بناء اختبار محكي المرجع لقياس الكفايات الرياضية في التفكير الهندسي لدى طلبة المرحلة المتوسطة وفق النموذج اللوجستي أحادي المعلم

*عبد الرحمن بن سالم الشهري**

ملخص

هدفت الدراسة إلى استخدام النموذج اللوجستي أحادي المعلم (نموذج راش) في بناء اختبار محكي المرجع لقياس الكفايات الرياضية في التفكير الهندسي، تم إعداد اختبار تحصيلي يقيس مستويات وقدرات التفكير الهندسي مكون من (38) فقرة من نوع الاختيار من متعدد، تكونت عينة الدراسة من (480) طالب وطالبة من طلبة الصف الثالث متوسط في محافظة الجوف، وأظهرت النتائج تحقق افتراضات النموذج اللوجستي أحادي المعلم ومطابقة (38) فقرة . وتم حذف (2) فقرتين غير مطابقات للنموذج كما أظهرت النتائج أن الفقرات كانت ضمن معايير المحكات المعتمدة في الأدب النظري للقياس، كما أظهرت النتائج لتقديرات أفراد عينة الدراسة استبعاد (3) أفراد غير مطابقين للنموذج وتراوحت معالم القدرة لأفراد عينة الدراسة بين (-1.389) الى (1.674).

الكلمات الدالة: اختبار محكي المرجع، نظرية الاستجابة للفقرة، نموذج راش أحادي المعلم، التفكير الهندسي.

المقدمة

تعد الاختبارات من أكثر أدوات القياس شيوعاً، وتعتمد كثير من القرارات على نتائج هذه الاختبارات، وهي مفيدة لجميع الأطراف التعليمية لما تقدمه من معلومات مفيدة لكل جهة، فتستخدم لتقويم تحصيل الطلبة واستعدادتهم وتشخيص مواطن الضعف والقوة عندهم أو انتقاء الطبة وتصنيفهم وارشادهم إلى أنواع التعلم المختلفة وضعهم في المسار التربوي المناسب وفي التوظيف والتنبؤ وتحديد مستويات التلاميذ و.....الخ.

ونظراً لأهمية القرارات المترتبة على نتائج الاختبارات فإنه لابد من بناء اختبارات تتصف بادقة والموضوعية في القياس بمعنى توفر القدرة والفاعلية للاختبارات في إعطاء بيانات ونتائج يمكن الوثوق بها تحت تغير أي ظروف، في التوصل إلى أحكام واستنتاجات وقرارات صحيحة لها ما يبررها، وعناصر فاعلية الاختبار تكون في جانبين مترابطين يتصل الأول بخصائص أو دلالات الاختبار بمجمله وتعبر عن درجة الثقة في العلامات أو البيانات المتحققة من تطبيقه وهي الخصائص السيكومترية الصدق والثبات (الكيلاني، وعدس والتقي، 2003).

ويشير عودة (2010) أن الاختبار التحصيلي يستخدم لتحديد مستوى مستوى تحصيل الطالب للمفاهيم والمهارات والمعارف المتضمنة في مادة دراسية تم تعلمها مسبقاً بصفة رسمية، من خلال إجاباته عن عينة من الأسئلة (الفقرات) التي تمثل محتوى تلك المادة (الدراسية) لذلك تعد عمليتي القياس والتقويم من المجالات المهمة في العلوم التربوية، حيث إنهما يهدفان إلى تطوير كل من المعلم وتحصيل المتعلم، وهما يعتمدان على تصميم وبناء أدوات تشخيصية أو تكوينية لقياس التحصيل عبد الهادي (2002)، ومن هذه الادوات بناء اختبار محكي المرجع للتحقق من كفاءة المتعلمين واكتسابهم للمعارف والمهارات وقياس مستويات تفكيرهم الرياضي والاتجاهات المرجوة (علام،2001).

وتعد الهندسة من أهم فروع الرياضيات لتي تعنى بالتفكير وتنميته لدى المتعلمين وحتى منذ القدم، كما ظهرت المبادئ وكأنها تتحدى العقل بقولها : "إذا لم تستطيع البرهان على أمر، فلا تقل أنك تعرفه " (عويضة، 2008)، وفي ذلك يشير هندام (1966) إلى أهمية الهندسة في تنمية أسلوب تفكيري سليم لدى الطلبة أثناء دراستها، ثم في استخدامه ففي مختلف شؤون الحياة.

فتنمية قدرة التلميذ على التفكير الهندسي من أهم أهداف تدريس الهندسة، لذا يجب أن يسعى المعلمون لاستخدام اختبارات قياسية موضوعية ومتسقة ودقيقة ومختلفة في قياس مستويات التفكير الهندسي لدى الطلبة لقياس نتاجات هذا االهدف (نوفل، 2010).

---

* وزارة التعليم، المملكة العربية السعودية. تاريخ استلام البحث ٢٠١٦/١/٢٣، وتاريخ قبوله ٢٠١٦/٤/٢.

© ٢٠١٨ عمادة البحث العلمي / الجامعة الأردنية. جميع الحقوق محفوظة.

مشكلة الدراسة

إن العمل في المجال التربوي يتطلب تحديد درجة معرفة وإتقان الطلاب للمهارات والمفاهيم الهندسية المختلفة المراد إكسابهم إياها والوصول بهم إلى مستويات تفكير رياضي وهندسي على درجة عالية من الفهم والاتقان، وكافة الخبرات التعليمية المقدمة إليهم، وما يمتلكون من أنماط التفكير، وما يستطيعون أداءه من أنشطة ومهام وتكليفات، وكذلك معرفة ميولهم واستعداداتهم واتجاهاتهم نحو القضايا المختلفة؛ وهذا كله يتطلب مقاييس في غاية الدقة والموضوعية والصدق والثبات والحساسية للفروق الطفيفة بينهم، وبما أن الفائدة التربوية المرجوة من تدريس الرياضيات والهندسة قد تتعذر ليس بسبب صعوبة المادة الدراسية وطبيعتها وحيويتها فقط، وأيضاً قياس نتاجاتها يتسم أيضاً بالصعوبة فبناء اختبار محكي المرجع لقياس الكفايات الرياضية في التفكير الهندسي لدى طلبة المرحلة المتوسطة يعد من الخطوات الهامة لتسهيل قياس النتاجات التعليمية من تدريس الهندسة، وبالتالي جاءت الدراسة الحالية لتجيب عن الأسئلة الآتية:

السؤال الأول: ما هي الخصائص السيكومترية للاختبار وفقراته التي يتم انتقاؤها لأغراض بناء اختبار محكي المرجع لقياس الكفايات الرياضية في التفكير الهندسية وفق نظرية الاستجابة للفقرة؟

السؤال الثاني: ماالقيم المتحققة لفقرات اختبار التفكير الهندسي (محكي المرجع) نحو قياس مستويات التفكر الهندسي وفق نظرية استجابة الفقرة؟

أهمية الدراسة

لم يعد هدف التعليم مقتصرا على التمييز بين الطلاب في القدرة أو المستوى، وإنما يسعى إلى إكسابهم مهارات معينة، وتحقيق أهداف محددة، بل وإتقان تلك المهارات، ولذا فإن على القائمين على التعليم بشكل عام، والمدرسة بشكل خاص تهيئة الظروف المناسبة التي تساعد على هذا النوع من التعليم، أو هذا النوع من الإتقان ((. Mastry Learning)، ولذلك برزت أهمية الدراسة في:

الأهمية النظرية: إضافة أدب تربوي للنظرية الحديثة للقياس وبالأخص نموذج راش أحادي المعلم، وكيفية إجراءات بناء الاختبارات وفق هذا النموذج، وبيان فاعلية وموضوعية ودقة نظرية الاستجابة للفقرة في بناء الاختبارات.

الأهمية التطبيقية : بيان كيف أن تقدير الخصائص السيكومترية وتقييمها للاختبار وفقراته يساعد في بناء اختبارات محكية المرجع من خلال انتقاء فقرات لها خصائص سيكومترية محددة بهدف تقدير علامات الطلاب على مجال محدد من خلال انتقاء فقرات ممثلة لمجال سلوكي محدد التفكير الهندسي، وبخصائص سيكومترية محددة، كما أن أهمية هذه الدراسة تنبع من أنها تسهم في تسليط الضوء على الاختبارات محكية المرجع، والتي لا تهتم بمقارنة أداء الطالب بأداء أقرانه، وإنما مقارنة أدائه بمستوى محدد من الإتقان، أو بتحقيق مجموعة من الأهداف الموضوعة مسبقا وقد تفيد الدراسة القائمين على بناء الاختبارات بوضع لهم منهجية علمية سليمة في بناء الاختبارات.

