Sana'a University Journal of Human Sciences
مجلة جامعة صنعاء للعلوم الإنسانية
Vol. 5 | No. 3 | Page 111 – 138 |2026 |
https://journals.su.edu.ye/index.php/jhs
ISSN : 2958-8677

Construction and Calibration of an Item Bank in the Educational Measurement and Evaluation Course Based on the One-Parameter Item Response Model for Students of the Faculty of Education – Sana'a University

Fatehi Ahmed Nasser Ibrahim Adhubaibi[1,*], Nagah Abdelrahim Mohamed Othman [1]

[1] Department of Measurement and Evaluation ,Faculty of Education - Sana'a University, Sana'a, Yemen.

***Corresponding author:** *fathiahmed1982@gmail.com*

Keywords

1. Item bank
2. Item response models
3. Measurement and Evaluation
4. Rasch model
5. Test equating
6. Calibration of Item Bank

Abstract:

The study aimed to construct and calibrate an item bank for the Educational Measurement and Evaluation course based on the one-parameter item response theory model (1PL). The study sample consisted of 119 third-year students from the Faculty of Education – Sana'a University, representing 37% of the total study population. The research instrument was an item bank covering the first three units of the textbook assigned to third-year students. Scientific procedures were followed in building the item bank, which comprised 112 multiple-choice items distributed across three test forms (A, B, and C). Each test form included 44 items, with 10 common anchor items shared among the three forms. The Bilog-MG3 software was used to verify the fit of the item bank data to the one-parameter model, calibrate the items based on their difficulty levels, and estimate the parameters of item difficulty, individual abilities, test information function, empirical reliability coefficients, and the reliability index. Additionally, SPSS software was used to test the assumption of unidimensionality. Among the most prominent findings: after verifying the assumptions of unidimensionality, local independence, and speededness, it was found that no items were excluded by the Bilog-MG3 program under the one-parameter model. However, based on the chi-square goodness-of-fit test, 16 items were excluded for not fitting the model. No participants were excluded from the three test forms. The item bank was successfully calibrated according to the one-parameter model, with a mean difficulty parameter of 0.65, a standard deviation of 0.907, and item difficulties ranging from -1.93 to 2.27, corresponding to ability estimates ranging from -2.20 to 3.37.

Sana'a University Journal of Human Sciences
مجلة جامعة صنعاء للعلوم الإنسانية
Vol. 5 | No. 3 | Page 112 – 138 | 2026
https://journals.su.edu.ye/index.php/jhs
ISSN : 2958-8677

بناء وتدريج بنك أسئلة في مقرر القياس والتقويم التربوي وفق نموذج الاستجابة للفقرة أحادي المعلم لطلبة كلية التربية– جامعة صنعاء

فتحي أحمد ناصر إبراهيم الضبيبي[1]*، نجاح عبد الرحيم محمد عثمان [1]
[1] قسم القياس والتقويم ، كلية التربية - جامعة صنعاء ، صنعاء ، اليمن.
*المؤلف: fathiahmed1982@gmail.com

الكلمات المفتاحية

1. بنك الأسئلة
2. نماذج استجابة الأسئلة
3. القياس والتقييم
4. نموذج راش
5. معادلة الاختبار
6. معايرة بنك الأسئلة

الملخص:

هدفت الدراسة إلى بناء وتدريج بنك أسئلة لمقرر القياس والتقويم التربوي وفق النموذج أحادي المعلم لنظرية الاستجابة للفقرة، وتكونت عينة الدراسة من (119) طالبًا وطالبة من طلبة المستوى الثالث بكلية التربية – جامعة صنعاء ويمثلون (37٪) من مجتمع الدراسة، وتكونت أداة الدراسة من بنك أسئلة في الثلاث الوحدات الأولى من الكتاب المقرر على طلبة المستوى الثالث بكلية التربية– جامعة صنعاء، حيث اتبعت الخطوات العلمية في إجراءات بناء بنوك الأسئلة، وتكون البنك من(112) فقرة من نوع الاختيار من متعدد، وزعت على ثلاث صور اختبارية (C,B,A) كل صورة تحتوي على (44) فقرة منها (10) فقرات مشتركة رابطة بين الثلاث الصور الاختبارية، واستخدمت الدراسة برنامج الحاسب الآلي Bilog mg3 للتحقق من ملاءمة بيانات بنك الأسئلة للنموذج أحادي المعلم وتدريج فقرات بنك الأسئلة وفقًا لصعوبتها، وتقدير معالم كل من: صعوبة الفقرات وقدرات الأفراد ودالة معلومات الاختبار، ومعاملات الثبات الإمبريقي ومؤشر الثبات، كما استخدمت برنامج الحاسب الآلي SPSS للتحقق من افتراضات أحادية البعد. ومن أبرز النتائج التي توصلت إليها الدراسة بعد التحقق من افتراضات نظرية الاستجابة للفقرة أحادية البعد والاستقلال المركزي والتحرر من عامل السرعة، تبين أن البرنامج المستخدم لم يحذف أي فقرة من النموذج أحادي المعلم، وفي الملاءمة الإحصائية لمربع كاي تم حذف (16) فقرة لم تلاءم النموذج أحادي المعلم، ولم يحذف أي فرد في الصور الاختبارية الثلاث، وتم تدريج بنك الأسئلة لمقرر القياس والتقويم التربوي بناءً على النموذج الأحادي المعلم، حيث بلغ بمتوسط حسابي لمعلم الصعوبة بعد تدريجه (0.65) وبانحراف معياري (0.907)، وتراوحت بين (−1.93−2.27) وتقابلان القدرة (3.37− 2.20−).

المقدمة:

إن تطوير منظومة القياس والتقويم بمختلف ميادينه يضمن قياس مخرجات التعلم المستهدفة وفقاً لمعايير عالمية معترف بها، ومما لا شك فيه أن تطوير نظم التقويم الجامعي يعكس مدى اهتمام الجامعة بتعديل المسار الجامعي نحو المعايير عالية الجودة والكفاءة، حيث إن تقويم مخرجات التعلم الجامعي يعتمد على الاختبارات بأشكالها وأنواعها كأدوات للقياس.

وتعد الاختبارات التحصيلية، والنفسية من أهم الوسائل الرئيسة لعملية القياس والتقويم، والتي تزود المؤسسات التعليمية، والتربوية بالمعلومات اللازمة عن مدى تحقق الأهداف التعليمية، بالإضافة إلى تقديم معلومات يمكن من خلالها اتخاذ قرارات مصيرية للعديد ممن يخضعون لهذه الاختبارات، سواء بالانتقاء أو التصنيف أو الإجازة أو التحويل إلى برامج علاجية أو تأهيلية، وترتبط دقة وسلامة هذه القرارات بمدى جودة وكفاءة هذه الاختبارات، وإذا كانت هذه الاختبارات لا تتميز بالدقة والموضوعية، والكفاءة المطلوبة فإن ذلك يؤدي إلى صدور قرارات غير صائبة؛ مما يترتب عليها إهدار العديد من القدرات وضياع حقوق العديد من الطلاب، الذين إن تيسرت لهم الوسائل والظروف الملائمة لأثبتوا قدراتهم وكفاءتهم في العديد من المجالات. (سيف، 2013)

وعلى الرغم من اعتماد مختلف مؤسسات البحث العلمي والمؤسسات التعليمية والأكاديمية المتخصصة على أساليب القياس التقليدي بصور واسعة الانتشار منذ فترة طويلة وحتى الآن (Embereton & Reise,2010) إلا أن هذه الأساليب عانت من بعض نواحي الضعف والقصور، ووجه إليها العديد من الانتقادات، والتي من أبرزها انعدام خطية القياس، وعدم وجود وحدة ثابتة للقياس بالإضافة إلى القياس في أكثر من بُعد، وأن معالم الفقرات والأفراد محكومة بعينة المفحوصين التي تطبق عليهم الاختبار، فعندما تكون عينة المفحوصين مرتفعة في مستوى القدرة نحصل على صعوبة منخفضة للفقرات، وإذا كانت عينة المفحوصين منخفضة في مستوى القدرة نحصل على صعوبة مرتفعة للفقرات، وبالتالي لا يمكن التنبؤ بأداء المفحوصين على فقرة اختبارية معينة.(Hambleton & Jones,1993 Ojerinde,2013)

وقد بذلت كثير من الجهود بهدف حل هذه المشكلات والتخلص منها (Reise & Waller ,2003)، وأسفرت هذه الجهود في منتصف القرن العشرين تطورات جوهرية في منهجيات القياس النفسي وطرق تصميم الاختبارات والمقاييس وتقنية تحليل البيانات المستمدة منها، وقدمت إطاراً مرجعياً لبناء المقاييس النفسية والتربوية، وطريقة تفسير الدرجات على هذه الاختبارات مقارنة بما قدمته النظرية الكلاسيكية في القياس(Mislevy & Bock ,1990; Van der Linden,2010) من خلال ظهور بعض الاتجاهات الحديثة للقياس التي حققت تقدماً حقيقياً واضحاً وملموساً في تحقيق الموضوعية المطلوبة في هذا الصدد، ومن أبرز هذه الاتجاهات نظرية الاستجابة للفقرةItem Response Theory. (David & Seung ,2010)، التي اعتبرت بمثابة

الثروة والمستقبل الزاهر للقياس النفسي والتربوي. (Anstas & Urbena ,2005)

وتُعد استقلالية القياس أحد أهم المزايا المرتبطة بنظرية الاستجابة للفقرة، وتعني أن تقدير معالم الفقرات (الصعوبة) يكون مستقلاً عن خصائص الأفراد الذين استخدموا في تقدير هذه المعالم (Sample Free)، وإن تقدير قدرة الفرد يكون مستقلاً عن عينة الفقرات التي تطبق عليه (Item Free)، ولتحقيق استقلالية القياس اهتم البحث السيكومتري بنظرية الاستجابة للفقرة بتطوير نماذج احتمالية تقوم على تحديد العلاقة بين أداء الفرد على فقرات الاختبار وبين السمات أو القدرات الكامنة التي تقف وراء هذا الأداء، وبتحقق هذه الخاصية أصبح بالإمكان القيام ببعض التطبيقات، مثل: معادلة درجة الاختبار، وبناء بنوك الأسئلة، والقياس التكيفي. (Reckase,2009)

مشكلة الدراسة:

تعد بنوك الأسئلة ركيزة أساسية في تطوير وتجويد عملية التعليم والتقويم في التعليم العالي، إذ توفر مخزون غني من الأسئلة القابلة للاستخدام المتكرر لتقييم التحصيل وتحديد المستويات المعرفية.

وعلى الرغم من الجهود المبذولة في تطوير بنوك الأسئلة على المستوى العالمي والعربي والمحلي، إلا أنه لا يتوفر بنك أسئلة في مقرر القياس والتقويم التربوي بكلية التربية في جامعة صنعاء ولا في غيره من مقررات الكلية. إذ يعد مقرر القياس والتقويم التربوي أحد المقررات المهمة التي تزود طلبة كلية التربية بالأدوات التربوية والمعرفية والمهارات اللازمة لتقويم تعلم المتعلمين بكفاءة عالية وتحقيق الأهداف التربوية المنشودة. فضلاً عن النقص الواضح في البيئة اليمنية في بنوك الأسئلة المبنية باستخدام نماذج نظرية الاستجابة للفقرة التي تلافي أوجه القصور في بنوك الأسئلة المبنية باستخدام نظرية القياس التقليدية.

وبناءً على ما تقدم تتحدد مشكلة البحث في: الحاجة إلى بناء بنك أسئلة في مقرر القياس والتقويم التربوي لطلبة كلية التربية ـ جامعة صنعاء وفق نظرية الاستجابة للفقرة.

