل) على نسبة مئوية (61.404%).
5. أما شرط (صياغة السؤال بما يناسب مستويات المتعلمين) فقد حصل على نسبة مئوية (58.596%). والجدول (5) يوضح هذه النتائج.

إن هذه النتائج تشير إلى تحقق هذا الشرط، وربما يرجع سبب ذلك إلى إن إعداد الأسئلة الامتحانية يدخل ضمن تخصص التدريسيين من جهة، ومن جهة أخرى قد يكون السبب إلى وعي التدريسيين وخبراتهم المتنوعة في التدريس الجامعي.

جدول (5)

نتائج التقييم بحسب معيار الصياغة الجيدة للأسئلة

ت	شروط معيار (الصياغة الجيدة للأسئلة)	عدد الأسئلة	التكرارات	النسبة المئوية %
1	صياغة السؤال بطريقة واضحة، بعيدة عن غريب اللغة، وتجنب الصيغ المفتوحة أو الناقصة.	285	238	83.509
2	صياغة السؤال بطريقة تدل على حدوث الناتج التعليمي المرغوب فيه.		205	71.93
3	صياغة السؤال بما يحقق فهما واحدا لا يقبل التأويل.		175	61.404
4	صياغة السؤال بما يناسب مستويات المتعلمين.		167	58.596
5	صياغة السؤال بطريقة تثير اهتمام المتعلمين.		184	64.561

رابعا: النتائج المتعلقة بمعيار توزيع الأسئلة على المستويات المعرفية لبلوم:

من خلال تحليل الأسئلة الامتحانية النهائية على وفق معيار توزيع الأسئلة على وفق المستويات المعرفية لبلوم تبين أنها قد توزعت على جميع المستويات إذ تراوحت النسب بين (43.86%) للمعرفة و(3.86) للتقويم، والجدول (6) يوضح هذه النتائج.

جدول (6)

نتائج التقييم بحسب معيار توزيع الأسئلة على المستويات المعرفية لبلوم

المستويات المعرفية لبلوم	عدد الأسئلة	التكرارات	النسبة المئوية %
المعرفة	285	125	43.86
الفهم		86	30.175
التطبيق		18	6.316
التحليل		30	10.526
التركيب		15	5.263
التقويم		11	3.86

إن الملاحظ من توزيع الأسئلة على المستويات المعرفية لبلوم، إن المستويات العليا للتصنيف لم تلق التغطية اللازمة في التوزيع، إذ أن النسب كانت متدنية ولا تليق بالتعليم الجامعي الذي يجب أن يكون تركيزه على هذه المستويات. وربما يعود سبب ذلك إلى تدني مستوى الطلبة عموما، ومنهم طلبة الجامعة، مما يضطر أعضاء هيئة التدريس إلى الاعتماد على الأسئلة ضمن المستويات المعرفية الدنيا؛ وعلى وجه الخصوص أن طبيعة التقويم المعتمدة بشكل رسمي في العراق تتمثل بالاعتماد على التقويم معياري المرجعي، مما يدفع عضو هيئة التدريس إلى التركيز على الأسئلة في المستويات الدنيا للحفاظ على توزيع اعتدالي لنتائج الطلبة، وتحقيق نسب نجاح معقولة.

خامسا: النتائج المتعلقة بمعيار الإخراج الجيد للأسئلة:

من خلال تحليل الأسئلة الامتحانية النهائية على وفق معيار شروط الإخراج الجيد لأسئلة تبين تحقق الشروط جميعها ي هذا المعيار، إذ تراوحت نسب التحقق بين (63.86% – 99.298%)، والجدول (7) يوضح ذلك.

جدول (7) نتائج التقييم بحسب معيار الإخراج الجيد للأسئلة

ت	شروط معيار (الإخراج الجيد للأسئلة)	عدد الأسئلة	التكرارات	النسبة المئوية %
1	الترتيب المنطقي للأسئلة مثل التدرج في الصعوبة.	285	182	63.86
2	ارتباط الأسئلة الفرعية بموضوع السؤال ومع بعضها.		221	77.544
3	خلو الأسئلة من الحشو.		265	92.982
4	خلو الأسئلة من الكلمات التي تحمل أكثر من معنى.		273	95.789
5	توزيع الدرجات على الأسئلة والفروع.		281	98.596
6	طباعة الأسئلة بشكل جيد ووضوح الكلمات.		283	99.298

إن تحقق هذا المعيار في الأسئلة الامتحانية قد يرجع سببه إلى وجود التعليمات الموجهة لأعضاء هيئة التدريس بضرورة الالتزام بشروط إخراج الأسئلة الامتحانية، كما قد يرجع أيضا إلى وعي التدريسيين بضرورة مراعاة شروط الإخراج الجيد للأسئلة.