مصطلحات الدراسة

اختبار محكي المرجع (Criterion- Referenced Test): هو ذلك الاختبار الذي يستخدم في تقدير أداء الفرد في نطاق سلوكي محدد تحديداً دقيقاً، ويمكن أن يكون الاختبار المحكي المرجع اختباراً انتقائياً يتألف من عينة من فقرات مأخوذة من مجال سلوكي محدد ولكن يمكن أن تحدد له نقطة قطع يميز عندها بين متقنين وغير متقنين (Hambleton & Swaminathan, 1985, P:61)

علامة القطع: هي نقطة على متصل علامات الاختبار تستخدم لتصنيف المفحوصين إلى فئتين تعكس مستويات الأداء المختلفة بالنسبة لهدف معين أو الأهداف المراد قياسها في الاختبار.

الخصائص السيكومترية :يقصد بالخصائص السيكومترية معالم الصعوبة والتمييز والتخمين للفقرة والصدق والثبات للاختبار.

نموذج راش (Rash,model) أحادي المعلم، هو أحد نماذج النظرية الحديثة للقياس ويفترض وجود قدرة واحدة تفسر أداء الفرد على الاختبار، كما تفترض تساوي جميع فقرات التاختبار في التمييز بين مستويات القدرة التي يقيسها الاختبار، وأن معل التخمين يساوي صفراً (Wright&stone,1979).

حدود الدراسة

– اقتصرت الدراسة على طلبة والطالبات الصف الثالث متوسط في محافظة الجوف الممكلة العربية السعودية للعام الدراسي

. 2016/2015

– اقتصرت الدراسة على مووضوعات الهندسة ومستوياتها ((القدرة التحليلية والقدرة التصورية والقدرة الاستدلالية والقدرة الشكلية.)

– اقتصرت الدراسة على تطبيق افتراضات نظرية استجابة الفقرة. وفق نموذج راش أحادي المعلم

– اقتصرت الدراسة على ثلاث برمجيات إحصائية متخصصة في بناء اختبار محكي المرجع وفق النظرية الحديثة في القياس والتي أمكن توفرها وهي (SPSS & BIGSTEPS & Bilog-MG3) على الترتيب.

الاطار النظري

لاقت الاختبارات محكية المرجع (Criterion _ referenced test : CRT) خلال العقود الماضية عناية فائقة طريقة لوضع الدرجات، ووصف أداء الطلبة على اعتبار أن الاختبارات معيارية المرجع (Norm_ referenced test: NRT) تعرضت لتساؤلات حول قدرتها في المساعدة على تجويد التعليم نظراً للتفسير التقليدي حول أداء الطلبة، إذ أن معنى العلامة في إطار التفسير المعياري المرجع يتم اكتسابه من خلال مقارنة أداء الطالب بأداء مجموعته التي ينتمي إليها، وعلى العكس من ذلك لقيت الاختبارات محكية المرجع رواجاً في أوساط التربويين، لأنها تتسم بالموضوعية والمساءلة وهي تعتمد على تقييم جودة أداء الطالب نفسه مقابل معايير محددة. وكما ورد في هامبلتون Hambleton,, 1988 ي إلى كن إجمال الصفات الأساسية للاختبارات محكية المرجع في ما يأتي:

1. تركز على مجموعة النتاجات السلوكية أو الاجرائية الأكثر بشكل تفصيلي ومحدد، وتعطي درجة (علامة) لكل من تلك الأهداف على حده.
2. تصميمها يكون على درجة عالية من الموافقة ((Adequacy) وصدق المحتوى لقياسها للهدف السلوكي .
3. يفسر أداؤها استناداً إلى محك معين يحدد مسبقاً ويمثل الحد الأدني المقبول للأداء والذي كثيراً ما يشير إلى مستوى الاتقان أو أقرب مستوى له. وهي تقيس الأداء المطلق للفرد.

وهناك العديد من المعايير التي استخدمت لاختيار فقرات أدوات القياس، والغالبية العظمى من هذه المعايير انبثق عن مفاهيم النظرية التقليدية أو الكلاسيكية في القياس (Classic Test Theory, CTT). فيما ذكر انستازي (Anastasi, as cited in Gruijter & kamp, 2005) أن النظرية الحديثة في القياس، أو ما تعرف بنظرية استجابة الفقرة (Item Response Theory, IRT) تشكل إطاراً علمياً جديداً ووثيقاً في اختيار الفقرات في الوقت الحالي، وهي تعالج الكثير من القضايا التربوية والنفسية بشكل أكثر فاعلية من النظرية التقليدية.

وتفترض نظرية استجابة الفقرة (IRT) أنه يمكن التنبؤ بأداء المفحوصين، أو يمكن تفسير أدائهم في اختبار نفسي أو تربوي، في ضوء خاصية مميزة لهذا الأداء تسمى السمة (Trait)، ويصعب ملاحظة هذه السمة مباشرة؛ لذلك يجب تقديرها أو الاستدلال عليها من أداء المفحوص الذي يمكن ملاحظته على مجموعة من فقرات المقياس أو الاختبار (Hambleton, Swaminathan & Rogers, 1991).

وتم تطوير نماذج مختلفة للنظرية الحديثة في القياس تعرف باسم نماذج السمات الكامنة (Latent Trait Models, LTM)، تهدف في مجملها إلى تحديد العلاقات بين أداء الفرد في الاختبار والسمة التي تكمن وراء هذا الأداء وتفسّره. وتقوم بعض نماذج نظرية استجابة الفقرة على افتراضات أساسية (Gruijter & kamp, 2005)، هي: افتراض أحادية البعد (Unidimensionality)، وافتراض الاستقلال الموضعي (Local Independence)، وافتراض المطابقة لمنحنى خصائص الفقرة (Item Characteristic Curve, ICC)، وافتراض التحرر من السرعة (Speediness).

وقد حدد (Hambleton &swaminath an,1985 ,Hambleton & Jones, 1993) المشار إليه في الشريفين، 2003) مجموعة الافتراضات الأساسية التي تقوم عليها نظرية الاستجابة للفقرة:

– **أحادية البعد:** أي أن هناك سمة (قدرة) واحدة تفسر أداء الطالب على الاختبار، أما في النماذج فيفترض وجود أكثر من قدرة للفرد يقيسها الاختبار.

– **افتراض منحى خصائص الفقرة (Item** Characteristics Curve**)**

احتمالات استجابة المفحوص عن الفقرة في مستويات القدرة المختلفة إجابة صحيحة، وأن صفة التراكمية تشير إلى زيادة احتمال الاجابة الصحيحة على الفقرة بازدياد قدرة المفحوص وتبدو في نماذج الاختبارات المعدة لقياس سمة واحدة بدلالة معلم

واحد الصعوبة، أو معلمين الصعوبة والتمييز للفقرة، أو ثلاثة معالم الصعوبة والتمييز والتخمين.

وتتم الاحتمالية من خلال وصف العلاقة بين تحصيل المفحوص على الفقرة وقرته المقاسة بالاختبار من خلال الاقتران التراكمي الصاعد (دالة الاستجابة للفقرة) .

ويعدّ نموذج المعلمة الواحدة (One Parameter Model) أو ما يسمى بنموذج راش نسبة للعالم (Rasch) من أكثر نماذج نظرية استجابة الفقرة استخداماً؛ ومن المميزات التي أعطت هذا النموذج أهمية خاصة، هو أنه عندما تتطابق البيانات مع هذا النموذج، فإن معالم الفقرات التي تتمثل في درجة صعوبتها يمكن تقديرها بشكل مستقل عن العينة المستخدمة، كما أن قدرات الأفراد يمكن تقديرها بشكل مستقل عن درجة صعوبة الفقرات (Hambleton, Swaminathan & Rogers, 1991).

