USAID
FROM THE AMERICAN PEOPLE

الطبعة الثانية

مجموعة أدوات تقييم القراءة في المراحل الدراسية الأولى

مارس 2016

تم إعداد هذا الإصدار ليراجع من قبل
الوكالة الأمريكية للتنمية الدولية.
وأعدته مؤسسة آر تي آي الدولية.

أداة تقييم القراءة في الصفوف الدراسية الأولى، الطبعة الثانية

صورة الغلاف: العاملين في مشروع آر تي آي، مبادرة التعليم بين الوكالة الأمريكية للتنمية الدولية ونيجيريا

أعدت هذا الإصدار مؤسسة آر تي آي الدولية لصالح الوكالة الأمريكية للتنمية الدولية، من خلال مشروع بيانات التعليم لصنع القرار (إيد داتا)، والهدف الاستراتيجي التعليمي 1 للقياس ودعم الأبحاث، أمر رقم - AID-OAA-12-BC-00003 (RTI Task 20).

آر تي آي الدولية.16 20. أداة تقييم القراءة للصفوف الدراسية الأولى، الطبعة الثانية، واشنطن دي سي، الوكالة الأمريكية للتنمية الدولية.

حقوق الطبع محفوظة © آر تي آي الدولية 2016

آر تي آى الدولية علامة مسجلة والاسم التجاري لمعهد أبحاث تراينجل.

هذا العمل خاضع للترخيص الدولي Creative Commons Attribution 4.0. للاطلاع على نسخة من هذا الترخيص، راجع الرابط:

http://creativecommons.org/licenses/by/4.0/. بموجب الترخيص، يمكن نسخ، توزيع، نقل، وتكييف هذا العمل بموجب الشروط التالية:

إعادة النشر – إذا قمت بنسخ وتوزيع هذا العمل بالكامل، دون أي تغييرات على المحتوى أو الأشكال التوضيحية، برجاء الإشارة إلى العمل كالتالي:

معاد إنتاجه على أساس العمل الأصلي المنشور من قبل مؤسسة آر تي آي الدولية ومرخص بموجب ترخيص Creative Commons Attribution 4.0 International License.

الترجمة – إذا قمت بترجمة هذا العمل، برجاء استخدام التوضيح التالي على إصدارك: **مترجم عن العمل الأصلي الذي نشرته مؤسسة آر تي آي الدولية ومرخص بموجب ترخيص Creative Commons Attribution 4.0 International License.**

التكييف – إذا قمت بتكييف هذا العمل، برجاء استخدام التوضيح التالي على إصدارك: **هذا تكييف للعمل الأصلي الذي نشرته مؤسسة آر تي آي الدولية ومرخص بموجب ترخيص Creative Commons Attribution 4.0 International License.**

شكر وعرفان

هذه الوثيقة تقدم نسخة معدلة من الطبعة الثانية من أدوات تقييم القراءة في المراحل الدراسية الأولى (EGRA) والتي أعدتها مؤسسة آر تي آي الدولية في مارس 2016 لصالح الوكالة الأمريكية للتنمية الدولية. ولم يكن بالإمكان إصدار هذه النسخة العربية من مجموعة الأدوات دون جهود السيد/ سمير حبيب الذي جمع المحتوى الفني من عدة فصول في النسخة الإنجليزية إلى دليل إرشادي ملائم لاحتياجات المستخدمين في السياقات الناطقة باللغة العربية. كما قدم السيد/ وائل صلاح مراجعات فنية وإفادات إضافية حول عدد من الأجزاء المعدلة والمترجمة. كما نتوجه بشكر خاص لشركة خدمات اللغات الأجنبية FLS لجهدها الرائع في ترجمة غالبية المحتوى الوارد في هذه الوثيقة.

هذه الأداة هي نتاج التعاون المستمر بين مجتمع ممتد من الباحثين، والمتخصصين، والمسؤولين الحكوميين، ومتخصصو التنمية التعليمية للنهوض بمسألة تقييم القراءة المبكرة والتحصيل بين أطفال المدارس الابتدائية في البلدان منخفضة الدخل.

بالرغم من صعوبة توجيه الشكر لجميع من ساهم في مراجعة وتحديث الأداة، فإنها ما كانت لتصبح ممكنة دون ميليسا تشيابيتا من إيه بي تي أسوشيتس، وراي آدامز وجوليت مندلوفيتز من المجلس الأسترالي لبحوث التعليم، وبوجا ريدي ناكامورا وزاكرو فوكميروفيتش من المعاهد الأمريكية للبحوث، وفتحي العشري من كريتيف آسوشيتس الدولية، وإلينا فينوجرادوفا من مركز تنمية التعليم، ومات سلون من ماثماتيكا لبحوث السياسات، وتوماز ألفاريز وعبد الله فردوس من نظم الإدارة الدولية، وروجر ستانتون من مجموعة أوبتيمال سوليوشنز، وكيلي بيتس، وكريس كاميسكي، ومارجريت دوبيك، وكارون هاردن، وسايمون كينج، وجيسيكا ميا، وإرين نيوتن، وأليسون بلبسين، وليلي بايبر، وسارة بوزيفارا، وجوناثان ستيرن من آر تي آي الدولية، وإليوت فريدلاندر وكارول دا سيلفا من هيئة إنقاذ الطفولة، وأجاليا زافيراكو من البنك الدولي. كما يحظى بالتقدير الخاص والشكر كل من تروفينا تشوتي من شركة أبحاث الجامعات، وجين بينبو من شركة أبحاث الجامعات، والعاملين بشبكة القراءة العالمية ممن ساهموا للمنظمة واستضافوا مختلف ورش عمل وندوات تقييم القراءة للصفوف الدراسية الأولى خلال 2015.

لم يكن إعداد الأداة الأصلية للتقييم ممكناً دون دعم المنظمات غير الحكومية وفرق تقييم القراءة في المراحل الأولى في وزارات التعليم في كل من أفغانستان، وبنجلاديش، ومصر، وجامبيا، وجيانا، وهايتي، وهندوراس، وجاميكا، وكينيا، وليبيريا، ومالي، ونيكاراجوا، والنيجر، وبيرو، والسنغال، وجنوب أفريقيا. ونتوجه بخالص الامتنان للمعلمين، والطلاب، وأسرهم لمشاركتهم وإيمانهم المستمر بمزايا التعليم. وبالمقابل، سنسعى جاهدين لتحسين نتائج القراءة لجميع الأطفال حول العالم.

تولت آمبر جوف مسؤولية التأليف الأولي للأداة الأصلية، مع مساهمات من لويس كراوتش، وإيمي مولكاي-ضن، ومارجريت كلارك. وقد أفادت الطبعة الثانية من مساهمات العديد من المشاركين في الندوات، والقادة، والمراجعين.[1]

الآراء الواردة في هذه الوثيقة تعبر عن مؤلفيها ولا تعكس بالضرورة آراء الوكالة الأمريكية للتنمية الدولية. برجاء توجيه الأسئلة أو التعليقات إلى بينيلوبي بيندر على البريد الإلكتروني pbender@usaid.gov.

[1] أسماء الأفراد والمنظمات المساهمة بشكل مباشر في الطبعة الثانية من الأداة موجه لها الشكر والعرفان في **الملحق** أ

جدول المحتويات

صفحة

شكر وعرفان iii
قائمة الأشكال التوضيحية x
قائمة المصطلحات xii
مصطلحات متعلقة بالقراءة xii
مصطلحات إحصائية xiv
مصطلحات منهجية xviii
1 مقدمة 1
1.1 لماذا نحتاج تقييم القراءة في المراحل الدراسة الأولى؟ 1
1.1.1 لماذا يجب تقييم **القراءة؟** 2
1.1.2 لماذا يجب التقييم **المبكر؟** 3
1.1.3 لماذا يجب التقييم **الشفهي؟** 5
1.1.4 موقع التقييم المبكر ضمن خيارات التقييم 6
1.2 إعداد أداة تقييم القراءة للمراحل الدراسية المبكرة 8
1.3 عمل الأداة 9
1.4 الأداة الأصلية والنسخة الثانية 10
2 أخلاقيات البحث والمراجعة الإلزامية من قبل مجلس المراجعة المؤسسي 13
2.1 ما هو مجلس المراجعة المؤسسي؟ 13
2.2 كيف تنطبق موافقة مجلس المراجعة المؤسسي على التقييم المبكر؟ 14
2.3 الموافقة والقبول الواعي للأفراد المشاركين 15
3 غرض التقييم واستخداماته 17
3.1 خلفية وعرض عام 17
3.2 التقييم المبكر كنظام تشخيصي 18
4 الإطار الفكري والأسس البحثية 22
4.1 ملخص المهارات اللازمة للقراءة الناجحة 22
4.2 الوعي الفونولوجي 23
4.2.1 الوصف 23
4.2.2 مقاييس الوعي الفونولوجي 25
4.3 أسس الهجاء والصوتيات وفك التشفير 25

4.3.1 الوصف 25
4.3.2 مقاييس المعرفة الهجائية ومهارات فك التشفير 28
4.4 المفردات واللغة الشفهية 30
4.4.1 الوصف 30
4.4.2 مقاييس المفردات 30
4.5 الطلاقة 31
4.5.1 الوصف 31
4.5.2 مقاييس الطلاقة 32
4.6 الفهم 33
4.6.1 الوصف 33
4.6.2 مقاييس فهم القراءة 33
5 تصميم دراسات التقييم 35
5.1 اعتبارات لتصميم دراسة التقييم 35
5.2 خيارات التصميم وفقاً للغرض من البحث 36
5.2.1 التقييمات الموجة وتقييمات الأداء كتصميمات بحثية 36
5.2.2 تقييمات الأثر كتصميم بحثي 37
6 تصميم أداة التقييم: إعداد التكييف وتعديله 42
6.1 ورشة عمل التكييف 42
عرض عام لاعتبارات تخطيط ورشة العمل 36
6.1.1 من يشارك؟ 44
6.1.2 ما هي المواد اللازمة؟ 45
6.1.3 ما المواد التي نحتاج إليها؟ 46
6.2 مراجعة المواد المشتركة للأداة 48
6.2.1 فهم الاستماع 50
6.2.2 التعرف على الحروف 52
6.2.3 قراءة الكلمات بلا معنى 57
6.2.4 طلاقة القراءة الشفهية 60
6.2.5 وعي الفهم الفونولوجي – التعرف على الأصوات الأولى والأخيرة أو تمييز أصوات الحروف 63
6.2.6 قراءة الكلمات المألوفة 68
6.3 مراجعة المكونات الإضافية للأداة 70
6.3.1 الإملاء 70
6.3.2 تقسيم المقاطع الصوتية 71
6.3.3 المتاهة وإغلاقها 73

6.4 أسباب عزل المكونات الأخرى للأداة 75
6.5 الترجمة وغيرها من الاعتبارات اللغوية 77
6.5.1 الترجمة مقابل التكييف 77
6.5.2 المقارنات عبر اللغات: إعدادات واعتبارات 78
6.6 استخدام أدوات من نفس اللغة عبر تطبيقات متعددة 81
6.7 أفضل الممارسات 82
7 استخدام تحصيل البيانات إلكترونياً 83
7.1 تحذيرات وقيود على تحصيل البيانات إلكترونياً 85
7.2 برامج تحصيل البيانات 86
7.3 اعتبارات اختيار الأجهزة وشرائها 86
7.4 الإمدادات اللازمة لتحصيل البيانات الإلكترونية والتدريب 87
8 تدريب المُقيم 88
8.1 تحديد المشاركين في التدريب 89
8.2 التخطيط لفعالية التدريب 91
8.3 مكونات تدريب المقيم 93
8.4 أساليب التدريب وأنشطته 93
8.5 زيارات للمدارس 95
8.6 عملية تقييم المقيم-المتدربس 98
8.7 قياس دقة المقيم 99
9 جمع البيانات الميداني: الاختبار التجريبي والدراسة الكاملة 104
9.1 إجراء التقييم بصورة تجريبية 104
9.1.1 بيانات الدراسة التجريبية ومتطلبات العينة 105
9.1.2 تحديد صلاحية الاختبار ومصداقيته 107
9.1.3 اعتبارات تتعلق بتوقيت الاختبار التجريبي 110
9.2 إجراءات الجمع الميداني للبيانات للدراسات الكاملة 111
9.3 اختيار الطلاب 115
9.3.1 الخيار الأول لعينات الطلاب: جدول العدد العشوائي 115
9.3.2 الخيار الثاني لعينات الطلاب: العينات المتقطعة 115
9.4 نهاية يوم التقييم: عرض موجز 119
9.5 رفع البيانات المحصلة ميدانياً 119
10 إعداد بيانات التقييم 121
10.1 تنقيح البيانات 121
10.2 تنفيذ المهام الفرعية للتقييم 123
10.2.1 <prefix>_ 124

10.2.2 <suffix> 125
10.3 المهام الفرعية المحددة بوقت زمني..... 126
10.4 المهام الفرعية غير المرتبطة بوقت زمني..... 127
10.5 المعادلة الإحصائية..... 128
10.6 إتاحة بيانات التقييم للجمهور 133
11 تحليل البيانات وإعداد التقارير..... 136
11.1 الإحصائيات الوصفية (غير الاستنباطية)..... 136
11.2 الإحصائيات الاستنباطية..... 137
11.3 أنواع التحليل الانحداري..... 139
11.4 تقارير تحليل البيانات 139
12 استخدام النتائج لإفادة الإجراءات..... 142
12.1 استراتيجية النشر..... 142
12.1.1 توصيل النتائج 143
12.1.2 مناهج النشر 146
12.2 وضع مقاييس خاصة بالبلدان..... 152
12.2.1 ما هي المقاييس؟ 154
12.2.2 معايير وضع المقاييس 156
12.2.3 عملية تحديد المقاييس 158
12.3 المحاذير والقيود 158
المراجع 161
ملحق أ: معلومات عن ورش تقييم مهارات القراءة للمراحل المبكرة 2015..... 176
ملحق ب: اعتبارات حجم العينة في تقييم القراءة في المراحل المبكرة 180
ملحق ج: المعاينة المركبة والعنقودية 188
ملحق د: المعاينة لتقييمات التأثيرات..... 190
ملحق هـ: تقييم الجودة التقنية للأداة..... 194
ملحق و: توصيات وإرشادات للمقارنات بين اللغات 197
ملحق ز: مقارنة برامج تحصيل البيانات..... 202
ملحق ح: مقارنة تعليمات التقييم الورقي مقابل الإلكتروني لتقييم القراءة في المراحل المبكرة 206
ملحق ط نموذج لجدول أعمال لتدريب أحد المقيمين 223
ملحق ى: تحليل البيانات والإرشاد الإحصائي لقياس دقة المقيمين 225
ملحق ك: خطط العينة لاختبار الثقة المتبادلة بين المقيمين على أساس ميداني..... 229
ملحق ل: نموذج لكتيب الأكواد 232
ملحق م: توصيات للمعادلة..... 234

ملحق ن: التوصيات الفنية المفصلة لملفات الاستخدام العام 238
ملحق س: تحليل بيانات التقييم 243
ملحق ع: أعراف الطلاقة في القراءة الشفهية للإنجليزية 246

قائمة الأشكال التوضيحية

صفحة

الشكل التوضيحي 1. مسارات القراءة للقراء المنخفضين والمتوسطين: القراءة الشفهية (مقاسة بالكلمات الصحيحة في الدقيقة) 3

الشكل التوضيحي 2. درجات الطالب للكلمات في الدقيقة، الصفين الأول والثاني 4

الشكل التوضيحي 3. الأنواع المختلفة للتقييمات: المتتاليات 7

الشكل التوضيحي 4. خريطة أماكن تطبيق تقييم القراءة للصفوف الأولى 9

الشكل التوضيحي 5. تطبيق أداة التقييم حول العالم: عدد البلدان، حسب السنة 10

الشكل التوضيحي 6. الدائرة المستمرة لتحسين تعلم الطلاب 19

الشكل التوضيحي 7. الاختلافات بين تنمية تكييف التقييم وتعديل التكييف 43

الشكل التوضيحي 8. عينة من جدول الأعمال: ورشة عمل تنمية تكييف التقييم أو تعديل التكييف 47

الشكل التوضيحي 9. استعراض المكونات الشائعة للأداة 50

الشكل التوضيحي 10. عينة: فهم الاستماع (باللغة الإنجليزية) 52

الشكل التوضيحي 11. حروف في اللغة الإنجليزية: شيوع الاستخدام 53

الشكل التوضيحي 12. عينة: التعرف على أصوات الحروف (لغة أسيبيمبا، زامبيا) 56

الشكل التوضيحي 13. عينة المهمة الفرعية للتعرف على أصوات الحروف (لغة ولوف، السنغال) .....??

الشكل التوضيحي 14. قراءة الكلمات بلا معنى (لغة أسيبيمبا، زامبيا) 59

الشكل التوضيحي 15. عينة: الطلاقة في القراءة الشفهية مع الفهم (الإنجليزية) 62

الشكل التوضيحي 16. عينة: الوعي الفونيمي – تحديد الصوت المبدئي (الإنجليزية) 65

الشكل التوضيحي 17. عينة: الوعي الفونيمي – تمييز أصوات الحروف (باهاسا، إندونيسيا) 67
الشكل التوضيحي 18. عينة: قراءة الكلمات المعتادة (البرتغالية، تيمور ليستي) 69
الشكل التوضيحي 19. عينة: الإملاء – كتابة الحروف (كريول، هايتي) 71
الشكل التوضيحي 20. عينة: تقسيم الفونيمات (البرتغالية، تيمور ليستي) 72
الشكل التوضيحي 21. عينة: متاهة (الإنجليزية، كينيا) 75
الشكل التوضيحي 22. إطار من فيديو مستخدم للتقييم..... 100
الشكل التوضيحي 23. بروتوكول العينات لمتابعة إمكانية الثقة أثناء العمل الميداني..... 102
الشكل التوضيحي 24. الاختلافات بين الاختبارات التجريبية للتقييم وتحصيل البيانات الكاملة 105
الشكل التوضيحي 25. معوقات مجموعات العينات 116
الشكل التوضيحي 26. قائمة تنقيح البيانات 122
الشكل التوضيحي 27. متغيرات المهام الفرعية للتقييم وأسماء المتغيرات للدرجات المرتبطة بوقت زمني محدد 124
الشكل التوضيحي 28. متغيرات الدرجات 126
الشكل التوضيحي 29. تصميم عينات التوازن المضاد 131
الشكل التوضيحي 30. إطار عمل للتواصل 144
الشكل التوضيحي 31. عرض عام للجماهير المحتملة 144

قائمة المصطلحات

مصطلحات مرتبطة بالقراءة:

الأبجدية: هي تحديد المبادىء التي تحكم تمثيل الحروف المكتوبة للأصوات المنطوقة بشكل منتظم ، والتي تُدمج معًا لتكوين كلمات لها معنى.

الدمج: قراءة مقطعين صوتين أو أكثر يكونان الكلمة ، ويبدأ المقطع الصوتي بصوت متحرك وساكن (مقطع قصير) ، أو صوتين متحركين وساكن (مقطع متوسط) ، أو ثلاث متحركات وساكن (مقطع طويل) ؛ وذلك لتكوين وقراءة كلمة صحيحة ، ولكن هذا يختلف عن الدايجراف digraph الذي يشير إلى حَرْفين مختلفين ينطقان صوتًا واحدًا . ؛ لأن الحروف تحافظ على أصواتها المميزة عند قراءتها ، وهذه الظاهرة صعبة الحدوث فى اللغة العربية.

المشتق (الاشتقاق): كلمة مشتقة من كلمة أخرى مثل: **اسم** الفاعل ،واسم المفعول ، وصيغة المبالغة ، واسم الزمان ، واسم المكان وغيرها ... .

علامات التشكيل (حركات الضبط): هي علامات أو حركات **توضع** فوق الحرف أو تحته ؛ لتحدد نطقه. (وتلك الحركات هى : الفتحة والضمة والكسرة، والتي تمثل الصوائت القصيرة بَ - بِ - بٌ).

الدايجراف : متوالية من حرفين : يشير إلى حَرْفين مختلفين ينطقان صوتًا واحدًا . ؛ لأن **الحروف** تحافظ على أصواتها المميزة عند قراءتها ، وهذه الظاهرة صعبة الحدوث فى اللغة العربية حيث إن الحروف تحافظ على أصواتها المميزة عند قراءتهاكما في اللغة الانجليزية (مثل ea في كلمة bread و ch كما في كلمة chin) لكن هذه الظاهرة لا توجد في اللغة العربية .

الطلاقة / الألية في القراءة : تمثل العلاقة الارتباطية بين فك رموز الكلمة (قراءة الكلمة) و الفهم ، وتشير الطلاقة إلى القدرة على قراءة الكلمات بسرعة و دقة قراءة تعبيرية.، و تساعد هذه المهارة القراء على توظيف المجهود الذهني في التوصل إلى المعني وترجمة الحروف أصوات ، و تحويل الأصوات إلي كلمات. وعند إتقان هذه المهارة ، يتمكن القراء من قراءة الكلمات بالسرعة الكافية التي تمكنهم من تركيز مجهودهم على فهم الكلمة والنص القرائى .

تحليل الطلاقة : هي مقياس للكفاءة القرائية عامةً ، و التي تعكس القدرة على القراءة بدقة و بسرعة .(انظر الطلاقة / الألية في القراءة).

الجرافيم: الوحدة الأساسية في النظام الأبجدي المكتوب والذي يمكنه تغيير معنى الكلمة ، وقد يتكون من حرف أو أكثر ، هذا و يمثل الجرافيم الفونيم (الصوت).و قد يتكون الجرافيم من حرف أو أكثر رُسمت عليه حركة غيرت من نطقها ودلالتها داخل الكلمة.أو من حرف و حركة مثل تـَ.

الصورة المتصرفة : تغير في بنية الكلمة وفقًا لتنوع السياق الثقافي للشخص ، والنوع الاجتماعي ، والزمن الواردة فيه الكلمة .

المورفيم : هو الوحدة الصرفية الدالة على معنى بحيث إن تغييرها يغير المعنى أو هو أصغر وحدة صوتية دالة على معنى ، وهي تختلف عن الكلمة ، حيث إن الكلمات يمكن تكوينها من أكثر من مورفيم. فمثلا كلمة قادمون تتكون من قادم و الواو و النون. وهناك مورفيم حر و مورفيم مقيد. المورفيم الحر: هو المورفيم الذي يقوم بنفسه دونما حاجة إلى غيره من المور فيمات، أي أنه يقوم وحده في الجملة، مثل: الأسماء، والأفعال، و المورفيم المقيد: هو المورفيم الذي لا يقوم بنفسه، إنما يكون جزءا من الكلمة، مثل: الضمائر المتصلة (كتبتُ).

بداية المقطع أو الوحدة الصوتية الأولى في المقطع: تشير إلى الصوت الأول في المقطع (الصامت) ، ثم يتبع بحركة .

علم الكتابة : التمثيل الكتابي لأصوات اللغة .

الفونيم: أصغر وحدة لغوية صوتية تغير معني الكلمة. مثال الكلمتان وَجَدَ و نَجَدَ متشابهتان ولكن تختلفان في صوت واحد، وهذا قد أدى إلى اختلاف المعنى .

الوعي الصوتي: هو إدراك أن الكلمات تتكون من أصوات، والقدرة على سماع و تعرف هذه الأصوات و إتقانها والتلاعب بها لتكوين مقاطع صوتية ثم كلمات جديدة .

المبدأ الأبجدي (الصوتيات): الممارسات التدريسية التي تؤكد على الطرق المنتظمة التي تربط بين الحروف المكتوبة والأصوات المنطوقة.

غلق المقطع: يشير إلى الوحدة الصوتية التي تأتي بعد بداية المقطع ، و قد تكون الحركة القصيرة التي تأتي بعد بداية المقطع أو الحركة الطوية أو الحركة الطويلة و صامت اللذان يشكلان معًا مقطعًا مغلقًا. وغلق المقطع هو جزء من المقطع الصوتى المكون للكلمة المراد قراءتها .

مصطلحات إحصائية :

أثر النتائج العالية : يحدث عندما يوجد حد أعلى ظاهري للقيم المحتملة لمتغير ما ، و يوجد تركيز كبير لنتائج المشاركين عنده أو بالقرب من هذا الحد. ومضاد هذا هو الحد الأدنى(أنظر أدناه) . على سبيل المثال، لو هناك مهمة فرعية من مهام تقييم مهارات القراءة EGRA سهلة جدًا بالنسبة لمعظم الأطفال ، فإن النتائج سوف تتركز بشدة عند نهاية المدى المسموح به ، ويحد من التنوع في الدرجات ، ويؤثر سلبًا على صدق الأداة نفسها.

إحصاء/ تعداد: عندما تشمل الدراسة كل أفراد المجتمع الأصلي (بمعنى أنه لم يتم تحديد عينة).

تحديد العينة العنقودية: هو أسلوب لتحديد العينة ، وذلك عند تقسيم المجتمع الأصلي إلى مجموعات (أو عناقيد) يتم تعيين العناقيد ، و بعد ذلك تُقيم كل البنود في العنقود (المجموعات) ، فعلى سبيل المثال، تُستخدم القائمة الكاملة لكل المدارس الابتدائية ؛ لتحديد عينة من 20 مدرسة، ثم تُقيم كل تلاميذ الصف الثالث الابتدائي في هذه المدارس المختارة .

العينة المختلطة/ العينة المركبة: أسلوب لاختيار العينة مشابه للعينة العنقودية (انظر أعلى)، و لكن تُحدد عينة من البنود الموجودة في الوحدات التي عينت. فعلى سبيل المثال، نستخدم قائمة كاملة بكل المدارس الابتدائية ؛ لاختيار عينة من 20 مدرسة، ثم تُأخذ عينة مكونة من 10 طلاب من الصف الثالث ، ويقيمون في المدارس المختارة.

فترة الثقة : مدى من القيم حول قيمة قيست من عينة تكشف درجة الدقة التي تعكس هذه القيمة . فالعينة قيمة المجتمع كله ، وتعكس فترة الثقة الأكبر دقة منخفضة. فعلى سبيل المثال ، إذا كان متوسط عمر عينة ما 36 عامًا ، فبالتالي، فترة الثقة الأصغر (من 35 إلي 37 عامًا) تقترح أن العمر المتوسط للعينة من الممكن أن يكون تقدير أكثر دقة للعمر المتوسط للمجتمع عنه إذا كانت فترة الثقة أكبر (تتراوح ما بين 34 إلي 38 على سبيل المثال).

العينة الملائمة: وتعرف أيضًا بالعينة المتاحة. فالعينة الملائمة هي عينة غير احتمالية تعتمد على جمع البيانات من أعضاء المجتمع الذين يسهل الوصول إليهم. (المتاحة بشكل ملائم) لا تسمح هذه الطريقة بالتعميمات ، و لها قيمة محدودة في أبحاث العلوم الاجتماعية.

أثر الحد الأدنى : يحدث عندما يوجد حد أدنى ظاهري للقيم المحتملة لمتغير ما ، و يوجد تركيز كبير لنتائج المشاركين عند أو بالقرب من هذا الحد. وهذا هو المقابل لأثر الحد الأعلى. (انظر أعلى). على سبيل المثال، إذا كان هناك مهمة فرعية في تقييم مهارات القراءة صعبة جدًا، فستتركز درجات التلاميذ بشدة عند الحد الأدنى للمدى المسموح (يتفق مع نسب عالية من درجة الصفر) ، و بالتالي تحد من التنوع في درجات التلاميذ ، و تؤثر سلبًا على صدق الأداة نفسها.

معامل ارتباط التصنيف المتداخل: هذا إحصاء وصفي يستخدم عندما تُجمع البيانات في عناقيد أو مجموعات ، و يتراوح بين 0 و 1 ، وهو يقدم قياسًا لدرجة التشابه بين أعضاء مجموعة ما في صفات قابلة للملاحظة. ويستخدم هذا الأسلوب في تحديد اتساق المقياس بين ملاحظين. من Fleiss (1981):

كا	شـدة
أ	ضـع
م	مـت
أ	ممـت

كابا : تقيس مدى حدوث تقديرين مختلفين لنفس العينة بالصدفة. وتتراوح قيم **كابا** بين-1.0 و 1.0. القيم الأعلى تبين احتمالية أقل للاتفاق العشوائي.

أقل أثر يمكن رصده: أقل أثر للمعالجة يمكن ملاحظته من البيانات بمعلومية حجم معين للعينة .

العينة غير المحتملة: إجراء لأخذ العينة ، و الذي تُحدد فيه العينة بدون استخدام نظرية الاحتمالية ، والأمثلة التي تتضمن الملائمة ،العينة المرجعية (كرة الثلج) ، و العينة الحصصية.

التقدير بنقطة : قيمة فردية أو حجم تأثير مشتق من بيانات العينة التي تعطي تقديرًا لقيمة أو حجم تأثير للمجتمع (انظر أسفل) ككل.

المجتمع: المجموعة النظرية للعينات (الأفراد و الوحدات) التي تعمم عليها نتائج الدراسة. فالعينة (انظر أسفل) و المجتمع يشتركان في خصائص متشابهة ، العينة جزء من مجتمع الدراسة.

تحليل القوة: يمكن استخدامه في حساب الحد الأدنى لحجم العينة المطلوب حتي يمكن رصد تأثير حجم معين، و يستخدم أيضًا في حساب الحد الأدنى لحجم التأثير الذي يحتمل رصده في دراسة باستخدام حجم عينة ما.

تقدير الدقة : حينما تسحب العديد من العينات من مجتمع ما، فإن تقدير الدقة يشير إلى مدى قرب تقديرات النقط (انظر أعلى) من العينات المختلفة إلى بعضها. كلَّما كانت تقديرات النقطة قريبة من بعضها كلَّما كان التقدير أكثر دقة.

العينة الاحتمالية : هذا مصطلح عام لكل العينات التي تُحدد وفقًا لنظرية الاحتمالات ، و تشمل طريقة الاختيار العشوائي ، والتي تتضمن الأمثلة الشائعة الاحتمالية المنتسبة إلي الحجم ، و العينة العشوائية البسيطة.

النزعة إلي تطابق الدرجات: يتضمن هذا الإجراء تطابق الملاحظات من المجموعة التجريبية و المجموعة الضابطة بخصوص الاحتمالية المقدرة للمشاركة في البرنامج (مع وجود مجموعة من الخصائص التي لوحظت). والغرض من هذا المدخل هو تحقيق التوازن بين المجموعة التجريبية و المجموعة الضابطة خاصة عندما لا تُستخدم التعيين العشوائي للمشاركة في البرنامج.

قطع الانحدار : تصميم شبه تجريبي للبحث يستخدم ؛ لتقدير أثار المعالجة في سياق غير تجريبي وذلك عن طريق استخدام نقطة قطع معينة. على سبيل المثال، إذا خُصص برنامج قراءة للطلاب الذين حصلوا على درجات أقل من 50في تقييم ما، فإن هذا المدخل سيركز على الطلاب الحاصلين على الدرجات الأقل والأعلى مباشرة من نقطة القطع (بناءً على الافتراض بأنهم طلاب متشابهين و لكن كل ما حدث هو أنهم عُينوا للمجموعة التجريبية – أولم يحدث – ، وذلك فقط بسبب العشوائية التقديرية لكونهم حصلوا على نقطة فوق أو أسفل نقطة القطع.

نسبة الاتفاق الخام: تقيس الي أي مدى يقدر المقيمون نفس القدير ، و نظرًا لقلة التفاصيل التي يقدمها هذا الأسلوب فقط، فلا يمكن وجود أي علامة مرجعية .وفي الوضع المثالي، فإن نسبة الاتفاق بين المقيمين ستكون في أعلى مستوى ممكن (قريبة من 100%) عند تقييم التلاميذ، ومع ذلك، وبغض النظر عن نسبة الاتفاق، لابد من الرجوع إلي إحصاء كابا لفهم جودة الاتفاق الاحصائي.

العينة: مجموعة العينة (الأفراد أو الوحدات) من مجتمع يُختارون للمشاركة دراسة.

إطار تحديد العينة : قائمة كل أفراد المجتمع التي يمكن الوصول إليها لدمجها في العينة.

وحدة تحديد العينة: أفراد العينة (انظر أعلى) : الوحدة التي تُجمع البيانات منها. على سبيل المثال، قد يكون الأفراد أو ربات المنزل وحدة تحديد العينة.

العينة العشوائية البسيطة: أبسط صور العينة الاحتمالية. فالعينة العشوائية البسيطة طريقة يكون فيها لدى كل فرد من المجتمع نفس درجة الاحتمالية أن يُختار للمشاركة في عينة الدراسة (انظر المصطلحات المائلة).

العينة المرجعية (كرة الثلج) : عينة غير احتمالية تُوظف المشاركين بشكل مبدئي في العينة لتقديم المعلومات اللازمة لتحديد مشاركين إضافيين في العينة.

القوة الإحصائية: احتمالية أن تجد الدراسة أثرًا للمعالجة مع العلم بوجود أثر للمعالجة سيُرصد .

الدلالة الإحصائية: احتمالية أن أثر المعالجة الذي وُجد في الدراسة ليس نتيجة الصدفة. كلَّما ازدادت الدلالة الاحصائية كلَّما زادت احتمالية كون أثر المعالجة الذي رُصد لا يرجع إلى الصدفة.

مصطلحات منهجية:

ثبات الاختبار: اتساق درجات التلاميذ الذين يحصلون عليها عند أخذ صورتين مختلفتين (طبق مرتين) ، ولكن شريطة أن يكونا متكافئتين من نفس الاختبار .

انحراف المقيم: ميل المقيم لتغيير ما يعد إجابة صحيحة أو مقبولة بمرور الوقت.

التآكل: الفقدان التدريجي لعينة الدراسة ، وغالبًا يحدث هذا في الدراسات التتابعية (انظر أسفل) عندما يتسرب أفراد العينة من الدراسة قبل استكمالها، على سبيل المثال ، بين دراسة الأساس ، ودراسة نصف المدة.

صدق المحتوى : يشير إلى درجة تمثيل البنود ؛ لقياس المهارات المستهدفة.
المجموعة الضابطة : العينة التي تُعين عشوائيًا ، والتي لا تخضع للمعالجة (التدخلات) ، و الذين تتشابه خصائصهم مع أفراد المجموعة التجريبية .

الصور المتكافئة للاختبار: الاختبارات التي يقصد الحكم عليها بالعلاقة مع كل صورة ، و هي مصممة بنفس المكونات ، و المهام الفرعية، إلخ...

مجموعة المقارنة: أفراد العينة الذين لا يتلقون المعالجة التجريبية (المداخلات) لكنهم مشابهون لأولئك الذين حصلوا على المعالجة ، والذين خصائصهم متشابهة مع أفراد المجموعة التجريبية الذين يتلقون المعالجة. (انظر أسفل).تُختار غالبًا على أساس تشابه خصائص معينة مع المجموعة التجريبية، و تسمى أيضًا مجموعة المقارنة المتناظرة.

الواقع المضاد: قياس ما كان سيحدث للمجموعة التجريبية في حالة عدم تلقي التدخل / المعالجة؛ ولأن الواقع المضاد غير معروف ، فهناك مجموعة متنوعة من الطرق الإحصائية التي تستخدم ؛ لمعرفة مجموعة الواقع المضاد التي تمثل ما كان سيحدث للمجموعة التجريبية إذا لم تكن قد حصلت على التدخل، وتُقارن المجموعة التجريبية مع الواقع المضاد للحصول على تقدير لأثر المعالجة.

التصميم ذو المقطع المستعرض: تصميم بحث يجمع البيانات من نفس العينة (انظر المصطلحات الإحصائية) مرة واحدة فقط، والتي يمكن مقارنة هذه التصميمات مع بيانات أخرى ، و التي تُجمع بشكل مستقل من نفس المجتمع على فترات مختلفة. ، وتقارن تحليل الاتجاه السائد لتلاميذ الصف الثالث في 2016 مع الصف الثالث في 2017.

صور الاختبار المتساوية : تشير إلى صور الاختبار التي تُطوع بواسطة عملية إحصائية ؛ لجعل الدرجات قابلة للمقارنة.

صور الاختبار المعادلة: اختبارات يستهدف أن تكون ذات صعوبة متساوية، و بالتالي يمكن استبدالها مباشرة بصورة مقابل صورة.

الصدق الظاهري: يشر إلى أي مدى يرى رأي الشخص في ما إذا كان الاختبار يغطي بشكل عام المفاهيم التي قصد مصممو الاختبار قياسها.
الدراسة التتابعية : تصميم للدراسة يتم فيه تتبع نفس وحدات العينة (انظر المصطلحات الاحصائية) على مدى فترة زمنية ، وتُجمع البيانات بشكل متكرر من نفس وحدات العينة.

تصميم اللقطة : نوع من الدراسات المقطعية (انظر أعلى) ، و التي تجمع البيانات فيه مرة واحدة ، ولا تجرى أي مقارنات على فترات زمنية.

الطبقية : عملية فصل الأعضاء أو أفراد العينة في مجتمع ما إلى مجموعات فرعية قبل أخذ العينة منهم. و يُستخدم هذا الأسلوب غالبًا في تحديد عينة ؛ لضمان حجم عينة كاف لكل مجموعة فرعية داخل المجتمع الأصلي .

المجموعة التجريبية: العينة التي تتلقى التدخل، و تسمى مجموعة التدخل.

1 مقدمة

1.1 لماذا نحتاج تقييم القراءة في المراحل الدراسة الأولى؟

تمكنت الدول حول العالم من زيادة معدل الالتحاق بالتعليم الابتدائي إلى معدلات لم تسبق في التاريخ. ويرجع الفضل إلى جهود " التعليم للجميع" (EFA)التي قادتها الأمم المتحدة وأهداف التنمية الألفية (MDGs) التي وصلت لقمتها في 2015. شهد العلم تحسنًا كبيرًا جدًا في معدلات الالتحاق بالمدارس الابتدائية. وصلت في بعض الأماكن إلى نفس المعدلات العالية في الدول المتقدمة ذات الدخل المرتفع. وصلت معدلات الالتحاق بالمدارس الابتدائية في الدول المناطق النامية حوالي 91%في 2015. وانخفض عدد الأطفال خارج المدرسة الابتدائية عالميًا إلي حوالي النصف في نفس الفترة الزمنية (الامم المتحدة ،2015) . صارت البيانات حول النتائج من الدول منخفضة الدخل والتي شاركت في تقييمات دولية تضمنت تطبيق اختبارات على تلاميذ الصفوف 1 – 3 من المرحلة الابتدائية متاحة الأن مقارنة نواتج التعلم على EdStats للبنك الدولي اليكترونيًا. ومع ذلك، تشير الأدلة إلى أن معدلات الالتحاق ازدادت ولكن مازال متوسط نواتج التعلم للطلاب منخفضًا في معظم الدول منخفضة الدخل. لخص حاليًا البنك الدولي الموقف ما يلي" هناك اتفاق في المجتمع الدولي أن تحقيق هدف التنمية الالفية للتعليم (MDG) يتطلب تحسينات في نواتج التعلم (البنك الدولي 2015b). تبنى الهدف 4 من أهداف ما بعد التنمية المستدامة 2015(برنامج التنمية للأمم المتحدة [UNDP، 2015) .

وهناك اتفاق واسع النطاق حول أهمية جودة التعليم بالنسبة للنمو والتطور الاقتصادي الوطني. "تكشف الأبحاث الحديثة أن التعلم و ليس عدد سنوات التحصيل الدراسي هو ما يسهم في النمو الاقتصادي للبلد حيث إن الزيادة بنسبة 10% في حصة التلاميذ من إتقان مهارات القراءة يُترجم إلى معدل نمو سنوي بنسبة 3% أعلى مما سيكون الوضع مختلفًا بالنسبة لهذا البلد. " (Hanushek & Woessman, 2009)كما في Gove & Wetterberg, 2011, pp. 1–2). وورد في النسخة الأولى من هذه الحقيبة في 2009 أن المقاييس الأكثر استخدامًا تمكنت من الكشف عما لم يتمكن التلاميذ في الدول منخفضة الدخل من معرفته ، ولكنها لم تتمكن من تأكيد ما عرفوه، وهذا يرجع غالبًا إلى أنهم حصلوا على درجات منخفضة

بحيث لم يتمكن من أن يحدد بدقة موقعهم على متصل المعرفة و المهارة. كما أن معظم التقييمات الوطنية و الدولية جرى تطبيقها تاريخيًا كاختبار ورقة و قلم على تلاميذ الصف الرابع الابتدائي و الصفوف الأعلى.

(وهذا يعني ، أنهم افترضوا أن التلاميذ كانوا قادرين على القراءة و الكتابة .) لم يكن ممكنًا دائمًا أن تخلص من نتائج هذه الاختبارات أن التلاميذ حصلوا على درجات متدنية لأنهم لا يمتلكون المعارف التي تقيسها هذه التقييمات أم أنه ليس لديهم المهارات الأساسية للقراءة والفهم. ومنذ عام 2010، ظهر اتجاه نحو تقييمات مهارات القراءة في الصفوف الأولى – و يرجع هذا بنسبة كبيرة إلى تأثير الوكالة الأمريكية للتنمية الدولية USAIDو البنك الدولي – و يشكل هذا علامة على تغير في الوعي لدى الباحثين في التعليم الدولي والمعنيين بخصوص الحاجة لمزيد من المعلومات من أبحاث تجريبية حول قدرة الأطفال الصغار على القراءة بفهم.

إن القدرة على قراءة وفهم نص بسيط هي واحدة من المهارات الأساسية التي يستطيع الطفل تعلمها. فبدون إتقان مهارات القراءة و الكتابة الأساسية ، تتضاءل فرصة الطفل للنجاة من الدائرة المتوارثة بين الأجيال من الفقر. ومع ذلك ، ففي كثير من البلدان لا يتمكن التلاميذ الملتحقون بالمدرسة في سن السادسة من قراءة وفهم نص بسيط. تشير الدلائل من الأبحاث إلى أن تعلم القراءة مبكرًا و بالقدر الكافي (مع التركيز على الفهم) أمر ضروري لتعلم القراءة جيدًا.

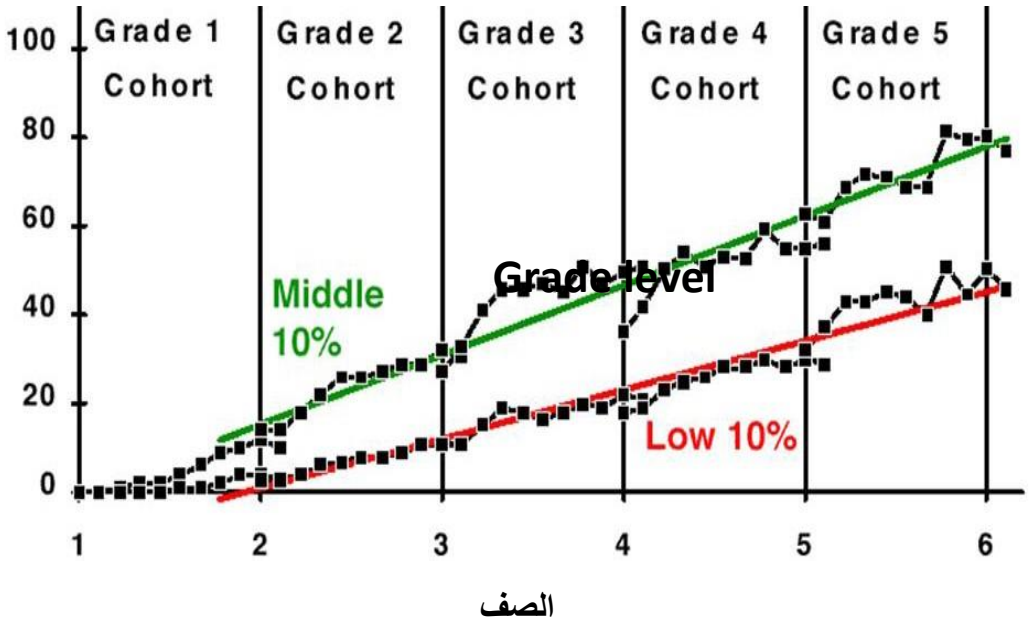
1.1.1 لماذا تقييم القراءة؟

تعتبر مهارات القراءة و الكتابة المبكرة الأساس الذي يحتاجه الأطفال ليكونوا ناجحين في كل المواد الدراسية في التعليم . يحتاج الأطفال أولًا إلى "تعلم القراءة " حتى يتمكنوا من " القراءة للتعلم". وهذا يعني كلما يتقدم التلاميذ في الصفوف المدرسية، يجري نقل المحتوى الدراسي لهم من خلال النص أكثر وأكثر، وتعتمد قدرتهم على اكتساب المعارف و المهارات الجديدة بشكل كبير على قدرتهم على القراءة واستخلاص المعني من النص. على سبيل المثال، الحساب مهارة مهمة، ولكن استخدام كتاب الحساب يحتاج إلى القدرة على القراءة. والتلاميذ مطالبون بشكل متزايد أن يظهروا تعلمهم من خلال الكتابة ،وهي مهارة مرتبطة بشكل تكاملي مع القراءة والفهم القرائي. وأيضًا وجود مستوى متدني من إجادة القراءة والكتابة يعيق بشدة قدرة الشخص على التعلم الذاتي و التعلم المستمر مدى الحياة الذي يمثل أهمية كبيرة خارج جدران حجرة الدراسة في عالم مسئوليات الراشدين.

1.1.2 لماذا التقييم المبكر؟

يصبح اكتساب مهارا القراءة والكتابة أكثر صعوبة كلما كبر التلاميذ، فالأطفال الذين لا يتعلمون القراءة في الصفوف الأولى القليلة من المحتمل أن يعيدوا نفس الصف الدراسي و في نهاية الأمر يتسربون من المدرسة. وهذا يعني إذا لم يكتسب التلاميذ المهارات الأساسية مبكرًا ستتسع الفجوات في نواتج التعلم أكثر بمرور الوقت (بين التلاميذ الذين اتقنوا المهارات الأساسية في القراءة و بين الذين لم يتقنوها بعد) (أنظر الشكل 1 و (see Adolf, Catts, & Lee, 2010; Daniel et al., 2006; Darney, Reinke, Herman, Stormont, & Ialongo, 2013; Scanlon, Gelzheiser, Vellutino, Schatschneider, & Sweeney, 2008; Torgesen, 2002). ترد نفس الصورة البلاغية التي توضح أن " الأغنياء يزدادون غنى والفقراء يزدادون فقرًا." في المناقشات التي تدور حول التباينات التي تحدث بين التلاميذ الذين يقرأون بطلاقة وبين الذين لا يقرأون بطلاقة وذلك من أجل الأطفال الغير قادرين على اكتساب مهارات القراءة و الفهم في الصفوف الأولى (Gove & Wetterberg, 2011).

الشكل 1: مسارات القراءة للقراء منخفضي ومتوسطي المستوى: جرى القراءة الشفوية من خلال عدد الكلمات الصحيحة في الدقيقة)

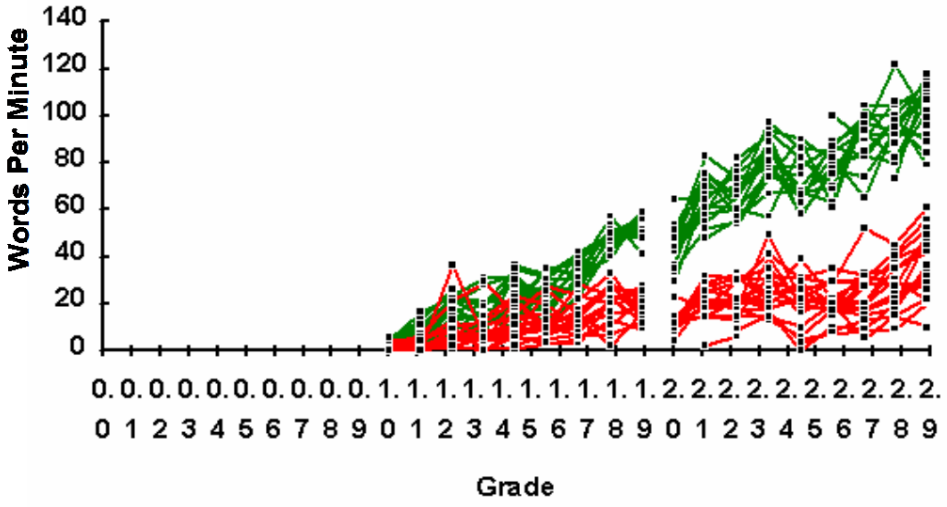


المصدر: Good, Simmons, & Smith, 1998, Figure 1.

على عكس العديد من المهارات مثل المشي و التحدث، فإن القدرة على القراءة لا تكتسب طبيعيًا بدون تدريس. تقترح الدراسات أنه بدون تدريس ذي جودة، فإن الطفل الذي يقرأ بشكل ضعيف في الصفوف الأولى سيستمر في القراءة بشكل ضعيف في الصفوف الأعلى، وهذا سيتطلب تدخلات للتدريس أكثر و أكثر، وذلك ليعوضوا وليلحقوا بالركب.(Juel, 1988) .

الشكل 2 يوثق مسار أداء التلميذ في الطلاقة في القراءة الشفوية لمجموعة من التلاميذ أثناء الصفين 1 و2 في الولايات المتحدة الأمريكية لم يتلقوا أي تدريس إضافي لتحسين القراءة. توضح مجموعة الخطوط في الجزء العلوي للجانب الأيسر من الرسم البياني نتائج التلاميذ الذين يستطيعون القراءة على الأقل عند 40 كلمة صحيحة في القيقة في نهاية الصف الأول. (كل وحدة على المحور الأفقي تمثل شهرًا في العام الدراسي.) كما هو موضح بالشكل 2، فإن الفجوة بين القراء الأكثر كفاءة و الأقل كفاءة تزداد بشدة عند نهاية الصف الثاني (الجانب الأيمن من الرسم البياني). فعند غياب التدخلات و المعالجة في الوقت المناسب ، فإن من المحتمل أن الفجوة الأولية في اكتساب القراءة تتسع و يصعب بشكل متزايد التغلب عليها.

الشكل 2: درجات التلاميذ بالصفين 1 و2 على الكلمات في الدقيقة



المصدر.(Good, Simmons, & Smith,1998, Figure 2 (grade 1) and Figure 3 (grade 2) :
ملحوظة: تشير الأرقام على المحور الأفقي إلى الصف (الصف الأغلى) و الشهر (الصف الأسفل)

كلما يعاني التلاميذ في المدرسة، كلما ازدادت احتمالية أن يُحبط التلاميذ ومن ثم يتسربون من

المدرسة فيخسرون أية فوائد ممكنة كان التعليم سيحققها فيما بعد في حياتهم. وعلى النقيض، كلما تعلم الأطفال بشكل أفضل، كلما بقوا في المدرسة لمدة أطول (Patrinos & Velez, 2009). توصلت إحدى الدراسات أن المؤشر الأقوى لتوقع إكمال المرحلة الابتدائية في السنغال كان مستوى نجاح التلميذ في القراءة في الصف الثاني (Glick & Sahn, 2010). فمن الفعال جدًا بالنسبة للأطفال كأفراد أو النظام التعليمي ككل أن يُعالج القصور في القراءة في الصفوف الأولى عنه في الصفوف الأعلى.

1.1.3 لماذا التقييم الشفوي؟

تتطلب الاختبارات التقليدية التحريرية (اختبار الورقة و القلم) من التلاميذ أن يكونوا قد اكتسبوا مهارات الطلاقة الأساسية في القراءة و الفهم القرائي. وإذا لم يكونوا قد اكتسبوا هذه المهارات (على سبيل المثال، إذا كانوا غير قادرين على قراءة السؤال أو كتابة الإجابة) ، فإن الاختبار لن يقيس بدقة ما يعرفه التلاميذ. و بلغة فنية، فإن النتائج ستعاني من الأثر الأدنى، مع وجود عدد كبير من التلاميذ الذين حصلوا على صفر. وفي مثل هذه الحالات، يخبرنا الاختبار التحريري (اختبار الورقة و القلم) فقط عما لا يعرفه الأطفال، و لكن ليس ما يعرفونه أو أين هم على مسار النمو.

وفي العديد من الدول، يتعين على التلاميذ اجتياز اختبار نهائي في نهاية الصف السادس لكي يحصلوا على شهادة اتمام المرحلة الابتدائية و؟ أو يلتحقوا بالمدرسة الثانوية (Braun & Kanjee, 2006). تُطبق أيضًا التقييمات الدولية في العديد من الدول مثل التقدم في دراسة القراءة الدولية و PIRLS(التي تُطبق على طلاب الصف الرابع الابتدائي) و برنامج التقييم الدولي للتلاميذ، أو PISA(التي تُطبق على الطلاب في سن 15 سنة ، و التي تطبق في الدول ذات الدخل المرتفع). وفي كلا نوعي التقييم، يُطلب من التلاميذ بشكل عام قراءة العديد من قطع القراءة و إجابة أسئلة اختيار من متعدد. إذا كانت مهارات التلاميذ في القراءة والفهم غير كافية لفهم الاختبار، فسيرسبون في هذا التقييم، ولكن البيانات الناتجة عن ذلك لن تكشف لماذا فشلوا. هل لم يملك التلاميذ المعرفة اللازمة للإجابة عن الأسئلة، أو مجرد أنهم كانوا غير قادرين على قراءة الأسئلة؟

تُعتبر الطلاقة في القراءة و الفهم من مهارات التفكير العليا في عملية اكتساب القراءة، وتقوم هذه المهارات على العديد من مهارات التفكير الدنيا و المهارات الأساسية مثل الوعي الفونولوجي و معرفة الأبجدية و قراءة الكلمات و المفردات، وغيرها و التي يمكن رصدها

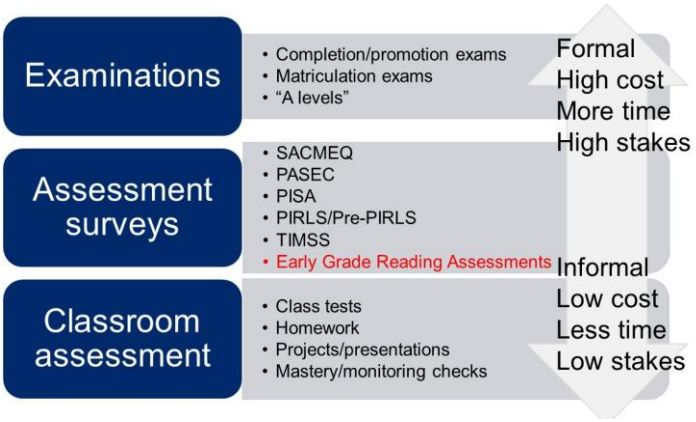
من خلال التقييم الشفوي بالتالي يستطيع التقييم الشفوي أن يمدنا من بمعلومات أكثر عما يعرفه

التلاميذ فعليًا و أين هم بالنسبة لعملية اكتساب القراءة في مراحل مبكرة تساعد أيضًا التقييمات الشفوية في رصد النمو المبكر بمعنى التغيرات التي لا يمكن رصدها بواسطة اختبار الورقة و القلم وهذا يشكل التقدم نحو اكتساب القراءة .

1.1.4 مكانة أداة تقييم مهارات القراءة في الصفوف الأولى (EGRA) بين خيارات التقييم

سيكون من المفيد أن نوضح الأنواع المختلفة للتقييم على متصل (كما في الشكل 3) وذلك لشرح أين تقع أداة تقييم مهارات القراءة في الصفوف الأولى EGRA بين خيارات التقييم. المتصل مقسم إلى ثلاثة أقسام هي : الامتحانات و دراسات التقييم و التقييمات التي تتم على مستوى حجرة الدراسة. يعرف Kanjeeكانجي (2009) الامتحانات على أنها العمليات المستخدمة لاختبار مؤهلات المرشحين (على سبيل المثال، الاختبارات الربع سنوية و اختبارات الترقي و اختبارات الالتحاق). وتُعد هذه الاختبارات طويلة وذات درجة رسمية عالية، وهي أيضًا تقييمات مقننة تُطبق على كل التلاميذ (وهذا يجعلها تستغرق وقتًا أكثر وعالية الكلفة). وعلى الناحية الأخرى من المدى، توجد التقييمات التي تتم على مستوى حجرة الدراسة، والتي يمكن تعريفها على أنها مقاييس تُستخدم للحصول على أدلة على معارف و مهارات و اتجاهات الأفراد من المتعلمين بغرض تزويد عمليتي التدريس و التعلم بمعلومات و تحسينها (Kanjee, 2009). وهذه التقييمات التي تتمتع بدرجة أقل من الرسمية تأتي في عدة صور مثل الاختبارات والتكليفات المنزلية و و المشروعات و العروض التقديمية. و يُقصد من التقييمات التي تتم على مستوى حجرة الدراسة وفقًا لتصميمها أن تكون قليلة التكلفة وأن تستغرق وقتًا أقل ، وأن تنطوي على مخاطر أقل (خاصةً عند مقارنتها بالامتحانات) .

الشكل 3: أنواع التقييمات متصل



المصدر: معدل من Kanjee (2009).

تصمم دراسات التقييم ولها غرض واحد هو الحصول على معلومات عن أداء التلاميذ ، وأيضًا عن النظام التعليمي ككل. يوجد بالإضافة إلى PIRLS و PISA العديد من التقييمات الدولية و الإقليمية الأخرى التي تتماشى مع هذا التصنيف مثل هذه التقييمات التي تنفذها Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), the Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC), the Laboratorio Latinoamericano de Evaluación de laCalidad de la Educación (LLECE), و the Trends in International Mathematics and Science Study (TIMSS). وهذه الاختبارات المرتبطة بهذه البرامج تتطلب عمليات تطوير طويلة الأمد وتعقيدات لغوية محلية و وإجراءات معقدة في تقدير الدرجات و والمعايرة حيث يُقصد من هذه الاختبارات قياس الاتجاهات العامة في التحصيل في إجادة القراءة و الكتابة وذلك من أجل عقد المقارنات بين الدول. بالإضافة إلى ذلك، يتطلب كل نوع من هذه التقييمات القدرة على القراءة الأساسية (على سبيل المثال ، يعتمد التقييم على قراءة قطعة) و التي قد تحد من ملائمة الاختبار وفائدته لقياس مهارات القراءة في الصفوف الأولى في الدول النامية (و ذلك بسبب أثر الحد الأدنى) و قد شهدت السنوات الأخيرة تطورًا في التقييمات الجديدة لمهارات القراءة في الصفوف الأولى لسد هذه الفجوة Pratham's Annual Status of Education (ASER) Report و أيضًا World Vision's Functional Literacy Assessment Tool [FLAT] . وعُرف عن هذه التقييمات أنها أصغر وأسرع وأقل تكلفة عند مقارنتها بالاختبارات الدولية (Wagner, 2011).

2.1 تطوير أداة تقييم مهارات القراءة في الصفوف الأولى EGRA:

وفي سياق الإجابة عن هذه الأسئلة حول تعلم التلاميذ و الاستثمارات المستمرة في التعليم للجميع، دعت أقسام التربية والمختصون في التنمية في البنك الدولي و الوكالة الأمريكية للتنمية الدولية و المؤسسات الأخرى إلى ابتكار مقاييس لنواتج قراءة التلاميذ تكون بسيطة و فعالة و قليلة الكلفة (Abadzi, 2006; Center for Global Development, 2006; Chabbott, 2006; World Bank: Independent Evaluation Group, 2006).

و بدأ العمل في تطوير أداة تقييم مهارات القراءة في الصفوف الأولى استجابةً لهذا الاحتياج و ضرورة وجود مقياس اكتساب القراءة في الصفوف الأولى شريطة أن يكون قليل الكلفة و فعالًا. فهي أداة بسيطة يمكنها تقديم تقارير عن المستويات الأساسية لتعلم التلاميذ و تتضمن تقييمًا للخطوات الأولى التي يأخذها اطلاب في تعلم القراءة. و تعاقدت الوكالة الأمريكية للتنمية الدولية USAIDمع معهد تراينجل للأبحاث RTI لتطوير أداة لمساعدة البلاد الشريكة مع الوكالة الأمريكية للتنمية الدولية في عملية قياس كيف يكتسب الأطفال في الصفوف الأولى مهارات القراءة وذلك القياس يكون بطريقة دقيقة و منتظمة. وفي نهاية الأمر ، كان الغرض هو تشجيع المزيد من الجهود لتحسين الأداء في هذه المهارات الأساسية باستخدام عملية تقييم يمكن تطويعها بسهولة لتلائم سياقات ولغات جديدة ، و نظام بسيط لتقدير الدرجات ، وتستخدم لقياس نواتج التعلم للتلاميذ، و لا تستغرق وقتًا كبيرا بالنسبة للأفراد الذين يُجرى عليهم التقييم.

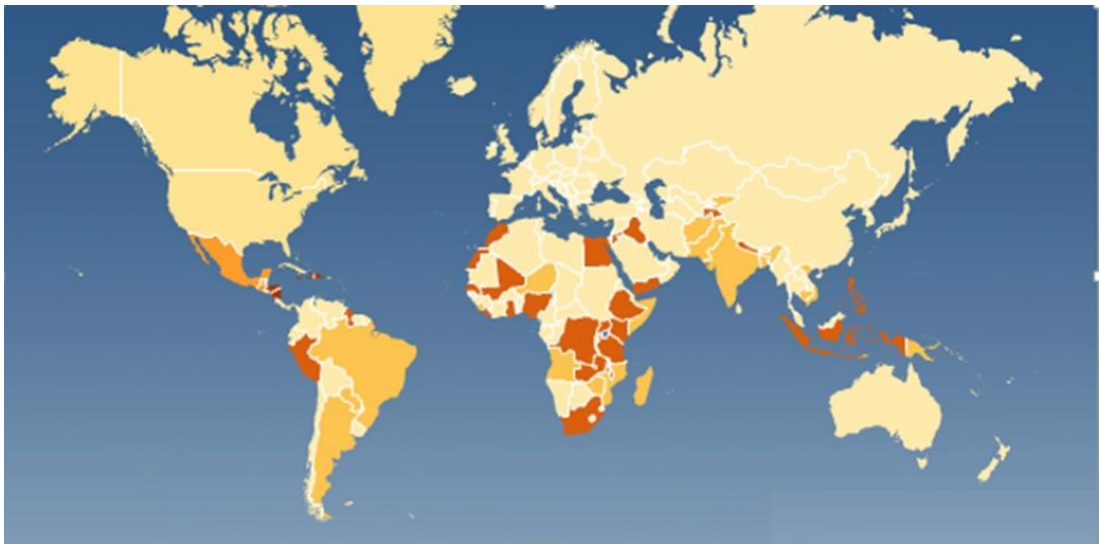
قامت منظمة معهد تراينجل للأبحاث بناء على مراجعة للأبحاث و تقييمات وأدوات القراءة بتطوير برتوكول لتقييم مهارات القراءة الأساسية لدي التلاميذ يُجرى شفويًا و بشكل فردي. قام علماء في المجال المعرفي و خبراء في مجال القراءة المبكرة و متخصصون في تصميم الأبحاث و خبراء في مجال التقييم بمراجعة الأداة المقترحة و قدموا تغذية راجعة و مصادقة على بروتوكول التقييم و صدق هذا الأسلوب في ورشة عمل مبدئية استضافتها الوكالة الأمريكية للتنمية الدولية و البنك الدولي ومنظمة معهد تراينجل للأبحاث في نوفمبر 2006. تضمنت ورشة العمل مشاركات من أكثر من اثنى عشر خبير من مجموعة متنوعة من الدول بالإضافة إلى 15 ملاحظًا من مؤسسات مثل الوكالة الأمريكية للتنمية الدولية و البنك الدولي مؤسسة وليام و فلورا هوليت و جامعة جورج واشنطن ووزارة التعليم من جنوب أفريقيا و مؤسسة بلان و أخرين. أداة تقييم مهارات القراءة EGRA أداة متاحة وجاهزة لدعم المستويات العليا ونشر للمعارف عن نواتج التعلم و القراءة. كان الغرض من هذا القرار هو

ضمان أن الجمهور من المتخصصين وغير المتخصصين في المجال يمكنهم الحصول على بيانات دقيقة وفي حينها و سهل الوصول إليها حول القراءة و الحساب في الصفوف الأولى في السياق الذي يعيشون فيه، و يكونوا قادرين على صنع القرارات ووضع السياسات.

3.1 الأداة تعمل

دعم البنك الدولي دراسة استطلاعية عام 2007 في السنغال للأداة المسودة(اللغة الفرنسية و الولوف) و جامبيا (اللغة الانجليزية) ، بينما دعمت الوكالة الأمريكية للتنمية الدولية دراسة استطلاعية في نيكاراجوا (اللغة الإسبانية). وعقب هذه الدراسات الاستطلاعية اتسع استخدام أداة تقييم مهارات القراءة في الصفوف الأولى من خلال العديد من الممولين والمنفذين للمشروعات و الدول و اللغات. أصبحت الوكالة الأمريكية للتنمية الدولية واحة من أكبر الرعاة لتطبيقات أداة تقييم مهارات القراءة في الصفوف الأولى من خلال عقد بيانات التعليم 2 EdData II .دعم بيانات التعليم EdData II وحده دراسة أداة تقييم مهارات القراءة في الصفوف الأولى في 23 دولة و 36 لغة في الفترة ما بين 206 و 2015(**انظر الشكل 4**)

الشكل 4 خريطة تطبيق أداة تقييم مهارات القراءة في الصفوف الأولى

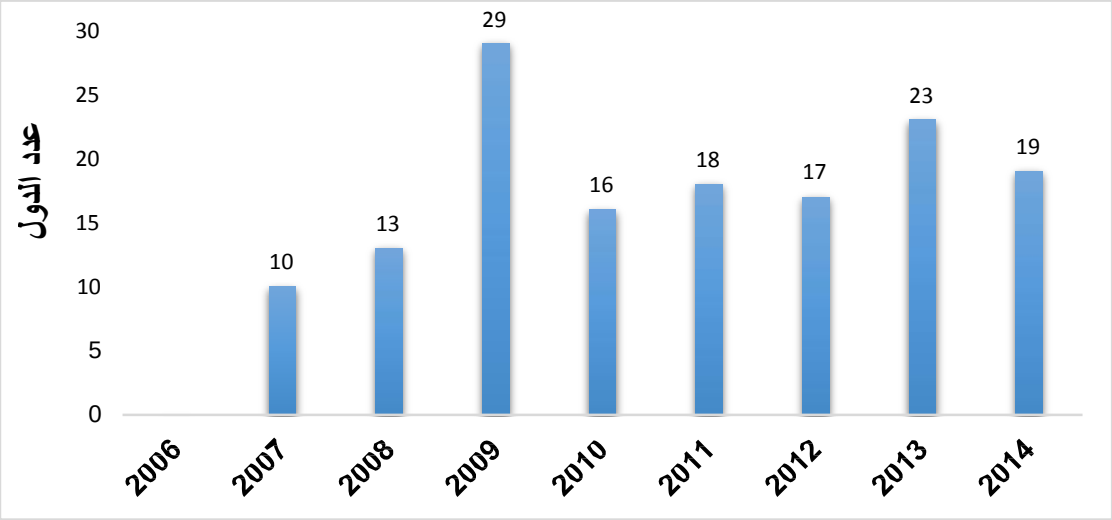


المصدر: الوقع الاليكتروني لبيانات التعليم EdData II الذي ينفذه منظمة معهد تراينجل للأبحاث ,
https://www.eddataglobal.org/countries/index.cfm

و بعد حوال 10 سنوات ومنذ سبتمبر 2015 من التطبيق الأول لأداة تقييم مهارات القراءة في الصفوف الأولى استخدمت الأداة أكثر من 30 مؤسسة في أكثر من 70 دولة. انتقل مدخل

تحسين القراءة في السنوات الأولى ليركز على تحسين تدريس اللغة الأم حيث طُوعت الأداة لتطبيقها في أكثر من 120 لغة مختلفة. وتتبع مشروع بيانات التعليم EdData II هذه التطبيقات من جانب الوكالة الأمريكية للتنمية الدولية (**أنظر الشكل 5**).

الشكل 5: التطبيق لأداة تقييم مهارات القراءة في الصفوف الأولى على مستوى العالم: عدد الدول والسنوات



المصدر: منظمة معهد تراينجال للأبحاث 2015

4.1 الحقيبة الأصلية للأداة والنسخة الثانية:

طلب البنك الدولي من منظمة معهد تراينجل للأبحاث في 2009 ان يطور حقيبة أو دليل المستخدم و الذي يمكن أن يكون دليلًا للدول للعمل بأداة تقييم مهارات القراءة في الصفوف الأولى في نواحي تطويع الأداة و العمل الميداني و تحليل النتائج ،وذلك بهدف تجميع الخبرات المتنوعة ووضع أسلوب مقنن لتقييم اكتساب الأطفال للقراءة المبكرة.

طلبت الوكالة الامريكية للتنمية الدولية من منظمة معهد تراينجل للأبحاث مراجعة الحقيبة حيث اتسع استخدام الأداة بين المنظمات والدول، وذلك يرجع إلى طبيعة أداة التقييم. عملت منظمة معهد تراينجل للأبحاث من خلال مشروع EdData II دعم القياس و الأبحاث الهدف 1 لاستراتيجية التعليم على تطوير النسخة الثانية من الحقيبة. تعكس هذه النسخة المراجعة التقدم الحادث منذ النسخة الأصلية للحقيبة نحو تحسين جودة البيانات المستخدمة لتحديث أجندة أهداف التنمية المستدامة.

بدأت عملية مراجعة الحقيبة في ديسمبر 2014. بدأت منظمة معهد تراينجل للأبحاث بتجميع الخبرات و المعلومات المتعلقة بأداة تقييم مهارات القراءة في الصفوف الأولى من أنشطة EdData II و الدراسات الأخرى التي تدعمها الوكالة الأمريكية للتنمية الدولية. تضمنت المصادر تقارير المشروعات و الأبحاث المنشورة وقصص الباحثين في مجال التعليم الدولي. وتضمنت المصادر أيضًا المواد التي تعكس أفضل الممارسات و الدروس المتعلمة من التطبيقات العديدة لأداة تقييم مهارات القراءة في الصفوف الأولى بالإضافة إلى تقارير تحليل البيانات عن نتائج الدول المختلفة و كذلك التطورات في مجال التكنولوجيا الحديثة في تخطيط و تنفيد تقييم مهارات القراءة في الصفوف الأولى. رُوجعت المعلومات ورُكزت في شكل عروض تقديمية ونشرات.

استخدمت بعد ذلك العروض التقديمية و النشرات في ورشة عمل لأداة تقييم مهارات القراءة في الصفوف الأولى بعنوان" تصميم و تنفيذ تقييمات مهارات القراءة في الصفوف الأولى: فهم الأساسيات" و التي استضافتها شبكة القراءة الكونية Global Reading Network كورشة عمل و ويبنار في مارس 2015.

بعد هذه الورشة في مارس حول أساسيات تقييم مهارات القراءة في الصفوف الأولى، عقدت ورشة أخرى لتحسين جودة بيانات التقييم في مايو 2015. استضافت شبكة القراءة الكونية Global Reading Network كورشة عمل و ويبنار ومولتها الوكالة الأمريكية للتنمية الدولية. قدم الخبراء من العديد نمن المنظمات العروض التقديمية حول تصميم أداة تقييم مهارات القراءة في الصفوف الأولى وإجراء التحليل وكتابة التقرير. تلت هذه العروض المناقشات بين المشاركين وأعضاء اللجان العارضين. جرت المناقشات للأفكار المطروحة نم قبل لجان العارضين.

كان هناك وجهين لهدف هذه الورش: التوسع في استخدام أداة تقييم مهارات القراءةEGRA عن طريق تقديم تدريب مفصل حول كيفية تنفيذ دراسة تقييم مهارات القراءةEGRA ، و كذلك تحسين جودة بيانات تقييم مهارات القراءةEGRA بعرض عمليات التقييم و التحليل. لمزيد من المعلومات انظر **الملحق** أ.

جاءت الخطوة التالية في عملية تحديث الحقيبة بعد ختام ورشتي العمل. تشكلت مجموعات العمل مؤلفة من الخبراء الفنيين. كانت مهمة مجموعات العمل مناقشة المجموعة النهائية للتوصيات و الاتفاق عليها و التي تتعلق بموضوعات الجلسات المقدمة ورشة عمل في مايو 2015.تضمنت هذه الجهود التعاونية الأفكار و الأمور التي أُتفق عليها من جانب المنظمات المتعددة التي تنفذ و التي تسعى لتقديم منهجيات محددة بشكل جيد لتخطيط و تنفيذ و تحليل

بينات تقييم مهارات القراءة EGRA.و تمثل هذه الحقيبة المحدثة نتاج ورش العمل هذه و مجموعات العمل الفنية. فهي تمثل المدخلات التعاونية للمنظمات العديدة والأفراد في مجالات التنمية و التعليم الدوليين.

استخدام الحقيبة:

صُممت هذه الحقيبة لكي يستخدمها موظفو وزارة أو أقسام التعليم والمهنيون في مجال تطوير التعليم. تسعى الحقيبة المكونة من 12 قسما إلى تلخيص مجموعة كبيرة من الأبحاث بطريقة يسهل الوصول إليها. يُقصد أيضًا من الإجراءات الموصوفة في هذه الحقيبة أن تُستخدم في كل تطبيقات تقييم مهارا القراءة EGRA التي تمولها الوكالة الأمريكية للتنمية الدولية ، وليس ذلك فقط بل يُؤمل أن تُنفذ في كل التطبيقات .

استخدام مجموعة الأدوات

بالإضافة إلى ذلك، فإن هذه النسخة العربية من مجموعة الأدوات تستهدف ناطقي اللغة العربية ممن يجرون تقييمات القراءة للمراحل الدراسية الأولى في السياقات العربية. وفي حين أن محتويات هذه الوثيقة مماثلة للبحث، والمنهجيات، والمعلومات الأخرى المقدمة في النسخة الإنجليزية، فإن عدة أقسام من هذا الدليل قد تم تعديلها عوضاً عن الترجمة المباشرة من الإنجليزية إلى العربية. ومن ثم، فإن الإرشادات، والشكل العام، والتصميم الخاص بالوثيقة قد يختلف في بعض المواضع.

بعد هذه المقدمة، يغطى القسم موضوع حماية العينة البشرية في البحث. فسم 3 عرض لأغراض واستخدامات EGRA. القسم 4 يقدم الإطار المفاهيم و أسس البحث (المقومات الأساسية للتقييم). القسم 5 يناقش الخيارات لتصميم الدراسة. يناقش القسم 6 الترتيبات لتطبيق التقييم. القسم 7 يعرض الجمع الاليكتروني للبيانات. يقدم القسم 8 المعلومات والإجراءات لتدريب المقيمين. يقدم القسم 9 النصح بخصوص جمع البيانات من الميدان لاستطلاع الأداة وكذلك الدراسة نفسها. يناقش القسم 10 الإجراءات الملائمة لبرتوكولات تنظيف و إعداد بيانات الاستبيان. القسم 11 يقدم عرضًا للتحليلات التي يتعين إجراؤها. واخيرًا، يقدم الجزء 12 الإرشادات حول تفسير النتائج، ووضع العلامات المرجعية، و بعض التوصيات الموجزة للحوار حول السياسات المتعلقة بتحسين تدريس القراءة و عرض النتائج على المدارس.

توجد مجموعة من المرفقات توضح النقاط الموجودة في النص مع لأمثلة و ورسوم توضيحية وتفاصيل فنية و إرشادات إحصائية.

2 أخلاقيات البحث والمراجعة الإلزامية لمجلس المراجعة المؤسسية

يجب على المؤسسات البحثية التي تتلقى تمويلاً فيدرالياً اتباع اللوائح الفيدرالية الأمريكية عند إجراء الأبحاث **الأخلاقية** وينص المبدأ الأساسي للإحصائيات الرسمية الخاص بالأمم المتحدة على أن: "البيانات الفردية التي تجمعها الوكالات الإحصائية للتجميع الإحصائي، سواء كانت تشير إلى شخصيات طبيعية أو اعتبارية، يجب أن تكون شديدة السرية وتستخدم حصرياً للأغراض الإحصائية."[5] وعلى جميع المنظمات الأمريكية التي تجري أبحاثاً تتضمن عناصر بشرية التشاور مع مجلس المراجعة المؤسسية قبل إجراء أي عملية مسح (22 قانون اللوائح الفيدرالية 225). وقد أدت محاكم نورمبيرج العسكرية إلى تشكيل مجالس المراجعة المؤسسية وحماية العناصر البشرية، وقد بدأت اللوائح الأمريكية فيما يتعلق بحماية العناصر البشرية في 1974.

على جميع المنظمات الأمريكية التي تجري أبحاثاً تتضمن عناصر بشرية التشاور مع مجلس المراجعة المؤسسية قبل أي عملية مسح.

1.2 ما هو مجلس المراجعة المؤسسية؟

تستخدم مجالس المراجعة المؤسسية مجموعة المبادئ الأساسية المبينة في "تقرير بيلمونت"، الذي أصدرته اللجنة الوطنية لحماية العناصر البشرية في الأبحاث الطبية الحيوية والسلوكية في الولايات المتحدة الأمريكية (1978)، لتوجيه مراجعتها للبروتوكولات المقترحة للأبحاث. ويحدد تقرير بيلمونت ثلاثة مبادئ أساسية:

- **احترام الأشخاص.** يجب معاملة العناصر المحتملة الخاضعة للبحث ككيانات مستقلة، لديها القدرة على النظر في البدائل، والاختيار، والعمل دون تأثير أو تدخل في غير محله من الآخرين.

[5] تعد السياسة الفيدرالية لحماية العناصر البشرية لازمة بموجب قانون اللوائح الفيدرالية الأمريكي، المادة 22 الجزء 225

- **المصلحة.** يتمثل المبدأين **الأساسيين** للمصلحة في: (1) عدم الإضرار، (2) الحماية من الضرر من خلال تعظيم الفوائد الممكنة والحد من الأضرار الممكنة.
- **العدالة.** يتطلب هذا المبدأ الأخلاقي النزاهة في توزيع أعباء **البحث** وفوائده. .

أصدرت الإدارة الأمريكية للأغذية والعقاقير ووزارة الصحة والخدمات البشرية الأمريكية خطوط إرشادية إضافية لتقييم العناصر البشرية عام 1981. وتتمثل المعايير التي وضعتها الوكالتين فيما يلي:

يجب تقييم البروتوكول للنظر في صحته وقيمته العلمية؛

1. يجب الحد من المخاطر إلى أقصى قدر ممكن؛
2. يجب اختيار العناصر بطريقة منصفة؛
3. يجب الحصول على الموافقة الواعية؛
4. يجب حماية الخصوصية والسرية؛
5. يجب متابعة الدراسة بشكل ملائم.

2.2 كيف تنطبق موافقة مجلس المراجعة المؤسسية على دراسات تقييم المراحل المبكرة للقراءة (إجرا)؟

كما سبقت الإشارة، فإن جميع المنظمات التي تحصل على الدعم عن طريق التمويلات الفيدرالية الأمريكية أو التي تخضع بصورة أخرى للوائح أي وزارة أو وكالة فيدرالية وتجري بحوثاً تتضمن عناصر بشرية يجب أن تتشاور مع أحد مجالس المراجعة المؤسسية والحصول على موافقته قبل إجراء البحث. يُعرف البحث بأنه "التحقيق المنهجي، بما في ذلك إعداد البحث، واختباره، وتقييمه، والمصمم لتطوير المعرفة العامة أو المساهمة فيها". وتتولى مجالس المراجعة المؤسسية مراجعة المشروعات البحثية التي تتضمن عناصر بشرية

يُعرف البحث بأنه "التحقيق المنهجي، بما في ذلك إعداد البحث، واختباره، وتقييمه، والمصمم لتطوير المعرفة العامة أو المساهمة فيها." - القانون الأمريكي للوائح الفيدرالية 22-225.

وتحدد درجة المخاطر التي قد تتعرض لها تلك العناصر نتيجة للمشاركة في البحث. ويوافق المجلس على الأنشطة البحثية أو يرفضها بناء على المراجعة الشاملة لبروتوكولات الدراسة والظروف التي يجري فيها البحث.

وفي الغالب تكون الأبحاث التي تشتمل على اختبارات تعليمية معفاة من متطلبات مجلس المراجعة المؤسسية (على أساس أن الاختبارات التي تجرى لا تختلف كثيراً عما يشهده الأطفال في بيئات مدارسهم الطبيعية). إلا أن مجلس المراجعة المؤسسية هو وحده القادر على حصول الدراسة المتعلقة بتقييم إجرا على وضعية الإعفاء. إذا ما قرر المجلس منح الإعفاء للدراسة بالكامل، قد لا تزال هناك حاجة للحصول على الموافقة المسبقة على أسئلة المسح على وجه الخصوص، إذا كانت المعلومات المجمعة أثناء الدراسة قد تعرض الطلاب أو المدرسين للمخاطر.

في حالة تقييم إجرا والتقييمات ذات الصلة بالأطفال الصغار، يجب أن تتاح الفرصة لكل بلد يصرح بإجراء دراسة إجرا بأن تراجع الهيئة الأخلاقية بالبلد بمراجعة شروط الدراسة وإصدار الموافقة بالمباشرة، أو طلب أي تعديلات لازمة لضمان تلك الموافقة (القانون الأمريكي للوائح الفيدرالية 22-225).

3.2 قبول الأفراد المشاركين وموافقتهم عن علم

دائماً ما يتضمن نموذج تقييم إجرا والوثائق الداعمة له قسم في البداية يوجه المقيمين بشأن كيفية طلب الموافقة (من البالغين) أو القبول (من الأطفال) من المختارين للمشاركة. وقبل تطبيق إجرا على الأطفال، يصف المقيمين أهداف الدراسة وتعريف الطلاب بأن التقييم دون هوية، ولن يؤثر على درجاتهم الدراسية، وسيستخدم لإدخال تحسينات على تعلم الأطفال في بلادهم للقراءة. لكل طفل خيار الموافقة الشفهية على الخضوع للاختبار أو رفض المشاركة، دون عواقب من أي نوع. إذا كانت استبيانات مدير المدرسة والمدرسين تجرى كجزء من الدراسة، يتم استيفاء عملية موافقة مماثلة – ولكن كتابية بدلاً من شفهية.

على الرغم من أن عملية القبول/الموافقة قد تكون غير شائعة لدى القرناء في البلد المضيف، غالباً ما تلقى العملية الترحاب من الطلاب والمدرسين، الذين يشيرون إلى شعورهم بالتمكين بمنحهم خيار المشاركة من عدمه. وقد أوضحت التجربة عبر العديد من عمليات تنفيذ إجرا حتى الآن أن عدد قليل من الطلاب أو المدرسين يرفضون المشاركة. وفي حال رفض أحد المشاركين

المؤهلين المشاركة، يتم اختيار مشارك آخر بشكل عشوائي. للحصول على المزيد من المعلومات عن مجالس المراجعة المؤسسية والأبحاث الأخلاقية ذات العناصر البشرية، بما في ذلك الأطفال، يرجى الاطلاع على الموقع الإلكتروني لمكتب حماية البحوث البشرية بوزارة الصحة والخدمات البشرية الأمريكية: http://www.hhs.gov/ohrp

3 الهدف من تقييم القراءة للمراحل الأولى واستخداماته

3-1 نبذة تاريخية وعرض عام

على الرغم من أنه بدا واضحا من البداية أن تقييم القراءة للمراحل الأولى سيركز على السنوات الأولى والمهارات التأسيسية للقراءة، فإن استخدامات النتائج كانت محل نقاش أوسع.