أهداف الدراسة:

تهدف الدراسة إلى بناء وتدريج بنك أسئلة في مقرر القياس والتقويم التربوي وفق نموذج الاستجابة للفقرة أحادي المعلم، ويتفرع من هذا الهدف الرئيس الأهداف الفرعية الآتية:

1. التأكد من ملائمة بيانات الدراسة (الفقرات، الأفراد) لنموذج الاستجابة للفقرة أحادي المعلم.
2. التعرف على تقدير معالم القدرة للأفراد والفقرات ودالة المعلومات للصور الاختبارية لبنك الأسئلة وفق نموذج الاستجابة للفقرة أحادي المعلم.
3. تدريج فقرات بنك الأسئلة في مقرر القياس والتقويم التربوي وفق نموذج الاستجابة للفقرة أحادي المعلم.

أهمية الدراسة:

تتمثل أهمية الدراسة في كونها تسهم في نشر الوعي بمفاهيم بنوك الأسئلة النظرية والتطبيقية، وفي توفيرها بنك أسئلة في مقرر القياس والتقويم التربوي باستخدام نموذج الاستجابة للفقرة أحادي المعلم يتميز بالدقة والموضوعية ويتمتع بخصائص سيكومترية جيدة، يمكن أن يستفيد منه أعضاء هيئة التدريس والطلبة وكليات التربية. كما سيسهم هذا البنك في:

- توفير اختبارات تحصيلية دقيقة وموثوقة لتقييم تحصيل الطلبة في مقرر القياس والتقويم التربوي.
- تمكين أساتذة مقرر القياس والتقويم التربوي من بناء اختبارات متكافئة بسهولة كبيرة.
- تشجيع الباحثين والمهتمين بالقياس والتقويم على استخدام نظرية الاستجابة للفقرة في أبحاثهم، وفي بناء أدواتهم وتحليل بياناتها.
- تطوير نظم تقويم أكثر فاعلية وكفاءة في التعليم العالي بعدة نواة لتطوير بنوك الأسئلة في مقررات التعليم العالي باستخدام نظرية الاستجابة للفقرة.

حدود الدراسة ومحدداتها:

تحددت الدراسة باستخدام نموذج الاستجابة للفقرة أحادي المعلم في بناء وتدريج بنك أسئلة في مقرر القياس والتقويم التربوي (الوحدات الثلاث الأولى) لطلبة كلية التربية بجامعة صنعاء.

التعريفات الاصطلاحية والإجرائية:

تناولت الدراسة الحالية المصطلحات الآتية:

بنك الأسئلة:

عرف علام(2002) بنك الأسئلة بأنه عبارة عن مجموعة من فقرات الاختبار يكون لها خصائص سيكومترية مميزة ومعلومة، حيث تودع هذه الفقرات في البنك بطريقة تيسر على الفاحصين سحب مجموعة منها لتستخدم في بناء اختبارات تناسب أغراضاً مختلفة، حيث يبدأ بنك الأسئلة بتحليل عدد كبير من الفقرات على أعداد كبيرة من الأفراد بهدف استبعاد الأفراد غير الملائمين وكذلك الفقرات غير الملائمة والإبقاء على ما هو ملائم، ثم دمج هذه الفقرات في تدريج واحد ينتهي بشبكة من عدة اختبارات (تشكل بنكاً للأسئلة) يمكن أن تغطى المدى الواسع من المتغير موضوع القياس (علام، 2002)

وعرفه الأحول (2021) بأنه مجموعة من الفقرات أعدت في ضوء معايير ومنهجية علمية محددة يكون لها خصائص سيكومترية مميزة ومقبولة وفقاً لأسس القياس الموضوعي لنظرية الاستجابة للفقرة تُودع في برامج الحاسوب بطريقة يسهل سحب مجموعات منها لتكوين صور اختبارية تناسب الأغراض المختلفة للقياس والتقويم التربوي (الأحول، 2021).

ويعرف بنك الأسئلة إجرائيًا في الدراسة الحالية بأنه: مجموعة الفقرات الاختبارية التي أعدت على وفق الخطوات المتعارف عليها في بناء بنوك الأسئلة لها خصائص سيكومترية مميزة ومقبولة وفقاً لأسس القياس الموضوعي لنموذج الاستجابة للفقرة أحادي المعلم، تقيس مستوى تحصيل الطلبة في الوحدات الثلاث الأولى لمقرر القياس والتقويم التربوي بكلية التربية- جامعة صنعاء. ويقاس مستوى التحصيل بالدرجة الكلية التي يحصل عليها المستجيب في أي صور اختبارية مسحوبة من بنك الأسئلة المعد في هذه الدراسة.

نموذج الاستجابة للفقرة أحادي المعلم:

اعتمدت الدراسة تعريف حبيش وآخرين (2018) أن نموذج الاستجابة للفقرة أحادي المعلم يفترض أن جميع الفقرات لها القدرة التمييزية نفسها للتمييز بين الأفراد، أي تكون جميع المنحنيات المميزة لفقرات الاختبار متوازية (متساوية في الميل) وتختلف عن بعضها البعض في نقاط التقائها بالمحور الأفقي الذي يمثل متصل القدرة أو السمة الكامنة، كما يفترض عدم لجوء الأفراد للتخمين العشوائي عند

إجابتهم عن فقرات الاختبار، وهذا يعني أن الفقرات تختلف في صعوبتها فقط. (حبيش وآخرون،2018)

معالم قدرة الأفراد:

عرف هامبلتون وآخرون (1991) معلم القدرة بأنه القيمة التي يتم استخراجها عند مقدار التفاوت بين درجة المستجيب الحقيقية ودرجته الملاحظة، ويستخدم الرمز(θ) للدلالة على قدرة المفحوص أو مقدار السمة التي يمتلكها، مع وجود مقدار من الاحتمالية(probability) للاستجابة الصحيحة عند كل مستوى قدرة يرمز لها بالرمز (θ) P، وتشير هذه القدرة إلى تعبير اصطلاحي قد يعرف بالتحصيل أو الاستعداد أو متغيرات الشخصية.

معالم الفقرات:

هي عبارة عن مجموعة القيم المستخرجة عند تطبيق نموذج الاستجابة للفقرة أحادي المعلم، وتتمثل في **معلم الصعوبة:** وهو عبارة عن نقطة على متصل القدرة تقابل احتمال (50%) للإجابة على الفقرة إجابة صحيحة. (Hambleton & Swaminathan & Rogers,1991)

دالة معلومات الاختبار:

عرف علام (2005) دالة معلومات الاختبار بأنها عبارة عن علاقة منحنية بين متغيرين هما مستويات القدرة التي يمثلها المحور الأفقي، والمعلومات المقدمة من الاختبار ككل والتي يمثلها المحور الرأسي، وهي تعبر عن كمية المعلومات المقدمة من المجموع الكلي لفقرات الاختبار (علام، 2005: 117).

وعرفها نور الدين (2002) بأنها مقدار ما تعطيه الفقرة أو الاختبار من معلومات عند المستويات المختلفة لمتصل القدرة، ويتم حسابها عن طريق مقلوب تباين الخطأ المعياري لتقدير القدرة (نور الدين،2002)

تدريج الفقرات:

عرف رايتو واستون (1979) في الشافعي (2014) تدريج الفقرات بأنه: انتظام مواقع الفقرات على متصل صعوبتها حول صفر تدريج متغير مستوى الصعوبة بحيث يعبر موقع كل فقرة (بُعدها عن نقطة صفر التدريج) عن مستوى صعوبتها (الشافعي، 2014).

ويعرف تدريج الفقرات إجرائيًا في الدراسة الحالية بأنه: سلسة التحليلات الإحصائية التي تجرى وفق نموذج الاستجابة أحادي المعلم باستخدام برنامج Bilog mg3على بيانات فقرات بنك الأسئلة الوحدات الثلاث الأولى لمقرر القياس والتقويم التربوي لطلبة كلية التربية بجامعة صنعاء لترتيب الفقرات على متصل صعوبتها حول صفر تدريج مشترك بحسب مستوى صعوبتها.

مقرر القياس والتقويم التربوي:

أحد مقررات إعداد المعلمين في كلية التربية بجامعة صنعاء يُدَّرس في المستوى الثالث، ويتضمن المقرر المعتمد في الكلية ثمانية فصول دراسية، هي: الفصل الأول (التعريف بالقياس والتقويم)، والفصل الثاني (المفاهيم الإحصائية الأساسية)، والفصل الثالث (الاختبارات التحصيلية "الجزء الأول")، الفصل الرابع (الاختبارات التحصيلية "الجزء الثاني)، الفصل الخامس (المعايير ووحدات القياس)، والفصل السادس (خصائص الاختبار الجيد)، والفصل السابع (اختبارات الذكاء والقدرات الخاصة)، الفصل الثامن (قياس الشخصية).

الإطار النظري:

نماذج نظرية الاستجابة للفقرة:

وهذه النماذج تتضمن مجموعة من الدوال الرياضية، تتصف بأنها احتمالية وليست حتمية، تهدف لتحديد علاقة رياضية بين أداء الفرد على الاختبار، وبين السمات أو القدرات التي تكمن وراء هذا الأداء وتفسره، أي أن العلاقة التي تحددها الدالة في ضوء هذه النماذج تسير وفق نظرية الاحتمالات، وتمثل هذه الدالة بعائلة من منحنيات الترجيح اللوغاريتمي Logistic Curves.

وأشهر هذه النماذج هي النماذج ثنائية التدريج التي تستخدم في تحليل الفقرات التي يجيب الفرد عنها إجابة صحيحة أو إجابة خاطئة، أو بالاختيار بين (نعم ، لا) وتكون الدرجة التي يحصل عليها الفرد إما (1) أو (صفر)، ويعتمد تحديد النموذج المناسب لتدريج فقرات المقياس أحادي البعد ثنائي الاستجابة على عدد معالم الفقرة التي تؤثر على إجابة الفرد عن الفقرة، فقد يتم تدريج الفقرات في ضوء مستوى صعوبتها (B) كما في النموذج اللوغاريتمي الأحادي المعلم، وقد يتم تدريج الفقرات في ضوء مستوى صعوبتها وقدرتها على التمييز (A) بين المستويات المختلفة من القدرة كما في النموذج اللوغاريتمي ثنائي المعلم، وقد يتم هذا التدريج في ضوء مستوى صعوبة الفقرة وقدرتها على التمييز ومستوى التخمين (C) كما في النموذج الثلاثي المعلم (علام ، 2005)

وفيما يلي شرح موجز للنموذج اللوغاريتمي أحادي المعلم المعتمد في الدراسة الحالية:

اقترح هذا النموذج العالم الدنماركي جورج راش (1960)، ويعد هذا النموذج أبسط نماذج الاستجابة للفقرة، ويهتم هذا النموذج بتحديد موقع الفقرة الاختبارية على ميزان صعوبة الفقرات التي تشكل الاختبار جميعاً، كما يهتم بتدريج مستويات قدرة الفرد باختبار معين على ميزان تعيير الفقرات نفسها (Thorindike,1991)، ويفترض أن جميع الفقرات لها القدرة التمييزية نفسها للتمييز بين الأفراد(العبدالله،2014،ص48) أي تكون جميع المنحنيات المميزة لفقرات الاختبار متوازية (متساوية في الميل) وتختلف عن بعضها البعض في نقاط التقائها بالمحور الأفقي الذي يمثل متصل القدرة أو السمة الكامنة، كما يفترض عدم لجوء الأفراد للتخمين العشوائي عند إجابتهم عن فقرات الاختبار، وهذا يعني أن الفقرات تختلف في صعوبتها (حبيش وآخرون،2018)، ويمكن الحصول على قدرة الفرد بمعرفة صعوبة الفقرة، واحتمال الإجابة الصحيحة عليها، وكذلك الحصول على صعوبة الفقرة من خلال معرفة قدرة الفرد، واحتمال الإجابة الصحيحة على الفقرة، بالإضافة إلى أن كلاً من صعوبة الفقرة ومستوى القدرة مستقلان، وهذا يسمح بتقدير متسق لكلٍ منهما؛ أي أن تقدير القدرة مستقل عن الفقرة المستخدمة في التقدير، وكذلك يمكن تقدير صعوبة الفقرة بغض النظر عن قدرة أفراد العينة المستخدمة، ويمكن الحصول على قدرة الفرد بمعرفة صعوبة الفقرة، واحتمال الإجابة الصحيحة عليها، وكذلك الحصول على صعوبة الفقرة من خلال معرفة قدرة الفرد، واحتمال الإجابة الصحيحة على الفقرة، بالإضافة إلى أن كلاً من صعوبة الفقرة ومستوى القدرة مستقلان، وهذا يسمح بتقدير مستقل لكلٍ منهما؛ أي أن تقدير القدرة مستقل عن الفقرة المستخدمة في التقدير، وكذلك يمكن تقدير صعوبة الفقرة بغض النظر عن قدرة أفراد العينة المستخدمة، وهذا الاستقلال في التقدير لهذه المعالم هو جوهر وأهم خاصية

في نظرية الاستجابة للفقرة، وقد سماها راش بالموضوعية المحددة(Objectivity Specific).(علام،2005)

افتراضات نظرية الاستجابة للفقرة:

1. **أحادية البعد:** والتي تعني أن أداء الفرد على الاختبار يمكن أن يعزى إلى سمة واحدة، وهذا الافتراض يصعب تحقيقه في كثير من الأحيان؛ لأن هناك عوامل أخرى ربما تؤثر في الأداء إضافة إلى قدرة الفرد، كالعوامل الشخصية أو المعرفية أو ظروف تطبيق الاختبار وتسمى الاختبارات التي تقوم على قياس سمة واحدة للأفراد اختبارات أحادية البعد.

ويشار إلى أن هذا الافتراض لا يتحقق على نحوٍ تام، حيث توجد عوامل كثيرة قد تؤثر على أداء الفرد في الاختبار مثل الدوافع، وقلق الامتحان، والقدرة على العمل بسرعة، والميل إلى تخمين الإجابة الصحيحة عندما يكون ذلك ممكناً، فجميع هذه العوامل وغيرها قد تؤثر في أداء الأفراد في الاختبار، بالإضافة إلى القدرة التي يسعى الاختبار لقياسها، ولكن بسبب وجود القدرة المسيطرة لدى جميع الأفراد، والتي يعزى إليها الأداء بشكل عام وقلة تأثير العوامل الأخرى وعدم ثبات تأثيرها، يرد القياس إلى هذه القدرة، وتعد هذه القدرة هي البعد الأساسي الأحادي وتعد العوامل الأخرى أخطاء في تحليل نماذج الاستجابة للفقرة. (Hambilton& De Gruaijtor,1983).

وإن التحقق من شرط أحادية البعد للبيانات المستمدة من الفقرات الاختبارية، يكون باستخدام برامج الحاسوب المتعلقة بتحليل المركبات الأساسية لمصفوفة البيانات الاختبارية، أو من خلال نتائج التحليل العاملي، والتي تعد فرضاً أساسياً في نماذج الاستجابة للفقرة أحادية البعد، ويجب أن يطبق الاختبار على عينة مماثلة في قدرتها التي يقيسها الاختبار، للعينة المقصودة في الدراسة (طيفور، 2007).

2. **الاستقلال الموضعي:** يتطلب هذا الافتراض أن يكون احتمال الإجابة الصحيحة أو (الخاطئة) للفرد على فقرة ما مستقلاً من أي فقرة أخرى عند مستوى معين للقدرة، والعوامل المؤثرة في احتمال استجابة الفرد على فقرة معينة هي فقط السمة الكامنة التي تقيسها الفقرة، ومعالم الفقرة التي يجيب عنها الفرد. (Finch, 2002)

وممكن القول إن هذا الفرض يجعل لنمط درجة الفرد الكلية في الاختبار أهمية كبيرة بحيث يمكن استخدامها في التقدير الإحصائي للسمة أو للقدرة تقديراً يتسم بالكفاءة، وعندئذٍ يمكن اعتبار أن فقرات الاختبار تقيس سمة أحادية البعد. (طيفور، 2007). ويمكننا القول إن فرضية الاستقلال الموضعي تتحقق ضمناً عند تحقق فرضية أحادية البعد. (العبد الله،2012)

فإذا كانت لدينا مجموعة فقرات تقيس قدرة واحدة (أحادية البعد) فعند مستوى محدد من هذه القدرة للأفراد، تكون استجاباتهم لهذه الفقرات مستقلة إحصائياً فإن درجات بعض الأفراد على هذه الفقرات ستكون أعلى من درجات أفراد آخرين وبما أن مستوى القدرة ثابت فإن هذا يدل على أن إجابات الأفراد ذوي الدرجة المرتفعة متأثراً بعامل آخر غير القدرة المقاسة، وهذا يعني أن الفقرات

تقيس قدرة أخرى وهذا تناقض مع شرط أحادية البعد الذي اعتبرناه محققاً في الأصل. (Bishop & Omar,2002).

وتحقيقاً لفرض الاستقلال الموضعي فإن ذلك يكون من خلال ضرورة اهتمام معد الاختبار عند بناء فقراته ألا تكون هذه الفقرات معتمدة بعضها على بعضها الآخر، وألا تقدم أي فقرة مفتاحاً يساعد في حل فقرة أخرى في الاختبار أو فهمها، فعندما يعتنى ببناء الفقرات الاختبارية بحيث لا تعتمد أي منها على الآخر، تصبح الفقرات مستقلة، وبالتالي يتحقق شرط الاستقلال الموضعي. (طيفور،2007)

3. **منحنى خاصية الفقرة:** وهو من الافتراضات الأساسية التي تقوم عليها نظرية الاستجابة للفقرة، التي تتناول سمة أحادية البعد، فمنحنى خاصية الفقرة هو دالة رياضية تربط بين احتمال نجاح الفرد في إجابة الفقرة، وبين السمة أو القدرة التي تقيسها الإجابة لمجموعة فقرات أو يقسها اختبار يحوي هذه الفقرات أي هو دالة غير خطية(Nonliner) لانحدار احتمال الإجابة الصحيحة عن الفقرة على مستوى السمة الكامنة أو القدرة التي يقيسها الاختبار، وتأخذ شكل منحنى الترجيح اللوغاريتمي الاحتمالي Logistic Curve، والاختلاف الرئيس بين نماذج نظرية الاستجابة للفقرة يعتمد على الصيغة الرياضية لمنحنى خاصية الفقرة. (Hambleton,1982) (Klein,1993) أي أن منحنى خاصية الفقرة أو المميزة للفقرة هو الخط البياني لاحتمال الإجابة الصحيحة لهذه الفقرة وذلك عند مستويات القدرة المقاسة. (Cohen &Swerdlik ,2005)

وهناك إجماع بأن منحنى خاصية الفقرة يمكن تحديده من خلال أربعة متغيرات هي: قدرة الفرد ومعلم الصعوبة والتي تقابل بنقطة على متصل (محور) السمة الكامنة أو القدرة يمر فيها منحنى خاصية الفقرة عند القيمة الاحتمالية (0.5) وبالتالي صعوبة الفقرة تقابل نقطة من متصل السمة أو القدرة يكون فيها احتمال الإجابة الصحيحة يساوي احتمال الإجابة الخاطئة ويساوي (0,5) (Klein,1993).

4. **التحرر من عامل السرعة:** تفترض نظرية الاستجابة للفقرة بأن لا يكون شرط السرعة عاملاً يؤثر في تطبيق الاختبار، أي أن الفرد لا يكون تحت تأثير عامل الوقت أثناء استجابته عن فقرات الاختبار (العبد الله ، 2012) وبالتالي إذا فشل الفرد في الإجابة عن الفقرة يرجع إلى قدرته المحدودة وليس بسبب عدم تمكن الفرد من أن يصل إلى هذه الفقرة بسبب الزمن المحدود.

مطابقة البيانات للنموذج:

تتميز النظرية الحديثة في القياس بقدرة جيدة على حل بعض المشاكل التي يمكن أن تبرز أثناء اختبار الأفراد أو قياس بعض السمات التي يمتلكونها، ولكن يرى هامبلتون وسوامنيثان وروجرز (Hambleton & Swaminathan & Rogers ,1991) أن هذه القدرة تتحقق فاعليتها عند وجود مطابقة بين النموذج المستخدم وبيانات الاختبار.

ونظراً لعدم قدرتنا على إثبات صحة أو خطأ هذه النظرية فإن ملاءمة هذه النماذج لمجموعة من البيانات يمكن أن يبرهن من خلال مطابقة النموذج للعينة Model Test Data Fit. والهدف من اختبار مطابقة

النموذج للعينة هو تحديد ما إذا كان النمط العام لتوزيع الدرجات في مصفوفة البيانات يتفق مع النمط المتوقع طبقاً للنموذج المستخدم. وعند مطابقة النموذج للعينة يجب البحث في ثلاثة أمور هي:

أ. تحديد فيما إذا كانت البيانات تحقق افتراضات النموذج أم لا، وأول هذه الافتراضات أحادية البعد.

ب. التأكد من أن الغرض المنشود من استخدام النموذج قد تحقق مثل (ثبات تقديرات معالم الفقرة والقدرة).

ج. تحديد مدى التقارب بين القيم المتوقعة والقيم الملاحظة كالمقارنة بين تقديرات النموذج الحقيقية للدرجات على الاختبار.

ويذكر هامبلتون واسوامنيثان(Hambleton& Swaminathan,1985) أنه عند وجود تطابق تام بين نماذج الاستجابة للفقرة وعينة البيانات فإنه:

- بافتراض وجود مجتمع كبير من الأفراد، تكون تقديرات معالم الفقرة مستقلة عن عينة الأفراد.
- بافتراض وجود مجموعة كبيرة من الفقرات التي تقيس السمة، تكون تقديرات القدرة للأفراد مستقلة عن عينة الفقرات التي طبقت عليهم.

والجدير بالذكر أن هناك طرقاً متنوعة لفحص مدى تحقق افتراضات النموذج، فعلى سبيل المثال يمكن استخدام رسم الجذور الكامنة (Plot of Eigenvalues) وملاحظة العامل الأول السائد والنسبة بينه وبين العامل الثاني للتحقق من افتراض أحادية البعد في جميع نماذج الاستجابة للفقرة ، كما يمكن استخدام الحد الأدنى من التخمين (Minimal Guessing) والتي تهتم بأداء الأفراد ذوي القدرة المتدنية على الفقرات الصعبة في الاختبار (Hambleton & Swaminathan, 1985) ومن الطرق الإحصائية المستخدمة لفحص حسن المطابقة (Goodness of Fit) لنماذج الاستجابة للفقرة، إحصائي كاي تربيع (Chi Square) الذي يأخذ الصيغة الرياضية التالية للمفحوص i:

$$x^2 = \sum_{k-1} \frac{[0\mathrm{i}\ (k) - Ei(K)]^2}{Ei(K)}$$

حيث: s عدد فئات k، وOi (k) التكرار الملاحظ للقيمة k، و(k) Ei التكرار المتوقع للقيمة k.

