

المرجع في التشخيص والتعرف على ذوي الاستثنائية

ذوي الإعاقة والموهوبين ومزدوجي الاستثنائية



د. عمر عبدالله الصمعاني

المرجع في التشخيص والتعرف على ذوي الاستثنائية

ذوي الإعاقة والموهوبين ومزدوجي الاستثنائية

د. عمر عبدالله الصمعاني

ح) عمر عبدالله الصمعاني ، ١٤٤٨ هـ

السمعاني ، عمر عبدالله صالح
المرجع في التشخيص والتعرف على ذوي الاستثنائية. /
السمعاني ، عمر عبدالله صالح -. الرياض ، ١٤٤٨ هـ

رقم الإيداع: ١٤٤٨/٧٠٠٨
ردمك: ٩٧٨-٦٠٣-٠٦-٧٧١٤-٦

المرجع في التشخيص والتعرف على ذوي الاستثنائية
ذوي الإعاقة والموهوبين ومزدوجي الاستثنائية

هذا الكتاب منشور بموجب رخصة المشاع الإبداعي: نسب المصنف، غير تجاري، منع الاشتقاق ٤.٠ دولي (CC BY-NC-ND 4.0). فيجوز نسخه وتوزيعه ومشاركته كاملاً أو مقتطعات منه، بأي وسيط أو صيغة، للأغراض التعليمية والبحثية غير التجارية، بشرط نسبته إلى مؤلفه دون تعديل في نصه. ونص الرخصة على: creativecommons.org/licenses/by-nc-nd/4.0/deed.ar

الموقع: alsamani.com

العلامات التجارية وأسماء المقاييس الواردة في الكتاب ملك لأصحابها، وتذكر لأغراض الوصف والتعليم وحدها. ولا يتضمن الكتاب بنود أي مقياس مقنن، والرسوم التي تمثل الاستثمارات تخطيطية.

الطبعة الأولى، ١٤٤٨ هـ / ٢٠٢٦ م



أساسيات القياس

ماذا تعني درجة ٧٨؟ وما الفرق بين الرتبة المئينة والنسبة المئوية؟ ولماذا يكتب الأخصائي بجوار الدرجة رقمين آخرين يحيطان بها؟ نقدم في هذا الفصل المفاتيح التي تُقرأ بها تقارير القياس على اختلافها، ونبين ما يدل عليه الرقم وما يقصر عن الدلالة عليه.

مشهد من اجتماع الفريق

تجلس أم ماجد في اجتماع المدرسة، وأمامها تقرير من ست صفحات. تقرأ فيه: «الدرجة الكلية ٧٢، الرتبة المئينة ٣، فترة الثقة ٦٥ إلى ٧٩». ترفع رأسها وتساءل المعلم: «هل يعني هذا أن ابني نجح بثلاثة في المئة؟ وهل لديه إعاقة أم لا؟»
يتردد المعلم. هو يعرف أن الرقم ليس نسبة نجاح، لكنه لا يعرف كيف يشرح ذلك، ولا لماذا كُتب في التقرير رقمان إضافيان حول الدرجة.

وسنعود إلى أم ماجد في ختام الفصل لنجيب عن سؤالها إجابة دقيقة. والمفاهيم التي نعرضها هنا تتردد في فصول الفئات جميعها، إذ هي اللغة المشتركة التي تُكتب بها تقارير التشخيص كافة.

أهداف الفصل

بعد دراسة هذا الفصل، يُتوقع أن يكون القارئ قادرًا على أن:

- ١ يصنّف نتيجة قياس معروضة بأنها معيارية المرجع أو محكية المرجع، ويعلل تصنيفه بالسؤال الذي تجيب عنه.
- ٢ يحوّل الدرجة بين أنواع الدرجات المعيارية الشائعة، ويحدد موقعها على المنحنى الطبيعي وتصنيفها الوصفي.
- ٣ يفسر الرتبة المئينية تفسيرًا صحيحًا، ويصحح عبارة تخلط بينها وبين النسبة المئوية أو العمر المكافئ.
- ٤ يحسب الخطأ المعياري للقياس وفترة الثقة والخطأ المعياري للفرق، ويستعملها في الحكم على درجة قريبة من درجة القطع أو على فرق بين درجتين.
- ٥ يحسب القيمتين التنبؤيتين الموجبة والسالبة من الحساسية والنوعية ونسبة الانتشار، ويفسر اختلافهما بين المسح العام والمجموعة المحالة.
- ٦ يقوم صلاحية أداة لطالب بعينه في ضوء مصادر أدلة الصدق الخمسة، والثبات، والمعايير، واللغة، وأدلة العدالة.
- ٧ يحدد المبدأ الأخلاقي أو المهني الذي يمسه موقف تقييمي معروض، ويقترح تصرفًا يتسق معه.

مقدمة

تصل نتائج القياس إلى المعلم والأسرة في صورة أرقام: درجة معيارية، ورتبة مئينية، وفترة ثقة، وتصنيف وصفي. والرتبة المئينية هي النسبة المئوية من أقران الطالب في عينة التقنين (المجموعة الممثلة التي استُخرجت منها معايير المقياس) الذين حصلوا على درجة أقل من درجته؛ فالرتبة ٨٤ تعني أن أداءه يفوق أداء ٨٤% منهم. ولكل رقم منها دلالة محددة وحدود لا يتجاوزها، والخطأ في قراءته يفضي إلى خطأ في القرار: كأن تُفهم الرتبة المئينية نسبةً للإجابات الصحيحة، كما فهمتها أم ماجد في المشهد الافتتاحي، أو تُعامل الدرجة رقمًا دقيقًا لا يحيط به خطأ، أو يُقطع بقرار حين تعبر فترة الثقة درجة القطع. فالحكم التشخيصي يعتمد في أكثر الفئات على موقع الطالب من أقرانه، وعلى دقة تقدير هذا الموقع، وعلى صلاحية الأداة لغرضها وللطالب الذي طبقت عليه.

نستهل الفصل بالتمييز بين المقاييس معيارية المرجع ومحكية المرجع (القسم ٢-١)، ثم نعرض المتوسط والانحراف المعياري (القسم ٢-٢)، والمنحنى الطبيعي والدرجات المعيارية وتصنيفاتها الوصفية (القسم ٢-٣)، والرتب المئينية ومزلقها (القسم ٢-٤). ونتناول بعد ذلك الثبات/الدقة والخطأ المعياري للقياس، وما يتصل بهما من فروق بين الدرجات وارتداد نحو المتوسط وأثر للأرضية والسقف (القسم ٢-٥)، ثم

الصدق والحساسية والنوعية والعدالة (القسم ٢-٦)، فالمعايير والتقنين وتكييف المقياس (القسم ٢-٧). ونطبق ذلك على قراءة جدول درجات حقيقي الشكل (القسم ٢-٨)، ونختم بالأخلاقيات والمسؤولية المهنية ومستويات التأهيل (القسم ٢-٩).

١-٢ المقاييس معيارية المرجع ومحكية المرجع

قبل أن نسأل «كم حصل الطالب؟» يجب أن نسأل: «مقارنة بماذا؟» فالدرجة وحدها لا معنى لها. ثمانون درجة قد تكون ممتازة في اختبار صعب، وضعيفة في اختبار سهل. ولهذا تنقسم المقاييس بحسب ما تُقارن به الدرجة إلى نوعين رئيسيين:

المقاييس محكية المرجع - Criterion-referenced

هل أتقن الطالب المهارة المطلوبة؟

تُقارن درجة الطالب بمستوى أداء محدد مسبقًا. والنتيجة تخبرنا بما يستطيع فعله وما لا يستطيع، بغض النظر عن أقرانه.

مثال: «قرأ ٤٠ كلمة في الدقيقة بدقة ٩٥%»، أو «يجمع عددين من منزلتين مع الحمل في ٨ من ١٠ مسائل».

المقاييس معيارية المرجع - Norm-referenced

أين يقع الطالب مقارنة بأقرانه؟

تُقارن درجة الطالب بدرجات مجموعة كبيرة من أقرانه في العمر أو الصف، تُسمى عينة التقنين. والنتيجة تخبرنا بموقعه بينهم: في المتوسط، أو أعلى منه، أو أدنى منه.

مثال: اختبارات الذكاء، ومقاييس السلوك التكيفي، ومقاييس التقدير مثل مقياس جيليام لتقدير التوحد، الإصدار الثالث (Gilliam Autism Rating Scale, GARS-3).

والنوعان يكمل أحدهما الآخر. فالتشخيص يحتاج غالبًا إلى المقاييس معيارية المرجع، لأن السؤال فيه «هل أداء الطالب بعيد عن أقرانه بدرجة غير معتادة؟». أما التخطيط للتدريس فيحتاج إلى المقاييس محكية المرجع، لأن السؤال فيه «ما المهارة التالية التي يجب أن نعلمها؟».

وقد تجمع الأداة الواحدة بين المنطقتين. ومثالها القياس المبني على المنهج، وهو مقاييس قصيرة متكررة تُتابع بها قراءة الطالب أو حسابه أسبوعيًا بعد أسبوع (ويعرفه الفصل الرابع ويفصل فيه). فهو مؤشر عام للنمو، تُفسر درجته معياريًا حين تُقارن بمعايير موسمية لأقران الصف، كالرتب المثينية في الخريف والشتاء والربيع، وتُفسر في الوقت نفسه بمقارنتها بهدف نمو محدد للطالب. فهو يجمع بين المرجعين معًا.

مثال يوضح الفرق

حصل سعد في الصف الرابع على الرتبة المئينية ٥٠ في اختبار مقنن للقراءة، أي أنه في منتصف أقرانه تمامًا. هذه معلومة معيارية المرجع، ولا تدل على وجود مشكلة. لكن معلمه لاحظ في اختبار محكي المرجع أنه لا يميز بين الحروف المتشابهة في الرسم إذا اختلفت نقاطها، مثل ب و ت و ث. وهذه معلومة لا يكشفها الاختبار الأول، وهي التي يحتاج إليها المعلم ليعرف ماذا يدرسه غداً.

٢-٢ المتوسط والانحراف المعياري

كل درجة معيارية مبنية على رقمين: **المتوسط (Mean)**، وهو مركز درجات عينة التقنين، و**الانحراف المعياري (Standard deviation, SD)**، وهو مقدار تشتت الدرجات حول ذلك المركز. فإذا كان الانحراف المعياري صغيراً كانت درجات الناس متقاربة، وإذا كان كبيراً كانت متباعدة.

ولتقريب فكرة الانحراف المعياري، تخيل صفين في كل منهما عشرون طالباً، ومتوسط الطول في الصفين ١٤٠ سنتيمتراً. في الصف الأول تتراوح الأطوال بين ١٣٥ و ١٤٥، وفي الثاني بين ١٢٠ و ١٦٠. المتوسط واحد، لكن التشتت مختلف. وطالب طوله ١٥٠ يُعد طويلاً جداً في الصف الأول، وعادياً نسبياً في الصف الثاني. الانحراف المعياري هو الذي يخبرنا بذلك.

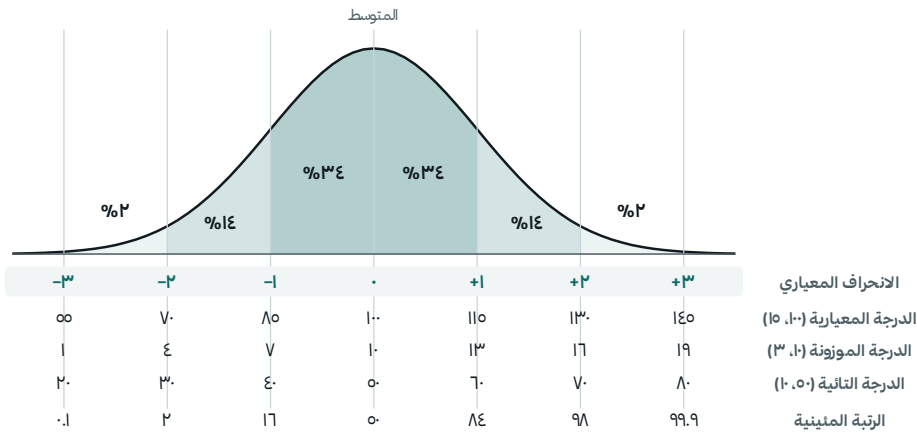
ومن هنا جاءت القاعدة التي يقوم عليها التشخيص: نقيس بُعد الطالب عن المتوسط بعدد الانحرافات المعيارية، بدل عدد النقاط. فالبعد بانحرافين معياريين عن المتوسط يعني الشيء نفسه في أي اختبار، مهما كانت طريقة حساب درجاته.

