١

الجذور والقضايا المهنية

تطالعنا المعاجم اللغوية بتعاريف عديدة لكلمة "اختبار" إلا أن المعنى الذي سيستخدم في هذا الكتاب هو معنى الاختبار النفسي ويعني مجموعة الأسئلة والمسائل والمهام المصممة لاستثارة ردود وإجابات يتم استخدامها لاحقاً في تحديد وقياس صفات الأفراد أو إمكاناتهم أو تحصيلهم. وغالباً ما تكون الاختبارات التي تحدد صفات الأشخاص اختبارات للشخصية أو مقاييس للسمات أو مؤشرات للقياس، وقد تم تناول موضوع اختبارات الشخصية بشكلٍ تفصيلي في كتابين سابقين من هذه السلسلة هما: سلالم التقدير وقوائم الرصد، والاستبانات ومقاييس السمات. وغالباً ما ينظر إلى هذا النوع من أدوات اختبار الشخصية على أنها مقاييس عاطفية لاإدراكية أو لاذهنية تقيم دوافع واهتمامات ومواقف وميول الأفراد بالإضافة إلى خصائص وصفات أساليبهم الشخصية. يتناول هذا الكتاب، على عكس الكتب الآنفة الذكر، أدوات قياس الشخصية المعنية بالقدرات الإدراكية للأفراد والتي تتألف من اختبارات تقيس عمليات ومنتجات الذهن من ذاكرة واستيعاب واستدلال وحل للمشكلات.

وعلى الرغم من أن معظم هذه الاختبارات تتطلب استخدام ورقة وقلم إلا أن بعضها يمكن أن يكون شفهياً أو اختباراً للأداء. وتقسم الاختبارات الإدراكية إلى اختبارات تحصيلية (Achievement Tests) تقيس درجة معرفة الفرد أو مهارته ضمن مادة علمية معينة أو مشروع إنتاجي معين وعند النقطة الزمنية التي يتم فيها الإختبار، واختبارات للاستعداد (Aptitude Tests) تقيس قدرة الأفراد على تعلم مهارة أو مادة علمية معينة. ويكمن الاختلاف الرئيسي بين اختبارات الاستعداد والاختبارات التحصيلية في كون الأخيرة أدوات تقيس الماضي والحاضر أو ما تمكن الفرد من تعلمه وتحصيله بينما تعنى اختبارات الاستعداد بقياس الإمكانيات المستقبلية للفرد أو ما يمكن للشخص تعلمه في حال توفر التعليم والتدريب المناسبين له. ونظراً لأن كلمة "قدرة" تشمل مفاهيم الذكاء والتحصيل والاستعداد، يشار عادةً إلى هذا النوع من المقاييس الإدراكية على أنها اختبارات قدرة، وعليه تعتبر اختبارات الذكاء اختبارات للقدرات العقلية بشكلٍ عام بينما يقتصر عمل اختبارات الاستعداد على قياس إمكانيات الأفراد في مجالات محددة فقط.

مما لا شك فيه أن كلمة "اختبار" كلمة مكروهة من قبل الملايين من الناس الذين يتعين عليهم كل عام أن يثبتوا درجة معرفتهم ومهارتهم في مادة دراسية ما أو في حرفة معينة أو في غيرهما من المجالات التي يتطلب نجاح أو استمرار الأفراد فيها تحقيقهم لمعاييرأداء معينة. ويرى الناس الامتحانات والاختبارات كعقبات وعوائق تحول بينهم وبين الحصول على الإنتماء الكامل لطبقة معينةأو مؤسسة أو مهنة أو أي مجموعة تخصصية يرغبون في الإنتماء إليها. ومن الجدير بالذكر أن هنالك امتحانات لا يترتب عليها إلتحاق شخصٍ ما (أو عدم إلتحاقه) ببرنامج تعليمي أو تدريبي أو مهني معين كامتحانات القدرة التي تهدف

إلى تصنيف الأشخاص ضمن البرنامج أو المنصب أو مجال العمل الأنسب لهم. وتهدف أنواع أخرى من الإختبارات إلى تحديد مصير بعض الأشخاص من حيث منحهم شهادة أو وثيقة تخرج أو غيرهما مما يوثق اجتياز هؤلاء الأشخاص وبنجاح مستوى الأداء المقبول. كما تستخدم أنواع أخرى من الإختبارات كأدوات تشخيصية لتحديد المشكلات أو الاحتياجات الخاصة عند بعض الأفراد أو المجموعات ومن ثم وضع خطط لمعالجتها.

الخلفية التاريخية

لقد قاس الإنسان ومنذ فجر التاريخ القدرات البشرية بطريقةٍ أو بأخرى وعلى الرغم من أن هذه "الاختبارات" لم تكن علميةً بمقاييس اليوم إلا أنها وبدون شك كانت على درجةٍ من الشرعية سمحت باستخدامها حقباً طويلة من الزمن. إن مجرد مراقبة أشخاصٍ يؤدون مهمةً ما يقودنا إلى ملاحظة درجات المهارة المتفاوتة التي ينفذ بها هؤلاء الناس المهام المسندة إليهم وقدرة البعض منهم على إداء المهمة بنجاح في حال فشل الآخرين في الإرتقاء إلى المستوى المطلوب، ويقودنا هذا التصنيف بدوره إلى ترتيب مؤديّ المهمة من الأكثر فشلاً إلى الأكثر نجاحاً وبالتالي إلى وضع نظامٍ لتقييم اختبارات الأداء. كما يمكننا زيادة درجة دقة وتناغم عملية التقييم هذه من خلال إخضاع جميع المختَبرين إلى ظروف الاختبار ذاتها من حيث المواد المستخدمة والموقع والتوقيت ووجود أشخاص آخرين، وبهذه الطريقة يتم وضع امتحان يشكل المقياس أو المعيار الذي يتم على أساسه تقييم أداء الأفراد للمهمة التي يتألف منها الاختبار. ولعل أول إختبارغير رسمي في تاريخ البشرية هو "اختبار التفاحة" المذكور في سفر التكوين من الإنجيل. وفي هذا الاختبار لم يكن هناك إلا ممتَحَنين هما آدم وحواء إلا أن عواقب فشلهما كانت وخيمة. ويذكر الإنجيل

مثالاً آخر عن الامتحانات إذ يخاطب الله شعيباً في سفر القضاة ٧-٣:٧ بقوله:

وقال الرب لجدعون: إن لديك أناساً كثيرين لأدفع مديناً في قبضتهم وذلك حتى لا تفاخر عليَّ إسرائيل بأنها خلصت نفسها بقوتها، والآن أعلن للناس بأنه من كان خائفاً ومرتعداً فليرجع وينصرف من جبل جلعاد. فرجع من الشعب اثنان وعشرون ألفا وبقي عشرة آلاف.

وقال الرب لجدعون: لم يزل الشعب كثيراً انزل بهم إلى الماء فسأمحصهم لك هناك؛ فالذي أقول لك عنه يذهب معك فهو يذهب معك وكل من أقول لك عنه هذا لا يذهب معك، فهو لا يذهب.

فنزل بالناس إلى الماء وقال الرب لجدعون: قم بفصل كل من لصق الماء بلسانه كما يلعق الكلب عن أولئك الذين يجثون على ركبهم للشرب. وكان عدد الذين لعقوا بيدهم إلى فمهم ثلاث مئة رجل وأما باقي الناس فجثوا على ركبهم لشرب الماء.

قال الرب لجدعون بالثلاث مئة رجل الذين لعقوا بألسنتهم أخلصكم وأدفع المدينيين ليدك، وأما سائر الناس فليذهبوا كل إلى مكانه.

يبين هذا المقطع أن الله يعلم أن الجنود الذين لعقوا الماء كالكلاب هم أكثر حذراً وتأهباً ولذلك فإن عدة مئاتٍ من هؤلاء الجنود أكفأ بكثير من جحافل جنودٍ أقل حذراً وأقل تأهباً.

وتعود جذور الاختبارات النفسية الرسمية إلى الصين عام ٤٠٠٠ قبل الميلاد إذ وضع إمبراطور الصين آنذاك نظاماً امتحاناً يهدف إلى تحديد قدرة الأفراد على استلام مناصب في الدولة من خلال امتحانات شفهية تقام كل ٣ سنوات.

وبعد ٢٠٠٠ عاماً من ذلك الإجراء، قامت أسرة هان الحاكمة بإخضاع موظفي الدولة إلى إختبارات كتابية في القانون المدني والشؤون العسكرية والزراعة والمالية والجغرافية والإنشاء والشعر. وفي عام ١٣٧٠ ميلادية تم إتخاذ الخطوة الأخيرة في تطور الامتحانات الوظيفية في الصين عندما تم (وتبعاً للفلسفة الكونفوشية) تقسيم الامتحانات إلى ثلاثة أقسام تبدأ باختبارات تمهيدية ينتقل الناجحون فيها إلى اختبارات إقليمية ينتقل الناجحون فيها بدورهم إلى مرحلة الامتحانات على مستوى الدولة، وكانت نسبة الناجحين في امتحانات المرحلة الثالثة لا تتجاوز ١٪ من عدد المتقدمين للامتحانات التمهيدية. وقد استوحت عدد من الدول الأوروبية كبريطانيا وفرنسا وألمانيا امتحاناتها الوظيفية من النموذج الصيني خلال القرن التاسع عشر، إلا أن الصينيين أنفسهم تخلوا عن نظام امتحاناتهم مع بداية القرن العشرين.

الجدول ١-١

أحداث مختارة من تاريخ الاختبارات النفسية والتعليمية

٢٢٠٠ ق.م. استخدام الصينيين للامتحانات الوظيفية
١٢١٩ اعتماد الجامعات الأوروبية لنظام الامتحانات الرسمية
١٥٤٠ بدء الامتحانات الكتابية في الجامعات
١٥٩٩ نشر اليسوعيين لقوانين الامتحانات الكتابية
١٨٣٣ بدء نظام الامتحانات التنافسية لوظائف الدولة في بريطانيا
١٨٤٥ استخدام لجنة مدرسة بوسطن لامتحانات مطبوعة لأول مرة تحت إشراف التربوي هوراس مان
١٨٦٤ إعداد مدرس بريطاني يدعى جورج فيشر لعدد من المقاييس مؤلفة من نماذج أسئلة وأجوبة بغرض استخدامها كمقياس لتقييم إجابات الطلاب على أسئلة الامتحانات المقالية

١٨٦٦ نشر أ. إ. سيجوين لأول نص يتناول موضوع التقييم ومعالجة التخلف العقلي

١٨٦٩ بدء عملية دراسة الفروقات بين الأفراد بطريقة علمية بعد نشر كتاب تصنيف البشر وفقاً لملكاتهم الطبيعية لفرانسيس غالتون

١٨٧٠ بدء نظام امتحانات التوظيف التنافسية في الولايات المتحدة

١٨٨٤ افتتاح فرانسيس غالتون لمعرض المختبر الانثروبومتري[1] للصحة الدولية

١٨٨٨افتتاح ج. م. كاتل مختبراً للامتحان في جامعة بنسلفانيا

١٨٩٣عرض جوزيف جاسترو لاختبارات المهارات الحسية - الحركية في المعرض الكولومبي في شيكاغو

١٨٩٧نشر ج. م. رايس لنتائج أبحاثه حول قدرات التهجئة عند طلاب المدارس في الولايات المتحدة

١٩٠٤وضع تشارلز سبيرمان لنظرية ثنائية العامل في القدرة الذهنية

١٩٠٥نشر النسخة الأولى من مقياس بينيه - سيمون للذكاء

١٩٠٨نشر النسخة المراجعة من مقياس بينيه - سيمون للذكاء ونشر اختبارات للحساب من قبل ج. ك. ستون و س. أ. كورتيس

١٩٠٨-١٩١٤ تطوير إ. ل. ثورندايك لاختبارات قياسية في الحساب والخط واللغة والتهجئة

١٩١٥تعميم أول اختبارمنشور متعدد الإختيارات وهو اختبار كنساس للقراءة الصامتة. نشر اختبار ستنكويست للقدرة الميكانيكية

١٩١٦نشر لويس تيرمان لمقياس ستانفورد - بينيه للذكاء.