وسيتم استخدام النموذج الوجستي راش (Rasch Model) أحادي المعلم في تحليل استجابات عينة الدراسة، إذ يعد هذا النموذج من أبسط هذه النماذج شيوعاً ويعبر عنه بالدالة الرياضية التي تربط بين احتمال الاجابة الصحيحة للمفحوص على الفقرة، ومعلمة صعوبة الفقرة ومنحنيات خصائص الفقرة التي تختلف فقد في موقعها على متصل السمة، كما يفترض هذا النموذج تساوي الفقرة التمييزية لجميع الفقرات وتساوي (1) صحيح، وأن معلمة التخمين تساوي صفراً لجميع الفقرات، والدالة الرياضية التي تمثل الاحتمالية يعبر عنها بالمعادلة الرياضية التالية :

$$Pi_{(\square)}=(eD()/\sqsubset x$$

حيث:

(□) p_i: **احتمال الاجابة الصحيحة لفرد قدرته** (□ عن الفقرة

D عامل تدريج (calingfactor) ويساوي (1.7)، = قدرة فرد ما، bi = معلم صعوبة الفقرة i، ai معلم تمييز الفقرة (Hambleton & Swaminathan, 1985)، ويعتمد نموذج راش على الفرق بين استجابة الطالب (القدرة التي يمثلها المفحوص والسمة التي يراد قياسها) (التقي،2009).

a: ميل منحنى خصائص لفقرة عند نقطة انعطاف المنحنى (معلم تميز الفقرة) ويفترض هذا النموذج أن تكون قيمة معامل التمييز في جميع فقراته ثابتة وفي الغالب مساوية للواحد الصحيح.

bi: النقطة التي تقع على متصل القدرة عندما يكون ميل منحنى خصائص الفقرة أكبرما يمكن (معامل الصعوبة).

وتؤدي دالة معلومات الاختبار دوراً مهما في في نظرية الاستجابة للفقرة، إذ يمكن من خلالها تحديد الخطأ المعياري في التقدير الذي يرتبط عكسياً مع الجذر التربيعي لدالة المعلومات، كما تتيح استخدام معلومات الاختبار انتاج اختبار يتصف بالدقة اعتماداً على التقديرات الدقيقة لمعلم القدرة، وفي الاختبارات المحكية خاصة (Gruijter & Kamb,2005).

ويتم تقييم الفقرات وفق الفقرات وفق نظرية الاستجابة للفقرة حسب مطابقتها النموذج الذي يصف العلاقة بين الاستجابة والقدرة باستخدام اختبار احصائي أو تحليل البواقي (Residuals)، ويتم التعرف على الفقرات الضعيفة من خلال مؤشرات التميز والتي إما أن تكون سالبة أو موجبة ضعيفة، وتقديرات مؤشرات الصعوبة وإلا تكوت الفقرات سهلة جداً أو صعبة جداً لدى مجموعة من المفحوصين (2005ResidualsChernyshenko, Stark, Chan Drasgow & Williams, 2001).

وتم تقدير القدرة للمفحوصين من خلال نمط استجابة المفحوص على فقرات الاختبار باستخدام طريقة الارجحية العظمى بإيجاد (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 طريق رياضية وبتقريب هذه القيمة من خلال تحسين عملية التقدير بدءاً من قيمة تحدد مسبقاً (التقي، 2009).

الخصائص السيكومترية في نظرية الاستجابة للفقرة

معلم صعوبة الفقرة (bi): ويعرف بأنه تقدير القدرة المناظرة لاحتمال الاجابة الصحيحة (.50) عندما يكون نقطة تقاطع منحى خصائص الفقرة مع المحور الصادي يساوي صفراً تقريباً، ويكون معلم التخمين للفقرة (C) يساوي صفراً.

أما إذا كانت قيمة المقطع الصادي لمنحنى خصائص الفقرة أعلى من الصفر فإن صعوبة الفقرة هي القدرة الممثلة على محور السينات التي تقابل احتمالية الإجابة الصحيحة في منتصف المسافة بين تقاطع المنحنى مع المحور الصادي والقمية (1) والاحتمالية تساوي (1+Cl2) لأن القيمة (1) هي أكبر احتمال الاجابة الصحيحة التي تعبر عن مستوى صعوبة الفقرة.

وعند تحويل علامات القدرة للمفحوصين إلى مقياس معين وبمتوسط قدرة يساوي صفر (α=صفر) وانحراف معياري يساوي واحد فإن مدى قيم معلم الصعوبة للفقرات يتراوح من (2- إلى +2)، فقيم (b) القريبة من (2-) تكون الفقرة سهلة جداً وقيم الفقرات

القريبة من (+2) تكون صعبة جدا (Hambleton&swamnathan,1985).

معلم التمييز للفقرة (iα) : ويعبر عن قدرة الفقرة على التمييز بين مستويات القدرة المختلفة ويتناسب معلم التمييز طردياً مع ميل منحنى خصائص الفقرة عند نقطة الانعطاف، وتزداد قيمة معامل تمييز الفقرة بزيادة شدة ميل الانحدار، ويعرف معلم تمييز الفقرة (α) نظرياً على مقياس مداه ( )، مما يشير إلى وجود فقرات ذات تمييز سالب وهذه يتم حذفها من الاختبار، وفقرات ذات تمييز موجب، وأن مدى معالم تمميز الفقرات يتراوح بين (صفر إلى 2)، وإذا كان معامل التمييز للفقرة أكثر من صفر دل ذلك على وجود علاقة موجبة بين الأداء على الفقرة والقدرة المقاسة بالاختبار (Hambleton & swamnathan,1985)، فالفقرات عالية التمييز يمكن أن تكون فقرات الاختبار جيدة في مدى ضيق من القدرة، والقدرات منخفضة التمييز تكون ضعيفة في مدى واسع من القدرة (Warm,1978).

معلم تخمين الفقرة (Ci): هو الاستواء السفلي لمنحنى خصائص الفقرة، ويتمثل في نقطة تقاطع المنحنى مع المحور الصادي ويعبر عن احتمال الاجابة الصحيحة عن الفقرة من المفحوصين ذوي القدرات المنخفضة على الفقراتالصعبة جدا في التخمين أو هو احتمال الاستجابة الصحيحة عند مستوى قدرة متدنية جداً، وهذا المعلم خاص بالفقرات الموضوعية ثنائية التدريج من نوع الاختبار من متعدد فقط.

الدراسات السابقة

قد أجريت العديد من الدراسات في بناء اختبارات محكية المرجع لقياس الكفايات الرياضية للتأكد من مدى اتقان الطلبة المهارات الرياضية المختلفة لدى الطلبة التي تنمي من تفكيرهم الرياضي والتحقق من الخصائص السيكومترية للاختبارات وفق النظرية الحديثة للقياس (نظرية الاستجابة للفقرة) نموذج راش أحادي المعلم تحديداً، ففي دراسة أجرتها دينا (Dena, 2003) هدفت إلى استخدام اختبار متعدد المستويات وفق نماذج استجابة الفقرة في البحوث لتطبيقية، وتم تطبيق اختبار متعدد المستويات على أطفال من مراحل عمرية مختلفة (12–18) عاماً، وتم حساب الخصائص السيكومترية للاختبارات المتعددة حسب نماذج نظرية استجابة الفقرة والمعادلة الرأسية للتنبؤ بالسمة التي يمتلكها المفحوص من خلال الاختبار المشترك، وأظهرت النتائج أن القدرة تزداد بزيادة الفئة العمرية مع ازدياد صعوبة الفقرة وتميزها.

دراسة بني ياسين (2004) التي هدفت إلى تقدير الخصائص السيكومترية لاختبار محكي المرجع في الكيمياء للصف الأول الثانوي العلمي وفق النظريتين التقليدية والحديثة في القياس، ولتحقيق ذلك تم بناء اختبار تحصيلي في كيمياء الصف الأول الثانوي العلمي من مقرر الفصل الأول، وقد تكون الاختبار من 52 فقرة من نوع الاختيار من متعدد من أربعة بدائل . وتكون مجتمع الدراسة من (2462) طالباً وطالبةً من الصف الأول ثانوي العلمي للعام 2002 – 2003 في مدارس مديرية التربية والتعليم لمنطقة اربد الأولى . أما عينة الدراسة فتكونت من 481 طالباً وتم اختيارهم بالطريقة العشوائية العنقودية، وطبق الاختبار بإعطاء العلامة (1) للإجابة الصحيحة والعلامة (0) للإجابة الخطأ . وقد ت تم التحقق من مطابقة الاستجابات لفقرات الاختبار مع النموذج الثلاثي المعلم بعد تحقيقها لافتراضات النموذج وتكون الاختبار بصورته النهائية من 48 فقرة طابقت النموذج بعد حذف ٤ فقرات لم تطابق النموذج . تم تقدير الخصائص السيكومترية للاختبار (الصدق والثبات) وفق النظرية التقليدية حيث بلغ معامل الصدق بدلالة المحك 84. ومعامل الاتساق الداخلي 90. (كرونباخ ألفا)، كما تم تقدير معامل الثبات للاختبار 99. ومعامل ثبات الأفراد 88. . تم تقدير قيم معالم الفقرات لكل فقرة من فقرات الاختبار وفق النظريتين التقليدية والحديثة في القياس باستخدام النموذجين الأحادي والثلاثي المعلم، فقد بلغ متوسط صعوبة الفقرات وفق النظرية التقليدية 62. ومتوسط تمييز الفقرات 43. كما بلغ متوسط صعوبة الأفراد وفق النموذج الأحادي المعلم صفراً (لوجيت) ومتوسط الخطأ المعياري 119. (لوجيت) ومتوسط تمييز الفقرات 77. ومتوسط الخطأ المعياري للتمييز 125. ومتوسط تخمين الفقرات 24..