الاستنتاجات:

1. إن تدريسيي قسم العلوم التربوية والنفسية لهم القدرة على إعداد الأسئلة الامتحانية على ضوء المعايير المطلوبة.
2. إن التراكمات السلبية للسلم التعليمي في الدراسة ما قبل الجامعية قد يكون سببا في الخلل الذي ظهر في ضعف تحقق بعض المعايير.
3. يحتاج أعضاء هيئة التدريس إلى دورات تدريبية لمعالجة حالات الضعف في مراعاة بعض معايير إعداد الأسئلة الامتحانية.

التوصيات:

1. ضرورة تهيئة القاعات الامتحانية بما يسمح للتدريسيين من تنويع الأسئلة بين الموضوعية والمقالية والموازنة بينهما.
2. عمل دورات تدريبية للتدريسيين بمشاركة مختصين في القياس والتقويم لغرض معالجة حالات الضعف في مراعاة بعض معايير إعداد الأسئلة الامتحانية، ولتدريبهم على سبل التعامل مع الظروف التي قد تكون عائقا أمام مراعاة معايير إعداد الأسئلة الامتحانية.
3. الارتقاء بمستوى التدريس بما يسهم في تحقيق المستويات المعرفية جميعها من أجل مراعاة شروط إعداد الأسئلة الامتحانية في هذا الجانب.

المقترحات :

1. إجراء دراسة لتعرف أسباب الأخفاق في مراعاة معايير الشمولية وتمثيل المستويات المعرفية لبلوم.
2. إجراء دراسة مماثلة لتقييم الأسئلة الامتحانية للمقررات الدراسية جميعها.

المصادر العربية:-

- أبو علام ، رجاء محمود . (2005) . **قرارات في المناهج** ، ط2 ، القاهرة : مكتبة النهضة المصرية .
- داود ، عزيز حنا وأنور حسين (1990).**مناهج البحث في التربية وعلم النفس** ،جامعة بغداد، مطبعة دار الحكمة للطباعة والنشر، بغداد.
- الدرندري، اقبال زين العابدين.(2010). **نحو تقييم افضل للتعلم: افضل الممارسات وأهم التحديات**. الرياض: جامعة الملك سعود.
- الدليمي ، احسان عليوي والمهداوي ، عدنان محمود . (2005) . **القياس والتقويم في العملية التعليمية** ، ط2 ، بغداد : دار الكتب والوثائق .
- الرشيدي، بشير صالح (2000). **مناهج البحث التربوي رؤية تطبيقية مبسطة**، دار الكتاب الحديث، بيروت، لبنان.
- بداري، هند أحمد (2004). **الثقافة العلمية في مصر (الصحافة نموذجاً)**، ط1، المكتبة الأكاديمية، مصر.
- حسن، طيبة سرمد محمد (2017). **تقويم أسئلة الكتاب لمادة القران الكريم والتربية الإسلامية للصف السادس الإعدادي في ضوء معايير بناء الأسئلة**، رسالة ماجستير غير منشورة، بغداد، الجامعة العراقية، كلية التربية.
- سند ، روبرت ، وآرثر كارين . (1985) . **الاستجواب الابداعي واساليب الاصغاء المتحسس** ، ترجمة : رؤوف عبد الرزاق ، ط2، الموصل : مطبعة الموصل .
- كحلوات ، احمد . (2000) . **خصائص الاختبار الجيد** . ورقة عمل مقدمة في اليوم الدراسي حول الامتحانات والتعيينات : جامعة القدس المفتوحة (1 – 5) .
- عودة، أحمد سليمان، وفتحي حسن ملكاوي .(1987). **أساسيات البحث العلمي في التربية والعلوم الإنسانية**، مكتبة المناف، الزرقاء، الأردن.
- عيسوي ، عبد الرحمن محمد . (1985) . **التقويم التربوي** ، مسقط : مطابع العتبة.