تم تصميم الأداة الأصلية لتقييم القراءة للمراحل الأولى لتكون مقياسا لـ"تشخيص النظام" باستخدام العينات. كان الهدف منها توثيق أداء الطالب من حيث مهارات القراءة في المراحل الأولى بحيث يتم إعلام الحكومات والمانحين بخصوص احتياجات التحسين في التلقين.
وبمرور الوقت، توسعت استخداماتها لتشمل كافة ما يلي، مع استخدامات مختلفة في سياقات مختلفة:

- توليد بيانات مرجعية حول اكتساب القراءة في المراحل الأولى في سنوات و/أو مناطق معينة.
- توجيه إعداد برامج التلقين من خلال تحديد المهارات الأساسية أو مساحات التلقين التي تحتاج للتحسين.
- الوقوف على التغيرات في مستويات القراءة عبر الوقت.
- تقييم مخرجات أو أثر البرامج المصممة بهدف تحسين القراءة في المراحل الأولى
- استكشاف فعالية تكلفة تصميمات البرامج المختلفة.
- تطوير مؤشرات ومقاييس للقراءة.
- العمل كمشخص للنظام (انظر قسم 3.2) لتقديم المعلومات لسياسة قطاع التعليم والتخطيط الاستراتيجي وتخصيص الموارد.

علاوة على ذلك، "فإن المهام الفرعية المتضمنة في تقييم القراءة للمراحل الأولى يمكن تطويعها للمدرسين للمساعدة في التدريس.[6] وكتقييم موجه، يمكن للمدرسين إما استخدام تقييم القراءة للمراحل الأولى في مجملها أو اختيار مهام فرعية لمراقبة تقدم الفصل الدراسي

والوقوف على الاتجاهات في الأداء وتطويع التلقين للوفاء باحتياجات الأطفال التعليمية"
(Dubeck&Gove، 2015: 2).

غير أنه، وتوضيحاً للأمور، فإن تقييم القراءة للمراحل الأولى في تصميمها الحالي لديها مساحات قصور. فهي ليست مصممة للعمل كمقياس للمحاسبية عالية المخاطر لتحديد تحسن درجات الطلاب أو تقييم المدرسين كل على حدة. فتقييم القراءة للمراحل الأولى مصممة لتكمل، وليس لتحل بديلاً عن، التقييمات الورقية الموجودة والمعتمدة على المقررات. يتكون تقييم القراءة للمراحل الأولى من عدة مهام فرعية تقيس المهارات التأسيسية التي يمكن أن تكشف عن نجاح لاحق في القراءة. غير أنه بسبب القيود المفروضة نظراً لمحدودية نطاق تركيز الأطفال وقوة تحملهم، تعجز أي أداة بمفردها، بما فيها تقييم القراءة للمراحل الأولى، عن قياس كل المهارات المطلوبة للطلاب للقراءة بإدراك وفهم. لا يهدف تقييم القراءة للمراحل الأولى لأن يصبح برنامجا تلقينياً وإنما أن يوجه البرامج التلقينية. لا يمكن لتقييم القراءة للمراحل الأولى بشكل كامل تحديد السلوكيات الأساسية أو التعليمية والتي يمكن أن تؤثر على قدرة الطالب على القراءة (Dubeck&Gove، 2015).

2.3 تقييم القراءة للمراحل الأولى كأداة لتشخيص النظام

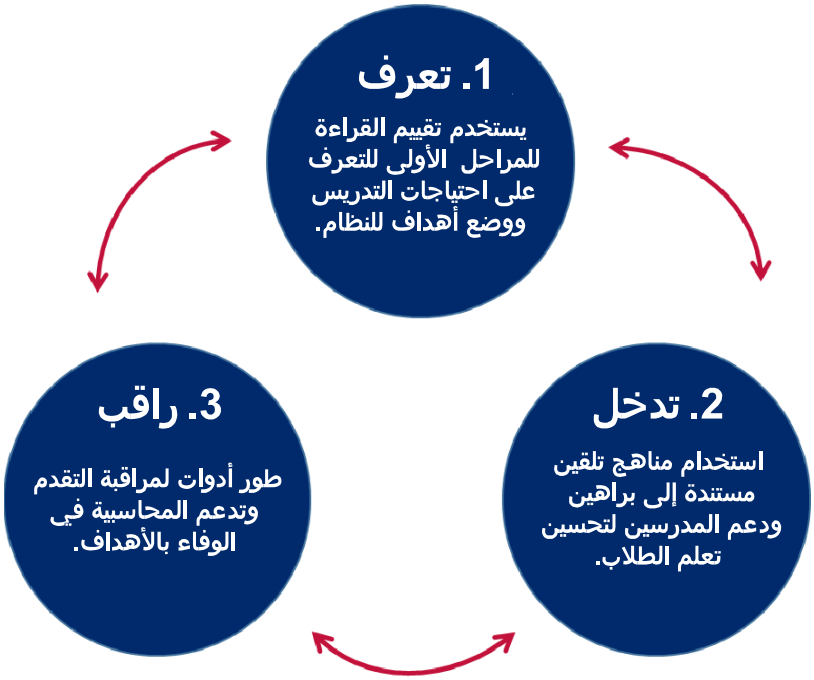
تم تصميم تقييم القراءة للمراحل الأولى المشخص للنظام، كما هو مفصل في هذا الدليل التدريبي، بحيث تصبح جزءا من الدورة الكاملة لدعم وتحسين التعلم. كما هو مبين في **شكل توضيحي 6**، يمكن استخدام تقييم القراءة للمراحل الأولى كجزء من منهج متكامل لتحسين قدرات الطلاب على القراءة، تكون الخطوة الأولى فيه هي التعرف على مساحات التحسين على مستوى النظام. بمقدور تقييم القراءة للمراحل الأولى وكذلك توليد بيانات مرجعية حول اكتساب القراءة في المراحل الأولى (Dubeck&Gove، 2015). يمكن أيضا تحديد المقاييس الرئيسية ووضع الأهداف للتطبيقات المستقبلية من خلال التطبيق الأولي لتقييم القراءة للمراحل الأولى. واعتمادا على نتائج التقييم، يصبح بالتالي بمقدور وزارات التعليم أو النظم

[6] استخدام تقييم القراءة للمراحل الأولى كتقييم موجه يعتمد على فصول الدراسة يمكن القيام به فقط مع التعديلات المحددة المطلوبة للأداة ولإجراءات سحب العينة. التقييمات المعتدة على فصول الدراسة قد تتطلب كذلك التطوير المهني للمدرسين مع توجيهات محددة بخصوص الإدارة وتفسير المهام الفرعية.

المحلية *التدخل* لتوجيه محتوى البرامج الجديدة أو القائمة باستخدام مناهج تلقين مستندة إلى براهين لدعم المدرسين في تحسين القدرات التأسيسية في القراءة. ومن ثم يمكن استخدام نتائج تقييم القراءة للمراحل الأولى لتوجيه تطوير كلا من تدريبات المدرسين وبرامج التطوير المهني.

بمجرد تنفيذ التوصيات، يمكن استخدام صور مختلفة من تقييم القراءة للمراحل الأولى لتتبع التطور ومكتسبات تعلم الطلاب عبر الوقت باستخدام *المراقبة* المستمرة، مع توقع أن مثل هذه العملية ستشجع المدرسين والقائمين على إدارة التعليم على ضمان تحقيق الطلاب للتقدم في اكتساب المهارات التأسيسية.

شكل توضيحي 6، الدورة المستمرة لتحسين تعلم الطلاب



تقييم القراءة للمراحل الأولى والتقييمات المعتمدة عليه يمكن استخدامها *للتعرف* على الاحتياجات *والتدخل ومراقبة* التقدم باتجاه تحسين مخرجات

عند العمل على مستوى النظام، فإن الباحثين ومسئولي التعليم يبدؤون في الغالب ببيانات على مستوى الطلاب، التي تم جمعها في عينات وحصلت على الوزن النسبي الملائم، حتى يتسنى

الوصول إلى نتائج حول أداء النظام (والطلاب المدرجين فيه). أساس هذا المنهج هو فهم أن الطرق التي يتعلم من خلالها الطلاب في مجملها هي انعكاس للتلقين الذي يتلقونه. من خلال استخدام أداء الطالب المتوسط من حيث الدرجات على مستوى النظام، يمكن للمسئولين تقييم المساحات داخل النظام التعليمي التي يواجه فيها الطلاب صعوبات في المعتاد ويمكن استخدام هذه المعلومات لتطوير مناهج التلقين الملائمة. وشأنها ككل التقييمات التي تهدف إلى تشخيص الصعوبات وتحسين مخرجات التعلم، لكي يكون المقياس مفيداً فإنه يجب: 1) يأخذ في الاعتبار توقعات ومقاييس الأداء، 2) يعتبر المهارات المرغوب في التمتع بها لاحقا، 3) وأن يكون من الممكن تعديله أو تحسينه وفق المهارات من خلال التلقين المعدل (Linan-Thompson&Vaughn، 2007). يفي تقييم القراءة للمراحل الأولى بهذه المتطلبات على النحو التالي:

أولاً، في العديد من دول الدخل المرتفع، بإمكان المدرسين (ومسئولي النظام) النظر إلى التوزيعات الوطنية الموجودة ومعايير الأداء لفهم أداء طلابهم مقارنة بمن سواهم. في الولايات المتحدة وأوروبا، يمكن لمسئولي النظام تحديد ما إذا كانت المدارس والمدرسين بحاجة لدعم إضافي من خلال مقارنة أداء المجموعة الفرعية للطلاب بالتوزيعات ومعدلات الأداء القومية. وبطريقة مماثلة، يمكن استغلال تقييم القراءة للمراحل الأولى من قبل دول الدخل المنخفض لتحديد المناطق (أو، إذا ما كان ممكنا من خلال العينة، المدارس) التي تستحق المزيد من الدعم، بما في ذلك تدريب المدرسين وغيرها من التدخلات. عند تصميم تقييم القراءة للمراحل الأولى لأول مرة كانت المشكلة التي تواجه دول الدخل المنخفض هي عدم توافر (بعد) مقاييس للأداء تعتمد على نتائج محلية. حاليا، تم إجراء التقييم في 12 دولة على الأقل لوضع مقاييس قومية أو إقليمية باستخدام بيانات تقييم القراءة للمراحل الأولى. ستتم مناقشة التفاصيل في قسم 12.2.

علاوة على ذلك، فإنه تم تطوير مهام تقييم القراءة للمراحل الأولى بشكل مقصود لتكون قادرة على التنبؤ بالتحصيل اللاحق في القراءة والعديد من التطبيقات لتقييم القراءة للمراحل الأولى في العديد من الدول واللغات أكدت على هذه الارتباطات المتوقعة. على الرغم من تأثير التباينات الصوتية والإملائية على معدل وتوقيت اكتساب القراءة، فإن كل المهارات التي يتم قياسها من قبل تقييم القراءة للمراحل الأولى كشف عن ارتباطها بمهارات القراءة في تهجئة الحروف الأبجدية. على سبيل المثال، معرفة العلاقة بين الأصوات والحروف التي تمثلها ذو علاقة تنبؤية بالنجاح في قراءة الكلمات. كذلك تم الكشف عن أن الطلاقة في القراءة الشفهية

تنبئ عن فهم القراءة. يتم قياس هذه المهارات في تقييم القراءة للمراحل الأولى وبالتالي يمكن الافتراض بثقة أن نتائج تقييم القراءة للمراحل الأولى تقدم لنا نتائج ذات معنى بخصوص الاتجاه الذي يتحرك باتجاهه الأطفال في إطار عملية تعلم القراءة.

ثالثاً، لا يقدم لنا تقييم القراءة للمراحل الأولى توقعات ذات معنى بخصوص الأداء المستقبلي فحسب، بل يمكن أن يوجه انتباهنا كذلك إلى التغييرات المطلوبة في التلقين. سيكون من غير المجدي قياس شيء لا نطمح إلى إحداث تغييرات فيه من خلال إدخال تعديلات على طريقة التلقين. ولذلك فإن تقييم القراءة للمراحل الأولى هو أداة تشخيص قيمة لتضمنه مقاييس للمهارات التي يمكن تحسينها من خلال التلقين.

4 الإطار المفاهيمي والأسس البحثية

إن الإطار المفاهيمي لاكتساب القراءة الذي تقوم عليه أداة تقييم مهارات القراءة في الصفوف الأولى يستمد اتجاهاته الأساسية من أعمال اللجنة الوطنية للقراءة (المعهد الوطني لصحة الطفل والنمو الإنساني2000)، و(2006) August and Shanahan ، ولجنة علاج صعوبات القراءة لدى الأطفال الصغار,(Snow, Burns, & Griffin, 1998) ،و أخرين. وتشير الدراسات السابقة الكثيرة حول القراءة إلى الحاجة لاكتساب مهارات أساسية من خلال التدريس المستهدف بغية تنمية قراء ناجحين مدى الحياة.

1.4 موجز للمهارات الأساسية للقراءة الناجحة:

الهدف النهائي لتعلم القراءة هو الفهم، أو عملية استخلاص وبناء المعني بشكل تلقائي من خلال التفاعل والانخراط مع اللغة المكتوبة" Snow & the RAND Reading Study Group, 2002, p. 11) ، وقد تبدو القراءة بالنسبة للقراء الماهرين أنها سهلة ولا تتطلب جهدًا كبيرًا، فهم يقرأون النص و يفهمونه بسرعة و سهولة لدرجة أنهم قد لا يكونون واعين بعملية الفهم نفسها. ومع ذلك، فإن الفهم عملية مركبة بالفعل أو سلوك مركب يكتمل بناءه عندما تُتقن مجموعة من المهارات الفرعية وتُستخدم بتلقائية.

ويعد اكتساب القراءة عملية نمائية (Chall, 1996). تُبنى المهارات العليا (كالطلاقة و الفهم)، على أساس المهارات الدنيا (كالوعي الصوتي ، ومعرفة الأصوات و الحروف ، وقراءة الكلمة) ، و المهارات الدنيا لها القدرة على التنبؤ بالتحصيل القرائي فيما بعد. ونتيجة لذلك، إذا كان الأطفال غير قادرين على قراءة قطعة وفهمها فلا يمكن قياس تقدمهم نحو اكتساب المهارات الدنيا التي تعد مهارات لازمة لتحقيق الغاية من عملية القراءة نفسها.

هناك اتفاق حول وجود خمسة مكونات لازمة لإتقان عملية القراءة وهي: الوعي الصوتي ، و المبدأ الأبجدي ، والمفردات ، و الطلاقة ، و الفهم (الاستيعاب) on (Armbruster, Lehr, & Osborn, 2003; Vaughn & Linan-Thompson, 2004). ، ولا تكفي المهارات المتضمنة داخل كل مكون بمفردها للتوصل إلى القراءة الناجحة، ولكن تعتمد هذه المهارات على بعضها البعض ؛ للوصول إلى الهدف النهائي ، وهو الفهم القرائي

المهارات الفرعية من أداة تقييم مهارات القراءة EGRA (ارجع إلى القسم 6) ، والتي تتماشي مع هذه المكونات للقراءة.ونظرًا لأن هذه المهارات تُكتسب وتُتعلم على مراحل ، وفي أى فترة زمنية ، فإنه من المحتمل أن بعض المهام الفرعية تحصل على أدنى أثر ، وذلك يعني أن معظم الأطفال في السنوات الأولى لن يستطيعوا تحقيق مستوى كاف من المهارة يسمح بتحليل النتائج)) ،أو الحصول على بعض هذه المهام الفرعية على أعلى أثر (وهذا يعني أن معظم الأطفال يحصلون على درجات عالية) ، و هذا يعتمد على مرحلة النمو للأطفال.

2.4 الوعي الصوتي:

1.2.4 الوصف

بمكن تعريف الوعي الصوتي على أنه " القدرة على رصد و امتلاك أو تحليل النواحي المسموعة للغة المنطوقة (والذى يتضمن القدرة على تمييز، و تقطيع الكلمة، و المقاطع و الأصوات) بعيدًا عن المعنى". (المركز الوطني لإجادة القراءة والكتابة للأسرة [NCFL], 2008, p. vii.مصطلحا الوعي الصوتي والوعي الفنولوجي يُستخدمان بالتبادل، حيث يشير إلى الوعي بالأصوات – الفونيم – وهو أصغر وحدة صوتية تميز معنى الكلمة في لغة ما. فعلى سبيل المثال صوتا /k/, /p/7, في اللغة الإنجليزية والصوت الاحتكاكي/ð/ (صوت th) هي الوحدات الصوتية التي تميز كلمة pat عن cat و that في اللغة المنطوقة.

وعلى نحو مشابه، في طرق الكتابة الأبجدية، يشير الجرافيم إلى اللغة المكتوبة وعلاقتها بالأصوات المسموعة -كما هو موضح بقائمة المصطلحات في أول الحقيبة، فهو الوحدة الأساسية في أي نظام كتابة أبجدي يمكن أن يغير معنى الكلمة. وقد يتكون الجرافيم من حرف أو أكثر أو حرف وحركة" (انظر تعريف حركات التشكيل في قائمة المصطلحات). وتتنوع اللغات في درجة التطابق بين الأصوات وطريقة كتابة الحروف التي تمثلها، ففي بعض اللغات كاللغة الإسبانية، فالجرافيم والفونيم بينهما علاقة تمثيل متطابق تقريبًا، ولكن في اللغة الإنجليزية الرابطة أكثر تعقيدًا. فعلى سبيل المثال، الصوت في اللغة الإنجليزية يمكن أن تمثله الحروف التالية: *c, k, ck, ch, qu* إلى آخره، تمامًا كالحرف *c* يمكن أن يمثل الصوت /k/ في كلمة وصوت /s/ في كلمة أخرى.

معرفتنا بالوعي الصوتي تظل في اللاوعي كما يعالج البشر اللغة الشفوية السريعة. وعلى ذلك، فإن تعلم القراءة (في اللفات الأبجدية)، ويطلب الربط بين الأصوات المفردة والحروف التي تمثلها، والتي تتطلب الوعي المقصود لأصوات اللغة والقدرة على التمييز بينها والتلاعب بها. (Gove & Wetterberg, 2011). ، ويساعد الوعي الصوتي الأطفال على تقطيع الكلمات إلى أصوات و دمج الأصوات لتكوين كلمات، وهي مهارات شفوية تمثل مهارات أولية ضرورية تؤدي إلى قراءة الكلمات وهجائها.

[7] Phonemes are traditionally written between slashes in the International Phonetic Alphabet (IPA). The full IPA chart is available for reference and use from http://www.internationalphoneticassociation.org/content/ipa-chart, under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2005 International Phonetic Association.

وتوصى نتائج الأبحاث أن معرفة الأطفال بأصوات الكلام تتطور بشكل تقدمي بدءً بالوحدات الأكبر، فعلى سبيل المثال على مستوى الكلمة، ثم التقدم نحو الوحدات الأصغر التي تتضمن المقطع والأصوات التي يبدأ بها وتلك التي تتمم المقطع ثم الفونيم. في الواقع، والحساسية نحو مستوى الفونيم، والذي يعد ضروريًا لقراءة الكلمة، والذي قد لا يبدأ في التطور مع بداية تدريس مهارات القراءة والكتابة .(Goswami, 2008 وكما توضح العديد من الدراسات أن الوعي الصوتي في لغات عديدة لديه القدرة على التنبؤ بالتحصيل في القراءة فيما بعد (Badian, 2001; Denton, Hasbrouck, Weaver, & Riccio, 2000; Goikoetxea, 2005; McBride-Chang & Kail, 2002; Muter, Holme, Snowling, & Stevenson, 2004; Wang, Park, & Lee, 2006).

2.2.4 قياس الوعي الصوتي:

تتضمن أدة تقييم مهارات القراءة EGRA عادة مقياسين على مستوى الفونيم (الوعي الصوتي). ويشمل مهام تعرف الصوت في بداية الكلمة، وفيه يقدم للتلاميذ كلمة تقرأ شفويًا، ويُطلب منهم تحديد الصوت الأول في الكلمة ونطقه، وكذلك مهمة تمييز الصوت في بداية الكلمة، وفيه يُقدم للتلاميذ ثلاث كلمات، ويُطلب منهم اختيار الكلمة التي تبدأ بصوت مختلف عن الكلمتين الأخرتين.

وتوجد مهمة فرعية اختيارية (بمعنى أنها ليست أصلية) استُخدمت من قبل، وهي تقطيع الكلمات إلى مقاطع ويُقدم فيها للتلاميذ كلمة تقرأ شفويًا، ويُطلب منهم على حدة نطق المقاطع التي تتألف منها الكلمة، ونظرًا لأن التقطيع مهارة مركبة (Linan-Thompson & Vaughn, 2007)، فإن هذه المهمة الفرعية توصلت لنتائج اتفقت مع الحد الأدنى للأثر في سياقات عديدة ذات أداء منخفض. ومع ذلك، قد تكون بديلًا ملائمًا عندما تعطى المهمة الفرعية لتعرف الأصوات في بداية الكلمة الأعلى أثر.

3.4 المبدأ الأبجدي/ الصوتيات، وفك الرموز

1.3.4 الوصف

يشير المبدأ الأبجدي إلى فهم الكلمات المكونة من أصوات (فونيم)، وأن الحروف (الجرافيم)

هي رموز تمثل أصوات. والمبدأ الأبجدي مفهوم مجرد، وأفضل طريقة لتدريسه للتلاميذ هي التدريس المباشر حتى يُوضح ما تمثله الرموز الموجودة على الصفحة في صورها الأولية، وعندما يعي التلاميذ أن الأصوات توضحها الحروف، فيمكنهم البدء في تعلم فك رموز الكلمات، وتتضمن المعرفة بالأبجدية معرفة أسماء الحروف المفردة، والملامح المميزة لها عند كتابتها وأي الأصوات التي تمثلها.

تدريس العلاقات بين الصوت و الحرف، و الحرف و الصوت طريقة تدريسية تُعرف بالصوتيات . وأوضحت نتائج الأبحاث أن معرفة الأبجدية لديها قدرة قوية على التنبؤ مقارنة بالتحصيل القرائي فيما بعد (Adams, 1990; Ehri & Wilce, 1985; Piper & Korda, 2010; Wagner, Torgesen, & Rashotte, 1994; Yesil-Dağli, 2011) بالنسبة للمتحدثين الأصليين وغير الأصليين لأي لغة (Chiappe, Siegel, & Wade-Woolley, 2002; Manis, Lindsey, & Bailey, 2004; Marsick & Watkins, 2001; McBride-Chang & Ho, 2005). ، ويعد واحدًا من الفروق الأساسية بين القراء الجيدين و القراء الضعاف هو قدرتهم على استخدام العلاقة بين الأصوات والحروف في فك رموز الكلمات الجديدة التى يقابلونها ، وقراءتها في أي نص ، و القدرة على ترميز (هجاء) الكلمات التي يكتبونها(Juel, 1991)

الأصوات والإملاء

تتنوع اللغات في درجة الصعوبة في النظام الصوتي. بعض اللغات لديها العديد من الأصوات أكثر من لغات أخرى. فالبعض لديه تركيبات أكثر تعقيدًا للمقطع (مجموعة من الصوامت في بداية المقطع وفي نهاية المقطع). فبعض اللغات لديها أيضًا كلمات أطول من غيرها...وبالمثل، تتنوع طرق الكتابة (نظام الهجاء في اللغة) في درجة الشفافية أو الاتساق في العلاقة بين الصوت والحرف.

العلاقة بين الحروف والأصوات في نظم الكتابة الشفافة بشكل كبيره علاقة واحد لواحد، وهذا ييسر اكتسابهم وذلك لأن كل حرف يمثل صوتًا واحدا نفس الصوت على نحو موثوق به بغض النظر عن الكلمة التي يرد فيها. وعلى النقيض، يمكن أن يُطلق على نظام الكتابة في اللغة الإنجليزية أنها غير شفافة أو عميقة؛ وذلك لأن كل حرف تقريبًا يمثل أكثر من صوت واحد، وكل صوت يمثله أكثر من حرف، وبالتالي يُصعب هذا عملية تحديد العلاقة بين الصوت والحرف بشكل كبير.

وبإيجاز، فإنه توجد تبعات لكلٍّ من الصعوبة النسبية للنظام الفونولوجي للغة ما ، و نظام الكتابة الخاص بها التي ترتبط بمعدل اكتساب المهارات الفرعية المتعلقة بها كالصوتيات، و على طرفي النقيض، فإن الطفل الذي يتعلم القراءة في لغة متسقة وشفافة ، ولديها مخزون من الأصوات قليل وتركيب بسيط للمقاطع ومتوسط قليل من طول الكلمات التى يكون في وضع أفضل في إتقان العلاقات بين الأصوات والحروف ، ومهارات قراءة الكلمة بشكل أسرع من الطفل الذي يتعلم القراءة في لغة ليست شفافة ، وبها الكثير من حالات عدم الانتظام و العديد من اِلأصوات والتركيبات المعقدة للمقاطع ، و متوسط كبير لطول الكلمات. وهذا واحد من الأسباب وراء عدم ملاءمة ومقارنة نتائج تقييم مهارات القراءة والعلامات المرجعية بين اللغات.

ووفقا لنموذج الطريق المزدوج لتعرف الكلمات (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Zorzi, 2010)، هناك طريقتان متميزتان ولكن ليستا الطريقتان الحصريتان التي يعالج بها البشر النص لتعرف الكلمات، وتعرفان على أنهما الطريقة المعجمية والمعجمية الفرعية.

تتطلب القراءة من خلال الطريقة المعجمية البحث عن الكلمة في المعجم العقلي الذي يحتوي على المعرفة بطرق الهجاء ونطق الكلمات الحقيقية، ويشير التعرف السريع للكلمة إلى أن الكلمات الموجودة على الصفحة مألوفة ويُتعرف عليها بسرعة، وذلك بسبب معرفة الحروف ونموذج هجائها. أما في الطريقة المعجمية الفرعية، فنحن نفك رموز (نقرأ) الكلمات بتحويل الحروف إلى أصوات باستخدام معرفتنا بالعلاقات بين الصوت والحرف، ودمج الأصوات لتكوين كلمات، ثم إدراك الكلمة بناءً على صيغتها الفنولوجية.

قد تكون الطريقة المعجمية أسرع بالنسبة للكلمات المألوفة، وهي ضرورية لمعالجة الكلمات التي تُكتب بشكل غير منتظم، ولكن الطريقة المعجمية الفرعية ضرورية لمعالجة الكلمات الجديدة وغير المألوفة. ويمكن قراءة كل الكلمات ويسهل التوصل إليها من خلال الطريقة المعجمية الفرعية في اللغات ذات نظام الكتابة المتسق. توظف أداة تقييم مهارات القراءة EGRA مهمة قراءة الكلمات التي بلا معنى (غير المألوفة)؛ لتقييم مهارات التلاميذ في فك الرموز (قراءة الكلمات) من خلال الطريقة المعجمية الفرعية.

2.3.4 قياس مهارات معرفة الأبجدية وفك الرموز:

تُقيم أداة تقييم مهارات القراءة في الصفوف الأولى EGRA معرفة الأطفال بالأبجدية بطرق عديدة، تبدأ المهمة الفرعية تعرف الحرف، وهي أحد مكونات أداة تقييم مهارات القراءة في الصفوف الأولى EGRA الأصلية. وتختبر المهمة الفرعية لتعرف الحروف قدرة الأطفال على تعرف الخصائص المتعلقة بشكل وصوت كل حرف، وتحدد الصوت بدقة أو اسم الحرف الذي تمثله. إما تُختار المهمة الفرعية لتعرف الحرف أوكلاهما، وهذا يتوقف على ما هو ملائم للسياق: تعرف اسم الحرف أو تعرف صوت الحرف.وفي كلا البديلين، يُقدم للأطفال قائمة مكتوبة تحتوي على الحروف (أو الحروف مع الحركات) بشكل عشوائي، ويُطلب منهم تسميه كل حرف أو قراءة كل صوت.

كانت المهمة الفرعية لتعرف اسم الحرف هي في الأصل المقياس الأكثر استخدامًا لمعرفة الأبجدية في التقييمات التي تستخدم أداة تقييم مهارات القراءة في الصفوف الأولى EGRA ،واتضح أن لديها القدرة على التنبؤ بقوة بالتحصيل في القراءة في اللغة الانجليزية فيما بعد .ومع ذلك، و بمرور الوقت أصبحت المهمة الفرعية لتعرف الصوت هي البديل الأكثر استخدامًا، حيث إن معرفة أصوات الحروف ترتبط بشكل مباشر بقدرة الأطفال على فك رموز (قراءة) الكلمات خاصًة في نظم الكتابة الشفافة (Ehri, 1998).

قد يختار مطورو أداة تقييم مهارات القراءة في الصفوف الأولىEGRA إضافة مقياس تعرف المقطع بالإضافة إلى أسماء الحروف وأصواتها. استخدمت هذه المهمة الفرعية في السياقات التي تحتوي اللغة فيها على مقاطع مفتوحة ، و تؤكد طريقة تدريس القراءة على التجمعات المقطعية.

الخطوة التالية الأعلى في درجة صعوبة المهارة بالنسبة للقراء هي أن يستخدم الأطفال إتقانهم للعلاقة بين الصوت والحرف في فك رموز (قراءة) كلمات غير مألوفة (لا معنى لها)، وفي المهمة الفرعية لقراءة الكلمات غير المألوفة (لا معنى لها) يُقدم للأطفال قائمة مكتوبة بكلمات لا معنى لها تتبع قواعد الكتابة للغة، ولكنها ليست كلمات حقيقية لها وجود في اللغة. ويُطلب من الأطفال قراءة أكبر عدد من الكلمات بصوت مرتفع بأكثر سرعة وأعلى درجة من الدقة يستطيعونها. ووفقًا لنموذج الطريق المزدوج، تتطلب هذه المهمة الفرعية من الأطفال أن يوظفوا مهاراتهم في فك الرموز بناءً على معرفتهم بخرائط العلاقة بين الصوت والحرف. فلابد للتلاميذ من أن يعتمدوا على فك الرموز حتى يتمكنوا من قراءة الكلمة، وذلك لأن الكلمات غير المألوفة ليس لها أي تخزين سابق في الذاكرة ككلمة كلية في الذاكرة طويلة الأمد يجري تقييمها بشكل مباشر.

تتشابه المهمة الفرعية لقراءة الكلمات المألوفة في الشكل مع المهمة الفرعية لقراءة الكلمات غير المألوفة ماعدا أنها تقدم للأطفال قائمة بها كلمات، ويُتوقع من الأطفال أن يكونوا قادرين على القراءة بما يتناسب مع مستوى صفهم، والتي من المحتمل أن يكوا رأوها من قبل. ووفقًا لنموذج الطريق المزدوج مروة اخرى، يُتوقع من الأطفال أن يُعالجوا الكلمات المألوفة -إذا كانت بالفعل مألوفة وبخاصة إذا كانت مكتوبة بطريقة غير منتظمة عن القاعدة الأساسية – بشكل مباشر باستخدام الطريقة المعجمية، وهذا يعني أنهم قد يتعرفون على الكلمات بسرعة أكثر من محاولة قراءتها صوتًا صوتًا.

أخيرًا، الإملاء حيث يُطلب من التلاميذ أن يستمعوا إلى أصوات الحروف وكلمات و/أو جمل قصيرة، ويكتبونها. وتقيس تلك المهمة الفرعية معرفة التلاميذ بالأبجدية وقدرتهم على سماع وتمييز الأصوات المفردة بمعزل أو في كلمات و تشفير الكلمات (كتابتها) بشكل صحيح. إذا قُدمت الجملة للتلاميذ، قد تقيس المهمة قدرتهم على استخدام قواعد كتابة الجملة كعلامات الترقيم، وأظهرت المهمة الفرعية تحديًا أما تحصيل الدرجات بطريقة مقننة في بعض السياقات. ولم تعد الإملاء جزءًا من الأداة الأصلية، ولكنها أُستخدمت في بعض الدول، وكانت ملائمة.

4.4 المفردات واللغة الشفوية

1.4.4 الوصف

إن الفهم القرائي يشمل أكثر من مجرد تعرف الكلمة. فنحن لكي نكوِّن المعني نحتاج إلى ربط هذه الكلمات معًا حتى نقرأها بتمثيلات المعنى المرتبط بالكلمة في أذهاننا، ويرتبط معنى الكلمات الفهم العام للغة الشفوية للأشخاص (Kamhi & Catts, 1991; Nation, 2005; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). وتشير المفردات إلى قدرة التلاميذ على فهم معنى الكلمات عندما نسمعها أو نقرأها (استقبالية)، وكذلك استخدامها عندما نتحدثها و نكتبها (إنتاجية)، ولذلك اقترح الخبراء في القراءة معرفة المفردات بنسبة 90% و95% من تلك الواردة في نص لازمة للفهم (Nagy & Scott, 2000) ، وليس مفاجئًا أنه في الدراسات التتبعية أوضحت النتائج بشكل متكرر أن المفردات تؤثر على التحصيل في القراءة ،وهي مؤشر قوي على التحصيل في القراءة (Muter et al., 2004; Roth, Speece, & Cooper, 2002; Share & Leiken, 2004).

2.4.4 قياس المفردات

بالرغم من أنه لاتوجد مهمة فرعية واحدة في الأداة الأصلية لتقييم مهارات القراءة في الصفوف الأولى EGRA تقيس المفردات بشكل مباشر، توجد مهمة فرعية اختيارية، وغير محددة بوقت تقيس المهارات الاستقبالية للغة من خلال كلمات مفردة أو عبارات ترتبط بأعضاء الجسم، والأشياء الشائعة والعلاقات المكانية. استُخدمت هذه المهمة الفرعية في سياقات قليلة، ولكن لم تكن من خلال عمليات المراجعة والتحقق من الصدق بواسطة نفس مجموعة الخبراء الذي قاموا بنفس الأشياء للمهام الفرعية الأخرى.

بالإضافة إلى ذلك، فإن فهم المسموع، وهو مهمة فرعية للأداة الأصلية لتقييم مهارات القراءة في الصفوف الأولى، تقيم الفهم في اللغة الشفوية العامة ومن ثم، بشكل غير مباشر المفردات الشفوية التي تُبنى جزئيًا. وبالنسبة لهذه المهمة الفرعية، يقرأ المقيمون للأطفال قصة قصيرة حول موضوع مألوف لديهم، ثم يطلبون من الأطفال الإجابة عن ثلاثة إلى خمسة أسئلة عما استمعوا إليه. فتستخدم المهمة الفرعية لفهم المسموع بشكل أولى بالعلاقة مع المهمة الفرعية للفهم القرائي (انظر الفهم، القسم 4. 6)؛ لتحديد ما إذا كانت صعوبات الفهم تنشأ أوليًا من مهارات القراءة الضعيفة أو من الفهم الضعيف العام للغة.

5.4 الطلاقة

1.5.4 الوصف

الطلاقة هي " القدرة على قراءة نص بسرعة ودقة قراءة تعبيرية" (NICHD, 2000 pp. 3–5)، ووفقًا لـ Snow and the RAND Reading Study Group (2002):

> يمكن فهم الطلاقة على أنها سابقة على / وتابعة للفهم. قد تعتمد بعض نواحي الطلاقة في القراءة التعبيرية على الفهم الدقيق للنص. ومع ذلك، فإن بعض مكونات الطلاقة – كالتعرف السريع، والفعال للكلمات، وعلى الأقل بعض النواحي للتحليل المرتبط بالمعاني (كمعالجة تركيب الجملة) من المتطلبات السابقة للفهم. (p. 13)

يُمكن أن نُعرف الطلاقة على أنها الجسر أو الربط بين تعرف الكلمة وفهم النص، بينما فك الرموز هو الخطوة الأولى نحو تعرف الكلمة، فلابد للقراء أن ينموا قدرتهم على فك الرموز إلى الحد الذي تصبح فيه آلية، عندئذ يتحول الانتباه من الحروف المفردة والكلمات إلى الأفكار نفسها التي يتضمنها النص ي نص مكرر (Armbruster et al., 2003; Hudson, Lane, & Pullen, 2005; LaBerge &Samuels, 1974). قد تكون الآلية مهمة جدًا بسبب حدود الذاكرة قصيرة المدى. فإذا فككنا الرموز ببطء شديد فإننا سنركز انتباهنا على كل كلمة، فلن يكون هناك مساحة كافية في الذاكرة العاملة للجملة كلها. فما يحدث هو أننا ننسى بداية النص عندما نصل إلى نهايته. فإذا لم نتمكن من أن نتابع النص كله في ذاكرتنا العاملة حالاً فلن نتمكن من استخلاص المعنى داخل النص (Abadzi, 2006; Hirsch, 2003).

الطلاقة مثل الفهم من مهارات التفكير العليا تتطلب المعالجة المركبة والمتناغمة لفك الرموز، وتعرف معنى الكلمات، ومعالجة تركيب الجملة والنحو، واستنتاج التتابع السريع للكلمات (Hasbrouck & Tindal, 2006)، وهي تتطور ببطء بمرور الوقت ومن خلال التعرض الكبير للنصوص وفك الرموز.

كشفت العديد من الدراسات أن الفهم القرائي يرتبط بالطلاقة خاصةً في المراحل الأولى (Fuchs, Fuchs, Hosp, & Jenkins, 2001) وللأفراد الذين يتعلمون القراءة في

اللغة التي يتحدثونها ويفهمونها. فعلى سبيل المثال، أوضحت اختبارات الطلاقة في القراءة الشفوية، والتي تُقاس عن طريق تقييمات ذات وقت محدد لعدد الكلمات الصحيحة في الدقيقة، أن هناك ارتباط (0.91) المهمة الفرعية للفهم القرائي، واختبار ستانفورد للتحصيل (Fuchs et al., 2001). وتؤكد البيانات من تطبيقات تقييم مهارات القراءة في الصفوف الأولى EGRA من سياقات ولغات متعددة العلاقة بين هذين العاملين constructs (Bulat et al., 2014; LaTowsky, Cummiskey, & Collins, 2013; Management Systems International, 2014; Pouezevara, Costello, & Banda, 2012; among many others). تتناقص أهمية أهمية الطلاقة كمقياس تحليلي في الصفوف التالية حيث يتعلم التلاميذ القراءة بطلاقة وكفاءة. وكلما يصبح التلاميذ أكثر كفاءة، وطلاقة. لذلك تصبح المفردات مؤشرًا أكثر أهمية للنجاح الأكاديمي (Yovanoff, Duesbery, Alonzo, & Tindall, 2005).

وهنا يطرح تساؤلاً وهو ما السرعة الكافية؟ بالرغم من أنه نظريًا يحتاج القراء إلى حد أدنى للطلاقة حتى يتمكنوا من فهم نص متصل، فإن العلامات المرجعية للطلاقة تتنوع وفقًا للصف الدراسي واللغة. فاللغة التي بها كلمات أقصر في المتوسط مثل اللغة الانجليزية و الإسبانية، والتي تسمح للتلاميذ قراءة كلمات أكثر في الدقيقة أكثر من اللغات مثل السواحيلي حيث تتكون الكلمات من 10 – 15 حرف أو 20 حرفًا. بمعنى أنه كلما ازداد طول الكلمات وكلما كونوا معنى كلما كانت الكلمات التي يحتاجونها أقل للقراءة في الدقيقة لتشير إلى كفاءتهم في القراءة.

2.5.4 قياس الطلاقة

نظرًا لأهمية الطلاقة للفهم، فإن المقياس الأكثر مباشرة في أداة تقييم مهارات القراءة في الصفوف الأولى EGRA I ، هو المهمة الفرعية للطلاقة في القراءة الشفوية مع الفهم، وهو مكون أساسي في الأداة. يُعرض على الأطفال قطعة قصيرة مكتوبة عن موضوع مألوف لديهم، ويُطلب منهم قراءتها بصوت مرتفع بسرعة وبدقة. تتألف الطلاقة من السرعة والدقة والقراءة التعبيرية. وتلك المهمة الفرعية للطلاقة في القراءة الشفوية مهمة محددة بزمن وتقيس السرعة والدقة فيما يتعلق بعدد الكلمات التي تُقرأ بشكل صحيح في الدقيقة. مع الأخذ في الاعتبار أن هذه المهمة الفرعية لا تقيس القراءة التعبيرية.

بالإضافة إلى هذه المهمة الفرعية للطلاقة في القراءة الشفوية، فإن العديد من المهام الفرعية في أداة تقييم مهارات القراءة في الصفوف الأول EGRA التي نوقشت أعلاه ذات وقت

محدد و تُعطى الدرجات وفقًا للسرعة و الدقة وعلاقتها بعدد الحروف الي تُقرأ بشكل صحيح (أو الأصوات والمقاطع) ، أو الكلمات في الدقيقة: تعرف اسم الحرف ، وتعرف صوت الحرف ، و قراءة الكلمات غير المألوفة ، وقراءة الكلمات المألوفة ؛ ولأن القراء يصبحون أكثر طلاقة بشكل متزايد حيث تنمو مهارات القراءة لديهم، فإن التقييمات ذات الوقت تساعد على رصد التقدم في هذه المقاييس كلها ،وتوضح أين هم التلاميذ في طريقهم إلى القراءة الماهرة.

6.4 الفهم

1.6.4 الوصف

الفهم هو الهدف النهائي للقراءة. فهو يساعد التلميذ على التوصل إلى المعنى مما يقرأون ويستخدمونه ليس فقط من أجل المتعة؛ ولكن لتعلم أشياء جديدة أيضًا، خاصة المحتوى الأكاديمي. والفهم القرائي مهمة مركبة بشكل كبير، وتتطلب من التلاميذ تكوين المعنى من النص. يعتمد الفهم القرائي على التناغم بين الدافعية و الانتباه و الاستراتيجيات و الذاكرة و معرفة الموضوع ، و المعرفة اللغوية ، و المفردات ، وفك الرموز ، و الطلاقة ، و أمور أخرى ، ولذلك تمثل مكونًا صعبًا لأي تقييم أن يقيسه مباشرة (Snow & the RAND Reading Study Group, 2002)

2.6.4 قياس الفهم القرائي

تقيس أداة تقييم مهارات القراءة في الصفوف الأول EGRA الفهم من خلال المهمة الفرعية للفهم القرائي، حيث يقرأ الأطفال قطعة بصوت عال من أجل الطلاقة في القراءة الشفوية. وبعد أن يقرأ الأطفال القطعة بصوت عال، يُطرح عليهم 3 – 5 أسئلة فهم مباشر وغير مباشر، والتي يمكن الإجابة عنها من خلال قراءة النص فقط. يمكن أن يُسمح باستخدام استراتيجية انظر ثانية – حيث يرجع إلى القطعة للإجابة عن الأسئلة-لتخفيف العبء عن الذاكرة، ولا تُستخدم في الأداة الأصلية. ويوجد بديل آخر لمهمة فرعية، وهي متاهة الفهم حيث يُطلب من الأطفال تعرف كلمة من بين بدائل عديدة؛ لإكمال الجملة وهذا البديل لديه القدرة الكامنة لقياس الفهم أكثر، ولكن لم يستخدم بشكل كبير في تطبيقات تقييم مهارات القراءة في الصفوف الأولى EGRAحتى الآن.

صُممت أداة تقييم مهارات القراءة في الصفوف الأولى EGRA؛ لتكون مقياسًا واسع النطاق، ومقننًا يعكس تصميمها أن الأبحاث لم تنتج حتى الآن وسيلة مؤكدة؛ لقياس مهارات التفكير العليا وبخاصة مهارات الفهم باتساق ودقة وبطريقة مقننة،

يمكن قبولها كمقياس يتمتع بالصدق والثبات. فبعض التقييمات البديلة المتوفرة كنشاط إعادة سرد القصة حيث يُطلب من التلاميذ أن يسردوا ما يتذكرونه من القصة، وستسمح للمعلم أن يفحص مستويات أعمق للفهم من خلال أحاديث التلاميذ التي سيجد بها صعوبة، وذلك من خلال طرح أسئلة مفتوحة على التلاميذ، والتقدير غير الموضوعي للدرجات ومع ذلك توجد بعض البدائل تحت الدراسة يُؤمل أن يستمر قياس الفهم يستمر في التحسن، حيث إن هذه المهارة واحدة من أهم مقاييس النجاح في القراءة.

5 تصميم دراسة تقييم مهارات القراءة في المراحل المبكرة EGRA

5-1 إرشادات تصميم دراسة تقييم مهارات القراءة في المراحل المبكرة

يؤسس هذا القسم تصميمات الدراسة والمبادئ لتقييمات باستخدام تصميم دراسات تقييم مهارات القراءة في المراحل المبكرة. ويُراعى في هذا البحث ما يُراعى في أي بحث من حيث غرض التقييم و أسئلة البحث التي تقوده ، و توجه التصميم للدراسة. تُراعى هذه الأسئلة و القرارات بالتعاون التام بين الممولين و شركاء الحكومة المضيفة لضمان إمكانية التنفيذ وقرارات ملائمة التصميم.

متطلبات عينات الدراسة لكل تصميمات الدراسة

تصميمات دراسة تقييم مهارات القراءة في المراحل المبكرة و العينات ترتبط معًا،ومع ذلك،كل العينات لدراسات تقييم مهرات القراءة في المراحل المبكرة بغض النظر عن تصميم الدراسة تتطلب مراعاة مايلي فيما يتعلق بالبيانات و التحليل و كتابة التقرير:

- تعريف صريح لمجتمع الدراسة موضوع الاهتمام موضحًا به كل الاستثناءات، و موثق بالتفصيل قبل تحديد العينة ،ويشمل;
- أوزان العينة محسوبة بشكل جيد(يستخدم لوزن العينة التي تؤخذ من المجتمع الأصلي)
- تحليل مركب للدراسة يكون مصممًا بشكل سليم (يوضح طريقة اختيار العين و وأثر العينة العنقودية)
- برنامج الكتروني إحصائي يمكنه تحليل العينة المركبة (على سبيل المثال SPSS, Stata, SAS)
- تحليلات استدلالية تستخدم أسلوب تحليل مركب سليم.

تعتبر الخطوة الأولى التي تحدد التصميم هي طرح السؤال التالي و الذي لايجب أن يُغفل نظرًا للتعجل في تطبيق الدراسة،و هو ما الغرض من دراسة تقييم مهارات القراءة في المراحل المبكرة؟ وتقع دراسات تقييم مهارات القراءة في المراحل المبكرة في واحدة من التصنيفات الثلاثة التالية:

1. **تقييم اللقطة:** للحصول على تشخيص لأداء التلاميذ في نقطة زمنية محددة.
2. **تقويم الأداء:** لتقييم ما إذا كانت التغييرات التي حدثت في أداء التلاميذ على مدى فترة زمنية محددة ، وبناء على تقييم مبدئي حدث و تقييم متابعة أخر تم.
3. **تقييم الأثر:** لتقييم أثر برنامج أو تدخل على أداء المتعلمين في فترة زمنية محددة بناء على مقارنة المجموعات التجريبية و المجموعات الضابطة.

يشرح القسم 5 – 2 هذه الأغراض و أكثر التصميمات الدراسة ملائمة و التي ترتبط مع كل نوع من الدراسة. يوجد وصف لكل بديل ن وكذلك المرفقات التي تقدم معلومات مفصلة حول العينة المتعلقة بكل تصميم.

2-5 تصميم الدراسة وفقًا للغرض منها

1.2.5 تقييمات اللقطة و تقييمات الأداء كتصميمات للدراسة

كما هوموضع أعلاه، يسعى تقييم اللقطة إلى عرض مؤشر محدد أو متغير محدد مثل الطلاقة في القراءة الشفوية ، في نقطة زمنية محددة، وتقوم دراسات تقييم الأداء بنفس الشئ و لكن في نقط زمنية متعددة. لاتسمح تقييمات اللقطة و تقييمات الأداء بعزو النتائج لتدخل محدد بعينه.

اعتبارات تُؤخذ في الاعتبار للعينة : تقييمات اللقطة و تقييمات الأداء

تسعى معظم تقييمات اللقطة لتقييم مهارات القراءة في المراحل المبكرة لتقدير القدرة على القراءة لمجتمع معين من السكان. تستخدم هذه الأنواع من التقييمات العينة العنقودية / المركبة (يوجد وصف للعينة في **الملحق ب ، ج)**

وعلى النقيض لذلك، يعتمد أسلوب تحديد العينة لتقييمات الأداء على أسئلة التقييم التي يُجاب عنهاو مستوى الموارد المتاحة للإجابة عن هذه الأسئلة. على سبيل المثال، إذا كانت أسئلة التقييم تسأل " ما هو أداء التلاميذ المستفيدين مقارنة بالعلامات المرجعية لأداء الوطني ؟ فمن المحتمل أنم يتضمن تقييم الأداء وجودة عينة عشوائية بسيطة أوعينة عنقودية / مركبة ، وذلك لضمان أن العينة ممثلة للمجتمع المستفيد كله. ولكن إذا كانت الموارد المتاحة محدودة، قد يتطلب تقييم الأداء كبديل أخذ عينة أصغر، عينة غير ممثلة و بعد ذلكك يُقدم شرح لحدود البيانات و و النتائج المستخدمة للاستدلال الإحصائي بالنسبة للمجتمع كله. قد يتم هذاللمراجعة الداخلية ، ولكن لا ينبغى أن يحدث أذا كانت الدراسة تسعى إلى الوصول إلى تعميمات بالنسبة للمجتمع موضوع الاهتمام.

2.2.5 تقييم الأثر كتصميم للدراسة

يختلف تقييم الأثر عن تقييم الأداء في أن تقييم الأثر يسعى إلى عزل الاثر لتدخل ما على ناتج معين من التأثيرات الأخرى وذلك بمقارنة النواتج لمجموعة تتلقى التدخلات (مجموعات متعددة تتلقى تدخلات مختلفة) بمجموعة أخرى تلعب دور المجموعة الضابطة.

و بكلمات أخرى، يسعى تقييم الأثر باستخدام تحديد الواقع المضاد (أنظر قائمة المصطلحات) إلى الكشف عن كم التغير الذي حدث في ناتج معين و الذي يمكن أن يُنسب و بثقة إلى برنامج معين. ويسعى عدد متزايد من تقييمات الأثر إلى فهم أثر تدخلات متعددة لتحسين القراءة المبكرة على درجات تقييم مهارات القراءة في المراحل المبكرة . يوجد نوعات لتصميمات دراسة تقييم الأثر – التصميم التجريبي و التصميم شبه التجريبي- حيث يوجد لهما وصف في المرع أسف. توجد المعلومات المفصلة عن تحديد العينة لتقييمات الأثر في **الملحق د**.

نوعا تقييمات الأثر

التصميمات التجريبية

يجب أن تبدأ التصميمات التجريبية (يُشار إليه أحيانًا بالمجموعات الضابطة التجريبية RCTs) قبل أن يبدأ التدخل. فهي تتطلب بيانات دراسة للأساس و مشاركين يُختارون عشوائيًا (أو مدارس ، إدارات ، أو نوع مشابه من الوحدات) ويصنفون إلى مجموعات مستفيدة وأخرى غير مستفيدة (مجموعات مقارنة) لابد أن يكون لكل الأفراد (أو الوحدات) أحتمالية التعيين إلى المجموعة المستفيدة أو المجموعة غير المستفيدة (المقارنة)، ولابد ان يكون حجم العينة كبيرًا بالقدر الذي يسمح بالقارنة بين المجموعتين مع وجود حد أدنى مقبول يمكن رصده لحجم الأثر(MDES انظر الملحقD).

التصميمات شبه التجريبية (QEDs)

تبدأ التصميمات شبه التجريبية عادة قبل بدء التدخل ، ولكن ليس من الضروري أن يتعين عليها ذلك، طالما توجد بيانات لدراسة الأساس لمجموعة المستفيدين و مجموعة المقارنة. وفي التصميمات شبه التجريبية ، فإن المشاركين في التدخل (أو المدارس، الإدارات ، أو أي نوع أخر للوحدة) لا يُعينون بشكل عشوائي إلى مجموعة المستفيدي و مجموعة غير المستفيدين. و عوضًا عن ذلك، يعين المشاركون أنفسهم إلى مجموعة التدخل، أو أن المنفذ يختار المستفيدينباستخدام مجموعة من معايير الاختيار. يتولد في كلا النوعين من بدائل الاختيار المذكورة نوعًا من التحيز في الاختار، وتحاول التصميمات شبه التجريبية تقليل أو التحكم في تحيز الاختيار من خلال مجموعة من الأساليب الإحصائية .

بينما تسمح الأساليب شبه التجريبية بعزو النتائج، إلا أنها تعتمد على افتراضات إحصائية قد لاتكون صحيحة دائمًا، وبالتالي،قدتكون أقل من التصميمات التحجريبية من حيث الثبات و المتانة(حيث يُفترض أنهما قد أُتما بشكل سليم)تستخدم بعض الأنواع الشائعة من التصميمات شبه التجريبية لدراسات تقييم الأثر في تقييم مهارات القراءة في المراحل المبكرة هي قطع الانحدار و تصميمات مزاوجة درجة الاتجاه(أنظر المصطلحات) .

مستويات التعيين لتنقييمات الأثر

إذا تم أخذ القرار باستخدام تقييم الأثر لقياس النتائج، فلابد من تحديد مستوى التعيين وماإذا كانت الدراسة ستكون تتبعية (تتبع نفس التلاميذ في فترة زمنية)، أو شبه تتبعي (تتبع نفس المعلمين أو المدارس في فترة زمنية) ، أو مابين القطاعات (إعداة تحديد عينة المدارس المختلفة و التنلاميذ في كل مرة تُجمع فيها البيانات) مرة أخرى، فإن هذه القرارات تعتمد على الغرض من الدراسة ، ةولكنها أيضًا تعتمد على الكيفية التي سيُنفذ بها التدخل (البرنامج ، المشروع، النشاط) الذي يجرى تقييمه.مايلي هو مجموعة متعددة من المستويات التي قد يحقق تدخل ما فوائد:

- **الإدارة ، النطاق او مستوى الوحدة الإدارية:** على سبيل المثال، تقديم تدريب للمعلمين لكل المعلمين في إدارة تعيمية معينة.
- **مستوى المجتمع:** على سبيل المثال،تنفيذ برامج توعية في المجتمع لمساعدة المجتمع على المشاركة في المدرسة أو البدء في إنشار مركز للقراءة للمجتمع.
- **مستوى المدرسة :** على سبيل المثال، تزويد المدارس بالكتب و المواد أو بعض الفوائد مباشرة على مستوى المدرسة ، تستهدف بعض المدارس في وحدة إدارية محددة و ليست كل الوحدات.
- **مستوى التلميذ:** على **سبيل** المثال،تقديم منح دراسية لبعض التلاميذ - دون غيرهم - أو دعم مادي.

قد توفر التدخلات فوائد على مستويات متعددة. على هذا النحو، من المهم عند تحديد مستوى التعيين للمجموعات المستفيدة والمقارنة لجعل التعييانات على أساس أعلى مستوى من التدخل الذي يخطط له المنفذ. على سبيل المثال، إذا كان البرنامج يخطط لتقديم تدريب المعلمين على "مستوى منطقة" وكذلك الكتب أو المواد في "مستوى المدرسة،" ثم "مستوى منطقة" سيتم تعيينه لمجموعة المستفيدين و مجموعة المقارنة. لهذا السبب، فإنه من الأهمية بمكان أن فرق تقييم العمل بشكل وثيق جدا إلى جانب المنفذين لتخطيط تصميم التقييم، والمنفذ يصمم التدخل.

بدائل تصميم الدراسة لتقييمات الأثر

بعد ذلك، يجب أن تقرر ما إذا سيتم تتبع الطلاب تتبعيا أو استخدام تصميم شبه تتبعي أو تصميم مستعرض. تمثل المعلومات التالية المعلومات المطلوبة للمساعدة في تحديد القرار الملائم:

1. ما الغرض من التقييم , و ما أسئلة التقييم؟

التصـــميم التتبعي: إذا كان أســـئلة البحث تتطلب فهمًا لتغيرات محددة لكل تلميذ يتلقى التدخل، فإن استخدام التصميمي التتبعي سيسمح للباحثين برصد أي تغيرات حدثت بشكل محدد في كل تلميذ مشـــارك في التدخل ومع ذلك، ان يتمكن الباحثون من عزو هذ النتائج إلى المجتمع الذي يتلقى التدخل. هذا النوع من التصـــميمتات هو الأفضـــل للدراســـاتن الاستطلاعية.ن وللاستمرار في التقييمات الداخلية للتدخل.

إذا كان الهدف من الســـؤال البحثي هو التحقق في التغييرات المعقدة التي تحدث في مدارس محددة، فإن الدراسة شبه التتبعية هي الأنسب، حيث يجري زيارة المدارس نفسها ولكن يتم أخذ عينة عشـــوائية من الطلاب داخل المدارس نفســـها. وهذا النوع من الدراســـة يســـمح للباحثين معرفة أي تغيرات في المدارس المحددة التي يجري تقييمها لكنها لن تسمح للباحثين تعميم النتائج التي توصـــلوا إليها إلى أعداد أكبر من المدارس في الدراســـة. هذا النوع من التصميم هو أفضل للدراسات الاستطلاعية، ولمواصلة التقييمات الداخلية للتدخل.[8]

التصـميم ذو المقطع المسـتعرض: إذا كان الغرض من السـؤال البحثي هو تقييم كيف يمكن لمجتمع (انظر المصـطلحات) من المدارس ومجتمع من الطلاب في تلك المدارس أن يتغير بسبب هذا التدخل، ثم هناك حاجة إلى سلسلة من عينات المقطع المستعرض، والتي يتم من خلالها أخذ عينات مختلفة تماما من المدارس والطلاب في كل مرة يتم فيها جمع البيانات. يتم أخذ عينـات من المـدارس والطلاب من داخـل مجتمع المـدارس والطلاب الـذين يتلقون التدخل.

وهذا التصميم لا يسمح للباحثين لكشف بحسم التغييرات الدقيقة التي وقعت داخل مدارس أو طلاب بعينها (لأنها مختلفة في كل مرة). ولكن التصـــميم ســـوف تســـمح للباحثين لتعميم التغيرات التي يمكن أن تعزى للمجتمع ككل. هذا النوع من الدراســـة هو أفضـــل للتقييم الخارجي لدراســـة التدخل، حيث يحاول الباحثون الفهم بشـــكل أفضـــل لأثر التدخل على المجتمع كله الذي حصـل على التدخل، بدلا من الأثرالذي أحدثه التدخل على مدارس محددة أو طلاب محددين.

[8] قد تكون التصميمات التتبعية وشبه التتبعية مفيدة لدراسات عدم التدخل، كما تسمح التصميمات للباحثين بمتابعة التغيرات لوحدة واحدة (على سبيل المثال، التلميذ) في الفترات التي لايتم فيها تقديم أي تدخل.

2. مادرجة سهولة تتبع نفس التلاميذ و المعلمين و المدارس ،...؟

على ســـبيل المثال، إذا كان الفريق الي يقيم برنامج في بلد ما حيث إما يتم تعيين لكل طالب رقم هوية يبقى معه / معها حتى عندما انه / انها تتحرك، أو على الأســر أن تســجل نفســها لدى الحكومة لأغراض الضرائب أو التعداد، فأنه قد لا يكون من الصعب جدا تعقب الطلاب بشكل فردي. إذا لم تكن هذه هي الحالة، رغم ذلك، وإذا كانت المجتمعات تميل إلى أن تكون متحركة جدا ومعدلات التســـرب تميل إلى أن تكون عالية جدا، فإن تتبع نفس الطلاب مع مرور الوقت يمكن مكلفًا جدا.

3. ما مستويات الموارد المتاحة للتبع التلاميذ و المعلمين و المدارس،...؟

حتى إذا كان التتبع من الســـهل نســـبيا في بلد ما، فإنه لا يزال عادة أكثر تكلفة واســـتهلاكا للموارد من أخذ لقطة لبعض الطلاب. على ســـبيل المثال، قد يكون من المســـتحيل أن يتم التتبع، وعلى هذا النحو، الإفراط في العينة كدراسة للأساس عادة ما يكون ضرورة.

4. مامستوى المتانة والدقة المطلوبة في النتائج؟

إذا كان المانح أو المنفذ يحتاج نتائج دقيقة عن التســـرب، على ســـبيل المثال، قد يكون من الضـــروري دراســـة تتبعية. ومع ذلك، إذا كانت معدلات التســـرب التقريبية كما وردت في تقريرالمعلمين أو المدارس ستخدم هذا الغرض، فقد تكون الدراسة المستعرضة كافية.

6 تصميم أداة تقييم مهارات القراءة في الصفوف الأولى إجراء التطويع للأداة وتعديل التطويع

يناقش هذا القسم الشكل والمتطلبات الضرورية؛ لتصميم أو تعديل أداة تقييم مهارات القراءة في الصفوف الأولى؛ لتلائم السياق الذي ستُطبق فيه الأداة. يهدف هذا الفصل تعريف القراء مختلف المهام الفرعية التي يمكن تضمينها في أداة تقييم مهارات القراءة في الصفوف الأولى وذلك من خلال تقديم وصف للمهام الفرعية والمبادئ التوجيهية لبناء الأداة.

1.6 ورشة عمل تطويع الأداة:

الخطوة الأولى في تطويع الأداة هي تنظيم ورشة عمل داخل البلد، وعادة تدوم حوالي خمسة أيام عمل. وهذا القسم الفرعي يستعرض خطوات إعداد وتنفيذ ورشة عمل تطويع أداة تقييم مهارات القراءة في الصفوف الأولى، ويُعطى تصورًا عامًا عن الموضوعات التي ستُغطى.

تُعقد هذه الورشة لتطويع الأداة في البلاد التي ستُطبق فيها في بداية عملية تطوير الاختبار (أو تعديله) لأدوات تقييم مهارات القراءة في الصفوف الأولى. فهي توفر فرصة للدول التي ستُطبق فيها؛ لبناء صدق المحتوى (انظر المصطلحات) في الأداة من خلال وجود مسؤولين حكوميين وخبراء المناهج الدراسية، وغيرها من الجماعات ذات الصلة لفحص المهام الفرعية لتقييم مهارات القراءة في الصفوف الأولى، وإصدار الأحكام حول مدى ملاءمة كل بند من بنود الأداة؛ لقياس مهارات القراءة المبكرة لتلاميذهم كما هو محدد في وثائق المناهج الدراسية أو غيرها من المبادئ التوجيهية لتوقعات أو معايير التعلم. وكجزء من عملية التطويع، فإن الأفراد المشاركين في ورشة العمل يطوعون أداة تقييم مهارات القراءة في الصفوف الأولى حسب الضرورة، ويعدون البنود الملائمة لكل مهمة فرعية في الاختبار. يضمن هذا الأسلوب بأن التقييم يتمتع بالصدق الظاهري (انظر المصطلحات). وبعد ورشة العمل، يجرى تجريب الأداة من خلال فرق عمل في أحد المدارس، وهذا أمر ضروري. وتناقش الاختبار التجريبي للأداة والعمل الميداني بالتفصيل في القسم 9.

[9] إلى أي مدى بنود أداة تقييم مهارات القراءة في الصفوف الأولى ممثلة للبناء الذي يتم قياسه، والذي يُعرف على أنه دليل مرتبط بمحتوى الاختبار (أي مهارات القراءة المبكرة في بلد معين).

للحصول على معلومات إضافية على الجودة الفنية للأداة، وثباتها بما في ذلك المبادئ التوجيهية لإجراء فحص لجودة الأداة من الناحية الفنية وثباتها، لذا يُرجى الاطلاع على المرفق D والقسم 9.1.2 من هذا الدليل.

أهداف ورشة العمل:

- تقديم خلفية بحثية أساسية عن مكونات الأداة لكل من المسؤولين الحكوميين، والمختصين في المناهج والتقييم المحليين.
- • تطويع الأداة لتلائم الظروف المحلية باستخدام المبادئ التوجيهية؛ لكتابة البنود الواردة في هذا الدليل، بما في ذلك:
  - ترجمة تعليمات الأداة.
  - تطوير نسخ بلغات ملائمة، إذا لزم الأمر.
  - تعديل مكونات الكلمة، وقطع القراءة؛ لتعكس الكلمات والمفاهيم المناسبة محليًّا وثقافيًّا.
- مراجعة إجراءات الموافقة المسبقة (الكبار)، أو موافقة (الأطفال)، ومناقشة أخلاقيات البحث والعمل مع الناس، ولا سيما وخاصةً مع الأطفال.

الشكل 7 يحدد بوضوح أكبر الفروق بين ورش عمل التطويع والتعديل. إذا كان يُوضع أداة تقييم مهارات القراءة في الصفوف الاولى لبلد معين لأول مرة، فإنه يعتبر تطويع للأداة؛ وأما إذا كان قد طُبق قبل ذلك أداة تقييم مهارات القراءة في الصفوف الاولى في البلد نفسها، فهي ورشة عمل لتعديل الأداة.

الشكل 7 الفروق بين ورش عمل تطويع أداة تقييم مهارات القراءة في الصفوف الأولى EGRA وتعديل

تعديلاداة موجودة	تعديل اداة جديدة
التحليل اللغوي	التحليل اللغوي (اختياري)
انتقاء البند	إعادة ترتيب البنود والتوزيع العشوائى و
التحقق من الأدوات	التحقق من الأدوات
الاختبار القبلى	الاختبار القبلى

1.1.6 نظرة عامة على اعتبارات تخطيط ورشة عمل:

سواء تصميم أداة تقييم مهارات القراءة في الصفوف الأولى الخاصة بكل بلد من البداية (تطوير) أو من النموذج الموجود (تعديل)، فإن فريق الدراسة بحاجة؛ للتأكد من أن الأداة مناسبة للغة أو اللغات المستهدفة لقياسها، ومستويات الصف للمشاركين في الدراسة، وأسئلة البحث التي في متناول اليد.

إن تطوير الأداة يتطلب اختيار المهام الفرعية المناسبة ، والبنود للمهام الفرعية. وتشتمل على الاعتبارات الآتية:

- يجب أن يسمح جدول الأعمال باختبار ميداني محدود للأداة كما يتطور، والذي يتضمن أخذ المشاركين (إما مجموعات فرعية أو كل الفريق) إلى المدارس القريبة؛ لاستخدام مسودة الأداة مع الطلاب. وهذا الاختبار الميداني يتيح للمشاركين اكتساب فهم أعمق للأداة، وهو بمثابة اختبار سريع للبنود؛ لتحديد أي تغييرات واضحة قد تكون هناك حاجة لها (مثل: المراجعات لاختيارات الإجابة الغامضة، أو المفردات الصعبة أكثر من اللازم).

- إن تحليل اللغة ضروري لصياغة البنود، والذي يمكن أن يتحقق في وقت مبكر، جنبًا إلى جنب مع ترجمة التعليمات، التي يجب أن تبقى مقننة وموحدة في جميع البلدان. لذا يجب على الخبراء، ومجلس المراجعة المؤسسي **IRB** مراجعة التعليمات؛ لضمان سلامتها من الناحية الأخلاقية ودقتها لكل قسم. لأغراض التقنين، ويجب أن تتاح لجميع الطلاب نفس الفرص بغض النظر عن المقيم أو السياق؛ لذلك، فإنه مطلوب للحفاظ على التعليمات واحدة في كل البلدان والسياقات.

- إذا لم يكن ممكنًا أن تعقد ورشة العمل في المنطقة الذي سيجري فيه الاختبار، يجب على فريق الدراسة اتخاذ الترتيبات اللازمة لاختبار ميداني بعد ذلك، أو العثور على مجموعة من الطلاب القريبين الذين يتحدثون اللغة، والذين هم على استعداد للمشاركة في الاختبار الميداني خلال ورشة العمل، وفي كلا البديلين، فإن فريق الاختبار الميداني يحتاج إلى رصد النتائج، وتقديم تقرير إلى المجموعة الكاملة.

- كتابة القصة هي في العادة الجزء الأكثر صعوبة في التطويع، ولذلك فمن المهم عدم ترك هذه المهمة الفرعية حتى اليوم الأخير، وتنطوي هذه الخطوة على سؤال يسأل الخبراء

- المحليين لكتابة القصص القصيرة باستخدام الكلمات المناسبة على مستوى الصف، وكذلك كتابة أسئلة الفهم المناسبة لتصاحب القصص. فالقصص والأسئلة غالبًا ما تحتاج إلى أن تترجم إلى اللغة الإنجليزية أو لغة أخرى للمراجعة من قبل خبراء إضافيين في القراءة في الصفوف الأولى، والمراجعة عدة مرات في لغة التقييم قبل أن يتمكنوا من وضعها في الصورة النهائية.

2.1.6 من يشارك؟

مجموعة مكونة من موظفي الحكومة، ومدربي المعلمين والمدرسين السابقين والحاليين، وخبراء اللغة من الجامعات المحلية توفر مزيجًا جيدًا من الخبرة والعناصر المهمة لعملية التطويع، ومع ذلك، وسيحدد عدد المشاركين في ورشة عمل التطويع توافر الموظفين الحكوميين للمشاركة. وينصح بوجودهم من أجل بناء القدرات والمساعدة على ضمان استدامة للتقييم. إن عدد المشاركين يعتمد في جزء منها على عدد اللغات التي تشارك في عملية التطويع لدراسة معينة، ولكن بصفة عامة، و30 مشاركًا هو الحد الأقصى الموصي به من المشاركين.

ومن بين المشاركين في ورشة العمل دائما1 :هم على النحو التالي:

1. خبراء اللغة: وذلك بهدف التحقق من التعليمات التي ترجمت ولتوجيه عملية مراجعة البنود المختارة، ولدعم كتابة قصة أو التعديلات .
2. الممارسون غير الحكوميين: أكاديميون (المتخصصون في القراءة، على وجه الخصوص)، والمعلمون الحاليون أو السابقون (مع تفضيل معلمي القراءة.
3. مسؤولون حكوميون: الخبراء في تطوير المناهج الدراسية، والتقييم.
4. السيكومتريون: خبراء في وضع الاختبارات.

من الناحية المثالية، ويشارك الموظفون الحكوميون الرئيسيون في كل عمليات التطويع بالكامل وتدريب المقيمين، وعملية تجريب (تتم خلال شهر واحد كلها، وهذا يتوقف على عدد من المدارس التي تمثل عينة الدراسة). هناك حاجة إلى وجود الاتساق بين المشاركين حتى يتقدم العمل إلى الأمام بوضوح ونزاهة، وفي نفس الوقت، يتم بناء القدرات، وتتحقق الاستمرارية .

ويقوم بتيسير ورشة العمل فريق مكون من اثنين على الأقل من الخبراء. يجب أن يكون ميسرو ورشة العمل متمكنين من مكونات ومبررات التقييم، وكذلك يكونون بارعين في العمل في مجموعة متنوعة من البلدان والسياقات.

- **خبير التقييم:** مسئول عن قيادة تطويع (إذا كانت عملية تطويع أو تعديل التطويع) للأداة، وفيما بعد يوجه تدريب المقيمين وجمع البيانات، لديه خلفية عن البحوث المسحية، وفي تصميم التقييمات الاختبارات. وتتضمن هذه الخبرة الإحصاء الأساسي، وخبرة في برامج إحصائية مثل Excel، و برامج إحصائية أخرى مثل SPSSو Stata.
- **خبير القراءة المبكرة:** مسئول عن البحث في مجال القراءة، والعمليات التربوية، والتدريس، ولديه خلفية في أدوات تقييم القراءة والتدريس.

3.1.6 ما المواد التي نحتاج إليها؟

مواد لورشة عمل التطويع تتضمن ما يلي:

- الورق، وأقلام الرصاص، وممحاة للمشاركين،
- وجهاز عرض LCD، ولوحة بيضاء، وورق قلاب (إذا كان ذلك ممكنًا، يجب أن يكون جهاز العرض LCD قادرًا على العرض على السبورة بالنسبة للتمارين التي تتطلب تقديرًا للدرجات)،
- ونصوص للقراءة وطنية أو محلية، ومناسبة لمستويات الصفوف الدراسية وللغات التي ستقيم (وهذه النصوص ستوضح المفردات المستخدمة في كتابة القصة وتحديد مستوى الصعوبة).
- نسخ ورقية من العروض التقديمية، ومسودة الأداة
- العرض التقديمية عن نتائج الأبحاث وثيقة الصلة بالقراءة وتقييم مهارات القراءة في الصفوف الأولى، وعملية تطوير الأداة والغرض منها والاستخدامات والخلفية البحثية.
- نماذج لقطع القراءة الشفوية وأسئلة الفهم وأسئلة فهم المسموع من بلدان أخري وفقًا لأداة تقييم مهارات القراءة في الصفوف الأولى، وذلك بهدف تعديلها.

يعرض الشكل 8 نموذجًا لأجندة ورشة عمل تطويع الأداة و الأبحاث.

الشكل 8: نموذج للأجندة: ورشة عمل تطويع أو تعديل التطويع لأداة تقييم مهارات القراءة في الصفوف الأولى

اليوم والزمن	اليوم 1	اليوم 2	اليوم 3	اليوم 4	اليوم 5
9:00-9:30 a.m.	ترحيب ومقدمة	مراجعة اليوم الأول	مراجعة اليوم الثاني	مراجعة اليوم الثالث	زيارة المدارس للتجريب الميداني للأداة والاستبيانات
9:30-10:30 a.m.	عرض عام للمشروع وسياق.	مراجعة النسخة الأولية من (على سبيل المثال: الكلمات غير المألوفة)	تطوير قطع فهم المسموع.	تعديل/ تطوير المهام الفرعية الإضافية والاستبيانات إذا أمكن.	
10:30-11:00 a.m	**راحة**				
11:00-12:30 p.m.	عرض عام لأداة : الغرض ومحتوى الأداة ونتائج الاستخدام.	تطوير قطع الطلاقة في القراءة الشفوية.	الاستمرار في وضع أسئلة قصص فهم المسموع.	تعديل /تطوير المهام الفرعية الإضافية والاستبيان إذا أمكن.	ملخص لزيارة المدرسة
12:30-1:30 p.m.	**الغذاء**				
1:30-3:00 p.m.	عرض عن اللغة: نظام الكتابة وقضايا تؤخذ في الاعتبار عند تطوير.	الاستمرار في قصص الطلاقة في القراءة الشفوية ووضع أسئلة.	مراجعة وتحديث استبيان الأداة.	مراجعة الأداة والتدريب على تطبيق الأداة للتجريب الميداني.	وضع الصورة النهائية للأداة
3:00-3:45 p.m.	**راحة**				
3:45-5:00 p.m.	مراجعة مسودة الأداة (على سبيل المثال الوعي الصوتي وأصوات	الانتهاء من القصص والاسئلة.	الانتهاء من القصص والأسئلة واستبيان التلاميذ إذا أمكن.	مراجعة الأداة والتدريب على تطبيق الأداة للتجريب الميداني.	ختام الورشة
الأهداف اليومية	*فهم الغرض من والمحتوى.*	*تطوير قطع الطلاقة في القراءة الشفوية.*	*كتابة قطع وقصص الفهم المسموع واستبيان التلاميذ.*	*تطوير المهام الفرعية والاستبيانات الإضافية.*	*وضع الصورة النهائية للأداة*

ملاحظة: ستعتمد مدة ورشة عمل التطويع والجلسات المحددة على عدة عوامل، بما في ذلك: وجود أداة EGRA مستخدمة للغة / بلد / صف معين،وعدد من المهام الفرعيةالتي ستفحص ، وعدد اللغات التي ستقيم، والحاجة إلى استبيانات وأدوات إضافية؛ والغرض والجمهور المستهدف من ورشة العمل.

2.6 مراجعة المكونات المشتركة للأداة:

كما نوقش في القسم 1، طور التصميم الأولي لأداة بدعم من خبراء من الوكالة الأمريكية للتنمية الدولية والبنك الدولي، ومنظمة معهد تراينجل للأبحاث RTI . وعلى مدى السنين، أثمرت استشارات الخبراء عن التطبيق الكامل لتقييم مهمات القراءة في الصفوف الأولى باللغة الانجليزية، والذي روجع وحُدِث باستمرار. والأداة الشائعة لنظم الكتابة التي تعتمد على شكل الحرف تحتوي على ست مهام فرعية منها المهام الفرعية الجوهرية، وهذه المهام الفرعية الأربعة هي:

1. فهم المسموع.
2. تعرف الحروف.[10]
3. قراءة الكلمات غير المألوفة.
4. الطلاقة في القراءة الشفوية مع الفهم.

ومن الجدير بالذكر فإن كل من هذه المكونات الأساسية (انظر **الجدول 9**) قد جُربت في العديد من اللغات حول العالم.

يمكن تقييم اثنين من المهام الفرعية المشتركة الإضافية التي غالبا ما تُضمن في الاختبار وهي الوعي الصوتي، وقراءة الكلمات المألوفة. ويمكن تقييم الوعي الفنولوجي (مجموعة فرعية من الوعي الصوتي) باستخدام إجراءات أخرى مختلفة، اعتمادًا على ما قد يكون مناسبًا في سياق معين. [11]

[10] ويمكن أن تتكون المهمة الفرعية ؛ لتعرف الحرف من أصوات الحروف أو أسماء الحروف. على الرغم من أن المهمة الفرعية لأصوات الحروف تُستخدم على نحو أكثر شيوعًا، قد تكون أسماء الحروف أكثر ملاءمة وهذا يعتمد على البلد واللغة التي تُطبق فيها الأداة. فهناك لغات قائمة على مقطع واحد قد تُضمن المهام الفرعية فيها لقراءة المقطع أكثر من تسمية الحروف وأصوات الحروف.

[11] المهام الفرعية الاختيارية مثل: الإملاء المتاهة، و التتمة تُستخدم في بعض الأحيان بالإضافة إلى المهام الفرعية الشائعة. فالرجاء الاطلاع على المناقشة حول هذه المهام الفرعية الإضافية في الفصل 1 (Gove and Wetterberg 2011)، وكذلك القسم 6.3 أدناه، "عرض مكونات الأداة الإضافية."

تضمنت تعليقات من الممارسين والنظراء المحليين طلبات للحد من عدد من المهارات التي تُختبر في أداة تقييم مهارات القراءة في الصفوف الأولى. وكما ذُكر أعلاه، فإن واحدًا من أغراض الأداة هو تقييم بطارية كاملة لمهارات القراءة الأساسية بشكل معقول؛ لتكون قادرة على تحديد المناطق التي تحتاج إلى تدريس إضافي. إذا كانت أداة تختبر فقط طلاقة القراءة الشفهية، فإن من المرجح أن العديد من البلدان ذات الدخل المنخفض ستشهد بشكل كبير أقل أثر ممكن. الحفاظ على عدد المهام الفرعية ليكون نحو ست مهمات تقريبًا؛ ليسمح للدول التي أداء طلابها في القراءة ضعيف جدا أن يحدد التقدم على الأقل في بعض مهارات القراءة.

ومن المهم أيضا أن نلاحظ أن الأداة والإجراءات المقدمة هنا قد عُرضت لتكون نقطة انطلاق معقولة؛ لتقييم القراءة المبكرة(NICHD, 2000; and Dubcek & Gove, 2015) ، وهذا يعني أيضًا أن المهارات التي تقيسها ضرورية لكنها غير كافية؛ لتحقيق القراءة الناجحة التي تغطي أداة عددًا كبيرًا من المهارات التنبئية منها، ولكن ليست كل المهارات أو المتغيرات التي تساهم في التحصيل القرائي. فعلى سبيل المثال، لا تقيس الخلفية المعرفية لدى الطفل والدافعية، والانتباه، واستراتيجيات الذاكرة، واستراتيجيات القراءة ، واستراتيجيات المفردات الإنتاجية ، وفهم أنواع متعددة للنصوص و إعادة السرد ، وغيرها. ولا يوجد نوع واحد من التقييم يمكن أن يغطي جميع المهارات الممكنة؛ لأنه سيكون طويلًا بشكل استثنائي، وستجعل الطلاب يشعرون بالتعب ويؤدون بشكل ضعيف. فلا ينبغي أن يُنظر إلى أداة على أنها مقدسة من حيث الأجزاء المكونة لها، ولكن يُوصى أن التنوع مطلوب، سواءً في مكونات المهمة أو في الإجراءات المنفذة للأداة، ويوجد ما يبرر ذلك بحيث يكون موثقًا من حيث الغرض واستخدام التقييم، وكذلك يُشارك هذا مع المجتمع الأكبر.

الشكل 9- مراجعة المكونات الشائعة للأداة:

المكون	مهارة القراءة المبكرة	المهارة تعكس قدرة التلاميذ على أن:
1. فهم المسموع:	فهم المسموع اللغة الشفوية	الاستجابة بطريقة صحيحة إلى أنواع مختلفة من الأسئلة، بم ا في ذلك أسئلة الفهم الحرفي والاستنتاجي حول النص الذي يقرأه المقيم لهم.
2. قراءة الكلمات غير المألوفة	فك رموز (قراءة) الكلمات:	إعطاء اسم الحرف (العلاقة بين الصوت و الرمز) من خلال قراءة كلمات لا معنى لها.
3. الطلاقة في القراءة الشفوية مع الفهم	الطلاقة في القراءة الشفوية	قراءة نص بدقة بجهد قليل، وبمعدل كاف
	الفهم القرائي	الاستجابة بشكل سليم لأنواع مختلفة من الأسئلة تتضمن أسئلة للفهم الحرفي والاستدلال ي حول نص قراؤه
4. تعرف الصوت الأول في الكلمة والصوت الأخير في الكلمة، أو تمييز صوت الحرف أو التقطيع إلى مقاطع صوتي	الوعي الفنولوجي	تحديد / تمييز بين بداية المقطع والصوت المتمم له في كلمات أو الصوت الأول أو الصوت الأخيري الكلمات، أو تقطيع الكلمات إلى أصوات وذلك عندما يقرأ الميسر الأصوات بصوت عال، ثم يتبعه التلميذ بعد ذلك.

1.2.6 فهم المسموع

ويشمل تقييم فهم المسموع قراءة قطعة بصوت عال من قبل المقيم، ثم يجيب الطلاب عن أسئلة الفهم شفويًا، أو عن عبارات تقريرية. يمكن أن تتضمن هذه المهمة الفرعية في بداية السلسلة؛ لتمهيد الأطفال للدخول في عملية التقييم وتعريفهم بلغة التقييم.

القيام باختبار فهم المسموع بشكل منفصل عن الفهم القرائي أمر مهم؛ لأنه يوفر معلومات حول ما يستطيع الطلاب فهمه بعيدًا عن تحدي فك رموز نص. فالطلاب الذين يجدون صعوبة في القراءة، أو لم يتعلموا بعد فك رموز (قراءة) نص مكتوب مازال لديهم لغة شفوية و مفردات، ومهارات واستراتيجيات للفهم يمكنهم إظهارها بعيًدا عن قراءة النص، وهذا يعطي صورة أكثر اكتمالًا بكثير عما بمقدور الطلاب عندما يتعلق الأمر بالفهم ، وقد تواجدت اختبارات فهم المسموع لبعض الوقت، وعلى وجه الخصوص، فقد استخدمت بمثابة تقييم بديل للأطفال المحرومين ، والذين لا تتوافر لديهم اللغة المكتوبة (الكتب) نسبيا لعدم وصول الطباعة في بلدانهم .(Orr & Graham, 1968).

القسم4 - فهم المسموع

هذا التمرين لا يعتمد التوقيت. **ليس هناك ورقة للتلميذ**. المسير يقرأ بصوت عال المقطع التالي مرة واحدة فقط، بتمهل (كلمة كل ثانية تقريبا). قل:

سأقرأ عليك قصة قصيرة بصوت عال، مرة واحدة. و بعد ذلك سأطرح عليك بعض الأسئلة. اسمع جيدا من فضلك وحاول أن تجيب قدر ما استطعت على الأسئلة. هل فهمت المطلوب منك؟

كُلَّ صَبَاحٍ، تَحْلُبُ أُمِّي بَقَرَتَهَا، وَ تُحَضِّرُ لَنَا الفُطُورَ. ذَاتَ يَوْمٍ، لَمْ تَجِدِ البَقَرَةَ فِي الإِسْطَبْلِ. اسْتَيْقَظْتُ فَلَمْ أَرَ فُطُوراً عَلَى الطَّاوِلَةِ. كَانَتْ أُمِّي قَدْ خَرَجَتْ لِلْبَحْثِ عَنْهَا عِنْدَ الجِيرَانِ، فِي الحَقْلِ، وَ قُرْبَ النَّهْرِ. غَيْرَ انَّهَا لَمْ تَعْثُرْ عَلَيْهَا. فَبَكَتْ حُزْنًا لأَنَّهَا كَانَتْ تُحِبُّهَا كَثِيرًا. فَلَمَّا عَادَتْ سَمِعَتْ ضَجِيجًا فِي المَطْبَخِ. إِنَّهَا البَقَرَةُ تَأْكُلُ جَزَرًا مِنْ سَلَّةِ الخُضَرِ !