ودراسة وفي دراسة قامت بها ويبرج (Wiberg, 2004) هدفت الدراسة إلى تقييم الاختبار النظري لاختبار رخصة القيادة السويدي باستخدام كل من النظرية التقليدية في القياس ونظرية الاستجابة للفقرة مع تركيز أكثر على النظرية الأخيرة وكان الهدف الأهم مقارنة نماذج نظرية الاستجابة للفقرة الأحادي المعلم والثنائي المعلم والثلاثي المعلم من أجل معرفة النموذج الأكثر ملائمة لنمذجة الفقرات، ويتكون اختبار رخصة القيادة السويدي من اختبار نظري واختبار عملي، حيث يتكون الاختبار النظري من (65) فقرة من نوع الاختيار من متعدد وهو اختبار محكي المرجع حيث تتكون كل فقرة من (2–6) بدائل واحدة منها فقط صحيحة وتصحح بطريقة ثنائية حيث يحصل المفحوص على علامة واحدة عن كل فقرة أجيب عنها بشكل صحيح، وتكونت عينة الدراسة

من (5404) مفحوصاً وطبق الاختبار عليهم في شهر كانون الثاني للعام 2004 وقد تم استخدام النظرية التقليدية في لقياس ونظرية الاستجابة للفقرة للحصول على إحصائيات الفقرة والأفراد والاختبار وقد تم حساب معامل الثبات ألفا للاختبار وكل من قيم الصعوبة والتمييز للفقرة وفق النظرية التقليدية في القياس، وتم استخدام نماذج نظرية الاستجابة للفقرة الثلاثة المعروفة للاستجابة الثنائية (النموذج الأحادي المعلم والنموذج الثنائي المعلم والنموذج الثلاثي المعلم)، أشارت النتائج إلى وجود اختلاف بين معاملات الارتباط الثنائي النقطي وهذا يعني وجود اختلاف في مدى تمييز الفقرات . وأن (53) فقرة لم تطابق النموذج الأحادي المعلم و (25) فقرة لم تطابق النموذج الثنائي المعلم وهو المفضل على بقية النماذج الأخرى . كشفت النتائج أن تقديرات قدرات المفحوصين موزعة توزيعاً طبيعياً تقريباً وتم مقارنة النماذج الثلاثة من حيث دالة المعلومات والخطأ المعياري في التقدير وتبين أن النموذج الثلاثي المعلم يعطي معلومات أكثر حول مستوى القدرة التي تناظر علامة القطع (-1).

وفي دراسة أجراها كل من شن و نوزاوا و زهانج ((Chon, Hee & Nozawa, Yuki &Zhng, Zu. (2006) هدفت إلى فحص مط**ابقة نماذج الاستجابة للفقرة أحاد**ية المعلم لبيانات ثنائية الاستجابة وذلك بفحص التفاوت في كفاءة عدة مقاييس في تحقيق المطابقة للنموذج، وتتلخص هذه الطريقة في استخدام التوزيعات التنبؤية اللاحقة للبيانات المتكررة في اشتقاق قيم محاكاة (Simulated Values) ومقارنتها بالبيانات المشاهدة (Observed Data) فإن كان الاختلاف بينهما منتظماً دل ذلك على عدم مطابقة النموذج . استخدم الباحثون النموذج اللوجستي ثلاثي المعلم في اشتقاق بيانات المحاكاة الخاصة باستجابات (1000) مفحوص على (15) فقرة ثنائية الاستجابة . لقد تم استخدام البيانات في فحص كفاءة أربعة مقاييس في الكشف عن عدم المطابقة وهي: مقياس نسبة العلامة المعبر عنها بنسبة الفقرات الصحيحة (Proportion Correct Item Score) أي نسبة الفقرات التي كانت الإجابة عنها صحيحة، مقياس معامل الارتباط الثنائي النقطي (Point Biserial Correlation Coefficient) والذي يعبر عن معامل التمييز للفقرة في النظرة الكلاسيكية، ومقياس نسبة الأرجحة في زوج من الفقرات (Item Pair Odds Ratio) ويعبر عن عدد المفحوصين الذين يستجيبون لفقرتين مختلفتين استجابةً صحيحة للفقرتين أو خطأ لكليهما، ومقياس توزيع العلامات الكلية (Totl Score Distribution) . حلصت الدراسة إلى أن المقاييس المختلفة تكشف عن مظاهر مختلفة من عدم المطابقة، كما تعتمد كفاءة هذه المقاييس في الكشف عن عدم المطابقة على السياق الذي تستخدم فيه، فمثلاً كان مقياس الارتباط الثنائي النقطي ومقياس نسبة الأرجحية في زوج من الفقرات فاعلين في الكشف عن عدم مطابقة النموذج الأحادي المعلم ولكن ليس للنموذج الثنائي المعلم . كما أظهرت نتائج توزيع العلامات الكلية أن النموذج اللوجستي الثنائي المعلم أفضل قليلاً في تحقيق المطابقة من النموذج الأحادي المعلم . ولم تظهر نتائج توزيع العلامات الكلية أن النموذج الثلاثي عند النهايات العلوية والسفلية لمنحنى خصائص الفقرة جيدة نسبياً، ولكنه كان نموذجاً غير مناسب في منتصف مدى التوزيع الكلي للعلامات . أما مقياس نسبة العلامة المعبر عنها بنسبة الفقرات الصحيحة فلم يكن فاعلاً في الكشف عن المطابقة .

أما المهاجري (2007) فقد أجرت دراسة هدفت إلى بناء إختبار محكي المرجع لقياس الكفايات الرياضية في حل المعادلات والمتباينات من الدرجة الأولى (بمتغير واحد وبمتغيرين) لطالبات المرحلة المتوسطة بمكة المكرمة ولتحقيق أهداف الدراسة قامت الباحثة ببناء اختبار محكي المرجع مكون من (50) فقرة، وتأكدت من خصائصه السيكومترية بأكثر من طريقة، حيث حللت الباحثة الأسئلة بستة طرق مختلفة، كما أوجدت الصدق الوصفي للاختبار ومعاملات الاتساق الداخلي للاختبار والصدق المرتبط بمحك. والصدق الوصفي ومعاملات الاتساق الداخلي للاختبار والصدق المرتبط بمحك، وتم تقدير ثبات الاختبار بأربع طرق مختلفة وهي : طريقة إعادة الاختبار وطريقة معامل ألفا كرونباخ وطريقة ليفنجستون وطريقة هاريس، كما طبق الاختبار على عينة عشوائية من طالبات الصف الثالث متوسط بمدارس مكة المكرمة للتعرف على مدى اتقانهن. وتوصلت الدراسة إلى انخفاض مستوى الطالبات الصف الثالث متوسط في حل المعادلات والتباينات والمسائل الحسابية والأنظمة بشكل واضح وكبير، وأن مستوى الطالبات في حل الأنظمة كان أكثر المستويات انخفاضاً بينما أفضلها كان في حل المتباينات والبقية كانت نتائجهم متقاربة.

وأجرى الطيطي (2011) دراسة هدفت الى بناء اختبار محكي المرجع لقياس الكفايات الرياضية لدى طلبة الصف السادس الأساسي، تم تطبيق الاختبار بصورته النهائية والمكون من (36) فقرة على عينة مكونة من (600) طالبا وطالبة موزعين على 15) (مدرسة ضمن نطاق مدارس تربية اربد الاولى، وأظهرت النتائج ووجد ان غالبية الفقرات تتراوح صعوبتها مابين (0.30) الى 0.86) (، اما بالنسبة للقدرة التمييزية فقد استخدم الباحث مؤشر لقياس القدرة التمييزية هو :معامل ارتباط بوينت بايسى ريال بين الفقرة والاختبار، كما وجد الباحث ان غالبة المموهات تميز بالاتجاه المعاكس لتمييز الفقرة. وقد توافرت للاختبار ثلاثة أنواع من الصدق هي :صدق المحتوى تحديد المجال السلوكي وصياغة الاهداف العامة، والأهداف التفصيلية، وبناء الفقرات، وتقرير المحكمين)، والصدق التلازمي

من خلال قيمة معامل كابا (Kappa Coefficient)، حيث بلغت قيمته (0.82) ، وصدق البناء من حساب معاملات الارتباط بين كل فقرة والدرجة الكلية على المجال (المهارة) الذي تنتمي اليه وقد تراوحت قيمته مابين (.30−.81) حيث بلغت – Livingston) (indexوتم التحقق من الثبات باستخدام معامل ليفنجستون (Livingston – index) حيث بلغت (.97).