٣-٢ المنحنى الطبيعي والدرجات المعيارية

حين نقيس صفة مثل الذكاء أو الطول لدى عدد كبير من الناس، تتوزع الدرجات غالباً على شكل جرس: أغلب الناس قريبون من المتوسط، ويقل عددهم كلما ابتعدنا عنه في أي اتجاه. هذا الشكل هو **المنحنى الطبيعي (Normal curve)**، ومن خصائصه نسب ثابتة لا تتغير:

- نحو ٦٨% من الناس يقعون بين انحراف معياري واحد تحت المتوسط وانحراف معياري واحد فوقه.
- نحو ٩٥% يقعون بين انحرافين معياريين تحت المتوسط وانحرافين فوقه.
- نحو ٢% فقط يقعون دون المتوسط بأكثر من انحرافين معياريين، ومثلهم تقريباً فوقه بأكثر من انحرافين معياريين.

والنقطة الأخيرة هي أصل الحدود التي تتكرر في هذا الكتاب. فحد «انحرافين معياريين دون المتوسط» يُستخدم في الإعاقة الفكرية، وحد «انحرافين معياريين فوق المتوسط» يظهر في كثير من تعريفات الموهبة. أي أن الفئتين تمثلان، بحسب هذا المعيار الإحصائي وحده، نحو ٢% من الناس في كل طرف.



شكل ١-٢: المنحنى الطبيعي. وتحت أنواع الدرجات الشائعة مصفوفة بحسب موقعها. كل عمود يمثل الموقع نفسه بلغات مختلفة: ٧٠ في الدرجة الكلية لاختبار ذكاء تعادل ٤ في اختبار فرعي، وتعادل ٣٠ في مقياس بالدرجة الثانية، وكلها تعني انحرافين معياريين دون المتوسط.

أنواع الدرجات المعيارية

تحوّل الاختبارات الدرجة الخام (**Raw score**). أي عدد الإجابات الصحيحة، إلى درجة معيارية (**Standard score**) لها متوسط وانحراف معياري ثابتان. والفائدة أن الدرجات المعيارية قابلة للمقارنة بين الأعمار والاختبارات. وأشهر أنواعها:

جدول ١-٢: أنواع الدرجات المعيارية الشائعة.

النوع	المتوسط	الانحراف المعياري	أين تصادفه؟
الدرجة الزائفة (z)	٠	١	في الحسابات والبحوث. وهي الأساس الذي تُشتق منه بقية الأنواع

النوع	المتوسط	الانحراف المعياري	أين تصادفه؟
الدرجة المعيارية	١٠٠	١٥	الدرجات الكلية والمؤشرات في اختبارات الذكاء والتحصيل ومقاييس السلوك التكيفي
الدرجة الموزونة	١٠	٣	الاختبارات الفرعية في مقاييس الذكاء، والمقاييس الفرعية في GARS-3
الدرجة التائية (T)	٥٠	١٠	مقاييس التقدير السلوكية والانفعالية، ومنها كثير من مقاييس اضطراب ضعف الانتباه وفرط الحركة (Attention-Deficit/Hyperactivity Disorder, ADHD)

ولتحويل أي درجة إلى أخرى، نحولها أولاً إلى درجة زائفة، ثم منها إلى النوع المطلوب:

$$z = (X - M) \div SD$$

الدرجة الزائفة = (درجة الطالب - المتوسط) ÷ الانحراف المعياري

مثال محلول

- ١ حصلت هند على درجة معيارية ٨٢ في اختبار ذكاء متوسطه ١٠٠ وانحرافه ١٥.
- ٢ الدرجة الزائفة: $100 - 82 = 18$ ، أي أنها دون المتوسط بانحراف معياري وخمس.
- ٣ لو عُبر عنها بالدرجة التائية لكانت: $50 + (18 \times 1.2) = 71.6$.
- ٤ ورتبتها المئينية نحو ٨٢، أي أن أداؤها أعلى من نحو ٨٢% من أقرانها في عينة التقنين.

⚠ انتبه إلى اتجاه الدرجة

في اختبارات القدرة، الدرجة الأعلى تعني أداءً أفضل. أما في كثير من مقاييس التقدير السلوكية، فالدرجة الأعلى تعني مشكلات أكثر. فدرجة تائية ٧٠ في اختبار قراءة ممتازة، وفي مقياس لفرط الحركة مقلقة. لذلك اقرأ دائماً في دليل المقياس أو في التقرير ماذا تعني الدرجة المرتفعة.

التصنيفات الوصفية

ترجم الاختبارات الدرجة المعيارية إلى وصف لفظي سهل فهمه. وهذا مثال من تصنيفات مقياس وكسلر لذكاء الأطفال، الإصدار الخامس (**Wechsler Intelligence Scale for Children, Fifth Edition, WISC-V**). مع التسمية الإنجليزية الرسمية لكل فئة. لأن الأخصائي يقرأ كثيرًا تقارير مكتوبة بها (Wechsler, 2014):

جدول ٢-٢: التصنيفات الوصفية للدرجات المركبة في مقياس وكسلر لذكاء الأطفال، الإصدار الخامس.

مدى الدرجة المعيارية	التصنيف الوصفي	النسبة التقريبية من الناس
١٣٠ فأكثر	مرتفع للغاية (Extremely High)	٢%
١٢٠-١٢٩	مرتفع جدًا (Very High)	٧%
١١٠-١١٩	متوسط مرتفع (High Average)	١٦%
٩٠-١٠٩	متوسط (Average)	٥٠%
٨٠-٨٩	متوسط منخفض (Low Average)	١٦%
٧٠-٧٩	منخفض جدًا (Very Low)	٧%
٦٩ فأقل	منخفض للغاية (Extremely Low)	٢%

ولهذا الجدول ثلاثة تنبيهات. الأول أن التسميات تختلف بين الاختبارات، وحتى بين إصدارات الاختبار الواحد. فالإصدار الرابع من المقياس نفسه كان يسمى المدى ٧٠-٧٩ «حدّيًا» (**Borderline**). ويسمى ما فوق ١٣٠ «متفوقًا جدًا» (**Very Superior**). فإذا قرأت في تقرير أجنبي عبارة **Very High** عن درجة ١٢٥، فاعلم أن فوقها فئة أعلى. والثاني أن فئة «المتوسط» هنا (٩٠-١٠٩) تصنيف وصفي خاص بمقاييس وكسلر، وهي أضيق من المدى ٨٥-١١٥ الذي يمثل انحرافًا معياريًا واحدًا حول المتوسط، وهو المدى المعتمد في رسوم هذا الكتاب. والثالث، وهو أهمها، أن التصنيف الوصفي ليس تشخيصًا. فعبارة «منخفض للغاية» تصف موقع الدرجة وحده، والحكم بالإعاقة الفكرية يحتاج إلى محكات أخرى كالسلوك التكيفي، أي مهارات الحياة اليومية التي يمارسها الطالب فعليًا في البيت والمدرسة (ويتناول الفصل السابع ذلك بالتفصيل).

٤-٢ الرتب المئينية ومزلقها

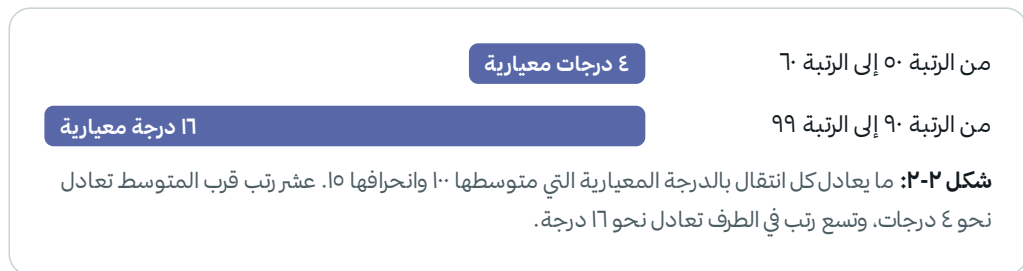
الرتبة المئينية (Percentile rank) أسهل الدرجات فهمًا في ظاهرها لدى الأسر، وأكثرها عرضة لسوء الفهم أيضًا. وهي تعني نسبة الأقران الذين حصلوا على درجة أقل من درجة الطالب في عينة التقنين. فالرتبة المئينية ٣٠ تعني أن الطالب تفوق على نحو ٣٠% من أقرانه، وأن نحو ٧٠% منهم تفوقوا عليه.

المزلق الأول: الخلط بينها وبين النسبة المئوية

الرتبة المئينية لا علاقة لها بنسبة الإجابات الصحيحة. فطالب أجاب عن ٨٠% من الأسئلة قد تكون رتبته المئينية ٢٠ إذا كان الاختبار سهلًا وأجاب أغلب أقرانه عن أكثر من ذلك. وهذا ما وقعت فيه أم ماجد حين ظنت أن ابنها «نجح بثلاثة في المئة».

المزلق الثاني: المسافات بين الرتب غير متساوية

لأن أغلب الناس يتجمعون حول المتوسط، فإن الانتقال بضع رتب مئينية قرب المتوسط يحتاج إلى فرق صغير في الأداء، بينما يحتاج الانتقال بالعدد نفسه من الرتب في الأطراف إلى فرق كبير. قارن بين هذين الانتقالين، وكلاهما بمقدار رتب متقاربة:



ولهذا لا يُطرح بعض الرتب المئينية من بعض، ولا يُحسب متوسطها، ولا يُقاس بها تقدم الطالب. فإذا انتقل طالب من الرتبة ٥ إلى الرتبة ١٠، فقد يكون تقدمه أكبر بكثير مما يوحي به الرقمان.

المزلق الثالث: العمر المكافئ والصف المكافئ

العمر المكافئ (Age equivalent) والصف المكافئ (Grade equivalent) درجتان تعبّران عن الأداء بعمر أو صف دراسي. فقد يقرأ المعلم في تقرير أن «عمر الطالب القرائي المكافئ ٧ سنوات و٣ أشهر» وعمره الفعلي عشر سنوات، فيظن أنه يقرأ مثل طفل في السابعة. وهذا فهم غير دقيق. فالعمر المكافئ يعني فقط

أن درجته الخام تساوي متوسط درجات الأطفال في ذلك العمر، ولا يعني أن مهاراته أو أخطائه تشبه مهاراتهم وأخطاءهم. كما أن هذه الدرجات لا تتوزع بانتظام، ولا يتبين منها هل الفرق كبير قياسًا إلى أقرانه أم لا. ولهذا توصي معايير الاختبار التربوي والنفسي بالحدز الشديد في استخدامها (AERA et al., 2014)، ويُفضل الاعتماد على الدرجة المعيارية والرتبة المئينية.

٥-٢ الثبات/الدقة والخطأ المعياري للقياس

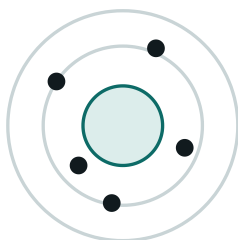
الثبات (Reliability) هو اتساق نتائج المقياس. فلو طُبق الاختبار على الطالب نفسه مرتين في ظروف متشابهة، هل يحصل على درجتين متقاربتين؟ والميزان الذي يعطي وزنًا مختلفًا في كل مرة لا فائدة منه مهما كان دقيقًا في المتوسط. وتستعمل معايير الاختبار التربوي والنفسي المصطلح المركب «الثبات/الدقة» (**Reliability/precision**)، لتشمل معاملات الثبات التقليدية وكل مؤشر آخر لدقة الدرجة. كالخطأ المعياري للقياس ودالة المعلومات التي سنعرضها في آخر هذا القسم (AERA et al., 2014).

أنواع الثبات

جدول ٢-٣: أنواع الثبات والسؤال الذي يجيب عنه كل منها.

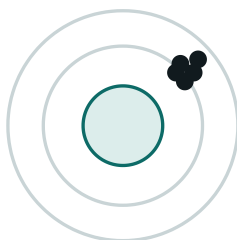
النوع	السؤال الذي يجيب عنه	كيف يُحسب
ثبات الإعادة	هل تبقى الدرجة مستقرة عبر الزمن؟	يُطبق الاختبار مرتين بفواصل زمني قصير، ويُحسب الارتباط بين التطبيقين
الاتساق الداخلي	هل تقيس بنود الاختبار الشيء نفسه؟	يُحسب من ارتباطات البنود في تطبيق واحد، ومن أشهره معامل ألفا كرونباخ (Cronbach's alpha)
الثبات بين المقدرين	هل يعطي مقدّران مختلفان الدرجة نفسها؟	يقدر شخصان الطالب نفسه، ويُقارن تقديرهما. وهو مهم جدًا في مقاييس الملاحظة والتقدير
ثبات الصور المتكافئة	هل تعطي صورتان من الاختبار نتائج متقاربة؟	تُطبق صورتان متكافئتان، ويُحسب الارتباط بينهما

ويُعبّر عن الثبات بمعامل تتراوح قيمته بين الصفر والواحد. وقد اقترح سالفيا وزملاؤه (Salvia et al., 2017) مستويات مرجعية شائعة الاستخدام: ٠.٦٠ على الأقل حين تُستخدم الدرجات لوصف مجموعات، و٠.٨٠ على الأقل لقرارات المسح، و٠.٩٠ على الأقل للقرارات المهمة المتعلقة بفرد، كالتشخيص والتصنيف.