تطوير آرثر أوتيس لأول اختبار جماعي للذكاء مبني على مراجعة تيرمان ستانفورد لمقياس سيمون - بينيه

١٩١٧ وضع واستخدام اختباري الجيش آلفا وبيتا وهما أول إختبارين جماعيين للذكاء عند المنتسبين للجيش الأمريكي

١٩١٩نشر مقياس سيشور للموهبة الموسيقية

١٩٢٣ نشر الطبعة الأولى من اختبار ستانفورد للتحصيل

(١) فرع من علم الأجناس البشرية يبحث في مقاييس الجسم البشري، (المترجم).

١٩٢٤ نشر امتحان آيوا للمحتوى الخاص بالمدارس الثانوية

١٩٢٦ إجراء أول امتحان للاستعداد الدراسي (SAT)

١٩٢٧ نشر جداول جيسيل الإرتقائية

١٩٣٢ نشر اختبار متروبوليتان للتحصيل

١٩٣٨ نشر أ. ك. بيروس لأول كتاب سنوي للمقاييس العقلية.

نشر امتحان القدرات الذهنية للمرحلة الإبتدائية ومصفوفات ريفين التقدمية وامتحان بندر الحركي - البصري على مبدأ الغشتالت

١٩٣٩ نشر مقياس ويكسلر - بلفيو لذكاء الأطفال

١٩٤١ تطوير وتطبيق امتحان تصنيف الطلاب الضباط في سلاح الجو على طيارين متدربين في سلاح الجو الأمريكي

١٩٤٧ نشر بطارية اختبارات الاستعداد Aptitude Test Battery التباييني

١٩٤٩ نشر مقياس ويكسلرلذكاء الأطفال

١٩٦٨ نشر مقياس ويكسلرلذكاء الأطفال ما دون سن المدرسة وأطفال المرحلة الإبتدائية (WPPSI)

١٩٧٠- اليوم، استخدام الحاسب بشكل متزايد لتصميم وإجراء الامتحانات ثم وضع درجاتها وتحليل وتأويل نتائجها

١٩٧٤نشرالنسخة المراجعة لمقياس ويكسلر لذكاء الأطفال (WISC-R) وبطارية اختبارات هالستد - رايتن النفسي - العصبي

١٩٨٠- اليوم، تطوير نظرية الإستجابة لفقرة الاختبار (IRT)

١٩٨٣ نشر بطارية كوفمان لتقييم الأطفال (K-ABC)

١٩٨٥ نشركتاب معايير الاختبارات النفسية والتعليمية

١٩٨٦ نشر مقياس ستانفورد - بينيه للذكاء (الطبعة الرابعة) (SB-IV)

١٩٩٥ نشر الطبعة ١٢ من الكتاب السنوي للمقاييس العقلية

١٩٩٧ نشر الطبعة الثالثة من مقياس ويكسلر لذكاء البالغين (WAIS-III)

كانت الدول الأوروبية أبطأ من الصين في عملية إجراء امتحانات رسمية كان أولها امتحانات القانون في جامعة بولونيا التي بدأت عام ١٢١٩، وقد كانت معظم هذه الامتحانات شفهية ولم تصبح رسمية إلا في القرن السادس عشر. يوضح الجدول ١-١ بعض الأحداث الهامة في تاريخ الامتحانات النفسية والتعليمية، ولعل من أهم هذه الأحداث أول استخدام لامتحانات مطبوعة في مدارس بوسطن عام ١٨٤٥، ووضع معايير موضوعية لتقييم الإجابات على الأسئلة المقالية عام ١٨٦٤، والعمل الرائد الذي بدأه فرانسيس غالتون عام ١٨٦٩ لتقييم الذكاء البشري والقدرات العقلية.

قياس الذكاء

تعتبر أبحاث وكتابات فرانسيس غالتون عن الذكاء في النصف الثاني من القرن التاسع عشر بداية عصر الاختبارات النفسية المعاصرة. وقد كان غالتون على قناعة بأن الذكاء هو قدرة عامة موروثة يمكن قياسها عن طريق اختبارات حسية - حركية بسيطة لمكونات الذكاء المختلفة. ومن المهم هنا الإشارة إلى أن فهم أبحاث غالتون عن الذكاء يتطلب منا أن نتذكر أن غالتون بدأ أبحاثه في وقت كانت فيه عمليات البحث والتجربة المخبرية تتطور بشكل سريع في مجالات العلوم التطبيقية، إلا أن الفهم الفلسفي لعمل العقل البشري والبحث العلمي في عمل العقل والجهاز العصبي كانا بدائيين للغاية. وقد كانت الفلسفة الغالبة عن العقل آنذاك ترى أن العقل يتطور عن طريق تعقيد وتحليل الإنطباعات الحسية، مما يعني أن العقل يبدأ بالحواس ومما يفسرقرار غالتون بقياس الذكاء عن طريق اختبارات حسية - حركية.

وفي ضوء ذلك؛ فإنه ليس من المستغرب أن غالتون اختار لقياس الذكاء والتميز بشكل خاص وسائل من اختبارات قياس حسية - حركية. وكانت اختبارات غالتون تتألف من اختبارات للسلوك مبنية على التمييز الحسي

واختبارات لأوقات ردود الفعل تجاه مثيرات سمعية وبصرية واختبارات لسرعة الحركة والقوة العضلية والإحساس بالألم والدقة السمعية والبصرية وتمييز الأوزان. كما حصل غالتون على أعداد كبيرة من القياسات الخاصة بكل من الأطوال والأوزان وطول ومحيط الرأس والمسافة بين الذراعين وطول الأصبع الأوسط وطول القسم الأسفل من الذراع وحجم استيعاب الرئتين للهواء وغيرها من الخصائص الجسدية. وقد أفادت النتائج التي خلص إليها كلارك ويسلر في دراسة أجراها عام ١٩٠١ على ما أسماه ج. ماكين كاتل (١٨٩٠) "باختبارات غالتون العقلية" بعدم وجود أية علاقة وثيقة بين مقاييس غالتون بعضها مع بعض ومع نتائج الاختبارات الجامعية. وقد أدى ذلك إلى إيقاف كافة المحاولات لقياس الذكاء عن طريق اختبارات "الأدوات النحاسية" التي كانت متوفرة في المختبرات النفسية آنذاك.

وعلى الرغم من أن فشل اختبارات غالتون وكاتل ثبط همم الباحثين في مجال اختبارات الذكاء البسيطة والنظريات المصاحبة إلا أن عدداً من الأبحاث المعاصرة (جونسون ١٩٨٢a، ١٩٨٢b) اقترح بأن التراجع عن البحث في الذكاء في ذلك الوقت فعلاً كان متسرعاً. فمثلاً؛ تعد علاقت الربط التي تقدر بحوالي ٤٠، والتي تربط كلاً من وقت رد الفعل المعقد والدرجات في امتحانات الذكاء علاقة ممكنة.

من جهة أخرى، لم يمض وقت طويل قبل ولادة اتجاه عملي لتقييم الذكاء، ففي عام ١٩٠٥ وضع الأخصائي النفسي الفرنسي ألفرد بينيه ومساعده ثيوفيل سيمون اختباراً للذكاء مبني على المهام التي يطلب من الأطفال في سن المدرسة أداؤها. ويطلق على هذا النوع من الاختبارات اليوم اسم "نسخة عن العمل" Job replica يتضمن الاختبار عينة من المهام التي يطلب عادة من الممتحن تأديتها. إن المهام الثلاثين التي يتضمنها مقياس بينيه - سيمون لاختبار الذكاء هذا قد تم رصدها في الجدول رقم ١-٢.

الجدول ١-٢

الاختبارات الثلاثون في مقياس بينيه - سايمون الأول للذكاء

١ . التناسق البصري

٢ . الإدراك الحسي

٣ . الإدراك البصري

٤ . إدراك الطعام

٥ . البحث عن الطعام ضمن ظروف غير موائمة

٦ . تنفيذ الأوامر البسيطة

٧ . التعرف على الأشياء بأسمائها

٨ . التعرف على الأشياء المصورة بأسمائها

٩ . تسمية الأشياء المصورة

١٠ . مقارنة خطوط الأشياء المختلفة على مستوى فوق عتبة الشعور

١١ . الذاكرة السمعية لأعداد مؤلفة من ثلاثة أرقام

١٢ . مقارنة أوزان الأشياء المختلفة على مستوى فوق عتبة الشعور

١٣ . التأثر بالآخرين

١٤ . تعريف الأشياء المألوفة

١٥ . تذكر الجمل

١٦ . تحديد الإختلافات بين الأشياء المألوفة الموجودة في الذاكرة

١٧ . تذكر صور الأشياء

١٨ . الرسم من الذاكرة

١٩ . الذاكرة السمعية لأعداد مؤلفة من أكثر من ثلاثة أرقام

٢٠ . تحديد التشابهات بين الأشياء المألوفة الموجودة في الذاكرة

٢١ . تمييز الخطوط

٢٢ . ترتيب خمسة أوزان

٢٣ . تحري الوزن المفقود

٢٤ . الأنغام

٢٥ . الكلمات المفقودة

٢٦ . تركيب الجمل

٢٧ . الإجابة على الأسئلة الإشكالية

٢٨ . حركة عقارب الساعة المتبادلة

٢٩ . الرسم من تصميم تم قصه من ورقة مطوية أربع طيات

٣٠ . التمييز بين المفاهيم المجردة

تجسد المهام المذكورة أعلاه فرضية "بينيه" بأن الذكاء يتألف من إمكانيات إدراكية متعددة تتغير مع تقدم الإنسان في العمر ويؤكد اختبار بينيه أيضاً على إيمانه بأن اختبارات الذكاء يجب أن تؤكد على القدرات الذهنية العليا. وعلى الرغم من أن معظم إختبارات بينيه تتألف من مهام مناسبة لطلاب المدارس (تنفيذ الأوامر البسيطة، المعرفة اللغوية، التعاريف، التعرف على الصور وإتمام الجمل) إلا أن بينيه استبعد كافة مقاييس المعرفة الإكاديمية من اختباره وركز على مقاييس الذكاء حصراً وذلك لأنه لم يربط الذكاء بأي علم يكتسب عن طريق الدراسة بل ربطه بالقدرة على الحكم والإستدلال والاستيعاب.

طبق بينيه وسيمون اختبار ذكائهم على عينة من الأطفال غير المتخلفين عقلياً أعمارهم ٣ و٥ و٧ و٩ و١١ عاماً بالإضافة إلى عينة من الأطفال والبالغين المتخلفين عقلياً. تم بعد ذلك ترتيب المهام تبعاً للعمر الذي يفترض فيه أن ينجز الطفل المهمة (مقياس العتبة العمرية)، ثم تم إجراء الاختبارات بدءاً بالمهمة ذات المقياس العمري الأدنى (المهمة ١) وانتهاءاً بالمهمة ذات المقياس العمري الأعلى (المهمة ٣٠). واستند بينيه وسيمون إلى النظام الفرنسي لتصنيف المتخلفين عقلياً

والذي يضم ثلاث فئات: الأغبياء والبلهاء والحمقى. وحدد بينيه وسيمون المهمة ٦ كالحد الأعلى للبالغين الأغبياء والمهمة ١٦ كالحد الأعلى للبالغين البلهاء والمهمة ٢٣ كالحد الإعلى للبالغين الحمقى.

نشرت نسخ مراجعة عن مقياس بينيه - سيمون عامي ١٩٠٨ و١٩١١ وحدثت تطورات أخرى على اختبارات الذكاء في بريطانيا والولايات المتحدة. وفي نسختي ١٩٠٨ و ١٩١١ من اختبار بينيه - سيمون تم تجميع المهام وفقاً للعمر المناسب لأدائها كما تمت إضافة مهام فرعية في حال نجاح نسبة من الأطفال تتراوح بين ٦٧٪ و ٧٥٪ في تجاوز هذه المهمة. هذه النسبة هي لأطفال من شريحة عمرية واحدة مع الأخذ بعين الاعتبار وجود نسبة قليلة من الأطفال الأصغر ونسبةعالية من الأطفال الأكبر يمكنها اجتياز هذه المهمة بنجاح. كان الاختبار يبدأ بتحديد العمر القاعدي Basal age للممتحن (وهو أعلى مستوى عمري للمهام الفرعية تمكن من تجاوزه بنجاح) ومن ثم تحديد العمر السقفي Ceiling age (وهو أدنى عمر فشل الممتحن في كافة مهامه الفرعية). ويحدد عمر الممتحن العقلي بربط كل مهمة فرعية ينجح فيها بعدد من الأشهر بين العمر القاعدي والعمر السقفي ثم يضاف الناتج إلى العمر القاعدي. وأخيراً، كانت تتم مقارنة العمر الناتج عن الاختبار بعمر الممتحن الحقيقي لتحديد ما إذا كانت قدراته العقلية أعلى أو أدنى من المعدل.

