

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة محمد لمين دباغين

كلية العلوم الإنسانية والاجتماعية

قسم علم النفس وعلوم التربية والأرطوفونيا

مطبوعة مقياس:

القياس النفسي والتربوي

المستوى: السنة الثالثة إرشاد وتوجيه.

السداسي الخامس

إعداد: د. عبد السلام دعيدش

ماي 2018

توصيف المقرر:

* المعارف المسبقة:

- يكون لدى الطالب معارف حول مبادئ الإحصاء.
- يكون لدى الطالب مفاهيم حول معنى القياس في المجال النفسي والتربوي.
- خصائص القياس النفسي والتربوي.
- *الأهداف العامة للمادة التعليمية: يتمكن الطالب في نهاية السداسي من معرفة أنواع الاختبارات، وطرق بنائها، ومجالات استخدامها وكيفية توظيفها لأغراض التشخيص والتوجيه، وحل مشكلات المفحوصين في الوسط المدرسي.
- *الأهداف التعليمية للمادة (أهداف إجرائية قابلة للتقييم):
- التمكن من مفاهيم القياس النفسي والقياس التربوي.
- التمكن من بناء اختبارات نفسية وتربوية من خلال اظهار الطالب مهارات في مراحل إعداد وبناء الاختبارات (صياغة محتوى الاختبارات ودراسة الخصائص الإحصائية لها).
- يتمكن الطالب من معرفة أهم الاختبارات والمقاييس التي تستخدم في مختلف مجالات القياس النفسي والتربوي
- القدرة على تفسير درجة الاختبار (المعايير) وتوظيفها في ميدان الارشاد والتوجيه.

فهرس المحاضرات:

- المحاضرة رقم 01: المحاضرة رقم 01: مفاهيم القياس ومستوياته.ص03
- المحاضرة رقم 02: الاختبارات النفسية والتربوية والمفاهيم المرتبطة بها.ص09
- المحاضرة رقم 03: تحليل بنود الاختبار (مؤشر الصعوبة، التمييز، الثبات)ص14
- المحاضرة رقم 04: ثبات الاختبار.....ص24
- المحاضرة رقم 05: ثبات الاختبار(تابع).....ص28
- المحاضرة رقم 06: صدق الاختبار.....ص32
- المحاضرة رقم 07: معايير الاختبارات ومحكات الأداء.....ص39
- المحاضرة رقم 08: قياس الذكاء (مفهومه وتطورات قياسه)ص41
- المحاضرة رقم 09: نظريات الذكاء.....ص46
- المحاضرة رقم 10: اختبارات الذكاء.....ص60
- المحاضرة رقم 11: قياس التحصيل.....ص68
- المحاضرة رقم 12: قياس الشخصية.....ص70
- المحاضرة رقم 13: قياس الشخصية(تابع).....ص74
- المحاضرة رقم 14: قياس الميول.....ص77
- قائمة المراجع.....ص80

. مقدمة عامة:

تعد الاختبارات النفسية أدوات لجمع البيانات والمعلومات حول الأفراد المختبرين بهدف تشخيصهم وتصنيفهم أو إصدار الحكم على تقدماتهم ومكانتهم النسبية مع الجماعة المعيارية، ولكي نستفيد من تلك الخدمات التي تقدمها هذه الاختبارات ينبغي ان ننتبه إلى حقيقة مهمة وهي أي أداة يمكن أن تكون مفيدة أو ضارة، بحسب كيفية بناؤها واستخدامها. فالعمليات الاختبارية تنمو بوتيرة متزايدة وتساهم بفعالية في مجالات كثيرة في حياتنا اليومية، إذ أضحت هذه الاختبارات تخدم وظائف مهمة في البحوث الأساسية مثل قياس التغيرات النمائية مدى الحياة داخل الفرد، وتأثير المتغيرات البيئية في الأداء الشخصي، ولكن هذا النمو قد صاحبه بعض التوقعات غير الواقعية وسوء الاستخدام، مما يفرض على المشتغلين في مجال الاختبارات معرفة كيفية تقييم الاختبارات والتمييز بينها، وطرق تجويدها وتفسير نتائجها. وسنتعرض في هذه المحاضرات إلى جملة من المفاهيم الواردة حول القياس النفسي والتربوي، والعمليات المصاحبة لبناء الاختبارات وتصنيفاتها المختلفة.

المحاضرة رقم 01: مفاهيم القياس ومستوياته.

1 – تعريف القياس النفسي والتربوي. يمكن القول أنّ القياس، بالمعنى الواسع، بأنه تعيين أرقام للأشياء أو الأحداث وفقا لقواعد. وحقيقة أن الأرقام يمكن تعيينها تحت قواعد مختلفة تؤدي إلى أنواع مختلفة من المقاييس وأنواع مختلفة من القياس. وتصبح المشكلة بعد ذلك هي تفسير: (أ) القواعد المختلفة لتعيين الأرقام، (ب) الخصائص الرياضية (أو بنية المجموعة) للمقاييس الناتجة، و (ج) العملية الإحصائية المطبقة على نمط القياس مع كل نوع من المقاييس (Stevens, 1946)

ويعرف القياس أيضا بأنه التحديد الرقمي للسلوك أو الخصائص وفق قواعد محددة، بحيث تتضمن عملية قياس أي خاصية: أساليب قياس دقيقة ومناسبة أو عمليات، وقواعد محددة لتعيين قيم عددية للخاصية أو لأداء الفرد على اختبار يقيس هذه الخاصية (Brown, 1984) وطبقا لهذه التعريفات نجد (Thorndike & others, 1991) أشار إلى أن القياس يتضمن ثلاث عمليات أو خطوات أساسية هي:

- تحديد الخاصية المراد قياسها أو تقديرها.
- تحديد مجموعة الاجراءات التي يمكن عن طريقها ابراز هذه الخاصية حتى يمكن ملاحظتها، فيسهل قياسها.
- بناء أو تصميم مجموعة من الأساليب أو الأدوات لترجمة هذه الملاحظات إلى مصطلح أو صيغ كمية أو قيم رقمية.

وفي ضوء ما تقدم يمكن أن نعرف القياس، على أنه التعبير الرقمي لمقدار امتلاك الفرد لخاصية ما أو قدرة معينة بمقياس معين، وفي إطار قواعد معينة.

2— مستويات القياس النفسي : لقد ذكرنا أن القياس هو التقدير الكمي لخاصية ما أو قدرة معينة، ويكون لهذا التقدير قواعد وخصائص رياضية تسمى مستوى القياس، والمستوى نظام لتعيين قيم أو درجات للخصائص موضع القياس، وقد حدد (Stevens, 1946) في مقال بعنوان "نحو نظرية مستويات القياس" أربع مستويات للقياس، المستوى التصنيفي، المستوى الرتبي، المستوى الفئوي (المجال)، المستوى النسبي.

1-2. المستوى الاسمي (التصنيفي) Echelles Nominales: يعتبر هذا المستوى من أبسط مستويات القياس، يناسب المتغيرات النوعية، ويتضمن تصنيفا أو تقسيما إلى فئات على أساس خاصية معينة أو عدة خصائص مميزة، حيث يجب أن توضع كل الأشياء التي يتم قياسها في فئات حصرية وشاملة (Laurent, 2010) يكون الهدف من عملية القياس في هذا المستوى هو التصنيف classification الذي يراعي الفروق النوعية بين الأفراد مثلا: ناجح / راسب. منبسط/ منطوي. واستخدام الأرقام هنا لا يعكس مقادير كمية. فهي تستخدم فقط لتمثيل فئات فريدة من نوعها وليس لها خصائص رياضية، مثلا نرسم لمجموعة الذكور بالرقم (1) ومجموعة الاناث بالرقم (2)، هذه الأرقام ليست لها دلالة كمية. متغيرات هذا المستوى دائما تكون منفصلة. ولا يسمح هذا المستوى بإجراء العمليات الحسابية الكلاسيكية ما عدا العد البسيط (حساب التكرارات).

2-2. المستوى الرتبي Echelles Ordinales: على غرار المستوى الاسمي، يُسمح لنا في هذا المستوى بتصنيف الأفراد بناء على خاصية معينة، كما يكون ترتيب الأفراد متاح أيضا في المستوى الرتبي، ويكون الترتيب تبعا لخاصية أو سمة ما على أساس التفاضل، والأرقام التي تنسب لمثل هذه المتغيرات تخدم غرضين أساسيين هما:

- تصنيف هذه المتغيرات في فئات أو مجموعات تدل عليها.
- بيان درجة الأفضلية من حيث مدى الامتلاك لسمة معينة، الأمر الذي يساعدنا في ترتيبها تنازليا أو تصاعديا.

الأرقام التي تنسب لمثل هذه المتغيرات لا تعكس بالضرورة مقادير كمية، كما أن المسافات بين الرتب (الأرقام) غير متساوية فعلى سبيل المثال: ترتيب الافراد حسب الطبقة الاجتماعية وتعد الرتب المئينية، ومكافئات العمر، ومكافئات الفرق الدراسية أمثلة لمستويات رتبية شائعة.

2-3. المستوى الفئوي أو المجال: L'échelle d'intervalle: يتضمن هذا المستوى خصائص المستوى التصنيفي والترتيبي المتمثلة في التصنيف والترتيب، ويتميز باستخدام الأرقام بصيغة كمية للتعبير عن مدى امتلاك السمة أو الخاصية أي الأرقام هنا تعكس معان كمية، لذا يعد أرقى مستويات القياس. ويمكن لنا المقارنة بينها على أساس كمي نظرا لتوفر وحدة قياس متساوية، فالفرق بين الدرجتين 2 و 3 في اختبار تحصيلي هو نفسه بين الدرجتين 3 و 4. ومعظم الاختبارات النفسية تصمم للحصول على درجات تقع في المستوى الفئوي (Pettersen, 2002)، ولتوضيح معنى تساوي الفروق بين الوحدات المتجاورة وخصائص هذا المستوى نقدم المثال التالي: في اختبار تحصيلي في مادة العلوم حصل الطالب (أ) على الدرجة 19، والطالب (هـ) على الدرجة 10، والطالب (س) على الدرجة 14. من خلال هذه الدرجات يمكن لنا أن نعرف أن الطالب (أ) حصل على أعلى الدرجات يليه الفرد (س) ثم الفرد (هـ)، وبما أن الدرجات تقع على مستوى فئوي فيمكننا القول إن الفرق بين درجتَي الفردين (أ) و (س) مساوٍ للفرق بين درجتَي الفرد (س) و (هـ). ونستطيع التعامل مع بيانات المستوى الفئوي باستخدام جميع العمليات الرياضية المعروفة (الجمع، الطرح والضرب، القسمة) في حين المستويين التصنيفي والترتيبي لا يمكن التعامل معهما باستخدام هذه العمليات. كما يُمكننا من استخدام معظم الإجراءات الإحصائية (مقاييس النزعة المركزية، التشتت) مع البيانات المحصل عليها. ونظرا لتوفر هذا المستوى على وحدة قياس ثابتة يجعله يتوفر على خاصية وجود الصفر، غير أنه ليس حقيقي مطلق فهو لا يعكس الغياب التام للسمة، فمثلا حينما يحصل شخص ما على درجة صفر في اختبار القدرة العددية فإن ذلك لا يدل على افتقاره كلية لهذه القدرة.

2-4. المستوى النسبي Échelle de ratio ou proportionnelle: يتميز هذا المستوى من القياس بخصائص المستوى الفئوي، بالإضافة إلى وجود صفر مطلق والذي يعكس الغياب التام أو الانعدام للسمة المراد قياسها (Pettersen, 2002) ومثال ذلك الطول، الكتلة، ويستخدم هذا النوع من القياس في العلوم الطبيعية، ونادرا جدا ما يستخدم في القياس النفسي كقياس زمن الرجوع

3- افتراضات القياس النفسي: نستعرض فيما يلي الافتراضات الأساسية التي ينطوي عليها القياس النفسي والتربوي، وهذه الافتراضات اقترحها (Cohen & Swerdlik, 2002) وهي تبسيط لبعض القضايا المعقدة للغاية في القياس النفسي والتربوي وتصنع جدلا مكثفا لدى خبراء القياس النفسي:

-**الافتراض الأول:** توجد تكوينات فرضية نفسية: التكوين الفرضي هو السمة أو الخاصية (متغير) التي يصمم الاختبار لقياسها، ويهتم ميدان القياس النفسي والتربوي بقياس عدد من التكوينات الفرضية، مثل: ذكاء الفرد، الميول الدراسية، التحصيل الدراسي، ولهذه التكوينات وجود.

-**الافتراض الثاني: امكانية قياس التكوينات الفرضية:** خبراء القياس يعتقدون أنه يمكن قياس التكوينات الفرضية النفسية والتربوية، وهذا استنادا إلى قبولنا بأن لهذه التكوينات وجود، حيث يقول (Cronbach) "إذا وجد شيء ما، فإنه يوجد بمقدار معين، وإذا وجد بمقدار معين، فإنه يمكن قياسه".

-**الافتراض الثالث: على الرغم من أنه يمكننا قياس التكوينات الفرضية، إلا أن قياسنا لا يكون تاما:** يعتقد خبراء القياس أنه يمكنهم قياس التكوينات الفرضية النفسية، ولكنهم يعترفون أيضا بأن عملية القياس ليست تامة. ويوضع ذلك عادة في إطار أخطاء القياس وتأثيراتها في ثبات الدرجات، وهناك جهود لتقدير وخفض تأثيرات القياس إلى أقل حد ممكن.

-**الافتراض الرابع: توجد طرق مختلفة لقياس أي تكوين فرضي:** توجد مداخل متعددة لقياس أي تكوين فرضي نفسي، ولننظر إلى مثال الميول الدراسية، فالميل الدراسي لدى المتعلم يمكن قياسه باستخدام عدد من المداخل المختلفة. فمثلا يجري اخصائي الارشاد النفسي والمدرسي مقابلة مع العميل ويسأله عن ميوله الدراسية، ويمكن أن يلاحظه في مواقف دراسية، كما يمكن توجيه موازين تقدير لمدرسيه تتناول الميول الدراسية للعميل، إلى جانب إمكانية تطبيق اختبارات ومقاييس الميول الدراسية لتقييم طبيعة الميول لدى العميل. هذه الإجراءات كلها تساعد الاخصائي في الفهم الجيد لميول العميل الدراسية.

-**الافتراض الخامس: جميع إجراءات القياس لها جوانب قوة ومحدوديات:** رغم تعدد مداخل قياس التكوينات الفرضية. نجد لها جوانب قوة ومحدودية، فقد ينتج عن تطبيق أحد مداخل القياس ثبات كبير لدرجاته، لكنه لا يقيس بعض مظاهر تكوين فرضي معين بنفس جودة مدخل آخر ينتج عنه درجات أقل ثباتا. فمن المهم جدا فهم جوانب القوة والضعف للإجراءات المستخدمة.

-**الافتراض السادس: ينبغي أن تكون المصادر المتعددة للمعلومات جزءا من عملية القياس:** القرارات المهمة لا يجب أن تعتمد على نتائج اختبار واحد أو إجراء قياس واحد، فمثلا عند اتخاذ قرار توجيه متعلم لتخصص دراسي معين، فإن معلومات مثل متوسط المعدل في المواد التي تصب في التخصص، وخطابات التوصية الصادرة من المدرسين، وكذا الأداء في مقياس الميول، ينبغي أخذها بعين الاعتبار.

-**الافتراض السابع: يمكن تعميم الأداء في الاختبارات على سلوكيات غير اختبارية:** عندما نطبق اختبارا فإننا لا نهتم عادة بأداء الفرد في الاختبار، وإنما بالقدرة على التعميم من الأداء في الاختبار على سلوكيات غير اختبارية.

4- أهداف القياس

- تحديد الخصائص الشخصية والنفسية والعقلية للإنسان، وتصنيفها بهدف التعرف على مختلف جوانبها وتبيين المتغيرات المتعلقة بها، وذلك للوصول إلى القوانين التي تحكم سلوكنا وقدراتنا العقلية بوصفنا

أفرادًا، وبالتالي سلوكنا الجمعي بظواهره النفسية والتربوية والاجتماعية. - الحصول على معلومات محددة تفيد المجتمع بمستوياته كافة العام والخاص والفردى فالمسئول في المجال التربوي والمجال النفسي والمجال الاجتماعي وغيرها من المجالات، يتعين عليه بحكم عمله الوفاء بمطالب معينة. فهو مطالب، في مجال التربية مثلا، بتوجيه الطلاب وفقًا لقدراتهم، ومطالب في الوقت نفسه بتشخيص الحالات غير السوية لتتلقى ما تتطلبه من علاج أو رعاية..

