International Journal for Research in Education

Volume 46
Issue 1 *Vol.(46), issue(1), January 2022*

Article 11

2022

Developing a Diagnostic Criterion-Referenced Mathematics Test in Solving Equations Using the Attribute Hierarchy Method

Dr. Amal Ahmad Alzoubi
yarmouk university, amal@yu.edu.jo

Follow this and additional works at: https://scholarworks.uaeu.ac.ae/ijre

Part of the Educational Assessment, Evaluation, and Research Commons

Recommended Citation

Alzoubi, Dr. Amal Ahmad (2022) "Developing a Diagnostic Criterion-Referenced Mathematics Test in Solving Equations Using the Attribute Hierarchy Method," *International Journal for Research in Education*: Vol. 46 : Iss. 1 , Article 11.
Available at: https://scholarworks.uaeu.ac.ae/ijre/vol46/iss1/11

This Article is brought to you for free and open access by Scholarworks@UAEU. It has been accepted for inclusion in International Journal for Research in Education by an authorized editor of Scholarworks@UAEU. For more information, please contact j.education@uaeu.ac.ae.

UAEU كلية التربية College of Education
جامعة الإمارات العربية المتحدة United Arab Emirates University
IJRE

المجلة الدولية للأبحاث التربوية
International Journal for Research in Education

المجلد (46) العدد (1) يناير 2022 - Vol. (46), issue (1) January 2022

Manuscript No. : 1719

Developing a Diagnostic Criterion-Referenced Mathematics Test in Solving Equations Using the Attribute Hierarchy Method

بناء اختبار تشخيصي محكي في الرياضيات في حل أنظمة المعادلات وفقا للاستراتيجية الهرمية للسمات

Received تاريخ الاستلام	Sep 2020	Accepted تاريخ القبول	Dec 2020	Published تاريخ النشر	Jan 2022

DOI : http://doi.org/10.36771/ijre.46.1.22-pp334-364

Dr. Amal A. Alzoubi
Yarmouk University
amal@yu.edu.jo

د. / آمال أحمد الزعبي
كلية التربية - جامعة اليرموك
amal@yu.edu.jo

حقوق النشر محفوظة للمجلة الدولية للأبحاث التربوية
ISSN : 2519-6146 (Print) - ISSN : 2519-6154 (Online)

Developing a Diagnostic Criterion-Referenced Mathematics Test in Solving Equations Using the Attribute Hierarchy Method

Abstract

The study aimed at developing a criterion test for solving equations systems according to the strategy of hierarchical constructive analysis for 10th-grade students. A diagnostic test was developed, consisting of 46 items distributed to six sub-skills. The test was administrated to 390 10th grade students in Irbid, were male 184, Jordan. The validity, reliability of the test, the validity, and efficiency of hierarchical construction of solving equations were verified. The study results revealed that the test might be good for diagnosing purposes in solving quadratic equations. The researcher suggested recommendations to universities, schools, and measurement centers to use this strategy in different subjects and apply the test to students.

Keywords: Diagnostic Criterion-Referenced Test, hierarchical model, cognitive diagnostic model, attribute hierarchy method, solving systems of equations.

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

بناء اختبار تشخيصي محكي في الرياضيات في حل أنظمة المعادلات وفقا للاستراتيجية الهرمية للسمات

مستخلص البحث

هدفت هذه الدراسة إلى تقديم استراتيجية لبناء اختبار تشخيصي محكي المرجع في الرياضيات لطلبة الصف العاشر الأساسي في موضوع حل أنظمة المعادلات باتباع الطريقة الهرمية. ولتحقيق هذا الهدف تم بناء اختبار تشخيصي لقياس مهارة "حل نظام مكون من معادلتين تربيعيتين بمتغيرين" قسّم إلى ست مهارات فرعية مكون في صورته النهائية من 46 فقرة. طبق الاختبار على 390 طالباً وطالبة منهم 184 ذكور في الصف العاشر الأساسي في إربد - الأردن ، وتم استخراج دلالات صدق وثبات الاختبار، وتم التحقق من صدق البناء الهرمي للاختبار وفاعليته. أظهرت نتائج الدراسة أن الاختبار الذي بني في هذه الدراسة يمكن أن يكون صالحا لأغراض التشخيص في حل المعادلات التربيعية. وأوصت الدراسة باستخدام هذه الاستراتيجية في المدارس والجامعات ومراكز القياس في موضوعات أخرى وذلك لقدرتها على التشخيص، كذلك تطبيق الاختبار الذي تم بناؤه في هذه الدراسة على الفئات المستهدفة من الطلبة.

الكلمات المفتاحية : اختبار تشخيصي محكي ، النموذج الهرمي للسمات ، نموذج معرفي تشخيصي ، الطريقة الهرمية للسمات، حل أنظمة المعادلات.

مقدمة البحث

يعد التقويم التربوي أحد المكونات الأساسية والمرتكزات المهمة التي تقوم عليها منظومة التربية، وتعد الاختبارات من أهم أدوات القياس النفسي والتربوي،وأكثرها شيوعا وذلك لما لها من أهمية في عملية التقويم، فهي تعطينا فكرة واضحة عن قدرات الطلبة التحصيلية، ومستوى نشاطهم ومن خلالها يمكن وضع الخطط العلاجية للطلبة الضعاف، وتصنيفهم وتحديد مستوى نجاحهم ورسوبهم، ومن خلال نتائجها يمكن أن تعدّل الأساليب وتقنيات التدريس(عبد الهادي،2002).

يعتمد التقييم التقليدي في بنائه بشكل نموذجي على التصنيفات المنطقية، ولائحة المواصفات، لكنه يفتقد النموذج المعرفي الصريح للبناء والعمليات المعرفية لتحصيل الطلبة(Snow &Madinah, 1991)؛ نتيجة لذلك فإن الدرجات الاختبارية الناتجة من التقييم التقليدي مربوطة بالمحتوى أكثر من العمليات المعرفية المقيسة بفقرات الاختبار؛ فهي تفشل في تزويد المعلمين بمعلومات متعلقة بنقاط القوة والضعف لمعرفة طلبتهم، التي تمكّنهم من اتخاذ قرارات تساعد الطلبة في النجاح في المواقف التربوية(Nichols, 1994) .

ويستخدم التقييم التقليدي لغرض يتمثل في قياس نجاح الطلبة في الاختبارات التي تأتي في نهاية وحدة دراسية أو فصل دراسي، يتلقى الطلبة من خلاله تغذية راجعة تخص الدرجة النهائية، التي تتسم بمؤشر عام عن مستوى إتقان المفحوص، وهذا النوع من التغذية الراجعة لا يعطي استدلالات محددة عن نقاط القوة والضعف للمفحوص، وبهذا فإنه يمكن النظر إلى التقييم التقليدي الذي غالبا يستخدم لأغراض التقويم الختامي أنه غير فعّال، وليس له أثر مباشر في التعلم (Anderson, 1998; Sadler, 1989; Struyven, et al., 2003, 2008)، ويزود بتغذية راجعة محدودة عن تحسن أداء المفحوص(Bailey, 1998) .

ونتيجة لهذه العيوب في التقييم التقليدي، قوبل هذا المنحى بالتحفظ من بعض مستخدمي الاختبارات، إضافة لذلك بدأ الطلب والبحث عن تقييم يمكن أن يزود بمعلومات مفيدة للمعلم والطالب، ويخبرهما بدرجة الإتقان المحصلة، واكتشاف مواطن الصعوبة التي يصادفها الطالب خلال تعلمه، واكتشاف الصعوبات التي تواجه المعلم من أجل البحث عن طرائق وإستراتيجيات ناجعة. ومن طرق التقويم المقترحة التي من الممكن أن تساعد في مثل هذه الحالات، التقويم التشخيصي الذي يجمع البيانات الخاصة بالمستوى التعليمي لدى الطلبة، التي تمكّن المعلم في تعرف مواطن القوة والضعف لديهم، ومن خلال معرفة هذه النقاط المهمة يتمكّن من تصميم نشاطات تناسب حاجة الطلبة في غرفة الصف. وتتطلّب استراتيجيات العلاج بالتقويم التشخيصي معرفة أمرين رئيسين، هما: بناء الاستراتيجية، ووضع الأهداف بعد قياس مستوى الأداء السابق للطلبة، وبعد إدراك هذين الأمرين يجب البحث عن الوسيلة العلاجية المناسبة لتحقيق الأهداف المرجوة (علام، 1995).

التقويم التشخيصي المعرفي :يعد التقويم التشخيصي المعرفي منحى جديدا نسبيا يهدف إلى التزويد بالتغذية الراجعة من خلال تقديم ملف مدون مفصل عن المهارات التي أتقنها المتعلم (DiBelloet al., 2007; Embretson, 1991, 1998; Hartz, 2002; Nichols et al., 1995; Tatsuoka, 1983).

وعلى الرغم من ضرورة البحث في عمليات تدوين درجات الاختبار التشخيصي، إلا أن هناك دراسات قليلة بحثت في استخدام تقارير التشخيص في سياق التعليم والتعلم. واستخدام التغذية الراجعة التشخيصية يحتاج لأن يفهم عن طريق الاهتمام بالاعتقادات المختلفة عن التعلم وطرق التدريس التي يقوم بها المعلمون، وآثارها المختلفة بالاعتماد على مستويات الإتقان للمتعلمين، وأنماط تعلمهم، أو سياق التعلم (Kunnan& Jang, 2011). وتحتاج التغذية الراجعة التشخيصية لأن تكون وصفية وتفسيرية بحيث تساعد المتعلمين في سد الفجوة بين مستوى الإتقان لديهم والمستوى المأمول منهم (Black &William, 1998).

وبالرغم من فائدة التقويم التشخيصي، إلا أنه لم يجد الاهتمام الكبير مقارنة باختبارات التحصيل؛ إذ تعد أدوات التقويم التشخيصي المتوافرة للمعلمين لاستخدامها في الغرفة الصفية قليلة جدا(Alderson, 2005). وهناك حاجة إلى المزيد من البحوث لتطوير الاختبارات التشخيصية التي تتضمن مهمات معرفية لتشخيص نقاط القوة والضعف للمتعلمين. وهذه الاختبارات تحتاج لأن تستند إلى تصميم منتظم مكون من عدة خطوات (Davidson & Lynch, 2002; Mislevy et al., 2003; Pellegrinoet al., 2001)

ويحتاج معد ومطبق الاختبارات التشخيصية المحكية خطوات دقيقة تساعده في تشخيص نقاط الضعف والقوة،وتعرف مصادر الأخطاء الناجمة عن سوء الفهم، أو عن عدم التمكن من الإجراءات أو العمليات التي تنطوي عليها الكفايات أو المهارات، وهذه الخطوات كالآتي:1- التعريف بالأهداف التعليمية التي تخدم كمحكات Criteria لمحتوى التشخيص 2- تصميم مهمات معينة تخبرنا عن مستويات الإتقان للمتعلم في ضوء الأهداف التعليمية3- تطوير نظام تصحيح ورصد درجات يسمح بالحصول على معلومات تشخيصية دقيقة 4- إعداد تقرير التشخيص للحصول على أكبر فائدة مرجوة (Nichols, 1994).

