توظيف نموذج راش في بناء اختبار تحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي بفلسطين

د.معزوز جابر علاونه/ جامعة القدس المفتوحة

ملخص

هدفت هذه الدراسة إلى توظيف أنموذج راش في بناء اختبار تحصيلي في منهاج التاريخ، ومن أجل تحقيق ذلك قام الباحث ببناء اختبار تحصيلي لمنهاج التاريخ، تكون بصورته الأولية من (136) فقرة من نوع الاختيار من متعدد بأربعة بدائل. وتم تطبيقه على عينة مؤلفة من (327) طالبا وطالبة من طلبة الصف الثاني الثانوي الأدبي في مدينة نابلس، من أجل تحليل البيانات وقد استخدمت البرمجية الإحصائية (winsteps)، والرزمة الإحصائية للعلوم الاجتماعية (SPSS).

وقد تم التحقق من مطابقة الاختبار التحصيلي لافتراضات أنموذج راش، إذ حذف (25) فرداً، لعدم مطابقتهم لافتراضات الأنموذج، بالإضافة إلى (5) فقرات من فقرات الاختبار، وقد التحقق من صدق الاختبار وثباته؛ وأشارت النتائج أنَ ثبات الأفراد الذين طبق عليهم الاختبار (0.80) في حين بلغ معامل ثبات فقرات الاختبار(0.73)، وبناء على النتائج أوصت الدراسة بإجراء دراسة مقارنة بين نظرية القياس الحديثة، والكلاسيكية في التحقق من خصائص الاختبارات التحصيلية.

الكلمات المفتاحية: نظرية الاستجابة للفقرة، أنموذج راش، اختبار تحصيلي.

Employing Rush Model for Test Construction of the History Curriculum of 12th Grade in Palestine

Abstract

This study aimed at employing Rasch model for constructing achievement test for the history Curriculum. To achieve, this objective a (136) multiple-choice item test of four alternative each was administered among a sample of (327) male and female students in the second secondary grade (Literary stream) in Nablus city. The statistical software (Winsteps) and SPSS were used to analyze the data. To fit the achievement test with Rasch Model, (25) students were excluded from the sample because they were found to be inconsistent with the model suppositions; in addition; (5) items were excluded. Furthermore, test validity and reliability were ensured and showed that the participants> reliability was (0.80) while test reliability was (0.73). In light of the results further comparative research is recommended to compare between the modern measurement theory and the classical theory to verify achievement test properties.

Keywords: Item Response Theory, Rasch Model, Achievement Test

مقدمة:

قام علماء القياس النفسي بمحاولات عدةّ في سبيل التوصل إلى الموضوعية والدقة في تقدير السلوك، فتحقيق الموضوعية والدقة في تقدير السلوك يُعدّ الخطوة الأولى في فهم الظاهرة السلوكية وضبطها والتحكم فيها، لذا، فبناء الأداة التي تصل إلى التقدير الموضوعي للسلوك من أهم أهداف العاملين في المجال السلوكي (الشواورة، 2013).

تعد الاختبارات التحصيلية من أكثر أدوات القياس شيوعاً واستخداماً في تقويم الطلبة، في معظم البلدان العربية والعالمية، وقد نالت عملية بنائها وتطويرها اهتماماً بالغاً من التربويين وصناع القرار في المجال التربوي، وذلك لما يمكن أن تقدمه من بيانات يمكن الاعتماد عليها في صنع القرارات التربوية السليمة، التي تهدف إلى تحسين عمليتي التعلم والتعليم (أبو جراد، 2011).

وقد كانت الاختبارات التحصيلية، وما زالت ، من أهم الوسائل وأكثرها فاعلية بالنسبة لأطراف العملية التعليمية التعلمية جميعها، لما تقدمه من معلومات مفيدة لهم، ولما تتمتع به من سهولة التنفيذ وبساطته، ولكن مع تقدم العلوم التربوية والنفسية، بشكل عام، وعلم القياس بشكل خاص، تبين قصور استخدام الاختبارات التحصيلية في قياس قدرات الطلبة التحصيلية، وتحديد مستوياتهم الدراسية، إذ وجهت لنظرية القياس التقليدية مجموعة من الانتقادات بسبب اعتمادها على معيار معين ووحيد في تفسير الدرجات، وهذا المعيار هو معيار الجماعة في الاختبارات جماعية المعيار، ومعيار المحك في الاختبارات محكية المرجع(علي، 2012).

وتعد الاختبارات التحصيلية أدوات لقياس مدى تمكن الفرد من الفهم والمعرفة والمهارة في محتوى علمي معين، أو مادة دراسية، أو تدريبية، أو محتوى مجموعة من المواد الدراسية في مرحلة معينة. وتُعدّ الاختبارات التحصيلية التي يعدها المعلمون لقياس مستوى تحصيل طلبتهم من أهم أدوات التقويم التربوي، وأكثرها شيوعاً في مجال التقويم التربوي، والوسيلة الأساسية المستخدمة في انتقاء الطلبة وتوجيههم، ولذا، فإن الاختبار التحصيلي يحتل مكانة مهمة، خاصة بين أدوات التقويم التربوي، لما لها من أهمية في صنع القرارات التربوية، إذ لا تقتصر على ترفيع الطلبة وترسيبهم بل تتعدى إلى تحديد مستوياتهم وتصنيفهم في فئات مختلفة حسب مستوياتهم (السحيمات، 2007).

ونظراً لأهمية الاختبارات بشكل عام، واختبارات التحصيل بشكل خاص، فقد اهتم الباحثون بإجراء الكثير من الدراسات والبحوث التي اهتمت ببناء فقرات الاختبارات وانتقائها بالاعتماد على نظرية القياس التقليدية، وكان نتيجتها قيام الكثير من العلماء مثل ثرستون (Thurstone,1995)، ولورد (Lord,1953) وبيرنبوم

(Birnbaum,1958) بتوجيه مجموعة من الانتقادات لنظرية القياس التقليدية، بسبب قصورها في تحقيق الموضوعية في القياس، والمتمثلة في اعتماد القياس على خصائص الاختبار المستخدم (الصدق، والثبات، والموضوعية)، بمعنى اختلاف نتيجة القياس باختلاف الاختبار المستخدم، وكذلك اختلاف نتيجة القياس باختلاف مستوى الأفراد الذين يطبق عليهم الاختبار، وبالإضافة إلى ذلك، عدم خطية القياس؛ أي عدم وجود وحدات متساوية على مدى المتصل (القدرة) الذي يتدرج عليه المتغير موضوع القياس (علام، 2010)، وهذا ما دفع الكثير من الباحثين والمهتمين بالقياس إلى بناء نظرية حديثة تتسم بالمرونة، والقدرة على تلافي المشكلات التي تعاني منها أساليب القياس التقليدية من جهة، وتحقيق الدقة والموضوعية من جهة أخرى.

وتعد نظرية القياس الحديثة، أو ما يُعرف بنظرية الاستجابة للفقرة طريقة موضوعية لنمذجة البيانات، بمعنى أنها تحاول نمذجة العلاقة بين المتغير غير الملاحظ ويسمى قدرة الفرد، واحتمال الإجابة الصحيحة على الفقرة وما يعرف بمتغير الملاحظ، ويعتمد صدق طرق نظرية الاستجابة للفقرة، في جزء كبير، على البيانات المستمدة من أداء الفرد (Harwell,1997)، ويشير أنجيبو (Ingebo,1993) إلى أن نظرية الاستجابة للفقرة تقوم على أساس الجمع بين متغيرين في علاقة رياضية واحدة، وتمثل هذه العلاقة أنموذجاً احتمالياً، وذلك نظراً لأن طبيعة البيانات تسمح بالتنبؤ باحتمال النجاح على فقرة من الفقرات المدرجة، كما توفر تقديرات لقدرات الأفراد بوحدات الأنموذج المستخدم.

وتتضمن نظرية الاستجابة للفقرة نماذج عدةّ تسمح بالقياس الموضوعي، حيث يتم التركيز على الفقرات الاختبارية، مما يسمح بإضافة مفردات، أو حذفها، أو تعديلها، دون أن يتأثر الاختبار ككل، مما يساعد في التغلب على مشكلات القياس النظرية والتطبيقية التي فشلت نظرية القياس الكلاسيكية في حلها (اسماعيل، 2007).

ويعدّ أنموذج راش (Rasch Model) من أهم نماذج نظرية السمات الكامنة وأبسطها، وأكثرها استخداماً في بناء الاختبارات وتحليل فقراتها، فهو من أكثر النماذج التي أجريت حوله بحوث متعددة للتحقق من خصائصه، وقدرته على بناء الاختبارات ومواجهة أوجه القصور الناتجة عن استخدام المدخل الكلاسيكي في القياس النفسي والتربوي (زكريا، 2009). ويمتاز هذا الأنموذج من غيره عن نماذج نظرية الاستجابة للفقرة بمجموعة من الخصائص، من أهمها: اعتماد الأنموذج على إعطاء استجابتين للفقرة، وهي إما صح أو خطأ ويعبر عنها بالأرقام (1 ، 0) وهذا من شأنه أن يصل بالمقياس إلى الموضوعية، أو تقترب استجابات المفحوصين من الدقة (مقارنة بما لو كانت أكثر من استجابتين) (Moral et al., 2006). كما يؤدي الأنموذج إلى منحنيات مميزة للفقرات تكون على شكل متوازيات، وكذلك فان هناك منحنيات مميزة للأفراد الذين يطبق

عليهم الاختبار، فالمنحنى المميز للفقرة هو عبارة عن دالة رياضية تربط بين احتمال الإجابة الصحيحة لفرد ما على فقرة اختبارية وبين القدرة التي يقيسها الاختبار (Vanderlinden,2001). كما أن صعوبة الفقرات لا تعتمد على تقديرات صعوبة الفقرات الأخرى المكونة للاختبار، كذلك لا تعتمد على تقديرات الأفراد اللذين يجيبون عن الاختبار، كما لا تعتمد تقديرات قدرة الفرد على تقديرات قدرة أي مجموعة من الأفراد الذين يؤدون الاختبار، أو على تقديرات صعوبة فقرات الاختبار (Avery L, Russell D, Raina P, Walter S, Rosenbaum P,2003) كذلك توفير برامج حاسوبية متعددة لمعالجة البيانات، وبأسعار مناسبة غير مكلفة (أبو هاشم، 2006).

ويعبر عن أنموذج راش بالدالة الرياضية التي تربط بين احتمال الاستجابة الصحيحة باقتران لوجستي للفرق بين برامتري قدرة الفرد وصعوبة الفقرة، كما يفترض هذا الأنموذج تساوي القوة التمييزية للفقرات، وأن معلمة التخمين تساوي صفراً لجميع الفقرات، والدالة الرياضية للأنموذج يعبر عنها في المعادلة الآتية:

$$p_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} \quad , \qquad (i = 1, 2, 3, \dots\dots, n) \dots\dots\dots\dots\dots\dots$$

حيث:

$p_i(\theta)$ احتمال الإجابة الصحيحة لفرد قدرته (θ) عن فقرة (i) حيث (i) تعني رقم الفقرة (١ ، ٢، ٣،....)

e تعني الأساس اللوغاريتمري وهو (2.718)

D عامل تدريج ويقدر بـ (١,٧)

(θ) قدرة فرد ما

(b_i) معلم صعوبة الفقرة (i)

ويعتمد هذا الأنموذج على الفرق بين القدرة (θ)، التي يمثلها المفحوص حرف (S)، في السمة المراد تقديرها (القدرة الكامنة) وراء إجابة المفحوص، ودرجة الفقرة (i)، التي يرغب المفحوص بالإجابة عنها، والتي يمثلها الرمز (b_i) (التقي، 2013).

ويعد أنموذج راش، الذي وضعه العالم الدنمركي جورج راش في عام (1960)، وطوره للتطبيق العملي العالم الأمريكي رايت (Wright)، من أهم نماذج نظرية القياس الحديثة، إذ يهدف هذا الأنموذج إلى تحقيق خاصية الموضوعية في القياس النفسي والتربوي (Harris,1989).

وتقيّم الفقرات الاختبارية حسب أنموذج راش، حسب مطابقتها للأنموذج الذي يصف العلاقة بين الاستجابة والقدرة باستخدام اختبار إحصائي أو تحليل البواقي (Residuals)، فيتم التعرف إلى الفقرات الضعيفة من خلال مؤشرات التمييز التي تكون سالبة أو موجبة، وتقديرات مؤشرات الصعوبة، وإلا تكون الفقرات سهلة جداً، أو صعبة جداً لدى مجموعة من المفحوصين، كما يتم أيضا تقدير قدرة المفحوصين من خلال ربط استجابة المفحوص على فقرات الاختبار باستخدام طريقة الأرجحية العظمى بإيجاد قيمة (θ) التي تعطي أرجحية عظمى للحصول على نمط الإجابة عن فقرات الاختبار، ويقوم هذا الأساس على إيجاد الأرجحية العظمى للقدرة (θ)، التي تتراوح بين (+∞ ، −∞) من خلال الاعتماد على طرق رياضية وبتقريب هذه القيمة من خلال تحسين عملية التقدير بدءاً من قيمة تحدد مسبقاً (التقي، 2013).

