الإطار الوطني لتقييم الطلبة في الأردن

2024

الإطار الوطني لتقييم الطلبة في الأردن

2024

إعداد فريق العمل

برئاسة الأستاذ الدكتور يوسف سوالمه

وعضوية كل من

الدكتور جهاد عناتي	الدكتور لؤي شواشرة
الأستاذ محمد شحادة	الدكتور محمد المومني
الدكتور جعفر الموسى	الدكتورة فاطمة يدك
الأستاذ محمود هزايمة	

وبمتابعة من

الدكتورة خولة حطاب والمهندس بشار كفافي والدكتورة سلام معايعة

جدول المحتوى

1. الإطار الوطني لتقييم الطلبة في الأردن 2024 4
2. الجزء الأول: الوضع الراهن لتقييمات الطلبة في النظام التعليمي الأردني 6
3. أولًا: التقييمات الصفية في الأردن(Classroom Assessments in Jordan) 7
4. واقع التقييمات الصفية في الأردن 7
5. الترتيبات المؤسسية واللوجستية/التقنية والتمويل للتقييمات الصفية في الأردن 13
6. أليات كتابة التقارير والإبلاغ عن النتائج في التقييمات الصفية في الأردن 13
7. التقييمات الصفية في الدراسات السابقة والتقارير الوطنية والدولية..... 14
8. الممارسات الفضلى في التقييمات الصفية 15

ثانيًا: التقييمات الوطنية Large-Scale National Assessment 15

9. ثالثًا: التقييمات الدولية International Large-Scale Assessment..... 25
10. رابعًا: امتحانات الثانوية العامة (التوجيهي) (General Secondary Examinations (Tawjihi)..... 34
11. الجزء الثاني: الإطار الوطني المقترح لتقييمات الطلبة في الأردن..... 49
12. اختبار الاستعداد المدرسي (الصف الأول) 56
13. التقييمات الصفية (جميع الصفوف من الصف الأول إلى الثاني عشر)..... 60
14. الخلاصة والمقترحات بخصوص التقييمات الصفية للطلبة في الأردن:..... 75
15. التقييمات الوطنية (صفوف الثالث والسادس والتاسع) 77
16. أولاً: التقييم الوطني للصف الثالث الأساسي 83
17. ثانيًا: التقييم الوطني للصف السادس الأساسي 85
18. ثالثًا: التقييم الوطني للصف التاسع الأساسي 87
19. الخلاصة والتوصيات بخصوص التقييمات الوطنية 88
20. التقييمات الدولية (صفوف الرابع والثامن والعاشر) 90
21. أولًا: الدراسة الدولية للتقدم في القرائية (PIRLS)..... 91
22. ثانيًا: دراسة التوجهات الدولية في الرياضات والعلوم _ الصف الثامن..... 93
23. **أدوات جمع البيانات** 95
24. ثالثا: البرنامج الدولي لتقييم الطلبة (PISA) في عمر 15 سنة. 97
25. الغرض من المشاركة في التقييمات الدولية 99
26. الأدوار والمسؤوليات في التقييمات الدولية..... 100
27. آليات كتابة التقارير والإبلاغ عن النتائج 101

28. التوصيات بخصوص التقييمات الدولية..... 103
29. امتحان الثانوية العامة (التوجيهي في الصفين الحادي عشر والثاني عشر)..... 104
30. الأدوار والمسؤوليات 111
31. إنشاء بنوك الأسئلة 112
32. معايرة فقرات الاختبار باستخدام نظرية استجابة الفقرة IRT 122
33. معادلة نماذج الاختبار 126
34. المقياس المرجعي للدرجات 129
35. إعداد الامتحانات 131
36. جدول مواصفات الامتحان Table of Examination Specification..... 133
37. جمع المعلومات وتحليلها وإصدار النتيجة 134
38. الخلاصة والتوصيات بخصوص امتحانات الثانوية العامة..... 135
39. الجزء الثالث: بناء القدرات وتوجيه السياسات: 136
40. بناء القدرات: 136
41. توجيه السياسات 139
42. المراجع: 141

الإطار الوطني لتقييم الطلبة في الأردن 2024

National Student Assessment Framework in Jordan 2024

يحتل نظام التعليم في الأردن مرتبة متقدمة في العالم العربي، وهناك إصرار على الاستمرار في مسيرة تطويره بكل عناصره ومحاوره لبناء نظام ذو جودة عالية، لأنه أساس الاستثمار في رأس المال البشري. وأكدت الورقة النقاشية السابعة لجلالة الملك عبد الله الثاني على نوعية التعليم الذي يساهم في إعداد أجيال قادرة على الإبداع والتفوق والمنافسة في المجالات العلمية المختلفة. ويؤدي التقييم التربوي والنفسي الجيد دورًا مهمًا في تطوير العملية التعليمية التعلمية، فهو الوسيلة والمعيار للحكم على فاعليتها، ويتطلب المراجعة والتحسين بشكل مستمر. ولتطويره آثار إيجابية على الطالب والأسرة والمجتمع، فالتطوير يقلل نسبيًا من المخاوف والأعباء الكبيرة لدى الطلبة وأسرهم، ويزيد من قيم العدالة والسعي نحو التميز والإبداع والتفكير الخلاق.

ولدى وزارة التربية رغبة أكيدة ودافع قوي لتطوير تقييمات الطلبة في المراحل الدراسية كافَّة انسجامًا مع الرؤى الملكية السامية لتطوير النظام التعليمي ليتلاءم مع المستجدات العالمية وتطور المعرفة ووسائل التكنولوجيا والاتصالات. لذلك تم تشكيل فريق عمل بقرار من معالي وزير التربية والتعليم بتاريخ 26 شوال 1445 الموافق 6 تموز 2024، لمشروع وضع إطار وطني لتقييم الطلبة في الأردن يغطي التقييمات الصفية والامتحانات العامة والتقييمات الوطنية والدولية. أوكل إليه المهام الآتية:

- دراسة الوضع الراهن لتقييمات الطلبة.
- عمل دراسة مفصلة للامتحانات العامة والتقييمات الوطنية والدولية والتقييمات الصفية وخارطة الطريق للتحسين.
- بناء القدرات في مجال التخطيط للتقييم الوطني واستخدام النتائج في وضع سياسات وزارة التربية والتعليم.

وقد عمل الفريق من خلال المنهجية الآتية:

1- الاطلاع على الوضع الراهن لتقييمات الطلبة، ومعرفة المعلومات التفصيلية حولها والسياسات والتعليمات التي تحكمها، وتحليلها ودراستها والاستفادة منها والبناء عليها.

2- مراجعة التقارير والدراسات السابقة المحلية والدولية التي تناولت تقييمات الطلبة بالبحث والتحليل سواء بصورة جزئية أو كلية، وتحديد التوجهات والتوصيات التطويرية المقترحة لتحسين الممارسات التقييمية.

3- الاطلاع على أفضل الممارسات الدولية فيما يتعلق بكل محور من محاور تقييم الطلبة وكيفية استخدام النتائج في تعزيز التعليم وتحسين التعلم.

4- توزيع استبانات وإجراء مقابلات لاستطلاع رأي مجموعة واسعة من المسؤولين والمهتمين وأصحاب المصلحة من معلمين ومشرفين ومديري إدارات ومدارس ومشاريع دولية داعمة (أساس، والدعم الفني) وأكاديمية الملكة رانيا لتدريب المعلمين وبعض عمداء كليات التربية وغيرهم.

5- تقديم المقترحات والأوراق من قبل أعضاء الفريق بما ينسجم مع المحاور والقضايا والمواضيع التي تم الاتفاق على مناقشتها وتحليلها.

6- الاستفادة من المعرفة النظرية والخبرة العملية للخبير الدولي الدكتور ادواردو كاسكالر، وما توصل إليه من استنتاجات وتوصيات في ضوء المعلومات التي جمعها في زيارتيه للأردن.

وتهتم الورقة الحالية بتطوير إطار وطني لتقييمات الطلبة على مستوى النظام التعليمي الأردني. وتتكون من ثلاثة أجزاء. يتناول الجزء الأول الوضع الراهن لتقييمات الطلبة في النظام التعليمي الأردني، ويتناول الجزء الثاني الإطار الوطني المقترح لتقييمات الطلبة، ويتناول الجزء الثالث تحديد الحاجات اللازمة لبناء القدرات، وتوجيه السياسات التي يجب تنفيذها من أجل تحقيق أفضل النتائج.

الجزء الأول: الوضع الراهن لتقييمات الطلبة في النظام التعليمي الأردني

تُراجع الأنظمة التعليمية الحيّة سياساتها باستمرار، وتعمل على تطوير خططها وبرامجها لتلبية احتياجات المجتمع المتجددة من جانب، ولتكون قادرة على التنافس عالميًا من جانب آخر. ويعد التقييم أحد الجوانب المهمة في عمليات التطوير التربوي التي تتطلب المراجعة والتحسين بشكل مستمر، ويوفر التقييم معيارًا صادقًا وموضوعيًا للحكم على فاعلية النظام التعليمي في أداء دوره، وللحكم على قدرات الطلبة، وكفايات المعلمين ومديري المدارس والمشرفين التربويين.

ونظرًا لأهمية التقييم في توجيه العملية التعليمية التعلمية نحو تحقيق أهدافها، اهتمت وزارة التربية والتعليم في الأردن ببناء إطار عام للتقييم منذ ثمانينات القرن الماضي في مشروع التطوير التربوي الأول عام 1987 بقرار مجلس التربية رقم 1988/88/34 الذي أقر الخطوط العريضة للمناهج التي تضمنت توجيهات عامة في أساليب التقييم، وبذلك ظهر أول إطار عام للتقييم وفق التوجهات التي ترى أن الإطار العام للتقييم جزء لا يتجزأ من وثيقة المنهاج.

ومع انطلاقة برنامج تطوير التعليم نحو الاقتصاد المعرفي (ERfKE, 2003–2014)، تم تبني منحى جديد في بناء وثيقة خاصة بالتقييم مستقلة عن وثيقة المنهاج، فقد اهتمت وزارة التربية والتعليم في الأردن بتطوير وتحديث استراتيجيات التقييم لأداء طلبتها بتبني التقييم الواقعي (Authentic Assessment). وتم ترتيب رحلات معرفية (Study Tours) لأعضاء إدارة الامتحانات والاختبارات إلى أشهر بيوت الخبرة في الولايات المتحدة الأمريكية (ETS) واستراليا (SEART) وبريطانيا (PERSON) وكندا؛ للاطلاع على الممارسات الفضلى لديهم، وتعاقدت وزارة التربية مع نخبة من أساتذة القياس والتقويم في الجامعات الأردنية والخبراء الأجانب، ونتج عن تلك الجهود إعداد "الإطار النظري لاستراتيجيات التقويم وأدواته–2004" الذي اعتبر إطارًا عامًّا للتقويم.

ومن أجل اقتراح إطار وطني لتقييمات الطلبة في الأردن لا بد من إجراء مسح للتقييمات الحالية، لتوفير المعلومات اللازمة للإجابة عن الأسئلة الآتية:

1. ما تقييمات الطلبة المطبقة حاليًا في نظام التعليم الأردني، وما غرض كل منها، وماذا تقيس؟
2. ما الترتيبات المؤسسية واللوجستية/التقنية ومستويات التمويل ومصادره لكل منها؟
3. ما آليات كتابة التقارير والإبلاغ عن النتائج في كل منها؟
4. ما الذي تقوله الدراسات السابقة والتقارير الوطنية والدولية –إن وجدت– عن كل منها؟
5. ماذا يقول الأدب التربوي والممارسات الدولية الفضلى فيما يتعلق بكل منها؟

أولًا: التقييمات الصفية في الأردن(Classroom Assessments in Jordan)

1. واقع التقييمات الصفية في الأردن

أ. التقييمات الصفية في المرحلتين الأساسية والثانوية (المسار الأكاديمي)

بدأ إصدار التشريعات الناظمة للتقييمات الصفية للطلبة في الأردن[1] عام 1921م مع انطلاقة التعليم الرسمي وتوسعه على حساب التعليم غير الرسمي (الكتاتيب)، وكان التقييم الصفي شفويًا أو تحريريًا باستعمال القلم والورقة(Pencil and Paper)، والأسئلة إنشائية، وقد تطورت التقييمات الصفية خلال النصف الثاني من القرن العشرين حتى انعقاد مؤتمر التطوير التربوي الأول[2] عام (1987)، وكانت النهاية العظمى لعلامة المبحث (100) موزعة على أربعة مكونات: الامتحان الأول 20%، والامتحان الثاني 20%، والامتحان النهائي للفصل الدراسي 40%، والمشاركة الصفية 20%. وتقدر علامة المشاركة من قبل المعلم بناءً على مجموعة من الملاحظات والتقييمات المستمرة للطالب أثناء الحصص الصفية، وأدائه في الواجبات المنزلية. وغالبًا ما كانت تتأثر علامة المشاركة بعوامل لا علاقة لها بتحقيق نتاجات التعلم كالحضور والغياب والنظافة والترتيب وغيرها، وتعاني من التضخم في القيمة وعدم الاتساق مع التقييمات الأخرى، وأشارت الأوراق المقدمة في المؤتمر إلى وجود عدة مشكلات في الامتحانات الصفية تتمثل بالتركيز على الحفظ والاستظهار، وعدم الاهتمام بالقدرات العقلية العليا من استقصاء وحل المشكلات، ولا توفر تغطية جيدة لمحتوى المبحث في الكتاب المدرسي، أي تفتقر لصدق المحتوى، ويعود ذلك إلى عدم امتلاك المعلمين لمهارات بناء الاختبارات التحصيلية الجيدة، وعدم معرفتهم بقواعد صياغة الأسئلة، والاعتماد بصورة مكثفة على الأسئلة الإنشائية.

وتم التركيز في مشروع التطوير التربوي على التنمية المهنية للمعلم في مجال إعداد الاختبارات التحصيلية، وعقدت ورشًا متخصصةً بإعداد الأسئلة الصفية بأنواعها المختلفة، ومن ضمنها الأسئلة التي تقيس مهارات التفكير العليا وحل المشكلات، وتم وضع خارطة طريق لتضمينها في امتحان الثانوية العامة[3] والامتحانات المدرسية بشكل متدرج بنسبة تصاعدية. وتم إعادة النظر في طرق تقدير علامة المشاركة، وتم الاعتماد عليها لتقييم المجال الانفعالي بما فيه من قيم واتجاهات وميول وسلوك اجتماعي.

واعتمدت التقييمات الصفية حتى مطلع القرن الحادي والعشرين على الاختبارات التي يعطيها المعلمون والمعلمات لطلبتهم بغرض الحصول على معلومات عن تحصيلهم لتقديمها لأولياء الأمور وغيرهم من المعنيين،

---

[1] الطراونة والمحاسنة والفهيقي (1934) (تقرير أحوال التعليم في إمارة شرق الأردن،

[2] (الوهر، محمود طاهر، العناتي، جهاد محمد (وثائق وقائع جلسات مؤتمر التطوير التربوي، تسجيلات صوتية لوقائع المؤتمر

[3] استراتيجيات التقويم وأدواته: الإطار النظري، إدارة الامتحانات والاختبارات، 2004.

وكانت تقيس مهارات ومفاهيم بسيطة يتم التعبير عنها بأرقام لا تقدم معلومات ذات قيمة عن تعلم الطالب وما حققه من نتاجات التعلم، وبالرغم من أن الطلبة هم محور عملية التقييم، فإنهم لم يشاركوا في تقييم أنفسهم.

وقد اهتم "برنامج تطوير التعليم نحو الاقتصاد المعرفي، (ERfKE I &II, 2003–2014) بالتقييمات الصفية، وتمثل ذلك بتبني التقييم الواقعي (Authentic Assessment) وإصدار "الإطار النظري لاستراتيجيات التقويم وأدواته–2004"[4]، وإعداد أدلة إجرائية ونشرات توجيهية متخصصة للمعلمين، وتدريب المشرفين والمعلمين على استخدام تلك الأدلة.

وتقوم التقييمات الصفية في الأردن على افتراض مفاده أنه كلما استطعت أن تعرف أكثر عن تعلم الطلبة وكيفيته تكون أكثر قدرة على تصميم الأنشطة الصفية لتوجيه عمليتي التعليم والتعلم[4]. فالتقييمات الصفية ليست مجرد إصدار حكم على مستوى الطالب وتصنيفه في فئة معينة، وإنما تتعداه إلى تقويم جوانب شخصيته، ومعرفة ما تعلمه، وما يمكنه القيام به من مهام، والعمليات التي استخدمها في تعلمه، واتجاهاته ومشاعره ودافعيته نحو التعلم، ومساعدته على فهم موقعه في عملية التعلم، وتوجيهه بشكل يمكنه من تحديد جوانب القوة لديه وإثرائها، وجوانب الضعف للعمل على معالجتها. ويتمركز نموذج التقييمات الصفية المستخدم في الأردن حول الطالب كما هو مبين في الشكل (1).

الشكل (1): نموذج التقييم الصفي المتمركز حول الطالب.

وتتعدد أنواع التقييمات الصفية المستخدمة في الأردن، فمنها التقييمات القبلية التي تستخدم في بداية العام الدراسي أو قبل البدء بتدريس وحدة معينة بهدف الكشف عن المهارات التي يجب على الطلبة امتلاكها قبل البدء بعملية التدريس. ومنها التقييمات التكوينية التي تستخدم أثناء التدريس بهدف تزويد المعلم والمتعلم

[4] نشرة إرشادية حول التقويم الصفي للعام الدراسي 2022/2023.

بتغذية راجعة من أجل تحسين العملية التعليمية التعلمية ومعرفة مدى تقدم التلميذ. وللتأكد من سلامة سير العملية التدريسية لابد من أن تكون التقييمات التكوينية دورية ومستمرة، وأخيرًا التقييمات الختامية التي تستخدم بعد الانتهاء من تدريس الوحدة أو الفصل الدراسي بجميع وحداته بهدف تقييم التحصيل وإعطاء العلامات من أجل اتخاذ قرارات النجاح والرسوب أو التخرج أو منح الشهادات.

وتُعِدُّ إدارة الامتحانات والاختبارات أسس النجاح والإكمال والرسوب لمرحلتي التعليم الأساسي والتعليم الثانوي الشامل، وتقوم بمراجعتها وتحديثها بشكل مستمر لتوضيح أنواع التقييم المعمول بها في مختلف المراحل وأسسها وطريقة تنفيذها، بالإضافة إلى توضيح كيفية توزيع العلامات للمباحث في مختلف الصفوف والمراحل الدراسية[5]. وبموجب هذه الأسس تتوزع العلامات للمباحث النظرية للصفوف الأساسية الثلاثة الأولى على أربع فترات تقييمية، لكل منها (25%) من علامة المبحث، ويُقوَّم أداء الطالب فيها باستخدام استراتيجيات التقييم الواقعي المتنوعة، ومن خلال تقييم الأدلة من أعمال الطالب على تحقيقه لنتاجات التعلم والتي يحتفظ بها في ملف أعمال الطالب، وتتكون نتيجة الطالب في المبحث للفصل الدراسي الواحد من مجموع علاماته التي حصل عليها في الفترات التقييمية الأَربع. ويحتفظ المعلم بسجل جانبي تُسجَّل فيه تقييمات الطالب، وينقل مُعدّلها إلى سجل العلامات المدرسية، وتتضمن خطة التقييم في الصفوف الثلاثة الأولى أن يقوم المعلم في بداية الفصل الدراسي بتَشْخيص القدرات القرائية والحسابية لدى طلبته باستخدام الأداة التشخيصية العامة والأداة التشخيصية الخاصة، ويستخدم المعلم المعلومات الناتجة عن الأداتين في وضع خطة علاجية لمعالجة جوانب الضعف لدى طلبته بالتعاون مع إدارة المدرسة، وأولياء الأمور، والمشرف التربوي المتخصص، وفِرَق التطوير في المدرسة.

وتجدر الإشارة إلى أن الطلبة في الصفوف من الأوّل الأساسي ولغاية الثالث الأساسي يُرفَّعون تلقائيًّا إلى الصف الأعلى شريطة عدم تجاوز نسبة الغياب المسموح بها، ويُوَزَّع الطّلبة في هذه الصفوف إلى فئتين: الطّلبة الذين حقّقوا الحدّ الأدنى من المهارات القرائيّة والمهارات الحسابيّة ويُقدَّم لهم المنهاج الاعتيادي في الصفّ اللاحق، والطّلبة الذين لم يُحقّقوا الحدّ الأدنى من المهارات القرائيّة والمهارات الحسابيّة ويُقدَّم لهم المنهاج الاعتيادي في الصفّ اللاحق بالإضافة إلى أنشطة تعلُّميّة علاجيّة يُعدّها معلم الصف، وتُعتَمد من مدير المدرسة والمشرف التربوي وتُنفّذ أثناء الدّوام الرسمي.

[5] أسس النجاح والإكمال والرسوب في مرحلتي التعليم الأساسي والتعليم الثانوي الشامل (الأكاديمي والمهني) للعام 2023/2024.

ووفقًا لأسس النجاح والإكمال والرسوب توزع العلامات للمباحث النظرية للصفوف (4–12) على أربعة مكونات يتضمنها سجل العلامات الرسمي، هي الامتحان الأول 20%، والامتحان الثاني 20 %، والامتحان النهائي للفصل الدراسي 40%، وأعمال الفصل 20%، وتتضمن أعمال الفصل تقييمات مستمرة معتمدة على الأداء (Performance Based Assessment) مثل التقييمات الواقعية الشفوية والكتابية، والواجبات المنزلية، والمشاريع الفردية والجماعية ومراجعة الذات والملاحظة والتواصل. ويتم فيها استعمال أدوات تقييم غير اختبارية مثل قوائم الرصد وسلالم التقدير، وسلم التقدير اللفظي (Rubric)، وملف الإنجاز (Student Portfolio).

وتتناول الأسس الأحكام الخاصة بذوي الإعاقة، والأحكام الخاصة بالمستفيدين من غرف مصادر التعلم، وعلامة النجاح، والسماح بإعادة الامتحان. كما تصف التقديرات في الشهادة المدرسية وصفًا كميًّا ونوعيًّا على النحو الآتي:

1- التقدير ممتاز، ويرصد للعلامة ضمن الفئة 90– 100 ٪ ويعني أن الطالب يمتلك المعارف والمهارات المطلوبة، ويحقق نتاجات التعلّم بشكل يفوق معايير المستوى التعليمي المُحدَّد.
2- التقدير جيّد جدًا، ويرصد للعلامة ضمن الفئة 80 – 89 ٪ ويعني أن الطالب يمتلك المعارف والمهارات المطلوبة، ويحقق نتاجات التعلّم ضمن معايير المستوى التعليمي المُحدَّد.
3- التقدير جيّد، ويرصد للعلامة ضمن الفئة 70 – 79 ٪ ويعني أن الطالب يمتلك معظم المعارف والمهارات المطلوبة، ويحقق نتاجات التعلّم ضمن معايير المستوى التعليمي المُحدَّد.
4- التقدير متوسط، ويرصد للعلامة ضمن الفئة 60 – 69 ٪ ويعني أن الطالب يمتلك بعض المعارف والمهارات المطلوبة، ويقترب من تحقيق نتاجات التعلم ضمن معايير المستوى التعليمي المُحدَّد.
5- التقدير مقبول، ويرصد للعلامة ضمن الفئة 50 – 59٪ ويعني أن الطالب يمتلك بعض المعارف والمهارات المطلوبة بمستوى محدود، ويحقق الحد الأدنى المطلوب من نتاجات التعلم ضمن معايير المستوى التعليمي المُحدَّد.
6- التقدير مُقصِّر، ويرصد للعلامة 49 ٪ فما دون، ويعني أن الطالب لا يمتلك الحدّ الأدنى المطلوب من المعارف والمهارات المطلوبة، ويتطلَّب خطة علاجيّة لإعادة توجيه تعلّمه في المسار الصّحيح.

ويكون الطالب ناجحًا في أيّ صفّ من الرّابع الأساسي وحتّى الصف الثاني عشر، إذا حصل على (50٪) من النهاية العظمى لعلامة كلّ مبحث، ومُكمِلاً إذا قصّر في ثلاثة مباحث فأقل، وعليه أن يتقدَّم

إلى امتحان الإكمال في المبحث أو المباحث التي قَصَّر بها قبل بدء العام الدراسي الجديد، ويكون الطالب ناجحًا في صفه بعد أداء امتحان أو امتحانات الإكمال والنجاح فيها أيًّ كان معدله قبل ذلك. ويكون الطالب راسبًا في أيّ صف إذا قصَّر في أربعة مباحث على الأقل، أو إذا قصَّر في أيّ امتحان من امتحانات الإكمال التي تقدَّم لها. ولا يجوز أن تتجاوز نسبة الرسوب (15%) في الصف الواحد، مهما كان عدد الشُّعَب في المدرسة، على أن تحتسب بعد أداء اختبار الإكمال؛ ولا تُحتسَب حالاتُ الرسوب بسبب الغياب ضمنَ هذه النسبة. وتحديد نسبة معينة للراسبين يتعارض مع التقييمات القائمة على المعايير ولا يتوافق مع المناهج الأردنية المطورة حديثًا والقائمة على المعايير.

وتستخدم علامات الطلبة المدرسية في صفوف السابع والثامن والتاسع الأساسية، لأغراض توزيعهم على مساري التعليم الثانوي الشامل الأكاديمي والمهني وفروعهما. ويُحسَب المُعدَّل المُعتمَد للتوزيع بتخصيص نسبة 30٪، 30٪، 40٪ من المُعدَّل العام لعلامات الطالب في صفوف السابع والثامن والتاسع الأساسية على الترتيب، ويؤخذ على ذلك عدم مراعاة رغبة الطلبة في اختيار المسار الذي يتناسب مع قدراتهم واهتماماتهم وميولهم.

ب. التقييم في برنامج بيتك(Business & Technology Education Council (BTEC

تم تطوير برامج ومؤهلات BTEC من قبل شركة التعلم الرائدة بيرسون Pearson التي تعمل في بلدان كثيرة من جميع أنحاء العالم حيث أن هذه البرامج والمؤهلات معترف بها دوليًا من قبل الحكومات وأصحاب العمل. وتم تصميمها لتعزيز المناهج الدراسية وإعداد المتعلمين لها لعالم العمل المتغير باستمرار، وتركز برامج BTEC على المتعلم، وتعتمد بنى وهياكل مرنة قائمة على الوحدات الدراسية، والتقييمات فيها قائمة بصورة أساسية على المشاريع والتعيينات. وتركز على التطوير الشامل للمهارات العملية والشخصية والتفكيرية المطلوبة للنجاح في العمل والتعليم العالي.

وفي المسار المهني في الأردن، يتم تطبيق تخصصات برنامج BTEC من المستوى الثاني على طلبة الصف العاشر، وتخصصات المستوى الثالث على طلبة الصفين الحادي عشر والثاني عشر. ويقوم كل مستوى على عدد كبير من الوحدات الدراسية، لكل وحدة دراسية عدد من ساعات التعلم الموجه Guided Learning Hours هو (30 or 60 GLH) في المستوى الثاني والعدد (60 or 90 or 120 GLH) في المستوى الثالث. وللحصول على مؤهل الدبلوم في المستوى الثاني يجب أن ينجح الطالب في 7 وحدات دراسية إلزامية لها 240 ساعة ووحدات أخرى اختيارية لها 240 ساعة، أي بمجموع كلي يساوي 480 ساعة تعلم موجه.

أما للحصول على مؤهل الدبلوم في المستوى الثالث، فيجب النجاح في 5 وحدات إلزامية لها 480 ساعة ووحدات أخرى اختيارية لها 240 ساعة، أي بمجموع كلي هو 720 ساعة.

ويصمم التقييم ليناسب غرض وهدف المؤهل، ويشمل مجموعة من أنواع وأساليب التقييم المناسبة للمؤهلات المهنية. وجميع التقييمات في الوحدات الدراسية داخلية وتعتمد على التعيينات، أي أن المعلم يضع التعيينات ويصححها، باستثناء بعض الوحدات الإلزامية حيث تضع بيرسون التعيينات ويقوم المعلم بتصحيحها. ويتم تقييم الطالب في الوحدة الدراسية وفقًا لثلاثة معايير في الأداء: معيار الاجتياز Pass (P)، ومعيار الجدارة Merit (M)، ومعيار التميز Distinction (D). ويعد الطالب الذي لا يحقق أيًّا من معايير النجاح الثلاثة بأنه غير مصنف Unclassified (U). ويتم تدقيق التقييمات للطلبة داخليًا للتأكد من دقتها. وتتصف التقييمات الداخلية بأنها مرنة وعادلة وتحقق المعايير الدولية، وهناك التزام بمبادئ التقييم الداخلي التي تتمثل بالتقييم من خلال التعيينات، وبتطبيق المعايير القائمة على الوحدات في اتخاذ قرارات التقييم، وبوجود فريق فعال للتقييم، والتنظيم الفعال، وإعداد المتعلم وتأهيله بصورة جيدة للتقييم. وتتاح الفرصة للطالب لإعادة تقديم أدلة محسنة عن تعلمه. ويمكن للطالب الذي لا يصل إلى مستوى النجاح إعادة التقييم.

ولإعطاء وصف كمي أدق لإنجاز الطالب يتم إعطاء نقاط للمعايير، ويبين الجدول 1 النقاط المعتمدة للمعايير في المستوى الثاني والمستوى الثالث. وتكون القيمة العظمى لعدد النقاط في المستوى الثاني 384 بينما تكون في المستوى الثالث 192.

الجدول 1: عدد النقاط للمعايير في المستويين 2،3 بدلالة عدد ساعات الوحدة الدراسية

	LEVEL 2		LEVEL 3		
Standard	30 GLH	60 GLH	60 GLH	90 GLH	120 GLH
Unclassified U	0	0	0	0	0
Pass P	12	24	6	9	12
Merit M	18	36	10	15	20
Distinction D	24	48	16	24	32

2. الترتيبات المؤسسية واللوجستية/التقنية والتمويل للتقييمات الصفية في الأردن

تتشابه الترتيبات المدرسية للتقييمات الصفية المطبقة على الطلبة في نظام التعليم الأردني عبر المدارس، حيث إن أسس النجاح والإكمال والرسوب تحدد خطة التقويم للعام الدراسي ومواعيد الامتحانات وتسليم النتائج واعتمادها، وتتولى كل مدرسة بشكل مستقل إعداد جدول الامتحانات وتوزيعه على الطلبة وتكليف المعلمين بإعداد الامتحانات وإجاباتها النموذجية وسحب العدد الكافٍ من نسخ الامتحان وتغليفها وحفظها في مكان تخزين آمن في المدرسة قبل بدء الامتحانات بوقت كافي، وتحدد إدارة المدرسة أسماء أعضاء لجان المراقبة وتعد جداول المراقبة وتتابع إجراءات تطبيق الاختبار وتصحيحه ورصد نتائجه. بالإضافة لاعتماد المدارس على ذاتها في التقييمات الصفية حيث إن هناك موارد ومواد متاحة على مستوى النظام لدعم المعلمين في إجراء التقييمات الصفية، بما في ذلك التدريب أثناء الخدمة الذي يقدمه المشرفون التربويون في مديريات التربية والتعليم. والتقييمات الصفية من مسؤولية المعلم وهي من ضمن مهام عمله.

وقد وفّر برنامج الدعم الفني[6] Technical Assistance Program TAP التدريب للمعلمين على ملاحظة أداء الطّلبة باستخدام التّقييم المعتمد على الأداء. ويوفر برنامج ASAS[7] الدعم للمعلمين في تنفيذ كل من الاختبارات التشخيصية والشهرية على مستوى كل صف من الصفوف الثلاثة الأولى والروضة Kg2. وينخرط هذا البرنامج بفعالية مع أكاديمية الملكة رانيا لتدريب المعلمين وكليات التربية في الجامعات الأردنية في دعم برامج إعداد المعلمين للصفوف المبكرة وتطوير خططها.

3. آليات كتابة التقارير والإبلاغ عن النتائج في التقييمات الصفية في الأردن

تؤكد التعليمات في أسس النجاح والإكمال والرسوب للصفوف (4–12) أن على المعلم أن يسجل تقييماته الصفية للطلبة في دفتر جانبي مع الملاحظات حول تلك التقييمات، وعليه نقل التقييمات إلى سجل العلامات الذي يعد وثيقة رسمية يجب المحافظة عليها، ويتم رصد نتائج الطلبة في جداول خاصة يتم اعتمادها من مديرية التربية والتعليم، وبعد اعتماد النتائج يتم إدخالها إلى نظام إلكتروني معتمد لدى وزارة التربية والتعليم، كما يتم تسجيل علامات الطلبة في الشهادات المدرسية التي ترسل إلى أولياء أمور الطلبة، وتستخدم النتائج لغرض ترفيع الطلبة إلى الصف اللاحق.

[6] Technical Assistance Program, USAID.

[7] USAID Early Grade Education Activity (ASAS)

4. التقييمات الصفية في الدراسات السابقة والتقارير الوطنية والدولية

أظهرت الدراسات والتقارير الدورية[8] التي استهدفت التقييمات الصفية المطبقة على الطلبة في نظام التعليم في الأردن أن الاختبارات من إعداد المعلم تتضمن العديد من نقاط الضعف حيث لا تلتزم بدرجة كبيرة بخصائص السؤال الجيد، وتركز على قياس المستويات المعرفية الدنيا (الحفظ والتذكر)، وتتضمن أسئلة صياغتها اللغوية ضعيفة وتتضمن أخطاء نحوية وإملائية، ولا تستند في بنائها إلى جداول مواصفات أو مخطط اختبار، كما يستخدم المعلمون أسئلة سنوات سابقة متداولة بين أيدي طلبتهم، وهذا يسهم في تضخم الدرجات. وأشارت نتائج الدراسات للاختبارات من إعداد المعلمين أنها تتسم بدرجة ضعيفة إلى متوسطة في مؤشرات صدقها وثباتها. ومن جهة أخرى أظهرت التقارير والدراسات ضعف استخدام معلمي المباحث المختلفة لاستراتيجيات التقويم الواقعي وأدواته بسبب ضعف كفاياتهم المهنية في هذا المجال، ولا يستندون إلى معايير موضوعية في تقدير العلامات لأعمال الفصل.

وأشار تقرير البنك الدولي SABER[9] أنه بالرغم من توافر العديد من الوثائق على مستوى النظام، وإرشادات تنفيذ أنشطة التقييم الصفي، والموارد المتاحة لدعم المعلمين في إجراء تلك الأنشطة، فإن ممارسات التقييم الصفي من قبل المعلمين ضعيفة وهي في مستوى "ناشئ". وأكد التقرير الإقليمي حول خريطة التقييم في الدول العربية[10] على هذا الضعف.

وبالرغم من الجهود الكبيرة، التي تبذلها إدارة الامتحانات والاختبارات في التدريب على بناء الاختبارات التحصيلية واستراتيجيات التقييم الواقعي إلا أنه ما زال "هناك ضعف في كفايات المعلمين في إعداد وتنفيذ وتحليل نتائج التقييمات الصفية التي تجري في المدارس". فقد أشارت الاستطلاعات[11] التي تم القيام بها لأغراض الدراسة الحالية أن التقييمات الصفية شكلية وتقليدية وتركز على الحفظ وغير تفاعلية وتقتصر على الأغلب على اختبارات الورقة والقلم. وهناك تباين كبير في الاختبارات التي تستهدف المحتوى نفسه عبر المدارس، والمؤسف أنه لا تتوافر معلومات عن درجة صدق وثبات وموضوعية وعدالة هذه التقييمات، كما أشارت إلى تدني مخرجات التدريب بسبب عدم وجود تنسيق بين الجهات التي تعنى في تدريب المعلمين وإدارة الامتحانات والاختبارات من جهة، وعدم إشراك المتخصصين في القياس والتقويم في التدريب على التقييمات الصفية من جهة أخرى. وهذا يحتم ضرورة مراجعة فرص التدريب الحالية في برامج إعداد المعلمين

[8] (2010-2024) تقارير إدارة الامتحانات والاختبارات لنتائج تحليل عينات من الامتحانات المدرسية

[9] SABER Country Report, 2014, Jordan

[10] أعد التقرير عدنان الأمين بتكليف من مكتب اليونسكو الإقليمي للتربية في الدول العربية والمنظمة العربية للتربية والثقافة والعلوم

[11] لأغراض الدراسة الحالية تم استطلاع آراء 338 مشرفًا ومشرفة و498 مديرًا ومديرة و2927 معلمًا ومعلمة وغيرهم

قبل الخدمة وفي برامج التدريب المهني أثناء الخدمة، وتحسين محتواها وجودتها، وتوسيع نطاقها ومدتها، مع توفير نماذج للتقييم والترويج لها وتشجيع المعلمين على استخدامها.

5. الممارسات الفضلى في التقييمات الصفية

تركز الممارسات الفضلى في التقييمات الصفية كما في برنامج International Baccalaureate (IB)، وبرنامج BTEC على منهجية التقييم المستمر والمتنوع من خلال مجموعة من الاستراتيجيات وأدوات التقييم التي يستخدمها المعلمون والطلبة، وتهتم بتحليل بيانات التقييم لإثراء التعليم والتعلم. وتتاح الفرص للطلبة في التقييمات الصفية لأن يوظفوا تعلمهم في تقييم فهمهم وأعمالهم وأعمال الآخرين. وتقوم التقييمات الصفية على توظيف استراتيجيات التقييم الواقعي وأدواته، ومن ضمنها التقييم المعتمد على الأداء والمشاريع واستخدام بطاقات الملاحظة ودفتر تسجيل الملاحظات.

وتقوم الممارسات الفضلى في التقييمات الصفية على الموازنة بين التقييمات التكوينية والتقييمات الختامية، وبالرغم من تعدد أشكال التقييمات التكوينية، فإن هدفها ليس إعطاء درجات للطلبة بل توفير التغذية الراجعة الضرورية لتوجيه التعليم وتحسين التعلم. وهي مستمرة وملازمة للتدريس، وإنها تهتم بالتقييم من أجل التعلم (Assessment for learning) وتهتم بالتقييم كتعلم (Assessment as learning) وهي تعلي من شأن الطلبة وتزيد ثقتهم بأنفسهم من خلال تقييم الطالب لأدائه ولأداء أقرانه. ويمكن الاعتماد على التقييمات الختامية لغرض إعطاء درجات للطلبة.

ثانيًا: التقييمات الوطنية Large-Scale National Assessment

تستهدف التقييمات الوطنية واسعه النطاق، الطلبة في صفوف الرابع والثامن والعاشر الأساسية للوقوف على مدى اكتساب الطلبة للمعارف والمهارات، ومقارنتها بما يجب أن يمتلكه الطلبة (تقييم محكي المرجع) وتستخدم نتائجها للمقارنة بين المدارس والمناطق (تقييم معياري المرجع) ولتحديد المدارس ذات الأولوية بالرعاية كالمدارس الصفرية التي لم ينجح فيها أحد أو متدنية الأداء.

1. أغراض الاختبارات الوطنية الراهنة وما تقيسه.

1. الاختبار الوطني لضبط نوعية التعليم National Test for Controlling the quality of Instruction

هو اختبار سنوي تعده وتنفذه مديرية الامتحانات والاختبارات في وزارة التربية والتعليم منذ عام 2000 لقياس مدى امتلاك الطلبة في الصفوف 4، 8، 10 لمهارات التعلّم الأساسيّة في عدة مباحث دراسية، ويهدف إلى توفير معلومات عن مستويات الطلبة لاستخدامها كخط قاعدي (Base Line) لدراسة أثر برامج التطوير

التربوي في إحداث النمو المنشود. وقد شهد تطويرًا في إجراءاته مع بدايات برنامج التطوير التربوي نحو اقتصاد المعرفة عام 2006م من حيث:

- تغير المباحث التي يستهدفها الاختبار حيث اقتصر على أربعة مباحث أساسية هي: اللغة العربية، واللغة الإنجليزية، والرياضيات، والعلوم.
- زيادة حجم العيّنة من عينة صغيرة 5% إلى عينة متوسطة 15%، إلى المسح الشامل.
- الانتقال من امتحان مستقل لكل مبحث إلى أوراق امتحانيّة تضم مبحثين، إلى تخصيص مادة عشوائيًا للطالب من المواد الأربع.
- تقديم الاختبار إلكترونيًّا جنبًا إلى جنب مع التقديم الورقي وبزيادة مضطردة للنسخ الإلكترونية عبر السنوات.
- استهداف الصف (4 أو8 أو10) كل ثلاثة أعوام عوضًا عن الاستهداف السنوي.
- الانتقال من استخدام موضوعات المحتوى وأهداف التعلم للمبحث كإطار مرجعي لتحديد مجال محتوى الاختبار (Test Domain) إلى استخدام محاور المحتوى ونتاجات التعلم، وإلى استخدام معايير ومؤشرات الأداء.
- التحول في غرض الاختبار من قياس التحصيل للصف المستهدف إلى قياس المهارات والكفايات الأساسية للمرحلة دراسية (الصف الرابع للحلقة الأولى في التعليم الأساسي، والصف الثامن للحلقة الثانية والصف العاشر للحلقة الثالثة).

وتجدر الإشارة إلى أنه عند بدء جلسة الامتحان الوطني لضبط جودة التعليم، يتم إعطاء طلبة الصف المستهدف، وعلى التناوب، امتحانًا من الامتحانات الأربعة (اللغة العربية، أو الرياضيات، أو اللغة الإنجليزية، أو العلوم). أي أن الطالب لا يعرف الامتحان الذي سيتقدم له قبل جلسة الامتحان. وهذا يتعارض مع أخلاقيات التقييم وعدالته، ومن المتوقع أن يؤثر سلبًا على أداء الطلبة، وعلى خصائص الاختبار وفقراته وقد تكون النتائج مضللة وغير صادقة.

ويهدف الاختبار الوطني لضبط نوعية التعليم إلى:

1. تزويد صانع القرار (مجالس ولجان التربية) بمعلومات عن جودة التعليم تساعد في اتخاذ قرارات التطوير المناسبة.

2. تزويد إدارة المناهج والكتب المدرسية وإدارة التدريب والإشراف التربوي بتغذية راجعة عن مدى إتقان الطلبة لمحاور ومهارات التعلم الأساسية، وجوانب الضعف في أداء الطلبة مما يساعد في إعادة النظر في محتوى الكتب المقررة وبرامج التنمية المهنية.
3. مساعدة المعلمين في تقييم الطلبة ومتابعة وتحديد مسار تقدمهم، وتزويدهم بمعلومات عن جوانب القوة والضعف في أداء طلبتهم تساعدهم في بناء الخطط لبرامج علاجية وإثرائية.
4. تزويد الطلبة بتغذية راجعة عن مدى إتقانهم لمحاور ومهارات التعلم الأساسية مما يوفر لهم معرفة موضوعية عن قدراتهم تساعدهم في تحديد مسارهم وخياراتهم المستقبلية.
5. مقارنة أداء الطلبة بالمرجعية/خط الأساس (Base Line) واستخدام ذلك كمؤشر على جودة ونوعية التعليم.

2. اختبار ضبط جودة التعلم (Lot Quality Assurance Sampling (LQAS

هو اختبار سنوي شرعت إدارة الامتحانات والاختبارات في وزارة التربية والتعليم عام 2021 بإعداده وتنفيذه بهدف ضبط جودة تعلم مهارات القراءة والحساب لطلبة الصفين الثاني والثالث باستخدام تقنية LQAS التي تسمح بتطبيقه على 19 طالبًا وطالبة (10 من الصف الثاني و9 من الصف الثالث أو العكس) في 19 مدرسة. ويطبق الاختبار ورقيًّا (paper and pencil) بصورة فردية في المهارات القرائية (Individual Testing) وبصورة جماعية (Group Testing) في الحساب. وهو أداة سريعة وغير مكلفة لجمع معلومات حول أداء الطلبة في الصفين الثاني والثالث في القراءة والحساب على مستوى المديرية والمدرسة، ويكتب تقرير حوله على مستوى المدرسة، ويناقش مجلس أولياء الأمور نتائج المدرسة ويقترح خططًا علاجية وتحسينية، وتتم مساءلة المدارس في ضوء نتائجها في هذا الاختبار من قبل أولياء الأمور والمديرية والوزارة.

ويهدف اختبار ضبط جودة التعلم (LQAS) إلى:

1. ضبط جودة تعلم مهارات القراءة والحساب لطلبة الصفين الثاني والثالث.
2. مراجعة الخطط الدراسية.
3. تحسين استراتيجيات التعليم والتعلم.
4. مساءلة المعلم والمدرسة.
5. تحديد مسار تقدم طلبة الصفين الثاني والثالث في إتقان مهارات القراءة والحساب.

ويؤخذ على هذا الاختبار ضعف تمثيل العينات للمجتمعات المستهدفة مما قد يحد من تعميم نتائجه.

3. اختبار المسح الوطني للقرائية (EGRA) والحساب (EGMA)

يُعد تمكين الطفل في الصفوف الثلاثة الأولى من القراءة باستيعاب وإجراء العمليات الحسابية بفهم من أكثر العوامل المتنبئة بنجاحه في المراحل اللاحقة. وبغرض قياس كفايات الطلبة في القراءة والحساب مقارنة بالمعايير الدولية قامت إدارة الامتحانات والاختبارات في وزارة التربية والتعليم وبدعم من الوكالة الأمريكية للتنمية الدولية (USAID) وإدارة التنمية الدولية (DFID)، ومعهد (RTI) (Research Triangle Institute) بإعداد اختبارين وتطبيقهما ورقيًّا وبصورة فردية، الأول للقرائية (EGRA) والثاني للحساب (EGMA) على طلبة الصف الثاني الأساسي والصف الثالث الأساسي للعام 2023. ويعاد تطبيقهما كل عامين باستخدام أسلوب العينات العشوائية، حيث يتم اختيار 20 مدرسة من كل محافظة، ومن كل مدرسة يتم اختيار عينة من 10 أفراد من الصف الثاني و10 أفراد من الصف الثالث. وفي ضوء نتائج المدرسة يتم بناء الخطط العلاجية على المدى القريب والبعيد؛ وتنفذ الأنشطة العلاجية لإكساب الطلبة المهارات الأساسية في القراءة والحساب.

وتتألف الصورة الأصلية من اختبار مهارات القراءة للصفوف الأولى (EGRA) الذي يتم إجراؤه شفويًّا من خمس مهام فرعية هي: معرفة أصوات الحروف، والتعرف على الكلمات المخترعة، والطلاقة في القراءة الشفوية لنص متصل، والفهم القرائي، وفهم الاستماع، وتألفت الصورة الأصيلة من اختبار مهارات الحساب للصفوف الأولى (EGMA) من ست مهام فرعية هي: التعرف على الرقم، ومقارنة الكميات، والتعرف على الرقم المفقود، والجمع والطرح من المستوى الأول، والجمع والطرح من المستوى الثاني، والمسائل اللفظية.

وقد كانت البدايات الأولى لمشروع مسح القرائية والحساب في العام الدراسي 2011 – 2012 ، وتضمن المسح عينة ممثلة على الصعيد الوطني تتألف من (156) مدرسة حكومية في الأردن و(3120) طالبًا وطالبة من الصفين الثاني والثالث الأساسيين. وأظهرت نتائج المسح أن نسبة الإتقان لمهارات القرائية والحساب بعد الصف الثالث الأساسي 13% تقريبًا، مما شكل صدمة على المستويين الشعبي والرسمي. واعتبرت وزارة التربية هذه النتيجة خطًا قاعديًا تقارن فيه القياسات اللاحقة بالاختبارين نفسيهما للحكم على نجاعة وفاعلية التدخلات العلاجية في الفترة 2012–2014.

وقد تم إطلاق مبادرة RAMP (Early Grade Reading and Math Project) بهدف تحسين مهارات القراءة والحساب للطلبة من صف الروضة الثانية(KG2) وحتى الصف الثالث الأساسي، وتم الاستمرار بتقييم فاعلية التدخلات العلاجية للمبادرة باختباري (EGRA) و (EGMA) كل عامين أي في السنوات 2016،

2018 ،2020، ثم تم توطين البرنامج في إدارة الامتحانات والاختبارات في العامين 2021، 2023، وسينفذ في العام 2025، بشكل فردي إلكتروني للقرائية وبشكل جماعي ورقي للحساب، على طلبة الصف الثاني فقط بعد أن كان يستهدف طلبة الصفين الثاني والثالث.

ويهدف اختبار المسح الوطني للقرائية (EGRA) والحساب (EGMA) إلى:

1. توفير أدوات قياس موثوقة لقياس مهارات الطلبة في القراءة والحساب في الصفوف الثلاثة الأولى الأساسية وتشخيص نقاط القوة ونقاط الضعف فيها.
2. مقارنة أداء الطلبة بالمرجعية/ خط الأساس (Base Line) للأداء على مهارات القراءة والحساب وقياس مقدار النمو والتحسن.
3. قياس أثر برامج التدخل العلاجي الخاصة بمبادرة (RAMP) في مهارات القراءة والحساب للطلبة من صف الروضة الثانية (KG2) وحتى الصف الثالث الأساسي.
4. مساءلة المعلم والمدرسة.
5. تحسين استراتيجيات التعليم والتعلم.
6. تزويد الطلبة من صف الروضة الثانية (KG2) وحتى الصف الثالث الأساسي بتغذية راجعة عن مدى إتقانهم لمهارات القراءة والحساب مما يوفر لهم معرفة موضوعية عن مهارات التعلم الأساسية لديهم.

4. الاختبار التقييمي الوطني

هو اختبار سنوي تعده وتنفذه إدارة الامتحانات والاختبارات في وزارة التربية والتعليم في الفصل الدراسي الثاني/ خلال شهر أيار من كل عام ويستهدف نتاجات التعلم في مبحثي اللغة العربية والرياضيات على جميع طلبة الصف الثالث الأساسي؛ إذ يتقدم نصف الطلبة لمبحث اللغة العربية والنصف الآخر لمبحث الرياضيات، ويطبق الاختبار جماعيًّا بصورة ورقية، والغرض الرئيس له هو تحديد جوانب القوة والضعف في أداء الطلبة، ووضع الحلول المناسبة لمعالجتها قبل الانتقال إلى الصف الرابع، ويُبنى على نتائج هذا الاختبار قرارات مهمة على مستوى المدرسة والمديرية والإدارات الفنية في مركز الوزارة، ويشبه إلى حد كبير في تصميمه وتنفيذه وكتابة التقرير عن نتائجه الاختبارات الوطنية لضبط جودة التعليم، ويصنف أداء الطلبة ضمن أربعة مستويات لكن بنقاط قطع تختلف عن نقاط القطع للاختبارات الوطنية لضبط جودة التعليم، وقد شهد تطويرًا في إجراءاته في العام 2024 حيث تم إضافة محور الاستماع إلى محاور اللغة العربية لأول مرة.

ويهدف الاختبار التقييمي الوطني إلى:

1. تعرف المستوى التعليمي لطلبة الصف الثالث الأساسي في مبحثي اللغة العربية والرياضيات.
2. تزويد المعلمين بمعلومات عن جوانب القوة والضعف عند طلبتهم تعينهم في بناء الخطط العلاجية في المهارات الأساسية في مبحثي اللغة العربية والرياضيات.
3. مراجعة المناهج والكتب المدرسية وتحسين استراتيجيات التعليم والتعلم في ضوء نتائج الطلبة.

5. اختبار الاستعداد الدراسي للروضة (Readiness to Learn Test)

لقد تم بناء صورة معربة لاختبار الاستعداد الدراسي للروضة باستخدام أداة النمو المبكر Early Development Instrument (EDI)، وتم التواصل مع جامعة مكماستر (McMaster) التي منحت وزارة التربية والتعليم حقوق استخدام الأداة وذلك بتمويل من اليونيسف، وبالتالي تم السماح للمركز الوطني لتنمية الموارد البشرية باستخدام أداة النمو (EDI) والتي تشخص لدى طفل الروضة الصحة الجسمية والكفاءة الاجتماعية والنضج الاجتماعي والتطور اللغوي والمعرفي ومهارات التواصل والمعرفة العامة حيث أن غرض الاختبار هو تقييم استعداد الطلبة بعد مرحلة رياض الأطفال للالتحاق الآمن بالتعليم الأساسي في الصف الأول، والكشف عن جوانب الضعف قبل الانتقال إلى الصف الأول الأساسي.