القياس التربوي وتقويم المعرفة التشخيصي: يمكن أن يوصف تقويم المعرفة التشخيصي بأنه اختبار تربوي لقياس العمليات المعرفية للمتعلم، والتعلم، والمعرفة، وتطور المهارات لأغراض تشخيصية (Ketterlin-Geller &Yovanoff, 2009). ويستخدم تقويم المعرفة التشخيصي للتحقق فيما إذا كان المفحوص يمتلك المعرفة والمهارة المطلوبة لحل مسائل في مجال ما. وبإعداد ملف إنجاز Protfile لنقاط القوة والضعف لمعرفة الطالب، فإن المعلم يمتلك وسائل المعالجة وتحسين خطط البرامج بحيث تحقق حاجات كل طالب بشكل فردي، حيث يعمل تقويم المعرفة التشخيصي على ضم نظريات المعرفة واتحادها مع نماذج الإحصاء كمدخل قيّم لعمل استنتاجات

عن نقاط القوة والضعف للمهارة المعرفية أو السمة للطلبة (Nichols et al. 1995; Nichols, 1994). ويقصد تاتسكووتاتسكو (Tatsuoka&Tatsuoka, 1997) بالسمة بأنها العمليات المعرفية والمعرفة المطلوبة لحل المشكلات في مجال الهدف.

ويعرّف هف وجودمان(Huff & Goodman, 2007, P.20)تقويم المعرفة التشخيصي بأنه "الممارسات المشتركة لاستخدام النماذج المعرفية للتعلم كقاعدة لتصميم مبادئ التقويم وتدوين نتائجه، وربطها بشكل مباشر مع عمليتي التعلم والتعليم".

أهمية تقويم المعرفة التشخيصي:ألقى تعريف هف وجودمان السابق لتقويم المعرفة التشخيصي الضوء على أهمية هذا النوع من التقويم كونه يزود بتغذية راجعة أكثر غنى وفائدة للطالب والمعلم،تساعد في تعرف مشكلات التعلم ومعالجتها(Gierl et al., 2007) . كما أن استخدام تقييم المعرفة التشخيصي يزود المدرسين بمعلومات بنائية Formative Information ترفع مستوى تحصيل الطلبة (Nichols & Joldersma, 2008).

وعندما تستخدم نتائج تقويم المعرفة التشخيصي بشكل صحيح، فإنها ستكون دليلا لتصميم التعليم العلاجي، وتسكين الطلبة (Ketterlin-Geller & Yovanoff, 2009). وتمكّن الطلبة من إعادة النظر في الاستراتيجيات التي يتبعوها في الدراسة بالاعتماد على التغذية الراجعة للتقويم، وبهذا فإن محصلة تقويم المعرفة التشخيصي قد يكون لها أثر دال في الفرص التربوية المتاحة للطلبة؛ حيث من الممكن أن تزود بمعلومات قيّمة تساعد المختصين في وضع علاج تكيفي .(De la Torre & Douglas, 2004)وقد أهمل البحث في عمليات التعلم في أغلب البرامج الاختبارية المعاصرة (Zhou, 2010). وكان الاهتمام نحو الأساليب السيكومترية لتحليل البيانات أكثر من الاهتمام بالمعرفة قيد الاهتمام وعمليات استجابة الطلبة (Leighton et al., 2004; Gorin, 2006). لهذا، فإن المعلومات عن أسباب ضعف تحصيل الطلبة، وكيف يمكن الإفادة من نتائج التقييم في تحسين عملية التعلم والتعليم محدودة. ويؤكد ذلك أوزبل وبرونر (Ausubel,1963; Bruner, 1964) بقولهما إن التتابع يعد متغيرا مهما في عملية التعلم؛ حيث يبدأ المتعلم بمهمات بسيطة، ويتقدم في زيادة تعقيد المهمة، وبهذا اقترحت استراتيجية تعليم مبنية على التتابع؛ إذ أشار وانغ (Wang ,1973) إلى ضرورة إيجاد منهاج معتمد على التعلم الهرمي.

وإن فهم معرفة الطلبة المكتسبة والعمليات المعرفية أساسي للتشخيص، وقد يعمل على تحسين صدق الاختبار وثباته. لهذا السبب، فإن القياس الصادق لتحصيل الطلبة وعمليات التعلم تساعدان في تحسين كل من عمليتي التعليم والتعلم، وبهذا يكون تقييم المعرفة طريق مهم لتحسين جودة وصدق تفسير درجات الاختبار (Cui & Leighton, 2009; Embretson & Gorin, 2001)، ويمكن أن ينظر إليه كواحد من أكبر التحديات في حقل القياس والتقويم (Snow & Lohman, 1989; Pellegrino et al., 2001).

ويرى ديبلو وآخرون (DiBello, et al., 2007) أن هناك حاجة للكثير من البحث والعمل لتطوير مهارات التقييم البنائي. ولجمع بيانات التشخيص، تبرز الحاجة إلى استخدام نموذج معرفي تشخيصيCDM, Cognitive Diagnostic Model يصمم لتقويم التشخيص المعرفي لقياس البنية المعرفية والمهارات العملية للمفحوص، هذه المهارات مهمة لقياس التقويم التشخيصي المعرفي؛ لأنها تسمح بتعرف نقاط القوة والضعف لبنية معرفة المفحوص، وبالتالي يؤدي إلى عمل استدلالات عن مهارات حل المشكلات (Lohman, 2000). ولسوء الحظ، فإن هذه الأنواع من الاستدلالات المستندة لقاعدة المعرفة يصعب عملها، إن لم يكن مستحيلا دون إطار تفسيري صريح؛ بسبب الحجم الضئيل للاستدلالات. وتؤدي النماذج المعرفية هذا الغرض عن طريق التزويد بإطار ضروري لربط الاستدلالات المستندة لقاعدة معرفية مع تفسيرات لدرجات اختبارية محددة (Gierl& Leighton, 2007; Leighton &Girl, 2007a, 2007b). وتتولد النماذج المعرفية من دراسة المعرفة والعمليات والاستراتيجيات المستخدمة من المفحوصين عند استجابتهم على الفقرات. وبتوفير عدة مصادر للبيانات والاجراءات التحليلية، يمكن أن يساعد في دراسة التفكير وحل المشكلات بما فيها التحليل المنطقي، وإصدار الأحكام، ودراسات التعميم، وتحليل الاختلاف بين المجموعات، والارتباطات وتحليل التغاير(Messick, 1989) . ويهدف النموذج المعرفي إلى توجيه تطوير الفقرات، ويسمح بدراسة تحصيل المفحوصين لسمة ما؛ ليمكّن الوصول إلى استنتاجات معرفية محددة عن معرفتهم، وعملياتهم واستراتيجياتهم. وتتشكل النماذج المعرفية بتركيبات مختلفة للسمات. ويعرّفها لايتون وجيرل (Leighton & Gierl, 2007a, P.6) بأنها "وصف مبسط لحل المرء لمشكلات على شكل مهمات مقننة بشكل مناسب من حيث الحجم أو مستوى التفصيل؛ كي تسهل عملية الشرح والتنبؤ في أداء الطلبة، يتضمن ذلك نقاط القوة والضعف". ويشير تعبير وصف مبسط المستخدم في تعريفلايتونوجيرل السابق إلى أن النموذج المعرفي يعمل على تمثيل مبسط لظاهرة كثيرة التعقيد، هذا التمثيل يعكس قضايا صغيرة لكنها مجموعة مفاتيح مطلوبة لفهم الظاهرة (Leighton & Gierl, 2011). وتتعلق النماذج المعرفية أيضا بكيفية تطوير الأشخاص بنيتهم المعرفية، بما في ذلك المفاهيم المشتركة مع المجال والإجراءات المستخدمة للتبرير وحل المشكلات (Pellegrino et al., 2001). وعرض كل من جيرل وآخرون(Gierl et al., 2009) أربع خصائص تعريفية للنماذج المعرفية لتقويم المعرفة التشخيصي، وهي: القابلية للتقسيم Granularity، والقابلية للقياس Measurability، وهرمية ترتيب مهاراتها Hierarchy of Ordered Skills،وارتباطها بالتعليم Instructional Relevancy.

إن تطبيق تقويم المعرفة التشخيصي طريق يجمع بين المعرفة والخصائص السيكومترية، هذه العملية شرحها نيكولز وجولدرسما(Nichols & Joldersma, 2008) كالآتي: إن قياس النموذج يربط بين المعرفة والمهارات والقدرات بأداء ملاحظ في الاختبار. وتكون طريقة إعداد التقارير اللفظية مناسبة لدراسة عمليات المعرفة البشرية، من هنا فإن النماذج المعرفية عادة يتم

توليدها والتحقق من صدقها عن طريق جعل المفحوصين يفكرون بصوت عال عندما يقومون بحل الفقرات للتعرف إلى المعلومات المطلوبة، ومهارة العمليات المختارة لهذه المهمات (Ericsson & Simon, 1993; Leighton, 2004; Leighton &Gierl, 2007b; Royer et al., 1993)، ويتم بعد ذلك تقويم النموذج عن طريق مقارنة مطابقة البيانات للنموذج عبر نماذج منافسة. وبعد التقويم والمراجعة يصادق على النموذج ويكون ثقة، وبالتالي يمكن تعميمه لمجموعات أخرى من المفحوصين ولمهمات أخرى.

ويتم تحديد النموذج المعرفي بحجم صغير، لأنه مصمم لتعظيم وإبراز أو تشكيل بنية معرفة محددة ومهارات العمليات التي تقع وراء التحصيل في الاختبار. ويتم تحديد النموذج الهرمي للسمة Attribute Hierarchy Model, AHM بشكل هرمي للعمليات المعرفية؛ لأن السمات يفترض أن تتشارك بشكل اعتمادي مكّونة شبكة من العمليات والمهارات المترابطة داخليا (Anderson, 1996; Dawson, 1998; Fodor, 1983; Mislevy, et al., 2003)، وهذا الافتراض حول السمة مهم لبناء الاختبار لأن الفقرات التي تقيس السمة يجب أن تحافظ على البنية المعرفية بشكل هرمي، وتقيس بشكل مباشر عمليات معرفية محددة.

