فحص الأداء التفاضلي لفقرات أداة تقييم الأداء التدريسي تبعاً لمتغير المستوى الدراسي في جامعة اليرموك باستخدام طريقة مانتل- هانزل العامة.

د. أيمن محمد أحمد فريحات* **د. علي يوسف أحمد البطوش****

تاريخ وصول البحث: ٢٥ / ١٠/ ٢٠١٨م تاريخ قبول البحث: ٢٣/ ١٢/ ٢٠١٨م

الملخص

هدفت الدراسة إلى فحص الأداء التفاضلي لفقرات أداة تقييم الأداء التدريسي في جامعة اليرموك، تبعاً لمتغير المستوى الدراسي من وجهة نظر الطلبة باستخدام طريقة مانتل–هانزل العامة. ولتحقيق أهداف الدراسة تمت الاستفادة من بيانات الأداة المعتمدة من المجالس المختصة في جامعة اليرموك. وقد تم تحليل استجابات (٧٣٤٣) طالبًا وطالبة من طلبة جامعة اليرموك على استبانة تقييم الأداء التدريسي لأعضاء هيئة التدريس، خلال الفصل الصيفي من العام الدراسي ٢٠١٥/٢٠١٦م، أظهرت النتائج وجود فقرة واحدة ذات أداء تفاضلي للمستوى الدراسي تنص على: "يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة"، بين المستوى الثاني والمستوى الرابع فأكثر ولصالح طلبة المستوى الرابع فأكثر، وعند حساب حجم الأداء التفاضلي للفقرة تبعاً للمستوى الدراسيفي الأداة المستخدمة في جامعة اليرموك،تبيّن أنّه لا يوجد أثر للفقرة ذات الأداء التفاضلي.وأظهرتنتائج التحليل العاملي التوكيدي أنّ حذف الفقرة ذات الأداء التفاضلي في الأداة المستخدمة في جامعة اليرموك تبعاً للمستوى الدراسي يسهم وبدلالة إحصائيةفي تحسين مؤشرات حسن المطابقة للأداة، وبالتالي في إعداد الفقرات.

وأوصت الدراسة بزيادة اهتمام الجامعات بأدوات التقييم للأداء التدريسي مع الأخذ بعين الاعتبار موضوع الأداء التفاضلي للفقرات عند إعدادها.

الكلمات المفتاحية: لأداء التفاضلي للفقرات، أداة تقييم الطلبة لأداء التدريس، طريقة مانتل-هانزل العامة، الأداء التفاضلي للمستوى الدراسي.

* أستاذ مساعد،قسم العلوم التربوية، كلية عجلون الجامعية، جامعة البلقاء التطبيقية

** وزارة التربية والتعليم.

Abstract

The study aimed at examining differential items functioning (DIF) in teaching performance evaluation at Yarmouk University according to the educational level from the students' perspective using the Mantel-Hansel method. To achieve the aims of the study, the data was obtained from Yarmouk University. This data set was collected during the summer semester of the academic year 2015/ 2016. The total number of the analyzed data was of (7343) questionnaire. The results showed that one item showed differential functioning according to students' Educational Level in favor for the fourth level students compared to second level students, when calculating the size of DIF for the item showed DIF according to Educational Level in the instrument used at Yarmouk University, it was found that there is no effect for the items showing DIF.

The results of the study also indicated that removing the item with DIF in the instrument adopted at Yarmouk University according to educational level significantly improves goodness of fit indices of the instrument and thus in the preparation of items. The study recommended the need for universities to pay more attention to assess the quality of instruments with respect to DIF when preparing these instruments.

Keywords: Differential items functioning, Student Evaluation, Performance of teaching evaluation.

المقدمة:

يعدّ موضوع تقييم أداء عضو هيئة التدريس في الجامعة من الموضوعات التي لاقت الاهتمام لدى الباحثين في البحوث العربية والبحوث الأجنبية ؛ ويعود ذلك الى أهمية هذا الموضوع وتأثيره المباشر في أداء الطلبة وتحصيلهم الدراسي، بالإضافة إلى الدور المهم والفاعل الذي يؤديه أعضاء هيئة التدريس في الجامعة في توجيه سلوك الطلبة وتعزيز نموهم

الشخصي والمعرفي، ولأهمية دور عضو هيئة التدريس في التعليم الجامعي.

إنّ موضوع تقييم الأداء التدريسي الجامعي سيظل عاملًا مهماً للأساتذة والطلبة والباحثين لعدة سنوات قادمة، وإنّ من الضروري التعرف إلى آراء الطلبة ووجهات نظرهم عند تقييم عضو هيئة التدريس في الجامعة؛ لأنّهم الذين يتعاملون ويتفاعلون مع عضو هيئة التدريس بشكل مباشر. وعلى الرغم من أنهم قد يتأثرون أحيانًا بعوامل شخصية وذاتية أو غير موضوعية في تقييمهم للأساتذة، إلا أنّ هناك شبه اتفاق عام بين أعضاء هيئة التدريس في الجامعة على حتمية التقييم، وذلك بناء على ما أشارت إليه نتائج الدراسات والبحوث في هذا المجال، ومن ثم تطوير هذا الأداء، الأمر الذي يؤدي إلى تمكين الجامعة من أداء رسالتها على الوجه الأمثل وتخليصها من كافة المعوقات التي تعترض مسيرتها(Tang,1997)[١].

وبشكل عام فإنّ عملية التقييم تعدّ مؤشرًا على مدى تحقيق الأهداف المرسومة، ومؤشراً على فعالية الأداء، وتسهم في تعرّف مواطن الخلل، ووضع اليد عليها تمهيدًا للبحث في إيجاد سبل المعالجة، كما تسهم في التعرف على نقاط القوةلتعزيزها وتنميتها، والعمل على إكسابها لجميع المعنيين بها كونها العنصر المؤثر في النجاح،ويعرف تقييم الأداء بأنه: العملية التي يتم من خلالها التعرف إلى الجوانب الإيجابية والجوانب السلبية الخاصة بتحقيق الأهداف وإنجاز معدلات الأداء المستهدفة[٢]، لذلك فانّ مؤسسات التعليم العالي قد تستخدم في تقويمها لأداء أعضاء هيئة التدريس أدوات عاجزة تؤدي إلى تقويم غير عادل للأداء التدريسي، لهذا الغرض، تم تحليل محتوى الأدوات المستخدمة في تقويم أعضاء هيئة التدريس والموجّهة إلى الطلبة، وأظهرت نتائج تحليل محتوى الأدوات المستخدمة في تقويم الأداء التدريسي لفقرات مضلّلة، وغير واضحة، وغيرموضوعية، وفقرات لم ترتبط مع الأداء التدريسي الذي تقيسه، ضرورة مراجعة الأدوات التقويمية المستخدمة في تقويم الأداء التدريسي- الجامعي، لأنّ أكثر المشكلات هي في محتوى أدوات التقويم المستخدمة (Tagomori,1993)[٣].

وحظيت قضية عدالة أدوات القياس باهتمام المختصين في مجال القياس النّفسي والتّربوي، ونظرًا لأهمية القرارات التي يجري اتخاذها استنادًا إلى نتائج تلك الأدوات من اختيار الأفراد للوظائف المختلفة، أو تصنيفهم تبعاً لقدراتهم أو تحصيلهم، أو توجيه الطّلبة مهنيًا، أو تحديد الصفوف المناسبة لهم، وتتصف أدوات القياس بالعدالة عندما تكون

صادقة بالنسبة لجميع الأفراد والمجموعات، وتعطي فرصاًمتساوية لجميع المفحوصين لإبداء المعرفة والمهارات المكتسبة ذات الصّلة بغرض أدوات القياس(Roever, 2005)[٤].

ويمكن اعتبار أدوات القياس متحيزة إذا كانت تحابي مجموعة من المفحوصين من نفس مستوى القدرة على حساب مجموعات أخرى، وبالتالي ينتهك مبدأ العدالة بين المفحوصين. وفي هذا السياق فإن التّحيز هو أحد مهدّدات صدق علامات أدوات القياس، فعندما تحابي إحدى الفقرات مجموعة من المفحوصين، فإنّ ذلك يعني أنّ هذه المجموعة لها أفضلية دون المجموعات الأخرى في التوصل إلى الإجابة الصحيحة على الفقرة (Pokropek&Kondratek, 2012)[٥].

التّحيز والأداء التّفاضلي للفقرات: يربكاميليوشيبارد Camilli& Shepard,[٦] 1994إن الفقرة تعدّ متحيزة عندما تحابي مجموعة عن مجموعة أخرى من نفس مستوى القدرة للسّمة المراد قياسها، ويرى كلّ من هامبلتون وروجرز Hambleton& Rogers, 1995 [٧] أنّ الفقرة تعدّ متحيزة إذا كان الفرق في المساحة بين منحنيات الخصائص في المجموعات المختلفة المتكافئة في القدرة ذات دلالة إحصائية، كأن يكونوا ذكوراً وإناثًا، أو أن يكونوا من أعراق مختلفة، أي أن احتمالات الاستجابة الصحيحة على الفقرة مختلفة للأفراد في المجموعات الفرعية من لهم نفس القدرة، أما هامبلتونوسوا مينثان وروجرز[٨] Hamblelton,Swaminthan& Rogers, 1991 فقد عرّفوا الفقرة التي تظهر أداءً تفاضلياً، بأنها: تلك الفقرة التي تكون دالة الاستجابةلها غير متماثلة في المجموعات الفرعية.

والأداء التّفاضلي للفقرة من منظور نظرية الاستجابة للفقرة يكون بفحص العلاقة بين الاحتمالات المشروطة للإجابة الصحيحة على الفقرة، وبين السمة الكامنة المراد قياسها، بحيث يتم مقارنة منحنيات خصائص الفقرة في المجموعتين الفرعيتين لمجتمع إحصائي لكشف الأداء التّفاضلي فيها والتي يجري رسمها وفق النموذج اللوجستي المستخدم(Gybles, 2004)[٩]، ويقسم نينغ ولان (Ning& lane,) [١٠] 1996 الأداء التّفاضلي في الفقرات ثنائية التّدريج إلى قسمين، هما:

أولاً: الأداء التّفاضلي المنتظم: حيث يمكن اعتبار الأداء التّفاضلي بأنه منتظم عندما تكون احتمالية الإجابة الصحيحة للفقرة دائما أكبر لمجموعة ما عند جميع مستويات القدرة، أي عندما لا يكون هناك تفاعل بين مستوى الأداء وعضوية المجموعة.

ثانياً: الأداء التّفاضلي غير المنتظم: يظهر عند حدوث تفاعل بين مستوى القدرة وعضوية المجموعة، أي أنّ فروق احتمال الإجابة الصحيحة ليست ثابتة عند

جميع مستويات القدرة، فمرّة يظهر الأداء التّفاضلي لصالح المجموعة المرجعية، ومرة أخرى يظهر لصالح المجموعة المستهدفة عند مستويات قدرة مختلفة، ويشير دورانز وهولاند (Dorans& Holland, 1993)[١١] أنّ الأداء التّفاضلي للفقرة يهتم في الخصائص السيكومترية لها من حيث الكيفية التي تعمل بها في المجموعات المختلفة، بينما يهتم تحيز الفقرة بالمعنى الاجتماعي بشكل كبير؛ وهذا يدل على وجود فرق بين مفهوم التّحيز والأداء التّفاضلي للفقرة.

وبشكل عام، فإنّ أي دراسة تتعلق بالأداء التّفاضلي للفقرة تتطلب إجراء ثلاث خطوات رئيسة تتمثل بالآتي (Uiterwijk&Vallen, 2005)[١٢]:

١- استخدام إجراءات إحصائية محدّدة للكشف عن الفقرات ذات الأداء التّفاضلي.

٢-معرفة العناصر المسؤولة عن ظهور الأداء التّفاضلي، ولتحقيق ذلك لا بد من استخدام طرق مختلفة للحصول على مصادر الأداء التّفاضلي للفقرات مع الأخذ بالاعتبار صعوبة تحديد العنصر المسبب للأداء التّفاضلي بشكل كامل.

٣- اتّخاذ قرار حول مصادر الأداء التّفاضلي فيما إذا كانت ذات علاقة ببنية الاختبار والسمة المقيسة.

وقد أشارت بينيتوونافاس آرا (Benito &Navas-Ara, 2000)[١٣] إلى أن أكثر الطرق شيوعا ضمن هذه المجموعة، هي:

١-طريقة مانتل-هانزل.

٢-طريقتي كأي تربيع للاستجابات الصحيحة، وكأي تربيع للاستجابات الصحيحة والخاطئة.

٣-النماذج الخطيّة اللوغاريتمية.

٤-طريقة الانحدار اللوجستي.

طريقة مانتل-هانزل:

قدمت زويكودونغو وغريما (Zwick, Donoghue, &Grima, 1993)[١٤] معادلات لاستخدام طريقة مانتل لتحديد الأداء التّفاضلي للفقرة في الفقرات متعدّدة التّدريج. وكما هو الحال مع طريقة مانتل-هانزل للفقرات ثنائية التّدريج، يجري أولًا مطابقة المجموعة المرجعية و المستهدفة من حيث السّمة أو العلامة الكلية للاختبار، ثم تنظم البيانات في جداول، وحساب طريقة ما نتل قائمة على مقارنة متوسطات المجموعات في الأداء على الاختبار ككل أو على فقراته المتكافئة, والفرضية الصفرية هي أنه ليس هناك ارتباط شرطي بين علامة الفقرة وعضوية المجموعة, Potenza &Dorans 1995[١٥].

بعد الاطلاع على الدّراسات السابقة والطرق المستخدمة في الكشف عن الأداء التّفاضلي، وجد الباحث أنّ الطريقة

الأقرب للدّراسة هي طريقة مانتل-هانزل، حيث تعدّ هذه الطريقة من الطرق المستخدمة في الكشف عن الأداء التّفاضلي في الوقت الحاضر وذلك لسهولة حساباتها وإجراءاتها، واستخدم الباحث طريقة تخدم غرض الدّراسة بالشكل التام والمطلوب وأكثر تطورا من طريقة مانتل-هانزل، وهي طريقة مانتل-هانزل العامة.

طريقة مانتل-هانزل العامة (Generalized Mantel-Haenszel)[١٦]:

تتعامل طريقة مانتل هانزل العامة مع فئات الفقرات كبيانات متعدّدة الاستجابات، ويتم استخدامها عندما تكون متغيرات العامل والفقرة ضمن المقاييس التصنيفية، الاسمية أو الترتيبية، فالبيانات الناتجة يمكن تلخيصها في الجداول المخصّصة Q بالأبعاد RX C، وفي الفقرات متعدّدة الاستجابات تقدم أكثر من استجابتين، ومن أهم أمثلتها فقرات مقياس ليكرت. وتحليل الأداء التفاضلي للفقرة في هذا النوع من الفقرات باستخدام أساليب مانتل هانزل العامة ومانتل هانزل (Fidalgo, 2005)[١٧].

وأشار فيدالغو ومادييرا Fidalgo & (Madeira, 2008)[١٨] إلى أنّ طريقة مانتل-هانزل العامة هي الطريقة الأكثر تطوراً والأكثر توسعاً من الطريقة التقليدية الشائعة، حيث تقوم الطريقة التقليدية على تنظيم البيانات في جدول توافق ثنائي البعد فقط (٢×٢)، بينما في طريقة مانتل-هانزل العامة يتم بتحليل (DIF) لجداول Q بالأبعاد RX C لذلك نستطيع تطبيق إحصائيات مانتل - هانزل العامة على تقييم (DIF) في جماعات كثيرة (R >2)سواء لفقرات ثنائية(C = 2) أو لفقرات متعدّدة الاستجابة (C > 2). الاهتمام الرئيسي في العلاقة بين تصنيفين اثنين، هما مستوى القدرة واستجابات فئات المتغير، حيث تنظم البيانات في جدول توافق متعدّد الأبعاد (R×C) الآتي:

Data Structure in the *h*th Stratum

	Response Variable Categories						
Factor Levels	1	2	·	*j*	·	*C*	Total
1	n_{h11}	n_{h12}	·	n_{h1j}	·	n_{h1C}	$N_{h1\cdot}$
2	n_{h21}	n_{h22}	·	n_{h2j}	·	n_{h2C}	$N_{h2\cdot}$
⋮	⋮	⋮	⋮	⋮	·	⋮	⋮
i	n_{hi1}	n_{hi2}	·	n_{hij}	·	n_{hiC}	$N_{hi\cdot}$
⋮	⋮	⋮	⋮	⋮	·	⋮	⋮
R	n_{hR1}	n_{hR2}	·	n_{hRj}	·	n_{hRC}	$N_{hR\cdot}$
Total	$N_{h\cdot1}$	$N_{h\cdot2}$	·	$N_{h\cdot j}$	·	$N_{h\cdot C}$	$N_{h\cdot\cdot}$

وتعطى معادلة مانتل-هانزل العامة بالصيغة الآتية:

1- $$Q_{CMH} = \left\{\sum_{h=1}^{H}(n_h - m_h)'A_h'\right\}\left\{\sum_{h=1}^{H}A_hV_hA_h'\right\}^{-1}\left\{\sum_{h=1}^{H}A_h(n_h - m_h)\right\}$$

2-$nh = (nh11, nh21, \ldots, nhRC)'$ $(CR \times 1)$

3-$mh=Nh\cdot\cdot(ph\cdot^{*}\otimes ph^{*}\cdot)$ $(CR \times 1)$

4-$Vh = N^2\, h\cdot\cdot /(Nh\cdot\cdot - 1)\{ (Dph\cdot^{*} - ph\cdot^{*}p'h\cdot^{*})\otimes(Dph^{*}ph^{*}\cdot p'h^{*}\cdot)\}$ $CR \times CR$)

حيث إن:

n_h: التكرارات المشاهدة.

m_h: التكرارات المتوقعة.

