دار نشر جامعة قطر
Qatar University Press

كلية التربية
College of Education
QATAR UNIVERSITY جامعة قطر

OPEN ACCESS
Submitted: 31 March 2018
Accepted Revised Article: 16 October 2018
Accepted: 15 November 2018

الدقة التشخيصيّة لقياسات الاستعداد العددي المبكّر القائمة على المنهاج في فرز تلاميذ الصفّ الأوّل المعرضين للفشل في الرياضيات

المؤلف المراسل: رحمة بنت صالح بن حديد العريمية
مشرفة تربويّة على برنامج صعوبات التعلّم مديريّة التربية والتعليم بمحافظة جنوب الشرقية
rahma.alerami@moe.om

محمود محمد إمام
أستاذ مشارك التربية الخاصة
جامعة السلطان قابوس
memam@squ.edu.om

يوسف عبد القادر أبو شندي
أستاذ مساعد القياس والتقويم
جامعة السلطان قابوس
yousefaaa@squ.edu.om

ملخص

تهدف هذه الدراسة إلى تصميم مجموعة من مقاييس المهارات الرياضية التي تقيس المعرفة بالرياضيات لدى تلاميذ الصف الأول – والتي يُطلق عليها «مقاييس الاستعداد العددي المبكر» – وتقييم القدرة التشخيصية لهذه المقاييس من ناحية فرز التلاميذ المعرضين للفشل في الرياضيات، حيث استُخدمت ثلاثة مقاييس للاستعداد العددي المبكر القائمة على المنهاج، وهي: مقياس تعرف العدد، ومقياس العدد المفقود، ومقياس التمييز بين الكميات.

وطُبّقت أدوات الدراسة على عينة من تلاميذ الصف الأول في محافظة جنوب الشرقية والبالغ عددهم 149 تلميذًا وتلميذة. وحُسب ثبات المقاييس بطريقتي الاتساق الداخلي وثبات استقرار التكافؤ. وكشفت المعالجات الإحصائية التي أُجريت على البيانات لاستخراج ثبات الاتساق الداخلي عن تراوح قيم معاملات كرونباخ ألفا ما بين 0.81 و0.97 للمقاييس الثلاثة المطبّقة. كما تراوحت قيم معاملات ارتباط بيرسون المحسوبة لاستخراج ثبات استقرار التكافؤ ما بين 0.72 و0.97 للمقاييس الثلاثة.

ووُجد الصدق العيني من خلال إعداد جداول المواصفات وإيجاد الصدق الظاهري من خلال آراء المحكّمين الذين اتفقوا بنسبة 100% على صلاحية المقاييس لتقييم مهارات الاستعداد العددي المبكّر. وقُدّرت الدقة التشخيصية لمقاييس الاستعداد العددي المبكر المطبّقة في الدراسة بإجراء تحليل ال ROC. وقد تراوحت قيم الحساسية ما بين 0.80 و0.91، ووجدت قيم النوعية ما بين 0.65 و0.70، وبلغت قيم التنبؤ الموجبة 0.57 في أعلى قيمها و0.47 في أقل قيمة لها، كما كانت قيم التنبؤ السالبة مرتفعة نسبيًا لجميع المقاييس وتراوحت ما بين 0.89 و0.96. وذلك بالإضافة إلى تراوح قيم الدقة التصنيفية ما بين 0.69 و0.74. ووضحت النتائج أنّ مقاييس الاستعداد العددي المبكّر لها القدرة على التنبؤ بأداء التلاميذ وفرز المعرّضين منهم للفشل في الرياضيات.

الكلمات المفتاحية: الدقة التشخيصية، الاستعداد العددي المبكّر، الفرز، الفشل في الرياضيات

لتوثيق هذه المقالة البحثية: العريمية ر.، إمام م. وأبو شندي ي.، «الدقّة التشخيصيّة لقياسات الاستعداد العددي المبكّر القائمة على المنهاج في فرز تلاميذ الصفّ الأوّل المعرضين للفشل في الرياضيّات»، مجلة العلوم التربوية، العدد 14، 2019

https://doi.org/10.29117/jes.2019.0012

© 2019، العريمية، إمام، أبو شندي، الجهة المرخص لها: دار نشر جامعة قطر. تم نشر هذه المقالة البحثية بواسطة الوصول الحر ووفقًا لشروط Creative Commons Attribution license CC BY 4.0. هذه الرخصة تتيح حرية إعادة التوزيع، التعديل، التغيير، والاشتقاق من العمل، سواء أكان ذلك لأغراض تجارية أو غير تجارية، طالما ينسب العمل الأصلي للمؤلفين.

دار نشر جامعة قطر
Qatar University Press

كلية التربية
College of Education
QATAR UNIVERSITY جامعة قطر

OPEN ACCESS
Submitted: 31 March 2018
Accepted Revised Article: 16 October 2018
Accepted: 15 November 2018

Diagnostic Accuracy of Early Numeracy Curriculum-Based Measurements in Screening for First Grade Students at Risk for Failure in Mathematics

Corresponding Author:
Rahma Saleh Hadid Al-Uraimi
Educational Supervisor in Learning Disability Program - Educational Directorate at Al-Sharqia Governorate
rahma.alerami@moe.om

Mahmoud Mohamed Emam
Associate Professor of Special Education, Sultan Qaboos University
memam@squ.edu.om

Yousof Abdelqader Abu Shindy
Assistant Professor of measurement and evaluation, Sultan Qaboos University
yousefaaa@squ.edu

Abstract

This study aims to examine the accuracy of three Early Numeracy Curriculum-Based Measurements (EN-CBMs). The measures were designed to assess three different skills of number knowledge, including number identification (NI), missing number (MN), and quantity discrimination (QD). The measures were applied to a sample comprising 149 first-grade students from Sharqia South Governorate. The results showed that the EN-CBMs have good internal consistency as indicated by Cronbach's Alpha coefficients ranging between 0.81 and 0.97. In addition, the measures showed good alternate forms reliability as indicated by correlations coefficients ranging between 0.72 and 0.97. The validity of the measures was assessed by both content and face validity. Content validity was based on the use of specification tables. Face validity exhibited 100% approval by a panel of experts. Receiver Operating Characteristic curve (ROC) analysis was used to examine the diagnostic accuracy of the EN-CBMs, using the Learning Disabilities Diagnostic Inventory (LDDI) as the outcome measure. Sensitivity values ranged from 0.80 to 0.91, specificity ranged from 0.65 to 0.70, Positive Predictive Value (PPV) ranged from 0.47 to 0.57, Negative Predictive Value (NPV) ranged from 0.89 to 0.91, and diagnostic accuracy ranged from 0.69 to 0.74. Overall, results showed that EN-CBM can predict can screen for students at risk for failure in mathematics.

Keywords: Diagnostic accuracy; Early Numeracy Curriculum-Based Measurement; Screening; Failing mathematics

لتوثيق هذه المقالة البحثية: العريمية ر.، إمام م. وأبو شندي ي.، «الدقّة التشخيصيّة لقياسات الاستعداد العددي المبكّر القائمة على المنهاج في فرز تلاميذ الصفّ الأوّل المعرضين للفشل في الرياضيّات»، مجلة العلوم التربوية، العدد 14، 2019

https://doi.org/10.29117/jes.2019.0012

© 2019، العريمية، إمام، أبو شندي، الجهة المرخص لها: دار نشر جامعة قطر. تم نشر هذه المقالة البحثية بواسطة الوصول الحر ووفقًا لشروط Creative Commons Attribution license CC BY 4.0. هذه الرخصة تتيح حرية إعادة التوزيع، التعديل، التغيير، والاشتقاق من العمل، سواء أكان ذلك لأغراض تجارية أو غير تجارية، طالما ينسب العمل الأصلي للمؤلفين.

مقدمة

تؤكد آراء عديد من التربويين والخبراء والباحثين الأثرِ الكبير لجودة عمليتي التعليم والتعلّم في سنوات الطفولة المبكرة لتطور الطفل في مراحله العمرية اللاحقة، ومن أجل ذلك أكدت أهمية رصد تعلّم التلاميذ وتقييمهم في تلك المرحلة المبكرة للمهارات الأساسية في القراءة والحساب (Brassard & Boehm, 2007). وفي هذا الصدد أشار التقرير النهائي للجنة الاستشارية للرياضيات (National Mathematics Advisory Panel, 2008) في الولايات المتحدة الأمريكية إلى أنّ الرياضيات التي يتعلّمها الأطفال منذ مرحلة رياض الأطفال وحتى الصفوف المتوسطة تزوّدهم بالركيزة الأساسية لتعلّم الجبر والرياضيات المتقدّمة.

ومن ثم كان لا بد من الاهتمام بالخبرات التعليمية في الرياضيات التي يتلقاها الطفل، ويتعرض لها منذ بداية التحاقه بالسلّم التعليمي الأكاديمي، وربطها بما اكتسبه من معارف رياضية في بيئته ومن خلال خبراته السابقة. ويُطلق على المهارات الرياضية المبكرة مسمى «مهارات الاستعداد العددي المبكر» (Early Numeracy Skills) (Hosp, Hosp, & Howell, 2016)، وهي تركّز على أحد أهم مكوّنات الرياضيات، وهو مكوّن الحس العددي (Number Sense)، حيث يشير مفهوم الحس العددي في صورته الأساسية إلى «القدرة على تعرف القيمة التي يمثلها العدد، والمرتبطة بالكميات الصغيرة، والسهولة في التعامل مع مهارات العد الأساسية، والكفاءة في تقدير الكمية المقترنة بالأعداد الصغيرة للأشياء، والعمليات الحسابية البسيطة». (National Mathematics Advisory Panel, 2008, p. 27)

وأولت الدراسات البحثية والأنظمة التعليمية الحديثة أهمية كبرى لعملية تقييم المهارات المبكرة في الرياضيات. وفي هذا الصدد ظهرت القياسات القائمة على المنهاج ((Curriculum-Based Measurement (CBM) كطريقة فاعلة ذات قاعدة علمية كبيرة. ولقد كان لدراسات المقاييس القائمة على المنهاج في مجال القراءة القدر الأكبر من مجمل البحوث والدراسات المنشورة (Tindal, 2013) في حين أنّ الدراسات التي ركّزت على القياسات القائمة على المنهاج في الرياضيات قليلة. وتشمل تلك القياسات في الرياضيات ما يلي: مقاييس الاستعداد العددي المبكر القائمة على المنهاج، مقاييس العمليات الحسابية القائمة على المنهاج، ومقاييس المفاهيم والتطبيقات الرياضية القائمة على المنهاج. وتُعتبر الدراسات التي أُجريت على مقاييس العمليات الحسابية أكثر مقارنة بمقاييس الاستعداد العددي المبكر، والمفاهيم والتطبيقات الرياضية وذلك في البيئة الغربية، أما في البيئة العربية فإنّ التركيز على القياسات القائمة على المنهاج بوجه عام سواء في القراءة أو الرياضيات محدود للغاية (Abu-Hamour, Al-Hmouz, & Kenana, 2013).

وفي البيئة الغربية أوجد الباحثون مجموعة واعدة من المقاييس القائمة على المنهاج في مهارات الاستعداد العددي المبكر. فمن أجل تحديد هذه المهارات طُورت مجموعة من المقاييس التي تغطي جوانب الحس العددي ومجالاته لدى الطفل. ويمكن إجمالها في مجموعات فرعية من المقاييس التي تستهدف تقييم عدد من المهارات، فتشمل: العدّ الشفوي (Oral Count-ing)، العدّ باللمس (Touch Counting) والتناظر (One-To-One Correspondence Counting)، تسمية الأرقام (Number Naming) أو تعرف الأعداد (Number Identification)، التمييز بين الكميات (Quantity Discrim-ination) أو المقارنة بين الكميات (Quantity Comparison)، العدد المفقود (Missing Number)، حقائق العدد (Number Facts)، وترتيب الكميات (Quantity Array).

وتتناول الدراسة الحالية إعداد وتصميم مجموعة من المقاييس القائمة على المنهاج التي تختبر امتلاك التلميذ لمهارات الاستعداد العددي المبكر. كما أنّها تركز على اختبار الدقة التشخيصية لهذه المقاييس باعتبارها أدوات تقييم للمهارات المبكرة لتعلّم الرياضيات، والتي يمكن من خلالها التنبؤ بالأداء اللاحق للمتعلّمين في تعلّم الرياضيات المتقدمة.

مشكلة الدراسة

على الرغم من التقدم الكبير الذي تشهده العملية التعليمية في مجال التقويم والتقييم في سلطنة عمان؛ إلا أنّه ومنذ إنشاء برنامج صعوبات التعلّم في العام الدراسي 2002/2001، عانى المتخصصون في مجال التربية الخاصة (Special Education) من ضعف أدوات التقييم في الميدان التربوي والتعليمي، حيث أشارت بعض التقارير والدراسات العربية (إمام وآخرون، 2013) والأجنبية (Emam & Kazem, 2016; Emam, Mohamed & Kazem, 2012) إلى اعتماد الميدان على عدد من الاختبارات الإدراكية التي تم تطبيقها وتقنينها على بيئات عربية أخرى غير البيئة العمانية.