ويعد نموذج التقويم الهرمي AHMمن أحدث نماذج المعرفة التشخيصيCDMs، والذي طوره لايتون وجيرل وهنكا(Leighton, Gierl, &Hunka) في عام2000، ويعمل على تنظيم السمات في علاقات خطية أو غير خطية (Li & Suen, 2013). ويمكن استخدامه للتعرف إلى مستويات إتقان الطلبة لمهارات معينة. وتعتمد هذه الطريقة على أن سمة واحدة ترتبط مع الأخرى بترتيب هرمي، بحيث أن سمات معينة تكون بمثابة المتطلب السابق للسمات الأخرى. ويجري فحص مطابقة البيانات للنموذج إحصائيا باستخدام مؤشرات إحصائية (علام،1995 Gierl et al., 2008;

وتعد الطريقة الهرمية للسمة طريقة سيكومترية لتصنيف استجابات المفحوصين على فقرات الاختبار إلى مجموعة من الأنماط التركيبية للسمة المترابطة بعناصر مختلفة من النموذج المعرفي للتحصيل. وتمثل السمة وصف للمعرفة الإجرائية أو التفسيرية اللازمة لإنجاز المهمة في المجال المحدد. ويجب أن يمتلك المفحوص هذه السمات للإجابة عن الفقرات بشكل صحيح. وتشكل هذه السمات مهارات معرفية مطلوبة ومرتبة هرميا لحل الفقرات؛ حيث تمثل هرمية السمات نموذجا معرفيا لمهمة التحصيل، وتؤدي الهرمية دورا أساسيا في توجيه الاختبار والتحليل السيكومتري بحيث تمتلك درجات الاختبار قيمة تشخيصية، ففي دراسة أجراها علي(2014) هدفت إلى بناء اختبار للقدرات العقلية وفقاً لاستراتيجية التحليل البنائي الهرمي لدى طلبة المرحلة الإعدادية ، ولتحقيق هذا الهدف تم اعتماد نظرية فرنون وتصنيفه الهرمي للقدرات العقلية .ولغرض التحليل الإحصائي لفقرات الاختبار واستخراج الخصائص السيكومترية لها تم تطبيق الاختبار على

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

عينة التحليل الإحصائي البالغ قوامها 400 طالباً وطالبة اختيروا بالطريقة العشوائية من مديريات تربية بغداد ، لحساب الخصائص السيكومترية للفقرات المتمثلة بالصعوبة والتمييز وارتباط الفقرة بالدرجة الكلية ، وتحقق الباحث من (الصدق ، والثبات) إذ تم التحقق من الصدق من خلال الصدق الظاهري وصدق البناء ، وأيضاً تم التحقق من صدق البناء الهرمي للقدرات العقلية باستخراج نسبة الاتساق ونسبة كفاية الترتيب ونسبة الاكتمال ونسبة الاتساق العكسي ، ونسبة الترتيب العكسي ، واستخراج درجة صدق البناء الهرمي ، وكانت القيم المستخرجة تدل على صحة البنية الهرمية . أما ثبات الاختبار فقد حسب بطريقة الفا كرونباخ.

أهداف الدراسة

تهدف هذه الدراسة إلى تقديم استراتيجية لبناء اختبار تشخيصي محكي المرجع باتباع الطريقة الهرمية. ولتحقيق غرض الدراسة جرى الإجابة عن الأسئلة الآتية:

1. ما الخصائص السيكومترية لفقرات الاختبار؟
2. ما الخصائص السيكومترية للاختبار ؟
3. ما فاعلية البناء الهرمي للاختبار؟
4. ما درجة مطابقة استجابات الطلبة على الاختبار مع الاستجابة المتوقعة من النموذج الهرمي الذي صمم للمحتوى؟

فروض الدراسة

تقوم هذه الدراسة على أن الطريقة الهرمية تفترض أن التزويد بالإجابة الصحيحة للفقرة الاختبارية يعتمد على إتقان المفحوص للسمة المعرفية المطلوبة كمتطلب سابق، وبكلمات أخرى، تفترض الطريقة الهرمية أن السمات المعرفية مبنية بطريقة هرمية.

محددات الدراسة

تتحدد نتائج هذه الدراسة بمدى جدية المفحوصين في الإجابة عن فقرات الاختبار، ودرجة ثبات الاختبار المعد لهذه الدراسة.

التعريفات الإجرائية للمتغيرات

النموذج الهرميHierarchy models. يعرّف بأنه تمثيل مرئي ينظم السمات المحددة بطريقة معينة؛ حيث يوضح ترتيب السمات في النموذج الهرمي المهارات ذات المتطلب السابق التي يحتاجها المفحوص للإجابة عن الفقرة بشكل صحيح متقن(Wang &Gierl, 2011) ، حيث تزود النماذج الهرمية بتمثيل لتركيبة المعرفة (Gierl et al., 2008) .

تقويم المعرفة التشخيصي. يعرّف بأنه "الممارسات المشتركة لاستخدام النماذج المعرفية للتعلم كقاعدة لتصميم مبادئ التقويم، وتدوين نتائجه، وربطها بشكل مباشر مع عمليتي التعلم والتعليم"(Huff & Goodman, 2007, P.20).

الطريقة والإجراءات

المنهجية

استخدم المنهج الوصفي في هذه الدراسة لملاءمته أهداف الدراسة وطبيعتها.

مجتمع الدراسة

اختير طلبة الصف العاشر الأساسي في المدارس الحكومية للعام الدراسي 2020/2019 في الأردن.

عينات الدراسة

العينة الرئيسية. اختيرت عينة عشوائية عنقودية على مستوى المدرسة، تكونت من 390 طالباً وطالبة منهم 184 ذكور، وذلك على النحو الآتي: أولاً أقاليم المملكة، وكان إقليم الشمال، ومن هذا الإقليم اختيرت إحدى المحافظات عشوائيا، فكانت محافظة إربد، ومن هذه المحافظة اختيرت إحدى المديريات عشوائيا، فكانت مديرية تربية إربد الأولى، ومن هذه المديرية أعدت قائمة بأسماء المدارس التي تحتوي على الصف العاشر الأساسي، ومن هذه القائمة اختيرت أربع مدارس بالطريقة العشوائية البسيطة (القرعة مع الإرجاع) بواقع مدرستين للذكور، ومدرستين للإناث.

العينة الاستطلاعية. تكونت من 97 طالبا وطالبة من المجتمع ومن خارج عينة الدراسة؛ حيث تم اختيار مدرسة للذكور وأخرى للإناث بالطريقة العشوائية البسيطة ثم تم اختيار شعبة من شعب الصف العاشر في كل مدرسة أيضا بالطريقة العشوائية البسيطة(القرعة) ، وجرى استخدامها لحساب معاملات الصعوبة والتمييز، وفعالية المموهات لفقرات الاختبار، كما تم التأكد من معاملات صدق الاختبار وثباته، والتأكد من وضوح فقرات الاختبار، ومقروئيتها، وتسلسل المهارات والعمليات بشكل هرمي، والتعرف إلى الأخطاء (تحليل أنماط الأخطاء تصنيف أنماط الأخطاء)، وكذلك تقدير الزمن المطلوب للإجابة عن فقرات الاختبار.

أداة الدراسة. بني اختبار تشخيصي لقياس مهارة "حل نظام مكون من معادلتين تربيعيتين بمتغيرين"،قسّم إلى ست مهارات فرعية تقيس حل نظام مكون من: 1) معادلتين تربيعيتين بمتغيرينx,y، 2) معادلة تربيعية بمتغيرينx,y، ومعادلة تربيعية يظهر فيها الحدxy مع حدود أخرى،3) معادلة تربيعية بمتغيرينx,y، ومعادلة تربيعية يظهر فيها الحدxy وحده، 4)معادلة تربيعية بمتغيرينx,y، ومعادلة تربيعية يظهر فيها التربيع على متغير واحد، 5) معادلة تربيعية يظهر فيها الحدxy وحده، ومعادلة كسرية بمتغيرينx,y، 6) معادلة تربيعية بمتغيرينx,y، ، ومعادلة يظهر فيها مربع كامل على أحد المتغيرين على الأقل.

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

واتبعت خطوات بناء الاختبار المحكية الواردة في كتب القياس من تحديد الغرض من الاختبار، وتحديد محتوى الاختبار تحديدا دقيقا ومفصلا، وصياغة الأهداف المتعلقة به، وعرض الأهداف على مجموعة من المحكمين لتقدير أولويات هذه الأهداف على مقياس ثلاثي التدريج (3=هام جدا، 2= هام ، 1= قليل الأهمية)، ومن ثم جرى مراجعة قائمة الأهداف وفقا لنتائج التقييم الواردة من المحكمين، وبناء تجمع فقرات من نوع الاختيار من متعدد لكل فقرة أربعة خيارات، وعرضت الأهداف والاختبار على مجموعة من المحكمين؛ للتحقق من أن كل فقرة تعد مقياسا صادقا للهدف المناظر له؛ وذلك عن طريق قيام المحكم بتحديد درجة موافقته على مطابقة الفقرة للهدف من خلال مقياس ثلاثي (3= متأكد جدا،2= متأكد،1=غير متأكد)، وطلب إلى المحكمين اقتراح التعديل الذي يرونه مناسبا على الفقرة، بعد ذلك عدّلت الفقرات في ضوء التغذية الراجعة من المحكمين، وطبق الاختبار على العينة الاستطلاعية، وجرى مناقشة الطلبة في الأخطاء التي ارتكبوها، وأعد الاختبار الذي اشتمل في صورته النهائية على 48 فقرة، وإعداده للتطبيق، حيث احتوى الاختبار على فقرات تمثل المهارات الفرعية اللازمة لتحقيق المهارة التي يهدف الاختبار إلى تحقيقها، وتترتب هذه الفقرات هرميا؛ بحيث تعتمد قدرة المفحوص في حل المهارة الأعلى على حله للمهارات الأدنى.

وبعد انتهاء المرحلة الأولى، وهي تطوير المعرفة التي كان الغرض منها التعريف بالسمات المحددة في الاختبار، وتطوير بنية هرمية تمثل هذه السمات، جرى فيها تحديد السمات المعرفية والبناء الهرمي المستخدم من المفحوصين في أثناء اجابتهم فقرات الاختبار؛ تضمنت الخطوة الاولى مراجعة المحكمين والمختصين للموضوع الدراسي قيد التقييم؛ وذلك عن طريق إعداد وصف تفصيلي للاختبار يتضمن تحديد المحتوى، حيث اختير موضوع أنظمة المعادلات، وذلك أن عدة دراسات أشارت إلى وجود ضعف عند الطلبة في التعامل مع حل المعادلات بكافة أنواعها (Carry et al., 1980; Kieran, 1992; Napaphun, 2012) ، واقتصرت الدراسة الحالية على موضوع واحد، وهو حل نظام مكون من معادلتين تربيعيتين بمتغيرين. وتم العمل على التحليل السيكومتري، حيث يتم اختبار الخاصية الهرمية إحصائيا (Wang & Gierl, 2011)، يتضمن التحليل استخراج نسبة الاتساق، ونسبة كفاية الترتيب، ونسبة الاكتمال.

وجرى بعد ذلك تحديد المجموعة التي سيطبق عليها الاختبار، باختيار طلبة الصف العاشر الأساسي، حيث يختم منهاجه موضوع أنظمة المعادلات بالنسبة لطلبة المرحلة الأساسية؛ أي ما يلزم أي طالب في أنظمة المعادلات أيا كان اختياره في المرحلة الثانوية ليكمل مسيرته العلمية. وجرى تحليل المحتوى، وتقصي المهارات الأساسية في موضوع حل نظام مكون من معادلتين تربيعيتين بمتغيرين، بالاستعانة بالكتاب المدرسي، ودليل المعلم للصف العاشر الأساسي، ومراجعة محتوى كتب الرياضيات للصفوف السابع والثامن والتاسع الأساسيين؛ وذلك لربط المعرفة، وتحديد المتطلبات السابقة من عناصر المعرفة الضرورية للطالب؛ كي يحقق مهارة حل نظام مكون من معادلتين تربيعيتين بمتغيرين، بالاستعانة بمجموعة من خبراء التحكيم الذين قاموا بتحليل الموضوع، وتنظيم المعرفة المتعلقة به تنظيما هرميا.