يعد مقياس ستانفورد - بينيه الذي نشره لويس تيرمان عام ١٩١٦ من أهم التعديلات على مقياس سيمون - بينيه والذي لم يكتف فيه تيرمان بمقارنة العمر العقلي بالعمر الفعلي للممتحن بل أدخل مبدأ الذكاء النسبي (ratio IQ) والذي يمكن تعريفه بنسبة العمر العقلي (بالأشهر) إلى العمر الحقيقي (بالأشهر) مضروباً بـ ١٠٠. وعليه فإن الأطفال ذوي معدلات الذكاء الطبيعية كانت معدلات ذكائهم ١٠٠، بينما تعدت معدلات ذكاء الأطفال الموهوبين الرقم ١٠٠ بشكل

ملحوظ وتدنت معدلات ذكاء الأطفال المتخلفين عقلياً دون ١٠٠ بشكل ملحوظ أيضاً.

أطلق اسم "الخبز والسمن لعلم النفس" على اختبارات الذكاء لأنها أتاحت الفرصة للعديد من علماء النفس لكسب المال بعيداً عن المختبرات والمكتبات ومدرجات الجامعات. فكثيراً ما كان علماء النفس هؤلاء يُرون وهم يتنقلون من مدرسة إلى أخرى ليختبروا ذكاء الأطفال قبل ضمهم إلى فصول تعليمية خاصة، كما كانوا يُشاهدون في عيادات إرشاد الأطفال حيث كانوا يحكمون ما إذا كان لمعدل ذكاء هؤلاء الأطفال المتدني دور في مشاكلهم السلوكية.

إن اختبار ستانفورد - بينيه اختباراً فردياً والذي خلافاً لاختبار المجموعات يعطى لكل فرد على حدة في وقت واحد، تم وخلال أقل من سنة من تاريخ نشر الاختبار تطوير اختبارين جماعيين للذكاء هما اختبارا الجيش ألفا وبيتا[٢]، وقد كان التركيز الرئيسي لاختبار ألفا هو على القدرة اللغوية إذ كان يتألف من: اختبار القدرة على اتباع الأوامر الشفهية واختبارالإستدلال الحسابي وإتخاذ القرارات العملية ومزاوجة الأضداد والمترادفات وترتيب الجمل المبعثرة وإكمال متسلسلات الأرقام والمتشابهات والمعلومات. وتجدر الإشارة هنا إلى أن الكثير من اختبارات القدرة الأكاديمية اليوم تحتوي على عناصر مشابهة. وعلى عكس اختبار ألفا، كان اختبار بيتا اختباراً لالغوياً يعطى للجنود الأميين أو غير الناطقين باللغة الإنجليزية وكان يتألف من عناصر بصرية - إدراكية حسية ومهارات حركية وذلك كمن يتتبع طريقاً الخروج في متاهة أو تصور عدد المباني

(٢) تم تطوير أول اختبار جماعي للذكاء من قبل بايل (١٩١٣). وكان اختبار بايل مؤلفاً من مقاييس للذاكرة المنطقية وذاكرة الحفظ عن طريق التكرار والقدرة على التعويض و بناء الكلمات وتداعي الأفكار وكان مرفقاً مع الاختبار معدلات للأداء تم الحصول عليها من عينات من طلاب المدارس. راجع بينتر الاختبار عام ١٩١٧ وأصبح اسمه اختبار بينينتر-باترسون لمقاييس الأداء.

في رسم ثلاثي الأبعاد. وقد خضع أكثر من مليوني مجند لاختباري ألفا وبيتا بغرض تصنيفهم وتوزيعهم على المهام العسكرية المختلفة خلال الحرب العالمية الأولى وبعدها مباشرة (يركيس، ١٩١٩). وفي العشرينيات، خضع الملايين من طلاب المدارس في الولايات المتحدة لاختبار الذكاء الوطني وغيره من الاختبارات الجماعية المقتبسة عن اختباري الجيش ألفا وبيتا، كما استخدم اختبارانهما الاختبار العام لتصنيف الجيش واختبار التأهيل للقوات المسلحة في سنواتٍ لاحقة ويعد هذان الاختباران منحدرين من سلالة الاختبارين ألفا وبيتا.

لم يتم استخدام البيانات الناتجة عن اختباري ألفا وبيتا في رفع كفاءة العمليات الحربية لأسباب متعددة من أهمهَا انعدام الثقة لدى السلطات العسكرية في الأبحاث العلمية، إلا أن نتائج هذه الاختبارات وغيرها التي أجريت على مجموعات من المهاجرين إلى الولايات المتحدة تم تفسيرها بشكلٍ خاطئ لإثبات فكرة الدونية الفكرية لمواطني أوروبا الشرقية بالمقارنة مع المهاجرين من أوروبا الغربية والشمالية (بريغهام، ١٩٢٣؛ يركيس، ١٩٢١). وقد أدت أدلة تم الحصول عليها لاحقاً إلى تراجع بعض علماء النفس (مثل بريغهام، ١٩٣٠) عن تصريحاتهم حول الجنسيات وفروقات الذكاء. وبشكلٍ عام فإن نتائج الدراسات الأولى واللاحقة أشارت إلى تباين نتائج الأفراد في اختبارات الذكاء الأمريكية تبعاً لتشابه حضارتهم الأم مع الحضارة الأمريكية المسيطرة (أو إختلافها عنها).

بالإضافة إلى الاختبارات ذات التركيز اللغوي مثل اختبار ستانفورد - بينيه وغيره من اختبارات الذكاء الجماعية، تم في العقدين الأولين من القرن العشرين اعداد الكثير من الاختبارات الفردية ومقاييس الأداء وبطاريات الاختباربهدف تحديد مستوى القدرات العقلية للأفراد. ولا تزال العديد من هذه الاختبارات متوفرة حتى اليوم مثل لوائح سيغوين ومكعبات نوكس وتصاميم كوهس المكعبة

ومتاهات بورتيوس. وتتطلب اختبارات الأداء هذه، على عكس غيرها من الاختبارات التي تعتمد على القدرات اللغوية الشفهية والكتابية، التعامل مع الأجسام الفيزيائية كالمكعبات والمتاهات والخرز والصور المفككة (puzzles).

الاختبارات التحصيلية

لم تكن أدوات قياس الذكاء هي الاختبارات الوحيدة التي تم نشرها في بداية القرن العشرين، ففي عام ١٩١٠ تم تداول نماذج لاختبارات قياسية في الحساب والخط واللغة والتهجئة وغيرها من المواد الدراسية المدرسية. وفي عام ١٩٢٣ نشر ت. ل. كيلي الطبعة الأولى من اختبار ستانفورد للتحصيل، وفي عام ١٩٢٤ نشر إ. ف. ليندكويست امتحان آيوا للمحتوى الخاص بالمدارس الثانوية بينما نشر مجلس التعليم العالي (College Board) أول طبعة من اختبار الاستعداد المدرسي (SAT) ذو أسئلة الاختبارات المتعددة. وقد ساهم نمط الاختبارات القياسية ذات أسئلة الاختبارات المتعددة في تطوير العديد من اختبارات التحصيل القياسية خلال العشرينيات والثلاثينيات من القرن الماضي، كما ساعد استخدام الآلة في تصحيح الاختبارات في الثلاثينيات في تطور هذا النوع من الاختبارات بشكل كبير. ومما زاد من أهمية هذه الاختبارات تجميع معايير الاختبارات التي أجريت على عينات مختلفة تمثل المجتمع من طلاب المدارس ثم إجراء أبحاث للتأكد من صدق وثبات هذه الاختبارات بالإضافة إلى تحليل فقراتها المختلفة.

اختبارات الاستعداد

شاعت في العشرينيات والثلاثينيات، بالإضافة إلى اختبارات الذكاء والتحصيل، اختبارات تقيس الاستعداد الفردي في مجالات التوظيف وانتقاء الطلاب المتقدمين إلى المدارس والكليات المختلفة. وعلى عكس اختبارات الذكاء التي كانت مرتبطة بالمدارس فقط، تم تطوير اختبارات الاستعداد بغرض اختيار

وتصنيف الموظفين في مجالات الأعمال والصناعة والجيش. كما اثبتت هذه الاختبارات جدواها في المجالات المهنية، ففي كتابه الرائد: كيف تختار مهنتك، اقترح فرانك بارسون نموذجاً لمطابقة الأشخاص مع العمل المناسب لهم بناءً على مواصفات هؤلاء الأشخاص ومتطلبات الوظيفة. وقد كان بارسون على قناعة أن اختيار المهن وتوزيع القوة العاملة عليها يمكن أن يصبح أكثر فاعلية إذا ما أتيحت الفرصة للأفراد لمعرفة إمكاناتهم وقابلياتهم واهتماماتهم بالإضافة إلى معرفة متطلبات المهنة أو الوظيفة. وقد تلقت حركة بارسون للإرشاد المهني التي انطلقت عام ١٩٠٩ دعماً كبيراً من التطورات التي طرأت على اختبارات الاستعداد واختبارات الميول واختبارات الشخصية في العشرينيات والثلاثينيات. كما لعب نشر المعجم الأول للعناوين المهنية (DOT) عام ١٩٣٩ دوراً هاماً في دعم حركة بارسون أيضاً إذ يعتبر هذا المعجم الذي كان يعدل دورياً منذ تاريخ نشره أول محاولة حقيقية لتوفير نظام شامل لتصنيف المهن وتفصيل مقتضيات الوظائف المختلفة.

تضمنت اختبارات الاستعداد الأولى تلك التي عملت على قياس القدرات النفسية - الحركية والقابلية الميكانيكية والكتابية (السكريتارية) إلا أن اختبارات الاستعداد الفني والموسيقي لم يتم إعدادها حتى الثلاثينيات. وكان كارل سيشور قد نشر مقياس الموهبة الموسيقية عام ١٩١٨ في حين طور جونسون أوكونور في العشرينيات مقياساً للمهارة الحركية وطور د. ج. باترسون وزملاؤه في جامعة مينيسوتا في أواخر العشرينيات وبداية الثلاثينيات اختبارات للقدرة الميكانيكية والمكتبية (السكريتارية). وقد استخدمت جميع هذه الاختبارات وغيرها في تقديم الإرشاد المهني وانتقاء الموظفين لعقدين كاملين.

بالإضافة إلى اختبارات الاستعداد الواحد، تم نشر بطارية اختبارات الاستعداد العام واختبارات الاستعداد التباييني وهي اختبارات لا تقيس

استعداداً معيناً فحسب بل عدداً من الاستعدادات في آنٍ واحد. وقد اثبتت هذه الاختبارات جدواها في اختيار الموظفين للمهام اللاإدارية كالكتاب وعمال خطوط التجميع في المصانع وعمال الصيانة. كما اثبتت اختبارات الاستعداد تفوقها على أدوات تقييم الأداء أثناء العمل في توقع نجاح الأفراد في البرامج التدريبية المختلفة. ومن أفضل المقاييس الخاصة بالاستعداد تلك التي تقيس السرعة والدقة والتي تستخدم في اختيار موظفين للأعمال المكتبية وفي تقييم الإدراك الميكانيكي والعلاقات الفراغية عند اختيارعمال خطوط الإنتاج وعمال الصيانة. كما ساهمت اختبارات الذكاء العامة في اختيار العاملين في مجالات مختلفة إذ أكد فيليب فيرنون (١٩٦٠) على أن الذكاء العام لا القابليات أو الاستعدادات الخاصة هو الذي يحدد النجاح المهني.