V_h: مصفوفة التّباين.

A_h: مصفوفة الاقترانات الخطيّة.

Q: هي إحصائياً لإبعاد (RXC) لتقييم (DIF) في فقرات ثنائية ومتعدّدة الاستجابات (C) وبشكل متزامن في مجموعات كثيرة (R >2) سواء لفقرات ثنائية (C=2) أو لفقرات متعدّدة الاستجابة (C >2). وأشار فيدالغو وماديرا Fidalgo & Madeira, 2008[19] إلى مبررات استخدام طريقةمانتل-هانزل العامة:

١- تعدّ أكثر فاعلية في الكشف عن الأداء التّفاضلي للفقرات متعدّدة الاستجابات.

٢- تعدّ أكثر تطوراً وأكثر توسعاً من طريقة مانتل-هانزل.

وبيّن بينفيلد Penfield,2001 أن استخدام مستوى ألفا المعدل حسب Bonferroni يعطي سيطرة جيدة على معدل خطأ النوع الأول، ويتم استخدام GMH بتحليل (DIF) ذات الجداول بأبعاد RXC لتقييم (DIF) بشكل متزامن في مجموعات كثيرة. وأنه في حالة رفض الفرضية الصفرية فسيكون من الضروري إجراء مقارنات زوجية باستخدام مستوى ألفا معدل لتحديد بين أي المجموعات يوجد (DIF) وكما أشار إليه من قبل (Penfield, 2001) لذلك فإنّ اكتشاف (DIF) بهذه المنهجية يجب أن **يشتمل ثلاث مراحل:**

-**المرحلة (١) إجراء** تحليل أول (إجراء ثنائي المراحل)، أو عدد من التحليلات (إجراء تجريبي)، مع إلغاء كل تلك الفقرات التي تم اكتشاف (DIF) فيها من مطابقة المتغير.

-**المرحلة (٢)** تحليل الفقرات باستخدام المتغير الذي تم تنقيته والذي تم الحصول عليه في المرحلة (١).

-**المرحلة (٣)** تطبيق نفس إحصائيات MH العامة على الفقرات التي تم كشف وجود (DIF) فيها في المرحلة الثانية، وهذه المرة تجري مقارنة المجموعات 2X2 باستخدام مستوى ألفا المعدل حسب Bonferroni.

الإطار النظري:

مشكلة الدّراسة وأسئلتها:

تناولت العديد من الدراسات الأداء التفاضلي وفقراته، واهتمّت بعض تلك

الدراسات بالكشف عن الفقرات ذات الأداء التفاضلي في تقييم الأداء التدريسي، بينما اهتمت دراسات أخرى بالمقارنة بين طرق الكشف عن الأداء التفاضلي في ضوء معايير محددة. وتكشف مراجعة الدراسات التربويّة العربيّة في هذا المجال عن عدم تناولها الأداء التفاضلي للفقرة تبعاً للمستوى الدراسي للطلبة (الأول، والثاني، والثالث، والرابع فأكثر).

ولا بد من معرفة أنّ الاختلاف في تقديرات الطلبة للأداء التدريسي الناتج عن اختلاف المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر) هو فرق ناتج عن الاختلاف الحقيقي في الأداء التدريسي أم ناتج عن وجود فقرات تعطي أفضلية لأعضاء هيئة التدريس تبعاً لمستوى تعليمي دون آخر، أوأنّ فهم الطلبة لفقرات أداة تقييم الأداء التدريسي المستخدمة مختلفة.

لذلك ستحدد الدراسة الحالية الفقرات التي تبدي الأداء التفاضلي في أداة تقييم الطلبة لأداء أعضاء الهيئة التدريسية في جامعة اليرموك تبعاً لمتغيرالمستوى الدراسي باستخدام طريقة مانتل- هانزل العامة، ومن ثمّ معرفة حجم الأداء التفاضلي لفقرات المستوى الدراسي للطلبة (الأول، والثاني، والثالث، والرابع فأكثر) لكلّ منها. فقد يكون الخلل ليس بالفقرة؛ لذلك لا نحتاج إلى شطب الفقرة. بالإضافة إلى ذلك فإننا سنعمل على دراسة ماذا يترتب على مؤشرات حُسن المطابقة بعد حذف الفقرات التي أظهرت أداءً تفاضلياً؛ وذلك من خلال الإجابة عن السؤالين الآتيين:

١-هل تظهر فقرات أداة تقييماللطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة أداءً تفاضليا تبعاً لمتغير المستوى الدراسي للطلبة؟

٢- ما أثر الفقرات ذات الأداء التّفاضلي علىمؤشرات حُسن المطابقة لأداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً لمتغير المستوى الدراسي للطلبة؟

هدف الدراسة:

هدفت هذه الدراسة إلى فحص الأداء التفاضلي تبعاً لمتغير المستوى الدراسي في فقرات متعددة الاستجابات, في أداة تقييم الطلبة لأداء أعضاء الهيئة التدريسية في جامعة اليرموك.

أهمية الدّراسة:

تبرز أهمية هذه الدراسة النظرية في معالجتها الأداء التفاضلي في كونها تبحث في الأداء التفاضلي لفقرات أداة تقييم أعضاء الهيئة التدريسية تبعاً لمتغير المستوى الدراسي للطلبة (الأول، والثاني، والثالث، والرابع فأكثر) على

الفقرة، وبذلك تضيف موضوعاً جديداً لم تتطرق إليه الدراسات العربية من قبل، أما الأهمية العملية تتمثل في تسليط الضوء على أهمية معرفة حجمالأداء التفاضلي للفقرة، وكيف يساعدنا على فهم الأداء التفاضلي للفقرات بشكل أفضل، وارشادات تساعد المهتمين على بناء الفقرات؛ لانتقاء أفضل الطرق في الكشف عن الأداء التّفاضلي لفقرات التقييم، وذلك للتخلّص من الفقرات التي تظهر أداءَ تفاضليًا لفقرات التقييم، أو تعديلها لتحقيق العدالة بين الأفراد في الاستجابة على الفقرات.

وتُساعد هذه الدراسة على تطوير أداء عضو هيئة التدريس في الجامعة وتحسين أدائه التدريسي من خلال التعرّف إلى جوانب القوة والنقاط الإيجابية لتعزيزها وتشخيص جوانبِ الضعف والقصور في أدائه لمعالجتها. ويُمكن أن تسهم الدراسة من خلال النتائج التي تسفر عنها في التعرف إلى الكفايات التدريسيّة الضرورية التي يجب أن يمارسها أعضاء هيئة التدريس في الجامعة لمساعدة الطلبة على تحقيق الأهداف التربوية المرجوّة، وكذلك أثر متغير المستوى التعليمي في تقييم الأداء التدريسي. ويمكن أن تساعد هذه الدراسة متخذي القرار في الجامعة على وضع برنامج أو تعميم دورة لرفع درجة تأهيل أعضاء هيئة التدريس وتبصيرهم بالمهارات اللازمة والكفايات التدريسية للنجاح في مهنتهم الجامعية وتقييمهم الذاتي, وإثراء المكتبة العربية بدراسات حول موضوع الأداء التفاضلي للفقرات في تقييم الأداء التدريسي، بالإضافة أيضًا إلى ندرة وجود دراسات تطرقت إلى احتمالية أن تحتوي الأدوات المستخدمة على فقرات ذات أداء تفاضلي حسب المستوى الدراسي للطلبة (الأولى والثانية، والثالثة، والرابعة وأكثر) في حدود علم الباحث.

التعريفات الإجرائية للدراسة:

-**الأداء التّفاضلي للفقرة:** وهو دلالة ومشتقة إحصائياً للتعبير عن الفروق في الاستجابة للفقرة بين مجموعتين من المفحوصين ممّن هم في نفس المستوى من القدرة.

-**الطّلبة:** هم طلبة البكالوريوس الذين يدرسون المساقات المختلفة خلال الفصل الدّراسي الصيفي ٢٠١٥/٢٠١٦م ذكوراً وإناثاً في جامعة اليرموك.

-**المستوى الدراسي:** هي السنوات الدراسية (الأولى، والثانية، والثالثة، والرابعة فأكثر) خلال مرحلة البكالوريوس في جامعة اليرموك.

-**الأداء التدريسي:** هو كلّ ما يقوم به عضو هيئة التدريس من أنشطة وعمليات، وإجراءات، وسلوكيات تعليمية تتعلّق بعملية التدريس داخل قاعة التدريس أو خارجها.

-**أعضاء هيئة التّدريس:** هم الأساتذة الذين يقومون بتدريس المساقات المختلفة

في جامعة اليرموك بمختلف رتبهم الأكاديمية.

-طريقة مانتل-هانزل العامة: هي طريقة إحصائية لفحص الأداء التفاضلي سواء لفقرات ثنائية الاستجابة (C=2) أو لفقرات متعدّدة الاستجابة (C > 2) في مجموعات كثيرة (R >2) وبشكل مُتزامن.

محددات الدّراسة:

١. اقتصرت هذه الدراسة في الكشف عن الأداء التفاضلي للفقرات من وجهة نظر طلبة البكالوريوس لجامعة اليرموكخلال العام الدّراسي ٢٠١٥ / ٢٠١٦م.

٢. اقتصرت هذه الدراسة على الكشف عن الأداء التفاضلي للفقرات تبعاً لمتغير المستوى الدراسي في أداة تقييم الطلبة لأداء أعضاء هيئة التدريس في جامعة اليرموك.

٣. اقتصرت هذه الدراسة على استخدام طريقة واحدة في الكشف عن الأداء التفاضلي للفقرات (DIF) وهي طريقة مانتل- هانزل العامة.

الدِّراسات السّابقة:

أجرى كلا من راجو ودراسغووسلند [٢٠] (Raju, Drasgow&Slind, 1993) دراسة هدفت إلى مقارنة ثلاث طرق للكشف عن الأداء التّفاضلي للفقرة، هي: طريقة فرق المساحة بين منحنيات خصائص الفقرة، وطريقة كأي تربيع، وطريقة مانتل - هانزل. حيث استخدم الجانب اللفظي من اختبار جيتس – ماكجينيتي [٢١] (Gates-Macginitie Reading Test) المكوّن من (٤٥) فقرة لكل فقرة منها خمسة بدائل. تم تطبيقه على (٨٣٩) طالبًا وطالبة من الفئة العمرية (١٠-١٢) سنة. ويتوزعون على مجموعات فرعية وفقًا لمتغيري الجنس والعرق. أظهرت نتائج الدّراسة فعالية طريقة مانتل-هانزل في الكشف عن أكبر عدد من الفقرات المتحيزة، كما أظهرت الدّراسة وجود درجة عالية من الاتفاق بين نتائج تطبيق طريقة فرق المساحة بين منحنيات خصائص الفقرة ونتائج تطبيق طريقة كاي تربيع في الكشف عن الفقرات المتحيزة، وأظهرت النتائج وجود اتفاق عال بين طريقة فرق المساحة بين منحنيات خصائص الفقرة وطريقة كاي تربيع من جهة، وطريقة مانتل-هانزل من جهة أخرى فقط عند اعتماد الجنس كأساس للكشف عن الأداء التّفاضلي للفقرة.

قام تاقوموري [٢٢] (Tagomori,1993) بدراسة هدفت إلى تحليل محتوى الأدوات المستخدمة في تقويم أعضاء هيئة التّدريس والموجهة إلى الطّلبة، وشكّلت المعلومات التي تم الحصول عليها من هذه الدّراسة الأساس للبحث والتطوير المستقبلي في تقويم التّدريس. وأظهرت نتائج تّحليل المحتوى لعينة من (٤٠٢٨) فقرة من (٢٠٠) أداة

مستخدمة في تقويم الأداء التّدريسي أنّ(٤,٦ %) من الفقرات كانت مضلّلة، وغير واضحة، وغير موضوعية،وأنّ (٢٤,٥%) من الفقرات الأخرى لم ترتبط مع الأداء التّدريسي الذي تقيسه، وأنّ(٥٨%) من (٢٠٠) أداة احتوت على فقرات مضلّلة، وموجبة الالتواء أو سالبة الالتواء، وخلصت الدّراسة ضرورة مراجعة الأدوات التقويمية المستخدمة في تقويم الأداء التّدريسي الجامعي، لأنّ أكثر المشكلات هي في محتوى أدوات التقويم المستخدمة.

قارن زويكودونغو وغريما [٢٣] (Zwick, Donoghue &Grima, 1993) الخطأ من النوع الأول وقدرة طريقة مانتل-هانزلو(GMH) في تحديد الأداء التّفاضلي الموحد المتوازن متدني الحركة (low shift) (أي الأداء التّفاضلي المتأثر فقط عبر فئات الفقرة المتدنية) والأداء التّفاضلي عالي الحركة (أي الأداء التّفاضلي المتأثر فقط بفئات الفقرات المرتفعة)، كما درسوا أيضا أداء طريقة (STNDES.DIF) في إظهار حجم الأداء التّفاضلي من خلال دراسة الوسيط والانحراف المعياري للقيمة الإحصائية. وفي محاكاة للاختبار المكوّن من (٢٥) فقرة، كانت (٢٠) فقرة ثنائية التّدريج بينما كانت (٥) فقرات متعدّدة التّدريج مع أربعة مستويات من العلامة. والفقرة الخامسة والعشرين كانت هي الفقرة الخاضعة للدراسة. وحجم الأداء التّفاضلي، وتوزيع قدرة المجموعة، وطريقة المطابقة تنوعت. وكان حجم العينة (٥٠٠) لكل مجموعة. وقد وجدوا أن طريقة مانتل كانت ذات قدرة أعلى من طريقة (GMH) في تحديد الأداء التّفاضلي الموحد، بينما كان للطريقة الأخيرة قوة أعلى للأداء التّفاضلي المتوازن، وكانت طريقة (THE STNDES.DIF) هي الوحيدة الحساسة للأداء التّفاضلي المتوازن. وقد قادت الفروقات في قدرة المجموعة إلى زيادة طفيفة في معدلات الخطأ من النوع الأول لكافة الطرق. لكن متوسط معدل الخطأ من النوع الأول لكل من طريقة مانتلوGMH كان أدنى من (0.05) في كافة الظروف. كما وجدوا أيضا أن الفقرة الخاضعة للدراسة يجب أيضا تضمينها في اختبار التكافؤ لتقليص الخطأ من النوع الأول.

وقدم سبراي وميلر [٢٤] (Spray & Miller, 1994) بمحاكاة اختبار من ٢٠ فقرة لمقارنة القوة والخطأ من النوع الأول لكل من طرق (LDFA) ومانتل، و (GMH) في تحديد الأداء التّفاضلي للفقرة في الفقرات متعدّدة التّدريج، ضمن عينة صغيرة (٥٠٠) وعينة كبيرة (٢٠٠٠). وكانت كافة الفقرات العشرين متعدّدة التّدريج، وكانت الفقرة الأخيرة هي الفقرة المدروسة، وتمت محاكاة الأداء التّفاضلي المتوازن والمنتظم وغير المنتظم لاختبار القوة. وكان معدل الخطأ من النوع الأول

لطرق التحقق من الأداء التّفاضلي قد تم تحديده بعدد مرات ملاحظة اختبار ذي دلالة للفقرات (١-١٩)، (لا يوجد أداء تفاضلي)، وكان معدل الطرق الثلاثة متدنياً ضمن ظروف الدّراسة هذه، وكان لطريقة (LDFA) قوة أعلى مقارنة بالطريقتين الأخريين في تحديد الأداء التّفاضلي المنتظم وغير المنتظم، كما كان لها قوة أعلى في تحديد الأداء التّفاضلي المتوازن في العينات الصغيرة.