-ماذا تفعل الأم كل صباح ؟	تحلب الام بقرتها	o	صحيح	o	غير صحيح	o	لا إجابة
ماذا حدث للام في الاسطبل ؟	لم تجد البقرة.	o	صحيح	o	غير صحيح	o	لا إجابة
لماذا لم يجد الطفل فطورا على الطاولة	لان الام لم تحلب البقرة.	o	صحيح	o	غير صحيح	o	لا إجابة
اين بحثت الام عن البقرة؟	بحثت عنها عند الجيران و في الحقل و قرب النهر	o	صحيح	o	غير صحيح	o	لا إجابة
لما بكت الأم ؟	لأنها تحب بقرتها لمْ تجد البقرة	o	صحيح	o	غير صحيح	o	لا إجابة
ماالذي احدث الضجيج في المطبخ؟	البقرة	o	صحيح	o	غير صحيح	o	لا إجابة

، ويشير الأداء الضعيف على أداة فهم المسموع أيضًا إلى أن الأطفال يفتقرون إلى المعرفة الأساسية للغة التي يُقيَّمون فيها، أو أن لديهم صعوبة في معالجة ما يسمعونه.

البيانات: يُمنح الطلاب الدرجات على عدد الإجابات الصحيحة التي يقدمونها عن الأسئلة المطروحة (من أصل العدد الإجمالي للأسئلة). ويتجنب مصممو الأداة الأسئلة التي يُجاب عنها فقط ب "نعم" أو "لا".

بناء البند. يعتمد على طول القطعة، وعلى المستوى واللغة الأولى للأطفال الذين يقيَّمون فيها، وعلى الرغم من أن معظم القطع يجب أن تكون حوالي 30 كلمة في الطول وذلك من أجل تقديم نص يكفي؛ لتطوير المادة من ثلاثة إلى خمسة أسئلة للفهم. تحكي القصة نشاطا أو حدثًا مألوفًا للأطفال. فيجب أن تكون الأسئلة مماثلة للأسئلة التي تُطرح في مهمة الفهم القرائي

(الموضح أدناه). وستكون معظم الأسئلة في مستوى الفهم الحرفي، والتي يمكن الإجابة عنها مباشرة من النص. واحد أو اثنين من الأسئلة استنتاجية، والتي تتطلب من الطلاب استخدام معرفتهم السابقة، وكذلك النص لإجابة السؤال.

الشكل 10 -نموذج للمهمة الفرعية لفهم المسموع:

2.2.6 تعرف الحرف

معرفة كيف تمثل مهمة تعرف الحروف الأسماء أو الأصوات مهارة مهمة يجب على التلاميذ إتقانها حتى يصبحوا قراءً ناجحين. فيدرس الطالب العلاقة بين الحروف والأصوات (المبدأ الأبجدي) من خلال المداخل القائمة على الصوتيات. فمعرفة الحروف هو مدخل شائع للتقييم، ويُستخدم في العديد من التقييمات للقراءة المبكرة، بما في ذلك الاختبار الشامل لمرحلة ما قبل المدرسة في المعالجة الصوتية و المكتوبة (Lonigan, Wagner, Torgesen, & Rashotte, 2002). ويمكن أن يشمل التقييم واحدًا منها أو كل من الخيارات المدرجة أدناه. وذلك لاختبار تعرف الصوت والحرف بطريقة صحيحة التي يحتاجها الطلبة حتى يكونوا قادرين على فك رموز(قراءة) الكلمات، وبمعنى آخر فإن معرفة الصوت الذي يمثله الحرف يساعد الطلاب على قراءة الكلمة. وفي الوقت نفسه، فإن الأبحاث، وخاصة تلك التي تركز على نظم الكتابة الأقل شفافية مثل اللغة الإنجليزية أظهرت أن معرفة أسماء الحروف لديها القدرة على التنبؤ بالتحصيل في القراءة فيما بعد فى مراحل متقدمة في أثناء تعلم تلك اللغات.

المدخل الأول-تعرف صوت الحرف:

في هذه المهمة الفرعية يطلب من التلاميذ نطق أصوات الحروف كلها من قائمة تُعطى لهم خلال فترة من الزمن مقدارها دقيقة واحدة.

وبالنسبة للحروف، توضع مجموعة كاملة من الحروف الأبجدية في ترتيب عشوائي، بحيث يُكتب عشرة حروف في كل صف واحد، وذلك باستخدام خط واضح، وكبير، ومألوف. أما عدد مرات تكرار الحرف يعتمد على تكرار الحرف في اللغة التي يُجرى التقييم بها. (فعلى سبيل المثال، انظر تكرار الحروف في اللغة الإنجليزية في **الشكل 11**). والتي تكون الأبجدية كاملة بناءً على أدلة من اللغات الأوروبية أن مهارات القراءة لدى الطلاب تتحسن فقط بعد أن يعرف ما يقرب من 80 % من الأبجدية (Seymour, Aro, & Erskine, 2003).

وتعتمد جداول تكرار الحروف على النص الذي تُحلل مفرداته. وهذه الجداول متوفرة للغات الإنجليزية، و الإسبانية ، و الفرنسية ، و لغات أبجدية دولية أخرى. ومطورو الاختبارات الذين يصممون أدوات تقييم مهارات القراءة لتلاميذ الصفوف الأولى في لغات أخرى يختارون عينة من 20 – 30 صفحة من كتاب ملائم لصف دراسي أو نصوص قراءة إثرائية ويحللون تكرار الحروف إليكترونيًّا، ثم يضعون جداول لتكرار الحروف المتشابهة.

الشكل 11: الحروف في اللغة الانجليزية: تكرار الاستخدام

E	12.02%	C	2.70%	Y	2.11%
A	8.12%	U	2.88%	W	2.09%
R	6.02%	D	4.32%	K	0.69%
I	7.31%	P	1.82%	V	1.11%
O	7.68%	M	2.61%	X	0.17%
T	9.10%	H	5.92%	Z	0.07%
N	6.95%	G	2.03%	J	0.10%
S	6.28%	B	1.49%	Q	0.11%
L	3.98%	F	2.30%		

المصدر: *English letter frequency (based on a sample of 40,000 words).* Ithaca, New York: Department of Mathematics, Cornell University. Retrieved September 2015 from http://www.math.cornell.edu/~mec/2003-2004/cryptography/subs/frequencies.html

يتطلب وضع جدول تكرارات الحروف كتابة عينة من الصفحات في برنامج معالج الكلمات، واستخدام الأمر "البحث". أدخل الحرف "أ" في مربع البحث "ابحث عن"، وسيظهر الحرف موضوع البحث مظللًا بلون في كل الصفحات. وفي حالة استخدام Microsoft Word، فسيظلل الحرف "أ" في كل مرة يظهر في الوثيقة، وسيقدم تقريرًا عن عدد المرات التي ظهر فيها الحرف. وسيكرر من يقوم بعملية التحليل هذه العملية لكل حرف من الحروف الأبجدية، وتسجيل العدد لكل حرف حتى تحسب نسبة التكرار لكل حرف، ويمكن أن تحسب كنسبة من إجمالي عدد الحروف في الوثيقة.

[12] More about Andika, including how to download this font, can be found on SIL's website: http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=andika

[13] Letter-frequency rates for French, German, Spanish, Portuguese, and others are available from University of California at Los Angeles Statistics Online Computational Resource, http://wiki.stat.ucla.edu/socr/index.php/SOCR_LetterFrequencyData#SOCR_Data (accessed September 18, 2015).

أما تضمين متواليات من حرفين أو أكثر، والتي تعطى صوتًا واحدًا (digraph)وكذلك الـ Diphthongs التي تشير الي وجود صوتين متحركين متجاورين يكونان مقطعًا واحدًا في هذه المهمة الفرعية يتنوع من لغة لأخرى و كذلك وفقًا لطبيعة المنهج و الصف الدراسي في الدولة التي يجرى فيها التقييم، و يتعرف مطورو الأداة على أي من هذه الظواهر اللغوية شيوعًا، و التي يتعين على الطلاب تعلمها من خلال تحليل التكرار لعينة من صفحات الكتب. تجري استشارة الخبراء في اللغة حول الحروف الأكثر ملائمة؛ لتضمينها في المهمة الفرعية، وذلك من خلال المنهج وتحليل التكرار اللغوي.

تحتاج قضية النطق إلى تناولها بشيء من الحساسية في هذه المهمة الفرعية، وغيرها من المهام. القضية ليست اختبار النطق السليم. فالتقييم يقيس ما إذا كان التلاميذ يستطيع تحديد صوت الحرف أم لا، متسامحًا مع النطق الذي قد يكون شائعًا في منطقة جغرافية معينة أو صيغة لغوية معينة. و هكذا، فإن اللهجات المحلية مقبولة عند الحكم على ما إذا كان نطق صوت الحرف سليمًا.

أما بالنسبة للحروف التي يمكن أن تعطى أكثر من صوت في بعض اللغات، فستكون الإجابات المتعددة مقبولة. وفى أثناء التدريب، يراجع المقيمون والمشرفون بمساعدة خبراء اللغة أوجه النطق الممكن لكل حرف ويتفقون معًا على الاستجابات المقبولة مراعين الاختلافات، واللهجات المحلية. (ففي اللغة الإنجليزية، توجد قائمة بكل الرموز والأشكال في هجاء الأبجدية الصوتية، ارجع الي الموقع http://westonruter.github.io/ipa-chart/keyboard/ للصول على نسخة من الجدول المعد بواسطة جمعية الصوتيات الدولية).

البيانات: تحتسب درجة الطفل على هذه المهمة الفرعية على أساس عدد أصوات الحروف التي تُقرأ بطريقة صحيحة في الدقيقة الواحدة. فإذا أكمل الطفل قراءة كل أصوات الحروف قبل نهاية الوقت، يُسجل وقت الانتهاء، وتحتسب الدرجة على هذا الوقت. وفي حال استخدام التقييم الورقي يضع المقيِّمون علامة الخط المائل (//) على أصوات الحروف التي تُقرأ بطريقة خطأ، ووضع ([) بعد الحرف الأخير الذي قراءه التلميذ، وتسجيل الوقت المتبقي على ساعة التوقيت عند الانتهاء من التمرين. أما في جمع البيانات الإلكتروني، فإن وضع العلامات والحسابات تُجرى بشكل آلي على شاشة الكمبيوتر اللوحي للمقيمين. وتستخدم ثلاث نقاط للبيانات لحساب مجموع أصوات الحروف الصحيحة في الدقيقة الواحدة:(clspm) وتقاس بالمعادلة الآتية:

عدد الأصوات الصحية في الدقيقة = عدد أصوات الحروف التي تعرف عليها التلميذ – عدد الاجابات الخاطئة /{(60 – الوقت المتبقي)/ 60}

كل من نقاط البيانات هذه يمكن أن تُستخدم أيضًا؛ لإجراء تحليلات إضافية. فعلى سبيل المثال، فإن المعلومات حول العدد الإجمالي للأصوات التي يحددها التلميذ تسمح بالتمايز بين الطالب الذي يقرأ 50 صوتًا في الدقيقة شريطة أن يكون نصفهم صحيحًا، ولكن يعطي أسماء نصفهم فقط بشكل صحيح وبين طالب ينطق 25 صوتًا في الدقيقة الواحدة، ولكن أسماء كل من هذه الحروف صحيحة.

لاحظ أن هذه المهمة الفرعية، فضلاً عن العديد من المهام الفرعية التي تأتى بعد ذلك ليست فقط في ذات الوقت، ولكن لها وقت محدد (أي يطلب من التلميذ التوقف بعد فترة محددة، سواء أتم المهمة، أو لم يتمها). وجود حدود للوقت يفيد في جعل التقييم يستغرق وقتًا أقل، وأيضًا أقل إرهاقًا لكل من الطفل والمقيم، كما أن الطفل لا يتعين عليه الاستمرار في المحاولة، لإتمام المهمة ببطء. وبالإضافة إلى ذلك، فإن التوقيت ضروري لقياس الطلاقة.

بناء البند: وتتكون هذه المهمة الفرعية من 100 بند. فالحروف الأبجدية توزع بشكل عشوائي، مع مراعاة وجود 10 حروف في كل صف أفقي، وتوزع بالتساوي على باقي الصفوف. وستُعرض معظم الرموز عدة مرات. فالنسب المئوية المحسوبة في التمرين المذكور أعلاه بمثابة مثال توضيحي لتكرارات المهمة الفرعية كما تظهر في ورقة العمل.

ليس من غير المألوف لأداة الحالية أن تحتاج إلى تعديلها لنسخ متوازية، فعلى سبيل المثال، يهدف مراقبة التحسن من دراسة الأساس إلى الدراسة النهائية. فإن البنود في هذه المهمة الفرعية قد يعاد ترتيبها أو إعادة عرضها بشكل عشوائي وذلك لوضع جداول جديدة – أي 10 صفوف مكونة من 10 حروف دون الحاجة إلى إعادة حساب التكرارات للحروف من جديد.

وفي هذه الحالات، ولضمان وجود نماذج متكافئة للاختبار، فمن المهم أن يحدث إعادة ترتيب فقط داخل الصفوف الفردية (من أجل الحفاظ على صعوبة المهمة الفرعية النسبية). وبعبارة أخرى، يظل كل بند في الجدول في نفس الصف الأفقي الذي تظهر فيه في الأداة السابقة.

الشكل 12 عينة لتصميم المهمة الفرعية لتعرف صوت الحرف من مصر والشكل 13 يوضح خليط الحروف المختارة لأداة في اللغة العربية

14 بينما ستحد إعادة الترتيب داخل الصفوف من إحداث تغييرات كبيرة في صعوبة المهمة الفرعية. لا يزال من الموصي به اختبار أثر الترتيب كلما أمكن ذلك.

الشكل 12 عينة لتصميم المهمة الفرعية لتعرف صوت الحرف منمصر

جزء 1- ب : تعرف أصوات الحروف

قدم للتلميذ/ للتلميذة ورقة الأحرف الموجودة في الصفحة الثانية من كراس التلميذ/ التلميذة وقل له :

هذه الصفحة مليئة بالأحرف. رجاء انطق صوت كل حرف. فعلى سبيل المثال اسم صوت الحرف "مرّ" (أشر الى الحرف " مَ ") هو " مَ " ، الآن حاول انت . أريني لي أنت صوت هذا الحرف (أطلب من التلميذ/ التلميذة أن ينطق صوت "لَ") . فإذا استطاع التلميذ/التلميذة أن ينطق صوت الحرف بشكل صحيح قل له "أحسنت"، صوت هذا الحرف " كُ " . أما إذا أخطأ التلميذ/التلميذة قل له صوت هذا الحرف " كُ ". الآن حاول مرة أخرى: أريني لي صوت هذا الحرف(اطلب من التلميذ/التلميذة تَعْرِف صوت "لَ") . فإذا استطاع التلميذ/التلميذة أن ينطق صوت الحرف بشكل صحيح ، قل له " أحسنت"، صوت هذا الحرف " لَ" . أما إذا أخطأ التلميذ/التلميذة قل له صوت هذا الحرف "لُ". هل تفهم ما هو مطلوب منك ؟

عندما أقول لك " ابدأ " أنطق أصوات الأحرف بأفضل طريقة ممكنه، سوف ألتزم الصمت وأستمع لما تقول إلا إذا احتجت مساعدتي، مستعد؟ ابدأ

يجب على التلميذ قراءة الحروف بالتشكيل لكي تحسب له صحيحه. رجاء لاحظ أن التلميذ /التلميذة يلتزم بالقراءة من اليمين الى اليسار إبتداء من السطر الاول. يجب على التلميذ قراءة الحروف بالتشكيل لكي تحسب له صحيحة

سجل مراعاة التوقيت عندما يبدأ التلميذ/ التلميذة في القراءة. تابع التلميذ/ التلميذة باستخدامك قلم الرصاص وضع إشارة " / " لأي صوت يخطئ التلميذ/ التلميذة في نطقه. احسب التصحيح الذاتي الذي يقوم به التلميذ/التلميذة على أنه صحيح . التزم الهدوء إلا عندما تكون استجابات التلميذ/التلميذة على النحو التالي: إذا تردد التلميذ/ التلميذة لمدة 3 ثوان، قل له صوت الحرف ثم أشر الى الحرف اللاحق، واطلب منه أن يستمر في نطق الأحرف. احسب هذا الصوت الذي نطقه التلميذ/ التلميذة على أنه غير صحيح.

بعد 60 ثانية. ضع علامة] عند الحرف الأخير الذي نطقة التلميذ/التلميذة ثم أطلب من التلميذ/التلميذة أن يتوقف .

10/	تِ	قَ	جُ	سِ	كُ	هَ	ر	غُ	و	يَ
20/	ش	ذْ	حَ	لُ	زَ	شُ	عَ	طَ	ضِ	ظ
30/	زُ	سَ	نَ	ر	يَ	خِ	مُ	ثُ	فِ	ن
40/	لِ	قُ	جِ	صُ	دَ	بِ	أُ	خُ	مِ	ثْ
50/	حَ	ذْ	دُ	بِ	وُ	فَ	صْ	عِ	طُ	أ

الوقت الذي تشير إليه الساعة _____

☐ تم إيقاف التدريب نظراً لأن التلميذ/ التلميذة لم يستطع قراءة السطر الأول

المدخل الثاني-تعرف اسم الحرف :

هذا المهمة الفرعية مشابهة جدًا في التركيب والتطبيق للمهمة الفرعية السابقة. بالرغم من أن معرفة أصوات الحروف من المتطلبات السابقة لفك رموز (قراءة) الكلمات، فإن العديد من الدراسات التي أُجريت في الولايات المتحدة والدول الأوروبية أكدت أن معرفة اسم الحرف أيضًا القدرة على التنبؤ بدرجة كبيرة عن التحصيل في القراءة فيما بعد. وفي العديد من البلدان، فإن الاختبارات الاستطلاعية لهذه المهمة الفرعية في أداة تقييم مهارات القراءة في الصفوف الأولى ، والتي تقيَّم معرفة أسماء الحروف لدى الأطفال أوضحت أن النتائج ذات أثر كبير -أي أن معظم التلاميذ حصلوا على درجات عالية في هذه المهمة -، ولذلك فإن المهام الفرعية الأخرى تستخدم غالبًا في النسخة النهائية من الأداة.

البيانات: كما في تمرين تعرف صوت الحرف، فإن درجة الطفل على هذه المهمة الفرعية تُحسب بناء على عدد الحروف الصحيحة المقروءة في الدقيقة.

بناء البند:

تُبنى المهمة الفرعية لاسم الحرف بناء على نفس تحليل التكرار كما سبق شرحه في إطار مهمة أصوات الحروف. وتُغلَّف كراسة أسئلة الطالب بالبلاستيك، والتي تحتوي على جميع الحروف الأبجدية، بحيث تكون موزعة عشوائيًّا، ويكتب 10 حروف في كل صف أفقي، وتضم المهمة الفرعية عشرة صفوف (مجموعها 100 حرفٍ)، وستعرض معظم الحروف عدة مرات، بناءً على تكرار الحروف المستخدمة في اللغة. ويمكن إعادة ترتيب البنود في هذه المصفوفة، وإعادة ترتيبها (إعادة الترتيب العشوائي)، لإعداد صور مكافئة للاختبار على الرغم من أنه يوصى باختبار أثر إعادة الترتيب.

انظر نموذج المهمة الفرعية لأسماء الحروف في الشكل 12 من أجل معرفة شكل المهمة الفرعية لتعرف اسم الحرف.

3.2.6 قراءة الكلمات غير المألوفة

قراءة الكلمات غير المألوفة هو مقياس لقدرة فك الرموز (أي الطريقة المعجمية الفرعية لمعالجة الكلمات كما جاء في القسم 4.3.1)، وهي تختلف عن الطريقة الكلية لتعرف الكلمة، أو طريقة احتفاظ الكلمة في الذاكرة، والطريقة المعجمية. التي يتعلم العديد من الأطفال في الصفوف الأولى بها ، وهى أن يحفظوا عن ظهر قلب أو أن يتعرفوا بمجرد النظر مجموعة

واسعة من الكلمات. فقد ارتبط استهلاك هذه الكلمات البصرية في حوالي سن العاشرة مع تراجع في تلاميذ الصف الرابع بالولايات المتحدة .(Hirsch, 2003) ويجب أن يجمع الأطفال بين مهارات فك الرموز والإدراك الكلي للكلمات حتى يكونوا قراءً ناجحين، والاختبارات التي لا تتضمن سؤالاً عن فك الرموز فإنه من الممكن أن يبالغ في تقدير قدرة الأطفال على قراءة الكلمات غير المألوفة حيث قد تكون الكلمات التي يُجرى اختبارها جزء من المفردات التي يتعرف عليها بصريًا.

البيانات: تحسب درجة الطفل بناء على عدد الكلمات غير المألوفة التي تُقرأ قراءة صحيحة في الدقيقة الواحدة. وتُجمع نفس أقسام المتغيرات كما تجمع للمهام الفرعية ذات الوقت المحدد إلكترونيًا لقراءة الكلمات غير المألوفة . من خلال عدد الكلمات المقروءة قراءة صحيحة، وعدد الكلمات التي قُرأت قراءة خطأ واحتساب الوقت المتبقي من تنفيذ المهمة.

بناء البند: يشمل هذا الجزء من التقييم قائمة من 50 كلمة غير مألوفة، تُكتب خمسة كلمات في صف واحدم مراعاة نماذج الحروف داخل الكلمات حسبما يتفق مع اللغة تتبع الكلمات غير المألوفة قواعد اللغة، وذلك باستخدام الحروف في مواضعها الصحيحة للكلمات المألوفة. وأيضا، تقتصر على تركيب الساكن والمتحرك كما في الكلمات المألوفة. (فعلى سبيل المثال، لا يبدأ المقطع العربي بساكنين مثال (صْ+بْ) تُستخدم الكتابة الواضحة والمسافات الكافية بين الكلمات. فيمكن إعادة ترتيب الكلمات في الصف الأفقي الواحد (إعادة الترتيب العشوائي للكلمات)، وذلك لإعداد صور متكافئة للاختبار، على الرغم من أنه يوصى باختبار أثر إعادة الترتيب.

الشكل 14 مثال للمهمة الفرعية لقراءة الكلمات غير المألوفة

جزء 4: قراءة كلمات غير مألوفة

قدم للتلميذ/ للتلميذة صفحة الكلمات التي ليس لها معنى في الصفحة الرابعه من نموذج التلميذ/التلميذة وقل له :

انظر إلى هذه الكلمة ، إنها كلمة ليس لها معنى معروف لديك : هذا مثال على ذلك : " لمش "

الآن حاول أنت : اقرأ لي هذه الكلمة " راك " إذا استجاب التلميذ/التلميذة بشكل صحيح، فقل أحسنت ، هذه الكلمة هي " راك" . أما إذا أخطأ التلميذ/التلميذة ، فقل هذه الكلمة هي " راك" .

الآن حاول مرة اخرى : اقرأ لي هذه الكلمة "لمك " إذا استجاب التلميذ/التلميذة بشكل صحيح، فقل أحسنت ، هذه الكلمة هي " لمك " . أما إذا أخطأ التلميذ/التلميذة ، فقل هذه الكلمة هي " لمك "

هل فهمت المطلوب ؟

هل أنت مستعد ؟ ابدأ .

شغل ساعة التوقيت عندما يبدأ " التلميذ/التلميذة في القراءة . من المهم وضع إشارة " / " لكل كلمة يقرؤها التلميذ/التلميذة بشكل غير صحيح .

احسب الكلمات التي يقوم التلميذ/التلميذة بتصحيح نفسه فيها على أنها صحيحة . التزم الهدوء إلا عندما تكون استجابة التلميذ/التلميذة على النحو التالي : إذا تردد التلميذ/التلميذة لمدة 3 ثوان ، اقرأ له الكلمة ثم أشر إلى الكلمة اللاحقة وأطلب من « أن يستمر في قراءة الكلمات. احسب الكلمة التي قرأتها على أنها غير صحيحة . اذا أخطاء التلميذ في خمس كلمات انتقل فوراً الى السؤال التالي

بعد 60 ثانية، ضع علامة] عند آخر كلمة قرأها التلميذ /التلميذة ثم قل " توقف"

ربر	ديص	هزن	نوين	حْناج	5
غلد	ألفاط	ضعب	دشين	بكح	10
أغو	سريلح	درمسه	ركف	نيد	15
مبون	صوج	عوسج	دريب	قرج	20
نيراج	دعن	فتماح	عقص	فرتع	25
نهذ	عكور	عمس	كربش	قعل	30
أمزى	قانره	ركس	تن	العيس	35
زاك	برت	خادي	رفيم	دبلَ	40
عبوج	دلب	كمس	باصح	رص	45
آخالا	ميلار	قرز	معالات	لصوت	50

الوقت الذي تشير اليه الساعه _____

تم إيقاف التدريب نظراً لأن التلميذ/ التلميذة لم يستطع قراءة السطر الأول ☐

4.2.6 الطلاقة في القراءة الشفوية مع الفهم

الطلاقة في القراءة الشفوية هي مقياس لكفاءة القراءة بشكل عام: كالقدرة على ترجمة الحروف إلى أصوات، وتوحيد الأصوات إلى كلمات، ومعالجة العلاقات وربط النص بالمعنى، وعمل الاستدلالات لاستكمال المعلومات الناقصة (Hasbrouck & Tindal, 2006).

عندما يترجم القراء الناجحون النص إلى لغة منطوقة، فهم يجمعون هذه المهام بدون جهد كبير؛ وذلك لأن الطلاقة في القراءة الشفوية ترصد هذه العملية، ويمكن استخدامها؛ لتمييز القدرة على القراءة الشاملة. وتوضيح اختبارات الطلاقة في القراءة الشفوية، كما تقيسها تقييمات ذات وقت محدد لعدد الكلمات التي تُقرأ قراءة صحيحة في الدقيقة، حيث إنه توجد علاقة ارتباط (0.91) مع المهمة الفرعية للفهم القرائي لاختبار ستانفورد للتحصيل Test (Fuchs et al., 2001; Piper

(Zuilkowski, 2015 &. ويشير الأداء الضعيف على أداة الفهم القرائي أن التلاميذ قد يجدون صعوبة في فك الرموز، وفي القراءة بطلاقة بدرجة كافية لفهم المفردات.

وZuilkowski، 2015). إن الأداء الضعيف على أداء الطلاب في القراءة والفهم يشيران إلى أن الطالب قد تواجهه مشكلة مع فك التشفير، أو مع القراءة بطلاقة بما فيه الكفاية للفهم، أو مع المفردات.

البيانات: تحتسب درجة التلاميذ بناء على عدد الكلمات التي تقرأ قراءة صحيحة في الدقيقة، وعدد أسئلة الفهم التي يُجاب عنها إجابة مقبولة. وسيكون هناك نوعان من درجات التلاميذ: درجات مرتبطة بعدد الكلمات التي قرأت قراءة صحيحة في الوقت المخصص، ونسبة أسئلة الاستيعاب أو الفهم المجاب عنها إجابة صحيحة بنهاية القطعة القرائية. تُجمع الثلاث أقسام للمتغيرات نفسها للمهام الفرعية الأخرى ذات الوقت المحدد إلكترونيًا: بإجمالي عدد الكلمات التي تُقرأ، وعدد الكلمات التي تقرأ قراءة خطأ، والزمن المتبقي. بالإضافة إلى ذلك، فإن نتائج أسئلة الفهم القرائي تُسجل إليكترونيًا، وتُدخل على قاعدة البيانات مع متغير الدرجة النهائية محسوبًا على أساس حصة الأسئلة المطروحة. ويوجه برنامج جمع البيانات الإليكتروني المقيمين إلى طرح الأسئلة التي ترتبط بالنص الذي قرأه الطفل (انظر تركيب الأسئلة، والفقرة التي جاءت تحت عنوان " بناء البند" أدناه).

بناء البند: لبناء مهمة فرعية للطلاقة في القراءة الشفوية مع الفهم، فإن مطورى الأداة يراجعون السرد الموجود في مواد القراءة للأطفال. فالقصة السردية لها بداية حيث يقدم المؤلف شخصيات القصة، ووسط حيث توجد

مشكلة القصة، ونهاية حيث الحدث الذي يسرد لحل المشكلة. فالقصة ليست مجموعة من الجمل المرتبطة بشكل ضعيف. يبلغ طول القصة 60 كلمة.

يُتجنب استخدام أسماء الشخصيات الشائعة في كتب التلاميذ، لأن التلاميذ قد يعطون استجابات آلية بناء على القصص المألوفة لديهم. ومع ذلك، يجب أن تكون أسماء الشخصيات من السياق ومن نفس اللغة. وكذلك أيضًا يوجد بالقصة شخصية واحدة أو شخصيتان، لتجنب أن تكون القصة معتمدة على الاستدعاء من الذاكرة، ويجب أيضًا أن تعكس الأسماء، والأماكن الثقافية المحلية.

يحتوي نص القصة على بعض المفردات المعقدة (على سبيل المثال، المشتقات) تركيبات للجمل. ويستخدم حجم كبير للخط يكون واضحًا مألوفًا، ويوجد مسافات جيدة بين السطور؛ وذلك لتسهيل القراءة للطلاب. ولا يوجد أي تضمين الصور.

وتضم القائمة المرتبطة لأسئلة الفهم أسئلة يمكن الإجابة عنها مباشرة من النص، وكذلك يوجد سؤال استنباطي واحد على الأقل يتطلب من التلاميذ الجمع بين المعرفة والخبرة من خارج النص للإجابة بشكل صحيح. وهذه الأسئلة الاستنتاجية لديها أكثر من إجابة صحيحة واحدة، ولكن يجب أن تكون إجابات منطقية استنادًا إلى النص والسياق.

وأسئلة الفهم الحرفي التي ترتبط مباشرة بالقراءة الشفوية لقطعة القراءة هي أسهل أنواع قياس الفهم. ويمكن وجود أسئلة للفهم الاستنتاجي في المهمة الفرعية تعلمنا ما إذا كان التلاميذ قادرين على ربط محتوى قطعة القراءة مع معرفتهم. سيوضح بروتوكول المهمة الفرعية أنواع الأجوبة التي قد تكون قد قيَّمت على أنها "صحيحة".

عند وضع صور متكافئة لهذه المهمة الفرعية لاستخدامها في تطبيقات متعددة لنفس الأداة في نفس اللغة (على سبيل المثال دراسة خط الأساس، وتقييم نصف المدة، والتقييم النهائي في أي من البلدان التي تُطبق فيها الأداة)، فمن المستحسن إجراء تغييرات بسيطة في القصة من أجل الحد من تأثير تسرب الاختبار، مع الحفاظ على صعوبة مماثلة للاختبار. فعلى سبيل المثال، يمكن استبدال أسماء من موضوعات القصة، والأحداث، والصفات مع بدائل مماثلة على مستوى الصف الدراسي.

الشكل 15: نموذج للمهمة الفرعية للطلاقة في القراءة الشفوية مع مكون الفهم القرائي

جزء 5: قـراءة قطـعة

قدم للتلميذ/ للتلميذة قطعة القراءة الموجودة في آخر صفحة من نموذج التلميذ/التلميذة وقل له :

أريد منك أن تقرأ هذه القطعة بصوت عال وعندما تنتهي من قراءتها سوف أسألك بعض الأسئلة عنها.

هل فهمت المطلوب ؟ عندما أقول ابدأ اقرأ القصة

هل أنت مستعد ؟ ابدأ

شغل ساعة التوقيت عندما تقول للتلميذ/للتلميذة " ابدأ" . ضع علامة " / " لكل كلمة يقرؤها التلميذ/التلميذة بشكل غير صحيح.

احسب الكلمات التي يقوم التلميذ/التلميذة بتصحيح نفسه فيها على أنها صحيحة . التزم الهدوء إلا عندما تكون استجابة التلميذ/التلميذة على النحو التالي: إذا تردد لمدة 3 ثوان ، فقرأ له الكلمة ثم أشر إلى الكلمة اللاحقة وشجعه على أن يستمر في قراءة الكلمات بقولك جهدك طيب ، استمر. احسب الكلمة التي تقرؤها أنت للتلميذ/ للتلميذة على أنها غير صحيحة

. بعد 60 ثانية، ضع علامة] عند آخر كلمة قرأها التلميذ /التلميذة ثم قل " توقف"

صَباحُ الخَيْر. أنا إسْمي سُعاد ، وعُمْري عَشْر سنوات	9
وأنا في الصَّفِ الرابع ، ولي أُخْتٌ صَغيرةٌ اسْمُها جَميلة وعُمْرُها خَمْس سنوات.	24
عِنْدما ألْعبُ مَعَ صَديقاتي تُحبُّ جَميلة أن تَلْعَبَ مَعَنا.	33
جَلَسْتُ مع صَديقاتي لِصنْع طائِرَة وَرَقيّة	39
جاءَت أُختي الصُّغْرى جَميلة وَقالَتْ لي: أُريدُ أنْ ألعبَ مَعَكِ	50
قُلْتُ لها : لا. أنْتِ صَغيرة، اجلسي وَانظُري.	58

الوقت الذي تشير إليه الساعة _____

تم إيقاف التدريب نظراً لأن التلميذ/ التلميذة لم يستطع قراءة السطر الأول. ☐

اسمح للتلميذ/للتلميذة بالاحتفاظ بقطعة القراءة عقب الانتهاء من قراءتها. بعد قراءة كل سؤال اعط التلميذ/التلميذة 15 ثانية على الأكثر للإجابة عن كل سؤال. ضع إشارة √ في العمود الأول إذا كانت الإجابة صحيحة وفي العمود الثاني إذا كانت الإجابة غير صحيحة. وفي العمود الثالث في حالة عدم الإجابة.

الآن سوف أسألك عدة أسئلة عن القطعة التي قرأتها ، حاول الإجابة عن الأسئلة بمكن الاستعانة بالنص.				
كم عمر سُعاد؟	(عشر سنوات)	صحيح ___	غير صحيح ___	لا إجابة ___
في أي صف سعاد ؟	(الصف الرابع)	صحيح ___	غير صحيح ___	لا إجابة ___
ما اسم أخت سعاد ؟	(جميلة)	صحيح ___	غير صحيح ___	لا إجابة ___
ماذا تصنع سعاد ؟	(طائرة ورقية)	صحيح ___	غير صحيح ___	لا إجابة ___
لماذا رفضت سعاد أن تلعب أختها جميلة معها؟	(لأنها صغيرة)	صحيح ___	غير صحيح ___	لا إجابة ___

5.2.6 الوعي الفنولوجي – تعرف الأصوات الأولى أو المبدئية والأصوات الأخيرة: تمييز صوت الحرف

كما هو موضح في القسم 4، "الإطار المفاهيمي والأساس البحثي" لكي نقرأ، علينا أن نحول الحروف التي نراها ونقرأها إلى أصوات، والأصوات إلى كلمات، والكلمات إلى معنى. وتتطلب الإدارة الناجحة لهذه العملية القدرة على العمل في الاتجاه المعاكس. أي من أجل فهم عملية الانتقال من الحروف إلى الأصوات إلى الكلمات، فإن الطلاب يجب أيضًا عليهم فهم أن الكلمات تتكون من الأصوات الفردية، وفهم عملية عزل (والتلاعب بـ) الكلمات إلى الأصوات.

أشار Stanovich (2000) وآخرون الى أن الأطفال الذين يدخلون المدرسة، ولديهم قليل من الوعي الفنولوجي يجدون صعوبة في اكتساب مهارة فك الرموز الأبجدية، وبالتالي يجدون صعوبة في تعرف الكلمات."

أوضح Stanovich (2000) وآخرون أن الأطفال الذين يدخلون المدرسة، ولديهم مستوى محدود من الوعي الفنولوجي يجدون صعوبة في اكتساب مهارة الترميز الأبجدية، وبالتالي تكون لديهم صعوبة في تعرف الكلمات (ص 393). وقد توصلت الأبحاث أن الوعي الصوتي يلعب دورا مهمًا في اكتساب القراءة. وتقييم الوعي الصوتي، وعمل تدخلات علاجية لنقاط الضعف لدي التلاميذ في الوعي الفنولوجي مهم جدًا، لتنمية مهارات اكتساب القراءة مستقبلاً.

تقيس أداة على نحو أكثر شيوعًا الوعي الفنولوجي (وهو أحد جوانب الوعي الصوتي) من خلال تحديد أو تمييز الأصوات الأولية أو الأصوات النهائية. وهذه الأساليب شائعة في اختبار القراءة المبكرة، وتشمل

- Dynamic Indicators of Basic Early Literacy Skills (DIBELS)
مؤشرات ديناميكية لمهارات إجادة القراءة والكتابة المبكرة (DIBELS)
https://dibels.uoregon.edu/ and Dynamic Measurement Group https://dibels.org/

- Test of Phonological Awareness, Second Edition Plus (TOPA-2+)
اختبار الوعي الفنولوجي الطبعة الثانية مزيدة (TOPA-2+)
https://www.linguisystems.com/products/product/display?itemid=10293

- Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)
الاختبار الشامل لمعالجة الفنولوجية الطبعة الثانية (CTOPP-2)
http://www.pearsonclinical.com/language/products/100000737/comprehensive-test-of-phonological-processing-second-edition-ctopp-2-ctopp-2.html#tab-details

المدخل الأول-تعرف الصوت الأول والصوت النهائي:

المدخل الأول لتقييم الوعي الصوتي هو أن يُطلب من الطلاب تحديد أول (أو آخر) صوت في مجموعة مختارة من الكلمات الشائعة. المثال في الشكل 16 يستخدم 10 مجموعات من كلمات بسيطة، ويُطلب من الطلاب على تعرف الصوت الأول في كل من الكلمات. ويقرأ المقيم كل كلمة بصوت عال مرتين قبل أن يطلب من الطالب تعرف الصوت.

البيانات: يسجل المقيِّم عدد الإجابات الصحيحة؟ هذه ليست مهمة ذات وقت محدد من التقييم، ولا تستخدم كراسة أسئلة للتلميذ.

بناء البند: كلمات بسيطة مختارة من قوائم الكلمات التي يدرسها الصفان الأول والثاني. إذا كان ممكنًا بالنسبة للغة ما، ويمكن استخدام كلمات من مقطع واحد حتى لا نرهق الذاكرة العاملة للطلاب.

الشكل 16-نموذج الوعي الصوتي – تحديد الصوت الأول في الكلمة:

SUBTASK 2. PHONEMIC AWARENESS | 🕮 X | Untimed

This is a listening exercise. I want you to tell me the first sound of each word. For example, in the word "pot", the first sound is /p/. I would like you to tell me the first sound you hear in each word. I will say each word two times. Listen to the word, then tell me the very first sound in that word.

Let's practice. What is the first sound in "mouse"? … "mouse"?

[If the child responds correctly, say:] **Very good, the first sound in "mouse" is /m/.**
[If the child does not respond correctly, say]: **Listen again: "mouse". The first sound in "mouse" is /mmm/.**

Now let's try another one: What is the first sound in "day"? … "day"?

[If the child responds correctly, say]: **Very good, the first sound in "day" is /d/.**
[If the child does not respond correctly, say]: **Listen again: "day". The first sound in "day" is /d/.**

Ready? Let's begin.

Read the instructions to the child and conduct the examples.

Read the prompt and then pronounce the word a second time. Pronounce each word slowly.

➲ If the child does not respond after 3 seconds, mark as "No response" and say the next prompt.

If the child responds incorrectly or does not respond to the first five words, say "Thank you!, discontinue this subtask, check the box at the bottom of the page, and go on to the next subtask.

	Item	Answer	Correct	Incorrect	No response
1.	**What is the first sound in "at"? …"at"?**	/a/			
2.	**What is the first sound in "so"? …"so"?**	/s/			
3.	**What is the first sound in "chalk"? …"chalk"?**	/ch/			
4.	**What is the first sound in "very"? …"very"?**	/v/			
5.	**What is the first sound in "car"? …"car"?**	/k/			
6.	**What is the first sound in "for"? …"for"?**	/f/			
7.	**What is the first sound in "man"? …"man"?**	/m/			
8.	**What is the first sound in "ox"? …"ox"?**	/o/			
9.	**What is the first sound in "yes"? …"yes?**	/y/			
10.	**What is the first sound in "go"? …"go"?**	/g/			

Exercise discontinued because the child had no correct answers in the first line

المدخل الثاني: تمييز صوت الحرف

يتطلب المدخل الثاني من التلاميذ أن يستمعوا إلى سلسلة من ثلاث كلمات، ويطلب منهم تحديد الكلمة التي تبدأ (أو تنتهي) بصوت يختلف عن الكلمات الأخرى في هذه السلسلة.

يشمل التصميم النموذجي 10 مجموعات من ثلاث كلمات. في كل مجموعة، كلمتين تبدأن بنفس الصوت، والصوت الأول للكلمة الثالثة مختلف. موضع الكلمة المختلفة يتنوع في المجموعة. يقرأ المقيم كل مجموعة من ثلاث كلمات بصوت عال ببطء، مرتين، ويطلب من الطفل اختيار الكلمة التي تبدأ بصوت مختلف. لأن هذا النوع من المهام قد يكون غير مألوف تماما للأطفال، فإن بروتوكول المهمة الفرعية يشمل مجموعات لممارسة الطفل ليحاول فيها قبل أن يبدأ المقيم الاختبار الفعلي.

البيانات: يسجل المقيم عدد الاجابات الصحيحة. هذه ليست مهمة ذات وقت محدد في التقييم.

بناء البند: كلمات بسيطة مختارة من قوائم الكلمات التي يدرسها تلاميذ الصفين الأول والثاني. كما في مدخل صوت الحرف الأول، ويمكن استخدام كلمات من مقطع واحد أو مقاطع قليلة حتى لا نرهق الذاكرة العاملة للتلاميذ.

الشكل 17: عينة للمهمة الفرعية تمييز صوت الحرف من أداة

جـزء 2: الوعـي الصوتـي

هذا الجزء لا يعتمد على التوقيت . **لا يوجد نموذج أو ورقة للتلميذ/ للتلميذة في هذا التمرين**، وبدلاً من ذلك فإن الفاحص يقوم بالقراءة بصوت عال للتلميذ/ للتلميذة.

اقرأ بصوت عال للتلميذ/للتلميذة واطلب منه/ التعرف على الصوت الأول في الكلمات .

يعتمد هذا التمرين على الاستماع.

الآن سوف أقول ثلاث كلمات ، التفق من بينهما تبدأ بنفس الصوت وواحدة فقط تبدأ بصوت مختلف، اريد ك أن تعرف أي كلمة هي التي تبدأ بصوت مختلف . فمثلا إذا قلت لك الكلمات التالية: **(مريم – منى – هدى)** فإن **"هدى"** هي التي تبدأ بصوت يختلف عن الميم وهو صوت **"هـ"** لأن **"مريم"** و**"منى"** يبدان بصوت **"مـ"**

دعنا نحاول ، قل لي ما الكلمة التي تبدأ بصوت مختلف

"تفاح" ، "برتقال" ، "تمر" أي من هذه الكلمات الثلاث يبدأ بصوت مختلف؟ إذا أجاب التلميذ/ التلميذة بشكل صحيح قل **"أحسنت"** الكلمة التي تبدأ بصوت مختلف هي**"برتقال"** . أما إذا أخطأ للتلميذ/ للتلميذة قل له أن الكلمة التي تبدأ بصوت مختلف هي **"برتقال"** لأنها تبدأ بحرف **"ب"**

دعنا نحاول مرة أخرى ، قل لي ما الكلمة التي تبدأ بصوت مختلف

"زرافة" ، "طاولة" ، "زهر" أي من هذه الكلمات الثلاث يبدأ بصوت مختلف؟ إذا أجاب التلميذ/ التلميذة بشكل صحيح قل **"أحسنت"** .الكلمة التي تبدأ بصوت مختلف هي **"طاولة"** لأنها تبدأ بصوت **"ط"**. أما إذا اخطأ التلميذ قل له الكلمة التي تبدأ بصوت مختلف هي "طاولة" لأنها تبدأ بصوت **"ط"**.الآن جاء دورك، قل لي ما الصوت الأول المختلف في الكلمات التالية:

هل فهمت المطلوب منك عمله ؟

إذا لم يستجب التلميذ/التلميذة بعد 10 ثوان ضع علامة خطأ ثم استمر . الفظ كل كلمة مرتين . (بمعدل كلمة في الثانية) ضع إشارة(√) عند المكان الصحيح

أ – تعرف الصوت الأول المختلف :

			صحيح	غير صحيح	لا إجابة
عصفور	صنبور	صندوق			
حديد	حليب	**جديد**			
قلم	**سمر**	قمر			
ناس	نار	**عام**			
لبن	جزر	جمل			
سحر	سمر	**شجر**			
طابور	**شاكوش**	طاووس			
ورقة	**ثمرة**	وردة			
فلفل	معرض	منزل			
بقرة	بكرة	**شجرة**			

6.2.6 قراءة الكلمات المألوفة:

وغالبا ما تُقيم مهارات القراءة للأطفال باستخدام قوائم قراءة كلمات غير مترابطة، وهذا يسمح لقياس أنقى من تعرف الكلمات ومهارات فك الرموز أكثر من قراءة كلمات نص متصل حيث يكون التلاميذ غير قادرين على تخمين الكلمة التالية من السياق عند قراءة قوائم كلمات غير مترابطة.

وفي هذا التقييم، فالكلمات المألوفة هي كلمات ذات تكرار عال مختارة من مواد القراءة، و كتب القصص للصفوف الأول، و الثاني، و الثالث في اللغة والسياق.

البيانات: هناك ثلاثة متغيرات تُجمع من حساب هذه النتيجة: عدد الكلمات التي قُرئت، وعدد الكلمات التي قُرئت خطأ والزمن المتبقي، وهو كما يجرى في مهمة تعرف صوت الحرف، وفي المهام ذات الوقت المحدد.

بناء البند:

تُوضع قوائم كلمات لهذه المهمة من كتب القراءة الوطنية من مستويات الصفوف التي ستشترك في الدراسة. تفيد التحليلات لتكرار الكلمات في هذه النصوص في اختيار 50 كلمة شائعة، ومألوفة، وبسيطة، وتمثل الأجزاء المختلفة للكلام (مثل الأسماء، والأفعال، إذا كان ذلك ممكنًا). إلى أقصى حد ممكن، وينبغي أن يكون نطق الكلمات لا لبس فيه ومألوفًا في اللغة أو للهجة ذات الصلة.

تُرتب الكلمات أفقيا مع تباعد مناسب وواضح، ومكتوبة بخط مألوف للتلاميذ. يوجد في المهمة 10 صفوف خمس كلمات في كل صف، وتظهر البنود في ترتيب عشوائي، ولا توضع البنود وفقًا للصعوبة، و الطول، و الترتيب الأبجدي ، وغيرها) في الشبكة. يمكن إعادة ترتيب البنود في الصفوف (يُعاد ترتيبها بشكل عشوائي)، وذلك لإعداد نماذج اختبار متكافئة، على الرغم من أنه ينصح اختبار أثار الترتيب.

اعتمادًا على خصائص اللغة، تتضمن الكلمات المختارة توازنًا بين الكلمات المألوفة ا التي يسهل قراءتها (فك الرمز) (مثل " لعب")، والكلمات البصرية (على سبيل المثال "، و "). وتعتمد أيضًا على طول الكلمة، والإملاء، و نمط الكتابة ممثلة لما هو موجود في القراءة في الصفوف الأولى ، والكلمات مكونة من مجموعة متنوعة من الحروف مع عدم وجود تكرارات غير متناسبة. ولا يشتمل أي بند من بنود المهمة على كلمات أعجمية من لغات أخرى قد تكون

مألوفة لدى التلاميذ. ولا تتضمن هذه المهمة كلمات مكونة من حرف واحد، وهذه سوف تدرج بالفعل في شبكة الحرف. ويمكن أن تكون ثلاث كلمات إضافية بمثابة كلمات أمثلة يستخدمها المقيمون؛ لتدريب الطلاب عليها. ويجب أن تكون الكلمات مماثلةً في مستوى الصعوبة للكلمات الموجودة في الشبكة. وينبغي استخدام نفس نوع الخط المستخدم في الكتاب المدرسي ، أو الاستعانة بكتب أخرى تستخدم نفس الخط ، والطباعة إذا لم يتوفر ذلك في كتاب المدرسة. انظر أيضا المناقشة الموجزة حول الخطوط في قسم تعرف صوت الحرف أ (القسم 6.2.2).

الشكل 18 نموذج للمهمة الفرعية لقراءة الكلمات المألوفة من أداة

جزء 3: قراءة الكلمات المألوفة

قم للتلميذ/للتلميذة نموذج الكلمات الموجودة في الصفحة الثالثة من النموذج الخاص بالتلميذ/التلميذة ، قل:

فيما يلي بعض الكلمات. أريد منك أن تقرأ لي جميع ما تستطيع قراءته من الكلمات المعروضة أمامك

هذا مثال على ذلك : "باب "

الآن حاول أنت : اقرأ لي هذه الكلمة " أمّي " إذا استجاب التلميذ/التلميذة بشكل صحيح، فقل أحسنت ، هذه الكلمة هي " أمّي " . أما إذا أخطأ التلميذ/التلميذة ، فقل هذه الكلمة هي " أمّي "

الآن حاول مرة أخرى اقرأ لي هذه الكلمة " صف" إذا استجاب التلميذ/التلميذة بشكل صحيح، فقل أحسنت ، هذه الكلمة هي " صف. أما إذا أخطأ التلميذ/التلميذة ، فقل هذه الكلمة هي " صف

. هل فهمت المطلوب ؟

هل أنت مستعد ؟ ابدأ

شغل ساعة التوقيت عندما يبدأ " التلميذ/ التلميذة في القراءة. ضع علامة " / " لكل كلمة يقرؤها التلميذ/التلميذة بشكل غير صحيح.

احسب الكلمات التي يقوم التلميذ/ التلميذة بتصحيح نفسه فيها على أنها صحيحة . التزم الهدوء إلا عندما تكون استجابة التلميذ/التلميذة على النحو التالي: إذا تردد التلميذ/التلميذة لمدة 3 ثوان ، اقرأ له الكلمة ثم أشر إلى الكلمة اللاحقة و اطلب منه أن يستمر في قراءة الكلمات. احسب هذه الكلمة التي قرأها على أنها غير صحيحة . إذا اخطأ التلميذ في خمس كلمات انتقل فوراً إلى السؤال التالي

بعد 60 ثانية ، ضع علامة] عند آخر كلمة قرأها التلميذ/ التلميذة. ثم قل " توقف

أحِبُّ	غُراب	ماما	هُوَ	يَزرَسْ	5/
بِنْت	شُبّاك	كَلام	ضَايط	أنتَ	10/
أرْنَبَ	بوق	ثَعْلَب	وَقَفَ	وَرْدَة	15/
صَفْراء	ساعَة	عَصافير	مَركب	سيكين	20/
غَزّال	زَرَعَ	حَديقَة	هَرَم	وَطَني	25/
نَار	عَصير	صَديقَتي	التّلاميذ	ديك	30/
عِنْدَما	الشّارع	لِماذا	لَبَن	نَلعب	35/
حَياتي	يَعيش	بَيت	الهُدوء	نُقود	40/
يُذاكِر	صاروخ	هُدْهُدْ	أصَابِع	شُكراً	45/
نَأكُلُهُ	رايحينكُما	الرِزْقُ	اسْمَع	النّافِذَة	50/

الوقت الذي تشير إليه الساعة _____

☐ تم إيقاف التدريب نظراً لأن التلميذ/ التلميذة لم يستطع قراءة السطر الأول.

6

3.6 عرض للمكونات الإضافية للأداة:

كما سبق وذكرنا أعلاه، توجد هناك مجموعة أقل شيوعًا من ناحية الاستخدام طُورت، وجُرِبت استنادًا على عوامل مثل: خصوصية اللغة، وأسئلة بحثية معينة وأسئلة متعلقة بالمنهج، ويوجد وصف لهذه العوامل بإيجاز فيما يلي:

1.3.6 الإملاء

وكثيرا ما يُستخدم تقييم الإملاء من قبل المعلمين؛ لاختبار

مهارات الفهم الشفوي والكتابة على حد سواء. كما نوقش في وقت سابق، ويمكن اختبار عملية القراءة في الاتجاه المعاكس: كقدرة الطلاب على سماع الأصوات بشكل صحيح، وكتابة الحروف، والكلمات المقابلة للأصوات التي يسمعونها مما يدل على نجاحهم في المبدأ الأبجدي (إيجاد العلاقة بين الرمز المكتوب والصوت)، وهناك عدد من حزم التقييم المقدمة من قبل المتخصصين في مجال تصميم الاختبار التجاري يقدمون للمعلمين تعليمات حول كيفية تطوير وتسجيل درجات تقييماتهم الخاصة. قدمت خدمة التقييم التربوي ومبادرة إجادة القراءة والكتابة لدى الأطفال نماذج للمعلمين – ولم يعد هذان النموذجان موجودين.)، والبحث من قبل جمعية القراءة الدولية ((Denton, Ciancio, & Fletcher, 2006). وكانت هذه المهمة الفرعية جزء من أداة الأصلية، ومع ذلك، حُذفت في وقت لاحق بسبب صعوبات في تقنين تقدير الدرجات والتطبيق. واعتبارا من عام 2015، أصبحت الإملاء تستخدم، ولكن ليس على نطاق واسع.

البيانات: يعطى التلاميذ الدرجة بناء على مقياس بسيط يقيس قدرة التلاميذ على كتابة الحروف بشكل سليم.

بناء البند: تُختار خمسة من الحروف الأكثر شيوعًا في اللغة التي تُقيم في هذه المهمة الفرعية.

الشكل 19 نموذج للمهمة الفرعية للإملاء في جزئيين يتضمنان كتابة حروف مفردة ثم كتابة كلمات قصيرة عديدة (في مقابل الجمل)

جزء 7: الإملاء

هذا التمرين لا يعتمد على التوقيت . اطلب من التلميذ/ التلميذة أن يكتب الجملة خلف أي صفحة تمرين

اقرأ الجملة التالية بصوت عال مرة واحدة (بمعدل كلمة في الثانية الواحدة) . ثم اطلب من التلميذ/ التلميذة الكتابة. ثم أعد الجملة مرة ثانية وببطء (أي بمعدل كلمة في الثانية الواحدة). توقف لمدة 5 ثوان ثم أعد الجملة مرة ثالثة في أثناء كتابة التلميذ/ التلميذة لها. لا تقرأ أكثر من ثلاث مرات.

قل للتلميذ/ للتلميذة

سوف أقرأ عليك بصوت عال جملة قصيرة ثم اطلب منك كتابتها. رجاء الاستماع بحرص في المرة الأولى ثم ابدأ بكتابتها . ثم ساعيد قراءة الجملة مرتين . هل فهمت المطلوب؟

شاهد أحمد العصفور الأزرق يطير عن الشجرة .

م	معيار التقدير	غير صحيح (صفر)	صحيح جزئيا (1)	صحيح (2)
	حجم الفراغ بين الكلمات مناسب	0 – 2 فراغات	3 – 5 فراغات	6 فراغات
1	شاهد			
2	أحمد			
3	العصفور			
4	الأزرق			
5	يطير			
6	عن			
7	الشجرة			

2.3.6 تقسيم المقاطع الصوتية:

تقسيم المقاطع الصوتية، وهو ما يعني في هذه الحالة تقسيم الكلمات إلى وحدات صوتية، وهي واحدة من المهارات الأكثر تعقيدًا في الوعي الصوتي، وينبغي التأكيد عليها في الصفوف الأولى (Linan-Thompson & Vaughn, 2007)

بل لديه القدرة على التنبؤ بمهارات التعلم في وقت لاحق، وتميل المهام الفرعية في أداة ، والمصممة بالشكل الذي يختبر تقسيم المقاطع الصوتية إلى أن يكون من الصعب تطبيقها، وأظهرت مشكلات متعلقة بأدنى تأثير (أي أن عددًا قليلاً جدًا من الطلاب قادرون على إكمال المهمة الفرعية).

والمهمة الفرعية الأصلية للوعي الصوتي، والتي تضمنت تجزئة المقاطع الصوتية. وبالنسبة لهذا الجزء من التقييم، يقرأ المقيم قائمة من 10 كلمات بسيطة مكونة من

مقطعين مثلا في المرة الواحدة، ويُطلب من الطلاب تحديد ونطق الأصوات في الكلمة حيث إن التقييم يُنفذ سماعيًا فلن يكون هناك أية أوراق أسئلة للطالب، وحذفت هذه المهمة فيما بعد من الأداة الأصلية بسبب الحد الأدنى للتأثير). وبالرغم من ذلك فقد استخدمت في عدد من السياقات في الآونة الأخيرة حيث يكون الطلاب أكثر ألفة مع الوعي الصوتى.

البيانات: يسجل المقيِّم عدد المقاطع الصوتية الصحيحة بالنسبة لعدد المقاطع التي حاول الطالب الإجابة عنها. هذا الجزء من التقييم ليس له زمنًا محددًا.

بناء البند: كلمات مكونة من مقطعين أو ثلاثة في بنية الكلمات التي تستخدم المقاطع الصوتية، والتي لا تحتوي على تراكيب لغوية معقدة، ويراعى فيها تنوع الأصوات التي تأتى في أول الكلمة.

الشكل 20: مثال للمهمة الفرعية لتقسيم المقاطع الصوتية

Sesaun 3b. Kualidade ba Fonema sira - 2

*Ida ne'e "LA'OS" ezersísiu atu marka ho tempu **NO LA EZISTE FORMULÁRIU BA LABARIK.** Lee ho lian maka'as kada liafuan ida-idak dala rua hafoin husu labarik atu dehan ninia son sira.*

Ida ne'e ezersísiu kona-ba rona. Ó hatene katak kada letra iha nia son ida-idak. Ezemplu, "casa", "c"-"a"-"s" -"a" bele rona "/c/ - /a/ - /s/- /a/". Ha'u sei temi liafuan ne'e dala rua, entaun rona didi'ak.

Ita koko. Son saída mak iha liafuan "par" – "par"?
[sé labarik hatán loos, dehan]: **Di'ak, son ba liafuan "par" mak /p/ /a/ /r/.**
[sé labarik hatán laloos, dehan]: **son ba liafuan "par" mak /p/ /a/ /r/.**
Agora ó fali. Dehan mai ha'u son iha liafuan "par". *[hein ba labarik to'o segundu 5 atu hatán].*

Ita koko fali seluk. Son saída mak iha liafuan "mar" – "mar"?
[sé labarik hatán loos, dehan]: **Di'ak, son ba liafuan "mar" mak /p/ /a/ /r/.**
[sé labarik hatán laloos, dehan]: **son ba liafuan "par" mak /p/ /a/ /r/.**
Agora ó fali. Dehan mai ha'u son iha liafuan "par". *[hein ba labarik to'o segundu 5 atu hatán].*

Agora ita komesa. Ha'u sei temi liafuan ida – ida dala rua. Rona ba liafuan ida-ida no dehan mai ha'u liafuan ne'e nia son. Ó hatene ona atu halo saidá?

*Pronunsia **DALA RUA** ho neineik kada liafuan ida-idak ninia konjuntu (liafuan ida segundu ida).*

*Marka ho barra (**/**) ba resposta ne'ebé laloos, no moos ba son ne'ebé sira la temi*

ou hakat liu.

Sé labarik la hatán ba liafuan ida iha segundu 5 nia laran, entaun marka son sira iha liafuan ne'e laloos no kontinua ba liafuan seluk.

Regra atu hapara ezersísiu: *sé labarik la fó resposta ida ka rua ne'ebé loos iha liafuan 5 dahuluk, dehan "**obrigada barak",** keta kontinua ezersísiu ne'e, marka kaixa mamuk ida iha okos hafoin pasa fali ba ezersísiu tuir mai.*

Son saída mak iha_____?_____/ [*repete liafuan ne'e dala rua*]]					
"lar"	/l/	/a/	/r/		_____/3
"era"	/e/	/r/	/a/		_____/3
"paz"	/p/	/a/	/z/		_____/3
"sal"	/s/	/a/	/l/		_____/3
"mal"	/m/	/a/	/l/		_____/3 **Linha 5**
"arroz"	/a/	/r/	/o/	/z/	_____/4
"casa"	/c/	/a/	/s/	/a/	_____/4
"fala"	/f/	/a/	/l/	/a/	_____/4
"massa"	/m/	/a/	/s/	/a/	_____/4
"carro"	/c/	/a/	/r/	/o/	_____/4

*Marka **X** iha kaixa tuir mai ne'e sé ezersísiu la kontinua tanba labarik la hatán loloos iha liafuan 5 dahuluk: ☐*

Servisu di'ak! Mai ita kontinua sesaun tuir mai.

3.3.6 المتاهة

تقييم المتاهة غير شائع في الفصول المدرسية باعتبارها وسيلة لاختبار الفهم، وتتطلب المتاهة تأليف فقرة موجزة عبارة عن نص متواصل، أو سرد، وهو الذي يحكي نوعًا من القصة.

ثم يقوم مصمم الاختبار باستبدال كلمة كل 15 كلمة يوضع مكانها سؤال من نوع الاختيار من متعدد من ثلاث كلمات ممكنة، وبعضها يكون له معنى في السياق بينما البعض الآخر غير ذي صلة بسياق الموضوع.

يقرأ الطالب هذه البدائل قراءة صامتة أو بصوت عال، ويختار كلمة واحدة من بين البدائل الثلاث من التي من شأنها أن تكمل العبارة بشكل أفضل.

يتطلب Cloze نفس النوع من حذف كلمة، ولكن يستبدل الكلمات الناقصة بمساحة فارغة، ويسمح للطالب بإعطاء الكلمات الناقصة. ولم يُفحص هذا المقياس على نطاق واسع في سؤال الـ Cloze، ويمكن أن يكون هذا السؤال صعبًا جدًا من حيث البناء بطريقة مناسبة لمستويات التلاميذ في القراءة، وبالرغم من ذلك فهو نوع جديد أو غير مألوف من المهام للتلاميذ، والتي يمكن أن تتسبب في أن يؤدوا بشكل ضعيف. ولا يوجد توفيت لهذه المهمة، ومع ذلك، فإن لها حدًا زمنيًّا يتعين على التلاميذ استكمال والانتهاء من المهمة الفرعية خلاله (عادة 3-5 دقائق).

البيانات: تُعطى الاستجابات درجة اعتماد على مفتاح يحدد الاستجابات المقبولة، وتُجمَّع الاستجابات الصحيحة من إجمالي عدد البنود.

بناء البند: الجملة الأولى والجملة الاخيرة من القصة كاملتان، ولا توجد بهما كلمات ناقصة. أما باقي الجمل ففي كل جملة توجد كلمة واحدة ناقصة. ومع ذلك، قد توجد عبارة قصيرة ناقصة من جملتين. وبالنسبة للجمل التي بها كلمة ناقصة، لا يجب أن تكون الكلمة الأولى. وكذلك توجد كلمة واحدة أو عبارة واحدة فقط ناقصة في الجملة ، و يوجد ثلاثة بدائل لكل كلمة ، أو عبارة ناقصة. فقد تكون البدائل معقولة، ولكن ليست ممكنة بالفعل. وتوجد إجابة واحدة فقط صحيحة في مجموعة الاستجابات. فعلى سبيل المثال، في الجملة التالية/ " كانت البحيرة كبيرة، ورأى السمكة (تسبح-تستلم-تحفر) وقد يجادل أحد ان السمكة قد تحفر في الرمال، ولكن الإجابة الأكثر معقولية هي " تسبح".

[15]يفضل أن تكون الكلمة التي ستستبدل ليست الكلمة الأولى في الجملة.

الشكل 21: متاهة الفهم (مصر)

Jane does not like to do homework. When she gets home from school (**cat, she, fly**) only wants to play. Jane tells (**table, her, red**) mother that a full day of (**scary, school, house**) and homework is too much! Her (**hen, to, mother**) tells Jane she needs to do (**her, book, run**) homework before playing. Jane tells her (**goat, mother, work**) that she is just a little (**child, leg, three**), so she needs more time to (**eats, hat, play**).

One day, Jane decides she will (**under, never, dog**) do homework again. She does not (**bring, throw, with**) her books home from school anymore. (**She, Class, Jump**) feels that she is on a (**shop, holiday, on**). Jane is happy. After two full (**weeks, chairs, up**), she takes her exams. Jane gets (**dirty, ear, poor**) marks. Her mother is very angry, (**and, in, hot**) Jane is very sad. She cries (**to, for, walk**) a long time.

Her big brother (**on, pulls, comes**) to see her. She tells him (**in, pen, about**) her homework and her very poor (**marks, pigs, fear**). The brother tells her that doing (**shirt, have, homework**) will help her a lot. Now, (**box, run, Jane**) knows that homework can help her (**get, drink, bus**) good marks. She now wants to (**fall, work, pink**) hard and do her homework.

4.6 الأسباب وراء استبعاد مكونات أخرى محتملة للأداة

خلال تطوير الأداة الأصلية، أثمرت مراجعة الأدب التربوي، والخبراء العديد من المقترحات؛ لإدراج مكونات إضافية للاختبار. وبعد مراجعة كل هذه الاقتراحات، صُمِمَت مجموعة من معايير الاختيار لمدى ملاءمة ومناسبة دمجها في الأداة. وكان الاعتبار الأساسي هو الإجابة عن هذا التساؤل: ما مدى جدوى اختبار تقييم مهارات القراءة للصفوف الأولى في التنبؤ بنجاح التلميذ في القراءة.

قبل كل شيء كان من بين هذه الاقتراحات إدراج مهمة فرعية قائمة على الصورة مثل: تلك الموجودة في اختبار مفردات الصورة Peabody (PPVT)، وهو اختبار متاح تجاريًّا من مجموعة التعلم لبيرسون. وبعض المتغيرات من أدوات تقييم القراءة في الصفوف الأولى، وتضمَّن ذلك نسخة تطبيقها، والتي نفذته مؤسسة بلان إنترناشيونال باللغة الفرنسية في غرب

أفريقيا؛ وبها صور لتحديد الالمام بالمفردات الشائعة (كأجزاء الجسم مثل: اليد، والرأس، ...)

وبالرغم من ذلك، فإن مطوري الأداة الأصلية لـ حذفوا الصور، أو بالأحرى اختبارات مفردات الصور لعدة أسباب سنطرحها على النحو التالي: (1) تُقاس المفردات بشكل غير مباشر في كل من فهم المسموع، وأثناء قراءة القطعة القرائية. (2) تطوير الصور كثيرًا مما يصطدم بقضايا حقوق الطبع (كاستخدامPPVT ، فعلى سبيل المثال، لم يُؤخذ في الاعتبار به كخيار و ذلك لأنه يحتاج إلى تصاريح حقوق التأليف والنشر في كل مرة سيُستخدم فيها أثناء تطبيق الأداة في أي بلد أو دولة)؛ (3) مواجهة صعوبة شديدة في ابتكار الصور التي من شأنها أن تكون ملائمة عالميًا لجميع الثقافات والسياقات ، أو الاعتماد على الرسامين من داخل تلك البلاد ؛ بهدف ابتكار العمل الفني الأصلي في إطار الفترة الوجيزة المخصصة لتطويع أداة .. وبالإضافة إلى ذلك، عند تطوير الصور محليًا فقد أثبتت التجربة أنه على الأقل ستوجد مشكلتان: أولًا-الصور في كثير من الأحيان ستكون ذات جودة منخفضة جدًا، مما يجعل من الصعب حتى للكبار من ذوي الخبرة أن يفسروا الصورة، وأن يجيبوا عن الأسئلة المطروحة عليها. ثانيًا، مع افتراض جودة الصورة العالية، وتطوير البنود الملاءمة للصورة، فإنه يبدو أن الأمر يحتاج إلى مهارة كبيرة بسبب صعوبة وجود صور تُفسَر وتُفهم بنفس التفسير وبطريقة مقبولة في العديد من الدول. القضايا المتعلقة بالارتباط الثقافي تجعل القياسات القائمة على الصورة صعبة جدًا في تطويرها وتقنينها.

وقد اشتق مكون آخر واختُبِرَ في البداية وحُذفَ من التقييم في وقت لاحق ماري كلاي Marie Clay (1993) كمفاهيم تقييم اللغة المكتوبة. وأوضحت التطبيقات المبكرة لمهمة فرعية تتطلب من الأطفال الإشارة إلى أين نبدأ القراءة (اتجاه المادة أو النص المقروء) ؟، وماذا نقرأ بعد ذلك؟، واتجاه القراءة، فقد أظهرت الأثر الأعلى للنتائج (أكمل جميع الأطفال تقريبًا المهمة بنجاح). ولذلك فإن النتائج المستسقاة، أو التي حُصل عليها في كل من الولايات المتحدة الأمريكية والبحوث الدولية، ذكر كتاب علم النفس أن الوعي باللغة المطبوعة يبدو ان له قدر محدود من القوة التنبؤية لمهارات القراءة فيما بعد. بل إنه يخدم أساسًا كمقياس موثوق في بياناته حول التعرض للطباعة والتمكن من القراءة و .(Paris & Paris, 2006). وفى ضوء هذه النتائج بالإضافة إلى ضعف الكفاءة وقيود الوقت، ولم تُتضمن في أداة كمهمة فرعية تحت عنوان " الوعي بالطباعة".

5.6 الترجمة واعتبارات لغوية أخرى

1.5.6 الترجمة مقابل التطويع

التوافق بين خبراء التعليم عندما يطور المقيِّمون أو يعدلِّون أداة، فمن غير المجدي ببساطة ترجمة الكلمات أو قطعة لنص متصل صدر بلغة مختلفة. بكل بساطة يمكن أن تؤدي الترجمة إلى استخدام كلمات غير مناسبة للغتهم الأم فقد تكون صعبة للغاية بالنسبة لمستوى تلاميذ نفس الصف.فعلى سبيل المثال ، عند ترجمة المهمة الخاصة بعزل المقاطع من اللغة الإنجليزية إلى العربية مثلا عند ترجمة كلمة Yesterday ، وهي مكونة من ثلاثة مقاطع صوتية في اللغة الإنجليزية ، تُترجم إلى كلمة (أمس) وهي مكونة من مقطعين صوتيين في اللغة العربية قد يؤدي هذا إلى مقارنة كلمة مكونة من ثلاثة مقاطع في لغة بكلمة أخرى مكونة من مقطعين في لغة أخرى ، وقد يُضعِف هذا من ثبات أداة التقييم وصدق مقارنة النتائج بين اللغات المختلفة ، كما سبق مناقشة هذا . ولكن العمل الدقيق في ورشة عمل التطويع يُنتج قطعًا أصيلة مساوية تقريبًا في الصعوبة للنص الذي يتوقع من التلاميذ قراءته داخل الصف.

ومن الأمور المهمة الذكر أنه يجب ترجمة التعليمات بلغة أقرب ما تكون لتعليمات الأداة الأصلية التي تركز على المعنى أكثر من الألفاظ.

وكما أشارت Penelope Collins في ورشة عمل تطوير أداة في عام 2006 إلى خبرة شخصية في التواصل مع قسم التعليم في جنوب افريقيا،

أنه بسبب الاختلافات اللغوية (والمقصود بها الكتابة والمورفولوجيا) فمن المهم جدًا أن تكون القطع المستخدمة تُكتب بشكل مستقل. ولا يمكن تحقيق التساوي بين القطع عن طريق الترجمة للقطع المكتوبة باللغة الإنجليزية إلى لغات مختلفة.

وأوضحت هذه الملحوظة الدراسة الاستطلاعية الأولية للقطع المكتوبة بالزولو. كانت قطع الزولو ترجمة للقطع المكتوبة باللغة الإنجليزية أي بنفس الطريقة والأسلوب والقواعد المستخدمة في اللغة الإنجليزية أو قريبة منها. وبالرغم من أنه يمكن توقع مستوى الطلاقة في القراءة الشفوية لدى الأطفال أن يكون مشابهًا لقوائم الكلمات المألوفة، والكلمات غير المألوفة، وقراءة القطع التي يقرأنها، فان التلاميذ الذين قرأوا من 20 – 30 كلمة قراءة صحيحة في الدقيقة بلغة الزولو لم يتمكنوا من قراءة القطعة على الاطلاق. وعند فحص القطع المكتوبة بلغة الزولو، وجد أن الكلمات بلغة الزولو كانت أطول بكثير من الكلمات التي استخدمت باللغة

الإنجليزية، وبالتالي فإن القطعة التي كتبت بلغة الزولو كانت صعبة جدًا على التلاميذ في الصف الأول أن يقرأها.

النص باللغة الانجليزية

> "John had a little dog. The little dog was fat. One day John and the dog went out to play. The little dog got lost. But after a while the dog came back. John took the dog home. When they got home John gave the dog a big bone. The little dog was happy so he slept. John also went to sleep."

النص بالزولو

> *IsiZulu*: "USipho wayenenja encane. Inja yakhe yayikhuluphele. Ngolunye usuku uSipho wayehamba nenja yakhe ukuyodlala. Inja yalahleka. Emva kwesikhathi inja yabuya. USipho waphindela ekhaya nenja yakhe. Emva kokufika ekhaya, uSipho wapha inja ekhaya ukudla okuningi. Inja yajabula kakhulu yaze yagcina ilele. NoSipho ngokunjalo wagcina elele."

2.5.6 المقارنة بين اللغات: الاستعدادات والاعتبارات

إن مسألة المقارنة بين اللغات والبلدان تمثل تحديًا من منظور التقييم، ولذلك فإن التقييمات التي أُجريت باستخدامs في سياقات مختلفة، أو في لغات مختلفة قد تكون قد استخدمت صورًا من الاختبار قابلة للمقارنة، وهذا يعني أن المقصد من الاختبارات أن يحكم عليها بالعلاقة مع بعضها البعض، ولذلك صُمِمت بنفس البنية والمهام الفرعية وما إلى ذلك. وهذا يعنى أن الصور نفسها لها نفس الغرض من القياس، ومع ذلك، ليس هناك افتراض التكافؤ (أي صعوبة متطابقة للبند في الإصدارات المختلفة).

وتشير الأبحاث إلى أن الفرق بين اللغات قد يكون في المقام الأول مسألة المعدل الذي يحققه الأطفال في الخطوات القليلة الأولى نحو اكتساب وتعلم القراءة (Seymour et al., 2003). وبغض النظر عن اللغة، فإن الأطفال الذين يتعلمون قراءة يتقدمون من كونهم ليسوا قراءً غير قادرين على قراءة الكلمات إلي كونهم قراءً جزئيًا (أي يمكنهم قراءة بعض البنود) إلى كونهم قراء (يمكنهم قراءة كل أو غالبية البنود). في اللغات ذات نظام الكتابة الهجائية الشفافة أو "الضحلة" (غالبًا ما تسمى لغات مكتوبة صوتيا)، فإن التقدم من خلال هذه المستويات هو (مجرد بضعة أشهر من التعلم) سريع جدا. في المقابل، وأمَّا اللغات ذات

التهجئة الأكثر تعقيدًا أو "أعمق"، فإن هذه العملية يمكن أن تستغرق عدة سنوات. ففي اللغة الإنجليزية، على سبيل المثال، فاستكمال الخطوات التأسيسية يتطلب سنتين أو أكثر، مع زيادة معدل اكتساب عدد قليل من البنود الجديدة مع كل شهر للتعلم. و بالمقارنة، هناك مجموعة أخرى من اللغات الأبجدية المنتظمة والشفافة مثل : الإيطالية والفنلندية واليونانية يستغرق الأمر نحو عام من التدريس للطلاب ؛ للوصول إلى مستوى مماثل من الكفاءة في القراءة (Seymour et al., 2003).

حيث إن اللغات لها مستويات مختلفة من الشفافية الهجائية، فإنه ليس من السهل القول بأن بلد "أ" (جميع الأطفال يقرأون فيه بآلية وتلقائية بنهاية الصف الثاني) يتفوق على البد "ب" (حيث يصل الأطفال إلى هذا المستوى نفسه فقط بنهاية الصف الثالث)، إذا كانت لغة البلد" لديها نظام كتابة أكثر شفافية من لغة البلد "ب".

بالإضافة إلى الشفافية، فإن معدل اكتساب مهارات القراءة يتأثر أيضًا بدرجة تعقيد نظام الكتابة الذي يتضمن نظام الكتابة عدد الرموز التي ينبغي تعلمها وتعقيدها البصري. فعلى سبيل المثال، فإن معظم نظم الكتابة في جنوب آسيا وجنوب شرق آسيا لديها ما يقرب من 500 مجموعة من الرموز التي لابد من اكتسابها(Nag & Perfetti, 2014) ، وبالتالي فإن عملية اكتساب القدرة على القراءة في هذه اللغات قد تستغرق 4-5 سنوات (Nag, 2007) مقارنة بـ 2-3 سنوات في نظم الكتابة الأبجدية العميقة وسنة واحدة في نظم الكتابة الأبجدية الضحلة.(Seymour et al., 2003)

ويعد التعقيد البصري أيضًا عاملا مهمًا في تعقيد نظام الكتابة، حيث تُعرض الحروف بيانيًّا لا خطيا، وتظهر الرموز طبيعيًّا في أدنى، أو فوق، أو من اليسار، أو إلى اليمين، أو إلى الصوت الذي يمثلونه بالتتابع. ففي بعض الحالات، كما في اللغة العربية، فإنها لا تظهر على الإطلاق بصريًّا، ويستدل عليها من خلال السياق. وهذه العوامل تؤثر أيضًا على معدل اكتساب العديد من مهارات القراءة، (e.g. Kandhadai & Sproat, 2010; Karanth, 2002)، وتحتاج إلى مراعاتها عند عقد المقارنة بين اللغات.

وهناك عامل أخر مهم في المقارنات بين اللغات، وهو الفروق في علم الأصوات؛ وذلك لأن بعض اللغات تمثل الأصوات في صورة مقاطع، والبعض فونيمات فقط، والبعض مزيج بينهما (Perfetti, 2003)، فمن المهم تحديد نوع مقياس الوعي الصوتي الأكثر مناسبة للتركيب الفنولوجي للغة.

ومع ذلك، فإن معرفة الصف الذي يحقق فيه الأطفال "تقدمًا في القراءة في العديد من الدول، على سبيل المثال، و مقارنة هذه الصفوف مع صفوف أخرى، يمكن أن يكون تحليلاً مفيدًا و تمرينًا عن السياسات لهذه البلدان، وكانت الحاجة الماسة إلى مثل هذا النوع من البيانات الذي كان سببًا وراء تطوير مقياس تقييم مهارات القراءة في الصفوف الأولى ((http://www.earlygradereadingbarometer.org/users/login),)،

كأداة تفاعلية طُورت بتمويل من الوكالة الأمريكية للتنمية الدولية USAID، حيث استخدمت مجموعات بيانات الفعلية من عشرات الدول؛ لتوليد رسوم بيانية تعرض أداء التلاميذ في القراءة حسب البلد، وهي متاحة للجمهور (الدخول مجانًا مطلوب لتعرف بيانات التلاميذ من موقع الوكالة الأمريكية للتنمية الدولية).

ولكي تكون عملية المقارنة بين اللغات منطقية، فإن على التربويين وصناع السياسات اتخاذ خطوتين مهمتين:

أولاً-لضمان الكفاية الفنية لأداة عبر اللغات على وجه التحديد، فلا بد من تطويع وتطوير الأداة، بدلاً من ترجمتها، ولتوضيح الاختلافات في العناصر الثقافية أو اللغوية (كما هو موضح في القسم 6.5.1 أعلاه).

ثانيًا-في حالة الرغبة في التحليل بين اللغات، فإن القائمين على تطويع وتحليل نتائج – كحد أدنى-إجراء فحص دقيق لما يلي:

الكفاية الفنية للتقييم بالعلاقة مع الهدف منه.

خصاص اللغة على سبيل المثال: العمق اللغوي أو تعقيد نظام الكتابة.

لكل مهمة فرعية، وذلك لفهم البناء العام، والخاص الذي تسعى المهمة لرصده

ولمزيد من التوصيات والارشادات حول كيفية تطويع ومقارنة نتائج بين اللغات، يُرجى الاطلاع على **الملحق F**.

6.6 استخدام الأدوات في اللغات عبر تطبيقات متعددة: ابتكار صور مكافئة من الاختبار

كما سبق أن ذُكر في هذا القسم، فإن التطويع يتطلب تعديلاً لأداة موجودة بالفعل سبق وضعها للغة معينة. وإذا لم يكن هناك قلق من جهة تسريب الاختبار (أي إذا كان لدى المعلمين فرصة محدودة للحصول على أدوات وإذا كان من غير المحتمل أن الطلاب سيألفون صورة معينة من صور التقييم)، فيمكن ببساطة استخدام نفس الأداة على فترات زمنية متعددة. ولكن، إذا كان الخوف من تسرب الأداة قائمًا، فسيكون من الضروري استخدام تقييمات متعددة (صور متكافئة من الاختبار)؛ لقياس التغيرات في الأداء. ولضمان المقارنات الصادقة للنتائج أن تُطبق عبر صور للتقييم / التطبيقات، إذن يجب أن تُعدل الأدوات بالطريقة التي ينتج عنها صور جديدة معادلة في الصعوبة قدر الإمكان للنموذج الأصلي. والصور المتكافئة للاختبار تشير إلى الاختبارات التي يقصد بها أن تكون متساوية في الصعوبة، (وبالتالي يمكن استبدالها مباشرة).

في الحالات التي تُحدد فيها صعوبة المهمة الفرعية من أداة (أ) والأداة (ب) في الاختبار البعدي. اتضح أنها غير مساوية أي لا يجب أن تكون على قدم المساواة، فينبغي تطبيق بعض إجراءات المساواة لشرح الفروق (انظر القسم 10.5). وتشير نماذج الاختبار المعادلة إلى النماذج التي عُدلت من قبل عملية إحصائية تتبع تطبيق الاختبار؛ لجعل الدرجات قابلة للمقارنة. ومع ذلك، توصي أفضل الممارسات بتعديل المهمة الفرعية بالحد من الحاجة إلى المساواة الإحصائية ما بعد التطبيق. ويوجد وصف لأساليب إعداد الصور المتكافئة قسم تطويع الأداة من حقيبة الأدوات (القسم 6)، ويمكن أن تشتمل على:

- إجراء تغييرات بسيطة في أسماء موضوعات القصة، والأحداث، والصفات، واستبدالها بما يعادلها على مستوى الصف.
- بالنسبة للمهام الفرعية التي تقدم للمتعلمين في كراسة الأسئلة للطالب، و التي على شكل شبكة، تُخلط البنود ضمن صفوف الشبكة، بحيث تكون خبرة المتعلم مع هذا التقييم تكون هي نفسها خبرة الاختبار السابق بغض النظر عن المدة التي شارك التلميذ خلالها في الاختبار من قبل.

[16] الأداة الكافية فنيًّا هي الأداة التي أظهرت انها تنتج نتائج ثابتة ، و تسمح بتوليد التحليلات الصادقة و بالتالي تتمتع بالثقة.

بالرغم من أن هذه الاساليب يُقصد بها الحد من الحاجة إلى المساواة، فإنها لا تضمن وجود صور متكافئة ولا تلغي الحاجة الي فحص التكافؤ بعد التجريب. وبالنسبة للحالات التي تستخدم هذه الأساليب، ولا ينتج عنها صورًا غير متكافئة للاختبار، فقد يتطلب هذا بعض الإجراءات الإحصائية. يناقش القسم 10.5 المنهجيات والتوصيات المحددة لمساواة الدرجات بعد معالجة البيانات وتحليلها.

7.6 أفضل الممارسات

ونظرًا لأن تقييم مهارات القراءة في الصفوف الأولى امتد إلى عشرات البلدان، وإلى لغات كثيرة، فهناك الكثير من الدروس المتعلمة، والتي تستحق ا أن تؤخذ في الاعتبار عند تخطيط، وتنفيذ كل من تطويع الأداة، وتعديل التطويع.

- **التعليمات.** مناقشة بروتوكول، أو التعليمات التي يتعين على المقيمين اتباعها أمر غير مثمر. وقد وضعت التعليمات بعناية بناءً على أدلة من الأبحاث والخبرات السابقة، ولا يجب تعديلها أبدًا. وبدلاً من ذلك فإن استخدام الوقت في الترجمة الدقيقة للتعليمات أمر بالغ الأهمية للتنفيذ الناجح.