- الاختيار والتصنيف، ويقصد به تحديد مستويات الأشخاص في سمات معينة وتصنيفهم وفقًا للمجال المناسب لكل منهم سواء تعلق ذلك بالنواحي العملية أم التعليمية -الكشف عن فعالية الجهاز الإداري أو التربوي في البرامج والأقسام العلمية والإدارية وغيرها، والتأكد من صحة القرارات التي اتخذت.

- التعرف على المستوى العلمي للطلبة في المهارات والقدرات الأساسية، وما قد يعثرها من تغير وتحول عبر السنين.

-تشخيص العملية التعليمية واكتشاف ما تعانيه من مشكلات آنية، وما قد يعثرها من عوائق مستقبلية، من مستوى المؤسسة الواحدة إلى التعليم على مستوى الدولة.

-تحديد مستويات أداء عناصر العملية التربوية من خلال الكشف عن أداء الطلبة أنفسهم.

-فحص الأهلية. ويقصد به تحديد ما إذا كانت تتوفر في الفرد الأهلية الشروط اللازمة لتولي مهمة معينة أو الانخراط في عمل معين مثل الطب والتعليم وغيرهما.

4-مبادئ القياس والتقويم : هناك مجموعة من المبادئ التي ينبغي مراعاتها عند تخطيط وتنفيذ عملية القياس والتقويم إذا أريد لهذه العملية النجاح في بلوغ أهدافها وهذه المبادئ هي:

أ- **الاتساق:** مع الأهداف المراد تحقيقها (تنوع في الوسيلة كي يتحقق الهدف) وأن مفهوم الاتساق

مع الهدف المراد تحقيقه ينطوي على ميزتين هما:

1. مدى إمكانية تناسق المنهج لتحقيق الهدف.

2. مدى إمكانية التنوع في الوسائل المستخدمة لتحقيق هذا الهدف.

أما بالنسبة للموضوع الأول فإنه يجب أن يكون هذا الاتساق مع الفلسفة التي يقوم على أساسها المنهج الذي يتم من خلاله تحقيق الهدف ، وهذا يعني انه هناك تناسقاً وانسجاماً في المنهج يساعدان على تحقيق الهدف أما بالنسبة للنقطة الثانية وهي إمكانية تنوع في الوسائل والأساليب من أجل تحقيق الهدف فإنه يعد من المطالب الأساسية في العملية التعليمية، فإذا الهدف معرفة مدى تحصيل الطالب فإن التدريسي يستخدم أسلوب الاختبارات بأنواعها ولكن إذا كان الهدف هو قياس ذكاء الطالب فإن استخدام التدريسي للاختبارات التحصيلية لا يكشف عن الهدف وكذلك الحال بالنسبة للميول والاتجاهات

والعلاقات الاجتماعية ومن شأن ذلك أن يساعد التدريسي على تطوير خبراته بدلا من إتباع الروتين في استخدام الاختبارات مما يجعل العملية التربوية أكثر فعالية.

ب الشمول: من المبادئ الرئيسية للقياس والتقويم أن يتميز التقويم بالشمولية بمعنى أن يشمل جميع جوانب الموضوع المراد تقييمه فمثلا إذا أردنا تقييم المناهج ومدى نجاحه وأثر هذا المنهاج على الطالب فإن التقويم يجب أن يشمل كل الجوانب التي لها علاقة بالموضوع أما إذا كان الهدف تقويم العملية التربوية أو التعليمية، بشكل عام فإنه يظهر من خلال ما يأتي:

- جميع أطراف وعناصر العملية التعليمية كالتدريسي والمساق الدراسي والقسم المعني وما فيه من خدمات.
- جميع جوانب الموضوع المقيم وبما أن موضوع التقويم هنا لا يقتصر على النواحي المعرفية وإنما يجب أن يشمل كافة الجوانب الأخرى كالشخصية والنمو الانفعالي والجسمي والعقلي واللغوي والاتجاهات والميول.

ج- التعاون: يجب أن يتم القياس بطريقة تعاونية يشارك فيها كل من يؤثر في العملية التربوية ويتأثر بها والمسؤولين في التربية وأفراد من البيئة كالخبراء ويجب أن تكون هناك فرص للتقويم الذاتي من جانب الطالب وكذلك من جانب التدريسي.

د- الاستمرارية: يجب أن يكون القياس والتقويم عملية تقدير مستمرة لمدى ما يحققه البرنامج التربوي من الأهداف المرسومة لعملية التربية حتى يتم تصحيح مسار عملية التعلم باستمرار وهذا يعني أن تكون الملاحظة اليومية متواصلة على مدار العام الدراسي ولا تقتصر على فترة محددة منه أي يجب أن تكون مواكبة لعملية التعلم فتلازمه وتسير جانبا إلى جنب معها.

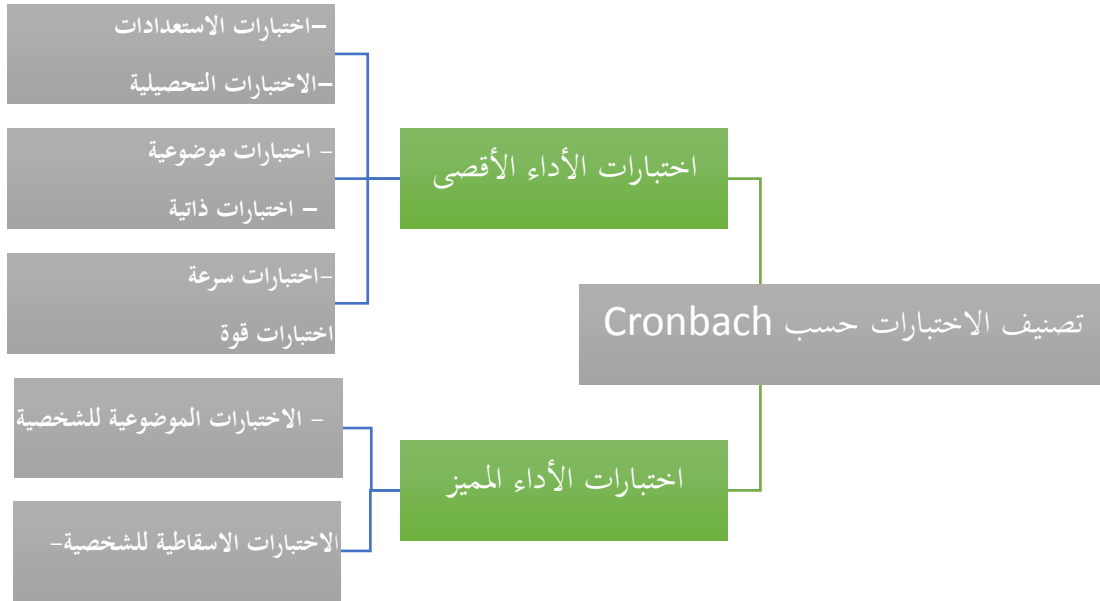
هـ- التشخيص والعلاج: يجب أن يكون القياس والتقويم تشخيصيا وعلاجيا في الوقت نفسه بمعنى أن يصف نواحي القوة ونواحي الضعف في عمليات الأداء وفي نتائج هذا الأداء يقصد تعزيز نواحي القوة والإفادة منها والعمل على علاج نواحي الضعف وتلافيها أو التقليل من حدته.

المحاضرة رقم 02: الاختبارات النفسية والتربوية والمفاهيم المرتبطة بها:

- 1-تعريف الاختبارات :** كلمة "اختبار" حسب (Murphy & Davidshofer, 2005) تشير إلى طريقة منتظمة ومعيارية في الحصول على معلومات حول بعض جوانب السلوك البشري. يشمل هذا التعريف كلا من الاختبارات التربوية والنفسية، ويغطي ثلاث خصائص محددة:
- الاختبار النفسي هو عينة من السلوك.
 - يتم الحصول على العينة في ظل ظروف موحدة.
 - هناك قواعد ثابتة لتسجيل النقاط، أو للحصول على معلومات كمية (رقمية) من عينة السلوك.
- _ الاختبار هو طريقة لقياس قدرة الشخص، معارفه أو أدائه في مجال معين (Brown, 1984)
- _ الاختبار هو مقياس موضوعي ومقنن لعينة من السلوك، والاختبارات النفسية تماثل الاختبارات في علم آخر، من حيث أن الملاحظات تُجرى على عينة صغيرة ولكن يتم انتقاؤها بعناية (Anastasi & Uraina, 1997)

- وردت في التعريفات السابقة مصطلحات التقنين والمعايير، الموضوعية، ربما تحتاج إلى توضيح أدق كون هذه المفاهيم مرتكزات أساسية في بناء وتطبيق الاختبارات وتفسير نتائجها.
- _ **التقنين:** انتظام وتوحيد الإجراءات عند تطبيق الاختبار وتقدير درجاته، وتكوين المعايير إحدى الخطوات المهمة في تقنين اختبار معين.
- _ **المعايير:** تكوين المعايير إحدى الخطوات المهمة في تقنين الاختبارات، والمعايير حسب Sax هي توزيع الدرجات في اختبار من الاختبارات المطبق على مجموعة مرجعية تم اختيار أفرادها بصورة تمثل المجتمع الأصلي (شحاتة، 2007).
- _ **الموضوعية:** المقصود بها تجنب الذاتية والابتعاد عن الأحكام الذاتية في تقدير وتفسير درجات الاختبار من طرف القائمين على الاختبار.

- 2-تصنيف الاختبارات النفسية والتربوية :** تختلف طريقة تصنيف الاختبارات باختلاف الأسس المتبعة، من حيث الغرض وطريقة البناء والتطبيق، وكذا طريقة تفسير النتائج، وقد ذكر Cronbach 1990 أنه يمكن تصنيف الاختبارات بعامة في مقاييس الأداء الأقصى أو في مقاييس الأداء المميز (Reynolds & Livingston, 2013)



1-2. **اختبارات الأداء الأقصى:** وتسمى أيضا اختبارات القدرة، وهي متضمنة اختبارات التحصيل، تُقَدَّر درجة مفردات الاختبار هنا على أساس أن إجابة المفردة إما "صحيحة" أو "خاطئة". تصمم هذه الاختبارات لقياس الحدود العليا للفرد وقدراته، أين يحث المختبرين على بذل أقصى قدرتهم في أدائهم، مثلا: نصمم اختبار لقياس مدى اتقان المتعلم لنطاق محتوى دراسي معين (قياس درجة التحكم في تقنيات الاعلام الآلي لطلبة الجامعة). وفي إطار اختبارات الأداء الأقصى توجد تقسيمات فرعية متعددة وهي:

أ- **الاختبارات التحصيلية واختبارات الاستعدادات:** تصمم الاختبارات التحصيلية لقياس معارف ومهارات فرد معين في نطاق محتوى قُدِّم له فيه تعليم معين أي تقدير ما حصله من منهج دراسي أو برنامج دراسي (عباس، 2001) أما اختبارات الاستعدادات تكون أكثر اتساعا في منظورها، وتصمم لقياس المهارات والقدرات المعرفية، والمعارف التي تجمعت لدى الفرد نتيجة لخبراته مدى حياته. غير أن الفرق بين اختبارات الاستعدادات والتحصيل هو فرق نسبي، حيث أن كلاهما يمكن استخدامه في ظروف معينة لتقرير تأثير الخبرة وللتنبؤ بالتحصيل (عباس، 2001)

ب. **الاختبارات الموضوعية والذاتية:** تشير الموضوعية إلى استبعاد أثر العوامل الذاتية أو الشخصية في عملية تقدير الدرجة في الاختبار، وتلغي احتمال ظهور أي خلاف بين المقدرين نظرا لتوافر مفتاح تصحيح يضمن التجرد من التحيزات الشخصية. ومفردات هذا النوع من الاختبارات تكون على شكل: اختيار من متعدد، الصواب والخطأ، المزوجة، أما الاختبارات الذاتية فهي تعتمد في تقدير درجاتها على الاحكام الذاتية للقائمين على عملية التقدير، ويطلق عليها عادة اختبارات المقال.

ج. **اختبارات السرعة واختبارات القوة:** تهدف اختبارات السرعة Speed test إلى معرفة الفروق في سرعة الأداء أو التنفيذ لدى المختبرين، فقراتها تكون سهلة نسبيا، غير أن الزمن المخصص للإجابة

يكون محددًا. في حين اختبارات القوة Power Test يكون الزمن المتاح كاف للإجابة على جميع فقراته.

2-2-اختبارات الأداء المميز : يطلق عليها أيضا اختبارات الشخصية، وهي أدوات لقياس الخصائص الانفعالية والدافعية والبيشخصية و الاتجاهات (Anastasi & Uraina, 1997)) حينما نستخدم أحد اختبارات الشخصية فإننا نكون مهتمين بمعرفة النمط أو الأسلوب الذي يتصرف وفقا له شخص معين في المواقف المختلفة، وليس هناك إجابة صحيحة أو خاطئة للعبارات التي تصف سلوكه (شحاتة، 2007)). لذا نجد البعض يستخدم مصطلحات مثل: مقياس Scale أو استبيان inventory على أدوات الاستجابة المميزة.

ويمكن تقسيم اختبارات الشخصية إلى اختبارات موضوعية واختبارات إسقاطية، هذ التقسيم يقوم على أساس طبيعة المثير الذي يقدم للمفحوص وشكل الاستجابة المتوقعة منه:

أ.الاختبارات الموضوعية للشخصية : هي تلك الاختبارات التي لا تتأثر بالأحكام الشخصية للقائم بتقدير درجات الاختبار، حيث يكون هناك اتفاق بين مقدري الدرجات بحكم وجود مفتاح التصحيح مسبقا، وبنية المثيرات محددة. إذ تستخدم عدد محدد من البدائل في الفقرة الواحدة و يقرر المختبر ما يراه منها منطبقا عليه (علام ، 2000) فمثلا: مقياس Rosenberg لتقدير الذات Self-Esteem Scale ، يشتمل على مفردات تستخدم عدد محدد من البدائل (موافق بقوة، موافق بدرجة بسيطة، بين الموافقة وعدمها، غير موافق بدرجة بسيطة، غير موافق بشدة) و تقابلها تقديرات تدرجية من (04) لموافق بقوة إلى صفر لبدل غير موافق بشدة ، مثل : "أشعر بأنني شخص له قيمة أو على الأقل مساوي للآخرين" فالفرد في هذه الفقرة يستجيب ببساطة في ضوء البدائل المحددة سابقا، فحينما يختار البديل "موافق بقوة" سيحصل في هذه الفقرة على درجة تقديرية تساوي (04)، وهذا ما يمنع من التحيز بين المقدرين عند التقيد بمفتاح التصحيح.

ب.الاختبارات الإسقاطية للشخصية : ترتكز هذه الاختبارات على تقديم مثيرات غير محددة البنية أو غامضة، وينتج عن هذه المثيرات تقريبا عددا لا متناهيا من الاستجابات من قبل المفحوصين، كما تتميز هذه الاختبارات أنها لا تفرض قيودا على طريقة أو كيفية استجابة المفحوص. أما عملية تقدير درجة الاستجابة تعتمد بصورة كبيرة على ذاتية القائم على الاختبار، وهذا هو الأساس الذي يميزها عن الاختبارات الموضوعية للشخصية. فمثلا "اختبار تفهم الموضوع" T.A.T الذي يستخدم في دراسة الشخصية، يقوم حول تقديم عدد من الصور الغامضة نوعا ما للمختبرين (عباس، 2001) ويطلب منهم تكوين أو إنشاء قصة تصف ما يدور بالصورة وتحدث عن أحوال الأشخاص والأحداث التي تجري

فيها، وبعد ذلك يقوم الفاحص بدراسة وتقدير استجابة المختبر وهذا التقدير يقترب بدرجة كبيرة إلى الذاتية، أي يتأثر بالأحكام الذاتية للمقربين.

يبقى هذا التصنيف الذي استعرضناه لمختلف أنواع الاختبارات غير شامل، بالرغم من تفرد به بالتصنيف الأساسي للأنواع الرئيسية للاختبارات النفسية والتربوية، حيث توجد تصنيفات أخرى يمكن اعتمادها أيضا ونورد أهمها:

3-2_ التصنيف حسب طريقة تفسير النتيجة :

_ الاختبارات معيارية المرجع، تركز على الوضع النسبي للفرد في علاقته بالآخرين.