أحداث تاريخية أخرى

ومن الجدير بالذكر وجود أحداث أخرى هامة في تاريخ الاختبار النفسي والتعليمي كنشر مراجعات لمقياس ستانفورد - بينيه للذكاء عامي ١٩٣٧ و١٩٦٠ ونشر مقاييس ويكسلر للذكاء (للأطفال والبالغين والأطفال دون سن المدرسة) ونشر مقاييس بايلي لتطور الرضع. ومن الأحداث الهامةفي تاريخ اختبارات التحصيل ظهور العديد من الطبعات المعدلة عن الاختبارات الموجودة وظهور اختبارات جديدة وتطويرعمليات التقنين ونمو مؤسستي "خدمة الاختبارات التعليمية" و"برنامج الاختبار للتعليم العالي في الولايات المتحدة" وغيرها من المؤسسات المهنية للاختبارات. ومن أهم نتائج البحث في هذا المجال التوصل إلى حقيقة أن اختبارات الاستعداد غير منحازة ضد الأقليات ولها درجة عالية من الصدق في توقع النجاح المهني للأفراد في مواقع متعددة (شميدت وهنتر، ١٩٧٧، ١٩٧٨، ١٩٨١).

ويجب التأكيد كذلك على التقدم في مجالي نظرية وطرق الاختبار النفسي والتعليمي، فمع نهاية العشرينيات كان تشارلز سبيرمان و ترومان كيللي وغيرهما من علماء النفس قد وضعوا أساسيات نظرية الاختبار الكلاسيكية. وقد تزامنت هذه التطورات مع مساهمات ومساعدات إحصائية في مجالي النظرية وطرق البحث قدمها كارل بيرسون ور. أ. فيشر وغيرهما. لقد كان للتطورات في تحليل العوامل وتحليل البنى الكامنة وتحليل المسارات ونظرية الإستجابة للفقرة وغيرها من التطورات والاكتشافات في مجال الاختبارات دور كبير في تطوير وتوسيع نظرية الاختبار النفسي وطرق البحث فيها خلال القرن العشرين. وبدورها فقد ساهمت ثورة التكنولوجيا وتقنية الكمبيوتر ليس في تحليل الامتحانات وتدوين نتائجها فحسب بل في إجراء هذه الامتحانات وتأويل معطياتها مما أخذ مساهمات الاختبارات النفسية في مجالي العلوم والتطبيق إلى أبعاد لم يتخيلها أكثر المؤيدين لهذه الاختبارات. وقد كان للحاسب أثر مشابه لأثر الإحصاء على عالم الاختبارات في النصف الأول من القرن العشرين.

تصنيف الاختبارات ومصادر المعلومات

كما أشرنا سابقاً فإن أدوات التقييم النفسي يمكن أن تصنف تبعاً إلى ما تهدف هذه الأدوات إلى قياسه كالقدرات الإدراكية أو العاطفية اللاإدراكية أو النفسية - الحركية، ويمكن تقسيم كلٍ من هذه الفئات إلى فئات فرعية أخرى إذ يمكن للاختبارات الإدراكية أن تقيس الذكاء والتحصيل والقدرات الخاصة بينما تقيس اختبارات الشخصية الطباع والصفات الشخصية وغيرها من الخصائص العاطفية. كما يمكن وصف اختبارات الذكاء والاستعدادات الخاصة بأنها اختبارات لطرق تفكير تجميعية وتفكيكية (إبداعية) أو اختبارات قدرة أكاديمية (مدرسية) أو اختبارات غير منحازة لحضارةٍ ما على حساب أخرى. ويمكن

تقسيم هذه الاختبارات بناءً على الأهداف التي تخدمها إلى اختبارات اختيار للمتقدمين وتحديد الأدوار(اختبارات التنسيب) واختبارات لتصنيف وتقييم الأداء واختبارات للجاهزية وأدوات للتحليل والتنبؤ. ويمكن لهذه الاختبارات أن تكون قياسية مما يقتضي استخدام معايير Norms ثابتة لتحليل النتائج أو غير قياسية (يضعها المدرسون). وهنالك أيضاً اختبارات للسلوك والإرتقاء واختبارات عصبية - نفسية واختبارات حركية - حسية واختبارات للسمع والنطق واختبارات تكييفية واختبارات المتشابهات واختبارات الفقرات الأساسية المتكررة واختبارات تسلسلية وغيرها. ويمكن تصنيف الاختبارات وفقاً للتوقيت الزمني الذي يتم فيه الاختبار ضمن العملية التعليمية أو التدريبية أو الشفائية أو العلاجية ويطلق على عمليات التقييم التي تفيد في تقييم التطور الحاصل أثناء تطبيق برنامجٍ ما عمليات التقدير التكويني (Formative Assessment) بينما يطلق على العمليات التي تقييم النواتج النهائية عمليات التقدير التراكمي (Summative Assessment).

طرق أخرى لتصنيف الاختبارات

يمكن تصنيف الاختبارات تبعاً لطرق وضع هذه الاختبارات والشكل الذي تتخذه وطرق إجرائها وطرق الإجابة على أسئلتها وطرق وضع درجاتها وطرق تحليل الدرجات أو تقييم الأداء.

تتنوع طرق وضع الاختبارات بتنوع الأهداف المراد تحقيقها منها ولكنها غالباً ما تنطلق من مخطط مفصل أو جدول للمواصفات يحدد محتوى الاختبار، ثم يتم إتخاذ قرارات متعلقة بطبيعة الأسئلة التي سيتضمنها الاختبار: أسئلة موضوعية (اختبارات متعددة أو أسئلة الخطأ والصواب أوأسئلة الإجابات القصيرة أو أسئلة المطابقة) أو أسئلة ذات إجابات مفتوحة (مقالية) أو أسئلة

تطبيق عملية. ويمكن ترتيب هذه الأسئلة ضمن عدد من الاختبارات الفرعية أو على شكل كتيب لولبي (أي أسئلة مرتبة بالتناوب حسب نوعها ومن الأسهل إلى الأصعب). وتتطلب الإجابة على أسئلة الاختبارات المتعددة قدرات معرفية واستيعابية وتطبيقية وتحليلية وتجميعية وتقييمية وغيرها من العمليات الإدراكية (بلووم وكاثول، ١٩٥٦).

ويمكن لأسئلة الاختبارات أن تكون بصرية (مطبوعة ضمن كتيب أو على بطاقات أو على شاشة الكمبيوتر) أو سمعية (يقرؤها المراقب الموجود في قاعة الاختبار أو تقدم على شريط مسجل). كما تطبع تعليمات الاختبار ضمن كتيب الاختبار أو على شاشة الحاسب أو تقرأ من قبل المراقب الموجود في القاعة أو يشار إليها بالإيماء. وقد تكون أسئلة الاختبارات في حد ذاتها أسئلة سهلة ولكنها عديدة لا يتمكن معظم الممتحنين من إكمالها ضمن الوقت المحدد للاختبار (اختبار سرعة) أو يكون الوقت المحدد للاختبار طويلاً جداً إلا أن الأسئلة صعبة(اختبار قوة).

يضع الممتَحَنون (ووفقاً لشكل الاختبار وأغراضه) إشارة أمام الإجابة الصحيحة على ورقة الأسئلة أو على ورقة إجابة منفصلة أويطبعون الإجابات مستخدمين لوحة مفاتيح الكمبيوتر أو يجيبون على الأسئلة شفهياً أو بالإشارة أو بالتطبيق أو بغيرها من الطرق التي تسمح لهم بالإجابة. ويمكن للإجابات أن تكون محددة بعدة خيارات (أسئلة الاختيار من المتعدد) أو مفتوحة كالأسئلة المقالية أو أسئلة المشاريع التطبيقية.

وتعتمد عملية منح الدرجات في الاختبارات المختلفة على كون الاختبار موضوعياً أو ذاتياً ففي الاختبارات الموضوعية هنالك عدد محدد من الإجابات ومفتاح للأجابات يمكن مقارنة جميع الإجابات به. وفي حال عدم وجود أي أخطاء مكتبية تبقى درجات الممتحن ثابتة حتى مع اختلاف واضعي الدرجات. أما في الاختبارات ذات الطابع الذاتي كتلك التي تتضمن أسئلة مقالية، تختلف

الدرجات المعطاة للممتَحَن باختلاف المصحح[٣] على الرغم من وجود تعليمات وإرشادات عامة تنظم عملية وضع الدرجات. وينطبق هذا الكلام على تقييم المشاريع والعروض التجريبية وغيرها. وبغض النظر عن أشكال الاختبارات أو أسئلتها أو عن كون الاختبارات موضوعية أو ذاتية، غالباً ما تعتمد درجات الاختبارات على معادلة ١ - صفر (أي تمنح درجة صفر للإجابة الخاطئة وعدد من الدرجات وغالباً ١ للإجابات الصحيحة)، كما يمكن أن توزع الدرجات على الأسئلة وفقاً لصعوبة أو أهمية السؤال. وفي بعض الاختبارات المقالية والمشاريع والتجارب العملية تعطى درجات للدقة ومحتوى الإجابة ولشكل الإجابة أيضاً (القواعد، الخط، الكفاءة، سلاسة الأسلوب... إلخ).

ومن الجدير بالذكر أن عملية وضع الدرجات ليست هي نهاية المطاف في عملية الاختبار إذ يجب تقييم العناصر التي منحت درجات والعناصر التي لم تمنح أية درجات في كل اختبار وتحليلها وفق معايير معينة (ذاتية أو موضوعية). وفي معظم الاختبارات المدرسية يقوم المدرس بوضع الدرجات وتقييم نتائج الاختبارات وتحليل معطياتها دون الإستعانة بأي شخصٍ آخر. فالمدرس إذاً هو واضع الدرجات وهو الذي يحدد التقدير النهائي للطالب تبعاً لمقاييس تجربته الخاصة، وقد يكتب هذا المدرس تعليقاً مفيداً على أداء الطالب في ذلك الاختبار أو لا يكتب. أما في الاختبارات التحصيلية الرسمية، يفرق المختصون بين الاختبارات التي تمنح الدرجات بناءً على مرجعية معيارية (Norm-referenced Assessment) والاختبارات التي تمنح الدرجات بناءً على مرجعية المحك (Criterion-referenced Assessment). ونظراً لأن الاختبارات

(٣) استخدمت كلمة مصحح هنا على الرغم من أن واضع الدرجات هي المصطلح الأنسب إذ أن الذي يضع الدرجات لا يصحح الإجابات بل يقييمها فقط. (المترجم).

معيارية المرجع يتم قياسها من خلال إجراء الاختبار على عينة من الأشخاص تمثل المجتمع، يتم تفسير (تحليل) درجات هذه الاختبارات بمقارنتها مع المعايير الناتجة عن عينة القياس. أما درجات الاختبارات المحكية المرجع (والتي تسمى أيضاً اختبارات الإتقان Masterytests)، فيتم تفسيرها عن طريق مقارنتها بأهداف أو مستويات محددة مسبقاً للأداء المقبول. ويقوم بتفسير نتائج الاختبارات عادةً أخصائيون في مادة الاختبار. كما يتم تفسير نتائج العديد من الاختبارات النفسية واختبارات القدرة والشخصية بواسطة الحاسب، ويحتوي برنامج تحليل الاختبارات بواسطة الحاسب (CBTI) على عدد من الجمل التي تصف إجابات معينة على الأسئلة أو الدرجات النهائية للاختبار بكامله، ومن ثم تتم طباعة هذه الجمل في وثيقة توضح أداء الفرد في الاختبار.