كما يتم حساب التكرار المتوقع للأفراد لبديل ما باستخدام العلاقة:

$$Ei(k) = N\int P(Vi = k\,|\theta = t)f(t)dt$$

حيث: $f(t)$ دالة كثافة القدرة، والتي تعد توزيعاً سوياً لأن تدريج دوال الاستجابة للفقرة يتم بالرجوع إلى ذلك التوزيع.

حيث تعد طريقة تمثيل المطابقة بيانياً (FitPolt) إحدى الطرق التي يتم استخدامها للمقارنة بين تقديرات القدرة كما يتوقعها النموذج (Model Prediction) وكما هي فعلاً في البيانات الحقيقية (Actual Data)، ويتم الاعتماد في هذه الطريقة على المقارنة بين تقديرات معلم الصعوبة الفقرة المتوقعة والمشاهدة بسبب وجود هذا المعلم في نماذج الاستجابة للفقرة الثلاثة.

وتلخص طريقة تمثيل المطابقة بيانياً (Fit Plot) بالخطوات الآتية:

- تقسيم متصل القدرة إلى فترات، مثلاً (25) فترة.
- تقدير القدرة لكل مفحوص ثم حساب العدد الكلي للمفحوصين في كل فترة.

- حساب النسبة الإمبريقية وهي عدد المفحوصين الذين كانت استجابتهم صحيحة مقسوماً على العدد الكلي للمفحوصين في كل فترة.

وتكمن المشكلة في هذه الطريقة في أنه يصعب أن تتساوى قدرة الفرد المقدرة مع قدرته الحقيقية بسبب الخطأ في التقدير وخاصة في الاختبارات القصيرة حيث تزداد احتمالية الخطأ في تقدير القدرة.

الأسئلة:

لقد أولت غالبية الدول العربية اهتماماً خاصاً بتطوير أساليب التقويم التربوي وأدواته، وتبذل في هذا المجال جهوداً متزايدة تحقيقاً لهذا الهدف، ومن بين الجهود التي تبذلها كثير من الدول مشروعات بنوك الأسئلة Item Banking بمفهومها المعاصر، فبنوك الأسئلة التي يتم إعدادها على أسس منهجية وسيكومترية متطورة بالاستعانة بالتقنيات المناسبة يمكن أن تسهم في تطوير أساليب تقويم المتعلمين وأدواته، وتحسين نوعية الأسئلة والمفردات الاختبارية، وتشكيل اختبارات تحصيلية ذات خصائص معينة من الرصيد المتجدد لمفردات البنك (علام، 1997: ص9).

وتعد بنوك الأسئلة أحد الاتجاهات الحديثة في مجال القياس السلوكي، وبصفة خاصة تلك التي يتم إعدادها في ضوء النماذج اللوغاريتمية المرتبطة بنظرية الاستجابة للفقرة، وعلى الرغم من أن هذه البنوك يعتمد عليها كأحد الوسائل المتطورة في تحقيق الموضوعية المرجوة في هذا المجال، إلا أن هناك بعض المتغيرات التي يمكن أن تؤثر في دقة تدرج فقرات هذه البنوك وكذلك دقة معادلة درجات الاختبارات المختلفة التي يمكن أن تسحب منها، ومن هذه المتغيرات عدم ملاءمة البيانات للنموذج المستخدم في (الشافعي، 2014: ص431)

أهداف بنوك الأسئلة :

- تحقيق الموضوعية في تقويم التحصيل العلمي للطلبة، ويعد هذا هو الهدف الأساسي لبنوك الأسئلة، ويعتمد هذا الهدف على مجموعة من العوامل التي تؤثر على بناء وتكوين بنك الأسئلة ، مثل فلسفة القياس التي يعتمد عليها، نوع القياس الذي سيبنى عليه (وفق النظرية التقليدية، أو نظرية الاستجابة للفقرة، ووفق النموذج المناسب من نماذج الاستجابة للفقرة الذي سيتم تحليل البيانات الإحصائية على ضوئه، ومدى الدقة في صياغة الفقرات (الأسئلة) وتحليلها ومواصفاتها الإحصائية والعلمية، ومدى تغطيتها لمجال القياس من حيث المحتوى والمستوى.
- إتاحة الفرصة لقياس عمليات عقلية عليا ومتنوعة بتنوع مستويات ومجالات المادة العلمية وذلك نظراً لاتساع المدى الذي يمكن أن يغطيه البنك من حيث المحتوى والمستويات العقلية المختلفة.
- توفير وقت وجهد عضو هيئة التدريس في بناء الاختبارات التحصيلية، وذلك بحصوله على اختبارات جديدة من بنوك الأسئلة طبقاً لمواصفات وأهداف المقرر الدراسي، وذلك من خلال فقرات لها خصائص سيكومترية وإحصائية بمواصفات علمية، مما يتيح لعضو هيئة التدريس التركيز على الجوانب الأخرى من العملية التعليمية.
- تخفيض عوامل القلق والرهبة والتوتر المصاحبة لمواقف الاختبارات، وذلك باستخدام عينات من

البنك للتقويم المستمر أثناء العام الدراسي حتى يألف الطلبة هذا النوع من الأسئلة مع ضمان عدم التأثر من سرية الاختبارات.

- مرونة القياس حيث يسهل تشكيل أي اختبار في أي وقت بسرعة وسهولة.
- سرعة الحصول على نماذج وصور متكافئة من الاختبارات لها نفس القيم الإحصائية، تضمن العدالة في التقويم.
- توفير تدريج يوضح تقديرات الصعوبة المقابلة لكل فقرة من فقرات البنك، وكذا تدريج آخر لقدرة الفرد المقابل لكل درجة كلية محتملة سواء لبنك الأسئلة أو لأي صورة اختبارية.
- إمكانية المقارنة الموضوعية بين مستوى تحصيل الطلبة مهما اختلفت الأعوام، ومهما اختلف المدرِس، ومهما اختلفت الشعب أو المدارس أو الجغرافيا؛ طالما والمنهج واحد والعينة أخذت من نفس المجتمع الجغرافي، وقد تم تدريج البنود ومعادلة صوره الاختبارية. ويعد هذا الهدف منعطفاً أساسياً في تاريخ القياس النفسي والتربوي، ونقلة نوعية باتجاه تحقيق الموضوعية. (بنك الأسئلة مصر، جامعة حلوان: ص100-104)

خطوات إعداد بنوك الأسئلة:

1. **التخطيط والتنسيق:** وتشمل تحديد المادة الدراسية، واختيار مجموعة من معلميها ذوي الخبرة الجيدة في بناء الاختبارات، وتدريبهم على كيفية بناء بنوك الأسئلة، وتحديد أغراض بنك الأسئلة، وتحليل محتوى المنهج، وصياغة الأهداف السلوكية، وإعداد جدول المواصفات لأسئلة البنك وتحديد حجمه.
2. **التنفيذ:** وتشمل تحديد أنواع الفقرات، وصياغتها، وتحكيمها من حيث مدى قياس الهدف السلوكي، وصحة المحتوى العلمي، وتحديد الصعوبة المتوقعة لكل فقرة، وفي ضوء نظرية الاستجابة للفقرة يجب تحديد ما إذا فقرة تجيب على فقرة أخرى أو يفهم منها إجابة لفقرةٍ ما، وهو ما يسمى بالاستقلال الموضعي كأحد افتراضات نظرية الاستجابة للفقرة، وبناء الصور الاختبارية، واختيار التصميم المناسب؛ والتطبيق الاستطلاعي، والتحليل الإحصائي الأولي، وتقييم التحليل، وإعداد الصياغة النهائية للأسئلة.
3. **التجريب والتحليل:** وتشمل التطبيق النهائي، ثم إعداد مفتاح التصحيح، وتحديد الزمن المتوقع للإجابة عن كل فقرة، ثم تجريب الفقرات، وتصحيحها وتحليلها، والتحقق من افتراضات نظرية الاستجابة للفقرة، والتأكد من ملاءمة البيانات لأي من نماذج الاستجابة للفقرة، والتأكد من الثبات الإمبريقي، واستخراج تقدير معالم الأفراد والفقرات ودالة المعلومات، وتدريج الفقرات.
4. **التصنيف والتخزين:** وذلك في ضوء المادة الدراسية، أو الصف الدراسي، أو الوحدة الدراسية، أو مستويات الأهداف السلوكية، بحيث يكون لكل فقرة رمز يميزها، بالإضافة إلى المعلومات المتعلقة بنوع الفقرة، وزمن الإجابة عنها، ومفتاح الإجابة، وخصائصها السيكومترية؛ بتقدير معالم الفقرات، ومن خلال التدريج النهائي للفقرات (معادلة الصور

الاختبارية)، وتاريخ التطبيق، وملاحظات المعلمين أثناء التطبيق.

5. **الاستخدام والتقييم:** وتشمل البدء بسحب الفقرات للتطبيق أو للتدقيق والمراجعة دورياً أو عند الحاجة وذلك بهدف تحسين الصياغة، أو حذف وإضافة فقرات، حسب التغيرات التي قد تطرأ على المناهج الدراسية، وفق خطط وآلية علمية تشمل التطوير والتحسين الدوري لتحقيق الجودة الشاملة لبنك الأسئلة وللفقرات. وذلك من خلال تكوين فريق لإدارة البنك، يقوم بالإشراف على إجراءات تكوينه، وعمل جدول زمني للسير في هذه الإجراءات، وتوزيع المهام على أعضاء هيئة التدريس، وعمل آلية لتدقيق أسئلة البنك وتحكيمها، ومتابعة تحديث المعلومات في البنك والإشراف على استعماله. (علام،2005)

معادلة الاختبارات في بنك الأسئلة:

يوفر بنك الأسئلة رصيدا كبيرا من الفقرات البنكية يمكن من خلالها سحب صور مختلفة ومتنوعة من الاختبارات التي تستهدف قياس نفس المتغير، حيث يمكن سحب صور متكافئة من هذه الاختبارات، كما يمكن سحب اختبارات مختلفة من حيث مستوى الصعوبة، وبغض النظر عن مستوى صعوبة الاختبار المسحوب فقراته من بنك الأسئلة فإنه يعطي نفس تقديرات القدرة التي يمكن الحصول عليها من اختبار آخر مسحوب من نفس البنك، ويختلف عن هذا الاختبار من حيث مستوى الصعوبة، كما أن تقديرات القدرة أيضاً لا تختلف أيضاً باختلاف عدد الفقرات المحتواه بأي صورة اختبارية شريطة تمثيلها لنفس نطاق المتغير المستهدف من القياس، وهذا يتفق مع ما تحققه نظرية الاستجابة للفقرة لأداة القياس من تحرر عن خصائص العينة Item Free، وكذلك تحرر العينة من أداة القياس Sample Free (Baker & klim,2004:91)، وهذا ما أكدته دراسة (الشافعي ، 1996) من عدم اختلاف تقديرات القدرة المشتقة من اختبار مسحوب من بنك أسئلة عدد فقراته (34) فقرة، من تقديرات القدرة المشتقة من نفس البنك عدد فقراته (64) فقرة، ويتضح مما سبق إمكانية معادلة درجات أي صورتين اختباريتين مسحوب فقراتهما من بنك الأسئلة وإن اختلفت هاتان الصورتان من حيث عدد الفقرات أو مستوى الصعوبة (الشافعي، 2014)

دراسات سابقة:

أ. دراسات تناولت تدريج الاختبارات والمقاييس وفق نماذج نظرية الاستجابة للفقرة:

ـ دراسة العكايلة(2017) والتي هدفت إلى استخدام النموذج ثنائي البارامتر كأحد نماذج نظرية الاستجابة للمفردة في تدريج مفردات اختبار التفكير الاستدلالي المعرفي على طلبة المرحلة الثانوية في مديرية تربية لواء الجامعة بالأردن، وذلك من خلال التحقق من افتراضات نظرية الاستجابة للفقرة لبيانات العينة على اختبار التفكير الاستدلالي، ومدى ملاءمة البيانات المستمدة من استجابات أفراد العينة على فقرات اختبار التفكير الاستدلالي لافتراضات نموذج الاستجابة للفقرة المستخدم في تدريج فقراته، حيث طبق الباحث اختبار التفكير الاستدلالي على عينة من طلبة المرحلة الثانوية بمديرية تربية لواء الجامعة في الأردن والبالغ عددها (226) فرداً، وتوصلت الدراسة إلى تحقق افتراضات نظرية الاستجابة للفقرة وهي أحادية البعد بقيمة بلغت(0.245)، والاستقلال الموضعي بقيمة تجاوزت (2) والتحرر من السرعة، وتوافقت بيانات استجابات أفراد العينة من

افتراضات نموذج نظرية الاستجابة للفقرة حيث كانت قيمة معلم التمييز للفقرات الملائمة لافتراضات النموذج بين (0.099 ـ 1.010) بمتوسط قدره (0.68)، فيما بلغت قيم معلم الصعوبة للفقرات الملائمة لافتراضات النموذج بين القيمتين (-1.264 ـ 2.453) بمتوسط قدره (-3.024).