_ الاختبارات محكية المرجع، تركز على مقارنة أداء الفرد بمحك أو مستويات أداء متوقعة

4-2- _ التصنيف حسب الجهة التي تعد الاختبار

_ اختبارات مقننة، يكون تطبيقها وتصحيحها، تفسير درجاتها بطريقة معيارية

_ اختبارات غير مقننة (من إعداد المعلم).

5-2_ التصنيف حسب أسلوب تقديم الإجابة :

_ اختبارات لفظية ن تكون فيها الإجابة تحريرية (ورقة وقلم) أو شفوية

_ اختبارات الأداء العملي، مثل إجراء تجربة علمية.

6-2_ التصنيف حسب طريقة التطبيق :

_ اختبارات جماعية (اختبارات الورقة والقلم).

_ اختبارات فردية (شفوية).

في ضوء هذه التصنيفات لمختلف أنواع وأشكال الاختبارات، يمكن لاختبار واحد أن يتبع

أكثر من تصنيف، فاختبار التحصيل يمكن أن يكون مقنن وموضوعي، يفسر تفسير محكي المرجع واختبار قوة من نوع أقصى الأداء.

3-خطوات بناء الاختيار :

الاختبارات النفسية في كثير من الأحيان قياسات تتمحور حول موضوع ما (أي سمة معينة) وتتبع بعض المبادئ التوجيهية الصارمة في عملية البناء، الاعداد، التصحيح والتفسير، والهدف الأساسي لتطوير مقياس جديد هو الحصول على قياس صادق وثابت، وقد ذكر (Singh, Junnarkar, &)

Crocker & Algina,) في كتابهم "القياس في علم النفس الإيجابي"، أن كل من (Kaur, 2016 (1986) حددا عشرة خطوات يمكن اتباعها في عملية البناء:

1. تحديد الغرض الأساسي الذي ستستخدم فيه درجات الاختبار (اختبار تنبؤي، تشخيصي، انتقائي...).
2. تحديد السلوكيات التي تمثل البناء النفسي أو نطاقه (تحويل البناء النفسي إلى مجموعة سلوكيات أو بنود خاصة) مراجعة الأبحاث، الملاحظات المباشرة، أحكام الخبراء، الأهداف التدريسية. استخدام تحليل المحتوى.
3. إعداد مجموعة مواصفات للاختبار تصف بدقة نسبة البنود أو الفقرات الممثلة لكل نوع من أنواع السلوكيات المحددة في الخطوات الثانية. أي ضمان عينة مناسبة من البنود داخل كل مجال من المجالات الرئيسية للسمة (عدد البنود يكون متناسب مع أهمية كل مجال).
4. إنشاء ملف أولي للبنود أو الفقرات، على مطوري الاختبار هنا اختيار الصياغة المناسبة للبنود وتجنب البنود التي تحمل الإيحاءات أو التأويل في الفهم، على العموم هنالك شكلين من الصياغة نجدهما أكثر انتشارا في الاختبارات النفسية والتربوية، بنود ذات الإجابة الثنائية (صحيح، خطأ)، (نعم، لا) وسلام التقدير على شكل مقياس ليكرت الثلاثي أو أكثر.
5. مراجعة البنود وتعديلها من حيث الدقة والصياغة والغموض ودرجة الملاءمة من خلال الاستعانة بالمختصين. إذ يجدر الأمر هنا الانتباه إلى استخدام لغة متداولة وبسيطة في الصياغة، ومراعاة المستوى القرائي للمختبرين، الأصل العرقي والمرحلة العمرية وكذا النوع (ذكور، إناث)
6. إجراء تجريب أولي للبنود على عينة من المفحوصين (15- 20 فرد)،
7. مراجعة البنود بعد التجريب الأولي وتعديلها إذا تطلب ذلك قبل طبعها نهائيا وتحليلها إحصائيا من حيث الصعوبة والتباين والتمييز.
8. تطبيق الفقرات على عينة كبيرة ممثلة لمجتمع المفحوصين (الأفراد الذين أعد الاختبار لهم)
9. تحديد الخصائص الإحصائية لدرجات بنود الاختبار (مؤشرات الصعوبة والتمييز، الثبات)، وحذف الفقرات التي لا تتفق مع المعايير المحددة مسبقا.
10. تصميم وإجراء دراسات الصدق والثبات للصيغة النهائية للاختبار للتحقق من مدى توفرها على الخصائص السيكومترية الضرورية.
11. وضع دليل الاختبار الذي يفيد في التطبيقات والتصحيح وتفسير الدرجات (مثل الجداول المعيارية، اقتراح توصيات مثل معايير الأداء، درجات القطع وغيرها) (Matlock-Hetzel, 1997)

الدراسة الاحصائية للاختبارات النفسية والتربوية

محاضرة رقم 03: تحليل البنود (مؤشر الصعوبة، مؤشر التمييز، الثبات)

1-تحليل البنود : بعد الانتهاء من الخطوات الأولى في بناء الاختبار ، والحصول على الصياغة الأولية لبنود الاختبار أو المقياس ، ونتائج التطبيق على عينة التقنين وفق الخطوات التي حددها (Crocker & Algina, 1986) تأتي مرحلة تحليل البنود وفحص الخصائص القياسية الإحصائية لها، من خلال استخدام مؤشرات ومعاملات نتحصل عليها من البنود، هذه المؤشرات والمعاملات تسمح للقائمين على بناء الاختبار بالحكم على قيمة وصلاحيته بنود الاختبار، والحصول على بنود تتوافر على الموصفات الإحصائية المطلوبة يمهد للحصول على اختبار جيد وموثوق فيه، وتتضمن هذه الإجراءات الإحصائية ثلاث عناصر (Legendre, 1998):

- التحقق من التطابق بين البنود والأهداف، ومعقولية خيارات الإجابة، بالإضافة إلى جوانب الصدق والثبات.

- حساب مؤشرات (المتوسط، الانحراف المعياري، مؤشر الصعوبة، التمييز ...)

- اتخاذ القرار بمراجعة البنود أو استبدالها أو الاحتفاظ بها. _

1-1. **مؤشر الصعوبة Indices de difficulté d'items**: يعد مؤشر صعوبة البند واحد من الأساليب الإحصائية الأكثر استخداماً وأهميةً في تحليل بنود الاختبارات، وهو تعبير عن نسبة المختبرين الذين أجابوا إجابة صحيحة على البند إلى المجموع الكلي للمختبرين. والذي نرسم له بـ P :

$$P = \frac{P_i}{n}$$

حيث: P مؤشر صعوبة البند P_i عدد الأفراد الذين نجحوا في الإجابة على البند، n العدد الكلي لأفراد المجموعة.

تتراوح قيمة P بين (0 و 1) فحينما تكون قيمة P مرتفعة جداً أكبر من 0.90، يشير إلى أنّ نسبة كبيرة من المختبرين نجحوا في الإجابة على البند (90%)، ومنه البند سهل جداً. وحينما تكون قيم P أقل من 0,20 البند سهل جداً (20%)، وتمثل القيمة (0,50) أن البند متوسط السهولة. ويتم تحويل قيمة p إلى نسبة مئوية، ووفقاً لـ (Thorndike, Cunningham, Thorndike, & Hagen, 1991) نقلاً عن (Quagrain & Ato , 2017) البنود التي تتراوح قيمتها بين 20 و 90% تعتبر جيدة ومقبولة. والتي تتراوح قيمتها بين 40 و 60% ممتازة، أما البنود ذات القيمة p (مؤشر الصعوبة) أقل من 20% (صعبة للغاية) وأكثر من 90% (سهلة جداً) تكون غير مقبولة أو تحتاج إلى تعديل. إن

البنود التي تتراوح قيمها بين 0,4 و0,6 تناسب اختبارات معيارية المرجع في حين الاختبارات المحكية المرجع تتطلب بنود تكون غالبية مؤشرات صعوبتها 0,90 فما أكثر.

يمكن الحصول أيضا على مؤشر السهولة للبند والذي نرسم له بـ Q من خلال قيمة مؤشر الصعوبة عن طريق المعادلة التالية:

$$Q=1-p$$

حيث: Q = نسبة عدد الأفراد الذين أجابوا إجابة خاطئة على البند.

تجدر الإشارة هنا إلى أن هذه المعادلة المعتمدة في حساب مؤشر صعوبة البند تصلح حينما تكون طريقة تصحيح البند ثنائية (1,0). ولعل حساب مؤشر الصعوبة في حالات الإجابات الثنائية (إجابة صحيحة أو إجابة خاطئة) يبدو بسيطا لأنه يقوم على عدّ الإجابات الصحيحة المسجلة في البند من خلال الأفراد المختبرين وتقسيم عدد الإجابات الصحيحة على المجموع الكلي للأفراد المختبرين. لكن هذا المؤشر لا يعكس بالضرورة وبدقة الإجابات الصحيحة الناتجة عن الكفاءة الحقيقية للمختبر في حالة البنود ذات الاختيار من متعدد أو صحيح/خطأ، حيث يمكن أن نسجل إجابات صحيحة عن طريق الصدفة أو الحظ والذي يطلق عليه التخمين وهو نجاح الفرد في البند دون ما يعرف الإجابة الصحيحة، للتقليل من أثر التخمين في الإجابة على البند وعزل الأسباب التي تساهم في زيادته هناك عدة إجراءات يمكن اتباعها:

- رفع عدد البدائل في البند، فكلما زاد عدد البدائل في البند يقل أثر التخمين، يكون أثر التخمين في البند الذي يحمل بديلين (0,50) وفي حالة أربع بدائل ينخفض ليصبح (0,20).
- إضافة بدائل مموهة تحمل عبارات "كل الإجابات صحيحة" "كل الإجابات خاطئة" "السؤال الحالي غير منطقي"
- استخدام طريقة الخصم (1-) في حالة الإجابة الخاطئة.

كما يمكن تصحيح مؤشر الصعوبة من أثر التخمين الذي أشرنا إليه سابقا، باعتماد معامل التصحيح لمؤشر الصعوبة من أثر التخمين عن طريق المعادلة التالية:

$$P^c = p - \left[\frac{1-p}{M-1} \right]$$

حيث: P^c مؤشر الصعوبة المصحح من أثر التخمين.

P مؤشر الصعوبة قبل التصحيح من أثر التخمين.

M عدد احتمالات الإجابة (البدائل).

أما حساب مؤشر الصعوبة في حالة عدم الإجابة على البند من طرف بعض أفراد العينة (البند المتروكة)، خاصة البنود التي تأتي في آخر الاختبار بسبب ضيق الوقت (اختبارات السرعة) أو في حالة اختبارات القوة إذ تتضمن أيضا عامل الزمن، فلا ينبغي أن نتعامل مع هذه البنود المتروكة بنفس منطق التحليل الذي نطبقه على البنود التي تأتي في بداية الاختبار. المعادلة التالية تسمح لنا بتقدير مؤشر صعوبة البند في هذه الحالة:

$$P = pin - k$$

K عدد الأفراد الذين تركوا البند (لم يجيبوا عن البند)

أهمية حساب مؤشر صعوبة البند: حساب مؤشر صعوبة البند يساهم في الحصول على بنود مناسبة ومميزة وتحسين بناء الاختبار بشكل عام، ونورد أدناه الفائدة العملية لهذا الإجراء في ضوء ما أشار إليه (Coaley , 2010).

- اختيار البنود المناسبة من حيث الصعوبة والسهولة تمهيدا للحصول على بنود تميز بين الفئة الأعلى والفئة الأدنى، بين الفئة المتمكنة وغير المتمكنة، فالقيم المتطرفة سواء التي أجاب عنها جميع المختبرين أو التي لم يتمكن أحد من الإجابة عليها يوصى بحذفها في الصيغة النهائية للاختبار، أو مراجعتها لعل بها خلل في الصياغة اللغوية أو مبهمة وغير مفهومة.

- يضمن لنا توازن في صعوبة بنود الاختبار.

- يفيدنا ذلك في ترتيب البنود من السهل إلى الصعب بهدف تحفيز المختبرين على الإجابة. وتحديد البنود التي توقف عندها المختبر ولم يجيب عنها بسبب صعوبتها أو ضيق الوقت، وهناك قاعدة نتبعها في ذلك تتمثل في التوقف عن احتساب البنود على أنها متروكة بسبب عدم معرفة الإجابة خاصة في اختبارات QCM واعتبارها متروكة بسبب ضيق الوقت وذلك بعد ثلاث بنود متتالية متروكة.

- تحديد الأخطاء الشائعة والإجابات غير المعتادة.

- استبعاد البنود المتحيزة وغير العادلة حيث يمكن أن تجيب مجموعة على بند معين في حين تظهر استجابات مختلفة تماما للمجموعة الأخرى.

1-2 مؤشر تمييز البند Indices de discrimination des items: يُعرّف مؤشر التمييز الخاص بالبند بأنه "... الدرجة التي تميّز بها بين الطلاب ذوي الإنجازات العالية والمنخفضة" (Linn & Gronlund, 1995) حيث يكون للأفراد ذوي القدرات العالية أفضلية في الإجابة على البند، فمثلا

إذا صمم اختبار لقياس فهم المنطوق، فتمييز المفردة هنا يعني قدرتها على إظهار الفرق بين التلاميذ ذوي مهارات جيدة في فهم المنطوق والتلاميذ ذوي مهارات ضعيفة في فهم المنطوق. فيجب أن يكون للبند إذن قدرة على التمييز بين الأفراد المختبرين في الخاصية المراد قياسها وإظهار الفروق الموجودة بينهم. وحسب Anastasi و Urbina طُورت أكثر من خمسين طريقة في تقدير مؤشر تمييز البند، ومعظم هذه المؤشرات تؤدي إلى قيم متقاربة (Reynolds & Livingston, 2013) ونستعرض فيما يلي أكثر المداخل استخداما في تقدير تمييز البند.

تتخذ معاملات التمييز قيما تتراوح ما بين (-1 و +1)، وتشير أي قيمة إلى دلالة معينة، حيث اقترح (Ebel & Frisbie , 1991) قيما مرجعية لتفسير معاملات التمييز:

0.40 أو أكثر : البند مميزا جيدا 0.30-0.39 : بند مميز

0.29-0.20 : بند أقل تمييزا 0.10-0.19 : بند محدود ينبغي تحسينه

أقل من 0.10: لا فائدة منه في الاختبار

- **طريقة المقارنة الطرفية:** تعد من الطرق الشائعة في حساب مؤشر تمييز البند، وتعتمد على الفرق بين أداء المجموعتين (الأعلى، الأدنى) وهناك طرق متعددة في اختيار المجموعتين، منها الطريقة التي تعتمد على ترتيب أفراد العينة بناءً على الدرجة الكلية في الاختبار ترتيبا تصاعديا، ثم اختيار (27%) من الفئة الأعلى و(27%) من الفئة الأدنى (Richard & Sheila , 1999) وبعدها نقوم بحساب صعوبة البند لكل فئة، وبعدها نطرح مؤشر صعوبة الفئة الأدنى من مؤشر صعوبة الفئة الأعلى باعتماد الصيغة الرياضية التالية:

$$D = P+ - P-$$

حيث: D = مؤشر تمييز البند.

$P+$ = مؤشر صعوبة البند في الفئة الأعلى. $P-$ = مؤشر صعوبة البند في الفئة الأدنى.

مثال: نفترض تطبيق اختبار القدرة اللغوية يتكون من (40) بند على عينة من التلاميذ (200)، وأردنا حساب مؤشر تمييز البند رقم 3. بعد تصحيح الاختبار والحصول على الدرجة الكلية لكل فرد، نقوم بترتيب الأفراد تصاعديا، ثم نحسب مؤشر الصعوبة للفئة الأعلى والذي يمثل (27%) ومؤشر الصعوبة للفئة الأدنى (27%) باتباع الطريقة التالية:

$$200 \text{ (العينة)} \leftarrow 100\%$$

54 → 27% من الفئة الأعلى.

54 → 27% من الفئة الأدنى.

حيث:

- عدد الذين أجابوا صحيح على البند رقم 3 من الفئة الأعلى 0,39. ومنه $P_{+=3954} = 0,72$

- عدد الذين أجابوا صحيح على البند رقم 3 من الفئة الأدنى 0,20. ومنه $P_{- =2054} = 0,37$

وبالتعويض في المعادلة السابقة:

$$D=0,72-0,37=0,35$$

قيمة مؤشر تمييز البند رقم 03 هي: 0,35 وهذا يدل على أن البند مميز ويمكن الاحتفاظ به في الصيغة النهائية للاختبار.