إعداد مواصفات الاختبار

1. تحديد الغرض والمحتوى للاختبار، وهو تشخيص مواطن الضعف في موضوع حل نظام مكون من معادلتين تربيعيتين بمتغيرين لطلبة الصف العاشر الأساسي.
2. تحليل موضوع حل نظام مكون من معادلتين تربيعيتين بمتغيرين تحليلا تفصيليا، ويظهر جدول 1 هذه المواصفات لتحليل المحتوى، حيث يمثل المستوى الأول المتطلب السابق لأشكال المستوى الثاني.

جدول1

تحليل تفصيلي لمحتوى الاختبار ومتطلباته السابقة مع أمثلة توضيحية

المستوى / التوصيف	مثال
المستوى الأول: م1	
1. جعل المتغير y موضوع المعادلة	أكتب y بدلالة x في المعادلة $2x+5y=3$
2. حل معادلة خطية بمتغير واحد	$2x-6=13$
3. حل معادلة تربيعية بمتغير واحد	$2x^2+4x-16=0$
4. حل معادلة من الدرجة الرابعة بمتغير واحد	$2x^4-6x^2=8$
5. كتابة معادلة خطية بمتغيرين	$2x+5y=3$
6. تحليل المعادلة التربيعية بمتغيرين	$x^2-y^2=64$
7. طريقة الحذف أو التعويض لحل المعادلات	حل المعادلتين $2x+5y=3$ و$3x-6y=13$
8. فك مربع كامل	$(x+y)^2=x^2+y^2+2xy$
9. إيجاد مضاعف مشترك أصغر لمقدارين جبريين خطيين بمتغير واحد	ما المضاعف المشترك الأصغر للمقدارين $x-1$، $2x+3$
10. جمع كسرين جبريين مقامهما مقدارين جبريين خطيين بمتغير واحد	$(3-2x)\1 + (5+x)\3$
11. حل معادلة كسرية خطية بمتغير واحد	$6=5 + (2-x)\2$
12. حل معادلة كسرية تربيعية بمتغير واحد	$3=5x^2 +(x^2\3)$
المستوى الثاني:م2	
1. حل معادلتين تربيعيتين بمتغيرين*	$3x^2+2y^2=14$ $4x^2-3y^2=13$
2. حل معادلة تربيعية بمتغيرين، ومعادلة تربيعية يظهر فيها الحدxy مع حدود أخرى	$2x^2-y^2=7$ $3x^2+xy=2$
3. حل معادلة تربيعية بمتغيرين، ومعادلة تربيعية يظهر فيها الحدxy وحده	$x^2+y^2=20$ $xy=8$
4. حل معادلة تربيعية بمتغيرين، ومعادلة تربيعية يظهر فيها التربيع على متغير واحد	$x^2+y^2=4$ $x^2-2y=44$
5. حل معادلة تربيعية يظهر فيها الحدxy وحده ومعادلة كسرية بمتغيرين	$xy=18$ $6\1=y\1- x\1$
6. حل معادلة تربيعية بمتغيرين ومعادلة تربيعية ومعادلة يظهر فيها مربع كامل على أحد المتغيرين على الأقل	$x^2-y^2=24$ $x^2=(y+4)^2$

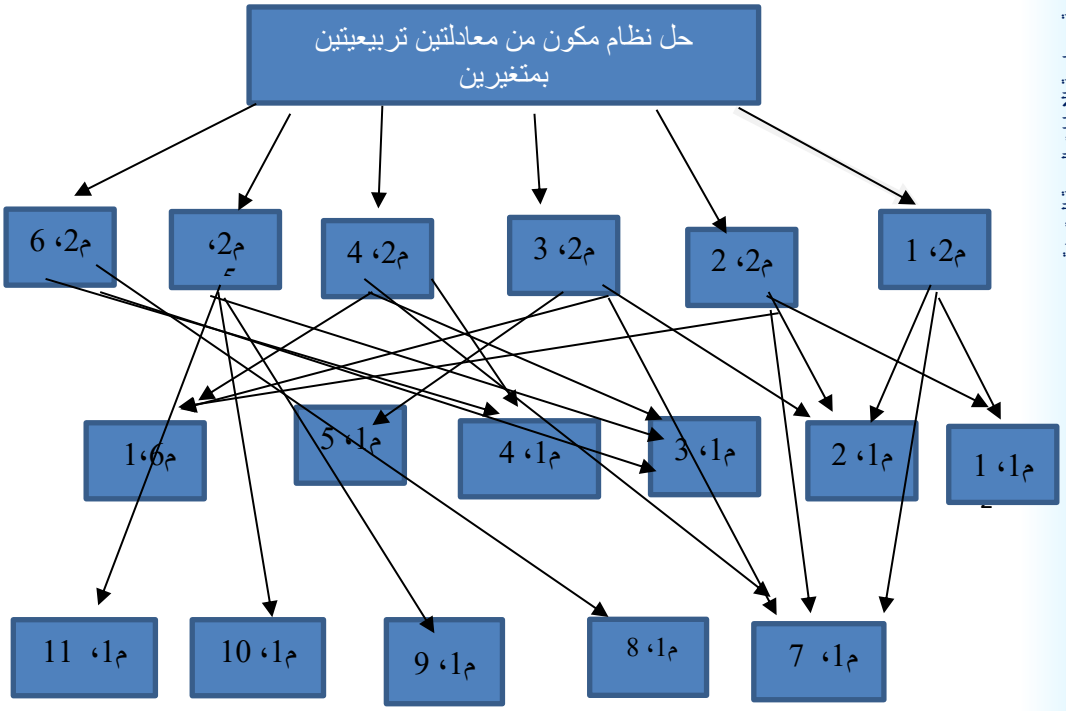
*تم التعبير عن كل مستوى فرعي بالرسم برمز المستوى م2، ورقم من 1 إلى 6

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

تم رسم تخطيط هرمي لهذا التحليل، ويبين شكل 1 البناء الهرمي لمهارة حل نظام مكون من معادلتين تربيعيتين بمتغيرين

شكل 1

البناء الهرمي لمهارة حل نظام مكون من معادلتين تربيعيتين بمتغيرين



ولتحديد مواصفات الفقرات، جرى إعداد مواصفات لكل فقرة تضمن الهدف الذي تقيسه الفقرة، وعدد الفقرات التي تقيس كل هدف، ووصف تفصيلي إجرائي للفقرة (المثير والاستجابة)، ووصف للمموهات، وعقدت جلسات لمناقشة آلية بناء الاختبار، وتسلسل الفقرات، وربطها بالأهداف التي أعدت لهذا الغرض، وعليه فقد أعدت الباحثة اختبار يقيس مهارة حل نظام مكون من معادلتين تربيعيتين بمتغيرين من ست اختبارات فرعية بلغ مجموع الفقرات 54 فقرة، بحيث خصص لكل مهارة فقرتان موزعة كالآتي:

الجزء الأول ويتكون من 3 اختبارات فرعية جميع الفقرات فيها تخلو من الحدxy

الاختبار الأول حل نظام مكون من معادلتين تربيعيتين، تكون الاختبار من ست فقرات تقيس مهارات: 1) حل معادلة تربيعية بمتغير واحد 2) حل معادلتين تربيعيتين بمتغيرين بالحذف أو التعويض 3) كتابة الحل باستخدام الأزواج المرتبة.

الاختبار الثاني حل نظام مكون من معادلة تربيعية بمتغيرين ومعادلة تربيعية يظهر فيها التربيع على متغير واحد، تكون الاختبار من ثماني فقرات تقيس مهارات :1) حل معادلة تربيعية بمتغير واحد ،2) حل معادلة من الدرجة الرابعة بمتغير واحد 3) حل معادلتين بمتغيرين بالحذف أو التعويض 4) كتابة الحل باستخدام الأزواج المرتبة.

الاختبار الثالث حل نظام مكون من معادلة تربيعية بمتغيرين، ومعادلة تربيعية يظهر فيها مربع كامل على أحد المتغيرين على الأقل، تكون الاختبار من عشر فقرات تقيس مهارات :1) حل معادلة خطية بمتغير واحد 2) فك المربع الكامل 3) حل معادلة تربيعية بمتغير 4) حل معادلتين بمتغيرين بالحذف أو التعويض 5) كتابة الحل باستخدام الأزواج المرتبة.

الجزء الثاني ويتكون من ثلاث اختبارات فرعية جميع الفقرات تحوي الحدxy

الاختبار الرابع حل نظام مكون من معادلة تربيعية بمتغيرين، ومعادلة تربيعية يظهر فيها الحدxy وحده،تكون الاختبار من اثنتي عشرة فقرة تقيس مهارات: 1) حل معادلة خطية بمتغير واحد 2) كتابة أحد المتغيرين بدلالة الآخر 3) حل معادلة تربيعية بمتغير 4) حل معادلة من الدرجة الرابعة بمتغير 5) حل معادلتين بمتغيرين بالحذف أو التعويض 6) كتابة الحل باستخدام الأزواج المرتبة.

الاختبار الخامس حل نظام مكون من معادلة تربيعية بمتغيرين ومعادلة تربيعية يظهر فيها الحدxy مع حدود أخرى، تكون الاختبار من ثماني فقرات تقيس مهارات :1) كتابة أحد المتغيرين بدلالة الآخر 2) حل معادلة تربيعية بمتغير 3) حل معادلتين بمتغيرين بالحذف أو التعويض 4) كتابة الحل باستخدام الأزواج المرتبة.

الاختبار السادس حل نظام مكون من معادلة كسرية بمتغيرين، ومعادلة تربيعية يظهر فيها الحدxy وحده، تكون الاختبار من ثماني فقرات تقيس مهارات: 1) كتابة أحد المتغيرين بدلالة الآخر 2) حل معادلة خطية بمتغير3) حل معادلة تربيعية بمتغير 4) إيجاد المضاعف المشترك الأصغر لمقدارين 5) كتابة الحل باستخدام الأزواج المرتبة.

1. جرى اختيار أربعة من المعلمين الأكفاء للمرحلة الأساسية العليا لمادة الرياضيات بمساعدة مشرفي الرياضيات في مديرية التربية والتعليم ومدراء المدارس، قاموا بعقد مناقشات مفصلة لكل فقرة للتعرف إلى المهارات التي يحتاجها المتعلم للإجابة عن الفقرة، وجرى تحليل كل مهارة من المهارات الأساسية إلى المهارات الجزئية، ثم قاموا بوضع تعريف لكل سمة، حيث جرى تطوير عدة بنى معرفية لتوضيح العلاقات الهرمية بين السمات من أجل التحليل السيكومتري، كما جرى تطبيق الاختبار على العينة الاستطلاعية، وطلب إليهم توضيح خطوات الحل؛ من أجل التأكد من تسلسل المهارات الفرعية.