مصادر للمعلومات عن الاختبارات

لا يوجد مصدر واحد يحتوي على قائمة ووصف لكل الاختبارات النفسية والتعليمية، إلا أن أشمل قوائم الاختبارات المطبوعة موجودة في كتاب الاختبارات المطبوعة ٤ (مورفي، كونولي، وإيمارا، ١٩٩١) وكتاب اختبارات (سويتلاند وكيسير، ١٩٩١). ويحتوي برنامج F-8 على قوائم خاصة بالاختبارات التحصيلية بينما يحتوي برنامج H-7 على قوائم باختبارات الاستعدادات العامة والخاصة. وتعد الكتب الثلاثة التالية من أهم المراجع الخاصة بالاختبارات غير المنشورة: دليل المقاييس العقلية غير المنشورة (غولدمان وميتشل ١٩٩٠، غولدمان وأوزبورن ١٩٨٥)، دليل المستهلك للاختبارات المنشورة (هاموميل، براون وبراينت، ١٩٩٢)، وفهرس الاختبارات المستخدمة في أطروحات الدراسات العليا (فابينو، ١٩٨٩). كما يمكن الإستفادة من قواعد البيانات الالكترونية مثل خدمة تجميع

الاختبارات التابعة لـ (ETS مجموعة الاختبارات وخدمة الاختبارات التعليمية في برينستون نيوجرسي) وقاعدة بيانات PsycINFO.

ولقراءة مقالات نقدية عن الاختبارات المختلفة يمكن العودة إلى كتب المقاييس العقلية السنوية بمجلداتها المختلفة (كونولي وإمبارا، ١٩٩٥، كونولي وكرامر، ١٩٨٩، وكرامر وكونولي، ١٩٩٢) وكتاب مقالات نقدية عن الاختبارات (كيسير وسويتلاند، ١٩٨٤-١٩٩٤) وكتاب الاختبارات في مجال التعليم: آراء النقاد والمقيمين (ليفي وغولدشتاين، ١٩٨٤) وكتاب اختبار الأطفال (ويفر، ١٩٨٤) وكتاب اختبار المراهقين (هارينفتون، ١٩٨٦) وكتاب اختبار البالغين (سفايرسنسكي، ١٩٨٥). وتحتوي الكتب التالية على قوائم ووصف لمقالات المجلات التخصصية المتعلقة بالاختبارات وغيرها من الفصول ضمن الكتب بشكل عام والكتب المعنية بالاختبارات بشكل خاص: ملخصات أطروحات الدراسات العليا في مجال علم النفس، فهرس التعليم، الفهرس الحالي للمجلات التخصصية التعليمية. كما تحتوي بعض المجلات التخصصية في علم النفس والتعليم والأعمال وغيرها على مقالات تتعلق بالاختبارات وعمليات الاختبار. وتنشر المجلات التخصصية التالية مقالات عن المهارات الإدراكية والحركية - النفسية: المجلة الأمريكية للأبحاث التعليمية ومجلة المقاييس النفسية التطبيقية ومجلة المقاييس النفسية والتعليمية ومجلة علم النفس السريري ومجلة علم النفس الاستشاري ومجلة المقاييس التعليمية ومجلة علم النفس المدرسي ومجلة السلوك المهني ومجلة القياس والتقييم في الإرشاد والتطوير ومجلة علم نفس الموارد البشرية ومجلة التقييم النفسي - التعليمي ومجلة التقييم النفسي ومجلة علم النفس في المدارس.

وتتوفر قوائم لبرامج الحاسب التي تستخدم في مجال الاختبار في كتاب استخدام الحاسب في علم النفس: دليل البرامج (ستولوف وكاوتش، ١٩٩٢)

وكتاب مصادر برامج الكمبيوتر النفسية (كروغ، ١٩٩٣)، ويحتوي الملحق (ج) من هذا الكتاب على قائمة بعناوين الشركات التي تسوق برامج الحاسب المستخدمة في عمليات الاختبار التعليمي والنفسي.

كما توفر القوائم التي تنشرها الشركات المتخصصة في إنتاج مواد الاختبارات مصدراً وإن كان محدوداً للمعلومات المتعلقة ببعض الاختبارات الخاصة. ويحتوي الملحق (ب) من هذا الكتاب على قائمة بعناوين وأرقام هواتف دور نشر وتوزيع مواد التقييم التعليمية والنفسية، وتتوفر معلومات تفصيلية عن معظم الاختبارات المنشورة في أدلة المستخدم التي ترفقها دور النشر عادةًبالاختبارات وأدوات القياس المختلفة.

المقاييس والاعتبارات الأخلاقية والقانونية

ساهمت عدة عوامل خلال العقود القليلة الماضية في زيادة أهمية أدوات القياس النفسية والتعليمية وفي توطيد الأسس العلمية لعملية الاختبار، ولعل أهم هذه العوامل هو تناول موضوع القياس النفسي والتعليمي في الآلاف من الكتب والمقالات والفصول ضمن الكتب والمئات من المحاضرات التي ألقيت في مؤتمرات إقليمية ودولية. وقد ساهم التأكيد على أخلاقيات ومعايير الاختبار من قبل هيئات رسمية كالإتحاد الأمريكي لعلم النفس والإتحاد الأمريكي للبحوث التعليمية والإتحاد الأمريكي للإرشاد وشؤون الموظفين في ضمان تطور ونمو مجال الاختبارات التعليمية والنفسية بطريقة أكثر انتظاماً وأقل فوضويةً من السابق. كما يجري حالياً تطويرعدد من الإجراءات وأدوات القياس السايكومترية التي لا تستهدف الربح المادي فحسب بل تحقق درجات عالية من الصدق والثبات ويتم تطبيقها بشكلٍ يضمن مصالح الناس.

تُجرى الاختبارات عادةً لعدة أغراض في المدارس والجامعات والعيادات والمحاكم ومجالات الأعمال والصناعة والدولة والجيش وغيرها من المجالات،

وتوفر معدلات الدرجات النهائية الناتجة عن هذه الاختبارات معلومات هامة تساعد في إتخاذ قرارات بشأن أشخاص وبرامج مختلفة (تعليمية أو غيرها). ويمكن تلخيص استخدامات نتائج الاختبارات النفسية والتعليمية في النقاط التالية:

- تصنيف وتوزيع واختيار الأفراد في المؤسسات التعليمية والمهنية.
- فَصْل أو ترقية أو الإبقاء على الطلاب في المدارس والموظفين في المؤسسات.
- منح الحرفيين أو أصحاب المهن تراخيص بمزاولة مهنهم.
- تشخيص المشكلات النفسية أو التعليمية والتعرف على احتياجات المصابين بغرض توفير المساعدة لهم عن طريق جلسات الإرشاد أوالعلاج أوغيرها من الإجراءات.
- تقييم فعالية التدريب أو العلاج أو غيرها.
- توفير أدوات لقياس متغيرات السيطرة والمتغيرات المستقلة وغير المستقلةللأبحاث في تأثير الإجراءات العلاجية على التغيرات السلوكية.
- توفير معلومات للأفراد تساعدهم على معرفة أنفسهم بشكلٍ أفضل.

وتعتمد درجة الوثوق في تحقيق أهداف الاختبارات المذكورة أعلاه على إجراء اختبارات صادقة وباستخدام إجراءات قياسية موحدة مع مراعاة تناسب الاختبارات مع الأفراد موضع الاختبار والأهداف المراد تحقيقها من الاختبار.

مقاييس الاختبار

يعد الدليلان التاليان: مقاييس للاختبارات النفسية والتعليمية (الإتحاد الأمريكي للبحوث التعليمية، ١٩٨٥) وتوجيهات للاختبارات المحوسبة وتفسير

نتائجها (الإتحاد الأمريكي لعلم النفس، ١٩٨٦) مصدرين هامين لمطوري ومستخدمي ومسوقي الاختبارات. ويوفر دليل "المقاييس" معايير لتقييم الاختبارات والإجراءات الاختبارية وآثار استخدام الاختبارات. كما يحتوي هذا الدليل على الأهداف المعيارية الأولية التي يجب على واضعي الاختبارات والقائمين على الاختبارات تحقيقها، بالإضافة إلى الأهداف المعيارية الثانوية التي يُشجع على تحقيقها مع العلم أن هذا قد لا يكون ممكناً في حالاتٍ معينة. وتقسم الأهداف المعيارية إلى أربع مجموعات تتعلق ببنية الاختبار وطريقة تقييمه وبالمعايير المهنية لاستخدام الاختبارات (في المجالات السريرية والتعليمية والإرشادية والتوظيفية ومجالات منح الشهادات والتراخيص وتقييم البرامج) وبالتطبيقات الخاصة (كاختبار الأقليات اللغوية والأشخاص ذوي الإعاقات) وبالإجراءات الاختبارية.

وبالإضافة إلى المبادئ المتعلقة بواجبات مستخدمي ومطوري الاختبارات الموجودة في كتاب المقاييس، يوفر كتاب توصيات للاختبارات المحوسبة قائمة مؤلفة من تسع توصيات لإجراء الاختبارات عن طريق الحاسب وتفسير نتائجها، كما يوفر٣١ توصية لمطوري خدمات الاختبارات المحوسبة (مثل: العوامل البشرية والخصائص السايكومترية والتصنيف ومصداقية تفسير الدرجات عن طريق الحاسب). ولم تقتصر المعايير والتوجيهات على مساعدة مطوري ومستخدمي الاختبارات فحسب بل ساهمت في حث الهيئات المسوقة للاختبارات على التأكد من أن الأفراد الذين يشترون الاختبارات منهم مؤهلون تماماً لاستخدامها. ومن الجدير بالذكر أن دستور الممارسات الاختبارية العادلة في مجال التعليم (فيرمر ودايموند وكامارا، ١٩٨٩)، والذي بذل جهداً كبيراً للعمل مع ناشري الاختبارات، ساهم في تحسين نوعية الإجراءات الاختبارية بما يخدم المصلحة العامة إذ يحدد ويقسم الدستور مسؤوليات مطوري ومستخدمي الاختبارات إلى أربع

فئات: تطوير واختيار الاختبارات المناسبة وتحليل النتائج والسعي لتحقيق العدالة وإبقاء الممتحنين على دراية بكل ما يتعلق بالاختبار.

تختلف مؤهلات مستخدمي الاختبارات مع إختلاف أنواع الاختبارات وأغراضها فعلى سبيل المثال إن المؤهلات المطلوبة لشراء اختبار فردي أشد صرامة من تلك المطلوبة لشراء اختبار جماعي، كما أن المؤهلات المطلوبة لشراء اختبارات الذكاء أعلى من تلك المطلوبة لشراء الاختبارات التحصيلية واختبارات الاستعدادات الخاصة. وقد نهجت عدة منظمات نهج الشركة المتحدة لعلم النفس والتي اعتمدت نظاماً خاصاً يحدد ما يطلب من الراغبين بشراء الاختبارات، ويقسم هذا النظام الاختبارات إلى ثلاثة مستويات (أ، ب، ج) مرتبة تصاعدياً من حيث المتطلبات إذ يمكن للمدارس وغيرها من المنظمات أن تشتري اختبارات من الفئة (أ) بعد تقديم رقم شراء رسمي أو رسالة رسمية موقعة من قبل شخص مسؤول عن المشتريات ضمن المؤسسة. وأما بالنسبة للأفراد الراغبين في شراء اختبارات من الفئتين (ب) و (ج) فيطلب منهم إرسال إثبات عن مؤهلاتهم قبل أن يسمح لهم بشراء المواد الاختبارية المصنفة ضمن هاتين الفئتين، فيشترط في الراغبين بشراء اختبارات من الفئة (ب) حيازتهم على درجة الماجستير على الأقل في علم النفس أو في التعليم أو ما يعادلها من التدريب المتعلق بالتقييم النفسي أو أن يكونوا أعضاء في أي إتحاد مهني يخضع أعضاؤه لتدريب مناسب في مجال التقييم. أما الراغبون في شراء اختبارات من الفئة (ج) فيجب أن يكونوا من حملة درجة الدكتوراة في التعليم أو علم النفس أو ما يعادلها من التدريب في مجال التقييم أو أن يكونوا حاصلين على شهادة أو إجازة تتطلب تدريباً وخبرة مناسبين في مجال التقييم النفسي. ولقد وضعت عدد من دور النشر (كدار نشر علماء النفس الاستشاريين وNCS للتقييم ومصادر التقييم النفسي وغيرها من دور النشر المذكورة في الملحق ب من هذا الكتاب) نظاماً مشابهاً لمؤهلات

الراغبين بشراء الاختبارات النفسية. وعلى الراغبين في شراء الاختبارات ملء استمارة لتحديد درجات مؤهلاتهم.