-طريقة معاملات الارتباط:

هناك ثلاث طرق تعتمد على معاملات الارتباط تستخدم في تقدير مؤشر تمييز البند، معامل الارتباط ثنائي التسلسل الخاص، معامل ثنائي التسلسل، ومعامل فاي. ويكون قياس تمييز البند بحساب درجة الارتباط بين متغيرين، واحد منهم ثنائي التقسيم (0,1) (درجات البند) والآخر متصل أي على مستوى مجال ويقصد به الدرجة الكلية للاختبار، وأهم مزايا هذه الطريقة استخدامها لدرجات كل المختبرين في حساب المعامل، عكس الطريقة السابقة التي تستخدم 54% فقط (27% أعلى، 27% أدنى) (Ebel & Frisbie , 1991)

الارتباط المرتفع بين درجة المفردة (البند) والدرجة الكلية في الاختبار يشير إلى أن البند يقيس نفس الخاصية (التكوين الفرضي) الذي يقيسه الاختبار ككل، كذلك البند يميز بين الأفراد ذوي الإنجاز المرتفع والإنجاز المنخفض في السمة.

-معامل الارتباط ثنائي التسلسل الحقيقي: يشير معامل الارتباط المستخرج بهذه الطريقة إلى القدرة التمييزية للبند، يعتمد على إيجاد الارتباط بين الأداء على البند والاختبار (الارتباط بين درجة البند والدرجة الكلية للاختبار). ويفترض أن الذين يجيبون عن البند إجابة صحيحة هم من فئة المتقنين بينما الذين يجيبون عن البند إجابة خاطئة هم من فئة الأفراد غير المتقنين، ويستخدم هذا النوع من معاملات الارتباط حينما يكون أحد المتغيرين يقع في المستوى الفئوي أو النسبي (الدرجة الكلية للاختبار) والآخر

يقع في المستوى الاسمي ذو تقسيم ثنائي حقيقي (صحيح، خطأ) للبند. والصيغة التالية تسمح لنا بالتقدير الكمي لتميز البند المعتمد على (Henning, 1987):

$$r_{pbis} = \frac{M_p - M_q}{S_x} \sqrt{pq}$$

حيث:

M_p = متوسط توزيع الدرجات الكلية للمجموعة التي أجابت إجابة صحيحة على البند.

M_q = متوسط توزيع الدرجات الكلية للمجموعة التي أجابت إجابة خاطئة على البند.

S_x = الانحراف المعياري للدرجات الكلية في الاختبار لجميع أفراد العينة.

P = ترمز إلى نسبة الأفراد الذين أجابوا صحيح على البند (مستوى صعوبة البند).

q = ترمز إلى نسبة ع الأفراد الذين أجابوا إجابة خاطئة على البند (مستوى سهولة البند).

أما في حالة تقسيم درجة البند تقسيما اصطناعيا، نستخدم معامل الارتباط ثنائي التسلسل.

مثال: طبق اختبار على عينة مكونة من (15) فردا، حصلنا على النتائج التالية في البند رقم 03. المطلوب حساب معامل ثبات البند:

الأفراد	الدرجة الكلية	البند رقم 03
1	25	1
2	23	1
3	18	0
4	24	0
5	23	1
6	20	0
7	19	0
8	22	1
9	21	1
10	23	1
11	21	0
12	20	0
13	21	1
14	21	1
15	22	1

جدول يصف توزيع الدرجات الكلية ودرجة البند

من الجدول نحسب مؤشر تمييز البند رقم 03 عن طريق صيغة r_{pbis} :

عدد الإجابات الصحيحة = 9

عدد الإجابات الخاطئة = 6

$$MP = [25 + 23 + 23 + 22 + 21 + 23 + 21 + 21 + 22] / 9 = 22.23$$

$$Mq = [18 + 24 + 20 + 19 + 21 + 20] / 6 = 20.33$$

$$S_x = 1.8$$

$$P = \frac{9}{15} = 0,60$$

$$q = \frac{6}{15} = 0,40$$

بالتعويض في المعادلة: $r_{pbis} = \frac{MP - Mq}{S_x} \sqrt{pq}$

$$r_{pbis} = \frac{22.23 - 20.33}{1.8} \sqrt{0.60 \times 0.40}$$

-معامل الارتباط فاي Phi: يستخدم معامل فاي Phi في تقدير تمييز المفردة عن طريق حساب الارتباط بين درجات البند محل دراسة صدقه (التجريب) حيث يستخدم في حالة بنود ثنائية التقسيم (صحيح، خطأ أو 0,1) ودرجات محك ثنائية أيضا وهو من الطرق التي يتم فيها تطبيق الاختبار مرة واحدة على مجموعة واحدة من الأفراد، كما يمكن استخدامه في تحديد درجة استقرار في استجابات البند نفسها ثنائية التصحيح للأفراد أنفسهم في موقف آخر، وهناك عدة صيغ لحساب Phi نذكر منها:

$$\phi = \frac{P_{jk} - P_j P_k}{\sqrt{P_j q_j P_k q_k}}$$

نفترض أن:

P_{jk} إلى النسبة المشتركة للأفراد الذين أجابوا على الفقرتين z و K إجابة صحيحة..

P_j نسبة الأفراد الذين أجابوا على البند z إجابة صحيحة.

P_k ترمز إلى نسبة الأفراد الذين أجابوا على البند K إجابة صحيحة.

q_j ترمز إلى نسبة الأفراد الذين أجابوا على البند z إجابة خاطئة.

q_k ترمز إلى نسبة الأفراد الذين أجابوا على البند K إجابة خاطئة.

مثال: طبق اختبار يقيس القدرة الابتكارية على 12 فردا، فكانت استجاباتهم على البند رقم 10 كما يلي:

الأفراد	البند محل التجريب J	البند المحكي K
1	1	0
2	1	0
3	1	0
4	0	1
5	0	1
6	1	1
7	1	1
8	0	0
9	0	0
10	0	1
11	1	1
12	1	0

المطلوب: أحسب معامل تمييز البند بطريقة الارتباط ϕ ؟

لحساب معامل الارتباط فاي

$$\phi = \frac{P_{jk} - P_j P_k}{\sqrt{P_j q_j P_k q_k}}$$

$$P_{jk} = \frac{3}{12} = 0.25$$

$$P_j = \frac{7}{12} = 0.58$$

$$P_k = \frac{6}{12} = 0.50$$

$$q_j = \frac{5}{12} = 0.42$$

$$q_k = \frac{6}{12} = 0.50$$

$$\phi = \frac{0.25 - 0.58 \times 0.50}{\sqrt{0.58 \times 0.42 \times 0.5 \times 0.50}} = -0.19$$

فرضا أن الأفراد المتقنين هم الذين يحصلون

على درجات كلية تساوي أو أكبر من 12.

ومنه البند غير مميز ويوصى بحذف

1-3. ثبات البنود:

يعتمد ثبات الاختبار اعتمادا مباشرا على ثبات بنوده، ويشير ثبات البند إلى استقرار استجابات الأفراد على نفس البند باختلاف الظروف. وهناك طرق متعددة في تقدير ثبات بنود الاختبار يمكن تلخيصها فيما يلي:

- طريقة التطبيق-إعادة التطبيق:

يمكننا استخدام معامل Phi لتحديد درجة استقرار استجابات ثنائية التصحيح في البند نفسها للأفراد في موقف آخر (Knapp , 2009) حيث تتطلب هذه الطريقة تطبيق البند على نفس الأفراد في فترتين مختلفتين وفقا للخطوات التالية:

- تطبيق بنود الاختبار على عينة من الأفراد.

- إعادة تطبيق نفس البنود على نفس عينة الأفراد بفواصل زمني.

- رصد استجابات الأفراد على كل بند من بنود الاختبار، حيث تسجل نتائج التطبيق الأول ونتائج التطبيق الثاني في شكل جدول تكراري.

- حساب معامل ارتباط فاي بين درجات الاستجابة في التطبيق والتطبيق الثاني.

يقدر معامل ثبات البند وفقا للصيغة التالية:

$$\phi = \frac{(bc - da)}{\sqrt{(a + b)(a + c)(c + d)(b + d)}}$$

حيث: a، b، c، d ترمز إلى خلايا الجدول الرباعي.

إعادة التطبيق

	1	0
1	a	b
0	c	d

التطبيق

- كما يمكن تطبيق معامل الارتباط الرباعي في حساب معامل ثبات البند وفقا للصيغة التالية:

$$r_t = \cos \left(\frac{180^\circ}{1 + \sqrt{\frac{ad}{bc}}} \right)$$

- كلما ارتفعت قيمة معامل ارتباط البند بين التطبيقين كلما دل على ثبات البند، والعكس كلما انخفضت قيمة معامل ارتباط البند دل على عدم استقرار أو ثبات البند.

- طريقة الاحتمال المنوالي:

تصلح هذه الطريقة لحساب ثبات البنود التي تعتمد إجاباتها على اختيار إجابة من إجابتين أو اختيار من إجابات متعددة، نتلخص صيغتها فيما يلي:

$$r_t = \frac{k}{k-1}$$

$$\left(f_{max} - \frac{1}{k} \right)$$

ترمز k إلى عدد بدائل الإجابة،

وترمز إلى أكبر احتمال نسبي في احتمالات الإجابة (أكبر تكرار نسبي).

محاضرة رقم 04: ثبات الاختبار

1- ثبات الاختبار:

"الثبات يشير إلى اتساق الدرجات التي يحصل عليها نفس الأفراد عند إعادة فحصهم بنفس الاختبار في مناسبتين مختلفتين، أو بمجموعتين مختلفتين من البنود المماثلة، أو تحت ظروف اختبارية متغيرة"

الثبات بمعناه الواسع يشير إلى مدى عزو الفروق الفردية في درجات الاختبار إلى فروق حقيقة في امتلاك الخاصية المراد قياسها. والمدى الذي تعزى به إلى درجة الخطأ، فالثبات في أبسط معانيه هو الحصول على نفس الدرجات تقريبا تحت ظروف متعددة، اما إحصائيا فيُعبر الثبات عن نسبة تباين الدرجة الملاحظة إلى تباين الدرجة الحقيقية، أو معامل الارتباط بين الدرجة الملاحظة والدرجة الحقيقية.

2- طرق تقدير ثبات الاختبار: هناك أربع طرق لتقدير ثبات الاختبار، ثلاثة منها تستخدم في تقدير ثبات الاختبارات التحريرية (اختبارات الورقة والقلم)، وهي على التوالي:

- طريقة التطبيق وإعادة التطبيق:

- طريقة الصور المتكافئة.

- مختلف طرق الاتساق الداخلي، وتتضمن أيضا طرق التجزئة النصفية

- ثبات تقدير المحكمين

1-2. طريقة التطبيق وإعادة التطبيق: وتسمى أيضا معامل الاستقرار عبر الزمن، تقوم هذه الطريقة على إجراء الاختبار على مجموعة من الأفراد، ثم إعادة تطبيق نفس الاختبار على نفس المجموعة في ظروف مماثلة، وذلك بعد مرور فترة زمنية محددة بين التطبيقين، بعدها نقوم بحساب معامل الارتباط بين درجات المختبرين في التطبيق الأول ودرجاتهم في التطبيق الثاني، وهذا ما نسميه معامل الثبات. تصلح هذه الطريقة في اختبارات السرعة واختبارات القوة، لكن تصبح عديمة الجدوى في حالة الاختبارات التي تهدف إلى قياس التذكر أو التكوينات الفرضية المرتبطة بهذه العملية العقلية. يجب على القائمين على تطبيق الاختبار الانتباه إلى بعض الإجراءات المحيطة بعملية التطبيق ونلخص أهمها:

- التأكد من ملاءمة طريقة تقدير الثبات لطبيعة الاختبار.

- الحرص على توحيد ظروف إجراء الاختبار.

- تفادي أن تكون المدة الزمنية الفاصلة بين التطبيقين طويلة، لما ينجر عنها من تأثيرات (عامل النضج، إمكانية حدوث تدريب...).

والصيغة التالية لمعامل ارتباط بيرسون تسمح لنا بحساب معامل الارتباط:

$$r = \frac{n \sum(xy) - (\sum x) (\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2] [n\sum y^2 - (\sum y)^2]}}$$

حيث تعبر: n عن عدد أفراد العينة.

و x عن درجات الأفراد في التطبيق الأول.

و y عن درجات الأفراد في التطبيق الثاني.

2-2. طريقة الصور المتكافئة: تفترض هذه الطريقة إنشاء صورتين متكافئتين من الاختبار، وهي اختبارات متشابهة إلى حد كبير وليست متطابقة، وحتى تكون الصيغتين متكافئتين يتطلب أن يكون محتواهما متماثلا من حيث عدد البنود وطريقة صياغتها ومستوى صعوبتها ولها نفس نطاق المحتوى (اسماعيل، 2004) ويكون تقدير معامل الثبات في هذه الحالة بتطبيق الصيغتين في على عينة واحدة و في جلسة واحدة، ولتقدير معامل الثبات إحصائيا نعتمد معامل الارتباط بيرسون، حيث تمثل (x) درجات الاختبار الذي نبحث عن ثباته، و (y) الصورة المكافئة له.

2-3. طريقة الاستقرار والتكافؤ: يتسنى لنا في هذه الطريقة الجمع بين مدخلي مع بعضهما في تقدير ثبات الاختبار الذي يعتمد على إجراء الاختبار مرتين، من خلال تطبيق النموذج الأول للاختبار وبعد مدة زمنية (من أسبوع إلى ثلاثة أشهر) نقوم بتطبيق الصيغة أو النموذج الثاني المكافئ للنموذج الأول، (Crocker & Aigina, 2008) ونقوم بعده بتقدير الثبات إحصائيا بالاعتماد على معامل الارتباط بيرسون. غير أن هذه الطريقة تتأثر هي الأخرى بأخطاء القياس للطريقتين السابقتين.

2-4. طرق التجزئة النصفية: يمكن الحصول على تقديرات لثبات الاختبار دون الحاجة إلى القيام بتطبيقين لنفس الاختبار على نفس الأشخاص، أو إيجاد صيغتين متكافئتين للاختبار، هذا النوع من الثبات يتم اشتقاقه من تقييم الاتساق الداخلي لفقرات الاختبار، وهناك عدة صيغ لتقدير ثبات الاتساق الداخلي لبنود الاختبار. يكون تطبيق الاختبار في هذه الطريقة مرة واحدة فقط، وبعدها نقوم بتقسيم الاختبار إلى نصفين (جزأين متساويين) وهناك عدة طرق لتقسيم الاختبار إلى نصفين منها طريقة البنود الفردية والزوجية، حيث يتضمن النصف الأول البنود الفردية (1،3،5...) ويتضمن النصف الثاني البنود الزوجية (2،4،6...)، عملية تقسيم الاختبار إلى نصفين تشترط تساوي بنود نصف الاختبار من حيث الطول والصعوبة ومحتوى البنود، ولخص كل من (Crocker & Aigina, 2008) هذه الشرو في أربع نقاط:

-تقسيم الاختبار إلى نصفين واختيار البنود الفردية لتشكل الاختبار الفرعي الأول، والبنود الزوجية لتشكل الفرعي الآخر.

- ترتيب البنود المشكلة للاختبار وفقا لمستوى صعوبتها ثم اختيار البنود الفردية لتؤلف الاختبار الفرعي الأول والبنود الزوجية لتؤلف الاختبار الفرعي الآخر.

- اختيار بنود كلا النصفين عشوائيا.

- تجزئة الاختبار إلى جزأين بحيث تكون هناك مزوجة بين بنود كلا الجزأين من حيث المحتوى.

من مميزات هذه الطريقة:

- تشابه ظروف التطبيق في الفقرات الفردية والزوجية كون الاختبار يطبق مرة واحدة فقط.

- لا يوجد أثر لعامل النضج وحدوث التدريب أو تعلم مهارات جديدة

- توفر الجهد والوقت على مصممي الاختبارات.