- **الاختبار القبلي والتجريب الاستطلاعي.** كل من هذه الخطوات أجزاء مهمة من العملية (انظر الجزء الأول من القسم 6، وكذلك القسم 9 من الحقيبة)، ويجب أن يُخطط له، وتُقدر الموازنة.

- **الاختبار القبلي والتجريب الاستطلاعي.** كل من هذه الخطوات أجزاء مهمة من العملية (انظر الجزء الأول من القسم 6، وكذلك القسم 9 من الحقيبة)، ويجب أن يُخطط له، وتُقدر الموازنة.

- **استخدام نفس بنود المهام الفرعية المتطابقة تقريبًا عبر صور متعددة من الأداة.** وتشير أفضل الممارسات إلى الحد من الحاجة إلى إجراءات المساواة الإحصائية التي تتم بعد التطبيق فكلما كان ذلك ممكنًا. يمكن أن تنتج الإجراءات القوية لتصميم الأداة صورًا قابلة للمقارنة بشكل كبير ، و التي يمكنها أن تخفف من الحاجة لإجراء المساواة.

7 استخدام الجمع الإلكتروني للبيانات

اعتبارا من 2010، بدأ باحثو تقييم القراءة للمراحل الأولى في الانتقال من الجمع الورقي إلى الجمع الإلكتروني للبيانات. يقلل الجمع الإلكتروني للبيانات من فرص حدوث أخطاء أو الحذف في البيانات كما يتيح الوصول للنتائج بسرعة أكبر.

الجمع الإلكتروني للبيانات يحسن ويدعم العمل الميداني.

أظهرت مقارنات الجمع الإلكتروني للبيانات بنظيره الورقي ميزات من حيث الفعالية والكفاءة. من خلال التوافر المتزايد للأجهزة المتنقلة غير المكلفة والاتصال عبر الإنترنت تمكن الباحثين من تحليل البيانات بشكل وقتي مما يسهم باستمرار في توفير الدعم للاقتناص الإلكتروني للبيانات (Walther et al.، 2011).

يتمثل أحد الفروق الرئيسية بين الجمعين الإلكتروني والورقي للبيانات في التخلص من الإدخال اليدوي للبيانات للاستمارات الورقية التي تم إكمالها إلى إحدى قواعد البيانات الإلكترونية، مما يقلل الوقت المطلوب وكذلك الأخطاء التي عادة ما تصاحب الإدخال اليدوي للبيانات من الورق، بالإضافة إلى الأخطاء الناجمة عن قيام المقيمين بصورة خاطئة أو غير دقيقة بوضع علامات على الاستمارات الورقية أو إغفال بعض الأسئلة. علاوة على ذلك، فإنه يمكن تحميل نتائج الجمع الإلكتروني للبيانات من الميدان، ما يمكن من العمل عليها وتحليلها في وقت أقصر. كما توفر هذه الخاصية فرصة للوقوف على الأخطاء وتداركها بينما لا يزال المقيمون في الميدان، ومن ثم فإن الجمع الإلكتروني للبيانات يحسن العمل الميداني ويدعمه.

من المهم العلم بأن الجمع الإلكتروني للبيانات لا يغير إجراءات التنفيذ الأساسية للتقييم. يظل الطفل يقرأ من استمارة ورقية مطبوع عليها حروف وكلمات، وتظل التوجيهات التي يقدمها الميسر كما هي. لم تتغير التعليمات للجمع الإلكتروني للبيانات فيما عدا الإشارة لكيفية وضع علامة الإجابات (مثال "ضع علامة" في مقابل "المس الشاشة").

كانت الأمثلة الأولى المعروفة لجمع البيانات المتنقل غير الشبكي والمصممة خصيصا لتقييم القراءة للمراحل الأولى iProSurveyor، والتي أعدتها مؤسسة بروديجي للنظم Prodigy Systems للاستخدام باللغة العربية في اليمن وبعدها في المغرب في 2011. [17] ونظام البرمجة Tangerine® الذي أعدته مؤسسة آر تي آي انترناشيونال RTI International في 2010 والذي خضع للاختبار في 2012.

حولت هذه البرامج أداة تقييم القراءة للمراحل الأولى، شاملة المهام ذات التوقيت، إلى سطح جهاز لوحي متنقل متفرد وبديهي، لا يتداخل مع الإجراءات الإدارية الأساسية الثنائية لتقييم القراءة للمراحل الأولى.[18] شمل مشروع iProSurveyor لتقييم القراءة للمراحل الأولى في اليمن 38 مدرسة عبر ثلاث محافظات أجريت فيها مقابلات مع 735 طالباً في الصفين الثاني والثالث. تمت تجربة Tangerineميدانياً للمرة الأولى في يناير 2012 تحت مظلة القراءة والرياضيات الأساسية للوكالة الأمريكية للتنمية الدولية (PRIMR) في كينيا والتي تم في إطارها الوقوف على 176000 نقطة بيانات من خلال عينة صغيرة قوامها 200 طالب من 10 مدارس يتم تقييمها للقراءة للمراحل الأولى للغة الإنجليزية والسواحيلية وتقييمها للرياضيات للمراحل الأولى (EGMA؛ Strigel، 2012). كشفت هذه الاختبارات الميدانية عن سهولة الاستخدام والكفاءات المتحققة كما تم التأكيد على الجمع الإلكتروني للبيانات كمنهج عملي يتفوق على جمع البيانات الورقي لتقييمات القراءة الشفهية (والرياضيات) ذات المكونات المرتبطة بالوقت.

[17] بموجب عقد من الباطن لمؤسسة RTI International ضمن المرحلة الثانية من مشروع إد داتا التابع للوكالة الأمريكية للتنمية الدولية EdData II (انظر Collins & Messaoud-Galusi, 2012; Prodigy Systems, 2011).

[18] لم يتم اعتبار أجهزة الكمبيوتر المحمولة بديلاً تكنولوجياً قابلا للتطبيق لهذا الغرض بالنظر للآثار المحتملة لهذه الطريقة بسبب رؤية التكنولوجيا في الفصول وبسبب القيود على استخدامها في سياقات معينة (غياب مصادر الكهرباء والرطوبة/الغبار والتنقل على الأقدام أو بالدراجات والمراكب وغيرها). توجد أيضاً نظم لجمع وإدخال البيانات لأجهزة الحاسب الآلي المكتبية أو المحمولة؛ فمثلا طورت eEGRA واستخدمت من قبل مؤسسة مركز تطوير التعليم (EDC) (http://eegra.edc.org/).

1-7 تحذيرات وقصور في الجمع الإلكتروني للبيانات

بشأن الجمع الإلكتروني للبيانات، يجب الانتباه لأوجه القصور التالية:

- **مخاطرة الأخطاء.** لا يعد الجمع الإلكتروني للبيانات محصناً ضد الأخطاء. تظل هناك مساحة لأخطاء الإدخال أو فقد البيانات.

- **اعتبارات التكلفة.** تحليلات التكلفة التي قامت الوكالة الأمريكية للتنمية الدولية بإجرائها في إطار مشروع **EdData II** بينت أن الكفاءات المتحققة باستخدام الجمع الإلكتروني للبيانات بالمقارنة بالأدوات الورقية تحدث في الغالب عندما يتم استخدام الأجهزة لعمليات عدة لجمع البيانات. قد لا يتحقق توفير في التكلفة إذا ما تم استخدام الأجهزة لجمع البيانات لمرة واحدة فقط.

- **الحاجة لنسخ احتياطية ورقية.** قد تظل فرق التقييم بحاجة لحمل أدوات ورقية احتياطية في حال تعطل الأجهزة الإلكترونية خلال العمل الميداني. وبالتالي فإنه يتم التعامل مع الأدوات الورقية خلال تدريب المقيمين جنباً إلى جنب مع البرامج الإلكترونية.

- **التعرض المحدود للتكنولوجيا.** يجب أن يضع المخططون في الاعتبار السياق المحلي/الإقليمي ومدى تأقلم المُقيمين مع التكنولوجيا عند الاستقرار على استعمال الجمع الإلكتروني للبيانات.

- **مسائل الأمن.** الفقد، والسرقة، والتعرض للتلف بالنسبة للأجهزة يخلق مساحة للخسارة المالية أو الأذى الشخصي وبالتالي فإن ضمان أمن وسلامة الأجهزة والمقيمين يتطلب تخطيطاً حريصاً.

- **محدودية البينة التحتية للاتصالات.** العثور على أو إقامة نقاط متنقلة بعيدة لرفع البيانات الميدانية قد يكون صعباً في بعض الدول أو المناطق.

- **محدودية القدرة المحلية.** إن تطويع الأداة للغات والنصوص المحلية واستدعاء المضمون في برنامج جمع البيانات محل الاختيار يفرض تحديات ذات صلة بالعملية. ويعد التواصل مع الشركاء المحليين ذوي الخبرة أساسي للتعرف على والتعامل مع القيود على القدرات المتعلقة بجمع البيانات.

عند وقوع الاختيار على الجمع الإلكتروني للبيانات بدلاً من الجمع الورقي، فإنه يجب على الباحثين التعامل مع الحاجة للحفاظ على أمن البيانات الرقمية؛ وفقاً للبرنامج المستخدم لجمع البيانات، فإن الوصول للنتائج الأولية قد يكون متاحاً للعديد من الأفراد. وحتى نقاط نظام تحديد المواقع العالمي (GPS) فإنه يجب استخدامها فقط لأغراض التحقق، وليس لتحديد مدارس بعينها. وكما هو الحال مع الأبحاث المعتمدة على الجمع الورقي، يجب القيام بكل ما يلزم لضمان الحفاظ على الخصوصية وأنه لا يمكن تعرض أي مدارس أو مدرسين أو طلاب لتبعات سلبية بسبب النتائج.

2-7 برامج جمع البيانات

توجد العديد من أدوات البحث المتنقل التي يمكن تطويعها لإدارة تقييم القراءة للمراحل الأولى. يعد برنامج Tangerine المجاني هو أحد الأدوات المستخدمة على نطاق واسع، وتم تنفيذه في أكثر من 60 مشروعاً في 36 دولة من قبل 27 منظمة حتى منتصف 2015 (انظر www.tangerinecentral.org). توجد مقارنة لمميزات Tangerine وعدد آخر من أدوات جمع البيانات الإلكترونية – Magpi، SurveyToGo، doForms، Droid Survey، Open Data Kit (ODK) وCommand Mobile - في ملحق ز، كما توجد عينة من الإرشادات الإلكترونية في مقابل الورقية في ملحق ح. وحتى كتابة هذه السطور، تعد iProSurveyor (للآيباد)، Tangerine وSurveyToGo المنصات الوحيدة التي لا توفر نظم إدخال عبر أجهزة الحاسب المكتبية أو المحمولة (انظر هامش رقم 18 في الصفحة السابقة) وتم تطويعها لتقييم القراءة للمراحل الأولى. يأخذ المنفذين بعين الاعتبار أي البرامج الأكثر ملائمة للسياق وطبيعة البيانات التي يتم تجميعها – لا سيما صيغة المصفوفة المميزة محددة الوقت للعديد من المهام الفرعية لتقييم القراءة للمراحل الأولى والحاجة لحساب العدد الكلي للعناصر التي تمت محاولتها (الدقة)، والعناصر الصحيحة في الدقيقة (الطلاقة). أين سيتم تخزين البيانات ومن سيقوم على إدارتها والقدرة الفنية يمكن أن توضع أيضاً في الاعتبار عند اختيار برنامج بعينه.

3-7 اعتبارات خاصة بشراء الأجهزة واختيارها

عند شراء الأجهزة للوفاء بمتطلبات جمع البيانات الإلكتروني لتقييم القراءة للمراحل الأولى، يجب على المنفذين اعتبار عوامل كالشحن والتخزين وإعادة استخدام العناصر. حتى عام 2015، كانت أجهزة الحاسب اللوحية (وليست المحمولة أو الهواتف النقالة أو الذكية) تعد

أنسب أنواع الأجهزة بسبب حجم الشاشة وسهولة التعامل وخفة الوزن وبالأخص طول عمر البطارية. وكحد أدنى، فإن الملحقات الإضافية يجب أن تشمل قلماً إلكترونياً وحافظة واقية وراوتر لاسلكي للجمع الفعال للبيانات والقدرة على إرسال النتائج يومياً.

يجب أن يوازن المنفذين بين مميزات وعيوب شراء أجهزة في الدولة التي سيتم فيها جمع البيانات في مقابل الشراء من خارجها. سيتطلب الشراء من الخارج توفير وقت إضافي كافي مقدماً تحسباً للشحن وإنهاء إجراءات الجمارك. قد يكون حمل الأجهزة باليد ممكناً من دولة لأخرى في حالة استخدام (أو إعادة استخدام) أعداد صغيرة من الأجهزة اللوحية وملحقاتها، إلا أنه يجب على الأفراد المعنيين التعرف على اللوائح الجمركية والرسوم المستحقة على استيراد الأجهزة، وفق السياق المحلي. على سبيل المثال، تطلب بعض الدول إثبات وجود خطط لإعادة تصدير الأجهزة بعد جمع البيانات قبل أن تعفيها من رسوم الاستيراد.

كما يجب أن يخطط المنفذون للتخزين الملائم لكافة الأجهزة والملحقات قبل جمع البيانات وبعده وخلال التدريب. يجب إيداع كافة الأجهزة والملحقات في موقع يمكن تأمينه لتجنب السرقة. كما يجب أن تكون منطقة التخزين محمية من الغبار والرطوبة والحرارة الشديدة. يجب ملاحظة أن عمر بطارية هذه الأجهزة قد يتأثر بعد فترات طويلة من عدم الاستخدام.

من الضروري، كجزء من عملية التنفيذ، وضع إجراءات واضحة لملكية الأجهزة والبرامج والبيانات والوصول إليها واستخدامها. من المعتاد (والموفر للنفقات كذلك) أن تتم إعادة استخدام الأجهزة من قبل المنفذ أو المنظمة الممولة أو أن يتم نقل ملكية العناصر التي تم شراؤها إلى المنظمات المحلية للاستمرار في استخدامها.

4-7 المستلزمات المطلوبة للجمع الإلكتروني للبيانات والتدريب

- أجهزة حاسب لوحية وشاحن لكل منها
- برنامج يحتوي نسخة إلكترونية من التقييم
- حوافظ للأجهزة اللوحية
- أقلام إلكترونية
- حقائب للمقيمين لحمل الأجهزة خلال العمل الميداني
- أجهزة راوتر متنقلة ووصلات اتصال مع اشتراك للبيانات
- عدد من الأجهزة الاحتياطية تحسباً للتلف أو الفقد

8 تدريب المقيم على تقييم القراءة للمراحل الأولى

يقدم هذا القسم توجيهات بخصوص إعداد وتنفيذ تدريب للمقيمين على تقييم القراءة للمراحل الأولى.

لاحظ أن هذا القسم لا يهدف لأن يعمل كدليل للمقيم أو المشرف وإنما كمرجع لمنظمي التدريب. تحتوي *ملاحظات توجيهية لتخطيط وتنفيذ تقييمات القراءة للمراحل الأولى* على تفاصيل إضافية لتدريب المقيمين وينصح بأن تصاحب هذه المادة (RTI International & International Rescue Committee، 2011). سيكون المقيمين الذين سيختبرون هذه الأداة بحاجة لخمسة أيام تدريبية تقريبا.[20] وسيتوقف طول التدريب على عوامل من قبيل عدد الأدوات التي سيتم إدارتها (مثلاً تقييم الرياضيات بجانب تقييم القراءة)، عدد المدربين المتاحين، عدد المتدربين، الخبرة السابقة للمتدربين، والميزانية والوقت المتاحين. على سبيل المثال، إذا ما كان عدد من المتدربين محدودي الطلاقة في لغة التدريب (لدرجة تتطلب وجود مترجم)، فإنه قد يكون من الحكمة إضافة يومين أو ثلاثة للجدول.

بالنسبة للتدريب على تقييم القراءة للمراحل الأولى الذي عقد في تنزانيا في 2013 وحضره 150 مقيما، كانت الأداة بلغتين مع عدد من المسوح الإضافية، وضم الفريق الفني خمسة أفراد: متخصص في اللغة 1، متخصص في اللغة 2، خبير في برامج جمع البيانات، مدير لوجيستي، ومنسق عام والذي ركز كذلك على اختبارات أداء المقيمن، التجهيز لورش العمل وتصميم المسوح والعلاقات مع المانحين.

لضمان فهم كافة المتدربين للهدف من العمل ودعمه، فإن عنصر أساسي على الأجندة كان مراجعة المبادئ والهدف وراء كل من مكونات أداة تقييم القراءة للمراحل الأولى. وتتمثل أهدافنا الرئيسية في:

[20] انظر قسم 9.1.3 حول مزايا وعيوب التوقيتات المحتملة المختلفة لعقد تدريب المقيمين بالنظر للجمع الاختباري والكلي للبيانات.

- تدريب قطاع من المقيمين ليقوموا بدقة وكفاءة على إدارة تقييم القراءة للمراحل الأولى، في نسختيه الورقية والإلكترونية.
- التعرف على ذوي المهارة للعمل كمقيمين لجمع البيانات.
- التعرف على أفراد مختارين للعمل كمشرفين خلال جمع البيانات وتدريبهم.

1-8 اختيار المشاركين في التدريب

من الضروري اختيار وتدريب عدد أكبر من المقيمين بـ 10-20% من العدد المبين في خطة العينة، وبلا شك، لن يفي بعض المشاركين بمعايير الاختيار وقد لا يستطيع آخرين إكمال التدريب لأسباب شخصية أو غيرها.

يمكن أن تتشكل فرق جمع البيانات من مسئولي التعليم و/أو مقيمين مستقلين تم إختيارهم لهذه العملية بعينها. وتتحدد المتطلبات والتفضيلات خلال مرحلة الاختيار، قبل التدريب، اعتمادا على الملابسات والأهداف المحددة.

يمكن النظر للمسئولين الحكوميين كمرشحين لتولي أدوار المقيمين أو المشرفين، غير أنه وحتى يتم اختيارهم للعمل الميداني، يجب أن يكونوا على نفس القدر من الوفاء بمعايير الأداء كغيرهم من المتدربين. يجب أن يؤكد الميسرين على معايير الاختيار في بداية التدريب. أحد الفوائد المتوقعة من إدماج المسئولين الحكوميين المؤهلين هو تحسين فرص الاستقبال اجيد لتحليل البيانات من قبل الحكومة عقب إعلان النتائج.

عنصر آخر يمكن اعتباره في مرحلة الاختيار هو مدى احتمالية وجود تعارض في المصالح لدى المرشحين – سواء في القطاع الخاص أو العام - قد ينتج عن السياق السياسي الحالي في البلاد.

تتمثل المعايير الهامة التي ينبغي للمخططين أخذها في الاعتبار عند تحديد من سيحضرون تدريب المقيمين في:

- قدرة المرشحين على القراءة والتحدث بطلاقة اللغات المطلوبة للتدريب وإدارة تقييم القراءة للمراحل الأولى؛
- الخبرات السابقة في العمل على إدارة تقييمات أو جمع البيانات؛

- الخبرة في العمل مع طلاب المراحل الأولى؛
- التوافر طيلة مرحلة جمع البيانات والقدرة على العمل في المناطق المستهدفة؛
- الخبرة والمهارة في استخدام الحاسب الآلي أو الأجهزة الإلكترونية المحمولة (الأجهزة اللوحية أو الهواتف الذكية).

سيقوم فريق التدريب باختيار القائمة النهائية للمقيمين اعتماداً على المعايير التالية. يجب أن يتم توصيل هذه المتطلبات للمتدربين منذ البداية كي يدركوا أن الاختيارات النهائية تعتمد على من هو الأكثر ملائمة للمهمة.

- **القدرة على إدارة تقييم القراءة للمراحل الأولى بدقة وكفاءة.** كل من يتم اختيارهم للعمل كمقيمين يجب أن يكشفوا عن درجة عالية من المهارة في إدارة التقييم، وهو ما يتضمن الإلمام بقواعد وإجراءات الإدارة، والقدرة على تسجيل ردود الطلاب بدقة، والقدرة على استخدام كل العناصر المطلوبة – مثل الجهاز اللوحي – لإدارة التقييم. يجب أن يكون المقيمين قادرين على إدارة عدة مهام في آن واحد، شاملة الاستماع للطالب وتسجيل النتائج وتشغيل الجهاز اللوحي.

- **القدرة على إقامة صلات جيدة مع الأطفال.** من المهم أن يتمكن المقيمين من التواصل بطريقة خالية من التهديد مع الطلاب الصغار. إقامة صلات إيجابية دافئة مع الطلاب يساعدهم على الأداء بأفضل قدر ممكن. وعلى الرغم من أن هذا العنصر من إدارة الاختبار يمكن تعلمه، لن يكون بمستطاع كافة المقيمين أن ينفذوه بمهارة.

- **القدرة على العمل كفريق في بيئة مدرسية.** لا يعمل المقيمين بمفردهم وإنما كجزء من فريق. ومن ثم فإنهم بحاجة للكشف عن قدرتهم على العمل بشكل جيد مع الآخرين لتحقيق كافة المهام خلال زيارة المدرسة. علاوة على ذلك، فإنهم بحاجة لاستعراض قدرتهم على العمل بشكل جيد في بيئة مدرسية والتي تتطلب اتباع بروتوكولات معينة واحترام المدرسة والعاملين بها والتعامل بشكل لائق مع الطلاب.

- **التوافر والمرونة.** كما هو مبين بعاليه، يجب أن يكون المقيمين متاحين طيلة فترة جمع البيانات وأن يستعرضوا قدرتهم على العمل في المواقع الميدانية المعينة. على سبيل المثال، فقد يكونون بحاجة لقضاء أسبوع في بيئة ريفية حيث قد تصعب الانتقالات وتتواضع اختيارات الإقامة.

من بين المتدربين، سيحدد الميسرين كذلك المشرفين لدعم والتنسيق مع المقيمين خلال جمع البيانات. يجب أن يفي المشرفين (والذين قد يعرفون كذلك بمنسقي جمع البيانات، أو مسمى مشابه)، أو يتجاوزوا، المعايير الموضوعة للمقيمين. علاوة على ذلك، فإنهم يجب أن:

- يتمتعوا بمهارات القيادة، لديهم خبرة في القيادة الفعالة لفريق، ويتمتعون باحترام زملائهم؛
- يكونوا منظمين ويهتمون بالتفاصيل؛
- على إلمام بإجراءات إدارة تقييم القراءة للمراحل الأولى بصورة تمكنهم من الإشراف على الآخرين والوقوف على الأخطاء في جمع البيانات؛
- لديهم معرفة/مهارات كافية بالأجهزة اللوحية لمساعدة الآخرين؛
- التعامل بطريقة لائقة مع مسئولي المدارس والطلاب.

يجب أن يوضح الميسرين كذلك هذه المواصفات مقدما للمتدربين وأي شركاء محليين لجمع البيانات. لن يكون المشرفين بالضرورة من ذوي المناصب الحكومية العليا أو ممن يتمتعون بدرجة ما من الأقدمية. قد يتمكن المسئولين ممن لا يوفون بالمعايير من القيام بمهام إشرافية أخرى كزيارات ميدانية غير معلنة. تبرز هذه المواقف عادة عندما يرغب مسئولي التعليم في القيام بدور ما في مراقبة والإشراف على جمع البيانات، وسواء تمكنوا من حضور تدريب المقيمين أو لا، فإن فوائد إدماجهم قد تكون فهما أكبر لعملية تقييم القراءة للمراحل الأولى وتقبل نتائجها.

2-8 الإعداد للتدريب:

يجب القيام بعدد من المهام الرئيسية قبل عقد التدريب وتشمل:

- **إعداد أداة تقييم القراءة للمراحل الأولى والمواد التدريبية.** الإنتهاء من محتوى الأدوات التي سيتم استخدامها خلال التدريب – سواء الإلكترونية أو الورقية، لكل اللغات. كما يجب تجهيز ونسخ مواد التوزيع الأخرى (كالجدول والنسخ الورقية من الاستبيان واستمارات التحفيز، ودليل المشرف).
- **شراء الأجهزة.** المواد والأجهزة التي يخطط لها ويشتريها المنظمون مقدما تتراوح ما بين الأجهزة اللوحية وحافظاتها لألواح الرسم وساعات التوقيت وشواحن الطاقة إلى هدايا

الطلاب. صمموا قائمة جرد لمتابعة حركة كافة المواد خلال التدريب وجمع البيانات لتقييم القراءة للمراحل الأولى.

- **إعداد الأجهزة.** للقائمين على جوانب الدعم التكنولوجي للتدريب، فإنه يجب إعداد الأجهزة لجمع البيانات بمجرد شرائها وهو ما يعني تحميل البرنامج والنسخ الإلكترونية من الأدوات على الأجهزة اللوحية وإعدادهم بشكل ملائم.

- **إعداد جدول ورشة العمل.** جهز نسخة أولية من الجدول وتشاركها مع الفريق المنفذ للورشة. للتدريب الموجه فقط لتقييم القراءة في المراحل الأولى، فإن مساحات المحتوى الرئيسية في الجدول تتضمن:

  - نبذة عامة عن أداة تقييم القراءة في المراحل الأولى (الغرض والمهارات التي يتم قياسها)
  - إدارة المهام الفرعية لتقييم القراءة في المراحل الأولى (البروتوكولات والعمليات، الممارسات المتكررة)
  - استخدام الجهاز اللوحي (التشغيل وتخزين ورفع التقييمات)
  - بروتوكولات سحب العينات والعمل الميداني.

  انظر **ملحق 1** لنموذج للجدول.

- **الاستقرار على فريق الميسرين.** تدريبات المقيم يتم تيسيرها من قبل مدربين اثنين على الأقل من ذوي المعرفة بتقييمات القراءة (وتقييم القراءة في المراحل الأولى على وجه الخصوص) وذوي خبرة في تدريب جامعي البيانات. لا يحتاج المدربين للتحدث باللغة الخاضعة للتقييم بالضرورة في حالة توافر خبير لغة محلي يستطيع التحقق من النطق الصحيح للحروف والكلمات، ويساعد في أي ترجمة يمكن أن تقوم إليها الحاجة خلال تيسير التدريب. غير أن المدربين يجب أن يتحدثوا اللغة التي ستعقد بها ورشة العمل بطلاقة. وفي حالة عقد التدريب بلغات مختلفة، فإنه يفضل توافر فريق من المدربين المتمرسين مع اعتبار عدد من المدربين الإضافيين.

3-8 مكونات تدريب المقيم

كما هو مبين في نموذج الجدول في ملحق 1، فإن تدريب المقيمين يتضمن عدة مكونات ثابتة. وبترتيب مماثل لما يلي، فإنه على الميسرين:

- دعوة المسئولين رفيعي المستوى بغرض التعبير عن إلتزامهم المعلن نحو تقييم القراءة للمراحل الأولى واهتمامهم بالنتائج.
- تقديم مشروع التقييم، أهمية القراءة في المراحل الأولى وماهية تقييم القراءة للمراحل الأولى وأساسيات إدارة الأداة.
- بيان الأهمية للبحث في مراقبة أداء المقيمين والمعايير التي سيتم إختيارهم وتقييمهم على أساسها.
- تقديم نبذة عن المهام الفرعية وبيان كيفية إدارتها.
- عرض وشرح أي أدوات إضافية ستتم إدارتها بجانب تقييم القراءة للمراحل الأولى.
- منح المشاركين الفرصة للمارسة في مجموعات من اثنين أو أكثر مع إشراف وتشجيع من جانب كبير المدربين. بعد عدة أيام من التدريب، رتب ليوم واحد على الأقل من التدرب مع الأطفال في سياق مدرسي.
- راقب، ساعد وأعد التدريب حسب الحاجة. تأكد من تكيف المتدربين مع كل من محتوى الاستبيان والأجهزة والبرنامج.
- قم بتقييم دقة المقيمين بطريقة رسمية (راجع قسم 8.7)، استخدم النتائج لمعالجة المشكلات قبل القيم بالاختيار النهائي لفريق المقيمين للجمع الرئيسي للبيانات.

4-8 طرق وأنشطة التدريب

البحوث في تعلم البالغين تشير إلى بعض أفضل الممارسات التي ينبغي استخدامها في تدريب المقيمين. وبعيدا عما إذا كان التدريب يتم لفريق من 20 أو 100 مقيم، فإن تصميم *جلسات تفاعلية* يعمل فيها المشاركون سويا ومع التكنولوجيا والأداة ستؤدي إلى تعلم أكثر فعالية.

والخبرات الدولية في تدريب المقيمين على تقييم القراءة في المراحل الأولى تشير إلى أنه مع زيادة عدد الفرص التي يتاح فيها للمشاركين ممارسة إدارة تقييم القراءة للمراحل الأولى، فإن تعلمهم سيكون أفضل في الإدارة الفعالة للأداة. علاوة على ذلك، فإن الأنشطة المتنوعة من يوم لآخر تتيح للمشاركين الفرصة للارتباط المعمق والمخرجات الأفضل. على سبيل المثال، فإن الأنشطة اليومية للتدريب على الجهاز اللوحي يمكن أن تشمل:

- عروض الميسرين
- الفيديوهات
- تدريبات المجموعة الكاملة
- تدريبات المجموعات الصغيرة
- التدريبات الثنائية
- عروض المتدربين

وخلال التدريب، يجب أن يغير الميسرين من تشكيل المجموعات الصغيرة والثنائية، وهو ما يمكن أن يشمل الجمع بين مقيم أكثر مهارة أو خبرة مع شخص أقل خبرة.

بعض الأفكار تتضمن منهج "الحبل المستدير" للتدرب على العناصر الأكثر حاجة للمراجعة (مثلا يجلس المشاركون في دائرة ويتبادلون الأدوار في النطق بسرعة بأصوات الحروف في أداة تقييم القراءة للمراحل الأولى) أو المحاكاة والتي يلعب فيها أحد الأشخاص دور مقيم يرتكب أخطاء أو لا يتبع الإجراءات السليمة وبعدها يتم سؤال المشاركين لمناقشة ما حدث وما الذي كان يجب على "المقيم" القيام به بدلا عن ذلك.

في حالة العمل على أكثر من لغة، ينصح بعقد هذه الأنشطة في نطاق مجموعات اللغات.

سيحتاج الميسرين كذلك لتوجيه المتدربين كذلك لقضاء بعض الوقت في التدرب على استعمال الجهاز اللوحي: القوائم المنسدلة، الخواص المميزة للمدخلات وغيرها.

عرض فيديوهات لتقييم القراءة للمراحل الأولى الذي تتم إدارته للمشاركين في ورشة العمل يمكن أن يساعدهم على فهم العملية والبروتوكولات قبل توافر الفرصو لهم لإدارتها بأنفسهم. هذه الفيديوهات – والتي ستتطلب تصريحات معينة وستحتاج لتسجيلها قبل التدريب – يمكن

أن تستخدم لاستعراض أفضل الممارسات والسيناريوهات الأكثر شيوعا في التطبيق. وبالتالي يمكن أن تستخدم كمنطلق مفيد للنقاشات والممارسات.

5-8 الزيارات المدرسية

يتضمن تدريب المقيمين دائما زيارة مدرسية واحدة على الأقل لإتاحة الفرصة للمقيمين لتجربة إدارة تقييم القراءة للمراحل الأولى للأطفال واستخدام التكنولوجيا في ظروف مشابهة لتلك التي سيواجهونها خلال الجمع الفعلي للبيانات. كما ستسمح هذه الزيارات لهم كذلك بممارسة إجراءات سحب عينة التلاميذ واستكمال التوثيق المطلوب بشأن الزيارة المدرسية.

للمساعدة على ضمان تحقق الفائدة من الزيارة المدرسية، فإنه على فريق قيادة التدريب:

- تحديد موعد لزيارة مدرسية واحدة على الأقل خلال التدريب (يفضل أن تكون اثنتين أو أكثر):
  - خطط لواحدة خلال منتصف التدريب والثانية قرب النهاية.
- حدد عدد المدارس المطلوبة:
  - حدد عدد المدارس بناء على عدد المتدربين وحجم المدارس القريبة وعدد الزيارات.
  - تجنب مداهمة المدارس بإحضار عدد كبير من الأفراد لمدرسة واحدة. عين ما لا يزيد عن 35-40 فرد لمدرسة كبيرة وعدد أقل من ذلك للمدارس الأصغر.
- حدد المدارس قبل بداية التدريب:
  - احصل على التصاريح المطلوبة وقم بتنبيه مسئولي المدرسة وجهز وسائل الانتقال وتحقق من أن المدارس ليست جزءا من عينة الجمع الكلي للبيانات (وإذا تعذر ذلك قم باستثناء مدارس التدريب من العينة النهائية).
- **جهز الفرق في اليوم السابق على الزيارة حتى يكونوا على علم بما سيجري:**
  - من حيث ترتيبات الانتقالات وإلى أي مكان يذهب كل واحد ومشرفي الفرق وعدد الطلاب لكل مقيم وعدد التقييمات التي سيتم إجرائها وما إلى ذلك.

- خلال الزيارة الثانية أو الثالثة، يمكن أن يصبح المشاركين أكثر راحة في العمل باستقلالية وسيستفيدون من تجربة الإدارة مع أكبر عدد ممكن من الأطفال خلال كل زيارة. كما أنه سيكون بمقدورهم كذلك ممارسة إجراءات سحب عينة التلاميذ وغيرها من عناصر جمع البيانات والتي يمكن ألا يكون قد سبق لهم تعلمها قبل الزيارة المدرسية الأولى.

- سيقوم كل مقيم بإدارة الأداة (الأدوات) مع ما بين 4-10 طلاب [21] لكل واحدة خلال كل زيارة مدرسية.

- من الأهمية بمكان عقب الزيارة عقد جلسة ختامية مع المشاركين، حيث تعطيهم الفرصة للتشارك كع المجموعة بخصوص ما يرون أنه تم أداؤه بشكل جيد وكذلك ما الذي شعروا بتحدي تجاهه. وعادة ما تثير الزيارة المدرسية قضايا جديدة وتمنح الفرصة للإجابة على أسئلة قد تكون تمت إثارتها خلال التدريب.

[21] عدد الطلاب الذي يمكن لكل مجمع بيانات أن يقيمه في المدرسة يتوقف إلى حد كبير على عدد المهام الفرعية لكل أداة والعدد الإجمالي للأدوات التي تتم إدارتها مع كل طالب (مثل اللغات أو المقررات المتعددة).

ملخص واجبات المدربين خلال الزيارات التدريبية للمدارس

- تحديد المشرفين من بين المتدربين
- مساعدة الفريق في البدايات حسب الحاجة
- ملاحظة المقيمين وتقديم المساعدة حسب الحاجة.
- مع الحصول على الإذن الملائم: التقاط صور أو تسجيل فيديوهات للمقيمين، للمزيد من التدريب والنقاش خلال الجلسة الختامية.
- إعاد ةالفصل/الموارد للشكل االذي كانت عليه قبل وصول الفرق.
- شكر ناظر المدرسة على الوقت والمشاركة.

ستطون هناك حاجة لمنطقة هادئة ومنفصلة داخل المدرسة ليتسنى للمشاركين التدرب على إدارة التقييمات. وبشكل مثالي، فإنه ينبغي أن يجلس المقيمين في مواجهة مقاعد الطفل عبر المكتب ويقوموا بإدارة الأداة. في حالة عدم توافر مكاتب، فإنه يمكن للأطفال الجلوس على مقعد موضوع بشكل قطري قليلا من المقيم.

خلال الزيارة المدرسية الأولى، من المفيد للمشاركين إجراء تقييم القراءة في مجموعات من اثنين حتى يتمكنوا من ملاحظة وتقديم تقييمات لبعضهما. العمل الثنائي يمكن أن يكون مفيدا أيضا بالنظر لأن المشاركين عادة ما يشعرون بالتوتر خلال أول مرة يقومون فيها بإجراء تقييم القراءة للمراحل الأولى مع طفل.

6-8 عملية تقييم المقيم-المتدرب

عملية التقييم الشفافة والمعايير الواضحة للتقييم تساعد كلا من الميسرين والمتدربين. تتضمن العملية المستخدمة لتقييم المقيمين خلال التدريب وسائل تقييم رسمية وغير رسمية. وكجزء من التقييم غير الرسمي، فإن الميسرين يلاحظون المتدربين عن كثب خلال ورشة العمل والزيارات المدرسية كما تشمل إجراء مقبلات فردية معها حسب الإمكان.

سيتوقع المتدربين تقييما لنقاط القوة والضعف لديهم خلال الورشة. توافر فريق مناسب ومؤهل من المدربين شيضمن انتظام ونوعية هذا التقييم. وبالمثل، فتوافر عدد كافي من المدربين سيسمح للتقييم بالتعامل مع حاجات المتدربين للمساعدة الخاصة وللاختيار المدقق للمشرفين.

الملاحظة الدقيقة للمقيمين تساعد على جمع بيانات عالية الجودة – وهو الهدف الرئيسي. وبالتالي فإنه أثناء تدرب المقيمين، يقوم الميسرين بالتنقل فيما بينهم للمراقبة وتدوين الملاحظات بخصوص أي مسألة يجب التعامل معها مع المجموعة الكلية.

يحتمل تقييم المقيمين وجوها عديدة ويأخذ في الاعتبار عدة عوامل، من بينها القدرة على:

- إدارة الأدوات بشكل صحيح وكفء، شاملة ومعرفة واتباع كل قواعد الإدارة.
- تسجيل كافة البيانات الديموغرافية والإجابات بدقة
- التعرف على الإجابات الصحيحة والخاطئة.
- استخدام الأجهزة بشكل صحيح وكفء، لا سيما الجهاز اللوحي.
- العمل بشكل جيد كجزء من فريق.
- الالتزام ببروتوكولات زيارة المدارس.
- خلق صلات مع التلاميذ ومسئولي المدارس.

خلال التدريب، فإن المشاركين أنفسهم سيتوقفون أمام ويتشاركون خبراتهم في استخدام الأداة. يجب أن يكون القائمين على التدريب على استعداد لتوضيح بروتوكول تقييم القراءة للمراحل الأولى (مثل التعليمات المتضمنة) اعتمادا على تجارب المقيمين سواء خلال ورشة العمل وخلال الزيارات المدرسية.

التقييم الرسمي للمقيمين أصبح ممارسة ممارسة معتمدة في العديد من المشروعات الممولة من المانحين وهي نتاج متوقع لبرنامج تدريب المقيم. القسم التالي سيفصل في قياس دقة المقيمين. يقيم المدربين درجة الاتفاق بين المقيمين المتعددين الذين يقومون بإدارة ذات الاختبار في نفس الوقت لنفس الطالب. هذا النوع من الاختبار أو قياس مهارات المقيمين يحدد قدرةالمتدربين على إدارة تقييم القراءة للمراحل الأولى بدقة.

7-8 قياس دقة المقيم

كجزء من عملية اختيار المقيم، يقيس قادة ورشة العمل دقة المقيمن خلال التدريب من خلال تقييم الدرجة التي يتفق فيها المقيمين في تقييمهم لنفس الملاحظة.

نظرة عامة حول التقييم الرسمي لقياس دقة المقيمين خلال التدريب

1. **تقييم واختيار المقيمين.** حدد نقطة تميز ويتم استبعاد المقيمين غير القادريين على تحقيقها من جمع البيانات. في تدريب تقييم القراءة للمراحل الأولى، يتم تحديد هذه النقطة بـ90% اتفاق مع التقييم الصحيح للطفل في التقييم النهائي للتدريب.

2. **تحديد الأولويات للتدريب.** هذه التقييمات الرسمية توضح المهام الفرعية والعناصر التي تمثل تحدي للمقيمين، والتي تمثل أيضا مساحات هامة للتحسين

هذا النوع من التقييم مهم بصورة خاصة لتحسين أداء المقيمين قبل نزولهم للميدان. كما يجب أن يستخدم كذلك لاختيار أفضل المقيمين أداء ليشكلوا فريق المقيمين النهائي للجمع الكلي للبيانات وكذلك البدلاء والمشرفين.

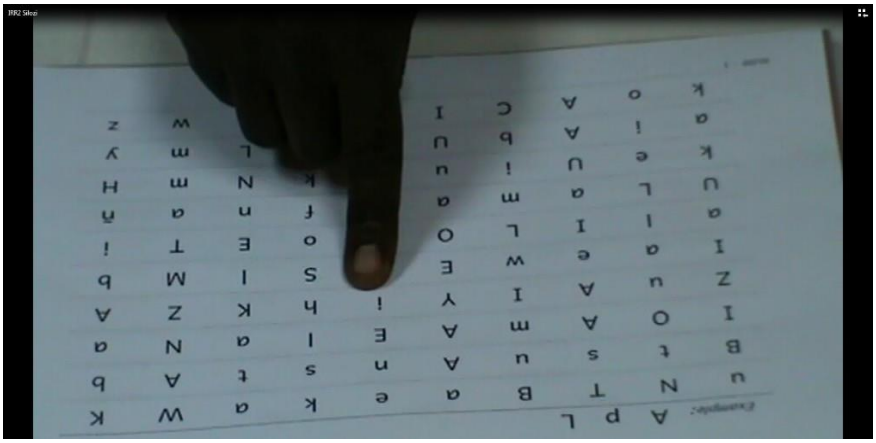
يضع فريق التدريب أداة منفصلة على الأجهزة اللوحية بغرض إجراء قياس لدقة المقيمين.

هناك طريقتان رئيسيتان لتوليد البيانات لحساب دقة المقيم:

1. في حالة تمكن قادة التدريب من الحصول على التصاريح اللازمة قبل ورشة العمل ولعمل تسجيلات صوتية أو مرئية للطلاب المشاركين في التدريب أو التقييمات الاختبارية (انظر **شكل توضيحي 22**)، فإنه يمكن عرض هذه التسجيلات للمجموعة

بينما يقوم **كل** المقيمين بتسجيل درجة التقييم كما لو كانوا يقومون بإدارة تقييم "حقيقي" للقراءة في المراحل الأولى. كما يقوم مقيم من ذوي الخبرة بعمل تقييم ويتم استخدام نتائجة على أنها المعيار الذهبي.

شكل توضيحي 22. لقطة من فيديو يستخدم للتقييم



2. يمكن أن يلعب المدربون أو المقيمون البالغون أدوار "الطالب" و"المقيم" أمام المجموعة الكبيرة (أو في تسجيل فيديو) في حين يعمل المقيمون على تقييم النشاط. فائدة السيناريو الأخير هي أن البالغين يمكن أن يقوموا عن عمد ودون غموض بإرتكاب عدة أخطاء في أي من المهمات الفرعية (مثل تجاوز أو تكرار كلماتأو أسطر، تغيير مستوى الصوت، التوقف لبعض الوقت للتوجيه لنقاط معينة وغيرها). يتم إعداد النص مسبقا شاملا الأخطاء المتعددة ويصبح هو المعيار الذهبي.

يقوم المتدربين بعد ذلك بتحميل كافة التقييمات على برنامج إكسيل أو غيره من برامج التحليل وتحليل النتائج بشكل مقارن. راجع **ملحق ي** للمزيد بخصوص تحليل البيانات والإرشادات الإحصائية حول قياس دقة المقيم.

بعد تقييم المقيم، تقوم الحاجة لتقليص البيانات لمحاولات المتدربين فقط خلال التقييمات بالإضافة إلى تقييم المعيار الذهبي.

في حالة عدم قيام فريق التدريب بوضع المعيار الذهبي لأي سبب، يقوم كبير المدربين بإعداده بعد ذلك ويضيف نتائجه قاعدة البيانات. علاوة على ذلك، فإنه يجب على فريق التدريب مراجعة إجابات القاعدة الذهبية للتأكد من أن كل ما تم تسجيله لكل رد للمعيار الذهبي يعكس الإجماع على الإجابات الصحيحة للتقييم. أحد المناهج الهامة هو مقارنة المعيار الذهبي بالرد الأكثر تكررا للمقيمين على مستوى كل عنصر.

وكما ذكر من قبل، فإن قياس دقة المقيمين مهم لأنه يساعد المدرب على تحديد المقيمين الذين تتجاوز نتائج تقييماتهم انحراف معياري واحد من المعيار الذهبي والذين قد يكونون بحاجة لتدريب أو دعم إضافي. كما يمكن أن يستخدم لتحديد ما إذا كانت المجموعة ككل بحاجة لمراجعة أخرى أو إعادة تدريب على بعض المهام الفرعية أو ما إذا كانت مهارات معينة (من قبيل الوقوف المبكر) تحتاج للمزيد من الممارسة.

إذا ما كشف تحليل التقييم الرسمي أداء منخفض باستمرار من قبل أحد المقيمين، وعدم تحسنه بعد التدريب والدعم الإضافيين، فإنه لن يكون باستطاعة المقيم المشاركة في العمل الميداني. مرة أخرى، راجع **ملحق ي** للمزيد من المعلومات بخصوص كيفية تقييم دقة بيانات المقيم.

بالإضافة لعملية تقييم المقيمين خلال التدريب، فإنه من المطلوب استمرار المقيمين في اختبار الاعتمادية والاتساق فيما بينهم (اعتمادية ما بين المقيمين) بمجرد نزولهم للميدان لجمع البيانات. يمكن أن تساعد اعتمادية ما بين المقيمين على تدعيم اعتمادية واتساق البيانات خلال جمعها وكذلك الحيلولة دون *انجراف المقيم* (انظر دليل المصطلحات). المزيد من المعلومات حول إجراءات اعتمادية ما بين المقيمين خلال عملية جمع البيانات واردة في **الشكل التوضيحي 23؛ ملحق ك** يحوي أشكال نصية تعرض لعدة خطط عينات لتنويع ثنائيات المقيمين.

شكل توضيحي 23. نموذج بروتوكول لمراقبة اعتمادية ما بين المقيمين خلال العمل الميداني

Save the Children.

بروتوكول لجمع بيانات اعتمادية ما بين المقيمين

جزء هام من أي عملية لجمع البيانات عند العمل على تقييمات فردية هو اختبار مدى اتفاق المقيمين فيما بينهم ومدى اعتمادية تقييماتهم للطلاب. في عالم مثالي، سيقوم المقيمون بتسجيل كل إجابة بذات الطريقة؛ إلا أنهم قد يختلفون فيما بينهم بخصوص تسجيل رد الطالب كصحيح أو غير صحيح. غير أنه يؤمل أن تقوم عملية التدريب والاختبار بمساعدة المقيمين على الاتفاق فيما بينهم بشكل متسق. غير أنه يظل من المهم قياس معدل الاتفاق باستمرار بين المقيمين، وهو ما يتم وفق الآلية التالية.

كل يوم، يعمل المقيمون كفريق لتقييم الطالب الأول في اليوم. على سبيل المثال، في حالة تكون فريق المقيمين من أربعة مقيمين، فإن اثنين منهم يشكلان فريقا. يتم تعيين الطلاب لعينات طبيعية عشوائيا. ينادي المقيم الأول (أ) على الطالب الأول للحضور للمكتب/الشجرة/موقع التقييم حيث يكون المقيم الثاني (ب) بالانتظار كذلك. يجلس المقيم ب بحيث لا (ي)تتمكن من رؤية ما الذي يكتبه المقيم أ. يجري المقيم أ التقييم بشكل طبيعي ويسأل الأسئلة الأولية ومهارات القراءة و/أو الحساب المختلفة بينما يسجل إجابات الطالب. يبدأ المقيم ب في عمل تقييم منفصل لنفس الطالب دون أن يقوم بسؤاله أي أسئلة، وإنما يقوم فقط بالاستماع والتسجيل. ومن ثم، فإنه يتم ***تسجيل تقييمين للطالب الأول من كل مدرسة***. سيقوم المقيمين ج و د باتباع نفس الإجراء مع طالب آخر، ويسجلان تقييمين آخرين للطالب الثاني الذي يتم تقييمه في المدرسة.

يجب أن يحرص المقيمين في التقييم على توضيح ما إذا كانوا يقومون بإدارة التقييم أم كانوا فقط يستمعون ويسجلون. وهو ما يتطلب فقرة على كل تقييم ليتمكن المقيمين من تسجيل هذه المعلومة

بعد الانتهاء من تقييم الطالب الأول، فإن مقيم أ يشكره على المشاركة ويعيده إلى صفه. وبعدها يقارن المقيمين أ و ب كيف قاموا بتقييمه، وأن يتناقشوا في العناصر التي اختلفوا بشأن تقييمها ويصلوا غلى اتفاق بخصوص الطريقة الملائمة التي كان يجب تسجيل العنصر وفقها. في حالة عدم تمكن الفريق الثنائي من الوصول لحل لخلافهم في التقييم، فإنه يجب تدوينها ومناقشتها مع مجموعة المقيمين في ختام اليوم. برجاء ملاحظة أنه **بمجرد قيام المقيمين أ و ب بتسجيل إجابة**

في تقييمهم، فإنه لا ينبغي تغييرها أو مسحها أو تصحيحها بعد مغادرة الطالب للمكان. فمن المهم الاحتفاظ بهذه النقاط الخلافية لأنها ستساعد على توفير معلومات حول اتفاق واعتمادية ما بين المقيمين. ومن الطبيعي تماما وجود بعض الاختلافات فيما بين المقيمين، ومن المهم قياس مساحة الاختلاف في عملية تحليل البيانات لأنها ستوفر معلومات بخصوص مدى تأثير أخطاء القياس على النتائج المعينة للقراءة عند الطلاب.

بمجرد انتهاء المقيمين من مضاهاة تقييماتهم، فإنهم يجب أن ينفصلوا وينادي كل منهم على الطالب التالي للتقييم الفردي.

في المدارس التالية، يجب تغيير الفرق وبالتالي يلعب أعضاء آخرون أدوار المقيمين المتحدثين/المستمعين ويتم تعيين كل مقيم مع شريك مختلف كل يوم.

© 2015 هيئة إنقاذ الطفولة. مستخدم بتصريح. كل الحقوق محفوظة.

9 الجمع الميداني للبيانات: الاختبار التجريبي والدراسة الكاملة

9-1 إجراء تقييم اختباري للقراءة في المراحل الأولى

الاختبار التجريبي هو دراسة مبدئية محدودة النطاق يتم إجراؤها قبل المسح الشامل. وتستخدم الدراسات الميدانية للقيام بتقييمات على مستوى العنصر لتقييم كل مهمة فرعية وكذلك صحة واعتمادية أداة تقييم القراءة للمراحل الأولى وأي استقصاءات رأي مصاحبة. كما أن الدراسات التجريبية يمكن أن تختبر أبعاد تطبيق الدراسة (التكلفة، الوقت، الإجراءات الفعالة، والتعقيدات المحتملة) والسماح للعاملين الذين سيقومون على تنفيذ الدراسة الكاملة بممارسة الإدارة في نطاق ميداني حقيقي.

ومن حيث تقييم الأدوات التي سيتم استخدامها خلال جمع البيانات، فإن الدراسة الاختبارية يمكن أن تضمن ملائمة المحتوى المتضمن في التقييم للجمهور المستهدف (على سبيل المثال، مصاغ بشكل واضح وملائم لغويا وثقافيا). كما أنها فرصة كذلك لضمان عدم وجود أخطاء مطبعية أو في الترجمة أو أي توجيهات بحاجة للمزيد من التوضيح.

لماذا يتم إجراء دراسة اختبارية لتقييم القراءة للمراحل الأولى؟

يمكن استخدام الاختبار التجريبي ل

- التأكد من اعتمادية وصحة الأداة من خلال التحليل السيكومتري.
- الحصول على البيانات لاستمارات مختلفة للأدوات، لأغراض المعادلة.[21]
- مراجعة إجراءات جمع البيانات من قبيل عمل الأجهزة اللوحية والأدوات الإلكترونية وكذلك إجراءات تحميل البيانات من الموقع.
- مراجعة جاهزية المواد.
- مراجعة الإجراءات اللوجستية، شاملة الانتقالات والاتصالات فيما بين فرق المقيمين والمنسقين الميدانيين وغيرهم من العاملين.

[21] في حالة الحاجة لوجود عدة نسخ لأداة لدراسة خط الأساس/خط النهاية، على سبيل المثال، فإن تجهيز واختبار الاستمارات الموازية في هذه المرحلة يساعد على تحديد ويحمل إمكانية تقليل الحاجة لمعادية البيانات بعد الجمع الكلي؛ راجع قسم 6.6

للمبادئ التوجيهية لخلق أدوات متوازية وقسم 10.5 للمبادئ التوجيهية حول المعادلة الإحصائية.

الترتيبات العملية للاختبار التجريبي تتشابه إلى حد كبير مع تلك المتوقعة للجمع الكلي للبيانات على الرغم من أن المهام الفرعية قد لا يمكن اختبارها كلها كما أن اعتبارات اختيار العينات الشاملة (كالمناطق والإدارات والمدارس وعدد التلاميذ لكل مرحلة) يتوقع أن تختلف.

شكل توضيحي 24 يبين الاختلافات الرئيسية بين الاختبار التجريبي والجمع الكلي للبيانات.

شكل توضيحي 24. الاختلافات بين الاختبار التجريبي لتقييم القراءة للمراحل الأولى والجمع الكلي للبيانات

الاختبار التجريبي	الجمع الكلي للبيانات	
إكمال التقييم الكلي للتلاميذ والمدارس في العينة	اختبار اعتمادية وصحة وجاهزية الأداة (الأدوات) وتوفر للمقيمين تدريب إضافي.	**الهدف:**
يأخذ في الاعتبار الأجندة الأكاديمية أو الاعتبارات الفصلية (الإجازات، الطقس) كما تأخذ في الاعتبار تعديلات ما بعد الاختبار التجريبي ومراجعات الأداة	يعقد بعد التكيف	**التوقيت:**
تعتمد على الجمهور المستهدف (السنة الدراسية، اللغة، المنطقة، وغيرها)	عينة مقبولة مأخوذة من الجمهور المستهدف للجمع الكلي للبيانات	**العينة:**
يتم تحليلها لمراجعة الأداة (الأدوات) حسب الحاجة	يتم تخزينها خلال عملية جمع البيانات (على سبيل المثال تحميلها على قاعدة بيانات خارجية) وتحليلها بعد جمع كافة البيانات.	**البيانات:**
لا يتم إدخال أي تغييرات على الأداة خلال جمع البيانات	يمكن إجرائها في ضوء تحليل البيانات مع إعادة الاختبار التجريبي على نطاق محدود بعد التغييرات.	**مراجعات الأداة:**

9-1-1 بيانات الدراسة التجريبية ومتطلبات العينة

لضمان كفاية بيانات التطبيق الاختباري للتحليل السيكومتري الذي يتم إجراؤه لتأسيس صحة واعتمادية الاختبار، يجب جمع 150 علامة على الأقل لا تكون قيمة أي منها صفرا ودون أي علامات مفقودة، على أن تكون هذه العلامات الغير صفرية في نطاق معقول ويمكن مقارنتها

بالعلامات غير الصفرية المتوقعة في الدراسة الكاملة. على الرغم من أن العينة الاختبارية للمدارس والتلاميذ سيتم اختيارها عشوائيا بشكل مثالي، فإنه في العادة يتم الحصول عليها من خلال *عينة مقبولة* (انظر دليل المصطلحات). وتوجد ثلاثة أسباب لذلك. أولها أن الغرض الرئيسي للتطبيق الاختباري هو التأكد من عمل الأدلة بشكل مناسب؛ وثانيها أنه لا يتم استخلاص أية نتائج اعتمادا على البيانات الاختبارية فيما يتعلق بالأداء الكلي للطلاب في الدولة ما يعني أن العينة لا يجب أن تكون ممثلة؛ وثالثها أن جمع البيانات باستخدام عينة مناسبة يمكن أن يتم بشكل أسرع وأقل تكلفة مقارنة بجمع البيانات باستخدام عينة عشوائية.

الطلاب والمدارس الواقعين ضمن العينة الاختبارية يجب أن يكونوا مشابهين للجمهور المستهدف في الدراسة الكلية. غير أنه لتقليل عدد العلامات الصفرية التي يتم جمعها بين النتائج التجريبية، فإنه يمكن للمقيمين أن يختاروا عمدا طلابا ذوي مستوى أعلى أو يمكن أن يستهدف المخططون ويزيدوا من مشاركة المدارس ذات المستوى الأعلى. في الدول حيث الأغلبية (70%-80%) من طلاب الابتدائية على علامات صفرية فإنه ستقوم الحاجة لاختيار العينة الاختبارية بحيث تكون كبيرة جدا وبشكل عشوائي للحصول على العلامات ال150 غير الصفرية. على سبيل المثال، إذا ما توقع أن 20% فقط من الحالات ستؤدي لعلامات غير صفرية، فإنه عينة اختبارية من 750 طالبا ستكون مطلوبة للحصول على العلامات ال 150 غير الصفرية المطلوبة للتحليل السيكومتري. غير أن زيادة نسبة المدارس الأعلى أداءا في العينة الاختبارية يمكن أن يقلل حجمها بشكل كبير.

للوقوف على كيفية عمل أداة تقييم القراءة للمراحل الأولى حال تطبيقها على مجموعة متنوعة من الطلاب، فإن البيانات الاختبارية التي تم الحصول عليها من خلال العينة المقبولة يجب أن تشمل تلاميذ منخفضي ومتوسطي ومرتفعي الأداء. برجاء ملاحظة أنه في حالة عدم توافر بيانات عن أداء المدارس، فإنه ينصح بمراجعة المعلومات الاجتماعية-الاقتصادية للمناطق الجغرافية المعينة واستخدامها لتعكس مستويات أداء المدارس. غير أنه من غير المحبذ أن تتضمن العينة المقبولة سنوات دراسية أكبر من تلك للجمهور المستهدف (على سبيل المثال السنة الخامسة بدلا من الثانية) نظرا لأن هؤلاء التلاميذ سيكونوا قد تعرضوا لمواد تعليمية مختلفة عن تلك لطلاب السنة المستهدفة كما أن نطاق العلامات غير الصفرية يمكن أن يختلف بشكل كبير.

أخيرا، العينة الاختبارية، على عكس عينة الدراسة الكلية لتقييم القراءة للمراحل الأولى والتي تقيد عدد الطلاب لكل سنة دراسية ولكل مدرسة في حدود 10-12 تلميذا، تنحو لدمج عدد

أكبر من الطلاب لكل مدرسة في العينة، وهي الطريقة التي تسمح بجمع أسرع لبيانات العينة من خلال عدد أقل من المقيمين. مرة أخرى، هذه الممارسة مقبولة نظرا لأن البيانات الناتجة لا تستخدم في الوصول لنتائج لمستويات الأداء الكلي في الدولة.

9-1-2 تأسيس صحة واعتمادية الاختبار

اعتمادية الاختبار. تعرف الاعتمادية بأنها الاتساق الكلي للمقياس. على سبيل المثال، قد تشير الاعتمادية للدرجة التي تتسق فيها علامات تقييم القراءة للمراحل الأولى عبر الوقت أو عبر مجموعات الطلاب. ويمكن عمل قياس مع مثال من الحياة اليومية مع الميزان. ففي حالة وضع كيس من الأرز على الميزان خمس مرات وتكون قراءة الميزان "20 كيلوجرام" في كل مرة، فبذلك يكون الميزان ينتج قراءات معتمدة. غير أنه في حالة أعطى الميزان قراءات مختلفة (على سبيل المثال، 19، 20، 18، 22، 16) لكل مرة يتم فيها وضع كيس الأرز عليه، فإنه لا يمكن الاعتماد عليه بالتالي.

صحة الاختبار. تشير الصحة لمدى استقامة المقاييس وبالتالي لملائمة الاستنتاجات أو القرارات المعتمدة على نتائج الاختبار. مرة أخرى، باستخدام مثال الميزان، فإنه حال وضع كيس أرز يزن 30 كيلو جراما على الميزان خمس مرات وفي كل مرة تكون القراءة "30" فإن الميزان يقدم نتائج ليست فقط اعتمادية وإنما صحيحة كذلك. أما في حالة أن تكون قراءة الميزان بشكل مستمر "20" في كل مرة يوضع فيها كيس الأرز الثلاثين كيلو، فإن نتائجه تكون ذات اعتمادية (بالنظر لاتساقها) ولكنها غير صحيحة.

المقياس الأكثر شيوعا لاعتمادية نتائج الاختبارات هو **ألفا كرونباخ** والذي يقيس الاتساق الداخلي للاختبار (الحزم الإحصائية ك SAS، SPSS و Stata يمكن أن تحسب قيمة هذا المعامل بيسر). غير أنه قد لا يكون المقياس الأكثر ملائمة لقياس اعتمادية المهام الفرعية أو للتطبيق على عناصر فردية فيها. وهذا بسبب أن أجزاء من تقييم القراءة للمراحل الأولى موقوتة. والمقاييس الموقوتة أو المحددة بوقت والتي يجب أن يجتاز فيها الطلاب العناصر بشكل خطي تؤثر على حساب معامل الألفا بطريقة تجعله تقديرا مبالغا فيه لاعتمادية نتائج الاختبار. وبالتالي فإن ألفا كرونباخ وما شابهها من المقاييس لا تستخدم لتقييم اعتمادية *المهام الفرعية* لتقييم القراءة للمراحل الأولى *كل على حدة*. فعلى سبيل المثال، لن يكون مناسبا حساب ألفا كرونباخ للمهمة الفرعية لقراءة اللاكلمات في تقييم القراءة للمراحل الأولى باعتبار كل

لاكلمة عنصرا. من ناحية أخرى، استخدام العلامة الملخصة (مثل نسبة العلامات الصحيحة أو الطلاقة) للمهام الفرعية وحساب الألفا الكلية لتقييم القراءة للمراحل الأولى (عبر كل المهام الفرعية) باستخدام هذه الأرقام يعد ضروريا .[22]

بخصوص ألفا كرونباخ أو غيرها من مقاييس الاعتمادية، فإنه كلما زادت قيمة المعامل أو الارتباط البسيط، فإن احتمالية تغير علامات تقييم القراءة للمراحل الأولى بسبب التغيرات اليومية العشوائية في ظروف الخاضع للاختبار أو بيئة الاختبار تقل. ومن ثم، فإن قيمة 0.7 أو أعلى تؤخذ على أنها مقبولة على الرغم من أن معظم تطبيقات تقييم القراءة للمراحل الأولى تحصل في العادة على قيمة معامل 0.8 أو أكثر. يوجد توصيف لعدد من الأنواع الأخرى لاختبارات الاعتمادية في ملحق ه.

بالإضافة للمقاييس الأساسية للاعتمادية المذكورة بأعلاه، فمن المفيد اختبار ما إذا كان التقييم أحادي الأبعاد (أي أنه يقيس تركيبا واحدا – مثل القدرة على القراءة في المراحل الأولى). أحد طرق القياس هي إجراء تحليل العامل الاستكشافي، وهذا النوع من التحليل يفترض وجود هيكل ضمني (كامن) في البيانات بهدف التعرف على العدد الكلي للهياكل. يمكن استخدام القيم الذاتية المرتبطة لتحديد ما إذا كان العامل الأول مسئول عن تباين كافي حتى يمكن اعتبار الاختبار الكلي أحادي البعد - أي أن الاختبار يقيس تركيبا كليا واحدا تمكن تسميته "القراءة في المراحل الأولى." وعلى الرغم من أنه لا توجد قيم محددة للقيم الذاتية، فإن الشكل البياني للنقاط يمكن استخدامه كتمثيل مرئي لتحديد ما إذا كانت توجد عدة هياكل أم لا (كوجود انفصال طبيعي يلي العنصر الأول مع مرتفع ذو قيم متراجعة). العديد من الحزم الإحصائية تتضمن إجراءات لتحليل العامل الاستكشافي. وكما هو الحال مع المقاييس الأخرى، فإن التحليل يجب أن يجرى فقط على المقاييس الملخصة للمهام الفرعية (مثل نسبة العلامات الصحيحة أو الطلاقة) وعلى تقييم القراءة للمراحل الأولى ككل، وليس على مدى صحة العناصر الفردية داخل المهام الفرعية. لمعظم تطبيقات تقييم القراءة للمراحل الأولى عنصر أول مسئول عن قدر كافي من التباين لافتراض أن التقييم بالفعل يقيس هيكلا واحدا كليا هاما.

بعد آخر للاعتمادية هو قياس الاتساق بين المقيمين ومدى الاتفاق بينهم (أي اعتمادية ما بين المقيمين) خلال عملية الجمع الميداني للبيانات. فإذا ما استمع مقيمين لنفس الطفل يقرأ قائمة كلمات من اختبار تقييم القراءة للمراحل الأولى، فهل يحتمل أن يقوما بتسجيل نفس العدد من

[22] يجب ملاحظة أنه يتم حساب هذه المقاييس أولا على البيانات الاختبارية للتأكد من اعتمادية الأداء قبل الإدارة الشاملة ولكن يتم إعادة حسابها على البيانات العملية (أو المسح الشامل) للتأكد من استمرار الاعتمادية المرتفعة.

من الكلمات على أنه تمت قراءته بصورة صحيحة؟ هذا النوع من قياس الاعتمادية يتضمن إدارة المقيمين للمسح في فرق ثنائية، يقوم أحدهما بإدارة التقييم في حين يقوم الآخر فقط بالاستماع ويقيم بشكل مستقل. تمكن مراجعة المزيد من الشرح لكيفية إدارة اعتمادية ما بين المقيمين في قسم 8، وخصوصا شكل توضيحي 23 "نموذج بروتوكول لمراقبة اعتمادية ما بين المقيمين خلال العمل الميداني." يمكن بعد ذلك قياس الاتفاق بين المقيمين من خلال تقدير معامل *كابا* كوهين (انظر دليل المصطلحات). هذا الإحصاء (والذي يأخذ بعين الاعتبار معامل تخميني) يعد تحسينا على نسبة الاتفاق بين المقيمين، ولكن يجب ذكر كلا القياسين. وبينما يثور جدال بشأن النقاط المحددة ذات المعنى لكابا كوهين، فإن المعلومات بشأن نقاط التميز لاتفاق المقيمين والمقاييس التي يرجع إليها في العادة بخصوص إحصاءات كابا يمكن الاطلاع عليها في ملحق ط، قسم ط.4.

للتحقق من صحة البناء، فإن الإحصاءات على مستوى العنصر يجب إنتاجها للتأكد من أن كافة العناصر تعمل كما ينبغي. تحليلات راش (والتي تعتمد على افتراض أحادية البعد) توفر معلومات حول صحة البناء بعدة طرق. أولا، يضع نموذج راش العناصر والطلب على نفس المستوى في القياس بالترتيب من السهل (الطلاب الأقل قدرة) إلى الصعب (الأعلى قدرة). ومن ثم، فإن ترتيب العناصر من الأقل للأكثر صعوبة هو التعريف الإجرائي للبناء. في حالة توافق هذا التعريف مع التصميم المقترح، فإنه توجد مؤشرات على صحة البناء. غير أنه في الحالات التي لا يكون فيها عناصر ممثلة للطلاب تقيم قدراتهم بدقة. وأخيرا، تحليلات راش تقيم أداء العنصر من خلال إحصاءات التوافق. فإذا ما كانت العناصر لا تقوم بقياس القدرة بدقة، أو تنتج "ضوضاء" فبالتالي فإن القيم الأكبر للإحصاءات (2.0 أو أكثر) تشير لوجود تباين وبالتالي تصبح بحاجة لإعادة التقييم. التقييمات التي تحتوي العديد من العناصر غير المتوافقة يشار إليها بأنها تعاني من *تباين غير مرتبط بالبناء*، والتي تعد محددا لصحة البناء. مخرجات نموذج راش يمكن أن تساعد واضعي الامتحان على تحديد ما إذا كانت العناصر تتعامل وفق ما هو متوقع أم لا وأي من العناصر (إذا ما وجدت) تنبغي إزالته أو مراجعتها بسبب عدم التوافق. من الضروري أن تجرى هذه التحليلات على كلا من البيانات الاختبارية (لبيانات الاختبار التشغيلية الأولية) وبيانات الدراسة الكاملة (لتحديد ما إذا كانت توجد عناصر معينة تجب إزالتها قبل إجراء التقييم).

خلال الفترة ما بين الاختبار التجريبي والجمع الكلي للبيانات، يقوم الإحصائيين والسيكوميتريين بتحليل البيانات واقتراحات أي تعديلات مطلوبة؛ يقوم متخصصوا اللغة

والمترجمين بعمل التصحيحات؛ وتحديث النسخ الإلكترونية من الأدوات وإعادة تحميلها على الأجهزة اللوحية؛ والتعامل مع أي مشاكل في الأجهزة وإعادة تدريب المقيمين والمشرفين على التغييرات.

3-1-9 اعتبارات متعلقة بتوقيت الاختبار التجريبي

يناقش هذا القسم ميزات ومثالب اختيارين لتوقيت الاختبار التجريبي أخذا في الاعتبار تدريب المقيم والجمع الكلي للبيانات.*

يمكن أن يتم إجراء الاختبار التجريبي قبل او بعد تدريب المقيم. وهناك مميزات وعيوب لكلا الاتجاهين ويتخذ القرار عادة في ضوء الترتيبات اللوجستية والسياق.

في حالة عدم توافر مقيمين من ذوي الخبرة (من تجربة سابقة مع التقييم)، فإنه قد يكون من الأفضل عقد الاختبار التجريبي عقب إنتهاء ورشة عمل تدريب المقيمين مباشرة. وبشكل عام فإن الاختبار التجريبي يستغرق ما بين يوم إلى يومين بعد تجهيز المقيمين المدربين. ميزة هذا المنهج هي أن الاختبار التجريبي، علاوة على توليد بيانات هامة حول الأدوات ذاتها، فإنه يقدم معلومات قيمة بخصوص أداء المقيمين. من يقومون على تحليل البيانات الاختبارية يمكن أن ينظروا لمؤشرات على ارتكاب المقيمين لأخطاء معينة من قبيل استعجال الأطفال أو السماح بأكثر من الوقت المسموح لأداء مهام معينة.

عيب الاختبار التجريبي عقب تدريب المقيمين هو أن الأدوات التي يتم استخدامها خلال تدريب المقيمين لا تكون في صورتها النهائية بالنظر لأنها لم تكن قد تم تجريبها. في العديد من الحالات، قد يأتي الاختبار المبكر الأقل رسمية للأدوات ليساهم في صقلهم بحيث لا يؤدي الاختبار التجريبي الرسمي لمراجعات كبيرة في الأداء. غير أنه في هذا السيناريو يجب أن يكون المقيمين على علم بأن الأدوات التي يتمرنون خلال التدريب قد تجرى عليها تغييرات بسيطة لاحقا خلال جمع البيانات. يجب أن يوضح المنفذ كافة التغييرات التي تحدث عقب الاختبار التجريبي بشكل كافي لكافة المقيمين قبل النزول للميدان.

عندما يجري الاختبار التجريبي بعد تدريب المقيمين مباشرة، فإنه ينصح بفترة فاصلة على الأقل لمدة أسبوعين ما بين الاختبار التجريبي والجمع الكلي للبيانات للسماح بتحليل البيانات الاختبارية ومراجعات الأدوات والطباعة وتحديث وسائط الجمع الإلكتروني للبيانات، وتوزيع المواد لفرق التقييم.

في حالات أخرى، يفضل عقد الاختبار التجريبي قبل تدريب المقيمين. في السياقات حيث تم عقد تقييم القراءة للمراحل الأولى قبل ذلك (ليس قبل أكثر من سنتين) وبالتالي يتوافر المقيمين ذوي الخبرة فإن تدريب تذكير موجز لمدة يوم أو يومين يمكن أن يكون كافيا للتجهيز للاختبار التجريبي. ميزة هذا الاختيار أنه يمكن الانتهاء من الأدوات (اعتمادا على تحليل البيانات من الاختبار التجريبي) قبل الشروع في تدريب المقيمين. واتساقا مع التوصية بعاليه، فإنه من الحكمة السماح بأسبوعين على الأقل بين الاختبار التجريبي وتدريب المقيمين بحيث يمكن تجهيز المواد ليس فقط للتدريب ولكن لجمع البيانات. في هذا السيناريو، يبدا جمع البينات بمجرد انتهاء التدريب.

* القسم المظلل من هذا القسم الفرعي مأخوذ مباشرة من Kochetkova and Dubeck (تحت النشر). © معهد اليونسكو للإحصاءات. مستخدم بتصريح. كافة الحقوق محفوظة.

2-9 إجراءات جمع البيانات ميدانيا للدراسات الكاملة

الانتقالات. كل فريق ستكون لديه سيارة لنقل المواد وللوصول لمدارس العينة قبل بداية اليوم الدراسي.

عدد التقييمات. تكشف الخبرة حتى الآن أن تطبيق تقييم القراءة للمراحل الأولى يتطلب ما بين 15 إلى 20 دقيقة للطفل. خلال الجمع الكلي للبيانات، يعني هذا أن فريق من ثلاثة مقيمين يمكن أن يكملوا 9-10 أدوات خلال ساعة أو حوالي 30 طفلا خلال ثلاث ساعات متواصلة.

التحكم في الجودة. من المهم التأكد من جودة الأدوات المستخدمة والبيانات التي يتم تجميعها. يجب أن يتبع المنفذين أفضل ممارسات البحث العامة:

- التأكد من أمن وسلامة الأطفال الخاضعين للاختبار، شاملة الحصول على موافقة الطفل.

- التأكد من كلية الأدوات (أي تفادي النشر العام).
- التأكد من جمع وإدارة والتقرير حول البيانات بمسئولية (الجودة والسرية وطمس الهويات[23])
- مراقبة بيانات اعتمادية ما بين المقيمين لتحسين جودة البيانات وتقليل فرصة "الانحراف" – (يعرف كذلك ب *انحراف المقيم* – انظر المصطلحات).
- الاتباع الصارم للتصميم البحثي.

الأدوات. توفير الأجهزة الكافية والمواد للمقيمين والمشرفين هي بعد هام آخر لمرحلتي جمع البيانات ميدانيا.

لجمع البيانات، تشمل الاحتياجات ما يلي:

- جهاز لوحي، كامل الشحن ومحمل بالنسخة الحالية للأداة
- كتاب مجلد يضم محفزات الطلاب، نسخة لكل مقيم (يتم استخدام نفس الكتاب لكل طالب يختبره المقيم)[24]
- ساعات توقيت أو مؤقتات زمنية (في حالة عدم عمل الأجهزة اللوحية واستخدام أدوات الدعم الورقية)
- أقلام رصاص وممحاة ولوحات
- أقلام رصاص أو غيرها من المستلزمات المدرسية الصغيرة لإعطائها للطلبة تقديرا لمشاركتهم (في حالة تأكد المخططين مسبقا أن القيام بذلك يتفق مع قواعد المانحين)

[23] طمس الهويات: سمعة تقييم القراءة للمراحل الأولى والأدوات الشبيهة تعتمد على موافقة المدرس/الطالب وضمان عدم التعريف بالأشخاص. في حالة اساءة استخدام البيانات – حتى الاختبارية منها – (مثل تسمية المدارس ومجازاتها) فإن هذا الأمر قد يقوض كامل المنهج للتقييم لصنع القرار في بلد أو منطقة معينة.

[24] بالنظر لأن لوحات تحفيز الطلاب سيتم استخدامها مع طلاب مختلفين، فإن التجليد، على الرغم من أنه ليس ضروريا في المطلق، فإنه يطيل عمر استمارات إجابة الطلاب (أوراق الحماية البلاستيكية في مجلدات يمكن ان تكون مفيدة كذلك).

الإشراف. من المهم الترتيب لمصاحبة مشرف لكل فريق من المقيمين. يوفر المشرفين مراقبة هامة للمقيمين ولعملية الجمع. كما أنهم قادرون كذلك على إدارة العلاقات مع مسئولي المدرسة ومصاحبة الطلاب من وإلى مكان الاختبار وتجديد احتياجات المقيمين، التواصل مع فريق الدعم والعمل كمقيم حال اقتضت الحاجة.

اللوجستيات. الاختبار التجريبي مفيد لتقدير الترتيبات والدعم اللوجستيين المخططين لعملية جمع البيانات. غير أن الجمع الكلي للبيانات يتضمن أبعادا إضافية أخرى للدراسة يجب التعامل معها قبل مغادرة المقيمين للميدان: التحقق من عينة المدارس وتحديد المواقع وترتيب الانتقال إلى/الإقامة حول المدارس. كما أن وجود خطة أولية مهم للغاية وستتضمن دائما قائمة بالتواريخ والمدارس وأرقام الاتصال بنظار المدارس وأسماء أعضاء الفريق. هذه القائمة يتم وضعها من قبل شخص ذو معرفة بالمنطقة. علاوة على ذلك، فإن احصائي الدراسة سيضع المعايير الاحصائية للعينة والبروتوكولات لاستبدال المدارس والمدرسين و/أو الطلاب ويقوم فريق التدريب بتوصيلهم للمقيمين. أخيرا، لمرحلة الجمع الكلي للبيانات، فإن المخططين ينظموا ويرتبوا لتوصيل مواد التقييم والأدوات كنسخ دعم للأدوات والأجهزة اللوحية وخطابات تفويض المدارس.

قبل التحرك للمدارس، المقيمين والمشرفين:

- مراجعة كافة المواد
- مناقشة إجراءات واستراتيجيات إدارة الاختبار لجعل الطلاب يشعرون بالراحة
- التأكد من شعور كافة المديرين بالراحة باستخدام ساعة توقيت أو ساعاتهم في حالة عدم عمل الأجهزة اللوحية.

عند الوصول للمدرسة، يقوم المشرف بتقديم فريق المقيمين لناظر المدرسة. في معظم الدول، يطلب خطاب موقع من الحكومة لإجراء الاختبار كما يشرح المشرف أيضا الغرض وأهداف التقييم ويشكر ناظر المدرسة على مشاركة المدرسة في تقييم القراءة للمراحل الأولى. يجب أن يؤكد المشرف للناظر أن الهدف من الزيارة **ليس** تقييم المدرسة أو الناظر أو المدرسين وأن كافة المعلومات ستظل مجهلة.

يجب أن يسأل المشرف الناظر عن فصل دراسي متاح أو حجرة مدرسين أو مكان هادئ لكل من المقيمين لإجراء التقييمات الفردية. ويتوجه المقيمين إلى المكان المعين ويجهزون مقعدين أو مكتبين، واحد للطالب والآخر للمقيم. من المفيد كذلك السؤال عما إذا كان يتوافر شخص في المدرسة للمساعدة خلال اليوم؛ كما يظل هذا الفرد كذلك مع الطلاب المختارين في المكان المحدد.

خلال اليوم الأول للتقييم، يرتب المشرف للمقيمين للعمل في فرق ثنائية لإدارة تقييم القراءة في المراحل الأولى بشكل متزامن لأول طالب يتم اختياره يقوم أحدهما بإدارتها في حين يلاحظ الثاني في صمت ويعطي الدرجة. هذا التقييم الثنائي – والذي يساعد على التأكد من جودة البيانات من خلال تقييم اعتمادية ما بين المقيمين بشكل مستمر – كما هو مبين في قسم 8.7 وملحق ي.

خلال اليوم المدرسي، وهو محور التركيز الرئيسي للأطفال في الدراسة. سيتم تدريب المقيمين على بناء صلات جيدة ولكن عادة ما سيكون الاختبار التجريبي هو أول فرصة يعملون فيها مع الأطفال. سيراقب المشرفين عن كثب للتأكد من عدم شعور أي من الطلاب بالضغط أو الضيق وأن المقيمين يخصصون بعض من الوقت لتأسيس علاقات إيجابية مع الطلاب قبل طلب الحصول على موافقتهم. أي نقاط رئيسية من ملاحظات المقيمين حول العمل مع الأطفال يتم تشاركها خلال ملخص الاختبار التجريبي للتأكد من أنه بمجرد خروج الفرق للعمل في الميدان، فإنهم سيكونون أكثر قدرة على التعامل مع الأطفال. بعض الأمور التي قد تكون بسيطة من قبيل التأكد من ضبط المقيمين هاتفهم المحمول على وضع الصمات قد تصنع فارقا للطلاب.

يجب أن يذكر المشرف المقيمين أنه في حالة عدم موافقة الطلاب على الخضوع للاختبار فإنه يجب صرفهم بلطف واختيار بديل وفق البروتوكول المتبع.

في حالة عدم توفير الناظر لمكان للنشاط، فإن فريق التقييم سيتعاون لإيجاد مكان هادئ (ملائم لتفاعل البالغ/الطفل) ويكون مناسبا لعلم التقييم. ينبغي أن يتوافر في المكان:

- إضاءة كافية للقراءة وللمقيمين لقراءة لمطالعة الجهاز اللوحي
- مكاتب موضوعة بحيث لا يمكن للطلاب أن ينظروا للخارج من شباك أو باب أو في مواجهة طلاب آخرين.

- مكاتب فارغة من كافة الأوراق والمواد (توضع مواد المقيمين على طاولة منفصلة حتى لا يتشتت الطالب)
- أن يكون بعيد عن مدى باقي الطلاب المختارين؛ بحيث لا يكون بوسعهم سماع أو رؤية الاختبار.

3-9 اختيار الطلاب

يقدم هذا القسم اختيارين لسحب عينة من بين الطلاب بمجرد وصول المقيمين للمدرسة المعنية. الأول يعتمد على القيد، والثاني يسمى العينة المنفصلة.

1-3-9 الاختيار الأول لعينة الطلاب: جدول الرقم العشوائي

في حالة توافر بيانات دقيقة ومحدثة حول قيد الطلاب ودرجاتهم وصفوفهم متاحة على المستوى المركزي قبل وصول فرق التقييم للمدارس، فإنه يمكن استخدام جدول الرقم العشوائي لخلق عينة الطلاب. خلق هذا الجدول قد يكون أكثر دقة من الناحية الإحصائية عن العينة المنفصلة، وبالنظر للاحتمالية العالية لعدم توافر شروط هذا السيناريو في سياقات العديد من الدول، فإنه يتم استخدام الخيار الثاني بشكل أكثر شيوعا.

2-3-9 الاختيار الثاني لعينة الطلاب: العينة المنفصلة

هذه الطريقة لسحب العينة تتضمن تأسيس عينة منفصلة لكل صف يتم تقييمه في المدرسة. الفكرة هي تحديد عينة منفصلة لاختيار الطلاب عشوائيا، انطلاقا من عدد الطلاب الحاضرين خلال يوم التقييم. هذه الطريقة تتطلب ثلاثة خطوات منفصلة.

الخطوة الأولى: اعتمادا على المشروع البحثي، يتم تحديد أي المجموعات ستشكل القاعدة لسحب العينة.

من المهم ملاحظة أنه يجب الانتهاء من الخطوة الأولى قبل وصول المقيمين للمدرسة. ويتم الوصول لهذا الحسم خلال فترات التخطيط الأولى لتصميم البحث وسحب العينة. خلال تدريب المقيم، سيتم تلقين المقيمين المرشحين لتطبيق منهاجية سحب العينة اعتمادا على تصميم البحث.

الهدف من الخطوة الأولى هو تحديد دور بيانات المدرس والدرجة (الدرجات) و/أو الصف (الصفوف) المطلوبة، والتوقعات للتقارير حول النتائج المنفصلة للأولاد والبنات. **شكل توضيحي 25** في الصفحة التالية يعرض الاعتبارات المطلوبة.

شكل توضيحي 25. محددات مجموعات العينة

تصميم البحث – بيانات المعلم:	لا يتضمن المسح بيانات المعلم والتي سيتم ربطها بالطلاب	يتضمن المسح بيانات المعلم لمدرس واحد في كل صف والذي سيتم ربطه ببيانات أداء الطلاب	يتضمن المسح بيانات المعلم لمدرسين مختلفين في كل صف والتي سيتم ربطها ببيانات أداء الطلاب
أساس سحب العينة – الفصل أو الصف:	مستوى الدرجة	مستوى الفصل – فصل لكل صف	مستوى الفصل – أكثر من فصل لكل صف

ملاحظات:

- قد تتضمن المسوح صف أو أكثر.
- بالإضافة لاختيار الفصول/الصفوف، فقد يحدد التصميم البحثي أن الطلاب يمكن اختيارهم وفق النوع (انظر الصف التالي)
- مواد المدارس للمقيمين تتضمن مجموعة من النرود لاختيار الفصل أو الفصول، في حالة وجود مدرسين متعددين للصف المعين. بروتوكول العين يحدد كيفية استخدام النرد.