ـ دراسة طوخي (2017)، والتي هدفت إلى استخدام نظرية الاستجابة للفقرة لتدريج مقياس جودة الحياة وفق نموذج راش وعمل تقديرات السمة المقابلة لكل درجة خام كلية محتملة على المقياس في صورته النهائية والتحقق من صدق وثبات المقياس في صورته النهائية وعمل المعايير المختلفة التي عن طريقها يمكن تفسير مستويات الأفراد على مقياس جودة الحياة المدرج، حيث طبق المقياس على عينة بلغت (600) طالب من طلاب جامعة أم القرى، وتوصلت الدراسة إلى تدريج فقرات مقياس جودة الحياة على ميزان تدريج خطي واحد والتي بلغت (46) فقرة بعد حذف (14) فقرة غير ملائمة لافتراضات النموذج، ويتمتع المقياس بالصدق والثبات، كما أمكن للدراسة حساب تقديرات الأفراد المقابلة لكل درجة خام كلية محتملة على المقياس، كما أمكن للدراسة حساب معايير الرتب المئينية والدرجات التائية التي تفسر تقديرات الأفراد على المقياس المدرج.

ب. دراسات تناولت بناء بنوك الأسئلة وفق نماذج نظرية الاستجابة للفقرة:

ـ دراسة ناجي (2023) والتي هدفت إلى بناء بنك أسئلة في مقرر القياس والتقويم لطلبة جامعة إب باستخدام النموذج اللوغاريتمي ثلاثي المعلم، تم فيها بناء تجمع من الفقرات (288) فقرة، مكون من أربع صور اختبارية، وتكونت عينة الدراسة من (490) فرداً من طلبة كلية التربية جامعة إب، واستخدمت الدراسة البرنامج (Xcalibere.4.2.2)، وبعد التحقق من افتراضات نظرية الاستجابة للفقرة، وأيضا تم التطابق بين جميع الفقرات مع النموذج اللوغاريتمي ثلاثي المعلم، كما تم التطابق بين جميع الأفراد والنموذج، دون حذف أي فرد، وبلغت نتائج تقدير المعالم بعد التدريجعلى النحو الآتي: إن قيم معلم التمييز تراوحت بين (0.262 – 1.966)، بمتوسط (0.895)، وتراوحت قيم معالم الصعوبة بين (-1.132 ، +1.539) بمتوسط (0.356)، وتراوحت قيم معلم التخمين بين (0.012، 0.380) بمتوسط (0.251).

ـ دراسة عالم وأبو هادي (2019): والتي هدفت إلى بناء بنك أسئلة في مادة الرياضيات للصف التاسع الأساسي في ضوء نظرية استجابة الفقرة باستخدام النموذج اللوغاريتمي ثلاثي المعلمة، حيث قامت الدراسة بتحليل المحتوى وتحديد مخرجات التعلم المستهدفة، ومن ثم بناء تجمع من الفقرات، وتشكيل جدول مواصفات لمخرجات التعلم بفقرات اختبارية بلغ عددها (210) فقرة من نوع الاختيار من متعدد موزعة في ستة نماذج متكافئة من حيث المحتوى تكون كل منها من (40) فقرة، منها (6) فقرات مشتركة في كل النماذج، وطبقت الدراسة الاختبار على عينة مكونة من (1200) فرد، حيث طبق كل نموذج على(200) فرد، وتوصلت نتائج تحليل النتائج إلى حذف(4) فقرات لعدم مطابقتها لافتراضات النموذج الثلاثي المعلمة، وحذف (4) أفراد لم يتطابقوا مع افتراضات النموذج.

ـ اتفقت الدراسة الحالية مع الدراسات السابقة ببناء وتدريج أداة الدراسة وفق نظرية الاستجابة للفقرة،

واختلفت مع بعض الدراسات السابقة في الأداة المستخدمة وفي البرنامج الحاسوبي المستخدم في التحليل والتدريج.

منهج الدراسة وإجراءاتها:

منهج الدراسة:

اتبعت الدراسة الحالية المنهج السيكومتري، والذي يتناسب مع إجراءات بناء بنك الأسئلة بكل خطواته وعملياته، وهذا المنهج يهتم بتصميم وتطوير أدوات القياس في مجالات نفسية وتربوية مختلفة, وذلك بهدف الوصول إلى تقديرات كمية لجوانب الأداء في لحظة ما. ويستند هذا المنهج على مسلمتين متداخلتين مفادهما أنه كل شيء موجود إنما يوجد بمقدار، وكل ما وجد بمقدار يمكن قياسه (فرج، 2006: 102- 103)

مجتمع الدراسة:

تكون مجتمع الدراسة من جميع طلبة كلية التربية ـ جامعة صنعاء الذين يدرسون مقرر القياس والتقويم التربوي في المستوى الثالث بالأقسام العلمية والأدبية من مرحلة البكالوريوس للفصل الدراسي الثاني من العام الجامعي (2023/ 2024م)، والبالغ عددهم (320) طالب وطالبة.

عينة الدراسة:

تكونت عينة الدراسة من (119) طالبا وطالبة من المستوى الثالث في التخصصات العلمية والأدبية من طلبة البكالوريوس بكلية التربية – جامعة صنعاء من مجتمع البحث، الذين وافق أساتذتهم لمقرر القياس والتقويم التربوي على تطبيق الاختبارات عليهم [جدول (1)]، بعد استبعاد استجابات ثلاثة طلبة كانت استجاباتهم نمطية غير صالحة للتحليل.

جدول (1)توزيع أفراد عينة الدراسة وفقًا لمتغير التخصص

القسم	رياضيات	إنجليزي	معلم حاسوب	معلم مجال	رياض أطفال	الإجمالي
العدد	28	43	32	8	8	119

أداة الدراسة:

تكونت أداة الدراسة من بنك أسئلة في الوحدات الدراسية الثلاث الأولى من مقرر القياس والتقويم التربوي بكلية التربية- جامعة صنعاء، ويتألف من (112) فقرة من نوع الاختيار من متعدد، موزعة على ثلاث صور اختبارية متكافئة في المحتوى، تحتوي الصورة (44) فقرة بينها (10) فقرات مشتركة (رابطة) في الصور الاختبارية الثلاث.

وقد تم بناء البنك في ضوء الخطوات الآتية:

الخطوة الأولى- التخطيط والتنسيق:

وشملت هذه الخطوة الآتي:

أ. **تحديد المحتوى المراد قياسه:** حيث تم الاستفسار من أعضاء هيئة التدريس (دكاترة ومعيدين) الذين يدرسون المقرر من الكتاب المعتمد لهذا المقرر وتبين اعتماد أغلبهم على كتاب "مدخل إلى القياس والتقويم التربوي" إعداد: أ.د. صباح العجيلي، وتم تحديد المحتوى الدراسي

المراد قياسه بالوحدات الدراسية الثلاث الأولى لذلك الكتاب [جدول (2)].

جدول (2) الفصول الدراسية الأربعة الأولى وموضوعاتها في كتاب مدخل إلى القياس والتقويم التربوي

الفصل وعنوانه	الموضوعات
الفصل الأول والثاني: "التعريف بالقياس والتقويم" والمفاهيم الإحصائية الأساسية	- القياس والتقويم والاختبارات والعلاقة بينهما
	- دور القياس والتقويم في العملية التعليمية.
	- أنواع التقويم.
	- أنواع الاختبارات عامة.
	- المتغيرات.
	- مستويات القياس.
الفصل الثالث والرابع: "الاختبارات التحصيلية، الجزء الأول والثاني	- التخطيط للاختبار والاعداد له.
	- أنواع الاختبارات التحصيلية الشفوية والأدائية
	- الاختبارات القائمة على أساس تزويد الإجابة.
	- الاختبارات القائمة على أساس الاختيار.
	- تصحيح الاختبارات من أثر التخمين.
	ـ التحليل الإحصائي لفقرات الاختبار.

ب. تحليل المحتوى المراد قياسه إلى مكوناته:

حلل محتوى الوحدات الثلاث الأولى من الكتاب إلى عناصر رئيسة وجد أنها ممثلة بموضوعات الوحدات المشار إليها بجدول (2).

ج. صياغة الأهداف السلوكية:

تم صياغة (116) هدفًا سلوكيًا اشتقت من محتوى العناصر الرئيسة وفق تصنيف بلوم وتم الاكتفاء بالمستويات الأربعة الأولى (المعرفة، الفهم، التطبيق، التحليل) وتم مراعاة الأسس العلمية لصياغة الأهداف السلوكية.

وعرضت قائمة الأهداف السلوكية بمستوياتها المعرفية على مجموعة من المختصين في مجال القياس والتقويم التربوي في الجامعات اليمنية والمراكز البحثية[1] وذلك لإبداء الرأي حول: مدى كفاية الأهداف السلوكية في قياس العناصر الرئيسة للمحتوى، ومدى مناسبة صياغتها، ومدى ارتباطها بالمستويات المعرفية المندرج تحتها.

وفي ضوء آراء المحكمين تم الإبقاء على (112) هدفاً سلوكيًا لاتفاقهم على تحقق النقاط السابقة فيها، وتم حذف (4) أهداف سلوكية

1 ا.د. محمد الصوفي، ود. صفيه الغمري، ود. عبد السلام الخميسي كلية التربية جامعة صنعاء، وأ.د صباح العجيلي جامعة الإسراء ود.مأمون البناء جامعة إب ود.قايد المنتصر جامعة حضرموت.

لتشابهها، كما تم أخذ ملاحظات المحكمين بعين الاعتبار عند إجراء التعديلات اللازمة.

د. إعداد الخارطة الاختبارية:

تم إعداد الخارطة الاختبارية الشاملة اعتمادًا على (112) هدفا سلوكيا كعينة ممثلة للمحتوى [جدول (3)].