وفي هذا المجال، أشار الكثير من الباحثين، إلى فاعلية أنموذج راش في بناء الاختبارات التحصيلية، فقد أشار فوربس (Forbes,1989)، في دراسته حول فاعلية استخدام أنموذج راش في بناء الاختبارات التحصيلية، في اللغة الانجليزية، والعلوم، والرياضيات إلى أنه يمكن إعداد اختبارات تحصيلية أكثر جودة باستخدام هذا الأنموذج، وتوصلت دراسة عبد الوهاب (2010) إلى ضرورة توعية المتخصصين في مجال القياس النفسي بأهمية استخدام نظرية الاستجابة للمفردة في تحليل مفردات الاختبارات والمقاييس المستخدمة ومحاولة تطويرها، كما أوصت علي (2012) في دراستها إلى استخدام أنموذج راش في بناء اختبارات تحصيلية.

مشكلة الدراسة:

يغلب على الاختبارات التحصيلية المستخدمة في المدارس العربية بشكل عام، والمدارس الفلسطينية بشكل خاص، اعتمادها عند بنائها واستخراج خصائصها السيكومترية على مبادئ نظرية القياس التقليدية وأسسها، ونتيجة للعيوب المنطقية المتعلقة بنتائج تحليل الاختبارات التحصيلية باستخدام نظريه القياس التقليدية، ونظراً للمشكلات الكثيرة التي كانت تحتويها الاختبارات التحصيلية التي بنيت وفق نظرية القياس التقليدية والتي تمثلت في حصول الفرد على علامتين مختلفتين إذا تم اختباره باختبارين مختلفين يقيسان نفس السمة، وذلك لاختلاف الخطأ المعياري في الحالتين، (Domino&Domino,2006) وتأثر أداء الفرد على نفس الاختبار وتحت نفس الظروف الاختبارية إذا أعيد تطبيقه عليه مرة أخرى وذلك بسبب وجود الخطأ العشوائي في الدرجة التي يحصل عليها (Sijitsma&Junker,2006)، كما تفترض نظرية القياس التقليدية أن الدرجة الملاحظة التي يحصل عليها في اختبار معين هي مجموع درجاته على جميع فقرات الاختبار، وبالتالي قد يحصل فردان على نفس الدرجة رغم اختلاف الفقرات التي أجابا عنها، كما أنه قد يتساوى الفرق بين درجتي فردين منخفضين في القدرة، أو السمة مع الفرق بين درجتي فردين مرتفعين في القدرة أو السمة التي يقيسها المقياس رغم اختلاف صعوبة المفردات في الحالتين (Baker, 2001)، كما أن اعتماد النظرية الكلاسيكية في تحليل المفردات على عينة الأفراد المطبق عليهم المقياس سواء عند حساب الثبات أو معاملات التمييز أو معاملات الصعوبة يؤدي إلى عدم توفر القياس الموضوعي لهذه السمة أو القدرة (Stage,2008) إذ يتأثر حساب الثبات مثلاً بدرجة تجانس العينة، كما تتأثر معاملات صعوبة المفردات باختلاف مستوى الأفراد المطبق عليهم الاختبار(Reeve,2004) في الوقت الذي تتعدد فيه عيوب النظرية الكلاسيكية في القياس تزداد مزايا نظرية الاستجابة للمفردة بنماذجها المختلفة، وتتمثل هذه المزايا في استقلال خصائص المفردات (مثل: معاملات الصعوبة والتمييز) عن عينة الأفراد المستخدمة في التحليل، وكذلك استقلال تقدير قدرات الأفراد عن عينة المفردات المكونة للمقياس، كما أن تفسير درجات الأفراد يتم في ضوء المفردات وليس في ضوء الجماعة المرجعية كما في النظرية الكلاسيكية، وتتحقق خصائص الميزان الفتري وربما القياس النسبي دون ضرورة أن يكون توزيع مستويات القدرة في المجتمع المستهدف اعتدالياً، ويتم تقدير الخطأ المعياري لكل مختبر على حده وليس خطأ معياري واحد لكل المختبرين، مع إمكانية المقارنة بين أداء الأفراد الذين اختبروا باختبارات مختلفة تقيس نفس السمة أو القدرة دون أن تكون هذه الاختبارات متوازية (علام،٢٠٠٥) وفي ظل التزايد الواضح في مزايا نظرية الاستجابة للمفردة وتعدد عيوب النظرية الكلاسيكية اتجه أغلب العلماء إلى استخدام مبادئ نظرية الاستجابة للمفردة في بناء اختبارات جديدة أو التأكد من مطابقة اختبارات معدة مسبقاً لافتراضات نماذج نظرية الاستجابة للمفردة، أو تطوير هذه الاختبارات لتتطابق مع تلك النماذج؛ وذلك من أجل تحقيق شروط القياس الموضوعي للسمة أو القدرة المراد قياسها، تلك الشروط المتمثلة في استقلال تقدير خصائص المفردات عن عينة الأفراد المطبق عليهم المقياس، واستقلال تقديرات الأفراد عن مجموعة المفردات المكونة للمقياس (الشرقاوي، وآخرون، 1996).

ولعل من أهم الحلول التي وضعت لمعالجة هذه المشكلات هو تطبيق نماذج نظرية القياس الحديثة التي اقترحها أصحاب هذه النظرية، ومن أشهر هذه النماذج وأبسطها أنموذج راش أحادي المعلمة، وبناءً على ما سبق، فقد ارتأى الباحث إجراء دراسة تعتمد على استخدام أنموذج راش في بناء اختبار تحصيلي في مادة التاريخ للصف الثاني الثانوي الأدبي بفلسطين، وقد قام الباحث باختيار منهاج التاريخ دون غيره لما له من مكانة عظيمة بين مختلف المناهج في المرحلة الثانوية في الفرع الادبي بفلسطين ومن هنا، فقد تحددت مشكلة الدراسة بتوظيف أنموذج راش في بناء اختبار تحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي بفلسطين، وبالتحديد إن مشكلة الدراسة تتلخص في الإجابة عن الأسئلة الآتية:

1. ما مدى مطابقة فقرات الاختبار التحصيلي محكي المرجع مع افتراضات أنموذج راش؟
2. ما تقدير قدرات الأفراد المقابلة لكل درجة كلية محتملة على الاختبار التحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي، حسب أنموذج راش؟
3. ما تقدير صعوبة فقرات الاختبار التحصيلي، في منهاج التاريخ للصف الثاني الثانوي الأدبي، حسب ملاءمتها لأنموذج راش؟
4. ما مدى صدق الاختبار التحصيلي وثباته في منهاج التاريخ للصف الثاني الثانوي الأدبي باستخدام أنموذج راش؟

أهداف الدراسة:

هدفت الدراسة إلى تحقيق الأهداف الآتية :

1. توظيف أنموذج راش في بناء اختبار تحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي بفلسطين.
2. مطابقة فقرات الاختبار التحصيلي محكي المرجع مع افتراضات أنموذج راش.
3. تقدير قدرات الأفراد المقابلة لكل درجة كلية محتملة على الاختبار التحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي حسب أنموذج راش.
4. تقدير صعوبة فقرات الاختبار التحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي حسب ملاءمتها لأنموذج راش.
5. التحقق من صدق الاختبار التحصيلي وثباته في منهاج التاريخ للصف الثاني الثانوي الأدبي باستخدام أنموذج راش.

أهمية الدراسة:

تكتسب هذه الدراسة أهميتها كونها من أولى الدراسات – في حدود علم الباحث – التي تناولت بناء اختبار تحصيلي في منهاج التاريخ باستخدام أحد نماذج نظرية الاستجابة للفقرة، ولما لهذه النظرية من أهمية في التغلب على الكثير من مشكلات القياس التي كانت تواجه نظرية القياس التقليدية، وتحقيقها لأهداف القياس الموضوعي، وتأتي أهمية هذه الدراسة من خلال استخدامها أنموذج راش الذي يحقق خاصية الصدق، من خلال استبعاد الأفراد غير الملائمين، والمفردات غير الملائمة، واستبقاء استجابات الأفراد الصادقة في تدريجها على متغير موضوع القياس، والمفردات الصادقة في تعريفها للمتغير موضوع القياس. بالإضافة إلى أن منهاج التاريخ للصف الثاني الثانوي، الذي يعد من المقررات الأساسية في هذا الصف.

حدود الدراسة:

اقتصرت الدراسة على الحدود الآتية:

1. طلبة الصف الثاني الثانوية الأدبي في مدينة نابلس، وذلك خلال الفصل الدراسي الأول من العام الدراسي 2016/2015م
2. الاختبار التحصيلي في منهاج التاريخ الفلسطيني للصف الثاني الثانوي الأدبي.
3. اختبار التاريخ المؤلف من (136) فقرة من نوع الاختيار من متعدد.
4. أحد نماذج نظرية القياس الحديثة أنموذج راش اللوغرتمي الذي يعتمد على معلم واحد.
5. إجراء التحليل الإحصائي باستخدام برمجية (winsteps) وبرنامج (SPSS)

مصطلحات الدراسة :

نظرية الاستجابة للفقرة: Item Response Theory

وهي نظرية تفترض أنه يمكن التنبؤ بأداء الأفراد، وتفسير أدائهم في اختبار تربوي معين بناءً على خاصية أو خصائص مميزة لهذا الأداء تسمى سمات، وتحاول هذه النظرية تقدير درجات الأفراد في هذه السمات، ونظراً لصعوبة ملاحظة هذه السمات أو قياسها بطريقة مباشرة، فإنه يجب تقديرها أو الاستدلال عليها من خلال استجابات الأفراد الملاحظة على الاختبار الذي تقيسه السمة، لهذا، يطلق عليها بعض العلماء نظرية السمات الكامنة، وتهدف هذه النظرية الى تقدير إحصاءات الفقرة والقدرة جميعها، وكلما كان هناك ملاءمة بين الأنموذج المستخدم ومجموعة البيانات، أدى ذلك إلى الحصول على تقديرات دقيقة لهذه الإحصاءات(عبد المسيح، 1990).

أنموذج راش: Rash Model

هو أحد نماذج نظرية الاستجابة للفقرة، قام بوضعه جورج راش وهو نموذج لوجستي ذو معلمة واحدة، ويعمل على تقدير احتمالية إجابة الفرد عن فقرة ما إجابة صحيحة، بدلالة قدرته، ومعامل صعوبة الفقرة بغض النظر عن عدد فقرات الاختبار وعدد أفراد العينة(حمادنه والبلاونه، 2015) واستخدم نموذج راش في تحليل فقرات الاختبارات التي تعطى فيها درجة خام واحدة عن كل استجابة صحيحة على المفردة، ودرجة خام صفر عن كل استجابة خاطئة (ابو جراد، 2014).

اختبار تحصيلي:

طريقة منظمة لقياس مستوى تحصيل الطلبة لمعلومات مادة دراسية ومهاراتها من خلال الإجابة عن مجموعة من الفقرات التي تمثل محتوى المادة الدراسية (عودة، 2010).

الدراسات السابقة:

بعد مراجعة مصادر المعلومات المختلفة، تبين للباحث أن هناك الكثير من الدراسات التي تناولت موضوع الدراسة بجوانبه المختلفة، سواء كان من خلال بناء اختبارات التحصيل، والخصائص السيكومترية أو توظيف نظرية القياس الحديثة في بناء الاختبارات التحصيلية، وتوظيف النظرية التلقليدية في بناء الاختبارات التحصيلية، غير أن الباحث لم يتوصل إلى أية دراسة تناولت توظيف أنموذج راش في بناء اختبار تحصيلي في التاريخ، سواء على المستوى المحلي أو العربي أو العالمي، وبناءً عليه فإن الباحث ارتأى عرض مجموعة من الدراسات السابقة ذات العلاقة بموضوع الدراسة، خاصة الدراسات التي تناولت نظرية القياس الحديثة، وخاصة أنموذج راش، و وفيما يأتي عرض هذه الدراسات:

ففي دراسة أجراها الطراونه (2016). هدفت الدراسة إلى بناء اختبار محكي المرجع في الإحصاء التحليلي لطلبة الدراسات العليا في كلية العلوم التربوية في الجامعات الأردنية الحكومية باستخدام نظرية الاستجابة للفقرة، ومن أجل تحقيق ذلك، قام الباحث باختيار عينة مؤلفة من (319) طالبا وطالبة من طلبة الماجستير في كليات العلوم التربوية في الجامعات الأردنية، وقام بتوزيع الاختبار التحصيلي عليهم، وأشارت نتائج الدراسة إلى أنَ مطابقة عينة الدراسة وفقرات الاختبار التحصيلي لافتراضات أنموذج راش، كما تم التحقق من افتراضات نظرية الاستجابة للفقرة. وقد تم حساب معاملات الصعوبة لفقرات الاختبار باستخدام البرنامج الإحصائي (Bigsteps) وقد تكون الاختبار بصورته النهائية من (36) فقرة، وأظهرت النتائج أن الاختبار يتمتع بدلالات صدق وثبات مقبولة.