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

2. تحديد خبراء التحكيم للاختبار: تكّون فريق التحكيم من ستة من أساتذة الجامعات؛ اثنان منهم مختصان في الرياضيات، واثنان في مناهج الرياضيات وأساليب تدريسها، وأستاذان جامعيان في تخصص القياس والتقويم وشهادة البكالوريوس لهما في الرياضيات، ولهما خبرة طويلة في بناء الاختبارات، إضافة إلى أربعة من معلمي الرياضيات الأكفاء للمرحلة الثانوية، ومشرفين اثنين للرياضيات يحملان مؤهلا تربويا بعد الدرجة الجامعية الأولى في وزارة التربية والتعليم.

وضوح تعليمات الاختبار وفقراته

بغرض معرفة وضوح كل فقرة من فقرات الاختبار ووضوح التعليمات وكيفية الإجابة، طبقت الفقرات بعد تعديلها على العينة الاستطلاعية المكونة من 97 طالباً وطالبة؛ وذلك للتأكد من وضوح فقرات الاختبارات ومقروئيتها، وكذلك تقدير الزمن المطلوب للإجابة عن فقرات الاختبار، وقد طلب منهم قراءة التعليمات والفقرات والاستفسار عن أي غموض من الممكن أن يواجه الطالب في أثناء الاستجابة، وقد أوضحت التجربة أن متوسط الوقت المستغرق للإجابة عن الاختبار 56 دقيقة تقريباً ضمن مدى أعلى وقت 74 دقيقة وأقل وقت 38 دقيقة.

طبقت الفقرات في صورتها النهائية على عينة الدراسة الرئيسية المكونة من 390 طالبا وطالبة، وصححت الأوراق وأعطيت كل إجابة صحيحة درجة واحدة أما الإجابة الخاطئة أو المتروكة فتعطى صفراً. ثم جمعت درجات الإجابات لتمثل الدرجة الكلية للطالب. ولأن الاختبار يتكون من 46 فقرة كانت أقصى درجة يمكن أن يحصل عليها الطالب 46 درجة وأدنى درجة صفر، وفرغت النتائج في جداول معدة لذلك من أجل التحليل الإحصائي.

عملية جمع ومعالجة وتحليل البيانات

1. -بعد بناء الاختبار التشخيصي وفق الأسلوب الهرمي مكوّن في صورته الأولية من 52 فقرة،
2. -جرى التحقق من صدق المحتوى للاختبار عن طريق عرضه مع مجموعة الأهداف التي يقيسها على مجموعة المحكمين.
3. التأكد من أن الفقرات مبنية بناء هرميا تراكميا باعتبارها مهارات أساسية يجب على طلبة الصف العاشر الأساسي إتقانها وفق الأسلوب الهرمي في بناء الاختبار، وجرى التحقق من صدق البناء الهرمي عن طريق:

- عرض البناء الهرمي المقترح على مجموعة المحكمين؛ للتحقق من صدق البناء الهرمي لعناصر الموضوع قيد الدراسة عن طريق العمل على الحصول على فهم مشترك لكل ما يتعلق بالفقرات، مثل: مواصفات الفقرة (المثير والبدائل)، وصدق محتوى الفقرة. تم بعد ذلك إعطاء حكم على البناء الهرمي المقترح من حيث درجة صحة العلاقات الهرمية وانتماؤها،

وأعطيت القيم 1، 2، 3 ، 4 على الترتيب، حيث تشير الأرقام إلى (غير صحيحة / غير منتمية)، (صحيحة / غير منتمية)، (غير صحيحة / منتمية)، (صحيحة / منتمية)، وتفريغ النتائج وفق التقسيمات الأربعة وفق جدول 2 الآتي

جدول 2

المقياس الرباعي للحكم على صدق البناء الهرمي

		المحكم 2	
		ارتباط ضعيف (1،2)	ارتباط قوي (3،4)
المحكم 1	ارتباط ضعيف (1،2)	A	B
	ارتباط قوي (3،4)	C	D

وطبقت معادلة1 (علام،1995؛ Gregory, 2004)، لتعطي درجة الارتباط

درجة الارتباط =(1-D/(A+B+C+D)) (1)

حيث تشير A إلى عدد الاتفاقات بين المحكم الأول والثاني على الضعف، في حين تشير D إلى عدد الاتفاقات بين المحكم الأول والثاني على القوة.

تم بعد ذلك حساب صدق البناء الهرمي عن طريق حساب معدل جميع الأزواج الممكنة وفق معادلة 1، وتراوحت معدلات تقديرات المحكمين بين 0.86 و0.92 على 48فقرة، في حين كانت التقديرات أقل من 0.80 لست فقرات، وبذلك تم حذف هذه الفقرات فأصبح الاختبار مكونا من 48 فقرة.

التحقق من صحة الفرضيات المتعلقة بمستوى التحليل في البناء الهرمي.

وللإجابة عن السؤال الأول ونصه" ما الخصائص السيكومترية لفقرات الاختبار؟"

جرى تطبيق الاختبار (نسخة بعد التحكيم) وعدد فقراته 48 على العينة الاستطلاعية المكونة من 97طالبا وطالبة؛ وذلك لحساب معاملات الصعوبة والتمييز للفقرات؛ حيث حسب معامل تمييز الفقرة عن طريق طرح نسبة الطلبة غير المتقنين الذين أجابوا عن الفقرة إجابة صحيحة من نسبة الطلبة المتقنين الذين أجابوا عن الفقرة إجابة صحيحة، وقد حددت الفئة غير المتقنة بالفئة التي لم تتجاوز علامات الطلبة فيها علامة القطع 71% وفقاً لطريقة ندلسكيNedlesky,s Method المعتمدة على التحكيم،حيث يطلب من كل محكم فحص كل فقرة في الاختبار، وأن يحدد من بين بدائل كل فقرة تلك التي يمكن أن يتجنب اختيارها الأفراد من ذوي الحد الأدنى للكفاية في المجال الذي يقيسه الاختبار، ويكون الحد الأدنى لمستوى اجتياز الفقرة هو مقلوب عدد البدائل المتبقية للفقرة (علام،1995).

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

ولحساب معامل صعوبة كل فقرة فقد تم حسابه عن طريق إيجاد نسبة الطلبة الذين أجابوا عن الفقرة إجابة صحيحة.

وجرى حساب فاعلية المموهات لكل فقرة، عن طريق طرح نسبة الطلبة غير المتقنين الذين اختاروا المموه ،ويوضح جدول 3 القيم المحسوبة لمعاملات الصعوبة والتمييز وفعالية المموهات لفقرات الاختبار الكلي48فقرة.

جدول 3

معاملات الصعوبة والتمييز وفاعلية المموهات لفقرات الاختبار الكلي(48 فقرة)

رقم الفقرة /الاختبار	معامل الصعوبة	معامل التمييز	فاعلية المموهات		
			الأول	الثاني	الثالث
1/1	0.42	0.51	-0.19	-0.22	-0.10
1/2	0.63	0.53	-0.20	-0.17	-0.16
1/3	0.58	0.49	-0.14	-0.23	-0.11
1/4	0.62	0.59	-0.21	-0.16	-0.20
1/5	0.55	0.47	-0.15	-0.22	-0.09
1/6	0.68	0.61	-0.25	-0.19	-0.17
2/7	0.65	0.55	-0.16	-0.24	-0.14
2/8	0.49	0.53	-0.21	-0.19	-0.14
2/9	0.66	0.54	-0.23	-0.20	-0.12
2/10	0.62	0.53	-0.18	-0.15	-0.20
2/11	0.59	0.46	-0.16	-0.18	-0.12
2/12	0.44	0.50	-0.12	-0.18	-0.20
2/13	0.69	0.46	-0.17	-0.20	-0.09
2/14	0.60	0.58	-0.21	-0.15	-0.21
3/15	0.48	0.54	-0.20	-0.18	-0.15
3/16	0.53	0.59	-0.20	-0.15	-0.23
3/17	0.66	0.60	-0.24	-0.19	-0.18
3/18	0.55	0.56	-0.17	-0.21	-018
3/19	0.23	0.04	0.03	-0.06	-0.009
3/20	0.67	0.61	-0.21	-0.20	-0.19
3/22	0.65	0.58	-0.23	-0.19	-0.17
3/23	0.61	0.52	-0.20	-0.19	-0.13
3/24	0.60	0.55	-0.16	-0.24	-0.14
4/25	0.50	0.56	-0.20	-0.18	-0.12
4/26	0.56	0.57	-0.20	-0.15	-0.21
4/27	0.68	0.59	-0.20	-0.18	-0.18
4/28	0.48	0.52	-0.15	-0.23	-0.13
294/	0.59	0.56	-0.20	-0.21	-0.15
304/	0.53	0.56	-0.19	-0.21	-0.16
314/	0.51	0.59	-0.16	-0.22	-0.21
5/ 32	0.55	0.53	-0.14	-0.23	-0.15
5/33	0.60	0.57	-0.21	-0.16	-0.20

رقم الفقرة /الاختبار	معامل الصعوبة	معامل التمييز	فاعلية المموهات		
			الأول	الثاني	الثالث
5/35	0.59	0.56	-0.19	-0.22	-0.13
5/36	0.68	0.61	-0.25	-0.19	-0.17
5/37	0.17	0.12	-0.06	-0.10	-0.05
5/38	0.65	0.51	-0.18	-0.19	-0.13
5/39	0.68	0.60	-0.23	-0.20	-0.18
5/40	0.63	0.55	-0.20	-0.19	-0.13
6 /41	0.65	0.68	-0.20	-0.18	-0.17
6 /42	0.54	0.61	-0.20	-0.18	-0.17
6 /43	0.63	0.78	-0.15	-0.10	-0.21
6 /44	0.59	0.60	-0.05	-0.13	-0.22
6 /45	0.60	0.62	-0.15	-0.18	-0.22
6 /46	0.49	0.59	-0.12	-0.20	-0.08
6 /47	0.61	0.64	-0.32	-0.11	-0.13
6 /48	0.67	0.42	-0.18	-0.21	-0.17

يبين جدول 3 أن معاملات الصعوبة للفقرات تراوحت بين 0.17 و0.69، وهي قيم تشير إلى فقرات ذات صعوبة متوسطة ومقبولة بناء على المعايير الواردة في الأدب التربوي والخاصة بالاختبارات (Lord, 1952; Thompson &Levitov, 1985)، ما عدا الفقرة 37 من الاختبار الخامس.

ويظهر جدول 3 أن قيم معاملات التمييز للفقرات تراوحت بين 0.04و0.78؛ مما يشير إلى أن جميع الفقرات تتمتع بمعاملات تمييز مقبولة بناء على المعايير الواردة في دراسة بيرك للاختبارات المحكية (Berk, 1982)، ما عدا الفقرة 19 من الاختبار الثالث.

وأظهرت النتائج الخاصة بفاعلية المموهات أن قيم المعاملات وقعت بين -0.05و-0.32؛ وهي قيم جميعها تعبر عن مموهات جذابة، ما عدا مموه للفقرة 19 كان 0.03.