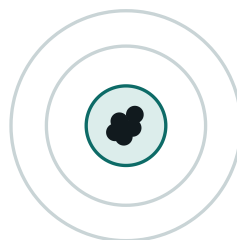
غير ثابت

النتائج متناثرة، فلا يمكن أن يكون صادقاً



ثابت غير صادق

النتائج متقاربة، لكنها تقيس شيئاً آخر



ثابت وصادق

النتائج متقاربة، وتصيب ما يُراد قياسه

شكل ٣-٢: العلاقة بين الثبات والصدق بتشبيه الرماية. الثبات شرط للصدق، لكنه لا يكفي وحده.

الخطأ المعياري للقياس

لا يوجد مقياس ثابت ثباتاً تاماً. فدرجة الطالب في أي يوم تتأثر بتعبه، وانتباهه، وتخمينه، وظروف الغرفة. ولذلك تُعد الدرجة التي نحصل عليها **تقديرًا** للدرجة الحقيقية (**True score**). أي الدرجة التي كان سيحصل عليها لو قيس دون أي خطأ. والخطأ المعياري للقياس (**Standard Error of Measurement, SEM**) رقم يخبرنا بمقدار التذبذب المتوقع في درجة الطالب لو أعدنا قياسه مرات كثيرة:

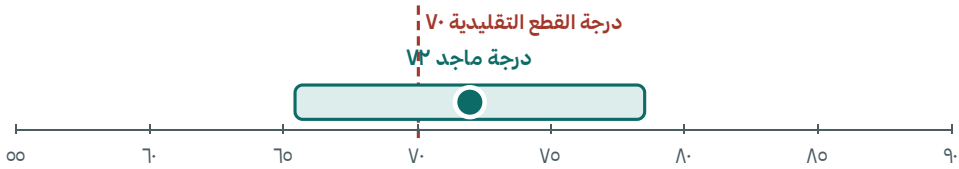
$$SEM = SD \times \sqrt{1 - r}$$

الخطأ المعياري للقياس = الانحراف المعياري × الجذر التربيعي لـ (١ - معامل الثبات)

ومنه نبني فترة الثقة (**Confidence interval**): مدى من الدرجات نثق بنسبة معينة بأن الدرجة الحقيقية تقع فيه. ولفترة ثقة ٩٥% نضرب الخطأ المعياري في ١.٩٦ تقريباً، ثم نضيف الناتج إلى الدرجة ونطرحه منها.

مثال محلول: درجة ماجد

- ١ حصل ماجد على درجة كلية ٧٢ في اختبار ذكاء انحرافه المعياري ١٥ ومعامل ثباته ٠.٩٥.
- ٢ الخطأ المعياري للقياس: $٣.٤ \approx ٠.٢٢٤ \times ١٥ = (٠.٩٥ - ١) \sqrt{١٥}$
- ٣ فترة الثقة ٩٥%: $٧٢ \pm (٣.٤ \times ١.٩٦) = ٧٢ \pm ٦.٦$ أي من نحو ٦٥ إلى نحو ٧٩.
- ٤ النتيجة: نثق بنسبة ٩٥% بأن درجة ماجد الحقيقية في هذا الاختبار تقع بين ٦٥ و٧٩. وهذا المدى يعبر درجة القطع ٧٠.



شكل ٢-٤: درجة ماجد وفترة الثقة حولها. النقطة فوق درجة القطع، لكن الشريط يمتد على جانبيها، فلا يصح أن نقول إنه «تجاوزها» أو «لم يبلغها» بناءً على الرقم وحده.

و**درجة القطع (Cut-off score)** هي الدرجة التي تفصل في قرار ما بين فئتين، كأن تفصل بين «يُحال» و«لا يُحال» في أداة مسح، أو بين «يستوفي المحك» و«لا يستوفيه» في محك تشخيصي. وسنستعمل هذا المصطلح في الكتاب كله.

ولهذا السبب بالذات تنص الجمعية الأمريكية للإعاقات الفكرية والنمائية (**American Association on Intellectual and Developmental Disabilities, AAIDD**) على مراعاة الخطأ المعياري للقياس حين تحدد القصور الدال في القدرة العقلية، فلا يُعامل الرقم ٧٠ على أنه فاصل قاطع (Schalock et al., 2021). فدرجة مثل ٧٢ أو ٧٤ لا تستبعد الإعاقة الفكرية تلقائيًا، ودرجة مثل ٦٨ لا تثبتها تلقائيًا. والذي يحسم هو اجتماع الأدلة، ولا سيما أدلة السلوك التكيفي.

♀ القاعدة التي تحملها معك

كل درجة في تقرير تشخيصي هي نقطة داخل شريط. وكلما كان المقياس أقل ثباتًا، اتسع الشريط. وحين يقع الشريط على جانبي درجة القطع، يصبح القرار قرارًا مهنيًا يستند إلى أدلة أخرى إلى جانب الرقم.

للمتخصص: لماذا لا تكون فترات الثقة في التقارير متماثلة حول الدرجة؟

الطريقة التي حسبنا بها فترة ماجد طريقة مبسطة، تبني الفترة حول الدرجة الملاحظة نفسها. أما كثير من الاختبارات الحديثة، ومنها مقياس وكسلر لذكاء الأطفال، فتبني الفترة حول **الدرجة الحقيقية المقدّرة (Estimated true score)**، وتستعمل في ذلك **الخطأ المعياري للتقدير (Standard error of estimation, SEE)** (Sattler, 2018):

$$T' = M + r(X - M)$$

$$SEE = SD \times \sqrt{r(1 - r)}$$

الدرجة الحقيقية المقدّرة = المتوسط + الثبات × (الدرجة الملاحظة - المتوسط)؛ والخطأ المعياري للتقدير = الانحراف المعياري × الجذر التربيعي لـ [(الثبات × (١ - الثبات))]

ولأن الثبات أقل من الواحد، تقع الدرجة الحقيقية المقدّرة دائماً أقرب إلى المتوسط من الدرجة الملاحظة. فلو طبقنا ذلك على ماجد لكانت درجته الحقيقية المقدّرة $100 + 0.95 \times (72 - 100) \approx 73.4$ ، والخطأ المعياري للتقدير نحو ٣.٣، فتمتد الفترة من نحو ٦٧ إلى نحو ٨٠. الفترة إذن ماثلة نحو المتوسط، ولا تتماثل حول ٧٢، لكنها تعبر درجة القطع ٧٠ في الطريقتين. ولهذا يرى القارئ في جداول التقارير فترات مثل «٨٤: ٧٨-٩٢»، فيظنها خطأً مطبعياً وهي ليست كذلك (انظر الجدول ٢-٨). وحين تقرأ تقريراً، ابحث في دليل الاختبار عن الطريقة التي بُنيت بها فتراته قبل أن تقارنها بفترات اختبار آخر.

📍 في السياق السعودي: درجات القطع في التنظيمات

ينص الدليل التنظيمي للتربية الخاصة على أن من شروط أهلية الطالب لبرامج الإعاقة الفكرية «ألا تزيد درجة ذكاء الطالب عن ٧٥ درجة حسب اختبار وكسلر، أو ٦٩ على اختبار ستانفورد بينيه (الصورة الخامسة)، أو ما يعادل أيًا منهما»، مع قصور في مجالين على الأقل من مجالات المهارات التكيفية، ويشترط «عدم الاكتفاء بنتائج مقياس واحد» وأن تُطبق الأدوات وتُفسّر «من قبل متخصصين مؤهلين» (وزارة التعليم، ١٤٣٦هـ). وكانت القواعد التنظيمية السابقة تحدد مدى «القابل للتعليم» بدرجة ذكاء بين ٥٥ و ٧٥ على وكسلر، أو بين ٥٢ و ٧٣ على ستانفورد بينيه (وزارة المعارف، ١٤٢٢هـ). ويُفهم اختلاف الرقمين في القواعد من اختلاف الانحراف المعياري للإصدارات القديمة من ستانفورد بينيه (١٦ بدل ١٥)، فكلاهما يقع على البعد نفسه تقريبًا من المتوسط. أما في الدليل الحالي، فللمقياسين في إصداريهما الحاليين انحراف معياري واحد (١٥)، ولم نجد في النص تعليقًا لاختلاف الرقمين (٧٥ و ٦٩). ولم نجد فيه كذلك نصًا يوجه إلى مراعاة الخطأ المعياري للقياس أو فترة الثقة عند تطبيق هذه الأرقام.

سؤال للتأمل: طالبة درجتها ٧٧ على مقياس وكسلر، وفترة الثقة حولها ٧٢-٨٣، ولديها قصور واضح في ثلاثة مجالات تكيفية. كيف يتعامل الفريق مع درجة القطع ٧٥ في ضوء ما تعلمته في هذا القسم؟

٢-٥-١ الفروق بين الدرجات

كثير من قرارات التشخيص تُبنى على فرق بين درجتين: بين القدرة والتحصيل في نموذج التباعد، أو بين مؤشرين في اختبار ذكاء، أو بين درجة الطالب قبل التدخل وبعده. وكل درجة من الدرجتين تحمل خطأها، فيجتمع الخطآن في الفرق. ولذلك يكون الفرق أقل دقة من كل درجة منهما، ويحتاج إلى أن يكون كبيرًا نسبيًا قبل أن نطمئن إلى أنه فرق حقيقي.

وحين نجد فرقًا بين درجتين نسأل سؤالين مختلفين. الأول سؤال **الدلالة الإحصائية**: هل الفرق أكبر من أن يُعزى إلى خطأ القياس وحده؟ والثاني سؤال **الندرة** أو **معدل الأساس (Base rate)**: ما نسبة من لديهم فرق بهذا الحجم أو أكبر في عينة التقنين نفسها؟ فقد يكون الفرق دالًا، أي حقيقيًا غير ناتج عن الخطأ، ومع ذلك شائعًا بين الأطفال العاديين، فلا يصح أن يُعد علامة على اضطراب. وأدلة الاختبارات الجيدة تعرض الأمرين في جدولين منفصلين.

١٢ للمتخصص: الخطأ المعياري للفرق

إذا كان للدرجتين الانحراف المعياري نفسه، يُحسب الخطأ المعياري للفرق (Standard error of the difference, SE_{diff}) هكذا:

$$SE_{diff} = SD \times \sqrt{(2 - r_1 - r_2)}$$

الخطأ المعياري للفرق = الانحراف المعياري × الجذر التربيعي لـ (٢ - ثبات الدرجة الأولى - ثبات الدرجة الثانية)

مثال محلول: فرق بين مؤشرين (من الجدول ٨-٢)

- ١ مؤشر الفهم اللفظي ٨٤ وثباته ٠.٩٤ ومؤشر البصري المكاني ١٢ وثباته ٠.٩٢. الفرق ١٨ درجة.
- ٢ الخطأ المعياري للفرق: $(0.92 - 0.94 - 2) \times 10 = -0.02 \times 10 = -0.2$ ، $0.6 \approx 0.37 \times 10$
- ٣ الفرق الحرج عند مستوى دلالة ٠.٠٥: $1.96 \times 0.6 \approx 1.18$. والفرق الملاحظ (١٨) أكبر منه، فهو دال إحصائياً.
- ٤ ثم نسأل عن الندرة. لنفترض، للتوضيح فقط، أن جدول معدل الأساس في الدليل يبين أن نحو ١٥% من أطفال عينة التقنين لديهم فرق بهذا الحجم أو أكبر بين هذين المؤشرين. فالفرق حقيقي، لكنه يظهر لدى طفل من كل سبعة تقريباً، فهو غير نادر، ولا يصلح وحده دليلاً على اضطراب.
- ٥ قارن ذلك بالفرق بين الذاكرة العاملة (٧٦، وثباتها ٠.٩٣) وسرعة المعالجة (٨٠، وثباتها ٠.٩٠): الخطأ المعياري للفرق $0.17 \times 10 = 1.7$ ، والفرق الحرج نحو ١.٢، والفرق الملاحظ ٤ فقط. فلا يصح أن نقول إن سرعة المعالجة «أقوى» من الذاكرة العاملة.

ولثبات درجة الفرق نفسها صيغة توضح المشكلة من زاوية أخرى: $r_{diff} = [(r_1 + r_2) \div 2 - r_{12}] \div (1 - r_{12})$ ، حيث r_{12} الارتباط بين الدرجتين. فلو كان ثبات المؤشرين ٠.٩٤ و ٠.٩٢، والارتباط بينهما ٠.٦٠، لكان ثبات الفرق $0.40 \approx (0.93 - 0.60) \div (1 - 0.60)$ ، وهو أدنى من ثبات كل منهما. وكلما ارتفع الارتباط بين الدرجتين انخفض ثبات الفرق بينهما. وهذا أصل النقد الموجه إلى القرارات المبنية على الفروق وحدها. كما في نموذج التباعد بين القدرة والتحصيل ونماذج أنماط القوة والضعف (ويعود إليها الفصل السادس).