أما دراسة المصري (2015) فقد هدفت إلى استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في البحث التربوي وفق النموذج اللوجستي أحادي المعلم، تم إعداد اختبار لقياس تحصيل الطبة في مساق البحث التربوي مؤلف من (48) فقرة من نوع الاختيار من أربعة بدائل، وبلغت عينة الدراسة (220) طالباً وطالبة من طلبة كلية التربية المحويت باليمن، وأظهرت النتائج تحقق افتراضات النموذج أحادي المعلم ومطابقة (42) فقرة من فقرات الاختبار لتوقعات النموذج، وتم حذف (4) فقرات غير مطابقة للنموذج كما بينت النتائج أن تقديرات معلمي الصعوبة والتمييز ضمن المحكات المقبولة في أدبيات القياس التربوي، كما أظهرت نتائج تقديرات قدرات الأفراد استبعاد (7) أفراد غير مطابقين للنموذج وتراوحت معالم القدرة لأفراد عنية الدراسة بين (-2.659) لوجيت و (2.812) لوجيت.

التعقيب على الدراسات السابقة

تبين من خلال عرض الدراسات السابقة ذات الصلة الاهتمام الكبير بالنظرية الحديثة للقياس لما أظهرت من دقة وموضوعية في بناء المقاييس والاختبارات والتحقق من صدقها وثباتها وقياس السمات الكامنة لدى الأفراد في إعداد الاختبارات مثل، (Wiberg, 2004) وقياس مستوى قدرات الأفراد ومن الدراسات التي تناولت بناء اختبار محكي المرجع كدراسة بني ياسين (2004) باستخدام نظريتي القياس الكلاسيكية والحديثة وبعض الدراسات استخدمت النظرية الحديثة للقياس مثل دراسة (Dena, 2003)، والنماذج اللوجستية ثنائية المعلم وثلاثية المعلم، ومن النتائج التي خرجت منها هذه الدراسات الطرق المناسبة لحساب معامل الثبات في الاختبارات محكية المرجع استخدام معامل ليفنجستون كدراسة الطيطي (2011)، إذ أكدت نتائج الدراسات إلى فاعلية النظرية الحديثة للقياس أكثر من النظرية التقليدية. فيما تشابهت الدراسة الحالية مع دراسة المصري (2015) في استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع وفق النموذج اللوجستي أحادي المعلم، لكنها اختلفت في تناولها الكفايات الرياضية في التفكير الهندسي، ونظراً لندرة الدراسات التي تناولت بناء وإعداد اختبارات وفق نظرية الاستجابة للفقرة باستخدام نموذج راش أحادي المعلم. فقد جاءت الدراسة الحالية محاولتها لبناء اختبار محكي المرجع للكفايات الرياضية في التفكير الهندسي لدى المرحلة المتوسطة في القريات لما لهذا النوع من الاختبارات من الأهمية في التحديد الدقيق للكفايات التي يمتلكها الطلبة .

الطريقة والاجراءات

مجتمع الدراسة

تكون مجتمع الدراسة من طلاب وطالبات الصف الثالث متوسط في مدارس التعليم العام التابعة لمحافظة الجوف في المملكة العربية السعودية للعام الدراسي 2016/2015، إذ بلغ عددهم (3476) طالباً وطالبة حسب إحصائيات مكتب التعليم في منطقة الجوف.

عينة الدراسة

تكونت عينة الدراسة من (480) طالباً وطالبة موزعين على (14) شعبة من مدارس المرحلة المتوسطة الذكور والاناث، تم اختيارهم بالطريقة العشوائية العنقودية والجدول (1) يبين توزيع أفراد عينة الدراسة :

جدول (1)
توزيع أفراد عينة الدراسة

الجنس	العدد	النسبة المئوية%
ذكور	255	%53.12
إناث	261	%54.37
المجموع	480	%100

أداة الدراسةا

تم بناء اختبار محكي المرجع في مادة الرياضيات وحدات الهندسة للصف الثالث متوسط للفصل الدراسي الأول (2016/2015) وفق الخطوات الآتية:

1. تم تحديد الغرض من الاختبار، وهو قياس تحصيل طلبة الصف الثالث متوسط في مادة الرياضيات موضوعات الهندسة ومدى تحقق الأهداف التربوية من دراسة موضوعات الهندسة.

2-تم تحليل منهاج الرياضيات المقرر من وزارة التعليم السعودية للصف الثالث متوسط للفصل الدراسي الأول، والمكون من وحدات هي (التفكير الهندسي والاستدلال المكاني)، واشتقت منه الأهداف السلوكية كنتاجات تعلم يفترض أن تتحقق عند المتعلم في مستوى الإتقان التام. وقد شملت عناصر البناء المفاهيمي له وموضوعات الهندسة للصف الثالث متوسط من مفاهيم وحقائق ونظريات وقوانين وتطبيقات ومستوى التفكير الهندسي وهي (القدرة التحليلية والقدرة التصورية والقدرة الاستدلالية والقدرة الشكلية.) .

3-تم إعداد جدول مواصفات يربط مستويات الأهداف بفئات المحتوى الدراسي موضوع الاختبار.

4-عرض جدول المواصفات على (4) محكمين مختصين في الرياضيات؛ اثنان يعملان مشرفين تربويين ويحملان درجة الماجستير في أساليب تدريس الرياضيات، وثالث يحمل درجة الدكتوراه في القياس والتقويم، ورابع يعمل مدرسا لمادة الرياضيات ذو خبرة واسعة، وكان الهدف من الاحتكام إلى المحكمين إعطاء حكم على نتائج تحليل المحتوى والأهداف كما وردت في جدول المواصفات ومدى تغطيتها للمجال السلوكى المحدد في المنهاج والكتاب المقرر .وقد تم تعديل بعض الأهداف المصاغة وبناءً على رأى المحكمين واستقر عدد الأهداف السلوكية على (18) هدفا.

5-تمت كتابة فقرات عن كل هدف تم تحديده، واختلف عدد الفقرات حسب حجم المادة المقررة لكل هدف، وبلغ عدد هذه الفقرات (44) فقرة وقد روعي أن تطابق الفقرة الهدف السلوكي الذي تقيسه من حيث المحتوي والمستوى المعرفي، كما روعي ألا تعتمد إجابة إحدى الفقرات على إجابة فقرة أخرى.

6. تم التحقق من صدق الاختبار بعرض الاختبار وجدول المواصفات على مجموعة من المتخصصين للحكم على مدى تمثيل الفقرات للمحتوى المراد قياسه ،ودقة صياغة الفقرات،

وقد أخذت أحكام المحكمين حول الأمور الآتية:

- مدى التوافق بين الفقرة والهدف الذي يفترض أنها تقيسه.
- مدى وضوح الصياغة اللغوية للفقرة.
- مدى شمول الاختبار للمجال السلوكي.

وأسفر التحكيم عن استبعاد (4) فقرات اتفق ٨٠% من المحكمين على استبعادها وتعديل صياغة بعض الفقرات، وتعديل صياغة بعض البدائل لبعض الفقرات، و تكون الاختبار الاختبار من (40) فقرة شكلت الصورة الاولية وأجمع المحكمون على أنها كافية وممثلة للمجال السلوكي الذي يقيسه الاختبار, وهو مستويات التفكير الهندسي

التجريب الأولي للاختبار:

تم تطبيق الاختبار على عينة بلغت (51) طالباً وطالبة من مجتمع الدراسة للتعرف على صعوبة الفقرات، وتحديد الزمن اللازم للتطبيق، والاطمئنان على وضوح صياغة الفقرات والاجابة عن تساؤلات الطلبة، واجراء التحليل لبيانات العينة الاستطلاعية باستخدام الحزمة الاحصائية (Spss).

وقد تم حساب معاملات الصعوبة بإيجاد نسبة الطلبة الذين أجابوا صحيحة على الفقرة من الطلبة الذين حاولوا الاجابة عن الفقرة، وحساب القدرة التميزية للفقرات بحساب معامل الارتباط الثنائي بوينت بايسيريال بين الدرجة على الفقرة والدرجة على الاختبار، والجدول (2) يوضح ذلك .