وفي هذا الصدد، تمّ القيام بمشروع بحثي استراتيجي في الفترة ما بين 2012 و2015، بتمويل من المكرمة السامية لجلالة السلطان قابوس، استهدف بناء إطار منهجي لتشخيص صعوبات القراءة في الحلقة الأولى من التعليم الأساسي في سلطنة عمان، واحتلت فيه القياسات القائمة على المنهاج في القراءة مكانًا رئيسًا. وكان من ضمن توصيات المشروع تطبيق التجربة نفسها في مجال الرياضيات، وذلك للحاجة إلى أدوات تقييم قائمة على دراسات إمبريقية تتصف بالثبات والصدق.

وقدمت الشحي (2016) أول دراسة في مجال القياسات القائمة على المنهاج في مجال العمليات الحسابية في سلطنة عمان تم فيها اختبار الدقة التشخيصية للقياسات القائمة على المنهاج في العمليات الحسابية في فرز صعوبات تعلّم الرياضيات في الصف الرابع بالحلقة الأولى من التعليم الأساسي. واستكمالًا لهذا التوجه البحثي والميداني اختير في هذا البحث اختبار الدقة التشخيصية للقياسات القائمة على المنهاج في مهارات الاستعداد العددي المبكر لأنها سوف تخدم الميدان التربوي وخاصة في الحلقة الأولى من التعليم الأساسي، ومن ثم سيدعم ذلك إجراءات التشخيص المبكرة والتدخل العلاجي الوقائي المبكّر.

وبعد مراجعة وثيقة لتقويم تعلّم التلاميذ في مواد المجال الثاني للصفوف 4-1 (المديرية العامة للتقويم التربوي، 2015) تم التوصل إلى واقع عدم وجود مقاييس مطبّقة في المؤسسات التعليمية في سلطنة عمان لتقييم مهارات الاستعداد العددي المبكر، كما هو الحال في باقي الدول العربية. والواقع أنّ فهم العلاقة بين القدرات العددية المبكرة ومهارات الحساب اللاحقة تُعد أمرًا هامًا في فهم طبيعة التطور في مجال الرياضيات. وكثُرت المناظرات حول الصلة ما بين المعرفة الإجرائية (Procedural) والمعرفة المفاهيمية (Conceptual knowledge) بالأعداد في سنوات التعليم المبكرة واللاحقة.

وهناك العديد من الاقتراحات فيما يخص هذا الشأن لخّصها دوكر (Dawker, 2005) في ثلاثة اقتراحات هي: أنّ مبادئ العدّ (Counting Principles) تُعتبر مكونًا رئيسًا لتطوير الدقة في إجراءات العدّ (Counting Procedures)؛ وأنّ التوسع في مبادئ العد يُعتبر كنتيجة لتجريد إجراءاته؛ وأنّ العملية تبادلية وتفاعلية في تطوير كل من المعرفة الإجرائية والمعرفة المفاهيمية للأعداد.

وقد ذكر هوسب وآخرون (Hosp et al., 2016) أنّ المعضلة الرئيسة في تطوير مقاييس الاستعداد العددي المبكّر القائمة على المنهاج تكمن في عدم التحديد الدقيق للمهارة التي تمثّل المنبئ الأكثر دلالة على الحس العددي؛ بمعنى آخر: أيّ مكوّن من مكوّنات القدرة العددية هو الأكثر أهمية في الاستدلال عليها؟ لهذا تُستخدم عدة مقاييس لأكثر من مهارة لتقييم مهارات الحس العددي. كما تُعدّ المهارات التي تقيسها مقاييس الاستعداد العددي المبكّر أساسية ليكتسب الطفل الطلاقة في التعامل مع الأعداد وإجراء العمليات الحسابية وحل المشكلات الحياتية.

ولا بدّ من تقييم مستمر ومنظّم للمخرجات التعليمية التي تنبثق من مرور التلميذ بخبرات أكاديمية تعليمية يكتسبها في بيئة التعلم. ووفقًا للتصريح الصادر عن الرابطة الوطنية لتعليم الأطفال الصغار (National Association for the

Education of Young Children) والذي نصّ على أنّ «الهدف والغاية من أيّة عملية تقييم لا بدّ أن يكون من أجل تحسين وتطوير مستوى الخدمات المقدّمة للطفل، ولتأكيد أنّ الطفل مستفيد فعلًا من الخبرات التعليمية التي يمر بها»، وذلك حسبما ورد في (Brassard & Boehm, 2007, p. 2). في هذا الإطار، من الضرورة بمكان إيجاد أدوات تقييم لها خصائص ومميزات تمنحها امتيازًا عن سواها من أدوات التقييم المتوافرة في الميدان والحقل التربوي والعمل على تطويرها.

ومن هذا المنطلق، طُورت المقاييس القائمة على المنهاج (Curriculum-Based Measurement). فقد أشار هوسب وآخرون (Hosp et al., 2016) إلى أنّ هذه المقاييس ليست مجرد تقييمات إضافية على المعلّمين تطبيقها إلى جانب الأساليب التقييمية المألوفة والمعمول بها؛ ولكنّها تُعدّ بديلًا عن الممارسات التقييمية التي قد يكون التربوي مطبّقًا لها بالفعل أو يحاول تجنّبها لأسباب عدة: كأن تكون معقدة في الأداء والتطبيق، أو من الصعب تكييفها في الأوضاع الأكاديمية المختلفة، أو بسبب استهلاكها للكثير من الوقت سواء في التطبيق أو التصحيح أو تفسير النتائج أو جميع ما سبق.

ولذا ركّزت هذه الدراسة على إعداد مقاييس للاستعداد العددي المبكّر قائمة على المنهاج وتتلاءم مع ما يدرسه تلاميذ الصف الأول الأساسي وفقًا للمنهج الدراسي في سلطنة عمان، ودراسة الخصائص السيكومترية لهذه المقاييس والكشف عن الدقة التشخيصية لها في فرز التلاميذ المعرضين للفشل في الرياضيات في الصف الأول الأساسي.

أسئلة الدراسة

تحاول هذه الدراسة الإجابة على الأسئلة الآتية:

- ما هي المعايير القياسية (Benchmark) لأداء تلاميذ الصف الأول الأساسي على مقاييس الاستعداد العددي المبكّر القائمة على المنهاج (EN-CBM)؟
- ما هي مستويات الدقة التشخيصية: الحساسية (Sensitivity) والتحديد النوعي (Specificity) لمقاييس الاستعداد العددي المبكّر؟
- هل توجد فروق دالة إحصائيًا عند مستوى دلالة $\alpha = 0.05$ في أداء تلاميذ الصف الأول الأساسي في مقاييس الاستعداد العددي المبكّر القائمة على المنهاج تُعزى لمتغيري النوع والالتحاق برياض الأطفال والتفاعل بينهما؟

أهمية الدراسة

تكمن أهمية الدراسة الحالية من الناحية النظرية في أنّها تلبي تطلعات الباحثين في مجال مرحلة الطفولة المبكرة بشكل عام وتعلّم الرياضيات في هذه المرحلة بشكل خاص، حيث تستهدف الدراسة موضوعًا حديثًا يتناول مقاييس الاستعداد العددي المبكّر القائم على المنهاج (Early Numeracy CBM) في سلطنة عمان. فبالرجوع إلى مصادر البحث وقواعد البيانات المفهرسة مثل دار المنظومة وشمعة بالإضافة إلى المراجع المتعددة لم يتم التوصل إلى دراسات وبحوث عربية منشورة ذات علاقة بهذا الموضوع.

كما تتمثّل أهمية الدراسة من الناحية التطبيقية في كونها تقدم مقاييس جديدة لتقييم مهارات الاستعداد العددي المبكّر ذات كفاءة تشخيصية ودقة تصنيفية عالية للميدان التربوي، وتستهدف تلاميذ الصف الأول للتعليم الأساسي. وهي في الوقت ذاته توسع قاعدة البيانات للدراسات والبحوث الإمبريقية التي استهدفت الدقة التشخيصية (Diagnostic Accuracy) لمقاييس الاستعداد العددي المبكر القائمة على المنهاج. إضافة إلى ذلك، فهي تقدم أدلة صدق وثبات للمقاييس، الأمر الذي يجعل للبيانات المستخرجة من تطبيقها موثوقية أكبر في اتخاذ القرارات التعليمية المرتبطة بفاعلية استراتيجيات التعليم وأساليبه المتبعة مع التلاميذ، أو بتقديم الدعم للذين يظهرون عجزًا في اكتساب المهارات الرياضية التي تم تقييمها.

أهداف الدراسة

- تصميم مقاييس الاستعداد العددي المبكر التي تقيس المهارات الرياضية ذات الصلة بالقدرة العددية المبكرة.
- تحديد الخصائص السيكومترية للمقاييس من حيث الثبات والصدق.
- تحديد معايير الدقة التشخيصية التي تمتلكها مقاييس الاستعداد العددي المبكّر في فرز التلاميذ المعرضين للفشل في الرياضيات، وذلك من خلال تحديد الحساسية والنوعية والدقة التصنيفية.

مصطلحات الدراسة

الدقة التشخيصية (Diagnostic Accuracy): هي قدرة مقاييس الاستعداد العددي المبكّر القائمة على المنهاج (Early Numeracy CBM) المطبّقة في الدراسة الحالية على التحديد الدقيق للأفراد الذين سيقع مستوى أدائهم عند أو أقل من الدرجة المئينية 25 (25th. Percentile) (Clarke et al., 2011).

المقاييس القائمة على المنهاج (Curriculum-Based Measurement (CBM)): وهي عبارة عن مجموعة من التقييمات المكررة والمستمرة التي تهدف إلى قياس ورصد التغيّر في تعليم وتقدّم التلاميذ خلال فترة زمنية قصيرة. وهي تقييمات موضوعية وصادقة وذات خطوات تطبيق معيارية، كما أنّها متصلة بصورة مباشرة بالمهارات المتضمّنة في المنهج الذي يدرسه التلميذ.

مقاييس الاستعداد العددي المبكرة القائمة على المنهاج (Early Numeracy CBM (EN-CBM)): هي مجموعة مطورة من المقاييس تُطبّق فرديًا على تلاميذ الصف الأول خلال دقيقة واحدة لتقييم قدرة التلميذ في مهارات الحس العددي. وتستهدف القدرة على: تعرف الأعداد (Number Identification)، إيجاد العدد المفقود (Missing Number) ضمن سلسلة من الأعداد، والتمييز بين الكميات (Quantity Discrimination).

الفرز (Screening): هو عملية تحديد التلاميذ الذين يظهرون ضعفًا في امتلاك مهارات الاستعداد العددي المبكّر المُطبّقة في الدراسة.

التلاميذ المعرضون للفشل في الرياضيات (Students at risk of failure in mathematics): هم التلاميذ الذين يقع مستوى أدائهم عند درجات القطع المناظرة للدرجة المئينية 25 (25th. Percentile) 25 أو أقل منه في مقاييس الاستعداد العددي المبكّر القائمة على المنهاج المطبقة.

حدود الدراسة

- **الحدود الموضوعية:** تقتصر الدراسة على موضوع مقاييس الاستعداد العددي المبكّر القائمة على المنهاج.
- **الحدود البشرية والمكانية:** تم تطبيق الدراسة على تلاميذ الصف الأول من التعليم الأساسي في محافظة جنوب الشرقية في سلطنة عمان.
- **الحدود الزمانية:** تم تنفيذ الدراسة خلال العام الأكاديمي 2017/2016.

الإطار النظري والدراسات السابقة

ارتبط ظهور القياسات القائمة على المنهاج واستخدامها بظهور مدخل الاستجابة للتدخل العلاجي (Response to Intervention RTI)، حيث يتضمّن هذا المدخل اتخاذ القرارات القائمة على البيانات (Data-Based Decisions) من أجل حلّ المشكلات التربوية المختلفة لنجاح مدخل الاستجابة للتدخل وجودته. ولا بدّ من توافر عدد من المكونات الأساسية التالية: استخدام التقييمات لاتخاذ قرارات المسح الشامل وقرارات رصد التقدّم، وتقديم التعليم والتدخلات ضمن نظام

متعدد المستويات بحيث يضمن حصول التلاميذ الأكثر احتياجًا للدعم على التعليم والمساندة بقدر يناسب ذلك الاحتياج. وتُعدّ المقاييس القائمة على المنهاج وسيلة ممتازة لاتخاذ القرارات المتعلّقة بالمسح الشامل ورصد تقدّم التلاميذ، وهذا ما يتمحور حوله مدخل الاستجابة للتدخل (Hosp et al., 2016).