- التل ، شادية احمد . (1988) . **دراسة تجريبية في تأثير استخدام الاسئلة التي تتطلب قدرات عقلية عليا في الاستيعاب القرائي** ، بحث مقدم الى ندوة طرائق التدريس ، بغداد : الجامعة المستنصرية .
- بركات ، زياد وصباح ، عبد الهادي . (2007) . **مدى تحقيق اسئلة الامتحانات النهائية في جامعة القدس المفتوحة للأهداف التعليمية تبعاً لهرم بلوم** ، الاردن : دار امجد
- بو حُمصي، عادل .(1996). تقييم الاسئلة الامتحانية، عمان: دار المسيرة.
- الياس ، وديع (2018) . تعريف التقييم www.aptksa.org
- الشواورة ، شادي يوسف (1995) **الاختيارات محكية المرجع بناء وتحليل (دراسة مقارنة)** ، الاردن : الكلية العلمية الاسلامية.
- عبيدات ، احمد سلمان (1989) . **اساسيات في تدريس الاجتماعيات وتطبيقاتها العلمية** ، ط2 ، عمان : مطبعة النور الانموذجية .
- العفوان ، نادية حسين . (1991) . **تقويم الاسئلة الامتحانية للصف الخامس العلمي لمادة الاحياء** ، رسالة ماجستير غير منشورة ، بغداد : جامعة بغداد ، كلية التربية .
- اقبال زين العابدين ، درندري . (2010) . **نحو تقييم موجه للتعلم : افضل الممارسات واهم التحديات** ، جامعة الملك سعود ، الرياض .
- زيتون ، كمال . (1997) . **التدريس : نماذجه ومهارته** ، الاسكندرية : المكتب العالمي للكومبيوتر والنشر والتوزيع .
- فان دالين ، ديو بولد (1985) .**مناهج البحث في التربية وعلم النفس**، ترجمة: محمد نبيل نوفل وآخرين، ط2، مكتبة الأنجلو المصرية، القاهرة، مصر.
- قطامي ، يوسف . (1998) . **سيكولوجية التعلم الصفي** ، عمان : دار الشروق .

www.qou.edu/hom/sciResearch /researchers pages/Ziad Barakat/r22-drziad Barakat . Pdf .

- الزغول ، عماد عبد الرحيم والمحاميد ، شاكر عقله . (2007) . **سيكولوجيه التدريس**.
- مجيد ، عبد الحسين رزوقي وعيال ، ياسين حميد . (2012) . **القياس والتقويم للطالب الجامعي** ، جامعة بغداد ، كلية التربية ابن رشد : مكتبة اليمامة للطباعة والنشر.

- العزاوي ، رحيم يونس كرو . (2008) . **المنهل في العلوم التربوية : القياس والتقويم في العملية التدريسية** ، الاردن : دار دجلة.
- حبيب عبد الحسين شاكر وعلوة ، عبد الرزاق شنين (2017) . تحليل وتقويم اسئلة الاختيارات النهائية للمواد الدراسية في الاقسام العلمية كلية التربية للبنات/جامعة الكوفة وفق معايير معتمدة ، **مجلة كلية التربية للبنات للعلوم الانسانية** ، العدد : 21 . (87-114) . www.qou- edu/hom/sciResearch/researchers pages/ziad Barakat/ r22-drziad Barakat/r22-drziad Barakat.pdf .

المصادر الاجنبية :-

- Marso , K & pigge , G . (2002) . **objective and subjective methods for evaluating conversation quality in multimedia serices** . www.in/ Itu – 2001 – 2004/com 09-921 . htm
  WWW.Qou.edu /hom/sciResearch/researcher pages/Ziad Barakal /r22-drziad Barakat – pdf
- T.A. Angelo & p.k. (1993) **. Crossing classroom Assessment Technic use** , 2nd edition scan Francisco : Jossey – Bass http://www.Tmcc . edu/vp/aa/assessment/support/cat/
- Good , Carter . v . (1973) . **Dictionary of Education** , Third Edition New york : Me grow Hill book co .
- Harlan , W. (2007) **. Assessment of Leering** . London : Sag to Student Motivation and Learning . Ph Delta kappa EDGE , 2(2) , 19 .
- Moulton – (2003) . Effects of level review and sequence of inserted questions of test processing **. Journal of Educational psychology ,** 73 , 427 – 436 .