بعد الانتهاء من الخطوات السابق ذكرها، نقدر ثبات الاختبار عن طريق حساب معامل الارتباط بين نصفي الاختبار، وهناك عدة طرق لحسابه:

أ- **صيغة Spearman-Brown**: تقوم معادلة (سبيرمان براون) على أساس أنه يمكن التنبؤ بمعامل ثبات الاختبار إذا علمنا معامل ثبات نصفه، ذلك معناه إذا استطعنا أن نقسم الاختبار إلى نصفين متكافئين، ويكون المتوسط والتباين للجزء الأول مساويا للمتوسط والتباين في الجزء الثاني، كذلك تساوي مستويات السهولة والصعوبة بين الجزأين فإننا إذا طبقنا هذه المعادلة نستطيع أن نتعرف على ثبات الاختبار كله. وتعتمد هذه الصيغة على حساب معامل الارتباط بيرسون بين نصفي الاختبار، ثم إدخال معامل الثبات باستخدام صيغة سبيرمان-براون للحصول على القيمة المصححة لمعامل ثبات الاختبار.

$$\rho_{xx'} = \frac{2 \rho_{AB}}{1 + \rho_{AB}}$$

حيث تشير $\rho_{xx'}$ إلى معامل الارتباط المعدل للاختبار الكلي و ρ_{AB} إلى معامل ارتباط نصفي الاختبار.

قد لا تصلح طريقة سبيرمان براون لحساب ثبات الاختبار في حالة غياب التكافؤ بين نصفي الاختبار خاصة عند اختلاف قيم التباين بشكل كبير بين النصفين، وكذلك لا تصلح في الاختبارات الموقوتة لأن كثرة الأسئلة المتروكة في آخر كل اختبار تؤثر على الارتباط بين الجزأين.

ب- **صيغة Rulon**: في ضوء الملاحظات التي سجلت على صيغة Spearman-Brown والمتمثلة في ضرورة تكافؤ نصفي الاختبار من حيث التباينات ومؤشرات الصعوبة والذي يصعب تحقيقه، كما أن غياب هذه الشروط أو تناقصها يعمل على تخفيض قيمة معامل الثبات، اقترح **Rulon 1939** طريقة بديلة لتقدير ثبات الاختبار من درجات نصفي الاختبار (Crocker & Aigina, 2008)، والميزة الأساسية لهذه الطريقة اعتمادها على تباين الفروق بين نصفي الاختبار ($D = A - B$) حيث تمثل A درجة الفرد على نصف الاختبار الأول و B درجة الفرد على نصف الاختبار الثاني. ويستخدم تباين فروق الدرجات σ_D^2 على أنه حساب لتباين الخطأ σ_E^2 في الصيغة المحددة لمعامل الثبات:

$$\rho_{xx'} = 1 - \frac{\sigma_D^2}{\sigma_E^2}$$

ج- **صيغة Guttman**: هي من طرق التجزئة النصفية التي تقوم على التطبيق لمرة واحدة، وتصلح لحساب الثابت عندما لا تتساوى الانحرافات المعيارية لجزئي الاختبار وتصلح أيضا عندما لا تتساوى هذه الانحرافات، وتتلخص هذه الطريقة في المعادلة التالية:

$$\rho_{xx'} = 2 \left[1 - \frac{(\sigma_x^2 + \sigma_{x'}^2)}{\sigma_T^2} \right]$$

σ_x^2 تباين درجات الاختبار الفردي، و $\sigma_{x'}^2$ إلى تباين درجات الاختبار الزوجي.

و σ_T^2 إلى تباين الاختبار الكلي

المحاضرة رقم 05: ثبات الاختبار (تابع)

5-2. طرق تعتمد على التباين المشترك للبنود: تعتمد الطرق التي تتصل بالاتساق الداخلي لحساب الثبات على الاتساق في أداء المختبرين على الاختبار من بند إلى آخر، وبالتالي فهي ليست بحاجة لتطبيق الاختبار أكثر من مرة، أو تقسيمه إلى نصفين متكافئين، وإنما يقسم إلى عدد كبير من الأجزاء، بحيث يتكون كل جزء من بند واحد من بنود الاختبار، وكلما زاد الاتساق بين هذه البنود زاد ثبات الاختبار ككل (أبو ناهية، 1994) وهناك أكثر من معادلة لحساب معامل الاتساق الداخلي، ونذكر ثلاث طرق الأكثر استخداما وانتشارا.

أ-صيغتي Kuder-Richardson 1937: تستخدم في حالة البنود ثنائية التصحيح أو الإجابة (1,0)، وقد اشتقت هذه الصيغة من طرف كيودر-وريتشاردسون (21KR20) (-KR-) كبديل لطرق التجزئة النصفية التي عجزت عن إعطاء نتيجة واحدة لاختبار معين. واتخذت هذه الأسماء من الخطوات المرقمة في الاشتقاق. حيث صيغة هي:

$$KR20 = \frac{k}{k-1} \left(1 - \frac{\sum p_i q_i}{\sigma_x^2}\right)$$

حيث:

ترمز k إلى عدد بنود الاختبار، و $\sum p_i q_i$ إلى مجموع نسب ضرب معامل الصعوبة في معامل السهولة، σ_x^2 إلى تباين الاختبار الكلي.

ج- وقد اشتقا Kuder-Richardson طريقة أخرى لحساب الاتساق الداخلي للاختبار أكثر سهولة من الصيغة السابقة، حيث لا تعتمد على حساب مؤشرات السهولة والصعوبة:

$$KR21 = \frac{k}{k-1} \left(1 - \frac{\mu(k-\mu)}{k \sigma_x^2}\right)$$

حيث:

μ = متوسط الدرجة الكلية،

k = عدد بنود الاختبار،

δ_x^2 = تباين الدرجة الكلية للاختبار.

في حالة تتساوى صعوبات بنود الاختبار فإن تقديرات الثبات في كلا من الصيغتين (KR20-) و (KR21-) تكون متساوية، في حين أنه عندما تختلف صعوبة البنود فإن تقدير الثبات عند استخدام صيغة (KR21-) يكون أقل من القيمة المحسوبة باستخدام الصيغة (Crocker & Aigina, 2008).

د-معامل ألفا Coefficient alpha:

وضعه كرونباخ Cronbach 1951 وطوره فيما بعد كل من (Kaiser & Michael, 1967; Novick & Lewis, 1975)، ويمكن اعتبار معامل ألفا كمتوسط لجميع الارتباطات الممكنة ذات الانقسام النصفى (Julius & Howard, 2010) ويستخدم لحساب الثبات في حالة الاختبارات التي ينتج عن بنودها درجات قيمها متعددة (0, 1, 2) وهو على النقيض من KR-20 الذي يستخدم في حالة التقسيم الثنائي للإجابة على البنود. يعد معامل ألفا القانون الإحصائي المفضل لتقدير موثوقية الاتساق الداخلي للاختبار ويعتمد على تطبيق مرة واحدة للاختبار. كما يصلح استخدامه في اختبارات الاستجابة المميزة.

صيغة الفا كرونباخ هي:

$$\Gamma_{\alpha} = \left(\frac{K}{K-1} \right) \left(1 - \frac{\sum \sigma_i^2}{\sigma^2} \right)$$

حيث: Γ_{α} = معامل الثبات ألفا. k = عدد بنود الاختبار

σ_i = تباين كل بند . \sum = مجموع تباينات البنود.

σ^2 = تباين الدرجات الكلية

و-**ثبات تقديرات المحكمين:** عندما يتم تقدير درجات الاختبار على أحكام ذاتية، يكون مهم جدا تقويم الاتفاق حينما يقوم أكثر من فرد بتقدير الدرجات، الاختبار في هذه الصيغة يطبق مرة واحدة ويقدر فردان كل على حدا درجات الاختبار، ومن ثم يتم حساب الارتباط بين الدرجات التي نحصل عليها، هذا النوع من تقدير الثبات يعكس بدرجة أساسية الفروق التي ترجع على الأفراد الذين يقدر درجات الاختبار. بالإضافة إلى المدخل الارتباطي يمكن أيضا تقويم اتفاق المحكمين بحساب النسبة المئوية من المرات التي يُعَيَّن فيها المحكّمان نفس الدرجات لأداء المختبرين (Reynolds & Livingston, 2013)، ويطلق عليه عادة "اتفاق المحكمين"

3-العوامل المؤثرة على ثبات درجات الاختبار:

-**طول الاختبار:** ثبات درجات الاختبار له علاقة بعدد البنود، فكلما زاد عدد بنود الاختبار زاد معامل الثبات، فطول الاختبار أحد المظاهر التي تؤثر بالتأكيد على تباين الدرجة الحقيقية والدرجة الملاحظ.

فالاختبارات ذات العدد المحدود من الأسئلة (البنود) تكون أكثر تأثراً بعوامل الصدفة (كامل، 1996) فحل مسألة عن طريق الصدفة في اختبار يشتمل على ثلاث مسائل أو أربعة يؤثر على الدرجة الكلية للاختبار، لكن إذا زاد عدد المسائل (البنود) إلى أربعين أو خمسين فإن حل مسألة واحدة بالصدفة لا يظهر لها تأثير واضح وكبير على الدرجة النهائية التي يحصل عليها، وطور Spearman-Brown معادلة تسمح لنا بالتحكم في ثبات الاختبار عن طريق زيادة عدد البنود، والصيغة العامة لهذه المعادلة هي:

$$\rho_{xx'} = \frac{k \times \rho_{jj'}}{1 + (k-1) \rho_{jj'}} -$$

بحيث: $\rho_{xx'}$ = عن الثبات المرغوب للاختبار المعدل

$\rho_{jj'}$ الثبات الأصلي للاختبار، و k : عدد مرات زيادة عدد البنود.

إنَّ حساب العلاقة بين طول الاختبار ومقدار الثبات يمكِّننا من التعرف بدقة على درجة الثبات التي يصل إليها الاختبار عندما نزيد عدد البنود بنسبة k . وجدير أن نتذكَّر أن زيادة عدد بنود الاختبار وجعله أطول ستترتب عليه زيادة في تكلفة كتابة البنود الإضافية والوقت المستهلك على تطبيق وتصحيح الاختبار، دون أن نغفل الحرص على أن تكون البنود التي نقوم بإضافتها موازية في المحتوى والصعوبة لبنود الاختبار الأصلي.

ولتحديد قيمة k وهي عدد مرات زيادة عدد بنود الاختبار نطبق المعادلة التالية.

$$k = \frac{\rho_{xx'}(1 - \rho_{jj'})}{\rho_{jj'}(1 - \rho_{xx'})}$$

$\rho_{xx'}$ = معامل الثبات المرغوب.

$\rho_{jj'}$ = معامل الثبات الأصلي.

زمن الاختبار: كثيرا ما يكون لعامل الزمن تأثير على درجة الثبات، فاختبارات السرعة الموقوتة بزمن محدد تتأثر درجات ثباتها وهذا ما أكدته أبحاث كل من Lindquist & Cook (السيد، 1989) كون هذا النوع من الاختبارات تظهر فيه نسبة معتبرة من البنود في آخر الاختبار يتركها المختبرين دون الإجابة عنها، وهذا ليس بسبب صعوبتها لكن بسبب عامل الزمن وتصحح عموما بدرجة صفر، وينجر عن هذه العملية ارتفاع في الارتباط بين البنود الأخيرة، ما يمنح اتساق أكبر لهذه البنود.

-**تقارب مستوى صعوبة البنود:** الاختبارات التي تتكون من بنود صعبة جدا أو سهلة جدا تكون درجات ثباتها ضعيفة، لأنها تؤدي إلى انتشار محدود للدرجات وتظهر فروقا طفيفة، كون تقدير الثبات يعتمد على تباين درجات الاختبار ومدى انتشارها.

-**موضوعية التصحيح:** تعد أحكام المصححين (الذاتية، التحيز) من العوامل المؤثرة على ثبات درجات خاصة في تقدير درجات اختبارات المقال، وبعض اختبارات الأداء أو مقاييس الشخصية لأن التصحيح

يتطلب أحكاما ذاتية حول استجابات المختبرين مما يؤثر تأثيرا كبيرا على ثبات التقديرات بحيث تنخفض قيمة معامل الثبات نتيجة ذلك.

-**تجانس عينة التصحيح:** يتأثر ثبات الاختبار بنوع عينة الأفراد التي يطبق عليه الاختبار (متجانسة أو غير متجانسة، فتجانس درجات المجموعة يؤدي إلى انخفاض درجة تشتتها وانحرافاتها عن المتوسط، وهذا بدوره يقلل من قيمة معامل الثبات الذي يقوم على حساب الانحرافات عن المتوسط صياغة البنود: البنود الغامضة والطويلة تؤثر على طريقة الإجابة، كذلك خلوه من الأخطاء يعزز ثبات الاختبار، فبعض البنود قد تشتمل على مؤثرات الإجابة مما ينجر عنها زيادة عامل التخمين في الحصول على الدرجة الصحيحة.

محاضرة رقم 6: صدق الاختبار

يعتبر الصدق من المعايير الفنية الرئيسية التي يقوم عليها بناء الاختبارات النفسية والتربوية، وقد مر مفهوم الصدق واستخداماته بعدة مراحل وتطورات، فالبدايات الأولى لاستخدام الصدق تركزت على تقييم ما إذا كان الاختبار يقيس ما أعدّ لقياسه، وهذا يعكس التعريف البسيط والكلاسيكي لمفهوم الصدق، غير أن التعريف الحديث يرتبط بالاستخدامات المختلفة لدرجات الاختبار وليس بالاختبار نفسه وعليه فإن الصدق هو صدق استخدام درجات الاختبار وليس صدق الاختبار. وقد عرّفه (Messick, 1988) على أنه "حكم تقييمي شامل، مبني على أدلة تجريبية ومبررات نظرية، على مدى كفاية وملاءمة الاستدلالات والإجراءات المستندة إلى درجات الاختبار" هذا المفهوم اعتمد فيما بعد من قبل وثيقة معايير العمليات الاختبارية النفسية والتربوية (AERA, APA, & NCIME, 1999) في اشتقاق التعريف التالي للصدق "على أنه الدرجة التي تؤيد بها الأدلة والنظرية تفسير درجات الاختبارات التي تتطلبها استخدامات مقترحة للاختبار". وعليه يمكن توضيح الصدق في النقاط التالية:

- أ- يشير الصدق إلى تفسير نتائج الاختبار وليس الاختبار نفسه.
- ب- يستنتج الصدق من الأدلة المتوفرة وليست المناسبة.
- ت- الصدق خاص باستخدام درجات الاختبار مثل استخدام الدرجات في اختيار عينة الأفراد أو توزيعهم أو تقويم أدائهم.
- ث- يعبر عن الصدق بدرجة وصفية مثل مرتفع أو متوسط أو منخفض وقد تستخدم الأرقام (معاملات الاتفاق أو الارتباط) لتوضيح الدرجة الوصفية.

وقبل التعرض لمصادر أدلة الصدق في ضوء وثيقة المعايير الصادرة من الجمعية الأمريكية للبحث في التربية، الجمعية الأمريكية لعلم النفس، المجلس القومي للتربية الأمريكي (AERA, APA, & NCIME, 1999) نستعرض أبرز التطورات التاريخية لمفهوم الصدق (النظرية الكلاسيكية والحديثة)، وهذا تقاديا لأي لبس يقع للباحثين ومطوري الاختبارات النفسية والتربوية بسبب اعتماد المراجع الصادرة قبل التحديث الأخير (AERA 1999) على التسميات الكلاسيكية لأنواع الصدق (صدق المحتوى، الصدق المرتبط بمحك، صدق التكوين الفرضي) ونذكر أن هذا التصنيف شائع الاستخدام إلى يومنا هذا. سيكون التطرق لنظرية الصدق الكلاسيكية والحديثة في ضوء المراجعات النظرية التي قدمها كل من (تيغزة، 2008) و (Reynolds & Livingston, 2013) و (Goodwin & Leech, 2003)

1- نظرية الصدق الكلاسيكية:

المرحلة الأولى: بدايات هذه النظرية كانت في النصف الأول من القرن العشرين، وتميزت هذه المرحلة بظهور أصناف عديدة وتسميات متباينة لأنواع الصدق، نجم عنه فوضى في التسميات (الصدق التنبئي، الصدق التلازمي الصدق الارتباطي...). طبعت هذه المرحلة بروز منظورين أساسيين على ميدان الصدق هما:

-الاتجاه الإمبريقي(التجريبي): والذي روج للصدق المحكي أو صدق الارتباط بالمحك، الصدق العملي.
-الاتجاه المنطقي: يعتبر المنطق والأحكام الفردية (الخبرة) كأطر مرجعية لتقدير الصدق. صدق المحتوى
انبثق عن هذا الاتجاه.

المرحلة الثانية: تميزت هذه المرحلة بصدور الدليل الإرشادي والفني للقياس والاختبارات والذي جاء نتيجة التغيرات التي طرأت على مفهوم الصدق في أوساط الباحثين وكذا تنويعا لانضمام أعضاء الجمعية الأمريكية لعلم النفس التطبيقي إلى الجمعية الأمريكية لعلم النفس.