أما الدراسة التي قام بها المصري (2015)، والتي هدفت إلى استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في البحث التربوي وفق الأنموذج اللوجستي أحادي المعلمة ، ومن أجل تحقيق ذلك، تم بناء اختبار مؤلف من (48) فقرة تم توزيعه على عينة بلغ حجمها (220) طالباً وطالبة، وأشارت النتائج إلى أن الاختبار حقق افتراضات الأنموذج أحادي المعلمة، كما تبين أن (42) فقرة مطابقة للأنموذج، كما أظهرت نتائج تقديرات قدرات الأفراد استبعاد (7) أفراد غير مطابقين للأنموذج، وتراوحت معالم القدرة لأفراد عينة الدراسة ما بين (2.659−) لوجيت و (2.812) لوجيت.

كما أجرى العنزي (2014) دراسة هدفت إلى تحديد علاقة الحكمة الاختبارية بتحصيل طلبة المرحلة الثانوية في دولة الكويت، في اختبار تحصيلي مبني وفق أنموذج راش، ومن أجل تحقيق ذلك قام الباحث ببناء اختبار تحصيلي حسب أنموذج راش تم توزيعه على عينة مؤلفة من (284) طالبا وطالبة، فأشارت النتائج إلى أن الاختبار التحصيلي تمتع بصدق وثبات مناسبين وذلك حسب أنموذج راش.

وقام أون (Onn, 2013) بدراسة هدفت إلى مقارنة بين النظرية التقليدية ونظرية الاستجابة للفقرة ، وذلك من حيث عدد الفقرات المنتقاة، ومعلمة الثبات، ومن أجل تحقيق هدف الدراسة، قام الباحث بإعداد اختبار تحصيلي في مادة الفيزياء تكون من (50) فقرة من نوع الاختيار من متعدد، قام بتوزيعه على عينة مؤلفة من (69) طالبا وطالبة من طلبة المدارس في نيجيريا ، وقد أشارت نتائج الدراسة إلى مطابقة (38) فقرة لنظرية الاستجابة للفقرة، وحذف (12) فقرة، كما أشارت النتائج إلى تدني معامل الثبات للاختبار وذلك وفق نظرية الاستجابة للفقرة، فقد بلغ معامل الثبات (0.67).

أما الدراسة التي أجراها هادي ومراد (2013) والتي هدفت إلى استخدام أنموذج راش في إعداد وتقنين اختبار في اللغة العربية لقبول الطلبة في الدراسات العليا بكليات العلوم الإنسانية بجامعة الكويت، ومن أجل تحقيق ذلك قام الباحث ببناء اختبار مؤلف من (130) فقرة، تم تطبيقها على عينة استطلاعية مؤلفة من (86) طالباً وطالبة، وأسفرت النتائج عن حذف (30) فقرة وتعديل (16)، وقام الباحثان بتطبيق الاختبار في صورته النهائية على عينة بلغ حجمها(335) طالباً وطالبة، وتم تحليل البيانات باستخدام أنموذج راش، وقد أسفرت نتائج التحليل باستخدام أنموذج راش، إلى تدريج فقرات الاختبار المؤلف من (100) فقرة، فقد تم حذف(26) فرداً و(19) فقرة وذلك لعدم ملاءمتهم للأنموذج، وأظهر التحليل عن معامل ثبات جيد للاختبار بلغ (0.81 و 0.99) لقدرات الأفراد وتراوح مدى القدرة بين (2+،2−) لوجيت وبمتوسط يساوي (0.5) وهو مستوى جيد.

وفي دراسة أجراها السامرائي والخفاجي (2012) هدفت إلى بناء اختبار تحصيلي في مادة علم النفس الخواص لطلبة أقسام العلوم التربوية والنفسية في الجامعات العراقية وفق نظرية السمات الكامنة ونماذج القياس المنبثقة عنها (أنموذج راش)، ومن أجل تحقيق هدف الدراسة قام الباحثان ببناء اختبار تحصيلي مؤلف من (160) فقرة من نوع الاختيار من متعدد، تم توزيعه على عينة مؤلفة من (349) طالبا وطالبة، فأشارت نتائج الدراسة إلى عدم مطابقة (35) فقرة من فقرات الاختبار لنتائج التحليل؛ فقد تم استبعادها، ومطابقة الفقرات المتبقية والبالغ عددها (125) فقرة متطلبات استقلالية القياس، كما أظهرت النتائج اقتراب معامل التمييز للاختبار من (1)، وكانت قيمة معامل التمييز من المدى المقبول والبالغة (1.106)، وبلغت قيمة معامل الثبات باستخدام أنموذج راش (0.957).

أما دراسة علي (2012) فقد هدفت إلى التعرف إلى فاعلية استخدام أنموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم في التربية، ومن أجل تحقيق ذلك، قام الباحث ببناء اختبار تحصيلي محكي المرجع مؤلف من (99) فقرة من نوع الاختيار من أربعة بدائل، وقد تم توزيعه على عينة مؤلفة من (426)طالبا وطالبة من طلبة كلية التربية في جامعة دمشق، وقد توصلت الدراسة إلى أنه تم حذف (10) فقرات من الاختبار ليصبح(89) فقرة ، وتم الحذف نتيجة وقوع هذه الفقرات خارج حدود المطابقة والتي حددت بـ(0.7-1.3)، كما أشارت النتائج إلى أن معامل ثبات الاختبار بلغ (0.98)، في حين بلغ معامل ثبات الأفراد(0.95). كما بينت النتائج أيضا، أن الاختبار يحقق افتراضات نظرية الاستجابة للفقرة، كما بينت النتائج أن قيمة دالة المعلومات للاختبار تقدم أكبر كمية معلومات عند مستويات القدرة المتوسطة.

كما أجرى الخياط (2012) دراسة هدفت إلى التحقق من درجة مطابقة اختبار تحصيلي، وفق أنموذج راش أحادي المعلمة، في الكشف عن مستوى المعرفة العلمية، في المهارات الرياضية لدى طلبة الصف الثامن الأساسي، ولتحقيق هدف الدراسة، تم تطبيق فقرات الاختبار على عينة من طلبة الصف الثامن الأساسي بلغ حجمها(599) طالباً وطالبة، ومن أجل تحليل البيانات، استخدم البرنامج الإحصائي (Rumm2010). وأشارت نتائج الدراسة إلى تمتع الاختبار المقدم بالصدق والثبات؛ فقد بلغت قيمة معامل الثبات للاختبار (0.90). أما معامل الصدق بدلالة المحك بين درجات الطلبة على الاختبار، ودرجاتهم في مادة الرياضيات فقد بلغت (0.87). كما أشارت النتائج أيضا إلى مطابقة (25) فقرة من فقرات الاختبار البالغ عددها (35) فقرة للبيئة المحلية، بحيث تمتعت بخصائص إحصائية مناسبة من حيث معامل صعوبة الفقرة.

وفي دراسة أجراها الربيعي (2012) هدفت إلى بناء اختبار تحصيلي في مادة الأحياء لدى طلبة الصف الأول متوسط وفق أنموذج راش، ومن أجل تحقيق هدف الدراسة قام الباحث ببناء اختبار تحصيلي في مادة الأحياء للصف الأول متوسط مؤلف من (40) فقرة من نوع الاختيار من متعدد، قام الباحث بتوزيعه على عينة مؤلفة من (200) طالباً وطالبة، وحللت نتائج الاختبار إحصائياً باستخدام التحليل العاملي؛ حيث أشارت نتائجه إلى أن الاختبار يحقق افتراض أحادي البعد، ثم تم تحليل البيانات باستخدام أنموذج راش، فقد تم تدريج الفقرات، حيث تم حذف (6) فقرات لم تحقق افتراضات الأنموذج، وأصبح الاختبار في صورته النهائية مكون من (34) فقرة.

أما دراسة الجبوري (2012) فقد هدفت إلى بناء اختبار تحصيلي لمادة الإشراف التربوي، وفق نظرية السمات الكامنة لطلبة معاهد إعداد المعلمين، ومن أجل تحقيق هدف الدراسة، قام الباحث بصياغة (153) فقرة اختبارية من الاختيار من أربعة بدائل، تم تطبيقه على عينة مؤلفة من (310) طالباً وطالبة، ومن أجل تحليل النتائج استخدم الباحث أنموذج راش الذي أشارت نتائجه إلى مطابقة (110) فقرة من فقرات الاختبار فقد تحققت خاصية استقلالية القياس التي تحقق موضوعية القياس، وامتاز الاختبار بالتمييز، حيث اقترب معامل تمييز الاختبار من (1) صحيح، كما بلغ معامل ثبات الاختبار (0.95).

وقامت الحموري (2011) بدراسة هدفت استقصاء إمكان استخدام نماذج الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم، تم بناء اختبار تحصيلي مؤلف من (23) فقرة من نمط أربعة بدائل تقيس العدد نفسه من الأهداف المقصودة طبق الاختبار على عينة مؤلفة من (284) طالبا وطالبة، وحللت البيانات الناتجة من الناتجة من الاستجابة للفقرة وفق نظرية الاستجابة للفقرة. وكشفت نتائج الدراسة عدم إمكانية استخدام هذه النظرية لبناء الاختبارات ذات العدد القليل من الفقرات.

أما شانج وآخرون (Chang et al., 2010) فقد أجرى دراسة هدفت إلى تحليل فقرات اختبار قرائي خاص بقبول طلبة كلية الطب بتايوان باستخدام أنموذج راش، ومن أجل تحقيق هدف الدراسة طبق الباحثون اختباراً مؤلفاً من (100) فقرة على عينة بلغ حجمها (36) فرداً، وأشارت نتائج الدراسة إلى أن متوسط قدرة المفحوصين كانت أعلى من متوسط صعوبة فقرات المقياس، كما أشارت النتائج إلى أن هناك (60) من أصل (100) فقرة دون مستوى قدرة المفحوصين، وعدم فاعلية أنموذج راش المستخدم في الدراسة في تحليل بيانات الاختبار القرائي.

وأجرى ادواردز والكوك (Edwards and Alcock,2010) دراسة هدفت إلى تحليل نتائج اختبار في الرياضيات في بريطانيا وفق أنموذج راش، ومن أجل تحقيق هدف الدراسة قام الباحثان ببناء اختبار مؤلف

من (11) فقرة، وقد تم توزيعه على عينة الدراسة المؤلفة من (164) طالباً وطالبة من طلبة السنة الأولى في المرحلة الجامعية، وقد طلب من أفراد العينة ترتيب فقرات الاختبار حسب مستوى صعوبتها، وبعد ذلك تم تحليل البيانات، التي تم الحصول عليها، باستخدام أنموذج راش، وأشارت النتائج إلى عدم مطابقة فقرات الاختبار لستة أشخاص فقط من عينة الدراسة، بعد إعادة ترتيب فقرات الاختبار حسب أنموذج راش.

وفي دراسة أجراها حمادنه (2009) هدفت إلى بناء اختبار تحصيلي محكي المرجع في الرياضيات وفق الأنموذج اللوجستي ثلاثي المعلم في الأردن، ومن أجل تحقيق ذلك قام الباحث ببناء اختبار تحصيلي في الرياضيات يقيس تحصيل الطالب في الإحصاء مكون من (28) فقرة من نوع الاختيار من متعدد من نوع أربعة خيارات، وقد تم تطبيق الاختبار على عينة مؤلفة من (411) طالباً وطالبة بالصف الأول الثانوي العلمي في الأردن، وقد أظهرت نتائج الدراسة إلى تحقق الاختبار لافتراضات نظرية الاستجابة للفقرة، كما أشارت النتائج إلى مطابقة (24) فقرة من فقرات الاختبار لتوقعات الأنموذج اللوجستي ثلاثي المعلمة، وحذف(4) فقرات لم تطابق الأنموذج، وأن تقديرات معالم الفقرات (الصعوبة والتمييز والتخمين) لها كانت مقبولة ضمن المحكات الواردة في أدبيات القياس التربوي.