جدول رقم (2)
معاملات الصعوبة مقدرة بنسبة النجاح على الفقرة، ومعاملات التميز مقدرة بمعامل الارتباط الثنائي لفقرات الاختبار في العينة الاستطلاعية

رقم الفقرة	معامل الصعوبة	القدرة التمييزية	رقم الفقرة	معامل الصعوبة	القدرة التمييزية	رقم الفقرة	معامل الصعوبة	القدرة التمييزية
1	.45	.442	15	.39	.527	28	.35	.421
2	.52	.435	16	.49	.435	29	.26	.387
3	.31	.386	17	.39	.381	30	.18	.21
4	.30	.347	18	.43	.448	31	.41	.375
5	.41	.438	19	.42	.503	32	.34	.343
6	.42	.516	20	.41	.486	33	.37	.382
7	.37	.532	21	.44	.42	34	.35	.443
8	.31	.515	22	.45	.43	35	.33	.465
9	.32	.582	23	.37	.47	36	.35	.454
10	.33	.546	24	.41	.532	37	.19	.318
11	.41	.371	25	.38	.343	38	.49	.435
12	.42	.358	26	.39	.501	39	.39	.381
13	.42	.34	27	.31	.515	40	.43	.448
14	.43	.35						

يلاحظ من الجدول (2) أن قيم معاملات الصعوبة لفقرات الاختبار تراوحت بين (.18) للفقرة (30) و(.52) للفقرة (2) . وبمتوسط الصعوبة لجميع الفقرات بلغ (.374)، اما قيم معاملات القدرة التمييزية فقد تراوحت بين (.21) للفقرة (30) و (.58) للفقره (9) وبمتوسط كلي لجميع معاملات التمييز بلغ (.431). وبناء على النتائج أعلاه فقد تم حذف الفقرات التي لها معامل صعوبة أقل من (.20) وهي الفقرة (30) والفقرة (37) . وهذا ما تفسره صعوبة الفقرة الذي يعرف بأنه تقدير القدرة المناظرة لاحتمال الاجابة الصحيحة حول (.50). وبذلك تتشكل الصورة النهائية للاختبار المطبق على عينة الدراسة من (38) فقرة (Hambletone & Swaminathan, 1985).

نتائج الدراسة ومناقشتها

الإجابة عن السؤال الأول والذي ينص على: "ماهي الخصائص السيكومترية للاختبار محكي المرجع في التفكير الهندسي وفق نموذج راش؟

أولاً: دلالات الصدق

الطريقة الاولى : صدق المحتوى: تم التحقق من صدق المحتوى بالاعتماد على تحليل المحتوى وبناء جدول المواصفات، وتأكد ذلك من خلال أحكام المحكمين من اعضاء الهيئة التدريسية كما سبق توضيحه في اجراءات إعداد الصورة الاولية للاختبار.

الطريقة الثانية: الصدق البنائي : تم تقديره بإجراء التحليل العاملي لاستجابات العينة الاستطلاعية على فقرات الاختبار بتحديد المكونات الرئيسة بايجاد قيم الجذور الكامنة للعوامل المستخلصة(Eigen Value) و نسبة التباين المفسر لكل عامل و النسبة التراكمية للتباين المفسر و الجدول (3) يوضح نتائج العينة لفقرات الاختبار.

وقبل التحقق من الصدق البنائي لا بدّ من التحقق من افتراض أحادية البعد (Unidimensionality). حيث تفترض نماذج نظرية استجابة الفقرة (IRT)، وجود قدرة واحدة تفسر أداء المفحوص في المقياس، ولذلك تسمى بالنماذج أحادية البعد. وللتحقق

من هذا الافتراض تم إجراء تحليل عاملي لبيانات المقياس المتعلقة باستجابات (480) طالبا وطالبة من طلبة الصف الثالث متوسط (38) فقرة، باستخدام البرمجية الإحصائية SPSS.

تم إجراء التحليل العاملي بطريقة تحليل المكونات الرئيسة (Principle Components Analysis, PCA) بناءً على درجات أفراد العينة على فقرات الاختبار، وتم تدوير المحاور باستخدام طريقة "Varimax"، لتقصي العوامل المسؤولة عن الأداء في الاختبار. وقد أظهرت نتائج التحليل العاملي تشبع فقرات الاختبار ب(11) عاملا، فسر العامل الأول منها 14.678% من التباين الكللي، وفسرت جميع العوامل 53.56% من التباين الكلي. ويبين الجدول (3) قيم الجذور الكامنة (Eigenvalues)، ونسبة التباين المفسر لكل عامل، وكذلك نسبة التباين المفسر التراكمية.

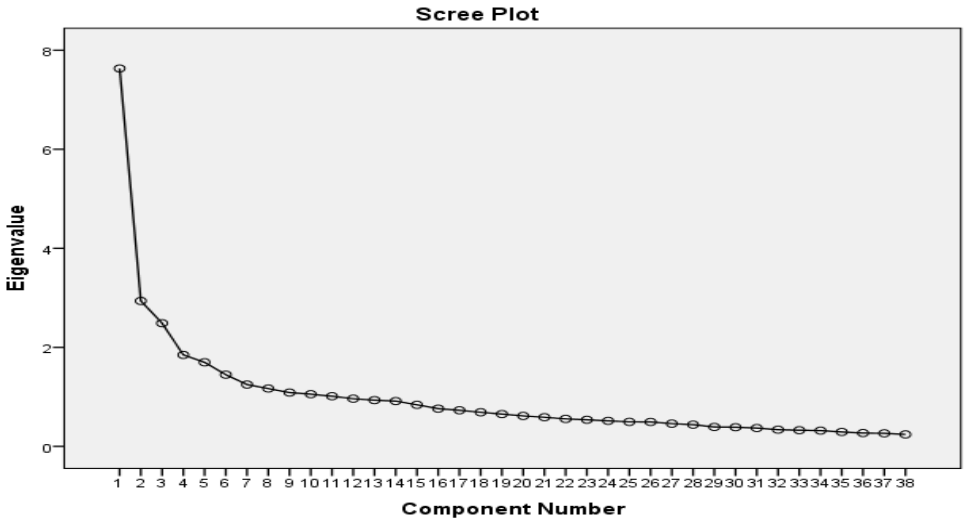
الجدول (3)

قيم الجذر الكامن ونسبة التباين المفسر من خلال العوامل التي تكوّن منها اختبار التفكير الهندسي محكي المرجع حسب نتائج التحليل العاملي تبعاً لاستجابات أفراد عينة الدراسة

العامل	قيمة الجدر الكامن	نسبة التباين المفسر	نسبة التباين%
1	5.561	15.789	15.817
2	1.921	3.991	19.808
3	1.632	3,924	23.732
4	1.541	3.715	27.447
5	1.420	3.628	31.075
6	1.353	3.577	34.652
7	1.318	3.399	38.051
8	1.214	3.160	41.211
9	1.167	3.015	44.226
10	1.158	2.893	47.117
11	1.135	2.811	49.930

يتبين من الجدول (3)، أن قيمة الجذر الكامن للعامل الأول بلغت ٥,٥٦٢ ويفسر ما نسبته ١٥,٨١٧% من التباين الكلي، وهي قيمة مرتفعة إذا ما قورنت مع قيم الجذور الكامنة لبقية العوامل، أما قيمة الجذر الكامن للعامل الثاني بلغت ١,٩٢١ ويفسر ما نسبته ٣,٩٩١% من التباين الكلي، بمعنى أن العامل الأول فسر ما يزيد عن ضعفي ما فسره العامل الثاني. أضف إلى ذلك، يلاحظ بأن نسبة التباين المفسر لكل من العوامل المتبقية متقاربة، بمعنى أنه يوجد شبه استقرار في نسب التباين المفسر لجميع العوامل باستثناء العامل الأول، وهذا مؤشر على تحقق افتراض أحادية البعد للمقياس (Hulin, Drasgow & Parson, 1983; Hattie, 1985)، أي أن المقياس يقيس سمة واحدة.