قام جرادات (٢٠٠٣) بدراسة هدفت إلى مقارنة بين طريقتي مانتل-هانزل وطريقة الصعوبة المحولة في الكشف عن تحيز الفقرات تبعاًلمتغيري الجنس والحقل الأكاديمي، ومدى الاتفاق بين الطريقتين. وقد استخدم الباحث اختبار كاليفورنيا لمهارات التفكير الناقد الذي تم تعريبه. تكونت عينة الدراسة من (٩٩٦) طالباً و(٧٨٩) طالبة. استخدم الباحث معامل الارتباط فاي وكاي تربيع لاختبار درجة اتفاق الطريقتين والفروق بين نسب الاتفاق. وقد أظهرت نتائج الدراسة أنّ طريقة الصعوبة المحولة كانت أفضل من طريقة مانتل-هانزل في الكشف عن الفقرات المتحيزة لمتغير الجنس، وكانت نسبة الفقرات المتحيزة (٣٩%)، (١٨%) على الترتيب، وأظهرت الدراسة أنّ هناك اتفاقًا بين الطريقتين في الكشف عن الفقرات المتحيزة، وكانت نسبة الاتفاق (٥٩%) للجنس، و(٦٥%) للحقل الأكاديمي.

وهدفت دراسة باي[٣٥] (Pae, 2004) إلى الكشف عن الأداء التّفاضلي في فقرات الجزء الإنجليزي من اختبار القبول الكوري للعام (١٩٩٨) تبعاً لمتغير المجال الأكاديمي (علوم إنسانية، علوم) وذلك باستخدام نسبة الأرجحية وطريقة مانتل - هانزل، حيث يضم الجزء الإنجليزي في هذا الاختبار (٥٥) فقرة منها (١٧) فقرة تقيس مهارة الاستماع و(٣٨) فقرة تقيس مهارة استيعاب المقروء، وطبق الاختبار على (٨٣٩٨٣٧) مفحوصًا، اختار الباحث منهم عينةً عشوائيةً مكونة من(١٤٠٠٠) مفحوص موزّعين بالتساوي على المجالين الأكاديميين، وأظهرت نتائج الدّراسة وجود (١٨) فقرة ذات أداء تفاضلي، فيها (٧) فقرات كانت أسهل للعلوم الإنسانية، وتم تحليل محتوى الفقرات التي أظهرت أداء تفاضليًا لمتغير المجال الأكاديمي في محاولة لتفسير نتائج الدّراسة.

وقام أوتيروجكوفالن[٣٦] (Uiterwijk&Vallen, 2005) بدراسة هدفت إلى الكشف عن مصادر تحيز البنية اللّغوية لفقرات اختبار التحصيل النهائي الخاص بالطّلبة المهاجرين إلى ألمانيا، ويقيس هذا الاختبار ثلاثة مجالات فرعية، هي: اللغة، والرياضيات، ومعالجة المعلومات. وأخذ (٥٤٠) فقرة من نوع

الاختيار من متعدد، موزعة بالتساوي على كل مجال، واستخدمت طريقة مانتل-هانزل، وطريقة الإجراءات التّحكيمية للكشف عن مصادر التّحيز المختلفة في الفقرات، وطبق الاختبار على عينة من الطّلبة تبعاً لمتغير العرق، وبينت نتائج الدّراسة وجود(٤٨) فقرة متحيزة لصالح الطّلبة الألمان الأصل و(٣٠) فقرة غير متحيزة، وذلك حسب طريقة مانتل - هانزل، وبينت الدّراسة أيضًا أنّ من أهم مصادر تحيز الاختبار: المكونات اللغوية للفقرات، بالإضافة إلى مهارات اللغة ومؤشرات الإجابة الصحيحة التي تتضمنها صيغ الفقرات المتحيزة، والاختلاف في مهارات الرياضيات، ووجود بعض الفقرات التي تحتوي على عناصر تتطلّب معرفة بالثقافة الألمانية، وأظهر أغلبية المحكّمين أنّ الفقرات المتعلقة بالرياضيات تتضمّن نصوصًاتتوسطها الأرقام قد تكون مصدرا للتّحيز.

أما دراسة عنابي ودوديين[٢٧] (Innabi&Dodeen, 2006) فقد هدفت إلى تحليل الفقرات التي تظهر أداءَ تفاضليا لمتغير الجنس في اختبار الرياضيات في الدراسة الدولية وتوجهات في الرياضيات والعلوم (Timss,١٩٩٠) في الأردن، حيث تكونت العينة من (٥٢٩٩) طالباً و(٢٨٢٩) طالبة من طلبة الصف الثامن الأساسي، وقد طبق الباحثان طريقة مانتل (٣٧) فقرة قد أظهرت أداءَ تفاضلياً بنسبة (٣٠%) من فقرات الاختبار، منها (١٧) فقرة لصالح الذكور و(٢٠) فقرة لصالح الإناث، وقد كانت أغلب الفقرات التي أظهرت أداءَ تفاضليًا لصالح الإناث تنتمي لموضوع الجبر والإحصاء، أما الفقرات التي كانت سلبية باتجاه الإناث تنتمي لموضوع الجبر والإحصاء، أما الفقرات التي كانت سلبية باتجاه الإناث فقد كان من موضوعاتها ما يتطلب المخاطرة في الحلّ، مثل موضوعات التقدير، التوقع، التقريب. وقد كانت أغلب الفقرات التي أظهرت أداءَ تفاضليًا لصالح الذكور تنتمي لموضوع القياس.

وهدفت دراسة الصباطي وشحتة (2007) إلى الكشف عن تأثير بعض المتغيرات في تقييم الطّلبة لأداء عضو هيئه التّدريس في الجامعة. وتكونت الدّراسة من (٤٢٧) طالباً وطالبة طبّقت عليهم أداة لأهم العوامل المؤثرة في تقييم الطّلبة لأداء عضو هيئة التّدريس في الجامعة. وهو من إعداد الباحثين بعد التحقق من صدقه وثباته. وقد تكونت الأداة من (٦٩) فقرة في شكلها النهائي تمثل تسعه أبعاد أساسية تؤثر في تقييم الطّلبة لأداء عضو هيئه التّدريس في الجامعة، وقد تم استخدام النموذج الخطي العام، واختبار (ت)، وتحليل التغاير متعدد المتغيرات التابعة (MANCOVA)، واختبار توكي، لدراسة تأثير كل من الجنس والتّخصص الأكاديمي والمستوى الدّراسي

والتفاعلات الثنائية والتفاعل الثلاثي بينها على درجات أبعاد أداة تقييم أداء عضو هيئة التّدريس في الجامعة من وجهه نظر الطّلبة. وقد دلّت النتائج على وجود اختلاف في الأبعاد المكوّنة لأداة تقييم الأداء الجامعي باختلاف الجنس والتّخصص الأكاديمي والمستوى الدّراسي، وقد أوصت الدّراسة بضرورة إعداد دورة لأعضاء هيئة التّدريس في الجامعة لتبصيرهم بالنواحي الإيجابية بواسطة أبعاد الأداة وكيفية تلافي نقاط الضعف، والنّواحي السلبية في أدائهم، ومحاولة إقناعهم بالدور الايجابي لمشاركة الطّلبة في عمليه التقييم، بالإضافة إلى التقييم الذاتي، وتقييم الزملاء والإداريين، ومن ثم تحسين الأداء التّدريسي الجامعي.

وفي دراسة قام بها حمادنه (2007) هدفت إلى مقارنة الأداء التفاضلي لفقرات اختبارات تحديد الكفاءة اللغوية في اللغة الانجليزية في الجامعات الأردنية (العلوم والتكنولوجيا، اليرموك، آل البيت) تبعاً لمتغيري الجنس وفرع الثانوية العامة، وفق طريقتي المساحة بين منحنيات خصائص الفقرة في النموذج اللوجستي ثلاثي المعلم ومانتل-هانزل. تم اختيار ثلاثة نماذج من اختبارات تحديد الكفاءة اللغوية، وجميعها من نوع الاختيار من متعدّد، وبلغ عدد الفقرات في الجامعات الثلاث (١٠٠،٥٠،٥٠) فقرة، إذإنها تقيس مهارات مشتركة في اللغة الانجليزية (الكتابة، استيعاب المقروء، القواعد والمفردات). تكوّنت عينة الدراسة من (١٩٣٥) طالباً وطالبة، وقد أظهرت نتائج الدراسة أنّ أكثر عدد من الفقرات التي أظهرت أداءَ تفاضلياً لمتغير الجنس كان في اختبار جامعة اليرموك، ولمتغير فرع الثانوية العامة في اختبار جامعة العلوم والتكنولوجيا. وأكثر المهارات التي أظهرت فقراتها أداءَ تفاضليا لمتغير الجنس هي مهارة القواعد والمفردات وكانت معظمها لصالح الإناث، ولمتغير فرع الثانوية العامة مهارة استيعاب المقروء، وكانت في معظمها لصالح الفرع العلمي.

وأجرى مبارك (2010) دراسة هدفت إلى الكشف عن الأداء التفاضلي في فقرات العلوم في الدراسة الدولية بيزا (٢٠٠٦)، تبعاً لمتغيرات الدولة واللغة والانتماء لمنظمة التنمية والتعاون الاقتصادي، واستخدم طريقتي مانتل-هانزل ونسبة الأرجحية. وبلغ عدد الفقرات في اختبار العلوم (٩٧) فقرة، وتكوّنت عينة الدراسة من (٤٠٠،٠٠٠) طالب وطالبة تم اختيارهم بطريقة عشوائية من مجتمع الدراسة. وقد أظهرت النتائج باستخدام الطريقتين (نسبة الأرجحية، مانتل-هانزل) وجود أداء تفاضلي، حيث كانت جميع الفقرات التي أظهرت أداءَ تفاضلياً لمتغير الدولة (الأردن مقابل غيرها) لصالح

مجموعة الدول الأخرى، ولمتغير اللغة (العربية مقابل الانجليزية) لصالح اللغة الانجليزية، ولمتغير الانتماء لمنظمة التنمية والتعاون الاقتصادي (تنتمي مقابل لا تنتمي) لصالح مجموعة الدول المنتمية للمنظمة. وقد وجد الباحث أن متوسط الأداء الأردني وترتيبه قد تحسن بعد حذف الفقرات التي أظهرت أداءً تفاضلياً، ولكنه بقي دون المتوسط الحسابي العام.

وهدفت دراسة الخرابشة و الجعافرة و الهباهبة و السعايدة(2012) إلى تعرف العوامل المؤثرة في تقويم أداء أعضاء هيئة التّدريس في كلية الأميرة عالية الجامعية بجامعة البلقاء التطبيقية الرسمية في الأردن من وجهة نظر طالبات الكلية. تكون مجتمع الدّراسة من طالبات البكالوريوس في الكلية وقد بلغ عددهن في العام الجامعي ٢٠١٠/٢٠١١م، (٢٦٨٦) طالبة موزعات على أربعة أقسام أكاديمية وسبعة تخصصات، واختار الباحثون عينة عشوائية طبقية من الأقسام جميعا شكلت نسبتها (9%) من مجتمع الدّراسة وانتهت العينة إلى (٢٢٩) طالبة. طوّر الباحثون أداة دراسة تضم (٥٣) عاملًا لاختبار أثرها في تقييم الطّالبات لأداء أعضاء هيئة التّدريس في كلية الأميرة عالية الجامعية، وتوصل الباحثون إلى (٦) عوامل منها كان تأثيرها بدرجة عالية جداً، و(٤٥) منها بدرجة عالية وعاملين فقط بدرجة متوسطة. وتوصلت الدّراسة إلى أن مجالات الدّراسة جميعها تؤثر بدرجة عالية في تقييم الطّالبات لأداء أعضاء هيئة التّدريس في الكلية، وجاء مجال الصفات الشخصية للمدرّس بالمرتبة الأولى، في حين جاء مجال علاقة المدرّس بالطّالبات والامتحانات في المرتبة الثّانية، وجاء مجال المهارات التّدريسية في المرتبة الرابعة وكانت درجة تأثيره عالية، وجاء مجال الواجبات الدّراسية بالمرتبة الخامسة، وجميعها كانت دالّة إحصائياً. وتوصلت الدّراسة إلى أنّه لا توجد فروق ذات دلاله إحصائية في تقييم طالبات الكلية لأداء أعضاء هيئه التّدريس تعزى لمتغيري: القسم الأكاديمي والمستوى الدّراسي وفي جميع مجالات الدّراسة.

وهدفت دراسة كركايا [٢٨](Karakaya, 2012) إلى الكشف عن الفقرات التي تظهر أداءَ تفاضليا تبعاً لمتغير الجنس في الاختبارات الفرعية في العلوم والتكنولوجيا والرياضيات والتي تم تطبيقها على الصفوف السادس والسابع والثامن لعام (٢٠٠٩)، ضمن اختبارات تحديد المستوى (LDE)، حيث يتكوّن الاختبار من خمسة مجالات فرعية (اللّغة التركيّة، الرياضيات، العلوم والتكنولوجيا، العلوم الاجتماعية، اللغات)، ويهدف الاختبار إلى معرفة مدى امتلاك الطلبة للمعارف والتحصيل العلمي في البرامج التعليمية في كلّ مستوى من الصفوف الأساسية، وتستخدم هذه الاختبارات لغايات تحديد الطلاب الذين

سوف يلتحقون بالمرحلة الثانوية. تم تطبيق الاختبار على عينة مكونة من ٢٢٦٢٤ طالبًا وطالبة، تم اختيارهم بصورة عشوائية، واستخدم الباحث طريقة مانتل-هانزل للكشف عن الفقرات ذات الأداء التفاضلي. أظهرت نتائج الدراسة أنّ هناك ثلاث فقرات ذات أداء تفاضلي لصالح الإناث واثنتين لصالح الذكور في اختبار العلوم والتكنولوجيا، وثلاث فقرات لصالح الذكور، وفقرة واحدة لصالح الإناث في اختبار الرياضيات. وتم عرض الفقرات ذات الأداء التفاضلي على خبراء في القياس لمعرفة ما إذا كانت هذه الفقرات تظهر تحيزًا قائمًا على الجنس. حيث أشار الخبراء إلى أنّ هذه الفقرات لا تظهر أي تحيز بالنسبة لمتغير الجنس. وأجرى إرحيل (2013) دراسة هدفت إلى الكشف عن أثر وجود الفقرات ذات الأداء التّفاضلي للجنس على دقة المعادلة العمودية لاختبار أوتيس لينون للقدرة الفعلية بمستوييه المتوسط والمتقدم، حيث تألفت عينه الدّراسة من (١٠٠١) طالب وطالبة من طلبة الصف السابع والثامن والتّاسع، مقسمين إلى (500) في مستوى القدرة المتوسط، و(٥٠١) في مستوى القدرة المتقدم، واستخدم في هذه الدّراسة طريقة مانتل-هانزل (MH) للكشف عن الفقرات ذات الأداء التّفاضلي في كلّ مستوى من مستويات الاختبار بشكل منفصل، وبعد ذلك تم إجراء المعادلة العمودية باستخدام برمجية (MG-Bilog) وفقاً للنموذج أحادي المعلمة، وأجريت المعادلة في ضوء الاحتفاظ بالفقرات جري ذات الأداء التّفاضلي، وفي ظل حذ ف الفقرات ذات الأداء التّفاضلي. وأشارت النتائج إلى وجود أثر للأداء التّفاضلي على دقة المعادلة العمودية، وقام الجعافرة (2013) بدراسة هدفت إلى تقييم فاعلية أداء أعضاء هيئة التّدريس في جامعة الزرقاء من وجهة نظر طلبتها في ضوء معايير جودة التّعليم، ولتحقيق هذا الهدف جرى بناء استبانة تألفت من (٥٠) فقرة، موزعة على خمسة مجالات رئيسة: التّخطيط للتّدريس، وتنفيذ التّدريس، والاتصال والتّواصل، والسّمات الشخصية، وتقويم تعلم الطّلبة. وبعد تحقق الصدق والثبات اللازمين، تم تطبيق الأداء على عينة الدّراسة المكونة من (٩١٠) طالب، بواقع (٦٠٠ طالب و٣١٠ طالبات)، من طلبة السنوات الثّانية والثّالثة والرابعة في الكليات الإنسانية والعلمية، أظهرت نتائج الدّراسة أنّ مستوى تقييم الطّلبة لأداء أعضاء هيئة التّدريس يساوي (٣,٥٧) وهو في المستوى المرتفع، وجاءت المتوسطات الحسابية للمجالات جميعها في المستوى المرتفع، عدا مجال "تقويم تعلم الطّلبة"، في المستوى المتوسط، وكشفت نتائج الدّراسة