وقد قام الحواري وعودة (2008) بدراسة هدفت إلى التعرف إلى الخصائص السيكومترية لصورة مختارة من اختبارات الرخصة الدولية لقيادة الحاسوب في الأردن، ومعادلة درجاته، وباستخدام أنموذج راش، أحادي المعلم، من نماذج نظرية الاستجابة للفقرة، وعلى عينة مؤلفة من (1200) مفحوص اختيروا بالطريقة العشوائية، وزعت على ثلاث مجموعات متكافئة، تم تطبيق نماذج المقاييس الفرعية وهي: المفاهيم الأساسية، واستخدام الحاسوب، وإدارة الملفات، ومعالجة الكلمات، وجداول البيانات، والعروض التقدمية، والمعلومات، والاتصالات؛ وقد أظهرت النتائج أن المعادلة وفق أنموذج راش كانت الأفضل، والأكثر فاعلية في معادلة درجات المفحوصين من خلال وضعهم على تدريج مشترك بالاعتماد على المعايير السابقة.

أما دراسة ويلسون وماك (Wilson and Mac,2007) فقد هدفت إلى قياس المهارات الرياضية باستخدام اختبار موضوعي من نوع الاختيار من متعدد حسب أنموذج راش، ومن أجل تحقيق هدف الدراسة تم تطبيق الاختبار على عينة مؤلفة من (566) طالبا وطالبة، وقد تم إعطاء معظم الطلبة مهارات في الجبر والحساب تعتمد على المهارات الرياضية، واستخدم أنموذج راش عند تحليل البيانات لاعتماد الاختبار، إضافة إلى وصف مستويات الطلبة في المهارات الرياضية، وقد أظهرت النتائج أن درجات الطلبة على المقياس تعتمد على الأكبر سنا ومعرفة بالمهارات الرياضية، إضافة إلى مستواهم السابق في القدرات الرياضية، كما أظهرت نتائج التحليل مطابقة البيانات للأنموذج.

وفي دراسة أجراها كوه، وهسويه، و وانج، ويو، و ونج ، و هسيه , Hsueh, Wang, Sheu, Yu, Wang, Koh & Hsieh, 2006) هدفت إلى التحقق من أن فقرات اختبار (ARAt) وفق نظرية الاستجابة للفقرة وباستخدام أنموذج راش، ومن أجل تحقيق ذلك تم تطبيق الاختبار على عينة مؤلفة من (351) فرداً في تايوان، وقد أظهرت نتائج الدراسة مطابقة (19) فقرة من فقرات الاختبار لأنموذج راش، إضافة إلى دعم النتائج لافتراض أحادية البعد للفقرات السابقة، كما أظهرت النتائج عدم مطابقة (4) فقرات من فقرات الاختبار لأنموذج راش.

التعليق على الدراسات السابقة:

يلاحظ من خلال عرض الدراسات السابقة، أن غالبية الدراسات استخدمت أنموذج راش في بناء الاختبارات مثل دراسات المصري (2015)، والعنزي (2014)، وهادي، ومراد (2013)، والسامرائي والخفاجي (2012) . كما أشارت نتائج بعض هذه الدراسات إلى مطابقة الاختبارات التحصيلية الى افتراضات أنموذج راش، مثل دراسة الطراونه (2016)، ودراسة علي (2012) ودراسة الربيعي (2012).

كما يلاحظ أيضا من خلال الدراسات السابقة، اختلاف في حجم العينات التي تم تطبيق الاختبارات عليها، غير أن غالبية الدراسات السابقة زاد حجم العينة فيها عن (200) فرد مثل دراسة الجبوري (2012)، ودراسة حمادنه (2009)، ودراسة ويلسون وماك (Wilson and Mac,2007)، ودراسة كوه وآخرين (Koh et al., 2006).

نجد أيضا تفاوت في الدراسات في عدد فقرات الاختبار التحصيلي فبعضها زاد عدد فقرته على (100) فقرة، مثل دراسة هادي، ومراد (2013)، ودراسة السامرائي والخفاجي (2012)، وبعضها الآخر قل عدد فقرات الاختبار التحصيلي فيها عن (30) فقرة مثل دراسة كوه وآخرين (Koh et al., 2006)

وتتميز الدراسة الحالية، عن باقي الدراسات، بأنها تعد الدراسة الأولى(في حدود علم الباحث)، التي تناولت موضوع توظيف أنموذج راش في بناء اختبار محكي المرجع في منهاج التاريخ بفلسطين. وقد استفادت الدراسة الحالية، من الدراسات السابقة، في تحديد المنهج المناسب، حيث إن جميع الدراسات، التي استخدمت أنموذج راش، استخدمت المنهج الوصفي، بالإضافة إلى الاستفادة في بناء الاختبار التحصيلي، والرجوع إلى المراجع ذات العلاقة بموضوع الدراسة، وتحديد المعالجات الإحصائية المناسبة، وطريقة عرض نتائج الدراسة.

المنهجية والإجراءات:

مجتمع الدراسة:

تكون مجتمع الدراسة من جميع طلبة الصف الثاني الثانوي الأدبي، في مدينة نابلس وقد بلغ عددهم (2541) طالباً وطالبة، بواقع (1015) طالباً، و(1526) طالبة، خلال العام الدراسي (2016/2015).

عينة الدراسة:

تكونت عينة الدراسة من (327) طالبا وطالبة، من طلبة الصف الثاني الثانوي الأدبي، في مدينة نابلس، وقد تم تحديد حجم العينة بناءً على معادلة ستيفن ثامبسون، كما جاءت في (بشماني، 2014)، والتي تنص على:

n	حجم العينة
N	حجم المجتمع
z	الدرجة المعيارية المقابلة لمستوى الدلالة 0.95 وتساوي 1.96
d	نسبة الخطأ وتساوي 0.05
p	نسبة توافر الخاصية والمحايدة = 0.50

وقد تم اختيار عينة الدراسة بناءً على طريقة العشوائية العنقودية من (12) مدرسة ثانوية للذكور والإناث من مدارس مديرية التربية والتعليم في مدينة نابلس، وقد تم اختيار جميع الطلبة في هذه المدارس بواقع (193) طالبة، و(134) طالبا.

أداة الدراسة:

قام الباحث ببناء اختبار تحصيلي لمنهاج التاريخ للصف الثاني الثانوي الأدبي، مستنداً على الخطوات العلمية المتبعة في بناء الاختبارات . وقد اعتمد الباحث، في بناء الاختبار، على الخطوات العلمية التي تناولها الإطار النظري ذو العلاقة بموضوع إعداد الاختبار التحصيلي وفق نظرية القياس التقليدية ونظرية القياس الحديثة مثل العتيبي (2014)، وعودة (2010)، والقدومي (2008)، وعلام (2007)، وعلام (2006)، والعبادي (2006)، وعلام (2002)، وفيما يلي عرضا لهذه الخطوات:

أولاً: تحديد محتوى الاختبار:

قام الباحث بتحديد محتوى الاختبار من خلال تحليل عناصر محتوى منهاج التاريخ للصف الثاني الثانوي الأدبي المقرر، من خلال وزارة التربية والتعليم الفلسطينية، وقد شمل المنهاج ثلاث وحدات دراسية وهي:

الوحدة الأولى: التغلغل والتنافس الأجنبي على الوطن والعالم في القرن العشرين.

الوحدة الثانية: الحربان العالميتان الأولى والثانية.

الوحدة الثالثة: المنظمات والتكتلات والأحلاف الدولية والإقليمية بعد الحرب العالمية الثانية.

ثانياً: صياغة الأهداف السلوكية:

قام الباحث بالاطلاع على الأهداف العامة الواردة في دليل المعلم لمنهاج التاريخ للصف الثاني الثانوي الأدبي، وبناء على الأهداف العامة، قام الباحث بصياغة الأهداف السلوكية لكل فصل من فصول المنهاج، وبحسب المستويات المعرفية الأولى، ضمن تصنيف بلوم للأهداف التربوية، في المجال المعرفي، وهي (تذكر، فهم واستيعاب، وتطبيق) وقد بلغ مجموع الأهداف السلوكية، التي تم صياغتها (145) هدفاً، ومن أجل التحقق من سلامة صياغة هذه الأهداف، قام الباحث بعرضها على لجنة متخصصة في المجال التربوي، وبخاصة القياس والتقويم والمناهج وطرق التدريس، ومدرسي منهاج التاريخ للصف الثاني الثانوي الأدبي والذي بلغ عددهم(15) محكماً، وقد تم تعديل صياغة الأهداف السلوكية لتتناسب مع طبيعة المحتوى فقد تم حذف (12) هدفا وإضافة (3) أهداف وإعادة صياغة (19) هدفاً فأصبح عدد الأهداف (136) هدفاً.

ثالثاً: إعداد جدول المواصفات:

في هذه الخطوة قام الباحث ببناء جدول المواصفات، وذلك بناء على عناصر المحتوى والأهداف السلوكية والقيم الوزنية لكل منهما، والجدول الآتي يوضح جدول المواصفات للاختبار التحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي:

جدول (1): جدول المواصفات الخاص بالاختبار التحصيلي لمنهاج التاريخ للصف الثاني الثانوي الأدبي

مستوى / الهدف الوحدة	معرفه	فهم	تحليل	المجموع	النسبة المئوية
الوحدة الأولى	29	14	14	57	0.42
الوحدة الثانية	12	15	15	42	0.31
الوحدة الثالثة	18	6	13	37	0.27
المجموع	59	35	42	136	1
النسبة المئوية	0.43	0.26	0.31	1	

رابعاً: بناء فقرات الاختبار التحصيلي:

في هذه الخطوة قام الباحث بإعداد الصيغة الأولية لفقرات الاختبار التحصيلي، وهي من نوع الاختيار من متعدد، وقد اعتمد الباحث هذا النمط من الأسئلة لشيوع استخدامها في الدراسات التي هدفت إلى توظيف أنموذج راش في بناء اختبار تحصيلي، وكذلك للموضوعية التي يحققها هذا النوع من الاختبار، وبخاصة في مجال تصحيحها. وقد قام الباحث بصياغة فقرات الاختبار من نوع الاختيار من أربعة بدائل، معتمداً في ذلك، على جدول المواصفات السابق؛ وقد تم صياغة (136) فقرة من نوع الاختيار من متعدد، وذلك بناءً على محتوى المنهاج وأهدافه السلوكية، وتم الأخذ بعين الاعتبار معايير صياغة فقرات الاختبارات من نوع الاختيار من متعدد الجيدة، وذلك كما جاءت في الأدب التربوي، وبعض الدراسات السابقة مثل دراسة ظاظا (2012).

خامساً: تحكيم المحتوى والأهداف السلوكية والفقرات:

من أجل التحقق من صلاحية الفقرات، قام الباحث بعرضها في صيغتها الأولية على لجنة من المتخصصين في مجال القياس والتقويم، ومناهج وطرق التدريس والتاريخ بلغ عددهم (10) محكمين من حملة الدكتوراة في القياس والتقويم ومناهج وطرق التدريس والتاريخ، كما تم عرض الاختبار على (5) مدرسين ممن يدرسون منهاج التاريخ للصف الثاني الثانوي الأدبي، وقد قام الباحث بتعديل وحذف واعادة بعض الفقرات حتى أصبح الاختبار في صيغته الاولية مؤلفاً من (136) فقرة.

سادساً: إعداد الاختبار في صيغته الأولية:

تكون الاختبار، في صيغته الأولية، من (136) فقرة موزعة على ثلاثة مستويات معرفية، وهي: تذكر، وفهم واستيعاب، وتطبيق، وقد اتبع الباحث التسلسل المنطقي لفقرات الاختبار حسب المحتوى التعليمي، وكان الهدف من ذلك تسلسل فقرات الفصل الواحد، وتجميع فقرات الفصل، وتتابع فصول منهاج التاريخ للصف الثاني الثانوي الأدبي لسهولة استرجاع المعلومات بالنسبة للطلبة خوفا من التشتت، وزيادة قلق الاختبار خلال تنقل الطالب في الإجابة عن فقرات الاختبار.

سابعاً: إعداد كراسة الاختبار وتعليمات الإجابة على ورقة الإجابة:

قام الباحث بإخراج الاختبار بحيث تضمن ثلاثة أقسام؛ تضمن القسم الأول تعليمات الاختبار، ومن أهم هذه التعليمات الإجابة عن فقرات الاختبار جميعها، ونقل الإجابة إلى ورقة الإجابة المخصصة في نهاية الاختبار، وزمن الاختبار. أما القسم الثاني، فقد تضمن فقرات الاختبار من نوع الاختيار من أربعة إجابات، والمكون من (136) فقرة، في حين تضمن القسم الثالث ورقة الإجابة للاختبار.