ويتعزز افتراض أحادية البعد، للمقياس المعدّ بالدراسة الحالية، ومن الجدير ذكره أنه تم استخدام الرسم البياني لقيم الجذر الكامن التي تم الحصول عليها من نتائج التحليل العاملي التي تظهر قيم الجذر الكامن لكل الأبعاد التي نتجت من خلال هذا التحليل. ويتضح من خلال الشكل رقم (١) أن العوامل الأول يتميز عن باقي العوامل من حيث قيم الجذور الكامنة المرتبطة به، كما يتضح أنه بعد العامل الأول أن الجذر الكامن للعامل الأول يطغى بشكل واضح على الجذور الكامنة لبقية العوامل، وأنه حصل نوع من الانكسار (Break Point) لقيم الجذور الكامنة للعوامل الباقية باستخدام طريقة البقايا المبعثرة Scree Plot method، وهذا مؤشر أيضاً على أحادية البعد لبيانات المقياس، ويستنتج من ذلك ان الاختبار يقيس سمة واحدة و بالاعتماد على ما اشار اليه ريكاسي (Reckase ,1979) بانه اذا فسر العامل الاول اكبر التباين، فإن هذا مؤشر على احادية البعد وبالاعتماد ايضا على مؤشر لورد (Lord , 1980) الذي ينص على ان الفقرات تحقق احادية البعد اذا كانت قيمة الجذر الكامن للعامل الاول كبيرة مقارنة بقيمة الجذر الكامل للعامل الثاني وهذا ما اكدته النتائج .



الشكل (1)

التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار على البيانات الكلية.

دلالات الثبات

تم تقدير دلالات الثبات كما يلي :

بإيجاد معامل الثبات بالطريقة النصفية وتصحيحه بمعادلة سبيرمان براوم فبلغ (0.90) . كما تم حساب معامل ثبات الاتساق الداخلي لفقرات الاختبار، بصورته النهائية المكون من (38) فقرة، باستخدام معادلة كرونباخ ألفا، وبلغ (0.92)، باستخدام برمجية (SPSS)، وتدل على تمتع المقياس بدرجة مقبولة من الثبات. وان القيمة العالية لمعامل الثبات المحسوب ب (KR-20) تؤكد ايضا تحقق احادية البعد حسب رأي كرونباخ والمشار اليه في (Hattie , 1985) و(حمادنة , 2009) اذ يرى ان معامل الثبات يعد مؤشرا جيدا لتحقق احادية البعد لانه يمثل متوسط كل من المعاملات النصفية الممكنة اضافة الى ان (KR – 20) تمثل القيمة لمتوقعة بالنسبة للتباين المفسر من العوامل المشتركة بين الفقرات عند ارتباط عينتين عشوائيتين من تجمع الفقرات الكلي .

وتستخدم دالة المعلومات (Information Function)، كمؤشر يمكن أن يستدل منها على معامل الثبات للاختبار وفق النظرية الحديثة في القياس، هو حيث أن منحنى دالة المعلومات للاختبار ينتج من تجميع منحنيات الفقرات فوق بعضها، وفق العلاقة الآتية:

$$I(\theta) = \sum_{g} Ig(\theta)$$

حيث $I(\theta)$ هي كمية المعلومات للاختبار عند مستوى القدرة θ، و $\sum Ig(\theta)$ هي مجموع دوال المعلومات لفقرات الاختبار عند مستوى القدرة θ، ولذلك فإن زيادة عدد الفقرات يعطي خطأً معياريّا صغيراً $S.E(\square)$ ونقصان قيمة الخطأ المعياري في تقدير القدرة θ يؤدي إلى زيادة كمية المعلومات التي يقدمها الاختبار وفق العلاقة الآتية:

$$I(\theta) = \frac{1}{\sqrt{S.E(\theta)}}$$

وعليه، فإن نقصان قيمة الخطأ المعياري في تقدير القدرة يؤدي إلى زيادة قيمة معامل الثبات. وما يميز نظرية استجابة الفقرة عن النظرية التقليدية في القياس، هو أن تقدير الثبات في النظرية التقليدية مرتبط بالعينة، بينما نظرية استجابة الفقرة تقدم لنا تقديراً للخطأ المعياري للقياس عند كل مستوى من مستويات القدرة، وبالإمكان من خلالها تحديد مدى مساهمة كل فقرة في تحديد دقة القياس. ولتطبيق ذلك، فقد تم استخدام البرمجية الإحصائية (Bilog- MG3) لتقدير كمية المعلومات للمقياس عند كل مستوى من مستويات القدرة المحدّدة، ورسم العلاقة بين قيم دالة المعلومات (Information Function)، والخطأ المعياري للتقدير (Standard Error) لفقرات الاختبار،

الإجابة عن السؤال الثاني والذي ينص على "ما القيم المتحققة لاحصائيات الفقرة للاختبار للعينة النهائية وفق نظرية الاستجابة للفقرة؟

تم تقدير قيم معالم فقرات الاختبار من تحليل استجابات افراد عينة الدراسة (480) عن فقرات الاختبار (38) فقرة . وهو عدد مناسب وفق ما اشار اليه (fan 1998), أنه يجب ان الايقل عن عدد أفراد العينة (200) وفق نموذج راش، باستخدام (ACER 301 ,) ويحذف البرنامج استجابات الافراد اللذين يحصلون على العلامة صفر على الاختيار او الحاصلين على العلامة التامة (38) على الاختبار , وكذلك الفقرات الذي يفشل جميع افراد العينة في الاجابة عليها جميعا, وكذلك الفقرات التي ينجح جميع افراد العينة في الاجابة عليها اجابة صحية، ولم يحصل اي فرد على العلامة الكاملة (38) ولا أي فرد حصل على العلامة صفر، وكذلك لم تكون هنالك فقرة اجاب عليها جميع افراد العينة اجابة صحيحة او اجابة خاطئة، ونتيجة لذلك لم تستبعد اجابات اي فرد من العينة وكذلك لم تستبعد اي فقرة من فقرات الاختبار. ويبين الجدول (4) معاملات صعوبة الفقرات الاخطاء المعيارية في تقدير معلم الصعوبة لكل فقرة من فقرات الاختبار وفق نموذج راش .

جدول (4)

معامل الصعوبة والاخطاء المعيارية لها مقدرة باللوجيت ومرتبة ترتيبا تنازليا حسب صعوبة الفقرة للاختبار في صورته النهائية

رقم الفقرة	معامل الصعوبة	الخطأ المعياري	رقم الفقرة	معامل الصعوبة	الخطأ المعياري	رقم الفقرة	معامل الصعوبة	الخطأ المعياري
8	1.674	0.109	6	.209	.09	35	1.174-	.092
23	1.281	0.107	3	.915	.09	7	1.225-	.092
10	1.128	.099	24	.082	.09	1	1.256-	.092
37	.1.007	.079	27	.051	.09	13	1.418-	.092
20	1.004	.096	34	.038	.09	16	1.462-	.092
26	1.002	.095	11	.024	.09	2	1.491-	.099
29	1.00	.094	14	.016	.09	4	1.862-	.099
28	.997	.794	36	.009	.09	25	1.590-	.099
18	.829	.791	32	.208-	.09	19	1.487-	.099
22	.784	.09	29	.334-	.091	12	1.405-	.112
21	.776	.09	31	.343-	.091	15	1.389-	.114
5	.534	.09						
المتوسط الحسابي			صفر			0.10		
الانحراف المعياري			1.164			0.032		

يشير الجدول (4) ان مدى قيم صعوبة الفقرات امتد من (-1.389) الى (1.674) لوجيت وبخطأ معياري يتراوح بين (0.109 – 114.) وبمتوسط حسابي يساوي صفر بخطا معياري (10.)

وبتطبيق معيار خصائص الفقرة وفق نظرية الاستجابة للفقرة فإن الفقرات التي يزيد معامل صعوبتها (+1.50) لوجيت تعد ضمن الفقرات الصعبة وان الفقرات التي يقل معاملها عن (1.50−) لوجيت تعد من الفقرات السهلة , وان الفقرات التي تتراوح قيم معامل صعوبتها بين (1.50+ , 1.50−) هي ضمن المدى المقبول لمعاملات الصعوبة، وعليه فان الفقرة (8) تعد صعبه لان معامل صعوبتها زاد عن (1.50+) وان الفقرتان (25، 4) فقرات سهلة لان معامل صعوبه كل منها زاد عن (1.50−)، واما بقية الفقرات فهي ضمن المدى المتوسط والمقبول بمعاملات صعوبة الفقرات حسب المعيار (1.50+ , 1.50−)

وللتحقق من مؤشرات حسن المطابقة لكل فقرة من فقرات الاختبار ومقارنتها بالمدى اللذي يمكن ان يقع ضمنه احصائي المطابقة بتحليل البواقي (FitResid) هو (2.50− الى 2.5 +) والجدول (5) يوضح ذلك .