عن وجود فروق ذات دلالة إحصائية تعزى لمتغير الكلية لصالح طلبة الكليات العلمية، وعدم وجود فروق ذات دلالة إحصائية تعزى لمتغير النوع الاجتماعي، ووجود فروق ذات دلالة إحصائية تعزى لمتغير المستوى الدّراسي على المستوى الكلي، لصالح طلبة السنة الثّانية ثم الرابعة وأخيراً الثّالثة، وكانت دراسة مبارك ومبارك[٢٩] (2013) تهدف إلى الكشف عن الأداء التّفاضلي في اختبار الرياضيات في الدّراسة الدولية بيزا (٢٠٠٦) تبعاً لمتغير اللغة باستخدام طريقة مانتل-هانزل. وبلغ عدد فقرات اختبار الرياضيات (٤٨) فقرة، وتقدم للاختبار(٤٠٠٠٠٠) مفحوص، جرى اختيارهم عشوائيًا من مجتمع الدّراسة. وبينت الدّراسة وجود(٢٧) فقرة أظهرت أداءً تفاضليًا لمتغير اللغة (العربية مقابل الإنجليزية) حسب طريقة مانتل-هانزل، أي ما نسبته (٥٦%)، والفقرات التي أظهرت أداءً تفاضليا كان (٧٤%) منها لصالح مجموعة اللغة الإنجليزية. وبينت الدّراسة أيضًا أن متوسط الأداء في الأردن وترتيبه قد تحسن بعد حذف الفقرات التي أظهرت أداءً تفاضليًا، والسبب قد يعود إلى أن فقرات الاختبار الأصلية كانت مكتوبة باللغة الإنجليزية، وترجمة تلك الفقرات ربما يغيّر معناها أو يؤدي إلى حذف معلومات ضرورية تجعل الاستجابة عليها صعبة، وفي دراسة عبيدات[٣٠] (2014) تم استقصاء الأداء التفاضلي لفقرات اختبار من نوع الاختيار من متعدّد تبعاً لدرجة مخاطرة المفحوصين، وذلك وفق طريقة راجو المطوّرة لحساب المساحة بين منحنيات خصائص الفقرة. وقد تم بناء اختبار تحصيلي مكوّن من (٤٥) فقرة لكل فقرة أربعة بدائل في مبحث الرياضيات للصف الأول ثانوي العلمي، وكان من ضمنها (١٠) فقرات مخاطرة لم تتضمّن البدائل الموضوعة لها أي إجابة صحيحة. تكونت العينة من (٢٠٠٨) طالب وطالبة من الصف الأول الثانوي العلمي في مدارس لواء قصبة إربد. تحت تعليمات التصحيح من أثر التخمين، تم استخدام فقرات المخاطرة في عملية تصنيف الأفراد المفحوصين إلى مجموعتين؛ مجموعة الطلبة منخفضي المخاطرة كمجموعة مرجعية، ومجموعة الطلبة مرتفعي المخاطرة كمجموعة مستهدفة. أظهرت النتائج وجود أداء تفاضلي للفقرات ما نسبته (٦٧%)، حيث كان (١٤%) من النوع المنتظم و(٥٣%) من النوع غير المنتظم. وعزا الباحث ظهور الأداء التفاضلي في الفقرات إلى اختلاف الأفراد المفحوصين في التفاعل مع تعليمات الاختبار تبعاً لاختلاف درجة مخاطرتهم، وبحسب بنية كل فقرة من فقرات الاختبار وخصوصيتها من حيث المهارة المطلوبة لحلّها. مما أوجد أنماطاً مختلفة من الاستجابات على الفقرات لدى

مجموعة الدراسة، وبالتالي كانت معالم الفقرات متباينة بين مجموعتي الدراسة؛ مما أدى إلى ظهور أداء تفاضلي في تلك الفقرات.

وفي دراسة العجلوني والسوالمة[٣١] 2016. هدفت لفحص الأداء التفاضلي لمموهات فقرات اختبار الرياضيات في التقييم الوطني لاقتصاد المعرفةحسب الجنس ومنطقة المدرسة. ولأغراض الدراسة تمّ استخدام البيانات التي وفرها المركز الوطني لتنمية الموارد البشرية. وتتضمن البيانات استجابات ١٥٥٨ طالباً وطالبة في الصف الخامس الأساسي على ٢٦ فقرة اختيار من متعدد تقيس مهارات أساسية في الرياضيات. وباستخدام طريقة نسبة الأرجحية تمّ الكشف عن الأداء التفاضلي لفقرات الاختبار ومموهات فقراته، **وكان من أبرز نتائج الدراسة:**

- تمّ الكشف عن ١٥ مموهاً ذا أداء تفاضلي للجنس؛ أي ما نسبته ١٩% من المموهات. وكان الأداء التفاضلي للمموهات لصالح الإناث في تسعة مموهات ولصالح الذكور في ستة مموهات. كما تمّ الكشف عن ١٥ مموهاً ذا أداء تفاضلي لموقع المدرسة؛ أي ما نسبته ١٩% من المموهات. وكان الأداء التفاضلي للمموهات لصالح طلبة مدارس المدن في تسعة مموهات ولصالح طلبة مدارس الريف في ستة مموهات.

- أجرى القرعان والكويتي[٣٢](2017) دراسة هدفت إلى تقصّي اثر حقل المعرفة (التخصص) للطالب كمصدر للأداء التفاضلي في الفقرات تقييم فعالية وجودة التعليم العالي من وجهت نظر الطالب ولتحقيق هدف الدراسة، تم الاستفادة من البيانات التي قامت جامعة الامام عبدالرحمن بن فيصل بجمعها لأغراض الحصول على الاعتماد الاكاديمي العام وقد تم تحليل (٣٦٤٥٩) استبانة في ثلاث كليات (التربية، والصحة، والهندسة) لاستخدام نموذج راش المعدل واظهرت النتائج وجود أربع فقرات أبدت أداءا تفاضلياً حسب الكلية، وأكد محتوى الفقرات احتمالية أن تكون تلك الفقرات متحيزة لكلية دون أخرى، كما بينت الدراسة انّ حذف الفقرات ذات أداء التفاضلي يسهم وبدالة إحصائية في تحسين صدق البناء للأداء.

-أجرى أبو شندي وكاظم (2018) دراسة هدفت الى معرفة الأداء التفاضلي لفقرات اختبار الرياضيات في برنامج التنمية المعرفية بسلطنة عمان تبعاً لمتغير الجنس باستخدام طريقتي مانتل هانزل ومنحنى خصائص الفقرة. ولتحقيق ذلك، تم الحصول على بيانات 27329 طالبة وطالبة من طلبة الصف السابع الاساسي الذين تقدموا لاختبار الرياضيات الذي يتكون من 25 سؤالاً في برنامج التنمية

المعرفية عام 2017 بواقع 13258 طالباً، و13801 طالبة. أظهرت النتائج أن طريقة منحنى خصائص الفقرة أكثر صرامة من طريقة مانتل هانزل في الكشف عن الاداء التفاضلي في اختبار الرياضيات باختلافمستويات الطلبة، حيث تراوحت بين ٤٨%إلى ٥٢ % في مانتل هانزل، وتراوحت بين ٥٦ % إلى ٨٠ % في طريقة منحنى خصائص الفقرة، وتراوحت نسبة التوافق بين المؤشرين بين ٣٦% إلى ٤٨% كما أظهرت النتائج تأثر عدد الفقرات التي أظهرت اداء تفاضلي لمتغير الجنس بحجم عينة المفحوصين، وعدم تأثر طريقة مانتل هانزل باختلاف مستوى قدرة المفحوصين. وفي ضوء هذه النتائج تم الخروج بعدد من التوصيات.

التّعقيب على الدّراسات السّابقة:

بعد الاطّلاع على الدراسات السّابقة التي أجريت لتقييم الأداء التدريسي لأعضاء هيئة التدريس تبعاً لمتغير المستوى الدراسي، يلاحظ اهتمام بعض الدراسات السابقة بدراسة الفروق بين المتغيرات التصنيفية، كالمستوى الدراسي وغيرها في تقييم الأستاذ الجامعي، ولم تهتم بالبنية العاملية بشكل عام لأداة القياس المستخدمة في تقييم أداء عضو هيئة التدريس، وبعد الاطّلاع أيضاً على الدراسات السّابقة التي أجريت للكشف عن الفقرات ذات الأداء التفاضلي تبعاً لمتغير المستوى الدراسي، تكشف مراجعة الدراسات التربويّة العربيّة في هذا المجال عن عدم تناولها للأداء التفاضلي للفقرة تبعاً لمتغير المستوى الدراسي سواء من حيث كون ذلك وسيلة للكشف عن الفقرات ذات الأداء التفاضلي للفقرة أم غا ية بحدّ ذاته، وبالرغم من كثرة الدراسات العربية التي كشفت عن الأداء التفاضلي للفقرة وباستخدام طرق متعدّدة، إلا أنه لا توجد أي دراسة عربية استخدمت طريقة مانتل- هانزل العامة للكشف عن الأداء التفاضلي للفقرات تبعاً لمتغير المستوى الدراسي التي سوف تستخدم في هذه الدراسة، أمّا بالنسبة للدراسات الأجنبية التي تطرّقت للأداء التفاضلي للفقرة فهي كثيرة جداً، إلا أنّ الدراسات التي اهتمت في الأداء التفاضلي تبعاً لمتغير المستوى الدراسي تعدّ قليلة جدًا إذ ما قورنت بالدراسات التي تناولت الأداء التفاضلي للفقرة، وقد استخدمت بعض هذه الدراسات طريقة واحدة للكشف عن الأداء التفاضلي للفقرات، تبعاً لمتغير المستوى الدراسي وبعضها الآخر استخدمت عدة طرق وتم المقارنة بين نتائج الطرق .لذلك ما يميز هذه الدراسة أنها سوف تكشف عن الأداء التفاضلي للفقرات في أداة تقييم الطلبة لأداء أعضاء الهيئة التدريسية في جامعة اليرموك باستخدام طريقة مانتل- هانزل العامة، تبعاً لمتغير المستوى الدراسي، وسوف

تحاول معرفة الأسباب الكامنة وراء ذلك التحيز إن وجد.

الطريقة والإجراءات:

منهج الدّراسة:

جرى استخدام المنهج الوصفي التحليلي للكشف عن الفقرات التي تُظهر أداءً تفاضليًا تبعاً لمتغير المستوى الدراسي في اداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك، بالإضافة إلى الكشف عن أثرها على مؤشرات حُسن المطابقة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك المستخدمة في هذه الدّراسة؛ وذلك لمناسبة وطبيعة هذه الدّراسة وأهدافها.

بيانات الدراسة:

توفرت بيانات الدراسة الحاليةبالحصول على استجابات الطلبة على أداة تقييم الاداء التدريسي- للمدرس في جامعة اليرموك, حيث بلغ عدد الاستجابات (١٩٤٠١) لطلبة جامعة اليرموك, ثم تم تحديد بيانات المستوى الدراسي للطالب في جامعة اليرموك، وبعد تنقيح البيانات من الاستجابات النمطية (أي الاستجابة الواحدة على جميع فقرات الأداة من قبل كل طالب أو طالبة) من العام الدّراسي ٢٠١٥/ ٢٠١٦م، بلغ عدد الاستجابات (٧٣٤٣) استجابة في جامعة اليرموك التي سوف يتم عليها التحليل، ويبين الجدول (1) عدد الاستجابات التي تم الحصول عليها تبعاً للمستوى الدراسي في جامعة اليرموك.

جدول١: توزيع عينة الدّراسة تبعاً للمستوى الدراسي في جامعة اليرموك

المستوى الدراسي	الذكور		الاناث		الكلي	
	#	%	#	%	#	%
الأول	490	6.7%	865	11.8%	**1355**	%18.5
الثاني	620	8.4%	1202	16.4%	**1822**	%24.8
الثالث	586	8.0%	1036	14.1%	**1622**	%22.1
الرابع فأكثر	985	13.4%	1559	21.2%	**2544**	%34.6
الكلي	**2681**	**%36.5**	**4662**	**%63.5**	**7343**	**%100.0**

المستوى الدراسي	الذكور		الاناث		الكلي	
	#	%	#	%	#	%
الأول	490	6.7%	865	11.8%	1355	%18.5
الثاني	620	8.4%	1202	16.4%	1822	%24.8
الثالث	586	8.0%	1036	14.1%	1622	%22.1
الرابع فأكثر	985	13.4%	1559	21.2%	2544	%34.6
الكلي	2681	%36.5	4662	%63.5	7343	%100.0

وتم اختيار كافة البيانات المتوفرة للدراسة في جامعة اليرموك، تبعاً لمتغير للمستوى الدراسيلكون عدد استجابات الطلبة في مجتمع جامعة اليرموكضمن متطلبات برنامج (GMHDIF) Differential Item Functioning Generalized Mantel-Haenszel التي تشترط ألا يزيد عدد أفراد عينة الدّراسة عن (١٠٠٠٠) طالباً وطالبة.

وتم حساب الأوساط الحسابية والانحرافات المعياريةلاستجابات الطلبة على أداة تقييم الطّالب للأداء التّدريسي للمدرسفي جامعة اليرموك تبعاً للمستوى الدراسي يبين الجدول (2) ذلك:

جدول 2: الأوساط الحسابية والانحرافات المعياريةلاستجابات الطلبة على أداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي(المستوى١، والمستوى٢، والمستوى٣، والمستوى٤ فأكثر)

المستوى الدراسي	الوسط الحسابي	الانحراف المعياري
مستوى ١	4.01	0.904
مستوى ٢	3.94	0.924
مستوى ٣	3.89	0.963
مستوى٤ فأكثر	3.84	0.967

يظهر من الجدول (٢) وجود اختلاف ظاهري في قيم الأوساط الحسابية تبعاً للمستوى الدراسي.

أداة الدّراسة:

تم تبنّي أداة تقييم الطّالب للأداء التّدريسي للمدرس الذي جرى بنائها من قبل لجنة متخصصة في جامعة اليرموك والمكونة من (٢٠) فقرة، من نوع ليكرت الخماسي (أبداً، نادراً، أحياناً، غالباً، دائماً)، وتم إعطاء الفقرات الدرجات (١، ٢، ٣، ٤، ٥) على الترتيب حسب الظهور، وبذلك تراوحت الدرجة الكلية على الأداة من (٢٠) وحتى (١٠٠) درجة، الملحق (أ).

وتم التحقق من الصدق العاملي لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للجنس والمستوى الدراسي، حيث أدخلت بيانات العينة للمستوى الدراسي(الاول، والثاني، والثالث، والرابع فأكثر) على الحاسوب ليصار إلى تحليلها باستخدام برنامج (SPSS)، حيث جرى أولاً حساب معامل الارتباط المصحح (Corrected Item-Total Correlation) لكل فقرة من فقرات أداة تقييم الطّلبة الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك ككل ومن ثم تبعاًللمستوىالدراسي(الاول، والثاني، والثالث، والرابع واكثر)، وكذلك حُسب معامل ثبات الاتساق الداخلي كرونباخ ألفا (Cronbach Alpha) لفقرات الأداة. وثانياً تم تطبيق التّحليل العاملي (Factor Analysis) بطريقة المكونات الأساسية (Principal Component Analysis) على استجابات أفراد عينة من جامعة اليرموك على فقرات أداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي (الأول، والثاني، والثالث، والرابع وأكثر).

دلالات الصدق والثبات لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموكتبعاً للمستوى الدراسي:

جرى التّحقق من دلالات المؤشرات السيكومترية للأداة على النّحو الآتي:

صدق البناء:

للتحقق من صدق البناء لفقرات الأداة عن طريق الاتساق الداخلي والصدق العاملي والذي يمكن التوصل إليه باستخدام التّحليل العاملي.