وبناء عليه فقد كانت جميع فقرات الاختبار تحقق المعايير المقبولة في الأدب التربوي لمؤشرات الصعوبة والتمييز، وفعالية المموهات، ما عدا الفقرتين 19، 37 من الاختبارين الثالث والخامس قد أظهرتا معاملات الصعوبة والتمييز لهما غير مقبولة وفق معايير الأدب التربوي للاختبارات المحكية(Berk, 1982) ؛ مما أدى إلى حذف الفقرتين، وبناء على ذلك أصبح عدد فقرات الاختبار الكلي 46 فقرة.

وللإجابة عن السؤال الثاني ونصه"ما الخصائص السيكومترية للاختبار؟"

مؤشر ثبات الاختبار: جرى التحقق من ثبات الاختبار عن طريق حساب معامل الاتفاق لسابكوفياك التي تعتمد على درجة اتساق تصنيف الأفراد في مجموعتين، إحداهما متمكنة، والأخرى

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

غير متمكنة باستخدام نتائج تطبيق الاختبار مرة واحدة، وقد بلغت قيمة معامل الاتفاق 0.84 عند علامة قطع 71%. وبهذا يمكن القول إن الاختبار يتمتع بدرجة ثبات جيدة.

واستخرجت مؤشرات الصدق الآتية:

1-الصدق الوصفي Descriptive Validity.

عرضت فقرات الاختبار بصيغتها الأولية على مجموعة من المحكمين المتخصصين لتقويم صلاحية الفقرات في قياس أنظمة المعادلات (فريق التحكيم في الدراسة). وللتحقق من أن المجال السلوكي تم تحديده بشكل واضح ودقيق، وأن مجموعة الفقرات التي تم إعدادها تعد كافية وممثلة للمجال استخدمت طريقة التحقق من تجانس الفقرات Item Homogeneity (Rovinelli&Hambleton,1977)؛ التي تهدف إلى التحقق من مدى اتفاق المحكمين حول تجانس الفقرات وفقا للمواصفات التفصيلية، وما إذا كانت كل فقرة تنتمي إلى المجال السلوكي الذي تحدده هذه المواصفات، وجرى التحقق من ذلك عن طريق تجميع الفقرات المتعلقة بالمواصفات التي جرى تحديدها، وصممت استمارة كتبت فيها المواصفات التفصيلية، ووضعت الفقرات أمام المواصفات المناظرة لها في عمود مقابل، ووضع عمود يحوي مربعين ليضع المحكم رأيه بالفقرة وفق التعليمات الآتية:

1. وضع الرقم +1 داخل المربع المناسب إذا كان متأكدا من أن الفقرة لا تنتمي إلى نطاق سلوكي آخر غير المحدد أمامها، والرقم صفر إذا كان غير متأكد من ذلك، والرقم -1 إذا كان متأكدا أن الفقرة تنتمي إلى أكثر من مجال سلوكي من المجالات المحددة في بقية المواصفات.
2. تكوين مصفوفة لكل فقرة بحيث تشتمل على صفوف عددها 12(بقدر عدد المحكمين)، وأعمدة عددها (بقدر عدد مجموعات المواصفات)؛ وبهذا تشكلت الخلايا الناتجة من تقاطع الصفوف والأعمدة لتشير إلى تقديرات المحكمين للفقرة في إحدى مجموعات المواصفات.
3. إيجاد ICI الذي يشير إلى مدى اتفاق الفقرة مع مجموعة المواصفات المتعلقة بها، وانتمائها إلى المجال السلوكي الذي تحدده المواصفات، وذلك وفق معادلة 2

$$ICIr=\frac{(n-1)\,\Sigma xrci-\Sigma xhci+\Sigma xrci}{2(n-1)\,k} \qquad (2)$$

حيث:

n: تشير إلى عدد مجموعات المواصفات

k: تشير إلى عدد المحكمين

الحد الأول في البسط يشير إلى مجموع تقديرات المحكمين للفقرة i على أنها تتفق مع مجموعة المواصفات r. ويشير الحد الثاني في البسط إلى مجموع جميع التقديرات المدونة في خلايا المصفوفة. فيما يشير الحد الثالث في البسط إلى مجموع التقديرات المدونة في العمود الذي يرمز

إلى مجموعة المواصفات (r) في المصفوفة، حيث تكون إضافة هذا الحد إلى الفرق بين الحدين الأول والثاني في البسط تمكن من تحديد مدى عدم اتفاق الفقرة (i) مع مجموعة المواصفات الباقية وعددها (n-1) (علام، 1995).

وقد أشارت النتائج إلى أن معاملات تجانس الفقرات الاختبارية تراوحت بين 0.86و 0.93مما يشير إلى أن الفقرات تتفق إلى حد ما مع مجموعة المواصفات المتعلقة بها.

وقام المعلمون في لجنة التحكيم بعقد مناقشات مفصلة لكل فقرة للتعرف إلى المهارات التي يحتاجها المتعلم للإجابة عن الفقرة، وجرى تحليل كل مهارة من المهارات الأساسية إلى المهارات الجزئية، ثم قاموا بوضع تعريف لكل سمة كآلية لتطوير بنية الترميز، حيث جرى تطوير عدة بنى معرفية لتوضيح العلاقات الهرمية بين السمات؛ من أجل التحليل السيكومتري، بعد ذلك عقدت جلسات لمناقشة تحليلات المعلمين.

وبعد الانتهاء من دراسة جميع الفقرات، جرى تحديد السمات وكتابتها بالاعتماد على العوامل الآتية: 1) ثبات المقدرين، بمعنى أنه جرى وصف السمات بطريقة يستطيع من خلالها المقدرين وبشكل مستقل استخدام الوصف لتطبيق السمة باتساق لفقرات الاختبار 2) تنظيم السمات بترتيب من السهل إلى الصعب 3) التعبير بشكل كلمات عن السمات بطريقة تكون ذات معنى للمعلمين، والطلبة والمعلمين. والنقطة الأخيرة تسمح بإجراء تغذية راجعة لمعرفة الطالب من نقاط قوة وضعف التي تتيح الفرصة في تحسين التعلم والتعليم، وكل سمة جرى اختبارها وكتابتها بطريقة محددة بشكل كاف بحيث تسمح وبشكل مستقل تطبيق وفهم السمة.

وفي اللقاء الثاني للمحكمين اتبعت طريقة جيرل وسيو(Gierl, & Cui, 2008) ؛ حيث أجري ما يأتي: 1) مراجعة السمات التي تم الاتفاق عليها في اللقاء الأول 2) إنشاء قائمة رصد مرمزة Coding Checklist؛ للتحقق من ثبات واتساق المقدرين، وتم تطوير قائمة الرصد للتزويد بأمثلة وتعريف مفصل لمكونات كل سمة. وبعد الانتهاء من قائمة الرصد، أعطي كل محكم نسخة من الاختبار وقائمة الرصد 3) الطلب إلى كل محكم وبشكل مستقل أن يحدد السمات المصاحبة لكل فقرة، وأن يظلل موقع السمة كما فكّر فيها، ويشرح ذلك لبقية المحكمين لإبداء الرأي بالموافقة أو عدم الموافقة، وبالنسبة للفقرات التي لم يتم الاتفاق عليها أجريت المناقشة، بحيث يبدأ المناقشة المحكم الذي استهل المناقشة في بداية الجلسة، بحيث يفكر المحكمين بصوت مسموع كيف توصلوا لإجابتهم، وكيف فكروا فيها.

- إعداد الاختبار بعد الانتهاء من تحديد المواصفات التفصيلية الشاملة.
- التحقق من صدق المحكمين للاختبار، بعرض الاختبار والمواصفات التفصيلية الشاملة على لجنة المحكمين، حيث تم التأكد من توافر الخصائص المناسبة لهذه الفقرات وبدائلها من

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

حيث الشكل والمضمون، وأن الفقرات مبنية بناء هرميا تراكميا باعتبارها مهارات أساسية يجب على طلبة الصف العاشر الأساسي إتقانها وفق الأسلوب الهرمي في بناء الاختبار، وقد رفضت أي فقرة كانت نسبة الاتفاق عليها أقل من 0.80.وبهذا تم التحقق من صدق المحتوى للاختبار التشخيصي المحكي المرجع المبني وفق الأسلوب الهرمي والمكون في صورته الأولية من 52 فقرة.

صدق القرار Decision Validity

للتحقق من صدق القرار اختارت الباحثة شعبتين من الصف العاشر الأساسي ومن خارج عينة الدراسة، واحدة للذكور وأخرى للإناث عدد الطلبة فيهم 78 منهم 41 طالبة. وأخضع طلبة الشعبتين لاختبار في الوحدة قبل أن يتعلموها، ثم قامت الباحثة بتدريب معلميْ هاتين الشعبتين على طريقة التدريس لوحدة المعادلات وفق البناء الهرمي في هذه الدراسة

وجرى تدريس الطلبة الجزء الخاص بالدراسة (حل نظام مكون من معادلتين تربيعيتين بمتغيرين) وفق الاستراتيجية المستندة إلى البناء الهرمي، وبعد انتهاء التدريس ولمقارنة أداء الطلبة على الاختبار (قبل التعليم وفق الطريقة الهرمية وبعد التعليم) طبّق الاختبار على طلبة الشعبتين وحسب معامل دقة القرار عن طريق جمع النسب المئوية للطلاب المتمكنين الذين اجتازوا الاختبار والنسبة المئوية للطلاب غير المتمكنين الذين لم يجتازوه، وفق معادلة 3

معامل دقة القرار=$\frac{A+D}{N}$ (3) (علام،1995)

حيث: Aتشير إلى عدد الطلاب المتمكنين الذين اجتازوا الاختبار.

و D تشير إلى عدد الطلاب غير المتمكنين الذين اجتازوا الاختبار.

و N تشير إلى عدد الطلبة الكلي.

وكانت قيمته 0.82 وهي قيمة مقبولة.

وبهدف التحقق من أن الاختبار يميز بين المجموعات المتمايزة، وأن الفقرات حساسة للتدريس، اختبرت الفرق بين معاملات الصعوبة للفقرات لعينة الطلبة قبل التدريس بالطريقة الهرمية وبعده باستخدام اختبار (ت) للعينات غير المستقلة، أشارت النتائج إلى أن هناك فروقا ذات دلالة إحصائية عند مستوى الدلالة 0.05بين معاملات الصعوبة لجميع فقرات الاختبار قبل وبعد عملية التدريس، باستثناء الفقرة 18، حيث كانت الفروق غير دالة إحصائيا. كما أشارت النتائج أيضا إلى أن قيم مؤشر حساسية الفقرة للتدريس تراوحت بين 0.14 و0.68، وهي قيم موجبة ومرتفعة وتدل على حساسية الفقرات لعملية التدريس، مما يوفر مؤشرا من مؤشرات صدق فقرات الاختبار محكي المرجع.

التحليل العاملي

استخدام التحليل العاملي للتأكد من أحادية البعد (Hattie,1985) كإجراء للتحقق من أن الاختبار يقيس سمة واحدة؛ مما يساعد في تفسير نتائج الاختبار. حيث أشارت نتائج التحليل

العاملي الاستكشافي إلى وجود عامل واحد فقط، قيمة الجذر الكامن له 2.459 تزيد قيمته على الواحد صحيح، وفسر هذا العامل 61.470% من التباين، وهي نسبة عالية، وجميع تشبعات الفقرات على العامل تزيد على 0.3، وبهذا فإن الاختبار يقيس سمة واحدة (Hattie, 1985) .