٢-٥-٢ الارتداد نحو المتوسط

حين يُعاد قياس من حصلوا على درجات متطرفة، يميل متوسط درجاتهم في المرة الثانية إلى الاقتراب من متوسط المجتمع. وهذه ظاهرة إحصائية تُسمى **الارتداد نحو المتوسط (Regression toward the mean)**، وسببها أن الدرجة المتطرفة تجمع غالباً بين قدرة مرتفعة (أو منخفضة) فعلاً وخطأ قياس صادف أن جاء في الاتجاه نفسه، والخطأ لا يتكرر بالضرورة في المرة الثانية. والدرجة الحقيقية المقدّرة التي رأيناها في مربع «للمتخصص» السابق تطبيق مباشر لهذه الفكرة.

ولهذه الظاهرة أثر عملي في موضعين. الأول في الموهبة: لو اخترنا الطلاب الذين تجاوزوا ١٣٠ في اختبار ثباته ٩٠٪، فإن الدرجة المتوقعة لطالب حصل على ١٣٥ في إعادة الاختبار نفسه نحو $100 + 0.90 \times 35 \approx 132$ ، وتنخفض أكثر إن كان الاختبار الثاني مختلفاً لا يرتبط بالأول إلا بنحو ٨٠٪ (نحو ١٢٨). فانخفاض درجات بعض من اختيروا عند إعادة التقييم متوقع، ولا يعني أن القرار الأول كان خطأ ولا أن الطالب «فقد» موهبته (Lohman & Korb, 2006). والثاني في الإعاقة وصعوبات التعلم: من اختيروا لدرجاتهم المنخفضة جداً سترتفع درجات بعضهم عند إعادة القياس ولو لم يتلقوا أي تدخل، فلا يصح أن يُنسب هذا الارتفاع كله إلى البرنامج دون مجموعة مقارنة.

٣-٥-٢ الأرضية والسقف وتدرج البنود

أرضية الاختبار (Floor) هي قدرته على التمييز بين المستويات المنخفضة جداً من الأداء، و**سقفه (Ceiling)** قدرته على التمييز بين المستويات المرتفعة جداً. فإذا كانت بنود الاختبار السهلة قليلة، حصل الطفل ذو الأداء المنخفض جداً والطفل ذو الأداء المنخفض على الدرجة الدنيا نفسها، ولم يعد الاختبار يفرق بينهما. وكذلك إذا كانت البنود الصعبة قليلة، تجمع المتفوقون عند أعلى درجة ممكنة. ويقف أدنى مدى الدرجة الكلية في كثير من اختبارات الذكاء الفردية عند نحو ٤٠، وأعلاه عند نحو ١٦٠. وهذا يكفي أغلب الأطفال، لكنه لا يكفي في موضعين: التمييز بين مستويات الشدة في الحالات الشديدة من الإعاقة الفكرية، وهو ما يعود إليه الفصل السابع، والتمييز بين الموهوبين بدرجة استثنائية، وهو ما يعود إليه الفصل الثاني عشر. ولهذا أصدر بعض الناشرين **معايير ممتدة (Extended norms)** لطرفي التوزيع.

ويرتبط بذلك **تدرج البنود (Item gradient)**، أي مقدار تغير الدرجة المعيارية حين تزيد الدرجة الخام نقطة واحدة. فإذا كانت إجابة صحيحة واحدة إضافية تنقل الطفل من درجة موزونة ٤ إلى ٧، فالاختبار خشن في هذا المدى، وخطأ صغير في التصحيح أو لحظة سهو تغير التصنيف كله. وقد اقترح براكن حدوداً دنيا للكفاية الفنية شاعت في الأدبيات بعده، منها أن تنزل الدرجة الخام «١» في الاختبار الفرعي بالطفل إلى ما دون المتوسط بأكثر من انحرافين معياريين حتى تكون الأرضية كافية، وألا تُحدث زيادة الدرجة الخام نقطة

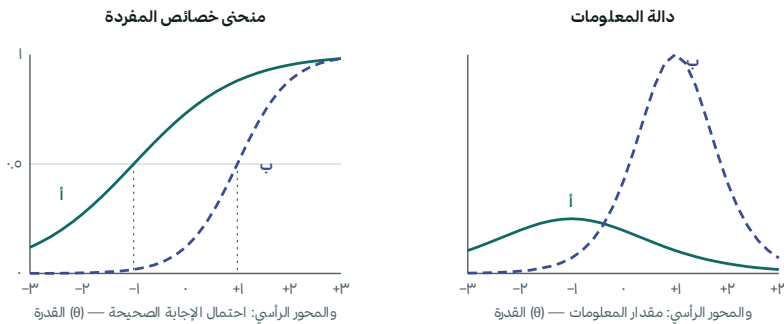
واحدة تغييرًا يتجاوز ثلث انحراف معياري (Bracken, 1987). وتظهر هذه المشكلة أكثر ما تظهر عند صغار السن، وعند الأطفال في أدنى الأعمار التي يغطيها الاختبار أو أعلاها. فإذا كان الطفل في طرف المدى العمري للاختبار، فافحص جداول المعايير قبل أن تثق بالدرجة.

٢-٥-٤ نظرية الاستجابة للمفردة

كل ما عرضناه حتى الآن ينتمي إلى النظرية الكلاسيكية للقياس، وهي تفترض خطأً معياريًا واحدًا لكل درجات الاختبار. لكن أغلب الاختبارات الحديثة بُنيت، كلها أو بعضها، بإطار آخر هو **نظرية الاستجابة للمفردة (Item Response Theory, IRT)**. ويكفي المعلم أن يعرف منها فكرتين: أن دقة الاختبار تتفاوت بتفاوت مستويات القدرة، فتكون أعلى حيث تتركز بنوده؛ وأن هذا هو ما يسمح بالاختبارات التكيفية المحوسبة، التي يختار فيها الحاسوب لكل طالب البنود المناسبة لمستواه.

١٤ للمتخصص: منحنى خصائص المفردة ودالة المعلومات

تصف النظرية كل بند بمنحنى خصائص المفردة (Item characteristic curve)، وهو يربط قدرة الطالب، ويُرمز لها بـ θ وتُعرض على مقياس يشبه الدرجة الزائفة، باحتمال أن يجيب عن البند إجابة صحيحة. وأهم معالم المنحنى معلمتان: الصعوبة (b)، وهي مستوى القدرة الذي يصبح عنده احتمال الإجابة الصحيحة ٥٠%، والتمييز (a)، وهو انحدار المنحنى عند تلك النقطة. وقد تُضاف معلمة ثالثة للتخمين في أسئلة الاختيار من متعدد (Embretson & Reise, 2000).



شكل ٢-٥: مفردتان افتراضيتان. المفردة «أ» (الخط المتصل) سهلة ومعتدلة التمييز، والمفردة «ب» (الخط المنقطع) صعبة وعالية التمييز. في الرسم الأيسر يبلغ احتمال الإجابة الصحيحة ٥٠% عند $\theta = 1$ للمفردة «أ» وعند $\theta = 0$ للمفردة «ب». وفي الرسم الأيمن تعطي كل مفردة أكبر قدر من المعلومات قرب صعوبتها، والمفردة الأعلى تمييزاً أغزر معلومات وأضيق نطاقاً.

ومن المنحنى تُشتق دالة المعلومات (Information function)، وهي تبين مقدار ما يسهم به البند في دقة تقدير القدرة عند كل مستوى. ويُجمع ما تعطيه البنود كلها في دالة معلومات الاختبار، ويساوي الخطأ المعياري لتقدير القدرة عند أي مستوى مقلوب الجذر التربيعي للمعلومات عنده. فالاختبار الذي تتركز بنوده حول المتوسط يقيس الطلاب المتوسطين بدقة عالية، وقيس الطرفين بدقة أقل، وهذه صياغة أدق لمشكلة الأرضية والسقف السابقة. أما الاختبار التكيفي المحوسب (Computerized adaptive testing, CAT)، فيبدأ ببند متوسط الصعوبة، ثم يقدر قدرة الطالب بعد كل إجابة، ويختار البند التالي من بنك البنود بحيث يعطي أكبر قدر من المعلومات عند ذلك

التقدير. فيصل إلى دقة مقبولة بنود أقل، ويواجه كل طالب بنودًا قريبة من مستواه. لكن صدق هذه الطريقة مرهون بمعايرة بنك البنود على عينة تشبه الطلاب الذين نقيسهم، وهو شرط يعود بنا إلى القسم ٧-٢.

٦-٢ الصدق

إذا كان الثبات يسأل «هل يقيس المقياس باتساق؟»، فالصدق (**Validity**) يسأل «هل يقيس ما يُراد قياسه فعليًا؟ وهل تصح القرارات التي نبنيناها على درجاته؟». والصدق في الفهم الحديث صفة لتفسير درجات المقياس في غرض معين، تتغير بتغير الغرض والمفحوص. فمقياس صادق في مسح أطفال في الثالثة قد لا يكون صادقًا في تشخيص مراهق. وتعد معايير الاختبار التربوي والنفسي الصدق مفهومًا واحدًا جامعا، يُسمى أحيانًا **الصدق البنائي**، أي مدى دعم الأدلة لتفسير الدرجة بأنها تمثل البناء المقصود. فالبناء هو ما تشهد له الأدلة كلها مجتمعة. وتجمع المعايير هذه الأدلة من خمسة مصادر (AERA et al., 2014):

جدول ٤-٢: المصادر الخمسة لأدلة الصدق في معايير الاختبار التربوي والنفسي، مع أمثلة من تكييف المقاييس للبيئة العربية.

مصدر الأدلة	السؤال	مثال
المحتوى	هل تغطي البنود جوانب البناء تغطية مناسبة، وهل تناسب البيئة التي تُطبق فيها؟	مقياس لاضطراب طيف التوحد تغطي بنوده جوانبه الأساسية في الدليل التشخيصي: صعوبات التواصل والتفاعل الاجتماعي، والسلوك النمطي والاهتمامات المحدودة. وفي النسخة العربية يراجع خبراء محليون كل بند: هل يصف سلوكًا يُتوقع ظهوره في البيت أو المدرسة السعودية؟
عمليات الاستجابة	هل يستعمل المفحوص أو المقدر العمليات التي يفترضها البناء حين يجيب؟	يُطلب من أطفال سعوديين أن يفكروا بصوت مسموع وهم يحلون بنود اختبار استدلال. أو يُسأل آباء عن فهمهم لعبارة في مقياس تقدير. فإذا تبين أن الطفل يخفق لأنه لا يعرف كلمة فصيحة في السؤال، فالبند يقيس المفردات لا الاستدلال

مصدر الأدلة	السؤال	مثال
البنية الداخلية	هل تترابط البنود والأبعاد على النحو الذي تفترضه النظرية؟	التحليل العاملي للنسخة العربية: هل تحتفظ ببنية الأصل؟ فقد انتهى تحليل عاملي للنسخة العربية من مقياس تقدير التوحد الطفولي إلى نموذج من ثلاثة عوامل يضم ١٤ بندًا (Alotaibi & Alotaibi, 2021). ودعمت دراسة عُمانية لنظام تقييم السلوك التكيفي (ABAS-3) نموذج العامل العام الواحد لدى أطفال ذوي إعاقة فكرية (Emam et al., 2020)
العلاقات بمتغيرات أخرى	هل ترتبط الدرجات بما يُتوقع أن ترتبط به، وتتنبأ بالمحك، وتميز بين المجموعات؟	ارتبطت الصورة السعودية من مقياس وكسلر لذكاء أطفال ما قبل المدرسة والمرحلة الابتدائية، الإصدار الثالث، بالصورة السعودية من مقياس ستانفورد بينيه بمعامل ٠.٨١ (Ejbara, 2024). وفي تحقق من النسخة العربية لقائمة التوحد المعدلة للأطفال الصغار في تسع دول عربية بلغت الحساسية ٠.٨٦ والنوعية ٠.٨٠ (Seif Eldin et al., 2008)
العواقب	ما آثار استعمال الدرجات في القرارات مقصودة كانت أو غير مقصودة؟	هل يؤدي استعمال مقياس في الترشيح لبرامج الموهوبين إلى تمثيل أقل من المتوقع لطلاب القرى أو لفئة بعينها؟ (ويعود إلى ذلك الفصل الرابع عشر)

والتقسيم القديم إلى «صدق المحتوى» و«صدق المحك» و«الصدق البنائي» ما زال يرد في كثير من أدلة المقاييس والكتب العربية. ويقابل «صدق المحك» فيه جزءًا من المصدر الرابع، ويقابل «الصدق البنائي» الصدق كله في الفهم الحديث. أما المصدران الثاني والثالث فهما أهم ما نسأل عنه حين نقوم نسخة معربة: هل يفهم الطفل العربي البند كما يُراد له؟ وهل احتفظت النسخة ببنية الأصل؟

الحساسية والنوعية، ولماذا تتغير قيمة النتيجة الموجبة

في أدوات المسح والتشخيص، يُعرض الصدق غالبًا، ضمن العلاقات بمتغيرات أخرى، بمفهومين: الحساسية (Sensitivity)، وهي نسبة من تكشفهم الأداة بحق ممن لديهم الحالة، والنوعية (Specificity)، وهي نسبة من تستبعدهم الأداة بحق ممن ليس لديهم الحالة. لكن هناك رقمين يهتمان المعلم والأسرة أكثر. الأول القيمة التنبؤية الموجبة (Positive predictive value): إذا جاءت

نتيجة الطالب موجبة، فما احتمال أن تكون لديه الحالة فعلاً؟ والثاني القيمة التنبؤية السالبة (Negative predictive value): إذا جاءت نتيجته سالبة، فما احتمال أن تكون الحالة غائبة فعلاً؟ وهذا الثاني هو أهم ما يُسأل عنه في المسح، لأن السالب يعني أن الطالب لن يُتابع.