وتُعّد المقاييس القائمة على المنهاج (CBM) نموذجًا للتقييمات القائمة على المنهاج (Curriculum-Based Assess-ments) التي تم تطويرها كما ورد في رايلي هيلر وكيلي فاين وشريفر (Riley, H. Kelly, V. & Shriver, 2005) على يد ستانلي دينو Stanley Deno وفيليس ميركن Phyllis Mirken، وهي عبارة عن مجموعة من التقييمات المكررة والمستمرة التي تهدف إلى قياس ورصد التغيّر في تعليم وتقدم التلميذ خلال فترة زمنية قصيرة، وكذلك تقييم تقدُّمه عبر الزمن في مجالات القراءة والرياضيات والعمليات الحسابية والإملاء والكتابة، وهي تقييمات موضوعية وصادقة وذات خطوات تطبيق معيارية، كما أنّها متصلة بصورة مباشرة بالمهارات المتضمّنة في المنهج الذي يدرسه التلميذ.

وقد كان المقصد الرئيسي من تطوير المقاييس القائمة على المنهاج استخلاص مواد من صُلب المنهج التعليمي. وتقوم فكرة إيجاد واستخدام المقاييس القائمة على المنهاج على الأساس الذي يَفترض أنّ التقييم واتخاذ القرارات أمران يعودان إلى المنهج (Riley, H. Kelly, V. & Shriver, 2005).

وأشار فيوكس وفيوكس (Fuchs & Fuchs, 2001) إلى أنّه تم تطوير المقاييس القائمة على المنهاج من أجل تزويد المهتمين بالمعلومات المتكررة بما يجب التركيز عليه في تدريس التلاميذ من المهارات المختلفة التي يتضمّنها المنهج.

وتختلف عدد مرات تطبيق المقاييس القائمة على المنهاج تبعًا للغرض التربوي من هذه المقاييس، وبالتالي لطبيعة القرارات التربوية التي سيتخذها التربويون. ففي حالات رصد التقدم العلمي للتلميذ يتطلّب الأمر إجراء تقييمات أسبوعية. في الدراسة التي أجراها هومبتون وآخرون (Hompton et al., 2012)، تم فحص ستة مقاييس مطبقة لمراقبة التقدم الرياضي لتلاميذ الصف الأول، والتي طُبّقت بصورة أسبوعية.

ومن بين المقاييس التي طبقتها الدراسة ما يلي: حقائق العدد، والتمييز بين الكميات، والعدد اللاحق. أظهرت نتائج الدراسة أنّ هناك تقدمًا ذا دلالة عبر فترات تطبيق المقاييس على تلاميذ الصف الأول. أما في حالات الفرز، فيتطلّب الأمر إجراء تقييمات متكررة ثلاث مرات خلال العام الدراسي، وهذا ما قامت به دراسة كلارك وآخرين (Clarke et al., 2011) حيث طُبّقت المقاييس خلال بداية ومنتصف ونهاية سنة دراسية.

ومن أجل جذب اهتمام المعلّمين وشدّ انتباههم في المدارس العامة، تم تقديم معلومات المقاييس القائمة على المنهاج على أنّها معلومات تشير إلى نتائج تلاميذ الفصل الذي يدرّسه المعلّم. وقد ركّز المعلّمون انتباههم على فئة مُختارة من التلاميذ الذين يحتاجون إلى قدر كبير من الدعم والمساندة.

ويُستفاد من هذه المقاييس بصورة كبيرة جدًا في كونها تقدّم ما يلي:

- تقييمًا أسبوعيًا على مستوى الفصل، حيث يقوم المعلّم بتطبيق المقاييس لتتبّع تقدم التلاميذ في امتلاك المهارات المتعلمة، وإتقان تطبيقات وممارسات المنهج. ويتم التطبيق الأسبوعي باستخدام صور متكافئة من الاختبارات التي تشتمل على المهارات المتضمنة في المنهج، على سبيل المثال، إذا كان المعلّم يدرِّس مادة الرياضيات لتلاميذ الصف الثالث فإنّه يقوم باختبار جميع تلاميذ الفصل على مقاييس تستهدف المشكلات الرياضية التي يتضمّنها منهج الصف الثالث وبشكل متوازٍ مع الوزن النسبي لها في هذا المنهج.

- تغذية راجعة عن أداء التلاميذ كل أسبوعين، حيث يتم استخدام برامج الحاسوب لتلخيص أداء التلاميذ واستعراضه بصور مختلفة تتمثل في مخططات بيانية توضّح درجات التلميذ بمرور الوقت، وملف الإنجاز الخاص بالمهارات التي أتقنها التلميذ خلال أسبوعين في كل مجال من مجالات المنهج.

ولقد اكتسبت مهارات الاستعداد العددي المبكر شأنًا كبيرًا واهتمامًا بحثيًا رائدًا في مجال تقييم القدرات الرياضية في الآونة الأخيرة، وذلك على الرغم من عدم وصول مستوى الوعي بقيمة اكتساب المهارات العددية المبكّرة بالنسبة للرياضيات، ذات القدر الذي وصل إليه مستوى الوعي بأهمية مهارات الاستعداد لاكتساب اللغة بالنسبة للقراءة (Hosp et al., 2016). وتركّز مقاييس الاستعداد العددي المبكّر القائمة على المنهاج (EN-CBM) على مجال مهم في الرياضيات، وهو ما يُمكن تسميته بالحس العددي (Number Sense).

ولا تزال البحوث والدراسات المتعلّقة بالاستعداد العددي المبكّر في مراحلها الأولى، إلا أنّ هناك مجموعة من المقاييس الواعدة التي طورتها تلك الدراسات. فقد طبّقت دراسة ميث وبيجيني وليري (Methe, Begeny & Leary, 2011) أربعة مقاييس منفصلة، طُبّقت بثلاث صور متكافئة على تلاميذ الصف الأول، ومن بينها مقاييس: التجميع بالخمسة، والترتيب حتى عشرة، والتفكيك، والحقائق اللفظية. وقد توصّلت الدراسة إلى أنّ بعض المقاييس ذات خصائص تقنية قوية.

كما بحثت عديد من الدراسات استخدام هذه التقييمات المستمرة والمتكررة خلال العام الدراسي، واختبرت مدى فاعليتها في فرز المتعلّمين ورصد تقدّمهم في مختلف المراحل الدراسية عامة وفي المرحلة الأساسية المبكّرة بشكل خاص. وقد استهدف لي وليمبكي ومور وجينسبرغ وباباس (Lee, Lembke, Moore, Ginsburg & Pappas, 2012) فحص الكفاءة التقنية لمقاييس الاستعداد العددي المبكّر والتي ضمت المقاييس التالية: العد الشفوي، وتعرف العدد، والعدد المفقود، والعدد اللاحق، وحقائق العدد. وطُبّقت هذه المقاييس على تلاميذ رياض الأطفال والصف الأول. وبينت هذه الدراسة أنّ المهمات التي تناولتها المقاييس المُطبّقة تقيس بالفعل مهارات الاستعداد العددي المبكر.

المهارات التي تقيسها مقاييس الاستعداد العددي المبكر القائمة على المنهاج

نظرًا لأهمية الاستعداد العددي المبكّر في تعلّم الرياضيات اللاحقة واستنادًا إلى صعوبة تحديد المهارة التي من شأنها أن تدلّ على امتلاك التلاميذ لهذه المهارة كان هناك تنوعٌ في المقاييس المُطبّقة في الدراسات التي بحثت في هذا المجال. فمن بين هذه المقاييس ما يلي: العدّ باللمس (One-to-one Corresponding Counting)، العدّ الشفوي (Oral Counting)، تسمية الأرقام (Number Naming)، والمقارنة بين الكميات (Quantity Comparison)، (Laracy, Hojnoski, & Dever, 2016 ; Floyd, Hojnosky & key, 2006).

وأضافت دراسة (Clarke & Shin, 2004) مقياس العدد المفقود (Missing Number) إلى مجموعة المقاييس في الدراسات السابقة واقتصار قياس مهارة العدّ على العدّ الشفوي فقط. وفي الدراسة الحالية، تم الاقتصار على أكثر ثلاثة مقاييس تطبيقًا وقدرةً على التعبير عن الاستعداد العددي المبكر.

التعرف على العدد (NI) Number Identification

لقد صُمّمت مقاييس تعرف العدد (NI-CBM) لتقييم الحس العددي لدى المفحوصين، والذي يُعدّ من أساسيات المهارات الرياضية. وهذه المهارة شأنها في الأهمية بالنسبة للرياضيات كشأن مهارة تسمية الحروف بالنسبة للقراءة، فهي مؤشر مهم، ويدل على الآلية (Automaticity) ويقيسها في آن واحد.

العدد المفقود Missing Number (MN)

تتطلب مقاييس العدد المفقود القائمة على المنهاج (MN-CBM) من التلميذ أن يتعرّف العدد المفقود ضمن نمط لسلسلة عددية مكوّنة من أربعة أعداد، يتم تزويد المفحوص بثلاثة أعداد ضمن السلسلة، وعلى التلميذ أن يتوصّل للعدد المفقود والمشار إلى موقعه بعلامة شرطه «ـ». وتُعدّ مهارة تعرّف النمط من المهارات الأساسية في علم الجبر. كما أنّها من تطبيقات مهارة الحس العددي.

التمييز بين الكميات Quantity Discrimination (QD)

يتطلّب هذا المقياس من المفحوص أن يحدد أي عدد من بين زوج من الأعداد هو الأكبر كمًّا. وتعتبر القدرة على تحديد المقادير التي تعبر عنها الأعداد مكونًا أساسيًا للحس العددي (Number Sense)، ومتطلبًا أوليًا كمهارة لإجراء العمليات الحسابية وحلّ المشكلات الرياضية.

وتتساوى المدة الزمنية اللازمة لتطبيق مقاييس الاستعداد العددي المبكّر القائمة على المنهاج في أربعة مقاييس: العدّ (Counting)، تعرف العدد (Number Identification)، العدد المفقود (Missing Number)، والتمييز بين الكميات (Quantity Discrimination). وبمجرد جلوس التلميذ على مقعد التقييم، وحلّ مقياس الاستعداد العددي المحدد بدقيقة واحدة، وتسجيل استجاباته على ورقة التقييم الخاصة بالفاحص سيكون قد انقضى من الوقت ما يعادل دقيقتين إلى ثلاث دقائق. وفي حال تطبيق التقييمات – بغرض المسح الشامل والذي ينفذ بأن يحل المفحوص ثلاث نسخ متكافئة من كل مقياس – فسيصبح زمن التطبيق المستغرق مقاربًا لخمس إلى ست دقائق.

الدقة التشخيصية للمقاييس المستخدمة في المسح الشامل

يُقصد من عملية المسح الشامل تحديد التلاميذ الذين يحتاجون للمساندة والدعم. فالتلميذ الذي يقدّم مستوى (ضعيفًا) في الأداء في مقاييس الفرز يحتاج إلى مزيد من المتابعة والتقييم لتقدمه التحصيلي وتكثيف التدريس إذا تطلب الأمر. وتسمح عملية المسح الشامل بمتابعة جميع التلاميذ وتتحقق من بقاء التلاميذ المجيدين على نفس المسار الصحيح، وهذا بطبيعة الحال يعكس أثر التدريس الجيد على أداء المتعلّمين ويتيح فرصة ملائمة للإدارات بتقييم الخدمات المقدمة لهؤلاء التلاميذ. وقد تساعد مقاييس المسح الشامل على فرز التلاميذ وتصنيفهم إذا كانوا ضمن إحدى الحالتين: الإيجابيات الحقيقية (True Positives) وتشمل التلاميذ الذين تمّ تصنيفهم بأنّهم معرّضون لخطر الفشل الدراسي وهم بالفعل كذلك، أو إذا كان المشخَّصون ضمن السلبيات الحقيقية (True Negatives)، وتشمل التلاميذ الذين تمّ تصنيفهم بأنّهم غير معرضين لخطر الفشل، وهم فعليًا ليسوا كذلك.

ولكن يجب التنويه إلى أنّ جميع مقاييس المسح الشامل قد تخطئ في فرز التلاميذ بإحدى الطريقتين: إما أن يقع بعض التلاميذ ضمن المحتاجين للمساعدة بينما هم ليسوا كذلك فعليًا، وهذا ما يُطلق عليه حالة الإيجابيات الخطأ (False Positives)، أو أن يصنف بعض التلاميذ بأنّهم ليسوا بحاجة للدعم والمساندة، بينما هم في حقيقة الأمر يحتاجون لذلك، ويُطلق على هذه الحالات حالة السلبيات الخطأ (False Negatives).