جدول (5)

قيمة احصائي المطابقة لقيم صعوبة فقرات الاختبار

رقم الفقرة	المطابقة لصعوبة الفقرات	رقم الفقرة	المطابقة لصعوبة الفقرات	رقم الفقرة	المطابقة لصعوبة الفقرات
1	.254-	15	.752	28	.950-
2	.325-	16	1.217-	29	.95-
3	1.65−	17	.670-	30	.156
4	1.18-	18	3.72	31	.806-
5	2.306	19	2.436	32	1.157-
6	1.139	20	1.431	33	.365
7	2.054	21	2.283	34	.569
8	1.108	22	.376	35	1.748−
9	.256-	23	.734	36	2.891
10	.325	24	1.417	37	.987
11	1.337	25	2.417-	38	3.343
14	.71-				

يلاحظ من الجدول (5) ان الفقرات (38، 36،18) تقع قيم مؤشرات المطابقة لها خارج المدى الذي يمكن ان يقع ضمنه احصائي المطابقة، لذا تعد هذه الفقرات غير مطابقة لنموذج راش احادي المعلم واما بقية الفقرات فهي مطابقة للنموذج .

كما يعطي النموذج ايضا تقديرات لقدرات الافراد، اللذين طبق عليهم فقرات الاختبار المبنية وفق هذا النموذج، يوضح الجدول (6) تقديرات الافراد والمناظرة للعلامة الخام في قياسها للاختيار في صورته النهائية مرتبة تنازليا .

جدول (6)

قدرات الأفراد في قياسها للاختبار بالصورة النهائية مرتبة تنازلياً

العلامة الخام	تكرار الأفراد	قيم القدرة المناظرة للعلامة الخام	العلامة الخام	تكرار الأفراد	قيم القدرة المناظرة للعلامة الخام	العلامة الخام	تكرار الأفراد	قيم القدرة المناظرة للعلامة الخام
38	0	3.923	24	15	.482	11	2	.026-
37	0	3.224	23	9	.417	10	1	.0148-
36	1	2.812	22	7	.364	9	5	.186-
35	3	2.240	21	3	.345	8	3	.262-
34	1	2.03	20	8	.163	7	2	.482-
33	1	1.793	19	11	.091	6	1	.674-
32	2	1.579	18	13	1.528	5	3	.785-
31	1	1.385	17	9	.081	4	1	1.154
30	1	1.07	16	8	.144	3	2	1.836
29	1	.965	15	7	.133	2	4	1.969-
28	1	.86	14	4	.123	1	2	2.034-
25	3	.567						

يتضح من الجدول (6) ان معالم القدرة الفعلية لافراد عينة الدراسة، قد تراوحت بين (2.659−) لوجيت التي يقابل العلاقة (ح) و (2.812) لوجيت التي تقابل العلامة الخام (36)، وكانت ادنى قدرة (-2.034) . وتقابل العلامة الخام (2) ، وأظهرت النتائج وجود (3) أفراد من العينة غير مطابقين لنموذج راش احادي المعلم لان قيم التقديرات لهم خارج مدى المطابقة (2.50− الى 2.50+) لوجيت .

التوصيات :

وفي ضوء لنتائج الدراسة يوصي الباحث بالاتي :

١. استخدام نظرية الاستجابة للفقرة لبناء اختبارات محكية المرجع في مختلف العلوم.

٢. بناء اختبارات معيارية المرجع وفق النظرية الحديثة للقياس.

٣. استخدام المقياس بصورته النهائية (38) فقرة لقياس مستويات التفكير الهندسي لدى طلبة الثالث متوسط، نظرا لتمتع المقياس المبني والمطابق لتوقعات نظرية استجابة الفقرة (IRT) بدلالات صدق وثبات مقبولة.

٤. التحقق من للخصائص السيكومترية للاختبار وفق النموذج اللوجستي ثنائي المعلم وثلاثي المعالم.

الاستناد على المنهجية المستخدمة في اعداد الاختبار في اعداد اختبارات لمساقات اخرى والتأكد من خصائصها السيكومترية.

المراجع

بني ياسين، ع (2004). الخصائص السيكومترية لاختبار محكي المرجع في الكيمياء لطلاب الصف الأول الثانوي العلمي مقدرة وفق النظريتين الكلاسيكية والحديثة في القياس . رسالة دكتوراة غير منشورة ، جامعة عمان العربية للدراسات العليا ، عمان، الأردن .

التقي ، أ (2009). النظرية الحديثة في القياس ، دار المسيرة للنشر والتوزيع والطباعة ، عمان،الأردن.

حمادنة، إ (2009). استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلم، مجلة العلوم التربوية والنفسية، المجلد (10)، العدد (2)216 − 230.

الشريفين، ن (2003). مدى تحقق معايير الفاعلية في معادلة اختبارين أحدها ثنائي التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس. رسالة دكتوراة غير منشورة ، جامعة عمان العربية للدراسات العليا ، عمان ، الأردن .

الطيطي، م (2011). بناء اختبار محكي المرجع لقياس الكفايات الرياضية في مادة الرياضيات للصف السادس الأساسي ، رسالة ماجستير غير منشورة، جامعة جدارا، الأردن.

عبد الهاي، ع (2002). القياس والاختبارات النفسية، أسس وأدوات المراجع في التربية وعلم النفسن سيد الناشر: دار الفكر.

علام، ص (2001). القياس والتقويم التربوي والنفسي، القاهرة: دار الفكر العربي.

عودة، أ (2010). القياس والتقويم في العملية التدريسية، اربد، الأردن: دار الأمل للنشر والتوزيع.

عويضة، سوار عبداللطيف (2008)، موسوعة علم الرياضيات، عمان : دار دجلة،

الكيلاني ، ز وعدس، ع والتقي، أ (2003). القياس والتقويم في التعلم والتعليم . ط2. جامعة القدس المفتوحة.

المصري، م (2015) الخصائص السيكومترية لاختبار محكي المرجع في البحث التربوي وفق نظرية الاستجابة للفقرة، مجلة كلية التربية الأساسية، المجلد 21، العدد 89، 701−730.

المهاجري، م (2007). بناء اختبار محكي المرجع لقياس الكفايات الرياضية في حل المعادلات والمتباينات من الدرجة الأولى (بمتغير واحد ومتغيرين) لطالبات المرحلة المتوسطة بمدارس مكة المكرمة الحكومية، رسالة ماجستر غير منشورة، المملكة العربية السعودية.

نوفل، محمد بكر (2010)، تطبيقات عملية في تنمية التفكير باستخدام عادات العقل، عمان: دار المسيرة.

هندام، يحيى حامد (1966)، تدريس الهندسة النظرية ومقومات البرهان المنطقي، القاهرة: دار النهضة العربية.

Chon, Hee & Nozawa, Yuki &Zhng, Zu. (2006). Posterior Predictive Checking of Uni-dimensional Item Theory Models. http://www.stat.ucl.ac.be/

Dena. A. paster (2003). The use of Multilevel Item response Theory Modeling in applied Research: Applied Measurement in Education, 16 (3) 223- 243.

Gruijter, D. and Kamp, L., (2005). Statistical Test Theory for Education and Psychology, Retrieved December 30, From: www.leidenuniv.nl/-gruijterdnmde, 2005.

Hambleton, R. (1988). Criterion-Referenced Measurement. In Keeves (Ed), Educational Research, Methodology and Measurement An International Hand book, Pergmon Press. New York (277-281).

Hambleton, R., Swaminathan, H. and Rogers, J., Fundamental of Item Response Theory, Second Edition, Newbury Perk California, 1991.

Hambletone, R. K., &Swaminathan, H. (1985). Item response theory: principles and applications. Boston , MA:

Warm,T.A.(1978) . Aprimer of Item Response theory Oktahoma. Oklahoma: U.S. cost Guard Institute 73/ 69.

Wiberg, Marie. (2004) . Classical Test theory vs. Item Response Theory an Evaluation of the theory test in the Swedish driving – licencetest . EM No. 50. Umea University .

Construct A Criterion-Referenced Test to Measure the Mathematical Competencies in Engineering Thinking for Middle School Student (According to Rash Model)

Abd Alrahman Bin Salem Al-Shahri*

ABSTRACT

The aim of this study was to construct a criterion-referenced test in geometric Thinking, using item-response theory (IRT), applying One-Parameter Logistic Model (Rasch Model). The test consisted of (38) Item with multiple choice. The subjects of study were (480) students in the third grade average in al-Jawf province, The results showed that all test items proved the assumption of the model except (4) items, that were deleted. The estimates of difficulty and discrimination parameters were acceptable. Results also showed that (3) subjects were excluded because they were misfit the model. Parameters estimates of subjects ranged between 1.674) and (1.389-) logit.

Keywords: criterion-referenced test, item-response theory, One-Parameter Logistic Model (Rasch Model), geometric Thinking.

* Ministry of Education, Kingdom of Saudi Arabia. Received on 23/1/2016 and Accepted for Publication on 2/4/2016.