وقد تم حساب معاملات الارتباط المُصحح للفقرات لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي، وذلك كما هو مبين في الجدول (3).

جدول3: معاملات الارتباط المُصحح للفقرات لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي.

الرقم	مضمون الفقرة	الارتباط المصحح بين الفقرة والأداة وفق المستوى الدراسي				الكلي
		الأول	الثاني	الثالث	الرابع فأكثر	
1	تمّ تحديد أهداف المساق بوضوح في خطة المساق	0.73	0.74	0.74	0.74	0.74
2	تمّ توضيح معايير تقييم الطلبة في المساق منذ البداية	0.78	0.77	0.76	0.77	0.77
3	يلتزم المُدرِّس بخطة المساق	0.76	0.76	0.76	0.77	0.76

الرقم	مضمون الفقرة	الارتباط المصحح بين الفقرة والأداة وفق المستوى الدراسي				الكلي
		الأول	الثاني	الثالث	الرابع فأكثر	
4	يُقدِّم المُدرِّس مادة المساق بأسلوب منظم تمكن الطالب من المتابعة	0.76	0.78	0.79	0.79	0.78
5	يلتزم المُدرِّس بحضور المحاضرات وبموعد بدئها وانتهائها	0.77	0.77	0.77	0.76	0.77
6	يحرص المُدرِّس على متابعة حضور الطلبة	0.76	0.78	0.79	0.82	0.79
7	يحرص المُدرِّس على السلوك والالتزام الجيد	0.76	0.75	0.78	0.80	0.78
8	يستجيب المُدرِّس لأسئلة الطلبة واقتراحاتهم	0.81	0.81	0.82	0.83	0.82
9	يتعامل المُدرِّس مع الطلبة بودٍّ واحترام	0.76	0.77	0.80	0.79	0.78
10	يُؤكِّد المُدرِّس على العدالة ويتجنب التحيز في تعامله مع الطلبة	0.78	0.82	0.81	0.82	0.81
11	يُشجِّع المُدرِّس الطلبة على التفكير الناقد	0.76	0.79	0.81	0.80	0.79
12	يستغل المُدرِّس وقت المحاضرات في تدريس مُنتج وفعَّال	0.77	0.77	0.76	0.79	0.77
13	يُظهر المُدرِّس مهارة عالية في مساعدة الطلبة على إتقان مادة المساق	0.80	0.81	0.84	0.83	0.82
14	الامتحانات والواجبات شاملة وعلى علاقة بما تم تدريسه	0.81	0.82	0.83	0.82	0.82
15	يلتزم المُدرِّس بتصحيح الامتحانات والتقارير والواجبات وإعادتها للطلبة	0.79	0.80	0.83	0.83	0.82

الرقم	مضمون الفقرة	الارتباط المصحح بين الفقرة والأداة وفق المستوى الدراسي				الكلي
		الأول	الثاني	الثالث	الرابع فأكثر	
16	يُناقش المُدرِّس إجابات الامتحانات والنتائج مع الطلبة	0.80	0.82	0.83	0.83	0.82
17	يحرص المُدرِّس على توفير بيئة صفية مريحة	0.80	0.82	0.82	0.82	0.82
18	يحرص المُدرِّس على التواجد في ساعات المكتب ويلتزم بمواعيده	0.65	0.70	0.70	0.75	0.71
19	يستثير المُدرِّس دافعية الطلبة للتعلم والتفاعل في النقاش الصفي	0.67	0.70	0.68	0.75	0.71
20	يُظهر المُدرِّس تمكناً وإلماماً بموضوعات وعناصر المساق	0.69	0.72	0.73	0.76	0.73
القيمة الصغرى		**0.70**	**0.70**	**0.75**	**0.71**	
القيمة العظمى						**0.65**

يُلاحظ من الجدول (3) أنَّ قيم معاملات ارتباط الفقرات المُصحح مع الأداة ككل قد تراوحت من (٠,٧١- ٠,٨٢) كما تراوحت قيم معاملات ارتباط الفقرات المُصحح مع الأداة من (٠,٦٥- ٠,٨١) وفقاً للمستوى الأول، فيما تراوحت قيم معاملات ارتباط الفقرات المُصحح مع الأداة من (٠,٧٠- ٠,٨٢) وفقا للمستوى الثاني، فيما تراوحت قيم معاملات ارتباط الفقرات المُصحح مع الأداة من (٠,٧٠- ٠,٨٣) وفقاً للمستوى الثالث، وأخيراً؛ تراوحت قيم معاملات ارتباط الفقرات المُصحح مع الأداة من (٠,٧٥- ٠,٨٢) وفقاً للمستوى الرابع فأكثر.

ثبات الأداة:

كما تم حساب ثبات أداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي باستخدام معادلة كرونباخ - ألفا α Cronbach's، وذلك كما هو مبين في الجدول 4.

جدول4: قيم معاملات ثبات الاتساق الداخلي لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي

الأداة وفق المستوى الدراسي	ثبات الاتساق الداخلي	عدد الفقرات
الأول	**.967**	**20**
الثاني	**.970**	**20**
الثالث	**.971**	**20**
الرابع فأكثر	**.974**	**20**
الكلي للأداة	**0.971**	**20**

يبين جدول (٤) إن ثبات الاتساق الداخلي لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموكتبعاً للمستوى الدراسي بلغت قيمته (٠,٩٧١) للأداة ككل، وتراوحت من (٠,٩٦٧-٠,٩٧٤) تبعاً للمستوى الدراسي.

وتم التحقق من الصدق العاملي لفقرات أداة تقييم الطلبة للأداء التدريسي في جامعة اليرموك باستخدام التّحليل العاملي (Factor Analysis) بطريقة المكونات الأساسية (Principal Component Analysis) على استجابات أفراد عينة المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر) في جامعة اليرموك على فقرات الأداة.

وتم إجراء التّحليل العاملي لاستجابات أفراد عينة طلبة المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر)، على فقرات أداة تقييم الطّالب للأداء التّدريسي في جامعة اليرموك، وتم الاقتصار على معرفة قيمة الجذر الكامن للعامل الأول والعامل الثاني للمستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر)، كما في الجدول (٥):

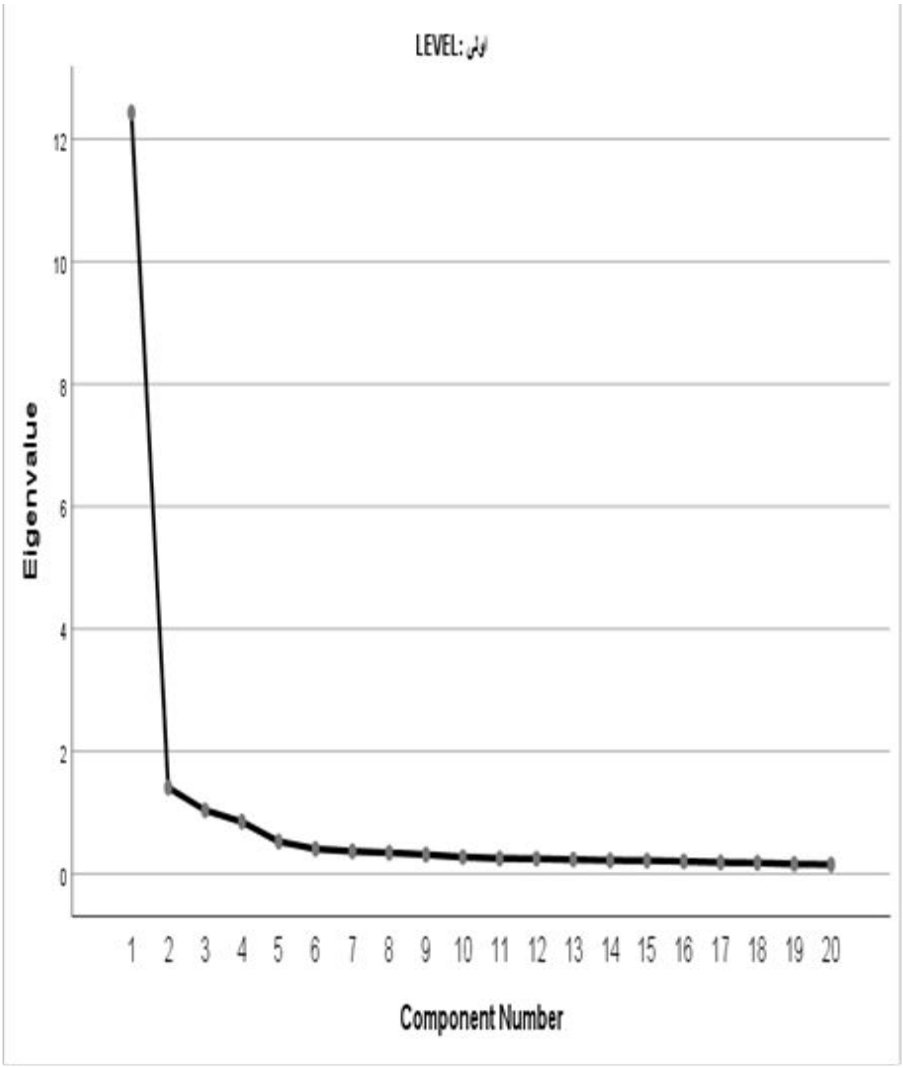
جدول 5 : التّحليل العاملي على فقرات أداة تقييم الطّالب للأداء التّدريسي- في جامعة اليرموك تبعاً للمستوى الدراسي مبينا قيمة الجذر الكامن للعامل الأول والعامل الثاني.

المستوى العامل التباين الكلي المفسر من (Total Variance Explained)

الدراسي		الجذور الكامنة المبدئية		
		الجذر الكامن	نسبة التباين المفسر%	نسبة التباين المفسر التراكمي%
الاول	1	12.433	62.166	62.166
	2	1.408	7.038	69.204
الثاني	1	12.786	63.928	63.928
	2	1.329	6.647	70.575
الثالث	1	13.028	65.142	65.142
	2	1.573	7.867	73.010
الرابع فأكثر	1	13.336	66.682	66.682
	2	1.498	7.489	74.171

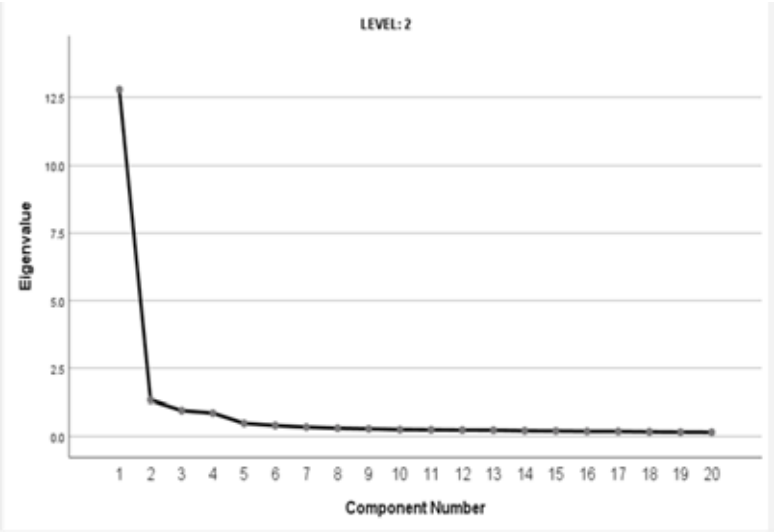
يتبين من الجدول (٥) أنّ قيمة الجذر الكامن للعامل الأول المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر) على الترتيب بلغت (12.433)، (12.786)، (13.028)، (13.336) ويُفسر ما نسبته (62.166%)، (63.928%)، (65.142%) (66.682 %)، على الترتيب من التّباين الكلي، وهي قيم مرتفعة إذا ما قورنت مع قيم الجذور الكامنة لبقية العوامل، أما قيمة الجذر الكامن للعامل الثّاني المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر) على الترتيب بلغت (1.408)، (1.329)،(1.573)، (1.498) ويفسر ما نسبته (7.038%)، (6.647%)، (7.867%) (7.489 %)، على الترتيب من التّباين الكلي، وتبعاً لما أورده لورد (lord, 1980) فإن أحادية البعد تتحقق إن كانت نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثّاني تزيد عن (2) وهذا الشرط متحقق وتساوي (8.401)، (9.621)،(8.282)،(8.902)على الترتيبب معنى أن العامل الأول فسر ما يزيد عن ضعفي ما فسره العامل الثّاني. أضف إلى ذلك، يلاحظ بأن نسبة التّباين المفسر- لكل من العوامل المتبقية متقاربة، بمعنى أنه يوجد شبه استقرار في نسب التّباين المفسر- لجميع العوامل باستثناء العامل الأول، وهذا مؤشر على تحقق افتراض أحادية البعد للأداة Hulin [٣٣] Drasgow& Parson, 1983; Hattie, 1985، أي أن الأداة يقيس سمة واحدة.

ويتعزز افتراض أحادي البعد، للأداة المعدّ بالدّراسة الحالية، بتمثيل الجذور الكامنة بيانياً باستخدام ما يعرف باختبار فحص العوامل (Scree plot) الذي يظهر في الشكل (1) والشكل(2) والشكل (3) والشكل(4).



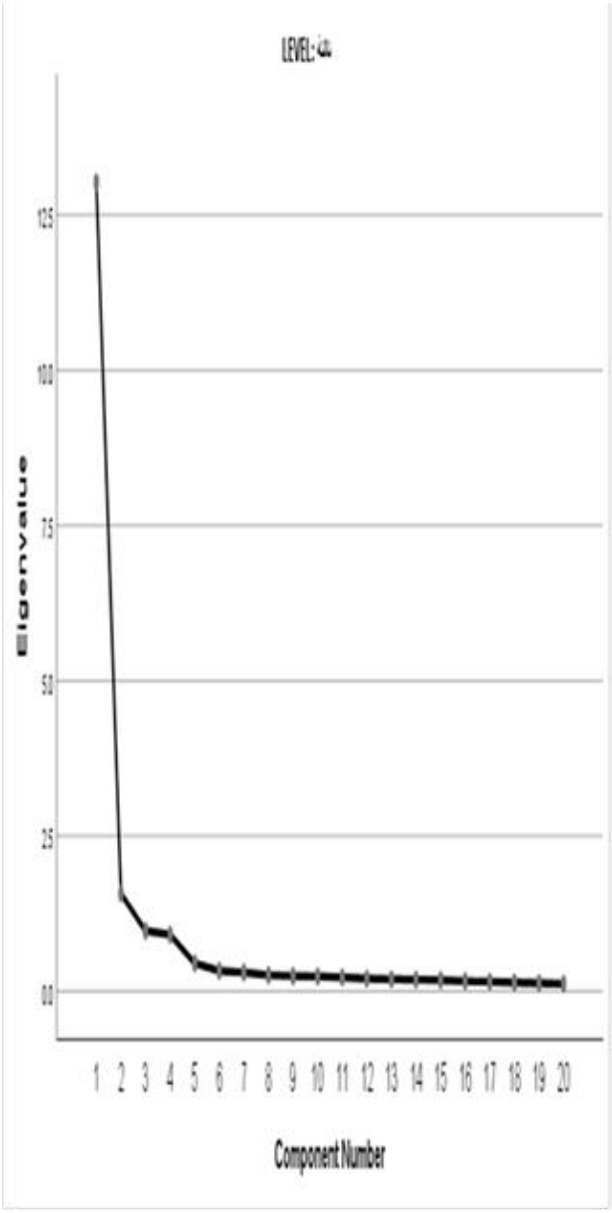
الشكل 1: التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك للمستوى الأول.

يتبين من الشكل (١)، أن الجذر الكامن للعامل الأول يطغى بشكل واضح على الجذور الكامنة لبقية العوامل على جميع المستويات الدراسية (الأول، والثاني، والثالث، والرابع فأكثر)، وهذا مؤشر أيضاً على أحادية البعد لبيانات المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر).



الشكل٢: التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة لأداة تقييم الطّالب لـلأداء التّدريسي للمدرس في جامعة اليرموك للمستوى الثاني.