ولتلافي الوقوع في مثل هذه الأخطاء التشخيصية يضع التربويون في اعتبارهم أمرين مهمّين وهما: رفع أعداد التلاميذ المصنّفين بدقة بأنّهم معرّضون للخطر (At Risk Students) من خلال ارتفاع قيمة حساسية المقياس (Measure's Sensitivity). ورفع أعداد التلاميذ المصنّفين بدقة بأنّهم ليسوا ضمن المعرّضين للخطر من خلال ارتفاع قيمة النوعية للمقياس (Measure's Specificity) كما ورد في (Johnson, Jenkins & Petscher, 2010).

وتعتمد قيم الحساسية والنوعية للمقاييس على درجات القطع الفاصلة التي يُفرز التلاميذ بناءً عليها. فعندما تكون درجة القطع مرتفعة، أي أنّ جميع المتعلّمين الحاصلين على درجات تساويها أو أقل منها يتم تصنيفهم بأنّهم معرّضون للخطر، فهذا يشير إلى أنّ للمقياس قيمة مرتفعة للحساسية (High Sensitivity)، حيث سيكون التلاميذ المتعثرون فعلًا ضمن المصنّفين كمعرّضين للخطر. ولكن سيترتب على ذلك أن يكون هناك عدد من التلاميذ غير المحتاجين بالفعل إلى المساندة والدعم ضمن من صُنّفوا بأنّهم معرّضون للخطر، أي أن للمقياس نوعية منخفضة (Low Specificity) وبصورة مماثلة.

ويمكن القول إنّه عندما تكون درجة القطع منخفضة سيعني ذلك أنّ حساسية المقياس منخفضة، فعديد من التلاميذ المحتاجين فعلًا إلى الخدمات والدعم لن يكونوا ضمن المصنّفين كمعرّضين للخطر، فنوعية المقياس في هذه الحالة تكون ذات قيمة مرتفعة، حيث أنّ التلاميذ الذين ليسوا بحاجة للمساعدة فعلًا لن يقعوا ضمن التلاميذ المعرّضين للخطر. وهنا لا بدّ من الانتباه والوعي بطبيعة العلاقة بين الحساسية والنوعية في المقاييس، فينبغي للتربويين وصنّاع القرار الانتباه إلى أثر ارتفاع أو انخفاض درجات القطع في فرز التلاميذ وتصنيفهم، وبالتالي اتخاذ القرارات السليمة حول نوعية الخدمات ودرجة الدعم والمساندة بناءً على نتائج الفرز، كما أنّه لا بدّ من تحديد كمّ المصادر والإمكانيات ودرجة توافرها عند اتخاذ القرارات التربوية المختلفة ليتم استغلالها وتوظيفها بالشكل الأمثل.

الإجراءات المنهجية للدراسة

منهج الدراسة: استخدمت الدراسة الحالية المنهج الوصفي الذي عرّفه (عبد المؤمن، 2008، ص. 287) بأنّه «أحد أشكال التحليل والتفسير العلمي المنظّم لوصف ظاهرة أو مشكلة محددة وتصويرها كميًا، عن طريق جمع بيانات ومعلومات معيّنة عن الظاهرة أو المشكلة وتصنيفها وتحليلها وإخضاعها للدراسة الدقيقة».

مجتمع وعيّنة الدارسة

مجتمع الدراسة: يشمل مجتمع الدراسة جميع تلاميذ الصف الأول من الحلقة الأولى في التعليم الأساسي (من الصف الأول إلى الصف الرابع) في محافظة جنوب الشرقية خلال العام الدراسي 2017/2016 والبالغ عددهم 4621 تلميذًا وتلميذةً (دائرة الإحصاء والمؤشرات، 2016). ويقدّم الجدول 1 وصفًا لمجتمع العيّنة.

الجدول 1

مجتمع الدراسة وفقًا لمتغيّر النوع

النوع	الذكور	الإناث	المجموع
العدد	2309	2312	4621
النسبة	49.9%	50.1%	100%

العينة: تم اختيار العينة بطريقة عشوائية من تلاميذ الصف الأول الأساسيّ في محافظة جنوب الشرقية. وقد اختير مكان انتقاء العينة نظرًا لتيسُّر عملية جمع البيانات في تلك المنطقة الجغرافية من السلطنة. وقد أُخذت مجموعة من العوامل بعين الاعتبار عند اختيار المدارس التي طُبّقت فيها أدوات البحث وأهمّها: سهولة الوصول إلى المدارس المشاركة، وذلك لاستغلال الوقت بالطريقة المثلى؛ وتعاون الطاقم الإداري والهيئة التدريسية مع الباحثين، وتسهيل إجراءات التطبيق وتوفير المكان المناسب لذلك، والحصول على موافقة من إدارات المدارس لإجراء البحث.

وبناءً على هذه العوامل واستيفائها، اختيرت عيّنة استطلاعية من مدرسة للتعليم الأساسي في ولاية صور بسلطنة عمان، وهي من المدارس التي وافقت على تطبيق الدراسة. وقد اشتملت العينة الاستطلاعية على 44 تلميذًا وتلميذةً اختيروا بطريقة عشوائية من تلاميذ الصف الأول بالمدرسة. أما فيما يتعلّق بالعينة الأساسية فقد وقع الاختيار العشوائي على مدرستين

للتعليم الأساسي في ولاية صور بسلطنة عمان من المدارس التي وافقت على تطبيق الدراسة. واشتملت العينة على 149 تلميذًا وتلميذةً تمّ اختيارهم بطريقة عشوائية من تلاميذ الصف الأول من المدرستين. ويوضّح الجدول 2 توزيع العينتين طبقًا لمتغيّر النوع.

الجدول 2

عينتا الدراسة الاستطلاعية والأساسية وفقًا لمتغيّر النوع

	العيّنة الاستطلاعية		العيّنة الأساسية	
النوع	ذكور	إناث	ذكور	إناث
العدد	19	25	76	73
المجموع	44		149	

أدوات الدراسة

اشتملت الدراسة على إعداد وتصميم مجموعة من المقاييس القائمة على المنهاج بناءً على مراجعة الأدبيات والدراسات (Laracy et al., 2016; Purpura., 2015; Hompton et al., 2012; Lee et al., 2012; Methe et al., 2011; Hojnoski et al., 2009; Clarke et al., 2008 & Floyd et al., 2006) التي بحثت الخصائص السيكومترية والدقة التشخيصية لتلك المقاييس. وفيما يلي المقاييس التي تمّ إعدادها:

أولًا: مقاييس الاستعداد العددي المبكر القائمة على المنهاج ((Early Numeracy (CBM)

بعد مراجعة الأدبيات (Laracy et al., 2016; Purpura., 2015; Hompton et al., 2012; Lee et al., 2012; Methe et al., 2011; Hojnoski et al., 2009; Clarke et al., 2008 & Floyd et al., 2006) التي تناولت دراسة مقاييس ومؤشرات الاستعداد العددي المبكّر لدى تلاميذ رياض الأطفال والصف الأول في السلّم التعليمي، تم الاعتماد على ثلاثة مقاييس لها قدرة تصنيفية مناسبة لتقييم التطور في القدرات الرياضية المبكّرة والحس العددي للطفل في مراحله التعليمية الأولى، وهي:

- **مقاييس التعرف على العدد (NI):**

تقيّم هذه المقاييس قدرة الطفل على تعرّف الأعداد المطبوعة ضمن الأرقام من 0 إلى 99 التي تُدرّس خلال السنة الدراسية الأولى ضمن مقرر مادة الرياضيات للصف الأول من الحلقة الأولى في التعليم الأساسي. ويتطلّب تطبيق هذا المقياس تسجيل الإجابة الشفوية التي يقدّمها المفحوص أثناء العرض العشوائي للأعداد المطبوعة خلال دقيقة واحدة تمثل مدة التطبيق. وتشتمل نسخة التلميذ من مقياس تعرف العدد القائم على المنهاج (MN-CBM) على 84 عددًا مرتبًا بصورة عشوائية مطبوعة على ثلاث أوراق بيض ذات مقاس A4 ومقسّمة بواقع 28 عددًا في كل ورقة، وذلك في سبعة صفوف وأربعة أعمدة من المربّعات المنفصلة التي تحوي بداخلها الأعداد.

ومن خصائص تطبيق مقاييس الاستعداد العددي المبكر (EN-CBM) أن يتم تقديمها للمفحوصين بصورة فردية لتقييم القدرات التي يستهدفها كل مقياس منها، وبهذا يكون تطبيق مقياس تعرف العدد القائم على المنهاج فرديًا على تلاميذ الصف الأول من أفراد عينة الدراسة.

- **مقاييس العدد المفقود القائمة على المنهاج (MN):**

تستهدف هذه المقاييس قدرة المفحوص على إيجاد العدد المفقود ضمن سلسة مكوّنة من أربعة أعداد متتابعة بنمط ترتيبي معيّن، مطبوع منها ثلاثة أرقام وفراغ ( ) مما يستدعي الكشف عن النسق الترتيبي. ومن ثم تعرف العدد المكمّل لهذا النمط والذي يلزم وضعه في الفراغ ويقدم المفحوص استجابات شفوية خلال زمن التطبيق المحدد بدقيقة واحدة فقط. يبلغ

عدد المفردات في نسخة المفحوص 63 مفردة مقسمة بواقع 21 مفردة وسبعة صفوف وثلاثة أعمدة في كل صفحة من الصفحات الثلاث للمقياس. وتُقدّم المقاييس مطبوعة على أوراق بِيضٍ بحجم A4، وتطبع كل سلسلة عددية داخل مستطيل، ويكون التقييم فرديًا في هذا النوع من المقاييس شأنه كشأن مقاييس الاستعداد العددي المبكّر الأخرى.

- **مقاييس التمييز بين الكميات (QD):**

يتطلّب هذا المقياس من المفحوص أن يحدد أي عدد من بين زوج من الأعداد هو الأكبر كمًّا. وتُعتبر القدرة على تحديد المقادير التي تعبّر عنها الأعداد مكونًا أساسيًا للحس العددي (Number Sense) ومتطلبًا أوليًا كمهارة لإجراء العمليات الحسابية وحلّ المشكلات الرياضية. يقيّم هذا المقياس قدرة الطفل على تحديد العدد الأكبر من بين زوج من الأعداد المطبوعة، ويبلغ عدد المفردات في نسخة المفحوص 63 مفردة مقسّمة بواقع 21 مفردة وسبعة صفوف وثلاثة أعمدة في كل صفحة من الصفحات الثلاث للمقياس. وتُقدّم المقاييس مطبوعة على أوراق بِيض بحجم A4، ويطبع كل زوج من الأعداد عددية داخل مستطيل.

ثانيًا: قائمة تشخيص صعوبات التعلّم ((Learning Disabilities Diagnostic Inventory (LDDI)

استُخدمت قائمة تشخيص صعوبات التعلّم كاختبار محك لقياس الدقة التشخيصية لمقاييس الاستعداد العددي المبكّر القائمة على المنهاج. وقد بيّن إمام وكاظم والمحرزي والمخمري (2017) أنّ قائمة تشخيص صعوبات التعلّم تُعدّ أحد مقاييس التقدير، والتي صُمّمت لمساعدة الأخصائيين النفسيين والمختصين في مجالات صعوبات التعلّم والنطق والتخاطب وغيرهم من ذوي الاختصاص في جوهر عملية تشخيص الاضطرابات وصعوبات التعلّم لدى التلاميذ في المرحلة التعليمية من الصف الأول إلى الصف السادس.

وقُننتْ القائمة على عينة من التلاميذ ذوي صعوبات التعلّم والبالغ عددهم 2152 تلميذًا في سلطنة عمان. وتتسم القائمة بخصائص سيكومترية من الصدق والثبات، ويعتمد تفسير الدرجات فيها على المئينيات (Percentile) والتُساعيّات المعيارية (Stanines).

وتتألف القائمة من ستة مقاييس تقديرية مستقلة تقيّم المجالات الستة لصعوبات التعلّم كما تمّ تحديدها من قبل المكتب الأمريكي للتعليم (US Office of Education)، واللجنة الوطنية المشتركة (National Joint Committee). وتضم مجالات صعوبات التعلّم ما يلي: الاستماع (listening)، التحدث (Speaking)، القراءة (Reading)، الكتابة (Writ-ing)، الرياضيات (Mathematics)، والاستدلال (Reasoning).

يتألف كل مقياس مستقل من 15 مفردة سهلة التصنيف. وقد تمّ توليد هذه المفردات بعد مراجعة وتمحيص للأدبيات والبحوث الإمبريقية المنهجية في مجال صعوبات التعلّم وبشكل خاص تلك التي تناولت مجال النفس العصبي لصعوبات التعلّم. وقد تم التحقق من مدى ملائمة السلوكيات التي تعبّر عنها تلك المفردات من خلال دراسة قام بها مجموعة من الباحثين والخبراء في الميدان.