وهذان الاحتمالان يعتمدان على الأداة وعلى مدى انتشار الحالة في المجموعة التي نفحصها معاً. ولنأخذ مثلاً افتراضياً بأرقام توضيحية: أداة مسح حساسيتها ٨٥% ونوعيتها ٩٥%. طُبقت مرة على ١٠٠ طفل في مسح عام تنتشر الحالة بينهم بنسبة ٢%. ومرة على ٢٠٠ طفل أُحيلوا لوجود مخاوف وتنتشر الحالة بينهم بنسبة ٣٠%:

جدول ٢-٥: الأداة نفسها (الحساسية ٨٥%، والنوعية ٩٥%) في مجموعتين تختلف فيهما نسبة الانتشار. أرقام افتراضية.

لديهم الحالة	مسح عام (١٠٠٠ طفل، الانتشار ٢%)	مجموعة محالة (٢٠٠ طفل، الانتشار ٣٠%)
لديهم الحالة	٢٠	٦٠
موجب صحيح (كشفتهم الأداة)	١٧	٥١
سالب خاطئ (فاتتهم الأداة)	٣	٩
ليست لديهم الحالة	٩٨٠	١٤٠
موجب خاطئ	٤٩	٧
سالب صحيح	٩٣١	١٣٣
القيمة التنبؤية الموجبة	$١٧ \div ٦٦ \approx ٢٦\%$	$٥١ \div ٥٨ \approx ٨٨\%$
القيمة التنبؤية السالبة	$٩٣١ \div ٩٣٤ \approx ٩٩.٧\%$	$١٣٣ \div ١٤٢ \approx ٩٤\%$

في المسح العام تكون ثلاثة أرباع النتائج الموجبة تقريباً خاطئة. لكن النتيجة السالبة موثوقة جداً. وفي المجموعة المحالة ينقلب الأمر: أغلب النتائج الموجبة صحيحة، لكن نحو طفل من كل ستة عشر ممن جاءت نتائجهم سالبة لديه الحالة فعلاً. ولهذا لا تُعد نتيجة المسح الموجبة تشخيصاً، ولا تُعد نتيجته السالبة في مجموعة عالية الخطر سبباً كافياً لإغلاق الملف.

١-٦-٢ العدالة والتحيز في القياس

تجعل معايير الاختبار التربوي والنفسي **العدالة (Fairness)** أساسًا ثالثًا للقياس بجوار الصدق والثبات/ الدقة، وتفرد لها فصلها الثالث بعنوان «العدالة في الاختبار» (AERA et al., 2014). والعدالة بهذا المعنى مسألة صدق في جوهرها: هل تعني الدرجة الشيء نفسه لطالب من مجموعة لغوية أو ثقافية أو جغرافية مختلفة، أو لطالب ذي إعاقة؟ ويُسمى الخلل في ذلك **تحيز القياس**، ويظهر في صور متعددة:

جدول ٦-٢: صور تحيز القياس وكيف تُفحص.

الصورة	معناها	كيف تُفحص؟
تحيز البناء (Construct bias)	البناء نفسه لا يُعرّف أو لا يظهر بالطريقة نفسها في المجموعتين، كأن يُقاس «السلوك الاجتماعي المناسب» بمعايير ثقافة واحدة	مراجعة الخبراء من الثقافتين، ومقارنة البنية العملية
الأداء التفاضلي للبند (Differential item functioning, DIF)	طالبان متساويان في القدرة من مجموعتين مختلفتين لا يتساوى احتمال إجابتهما عن بند بعينه	مقارنة أداء المجموعتين على كل بند بعد مطابقتها في الدرجة الكلية أو القدرة المقدّرة، ثم يراجع الخبراء البنود التي يظهر فيها الفرق
عدم تكافؤ القياس (Measurement non-invariance)	العلاقة بين البنود والبناء تختلف بين المجموعتين، فلا تصح مقارنة درجتهما مباشرة	اختبارات متدرجة في التحليل العاملي التوكيدي متعدد المجموعات: تماثل البنية، ثم تماثل التشبعات، ثم تماثل المقاطع
التحيز التنبئي (Predictive bias)	الدرجة نفسها تتنبأ بأداء لاحق مختلف في المجموعتين	مقارنة خطوط الانحدار بين الدرجة والمحك في المجموعتين: اختلاف الميل أو المقطع دليل على التحيز

والفرق في متوسطات المجموعات قد يكون حقيقيًا؛ فقد تختلف المجموعتان فعليًا في المهارة المقيسة لأسباب تتعلق بفرص التعلم، والمقياس يصف هذا الفرق بدقة. والتحيز هو أن تعني الدرجة نفسها شيئًا مختلفًا باختلاف المجموعة. ولا يكفي فحص الأداة؛ فالعدالة تشمل أيضًا إتاحة البناء لكل مفحوص، كأن يُقدّم للطالب ضعيف البصر الاختبار بخط مكبر ما دام ذلك لا يغير ما يقيسه. أما العدالة في مراحل الإحالة والتحديد، ومؤشرات التمثيل النسبي، فيعود إليها الفصل الرابع عشر.

٧-٢ المعايير والتقنين وتكييف المقياس

الدرجة المعيارية تعني «موقع الطالب بين أقرانه في عينة التقنين». وعينة التقنين (Norm sample) هي المجموعة التي تُطبق عليها المقياس لاستخراج معاييره. فإذا كانت عينة التقنين لا تشبه الطالب، فقدت الدرجة معناها. ولذلك نسأل عن أي مقياس ثلاثة أسئلة قبل أن نستعمله:

1. من في عينة التقنين؟ هل تمثل الأطفال الذين نقيسهم في عمرهم وبيئتهم ولغتهم؟ اختبار قُنن على أطفال أمريكيين يقارن الطفل السعودي بأطفال أمريكيين بدل أقرانه في بيئته.
2. متى قُنن؟ ارتفعت متوسطات الأداء في اختبارات الذكاء عبر الأجيال في كثير من الدول، وتُعرف هذه الظاهرة بأثر فليين (Flynn effect) (Flynn, 1984). وقد قَدَّرَها تحليل بعدي، أي دراسة تجمع نتائج دراسات كثيرة إحصائيًا، بنحو نقطتين إلى ثلاث نقاط معيارية في العقد (Trahan et al., 2014). لكن الظاهرة ليست كونية ولا ثابتة الاتجاه، فقد رُصد تباطؤها أو انعكاسها في بعض الدول في العقود الأخيرة (Bratsberg & Rogeberg, 2018). وحيث يوجد الأثر، تضخم المعايير القديمة درجات الأطفال اليوم، فتخفي إعاقة موجودة في طرف، وتزيد في الطرف الآخر أعداد من يُحدِّدون موهوبين. والدرس العملي واحد في الحالتين: اعرف سنة جمع المعايير، وتعامل بحذر مع الدرجات القريبة من درجات القطع حين تكون المعايير قديمة.
3. بأي لغة وثقافة؟ الترجمة وحدها لا تكفي. فالسؤال عن «الثلاث» أو «عيد الشكر» مألوف لطفل في بيئة غريب في أخرى.

من الترجمة إلى التقنين

ولتبسيط الطريق الذي يقطعه المقياس من ثقافة إلى أخرى، نميز في هذا الكتاب بين ثلاثة مستويات. وهذا التقسيم تبسيط تعليمي من المؤلف، مبني على ما تطلبه الإرشادات الدولية (انظر المربع التالي):

جدول ٧-٢: ثلاثة مستويات لنقل المقياس من ثقافة إلى أخرى (تبسيط من المؤلف).

المستوى	ما يحدث	ما يمكن فعله بالدرجات
الترجمة	نقل البنود إلى العربية حرفيًا أو قريبًا من ذلك	لا يصح استعمال المعايير الأصلية للحكم على الطفل

المستوى	ما يحدث	ما يمكن فعله بالدرجات
تكييف المقياس	تعديل البنود والأمثلة لتناسب الثقافة، مع الترجمة العكسية (إعادة ترجمة النسخة العربية إلى اللغة الأصلية ومقارنتها بالأصل)، ومراجعة الخبراء، وتحليل البنود، والتحقق من تكافؤ القياس بين النسختين ومن خلو البنود من الأداء التفاضلي (انظر ١-٢-٦)	أفضل من الترجمة، لكن المعايير الأصلية تبقى غير مناسبة
التقنين المحلي	تطبيق النسخة المكيفة على عينة ممثلة من المجتمع المحلي، واستخراج معايير جديدة، ودراسة الثبات والصدق في ذلك المجتمع	تصبح الدرجات المعيارية ذات معنى لأطفال ذلك المجتمع

٥٨ للمتخصص: ما الذي تقوله إرشادات الهيئة الدولية للاختبارات فعلاً؟

تتألف إرشادات الهيئة الدولية للاختبارات في ترجمة الاختبارات وتكييفها، في إصدارها الثاني، من ثماني عشرة إرشادة موزعة على ست مجموعات، تتبع مراحل العمل، على خلاف مستويات النقل في جدولنا (International Test Commission, 2017):

1. **الشروط المسبقة:** الحصول على إذن صاحب الحقوق، والتحقق من أن البناء المقيس له معنى في الثقافة الجديدة قبل البدء.
2. **التطوير:** الترجمة وتكييف البنود بفريق مؤهل في اللغتين والموضوع، وتجربة البنود على عينات صغيرة.
3. **التأكيد:** جمع الأدلة التجريبية على تكافؤ النسختين: تكافؤ القياس، وفحص الأداء التفاضلي للبنود، والثبات، والصدق في المجتمع الجديد.
4. **التطبيق:** ضبط تعليمات التطبيق وظروفه بما يقلل مصادر التحيز في المجتمع الجديد.
5. **مقاييس الدرجات والتفسير:** لا تُقارن الدرجات بين المجتمعين إلا بقدر ما يسمح به مستوى التكافؤ الذي ثبت.
6. **التوثيق:** توثيق كل تغيير أجري على الأصل، وتقديم إرشادات لمن يستعمل النسخة الجديدة.

فالتقنين المحلي في جدولنا يقابل تقريبًا اكتمال مرحلتي «التأكيّد» و«مقاييس الدرجات» مع استخراج معايير جديدة. أما نسخة لم تمر بمرحلة التأكيّد، فلا يكفي أن تكون ترجمتها متقنة.

📍 في السياق السعودي: أين تقف معايير المقاييس المستعملة؟

تتفاوت حالة المقاييس المستعملة في المملكة تفاوتًا كبيرًا بين المستويات الثلاثة. ومن أمثلة ما وجدناه منشورًا:

- **مقياس وكسلر لذكاء الأطفال:** التقنين السعودي الذي تذكره الأدبيات هو تقنين الصورة المعدلة (WISC-R) للأعمار ٦-١٦ ضمن مشروع برنامج الكشف عن الموهوبين ورعايتهم بمدينة الملك عبدالعزيز للعلوم والتقنية (النافع وآخرون، ٢٠٠٠، كما ورد في Ejbara, 2024). ولم نجد تقنيًا سعوديًّا منشورًا للإصدارين الرابع أو الخامس عند إعداد الكتاب. ومعنى ذلك في ضوء ما سبق أن معايير الإصدار المقتن محليًا قديمة، وأن معايير الإصدار الحديث غير محلية.
- **مقياس وكسلر لذكاء أطفال ما قبل المدرسة والمرحلة الابتدائية، الإصدار الثالث:** صورة سعودية طبقت على ٧٨٨ طفلًا من ٤ سنوات إلى ٧ سنوات و٣ أشهر من منطقة جدة التعليمية بين عامي ٢٠١٨ و٢٠٢٢، واشتُقت من الصورة الأردنية (Ejbara, 2024). وهي عينة كبيرة، لكنها من منطقة واحدة.
- **مقياس ستانفورد بينيه، الصورة الخامسة:** تُستعمل له نسختان خليجيتان قارنت بينهما دراسة في منطقة الحدود الشمالية (محمد، ٢٠١٦). ولم نجد تفاصيل عيني التقنين فيهما في مصدر منشور.
- **مقياس جيليام لتقدير التوحد، الإصدار الثالث:** نسخة عربية دُرست على ١٧٨ مشاركًا في أعمار ٣ إلى ٢٢ سنة من بيانات تعليمية وإكلينيكية في المملكة، وبلغ معامل ألفا كرونباخ فيها ٠.٩٧ (Alasmari et al., 2024). وهي دراسة لخصائص النسخة، والمعايير الوطنية خطوة لاحقة.