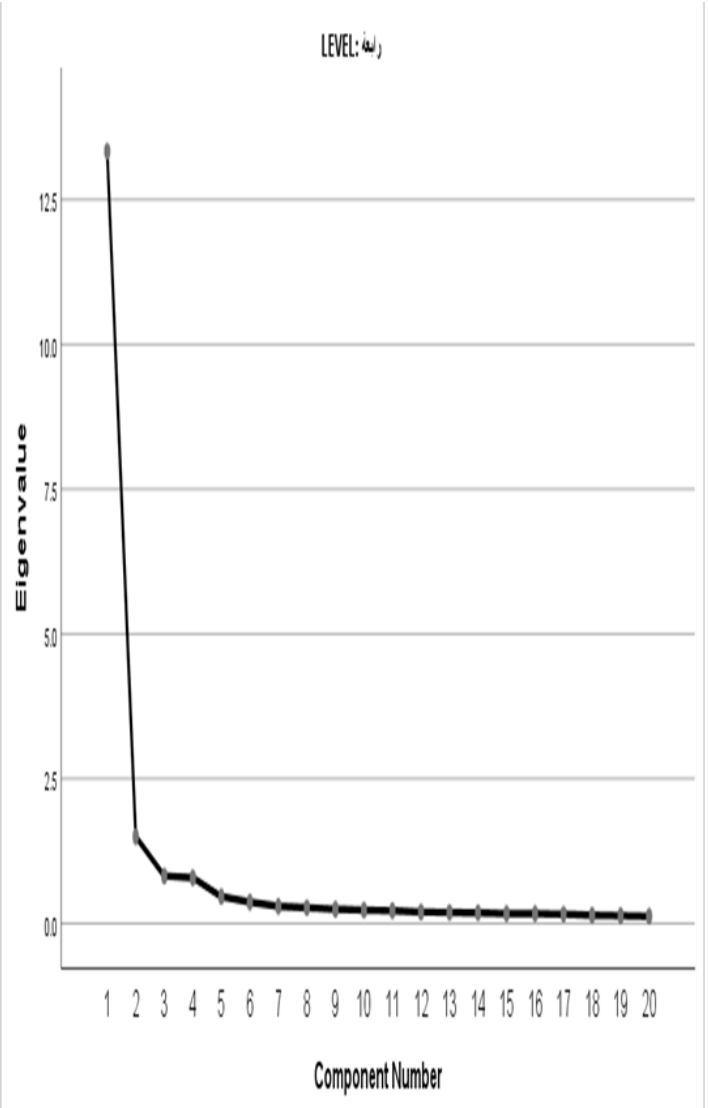
يتبين من الشكل (٢)، أن الجذر الكامن للعامل الأول يطغى بشكل واضح على الجذور الكامنة لبقية العوامل على جميع المستويات الدراسية (الأول، والثاني، والثالث، والرابع فأكثر)، وهذا مؤشر أيضاً على أحادية البعد لبيانات المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر).



الشكل ٣: التمثيل البياني لقيم الجذور الكامنة للعوامل المكونةلأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك للمستوى الثالث.

يتبين من الشكل (٣)، أن الجذر الكامن للعامل الأول يطغى بشكل واضح على الجذور الكامنة لبقية العوامل على جميع المستويات الدراسية (الأول، والثاني،

والثالث، والرابع فأكثر)، وهذا مؤشر أيضاً على أحادية البعد لبيانات المستوى الدراسي (الأول، والثاني، والثالث، والرابع فأكثر).



الشكل ٤: التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك للمستوى الرابع فأكثر.

يتبين من الشكل (٤)، أن الجذر الكامن للعامل الأول يطغى بشكل واضح على الجذور الكامنة لبقية العوامل على جميع المستويات الدراسية (الأول، والثاني، والثالث، والرابع فأكثر)، وهذا مؤشر أيضاً على أحادية البعد لبيانات المستوى الدراسي(الأول، والثاني، والثالث، والرابع فأكثر).

إجراءات الدّراسة:

نفذت الدّراسة وفقا للخطوات الآتية:

١- تبنِّي أداة تقييم الطّالب للأداء التّدريسي- للمدرس الذي تم بنائها من قبل مركز الاعتماد وضمان الجودة في جامعة اليرموك المكون من (٢٠) فقرة.

٢- التّحقق من الخصائص السيكومترية لأداة تقييم الطّالب للأداء التّدريسي- للمدرس في جامعة اليرموك.

٣-الحصول على كتاب تسهيل مهمة من عميد كلية التّربية إلى رئيس جامعة اليرموك.

٤-الحصول على كتاب تسهيل مهمة من رئيس جامعة اليرموك مُوّجه لمدير مركز الحاسب الالي في جامعة اليرموك بهدف تأمين البيانات اللازمة للباحث.

٥-تنقيح ملف بيانات جامعة اليرموك من استجابات الأفراد الثابتة (تأخذ تدريج واحد فقط) خلال كافة فقرات الأداة، بحيث أصبح العدد النهائي لمجتمع جامعة اليرموك (٧٣٤٣) من أصل (١٩٤٠١).

٦-إجراء المعالجات الإحصائية المناسبة باستخدام برنامج GMHDIF والرزمة الإحصائية للعلوم الاجتماعية (SPSS)، وبرنامج Excel وبرنامج (DIFAS 5.0) واختبار بونيفروني والتّحليل العاملي التوكيدي باستخدام برنامج (AMOS).

المعالجات الإحصائية:

للإجابة عن أسئلة الدّراسة، اُستخدمت المعالجات الإحصائية الآتية:

للإجابة عن السؤال الأول؛ تم استخدام طريقة مانتل-هانزل العامة للكشف عن الفقرات التي تظهر أداءَ تفاضلياً لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك، **تبعاً متغير المستوى الدراسي على مرحلتين:**

-**المرحلة الأولى:** وهي مرحله أولية تم فيها الكشف عن الفقرات ذات الأداء التّفاضلي، واعتمدت على جميع استجابات الطّلبة.

-**المرحلة الثّانية:** وهي مرحله نهائية (تنقيحية)هدفها الكشف عن الفقرات ذات الأداء التّفاضلي بشكل نهائي (بصورتها النهائية)، حيث تم إسقاط استجابات الطّلبة للفقرات ذات الأداء

التّفاضلي، واعتمدت على استجابات الفقرات الخالية من الأداء التّفاضلي.

١-أصبحت أداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك، تبعاً لمتغير المستوى الدراسي مُتحررةً من الفقرات التي تُبدي أداءً تفاضلياً.

وقام الباحث بحساب حجمالأداء التفاضلي للفقرات في أداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً لمتغير المستوى الدراسي، باستخدام برنامج (DIFAS5.0) (Penfield, 2012) من خلال حساب مؤشر[٣٤] (Liu-Agresti Cumulative Common Log-Odds Ratio: L-A LOR) (Liu and Agresti, 1996).

مؤشر L-ALOR يدل على اتجاه وحجم الأداء التفاضلي للفقرة بحيث:

-الاتجاه: إذا كانت (L -ALOR> 0) يكون DIFلصالح المجموعة المرجعية (الإناث مثلاً): إذا كانت (L -A LOR 0 >) يكون DIF لصالح المجموعة المستهدفة (الذكور مثلاً).

-الحجم: ويتم حساب حجم الأداء التفاضلي للفقرة من القيمة المطلقة لـ L A LOR-، والتأثيرات تتراوح من الناحية النظرية من -١ إلى +١. وغالباً ما تقيم هذه التأثيرات حسب المعايير المحددة للكشف عن الاداء التفاضلي:

1- 0.43 > | L -A LOR |≥0 لا يوجد تأثير للأداء تفاضلي(Non DIF).

2- 0.64≤ | L -A LOR | ≥ 0.43 يكون تأثير الأداء التفاضلي متوسط (FlaggedDIF).

3- 0.64 < | L -A LOR | يكون تأثير الأداء التفاضلي كبير(Large DIF). (Penfield, 2007).

ولتحديد مصدر الأداء التفاضلي للفقرات في أداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي؛ فقد تم إجراء اختبار بونفيرروني (Bonferroni) للمقارنات الثنائية البعدية المُتعددة.

٤- للإجابة عن السؤال الثّاني؛ فقد تم حساب مؤشرات حُسن المطابقة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً لمتغير المستوى الدراسي قبل الكشف عن الفقرات التي تُبدي أداءً تفاضليا وبعد إسقاط الفقرات التي تُبدي أداءً تفاضلياً، تم الكشف عن أثر الفقرات التي تُبدي أداءً تفاضليا على مؤشراتحُسن المطابقة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً لمتغير المستوى الدراسي وذلك باستخدام التّحليل العاملي التوكيدي برنامج أموس(AMOS).

عرض النّتائج:

أولاً: النتائج المتعلقة بسؤال الدّراسة الأول الذي نصّ على: **"هل تظهر فقرات أداة تقييمالطّلبة للأداء التّدريسي- للمدرس في جامعة اليرموك في عينة الدراسة أداء تفاضلياً تبعاً لمتغير المستوى الدراسي للطلبة؟"**.

للإجابة عن سؤال الدّراسة الأول؛ تم استخدام طريقة مانتل-هانزل العامة (Generalized Mantel-Haenszel Method) التي يوفرها برنامج (GMHDIF) للكشف عن الفقرات التي تُظهر أداء تفاضلياً في أداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً للمستوى الدراسي، **وذلك على النحو الآتي:**

للقيام بهذه المهمة يعمل البرنامج على مرحلتين، حيث يقوم في المرحلة الأولى بحساب مجموع درجات كل طالب وطالبة لكافة فقرات الأداة في جامعة اليرموك تبعاً للمستوى الدراسي الأكاديمية للطالب البالغ عددها (٢٠) فقرة بهدف تحديد مدى درجات الطّلبة على كافة فقرات الأداة، حيث تبين أنَّ توزع درجات الطّلبة على كافة فقرات الأداة قد تراوح من ٢١ وحتى ٩٩كما هو مبين في ملحق (2) وبالاعتماد على مدى درجات الطّلبة على كافة فقرات الأداة تم إنشاء ما مقداره [(٩٩-٢١)+١=٧٩] طبقة تمثل قدرات الطّلبة مع مراعاة إسقاط الطّلبة ذوي الدرجات التي لم يرصد عليها سوى طالب أو طالبة فقط إن وجدت حسب اشتراطات برنامج (GMHDIF)، ثم تم حساب قيمة Q_{MH} للكشف عن الفقرات التي تُظهر أداءً تفاضلياً في أداة تقييم الطلبة للأداءالتّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي، وذلك كما هو مبين في الجدول (6).

جدول6: نتائج المرحلة الأولى للكشف عن الفقرات التي تُبدي أداءً تفاضلياًعلى كافة فقرات أداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً للمستوى الدراسي:

الرقم	مضمون الفقرة	نتائج المرحلة الأولى		
		Q_{MH}	درجة الحرية	الدلالة الإحصائية
1	تمّ تحديد أهداف المساق بوضوح في خطة المساق	2.6340	3	0.4516
2	تمّ توضيح معايير تقييم الطّلبة في المساق منذ البداية	2.8623	3	0.4133
3	يلتزم المُدرّس بخطة المساق	6.1618	3	0.1040
4	يُقدِّم المُدرِّس مادة المساق بأسلوب منظم تمكن الطّالب من المتابعة	2.8229	3	0.4197
5	يلتزم المُدرِّس بحضور المحاضرات وبموعد بدئها وانتهائها	7.6450	3	0.0539
6	يحرص المُدرِّس على متابعة حضور الطّلبة	7.1847	3	0.0662
7	يحرص المُدرِّس على السلوك والالتزام الجيد	4.5506	3	0.2078
8	يستجيب المُدرِّس لأسئلة الطّلبة واقتراحاتهم	5.6321	3	0.1309
9	يتعامل المُدرِّس مع الطّلبة بودٍّ واحترام	4.7968	3	0.1873
10	يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة	*16.7750	3	*0.0008
11	يُشجِّع المُدرِّس الطّلبة على التفكير الناقد	*11.9380	3	*0.0076
12	يستغل المُدرِّس وقت المحاضرات في تدريس مُنتج وفعَّال	1.4076	3	0.7037
13	يُظهر المُدرِّس مهارة عالية في مساعدة الطّلبة على إتقان مادة المساق	9.5547	3	0.0228
14	الامتحانات والواجبات شاملة وعلى علاقة بما تم تدريسه	5.8428	3	0.1195
15	يلتزم المُدرِّس بتصحيح الامتحانات والتقارير والواجبات وإعادتها للطلبة	2.8135	3	0.4213
16	يُناقش المُدرِّس إجابات الامتحانات والنتائج مع الطّلبة	4.5147	3	0.2110
17	يحرص المُدرِّس على توفير بيئة صفية مريحة	2.7722	3	0.4281
18	يحرص المُدرِّس على التواجد في ساعات المكتب ويلتزم بمواعيده	2.0449	3	0.5631

الرقم	مضمون الفقرة	نتائج المرحلة الأولى		
		Q_{MH}	درجة الحرية	الدلالة الإحصائية
19	يستثير المُدرِّس دافعية الطّلبة للتعلم والتفاعل في النقاش الصفي	5.0356	3	0.1692
20	يُظهر المُدرِّس تمكنًا وإلمامًا بموضوعات وعناصر المساق	0.8442	3	0.8389

* دالة عند مستوى α=0.01

يُلاحظ من الجدول (٦) وجود فقرتان تُبدي أداءً تفاضلياً أوَّلياً في أداة تقييم الطلبة للأداء التدريسيفي جامعة اليرموك تبعاً للمستوى الدراسي وهي الفقرتان (١٠، و١١) اللتين نصتا على الترتيب على "يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة، يُشجِّع المُدرِّس الطّلبة على التفكير الناقد"، ولتنفيذ المرحلة الثّانية؛ فقد تم إسقاط استجابات الطّلبة على هذه الفقرات في أثناء حساب مجموع درجات كل طالب وطالبة لبقية فقرات أداة تقييم الطلبة للأداء التدريسي- في جامعة اليرموك تبعاً للمستوى الدراسي البالغ عددها (١٨) فقرات بهدف تحديد مدى درجات الطّلبة لبقية فقرات الأداة في جامعة اليرموك التي لا تبدي أداءً تفاضلياً أوّلياً تبعاً للمستوى الدراسي، حيث تبين أنَّ توزع درجات الطّلبة على بقية فقرات أداة تقييم الطلبة للأداء التدريسي- في جامعة اليرموك تبعاً للمستوى الدراسي قد تراوحت من ١٨ وحتى ٩٠ كما هو مبين في ملحق (٣) وبالاعتماد على مدى درجات الطّلبة على بقية فقرات الأداة تم إنشاء ما مقداره [(٩٠-١٨)+١=٧٣] طبقة تمثل قدرات الطّلبة مع مراعاة إسقاط الطّلبة ذوي الدرجات التي لم يرصد عليها سوى طالب أو طالبة فقط إن وجدت حسب اشتراطات برنامج (GMHDIF)، ثم تم الكشف عن الفقرات التي تُظهر أداء تفاضلياً بشكل نهائي من أصل (٢٠) فقرة في أداة تقييم الطلبة للأداء التدريسي في جامعة اليرموك تبعاً للمستوى الدراسي، وذلك كما هو مبين في الجدول (٧).

جدول٧: نتائج المرحلة الثّانية للكشف عن الفقرات التي تُبدي أداءً تفاضلياً على كافة فقرات أداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً للمستوى الدراسي:

الرقم	نتائج المرحلة الثّانية		
	Q_{MH}	درجة الحرية	الدلالة الإحصائية
1	2.8574	3	0.4141
2	1.6666	3	0.6444
3	7.4918	3	0.0578
4	1.4986	3	0.6826
5	7.0822	3	0.0693
6	6.0819	3	0.1077
7	3.7065	3	0.2949
8	3.8363	3	0.2797
9	7.1148	3	0.0683
10	*16.4050	3	*0.0009
11	10.0220	3	0.0184
12	1.9474	3	0.5834
13	9.7309	3	0.0210
14	5.6684	3	0.1289
15	3.2785	3	0.3507
16	5.5152	3	0.1377
17	2.9037	3	0.4067
18	2.0436	3	0.5634
19	5.7051	3	0.1269
20	0.4606	3	0.9275

*

يلاحظ من الجدول (٧) وجود فقرة واحدة تُبدي أداءً تفاضلياً بشكل نهائي في أداة تقييم الطلبة للأداء التدريسي في جامعة اليرموك تبعاً للمستوى الدراسي ذات الرقم (١٠) التي نصت على "يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة".