ولا تعد القائمة التشخيصية لصعوبات التعلّم (LDDI) مقياسًا كميًا للقدرة أو للإنجاز في المجال المُقيّم، ولكنّه يصف النمط الذي يظهره التلميذ في مهارة محددة في المجال المُستهدف (على سبيل المثال، القراءة والكتابة) والمترافقة مع صعوبات التعلّم المحددة (صعوبات القراءة (Dyslexia)، صعوبات الكتابة (Dysgraphia)).

ويساعد استخدام القائمة على تحول المسار التشخيصي للأفراد ذوي صعوبات التعلّم من مجرد إيجاد درجة التلميذ في اختبار محكي إلى دراسة نمط المهارة التي يمتلكها، لا سيما تلك المهارات ذات الصلة الوثيقة بصعوبات التعلّم النوعية (Specific Learning Disabilities). لقد قام (إمام وآخرون، 2017) بتكييف ومواءمة هذه القائمة من خلال ترجمته للغة العربية وتقنينه على البيئة العمانية.

وعدّل الباحثون بعضًا من المفردات لتتناسب مع خصائص اللغة العربية من حيث السياق اللغوي والصياغة والبناء كالمفردة 11 من مقياس الاستماع والمفردة 13 من مقياس الكتابة. كما أدرج المؤلفون أمثلة توضيحية في مواقع مختلفة من المقياس حتى يكون أكثر وضوحًا بين يدي المستفيدين من هذه القائمة، الأمر الذي زاد من صدق النسخة العربية من قائمة التشخيص لصعوبات التعلّم. وقد طُبّقت القائمة على عيّنة ممثّلة للمجتمع العماني بلغت 1250 من تلاميذ الحلقة الأولى من التعليم الأساسي من الصف الأول وحتى الصف السادس ممن تتراوح أعمارهم بين السبع سنوات إلى 12 سنة. وقد أظهرت النتائج تمتع المقياس بخصائص سيكومترية جيدة.

فمن ناحية الصدق، قام إمام وآخرون (2017) بإجراء تحليل عاملي توكيدي للقائمة للتأكد من صدق البناء وسلامة بنية المقياس، حيث اُختيرت ثلاثة نماذج وهي: نموذج العامل الواحد، ونموذج العوامل الستة المستقلة، ونموذج العوامل الستة المرتبطة. وأظهرت نتائج التحليل العاملي التوكيدي أنّ مؤشرات حسن المطابقة دعمت نموذج العوامل الستة المرتبطة. فقد كانت قيمة النسبة بين مربع كاي ودرجة الحرية أصغر القيم 2.934، وكانت قيمة جذر متوسط مربعات خطأ التقريب (RMSEA) أصغر القيم 0.046، وحققت قيمة مؤشر المطابقة المقارن (CFI) القيمة الكاملة 1.00، وحققت قيمة مؤشر حسن المطابقة (GFI) أكبر القيم 0.80. ومن ناحية الثبات، أشارت دراسة إمام وآخرين 2017 إلى تمتّع القائمة بقيم ثبات مرتفعة باستخدام طريقة الاتساق الداخلي. وجدت القيم التقديرية للثبات في المقاييس التقدير الفرعية جيدة في مقاييس: الاستماع 0.85 =α، التحدث 0.93=α، القراءة 0.91=α، الكتابة 0.92=α، الرياضيات 0.95=α، والاستدلال .α=0.94

الخصائص السيكو مترية لمقاييس الاستعداد العددي المبكّر القائمة على المنهاج

الصدق

تقدير صدق مقاييس الاستعداد العددي المبكّر القائمة على المنهاج: يبحث الصدق في مستوى أداء الفرد في الحالات التي تعتبر الاختبار عيّنة للسلوك المُراد قياسه. ويرتبط الصدق وفق ما ذكر عودة (1999) بالإجابة عن السؤال الآتي: إلى أي حد يكون الاختبار قادرًا على قياس مجال محدد من السلوك؟ ونظرًا لكون المهارات المستهدفة في مقاييس الاستعداد العددي المبكّر مهارات محددة وواضحة ومباشرة، فإنّ لهذا النوع من أنواع الصدق أهمية أكثر من غيره.

صدق المحتوى (Content Validity): يعرّف صدق المحتوى بأنّه «الصدق الذي يتم عن طريق إجراء تحليل منطقي لمواد القياس وفقراته وبنوده لتحديد مدى تمثيلها لموضوع القياس، أي أن يقوم الفاحص بفحص مضمون الاختبار فحصًا دقيقًا منظمًا». (مجيد، 2010)

مؤشرات صدق المحتوى:

الصدق العيني (Sampling Validity): تم اتباع مجموعة من الخطوات الأساسية لتحري صدق المحتوى وهي:

- التحديد الدقيق للمجال المعرفي الذي تقيسه مقاييس الاستعداد العددي المبكّر المُطبّقة في البحث، فهذه المقاييس تعكس فهم التلاميذ للقدرة العددية أو ما يُطلق عليه الحس العددي (Number Sense) في السنة الدراسية الأولى وما قبلها في رياض الأطفال (Clarke et al., 2008).
- تحليل وكتابة محتوى المادة المراد اختبار العينة بها. وهذا المحتوى يتمثل في محور الأعداد في منهج الرياضيات للصف الأول الأساسي.

- كتابة الأهداف المُراد تحقيقها وترتيبها حسب الأهمية والارتباط بمجال المقاييس القائمة على المنهاج، وتحديد وزن كل قسم منها وهو يمثل النسبة المئوية لعدد الصفحات التي تناولت الهدف.
- تحديد الوزن النسبي لمستويات الأهداف في المجال المعرفي: معرفة، وفهم، وتطبيق، وتحليل، وتركيب، وتقويم. قد تمّ فرز الأهداف المرتبطة بمجال تقييم مقاييس الاستعداد العددي المبكّر من أدلّة المعلّم للرياضيات (ستة أجزاء)، ومن الأجزاء الستة للكتاب المدرسي للتلميذ في الصف الأول الأساسي.
- تحديد وزن وأهمية كل قسم وذلك بضرب النسبة المئوية للقسم في النسبة المئوية لمستويات الأهداف.
- صياغة فقرات المقياس بحيث تتلاءم وتنسجم في عددها مع الأوزان النسبية المحسوبة كما ذكر في الخطوة السابقة.

ووفقًا لمعايير تطبيقات مقاييس الاستعداد العددي المبكر القائمة على المنهاج فإنّ طول الاختبار أي عدد الفقرات وزمن التطبيق وإجراءات التطبيق والتصحيح تُعدّ موحدة وثابتة.

وقد أشار (عودة، 1999) إلى أنّ جدول المواصفات يوفر درجة مقبولة من صدق محتوى الاختبار أو صدق تمثيل عينة الفقرات للأهداف إذا تمّ تقسيم الموضوعات وأوزانها والمستويات وأزانها على أسس منطقية.

الصدق الظاهري (Face Validity): تمّ إيجاد الصدق الظاهري من خلال حكم مختص على درجة قياس الاختبار للسمة المقاسة (عودة، 1999). ومن أجل الابتعاد عن الذاتية، تمّ عرض المحتوى والأهداف وفقرات الاختبار على لجنة مكونة من خمسة محكّمين من الأساتذة الجامعيين في جامعة السلطان قابوس من ذوي التخصصات المختلفة لمعرفة آرائهم في مدى مطابقة المحتوى والأهداف وفقرات المقياس لبعضها بعضًا. وقد بلغت نسبة الاتفاق بين المحكمين 100%.

الثبات

تقدير ثبات صدق مقاييس الاستعداد العددي المبكّر القائمة على المنهاج: تم تقدير الثبات الذي ذكر عمر وفخرو والسبيعي وتركي (2010) أنّه يسمح بمعرفة درجة الدقة التي في ضوئها تفسّر الدرجات التي يتمّ الحصول عليها بطرق مختلفة. ومن الطرق التي طُبّقت في هذه الدراسة ما يلي:

الاتساق الداخلي (Internal Consistency): تعتمد هذه الطريقة على استخدام كل من تباين فقرات الاختبار وتباين الاختبار ككل لتقدير الثبات وهي من الطرق الشائعة لتعيين الثبات حيث يتمّ حساب الثبات من خلال تطبيق نسخة من المقياس لمرة واحدة واستخدام الدرجات الجزئية أو الكلية على فقرات الاختبار في تلك النسخة لتقدير ثبات الاختبار. وتمّ استخدام معادلة ألفا كرونباخ (Cronbach's Alpha) حيث تمثل هذه المعادلة الصيغة العامة لحساب ثبات الاختبار من خلال التجانس الداخلي وتتمثل في اللجوء إلى عدد الفقرات ومجموع تباينات الفقرات المكوّنة للاختبار وكذلك تباين الاختبار ككل (محاسنة، 2013). وأشارت قيم المعاملات التي تتراوح ما بين 0.81 في مقياس العدد المفقود و0.97 في مقياس تعرف العدد إلى أنّها أعلى من مستوى العتبة المقبول (Acceptable Threshold) لمعاملات الثبات والذي يساوي 0.70 (Lee et al., 2012). إنّ أعلى القيم للمعاملات كانت لمقياس تعرف العدد (Number Identification)، فقد وُجد الثبات مستقرًا في جميع التطبيقات الثلاثة، وأقلّها كان لمقاييس العدد المفقود (Missing Number) في التطبيق الأول ثم بدأت معامل الثبات لهذا المقياس يزداد في التطبيقات الآتية بصورة جيدة.

وقد ذكر كلارك وشين (Clarke & Shinn, 2004) أنّه قد تمّ وضع معايير لتقييم معاملات الثبات للاختبارات والمقاييس وتحديد مدى القيم المقبولة تبعًا للغرض المقصود من تطبيق الأدوات في إطار اتخاذ القرارات التربوية. وتمّ تحديد هذه المعايير من قبل سالفيا ويسلدايك (Salvia & Ysseldyke)، وهي كالآتي:

- أن تكون قيم الثبات ≥ 0.90 لغرض اتخاذ قرارات تربوية تتعلّق بالحالات الفردية للتلاميذ (Individual Students).
- أن تكون قيم الثبات ≥ 0.80 لغرض اتخاذ قرارات تربوية ترتبط بالمسح (Screening) الشامل للتلاميذ.

- أن تكون قيم الثبات ≤ 0.60 لاتخاذ قرارات تربوية ترتبط بتدريس مجموعة من التلاميذ.

ونظرًا لاستخدام مقاييس الاستعداد العددي المبكّر (EN-CBM) في هذا البحث لأغراض المسح توافقت النتائج التي تمّ الحصول عليها لمعاملات الثبات مع المعايير العامة التي تعتمد عليها البحوث والدراسات كما ورد في النقطة الثانية من تسلسل تقييم معاملات الثبات وفق المعايير.

معامل ثبات استقرار التكافؤ (Stability Equivalence): أشار عودة (1999) إلى أنّ فكرة هذه المعامل تقوم على حساب الارتباط بين علامات عيّنة البحث على صورتين متكافئتين للاختبار مع وجود فاصل زمني طويل نسبيًا بين تطبيق الصورة الأولى والصورة الثانية. ويُطلق على ثبات الصور المتكافئة في المقاييس القائمة على المنهاج مسمى ثبات الصور البديلة (Alternate forms Reliability). فقد ذكر (عمر وفخرو والسبيعي وتركي، 2010) أنّه لا بدّ من وجود أكثر من صورة مستقلة للمقياس بحيث تشترك في المحتوى وتغطي نفس المدى من الصعوبة وبنفس العدد من المفردات، إضافة إلى التكافؤ في التعليمات وزمن التطبيق والأمثلة التوضيحية.

وتتسم صور مقاييس الاستعداد العددي المبكّر بأنّها ذات طبيعة متكافئة، حيث تمّ تصميم صورتين (أ و ب) لكل مقياس من المقاييس الثلاثة: تعرف العدد، والعدد المفقود، والتمييز بين الكميات لكل فترة من فترات التطبيق الثلاثة المحددة. فباختلاف المحتوى الذي يدرسه التلميذ في كل فترة كان لا بدّ من تصميم المقاييس لتناسب ما تمّ تدريسه عبر المقرر الدراسي. كما أنّ المبدأ الأساسي الذي تقوم عليه تطبيقات المقاييس القائمة على المنهاج هو: «أن نختبر ما تمّ تدريسه» هوسب وآخرين (Hosp et al., 2016)، وعليه قد استهدفت المقاييس التي طُبّقت في شهر فبراير الأعداد من 0 إلى 20 وأُجريت بدايةً على العيّنة الاستطلاعية ن=44 الصورة أ من كل مقياس، وبعد ثمانية أيام على الأقل من هذا الإجراء، تمّ اختبارهم في الصورة ب من كل مقياس، ومن ثم حساب معامل الثبات الذي يُطلق عليها في هذه الحالة مسمى معامل التكافؤ (Equivalence Coefficients) والذي تمثّلها معامل الارتباط بين درجات التلميذ في الصورتين أ وب من كل مقياس. تمثّلت الخطوة الآتية من تطبيق الفترة الأولى على اختبار العيّنة الأساسية ن= 149 في الصورة أ من كل مقياس.