سؤال للتأمل: أي هذه الأمثلة يصل إلى مستوى «التقنين المحلي» في الجدول ٢-٧؟ وماذا ينقص كلاً منها ليصلح لطالب في مدينتك؟

❓ مفاهيم خاطئة شائعة

- «النسخة المترجمة بدقة تحمل صدق الأصل وثباته ومعاييره.» الصدق والثبات والمعايير خصائص لاستعمال الدرجات في مجتمع بعينه، ولا تنتقل مع الترجمة. لا بد من أدلة جديدة من المجتمع الجديد.
- «ما دام معامل الثبات مرتفعًا فالمقياس صادق.» الثبات شرط للصدق لا دليل عليه (انظر الشكل ٢-٣). فالمقياس قد يقيس باتساق شيئًا غير ما نريد.
- «كل فرق كبير بين درجتين يدل على مشكلة.» الفرق قد لا يكون دالًا إحصائيًا، وقد يكون دالًا لكنه شائع في عينة التقنين (انظر ٢-٥-١).

⚠ قبل أن تثق بمقياس معرب

- هل النسخة مرخصة من الناشر الأصلي؟
 - هل أُجريت لها دراسة تكييف منشورة، أم تُرجمت فقط؟ وهل فُحص فيها تكافؤ القياس والأداء التفاضلي للبنود؟
 - هل لها معايير محلية؟ ومتى استُخرجت؟ ومن أي مناطق كانت العينة؟
 - هل نُشرت معاملات ثباتها وأدلة صدقها في البيئة المحلية، ومن أي المصادر الخمسة جاءت هذه الأدلة؟
- إذا كانت الإجابة عن أغلب هذه الأسئلة بالنفي، فاستعمل نتائج المقياس بحذر شديد، واقصرها على الوصف دون التشخيص، واذكر ذلك صراحة في التقرير.

٨-٢ قراءة جدول درجات حقيقي الشكل

تعرض تقارير اختبارات الذكاء عادة جدولًا يلخص الدرجات المركبة. ونعرض في الجدول ٨-٢ قيمًا افتراضية لطالب في العاشرة، على نمط مؤشرات مقياس وكسلر لذكاء الأطفال، الإصدار الخامس:

جدول ٨-٢: درجات افتراضية لطالب في العاشرة على نمط مؤشرات مقياس وكسلر لذكاء الأطفال، الإصدار الخامس.

المؤشر	الدرجة المعيارية	الرتبة المئينية	فترة الثقة ٩٥%	التصنيف الوصفي
الفهم اللفظي	٨٤	١٤	٩٢-٧٨	متوسط منخفض
البصري المكاني	١٠٢	٥٥	١١٠-٩٤	متوسط
الاستدلال السائل	٩٨	٤٥	١٠٥-٩١	متوسط
الذاكرة العاملة	٧٦	٥	٨٥-٧٠	منخفض جدًا
سرعة المعالجة	٨٠	٩	٩١-٧٣	متوسط منخفض
الدرجة الكلية للذكاء	٨٦	١٨	٩٢-٨١	متوسط منخفض

بُنيت فترات الثقة في الجدول حول الدرجة الحقيقية المقدّرة، كما في المقياس الأصلي، ولذلك لا تماثل حول الدرجة (انظر مربع «للمتخصص» في ٢-٥). وقد افترضنا لها معاملات ثبات هي: ٠.٩٤ للفهم اللفظي والاستدلال السائل، و٠.٩٢ للبصري المكاني، و٠.٩٣ للذاكرة العاملة، و٠.٩٠ لسرعة المعالجة، و٠.٩٦ للدرجة الكلية للذكاء (Full Scale IQ, FSIQ). فكيف يقرأ المعلم هذا الجدول؟

1. انظر في الدرجة الكلية وحدود دلالتها أولاً. الدرجة الكلية ٨٦، وفترتها ٨١-٩٢، في المدى المتوسط المنخفض. وهي أثبتت درجات الجدول وأقواها تنبؤًا بالتحصيل.

2. ثم افحص الفروق من حيث دلالتها وندرتها قبل تفسير النمط. الفرق بين أعلى مؤشر وأدناه ٢٦ درجة. والسؤال الذي يعنينا: هل هو دال إحصائيًا؟ وهل هو نادر في عينة التقنين؟ وقد رأينا في ٢-٥ أن الفرق بين الفهم اللفظي والبصري المكاني دال لكنه غير نادر، وأن الفرق بين الذاكرة العاملة وسرعة المعالجة غير دال أصلاً. فإذا تبين أن الفروق دالة ونادرة معًا، صار للحديث عن «نمط» معنى، ولم تعد الدرجة الكلية وحدها تصف الطالب وصفًا كافيًا.

3. لاحظ نقاط القوة. قدرته على الاستدلال والتفكير البصري في المتوسط. وهذه معلومة تعليمية مهمة: يمكن تقديم المعلومات الجديدة بصريًا، والاعتماد على قدرته على الاستنتاج.

4. لاحظ نقاط الضعف وما قد تعنيه في الصف. الذاكرة العاملة وسرعة المعالجة منخفضتان. فقد يفقد التعليمات الطويلة، ويحتاج إلى وقت أطول لإنهاء المهام، وقد يبدو غير منتبه وهو في الحقيقة لا يجاري إيقاع الصف. وهذه فرضيات تنتظر التحقق بالملاحظة في الصف.

5. اقرأ فترات الثقة. فترة الذاكرة العاملة تمتد من ٧٠ إلى ٨٥، فالقول إنها «منخفضة جدًا» في تصنيف المقياس صحيح للدرجة الملاحظة، لكن الفترة تمتد إلى المدى المتوسط المنخفض، فمقدار

الانخفاض غير محسوم.

6. **اسأل: ماذا يعني هذا للتشخيص؟** هذا الجدول وحده لا يشخص شيئاً. لكنه يدعو إلى النظر في احتمال صعوبات التعلم أو اضطراب ضعف الانتباه وفرط الحركة، وإلى جمع أدلة أخرى من التحصيل والملاحظة ومقاييس الانتباه.

١٤ للمختص: جدل علمي حول تحليل ملف القدرات

يدور في علم النفس المدرسي جدل قائم حول قيمة **تحليل ملف القدرات**. أي تفسير نمط القوة والضعف بين المؤشرات. فمن جهة، يرى ناقده أن المؤشرات لا تضيف إلا قليلاً إلى الدرجة الكلية في التنبؤ بالتحصيل، وأن ثبات الفروق بينها أدنى كثيراً من ثبات الدرجة الكلية، وأن الأنماط لم تثبت فائدتها في اختيار تدخل تعليمي بعينه (McGill et al., 2018). ومن جهة أخرى، يرى المدافعون عنه أن التفسير المنضبط للمؤشرات، حين يقوم على فروق دالة ونادرة وتدعمه أدلة من خارج الاختبار، يعين على فهم الطالب وتوجيه التقييم اللاحق، ولا سيما مع مزدوجي الاستثنائية الذين قد تخفي الدرجة الكلية تباين قدراتهم (Kaufman et al., 2016). والموقف الذي يأخذ به هذا الكتاب وسط بينهما: تبدأ بالدرجة الكلية وحدود دلالتها، ولا تنتقل إلى تفسير النمط إلا بعد فحص الفروق إحصائياً، وتعامل مع النمط بوصفه فرضية تحتاج إلى أدلة أخرى. ويعود الفصل السادس إلى هذا الجدل في نماذج أنماط القوة والضعف.

٩-٢ الأخلاقيات والمسؤولية المهنية

يحتكم من يقيس الأطفال إلى نوعين من المراجع ينبغي ألا يُخلط بينهما. الأول **مرجع فني**، أهمه معايير الاختبار التربوي والنفسي التي تصدرها ثلاث جمعيات أمريكية (AERA et al., 2014). وهي تحدد ما ينبغي أن تستوفيه الاختبارات واستعمالها من صدق وثبات/دقة وعدالة، ويخصص فصلها الثالث للعدالة في الاختبار. ويكملها في الاستعمال الدولي إرشادات الهيئة الدولية للاختبارات في استعمال الاختبارات، وهي تفصل الكفايات التي يحتاج إليها من يستعمل الاختبارات (International Test Commission, 2001). والثاني **موثوق أخلاقية** تصدرها الهيئات المهنية لأعضائها، ومنها المبادئ الأخلاقية ومدونة السلوك للجمعية الأمريكية لعلم النفس (American Psychological Association, 2017). والمعايير المهنية للرابطه الوطنية للأخصائيين النفسيين المدرسين، وفيها مبادئها الأخلاقية (National Association of School Psychologists, 2020). والأخلاقيات والمعايير المهنية لمجلس الأطفال غير العاديين (Council for Exceptional Children, 2015). ومن المبادئ التي تتفق عليها هذه المراجع ما يمس عمل المعلم مباشرة:

- **الموافقة المستنيرة:** تعرف الأسرة سبب التقييم، وما سيُطبق، ومن سيطلع على النتائج، وتوافق قبل البدء. وفي المملكة يشترط الدليل التنظيمي للتربية الخاصة موافقة ولي الأمر خطيًا بعد إشعاره رسميًا (وزارة التعليم، ١٤٣٦هـ)، وتنص القواعد التنظيمية على أخذ إذنه «لغرض التشخيص» وشرح حالة ابنه له «بلغة مفهومة» (وزارة المعارف، ١٤٢٢هـ).
 - **السرية:** نتائج الطالب لا تُناقش في غرفة المعلمين، ولا تُرسل في مجموعات المراسلة، ولا يطلع عليها إلا من يحتاج إليها لخدمة الطالب. وفي المملكة يسري نظام حماية البيانات الشخصية الصادر بالمرسوم الملكي رقم (م/١٩) عام ١٤٤٣هـ، وتُعد البيانات الصحية فيه من البيانات الحساسة.
 - **أمن الاختبار:** لا تُصور صفحات الاختبار، ولا يُدرب الطالب على بنوده، فذلك يُفسد صدق نتائجه ونتائج غيره.
 - **الاستعمال المناسب:** لا يُستعمل مقياس لغرض لم يُصمم له، ولا يُبنى قرار مهم على درجة واحدة. وهذا ما ينص عليه الدليل التنظيمي بعبارة «عدم الاكتفاء بنتائج مقياس واحد».
 - **الإنصاف:** يراعي المقيم لغة الطالب وثقافته وإعاقاته الأخرى، فطالب ضعيف البصر لا يُحكم على ذكائه باختبار بصري (انظر ٦-٢-١).
 - **حدود الكفاءة:** لا يطبق المقيم أداة لم يُدرَّب عليها، ولا يفسر نتائج طالب من خلفية لغوية أو ثقافية لا يعرفها دون استشارة من يعرفها. والكفاءة الثقافية لا تُكتسب بقراءة دليل، ومن حدودها أن يعرف المقيم متى يحتاج إلى مترجم مؤهل أو زميل من المجتمع نفسه.
 - **تعارض المصالح:** حين يكون من يشخص هو نفسه من يقدم الخدمة المدفوعة التي يوصي بها، أو حين يرتبط تقييم المعلم لطالبه بحوافز أو بعدد المقبولين في برنامجه، يتأثر الحكم ولو بغير قصد. فيُفصح عن ذلك، ويُطلب رأي مستقل حين يمكن.
- ويقع بين هذه المبادئ توتر يتكرر في الواقع: **حق الأسرة في الاطلاع** في مقابل **أمن الاختبار**. فلأسرة حق في أن تعرف كل ما يتعلق بتقييم ابنها وبرنامجه، وفي أن تطلب إعادة التشخيص إذا رأت أنه لم يوضع في البرنامج المناسب (وزارة المعارف، ١٤٢٢هـ). وفي المقابل، فإن نسخ بنود الاختبار وتسليمها يعرض الاختبار للانتشار، وقد يخالف حقوق الناشر. والحل الذي تميل إليه المواثيق المهنية أن تُشرح النتائج للأسرة تفصيلًا، وأن تُعرض عليها إجابات ابنها وطريقة تصحيحها في حضور المختص، دون تسليم كراسة الاختبار أو نسخ بنوده.

مستويات التأهيل: من يطبق ماذا؟

يصنف كثير من الناشرين مقاييسهم في مستويات تأهيل تحدد من يحق له شراؤها واستعمالها، ويختلف التفصيل بين ناشر وآخر، لكن الفكرة العامة واحدة: مستوى أول لأدوات يستطيع المعلمون وغيرهم تطبيقها بعد قراءة الدليل، ومستوى ثانٍ يتطلب تدريبًا في القياس أو ترخيصًا مهنيًا، ومستوى ثالث

للأدوات المعقدة كاختبارات الذكاء الفردية وأدوات التشخيص، ويتطلب دراسات عليا في مجال متخصص أو ترخيصًا مهنيًا متقدمًا. ومن أمثلة هذا المستوى الأخير مقياس وكسلر لذكاء الأطفال. أما ما تنص عليه الوثائق السعودية التي اطلعنا عليها، فيلخصه الجدول التالي:

جدول ٩-٢: من يطبق ماذا في المملكة بحسب الوثائق المنشورة التي اطلعنا عليها عند إعداد الكتاب.