ويطبق الاختبار كل أربع سنوات من قبل المركز الوطني لتنمية الموارد البشرية بالتعاون مع وزارة التربية والتعليم ويستهدف قياس استعداد عينة عشوائية (Sample) من أطفال الروضة في الأردن للتعلم، وتحديد مستويات الاستعداد للتعلم في المدرسة حسب الأبعاد الفرعية لأداة النمو المبكر، وتحديد الأبعاد التي يظهر الأطفال ضعفًا فيها، والتعرف إلى العوامل التي تفسر التباين في الاستعداد للتعلم، وتقدير الأهمية النسبية لتلك العوامل، وتتبع التغير في مستويات الاستعداد للتعلم عبر الزمن وذلك للأعوام 2010، 2014، 2018. ويطبق الاختبار ورقيًّا وبصورة فردية.

ويهدف اختبار الاستعداد الدراسي للروضة إلى:

1. قياس الاستعداد للتعلم لدى الأطفال في الأردن.
2. تحديد مستويات الاستعداد للتعلم في المدرسة حسب الأبعاد الفرعية لأداة النمو المبكر.
3. تحديد الأبعاد الفرعية لأداة النمو المبكر التي يظهر الأطفال ضعفًا فيها.
4. التعرف إلى العوامل التي تفسر التباين في الاستعداد للتعلم.

5. تقدير الأهمية النسبية لتلك العوامل.
6. تتبع التغير في مستويات الاستعداد للتعلم عبر الزمن.

2. الترتيبات المؤسسية واللوجستية/التقنية والتمويل ومصادره للتقييمات الوطنية الراهنة.

تتشابه الترتيبات المؤسسية للتقييمات الوطنية المطبقة على الطلبة في نظام التعليم الأردني حيث تشرف على تنفيذها إدارة الامتحانات والاختبارات. وتحدد الخطة السنوية لعمل إدارة الامتحانات والاختبارات، الاختبار المستهدف والخطة الزمنية لتنفيذه. وتصدر لجنة الامتحانات في وزارة التربية والتعليم كتاب التكليف، وتعتمد الخطة الزمنية للتنفيذ. وتتولى إدارة الامتحانات والاختبارات تنفيذ ما ورد في كتاب التكليف في ضوء تعليمات إعداد الاختبارات الوطنية، حيث تُعِدُّ مخطط تدفق إجراءات الاختبار باستعمال نماذج الأيزو المصممة لضبط الجودة، وتتابع بواسطتها إجراءات إعداد الاختبار وتطبيقه وتصحيحه وتحليل نتائجه وكتابة التقرير عنه وتوزيعه على الجهات ذات العلاقة ومناقشته مع المعنيين في وزارة التربية والتعليم وغيرها.

ويحدد مخطط تدفق إجراءات إعداد الاختبار الوطني مهام أقسام المديرية المختلفة ومنها الترتيبات التقنية الخاصة بالاختبارات الإلكترونية ولجان العمل لكل إجراء، حيث تصدر المديرية كتب التكليف وكتب حجز المخصصات المالية، وكتب طرح عطاءات اللوازم لمتطلبات الطباعة والتغليف والحفظ، وتوفير مستلزمات خطة الأزمات والطوارئ والخطط البديلة للامتحان، وتُعَدُّ موازنة وزارة التربية والتعليم هي المصدر الوحيد والمستدام لتمويل التقييمات الوطنية المطبقة على الطلبة في نظام التعليم الأردني وتغطية كل نفقاتها.

ومن جهة أخرى يقوم المركز الوطني لتنمية الموارد البشرية (NCHRD) بالترتيبات المؤسسية واللوجستية/التقنية وتوفير التمويل لاختبار الاستعداد الدراسي للروضة منذ العام 2010. ووفرت منظمة اليونيسيف ووزارة التربية والتعليم التمويل والدعم الفني والتقني في مرحلة إعداد الصورة الأردنية للاختبار الذي تم تسكينه في المركز الوطني لتنمية الموارد البشرية، وتوالى تطبيق الاختبار من قبل المركز للعامين 2014، 2018، وتوقف التطبيق بسبب جائحة كورونا ولم يستأنف بعد.

وتدعم الوكالة الأمريكية للتنمية الدولية (USAID) من خلال برنامج ASAS[12] العديد من التقييمات والاختبارات التشخيصية للصفوف المبكرة، كما أجرى برنامج الدعم الفني[13] Technical Assistance Program TAP دراسة تشخيصية لمهارات الطلبة في اللغة العربية والرياضيات للصفوف (4–11). وعلاوة

[12] USAID Early Grade Education Activity (ASAS)

[13] Technical Assistance Program, USAID.

على ذلك أُستثمر TAP في تدريب العاملين في إدارة الامتحانات والاختبارات على كيفية بناء الاختبارات التشخيصية وإجرائها وتحليل نتائجها وتفسيرها وإعداد التقارير حولها.

3. آليات كتابة التقارير والإبلاغ عن النتائج في التقييمات الوطنية في نظام التعليم الأردني

يوجد لدى إدارة الامتحانات والاختبارات نموذج موحد يبين آليات كتابة التقارير عن النتائج في التقييمات الوطنية المطبقة على الطلبة في نظام التعليم الأردني، ويتكون التقرير من أربعة أجزاء بالإضافة إلى المقدمة والملخص هي: الجزء الأول (الإطار النظري) ويتضمن التمهيد، وتعريف بالاختبار، وأهدافه، ومراحل تطوره، والجزء الثاني (الطريقة والإجراءات) ويصف مجتمع الدراسة وعينتها، وخطوات إعداد الاختبار وخصائصه السيكومترية، وإجراءات الدراسة، ومؤشرات الأداء ومعايير تفسير الدرجات، والجزء الثالث (نتائج الدراسة ومناقشتها) ويعرض نتائج كل مبحث على حدة وتفسير محكي المرجع وآخر معياري المرجع للأداء على محاور التعلم ومهارات التعلم الأساسية للنسخة الورقية والإلكترونية، ثم يتم عرض ومناقشة نتائج المديريات بالأسلوب نفسه علمًا بأن العلامات المعتمدة في التقارير هي العلامات الخام أو العلامات الخام محولة إلى نظام مئوي (التدريج من 0 إلى 100)، وأخيرًا الجزء الرابع: التوصيات.

ويتم تحليل النتائج على مستوى الطالب والصف والمدرسة وتزويد المديريات بنسخ إلكترونية (CD) من النتائج، تقدم تقارير عن أداء الطالب على محاور التعلم ومهاراته الأساسية، وتعرض نتائج الصف والمدرسة بالأسلوب نفسه. وتشير التقارير الخاصة بنتائج الطلبة على الامتحانات الوطنية المختلفة إلى تدني الوسط الحسابي لأداء الطلبة بصورة كبيرة، وقد يعزى ذلك في أحد جوانبه إلى حقيقة لا يمكن تجاهلها تتعلق بضعف الدافعية وعدم الجدية لدى الطلبة لكون العلامات في هذه الامتحانات لا تحتسب في تقييم الطالب. ويضاف إلى ذلك، أن الطالب لا يعرف المبحث الذي يستهدفه الامتحان إلا ساعة التقدم للامتحان، أي أنه لا يوجد تهيؤ عقلي واستعداد مسبق للامتحان لدى الطلبة، والتباين في أداء الطلبة يصعب تفسيره لعدم جمع معلومات عن المتغيرات السياقية التي تتعلق بالطالب ومدرسته ومعلمه والتي قد تؤثر في التحصيل.

ومن جهة أخرى لا تختلف آليات كتابة التقارير والإبلاغ عن النتائج في اختبار الاستعداد الدراسي للروضة عن التقييمات الوطنية التي تشرف عليها إدارة الامتحانات والاختبارات، حيث ينتهي دور المركز الوطني لتنمية الموارد البشرية بإرسال تقرير الدراسة للوزارة، وتتولى الوزارة الإبلاغ عن النتائج.

4. التقييمات الوطنية في الدراسات والتقارير الوطنية والدولية.

أظهرت الدراسات والتقارير الوطنية والدولية التي تناولت التقييمات الوطنية أن الأطر المستخدمة تساير في جودتها أطر التقويم في الدول المتقدمة ولكنها تواجه تحديات لا بد من الوقوف عندها، فقد أشارت استراتيجية تنمية الموارد البشرية (2015–2025) إلى أنه ما زال "هناك عدم توافق بين التقييمات الوطنية والتقييمات الدولية التي تجري في المدارس، وهناك غياب للتنسيق بين هذه التقييمات وتداخل في أهدافها، وعدم القدرة على القياس الدقيق لأداء الطلبة من بداية الصف الأول وحتى نهاية الصف الثاني عشر. وتشير الاستراتيجية أيضًا إلى ضرورة تطوير نظام تقويم تكويني وختامي في آن واحدٍ، يقيس بدقة أداء جميع الطلبة وتطورهم حسب المراحل الدراسية". وأوصت بضرورة وضع إطار عمل وسياسات تقويم لمرحلتي التعليم الأساسي والثانوي لضمان توافق جميع التقييمات لكي يتمكن المعنيون من تتبع وقياس تطوّر النظام وأداء الطلبة المتلقين بدقة في كل مرحلة من مراحل تعلّمهم.

وفي مجال الجودة، أشارت الخطة الاستراتيجية لوزارة التربية والتعليم (2018–2022) إلى وجود تحديات تتعلق بعدم كفاية التقييم الفعلي وعدم مقدرته على قياس أداء الطلبة ومهاراتهم ومعارفهم بشكل شمولي. واستجابة لذلك، أفْرَدَتِ الوزارة ضمن خطتها الاستراتيجية (2018–2022) مكوّنًا فرعيًّا خاصًا بتطوير ومراجعة نظام التقييم الوطني للوصول إلى تقييم شامل لمهارات الطلبة وسلوكياتهم وقدراتهم ومعارفهم وبما يتماشى مع المعايير الدولية، وشملت الخطة التنفيذية لهذا المكون الأنشطة الآتية:

- تطوير استراتيجية تقييم وطنية.
- تحديد مراحل التقييم بما يتماشى مع هيكل سلم التعليم الجديد (نهاية المرحلة الأساسية الدنيا، نهاية المرحلة الأساسية العليا، ونهاية المرحلة الثانوية).
- اعتماد مؤشرات أداء للطلبة في كل مرحلة.
- إعداد امتحانات معيارية للصفين الثالث والتاسع وفقًا لمعايير دولية وتطبيقها على المستوى الوطني.
- تنفيذ الاختبار التشخيصي للصف الثالث الأساسي في مجالي القراءة والرياضيات في الصفوف المبكرة.
- تطوير معايير وأسس المتابعة والمساءلة للمشاركين في التقييمات الوطنية وتدريب الموظفين على استخدامها وتطبيقها في المدارس.

وقد أشار تقرير البنك الدولي[14] فيما يتعلق بالتقييمات الوطنية، إلى أن هناك حاجة إلى المزيد من الكوادر المتخصصة للمساعدة في تنفيذ أنشطة التقييم وتقوية آليات ضمان الجودة، وقد أوصى التقرير بمضاعفة الجهود داخل الوزارة لتنمية الموارد البشرية لضمان تحسين التوافق والاتساق في الاختبار مع الإصلاح المستمر للمناهج الدراسية، وإدخال ممارسات مبتكرة لتقييم الطلبة، ودمج تكنولوجيا المعلومات والاتصالات في المناهج الدراسية والتقييمات، وإتاحة الفرص على نطاق واسع للمعلمين للتعرف على التقييمات الوطنية، وتقديم ورش عمل لهم حول المجالات التي تقيسها وكيفية قراءة نتائجها واستخدامها، وقد أكد التقرير الإقليمي حول خريطة التقييم في الدول العربية 2014 [15] على النقاط الواردة في تقرير البنك الدولي، وأضاف إليها أن الأداء على الاختبارات الوطنية لا يحتسب في درجات الطلبة وليس له أي اعتبار في نجاحهم مما يقود إلى اللامبالاة وعدم الجدية، ويقتصر ضبط جودة الاختبارات على المراجعات الداخلية، وقد صنف الأردن في المستوى "ناشئ" وهو المستوى نفسه الذي حدده تقرير البنك الدولي للأردن في التقييمات الوطنية واسعة النطاق، وبالرغم من مضي حوالي عشر سنوات على التقريرين، إلا أن التوصيات والملاحظات التي وردت فيهما لا زالت قائمة، وهذا يتطلب مضاعفة الاهتمام بالتقييمات الوطنية ووضعها في الإطار الصحيح وبالفعل، فقد أشار 71% من المعلمين والمعلمات الذين تم استطلاعهم أنه لم تتح لهم فرص الاطلاع على تقارير الاختبارات الوطنية، مما يعني أنه لا يمكن الجزم بتحقيق تلك الاختبارات لأهدافها.

4. الممارسات الدولية الفضلى في الاختبارات الوطنية

تشير دراسة حديثة[16]، إلى أن التقييم عالي الجودة، والمناهج المطورة، وطرق التدريس الحديثة تشكل المثلث الذهبي للنظام التعليمي ذي الأداء العالي، وعند توافرها في بيئة مُمكنة، سيكون لها تأثير عميق على نتائج التعلم، وينصب تركيز المنهاج على تطوير الطلبة مجموعة واسعة من المهارات ذات الصلة بالمبحث التي تدعم تقدمهم في المدرسة وعلى نطاق أوسع كمواطنين عالميين، أي تحدد نتاجات التعلم عند مستوى مرتفع نسبيًّا للمهارات التي يتوقع من الطلبة أن يكونوا قادرين عليها ومكتوبة بطريقة تساعدهم في تطوير مهارات التفكير العليا التي تعدهم بشكل أفضل للحياة القادمة، ويتوقع أن يكون هناك تأثير إيجابي لذلك على كل من تقييمات المعلمين لطلبتهم في الفصول الدراسية والتقييمات الوطنية واسعة النطاق.

---

[14] JORDAN Student Assessment /SABER Country Report 2014

[15] أعد التقرير عدنان الأمين بتكليف من مكتب اليونسكو الإقليمي للتربية في الدول العربية والمنظمة العربية للتربية والثقافة والعلوم

[16] A suggested assessment framework for CBSE science, mathematics and English for classes 6 to 10. 2021

وقد اكتسبت تقييمات التعلم واسعة النطاق أهمية كبيرة في العقود الماضية[17] باعتبارها شكل من أشكال الاختبارات الموحدة الوطنية التي توفر لمحة سريعة عن التحصيل التعليمي لمجموعة من المتعلمين في سنة معينة وفي عدد محدد من مجالات التعلم. وقد شجع ذلك التركيز على التعلم العادل والفعال والملائم للجميع والمتأصل في خطة التنمية المستدامة لعام 2030 والتركيز على استخدام البيانات لتحسين السياسات والاستراتيجيات، وقد أتاح التحول الرقمي في الاختبارات التعليمية العديد من الفرص الجديدة للتكنولوجيا لتعزيز التقييمات واسعة النطاق[18]، يتضمن ذلك إمكانية جمع بيانات حول عملية استجابة المتقدم للاختبار بشكل روتيني وعلى نطاق واسع.

ثالثًا: التقييمات الدولية International Large-Scale Assessment

تعود جذور الدراسات الدولية إلى منتصف القرن العشرين، إذ بدأت فكرة مقارنة أداء نظم تربوية مختلفة باستخدام اختبارات مقننة تطبق على عينات كبيرة، وتم بالفعل تنفيذ أول دراسة دولية عام 1964 لقياس أداء طلبة من 12 دولة في الرياضيات، وقد واجهت هذه الدراسات نقدًا واسعًا لأسباب تتعلق بالمنهجية ونطاق القياس ومحدودية الفائدة، كما لم تلق اهتمامًا على مستوى صناع القرار وواضعي السياسات.

وتزايد زخم الاهتمام في التقييمات الدولية منذ منتصف تسعينيات القرن الماضي بسبب عولمة التعليم بشكل عام والتقييم على وجه التحديد، وزيادة الإمكانات التقنية العالمية على إجراء تقييمات واسعة النطاق، وتزايد أهمية رأس المال البشري في التنمية الاقتصادية (Kamens, 2013; Lockheed, 2010). وتتنوع مسوغات المشاركة في الدراسات الدولية، ومنها الآتي (Addey & Sellar, 2019):

1- توليد أدلة موثوقة لصنع السياسات (OECD, 2014; IEA, 2015).
2- تتيح الفرص لبناء القدرات التقنية وتطوير التقييمات الوطنية من خلال اكتساب منهجيات إحصائية مبتكرة ومتقدمة (مثل نظرية استجابة الفقرة)، لتتماشى مع المعايير الدولية، كما تشجع الدول على الربط بين عمليات التقييم الدولية والوطنية (Addey, 2017; OECD, 2013; Phillips & Jiang, 2011).
3- استخدامها من قبل الجهات المانحة كمعيار على تحسن النظام التربوي ودليل على الالتزام بالمساءلة والشفافية وهي شروط لتقديم التمويل والمساعدات (Lockheed et al. 2015).

[17] UNESCO. (2019). The promise of large scale learning assessments: Acknowledging limits to unlock opportunities.
[18] Maddox, B. (2023). The uses of process data in large-scale educational assessments. OECD Education Working Papers No 286.

4- تعزيز العلاقات الدولية من خلال الاطلاع على النظم التعليمية للدول ذات الأداء العالي. وأصبحت المشاركة في هذه الدراسات مطلبًا سياسيًا عالميًا، يرسل عدم المشاركة إشارة مفادها أن الدولة غير ملتزمة بشكل كاف بتطوير نظامها التربوي (Addey & Sellar, 2019).

5- استجابة لضغوطات وسياسات وطنية من أجل التحسين أو مسألة رأي عام (Addey, 2015).

6- قياس جودة النظام التربوي في القدرة على المنافسة (Klees, 2021; Kamens, 2015).

7- تطوير المناهج وطرق التدريس والتقييم (Lockheed, 2013).

1. مشاركة الأردن في التقييمات الدولية

حرص الأردن على المشاركة في التقييمات الدولية، ويبين الجدول 2 التقييمات التي يشارك فيها الأردن.

جدول 2: التقييمات الدولية التي يشارك بها الأردن

اسم التقييم	غرض التقييم	المباحث المستهدفة	دورة التقييم	أسلوب التطبيق		الفئات المستهدفة
				المشاركون	صورة الاختبار	
دراسة التوجهات الدولية في الرياضيات والعلوم (TIMSS)	تقييم تحصيل الطلبة في الرياضيات والعلوم	الرياضيات والعلوم	أربع سنوات	عينة	جماعي/إلكتروني	الصفين 4، 8
الدراسة الدولية للتقدم في القرائية (PIRLS)	تقييم تحصيل الطلبة في الاستيعاب القرائي	اللغة العربية	خمس سنوات	عينة	جماعي / ورقي	الصف 4
البرنامج الدولي لتقييم الطلبة (PISA)	قياس الكفايات في القراءة والعلوم والحساب مقارنة بالمعايير الدولية	القرائية والعلوم والرياضيات	ثلاث سنوات	عينة	جماعي/الكتروني	عمر 15 سنة

2. الغرض من التقييمات الدولية التي يشارك فيها الأردن.

شارك الأردن في جميع دورات TIMSSمنذ 1999. وتركز TIMSS على قياس التحصيل التعليمي في الرياضيات والعلوم بشكل فعال في الصفين الرابع والثامن الأساسيين، وتجمع معلومات مفصلة حول المناهج وتنفيذها، إلى جانب معلومات واسعة النطاق حول العوامل السياقية في المدرسة والبيت المرتبطة بالتعلم وتحصيل الطلبة. وتوفر TIMSS بيانات مهمة عن اتجاهات تحصيل الطلبة مع مرور الوقت، ومنذ عام 1995، أصبحت بيانات TIMSS أداة لا تقدر بثمن للدول المشاركة لاتخاذ قرارات مبنية على الأدلة بشأن

أنظمتها وسياساتها التعليمية، ويمكن للدول مقارنة نتائجها عبر الدورات المختلفة التي شاركت فيها لقياس النمو، وتقييم تأثير سياسات التعليم والمناهج التعليمية، ومقارنة نتائجها بنتائج الدول الأخرى.

وتعد (PISA) من الدراسات الدولية الأخرى التي تمت المشاركة فيها، وهو برنامج طورته وتديره منظمة التعاون الاقتصادي والتنمية (OECD)، لاختبار مهارات ومعرفة الطلبة البالغين من العمر 15 عامًا في القراءة والرياضيات والعلوم، وهي بمثابة دراسة استقصائية لتقييم مدى اكتساب الطلبة للمعرفة والمهارات الأساسية الضرورية للمشاركة الفاعلة في الحياة الاجتماعية والاقتصادية، وتقدم (PISA) أدلة مقارنة حول أداء أنظمة التعليم في الدول المشاركة وجودة التعليم والمساواة، ورفاهية الطلبة، وطموحاتهم الحياتية، والمناهج الدراسية، وبيئة التعلم والمواد المدرسية وغيرها من المعلومات السياقية، ويتم إجراء (PISA) كل 3 سنوات باستثناء تأخير في الدورة الثامنة (من 2021 إلى 2022) بسبب جائحة كورونا. وفي كل دورة، يتم أخذ مجال واحد محدد باعتباره المجال الرئيس، ويشغل حوالي ثلثي وقت الاختبار، أما باقي وقت الاختبار فهو مقسم بين المجالين "الثانويين" الآخرين.

كما شارك الأردن في دراسة PIRLS (PIRLS Progress In International Reading Literacy Study,) لأول مرة عام 2021 باختبار الورقة والقلم، وسيشارك الأردن في دورة عام 2026 بصورتها الإلكترونية. ويتم إجراء دراسة (PIRLS) كل خمس سنوات منذ عام 2001، ويُعترف بها كمعيار عالمي لتقييم الاتجاهات في التحصيل القرائي في الصف الرابع، حيث توفر (PIRLS) بيانات مقارنة دولية حول مدى جودة الاستيعاب القرائي للطلبة وتقدم معلومات ذات صلة بالسياسات لتحسين التعلم والتدريس، وجاءت الدراسة استجابة للتحول من تعلم القراءة إلى القراءة من أجل التعلم، ويوفر تقييم تحصيل القراءة في هذه المرحلة الحاسمة للمعلمين وواضعي السياسات رؤى أساسية حول فعالية نظامهم التعليمي ويساعد على تحديد مجالات التحسين. ويقوم الاختبار بتقييم عمليات التركيز على المعلومات واسترجاعها، وإجراء استنتاجات مباشرة، وتفسير الأفكار والمعلومات ودمجها، وتقييم ونقد المحتوى والعناصر النصية، وتوفر استبانات (PIRLS) الخاصة بالمدرسة والمعلمين والطلبة وأولياء الأمور معلومات واسعة النطاق حول العوامل السياقية في البيت والمدرسة المرتبطة بتعليم وتعلم القراءة.

وتفيد المشاركة في التقييمات الدولية في تحقيق الأغراض الآتية:

- تقييم تحصيل الطلبة في المباحث والصفوف المستهدفة في تلك التقييمات، وتتبع التغيير في تحصيل الطلبة مع الوقت.
- مقارنة أداء الطلبة الأردنيين في المباحث المستهدفة مع أداء طلبة الدول الأخرى.

- الكشف عن الفروق في التحصيل بين مجموعات فرعية مختارة حسب الجنس، ونوع المدرسة، والمنطقة الجغرافية، والسلطة المشرفة، والمستوى الاقتصادي الاجتماعي وغيرها.
- الحصول على معلومات سياقية حول المدارس والمعلمين والطلبة، بما في ذلك الموارد، ومواقف المعلمين والطلبة، من خلال استبانات موجهة للمدرسة والمعلمين والطلبة وأولياء أمورهم.
- تحديد احتياجات التدخل على مستوى السياسات، وتقييم آثار التدخل مع مرور الوقت.
- التعرف على الإمكانات البحثية في مجال التحصيل المقارن للموضوع.

3. الترتيبات المؤسسية واللوجستية للتقييمات الدولية في الأردن:

يتولى المركز الوطني لتنمية الموارد البشرية (NCHRD) إدارة التقييمات الدولية في الأردن، ويقع على إجراء الترتيبات اللازمة لهذه الدراسات التنسيق بين الجهات الدولية والمؤسسات التربوية المختلفة (وزارة التربية والتعليم، والثقافة العسكرية، ووكالة الغوث الدولية، والمدارس الخاصة)، وجمع البيانات، وإصدار التقارير، بهدف إعانة واضعي السياسة التربوية وطنيًا على تطوير النظام التربوي الأردني والارتقاء بجودته.

ويقوم المركز بمراجعة إطار التقييم لكل دراسة، وإجراء مراجعة علمية لفقرات الاختبارات والاستبانات المعدة من قبل الجهات الدولية المشرفة على تلك التقييمات للكشف عن التحيز والغموض والتعقيد في الفقرات ومدى ملاءمتها للمناهج الدراسية في الأردن، كما يقوم المركز بترجمة الفقرات من اللغة الإنجليزية إلى اللغة العربية بالاعتماد على مجموعة من المختصين في المجال المعني وسبق لهم المشاركة في إعداد المناهج والكتب المدرسية في الأردن بناء على تزكية من المركز الوطني لتطوير المناهج، كما يتم إجراء تدقيق لغوي من خلال متخصص في اللغة العربية، ويقوم المركز بإرسال الفقرات إلى الجهات الدولية التي تقوم بالتحقق من جودة عمليات الترجمة والتدقيق.

ويشكل وزير التربية والتعليم اللجنتين التوجيهية والفنية برئاسة عطوفة الأمين العام للشؤون التعليمية وعضوية المركز الوطني لتنمية الموارد البشرية، ومديرية التربية والتعليم والثقافة العسكرية، ومديرية التربية والتعليم لوكالة الغوث الدولية، وإدارة التعليم الخاص، وإدارة الإشراف والتدريب التربوي، ومركز الملكة رانيا العبدالله لتكنولوجيا التعليم والمعلومات،حيث تعمل اللجنة الفنية على تطوير خطط الاستعداد للتقييمات الدولية وإقرارها من قبل اللجنة التوجيهية قبل تعميمها على الميدان التربوي، ويقع على عاتق إدارة الاشراف والتدريب

التربوي في وزارة التربية والتعليم ومديريات التربية والتعليم في السلطات التربوية الأخرى مسؤولية مراقبة تنفيذها، وتتضمن خطة الاستعداد المجالات الرئيسة الآتية: تجهيز مدارس العينة والطلبة للاختبار، الجاهزية التكنولوجية لمدارس العينة، والدعم الفني من ضباط الارتباط في مركز الوزارة والميدان، وعملية جمع البيانات.

وتقوم وزارة التربية والتعليم بتزويد المركز ببيانات تفصيلية حول مجتمع الدراسة المستهدف بالاعتماد على البيانات المتوافرة على منظومة EMIS. ويتم التشارك بين وزارة التربية والتعليم والسلطات التربوية الأخرى والمركز الوطني لتنمية الموارد البشرية لإعداد إطار المعاينة وتحديد الطبقات الصريحة والضمنية لضمان تمثيل مجتمع الدراسة المستهدف وذلك من خلال عقد اجتماعات دورية للجنة التوجيهية واللجنة الفنية المعنيتين بالاستعداد للدراسات الدولية.

واستمرت مشاركة الأردن في التقييمات الدولية باستخدام اختبارات الورقة والقلم حتى عام 2022، وتحول بعدها للمشاركة في الاختبارات بصورة محوسبة، لذا تقوم الجهة الدولية المشرفة على الدراسة بتزويد المركز بمواصفات الأجهزة التي تضمن نجاح تطبيق الاختبار، ويقوم مركز الملكة رانيا العبدالله لتكنولوجيا التعليم والمعلومات بالتحقق من جاهزية الأجهزة في مختبرات مدارس العينة.

ويقوم المركز بالتعاون مع وزارة التربية والتعليم بتنظيم عملية جمع البيانات من الميدان التربوي وبما يتوافق مع المعايير التي تحددها الجهات الدولية لهذه الغاية، وذلك لضمان توحيد إجراءات التطبيق في جميع الدول المشاركة، مما يساعد على الحصول على بيانات موثوقة قابلة للمقارنة، إذ يتم جمع البيانات في الأردن بواسطة مشرفين تربويين عاملين في وزارة التربية والتعليم بعد تدريبهم على إجراءات عملية التطبيق في المركز من قبل المنسق الوطني للدراسة المعنية، ويتم تزويدهم بالمواد والأدوات اللازمة لنجاح عملية التطبيق. ويتم إعداد فريق ضباط جودة وطني من مشرفين عاملين في مركز الوزارة لضمان جودة عملية التطبيق، كما تقوم الجهات الدولية بتعيين ضباط جودة دوليين.

4. تمويل التقييمات الدولية في الأردن

يقوم الأردن بتغطية رسوم المشاركة في الدراسات الدولية، وتكييف أدواتها للبيئة الأردنية، وعمليات التصحيح للاختبارات وإدخال البيانات، وحضور اجتماعات الخبراء الدولية بمنحة من الوكالة الأمريكية للتنمية الدولية (USAID)، وتساهم وزارة التربية والتعليم من خلال ميزانيتها بعملية فحص الجاهزية التكنولوجية لمختبرات مدارس العينة وعملية جمع البيانات من الميدان التربوي، ويساهم المركز الوطني لتنمية الموارد

البشرية بعملية تنظيف البيانات وتجهيزها وإرسالها للجهات الدولية، وإعداد التقارير من ميزانيته الخاصة. ويبين الجدول 3 رسوم المشاركة والتكلفة المحلية التقديرية للتقييمات الدولية التي يشارك فيها الأردن بالدينار.

جدول 3: رسوم المشاركة والتكلفة المحلية التقديرية للتقييمات الدولية بالدينار الأردني

الدراسة	رسوم المشاركة	التكلفة المحلية	المجموع	التكلفة / سنويا
TIMSS /صف 4	272000	100000	372000	93000
TIMSS /صف 8	272000	100000	372000	93000
PIRLS	272000	100000	372000	74400
PISA	131670	100000	231670	77223
المجموع	947670	400000	1347670	337623

5. آليات كتابة التقارير والإبلاغ عن نتائج التقييمات الدولية في الأردن

تقوم الجهة الدولية المشرفة على كل دراسة بإعداد وإصدار تقرير دولي مفصل يعرض نتائج الدول المشاركة في الدراسة، ويتضمن التقرير متوسطات الدرجات لكل مجال من مجالات الدراسة ولكل دولة مشاركة وفق نظام مقارنة مرجعي بوسط حسابي يساوي 500 وانحراف معياري يساوي 100. ويقدم التقرير عرضًا مفصلًا لاتجاه النتائج عبر دورات الدراسة المختلفة، مما يساعد الدول المشاركة على مراقبة التغير في تحصيل طلبتهم عبر الزمن، وتوفر مثل هذه التقارير تحليلات معمقة لمحاولة ربط التحصيل بالعديد من المتغيرات الديموغرافية والسياقية مثل، الجنس، الوضع الاجتماعي الاقتصادي، وغيرها.

ويقدم التقرير تفسيرًا لمتوسطات درجات الدول، مقارنة بمستويات أداء دولية محددة مسبقًا على مقياس التحصيل، ويتم وصف أداء الطلبة في كل مستوى من مستويات الأداء بدلالة محتوى الفقرات التي أجاب عنها الطلبة الذين وصلوا لذلك المستوى، وتوفر هذه التقارير معلومات غنية حول سياقات المنزل والمدرسة والصف الذي يتعلم فيه الطلبة للمساعدة في تحسين تدريس الرياضيات والعلوم والقراءة.

ويقوم المركز الوطني لتنمية الموارد البشرية بإعلان نتائج كل دراسة عقب الإعلان الرسمي من الجهة الدولية المشرفة على الدراسة من خلال مؤتمر صحفي تحت رعاية وزير التربية والتعليم وبحضور الشركاء من جميع القطاعات والمؤسسات التربوية المختلفة، ودعوة مجموعة من الخبراء التربويين المهتمين ووسائل الإعلام. و ثم يتم نشر التقارير الوطنية والبيانات المتعلقة بالدراسة على الموقع الإلكتروني للمركز الوطني لتنمية الموارد البشرية، وتتاح البيانات لأصحاب المصلحة والباحثين، وتتناول التقارير الوطنية بشيء من التفصيل مستويات

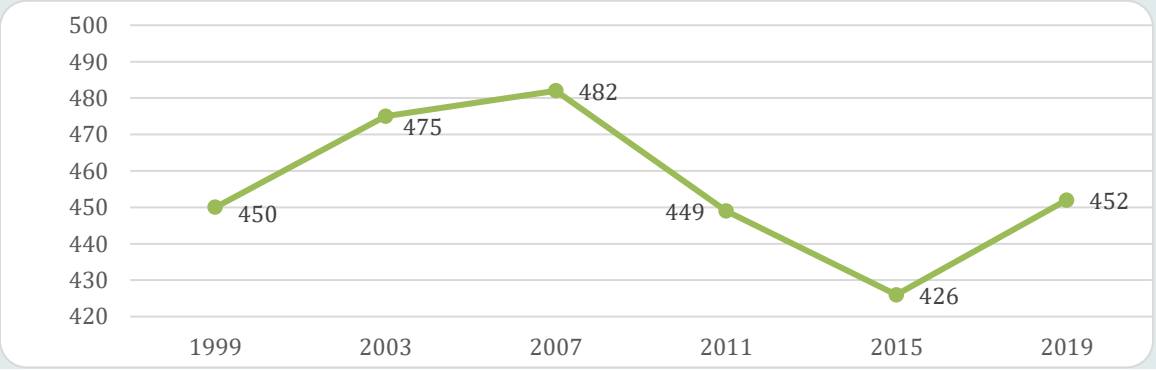
أداء طلبة الأردن مقارنة بأداء الدول العربية والدول الأخرى المشاركة في الدراسة، وبعض التحليلات الثانوية على المستوى الوطني وفقًا لمتغيرات جنس الطالب، والسلطة التربوية، وموقع المدرسة، وعدد فترات دوام المدرسة، وإقليم المدرسة، وطبقة المدرسة (وزارة التربية والتعليم، ووكالة الغوث الدولية، والثقافة العسكرية، والتعليم الخاص، ومدارس السوريين). وتوفر هذه التقارير تحليلات لربط التحصيل بالعديد من المتغيرات الديموغرافية والسياقية مثل، الجنس، الوضع الاجتماعي الاقتصادي، وغيرها من خلال مقاييس مشتقة من استبانات الطالب والمعلم والمدرسة وولي الأمر. ويُعد المركز الوطني لتنمية الموارد البشرية أدلة إرشادية موجهة لمعلمي اللغة العربية والرياضيات والعلوم لمعالجة أخطاء التعلم عند الطلبة في ضوء نتائجهم على هذه الاختبارات.

ويقوم المركز الوطني لتنمية الموارد البشرية بتوجيه دعوة لعمداء كليات التربية في الجامعات الحكومية والخاصة الأردنية لزيارة المركز والاطلاع على البيانات التي توفرها الدراسات الدولية، إذ يتم تزويدهم بالبيانات المحدثة لجميع الدراسات التي يشارك فيها الأردن، وذلك لدعوة الأساتذة وطلبة الدراسات العليا في هذه الجامعات إلى إجراء المزيد من الدراسات والأبحاث لهذه البيانات للمساهمة في تطوير مناهج الرياضيات والعلوم، وبرامج إعداد المعلمين وتحسين البيئة المدرسية للنهوض بمستويات أبنائنا الطلبة والارتقاء بجودة نظامنا التربوي في الأردن.

6. أداء الطلبة الأردنيين في التقييمات الدولية:

في جميع التقييمات الدولية يقل الوسط الحسابي لأداء طلبة الأردن عن الوسط الدولي بصورة دالة إحصائيًا، وتراجع الأداء في آخر دورة لكل منها، ويبين الشكل 2 الأوساط الحسابية للطلبة الأردنيين في العلوم في دراسة TIMSSعبر دورات الدراسة التي سبقت دورة 2023.

شكل 2: أوساط الطلبة الأردنيين في العلوم في دراسة *TIMSS* عبر دورات الدراسة



7.التقييمات الدولية في التقارير والدراسات السابقة

تشير التقارير الصادرة من المؤسسات والمنظمات الراعية للاختبارات الدولية (مثل(OCED) أنه على الرغم من التحسن الكبير في منهجيات وجودة التقييمات الدولية، فإنه لا زال هناك انتقادات توجه لها، فقد تزايد لجوء بعض الدول لتبني سياسات دول حققت مراتب متقدمة في هذه التقييمات دون الأخذ بعين الاعتبار السياقات التربوية والثقافية لهذه الدول، إضافة إلى توجه بعض الدول لتدريب الطلبة على الاختبارات بهدف تحسين رتبتها ونتائج طلبتها فقط دون الاهتمام بالتحسن الفعلي لنظمها التربوية، ومحدودية الاستفادة من نتائج هذه التقييمات للقصور في فهم وتفسير النتائج ضمن سياقاتها السليمة، مما قاد إلى استنتاجات خاطئة، كما يرافق التقييمات الدولية واسعة النطاق صعوبة التعامل مع التنوع الكبير في الثقافات واللغات والأعراق بين الدول المشاركة في هذه التقييمات، مما يخلق بعض المشاكل المتعلقة بالتحيز الثقافي وصعوبات بعمليات الترجمة ونقل المعاني والسياقات الثقافية، مما قد يؤثر على نتائج التقييمات وقابلية المقارنة (Bray at all., 2020).

وفي بعض الأحيان تختار بعض الدول عدم المشاركة في الدراسات الدولية أو تنسحب منها لأسباب أيدولوجية تتعلق بإعلان موقف بعدم الانتماء سياسيًّا أو اقتصاديًّا لمجموعات اقتصادية أو سياسية معينة، أو تربوية وثقافية، كون بعض هذه الدراسات غير مناسبة للسياقات التعليمية فيها، أو ضعف مواردها وقدراتها، أو المخاوف من الحصول على نتائج سيئة وما قد تسببه من حرج أمام المجتمع الدولي (Addey & Sellar, 2019; Ferrer & Fiszbein, 2015).

وفي الأردن أشارت أكثر من دراسة (مبارك، 2006؛ مبارك ومبارك، 2006) إلى وجود أداء تفاضلي في فقرات اختبار (PISA) في الرياضيات والعلوم لمتغير (اللغة العربية مقابل الإنجليزية) وكان معظمها لصالح اللغة الإنجليزية، حيث من الممكن أن تؤدي ترجمة الفقرات من الأصل باللغة الإنجليزية إلى اللغة العربية بتغيير في خصائص الفقرات ومحتواها، كما كشفت السعيد (2020) عن أسباب تدني أداء الطلبة الأردنيين في الدراسات الدولية من وجهة نظر المعلمين والمشرفين، وتعود تلك الأسباب إلى الطلبة أو المعلمين أو أولياء أمور الطلبة أو المنهاج، فالطالب لا يهتم بالاختبارات الدولية لكونها لا ترتبط بنتيجته المدرسية، ودافعيته للاستجابة عليها ضعيفة، ولا يمتلك الخبرة للتعامل مع الأسئلة المتضمنة في التقييمات الدولية، والمعلم لا يتابع قضايا التقييمات الدولية لزيادة العبء التدريسي له، ولا يوجه اهتمام طلبته نحو مصادر لها علاقة بالاختبارات الدولية، وهو غير مطلع بصورة كافية عليها، وبالفعل أشار 75% من المعلمين الذين تم استطلاع رأيهم إلى

أنه لم تتح لهم فرص الاطلاع على تقارير الاختبارات الدولية، ولا يهتم أولياء أمور الطلبة بأداء أبنائهم في التقييمات الدولية، ويعتقدون أن التركيز عليها هو مضيعة للوقت وعلى حساب نتائج أبنائهم.

وتخلو المناهج القديمة من المهام التقييمية التي تحاكي التقييمات الدولية، ومن الأنشطة الإثرائية التي تركز على العمليات العقلية العليا، ويبين كل من تقرير البنك الدولي 2014 والتقرير الإقليمي حول خريطة التقييم في البلدان العربية 2014 أن تقييم المشاركة الأردنية في التقييمات الدولية في مستوى "ناشئ Emerging"، وأوصوا بمشاركة إدارة الامتحانات والاختبارات في تحليل واستخدام بيانات التقييمات الدولية، لرفع مستوى الوعي بين المشرفين والمعلمين بهدف تلك التقييمات وقيمتها المضافة في تحسين التقييمات الصفية.

8. الممارسات الفضلى في التقييمات الدولية

هدفت بعض الدراسات (Teig & Steinmann, 2023; Wang & Yang, 2024) التعرف على الممارسات التقييمية الفضلى في الأنظمة التعليمية ذات الأداء العالي، وتم استخلاص مجموعة من الممارسات الفضلى بهذا الخصوص وهي:

1- تبني سياسات تعليمية واضحة ومستقرة.
2- بناء المهارات التقنية وتعزيز التطوير المهني للمعلمين والقادة التربويين بشكل مستمر.
3- مراجعة مستمرة للمعايير والمناهج وتحديد المعارف والكفايات التي ينبغي أن يكتسبها الطلبة ويكونوا قادرين على ممارستها في الحياة.
4- التوزيع الأمثل للموارد المالية وغير المالية ومنح المدارس إمكانية الوصول إلى البنية التحتية المادية والموارد التعليمية والتكنولوجيا بشكل عادل.
5- التأكيد على إجراء تقييمات للطلبة على المستوى الوطني والدولي لمواد ومراحل دراسية مختلفة لوضع مقاييس لتطور الطلبة والممارسات التعليمية.
6- الاستخدام الفعال للبيانات، إذ تلعب التقييمات الدولية واسعة النطاق دورًا مهمًّا في استكمال مصادر المعلومات الأخرى لإعلام أصحاب القرار وواضعي السياسات بالجوانب المختلفة للنظام التعليمي.

رابعًا: امتحانات الثانوية العامة (التوجيهي) General Secondary Examinations Tawjihi

تجري وزارة التربية والتعليم امتحانًا عامًا للطلبة في مناهج التعليم الثانوي الشامل في نهاية الصف الثاني عشر يعرف باسم امتحان شهادة الدراسة الثانوية العامّة (التوجيهي)، ويشكل الامتحان مَعلَمًا مهمًا في النظام التعليمي الأردني، وتعود بداياته إلى عام 1962. وقد شهد هذا الامتحان عبر الزمن تعديلات متعددة تتعلق بإدارته وتعليماته حتى وصل إلى ما هو عليه الآن، وقد تركزت التعديلات على عدد مرات الامتحان في السنة أو عدد المباحث التي يتقدم فيها الطلبة في فروع الثانوية المختلفة، والنهايات الصغرى والعظمى لها، وما يحتسب منها في معدل الطالب وأوزانها، ولدى وزارة التربية والتعليم رغبة أكيدة ودافع قوي لتطوير امتحان الثانوية العامة انسجامًا مع الرؤى الملكية السامية لتطوير النظام التعليمي ليتلاءم مع المستجدات العالمية، وتطور المعرفة، ووسائل التكنولوجيا والاتصالات، بما يتفق مع الاستراتيجية الوطنية لتنمية الموارد البشرية. وهذا يتطلب دراسة واقع الامتحان الحالي، ومعرفة ما له وما عليه، وكيف يمكن الاستجابة لما تثيره التقارير والدراسات حوله، وأين هو من الممارسات الدولية الفضلى في هذا المجال، ويعرض هذا الجزء من التقرير البيانات والأدلة حول امتحان الثانوية العامة وملامحه.

1. الامتحان في التشريعات التربوية:

تنص الفقرة أ من المادة (29) من قانون التربية والتعليم رقم (3) لسنة 1994م على ما يلي: "تجري الوزارة امتحانًا عامًا للطلبة في مناهج التعليم الثانوي الشامل، تمنح للناجح فيه شهادة الدراسة الثانوية العامة مُبَيَّنًا فيها نوع التخصص". وتنص المادة (30) من القانون على أن "قرارات الوزارة المتعلقة بإجراء الامتحان العام ونتائجه قطعية ولا تخضع للطعن أمام أي مرجع قضائي أو إداري". وهذا يعني أن الجهة المخولة قانونيًّا بتطوير الامتحان وإدارته هي وزارة التربية والتعليم. وتتولى إدارة الامتحانات والاختبارات في الوزارة هذه المهمة. وحيث إن تعليمات الامتحان وإجراءات عقده وإجراءات الأمن والسلامة محكومة بالقانون وبإشراف من مجلس الامتحان برئاسة معالي وزير التربية والتعليم، ولجنة الامتحان برئاسة عطوفة الأمين العام للشؤون التعليمية، ولجان الامتحانات المحلية برئاسة مديري التربية والتعليم.

وتشرف إدارة الامتحانات والاختبارات على تنفيذ الترتيبات المؤسسية واللوجستية الخاصة بالامتحان. وفي كل عام دراسي تقوم لجنة الامتحان العام بإقرار الإطار العام للامتحان وخطته[19]، وتحدد كل القضايا التي تتعلق بسيره، وتحدد الإجراءات الإدارية والفنية الكفيلة بعقده، وتتابع إجراءات تنفيذه وتصحيحه، وإقرار نتائجه. وتتضمن الخطة مواصفات أوراق الامتحان، وأسس اعتماد مراكز الامتحان، وأسس اختيار واضعي الأسئلة

[19] تعليمات امتحان شهادة الثانوية العامة رقم (8) لسنة 2017 وتعديلاتها

ورؤساء قاعات الامتحان ومساعديهم والمراقبين، وأسس اختيار رؤساء لجان التصحيح ومراكز التصحيح، وبرنامج الامتحان. ويتم توفير تمويل منتظم للامتحان من خلال رسوم الاشتراك في الامتحان ومن قبل الحكومة الأردنية حيث بلغت قيمة نفقات الامتحان للعام 2024 ما مجموعه 20815000 دينار أردني[20]. وتتوفر لإدارة الاختبارات والامتحانات كل التسهيلات الإدارية والفنية لعقد الامتحان وضمان سير إجراءاته ضمن ظروف موحدة وبجودة عالية.

وفي نيسان عام 2017 تأسس المركز الوطني لتطوير المناهج بإرادة ملكية سامية استجابة لتوصيات الاستراتيجية الوطنية لتنمية الموارد البشرية 2016-2025 وفق النظام رقم 33 لسنة 2017 وتعديلاته[21]. ويركز نظام المركز وتعديلاته في المادة (4) على تطوير المناهج والكتب المدرسية والامتحانات وفقًا لأفضل الأساليب الحديثة وبما يتناسب مع احتياجات المملكة ومسيرة التعليم الأمثل وفلسفة التربية والتعليم وأهدافها من خلال:

1. تطوير الاختبارات وامتحان شهادة الثانوية العامة وغيرها من التقييمات المستحدثة للطلبة.
2. إنشاء بنوك أسئلة لأغراض الاختبارات والامتحانات الوطنية والدولية والتقييم.

وهذا النظام يخول للمركز الوطني لتطوير المناهج الدخول على خط تطوير امتحان الثانوية العامة وإدارته. وهذا واقع جديد جاء في سياق الإصلاح الشامل للنظام التعليمي في الأردن. وأمام هذا الوضع، ينبغي تحديد الأدوار والمسؤوليات لكل الشركاء في تطوير الامتحان وإدارته.

2. الغرض من امتحان الثانوية العامة:

الغرض الرئيس لامتحان الثانوية العامة كما جاء في قانون التربية والتعليم هو منح شهادة الدراسة الثانوية للناجحين. أي أن غرضه الأساسي هو تقييم حصيلة وخبرات التعلم في نهاية المرحلة الثانوية ومنح من يحقق متطلبات ومعايير النجاح شهادة الدراسة الثانوية العامة. وهذا يتطلب تحديد متطلبات ومعايير النجاح ومن ضمنها علامة النجاح. وتحديد هذه المتطلبات من صلاحيات مجلس الامتحان وتكون في ضوء خطة دراسية مقرة من مجلس التربية.

ومن ناحية أخرى فقد تم استخدام الامتحان ولا زال يستخدم لأغراض القبول في الجامعات الأردنية الرسمية والخاصة، وتشير السياسة العامة لقبول الطلبة الأردنيين في الجامعات الأردنية لمرحلة البكالوريوس

[20] مشروع قانون الموازنة العامة للسنة المالية 2024

[21] نظام المركز الوطني لتطوير المناهج رقم 33 لسنة 2017 وتعديلاته

للعام الجامعي 2024/2023 الصادرة استناداً لأحكام المادة (6) البند (6) من قانون التعليم العالي والبحث العلمي رقم (17) لسنة 2018 وتعديلاته، بموجب قرار مجلس التعليم العالي رقم (308) تاريخ 6 /2023/7 إلى أنه[22] يتم قبول الطلبة في الجامعات الأردنية، في برنامج البكالوريوس، وفق خياراتهم، بحسب تسلسل معدلاتهم في شهادة الثانوية العامة، أو ما يعادلها، ووفق حد أدنى للمعدل في الكليات والتخصصات المختلفة".

أي أن المعدل في امتحان الثانوية العامة هو المعيار الوحيد للقبول في الجامعات الأردنية، ونظرًا لأن القبول مشروط بالحصول على شهادة الدراسة الثانوية العامة الأردنية أو ما يعادلها فإن النجاح (وجود علامة النجاح) في امتحان الثانوية العامة يلزم لغرض الالتحاق بالجامعات الأردنية.

3. متطلبات النجاح في الامتحان عبر السنين:

تُمنح شهادة الدراسة الثانوية العامة في أنواع التخصص المختلفة للطالب الذي يحقق متطلبات النجاح في التخصص، ويقوم مجلس الامتحان[23] بتحديد المباحث التي يتقدم لها الطلبة وأوزانها وعلامة النجاح. وقد جرى تغيير في متطلبات النجاح عبر السنوات الماضية، ويبين الجدول 4 متطلبات النجاح في سنوات مختلفة.

جدول 4: متطلبات النجاح في امتحان الثانوية العامة في سنوات مختلفة.

السنة	عدد المباحث التي يتم التقدم لها	نسبة النجاح من النهاية العظمى	النهايات العظمى للمباحث	متطلبات النجاح	ملاحظات
1964	9	40%	100–300	6 مباحث	نسبة النجاح 25% في البقية
2002	10	50%	80–200	10 مباحث	
2011	10	50%	100–200	10 مباحث	يوجد لبعض المباحث مستويان
2021	8	40%	200	8 مباحث	نسبة النجاح 50% لأغراض الجامعات

وبفحص البيانات في الجدول 4 يتضح الآتي:

- ✓ تفاوت في عدد المباحث وموضوعاتها من فترة إلى أخرى.
- ✓ وجود أكثر من مستوى لبعض المباحث في فترة معينة مما يزيد في عدد الامتحانات.

[22] السياسة العامة لقبول الطلبة الأردنيين في الجامعات الأردنية بموجب قرار مجلس التعليم العالي رقم 308 تاريخ 6/7/2023

[23] مجموعة من القيادات التربوية والخبراء برئاسة وزير التربية والتعليم

- ✓ عدم استقرار أوزان المباحث كما تعكسها النهايات العظمى لها.
- ✓ عدم استقرار علامات النجاح (النهايات الصغرى) حيث تختلف من فترة إلى أخرى.
- ✓ عدم استقرار متطلبات النجاح وكيفية حساب المعدل.

4. نظرة عامة على نتائج امتحان شهادة الدراسة الثانوية العامة:

يبين الجدول 5 أعداد المسجلين والمتقدمين والناجحين في امتحان شهادة الدراسة الثانوية العامة.

جدول 5: أعداد المسجلين والغائبين والمتقدمين والناجحين والراسبين في امتحان شهادة الدراسة الثانوية العامة

السنة	عدد المسجلين	الغائبون والمحرومون		المتقدمون		الراسبون من المتقدمين		الناجحون من المتقدمين	
		العدد	النسبة	العدد	النسبة	العدد	النسبة	العدد	النسبة
2016	**127872**	**47125**	**36.85%**	**80747**	**63.15%**	**47687**	**59.06%**	**33060**	**40.94%**
2021	**207311**	**21505**	**10.37%**	**185806**	**89.63%**	**72758**	**39.16%**	**113048**	**60.84%**

يلاحظ من الجدول 5 تزايد في عدد المسجلين للامتحان بصورة كبيرة جدًا في العام 2021 عما كان عليه في عام 2016. حث كانت نسبة الزيادة (62.12%)، وهذا يشكل تحديًا كبيرًا على إدارة الامتحانات والاختبارات من جهة وصعوبة في توفير اللوجستيات المناسبة من جهة أخرى. كما يزيد من كلفة الامتحان الذي يعتمد على التمويل الحكومي ورسوم الامتحان. ويلاحظ كذلك، انخفاضًا كبيرًا في عدد الغائبين والمحرومين في عام 2021 عما كانت عليه في عام 2016 بالرغم من الزيادة الكبيرة في عدد المسجلين، فقد انخفضت نسبتهم من 37% إلى حوالي 10%، وهذا شيء إيجابي يشير إلى تحسن في استعدادات الطلبة للامتحان أو انخفاض رهبة الامتحان والخوف من الفشل فيه. كما ارتفعت نسبة نجاح المتقدمين من 41% إلى 61%. وربما يكون من الأسباب الكامنة وراء النسبة العالية من الفشل في 2016 هو الفشل في مبحث واحد على الأقل حتى وإن كان خارج اهتمامات الطالب وميوله، ففي عام 2016 وحده، هناك ما مجموعه 94812 طالبًا وطالبة بين غائب وغير محقق لمتطلبات النجاح ونسبتهم (74.15%) من إجمالي عدد المتقدمين مقابل (94263) طالبًا وطالبة ونسبتهم (45.47%) للعام 2021 م. فإن هذه الأرقام صادمة وتزيد من حجم القلق النفسي والهدر التربوي ولها تبعات اجتماعية واقتصادية، وتثير الشكوك حول مصداقية الامتحان وبوجود تشدد

في الامتحان للعام 2016 م مقارنة بالعام 2021م سواء من ناحية المحتوى وصعوبة الأسئلة أو من ناحية الإجراءات.