المجموعة (المجموعات) التي يجب اختيار العينة (العينات)من بينها:	إما: كافة الطلاب في كل صف (بغض النظر عن النوع) أو: كافة الطلاب الذكور في كل صف، و كل الطالبات في كل صف	إما: • كافة الطلاب من كل من الفصول المختارة في كل صف (بغض النظر عن النوع) أو: • كافة الطلاب الذكور في كل من الفصول المختارة، و • كل الطالبات في كل من الفصول المختارة

الخطوة الثانية: تحديد عدد الطلاب الذين يتم اختيارهم من كل مجموعة: ن

الخطوة الثانية تتضمن القيام بحسابات اعتمادا على العدد الإجمالي للطلاب المكونين للعينة لكل مدرسة وعدد المجموعات المتضمنة.[25]

توضيح: إذا كان العدد الكلي للطلاب في العينة 20 للمدرسة ويتم اختيارهم من فصل واحد في كلا الصفين (مثل الصفين الثاني والثالث) وفقا للنوع، ثم أربع مجموعات وخمسة طلاب (20÷4) يتم اختيارهم من كل مجموعة على النحو التالي:

1. 5 طلاب ذكور من الفصل المختار للصف الثاني
2. 5 طالبات من الفصل المختار للصف الثاني
3. 5 طلاب ذكور من الفصل المختار للصف الثالث
4. 5 طالبات من الفصل المختار للصف الثالث

[25] انظر الملحقين ب و ج وقسم 5 للمزيد من المعلومات حول تصميم العينة.

الخطوة الثالثة: اختيار عدد ن عشوائيا من الطلاب من كل مجموعة

الهدف من هذه الخطوة هو اختيار الأطفال المعينين ليتم تقييمهم. الإجراء الموصى به هو:

1. دع الطلاب يشكلون خط مستقيما خارج الفصل.

- إذا كان التقييم لطلاب من صفوف مختلفة، ابدأ مع الأطفال من الصفوف الأصغر في بداية اليوم.

2. قم بعدد الأطفال في الخط: م.
3. قم بقسمة م على *ن* (من الخطوة الثانية) وقم بتقريب الإجابة لأقرب رقم صحيح: ب.[26]
4. ابتدءا من أحد طرفي الخط، قم باختيار طفل عشوائيا من الأطفال ب الأول، وبعدها قم بالعد واختار كل طفل ب بعد ذلك.

<u>توضيح:</u> لاختيار *ن* =8 أطفال من مجموعة ما:

1. يوجد 54 طفل في الخط (*ن* =54)
2. احسب ب: *54* ÷8 = 6.75؛ مقربة ب =7
3. اختار طفل من أول ب =7 [26]– على سبيل المثال الطفل رقم 3
4. اختار كل طفل سابع ابتداءا من طفل رقم 3:

3; 10; 17; 24; 31; 38; 45; 52

لاحظ أن هذه العملية ستسفر عن اختيار 9 أطفال – ويكون التاسع بديلا في حالة عدم رغبة أي من الأطفال الآخرين في المشاركة. في المثال السابق، ينبغي أن يستمر المقيم في العد حتى نهاية الخط ثم يعود للبداية ليختار الطفل السابع التالي (والذي سيكون الخامس من أول الخط).

[26] تعرف هذه العملية "بالبداية العشوائية"

بمجرد قيام المقيمين بإدارة تطبيق القراءة للمراحل الأولى لكل الطلاب في المجموعة الأولى (الموضحة في الخطوة الثانية)، يكرر فريق التقييم خطوة 3 لاختيار الأطفال من المجموعة الثانية. يتأكد المشرف من أن المقيمين دائما يعملون مع طالب ليقيموه حتى لا يضيع الوقت خلال التقييم.

4-9 ختام يوم التقييم: إتمام التقييم

في حدود الممكن، يجب إتمام كل المقابلات في مدرسة معينة خلال اليوم الدراسي. إلا أنه يجب وضع خطة طوارئ في بداية اليوم ومناقشتها مسبقا مع المقيمين والمشرفين كأفضل تطبيق ملائم في ضوء الظروف المحيطة. في حالة كانت المدرسة فترة واحدة ولم يتم الانتهاء من بعض التقييمات قبل نهايتها، سيلتقي المشرف بباقي الطلاب ويسألهم الانتظار بعد نهاية اليوم المدرسي. في هذه الحالة، ويقوم المدرسين أو مدير المدرسة بتنبيه أولياء الأمور أن بعض الطلاب سيتأخرون في العودة للمنزل.

في حدود الممكن، يجب إتمام كل المقابلات في مدرسة معينة خلال اليوم الدراسي. إلا أنه يجب وضع خطة طوارئ في بداية اليوم ومناقشتها مسبقا مع المقيمين والمشرفين كأفضل تطبيق ملائم في ضوء الظروف المحيطة. في حالة كانت المدرسة فترة واحدة ولم يتم الانتهاء من بعض التقييمات قبل نهايتها، سيلتقي المشرف بباقي الطلاب ويسألهم الانتظار بعد نهاية اليوم المدرسي. في هذه الحالة، ويقوم المدرسين أو مدير المدرسة بتنبيه أولياء الأمور أن بعض الطلاب سيتأخرون في العودة للمنزل.

5-9 رفع البيانات المجمعة ميدانيا

بافتراض جمع البيانات إلكترونيا (وهو أفضل الممارسات الموصى بها حاليا – انظر قسم 7)، يوفر المنظمون الوسائل للمقيمين التي ينقلون عبرها البيانات لسيرفر مركزي كل يوم لتفادي أي فقد محتمل للبيانات (في حالة كسر أو فقد أي جهاز محمول). في حالة عدم إمكانية ذلك، يتم تطبيق إجراءات الدعم. الإجراءات لضمن رفع أو تخزين البيانات بشكل ملائم واحدة لكلا من التطبيق الاختباري والجمع الكلي للبيانات. ويعد التطبيق الاختباري فرصة هامة للتأكد من أن هذه الإجراءات تعمل بشكل جيد.

مزايا الرفع والمراجعة الدوريين للبيانات

خلال جمع البيانات، يمكن أن يساعد الرفع والمراجعة الدوريين في الوقوف على أي أخطاء قبل الانتهاء من جمع البيانات وتفادي إعادة جامعي البيانات مرة أخرى للميدان بعد أسابيع من جمع البيانات. كذلك، فإن الرفع اليومي يمكن أن يساعد على تفادي خسارة كميات كبيرة من البيانات في حالة فقد جهاز لوحي أو سرقته أو تحطمه. تمكن مراجعة البيانات للتأكد من أنه يتم تقييم الدرجة الصحيحة وأن المقيمين يذهبون للمدارس في العينة وأن العدد الصحيح من الطلاب يتم تقييمه وللوقوف كذلك على أي مظاهر لعدم الاتساق. الاتصال المستمر والتحديثات لإعلام فريق المشروع بموعد بدء جمع البيانات ومتى يراجع محللو البيانات ما تم رفعه وما إذا كان هناك أي تأخير أو أسباب تحول دون رفع البيانات بشكل يومي يمكن أن تساعد في مراجعة البيانات وكذلك في معرفة أي النتائج يمكن توقعها ومتى.

سيرسل المقيمون بياناتهم للسيرفر المركزي عبر الإنترنت اللاسلكي، سواء من خلال شبكة لاسلكية في مكان عام أو مقهى إنترنت أو بشراء بيانات محمول (3G). عند التخطيط لجمع البيانات، يجب أن ينظر المخططون لعوامل كشبكات المحمول المتوافرة والتوافق بسن أجهزة الراوتر والمودم اللاسلكية والقدرات الفنية للمقيمين ويسعون لتوفير أفضل الحلول العملية والمعتمدة. خلال التطبيق الاختباري، يتدرب المقيمون على رفع وتخزين البيانات باستخدام الطريقة المختارة. يتحقق محلل بيانات من الرفع الفعلي للبيانات على السيرفر ويراجع قاعدة البيانات للوقوف على اي أخطاء فنية (مثل أسماء المتغيرات المتداخلة) قبل الانتقال في الجمع الكلي للبيانات.

إجراءات التخزين للجمع الإليكتروني للبيانات تتضمن توافر نسخ ورقية من الأداة متاحة لاستخدام جامعي البيانات. بعد كل تقييم يتم إكماله ورقيا، يراجع المشرف الاستمارة للوضوح والاكتمال (أي لا يوجد كود مدرسي مفقود أو علامات غامضة). المشرف أو الشخص المعين مسئول عن الاحتفاظ بالاستمارات الكاملة منظمة وآمنة من الضياع أو التلف ولا يصل غليها سوى المخولين بذلك.

10 إعداد بيانات أداة تقييم القراءة للصفوف الدراسية الأولى

يغطي هذا القسم عملية تنظيف وإعداد بيانات أداة تقييم مهارات القراءة للصفوف الأولى. بمجرد أن بجمع البيانات ينبغي إعادة ترميزها وصياغتها؛ لعمل ملخصات ومتغيرات شديدة الاختصار. لاحظ أن هذا القسم يفترض أن التعديلات على أخطاء أخذ العينات من تصميم المسح طُبقت بشكل مناسب.

تتكون كل استطلاعات أداة تقييم مهارات القراءة للصفوف الأولى تقريبًا شكل من أشكال العينات الطبقية متعددة المراحل. ولذا يجب الحرص على العناية بمراقبة، وتحقق، وتحرير، ودمج، ومعالجة البيانات جيدًا من أجل تحليلها نهائيًّا. ويجب أن تتم هذه العمليات عن طريق اثنين من الإحصائيين المتمرسين (بحد أقصى)، ليقوم أحدهما بالقيام بهذه الخطوات بينما يقوم الآخر بالتحقق من صحتها. وبمجرد أن تُعالج البيانات ويُنتهى منها يستطيع أي شخص ذا خبرة في استكشاف العينات المركبة والبيانات الهرمية تعرف أهداف البحث، وكذلك الاستبيانات والتقييمات، ومنهجية العينة، وهياكل البيانات، ومن ثم تحليل البيانات.

يفترض هذا القسم أن الإحصائي أو الإحصائيين الذين يقومون بمعالجة البيانات لديهم خبرة واسعة في معالجة العينات المعقدة، وهياكل البيانات الهرمية، ويستطيع أن يعطينا بعض تفاصيل معالجة البيانات الخاصة بأداة تقييم مهارات القراءة للمراحل الأولى.

10-1 تنظيف البيانات

إن تنظيف البيانات التي جُمعت خطوة بالغة الأهمية قبل تحليل البيانات، ومرة أخرى؛ يجب أن نعرف أن تنظيف البيانات ومراقبتها يتم عن طريق إحصائي ذي خبرة في مثل هذا النوع من معالجة البيانات.

وتُراقب جودة البيانات في أثناء جمعها. وباستخدام تقارير وجداول جمع البيانات من فريق العمل الميداني، وسيكون الإحصائي قادرًا على أن يطابق البيانات التي تُحمَّل على الأرقام المتوقعة للتقييمات لكل مدرسة، ولغة، ومنطقة، أو عينة أخرى. وخلال هذا الوقت، سيتمكن

الإحصائي المسئول عن المراقبة من التواصل مع العاملين في الميدان؛ لتصحيح أية أخطاء حدثت في أثناء إدخال البيانات، ومن أجل ضمان تنفيذ عدد مناسب من التقييمات في المدارس الصحيحة، وفي الأيام المخصصة للتقييم. ويشكل جمع البيانات من ثلاث مصادر أو أكثر من مصدر المعلومات المحددة جانبا مهمًا للتأكد من صحة عينات كبيرة بما فيه الكفاية لأغراض الدراسة، ولا شك أن القدرة على تحديد وتصحيح أي من هذه التناقضات سيساعد على تنظيف وتنقية البيانات، وسيضمن أيضًا أنه لن تُؤجل عملية جمع البيانات، أو تكرارها بسبب بعض الأخطاء الصغيرة.

الملحق 26 هو قائمة قصيرة ليستخدمها الإحصائيون في أثناء عملية التنظيف؛ للتأكد من أن جميع بيانات أداة تقييم مهارات القراءة للمراحل الأولى قد نُظفت تماما، وبشكل موحد لأغراض تحليل البيانات.

الملحق 26: قائمة تنظيف البيانات

- **مراجعة التقييمات الغير مكتملة.**
  تحقق من التقييمات غير المكتملة؛ لتحديد مستوى اكتمالها وملاءمتها؛ للبقاء في البيانات النهائية. سيكون لكل مشروع معايير متفق عليه لاتخاذ هذه القرارات. على سبيل المثال، يمكن الإبقاء على التقييمات التي لم ينتهي منها بشكل كامل، إذا كان ذلك ضروريًّا لأغراض حجم العينة، واستخدام المعلومات غير الكاملة، أو يمكن التحقق من دقة التقييمات، والتأكد من أنها لا تخلو من أي معلومات تحديد مهمة.
- **التخلص من أي تقييمات "تجريبية" انتهت قبل بدء جمع البيانات الرسمية.**
  تحقق من أن كافة التقديرات الواردة في النسخة "النظيفة" من البيانات المستخدمة للتحليل حقيقية، وتحققت خلال جمع البيانات الرسمية.
- **التأكد من أن جميع التقييمات ترتبط بالمعلومات المدرسية المناسبة لتحديد الهوية.**
  احذف أو أزل أي تقييمات لم تُحدد هويتها بشكل مناسب، أو تعاون مع الفريق الميداني؛ للتأكد من أن تُحدد هوية أي تقييمات غير محددة بدقة وبشكل مناسب.
- **ضمان الحصول على موافقة الطفل وتسجيلها لكل ملاحظة**
  احذف أو أزل أية تقييمات تمت بدون حصول المقيم على موافقة الطفل على تقييمه وتسجيل الموافقة.

- **حساب كل نتائج المهام الفرعية سوآءا كانت موقوتة أو غير موقوتة**
يمكن الاطلاع على المعلومات الخاصة بالمهام الفرعية الموقوتة والغير موقوتة في القسم 2.10
- **التأكد من أن جميع نتائج المهام الفرعية الموقوتة تقع ضمن نطاق مقبول وواقعي.**
قد يرتكب المقيمون بعض الأخطاء في أثناء جمع البيانات، أو تحدث أعطال ببرامج جمع البيانات قد تؤدي إلى وجود بعض النتائج غير المعقولة. لذا يجب التحقق من أي نتائج عالية بشكل غير طبيعي، والتحقق من أنها واقعية بالنسبة للتلميذ الذي يُقيم (على أساس أداء الطفل في المهام الفرعية الأخرى)، وأنها لم تنتج عن خطأ ما. فمن الضروري إزالة أي نتائج مبالغ فيها ناتجة عن أخطاء في التقييم، حتى لا يفسد تحليل البيانات. ليس من الضروري إزالة كل الملحوظات الخاصة بهذا التلميذ تحديدًا؛ لأن هذا من شأنه أن يؤثر على حجم العينة الخاضعة للتحليل في المهام الفرعية الأخرى. ببساطة أزل أي نتائج من المهمة الفرعية التي يظهر بها الخطأ

2-10 معالجة المهام الفرعية؛ لتقييم مهارات القراءة بالمراحل الأولى

يبدأ هذا القسم بتسميات للمتغيرات والمهام الفرعية الشائعة في أداة تقييم مهارات القراءة للمراحل الأولى. ثم تناقش المعلومات التي يجب جمعها خلال التقييم، وكيفية اشتقاق بقية المتغيرات المطلوبة من المتغيرات الأولية التي جُمعت. لاحظ أن **الملحق L** من الدليل هو نموذج لكتيب الرموز للمتغيرات بمجموعة البيانات في أداة تقييم مهارات القراءة للمراحل الأولى.

ونجد أن أسماء المتغيرات الخاصة بأداة تقييم مهارات القراءة للمراحل الأولى تتخذ الشكل التالي:

<السابقة > _ <الأساسية > <لاحقة>

أمثلة:

e_letter_sound1
انجليزية_الحرف_الصوت1
e_letter_sound2
انجليزية_الحرف_الصوت2
e_letter_sound_time_remain

انجليزية _الحرف_الصوت_الزمن_البقاء

للحفاظ على الاتساق الداخلي وعبر استطلاعات أداة تقييم مهارات القراءة للمراحل الأولى، من المهم تسمية متغيرات المهمات الفرعية بنفس الأسماء، ويقدم **الملحق 27** قائمة بأسماء المتغيرات بالمهام الفرعية الخاصة بأداة تقييم مهارات القراءة للمراحل الأولى. فضلاً عن أسماء المتغيرات للنتائج الموقوتة (إذا ضُبِط للمهام الفرعية توقيت).

الملحق 27. مسميات المتغيرات بالمهام الفرعية لأداة "تقييم مهارات القراءة للمراحل الأولى" وأسماء متغيرات النتائج الموقوتة

اسم المهمة الفرعية الموقوتة	اسم المتغير الموقوت بالمهمة الفرعية	اسم المهمة الفرعية	اسم المتغير بالمهمة الفرعية
أسماء الحروف الصحيحة في الدقيقة	clpm	اتعرف الحروف (أسماء)	الحرف
أصوات الحروف الصحيحة في الدقيقة	clspm	تعرف الحروف (أصوات)	صوت الحرف
الكلمات الصحيحة في الدقيقة	cwpm	قراءة الكلمات المألوفة	الكلمة المألوفة
	cnonwpm	قراءة الكلمات المزيفة	الكلمة المخترعة
طلاقة القراءة الجهرية	orf	طلاقة القراءة الجهرية	القراءة الجهرية
		الاستيعاب القرائي	الاستيعاب القرائي
		الاستماع والفهم	الاستماع والفهم
المقاطع الصوتية الصحيحة في الدقيقة	cspm	التعرف على المقاطع (الأصوات	المقاطع الصوتية
		المفردات الجهرية	المفردات الجهرية
		المفردات	المفردات
		المتاهة	المتاهة
		الإملاء	الإملاء

1.2.10 <السوابق> _

إذا قُيِّم الطالب في أكثر من لغة واحدة، فمن المهم أن يميز بين اللغات عن طريق سابقة. وتحتاج اللغات الثانوية إلى البادئة، مثل E_ للغة الإنكليزية أو F_ للغة الفرنسية.

ملاحظة حول الفقرات المتعددة: في العديد من الدراسات التجريبية، هناك أكثر من نسخة من نفس المهمة الفرعية. فعلى سبيل المثال، قد يكون هناك ثلاثة إصدارات مختلفة من الفقرات التي تقيس مدى طلاقة القراءة الجهرية، وكذلك ثلاث مجموعات مختلفة من أسئلة الفهم. وفي هذه الحالات، تكون السوابق هي الحرف الذي يرمز للغة، ورقم المهمة الفرعية المختلفة. وهكذا فبالنسبة للغة الإنجليزية سيكون اسم المتغير e1_oral_read<suffix>, e2_oral_read<suffix>, e3_oral_read<suffix>؛ للمساعدة في معرفة فقرة القراءة التي يشير إليها المتغير.

2.2.10 <اللواحق> _

وستؤدي المهام الفرعية الخاصة بأداة تقييم مهارات القراءة للمراحل الأولى إلى جمع البيانات لكل عنصر أصاب فيه طالب، أو أخطأ، أو لم يحاول لأن الوقت كان قد نفد. بمعنى، أنه في المهمة الفرعية الخاصة بتعرف الحروف (الأصوات)، على سبيل المثال، فسيكون للبيانات متغير لكل بند من بنود الاختبار. من هذه المعلومات، يمكن حساب كل متغيرات النتيجة غير الموقوتة. وتشير اللواحق إلى رقم البند بالمهمة الفرعية، وملخص النتيجة.

وستكون اللاحقة هي رقم البند في المهمة الفرعية، أو أية متغيرات إضافية مرتبطة بهذه المهمة الفرعية (مثل: _auto_stop، _attempted، _time_remain)، ويمكن أن تكون اللاحقة هي رقم البند الموجود في المهمة الفرعية. فعلى سبيل المثال، إذا كان هناك خمسة بنود في قسم الاستيعاب القرائي باللغة الإنجليزية، فإن أسماء المتغيرات ستكون e1_read_comp1، e1_read_comp2، e1_read_comp3، e1_read_comp4، e1_read_comp5، e1_read_comp_attempted.

لاحظ أن أسماء المتغيرات هذه لا تحتوي على شرطة سفلية "_" بين الكلمة الأساسية، ورقم اللاحقة 1-5. لذا فإن المتغيرات لن تكون e_read_comp_1, e_read_comp_2, e_read_comp_3, e_read_comp_4, e_read_comp_5. أما المتغيرات الغير بندية فنجد شرطة سفلية "_" بين الكلمة الأساسية واللاحقة. فمتغيرات أداة تقييم مهارات القراءة للمراحل الأولى تسمى e_read_comp_attempted and e_read_comp_score.

الملحق 28 يحتوي على بعض الأمثلة على كيفية تسمية المتغيرات الخاصة بأداة تقييم مهارات القراءة للمراحل الأولى، على أساس اللغة وعدد الأقسام المتكررة داخل الأداة.

الملحق 28 مسميات اللواحق لمتغيرات البنود والنتائج

القيم الممكنة	تسمية لاحق المتغير	اللاحقة
0 "غير صحيح" 1 "صحيح" ناقص "لم يتم سؤاله/لم يحاول"	رقم البند	#1
0-عدد البنود في المهمة الفرعية	صافي المجموع	_score
0-عدد البنود في المهمة الفرعية	إجمالي البنود التي حاول الطلاب الإجابة عنها	_attempted
0-100	نسبة الصواب	__score_pcnt
0 " النتيجة أكبر من صفر" 1 "النتيجة تساوي صفر"	مؤشر النتيجة صفر	_score_zero
0-100	نسبة الصواب في المحاولة	_attempted_pcnt

يُحسب موجز المتغيرات التالية فيما بعد:

- **score_**: مجموع الإجابات الصحيحة على البنود (التي تم ترميزها برقم 1)
- **_ attempted**: عدد الإجابات الصحيحة والغير صحيحة والتي رُمزت برقم 1 أو 0
- **_score_pcnt**: نتيجة _المهام الفرعية مقسومة على رقم البنود الممكنة في المهمة الفرعية.
- **_score_zero**. : نعم (يتم تسجيلها كرقم 1) إذا كانت نتيجة الطالب صفر، غير ذلك، لا (يتم تسجيلها كرقم 0)
- **_attempted_pcnt**.: _ النتيجة مقسومة على_ عدد البنود التي حاول الطالب أن يجيب عنها

3-10 المهام الفرعية الموقوتة

صُممت المهام الفرعية الموقوتة في أداة تقييم مهارات القراءة للمراحل الأولى؛ لتكون محسوبة بمعدل الدقيقة، ويجب أن تُكود الإجابات (مثل: الحروف والكلمات المنفردة)؛ لتكون صحيحة

أو *غير صحيحة* أو *بدون إجابة* أو *لم يجب*. ويجب أن يفرق من يقوِّم ميدانيًّا بين *غير صحيح* (المكودة برقم0) و*بدون إجابة*، حيث لن يكون من الممكن تحليل البنود التي حاول الطلاب الإجابة عليها إذا لم تتم هذه التفرقة.

وبالإضافة إلى إجابات البنود يجب أن تتضمن المتغيرات الملخصة التالية في البيانات الأولية للمهام الفرعية الموقوتة.

1. **الوقت المتبقي من المهمة الفرعية (Subtask)**: هو الوقت المتبقي من المهمة الفرعية إذا أنهى الطالب المهمة قبل أن ينتهي الزمن المخصص لها. هذا المتغير المختصر سيُستخدم لحساب معدل الدقيقة، ويُسجل بالثواني. وعادة ما يكون الزمن المحدد للانتهاء من المهمة الفرعية الموقوتة 60 ثانية مع حذف الوقت المستغرق للانتهاء من المهمة الفرعية.

2. **الإيقاف الأوتوماتيكي للمهمة الفرعية:** من أجل التقدم بفاعلية في مراحل هذا التقييم، ولتلافي توقف الطلاب لمدة طويلة محاولين الإجابة على أسئلة من الواضح عدم معرفتهم لإجاباتها، يتوقف التقييم بعد أن يفشل الطالب في الإجابة على البنود القليلة الأولى_ عادة أول 10 بنود (أو أقل). يتلقى الطالب الذي يفشل في الإجابة قبل التوقف الأوتوماتيكي للتقييم الرمز 1 لهذه المهمة الفرعية، وبما أن 1 يعني نعم يتوقف الطالب أوتوماتيكيا. وهذه النتيجة تكون للمهمة الفرعية ككل، ولا تُسجل على مستوى البند الواحد.

ومن أجل عمل متغيرات مختصرة يُوضع 1 على الإجابات الصحيحة للبنود المنفردة، وصفر على الإجابات الغير صحيحة، و*ناقصة* على الطلاب الذين لم يجيبوا.

وعادة ما نشير لمعدل الدقيقة بمعدل الطلاقة. وغالبا ما تُدار المهام الفرعية الموقوتة لمدة تزيد على ال60 ثانية. بحيث إن الطلاب الذين انتهوا من الإجابة على البنود الخاصة بالمهمة الفرعية، أو قراءة الفقرة قبل انتهاء الوقت المحدد سيحصلون على درجة طلاقة أعلى من نتيجتهم الأولية. ووحدة القياس النهائية هي إما الحروف الصحيحة، أو الكلمات الصحيحة في الدقيقة الواحدة.

ويتم حساب المعدل لكل دقيقة باستخدام المعادلة التالية:

$$\text{المهمة_في_ الدقيقة ()} = \frac{\text{نتيجة المهمة الفرعية e)}}{\text{الزمن المحدد والمتبقي للمهمة الفرعية ()}} \times 60$$

4-10 المهام الفرعية الغير موقوتة:

كما هو الحال مع المهام الفرعية الموقوتة، تحتاج هذه الإجابات على البنود إلى أن تُكود؛ لتكون *صحيحة* أو *غير صحيحة* أو *بدون إجابة/لم يجب*. ومن أجل عمل متغيرات مختصرة تُوضع 1 للإجابات الصحيحة، وصفر للإجابات الغير صحيحة، وناقصة *لبدون* إجابة/لم يجب.

ملحوظة على نشاط الاستيعاب القرائي:

وكما هو الحال في الممارسات القياسية، إذا احتسب الاستيعاب القرائي على نفس الفقرة التي تُقيم القراءة الجهرية بناءً عليها، نُقيم الطلاب على عدد أسئلة الاستيعاب القرائي التي أجابوا عليها عن الفقرة التي تمكنوا من قراءتها.

فعلى سبيل المثال، إذا كانت هناك خمسة أسئلة تقيم الاستيعاب القرائي تعتمد على قراءة الفقرة حتى الكلمات التاسعة، والسابعة عشر، والثامنة والعشرون، والثانية والأربعون، والخامسة والخمسون بالترتيب وقرأ الطالب الفقرة حتى الكلمة الثالثة والثلاثون إذا فسيقيَّم الطالب على أساس أول ثلاثة أسئلة فقط. وستصحح الأسئلة وتوضع علامات: *صحيح وغير صحيح وبدون إجابة*، أما آخر سؤالين فستوضع عليهم علامة لم *يطرح السؤال*.

وبالرغم من أن المعيار القياسي قد يتغير بتغير السياق إلا أنه بوجه عام يمكننا اعتبار الطالب يقرأ بطلاقة ويفهم ما قرأه إذا قرأ فقرة كاملة، وتمكن من الإجابة على 80 بالمئة أو أكثر من أسئلة الاستيعاب القرائي بشكل صحيح. ولحساب هذا، صُمم متغير مختصر جديد، وهو القراءة_الفهم_نتيجة_نسبة 80 بالمئة (**read_comp_score_pcnt80**)، والذي يُعد صحيحًا (ويوضع عليه كود1) إذا حصل على نتيجة تعادل 80 بالمئة، أو أكثر في الاستيعاب القرائي، وغير ذلك يعد غير صحيح (ويوضع عليه كود صفر).

5-10 المساواة الإحصائية

المساواة هي إجراء إحصائي يستخدم لتحويل النتائج من نماذج متعددة من الاختبارات لنفس المقياس المشترك، وتقوم عملية التحويل هذه بضبط وتعديل أية صعوبات تحدث بسبب الاختلاف بين النماذج المختلفة بحيث يمكن لنتيجة أحد النماذج أن ترتبط بما يوازيها في نموذج آخر. ونتيجة لذلك تتيح المساواة تقدير النتيجة التي كان سيحصل عليها الأطفال إذا كانوا قد خضعوا لشكل آخر من الاختبارات (هولاند ودورانز 2006، كولين وبرينان 2004).

وقد أظهرت الأبحاث التي أجريت على المساواة الإحصائية في العينات الصغيرة، والمناسبة تقريبًا كل أنواع المساواة الخاصة بأداة تقييم مهارات القراءة للمراحل الأولى أنه عندما يكون هناك فروقات حقيقية بين نتائج المهام الفرعية على نموذجين من نماذج الاختبارات أقل من 10/1 من الانحراف المعياري الذي يمكن أن يتعدى خطأ المساواة فيه إلى تحيز عدم المساواة. (هانسون و زنج وكولتون 1994، سكاجز 2005). ولهذا لا ينصح باستخدام المساواة في العينات الصغيرة عندما لا يزيد الفرق في النتائج عن 10/1 من الانحراف المعياري.

وعندما تكون المساواة ضرورية، هناك بعض الاعتبارات المهمة التي يجب أن تؤخذ في الاعتبار.

أول نقطة هي أن مطوري هذه الأداة يجب أن يضعوا في اعتبارهم ويدركوا مدى ملاءمة المهام الفرعية للمساواة، وهناك أربع تقنيات يمكن استخدامها للمساواة الإحصائية وهي *المساواة عن طريق عنصر مشترك، والمساواة عن طريق شخص مشترك ونظرية المساواة عن طريق الاختبار التقليدي، ونظرية المساواة عن طريق الإجابات على البنود ويأتي التفصيل على النحو التالي:*

المساواة عن طريق عنصر مشترك: تُستخدم هذه التقنية عندما تكون الأدوات أو المهام الفرعية تحتوي على بعض العناصر المشتركة في كل نماذج الاختبار**.** وهذه العناصر المشتركة (التي تُعرف باسم عناصر الربط) يجب أن تمثل ما لا يقل عن 20٪ إلى 25٪ من مجموع عناصر التقييم ، وتمثل نسخة مصغرة من التقييم الشامل من حيث الصعوبة والاختلاف.[27] ، ومن المهم أيضا التأكد من أن عناصر الربط تحتفظ بمكانها في الاختبارات المختلفة ، فعلى سبيل المثال إذا كان أحد عناصر الربط في نموذج الاختبار أ هو العنصر الخامس ، فيجب أن يكون هو العنصر الخامس أيضًا في نموذج الاختبار ب. أما ما يتبقى من العناصر الأخرى (بخلاف عناصر الربط) ، فإمَّا يتغير ترتيبها عن الترتيب الذي يوجد في الأداة الأصلية ، أو تستبدل بعناصر أخرى جديدة تمامًا.

والمبدأ الأساسي وراء المساواة عن طريق العنصر المشترك هو أن درجة صعوبة عنصر الربط يجب أن تكون متطابقة في كل نماذج التقييم. لذلك تُعدل النتائج؛ لتقوم بحساب صعوبة الاختبار ككل بناءً على الدرجة الفرعية الخاصة بعناصر الربط. وهناك العديد من الطرق للقيام بعملية المساواة عن طريق العنصر المشترك (بما فيها المساواة المتسلسلة، وما بعد التقسيم الطبقي) لكن اتساع وعمق المعلومات اللازمة لتغطية هذه الموضوعات خارج نطاق هذا الدليل.

وأخيرًا، فالمساواة عن طريق العنصر المشترك الأفضل للمهام الفرعية التي تحتوي على عناصر كافية (الحد الأدنى الموصي به 20-25 عنصر)، بسبب انخفاض احتمالية الخطأ الإحصائي (على فرض صغر حجم العينة).

المساواة عن طريق الأشخاص المشتركين: وتعرف باسم تصميم أحادي المجموعة ، أو تصميم مجموعة المعادل العشوائي، وتُستخدم هذه الطريقة عندما تُصمم أدوات أو مهام فرعية ؛ لقياس بنيات مماثلة ، لكن لا تحتوي على عناصر الربط ، وهذا هو حاليا النوع الأكثر شيوعا من المساواة الذي تم إجراؤه على أداة تقييم مهارات القراءة للمراحل الدراسية الأولى ؛ لأنها لا تتطلب الإلمام بإجراءات عملية المساواة في مرحلة تصميم الدليل ، ومن أجل تطبيق هذا النهج تُجرب عدة نماذج من أداة تقييم مهارات القراءة للصفوف الأولى على عينة من الطلاب بحيث يُعرض على كل منهم كل النماذج. والمبدأ الأساسي هو أن الاختلافات في درجات الاختبارات في نماذج التقييم المختلفة يمكن أن ينظر إليها كاختلافات في صعوبة الاختبار على العكس من قدرة الطالب، بما أن نفس الطلاب يخضعون لكل النماذج، وهذا النهج ضروري لفقرة طلاقة القراءة الجهرية لأداة تقييم مهارات القراءة للصفوف الدراسية الأولى" بما أنه ليس من الممكن إنشاء عناصر ربط لهذه المهمة الفرعية، وبما أن المعلومات على مستوى كل عنصر غير ذات صلة، وهو شرط أساسي لعملية المساواة عن طريق نظرية مساواة الإجابات كما هو مبين أدناه.

الخطوات اللازمة لعملية المساواة التجريبية عن طريق الأشخاص المشتركين

من أجل تحقيق أقصى قدر من الكفاءة، وتحقيق أقصى فائدة؛ لتصميم المساواة عن طريق الأشخاص المشتركين، لذا يجب أن يُستخدم السيناريو التالي خلال المرحلة التجريبية حيث يوجد ما يكفي من الوقت (والبصيرة)؛ لإنشاء عدد كبير من الفقرات المتوازية، والتمويل الكافي لإجراء التجربة مع ما لا يقل عن 500 طالب وطالبة.[28]

في هذا السيناريو، نقترح أن يقوم مطورو أداة تقييم مهارات القراءة بالصفوف الأولى بإنشاء 10فقرات للاستيعاب القرائي وخمسة أسئلة على كل منها (10مجموعات)، باستخدام حكم الخبراء؛ لتكون موازية بقدر الإمكان لبداية العملية، وسُيعرض على كل عينة من الطلاب 3فقرات منفصلة (وأسئلة الفهم المصاحبة لها). ويمكن للتصميم (نظريًّا) أن يبدو كما هو مبين في **الملحق 29** (مع 10 نماذج ل 3 مجموعات، و15 سؤالا، لكل منهم).

الملحق 29 تصميم عينة الموازنة

عدد الطلاب	المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة	نماذج الاختبار التجريبية بالحروف
50	1	2	4	أ
50	2	3	5	ب
50	3	4	6	ج
50	4	5	7	د
50	5	6	8	ه
50	6	7	9	و
50	7	8	10	ز
50	8	9	1	ح
50	9	10	2	ط
50	10	1	3	ك
500				

[27] هناك بعض الخلاف على النسبة المطلوبة تحديدًا لعناصر الربط، لكن نسبة 20 إلى 25 بالمئة عادة ما تكون هي المعيار.

[28] يمكن لهذه التجربة المنفردة أن تحل محل التجارب العديدة على 150-200 طالب وطالبة (وهو شيء غير مألوف في العمل التنموي). إنها ببساطة مسألة التكاليف مقابل الفوائد وقيمة الحصول على 10 فقرات تم تقييمها

في هذا التصميم، تظهر كل قطعة في كل مجموعة (الأولى، والثانية، والثالثة)، وكل قطعة تظهر مع ست قطع أخرى، ويتغير ترتيب القطع من أجل تقليل آثار الترتيب، ويتطلب هذا النهج عينة من 500 طالب وطالبة تُقسم عشوائيًّا إلى 10 عينات فرعية يتلقى كل منها نموذج من 10 نماذج اختبار). وبالتالي، فإنه من الممكن الحصول على تدابير قوية للصعوبة النسبية لكل بند ومجموعة، ثم تتطابق مجموعات من أجل الحصول على أقصى قدر من القابلية للمقارنة لما قبل وبعد الاختبار، والثقة في أن التغييرات في الدرجات والنتائج على مستوى العينة سيكون مجديًّا.

المساواة عن طريق نظرية الاختبار التقليدي: تقوم نماذج المساواة التي تعتمد على نظرية الاختبار التقليدي بإنشاء العلاقات بين الدرجات الكلية للنماذج المختلفة. وهذا هو النهج الأكثر "تقليدية" لمساواة الاختبارات، وهو أكثر المناهج المألوفة للمساواة مع العينات الصغيرة.، وتتضمن مناهج المساواة عن طريق نظرية الاختبار التقليدي المساواة المتوسطة والخطية والقوسية والمساواة عن طريق النسبة المئوية، وهذا الدليل لا يقدم تفسيرات متعمقة لكل نهج، ولكن هناك توصيات إضافية بشأن هذه المناهج في **الملحق م**.

والمساواة عن طريق نظرية الاختبار التقليدي مفيدة للبيانات الخطية، وللاستخدام في العينات الصغيرة، لكنها لا يُنصح بها للمهام الفرعية ذات العناصر القليلة نسبيًّا (أقل من 10). وبالنسبة للمهام الفرعية ذات العناصر الأقل من 10-25 عنصر، قد يكون من الممكن استخدام منهج المساواة القبلية عن طريق نظرية الاختبار التقليدي عن طريق تجريب عدة نماذج من الاختبار وُضعت حديثًا جنبًا إلى جنب مع النماذج الأساسية ومقارنة الإحصاءات على مستوى البند عبر النماذج، وفي سياق المهام الفرعية لأداة تقييم مهارات القراءة للصفوف الدراسية الأولى"، نجد أن هذا هو النهج الأكثر إفادة لمساواة طلاقة القراءة الجهرية.

نظرية المساواة عن طريق الإجابة على العناصر: وهي تقوم على مبدأ إقامة علاقات المساواة من خلال النماذج التي تربط المتغيرات الملحوظة والخفية، وهذا النهج لديه ميزة استخدام نفس النموذج الرياضي لخصائص الأشخاص وخصائص الأدوات، وكذلك فإن هذا النوع من المساواة يتميز بكونه أكثر توافقًا مع طبيعة الاختبار مع توفير فرص لمساواة المهام الفرعية مع عدد قليل من البنود. لكن بالرغم من ذلك، فالمساواة عن طريق الإجابة على العناصر من الناحية الإجرائية ولمفاهيمي معقدة، وتتطلب عينات أكبر بكثير من تلك التي تطلبها المساواة عن طريق نظرية الاختبار التقليدي (باستثناء نموذج راش الذي يتطلب نفس حجم عينة الاختبار التقليدي، وهو ما يقرب من 100-150 من المشاركين).

لذلك، فالمساواة عن طريق الإجابة على العناصر مفيدة جدًا لمرحلة المساواة، وهذا يعنى أن المساواة البعدية عن طريق البيانات التشغيلية، أو بيانات المسح الكاملة-بالمقارنة بالمساواة القبلية التي تتم عن طريق البيانات التجريبية، عندما تكون الخبرة الفنية كافية والقدرات متاحة. وفي الجزء الأكبر من عمل أداة تقييم مهارات القراءة للصفوف الدراسية الأولى، يمكن للمساواة عن طريق الإجابة على العناصر (عن طريق نماذج راش) أن تكون مفيدة بشكل خاص للمساواة القبلية في المهام الفرعية ذات العناصر القليلة، وهو قصور في نهج المساواة عن طريق نظرية الاختبار التقليدي طالما أن هذه المهام الفرعية لديها بيانات مفيدة على مستوى العناصر. وتشمل هذه المهام الفرعية الاستيعاب القرائي، والاستيعاب السمعي، والإملاء، والمفردات، والمتاهة.

وقد ذكرت المجالات الخاضعة لمزيد من الدراسة فيما يخص إجراءات اختبار المساواة في الملحق م.

6-10 جعل بيانات أداة "تقييم مهارات القراءة للصفوف الأولى" متاحة للجمهور

ومن المتوقع أن تقوم الوكالة الأمريكية للتنمية بجعل ملفات الاستخدام العام التي تحتوي على بيانات تقييم مهارات القراءة للمراحل الأولى متاحة للجمهور من خلال التحليل الثانوي لنتائج تتبع بوابة التعليم، ومكتبة بيانات التنمية وبشكل متزايد عن طريق منصة إحصائيات التعليم الخاصة بالبنك الدولي، وقد يختار الممولون الآخرون اتباع عملية مماثلة لجمع البيانات التي يقومون برعايتها.

وتُنظف ملفات الاستخدام العام والانتهاء منها وإزالة البيانات الشخصية من مجموعات البيانات المعدة للاستخدام العام، وتحتوي مجموعات البيانات هذه على كل المتغيرات المناسبة اللازمة؛ لتحليل صحيح للبيانات، ولكن تُخفى جميع المعلومات التعريفية لحماية هويات الأفراد والمؤسسات. ويعرف تحديث الوكالة الأمريكية للتنمية لتقارير إرشاد البيانات الشخصية التي نُظِفت، وانتهت، وأُزيلت كالتالي:

- **تنظيف البيانات:** فحص المنفذ وتصحيحه لأي أخطاء واضحة أو تناقضات في المحتوى، وكذلك المعلومات الناقصة، الخ

- **البيانات النهائية:** وتشمل مجموعات البيانات أي مؤشرات مشتقة أو ثانوية يستخدمها المنفذ لحساب قيم المؤشرات التي ضُمِّنت في التقارير. ويكون المنفذ قد انتهى من معالجة مجموعة البيانات، ولا يتوقع وجود تغييرات أخرى.
- **إزالة المعلومات الشخصية من البيانات:** وقد اتخذت خطوات لحماية الخصوصية، وإخفاء هوية الأفراد والمدارس المرتبطة بهذا التقييم. عملت المنظمة المنفذة مع مجلس المراجعة المؤسسية لضمان أن المشاركين في التقييم تمت حمايتهم كما ينبغي.

ويتضمن الملحق N مزيدا من التوصيات والمبادئ التوجيهية بشأن كيفية تنظيف، ووضع اللمسات النهائية، وإزالة المعلومات الشخصية من البيانات التي يمكن توزيعها على نطاق واسع على جميع الجماهير. وبمجرد أن تنشئ ملف الاستخدام العام لمجموعة البيانات، تكون هناك حاجة إلى معلومات موثقة جيدًا؛ لمساعدة المستخدم العام على الاطلاع على البيانات. وبالنسبة لعمليات مسح أداة تقييم مهارات القراءة للصفوف الأولى التي أُجريت بتمويل من الوكالة الأميركية، تتوفر المعلومات التالية للمستخدمين:

- معلومات أساسية، مثل: تحديد السكان المستهدفين بما في ذلك مصدر الإطار الذي يختار من خلاله العينة، ووصف تصميم العينة، ووقت جمع البيانات.
- جميع الوثائق ذات الصلة، بما في ذلك الاستبيانات والأدوات المستخدمة في التقييم. (إذا أُجري تقييم مهارات القراءة للصفوف الأولى لتقييم أثر البرنامج، ولا يتم الإفراج عن الاستبيانات إلا بعد اكتمال البرنامج، وذلك بما لا يضر المواد التي يمكن أن تستخدم في دراسات تقييم مهارات القراءة للمراحل الأولى في المستقبل.)
- تقرير تحليل البيانات الخطية المقدم إلى الوكالة والذي تم الموافقة عليه.

ويجب أن يتعرف المستفيدون على أهمية توثيق أسماء وأوصاف المتغيرات الأساسية جنبًا إلى جنب مع الإعدادات اللازمة لتحليل البيانات كما ينبغي. وتُضمن مبادئ توجيهية محددة بشأن كيفية توثيق أسماء وتوصيفات المتغيرات لأغراض تحليل البيانات في الملحق N.

في الوقت الذي كان يتم فيه كتابة هذا الإصدار من الدليل، كان هناك عدة مستودعات بيانات خاصة بالوكالة الأمريكية لاتزال في مرحلة التطوير، إلا أنه كان من الضروري أن تكون بيانات تقييم القراءة للصفوف الأولى متوفرة للجمهور. لذا ينصح المنفذون الذين جمعوا بيانات تقييم مهارات القراءة للصفوف الأولى أن:

- ينشروا ملف الاستخدام العام في موقع يسهل الوصول إليه عبر الإنترنت، مصحوبًا بوثائق يمكن تحديد مواقعها بسهولة (على سبيل المثال، تجمع جميع العناصر الموجودة في ملف مضغوط واحد، أو يحتوي الموقع على روابط لهذه الوثائق).
- ينشئ ملف الاستخدام العام باستخدام ملف بيانات لا يخضع لحقوق الملكية، وعند الإمكان، ملف خاضع لحقوق الملكية جاهز للتحليل (أي، مع تصميم المسح المعقد المحدد مسبقا).
- ولإنشاء ملف غير خاضع لحقوق الملكية يمكن إنشاء ملف نصي من نوعية (CSV) أو المتغيرات المفصولة بفواصل.
- ولإنشاء ملف خاضع لحقوق الملكية يمكن إنشاء ملف Stata .dta أو ملف SPSS.sav (مع ملف SPSS.csaplan).

11 تحليل البيانات وإعداد التقارير

يقدم هذا الجزء من الأداة عرض موجز لأنواع تحليل البيانات المتعلقة بمختلف تصميمات الأبحاث، وكذلك المكونات التي يلزم توافرها في تقارير تقييم القراءة في المراحل الأولى.

عند تحليل بيانات تقييم القراءة في الصفوف الأولى، يجب أن يستخدم الباحثون إحصاءات توصيفية و/أو استنباطية لوصف البيانات، وفحص الأنماط، واستخلاص الاستنتاجات. إلا أنه من الضروري فهم الاختلافات بين هذين النوعين من الإحصائيات، وكذلك الغرض من كل منها وقيمته.

11-1 الإحصائيات التوصيفية (غير الاستنباطية)

تستخدم الإحصائيات التوصيفية (أو غير الاستنباطية) ؛ لوصف البيانات وتلخيصها – غالباً بهدف رؤية أي الأنماط قد تظهر، ولا تسمح الإحصائيات التوصيفية باستخلاص الاستنتاجات فيما وراء البيانات، كما لا يمكن اختبار الافتراضات البحثية ، ويتمثل الغرض الأساسي من التحليل الوصفي في تقديم البيانات بطريقة مفيدة تسمح بسهولة التفسير (على العكس من مجرد تقديم بيانات خام) ، وتعد التدابير الأكثر شيوعاً في التحليلات الوصفية هي التكرار، وقياس الاتجاهات المركزية (مثل الوسائل والسبل) ، ووسائل الانتشار (مثل : الانحرافات القياسية ، والنطاقات المحدودة).

علاوة على ذلك، وكما يشير المصطلح ، فإن الإحصائيات التوصيفية تستخدم فقط لوصف عينات البيانات. فغالبًا أثناء تقييم القراءة في الصفوف المبكرة، تُختار العينات بحيث تكون ممثلة لأكبر عدد من السكان. وفي تلك الحالات، تعتمد مراحل التكرار، والوسائل، وغيرها المستخدمة على البيانات الموزونة ، ومن ثم تصبح فعلياً إحصائيات استنباطية. فتُعد التقارير عن الإحصائيات التوصيفية فقط بالنسبة للدراسات المصممة لعدم استخلاص نتائج ما وراء العينات، أو كتكرار غير موزون، أو وسائل غير موزونة لبيانات الاستبانات المعقدة.

وأخيراً، في غياب الإحصائيات غير الاستنباطية، من الضروري وصف العينة بشكل كامل وفقاً لمستوى التصنيف الذي سيتم تحليله وإعداد التقرير بشأنه. على سبيل المثال، إذا كانت درجات الطلاب في التقرير ستُصنف وفقاً للغة والدرجة، وتتضمن إحصائيات العينة التوصيفية هذه المستويات من التصنيف.

ومن أمثلة الإحصائيات التوصيفية المفيدة في إعداد تقارير تحليل القراءة في المراحل الأولى هي التكرار ، والخصائص السكانية الأساسية للعينة، وكذلك الوسائل الموزونة في المهام الفرعية لجميع مستويات التصنيف.

2-11 الإحصائيات الاستنباطية

تتيح الإحصائيات الاستنباطية للمقيمين الوصول إلى استنتاجات بشأن السكان بالكامل بناء على عينة من أولئك السكان، والوصول إلى استنتاجات بشأن الافتراضات الخاصة بالمقاييس السكانية ومقارنة فئتين مختلفتين من السكان (على سبيل المثال مجموعتي العلاج والرقابة). وتعد الإحصائيات الاستنباطية مهمة لكل من تقييم الأثر ، والمقتطفات لتقييم القراءة في الصفوف الأولى ؛ لإعداد التصريحات الخاصة بالتعليم لبلد أو إقليم بأكمله وفقاً لعينة الطلاب أو المدارس في ذلك البلد أو الإقليم. ويعتمد نوع الإحصائيات الاستنباطية اللازمة لطالب ما على تصميم التقييم، على النحو التالي:

- **التصميمات التجريبية (أو التجارب العشوائية المحكمة).** يمكن مقارنة مجموعتين باستخدام إحصائيات كلية لمجموعات زوجية ؛ لتحديد ما إذا كانت النتائج النهائية أعلى للمشاركين في مجموعات المعالجة عن المشاركين في مجموعة التحكم. إذا كان التوزيع العشوائي ناجحًا وحجم العينة كبير بما يكفي، ولذلك ليس من الضروري الوضع في الاعتبار الاختلافات بين مجموعات المعالجة ، والتحكم فيما يتعلق بالعوامل الديمغرافية (السكانية) أو درجات الخط الأساسي، حيث إن كلا المجموعتين ستكونان متطابقتان بسبب عملية التوزيع العشوائي.
- **تصميمات شبه تجريبية باستخدام التصميمات الطولية (تتبع الطلاب الأفراد بمرور الوقت) ، أو التصميمات شبه الطولية (تتبع المعلمين ، أو المدارس بمرور الوقت).** إذا ما استوفت كافة الشروط التالية ، يمكن للمقيمين أيضاً استخدام الإحصائيات المجمعة بهدف والحصول على الدرجات ؛ لإظهار التغييرات بمرور الوقت وذلك من خلال

مجموعات المعالجة والمقارنة التي كانت منسقة بشكل جيد عند الخط الأساسي. ولم تكن هناك أيضًا اختلافات كبيرة بين المجموعتين، وتم تتبع المدارس أو الطلاب بمرور الوقت في كلى المجموعتين من خلال التصميم الطولي، وبخلاف ذلك، تعد التحليلات شبه التجريبية (مثل تلك المبينة أعلاه) ضرورية ؛ لقياس النمو بدقة و/أو مدى فعالية البرنامج.

- **التصميمات شبه التجريبية باستخدام تصميمات تقاطعية.** لمقارنة مجموعتين، يجب أن يستخدم المقيمون منهج التحليل شبه التجريبي مثل : الاختلاف في الاختلافات، وعدم استمرارية الانحدار، ومتغيرات الأدوات، أو التحليل الانحداري (يفضلاستخدامه للتوازن بين المتغيرات). الاختلاف في الاختلافات يطرح النتائج الأساسية من النتائج النهائية (لتشكيل درجات الكسب) لكل من مجموعات المعالجة والمقارنة ، وبعدها يطرح درجة الكسب لمجموعة المقارنة من درجة الكسب لمجموعة المعالجة للحصول على تأثير المعالجة. ويهدف هذا المنهج إلى تحديد الاختلافات الأساسية ، ولكنه يعتمد أيضاً على الافتراض بأن تفاصيل الأداء قبل الحد الأساسي كانت متسقة عبر المجموعات. ومن المفيد في الغالب الموازنة بين هذا المنهج وإجراءات تنسيقية (مثل تنسيق درجات الاتجاهات) ؛ لضمان أن المجموعتين متشابهتان إلىحد بعيد . عند استخدام تحليل الانحدار، يجب أن يدرج المقيمون دميات الوقت والعلاج وكذلك دمية تفاعل للوقت x العلاج لتحديد الأثر (وهو بالأساس منهج الانحدار ؛ لتقدير الاختلاف في الاختلافات). كما يمكن تقديم متباينات مستقلة أخرى للتحكم في الخلافات بين المجموعات ، وهو أمر مهم للموازنة على الخط الأساسي ، وكذلك للتغيرات بمرور الوقت. وتقديرات حجم الأثر متضمنة ؛ لتحليل اختلاف الاختلافات (راجع **ملحق ط، معرض ط-1** لمثال على تحليل اختلاف الاختلافات) .

بغض النظر من نوع التصميم المستخدم، على المقيمين اختبار الموازنة بين مجموعات المعالجة ، والرقابة/المقارنة عند الخط الأساسي، والنظر في المتغيرين الأساسيين الناتجين ، والمؤشرات الأساسية لضمان إمكانية المقارنة بين المجموعتين. إن لم تكن المجموعات قابلة للمقارنة، يجب أن ينظر المقيمون في استخدام تقنيات المقاربة، مثل : تنسيق درجات الاتجاهات، وذلك لتحسين قوة التصميم والتحليلات. فإذا كانت البيانات الثانوية متاحة قبل جمع البيانات الأساسية، يمكن استخدام تلك البيانات لاختيار مدارس، أو مدرسين، أو طلاب قابلين للمقارنة لجمع بيانات الخط الأساسي. إلا أنه في حال قصور البيانات الثانوية، يمكن لفريق التقييم النظر في زيادة حجم العينة على الأقل لوحدة المقارنة ؛ لضمان توافر التنسيق الجيد لكل وحدة معالجة، بافتراض توافر الموارد الملاءمة للقيام بذلك. يوجد جدول موازنة في جميع

التقارير التي تستخدم تصميم شبه تجريبي ، ويمكن إدراجه في التقارير التي تستخدم التصميمات التجريبية أيضاً. ويعد هذا الجدول ضرورياً لإظهار الموازنة من خلال التدابير ، وللحد من بعض المخاوف بشأن عدم حيادية أو تحيز الاختبار الاختيار.

علاوة على ذلك، بالنسبة لجميع أنواع التصميم، يجب ضمان الصلاحية الداخلية من خلال فحص آثار القمع، والوفيات، والانتشار، والتاريخ.

3-11 أنواع التحليل الانحداري

يعد الانحدار أكثر السبل شيوعاً ؛ لتحليل العلاقات والقيم المتوقعة للمتغيرات في بيانات تقييم القراءة للمراحل الأولى، ومن الضروري النظر بإيجاز في الأنواع المختلفة للتحليلات الانحدارية التي يمكن إجراؤها. فالتحليل العادي يعمل جيداً مع بيانات التقييم التي توزعت في المعتاد للقيم المتبقية، فعندما يُستخدم متغير مستمر مثل درجات إجادة القراءة الشفهية.

إلا أن العديد من البلدان النامية لديها درجات اختبار تتجمع حول صفر، مما يجعل توزيع الدرجات غير عادل إلى حد كبير. عند التعامل مع مثل تلك البيانات، والانحراف المحتمل أو المنطقي، والذي يسمح للمقيمين دراسة النتائج الرقمية التى تفى بالمقاييس المحلية لقدرة الطالب على القراءة ، أو إذا ما كان الطالب يحصل على صفر في مهمة قراءة فرعية بعينها.

4-11 إعداد التقارير بشأن تحليل البيانات

للغرض من تحليل بيانات تقييم القراءة في المرحلة الأولى جانبين: تحسين فعالية البرامج وتقديم النتائج للعملاء، والمنظمات الشريكة والمسئولين الحكوميين من خلال النشرات الموجزة ، وتقارير البرامج الكاملة باعتبار أن الأهداف والجماهير المختلفة للتقارير ستشكل هيكل ومحتوى تلك التقارير، فإن المبادئ الإرشادية التالية تعد ضرورية ، وهى على النحو التالي :

1. **الأهداف والقيود:** يجب أن يحدد التقرير بوضوح أهداف الدراسة ، وأي قيود عليها.
2. **سهولة اللغة:** يجب تقديم النتائج الأساسية بوضوح، وتركيز، وباستخدام لغة غير فنية.
3. **التوضيح البصري للبيانات:** يجب استخدام التوضيح البصري للبيانات ؛ لتيسير فهم النتائج للجمهور العام، وتعد أشكال التوضيح البصري "مستقلة"، بحيث يكون التوضيح

البصري قابلاً للتفسير دون أن يحتاج الجمهور لقراءة المزيد من النصوص (راجع ملحق س، **عرض س-2** ؛ للاطلاع على مثال على التوضيحات المستخدمة لبيانات التقرير البصري).

4. **التحليلات الوصفية والاستنباطية:** يقدم التقرير الأساسي ملخص لنتائج التحليل الوصفي للبيانات، بما في ذلك توزيع الوسائط والتوزيعات المجمعة ، وتستخدم التحاليل الاستنباطية لتصميم الموازين، والأخطاء القياسية لإعداد التصميم المعقد للمسح (إن كان ملائماً).

5. **توزيع الدرجات:** لكل طالب درجة تقديرية واردة في التقرير، لذا يجب أن يتضمن التوضيح البصري أشكال لتوزيع الدرجات (راجع ملحق س، **عرض س-3**). وذلك يدعم تفسير القارئ للتقدير المقدم، فعلى سبيل المثال، في حين أنه بالإمكان تقديم الدرجة المتوسطة، يوضح التوزيع المرافق كيف أن التقدير يعد "ممثلاً" لدرجات الطالب. ويعد ذلك مهمًا على وجه الخصوص إذا كان توزيع درجات الطالب غير عادي. ففي بعض الحالات، قد يكون من المفيد تقديم متوسط درجات الطالب بالإضافة إلى متوسط الدرجات والتوزيعات.

6. **مستويات التقسيم:** فنتائج تقسيم البيانات وفقاً للجنس، والصف الدراسي، واللغة، وغيرها من المتغيرات المفيدة يجب توصيفها على النحو الملاءم لتصميم البحث.

7. **إعداد التقارير عن جميع النتائج.** كلما أجريت اختبارات مقارنة الوسائط للمقارنة بين مواد المجموعات (مثل : الجنس ، أو اللغة)، أو إجراء تحليلات إحصائية بمتغيرين أو أكثر (كعوامل الارتباط) ؛ لدراسة العلاقة بين المتغيرات المختلفة، فيجب إعداد تقارير عن النتائج حتى وإن لم تكن مهمة إحصائياً.

8. **التأسيس للتقديرات الاستنباطية:** فيجب أن يتاح جميع تقارير التقديرات الاستنباطية (دون الاقتصار على الوسائط، والمتوسط، والمعدل ، والنسب):

    - والدقة – سواء بمعدل ثقة بنسبة 95% للتقديرات ، أو درجة ، أو تقدير للمقارنات بالإضافة إلى الأخطاء القياسية.
    - وأيضًا حجم العينة.

9. **أحجام الآثار:** كلما قدمت نتائج مقارنات البيانات للمجموعات (كالاختلافات بين الخط الأساسي والنهائي، أو بين الصبية والفتيات، أو بين طلاب المدارس الريفية وطلاب المدارس الحضرية)، فيجب الإفادة بحجم الأثر للاختلاف.

10. **المعادلة:** في التصميمات التجريبية وشبه التجريبية، فيجب التأسيس لمعادلة الخطوط الأساسية (What Works Clearinghouse, 2015).

ملاحق مطلوبة

يجب أن يدرج الباحثون تفاصيل المنهجية ، والنتائج للتحليل في الملاحق، والتي قد تكون مطولة إلى حد كبير ، ومكتوبة بلغة فنية. فيجب إدراج الملاحق التالية:

تفاصيل المنهجية، والأساليب، وجمع البيانات

- أهداف الدراسة .
- التصميم .
- أساليب جمع البيانات وعملياتها .
- أدوات جمع البيانات .
- أسلوب ونتائج المعادلة، إذا ما استُخدمت أدوات مختلفة في مر احل الدراسة المختلفة .
- معايير العينة ووقفها (للدراسات المطولة).
- تفاصيل الوزن .
- الحدود .
- نتائج تحليل إمكانية الاعتماد على الاختبار (ألفا كرونباخ ، وعلاقات الترابط بين البند والإجمالي).
- عامل الترابط بين الفئات .

تفاصيل التحليلات غير المتضمنة في التقرير الأساسي:

- وصف العينة .
- تفاصيل التحليلات الوصفية.
- تفاصيل التحليلات ذات المتغيرين ومتعددة المتغيرات.

12 استخدام النتائج لتفعيل الإجراءات

12-1 استراتيجية النشر والتعميم

إن الهدف النهائي لأداة "تقييم مهارات القراءة في المراحل المبكرة" هو تحسين عملية تعليم القراءة وتحقيق النتائج المرجوة منها. ومن المعروف أن إجراء التقييم وحده لايكفي لتحقيق هذا الهدف. ويجب أن يتم استخدام النتائج لإعلام السياسات وممارسة التدريس والدعم التعليمي داخل وخارج المدرسة، واستخدام الموارد لسد الثغرات في النظام. وسواءا كان الحل في صورة تدريب المعلمين ليقوموا باستخدام أساليبا أفضل في التدريس، أو في شراء الكتب لتوزيعها على المدارس والفصول الدراسية، أو الحشد المجتمعي فلا شك أن الحوار والإجراءات التي تتبع "تقييم مهارات القراءة في المراحل المبكرة" لا تقل أهمية عن جمع البيانات.

كان الدافع لخلق أداة تقييم مهارات القراءة في المراحل المبكرة هو تسهيل الوصول للمعلومات اللازمة في الوقت المناسب لإخبارنا بمدى التحسن في عملية التعلم في الدول ذات الدخل المنخفض
دوبيك وجوف (2015)

وينطوي التأكد من أن النتائج تتم ترجمتها إلى إجراءات عملية على نهج متعدد المستويات بدءا من جهود التخطيط والتنفيذ، والتي تم مناقشتها في أماكن أخرى في هذا الدليل (على سبيل المثال، تعريف واضح للغرض من التقييم، وأخذ العينات بعناية من السكان الذي سيتم تقييمهم، بما في ذلك الأدوات التكميلية المناسبة مثل الملاحظات الصفية أو الاستبيانات، وإشراك أصحاب المصلحة في التخطيط والتنفيذ).

ويركز تحليل بيانات ما بعد التنفيذ على الأسئلة البحثية القابلة للتنفيذ من خلال التحليل الدقيق.

ويركز التحليل على تنفيذ ما بعد البيانات على أسئلة بحثية قابلة للتنفيذ من خلال التحليل الدقيق. وأخيرا، يجب إبلاغ النتائج إلى الجماهير المناسبة بطريقة مناسبة ثقافيا وسياقيا من أجل دعم التفاهم والعمل.

وبالرغم أن هذا القسم يركز على وضع استراتيجية للنشر والتعميم، إلا أنه يحث القراء أيضا على أن يتذكروا أن مصداقية _وبالتالي الاستخدام الفعال_ النتائج سيعتمد بشكل كبير على التنفيذ الدقيق للخطوات السابقة في جمع المعلومات.

وقبل التخطيط لأنشطة نشر النتائج، يجب على المنفذين أن يفكروا ويجدوا إجابات لهذه الأسئلة (على الأقل): من؟ ماذا؟ وكيف؟

1. **من** سيستخدم هذه البيانات

ولكل نوع من أنواع الجماهير التي سيحددها السؤال السابق:

2. **ما** هي البيانات التي يجب تقديمها؟
3. **ما هو** نوع المعلومات الأمثل (وما هو الشكل الأفضل لها) التي يستطيع الجمهور أن يتعامل معها؟
4. **كيف** سيستخدمون هذه البيانات؟

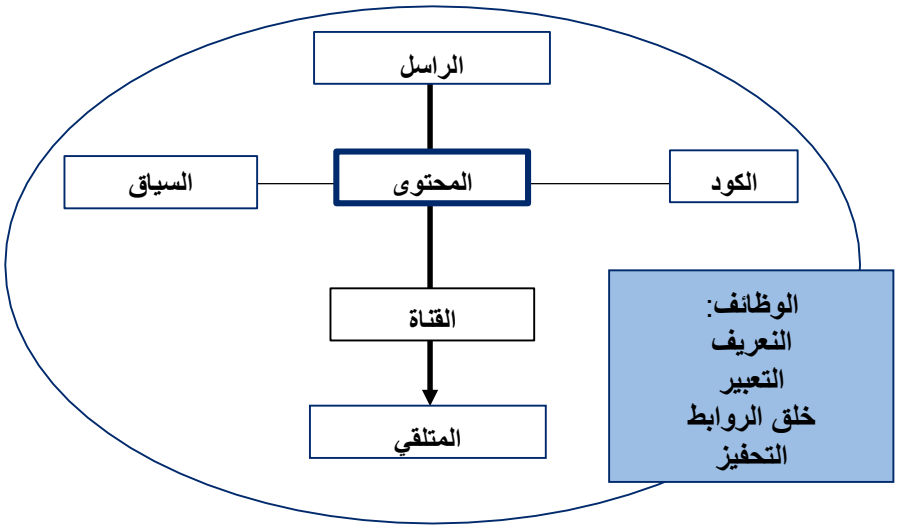
ستساعد هذه الأسئلة على تشكيل الطريقة التي سيتم بها إعداد ومشاركة النتائج بين أصحاب المصلحة المختلفين (على سبيل المثال: المجتمعات المحلية والمسؤولين على مستوى المدارس، والباحثين، والحكومة، ومسؤولي الوزارة والمعلمين وأولياء الأمور). ولاتحتاج منظمة تنفيذ أداة تقييم مهارات القراءة في المراحل المبكرة دائما إلى نشر النتائج على نطاق واسع إذا تمكنت من الوصول إلى الأشخاص المعنيين (في الوقت المناسب وبالرسالة المناسبة). عندما يكون الوقت أو الميزانية محدودة؛ يكون التركيز على الوصول إلى الأفراد المؤثرين ومن لهم القدرة على التواصل مع صانعي القرار والمعلمين.

12-1-1 توصيل النتائج

والملحق 30 هو تذكرة أساسية للعناصر المشاركة في كل أنواع الاتصالات. فمن السهل التركيز على محتوى الرسالة المراد توصيلها، لكن وضوح وتأثير هذه الرسالة يتم تشكيله عن طريق السياق الذي توجد به، "الكود" (اللغة والأسلوب) الذي ستكتب به والقناة أو وسيلة التواصل (مطبوعة أو شفهية أو رقمية) التي يتم إرسالها عن طريقها. وهذا يعني أن الراسل يجب عليه أن يعرف الجمهور ويتم تعريفه بالوسيلة المعتاد عليها الجمهور للوصول للمعلومات ومعالجتها –بما فيها مهارات القراءة والكتابة الأساسية والمامهم بالتصورات المرئية مثل الجداول والرسوم البيانية. فعلى سبيل المثال قد ترحب بعض الجماهير بالمعلومات التقنية المفصلة في حين يفضل البعض الآخر استخدام النتيجة في عرض قصة ترسم صورة لما تعنيه في سياق معين. وكذلك فلاشك أن مكان الوصول لهذه المعلومات ومن ينشرها يمكن أن يغير

أيضا من طريقة تلقي الرسالة وترجمتها.

ملحق 30: إطار التواصل



المصدر: مقتبس من جاكوبسون (1960)

الملحق 31 من *المذكرات التوجيهية لتخطيط وتنفيذ أداة تقييم مهارات القراءة في المراحل المبكرة* (معهد البحوث المثلثية ولجنة الإنقاذ الدولية، 2011 ص82)، يقدم لنا لمحة عامة عن الجماهير المحتملة التي يمكن استهدافها لنشر نتائج تقييم مهارات القراءة للصفوف الدراسية الأولى.

الملحق 31. لمحة عامة عن الجماهير المحتملة

المستوى	الجمهور	الملائمة
الدولي	الجهات المانحة	يمكن للجهات المانحة أن تساعد في دعم جهود الدعوة، والتدخلات التجريبية للقراءة الإبداعية ورفع مستوى أفضل الممارسات. ولاشك أن دعم الجهات المانحة بالغ الأهمية، بالنظر إلى محدودية موارد وزارات التعليم في العديد من الدول النامية التي يمكن إنفاقها في مجال الإبداع والابتكار. ويمكن أن تقوم الجهات المانحة بجمع نتائج تقييم مهارات القراءة للصفوف الدراسية الأولى وخاصة عندما لاتستطيع أو لاتريد الحكومات أن تقوم بهذا الدور.
	الأكاديميين أو الممارسين	غالبا ما يهتم الأكاديميون وغيرهم من الممارسون بنتائج المسح الخاص بتقييم مهارات القراءة للصفوف الدراسية الأولى لأنها تقدم لهم معلومات قيمة عن إنجازات التعلم والقضايا التربوية والتعليمية في العديد من السياقات بالقدر الكافي وكذلك فهي

		تساعد على تحديد أفضل الممارسات.
المحلي	وزارة التعليم	ويمكن أن تشجع عملية المسح والنتائج الخاصة بتقييم مهارات القراءة للصفوف الدراسية الأولى المسئولين الحكوميين على بذل المزيد من الجهود للتأكيد على القراءة كمهارة أساسية. ويمكن أن تكون نتائج تقييم مهارات القراءة للمراحل الدراسية الأولى حافزا على مجموعة من الإجراءات الحكومية الخاصة بالقراءة للمراحل الدراسية الأولى ويمكن أيضا أن تساهم في زيادة الاهتمام بإدراج تقييم مهارات القراءة للمراحل الدراسية الأولى في التقييمات التعليمية الوطنية. وفي بعض الحالات تكون الحكومة (أو العناصر الرئيسية بها) مقتنعة بالفعل بالحاجة إلى العمل ويقوم تقييم مهارات القراءة في المراحل المبكرة بالتأكيد على هذا الأمر ويزيد من دقة المعرفة القائمة بالفعل، ويسهم كذلك في تشكيل وصياغة التدخلات التالية.
	السلطات المسئولة عن الميزانية	قد تساعد نتائج تقييم مهارات القراءة في المراحل المبكرة على إقناع السلطات الحكومية المسئولة عن تخصيص التمويل بتوجيه المزيد من الموارد العامة لوزارات التعليم من أجل تعزيز القراء ة في المراحل المبكرة، وهذا لن يحدث إلا عندما تكون السلطات المالية مقتنعة بأن الحكومة وشركاؤها لديهم استراتيجية قابلة للتطبيق لمعالجة الوضع الذي يكشفه تقييم مهارات القراءة في المراحل المبكرة. وهكذا يجب أن تكون فكرة أن هناك طرق لتحسين النتيجة جزءا من التواصل مع هذا النوع من السلطات.
	نقابات المعلمين	إن دعم المعلمين للقراءة الصف المبكرة والتدخلات ذات الصلة أمر بالغ الأهمية. ويمكن أن يؤثر التواصل والتعاون مع النقابات على تصورات المعلمين حول تقييم مهارات القراءة في المراحل المبكرة كمنصة لإحداث تغيير إيجابي بدلا من أن تكون وسيلة لانتقاد أداء المعلم. ويمكن أن تساعد نقابات المعلمين على نقل أهمية القراءة في المراحل الأولى، والتوقعات المناسبة لكل صف من الصفوف الدراسية والنتائج الرئيسية للمعلمين وكذلك للجماهير الأخرى
	المجتمع المدني ووسائل الإعلام	ويمكن للمجتمع المدني ووسائل الإعلام أن يساعدا على زيادة الوعي والضغط على صناع القرار (مثل الحكومة)، وفي بعض الحالات، تعزيز الاستدامة
الإقليمي	رؤساء الإدارات والمصالح	إن السلطات التعليمية المحلية والإقليمية جمهور وشريك مهم ، وخصوصا في الحالات التي تكون فيها الخدمات التعليمية لامركزية. ويمكن تحفيز كل من المنافسة والتعاون من خلال نتائج تقييم مهارات القراءة للمراحل الأولى.
المجتمع المحلي	زعماء المجتمع المحلي	يمكن لقادة المجتمع المحلي المساعدة على رفع مستوى الوعي حول القراءة و الممارسات الجيدة بين أفراد المجتمع وخاصة الأباء، وكذلك ممارسة الضغط على السلطات المحلية والمدارس.
	الأباء	للأباء تأثيرا كبيرا على عادات القراءة للأطفال وتوعيتهم. ويمكن للتدخلات التي تستهدف الآباء أن ترفع مستوى الوعي

		حول توقعات القراءة المناسبة لكل فئة عمرية وتشجيع القراءة في المنزل، وزيادة الضغط على المدارس وصانعي السياسات لوضع القراءة في قائمة الأولويات
	المجتمع المدني	ويمكن للمجتمع المدني دعم النشر والأنشطة على مستوى المجتمع المحلي ويساعد أيضا في زيادة المسؤولية على المستوى الشعبي.
المدرسة	مديري المدارس	يحتاج مديرو المدارس أن يكونوا على دراية بأهمية القراءة في المراحل الأولى وكذلك الممارسات الجيدة في تعليم القراءة من أجل توفير أفضل دعم للطلبة والمعلمين، وأولياء الأمور. ويمكن أيضا لمديري المدارس المشاركة في التدخلات
	المعلمين	يعتبر المعلمون جمهورا بالغ الأهمية وهدفا للتدخلات ورفع الوعي بشأن القراءة للمراحل الدراسية الأولى والتوقعات الخاصة بكل فئة عمرية والممارسات الجيدة لتعليم القراءة.

المصدر: معهد البحوث المثلثية ولجنة الإنقاذ الدولية، 2011 ص82 (جدول 6.2)

2.1.12 مناهج النشر

تخص النتائج الخاصة بتقييم مهارات القراءة في المراحل المبكرة صانعي السياسات والمسئولين بالدرجة الأولى وخاصة عند استخدام هذا التقييم لتشخيص النظام الوطني أو عند ربط النتائج بتقييم الأثر لابتكار تعليمي موسع أو قابل للتوسع. وإذا كان صحيحا أن العمليات السياسية يمكن أن "تترجم الإرادة الشعبية في السياسات العامة ووضع القواعد لتقديم الخدمات بكفاءة وبشكل فعال لجميع أفراد المجتمع" (كروش وونكلير 2008 ص3)، إذا فمعالجة القضايا أو التحديات في النظام التعليمي (وهي خدمة هامة دون شك) يبدأ بالإرادة السياسية وهناك افتراض أساسي هو أن البيانات التي تظهر نتائج التعلم في نظام التعليم يمكن أن تستخدم لتحفيز الإرادة السياسية.

"على المدى الطويل، تعتبر زيادة قدرة الدولة على استخدام المعلومات لإبلاغ التعليمات أمرا بالغ الأهمية لتحسين عملية التعلم لنحو 250 مليون طفل في جميع أنحاء العالم من الذين لا يكتسبون المهارات الأساسية."

- اليونسكو، 2014

ومن أفضل الممارسات التحقق من صحة النتائج مع الحكومة (أو غيرها من العملاء / أصحاب المصلحة) قبل نشرها على نطاق أوسع. ويمكن أن يتم ذلك من خلال حدث يستمر لمدة يوم

واحد أو يومين (غالبا ما يسمى **" حوار السياسات "**) يجمع إما مجموعة مختارة من المسؤولين الحكوميين أو تمثيل أوسع من مجموعات أصحاب المصالح المتعددة. ولاشك أن الحضور بشكل شخصي يعطي أصحاب المصلحة الفرصة لطرح الأسئلة وتقديم المعلومات الأساسية الخاصة بكل بلد مما يمكن أن يثري تفسير النتائج ويحسن التقارير النهائية. ويشمل هذا النوع من ورش العمل عروضا للنتائج التي توصل إليها الباحثون و تصريحات السياسة والملاءمة من قبل مسئولي التعليم أو الوكالات التي تقوم بتنفيذ برامج تحسين القراءة و شهادات من جامعي البيانات الميدانية وإنشاء مجموعات عمل لمناقشة نتائج وخطوات العمل. يمكن أن يتبع هذه الورشة الوطنية اجتماعات إقليمية أو مجتمعية.

وبطبيعة الحال، إذا تم تطوير تطبيق تقييم مهارات القراءة في المراحل المبكرة مع الحكومة في المقام الأول، وكانت هناك علاقة وثيقة بين خبراء هذا التقييم والخبراء الحكوميين فستمضي عملية التحقق من الصحة بشكل أفضل كثيرا. وهكذا، ففي معظم (وليس كل) الحالات، سيقلل التعاون الوثيق من البداية من احتمالات الحوار المثير للجدل أثناء التحقق من الصحة.

وبصرف النظر عن ورش العمل الوطنية والإقليمية، تشمل استراتيجيات النشر الأخرى إعداد ونشر تقارير رقمية أو مطبوعة والنشرات واللافتات والرسوم البيانية والاجتماعات المجتمعية أو الوسائط المتعددة (البرامج الإذاعية، مقاطع الفيديو، الأفلام الوثائقية). ويعتبر استخدام التسجيلات الصوتية أو تسجيلات الفيديو خلال ورش العمل أو بشكل مستقل كأحد استراتيجيات النشر، وسيلة واضحة ومقنعة لرؤية الفروق بين القارئ الضعيف (الطفل الذي يقرأ، على سبيل المثال 10-15 كلمة في الدقيقة بدون فهم) والقارئ الجيد (الطفل الذي يقرأ على سبيل المثال 60 كلمة في الدقيقة، مع الفهم). ومن ثم، فوضع النتائج الكمية يكون أسهل كثيرا في هذا الإطار المرجعي.

نموذج دراسة حالة: إصلاح سياسات على المستوى الوطني في اليمن

في بداية عام 2012، طلبت وزارة التعليم اليمنية من برنامج تحسين معيشة المجتمع الممول من قبل الوكالة الأمريكية للتنمية الدولية والمنفذ من قبل منظمة إبداع الدولية، دعم تطوير نهج جديد لتعليم القراءة في الصفوف الابتدائية. وفي مارس 2012، تم تقديم نتائج تقييم مهارات القراءة في المراحل الأولى اجرته بيانات التعليم لصنع القرار (إدداتا)2 في عام 2011في ثلاثة محافظات في إطار حوار السياسات الذي ضم مجموعة كبيرة من أصحاب المصلحة بما فيهم وزير التعليم. وقد قامت منظمة إبداع الدولية_ومقرها صنعاء_ بتسهيل عملية استعراض نتائج تقييم مهارات القراءة للصفوف الدراسية الأولى عندما كان يتم وضع برنامج القراءة للصفوف الدراسية الأولى في اليمن.

عززت النتائج السيئة لتقييم مهارات القراءة للصفوف الدراسية الأولى والذي أجري في عام 2011 (27% من طلاب الصف الثالث لم يتمكنوا من قراءة كلمة واحدة باللغة العربية) من عزم وزارة التعليم لإصلاح طريقة تدريس قراءة اللغة العربية في المراحل المبكرة. وأعطت الوزارة الأولوية لإصلاح القراءة في الصفوف الدراسية الأولى بهدف وضع أساس لخطة تضع نهاية لسنوات من التأخر في قطاع التعليم. وخلال هذه الفترة كان لوزير التعليم دورا رئيسيا في تعبئة وزارة التعليم حول منهج القراءة في الصفوف الدراسية الأولى في اليمن. وتم تنفيذ هذا البرنامج بدعم من الوكالة الأمريكية للتنمية على سبيل التجربة في 310 مدرسة خلال العام الدراسي 2012-2013، ودليلا على نجاح هذا المنهج تم التوسع في تطبيق هذه التجربة في 800 مدرسة في العام الدراسي 2013-2014 و1200 مدرسة في العام الدراسي 2014-2015.

وفي خلال الحوار الوطني في عام 2012، كانت هناك عدة فصائل تريد أن توقف أيه تغييرات في المناهج الدراسية حتى الانتهاء من وضع الدستور وكانت هذه محاولة لضمان إتاحة الفرصة لجميع الأطراف للمشاركة في وضع أي مناهج جديدة. وكان الاستثناء الوحيد لوقف تنقيح المناهج الدراسية، وفقا لما اتفق عليه من قبل الوفود المشاركة في الحوار الوطني، المناهج الدراسية للصفوف الأول حتى الصف الثالث ، نظرا لأن برنامج القراءة للصفوف الدراسية الأولى في اليمن كان قد بدأ بالفعل في جذب الاهتمام لأنه أدى إلى تحسينات كبيرة في قراءة طلاب الصفوف الأولى، وكذلك أدى إلى تطوير مهارات المعلمين وتحفيزهم ، وحسن من مشاركة أولياء الأمور.

وقد أصدر وزير التربية والتعليم عددا من القرارات لضمان نجاح برنامج القراءة للصفوف الدراسية الأولى في اليمن الجديد، وكان أحدها مصمما لضمان الشفافية والجودة في اختيار المدربين؛ فبدلا من أن

تقوم الوزارة بتعيين مدربين من المناطق والمحافظات تم ترتيب عملية اختيار صارمة. وقد شملت تقديم الطلبات والاختيار على أساس مجموعة من المعايير المناسبة لتعليم قراءة اللغة العربية في الصفوف الأولى.

وكان أحد القرارات التي أصدرها الوزير أيضا زيادة وقت القراءة الإجبارية باللغة العربية للصفوف من الأول حتى الثالث في اليوم الدراسي من 5 ل70 دقيقة خمسة أيام في الأسبوع في المدارس التي تنفذ برنامج القراءة للصفوف الدراسية الأولى في اليمن. وأصدر الوزير قرارا آخر للتأكد من أن المعلمين الذين يحضرون التدريب على برنامج القراءة للصفوف الأولى في اليمن هم معلمون يدرسون بالفعل لهذه الصفوف. وقد أراد الوزير بهذا أن يتجنب موقفا متكررا حيث يتم اختيار المعلمين المفضلين من قبل مدير المدرسة بغض النظر عن كونهم يدرسون صفوفا أو موادا لايستهدفها التدريب.

وأخيرا، تم اتخاذ قرارا سياسيا للتوسع في تنفيذ برنامج القراءة للصفوف الأولى في اليمن في كل أنحاء البلاد في خلال العام الثاني من تجربة هذا البرنامج. وفي نفس الوقت الذي كانت تنفذ فيه منظمة إبداع الدولية برنامج القراءة للصفوف الأولى في اليمن بتمويل من الوكالة الأمريكية للتنمية في أكثر من 800 مدرسة كانت الوكالة الألمانية للتنمية تنفذه في 72 مدرسة بتمويل من البنك الدولي، وكانت الوزارة تنفذه في 200 مدرسة. وقد ساهم البنك الدولي في تقديم الدعم للوزارة من أجل التوسع على الصعيد الوطني ليتم تطبيقه في 14700 مدرسة إضافية في العام الدراسي 2014-2015. أي أنه بحلول العام الدراسي 2014-2015 بعد سنتين من بدء تجربة برنامج القراءة للصفوف الأولى في اليمن كانت كل مدارس اليمن البالغ عددها 16000 مدرسة تطبق برنامج القراءة للصفوف الدراسية الأولى _بما فيها معيار70 دقيقة قراءة كل يوم_في كل فصول الصف الأول

المصدر: مقتبس من دو بليسيس، العشري، تيتيان (يصدر قريبا)

المستوى المحلي

إن المواطنين الذين يجرون اختبار برنامج القراءة للصفوف الأولى في اليمن بأنفسهم (أو حتى يطلبون من الأطفال أن يقرأون لهم)، سرعان ما يدركون أن الأطفال لايستطيعون القراءة ولهذا يشعرون بالرغبة في اتخاذ الإجراءات اللازمة ليكونوا جزءا من التغيير. وكثيرا ما يلاحظ أفراد المجتمع مشكلة خطيرة وقصورا شديدا في مهارات القراءة بين الأطفال في المدارس. وقد ساعد تقييم مهارات القراءة للصفوف الدراسية الأولى على تحفيز ذلك في بعض

الأحيان، ولكن في أحيان أخرى يكون هذا رد فعل لمخاوف تم التعبير عنها بالفعل مسبقا. وفي الآونة الأخيرة، تم استخدام تقييم القراءة في المراحل المبكرة لكسب حماسة المجتمع حول قضايا التعليم والحركات الشعبية الأخرى لتطوير الوعي من خلال دليل على مهارة الطلاب في القراءة (أو عدمها). وتشكل الحاجة إلى تعبئة المجتمع والوعي المحلي "خطوة هامة نحو زيادة الطلب على الإصلاحات التعليمية التي تزيد من معرفة القراءة والكتابة" (غوف وسيفيليش، 2011، ص 45).