ثامناً: التجربة الاستطلاعية:

قام الباحث، بهذه الخطوة، من أجل التعرف إلى مدى وضوح التعليمات، وطريقة الإجابة، ومتوسط الزمن اللازم للإجابة عن الاختبار، كذلك التأكد من وضوح صياغة فقرات الاختبار للفئة المستهدفة، ومن أجل تحقيق هذه الأهداف، قام الباحث بتطبيق الاختبار على شعبتين صفيتين بواقع (60) طالباً وطالبة من مجتمع الدراسة، ولكن من خارج عينة الدراسة الفعلية، وقد قام الباحث بتوزيع كراسة الاختبار على الطلبة، وطلب منهم قراءة تعليمات الاختبار جيداً قبل البدء بالإجابة عن فقرات الاختبار، وبعد التأكد من وضوح التعليمات، وطريقة الإجابة طلب الباحث من الطلبة البدء بالإجابة عن فقرات الاختبار حتى يستطيع تحديد متوسط الزمن اللازم لتطبيقه، وقد تم تسجيل الزمن الذي استغرقه كل طالب للإجابة، فقد تراوح بين (60 دقيقة الى 120 دقيقة) وبمتوسط (90) دقيقة.

تاسعاً: تجربة التحليل الإحصائي:

في هذه المرحلة قام الباحث بتنفيذ الخطوات الآتية :

1. اختيار عينة التحليل الإحصائي، فقد تكونت عينة الدراسة من (327) طالبا وطالبة.
2. توزيع الاختبار على عينة الدراسة الفعلية.

3. تصحيح أوراق الإجابة، وفي هذه الخطوة قام الباحث بتصميم مفتاح تصحيح مثقب من أجل تصحيح أوراق الإجابة، وقد تم التصحيح بإعطاء علامة واحدة(1) للإجابة الصحيحة، وصفر(0) للإجابة الخاطئة، وبذلك تمثل علامة الطالب النهائية مجموع العلامة التي يحصل عليها في الاختبار (المجموع الكلي لعدد فقرات الاختبار التي أجاب عنها الطالب إجابة صحيحة).

4. إدخال العلامات إلى الحاسب الآلي.

5. تحقيق افتراض أحادية البعد.

نتائج الدراسة ومناقشتها:

أولاً : النتائج المتعلقة بالسؤال الأول الذي ينص على:

ما مدى مطابقة فقرات الاختبار التحصيلي محكي المرجع مع إفتراضات أنموذج راش؟

من أجل الإجابة عن هذا السؤال، تم التحقق من إفتراضات أنموذج راش، حيث قام الباحث بإدخال البيانات إلى البرمجية الإحصائية (wensteps)، فقد تم إدخال البيانات المتعلقة باستجابات أفراد العينة والبالغ عددهم (327) طالبا وطالبة عن فقرات الاختبار التحصيلي والمكون من (136) فقرة وذلك وفق تصحيح الإجابات الصحيحة والخاطئة التي اعتمدت (صفر) للإجابة الخاطئة و(واحد) للإجابة الصحيحة، وقد تم استخراج نتائج الدراسة للسؤال الأول حسب الخطوات الآتية:

1. مؤشرات حسن المطابقة الخاصة بأفراد عينة الدراسة (Persons Fit):

من أجل التعرف إلى مؤشرات مطابقة أفراد عينة الدراسة، تم تحديد قدرة كل فرد من أفراد العينة، كما تم حساب الخطأ المعياري في قياس قدرة أفراد العينة، كذلك قام الباحث بحساب إحصائي المطابقة الداخلية (INFIT) ويُعدّ هذا المؤشر مؤشراً إحصائياً لسلوك أفراد العينة غير المتوقع الذي يؤثر في الاستجابات عن الفقرات التي تكون قريبة من مستوى معين، كما تم حساب مؤشر إحصائي المطابقة الخارجية (OUTFIT) وذلك لكل تقدير من هذه التقديرات، وإحصائي المطابقة الخارجية يعد مؤشراً أكثر حساسية للسلوكيات غير المتوقعة من الأفراد عن الفقرات التي تبتعد عن مستوى قدراتهم، غير أنّ له مؤشرات مشابهة من إحصائي المطابقة الداخلية (INFIT) ونتائج الجدول (2) توضح المتوسط الحسابي والانحراف المعياري والخطأ المعياري لكل تقدير من تقديرات قدرات أفراد عينة الدراسة، في قياس قدراتهم، وكذلك متوسطات المربعات لإحصائي المطابقة الخارجية والداخلية وقيم إحصائيات المطابقة (ZSTD) لإحصائي المطابقة الداخلية والخارجية.

جدول (2): المتوسط الحسابي والانحراف المعياري لكل تقدير من تقديرات قدرات الأفراد والخطأ المعياري في قدرة الأفراد، ومتوسطات المربعات (MNSQ) لإحصائي المطابقة الداخلية والخارجية وقيم إحصائيات المطابقة (ZSTD) لإحصائي المطابقة الداخلية والخارجية

	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية (INFIT)		إحصائي المطابقة الخارجية (OUTFIT)	
			متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)	متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)
المتوسط الحسابي	1.41	0.25	1	0.00	0.99	0.00
الانحراف المعياري	1.02	0.09	0.05	0.90	0.10	1

يلاحظ، من خلال البيانات الواردة، في الجدول السابق، المتوسط الحسابي لمتوسطات المربعات (MNSQ) الداخلية، والخارجية فقد كانت على الترتيب (1.00) و (0.99)، وهذه القيم قريبة من الواحد صحيح، وهو الوضع المثالي الذي يتوقعه الأنموذج، كما نلاحظ أيضا من خلال البيانات الواردة في الجدول السابق، أن متوسط قيم إحصائي المطابقة الداخلية والخارجية للإحصائي (ZSTD) قد بلغت (0.00)، وهذه القيمة أيضا تقترب من القيم المثالية التي يفترضها الأنموذج وهي (صفر، 1) على الترتيب، وبعد تفحص قيم إحصائي المطابقة الخارجية الموزون للأفراد فقد تبين أن (25) فرداً من أفراد عينة الدراسة غير مطابقين للأنموذج، وذلك حيث تجاوزت قيم المطابقة الداخلية والخارجية لهم عن (2+) أو(2−)، فقد قام الباحث باستبعادهم من التحليل، وهذا مؤشر للباحث أن قيم متوسط المربعات المناظرة لقدراتهم تزيد عن الواحد صحيح أو أن قيم إحصائي المطابقة الخارجية المقابلة لقدراتهم تزيد عن (2+) وذلك كما أشار كلٌ من الأستير، وهيتشينسو (Alastair and Hutchinson,1987) بأنه إذا كانت قيمة المؤشرات الإحصائية تزيد أو تنقص عن (+2) أو (2−) فإن قدرة الفرد تعد غير متطابقة مع قدرات مجموعة الأفراد، وبناءً على ذلك، فان هؤلاء الأفراد غير مطابقين للأنموذج، ويجب استبعادهم واستكمال التحليل. والجدول (3) يوضح بيانات الطلبة غير المطابقين للأنموذج:

جدول (3): إحصائي المطابقة الداخلية والمطابقة الخارجية للأفراد غير المطابقين مع أنموذج راش

رقم الطالب	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية INFIT		إحصائي المطابقة الخارجية OUTFIT	
			MSQ متوسط المربعات	ZSTD قيمة الإحصائي	MSQ متوسط المربعات	ZSTD قيمة الإحصائي
13	0.15	0.17	0.9	−3.7	0.89	−3.72
2	0.18	0.17	0.9	−3.44	0.89	−3.48
54	0.27	0.18	0.9	−3.09	0.89	−3.08
5	0.4	0.18	0.88	−3.03	0.86	−3.17
19	0.06	0.17	0.92	−2.93	0.92	−2.93
56	−0.06	0.17	0.93	−2.83	0.92	−2.82
29	0.27	0.18	0.91	−2.82	0.9	−2.83
49	−0.27	0.18	0.91	−2.79	0.91	−2.45
58	0.43	0.18	0.88	−2.77	0.87	−2.87
59	0.15	0.17	0.92	−2.69	0.92	−2.74
110	0.4	0.18	0.9	−2.56	0.89	−2.65
52	−0.18	0.17	0.93	−2.51	0.92	−2.5
125	0.15	0.17	0.93	−2.4	0.93	−2.48
22	−0.43	0.18	0.9	−2.39	0.89	−2.48
121	0.33	0.18	0.91	−2.38	0.9	−2.47
62	0.21	0.17	0.93	−2.13	0.93	−2.13
16	0.55	0.18	0.89	−2.11	0.87	−2.43
15	0.03	0.17	0.95	−2.01	0.94	−2.03
130	0.69	0.18	1.13	2.05	1.17	2.38
195	0.46	0.18	1.1	2.09	1.11	2.17
205	−0.59	0.18	1.12	2.2	1.15	2.4
250	0.15	0.17	1.07	2.36	1.08	2.48
249	−0.03	0.17	1.07	2.48	1.07	2.52
89	−0.3	0.18	1.15	4.1	1.17	4.24
104	0	0.17	1.13	4.66	1.13	4.43

يلاحظ، من خلال البيانات الواردة، في الجدول السابق، قيم المطابقة الداخلية (INFT) للأفراد غير المطابقين مع بيانات أنموذج راش فقد تراوحت قيمة الإحصائي ZSTD عليها ما بين (3.7- 4.66)، كما نلاحظ أيضا، من خلال الجدول السابق، قيم إحصائي المطابقة الخارجية للأفراد غير المطابقين مع بيانات أنموذج راش فقد تراوحت قيمة الإحصائي ZSTD عليها ما بين (3.72- 4.43).

١. مؤشرات حسن المطابقة الخاصة بمعالم صعوبة فقرات القائمة الدراسة (Item- Fit) بعد استبعاد الأفراد غير المطابقين للأنموذج:

لقد تم استبعاد (25) طالباً وطالبة، وهم الطلبة الذين لم تتطابق استجاباتهم مع توقعات الأنموذج فقد قام الباحث بإعادة التحليل الإحصائي مرة أخرى لمعرفة مدى مطابقة فقرات القائمة للأنموذج (Item-Fit)، قد جرى تقدير معلم الصعوبة لكل فقرة والخطأ المعياري في قياس هذا المعلم، كما تم استخراج إحصائي المطابقة الداخلية، وإحصائي المطابقة الخارجية، وذلك لكل معلم من معالم الصعوبة، والجدول (4) يوضح المتوسط الحسابي والانحراف المعياري لكل من تقديرات معالم الصعوبة والخطأ المعياري في قياس هذه التقديرات، بالإضافة إلى الإحصاءات المطابقة الداخلية والخارجية للفقرات.

جدول (4): المتوسط الحسابي والانحراف المعياري لكل تقدير من تقديرات معالم والخطأ المعياري في معالم صعوبة فقرات الاختبار التحصيلي، ومتوسطات المربعات (MNSQ) لإحصائي المطابقة الداخلية والخارجية وقيم إحصائيات المطابقة (ZSTD) لإحصائي المطابقة الداخلية والخارجية

	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية (INFIT)		إحصائي المطابقة الخارجية (OUTFIT)	
			متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)	متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)
المتوسط الحسابي	0.00	0.14	1.00	0.00	0.99	0.0
الانحراف المعياري	0.32	0.01	0.06	0.9	0.13	1.00

يلاحظ، من خلال البيانات الواردة، في الجدول السابق، أن المتوسط الحسابي لمتوسطات المربعات، كانت قريبة جداً من الواحد صحيح، وهذا هو الوضع المثالي كما يتوقعه الأنموذج، كما نلاحظ أن متوسط قيم إحصائي المطابقة الداخلية بلغ (0.00)، والانحراف المعياري (0.9) وهذه القيمة قريبة من القيم المثالية

التي يفترضها الأنموذج، وهي (0، 1) على الترتيب، وتبين أيضا، من خلال الجدول السابق، أن متوسط قيم إحصائي المطابقة الخارجية، بلغ (0.0) والانحراف المعياري لها بلغ (1.0)، وتعد هذه القيم مثالية، وتقترب من القيم المثالية التي يفترضها الأنموذج.