جدول (3) الخارطة الاختبارية الشاملة للصورة الاختبارية

الفصل	المحتوى	عدد المحاضرات	المستوى المعرفي للأهداف				مج ٪100
			يتذكر 25% ٪25	يفهم 37% ٪37	يطبق 24% ٪24	يحلل ٪14	
الأول والثاني	مفاهيم أساسية في القياس والتقويم	3 (0,33)	9	12	6	6	33
الثالث والرابع	الاختبارات التحصيلية وتحليل الفقرات إحصائياً	6 (0,67)	19	29	21	10	79
	مج	9	28	41	27	16	112

الخطوة الثانية – التنفيذ:

وشملت هذه الخطوة الآتي:

أ. صياغة الفقرات والتحقق من صلاحيتها:

تم صياغة الفقرات من نوع الاختيار من متعدد بالاعتماد على الكتاب المقرر، وبالاستعانة ببعض المراجع والدراسات العلمية ذات العلاقة كدراسة علام (2003) ودراسة جاسم (2009) وكتيب الهيثي والصوفي (2000)، وروعي مطابقة الفقرة للهدف السلوكي الذي تقيسه من حيث المحتوى والمستوى المعرفي، كما روعي في صياغتها شروط صياغة فقرات الاختيار من متعدد.

وتم إعداد ثلاث صور اختبارية متكافئة في المحتوى (A، B، C) تمثل المحتوى المراد قياسه ببنك الأسئلة في ضوء الخارطة الاختبارية المبينة بجدول (3). كل صورة اختبارية تحتوي (34) فقرة وهدفًا سلوكيًا غير مكررة، بالإضافة لـ(10) فقرات وأهداف سلوكية مشتركة (رابطة) في الصور الثلاث، ومتوسطة في صعوبتها وبنسبة 23٪ تقريباً من مجموع فقرات الصورة الواحدة البالغ عددها (44) فقرة.

ولكل فقرة اختبارية أربعة بدائل إجابة (أ، ب، ج، د) ثلاثة منها خاطئة وواحد فقط صحيح.

وعرضت الأهداف السلوكية وفقرات الصور الاختبارية الثلاث بصورتها الأولية على مجموعة من المختصين في القياس والتقويم التربوي[2]، وذلك لإبداء رأيهم وملاحظاتهم حول: مدى سلامة صياغة الفقرات ووضوحها لغويًا وعلميًا، ومدى قياسها للأهداف المعدة لقياسها. وقد حصلت جميع الفقرات على نسبة اتفاق (80٪) فأكثر، اضافة إلى إعادة صياغة بعض الفقرات لتناسب الهدف المعد لها أو لضمان سلامتها ووضوحها اللغوي بناء على ملاحظات المحكمين.

ب. تجهيز الصور الاختبارية بصورتها شبه النهائية

من خلال:

ـ إعداد تعليمات خاصة بالصور الاختبارية توضح بشكل مبسط ومختصر كيفية الإجابة على فقرات الاختبار، إذ طلب من الطلبة وضع دائرة على البديل الذي يمثل الإجابة الصحيحة، وتم حثهم على بذل أقصى جهدهم، وتوخي الدقة في حل جميع الفقرات الاختبارية.

ـ إخراج الصور الاختبارية الثلاث بصيغتها شبه النهائية، وقد تألفت الصورة الاختبارية الواحدة من (44) فقرة اختبارية لكل منها أربعة بدائل إجابة، وتعليمات للإجابة.

الخطوة الثالثة – التجريب والتحليل:

وشملت هذه الخطوة الآتي:

أ ـ التجربة الاستطلاعية: اعتمدت الدراسة التطبيق على أول مجموعة من الطلبة البالغ عددها (32) طالبا وطالبة من طلبة قسم معلم حاسوب كعينة استطلاعية نظرًا لأن عدد عينة الدراسة قليل لا تسمح بالاستغناء عن أي فرد فيها كعينة استطلاعية، وكان الغرض من ذلك:

- التأكد من وضوح التعليمات وصياغة الفقرات.
- معرفة الوقت الذي تستغرقه الإجابة عن جميع الفقرات في الصورة الاختبارية.
- التعرف على أهم الصعوبات والمعوقات التي قد تحدث أثناء التطبيق لتلافيها.
- وقد تبين وضوح التعليمات وجميع الفقرات باستثناء فقرتين، تم توضيحهما لهذه المجموعة وغيرها من المجموعات عند التطبيق كما تم إعادة صياغتهما. أما بالنسبة لزمن الإجابة فقد استغرق الطلبة ما بين (1:30 – 2) ساعة بمتوسط (1:45) ساعة من الإجابة على فقرات الصورة الاختبارية، علمًا بأنه لا يوجد زمن محدد للإجابة.

ب ـ تطبيق الصور الاختبارية لبنك الأسئلة على عينة الدراسة الأساسية والتصحيح، وتم ذلك من خلال:

ـ التنسيق مع مدرسي مقرر القياس والتقويم بكلية التربية – جامعة صنعاء لإجراءات تطبيق الصور الاختبارية لبنك الأسئلة.

ـ تجهيز الصور الاختبارية لبنك الأسئلة بحسب عدد الطلبة في كل تخصص.

ـ تطبيق الصور الاختبارية للبنك على عينة الدراسة الأساسية كاختبار نصفي درجاته معتمدة من أستاذ المقرر لضمان جدية الطلبة في الاستجابة للصور الاختبارية، وقد كان التطبيق خلال الفصل الدراسي الثاني للعام الدراسي 2024م/2025م .

2 ا.د. محمد الصوفي، ود. صفيه الغمري، ود. عبد السلام الخميسي كلية التربية جامعة صنعاء، وأ.د صباح العجيلي جامعة الإسراء ود.مأمون البناء جامعة إب ود.قايد المنتصر جامعة حضرموت.

ـ تصحيح أوراق الاختبارات يدوياً، وقد أعطيت درجة واحدة للفقرة عند اختيار بديل الإجابة الصحيح، والدرجة (صفر) للفقرة عند اختيار بديل إجابة خاطئ، ومن ثم فإن درجة الطالب الكلية في الاختبار تساوي عدد الفقرات الاختبارية التي أجاب عنها بصورة صحيحة. أما أعلى درجة يمكن أن يحصل عليها الطالب في الصورة الاختبارية فتساوي (44) درجة وأقل درجة تساوي (صفرا).

ـ استبعاد استجابات (3) من الطلبة كانت استجاباتهم نمطية غير صالحة للتحليل.

ـ إدخال البيانات الخام للمستجيبين إلى الحاسوب من خلال برنامجي Bilog mg3 & SPSS.

ج ـ التحقق من الخصائص السيكومترية للصور الاختبارية المكونة لبنك الأسئلة:

تم التحقق من الصدق الظاهري لفقرات بنك الأسئلة وفق الخطوات السابقة لبناء بنك الأسئلة، ويتحقق صدق الاختبار في نموذج الاستجابة للفقرة أحادي المعلم إذا تحققت افتراضاته، وتم ملاءمة الفقرات واستجابات الأفراد مع النموذج. وقد تحقق ذلك للصور الاختبارية الثلاث كما هو موضح لاحقاً في خطوات بناء بنك الأسئلة.

وللتحقق من ثبات الصور الاختبارية الثلاث المكونة لبنك الأسئلة وفق نموذج الاستجابة للمفردة أحادي المعلم:

ـ مؤشر الثبات للاختبار Test Reliability Index: حيث يشير مؤشر الثبات إلى ثبات قدرة الأفراد وفق السمة المقاسة. وقد قام الباحث بحساب مؤشر الثبات للصور الاختبارية (A ، B ، C) ووفق النموذج الأحادي المعلم، باستخدام برنامج(bilog mg3)

ـ الثبات الإمبريقي للاختبار Test Empirical Reliability : حيث يشير الثبات الإمبريقي إلى مدى اقتراب القدرة التي تم تقديرها من خلال نماذج الاستجابة للفقرة من القدرة الحقيقية للأفراد، ولقد قام الباحث بحساب الثبات الإمبريقي للصور الاختبارية (A ، B ، C) وفق النموذج أحادي المعلم (راش)، باستخدام البرنامج الإحصائي(Bilog mg3).

حسب (الثبات الامبريقي، ومؤشر الثبات) كما هو مبين في جدول (4).

جدول (4) معامل الثبات الامبريقي ومؤشر الثبات

الإحصاء	الصور الاختبارية		
	A	B	C
معامل الثبات الامبريقي	0.707	0.811	0.653
مؤشر الثبات	0.706	0.806	0.822

يتضح من الجدول (5) أن معامل الثبات الامبريقي، ومؤشر الثبات كانا جيدين ومناسبين للصور الاختبارية الثلاث.

د ـ التحقق من افتراضات نموذج الاستجابة للفقرة أحادي المعلم:

- **التحقق من فرض أحادية البعد*:**

اعتمدت الدراسة عددا من المؤشرات على أحادية البعد لبيانات الصور الاختبارية المكونة لبنك الأسئلة، كما أشار إلى ذلك هاتي (Hattie,1985)، وقد تحقق افتراض أحادية البعد (أي أن هناك سمة كامنة واحدة تقيسها الصور الاختبارية الثلاث المكونة لبنك الأسئلة هي المسؤولة عما يحدث من تباين في درجات عينة الدراسة) من خلال هذه المؤشرات:

- حسن مطابقة الفقرات في الصور الاختبارية الثلاث للنموذج أحادي المعلم، وتقاس جودة الفقرة بمدى ملاءمتها للنموذج (Greene,1993)، وكما هو مبين تالياً.
- حسن مطابقة الأفراد في الصور الاختبارية الثلاث للنموذج أحادي المعلم، حيث إن النموذج لم يحذف أي فرد من أفراد العينة وهذا يدلُ على قوة وجودة المطابقة للنموذج، وكما هو مبين تالياً.
- تدريج الصور الاختبارية الثلاث بفقرات مشتركة وفق النموذج أحادي المعلم على متصل السمة (القدرة، الصعوبة).
- **التحقق من فرض الاستقلال الموضعي:**

أشار همبلتون وسوامنثان Hambleton& (Swaminathan,1985 إلى أن فرض الاستقلال الموضعي يتحقق بتحقق فرض أحادية البعد. وبالتالي ما دام فرض أحادية البعد متحققا فإن فرض الاستقلال الموضعي متحقق.

ـ التحقق من فرض المنحنى المميز للفقرة:

للتحقق من هذا الفرض استخدم برنامج BILOG–MG3 لرسم المنحنيات المميزة للفقرات، حيث رسم منحنى مميز خاص لكل فقرة من الفقرات أخذت الشكل اللوغاريتمي الاحتمالي والتي تبين أنها متوازية.

ـ التحقق من فرض التحرر من عامل السرعة:

لتحقيق هذا الفرض تم إعطاء الطلبة الوقت الكافي لقراءة فقرات الاختبار واختيار الإجابة وفق قدراتهم، حيث ترك الوقت مفتوحًا أمامهم للإجابة. ولذلك يمكن القول إن إخفاق الطلبة في الإجابة عن بعض فقرات الصور الاختبارية يرجع أساساً إلى قدراتهم فقط.

هـ ـ التأكد من ملاءمة بيانات الدراسة (الفقرات، الأفراد) لنموذج الاستجابة للفقرة أحادي المعلم. (مبينة في نتائج الدراسة)

و ـ التعرف على تقدير معالم القدرة للأفراد والفقرات ودالة المعلومات للصور الاختبارية لبنك الأسئلة وفق نموذج الاستجابة للفقرة أحادي المعلم. (مبينة في نتائج الدراسة).

نتائج الدراسة ومناقشتها:

الهدف الفرعي الأول:

ينص الهدف الفرعي الأول على: **التأكد من ملاءمة بيانات الدراسة (الفقرات، الأفراد) لنموذج الاستجابة للفقرة أحادي المعلم:**

ـ التحقق من ملاءمة استجابات الأفراد على الفقرات الاختبارية لنموذج الاستجابة للفقرة أحادي المعلم: وتم ذلك باستخدام برنامج Bilog mg3 من خلال مطابقة الأفراد FIT PROB والنتائج مبينة في الجدول (6).