فيما يتعلّق بالفترة الثانية من فترات التطبيق والتي كانت خلال شهر مارس، فقد كان المحتوى المستهدف تقييمه هو الأعداد من 0 إلى 49، وقد تمّ اتباع الخطوات بالترتيب المعمول به في التطبيق الأول. والأمر كذلك في التطبيق الأخير الذي تمّ في نهاية شهر إبريل إلا أنّ الفرق بينه وبين التطبيقات السابقة كان في المحتوى المستهدف الذي اشتمل على نطاق أوسع من الأعداد وهو من 0 إلى 99. وقد وُجد أنّ أعلى القيم لمعاملات الثبات كان لمقاييس تعرف العدد إذ بلغ أعلى قيمه خلال التطبيق الثالث في شهر إبريل. فبحساب معامل الارتباط بيرسون بين درجات العيّنة على الصوتين المكافئتين «أ» و«ب» وُجد أنّه يساوي 0.97 وهي قيمة مرتفعة، ويمكن الاعتماد عليها في اتخاذ القرارات المرتبطة بالحالات الفردية للتلاميذ.

كما تمت الإشارة سابقًا يجب تصنيف معاملات الثبات حسب الغرض والاستخدام الذي يستهدفه الباحث. كما أنّ أقل قيم لمعاملات الثبات تمتلكها مقاييس العدد المفقود في التطبيق الأول والتمييز بين الكميات في التطبيق الثاني، حيث بلغت هذه القيم 0.72 و0.73 على التوالي، إلا أنّه وبصورة عامة يمكن اعتبار أنّ قيم معاملات الثبات لهذه المقاييس تجعلها صالحة لاتخاذ القرارات التربوية المرتبطة بعملية المسح (Screening)، إذ إنّها غالبا ≤ 0.80.

نتائج الدراسة ومناقشتها

نتائج السؤال الأول: ما هي المعايير القياسية (Benchmark) لأداء تلاميذ الصف الأول الأساسي على مقاييس الاستعداد العددي المبكّر القائمة على المنهاج (EN-CBM)؟

تمّ إجراء مجموعة من الحسابات للتوصل إلى المعايير القياسية لأداء تلاميذ الصف الأول الأساسي في مقاييس الاستعداد

العددي المبكر المطبقة في البحث. إنّ الغرض الرئيس من تطبيق هذه الأدوات هو عملية المسح الشامل (Universal Screening) والذي يقضي بأن يتم تقييم جميع التلاميذ لثلاث مرات خلال العام الدراسي باستخدام أدوات تقييم بنفس مستواهم الصفي. وهذا التقييم كما أشار إليه (Hosp et al., 2016) يتمّ خلال الفصول الشتاء والربيع والخريف من العام الدراسي. وقد تمّ تعديل فترات التطبيق لتتلاءم مع ظروف التنفيذ والسير وفق المخطط الزمني لإجراءات الدراسة الحالية، وقد عُدّلت فترات التطبيق لتكون جميعها خلال الفصل الدراسي الثاني من العام الدراسي في أشهر فبراير ومارس وإبريل. وتكمن أهمية البيانات التي يتمّ جمعها في عملية المسح الشامل في أنّها:

- تحدد ما إذا كان المفحوص يعاني فعلًا من الفشل بحلول نهاية العام الدراسي.
- توضّح مدى تغطية الإجراءات والأساليب التدريسية لاحتياجات التلاميذ اللازمة من أجل اكتساب المهارات الأكاديمية المرجوة.

إنّ التحقق من بيانات عملية المسح هي ما يقود بصورة مباشرة إلى إيجاد المعايير القياسية لأداء المفحوصين. فقد أشار (Clarke et al., 2011 & Methe et al., 2011) إلى أنّ التلاميذ الذين يعانون من مشكلات وصعوبات في اكتساب مهارات الرياضيات هم الواقعون أسفل الدرجة المئينية 25 (25th. Percentile). وبناء على ذلك تمّ تحديد درجات القطع (Cut-off Scores) التي تمّ تعريفها من قِبل هامبلتون (Hambleton) كما ورد في (مجيد، 2007) بأنّها «نقطة على متصل درجات الاختبار تُستخدم لتصنيف التلاميذ إلى فئتين تعكس مستويات الأداء المختلفة بالنسبة لهدف معين أو الأهداف المراد قياسها في الاختبار». كما حددت دراسة (Laracy et al., 2016) التلاميذ المعرضين للخطر بأنّهم الواقعون أسفل الدرجة المئينية 40 (40^{th}. Percentile) أو أسفل الدرجة المئينية 25 (25^{th}. Percentile)، حيث حللت النتائج في عيّنة البحث عند هاتين الدرجتين حسب الفترات المختلفة للتطبيق.

من خلال الاتفاق بين الدراسات السابقة على فرز المفحوصين وفق الدرجة الفاصلة عند الدرجة المئينية 25 (25^{th}. Percentile)، فتمّ الاستناد إلى تحديد الدرجة الفاصلة (Cut-off Score) المقابلة للدرجة المئينية 25 (25^{th}. Per-centile) في كل فترة من فترات التطبيق المحددة للمقاييس الثلاثة.

وتمّ تحديد التلاميذ المعرضين لخطر الفشل الأكاديمي (At Risk Students) في مهارات الاستعداد العددي المبكّر حسب درجات القطع بأنّهم التلاميذ الحاصلين على درجة ≥ 23.50 في مقياس تعرف العدد، ودرجة ≥ 7 في مقياس العدد المفقود، ودرجة ≥ 10 في مقياس التمييز بين الكميات. كما أنّهم يمتلكون الحد الأدنى من المهارة مما يقتضي إلى اتخاذ أحد القرارات الآتية: إما أن يتم اجراء تغيير في عملية التدريس أو أن يتم إجراء مزيد من التقييمات التشخيصية وفقًا لاحتياجات المفحوص (Hosp et al., 2016). لقد حددت الدراسة المعايير القياسية لأداء تلاميذ الصف الأول الأساسي في مقاييس الاستعداد العددي المبكّر وفق البيانات التي سُجّلت في نهاية العام الدراسي. لقد صنف لاراسي وآخرون (Laracy et al., 2016) التلاميذ الواقعين أسفل الدرجة المئينية 25 (25th. Percentile) وأسفل الدرجة المئينية 40 (40th. Percentile) واعتبارهم معرضون لخطر الفشل في مهارات الاستعداد الرياضي المبكّر. وقد حُددت درجات القطع للمئين 25 في التطبيقين الثاني والثالث، وكذلك درجات القطع عند المئين 40 في التطبيقين الأول والثاني. وبأخذ الدرجات الفاصلة للتلاميذ الواقعين أسفل الدرجة المئينية 25 (25th. Percentile) في التطبيق الثالث اتضح أنّ درجة القطع لمقياس التمييز بين الكميات تساوي 23، وهي درجة أكبر من الدرجة الفاصلة التي حددتها الدراسة الحالية. كما أنّ درجة القطع لمقياس تسمية الأرقام (Number Naming) تساوي 11، علمًا بأنّ هذا المقياس يقيم القدرة على تعرف الأعداد المطبوعة على ورقة الاختبار شأنه في ذلك شأن مقياس تعرف الأعداد (Number Identification) المُطبّق في الدراسة الحالية. ويرى الباحثون أنّ اختلاف درجات القطع أمر مُتوقّع لعدة أسباب منها: اختلاف المفحوصين في بيئات التعلّم والخبرات التدريسية التي يتعرضون إليها وطرق التدريس واستراتيجياته المتباينة، بالإضافة إلى اختلاف نماذج التدخلات العلاجية المعمول بها لمساعدة التلاميذ ذوي الاحتياج للدعم والمساندة.

نتائج السؤال الثاني: ما هي مستويات الدقة التشخيصية: الحساسية (Sensitivity)، والخصوصية (Specificity) لمقاييس الاستعداد العددي المبكّر؟

تُعّد مقاييس الاستعداد العددي المبكّر (EN-CBM) من الاختبارات الممثلة (Representative Testing) أي أنّ السلوك المقاس بها يماثل أو يشابه إلى حد كبير السلوك المرتقب مثل درجة امتلاك التلميذ لمهارات الاستعداد العددي المبكّر (Early Numeracy Skills) وتحصيله الدراسي في الرياضيات في الصفوف اللاحقة. وقد أشار (عودة، 1999) إلى أنّ الاختبار الذي يعمل كممثل لا بدّ أن يعمل كمتنبئ (Predictor) والعكس ليس صحيحًا بالضرورة. ويتركز الاهتمام على الدقة التنبّئية التي تحددها معامل الارتباط بين الاختبار ومحك معيّن. وقد تمّ جمع البيانات لإيجاد العلاقة بين درجات التلاميذ في مقاييس الاستعداد العددي المبكّر وبين درجاتهم في مقياس التصنيف وهو القائمة التشخيصية لصعوبات التعلّم (LDDI)، والذي تمّ تطبيقه من قبل المعلّمات بالتزامن مع التطبيق الأخير للمقاييس خلال شهر إبريل 2017.

ويعد القيام بقياس وتحليل استجابة التلاميذ لأدوات التقييم المعدّة لأغراض المسح من الركائز الأساسية للعمل حسب مدخل الاستجابة للتدخل (RTI) في اتخاذ القرارات التدريسية. فمن خلال هذه الخطوة يتم تحديد التلاميذ الذين يظهرون مستويات أداء غير كافية، بمعنى آخر التلاميذ الذين يعانون من مشكلات في التمكّن من المهارات المقيّمة أو التلاميذ المعرضين لخطر الفشل في المهارة (At Risk Students) وبالتالي تقديم خدمات تدريسية إضافية لهم (Laracy et al., 2016).

وتتطلّب العملية التشخيصية للتلاميذ وحصر ذوي الاحتياج للمساعدة أن تكون أدوات التقييم (Assessment Measures) ذات دقة مرتفعة. ويمكن إجراء عديد من المعالجات والحسابات الإحصائية لتحديد الدقة (Accuracy) لهذه الأدوات والمقاييس. وتشمل هذه المعالجات إجراء تحليل الخصائص العملياتية للمستقبل (Receiver Operator Characteristic ROC) من أجل حساب الدقة التصنيفية (Classification Accuracy)، والحساسية (Sensitivity)، والخصوصية (Specificity)، والإيجابيات الحقيقية (True Positives)، والسلبيات الحقيقية (True Negatives)، الإيجابيات الخاطئة (False Positives)، والسلبيات الخاطئة (False Negatives)، (Johnson, Jenkins & Petscher, 2010).

يقع التلاميذ الذين تم تحديدهم بدقة (Accurately Identified Students) ضمن فئتين هما:

- الإيجابيات الحقيقية ((True Positives (TP): وهي الحالات التي صُنّفت بأنّها معرضة لخطر الفشل (At Risk) وهي بالفعل كذلك.
- السلبيات الحقيقية ((True Negative (TN): وهي الحالات التي صُنّفت بأنّها ليست معرضة لخطر الفشل (Not At Risk) وهي بالفعل كذلك.

بينما يقع التلاميذ الذين تم تحديدهم وفرزهم بصورة خاطئة (Misclassified Students) أيضًا ضمن فئتين:

- فئة السلبيات الخاطئة ((False Negatives (FN): هي الحالات التي صُنّفت بأنّها غير معرضة لخطر الفشل وفق أدوات المسح والتي أظهرت مستويات ضعيفة على المحك فيما بعد.
- فئة الإيجابيات الخاطئة ((False Positives (FP): هي الحالات التي صُنّفت بأنّها معرضة لخطر الفشل والتي أحرزت نتائج مرضية لاحقًا في المحك.