حتى الآن، تم استخدام تطبيقات تقييم مهارات القراءة في المراحل المبكرة في المقام الأول لتوليد المناقشة على المستوى الوطني، وتحفيز الوزارات على العمل. وتهدف نتائج التقييم للإبلاغ عن أدنى مستوى أو شريحة من العينة (غالبا على المستوى الوطنى ولكن في بعض الأحيان قد يتضمن ذلك مستوى الإقليم والمنطقة). ولأن هذه التقييمات يتم إجرائها على بعض العينات فلا يمكن مشاركة نتائج التقييمات الخاصة بكل مدرسة على حدة. وكذلك فسيصبح الأمر مكلفا إذا قمنا بإجراء تقييما لكل مدرسة (ولعدد كبير من الأطفال في كل مدرسة) للحصول على نتائج لكل مدرسة على حدة. . ومع ذلك، ونظرا للفائدة المزدوجة لرفع الوعي المجتمعي وتشجيع المشاركة المجتمعية في قضايا تنمية الإلمام بالقراءة والكتابة والقراءة للصفوف الدراسية الأولى؛ فهناك العديد من الاستراتيجيات التي يمكن أن يدرسها الممارسون.

وأول استراتيجية موصى بها هي انتاج نتائج تقييم مهارات القراءة للصفوف الأولى باختصار لمشاركتها مع المدارس وكبار رجال المجتمع للحصول على نقاط للمناقشة على حالة الإلمام بالقراءة بشكل عام (ليس بشكل خاص بكل مدرسة ولكن خاص بالشريحة التي توجد بها المدرسة). وهذا التقرير سيكون مصحوبا بشرح عن ما يربط كل مهمة فرعية بالتعليمات ومايمكن للمعلمين أن يقوموا به لتحسين نتائج الطلاب.

وكذلك يمكن مشاركة عينة من خطط الدروس والأنشطة المقترحة مع المدارس لبيان كيف يمكن لأصحاب المدرسة في مجتمع المدرسة اتخاذ اجرائات محليا. وثانيا الحصول على نتائج الإلمام بالقراءة على المستوى المدرسي والإبلاغ بها، ويمكن للممارسين أن يستخدموا أدوات أخرى لتقييم مدى الإلمام بالقراءة مثل تلك المستخدمة من قبل مؤسسة براثام من أجل التقرير السنوي لمستوى التعليم (http://www.asercentre.org/p/141.html) وكذلك البروتوكولات التي تنفذها مجموعات من أجل تقييم مستوى الإلمام بالقراءة والتي تطورها نماذج أخذ عينات

عشوائية لضمان الجودة (انظر دراسة الحالة بالأسفل وكذلك باتشيلدير بيتس مولكاي-دان وستيرن 2015، ومولكاي-دان وفالاديز وكوميسكي وهارتويل 2013 وأيضا فالاديز ومولكاي-دان وسام-بوصمان، 2014).

مثال دراسة حالة : استخدام نظام الرصد عن طريق أخذ عينات العشوائية لمراقبة الجودة بشكل تجريبي ومستمر في غانا

أنشأت غانا مجلس التفتيش الوطني الذي يهدف إلى تطوير الأدوات اللازمة لمراقبة نوعية التعليم كجزء من برنامجها الوطني لتسريع محو الأمية وتعليم القراءة بها. وتم استخدام نظام أخذ العينات العشوائية لضمان الجودة بشكل تجريبي في غانا كوسيلة لاختبار هذه الأداة لمراقبة جودة التعليم وتحديد المناطق التي تحتاج لمزيد من الدعم على المستوى المحلي. وهذه الأنشطة التجريبية والتي تم تصميمها من أجل تحسين مخرجات التعليم تضمنت استخدام تقييم مهارات القراءة للصفوف الأولى من أجل قياس مهارات القراء للطلاب.

وتبين نتائج الدراسة التجريبية لأخذ العينات العشوائية لضمان الجودة أن هناك مخاوف من انخفاض درجات الطلاب في مجال القراءة في المدارس التي كانت ضمن العينة، لكن بالرغم من ذلك فمهارات القراءة الأساسية – والتي تم تقييمها عن طريق تقييم مهارات القراءة في المراحل الأولى – أتاحت لمنهج أخذ العينات العشوائية لضمان الجودة أن يفرق بين نتائج التقييم من مدرسة لأخرى. وكان منهج أخذ العينات العشوائية لضمان الجودة يهدف إلى تصنيف المناطق إلى تلك التي "كانت على مستوى التوقعات" وتلك التي "كانت أقل من مستوى التوقعات". وهذه التصنيفات "كانت بناءا على ما إذا كان 80% من المدارس قد حققت سمات الفائدة المحددة في مجموعة من المؤشرات التي لها علاقة بأداء المعلم وإنجاز الطالب (مولكاي-دان 2013 ص9)

بعد جلسة حوار السياسات في غانا عن الأنشطة التجريبية لأخذ العينات العشوائية لضمان الجودة والعمل الأساسي لتقييم مهارات القراءة للصفوف الأولى تم نشر النتائج في صورة "المنتديات الموجهة للمجتمع"، وقد تم استخدام هذه المنتديات لنشر النتائج على المستوى المحلي. وفي أثناء هذه المنتديات أثارت الجهات المعنية المحلية الحاجة إلى المراقبة المستمرة على مستوى المناطق. وبناءا على هذا الاهتمام والنجاح الذي حققه منهج أخذ العينات العشوائية لضمان الجودة في تقييم أداء المنطقة والمدرسة بطريقة فعالة وموفرة، تلقى هذا المنهج تمويلا إضافيا من أجل التوسع بشكل أكبر في عدة مناطق بغانا.

2.12 وضع معايير قياسية خاصة بكل بلد

وإحدى مزايا تقييم مهارات القراءة للصفوف الأولى هي أن العلم ورائها يتوافق بشكل جيد مع مفهوم رجل الشارع العادي لمعنى القراءة: وهو "معرفة الحروف الأبجدية" والقدرة على القراءة دون تردد وبسرعة معقولة والقدرة على الإجابة على بعض الأسئلة عن ما تم قرائته.

وهكذا فالقدرة على معرفة عدم قدرة الأطفال على التعرف على الحروف الهجائية أو استطاعتهم قرائتها ولكن ببطء شديد هو شئ يستطيع معظم الأفراد اكتشافه. ويعتبر الاعتماد على البيانات التي ينتجها تقييم مهارات القراءة للصفوف الأولى (أو أي نوع آخر من من التقييمات الفردية الشفوية للصفوف الدراسية الأولى) طريقة سليمة يمكن أن تخبرنا بأبسط الطرق إن كانت المدارس تعمل جيدا من أجل الطلاب أم لا.

ومع ذلك، فمن أجل تركيز انتباه واضعي السياسات والمسؤولين على مسألة كيفية تعلم الطلاب القراءة، فمن المفيد أن نكون قادرين على قياس النتائج بطريقة أو بأخرى. ولاشك أن المعايير القياسية مفيدة بشكل خاص بالنسبة للقراءة، حيث أنها تضع التوقعات والمعايير للأداء فيما يخص القراءة. فهناك حاجة إلى معايير لقياس التقدم المحرز في أي دولة أو سياق. ويمكن استخدام المعايير القياسية السليمة لنترجم بسهولة هدفا محددا إلى مجموعة من قياسات التقدم الذي تم إحرازه في نقطة زمنية محددة. فعلى سبيل المثال، إذا كان الهدف هو تعلم كل الأطفال القراءة بنهاية الصف الثالث، فيمكن لمعيار القياس أن يبين نسبة الطلاب الذين حققوا مستويات مختلفة من القدرة على القراءة في صف محدد وسنة محددة لبيان ما إذا كان قد تم إحراز تقدما في سبيل الهدف الرئيسي. وبالإضافة إلى ذلك، فقد تم اكتشاف أن المعايير القياسية مفيدة عندما تستخدم كوسيلة للتواصل علنا بشأن التقدم (مثل، بطاقات تقرير المدرسة أو المراقبة وإعداد التقارير على المستوى الوطني).

وتسمح المعايير القياسية بتطبيق توقع مشترك وقابل للقياس في جميع أنحاء الدولة أو على جميع السكان على الصعيد الوطني، وكذلك السماح بصنع قرارات لامركزية بشأن كيفية مساعدة الأطفال على تحقيق هذه الأهداف. وتعمل نفس هذه القياسات الموضوعية أيضا كآلية للمسائلة لتحميل المدارس _والمعلمين أحيانا_ مسئولية التحصيل العلمي للطلاب. وتشير الدراسات إلى أن نظم التقييم العالية المخاطرة تؤثر بالفعل على سلوك المعلم والمسؤول، ولكن ليس بطريقة ثابتة ويمكن التنبؤ بها. ولذا يجب توخي الحذر عند تطوير معايير القياس لضمان أن نظام التعليم يمكن أن يستخدمهم لقياس مدى التقدم وتحديد المناطق التي تحتاج لمزيد من الجهود بدلا من استخدامهم لتقسيم المسئولية عن العواقب عالية المخاطر.

دراسة حالة: وضع معايير قياسية وطنية في كينيا

اعتبارا من شهر نوفمبر 2015 قامت الوكالة الأمريكية للتنمية ، من خلال مشروع إدداتا 2 (بيانات التعليم لصنع القرار) بتمويل مجموعة من ورش العمل الخاصة بوضع معايير القياس وتحديد الأهداف في 12 دولة هي: مصر، غانا، الأردن، كينيا، ليبيريا، ملاوي، مالي، باكستان، الفلبين، تنزانيا، الضفة الغربية، وزامبيا . وفي كل من هذه الدول، استخدمت بيانات القراءة في الصفوف الأولى للمساعدة في وضع معايير القياس.

من أغسطس 2011 إلى عام 2014، قامت وزارة التربية والتعليم والعلوم والتكنولوجيا بكينيا بتنفيذ مبادرة الرياضيات والقراءة للمرحلة الابتدائية لتحسين المهارات الأساسية في القراءة بين الطلاب. وقد استلهم تصميم هذه المبادرة من خلال تجربة تحسين القراءة التجريبية في منطقة ماليندي في كينيا، التي قامت بها مؤسسة آغا خان و معهد البحوث المثلثية في عام 2007 (معهد البحوث المثلثية الدولي، 2008).

وفي خلال مبادرة الرياضيات والقراءة للمرحلة الابتدائية تم قياس مهارات القراءة لدى مجموعة من الطلاب الذين تم اختيارهم بشكل عشوائي من المدارس التي طبقت فيها هذه المبادرة ومدارس أخرى لم تطبق بها. ولأن هذه المبادرة كان قد تم تصميمها لتكون تجربة عشوائية محكومة لذا لم يكن من الصعب تحديد أثر المبادرة على التعلم. ثم تم استخدام البيانات التي تم جمعها من تقييم مهارات القراءة للصفوف الأولى لإبلاغ القرارات السياسية الوزارية بشأن الاستثمار في طرق التدريس الخاصة التي من شأنها أن تؤدي إلى تحسين النتائج عبر التصنيفات القائمة على النوع الاجتماعي والمستوى الاجتماعي والاقتصادي. بالإضافة إلى ذلك، دعت وزارة التعليم والعلوم والتكنولوجيا الكينية مبادرة الرياضيات والقراءة للمرحلة الابتدائية " إلى تنفيذ برامج بحثية موحدة من خلال العمل مع الجمعية الوطنية لمجلس الامتحانات الكينية (KNEC) في وضع معايير للقراءة والكتابة والحساب" (معهد البحوث المثلثية الدولي، 2014ب، ص47). وعلاوة على ذلك، تم إنشاء أداة استحقاق كينية لجودة التعليم من أجل قياس مستوى التعلم في الصفوف الدراسية الأولى. وقد أنشأت مجموعة عمل مبادرة الرياضيات والقراءة للمرحلة الابتدائية المختصة بالمراقبة والتقييم هذه الأداة بناءا على أدوات تقييم مهارات القراءة للصفوف الدراسية

الأولى المستخدم في مبادرة الرياضيات والقراءة للمرحلة الابتدائية ثم قامت وزارة التعليم والعلوم والتكنولوجيا الكينية بتضمين هذه الأداة في مستويات النتائج المعيارية. وقد قدم موظفو مبادرة الرياضيات والقراءة للمرحلة الابتدائية للجنة توجيه الجمعية الوطنية لمجلس الامتحانات الكينية تصميم البحوث والنتائج الأساسية والتوصيات ونتائج وضع المعايير القياسية. وخلال هذا العرض تمكنت مبادرة الرياضيات والقراءة للمرحلة الابتدائية من إظهار مستوى دقتها في قياس مستوى تعلم الطلاب الذي أثر على المعايير القياسية. وفي تدريب تم خلال اجتماع الجمعية الوطنية لمجلس الامتحانات الكينية، طلب من أعضاء لجنة التوجيه تحديد معايير قياسية مناسبة لفهم وطلاقة الطلاب باستخدام البيانات الأساسية التي جمعتها مبادرة الرياضيات والقراءة للمرحلة الابتدائية من تقييم مهارات القراءة للصفوف الدراسية الأولى. و في النهاية تم تعديل أداة الإستحقاق لدمج مستويات التهديف القياسي لوزارة التعليم والعلوم والتكنولوجيا الكينية.

1.2.12 ما هي المعايير القياسية؟

تم تعريف المعايير القياسية بأنها "معيار أو نقطة مرجعية تتم من خلالها مقارنة الأشياء أو تقييمها" (قواميس أوكسفورد على الشبكة http://www.oxforddictionaries.com . وأيضا " معيار للأداء عند نقطة معينة (علامة فارقة)" وهي أيضا "معيار مرجعي ونتيجة مستهدفة مستمد من التجربة و يمثل التقدم في القراءة بشكل مناسب" (مجموعة القياس الديناميكي، شركة محدودة ، 2010، ص1)

ولأغراض خاصة بهذا الدليل نجد أن كلمة "المعيار القياسي" مرادفة لكلمة "المعيارأو النموذج" في أنها تحدد المستوى المطلوب من الأداء الذي يمكن تحقيقه عند نقطة معينة من الزمن.وهكذا، فإن التقييم القياسي" هو اختبار تشخيصي يتم على فترات منتظمة ويستخدم لتقييم الطلاب ومعرفة ما إذا كانوا يتقدمون على الطريق الصحيح نحو تحقيق المعايير المطلوبة أم لا. ويمكن أيضا "للنتائج المعيارية" أن يتم إنشائها في نقاط فاصلة لتساعد على تفسير معنى نتيجة معينة فعلى سبيل المثال: فإن تحديد نقاط فاصلة "مبتدئ" و"متوسط" و"متقن" يمكن أن تساعد في تحديد ملامح الطالب بناءا على تعريف الإتقان التام أو الجزئي.

ويمكن للمعايير القياسية أن ترتبط أيضا "بالأهداف والغايات" التي تحدد التوقعات للسكان فعلى سبيل المثال إذا حدد المعيار القياسي مدى صعوبة الهدف المطلوب تحقيقه، فسيحدد الهدف عدد الأطفال القادرين على تحقيقه. فعلى سبيل المثال سيتمكن 60% من الطلاب من تحقيق

المعيار القياسي في السنة الأولى ، وسيتمكن 80% من الطلاب من تحقيقه في السنة الثانية " فلاشك أن تحديد الأهداف يعد أمرا بالغ الأهمية عندما يكون مستوى الأداء منخفضا. فالهدف هنا يحدد خطوة انتقالية في سبيل تحقيق الهدف.

وكما أوضحنا سابقا (القسم 1.12 الخاص بالنشر)، في أنشطة التواصل، تكون الرسائل مؤثرة فقط عندما يفهمها الجمهور المراد الوصول اليه. فتقديم نتائج تقييم مهارات القراءة للصفوف الأولى دون وجود نقطة مرجعية يكون عادة غير مؤثر في البيئات الي تكون قياسات الطلاقة بها (20 كلمة صحيحة في الدقيقة)غير مألوفة ويتم الإبلاغ عن التقييمات عادة في صورة نسبة من الإجابات الصحيحة. والمعيار القياسي هو نقطة مرجعية تقوم بتفسير الأداء لأنها توفر مستوى متوقع من الإنجاز. وفي حالة المعايير القياسية التعليمية، فهي تضيف خصوصية لأهداف المناهج الدراسية الواسعة مثل "يجب أن يكون قادرا على القراءة بطلاقة" بالقول بدلا عن ذلك " "يجب أن يكون قادرا على القراءة بمعدل 40 كلمة صحيحة في الدقيقة الواحدة بحلول نهاية الصف الثاني. "ومع ذلك يجب أن تستند هذه التوقعات على واقع الدولة أكثر من الاعتماد على توقعات من دول أو لغات أخرى. ويمكن استخدام بيانات تقييم مهارات القراءة للصفوف الدراسية الأولى لتحديد المعايير القياسية ويمكن للإدارات اللاحقة توليد البيانات لتقييم الأداء مع مرور الوقت وفقا لتلك المعايير. لأغراض المقارنة، الملحق P يقدم قواعد القراءة الشفوية بطلاقة الخاصة باللغة الإنجليزية.

التعاريف

- الغاية هو طموح على المدى الطويل، ربما بدون قيمة عددية
  الغاية: يجب أن يتمكن جميع أطفالنا من القراءة

- متري هو، وحدة موثوقة صالحة للقياس
  متري: "الكلمات الصحيحة في الدقيقة الواحدة في قراءة المقطع"

- المعيار القياسي هو الخطوة العددية نحو الهدف، وذلك باستخدام المتري
  المعيار القياسي: 45 الكلمات الصحيحة في الدقيقة الواحدة، فهم 80٪ من ما قرأوا

- الهدف هو متغير باستخدام المعيار القياسي

الهدف: وصول ٪ من الأطفال إلى أو فوق المعيار القياسي، أو المتوسط الذي يحققه الأطفال، باستخدام المتري.

المصدر: لاتوسكي (2014)

2.2.12 معايير وضع المعايير القياسية

من أجل وضع المعايير القياسية يمكن استخدام عملية تجمع بين التحليل الإحصائي لبيانات الطلاب مع مرور الوقت ومعلومات إضافية مثل بحث عن وسيلة تعلم الأطفال القراءة والتجربة في أماكن أخرى ورؤى من العلوم المعرفية، ومعرفة السياقات المحلية.

وقد تتغير المعايير القياسية بمرور الوقت تماشيا مع تحسن مستوى أداء الطالب، وهناك العديد من الطرق لوضع المعايير القياسية لكن أهم المعايير لوضع معايير قياسية جيدة تتضمن:

- المعايير القياسية طموحة، ولكن يجب أن تكون واقعية وقابلة للتحقيق.
- وهي لاتخضع لتضخم النتيجة (أي أن زيادة النتائج لا تعمم على قياسات أخرى لنفس المحتوى لأنها تعكس في المقام الأول أنشطة اختبار تحضيرية محددة وموجهة نحو اختبار محدد (هاملتون وستيتشر ويوان 2008)
- يجب أن تكون المعايير القياسية قادرة على تحديد الطلاب الذين من المحتمل أن يفشلوا في تحقيق مستوى مستقل للقراءة. وترتبط المعايير القياسية بنقطة زمنية محددة (بداية العام، نهاية العام، الصف الدراسي وهكذا) والمعايير القياسية اللاحقة يتم استنباطها بالاستناد على احتمالية أن الأطفال الذين يحققون المعيار القياسي الأول سيتمكنوا من تحقيق المعيار القياسي التالي (تحت الظروف التعليمية الحالية) (مجموعة القياس الديناميكية، شركة محدودة 2010)

لا توجد درجات نجاح محددة للاختبار لكنها تكون دفاعية بشكل أو بآخر. وتستند القابلية للدفاع إلى حد كبير على الطريقة التي استخدمت لوضع المعايير. ثانيا: لا توجد طريقة مثلى أو صحيحة لوضع المعايير وإنما هي مجموعة من الأساليب التي قد تكون بشكل أو بآخر مناسبة لحالة معينة ".

فيرارا، بيري وجونسون 2008

- تستند المعايير القياسية على الأبحاث التي تدرس الصحة التنبؤية لنتيجة ما على مقياس ما في نقطة زمنية محددة مقارنة بالمقاييس التالية وتقييم النتائج الخارجية. إذا حقق الطالب هدف المعيار القياسي فمن المرجح أن يتمكن من تحقيق إنجازات القراءة التالية إذا تلقى توجيها قائما على الأبحاث من منهج دراسي جيد (مجموعة القياس الديناميكية، شركة محدودة 2010)

- وأفضل أنواع البيانات التي يمكن استخدامها هي درجات الاختبارات الخاصة بمن يخضعون لاختبارات حقيقية والذي تم الحكم على أدائهم من قبل حكام مؤهلين (زيكي وبيري 2006)

- وترتبط المعايير القياسية معا بشكل مناسب عبر الصفوف الدراسية لتجنب الخطأ في تصنيف الطلاب أو التقارير المضللة للجهات المعنية على سبيل المثال في حين أنه قد يكون من المناسب تعيين نقطة فاصلة أعلى لتعريف طالب الصف الثاني المتقدم المستوى بدلا من تعريف طالب الصف الثالث ذو المستوى العادي لكن العكس غير صحيح (زيكي وبيري 2006).

وتستند جميع المعايير في نهاية المطاف على القواعد والأعراف، أو الأحكام الصادرة عن ما يجب أن يكون الطفل قادرا على القيام به (زيكي وبيري 2006). ويمكن لكل دولة أن تضع المعايير القياسية الخاصة بها بالنظر إلى مستوى الأداء في المدارس المعروفة بحسن أدائها أو تلك التي تحصل على نتائج جيدة في تقييم مهارات القراءة للصفوف الأولى ولكن لاتمتلك أيه مزايا اجتماعية أو اقتصادية أو مستوى غير مستدام من استخدام الموارد.

وهذه المدارس عادة ما تسفر عن معايير قياسية متطلبة بشكل معقول لكن من الممكن تحقيقها بشكل واضح حتى من قبل الأطفال الذين لا يتمتعون بمزايا اجتماعية أو اقتصادية كبيرة أو في المدارس التي لا تمتلك موارد كبيرة طالما تمتلك هذه المدارس تدريسا على مستوى جيد. فعلى سبيل المثال اختارت الدراسة التي اجريت في عام 2001 "التقدم في تعليم القراءة دوليا " أربعة نقاط فاصلة على مقياس تعلم القراءة المشترك وسميت بالمعايير القياسية الدولية. وقد تم اختيار هذه المعايير القياسية لتتوافق مع عدد النقاط التي تم إحرازها والتي تساوي أو تفوق الربع الأسفل أو المتوسط أو الأعلى وكذلك أعلى 10 في المئة من طلاب الصف الرابع في عينة الدراسة الدولية التي أجريت عام 2001 "التقدم في تعليم القراءة دوليا" (معهد علوم التربية، بدون تاريخ).

3.2.12 عملية وضع المعايير القياسية

كما ذكر في واحدة من دراسات الحالة أعلاه، فقد دعم مشروع ادداتا (بيانات التعليم لصنع القرار) الثاني التابع للوكالة الأمريكية للتنمية صياغة المعايير القياسية لأداء القراءة في 12 دولة في نوفمبر 2015. في هذه الدول، تم استخدام عملية متسقة للمساعدة في تحديد مستويات مقبولة من الأداء في عدة مجالات من تنمية مهارات القراءة في الصفوف الدراسية المختلفة. وفيما يلي المبادئ التوجيهية الموضوعة استنادا إلى بعض الدروس المستفادة من خلال العمل في 12 دولة.

الخطوة الأولى: ابدأ بمناقشة مستوى الفهم المقبول أثناء القراءة والذي يدل على فهم كامل لنص معين. وقد استقرت معظم الدول على 80٪ أو أعلى (4 أو أكثر إجابات صحيحة من 5 أسئلة) كمستوى مطلوب من الفهم.

الخطوة الثانية: بالنظر إلى معيار الفهم أثناء القراءة يتم استخدام تقييم مهارات القراءة للصفوف الأولى لعرض مجموعة من نتائج طلاقة القراءة الشفهية _والتي تقاس بعدد الكلمات الصحيحة في الدقيقة_ والتي حققها الطلاب الذين تمكنوا من تحقيق المستوى المطلوب من الفهم. ولهذا فسيكون هناك حاجة للمناقشة لتحديد القيمة في هذا النطاق الذي تم طرحه كمعيار قياسي. بدلا من ذلك يمكن لهذا النطاق أن يشير إلى مستويات تنمية المهارات المقبولة لمستوى "الإتقان" أو تلبية معيار مستوى الصف (على سبيل المثال 40-50 كلمة في الدقيقة).

الخطوة الثالثة: بعد تحديد المقياس المعياري لطلاقة القراءة الشفوية ، تتيح العلاقة بين طلاقة القراءة الشفوية وفك التشفير (القراءة بدون كلمات)تحديد متوسط معدل القراءة بدون كلمات الذي يتوافق مع مستوى معين من طلاقة القراءة الشفوية.

الخطوة الرابعة: ثم تستمر العملية بنفس الطريقة في كل مجالات المهارات اللاحقة

3.12 تنبيهات وقيود

في بعض السياقات تكون ردود الأفعال تجاه تقييمات القراءة من نوعية تقييم مهارات القراءة في المراحل الأولى ليست واضحة. فبعض المعلقين في بعض الدول يشككون في جدوى طلاقة

القراءة الشفوية كمؤشر تمهيدي للتعليم العام أو حتى القراءة. بل وقد يتسائلون عن سبب احتواء التقييم على بنود أو أشكال لاتعكس التدريس في الفصول بشكل مباشر (على سبيل المثال: قراءة كلمات مخترعة). ولهذا من المهم الوصول للخلفية الأدبية التي تشرح المنطق السائد، وقد أشير لبعضها في هذا الدليل، وأيضا في الموقع الإلكتروني للجمعية الدولية لتعليم القراءة (www.reading.org) وفي صفحة المعهد الوطني الأمريكي لصحة الطفل، وصفحة لجنة القراءة الوطنية للتنمية البشرية وعلى الموقع الإليكتروني لمركز للتعليم والتعلم في جامعة ولاية أوريغون، والمؤشرات الحيوية لمهارات القراءة المبكرة الأساسية (DIBELS, http://dibels.uoregon.edu).

و في حالات أخرى، يبدو أن الجماهير المحتملين يرون أن جهود تقييم مهارات القراءة للصفوف الأولى تحاول أن تنقل المفهوم القائل بأن " القراءة هي كل ما يهم". في هذه الحالات من المهم أن نلاحظ أن القراءة هي في الواقع مهارة تأسيسية هامة تؤثر على النجاح الأكاديمي عبر المناهج الدراسية، وأيضا أن القراءة هي علامة جيدة لجودة المدارس بوجه عام. ومع ذلك، لا يقوم هذا الجهد على افتراض أن القراءة هي كل ما يهم.

وبشكل عام، فإن أي محاولة لقياس جودة التعليم أو التعلم، يخضع لهذا النوع من النقاشات المعروفة. ومن خلال التجربة المتراكمة والمكتسبة من خلال تطبيق أداة تقييم مهارات القراءة في المراحل المبكرة وما يشبهها من أدوات، يبدو أن المعلمين والمعنيين بدعم المعلمين بشكل مباشر، والمسؤولين رفيعي المستوى يدركون على الفور قيمة تقييم مهارات القراءة في المراحل المبكرة، في حين أن بعض أصحاب النظريات حول المناهج الدراسية أو القراءة لديهم بعض المخاوف بشأن احتمال التبسيط المفرط. فإنه من المهم أن نفهم أن الاستخدام العملي لتقييم مهارات القراءة في المراحل الأولى واستراتيجيات التحسين المستمدة منها ينبغي أن ينظر لها فقط كمدخل أو نقطة بداية. يمكن استخدام البيانات كمثال على ما يمكن تحقيقه من خلال التركيز ومتابعة نتائج محددة. ومن ثم يمكن تطبيق الدرس الأساسي على جوانب أخرى للتعليم والتعلم.

وقد تكون مقاومة منهجية تقييم مهارات القراءة للصفوف الأولى أقوى كلما كانت النتائج أضعف، ولهذا فمن المهم إجراء التقييم وتحليل النتائج بدقة وموضوعية. وكذلك يمكن للاستبيانات الإضافية السياقية (الطالب، المعلم، وخصائص المدرسة؛ والملاحظات الصفية، وغيرها) أن تساعد في تفسير نتائج الأداء، لكن لاشك أن تنفيذهم وتحليل البيانات الناتجة سيضيف لتكاليف التنفيذ. وعندما يتم ربط أدوات المسح الإضافية بنتائج تقييم مهارات القراءة

للصفوف الأولى، يجب أن يولي المنفذون اهتماما دقيقا لحجم العينة والدلالة الإحصائية وتجنب عمل علاقة مع السببية حتى يتم إجراء المزيد من البحوث.

وكانت مبادرة التعليم العالمية الأولى التي بدأت في سبتمبر 2012 من قبل الأمين العام للأمم المتحدة فد بينت الحاجة للتقييم الفعال لأداء الطلاب من أجل تحسين نظم التعليم. فالتقييم عن كثب ومراقبة مدى فعالية عمل النظام يمكن أن يؤدي إلى التأثير على السياسة، كما أنه يعطي المسؤولين وصناع القرار الفرصة "لاستخدام المعلومات لتوجيه الدعم والموارد لإيجاد حلول فعالة (مكتب الأمين العام للأمم المتحدة، 2012، ص 19).

المراجع

Abadzi, H. (2006). *Efficient learning for the poor.* Washington, DC: The World Bank. https://openknowledge.worldbank.org/handle/10986/7023

Adolf, S. M., Catts, H. W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities*, *43*(4), 332–345. http://dx.doi.org/10.1177/0022219410369067

Abadzi, H. (2012). *Developing cross-language metrics for reading fluency measurement: Some issues and options.* Global Partnership for Education working paper. Washington, DC: World Bank. Retrieved from https://www.academia.edu/3484052/Developing_Cross-Language_Metrics_for_Reading_Fluency_Measurement_Some_issues_and_options._World_Bank_Global_Partnership_for_Education_working_paper

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing: An Interdisciplinary Journal, 13,* 147–157.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Adolf, S. M., Catts, H. W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities*, *43*(4), 332–345. http://dx.doi.org/10.1177/0022219410369067

Armbruster, B. B., Lehr, F., & Osborn, J. (2003). *Put reading first: The research building blocks of reading instruction.* Washington, DC: Center for the Improvement of Early Reading Achievement (CIERA).

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners.* Prepared by the Center for Applied Linguistics and SRI International for the Institute of Education Sciences and the Office of English Language Acquisition, US Department of Education; and the US National Institute of Child Health and Human Development. Washington, DC: Lawrence Erlbaum Associates and the Center for Applied Linguistics.

Ayari, S. (1996). Diglossia and illiteracy in the Arab world. *Language, Culture and Curriculum, 9,* 243–253.

Badian, N. A. (2001). Phonological and orthographic processing: Their roles in reading prediction. *Annals of Dyslexia, 51,* 179–202.

Batchelder, K., Betts, K., Mulcahy-Dunn, A. & Stern, J. (2015). Lot quality assurance sampling (LQAS) pilot in Tanzania: Final report. Prepared for USAID under the EdData II project, Task Order No. AID-OAA-12-BC-00003 (RTI Task 20, Activity 5). Research Triangle Park, NC: RTI International.

Braun, H., & Kanjee, A. (2006). Using assessment to improve education in developing nations. In H. Braun, A. Kanjee, E. Bettinger, & M. Kremer (Eds.), *Improving education through assessment, innovation, and evaluation* (pp. 1–46). Cambridge, MA: American Academy of Arts and Sciences. Retrieved from https://www.amacad.org/publications/braun.pdf

Bulat, J., Brombacher, A., Slade, T., Iriondo-Perez, J., Kelly, M., & Edwards, S. (2014). *Projet d'Amélioration de la Qualité de l'Education (PAQUED): 2014. Endline report of Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA).* Prepared for USAID under Contract No. AID-623-A-09-00010. Washington, DC: Education Development Center and RTI International.

Center for Global Development. (2006). *When will we ever learn? Improving lives through impact evaluation.* www.cgdev.org/files/7973_file_WillWeEverLearn.pdf

Chabbott, C. (2006). *Accelerating early grades reading in high priority EFA Countries: A desk review.* http://www.equip123.net/docs/E1-EGRinEFACountriesDeskStudy.pdf

Chall, J. (1996). *Stages of reading development* (2nd ed.). Fort Worth, TX: Harcourt-Brace.

Chiappe, P., Siegel, L., & Wade-Woolley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading, 6*(4), 369–400.

Clay, M. M. (1993). *An observation survey of early literacy achievement.* Ortonville, MI.: Cornucopia Books.

Collins, P., & Messaoud-Galusi, S. (2012). *Student performance on the Early Grade Reading Assessment (EGRA) in Yemen* [English version; also available in Arabic]. Report prepared for USAID under the EdData II project, Task Order EHC-E-07-04-00004-00 (RTI Task 7). Research Triangle Park, NC: RTI International. http://pdf.usaid.gov/pdf_docs/PNADZ047.pdf

Coltheart M., Rastle K., Perry C., Langdon R., & Ziegler J. C. (2001). DRC: a dual-route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Crouch, L., & Korda, M. (2008). *EGRA Liberia: Baseline assessment of reading levels and associated factors.* Report prepared for the World Bank under Contract No. 7147768. Research Triangle Park, NC: RTI International.

Crouch, L., & Winkler, D. (2008). Governance, management, and financing of Education for All: Basic frameworks and case studies. Background paper commissioned for the *Education for All global monitoring report 2009: Governance, management and financing of education for all.* Research Triangle Park, NC: RTI International. unesdoc.unesco.org/images/0017/001787/178719e.pdf

Cunningham, P.M., & Allington, R. L. (2015). *Classrooms that work: They can all read and write* (6th ed.). Boston, MA: Pearson.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities*, *39*(6), 507–514. http://dx.doi.org/10.1177/00222194060390060301

Darney, D., Reinke, W. M., Herman, K. C., Stormont, M., & Ialongo, N. S. (2013). Children with co-occurring academic and behavior problems in first grade: Distal outcomes in twelfth grade. *Journal of School Psychology*, *51*(1), 117–128. http://dx.doi.org/10.1016/j.jsp.2012.09.005

Denton, C. A., Ciancio, D. J., & Fletcher, J. M. (2006). Validity, reliability, and utility of the observation survey of early literacy achievement. *Reading Research Quarterly, 41*(1), 8–34.

Denton, C. A., Hasbrouck, J. E., Weaver, L. R., & Riccio, C. A. (2000). What do we know about phonological awareness in Spanish? *Reading Psychology, 21,* 335–352.

Dubeck, M. M., & Gove, A. (2015). The early grade reading assessment (EGRA): Its theoretical foundation, purpose, and limitations. *International Journal of Educational Development, 40,* 315–322. http://dx.doi.org/10.1016/j.ijedudev.2014.11.004

du Plessis, J., El-Ashry, F., & Tietjen, K. (Forthcoming). Oral reading assessments in Yemen: Turning bad news into a national reform. In *Understanding what works in oral reading assessments.* Montreal: UNESCO Institute for Statistics (UIS).

Dynamic Measurement Group, Inc. (2010). *DIBELS® Next benchmark goals and composite score.* https://dibels.org/papers/DIBELSNextBenchmarkGoals.pdf

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum Associates.

Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly*, 20(2), 163–179.

Ferguson, C. A. (1959). Diglossia. *Word, 15,* 325–340.

Ferrara, S, Perie, M., & Johnson, E. (2008). Matching the judgmental task with standard setting panelist expertise: The item-descriptor (ID) matching method. *Journal of Applied Testing Technology, 9*(1), 1–22.

Filmer, D., Hasan, A., & Pritchett, L. (2006). *A millennium learning goal: Measuring real progress in education.* Washington, DC: World Bank. Retrieved from http://dx.doi.org/10.2139/ssrn.982968

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.) New York: John Wiley.

Fuchs, L., Fuchs, D., Hosp, M. K., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239–256.

Gambrell, L. B., & Morrow, L. M. (Eds). (2014). *Best practices in literacy instruction* (5th ed.). New York, NY: Guildford.

Glick, P., & Sahn, D. E. (2010). Early academic performance, grade repetition, and school attainment in Senegal: A panel data analysis. *The World Bank Economic Review, 24*(1), 93–120.

Goikoetxea, E. (2005). Levels of phonological awareness in preliterate and literate Spanish-speaking children. *Reading and Writing, 18,* 51–79.

Good, R. H., Simmons, D. C., & Smith, S. (1998). Effective academic intervention in the United States: Evaluating and enhancing the acquisition of early reading skills. *School Psychology Review, 27,* 45–56.

Goswami, U. (2008). The development of reading across languages. *Annals of the New York Academy of Sciences, 1145,* 1–12.

Gove, A., & Cvelich, P. (2011). *Early reading: Igniting education for all. A report by the Early Grade Learning Community of Practice* (rev. ed). Research Triangle Park, NC: RTI International. http://www.rti.org/publications/abstract.cfm?pubid=17099

Gove, A., & Wetterberg, A. (2011). The Early Grade Reading Assessment: An introduction. In A. Gove & A. Wetterberg (Eds.), *The Early Grade Reading Assessment: Applications and interventions to improve basic literacy* (pp. 1–37). Research Triangle Park, NC: RTI Press. http://www.rti.org/pubs/bk-0007-1109-wetterberg.pdf

Gove, A., & Wetterberg, A. (Eds.). (2011). *The Early Grade Reading Assessment: Applications and interventions to improve basic literacy.* Research Triangle Park, NC: RTI Press. http://www.rti.org/pubs/bk-0007-1109-wetterberg.pdf

Hamilton, L. S., Stetcher, B. M., & Yuan, K. (2008). *Standards-based reform in the United States: history, research, and future directions.* Prepared under National Science Foundation Grant No. REC-0228295. Santa Monica, CA: RAND Corporation. http://www.rand.org/content/dam/rand/pubs/reprints/2009/RAND_RP1384.pdf

Hanson, B. A., Zeng, L., & Colton, D. (1994). *A comparison of presmoothing and postsmoothing methods in equipercentile equating* (ACT Research Report 94-4). Iowa City, IA: ACT.

Hanushek, E. A., & Woessman, L. (2009). *Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation.* Working Paper 14633. Cambridge, MA: National Bureau of Economic Research.

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*(7), 636–644.

Hirsch Jr., E. D. (2003). Reading comprehension requires knowledge of words and the world: Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores. *American Educator* (Spring), 10–44.

Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), Educational measurement (4th ed., pp. 187–220).Westport, CT: Praeger.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127–160.

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher, 58*(8), 702–714.

Institute of Education Sciences, National Center for Education Statistics [US]. (n.d.). *International comparisons in fourth-grade reading literacy: Reading literacy by benchmarks* (Web page). http://nces.ed.gov/pubs2004/pirlspub/5.asp

Jakobsen, R. (1960). Closing statements: Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in language* (pp. 350–377). Cambridge, MA: MIT Press.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology 80*(4), 437–447.

Juel, C. (1991). Beginning reading. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 759–788). New York: Longman.

Kamhi, A.G., & Catts, H. W. (1991). Language and reading: Convergences, divergences, and development. In A. G. Kamhi & H. W. Catts (Eds.), *Reading disabilities: A developmental language perspective* (pp. 1–34).Toronto, Ontario, Canada: Allyn & Bacon.

Kandhadai, P., & Sproat, R. (2010). Impact of spatial ordering of graphemes in alphasyllabic scripts on phonemic awareness in Indic languages. *Writing Systems Research*, *2*(2), 105–116.

Kanjee, A. (2009). *Assessment overview* [Presentation]. Prepared for the first READ Global Conference, "Developing a Vision for Assessment Systems," Moscow, October 1, 2009. http://www.worldbank.org/content/dam/Worldbank/document/Program/READ/Events/READ-conference-2009/READ_GC_Presentation_5_AKanjee_Eng.pdf

Karanth, P. (2002). Reading into reading research through nonalphabetic lenses: Evidence from the Indian languages. *Topics in Language Disorders, 22*(5), 20–31.

Kleinman, L., Leidy, N. K., Crawley, J., Bonomi, A., & Schoenfeld, P. (2001). A comparative trial of paper-and-pencil versus computer administration of the quality of life in reflux and dyspepsia (QOLRAD) questionnaire. *Medical Care 39,* 181–189.

Kochetkova, E., & Dubeck, M. (In press). Assessment in schools. Chapter in *Understanding what works in oral reading assessments*. Montreal: UNESCO Institute for Statistics (UIS).

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking.* New York, NY: Springer-Verlag.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293–323.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

LaTowsky, R. (2014). *Towards possible early grade reading benchmarks for the West Bank* (Presentation slides). Prepared for USAID under the Education Data for Decision Making (EdData II) project, Measurement and Research Support to Education Strategy Goal 1, Task Order No. AID-OAA-12-BC-00003 (RTI Task 20). Research Triangle Park, NC: RTI International. https://www.eddataglobal.org/countries/index.cfm?fuseaction=pubDetail&ID=778

LaTowsky, R.J., Cummiskey, C., & Collins, P. (2013). *Egypt grade 3 Early Grade Reading Assessment baseline. Draft for review and comment.* Prepared for USAID under the Education Data for Decision Making (EdData II) project, Data for Education Programming in Asia and the Middle East (DEP-AME) task order, Contract No. AID-278-BC-00019. Research Triangle Park, NC: RTI International.

Linan-Thompson, S., & Vaughn, S. (2004). *Research-based methods of reading instruction: Grades K-3.* Alexandria, VA: Association for Supervision and Curriculum Development.

Linan-Thompson, S., & Vaughn, S. (2007). *Research-based methods of reading instruction for English-language learners: Grades K–4*. Alexandria, VA: Association for Supervision and Curriculum Development.

Lonigan, C., Wagner, R., Torgesen, J. K., & Rashotte, C. (2002). *Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP).* Tallahassee: Department of Psychology, Florida State University.

Management Systems International (MSI). (2014). Early Grade Reading Assessment baseline report. Balochistan province. Prepared for USAID under the Monitoring and Evaluation Program (MEP), Contract No. AID-391-C-13-00005. Washington, DC: MSI. http://pdf.usaid.gov/pdf_docs/PA00KB9N.pdf

Manis, F. R., Lindsey, K. A., & Bailey, C. E. (2004). Development of reading in grades K–2 in Spanish-speaking English language learners. *Learning Disabilities Research and Practice, 19*(4), 214–224.

Marsick, V. J., & Watkins, K. E. (2001). Informal and incidental learning. *New Directions for Adult and Continuing Education, 89,* 25-34. http://tecfa.unige.ch/staf/staf-k/borer/Memoire/incidentiallearning/incidentiallearning.pdf

McBride-Chang, C. & Ho, C. S.-H. (2005). Predictors of beginning reading in Chinese and English: A 2-year longitudinal study of Chinese kindergarteners. *Scientific Studies of Reading, 9,* 117–144.

McBride-Chang, C., & Kail, R. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development, 73,* 1392–1407.

Mulcahy-Dunn, A., Valadez, J. J., Cummiskey, C., & Hartwell, A. (2013). *Report on the pilot application of lot quality assurance sampling (LQAS) in Ghana to assess literacy and teaching in primary grade 3.* Prepared for USAID under the EdData II project, Task Order No. EHC-E-07-04-00004-00 (RTI Task 7). Research Triangle Park, NC: RTI International.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundation of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40,* 665–681.

Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading, 30*(1), 7–22.

Nag, S. (2014). Akshara-phonology mappings: the common yet uncommon case of the consonant cluster. *Writing Systems Research, 6,* 105–119.

Nag, S., & Perfetti, C. A. (2014). Reading and writing: Insights from the alpha-syllabaries of South and Southeast Asia. *Writing Systems Research, 6*(1), 1–9.

Nagy, W. E., & Scott, J. (2000). Vocabulary processes. In M. E. A. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr, (Eds.), *Handbook of reading research* (Vol. III, pp. 269-284). Mahwah, NJ: Erlbaum.

Nation, K. (2005). Connections between language and reading in children with poor reading comprehension. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 41–54). Mahwah, NJ: Erlbaum.

National Center for Family Literacy (NCFL) [US]. (2008). *Developing early literacy: Report of the national early literacy panel. A scientific synthesis of early literacy development and implications for intervention.* Prepared under inter-agency agreement IAD-01-1701 and IAD-02-1790 between the Department of Health and Human Services and the National Institute for Literacy. Washington, DC: National Institute for Literacy. https://www.nichd.nih.gov/publications/Pages/pubs_details.aspx?pubs_id=5750

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, US Department of Health, Education and Welfare (DHEW). (1978). *Belmont Report: Ethical principles and guidelines for the protection of human subjects of research.* Report of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. DHEW Pub. No. (OS) 78-0012. Washington, DC: United States Government Printing Office. http://videocast.nih.gov/pdf/ohrp_belmont_report.pdf

National Institute of Child Health and Human Development (NICHD) [US]. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: NICHD. https://www.nichd.nih.gov/publications/pubs/nrp/Pages/smallbook.aspx

Nielsen, D. (2014). *Early grade reading and math assessments in 10 countries: Dissemination and utilization of results—a review.* Report prepared for USAID under the Education Data for Decision Making (EdData II) project, Measurement and Research Support to Education Strategy Goal 1, Task Order No. AID-OAA-BC-12-00003 (RTI Task 20). Research Triangle Park, NC: RTI International. http://pdf.usaid.gov/pdf_docs/PA00K8RP.pdf

Office of the United Nations Secretary-General. (2012). *Global Education First Initiative: An initiative of the United Nations Secretary-General.* New York: United Nations. http://www.globaleducationfirst.org/files/GEFI_Brochure_ENG.pdf

Optimal Solutions Group, LLC. (2015). *Secondary Analysis for Results Tracking (SART) data sharing manual, USAID Ed Strategy 2011–2015, Goal 1.* Prepared for USAID under the Secondary Analysis for Results Tracking (SART) project, Contract AID-OAA-C-12-00069. Location: Optimal Solutions. Retrieved from https://sartdatacollection.org/images/SARTDataSharingManualFeb2015.pdf

Orr, D. B., & Graham, W. R. (1968). Development of a listening comprehension test to identify educational potential among disadvantaged junior high school students. *American Educational Research Journal, 5*(2), 167–180.

Paris, S. G., & Paris, A. H. (2006). Chapter 2: Assessments of early reading. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development, 6th Edition* (Vol. 4: Child Psychology in Practice). Hoboken, New Jersey: John Wiley and Sons.

Patrinos, H. A, & Velez, E. (2009). Costs and benefits of bilingual education in Guatemala: A partial analysis. *International Journal of Educational Development, 29*(6), 594–598.

Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading, 7*(1), 3–24.

Perfetti, C. A., & Dunlap, S. (2008). Learning to read: General principles and writing system variations. In K. Koda & A. Zehler (Eds.), *Learning to read across languages* (pp. 13–38). Mahwah, NJ: Erlbaum.

Piper, B., & Korda, M. (2010). *EGRA Plus: Liberia. Program evaluation report.* Prepared for USAID/Liberia under the Education Data for Decision Making (EdData II) project, Early Grade Reading Assessment (EGRA): Plus Project, Task Order No. EHC-E-06-04-00004-00 (RTI Task 6). Research Triangle Park, NC: RTI International. http://pdf.usaid.gov/pdf_docs/pdacr618.pdf

Piper, B., & Mugenda, A. (2014). *USAID/Kenya Primary Math and Reading (PRIMR) Initiative: Endline impact evaluation.* Prepared under the USAID EdData II project, Task Order No. AID-623-M-11-00001 (RTI Task 13). Research Triangle Park, NC: RTI International. http://pdf.usaid.gov/pdf_docs/pa00k27s.pdf

Piper, B., & Zuilkowski, S. S. (2015). The role of timing in assessing oral reading fluency and comprehension in Kenya. *Language Testing* [online publication]. http://dx.doi.org/10.1177/0265532215579529

Prodigy Systems. (2011). *EGRA Yemen with iProSurveyor* [Presentation slides]. Sana'a: Prodigy Systems.

Pouezevara, S., Costello, M., & Banda, O. (2012). *Malawi National Early Grade Reading Assessment survey. Final assessment – November 2012.* Prepared for USAID under the Malawi Teacher Professional Development Support (MTPDS) program, Contract No. EDH-I-00-05-00026-02; Task Order No. EDH-I-04-05-00026-00. Washington, DC: Creative Associates International, RTI International, and Seward, Inc. http://pdf.usaid.gov/pdf_docs/PA00JB9R.pdf

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M.S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2,* 31–74.

Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research, 95,* 259–272.

RTI International. (2008). *Early grade reading Kenya: Baseline assessment. Analyses and implications for teaching interventions design. Final report.* Prepared for USAID under the EdData II project, Task Order No. EHC-E-01-04-00004-00 (RTI Task 4). Research Triangle Park, NC: RTI International. http://pdf.usaid.gov/pdf_docs/PNADL212.pdf

RTI International. (2011). *EGRA Plus: Liberia. Final report: October 2008–January 2011.* Prepared for USAID/Liberia under the EdData II Project, Task Order No. EHC--E-06-04-00004-00 (RTI Task 6). Research Triangle Park, NC: RTI International. http://pdf.usaid.gov/pdf_docs/PNADZ817.pdf

RTI International. (2014a). *Codebook for EGRA and EGMA* [Excel spreadsheet]. Research Triangle Park, NC: RTI. Retrieved from https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=389

RTI International. (2014b). *USAID/Kenya Primary Math and Reading (PRIMR) Initiative: Final report.* Prepared for USAID under the EdData II project, Task Order No. AID-623-M-11-00001. Research Triangle Park, NC: RTI. http://pdf.usaid.gov/pdf_docs/PA00K282.pdf

RTI International. (2015). *EGRA tracker.* Prepared for USAID under the EdData II project, Contract No. EHC-E-00-04-00004-00. Research Triangle Park, NC: RTI. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=188

RTI International & International Rescue Committee (IRC). (2011). *Guidance notes for planning and implementing EGRA.* Research Triangle Park, NC: RTI and IRC. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=318

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: the case of Arabic diglossia. *Applied Psycholinguistics, 24,* 115–135.

Scanlon, D. M., Gelzheiser, L. M., Vellutino, F. R., Schatschneider, C., & Sweeney, J. M. (2008). Reducing the incidence of early reading difficulties: Professional development for classroom teachers versus direct interventions for children. *Learning and Individual Differences*, *18*(3), 346–359. http://dx.doi.org/10.1016/j.lindif.2008.05.002

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134*(4), 584–615.

Share, D. L., Jorm, A., Maclearn, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Education Psychology, 76,* 1309–1324.

Share, D. L., & Leikin, M. (2004). Language impairment at school entry and later reading disability: Connections at lexical versus supralexical levels of reading. *Scientific Studies of Reading, 8,* 87–110.

Skaggs, G. (2005). Accuracy of random groups equating with very small samples. *Journal of Educational Measurement, 42,* 309–330.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Prepared on behalf of the Committee on the Prevention of Reading Difficulties in Young Children under Grant No. H023S50001 of the National Academy of Sciences and the U.S. Department of Education. Washington, DC: National Academy Press.

Snow, C., & the RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Research prepared for the Office of Educational Research and Improvement (OERI), U.S. Department of Education. Santa Monica, CA: RAND Corporation.

Spencer, L. H., & Hanley, J. R. (2003). Effects of orthographic transparency on reading and phoneme awareness in children learning to read in Wales. *British Journal of Psychology, 94*(1), 1–28.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–406.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers.* New York: Guilford Press.

Stern, J. & Nordstrum, L. (2014). *Indonesia 2014: The National Early Grade Reading Assessment (EGRA) and Snapshot of School Management Effectiveness (SSME) survey.* Prepared for USAID/Indonesia under the Education Data for Decision Making (EdData II) project, Task Order No. AID-497-BC-13-00009 (RTI Task 23). Research Triangle Park, NC: RTI International. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=680

Strigel, C. (2012). *Tangerine™—Electronic data collection tool for early reading and math assessments. January 2012 – Kenya field trial report: Summary.* Research Triangle Park, NC: RTI International. www.rti.org/files/tangerine_report_0112.pdf

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, *40*(1), 7–26. http://dx.doi.org/10.1016/s0022-4405(01)00092-9

United Nations. (2015). *The Millennium Development Goals report 2015.* New York: United Nations. http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf

United Nations Development Programme (UNDP). (2015). *Sustainable Development Goals (SDGs)* [Web page]. Retrieved from http://www.undp.org/content/undp/en/home/mdgoverview/post-2015-development-agenda.html

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2014). *Education for All Global Monitoring Report 2013/4. Teaching and learning: Achieving quality for all.* Paris: UNESCO. http://en.unesco.org/gem-report/report/2014/teaching-and-learning-achieving-quality-all#sthash.n1q0vitl.dpbs

United States Agency for International Development (USAID). (2012). *How-to note: Preparing evaluation reports.* Monitoring and Evaluation Series, No. 1, Version 1.0. Washington, DC: USAID. Retrieved from https://www.usaid.gov/sites/default/files/documents/1870/How-to-Note_Preparing-Evaluation-Reports.pdf

Valadez, J. J., Mulcahy-Dunn, A., & Sam-Bossman, E. (2014). *Using lot quality assurance sampling to monitor impact of early grade reading programs* [87-slide training presentation plus handouts]. Prepared under the EdData II project, Task Order No. AID-OAA-12-BC-00003 (RTI Task 20), for a USAID-hosted webinar based in Washington, DC, July 9–10, 2014. Research Triangle Park, NC: RTI International. https://www.eddataglobal.org/reading/index.cfm?fuseaction=pubDetail&ID=602

Vaughn, S., & Linan-Thompson, S. (2004). *Research-based methods of reading instruction grades K-3.* Alexandria, VA: Association for Supervision and Curriculum Development.

Wagner, D.A. (2011). *Smaller, quicker, cheaper: Improving learning assessments for developing countries.* Paris: UNESCO International Institute of Educational Planning (IIEP) and Fast Track Initiative/World Bank. http://unesdoc.unesco.org/images/0021/002136/213663e.pdf

Wagner R. K., Torgesen J. K., & Rashotte C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology, 30,* 73–87.

Walther, B., Hossin, S., Townend, J., Abernethy, N., Parker, D., & Jeffries, D. (2011). Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. *PLoS One, 6*(9), e25348. http://dx.doi.org/10.1371/journal.pone.0025348

Wang, M., Park, Y., & Lee, K. R. (2006). Korean-English biliteracy acquisition: Cross-language phonological and orthographic transfer. *Journal of Education Psychology, 98,* 148–158.

What Works Clearinghouse. (2015). *Procedures and standards handbook, version 3.0.* Washington, DC: Institute of Education Sciences, US Department of Education. http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf

World Bank. (2015a). *EdStats dashboards: Learning outcomes dashboard* [Web page]. Washington, DC: World Bank. http://datatopics.worldbank.org/education/wDashboard/tbl_index.aspx

World Bank. (2015b). *Learning outcomes* [Web page]. Washington, DC: World Bank. http://go.worldbank.org/GOBJ17VV90

World Bank: Independent Evaluation Group. (2006). *From schooling access to learning outcomes—An unfinished agenda: An evaluation of World Bank support to primary education.* Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/7083

Yesil-Dağli, Ü. (2011). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Childhood Research Quarterly, 26*(1), 15–29.

Yovanoff, P., Duesbery, L., Alonzo, J., & Tindall, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement, Fall*, 4–12.

Zieky, M., & Perie, M. (2006). *A primer on setting cut scores on tests of educational achievement.* Princeton, NJ: Educational Testing Service. https://www.ets.org/Media/Research/pdf/Cut_Scores_Primer.pdf

Zimmerman, R. (2008). *Digital data collection demonstration white paper. A comparison of two methodologies: Digital and paper-based.* Prepared for USAID under the Educational Quality Improvement Program 1 (EQUIP1), Cooperative Agreement No. GDG-A-00-03-00006-00. Washington, DC: American Institutes for Research. http://www.equip123.net/docs/e1-DigitalDataCollection.pdf

Zorzi M. (2010). The connectionist dual process (CDP) approach to modelling reading aloud. *European Journal of Cognitive Psychology, 22,* 836–860.

الملحق أ: معلومات عن ورش تقييم مهارات القراءة للمراحل المبكرة 2015

المصدر: الموقع الإليكتروني لمشروع "إدداتا 2"، أخبار وأحداث
https://www.eddataglobal.org/news/index.cfm

أ-1 ورشة عمل عن تصميم وتنفيذ تقييم مهارات القراءة للمراحل المبكرة: فهم الأساسيات

24 مايو 2015-كيلي بيتس

في الفترة من 2 إلى 4 مارس 2015، قام فريق العمل الفني لمعهد البحوث المثلثية بعمل ورشة عمل بعنوان "تصميم وتنفيذ تقييم مهارات القراءة للمراحل المبكرة: فهم الأساسيات". وقد تم استضافة الورشة من قبل شبكة القراءة العالمية في منظمة يو آر سي في بيثيسدا بولاية ميريلاند. وقد ترأس ورشة عمل تقييم مهارات القراءة للمراحل المبكرة التي تهدف لتعليم المشاركين سواء في قاعات الدرس أو على الإنترنت أساسيات تصميم وتنفيذ أداة تقييم مهارات القراءة للمراحل المبكرة، كيت باتشيلدير وأليسون بيفليبسن وسارة بوزيفارا.

وكان أساس المنهج المدرس في ورشة العمل دليل أداة تقييم مهارات القراءة للمراحل المبكرة والملاحظات الإرشادية. وقد تم إدخال المعلومات والممارسة الميدانية في العديد من الجلسات. وقد قام ريتشارد فورمارورمن مركز التعليم والبحث (EARC) بغانا وإيفا يوسف من مركز ميرياد للأبحاث بأندونيسيا _وكلاهما مقاول فرعي لمعهد البحوث المثلثية_ بمشاركة قصص وتجارب خاصة من سياقاتهم الشخصية.

وقد ترأست الجلسة الافتتاحية خبيرة التعلم بمعهد البحوث المثلثية مارجريت "بيجي" دوبيك ومديرة البحث بمعهد البحوث المثلثية أمبير جوف التي قدمت معلومات عن تاريخ لأداة تقييم مهارات القراءة للمراحل المبكرة ونظرة عامة له. وطوال فعاليات ورشة العمل، تلقى المشاركون التوجيه اللازم وتعلموا الكثير عن أفضل الممارسات، وشاركوا في الممارسات التفاعلية على عدة جوانب مختلفة من تصميم وتنفيذ أداة تقييم مهارات القراءة للمراحل

المبكرة. ومن بين الموضوعات التي تمت تغطيتها تصميم البحث وإطار عملية أخذ العينة وكذلك التكيف والإدارة وتسجيل الدرجات والتقاط البيانات وجمع البيانات الإليكترونية وتدريب جامعي البيانات واختيارهم وتقييمهم والإصدارات التجريبية وجمع البيانات ونشرها.

وقامت إلينا فينوجرادوفا من مركز تنمية التعليم (EDC) بتقديم عرضا موجزا عن التقاط البيانات إليكترونيا، وبعد ذلك تلقى المشاركين تدريبا عمليا على الممارسة على أدوات مثل "SurveyToGo" وTangerine®. وكذلك قام بن سيلا وكريستي فيلساك من الوكالة الأمريكية للتنمية بتقديم عرضا كجزء من ورشة العمل.

وأتاحت الجلسة الختامية التي كانت بعنوان "التنفيذ: التدريب هو الطريق للكمال" للمشاركين أن يجمعوا ويلخصوا المادة العلمية والمعلومات التي اكتسبوها من خلال الجلسات السابقة ليبدأوا في تخطيط وتصميم تقييم مهارات القراءة للمراحل المبكرة.

وقد تم تمويل الورشة من قبل أمر المهمة الخاص بمشروع "إدداتا2" التابع للمعونة الأمريكية بواشنطن والمسمى "دعم قياس وبحث استراتيجية التعليم الهدف الأول".

أ-2 الفريق الفني يشارك في تحسين جودة بيانات تقييم مهارات القراءة

16 يونيو2015-كيلي بيتس

استضافت شبكة القراءة العالمية يومين من جلسات النقاش حول موضوعات متقدمة تتعلق بتصميم وإدارة وتحليل وإعداد التقارير عن تقييم مهارات القراءة للمراحل المبكرة. وقدم الفريق الفني لمعهد البحوث المثلثية عرضا على كل موضوع من الموضوعات جنبا إلى جنب مع أعضاء اللجان من الخبراء من مختلف المنظمات. وقد تم عقد هذا الحدث في منظمة "يو آر سي". مايو 27-28 2015.

وقد تم تمويل هذا الحدث من قبل الوكالة الأمريكية للتنمية الدولية، وهي تسعى لوضع مبادئ توجيهية معززة لتصميم وإدارة وإعداد التقارير لتقييم مهارات القراءة للمراحل المبكرة. وهذه السلسلة من العروض والمناقشات تم تصميمها لكي تجمع المنظمات وأعضاء شبكة القراءة العالمية لمشاركة التجارب والخبرات. وقد حضر هذا الحدث 30 مشاركا بصفتهم الشخصية وشارك عددا كبيرا من مختلف المنظمات التي تطبق أداة تقييم مهارات القراءة للمراحل

المبكرة عن طريق شبكة الانترنت من خلال برنامج الندوات عبر الانترنت "ويب اكس". وكان المشاركون يملكون خبرة كبيرة (سواء حالية أو سابقة) في تخطيط و/أو إدارة و/أو إعداد التقارير تتيح لهم المشاركة في العروض التقنية والمناقشات الغنية التي ميزت الحدث.

واستغرقت كل جلسة ساعتين وتضمنت عروضا من المتحدثين الخبراء فضلا عن المناقشات التي تمت بين المتحدثين والمشاركين. وكل مناقشة ركزت على موضوع العرض السابق وعادة ما كانت تتألف من أسئلة للتوضيح واقتراحات وتعليقات عامة. وفي نهاية كل مناقشة كان ميسر الجلسات ينهي الجلسة بإلقاء الضوء على المجالات التي أجمع عليها كلا من المتحدثين والمشاركين أو المجالات التي تتطلب المزيد من المناقشة والاهتمام.

وفي افتتاح هذا الحدث يوم الأربعاء 27 مايو رحب فريق عمل منظمة يو آر سي والوكالة الأمريكية للتنمية بالحضور وقاموا بتقديم المقدمين. وغطت الجلسة الأولى أطر تصميم البحث وأخذ العينات. وقدم كريس كاميسكي من معهد البحوث المثلثية مات سولان (مركز أبحاث السياسة الرياضية) وإلينا فينوجرادوفا (مركز تنمية التعليم). وترأست المناقشة الميسرة ميليسا تشيابيتا من "سوشيال إمباكت".

وتضمنت الجلسة الثانية الخاصة بوضع أدوات تقييم مهارات القراءة مماثلة عبر اللغات عروضا من متحدثين خبراء قام بتقديمها مارجريت "بيجي" دوبيك من معهد البحوث المثلثية وكارول دي سيلفا من منظمة أف اتش آي360 وفتحي العشري من منظمة المشاركون الإبداعيون (كرييتف أسوشييتس). وقاد بوجا ريدي ناكامورا (من المعاهد الأمريكية للأبحاث AIR) المناقشة حول هذا الموضوع.

وركزت الجلسة الأخيرة في اليوم الأول للحدث على الموثوقية بين المقيمين. وقد قام بتقديمها سيمون كينج (من معهد البحوث المثلثية RTI) وجيف دافيز من (أنظمة الإدارة الدولية MSI) وعبد الله فردوس من (المعاهد الأمريكية للأبحاث AIR). وقام فتحي العشري بتسهيل المناقشة التالية بين المتحدثين والمشاركين.

وبدأ اليوم الثاني من الحدث _الخميس 28 مايو _بجلسة عن إعداد وتحليل البيانات الخاصة بتقييم المهارات الرياضية للمراحل المبكرة(EGMA)، وقام بتسهيلها أجيا زافيراكو من الشراكة العالمية للتعليم (GPE). وتضمن المتحدثون في هذه الجلسة سيمون كينج (من معهد البحوث المثلثية RTI) وإلينا فيناجرادوفا EDCوميليسا تشيابيتا (من منظمة سوشيال إمباكت (Social Impact).

وقام بتقديم الجلسة الثانية التي كانت عن مساواة أداة تقييم مهارات القراءة للمراحل المبكرة عبر التطبيقات بنفس اللغة جوناثان ستيرن من معهد البحوث المثلثية وجيف دافيز من أنظمة الإدارة الدولية MSI) وزاركو فوكميروفيك من (المعاهد الأمريكية للأبحاث AIR). وقام أللا بيريزنر من المجلس الأسترالي لأبحاث التعليم بقيادة المناقشة التي تلتها.

وقد قدم جيف دافيز وتوماس ألفاريس من (أنظمة الإدارة الدولية MSI) عرضا موجزا لمناقشة الهدف الأول وهو استراتيجية التعليم والتحسينات المنهجية المقترحة الحالية. وبعد العرض، وتولى بينيامين سيلا من (الوكالة الأمريكية للتنمية) تنظيم طرح الأسئلة وإعطاء الإجابات بين الجمهور والمتحدثين.

وتضمنت الجلسة الختامية التي قامت جيل ميكس من (تشيمونيكس) بتسهيلها، ثلاثة عروض عن كيفية عمل ملفات الاستخدام العام لمجموعات البيانات، وقام بتقديمها كريس كاميسكي وكيلي بيتس من (معهد الأبحاث المثلثية RTI) وتوماس ألفاريس من (أنظمة الإدارة الدولية MSI) وروجير ستانتون من (أوبتيمال سولوشنز).

وعززت جلسات النقاش التي جرت على مدار يومين النقاشات التقنية وأثارت عدة مسائل وقضايا من أجل مزيد من البحث والدراسة لموضوعات متعددة فيما يخص إدارة أداة تقييم مهارات القراءة للمراحل المبكرة. ويعتزم أعضاء الفريق المشارك العمل معا لتقديم التوصيات ووضع مبادئ توجيهية بشأن كل موضوع من الموضوعات التي تم عرضها. وقد تم الإشادة بهذا الحدث، وأجمعت الغالبية العظمى من المشاركين سواءا على شبكة الإنترنت أو من داخل قاعات الدرس على أن محتوى الجلسات كان مفيدا للغاية ومشجعا للمشاركة والتفاعل وأدى إلى المشاركة الفعالة في المناقشات.

الملحق ب: اعتبارات حجم العينة في تقييم القراءة في المراحل المبكرة

ب-1 المقدمة

يضع هذا الملحق بعض الاعتبارات الأساسية الخاصة بعينات تقييم القراءة في المراحل المبكرة. وهي مصممة لاطلاع موظفي وزارة التعليم والجهات المانحة، أو غيرها من الجهات الفاعلة المهتمة بإنشاء تقييم القراءة في المراحل المبكرة على حسابات ومتطلبات الحجم الخاصة بالعينات.

ب-2 منهج أخذ العينات

وسوف يؤثر المنهج المستخدم في أخذ العينات على متطلبات حجم العينة، ومع تساوي الأمور الأخرى؛ نجد أن اختيار الطلاب بشكل عشوائي من قائمة وطنية سيتطلب حجما أصغر من العينة بينما أن *العينات العنقودية* تطلب أحجاما أكبر من العينات إلى حد ما، وبالرغم من أن الأمر قد يبدو متناقضا؛ إلا أن اختيار عينات نقية بسيطة يعتبر مكلفا نسبيا بالمقارنة بمناهج أخذ العينات الأخرى، فعلى سبيل المثال؛ إذا حاول أحدهم اختيار عينة عشوائية بسيطة من 400 طفل، فقد يكون مضطرا إلى الذهاب لحوالي 400 مدرسة ثم اختبار طفلا واحدا من كل مدرسة مما سيزيد من مصاريف النقل والعمالة بشكل كبير.[29]

وبالإضافة إلى ذلك، فسيحتاج المرء بشكل أساسي إلى قائمة بكل أطفال المدارس في البلاد وأماكنهم للحصول على عينة عشوائية بسيطة من الأطفال، ومثل هذه القوائم غالبا ما تكون غير موجودة في معظم البلاد، أما بالنسبة للعينات العنقودية المدرسية، فيتم اختيار المدارس أولا ثم اختيار الطلاب داخل المدارس (المجموعات)، لذلك فأخذ عينات عشوائية من المدارس

[29] سيكون هناك حاجة إلى الذهاب فقط إلى حوالي 400 مدرسة لأنه تبعا للقرعة واعتمادا على إجمالي عدد المدارس في البلاد ربما يتم اختيار أكثر من طفل من نفس المدرسة. فعلى سبيل المثال، فعند اختيار عينة بسيطة عشوائية من 400 طفل في دولة بها 500 مدرسة فقط ففي الغالب سيتم اختيار أكثر من طفل من كل مدرسة بينما سيختلف الأمر إذا كان في الدولة 80000مدرسة على سبيل المثال

أولا ثم اختيار الأطفال يقلل من تكاليف السفر والوقت الذي يستغرقه ويلغي الحاجة للاعتماد على القوائم الوطنية للطلاب، وبما أن جزءا كبيرا من تكلفة المسوحات يتمثل في تكلفة الوصول للمدارس بالدرجة الأولى، لذلك قد يقوم القائم على هذه المسوحات باختبار أكبر كم من الطلاب في زيارة من يوم واحد كوسيلة لزيادة حجم العينة بتكلفة منخفضة نسبيا، (راجع الملحق C من هذا الدليل لمعرفة المزيد عن أخذ العينات العنقودية والمركبة.)

وقد أظهرت التطبيقات السابقة لتقييم مهارات القراءة للمراحل الأولى أنه من الممكن لمقيم واحد أن يقوم باختبار ما بين 4 و 10 أطفال في يوم دراسي واحد، وهذا يتوقف على عدد المهام الفرعية والأسئلة لكل طالب[30]، وإذا افترضنا، *فقط على سبيل المثال* ، اننا سنختار عينة من 15 طفلا من كل مدرسة ، فسيتطلب اختيار عينة من 400 طفل زيارة نحو 27 مدرسة - وهو دون شك خيار أكثر توفيرا من الحاجة إلى زيارة 400 مدرسة أو نحو ذلك، (قد تختلف العينة الفعلية المطلوبة من الأطفال في كل مدرسة بناء على خصائص كل بلد)، ولذلك، فمن المستحسن استخدام نهج العينات العنقودية،

ومع ذلك، فإن تطبيق النهج العنقودي يؤدي إلى فقدان الواقعية لأن الاختلافات بين الأطفال عادة ما تكون أقل داخل المدرسة الواحدة بينما تزيد هذه الاختلافات بين الأطفال الذي يمثل كل منهم مدرسة مختلفة، وهنا يلعب معامل الارتباط بين المجموعات دوره ، في أن الأطفال في المدرسة الواحدة عادة ما ينتمون إلى طبقة اجتماعية واحدة، أو تكون لديهم نفس نقاط القوة والضعف في اللغة ، أو يكون معلميهم بنفس الكفاءة أو يتعرضون لنفس الممارسات الإدارية بدرجة أكبر من الأطفال من مدارس مختلفة، وهكذا، فإن الاختلاف الحقيقي أو السكاني بين الأطفال لا يعطى نفس الأهمية إذا كان أحدهم يستخدم منهج أخذ العينات العنقودية ،وهكذا نجد أن توفير تكلفة النقل وتوفير العمالة يأتي على حساب الوصول لمعلومات عن التباين بين الطلاب، وبالتالي، ما لم يتم عمل بعض التعديلات ، سنفتقد الدقة بشكل ما، ويطلق على هذا الإجراء *تأثير التصميم* ويمكن استخدامه لتعديل حجم العينة بحيث يتم تعويض فقدان عنصر التباين والاختلاف الذي أدى له استخدام منهج أخذ العينات العنقودية،

[30] يعتمد عدد الأطفال الذين يمكن مقابلتهم على نسخة أداة تقييم مهارات القراءة للمراحل الأولى التي يتم تطبيقها وعدد اللغات التي يجري تطبيق التقييم عليها، وكذلك ما إذا كانت أداة تقييم مهارات القراءة للمراحل الأولى هي جزء من أبحاث أخرى تجري في المدرسة

هناك أربع بنود سيتم تضمينها في حساباتنا لحجم العينة، وهي:

1. *التباين* في درجات القراءة لدى الطلاب (أو أي متغير آخر في تقييم القراءة للمراحل الأولى عند الرغبة)، سواء التباين الكلي أو التباين في كل مدرسة على حدة أو التباين بين المدارس،
2. *عرض فترة الثقة* التي يحددها الباحث
3. *مستوى الثقة* الذي يحدده الباحث (عادة 95%)
4. *تأثير التصميم* الذي يسببه تطبيق منهج العينة العنقودية

ب-3 حساب حجم العينة لفترة ثقة ومستوى ثقة معين

قد يتم تمثيل حجم العينة المطلوبة بالمعادلة التالية:

$$n = 4\left(\frac{CLTvalue\ DEFT\ SD}{CI_{Width}}\right)^2$$

where:

n هو الحجم المطلوب للعينة

CLtvalue (قيمة فترة الثقة) هي القيمة-ت المرتبطة بمستوى الثقة الذي تم اختياره (عادة ما تكون 1،96 ل95%)

DEFTهو الجذر التربيعي لتأثير التصميم (DEFF)

SD هو الانحراف المعياري المقدر، وهو قياس التغير في المتغير المختار.

CI width= العرض المحدد من قبل الباحث لفترة الثقة والرقم4 مأخوذ من المعادلة الأساسية لفترة الثقة.[31]

كما نرى من هذه المعادلة، فزيادة مستوى الثقة وتأثير التصميم والتباين (كما يقيسه الانحراف المعياري) تعمل جميعا لزيادة حجم العينة المطلوبة (n)، وعلى العكس، أي زيادة في عرض فترة الثقة تقلل من متطلبات حجم عينة وأيضا تقلل من الدقة، من حيث التعريف.

[31] وهذه المعادلة مشتقة من المعادلة التقليدية لحساب فترة الثقة $X \pm CLtvalue\ \frac{SD\ DEFT}{\sqrt{n}}$ ويكون التعبير على يمين علامة ± هو العرض الأحادي الجانب. وهكذا يكون العرض الكلي الثنائي الجانب $Width = 2CLTvalue\ \frac{SD\ DEFT}{\sqrt{n}}$.وستقوم المعالجة الجبرية باستخلاص واحدة من المعادلة المستخدمة في النص الأساسي وستظهر سبب تحول 2 إلى 4

ومن أجل وضع توصيات حجم العينة، يتم حساب الجذر التربيعي لتأثير التصميم (DEFT هو الجذر التربيعي لDEFF) وكذلك حساب الانحراف المعياري (SD) باستخدام بيانات من بعض تطبيقات تقييم القراءة للمراحل الأولى السابقة.

ويتم حساب تأثير التصميم كما يلي:

$$DEFF = 1 + (clustersize - 1)ICC,$$

وبالتعويض في المعادلة السابقة يكون:

Clustersize هو متوسط حجم المجموعة (عدد الأطفال الذين انضموا للعينة في كل مدرسة[32])، و

ICC هي معامل التلازم بين المجموعات

وتزيد الزيادات في حجم المجموعة أو في معامل التلازم بين المجموعات من تأثير التصميم، فإذا قمنا بأخذ عينات من الطلاب باستخدام المعاينة العشوائية البسيطة إذا سيكون حجم المجموعة clustersize يساوي 1(طفل واحد من كل مدرسة) وسيكون معامل الارتباط بين المجموعات يساوي صفر لأنه لم يكن هناك طلاب أخرين في المدرسة ضمن العينة لمقارنتهم، وسيكون تأثير التصميم يساوي1. وهذا يعني أن المعاينة العنقودية لا تؤثر على التباين التقديري إذا كان حجم المجموعة cluster sizeيساوي 1 فقط.

ويعتبر معامل الارتباط بين المجموعات مقياسا لمدى التباين بين المدارس وكذلك داخل كل مدرسة على حدة. وأحد الوسائل البديهية للتفكير في هذا الأمر أنه يشير إلى أن احتمالية العثور على ملاحظتين متماثلتين في المجموعة ترتبط بإيجاد ملاحظتين متطابقتين تم اختيارهما عشوائيا. على سبيل المثال فإذا كان معامل الارتباط بين المجموعات يساوي 0،41 يشير إلى زيادة احتمالية إيجاد طالبين بنفس مستوى الطلاقة في القراءة بنسبة 41% في نفس المجموعة (المدرسة) أكثر من احتمالية إيجاد طالبين بنفس مستوى الطلاقة في القراءة تم اختيارهم بشكل عشوائي من أي مدرستين.

[32] بافتراض أن المدارس هي أول مكان لأخذ العينات. في بعض المسوحات، يتم أخذ العينات في البداية من المناطق الجغرافية.

هناك عدة مفاهيم لمعامل الارتباط بين المجموعات في الكتابات. فمعامل الارتباط بين المجموعات في هذا السياق يتبع الاستخدام في برنامج ستاتا ويتم حسابه كما يلي:

$$\mathrm{ICC} = \frac{MSE_{between} - MSE_{within}}{MSE_{between} + (clustersize1)\, MSE_{within}}$$

يقيس *MSEwithin* مدى التباين داخل كل مدرسة (أو مجموعة). وحسابيا، *MSEwithin* هي مجموع مربع الانحرافات بين كل طفل ومتوسط باقي المجموعة (المدرسة) مقسوما على عدد الأطفال الإجمالي ناقص عدد العينات العنقودية. ورمزيا

$$\mathrm{MSE}_{\text{between}} \frac{\sum_{j-1}^{\text{cluster}} n_j (\bar{\bar{X}}_j - \tilde{X})^2}{\text{cluster} - 1}$$

و

$$\mathrm{MSE}\textit{within} \frac{\sum_{j-1}^{\text{cluster}} \sum_{i \in j=1}^{n} (X_{ij} - \bar{X}_j)^2}{\sum_{j-1}^{\text{cluster}} n_j - \text{cluster}}$$

وبالتعويض في المعادلة السابقة : يكون

$\tilde{X}$ هو المتوسط الحسابي الكلي

j هو الرقم القياسي للمجموعات

i □□*j* هو الرقم القياسي للطفل القياسي (j) في المجموعة

$\bar{\bar{X}}_i$ هو المتوسط الحسابي للمجموعة (أو المدرسة) القياسية (j)

Cluster عدد المجموعات أو الرقم القياسي لآخر عينة

Nj هو حجم العينة القياسية أو الرقم القياسي لآخر عضو في العينة القياسية (j)

ويمكن استخدام إجراء تحليل التباين لحساب كل من MSEwithinو .MSEbetween باستخدام برنامج Excel

الملحق ب-1 يوضح مجموعة من التقديرات لكلا من معامل الارتباط بين المجموعات والجذر التربيعي لتأثير التصميم لبضعة حالات على وجه الخصوص وكذلك تأثير هذه المتغيرات على عدد المدارس (المجموعات) وحجم العينة الُإجمالي. ويفترض أن يكون الانحراف المعياري 29 في كل الحالات لأغراض التوضيح، وتم تحديد عرض فترة الثقة الإجمالي(عرض ثنائي الجانب) ليساوي 10، وتم استخدام مستوى ثقة يساوي 95%. أما معامل الارتباط بين المجموعات والجذر التربيعي لتأثير التصميم وحجم العينة هم قيم حقيقية من دراسات أداة تقييم مهارات القراءة للمراحل الأولى.

معامل الارتباط داخل المجموعات والجذر التربيعي لتأثير التصميم بشكل تقديري في مجموعة مختلفة من البلاد والصفوف الدراسية وموضح متوسط حجم المجموعة في كل حالة

الدولة	معامل الارتباط داخل المجموعات	الجذر التربيعي لتأثير التصميم	حجم المجموعة	n
الدولة أ، الصف الثالث	17،0	2،1	75،3	198
الدولة ب، الصف الثاني	22،0	3،2	20	698
الدولة ج، الصف الثالث	25،0	6،1	57،7	356
الدولة د، الصف الثالث	47،0	3،2	05،10	708
الدولة ه، الصف الثاني	48،0	8،1	35،5	416

المصدر: قام بعمل هذه الحسابات معهد البحوث المثلثية عن طريق مجموعة مختلفة من مسوحات أداة تقييم مهارات القراءة للمراحل الأولى

وتتأثر تأثيرات التصميم في الملحق ب-1 بمعامل الارتباط بين المجموعات وأيضا بحجم المجموعة، وكما نرى في معادلة الجذر التربيعي لتأثير التصميم، يؤثر كلاهما في تأثير التصميم. ففي الدولة ب، على سبيل المثال، نرى أن معامل الارتباط داخل المجموعات مرتفعا قليلا (2،3) حتى وإن كان معامل الارتباط داخل المجموعات منخفضا (0،22) لأن حجم المجموعة 20، وهكذا يقمع المرء كثيرا من التباين عن طريق أخذه العديد من الأطفال من مدارس محددة، وفي الدولة د كان سبب ارتفاع الجذر التربيعي لتأثير التصميم هو ارتفاع معامل الارتباط داخل المجموعات، وفي الدولة أ كان الجذر التربيعي لتأثير التصميم في أدنى مستوياته لأن كلا من حجم المجموعة ومعامل الارتباط داخل المجموعات كانا منخفضين، وكان تأثيرهما على حجم العينة المطلوب واضحا. وفي الدولة أ، سنحتاج فقط إلى عينة من 198 طفلا فقط (من حوالي 53 مدرسة)، بينما في الدولة د سنحتاج إلى عينة من 708 طفل من 70 مدرسة.

ب-4 حجم العينات الموصى به لفترات الثقة

في تحديد حجم العينة الفعلية الموصى بها، ستكون الاختلافات "الملموسة" بين صفوف الطلاب مطلبا معقولا، فعلى سبيل المثال، يجب تكون فترات الثقة الكلية ضيقة بما فيه الكفاية لدرجة لا تتداخل معها فترات الثقة للصفوف متجاورة، فإذا عرفنا أن متوسط فرق الدرجات بين الصفوف هو 14، سيكون عرض 14 معقولا.

إذا افترضنا أن العرض 14،0 ومعامل الارتباط 0،45 وحجم العينة 12 ومعدل الانحراف المعياري 29 إذا فحجم العينة العنقودية "الصحيح" سيساوي 409 طفل، لاحظ أن مع تقدم الصف ترتفع درجات الطلاب، ونتيجة لذلك تحتاج الصفوف الأكبر عموما إلى أحجام أكبر من العينات لتحقيق نفس مستوى الدقة،

ونظرا للاختلافات التي عادة ما تكون صغيرة جدا بين البنين والبنات_ في متوسط الأداء للمهام الفرعية لأداة تقييم مهارات القراءة للصفوف المبكرة (و / أو بالنظر إلى أن الفروق بين الجنسين تختلف اختلافا كبيرا بين البلاد، على عكس التطور الثابت للصف)، وبالنظر إلى معادلة حجم العينة، فإنه من الواضح أنه ستكون هناك حاجة*لعرض* صغير جدا لكشف الفروق بين الجنسين، وبالتالي سنحتاج لحجم العينة كبير جدا: حوالي 7000 طالب، يبدو أنه من الحكمة القبول بفكرة أنه من غير المحتمل أن تلتقط أحجام العينات المعقولة فروقا ملموسة إحصائيا بين الذكور والإناث، وهذا يسلط الضوء، بشكل عابر، على أهمية التمييز بين الفروق الكبيرة الجوهرية والفروق ذات الدلالة الإحصائية. وبشكل عام، إذا كان هناك أي فرق على الإطلاق بين اثنين من الشرائح السكانية، حتى لو لم يكن واضحا ومثيرا للاهتمام من الناحية الموضوعية، يمكن للباحثين أن "يجبروه" ليصبح ذا دلالة إحصائية من خلال الاعتماد على عينة ضخمة. وبعبارة أخرى، فإن الاختلافات الصغيرة الهامشية يمكن تحديدها لتكون ذات دلالة إحصائية عن طريق حجم العينة الكبير. والحكم الذي يمكن أن نصل إليه هنا، هو أن الفروق بين الجنسين في تقييم القراءة للصفوف الدراسية الأولى تميل إلى أن تكون صغيرة لدرجة أنها لا تظهر سوى في العينات الكبيرة جدا ذات الدلالة الإحصائية.