-1954 صدور الدليل الأول (AERA, APA) بعنوان "التوصيات الفنية للاختبارات النفسية وفنيات أدوات التشخيص" قدم هذا الدليل تصنيف رباعي لأنواع الصدق:

أ- صدق المضمون.

ب- الصدق التنبئي.

ت- الصدق التلازمي.

ث- صدق التكوين الفرضي أو المفهوم.

يعتبر صدق المفهوم إضافة جديدة في ميدان القياس في تلك الفترة، وكان وراء هذا المفهوم وتوضيح دلالاته فيما بعد كل من Cronbach و Meehl 1954.

- 1966 صدور الدليل الإرشادي الثاني المنقح (AERA, APA): قلص هذا الدليل التصنيف الرباعي للصدق إلى تصنيف ثلاثي:

أ- صدق المضمون.

ب- صدق المحك.

ت- صدق المفهوم.

ظهر في هذا التصنيف تسمية صدق المحك كتسمية جديدة تضم (الصدق التنبئي، والصدق التلازمي).
جدير أن نشير إلى أن هذا التصنيف الثلاثي لأنواع الصدق هو المنتشر في ميدان التأليف وتطوير الاختبارات، ولازال يستخدم ويعتمد بشكل واسع في ميدان التدريس أيضا.

- 1974 الدليل الإرشادي الثالث للقياس: هذا الدليل يشير إلى أنواع الصدق باستخدام لفظ "جوانب الصدق" وعلى هذا فالصدق ليس صفة أو خاصية تقاس قياسا مباشرا، وإنما نستنتج الصدق استنتاجا من دراسة جوانب أو أبعاده الثلاثة، وهذه الأبعاد مترابطة فيما بينها من الناحية الوظيفية والمنطقية، وبغية تسهيل مناقشتها عمليا يمكن دراستها على انفراد.

- 1985 الدليل الإرشادي الرابع: أبرز توجهات هذا الدليل اعتبار الصدق وحدة واحدة لا تتجزأ، وتبني تعريف جديد للصدق يحمل تغييرا في المفهوم والاستخدام "الصدق يدل على ملاءمة ودلالة وجدوى الاستنتاجات الخاصة المستمدة من الدرجات" واستعمال لفظة "أدلة الصدق" (أدلة صدق المحتوى، أدلة صدق المحك، أدلة صدق المفهوم) بدل أنواع الصدق أو جوانب الصدق التي كانت متداولة سابقا. كما

أبقى على التصنيف الثلاثي لأنواع الصدق (تيغزة، 2008). هذا الدليل كشف عن توجهات جديدة في تقدير الصدق والذي يعتمد على تجميع البيانات أو الأدلة وتراكمها لتدعيم الاستنتاجات المستمدة من درجات الاختبار، فهنا يتم تقدير صدق الاستنتاجات التي تكشف عن أوجه استعمال معينة للاختبار وليس تقدير صدق الاختبار ذاته.

2- نظرية الصدق الحديثة:

التمعن في تطور مفاهيم الصدق وتصوراته واستخداماته عبر المراحل التاريخية التي استعرضناها سابقاً، يجعلنا ننتبه إلى أنّ بوادر التغيير بدأت تظهر في الإصدار الأول لدليل معايير العمليات الاختبارية النفسية والتربوية (1954) إلى غاية الإصدار الخامس (1999) والذي حمل تطورات حدثت على مستوى مفهوم الصدق والذي أشرنا إليه في التعريف المقدم للصدق، وكذا اعتبار مفهوم الصدق موحد ولا يتجزأ إلى أقسام أو أنواع. ويعتبر 1994 Messck من مهندسي فكرة وحدة الصدق، وأن هذه الوحدة يتضمنها صدق المفهوم والذي يعنى بدمج متكامل لمصادر أدلة الصدق التي تستهدف تفسير وتأويل وتبيين دلالة درجات الاختبار، وهذه النظرة الوحدوية للصدق لا تتنافى مع وجود أدلة متميزة في تصنيف مصادر أدلة الصدق وطبيعتها التي صنّفها إلى ست جوانب، ولخصها فيما بعد الدليل الإرشادي الخامس في خمسة مصادر لأدلة الصدق والتي نركز عليها في عرض مصادر أدلة الصدق في ضوء النظرية الحديثة لصدق الاختبارات.

*مصادر أدلة الصدق:

-أدلة تعتمد على محتوى الاختبار:

وثيقة المعايير ذكرت بأنه يمكن الحصول على أدلة قيمة بفحص العلاقة بين محتوى الاختبار والتكوين الفرضي (النطاق الذي صمم الاختبار من قياسه. فصدق المحتوى نحصل عليه من خلال من التحليل المنطقي / التقديري للبنود وصيغة المقياس أو الاختبار وحسب (Goodwin & Leech, 2003) يستند هذا النوع من أدلة الصدق على التحليلات المنطقية وتقييمات الخبراء لمحتوى الاختبار، والتي تشمل البنود من حيث المهام والصيغة والصياغة والعمليات المطلوبة من المختبرين. بشكل عام، فإنه يعالج أسئلة حول المدى الذي يمثل فيه محتوى القياس نطاق محتوى محددًا.

أنشطة التحقق من الصدق حسب (Goodwin & Leech, 2003)

1. القيام بتحليلات منطقية ومراجعة الخبراء لفحص مدى تمثيل محتوى الاختبار لنطاق المحتوى (التناظر بين محتوى الاختبار والتكوين الفرضي)
- 2 التحليلات المنطقية ومراجعة الخبراء لفحص مدى ملائمة البنود أو المهام أو الأجزاء الفرعية من المقياس لتعريف التكوين الفرضي / أو الغرض من التدبير.
- 3 التحليلات المنطقية ومراجعة الخبراء لفحص نجاعة وأهمية ووضوح البنود، وابتعادها عن التحيز.

4 - التحليلات المنطقية ومراجعة الخبراء لفحص ما قد ينتج عنه نقص التمثيل أو الجوانب غير ذات الصلة من القياس لمزايا غير عادلة لمجموعات فرعية واحدة أو أكثر من المحبيين. وقد أورد (Holmbeck & Devine, 2009) منهجا أكثر تفصيلا لجمع أدلة صدق المحتوى. يمكن الرجوع للتفصيل أكثر. نلاحظ أن جميع أدلة صدق المحتوى في عاداتها تكون نوعية لكن هذا لا يمنعنا من اتخاذ خطوات لتقرير النتائج بطريقة أكثر كمية ففي هذا السياق أشار (Reynolds & Livingston, 2013) إلى أنه يمكن للباحثين تقرير عدد الخبراء ومؤهلاتهم، عدد مرات مراجعة التقييم وتعليقاتهم عليه، ودرجة اتفاقهم على قضايا تتعلق بالمحتوى، ولأثر هذه الجهود نعتمد معادلة Lawche1975 التي تعكس درجة الاتفاق بين الخبراء الذين يصدرون احكاما تتعلق بالمحتوى، وكذا تحليل الموازين متعددة الأبعاد.

-أدلة تعتمد على عمليات أو سيرورات الاستجابة:

يذكر Whiston2009 أن أدلة الصدق التي تستند على عمليات الاستجابة تركز على ما إذا كانت طريقة أداء واستجابة المفحوص تتوافق مع التكوين موضع القياس، مثلا في تقويم ما يحاول قياس الاهتمام المهني، سوف يفحص مطورو الاختبار ما إذا كان الأفراد يجيبون على البند بناءً على ما يرغبون القيام به بدلا من ما يمكنهم فعله" (Whiston, 2009) فمدى التوافق بين الأنشطة العقلية والانفعالية التي يتكون منها المفهوم، وبين الأنشطة العقلية والانفعالية التي يمارسها الفرد عند إجابته على فقرات الاختبار تمثل إحدى المصادر الهامة لبيانات الصدق. ومن الإجراءات التي يمكن أن تتيح هذا النوع من البيانات عن الصدق ما يلي:

-تحليل استجابات المفحوصين بطريقة التفكير المسموع أثناء حل المشكلة، أو يصفوا بعد حل المشكلة العمليات والاستراتيجيات التي استخدموها في الوصول إلى الحل، وذلك بمعرف العمليات (المهارات والاستراتيجيات) التي وظفها المفحوصون أثناء أدائهم للمهام التقييمية لمطابقتها مع العمليات التي يحددها التكوين الفرضي المراد قياسه.

-فحص المنطق الذي قامت عليه إجابات الأفراد، وكذا طرق إجاباتهم (تحليل أسبابها ودواعيها) على بنود الاختبار، وذلك بتحليل أوجه التفسير أو التعليقات التي يحررها المختبرين، أو يصرحون بها شفويا، لتعليل المنطق والطرق التي استعملوها.

-بالنسبة للتقويم البديل الذي يتبنى فقرات أو مهمات أو أنشطة تقييمية واقعية، فلا بد من الاطمئنان بأن المقدرين يطبقون محكات تقدير الدرجات بطريقة دقيقة، وأت تقديراتهم لأداء المختبرين لم تتأثر بعوامل خارجية كالتحيز، أو تأويل محكات التصحيح تأويلا خاطئا (تيغزة، 2008).

-إجراء مقابلة مع المختبرين لمعرفة عمليات واستراتيجيات استجاباتهم، وتدوين مؤشرات سلوكية، مثل: زمن الاستجابة، أو حركات العين، أو حتى تحليل أنواع الأخطاء التي يقعون فيها.

-أدلة تعتمد على البنية الداخلية:

فحص البنية الداخلية للاختبار معين تحدد ما إذا كانت العلاقات بين مفردات الاختبار وعناصره متسقة مع التكوين الفرضي المقترح تفسير درجاته (AERA, APA, & NCIME, 1999) فحسب هذه الوثيقة فإن تحليلات البنية الداخلية للاختبار يمكن أن تشير إلى مدى تطابق العلاقات بين بنود الاختبار ومكوناته مع التكوين الفرضي الذي تقوم عليه تفسيرات الدرجات. وهناك عدة طرق كمية لتحليل البيانات:

-تقدير الارتباط بين البنود والدرجة الكلية للمقياس.

-تقدير الارتباط بين مجالات التكوين.

-التحليل العاملي الاستكشافي.

-التحليل العاملي التوكيدي.

-الاتساق الداخلي للبنية الداخلية للاختبار.

المظهر الرئيس لهذا النوع من أدلة الصدق هو أن الباحثين يفحصون إمبريقيا البنية الداخلية للاختبار، ويقارنوا ذلك ببنية التكوين الفرضي موضع القياس، هذا النوع من أدلة الصدق تم تضمينه تقليديا تحت عنوان صدق التكوين الفرضي (Reynolds & Livingston, 2013).

-أدلة تعتمد على العلاقات بمتغيرات أخرى أو أدلة البنية الخارجية:

نحصل على أدلة صدق البنية الخارجية للاختبار أو المقياس من خلال فحص ودراسة علاقة درجات الأداة موضع الاهتمام بدرجات متغيرات خارجية (مقاييس، أو محكات)، ودراسة ما إذا كان مستوى هذه العلاقات وطبيعتها أو اتجاهها ينسجم ودلالة المفهوم المقاس. ويندرج ضمن هذا النمط معظم التسميات الكلاسيكية للصدق (الصدق التقاربي والتمييزي، الصدق المحكي سواءً التلازمي أو المحكي)، وأهم إجراءات الحصول على أدلة الصدق القائمة على البنية الخارجية ما يلي (Goodwin & Leech, 2003) (تيعزة، 2008) وهذه الأدلة وردت في التصنيفين الصادرين عن (AERA, APA, NCIME, 1985, 1999):

1 - الدراسة الارتباطية لقوة واتجاه العلاقة بين درجات الاختبار المراد تقدير صدق نتائجه وبين متغيرات خارجية تكون محكية، ويعرف هذا النوع من الصدق، بالصدق التلازمي في النظرية الكلاسيكية.

2. الدراسة الارتباطية لمعرفة مدى قدرة الدرجات التي تم الحصول عليها من المقياس على التنبؤ مع المحكات الخارجية التي يتم قياسها في وقت لاحق.، ويعرف هذا الإجراء في نظرية الصدق الكلاسيكية بالصدق التنبؤي. فالدراسات التلازمية تكون ملائمة بدرجة كبيرة عندما يكون الغرض من الاختبار تحديد الوضعية الراهنة للفرد المختبر ن في مقابل التنبؤ بنتائج مستقبلية (Anastasi & Uraina, 1997) ويستخدم في تقدير علاقة درجات الاختبار بمحك طرق متعددة تعتمد على معاملات الارتباط بين الاختبار التنبؤي واختبار المحك، ومعادلة الانحدار في التنبؤ بدرجات الاختبار المحك، إضافة إلى طريقة جداول التوقع، وطريقة نظرية المنفعة.

3-دراسة قوة العلاقة الارتباطية، واتجاهها بين درجات المقياس بالمتغيرات الأخرى (كأن تكون اختبارات أخرى تقيس المفهوم ذاته) التي ينبغي أن يرتبط بها ارتباطا قويا بناء على المنطق النظري الذي قام عليه المقياس. وسمي هذا النوع من البيانات في نظرية الصدق الكلاسيكية بالصدق التقاربي.

4-دراسة قوة العلاقة الارتباطية، واتجاهها بين درجات المقياس بالمتغيرات الأخرى (كأن تكون اختبارات أخرى تقيس مفاهيم أخرى مختلفة) التي ينبغي أن يرتبط بها ارتباطا ضعيفا (قوة العلاقة ضعيفة لا يعتد بها رغم وجود دلالة إحصائية أحيانا) أو ارتباطا غير دال إحصائيا، بناء على المنطق النظري الذي قام عليه المقياس. وسمي هذا النوع من البيانات في نظرية الصدق الكلاسيكية بالصدق التمايزي.

5-طريقة تعدد السمات وتعدد الطرق استخدمت في النظرية الكلاسيكية للدلالة على صدق المفهوم، أو الصدق التقاربي والصدق التمايزي.

6-الدراسة التقييمية لفعالية عملية الانتقاء المهني وقرارات الوضع أو التسكين.

7- تقدير مدى موافقة المحكات وجودتها، بما في ذلك تقدير ثبات درجاتها وصدقها، والصعوبات العملية عند استعمالها، وإمكانية توفرها. وتتأتى أهمية هذه المحكات من إمكانية الحكم على صدق الأداة بناء على مدى ارتباطها بها.

-أدلة تعتمد على عواقب العملية الاختبارية:

هذا النوع من أدلة الصدق يتعلق بالنتائج المتوقعة وغير المتوقعة للقياس على حد سواء الإيجابية والسلبية. وهذه البينة لم تذكر سابق في دليل معايير العمليات الاختبارية النفسية والتربوية إلا في إصدار 1999. متأثرا بأعمال (Messick 1989) وتُدعى أدلة عواقب العملية الاختبارية أحيانا بصدق النتائج أو المترتبات.

"عادة يتم بناء المقياس بأمل تحقيق بعض المكاسب من استخدام الدرجات، من بين المكاسب الممكن تحقيقها اختيار العلاجات الفعالة لمرض ما، وضع الموظفين في الوظائف المناسبة، منع الأفراد غير المؤهلين من الوصول إلى وظائف، أو تحسن الممارسات التدريسية الصفية. الهدف الأساسي لتقدير صدق درجات المقياس تحديد ما إذا كان تحقيق هذه الفوائد أو المكاسب المنتظرة من عملية القياس يمكن تحقيقها" (AERA, APA, & NCIME, 1999)

هناك العديد من الباحثين والمدرّبين في دورات القياس، يرحبون بالتغيرات الطارئة في دراسة ومعالجة الصدق في ضوء المعايير الجديدة (أدلة الصدق). إن النظرة الثلاثية الكلاسيكية للصدق الواردة في كتب القياس لعدة سنوات، تعد مشكلة لعديد من الأسباب المترابطة. أولا تقسيم الصدق لثلاثة أنواع أو أربعة يشكل مصدر خلط بين صدق التكوين الفرضي والأنواع الأخرى للصدق لدى القائمين على دراسة صدق الاختبار، حيث يتولد لديهم قلق بشأن قدرتهم على التحديد الصحيح لنوع الصدق الملائم. وبذلك يفقدون النظرة الكلية والتكاملية للصدق وتضيع جهودهم في السعي إلى الفرز بين أنواع الصدق بدل البحث في قضايا مهمة متعلقة بالصدق. المشكلة الثانية في النظرة الكلاسيكية للصدق هي أنه يشجع على اتباع نهج

"قائمة مرجعية" للمراجعة عند قيام الباحثين بتجميع أدلة الصدق الموجود في دليل الاختبار أو المقالات البحثية. ثالثاً، يميل مخطط التصنيف الثلاثي للصدق إلى الترويج للاعتقاد الخاطئ بأن الصدق هو خاصية ثابتة للمقياس أو الاختبار، وليس خاصية لعينة (Tracey & Glidden-Tracey, 1999). تؤدي هذه المشكلة إلى استخدام تعبيرات غير مناسبة للصدق على سبيل المثال استخدام عبارات مثل، "هذا الاختبار صادق"، وقد ثبت أن هذا الاختبار صادق،" لقد أثبت صدق هذا الاختبار"، وهكذا دواليك. رابعاً، هذه الطريقة في التفكير في الصدق تتناقض مع المفاهيم التي مفادها أن الصدق هي مفهوم وحدوي وأن صدق التكون هي في الواقع "كل متكامل، وأخيراً، فإنه يسمح بشكل شمولي للغاية للباحثين ومطورو الاختبارات بالنظر لجميع أنواع الصدق التقليدية على نفس القدر من الأهمية في جمع مجموعة من الأدلة لدعم الاستدلالات التي سيتم إجراؤها من نتائج الاختبارات.