ومن أجل حساب دقة أدوات البحث في فرز التلاميذ المعرّضين لخطر الفشل، تمّ القيام بخطوات يمكن تلخيصها في النقاط الآتية:

- استخراج علامات التلاميذ في مقاييس الاستعداد العددي المبكّر: التعرف على العدد، العدد المفقود، التمييز بين الكميات، المُطبّقة خلال شهر أبريل أي علاماتهم في التطبيق الثالث.
- تصنيف التلاميذ حسب درجة القطع عند المئين 25 إلى فئتين: المعرضون للفشل (At Risk)، والعاديون (Normal) في كل من المقاييس الثلاثة على حدة.
- إعادة تعريف قيم الدرجات الخام بحيث يُعطى الرقم 1 للدرجات التي تكون < درجة القطع (Cut-off Score)، ويُعطى الرقم 0 للدرجات التي تكون ≤ درجة القطع.
- فرز التلاميذ حسب تصنيفهم تبعًا لقائمة تشخيص صعوبات التعلّم (LDDI) كما ورد في دليل التطبيق الخاص بالقائمة (إمام وآخرون، 2017)، والتي طُبّقت كمقياس للتصنيف لمقارنة نتائج الفرز المستخرجة مع نتائج الفرز عبر مقاييس الاستعداد العددي المبكّر.
- إعادة تعريف تصنيفات التلاميذ حسب مقياس التصنيف بحيث يُعطى الرقم 1 للمفحوصين الذين أشارت نتائجهم إلى أنّهم ذوو صعوبات تعلّم، ويُعطى الرقم 0 لمن هم غير ذلك.
- فرز علامات التلاميذ وفق تصنيفهم في الأداة التقييمية ومقياس التصنيف وتوزيعهم في الفئات الأربع في مصفوفة التصنيف (Confusion Matrix) التي تظهر الانتماء الفعلي للمفحوص حسب أدوات الفرز: مقاييس الاستعداد العددي المبكّر والانتماء المتنبئ به في مقياس التصنيف.
- استخراج خصائص الدقة من خلال المعادلات التي يوضّحها الجدول 3.

يوضّح الجدول 3 كيفية حساب نتائج الفرز تبعاً لمقياس الاستعداد العددي القائم على المنهاج وقائمة تشخيص صعوبات التعلّم.

الجدول 3

نتائج الفرز تبعًا لمقياس الاستعداد العددي المبكّر القائم على المنهاج ومقياس التصنيف (LDDI)

		مقياس الاستعداد العددي المبكّر	
		المعرضون لخطر الفشل (At Risk)	غير المعرضين لخطر الفشل (Not At Risk)
محك قائمة تشخيص صعوبات التعلّم (LDDI)	المعرضون لخطر الفشل (At Risk)	إيجابيات حقيقية TP	إيجابيات خاطئة FP
	غير المعرضين للفشل (Not At Risk)	سلبيات خاطئة FN	سلبيات حقيقية TN
الحساسية (Sensitivity)= TP/(TP+FN)			
النوعية (Specificity) = (TN/ (TN+ FP)			
الدقة التشخيصية (Classification Accuracy)= (TP+ TN) / (TP+ FP+ FP+ FN)			
القيمة التنبؤية الموجبة (Positive Predictive Value)= (TP/ (TP+ FP)			
القيمة التنبؤية السالبة (Negative Predictive Value)= (TN/ (TN+FN)			

أُجرى تحليل ROC لتحديد الدقة التشخيصية لأدوات البحث من خلال المساحة تحت المنحنى (AUC) لكل مقياس. ومن خلال التحليل، نتوصل إلى مجموعة من المعلومات: حيث تمثل الزاوية في أقصى اليمين إحداثيات (التنبيهات الخاطئة (False Alarms) والحساسية (Sensitivity))، وكلّما اتجه المنحنى إلى الزاوية في أقصى اليسار دلّ ذلك على أنّ للمقياس أو الأداة قدرة تمييزية مثالية، وكلّما ابتعد المنحنى عن الخط القُطْري دلّ ذلك أيضًا على قدرة تمييزية أفضل.

وذكر (Youngstrom, 2014) أنّه تمّ وضع معايير قياسية (Benchmarked Measures) لتقييم المساحة تحت المنحنى (AUC) من قِبل Swets كالآتي:

تُعتبر القيمة ≤ 0.90 ممتازة والقيمة ≤ 0.80 جيدة. أما القيمة ≤ 0.70 فهي مقبولة، كما أنّ القيمة ≥ 0.70 ضعيفة. لقد تراوحت قيم المساحة تحت المنحنى ما بين 0.81 و0.83. وتراوحت قيم الحساسية (Sensitivity) ما بين 0.80 لمقياس العدد المفقود و0.91 لمقياس التعرف على العدد، بينما تراوحت قيم النوعية ما بين 0.65 لمقياس العدد المفقود و0.70 لمقياس التمييز بين الكميات.

وبصورة عامة، فمن الضروري أن تمتلك الأدوات والمقاييس التي تهدف لتحري وفحص الحالات خصائص قوية لقيم النوعية (Specificity)، أما المقاييس التي يكون الغرض من استخدامها تقليص عدد الحالات السلبية فينبغي أن يكون لها قيم حساسية مرتفعة (Laracy et al., 2016) (Sensitivity).

وبحساب قيم القدرة التنبؤية الإيجابية (PPV)، فقد بلغت أعلى قيمها عند مقياس التمييز بين الكميات وتساوي 0.57. وتشير هذه القيمة إلى أنّ التلميذ الذي يقع أسفل الدرجة المعيارية في مقياس التمييز بين الكميات لديه احتمالية تصل إلى 57% أن يقع ضمن التلاميذ المصنّفين بأنّهم ذوو صعوبات تعلّم حسب المحك (LDDI). وأقل قيمة تنبؤيه موجبة بلغت 0.47 لمقياس العدد المفقود وتفسر كذلك بأنّ التلميذ الواقع أسفل العلامة المعيارية المئين 25 فإنّه من المرجح أنّ يقع ضمن التلاميذ ذوو صعوبات التعلّم في قائمة (LDDI) بنسبة 47%.

وفيما يتعلّق بقيم القدرة التنبئية السلبية (NPV) والتي تعبّر عن نسبة التنبؤ بالحالات السلبية الحقيقية من بين المصنّفين بأنّهم غير معرضين لفشل الرياضيات، فكانت قيمها مرتفعة. فبلغت أعلى قيمة لها 96.2% لمقياس التعرف على العدد، ووجدت مساوية لـ 89.9% لمقياسي العدد المفقود والتمييز بين الكميات. بلغت أعلى قيمة للدقة التشخيصية 74.5% لمقياس التمييز بين الكميات وتلتها الدقة التشخيصية لمقياس التعرف على العدد والتي بلغت 73.8%. وأخيرًا جاءت الدقة التشخيصية لمقياس العدد المفقود والتي بلغت 69.8%.

وتعبّر قيمة المساحة تحت المنحى للأدوات عن دقة جيدة حيث إنّها أعلى من الحد الأدنى المقبول 0.7 (Youngstrom, 2014). كما تعتبر المساحة تحت المنحنى مقياسًا عامًا للصدق التنبؤي (Predictive Validity) وتشير إلى أنّه عند الاختيار العشوائي للمفحوصين الحاصلين على درجات أعلى من الدرجة المعيارية في مقاييس المخرجات، فإنّهم سيحرزون درجات أعلى في الأداة المحكية، وهي أعلى من الدرجة التي سيُحرزها المفحوص الحاصل على علامة أقل من الدرجة المعيارية (Hanley & McNeil, 1982).

وقد كشفت دراسة (لارسي وآخرون (Laracy et al., 2016) أنّ قيم المساحة تحت المنحنى لجميع المقاييس المطبقة في الدراسة قد تراوحت بين 0.69 و0.82. ووُجدت أقوى هذه القيم عند استخدام مقياس (Quantity Comparison) للتنبؤ بأداء العينة في مقياس التمييز بين الكميات (Quantity Discrimination) فكانت القيم محصورة ما بين 0.77 و0.82. كما وُجدت خصائص الدقة للقدرة التنبؤية لأداء المفحوصين أسفل المئين 25 كالآتي: قيم الحساسية محصورة بين 0.44 و0.69، قيم النوعية محصورة بين 0.70 و0.91، القدرة التنبؤية الإيجابية (PPV) تقع ضمن المدى من 0.35 و0.60، والقدرة التنبؤية السلبية (NPV) تراوحت ما بين 0.83 و0.90. أما قيم الدقة التشخيصية فقد تراوحت ما بين 0.68 و0.81.

كما أشارت دراسة ميث وآخرين (Methe et al., 2011) إلى أنّ قيم المساحة تحت المنحى للمقاييس المطبقة على عينة المفحوصين من الصف الأول والمحكّات المعيارية المختلفة، قيمها أعلى من المستوى المقبول فقد تراوحت ما بين 0.72 و0.83.

وأجرت دراسة (Clarke et al., 2011) تحليلَ ROC على علامات العيّنة في المقاييس القائمة على المنهاج ومقياس (TerraNova-3). وتمّ استخدام المئين 25 كدرجة قطع. وظهرت خصائص الدقة التنبؤية من خلال التحليل عبر قيم الحساسية التي بلغت 0.96، والنوعية التي بلغت 0.81، وكذلك وُجدت الدقة التشخيصية مساوية للقيمة 0.83.

نتائج السؤال الثالث: هل توجد فروق دالة إحصائيًا عند مستوى دلالة α = 0.05 في أداء تلاميذ الصف الأول الأساسي في مقاييس الاستعداد العددي المبكّر القائمة على المنهاج تُعزى لمتغيّري النوع والالتحاق برياض الأطفال والتفاعل بينهما؟

أُجرى تحليل التباين الثنائي (Two-Way ANOVA)، والذي يفترض وجود متغيّرين (النوع والالتحاق برياض الأطفال) يؤثران على المتغيّر التابع: درجات المفحوصين في المقياس. فيتمّ اختبار تأثير كل متغيّر على حدة، وكذلك التفاعل بين هذين المتغيّرين إن وجد. تم القيام بإجراء تحليل التباين الثنائي باستخدام برنامج الحزم الإحصائية للعلوم الاجتماعية (SPSS) من خلال التحليل (Univariate) الموجود ضمن النموذج الخطي العام (General Linear Model). لقد كشف تحليل التباين لدى الثنائي (Two-Way ANOVA) بأنّه لا توجد فروق دالة إحصائيًا عند مستوى دلالة α= 0.05 على أداء المفحوصين في مقاييس الاستعداد العددي المبكر تُعزى لمتغيّري الالتحاق برياض الأطفال لما قبل المدرسة والنوع، أو التفاعل فيما بينهما.

ويرجّح الباحثون أن عدم وجود فروق ذات دلالة يعود لأسباب ترتبط بنوعية الخدمات التربوية التي تقدمها رياض الأطفال لمرتاديها. فهي مرحلة لا تقع ضمن السلّم التعليمي في سلطنة عمان. ونظرًا لأن هذه المرحلة تخلو من أية مناهج رسمية للأطفال في هذه المرحلة العمرية فإنه لا يمكن الجزم بأنّ هؤلاء المفحوصين قد تعرضوا لمهارات ترتبط بالقدرة العددية.

وقد أشار الزغلول (2009) إلى أنّ انتقال أثر التعلم يشير إلى الإفادة في الموقف التعليمي اللاحق مما تم تعلمه من الموقف التعليمي السابق. ويتأثر هذا الانتقال بعدة عوامل منها: الخبرة السابقة، والممارسة، والتشابه والتماثل في الخبرات والزمن اللازم لتعلم الخبرة والفاصل الزمني الذي يفصل تعلّم المهمة الثانية عن الأولى. ولم يقع ضمن العوامل السابقة عامل النوع كمؤثر على انتقال أثر التعلّم والذي يُقاس بأداء التلميذ على المقاييس المطبقة في هذه الدراسة لذلك لم يظهر فرق بين الذكور والإناث في الأداء والنتائج التي حصل عليها المفحوصون في مقاييس الاستعداد العددي المبكّر القائمة على المنهاج.

ويمكن تفسير نتائج الدراسة في ضوء اعتماد النجاح الأكاديمي للتلاميذ في التخصصات العلمية المختلفة كالعلوم والتكنولوجيا والهندسة وغيرها على تطوير مهارات أساسية في الرياضيات، حيث تُعّد مهارات القدرة العددية – أو التي يُطلق عليها مهارات الاستعداد العددي المبكر – أول مهارة يطورها الطفل بطريقة طبيعية تطورية خلال نموه، أو من خلال التعلّم المقصود الرسمي الذي يتلقاه في مؤسسات التعليم.

فالعلاقة وثيقة جدًا بين التمكّن من مهارات الاستعداد العددي المبكّر وتطوير مهارات الرياضيات اللاحقة بشكل عام. فمن الضروري أن يتمّ الكشف عن الصعوبات والمشكلات التي قد يتعرض لها الطفل في مهارات القدرة العددية على وجه

الخصوص في سبيل تعزيز التحصيل والإنجاز في الرياضيات في المراحل الدراسية فيما بعد (Leracy et al., 2016).