وفي نهاية عام 2015، تم إجراء تقييم القراءة للمراحل الأولى في العديد من الدول. وعندما يتم إجراء المزيد من التقييمات في البلاد التي اكتملت فيها التقييمات بالفعل، يمكننا استخدام بيانات ملف الاستخدام العام من التقييم الأول للحصول على تقديرات أكثر دقة للتقييم الجديد، ويمكن طلب بيانات ملف الاستخدام العام الخاص ببيانات التعليم لصانعي القرار2من الموقع الإلكتروني لبيانات التعليم لصانعي القرار(إدداتا2)،
(https://www.eddataglobal.org/datafiles/index.cfm?fuseaction=datafilesIndex).

فعلى سبيل المثال، استخراج معامل الارتباط داخل المجموعات والانحراف المعياري من مجموعة البيانات الخاصة بتقييم مهارات القراءة الذي أجري في زامبيا في عام 2012، أتاح تقديرا أكثر دقة لحجم العينة لتقييم مهارات القراءة للصفوف الأولي الوطني الذي أجري في عام 2014، ومن خلال برنامج "ستاتا" الإحصائي يمكن استخدام الأمر *loneway* لتحديد معامل الارتباط داخل المجموعات، ويمكن استخدام الأمر *summarize* لتحديد معدل الانحراف المعياري.

ويجب الوضع في الاعتبار الاختلافات الأساسية بين السكان الذين يستهدفهم تقييم مهارات القراءة للمراحل الأولى السابق والحالي. فإذا تم تقييم لغات أو مناطق مختلفة أو صفوف مختلفة، قد لا يكون معامل الارتباط بين المجموعات ومعدل الانحراف المعياري على المستوى المطلوب من الدقة في التقييمات المستقبلية. ولكن إذا أمكن التخلص من هذه العقبات، فسيتيح استخدام البيانات القديمة الفرصة لحساب تقديرات معقولة لحجم العينة. ويوضح **الملحق ب-2** كيف يتم تنويع عدد الطلاب الداخلين في العينة من 10 إلى 22(في نقطتين للزيادات) وتنويع عرض فترات الثقة بين 10 و12 و14 .وباستخدام قيم ثابتة تساوي26 لمعدل الانحراف المعياري و0،45 لمعامل الارتباط بين المجموعات يمكننا بعد ذلك حساب الجذر التربيعي لتأثير التصميم ومن ثم حساب عدد الطلاب والمدارس المطلوبة. وكما هو موضح في الملحق ب-2 نجد أن هناك قليل من المنفعة في جعل العينة تحتوي على عدد أكبر من الطلاب في كل مدرسة وعدد أقل من المدارس لأن معامل الارتباط بين المجموعات سيكون مرتفعا، فسوف يوفر عدد 10 طلاب من كل مدرسة و52 مدرسة نفس مستوى الدقة لعدد 22 طالب من كل مدرسة و49 مدرسة، بفارق 561 طالب (1086 ناقص 525). وهكذا، يمكن القول أنه بتقييم عددا أكبر من الطلاب يمكننا أن نذهب لعدد أقل من المدارس (أقل ب3مدارس فقط)، مما لن يعود بأي منفعة مادية دون شك.

الملحق ب-2. العدد التقديري للطلاب والمدارس المطلوبة بناء على تنويع أعداد الطلاب لكل مدرسة وكذلك عرض فترة الثقة وتثبيت معامل الارتباط بين المجموعات معدل الانحراف المعياري

القيم المتغيرة		القيم الثابتة		النتائج			
عدد طلاب العينة في المدرسة	95% ثقة	الانحراف المعياري	معامل الارتباط بين المجموعات	الجذر التربيعي لتأثير التصميم	مجموع عدد الطلاب		عدد المدارس
52	525	25،2	45،0	26	±5		10
52	618	44،2	45،0	26	±5		12
51	712	62،2	45،0	26	±5		14
50	805	78،2	45،0	26	±5		16
50	899	94،2	45،0	26	±5		18
50	992	09،3	45،0	26	±5		20
49	086،1	23،3	45،0	26	±5		22
36	364	25،2	45،0	26	±6		10
36	429	44،2	45،0	26	±6		12
35	494	62،2	45،0	26	±6		14
35	559	2.78	0.45	26	±6		16
35	624	2.94	0.45	26	±6		18
34	689	3.09	0.45	26	±6		20
34	754	3.23	0.45	26	±6		22
27	268	2.25	0.45	26	±7		10
26	315	2.44	0.45	26	±7		12
26	363	2.62	0.45	26	±7		14
26	411	2.78	0.45	26	±7		16
25	458	2.94	0.45	26	±7		18
25	506	3.09	0.45	26	±7		20
25	554	3.23	0.45	26	±7		22

ب-5 اختبار الفرضيات مقابل فترات الثقة: آثار أخذ العينات

عند اتخاذ القرار حول حجم العينة، يجب أن نأخذ عاملا واحدا في الاعتبار وهو ما إذا كان أساس المقارنات بين المجموعات (على سبيل المثال: بين مستويات طلاقة القراءة في الصفوف المختلفة) فترات ثقة غير متداخلة أم اختبارات الفرضيات أحادية الجانب. ومن الإجراءات الشائعة تقديم المربعات الصغرى المقيدة للمتغيرات الرئيسية وتوضيح أن المربعات الصغرى المقيدة الغير متداخلة خطوة أولى مفيدة لرؤية ما إذا كانت الاختلافات بين المجموعات واضحة أم لا.

وعادة ما يتم هذا لأن الباحث لا يعرف مقدما أيا من التناقضات أو اختبارات الفرضيات

سيكون مثيرا للاهتمام بشكل أكبر. وبهذا المعنى، نجد أن تقديم المربعات الصغرى المقيدة للمتغيرات الرئيسية في أداة تقييم مهارات القراءة للصفوف المبكرة، يبدو إجراء حكيما. وبالإضافة إلى ذلك، فبوجه عام، يهتم القراء المهتمين بالأمر بشكل كبير بالمعايير الفعلية التي يتم تقديرها (في المستويات المتوسطة من الطلاقة، على سبيل المثال) ومداها المحتمل وربما يهتمون بشكل أقل بما إذا كانت الاختلافات بين المجموعات السكانية المعنية واضحة بشكل إحصائي.

ولكن محاولة تضييق المربعات الصغرى المقيدة بشكل لا يجعلها تتداخل ومن ثم الكشف عن الفرق بين المتوسطات الحسابية يحتاج لأحجام أكبر من العينات. وقد يتطلب إجراء اختبارات الفرضية أحادية الجانب عينات أصغر حجما. ومن ناحية أخرى فاختبارات الفرضيات أصعب في التفسير ربما لأنها تجذب الانتباه بشكل زائد إلى "المعنوية الإحصائية" أو "الدلالة الإحصائية" وبعيدا إلى حد ما عن المعايير قيد البحث.

وعلاوة على ذلك، فيمكن تحقيق بعض التوفير في إجراء اختبارات الفرضية إذا كانت هذه الاختبارات أحادية الجانب.

وهناك بعض الجدال في الكتابات التقييمية على الشروط التي تبرر اختبار الفرضيات أحادي الجانب. وبالرغم أن الجدل لم يتوصل لنتيجة قاطعة، إلا أنه قد يكون من المفيد أن نشير إلى القضايا المطروحة.

بوجه عام، تقوم اختبارات الفرضيات بافتراض فرضية "العدم" أو الفرضية "الصفرية" التي على سبيل المثال (مستخدمين الطلاقة كمثال) تقول بأن مستوى الطلاقة لصف معين يساوي مستوى الطلاقة للصف السابق أو أن مستوى الطلاقة بعد التدخل هو نفسه قبل التدخل. ثم يتم افتراض فرضية بديلة. وأحد أشكال الفرضية البديلة هو أن مستوى الطلاقة في صف متقدم مختلف عن مستوى الطلاقة في صف سابق أو أن مستوى الطلاقة بعد التدخل يختلف عن مستوى الطلاقة قبل التدخل. ولاختبار هذه الفرضية، يمكن إجراء اختبار فرضيات "ثنائي الجانب". وهذا أمرا شائعا عندما يكون أحدهم مهتما بالتحليل الاستكشافي عندما يكون متغيرا معينا (مثل مدى ريفية المكان أو خبرة المعلم ...وهكذا) له تأثير إيجابي أو سلبي على شيئا آخر (فربما تتأثر درجات الاختبار بمدى ريفية المنطقة، ولا يملك الفرد سببا بديهيا قويا لاختبار فرضية تسير في اتجاه معين)وفي معظم تطبيقات تقييم مهارات القراءة للمراحل الأولى، من المنطقي أن يعتقد الفرد أن معظم الفرضيات يتم اختبارها أو أن معظم التصريحات التي يود إصدارها أحادية الاتجاه. لهذا يمكن للمرء أن يجري اختبار فرضيات أحادي الاتجاه للتوفير في حجم العينة. وإذا كانت هناك أسباب وجيهة للاعتقاد في أن التحليل يحتاج لأن يكون

ذا طبيعة أكثر استكشافية ووصفية، إذا يجب استخدام اختبار الفرضيات ثنائي الاتجاه.

وعند الرغبة، يمكن عرض فترات الثقة جنبا إلى جنب مع اختبارات الفرضيات. والغرض من تقديم المربعات الصغرى المقيدة هو تعزيز التركيز على المعيار المذكور مثل الطلاقة الشفوية في قراءة النصوص المتصلة.

لكن يجب ملاحظة أنه إذا كانت أحجام العينات كبيرة بما فيه الكفاية لتتيح الكشف عن الفروق في اختبارات الفرضيات أحادية الجانب، ستميل المربعات الصغرى المقيدة لتكون عريضة نسبيا. وبالتالي، فإن منهج تقييم القراءة للمراحل الأولى يقرر أولا ما إذا كانت اختبارات الفرضيات أحادية الجانب مقبولة أم لا، مع ملاحظة أن هذا قد يعني مربعات صغرى مقيدة أعرض قليلا. وتلقي المناقشة التالية الضوء على هذه الموضوعات.

فلنفرض أن لدينا متوسط عينة $\overline{X}_1$ و$\overline{X}_2$، ولتبسيط الأمر فلنفترض أن كليهما لهما نفس الأخطاء المعيارية التقديرية، وهكذا فالخطأ المعياري 1 = الخطأ المعياري2 = الخطأ المعياري (. SE□□ SE_2 □□SE_1). ونفترض أيضا، من دون خسائر كبيرة في العمومية، أن هذا يرجع إلى الانحرافات المعيارية المتساوية وأحجام العينات المتساوية. [1] ومن أجل هذه المناقشة سنحافظ على اختبارات 5% و95% من المربعات الصغرى المقيدة. ومن المفترض أن تكون الإحداثيات الرأسية لاختبار-ت للدرجات المناسبة من الحرية. وال95% من المربعات الصغرى المقيدة هي:

$$\overline{X}_1 \pm t_{.025}\ SE$$

$$\overline{X}_2 \pm t_{.025}\ SE\ ,$$

حيث t 0.25 هي الإحداثي الرأسي ت المطلوب لاختبار 5% ثنائي الجانب مع درجة ملائمة من الحرية. وشرط عدم تداخل المربعات الصغرى المقيدة لكل متوسط يساوي شرط

[1] وفي الواقع فإن معظم الانحرافات المعيارية والأخطاء المعيارية تختلف عن بعضها البعض. وافتراض حجم العينة ومساواة الانحرافات المعيارية في هذا الشرح فقط بغرض التبسيط

$$\overline{X}_1 + t_{.025}SE < \overline{X}_2 - t_{.025}SE$$

أو

$$\overline{X}_2 - \overline{X}_1 > t_{.025}SE + t_{.025}SE = 2t_{.025}SE$$

وإذا كان المتوسط التقديري الأول أصغر من الثاني وعلى نحو مماثل ولكن بعلامات مختلفة X الثاني أصغر أو بشكل أكثر عمومية:

$$\left|\overline{X}_1 - \overline{X}\right|_2 > 2t_{.025}\,SE$$

ولأن المربعات الصغرى المقيدة متماثلة حول المتوسط ولها نفس العرض، مع مراعاة تساوي الأخطاء المعيارية ودرجات الحرية (المأخوذة من n)

لكن شرط عدم تداخل فترة الثقة للفرق مع صفر مساوي لشرط

$$\left|\overline{X}_1 - \overline{X}_2\right| > 1.41\,t_{.025}\,SE$$

بسبب معادلة الانحراف المعياري للفرق بين المتوسطات، كما يلي، ونظرا لافتراض الانحرافات المعيارية

المتساوية والعينات المتساوية:

$$SD_{diff} = \sqrt{\frac{SD^2}{n_1} + \frac{SD^2}{n_2}} = \sqrt{2\frac{SD^2}{n}} = 1.41SD$$

لاحظ أن نسبة 2 إلى 1.41 هو 1.41، حيث أن أي رقم مقسوما على الجذر التربيعي له يساوي جذره التربيعي. وهذا يعني أنه في الحالة الأولى، قد نحتاج لخطأ معياري أصغر منه في الحالة الثانية لكي نتلافى تداخل المربعات الصغرى المقيدة-أصغر بنسبة 1.41. وإذا كان

الخطأ المعياري =الانحراف المعياري/ $\sqrt{n}$ ($SE = SD/\sqrt{n}$)، والخطأ المعياري الأصغر بنسبة1.41 يتطلب حجم عينة أكبر بنسبة الضعف كالتالي

$$\frac{SE}{1.41} \quad \frac{SD}{1.41\sqrt{n}} \quad \frac{SD}{\sqrt{2n}}$$

وتساعد الاختبارات الفورية التالية المأخوذة من برنامج "ستاتا" (باستخدام أمر تي تيست "ttesti") على الشرح. ومن أجل شرح المبدأ الأساسي المتعلق بالفرق بين فترات الثقة واختبارات الفرضيات، ونركز هنا على إحدى الحالات التي يكون فيها تأثير التصميم=1. *والإجراء* المستخدم هنا هو للتباينات غير المتساوية _بالرغم أنه في الواقع ولتبسيط الشرح، كانت مدخلات الانحرافات المعيارية في الإيضاحات متساوية. لاحظ أن أمر "ttesti" لا يمكن استخدامه لمعظم تحاليل تقييم القراءة للصفوف المبكرة لأنه لا يقوم بتعديل الأخطاء المعيارية لتفسير تصميم العينة العنقودية.

أولا، نرى هنا حالة لا تتداخل فترة الثقة للفرق بين المتوسطين مع الصفر ولكن تكاد أن تفعل كما هو موضح في الجزء الأسفل المظلل. لاحظ أن برنامج "ستاتا" يقدم المربعات الصغرى المقيدة لكل متغير وفترة الثقة للفرق بين المتغيرات وكذلك كل اختبارات الفرضيات ذات الصلة بالفرق بين المتغيرات

اختبار ت 25 20 29 25 34 29، غير متساوية

عينتين من اختبار ت بمتغيرات غير متساوية

	ملاحظة	المتوسط	الخطأ المعياري	الانحراف المعياري	(95% فترة ثقة)	
س	34	20	4،973459	29	9،881422	30،11858
ص	34	34	4،973459	29	23،88142	44،11858
جمع	68	27	3،593661	29،63409	19،82702	34،17298
- الفرق		14-	7،033533		-28،0429	0،042902

الفرق= متوسط (س)-متوسط (ص) ت t= 1،9905-

الفرضية الصفرية Ho: الفرق=0 درجات الحرية لساتيرثوايت= 66

الفرضية البديلة Ha : الفرق <0 الفرضية البديلةHa: الفرق !=0 الفرضية البديلة Ha: الفرق >0

القيمة الاحتمالية Pr ((T < t))= 0،0253 القيمة الاحتمالية Pr(|T| > |t|) =0.0507 القيمة الاحتمالية Pr ((T > t))= 0،9747

تتداخل المربعات الصغرى المقيدة بشكل كبير كما هو ملاحظ في الجزئيين العلويين المظللين، لكن فترة الثقة لا تتداخل مع الصفر (بالرغم من أنها تكاد تفعل عن طريق التصميم) كما يمكننا أن نلاحظ في الجزء السفلي المظلل. إلا أن هذه هي الطريقة الصحيحة لترجمة شرط الكشف عن الفرق بين المجموعات. ولتجنب تداخل المربعات الصغرى المقيدة للمتوسطات نفسها يمكن أن يقوم المرء بمضاعفة حجم العينة.

ويوضح الاختبار التالي أنه مع مضاعفة حجم العينة تكاد المربعات الصغرى المقيدة أن تبتعد عن التداخل كما هو موضح في الأجزاء العليا المظللة.

اختبار-ت 25 20 29 25 34 29، غير متساوية

عينتين من اختبار ت بمتغيرات غير متساوية

	ملاحظة	المتوسط	الخطأ المعياري	الانحراف المعياري	(95% فترة ثقة)	
س	69	20	3،49119	29	13،03344	26،96656
ص	69	34	3،49119	29	27،03344	40،96656
جمع	138	27	2،531281	29،73582	21،99457	32،00543
- الفرق		14-	4،937288		-23،76379	-4،236213

الفرق= متوسط (س)-متوسط (ص) ت t= 2،8356-

الفرضية الصفرية Ho: الفرق=0 درجات الحرية لساتيرثوايت= 136

الفرضية البديلة Ha : الفرق <0 الفرضية البديلةHa: الفرق !=0 الفرضية البديلة Ha: الفرق >0

القيمة الاحتمالية Pr ((T < t))= 0،0026 القيمة الاحتمالية (Pr(|T| > |t| =0.0053 القيمة الاحتمالية
Pr ((T > t)) =0،9974

لكن لاشك أن مضاعفة حجم العينة يعتبر ثمنا باهظا (وغير ضروري) من أجل الحصول على مربعات صغرى مقيدة غير متداخلة للمتوسطات أكثر من فترات ثقة غير متداخلة مع الصفر للفرق بين المتوسطات. ويمكن رؤية هذا من خلال حقيقة أن فترة الثقة للفرق بين المتوسطات بعيدة جدا عن الصفر (التظليل الأوسط)، أو من خلال حقيقة أن اختبار الفرضيات ثنائي الاتجاه للفرق بين المتوسطين يعطي قيمة احتمالية أقل بكثير من 5% (التظليل السفلي).

إلا أن للفرد حرية أكثر قليلا. وتتحقق زيادة كفاءة اختبارات الفرضيات عن مفهوم "فترات الثقة الغير متداخلة" غالبا إلى طرح المشكلة في صورة اختبار فرضيات. ولكن، عند الرغبة أو عند وجود مبرر بديهي، يمكن اكتساب مزيد من الكفاءة عن طريق افتراض اختبار فرضيات أحادي الجانب. لاحظ أنه في النموذج المطبوع الأول من برنامج "ستاتا" بالأعلى، بالرغم من

أن فترة الثقة للفرق تكاد تلمس الصفر إلا أن اختبار الفرضيات *أحادي الجانب* شديد القوة بالنسبة لاختبار 5%. لأن 95% فترة ثقة للفرق يكاد يلمس الصفر، وقيمة الاحتمالية لاختبار ثنائي الجانب تساوي بالفعل حوالي 0،05 (أو ما يقارب ذلك)، فكما قد يتوقع الفرد عندما يعطى المعادل بين اختبار الفرضيات ثنائي الجانب وفترة الثقة لفرق بين المتوسطات لا يحتوي على صفر. لكن القيمة الاحتمالية لاختبار الفرضيات أحادي الجانب في التجربة الأولى بالأعلى كانت تساوي فقط 0،025 (في الحقيقة 0،249)، وهكذا لدينا درجات من الحرية يمكننا أن نستغني عنها إذا كان كل ما نريد هو اختبار 5%.. وبما أن قيمة t في اختبار فرضيات 5% أحادي الجانب تساوي 1،67(أو ما يقارب ذلك، ل n الكبيرة) بينما تلك المطلوبة في الاختبار ثنائي الجانب يجب أن تساوي حوالي 1،96، ويمكن أن نجعل العينة أصغر بنسبة تصل لحوالي

$$\sqrt{\frac{1.67}{1.96}} = 0،73$$

وفي الواقع، فنحن نشترط فقط

$$\left|\overline{X}_1 - \overline{X}_2\right| > 1.41\, t_{05}\, SE$$

ولاختبار ت أحادي الجانب، ت (t)≈ 1.67 مع أي nمرتفعة بشكل معقول

ويوضح اختبار "ستاتا" الفوري التالي أنه عند تقليل حجم العينة في المجموعة الأولى من النتائج بنسبة 0،73 من 34 أو 25، إذا سيكون لاختبار الفرضيات أحادي الجانب قيمة احتمالية أقل بقليل من 0،05 حسب الحاجة (التظليل السفلي). والآن تتداخل المربعات الصغرى المقيدة تماما (التظليل العلوي). وفترة الثقة 95% للفرق يتداخل مع الصفر لأن اشتراط عدم تداخل فترة الثقة للفرق مع الصفر سيكون مساويا لاختبار الافتراضات الثنائي الجوانب.

اختبار ت 25 20 29 25 34 29، غير متساوية

عينتين من اختبار ت بمتغيرات غير متساوية

	ملاحظة	المتوسط	الخطأ المعياري	الانحراف المعياري	(95% فترة ثقة)	
س	25	20	5،8	29	8،029388	31،97061
ص	25	34	5،8	29	22،02939	45،97061
جمع	50	27	4،180518	29،56073	18،59893	35،40107
- الفرق		14-	8،202439		-30،49211	2،492108

الفرق= متوسط (س)-متوسط (ص) ت t= -1،7068

الفرضية الصفرية Ho: الفرق=0 درجات الحرية لساتيرثوايت= 48

الفرضية البديلة Ha : الفرق >0 الفرضية البديلةHa: الفرق !=0 الفرضية البديلة Ha: الفرق <0

القيمة الاحتمالية Pr ((T < t))= 0،0472 القيمة الاحتمالية Pr(|T| > |t|) =0.0943

القيمة الاحتمالية Pr ((T > t))= 0،9528

ومع أخذ كلا العاملين معا، نجد أن حجم العينة المطلوب لاختبار الفرضيات أحادي الاتجاه جوالي36،0من المطلوب لعمل مربعات صغرى مقيدة غير متداخلة (ثنائية الجانب) على المتوسطين.

لاحظ أنه إذا تمت مضاعفة معدل الانحراف المعياري بكفاءة عن طريق الجذر التربيعي لتأثير التصميم الذي يساوي 2،44 (نتيجة نفس الافتراضات التي استخدمت لوضع حجم العينة بحيث كانت فترة الثقة تساوي 409 أي معامل ارتباط بين المجموعات يساوي 0،45 وحجم مجموعات يساوي 12) إذا فحجم العينة المطلوب لاختبار 5% يزيد ليصل إلى 2،44² مضروبا في 25 أو 148.

اختبار ت 148 20 70،7 148 34 70،7، غير متساوية

عينتين من اختبار ت بمتغيرات غير متساوية

	ملاحظة	المتوسط	الخطأ المعياري	الانحراف المعياري	(95% فترة ثقة)	
س	148	20	5،811504	70،7	8،515112	31،48489
ص	148	34	5،811504	70،7	22،51511	45،48489
جمع	296	27	4،122578	70،92751	18،88661	35،1139
- الفرق		14-	8،218708		30،17496-	2،174957

الفرق= متوسط (س)-متوسط (ص) ت t= 1،7034-

الفرضية الصفرية Ho: الفرق=0 درجات الحرية لساتيرثوايت= 294

الفرضية البديلة Ha : الفرق <0 الفرضية البديلةHa: الفرق !=0 الفرضية البديلة Ha: الفرق >0

القيمة الاحتمالية Pr ((T < t))= 0،0448 القيمة الاحتمالية Pr(|T| > |t|) = 0،0895
القيمة الاحتمالية Pr ((T > t))= 0،9552

وتسمح هذه العوامل ببعض الاقتصاد في حجم العينة عن طريق اختبار أحادي الجانب بعكس فترات الثقة الغير متداخلة. إلا أن هناك ضغطا معاكسا وهو أخذ عنصر القوة في الاعتبار. فمع أخذ عنصر القوة في الاعتبار، بافتراض أن القوة تساوي 0،8 واختبار افتراضات 5% وتقديم مفهوم أن معدلات الانحراف المعياري قد تكون مختلفة، يكون حجم العينة لاختبار افتراضي أحادي الجانب هو

$$n = \frac{(SD_1^2\, DEFF + SD_2^2\, DEFF)(0.85 + 1.67)}{DIFF^2}$$

وبالتعويض في المعادلة السابقة يكون:

0،85 هو القيمة ت (t) لقوة تساوي 0،8 و

1،67 هو القيمة ت (t) لاختبار 5% (لدرجات حرية تساوي 60% وهو رقم منخفض نسبيا) و DIFF هو الفرق بين الدرجات

باستخدام نفس المعايير لفترة الثقة، أي تأثير التصميم يساوي 5،595 (الجذر التربيعي لتأثير التصميم يساوي 2،44) (وهذا يرجع لمعامل الارتباط بين المجموعات الذي يساوي 0،45 وحجم مجموعة يساوي 12) ومعدلات انحراف معياري تساوي 29 (بمعنى أنه بالنسبة لهذا المثال كانت معدلات الانحراف المعياري متساوية ، لكن استخدام المعادلة يتيح وجود معدلات انحراف معياري مختلفة) وفرق يساوي 14 يكون حجم العينة المطلوب 324. وفي الحالة الأكثر تشاؤما حيث تكون معدلات الانحراف المعياري تساوي 50 ولكن يكون مسموحا للفرق بأن يساوي 20، يكون حجم العينة المطلوب 472. وفي كلا الحالتين تجد أنهما أقل قليلا من ما نحتاجه لفترة ثقة تساوي 95%.

وهكذا يمكننا أن نستنتج من كل ما سبق أن الاختبارات ثنائية الجانب كانت أكثر ملائمة ، وهكذا تكون المعادلة الصحيحة هي:

$$n = \frac{(SD_1^2\, DEFF + SD_2^2\, DEFF)(0.85 + 2)}{DIFF^2}$$

وفي هذه الحالة وباستخدام نفس الافتراض بالأعلى، سيكون حجم العينة بمعدل انحراف معياري يساوي 29 هو 414 ومع معدل الانحراف المعياري الأكثر تشاؤما والذي يساوي 50 ولكن مع فرق يساوي 20سيكون حجم العينة المطلوب 603.

ب-6 ملخص أحجام العينات بناء على فترات الثقة واختبارات الفرضيات

الملحق ب-3 يلخص مجموعة من المقترحات حول أحجام العينات، إلا أن إذا كانت بيانات "تقييم مهارات القراءة للمراحل الأولى" القديمة لهذه الدولة متاحة، يكون لهذه البيانات الأولوية في الاستخدام لتقدير معدلات الانحراف المعياري و معامل الارتباط بين المجموعات الأكثر ملائمة. ويفترض هذا الملحق معدل انحراف معياري يساوي 29 ومعامل ارتباط بين المجموعات يساوي 0،45 (وهو أعلى ما وجد في دراسات "تقييم مهارات القراءة للمراحل الأولى") وحجم *المجموعة* (عدد الأطفال الذين ضمتهم العينة من كل مدرسة) يساوي 10. وفي حالة اختبار الفرضيات، يفترض أن تكون القوة تساوي 0،8. وفي كل حالة، كان يتم حساب عدد المدارس المطلوبة عن طريق قسمة حجم العينة على 10.

الملحق ب-3. ملخص أحجام العينة تبعا لعدة اعتبارات

	حجم العينة	عدد المدارس
مستوى ثقة 90%		
منهج فترة الثقة		
عرض ثنائي الجوانب للفترة:10	475	48
عرض ثنائي الجوانب للفترة:15	211	22
منهج اختبار الفرضيات أحادي الجانب:		
الحد الأدنى للفرق الذي تم الكشف عنه:10	390	39
الحد الأدنى للفرق الذي تم الكشف عنه:15	173	18
منهج اختبار الفرضيات ثنائي الجانب:		
الحد الأدنى للفرق الذي تم الكشف عنه:10	539	54
الحد الأدنى للفرق الذي تم الكشف عنه:15	239	24
مستوى الثقة 95%		
منهج فترة الثقة:		
عرض ثنائي الجوانب للفترة:10	680	68
عرض ثنائي الجوانب للفترة:15	303	31
منهج اختبار الفرضيات أحادي الجانب :		
الحد الأدنى للفرق الذي تم الكشف عنه:10	539	54
الحد الأدنى للفرق الذي تم الكشف عنه:15	239	24
منهج اختبار الفرضيات ثنائي الجانب		
الحد الأدنى للفرق الذي تم الكشف عنه:10	689	69
الحد الأدنى للفرق الذي تم الكشف عنه:15	306	31

Source: Calculated by RTI International.

ب-7 أخذ العينات والترجيح

بوجه عام، من أجل أخذ عينات من المدارس يعتبر منهج المعاينة مع احتمال تناسبي مع الحجم منهج المعاينة المفضل والأكثر استخداما. وباستخدام هذا المنهج يتم اختيار المدارس للمرحلة الأولى من المعاينة حيث يكون احتمال اختيار كل مدرسة يتناسب مع عدد الطلاب في المدرسة (أو الصف) مقسوما على عدد المدارس في المنطقة أو البلد المطلوبة. واحتمال اختيار الطلاب في المدرسة هو المرحلة الثانية من أخذ العينات، حيث تكون احتمالية اختيار كل طالب هو عدد الطلاب المطلوب اختيارهم مقسوما على عدد الطلاب بالمدرسة أو الصف. وبالتالي، فالاحتمالية الكلية لاختيار الطلاب هي حاصل ضرب الاحتمالين السابقين. بينما تقدم المراحل الفردية احتماليات اختيار مختلفة لوحدات المعاينة لهذه المرحلة، يقدم حاصل ضرب المرحلتين احتماليات اختيار كلية متساوية عندما يكون عدد الطلاب المختارين من كل مدرسة متساويا.

حتماليات الاختيار الكلية تساوي

احتمالية المرحلة الأولى x احتمالية المرحلة الثانية

كما يلي

$$\frac{\#\ of\ pupils\ in\ school\ \times \#\ of\ schools\ selected}{number\ of\ pupils\ in\ region/country} \times \frac{number\ of\ pupils\ selected\ in\ school}{number\ of\ pupils\ in\ school}$$

ويمكن اختصارها كالتالي:

$$\frac{\cancel{\#\ of\ pupils\ in\ school}\ \times \#\ of\ schools\ selected}{number\ of\ pupils\ in\ region/country} \times \frac{number\ of\ pupils\ selected\ in\ school}{\cancel{number\ of\ pupils\ in\ school}}$$

لتصبح

$$\frac{\#\ of\ schools\ selected\ \times number\ of\ pupils\ selected\ in\ school}{number\ of\ pupils\ in\ region/country}$$

والعناصر المرجحة في النهاية هي معكوس احتمالات الاختيار الكلية. وإذا كانت العناصر التي تم ترجيحها متساوية فيلي ذلك تساوي المتوسط المرجح مع المتوسط غير المرجح. ولكن بالرغم من أن هذا هو – نظريا - ما يجب أن يحدث، إلا أن هذا لا يحدث عادة بسبب قلة عدد الطلاب الذين تضمنتهم العينة في بعض المدارس والطبقية في المدارس واستبدال المدارس...الخ. ولكن بالرغم من ذلك فالمعاينة مع احتمال تناسبي مع الحجم ينتج عنها ترجيح بعض العناصر القريبة من بعضها البعض للتقليل من التحيز في اختيار العينات.

الملحق ج: المعاينة المركبة والعنقودية

بالنسبة للمسوحات التعليمية الكبيرة، لاشك أن اختيار الطلاب بشكل عشوائي من مجموعة سكانية مستهدفة (أي استخدام منهج المعاينة البسيطة) أمر غير عملي وغير فعال من حيث التكلفة. فاستخدام هذا المنهج يتطلب وجود قائمة حديثة تتضمن كل طالب في المجموعة السكانية المستهدفة ومعظم وزراء التعليم لايملكون مثل هذه المعلومات، وحتى إن كانوا يملكونها فلن يكون استخدام هذا النهج فعالا من حيث التكلفة عندما تقوم فرق التقييم بالسفر والانتقال لمدرسة بعينها من أجل تقييم طالب واحد أو اثنين فقط. لكن بدلا من ذلك يمكنهم أختيار المدارس بشكل عشوائي لتكون ضمن العينة ثم اختيار مجموعة من الطلاب من كل من المدارس المختارة وهذا دون شك أكثر عملية وفعالية من حيث التكلفة. ومنهج المعاينة هذا يطلق عليه "عينة معقدة على مستوى المدارس". وتكون معظم عينات تقييم مهارات القراءة للصفوف الأولى على مستوى المدارس. وعادة ما يتم اتباع هذا المنهج في اخذ العينات عن طريق أخذ العينات من المدارس ككل ثم أخذ العينات من الطلاب (وأحيانا يتم بداية أخذ العينات من المناطق الجغرافية كالأحياء على سبيل المثال ثم المدارس ثم الطلاب). وبغض النظر عن هذه المراحل الخاصة بعملية المعاينة يتم عمل نوعا من تجميع الطلبة في المدارس التي تضمها العينة.

وتتضمن العديد من دراسات أداة تقييم مهارات القراءة للصفوف الأولى *استدلال إحصائي*، أي أنه يتم اختيار عينة من الطلاب من مجموعة سكانية مستهدفة ويتم استخدام نتائج الطلاب الذين تضمنتهم العينة كتقديرات تساعد على استنتاج نتائج باقي المجموعة السكانية. ويسمى هذا المنهج *المنهج الاستدلالي البارامتري* وهو عادة ما يضم نوعين من التقديرات: 1) *التقدير بنقطة* أو قيم منفردة يتم حسابها من البيانات لتمثيل بارامترات غير معروفة و2) *تقديرات الدقة*-أو مجموعة القيم المحتملة (يتم تعريف التقدير بنقطة وتقديرات الدقة بشكل أوضح بالأسفل)

واستنادا إلى هذه التقديرات وكذلك درجات الحرية، يمكن حساب فترة ثقة تساوي 95٪ ويمكن للتحليل الإحصائي الرسمي أن يمضي قدما. لاحظ أن كلا النوعين من التقديرات تتأثر مباشرة بكيفية أخذ العينة.

وإذا لم يتم منهج أخذ العينة في الاعتبار، ستفترض البرامج الإحصائية أن الطلاب تم اختيارهم من قبل العينة العشوائية البسيطة. وسيتسبب هذا في أن تصبح كل التقديرات بنقطة الخاصة ببارامترات السكان متحيزة، كما أن ذلك سيجعل دقة كل التقديرات منخفضة بشكل غير ملائم. وعند اجتماعهما معا، قد يؤدي هذا بالباحث إلى الاستنتاج الخاطئ بأن هناك فروق ذات دلالة إحصائية بين المجموعات الفرعية في حين أنه لا توجد أي فروق كما هو مبين في **الملحق ج-1**.

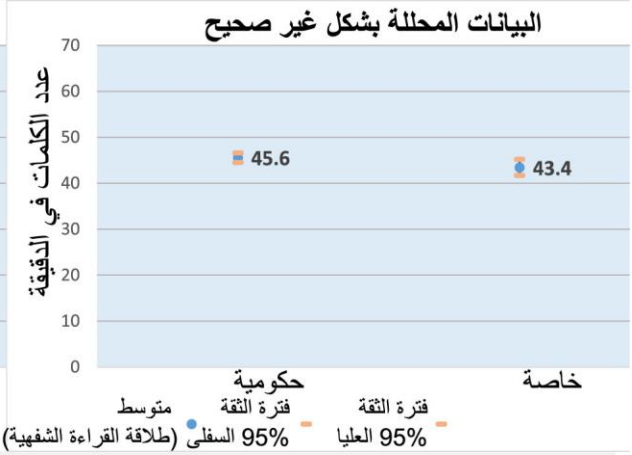
الملحق ج-1 بيانات محللة بشكل صحيح مقابل بيانات محللة بشكل غير صحيح

تم تحليل طلاقة القراءة الشفهية عن طريق نوع المدرسة. ويعرض الرسم البياني على اليسار البيانات التي تم تحليلها بعد إعداد العينة بشكل مناسب. أما الرسم البياني على اليمين فيعرض البيانات التي تم تحليلها دون إعداد العينة بشكل مناسب. أي أنه لم يتم ترجيح البيانات، وسمح الباحث لبرنامج التحليل الإحصائي بافتراض أنه الطالب قد انضم للعينة باستخدام منهج العينة العشوائية البسيطة.

ومتوسط تقديرات طلاقة القراءة الشفهية (عدد الكلمات في الدقيقة) في التحليل الغير صحيح متحيز للسكان ، لأنه يفترض أن 25%فقط من عدد الطلاب كان من منطقة جافا-بالي بأندونيسيا، بينما في الواقع 55%من عدد الطلاب كان من جافا-بالي. ولذلك تمنح العينة تمثيلا أقل لعدد الطلاب في جافا-بالي مما تسبب في تحيز متوسط التقديرات.

وعلاوة على ذلك، لم يأخذ برنامج التحليل في اعتباره أيا من تأثيرات التصميم التي تأتي مع العينات المركبة لأنه قد تم السماح له بافتراض المعاينة العشوائية البسيطة. وتسبب هذا في أن يصبح الخطأ المعياري منخفضا للغاية مما أدى إلى زيادة الثقة بشكل مبالغ فيه في متوسط التقدير المتحيز. وربما قام هذا الباحث باستنتاج خاطئ عن وجود فرق كبير وثابت في قدرة القراءة لدى الطلاب في المدارس الحكومية مقارنة بالمدارس الخاصة، في حين أن البيانات المحللة بشكل صحيح تشير إلى أن الأمر لم يكن كذلك.

البيانات المحللة بشكل صحيح
عدد الكلمات في الدقيقة
70
60
50
40
30
20
10
0
51.2
53.8
حكومية
خاصة
متوسط (طلاقة القراءة الشفهية)
فترة الثقة 95% السفلى
فترة الثقة 95% العليا


البيانات المحللة بشكل غير صحيح
عدد الكلمات في الدقيقة
70
60
50
40
30
20
10
0
45.6
43.4
حكومية
خاصة
متوسط (طلاقة القراءة الشفهية)
فترة الثقة 95% السفلى
فترة الثقة 95% العليا

الملحق د: المعاينة لتقييمات التأثيرات

يصف هذا الملحق كيفية استخدام حسابات القوة الإحصائية في تحديد العينة اللازمة لتقدير التأثيرات.

لا شك أن الانتقال من تقديرات النقطة لتقدير التأثيرات له تأثير كبير على اعتبارات حجم العينة. ويوضح حساب القوة الإحصائية كيف يمكن لتقييم ما أن يميز بين التأثيرات الحقيقية والفروق التي تحدث بمحض الصدفة. وهذا يساعد على الإجابة على سؤالين مترابطين:

1. **إذا كان لدينا حجم عينة ما، ما هو حجم التأثير الذي نحتاجه لكي يكون للدراسة فرصة معقولة لملاحظة فرق ذو دلالة إحصائية بين المجموعة المرجعية والمجموعة المعالجة؟** وهذا سؤال مناسب في حالة كان لدينا عينة معدة بالفعل ولا يمكن تغييرها. على سبيل المثال، ربما تقتصر دراسة ما على دراسة 200 مدرسة إجماليا، وستسمح حسابات القوة الإحصائية للباحث أن يحدد حجم التغيير المطلوب لملاحظته من أجل أن يكون واثقا من أن التغيير له دلالة إحصائية (أي حقيقي). في هذا الحالة، يبدأ الباحث بالعينة المتاحة ثم يقوم بحساب حجم التأثير المطلوب ليكتشفه بثقة.

2. **إذا كان لدينا حجم التأثير، فما هو حجم العينة المطلوب لكي نضمن إلى حد كبير أن الدراسة ستتمكن من كشف حجم التأثير الذي يجب أن ينتج عن التدخل؟** هنا إذا كان هناك مرونة في العينة يمكن أن تساعد حسابات القوة الإحصائية في تحديد حجم العينة المطلوب من أجل ملاحظة تغيير ذا دلالة إحصائية في نتيجة ملحوظة ذات حجم معين. على سبيل المثال قد يعرف الباحثون أو الوكالة الأمريكية للتنمية أو الشركاء المنفذون سواء من نظرية التغيير أو من دراسات لتدخلات مماثلة أنهم يتوقعون تغييرا معينا في النتيجة نتيجة للتدخل، وفي هذه الحالة سيبدأ الباحث بالتأثير المتوقع ثم يحدد أي عينة ستكون ضرورية لكشفها بثقة

ومن أجل تحديد حجم العينة المناسب للتقييم، يقوم المقيمون عادة بحساب الحد الأدنى للتأثيرات التي يمكن اكتشافها (MDI) والذي يمثل أصغر تأثيرات البرنامج الحقيقية- متوسط الفروق بين المجموعة المرجعية والمجموعة المعالجة- وهذا يمكن ان يؤدي إلى تقييم ذا دلالة إحصائية ذات احتمالية عالية إذا كنا نعرف حجم العينة. ومن المعتاد أن يتم توحيد الحد الأدنى من التأثيرات التي يمكن اكتشافها (MDI) في صورة وحدات حجم التأثير، أي مثل النسب المئوية

للانحراف المعياري لقياسات النتائج.

وفي هذه الحالة يطلق على الحد الأدنى الموحد للتأثيرات التي يمكن كشفها (MDI) حجم الحد الأدنى للتأثيرات التي يمكن كشفها (MDES). وتحجيم تقديرات التأثير في وحدات الانحراف المعياري ييسر المقارنة بين النتائج عبر النتائج التي يتم قياسها على مستويات مختلفة.

و رياضيا، يمكن التعبير عن معادلة الحد الأدنى للتأثيرات التي يمكن اكتشافها (MDI)على النحو التالي:

الحد الأدنى الموحد للتأثيرات التي يمكن كشفها (MDI) = العامل *SE (التأثير)

وبالتعويض في المعادلة يكون) "SE التأثير)" هو الخطأ المعياري لتقدير التأثير و"العامل" هو ثابت يمثل دالة الدلالة ومستويات القوة الإحصائية. ويزيد العامل بزيادة مستوى القوة ونقص مستوى الدلالة. وهكذا، فإن الحد الأدنى الموحد للتأثيرات التي يمكن كشفها (MDI) يرتفع عندما نسعى للحد من فرص عمل أخطاء من النوع1 و2. ويختلف الخطأ المعياري (التأثير) تبعا لتصميم تقييم التأثير. وبوجه عام تقلل العينات الأكبر من الخطأ المعياري (التأثير) وبالتالي يجعل الحد الأدنى الموحد للتأثيرات التي يمكن كشفها" (MDI) أكثر قوة". والقوة الأكبر مطلوبة لأن التقييم من المرجح أن يكتشف التأثيرات المعقولة بالرغم من أن القوة الأكبر قد تكلفنا ثمنا باهظا
ومن أجل الحصول على حجم الحد الأدنى للتأثيرات التي يمكن كشفها (MDES) نقوم بقسمة الحد الأدنى الموحد للتأثيرات التي يمكن كشفها (MDI) على)SD النتيجة) أي الانحراف المعياري لقياس النتيجة وهكذا تكون المعادلة:

MDES = MDI / SD(outcome)

والحد الأدنى للتأثيرات التي يمكن كشفها (MDES) هو دالة للخطأ المعياري لتقدير النتيجة ومستوى الدلالة المفترضة ومستوى القوة المفترض. ومستوى الدلالة هو احتمالية عمل خطأ النوع الأول ومثل هذا الخطأ هو نتيجة إيجابية خاطئة أي الاستنتاج الخاطئ بوجود تأثير في حين أن هذا غير صحيح. ومستوى التأثير التقليدي هو 5بالمئة. ومستوى القوة هو واحد ناقص احتمالية خطأ النوع الثاني، ومثل هذا الخطأ هو نتيجة سلبية كاذبة تفشل في اكتشاف تأثير موجود بالفعل.

وعادة ما تحاول التقييمات تحقيق 80 بالمئة من القوة. والهدف هو الحصول على حد أدنى صغير للتأثيرات التي يمكن كشفها (MDES) بحيث أنه إذا انتجت الدراسة تأثيرا أكبر من الحد الأدنى للتأثيرات التي يمكن كشفها (MDES) يطلق عليها في هذه الحالة ذات دلالة إحصائية ونستطيع الوثوق في صحتها. وكل الدراسات تحتوي على هذه العملية الحسابية مع توثيق المعادلة بدقة من أجل المساعدة على اتخاذ قرارات بشأن حجم العينة وتكوينها.

عوامل أخرى. هناك العديد من العوامل التي يجب أن يضعها المرء في اعتباره عند إعداد العينة لتقدير النقطة وتلك العوامل موجودة أيضا في حسابات القوة الإحصائية لتقديرات التأثير. وأحد العوامل ذات الصلة الوثيقة والتي تمثل أهمية كبيرة للباحثين الذين يسعون لتقدير التأثيرات باستخدام أداة تقييم مهارات القراءة للمراحل المبكرة ما إذا كانت العينة مجمعة. والعديد من الدراسات التي تستخدم هذه الأداة تستخدم العينات العنقودية أو المجمعة حيث يتم في البداية اختيار إما بعض المدارس أو المجموعات السكانية وتقسيمها لمجموعات مرجعية ومعالجة ثم يتم اختيار مجموعة من الأفراد داخل كل مجموعة من أجل اختبارهم.

وفي هذا الموقف، يفترض المرء أن الأفراد داخل كل مجموعة يتشاركون بعض أوجه التشابه. على سبيل المثال، فالأطفال المتواجدون في نفس الفصل بعيدا عن أية تدخلات_ غالبا ما يقوم نفس المعلم بتدريسهم. وهذا يقلل من التنوع الفردي داخل المجموعة الذي يقلل بدوره من مشاركة كل فرد في تقدير التأثير. ويطلق على قياس مدى ترابط نتائج الأفراد داخل المجموعات *بمعامل الارتباط داخل المجموعات (ICC)*

. وفي حالة كان *معامل الارتباط داخل المجموعات* أقرب إلى *1* فستكون لإضافة أفرادا أكثر تأثيرا ضعيفا أو معدوما على الحد الأدنى للتأثيرات التي يمكن كشفها (MDES) وبدلا من ذلك سيحتاج الباحث أن يضيف مجموعات إضافية لتقليل الحد الأدنى للتأثيرات التي يمكن كشفها (MDES). وفي الواقع فإنه في الدراسات العنقودية ذات معامل ارتباط بين المجموعات (ICC) مرتفع، يكون لإضافة أفراد إلى كل مجموعة الحد الأدنى من التأثير الإيجابي في أحسن الأحوال. ومن العوامل الأخرى التي تؤثر على حسابات القوة الإحصائية هو عدد التناقضات (أو المجموعات المعالجة أو مجموعات الدراسة) وما إذا كان الاختبار وحيد الذيل أو ثنائي الذيل (يسعى الاختبار وحيد الذيل إلى تقدير التأثير في اتجاه واحد فقط)

وفي النهاية، من المهم أن نتذكر أن تحليل المجموعة الفرعية سيكون له أثر على حجم العينة أيضا. فعلى سبيل المثال، قد يكون من الملائم أن نفهم كيف يمكن لتدخل معين أن يؤثر على الأولاد والبنات كل على حدة. وفي هذه الحالة، يتم عمل حسابات القوة على مستوى

المجموعات الفرعية. لاحظ أنه في هذه الحالة تكون كل مجموعة فرعية كبيرة بما فيه الكفاية لاكتشاف التأثيرات بين أعضاء المجموعة فقط في مجموعات المعالجة والمرجعية. بمعنى أنه كلما زاد عدد المجموعات الفرعية التي يقوم الباحث بتحليلها (أي التقسيمات المطلوبة من قبل الوكالة الأمريكية للتنمية أو بعض أصحاب المصلحة الآخرين) كلما كان احتاج الأمر لأحجام أكبر من العينات.

الملخص. وخلاصة الأمر أن الباحث سيقوم باستخدام المبادئ التالية لحسابات القوة الإحصائية لمعرفة حجم العينة.

1. كلما زاد حجم العينة كلما ارتفعت القوة (وبالنسبة للدراسات العنقودية كلما أزداد عدد المجموعات، كلما ارتفعت القوة)
2. ترتفع القوة عندما يكون الانحراف المعياري للنتيجة صغيرا أكثر من ارتفاعها إذا كان الانحراف المعياري كبيرا.
3. وكلما زاد حجم التأثير كلما صار من المرجح أن يجد التقييم تأثيرا ذا دلالة
4. هناك مفاضلة بين مستوى الدلالة والقوة، فكلما انخفض مستوى الدلالة كلما انخفضت القوة.
5. ترتفع القوة في الاختبار وحيد الذيل أكثر من الاختبار الثنائي الذي لطالما كان الاتجاه المفترض صحيحا.

ملحق هـ: تقييم الجودة التقنية للأداة

من المهم أن نقوم بتقييم الجودة التقنية لأي أداة يتم استخدامها لقياس إنجاز الطالب. وينطبق هذا على أداة تقييم مهارات القراءة للمراحل المبكرة. وتستمد الإجراءات في المتبعة لإجراء هذه الفحوصات من مجال القياس النفسي. وكما هي العادة، ركزت هذه الإجراءات على مفهومين أساسيين: الموثوقية والصلاحية. ويجب أن تتضمن الفرق التي يقوم بإجراء تقييم مهارات القراءة للمراحل المبكرة شخصا على دراية بالقياس النفسي ليقوم بعمل الفحوصات الضرورية. ويتم شرح وتوضيح الموثوقية والصلاحية في القسم 9 . 1. 2. وفيما يلي أنواع إضافية من التحليلات التي يمكن استخدامها لإرساء موثوقية وصلاحية الأدوات.

هـ- 1 اختبارات الموثوقية

منهج الاختبار وإعادة الاختبار هو إجراء بديل لاختبار الموثوقية. هذا النهج، والذي يمكن أن يتم كجزء من تجريب أداة تقييم مهارات القراءة للمراحل المبكرة، ينطوي على تطبيق هذه الأداة على نفس المجموعة من الطلاب في وقتين مختلفين (يفصل بينهما أسبوع أو نحو ذلك). ويجب أن يمثل الطلاب الذين يتم اختيارهم السكان المستهدفين من حيث المواصفات الرئيسية مثل النوع والسن والحالة الاجتماعية والاقتصادية وظروف النشأة والقدرات المعرفية وهكذا. ويمثل معامل الموثوقية للاختبار وإعادة الاختبار الارتباط بين نتائج الطلاب في المرتين التي تم فيهما التقييم. ويتم إنشاء هذه الارتباطات بشكل مثالي من خلال ملخص الإجراءات للمهام الفرعية (نسبة الصواب، الطلاقة، الخ) لأنه إذا تم حساب هذه الارتباطات على مستوى البنود منفردة داخل المهام الفرعية سيظهر نفس الانحياز المتزايد الذي يظهر عندما يكون تقييم مهارات القراءة للمراحل المبكرة موقوتا في "الاختبار وإعادة الاختبار" كما في ألفا كرونباخ.

ويمكن توضيح الانحياز المتزايد من خلال مثال عددي. فلنفرض أن لدينا نتائج طفلين أو نتائج نفس الطفل ولكن مع وجود فاصل زمني_ كما ورد في المرفق هـ-1 حيث ترمز الأرقام صفر وواحد إلى "غير صحيح" و"غير صحيح" على التوالي. في كل من الحالتين لم يتمكن الطفل من الاستمرار بعد الكلمة الخامسة، وإذا حاول المحلل أن يحسب معامل الارتباط للكلمات من 1 ل5 سيكون معامل الارتباط منخفضا جدا (0،17). في حين إذا اعتبر المحلل الكلمات بعد الكلمة الخامسة غير صحيحة ثم قام بحساب معامل الارتباط لل10 كلمات جميعها سيكون

معامل الارتباط 0،38 وهو أعلى بكثير ولكنه مستمد بشكل غير صحيح لأن كل الأصفار بعد الكلمة الخامسة قد عززت معامل الارتباط بشكل مصطنع.

المرفق هـ-1 نتائج المهام الفرعية للعينة لحساب الانحياز المتصاعد

الكلمة	الطفل الأول	الطفل الثاني(أو الطفل الأول بعد فترة من الزمن)
1	0	**0**
2	1	**1**
3	0	**1**
4	1	**0**
5	0	**0**
6	0	**0**
7	0	**0**
8	0	**0**
9	0	**0**
10	0	**0**

وهناك اعتبارين يجب أن نكون مدركين لهما عند استخدام منهج الاختبار وإعادة الاختبار للاختبارات التعليمية:

- الأول: إذا كانت النقاط الزمنية متباعدة للغاية فهناك احتمال أن الطلاب قد تعلموا كما كبيرا من المعلومات وأن عدم الموثوقية سوف يكون في الواقع مقياسا لتحسن أداء الطلاب.

- الثاني: تقليل مقدار الوقت بين مرات إجراء التقييم يمكن أن يقلل من تأثير التعلم لكن يزيد من سهولة التأثر بتأثيرات التأجيل. أي أن درجات التقييم الثاني الذي تم إجراؤه تتأثر بحقيقة أن الاختبار قد أجري لنفس الطلاب مؤخرا

وهناك إجراء آخر لقياس الثبات أو الموثوقية وهو **ثبات الأشكال المتوازية.** ويستخدم هذا المنهج شكلين متماثلين من أداة تقييم مهارات القراءة للمراحل المبكرة. وفي هذه الحالة يقوم هذا الإجراء باستخدام الشكل الأول لاختبار كل طالب ثم يتبع ذلك استخدام الشكل الثاني لاختبار نفس الطلاب. وينصح بعكس ترتيب تطبيق الأشكال على النصف الثاني من المجموعة المختارة. ويوفر لنا معامل الارتباط بين مجموعتي النتائج قياسا لدرجة الثبات أو الموثوقية لنتائج تقييم مهارات القراءة للمراحل المبكرة عبر أشكال الاختبار. وتزداد أهمية هذا المنهج عندما يتم إنشاء أشكال متعددة من التقييم لقياس الدرجات عند نقاط زمنية متعددة (مثل النقطة الأساسية، النصفية، والختامية). ومع ذلك فمن المهم أن نتذكر أن وجود ارتباط قوي بين شكلي

الاختبار لا يضمن تساويهما أو أن عملية المساواة لن تكون ضرورية. وينطبق نفس التحذير الخاص بتطبيق الارتباطات على أي شيء آخر غير القياسات الموجزة للمهام الفرعية هنا أيضا.

هـ.2 اختبارات الصلاحية

وتتعلق الأدلة المرتبطة بالمعيار بمدى قوة العلاقة (الارتباط) بين نتائج اختبار تقييم مهارات القراءة ونتائج الإجراءات الأخرى خارج هذا الاختبار. وبوجه عام سيشمل هذا الاختبار على النظر إلى العلاقة بين نتائج تقييم مهارات القراءة للمراحل المبكرة ونتائج بعض الإجراءات ذات معيار معين الذي من المتوقع أن يتوقعه الاختبار (مثل نتائج الفهم القرائي في الصفوف التالية)، وكذلك العلاقات باختبارات أخرى التي من المفترض أن تقيس نفس البنيات أو ما يتعلق بها (مثل نتائج الطلاب الخاصة باختبارات مهارات القراءة المبكرة الأخرى).قد يتم جمع البيانات عن هذه الإجراءات في نفس وقت جمع بيانات تقييم مهارات القراءة للمراحل المبكرة، ويمكن أن يتم جمعها في وقت لاحق (لكن يتم جمع هذه البيانات من نفس الطلاب). وهذا النوع من دليل الصلاحية سيكون من الصعب جمعه في البلاد التي لا تمتلك الكثير من القياسات المعيارية لنتائج تعلم الطلاب. ولكن بالرغم من هذا يجدر الأخذ في الاعتبار أن البحث الواسع النطاق في البلاد الأخرى قد أثبت أن الأدوات من نوع تقييم مهارات القراءة للمراحل المبكرة تظهر علاقات قوية (0،7 واعلى) لأنواع الإجراءات الخارجية المقدمة كأمثلة في هذه الفقرة.

ويوصي بعض مطوري الاختبارات أن يتم جمع نوعا إضافيا من الأدلة كجزء من إقرار صلاحية الاختبار: **دليل على النتائج المترتبة على استخدام نتائج الاختبار** على المتقدمين للاختبار وأصحاب المصلحة الآخرين. ويشمل هذا الإجراء جمع البيانات لتحديد ما إذا كانت الآثار المفيدة المرجوة من الاختبار تتحقق أم لا (على سبيل المثال، في حالة تقييم مهارات القراءة للمراحل المبكرة، تشمل الفوائد المرجوة تزويد صانعي القرار بنتائج على مستوى النظام بشأن مهارات القراءة المبكرة بحيث يمكنهم استهداف الموارد والتدريب على نحو أكثر فعالية). كما أنه يتضمن أيضا جمع الأدلة عن أية عواقب سلبية غير مقصودة ناتجة عن استخدام نتيجة الاختبار (على سبيل المثال، معاقبة المدارس التي تؤدي بشكل سيئ في تقييم مهارات القراءة للمراحل المبكرة من خلال حجب الموارد عنهم) واتخاذ خطوات لمنع تكرار هذه النتائج السلبية.

الملحق و: توصيات وإرشادات للمقارنات بين اللغات

F.1 توصيات حول طبيعة نظم الكتابة

للمساعدة في إجراء مقارنات معقولة بين اللغات، يتعين على القائمين بتطويع أداة EGRA امتلاك فهم عميق لخصائص نظم الكتابة للغة التي يعملون بها.

ولتحسين جودة المقارنات بين اللغات، علينا ان نعرف ما إذا كان نظام للغة التي يجرى التقييم فيها لغة مقطعية او أبجدية (اللاتينية والأبجدية اللاتينية) أو alphasylabic (مجموعة من الحروف تمثل صوتا واحدة)

مجموعة الارشادات التالية موصى بها وفقا لنوع اللغة.

F.1.1 اللغات الرومانية اللاتينية

عند التعامل مع اللغات الابجدية الرومانية:

1. حدد ما إذا كان العمق الكتابي للغة موضع العمل شفاف أم غير شفاف.

- تشير الابحاث إلى أن الاطفال الذين يتعلمون القراءة في اللغات غير الشفافة قد يتعلمون فك الرموز أسرع من الأطفال الذين يتعلمون القراءة في لغات ذات نظام كتابي عميق (Spencer & Hanley, 2003). عمق نظام الكتابة مرتبط أيضا بدرجة سرعة أو بطء تنمية الفهم. (E.g. Share, 2008).

2. اعرف التركيب المقطع في اللغة موضع الدراسة

- اللغات ذات التركيب المقطعي المعقد (صامت – حركة طويلة – صامت – صامت مثل كلمة بنتْ مثل او في الانجليزية صامت – صامت – متحرك – صامت – صامت مثل كلمة starts يستغرق وقتا اطول في التعلم أكثر من اللغات ذات المقاطع البسيطة

3. أعرف تأثير طول الكلمة على المقارنات بين اللغات

4. التعرف على الكلمات القصيرة أسرع من التعرف على الكلمات الطويلة. على سبيل المثال اللغات الاشتقاقية – التي تربط مورفيمات عديدة – واللغات غير الاشتقاقية.

- اعرف العلامات المكتوبة بالنسبة للغات اللحنية والتي تؤثر على الفهم في حين أن ذلك امر غير مهم بالنسبة للغات غير اللحنية.

F.1.2 اللغات المقطعية الابجدية

مثل اللغة الهندية والتايلاندية

1. اعرف عدد المكونات الصوامت والصوائت (الحركات الصوتية) في كل عنقود مقطعي (مثل *akshara*) read (Nag & Perfetti, 2014) وهذا يؤثر على القدرة على القراءة.
2. اعرف أن نوع الحركة الصوتية تؤثر على القدرة على القراءة. (Nag, 2014)
3. اعرف ان استقامة المكونات الصوتية داخل عنقود المقطع سيؤثر على القدرة على القراءة بسهولة.
4. اعرف أن عدد سنوات التدريس المطلوبة للوصول الي نفس المستوى من الطلاقة في اللغات الاسيوية والشرق اسيوية حوالي خمس سنوات وذلك بسبب وجود مجموعة كتابية كبيرة والتي يتعين اكتسابها، وذلك مقارنة بثلاث سنوات في اللغة الإنجليزية.
(Nag, 2007)

F.1.3 اللغات الابجدية غير الرومانية

مثل العربية والعبرية

1. اعرف ما إذا كان العمق لنظام الكتابة ضعيف (كلمات اللغة العربية والعبرية بها أصوات متحركة) او عميقة (كلمات العربية والعبرية ليس بها اصوات متحركة)

2. توجد ثنائية لغوية في اللغة العربية واضحة.

- الثنائية اللغوية تشير الي الحالة التي يكون فيها نوعان من اللغة مستخدمة لوظائف اجتماعية محددة. الوضوح الوظيفي اللغوي الاجتماعي والفاصل اللغوي (الفنولوجي والمرتبط بالعاني والمعجمي) بين صورتي اللغة العربية من الواضح كما يُعتقد أنه يعوق أو على الاقل يبطئ من الاكتساب الأولي للقراءة (Abu- Rabia, 2000; Ayari, 1996) .

3. ترتبط حالة الثنائية اللغوية للغة العربية بالعمق الكتاب والطلاقة في القراءة.

4. ترتبط الأصوات المتحركة (الصوائت) بالصوامت.

5. أوضحت الدراسات التي اجريت حول قراءة الكلمات في اللغة العربية واللغة العبرية لا ترتبط بالدقة في القراءة. .(Saiegh-Haddad, 2003)

6. أشكال الحروف لها اهمية في بعض نظم الكتابة (على سبيل المثال (اللغة العربية) حيث لا يرى الاطفال اشكال الحروف منفصلة.

F.2 توصيات بخصوص اللغة الشفوية:

توصيات بخصوص اللغة الشفوية:

بغض النظر عن الرغبة في إجراء مقارنات عبر اللغات ،لابد أن تطويع أداة EGRA يراعي النواحي المتعددة للغة الشفوية مثل الاختلافات بين اللهجات و ووجود الثنائية اللغوية ووضوح التعليمات و مستويات صعوبة المحتوى للمهام الفرعية للوعي الصوتي و الاستماع و المفردات.

وبالنسبة للذين يركزون على المقارنات بينن اللغات، فمن المهم:

التأكد من أن قطع القراءة الشفوية في اللغات المختلفة لها مستوى مشابه من الصعوبة.
التأكد من أن المفردات تقيس نفس معنى الكلمة أو التركيب في كلا اللغتين.

F.3 توصيات حول معرفة قواعد الكتابة

يمكن لمحتوى المهام الفرعية المصممة لتقيس المعرفة بقواعد الكتابة أن تكون مضبوطة حتى يمكن ان يكون هناك درجة من امكانية المقارنة بين اللغات.

تتبع المقارنات بين اللغات معدل الطلاقة لدى التلاميذ الذين يجري تقييمهم في اللغات المختلفة حيث يقومون بتعرف بنود ملائمة للصف الدراسي حسبما يحدد التكرار الموجود في كتب الصف الخاصة بهم.

F.4 توصيات حول قراءة النص المتصل

ضمان الكفاية الفنية والمقارنة الأساسية لقطع القراءة لنص متصل في تطبيقات التقييم في اللغات المتنوعة يتطلب عدة اعتبارات:

1. القطعة تكون نصًا أصليًا أُعد خصيصًا للتقييم.
2. تناقش القطعة موضوعًا يناسب أعمار التلاميذ في شكل مألوف للنص للحد من تأثير الخلفية المعرفية على الفهم.
3. يحتوي النص في كلا اللغتين العناصر الفنية للقصة وموضوعا مألوفا لكلا المجموعتين في اللغتين وذلك للوصول إلى أفضل مقارنة بين اللغات.
4. تتجنب القطعة استخدام الكلمات الغامضة مثل:

- كلمة لهاء هجاء معين يمكن أن تحتمل أكثر من معنى.
- كلمة لها أكثر من هجاء وتعطى نفس المعنى.

F.5 توصيات حول متعلمي اللغة الثانية أو المتعلمين متعددي اللغات:

عند عقد المقارنات بين اللغات، تأكد انها تتم بين نفس التصنيف اللغوي. على سبيل المثال، أذا عقد اختبار بين مجموعة من متحدثي اللغة الانجليزية فقط أو متحدثي اللغة الانجليزية كلغة اولى، فالمقارنة لا تتم مع مجموعة اللغة الانجليزية كلغة ثانية.

الاكتساب التلقائي للغة (تعلم لغتين أو أكثر من الميلاد أو سن مبكر) ممكن ولذلك يمكن أن يكتسب الطفل لغتين كلغة أولى.

يوجد إمكانية لنقل المهارات (معظم مهارات فك الرموز يمكنها أن تنتقل بين نظم الكتابة المتشابهة) عندما يقرأ الأطفال في لغة إضافية اخري غير لغته الأصلية.

1. إذا كان لطفل يتعلم لغة ثانية (او فيما بعد) دون تدريس كاف في اللغة الاولى، تشير نتائج الدراسات إلى أن الطفل قد يستغرق وقتًا أطول للوصول إلى الكفاية في القراءة في هذه الحالات.

مراجع الملحق

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing: An Interdisciplinary Journal, 13,* 147–157.

Ayari, S. (1996). Diglossia and illiteracy in the Arab world. *Language, Culture and Curriculum, 9,* 243–253.

Ferguson, C. A. (1959). Diglossia. *Word, 15,* 325–340.

Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading, 30*(1), 7–22.

Nag, S. (2014). Akshara-phonology mappings: the common yet uncommon case of the consonant cluster. *Writing Systems Research, 6,* 105–119.

Nag, S., & Perfetti, C. A. (2014). Reading and writing: Insights from the alphasyllabaries of South and Southeast Asia. *Writing Systems Research, 6*(1), 1–9.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: the case of Arabic diglossia. *Applied Psycholinguistics, 24,* 115–135.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin, 134*(4), 584–615.

Spencer, L. H., & Hanley, J. R. (2003). Effects of orthographic transparency on reading and phoneme awareness in children learning to read in Wales. *British Journal of Psychology, 94*(1), 1–28.

الملحق ز: مقارنة بين برامج جمع البيانات

الملامح	أو دي كي	درويد سيرفاي	دو فورمز	سيرفاي تو جو	ماجبي	تانجارين	كوماند موبايل
السعر	مجاني	$60 لشهر واحد، $280 لستة أشهر، $400لسنة واحدة. عدد غير محدود من الأجهزة وإمكانية رفع 3000 نتيجة كل شهر	9، $95لشهر واحد، $99،95لسنة واحدة وجهاز واحد. ولنسخة المحترفين $14،65 لشهر واحد/$149،95 لسنة واحدة وجهاز واحد لنسخة الإرسال	0، $10إلى 0،15 لكل نموذج مكتمل، حسب وحدة التخزين.قد يتم تطبيق رسوم إضافية لإرسال/تخزين صور ومقاطع فيديو ومقاطع صوتية عالية الدقة	ترخيص مجاني يصل ل6000 نموذج/ مكتمل في العام وبالنسبة لوحدة التخزين الأعلى السعر يختلف		النسخة القياسية $24.999لمدة ثلاثة شهور، $239.99 للسنة الواحدة.النسخة المتقدمة: $69.99 لشهر واحد
التوافق							
أندرويد	نعم	نعم	نعم	نعم	نعم	نعم	نعم
آي أو اس IOS	نعم،دعم طرف ثالث	نعم،باستخدام تطبيقا من فصلا يسمى isurvey	نعم	لا	yes	لا	نعم
هاتف ويندوز	لا	لا	متوفر قريبا	نعم	لا	لا	نعم
سيمبيان	لا	لا	لا	لا	نعم	لا	لا
بلاكبيري	لا	لا	لا	لا	لا	لا	لا
رسائل قصيرة	لا	لا	لا	لا	نعم	لا	لا
تطوير الأداة							
يعتمد على النماذج (لا يحتاج لخبرة مبرمج)	نعم	نعم	نعم	نعم	نعم	نعم	نعم
إجراء تعديل على الأداة دون اتصال بالإنترنت	نعم	لا	نعم	لا	نعم، من الممكن تعديل الاداة دون الاتصال بالانترنت في برنامج إكسيل أو صيغة اكس فورمز ثم رفعها بعد ذلك	نعم	لا
كود موحد متوافق، توافق لغوي ونصي واسع	نعم	نعم	نعم	نعم	نعم	نعم	نعم
واجهة مستخدم لتوطين اللغة	نعم	نعم، يمكن الاختيار من بين 15 لغة	الإنجليزية والأسبانية والروسية	نعم	الاختيار من بين خمسة لغات مختلفة	نعم	غير مذكور

نماذج أساسية للمهام	لا	لا	لا	لا	لا	نعم	لا	
نماذج أساسية للمهام	لا	لا	لا	لا	لا	نعم	لا	
هل من الممكن عمل أداة تقييم مهارات القراءة للمراحل المبكرة	نعم تم هذا من قبل	نعم (لكن هذا لن يكون سهلا لأن هذا الجدول الشبكي يستهدف عدة صفوف من الأسئلة بنفس اختيارات الإجابات المتعددة) أي أنه ليس من الممكن تصنيف البنود في الشبكة	نعم، مع أي من النسخ المدفوعة	نعم، ولكن ليس بدون بعض التدريب	لا، ليس بدون التعاقد على خدمات التخصيص	نعم	النسخة التجريبية غير متاحة على الانترنت	
نماذج مطبوعة	نعم	لا	لا	نعم	لا	نعم	لا	

الملامح	أو دي كي	درويد سيرفاي	دو فورمز	سيرفاي تو جو	ماجبي	تانجارين	كوماند موبايل
جمع البيانات							
جمع البيانات دون الاتصال بالانترنت	نعم	نعم	نعم	نعم	نعم	نعم	نعم
بيانات نصية أو رقمية	نعم	نعم	نعم	نعم	نعم	نعم	نعم
بيانات المسوحات الموقوتة	نعم	لا	نعم	نعم	لا	نعم	لا
الجداول الشبكية	نعم	نعم	نعم	نعم	لا	نعم	نعم
الإجابات المنفردة أو إجابات الاختيار من متعدد	نعم	نعم	نعم	نعم	نعم	نعم	نعم
موقع الجي بي اس	نعم	نعم	نعم	نعم	نعم	نعم	نعم
لقطة شاشة	نعم	نعم	نعم	نعم	نعم	نعم	نعم
كاميرا	نعم	نعم	نعم	نعم	لا	لا	نعم
فيديو	نعم	لا	نعم	نعم	لا	لا	نعم
صوت	نعم	لا	نعم	نعم	لا	لا	نعم
باركود	نعم	نعم	نعم	نعم	لا	لا	نعم
توقيع	نعم	نعم	نعم	نعم	لا	لا	نعم
ملامح المنطق							
تخطي المنطق	نعم	نعم	نعم	نعم	نعم	نعم	نعم
التأكد من صحة	نعم	لا	نعم	نعم	نعم	نعم	نعم
العرض بشكل شرطي	نعم	لا	نعم	نعم	نعم	نعم	نعم
تكرار حلقي	نعم	لا	نعم	نعم	نعم	لا	لا
تفرع الأسئلة	نعم	لا	نعم	نعم	نعم	نعم	لا
رفع البيانات							
واي فاي	نعم	نعم	نعم	نعم	نعم	نعم	نعم
خلوي	نعم	نعم	نعم	نعم	نعم	نعم	نعم
كابل	نعم	نعم	نعم	نعم	نعم	لا	نعم
النسخ الاحتياطي من	نعم، مع دعم	لا	لا	لا	نعم، عن طريق كارت	نعم	لا
تصدير البيانات إلى							
إكسيل	نعم	نعم	نعم	نعم	نعم	نعم	نعم
RSS	لا	لا	لا	لا	لا	لا	نعم
SPSS	لا	نعم	لا	نعم	لا	نعم	لا
وورد	لا	لا	نعم	نعم	نعم	لا	لا

أكسس	لا	لا	نعم	نعم	نعم	لا	لا
XML	نعم	لا	نعم	نعم	نعم	لا	لا
HTML	لا	لا	نعم	لا	لا	لا	لا
PDF	لا	لا	نعم	لا	نعم	لا	لا
وثائق جوجل	نعم	لا	نعم	لا	لا	لا	لا
اوبن اوفيس	لا	لا	نعم	لا	لا	لا	لا
تخزين البيانات							
حجم التخزين المتضمن	تخزين سحابي غير محدود	تخزين سحابي غير محدود	تخزين سحابي غير محدود	20 ميجابايت للمرفقات ويمكن شراء المزيد.تخزين سحابي غير محدود	تخزين سحابي غير محدود		250 جيجا بايت
التشفير أثناء النقل؟	نعم	نعم، تشفير SSL	نعم، تشفير SSL	نعم بتكلفة إضافية	مشفرة لمعيار 256 بت	لا	نعم
مصدر مفتوح	نعم	نتائج API متوفرة عند	APIمتوفر ولكن ليس	لا	لا، ولكن APIمتوفر	نعم	لا
وضع Kiosk؟	نعم ، مع دعم	نعم	لا	نعم	نعم	لا	لا

ملحق ح: مقارنة تعليمات التقييم الورقي مقابل الإلكتروني لتقييم القراءة في المراحل المبكرة

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
▪ **تعليمات عامة**		
• أسس لحوار يتسم باللعب والمرونة مع الطفل من خلال المحادثات القصيرة (راجع أمثلة الموضوعات أدناه). يجب أن ينظر الطفل للتقييم بكونه لعبة تقريباً يمكنه الاستمتاع به بدلاً من مجرد كونه اختباراً. استغل هذا الوقت في تحديد اللغة التي يعد الطفل أكثر راحة في التعامل بها. اقرأ بصوت عال وببطء ووضوح **فقط** الأقسام المبينة في الأطر.	أسس لحوار يتسم باللعب والمرونة مع الطفل من خلال المحادثات القصيرة (راجع أمثلة الموضوعات أدناه). يجب أن ينظر الطفل للتقييم بكونه لعبة تقريباً يمكنه الاستمتاع به بدلاً من مجرد كونه اختباراً. استغل هذا الوقت في تحديد اللغة التي يعد الطفل أكثر راحة في التعامل بها. اقرأ بصوت عال وببطء ووضوح **فقط** الأقسام المبينة في الأطر.	صباح الخير . أسمى ..... و أقيم في ..... أحب أن اخبرك قليلًا عن نفسي . (عدد الأطفال و أعمارهم ، الرياضة المفضلة إليك ، البرنامج التليفزيوني أو الإذاعي ، ...) 1. **ماذا تحب ان تفعل عندما لاتكون في المدرسة؟** (انتظر قليلًا للإجابةإذا كان التلميذ مترددًا ، اطرح السؤال رقم 2، وإذا شعرت بأن الطفل مرتاح، انتقل إلى الموافقة الشفوية) **ماذا تحب أن تلعب؟**
التقييم الشفهى		
قل للطالب ما يلى : • **دعني أخبرك لماذا أنا اليوم هنا.** • **نحاول في وزارة التربية و التعليم أن نفهم كيف يتعلم الأطفال القراءة . حدث اختيارك بمحض الصدفة.** • **نحتاج الى مساعدتك في إنجاز هذا العمل ، فإذا لم ترغب في ذلك فلك مطلق الحرية في عدم المشاركة.** • **سوف نلعب معاً لعبة القراءة. سوف أطلب منك أن تقرأ بصوت عال أحرفاً وكلمات وقصة قصيرة.** • **باستخدام ساعة التوقيت هذه ، سوف أعرف المدة التي استغرقتها في القراءة.** • **هذا ليس اختباراً ولن يؤثر على درجاتك المدرسية.**	إن لم يتم الحصول على موافقة شفهية، اشكر الطفل، واضغط "لا"، ثم اضغط "حفظ" و"ابدأ تقييم آخر."	إن لم يتم الحصول على موافقة شفهية، اشكر الطفل وانتقل للطفل التالي، باستخدام نفس النموذج.

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
• **لن أكتب اسمك وبالتالي لا أحد يمكنه أن يعرف أن هذه هي استجاباتك.** • **مرة أخرى ، إذا لم ترغب في المشاركة أو لم ترغب في الإجابة عن سؤال ما فلك ذلك .** • **هل لديك أي أسئلة تريد أن تطرحها على.؟** • **هل يمكننا أن نبدأ ؟**		
التوجيه المطبوع		
اعرض على الطفل جزء الفقرة في الصفحة الأخيرة في تقييم الطالب. اقرأ التعليمات للطفل. سجل إجابة الطفل قبل الانتقال إلى التوجيه التالي.	اعرض على الطفل جزء الفقرة في الصفحة الأخيرة في تقييم الطالب. اقرأ التعليمات للطفل. سجل إجابة الطفل قبل الانتقال إلى التوجيه التالي.	لا أريدك أن تقرأ هذا الأن. ***فقط أشر لي من أين تبدأ القراءة . ثم حرك إصبعك في الاتجاه الصحيح للقراءة .ثم ضع إصبعك على آخر كلمة.***
التعرف على أصوات الحروف		
اعرض للطفل ورقة الحروف في كتيب تحفيز الطالب أثناء قراءة التعليمات التالية للطفل. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الحرف الأول. • تابع باستخدام قلمك الرصاص وضع علامة بوضوح على أي حروف غير صحيحة بعلامة (/).	اعرض للطفل ورقة الحروف في كتيب تحفيز الطالب أثناء قراءة التعليمات التالية للطفل. ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الحرف الأول. تابع على شاشتك وضع علامة على أي حروف غير صحيحة من خلال لمس الحرف غير الصحيح حيث سيتحول إلى اللون الأزرق. ضع علامة على التصحيحات الذاتية بكونها صحيحة من خلال لمس الحرف مرة أخرى – سيتحول إلى اللون الرمادي.	**هذه الصفحة مليئة بالأحرف. رجاء انطق صـــوت كل حرف مع الحركة. فعلى سبيل المثال اسم الحرف مع الحركة "سَ"** (أشر الى الحرف " سَ ") **هو " سَ " ، الآن حاول انت . أنطق لي أنت صوت هذا الحرف** (أطلب من التلميذ/ التلميذة أن يقراء حرف "كُ"). فإذا استطاع التلميذ/التلميذة أن يقراء الحرف مع الحركة بشكل صحيح فقل له **"أحسنت"، صوت الحرف " كُ "** . أما إذا أخطأ التلميذ/التلميذة فقل له **صوت هذا الحرف " كُ ". الآن حاول مرة أخرى: انطق لي صوت هذا الحرف** (اطلب من التلميذ/التلميذة تعُرف صوت "لُ"). فإذا استطاع التلميذ/التلميذة أن يقراء الحرف مع الحركة بشكل صحيح ، فقل له **" أحسنت"، صوت هذا الحرف " لُ"** . أما إذا أخطأ التلميذ/التلميذة فقل له **القراء الصحيحة لهذا الحرف "لُ". هل تفهم ما هو مطلوب منك ؟** **عندما أقول لك " ابدأ " أقراء الحروف بأفضل طريقة ممكنه، سوف ألتزم الصمت وأستمع لما تقول مستعد؟**

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
• احسب التصحيحات الذاتية بكونها صحيحة. إذا كنت قد علمت بالفعل على الكلمة التي تم تصحيحها ذاتياً بأنها غير صحيحة، ضع دائرة حول العلامة (ø) واستمر. • إذا ما أسقط التلميذ سطراً كاملاً، ارسم خط على هذا السطر في البروتوكول. • التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للحرف التالي وقل "من فضلك استمر." ضع علامة غير صحيح على الحرف الذي تم إسقاطه. • إذا ما قدم الطفل اسم الحرف بدلاً من صوته، انطق صوت الحرف، وقل: "من فضلك اخبرني **بصوت** الحرف." يمكن استخدام هذا التنبيه مرة واحدة فقط أثناء المهام الفرعية. عند وصول ساعة ضبط الوقت إلى صفر، قل "توقف". ضع علامة على الكلمة الأخيرة التي تمت قراءتها بعلامة ( ]). قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في السطر الأول (10 بنود)، قل "شكرا!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية. **[في نهاية الصفحة، أدخل الأسطر التالية]** الوقت المتبقي على ساعة ضبط الوقت عند الانتهاء ☐ (عدد الثواني) ☐ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول	التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للحرف التالي وقل "من فضلك استمر." ضع علامة غير صحيح على الحرف الذي تم إسقاطه. إذا ما قدم الطفل اسم الحرف بدلاً من صوته، انطق صوت الحرف، وقل: "من فضلك اخبرني **بصوت** الحرف." يمكن استخدام هذا التنبيه مرة واحدة فقط أثناء المهام الفرعية. إذا ما انتهى الوقت قبل قراءة البند الأخير، ستصدر الشاشة ضوء أحمر وسيتوقف العد الزمني. اطلب من الطفل التوقف. ضع علامة على آخر حرف تمت قراءته من خلال لمسه بحيث يظهر قوساً أحمر. بعدها اضغط "التالي". إذا ما وصل الطفل إلى البند الأخير قبل إضاءة الشاشة باللون الأحمر، أوقف ساعة الضبط بمجرد قراءة الطفل للحرف الأخير. المس الحرف الأخير بحيث يظهر القوس الأحمر. بعدها اضغط "التالي." قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في السطر الأول (10 بنود)، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكرا!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
التعرف على أصوات الحروف		
اعرض للطفل ورقة الحروف في كتيب تحفيز الطالب أثناء قراءة التعليمات التالية للطفل. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الحرف الأول. • تابع باستخدام قلمك الرصاص وضع علامة بوضوح على أي حروف غير صحيحة بعلامة (/). • احسب التصحيحات الذاتية بكونها صحيحة. إذا كنت قد علمت بالفعل على الكلمة التي تم تصحيحها ذاتياً بأنها غير صحيحة، ضع دائرة حول العلامة (ø) واستمر. • إذا ما أسقط التلميذ سطراً كاملاً، ارسم خط على هذا السطر في البروتوكول. • التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للحرف التالي وقل "من فضلك استمر." ضع علامة غير صحيح على الحرف الذي تم إسقاطه. • إذا ما قدم الطفل اسم الحرف بدلاً من صوته، انطق صوت الحرف، وقل: "من فضلك اخبرني **<u>بصوت</u>** الحرف." يمكن استخدام هذا التنبيه مرة واحدة فقط أثناء المهام الفرعية. عند وصول ساعة ضبط الوقت إلى صفر، قل "توقف". ضع علامة على الكلمة الأخيرة التي تمت قراءتها بعلامة ( ]). قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في السطر الأول (10 بنود)، قل "شكرا!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية.	اعرض للطفل ورقة الحروف في كتيب تحفيز الطالب أثناء قراءة التعليمات التالية للطفل. ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الحرف الأول. تابع على شاشتك وضع علامة على أي حروف غير صحيحة من خلال لمس الحرف غير الصحيح حيث سيتحول إلى اللون الأزرق. ضع علامة على التصحيحات الذاتية بكونها صحيحة من خلال لمس الحرف مرة أخرى – سيتحول إلى اللون الرمادي. التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للحرف التالي وقل "من فضلك استمر." ضع علامة غير صحيح على الحرف الذي تم إسقاطه. إذا ما قدم الطفل صوت الحرف بدلاً من اسمه، قل: "من فضلك أخبرني **<u>باسم</u>** الحرف." يمكن استخدام هذا التنبيه مرة واحدة فقط أثناء المهام الفرعية. إذا ما انتهى الوقت قبل قراءة البند الأخير، ستصدر الشاشة ضوء أحمر وسيتوقف العد الزمني. اطلب من الطفل التوقف. ضع علامة على آخر حرف تمت قراءته من خلال لمسه بحيث يظهر قوساً أحمر. بعدها اضغط "التالي". إذا ما وصل الطفل إلى البند الأخير قبل إضاءة الشاشة باللون الأحمر، أوقف ساعة الضبط بمجرد قراءة الطفل للحرف الأخير. المس الحرف الأخير بحيث يظهر القوس. بعدها اضغط "التالي."	هذه الصفحة مليئة بالحروف. من فضلك، سم لي ماتستطيع تسميته من حروف. لا تنطق صوت الحرف بل اسم الحرف. فعلى سبيل المثال، اسم هذا الحرف (اشر الي الحرف " م ") اسم هذا الحرف : ميم" الأن حاول انت. (اطلب من التلميذ أن يتعرف اسم الحرف "ب") . إذا استطاع أن يسمي الحرف بشكل صحيح فقل له"أحسنت".اسم هذا الحرف "باء". أما إذا اخطاء التلميذ فقل له اسم هذا الحرف هو "باء.الأن حاول مرة اخرى : سم لي هذا الحرف(اطلب من التلميذ أن يتعرف اسم الحرف "ل"). فإذا استطاع أن يسمي الحرف بشكل صحيح فقل له"أحسنت".اسم هذا الحرف "لام" . أما إذا اخطاء التلميذ فقل له اسم هذا الحرف هو "لام". هل تفهم ماهو مطلوب منك ؟ عندما أقول لك "ابدأ" سم الأحرف، سألتزم الصمت و استمع لما تقولإلا إذا احتجت مساعدتي، مستعد؟ ابدأ. بعد الانتهام من المهمة ، قل للطالب" أحسنت. جهد ممتاز! لننتقل إلى الجزء التالي.