المبادئ والقواعد الأخلاقية

وضع دستور الممارسات الاختبارية العادلة في مجال التعليم (فيرمر ودايموند وكامارا، ١٩٨٩) لائحة بعدد من الأخلاقيات المرتبطة بمطوري ومستخدمي أدوات الاختبار، بينما حددت وثيقة أخرى هي "المبادئ الأخلاقية لعلماء النفس ودستور سلوكهم" (الاتحاد الأمريكي لعلم النفس، ١٩٩٢) سلوك علماء النفس كعلماء وكممارسي مهنة وكمحافظين على المصلحة العامة. وتؤكد هذه الوثيقة على أن التقييم والتشخيص الرسميين يجب أن يتما في ظروف رسمية فقط وعلى أيدي مستخدمين للاختبارات مدربين وأكفاء. وينص هذا الدستور الأخلاقي على استخدام طرق علمية في تصميم وإختيار الاختبارات المناسبة لمجموعة معينة من السكان. ثم يتم بعد ذلك إجراء الاختبارات ضمن ظروف آمنة ثم توضع الدرجات وتحلل النتائج بطريقة حكيمة.

ومن الجدير بالذكر أن توفر اختبارات مصممة بدقة وحذر ووجود معايير ثابتة لتوزيعها لا يضمن أن عمليتي إجراء الاختبارات النفسية والتعليمية وتحليل نتائجها تتمان بشكل صحيح. وعلى الرغم من أن علماء النفس يعترفون بأن مسؤولية استخدام الاختبارات بشكل صحيح تقع على عاتق مستخدمي الاختبارات أنفسهم، تنقص العديد من المرشدين الاجتماعيين وأصحاب العيادات وغيرهم من العاملين في مجال الخدمات البشرية، المعرفة والمهارات اللازمة لإجراء وتحليل بعض أدوات القياس السايكومترية. ويجب إعلام أولئك الذين لم يحصلوا على التدريب الكافي بالنقص المعرفي لديهم ويجب بالتالي تشجيعهم على الإلتحاق بدورات تدريبية أو العمل مع المختصين والخبراء والرجوع إلى مصادر المعلومات المتوفرة بغرض تدعيم إمكاناتهم. كما يجب على القائمين على

الاختبارات أن يكونوا على درجة من الحساسية والدراية بلمتطلبات وحاجات المتحنين والمجتمع بشكل عام. ويعرض كتاب "الاستخدام المسؤول للاختبارات" دراسات لعدة حالات تم فيها إجراء عدد من الاختبارات ووضع درجاتها ثم تحليلها بشكل خاطئ. ومن الجدير بالذكر أن الأمثلة المطروحة في الكتاب الآنف الذكر لا تقتصر على مادة دراسية واحدة بل تشمل عدداً لا بأس به من العلوم والمقررات.

ومن أهم المبادئ التي يتوجب على القائمين على الاختبارات والممتَحَنين معرفتها مبدأ الموافقة المطلعة ومبدأ السرية المهنية. لطالما ارتبط مبدأ الموافقة المطلعة بعالم الطب فقط حيث يعطي المرضى موافقتهم على إجراء عملية جراحية ما لهم (أو غيرها من المعالجات) أو على المشاركة بدراسة سريرية ما بعد الإطلاع على جميع الحقائق الطبية والمخاطر المرتبطة بما هم مقدمون عليه. أما استخدام هذا المصطلح في مجال علم النفس وفي مجال الاختبارات النفسية تحديداً فيشير إلى موافقة الفرد على المشاركة في اختبار نفسي أو المشاركة في دراسة نفسية بعد الإطلاع على جميع الحقائق والمخاطر التي يمكن أن تترتب على هذه المشاركة. وقبل إجراء معظم الاختبارات النفسية الرسمية يتم الحصول على الموافقة المطلعة للممتحن أو ممثله القانوني. ويفترض هنا أن المعلومات الشخصية التي يتم الحصول عليها من هذه الاختبارات لا تتم مناقشتها إلا مع الممتحن نفسه فلا يطلع عليها أي طرف ثالث دون موافقة الممتحن. وعليه تصبح المعلومات "مصانة" أي محمية من أن يطلع عليها من لا يملك الحق القانوني بالإطلاع. ويوضح الصندوق ١-١ نموذجاً للإستمارة التي يملؤها الممتحن للموافقة المطلعة على المشاركة في اختبار نفسي. تفيد الاستمارة بأن الممتحن على دراية بالأهداف الأساسية للاختبار وأن بإمكانه الإنسحاب في أي وقت وأن نتائج الاختبار لن يطلع عليها غيره دون موافقته. إذن، تشير الاستمارة إلى أنه تم

إعلام الممتحن بهدف الاختبار وبمن يمكنه الحصول على نتائجه وكيف سيتم استخدام هذه النتائج.

يعود أصل مبدأ السرية المهنية إلى معلومات يأتمن المدعي أو المدعى عليه في قضية ما محاميه أو طبيبه أو رجل الدين أو الزوج (أو الزوجة) عليها، وتصبح هذه المعلومات محصنة من الإدلاء بها أمام القضاء إذا أراد الموكل أو المريض أو التائب أو الشريك في الزواج هذه الحصانة.

الصندوق ١-١

الموافقة المطلعة على اختبار نفسي

أنا الموقع أدناه ---------- أعطي موافقتي الطوعية على أن أكون أحد المشاركين في الاختبار النفسي الذي تجريه ----------. كما أقر بأنني قد أُعطيت تفسيراً كاملاً وواضحاً عن الطبيعة العامة للاختبار وأهدافه المحددة والأسباب التي سأختبر لأجلها، وأنني قد أُعلمت بأنواع الاختبارات وغيرها من الإجراءات التي سيتم إجراؤها والطريقة التي سيتم استخدام نتائج الاختبار فيها.

وأنا على علم أنه من غير الممكن لمجري الاختبار أن يوضح جميع أوجه الاختبار قبل إتمامه بكامله. كما أنني على معرفة بأنه يمكنني الإنسحاب من الاختبار في الوقت الذي أشاء دون أن أتعرض لأي خسارة أو جزاء. وأعلم كذلك أنه سوف يتم إعلامي بنتائج الاختبار عند توفرها ولن يطلع عليها أي طرف آخر دون موافقتي. كما أود طلب إرسال نسخة من النتائج إلى العنوان التالي:

اسم الممتَحَن ---------- التوقيع ----------

التاريخ ---------- توقيع مجري الاختبار ----------

ويستخدم مصطلح السرية المهنية في السياق النفسي بمعنى أن المعلومات الناتجة عن الاختبارات النفسية يمكن إطلاع أطراف ثالثة عليها بطلب رسمي فقط، وحتى في حال موافقة الممتحن على إطلاع أطراف ثالثة على نتائج الاختبار من الممكن أن تبقى المعلومات محصنة من الناحية القانونية بحيث لا يطلع أحد غير الممتحن، أو في حال كون الممتحن قاصراً يمكن لكل من ولي أمره أو وكيله القانوني أو طبيبه أو طبيبه النفسي، الاطلاع على المعلومات الناتجة عن الاختبار.

ويوجد العدد من القوانين (مثل: الحقوق التعليمية للعائلات وقانون الخصوصية لعام ١٩٧٤) التي تنص على سرية المعلومات الناتجة عن الاختبارات النفسية والتعليمية المختلفة، وأن نتائج الاختبارات وملفات الطلاب المحفوظة لدى المدارس والمؤسسات التعليمية التي تتلقى دعماً مادياً من الموارد الفيدرالية لا يمكن الإطلاع عليها إلا بموافقة الطالب أو وليه الخطية وبطلب شخصي من الجهة الراغبة في الإطلاع على هذه الوثائق. أما بالنسبة للوالدين والعاملين في المدرسة الذين لهم "اهتمام تعليمي مشروع" ببعض الطلاب، يسمح لهم هذا القانون والفقرات ٩٤-١٤٢ من القانون العام في حالة الطلاب المعاقين، بمراجعة ملفات هؤلاء الطلاب.

وبالإضافة إلى حَقَّيّ الموافقة المطلعة والسرية المهنية، يحق للأشخاص المختبرين نفسياً أن توصف حالتهم بشكل صحيح ولكن بمصطلحات تخلو من التشهير والإهانة إذ يتم اختيار مصطلح لوصف الحالة خالٍ من الدلالات السلبية، فعلى سبيل المثال إن استخدام مصطلحات مثل متخلف عقلياً أو ضعف عقلياً له وقع أقل إدانةً من مصطلحات غبي أو أبله أو أحمق، ولذلك يفضل استخدام المصطلحين الأوليين. كما أن للأشخاص المعاقين جسدياً الحق في أن يتم اختبارهم باختبارات سايكومترية مصممة بشكل خاص للشريحة السكانية التي ينتمون إليها والتي تم قياسها وتوثيقها استناداً لعينات من الأشخاص مختارة من تلك الشريحة السكانية. ويعد هذا أحد تعاريف العدالة أوالتحرر من التحيز في الاختبار.

القضايا القانونية في الاختبارات الوظيفية

ليس من المستغرب أن عدد القضايا القانونية المتعلقة بالاختبارات التعليمية والنفسية قد ارتفع بشكل كبير نتيجة الاستخدام الواسع النطاق لهذه الاختبارات في مجالات التوظيف والتعليم والعيادات النفسية ونتيجة لكون نتائج هذه الاختبارات ذات تأثير عميق على حياة الأفراد ونتيجة للتخوف على حقوق النساء والأقليات العرقية.

الجدول ١-٣

قضايا وأحكام قانونية هامة متعلقة بالاختبارات النفسية ضمن مجالي التوظيف والتعليم

١٩٦٣ **مايرت ضد موتورولا:** إن اختبار التوظيف رقم ١٠ الذي تستخدمه شركة موتورولا لا يدخل تحت مظلة تكافؤ الفرص ولا يأخذ بعين الاعتبار المجموعات البشرية المحرومة والفقيرة. كما أن هذا الاختبار قديم بالٍ لا يتماشى مع الوقت الحاضر وقد تم قياسه بناءً على معايير norms تم الحصول عليها عن

طريق إجراء الاختبار على عينة من الناس تنتمي إلى الفئات ذات الحظوة في المجتمع.

١٩٦٧ **هوبسون ضد هانسن:** تميز اختبارات القابلية الجماعية ضد أطفال الأقليات وعليه لا يمكن استخدام هذه الاختبارات لتوزيع الطلاب على "مسارات" مختلفة للقدرة.

١٩٧٠ **ديانا ضد مجلس التعليم العام:** لا يمكن استخدام إجراءات الاختبار التقليدية لتصنيف الأطفال من أصل مكسيكي على أنهم "معاقون يمكن تعليمهم" بل يجب استخدام إجراءات خاصة (كأساليب التقييم المعتمدة على لغتين) لاختبار أطفال الأقليات.

١٩٧١ **غريغس وغيره ضد شركة كهرباء ديوك:** يمنع استخدام إجراءات التوظيف التي تهدف لاستبعاد الزنوج والتي لا يمكن أن تبرر علاقتها بالأداء الوظيفي.

١٩٧٢ **غوادالوبه ضد مقاطعة تيمب الإبتدائية:** يجب أن يتم اختبار الأطفال بلغتهم الأولى ويجب، بناءً على ذلك، حذف عدد من الأسئلة غير العادلة من اختبارات الذكاء التي تقرر انضمام الأطفال إلى صفوف "المعاقين الممكن تعليمهم".

١٩٧٣ **الولايات المتحدة ضد شركة كهرباء جورجيا:** تم تعزيز سلطة توصيات الـ EEOCC في تحديد صدق اختبارات التوظيف. ويجب على الشركات المختلفة أن تثبت أن إجراءاتها لاختيار الموظفين تقيس المهارات اللازمة لتأدية العمل.

١٩٧٥ **شركة الورق ألبرمال ضد مودي:** لا يمكن الاعتماد على تصنيفات المشرفين في اثبات صدق اختبارات التوظيف وحتى في حال كون الاختبار صادق ولكنه يؤثر بشكل سلبي على توظيف من ينتمون إلى جماعات بشرية معينة يتعين على المؤسسة بذل قصارى جهدها لإيجاد وسيلة اختيار أقل تحيزاً.