كما أنّ هناك ارتباطًا وثيقًا بين التقليل من وجود العجز في اكتساب مهارات الرياضيات في المرحلة الدراسية وبين أنظمة التقييم التي تستخدم أدوات دقيقة في تحديد التلاميذ الذين قد يعانون من مشكلات وتعثر في إحراز التقدم العلمي المطلوب. فكلما كانت الأدوات رصينة ودقيقة وذات خصائص تصنيفية مرتفعة كان عدد المخرجات التعليمية التي تعاني من مشكلات التعلّم أقل. ولهذا كان الاهتمام في هذه الدراسة موجهًا لتقديم أدوات تقييم بخصائص سيكومترية قوية، وذات دقة تصنيفية مرتفعة. ومن أجل تحقيق ذلك، تمّ تقييم مقاييس الاستعداد العددي المبكّر القائمة على المنهاج – التي صُمّمت واستُخدمت لأغراض المسح الشامل الذي يهدف إلى التعرف على الطلبة الذين يعانون بالفعل من مشكلات في اكتساب مهارات الاستعداد العددي المبكر – للعمل على تحديد أوجه المساعدة والدعم الممكن تقديمها في ضوء نقاط الضعف والاحتياج لدى هؤلاء الطلبة.

أظهرت التحليلات الإحصائية قيما مرتفعة للثبات. فكانت معامل كرونباخ ألفا بقيمة 0.97 في مقياس تعرّف العدد، 0.88 في مقياس العدد المفقود، و0.83 في مقياس التمييز بين الكميات. تُعدّ هذه القيم مناسبة جداً إذا كان الغرض من استخدام الأدوات هو عملية المسح الشامل أو الفرز.

وتُعدّ خصائص الدقة مكونًا رئيسيًا آخر في سبيل دعم أدوات التقييم والنظام التقييمي بصورة عامة، حيث أنّ ضعف الأدوات في خصائص الدقة (الحساسية والنوعية والدقة التشخيصية) يؤدي إلى أخطاء فادحة في التصنيف. وبالتالي في اتخاذ القرارات المرتبطة بذلك. فعلى سبيل المثال، إنّ أدوات التقييم ذات الحساسية المنخفضة (Low Sensitivity) تقود إلى أخطاء في تصنيف التلاميذ الذين يحتاجون حقًا للمساعدة. ويترتّب على ذلك حرمان عديد من المفحوصين المعرضين للخطر (At Risk Students) من الدعم والخدمات والتدريس المكثّف الذي يحتاجون إليه. وكذلك فإنّ أدوات التقييم ذات النوعية المنخفضة (Low Specificity) تزيد من درجة الخطورة لدى الحالات الإيجابية الخاطئة (FN) فتزيد بذلك من فرص تقديم خدمات وتدخلات علاجية لتلاميذ ليسوا بحاجة إليها في الواقع. لذا دعت هذه الاعتبارات لظهور توصيات لأن تكون قيم الحساسية والنوعية في أدوات التقييم من 90% فأعلى (Laracy et al., 2016).

وبحساب خصائص الدقة لمقاييس الاستعداد العددي المبكّر القائمة على المنهاج وقائمة تشخيص صعوبات التعلّم تراوحت قيم الحساسية ما بين 0.48 لمقياس تعرّف العدد و0.57 لمقياس التمييز بين الكميات؛ بينما تراوحت قيم النوعية ما بين 0.90 و0.96 لمقياسي العدد المفقود والتمييز بين الكميات. وبسبب قيم النوعية المرتفعة، فيمكن الاعتماد على هذه المقاييس في تحري الحالات الإيجابية وتمحيصها.

إنّ المقاييس التي تهدف إلى الكشف الدقيق عن الحالات الإيجابية – أي المفحوصين المعرضين للخطر– هي المقاييس التي تمتلك قيم عالية للنوعية (Specificity). أما المقاييس التي تهدف إلى تقليص عدد الحالات السالبة هي تلك المقاييس ذات القيم المرتفعة للحساسية (Sensitivity). وأخيرا يمكن القول أنّ أدوات ومقاييس الاستعداد العددي المبكّر القائمة على المنهاج، هي أدوات فاعلة في عملية التقييم للمفحوصين في الصف الأول الأساسي. فكما تبيّن أنّها ذات خصائص سيكومترية ملائمة من حيث الصدق والثبات لعملية التقييم لأغراض المسح الشامل. وقد اتضح أيضا امتلاكها لخصائص الدقة التشخيصية التي تمكّن المعلمين والتربويين والمهتمين من تطبيقها لفرز التلاميذ المعرضين لفشل الرياضيات.

التضمينات التربوية

في ضوء نتائج الدراسة يمكن القول بأنّها اشتملت على عدد من التضمينات التربوية، وهي التالية:

- من الممكن إدراج المقاييس القائمة على المنهاج ضمن أساليب التقييم المدرجة ضمن وثيقة تقويم تعلّم التلاميذ لمواد المجال الثاني وذلك من للصف الأول إلى الصف الرابع.
- يتمّ إعداد وتدريب المعلمين على العمل وفق معايير إعداد وتطبيق المقاييس القائمة على المنهاج ضمن برامج الإنماء المهني.
- يجب تنظيم دورات تدريبية لمعلمات المجال الثاني في الحلقة الأولى من التعليم الأساسي لإعداد مجموعة واسعة من المقاييس مهارات الاستعداد العددي المبكّر وتكوين بنك من أوراق العمل الجاهزة للتطبيق.
- من المهم توظيف هذه المقاييس وتفعيلها من قبل المعلمات في مدارس الحلقة الأولى على تلاميذ الصف الأول، فهي تكشف – بالفعل – عن مشكلات في مهارات مختلفة وهذا يمكّن المعلم من اتخاذ قرارات تدريسية تساعد التلاميذ وتدعمهم وتعزز تعلّمهم.
- تُستخدم مقاييس الاستعداد العددي المبكر كأدوات تقييم للتلاميذ المرشحين للإدراج ضمن برنامج صعوبات التعلّم. فتبعًا للنظام التشخيصي المعمول به في مدارس الحلقة الأولى من التعليم الأساسي في سلطنة عمان فإنّه يتمّ تشخيص تلاميذ الصف الأول في نهاية الفصل الثاني من العام الدراسي، وباستخدام نتائج مقاييس الاستعداد العدديّ المبكّر يمكن تقييم قدرة التلاميذ في المهارات الأساسية ووضع الأهداف التعليمية المناسبة لتدريسهم.

المراجع

إمام، محمود كاظم، علي المحرزي، راشد والمخمري، صالح. (2017). تقنين قائمة تشخيص صعوبات التعلّم النوعية لدى تلاميذ الصفوف من 6-1 في سلطنة عمان. **مجلة التربية الخاصة**، *5*(20)، 137-8. جامعة الزقازيق، كلية التربية.

إمام، محمود محمد المعمري، وطفة الشوربجي، سحر كاظم، علي مهدي المنذري، ريا حمدان، أحمد حسن القارسي جلال والمسكري، زيانة. (2013). التشخيص للعلاج أم العلاج للتشخيص: مدخل الاستجابة للعلاج في تشخيص الطلاب ذوي صعوبات القراءة في سلطنة عمان. **مجلة كلية التربية**، *2*، 52-27. جامعة بني سويف.

دائرة الإحصاء والمؤشرات. (2016). **الكتاب السنوي للإحصاءات التعليمية**. سلطنة عمان: وزارة التربية والتعليم.

الزغلول، عماد عبد الرحيم. (2009). **مبادئ علم النفس التربوي**. عمّان، دار المسيرة للنشر والتوزيع والطباعة.

الشحي، صفية بنت عبد الله. (2016). **الدقة التشخيصية لقياسات العمليات الحسابية القائمة على المنهج في فرز التلاميذ المعرّضين لصعوبات تعلُّم الرياضيات في الصف الرابع الأساسي**. رسالة ماجستير غير منشورة. سلطنة عمان، جامعة السلطان قابوس.

عبد المؤمن، علي معمر. (2008). **مناهج البحث في العلوم الاجتماعية الأساسيات والتقنيات والأساليب**. القاهرة، المجموعة العربية للتدريب والنشر.

عمر، محمود أحمد فخرو، حصّة عبدالرحمن السبيعي، تركي وتركي، آمنه عبد الله. (2010). **القياس النفسي والتربوي**. عمّان، دار المسيرة للنشر والتوزيع والطباعة.

عودة، أحمد. (1999). **القياس والتقويم في العملية التدريسية**. إربد، دار الأمل.

مجيد، سوسن شاكر. (2007). **أسس بناء الاختبارات والمقاييس النفسية والتربوية**. عمّان، ديبونو للطباعة والنشر والتوزيع.

مجيد، سوسن شاكر. (2010). **الاختبارات النفسية**. نماذج. عمّان، دار صفاء للنشر والتوزيع.

محاسنة، ابراهيم محمد. (2013). **القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة.** عمّان، دار جرير للنشر والتوزيع.

المديرية العامة للتقويم التربوي. (2015). **وثيقة تقويم تعلم التلاميذ مواد المجال الثاني للصفوف 4-1.** سلطنة عمان: وزارة التربية والتعليم.

Abu-Hamour, B., Al-Hmouz, H., & Kenana, M. (2013). The effect of short vowelization on curriculum-based measurement of reading fluency and comprehension in Arabic. *Australian Journal of Learning Difficulties, 18*(2), 181-197.

Brassard, M., & Boehm, A., (2007). *Preschool assessment: Principles and practices*. New York, The Guilford Press.

Clarke, B., & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33*(2), 234.

Clarke, B., Baker, S., Smolkowski, K., & Chard, D. J. (2008). An Analysis of Early Numeracy Curriculum-Based Measurement Examining the Role of Growth in Student Outcomes. *Remedial and Special Education, 29*(1), 46-57.

Clarke, B., Nese, J. F., Alonzo, J., Smith, J. L. M., Tindal, G., Kame'enui, E. J., & Baker, S. K. (2011). Classification accuracy of easy CBM first-grade mathematics measures: Findings and implications for the field. *Assessment for Effective Intervention, 36*(4), 243-255.

Dowker, A. (2005). Early identification and intervention for students with mathematics difficulties. *Journal of learning disabilities, 38*(4), 324-332.

Emam, M. M., & Kazem, A. M. (2016). Visual motor integration as a screener for responders and non-responders in preschool and early school years: implications for inclusive assessment in Oman. *International Journal of Inclusive Education, 20*(10), 1109-1121.

Emam, M., Al-Maamary, W., Mohamed, A., Kazem, A., (2012). *Development of an optimal framework for the identification and intervention of reading disabilities in key stage 1 school in Oman. A strategic project manuscript* (unpublished), Oman, Sultan Qaboos University.

Floyd, R. G., Hojnoski, R., & Key, J. (2006). Preliminary evidence of the technical adequacy of the preschool numeracy indicators. *School Psychology Review, 35*(4), 627.

Fuchs, L. S., & Fuchs, D. (2001). Principles for sustaining research-based practice in the schools: A case study. *Focus on Exceptional Children, 33*(6), 1.

Hampton, D. D., Lembke, E. S., Lee, Y. S., Pappas, S., Chiong, C., & Ginsburg, H. P. (2012). Technical adequacy of early numeracy curriculum-based progress monitoring measures for kindergarten and first-grade students. *Assessment for Effective Intervention. 37*(2) 118-126.

Hanley, J. A., & McNeil, B. J. (1982). The Meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*, 29-36.

Hojnoski, R. L., Silberglitt, B., & Floyd, R. G. (2009). Sensitivity to growth over time of the

preschool numeracy indicators with a sample of preschoolers in Head Start. *School Psychology Review, 38*(3), 402-418.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2016). *The ABCs of CBM: A practical guide to curriculum-based measurement*, 117-141. Guilford Publications.

Johnson, E. S., Jenkins, J. R., & Petscher, Y. (2010). Improving the accuracy of a direct route screening process. *Assessment for Effective Intervention, 35*(3), 131-140.

Laracy, S. D., Hojnoski, R. L., & Dever, B. V. (2016). Assessing the Classification Accuracy of Early Numeracy Curriculum-Based Measures Using Receiver Operating Characteristic Curve Analysis. *Assessment for Effective Intervention, 41*(3), 172-183.

Lee, Y. S., Lembke, E., Moore, D., Ginsburg, H. P., & Pappas, S. (2012). Item-level and construct evaluation of early numeracy curriculum-based measures. *Assessment for Effective Intervention, 37*(2), 107-117.

Methe, S. A., Begeny, J. C., & Leary, L. L. (2011). Development of conceptually focused early numeracy skill indicators. *Assessment for Effective Intervention, 36*(4), 230-242.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. US Department of Education.

Purpura, D. J., Reid, E. E., Eiland, M. D., & Baroody, A. J. (2015). Using a brief preschool early numeracy skills screener to identify young children with mathematics difficulties. *School Psychology Review, 44*(1), 41-59.

Riley-Heller, N., Kelly-Vance, L., & Shriver, M. (2005). Curriculum-based measurement: Generic vs. curriculum-dependent probes. *Journal of Applied School Psychology, 21*(1), 141-162.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education*, 1, 1-29.

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistic for pediatric psychology: We are ready ROC. *Journal of Pediatric Psychology, 39*, 204-221.