ويتبين من الجدول (٧) أنّ أكبر قيمه ل Q_{MH} للفقرة (١٠) وتنص على: "يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في

تعامله مع الطّلبة" تساوي (16.4050). علما بأن Q_{MH} للفقرات تتراوح بين(0.4060 -16.4050).

وتم حساب حجم الأداء التفاضلي للفقرات في أداة تقييم الطلبة للأداء التدريسي في جامعة اليرموك تبعاً للمستوى الدراسي، باستخدام برنامج (DIFAS 5.0) (Penfield, 2012) من خلال حساب مؤشر (Liu-AgrestiCumulative Common Log-Odds Ratio : L-A LOR)[٣٥] (Liu &Agresti, 1996)، وذلك كما هو مبين في الجدول (8).

جدول ٨: حجم الأداء التفاضلي للفقرات في أداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً للمستوى الدراسي

مضمون الفقرة	حجم الأداء التفاضلي للفقرات									
	الاول-الثاني		الاول-الثالث		الاول-الرابع فاكثر		الثاني-الثالث	الثاني-الرابع فاكثر		الثالث-الرابع فأكثر	
	L-A LOR	حجم الأداء التفاضلي	L-A LOR	حجم الأداء التفاضلي	L-A LOR	حجم الأداء التفاضلي	L-A LOR	حجم الأداء التفاضلي	L-A LOR	حجم الأداء التفاضلي
يُؤكّد المُدرّس على العدالة ويتجنب التّحيز في تعامله مع الطلبة	0.0-71	Non DIF	0.08	Non DIF	0.214	Non DIF	0.106	Non DIF	0.291	Non DIF	0.177	Non DIF

يتبين من الجدول (8) أن الفقرة(١٠) التي كانت ذات أداء تّفاضلي في مانتل - هانزل العامة، لا يوجد لها أثر ولجميع مقارناتها الثنائية تبعاً للمستوى الدراسي لأن قيمة حجم الأداء التفاضلي للفقرة و لجميع المقارنات الثنائية تساوي (0.177, 0.291, 0.106, 0.214, 0.08) وهي أقل من 0.43حسب مؤشر القيمة المطلقة لـ L A LOR-، وهي الفقرة التي نصت على "يُؤكّد المُدرّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة".ومن هنا يتبين أنّفقرة ذات الأداء التفاضلي لأداة تقييم الطّالب للأداء التّدريسي- للمدرس التي تم بناؤها من قبل لجنة متخصصة في الجامعة اليرموك لا يوجد لها أثر تبعاً للمستوى الدراسي.

ولتحديد المستوى الدراسي مصدر الأداء التفاضلي للفقرة (١٠) التي تنص على "يُؤكّد المُدرّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة" في أداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً للمستوى الدراسي؛ فقد تم إجراء اختبار بونفيرروني (Bonferroni) للمقارنات

الثنائية البعدية المُتعددة، وذلك كما هو مبين في الجدول (٩).

جدول٩: نتائج اختبار بونفيرروني (Bonferroni) للمقارنات الثنائية البعدية المُتعددة للفقرات التي أبدت أداءً تفاضلياً في أداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً للمستوى الدراسي:

مضمون الفقرة	نتائج اختبار Bonferroni للمقارنات الثنائية المتعددة											
	الاول-الثاني		الاول-الثالث		الاول-الرابع فأكثر		الثاني-الثالث		الثاني-الرابع فاكثر			الثالث-الرابع فأكثر
	Q_{MII}	الدلالة الاحصائية	Q_{MII}	الدلالة الاحصائية	Q_{MII}	الدلالة الاحصائية	Q_{MII}	الدلالة الاحصائية	Q_{MII}	الدلالة الاحصائية	Q_{MII}	الدلالة الاحصائية
يُؤكّد المُدرّس على العدالة ويتجنب التّحيز في تعامله مع الطلبة	.197	0.657	1.184	0.277	8.224	0.004	2.565	0.109	*15.465	0.000		3.784	0.052

يلاحظ من الجدول (9) أنَّ الفقرة (١٠) التي نصَّت على"يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة"، وقد أبدت أداءً تفاضلياً لصالح طلبة المستوى الرابع فاكثر في جامعة اليرموك مقارنةً بطلبة المستوى الثاني.

وللتمكن من تفسير الأسباب الكامنة وراء وجود أداء التفاضلي للفقرة ذات الرقم (١٠) لصالح طلبة المستوى الرابع فأكثر مقارنةً بطلبة المستوى الثاني في جامعة اليرموك، فقد تم حساب إحصائي الباقي المعياري المُعدَّل لتفاعل استجابات الطلبة على تدريجات الفقرات ذات الأداء التفاضلي تبعاً للمستوى الدراسي في جامعة اليرموك، وذلك كما هو مبين في الجدول (10).

جدول10: **نتائج إحصائي الباقي المعياري المُعدَّل لتفاعل استجابات الطلبة على تدريجات الفقرات ذات الأداء التفاضلي تبعاً للمستوى الدراسي في جامعة اليرموك.**

المستوى الدراسي	الفقرة (١٠)يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة				
	أبدًا	نادرًا	أحيانًا	غالبًا	دائمًا
مستوى أول	-0.5	-3.3	-2.1	-0.8	4.1
مستوى ثاني	-2.0	-1.1	-0.5	-1.4	3.1
مستوى ثالث	0.5	-0.5	0.5	0.5	-0.9
مستوى رابع فأكثر	1. 7	4.1	1.7	1.4	-5.4

يلاحظ من الجدول (10) أنَّ السبب الكامن وراء الأداء التفاضلي لنتائج الفقرة (١٠) في أداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً للمستوى الدراسي لصالح استجابات طلبة المستوى الرابع فأكثر مقارنةً باستجابات طلبة المستوى الثاني؛ وأنَّ التكرارات الملاحظة على التّدريجات المنخفضة (أبدًا، نادرًا، أحياناً) للفقرة(١٠) لدى طلبة المستوى الرابع فأكثر قد اختلفت جوهريًا عن تكراراتها المتوقعة وفق إحصائي الباقي المعياري المُعدَّل عند مستوى الدلالة (α=0.05)؛ مما يُعني وجود عدم إقرار لدى طلبة المستوى الرابع فأكثرحول مضامين الفقرة (١٠) مقارنةً بحيادية أو إقرار طلبة المستوى الثاني؛ مما استوجب دراسة مضامين الفقرة (١٠) التي نصّت على: يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة.

ثانيًا: النتائج المتعلقة بسؤال الدّراسة الثاني الذي نصّ على: **"ما أثر الفقرات ذات الأداء التّفاضلي على مؤشرات حُسن المطابقة لأداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً لمتغير المستوى الدراسي للطلبة؟".**

تم استخدام التّحليل العاملي التوكيدي Confirmatory Factor Analysis (CFA)، وهو حالة خاصة من نموذج المعادلة البنائية Structural Equation Modeling (SEM)، ولمعرفة أثر الفقرات ذات الاداء التفاضلي تبعاً لمتغير المستوى الدراسي،وذلك في ضوء عدد من المؤشرات الإحصائية لحُسن المطابقة حيث تقوم بتزويدناً عن مدى مطابقة النموذج للبيانات وهي كالآتي:

Chi-square: $\mathbf{C}_{min}$ (χ^2)
NCPEstimated Non-centrality Parameter:
F_{MIN}Minimum Fit Function Value:
F_0Population Discrepancy Function Value :

: RMSEARoot Mean Square Error Of Approximation

: ECVIExpected Cross-Validation Index

ويتم الاعتماد على أكثر من مؤشر لحُسن المطابقة وليس الاعتماد على مؤشر واحد فقط؛ لذا تم حساب مؤشرات حُسن المطابقة (C_{min} (χ^2), NCP, F_{MIN}, RMSEA, ECVI) التي يتوافر لها فترة ثقة بهدف الكشف عمّا إذا كان يوجد تقاطع من عدمه بين فترتي الثقة لمؤشرات حُسن المطابقة لأداة تقييم الطّلبة للأداء التّدريسي في جامعة اليرموك تبعاً لمتغير المستوى الدراسي.

وتم حساب مؤشرات حُسن المطابقة لأداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي للكشف عن أثر الفقرات التي تُبدي أداءً تفاضلياً، وذلك كما هو مبين في الجدول(11):

جدول 11: نتائج مؤشرات حُسن المطابقة لأداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك قبل وبعد الكشف عن الفقرات التي تُبدي أداءً تفاضلياً تبعاً للمستوى الدراسي:

مؤشرات حُسن المطابقة للأداة		جامعة اليرموك		النتيجة
		قبل الحذف	بعد الحذف	
عدد الفقرات		20	19	
عدد العزوم المميزة للعينة		210	190	
C_{MIN}	عدد المعالم	40	38	
	C_{MIN} (χ^2)	2716.240	2585.056	131.184
	درجة الحرية (عدد العزوم المميزة للعينة-عدد المعالم)	170	152	18
	احتمالية الخطأ	0.000	0.000	0.000
	C_{MIN}/DF	15.978	17.007	
NCP		2546.240	2433.056	
	الحد الأدنى لفترة الثقة ٩٠	2381.109	2271.895	يوجد تقاطع
	الحد الأعلى لفترة الثقة ٩٠	2718.721	2601.572	
F_{MIN}		0.370	0.352	
	F0	0.347	0.331	يوجد تقاطع
	الحد الأدنى لفترة الثقة ٩٠	0.324	0.309	
	الحد الأعلى لفترة الثقة ٩٠	0.370	0.354	

RMSEA	0.045	0.043	
الحد الأدنى لفترة الثقة ٩٠	0.044	0.042	يوجد تقاطع
الحد الأعلى لفترة الثقة ٩٠	0.047	0.045	
احتمالية الخطأ	0.000	0.000	
ECVI	0.381	0.362	
الحد الأدنى لفترة الثقة ٩٠	0.358	0.340	يوجد تقاطع
الحد الأعلى لفترة الثقة ٩٠	0.404	0.385	
MECVI	0.381	0.362	

يلاحظ من الجدول (11) أنَّ نتائج مؤشرات حُسن المطابقة (C_{min} ,(χ^2) NCP, F_{min}, ECVI) لأداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموكتبعاً للمستوى الدراسيتُظهر وجود فرق جوهري عند مستوى الدلالة (.05= α)، بين فترتي الثقة لمؤشرات حُسن المطابقة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي قبل وبعد الكشف عن الفقرات التي تبدي أداءً تفاضلياً؛ لصالح مؤشرات حسن المطابقة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي بعد حذف الفقرة التي تبدي أداءً تفاضلياً. مما يُعني أنَّ الأداة بعد تحررها من الفقرة التي تُبدي أداءً تفاضلياً أصبح يميل إلى حُسن المطابقةأفضل من ذي قبل (أي بوجود الفقرة ذات الأداء التّفاضلي).

مناقشة النتائج والتّوصيات:

أولًا: مناقشة النتائج المتعلقة بسؤال الدّراسة الأول الذي نصّ على: **"هل تظهر فقرات أداة تقييم الطلبةللأداء التّدريسي للمدرّس في جامعة اليرموك في عينة الدراسة أداءً تفاضليًا تبعاً لمتغير المستوى الدراسي للطلبة؟"**

أظهرت نتائج سؤال الدّراسة الأول؛ والمتعلقة بفقرات أداة تقييم الطلبة للأداء التّدريسي للمدرس في جامعة اليرموك تبعاً للمستوى الدراسي، وجود فقرة واحدة تُبدي أداءً تفاضليًا بشكل نهائي(المرحلة الثانية) في أداة تقييم للأداء التّدريسي للمدرّسفي جامعة اليرموك تبعاً للمستوى الدراسي باستخدام طريقة مانتل - هانزل العامة، وكانت لصالح طلبة المستوى الرابع فأكثر مقارنة بطلبة المستوى الثاني وهي الفقرة (١٠) ذات الأداء التّفاضلي التي تنص على: ("يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة).

وقام الباحث بحساب حجم الأداء التفاضلي للفقرة التي أبدت أداءً تفاضلياًفي أداة تقييم الطلبة للأداء التّدريسي للمدرّس في جامعة اليرموك تبعاً للمستوى الدراسي، باستخدام برنامج (DIFAS 5.0) من خلال حساب مؤشر Liu-Agresti Cumulative Common (Log-Odds Ratio: L-A LOR) وتبين أنه لا يوجد أثر للفقرة (١٠) التي كانت ذات أداء تفاضلي في مانتل - هانزل العامة.

وللتمكن من تفسير الأسباب الكامنة وراءوجود أداء تفاضلي للفقرة (١٠) التي تنص على("يُؤكّد المُدرّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة) في أداة تقييم الطلبة للأداء التّدريسي للمدرّس في جامعة اليرموك تبعاً للمستوى التعليمي، وبالنظر للجدول (١٥) نلاحظ أنّ أكبر قيمه ل Q_{MH} هي للفقرة ١٠ التي تنص على: ("يُؤكّد المُدرّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة) وتساوي (16.4050) وكانت لصالح طلبة المستوى الرابع فأكثر مقارنة بطلبة المستوى الثاني.

ويمكن عزو هذه النتيجة إلى أنّ طلبة المستوى الرابع فأكثر يكون لديهم اتجاهإيجابي نحو المُدرّس بدرجه أكبر من خلال خبرتهم الدراسية، وأن طلبة المستوى الأوليكونونطلاباً مستجدين في الجامعة،وليس لديهم القدرة والخبرة في تقييم المُدرّسين، وتكون صورة إدراكهم للمدرسغير دقيقة أو واقعية.

ومن جهة أخرى فإنّ هؤلاء الطلبةذوي المستوى الرابع يكونون على وشك التخرج، وبالتالي هم على معرفة كبيرة بالمُدرّسين وأساليبهم وجميع ظروفهم، ويقارنون بينهم طوال سنواتهم الدّراسية، فلا يتأثرون ببعض المتغيرات، وهذا ربما يؤدي إلى تقييمهم لأداء المُدرّس بصورة حقيقيةوفعاله، ولا سيما في إعطاء صورة حقيقية لما يجب أن يكون عليه المدرّس من العدالة وتجنّب التّحيز في تعامله مع الطّلبة, بالإضافة إلى إحساس طلبة المستوى الرابع فأكثر أنهم أكثر نضجاً و خبرة ووعياً من طلبة المستوى الثاني في التقييم وإعطاء الحكم الواقعي، وأنهم يميلون إلى إبراز نواحي الإيجابية في الأداء التدريسي، وإعطاء صورة مشعة عن الأداء التدريسي في الجامعة، لإحساسهم بالانتماء للجامعة والقسم، مما يؤدي إلى تقييم إيجابي للجوانب التي يرونها في المُدرّس.

وقد يرجع ذلك أيضاً إلى أن طلبة المستوى الثاني لم يمروا بخبرات تقييمية متعدّدة كما هو الحال بالنسبة لطلبة المستوى الرابع فأكثر، من حيث العلامات المتدنية أحياناً والعلامات المرتفعة أيضا في بعض المواد الدراسية ،فتراكم المعرفة عند الطلبة الذين عاشوا تلك الخبرات من

المستوى الرابع يؤدي إلى معرفة المُدرِّس من عدالته وعدم تحيزه للطلبة.

فإن طلبة المستوى الرابع فأكثر يتم ارتباطهم بالدراسة العملية بشكل أكبر من خلال مساقات المختبر والمساقات العملية التي تمكن الطلبة من تطبيق المعارف الجديدة والمهارات وبالإضافة إلى خبراتهم في السنوات السابقة قد تقود إلى الإحساس بالمهارة، مما يجعلهم أكثر قرباً من المُدرِّس وتقود بالتالي إلى الاحساس بالرضا نحو المدرس مما يؤدي إلى تفاعلات ايجابيه بصوره مرضيه بحيث ينعكس ايجابيا على تقييم الأداء من قبل الطلبة لهؤلاء المُدرِّسين. ولابد للإشارة أن طلبة الهندسة يدرسون خمسة سنوات دراسية أي في المستوى الرابع فأكثر ويكثر في دراستهم استخدام المختبرات ويكونون على قرب كبير من مدرسيهم ويرون فيهم العدالة وعدم التحيز في التعامل معهم.