5. درجة صعوبة أسئلة امتحان شهادة الدراسة الثانوية العامة:

بالاعتماد على نتائج الطلبة في الدورة الشتوية 2016 (السوالمه وآخرون، 2017)، تبين أن كل امتحان في مباحث التربية الإسلامية، واللغة العربية (المستوى 3)، واللغة الإنجليزية (المستوى 3)، والرياضيات (المستوى 3)، والفيزياء (المستوى 3)، يتضمن خمسة أسئلة من النوع المقالي، وحسب النظرية التقليدية في القياس، تعرف درجة صعوبة السؤال المقالي بنسبة ما يحصله الطلبة في المتوسط من العلامة الكاملة للسؤال. أي أن صعوبة السؤال هي الوسط الحسابي لعلامات الطلبة على السؤال مقسومًا على العلامة الكاملة للسؤال. فمثلاً، إذا كان الوسط الحسابي لعلامات الطلبة على سؤال علامته الكاملة 15 هو 12 فإن درجة صعوبته 0.80. ويكون الامتحان أكثر فاعلية عندما تتوزع صعوبة الأسئلة ضمن المدى 0.30 إلى 0.80. ويبين الجدول 6 درجات الصعوبة لأسئلة مباحث مختارة في الدورة الشتوية 2016.

جدول 6: درجات الصعوبة لأسئلة مباحث مختارة في الدورة الشتوية 2016

المبحث / السؤال	التربية الإسلامية	لغة عربية/ م 3	لغة إنجليزية/ م 3	رياضيات علمي/ م3	فيزياء علمي/ م3
الأول	**0.74**	**0.72**	**0.43**	**0.54**	**0.52**
الثاني	**0.69**	**0.64**	**0.49**	**0.55**	**0.51**
الثالث	**0.72**	**0.49**	**0.37**	**0.56**	**0.51**
الرابع	**0.67**	**0.36**	**0.33**	**0.43**	**0.49**
الخامس	**0.61**	**0.55**	**0.27**	**0.23**	**0.44**

يلاحظ من الجدول 6 أن درجات الصعوبة في كل الامتحانات لا توازي مستويات قدرات المفحوصين، وتفتقر للمدى الواسع في الصعوبة، فامتحان التربية الإسلامية يفتقر لأسئلة تتحدى الطلبة ذوي القدرات العالية وتميل أسئلته للسهولة، وبقية الامتحانات تفتقر للأسئلة التي تناسب ذوي القدرات المتدنية وتميل أسئلتها نحو الصعوبة خاصة في اللغة الإنجليزية والرياضيات. وتفسر هذه البيانات تدني النتائج في اللغة الإنجليزية والرياضيات والفيزياء. ولم يتم توفير تقديرات لمعاملات ثبات الامتحانات في المباحث المختلفة، ويقصد بثبات الاختبار دقة الدرجات ومدى خلوها من أخطاء القياس.

6. امتحان شهادة الدراسة الثانوية العامة في الدراسات والتقارير:

تناولت العديد من الدراسات (الخطايبة، 2002؛ الخوالدة وآخرون، 2007؛ السوالمه وآخرون، 2017؛ عبد العزيز، 2011؛ عدس وآخرون، 2001) صدق المحتوى للامتحان في مباحث بعينها. وتجدر الإشارة إلى أن أقرب دراسة تعود إلى العام 2017، وعليه فإنه لا يعول كثيرا على نتائج تلك الدراسات هذه الأيام حيث كانت الغالبية العظمى للأسئلة في كثير من المباحث وقت إجراء تلك الدراسات إنشائية بينما هي اختيار من متعدد حاليا. ويقصد بصدق المحتوى للامتحان مدى تمثيل أسئلته لمجالات محتوى المبحث من جهة ونتاجات التعلم المنوي قياسها من جهة أخرى. ويتحقق صدق المحتوى عندما يطابق الاختبار جدول المواصفات. وقد أكدت تلك الدراسات أن التركيز في الامتحانات على المستويات المعرفية الدنيا (التذكر والفهم) وقليلا ما تركز على التطبيق ونادرا ما تركز على مهارات التفكير العليا من تحليل وتركيب وتقويم.

ويقصد بالصدق التنبؤي لامتحان ما قدرته وفاعليته على التنبؤ بنتيجة أو سلوك مستقبلي معين. وينبغي أن يتصف الامتحان المستخدم لغرض القبول الجامعي بالقدرة على التنبؤ بالنجاح والتحصيل الجامعي. وفيما يتعلق بامتحان الثانوية العامة في الأردن، تشير مجموعة من الدراسات (الثوابية، 1994؛ جرادات، 2003؛ الرشدان، 2002؛ غرايبة،1991؛ اللطايفة،1998؛ فريحات، 2002؛ القوابعة، 2002) إلى ضعف القدرة التنبؤية لمعدل الثانوية العامة في التحصيل الأكاديمي الجامعي، وأن معدل الثانوية العامة لا يصلح معيارا منفردا لقبول الطلبة في الجامعات.

ويشير الصمادي وآخرون (2010) إلى ضعف العلاقة بين معدل الثانوية العامة وامتحان الكفاءة الجامعية في معظم التخصصات. ويشير العناتي (2014) إلى ارتباط ضعيف جدا (r= 0.07) مع الأداء الوظيفي. وبمقارنة القدرة التنبؤية لمجموعة من الاختبارات بالتحصيل الجامعي في الأردن احتل امتحان الثانوية العامة الأردني المرتبة الثالثة بعد الامتحانات العامة السويسرية IB التي احتلت المرتبة الأولى والامتحانات البريطانية IGCSE التي احتلت المرتبة الثانية ومتقدما على الامتحانات الامريكية SAT التي جاءت في المرتبة الرابعة (ظاظا والمشايخ، 2016). وبالرغم من صعوبة تفسير هذه النتائج وصعوبة تحديد العوامل المؤثرة فيها فإنها تؤشر إلى ضرورة عدم اقتصار القبول الجامعي على معدل الثانوية العامة وأن تكون هناك معايير أخرى معه.

ويصف تقرير SABER[24] كل ما يتعلق بامتحان الثانوية العامة من حيث الغرض، والجهة المنفذة، والتمويل، وإجراءات التطبيق، والتصحيح، والإبلاغ عن النتائج، وغيرها. وقد صنف الامتحان في المستوى

[24] SABER Jordan Report, World Bank, 2014.

الثالث "منجز" (Established)، أي يحقق الحد الأدنى من المعايير. وقد سجلت بعض الملاحظات على الامتحان التي لا زالت قائمة لغاية الآن، تتعلق هذه الملاحظات بالاعتماد على آلية واحدة لضبط جودته، وهي المراجعة الداخلية. أي أنه يفتقر للمراجعة الخارجية، والاعتماد الخارجي، والتدقيق الخارجي، والتجريب الميداني. وهناك دعوة لمواصلة العمل المستمر لدمج مهارات الاقتصاد المعرفي الجديدة في امتحان التوجيهي، وإيجاد خيارات أمام الذين لا يكون أداؤهم جيدا في الامتحان حتى يتمكنوا من الاستمرار في التعليم أو التدريب وتنمية امكاناتهم وقدراتهم على المساهمة المجتمعية، وضمان عدالة الامتحان من خلال تنفيذ تدابير وقائية تحد من السلوكيات غير المرغوبة التي تصاحب تنفيذ الاختبار.

كما أكد التقرير الإقليمي حول خريطة التقييم في الدول العربية[25]، على وجود سياسة واضحة للامتحان وقيادة قوية وتمويل منتظم وهياكل تنظيمية قوية وموارد بشرية فعالة. وكانت الملاحظات على الامتحان تتعلق بالاعتماد على المراجعة الداخلية لضبط الجودة. والافتقار للمراجعة الخارجية، والاعتماد الخارجي، وعدم توفر التجريب الميداني، وعدم وضوح آلية رصد النتائج ومن المحتمل استخدامها على نحو غير ملائم، وعدم وجود وثيقة اطارية تشرح ما الذي يقيسه الامتحان، وعدم توفر تقارير حول مكامن القوة والضعف في الأداء على الامتحان.

ورصد تقرير مسح تقييمات الطلبة في الأردن[26] بالاعتماد على بعض التعليقات في الصحف والمداولات في مجموعات تركيز من طلبة ومعلمين ومشرفين وأولياء أمور وأساتذة جامعات، مجموعة من الملاحظات بخصوص امتحان الثانوية العامة:

- له تأثير كبير على الطلبة، ويعطونه الكثير من الاهتمام على حساب التفكير الناقد.
- 89% من المشاركين في مجموعات التركيز يعتقدون بأنه مهم جدا.
- تقديره من قبل المشاركين في مجموعات التركيز هو 9 على تدريج من 1–10 لمتصل الأهمية، وهو الأكثر أهمية بين كل التقييمات الأخرى.
- يتضمن امتحان الثانوية العامة الكثير من الموضوعات التي تتطلب الكثير من الحفظ.
- يسبب امتحان الثانوية العامة العديد من المشكلات والضغوط النفسية، وتم تشبيهه من بعض المشاركين في مجموعات التركيز بفلم رعب يشاهده جميع الناس ويعيشونه.
- لا يشجع مهارات التفكير العليا.

[25] أعد التقرير عدنان الأمين بتكليف من مكتب اليونسكو الإقليمي للتربية في الدول العربية والمنظمة العربية للتربية والثقافة والعلوم

[26] This report is a product of World Education, Inc. (WEI) in collaboration with National Center for Human Resources Development (NCHRD) in Jordan

7. آخر التحديثات في خطة امتحان شهادة الثانوية العامة:

في آب / أغسطس 2021، تم تشكيل لجنة تطوير امتحان الشهادة الثانوية العامة[27] لمراجعة "التوجيهي" وتطويره. وقد سلّط تقرير اللجنة الضوء على الحاجة إلى استمرار إجراء امتحان "التوجيهي" وأكد على الغرض المزدوج للامتحان والمتمثل في منح شهادة إنهاء الدراسة في المرحلة الثانوية والقبول في مؤسسات التعليم العالي. وقد حددت اللجنة مجموعة من المبادئ لتوجيه عملية تطوير الامتحان، تمثلت بالآتي:

- الامتحان ضروري ويجب استمراره.
- الاتساق مع إصلاح النظام التعليمي، والانسجام مع الاقتصاد المعرفي ومهارات القرن الحادي والعشرين.
- أن يقوم على العدالة، والشفافية، وتكافؤ الفرص، ومجموعة أخرى من القيم.
- استخدام نظرية الاستجابة للفقرة في بناء الامتحان ومعايرة أسئلته وتحليل نتائجه.
- استخدام سلم موحد للعلامات scale score لإعطاء معنى للنتيجة وتحقيق درجة معقولة من العدالة والقابلية للمقارنة ضمن المباحث وعبر الدورات والسنوات.
- تقليل عدد المباحث وطول مدة الامتحانات مع تغليب الكيف على الكم.
- تغطية شاملة لنتاجات التعلم باستخدام أدوات تقييم متنوعة.
- معالجة التبعات المجتمعية للامتحان من توتر وضغط نفسي لدى الطلبة وأسرهم، وتخفيف حالات الشعور بالظلم.
- إعطاء مزيد من المرونة في التقدم لموضوعات الامتحان وعدد مرات التقدم لها.
- تحسين الخصائص السيكومترية للاختبار، وضمان درجة عالية من الصدق والثبات لنتائجه، وتحسين القدرة التنبؤية بالنجاح في التعليم الجامعي.

وتماشيا مع هذه المبادئ، اقترحت اللجنة إجراء الإصلاحات الآتية على ترتيبات الامتحان الحالية:

- دمج بيانات التقييم المدرسي في إجراءات القبول بالجامعة.
- إعادة هيكلة جدول الامتحانات بحيث يتم إجراء الامتحانات على مدى عامين.
- إنشاء بنوك أسئلة عالية الجودة لإتاحة قدر أكبر من المرونة في تحديد مواعيد الامتحانات.
- نقل مسؤولية وضع الامتحانات وإدارتها إلى المركز الوطني لتطوير المناهج.
- تطبيق نظام قياس يضمن عدالة نتائج التقييم واتساقها.

[27] قرار وزير التربية رقم 29788/120/14 تاريخ 2021/8/12 بالاستناد إلى قرار مجلس التربية رقم 2021/127 في الجلسة رقم 5 تاريخ 3/8/2021

- مراجعة سياسة القبول في الجامعات للحد من الاعتماد على نتائج "التوجيهي".

وهذه التوصيات مرحب بها ومن شأنها أن تساهم في تحسين مرونة امتحان "التوجيهي" وجودته، لكن تبقى هناك أسئلة حول ما إذا كانت الإصلاحات المقترحة ستعالج معدل الرسوب المرتفع في امتحان "التوجيهي"، وتضمن أن يقيّم امتحان "التوجيهي" بشكل مناسب الكفاءات التي يحتاجها اقتصاد المعرفة والمهارات التقنية العالية المستوى المطلوبة من القوى العاملة بما فيها مهارات القرن الحادي والعشرين.

وعلى ضوء هذه التوصيات تم إقرار خطة دراسية جديدة للمرحلة الثانوية. بمقتضى هذه الخطة في المسار الأكاديمي، يدرس الطالب في الصف الحادي عشر سبعة مباحث، وبعد النجاح فيها مدرسيا، يستطيع التقدم للجزء الأول من امتحان الثانوية العامة في أربعة مباحث منها، هي: التربية الإسلامية، واللغة العربية، واللغة الإنجليزية، وتاريخ الأردن. ويدرس الطالب في الصف الثاني عشر المهارات الرقمية وست مباحث أخرى، وبعد النجاح فيها مدرسيا، يستطيع التقدم للجزء الثاني من امتحان الثانوية العامة في أربعة مباحث يختارها حسب حقل التخصص الذي يرغب بدراسته في الجامعة وعلى النحو الآتي:

1. الحقل الصحي: كيمياء وعلوم حياتية ولغة إنجليزية متقدم + مبحث اختياري (من بين 11 مبحث)
2. الحقل الهندسي: رياضيات وفيزياء+ مبحثان اختياريان على أن يكون أحدهما مبحث علمي.
3. حقل العلوم والتكنولوجيا: رياضيات+ 3 مباحث اختيارية منها مبحثان علميان.
4. حقل اللغات والعلوم الاجتماعية: عربي تخصص وإنجليزي متقدم + مبحثان اختياريان على أن يكون أحدهما في العلوم الإنسانية والاجتماعية.
5. حقل القانون والعلوم الشرعية: عربي تخصص وتربية إسلامية تخصص + مبحثان اختياريان على أن يكون أحدهما في العلوم الإنسانية والاجتماعية.
6. حقل الأعمال: رياضيات الأعمال وثقافة مالية وإنجليزي متقدم + مبحث اختياري (من بين 11 مبحث)

ويحتسب المعدل للطالب لغايات القبول الجامعي بتخصيص 30% من المعدل لمباحث الثقافة العامة المشتركة في الصف الحادي عشر (الجزء الأول من الامتحان) و70% للمباحث الاختيارية الأربعة في الصف الثاني عشر (الجزء الثاني من الامتحان). وأظهر التطبيق الفعلي للخطة مجموعة من السلبيات في الخطة تتعلق بعدم مناسبة الحصص للمباحث المختلفة، والتفاوت في توزيع المباحث على الصفين (11 في صف 11، 15 في صف 12)، والتفاوت غير المبرر في أوزان المباحث في المعدل (10% للغة العربية مقابل

17.5% لعلم النفس والاجتماع مثلًا)، وعدم التوازن بين المباحث العلمية والإنسانية في كل من الصفين، وعدم إعطاء الرياضيات ما تستحقه من اهتمام، وزيادة كبيرة في عدد الامتحانات التي يجب بنائها وتطويرها (18 امتحان). الأمر الذي يتطلب إعادة النظر في الخطة الدراسية وتطويرها بما يخدم النظام التعليمي ويأخذ حاجات الطلبة وقدراتهم بالاعتبار.

8. الممارسات الدولية في الامتحانات العامة

تتفاوت الممارسات الدولية فيما يتعلق بالامتحانات العامة في نهاية مرحلة الدراسة الثانوية العامة وبداية مرحلة الانتقال للتعليم العالي. فهناك دول تجري امتحانات نهاية المرحلة الثانوية التي تتولاها وزارة التربية والتعليم أو مراكز مستقلة وتهدف إلى الوقوف على مستوى التعلّم الذي تحقق للطلبة بعد الانتهاء من المرحلة الثانوية، ويمنح الطالب بناءً على نتيجته شهادة تفيد إتمامه هذه المرحلة، كما هو معمول في فرنسا، ومصر على سبيل المثال. وتلجأ بعض الدول إلى امتحانات نهاية المرحلة الثانوية والأداء المدرسي معا، وتهدف هذه السياسة إلى التغلب على جوانب القصور الناتجة عن اعتماد القرار المتعلق بتحديد مستقبل الطالب على مجموعة من الامتحانات التي تعقد في فترة زمنية محددة وفي ظروف قد تكون غير مناسبة لبعض الطلبة وتحول دون إظهار قدراتهم الواقعية، كما هو معمول في أستراليا. وتعتمد دول أخرى على امتحانات القبول للتعليم الجامعي التي تتولاها جهات مركزية مستقلة في الدولة وتهدف إلى الوقوف على قدرات الطالب العامة واستعداده للتعليم الجامعي كالقدرة على التفكير والاستدلال بأنواعه كما هو معمول في اليابان والسعودية. وتعتمد بعض الدول على امتحانات القبول للتعليم الجامعي والأداء المدرسي معا. وتهدف هذه السياسة إلى إحكام الربط بين مخرجات التعليم الثانوي والمتطلبات الضرورية كمدخلات للتعليم الجامعي كما هو معمول في الولايات المتحدة.

وهذه الاختلافات بين الدول تدلل على أنه لا توجد وصفة معينة تناسب أوضاع جميع الدول. وهناك أسباب متعددة يمكن أن تفسر هذه الاختلافات، مثل مستوى مركزية أنظمة التعليم فيها ونموذج التقييم المعتمد فيها.[28] وتهيمن ثلاثة نماذج تقييم منفصلة على التقييم في العملية التعليمية التعلمية: الأول هو النموذج السيكومتري، الذي يهتم بقياس سمة واحدة، ويهتم بالصدق والثبات، ويركز بشكل كبير على الفقرات التي تميز بين الطلبة، ويصنف الطلبة بناءً على درجاتهم، وتكون الدرجات ذات توزيع طبيعي. والثاني هو النموذج القائم على النتائج، ويهتم بالكفاءة أو التمكن من مجموعة محددة من المهارات، ولا يفترض التوزيع الطبيعي للدرجات،

[28] https://unesdoc.unesco.org/query?q=Series:%20%22UNESCO-IESALC%20policy%20insight%22&sf=sf:*

ويميل إلى الارتباط بقرارات النجاح/الرسوب. والثالث هو التقييم القائم على المنهاج، ويهتم في المقام الأول بتقييم مدى اكتساب الطالب للمعرفة والمهارات المتضمنة في المنهاج من أجل التقدم إلى المستوى التالي من نظام التعليم، فهو يركز بشدة على صدق المحتوى، ويعطي نتائج على شكل درجات، ولا يفترض التوزيع الطبيعي للدرجات. وهو نهج التقييم عالي المخاطر السائد في معظم أنظمة التعليم.

وتستخدم الاختبارات العامة في العديد من الأغراض مثل إصدار الشهادات، والاختيار، وتحفيز الطلبة، ومراقبة أنشطة المدارس، وتوفير المعلومات التي يمكن الاستفادة منها في إدارة النظام التعليمي، ومساءلة المدارس أو المعلمين عن مستويات إنجاز الطلبة، وإضفاء الشرعية على العضوية في المجتمع العالمي. وقد حدد البنك الدولي في إصداره المعنون "مراجعة الامتحانات العامة 2020"[29] سمات محددة عالميا ومتفق عليها فيما يخص الامتحانات العامة، أبرزها:

- يتم وضعها وضبطها من قبل جهة رسمية وطنية غير المدارس التي يلتحق بها الطلبة.
- السلطة القائمة على إدارة الامتحانات العامة هي جهة حكومية أو مؤسسة وطنية.
- تعتمد الامتحانات العامة على مناهج دراسية مقررة أو مواضيع دراسية محددة، مثل اللغات، والرياضيات، والعلوم الطبيعية، والدراسات الاجتماعية، وتكنولوجيا المعلومات والاتصالات، والثقافة الدينية.
- يتم تطبيقها في بيئة اختبارية موحدة وظروف خاضعة لدرجة عالية من الرقابة والضبط.
- يحصل الطالب كنتيجة للأداء في الامتحان على درجة كلية أو درجة منفصلة في كل مادة يتقدم لها.

كما أطلق معهد اليونسكو الدولي للتعليم العالي في أمريكا اللاتينية ومنطقة البحر الكاريبي IESALC (UNESCO, 2024) رؤيته السياسية الرابعة للدول التي تتطلب في تشريعاتها تنظيم امتحانات وطنية للطلبة لغرض القبول في التعليم العالي، وتطبق هذه الامتحانات عادة، إما عند إكمال المرحلة الثانوية أو عند الالتحاق بالتعليم العالي، ويمكن استعمالها كوسيلة محايدة وموحدة لتقييم المرشحين للقبول، ولتعزيز الشفافية والمساءلة العامة، ولضمان تلبية جميع الطلبة لمعايير أكاديمية محددة عند الالتحاق بالتعليم العالي. ويمكن استخدام هذه الامتحانات أيضًا لمراقبة أداء نظام التعليم الوطني. ويؤخذ على هذه الامتحانات أنها مكلفة وتنفيذها صعب وقد يتطلب جهودًا وطنية كبيرة، وأنها تحد من استقلالية مؤسسات التعليم العالي في عملية القبول. علاوة على

29 World Bank, Public examinations examined, 2020

ذلك، قد تؤدي الامتحانات الوطنية عن غير قصد إلى إدامة عدم المساواة الاجتماعية والاقتصادية، حيث يتمتع الطلبة من خلفيات مميزة في كثير من الأحيان بإمكانية أكبر للوصول إلى الموارد والدعم للتحضير للاختبار.

وفيما يلي أمثلة على الممارسات المتبعة في بعض دول العالم:

- **التجربة السنغافورية**

تعتمد سنغافورة امتحان (GCE, O–Level)، ويعقد سنويًا ويقوم على المناهج المدرسية، ويستخدم لغرض منح شهادة إكمال المرحلة الثانوية، التي تُعدّ بمثابة مؤهل أكاديمي معترف به لمتابعة التعليم العالي والبحث عن عمل، ويتراوح عدد الموضوعات التي يتقدم لها الطالب في الامتحان بين 6 موضوعات بحد أدنى و9 موضوعات بحد أعلى لمن يرغب بزيادة عددها، والموضوعات الإجبارية هي اللغة الإنجليزية، والرياضيات، واللغة الأم للمواطن (يجوز للأجنبي اختيار لغة أخرى)، والعلوم الإنسانية مجتمعة، والعلوم (تجميع موضوعين من بين الفيزياء والكيمياء والأحياء)، وموضوع اختياري يختاره الطالب من عدة موضوعات.

ويتم تصنيف أداء المتقدمين للامتحان في كل موضوع في تسع فئات مرتبة تنازليًا من الأعلى إلى الأدنى (A1= 1،A2= 2، B3= 3، B4= 4، C5= 5، C6= 6، D7= 7، E8= 8 ، F9 =9)، ولاجتياز أي موضوع، يجب على الطالب الحصول على الفئة C6 على الأقل، أي الحصول على الدرجة 6 أو5 أو4 أو3 أو2 أو 1. وتحسب للمتقدم علامة إجمالية تعرف باسم (L1R5) وذلك بجمع درجاته في 6 موضوعات أحدها في اللغة و5 موضوعات أخرى ذات صلة. وتُعتمد الدرجة الإجمالية في القبول لمؤسسات التعليم العالي. وهناك تحول تدريجي إلى الامتحانات الإلكترونية، والأسئلة خليط بين أسئلة الاختيار من متعدد والإجابات الإنشائية القصيرة.

- **التجربة الكورية**

تستخدم كوريا الجنوبية اختبار القدرة الدراسية للكلية CSAT، الذي يطوره ويديره المعهد الكوري للمناهج والتقييم (KICE) سنويًا في ظل قيود صارمة، ومدة الامتحان 8 ساعات متواصلة يتخللها استراحات قصيرة. وهو مصمم لتقييم القدرة الأكاديمية المطلوبة للقبول الجامعي بالاعتماد على المناهج المدرسية. وجميع الأسئلة في الموضوعات المختلفة من نوع الاختيار من متعدد عدا 9 أسئلة في الرياضيات ذات إجابة قصيرة، ويتضمن الامتحان 6 موضوعات هي الرياضيات والتاريخ واللغة الكورية واللغة الإنجليزية وموضوعين اختياريين إما إنسانين أو علميين. ويستخدم المنحنى التساعي stanine curve في إعطاء الدرجات باستثناء اللغة الإنجليزية

والتاريخ حيث يستخدم سلم العلامات المطلق. ولا يوجد نجاح ورسوب في الامتحان، حيث ترصد لكل مادة في كشف العلامات درجة، ورتبة مئينية، وعلامة معيارية.

- **التجربة البولندية**

يتناول امتحان التخرج في بولندا ثلاثة اختبارات هي اختبار اللغة البولندية واختبار الرياضيات واختبار اللغة الأجنبية. ويتعين على جميع الطلبة اعتبارا من عام 2023 إظهار مهارات ومعرفة تتجاوز بكثير ما كانت عليه في السابق. فمنهاج اللغة البولندية يتضمن حاليًّا 50 عنوانًا مقابل 15 عنوانًا اعتادوا عليها. والامتحانات الكترونية، وتفتح الإمكانات الإلكترونية الباب أمام تنويع أسئلة الرياضيات من خلال استخدام محرر معادلات متقدم يدعم جميع أنواع العمليات التي تلزم في الامتحان، والواجبات المنزلية الإلكترونية، واختبار مهارتي الاستماع والتحدث، وتجدر الإشارة إلى أن علامة النجاح هي الحصول على 30% من النقاط المخصصة لكل اختبار.

- **التجربة الفنلندية**

يتم إجراء امتحان شهادة الثانوية العامة الفنلندي في نهاية التعليم الثانوي للتأهل للالتحاق بالجامعة. ويُطلب من كل ممتحن اجتياز خمسة اختبارات منها الجزء الإلزامي الوحيد هو اختبار اللغة الأم، وعلى الطالب اختيار أربع مواد أخرى هي: لغة محلية أخرى، لغة أجنبية، الرياضيات، وموضوع اختياري يتم اختياره من مباحث: الدين، والتربية على الأخلاق، والتاريخ الأخلاقي، والفلسفة، وعلم النفس، والتاريخ، والتربية المدنية، والفيزياء، والكيمياء، والأحياء، والجغرافيا، والتثقيف الصحي.

ويتم إجراء الاختبار في الربيع، وفي الخريف في المدارس، ويقوم المعلمون بتصحيح الأوراق وإرسال الأوراق التي تم تقييمها إلى المجلس الوطني الذي يقوم بعد ذلك بإعادة تقييم كل ورقة، وتختلف درجة كل اختبار باختلاف الموضوع، ويتم تصنيف الأداء وفقًا للتوزيع الطبيعي إلى سبع درجات لفظية هي "غير معتمد"، "موافق عليه"، "موافق عليه بكل سرور"، "موافق عليه مع الثناء"، "موافق عليه مع الكثير من الثناء"، "موافق عليه مع الثناء الاستثنائي"، و"المحمود". بشكل عام، يلزم الحصول على الدرجة "موافق عليه" على الأقل لاجتياز الاختبار، وتقابل هذه الدرجات النسب الآتية على الترتيب: 5%، 15%، 20%، 20%، 20%، 15%، 5%. أي ينجح 95% من الطلبة، وهناك نظام تعويضي يُمكِّن الذين يحصلون على درجة "غير معتمد" من النجاح إذا كان أداؤهم في الموضوعات الأخرى جيدًا.

التجربة المصرية

يتم إجراء امتحان الثانوية العامة المصري بعد إنهاء الصف الثالث الثانوي ويكون في ثلاث شعب هي شعبة العلوم وشعبة الرياضيات والشعبة الأدبية، ويدخل في المجموع 5 مواد كما هو مبين في الجدول 7.

الجدول 7: المواد التي تدخل في المجموع في امتحان الثانوية العامة المصرية

شعبة العلوم		شعبة الرياضيات		الشعبة الأدبية	
المبحث	العلامة	المبحث	العلامة	المبحث	العلامة
اللغة العربية	80	اللغة العربية	80	اللغة العربية	80
اللغة الأجنبية	60	اللغة الأجنبية	60	اللغة الأجنبية	60
الأحياء	60	الرياضيات	60	التاريخ	60
الكيمياء	60	الكيمياء	60	الجغرافيا	60
الفيزياء	60	الفيزياء	60	الإحصاء	60
المجموع	**320**	المجموع	**320**	المجموع	**320**

- **التجربة الإماراتية**

يعتمد قبول الطلبة في الجامعات الإماراتية، وفي العديد من الجامعات العالمية على النتائج التي يحصلون عليها في امتحانات الثانوية العامة، وتجرى هذه الامتحانات في الإمارات في ثلاثة فصول دراسية يتكون منها العام الدراسي، أي يتوزع حجم المادة التي يختبر فيها الطالب على الفصول الثلاثة. ويوجد ثلاثة مسارات هي المسار العام والمسار المتقدم ومسار النخبة، ويختبر الطلبة في المسارات الثلاثة في ثمانية مواد العلامة الكاملة لكل منها 100، المواد السبعة الأولى في المسارات الثلاثة هي اللغة العربية، واللغة الإنجليزية، والتربية الإسلامية، والدراسات الاجتماعية، والرياضيات، والفيزياء، والكيمياء، أما المادة الثامنة فهي العلوم الصحية في المسارين العام والمتقدم والعلوم الحياتية في مسار النخبة، ويحتسب المعدل للطالب بإعطاء وزن 35% للفصل الأول (25% للامتحانات+ 10% للعلامة المدرسية)، ووزن 30% للفصل الثاني (20% للامتحانات+10% للعلامة المدرسية)، ووزن 35% للفصل الثالث (25% للامتحانات+10% للعلامة المدرسية)، أي أن هناك ما مجموعة 30% من المعدل للعلامات المدرسية.

من خلال مراجعة أنظمة التقييم السابقة يمكن رصد النقاط الأساسية الآتية:

- الامتحانات العامة هي امتحانات تخرج Exit Exam لمنح شهادة أو امتحانات قبولِ Admission Exam للجامعات، ويجوز أن يكون الامتحان المعتمد هو امتحان قبول وتخرج في آن واحد.
- عدد المباحث المتضمنة في الامتحانات تتراوح بين (2–8) وما يحتسب في المجموع أو المعدل هو 6 مباحث على الأكثر، باستثناءات قليلة (الإمارات مثلًا).
- تتضمن بعض الأنظمة مباحث اختيارية يختارها الطلبة وفقًا لرغباتهم وقدراتهم.
- لا توجد علامة نجاح في امتحانات القبول، ونسب النجاح في امتحانات التخرج تزيد عن 95%.
- تستخدم معظم الامتحانات الدولية سلمًا موحدًا للعلامات (تدريج قياسي)، مما يجعل النتائج قابلة للمقارنة العادلة عبر الموضوعات وعبر السنوات دون التأثر باختلاف صعوبة الأسئلة.
- توجد معايير متعددة للقبول الجامعي في كثير من الأنظمة.
- تشرف على الامتحان وتديره مراكز مستقلة في العديد من دول العالم.
- تركز الامتحانات العامة على المباحث الأساسية المتمثلة باللغة الأم واللغة الأجنبية والرياضيات.
- تحظى العمليات العقلية العليا ومهارات التفكير وإنتاج المعرفة بتركيز واضح في كثير من الأنظمة.
- تعتمد بعض الدول على تقديم الاختبارات إلكترونيًا، وتعقد الامتحان أكثر من مرة واحدة في السنة.
- تتضمن الامتحانات العامة أسئلة الاختيار من متعدد بصورة كبيرة.

9. ملاحظات مصاحبة لامتحان شهادة الدراسة الثانوية العامة في الأردن:

- كثرة عدد الموضوعات التي يمتحن فيها الطالب، وطول مدة الامتحان.
- المعدل للامتحان هو المعيار الوحيد للقبول الجامعي أي أن الامتحان مصيري.
- نسب النجاح في الامتحان متدنية وتترك نسبة عالية من الطلبة لمصير مجهول.
- قلق وتوتر ورهبة وضغط نفسي لدى الطلبة وأسرهم.
- تفشي ظاهرة التدريس الخصوصي، وهناك بعض السلوكيات غير المرغوبة التي تصاحب تنفيذ الاختبار، وتضر بالعدالة، وتترك أثارا سلبية على مدخلات التعليم العالي.
- تفشي ظاهرة البحث عن بدائل خارجية متدنية المستوى لامتحان الثانوية العامة.
- قلة الاهتمام بالمهارات الأكاديمية العليا، وعدم تجريب الفقرات ومعرفة خصائصها مسبقًا.
- تفاوت وتباين في الأداء بين دورات الامتحان.
- الامتحان ورقي وتقليدي ومكلف.

الجزء الثاني: الإطار الوطني المقترح لتقييمات الطلبة في الأردن

Proposed National Framework of Students Assessments in Jordan

أكدت منظمة OECD ضمن السعي نحو تضافر الجهود من أجل تعلم أفضل (Synergies for Better Learning) وبناءً على مراجعة تجربة 28 دولة في بناء أطر التقويم[30]، على عدة عوامل تدفع إلى زيادة استخدام التقويم والتقييم، ومنها:

- ✓ زيادة الطلب على التقييم الفعَّال والمنصف، ذو الجودة العالية بما يكفل جودة التعليم في إعداد جيل من المتعلمين قادر على مواجهة التحديات الاقتصادية والاجتماعية.
- ✓ الاتجاه في التعليم نحو استقلالية المدرسة بشكل متزايد يزيد الحاجة إلى مراقبة أدائها.
- ✓ إدخال تحسينات على تكنولوجيا المعلومات مما يسمح بتطوير التقييم والتقويم.
- ✓ الحاجة إلى تقييم الطلبة واسع النطاق والتقييم الفردي بطريقة تسهّل المشاركة وإدارة البيانات.
- ✓ زيادة الاعتماد على نتائج التقييم لاتخاذ القرارات القائمة على الأدلة.

ويعد إطار التقييم جزءًا من نظام التعليم الشامل، ويتم تصميمه بما يتماشى مع أهداف نظام التعليم، ويسعى إلى تحسين نتائج الطلبة، ويتم تطويره في ضوء السياق الأوسع الذي تحدده سياسات التعليم السائدة، ومنها مستوى اللامركزية في الإدارة التعليمية، وتقاليد المجتمع وثقافته وقيمه في مجال التعليم. وتركز الحكومات وواضعو السياسات التعليمية بشكل متزايد على تقييم الطالب (Student Assessment)، باستعمال أدوات متنوعة، وتُستخدم بيانات التقييم في فهم أفضل لمدى جودة تعلم الطلبة، ولتوفير المعلومات للآباء والمجتمع ككل حول الأداء التعليمي، ولتحسين المدرسة وقيادتها وممارسات التدريس فيها.

وقد قدمت منظمة OECD في عام 2013 نظرة عامة حول التقييم والتقويم في مجال التربية، وضمنت ذلك إطارا عاما للتقييم والتقويم بمكوناته المختلفة من تقييم الطلبة Student assessment، وتقدير المعلمين Teacher appraisal، وتقويم المدارس School evaluation، وتقدير قادة المدارس School leader appraisal، وتقويم النظام التعليمي Educational system evaluation وما بينها من ارتباطات متبادلة. وتم التفريق بين مصطلحات التقييم Assessment، والتقدير Appraisal، والتقويم Evaluation. وقد تم استخدام مصطلح التقييم للإشارة إلى الأحكام المتعلقة بالتقدم الفردي للطالب وتحقيق أهداف التعلم، ومصطلح التقدير للإشارة إلى الأحكام المتعلقة بأداء المهنيين على مستوى المدرسة من معلمين وقادة المدارس، ومصطلح التقويم للإشارة إلى الأحكام المتعلقة بفعالية المدارس والأنظمة التعليمية والسياسات والبرامج.

[30] OECD Reviews of Evaluation and Assessment in Education

وتعد عمليتا التقييم والتقويم أمرًا بالغ الأهمية لبناء أو إصلاح أي نظام تعليمي، وهما محور أساسي في أي نظام تعليمي يسعى إلى التحسين. ومن منطلق الاهتمام بتقييمات الطلبة في الأردن لأغراض المشروع الحالي، فإن هناك حاجة إلى إطار وطني لتقييم الطلبة من أجل بناء تقييمات عالية الجودة للطلبة وإدارتها بفعالية. والتقييمات التي تتمتع بالصدق والثبات والموضوعية والعدالة ضرورية لتحسين الممارسات التعليمية وتعلم الطلبة على جميع المستويات.

وقد شهدت كثير من الأنظمة التعليمية تحولات في أنظمة التقييم، حيث ازداد التركيز على التقييم من أجل التعلم (Assessment for learning)، وزاد الاهتمام بالتقييم المعتمد على الأداء (Performance based assessment) وبالتقييم الواقعي (Authentic assessment) وبالتقييم القائم على المعايير (Standards based assessment) وبالتقييم القائم على الكفايات (Competency based assessment) كممارسات فضلى في تقييم الطلبة. ويركز التقييم من أجل التعلم على قياس نتاجات التعلم (Learning Outcomes)، ويوصف بأنه تقييم تكويني وواقعي ومحكي المرجع، يرافق العملية التعليمية التعلمية، ويوفر التغذية الراجعة الفورية لتوجيه التعليم وتحسين التعلم. ويتضمن أشكالًا متعددة من التقييمات التي تراعي الفروق الفردية، مثل التقييم الواقعي والتقييم الذاتي وتقييم الأقران.

وفي الأردن، بذلت وزارة التربية والتعليم من خلال الإدارات المختلفة وبشكل خاص إدارة الامتحانات والاختبارات جهودًا كبيرة لتطوير التقييمات في الأردن، إلا أنها لا زالت تركز على تقييم التعلم، ومحدودة النطاق، ولا تراعي احتياجات المتعلمين والفروق الفردية بينهم. كما أن هناك شواهد على التدريس من أجل الامتحان وليس من أجل التعلم، وقد أكدت الخطة الاستراتيجية لوزارة التربية والتعليم (2018–2022) على ضرورة تطوير ومراجعة نظام التقييم الوطني للوصول إلى تقييم شامل لمهارات الطلبة وسلوكياتهم وقدراتهم ومعارفهم وبما يتماشى مع المعايير الدولية. وأكدت الخطة التنفيذية ضرورة العمل على تطوير استراتيجية تقييم وطنية، وتحديد مراحل التقييم بما يتماشى مع هيكل سلم التعليم الجديد (نهاية المرحلة الأساسية الدنيا والمتوسطة والعليا، ونهاية المرحلة الثانوية)، واعتماد مؤشرات أداء للطلبة في كل مرحلة وللمباحث الأساسية وإعداد امتحانات معيارية موحدة للصفين الثالث والتاسع وفقًا لمعايير دولية، وتطبيقها على المستوى الوطني وتنفيذ الاختبارات التشخيصية في مجالي القراءة والرياضيات للصفوف المبكرة، وتطوير معايير وأسس المتابعة والمساءلة للمشاركين في التقييمات الوطنية، وتدريب الموظفين على استخدامها وتطبيقها في المدارس.

واستجابة إلى هذه الخطة، وتماشيًا مع حركة تطوير التقييم واستراتيجياته على المستوى الدولي، وتطوير مناهج دراسية أردنية جديدة قائمة على المعايير، وإقرار خطة دراسية جديدة للمرحلة الثانوية، وجعل التشعيب

إلى التعليم الأكاديمي والمهني بعد الصف التاسع، ولمعالجة جوانب الخلل والضعف المصاحبة للتقييمات الأردنية الحالية، فإنه من الضروري تطوير إطار وطني شامل ومتماسك لتقييم الطلبة في الأردن.

وينظر إلى الإطار الوطني لتقييم الطلبة باعتباره الخطة التي توجه تقييمات الطلبة وتحدد المحتوى الذي سيتم تقييمه في مراحل التعليم العام بشكل يضمن التكامل والتماسك بين مكونات التقييم، ومواءمته مع المعايير الوطنية للمناهج الدراسية، ويحقق التوافق بين سياسة التقييم الوطنية والأهداف التعليمية للدولة، وبما يخلق رؤية موحدة ومتماسكة للتقييم بين الجهات المختلفة سواء المسؤولة عن تصميم التقييم أو المنفذة لعملياته. والهدف الرئيس للإطار الوطني لتقييم الطلبة هو وصف كيفية تنظيم تقييمات جيدة لتعلم الطلبة، ويصف كيف يتم تنظيم التقييمات الخارجية والدولية وربطها بأشكال أخرى من التقييمات، وتستخدم تقييمات تحصيل الطلبة في اتخاذ القرارات وعمل التحسينات على المستوى الفردي وعلى مستوى النظام، لذلك يجب أن تكون التقييمات ذات جودة عالية من حيث الصدق ووضوح الغرض والثبات والعدالة والأمان والاستخدام الصحيح.

وقدمت دراسة البنك الدولي[31] تصورًا لإطار تقييم يرشد إلى كيفية بناء إطار أكثر فعالية لتقييم الطلبة وتم تنظيمه حول بعدين رئيسين لأنظمة التقييم، هما: أنواع/أغراض أنشطة التقييم، وجودة تلك الأنشطة، وهناك ثلاثة أنواع رئيسة لأنشطة التقييم، لكل منها أغراضه ويتناول احتياجات مختلفة للمعلومات، هذه الأنواع الثلاثة الرئيسة هي: التقييم الصفي، والامتحانات العامة، والتقييمات واسعة النطاق على مستوى النظام، ويتم تقييم جودة أنشطة التقييم بالاعتماد على مدى كفاية ما توفره من معلومات لدعم عملية اتخاذ القرار، وهناك ثلاثة محركات رئيسة لجودة المعلومات في نظام التقييم، هي: السياق التمكيني Enabling context، ومواءمة النظام System alignment، وجودة التقييم Assessment quality.

ويشير السياق التمكيني إلى السياق الذي يتم فيه النشاط التقييمي، وإلى أي مدى يفضي ذلك السياق إلى التقييم أو يدعمه، ويغطي قضايا مثل الإطار التشريعي أو السياسي لأنشطة التقييم؛ والهياكل المؤسسية والتنظيمية المسؤولة عن تصميم التقييم أو تنفيذه أو استخدام نتائجه؛ وتوافر مصادر تمويل كافية ومستقرة ووجود متخصصين مدربين بصورة جيدة على جميع عمليات التقييم، وتشير مواءمة النظام إلى مدى توافق أنشطة التقييم مع بقية عناصر النظام التعليمي، ويتضمن ذلك درجة التطابق بين أنشطة التقييم وكل من أهداف التعلم، والمعايير، والمناهج الدراسية، وتدريب المعلمين أثناء الخدمة وقبلها، وتشير جودة التقييم إلى الجودة السيكومترية لأدوات وعمليات وإجراءات الأنشطة التقييمية، وتغطي قضايا مثل تصميم أنشطة التقييم وتنفيذها،

[31] Jordan Student assessment, SABER Country Report, 2014

وتحليل استجابات الطلبة لتلك الأنشطة وتفسيرها ومدى ملاءمة عملية الإبلاغ عن نتائج التقييم واستخدام تلك النتائج.

ويعمل الإطار الوطني لتقييم الطلبة كمظلة لجميع أنشطة التقييم، ويساعد في إنشاء رؤية مشتركة ومتماسكة لتقييم الطلبة، ويوضح الغرض والأساليب والعلاقات بين التقييمات الصفية والتقييمات الوطنية والامتحانات العامة، ويضمن المساهمة المباشرة لهذه التقييمات في تحسين نتائج الطلبة من خلال توجيه التدريس وتحسين التعلم، ويوفر الإطار معلومات أساسية مهمة يحتاجها مطورو الاختبارات من أجل تطوير اختبارات صادقة وثابتة وعادلة. وتتلخص الملامح الأساسية في الإطار بالآتي:

- وصف واضح لأغراض وأهداف التقييم، وتعريف واضح لكل مجال من مجالاته.
- وصف أنواع التقييمات والنسب للمعرفةو المهاراتو القيمو الاتجاهات، ونتاجات التعلم.
- تحديد المعايير التي يتم على أساسها تقييم أداء الطلبة مثل معرفة المحتوى، وإتقان المهارات الأساسية، والتفكير الناقد، والإبداع، والتعاون وغيرها من كفايات القرن الحادي والعشرين، أي أنه يوجه عملية تطوير الأدوات والاختبارات.
- تحديد العوامل السياقية التي ترتبط بالتحصيل للمساعدة في تفسير النتائج والتباينات فيما بينها، ويوجه عملية تطوير أدواتها.
- تصميم الاختبار، بما في ذلك شكل الفقرة، ومدة الاختبار، ونماذجه، وعدد الفقرات في كل نموذج، وعدد الفقرات المشتركة بين النماذج المختلفة.
- توفير لغة مشتركة في مناقشة التقييمات ونتائجها.
- الإدماج الشامل من خلال توفير التسهيلات المناسبة لذوي الاحتياجات الخاصة.

ويحدد الإطار الوطني لتقييم الطلبة الكفايات الرئيسة التي يتم قياسها، ويضمن المواءمة بين الكفايات التي تسعى المناهج الأردنية إلى إكسابها للطلبة وأدوات التقييم التي تقيسها، كما يحدد الأهداف والأدوار والمسؤوليات في العمليات التقييمية وآليات استخدام نتائجها في وضع سياسات وزارة التربية والتعليم. ويغطي الإطار أنشطة التقييم الصفي (التشخيصية والتكوينية، والختامية)، والامتحانات العامة، والتقييمات واسعة النطاق (الوطنية والدولية)، ويجب أن تكون تقييمات الطلبة وسيلة مفيدة في دعم تعلمهم وتحسينه، وتقديم تقارير منتظمة عن التقدم الفردي لكل منهم، ومدى نجاحه في تحقيق الكفايات المحددة في المنهاج، ووضع معايير الإنجاز لكل مستوى تعليمي ومراقبتها، وإجراء المقارنات وإصدار الشهادات وتوجيه الطلبة لمزيد من التعليم.

وقد تبين من الجزء الأول من هذا التقرير وكما هو مبين في الشكل 3، التوزيع غير العادل للتقييمات المختلفة في الأردن على الصفوف المختلفة؛ إذ ازدحمت التقييمات المختلفة في الصفوف الثاني والثالث والرابع. وكانت التقييمات مكثفة في هذه الصفوف، وهذا يثقل كاهل المعلمين والطلبة، ويكون على حساب وقت التدريس ويُحوِّل التقييم من وسيلة لغاية من حيث لا نقصد. كما يغلب على التقييمات الصفية استخدام اختبارات الورقة والقلم، وهناك تداخل في أغراض التقييمات المستخدمة للصف نفسه، ولا يمكن مقارنة النتائج سواء للفرد نفسه في الأوقات المختلفة أو بين الأفراد، ولا تنسجم مع الخطة الدراسية الجديدة في المرحلة الثانوية حيث أصبح التفريع إلى الأكاديمي والمهني في نهاية الصف التاسع، وأصبح امتحان الثانوية العامة في سنتين. ويبدو أن لكثرة الاختبارات وعدم استخدام نتائجها على النحو الأمثل انعكاسات سلبية على نظرة الطلبة لها، ويتمثل ذلك في عدم الاستعداد الكافي لها وعدم بذل الجهد واللامبالاة في الأداء عليها.

الشكل 3: التقييمات الحالية في النظام التعليمي الأردني

وبالاعتماد على خبرات أعضاء الفريق المكلف بتطوير الإطار، والممارسات التقييمية الفضلى في الأنظمة التعليمية ذات الأداء العالي، والتوصيات والمقترحات من الدراسات والتقارير السابقة، ونتائج التحليل الكيفي للوثائق ذات الصلة، وتحليل نوعي للبيانات التي تم جمعها من استطلاعات رأي 2927 معلمًا ومعلمة و498 مديرًا ومديرة مدرسة و338 مشرفًا تربويًا و74 رئيس ورئيسة مجلس تطوير، والمعلومات التي تم الحصول عليها من المقابلات مع مديري عدة إدارات في وزارة التربية والتعليم ومع الجهات الدولية الداعمة، ومع الجهات المسؤولة عن برامج إعداد المعلمين وتأهيلهم في الجامعات الأردنية وأكاديمية الملكة رانيا لتدريب المعلمين وما تضمنه الإطار العام للمناهج الأردنية المطورة من موجهات، وما تضمنته الكتب المدرسية من تقييمات، وما تضمنه تقرير الخبير الدولي[32] من استنتاجات وتوصيات، فإن الفريق يقدم الإطار الوطني لتقييم تعلم الطلبة من الصف الأول الأساسي إلى الصف الثاني عشر على النحو المبين في الشكل 4.

الشكل 4: الإطار المقترح للتقييمات في النظام التعليمي الأردني

[32] Cascallar, E. (2024). NSAF- Initial observations and recommendations

وتتلخص أبرز ملامح هذا المقترح في أنه اهتم في قياس الاستعداد للالتحاق في المدرسة في الصف الأول، وقلل عدد الصفوف التي تقتصر فيها التقييمات على التقييم الصفي من ستة صفوف (1،5،6،7،9،11) إلى ثلاثة صفوف (2،5،7)، ووزع التقييمات بصورة أكثر عدالة على الصفوف مع مراعاة أن يكون عدد التقييمات للصف الواحد على الأكثر تقييمين بجودة عالية وبمواصفات سيكومترية جيدة، والتأكيد على جودة التقييمات بصورة أكثر من عددها، وراعى عدم إضافة تكاليف عالية، ويتمتع بمرونة تسمح بإضافة تقييمات جديدة وحذف وتعديل تقييمات قائمة وفقًا للظروف والمستجدات والإصلاحات المستقبلية خاصة ما يتعلق بالتقييمات التي ترتبط بمشاريع مدعومة من الجهات المانحة.

ويهدف هذا الإطار إلى إنشاء نظام متماسك للتقييم يزود أصحاب المصلحة ببيانات تقييم صادقة ودقيقة مع معلومات سياقية أخرى تساعد في تفسيرها. وتستخدم بيانات التقييم في مساعدة الطلبة على التقدم في التعلم وتوفر لهم الدعم الذي يحتاجون إليه، وفي إصدار أحكام أكثر عدالة حول أداء الطلبة، والسماح لواضعي السياسات بتحديد مجالات نظام التعليم التي يمكن تحسينها. ويعتمد الإطار على مبادئ أساسية تشكل الأساس في النظام التعليمي، لإنصاف الطلبة وتقييم إنجازاتهم. وتتمثل هذه المبادئ في التوافق بين التقييمات والمناهج الدراسية، وتمثيل محتوى الموضوعات الدراسية وما تتضمنه من كفايات ونتاجات تعلم، ووضوح أغراضها، وتنوعها، وذات معايير واضحة، وآمنة، وشفافة وعادلة، وقابلة للتفسير، ويمكن الاعتماد عليها، وقابلة للمقارنة، وتوفر التغذية الراجعة البناءة وفي الوقت المناسب، ومرنة بحيث يمكن التحكم فيها وتعديلها من أجل تحقيق الهدف المنشود، وأن تجرى وفقًا لأعلى المعايير الأخلاقية والمسؤولية والمساءلة.