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
[في نهاية الصفحة، أدخل الأسطر التالية] الوقت المتبقي على ساعة ضبط الوقت عند الانتهاء ☐ (عدد الثواني) ☐ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول	قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة لأول 10 بنود، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكرا!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	
الوعي الصوتي – **المهمة الأولى: التعرف الميدئى على الأصوات**		
هذا التدريب غير محدد الوقت وليس هناك ورقة للطالب. اخفي كتيب تحفيز التلميذ عن مرأى الطفل. اقرأ التعليمات للطفل ونفذ الأمثلة. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** اقرأ بداية الكلمة وبعدها انطق الكلمة مرة ثانية. انطق كل كلمة ببطء. اعتبر الإجابة الصحيحة أي صوت منفرد فقط. ضع علامة على الإطار الذي يتماشى مع إجابة الطفل. إن لم يستجب الطفل بعد 3 ثواني، ضع علامة "لا إجابة" واقرأ بداية الكلمة التالية. قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في البنود الخمس الأولى، قل "شكرا!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية.	هذا التدريب غير محدد الوقت وليس هناك ورقة للطالب. اقرأ التعليمات للطفل ونفذ الأمثلة. اقرأ بداية الكلمة وبعدها انطق الكلمة مرة ثانية. تابع على شاشتك وضع علامة على كل بند "صحيح" أو "خطأ". سيتحول اختيارك إلى اللون الأصفر. إن لم يستجب الطفل بعد 3 ثواني، ضع علامة "لا إجابة" واقرأ بداية الكلمة التالية. قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة لأول 5 بنود، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكراً!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	هذا تمرين للاستماع. أريدك أن تخبرني ما الصوت الأول في كل كلمة. على سبيل المثال، كلمة : مدرسة" الصوت الأول بها /مـ/ . ريدك أن تخبرني ما الصوت الأول الذي تسمعه في كل كلمة. سأقرأ كل كلمة مرتين، فتخبرني الصوت الأول في هذه الكلمة. الأن حاول انت. ما الصوت الأول في " قَلَم". . إذا استجاب التلميذ بشكل صحيح فقل له"أحسنت". الصوت الأول في " قَلَم". /قـ /. أما إذا اخطاء التلميذ فقل له الصوت الأول في " قَلَم". /قـ /. الأن حاول مرة اخرى : ما الصوت الأول في " مُتْحَفُ". . إذا استجاب التلميذ بشكل صحيح فقل له"أحسنت". الصوت الأول في " مُتْحَفُ".. /مُ /. أما إذا اخطاء التلميذ فقل له الصوت الأول في " مُتْحَفُ" /مُ /. هل تفهم ماهو مطلوب منك ؟ ابدأ. بعد الانتهام من المهمة ، قل للطالب" أحسنت. جهد ممتاز! لننتقل إلى الجزء التالي

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
[في نهاية الصفحة، أدخل الأسطر التالية] ☐ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول		
الوعي الصوتي - المرحلة الثانية: تقسيم المقاطع الصوتية		
هذا التدريب غير محدد الوقت وليس هناك ورقة للطالب. اخفي كتيب تحفيز التلميذ عن مرأى الطفل. اقرأ التعليمات للطفل ونفذ الأمثلة. اقرأ بصوت عال كل كلمة مرتين ودع الطفل ينطق الأصوات. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** انطق كل كلمة ببطء. لا تقطع الكلمة إلى أصوات فردية. قل كلمة كلمة **مرتين** فقط. إذا ما ذكر الطفل اسم الحرف بدلاً من صوته، قل: "من فضلك أخبرني بالأصوات في الكلمات وليس الحروف." يمكن تقديم هذه النصيحة مرة واحدة أثناء التدريب.	خفي كتيب تحفيز التلميذ عن مرأى الطفل. اقرأ التعليمات للطفل ونفذ الأمثلة. اقرأ بصوت عال كل كلمة مرتين ودع الطفل ينطق الأصوات. انطق كل كلمة ببطء. لا تقطع الكلمة إلى أصوات فردية. قل كلمة كلمة **مرتين** فقط. المس الشاشة فقط عند كل صوت **غير صحيح** – ستتحول إلى اللون الأصفر. إذا ما قدم الطفل اسم الحرف بدلاً من صوته، قل: "من فضلك اخبرني بأصوات الكلمة وليس حروفها" يمكن استخدام هذا التنبيه مرة واحدة فقط أثناء المهام الفرعية.	هذا تمرين للاستماع. أريدك أن تخبرني أنت تعرف أن الكلمة تتكون من مقاطع صوتية . على سبيل المثال، كلمة : زرَعَ" تتكون من المقاطع الصوتية /زَ/ ، /رَ/ ، /عَ/. .أريدك أن تخبرني المقاطع الصوتية التي تتكون منها الكلمة.. الأن حاول انت. ما المقاطع الصوتية في كلمة " سَريعُ"؟ . إذا استجاب التلميذ بشكل صحيح فقل له"أحسنت". المقاطع الصوتية في " سَريعُ" هي /سَ/ ، /ريـ/ ، /عُ/ . أما إذا اخطاء التلميذ فقل له المقاطع الصوتية في " سَريعُ" هي /سَ/ ، /ريـ/ ، /عُ/. الأن حاول مرة اخرى : ما المقاطع الصوتية في كلمة "طَبيبُ"؟. إذا استجاب التلميذ بشكل صحيح فقل له"أحسنت". الأصوات في "طَبيبُ" هي /طَ/، /بِيـَ/، /بُ/. أما إذا اخطاء التلميذ فقل الأصوات في طَبيبُ" هي /طَ/، /بِيـَ/، /بُ/. هل تفهم ماهو مطلوب منك ؟ ابدأ.

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
ضع علامة (/) على كل مقطع غير صحيح، وكذلك أي مقطع لا ينطقه الطفل. إن لم يجب الطفل بعد 3 ثواني، ضع علامة غير صحيح على جميع الأصوات وانتقل إلى الكلمة التالية. إذا قدم الطفل جميع الأصوات الصحيحة، ضع علامة على إطار "الإجابات كلها صحيحة." قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في البنود الخمس الأولى، قل "شكرا!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية. **[في نهاية الصفحة، أدخل الأسطر التالية]** □ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول	إن لم يجب الطفل بعد 3 ثواني، ضع علامة غير صحيح على جميع الأصوات وانتقل إلى الكلمة التالية. إذا قدم الطفل جميع الأصوات الصحيحة، ضع علامة على إطار "الإجابات كلها صحيحة." قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة لأول 5 بنود، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكراً!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	
التعرف على المقاطع اللفظية		
اعرض على الطفل ورقة المقاطع اللفظية في كتيب تحفيز الطالب. اقرأ التعليمات للطفل ونفذ الأمثلة. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** ابدأ ساعة تحديد الوقت عندما يقرأ الطفل المقطع الأول.	اقرأ التعليمات. تأكد من أن تظهر للطفل صفحة المقاطع اللفظية، مع توضيح أول مقطع لفظي عندما تقول "ابدأ." ابدأ ساعة ضبط الوقت عندما يقرأ الطفل المقطع الأول. تابع على شاشتك وضع علامة على أي مقطع غير صحيح من خلال لمس المقطع حيث سيتحول إلى اللون الأزرق.	

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
• تابع باستخدام قلمك الرصاص وضع علامة بوضوح على أي مقاطع غير صحيحة بعلامة (/). • احسب التصحيحات الذاتية بكونها صحيحة. إذا كنت قد علمت بالفعل على المقطع الذي تم تصحيحه ذاتياً بأنه غير صحيح، ضع دائرة حول العلامة (ø) واستمر. • إذا ما أسقط التلميذ سطراً كاملاً، ارسم خط على هذا السطر في البروتوكول. • التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للحرف التالي وقل "من فضلك استمر." ضع علامة غير صحيح على الحرف الذي تم إسقاطه. • إذا ما قدم الطفل اسم الحرف بدلاً من صوته، انطق صوت الحرف، وقل: "من فضلك استمر." علم على المقطع الذي تم إسقاطه بأنه غير صحيح. عند وصول ساعة ضبط الوقت إلى صفر، قل "توقف". ضع علامة على الكلمة الأخيرة التي تمت قراءتها بعلامة ( ]). قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في السطر الأول (10 بنود)، قل "شكراً!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية. **[في نهاية الصفحة، أدخل الأسطر التالية]** الوقت المتبقي على ساعة ضبط الوقت عند الانتهاء ☐ (عدد الثواني) ☐ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول	ضع علامة على التصحيحات الذاتية بكونها صحيحة من خلال لمس المقطع مرة أخرى – سيتحول إلى اللون الرمادي. التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للمقطع التالي وقل "من فضلك استمر." ضع علامة غير صحيح على المقطع الذي تم إسقاطه. إذا ما انتهى الوقت قبل قراءة البند الأخير، ستصدر الشاشة ضوء أحمر وسيتوقف العد الزمني. اطلب من الطفل التوقف. ضع علامة على آخر مقطع تمت قراءته من خلال لمسه بحيث يظهر قوساً أحمر. بعدها اضغط "التالي". إذا ما وصل الطفل إلى البند الأخير قبل إضاءة الشاشة باللون الأحمر، أوقف ساعة الضبط بمجرد قراءة الطفل للمقطع الأخير. المس المقطع الأخير بحيث يظهر القوس الأحمر. بعدها اضغط "التالي." قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة لأول 10 بنود، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكراً!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
تقسيم المقاطع اللفظية		
هذا التدريب غير محدد الوقت وليس هناك ورقة للطالب. اخفي كتيب تحفيز التلميذ عن مرأى الطفل. اقرأ التعليمات للطفل ونفذ الأمثلة. اقرأ بصوت عال كل كلمة مرتين واطلب من الطفل قول المقاطع. تذكر عند الممارسة المقاطع "المستقطعة" مثل "با". **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** انطق كل كلمة ببطء. لا تقسم الكلمة إلى مقاطع فردية. قل كلمة كلمة **مرتين** فقط. إذا ما ذكر الطفل الكلمة بدلاً من المقطع، قل: "من فضلك أخبرني بجميع المقاطع في الكلمة." يمكن تقديم هذه النصيحة مرة واحدة أثناء التدريب. ضع علامة (/) على كل مقطع غير صحيح، وكذلك أي مقطع لا ينطقه الطفل. إن لم يجب الطفل بعد 3 ثواني، ضع علامة غير صحيح على جميع الأصوات وانتقل إلى الكلمة التالية. إذا قدم الطفل جميع الأصوات الصحيحة، ضع علامة على إطار "الإجابات كلها صحيحة." قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في السطر الأول (5 بنود)،	هذا التدريب غير محدد الوقت وليس هناك ورقة للطالب. اخفي كتيب تحفيز التلميذ عن مرأى الطفل. اقرأ التعليمات للطفل ونفذ الأمثلة. اقرأ بصوت عال كل كلمة مرتين واطلب من الطفل قول المقاطع. تذكر عند الممارسة المقاطع "المستقطعة" مثل "با". انطق كل كلمة ببطء. لا تقسم الكلمة إلى مقاطع فردية. قل كلمة كلمة **مرتين** فقط. المس الشاشة فقط لكل مقطع **غير صحيح** – سيتحول إلى اللون الأصفر. إذا ما ذكر الطفل الكلمة بدلاً من المقطع، قل: "من فضلك أخبرني بجميع المقاطع في الكلمة." يمكن تقديم هذه النصيحة مرة واحدة أثناء التدريب. إن لم يجب الطفل بعد 3 ثواني، ضع علامة غير صحيح على جميع الأصوات وانتقل إلى الكلمة التالية. إذا قدم الطفل جميع الأصوات الصحيحة، ضع علامة على إطار "الإجابات كلها صحيحة." قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة لأول 5 بنود، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكراً!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
ل "شكرا!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية. **[في نهاية الصفحة، أدخل الأسطر التالية]** ☐ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول		
قراءة الكلمات المعروفة		
اعرض على الطفل ورقة الكلمة في كتيب تحفيز الطالب. اقرأ التعليمات للطفل ونفذ الأمثلة. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الكلمة الأولى. • تابع باستخدام قلمك الرصاص وضع علامة بوضوح على أي كلمة غير صحيحة بعلامة (/). • احسب التصحيحات الذاتية بكونها صحيحة. إذا كنت قد علمت بالفعل على الكلمة التي تم تصحيحها ذاتياً بأنها غير صحيحة، ضع دائرة حول العلامة (Ø) واستمر. • إذا ما أسقط التلميذ سطراً كاملاً، ارسم خط على هذا السطر في البروتوكول. • التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للكلمة التالية وقل "من فضلك استمر." ضع علامة غير صحيح على الكلمة التي تم إسقاطها. عند وصول ساعة ضبط الوقت إلى صفر، قل "توقف". ضع علامة على الكلمة الأخيرة التي تمت قراءتها بعلامة ( ]).	اعرض على الطفل ورقة الكلمة في كتيب تحفيز الطالب. اقرأ التعليمات للطفل. ابدأ ساعة ضبط الوقت عندما يقرأ الطفل الكلمة الأولى. تابع على شاشتك وضع علامة على أي كلمة غير صحيحة من خلال لمس الكلمة غير الصحيحة حيث ستتحول إلى اللون الأزرق. ضع علامة على التصحيحات الذاتية بكونها صحيحة من خلال لمس الكلمة مرة أخرى – ستتحول إلى اللون الرمادي. التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للكلمة التالية وقل "من فضلك استمر." ضع علامة غير صحيح على الكلمة التي تم إسقاطها. إذا ما انتهى الوقت قبل قراءة البند الأخير، ستصدر الشاشة ضوء أحمر وسيتوقف العد الزمني. اطلب من الطفل التوقف. ضع علامة على آخر حرف تمت قراءته من خلال لمسه بحيث يظهر قوساً أحمر. بعدها اضغط "التالي". إذا ما وصل الطفل إلى البند الأخير قبل إضاءة الشاشة باللون الأحمر، أوقف ساعة الضبط بمجرد قراءة الطفل للكلمة الأخيرة. المس الكلمة الأخيرة بحيث يظهر القوس. بعدها اضغط "التالي."	**انظر إلى هذه الكلمات . لأريدك أن تقرأ أكبر عدد من هذه الكلمات . لا تتهجأ الكلمات، بل اقراءها. على سبيل المثال ، هذه الكلمة " رَسَمَ" .** **الآن حاول أنت : اقرأ لي هذه الكلمة " مَدْرَسَة "** إذا استجاب التلميذ/التلميذة بشكل صحيح، **فقل أحسنت ، هذه الكلمة هي " مَدْرَسَة " .** أما إذا أخطأ التلميذ/التلميذة ، **فقل هذه الكلمة هي " مَدْرَسَة " .** **الآن حاول مرة اخرى : اقرأ لي هذه الكلمة " مُتْحَفُ "** إذا استجاب التلميذ/التلميذة بشكل صحيح، **فقل أحسنت ، هذه الكلمة هي " مُتْحَفُ "** أما إذا أخطأ التلميذ/التلميذة ، **فقل هذه الكلمة هي " مُتْحَفُ "** **هل فهمت المطلوب ؟ هل أنت مستعد ؟ ابدأ .** بعد الانتهام من المهمة ، قل للطالب" أحسنت. جهد ممتاز! لننتقل إلى الجزء التالي.

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في السطر الأول (5 بنود)، قل "شكرا!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية. **[في نهاية الصفحة، أدخل الأسطر التالية]** الوقت المتبقي على ساعة ضبط الوقت عند الانتهاء ☐ (عدد الثواني) ☐ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول	قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة لأول 5 كلمات، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكراً!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	
قراءة الكلمات بلا معنى		**Nonword Reading**
اعرض للطفل استمارة الكلمات في كتيب تحفيز الطالب. اقرا التعليمات للطفل ونفذ الأمثلة. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الكلمة بلا معنى الأولى. • تابع باستخدام قلمك الرصاص وضع علامة بوضوح على أي كلمة غير صحيحة بعلامة (/). • احسب التصحيحات الذاتية بكونها صحيحة. إذا كنت قد علمت بالفعل على الكلمة التي تم تصحيحها ذاتياً بأنها غير صحيحة، ضع دائرة حول العلامة (Ø) واستمر. • إذا ما أسقط التلميذ سطراً كاملاً، ارسم خط على هذا السطر في البروتوكول. التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للكلمة التالية وقل "من فضلك استمر." ضع علامة غير صحيح على الكلمة التي تم إسقاطها.	اعرض للطفل استمارة الكلمات بلا معنى في كتيب تحفيز الطالب. اقرا التعليمات للطفل. ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الكلمة بلا معنى الأولى. تابع على شاشتك وضع علامة على أي كلمة بلا معنى غير صحيحة من خلال لمس الكلمة حيث ستتحول إلى اللون الأزرق. ضع علامة على التصحيحات الذاتية بكونها صحيحة من خلال لمس الكلمة بلا معنى مرة أخرى – ستتحول إلى اللون الرمادي. التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للكلمة التالية وقل "من فضلك استمر." ضع علامة غير صحيح على الكلمة التي تم إسقاطها.	**انظر إلى هذه الكلمة ، إنها كلمة ليس لها معنى معروف لديك : هذا مثال على ذلك : " ماصِحُ "** **الآن حاول أنت : اقرأ لي هذه الكلمة " بيلاذُ "** إذا استجاب التلميذ/التلميذة بشكل صحيح، **فقل أحسنت ، هذه الكلمة هي " بيلاذُ " .** أما إذا أخطأ التلميذ/التلميذة ، **فقل هذه الكلمة هي " بيلاذُ " .** **الآن حاول مرة اخرى : اقرأ لي هذه الكلمة " ناسِبَ "** إذا استجاب التلميذ/التلميذة بشكل صحيح، **فقل أحسنت ، هذه الكلمة هي " ناسِبَ".** أما إذا أخطأ التلميذ/التلميذة ، **فقل هذه الكلمة هي " ناسِبَ "** **هل فهمت المطلوب ؟ هل أنت مستعد ؟ ابدأ .** بعد الانتهاء من المهمة ، قل للطالب" أحسنت. جهد ممتاز! لننتقل إلى الجزء التالي.

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
عند وصول ساعة ضبط الوقت إلى صفر، قل "توقف". ضع علامة على الكلمة الأخيرة التي تمت قراءتها بعلامة ( ]). قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة في السطر الأول (5 بنود)، قل "شكرا!"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل، وانتقل إلى المهمة الفرعية التالية. **[في نهاية الصفحة، أدخل الأسطر التالية]** الوقت المتبقي على ساعة ضبط الوقت عند الانتهاء ☐ (عدد الثواني) ☐ تم وقف التدريب لأن الطفل لم يقدم أي إجابات صحيحة في السطر الأول	إذا ما انتهى الوقت قبل قراءة البند الأخير، ستصدر الشاشة ضوء أحمر وسيتوقف العد الزمني. اطلب من الطفل التوقف. ضع علامة على آخر كلمة بلا معنى تمت قراءته من خلال لمسه بحيث يظهر قوساً أحمر. بعدها اضغط "التالي". إذا ما وصل الطفل إلى البند الأخير قبل إضاءة الشاشة باللون الأحمر، أوقف ساعة الضبط بمجرد قراءة الطفل للكلمة بلا معنى الأخيرة. المس الكلمة بلا معنى الأخيرة بحيث يظهر القوس. بعدها اضغط "التالي." قاعدة التوقف المبكر: إذا لم يقدم الطفل أي إجابة صحيحة لأول 5 كلمات، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "شكراً!"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	
إجادة القراءة الشفهية/قراءة الفقرات		
اعرض على الطفل القصة في كتيب تحفيز الطالب. اقرأ التعليمات للطفل. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الكلمة الأولى. • تابع باستخدام قلمك الرصاص وضع علامة بوضوح على أي كلمة غير صحيحة بعلامة (/). • احسب التصحيحات الذاتية بكونها صحيحة. إذا كنت قد علمت بالفعل على الكلمة التي تم تصحيحها ذاتياً بأنها غير صحيحة، ضع دائرة حول العلامة (Ø) واستمر.	اعرض للطفل القصة في كتيب تحفيز الطالب. اقرا التعليمات للطفل ونفذ الأمثلة. ابدأ ساعة تحديد الوقت عندما يقرأ الطفل الكلمة الأولى. تابع على شاشتك وضع علامة على أي كلمة غير صحيحة من خلال لمس الكلمة غير الصحيحة حيث ستتحول إلى اللون الأزرق. ضع علامة على التصحيحات الذاتية بكونها صحيحة من خلال لمس الكلمة مرة أخرى – ستتحول إلى اللون الرمادي. التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للكلمة التالية وقل "من فضلك استمر." ضع علامة غير صحيح على الكلمة التي تم إسقاطها.	**أريد منك أن تقرأ هذه القطعة بصوت عالٍ ،و بسرعة، وعندما تنتهي من قراءتها سوف أسألك بعض الأسئلة عنها.** • **هل فهمت المطلوب ؟ عندما أقول "ابدأ" أقرأ القصة.**

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
• التزم الهدوء، إلا إذا تردد الطفل لعدد 3 ثواني. أشر للكلمة التالية وقل "من فضلك استمر." ضع علامة غير صحيح على الكلمة التي تم إسقاطها. • إذا ما أسقط التلميذ سطراً كاملاً، ارسم خط على هذا السطر في البروتوكول. عند وصول ساعة ضبط الوقت إلى صفر، قل "توقف". ضع علامة على الكلمة الأخيرة التي تمت قراءتها بعلامة ( ]). بعد انتهاء الطفل من القراءة، ابعد القطعة من أمام الطفل. قاعدة التوقف المبكر: إذا لم يقدم الطفل أي كلمة صحيحة في السطر الأول من النص، قل "**شكراً!**"، وأوقف هذه المهمة الفرعية، وضع علامة على الإطار في الأسفل. لا تطرح أي أسئلة لقياس الفهم.	إذا ما انتهى الوقت قبل قراءة البند الأخير، ستصدر الشاشة ضوء أحمر وسيتوقف العد الزمني. اطلب من الطفل التوقف. ضع علامة على آخر كلمة تمت قراءته من خلال لمسه بحيث يظهر قوساً أحمر. بعدها اضغط "التالي". إذا ما وصل الطفل إلى البند الأخير قبل إضاءة الشاشة باللون الأحمر، أوقف ساعة الضبط بمجرد قراءة الطفل للكلمة الأخيرة. المس الكلمة الأخيرة بحيث يظهر القوس. بعدها اضغط "التالي." ابعد النص من أمام الطفل. قاعدة التوقف المبكر: إذا لم يقرأ الطفل أي كلمة صحيحة من بين الكلمات اللازمة لإجابة أول سؤال لقياس الفهم، ستصدر الشاشة نور أحمر متقطع وسيتوقف العد الزمني. قل "**شكراً!**"، وأوقف هذه المهمة الفرعية، وانتقل إلى المهمة الفرعية التالية.	
استيعاب القراءة		
اسأل الطفل فقط الأسئلة المتعلقة بقراءة النص. يجب أن يقرأ الطفل **النص بالكامل** والذي يتماشى مع السؤال المطروح. ضع علامة على إجابة الطفل وانتقل إلى السؤال التالي. ضع علامة صح فقط على الإجابات المتشابهة أو المماثلة في المعنى للأجوبة المقدمة بجوار كل سؤال. إن لم يجب الطفل على السؤال بعد مرور 10 ثوان، انتقل إلى السؤال التالي. لا تكرر السؤال.	قبل طرح الأسئلة، نح القطعة جانباً من أمام الطفل. اطرح على الطفل جميع الأسئلة الظاهرة على الشاشة. وهي مصفوفة أتوماتيكياً مع ما قرأه الطفل في قطعة القراءة الشفهية. ضع علامة على كل سؤال سواء "صحيح" أو "خطأ". سيتحول اختيارك إلى اللون الأصفر. إن لم يجب الطفل بعد مرور 10 ثوان، ضع علامة "لا إجابة" وانتقل للسؤال التالي.	الأن سأطرح عليك بعض الأسئلة عن القصة التي قرأتها. اجب عن الاسئلة على قدر الإمكان.

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
اقرأ التعليمات للطفل. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]**	الأجوبة ذات المعنى المشابه للأجوبة المقدمة يجب اعتبارها صحيحة. إذا قال الطفل "لا أعلم"، ضع علامة "غير صحيح".	
الاستيعاب السمعي		
نح كتيب تحفيز الطالب عن مرأى الطفل. اقرأ التعليمات للطفل. **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** هذه المهمة الفرعية ليست مرتبطة بوقت محدد. اقرأ القطعة بالكامل بصوت عال للطفل **مرة واحدة فقط**. اقرأ ببطء (نحو كلمة واحدة في الثانية). اطرح جميع الأسئلة. لا تسمح للطفل بالنظر في القطعة أو الأسئلة. ضع علامة على إجابة الطفل وانتقل إلى السؤال التالي. ضع علامة صح فقط على الإجابات المتشابهة أو المماثلة في المعنى للأجوبة المقدمة بجوار كل سؤال. إن لم يجب الطفل على السؤال بعد مرور 10 ثوان، انتقل إلى السؤال التالي. لا تكرر السؤال.	نح كتيب تحفيز الطالب عن مرأى الطفل. اقرأ التعليمات للطفل. هذه المهمة الفرعية ليست مرتبطة بوقت محدد. اقرأ القطعة بالكامل بصوت عال للطفل **مرة واحدة فقط**. اقرأ ببطء (نحو كلمة واحدة في الثانية). اطرح جميع الأسئلة. لا تسمح للطفل بالنظر في القطعة أو الأسئلة. ضع علامة على كل سؤال سواء "صحيح" أو "خطأ". سيتحول اختيارك إلى اللون الأصفر. إن لم يجب الطفل بعد مرور 10 ثوان، ضع علامة "لا إجابة" وانتقل للسؤال التالي. الأجوبة ذات المعنى المشابه للأجوبة المقدمة يجب اعتبارها صحيحة. إذا قال الطفل "لا أعلم"، ضع علامة "غير صحيح".	**سوف أقرأ لك النص مرة واحدة ثم أوجه لك بعض الأسئلة . رجاء الاستماع بحرص والإجابة عن الأسئلة على أفضل وجه ممكن. هل فهمت المطلوب ؟** **بعد الانتهام من المهمة ، قل للطالب" أحسنت. جهد ممتاز! لننتقل إلى الجزء التالي.**
الإملاء		
سأقرأ عليك جملة قصيرة . من فضلك استمع جيدًا. سأقرأ الجملة كلها مرة واحدة. بعد ذلك، سأقرأ الجملة على أجزاء حتى تستطيع أن تسمع جيدًا.سأقرأ اجملة مرة ثالثةحتي يمكنك مراجعة الكلمات. **هل فهمت المطلوب ؟ الانتهام من المهمة ، قل للطالب" أحسنت. جهد ممتاز!** لننتقل إلى الجزء التالي.	اقلب نموذج إجابة الطفل إلى آخر صفحة مسطرة للكتابة وضعها أمام الطالب. سيكتب الطفل جملة الإملاء على الصفحة المسطرة من نموذج الإجابة.	اقلب نموذج إجابة الطفل إلى آخر صفحة مسطرة للكتابة وضعها أمام الطالب. سيكتب الطفل جملة الإملاء على الصفحة المسطرة من نموذج الإجابة.

تعليمات للمقيم: التقييم الورقي	تعليمات للمقيم: التقييم الإلكتروني	
في نموذج تحفيز الطالب، انتقل إلى الصفحة الأخيرة، حيث ستجد نفس التعليمات مثل الموضحة أدناه. قدم قلم رصاص للطفل. قل للطفل: **[ادخل التعليمات في العامود الأيمن لقراءتها للطفل]** اقرأ الجملة التالية **مرة واحدة**، بمعدل كلمة في الثانية. **[ادخل جملة الإملاء]** بعدها اقرأ الجملة مرة ثانية، وتوقف لمدة 10 ثواني بين مجموعات الكلمات. **[ادخل جملة الإملاء، بحيث تكون الكلمات متفرقة بعض الشيء للإشارة إلى التوقف فيما بينها.]** بعد 15 ثانية، اقرأ الجملة كاملة مرة أخرى. **[ادخل نفس جملة الإملاء]** انتظر نحو 15 ثانية إضافية لينتهي الطفل من الكتابة، بعدها أوقف المهمة الفرعية.	انتقل إلى الصفحة الأخيرة من نموذج تحفيز الطالب، حيث ستجد نفس التعليمات مثل الموضحة أدناه. قدم قلم رصاص للطفل. اقرأ الجملة **مرة واحدة**، بمعدل كلمة في الثانية. بعدها اقرأ الجملة مرة ثانية، وتوقف لمدة 10 ثواني بين مجموعات الكلمات. بعد 15 ثانية، اقرأ الجملة كاملة مرة أخرى. انتظر نحو 15 ثانية إضافية لينتهي الطفل من الكتابة، بعدها أوقف المهمة الفرعية.	

الملحق ط نموذج لجدول أعمال لتدريب أحد المقيمين

تدريب جامعي البيانات لأداة تقييم مهارات القراءة للمراحل المبكرة

اليوم والوقت	اليوم الأول	اليوم الثاني	اليوم الثالث	اليوم الرابع	اليوم الخامس	اليوم السادس
الأهداف اليومية	• فهم الهدف من تقييم مهارات القراء للمراحل المبكرة • التمكن من تطبيق الإدارة وقواعد تسجيل النتائج على الورق	• فهم وظائف الكمبيوتر اللوحي وإدارته • التمكن من رفع البيانات	• تحسين مهارات ادارة الاختبار • التعرف على إدارة الاستبيان	• صقل مهارات إدارة تقييم مهارات القراءة للمراحل المبكرة ودقة تسجيل النتائج	• صقل مهارات إدارة تقييم مهارات القراءة للمراحل المبكرة ودقة تسجيل النتائج	تدريب المشرفين • استعدادات الفريق
8:30-9:00 ص	• الترحيب/ المقدمات	• استعراض اليوم الأول	**الزيارة المدرسية الأولى: التدريب على تقييم مهارات القراءة للمراحل المبكرة**	**الزيارة المدرسية الثانية: تقييم مهارات القراءة للمراحل المبكرة+ استطلاعات**	**الزيارة المدرسية الثالثة: تقييم مهارات القراءة للمراحل المبكرة+ استطلاعات**	**تدريب المشرفين** • **استعدادات الفريق لجمع البيانات**
9:00-10:30 ص	• نظرة عامة على أداة تقييم مهارات القراءة في الصفوف الأولى: الهدف، الأداة، المحتوى • هدف تقييم مهارات القراءة للمراحل المبكرة في هذا السياق	• نظرة عامة على الوظائف الأساسية للكمبيوتر اللوحي				
10:30-11 ص	**راحة**	**راحة**				
11:00-1:00م	• نظرة عامة على الأداة • شرح المهام الفرعية والتدريب عليها	• التدريب على تقييم مهارات القراءة للصفوف الأول على الكمبيوتر اللوحي (المجموعات الصغيرة)				
1:00-2:00 م	**تناول الغداء**					
2:00-3:30 م	• استكمال شرح المهام الفرعية والتدريب عليها • استبيانات التلاميذ	• قضايا وظائف المجموعات • رفع البيانات	• استخلاص معلومات الزيارات المدرسية • أدوات مسح إضافية إذا تمت إدارتها	• استخلاص معلومات الزيارات المدرسية • مناقشة نتائج قياس دقة المقيم 2	• استخلاص معلومات الزيارات المدرسية • مناقشة دقة المقيم	

اليوم والوقت	اليوم الأول	اليوم الثاني	اليوم الثالث	اليوم الرابع	اليوم الخامس	اليوم السادس
				• التدريب على تقييم مهارات القراءة للمراحل المبكرة على أجهزة الكمبيوتر اللوحية بشكل ثنائي (المهام والقضايا الرئيسية)	• نتائج الإجراء الثاني • **لوجستيات جمع البيانات**	
3:30-3:45 م	**راحة**					
3:45-5:30م	• استكمال التدريب الاجمالي والتدريب المجموعات الصغيرة والتصحيح	• **إجراءات معاينة تقييم مهارات القراءة للمراحل المبكرة** • لوجستيات الزيارات المدرسية	• التدريب على تقييم مهارات القراءة للمراحل المبكرة على أجهزة الكمبيوتر اللوحية بشكل ثنائي (المهام والقضايا الرئيسية) • **قياس دقة المقيم** • نظرة عامة للوجستيات الزيارات المدرسية	• **قياس دقة المقيم الثاني**	• **قياس دقة المقيم الثالث**	

يعتمد عدد أيام التدريب ومحتوى الدورات إلى حد كبير على عدد من الأدوات التي ستتم إدارتها (أداة تقييم مهارات القراءة للمراحل المبكرة بالإضافة إلى بعض الاستبيانات أخرى، أو في لغات متعددة) وعدد المقيمين الذي يتم تدريبهم ومستوى خبرتهم. وإذا تعلم المقيمون إدارة تقييم مهارات القراءة للمراحل المبكرة بلغتين سيحتاج تدريبهم على هذا التقييم وقتا أطول. ونتيجة ينصح بتقليل عدد الزيارات المدرسية إلى زيارتين فقط لتوفير مزيد من الوقت أثناء الورشة لتعلم الأداة.

الملحق ى: تحليل البيانات والإرشاد الإحصائي لقياس دقة المقيمين

يقدم لنا هذا الملحق بعض التفاصيل عن إدارة البيانات لقياس دقة المقيمين ويتضمن كذلك بعض المصطلحات والإرشادات الإحصائية.

ى.1 إعداد البيانات

المرفق ى-1 هو مثال يوضح الأماكن التي يوجد بها فرق بين المعيار الذهبي والنمط على مستوى البند (وهي موضحة بالخلايا المظللة). وإذا حدث هذا يقوم الفريق بالبحث لمعرفة السبب. ومن التفسيرات الممكنة أن المعيار الذهبي قد يكون غير دقيق أو أنه كان هناك مشكلة بالأداة أو أنه كان هناك مشكلة في تفسير المتدربين لهذا البند وكان هو محل التركيز لمزيد من التدريب.

المرفق ى-1: نموذج لمخرج من برنامج مايكروسوفت إكسيل يقارن بين المعيار الذهبي و استجابة المقيم النمطية

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	enumerator	non_word_time_remain	non_word_attempted	non_word1	non_word2	non_word3	non_word4	non_word5	non_word6	non_word7	non_word8	non_word9	non_word10	non_word11	non_word12	non_word13	non_word14
2	GoldStdirr1	0	41	0	1	1	1	1	0	1	1	1	0	1	0	1	1
3	mode	0	41	1	1	0	1	1	1	1	1	1	0	1	1	1	1
4	mode vs. GS	.	.	1	.	1	.	.	1	.	.	.	.	.	1	.	.
5																	
6	aloreirr1	0	41													1	1
7	apanjirr1	0	42													1	1
8	ashooirr1	0	42													1	0
9	dmtitirr1	0	40													1	1
10	hseleirr1	0	41													1	1
11	ikiwairr1	0	41													1	1
12	jmasairr1	0	41													1	1
13	jurasirr1	0	41	1	1	0	1	1	1	1	1	1	0	1	1	1	1
14	kkahairr1	0	42	0	1	0	1	1	1	1	1	1	0	1	1	1	1

تأكد من صحة المعيار الذهبي عن طريق مقارنته باستجابة المقيمين النمطية

ى.2 تحليل البيانات

يتم بعد ذلك حساب نسبة الاتفاق عن طريق المهمة الفرعية. وهذا الإجراء هو الاتفاق بين تقييم المقيم لأحد الأطفال والتقييم الصحيح لنفس الطفل. ولحساب نتيجة المقيم (لكل مهمة فرعية وللتقييم ككل)، يقوم المسئول عن التدريب بمطابقة عدد الاتفاقات مع المعيار الذهبي ويعبر عن هذا الرقم في صورة نسبة مئوية من عدد البنود في المهمة الفرعية/التقييم كما هو موضح في **المرفق ى-2**

المرفق ى-1: نموذج لمخرج من برنامج مايكروسوفت إكسيل لحساب نسبة الاتفاق مع المعيار الذهبي، بالمهمة الفرعية

enumerator	Non word	non_word_time_remain	non_word_attempted	non_word1	non_word2	non_word3	non_word4	non_word5	non_word6	non_word7	non_word8	non_word9	non_word10	non_word11	non_word12	non_word13	non_word14
Average	88%	95%	59%	32%	100%	14%	100%	100%	0%	100%	100%	100%	73%	100%	0%	100%	95%
aloreirr1	91%	1	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1
apanjirr1	81%	1	0	0	1	0	1	1	0	1	1	1	1	1	0	1	1
ashooirr1	75%	1	0	0	1	0	1	1	0	1	1	1	1	1	0	1	0
dmtitirr1	89%	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	1
hseleirr1	91%	1	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1
ikiwairr1	91%	1	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1
jmasairr1	89%	1	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1
jurasirr1	91%	1	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1
kkahairr1	89%	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	1
lkayoirr1	85%	1	0	1	1	0	1	1	0	1	1	1	0	1	0	1	1
mkyejirr1	79%	1	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1
mndolirr1	93%	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1
mpaziirr1	91%	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1
mramairr1	91%	1	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1
nkibonairr1	79%	0	0	1	1	0	1	1	0	1	1	1	0	1	0	1	1

باستخدام الصيغة، يتم الحساب على النحو التالي:

$$\text{نتيجة المهمة الفرعية للمقيم (\%)} = \frac{\text{عدد الاتفاقات مع المعيار الذهبي}}{\text{عدد بنود المهمة الفرعية}}$$

ويمكن أيضا حساب متوسط الاتفاقات على مستوى البند من خلال المقيمين باستخدام الصيغة التالية:

$$\text{الاتفاق على مستوى البنود(\%)} = \frac{\text{عدد الاتفاقات مع المعيار الذهبي للبند}}{\text{عدد الاستجابات (المقيمين) للبند}}$$

وحتى إذا كان بالمعيار الذهبي بعض العناصر المفقودة بسبب أن "الطفل" لم يكمل كل بنود المهمة الفرعية، يجب أن تتضمن نتائج الاتفاق التي يضعها المقيم اتفاقا مع البنود المفقودة.

وبالنسبة للمهام الفرعية الموقوتة مثل طلاقة القراءة الشفهية وعدد أصوات الحروف الصحيحة في الدقيقة، إذا تمكن الطفل من اكمال المهمة الفرعية في الوقت المطلوب، من المهم أن يدون المقيم قراءة دقيقة للوقت الذي استغرقه الطفل لإكمال هذه المهمة. إذا كان المقيم في حدود الثانيتين، يمكن اعتباره متوافقا مع المعيار الذهبي ثم يتم حساب نسبة الاتفاق الكلية من خلال كل متغيرات الوقت المتبقي.

نسبة الاتفاق الكلية عن طريق المقيم هي متوسط الاتفاقات بين المهمة الفرعية ونسبة الوقت المتبقي. ويتم حساب نسبة الاتفاق الكلية للتقييم كمتوسط لنسبة المقيم الكلية.

وهكذا يتم الإبلاغ عن ملخص النتائج الخاص بكل تقييم ويتضمن ما يلي:

- عن طريق المقيم: نسبة الاتفاق بين المهمة الفرعية و التقييم ككل
- متوسط نسبة الاتفاق الكلية
- نسبة الاتفاق الكلية للمهمة الفرعية

ى. 3 مصطلحات وتعريفات إحصائية

نسبة الاتفاق الأولية

تقيس درجة إصدار المقيمين لنفس الحكم

كابا

تقيس إلى أي مدى يمكن أن يكون قد صدر تقييمين مختلفين لنفس الموضوع بالصدفة.
وتتراوح قيم كابا من -1،0 وحتة 1،0. وتشير القيم الأعلى لاحتمالية أقل للاتفاق بالصدفة.

معامل الارتباط بين المجموعات

تصف اتساق الدرجات الممنوحة للطلبة من قبل مقيمين مختلفين. وتتراوح قيم معامل الارتباط بين المجموعات من 0،0 حتى 1،0. وتشير القيم الأعلى لاتفاق أكبر بين المقيمين

ى.4 المعايير القياسية لاتفاق المقيمين

نسبة الاتفاق الأولية

نظرا لعدم وجود التفاصيل التي يتم إنشاؤها فقط من خلال هذه الإحصائية، لايوجد معيار قياسي ممكن. وتبذل الجهود من أجل الوصول بنسبة الاتفاق بين المقيمين لأعلى مستوى ممكن (أقرب إلى 100٪) عند تقييم الطلاب. ومع ذلك، بغض النظر عن نسبة الاتفاق، يجب أن يشير المقيمون لإحصاءات كابا لمعرفة مدى جودة إحصائية نسبة الاتفاق.

كابا

الاختيار الأول

المصدر: لانديس وكوش (1977)

إحصاء كابا	قوة الاتفاق
أقل من 0،0	ضعيف
0،0 إلى 0،20	طفيف
0،21 إلى 0،40	مقبول
0،41 إلى 0،60	متوسط
0،61 إلى 0،80	كبير
0،81 إلى 1،00	شبه مثالي

الاختيار الثاني

المصدر: فليس (1981)

إحصاء كابا	قوة الاتفاق
أقل من 0،40	ضعيف
0،40 إلى 0،75	متوسط إلى جيد
أكبر من0،75	ممتاز

مراجع ملحق ى

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.) New York: John Wiley.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

الملحق ك: خطط العينة لاختبار الثقة المتبادلة بين المقيمين على أساس ميداني

يصاحب هذا الملحق بروتوكول "انقذوا الأطفال" في القسم رقم 8 . 7 والذي يصف كيفية تقييم الثقة بين المقيمين بشكل مستمر أثناء مسح تقييم مهارات القراءة للمراحل المبكرة. وتوضح الجداول أدناه طريقة منهجية لتنويع ثنائيات التقييم في أول تقييم في اليوم في كل مدرسة، لمجموعة تتكون من ثلاثة أو أربعة أو خمسة مقيمين. وبينما يختلف إجمالي حجم العينة المطلوب لاختبار الثقة المتبادلة بين المقيمين حسب تصميم المسح (أي إجمالي عدد المدارس والطلاب الذين يتم تقييمهم) ينصح باختبار الطلاب بشكل ثنائي عندما يكون عدد الطلاب 150طالب على الأقل. فعندما يكون حجم العينة أقل من 100 لاختبار الثقة بين المقيمين فمن المرجح ألا ينتج أية معلومات مفيدة.

كل الرجداول والنصوص الداعمة:

© 2015 انقذوا الأطفال. مصرح باستخدامها. جميع الحقوق محفوظة

Save the Children

انقذوا الأطفال

تصورات لاختبارات الثقة بين المقيمين من خلال عدد المقيمين

جدول لثلاثة مقيمين

المدرسة	تقسيم المقيمين	يقيم ويسجل	يستمع ويسجل
المدرسة 1	(أ) و (ب) (ج)	المقيم أ المقيم ج	المقيم ب ----
المدرسة 2	(ب) و(ج) (أ)	المقيم ب المقيم أ	المقيم ج ----
المدرسة 3	(ج) و(أ) (ب)	المقيم ج المقيم ب	المقيم أ ----
المدرسة 4	(أ) و(ج) (ب)	المقيم أ المقيم ب	المقيم ج
المدرسة 5	(ب) و (أ) (ج)	المقيم ب المقيم ج	المقيم أ ----
المدرسة 6	(ج) و(ب) (أ)	المقيم أ المقيم ج	المقيم ب ----
الخ...			

جدول لأربعة مقيمين

المدرسة	تقسيم المقيمين	يقيم ويسجل	يستمع ويسجل
المدرسة 1	(أ) و (ب) (ج)و (د)	المقيم أ المقيم ج	المقيم ب المقيم د
المدرسة 2	(أ) و(ج) (ب) و(د)	المقيم أ المقيم ب	المقيم ج المقيم د
المدرسة 3	(أ) و(د) (ب)و(ج)	المقيم أ المقيم ب	المقيم د المقيم ج
المدرسة 4	(ب)و(أ) (د)و (ج)	المقيم ب المقيم د	المقيم أ المقيم ج
المدرسة 5	(ج) و (أ) (د)و(ب)	المقيم ج المقيم د	المقيم أ المقيم ب
المدرسة 6	(د) و(أ) (ج)و(ب)	المقيم د المقيم ج	المقيم أ المقيم ب
الخ...			

جدول لخمسة مقيمين

المدرسة	تقسيم المقيمين	يقيم ويسجل	يستمع ويسجل
المدرسة 1	(أ) و (ب)	المقيم أ	المقيم ب
	(ج)و (د)	المقيم ج	المقيم د
	(ه)	المقيم ه	---------
المدرسة2	(ه) و(أ)	المقيم ه	المقيم أ
	(ب) و(ج)	المقيم ب	المقيم ج
	(د)	المقيم د	---------
المدرسة 3	(د) و(ه)	المقيم د	المقيم ه
	(أ)و(ج)	المقيم أ	المقيم ج
	(ب)	المقيم ب	---------
المدرسة 4	(ج)و(ه)	المقيم ج	المقيم ه
	(ب)و (د)	المقيم ب	المقيم د
	(أ)	المقيم أ	---------
المدرسة 5	(أ) و (د)	المقيم أ	المقيم د
	(ه)و(ب)	المقيم ه	المقيم ب
	(ج)	المقيم ج	---------
المدرسة 6	(ب) و(أ)	المقيم ب	المقيم أ
	(د)و(ج)	المقيم د	المقيم ج
	(ه)	المقيم ه	---------
المدرسة 7	(أ) و(ه)	المقيم أ	المقيم ه
	(ج)و(ب)	المقيم ج	المقيم ب
	(د)	المقيم د	---------
المدرسة 8	(ه) و(د)	المقيم ه	المقيم د
	(ج) و(أ)	المقيم ج	المقيم أ
	(ب)	المقيم ب	---------
المدرسة 9	(ه) و (ج)	المقيم ه	المقيم ج
	(د) و (ب)	المقيم د	المقيم ب
	(أ)	المقيم أ	---------
المدرسة 10	(د) و (أ)	المقيم د	المقيم أ
	(ب) و (ه)	المقيم ب	المقيم ه
	(ج)	المقيم ج	---------
الخ...			

إذا كان فريق التقييم يتألف من عدد فردي من المقيمين، يجب أن يتم تبادل أفراد الفرق كل يوم، بحيث يتم ترك شخصا واحدا في كل يوم دون أن يشترك في اختبار الثقة بين المقيمين لهذا اليوم. يجب إنشاء جدول للتقييم مثل الموضحين أعلاه، ومن المهم أن يتم إنشاء هذا الجدول لتلافي أي لبس من جانب المقيمين.

لأي أسئلة أو استفسارات اتصل بفريق البحث بإدارة التعليم وحماية الطفل من خلال البريد الإليكتروني learningassessment@savechildren.org.

الملحق ل: الكتاب الكودي للعينة

القسم: السكاني	الصيغة	اسم التعريف	تعريف القيمة	تعريف المتغير
Country	سلسلة	—	(أكبر متغير جغرافي)	في أي بلد تم عمل هذا التقييم؟
Project	سلسلة			ما هو المشروع الذي تم تنفيذه داخل البلد؟
Year	عدد صحيح (2020-2000)	—	—	في أي سنة تم عمل هذا التقييم؟
Month	ترتيبي (1-12)	month	1يناير2فبراير...12 ديسمبر	في أي شهر تم عمل هذا التقييم؟
Date	صيغة التاريخ	—	—	ما هو تاريخ هذا التقييم؟
State	اسمي	state	قائمة خاصة بالبلد (ثاني أكبرمتغير جغرافي أسفل البلد)	في أي ولاية تقع مدرسة الطالب؟
Region	اسمي	region	قائمة خاصة بالبلد (ثالث أكبرمتغير جغرافي أسفل الولاية)	في أي منطقة تقع مدرسة الطالب؟
District	اسمي	district	قائمة خاصة بالبلد (أصغرمتغير جغرافي أسفل المنطقة)	في أي حي تقع مدرسة الطالب؟
School_name	سلسلة	school	قائمة خاصة بالبلد	ما اسم مدرسة الطالب؟
School_code	عدد صحيح	—	قائمة خاصة بالبلد	رمز المدرسة داخل البلد
EMIS	عدد صحيح	—	—	رمز نظام معلومات إدارة التعليم
School_type	اسمي	school_type	ضع تعريف القيمة تبعا للمشروع	ما نوع المدرسة التي يدرس بها الطالب؟
Treatment	ثنائي	treatment	0 "المرجع" 1 "العلاج الجزئي" 2 "العلاج الكامل"	ما هو مستوى العلاج الذي تتلقاه المدرسة
Treat_year	ترتيبي (0-12)	—	—	كم سنة أمضتها المدرسة في تلقي العلاج؟
Treat_phase	ترتيبي (1-6)	treat_phase	ضع تعريف القيمة تبعا للمشروع	في أي مرحلة من مراحل الدراسة يوجد طالب المدرسة التي تتلقى العلاج

القسم: السكاني	الصيغة	اسم التعريف	تعريف القيمة	تعريف المتغير
Urban	ثنائي	urban	0 ريفي 1حضري	هل تقع المدرسة في منطقة حضرية؟
Shift	ترتيبي (0-2)	shift	0"لا توجد فترات" (يوم كامل) 1 صباحية 2 مسائية 3 التناوب	هل يدرس الطالب في المدرسة بنظام الفترات؟
Dbl_shift	ثنائي	yes/no	0لا 1 نعم	هل تعمل المدرسة على فترتين؟
Admin	اسمي	admin	قائمة خاصة بالبلد	من أدار الاختبار؟ (الرقم الكودي)
Admin_name	سلسلة	—	---	من أدار الاختبار؟
ID	سلسلة	—	يجب أن يكون فريدا و لا نظير له!!!	رقم هوية الطالب الفريد
Grade	عدد صحيح (-1 8)	grade	1 الأول ،2الثاني ،3 الثالث، 4 الرابع، 5 الخامس، 6 السادس، 7 السابع، 8 الثامن	ما هو الصف الملتحق به الطالب؟
Level	عدد صحيح	—	*مثل الصف، ولكن بالنسبة للطلاب الذين لاتكون أعمارهم تقليدية*	هل هي طالبة (أنثى)؟
Section	عدد صحيح	—	قائمة خاصة بالبلد	هل فصل الطالب يضم أكثر من صف دراسي؟
Female	ثنائي	female	0ذكر 1 أنثى	ما اسم معلم الطالب؟
Multigrade	ثنائي	yes/no	0لا 1نعم	كم يبلغ عمر الطالب؟
Teacher	عدد صحيح	teacher	قائمة خاصة بالبلد	وقت بدء التقييم؟
Age	عدد صحيح (18-5)	—	—	وقت نهاية التقييم؟
Start_time	الوقت (الساعة:الدقيقة)	—	—	كم دقيقةاستغرقها التقييم لاستكماله؟
End_time	الوقت (الساعة:الدقيقة)	—	—	لغة التقييم
Assess_time	الوقت (الدقيقة)	—	—	هل أعطى المشارك في التقييم موافقته على استكمال التقييم؟
Language	عدد صحيح	language	استخدم أكواد ISO 639-3	
consent	ثنائي	yes/no	0لا 1نعم	

ملحق م: توصيات للمساواة

في ورشة العمل التي أقامتها الوكالة الأمريكية للتنمية في 2015 بعنوان "تحسين جودة بيانات أداة "تقييم مهارات القراءة للمراحل المبكرة": مشاورة من أجل إطلاع الوكالة الأمريكية للتنمية على بعض الإرشادات بشأن إدارة تقييمات القراءة للمراحل المبكرة "، قام المتحدثون من الخبراء والمشاركون بالورشة بمناقشة موضوع مساواة المهام الفرعية لنفس اللغة لعدة نماذج من الأداة. ويوجد بالأسفل التوصيات التقنية التفصيلية للجنة المساواة وكذلك بعض النقاط التي تحتاج لمزيد من المشاورات.

م.1 التوصيات

1. **بالنسبة للمهام الفرعية ذات البنود القليلة (من 10إلى 25 على سبيل المثال)، قم بتجربة نماذج متعددة أعدت حديثا جنبا إلى جنب مع النماذج الأساسية.** ثم قارن بين الإحصائيات على مستوى البند الواحد عبر النماذج واستخدم هذه المعلومات (القيم الاحتمالية ومعامل بوينت بايسيريال) من أجل إنشاء نماذج منتصف المدة /وختامية تحاكي بشكل وثيق إحصائيات النموذج الأساسي. وهذا نهج مبسط للأشخاص العاديين ما قبل المساواة باستخدام منهج الاختبار التقليدي.

2. **لا تقم بتطبيق مناهج المساواة باستخدام منهج الاختبار التقليدي على المهام الفرعية ذات البنود القليلة.** ومبرر هذا واضح بالنسبة للمهام الفرعية التي يكون عدد بنودها من 3 إلى 5 بنود (مثل الاستماع والقراءة والفهم) ولكنها تصبح نقطة تحتاج مزيد من المناقشة عندما يكون عدد بنودها من 10 إلى 25 بند. بينما قد يكون من الممكن مساواتها باستخدام مناهج التجاوب مع المفردات اللغوية، وهي تتطلب حجم عينات لا يقل عن 500 حتى 1000طالب لنماذج أكثر تعقيدا (أي اثنين أو ثلاثة من النماذج الباراميترية). وتتطلب نماذج راش ومناهج الاختبار التقليدي أحجام عينات مماثلة ليكون الاختيار بينهم مسألة تحقيق الافتراضات (ومدى ملائمة تحليل البيانات على مستوى البنود)

3. **للحصول على بيانات خطية وعينات صغيرة استخدم مناهج المساواة عن طريق الاختبار التقليدي.** وفي هذه الحالات يمكن استخدام التجريب على الأشخاص العاديون على فقرات طلاقة القراءة الشفهية ويمكن تطبيق مناهج المساواة المتوسطة أو الخطية

(واختيار أحدهما يكون تبعا للملائمة البصرية والتحيز والخطأ). مع البيانات غير الخطية، تصبح العملية أكثر تعقيدا. وهذا أكثر ملائمة من المساواة عن طريق نظرية التجاوب مع المفردات اللغوية، بالنظر إلى أن إجراء طلاقة القراءة الشفوية يعطي نتيجة إجمالية (بدون بيانات مفيدة على مستوى البنود).

4. **تأكد من أن العينات التجريبية والتشغيلية متماثلتين قدر الإمكان.** وبما أن كثير من مناهج تقييم مهارات القراءة للمراحل المبكرة تعتمد على الأشخاص العاديون أو التجربة على عينات عشوائية معادلة (وخاصة لطلاقة القراءة الشفوية)، فيجب أن تمثل النسخة التجريبية النسخة التشغيلية قدر الإمكان بحيث تكون تعديلات المساواة التي يتم تطبيقها على النسخة التجريبية مناسبة للبيانات التشغيلية.

5. عند استخدام المئينات المتساوية لمساواة بيانات طلاقة القراءة الشفهية الغير خطية، **تأكد من أن العينة تتضمن طلاب حققوا كل الدرجات الممكنة-**مما يتطلب عادة حجم عينة أكبر من تلك التي يتم استخدامها في التجربة على الأشخاص العاديون للدراسات القائمة على تقييم مهارات القراءة للمراحل المبكرة.

مرفق م-1 يعرض بعض التوصيات فيما يخص المتغيرات الموجزة لأداة تقييم مهارات القراءة للمراحل المبكرة التي يمكن أو لا يمكن مساواتها باستخدام مناهج المساواة التقليدية مع العينات الصغيرة. ولايزال هناك مجالا للنقاش في هذا الجدول وخاصة فيما يخص النتائج الصفرية والمهام الفرعية التي تحتوي على ما بين 10 و25 بند وكذلك نسبة الإجابات الصحيحة.

مرفق م-1 المتغيرات الموجزة لأداة تقييم مهارات القراءة للمراحل المبكرة المراد مساواتها (توصيات)

المهام الفرعية لتقييم مهارات القراءة للمراحل المبكرة	عدد البنود	النتائج الموقوتة	الدرجة //النتيجة الصفرية	الدرجة	الدرجة بالنسبة المئوية	عدد الأسئلة التي حاول الطالب الإجابة عليها	نسبة الإجابات الصحيحة
مقطع صوتي/لفظي	~ 20	التجاوب مع المفردات اللغوية - راش	لا	لا	لا	لا	لا
مفردات	10-5		لا	لا	لا	لا	لا
أسماء الحروف	100	عنصر تثبيت	لا	عنصر تثبيت	عنصر تثبيت	لا	لا
أسماء الأصوات	100	عنصر تثبيت	لا	لا	عنصر تثبيت	لا	لا
الكلمات المألوفة	50	عنصر تثبيت	لا	عنصر تثبيت	عنصر تثبيت	لا	لا
كلمات زائفة	50	عنصر تثبيت	لا	عنصر تثبيت	عنصر تثبيت	لا	لا
طلاقة القراءة الشفوية	~50	الأشخاص العاديون (المئينات المتساوية)	لا	الأشخاص العاديون (المئينات المتساوية)	الأشخاص العاديون (المئينات المتساوية)	لا	لا
الفهم القرائي	~5	التجاوب مع المفردات اللغوية – راش*	لا	لا	لا	لا	لا
الفهم السمعي	~5	التجاوب مع المفردات اللغوية – راش*	لا	لا	لا	لا	لا
إملاء	15-10	التجاوب مع المفردات اللغوية – راش	لا	لا	لا	لا	لا
متاهة	15-10	التجاوب مع المفردات اللغوية – راش	لا	لا	لا	لا	لا

* = يحتاج لمزيد من البحث
ملحوظة: هذه المناهج موصى بها فقط عندما تكون النسخة التجريبية تتبع توزيعا مماثلا للتوزيع الخاص بالمسح الكامل (يتم هذا عن طريق المعاينة العشوائية)

م.2 نقاط تحتاج لمزيد من البحث

1. **ما هي المناهج التي يمكن استخدامها لمساواة المهام الفرعية للفهم القرائي (أو المهام الفرعية الأخرى التي بها حوالي 5 بنود)؟** لا تجدي المناهج الاختبار التقليدية نفعا في هذه الظروف، لكن المناهج التي تقوم على الخطوات أو بعض المناهج الغير خطية يجب أن يتم بحثها. أما منهج التجاوب مع المفردات اللغوية فيحتاج لمزيد من البحث لتحديد مدى جدواها مع أقصر المهام الفرعية.

2. **ما هي المناهج التي يمكن استخدامها لمساواة فقرات طلاقة القراءة الشفوية عندما يكون هناك دليل على علاقات غير خطية عبر النماذج؟** هناك دليل واعد على ملائمة كلا من قوس الدائرة و المساواة عن طريق المئينات المتساوية في هذه المواقف، ولكن بالرغم

من ذلك توجد بعض القيود على كل منهما لذا فهما يحتاجان لمزيد من البحث.

3. **ما هي الأولوية النسبية إذا فاضلنا بين منهجي المئينات المتساوية والاختبار التقليدي؟** تتضمن هذه المسائل الخبرة التقنية وحجم العينة وإجراءات التجربة وهكذا. في نهاية الأمر، عندما يتم تسجيل البيانات على مستوى البند، ينبغي تفضيل تحليل راش للعينات صغيرة.

4. **كيف يجب التعامل مع النتائج الصفرية أثناء المساواة؟** هل يجب استبعاد الطلاب الحاصلين على نتيجة صفرية في كل نماذج التقييم من حسابات المساواة (أم يجب استبعاد أي طالب يحصل على نتيجة صفرية في أيا من نماذج التقييم)؟ هل من الممكن أن تكون نتيجة أحد نماذج الاختبار صفر ولكن يكون لدينا تعديل مساواة نتيجته أعلى من صفر لنفس الطالب؟ ما مدى اعتمادية التعامل مع النتائج الصفرية وتطبيقها على منهج المساواة؟

5. **ما هي الآثار المترتبة على استخدام البيانات التجريبية مقابل البيانات التشغيلية للمساواة؟** في معظم الحالات، نحن تقتصر على استخدام البيانات التجريبية لمساواة طلاقة القراءة الشفوية، ولكن ما هي الاولويات النسبية في الظروف التي تسمح لنا باستخدام كلا المنهجين؟ وبما أن ما بعد المساواة (أي البيانات التشغيلية) في الغالب توفر علاقات مساواة يمكن الاعتماد عليها بشكل أكبر، إذا فهل هناك أي مبرر للاعتماد على ما قبل المساواة (أي البيانات التجريبية) عندما يكون كلا الخيارين متاحين؟

6. **كيف يمكن تحليل النتائج التي تمت مساواتها؟** إذا تمت مساواة النتائج بشكل منفصل، هل يجب أن يتم تحليلهم بشكل منفصل أيضا (وبالتالي يلغي التحليل الشامل/المشترك)؟ إذا تمت مساواة النتائج الأولية، هل يمكن تحليل النسبة الصحيحة من البنود التي حاول الطلاب حلها؟

7. **يجب التحقيق في آثار الترتيب للاختبارات التي يتم تصنيف فيها البنود في صفوف.** فتجميع البنود الصعبة معا بشكل غير متعمد قد يؤثر على نتائج الاختبار. وهذا يحتاج لمزيد من البحث.

8. **هناك حاجة لاستكشاف الأبعاد بين المهام الفرعية.** إذا كان هناك مستوى معقول وواضح من أحادية الأبعاد إذا فمن الممكن عن طريق مساواة بعض المهام الفرعية استقراء البعض الآخر. خلاف ذلك يجب أن تقتصر المساواة على مستوى المحتوى (أي المهام الفرعية) مما قد يحد من التعميم فيما يتعلق بإنجاز القراءة الإجمالي.

الملحق ن: توصيات تقنية مفصلة بشأن ملفات الاستخدام العام

القسم 10 .6 من هذا الدليل يوضح الخطوات المطلوبة قبل إتاحة بيانات تقييم مهارات القراءة للمراحل المبكرة للجمهور. ويعرض هذا المرفق التوصيات الفنية التفصيلية الإضافية من اللجنة بشأن ملفات الاستخدام العام المتفق عليها في ورشة عمل الوكالة الأمريكية للتنمية التي أقيمت في 2015 بعنوان "تحسين جودة بيانات تقييم مهارات القراءة للمراحل المبكرة: استشارة لاطلاع الوكالة الأمريكية للتنمية على بعض التوجيهات بشأن إدارة تقييم مهارات القراءة للمراحل المبكرة"

ن.1 توصيات محددة لتنظيف البيانات ووضعها في صورتها النهائية وإخفاء هوية أصحابها

ن.1.1 التنظيف

1. وتعزز الوكالة الأمريكية للتنمية من فوائد اعتماد دليل تكويد رئيسي للشركاء المقيمين /المنفذين

2. ويستخدم دليل التكويد المعد من قبل مشروع بيانات التعليم لصانعي القرار (إيدتا 2) التابع للوكالة الأمريكية للتنمية [34] كأساس لدليل التكويد الرئيسي. وتوصى اللجنة بإنشاء دليل تكويد للأدوات التكميلية، مثل المعلومات الديموغرافية في الاستبيان الطلابي

3. كلما كان ذلك ممكنا، يتم تعريف أسماء المتغيرات ب 12 حرفا كحد أقصى وتعريف العلامات المتغيرة ب 80 حرف كحد أقصى .

4. وملفات الاستخدام العام هي ذاتية الوصف ذات بيانات فئوية وتقوم باستخدام الفئات كقيم بدلا من الكود الرقمي.

[34] التكويد (مركز البحوث المثلثية، 2014) في صورة جداول إكسيل متاح من خلال الموقع https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=389

5. من أجل تجنب التعميمات التي لا أساس لها، يجب إزالة البيانات الخاصة بكافة المناطق الجغرافية التي لم تستخدم لأغراض أخذ العينات وينتمي لها عدد قليل جدا من المدارس في العينة للحصول على تقديرات الدقة المناسبة (مثل المنطقة، منطقة العد، المكان، والحي)

ن.1. 2 التنظيف

في الحالات التي استخدمت منهجية العينة المركبة ، يجب أن يتم تعيين مواصفات المسح في ملف البيانات الخاص بملف الاستخدام العام للحد من أخطاء صياغة النماذج من قبل المستخدم العام، كلما كان ذلك ممكنا.

من أجل التخفيف من أخطاء صياغة النماذج خلال تحليل البيانات من قبل المستخدم العام، يتم استخدام مجموعة تحليل البيانات كأساس لملف الاستخدام العام. و يقوم الباحثون بدمج ملفات البيانات (على سبيل المثال، بيانات المعلم والطالب) أو إلحاقها (على سبيل المثال، البيانات الأساسية والختامية) حيثما كان ذلك ممكنا ومناسبا لتجنب احتياج المستخدم العام للتعامل مع عدة ملفات للبيانات.

ن.1 .3 إزالة بيانات الهوية

يجب إزالة كل المعلومات الشخصية من مجموعات البيانات قبل وضعها في متناول الجمهور. وفيما يلي توصيات عامة لإزالة وإخفاء الهوية من أي معلومات تعريفية.[35]

1. يجب إزالة المعلومات التي تحدد الهوية الشخصية مثل الاسم وعنوان المنزل ورقم الهاتف والرقم القومي.
2. يجب إزالة أسماء المدارس وأسماء أي من المؤسسات والأفراد الأخرى التي قد تم جمعها خلال عملية جمع البيانات.
3. يجب إزالة كل المعلومات المستخدمة في الاتصال والعثور على المدارس أو المؤسسات (مثل العنوان ورقم الهاتف واسم مدير المدرسة، وإحداثيات تحديد الموقع((GPS)
4. قد تتضمن البيانات المستخدمة لأغراض أخذ العينات معلومات شخصية: يجب إخفاء هوية البيانات وليس تدميرها. فمن المهم للحفاظ على مجموعة محددة من البيانات

لمطابقة قيم المتغيرات مجهولة المصدر مع قيم المتغيرات معروفة المصدر.

5. يجب إخفاء هوية كل المتغيرات التي تحتوي على الرموز الرسمية للبلاد (على سبيل المثال، المدرسة أو رمز المؤسسة أو رمز المعلم)

ن.2 نشر بيانات ملف الاستخدام العام

ومن المتوقع أن تقوم الوكالة الأمريكية للتنمية بإتاحة ملفات الاستخدام العام التي تحتوي على بيانات تقييم القراءة للمراحل المبكرة للجمهور من خلال التحليل الثانوي لتتبع نتائج بوابة التعليم (SART Ed) ومكتبة بيانات التنمية (DDL).

ومن أجل تسهيل استكشاف الجمهور للبيانات، توصي اللجنة أنه بالإضاقة لملفات البيانات يجب توفير وثائق مصاحبة أيضا.

وإلى جانب ملف الاستخدام العام يجب أيضا نشر المعلومات الموثقة جيدا التي تساعد الجمهور أو المستخدم العام على التآلف مع البيانات التي يجب أن يتم نشرها. ويجب توفير المعلومات التالية للمستخدمين:

1. تقرير تحليل البيانات المكتوبة الذي تم تسليمه للوكالة الأمريكية للتنمية والتصديق عليه من قبلها.
2. كل الاستبيانات وأدوات التقييم المستخدمة في جمع البيانات (ويجب أن يتم تسليمها فقط بعد انتهاء المشروع حتى لا يتم الكشف عن المواد التي قد يتم استخدامها في دراسات مستقبلية في نفس المشروع)
3. المعلومات الخلفية وكل الوثائق ذات الصلة

بالإضافة إلى اشتراط وجود التقارير و أدوات جمع البيانات ، تؤكد الوكالة الأمريكية للتنمية مجددا لشركاء التقييم والشركاء المنفذين على أهمية توثيق أسماء وأوصاف المتغيرات الرئيسية والإعدادات اللازمة لتحليل البيانات بشكل مناسب، بما في ذلك:

[35] للاطلاع على المناقشة المفصلة، انظر الملحق أ من مجموعة اوبتيمال سولوشنز (2015).

1. وضع تعريف واضح للسكان المستهدفين، بما في ذلك مصدر إطار القائمة التي تم استخدامها لاختيار العينة. وتشير الوثائق إلى أن العدد الإجمالي للمدارس وعدد تقديري من الطلاب الذين من المفترض أن تمثلهم العينة. وهذه الأرقام تتطابق أيضا مع تقديرات البيانات المرجحة. وإذا كان المسح ينطوي على تدخل / مراقبة، يتم الإبلاغ عن الأرقام من قبل التدخل / المراقبة.

2. المتغيرات المطلوبة لتحليل البيانات المركبة بشكل مناسب استنادا إلى منهجية العينة (على سبيل المثال: لكل مرحلة من مراحل المعاينة: كل البنود التي تمت معاينتها والمتغير الطبقي ومتغير تصحيح المجتمع المحدود وكذلك متغير الترجيح النهائي)

3. متغيرات من أجل تصميم البحث (على سبيل المثال: العلاج والسنة والجماعة إذا كان تصميم البحث عبارة عن تقييم مرحلي لتقييم التأثير)

4. شرح لإعدادات منهجية العينة (على سبيل المثال، معالجة التباين)، بحيث يمكن استخدام خصائص تصميم المسح بشكل مستقل عن البرمجيات الاحتكارية.

5. دليل تكويد كامل يحتوي على:

أ قائمة بكل المتغيرات في مجموعة البيانات.

ب كل مسميات المتغيرات وأشكالها، وكذلك مسميات القيم (إذا اقتضى الأمر)

ج وصف منهجي للعملية الحسابية التي تم استخدامها لتوليد المتغيرات المحسوبة (على سبيل المثال، طلاقة القراءة الشفهية).

د العدد الإجمالي للملاحظات في مجموعة البيانات.

بينما مستودعات البيانات المخطط لها والخاصة بالوكالة الأميركية للتنمية مازالت قيد الإنشاء، جعل الشركاء المنفذين والمقيمون ملفات الاستخدام العام وبيانات تقييم القراءة للمراحل المبكرة متاحة للجمهور.

1. انشر ملف الاستخدام العام في موقع يسهل الوصول إليه عبر الإنترنت، مصحوبا بالوثائق الخاصة به على النحو المشار إليه أعلاه (أي، وضع جميع العناصر في ملف مضغوط واحد أو يجب أن يحتوي الموقع على روابط لهذه الوثائق).

2. إنشاء ملف الاستخدام العام باستخدام ملف بيانات غير احتكاري إنشاء ملف البيانات الخاصة، عندما يكون ذلك ممكنا.

3. للحصول على ملف غير احتكاري، قم إنشاء ملف CSV (ملف قيم مفصولة بفاصلة)

4. للحصول على ملف غير احتكاري، يمكن أن نقوم بإنشاء ملف Stata .dta و/أو ملف SPSS .sav (جنبا إلى جنب مع ملف SPSS.csaplan).

وتشجع اللجنة أيضا الوكالة الأمريكية للتنمية على النظر في وضع مبادئ توجيهية لتقارير التقييم استنادا إلى تقارير تقييم مهارات القراءة للصفوف المبكرة المماثلة ، على غرار التوجيهات العامة للوكالة الأمريكية للتنمية لإعداد تقارير التقييم (الوكالة الأمريكية للتنمية، 2012). لأن هذا من شأنه أن يجعل من السهل على الجمهور تحديد موقع المعلومات الواردة في التقارير. وكذلك فإنه يضمن أنه يمكن الاطلاع على نفس المعلومات الأساسية عبر التقارير.

مراجع لملحق ن

Optimal Solutions Group, LLC. (2015). *Secondary Analysis for Results Tracking (SART) data sharing manual, USAID Ed Strategy 2011–2015, Goal 1.* Prepared for USAID under the Secondary Analysis for Results Tracking (SART) project, Contract AID-OAA-C-12-00069. Location: Optimal Solutions. Retrieved from https://sartdatacollection.org/images/SARTDataSharingManualFeb2015.pdf

RTI International. (2014a). *Codebook for EGRA and EGMA* [Excel spreadsheet]. Research Triangle Park, NC: RTI. Retrieved from https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=389

United States Agency for International Development (USAID). (2012). *How-to note: Preparing evaluation reports.* Monitoring and Evaluation Series, No. 1, Version 1.0. Washington, DC: USAID. Retrieved from https://www.usaid.gov/sites/default/files/documents/1870/How-to- Note_Preparing-Evaluation-Reports.pdf

الملحق س: تحليل بيانات تقييم مهارات القراءة للصفوف الأولى

يجب أن يتم تقديم مرئية لتقسيم النتيجة لكل نتيجة طالب من الطلاب مثل تلك المبينة في المرفقات س-1 وحتى س-3 من أجل دعم فهم القارئ للتقدير الذي تم تقديمه.

مرفق س-1 نموذج لتحليل الاختلافات في الفارق

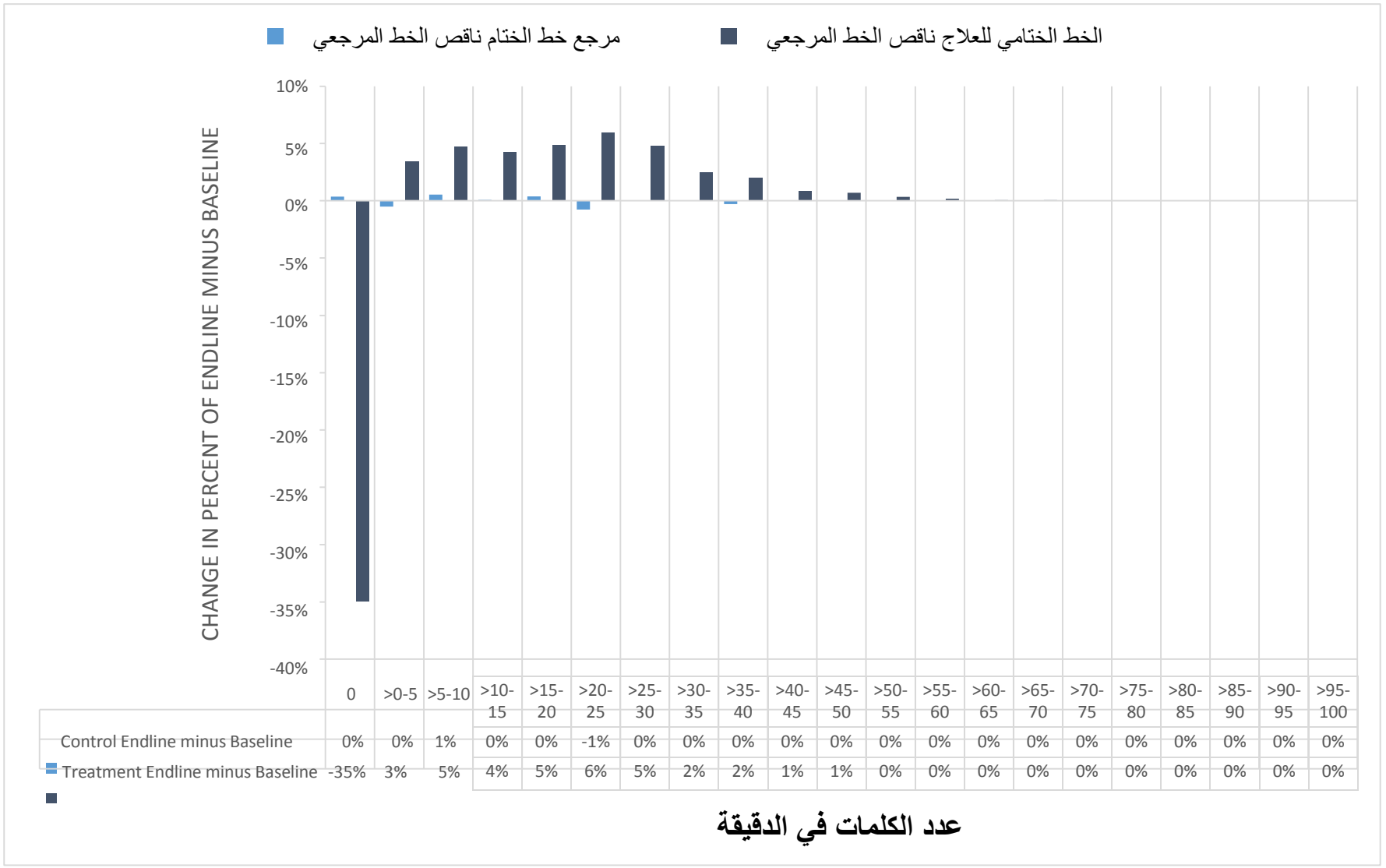
	خط الأساس					الخط الختامي								
العلاج	متوسط الطلاقة (عدد الكلمات في الدقيقة)	الخطأ المعياري	عدد الطلاب الذين يتضمنهم العينة	إحصاء تي	القيمة المحتملة	العلاج	متوسط الطلاقة (عدد الكلمات في الدقيقة)	الخطأ المعياري	عدد الطلاب الذين يتضمنهم العينة	إحصاء تي	القيمة المحتملة	الاختلاف في الفارق	القيمة المحتملة	القيمة المحتملة
المرجع	4،5	0،6	656	---	---	9،5	1،6	475	---	---	---	---	---	---
التدخل	5،2	1،2	349	0،510	0،611	11،7	1،1	480	1،189	0،236	1،5	0،490	0،12	

العمليات الحسابية:

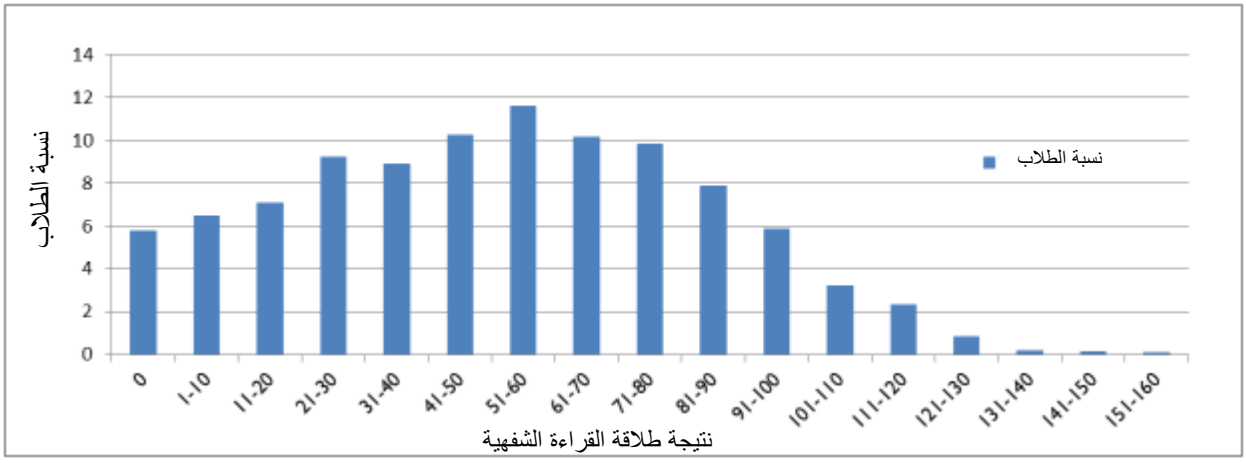
الاختلاف في الفارق: (متوسط علاج الخط الختامي- متوسط علاج خط الأساس)-(متوسط مرجع الخط الختامي- متوسط مرجع خط الأساس)

حجم التأثير (فارق كوهين): الاختلاف في الفارق/تجمع الانحراف المعياري

المرفق س-2 نموذج للمقارنة التوزيعية بين الفرق بين المرجع والعلاج



المرفق س-3 توزيع طلاقة القراءة الشفهية-أندونيسيا، 2013



المصدر: Stern, J. & Nordstrum, L. (2014). *Indonesia 2014: The National Early Grade Reading Assessment (EGRA) and Snapshot of School Management Effectiveness (SSME) survey*. Prepared for USAID/Indonesia under the Education Data for Decision Making (EdData II) project, Task Order No. AID-497-BC-13-00009 (RTI Task 23). Research Triangle Park, NC: RTI International. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=680

ملحق ف: بيانات هاسبروك وتيندال للطلاقة في القراءة الشفهية

"الجدول المبين أدناه يقدم مثال على أعراف طلاقة القراءة الشفهية اعتماداً على الأبحاث التي أجريت بالنسبة للغة الإنجليزية. يندرج هذا الجدول في مجموعة أدوات اللغة العربية كمثال على ما يمكن تحقيقه في لغات أخرى."

Hasbrouck & Tindal Oral Reading Fluency Data

This table shows the oral reading fluency rates of students in grades 1 through 8, based on an extensive study conducted by Jan Hasbrouck and Gerald Tindal. The results of their study are published in a technical report entitled, "Oral Reading Fluency: 90 Years of Measurement," which is available on these websites:

- **ERIC website:** eric.ed.gov/?id=ED531458
- **BRT website:** www.brtprojects.org/publications/technical-reports

This table can help you assess the oral reading fluency of your students relative to their peers. Students scoring 10 or more words below the 50th percentile using the average score of two unpracticed readings from grade-level materials need a fluency-building program. Teachers can also use the table to set long-term fluency goals for struggling readers.

For more information:

- **Essential Components of Reading:** readnaturally.com/components
- **Correlation Between Oral Reading Fluency and Overall Reading Achievement:** readnaturally.com/correlation
- **Read Naturally Tools for Assessing Fluency:** readnaturally.com/assessment-tools
- **Read Naturally Intervention Programs That Develop Fluency:** readnaturally.com/fluency-interventions

Grade	Percentile	Fall WCPM*	Winter WCPM*	Spring WCPM*	Avg. Weekly Improvement**
1	90		81	111	1.9
	75		47	82	2.2
	50		23	53	1.9
	25		12	28	1.0
	10		6	15	0.6
2	90	106	125	142	1.1
	75	79	100	117	1.2
	50	51	72	89	1.2
	25	25	42	61	1.1
	10	11	18	31	0.6
3	90	128	146	162	1.1
	75	99	120	137	1.2
	50	71	92	107	1.1
	25	44	62	78	1.1
	10	21	36	48	0.8
4	90	145	166	180	1.1
	75	119	139	152	1.0
	50	94	112	123	0.9
	25	68	87	98	0.9
	10	45	61	72	0.8
5	90	166	182	194	0.9
	75	139	156	168	0.9
	50	110	127	139	0.9
	25	85	99	109	0.8
	10	61	74	83	0.7
6	90	177	195	204	0.8
	75	153	167	177	0.8
	50	127	140	150	0.7
	25	98	111	122	0.8
	10	68	82	93	0.8
7	90	180	192	202	0.7
	75	156	165	177	0.7
	50	128	136	150	0.7
	25	102	109	123	0.7
	10	79	88	98	0.6
8	90	185	199	199	0.4
	75	161	173	177	0.5
	50	133	146	151	0.6
	25	106	115	124	0.6
	10	77	84	97	0.6

*WCPM = Words Correct Per Minute

www.readnaturally.com

**Average words per week growth

United States Agency for International Development
Officeof Education
Bureau for Economic Growth, Education, and Environment (E3)
1300 Pennsylvania Avenue, N.W.
Washington, DC 20523, USA
www.USAID.gov