١٩٧٦ **واشنطن ضد دايفيس:** تم توسيع قاعدة المعايير التي يصادق من خلالها على اختبارات اختيار الموظفين لتضم برامج تطوير الأداء التدريبية المقدمة أثناء الخدمة.

١٩٨٠ **لاري ف. ضد رايلز:** إن اختبارات الذكاء الفردية متحيزة ضد السود وعلى إدارة ولاية كاليفورنيا التعليمية أن تتوقف عن استخدام هذه الاختبارات في عمليات التشخيص والتصنيف التعليمية.

١٩٨٠جمعية الآباء الناشطين في التعليم الخاص (PASE) ضد هانون: إن استخدام اختبارات WISCو -R WISC واختبار ستانفورد - بينيه مع مراعاة المعايير المنصوص عليها قانونيا لتحديد البرنامج التعليمي المناسب لطفل ما لا يميز ضد الأطفال السود.

١٩٨١ديبرا ب. ضد تيرلينغتون: وقد أكد هذا الحكم حكماً سابقاً صدر عن محكمة فيدرالية يفيد بأن برنامج اختبارات ولاية فلوريدا لاختبار الحد الأدنى من الكفاءة هو برنامج غير دستوري لأنه يخالف بند الحماية المتكافئة التي ينص عليها الدستور.

١٩٨٥ NAACP في ولاية جورجيا ضد ولاية جورجيا: لا تميز اختبارات الذكاء ضد الأطفال السود ولا يشكل وجود أعداد كبيرة من الأطفال السود ضمن صفوف "المعاقين الممكن تعليمهم" دليلاً على التمييز العنصري في هذه الاختبارات.

١٩٨٨واتسون ضد بنك وتروست فورت وورث: يمكن المصادقة على إجراءات التوظيف الشخصية كالمقابلات ولكن يمكن للموظفين الإدعاء على أرباب العمل على أساس "التأثير السلبي" الناتج عن الترقيات المبنية على المقابلات.

١٩٨٩شركة واردس كوف للتغليف ضد أنتونيو: عكست فكرة أساسية في قضية غريغس ضد شركة كهرباء ديوك إذ حملت الموظفين أنفسهم مسؤولية إثبات عدم صدق أو ثبات الاختبارات النفسية المستخدمة في قرارات الترقية.

ولقد كان التحيز في الاختبارات النفسية كأدوات للاختيار والتنسيب في المؤسسات السبب الذي استندت إليه العديد من القضايا القانونية في العقود الأربعة الأخيرة (انظر الجدول ١-٣). ومن أول القضايا هي قضية مايرت ضد شركة موتورولا والتي استخدمت في اختيارها للموظفين اختباراً تم قياسه على عينة اختبار مؤلفة من البيض فقط وبناءً عليه حكمت المحكمة بأن الاختبار منحاز ضد الفئات المحرومة حضارياً. ونتيجةً لذلك أصدر الكونغرس الأمريكي قانون الحقوق المدنية لعام ١٩٦٤ والذي يمنع البند الثامن منه التمييز على أساس العرق واللون والأصل والجنس والدين. وقد وضع هذا القانون تحت

الاختبار في قضية غريغس ضد شركة كهرباء ديوك التي رفعت للمحكمة العليا إذ شككت هذه القضية بإجراءات شركة كهرباء ديوك في اختيار الموظفين والتي تشترط شهادة ثانوية بالإضافة إلى حد أدنى من الدرجات في اختبار ووندرليك للموظفين واختبار الإسيعاب الميكانيكي. وقد أتى حكم المحكمة ليمنع استخدام أي إجراء توظيفي ليس له علاقة بالأداء الوظيفي يهدف إلى استبعاد الأفارقة الأمريكيين. ورغم اعتراف المحكمة العليا بفائدة الاختبارات النفسية في عملية اختيار الموظفين إلا أنها أكدت على أن هذه الاختبارات لا يمكن أن تعطى سلطة الإدارة إلا في حالات ارتباطها المباشر بقياس الأداء الوظيفي.

وقد هدفت قرارات المحكمة العليا في قضية موتورولا وقضية الولايات المتحدة ضد شركة كهرباء جورجيا وقضية شركة ألبرمار للورق ضد مودي إلى مطالبة أرباب العمل توضيح وتوثيق العلاقة بين المهارات التي تقيسها الاختبارات المختارة وبين الوظائف الشاغرة. ولكن، وكما يوضح قرار المحكمة في قضية واشنطون ضد دايفيس، لا يشترط في الاختبارات أن يصادق عليها بمقارنتها مع الأداء الوظيفي الفعلي فالأداء ضمن برنامج تدريبي مناسب هو معيار مناسب للمصادقة على اختبار مصمم لاختيار الموظفين.

وقد أدى الفشل في تطبيق قانون الحقوق المدنية لعام ١٩٦٤ في مواقف متعلقة بالتوظيف إلى صدور قانون تكافؤ الفرص الوظيفية لعام ١٩٧٢ وإلى تشكيل مجلس تنسيق تكافؤ الفرص الوظيفية (EEOCC). وقد ساهمت وثيقة التوصيات الموحدة لإجراء ات اختيار الموظفين والتي وضعها مجلس EEOCC في التعريف بالإجراءات التي يتوجب على أرباب العمل والمنظمات العمالية ووكالات التوظيف استخدامها لتوضيح صدق إجراءات اختيارهم للموظفين والتي يمكن أن تكون متحيزة ضد أعضاء أي جماعة من عرق أو لون أو جنس ما. إضافة إلى ذلك، وكما جاء في الحكم في قضية شركة ورق البمارل ضد مودي، فحتى عندما

يكون إجراء الاختبار تم إثبات صدقه ولكن له تأثير سلبي على توظيف جماعات بشرية ما يتعين على المؤسسة أن تحاول إيجاد إجراء اختيار آخر أقل تحيزاً. ويعرّف مبدأ "التأثير المضاد (السلبي)" قانوناً وفق قاعدة الأربعة أخماس على النحو التالي: تعد إنخفاض نسبة إختيار مجموعة بشرية ما عن أربعة أخماس من نسبة المجموعة ذات نسبة الإختبار الأعلى دليلاً على وجود حالة من التأثير المضاد.

وقد أدت إلتباسات متعلقة بمعنى المعايير التي لها علاقة بالعمل وما يعتبر دليلاً مقبولاً على صدق اختبار ما إلى تعديلات في توصيات الـ EEOCC لاختيار الموظفين عام ١٩٧٨، وطالبت التوصيات المعدلة كما كان حالها قبل التعديل أرباب العمل بتبرير استخدام اختبارات وإجراءات إختيار تستبعد أعداداً نسبها غير مقبولة من المنتمين إلى الأقليات أو من النساء. ووضحت التعديلات إلى حدٍ ما أساليب الادّعاء التي يمكن أن يستخدمها أرباب العمل لإثبات صدق إجراءات الاختيار التي يستخدمونها. واستناداً إلى الحكم في هذه القضية وغيرها من القضايا (مثل قضية شركة ووادرس كوفر للتغليف ضد أنتونيو) تم تشريع قانون الحقوق المدنية لعام ١٩٩١ والذي (١) وضح شرط كون عبء إثبات صدق الاختبار يقع على عاتق رب العمل. (٢) ومنع بالقانون استخدام الدرجات الفاصلة التفاضلية التي تم تحديدها على أساس العرق أو الجنس أو الخلفية العرقية. وعليه فقد تم منع أي شكل من أشكال "تعديل الدرجات" المبني على أساس العرق أو اللون أو الدين أو الجنس أو القومية بموجب الفقرات ١٠٢-١٦٦ من الفقرة ١٠٦ من القانون العام. كما منع القانون ما يسمى بالمعايرة العرقية وهي إجراء تحدد فيه معايير الاختبارات لكل عرق بمعزل عن الأعراق الأخرى وتقارن درجات المتقدمين مع أولئك من الذين ينتمون إلى نفس المجموعة البشرية وذلك قبل إتخاذ قرارات توظيف بشأنهم.

القضايا القانونية في مجال الاختبارات التعليمية

تعرضت الاختبارات في مجال التعليم إلى انتقادٍ حاد، شأنها كشأن اختبارات التوظيف، بسبب تحيزها ضد الأقليات والنساء على وجه الخصوص. ولعل أهم القضايا المتعلقة بالتحيز ضد الإناث هي تلك المتعلقة باختبار القابلية المدرسية (التقدير في الوقت الحاضر) SAT Scholastic uptitude (assessment now) الذي يتفوق فيه عادةً الذكور. ونظراً لأن المنح الجامعية تعتمد بشكل أساسي على نتائج اختبار SAT، فإن أعدداً أكبر من الذكور يحصلون على هذه المنح (شيا، ١٩٩٤). وتبقى العلاقة بين تفاوت معدلات درجات الذكور والإناث في اختبار SAT وغيره من اختبارات القبول الجامعي (وبخاصة في الرياضيات) وطبيعة الاختبارات بحد ذاتها علاقة غير واضحة ومحفوفة بالجدل، تلعب المدارس والعوامل البيولوجية وغيرها من عوامل البيئة المحيطة بالطلاب دوراً فيها. فقد أثبتت الدراسات أنه لا يوجد أي مجال أكاديمي يتفوق فيه الذكور بشكل ملحوظ على الإناث (وو، ١٩٩٧) إلا أن نتائج اختبارات SAT تشير إلى عكس ذلك.

وبالإضافة إلى الاختلافات بين الذكور والإناث في نتائج اختبارات SAT، تتباين نتائج الأفراد من الفئات العرقية المختلفة إذ يحرز الأمريكيون من أصل أفريقي نتائج أدنى في قسمي الاختبار المتعلقين باللغة (SAT-V) والرياضيات (SAT-M) بشكل ملحوظ من أقرانهم البيض. إلا أن درجة كون هذه الاختلافات نتيجة لتحيز في الاختبار ذاته هي قضية لا تزال موضع للنقاش وبدون حل واضح. ومن القضايا الأخرى المتعلقة باختبار SAT مطالبة الطلاب الرياضيين بمعدل ٧٠٠ درجة في اختبار SAT كحد أدنى أو ١٧ درجة في اختبار الجامعات الأمريكي (ACT) ومعدل نهائي لا يقل عن ٢,٥. ومما أثار الكثير من الجدل وأدى أحياناً إلى دعاوى قضائية استخدام اختبارات الذكاء في تحديد مَن مِن طلاب المدارس يتوجب إلحاقه بفصول التعليم الخاص.

لطالما قامت السلطات التعليمية في الولايات المتحدة بالاعتماد على نتائج طلاب المدارس في اختبارات الذكاء المختلفة (الجماعية والفردية) في توزيع هؤلاء الطلاب على فصول التربية الخاصة. وفي واقع الأمر، يمكن القول أن هذا التوزيع هو أحد الأهداف الرئيسة لاستخدام اختبارات الذكاء في القرن العشرين. وتكمن المشكلة، التي لفتت النظر إليها عملية الدمج بين الأعراق والأجناس في المدارس، في وجود أعداد كبيرة (وغير متناسبة مع نسب المجموعات البشرية الأخرى) من الأفارقة الأمريكيين في عداد هذه الفصول الخاصة بسبب تدني معدلات درجاتهم في اختبارات الذكاء. وقد جاء قرار المحكمة في واحدة من أول القضايا التي أثارت هذا الموضوع (قضية هوبسون ضد هانسون) ليمنع استخدام اختبارات الذكاء الجماعية المتحيزة ضد الأفراد المنتمين إلى الأقليات في توزيع طلاب المدارس على مسارات مختلفة للقدرة. وفي قضيةديانا ضد مجلس التعليم العام صدر حكم بعدم جواز استخدام إجراءات الاختبار التقليدية لتصنيف الأطفال من أصل مكسيكي في ولاية كاليفورنيا على أنهم "معاقون يمكن تعليمهم" بل يجب اختبار أطفال الأقليات بلغتهم الأم. وقد أكدت قضية غوادالوبه ضد مقاطعة تيمب الإبتدائية على أنه يجب أن يتم اختبار الأطفال بلغتهم الأولى كما طالبت المحكمة بحذف عدد من الأسئلة غير العادلة من اختبارات الذكاء وأن يكون الحد الأدنى من الدرجات الذي يقرر إلتحاق الطلاب بفصول التعليم الخاص أدنى بدرجتين من المعدل الوسطي للاختبار وأن تستخدم مقاييس للسلوك التكيفي في تقرير ما إذا كان الطفل متخلفاً عقلياً أم لا.