وإذا ما أُحسن بناء مجموعة التقييمات المقترحة في الإطار وتنفيذها بصورة سليمة فإنها ستؤدي وظائف جمّة للطالب وولي أمره والمعلم والنظام التعليمي وكل أصحاب المصلحة. فهي تقيس التقدم لدى الطالب في مساره التعليمي في الصف الواحد أو عبر الصفوف، وتدعم تعلمه وتزيد من فرص نجاحه وتقدمه. وتزود أولياء الأمور بمعلومات دقيقة حول مسيرة التعلم لأبنائهم، وتوضح نقاط قوتهم وضعفهم وأين هم من تحقيق نتاجات التعلم. وتوفر المعلومات الكافية للطلبة والأهل التي تمكنهم من اتخاذ قرارات مستنيرة حول مسارهم التعليمي. وتساعد في تحديد احتياجات التعلم في وقت مبكر مما يُمكن من التخطيط لمجموعة من التدخلات العلاجية المناسبة وتنفيذها في الوقت المناسب. ويتطلب إعداد التقييمات المقترحة تحديد الأولويات التقييمية التي يجب أن تتناولها برامج التطوير المهني للمعلمين والقائمين عليها. ويجب أن تقلص برامج التطوير المهني التباينات في المعرفة التقييمية النظرية وأساليب التقييم وأدواته بين المعلمين في المدارس المختلفة والمباحث المختلفة.

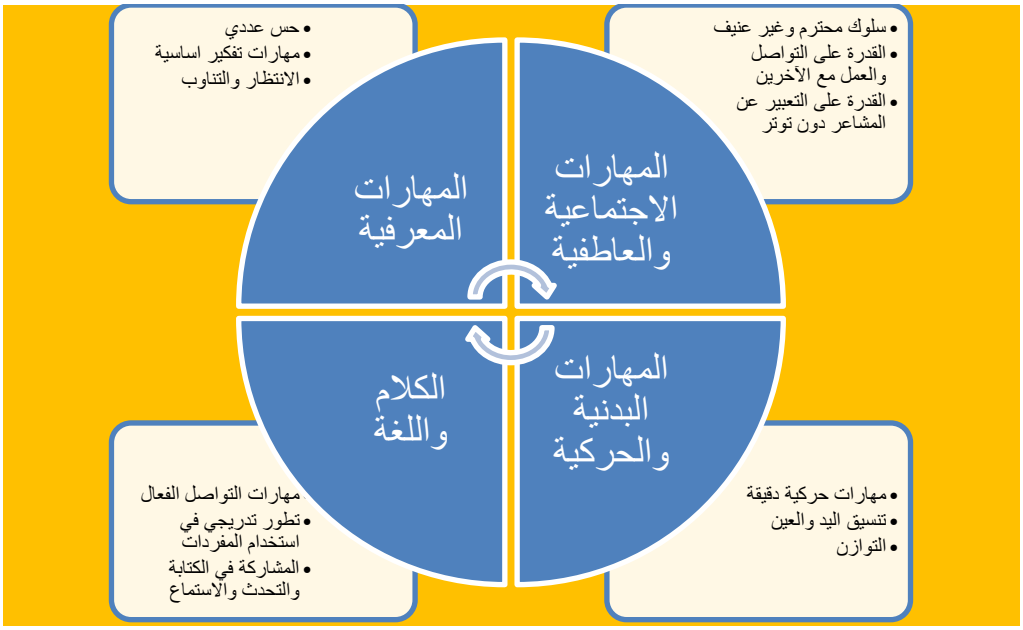
وعلى المستوى الوطني، فإن التقييمات المقترحة تمكن من إجراء مقارنات عادلة بين المدارس تأخذ باعتبارها السياقات الخاصة بتلك المدارس وطلبتها والمعلمين فيها، كما ستوفر أدلة أكثر موثوقية لمراقبة أفضل للنظام التعليمي مقارنة مع أنظمة عالمية وإقليمية أخرى. وفيما يلي وصف تفصيلي للتقييمات التي يتضمنها الإطار الوطني المقترح لتقييم الطلبة في الأردن.

1. اختبار الاستعداد المدرسي (الصف الأول)

تشير الدراسات التربوية إلى أن الاستثمار في مرحلة الطفولة المبكرة يقلل من الفوارق في التحصيل بين المجموعات المختلفة، خصوصًا بين الأطفال المحرومين اقتصاديًا واجتماعيًا ونظرائهم من الأطفال الآخرين. وتُعد السنوات الخمس الأولى من حياة الطفل فترة حاسمة لتطوره مدى الحياة. فهذه السنوات تمهد الطريق للنمو المستقبلي للأطفال والنجاح في المدرسة والحياة. وعند بلوغ الطفل سن الست سنوات، يصبح ذهابه إلى المدرسة أمرًا ضروريًا، مما يجعل تهيئة الأطفال للتعلم في المدرسة أمرًا بالغ الأهمية.

ويشمل نطاق استعداد الطفل للمدرسة الكفايات والمهارات الجسدية، والاجتماعية، والانفعالية، واللغوية، والمعرفية التي يحتاجها للمشاركة والاستفادة من خبرات التعلم المبكرة، والنجاح في المدرسة، والتقدم إلى مراحل لاحقة من التعلم. ويبين الشكل 5 الأبعاد الرئيسة في استعداد الطفل الشامل للمدرسة. ويُعد الاستعداد المدرسي مقياسًا لمدى جاهزية الطفل للنجاح في المدرسة، سواء من الناحية المعرفية أو الانفعالية، ويشير إلى أن الطفل قد بلغ مرحلة معينة من تطوره حيث يكون التعليم الرسمي مفيدًا له (Ounce prevention fund, 2017).

الشكل 5: الأبعاد الرئيسة في استعداد الطفل الشامل للمدرسة



وحظيت الطفولة المبكرة باهتمام بارز في الأردن، خاصة في إطار مشروع تطوير التعليم نحو الاقتصاد المعرفي (ERfKE, 2003–2014) . وتم تحقيق إنجازات ملموسة تمثلت في التوسع في بناء رياض الأطفال وتدريب معلمات رياض الأطفال، وتطوير المناهج، ودعم برامج الرعاية الوالدية، ووضع المعايير اللازمة لإنشاء رياض الأطفال، والمعايير النمائية للأطفال، وهذا يساعد في جعل الأطفال أكثر استعدادًا للتعلم المدرسي.

وهناك توافق متزايد على أن تكون النقطة الرئيسة التي يتم عندها إجراء فحوصات حول تطور المهارات الأساسية الضرورية لمعرفة القراءة والكتابة والحساب في بداية التعليم الإلزامي، وتستخدم هذه الفحوصات بالإضافة لذلك، في مراقبة النمو البدني والسلوكي والمعرفي بصورة رسمية أو غير رسمية، للتأكد من تلبية الطفل للتوقعات الطبيعية، وإذا لم يكن كذلك، لتحديد احتياجات التدخل اللازمة.

وقد استخدم المركز الوطني لتنمية الموارد البشرية أداة لقياس استعداد الأطفال للمدرسة، حيث كان الاستخدام يقوم على أساس العينات من جهة ويتم في سنوات متفرقة غير متتابعة من جهة أخرى، كما استخدم برنامج أساس USAID Early Grade Education Activity (ASAS) أداة لقياس الاستعداد المدرسي مع الأطفال في السنة الأخيرة في رياض الأطفال. ويتضح من هذه الاستخدامات، أنها لا تشمل جميع الأطفال الذين يلتحقون بالمدرسة، بل عينة من الأطفال وفي سنوات معينة وبخصائص معينة. وهذا يفيد في تحديد حاجات الأطفال عمومًا، لكنه لا يحدد الحاجات الفردية للأطفال، حيث تقتضي الضرورة تحديد الحاجات الفردية لكل طفل، وتَدَّخُل بَنَّاء من قبل المعلم/المعلمة لمساعدته على الانخراط السوي في خبرات التعلم.

غرض الاختبار

الغرض من الاختبار هو تحديد مجالات عدم الاستعداد لدى الطفل من أجل التخطيط للتدخلات اللازمة لمساعدته في الوصول إلى المستوى المطلوب من الاستعداد، وتنفيذ تلك التدخلات. ولا يجوز استخدامه كمعيار لقبول الطفل من عدمه، وحاليًا، لا يلتحق جميع الأطفال الصغار في الأردن برياض الأطفال. ويتم مراقبة نمو وتطور الذين يلتحقون بها، من قبل معلمات رياض الأطفال في المجالات المعرفية، والسلوكية، والعاطفية. ويتم تقديم الدعم المناسب لهم عند الضرورة. ومع ذلك، لا توفّر هذه العملية معلومات للمدارس حول استعداد جميع الأطفال للتعليم الرسمي في الصف الأول.

وعليه، فإن الغرض الأساسي للاختبار هو تقييم استعداد الطفل للمشاركة في التعليم الأساسي مقابل الحد الأدنى من مستويات المعايير، ويُستخدم لتحديد مستوى تطور الطفل في مجالات حيوية، مثل القدرات المعرفية، والمهارات الاجتماعية، والنمو العاطفي، والمهارات اللغوية، والمهارات الحركية. ويساعد ذلك في:

1. تحديد الأطفال الذين يحتاجون إلى دعم إضافي، إذ يمكن تحديد الأطفال الذين يواجهون صعوبات في بعض المجالات، مما يسمح بتقديم تدخلات مبكرة لمساعدتهم على النجاح في المدرسة.

2. تكييف الأنشطة التعليمية وفقًا لاحتياجات كل طفل، وبما يعزز تعلمه (Janus & Offord, 2007).

3. تقليل الفجوات في التحصيل بين الأطفال من خلفيات اجتماعية واقتصادية مختلفة، من خلال تقديم الدعم اللازم للأطفال المحتاجين (Ounce Prevention Fund, 2017).

4. مساعدة وزارة التربية والتعليم والمدارس في وضع استراتيجيات طويلة الأمد لتطوير التعليم بناءً على احتياجات الأطفال (UNICEF, 2016).

محتوى الاختبار

يجب أن يغطي اختبار تقييم استعداد الطفل للمدرسة مجموعة من الكفايات والمهارات التي يمكن تقييمها عند الأطفال والتي تعد ضمن التوقعات الطبيعية، ويبين الجدول (8) المجالات الرئيسة والفرعية لاختبار تقييم الاستعداد.

جدول (8): المجالات الرئيسة والفرعية لاختبار تقييم الاستعداد

المجالات الرئيسة	المجالات الفرعية
الصحة الجسمية	مهارات حركية دقيقة التوازن تناسق اليد والعين
الكفاية الاجتماعية	المسؤولية والاحترام الثقة بالنفس
النضج الانفعالي	السلوك التعاوني والاجتماعي التعبير عن المشاعر
التطور المعرفي واللغوي	حس عددي مهارات تفكير أساسية استخدام المفردات الكتابة والتحدث والاستماع
مهارات التواصل والمعرفة العامة	القدرة على التواصل الفعال مع الآخرين المعرفة العامة المتعلقة بالحياة اليومية.

وفي ظل التحول الرقمي في التعليم، قد يكون من المهم تقييم مستوى استعداد الأطفال لاستخدام التكنولوجيا في التعلم.

الأدوار والمسؤوليات

تتولى إدارة الامتحانات والاختبارات في وزارة التربية والتعليم مسؤولية بناء وتطوير أداة تقييم استعداد الطفل للتعلم المدرسي مع دليل استخدامها. وتقع على معلم/معلمة الصف الأول الأساسي مسؤولية تطبيقها. وللقيام ببناء الأداة، يتم تكليف فريق من أخصائي القياس ومشرفي المرحلة الأساسية الدنيا لتولي هذه المهمة. ويمكن لهذا الفريق بناء الأداة وتطويرها بالاستفادة من الأدوات التي سبق تطبيقها في الأردن وغيرها من الأدوات العالمية حسب الخطوات والمعايير المعتمدة في بناء الأدوات. ويتم عمل مجموعات تركيز من معلمي ومعلمات الصف الأول للاستفادة من خبراتهم وما يوفرونه من تغذية راجعة في تطوير الاداة وإخراجها بصورتها النهائية. وتقوم إدارة الامتحانات والاختبارات من خلال أخصائي القياس فيها وبالتعاون مع إدارة الإشراف والتدريب التربوي بتوفير التدريب المناسب لمعلمي ومعلمات الصف الأول الأساسي على كيفية استخدام الأداة في تحديد احتياجات الطلبة. ويتم توفير الأداة للمعلمين والمعلمات المعنيين في المدارس بأي طريقة مناسبة (إلكترونيًا أو ورقيًا) وتكليفهم بتطبيقها على الأطفال خلال الشهر الذي يسبق الفصل الدراسي الأول، لتحديد حاجات الأطفال اللازمة للنجاح في تجربة التعلم في المدرسة، والعمل على تلبية تلك الحاجات، وتوفير الدعم الذي يحتاجونه قبل بدء العام الدراسي، من خلال برامج علاجية تنفّذ في المدرسة من قبل معلم/معلمة الصف الأول الأساسي.

المجتمع المستهدف

يستهدف التقييم جميع الطلبة المسجلين في الصف الأول في مدارس الأردن جميعها، بما في ذلك الأطفال ذوي الإعاقة بعد أخذ تلك الاحتياجات بعين الاعتبار.

جمع البيانات

يتم جمع البيانات الناتجة عن التقييم من قبل المدرسة ويجب الاحتفاظ بها وادخالها على نظام إدارة المعلومات التربوية (EMIS). وينبغي توفير إمكانية الوصول إلى البيانات لتقييم الاحتياجات على المستوى الوطني فيما يتعلق بتدخلات دعم التعلم الإضافية للطلبة الذين ينقصهم الاستعداد الكافي.

استخدام النتائج

يجب التأكد من فهم المعلمين بوضوح أهداف التقييم ويدركون أن نتائجه لن تُستخدم لأغراض المساءلة كما لن تستخدم لغرض الالتحاق بالمدرسة، وإنمّا لتحديد الطلبة الذين يحتاجون لدعم إضافي. ويتم توجيه الطلبة

الذين لا يظهرون مستوى كافٍ من المهارات والكفايات المطلوبة للحصول على دعم تعليمي فردي إضافي حتى يصلوا إلى المستوى المطلوب من التطور. ويمكن للمعلمين والمدرسة الاستفادة من النتائج على النحو الآتي:

- تحديد الاحتياجات الفردية: يمكن استخدام النتائج لتحديد الأطفال الذين يحتاجون إلى دعم إضافي أو تدخلات تربوية مبكرة.
- التخطيط للتدريس: يمكن تكييف الأنشطة والبرامج التعليمية بناءً على احتياجات الأطفال المختلفة، سواء من حيث المهارات الجسدية أو المعرفية أو الانفعالية.
- الدعم الأسري: تقديم توصيات للأهل حول إشراكهم في كيفية دعم الطفل في المنزل من خلال الأنشطة التعليمية والترفيهية.

التوصيات بخصوص اختبار الاستعداد:

هذا الاختبار حديث على النظام التعليمي الأردني، لذلك لا بد من عمل الآتي:

أ. تحليل نتائج الاختبار للتحقق من خصائصه السيكومترية من جهة ومعرفة العوامل التي ترتبط بأداء الطلبة من جهة أخرى.

ب. توفير التدريب المناسب لمعلمي ومعلمات الصف الأول على كيفية استخدام نتائج الاختبار في تحديد حاجات الطلبة وبناء برامج علاجية لتلبية تلك الحاجات.

ت. دراسة أثر البرامج العلاجية المنفذة في تعلّم الطلبة.

2. التقييمات الصفية (جميع الصفوف من الصف الأول إلى الثاني عشر)

Classroom Assessment of students in Jordan

يُعد التقييم الصفي أحد أقوى الأدوات للكشف عمّا يعرفه الطلبة ويفهمونه، وما عليهم فعله والاستجابة له، والربط بين التقييم والتعليم أمر بالغ الأهمية للتعلم الفعّال، وهو من مهام وعمل المعلم. ولبناء أدوات التقييم الصفّي الجيدة لا بد للمعلم من أن يبدأ بالمعايير ويحلّل المهارات والمعارف المطلوب من الطلبة إتقانها، ثم يقوم بتصميم استراتيجيات تعليمية تجذب الطلبة وتربطهم بالمعايير، ثم يراقب تقدم الطلبة نحو الإتقان من خلال التقييمات التكوينية المتكررة والتغذية الراجعة الوظيفية. توفر التقييمات الصفيّة التكوينية للطلبة بيانات كمية ونوعية يمكن توظيفها بفاعلية لتحسين تعلمهم. ويعتمد التقييم الصّفي بصورة رئيسة على التقييم التكويني للحصول على معلومات مستمرة عن الطلبة أثناء التعلم تفيد في تعزيز التدريس وتحسين التعلم، وعلى التقييمات الصفية الختامية في إعطاء الدرجات وإصدار الأحكام بالنجاح والترفيع إلى الصفوف اللاحقة.

ويُنظر للتقييمات الصفية باعتبارها الأقرب إلى ما يجري داخل الغرفة الصفية، لذلك تعد الأفضل في تقييم تعلم الطلبة والحكم على فاعلية العملية التعليمية التعلمية ونجاحها في تحقيق غاياتها وأهدافها. وتقع مسؤولية تنفيذ التقييمات الصّفية على المعلم، ويتوقع منه أن ينوع في أساليب التقييم وأدواته وسياقاته وإجراءاته للحصول على بيانات صادقة ودقيقة عن تعلم الطلبة، ومن الضروري تنفيذ إصلاحات على التقييمات الصفية للطلبة في الأردن في مراحل التعليم المختلفة، وضمان التزامها بمعايير التقييم الصفي.

وتدور معايير التقييم الصفي حول ثلاثة مجالات: يتعلق المجال الأول بالأسس (Foundations) التي تقوم عليها التقييمات الصفية، ويتضمن ستة معايير، هي الغرض من التقييم، وتوقعات التعلم، وتصميم التقييم، وإشراك الطلبة في التقييم، وبناء التقييم، وإعلام الطلبة وأولياء الأمور بنتائج التقييم. ويتعلق المجال الثاني بالاستخدام (Use) للتقييمات الصّفية ونتائجها، ويتضمن خمسة معايير، هي تحليل أداء الطلبة، والتغذية الراجعة الفعّالة، والمتابعة التعليمية، والدرجات والتعليقات الموجزة، وكتابة وإعداد التقارير. ويتعلق المجال الثالث بجودة (Quality) التقييمات، ويتضمن خمسة معايير للجودة هي مراعاة التنوع بأشكاله المختلفة، ومراعاة الظروف والحاجات الخاصة، وعدالة التقييم وعدم تحيزه، والصدق والثبات، والتفكير بالتقييم ومراقبته بهدف تحسين جودته.

وهذا يتطلب توفير التدريب المناسب للمعلم حول كيفية تحقيق معايير التقييم الصفي أو توفير تقييمات جاهزة له وفقًا لتلك المعايير، والتدريب يكون قبل الخدمة في برامج الإعداد والتأهيل، وأثناء الخدمة من خلال برامج النمو المهني. ففي برنامج BTEC، يصمم التقييم ليناسب غرض وهدف المؤهل، ويشمل مجموعة من التقييمات المناسبة للمؤهلات المهنية. وجميع التقييمات في الوحدات الدراسية داخلية، أي أن المعلم يضع التعيينات ويصححها، باستثناء بعض الوحدات الإلزامية حيث تضع بيرسون التعيينات ويقوم المعلم بتصحيحها. ويمكن الاستفادة من هذه الممارسات التقييمية في التقييمات الصفية على نطاق أوسع، وبخاصة في تقييم الكفايات مثل التواصل والاتصال ومهارات الاستماع والمحادثة والتعبير.

وتشتمل ممارسات التقييم الصفي الفعّالة على أشكال متعددة في التقييمات الصّفية من مثل التقييمات التكوينية التي تستخدم للتحقق المنتظم من فهم الطلبة للمفاهيم والحقائق والنظريات الأساسية، والتقييمات التفاعلية التي تسمح للطلبة إظهار ما يعرفونه من خلال المناقشات الصّفية والمناظرات، والاختبارات الشفهية والعروض التقديمية، والتقييمات الذاتية التي تشجع الطلبة على التفكير في تعلمهم وتتبع تقدمهم، وتقييمات الأقران التي تقدم تغذية راجعة بناءة تساعد الطلبة على التحسن، وتقييمات الأداء والواجبات المنزلية والتقييمات القائمة على التكنولوجيا والتقييمات الختامية.

أغراض التقييمات الصّفية

تركز التقييمات الصّفية على ثلاثة أغراض مترابطة، هي: التقييم من أجل التعلم Assessment for learning، والتقييم كتعلم Assessment as learning، وتقييم التعلم Assessment of learning. ويصمم التقييم من أجل التعلم لتزويد المعلمين والطلبة بالمعلومات اللازمة لتعديل أنشطة التدريس وتحسين التعلم، فهو يقر بأن الطلبة يتعلمون بطرق خاصة، وأن لديهم أنماط تعلم سائدة، فهو يتطلب تصميمًا دقيقًا من جانب المعلمين بحيث يستخدمون المعلومات الناتجة لتحديد ما يعرفه الطلبة، ولاكتساب نظرة ثاقبة حول كيفية تطبيق ما يعرفونه. ويمكن للمدرسين أيضًا استخدام هذه المعلومات لتبسيط التدريس واستخدام الموارد والتغذية الراجعة في مساعدة الطلبة على تطوير تعلمهم.

أما التقييم كتعلم فهو عملية تطوير ودعم ما وراء المعرفة لدى الطلبة. وتتطلب العملية الجوهرية التنظيمية فيما وراء المعرفة أن يكون الطالب مُقيّمًا نشيطًا ومتفاعلًا وناقدًا. ويُمكّنه ذلك من فهم المعلومات وربطها بالمعرفة السابقة واستخدامها في التعلم الجديد. ويحدث ذلك عندما يراقب الطالب تعلّمه ويستخدم التغذية الراجعة من هذه المراقبة لإجراء التعديلات والتغييرات الرئيسة فيما يفهمه. وعلى المعلم مساعدة طلبته على التطور وتقييم ذاته، ويوفر لهم فرص التفكير فيما يتعلمونه، ويشجعهم على التحليل النقدي لتعلمهم.

وبالمقابل فإن تقييم التعلم تلخيصي بطبيعته، ويستخدم لتأكيد ما يعرفه الطلبة وما يمكنهم فعله، مقارنة بنتاجات التعلم في المنهاج و/أو أداء الآخرين و/أو أدائه السابق. وينبغي التركيز على جودة هذه التقييمات والتأكد من أنها تقدم بيانات دقيقة وسليمة عن كفاءة الطلبة، بحيث يمكن لمتلقي المعلومات استخدامها في اتخاذ قرارات معقولة ويمكن الدفاع عنها.

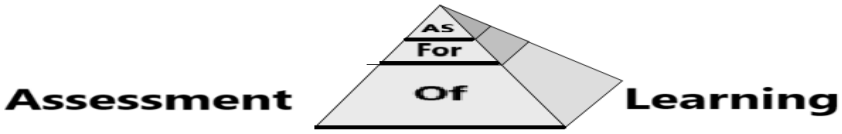
وتخدم التقييمات من أجل التعلم، والتقييمات كتعلم، وتقييمات التعلم أغراضًا قيمة ومختلفة[33]. ومع ذلك، ليس من السهل دائمًا تحقيق التوازن الصحيح في استخدامها. فإذا كان الهدف هو تعزيز التعلم لجميع الطلبة، فإن دور التقييم من أجل التعلم والتقييم كتعلم يأخذ مكانة أعلى بكثير من تقييم التعلم.

وتقليديًا، ركز التقييم الصفي بصورة كبيرة – كما هو موضح في الهرم الأول في شكل 6– على تقييم التعلم، أي قياس التعلم بعد حدوثه، واستخدام المعلومات لإصدار أحكام حول أداء الطلبة، وإبلاغ الآخرين بهذه

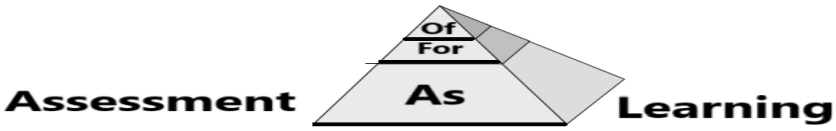
[33] Rethinking Classroom Assessment with Purpose in Mind: Assessment *for* Learning, Assessment *as* Learning and Assessment *of* learning, 2006.

الأحكام. وبصورة قليلة على التقييم من أجل التعلم، عندما يهتم المعلم بعمليات التشخيص والتقييم التكويني والتغذية الراجعة في مراحل مختلفة من عملية التدريس والتعلم. وبصورة نادرة على التقييم كتعلم، حيث يكون الطلبة محللين نقديين لتعلمهم. وعلى الرغم من أن بعض المعلمين قد أدرجوا التقييم الذاتي في برامجهم، إلا أن القليل منهم استخدم التقييم الذاتي بشكل منهجي أو صريح لتطوير قدرة الطلبة على تقييم تعلمهم وتكييفه. ونرى أنه من الأفضل إعادة تشكيل التوازن بين التقييمات الثلاثة في سياسات تقييم الطلبة الأردنيين لتكون كما في الهرم الثاني في شكل 6، حيث يتم التأكيد على التقييم باعتباره تعلمًا، والتقييم من أجل التعلم، لكن يبقى هناك دور مهم يلعبه تقييم التعلم عندما تكون هناك حاجة إلى أحكام ختامية.

شكل 6: التوازن في أغراض التقييمات الصفية



الهرم التقليدي للتقييمات



الهرم المقترح للتقييمات

ويحدد الغرض من التقييم كيفية بناء التقييم واستخدامه. إذا كان الغرض هو تعزيز التعلم، فيجب أن يمنح التقييم الطلبة فرصة لجعل تعلمهم واضحًا دون قلق أو انتقاد. وإذا كان الغرض هو التحقق من التعلم لإعداد التقارير، فيجب على المعلمين أن يهتموا بشكل خاص بجودة التقييم، وكيف يمكن استخدامه من قبل الآخرين. ومن المهم جدًا للمعلمين أن يفهموا أغراض التقييم الثلاثة، وأن يدركوا الحاجة إلى تحقيق التوازن فيما بينها، وأن يعرفوا أيًّا منها يستخدمون ولماذا، ويستخدمونها جميعهم بحكمة.

وتستهدف التقييمات الصَّفية جميع الطلبة في الصفوف المختلفة خلال مسيرتهم التعليمية. وقد تبدأ التقييمات الصفية قبل عملية التعليم، وتستمر مع استمرارها، وتنتهي بعد انتهائها. وعليه، فإن التقييمات الصفية تستعمل بشكل منتظم في المراحل التعليمية لأغراض التقييم التشخيصي والتكويني والختامي. وينفذ التقييم القبلي قبل البدء بعملية التدريس، وهو تقييم تشخيصي يهدف إلى الكشف عن المهارات الضرورية اللازمة للتعلم

الجديد. ويساعد التقييم التشخيصي في كشف معيقات التعلم الجديد وتوفير متطلبات تجاوزها، وتصحيح الأخطاء الشائعة حول الموضوع لدى الطلبة، وتحفيز الطلبة على التعلم. كما يساعد في التعليم المتمايز Differential Teaching الذي يراعي الفروق الفردية لدى الطلبة ويراعي احتياجات ذوي الاحتياجات الخاصة. وتتمثل أدوات التقييم التي يمكن استخدامها في التقييم التشخيصي بالاختبارات التشخيصية واليوميات ومسوحات الطلبة والمقابلات والتأملات والمناقشات الصفية والخرائط الذهنية والمخططات.

وتوفر التقييمات التكوينية بيانات كمية ونوعية يتم توظيفها بفاعلية لدعم عملية التعليم وتحسين تعلم الطلبة. وتستخدم التغذية الراجعة التي توفرها التقييمات التكوينية في متابعة تقدم الطالب نحو تحقيق الأهداف وتوجيه التعليم. وتتمثل أدوات التقييم التي يمكن استخدامها في التقييمات التكوينية بالاختبارات القصيرة وبطاقات الخروج Exit cards، والواجبات الصفية، واليوميات حول الموضوع، وملفات الطلبة، وتقارير التقدم والمناقشات الصفية وما يتخللها من أسئلة وإجابات. ولا بد من الحرص على أن تكون هذه الأنشطة ممتعة وقليلة المخاطر، يرتاح لها الطلبة، ويكونوا أكثر استعدادًا للاستجابة وأقل رهبة من ارتكاب أخطاء، مما يساعدهم في فهم كيفية تعلمهم. ومن المهم استخدام البيانات والمعلومات المتجمعة في اتخاذ قرارات حول من يحتاج إلى دعم إضافي، وما الدعم الذي يجب تقديمه ومتى يقدم. وعليه، فإن التقييمات التشخيصية والتكوينية تساعد في توفير معلومات حقيقية تدعم التعليم والتعلم في الحصص الصفية. وهي تجسد التقييمات من أجل التعلم.

أما التقييمات الختامية فتستخدم عند الانتهاء من عملية التعليم والتعلم في تحديد مدى تحقيق الطلبة للمعايير. وتجسد التقييمات الختامية تقييم التعلم، وتستخدم لتقييم الفهم الكلي للطالب وتقدير درجته، وإصدار الأحكام المتعلقة بنجاحه وترفيعه للصف الأعلى من خلال تبني نظام موضوعي دقيق لرصد الدرجات وتوثيقها. ويجب أن تتواءم التقييمات الختامية مع نتاجات التعلم والمؤشرات الدالة عليها. وتتمثل أدوات التقييم الختامي بالاختبارات النهائية والمشاريع الفردية أو الجماعية، والتقارير، وملفات الإنجاز.

وينبغي أن تكون التقييمات الصفية متوازنة[34] Balanced Assessment ، أي تراعي الأنواع المختلفة من التقييمات وتمارس كل منها في الغرفة الصفية من قبل المعلم بوزن مناسب في ضوء مستوى نمو الطلبة. فهناك حاجة للتقييمات التكوينية بأدواتها وأساليبها المختلفة من أجل تحسين تعلم الطلبة، وهناك حاجة للتقييمات الختامية بأدواتها وأساليبها المختلفة بما فيها الامتحانات والتعيينات والمشاريع خاصة في المرحلة الأساسية العليا والمرحلة الثانوية، بغرض اتخاذ قرارات النجاح والإكمال والرسوب. ويفيد التنويع في التقييمات الصفية في تعزيز الصدق والعدالة والثبات.

[34] Smarter Balanced Assessments Concortium,2012

ويتم تقييم الطلبة في الصفوف الثلاثة الأولى باستعمال استراتيجيات وأدوات تقييم تعتمد على الأداء والواجبات بأنواعها واستخدام التقييمات البديلة بغير الاختبارات، ولا يتقدم الطالب لامتحانات ختامية في نهاية الفصل الدراسي، ولا رسوب في هذه الصفوف، ويعطى الطالب شهادة مدرسية توفر وصفًا نوعياً لما يعرف وما حقق من نتاجات التعلم في المباحث المختلفة. ويبين الجدول 9 الأوزان النسبية المقترحة لنوع التقييمات الصفية المناسبة حسب الصف.

جدول 9: الأوزان النسبية لنوع التقييمات الصفية المناسبة حسب الصف.

الصفوف	الوزن النسبي للمكونات	
	التقييم البديل بغير الاختبارات	الاختبارات
1-3	100%	0%
4-6	70%	30%
7-9	60%	40%
10-12	50%	50%

ويؤكد الجدول على ممارسة التقييمات البديلة في الصفوف الثلاثة الأولى، في حين يتم تقييم الطلبة في الصفوف (4-12) بالتقييمات البديلة بنسب متناقصة والاختبارات بنسب متصاعدة. ومن الضروري أن تستند الاختبارات الختامية الفصلية في بنائها إلى مخطط (Blueprint) معلوم للطلبة، ويشمل جميع الوحدات التي تعلمها الطالب بعدالة، وبصورة متوازنة اعتمادًا على عدد الحصص المقررة لها، أو أي معيار يمكن تبريره والدفاع عنه. كما ينبغي أن توفر نتائج صادقة وثابتة.

الأدوار والمسؤوليات

تشتمل الكفايات التقييمية على كيفية إعداد أو انتقاء أدوات التقييم الصفي المناسبة، وكيفية تهيئة الطلبة للتقييم، وكيفية إدارة التقييم وتصحيحه، ورصد الدرجات وإصدار الشهادات وتحليل النتائج وتفسيرها، وهذه المهام المختلفة هي من مهام المعلم وتقع على عاتقه مسؤولية أن تكون تقييماته صادقة وثابتة وشفافة وعادلة. وهذا يتطلب أن يكون المعلم على درجة عالية من الإعداد والتأهيل في مجال القياس والتقويم سواء أثناء الخدمة أو قبلها، وتتحمل الجامعات الأردنية والجهات التعليمية الأخرى التي تطرح برامج خاصة بإعداد وتأهيل المعلمين مسؤولية مراجعة برامجها لضمان اشتمالها على مساقات نظرية وعملية في مجال القياس والتقويم. وكذلك ينبغي أن تتحمل الوزارة ممثلة بإدارة الإشراف والتدريب التربوي توفير البرامج التدريبية الخاصة التي تسهم في النمو المهني للمعلم في مجال التقييمات الصفية وما يتصل بها من كفايات، على أن توكل عملية التدريب لمتخصصي

القياس والتقويم في الوزارة، ولا بد من متابعة المعلمين والإشراف عليهم خلال ممارستهم وتنفيذهم للتقييمات الصّفية من قبل المشرفين التربويين وأخصائيي القياس والتقويم (نقترح وجوب تعيين أخصائي قياس وتقويم واحد على الأقل في كل مديرية تربية وتعليم) ومديري مدارسهم، وتقديم ما يحتاجون إليه من مساعدة.

وفيما يلي مجموعة من التوجيهات والتوصيات التي تساهم في تطوير التقييمات الصفية وتحسين جودتها:

1. ينبغي أن يستند التقييم الصفي إلى معايير الأداء ومؤشراته

إن تحديد معايير ومؤشرات أداء تصف بطريقة إجرائية أهداف التقييم والكفايات الرئيسة التي سيتم قياسها مقابل مجموعة التقييمات والامتحانات الصفية، يضمن المواءمة بين نتاجات التعلم التي تسعى المناهج الأردنية إلى إكسابها للطلبة وأدوات التقييم التي تقيس تلك النتاجات والكفايات المرتبطة بها، كما يعمل على ربط التقييمات الصفية بغيرها من التقييمات الوطنية والدولية. لذلك لا بد من تحديث أدلة معايير الأداء ومؤشراته (PSI Performance Standards & Indicators ,) لتتلاءم مع المناهج الجديدة في المباحث الأساسية الأربعة (اللغة العربية، اللغة الإنجليزية، الرياضيات، والعلوم وفروعها) ولجميع الصفوف. وتساعد هذه الأدلة المعلمين على كتابة فقرات اختبارية تغطي مستويات الأداء المختلفة وتقيس نتاجات التعلم المختلفة بما فيها مهارات القرن الحادي والعشرين وتحاكي أسئلة التقييمات الدولية، كما تسمح بإدماج التكنولوجيا في التقييمات المختلفة. ويمكن الاعتماد على معايير الأداء ومؤشراته من قبل الجهات المعنية في وزارة التربية والتعليم والمركز الوطني لتطوير المناهج في بناء بنك مركزي للمهام التقييمية (Central Bank of Assessment Tasks) في المباحث المختلفة ولكل المراحل التعليمية. ويستطيع المعلم أن يسحب من البنك ما يحتاج من مهام تقييمية للاستخدامات الصفية، ويمكن للطلبة سحب ما يحتاجون إليه من مهام تقييمية لاستخدامها في التقييم الذاتي. ويشكل هذا البنك في حال وجوده مصدرًا مهمًا للتقييمات الصفية تحت الطلب (On-Demand Assessments)، وبذلك تكون مصادر التقييمات الصفية متوافرة للمعلمين، وموحدة عبر المدارس، وتكون دقيقة وموثوقة، وتتصف بخصائص سيكومترية جيدة، لأنه يتم إعدادها من قبل فريق من خبراء المحتوى والقياس والتقويم بالاعتماد على نتاجات التعلم والكفايات المرتبطة بها، بما فيها كفايات القرن الحادي والعشرين، وتم تجريبها لضمان تمتعها بخصائص سيكومترية جيدة.

2. ينبغي أن يراعي التقييم الصفي المبادئ الأساسية للتقييم النفسي والتربوي

ينبغي أن تراعي التقييمات الصفية للطلبة في الأردن مبادئ التقييم الأساسية. فيجب أن تكون صادقة، فالصدق يضمن أن المهام التقييمية والمعايير المرتبطة بها تقيس بشكل فعَّال مدى تحقيق الطلبة لنتاجات التعلم المقصودة على المستوى المناسب. ويجب أن تكون ثابتة ومتسّقة، وهذا يتطلب عمليات واضحة ومتسّقة في تحديد المهام التقييمية وتصحيحها بموضوعية. ويجب أن تكون شفافة وعادلة ومعلنة، أي أن المعلومات الواضحة والصريحة عنها متوافرة للطلبة وأولياء أمورهم وغيرهم، ويمكن الوصول إليها بسهولة، وتراعي الفروق الفردية. ويجب أن تكون مستمرة حيث تحدث قبل وأثناء وبعد عملية التعليم والتعلم، ومتنوعة وشاملة ومتعددة المصادر. ويتطلب التنوع استخدام أساليب وأدوات تقييم متنوعة في مواقف وسياقات مختلفة. ويتطلب الشمول توفير تغطية كافية لكل نتاجات التعلم في المجالات المعرفية والانفعالية، ومن مصادر متعددة تشمل الطالب نفسه وأقرانه والمعلمين وأولياء الأمور. فلا يجوز أن تقتصر على اختبارات القلم والورقة (Pencil and Paper)، ولا على قياس المستوى الأدنى من نتاجات التعلم المتعلقة بالمعرفة القائمة على الحفظ والتذكر، بل يجب أن تتعداها لقياس مهارات التفكير وحل المشكلات. ولا بد من توثيق مؤشرات الأداء الدالة على تحقق نتاجات التعلم لدى الطلبة، فقد تكون مؤشرات الأداء كتابية تتمثل في عينات من الأنشطة التي نفذها المتعلم من أوراق عمل ومشاريع فردية أو جماعية، واختبارات القلم والورقة، وقد تكون مؤشرات الأداء أدائية تتطلب استخدام التقييمات القائمة على الأداء (Performance-based Assessment) التي تتمثل في استخدام النماذج، والمجسمات، والأجهزة، والأدوات، والخرائط، والجداول، والوسائل السمعية والبصرية في تحقيق نتاجات التعلم العليا(الكفايات المعرفية)، ويتطلب تقييم الأداء استخدام أساليب متنوعة لتوثيق بيانات التقويم من خلال توظيف المناسب من طرق التسجيل (Recording Strategies) مثل ملفات الأعمال للطالب (Student Portfolio) بأنواعها الورقية والإلكترونية، وتلك التي تتضمن جميع الأعمال أو أفضل الأعمال، ويوميات الطالب (Student Journal)، وسجل وصف سير التعلم (Learning Log)، والسجلات القصصية (Anecdotal Records).

3. ينبغي أن يهتم التقييم الصفي بقياس الكفايات

الكفايات هي مزيج من المعرفة والمهارات والقدرات والسلوكيات التي يكتسبها الأفراد ويطبقونها لأداء المهام أو الأنشطة بفعالية. وتُقيم التقييمات القائمة على الكفايات مدى قدرة الطلبة على تطبيق معارفهم في

سياقات عملية، وإظهار مهارات محددة ذات صلة بمجال معين. وتعمل على تحويل التركيز من الحفظ عن ظهر قلب إلى إظهار المهارات، وبدلاً من الاعتماد على الاختبارات الكتابية فقط، يتم تقييم الطلبة من خلال المهام القائمة على الأداء، والمشاريع، وعمليات المحاكاة، والتطبيقات في العالم الحقيقي. تتطلب هذه التقييمات من الطلبة تطبيق معارفهم، والتفكير النقدي، وحل المشكلات، والتواصل بشكل فعّال، والتعاون، وإظهار المهارات الأخرى ذات الصلة.

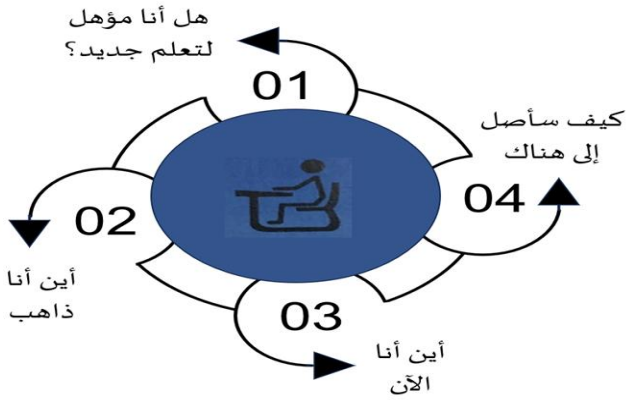
وتتضمن التقييمات القائمة على الكفايات عادةً مجموعة من الفقرات الأساسية التي تتمثل بوجود نتاجات واضحة ومحددة للتعلم، والتقييمات الواقعية التي تسمح للطلبة بتطبيق معارفهم ومهاراتهم في سياق مفيد، ووجود قواعد ومعايير شفافة للتصحيح تحدد التوقعات لكل كفاية، وتوفر مبادئ توجيهية واضحة للتقييم، والتقييم المستمر مع استمرار عملية التعلم مما يسمح بتلقي التغذية الراجعة وإجراء التحسينات بمرور الوقت.

ويجب الاهتمام بقياس مجموعة واسعة من الكفايات مثل معالجة المعلومات، والتواصل، والعمل الجماعي، والتحليل النقدي والإبداع وغيرها من مهارات القرن الحادي والعشرون، في سياقات تعلم تحاكي بيئات العمل، وتعزز ربط المعرفة بالحياة. ويتطلب ذلك استخدام العديد من إجراءات التقييم، ويشمل ذلك ملاحظات المعلم الرسمية وغير الرسمية للطالب، والاختبارات التحريرية، ومتابعة أداء الطالب على الواجبات المنزلية، وأعمال المختبر، والأوراق البحثية، والمشروعات، وسبر غور إجابة الطالب عن الأسئلة الشفوية، وتحليل السجلات المدرسية للطالب. ويمكن للمعلم في هذا المجال، انتقاء مواقف وسياقات تقييمية شبيهة بالمواقف التي تطرحها التقييمات الدولية، والتعرف على إمكانات الطلبة في تناول تلك المواقف وإيجاد الحلول لها.

4. ينبغي الاهتمام بالتقييم من أجل التعلم ((Assessment for learning (Afl)

يوفر التقييم من أجل التعلم أكبر قدر من المعلومات حول تعلم الطلبة وكيفيته. وتكون هذه المعلومات بمثابة تغذية راجعة فعّالة تفيد المعلم في توجيه وتعديل طريقة التعليم، وتفيد الطالب في تحسين عملية تعلمه. ويهدف التقييم من أجل التعلم إلى سد الفجوة بين الوضع الحالي للمتعلم وما يريد أن يصل إليه في تعلمه وإنجازه. وتظهر الأبحاث أنه يساعد الطلبة عمومًا وذوي التحصيل المنخفض خصوصا على تعزيز تعلمهم. ويبين الشكل 7 عمليات التقييم من أجل التعلم (Black & Wiliam,2006).

الشكل 7: عمليات التقييم من أجل التعلم



وينبغي مساعدة الطالب كمحور لعملية التعلم والتعليم على فهم موقعه في العملية التعليمية التعلمية وبدء التعلم من حيث (هو فعلًا) لا من حيث (يفترض معلمه أن يكون) فنتاجات التعلم هرمية وتحقيق بعضها يتطلب تحقيق نتاجات تعلم أخرى سابقة لها، ويجب أن يعمل المعلم على قضاء جزء من الوقت في بداية كل فصل دراسي أو وحدة دراسية بعمل تقييم تشخيصي للطلبة لمعرفة مستواهم وما لديهم من معرفة سابقة تلزم للموضوعات المراد تناولها، ويستثمر نتائج التشخيص في توفير بيئة تعلم تضمن تجسير الفجوة بين نتاجات التعلم السابق ونتاجات التعلم الحالي في البنية المعرفية للمتعلم، ويتمثل دور المعلم في تقييم استعداد الطالب للتعلم الجديد (Readiness Assessment). وأن يستثمر نتائج التقويم القبلي (Prerequisites Evaluation) في اتخاذ قرارات تعليمية جوهرية، مثل إعادة التركيز على نتاجات التعلم السابق التي لم يتقنها الطلبة وتعديل خطة التدريس بما يناسب حاجات الطلبة، أو يعمل على تقسيم الطلبة في مجموعات تعلم متجانسة حسب درجة التمكن وطريقة التعلم (Placement Assessment).

كما ينبغي تعريف المتعلم بنتاجات التعلم التي ينبغي عليه السعي لتحقيقها، وبمعايير تقييم تحدد مدى اقترابه من تحقيقها، أي تعريف الطالب بمعايير المحتوى التي تحدد ما ينبغي عليه معرفته، وبمعايير الأداء التي تصف كيف يعد الأداء الجيد جيدًا بما فيه الكفاية، ولا بد من تحقيق الطلبة لنتاجات التعلم ضمن بيئة تعلم تتيح التقييم الذاتي وتقييم الأقران. وتستمر العملية التقييمية حتى الحصول على شواهد وأدلة تبين للمعلم أن الطلبة قد تعلموا نتاجات تعلم الوحدة الدراسية والكفايات المرتبطة بها. والتقييم من أجل التعلم متعدد الأبعاد وتكويني في طبيعته، ومدمج في المنهاج، وواقعي، ومرن. ويفيد في تحسين نتائج المتعلمين، ويزيد من ثقتهم بقدرتهم على الوصول إلى أهدافهم بالعمل الجاد، ويزيد من الاستقلالية، ويغير من ثقافة الفصل الدراسي. ويساعد التقييم من أجل التعلم في إنشاء فصل دراسي داعم وتعاوني يشعر فيه الجميع (المعلم وطلبته) بالقدرة على تجربة أشياء جديدة دون القلق من احتمال الفشل، ويتعلم الطلبة من أخطائهم.

5. ينبغي إشراك الطلبة في التقييم من خلال التقييم الذاتي وتقييم الأقران

Peer and self-assessment for students

التقييم الذاتي عملية حيوية يشارك فيها الطلبة بتقييم أدائهم استنادًا إلى معايير محددة. ويعزز التقييم الذاتي الاستقلالية في التعلم وتنمية مهارات التفكير والتنظيم الذاتي[35]. ويحدث التقييم الذاتي عندما يقوم المتعلم بتقييم أدائه. ومع الممارسة يتعلم المتعلم التفكير بموضوعية في تقدمه وتنمية مهاراته ونقد ذلك، وكيفية تحديد الفجوات في فهمه وقدراته، وكيفية تحسين أدائه، والتعلم بشكل مستقل والتفكير النقدي. ويفيد التقييم الذاتي في تطوير مهارات التعلم التي سيحتاجها الطلبة لتحقيق النتاجات المستهدفة، ولجعلهم أكثر وعيًا ومسؤولية عن عمليات التعلم الخاصة بهم. ويمكن للمعلم تنظيم التقييم الذاتي من خلال توجيه الأسئلة الآتية للطالب:

ما الدرجة التي تعتقد أنها عادلة للعمل الذي قدمته، ولماذا تعتقد ذلك؟

ما أفضل ما قمت به في مهمة التقييم هذه؟

ما الذي قمت به على الأقل بشكل جيد في مهمة التقييم هذه؟

ما الجزء الأصعب الذي وجدته؟

ما أهم شيء تعلمته أثناء القيام بمهمة التقييم هذه؟

إذا كان لديك المزيد من الوقت لإكمال المهمة، ما الذي ستغيره (إذا كان هناك أي شيء)، ولماذا؟

وعند استخدام التقييم كوسيلة للتعلم فإن التقييم الذاتي وتقييم الأقران يعد نهجًا فعالًا لتعزيز تعلم الطلبة. فتعليم الطلبة كيفية تقييم عملهم وعمل أقرانهم له العديد من الفوائد، فهو يعزز فهمهم لتعلمهم، ويوفر فرصًا للتحليل النقدي لجهودهم الخاصة، مما يشجعهم على أن يصبحوا متعلمين أكثر استقلالية (Autonomous Learners). ويعتمد التقييم الذاتي على دعم المعلمين الذين يقدمون الإرشادات والتوجيهات المناسبة. ويكون أكثر فعالية عندما تكون المعايير واضحة ويتم دعمه بأدوات تكنولوجية، وعندما يتم دمجه في المناهج.

ويحدث التقييم الذاتي عندما يسأل المتعلم نفسه بعد عملية التعلم، "ما الذي تعلمته؟"، "وما الذي أحتاجه لأتعلم أكثر؟"، "وهل أتعلم بالطريقة الصحيحة؟"، "وكيف يمكنني أن احكم على تعلمي؟". ويحدث تقييم الأقران عندما يقوم الطلبة بالتفكير في عمل أقرانهم، في ضوء معايير النجاح وتحقيق نتاجات التعلم، وتقديم ملاحظات بناءة لهم. فالتقييم الذاتي وتقييم الأقران يُمكّن الطلبة من أن يكونوا مسؤولين عن تعلمهم وأهدافهم

[35] Pandora, Brown and Strijbos. (2016). The future of student self-assessment.

التعليمية المستقبلية. وينبغي تعزيز الوعي بأهمية التقييم الذاتي كجزء من العملية التعليمية الشاملة، وتطوير أدوات ومنصات رقمية تدعمه بطرق مبتكرة وفعالة وتساعد الطلبة على تحسين أدائهم الأكاديمي.

6. ينبغي الاهتمام بجودة التقييم الصفي

وفقًا لمعايير الاختبارات التربوية والنفسية، يعد الصدق والثبات والعدالة من السمات الحاسمة لجودة أي تقييم. ومعظم هذه السمات إحصائية بطبيعتها، وتقديرها يحتاج إلى توافر مجموعة كبيرة من البيانات، كما هو الحال في التقييمات واسعة النطاق، ولا تنطبق المقاييس الإحصائية بطريقة مماثلة على التقييمات الصفية، لأن عدد الطلبة في التقييم الصفي محدود.

ويرتبط الصدق في التقييمات الصفية بتحقيقها لأغراضها، ويكون التقييم الصفي التكويني صادقًا بقدر ما يوفر تغذية راجعة مفيدة حول عمليتي التعليم والتعلم. لذلك، يرتبط الصدق هنا بصحة الاستدلالات التي يتم الوصول إليها حول نتاجات تعلم الطلبة، ويكون التقييم الصَّفي الختامي صادقًا بقدر تمثيله لمحتوى الكتب المدرسية والمهارات والكفايات المستهدفة في المناهج. ويتم ضمان ثبات التقييمات الصَّفية باستخدام إجراءات موحدة في تقييم أداء الطلبة، ومشاركة توقعات التقييم المحددة بوضوح مع الطلبة. وتتحقق عدالة التقييم بالاستجابة للخصائص الفردية للطلبة وتوفير الفرص لهم لإثبات تحقيقهم لنتاجات التعلم بشكل أفضل دون أي تحيز صريح أو ضمني من المعلم. ولكي تكون التقييمات عادلة، يجب أن يكون الطلبة على دراية بنتاجات التعلم ومحتوى التقييم وكيفية استخدام نتائجه.

وإجراء التقييمات بشكل صحيح، يوفر المعلومات (البيانات) الأكثر ثباتًا وصدقًا وعدالة عن كل طالب. وتركز جودة التقييمات في الفصول الدراسية على حصول الطلبة على تغذية راجعة مفيدة في تحسين تعلمهم وفي الوقت المناسب، وإجراء مناقشات حول عملهم دون القلق بشأن الأداء والدرجات. كما تركز على بناء علاقة ودية بين المعلم وطلبته، يدركون من خلالها أنه يهتم حقًا بتعلمهم، ويشركهم ويمنحهم الفرصة لتحمل مسؤولية تعلمهم.

ويتطلب تحقيق مؤشرات الصدق والثبات والعدالة تحديد الغرض من التقييم، وما الذي يتم تقييمه، وكيف يتم تقييمه، وكيف يمكن التعبير عن النتائج واستخدامها. ويعتمد ذلك على كفاءة المعلم في تصميم وبناء المهام التقييمية. كما يجب أن يتناول الحكم على جودة التقييم الصَّفي أثر النتائج على المتعلم ومدى إسهامها في تحسين مستوى تعلمه، فنموذج التقييم الذي له آثار عكسية على التعلم أو على الطلبة، تكون نتائجه غير

صادقة، ويقع على عاتق المعلم التحقق من صدق نتائج أدوات التقييم الصَّفي كالواجبات المنزلية والاختبارات القصيرة والمشاريع. وهناك مجموعة من المعايير التي ترتبط بصدق نتائج التقييمات الصَّفية تتعلق بتمثيل المحتوى، وتمثيل المهارات وعمليات التفكير ، والثبات والموضوعية، والعدالة للفئات كافة، والاتساق مع التقييمات الأخرى، والاستخدام الفعَّال للنتائج. ويعرض الجدول 10 معايير التحقق من صدق نتائج أدوات التقييم الصفي ومؤشراتها.