وللإجابة عن السؤال الثالث ونصه:" ما فاعلية البناء الهرمي للاختبار؟ "

قامت الباحثة بحساب نسبة الطلبة الذين حققوا المهارة الرئيسة ممن تلقوا التعليم لهذه المهارة وذلك وفق الخطوات الآتية:

- اختيار شعبتين من شعب الصف العاشر الأساسي ومن خارج عينة الدراسة إحداهما للذكور والأخرى للإناث عدد الطلبة فيهما 73 منهم 38 طالبة.
- دربت الباحثة معلميْ هاتين الشعبتين على طريقة تدريس الطلبة لوحدة المعادلات وفق البناء الهرمي في هذه الدراسة.

بعد انتهاء التدريس طبق الاختبار على طلبة الشعبتين وحسبت نسبة الطلبة الذين حققوا المهارة الرئيسة أي الذين اجتازت علاماتهم على الاختبار علامة القطع التي حددت في الدراسة، فكانت 86%، وتعد هذه القيمة مؤشرا جيدا على فعالية البناء الهرمي(Walbesser, 1968).

وللإجابة عن السؤال الرابع من أسئلة الدراسة ونصه: "ما درجة مطابقة استجابات الطلبة على الاختبار مع الاستجابة المتوقعة من النموذج الهرمي الذي صمم للمحتوى؟

جرى التحقق من صحة الفرضيات المتعلقة بمستوى التحليل في البناء الهرمي؛ أي الفرضيات المتعلقة باعتمادية نواتج التحليل البنائي الهرمي ، أي اعتماد كل مستوى من مستويات النواتج على الآخر وذلك بالاعتماد على أسلوب والبسر(Walbesser, 1968)، ويقوم الأسلوب على التحقق من البنية الهرميةكمياً بالاستناد إلى أحكام المحكمين، وهنا يتم صنع فرضيات تقوم على أن الطالب ليتمكن من قيامه من الكفاية الرئيسة عليه أن يحقق الأهداف القبلية وهي قيامه بالأهداف المساعدة للوصول إلى إنجاز الكفاية الرئيسة بنجاح، ويتم ذلك وفق الإجراءات الآتية:

تم حساب التكرارات الآتية:

- الذين حققوا الهدف الرئيس الختامي الذي يقيسه الاختبار، وحققوا النجاح في جميع الأهداف التي تقيسها الاختبارات التي تعد متطلب سابق له تم تمثيلهم بالزوج (1،1).
- الذين حققوا الهدف الرئيس الختامي الذي يقيسه الاختبار، ولم يحققوا النجاح في أحد الأهداف على الأقل من تلك التي تقيسها الاختبارات التي تعد متطلب سابق له تم تمثيلهم بالزوج (1،0).
- الذين لم يحققوا الهدف الرئيس الختامي الذي يقيسه الاختبار، ولكن حققوا النجاح في جميع الأهداف التي تقيسها الاختبارات التي تعد متطلب سابق له تم تمثيلهم بالزوج (0،1).

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

- الذين لم يحققوا الهدف الرئيس الختامي الذي يقيسه الاختبار، ولم يحققوا النجاح في أحد الأهداف على الأقل من تلك التي تقيسها الاختبارات التي تعد متطلب سابق له تم تمثيلهم بالزوج (0،0) .

ومثال على ذلك عندما تكون المهارة الرئيسة هي حل معادلة تربيعية بمتغيرين ومعادلة تربيعية ومعادلة يظهر فيها مربع كامل على أحد المتغيرين على الأقل، فحتى يكون الطالب متقنا لهذه المهارة، عليه أن يتقن المهارات الفرعية الآتية 1) حل معادلة خطية بمتغير واحد 2) فك المربع الكامل 3) حل معادلة تربيعية بمتغير 4) حل معادلتين بمتغيرين بالحذف أو التعويض 5) كتابة الحل باستخدام الأزواج المرتبة.

فإذا كان الطالب متقنا للمهارات الفرعية الخمس و متقنا للمهارة الرئيسة يعطى الزوج (1،1) ،وإذا لم يكن متقنا للمهارة الرئيسة ولكنه متقنا للمهارات الخمس الفرعية يعطى الزوج(0، 1)، وبين شكل 2 التكوين الهرمي المعتمد على مستويين لإتقان المهارة الرئيسة.

شكل2

التكوين الهرمي المعتمد على مستويين لإتقان المهارة الرئيسة.

يبين جدول 4 النواتج الممكنة.

جدول 4

النواتج الممكنة للعلامات المتعلقة بالمهارات الرئيسة والفرعية

	اتقان المهارة					
مستوى المهارة	1	1	0	0	0	0
الرئيسي	1	1	0	0	0	0
الفرعي	11111	11110	0،1،1،1،1	11101	1،1،0،0،0	
الزوج الناتج	(1،1)	(1 ،0)	(0،0)	(0،0)	(0،0)	(0،0)

تم استخراج تكرار إجابات طلبة الصف العاشر الأساسي لعينة الدراسة الرئيسة البالغ عددها 390 طالباً وطالبة، ويوضح جدول 5 تكرارات الأزواج المرتبة.

جدول 5

تكرار الأزواج المرتبة للمستويات الهرمية الستة

مكونات الاختبار	تكرار الزوج المرتب			
	(1،1)	(0،1)	(1،0)	(0 ،0)
الأول	270	44	56	20
الثاني	277	43	39	31
الثالث	279	41	42	28
الرابع	280	44	34	32
الخامس	278	47	38	27
السادس	287	48	33	22

بعد استخراج تكرار إجابات الطلبة، ومن أجل التحقق من صدق البنية الهرمية لنواتج التحليل الهرمي قامت الباحثة بالتحقق من صحة الفرضية التي تنص على: للتحقق من الهدف الرئيس الختامي على الطالب تحقيق الأهداف المساعدة له (المتطلبات السابقة)، تم حساب نسب الاتساق ونسب كفاية الترتيب ونسبة الاكتمال باستعمال المعادلات الآتية:

- نسبة الاتساق Consistency Ratio وتعطى بمعادلة 4

$$\text{نسبة الاتساق} = \frac{\mathrm{N1,1)}}{N(O,1)+N(1,1)} \qquad (4)$$

حيث تقوم نسبة الاتساق على اختبار فيما إذا كان اتقان الطالب للمهارة الرئيسة يتضمن اتقانه جميع المهارات الفرعية التي تعد متطلب سابق لها.

- نسبة كفاية الترتيب Adequacy Ratio وتعطى بالمعادلة 5

$$\text{نسبة كفاية الترتيب} = \frac{\mathrm{N(1,1)}}{N(1,0)+N(1,1)} \qquad (5)$$

حيث تقوم نسبة كفاية الترتيب على اختبار فيما إذا كان إتقان الطالب للمهارات الفرعية أو بعضها التي تعد متطلب سابق للمهارة الرئيسة يتضمن اتقانه للمهارة الرئيسة.

- نسبة الاكتمال Completeness Ratio وتعطى بمعادلة 6

$$\text{نسبة الاكتمال} = \frac{\mathrm{N(1,1)}}{N(O,0)+N(1,1)} \qquad (6)$$

حيث تقوم نسبة الاكتمال على اختبار درجة اعتماد كل مستوى فرعي من مستويات الأهداف على الآخر.

وعرضت نتائج تحليل البناء الهرمي للاختبار (نسب الاتساق ونسب كفاية الترتيب ونسبة الاكتمال) في جدول 6.

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022

جدول 6

نتائج تحليل البناء الهرمي للاختبار

مكونات الاختبار	نسبة		
	الاتساق	كفاية الترتيب	الاكتمال
الأول	0.86	0.83	0.93
الثاني	0.87	0.88	0.90
الثالث	0.87	0.87	0.91
الرابع	0.86	0.89	0.90
الخامس	0.86	0.88	0.91
السادس	0.86	0.90	0.93

وبعد حساب النسب السابقة كانت جميع قيم نسبة الاتساق ونسبة كفاية الترتيب ونسبة الاكتمال أعلى من 0.85، وجميع هذه القيم تفيد بأنها قيم مناسبة وتحقق صدق البناء الهرمي المتعلق بحل نظام مكون من معادلتين تربيعيتين بمتغيرين لأنها أكبر من 0.85؛ حيث يفيد الأدب التربوي أن الترتيب الهرمي المتعلق بفرضية معينة يكون صادقا إذا بلغت كل من نسبة الاتساق، ونسبة كفاية الترتيب، ونسبة الاكتمال 0.85على الأقل. وينبغي أن تتحقق هذه القيمة لجميع فرضيات البناء الهرمي المتعلق بكفاية أو مهارة معينة. وعلى الرغم من أن هذه القيمة اعتبارية إلا أنه يمكن اعتمادها قيمة معقولة لأنها تفترض أن 85% من مجتمع الأفراد المختبرين يتبين لهم أن الأهداف المساعدة تعد متطلبات ضرورية وكافية لتحقيق الكفاية الرئيسة التي تعد بمثابة الهدف الختامي وبذلك يمكن اعتبار الترتيب الهرمي لهذه الأهداف صادقاً(Walbesser, 1968)، واتفق فيلبس (Phillips ,1971) مع والبسر في نسبة الاتساق، لكنه عدّل المعيار على نسبة كفاية الترتيب ليكون 0.70 ، ومعيار نسبة الاكتمال ليكون 0.50، وبهذا يمكن القول أن القيم التي جرى الحصول عليها تحقق المعايير الواردة في الأدب التربوي، وأن الاختبار يحقق خاصية البنية الهرمية.

يمكن القول أن قد تم تقديم استراتيجية لبناء اختبار تشخيصي محكي تضمن تقديم موصفات تفصيلية وشاملة للمجال السلوكي وفقرات الاختبار بدءا من تحديد الغرض من الاختبار ، ومن ثم تحديد المحتوى والفئة المستهدفة ،ثم بناء فقرات الاختبار ، والتحقق من صدق التحليل الهرمي، والتحقق من صدق البناء، وتقديم مؤشرات على صدق وثبات الاختبار، وعليه يمكن القول إن الاختبار الذي بني في هذه الدراسة يمكن أن يكون صالحا لأغراض التشخيص في حل المعادلات التربيعية.

التوصيات

توصي الباحثة باستخدام هذه الاستراتيجية في المدارس والجامعات ومراكز القياس في موضوعات ومواد أخرى؛ وذلك لقدرتها على التشخيص، كذلك تطبيق الاختبار الذي جرى بناؤه في هذه الدراسة على الفئات المستهدفة من الطلبة.

تضارب المصالح

أقر بعدم وجود تضارب في المصالح فيما يتعلق بالبحث، والملكية الفكرية، ونشر هذا البحث.

المراجع

عبد الهادي، نبيل. (2002). *القياس والتقويم التربوي واستخداماته في مجال التدريس الصفي*. دار وائل للنشر.

علام، صلاح الدين. (1995). *الاختبارات التشخيصية محكية المرجع في المجالات التربوية والنفسية والتدريبية*. دار الفكر العربي.