وبعد تفحص قيم إحصائي المطابقة الداخلية والخارجية، لفقرات الاختبار التحصيلي، بصورته الأولية البالغة (136) فقرة، تبين أن عدد الفقرات غير المطابقة للأنموذج بلغت (4) فقرات، فقد تجاوزت قيم إحصائي المطابقة لها (+2) و (−2)، كما تجاوزت قيم متوسط المربعات لها الواحد صحيح، وهذا مؤشر أن هذه الفقرات غير مطابقة لتوقعات الأنموذج (Linacre and Wright,1993) والفقرات التي تجاوزت قيم إحصائي المطابقة عن (+2) و(−2) كانت على التوالي (133 ,126 ,135, 1 ,136)، وبناء على عدم مطابقة هذه الفقرات للأنموذج، فقد تم استبعادها من التحليل. والجدول (5) يوضح نتائج الفقرات غير المطابقة للأنموذج:

جدول (5):إحصائي المطابقة الداخلية والمطابقة الخارجية للفقرات غير المتطابقة مع أنموذج راش

رقم الفقرة	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية INFIT		إحصائي المطابقة الخارجية OUTFIT	
			متوسط المربعات MSQ	قيمة الإحصائي ZSTD	متوسط المربعات MSQ	قيمة الإحصائي ZSTD
136	0.57	0.14	1.25	3.84	1.53	4.69
1	1.19	0.13	1.11	2.53	1.15	2.08
135	0.53	0.14	1.14	2.17	1.26	2.47
126	0.15	0.15	1.16	1.96	1.3	2.23
133	0.15	0.15	1.15	1.8	1.33	2.4

يلاحظ، من خلال البيانات الواردة، في الجدول السابق، قيم المطابقة الداخلية (INFT) لفقرات الاختبار التحصيلي في التاريخ غير المتطابقة مع بيانات أنموذج راش، وقد تراوحت قيمة الإحصائي ZSTD عليها بين (1.8− 3.84)، كما نلاحظ أيضا، من خلال الجدول السابق، قيم إحصائي المطابقة الخارجية للفقرات غير المتطابقة مع بيانات أنموذج راش، فقد تراوحت قيمة الإحصائي ZSTD عليها بين (2.4 − 4.69)، ويظهر من خلال البيانات أيضا، أن قيم المطابقة الداخلية للفقرتين (126, 133) كانت أقل من (2) وهي ضمن القيم المطابقة مع بيانات الأنموذج غير أن قيم المطابقة الخارجية لها كانت أكبر من (2) وبناء على ذلك تم استبعاد هذه الفقرات.

2. التحقق من افتراض أحادي البعد:

تركز نظرية الاستجابة للفقرة على توافر عدة افتراضات في أداء الأفراد المراد تطبيق أداة الدراسة عليهم، ومن هذه الافتراضات، افتراض أحادية البعد؛ بمعنى وجود قدرة واحدة تفسر أداء الأفراد المراد تطبيق أداة الدراسة عليهم، ولهذا سميت هذه النظرية بنماذج أحادية البعد، ومن أجل التحقق من ذلك، استخدم الباحث التحليل العاملي باستخدام طريقة المكونات الأساسية، التي تعد من أكثر الطرق استخداماً في التحقق من فرض أحادي البعد في نظرية الاستجابة للفقرة، ومن أجل استخراج التحليل العاملي للاختبار التحصيلي، تم حساب المتوسطات الحسابية لكل مستوى من مستويات الأهداف لكل وحدة من وحدات مقرر التاريخ للصف الثاني الثانوي الأدبي، وذلك حسب ما تم إعدده في جدول المواصفات، وبعد ذلك استخدم الباحث البرنامج الإحصائي للعلوم الاجتماعية (SPSS)، وقد أشارت نتائج التحليل العاملي إلى وجود عامل واحد، وبعد تدوير العامل على محاور متعامدة بطريقة الفايرمكس (Varimax) لكايزر تم الحصول على عاملين اثنين فسر العامل الأول منهما عاملاً عاماً واحداً، وبذلك تكون نتائج التحليل العاملي قد أفرزت عاملاً عاماً يقيسه الاختبار التحصيلي بشكل أساسي، ويتضح ذلك من خلال تشبع المستويات بالعامل العام، وبذلك يكون قد تحقق فرض أحادي البعد في الاختبار التحصيلي لمنهاج التاريخ للصف الثاني الثانوي، الذي يعد أحد فروض نظرية الاستجابة للفقرة والجدول (6) يوضح نتائج التحليل العاملي للاختبار التحصيلي:

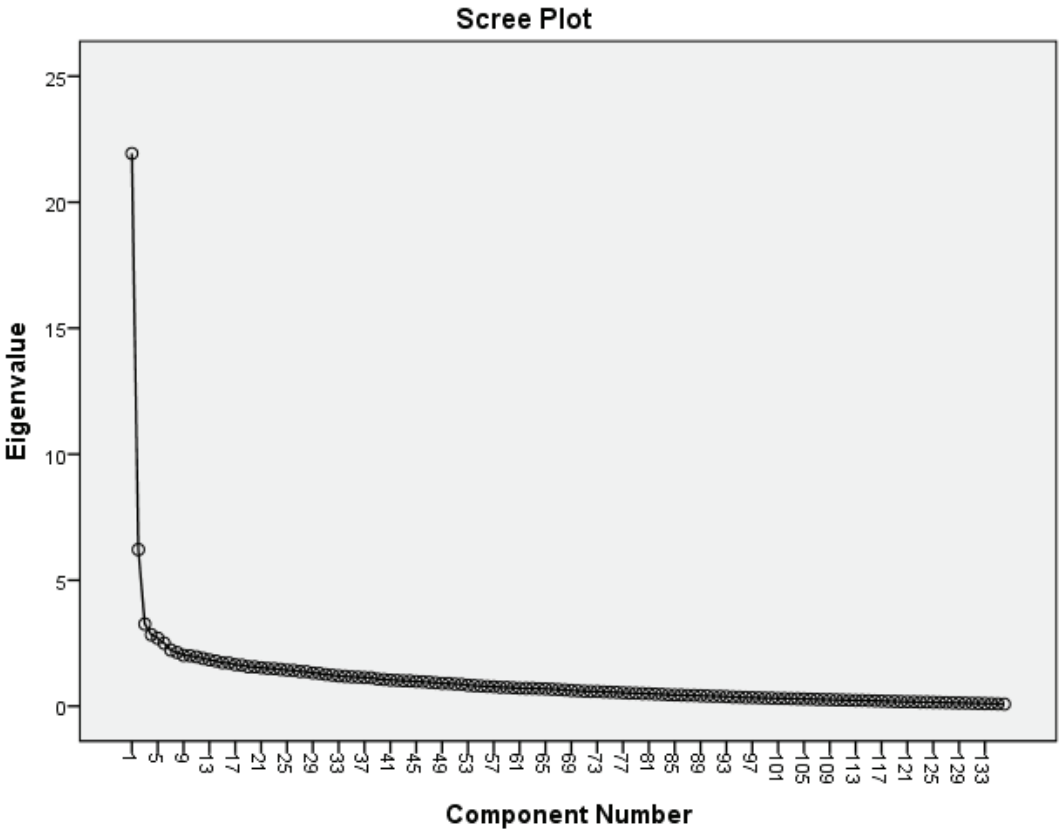
جدول (6): نتائج التحليل العاملي لإستجابات أفراد عينة الدراسة (327)طالبا وطالبة على الاختبار التحصيلي (136) فقرة.

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي	العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
1	21.93	16.13	16.13	23	1.49	1.10	51.22
2	6.21	4.57	20.70	24	1.46	1.07	52.30
3	3.26	2.40	23.09	25	1.43	1.05	53.35
4	2.84	2.09	25.18	26	1.42	1.04	54.39
5	2.70	1.98	27.16	27	1.37	1.01	55.40
6	2.50	1.84	29.00	28	1.37	1.01	56.41

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي	العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
7	2.23	1.64	30.64	29	1.32	0.97	57.38
8	2.13	1.57	32.21	30	1.30	0.96	58.33
9	2.01	1.48	33.69	31	1.26	0.93	59.26
10	2.00	1.47	35.16	32	1.23	0.90	60.16
11	1.96	1.44	36.60	33	1.20	0.88	61.05
12	1.89	1.39	38.00	34	1.19	0.87	61.92
13	1.83	1.35	39.35	35	1.17	0.86	62.78
14	1.79	1.32	40.66	36	1.15	0.85	63.63
15	1.73	1.27	41.93	37	1.14	0.84	64.46
16	1.72	1.26	43.20	38	1.13	0.83	65.29
17	1.65	1.21	44.41	39	1.08	0.79	66.08
18	1.63	1.20	45.60	40	1.06	0.78	66.86
19	1.57	1.15	46.75	41	1.04	0.76	67.62
20	1.56	1.15	47.90	42	1.03	0.76	68.38
21	1.53	1.12	49.02	43	1.01	0.74	69.12
22	1.50	1.10	50.12	44	1.00	0.74	69.86

نلاحظ، من خلال البيانات الواردة في الجدول السابقة، أن قيم جذور العوامل الكامنة التي زادت على الواحد صحيح فسرت ما نسبته (69.86%) من تباين استجابات أفراد عينة الدراسة، وقد بلغت قيمة الجذر الكامل للعامل الأول (21.93)، وتُعدُّ هذه القيمة مرتفعة إذا ما قورنت بالعامل الثاني الذي بلغ (6.21)، أما بقية الجذور الكامنة للعوامل الأخرى فقد كانت منخفضة ومتقاربة، وهذا يعد مؤشراً على وجود عامل سائد يرجح أحادية البعد لأغراض التحليل باستخدام نماذج الاستجابة للفقرة، أما بالنسبة للتباين الذي فسره كل عامل فنجد ان نسبة التباين المفسره بالعامل الأول بلغت (16.3%) من التباين الكلي، بينما كانت نسبة التباين المفسرة في

العامل الثاني (4.57%) من التباين الكلي، ويلاحظ ان نسبة التباين المفسرة في العامل الأول عالية بالمقارنة مع نسبة التباين المفسرة في العامل الثاني، كما يلحظ ان الفروق في نسبة التباين المفسرة لكل العوامل المتبقية قليلة جداً ومتقاربة وهذا يؤكد تحقق أحادية البعد، وقد أشار في هذا المجال كل من هامبلتون وسوامينثون (Hambleton and Swaminathan,1985)، وريكاسي (Reckase,1979) أنه إذا بلغت نسبة قيمة الجذر الكامن للعامل الأول الى قيمة الجذر الكامن للعامل الثاني (3.56)، وبما أن هذه النسبة أكبر من (2) فيعد هذا مؤشراً على أن العامل الأول هو عامل مسيطر، وهذا مؤشر يؤكد تحقق افتراض أحادي البعد للاختبار التحصيلي. ويتعزز افتراض أحادية البعد للاختبار من خلال تمثيل الجذور الكامنة بيانياً باستخدام ما يعرف بمخطط ساكري (Scree Plot) الذي يظهر في الشكل الآتي:



الشكل (1): التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار التحصيلي على البيانات الكلية.

يتبين من الشكل السابق رقم(١) أن الجذر الكامن للعامل الأول يطغى، بشكل واضح، على الجذور الكامنة لبقية العوامل، وهذا مؤشر أيضا على أحادية البعد لبيانات الاختبار.

3. التحقق من افتراض الاستقلال الموضعي:

أشار هامبلتون وسواميناثان (Hambleton &swaminathan,1985) إلى أن افتراض الاستقلال الموضعي في الاختبار يتحقق من خلال تحقق افتراض أحادي البعد، وهذا يعني أن افتراض الاستقلال الموضعي قد تحقق في هذا الاختبار.