جدول 10: معايير التأكد من صدق نتائج أدوات التقييم الصفي ومؤشرات الأداء عليها

المعايير	مؤشرات الأداء على صدق التقييم الصفي
1. تمثيل المحتوى وأهميته	1. يركز على ما تم تدريسه. 2. يمثل محتوى المنهاج المدرسي الرسمي. 3. يتضمن المحتوى المجدي تعلمه. 4. يمثل الاتجاه المعاصر حول تعلم المادة الدراسية.
2. تمثيل المهارات وعمليات التفكير	1. يتطلب من الطالب استعمال العديد من مهارات التفكير وبشكل متكامل. 2. يمثل مهارات وعمليات التفكير المنصوص عليها في المنهاج المدرسي الرسمي. 3. يحتوي مهام لا يمكن للطالب إكمالها دون استخدام مهارات التفكير المطلوبة. 4. يسمح بوقت كافٍ للطلبة لاستعمال مهارات وعمليات تفكير مركبة ومعقدة.
3. الاتساق مع أدوات التقييم الصفي الأخرى	1. يعطي نمطًا من النتائج يتسق مع نتائج التقييمات الصفية الأخرى. 2. يحتوى على أسئلة متدرجة في الصعوبة وبوسط متوسط للصعوبة.
4. الثبات والموضوعية	1. استخدام طريقة منظمة لكل الطلبة من حيث منحهم الدرجات وترتيب الإتقان. 2. يوفر لكل طالب فرصًا عديدة للتعبير عن قدراته في كل نتاج تعلم يتم تقييمه.
5. العدالة لمختلف الفئات	1. خاليًا من التحيز. 2. يراعي الطلبة ذوي الإعاقة أو صعوبات التعلم. 3. أن تكون التقييمات مفهومة للجميع. 4. ضمان التكافؤ في التقييمات عبر السنين
6. الاقتصادية والفعالية وإمكانية التطبيق	1. يتطلب من الطلبة وقتًا معقولًا للإعداد والتقدم له. 2. يمثل استخدامًا ملائمًا لوقت الطالب والمعلم في الحصة.
7. التنويع في أشكال التقييمات	1. استخدام أشكال متعددة من التقييمات لإصدار أحكام وقرارات مهمة.

7. ينبغي أن تهتم التقييمات الصَّفية بقياس مهارات التفكير العليا

للتقييم الصَّفي دور أساسي، كنماذج التعلم، في اكتساب المتعلمين لمهارات التفكير العليا، وامتلاك القدرة على التفكير أمر محوري في التعلم، لأنها تشجع النمو الفكري وتعزز التحصيل الأكاديمي الأفضل. وعلى المعلم تطوير التقييمات التي تتطلب القدرة على التفكير. ويعتمد ذلك على كفاءته في تصميم وبناء فقرات الاختبار أو المهام التقييمية. وقد أظهرت دراسات كثيرة بأنه نادرًا ما يتم طرح الأسئلة التي تقيس القدرة على التفكير عالي المستوى في التقييمات الصفية.

وتصنف بروكهارت[36]تعريفات التفكير عالي المستوى في ثلاث فئات: (1) فئة تُعرفه بدلالة النقل Transfer، و(2) وفئة تُعرفه بدلالة التفكير النقدي Critical Thinking، و(3) وفئة تُعرفه بدلالة حل المشكلة Problem Solving. ويتطلب قياس مهارات التفكير المتصلة بالنقل تصميم مهام تقييمية تتطلب من الطلبة ليس فقط أن يتذكروا ما تعلموه، بل أيضًا فهمه وأن يكونوا قادرين على استخدامه في مواقف جديدة وبما يعزز التعلم ذي المعنى. وتتضمن فئة التفكير النقدي تعريفات تشير إلى "التفكير التأملي المعقول" الذي يركز على تقرير ما يجب الاعتقاد به أو فعله، و"التفكير الفني" Artful Thinking، الذي يتضمن الاستدلال، والتساؤل والتحقيق، والملاحظة والوصف، والمقارنة والربط، والتركيب، واستكشاف وجهات النظر، والحكم على مصداقية الأشياء أو المصادر، وتحديد الافتراضات والتعميم والتحيز. كما ينبغي أن تؤكد التقييمات على حل المشكلات باعتبارها مهارات عقلية عليا. ويحتاج حل المشكلة إلى استخدام واحدة أو أكثر من عمليات التفكير العليا مثل التعلم مع الفهم، والتقييم النقدي للأفكار، وصياغة البدائل الإبداعية، والتواصل بشكل فعال. وهذا ما يجب أن تبنى حوله بعض المهام التقييمية.

فعلى المعلم أن يبني تقييمات لمهارات التفكير العليا. وعليه أن يكتسب مهارة كيفية تطوير واستخدام أسئلة الاختبار والتقييمات الأخرى التي تكشف مدى قدرة الطلبة على التحليل والتفكير وحل المشكلات والتفكير بشكل إبداعي. والإطار العملي في تقييم مهارات التفكير العليا يتضمن الخطوات التالية: (1) التحديد الواضح لنوع التفكير الذي تريد تقييمه؛ (2) وتصميم مهمة أو فقرة اختبار تتطلب من الطلبة إظهار نوع التفكير الذي تبحث عنه؛ (3) وتحديد كيفية تفسير النتائج أو تقييمها أو تسجيلها؛ و (4) وتوفير النوع المناسب من المواد لمساعدة الطلبة على إظهار تفكيرهم.

[36]Brookhart, S. (2010), How to Assess Higher Order Thinking Skills in Your Classroom, ASCD.

7. ينبغي أن يهتم التقييم الصفي بالتقييم الواقعي (Authentic Assessment)

لا بد من أن تتبنى التقييمات الصفية للطلبة التقييم الواقعي (Authentic Assessment) الذي يعكس إنجازات الطلبة ويقيسها في مواقف حقيقية، ويجعلهم ينغمسون في مهمات ذات قيمة ومعنى بالنسبة لهم، فتبدو التقييمات الصفية لهم كنشاطات تعلم، يمارسون فيها مهارات التفكير العليا، ويوائمون بين مدى واسع من المعارف لبلورة الأحكام أو اتخاذ القرارات أو حل المشكلات الحياتية التي يعيشونها في الواقع. وبذلك تتطور لديهم القدرة على التفكير التأملي (Reflective Thinking) الذي يساعدهم على معالجة المعلومات ونقدها وتحليلها؛ فهو يوثق الصلة بين التعلم والتعليم ويعمل على توجيه التعليم بما يساعد الطالب على التعلم مدى الحياة، وفي الأدب التربوي، هناك مرادفات أخرى للتقييم الواقعي مثل تقييم الأداء Performance Assessment أو التقييم البديل Alternative Assessment أو التقييم المباشر Direct Assessment. ويقوم التقييم الواقعي على عدد من الأسس والمبادئ التي يجب مراعاتها عند تطبيقه. ولعل أبرز هذه المبادئ ما يأتي:

1. التقييم الواقعي إجراء يرافق عمليتي التعليم والتعلم ويربطهما معاً بقصد تحقيق كل طالب لمستويات الأداء المطلوبة، ويوفر التغذية الراجعة الفورية حول إنجازات الطالب بما يكفل تصويب مسيرته التعليمية ومواصلة عملية التعلم. فهو يهتم بجوهر عملية التعلم، أي يركز على ما يهم حقا. وهو بذلك تقييم تكويني يستند إلى عدد من المعايير، ويجعل تمكّن الطالب منها هدفاً منشوداً لعملية التعليم والتعلم. فالعمليات العقلية ومهارات التقصي والاكتشاف هي غايات يجب رعايتها عند الطلبة والتأكد من اكتسابهم لها، ولا يتسنى ذلك إلا بإشغالهم بنشاطات تستدعي حل المشكلات وبلورة الأحكام واتخاذ القرارات بما يتناسب ومستوى نضجهم.
2. التقييم الواقعي يقتضي أن تكون المشكلات والمهام المطروحة للدراسة والتقصي واقعية، وذات صلة بشؤون الحياة العملية التي يعيشها الطالب في حياته اليومية.
3. أداء الطلبة هو مادة التقييم الواقعي وليس حفظهم للمعلومات واسترجاعها، ويقتضي ذلك أن يكون التقييم الواقعي متنوعًا في أساليبه وأدواته لمراعاة الفروق الفردية بين الطلبة، وألا تحتل الاختبارات فيه سوى حيز ضيقٍ، دون قلق أو رهبة.
4. يقوم التقييم الواقعي على التعاون بين الطلبة. ولذلك فإنه يتبنى أسلوب التعلم في مجموعات متعاونة. ويهيئ لجميع الطلبة فرصًا أفضل للتعلم، ويهيئ للمعلم فرصة تقييم أعمال الطلبة ومساعدة ذوي الإحتياجات الخاصة بينهم وفق الاحتياجات اللازمة لكل حالة.

5. التقييم الواقعي محكي المرجع يقتضي تجنب المقارنات بين الطلبة والتي تعتمد أصلاً على معايير أداء الجماعة والتي لا مكان فيها للتقويم الواقعي.

8. ينبغي التخطيط للتقييم الصفي

تعد خطط التدريس غير مكتملة ما لم تشتمل على خطط التقييم. وعند التخطيط للتقييم يجب أن يؤخذ في الاعتبار طول الفترة التي سيغطيها التقييم (درس أو وحدة أو فصل دراسي) لأن الخطط التي تستهدف فترة طويلة (فصل دراسي) تكون أقل تفصيلًا من الخطط التي تستهدف فترة قصيرة (وحدة أو درس)، وأن صدق النتائج موقفي ونسبي، ولا توجد طريقة واحدة للتقييم تعطي نتائج تامة الصدق، لذا يتعين على المعلم استعمال أكثر من طريقة لتقييم إنجاز الطلبة.

ويجب أن تكون خطة التقييم لفصل دراسي متسلسلة وفق تسلسل وحدات الكتاب، وتحتوي الموضوعات العامة التي ستدرس، وتظهر نتاجات التعلم التي سيحققها الطلبة لكل وحدة، والإطار الزمني لتنفيذها، وأنشطة التقييم التكويني، والتقييم الختامي، ووزن كل منها في الوحدة، وأخيرًا أوزان الوحدات التي يشملها الفصل الدراسي في العلامة الكلية للمبحث. وعلى المدرس عند التخطيط للتقييم الصفي أن يضع ما يلي في الاعتبار:

- بناء ثقافة تعزز التعاون بينه وبين طلابه.
- إشراك الطلبة في عملية التعلم والتقييم.
- وضوح نتاجات التعلم لدى الطلبة.
- ملاءمة التقييم للغرض.
- أن يكون الطلبة على دراية بكيفية التقييم ومعاييره.
- متابعة التقدم الفردي لكل طالب.
- الموازنة بين المعرفة والتطبيق ومهارات التفكير العليا.
- مراعاة معايير كتابة الأسئلة الجيدة.
- الاستفادة من أفضل الممارسات في تصميم تقييمات واقعية.

الخلاصة والمقترحات بخصوص التقييمات الصَّفية للطلبة في الأردن:

نظرًا لأهمية التقييم الصفي، ولكونه الأكثر ارتباطًا بتعلم الطلبة، فإنه يتطلب أن يكون المعلم على درجة عالية من التأهيل، لتنفيذ التقييمات الصفية بجودة عالية، تضمن تحقيق أغراضها التشخيصية والتكوينية والختامية. لذلك نوصي بالآتي:

1. دعوة مؤسسات التعليم العالي ومؤسسات التدريب وهيئة الاعتماد من أجل مراجعة الخطط الدراسية لبرامج إعداد المعلمين قبل الخدمة التي تطرحها على مستوى البكالوريوس أو الدبلوم أو الماجستير في التخصصات التربوية لتضمينها مساقات تتعلق بالجوانب النظرية والعملية في مجال القياس والتقويم حيث تبين للفريق أن الخطط الحالية تفتقر لمثل هذه المساقات.

2. تدريب المعلمين على كيفية تقييم المهارات العقلية العليا لدى طلبتهم، وتشجيعهم على التركيز على تقييم هذه المهارات لدى طلبتهم، ما يقيّمه المعلم وكيف يقيمه ويوصل بصورة قوية لطلبته ما يقدّره حقًا في تعلمهم.

3. تشجيع المعلمين ضمن كل مديرية على تبادل التقييمات فيما بينهم مما يقلّل من التباين في الممارسات التقييمية الصَّفية، ويَجْدر بمشرف القياس متابعة ذلك.

4. تشجيع المعلمين على الاستفادة من تجربة تقييم الطلبة في برنامج BTEC، وذلك بزيادة جرعة استخدام التعيينات بأشكالها المختلفة سواء الفردية أو الجماعية وقياس مهارات التفكير العليا وحل المشكلات، وبشكل خاص في صفوف المرحلة الأساسية العليا والمرحلة الثانوية.

5. إدماج التكنولوجيا في التقييمات المختلفة، وبناء بنك مركزي للمهام التقييمية، والاستفادة من ذلك في قياس مجموعة واسعة من الكفايات مثل معالجة المعلومات والعمل الجماعي والتحليل النقدي وغيرها من مهارات القرن الحادي والعشرين.

6. تشجيع المعلمين وتدريبهم على كيفية الاستفادة من التقييمات الدولية في ممارساتهم التعليمية والتقييمية من خلال مساعدتهم في تقديم المواقف والسيناريوهات التي تحاكي ما هو موجود في تلك التقييمات، ومتابعة أداء الطلبة فيها.

7. تطوير الوزارة لبرامج تدريبية للمعلمين أثناء الخدمة في مجالات القياس والتقويم، واعتماد تلك البرامج ضمن جهود الوزارة في تنمية النمو المهني للمعلمين وضمن متطلبات نظام رتب المعلمين.

8. تعيين أخصائي قياس وتقويم واحد على الأقل في كل مديرية من مديريات التربية والتعليم من حملة درجة الماجستير أو الدكتوراه في تخصص القياس والتقويم تكون مهمته متابعة عمل لجان التقييم في المدارس والممارسات التقييمية للمعلمين وتوفير المساعدة التي يحتاجون إليها.

9. تشكيل لجنة للتقييم في كل مدرسة ممن لديهم الخبرة الكافية في الممارسات التقييمية، وتتكون من 3 معلمين[37] على الأقل وحسب حجم المدرسة، برئاسة مدير المدرسة تكون مهمتها مساعدة المعلمين في تقييماتهم للطلبة والإشراف على تنفيذ التقييمات الوطنية والدولية على مستوى المدرسة.

[37] يتم تخفيض النصاب لهم

10. إعادة النظر في أسس النجاح والرسوب المعمول فيها حاليًا لتواكب التقييمات القائمة على المعايير والنتاجات، أي لا يجوز التحديد المسبق لنسبة الناجحين.
11. اعتماد نماذج موحدة للأدوات التي يستخدمها المعلم في بناء الاختبارات (تحليل محتوى وجداول مواصفات لاختبارات كل مبحث لكل صف، ...) من قبل الوزارة والمركز الوطني لتطوير المناهج، ويتأكد من تطبيقه كل من لجنة التقييم في المدرسة ومشرف القياس والمشرفون المختصون في المديرية.

3. التقييمات الوطنية (صفوف الثالث والسادس والتاسع)

National Assessments (3rd, 6th, 9th Grades)

التقييم الوطني هو مسح للمدارس والطلبة والمعلمين، مصمم لتقديم أدلة على مستوى نظام التعليم، فيما يتعلق بإنجازات الطلبة في مرحلة معينة من التعليم، في مجالات دراسية محددة، وتلعب التقييمات الوطنية دورًا حاسمًا في إظهار فعالية النظام التعليمي. لذلك لا بد من إجراء تقييمات وطنية وتخصيص الموارد لها، واستثمار نتائجها في إصلاح التعليم وتحسين التعلم، ويكمن الاهتمام الأساسي في التقييمات الوطنية فيما تخبرنا به عن أداء نظام التعليم، وقياس النمو الفردي للطلبة، وتزداد الفائدة من التقييم الوطني، طالما كانت النتائج قابلة للمقارنة ويمكن استخدامها في تحليل التوجهات[38] .

وتؤثر قضايا الجودة Quality والإنصاف Equity والتوفر Provision في التقييمات الوطنية على عملية تصميمها، ومجال التركيز فيها، والمجتمع الذي تستهدفه. فهل يجب التركيز على محتوى المنهاج أم المهارات الحياتية؟ وما المجموعات الفرعية التي سيتم مقارنتها؟ وما آثار بعض السياسات المتبعة؟ وهناك قرارات مهمة يجب اتخاذها في التقييمات الوطنية تتعلق بمن يجب تقييمهم (المجتمع أم عينة ممثلة عنه)، وفي أي مرحلة (أي صف أو أي عمر)، وما الذي يجب تقييمه (نتاجات تعلم في منهاج معين أم الكفايات في مهارات الحياة)، وما دورية التقييم وتكراره (سنوي أو كل 3 سنوات)، وكيف يتم التعبير عن النتائج واستخراج التقارير حولها (العلامة الخام أم العلامة القياسية). وتتمثل الأغراض الأساسية للتقييمات الوطنية في الأردن في تسليط الضوء على الفجوات في الإنجاز، وتتبع التقدم الوطني بمرور الوقت، ومقارنة تحصيل الطلبة داخل الأردن على مستوى المدرسة والمديرية والإقليم وغيرها من المجموعات الفرعية للطلبة، ويمكن استخدامها بحذر شديد لأغراض مساءلة المعلم والمدرسة.

[38]National and international assessments of student achievement (Guideance note: A DFID paper)

ولتحقيق هذه الأغراض ينبغي أن تكون أداة التقييم الوطني معيارية موحدة، وكذلك إجراءات التنفيذ والتصحيح، وأن تستخدم العلامة القياسية (Scale Score) للتعبير عن نتائج الطلبة. والعلامة القياسية هي علامة معيارية على مقياس موحد ضمن مدى متفق عليه بوسط حسابي محدد وانحراف معياري محدد، وتستخدم في عمل مقارنات مباشرة بين الطلبة، وعبر مجموعات فرعية منهم، وبمرور الوقت. والعلامة القياسية أفضل من العلامة الخام لكون معنى العلامة الخام يختلف باختلاف الاختبار أو حتى باختلاف نماذج الاختبار نفسه لاختلافها في صعوبة الفقرات، لذلك لم يعد الاهتمام بعدد الفقرات التي يجيبها المتعلم بصورة صحيحة كافيًا للتعبير عن أدائه، بل يتعدى ذلك إلى نمط الاستجابات الذي يأخذ صعوبة الفقرات بالاعتبار، وتجدر الإشارة في تحليل البيانات وتطوير أدوات التقييم، إلى أن نظرية الاختبار التقليدية وخصائص الفقرة الشائع استخدامها لها محددات تحول دون إنتاج مقياس ذي معنى للتحصيل. لذلك، لا بد من استخدام نظرية الاستجابة للفقرة (IRT). وهي نموذج رياضي متطور يسمح بوضع قدرة الفرد على مقياس ذي معنى يأخذ بالاعتبار صعوبة وسهولة الفقرات المتضمنة في الاختبار، كما تفيد في معايرة الفقرات وتقدير خصائصها وفي معادلة النماذج المختلفة للاختبار وبناء المقياس الموحد للعلامات القياسية.

وفي معرض تفسير التباين في الدرجات لأي تقييم، يجب توفير معلومات سياقية تتعلق بالمتعلم ووضع أسرته وبيئة مدرسته ومعلمه، ويتطلب ذلك تصميم وبناء استبانات خاصة لجمع المعلومات من الطالب والمعلم ومدير المدرسة وولي الأمر، كما يجب استخدام معايير الأداء لوصف أداء الأفراد، ومعيار الأداء هو تعريف متفق عليه لمستوى معين من الأداء في مجال معين من المحتوى ويتحدد بدلالة ما يعرفه الفرد أو يستطيع عمله، ويجب تحديد درجات القطع التي تفصل بين أي مستويين متتالين للأداء بطريقة علمية باستخدام أحد الأساليب المعتمدة عالميًا (مثل أسلوب أنجوف المعدل Modified Angoff Method).

ويتم تصنيف نتائج التقييم في مستويات للأداء، التي قد تختلف في عددها ومسمياتها من نظام تقييمي إلى آخر، وتبنى هذه المستويات بشكل هرمي على المتصل الكامن الذي يمثل السمة المقاسة، ويكون كل مستوى مصحوبًا بوصف نوعي لما سيكون عليه الأداء في ذلك المستوى، وقد قامت مجموعة من المؤسسات العالمية[39] بإعداد إطار الكفاءة العالمي (Global proficiency framework, GPF)، ليصف الحد الأدنى العالمي لمستويات الكفاءة المتوقعة من الطلبة. ويعد إطار الكفاءة العالمي مرجعًا عالميًا مفيدًا في فهم ما يجب

[39] USAID, WB, UNESCO UIS, UK's FCDO and others, 2019.

أن يكون الطلبة قادرين على تحقيقه في نموهم على أساس عالمي في ضوء المواد الدراسية المناسبة والتعليم المناسب، وفي مراجعة كل من المعايير والمناهج والمواد الدراسية وتدريب المعلمين والتقييمات المختلفة للطلبة.

وانسجامًا مع إطار الكفاءة العالمي نقترح استخدام ثلاثة مستويات هرمية للنجاح في إطار التقييمات الوطنية، وهي مستوى الإتقان الجزئي، ومستوى الإتقان التام، والمستوى المتقدم. ويصنف الذين لا يصلون إلى هذه المستويات ضمن المستوى المبتدئ (المستوى غير المرضي والذي يحتاج الطالب فيه إلى تحسين تعلمه). ويبين الشكل 8 هذه المستويات للأداء. وتبدأ هذه المستويات من الإنجاز المنخفض إلى الإنجاز المرتفع.

الشكل 8: مستويات الأداء الهرمية

وفيما يلي وصف لهذه المستويات.

المستوى المبتدئ: لا يفي المتعلم بالحد الأدنى من الكفاءة، ويفتقر المتعلمون إلى المعرفة والمهارات الأساسية، ونتيجة لذلك لا يمكنهم إكمال المهام الأساسية المطلوبة.

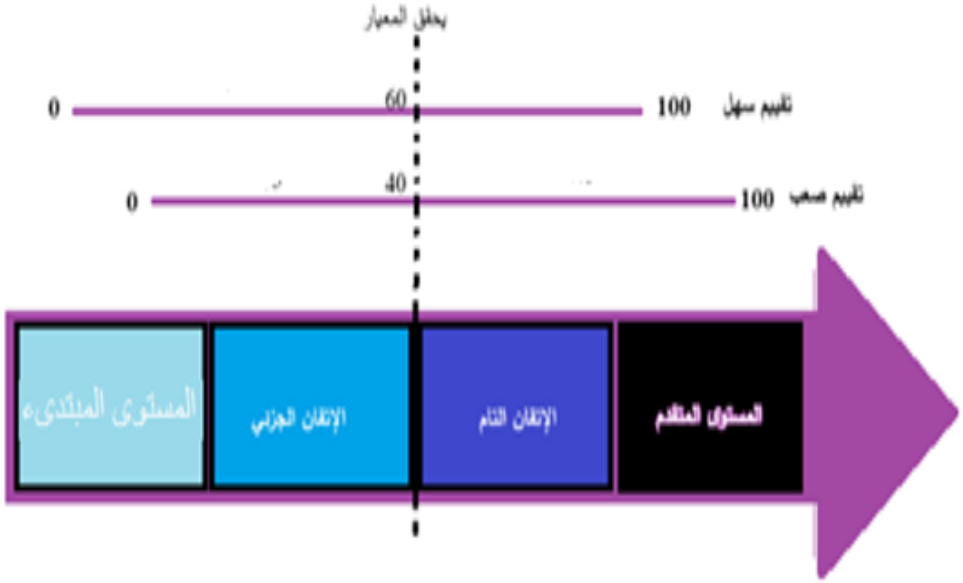
مستوى الإتقان الجزئي: يفي المتعلم جزئيًا بالحد الأدنى من الكفاءة، ويكون لدى المتعلمين معرفة ومهارات محدودة، ونتيجة لذلك، يمكنهم إكمال المهام الأساسية المطلوبة جزئيًا، وقليلا ما تكون لديهم القدرة على نقل معرفتهم ومهاراتهم إلى سياقات جديدة.

مستوى الإتقان التام: يفي المتعلم بالحد الأدنى من الكفاءة، ويكون لدى المتعلمين المعرفة والمهارات الكافية، ونتيجة لذلك يمكنهم إكمال المهام الأساسية المطلوبة بنجاح، وتكون لديهم القدرة على نقل معرفتهم ومهاراتهم إلى سياقات جديدة مماثلة لسياق التعلم.

المستوى المتقدم: يتجاوز المتعلم الحد الأدنى من الكفاءة، ويكون لدى المتعلمين معرفة ومهارات فائقة، ونتيجة لذلك يمكنهم إكمال المهام المعقدة بنجاح، وتكون لديهم القدرة على نقل أثر التعلم إلى سياقات جديدة مختلفة عن سياق التعلم.

تمثل هذه المستويات مقياسًا مشتركًا، يمكن استخدامه من أجل مقارنة نتائج تقييم الطلبة عبر المديريات والمدارس، وفي تتبع نتائج التقييم بمرور الوقت لرصد التقدم المحرز، ولتحديد المعايير سيتم استخدام أسلوب أنجوف المعدل الذي يركز على فقرات الاختبار، ومن المرجح أن يكون لتقييمين مختلفين معايير مختلفة للحد الأدنى من الكفاءة بسبب اختلاف صعوبتهما. وسيكون معيار الحد الأدنى لأي مستوى أداء، أعلى في التقييم السهل منه في التقييم الصعب. فمثلًا إذا كان هناك تقييمان الأول سهل ويتكون من 100 فقرة والثاني صعب ويتكون من 100 فقرة، وهناك شخص يمتلك الحد الأدنى لمستوى الإتقان التام، فإن احتمال إجابته لفقرات التقييم السهل أعلى من احتمال إجابته لفقرات التقييم الصعب. وإذا افترضنا أن الوسط الحسابي لاحتمالات إجابة فقرات التقييم السهل هو 0.60 والوسط الحسابي لاحتمالات إجابة فقرات التقييم الصعب هو 0.40، فإن معيار الإتقان التام في الاختبار السهل هو إجابة 60 فقرة وفي الاختبار الصعب هو إجابة 40 فقرة، كما هو مبين في الشكل 9.

الشكل 9: معيار الحد الأدنى للإتقان التام في تقييم سهل وما يعادله في تقييم صعب.



وهذا يدلل على وجود أثر لصعوبة فقرات التقييم في تحديد المعايير المطلوبة للنجاح، لذلك يجب إزالة مثل هذا التأثير من خلال معادلة الدرجات على النماذج المختلفة للامتحان.

ونظرًا لكون تحصيل الطلبة، من المؤشرات الأساسية في تقييم نوعية التعليم، فقد أعدت وزارة التربية والتعليم ممثلة بإدارة الامتحانات والاختبارات امتحانات وطنية ضمن خطة مدروسة لصفوف معينة (4، 8، 10) في مباحث محددة بهدف ضبط نوعية (جودة) التعليم، عن طريق قياس مهارات التعلم الأساسية المتوقع أن يمتلكها الطلبة في عينات مختارة من هذه الصفوف (وزارة التربية والتعليم،2007). وقد يكون التبرير المحتمل لاختيار هذه الصفوف بعينها هو كونها مستهدفة في التقييمات الدولية (تيمس، وPISA)، مما يُمكّن من إجراء مقارنة بين نتائج التقييمات الوطنية وما يناظرها من نتائج التقييمات الدولية، وبالتأكيد، مثل هذه المقارنة غير ممكنة لأن المقارنة العادلة تشترط استخدام نفس أداة التقييم أو استخدام المجموعة نفسها من الأفراد في كلا النوعين من التقييمات، وكلا الشرطين غير متحقق، وعلاوة على ذلك تختلف التقييمات الوطنية عن التقييمات الدولية في طريقة التعبير عن النتائج، حيث تستخدم التقييمات الوطنية العلامات الخام بينما تستخدم التقييمات الدولية العلامات القياسية، وتقوم الاختبارات الوطنية على أساس العينات، وهذا مناسب للغرض المعلن وهو تقييم نوعية التعليم، ويوفر بيانات تقييمية على مستوى النظام التعليمي والمديرية، لكنه لا يوفر بيانات تقييمية على مستوى كل من المدرسة والطالب، ولا يفيد في قياس التقدم لدى الطالب، ولا يمكن استخدام النتائج لأغراض المساءلة سواءً للمدرسة أو للمعلم.

لذلك نقترح أن تكون الصفوف المستهدفة في التقييمات الوطنية هي الصف الثالث والصف السادس والصف التاسع الأساسية. وتم اختيار هذه الصفوف لكونه تم إقرار أن يكون التشعيب إلى المسارين الأكاديمي والمهني بعد الصف التاسع الأساسي. كما أن الفجوة بين الصفوف المختارة هي 3 سنوات وهي تعادل الفجوة بين دورات الامتحانات التي سيتم تطبيقها كل ثلاث سنوات، ويخدم هذا النهج غرضين في وقت واحد، الغرض الأول هو إمكانية تتبع المجموعة نفسها من الطلبة من الصف الثالث حتى الصف التاسع في المبحث نفسه، وعليه يمكن تقييم التقدم الحاصل من مستوى تعليمي إلى آخر، والغرض الثاني هو إمكانية تنفيذ تدخلات علاجية بصورة منتظمة وتتبعها عبر السنوات لأغراض تقويمية.

وبالرغم من أن الاختبارات القائمة على أساس العينة أقل تكلفة، فإننا نرى أن تقوم جميع التقييمات الوطنية على التعداد الشامل لطلبة الصف المستهدف. وهذا ضروري إذا كان صانعو السياسات/ القرارات يسعون للحصول على معلومات حول الطلبة في المدارس الأردنية جميعهم من أجل توفير البيانات التشخيصية لكل مدرسة (أي على مستوى المدرسة) ولكل طالب (أي على مستوى الطالب)، والتخطيط لتدخلات إصلاحية

أو اتخاذ قرار بشأن تخصيص الموارد للمدارس التي تعاني من مشاكل، ولإعلام المجتمعات المحلية بالأداء الفردي لكل مدرسة، وإعلام أولياء الأمور بأداء ابنائهم، ولوضع نظام للمساءلة. ويجب أن يكون صانعو السياسات/القرارات حذرين جدًا في استخدام نتائج التقييمات الوطنية في تقويم المدارس والمعلمين وعلى دراية بالعواقب السلبية المحتملة الناتجة عن فرض العقوبات اعتمادًا على نتائج أداء الطلبة. ويبين الجدول 11 التقييمات الوطنية المقترحة حسب الصف.

الجدول 11: التقييمات الوطنية المقترحة حسب الصف

الصف	المباحث	النطاق	الدورية	السنة								
				2025	2026	2027	2028	2029	2030	2031	2032	2033
الثالث	لغة عربية	المجتمع	سنوي	●	●	●	●	●	●	●	●	●
	رياضيات	المجتمع	سنوي	●	●	●	●	●	●	●	●	●
السادس	لغة عربية	المجتمع	سنوي	●	●	●	●	●	●	●	●	●
	لغة إنجليزية	المجتمع	3 سنوات	●			●			●		
	رياضيات	المجتمع	3 سنوات		●			●			●	
	علوم	المجتمع	3 سنوات			●			●			●
التاسع	لغة عربية	المجتمع	سنوي	●	●	●	●	●	●	●	●	●
	لغة إنجليزية	المجتمع	سنوي	●	●	●	●	●	●	●	●	●
	رياضيات	المجتمع	سنوي	●	●	●	●	●	●	●	●	●
	علوم	المجتمع	سنوي	●	●	●	●	●	●	●	●	●

إجراءات تطوير التقييمات الوطنية وإدارتها:

تتلخص إجراءات تطوير التقييمات الوطنية وإدارتها بما يلي:

1. إعداد الدليل الشامل للتقييم الوطني المعني، ويتضمن ذلك الغرض من التقييم، ونتاجات التعلم أو الكفايات المراد قياسها، وبما يتلاءم مع المنهاج الدراسي ومهارات القرن الحادي والعشرين. كما يتضمن مخطط الامتحان وجدول مواصفاته.

2. بناء بنوك أسئلة خاصة بالتقييمات الوطنية، مع مراعاة المعايير الدولية في بناء البنوك حيث تتطلب تحليلًا للمحتوى وتحديدًا لنتاجات التعلم والكفايات المستهدفة في المناهج الأردنية، وكتابة الفقرات وفق الإرشادات والقواعد المعتمدة عالميًا، ومراجعتها وتدقيقها وتجريبها وتقدير خصائصها والاحتفاظ بالفقرات التي تحقق المعايير المعتمدة من حيث الصعوبة والتمييز وغيرها من المعايير.

3. تصميم التقييم المعني وبناء أدواته، ويتضمن ذلك اختيار الفقرات وبناء نماذج الاختبار وفق جدول المواصفات، وتحديد مستويات الأداء. كما يتضمن إعداد الاستبانات اللازمة لجمع البيانات والمعلومات السياقية حول الطالب ومدرسته وأسرته التي يتوقع ارتباطها بالتحصيل.

4. عمل الترتيبات الخاصة بإخراج أدوات التقييم وتطبيقها، ويتضمن ذلك ترتيب الفقرات في الاختبار، والتحقق من مطابقة نماذج الاختبار لجدول المواصفات، وكتابة تعليمات الاختبار، وإجراءات التطبيق للاختبار في ظروف معيارية موحدة، مع ضرورة مراعاة احتياجات ذوي الاحتياجات الخاصة.

5. تحضير البيانات وتجهيزها، ويتضمن ذلك تصحيح الاختبار بموضوعية عالية، ويتم تصحيح فقرات الاختيار من متعدد إلكترونيًا، وتجهيز الاستجابات لجميع فقرات الاختبار. كما يتضمن جمع البيانات الخاصة بالمعلومات السياقية حول الطالب وأسرته ومدرسته التي تفيد في تفسير التباينات في أداء الطلبة والمدارس.

6. تحليل البيانات ومعادلة النماذج وتدريجها على سلم موحد باستخدام نظرية استجابة الفقرة. واستخراج النتائج، ويتضمن ذلك استخراج الإحصاءات الوصفية للفقرات وللاختبار وتحديد العلاقات بين التحصيل والمعلومات السياقية، وكذلك مؤشرات ثبات الاختبار وغيرها من المؤشرات السيكومترية.

7. كتابة التقارير على المستويات كافة ونشرها وإيصالها للمعنيين مع توضيح لكيفية استخدام ما تتضمنه من معلومات في رسم السياسات وتحسين تعلم الطلبة وأداء المدارس.

8. إتاحة فرصة الوصول للبيانات من قبل الباحثين وطلبة الدراسات العليا دون الكشف عن هوية الطلبة وذلك لأغراض بحثية وتقييمية.

أولاً: التقييم الوطني للصف الثالث الأساسي

اختبار سنوي في مبحثي اللغة العربية والرياضيات. ويشتمل اختبار كل مبحث على نموذجين متكافئين في كل منهما 50 فقرة اختيار من متعدد (30 فقرة فريدة خاصة، 15فقرة مشتركة، 5 فقرات للتجريب). وتكون مدة اختبار الرياضيات ساعتين، ومدة اختبار اللغة العربية ساعتين من ضمنها 15 دقيقة لتعبئة استبانة الطالب. ويستهدف التقييم جميع طلبة الصف الثالث الأساسي. وسيُعد المركز الوطني للمناهج بالتعاون مع المختصين في إدارة الامتحانات والاختبارات في وزارة التربية والتعليم الاختبار وفقا لمعايير بناء الاختبارات

والمقاييس النفسية. وسيغطي التقييم المهارات الأساسية في المبحثين مع تغطية مناسبة لمهارات التفكير العليا فيهما وبما لا يقل عن 80% من مجمل الاختبار لتلك المهارات ومستوى التطبيق.

وسيتم إعلام الطلبة ومعلميهم والمهتمين كافة بمخطط كل اختبار، بما يتضمنه من المجالات الرئيسة للمحتوى الذي يقيسه، ومجالاته الفرعية، ونتاجات التعلم، وسيتم انتقاء فقرات كل اختبار من بنك أسئلة خاص بذلك الاختبار، أعدت فقراته وتم مراجعتها وتدقيقها وتجريبها ومعايرتها باستخدام نظرية الاستجابة للفقرة، وفقًا لجدول مواصفات محدد، يبين عدد فقرات الاختبار وتوزيعها حسب العمق المعرفي للفقرة ومستوى صعوبتها. وستشرف إدارة الامتحانات والاختبارات على تطبيق التقييم وتصحيحه، وتتولى لجنة التقييم في المدرسة (لجنة يتم تشكيلها في كل مدرسة) تطبيقه خلال الشهر الأخير من الفصل الدراسي الثاني لتشكل العلامات عليه 50% من علامة الاختبار النهائي في المبحث، وهذا يزيد من اهتمام الطلبة وجديتهم، وسيتم تصحيح الاختبار إلكترونيًا أو باستخدام الماسح الضوئي، وإجراء معادلة متزامنة للدرجات في النماذج المختلفة باستخدام نظرية الاستجابة للفقرة من أجل معادلة الدرجات ووضعها على مقياس موحد (بوسط حسابي 150 وإنحراف معياري 10). ويتم إبلاغ كل مدرسة بنتائج طلبتها وتوزيعهم حسب مستويات الأداء لاعتمادها والاستفادة منها في تحديد مجالات التحسين.

وتجدر الإشارة إلى أنه يمكن تطبيق الاختبار ورقيًا أو إلكترونيًا حسب ما تسمح به الظروف وتوافر البنية التحتية المناسبة في المدارس. ويوكل للجنة التقييم في المدرسة جمع المعلومات السياقية عن الطلبة والمدرسة والمعلمين. وسيتم تحليل بيانات التقييم والكشف عن العوامل المرتبطة بها والتحقق من ثبات الاختبار وصدقه، وكتابة تقرير وافٍ عن النتائج على مستوى الطالب والمدرسة والمديرية والوزارة. وترسل التقارير إلكترونيًا للجهات المعنية مقرونة بإيضاحات حول استخدام تلك النتائج. وستقوم إدارة الامتحانات والاختبارات وإدارة التدريب والإشراف التربوي بعقد ورش للمشرفين لمناقشة محتوى تلك التقارير وآليات توظيف نتائجها في تحسين التعليم والأداء المدرسي.

غرض الاختبار واستخدام النتائج

غرض الاختبار هو التصديق على التحصيل الأكاديمي للطلبة في مرحلة التعليم الأساسية المبكرة، وتوفير بيانات تقييمية على مستوى كل من الطالب والمدرسة والمديرية والنظام. وتصف البيانات التقييمية على مستوى الطالب مستوى الأداء له في المجالات الرئيسة لمحتوى المبحثين، مما يساعد في تحديد مجالات قوته والمجالات التي يحتاج فيها للتحسين، ويسمح للأهل والمدرسة والمجتمعات المحلية التعاون في التخطيط لبرامج

صيفية قصيرة، وتنفيذها للتغلب على نواحي الضعف، كما تعمل هذه التقييمات كخط أساس لقياس تقدم التلميذ في التعلم اللاحق.

وتفيد البيانات التقييمية لكل مدرسة في تحديد مجالات التحسين التي ينبغي العمل عليها من أجل تحسين الأداء المستقبلي لطلبتها، وتزوَّد البيانات التقييمية على مستوى المديرية والنظام متخذي القرار بمعلومات عن جودة التعليم على مستوى النظام، مما يساعدهم على وضع السياسات التقييمية واتخاذ قرارات التطوير المناسبة اعتمادًا على الأدلة المتوافرة من النتائج وبما يعزز ثقافة رسم السياسات واتخاذ القرارات القائمة على الأدلة، بالإضافة إلى إبراز الجوانب أو التدخلات التي يجب التركيز عليها أثناء ممارسة عملية التعليم في كل مدرسة. وهذا يتطلب أن توفر التقارير الخاصة بالنتائج معلومات تفصيلية عن المجالات الرئيسة للمحتوى وما يعرفه الطالب ويستطيع عمله.

ثانيًا: التقييم الوطني للصف السادس الأساسي

اختبار في المباحث الأساسية الأربعة (اللغة العربية واللغة الإنجليزية والرياضيات والعلوم) للصف السادس الأساسي. يعقد امتحان اللغة العربية بصورة سنوية بينما الامتحانات الأخرى تعقد كل ثلاث سنوات. أي أن جميع طلبة الصف السادس في كل سنة يتقدمون لاختبارين: اختبار في اللغة العربية، واختبار آخر في أحد المباحث الثلاثة الأخرى (لغة إنجليزية أو الرياضيات أو العلوم)، وتكون مدة كل اختبار ساعتين، من ضمنها في امتحان اللغة العربية 15 دقيقة لتعبئة استبانة الطالب. ويشتمل كل اختبار على نموذجين متكافئين في كل منهما 50 فقرة (30 فقرة فريدة، 15 فقرة مشتركة، 5 فقرات للتجريب) من نوع الاختيار من متعدد. ويستهدف التقييم جميع طلبة الصف السادس الأساسي.

وسيُعد المركز الوطني لتطوير المناهج بالتعاون مع المختصين في إدارة الامتحانات والاختبارات في وزارة التربية والتعليم كل اختبار وفقًا لمعايير بناء الاختبارات والمقاييس النفسية. ويقيس كل اختبار المهارات الأساسية في المبحث بحيث يكون 80% من الفقرات لمستوى التطبيق ومهارات التفكير العليا. وسيتم إعلام الطلبة ومعلميهم والمهتمين كافة بمخطط كل اختبار، بما يتضمنه من المجالات الرئيسة للمحتوى الذي يقيسه، ومجالاته الفرعية ونتاجات التعلم، وسيتم انتقاء فقرات كل اختبار من بنك أسئلة خاص بذلك الاختبار، أعدت فقراته وتم مراجعتها وتدقيقها وتجريبها ومعايرتها باستخدام نظرية الاستجابة للفقرة، وفقًا لجدول مواصفات محدد، يبين عدد فقرات الاختبار وتوزيعها حسب العمق المعرفي للفقرة ومستوى صعوبتها. وستشرف إدارة الامتحانات والاختبارات على تطبيق التقييم وتصحيحه، وتتولى لجنة التقييم في المدرسة (لجنة يتم تشكيلها في كل مدرسة)

تطبيقه خلال الشهر الأخير من الفصل الدراسي الثاني لتشكل العلامات عليه 50% من علامة الاختبار النهائي في المبحث، وهذا يزيد من اهتمام الطلبة وجديتهم. وسيتم تصحيح الاختبار إلكترونيًا أو باستخدام الماسح الضوئي وإجراء معادلة متزامنة للدرجات في النماذج المختلفة باستخدام نظرية الاستجابة للفقرة من أجل معادلة الدرجات ووضعها على مقياس موحد (بوسط حسابي 150 وانحراف معياري 10)، ويتم إبلاغ كل مدرسة بنتائج طلبتها وتوزيعهم حسب مستويات الأداء لاعتمادها والاستفادة منها في تحديد مجالات التحسين.

وتجدر الإشارة إلى أنه يمكن تطبيق الاختبار ورقيًا أو إلكترونيًا حسب ما تسمح به الظروف وتوافر البنية التحتية المناسبة في المدارس. ويوكل للجنة التقييم في المدرسة جمع المعلومات السياقية عن الطلبة والمدرسة والمعلمين. وسيتم تحليل بيانات التقييم والكشف عن العوامل المرتبطة بها والتحقق من ثبات الاختبار وصدقه، وكتابة تقرير وافٍ عن النتائج على مستوى الطالب والمدرسة والمديرية والوزارة. وترسل التقارير إلكترونيًا للجهات المعنية مقرونة بإيضاحات حول استخدام تلك النتائج. وستقوم إدارة الامتحانات والاختبارات وإدارة التدريب والإشراف التربوي بعقد ورش للمشرفين لمناقشة محتوى تلك التقارير وآليات توظيف نتائجها في تحسين التعليم والأداء المدرسي.

غرض الاختبار واستخدام النتائج

غرض الاختبار هو التصديق على التحصيل الأكاديمي للطلبة في نهاية المرحلة الأساسية المتوسطة، وتوفير بيانات تقييمية على مستوى كل من الطالب والمدرسة والمديرية والنظام. تصف البيانات التقييمية على مستوى الطالب مستوى الأداء له في المجالات الرئيسة والفرعية لمحتوى المبحثين، مما يساعد في تحديد مجالات قوته والمجالات التي يحتاج فيها للتحسين، ويسمح للأهل والمدرسة والمجتمعات المحلية التعاون في التخطيط لبرامج صيفية قصيرة، وتنفيذها للتغلب على نواحي الضعف، كما يمكن الحكم على التقدم الذي أحرزه الطالب مقارنة بما كان عليه في الصف الثالث الأساسي. وتفيد البيانات التقييمية لكل مدرسة في تحديد مجالات التحسين التي ينبغي العمل عليها من أجل تحسين الأداء المستقبلي لطلبتها.

وتزود البيانات التقييمية على مستوى المديرية والنظام متخذي القرار بمعلومات عن جودة التعليم على مستوى النظام، مما يساعدهم على وضع السياسات التقييمية واتخاذ قرارات التطوير المناسبة اعتمادًا على الأدلة المتوافرة من النتائج، بالإضافة إلى إبراز الجوانب أو التدخلات التي يجب التركيز عليها أثناء ممارسة عملية التعليم في كل مدرسة. وهذا يتطلب أن توفر التقارير الخاصة بالنتائج معلومات تفصيلية عن المجالات الرئيسة للمحتوى وما يعرفه الطالب ويستطيع عمله.

ثالثًا: التقييم الوطني للصف التاسع الأساسي

اختبار سنوي في مباحث اللغة العربية واللغة الإنجليزية والرياضيات والعلوم. ويتكون كل اختبار من نموذجين متكافئين يتضمن كل منهما 50 فقرة (30 فقرة فريدة، 15 فقرة مشتركة، 5 فقرات للتجريب). ومدة كل اختبار ساعتين منها في اختبار اللغة العربية 15 دقيقة لتعبئة استبانة الطالب.

وسيتم إعلام الطلبة ومعلميهم وكافة المهتمين بمخطط كل اختبار، بما يتضمنه من المجالات الرئيسة للمحتوى الذي يقيسه، ومجالاته الفرعية، ونتاجات التعلم. وسيتم انتقاء فقرات كل اختبار من بنك أسئلة خاص بذلك الاختبار، أعدت فقراته وتم مراجعتها وتدقيقها وتجريبها ومعايرتها باستخدام نظرية الاستجابة للفقرة، وفقا لجدول مواصفات محدد، يبين عدد فقرات الاختبار وتوزيعها حسب العمق المعرفي للفقرة ومستوى صعوبتها. وستشرف إدارة الامتحانات والاختبارات على تطبيق التقييم وتصحيحه، وتتولى لجنة التقييم في المدرسة (لجنة يتم تشكيلها في كل مدرسة) تطبيقه خلال الشهر الأخير من الفصل الدراسي الثاني لتشكل العلامات عليه 50% من علامة الاختبار النهائي في المبحث، وهذا يزيد من اهتمام الطلبة وجديتهم. وسيتم تصحيح الاختبار إلكترونيًا أو باستخدام الماسح الضوئي، وإجراء معادلة متزامنة للدرجات في النماذج المختلفة باستخدام نظرية الاستجابة للفقرة من أجل معادلة الدرجات ووضعها على مقياس موحد (بوسط حسابي 150 وانحراف معياري 10). ويتم إبلاغ كل مدرسة بنتائج طلبتها وتوزيعهم حسب مستويات الأداء لاعتمادها والاستفادة منها في تحديد مجالات التحسين.

وتجدر الإشارة إلى أنه يمكن تطبيق الاختبار ورقيًا أو إلكترونيًا حسب ما تسمح به الظروف وتوافر البنية التحتية المناسبة في المدارس. ويوكل للجنة التقييم في المدرسة جمع المعلومات السياقية عن الطلبة والمدرسة والمعلمين. وسيتم تحليل بيانات التقييم والكشف عن العوامل المرتبطة بها والتحقق من ثبات الاختبار وصدقه، وكتابة تقرير وافٍ عن النتائج على مستوى الطالب والمدرسة والمديرية والوزارة. وترسل التقارير إلكترونيًا للجهات المعنية مقرونة بإيضاحات حول استخدام تلك النتائج. وستقوم إدارة الامتحانات والاختبارات وإدارة التدريب والإشراف التربوي بعقد ورش للمشرفين لمناقشة محتوى تلك التقارير وآليات توظيف نتائجها في تحسين التعليم والأداء المدرسي.

غرض الاختبار واستخدام النتائج

غرض الاختبار هو التصديق على التحصيل الأكاديمي للطلبة في الصف التاسع الأساسي، وتوفير بيانات تقييمية على مستوى كل من الطالب والمدرسة والمديرية والنظام. تصف البيانات التقييمية على مستوى

الطالب مستوى الأداء له في المجالات الرئيسة لمحتوى المباحث الأربعة، مما يساعد في تحديد مجالات قوته والمجالات التي يحتاج فيها للتحسين، ويسمح للأهل والمدرسة والمجتمعات المحلية التعاون في التخطيط لبرامج صيفية قصيرة وتنفيذها للتغلب على نواحي الضعف، وتفيد البيانات التقييمية لكل مدرسة في تحديد مجالات التحسين التي ينبغي العمل عليها من أجل تحسين الأداء المستقبلي لطلبتها.

وتزود البيانات التقييمية على مستوى المديرية والنظام متخذي القرار بمعلومات عن جودة التعليم على مستوى النظام، مما يساعدهم على وضع السياسات التقييمية واتخاذ قرارات التطوير المناسبة اعتمادًا على الأدلة المتوفرة من النتائج، بالإضافة إلى إبراز الجوانب أو التدخلات التي يجب التركيز عليها أثناء ممارسة عملية التعليم في كل مدرسة، وهذا يتطلب أن توفر التقارير الخاصة بالنتائج معلومات تفصيلية عن المجالات الرئيسة للمحتوى وما يعرفه الطالب ويستطيع عمله.

ونعتقد أن ميول الطالب ورغبته هما الأساس في اختياره للمسار التعليمي في المرحلة الثانوية (أكاديمي، مهني (BTEC)). وإذا كان عدد الراغبين بالمسار المهني أكثر من عدد الشواغر المتاحة، فإنه يمكن استخدام نتائج الاختبار لغرض الالتحاق بالمسار المهني بحيث يكون للاختبار وزن 20% في تصنيف الطلبة و80% لمعايير أخرى مثل رغبة الطالب وميوله_ يطبق مقياس للميول_ وتحصيله المدرسي في صفوف السابع والثامن والتاسع.

الخلاصة والتوصيات بخصوص التقييمات الوطنية

من أجل تحقيق التقييمات الوطنية لأغراضها المنشودة وعلى النحو المأمول، فإنه يجب تنفيذ التوصيات الآتية دون تردد:

1. استخدام نظرية استجابة الفقرة في بناء وتطوير التقييمات الوطنية، ومعايرة فقراتها وتقدير خصائصها، ومعادلة نماذجها، واستخدام العلامة القياسية للتعبير عن النتائج وفي كتابة التقارير، وأن تكون على مقياس معياري موحد بوسط حسابي 150 وانحراف معياري 10. ويتطلب ذلك على مستوى إدارة الامتحانات والاختبارات والمركز الوطني لتطوير المناهج، توفير تدريب عالي المستوى للفريق العامل في مجال الاختبارات على كيفية استخدام نظرية استجابة الفقرة (Item Response Theory (IRT)) في معايرة فقرات الاختبار، وتقدير قدرات الطلبة، ومعادلة النماذج المختلفة للاختبار.