جدول (6) مطابقة الأفراد للنموذج أحادي المعلم

الصورة	مطابقة الأفراد FIT PROB	
	أدنى دلالة إحصائية لمربع كاي X^2	أعلى دلالة إحصائية لمربع كاي X^2
A	0.03	0.54
B	0.04	0.98
C	0.01	1.00

يتضح من الجدول (6) أنه في الدلالة الإحصائية للمطابقة (FIT PROB) ظهر أن جميع المستجيبين ضمن الدلالة المقبولة إحصائياً لمربع كاي X^2(0.01 >). وبذلك أسفر التحليل عن عدم حذف أي من الأفراد في الثلاث الصور الاختبارية والإبقاء عليهم جميعًا.

ـ التحقق من مطابقة الفقرات لنموذج الاستجابة للفقرة أحادي المعلم:

بحساب قيم مربع كاي كما هو مبين في جدول (7).

جدول (7) قيم مربع كاي

الإحصاء	A		B		C	
	Q	Sig	Q	Sig	Q	Sig
المتوسط	2.4	0.59	2.52	0.57	2.15	0.54
الانحراف المعياري	2.6	0.34	2.14	0.30	1.94	0.31
أكبر قيمة	11	0.99	10.1	0.99	8.00	0.98
أقل قيمة	0.1	0.00	0.1	0.00	0.1	0.00
الفقرات غير الملائمة	1، 4 ، 11 ، 17، 25، 28، 37		2، 21، 29، 38، 42		1، 29، 33، 42	

يتضح من الجدول (7) أن الفقرات غير المطابقة لنموذج الاستجابة أحادي المعلم هي سبع فقرات في الصورة الاختبارية (A)، وخمس فقرات في الصور الاختبارية (B)، وأربع فقرات في الصور الاختبارية (C)، لذلك تم استبعاد هذه الفقرات، وبالتالي أصبح إجمالي عدد الفقرات للصورة الاختبارية الواحدة (37 – 40) فقرة، وبذلك تحقق حسن مطابقة الفقرات للنموذج.

الهدف الفرعي الثاني:

ينص الهدف الفرعي الثاني على: **التعرف على تقدير معالم القدرة للأفراد والفقرات ودالة المعلومات للصور الاختبارية لبنك الأسئلة وفق نموذج الاستجابة للفقرة**

أحادي المعلم: تم استخدام برنامج (Bilog mg3) في تقدير معالم فقرات الصور الاختبارية كما هو مبين في جدول (8).

جدول (8) معالم قدرات الأفراد وفقرات الصور الاختبارية ودالة المعلومات

الصورة	الاحصاء	الصعوبة	القدرة	دالة المعلومات
A	المتوسط	0.19	0.00	2.509
	الانحراف المعياري	1.29	0.84	
	أكبر قيمة	2.79	2.14	
	اقل قيمة	-2.73	1.34-	
B	المتوسط	0.14	0.00-	4.501
	الانحراف المعياري	1.06	0.90	
	أكبر قيمة	1.77	2.51	
	أقل قيمة	-2.77	1.45-	
C	المتوسط	0.09	0.00	4.971
	الانحراف المعياري	1.61	0.86	
	أكبر قيمة	2.89	3.33	
	أقل قيمة	-3.69	1.33-	

من الجدول (8) نجد في الصورة الاختبارية (A) أن قيم معلم الصعوبة تمتد بين (2.79، 0.73-) وبمتوسط حسابي (0.19) وانحراف معياري (1.29)، وفي الصورة الاختبارية (B) أن قيم معلم الصعوبة تمتد بين (1.77، 2.77-) وبمتوسط حسابي (0.14) وانحراف معياري (1.06)، وفي الصورة الاختبارية (C) أن قيم معلم الصعوبة تمتد بين (2.89،3.69-) وبمتوسط حسابي (0.09) وانحراف معياري (1.61). وأن قيم القدرة في الصورة الاختبارية (A) تمتد بين (2.14،1.34-) وبمتوسط حسابي (0.00) لوجت، وبانحراف معياري (0.84) وفي الصورة الاختبارية (B) أن قيم القدرة تمتد بين (2.51،1.45-) وبمتوسط حسابي(0.00-) وبانحراف معياري (0.90) وأن قيم القدرة في الصورة الاختبارية (C) تمتد بين (3.33،1.33-) كما يتضح أن أعلى قيمة لدالة معلومات الاختبار للصور الاختبارية (C)، وأقل قيمة للصورة (A).

الهدف الفرعي الثالث:

ينص الهدف الفرعي الثالث على: "**تدريج فقرات بنك الأسئلة في مقرر القياس والتقويم التربوي وفق نموذج الاستجابة للفقرة أحادي المعلم**".

من خلال التدريج المشترك لفقرات الصور الاختبارية الثلاث المكونة لبنك الأسئلة في ضوء الخطوات الآتية:

بعد ملاءمة البيانات للنموذج أحادي المعلم، والحصول على معالم الفقرات المشتركة بين الصور الاختبارية باستخدام البرنامج الإحصائي (Bilog mg3)، واستخدم الباحث الطريقة الخطية في عملية تدريج الفقرات على مقياس عام مشترك، حيث استخدمت المعادلات الخطية التالية في عملية تحويل خطي لمعالم الفقرات إلى مقياس عام مشترك:

بافتراض أن معامل الارتباط بين by وbx يساوي واحد،

فإن: $\alpha = \sigma by / \sigma bx$

ولتحويل النماذج الاختبارية إلى مقياس عام مشترك، تم تثبيت مقياس الصعوبة المقدرة بوحدة اللوجيت للصورة الاختبارية (B)، ونقل مقياس الصعوبة للصورتين الاختباريتين (A ، C) إلى نفس المقياس من خلال معادلات خطية تم اشتقاقها باستخدام صعوبة الفقرات المشتركة بين الصور الاختبارية الثلاث، التي كانت عشر فقرات مكررة في كل صورة اختبارية، وبحساب متوسط (صعوبة الفقرات والانحراف المعياري) للفقرات المشتركة بين الصور الاختبارية (A ، B ، C) من خلال تحويل صعوبة الفقرة في الصورتين الاختباريتين (A ,C) إلى مقياس الصورة الاختبارية (B)، من خلال المعادلة التالية:

$$b_{j21} = 0.807\ b_{j1} + 0.023$$

حيث b_{j1} هي صعوبة الفقرة على مقياس الصورة الاختبارية (A)، وb_{j21} هي صعوبة الفقرة بعد تحويلها إلى مقياس الصورة الاختبارية (B) من خلال المعادلة:

$$\mathbf{b_{k21} = 0.657\ b_{k1} + -0.05}$$

$$b_{k21} = 0.657\ b_{k1} + -0.05$$

حيث b_{k1} هي صعوبة الفقرة على مقياس الصورة الاختبارية (C)، وb_{k21} هي صعوبة الفقرة بعد تحويلها إلى مقياس الصورة الاختبارية (B). والنتائج مبينة بجدول (9)

جدول (9) التدريج المشترك لفقرات الصور الاختبارية الثلاث المكونة لبنك الأسئلة

الفقرة	(b) بعد التدريج	القدرة عند أقصى معلومة θ	الفقرة	(b) بعد التدريج	القدرة عند أقصى معلومة θ
2	2.27	3.37	31	0.28	0.38
3	0.41	0.57	32	0.28	0.38
5	0.66	0.96	33	−1.62	−2.46
6	−0.47	−0.75	34	0.41	0.57
7	−0.34	−0.56	35	1.88	2.78
8	−0.21	−0.37	36	−0.34	−0.56
9	1.54	2.27	38	−0.74	−2.46
10	−0.60	−0.94	39	0.66	0.96
12	0.80	1.16	40	1.88	2.78
13	−0.34	−0.56	41	−0.74	−1.14

0.77	0.53	42	−1.99	−1.30	14
1.16	0.80	43	2.27	1.54	15
1.58	1.08	44	−1.35	−0.87	16
0.40	0.35	1	−0.94	−0.60	18
−0.58	−0.51	3	0.57	0.41	19
−0.43	−0.38	4	0.77	0.53	20
0.25	0.22	5	−1.99	−1.30	21
1.83	1.60	6	1.37	0.94	22
−0.15	−0.14	7	−0.75	−0.47	23
−1.01	−0.89	8	1.16	0.80	24
−0.72	−0.63	9	0.96	0.66	26
0.25	0.22	10	0.19	0.15	27
0.69	0.60	11	−1.14	−0.74	29
0.40	0.35	12	−0.75	−0.47	30

−0.94	−0.73	8	−0.16	−0.14	
0.36	0.36	9	0.40	0.35	14
−1.69	−1.36	10	1.47	1.29	15
1.05	0.94	11	−1.01	−0.89	16
0.36	0.36	12	0.11	0.10	17
−1.69	−1.37	13	0.40	0.35	18
1.65	1.44	14	−1.01	−0.89	19
0.20	0.22	15	1.47	1.29	20
−0.43	−0.31	16	1.65	1.44	22
0.20	0.22	17	0.11	0.10	23
0.20	0.22	18	1.14	1.00	24
−0.11	−0.04	19	−1.16	−1.02	25
0.70	0.64	20	1.30	0.14	26
−1.48	−1.19	21	−0.43	−0.38	27
1.24	1.10	22	0.69	0.60	28
1.11	−0.88	23	0.11	0.10	30
1.24	1.10	24	0.99	0.86	31

32	0.73	0.83	25	0.64	0.70
33	0.60	0.69	26	1.63	1.87
34	0.86	0.99	27	1.10	1.24
35	0.35	0.40	28	0.08	0.04
36	0.22	0.25	30	-1.03	-1.29
37	-1.02	-1.16	31	1.27	1.44
39	-0.02	-0.02	32	0.22	0.20
40	1.77	2.02	34	0.22	0.20
41	-1.02	-1.16	35	1.27	1.44
43	-0.38	-0.43	36	1.63	1.87
44	1.44	1.65	37	-0.04	-0.11
2	1.84	2.12	38	-0.45	-0.60
3	-0.58	-0.77	39	0.08	0.04
4	-1.74	-2.13	40	-0.58	-0.77
5	-0.88	-1.11	41	-0.73	-0.94
6	-1.03	-1.29	43	-0.45	-0.60
7	-1.03	-1.29	44	1.44	1.65

نلاحظ من الجدول (9) فاعلية نموذج الاستجابة للفقرة أحادي المعلم في تدريج الفقرات الاختبارية حيث امتدت معالم صعوبة فقرات بنك الأسئلة بعد تدريجه من (1.93-) إلى (2.27) لوجيت بمتوسط حسابي (0.65) وهي قيمة تقترب من الصفر مما يدل على أن فقرات البنك متوسطة الصعوبة في الغالب، وأن احتمال الإجابة الصحيحة عليها يقترب من (0.50)؛ بمعنى أن البنك يستهدف قدرات الأفراد بشكل جيد. وأعلى قيمة كانت للفقرة رقم (1) بقيمة (2.27) التي تقابل القدرة (3.37) وأدنى قيمة كانت للفقرة (61) بقيمة (1.93-) التي تقابل القدرة (2.20-).

وتشير النتائج السابقة إلى توفر بنك أسئلة وفق نموذج الاستجابة للمفردة مكون من (98) فقرة يقيس مستوى التحصيل في الوحدات الثلاث الأولى من مقرر القياس والتقويم التربوي لطلبة كلية التربية – جامعة صنعاء، تتوافر فيه شروط القياس الموضوعي.