4. تساوي معاملات التمييز:

قام الباحث بحساب معاملات التمييز لفقرات الاختبار، من خلال حساب معامل الارتباط الثنائي، بين أداء الطلبة، على كل فقرة وأدائهم على الاختبار ككل، وهو يُعدُ مؤشراً لدرجة تمييز الفقرة والجدول (7) يبين نتائج ذلك:

جدول (7): معاملات التمييز لفقرات الاختبار

رقم الفقرة	معامل التمييز	رقم الفقرة	معامل التمييز	رقم الفقرة	معامل التمييز	رقم الفقرة	معامل التمييز
1	0.3	35	0.52	69	0.35	103	0.28
2	0.33	36	0.44	70	0.36	104	0.36
3	0.3	37	0.46	71	0.27	105	0.28
4	0.37	38	0.47	72	0.26	106	0.36
5	0.47	39	0.49	73	0.3	107	0.4
6	0.42	40	0.47	74	0.33	108	0.36
7	0.38	41	0.52	75	0.43	109	0.28
8	0.33	42	0.4	76	0.32	110	0.25
9	0.28	43	0.47	77	0.29	111	0.28
10	0.34	44	0.44	78	0.31	112	0.25
11	0.36	45	0.48	79	0.35	113	0.27
12	0.3	46	0.46	80	0.33	114	0.32
13	0.41	47	0.43	81	0.33	115	0.26

0.26	116	0.36	82	0.52	48	0.45	14
0.27	117	0.29	83	0.39	49	0.41	15
0.29	118	0.34	84	0.48	50	0.43	16
0.27	119	0.33	85	0.37	51	0.42	17
0.26	120	0.36	86	0.5	52	0.43	18
0.27	121	0.29	87	0.42	53	0.4	19
0.26	122	0.32	88	0.42	54	0.43	20
0.27	123	0.31	89	0.43	55	0.49	21
0.29	124	0.38	90	0.31	56	0.42	22
0.25	125	0.29	91	0.4	57	0.38	23
0.28	126	0.28	92	0.38	58	0.48	24
0.26	127	0.29	93	0.39	59	0.49	25
0.28	128	0.29	94	0.4	60	0.52	26
0.27	129	0.35	95	0.41	61	0.46	27
0.32	130	0.28	96	0.41	62	0.43	28
0.36	131	0.33	97	0.33	63	0.41	29
0.36	132	0.3	98	0.33	64	0.42	30
0.25	133	0.29	99	0.4	65	0.44	31
0.25	134	0.33	100	0.35	66	0.43	32
0.43	135	0.32	101	0.25	67	0.48	33
0.26	136	0.4	102	0.4	68	0.42	34

يلاحظ، من خلال البيانات الواردة، في الجدول السابق، أن معاملات التمييز لفقرات الاختبار التحصيلي تراوحت بين (0.52) و (0.25)؛ أي أن مدى معامل التمييز الذي تقع فيه الفقرات هو (0.15)، وحتى يتحقق افتراض تساوي معاملات التمييز ومطابقتها للأنموذج يجب أن تكون جميع قيم معاملات التمييز واقعه بين المدى المبين في المعادلة التالية:

المدى = المتوسط الحسابي لمعاملات التمييز (± 0.15)

بلغ المتوسط الحسابي لمعاملات التمييز (0.37)، وتراوح المدى حدود المطابقة بين (0.52-0.25) وكما نلاحظ، من البيانات في الجدول السابق، فان جميع قيم معاملات الفقرات وضعت ضمن هذا المدى، وهذا مؤشر على تحقق افتراض تساوي معاملات التمييز (Hambleton &swaminathan,1985)

5. التحقق من افتراض التحرر من السرعة:

قام الباحث بالتحقق من هذا الافتراض من خلال إعطاء الطلبة وقتا كافيا للإجابة عن فقرات الاختبار، حيث استطاع جميع الطلبة إنهاء الإجابة عن الاختبار قبل الوقت المحددة وهو (ساعتان) على اعتبار أن جلسة امتحان الطلبة في هذا الصف هي ساعتين، وبعد تدقيق الإجابات تبين أن جميع الطلبة قاموا بالإجابة عن جميع الفقرات، وبذلك يكون الباحث قد تأكد من أن عامل السرعة لا يؤدي دوراً في الإجابة عن فقرات الاختبار، وهذا يعني أن إخفاق الطلبة في الإجابة عن فقرات الاختبار يعود إلى تدني قدراتهم وليس إلى عامل السرعة في الإجابة.

نتائج السؤال الثاني الذي ينص على:

ما تقدير قدرات الأفراد المقابلة لكل درجة كلية محتملة على الاختبار التحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي حسب أنموذج راش؟

قام الباحث باستخدام البرمجية الإحصائية (winsteps)، من أجل الحصول على تقديرات نهائية متحررة لكل من قدرات الأفراد وصعوبة الفقرات ، فقد بلغ عدد الأفراد (302) فرداً، و(131) فقرة، والجدول (8) يبين نتائج تحليل القيم المتحررة لقدرات الأفراد:

جدول (8) : نتائج التحليل للقيم المتحررة لقدرات الأفراد على الاختبار التحصيلي في التاريخ (عدد الأفراد= 302 فرد، عدد الفقرات، 131 فقرة)

	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية (INFIT)		إحصائي المطابقة الخارجية (OUTFIT)	
			متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)	متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)
المتوسط الحسابي	1.55	0.26	1.00	0.00	0.99	0.0
الانحراف المعياري	1.02	0.10	0.03	0.4	0.09	0.6

يَلاحظ، من خلال البيانات، الواردة في الجدول السابق، أن متوسط توزيع القدرات النهائية لقدرات أفراد عينة الدراسة، بعد استبعاد الأفراد غير المطابقين، بلغ (1.55) وحدة لوجت، والانحراف المعياري لها بلغ (1.02) وحدة لوجيت، وهي تقترب من الوضع المثالي للأنموذج، بالإضافة إلى أن هذه القيمة متدنية، وتشير إلى دقة تحديد موقع الأفراد على متصل السمة.

نتائج السؤال الثالث الذي ينص على:

ما تقدير صعوبة فقرات الاختبار التحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي حسب ملاءمتها لأنموذج راش؟

قام الباحث باستخدام البرمجية الإحصائية (winsteps)، من أجل الحصول على نتائج تحليل القيم المتحررة لصعوبة فقرات الاختبار التحصيلي في التاريخ المؤلف من (131) فقرة.، والجدول (9) يبين نتائج تحليل القيم المتحررة لصعوبة فقرات القائمة (131) فقرة.

جدول (9) : نتائج تحليل القيم المتحررة لصعوبة فقرات الاختبار التحصيلي في التاريخ (عدد الأفراد302 فرداً، و131 فقرة)

	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية (INFIT)		إحصائي المطابقة الخارجية (OUTFIT)	
			متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)	متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)
المتوسط الحسابي	0.0	0.10	1.00	0.00	0.99	0.00
الانحراف المعياري	0.27	0.01	0.06	0.6	0.13	0.9

يلاحظ، من خلال البيانات، الواردة في الجدول السابق، أن متوسط القيم التقديرية المتحررة لصعوبة الفقرات تتوزع بمتوسط حسابي قدره (صفر) لوجيت، وانحراف معياري قدرة (0.27) وحدة لوجت، وهذا يعني أن هناك اتساقاً في تدرج صعوبة فقرات الاختبار التحصيلي في التاريخ ، بمعنى أن هذا الاختبار يقيس مدى قدرات متوسطة.

نتائج السؤال الرابع والذي ينص على:

ما مدى صدق الاختبار التحصيلي وثباته في منهاج التاريخ للصف الثاني الثانوي الأدبي باستخدام أنموذج راش؟

وقد تمت الإجابة عن هذا السؤال من خلال استخراج معاملات صدق الاختبار وثباته، وقد تم ذلك من خلال الخطوات الآتية:

أولاً: صدق الاختبار:

تم التحقق من صدق الاختبار من خلال عدة طرق، وهي على النحو الآتي:

1. تتحقق موضوعية القياس بتحقق افتراضات أنموذج راش بصيغته النهائية، كما تتحقق موضوعية تفسير نتائج الاختبار أيضا بالطريقة نفسها، وقد تم التحقق من ذلك من خلال عملية حذف الفقرات غير المطابقة لافتراضات الأنموذج، كما تم حذف الأفراد غير المطابقين مع الأنموذج، ويعد التحقق من افتراضات الأنموذج تقويما لمدى صدق الاختبار في تحقيق موضوعية القياس (Stenner,1990).
2. التحقق من صدق المحكمين للاختبار التحصيلي، وقد تم توضيح ذلك تحت عنوان تحكيم المحتوى والأهداف السلوكية والفقرات، وذلك من خلال عرض الاختبار على لجنة من المحكمين.
3. قام الباحث باستخراج مؤشرات الصدق المتعلقة بالاختبار التحصيلي من خلال حساب معامل الارتباط بيرسون(Person Correlation) بين أداء الأفراد على كل بعد من أبعاد الاختبار التحصيلي، والدرجة الكلية للاختبار وذلك باستخدام برنامج (SPSS) ونتائج الجدول (10) توضح ذلك:

جدول (10): معامل الارتباط بيرسون بين أداء الأفراد على كل بعد من أبعاد الاختبار التحصيلي.

	2	3	4	5	6	7	8	9	10
1	**0.64	**0.62	**0.58	**0.54	**0.49	**0.49	**0.35	**0.32	**0.71
2		**0.75	**0.67	**0.50	**0.47	**0.50	**0.34	**0.32	**0.74
3			**0.67	**0.50	**0.48	**0.49	**0.35	**0.28	**0.73
4				**0.59	**0.47	**0.46	**0.31	**0.32	**0.72
5					**0.72	**0.65	**0.49	**0.50	**0.80

6								**0.76	**0.82
7								**0.69	**0.84
8									**0.73
9									**0.69

** دال إحصائياً عند مستوى الدلالة (0,01).

يلاحظ، من خلال البيانات، الواردة في الجدول السابق، قيم معاملات الارتباط بين، الأداء، على كل بعد من أبعاد الاختبار التحصيلي مع بعضها بعضاً والدرجة الكلية، وكانت هذه المعاملات جميعها دالة إحصائياً، وتشير إلى وجود علاقة ارتباطية قوية فيما بينها، وفقد بلغ مستوى الدلالة الإحصائية عليها جميعاً أقل من (0.01).

ثانياً: ثبات الاختبار:

يشير معامل الثبات، في نظرية الاستجابة للفقرة، إلى دقة تقدير مواقع الأفراد والفقرات على متصل السمة المراد قياسها، ويمكن تحديد مدى هذه الدقة في التقدير بحساب معامل الفصل بين الفقرات (Item Separation Index) الذي يشار له بالرمز (Gi)، قد بلغ (2.7)، فيما بلغت قيمة معامل الفصل بين الأفراد(Person Separation Index) الذي يرمز له بالرمز(3,92) (Gp) وهذه القيم المتعلقة بمعامل الفصل بين الفقرات والأفراد تعبر عن النسبة بين الانحراف المعياري للقيم التقديرية لقدرات الأفراد وصعوبة الفقرات، ومتوسط الخطأ المعياري وزيادة هذه القيم عن (2) يُعدّ دليلاً على أن عينة الأفراد كافية للفصل بين قدرات الأفراد والفقرات، وبناءً على القيم السابقة، يمكن الحصول على معاملات الثبات لكل من الأفراد والفقرات من خلال استخدام الصيغة الرياضية الآتية (G2 1+/R=G2)، حيث تعني (R) معامل الثبات، وتعني(G) معامل الفصل (Masters and Wright,1982) وقد بلغ معامل ثبات الفقرات والأفراد على التوالي(0.73) ،(0.80)، وتشير هذه القيم الى كفاية عينة الأفراد في الفصل بين الفقرات في تعريف متصل السمة الذي تقيس هذه الفقرات، وتدل القيمة الثانية على كفاية عينة الفقرات في الفصل بين الأفراد في التمييز بين أبعاد الاختبار التحصيلي لدى عينة الدراسة من طلبة المرحلة الثانوية في مدينة نابلس. ومن الملاحظ أن قيمة معامل ثبات الأفراد كانت أعلى من قيمة معامل ثبات الفقرات، غير أنها تُعدّ مناسبة.

ومن أجل تحديد عدد الطبقات الإحصائية المتعلقة بالفقرات والأفراد، فقد استخدم الباحث المعادلة الآتية: 3/H=4G+1 ويشير الرمز(H) الى عدد الطبقات الإحصائية لكل من الفقرات والأفراد، وبناء على المعادلة السابقة فقد بلغ عدد الطبقات المتعلقة بالفقرات (3,93)، وتشير هذه القيمة إلى قدرة الفقرات على إظهار

الفروق الفردية بين الأفراد من حيث امتلاكهم للسمة، كما بلغ عدد الطبقات الإحصائية المتعلقة بالأفراد أيضا باستخدام المعادلة نفسها (5.56) وتشير هذه القيمة إلى أن هناك خمسة مستويات لدى الطلبة الذين طبق عيهم الاختبار التحصيلي.

وفي الختام فقد هدفت هذه الدراسة إلى توظيف أنموذج راش في بناء اختبار تحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي بفلسطين، وقد اعتمد الباحث في ذلك على الدراسات السابقة في هذا المجال واتفقت غالبية نتائجها مع نتائج الدراسة الحالية مثل نتائج دراسات الطراونه (2016)، والعنزي (2014)، وهادي ومراد (2013)، والسامرائي والخفاجي (2012)، وتوصلت جميعها إلى أن الاختبارات التحصيلية المستخدمة فيها تمتعت بدرجة مقبولة من الصدق والثبات. كما أنها قامت بحذف مجموعة من الأفراد والفقرات حتى تصبح ملائمة لنموذج راش، واتفقت أيضا مع ما توصلت اليه نتائج دراسة الحموري (2011) في أن نظرية الاستجابة للفقرة غير مناسبة لبناء اختبارات تحصيلية ذات عدد قليل من الفقرات، وخاصة أن الدراسة الحالية اعتمدت على اختبار تحصيلي مؤلف من (131) فقرة.

واختلفت نتائج الدراسة الحالية مع ما توصلت إليه دراسة أون (Onn,2013) التي توصلت إلى تدني معامل الثبات الاختبار التحصيلي المستخدم فيها.