2. التحقق من أن التقييمات الوطنية تتماشى مع محتوى المنهاج الدراسي ومع ما تم تدريسه بالفعل، وأن تقيس مهارات التفكير العليا، وأن تحاكي الاختبارات الدولية في التركيز على قياس القدرة على تطبيق المعرفة المتعلمة في سياقات حياتية واقعية، وهذا ضروري من أجل أن تكون التقييمات الوطنية عالية الجودة.
3. تحديد البيانات السياقية التي سيتم جمعها، ووصف إطار جمعها، وبناء أدواتها من استبانات وغيره، ومناقشة محتواها، ووصف أدواتها، وكيفية استخدامها في تفسير نتائج التقييمات الوطنية.
4. بناء بنوك الأسئلة الخاصة بالمباحث الأساسية للصفوف المختلفة في السنوات القادمة حيث الأولوية الأولى لامتحانات الثانوية العامة، والتحول التدريجي إلى الاختبارات الإلكترونية وحسب ما تسمح به البنية التحتية.
5. استخدام ثلاثة مستويات للأداء الناجح: المتقدم، والإتقان التام، والإتقان الجزئي، وتصنيف الطلبة الذين لا يحققون أيًّا منها في المستوى غير المرضي "المبتدئ". وتحديد علامات القطع باستخدام أسلوب علمي (مثل أسلوب أنجوف المعدل)، ووصف تفصيلي لكل مستوى. ويتطلب ذلك تدريب القائمين على التقييمات الوطنية على كيفية استخدام أساليب تحديد علامات القطع.
6. أن تضع كل مدرسة خطة الاستفادة من نتائج التقييمات الوطنية، وتتولى مهمة تنفيذها وبمتابعة من مشرف القياس والتقويم في المديرية، وتوفر الأدلة المناسبة على تحقيق الغرض منها. وإدارة البيانات يجب أن تكون موحدة وتشمل طريقة جمعها، وتنقيحها من خلال البحث عن التناقضات والأخطاء والقيم المتطرفة.
7. تشكيل لجنة تقييم في كل مدرسة تتكون من 3 أفراد على الأقل وبما يتناسب مع حجم المدرسة (مدير المدرسة أو من ينيبه ومعلمين[40] لديهم المعرفة والخبرة في الممارسات التقييمية) لمساعدة المعلمين في تقييماتهم الصفية وللإشراف الداخلي على تطبيق التقييمات الوطنية والدولية.
8. ألّا تبقى التقييمات على شكل تقارير أو سجلات او سطور في قواعد البيانات، بل يجب أن تسهم في تحقيق أهداف النظام وتوجيه سياسات تعليمية وتقييمية أكثر فعالية في تحسين التعلم وضمان جودة نظام التعليم وضمان الإنصاف فيه والمساءلة.
9. التقييمات المستحدثة مستقبلًا يجب أن تكون متسقة مع إطار التقييمات الوطنية، من خلال احتوائها في التقييمات القائمة أو إحلالها بدلًا منها، مع مراعاة أن تكون على مستوى المجتمع وليس العينة.

[40] مقابل تخفيض النصاب لهم

4. التقييمات الدولية (صفوف الرابع والثامن والعاشر)

International assessments

التقييمات الدولية لتحصيل الطلبة مماثلة في كثير من النواحي للتقييمات الوطنية، فهي مسح للمدارس والطلبة بهدف توفير الأدلة على مستوى النظام التعليمي، وحول إنجازات الطلبة في مرحلة معينة من التعليم في مجالات محددة من المناهج، وكل منها يسمح بتتبع التغير مع مرور الوقت، ويمكن للدول المشاركة في التقييمات الدولية استخدام النتائج لإجراء تحليلاتها الخاصة بها.

والفرق الأكثر وضوحًا بين التقييمات الوطنية والتقييمات الدولية هو أن التقييمات الدولية توفر معلومات حول نظام التعليم مقارنة بنظام تعليمي آخر أو مجموعة من الأنظمة التعليمية في المنطقة أو العالم. ومع ذلك، فإن المعلومات التي تقدمها التقييمات الدولية قد تكون محدودة القيمة بالنسبة لدولة معينة، ويحصل ذلك عندما يكون إطارها غير مناسب لتلك الدولة، ويجب النظر للتقييمات الدولية لتحصيل الطلبة باعتبارها مكملة للتقييمات الوطنية وليست بديلًا عنها.

ويتطلب قرار مشاركة الأردن في أي تقييم دولي أو الاستمرار بالمشاركة فيه، عمل دراسة دقيقة لمدى ملاءمته للأردن، إذ لا بد من تقييم إطار التقييم الدولي وتحديد مدى مطابقته للمناهج الأردنية أو الكفايات التي يستهدفها النظام التعليمي الأردني، والنظر بالتكلفة مقابل المنفعة، ومدى الاستفادة من نتائج التقييم في تعزيز النظام التعليمي وتحسين تعلم الطلبة، وقد تمت الاستفادة في الأردن من التقييمات الدولية للطلبة في تطوير المناهج والكتب المدرسية حيث تم تضمينها لمهام تقييمية مشابهة أو تحاكي المهام التقييمية في التقييمات الدولية. وينصح باستمرار مشاركة الأردن في الدراسات الدولية المبينة في الجدول 12 مع مراعاة أن تساهم الجهات والمؤسسات التعليمية الأردنية في تطوير هذه التقييمات حيثما أتيح لها ذلك.

الجدول 12: التقييمات الدولية التي يوصى باستمرار بالمشاركة فيها

الفئة المستهدفة	اسم التقييم	غرض التقييم	المباحث المستهدفة	دورة التقييم	أسلوب التطبيق	
					الشمول	صورة الاختبار
الصف الرابع	الدراسة الدولية للتقدم في القرائية (PIRLS)	تقييم الاستيعاب القرائي	اللغة العربية	5 سنوات	عينة	جماعي / الكتروني
الصف الثامن	دراسة التوجهات الدولية في الرياضيات والعلوم (TIMSS)	تقييم التحصيل	الرياضيات والعلوم	4 سنوات	عينة	جماعي/ الكتروني

الفئة المستهدفة	اسم التقييم	غرض التقييم	المباحث المستهدفة	دورة التقييم	أسلوب التطبيق	
					الشمول	صورة الاختبار
عمر 15	البرنامج الدولي لتقييم الطلبة (PISA)	قياس الكفايات	القرائية والعلوم والرياضيات	3 سنوات	عينة	جماعي/ الكتروني

أولًا: الدراسة الدولية للتقدم في القرائية (PIRLS)

(Progress in International Reading Literacy Study, PIRLS)

تطبق دراسة PIRLS كل خمس سنوات منذ عام 2001، ويُعترف بها كمقياس عالمي لتقييم التوجهات في التحصيل القرائي في الصف الرابع الأساسي حيث يحدث التحول من تعلم القراءة إلى القراءة من أجل التعلم. وشارك الأردن في هذه الدراسة لأول مرة عام 2021 باختبار الورقة والقلم، وسيشارك في دورة عام 2026 بصورتها الإلكترونية، ويوفر تقييم تحصيل القراءة للمعلمين وصانعي السياسات رؤى أساسية حول فعالية النظام التعليمي، ويساعد على تحديد مجالات التحسين. ويشكّل غرضا القراءة وعمليات الفهم الأربع المبينة في جدول 13 الأساس لمحاور التقييم في دراسة PIRLS.

الجدول 13: النسب المئوية لأغراض القراءة وعمليات الفهم في دراسة *PIRLS*

أغراض القراءة	
اكتساب الخبرة الأدبية	50%
الحصول على المعلومات واستخدامها	50%
عمليات الفهم	
التركيز على المعلومات الواردة في النص واسترجاعها	20%
بناء استدلالات مباشرة	30%
تفسير الأفكار والمعلومات ودمجها	30%
تقييم ونقد محتوى النص وعناصره	20%

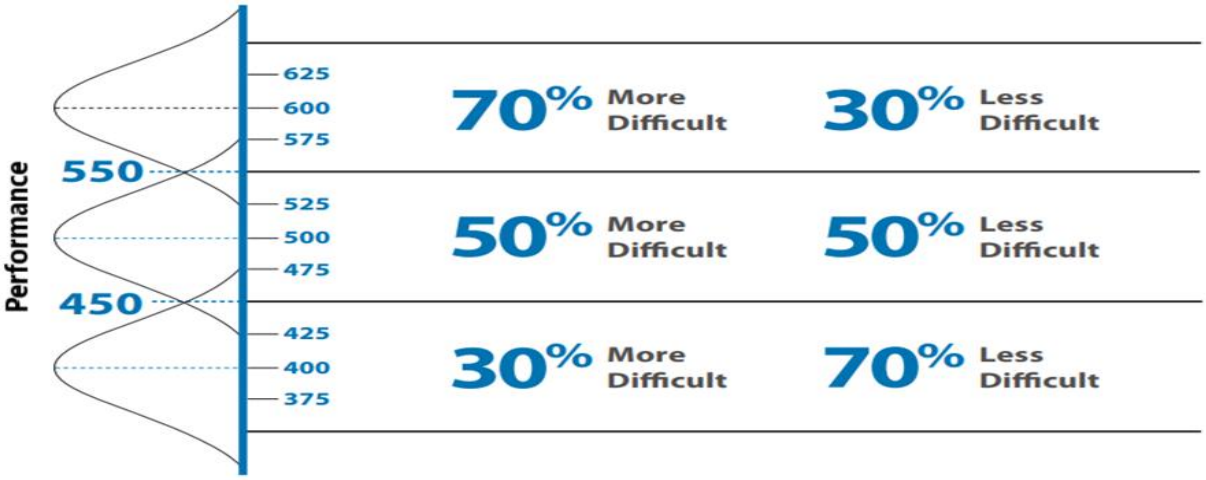
وتختلف النصوص والمهام التقييمية من حيث الطول، وتعقيدها النحوي، وتجريد أفكارها، وهيكلها التنظيمي، ومستوى الإدراك المعرفي. وتصنف النصوص في ثلاثة مستويات من الصعوبة: صعبة، ومتوسطة الصعوبة، وسهلة. وتخضع النصوص في دراسة PIRLS لمراجعة شاملة ودقيقة، لضمان ملاءمتها لمستوى

قراءة الطلبة، وأن تكون واضحة ومتماسكة، ومتحررة من تأثير الثقافات، وذات محتوى ممتع وجاذب لمجموعة واسعة من الطلبة، وقائمة على أساس كافٍ لتقييم جوانب عمليات الفهم الأربع.

وتُختار عينة الدراسة الأردنية وفقًا لدليل المعاينة الموحد في كل الدول بالاعتماد على قاعدة البيانات التربوية الأردنية، حيث تكون المدرسة هي وحدة المعاينة ومن كل مدرسة مختارة يتم اختيار شعبة واحدة عشوائيًا من شعب الصف الرابع. ويقوم NCHRD بالتشاور مع المعنيين في وزارة التربية والتعليم والسلطات التربوية بتحديد الطبقات الصريحة والضمنية لمجتمع الدراسة لضمان أن تكون العينة ممثلة لكامل المجتمع. ويتم جمع معلومات كاملة عن المدرسة ومنها الطبقة الصريحة للمدرسة (الثقافة العسكرية، سوريون، وزارة التربية والتعليم، وكالة الغوث الدولية، التعليم الخاص).

وتتكون أدوات الدراسة من كتيبات الأسئلة واستبانة الطالب واستبانة ولي الأمر واستبانة المعلم واستبانة مدير المدرسة. وتستخدم PIRLS التصميم التكيفي الجمعي للتقييم (Group Adaptive Assessment Design) حيث يوفر هذا التصميم قياسًا أفضل لتحصيل الطلبة من خلال المواءمة بين صعوبة التقييم ومستوى تحصيل الطلبة عبر الدول المشاركة. ويقوم هذا التصميم بتنظيم النصوص في مستويين من الكتيبات: الكتيبات الأكثر صعوبة وتضم نصوصًا صعبة وأخرى متوسطة الصعوبة، والكتيبات الأقل صعوبة وتضم نصوصًا سهلة وأخرى متوسطة الصعوبة. وتوزع الكتيبات على الدول وفق التصميم المبين في شكل 10.

شكل 10: تصميم توزيع الكتيبات وفق مستويات أداء الدول



مدة الاختبار

يتم تعيين كل طالب بشكل عشوائي إلى كتيب اختبار من بين كتيبات الاختبار. إذ يطلب من كل طالب الإجابة على كتيب اختبار الطالب المكون من جزأين أحدهما يحتوي نص أدبي والجزء الأخر نص

معلوماتي، يليه استبانة الطالب. ويتكون الاختبار من جلستين مدة كل منهما 40 دقيقة لكل جزء، مفصولة بفترة استراحة قصيرة، ويلي ذلك جلسة مدتها 30 دقيقة لاستبانة الطالب.

التحليل والمعايرة

يقع على عاتق كل دولة مشاركة في الدراسة عملية جمع البيانات من الميدان التربوي وادخالها في ذاكرة الحاسوب باستخدام برمجية خاصة تعدها الجهة الدولية لهذه الغاية. يقوم NCHRD بعملية تنظيف البيانات والتأكد من صحتها قبل إرسالها للجهة الدولية التي يقع على عاتقها معايرة وتحليل البيانات لجميع الدول المشاركة في الدراسة، إذ تستخدم الجهة الدولية في عمليات المعايرة والتحليل نظرية استجابة الفقرة. وتتم عملية المعايرة بشكل متزامن لبيانات أي دورة مع بيانات الدورة السابقة وذلك لقياس التوجهات من دورة إلى أخرى. وتستخدم دراسة PIRLS مقياس للقرائية بمتوسط حسابي 500 وانحراف معياري 100 لمتوسطات الأداء العام لطلبة الدول المشاركة في الدراسة. وتتراوح معظم متوسطات الدول من 300 إلى 700. وتستخدم دراسة PIRLS أربع نقاط مرجعية على طول مقياس التحصيل كمعايير دولية للتحصيل – معيار متقدم (625)، ومعيار مرتفع (550)، ومعيار متوسط (475)، ومعيار منخفض (400) مع وصف تفصيلي لكل منها.

ثانيًا: دراسة التوجهات الدولية في الرياضات والعلوم _ الصف الثامن

(Trends in international Mathematics and science study, TIMSS)

دراسة دولية رائدة تجرى كل 4 سنوات لتقييم تحصيل طلبة الصف الثامن في الرياضيات والعلوم. وتوفرTIMSS للدول المشاركة معلومات حاسمة لصنع سياسات تعليمية فعَّالة. ومن خلال قياس اتجاهات التحصيل منذ عام 1995، مكنت بيانات TIMSS الدول في جميع أنحاء العالم من اتخاذ قرارات قائمة على الأدلة لتحسين السياسات التعليمية المتعلقة بتدريس وتعلم الرياضيات والعلوم. وتقوم الدراسة بجمع معلومات مفصلة حول المناهج وتنفيذها، إلى جانب معلومات حول سياقات التعليم. وتستندTIMSS إلى أطار شامل يقوم على المناهج الوطنية في الدول المشاركة لوصف المعرفة والمهارات المتوقعة من الطلبة في الصف الثامن الأساسي.

ويلعب منسقو البحوث الوطنية دورًا مهمًا في المساعدة في تطوير أسئلة الاختبار والاستبانات حول العوامل السياقية في المدرسة والبيت المرتبطة بالتعلم وتحصيل الطلبة، وإدارة التقييم، والإبلاغ عن النتائج وتفسير النتائج ضمن سياقهم الوطني، وتوفر تقييماتTIMSS وصفًا موثوقًا لكيفية أداء طلبة الصف الثامن الأساسي في الرياضيات والعلوم، ويمكن للبلدان التي تشارك في دورات متعددة من هذه الدراسة أن تراقب

الاتجاهات في تحصيل الطلبة مع تقييم التغييرات التي حدثت في المناهج والجوانب الأخرى للتعليم التي تؤثر على التعلم.

المجتمع المستهدف وطريقة اختيار العينات _ الصف الثامن الأساسي

تستهدف دراسةTIMSS المدارس التي تتضمن طلبة في الصف الثامن في جميع الدول المشاركة في الدراسة. إذ يُعتبر الصف الثامن مرحلة انتقالية حيث يبدأ الطلبة في التعامل مع مواضيع أكثر تعقيدًا وتخصصًا مما يجعله نقطة حرجة لتقييم جاهزيتهم للتعليم الثانوي والعالي. وتختار عينة الدراسة بطريقة مماثلة تمامًا لما هو معمول به في دراسة PIRLS.

تصميم الدراسة

يتم إعداد إطار تقييمTIMSS في الرياضيات والعلوم للصف الثامن في كل دورة من دورات انعقادها بطريقة مشابهة لما كان عليه في الدورة السابقة، وذلك لضمان دراسة الاتجاه عبر دورات الدراسة المختلفة. وتركز دراسةTIMSS على تقيم مواقف حياتية من خلال الرياضيات والعلوم، حيث إن أكثر من نصف الفقرات (60%– 65%) تتطلب من الطالب استخدام مهارات التطبيق ومهارات التفكير وحل المشكلات. ويُنظم إطار تقييم كل من الرياضيات والعلوم بدلالة المحتوى والمهارات العقلية المستهدفة. ويبين الجدول 14 مجالات المحتوى في الرياضيات والعلوم ونسبها.

الجدول 14: مجالات المحتوى في الرياضيات والعلوم ونسبها.

مجالات الرياضيات	النسبة المئوية	مجالات العلوم	النسبة المئوية
الأعداد	30%	الأحياء	35%
الجبر	30%	الكيمياء	20%
الهندسة والقياس	20%	الفيزياء	25%
البيانات والاحتمالات	20%	علوم الأرض	20%

وفيما يتعلق بالمهارات العقلية، فقد تم تخصيص 35% للمعرفة والفهم، و35% للتطبيق، و30% لمهارات التفكير العليا.

أدوات جمع البيانات

تستخدم دراسةTIMSS أربع استبانات لتوفير بيانات سياقية ترتبط بتحصيل طلبة الصف الثامن في العلوم والرياضيات:

- استبانة معلم العلوم، لجمع بيانات حول خصائص الصف، والوقت المخصص لتدريس العلوم، والأساليب التعليمية، ويتم من خلال هذه الاستبانة جمع بيانات عن خصائص المعلم أيضًا، مثل الرضا الوظيفي، والمستوى التعليمي للمعلم، وأنشطة التنمية المهنية الحديثة وتحتاج هذه الاستبانة من 35 إلى 40 دقيقة لإكمالها.
- استبانة معلم الرياضيات، لجمع بيانات حول خصائص الصف، والوقت المخصص لتدريس الرياضيات، والأساليب التعليمية، ويتم من خلال هذه الاستبانة جمع بيانات عن خصائص المعلم أيضًا، مثل الرضا الوظيفي، والمستوى التعليمي للمعلم، وأنشطة التنمية المهنية الحديثة وتحتاج هذه الاستبانة من 35 إلى 40 دقيقة لإكمالها.
- استبانة مدير المدرسة، لجمع بيانات عن خصائص المدرسة، مثل بيئة المدرسة، وتوافر الموارد ومصادر التكنولوجيا، وتحتاج هذه الاستبانة من 30 إلى 35 دقيقة لإكمالها.
- استبانة الطالب، لجمع بيانات عن بيئة منزل الطالب، مثل المكتب في المنزل وموارد التعلم الأخرى الموجودة في المنزل، خبرات الطلبة في المدرسة، بما في ذلك الشعور بالانتماء للمدرسة وما إذا كانوا ضحايا للتنمر، تحتاج هذه الاستبانة من 30 إلى 35 دقيقة لإكمالها.

وتستخدِمTIMSS التصميم التكيفي الجمعي كما في دراسة PIRLS. وتوزع الكتيبات على الدول المشاركة حسب أدائها في الدورة السابقة وفق الشكل 10.

مدة الاختبار

يتم تعيين كل طالب بشكل عشوائي إلى كتيب اختبار من بين كتيبات الاختبار. إذ يطلب من كل طالب الإجابة على كتيب اختبار الطالب المكون من جزأين أحدهما رياضيات والأخر علوم، يليه استبانة الطالب. حيث يتكون الاختبار من جلستين مدة كل منهما 45 دقيقة لكل جزء، مفصولة بفترة استراحة قصيرة، ثم تليها جلسة مدتها 30 دقيقة لاستبانة الطالب.

التحليل والمعايرة

يقع على عاتق كل دولة مشاركة في الدراسة عملية جمع البيانات من الميدان التربوي وإدخالها في ذاكرة الحاسوب باستخدام برمجية خاصة تعدها الجهة الدولية لهذه الغاية. يقوم NCHRD بعملية تنظيف البيانات والتأكد من صحتها قبل إرسالها للجهة الدولية التي يقع على عاتقها معايرة وتحليل البيانات لجميع الدول المشاركة في الدراسة، إذ تستخدم الجهة الدولية في عمليات المعايرة والتحليل نظرية استجابة الفقرة.

يتولى مركز الدراسات الدولي لتيمس وPIRLS (ISC)، المسؤول عن تطوير وإدارة تقييماتTIMSS، عملية المعايرة للبيانات، إذ تتم عملية المعايرة بشكل متزامن لبيانات الدورة الحالية مع بيانات الدورة السابقة وذلك لقياس الاتجاهات من دورة إلى أخرى. توفر دراسةTIMSS مقياس رئيس بمتوسط حسابي 500 وانحراف معياري 100 لمتوسطات الأداء العام لطلبة الدول المشاركة في الدراسة على اختباري العلوم والرياضيات، بالإضافة إلى مقاييس فرعية حسب مجالات المحتوى لاختبار العلوم (الأحياء، الكيمياء، الفيزياء، علوم الأرض)، ومجالات الرياضيات (الأعداد، الهندسة، الجبر، البيانات والاحتمالات). ومقاييس فرعية لمتوسطات الأداء حسب المهارات العقلية (المعرفة، التطبيق، التفكير).

تُغطي تقييماتTIMSS مجموعة واسعة من المواضيع في الرياضيات والعلوم. ونظرًا لهذه التغطية الواسعة، تستخدمTIMSS تصميم (Matrix–sampling booklet) بحيث يتم إعطاء كل طالب مجموعة فرعية فقط من فقرات الرياضيات والعلوم. ومن أجل توفير تقديرات غير متحيزة لتحصيل الطلبة وعلاقته بالمتغيرات السياقية، تستخدم النموذج المجتمعي للانحدار الكامن (Latent Regression Population Model) للحصول على قيم معقولة (Plausible Values) تمثل الكفاءة في الرياضيات والعلوم لجميع الطلبة.

يتم الإبلاغ عن متوسط درجات تحصيل الدول المشاركة على مقاييس تحصيل الرياضيات والعلوم، حيث تتراوح معظم المتوسطات من 300 إلى 700. يوفر متوسط التحصيل لمستخدمي البيانات معلومات حول كيفية مقارنة التحصيل بين الدول المشاركة وما إذا كانت النتائج تتحسن أو تتراجع بمرور الوقت.

توفر معايير TIMSS الدولية معلومات حول ما يعرفه الطلبة وما يمكنهم فعله في نقاط مرجعية مختلفة على طول مقاييس التحصيل، وبشكل أكثر تحديدًا، حددتTIMSS أربع نقاط مرجعية على طول مقاييس

التحصيل لاستخدامها كمعايير دولية للتحصيل – معيار متقدم (625)، ومعيار مرتفع (550)، ومعيار متوسط (475)، ومعيار منخفض (400).

ثالثًا: البرنامج الدولي لتقييم الطلبة (PISA) في عمر 15 سنة.

(Program for International Student Assessment, PISA)

PISA، هو برنامج أنشأته منظمة التعاون الاقتصادي والتنمية (OECD) ويُنفَّذ كل ثلاث سنوات، واعتبارا من دورة العام 2029 ستنفذ الدورة كل أربع سنوات، لاختبار مهارات ومعرفة الطلبة البالغين من العمر 15 عامًا في (الرياضيات والقراءة والعلوم) ، فضلًا عن تقييم كفاءة الطلبة في المجال الابتكاري، الذي تمثَّل في التفكير الإبداعي عام 2022م. وتعد PISA بمثابة دراسة استقصائية للطلبة لتقييم مدى اكتسابهم للمعرفة والمهارات الأساسية الضرورية للمشاركة الفاعلة في الحياة الاجتماعية والاقتصادية. وتقدم PISA أدلة مقارنة لأداء الأنظمة التعليمية في أكثر من 80 دولة فيما يتعلق بجودة التعليم، والمساواة، ورفاهية الطلبة، وطموحاتهم الحياتية، والمناهج الدراسية، وبيئة التعلم والمواد المدرسية.

تُعرَّف PISA بأنَّها برنامج مستمر يُقدِّم رؤى لسياسات التعليم وممارساته، ويساعد على مراقبة الاتجاهات في اكتساب الطلبة ما يَلزم من مهارات ومعارف في الدول المشاركة. وبناءً على النتائج التي تخرج بها دراسة PISA، يُمكِن لواضعي السياسات مقارنة المهارات والمعارف التي يكتسبها الطلبة في بلدانهم، مع المهارات والمعارف التي اكتسبها الطلبة في البلدان الأُخرى، وتحديد السياسات والممارسات التي أفضت إلى تحسن النظم التعليمية في بعض الدول والإفادة منها.

المجتمع المستهدف وطريقة اختيار العينات

تُقيِّم دراسة PISA الطلبة الذين تتراوح أعمارهم بين سِنّ الخامسة عشرة وثلاثة أشهر وسِنّ السادسة عشرة وشهرين وقت التقييم، وغالبًا ما يكونون في الصفين التاسع والعاشر. وتتخذ خطوات وإجراءات مُحدَّدة لاختيار عيِّنات الطلبة وَفق الأساليب العلمية المُتَّبعة، فيتم اختيار المدارس المشاركة، ثمَّ تختار الطلبة منها؛ شرط بلوغهم سِنَّ الخامسة عشرة، والتحاقهم بالمدارس بصورة نظامية ورسمية.

يشرف الائتلاف الدولي المسؤول عن دراسة PISA على اختيار عيِّنات المدارس في الأردن، شأنه في ذلك شأن جميع الدول المشاركة. واشتملت هذه العيِّنات على قائمة كاملة تحوي جميع المدارس التي تضمُّ طلبة مُؤهَّلين في الأردن، بناءً على القوائم الكاملة التي تقدَّمها وزارة التربية والتعليم، وتشمل طلبة هذه المدارس ممَّنْ

هم في سِنِّ الخامسة عشرة، مُعتمِدةً في ذلك على نظام المعلومات التربوية في مركز الملكة رانيا العبدالله لتكنولوجيا التعليم والمعلومات. ويخضع اختيار العيّنات لإجراءات صارمة تَحْكُمها المعايير الفنية لدراسة PISA، وتهدف إلى التأكُّد أنَّ النتائج موثوقة وصحيحة وقابلة للمقارنة.

تصميم الدراسة

تُعنى كل دورة من دورات PISA بقياس كفاءة الطلبة وقدراتهم في الرياضيات والقرائية والعلوم، مع تركيز خاص على أحد هذه المجالات الثلاثة في كل دورة. وقد كانت الرياضيات محور الاهتمام والتركيز في دورة PISA 2022 بوصفها المجال الرئيس. وتؤكد الأُطر الخاصة بالمجالات الثلاثة على قدرة الطلبة على تطبيق المعرفة والمهارات في سياقات الحياة الواقعية: إذ يحتاج الطلبة إلى إظهار قدرتهم على التحليل والتفكير والتواصل الفاعل أثناء تحديدهم للمشكلات وتفسيرها وحَلِّها ضمن مجموعة مُتنوِّعة من المواقف. وتستخدم PISA تصميم الاختبارات التكيفية المتعددة الهجينة بهدف تحسين دقة القياس من خلال الموائمة بين صعوبة الفقرات وقدرة المفحوصين، إذ يطلب من كل طالب الإجابة على اختبار الطالب المكون من مجالين (رياضيات وعلوم، أو رياضيات وقرائية، أو رياضيات وتفكير إبداعي، أو علوم وقرائية، أو علوم وتفكير إبداعي، أو قرائية وتفكير إبداعي) يليه استبانة الطالب. حيث يتكون الاختبار من جلستين مدة كل منهما 60 دقيقة لكل جزء، مفصولة بفترة استراحة قصيرة، ثم تليها جلسة مدتها 35 دقيقة لاستبانة الطالب.

ويهدف اختبار العلوم إلى تقييم مدى قدرة الطلبة على تطبيق المعرفة العلمية في سياقات الحياة الواقعية. ويتضمن ذلك شرح الظواهر العلمية وتفسيرها، والاستقصاء العلمي وتفسير الأدلة العلمية ونقدها، واستخدام المعلومات العلمية في اتخاذ القرارات المتعلقة بقضايا عالمية أو محلية أو شخصية ذات صلة بالعلوم. وتتكون المعرفة العلمية من ثلاثة عناصر متمايزة، لكنها مترابطة: معرفة المحتوى (Content knowledge) بنسبة 43%، والمعرفة الإجرائية (Procedural knowledge) بنسبة 30%، والمعرفة المعرفية (Cognitive knowledge) بنسبة 27%.

ويهدف اختبار الرياضيات إلى تقييم قدرة الطلبة على تطبيق التفكير المنطقي الرياضي وتوظيف الرياضيات لحل المشكلات في مجموعة متنوعة من السياقات الواقعية. وتتكون المعرفة الرياضية من ثلاثة عناصر متمايزة، لكنها مترابطة: الاستدلال الرياضي (سواء الاستدلال الاستنتاجي أو الاستقرائي) وحل المشكلات، ومعرفة المحتوى الرياضي الذي يتم استهدافه في التقييم، والسياقات التي تتواجد فيها فقرات التقييم، مع مهارات مختارة من مهارات القرن الحادي والعشرين.

وتُعرف PISA الاستيعاب القرائي على أنه: قدرة الطلبة على فهم النصوص المكتوبة، واستخدامها، وتأملها وتقييمها والتعامل معها؛ تحقيقًا لأهدافهم، وتنميةً لمعارفهم وإمكاناتهم، وتحفيزًا لهم على المشاركة في خدمة المجتمع. ويقيس اختبار القراءة أربع عمليات أساسية: القراءة بطلاقة للكلمات والنصوص، وتحديد المعلومات، والفهم، والتقييم والتأمل.

التحليل والمعايرة

تُوفِّر دراسة PISA تقييمًا على مستوى النظام التربوي، يتيح المقارنة الدولية بين أنظمة التعليم في الدول المختلفة، وذلك بتخصيص فقرات مشتركة لجميع البلدان المشاركة، ويتم معايرة الفقرات بصورة متزامنة لجميع الدول باستخدام نظرية استجابة الفقرة لإنشاء مقياس مُوحَّد لكل مجال من مجالات الدراسة بوسط حسابي (500) نقطة، وانحراف معياري قَدْره (100). ويتم تصنيف أداء الأفراد في 6 مستويات للإتقان، وفي كل مستوى من مستويات الأداء، تُوضِّح الأوصاف ما يستطيع الطالب عمله بنجاح.

الغرض من المشاركة في التقييمات الدولية

تفيد المشاركة في التقييمات الدولية في تحقيق الأغراض الآتية:

- تقييم تحصيل الطلبة الأردنيين في المباحث والصفوف المستهدفة في تلك التقييمات، وتتبع التغير في تحصيلهم مع الوقت، وفقا للمعايير الدولية.
- تنمية القدرات المحلية لتلبية المعايير الدولية في الممارسات التقييمية.
- تحسين أداء النظام التعليمي الأردني من خلال تطبيق البيانات التشخيصية التي توفرها التقييمات الدولية وتعريضه للتدقيق الخارجي.
- فرصة الاستفادة منها وتوظيفها في سياق التقييمات الوطنية من حيث التصميم والتحليل وتحديد المستويات وتفسير النتائج في ضوء المعلومات السياقية حول المدارس والمعلمين والطلبة، بما في ذلك الموارد، ومواقف المعلمين والطلبة، من خلال استبانات موجهة للمدرسة والمعلمين والطلبة وأولياء أمورهم.
- تحديد احتياجات التدخل على مستوى السياسات، وتقييم آثار التدخل مع مرور الوقت.
- التعرف على الإمكانات البحثية في مجال التحصيل المقارن للموضوع.

• الالتزام بالمعايير الفنية العالية في تصميم التقييم، والأدوات، وأخذ العينات، والتحليل، وإعداد التقارير.

الأدوار والمسؤوليات في التقييمات الدولية

تتشعب وتتوزع الأدوار والمسؤوليات في التقييمات الدولية بين الجهة الدولية المسؤولة عن إعداد التقييم وتطويره وبين الدول المشاركة في الدراسة ويمكن تلخيصها بالآتي:

أولًا: المستوى الدولي

يقع على عاتق الجهة الدولية المسؤولة عن الدراسة الأدوار والمسؤوليات الآتية:

1- تطوير إطار التقييم لكل دورة من دورات الدراسة لمواكبة التطور والتغير المستمر في العملية التعليمية وجعله أكثر جاذبية للطلبة المشاركين في الدراسة، وتقوم ببناء فقرات الاختبار، والاستبيانات السياقية من خلال التشاور مع خبراء المناهج والتعليم والتقييم والمنسقين الوطنيين في الدول المشاركة في الدراسة.

2- إعداد دليل المعاينة لتوحيد إجراءات اختيار العينات للدول المشاركة في الدراسة، إضافة لاختيار العينات بالتشاور مع المنسقين الوطنيين للدول المشاركة في الدراسة.

3- تنسيق تنفيذ الدراسة بشكل عام عبر الدول المشاركة جميعها، لضمان تنفيذ الدراسة وفقًا لإجراءات محددة تضمن بيانات موثوقة وقابلة للمقارنة، إذ تقوم الجهة الدولية بتزويد الدول المشاركة بأدلة إرشادية لضمان إجراءات موحدة لاختيار العينات وعمليات الترجمة والتطبيق والتصحيح، إضافة إلى اجتماعات دورية تعقدها للمنسقين الوطنيين للدول المشاركة في الدراسة.

4- تعيين ضباط جودة دوليين للتأكد من أن إجراءات التطبيق تتم بطريقة موحدة عبر الدول المشاركة في الدراسة جميعها.

5- التصحيح الالي لفقرات الاختبار الموضوعية.

6- تحليل البيانات، والمعايرة، وإعداد التقارير، وإعلان النتائج. ويتم استخدام برمجيات إحصائية متقدمة لمراعاة تصميم العينة وضمان قابلية المقارنة بين النتائج عبر الدول المشاركة.

ثانيًا: المستوى الوطني

يتولى المركز الوطني لتنمية الموارد البشرية (NCHRD) تاريخيًا إدارة التقييمات الدولية في الأردن وذلك منذ مشاركته الأولى عام 1991، وبالرغم من أن نظام المركز الوطني لتطوير المناهج يؤكد على أنه

الجهة المسؤولة عن إدارة تقييمات الطلبة في الأردن، فإنه يُوصى في هذه المرحلة بأن يستمر (NCHRD) بإدارة التقييمات الدولية، على أن يتم تعزيز التعاون والشراكة مع كل من المركز الوطني لتطوير المناهج وإدارة الامتحانات والاختبارات في وزارة التربية والتعليم، وينبغي أن يقوم (NCHRD) بالأدوار الآتية:

1- تعيين منسق وطني للدراسة مسؤولًا عن التنسيق بين الجهة الدولية والسلطات التربوية في الأردن لتنفيذ الدراسة، ويقع على عاتقه مجموعة من المسؤوليات ومنها: مراجعة إطار التقييم للدراسة، وإعداد إطار المعاينة، والمشاركة في كتابة فقرات الاختبار ومراجعتها ومراجعة استبانات السياق بالتعاون مع المختصين في وزارة التربية والتعليم والمركز الوطني لتطوير المناهج.

2- المشاركة في اللجنتين التوجيهية والفنية التي تشكل برئاسة عطوفة الأمين العام للشؤون التعليمية، ومديرية التربية والتعليم والثقافة العسكرية، ومديرية التربية والتعليم لوكالة الغوث الدولية، وإدارة التعليم الخاص، وإدارة الإشراف والتدريب التربوي، ومركز الملكة رانيا العبد الله لتكنولوجيا التعليم والمعلومات. والتي تعمل على تطوير خطط الاستعداد للتقييم. ويقع على عاتق إدارة الإشراف والتدريب التربوي في وزارة التربية والتعليم ومديريات التعليم في السلطات التربوية الأخرى مسؤولية مراقبة تنفيذها، وتتضمن خطة الاستعداد المجالات الرئيسة الآتية: تجهيز مدارس العينة والطلبة للاختبار، والجاهزية التكنولوجية لمدارس العينة، والدعم الفني من ضباط الارتباط في مركز الوزارة والميدان، وعملية جمع البيانات.

3- متابعة عمليات ترجمة الاختبارات والاستبانات والأدلة من اللغة الإنجليزية إلى اللغة العربية بحيث تراعي الترجمة البيئة الأردنية والمعايير التي تضعها الجهة الدولية، وترسل الترجمة الأردنية إلى الجهة الدولية، لمراجعتها وإعادتها إلى الأردن لإدخال التعديلات المطلوبة على الترجمة.

4- جمع البيانات من الميدان التربوي بالتعاون مع وزارة التربية والتعليم وبما يتوافق مع المعايير التي تحددها الجهة الدولية لهذه الغاية، وذلك لضمان توحيد إجراءات التطبيق في جميع الدول المشاركة.

5- تصحيح أسئلة المقال.

6- إدخال البيانات التي تجمع من الميدان في ذاكرة الحاسوب باستخدام برمجية خبير إدارة البيانات Data Management Expert, DME).

7- إعلان النتائج وإصدار التقارير، ودعم الدراسات التربوية المختلفة المستفيدة من نتائج هذه الدراسة، بهدف إعانة واضعي السياسة التربوية على تطوير النظام التربوي الأردني والارتقاء بجودته.

آليات كتابة التقارير والإبلاغ عن النتائج

التقارير الدولية

تقوم الجهة الدولية المشرفة على كل دراسة بإعداد وإصدار تقرير دولي مفصل يعرض نتائج جميع الدول المشاركة في الدراسة، ويتضمن التقرير متوسطات الدرجات لكل مجال من مجالات الدراسة ولكل دولة مشاركة وفق نظام مقارنة مرجعي بوسط حسابي يساوي (500) وانحراف معياري يساوي (100)، ويوفر التقرير تحليلات معمقة للتحصيل وعلاقته بالعديد من المتغيرات الديموغرافية والسياقية مثل، الجنس، الوضع الاجتماعي الاقتصادي، وغيرها، ويتضمن التقرير مستويات الأداء وأوصافها بدلالة محتوى الفقرات التي أجاب عنها طلبة كل مستوى.

التقارير الوطنية

توصى اللجنة الفنية بالاستمرار بإعداد التقارير الوطنية لكل دراسة، مما سيسمح لأصحاب المصلحة باستخدام معلومات المسح عند مراجعة المناهج الدراسية وبرامج تدريب المعلمين وغيرها. ونوصي بإعداد تقارير مختصرة للمعلمين تتضمن ملخصًا للمعلومات ذات الصلة من نتائج المسوحات المختلفة، وينبغي أن تركز هذه التقارير على مجالات معينة من المعرفة والمهارات من المنهاج الدراسي، وتقديم أمثلة على فقرات الاختبار مع بيانات على مستوى الفقرة وتعليقات توضيحية، حيث ستساعد التقارير المختصرة في زيادة الوعي بالتقييمات الدولية في الأردن وتوفير مواد للتدريب في الصفوف الدراسية المختلفة.

وينبغي الاستفادة من التسهيلات التي تقدمها الجهات الدولية المسؤولة عن الدراسات للأردن في تعزيز قدرات العاملين في المركز الوطني لتطوير المناهج وفي إدارة الامتحانات والاختبارات على تحليل البيانات، وتفسير نتائج الدراسات، وكتابة التقارير، وكيفية الاستفادة منها.

ويجب أن تخاطب التقارير الوطنية الطلبة، وأولياء الأمور، والمُعلِّمين ومديري المدارس، والأوساط الأكاديمية، والمجتمع المدني، ووسائل الإعلام، والحكومة. وينبغي أن يوفر التقرير الوطني تقييمًا لمدى قدرة الطلبة على تطبيق ما تعلَّموه، ومقارنة منهجية بين الأردن والبلدان الأُخرى المشاركة في الدراسة من حيث النتائج التعليمية، والموارد المتوافرة، والفرص المتاحة، ومناقشة التحصيل المعرفي في ضوء المعلومات السياقية المختلفة.

كيفية استخدام النتائج

تستخدم التقييمات الوطنية والدولية في تقييم النظام التعليمي. فهي توفر أدلة وتقارير حول النتائج التي تعكس حالة وجودة التعليم في مستوياته المختلفة. وتشير ممارسات تقييم النظام الحالية في الأردن، رغم الخبرة المكتسبة على مر السنين، إلى أن تقييم النظام بشكل عام ليس مرتبطًا بشكل منهجي بتقييم الطلبة، ولا توجد

ممارسة موحدة لاستخدام النتائج في صنع السياسات التعليمية، ولم يتم استغلال الفرصة الجيدة المتاحة من المشاركة في التقييمات الدولية لتقييم نظام التعليم وإعداد التقارير واستخدام النتائج في صنع السياسات التعليمية.

يجب أن تدعم التقييمات الدولية تقييم النظام التعليمي وتوجيهه، ولا ينبغي النظر لنتائج هذه التقييمات باعتبارها ختامية، وينبغي استخدام نتائج التقييم بطريقة تكوينية، للمساعدة على تحسين النظام التعليمي. ويجب طرح السؤال: "لماذا نحصل على هذه النتائج؟" أي لا بد من تحديد العوامل التي تعيق تقدم تحصيل الطلبة، إضافة إلى مراجعة وتحديث المناهج وأساليب التدريس والتقييم. ومن الضروري أن يكون لدى الأردن نظام متكامل لإدارة البيانات، يضمن أن تكون البيانات من مختلف التقييمات الدولية والوطنية متوافرة ومخزنة بشكل مناسب، مما يتيح لأصحاب المصلحة والباحثين وطلبة الدراسات العليا في الجامعات الأردنية الوصول إلى البيانات لاستخدامها في تطوير النظام التعليمي ولأغراض بحثية.

ولتحسين أداء الأردن في التقييمات الدولية، ينبغي تطوير التقييمات الوطنية لتحاكي وتتوافق مع أطر التقييم للتقييمات الدولية، مما يعزز الاتساق بين التقييمات الوطنية والدولية، إذ ينبغي أن تصمم التقييمات الوطنية لتعكس المهارات العملية والحياتية التي تركز عليها التقييمات الدولية، وصياغة الفقرات فيها على نحو مماثل لما هو في التقييمات الدولية، وينبغي استخدام نتائج التقييمات الدولية لمراقبة الأداء على مستوى الأقاليم والسلطات التعليمية المختلفة.

التوصيات بخصوص التقييمات الدولية

أولًا: ينبغي أن تزود المؤسسات التعليمية الأردنية (NCHRD، وزارة التربية والتعليم، وNCCD) الجهات الدولية المشرفة على التقييمات الدولية بفقرات اختبارية تستند إلى أطر التقييم ومراعية للمناهج الأردنية.

ثانيًا: ينبغي أن تشارك المؤسسات التعليمية الأردنية في تحكيم الفقرات التي تعدها الجهات الدولية من حيث ملاءمتها للبيئة الأردنية واقتراح التعديلات المناسبة عليها أو المطالبة بحذفها.

ثالثًا: تزويد المدارس بالبنية التحتية التكنولوجية الحديثة وتهيئة الظروف لاستخدامها وتمكين الطلبة من المهارات الرقمية والحاسوبية والتي أصبحت عنصرًا أساسيًا ومهمًا في التقييمات الدولية وفي مختلف مناحي الحياة.

رابعًا: إعادة النظر في طبقات المجتمع الطلابي المستهدف التي على أساسها يتم اختيار العينات، فمثلًا يمكن اعتبار مدارس الملك عبد الله الثاني للتميز المنتشرة في الأردن طبقة قائمة بذاتها لضمان أن تكون ممثلة في عينة الدراسة.

خامسًا: إعادة النظر بالاستعدادات للتقييمات الدولية، مع التركيز أكثر على زيادة جدية الطلبة ودافعيتهم، من خلال التوعية بأهمية هذه التقييمات وتثقيف المعلمين وتدريبهم حول التقييمات الدولية لتمكين الطلبة من تقديم أفضل أداء ممكن على هذه التقييمات من خلال الآتي:

– تزويد المعلمين والطلبة بالفقرات الاختبارية من التقييمات الدولية التي تم السماح بالاطلاع عليها، مع فقرات أخرى يتم إعدادها لتحاكي تلك الفقرات.

– تدريب الطلبة على هذه الاختبارات قبل عامين من بدء المسح الرئيس لهذه التقييمات والتأكد من امتلاكهم للمهارات الحاسوبية والرقمية التي يحتاجونها في التقدم لهذه الاختبارات.

- العمل على نظام حوافز للطلبة المشاركين في الدراسات الدولية وخصوصًا لذوي الأداء المرتفع.

– مراجعة مستمرة للمعايير والمناهج وأساليب التقييم بالاستفادة مما تقدمه التقييمات الدولية.

– تعزيز التطوير المهني للمعلمين لتمكينهم من تبني أساليب تدريس حديثة تركز على التفاعل والتحليل. كما أن دمج التكنولوجيا في التعليم وتوفير بيئة تعليمية محفزة وتعاونية يمكن أن يسهم في تحسين نتائج الطلبة.

– جعل مهنة التعليم مهنة جاذبة من خلال نظام حوافز ومكافآت مرتبط بالأداء، وتعيين معلمين وقادة تربويين مؤهلين وتطويرهم بشكل مستمر.

– التوزيع الأمثل والعادل للموارد المالية والمادية والبشرية وتوفير بنية تحتية وموارد تعليمية وتكنولوجية للمدارس.

– الاستفادة من بيانات التقييمات الدولية من خلال إجراء تحليلات معمقة للبيانات، إذ لا تزال المراكز والجهات التي تهتم بالبحث التربوي في الأردن غير مهتمة بإنتاج بحوث تربوية موجهة نحو السياسات.

– إنشاء وحدة خاصة في إدارة الإشراف والتدريب التربوي لمتابعة خطط الاستعداد للمشاركة في التقييمات الدولية.

5. امتحان الثانوية العامة (التوجيهي في الصفين الحادي عشر والثاني عشر)

يعد امتحان الثانوية العامة (التوجيهي) مَعلمًا أساسيًا في النظام التعليمي الأردني، ويستهدف في المرحلة الحالية الطلبة بعد إكمالهم للمرحلة الثانوية في مسارهم التعليمي، أي بعد إنهاء الصف الثاني عشر، وتتولى

وزارة التربية والتعليم ممثلة بإدارة الامتحانات والاختبارات تصميم الامتحانات وإعدادها والإشراف على عمليات إدارتها وتصحيحها واستخراج نتائجها وتسليم كشوفات العلامات للمتقدمين لها، وتستمر الامتحانات في الدورة الواحدة لفترة طويلة تتعدى ثلاثة أسابيع، تبذل فيها أجهزة الدولة المختلفة مع وزارة التربية والتعليم بكل كوادرها من معلمين ومشرفين وغيرهم جهودًا كبيرة لإنجاح عملية تنفيذ الامتحانات، كما أن عملية تصميم الامتحانات وإعدادها وإدارتها تستغرق شهورًا عديدة وترتب على الوزارة تكاليف مالية عالية.

ولدى الناس توقعات عالية جدًا لامتحانات الثانوية العامة، فعلى أقل تقدير، يتوقعون أن تتوافق الاختبارات بشكل جيد مع ما تم تدريسه من حيث المضمون وطريقة طرح الأسئلة، وأن تكون الاختبارات عادلة، وتعطي فرصة لجميع الطلبة بما فيهم ذوي الاحتياجات الخاصة، ويتوقعون أن تكون العمليات الميدانية تنسيقًا وإدارّة موحدًة وعادلًة، وأن تكون عمليات تصحيحها موضوعية ولا تخضع للأحكام الشخصية، وأن تفي بالمعايير العلمية الصارمة القائمة على الصدق والثبات والعدالة والأمان العالي.

وقد دعت مجموعة من التقارير [41] إلى إصلاح نظام التوجيهي، وقد أبرزت الدعوات لإصلاح التوجيهي ما يأتي:

- يصنف الامتحان ضمن الامتحانات عالية المخاطر ويتطلب ذلك مستوى عاليًا جدا من الجودة في صدق النتائج وثباتها وعدالتها من جهة، ويفرض مستوى عاليًا جدًا من التوتر والقلق على المتعلمين والمعلمين والمدارس والأسر من جهة أخرى. وتتمثل سلبيات الامتحان في تثبيط عزيمة الطلبة وإجهادهم، وتضييق المنهاج الدراسي، وتعزيز التدريس للامتحان مما يقوض احترافية المعلم ويقيد عملية الإبداع في التدريس، ويرتب تكاليف باهظة على الأهل، ويعزز الفساد والغش.
- يركز الامتحان على حفظ الحقائق والمعلومات الأساسية ولا يقيس مجموعة كاملة من المهارات المطلوبة للتعليم العالي وعالم العمل.
- لا يتوافق مع الأطر العامة والخاصة للمناهج الجديدة، خاصة فيما يتعلق بتغطية مهارات القرن الحادي والعشرين مثل الاتصال وإدارة المعلومات وحل المشكلات والتفكير الناقد، فهو لا يعطيها الوزن الذي يتناسب مع أهميتها.

[41] Obediate, O. and Dawani, Z., 2014; Ababneh, I., Lebdih, K. and Tweissi, A, 2014; Venture, 2017; Najdawi, Z, et al, 2022; Kuttab, D, 2017

- يعمق عدم المساواة في مجتمع الطلبة المستهدفين من حيث توافر الفرص لفئات معينة منهم للمشاركة في التعليم والتدريب الخاص بالامتحان وعدم توافرها لفئات أخرى خاصة في المناطق الريفية والفقيرة، ويحد من إمكانية وصول المحرومين اقتصاديًا إلى التعليم العالي.
- يمزج بين غرضين فهو يستخدم لغرض التخرج من المرحلة الثانوية ومنح شهادة الثانوية العامة من جهة، ولغرض الالتحاق بالتعليم العالي من جهة أخرى، وهذا يقلل من فاعليته عنها فيما لو استخدم لغرض واضح منهما. ومن الناحية العملية فإنه اختبار قبول للتعليم العالي، ويجب أن يصمم على هذا الأساس.
- نسب الرسوب عالية في الامتحان، ويترتب على ذلك وجود عدد كبير من الشباب أمام مصير مجهول بلا فرص متاحة لهم، وفي ذلك هدر اقتصادي وتعليمي، علما بأن الاختبارات التي تستخدم لأغراض الالتحاق بالتعليم العالي لا تتطلب وجود معيار للنجاح.
- إجراءات الامتحان معقدة ومكلفة وتقليدية، ويستدعي ذلك تبسيطها وتحديثها وتقليل تكلفتها من خلال التدرج في تنفيذه إلكترونيًا وربطه بنتاجات التعلم والكفايات المستهدفة بالاعتماد على بنوك الأسئلة.
- يستخدم معيارًا وحيدًا للالتحاق بالتعليم العالي، ولا بد من مراجعة فلسفة اختيار الطلبة للالتحاق بالتعليم العالي بحيث تأخذ في الاعتبار نطاقًا أوسع من القدرات والمعايير وعدم اقتصارها على معدل العلامات في الثانوية العامة.

وعلى الرغم من أن العاملين في إدارة الامتحانات والاختبارات قد اكتسبوا خبرة عالية في عملية إدارة امتحانات التوجيهي، فإن هناك العديد من النقاط التي لا زالت بحاجة إلى المعالجة، واحدة من التحديات الرئيسة تتمثل في منع وتقليل الغش الذي يرافق الامتحانات وبصور مختلفة تهدد عدالتها، ومن التحديات الأخرى، كبر حجم الجهود التي تبذل من قبل قطاعات واسعة في الدولة لضبط عملية تطبيق الامتحانات واستخراج نتائجها، وعدم توافر مؤشرات دقيقة حول جودة الامتحانات لعدم وجود تحليلات معمقة لنتائجها من ناحية وعدم توفير البيانات للباحثين لتحليلها من ناحية أخرى، لذلك يجب التركيز بشكل أكبر على تحليلات نتائج الامتحانات، بطريقة تجعل من الممكن استخدامها في صنع السياسات التعليمية وتحسين جودة التعليم، وينبغي أن تكون النتائج قابلة للمقارنة العادلة عبر الدورات والسنوات المختلفة للامتحان في المبحث نفسه. ويمكن مقارنة نسب الطلبة في مستويات الأداء للمباحث المختلفة.