وقد صدر حظر في قضية لاري ب. ضد رايلز على استخدام اختبارات الذكاء في إلحاق الأطفال بفصول التعليم الخاص في مدارس ولاية كاليفورنيا الحكومية، إلا أن حكماً مغايراً صدر في قضية جمعية الآباء الناشطين في

التعليم الخاص (PASE) ضد هانون حيث قررت المحكمة أن "استخدام اختبارات WISCو WISC-R واختبار ستانفورد - بينيه مع مراعاة المعايير المنصوص عليها قانونيا لتحديد البرنامج التعليمي المناسب لطفل ما... لا يميز ضد الأطفال السود" (PASE ضد هانون، ١٩٨٠، ص. ٨٨٣). وصدر حكم مشابه - أن اختبارات الذكاء لا تميز ضد الأطفال السود - في قضية NAACP في ولاية جورجيا ضد ولاية جورجيا. وبناءً عليه تم رفع الحظر على استخدام اختبارات الذكاء في ولاية كاليفورنيا بحجة أن هذا غير عادل في حق الآباء من الأفارقة الأمريكيين الراغبين في إلحاق أطفالهم المعاقين عقلياً بفصول التعليم الخاص (بريداماير، ١٩٩١).

هذا وتستمر الأسئلة والتساؤلات في المستقبل حول الطبيعة العادلة أو التحيزية لاختبارات اختيار الموظفين وقرارت اختيار وتوزيع الطلاب على فصول القدرات المختلفة في مجال التعليم. ومن الجدير بالذكر أن هنالك وجوه اقتصادية وتعليمية وسياسية وعلمية واجتماعية متعلقة بهذه القضايا وغالباً ما تفشل قرارات المحكمة في هذه القضايا في حل المشكلة. ومن المؤكد أن استخدام الاختبارات لن يتوقف إلا أنه، ولربما، ستصبح الاختبارات أكثر حساسية تجاه الإختلافات بين المجموعات البشرية المختلفة أو ستتخذ أشكالاً جديدة. ويجب على مستخدمي الاختبارات أن يكونوا على دراية بأن أهداف ومخططات الأفراد تختلف بإختلافهم. وإنه غالباً ما تكون الحلول الوسط هي خير الحلول لمثل هذا النوع من النزاعات الاجتماعية، وعلى الرغم من أن الإتفاق بين المتنازعين هو بلاشك إتفاق مؤقت إلا أنه ربما يبقى أفضل من السعي لحل على أساس كل شيء أو لاشيء إذ يؤدي هذا إلى تطويل وتعقيد عملية الصراع.

الخلاصة

يركز هذا الكتاب على تطوير وتطبيق اختبارات القدرات الإدراكية والمهارية. ومن أهم الاختبارات التي يهتم بها الكتاب الاختبارات التحصيلية واختبارات الذكاء واختبارات الاستعدادات الخاصة. كما يهتم الكتاب بالمشاريع والمنتجات وغيرها من أشكال الأداء التي لا تختبر بالشكل التقليدي إلا أنها تعمل على تقييم المهارات في مجالات إدراكية ونفسية - حركية.

ويعود استخدام الاختبارات والامتحانات بغرض اختيار موظفي المؤسسات الحكومية إلى أكثر من ٤٠٠٠ عام، إلا أن الاختبارات الشفهية والكتابية الرسمية لم تستخدم في الجامعات إلا منذ عدة مئات من الأعوام. ولم يتم الاتجاه إلى التعامل مع الاختبارات النفسية بطريقة علمية مستندة إلى الأبحاث إلا في النصف الثاني من القرن التاسع عشر. وكان فرانسيس غالون وألفرد بينيه وغيرهما من علماء النفس من رواد البحث في مجال قياس الذكاء. وعلى حين أثبتت اختبارات غالتون السايكومترية أنها غير عملية، استخدمت مقاييس ستانفورد - بينيه وما أتى بعدها بشكل واسع في تشخيص حالات التخلف العقلي وقياس القدرة العقلية بمختلف درجاتها. كما تم استخدام اختبارات الذكاء الجماعية مثل اختباري الجيش ألفا وبيتا بشكل أوسع من استخدام مقاييس الذكاء مثل ستانفورد - بينيه وشبيهاتها.

كما شهدت السنوات الأولى من القرن العشرين بالإضافة إلى اختبارات الذكاء الفردية والجماعية المقننة ظهور اختبارات مقننة للتحصيل والاستعدادات الخاصة. وقد استخدمت اختبارات التحصيل في المدارس والجامعات كمقاييس لما قد أحرزه الطالب في مادة أكاديمية أو في مقرر تعليمي ما، كما استخدمت

اختبارات الاستعدادات الخاصة مثل الأعمال المكتبية والمهارات النفسية - الحركية والميكانيكية في مجال اختيار الموظفين بشكل خاص.

وقد أدى النمو الذي شهدته "صناعة" الاختبارات النفسية والتعليمية منذ العشرينيات إلى تأسيس العديد من دور النشر المسوقة لمواد الاختبار مثل خدمة الاختبارات التعليمية (وهي أكبر هذه الدور) والدار الوطنية لأنظمة الحاسب ومكتب اختبارات كاليفورنيا وشركة علم النفس المتحدة والبرنامج الأمريكي للاختبارات الجامعية ودار نشر ريفرسايد (راجع الملحق ب والملحق ج). وتسوق دور النشر هذه وغيرها من المؤسسات العديد من أدوات المقاييس السايكومترية في مجال كل من الإدراك والشعور والمجال الحسي - الحركي عند الأفراد. وقد استخدمت هذه الاختبارات بشكل واسع وغالباً ما كان لها تأثيرات مثيرة للجدل على البرامج التعليمية والتوظيفية والإرشادية والعلاجية. كما ترعى وتعد بعض المنظمات الحكومية والخاصة كاتحاد الجامعات الطبية الأمريكية، والمجلس الوطني لاختبارات المحامين، والمفوضية الأمريكية للوظائف الحكومية، اختبارات غير تجارية. ومن ضمن الأعداد الكبيرة من المصادر المتوفرة عن الاختبارات المنشورة وغير المنشورة المصادر التالية: كتاب الاختبارات المنشورة، وكتاب الاختبارات، والكتب السنوية للمقاييس العقلية، وكتاب مقالات نقدية عن الاختبارات بالإضافة إلى عدد من قواعد البيانات الالكترونية والمقالات في المجلات التخصصيةوالاختبارات بحد ذاتها ودلائل المستخدم التي ترفق بالاختبارات. وعلى الرغم من أن معظم الاختبارات المذكورة في المصادر أعلاه هي اختبارات تجارية منشورة، هنالك نسبة كبيرة من الاختبارات غير المنشورة والتي تطورها الهيئات الحكومية والشركات الخاصة بهدف استخدامها داخلياً.

كما يلتزم مستخدمو الاختبارات المحترفين ودور النشر ذات السمعة الطيبة بمقاييس أخلاقية حددتها منظمات رسمية كالاتحاد الأمريكي لعلم النفس والاتحاد الأمريكي للأبحاث التعليمية والاتحاد الأمريكي للتوظيف والإرشاد. ولعل من أهم المصادر الجديرة بالذكر هنا المصادر التالية: مقاييس الاختبارات النفسية والتعليمية، وتوصيات للاختبارات المحوسبة وتفسير نتائجها، و"دستور الممارسات الاختبارية العادلة في مجال التعليم والمبادئ الأخلاقية لعلماء النفس ودستور سلوكهم". وقد أكدت جميع هذه المصادر على ضرورة تسويق واستخدام الاختبارات ضمن درجة من الحساسية للخصائص السايكومترية لهذه الاختبارات (من حيث الصدق والثبات وإجراءات القياس... وغيرها)، ولكفاءات القائمين عليها، ولمصلحة الممتحنين والمصلحة العامة. كما يجب أن لا يتم اختبار أي فرد دون موافقته أو موافقة وكيله القانوني، ويجب التعامل مع النتائج التي يتوصل إليها الاختبار ضمن مبادئ السرية والخصوصية المنصوص عليها، وأخيراً يتم التعبير عن الحالات التي تشخصها الاختبارات بطريقة تخلو من التشهير والتجريح.

وقد ركزت الدعاوى القضائية في مجال استخدامات الاختبارات النفسية والتعليمية في سياقي التعليم والتوظيف على نزاعات متعلقة باختيار وتوزيع الموظفين ضمن المؤسسات المختلفة. وقد تناول هذا الفصل من الكتاب بعض الدعاوى القضائية الهامة والتي كانت أسبابها في مجال التوظيف تعود إلى إدعاءات بفقدان العدل في سياسات اختيار الموظفين المنتمين إلى الأقليات، فيما كانت أسباب القضايا في مجال التعليم مبنية على إدعاءات بالتوزيع غير العادل لأطفال الأقليات على صفوف التعليم الخاص. وغالباً ما كانت اختبارات الذكاء هي السبب الرئيسي للجدل والنزاع.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

القدرة

الاختبار التحصيلي

التأثير السلبي (المضاد)

المقاييس العاطفية

المعيار الموافق للعمر

اختبار الاستعداد

العمر القاعدي

العمر السقفي

القدرات الإدراكية

تفسير (تحليل) نتائج الاختبارات عن طريق الحاسب (CBTI)

العمر العقلي

اختبار مبني على مرجعية المحك

المعايير

الاختبار الموضوعي

السؤال المفتوح

اختبارات الأداء

اختبار القوة

المعلومات المحصنة من الإفشاء

اختبار مبني على مرجعية المعيار

العدالة

السؤال محدد البدائل

التقييم التكويني

قاعدة الأربعة أخماس

الاختبار الجماعي

الاختبار الفردي

الموافقة المطلعة

الذكاء

تشخيص الحالة دون تجريح

اختبار الاتقان

المهارات الحركية – النفسية

المعايرة المبنية على العرق

معدل الذكاء النسبي IQ

اختبار السرعة

شكل الدليل اللولبي

الاختبار القياسي

التقييم التراكمي

القدرات اللغوية

المراجع المقترحة:

Bowman, M. L. (1989). Testing individual differences in Ancient China. *American Psychologist, 44,* 576–578.

Brown, D. C. (1994). *Subgroup norming: Legitimate testing practice or reverse discrimination?* Washington, DC: American Psychological Association.

Doyle, K. O., Jr. (1974). Theory and practice of ability testing in Ancient Greece. *Journal of the History of the Behavioral Sciences, 10,* 202–212.

Forrest, D. W. (1974). *Francis Galton: The life and work of a Victorian genius.* London: Elek.

Goldstein, G., & Hersen, M. (1990). Historical perspectives. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 3–17). New York: Pergamon Press.

Robertson, G. J. (1992). Psychological tests: Development, publication, and distribution. In M. Zeidner & R. Most (Eds.), *Psychological testing: An inside view* (pp. 159–214). Palo Alto, CA: Consulting Psychologists Press.

Rosenbach, J. H., & Rusch, R. R. (1991). IQ and achievement: 1930s to 1980s. *Psychology in the Schools, 28* (4), 304–309.

Sandoval, J., & Irvin, M. G. (1990). Legal and ethical issues in the assessment of children. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 86–104). New York: Guilford Press.

Sokal, M. M. (1987). *Psychological testing and American society 1890–1930.* New Brunswick, NJ: Rutgers University Press.

Thorndike, R. M., & Lohman, D. F. (1990). *A century of ability testing.* Chicago: Riverside.

Wyche, L. G., & Novick, M. R. (1992). Standards for educational and psychological testing: The issue of testing bias from the perspective of school psychology and psychometrics. In A. K. H. Burlew, W. C. Banks, H. P. McAdoo, & D. A. ya Azibo (Eds.), *African American psychology: Theory, research, and practice* (pp.229-235). Newbury Park, CA: Sage.