الدور	ما تنص عليه الوثائق	ما لم نجد له نصًا
معلم التعليم العام	عضو في فريق القياس والتشخيص بالمدرسة (وزارة التعليم، ١٤٣٦هـ)	الأدوات التي يحق له تطبيقها. ودوره العملي الإحالة والملاحظة والأدوات غير الرسمية
معلم التربية الخاصة	عضو في فريق القياس والتشخيص (وزارة المعارف، ١٤٢٢هـ؛ وزارة التعليم، ١٤٣٦هـ)	قائمة بالمقاييس المقننة التي يطبقها
الأخصائي النفسي في المدارس والمعاهد	تسند إليه القواعد التنظيمية، باسم «معلم التدريبات السلوكية (الأخصائي النفسي)»، القياس والتشخيص بمقاييس الذكاء المقننة ومقاييس السلوك التكيفي والأساليب غير الرسمية، وإعادة التشخيص عند الحاجة (وزارة المعارف، ١٤٢٢هـ). ويشترط الدليل أن يطبق الأدوات ويفسرها «متخصصون مؤهلون»، وأن تكون «معتمدة من قبل وزارة التعليم» (وزارة التعليم، ١٤٣٦هـ)	قائمة منشورة بالمقاييس المعتمدة، ومعايير التأهيل المطلوبة لكل مقياس
الأخصائي النفسي في القطاع الصحي	يُصنّف ويُرخّص لدى الهيئة السعودية للتخصصات الصحية. ففي الطبعة التي اطلعنا عليها من دليل التصنيف يُصنّف حامل البكالوريوس «أخصائيًا مساعدًا» يعمل تحت إشراف، ويُصنّف «أخصائي علم نفس» بعد خبرة إكلينيكية لا تقل عن ثلاث سنوات (الهيئة السعودية للتخصصات الصحية، ٢٠٠٩). وقيس اختبار الترخيص في علم النفس، ضمن ما يقيسه، تطبيق المعرفة في التقييم النفسي والتشخيص (الهيئة السعودية للتخصصات الصحية، ٢٠٢٤)	الطبعة الحالية من دليل التصنيف

الدور	ما تنص عليه الوثائق	ما لم نجد له نصًا
الطبيب	يشارك في فريق القياس والتشخيص بالمدرسة «إن وجد» (وزارة التعليم، ١٤٣٦هـ)	نص عام يحدد التشخيصات التي يختص بها الطبيب دون غيره في الأنظمة التعليمية

ولم نجد نصًا منشورًا في المملكة يقابل مستويات التأهيل التي يعتمد عليها الناشرون في شراء الاختبارات النفسية المقننة وتطبيقها. وإلى أن يصدر مثل هذا النص، يبقى المرجع العملي هو اشتراطات الناشر نفسه، ومبدأ حدود الكفاءة السابق.

📍 في السياق السعودي: معضلتان أخلاقيتان

المعضلة الأولى: الموافقة بالسكوت. ينص الدليل التنظيمي على أن ولي الأمر إذا لم يرد خلال مدة لا تتجاوز أربعة عشر يومًا من تاريخ الإشعار الخطي «يعتبر ذلك إذنًا بالموافقة» (وزارة التعليم، ١٤٣٦هـ). والنص يحمي حق الطالب في ألا يتأخر تقييمه بسبب إهمال الإشعار. لكن مبدأ الموافقة المستنيرة يفترض أن الأسرة فهمت ما ستوافق عليه. فماذا يفعل الفريق إذا علم أن الأب لا يقرأ، أو أن الإشعار أُرسِل مع الطالب ولم يصل؟ يسمح النص بالبداية، وتقتضي الأخلاقيات محاولة الوصول إلى الأسرة بطريقة أخرى، وتوثيق هذه المحاولة، قبل الاكتفاء بالسكوت.

المعضلة الثانية: الأم التي تطلب البنود. تطلب أم نسخة من كراسة اختبار الذكاء الذي طُبِق على ابنها، لأنها تريد أن «تدربه على الأسئلة التي أخطأ فيها» قبل إعادة التقييم في العام القادم. للأم حق في الاطلاع والفهم، وللاختبار أمن يحمي صدق نتائج ابنها نفسه في المرة القادمة. كيف تصوغ ردًا يحترم الحقيقتين؟

🏠 للمعلم: عودة إلى أم ماجد

يمكن أن يجيب المعلم الآن عن سؤال أم ماجد بهذه الصيغة: «الرقم ٣ ليس درجة نجاح، بل يعني أن أداء ماجد في هذا الاختبار أعلى من نحو ٣% من أقرانه في العمر نفسه. أي أنه يحتاج إلى دعم واضح. أما سؤالك عن الإعاقة، فالاختبار وحده لا يجيب عنه. درجته قريبة من الدرجة التي يُعتمد عليها في القرار، والشريط حولها يمتد على جانبها، ولذلك سيجمع الفريق معلومات عن مهاراته اليومية في البيت والمدرسة قبل أي حكم.»

خاتمة

ينهض ما عرضناه في هذا الفصل على فكرة جامعة: الدرجة تقدير تقريبي لموقع الطالب بين أقرانه. فالدرجات المعيارية على اختلافها لغات لموقع واحد على المنحنى الطبيعي، والرتبة المئينية نسبة من تفوق عليهم الطالب، وكل درجة يحيط بها خطأ تبين فترة الثقة مداه، والفرق بين درجتين أقل دقة من كل منهما. ويترتب على ذلك أن القرار الذي تعبر فيه فترة الثقة درجة القطع قرار مهني يحتاج إلى أدلة أخرى. وأن الفرق بين درجتين لا يُبنى عليه تفسير إلا إذا كان دالاً ونادراً، وأن الدرجة المتطرفة يُتوقع أن تترد نحو المتوسط عند إعادة القياس.

ولا تكفي الدقة وحدها؛ فصلاحيّة الدرجة لغرضها مسألة صدق تُجمع أدلته من مصادر متعددة، وصلاحيّتها للطالب مسألة معايير وعدالة، إذ لا تصلح المعايير إلا إذا شابهت عينها الطالب وكانت حديثة بما يكفي، ولا تنقل الترجمة الدقيقة صدق الأصل ولا معايير. ويبقى أمام المقيّم كلما استعمل أداة مترجمة أو معايير من بيئة أخرى سؤال لا تحسمه الحسابات: هل تشبه عينة التقنين الطالب الذي أمامه؟ والإجابة عنه مرهونة بما يتوافر من أدلة فنية، وبما يصرح به التقرير من حدود.

١٠٢ الخلاصة والمسرد والأنشطة

خلاصة الفصل

- المقاييس معيارية المرجع تقارن الطالب بأقرانه وتخدم التشخيص، والمقاييس محكية المرجع تقارنه بمستوى إتقان وتخدم التدريس، وبعض الأدوات كالقياس المبني على المنهج تُفسّر بالطريقتين.

- الدرجات المعيارية كلها لغات مختلفة لموقع واحد على المنحنى الطبيعي، ويمكن التحويل بينها عبر الدرجة الزائدية. والتسميات الوصفية تختلف بين الاختبارات وإصداراتها.
- انحرافان معياريان دون المتوسط أو فوقه يحدان نحو ٢% من الناس في كل طرف، ومنهما تأتي درجات قطع كثيرة في التشخيص.
- الرتبة المئينية نسبة من تفوق عليهم الطالب، لا نسبة إجاباته الصحيحة، ومسافاتها غير متساوية.
- كل درجة تقدير يحيط به خطأ، وفترة الثقة تبين مداه. وحين تعبر الفترة درجة القطع يصبح القرار مهنيًا يستند إلى أدلة أخرى.
- الفرق بين درجتين أقل دقة من كل منهما، ويُسأل عنه سؤالان: هل هو دال إحصائيًا؟ وهل هو نادر في عينة التقنين؟
- الدرجات المتطرفة ترد نحو المتوسط عند إعادة القياس، ودقة الاختبار تضعف قرب أرضيته وسقفه.
- الصدق مفهوم واحد جامع تُجمع أدلته من خمسة مصادر، والعدالة أساس ثالث بجواره وجوار الثبات/الدقة. والقيمة التنبؤية للنتيجة تتغير بتغير نسبة الانتشار.
- المعايير لا تصلح إلا إذا كانت عينتها تشبه الطالب وحديثه بما يكفي، والترجمة الدقيقة لا تنقل صدق الأصل ولا معاييره.
- معايير الاختبار وثيقة فنية، والمواثيق الأخلاقية للهيئات المهنية مرجع مستقل عنها، والتنظيمات المحلية تضيف إليهما ما يلزم المقيّم في بلده.

مسرد المصطلحات

عينة التقنين	المجموعة التي طبق عليها المقياس لاستخراج معاييره. ويُقارن بها درجة كل طالب لاحقًا.	Norm sample
الدرجة الخام	عدد الإجابات الصحيحة أو مجموع النقاط قبل تحويله، ولا معنى له وحده.	Raw score
الدرجة المعيارية	درجة محولة لها متوسط وانحراف معياري ثابتان، أشهرها ما متوسطه ١٠٠ وانحرافه ١٥.	Standard score

النسبة المئوية من أقران الطالب في عينة التقنين (من عمره أو صفه) الذين حصلوا على درجة أقل من درجته: فالرتبة المئينية ٨٤ تعني أن درجته أعلى من درجات ٨٤% منهم، وهي غير النسبة المئوية للإجابات الصحيحة.	الرتبة المئينية Percentile rank
مقياس لتشتت الدرجات حول متوسطها.	الانحراف المعياري Standard deviation
الدرجة التي تفصل في قرار ما بين فئتين، كالإحالة وعدمها أو استيفاء المحك وعدمه.	درجة القطع Cut-off score
اتساق نتائج المقياس عبر الزمن أو البنود أو المقيدين، ودقة الدرجة عمومًا.	الثبات/الدقة Reliability/precision
تقدير لمقدار تذبذب درجة الفرد حول درجته الحقيقية بسبب عدم الثبات التام.	الخطأ المعياري للقياس Standard Error of Measurement, SEM
مدى من الدرجات يُتوقع بنسبة محددة أن تقع فيه الدرجة الحقيقية.	فترة الثقة Confidence interval
مقدار الخطأ المتوقع في الفرق بين درجتين، ومنه يُحسب أصغر فرق يُعد دالًا إحصائيًا.	الخطأ المعياري للفرق Standard error of the difference
نسبة من يظهر لديهم فرق أو نمط معين في عينة التقنين، ومنها يُعرف هل الفرق نادرًا أم شائع.	معدل الأساس Base rate
ميل الدرجات المتطرفة إلى الاقتراب من المتوسط عند إعادة القياس.	الارتداد نحو المتوسط Regression toward the mean
قدرة الاختبار على التمييز بين المستويات المنخفضة جدًا والمرتفعة جدًا من الأداء.	الأرضية والسقف Floor and ceiling

نظرية الاستجابة للمفردة Item Response Theory, IRT	إطار قياسي يصف كل بند بمنحنى يربط القدرة باحتمال الإجابة الصحيحة، وتختلف فيه دقة القياس باختلاف مستوى القدرة.
الصدق Validity	مدى دعم الأدلة لتفسير درجات المقياس في غرض معين، وتُجمع أدلته من خمسة مصادر.
الأداء التفاضلي للبند Differential item functioning, DIF	اختلاف احتمال الإجابة عن بند بين مجموعتين متساويتين في القدرة.
تكافؤ القياس Measurement invariance	أن يقيس المقياس البناء نفسه بالطريقة نفسها في مجموعتين، فتصح المقارنة بين درجتهما.
أثر فلين Flynn effect	تغير متوسطات الأداء في اختبارات الذكاء عبر الأجيال، وقد كان ارتفاعاً في أغلب القرن العشرين، مما يجعل المعايير القديمة تضخم الدرجات حيث يوجد.
الاختبار معياري المرجع Norm-referenced test	اختبار تُفسر فيه درجة الطالب بمقارنتها بدرجات أقرانه في عينة التقنين.
الاختبار محكي المرجع Criterion-referenced test	اختبار تُفسر فيه درجة الطالب بمقارنتها بمستوى أداء أو إتقان محدد مسبقاً.
العمر المكافئ Age equivalent	العمر الذي يساوي متوسط الدرجات الخام لأطفاله درجة الطالب الخام، ولا يعني أن مهاراته تشبه مهارات أطفال ذلك العمر.
الحساسية Sensitivity	نسبة من تكشفهم الأداة بحق ممن لديهم الحالة.
النوعية Specificity	نسبة من تستبعدهم الأداة بحق ممن ليس لديهم الحالة.
القيمة التنبؤية الموجبة Positive predictive value	احتمال أن تكون الحالة موجودة فعلاً لدى من جاءت نتيجته موجبة، ويتأثر بمدى انتشار الحالة في المجموعة المفحوصة.