وتعتمد امتحانات الثانوية العامة وما تتضمنه من فقرات على الكتب المدرسية لا على نتاجات التعلم، فكثير من الأسئلة منقولة بالنص من الكتاب المدرسي، وهذا يحول دون كتابة الفقرات القائمة على الكفايات

وفق ما يرشد إليه الإطار العام للمناهج، لذلك يغلب على الفقرات قياس مستويات معرفية دنيا في مستويي التذكر والفهم، وخصائص الفقرات المضمنة للامتحانات غير معلومة مسبقًا لكونه لم يتم تجريبها وتقدير خصائصها، وكثيرًا ما نفاجأ بوجود أخطاء علمية أو مطبعية في بعض الفقرات، وتشكل هذه الأخطاء عامل تشويش وإرباك لسير جلسة الامتحان سواء تم تصويبها خلال جلسة الامتحان أو لم يتم تصويبها، وفي العادة، يتم تكليف عدد محدود جدا من المشرفين التربويين لإعداد الامتحان بكامل فقراته، ولا توجد أي ضمانات بالتزامهم بالمعايير المعتمدة عالميا في كتابة الفقرات، ولا زالت النظرية التقليدية في القياس هي المستخدمة حيث يكون الاهتمام بعلامة الطالب الكلية على الامتحان أي مجموع ما حصله من علامات على أسئلة الامتحان أو عدد الفقرات التي أجابها بصورة صحيحة.

وتستخدم العلامات الخام في التقارير وفي كشوفات العلامات وفي استخراج المجموع العام والمعدل المئوي للطالب في مباحث الامتحان جميعها، وتكون العلامات الخام مرتفعة في الاختبارات السهلة ومنخفضة في الاختبارات الصعبة، والعلامة الخام لا تقدم الصورة الحقيقية عن أداء المفحوص، لأنها بحد ذاتها ليس لها معنى، ويختلف المعنى لها باختلاف صعوبة وسهولة فقرات الامتحان، واختلاف صعوبة الامتحانات عبر النماذج المختلفة للامتحان وعبر الدورات والسنوات المختلفة قضية لا يمكن تجنبها مهما كانت الجهود المبذولة والحرص الذي توليه الوزارة في بناء الامتحانات وإدارتها، ويُخِلُّ ذلك في ضمان قيم العدالة التي تسعى الوزارة جاهدة إلى تحقيقها بين الطلبة، فالطالب الذي يتقدم لاختبار صعب تتأثر علامته بصورة غير عادلة بصعوبة ذلك الامتحان، والطالب الذي يتقدم لاختبار سهل يكتسب ميزة على الذين تقدموا للاختبار الصعب. والمقارنة المباشرة بين العلامات الخام تقود إلى إصدار أحكام خاطئة بسبب أن العلامات الخام غير قابلة للمقارنة، وبالتالي لا تخدم الغرض الرئيس للتقييم والمتمثل بترتيب الأفراد حسب امتلاكهم للقدرة أو السمة المقاسة.

وتجدر الإشارة إلى أن المؤسسات الدولية الكبرى المعنية بتطوير الامتحانات وإدارتها لا تستخدم العلامات الخام في التقارير الخاصة بالنتائج، لأن معناها يختلف عبر النماذج والأوقات المختلفة للامتحان، ولعدم قابليتها للمقارنة عبر الامتحانات المختلفة، ولمعالجة هذه السلبيات فإنه يجب استخدام نظرية الاستجابة للفقرة، وتهتم هذه النظرية بنمط استجابات الفرد على الفقرات وليس بعدد الفقرات التي أجابها بصورة صحيحة. كما تسمح بمعايرة الفقرات وتقدير خصائصها دون التأثر بخصائص الأفراد الذين أجابوا عنها، وتسمح بتقدير قدرات الأفراد دون التأثر بخصائص الفقرات التي أجابوها وبعددها.

باستخدام نظرية الاستجابة للفقرة في معايرة الفقرات، ومن خلال عمليتي التدريج والمعادلة Scaling and equating processes يمكن اعتماد تدريج موحد للعلامات في الامتحانات المختلفة، يسمح بإجراء

مقارنات عادلة وقياس التقدم لدى الفرد نفسه عبر الزمن، كما أن الاعتماد على بنوك الأسئلة ذات البنية الجيدة وتطبيق الامتحانات وإدارتها إلكترونيًا جنبًا إلى جنب مع استخدام نظرية الاستجابة للفقرة سيفيد حتمًا في ضمان صدقها وثباتها وعدالتها ورفع جودتها إلى المعايير المهنية الدولية.

وبناء على ذلك فإن التغييرات والتحسينات في امتحانات الثانوية العامة ينبغي أن تشمل بنية الاختبارات وتصميمها وتطويرها وإعداد نماذجها بالاعتماد على بنوك أسئلة رصينة بحيث يمكن معادلة النماذج المختلفة باستخدام منهجيةIRT، وكذلك التحول إلى نهج التقييمات القائمة على المعايير والكفايات، حيث تم التأكيد على ذلك في الإطار العام للمناهج الأردنية، ويسمح التقييم القائم على المعايير للمتعلمين والمتعلمات بالحكم على جودة عملهم بمقابل تلك المعايير، ويتم تعيين الدرجات للطلبة كمقياس لإنجازهم لمحتوى الموضوع ونتاجات التعلم فيه، كما ينبغي تحسين معايير الجودة للاختبارات واستخدام تصاميم معادلة أفضل وإجراءات أفضل في وضع المعايير التي تمكن من الوصول إلى قرارات صحيحة، وهذا يتفق مع اقتراحات سابقة[42] بإجراء الإصلاحات التالية على ترتيبات الامتحان الحالية:

- إعادة هيكلة جدول الامتحانات بحيث يتم إجراء الامتحانات على مدى عامين.
- إنشاء بنوك أسئلة عالية الجودة لإتاحة قدر أكبر من المرونة في تحديد مواعيد الامتحانات.
- نقل مسؤولية وضع الامتحانات وإدارتها إلى المركز الوطني لتطوير المناهج.
- تطبيق نظام قياس يضمن عدالة نتائج التقييم واتساقها.
- مراجعة سياسة القبول في الجامعات للحد من الاعتماد على نتائج "التوجيهي ودمج بيانات التقييم المدرسي وغيرها من المعايير في إجراءات القبول بالجامعة.

وقد تم تنفيذ بعض هذه الإصلاحات، ويجري حاليًا العمل على تنفيذ إصلاحات أخرى، فقد تم استحداث خطة دراسية جديدة للمرحلة الثانوية، حيث تم تشعيب الطلبة في نهاية الصف التاسع إلى مسارين، هما المسار الأكاديمي والمسار المهني BTEC، وفي المسار الأكاديمي تم إلغاء الفرع العلمي والفرع الأدبي، ويدرس جميعهم طلبة الصفين العاشر والحادي عشر المباحث الدراسية نفسها، وتتاح لطلبة الصف الثاني عشر فرصة اختيار المباحث بحسب حقل التخصص الذي يرغبون بدراسته في الجامعة. وهناك ستة حقول يمكن للطالب اختيار أي منها، وهي: الصحي، والهندسي، والعلوم والتكنولوجيا، واللغات والعلوم الاجتماعية، والقانون والعلوم

[42] لجنة تطوير امتحان شهادة الثانوية العامة، 2021

الشرعية، والأعمال، وفي المسار المهني يلتحق الطلبة في أحد تخصصات برنامج BTEC، حيث يتم تطبيق المستوى الثاني في الصف العاشر، والمستوى الثالث في الصفين الحادي عشر والثاني عشر.

وتم إعادة هيكلة امتحانات الثانوية العامة لتكون في عامين، أي يتم تنفيذها على مرحلتين: تبدأ المرحلة الأولى بعد إنهاء الصف الحادي عشر حيث يمكن للطالب التقدم للجزء الأول من الامتحان في أربعة مباحث إجبارية مشتركة مطلوبة من جميع الطلبة في المسار الأكاديمي، وهي: التربية الإسلامية للمسلم، واللغة العربية، واللغة الإنجليزية، وتاريخ الأردن، وتبدأ المرحلة الثانية بعد إنهاء الصف الثاني عشر حيث يستطيع الطالب التقدم للجزء الثاني من امتحان الثانوية العامة في أربعة مباحث يختارها حسب حقل التخصص الذي يرغب بدراسته في الجامعة على النحو الآتي:

1. الحقل الصحي: كيمياء وعلوم حياتية ولغة إنجليزية متقدم + مبحث من اختيار الطالب.
2. الحقل الهندسي: رياضيات وفيزياء+ مبحثان من اختيار الطالب على أن يكون أحدهما مبحث علمي.
3. حقل العلوم والتكنولوجيا: رياضيات+ 3 مباحث من اختيار الطالب منها مبحثان علميان.
4. حقل اللغات والعلوم الاجتماعية: عربي تخصص وإنجليزي متقدم + مبحثان من اختيار الطالب على أن يكون أحدهما في العلوم الإنسانية والاجتماعية.
5. حقل القانون والعلوم الشرعية: عربي تخصص وتربية إسلامية تخصص + مبحثان اختياريان على أن يكون أحدهما في العلوم الإنسانية والاجتماعية.
6. حقل الأعمال: رياضيات الأعمال وثقافة مالية وإنجليزي متقدم + مبحث من اختيار الطالب.

وبدأ العمل على إنشاء منظومة متكاملة لبنوك الأسئلة بالتعاون بين وزارة التربية والتعليم والمركز الوطني لتطوير المناهج؛ وهناك نقل تدريجي لمسؤولية وضع الامتحانات وإدارتها إلى المركز الوطني لتطوير المناهج، ليتم إعدادها وفقًا لأفضل المعايير العالمية باستخدام نظرية الاستجابة للفقرة IRT، وبالاعتماد على بنوك أسئلة متطورة تتيح تطبيق الامتحان في أكثر من دورة خلال السنة. ويبين الشكل 11 الإيجابيات المتوخاة من الخطة الجديدة والتوجه الجديد في امتحانات الثانوية العامة.

الشكل 11: الإيجابيات المتوخاة من التوجه الجديد في امتحان الثانوية العامة.

وبالرغم من هذه الإيجابيات فإن مجموع ما يتقدم له الطالب هو ثمانية امتحانات، وهو رقم لا زال مرتفعًا نسبيًا مقارنًة بكثير من دول العالم والإقليم، لذلك لا بد من إعادة النظر في الخطة الدراسية للمرحلة الثانوية، من خلال العمل على أن يكون عدد المباحث متساويًا في الصفين الحادي عشر والثاني عشر (13 مبحث لكل منهما)، ومراعاة التوازن بين المباحث الإنسانية والمباحث العلمية في كل منهما، وأن تكون الحصص الصفية مناسبة للمباحث المختلفة، ومعالجة الخلل في الأوزان المعطاة للمباحث، وإعطاء الرياضيات ما تستحقه من اهتمام، الأمر الذي يتطلب إعادة النظر في الخطة الدراسية وتطويرها بما يخدم النظام التعليمي ويأخذ حاجات الطلبة وقدراتهم بالاعتبار، ويمكن إجراء ذلك بسهولة دون إحداث تغيير جوهري في الخطة، فمثلًا من خلال نقل مبحثين من الصف الثاني عشر إلى الصف الحادي عشر تعالج قضيتي عدم التساوي في المباحث والتوازن بين المباحث العلمية والإنسانية، و يجب التفكير جديًا في المستقبل القريب بتقليل عدد المباحث التي يتقدم لها الطالب وتقليل عدد المباحث التي تحسب في المجموع العام أو المعدل للطالب.

غرض الامتحان

غرض الامتحان من الناحية القانونية هو منح شهادة الدراسة الثانوية للناجحين، أي أن الغرض الأساسي لامتحان الثانوية العامة هو تقييم حصيلة وخبرات التعلم في نهاية المرحلة الثانوية، ومنح من يحقق متطلبات ومعايير النجاح شهادة الدراسة الثانوية العامة، كما يستخدم الامتحان لأغراض القبول في الجامعات

الأردنية، فالهدف من الامتحان هو تقييم التحصيل الأكاديمي للطلبة وفقًا للخطة الدراسية الجديدة للمرحلة الثانوية وبالاعتماد على المناهج الجديدة التي تم تطويرها في المركز الوطني لتطوير المناهج، وعليه فإن الغرض الرئيس للامتحان هو التصديق على التحصيل الأكاديمي للطلبة في نهاية التعليم الثانوي، مما يمكنهم من الالتحاق بالتعليم العالي أو سوق العمل، ويرتب استخدام امتحانات التوجيهي كمعيارًا وحيدًا حاليًا لغرض القبول في مؤسسات التعليم العالي على الطلبة وأولياء أمورهم الكثير من الضغوط للحصول على درجات مرتفعة، وينفق الآباء المقتدرين للطلبة أموالًا طائلًة على التدريس الخاص لأبنائهم بأشكاله المختلفة، وفي هذا تعزيز لعدم تكافؤ الفرص بين الطلبة، ويتعزز عدم تكافؤ الفرص أكثر عندما يبحث آباء الطلبة الميسورين عن امتحانات وفرص بديلة أسهل لأبنائهم في دول أخرى.

وبناء على ذلك فإنه يجب تعديل التشريعات الخاصة بامتحانات الثانوية العامة وحل إشكالية الغرض المزدوج لها بشكل صريح، والتأكيد على استخدامها لغرض القبول والالتحاق في مؤسسات التعليم العالي، على ألا تكون المعيار الوحيد في ذلك. كما يجب في عمليات إعدادها وتطويرها واستخراج نتائجها مراعاة كونها معيارًا من مجموعة معايير للالتحاق بالتعليم العالي، وعليه فإن امتحانات الثانوية العامة تكون صادقة بقدر ما تتنبأ بالتحصيل الجامعي أو الأداء الوظيفي في سوق العمل.

الأدوار والمسؤوليات

تقع المسؤولية عن مواصفات وإنشاء وإدارة امتحانات الثانوية العامة حاليًا على وزارة التربية والتعليم ممثلة بإدارة الامتحانات والاختبارات، وتتعزز جودة الامتحانات مع زيادة الوضوح والاتساق والتماسك في نظام إعدادها وتطويرها وإدارتها وتقديمها بإنصاف وفعالية في بيئة اختبارية موحدة، ومراقبة الأداء عليها، والحصول على نتائج تتمتع بدرجة عالية من الصدق والثبات، وإعداد التقارير التفصيلية المتعمقة حولها، وفي ضوء التوجه الجديد لإصلاح النظام التعليمي وتقييمات الطلبة ومن ضمنها امتحان التوجيهي، فإنه يجب تحديد الأدوار والمسؤوليات للجهات الفاعلة في إعداد الامتحانات وإدارتها بوضوح تام، ويتطلب ذلك قدرًا كبيرًا من التعاون والاتصال المفتوح بين تلك الجهات للاتفاق على دور كل منها في تنفيذ الإجراءات المتصلة بالامتحانات، تخطيطًا وإعدادًا وإخراجًا وتطبيقًا وتصحيحًا وتحليلًا للنتائج، وينبغي تحديد دور كل منها في عملية إعداد الامتحانات وتطويرها، وتطبيقها في بيئة اختبارية موحدة وضمن ظروف تخضع لدرجة عالية من الرقابة والضبط. ويجب أن تتعاون الجهات المعنية في إطلاق حملة إعلامية عامة لبناء ثقة المجتمع على نطاق أوسع بالتحديثات التي تخص الإجراءات والتدابير الجديدة التي تخص امتحانات الثانوية العامة.

وفي هذا الإطار تم التوقيع على مذكرة تفاهم بين وزارة التربية والتعليم والمركز الوطني لتطوير المناهج، تؤكد على رغبة الفريقين في التعاون المُشترك لإنشاء "منظومة بنوك أسئلة" تدعم تحسين كفاءة وجودة جميع عمليات بناء الاختبارات المختلفة وتطويرها، وبما يخدم تحسين النظام التعليمي ومن ضمنه تقييم الطلبة. وبموجب المذكرة يتولى المركز الوطني لتطوير المناهج إنشاء بنوك الأسئلة وفق المعايير الدولية، بما فيها من عمليات تأليف الفقرات الاختبارية ومراجعتها وتدقيقها سيكومتريًا ولغويًا، ويتولى المركز كذلك، استخدام النظرية الحديثة في القياس في تحليل ومعايرة تلك الفقرات بعد التجريب الميداني لها وتقدير معلماتها، وتخزين المناسب منها في البنك، كما يتولى بالاعتماد على بنك الأسئلة، إعداد نماذج الامتحانات وتجهيزها للتطبيق الفعلي، وتتولى وزارة التربية كل ما يتعلق بجدولة الامتحانات وتسجيل الطلبة وتطبيق الامتحانات وتصحيحها واستخراج النتائج والإبلاغ عنها.

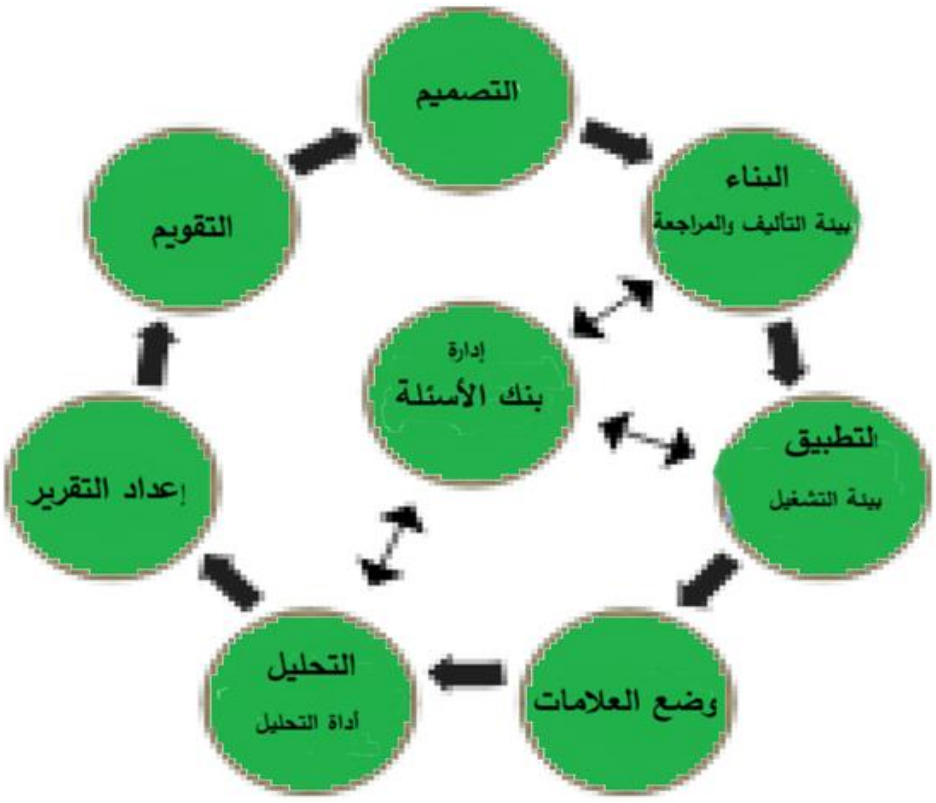
إنشاء بنوك الأسئلة

لا تشكل أي مجموعة من الأسئلة (الفقرات) بحد ذاتها بنكًا للفقرات، فبنك الفقرات عبارة عن مجموعة من الفقرات التي تم إعدادها بعناية وفق القواعد والإرشادات المتعارف عليها عالميًا، والتي تمت معايرتها بشكل مشترك، وتقيس موضوع مشترك، وبالتالي توفر معًا التعريف الإجرائي لذلك الموضوع، أي أن بنك الفقرات هو قاعدة بيانات مبرمجة تتضمن مجموعة كبيرة من الفقرات التي تم معايرتها والمصممة لقياس بنى محددة، ومنظمة في البرمجية وفق هيكلية تتيح عمل الآتي:

1. تطوير نماذج متكافئة من الاختبار تتمتع بالصدق والثبات.
2. ترتيب الفقرات في نموذج الاختبار بالصورة المرغوبة.
3. إتاحة فرصة تقديم الامتحانات إلكترونيًا في حال توافر البنية التحتية والتدريب المناسب للطلبة والمعلمين وفني المختبرات على استخدام التكنولوجيا في التعليم والتقييم.
4. ضمان عدم الإفراط في استخدام أي فقرة.
5. تتبع أداء الفقرات في مرات استخدامها المختلفة.
6. تصنيف الفقرات وتخزينها مع بياناتها الوصفية وفق هوية محددة تسهل عملية الوصول إليها.
7. تعقب المراجعات الدورية للفقرات.
8. فحص خصائص الاختبار قبل تطبيقه.
9. حذف أو إضافة فقرات للحصول على المواصفات وكمية المعلومات التي نريد.
10. إتاحة فرصة إجراء الاختبارات التكيفية مستقبلًا، أي تكييف الاختبارات لتناسب قدرة كل طالب.

وبات استخدام بنوك الأسئلة الإلكترونية سمة لأنظمة التقييم الوطنية في كثير من دول العالم، ويضمن استخدام نظرية الاستجابة للفقرة تطوير بنوك الأسئلة واستدامتها، وتحتل بنوك الأسئلة موقعًا مركزيًا في التقييم الرقمي، ولا بد من وجود برنامج (منصة) لتسهيل عملية التقييم الرقمي، ويمكن التمييز بين المكونات الأساسية في التقييمات الإلكترونية والتي تتمثل ببيئة تأليف الفقرات، وبنك الفقرات، وبيئة التطبيق، وأدوات التحليل، ويبين الشكل 12 مكونات نظام الاختبار الرقمي ضمن دورة التقييم.

الشكل 12: مكونات نظام الاختبار الرقمي ضمن دورة التقييم



يلاحظ من الشكل أن بنك الفقرات هو المكون المركزي الذي يتم فيه تخزين الفقرات، ويتم تحرير الفقرات وتطبيقها وتحليلها في أنظمة أخرى، وتقوم تلك الأنظمة بإرجاع الخصائص الإحصائية لها إلى بنك الفقرات، ويلعب بنك الفقرات دورًا مهما في المراحل الآتية لدورة التقييم:

- تصميم الاختبار (المعروف أيضًا بمواصفات الاختبار)، يتعلق بجوانب مثل الغرض من الاختبار، وما يقيسه، والاختيار بين التقييم التكويني أو الختامي، وطريقة التقييم، ويتم استخدام جدول المواصفات، أو مخطط الاختبار لكتابة الفقرات واختيار اللازم منها، ويحدد جدول المواصفات عدد الفقرات وأنواعها والموضوعات التي تغطيها والمستويات المعرفية لها.
- تطوير (بناء) الفقرات، يتم تنفيذه في بيئة التأليف الخاصة بنظام التقييم، وسيتم تضمين مفتاح الإجابة في حالة الأسئلة المغلقة والإجابات النموذجية في حالة الأسئلة المفتوحة.
- إدارة الاختبار، توفره برمجية التقييم في بيئة التشغيل، وتتحقق بيئة التشغيل من هوية الطالب عن طريق إجراء تسجيل الدخول، وتقدم للطالب فقرات الاختبار وتخزن إجاباته عليها.
- التصحيح ووضع العلامات، توفره برمجية التقييم للإجابات عن الأسئلة منتقاة الإجابة دون تدخل بشري بالاعتماد على مفتاح الإجابة الصحيحة الذي يتم تضمينه مسبقًا من قبل مطور الاختبار، ويتم تصحيح الإجابات عن الأسئلة المفتوحة بتدخل بشري وبالاعتماد على معايير خاصة للتقييم (Rubrics) أو بمساعدة برامج داعمة.
- التحليل للاختبار، تتضمن بعض البرامج أدوات خاصة لتحليل الاختبار، كما يمكن تنزيل ملف البيانات وإعداده في Excel أو SPSS، وتحليله خارجيًا، ويتم تحليل نتائج الاختبار من أجل تحديد خصائص الاختبار وتحديد الفقرات ذات الجودة المشكوك فيها، وهناك برمجيات خاصة (مثل Parscale أو IRTpro) باستخدام نظرية استجابة الفقرة في معايرة الفقرات ومعادلة الدرجات في النماذج المختلفة وتدريجها.

والبنوك الرقمية للفقرات ما هي إلا أدوات يمكن من خلالها زيادة جودة الاختبارات وخفض التكاليف وزيادة الأمان، وتسمح بمشاركة عدد كبير من الخبراء والمختصين في تطوير الفقرات، مما يوفر لكل فقرة خبرات أكثر ووقت أكثر لتطويرها ومراقبة جودتها، ويوفر مواءمة أفضل للأهداف والعمليات، وينصب التركيز في بنك الفقرات على تطوير الفقرات وإدارتها. ويتطلب إنشاء بنك الفقرات وجود منصة أو نظام خاص به، ونظام بنك الفقرات هو نظام رقمي يستخدم لتخزين وإدارة مجموعات الفقرات، وقد يحتوي على العديد من بنوك الفقرات، التي يكون لكل منها هيكلية وبنية خاصة تسهل عملية استخدامها، ويتم تنظيم الفقرات وما يتعلق بها من معلومات في بنك الفقرات على أساس المحتوى أو ناتج التعلم الذي تقيسه أو مستوى المهمة المتضمنة،

وتشتمل المعلومات المرتبطة بالفقرة على مُعرِف الفقرة ونصها ومحتواها ونوعها ومستواها وخصائصها الإحصائية بحسب نظرية القياس التقليدية ونظرية الاستجابة للفقرة.

وبدعم من مشروع الدعم الفني TAP الممول من USAID، فقد تم توفير منظومة Assessment Gourmet الإلكترونية الخاصة ببنوك الفقرات. والمنظومة من تطوير شركة المجموعة المتقدمة لتكنولوجيا المعلومات AGI، التي أدارت ونفذت مجموعة من الورش التدريبية على استخدامها للعديد من كوادر الوزارة والمركز الوطني لتطوير المناهج، ومحليًا فقد تم إخضاع البرمجية بجهود مقدرة من قبل وزارة الاقتصاد الرقمي والريادة والجمعية العلمية الملكية لفحوصات تتعلق بأدائها Performance test وقدرتها على التحمل Stress test، وأمانها Security test.

وتعرف هذه المنظومة من الآن فصاعدًا باسم المنظومة الأردنية لبنوك الأسئلة، وهي منظومة إلكترونية متكاملة لإدارة الاختبارات، ويمكن استخدامها على نطاق واسع لتغطي كافة أشكال التقييمات، وبشكل خاص امتحانات الثانوية العامة كأولوية أولى. وروعي في بنائها معايير بناء الاختبارات التربوية والنفسية كافة [43]، والمكونات الأساسية في نظم بنوك الأسئلة، ومراحل تطوير الاختبارات المقننة باستخدام النظرية الكلاسيكية للاختبارات ونظرية الاستجابة للفقرة، ويلخص الشكل 13 مراحل بناء الاختبارات التي تراعيها المنظومة.

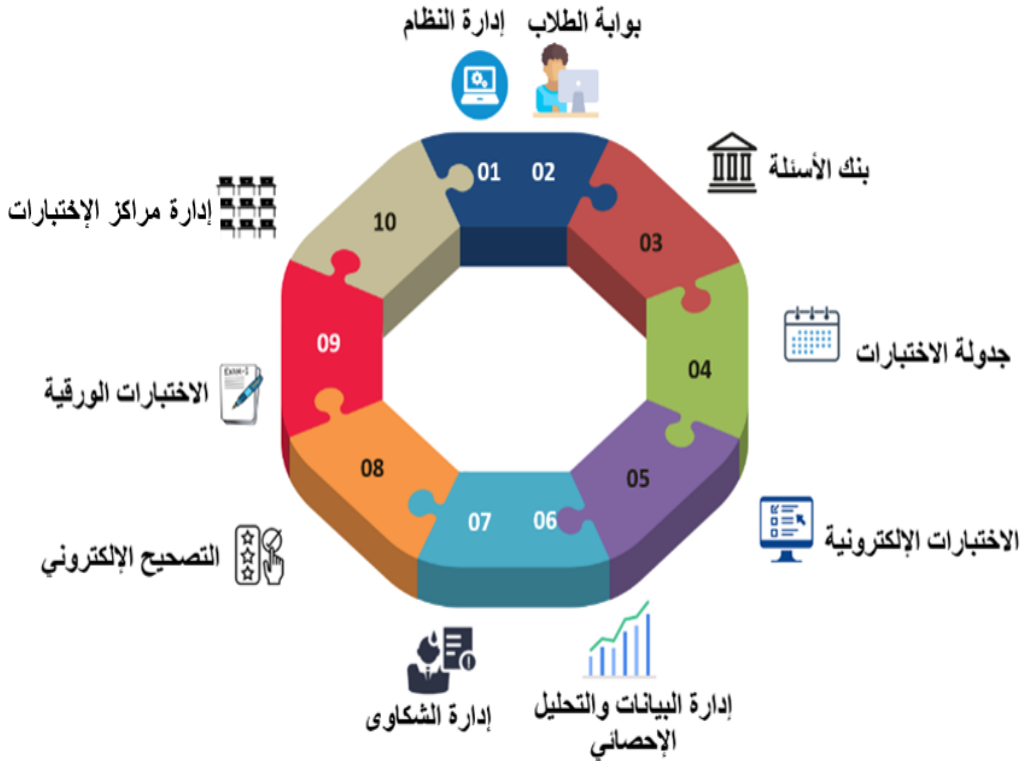
الشكل 13: مراحل بناء الاختبارات التي تراعيها المنظومة

[43] Standards for Educational and Psychological Testing, 2014.

وهذه المراحل متضمنة في معظم أنظمة التقييم العالمية، حيث يتم توفير المعلومات عن تصميم الاختبار ومخططه وجدول مواصفاته، ثم كتابة الأسئلة ومراجعتها وتدقيقها وتجريبها ومعرفة خصائصها، وإنشاء البنك ضمن هيكلية تسهل استخدام الفقرات في عمل النماذج الاختبارية، وتجهيزها للتطبيق الفعلي على الطلبة، ثم تصحيح استجابات الطلبة وإجراء التحليلات الإحصائية المناسبة لتعرف خصائص الاختبار وفقراته ومستوى أداء الطلبة.

وتتكون المنظومة من مكونات عدة تعمل بشكل متكامل ومتوافق، حيث إن المكونات جميعا قد تم بنائها باستخدام نظام الوحدات المنفصلة والتي يمكن فصلها واستخدامها بشكل منفصل في معظم الأحيان وفقًا للحاجة، ويوضح الشكل 14 الوحدات الرئيسة للمنظومة الأردنية لبنوك الأسئلة.

الشكل 14: الوحدات الرئيسية للمنظومة الأردنية لبنوك الأسئلة



وتتميز المنظومة باعتمادها على تكنولوجيا الويب والوحدات المنفصلة، والتي تمكن من إضافة مكونات أخرى مستقبلًا دون فقد أي بيانات، كما تعمل بواجهة ثنائية اللغة (إنجليزي – عربي) مع إمكانية إضافة لغات

أخرى مستقبلًا، كما تدعم متصفحات الويب الشائعة مثل **Chrome** و**Firefox** و**Safari** و **Microsoft EDGE**، والأجهزة اللوحية والجوالات الذكية، وتتكون المنظومة من أربعة مكونات رئيسة، تتمثل بإدارة المنظومة وبنوك الأسئلة، وإدارة الامتحانات وبوابة الطلبة، وتتولى وزارة التربية والتعليم القيام بالأدوار والمهام المرتبطة بجميع المكونات باستثناء مكون بنوك الأسئلة الذي يتولاه المركز الوطني لتطوير المناهج.

وتشكل إدارة المنظومة المكون الأول فيها، ويتم من خلالها تعريف إعدادات المستخدمين وصلاحياتهم، وتعريف البرامج الدراسية، والمباحث، والطلبة، وإعدادات المنظومة، كما تتضمن ربط قاعدة بيانات المنظومة بنظام إدارة معلومات الطلبة، ويمكن كذلك التعريف بإعدادات المؤسسة والتحكم بها، مثل بروفايل المؤسسة من بيانات أساسية وشعار وموقع المؤسسة، إلى جانب تحديد مدير البنك وإعدادات حقول الطلبة والتكامل مع الأنظمة الأخرى، ويهتم هذا المكون بتعريف مراكز وقاعات الامتحانات وتجهيزاتها وعدد المقاعد فيها، وغيرها من الإعدادات.

وتعد إعدادات الأمان من أهم هذه الإعدادات، نظرًا لأهمية الاختبارات من جهة، ولتطور وسائل الغش وطرق اختراق الأنظمة من جهة أخرى، لذلك توفر المنظومة بيئة آمنة لبناء الاختبارات وتنفيذها عبر وسائل عدة للحفاظ على سرية الاختبارات وأمانها، ومنها: التحكم في صلاحيات المستخدمين وأدوارهم، وتشفير كلمات السر والبيانات، واستخدام نظام اتصالات مؤمن ومشفر Secure Connection، واستخدام أساليب عدة للتأكد من هوية المستخدم، واستخدام مفتاح مشفر يمنع الاستخدام غير الرسمي للمنظومة. وسيتم وضع المنظومة على السحابة الخاصة بالحكومة الأردنية التي توفر كل الاحتياطات الضرورية للحفاظ على أمن فقرات البنك.

ويُعد بنك الفقرات المكون المركزي للمنظومة، وفيه يتم تخزين الفقرات. ويتولى المركز الوطني لتطوير المناهج إدارة هذا المكون. وللقيام بهذا الدور تم تشكيل اللجنة الاستشارية العليا لبنوك الأسئلة والامتحانات، ولكل بنك تم تشكيل لجنة الامتحان وتم تعيين مدير لكل بنك من حملة الدكتوراه في موضوع البنك أو في مناهج وتدريس الموضوع أو في مجال القياس والتقويم، وتعيين مطور للامتحان، كما تم تدريب مجموعات من المختصين في موضوع الاختبار المشهود لهم بالخبرة والكفاءة من المشرفين التربويين وأعضاء المناهج ومشرفي القياس والتقويم وغيرهم على كيفية تأليف الفقرات ومراجعتها وفق المعايير الدولية.

ولكل مبحث سيقوم مجموعة من المختصين بالمبحث ببناء مخطط الامتحان، وجدول مواصفاته. وسيتم تكليف مجموعتين من المشرفين: الأولى لتأليف الفقرات والثانية لمراجعتها والتحقق من صحتها العلمية ومن أن قواعد التصحيح ومفاتيح الإجابة لها صحيحة، وسيتم تكليف مجموعة من المتخصصين في القياس والتقويم لمراجعة الفقرات سيكومتريًّا للتأكد من الالتزام بقواعد وإرشادات كتابة الفقرات، وإزالة أي مؤشرات قد تجعلها

تتحيز لفئة معينة من الأفراد، والتأكد من أنها لا تتعارض مع عادات وتقاليد المجتمع. كما يتم تدقيق الفقرات لغويًا من قبل متخصص في لغة الامتحان، وكل هذه الإجراءات هي لضبط جودة الفقرات.

وتتمثل الخطوة الأولى في هذا العمل الوطني بتحليل محتوى الكتب المدرسية والإطار الخاص للمناهج من قبل مجموعة من المشرفين والمعلمين، من أجل تحديد معايير المحتوى ونتاجات التعلم والكفايات التي يجب قياسها، وتطوير إطار مرجعي للاختبار تقره لجنة الامتحان وتصادق عليه اللجنة الاستشارية العليا، والإطار المرجعي للاختبار بمثابة دليل للاختبار، ويتضمن خلفية عن الاختبار والغرض منه، ومجالات المحتوى، ونتاجات التعلم المستهدفة بالقياس، والتي تركز على المتعلم، وتكون محددة، وممكنة التحقيق، وقابلة للقياس. كما يتضمن المستويات المعرفية من معرفة وتطبيق ومهارات التفكير العليا وأوزانها، ونوع الفقرات، وجدول المواصفات، وعمليات تطوير الاختبار، وسيتم إعلان مخطط الاختبار للطلبة والمعلمين وكل المهتمين تحقيقًا للعدالة ومبدأ الشفافية.

ويتولى مدير البنك مهمة إنشاء البنك على المنظومة من خلال إدخال كود البنك واسمه، وعمل إعدادات له، وبناء هيكله، ويحدد مجموعة المستخدمين له وصلاحيات كل منهم، ويقوم بإنشاء جلسات خاصة على المنظومة لتأليف الفقرات وجلسات أخرى لتحكيمها ومراجعتها علميًا وسيكومتريًا وتدقيقها لغويًا.

ويقوم مؤلفو الفقرات بتأليف فقرات تقيس نتاجات التعلم التي تم تعيينها لهم، ويتم مراجعة الفقرات وفق إجراءات محددة تضمن المحافظة على سريتها وأمانها، وتكتب فقرات الاختيار من متعدد وتراجع، وفقًا للنماذج والشروط الآتية:

- يجب ألا تكون الفقرة منقولة حرفيًا من الكتاب المقرر أو شبكة الإنترنت أو أي مصدر آخر.
- أن ينسجم محتوى الفقرة مع محتوى الكتاب المدرسي المقرر.
- أن تقيس الفقرة بوضوح نتاج تعلمي محدد.
- أن تكتب الفقرة بلغة واضحة ومفردات لا تحتمل أكثر من معنى.
- أن يكون للفقرة بديل واحد صحيح.
- أن تُظهِر مقدمة (متن) الفقرة المطلوب بدقة قبل قراءة البدائل.
- أن تكون بدائل الفقرة جاذبة ومتجانسة في الطول والمحتوى.

- ألا تكون بدائل الفقرة متداخلة أو مترادفة.
- ألا تعتمد إجابة فقرة على إجابة فقرة أخرى.
- أن تكون الفقرات متنوعة في العمق المعرفي (معرفة وفهم، تطبيق، مهارات عليا).
- تجنب استخدام البدائل "جميع ما ذكر" أو "لا شيء مما ذكر" أو "a+b".

المركز الوطني لتطوير المناهج

نموذج كتابة فقرات الاختيار من متعدد

الصف	يكتب الصف
المبحث	يكتب اسم المبحث
المحور (الوحدة)	يكتب عنوان الوحدة
الدرس	يكتب عنوان الدرس
نتاج التعلم ورمزه	يكتب الرمز والنتاج
الرقم المؤقت للفقرة	يكتب رقم مؤقت للفقرة
نص الفقرة	يكتب نص الفقرة
بدائل الفقرة	**a b c d**
مفتاح الإجابة	يكتب رمز الإجابة الصحيحة
وصف أو عنوان الفقرة	تكتب جملة قصيرة تعبر عن الفقرة
كلمات مفتاحية	تكتب كلمات مفتاحية تساعد في البحث عن الفقرة
نسبة الطلبة الذين تتوقع أن يجيبوا عن الفقرة إجابة صحيحة	○ 80% فأكثر ○ 60% _79% ○ 40%–59% ○ أقل من 40%
المستوى المعرفي للفقرة	○ معرفة وفهم ○ تطبيق ○ مهارات وقدرات عقلية عليا
الوقت الذي تتوقع أن يحتاجه الطالب المتوسط للإجابة	○ أقل من دقيقة ○ بين دقيقة ودقيقتين ○ بين دقيقتين وثلاث دقائق ○أكثر من ثلاث دقائق
رمز مؤلف الفقرة	يكتب الرمز الذي يعطى للمؤلف من قبل المركز
تاريخ كتابة الفقرة	يكتب تاريخ كتابة الفقرة
تعليق كاتب الفقرة	تكتب الملاحظات التي يرغب كاتب الفقرة بعرضها مع الفقرة

نموذج مراجعة فقرات الاختيار من متعدد

الرقم المؤقت للفقرة	يكتب الرقم المؤقت للفقرة	
رمز المراجع	يكتب رمز المراجع المعطى من المركز	
تاريخ المراجعة	يكتب تاريخ المراجعة	
هل الفقرة تكرار طبق الأصل لفقرة سابقة قمت بمراجعتها؟	○ نعم ○ لا	إذا كانت الإجابة نعم فاكتب الرقم المؤقت للفقرة السابقة
هل يوجد أخطاء نحوية ولغوية ومطبعية في الفقرة؟	○ نعم ○ لا	إذا كانت الإجابة نعم فاكتب التصحيح للخطأ
هل يوجد أخطاء علمية في الفقرة؟	○ نعم ○ لا	إذا كانت الإجابة نعم فاكتب التصحيح للخطأ
هل مفتاح الإجابة صحيح؟	○ نعم ○ لا	إذا كانت الإجابة لا فاكتب الإجابة الصحيحة أو رمزها
هل تقيس الفقرة نتاج التعلم المحدد؟	○ نعم ○ لا	
هل المستوى المعرفي المحدد للفقرة صحيح؟	○ نعم ○ لا ما الصحيح؟ ○ معرفة وفهم ○ تطبيق ○ مهارات وقدرات عقلية عليا	
ما توصيتك بخصوص الفقرة؟	○ ترفض لعدم صلاحيتها ○ تقبل بعد التعديل عليها (يحدد التعديل المطلوب) ○ تقبل وترسل للمدقق السيكومتري	
تعليق المراجع	تكتب ملاحظة المراجع التي يرغب بإرسالها إلى كاتب الفقرة	

نموذج المراجعة السيكومترية للفقرات

الرقم المؤقت للفقرة	يكتب الرقم المؤقت للفقرة
رمز المراجع السيكومتري	يكتب الرمز للمراجع السيكومتري المعطى له من المركز
تاريخ المراجعة	يكتب تاريخ المراجعة
هل راعي مؤلف الفقرة القواعد الاسترشادية في كتابة الفقرة؟	○ نعم ○ لا
هل يوجد اتفاق بين المراجع الأول والثاني في مفتاح الإجابة؟	○ نعم ○ لا
هل يوجد اتفاق بين المراجع الأول والثاني في التوصية؟	○ نعم ○ لا
هل الفقرة خالية ظاهريا من أشكال التحيز المختلفة؟	○ نعم ○ لا
في حال تم إعادة الفقرة إلى كاتبها، هل قام بالتعديلات المطلوبة؟	○ نعم ○ لا
ما توصيتك بخصوص الفقرة؟	○ تقبل وتدرج في وضعية التجريب الميداني ○ حل الإشكالات إن وجدت بالتشاور مع المؤلف للفقرة و/أو المراجع ثم تقبل وتدرج في وضعية التجريب الميداني ○ تحذف ولا تدرج في وضعية التجريب الميداني
التعليق	يكتب تعليق المراجع السيكومتري الذي يرغب بإضافته لبيانات الفقرة

ويقوم مطور الاختبار بتكوين نماذج للاختبار ونشرها على المنظومة، وفق جدول المواصفات، بالاعتماد على الفقرات التي أجيزت في كل عمليات المراجعة والتدقيق، لغرض تجريب الفقرات جميعها، وسيقوم المركز الوطني لتطوير المناهج باختيار أحد تلك النماذج ليكون النموذج المرجعي ومعايرة فقراته من أجل إنشاء

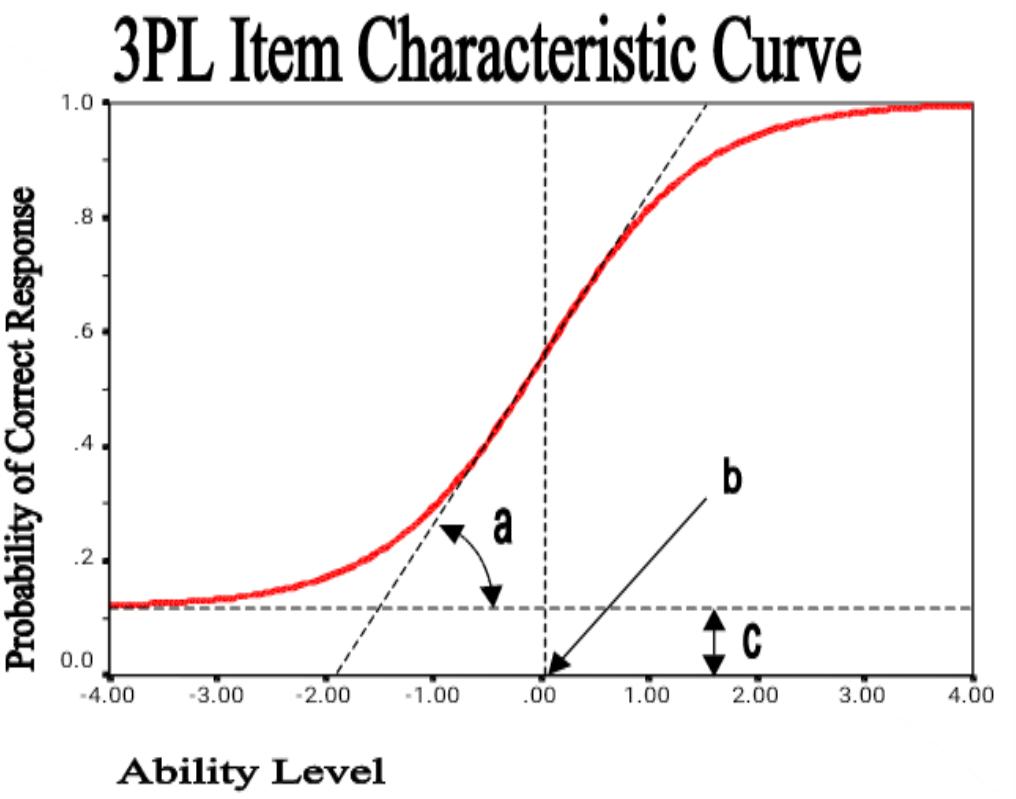
المقياس المرجعي، وسيتم تضمين النماذج التجريبية المتعددة فقرات أساسية مشتركة من النموذج المرجعي لربط النماذج بعضها ببعض.

وستقوم إدارة الامتحانات والاختبارات في وزارة التربية بتطبيق تلك النماذج على عينات من الطلبة – ليس من الضروري أن تكون متكافئة– وتصحيحها، وسيتم تحليل البيانات من النماذج المختلفة بحيث تتم معايرة الفقرات وفقًا للمقياس المرجعي لإنشاء بنك فقرات معايرة، من خلال انتقاء الفقرات التي تفي بالمعايير المطلوبة، وتخزينها مع خصائصها الإحصائية والوصفية في البنك، وسيتكون نموذج أي اختبار من 55 فقرة اختيار من متعدد لكل منها أربعة بدائل (30 فقرة منفردة، 20 فقرة مشتركة، 5 فقرات للتجريب) تختار من البنك حسب جدول المواصفات، ويستثنى من ذلك امتحانات اللغة العربية واللغة الإنجليزية حيث تتضمن الكتابة حول موضوع محدد.

معايرة فقرات الاختبار باستخدام نظرية استجابة الفقرة IRT

تعتمد نماذج القياس جميعها على مجموعة من الافتراضات الكامنة وراء استجابة الشخص للفقرة، ويصف منحنى خاصية الفقرة (Item characteristic curve ICC) العلاقة بين قدرة الفرد في السمة التي يتم قياسها واحتمال إعطاء استجابة صحيحة على الفقرة، وتقدم نماذج IRT افتراضات واضحة حول الشكل الرياضي لمنحنى خصائص الفقرة، ويوصف الشكل بأنه منحنى لوجستي، يحدد احتمال إعطاء إجابة صحيحة للفقرة بدلالة قدرة الفرد ومعلمات الفقرة، ويفترض النموذج ثلاثي المعلمات PL–3، أن احتمال إعطاء استجابة صحيحة للفقرة هو دالة لمستوى قدرة الشخص ومعلمات الفقرة الثلاثة: صعوبة الفقرة، وقدرتها التمييزية، واحتمال تخمينها، ويبين الشكل 15 منحنى خصائص الفقرة حسب النموذج ثلاثي المعلمات.

الشكل 15: منحنى خصائص الفقرة بحسب النموذج ثلاثي المعلمات



يمثل المحور الأفقي في الشكل قدرة المتقدم للاختبار، ويمثل المحور الرأسي احتمال تقديم إجابة صحيحة للفقرة. ويكون ارتفاع المنحنى أقل ما يمكن في الطرف الأيسر لدى المتقدمين للاختبار من ذوي القدرة المنخفضة. ومع زيادة قدرة المتقدم للاختبار يرتفع المنحنى، ويكون الارتفاع تدريجيًا في البداية، ثم يرتفع بشكل أسرع، ثم يستقر تدريجيًا مع اقتراب الاحتمال من 1.00 لدى المتقدمين للاختبار من ذوي القدرة العالية.

وتوضح معلمة الصعوبة b موقع النقطة على متصل السمة (القدرة) المقاسة التي تقابل نقطة انعطاف منحنى خاصية الفقرة ICC. وتشير قيم b الأكبر إلى الحاجة إلى قدرة أكبر لإجابة الفقرة بشكل صحيح، أي أن الفقرة أصعب، ومن الناحية العملية تتراوح قيم b بين القيمتين 3− و3+ ، وميزة هذا المؤشر للصعوبة هو أنه يتم التعبير عنه على مقياس قدرة الفرد نفسه، وسيتم الاحتفاظ بالفقرات التي تكون قيم b لها بين 2.5− و2.5+.

وتتناسب معلمة التمييز a مع ميل المنحنى ICC عند نقطة الانعطاف، وتفيد الفقرات التي منحنياتها أكثر انحدارًا في تصنيف الطلبة بشكل أكثر وضوحًا إلى مستويات قدرة مختلفة، ويشير ميل المنحنى ICC عند أي نقطة قدرة إلى القدرة التمييزية للفقرة عند تلك النقطة، ومن الناحية العملية تتراوح قيم a بين 0 و3.

وسيتم الاحتفاظ بالفقرات التي تزيد فيها قيمة a على 0.60 مع مراعاة ألّا يقل الوسط الحسابي لقيم التمييز عن 1. وتشير معلمة التخمين c إلى احتمال إجابة الفقرة بشكل صحيح من أفراد بمستويات قدرة منخفضة للغاية.

وتقدم نماذج نظرية **IRT** فوائد ومزايا عديدة في تصميم الاختبار، وبناء نماذج الاختبار، وتحليل النتائج ومعايرة الفقرات، والمعايرة في IRT، هي عملية تحديد منحنيات الاستجابة لمجموعة الفقرات التي تقيس نفس القدرة، وتقدير القيم الإحصائية لمعلماتها، ويعني ذلك تقدير صعوبتها في النموذج ذو المعلمة الواحدة، وتقدير صعوبتها وتمييزها في النموذج ذو المعلمتين، وتقدير صعوبتها، وتمييزها، واحتمال تخمين الإجابة الصحيحة لها في النموذج ثلاثي المعلمات، ولأغراض معايرة فقرات امتحانات الثانوية العامة سيتم استخدام نموذج نظرية استجابة الفقرة متدرج الاستجابة ثنائي المعلمة Two Parameter IRT Graded Response Model، لكونها تفيد في التعامل مع فقرات الاختيار من متعدد ومع الفقرات ذات الإجابة القصيرة.

ولتقدير هذه المعلمات نحتاج إلى استجابات المتقدمين للاختبار على الفقرات، ونحتاج إلى ربط جميع الفقرات المراد معايرتها معًا، ويتم الربط بين الفقرات بصورة مباشرة أو غير مباشرة، ويحدث الربط المباشر بين فقرتين إذا استجاب عليهما عدد كافٍ من الطلبة، ويحدث الربط غير المباشر بين مجموعتين من الفقرات التي لم يتم الاستجابة لها من قبل نفس الافراد، من خلال ربطها بفقرات مشتركة تم الاستجابة عليها من قبل الأفراد نفسهم.

ويبين الجدول 15 مثالًا على ربط الفقرات بصورة مباشرة أو غير مباشرة، في هذا المثال ترتبط الفقرات 1–20 بصورة مباشرة بالفقرات 21–40 لأن المتقدمين للاختبار 1–1000 استجابوا على هاتين المجموعتين من الفقرات، وبالفقرات 41–60 لأن المتقدمين للاختبار 1001–2000 استجابوا على هاتين المجموعتين من الفقرات، وبالفقرات من 61–80 لأن المتقدمين للاختبار 2001–3000 استجابوا على المجموعتين من الفقرات. ولا يوجد ارتباط مباشر بين مجموعة الفقرات 21–40 ومجموعة الفقرات 41–60، لكن بينهما ارتباط غير مباشر من خلال ارتباط كل منهما بمجموعة الفقرات 1–20. وسيتم البناء على هذا المثال في معايرة فقرات كل بنك من بنوك الأسئلة الخاصة بامتحانات الثانوية العامة.