علي ، سعدي .(2014).*بناء اختبار للقدرات العقلية وفقاً لاستراتيجية التحليل البنائي الهرمي لدى طلبة المرحلة الإعدادية*. رسالة ماجستير غير منشورة ، جامعة بغداد، العراق

AbedAlhadi, N. (2002*).Measurement and evaluation and its utility in classroom teaching domain* (in Arabic). Dar Wa'al for Distribution.

Alderson, J. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. A&C Black.

A'llam, S. (1995). *Diagnostic criterion domain tests in educational, psychological and training domains* (in Arabic). Arabian Dar Alfeker.

Ali, S. (2014).*Constructing a test for mental abilities according to the strategy of hierarchic constructive analysis for high school stage students* (in Arabic). Unpublished thesis, University of Iraq.

Anderson, J. (1996). ACT: A simple theory of complex cognition. *American Psychologist, 51*(4), 355–365365. https://doi.org/10.1037/0003-066X.51.4.355

Anderson R. (1998). Why talk about different ways to grade? The shift from traditional assessment to alternative assessment. *New Directions for Teaching and Learning, 74*, 5-16. https://doi.org/10.1002/tl.7401

Ausubel, D. (1963). *The psychology of meaningful verbal learning*. Grune& Stratton.

Bailey, K. (1998). *Learning about language assessment: dilemmas, decisions, directions. Heinle* & Heinle Publishers.

Berk, R. (1982). *Criterion-referenced measurement: The state of the art.* The Johns Hopkins University Press.

Black, P. & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-71. https://doi.org/10.1080/0969595980050102

Bruner, J. S. (1964). The course of cognitive growth. *American Psychologist, 19*(1), 1–15. https://doi.org/10.1037/h0044160

Carry, L., Lewis, C. &Bernard, J. (1980).*Psychology and equation solving: An information processing study*. Austin: University of Texas, Department of curriculum and instruction, NSF final report, 78-22293.

Cui,Y.& Leighton.(2009).The hierarchy consistency index: Evaluating person fit for cognitive diagnostic assessment. *Journal of Education Measurement, 46*,429- 449. https://doi.org/10.1111/j.1745-3984.2009.00091.x

Davidson, F. & Lynch, B. (2002). *Test craft: A teacher's guide to writing and using language test specifications*. Yale University Press.

Dawson, M. (1998). *Understanding cognitive science*. Blackwell

De la Torre, J., & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika, 69*(3), 333-353. https://doi.org/10.1007/BF02295640

DiBello, L., Roussos, L., & Stout, W. (2007). Cognitive Diagnosis Part I. In C. R. Rao, & S. Sinharay (Eds.), *Handbook of Statistics* (Vol. 26): Psychometrics (pp. 979-1030). Elsevier.

Embretson, S. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika, 56*, 495–516. DOI: 10.1007/BF 02294487

Embretson, S. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods, 3*(3), 380- 396. https://doi.org/10.1037/1082-989X.3.3.380

Embretson, S. & Gorin, J. (2001). Improving construct validity with cognitive psychology principles. *Journal of Educational Measurement, 38*(4), 343-368. https://doi.org/10.1111/j.1745-3984.2001.tb01131.x

Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data*. The MIT Press.

Fodor, J. (1983). *The modularity of mind*. MIT Press. https://doi.org/10. .7551/mitpress/4737.001.0001

Gierl, M., & Cui, Y. (2008). Defining characteristics of diagnostic classification models and the problem of retrofitting in cognitive diagnostic assessment. *Measurement: Interdisciplinary Research & Perspective, 6*(4), 263-268. From: http://dx.doi.org/10.1080/15366360802497762

Gierl, M., & Leighton, J. (2007). Linking cognitively based models and psychometric methods. In C. R. Rao & S. Sinharay (Eds.) *Handbook of statistics: Psychometrics*, Volume 26 (pp. 1103–1106). Elsevier.

Gierl, M., Leighton, J., & Hunka, S. (2007). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications*. (Pp.242–274). Cambridge University Press.

Gierl, M., Alves, C., Roberts, M., & Gotzmann, A. (2009, April). *Using judgments from content specialists to develop cognitive models for diagnostic assessments*. In J. Gorin (Chair), How to build a cognitive model for Educational Assessments. Paper presented at the annual meeting of the National Council on Measurement in Education San Diego, CA.

Gierl, M., Zheng, Y., & Cui, Y. (2008). Using the attribute hierarchy method to identify and interpret cognitive skills that produce group differences. *Journal of Educational Measurement, 45*(1), 65 – 89.

Gorin, J. (2006). Test design with cognition in mind. *Educational Measurement: Issues and Practice, 25*(4), 21-35. https://psycnet.apa.org/doi/10.1111/j.1745-3992.2006.00076.x

Gregory, R. (2004). *Psychological testing: History, principles, and applications*. Allyn & Bacon.

Hartz, S. (2002). A bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality. Unpublished doctoral dissertation, University of Illinois, Champaign. Available from: https://www.researchgate.net/publication/

Hattie, J. (1985) Methodology Review: Assessing unidimensionality of tests and items. Applied Psychological Measurement, 9,139-164 from: https://hdl.handle.net/11299/102073

Huff, K., & Goodman, D. (2007). *The demand for cognitive diagnostic assessment*. In J. P. Leighton& M. J. Gierl (Eds.), Cognitive diagnostic assessment for education: Theory and applications (pp. 19-60). Cambridge University Press.

Ketterlin-Geller, L., &Yovanoff, P. (2009). Cognitive diagnostic assessment in mathematics to support instructional decision making. *Practical Assessment, Research, & Evaluation, 14* (16), 1-11.

Kieran, C. (1992). *The learning and teaching of school algebra*. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 390–419). Macmillan.

Kunnan, A. & Jang, E. (2011). *Diagnostic feedback in language assessment*. In M. Long & C. Doughty (Eds.), the handbook of language teaching (pp. 610-625). Wiley-Blackwell.

Leighton, J. (2004). Avoiding misconceptions, misuse, and missed opportunities: The collection of verbal reports in educational achievement testing. *Educational Measurement: Issues and Practice, 23*, 6–15. https://psycnet.apa.org/doi/10.1111/j.1745-3992.2004.tb00164.x

Leighton, J., & Gierl, M. (2007a). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice, 26*, 3–16. https://doi.org/10.1111/j.1745-3992.2007.00090.x

Leighton, J., &Gierl, M. (2007b). *Verbal reports as data for cognitive diagnostic assessment*. In J. P. Leighton & M. J. Gierl (Eds.), Cognitive diagnostic assessment for education: Theory and applications. (pp. 146– 172). Cambridge University Press.

Leighton, J., & Gierl, M. (2011). *The learning sciences in educational assessment: The role of cognitive models*. Cambridge University Press.

Leighton, J., Gierl, M., &Hunka, S. (2004). The attribute hierarchy method for cognitive assessment: A variation Tatsuoka's Rule-Space Approach. *Journal of Educational Measurement, 41*(3), 205–237. http://www.jstor.org/stable/1435314

Li, H., Suen, H. (2013). Constructing and validating a Q-matrix for cognitive diagnostic analyses of a reading test. *Educational Assessment, 18*(1), 1–25. https://doi.org/10.1080/10627197.2013.761522

Lohman, D. (2000). *Complex information processing and intelligence*. In R.J. Sternberg (Ed.), Handbook of intelligence (pp. 285–340). Cambridge University Press.

Lord, F. (1952). *A theory of test scores (Psychometric Monograph No. 7).Richmond, VA: Psychometric Corporation*. Retrieved from: http://www.psychometrika.org/journal/online/MN07.pdf

Messick, S. (1989). *Validity*. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13–103). New York: American Council on Education, Macmillan.

Mislevy, R. (1995). *Probability-based inference in cognitive diagnosis*. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), cognitively diagnostic assessment (p. 43–72). Hillsdale, Lawrence Erlbaum Associates.

Mislevy, R., Steinberg, L., & Almond, R. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3–62. http://dx.doi.org/10.1207/S15366359MEA0101_02

Napaphun, V. (2012). Relational thinking: Learning arithmetic in order to promote algebraic thinking. *Journal of Science and Mathematics Education in Southeast Asia, 35*(2), 84-101.

Nichols, P. (1994). A framework for developing cognitively diagnostic assessments. *Review of Educational Research, 64*(4), 575–603. https://doi.org/10.2307/1170588

Nichols, P. & Joldersma, K. (2008). Review of cognitive diagnostic assessment for education: Theory and applications. *Journal of Educational Measurement, 45*(4), 407–411

Nichols, P., Chipman, S., & Brennan, R. (Eds.). (1995). *cognitively diagnostic assessment*. Erlbaum.

Pellegrino, J., Chudosky, N., & Glaser, R. (Eds.) (2001). *Knowing what students know: The Science and Design of Educational Assessment*. National Academy Press

Phillips, E. (1971). *Validating learning hierarchies for sequencing mathematical tasks*. Doctoral dissertation, Purdue University Ann Arbor, Michigan, University Microfilms, No. 72-8008

Royer, J., Cisero, C., & Carlo, M. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research, 63*, 201-243. https://psycnet.apa.org/doi/10.2307/1170473

Rovinelli, R., & Hambleton, R. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal for Educational Research, 2*, 49-60.

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144.

Snow, R. &Lohman, D. (1989). *Implications of cognitive psychology for educational measurement*. In R. L. Linn (Ed.), Educational measurement (3rd Edition), (pp. 263–330). Macmillan.

Snow, R. & Mandinach, E. (1991). *Integrating instruction and assessment: A research and development agenda* (RR91-8). Educational Testing Service.

Struyven, K., Dochy, F., & Janssens, S. (2003). *Students 'perceptions about new modes of assessment: a review*. In M. Segers, F. Dochy, & E. Cascallar (Eds), Optimizing new modes of assessment: In search of qualities and standards. Kluwer.

Struyven, K., Dochy, F., &Janssens, S. (2008). The effects of hands-on experience on students' preferences for assessment methods. *Journal of Teacher Education, 59*(1), 69-88.

Tatsuoka, K. (1983). Rule-space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20, 34-38.

Tatsuoka, K., & Tatsuoka, M. (1997). Computerized adaptive diagnostic testing: Effect on remedial instruction as empirical validation. *Journal of Educational Measurement, 34*, 3-20.

Thompson, B., & Levitov, J. E. (1985). Using microcomputers to score and evaluate test items. *Collegiate Microcomputer, 3*, 163-168.

Wang, M. (1973). Psychometric studies in the validation of an early learning curriculum. *Child Development, 44*, 54-60.

Wang, C., Gierl, M. (2011). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills in critical reading. *Journal of Educational Measurement, 48*(2), 165–187.

Zhou, J. (2010). *Estimating attribute-based reliability in cognitive diagnostic assessment*. Unpublished Doctoral Dissertation. University of Alberta: Edmonton, Alberta, Canada.

Walbesser, H. (1968). *An evaluation model and its application*. American Association for the Advancement of Science.

المجلة الدولية للأبحاث التربوية - جامعة الإمارات العربية المتحدة
المجلد (46) العدد (1) يناير 2022