احتمال أن تكون الحالة غائبة فعلاً لدى من جاءت نتيجته سالبة.	القيمة التنبؤية السالبة Negative predictive value
تعديل بنود المقياس وأمثله لتناسب ثقافة جديدة مع دراسة ذلك علمياً. وهو أوسع من الترجمة وأضيق من التقنين المحلي. ويختلف عن «التكيف» بمعنى التسهيلات المقدمة للطلاب أثناء الاختبار.	تكيف المقياس Test adaptation

أسئلة للمراجعة

- صنّف كل نتيجة مما يلي بأنها معيارية المرجع أو محكية المرجع. وعلّل: (أ) «قرأ ٤٢ كلمة صحيحة في الدقيقة، وهدفه في نهاية الفصل ٥٠». (ب) «رتبته المئينية ٣٠ في اختبار مقنن للقراءة». (ج) «أتقن ٧ من ١٠ مهارات في قائمة مهارات الجمع». (الهدف ١)
- حوّل الدرجة التائية ٣٥ إلى درجة معيارية متوسطها ١٠٠ وانحرافها ١٥، واذكر تصنيفها الوصفي في الجدول ٢-٢. (الهدف ٢)
- حصلت طالبة على الرتبة المئينية ٩٨ في اختبار رياضيات. ماذا يعني ذلك؟ وماذا لا يعني؟ ولماذا لا يصح القول إن طالباً انتقل من الرتبة ٤٥ إلى ٥٥ تقدم بقدر طالب انتقل من ٨٥ إلى ٩٥؟ (الهدف ٣)
- مقياسان للانتباه: ثبات الأول ٠.٧٠ وثبات الثاني ٠.٩٢. أيهما تفضل لقرار تشخيصي؟ ولماذا؟ احسب الخطأ المعياري للمقياس لكل منهما إذا كان الانحراف المعياري ١٠. (الهدف ٤)
- في الجدول ٢-٨، احسب الخطأ المعياري للفرق بين الاستدلال السائل (ثباته ٠.٩٤) والذاكرة العاملة (ثباتها ٠.٩٣)، ثم احكم: هل الفرق بينهما دال عند مستوى ٠.٠٥؟ وما السؤال الذي يبقى قبل أن تفسره؟ (الهدف ٤)
- أداة مسح حساسيتها ٩٠% ونوعيتها ٩٠%، طبقت على ٤٠٠ طفل تنتشر بينهم الحالة بنسبة ٥%. احسب القيمتين التنبؤيتين الموجبة والسالبة، واطرح لأسرة ماذا تعني النتيجة الموجبة لابنها. (الهدف ٥)
- نسخة عربية من مقياس تقدير سلوكي تُستعمل في مركز تشخيص بمعاييرها الأمريكية. اكتب لكل مصدر من مصادر أدلة الصدق الخمسة سؤالاً واحداً تطرحه على من أعد النسخة، واذكر المخاطر المحتملة على طالب سعودي. (الهدف ٦)
- أرسل زميل في مجموعة مراسلة المعلمين صورة لصفحة الملخص في تقرير تشخيصي لطالب، يسأل عن معنى إحدى الدرجات. ما المبادئ التي يمسها هذا التصرف؟ وكيف تجيبه؟ (الهدف ٧)

حالة للتطبيق

نورة في الثالثة عشرة، في الصف الأول المتوسط. أحوالها معلمتها لضعف واضح في التحصيل منذ سنوات. طُبق عليها اختبار ذكاء فردي معرب، جُمعت معاييرها قبل نحو ثماني عشرة سنة على عينة من منطقة غير منطقتها. فكانت درجتها الكلية ٧٤، ومعامل ثبات الدرجة الكلية ٠.٩٦. وكان مؤشر الفهم اللفظي ٨٦ (ثباته ٠.٩٤)، ومؤشر الاستدلال ٧٢ (ثباته ٠.٩٠). وتشير استمارة سلوك تكفي أجابت عنها المعلمة وحدها إلى قصور في مجال واحد. ويرى أحد أعضاء الفريق أن «الدرجة دون ٧٥، فالطالبة مؤهلة لبرنامج الإعاقة الفكرية».

احسب فترة الثقة للدرجة الكلية، والخطأ المعياري للفرق بين المؤشرين. ثم احكم: هل تكفي الأدلة لقرار الأهلية؟ علل حكمك بثلاثة أمور على الأقل تعلمتها في هذا الفصل، واذكر الأدلة الناقصة التي تطلبها قبل القرار. (الأهداف ٤ و٦ و٧)

أنشطة التعلم

1. **مسطرة الدرجات:** ارسم على ورقة منحنى طبيعيًا، وضع عليه درجات خمسة طلاب افتراضيين من أنواع مختلفة: معيارية، وتائية، وموزونة، ورتبة مئينية. ثم رتبهم من الأدنى إلى الأعلى. *الناتج:* رسم واحد مع جدول تحويل. *الزمن:* نحو ٣٠ دقيقة. (الهدفان ٢ و٣)
2. **ترجمة للأسرة:** أعد صياغة الجدول ٢-٨ في فقرة من خمس جمل تفهمها أسرة لم تدرس القياس، تذكر فيها فترة الثقة وما لا يعنيه الجدول. وإن أردت أن تتدرب على تقرير حقيقي، فليكن ذلك بموافقة الأسرة وإدارة المدرسة، وبعد إخفاء كل ما يدل على الطالب، وتحت إشراف مختص، ودون أن يُستعمل ناتج التدريب في أي قرار. *الناتج:* فقرة واحدة. *الزمن:* نحو ٢٠ دقيقة. (الأهداف ٣ و٤ و٧)
3. **فحص دليل مقياس:** اختر مقياسًا يُتاح لك دليله الفني، أو دراسة منشورة عن نسخة عربية من مقياس، واستخرج منها: وصف عينة التقنين، وسنة جمع المعايير، ومعاملات الثبات، وأدلة الصدق مصنفة بحسب المصادر الخمسة، وأي فحص للأداء التفاضلي أو لتكافؤ القياس. ثم احكم: هل يصلح لطالب في مدينتك؟ *الناتج:* جدول من صفحة واحدة. *الزمن:* نحو ساعتين. (الهدف ٦)

مراجع الفصل

- محمد، ح. أ. ع. (٢٠١٦). الفروق بين النسختين الخليجيتين لمقياس ستانفورد بينيه الصورة الخامسة في تقييم وتشخيص الذكاء والقدرات العقلية لدى ذوي الإعاقة بمنطقة الحدود الشمالية. *مجلة جامعة جازان للعلوم الإنسانية*, ٥ (ملحق), ٢٣٦-٢٥٤. <https://search.mandumah.com/Record/1103082>
- نظام حماية البيانات الشخصية، مرسوم ملكي رقم (م/١٩) (٢٠١٧/٢/١٤هـ)، المعدل بالمرسوم الملكي رقم (م/١٤٨) (١٤٤٤/٩/٥هـ).
- الهيئة السعودية للتخصصات الصحية. (٢٠٠٩). *دليل التصنيف المهني للممارسين الصحيين (ط٥)*. https://cams.ksu.edu.sa/sites/cams.ksu.edu.sa/files/attach/Professional_Registration_Classification_Manual_vr.6.pdf
- Saudi Psychology Licensure Examination (SPSLE): *Applicant guide*. <https://scfhs.org.sa/sites/default/files/2024-03/SPSLE%20Applicant%20Guide%202024.pdf>
- وزارة التعليم. (١٤٣٦هـ). *الدليل التنظيمي للتربية الخاصة (الإصدار الأول للعام الدراسي ١٤٣٦-١٤٣٧هـ)*. <https://moe.gov.sa/en/aboutus/aboutministry/RPRLibrary/الدليل%20التنظيمي%20للتربية%20الخاصة.pdf>
- وزارة المعارف. (١٤٣٢هـ). *القواعد التنظيمية لمعاهد وبرامج التربية الخاصة*. <https://moe.gov.sa/en/aboutus/aboutministry/RPRLibrary/القواعد%20التنظيمية%20لمعاهد%20وبرامج%20التربية%20الخاصة%20201422%20هـ.pdf>
- Alasmari, M., Alduais, A., Qasem, F., Almaghlouth, S., & AlAmri, L. (2024). Predicting language, cognition, and social skills in individuals with autism spectrum disorder using the Arabic validation of Gilliam Autism Rating Scale–Third Edition. *Applied Neuropsychology: Child*. Advance online publication. <https://doi.org/10.1080/21622965.2024.2417812>
- Alotaibi, B. M., & Alotaibi, A. A. (2021). Exploratory and confirmatory factor analyses of the Arabic version of the Childhood Autism Rating Scale. *Research in Autism Spectrum Disorders*, 86, Article 101827. <https://doi.org/10.1016/j.rasd.2021.101827>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct* (2002, amended effective June 1, 2010, and January 1, 2017). <https://www.apa.org/ethics/code>

- Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*, 5(4), 313–326.
- Bratsberg, B., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences*, 115(26), 6674–6678.
- Council for Exceptional Children. (2015). *What every special educator must know: Professional ethics and standards* (7th ed.). CEC.
- Ejbara, M. A. (2024). Psychometric characteristics of a Saudi version of Wechsler Intelligence Scale for Preschool and Young Children. *Dirasat: Educational Sciences*, 51(2), 35–52. <https://doi.org/10.35516/edu.v51i2.6661>
- Emam, M. M., Al-Sulaimani, H., Omara, E., & Al-Nabhany, R. (2020). Assessment of adaptive behaviour in children with intellectual disability in Oman: An examination of ABAS-3 factor structure and validation in the Arab context. *International Journal of Developmental Disabilities*, 66(4), 317–326. <https://doi.org/10.1080/20473869.2019.1587939>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum.
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95(1), 29–51.
- International Test Commission. (2001). International guidelines for test use. *International Journal of Testing*, 1(2), 93–114.
- International Test Commission. (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). <https://www.intestcom.org>
- Kaufman, A. S., Raiford, S. E., & Coalson, D. L. (2016). *Intelligent testing with the WISC-V*. Wiley.
- Lohman, D. F., & Korb, K. A. (2006). Gifted today but not tomorrow? Longitudinal changes in ability and achievement during elementary school. *Journal for the Education of the Gifted*, 29(4), 451–484.
- McGill, R. J., Dombrowski, S. C., & Canivez, G. L. (2018). Cognitive profile analysis in school psychology: History, issues, and continued concerns. *Journal of School Psychology*, 71, 108–121.

- National Association of School Psychologists. (2020). *The professional standards of the National Association of School Psychologists*. NASP.
- Salvia, J., Ysseldyke, J. E., & Witmer, S. (2017). *Assessment in special and inclusive education* (13th ed.). Cengage Learning.
- Sattler, J. M. (2018). *Assessment of children: Cognitive foundations and applications* (6th ed.). Jerome M. Sattler, Publisher.
- Schalock, R. L., Luckasson, R., & Tassé, M. J. (2021). *Intellectual disability: Definition, diagnosis, classification, and systems of supports* (12th ed.). AAIDD.
- Seif Eldin, A., Habib, D., Noufal, A., Farrag, S., Bazaid, K., Al-Sharbaty, M., Badr, H., Moussa, S., Essali, A., & Gaddour, N. (2008). Use of M-CHAT for a multinational screening of young children with autism in the Arab countries. *International Review of Psychiatry*, 20(3), 281–289. <https://doi.org/10.1080/09540260801990324>
- Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin*, 140(5), 1332–1360.
- Wechsler, D. (2014). *Wechsler Intelligence Scale for Children* (5th ed.). Pearson.

متى يصح الحكم بأن لدى الطالب إعاقة أو اضطراباً أو موهبة أو استثنائية مزدوجة، وعلى أي أساس؟

يتتبع هذا الكتاب مراحل التشخيص والتعرف تتبعاً منهجياً، من الإحالة وجمع الأدلة إلى تفسيرها وبناء القرار عليها. ويمضي إلى ما وراء تعريف الفئات وعرض محكاتها، فيتناول طبيعة أدوات القياس، وإجراءات تطبيقها، والمؤهلات اللازمة لمن يطبقها، وأصول تفسير نتائجها، والحد الذي تكفي عنده الأدلة لبناء الحكم.

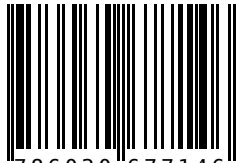
ويجمع الكتاب بين ميدانين تتناولهما المراجع عادة متفرقين: الإعاقات والاضطرابات، والموهبة والاستثنائية المزدوجة. ويقرأ المعايير الدولية والتجربة الأمريكية في ضوء التنظيمات السعودية المنشورة، ويبين حال النسخة العربية من كل مقياس يعرضه.

ويقع في خمسة عشر فصلاً تتناول أسس القياس، وأساليب جمع الأدلة الرسمية وغير الرسمية والبديلة، ونماذج التعرف، ثم فئات الاستثنائية فئةً فئةً، مدعومة بمسارات للقرار وبطاقات للمقاييس وحالات تطبيقية، وتختتم بالعدالة في التشخيص، وكتابة التقرير، وبناء الخطة التربوية الفردية.

د. عمر عبدالله الصمغاني

المرجع في التشخيص والتعرف على ذوي الاستثنائية

ISBN 978-603-06-7714-6



9 786030 677146