3- العلاقة بين الصدق والثبات:

يرتبط مفهوم الصدق بمفهوم الثبات ارتباطاً وثيقاً، حيث لا يمكن أن تتصف درجات الاختبار بالصدق في استخدامات معينة دون أن تكون درجاته متسقة، كما أن مؤشر الثبات يعدّ الحد الأعلى لقيمة معامل صدق درجات الاختبار. وتؤثر قيم معامل الثبات في درجة العلاقة بين الاختبار التنبؤي واختبار المحك، كما تسهم قيمة معامل استقرار درجات الاختبار في زيادة الثقة بصدق التكوين الفرضي للاختبار لأنه يهتم بالثبات النسبي للسمات.

كما يقدم معامل الاتساق الداخلي دليلاً آخر على صدق التكوين الفرضي، لذلك يرى كرونباخ وزملائه Cronbach, Glaser & Rajaratnam أن الصدق والثبات مفهومان مترابطان ويمكن أن يندرجا تحت اسم "مقاييس إمكانية التعميم" فالفرق الرئيسي بينهما يكمن في الأبعاد التي نريد التعميم عليها. لذلك تعدّ نظرية إمكانية التعميم التي قام بصياغتها كرونباخ وزملائه من التطورات المعاصرة المهمة التي ساهمت في إبراز التكامل بين مفهومي الصدق والثبات.

محاضرة رقم 7: معايير الاختبارات ومحكات الأداء

يعد مفهوم معايير الاختبارات من المفاهيم الأساسية المتعلقة بتفسير درجات الاختبارات مرجعية الجماعة أو المعيار، فالدرجة التي يحصل عليها فرد في اختبار ما والتي تسمى الدرجة الخام لا يكون لها معنى ويصعب تفسيرها ما لم يتم اسنادها إلى نظام مرجعي. فهذا النظام هو الذي يسمح باستخلاص معلومات مفيدة من درجات الاختبار.

والنظام المرجعي المناسب لهذا النوع من الاختبارات يعتمد على استخدام معلومات يتم الحصول عليها من جماعة معينة تسمى الجماعة المرجعية، وهذه الجماعة تكون خصائصها محددة ومعلومة لمستخدمي الاختبارات، لذلك تسمى الاختبارات التي تستخدم هذا النظام الاختبارات مرجعية الجماعة أو المعيار. وهناك نظاما آخر لا يعتمد على جماعة مرجعية، وإنما يعتمد على محك أداء متوقع، ويتحدد هذا المحك في ضوء الكفايات التي تشتمل على مهارات ومعارف متوقعة ينبغي أن يمتلكها الفرد، ونظرا لأن تفسير درجات الاختبارات في هذه الحالة يستند إلى هذا المحك فإن هذه الاختبارات عندئذ تسمى بالاختبارات مرجعية المحك. وفيما يلي تفصيل لهذه المحكات:

أولاً: معايير الاختبارات مرجعية الجماعة: المعايير عبارة عن مجموعة من الدرجات المشتقة بطرق إحصائية معينة من الدرجات الخام، وقبل أن نوضح أنواع المعايير التي تستخدم في تفسير الدرجات ينبغي أن نلقي الضوء على أنواع الجماعات المرجعية وخصائصها وكيفية انتقاؤها التي يستند في التوصل إلى هذه المعايير:

1.1 **أنواع الجماعات المعيارية:** بما أن معايير الاختبارات مرجعية الجماعة أو المعيار تعتمد على الجماعة المعيارية التي نستمد منها هذه المعايير. ينبغي أن تكون خصائص هذه الجماعة محددة تحديدا جيدا مما يمكننا من التفسير الدقيق لدرجات الاختبار. إضافة إلى هذا يجب أن تكون خصائص الجماعة المعيارية مشابهة لخصائص الأفراد الذين سيطبق عليهم الاختبار فيما بعد من أجل تقويم السمة التي يقيسها لديهم في ضوء مركزهم النسبي.

ومن هنا نطرح التساؤل التالي: ما هي الخصائص التي ينبغي، تتوافر في الجماعة المرجعية؟ والحقيقة أن هذه الخصائص تتباين بحسب المعايير المستخدمة، فالمعايير الوطنية تختلف عن المعايير المحلية وكلاهما يختلف عن المعايير المتعلقة بموقع واحد، مثل مدرسة معينة وكذلك معايير فئات خاصة، ونوضح فيما يلي هذه المعايير:

أ- **المعايير الوطنية:** تفيد هذه المعايير في تقويم البرامج التعليمية باستخدام الاختبارات التحصيلية المقننة، وذلك للتحقق مما إذا كان أداء طلاب نهاية مرحلة تعليمية معينة أو فرقة دراسية أعلى أو أدنى من المتوسط على المستوى الوطني (تقوم مؤسسات مختصة بإعداد هذه الاختبارات نظرا لتكاليها ووقتها وجهدها). فالجماعة المعيارية الوطنية يفترض أنها بمثابة عينة ممثلة ممثلة لطلاب

الوطن في مرحلة تعليمية معينة، يؤخذ في الاعتبار أثناء انتقائها الخصائص الديمغرافية المختلفة ، مثل المناطق الجغرافية ، الفرق الدراسية ، حجم المدرسة و موقعها ، الحالة الاجتماعية و الاقتصادية و النوع و التركيبة الطلابية ، وفيما بعد يتم سحب أو اختيار العينة باستخدام أساليب المعاينات العشوائية أو المعاينات العشوائية الطبقيّة التي تراعي الخصائص أو المتغيرات السابقة .و يجب أن تكون العينة المنتقاة كبيرة بدرجة كافية لضمان ملاءمة المعايير للمجتمع المستهدف و تقليل أخطاء المعاينات.

ب- **المعايير المحلية:** وتسمح هذه المعايير بمقارنة أداء الطالب بأقرانه في المنطقة التي يعيش فيها، ولا تسمح بالمقارنة خارج هذه المنطقة.

ت- **معايير مؤسسة أو مدرسة معينة:** تفيد هذه المعايير في مقارنة متوسط أداء طلاب فرقة دراسية معينة في إحدى المدارس لمتوسط أداء الفرقة المناظرة في مدارس أخرى.

ث- **معايير فئات خاصة:** تحتاج بعض الاختبارات إلى معايير لفئات خاصة من ذوي الاحتياجات العقلية أو الجسمية، أو لفئة من الأفراد.

2-1- خصائص عامة للجماعة المعيارية : على الرغم من اختلاف الجماعات المرجعية تبعاً لطبيعة هذه الجماعات ومدى اتساع عيناتها، إلا أنه لا ينبغي أن تتوافر فيها بعض الخصائص العامة لكي تكون هذه الجماعة مناسبة للمعايير التي تستمد منها، ومن أهم هذه الخصائص التي حددتها الجمعية الأمريكية لعلم النفس :

- تمثيل عينة الأفراد للمجتمع المستهدف: عند انتقاء عينة من الأفراد من أجل بناء المعايير ينبغي مراعاة أن تكون هذه العينة ممثلة تمثيلاً كافياً للمجتمع المستهدف (ويفضل هنا استخدام العينات العشوائية الطبقيّة النسبية).

حدثة المعايير: معايير الاختبارات تتغير بتغير خصائص الأفراد بمرور الزمن، وبخاصة إذا كانت تقيس خصائص تتأثر بما يطرأ على البيئة، مثلاً المفردات اللغوية تنمو وتتغير دلالتها ... فالجماعة المعيارية التي استمدت منها معاير اختبار معين منذ سنوات لا تصبح مناسبة للاختبار في الوقت الحاضر.

مجالات القياس

نتطرق في هذا المحور لأهم مجالات القياس النفسي والتربوي خاصة الرئيسية منها، مستعرضين فيه التراث النظري في كل مجال وأهم تطورات قياسه، وخطوات تطوير أدوات القياس.

محاضرة رقم 8: قياس الذكاء (مفهومه وتطورات قياسه)

2- مفهوم الذكاء : على الرغم من وجود تاريخ طويل من الأبحاث والنقاشات حول الذكاء، إلا أنه لا يوجد تعريف موحد للذكاء حتى الآن. وقد دفع هذا بالبعض إلى الاعتقاد بأن الذكاء يمكن وصفه بشكل تقريبي فقط، ولا يمكن تعريفه بالكامل أو بشكل دقيق. لكن ألا يمكن اعتبار هذا التوجه جد متشائم، فلو نتفحص عديد التعريفات المصاغة فستتظاهر لنا تشابهات واضحة وقوية بينها، ففي واقع الأمر هذه التعريفات حين يتم تفسيرها بشكل مناسب نجدها تقول نفس الشيء ولكن بصياغات أو بمصطلحات مختلفة. وقد أورد الباحثان (Legg & Hutter, 2007) سبعين تعريفا وردوا في مختلف المراجع والدوريات البحثية في مقال بعنوان " مجموعة من تعاريف الذكاء " وهي الأكثر شيوعا وانتشارا وراع أن تكون التعريفات من جهة المنظمات أو القواميس والموسوعات وكذا من جهة الأخصائيين النفسانيين. وسوف نكتفي بالتعريفات التالية المقتبسة من مختلف المراجع:

- "الذكاء هو القدرة على اكتساب المعرفة والمهارات وتطبيقها" (Soane & Hawke, 2008)

- القدرة على التعلم أو الفهم أو التعامل مع حالات جديدة أو صعبة... أو محاولة الاستخدام الماهر للعقل.
- القدرة على تطبيق المعرفة للتحكم في البيئة أو التفكير بشكل تجريدي وفق معايير موضوعية (مثل الاختبارات). Merriam-Webster Online Dictionary, 2006.

-تعريف Wechsler "القدرة الكلية للفرد على العمل الهادف، والتفكير المنطقي والتفاعل الفعال مع البيئة"
-تعريف Stoddard "الذكاء هو القدرة على القيام بنشاطات تتميز بالصعوبة، والتعقيد والتجريد، التكيف للوصول للهدف، والقيمة الاجتماعية والابتكارية، والاقتصاد في الجهد والوقت. والقدرة على الاستمرار في الظروف التي تتطلب الطاقة العقلية ومقاومة العوامل الانفعالية"

حسب علماء النفس صياغة تعريف بسيط وشامل للذكاء يقبل به الجميع ليس أمرا سهلا، فقد تعددت التعريفات واختلفت باختلاف المفهوم الذي يكونه كل واحد منهم، ويُرجع Sattler غموض مفهوم الذكاء وصعوبة الدقة في تحديده إلى كون الذكاء صفة وليس كينونة، بمعنى أنّ الذكاء لا وجود له في حد ذاته وإنما هو نوع من الوصف نعت به فردا معينا عندما يسلك بطريقة معينة في وضع معين. (محمد عامر والمصري، 2018)

وفي إجابة على سؤال طرحه (Santrok, 2014) حول ماهية الذكاء، أورد ما يلي:

-القدرة على حل المشكلات والتكيف مع تجارب الحياة اليومية والتعلم منها.

-القدرة على حل المشكلات.

-القدرة على التكيف والتعلم من الخبرات.

- يتضمن خصائص مثل الإبداع ومهارات التعامل مع الآخرين.
- القدرة العقلية التي تمكن الشخص من التكيف مع، أو تشكيل، أو اختيار بيئة الشخص.
- القدرة على الحكم، الفهم، والعقل.
- القدرة على فهم الأشخاص والأشياء والرموز والتعامل معهم.
- القدرة على التصرف بشكل هادف، والتفكير بعقلانية، والتعامل بفعالية مع البيئة.

2-لمحة تاريخية حول تطورات قياس الذكاء: نستعرض مختلف تطورات قياس الذكاء عبر قرن من الزمن في ضوء القراءات لمختلف المراجع التي اهتمت بالموضوع (علام، 2000) (طه، 2006) يعود الاهتمام بالذكاء والقدرة العقلية المختلفة إلى العصور القديمة، فقد أثارت الفروق بين الناس في هذه الجوانب وانعكاساتها على مختلف مناحي الحياة انتباه المجتمعات. ولعل أول قياس للذكاء والقدرة يرجع إلى الصينيين القدامى (2000 سنة قبل الميلاد) أين تم استخدام مجموعة من الاختبارات من طرف السلطات المحلية في الأقاليم المختلفة بتوجيه من الامبراطور لاختيار أفضل المرشحين للوظائف المختلفة على المستوى المحلي والقومي (تقيس مختلف الاستعدادات بدءاً من مستوى القراءة والكتابة، مستوى مهارة القراءة والكتابة، المهارة في العمليات الحسابية والمعرفة بالقانون).

وتقريباً في نفس الفترة الزمنية قدم أفلاطون في محاورة الجمهورية نظرية حول القدرات العقلية، التي تقترض تقسيم الناس إلى ثلاث فئات (الحكام -الفلاسفة -وهم أعلى في سلم القدرات العقلية، يليهم القادة العسكريون ويأتي الفلاحون والعمال في أدنى السلم) وحسب أفلاطون هذا التقسيم محدد وراثياً، وغير قابل للتغيير. كما ربط أرسطو بين الوظائف العقلية والمخ، وقدم تصوراً للذكاء يقوم على افتراض أنه يتكون من ثلاث جوانب: الجانب الأول نظري، يتعلق بفهم الجوانب المجردة، أما الثاني فهو عملي يتصل بالمهارة في الحياة العملية، وأجانب انتاجي يتعلق بالقدرة على الابداع والابتكار.

أما في الحضارة الإسلامية فأكثر المحاولات في دراسة الذكاء مباشرة هي دراسات "ابن الجوزي" (1110-1192) وذلك في كتابه "الأذكياء"، حيث ناقش ابن الجوزي في كتابه قضايا البيئة والوراثة والعلاقة بين التكوين الجسمي والذكاء وأثر المناخ في الذكاء، وظهر اهتمام ابن الجوزي أكثر بجانبين في دراسته للذكاء هما الذكاء اللفظي والذكاء العملي.

وكان الفكر الغربي حتى أواخر القرن الثامن عشر خالياً من التفكير في طبيعة الذكاء كما يعرف اليوم، أو حتى كما عرفته الحضارات السابقة. وأول إرهابات دراسة طبيعة الذكاء حدثت في أواخر القرن الثامن عشر وأوائل القرن التاسع عشر على يد اثنين من الأطباء النفسيين الفرنسيين (Jean-Étienne Esquirol 1772-1840) و (Jean Marc Itard 1775-1838) ويعتبر Esquirol أول من ميز بين المرض العقلي والتخلف العقلي وحاول أن يضع تصنيف لفئات التخلف العقلي بدلا من اعتباره ظاهرة كلية متجانسة، وأشار أيضا إلى أهمية الجوانب اللغوية في الذكاء واعتبر القصور اللغوي أهم مؤشرات الضعف العقلي. أما Itard فقد

أسس لما يسمى بالتربية الخاصة، في حين عمل Edward Seguin على اعتماد منهج التدريب الحسي للأطفال المعاقين عقليا في فرنسا وطور اختبار غير لفظي للذكاء.