فيجد طلبة المستوى الرابع فأكثر مجالاً خصباً لتكوين علاقات طيبة مع مُدرِّسيهم بشكل أكبرمن خلال مساقات المختبر والمساقات العملية التي تمكن الطلبة من تطبيق المعارف الجديدة والمهارات، وبالإضافة إلى خبراتهم في السنوات السابقة التي قد تقود إلى الإحساس بالمهارة، ممّا يجعلهم أكثر قرباً من المُدرِّس، وتقود بالتالي إلى الإحساس بالرضا نحو المدرس، مما يؤدي إلى تفاعلات إيجابية بصورة مرضية تنعكس إيجاباً على تقييم الأداء من قبل الطلبة لهؤلاء المُدرِّسينولا سيما في العدالة وعدم التحيز للطلبة. فيما يرى طلبة المستوى الثاني عدم إشراكهم في مجريات الدرس وعدم سماعهم أو التفاعل مع آرائهم أو تكوين علاقات معهم.

إن طلبة المستوى الرابع فأكثر لديهم اتجاه إيجابي نحو تقييم المُدرِّس بصورة عامة على معظم أبعاد مقياس التقييم، وقدرة على إصدار أحكام تقييمية واقعية بشكل دقيق، ولديهم ثقة بدرجة كبيرة في المُدرِّسوعدالته وعدم تحيزه في تعامله مع الطلبة، وحرصهم على تكوين علاقات إنسانية بدرجه أكبر من المستوى الثاني الذي يزداد لديهم جانب الخجل الاجتماعي.

وهناك ارتباط مهمّ بين حجم الطلبة في القاعة الجامعية وتقييمات الطّلبة نحو المُدرِّس؛ فإنّ عدد طلبة المستوى الرابع فأكثر في القاعة الجامعية أقلّ مقارنة بالمستوى الثاني وبشكل ملحوظ،فكلما قلّ أعداد الطلبة زاد إشراكهم في مجريات الدرس وسماع آرائهم واحترامها، ممّا يحسن تقييم الطلبة لعضو هيئة التدريس وخصوصاً من حيث العدالة وتجنّب التحيز في معاملته مع الطلبة. علمًا أنه في المساقات الاجبارية يكون عدد الطلبة أقلّ خصوصاً في المستوى الرابع فأكثر، مما يزيد الأفضلية لطرح الأسئلة على المدرّسين لتسهيل التعلم في سياق تطبيقي،

ومناقشة الطبيعة الانسانية والنمو، قد يعزز عملية بناء علاقة إيجابية بين المدرس والطالب. وبذلك، فإنّ العلاقة الإيجابية، تقود إلى تجنب التحيز في تعامل المدرّس مع الطلبة وإلى تقييمات أعلى للمُدرِّس.

إنّ الطلبة في المستوى الرابع فأكثر يتم تدريسهم من قبل الأستاذ الدكتور، والأستاذ المشارك، والأستاذ المساعد،من يمتلكون سماتالانبساط، والدعم، والكفاءة والقيادة والموضوعية، وغياب القلق، والهيمنة، والكفاءة العقلية، والدفء والخبرة، ويمكنهم اجتذاب الطّلبة نحو تقديم تقييمات إيجابية لها علاقة مهمة بتقييمات الطّلبة في العدالة وتجنب التحيز في التعامل مع الطلبة.

ثانياً: مناقشة النتائج المتعلقة بسؤال الدّراسة الثاني الذي نصّ على: **"ما أثر الفقرات ذات الأداء التّفاضلي على مؤشرات حُسن المطابقةلأداة تقييم الطّلبة للأداء التّدريسي للمدرس في جامعة اليرموك في عينة الدراسة تبعاً لمتغير المستوى الدراسي للطلبة؟".**

أظهرت النتائج وجود فرق جوهري عند مستوى الدلالة(α=.05)، بين فترتي الثقة لمؤشرات حُسن المطابقة لأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموك تبعاًللمستوى الدراسي قبل وبعد الكشف عن الفقرات التي تبدي أداءً تفاضلياً ؛ لصالح مؤشرات حُسن المطابقةلأداة تقييم الطّالب للأداء التّدريسي للمدرس في جامعة اليرموكتبعاً للمستوى الدراسيبعد حذف الفقرة التي تبدي أداءً تفاضلياً.

مما يعني أنَّ اداة تقييم الطّلبة للأداء التّدريسي للمدرّس في جامعة اليرموك تبعاًللمستوى الدراسي بعد تحررها من الفقرة التي تُبدي أداءً تفاضلياً أصبح يميل إلى حُسن المطابقة أكثر من قبل تحررها منها. وحسبما يشير باي[٣٦](Pae, 2004)، فإنَّ وجود فقرات ذات أداء تفاضلي في أداة تحدُّ من صدقه الداخلي وبالمحصلة تحدُّ من ثباته؛ وذلك لحساسية طلبة مستوى تعليمي تجاه مضامينها أكثر من أو أقل من حساسية المستويات الدراسية الأخرى تجاهها، على الرغم من تكافؤ قدراتهم، حيث إنّ الفقرات التي تبدي أداءً تفاضلياً تحدُّ من أحادية البعد للأداة فهي تشير إلى تعددية في البعدية، وأنَّها تنتهك خاصية اللاتغاير (Invariance). لذلك في حال التّخلص من الفقرات ذات الأداء التفاضلي تميل الأداة بفقراتها إلى أحادية البعد، وهذا يؤدي إلى مطابقة النموذج للبيانات لفقرات الأداة بعد حذف الفقرات ذات الأداء التّفاضلي أفضل من ذي قبل (أي بوجود الفقرات ذات الأداء التّفاضلي).

وذلك بسبب ظهور الفقرة ذات الأداء التفاضلي ذات الرقم (١٠) على أنَّها فقرة

صعبة بالنسبة لطلبة المستوى الثاني بدلالة أنَّ معظم استجاباتهم على تدريجاتها قد كانت منخفضة، مما يستدل منه على انخفاض قدراتهم إزاء تقييمهم لمضامين هذه الفقرة، وتتفق نتائج هذه الدراسة مع نتائج دراسات[٣٧](ارحيل,2013؛ مبارك, 2010؛مبارك ومبارك,2013)[٣٨] التي أظهرت أن هناك تحسناً بعد حذف الفقرات التي أظهرت أداءَ تفاضلياً.

ومما سبق من GMHDIF للكشف عن الفقرات التي تُبدي أداءً تفاضلياً على كافة فقرات أداة تقييم الطلبة للأداء التدريسي في جامعة اليرموك تبعاً للمستوى التعليمي تبين لنا أن أعلى قيمة لـQ_{MH} هي الفقرة (١٠) التي تنص على ("يُؤكِّد المُدرِّس على العدالة ويتجنب التّحيز في تعامله مع الطّلبة).

يتبيّن مما سبق أنّ هناك فقرة واحدة للأداء التّفاضلي في فقرات أداة تقييم الطلبة للأداء التدريسي في جامعة اليرموك تبعاً للمستوى الدراسي باستخدام طريقه مانتل- هانزل العامة.

وعندما تم حساب حجم الأداء التفاضلي للفقرة ذات الأداء التفاضلي للمستوى الدراسي تبين أنّه لا يوجد أثر لها.

التوصيات:

وبناءَ على نتائج الدّراسة يوصي الباحث بما يلي:

١-إعداد برنامج توعوي للطلبة يساعد على زياده فاعليتهم ومشاركتهم في عمليه تقييم الاداء الجامعي مما يؤدي الى زياده تحسين العملية التقييمية بالجامعة.

٢-اهتمام الجامعات بشكل أكبر بأدوات تقييم الطالب للأداء التدريسي- للمدرّس مع الأخذ بعين الاعتبار موضوع الأداء التفاضلي للفقرات عند إعدادها.

٣-إجراء دراسات عن الأداء التفاضلي لفقرات أداة تقييم من النوع المقالي، فلا توجد أي دراسة عربية تطرّقت لهذا الموضوع.

الهوامش:

(١)إرحيل، راتب. (2013). **أثر وجود أداء الفقرات التفاضلي للجنس في دقه المعادلة العامودية لاختبار أوتيس لينون للقدرة العقلية**. رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد- الأردن .المجلة الدولية التربوية المتخصصة. المجلد ٢، العدد٨، (٧٥٤-٧٧١).

(٢)جاد الرب، سيد محمد (2009)، **استراتيجيات تطوير وتحسين الأداء: الأطر المنهجية والتطبيقات العملية**، الطبعة الاولى، دار الفكر العربي: القاهرة.

(٣)الجعافرة، عبدالسلام .(2013). فاعلية أداء أعضاء هيئة التدريس في جامعة الزرقاء من وجهة نظر طلبتها في ضوء معايير جودة التعليم. **مجلة العلوم التربوية، (٤٢)**١، ١٣٩-١٥٥.

(٤)جرادات، عبد الله. (2003). **مقارنة بين طريقتي مانتل هانزل وطريقة الصعوبة المحولة في الكشف عن تحيز الفقرات.**

رسالة ماجستير غير منشورة. جامعة مؤتة: الكرك.

(٥)حمادنه، إياد.(2007). **الأداء التفاضلي لفقرات اختبارات "تحديد الكفاءة اللغوية" في اللغة الانجليزية في الجامعات الأردنية: دراسة مقارنة**. رسالة دكتوراه غير منشورة، جامعة اليرموك: اربد، الأردن.

(٦)الخرابشة، عمر والجعافرة، أسمى والهباهبة، عبدالله والسعايدة، ناجي. (2012). العوامل المؤثّرة في تقييم أداء أعضاء هيئة التدريس في كلية الأميرة عالية الجامعية في جامعة البلقاء التطبيقية في الأردن من وجهة نظر الطالبات. بحث منشور في **المجلة الدولية للأبحاث التربوية. جامعة الإمارات العربية المتحدة. (٣١)**٩، ٦١-٨٨.

(٧)الصباطي، إبراهيم ومحمد، شحته. (2007). دراسة تأثير بعض المتغيرات في تقييم طلبة الجامعة لأداء عضو هيئة التدريس. **المجلة العلمية في العلوم الإنسانية والإدارية، (٨)**١، ٩٧-١٧٦.

(8)Hambleton,R., &Swaminathan, H.,& Rogers, J.(1991). **Fundamentals of Item Response Theory**. Newbury Perk Califo rnia: Stage Publications.

(٩)عبيدات، عبدالله.(2014). **استقصاء الأداء التفاضلي لفقرات اختبار اختيار من متعدد تبعاً لدرجة مخاطرة المفحوصين.** أطروحة دكتوراه غير منشورة، جامعة اليرموك: اربد، الأردن.

(١٠)العجلوني، جهاد والسوالمة، يوسف (٢٠١٦). فحص الأداء التفاضلي لمموهات اختبار الرياضيات في التقييم الوطني الأردني لاقتصاد المعرفة وفقا لجنس الطالب وموقع المدرسة. **مجلة جامعة النجاح للأبحاث (العلوم الإنسانية)، ٣٠**(١١)، ٢٢٦٥-٢٢٨٦.

(١١)مبارك، مأمون ومبارك، وائل. (2013). الأداء التفاضلي لاختبار الرياضيات في الدراسة الدولية بيزا ٢٠٠٦. **مجلة العلوم التربوية، ٢٥** (٢)، ٢٤١ – ٢٦٧.

(١٢)مبارك، وائل.(2010). **الأداء التفاضلي لفقرات اختبار العلوم في الدراسة الدولية بيزا ٢٠٠٦**. رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

(13) Alquraan,M.,&Alkuwaiti, A. (2017). Differential Item Functi-oning in Students Rating Of Teaching: Effec -tiveness Surveys in Higher Education According to Academic Disciplines: Data From aSaudi University. **Journal of Educational and Psychological** Sultan Qaboos Univer- sity, 11 (4), 770-780.

(14) Benito, J., Navas -Ara, M. (2000) A comparison of X2, RFA and IRT Based procedures in the Detection of DIF. **Quality & Quantity 7**(34):17-31.

(15)Camilli,G., &Sheparad,L. (1994). **Methods For Identifying Bias Test Item**. Thousand Oasks, CA**:** Stage publication, USA.

(16)Dorans, N., & Holland, E. (1993). **DIF Detection and Description Mantel-Haenszel Standardization.** Educational testing service princetion, N.J. (QAT24225) ED287526.

(17)Fidalgo, A. M. (2005).Mantel–Haenszel Methods In B.S Everitt& D.C . Howell (Eds), **Encyclopedia of**

.**Applied Measurement in Eduction, 5**(14), 235-259.
(25)Pae,T. (2004). DIF for examinees with different acade- mic background. **Language Test- ing, 21**(1), 53-72.
(26) Penfield, R. D. (2007). An approach for categorizing DIF in polytomous items. **Applied Measure- ment in Education, 20**, 335-355. doi: 10.1080/08957340701431435.
(27)Penfield, R. D. (2012). **DIFAS 5.0 Differential Item Functioning Analysis System: User's manual** .Unpublished manuscript.
(28)Pokropek, A.&Kondratek, B.(2012). Test equating Definitions and examples of applications.**EDUKACJA, 120**(4), pp 52-71.
(29)Potenza, M.,&Dorans, N.(1995). DIF assessment for polytomously scored items: aframe work for classification and evaluation. **Applied Psychological Measurement, 11** (19), 23-27.
(30) Raju, N. S.,Drasgow, F ., &Slind, J. (1993), An Empirical Comparison Of The Area Methods, Lord's Chi-Square Test, And The Mantel-Haenszel Technique For Assessment Differential Item Function, **Educational And Psychological Measurement**, **53**(1) 301-314.
(31)Roever C. (2005). **That's not fair Fairness, bias, and Differential Item Functioning in Language Testing**. **http://www2.hawaii.edu/~roever/brownbag.pdf 25-6-2007**
statistics in behavioral science vol 3,pp1120-1126. Chichester, UK: John Wiley.& Sons .
(18)Fidalgo, A., & Madeira, J. (2008).Generalized Mantel–Haenszel Methods For Differential Item Functioning Detection. **Educational and Psychological measurement, 68**(6), 940-958.
(19)Gybles, J. (2004). The performance of some observed and unobserved conditional invariance techniques for the detection of differential item functioning. **Quality & Quantity, (38) 7**, 681-702.
(20) Hambleton,R., &Swaminathan, H.,& Rogers, J.(1991). **Fundamentals of Item Response Theory**. Newbury Perk California: Stage Publications.
(21) Innabi, H.& Dodeen, H. (2006). Content Analysis of Gender-related Differential Item Functioning TIMSS Items in Mathematics in Jordan**, School Science and Mathematics**. **(106)**8, 147-189.
(22)Karakaya, I. (2012). An Investig- tion of Item Bias in Science & Technology Subtests and Mathematic Subtests in Level Determination Exam (LDE)**. Educational science: theory &practice, (12)**1. Winter.222-229©2012 Educati- onal Consultancy and Research Center wwwe.edam.com.tr/estp.
(23)Pae,T. (2004). DIF for examinees with different academic background. **Language Testing, 21**(1), 53-72.
(24)Penfield, R. D. (2001). Assessing Differential Item FunctioningAmong Multiple Groups: Acomparison of threeMantel–Haenszelprocedures

(32) Shindi, Y., &Kazem,A. (2018). Sex Differential Item Functioning for Mathematics test in Cognitive Dev elopment Program in Sultanate of Oman by Mental-Haenszel and Item Characteristic Curve Methods. **International Journal of Learning Management Systems**, Sys. 6, No. 2,pp 61-73.

(33)Seldin, P. (1989). Using student feedback to improve teaching, New directions for teaching and learning. **The Department Chair Person's Role in Enhancing College Teaching.(37)**9, PP. 89-97.

(34)Spray, J.,&Miller,T.(1994). **Identify ing non uniform DIF in polytomously scored test items.** American College Testing Research Report Series 94-1, Iowa city,Iowa.

(35)Tagomori, Harry T. (1993)**. A Content Analysis Of Instruments Used For Student Evaluation Of Faculty In Schools At Universities And Colleges Accredited By The National Councit For Accreditiation OF Teacher Education (NCATE) (Student-Rated). DAI-A54/**05,P1629 (OrderNo. AAC9325732).

(36)Tang, T.L. (1997). Teaching evalua- tion at a public institution of higher

education: Factors related to the overall teaching effectiveness, **Public Personal Management,Vol. 26** , No. 3 , pp. 379-389.

(37)Uiterwijk, H&Vallen, T.(2005). Linguistic Sources Of Item BiasFor Second Generation Immigrants In Dutch Tests. **Language Testing. (22)**2, 211-234.

(38) Zwick, R., Donoghue, J.&Grima, A. (1993). Assessment of differential item functioning for performance tasks. **Journal of Educational Measurement, 15**(30),233-251.