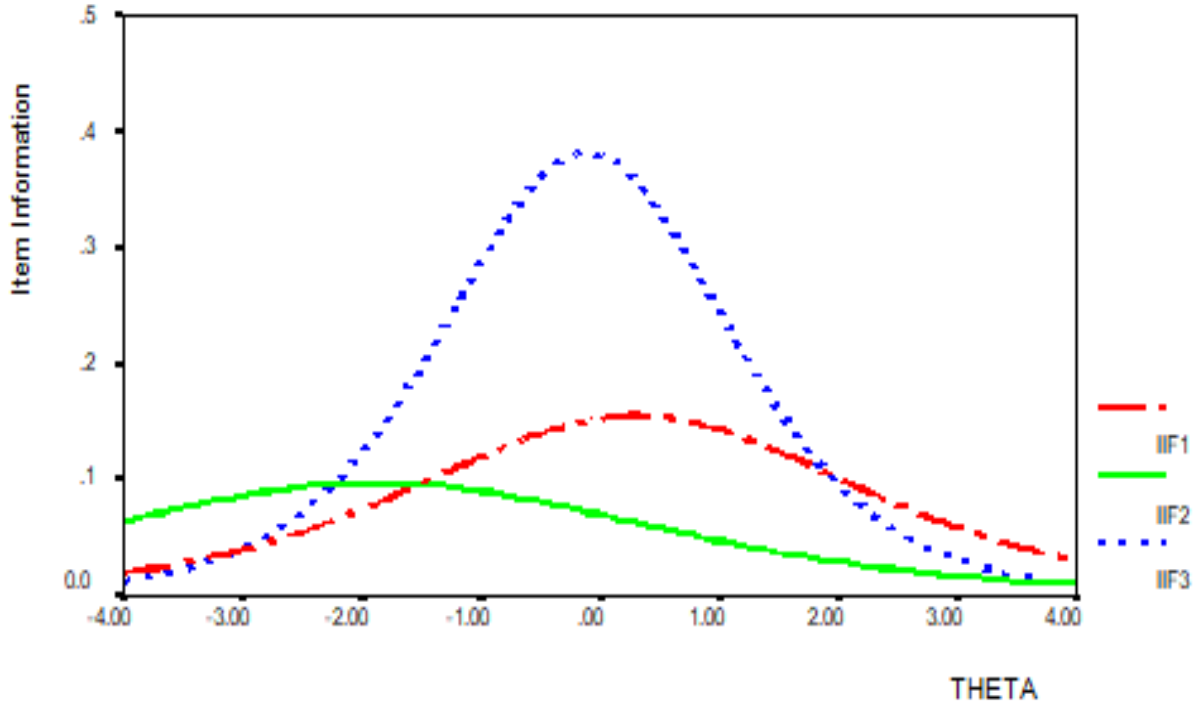
الجدول 15: ربط الفقرات بصورة مباشرة وغير مباشرة

المتقدمون للاختبار	الفقرات 1–20	الفقرات 21–40	الفقرات 41–60	الفقرات 61–80
1000- 1	توجد استجابات	توجد استجابات	–	–
2000 - 1001	توجد استجابات	–	توجد استجابات	–
3000 - 2001	توجد استجابات	–	–	توجد استجابات

وتمتاز نظرية الاستجابة للفقرة في أن تقدير قدرة الطالب لا يعتمد على مجموعة معينة من فقرات الاختبار، بمعنى أن قدرات الأفراد متحررة من خصائص الفقرات. وكذلك، فإن تقديرات قيم معلمات الفقرة لا تعتمد على الأشخاص الذين يستجيبون للفقرات، وبالتالي فإن معلمات الصعوبة والتمييز والتخمين لفقرة ما ستكون هي نفسها بغض النظر عن مجموعة الطلبة الذين يستجيبون لها، أي أن تقديرات معلمات الفقرات متحررة من خصائص الأشخاص الذين يجيبون عنها.

كما توفر نماذج IRT تقديرات للخطأ عند كل نقطة على متصل القدرة، وتتيح هذه الميزة لمطوري الاختبار بناء أدوات أكثر دقة وبالأخص عند درجات القطع حيث سيتم اتخاذ معظم القرارات، ويتم ذلك من خلال الاعتماد على ما يعرف بدالة معلومات الاختبار Test Information Function(TIF) الناتج عن تجميع دوالّ معلومات الفقرات Item Information Function(IIF) المتضمنة فيه. ويبين الشكل 16 دوالً معلومات لثلاث فقرات.

الشكل 16: دوالً معلومات لثلاث فقرات



وتوضح دالة معلومات الفقرة قيمة مساهمة الفقرة في تقييم القدرة عند مستويات معينة، ويلاحظ من الشكل أن الفقرة الثانية عند القدرة (3.0−) توفر كمية معلومات أعلى من الفقرتين الأولى والثالثة، وأن الفقرة الثالثة توفر معلومات أعلى من الفقرتين الأولى والثانية في مدى القدرة من 2.0− إلى 2.0+، وعليه فإن الفقرة الأولى هي الأكثر مساهمة في تقدير القدرات العالية (أكبر من 2)، وأن الفقرة الثانية هي الأكثر مناسبة في تقدير القدرات المنخفضة (أقل من 2−) وأن الفقرة الثالثة هي المساهم الأكبر في تقدير مدى واسع من القدرات (بين 2− و2+).

معادلة نماذج الاختبار

المعادلة هي العملية الإحصائية التي تستخدم لضبط الاختلافات في صعوبة نماذج الاختبار التي تبنى وفقًا لجدول المواصفات نفسه، وتتصف المعادلة بخاصية التماثل، أي أن الدالة المستخدمة لتحويل الدرجات على النموذج X إلى مقياس النموذج Y هي معكوس الدالة المستخدمة لتحويل الدرجات على النموذج Y إلى مقياس النموذج X، وهناك أكثر من تصميم للمعادلة، أهمها تصميم المجموعات المتكافئة وتصميم الفقرات المشتركة، وفي تصميم المجموعات المتكافئة يتقدم طلبة كل مجموعة لنموذج من نماذج الاختبار، وعندما يكون من الصعب ضمان تكافؤ المجموعات نستخدم تصميم الفقرات المشتركة (الجذع المشترك A)، ويمثل

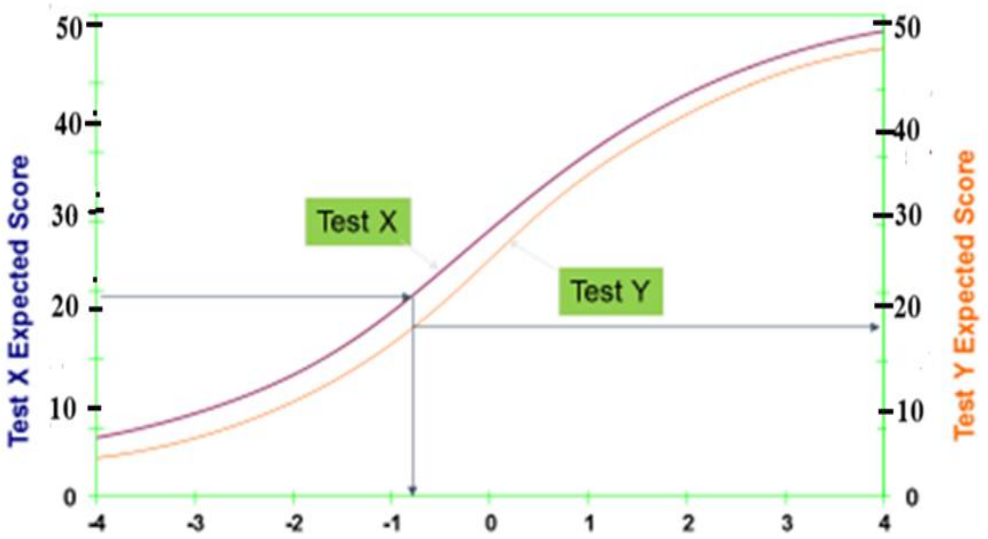
الجذع المشترك صورة مصغرة عن الاختبار يتم تضمينها لكل نموذج من نماذج الاختبار. وقد تدخل الفقرات المشتركة في حساب الدرجات (تضمين داخلي)، وقد لا تدخل في حساب الدرجات (تضمين خارجي).

وتتيح المعادلة لنماذج الاختبار فصل صعوبة نموذج الاختبار عن قدرة الطالب. ويفيد ذلك في تعديل درجات الطلبة للتحرر من الاختلافات في صعوبة نماذج الاختبار، إجراء مقارنات عادلة في أداء الطلبة عبر أوقات الاختبار المختلفة مما يوفر الصدق الداخلي لدراسة أثر التدخلات أو السياسات الجديدة، وضمان العدالة في حصول الطلبة على الدرجات المناسبة بغض النظر عن النموذج الذي تم أخذه.

والمعادلة من خلال نماذج IRT أقوى من المعادلة اعتمادًا على النظرية التقليدية في القياس، وتتم المعادلة بعدة أساليب، مثل معايرة الفقرات في النماذج المختلفة إلى المقياس المرجعي المشترك أو تحويل الدرجات إلى المقياس المرجعي بالاعتماد على منحنى خصائص الاختبار.

وهناك ميزة مفيدة لنماذج IRT، هي أن منحنى خاصية الاختبار (TCC) Test characteristic curve يمكن بناؤه كمجموع منحنيات خصائص الفقرات ICCs المتضمنة في الاختبار، ويمكن استخدام هذا المنحنى للتنبؤ بدرجة الفرد المتوقعة عند مستويات القدرة المختلفة، وعندما يتم تطوير نموذجين للاختبار نفسه، فإنه يمكن معادلة درجاتهما بسهولة من خلال استخدام منحنيات خصائص الاختبار TCCs (طريقة معادلة الدرجات الحقيقية)، كما هو مبين في الشكل 17.

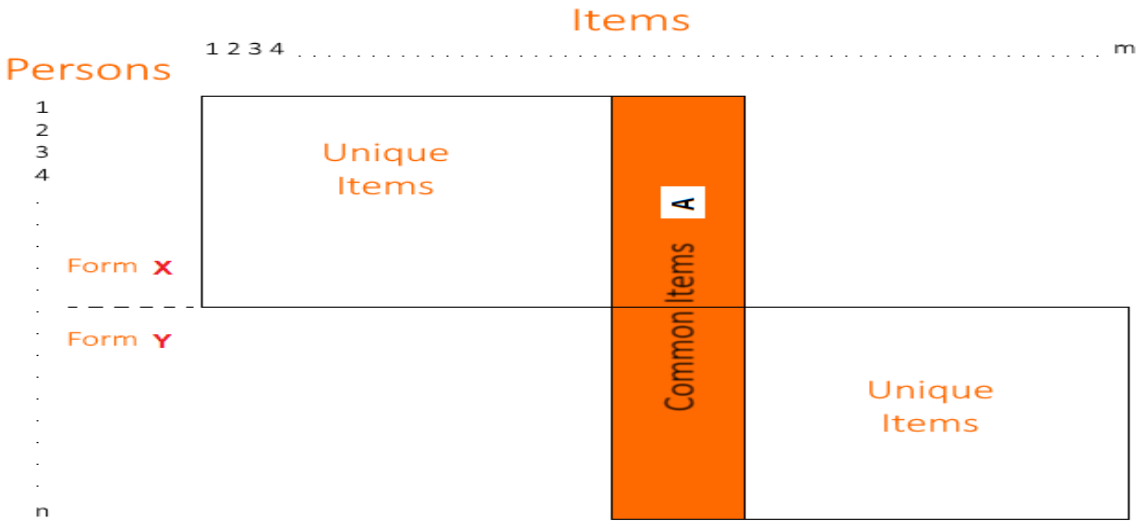
الشكل 17: معادلة نموذجي اختبار باستخدام منحنى خصائص الاختبار



ويبدو من الشكل أن النموذج Y أصعب قليلا من النموذج X، لذلك الدرجة 21 على النموذج X تعادل الدرجة 18 على النموذج Y، لاحظ أن كلتا الدرجتين ارتبطتا مع تقدير القدرة (0.8−).

وهناك طريقة أخرى لمعادلة نماذج الاختبار وهي المعايرة المتزامنة لفقرات النماذج جميعها، وفيها يتم دمج مجموعات البيانات الخاصة بنماذج الاختبار معًا فيما يعرف بالمصفوفة المتفرقة Sparse Matrix، بعد ذلك، يتم استخدام هذه المصفوفة لمعايرة الفقرات جميعها في النماذج جميعها باستخدام أي من برامج معايرة IRT، هذه المعايرة تعمل على معادلة النماذج جميعها، وتضع الفقرات جميعها والأشخاص جميعهم على مقياس مشترك. وتكون المعادلة في هذه الحالة وفق تصميم اختبار الجذر المشترك للمجموعات غير المتكافئة في المعادلة Non Equivalent Anchor Test Design (NEAT) والمبين في الشكل 18، وتقوم هذه المعايرة بمعادلة النماذج المختلفة تلقائيًا ويمكن استخدام النتائج الناتجة، وتحويلها إلى التدريج المرغوب.

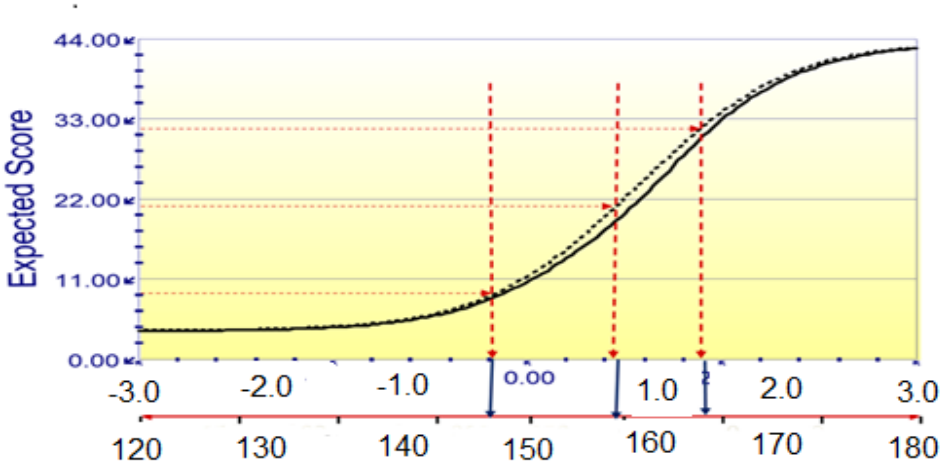
الشكل 18: تصميم الجذر المشترك للمجموعات غير المتكافئة في المعادلة.



المقياس المرجعي للدرجات

يمكن استخدام منحنى خاصية الاختبار لتحويل الدرجات إلى السلم المرغوب، ويتم التحويل بصورة خطية بضرب قدرة الفرد المقدرة ø بعدد معين a وإضافة عدد آخر b للناتج. تختار هذه الأعداد بطريقة تضمن الحصول على المقياس المرجعي المرغوب فيه من حيث الوسط الحسابي والانحراف المعياري كما في الشكل 19.

الشكل 19: التحويل من الدرجة على الاختبار إلى تدريج مرجعي بوسط 150 وانحراف معياري 10.



ويلاحظ من الشكل أن الدرجات المتوقعة 9، 21، 31 على النموذج X المنقط تذهب إلى الدرجات 147، 158، 165على التدريج المرجعي الذي وسطه 150 وانحرافه المعياري 10. التحويل يتم من خلال ضرب ø بالعدد 10 وإضافة العدد 150. في هذا التحويل من المحتمل جدا ألاً يحصل أي طالب على الدرجة 180 أو الدرجة 120. هذا التوزيع للدرجات طبيعي بقيم ثابتة لمعالمه، لذلك يتوقع أن يحصل حوالي 16% من الطلبة على درجات تزيد على أو تساوي 160، وحوالي 2.3% من الطلبة على درجات تزيد على أو تساوي 170.

ويمكن إجراء تحويل يضمن الحصول على الدرجتين 180 لمن يحصل على أعلى تقدير للقدرة و120 لمن يحصل على أدنى تقدير للقدرة، ويتم ذلك بتطبيق معادلة خاصة[44] لتحويل قيم ø إلى درجة قياسية تمتد من القيمة الصغرى SSl إلى القيمة الكبرى SSh بزيادة مقدارها h. وهذا التحويل يضمن أن تكون الزيادة على التدريج محددة، والمعادلة هي:

$$SS(ø) = h * R(\frac{1}{h}\left[SS_l + (\ ø - ø_l)\frac{ss_h - ss_l}{ø_h - ø_l}\right]$$

حيث

$SS(ø)$ العلامة القياسية التي تقابل القدرة ø

SS_l القيمة الصغرى للتدريج المرغوب

SS_h القيمة العظمى للتدريج المرغوب

$ø_l$ أقل تقدير للقدرة

$ø_h$ أعلى تقدير للقدرة

h مقدار الزيادة على التدريج المرغوب

R تدوير المقدار إلى أقرب عدد صحيح

[44]Choe, E & Han, K. (2022). Constructing a Robust Score Scale from IRT scores with informed Boundaries.

إعداد الامتحانات

تتطلب امتحانات الثانوية العامة، مستوى عاليًا من الصدق والثبات والأمن لضمان العدالة للطلبة وتوفير الثقة في النتائج، ولتحسين أمن الامتحان وتقليل احتمالية حدوث تسريب لأوراق الأسئلة، تقوم إدارة الامتحانات والاختبارات حاليًا بإجراءات صارمة لضبط عملية إعداد أوراق الامتحان وإدارة الامتحان. تتمثل هذه القيود والضوابط بتكليف عدد محدود جدًا من المشرفين الموثوقين لإعداد ورقة الامتحان بكامل أسئلتها، وتقديم مهام وخدمات لوجستية معقدة تتعلق بإخراج ورقة الامتحان وطباعتها وتغليفها وتوزيعها وحفظها وتطبيقها في جدول زمني طويل للغاية، ومع ذلك فإنه يصعب ضمان جودة كافية في ورقة الامتحان من ناحية وفي عملية تطبيقها من ناحية أخرى، وبالرغم من الاعتماد على نموذج واحد للامتحان في كل دورة للامتحان، فإنه يتم اكتشاف أخطاء علمية أو مطبعية في بعض الأسئلة، وتكون هناك احتمالية عالية للغش، ويحدث تسريب للأسئلة خلال جلسة الامتحان، مما يعطي فرصة لنقد الامتحان والتشكيك في نتائجه.

للتغلب على كل هذه التحديات، لا بد من إعداد نماذج متكافئة لكل امتحان، وتطبيق تلك النماذج بصورة إلكترونية بقدر ما تسمح به البنية التحتية للمراكز وقاعات الامتحان، ويتم تكوين النماذج المتكافئة للامتحان بالاعتماد على جدول المواصفات نفسه لضمان صدقها من ناحية وتكافؤها من ناحية أخرى، وتكوين النماذج المتكافئة وفق النظرية التقليدية للقياس، وفي غياب معرفة خصائص الفقرات عملية صعبة جدًا وتكاد تكون مستحيلة، إلا أن استخدام النظرية الحديثة في القياس (نظرية الاستجابة للفقرة) في بناء بنوك الأسئلة يساهم إلى حد كبير في إعداد النماذج المتكافئة للامتحان.

ويعد بنك الأسئلة نقطة البداية ونقطة النهاية في دورة تطوير الامتحان، أي لا بد من بناء البنك أولًا وفقًا لمجموعة من المعايير الدولية، ولا بد من المحافظة على استدامته وتزويده بفقرات جديدة مجربة في كل دورة امتحان، ويتم سحب فقرات (أسئلة) النماذج وفق مواصفات محددة من بنك مؤسس جيدا لأسئلة الامتحان، تم تطويره وهيكلته بصورة تبين ما الذي تقيسه كل فقرة، وما مستواها، وما العمق المعرفي لها، وما صعوبتها، وما المعلومات التي تقدمها في قياس السمة المستهدفة في القياس، وغيرها من الخصائص، وهذا يتيح إمكانية تفحص خصائص كل نموذج قبل استخدامه الفعلي.

ويحتوي بنك الأسئلة على مجموعة كبيرة جدًا من الفقرات المصاغة حسب القواعد والإرشادات المعتمدة عالميًا في كتابة الفقرات، والتي تم بناؤها وفقًا لجدول مواصفات متفق عليه وتم مراجعتها علميًا وسيكومتريًا، وتدقيقها لغويًا، وهي معلومة المواصفات والخصائص لأنه تم تجريبها ميدانيًا على عينة من المجتمع المستهدف في ظروف اختبارية طبيعية، وتفي خصائصها الإحصائية بالمعايير المعتمدة لإيداعها في البنك، وبعد تصميم

النماذج يتم تطبيقها في ظل ظروف معيارية موحدة، وتجمع الاستجابات وتصحح وتحلل، ويتم تقدير معلمات الفقرات في النماذج المختلفة بصورة متزامنة، ثم يتم إيداع تلك القيم في البنك، وهذا يتيح تتبع سلوك الفقرات وما يحصل عليها من تغييرات، واتخاذ قرارات حول الاستمرار باستخدامها من عدمه، كما سيتم تدريج علامات الطلبة على السلم المعتمد في كشوفات العلامات والتقارير.

ويشرف على بناء بنوك الأسئلة لجنة استشارية عليا[45]، ولكل امتحان من امتحانات الثانوية العامة سيتم إنشاء بنك آمن للفقرات وفق هيكلية تسمح بسحب الفقرات وفق المواصفات المطلوبة للحصول على امتحان تتصف نتائجه بالصدق والثبات والعدالة، وهناك العديد من الجهات[46] التي تشترك في بناء البنك دون أن يكون لها الحق بالاطلاع على الفقرات إلا بالحد الذي تسمح به مشاركتها، ولكل بنك هناك لجنة[47] تقر هيكلية البنك والبيانات التي يجب تضمينها مع الفقرات والشروط التي يجب أن تتوافر في مؤلفي الفقرات والمراجعين لها ومعايير الاحتفاظ بها، كما تقر الإطار العام للامتحان، وجدول مواصفاته، وعدد الفقرات وطبيعتها، وعدد النماذج، وزمن الامتحان، وشكل الامتحان، وكيفية التعامل مع التظلمات، ويقوم مطور الاختبار بتضمين كل نموذج من نماذج الاختبار فقرات جديدة بقصد التجريب ولا تدخل في حساب الدرجات، وذلك من أجل إثراء البنك واستدامته.

ويتوقع أن تكتمل المرحلة الأساسية الأولى في كتابة الفقرات لبنوك الأسئلة في مباحث اللغة العربية واللغة الإنجليزية والتربية الإسلامية وتاريخ الأردن في نهاية شهر أيار 2025، وسيتم تجريب الفقرات جميعها في بيئة اختبارية فعلية عندما تتاح أول فرصة لتطبيق الامتحان إلكترونيًا ، وفق تصميم يسمح بمعايرة الفقرات وتعريف مقياس الدرجات ومعادلة نماذج الاختبار جميعها بصورة متزامنة وفق تصميم الفقرات المشتركة الداخلية للمعادلة، وستبدأ مرحلة كتابة الأسئلة لبقية امتحانات الثانوية العامة عند الانتهاء من تأليف الكتب المدرسية الخاصة بها.

وتجدر الإشارة إلى أن بنوك الأسئلة تسمح بالتعامل مع أشكال الفقرات كافة، وتسمح باستخدام الصور والخرائط والأشكال والتسجيلات الصوتية والفيديوهات والقوانين والمعادلات وغيرها، ويسمح في مجال اللغات بقياس مهارتي الاستماع والتحدث، كما توفر إمكانية عقد الامتحانات في عدة دورات خلال العام الدراسي،

[45] يشكل اللجنة المركز الوطني لتطوير المناهج برئاسة رئيس المجلس الأعلى للمركز

[46] اللجنة الاستشارية للبنك، لجنة الامتحان، مدير البنك، مطور الامتحان، مجموعة كبيرة من مؤلفي الأسئلة، مجموعة من مراجعي الأسئلة، ومراجع سيكومتري

[47] يشكل اللجنة المركز الوطني لتطوير المناهج برئاسة مدير وحدة القياس والتقويم في المركز

وفي عدة جلسات في الدورة الواحدة، ويقل عدد هذه الجلسات مع اكتمال البنية التحتية التي تلزم للتطبيق الإلكتروني للامتحانات.

جدول مواصفات الامتحان Table of Examination Specification

يتم إعداد أي امتحان من امتحانات الثانوية العامة وفقًا للمواصفات المعتمدة والمعلنة في دليل الامتحان والغرض الرئيس من مواصفات ورقة الامتحان هو توصيل مخطط التقييم بوضوح واتساق لكل من اللجنة الاستشارية المشرفة ولجنة الامتحان ومطور الامتحان وأصحاب المصلحة الآخرين من طلبة ومعلمين وأولياء الأمور وغيرهم من المهتمين، وتوفر مواصفات الامتحان معلومات أساسية تتعلق ببنية الامتحان ومدته وعدد فقراته وأنواعها وصعوبتها والعمق المعرفي لها، وعلامة كل منها، وناتج التعلم المقاس، إضافة إلى التوازن عبر مستويات الأهداف أو مجالات المهارة.

وتختلف مواصفات الامتحان ومستوى التفاصيل فيها من مبحث إلى آخر، وذلك لاختلاف الخصائص المحددة لكل مبحث، ويجوز تعزيز مواصفات بعض الامتحانات بأدلة توضيحية، قد تختلف في نوع وكمية المعلومات التي تتضمنها، فمثلًا قد يحتوي دليل اللغة الإنجليزية على إرشادات مفصلة نسبيًا لمؤلفي الفقرات، ترشدهم إلى كيفية اختيار نص القراءة، بينما يتضمن دليل الرياضيات قائمة بالمهام التي سيتم استخدامها من أجل فحص وتقييم مهارات التفكير العليا لدى الطلبة، ويبين الجدول 16 نموذج لجدول مواصفات لامتحان من 50 فقرة.

الجدول 16: نموذج لجدول مواصفات لامتحان مكون من 50 فقرة اختيار من متعدد.

نتاجات التعلم	الوزن النسبي للوحدات%	الوزن النسبي للدروس%	عدد الفقرات	الصعوبة		
				منخفضة	متوسطة	مرتفعة
1	55	16	8	2	3	3
2		10	5	2	2	1
3		9	4	2	1	1
4		12	6	1	2	3
5		8	4	1	1	2
1	45	12	6	2	2	2
2		9	5	2	2	1
3		12	6	0	4	2
4		12	6	2	2	2
	100	100	50	14	19	17

العمق المعرفي(مرتفعة)		
R	A	K
2	1	
1		
	1	
2	1	
	1	1
1	1	
	1	
	2	
1		1
7	8	2

العمق المعرفي(متوسطة)		
R	A	K
1	1	1
1		1
	1	
1	1	
1		
	1	1
1	1	
1	1	2
1	1	
7	7	5

العمق المعرفي(منخفضة)		
R	A	K
1	1	
	1	1
1		1
		1
		1
	1	1
	1	1
	1	1
2	5	7

يشير العمود الأول من اليمين في الجزء العلوي من الجدول إلى أهداف التعلم أو نتاجات التعلم المنوي قياسها بالامتحان، ويبين العمود الثاني مجالات المحتوى الرئيسة (الوحدات الدراسية) والوزن النسبي لكل منها، فالوزن للمجال الأول في المحتوى هو 55%، وتحدد الأوزان النسبية وفقًا لمعيار محدد متفق عليه، فقد يعكس الوزن النسبي للمجال الرئيس أهميته أو حجمه أو الوقت المستغرق في تعليمه، ويتم توزيع الوزن النسبي لمجال المحتوى الرئيس على المحتويات الفرعية كما في العمود الثالث، وفي العمود الرابع عدد الفقرات، وهو عبارة عن ناتج ضرب الوزن النسبي في العدد الكلي لفقرات الاختبار (50 فقرة هنا)، وتبين الأعمدة الثلاثة الأخيرة توزيع الفقرات حسب درجة الصعوبة، وتبين الأجزاء السفلية توزيع الفقرات حسب الصعوبة والعمق المعرفي أو درجة التعقيد (K المعرفة، A التطبيق، R التفكير).

ويفيد جدول المواصفات في تحقيق صدق المحتوى للامتحان من خلال التركيز في الامتحان على الفقرات المهمة في المحتوى، وفي الوقت نفسه يرشد مطور الاختبار لانتقاء الفقرات من بنك الأسئلة وفقًا للمواصفات المحددة والتي تتمثل بنتاجات التعلم والمحتوى والصعوبة والعمق المعرفي، وعليه فإن جدول المواصفات أداة فعالة في التخطيط لاختبار جديد وفي تقييم الاختبار وفي تشكيل نماذج الاختبار. إن الحفاظ على ثبات بنية الاختبار أمر ضروري للغاية من أجل الحفاظ على قابلية المقارنة بين الاختبارات في كل مجال من مجالات الموضوع، وذلك من خلال الحفاظ على نفس المقياس لجميع أشكال الاختبار.

جمع المعلومات وتحليلها وإصدار النتيجة

في المرحلة الحالية سيتم استخدام أسئلة اختيار من متعدد في الامتحانات جميعها، ويمكن استخدامها لقياس مستويات معرفية متنوعة، ويفضل تطبيق الامتحانات إلكترونيًا بشرط توفر البنية التحتية التكنولوجية

اللازمة، وتدريب الطلبة على استخدامها. وفي مجال اللغة العربية واللغة الإنجليزية سيكون هناك جزء ورقي يتعلق بكتابة موضوع الإنشاء، ولا ينصح في المرحلة الحالية قياس مهارتي الاستماع والتحدث إلكترونيًا، وننصح أن يتم قياسهما مدرسيا. ويتم تسجيل إجابات الطلبة لفقرات الامتحانات والاستبانات السياقية إلكترونيًا، وتصحح إلكترونيًا. أما موضوع الإنشاء فيصحح يدويًا أي بتدخل بشري. وسيتم تحليل استجابات الطلبة باستخدام نظرية الاستجابة للفقرة لجميع نماذج الاختبار بصورة متزامنة مع مجموعة من الفقرات المشتركة المختارة بعناية من نموذج مرجعي بمواصفات محددة. وبذلك يتم معايرة الفقرات وتدريج جميع النماذج وفقا للتدريج المرجعي المعتمد بوسط حسابي 150 وانحراف معياري 10. ويتم تقدير الثبات للدرجات وتحديد علاقتها بالمعلومات السياقية الخاصة بالطالب وأسرته ومعلمه ومدرسته، ويبلغ الطالب إلكترونيًا بنتيجته، وبناء على طلبه يمكنه طباعة شهادة (كشف) بعلامته في كل امتحان تقدم له، ويبين الكشف علامته على التدريج المرجعي المعتمد (بوسط 150 وانحراف معياري 10)، ومستوى الأداء الذي يصنف فيه (مبتدئ، إتقان جزئي، إتقان تام، متقدم)، ودورة الامتحان التي تقدم لها، ولن يكون هناك رسوب في الامتحان.

وسيعقد الامتحان لكل مبحث في عدة دورات للامتحان في كل عام (3–4 دورات بين شهري آذار وكانون الأول) بشرط توافر الأجهزة التكنولوجية الكافية، وأن بنك الأسئلة يحتوي على ما يكفي من الأسئلة المقبولة والمختبرة مسبقًا ليكون قادرًا على دعم تطوير وإنشاء نماذج متعددة، مما يعطي للطالب مزيد من الحرية في اختيار الدورة التي تتناسب مع ظروفه واستعداداته، وله حرية اختيار المباحث ضمن خطته الدراسية التي سيتقدم لامتحاناتها في تلك الدورة. ويمكن للطالب في الدورة التي سجل لها التقدم للامتحان في مبحث واحد إن أراد أو التقدم لجميع المباحث في خطته. وبالرغم من أنه لا يوجد رسوب في الامتحانات، فإنه يجوز للطالب توزيع امتحاناته على عدة دورات ضمن فترة زمنية لا تزيد على ثلاث سنوات، كما يجوز له إعادة تقديم أي امتحان أكثر من مرة خلال تلك المدة لرفع علامته، وله حرية استخدام العلامة الأعلى لأغراض العمل أو الالتحاق بمؤسسات التعليم العالي. وستكتب التقارير حول كل امتحان على مستوى المدرسة والمديرية والنظام. وستستخدم المعلومات في تلك التقارير بطريقة تكوينية، لتحديد مجالات التحسين وتوجيه السياسات واتخاذ القرارات القائمة على الأدلة.

الخلاصة والتوصيات بخصوص امتحانات الثانوية العامة

لتنفيذ امتحانات الثانوية بصورة دقيقة وموضوعية وبعدالة، نوصي بما يأتي:

1. تعديل التشريعات التربوية فيما يتعلق باستخدام امتحانات الثانوية العامة لأغراض الالتحاق بمؤسسات التعليم العالي، مع التأكيد على ألا يكون المعيار الوحيد في ذلك.
2. تقليل عدد الامتحانات التي يتقدم لها الطالب لتكون 6 امتحانات على الأكثر.
3. عدم الاستعجال بالذهاب إلى الاختبارات الرقمية إلا بعد الاطمئنان على مناسبة البنية التحتية، وأن جميع الطلبة والقائمين على الامتحانات وفنيي المختبرات قد خضعوا للتدريب الكافي المتعلق باستخدام منظومة بنوك الأسئلة.
4. المحافظة على أمان بنوك الأسئلة وحمايتها وصيانتها وإدامتها وإثرائها بفقرات جديدة، وأن توفر وزارة الاقتصاد الرقمي الضمانات بعدم اختراق المنظومة كونها الجهة المسؤولة والمشرفة على السحابة الحكومية التي ستحمل عليها البرمجية.
5. تأجيل استخدام الأسئلة مصاغة الإجابة في الاختبارات الرقمية إلى مرحلة لاحقة.
6. تأجيل استخدام الاختبارات التكيفية إلى مراحل قادمة.
7. الاختبارات الأدائية والتدريب الميداني ومهارتي التحدث والاستماع يتم اختبارها مدرسيا.
8. تقديم التسهيلات اللازمة لذوي الحاجات الخاصة.
9. تحديد مستويات الأداء بحسب أسلوب أنجوف المعدل.
10. إجراء دراسة علمية واسعة تتعلق بالقدرة التنبؤية لامتحانات الثانوية العامة بالتحصيل في مؤسسات التعليم العالي وفي التخصصات المختلفة.

الجزء الثالث: بناء القدرات وتوجيه السياسات

بناء القدرات:

بعد نشر مسودة الإطار الوطني لتقييم الطلبة، سيكون من الضروري القيام بعملية تطوير مهني واسعة النطاق من أجل إكساب الجهات الفاعلة الرئيسية المهارات والفهم والكفاءات من أجل تنفيذ التقييمات التي تضمنها الإطار بطريقة عادلة ومنصفة، ويقوم نظام التقييم المقترح على قاعدة واسعة من الأدلة والبيّنات الحقيقية حول أداء الطلبة إضافة إلى نتائج الاختبارات الوطنية الموحّدة، ويحتاج أولياء الأمور وأرباب العمل ومؤسسات التعليم العالي إلى فهم التحولات والتغييرات في نظام التقييم الوطني، وهذا يستوجب قيام الوزارة بحملة توعية إعلامية تشرح ما يحدث وتبين الغاية من ذلك.

ويجب أن يتم بناء القدرات على كافة مستويات النظام من مركز وزارة التربية إلى مديريات التربية وعلى مستوى المدارس والمعلمين، ويتعين على المعلمين الإلمام بالمعرفة التقييمية الأساسية لكي يطبقوا الكفاءات الرئيسة في سياق التقييمات الصفية التي يتولون عملية تنفيذها وإدارتها، ويعد الضعف في الإلمام بالتقييم أحد الأسباب الرئيسة وراء عدم إنشاء أنظمة تقييم متوازنة. والضعف موجود على المستويات كافة ومن بينها الصفوف الدراسية حيث تكون الحاجة ملحة إلى إجراء المزيد من التطوير المهني للمعلمين. وتتسم برامج التطوير المهني المستمرة بأنها أكثر استدامة وترتكز على المدرسة ومجتمعات ممارسة التعلم في المدارس.

وتقوم معظم التقييمات للطلبة على مدى إلمام المعلمين بالتقييم، إذ إنهم يتولّون إجراء تقييمات تشخيصية، وإعداد الطلبة للتقييمات الوطنية، وتصميم التقييم الصفي وتنفيذه، وتفسير وتطبيق المعلومات التي يتم الحصول عليها من التقييمات المختلفة.

ويتمتع المعلمون الذين لديهم مستوى عالٍ من الإلمام بالتقييم بأن لديهم معرفة تفصيلية وقوية بأدوات التقييم والمهارات ذات الصلة بإصدار أحكام حول الطلبة، ولديهم القدرة على تفسير نتائج التقييم بشكل فعال، والقدرة على تطبيقها في دعم تعلم الطالب أو لتقييم برنامج التعلم وتحسينه لمجموعات مستقبلية من الطلبة، ومساعدة الطلبة على الاستعداد بشكل فعال للتقييمات الوطنية والدولية، واستخدام التقييمات التكوينية في الصفوف الدراسية، وتحسين جودة التقييمات الصفية. ويتطلب إدماج التكنولوجيا في نظام التقييم واستخدام النظرية الحديثة في إنشاء بنوك الأسئلة ومعايرة الفقرات ومعادلة النماذج ووضعها على تدريج موحد إكساب التلاميذ والمعلمين ومشرفي القياس والتقويم سواء في الوزارة أو المركز الوطني لتطوير المناهج، والفرق الفنية في مركز الملكة رانيا العبد الله لتكنولوجيا التعليم والمعلومات مهارات تكنولوجيا المعلومات والاتصالات، ويبين الجدول 17 الحاجات التدريبية والأطراف المستهدفة.

الجدول 17: الحاجات التدريبية والمستهدفون ببناء القدرات.

م	الحاجات التدريبية	المستهدفون ببناء القدرات				
		المدرسة (المعلم، منسق المبحث، مدير المدرسة، مساعد المدير، لجنة التقييم)	مديرية التربية والتعليم (مدير التربية والتعليم، مدير الشؤون التعليمية والفنية ، المشرف التربوي في الميدان، مشرف القياس)	مركز الوزارة		الشركاء والمعنيون من أصحاب المصلحة (عمداء كليات التربية، امناء عامون للتعليم العالي، أكاديمية الملكة رانيا، ...)
				المديريات الفنية (مشرف القياس، مؤشرات الأداء/عضو مناهج، إشراف، تدريب،...)	القيادات العليا (وزير التربية، الأمناء العامين، مديري الإدارات)	
1	التعريف بالإطار الوطني للتقييم	●	●	●	●	●
2	التقييم الصفي (التقييم الذاتي، تقييم الرفاق، التقييم الشفوي، ...)	●	●	●		
3	استراتيجيات التقييم من أجل التعلم	●	●	●		
4	تقييم المهارات العقلية العليا والقدرات ما وراء المعرفية	●	●	●		
5	تقييم الأداء (دراسة الحالة، المشاريع،...)	●	●	●		
6	التقييم التشخيصي كأداة للتعلم الفعال	●	●	●		
7	نتاجات التعلم ومعايير ومؤشرات الأداء وتوظيفها	●	●	●		
8	بناء الاختبار (تحليل المحتوى التعليمي، بناء مخطط الاختبار، صياغة الأسئلة، اخراج ورقة الامتحان...)	●	●	●		
9	استخدام نظرية استجابة الفقرة (بنوك الأسئلة، الاختبارات الإلكترونية، الأداء التفاضلي والكشف عن التحيز، ...)		مشرف القياس	مشرف القياس		
10	إدارة الاختبار	●	●	●	●	
11	أنظمة إعطاء الدرجات	●	●	●	●	●

م	الحاجات التدريبية	المستهدفون ببناء القدرات				
		المدرسة (المعلم، منسق المبحث، مدير المدرسة، مساعد المدير، لجنة التقييم)	مديرية التربية والتعليم (مدير التربية والتعليم، مدير الشؤون التعليمية والفنية ، المشرف التربوي في الميدان، مشرف القياس)	مركز الوزارة		الشركاء والمعنيون من أصحاب المصلحة (عمداء كليات التربية، امناء عامون للتعليم العالي، أكاديمية الملكة رانيا، ...)
				المديريات الفنية (مشرف القياس، مؤشرات الأداء/عضو مناهج، إشراف، تدريب،...)	القيادات العليا (وزير التربية، الأمناء العامين، مديري الإدارات)	
12	تحليل نتائج الاختبار	●	●	●		
13	كتابة التقارير عن الاختبارات	●	●	●	●	●
14	كيفية الاستفادة من نتائج التقييمات في تحسين تعلم الطلبة	●	●	●	●	●

توجيه السياسات

سياسة دمج ذوي الاحتياجات الخاصة

يجب أن تراعي التقييمات الصفية والامتحانات الوطنية والدولية المستخدمة في الأردن احتياجات ذوي الاحتياجات الخاصة سواء كانت سمعية أو بصرية أو غيرها، وتكييف التقييمات لهم بطريقة تتيح لهم إثبات تعلمهم، ومع استخدام الاختبارات الرقمية تزداد فرص إدماج ذوي الاحتياجات الخاصة في التقييمات المختلفة.

سياسة الوصول إلى البيانات

يجب إنشاء قاعدة بيانات واسعة حول نتائج الطلبة في التقييمات الوطنية وامتحانات الثانوية العامة والدراسات الدولية، وتضم المعلومات السياقية الخاصة بالطلبة ومدارسهم وأسرهم، وإتاحة فرصة وصول الباحثين وطلبة الدراسات العليا والمهتمين إلى البيانات لتحليلها وتفسيرها وإجراء الأبحاث والدراسات المتعمقة حولها.

سياسة تحليل البيانات

يغلب على كثير من التحليلات للبيانات السطحية، عدم استخدام الأساليب الإحصائية المناسبة، وعدم مراعاة طبيعة البيانات (كمية أو كيفية)، وهذا يستوجب إعطاء المزيد من الاهتمام بتحليل البيانات واستخدام أساليب إحصائية متقدمة سواء في الامتحانات الوطنية أو الدولية.

سياسة القبول في مؤسسات التعليم العالي

يستخدم المعدل في امتحانات الثانوية العامة معيارًا وحيدًا للقبول في مؤسسات التعليم العالي، ويحتسب المعدل بإعطاء 30% للمباحث المشتركة الأربعة و70% للمباحث التخصصية والاختيارية، وهذا التوزيع لا يعكس الأهمية النسبية للمباحث بطريقة متوازنة من ناحية، ويتعارض مع التوجهات العالمية التي تؤكد على ضرورة تعدد المعايير من جهة أخرى، كما أنه يعزز لدى الطلبة عدم الارتباط بالمدرسة والتعويض عن ذلك بالتدريس الخصوصي والذهاب إلى المراكز، لذلك يجب إعادة النظر في سياسة القبول في مؤسسات التعليم العالي. ويجب أن يكون معدل الثانوية العامة واحدًا من عدة معايير للقبول في مؤسسات التعليم العالي.

سياسة إنشاء مركز مستقل متخصص في الاختبارات

يوجد في كثير من دول العالم هيئة مستقلة حكومية تعنى بكل أشكال التقييمات، ويتجمع في هذه الهيئة المختصون وأصحاب الخبرات العالية في مجال القياس والتقويم، مما ينعكس إيجابًا على جميع التقييمات التي تتولى مسؤولية بنائها وإدارتها وتطويرها، من خلال تبني أفضل الممارسات والمعايير الدولية. ويوفر ذلك درجة عالية من الحيادية والعدالة والصدق والثبات والموضوعية والشفافية في تلك التقييمات، دون وجود تداخلات بين تلك التقييمات سواء من حيث الأغراض أو المحتوى، ودون وجود تشابك في الأدوار والمسؤوليات، وعليه ننصح باستحداث هيئة أو مركز وطني يُعنى بكل التقييمات في الأردن تصميمًا وتطويرًا وتحليلًا للنتائج.

سياسة التوزيع العادل للموارد والمصادر

تؤشر الفروق في النتائج على الاختبارات الدولية والوطنية بين الذكور والإناث وبين المدن والقرى إلى ضرورة إعادة النظر في توزيع الموارد والمصادر المتاحة بعدالة على الفئات المختلفة، وأن توجه المصادر بصورة أكبر إلى حيث توجد الحاجة.

سياسة تمهين التعليم

هناك مؤشرات متعددة على ارتباط الأداء في التقييمات المتعددة بأداء المعلمين. ويحدد أداء المعلم السقف الأعلى لأداء أي نظام تعليمي، لذلك لا بد من تمهين التعليم والعمل على أن تكون جاذبة ولا بد من العمل الجاد على إعداد المعلمين وتمكينهم قبل الخدمة، ولا بد من تنميتهم مهنيًا أثناء الخدمة.

المراجع

- الثوابية، أحمد. (1994). فاعلية العلامات في الصفوف الثلاثة الأخيرة كمتنبئ بتحصيل الطالب في امتحان شهادة الدراسة الثانوية العامة والتحصيل الجامعي لدى عينة من الطلبة الملتحقين بالجامعة الأردنية. رسالة ماجستير غير منشورة، الجامعة الأردنية، عمان، الأردن.
- جرادات، ضرار. (2003). تقييم القدرة التنبؤية لمعدل امتحان الثانوية العامة بالمعدل التراكمي عند التخرج من الجامعة: دراسة حالة. أبحاث اليرموك: سلسلة العلوم الإنسانية والاجتماعية 19(1)، 383-400.
- الرشدان، محمود. (2002). الترابط بين المعدلات التراكمية في الجامعات الأردنية والمعدلات في الثانوية العامة، مجلة اتحاد الجامعات العربية،40 ،. 203 – 162
- السعيد، دجانة. (2020). أسباب تدني نتائج الطلبة ذوي عمر 15 عاما في البرنامج الدولي لتقييم الطلبة "Pisa" لمادة الرياضيات من وجهة نظر المعلمين والمشرفين التربويين في محافظة العاصمة عمان. رسالة ماجستير، جامعة الشرق الأوسط، عمان، الأردن.
- الصمادي، يحيى وظاظا، حيدر وغرايبة، عايش واليونس، يونس. (2010). معدل الثانوية العامة والمعدل التراكمي الجامعي بصفتهما متنبئين بمستوى تحصيل طلبة الجامعات الأردنية في امتحان الكفاءة الجامعية. المجلة الأردنية في العلوم التربوية،(2)6 ،. 147-159
- ظاظا، حيدر والمشايخ، جهاد. (2016). مقارنة القدرة التنبؤية لمعدل الثانوية العامة الأردنية ومعدلات الثانوية للبرامج الأجنبية للطلبة الأردنيين بالمعدل التراكمي في الجامعات الأردنية الحكومية. دراسات العلوم التربوية، 43(3)،. 637-661
- العناتي، جهاد. (2014). القدرة التنبؤية لمعدل العلامات الجامعية ومعدل الثانوية العامة في الأداء الوظيفي لدى موظفي القطاع العام في المملكة الأردنية الهاشمية. المجلة الدولية التربوية المتخصصة، 3(3)،. 49-69
- غرايبة، عايش. (1991). مقارنة فعالية نظام قبول مقترح في الجامعة الأردنية يقوم على أساس أوراق الثانوية العامة مع فعالية نظام القبول الحالي. رسالة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن.
- فريحات، أيمن. (2002). مقارنة القدرة التنبؤية لامتحان شهادة الثانوية العامة بين الجامعات الحكومية والخاصة. رسالة ماجستير غير منشورة، جامعة اليرموك، اربد، الأردن.

- القوابعة، محمد. (2002). القدرة التنبؤية لامتحان شهادة الدراسة الثانوية العامة في التحصيل الأكاديمي الجامعي. رسالة ماجستير غير منشورة، جامعة مؤتة، الكرك، الأردن.
- اللطايفة، محمد علي. (1998). تعظيم الصدق التنبؤي لامتحان شهادة الثانوية العامة في التحصيل الجامعي. رسالة ماجستير غير منشورة، جامعة اليرموك، اربد، الأردن.
- Ababneh, I., Lebdih, K. and Tweissi, A. (2014). Mapping of Assessments in Jordan Monitoring and Evaluation Partnership [MEP] Project, World Education Inc.

- Addey, C. (2015). Participating in international literacy assessments in Lao PDR and Mongolia: A global ritual of belonging. Literacy as numbers: Researching the politics and practices of international literacy assessment, 147-164.
- Addey, C. (2017). Golden relics & historical standards: how the OECD is expanding global education governance through PISA for Development. Critical Studies in Education, 58(3), 311-325.
- Addey, C., & Sellar, S. (2019). Rationales for (non) participation in international large-scale learning assessments. Education Research and Foresight: Working Papers. UNESCO. Paris, France.
- Bray, M., Kobakhidze, M. N., & Suter, L. E. (2020). The challenges of measuring outside-school-time educational activities: Experiences and lessons from the Programme for International Student Assessment (PISA). Comparative Education Review, 64(1), 87-106.
- Darling-Hammond, L. (2014). What can PISA tell us about U.S. education policy? New England Journal of Public Policy, 26(1), 1-14.
- Ferrer, G., & Fiszbein, A. (2015). What has happened with learning assessment systems in Latin America? Lessons from the Last Decade of Experience. Washington, DC: World Bank.
- Huang, J., Tang, Y., He, W., & Li, Q. (2019). Singapore's School Excellence Model and student learning: evidence from PISA 2012 and TALIS 2013. Asia Pacific Journal of Education, 39(1), 96-112.
- Janus, M., & Offord, D. R. (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportment, 39(1), 1.

- Kamens, D. H. 2013. Globalization and the emergence of an audit culture: PISA and the search for 'best practices and magic bullets. H. D. Meyer and A. Benavot (eds), PISA, Power, and Policy: The Emergence of Global Educational Governance. Wallingford, UK, Symposium Books.
- Kamens, D. H. (2015). A maturing global testing regime meets the world economy: Test scores and economic growth, 1960–2012. Comparative Education Review, 59(3), 420-446.
- Klees, S. J. (2021). Economics, development, and comparative and international education. In Comparative and international education: Leading perspectives from the field (pp. 55-70).
- Kuttab, D. (2017). How to reform the education system? The Jordan Times, 17 May 2017.
- Leino, K., Nissinen, K., & Sirén, M. (2022). Associations between teacher quality, instructional quality and student reading outcomes in Nordic PIRLS 2016 data. Large-Scale Assessments in Education, 10(1), 25. https://doi.org/10.1186/ s40536-022-00146-4.
- Lockheed, M. (2013). Causes and consequences of international large-scale assessments in developing countries. H. D. Meyer and A. Benavot (eds), PISA, Power, and Policy: The Emergence of Global Educational Governance. Wallingford, UK, Symposium Books.
- Lockheed, M., Prokic-Breuer, T. and Shadrova, A. (2015). The Experience of Middle-Income Countries Participating in PISA 2000-2015. Washington DC/Paris, The World Bank/OECD Publishing.
- Mourshed, M., Chijioke, C., & Barber, M. (2011). How the worlds most improved school systems keep getting better. Voprosy obrazovaniya/educational studies Moscow, (1), 7-25.
-
- Najdawi, Z., Kardan, R., Zuraik, I., Al Shobaki, Z., Alarood, S., & Dardas Latefa . (2022). Depressive symptoms among students pursuing the general secondary education certificate examination (Tawjihi): A national study. Published online: International Journal of Mental Health, 53(3).
- Nilsen, T., Kaarstein, H., & Lehre, A. C. (2022). Trend analyses of TIMSS 2015 and 2019: School factors related to declining performance in mathematic.
- Obeidat, O and Dawani, Z. (2014). Disseminating and Using Student Assessment Information in Jordan, SABER, The World Bank.

- OECD (2013). PISA 2012 results in focus: What fifteen-year-olds know and what they can do with what they know. Paris: OECD Publishing.
- OECD (2014), PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014), OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208780-en.
- Ounce Prevention Fund. (2017). Starting Smart: How Early Experiences Affect Brain Development. (2nd ed.). Available at: [https://www.theounce.org/wpcontent/uploads /2017/03/StartingSmart](https://www.theounce.org/wp-content/uploads/2017/03/StartingSmart).
- Reforming Tawjihi in Venture, August 24, 2017
- Phillips, G. W. and Jiang, T.)2011, April (. Using PISA to Internationally Benchmark State Performance Standards. Paper presentation at the annual conference of the American Educational Research Association, New Orleans, Louisiana 2011.
- Rohatgi A., Hatlevik O. E., Björnsson J. K. (2022). Supportive climates and science achievement in the Nordic countries: Lessons learned from the 2015 PISA study. Large-Scale Assessments in Education, 10, 12. https://doi.org/10.1186/s40536-022-00123-x
- Takayama, K., Waldow, F., & Sung, Y.K. (2013). Finland has it all? Examining the media accentuation of 'Finnish education' in Australia, Germany, and South Korea. Research in Comparative and International Education, 8, 307-325.
- Teig, N., & Steinmann, I. (2023). Leveraging large-scale assessments for effective and equitable school practices: the case of the Nordic countries. Large-Scale Assessments in Education, 11(1), 21.
- UNICEF. (2016). Investing in early childhood development essential to helping more children and communities thrive, new Lancet Series finds
- Wang, H., & Yang, S. (2024). Research on student assessment and education reform based on Singapore's excellent school model. Educational Administration: Theory and Practice, 30(3).