التوصيات والمقترحات:

بناءً على نتائج الدراسة، فإن الدراسة توصي بما يأتي:

1. اعتماد الاختبار التحصيلي بصورته النهائية البالغة (131) فقرة؛ لقياس تحصيل طلبة الصف الثاني الثانوي الأدبي بفلسطين؛ لما يتحقق له من مؤشرات صدق وثبات موثوقة وفق النظرية الحديثة في القياس.
2. إجراء دراسة مقارنة بين نظرية القياس الحديثة والكلاسيكية للتحقق من خصائص الاختبار التحصيلي.
3. التحقق من خصائص الاختبار التحصيلي باستخدام نماذج نظرية الاستجابة للفقرة الأخرى بخاصة الأنموذج ثنائي المعلمة، والأنموذج ثلاثي المعلمة.
4. استخدام أنموذج راش في بناء اختبارات تحصيلية لمناهج دراسية مختلفة للصف الثاني الثانوي الأدبي بفلسطين.

المراجع العربية:

- أبو جراد، حمدي. (2011). مدى التزام المعلمين بتحليل نتائج الاختبارات التحصيلية وعلاقته باتجاهاتهم نحوها. مجلة جامعة الأزهر بغزة، سلسلة العلوم الإنسانية، م(13)، ع(2) ص ص:89-106.
- أبو هاشم، السيد. (2006). مقارنة بين النظرية التقليدية وأنموذج راش في اختيار فقرات قائمة مداخل الدراسة لدى طلاب الجامعة. مجلة كلية التربية، جامعة الزقازيق، ع(52)، ص ص:1-53.
- إسماعيل، ميمي. (2007). الخصائص السيكومترية لاختبار القدرة العقلية باستخدام أنموذج راش لدى طلبة المرحلة الثانوية العامة. رسالة ماجستير غير منشورة، جامعة الزقازيق، الزقازيق، مصر.
- بشماني، شكيب. (2014). دراسة تحليلية مقارنة للصيغ المستخدمة في حساب حجم العينة العشوائية. مجلة جامعة تشرين للبحوث والدراسات العلمية: سلسلة العلوم الاقتصادية والقانونية، م(63)، ع (5)، ص ص: 85-100.
- التقي، أحمد. (2013). النظرية الحديثة في القياس. ط(3)، دار المسيرة للطباعة والنشر، عمان الاردن.
- الثبيتي، سالم. (2014). بناء اختبار تشخيصي مرجعي المحك لقياس مهارات المعلمين في بناء الاختبارات التحصيلية لمقرر الحاسب الآلي. رسالة ماجستير غير منشورة، جامعة أم القرى، مكة ، المملكة العربية السعودية.
- الجبوري، رشيد. (2012). بناء اختبار تحصيلي لمادة الإشراف التربوي على وفق نظرية السمات الكامنة لطلبة معاهد إعداد المعلمين. مجلة الأستاذ، ع(203)، ص ص:1392-1420
- حمادنه، إياد. (2009). استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في الرياضيات وفق الأنموذج اللوجستي ثلاثي المعلم. مجلة العلوم التربوية والنفسية، م(10)، ع(2)، ص ص: 216-230.
- حمادنه، إياد والبلاونه، فهمي. (2015). بناء مقياس اتجاهات معلمي الرياضيات نحو حل المسالة الرياضية وفق نموذج سلم التقدير المنبثق عن نموذج راش. مجلة اتحاد الجامعات العربية للتربية وعلم النفس، م(13)، ع(3)، ص ص: 223-247.
- الحموري، هند. (2011). دراسة استكشافية لملائمة نماذج الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم. مجلة العلوم التربوية والنفسية، جامعة البحرين، م(12)، ع(2)، ص ص: 47-81.
- الحواري، أروى وعودة، أحمد. (2008). الخصائص السيكومترية لصورة مختارة من اختبارات الرخصة

الدولية لقيادة الحاسوب في الأردن، ومعادلة درجاته. المجلة الأردنية للعلوم التربوية، جامعة اليرموك، م(4)، ع(4). ص ص: 297-319.

- الخياط، ماجد. (2012). مطابقة اختبار تحصيلي وفق أنموذج راش أحادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الأساسي. مجلة جامعة الأقصى، سلسلة العلوم الإنسانية، م(16)، ع(1)، ص ص: 87-111.
- الربيعي ، ياسين. (2012). بناء اختبار تحصيلي على وفق أأنموذج راش في مادة الأحياء لدى طلبة الصف الأول متوسط . مجلة الباحث، جامعة كربلاء، م(2)، ع(2)، ص ص:211-256.
- زكريا، علي. (2009). الخصائص السيكومترية لاختبار (اوتيس- لينون) للقدرة العقلية مقدرة وفق القياس الكلاسيكي وأنموذج راش. رسالة ماجستير غير منشورة ، جامعة أم القرى، مكة، المملكة العربية السعودية.
- السامرائي، محمد والخفاجي، أحمد. (2012). بناء اختبار تحصيلي محكي المرجع في مادة علم النفس لطلبة أقسام العلوم التربوية والنفسية. مجلة الأستاذ، ع(203)، ص ص: 964-1002
- السحيمات، منير. (2007). بناء اختبار تحصيلي في الرياضيات متعدد المستويات للصفوف الثامن والتاسع والعاشر الأساسي في الأردن. رسالة ماجستير غير منشورة، جامعة مؤته، مؤته، الأردن.
- الشرقاوي، أنور، والشيخ سليمان، وكاظم، أمينة، وعبد السلام، نادية. (1996). اتجاهات معاصرة في القياس والتقويم النفسي والتربوي. مكتبة الأنجلو المصرية، القاهرة، مصر.
- الشواورة، ياسين. (2013). دراسة مقارنة بين أنموذجي التقدير الجزئي والاستجابة المتدرجة في معادلة درجات الاختبارات. رسالة دكتوراة غير منشورة، معهد الدراسات التربوية، جامعة القاهرة، القاهرة، مصر.
- الطراونه، صبري. (2016). بناء اختبار محكي المرجع في الإحصاء التحليلي لطلبة الدراسات العليا في كلية العلوم التربوية في الجامعات الأردنية الحكومية باستخدام نظرية الاستجابة للفقرة. مجلة العلوم التربوية والنفسية، جامعة البحرين، م(17)، ع(1)، ص ص: 515-542
- ظاظا، حيدر. (2012). الكشف عن مدى انتهاك قواعد صياغة فقرة الاختيار من متعدد في أسئلة الثانوية العامة في الأردن. المجلة الأردنية في العلوم التربوية، م(8)، ع(1)، ص ص:81-91.
- العبادي، رائد. (2006). الاختبارات المدرسية. مكتبة المجتمع العربي، عمان، الأردن.
- عبد المسيح، عماد.(1991). استخدام نموذج راش اللوغريتمي أحادي البارامتر في تحليل مفردات الاختبارات المعرفية مرجعية المعيار ثنائية القطب. مجلة البحث في التربية وعلم النفس، جامعة المنيا، ع(4)، ص ص: 443-475.

- عبد الوهاب، محمد. (2010). استخدام نماذج الاستجابة للمفردة الاختبارية في تدريج مفردات بعض الاختبارات المعرفية. رسالة دكتوراه غير منشورة، جامعة المنيا، المنيا، مصر.
- علام، صلاح الدين. (2006). القياس والتقويم التربوي والنفسي: أسسه وتطبيقاته، وتوجهاته المعاصرة. دار الفكر العربي، القاهرة، مصر.
- علام، صلاح الدين. (2007 أ). الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية. دار الفكر العربي، القاهرة، مصر.
- علام، صلاح الدين. (2007 ب). القياس والتقويم التربوي والنفسي: أسسه وتطبيقاته، وتوجهاته المعاصرة. دار الفكر العربي، القاهرة، مصر.
- علام، صلاح. (2005). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. دار الفكر العربي، القاهرة، مصر.
- علي، نداء. (2012). فاعلية استخدام أنموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم. رسالة دكتوراة غير منشورة، جامعة دمشق، دمشق، سوريا.
- العنزي، سامي. (2014). علاقة الحكمة الاختبارية بتحصيل طلبة المرحلة الثانوية في دولة الكويت في اختبار تحصيلي مبني وفق أنموذج راش. مجلة عالم التربية، ع(47)، ص ص: 345-369.
- عودة، أحمد.(2010). القياس والتقويم في العملية التدريسية. دار الأمل للنشر والتوزيع، إربد، الأردن.
- القدومي، عبد الناصر. (2008). الاختبارات التحصيلية وطرق إعدادها. جامعة النجاح الوطنية، مركز القياس والتقويم.

https://www.najah.edu/media/cms_page_media/588/lecture2.pdf.

- المصري، محمد. (2015). الخصائص السيكومترية لاختبار محكي المرجع في البحث التربوي وفق نظرية القياس الحديثة. مجلة كلية التربية الأساسية، الجامعة المستنصرية، م(21)، ع(89)، ص ص: 701-728.
- هادي، فوزية ومراد، صلاح. (2013). استخدام أنموذج راش في إعداد وتقنين اختبار في اللغة العربية لقبول طلبة الدراسات العليا لكليات العلوم الإنسانية في جامعة الكويت. المجلة التربوية، ع(106)، ج(1)، ص ص: 15-53.

المراجع الأجنبية:

- Avery L, Russell D, Raina P, Walter S, Rosenbaum P. (2003). Rasch analysis of the Gross motor function measure: Validating the assumption of the Rasch model to create an interval – level measure, Arch phys Med Rehebil, Vol. 84, PP. 97-705.
- Baker,F.(2001). The Basics of Item Response Theory (2nd ed.). The ERIC Clearinghouse on Assessment and Evaluation,USA.
- Chan, k. and et. al, (2010). Item analysis for the writing test of Taiwanese board certification examination in anesthesiology using the Rasch model. British Journal of Anaesthesia, vol(104), no(6)m, pp:717-722.
- Domino,G.& Domino,M.(2006): Psychological Testing: An Introduction. (2nd ed.) Library of Congress, New York, USA.
- Edwards, H. and ASlcock, l. (2010). Using Rasch analysis to identify uncharacteristic responses to undergraduate assessment . Teacher Mathematics and its Application,(29), pp:165-175.
- Forbes, D.(1986). The Rasch model asa practical and effective procedure for Educational measurement :Paper presented at their Annual Meeting of NCME. San Francisco.
- Hambletone, R. and Swaminathan, H. (1985). Item response theory: principles and applications. Boston , MA.
- Harris, D. (1989). Comarison of 1- , 2-, and 3- parameter IRT models. Educational Measurement: Issue and Practice . 8: 35-41.
- Harwel, M. (1997). Analyzing the result of monte carlo studies in item response theory. Education and psychological Measurement, vol(57), No(2), p p: 266-279.
- Ingebo, G.(1993). Examining item response theory: Consistency of rash calibration in basic skill item banks, ERIC, no: ED 408307.

- Koh C, Hsueh I, Wang W, Sheu C, Yu T, Wang C, Hsieh C. (2006). Validation of the action research arm test using item response theory in patients after stoke. Journal rehabilitation medicine (38), pp:375-380.
- Masters. G. and wright B.(1982). Rating scale analysis. Ahand book for research measurement Chicago: MESA press.
- Moral, F. and et al. (2006): Mapping and hazard assessment of atmospheric pollution in a medium sized urban area using the Rasch model and geo statistic techniques, Atmospheric Environmental, Vol. 40, pp. 1408 – 1418.
- Onn, D. (2013). Classical test theory versus item response theory: An evaluation comparability of item analysis result. Joint admissions and matriculation board, 1-23.
- Reckase, M.(1979). Unifactor latent trait models applied to multifactor tests: Results and implications. Journal of Educational Statistics, vol(4), no(3), pp: 207-230.
- Reeve,B.(2004).An Introduction To Modern Measurement Theory. Division Of Cancer Control And Population Science, National Cancer Institute.
- Sijetsma,K.Junker,B.(2006). Item Response Theory: Past Performance, Present Development, And Future Expectations. Behavior metrika, 33(1),75-102.
- Stage,C.(2003). Classical Test Theory Or Item Response Theory: The Swedish Experience. Spanish, Centro de Estudios Públicos, Santiago, Chile.
- Stenner, G. (1990). Objectivity specifc and general. Rasch Measurement Tran actions, vol(4), no (3),pp: 111-118.
- Vanderlinden, W. (2001). Applying the Rasch model. International Journal of Testing, Vol. 3, No. 4, pp. 319 – 326
- Wilson , T. and Mac, G. (2007). Counting on basics: Mathematical skill among tertiary entrants. International journal of mathematical education in science and technology, vol (38), no(1), pp:19-41.