التوصيات:

- إنشاء مركز للقياس والتقويم في جامعة صنعاء أسوة بالجامعات العربية والأجنبية، يهتم ببناء بنوك أسئلة للمقررات الجامعية.
- توحيد مساقات ومفردات المقررات الجامعية في كل كليات الجامعة ولكل المراحل الجامعية (بكالوريوس،

ماجستير، دكتوراه)، لما له من أهمية خاصة عند الحاجة لبناء بنوك أسئلة أو بناء اختبارات تحصيلية مقننة، وأيضا ليتمكن متخصصو القياس والتقويم وأصحاب القرار في الجامعة من إجراء المقارنات العلمية بين الأقسام والفروع المختلفة للجامعة.

- توفير البرامج الإحصائية الحديثة التي تتعلق بنماذج الاستجابة للفقرة، لما تتميز به هذه النظرية من دقة إحصائية وجودة موضوعية في التشخيص وتقدير القدرات. وتدريب أعضاء هيئة التدريس وطلبة الدراسات العليا على تلك البرامج.

- التأكد من ملاءمة البيانات للنموذج المستخدم وافتراضاته عند بناء بنوك أسئلة أو إعداد اختبارات تحصيلية وفق نماذج الاستجابة للفقرة وذلك لتحقيق الدقة الإحصائية في تقدير معالم قدرات الأفراد والفقرات.

- الاعتماد على النموذج الأحادي المعلم في إعداد الاختبارات التحصيلية عندما تكون العينة صغيرة كعينة هذه الدراسة.

- الإفادة من بنك الأسئلة المعد بهذه الدراسة في تحديد مستوى تحصيل طلبة كلية التربية بجامعة صنعاء في الوحدات الثلاث الأولى لمقرر القياس والتقويم التربوي.

المقترحات:

- إجراء المزيد من الدراسات لبناء بنوك أسئلة في كافة المقررات الدراسية في كلية التربية وغيرها من الكليات بجامعة صنعاء باستخدام نماذج الاستجابة للفقرة لما يحققه من شروط القياس الموضوعي.

- المقارنة بين نماذج الاستجابة للفقرة (أحادي، ثنائي، ثلاثي) المعلم في تقدير معالم بنك الأسئلة المعد في هذه الدراسة وغيره من الاختبارات والبنوك المعدَة للمقررات الدراسية في كليات الجامعة.

قائمة المصادر والمراجع:

أولاً: المراجع باللغة العربية:

[1] إبراهيم، نسرين وآخرون (2009). مقارنة بين النماذج اللوغاريتمية لنظرية الاستجابة للمفردة من حيث ملاءمتها للبيانات. **مجلة البحث العلمي في التربية**. العدد (10) الجزء الرابع.

[2] جاسم، خالد (2009). بناء بنك أسئلة في القياس والتقويم التربوي. **أطروحة دكتوراه غير منشورة**. كلية ابن رشد، جامعة بغداد.

[3] حامد، شيرين (2008). أثر نموذج الاستجابة للفقرة وتعدد الأبعاد وطريقة المطابقة في تقدير معالم الأفراد والفقرات، **أطروحة دكتوراه غير منشورة**، جامعة عمان العربية، كلية الدراسات العليا الأردن.

[4] حبيش، بشير وآخرون (2022): نظرية الاستجابة للمفردة الاختبارية (افتراضاتها ونماذجها)، **مجلة العلوم الإنسانية والاجتماعية**، مجلد 8 العدد 1، ص 200.

[5] حبيش، بشير وآخرون (2018). النظرية الحديثة في القياس (تحليل راش نموذجاً). **مجلة العلوم الاجتماعية ـ جامعة الأغواط، مجلد 07 العدد 31.**

[6] الدوسري، راشد (2001). معادلة الاختبارات: مفهومها، وطرقها، ومشكلات تطبيقها، **مجلة العلوم التربوية والنفسية**. جامعة البحرين مركز النشر العلمي، مجلد 2، العدد، ص: 123 – 137.

[7] سيف، محمد (2013). أثر عدد فئات المفردات المشتركة ومستوى صعوبتها على تدريج مفردات بنك الأسئلة المصمم باستخدام نموذج التقدير الجزئي، **اطروحة دكتوراه غير منشورة**، معهد الدراسات التربوية، جامعة القاهرة.

[8] الشافعي، محمد (2011): تأثير نماذج التحليل اللوغاريتمية الأحادية والثنائية والثلاثية البارامتر على تقديرات دوال معلومات المفردات الاختبارية وعلاقتها بدقة تقديرات القدرة: دراسة مقارنة. **المجلة المصرية للدراسات النفسية**، مجلد 21، العدد 73.

[9] الشافعي، محمد (2014). تأثير انتهاك بعض افتراضات نظرية الاستجابة للمفردة على تدريج بنك الأسئلة ودقة معادلة درجات الاختبارات البنكية المسحوبة . **مجلة كلية التربية** (جامعة بنها) . مصر، مجلد 25، العدد 98، ص: 437-440.

[10]الهيثي، خلف؛ والصوفي، محمد (2002) **دليل المعلم في تقويم الطلبة**، وزارة التربية والتعليم ـ اليمن.

[11]طيفور، مصطفى (2007). دراسة مقارنة لنماذج نظرية الاستجابة للمفردة في معادلة درجات الاختبار، **أطروحة دكتوراه غير منشورة**، جامعة القاهرة، معهد الدراسات التربوية.

[12]عالم، توفيق؛ وأبو هادي، سوسن (2019). بناء بنك أسئلة في مادة الرياضيات للصف التاسع الأساسي في ضوء نظرية استجابة الفقرة والتحقق من جودة توظيف برنامج حاسوبي لإدارته. **مجلة أبحاث**، مجلد (1)، العدد (15).

[13]العبد الله، زياد (2012): أثر بعض طرق التقدير على دقة تقدير المعالم ضمن نماذج الاستجابة للمفردة متعددة التدريج. **اطروحة دكتوراه غير منشورة:** معهد الدراسات التربوية، جامعة القاهرة.

[14]العجيلي، صباح (2011) مدخل إلى القياس والتقويم التربوي (ط.5) مركز التربية للطباعة والنشر.

[15]علام، صلاح الدين (1984). تحليل بيانات الاختبارات العقلية باستخدام نموذج راش اللوغاريتمي الاحتمالي (دراسة تجريبية) **المجلة العربية للعلوم الإنسانية**، جامعة الكويت العدد 117، ص: 100- 124.

[16]علام، صلاح الدين (1997). استخدام النماذج الإحصائية السيكومترية في تصميم أنظمة بنوك الأسئلة وبنائها، **المنظمة العربية للتربية والثقافة والعلوم ـ إدارة التربية** مج 17 العدد (2).

[17]علام، صلاح الدين (2003). **اختبار تشخيصي مرجعي المحك في أساسيات القياس والتقويم التربوي والنفسي** (كراسة الأسئلة). القاهرة: دار الفكر العربي.

[18]علام، صلاح الدين (2005). **نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**. القاهرة: دار الفكر العربي.

[19]مركز القياس والتقويم (2019) **دليل إعداد بنوك الأسئلة**، جامعة قناة السويس، جمهورية مصر العربية.

[20]مركز القياس والتقويم (2019) الدليل الإرشادي لإعداد بنوك الأسئلة بالجامعات، وحدة إدارة مشروعات تطوير التعليم العالي ـ وزارة التعليم العالي، جمهورية مصر العربية.

[21]مشروع تطوير نظم تقويم الطلاب والامتحانات، جامعة حلوان ص100-114.

[22]محمود، سومية (2012). استخدام نموذج راش في بناء بنك أسئلة لقياس التحصيل في مقرر سيكلوجية التعلم لدى طلاب كلية التربية بالمنيا . **أطروحة دكتوراه منشورة** فلسفة في التربية تخصص علم النفس التربوي، كلية التربية، جامعة المنيا.

[23]الكيلاني، عبدالله؛ والبرصان، إسماعيل (2017). **النظرية والتطبيق في نظرية الاستجابة للفقرة**، مترجم، دار جامعة الملك سعود للنشر.

[24]ناجي، ماجد (2023). بناء بنك أسئلة في مقرر القياس والتقويم التربوي لطلبة جامعة إب باستخدام النموذج اللوغاريتمي ثلاثي المعلم. **أطروحة دكتوراه غير منشورة**، جامعة الملك سعود، كلية التربية، قسم علم النفس.

[25]نور الدين، أمين (2002): فاعلية استخدام الاختبار الموائم باستخدام الحاسب في تقدير قدرة الأفراد وتحديد الخصائص السيكومترية للمقياس. **أطروحة دكتوراه غير منشورة**، كلية التربية، جامعة عين شمس.

[26]هياجنة، علي (2011). أثر طريقة انتقاء الفقرات في دالة معلومات الاختبار. **أطروحة دكتوراه غير منشورة**. جامعة اليرموك، كلية التربية، الأردن.

[27]فرج، صفوة (2006): **القياس النفسي**، القاهرة: مكتبة الأنجلو المصرية.

[28]فريحات، أيمن (2014): أثر عدد البدائل في دقة تقدير معالم فقرات الاختبار وفقا لنماذج الاستجابة للفقرة، **أطروحة دكتوراه غير منشورة**، جامعة اليرموك، كلية التربية . الأردن.

[29]المحرزي، راشد (2015). المفاضلة بين الدرجات المكافئة لنماذج اختبار القدرات العامة باستخدام طرق المعادلة الكلاسيكية في تصميم المفردات المشتركة بجماعات غير متكافئة. **مجلة العلوم التربوية والنفسية**، مجلد 16، العدد 3، ص: 398.

ثانياً: المراجع باللغة الإنجليزية:

[1] [1]Baker, F (2001): The Basics of Item Response Theory Eric clearing house on Assessment and Evaluation.

[2] [2]Crocker, L & Algina, J. (1986) Introduction to classical and modern test theory. New York: Holt, Rinehart,83(3), 154-159.

[3] [3]Embertson, S. E & Reise, S. P (2000). Item Response Theory for Psychologist. New jersey: Lawrence Erlbaum Associates, publishers.

[4] [4] Greene, W. H. (1993) Econometric Analysis, Second Edition. New York: Macmillan.

[5] [5]Hambleton, R & Swaminathan, H. (1985): Item response theory: Principles and applications. Boston: Kluwer.

[6] [6]Hambleton, R & Swaminathan. H, and Rogers H. (1991). Fundamentals of item response Theory. International Educational and professional. sage publication, Inc Newbury park, USA.

[7] [7] Hattie, J. (1985). Methodology review: Assessing One-dimensional of tests and items. Applied Psychological Measurment,9, 139-164.

[8] [8] Kolen, M & Brennan, R (2014) Test equating Scaling and Linking (3rd ed) New York: Springer.

[9] [9]Lord, F. (1980) Applications of item response theory to practical testing problems. Hillsdale: Lawrence Erlbaum Associates.

[10] [10]Mislevy, R. j & Bock, R. D. (1990). Bilog3: Item analysis and test scoring with binary logistic models (2nd ed). Scientific software, inc.

[11] [11] Nakamura, Y. (2006). Rasch Measurement and Item Banking Theory And Practice Mahwah, NJ: Erlbaum.

[12] [12] Pomplun, M., Omar, H. & Custer, M. (2004). A comparison of Winsteps and Bilog-mg for vertical scaling with the Rasch model. Educational and Psychological Measurement, 64(4), 600-616.

[13] [13]Reckase, M. (2009). Multidimensional Item response Theory. Dordrecht Heidelberg.

[14] [14]Reeve, B.B (2002) An Introduction to modern measurement theory Division of cancer control and population sciences, National cancer Institute.

[15] [15]Thorndike, R. (1997) Measurement and evaluation in psychology and evaluation. London Prentice-Hall.