وتبقى هذه الجهود متفرقة موجّهة لمساعدة بعض ذوي الحاجات الخاصة في إطار طبي ولا يعتبر قياسا منظما للذكاء، كما شهد القرنان الثامن والتاسع عشر ظهور بعض المحاولات لدراسة قدرات الانسان من خلال استخدام مؤشرات التكوين الجسماني في قياس الذكاء والمعروفة بنظرية الفراسة (phrénologie) على يد (F.Gall 1758-1820). والذي حاول إيجاد علاقة بين نتوءات الجمجمة وارتفاع أو انخفاض المناطق المختلفة بها من جهة، والذكاء وسمات الشخصية من جهة أخرى لكنها سرعان ما تعرضت للنقد واعتبرت نوعا من الدجل والاحتيال وليس علما حقيقيا. وتلتها جهود (Samuel Morton 1799-1851) في الولايات المتحدة الأمريكية و Paul Broca 1824-1880 في فرنسا والذي حدد المنطقة الدماغية المسؤولة عن فهم اللغة، وأسس لما يعرف آنذاك علم قياس الجمجمة أو إيجاد علاقة بين حجم الجمجمة ووزن المخ من ناحية والذكاء من ناحية أخرى. تعرضت هذه المحاولات أيضا للنقد واعتبرت إساءة في قياس الانسان وتحيز مسبق يعكس الأوضاع الاجتماعية والسياسية السائدة آنذاك.

الدراسة العلمية الحديثة للذكاء والقدرات العقلية بدأت في أواخر القرن التاسع عشر وأوائل القرن العشرين على يد Francis Galton 1848 الذي أسس ما عرف فيما بعد بالمدخل السيكو فيزيقي Psychophysical في قياس الذكاء و Binet et Simon 1905 صاحب المدخل السيكو متري في قياس الذكاء. حيث طور Galton مجموعة متنوعة من الاختبارات (17 مقياسا) لقياس الوظائف العقلية الأساسية مثل: التمييز الحسي وسرعة الإدراك، مسترشداً بافتراض أن الاختلافات في القدرة العقلية تنجم عن الكفاءة التفاضلية للدجاج العصبي المركزي. وأول من استخدم مفهوم الاختبار العقلي «Teste Montale» عام 1890 واعتبر القدرة العقلية تنعكس في التمييز الحسي وزمن رد الفعل وغير ذلك من الوظائف النفسية البسيطة.

في بداية القرن العشرين عام 1905 نُشر أول مقياس للذكاء من طرف Binet et Simon يشتمل على 30 مشكلة مرتبة تصاعدياً-حسب درجة صعوبتها- وتتطلب القدرة على الحكم والفهم، والتعليل. طُور المقياس لتمييز الأطفال الذين يعانون من صعوبات فكرية ولم ينجحوا في النظام المدرسي العمومي بفرنسا لوضعهم في صفوف دراسية خاصة بدل التمدريس مع العاديين، وخضع الاختبار إلى مراجعات في عام 1908 و1911 على يد واضعيه ليحتوي على عدد أكبر من الاختبارات الفرعية، وأصبح قياس الذكاء يستند على قياس الوظائف العقلية العليا بدلا من الوظائف الحسية البسيطة. كما اقترح هذا المقياس مفهوم العمر العقلي (Age Montale) والذي يحدد أداء الفرد بالنسبة لأقرانه في نفس المجموعة العمرية استنادا إلى درجة المقياس، ويتم تحديد العمر العقلي عن طريق طرح العمر الزمني من العمر العقلي (العمر العقلي - العمر الزمني) فإذا كان العمر العقلي للطفل أقل من العمر الزمني فإن ذكاء الطفل يعتبر أقل من المتوسط ويوزع على فصول التربية الخاصة. وفي العام 1912 اقترح العالم النفساني الألماني Stern William

نسبة الذكاء، من خلال قسمة العمر العقلي على العمر الزمني وضرب الناتج في 100. وأصبح يعرف فيما بعد بـ IQ.

وجدير أن نشير هنا إلى أنّ Binet أصرّ على اعتبار نسبة الذكاء مجرد مرشد امبريقي ذو هدف عملي محدد، وهو التعرف على الأطفال الذين هم بحاجة إلى المساعدة لتحقيق أكبر استفادة ممكنة من قدراته مهما كانت محدودة.

وثقل مقياس Binet إلى الولايات المتحدة الأمريكية وترجم إلى اللغة الإنجليزية أول مرة بواسطة Goddard لكن هذا الأخير حول هدف الاختبار إلى تصنيف الأفراد، بحيث يفتح المجال أمام ذوي القدرات المرتفعة ويعزل ذوو القدرات المنخفضة.

وتبقى جهود Lewis Terman بجامعة Stanford بالولايات المتحدة الأمريكية هي التي أدت إلى تطوير المقياس وانتشاره بالاسم الذي أصبح يعرف به في تراث القياس النفسي وهو مقياس "ستانفورد بنيه لقياس الذكاء" وكان اهتمامه بقياس الذكاء لغرض تصنيف الأفراد، إذ يرى أنّ الطريقة الوحيدة الفعالة للتعامل مع ضعاف العقول هي بالحجز المؤسسي الكامل. تمت مراجعة المقياس (1937) ليشمل على صيغتين متكافئتين وأعيد مراجعته (1960) ليصبح صيغة واحدة، وتوسع ليشمل الراشدين كما استبدل مفهوم نسبة الذكاء (Qi) بمفهوم نسبة الذكاء الانحرافية (DQI) باقتراح من وكسلر.

كون مقياس Binet يعد من المقاييس الفردية، اتجهت جهود مجموعة من علماء النفس إلى بناء الاختبارات الجماعية، خاصة مع دخول الولايات المتحدة في الحرب العالمية الأولى 1917، أين ظهرت الحاجة إلى تصنيف أكثر من مليون مجند لتخصصات معينة أو لتدريب معين أو تسريحهم، في هذا السياق اقترحت الجمعية الأمريكية لعلم النفس (APA) تحت قيادة Robert Yerkes تطوير اختبارين يمكن تطبيقهما جماعيا، اختبارات الجيش Alpha موجهة للمجندين المتمكنين من اللغة الإنجليزية قراءةً وكتابةً واختبارات الجيش Beta تم تصميمه للمجندين الأميين والذين لديهم ضعف أو عدم اتقان اللغة الإنجليزية وهو اختبار غير لفظي، إلى جانب تصميم اختبار ثالث وهو مقياس الأداء الفردي للجيش لأولئك المجندين الذين فشلوا في كلا الاختبارين السابقين.

في عام 1939 قام David Wechsler كبير علماء النفس في مستشفى Bellevue بتقديم "مقياس الذكاء وكسلر بيلفو" وهو من الاختبارات الفردية ومناسبا للراشدين، واستخدم كل من الاختبارات اللفظية والأدائية في نفس البطارية، وطوره فيما بعد (1955) ليشمل قياس ذكاء الراشدين والمراهقين. أما الصورة الثانية من Wechsler كانت مصممة لقياس ذكاء الأطفال من سن الخامسة إلى سن الخامسة عشرة (1946) وتطورت إلى مقياس وكسلر لذكاء الأطفال العام (Willis, Dumont, & Kaufman, 2015) وأبرز مساهماته هو قدرته على اعتماد على العينة المعيارية، وكذلك صلاحيته للمواقف المدرسية والعيادية. وأحدث إصدار لمقياس Wechsler لذكاء الأطفال (wisc-5) سنة 2014 وتم تقنيته لتطبيقه في البيئة العربية، أما الإصدار الخاص بمقياس ذكاء الراشدين (WAIS-4) سنة 2008 (نسخة أمريكية).

على مدار العقود التي تلت أبحاث وكسلر دخل العديد من الأخصائيين النفسانيين ميدان الاختبارات المعرفية، ففي نهاية الخمسينيات بدأ Richard Woodcock بتطوير اختبارات تقيس قدرات التعلم الخاصة، وكان الاختبار الأول الذي تم تطويره هو اختبار التعلم البصري والسمعي، وهو مقياس لقدرة الفرد على تعلم القراءة، وبعد فترة وجيزة جاء اختبار التحليل التجميعي كمقياس لقدرة الفرد على تعلم الرياضيات، وخضعت لمراجعة سنة 1977 لتعرف فيما بعد بـ "البطارية النفسية التربوية Johnson-Woodcock" أما الاستخدامات الرئيسية لها فتمثلت في تحديد الهوية الفردية، التشخيص، الاختيار والتنسيب، تخطيط البرامج الفردية، التنبؤ وتأكيد الحاجة للإرشاد، تقييم البرامج، البحث والتكوين (Shinn, Marston, & Ysseldyke, 1980/June/July) وأسفرت في النهاية عن 12 اختباراً شملت كل الوظائف اللفظية وغير اللفظية التي تقع على سلسلة متصلة من عمليات بسيطة إلى معقدة، وصدرت الطبعة الثالثة لهذه البطارية سنة 1989 (WJ III COG).

في عام 1983 تم تطوير بطارية (K-ABC) من طرف باحثين أمريكيين هما Nadeen Kaufman وزوجته Alan وهي عبارة عن اختبار ذكاء ومعارف، تطبق فردياً على الأطفال بين سنتين ونصف واثني عشرة ونصف، ويقوم على ثنائية متزامن-متسلسل كما يحتوي على 16 اختباراً فرعياً (بوفولة، 2009). بعد عقد من الزمن (1993) تابعا Kaufman إصدار طبعة منقحة لبطارية (K-ABC-II) وبطارية أخرى لقياس ذكاء المراهقين والراشدين (KAIT) وتتكون البطارية من مقياس متبلور ومقياس سائل، وفي نفس الحقبة تقريباً التي تم فيها تطوير بطارية (K-ABC) في الولايات المتحدة، صممت النفسانية البريطانية (Colin Elliott) المقاييس البريطانية للقدرات (BAS, British Ability Scales) وأدى تطوير وتقنين صورة (BAS) بالولايات المتحدة عام 1990 إلى ظهور النسخة الثانية لها تحت مسمى DAS-II وحققت هذه النسخة العديد من التطورات والابتكارات في التقييم المعرفي (Willis, Dumont, & Kaufman, 2015).

في 1997 تم تصميم الاختبار الشامل للذكاء غير اللفظي CTONI من طرف الباحثين (Hammill, Pearson, & Wiederholt, 1997) وطور لإعطاء الفاحصين وسيلة لتقييم مهارات التفكير المنطقي غير اللفظي للأفراد من سن 0-6 إلى 11-89. وتلت مقاييس أخرى كمقياس Cecil R. Reynolds & R.W. Kamphaus 2003 وهو من الاختبارات الفردية يستغرق تطبيقها 25-35 دقيقة و تطبق بشكل فردي وتصلح للأفراد الذين تتراوح أعمارهم بين 3 سنوات و 94 سنة. وغيرها من المقاييس، نظام التقييم المعرفي (CAS). الملاحظ أن الذكاء الإنساني قد استغرق زمناً طويلاً وتطلب جهداً أكثر من أي موضوع آخر من موضوعات القياس النفسي، وفي دراسة الذكاء يتفاعل دافع الاستطلاع والسياقات الاجتماعية والثقافية والسياسية مع المطالب العلمية والتربوية، وبسبب الحاجة العلمية في المجتمعات المتمدنة أصبح الأمر أكثر إلحاحاً لإيجاد طريقة ما لحساب ذكاء الأفراد بصورة دقيقة كما تقاس في الاختبارات الأخرى، عموماً يمكن أن نشير إلى ما ذكره (Willis, Dumont, & Kaufman, 2015) حول وصف تاريخ اختبار الذكاء بأنه

بحث عن نتيجتين: (أ) توفير مقاييس موثوق بها تقيس الذكاء العام أو القدرة العقلية العامة g و (ب) جرد وقياس القدرات الهامة المتضمنة في القدرة العامة g.

محاضرة رقم 9: نظريات الذكاء

3-نظريات الذكاء : بصرف النظر عن تعريف "الذكاء"، فقد كان علماء النفس مهتمين بمعرفة بنية الذكاء. إذ قاموا بتحليلات في محاولة لتحديد العوامل الأساسية المكونة له. كان الغرض من هذه التحليلات هو اكتشاف عناصر الذكاء أو مكوناته إن أمكن، ليس فقط من أجل فهم نظري أفضل لهذه العملية المعقدة، ولكن أيضًا لمعرفة ما قد تكون الآثار المترتبة على تصميم وبناء الاختبارات العقلية، وطور علماء النفس والفلاسفة نظريات تسعى إلى فهم ديناميكية وطبيعة الذكاء.

وقام (Gardner, 2011) بتصنيف نظريات الذكاء في أربعة أنواع رئيسية: (1) نظريات القياس النفسي. (2) النظريات المعرفية. (3) النظريات المعرفية - السياقية. و(4) النظريات البيولوجية. ويندرج تحت كل تصنيف مجموعة من النظريات.

1-3. نظريات القياس النفسي للذكاء: تستند نظريات القياس النفسي للذكاء على دراسة الفروق الفردية بين الأفراد: وبشكل خاص الفروق الفردية في الأداء على الاختبارات ذات المكون المعرفي (Gardner, 2011) ففي حالة إجراء بحث نموذجي يستخدم هذه المقاربة، يتم تعريف عدد كبير من الأشخاص لعدد من الاختبارات التي تقيس القدرات المعرفية (مثل المفردات، سلسلة الأرقام، السرعة الإدراكية، المعارف العامة...) نحصل من هذه الاختبارات على نتائج أو درجات مترابطة، هذه الأبعاد الكامنة تشكل عادة أساس نظرية الذكاء الناتجة. ونستعرض وصف النظريات النفسية للذكاء بناءً على التسلسل التاريخي.

أ-نظرية العاملين لسبيرمان: Spearman tow -Factores Théory يعد النفساني الإنجليزي سبيرمان أول من طور نظرية الذكاء على أساس نفساني (psychométrie) عام 1904. يرى سبيرمان أن الذكاء يمكن فهمه من خلال عاملين مهمين، الأول هو العامل العام (G) وهو قدرة عامة واحدة أو قدرة مشتركة تتدخل في كافة الأنشطة العقلية ويتميز العامل العام (أ) بأنه قدرة فطرية شاملة. (ب) طاقة عقلية عامة. (ج) ثابت وموجود عند كل فرد، ارتباطه بالقدرات الأخرى يظل بنفس المقدار (د) يختلف مقدار "g" من فرد إلى آخر.

(هـ) يتم استخدامه في جميع أنشطة الحياة. (و) زيادة "g" عند الفرد، يقابله زيادة النجاح في الحياة. (ز) إنه قيمة مجردة وليست شيئاً ملموساً، والعامل الثاني هو العامل الخاص (S) وهو مجموعة من القدرات الخاصة، حيث يتطلب إنجاز أية مهمة توظيف قدرات خاصة (S) إضافية لإنجاز تلك المهمة، وتتميز (S) بالخصائص التالية: (أ) يتم تعلمها واكتسابها في البيئة (ب) أنها تختلف من نشاط إلى آخر في نفس الفرد. (ج) الأفراد يختلفون في مقدار القدرة (S).

ويشير سبيرمان Spearman إلى أن العوامل الخاصة تتأثر بالتربية والممارسة، أما العامل العام فهو فطري ولا يقبل التنمية والتعديل ولا يتأثر بالبيئة، وينمو طبيعياً حتى يبلغ مداه في سن 18 (مقدم، 2003) وهذا يمثل إشارة صريحة إلى مسألة تأثير الوراثة والبيئة في النشاط العقلي.

ب- نظرية الروابط Bond Theory: تسمى أيضاً نظرية العوامل المتعددة، نتيجة الانتقادات التي تعرضت لها نظرية سبيرمان ترتب على ذلك ظهور نظريات آمنت بوجود عوامل للتركيبية العقلية لكن هذه العوامل تختلف من حيث النوعية والكمية. وبعد غودفري طومسون Godfrey Thomson من الأوائل الذين تعرضوا بالنقد لنظرية سبيرمان، فهو يتصور أن العامل العام هو نتيجة لمجموعة من الروابط التي تعمل كنموذج في وقت واحد لإنجاز المهام العقلية المستقلة المتداخلة. ويتضح أن درجة الارتباط بين كل اختبار وآخر تتوقف على درجة المهارات العقلية المشتركة بينها وبين العمليات المترابطة (الجاسم، 2010).

وطور أيضاً Thorndike 1925 نظريته حول الذكاء (نظرية العوامل المتعددة)، التي تقوم فكرتها على وجود وصلات شبيهة بالوصلات الفيسيولوجية في المخ وأن الفروق بين الأفراد ناتجة عن عدد الوصلات في المخ، والذكاء عند ثورندايك عبارة عن عناصر متعددة منفصلة ومستقلة عن بعضها البعض في بعض المظاهر. فالوصلات بين الأداء لا تفسر على أساس العامل العام ولكن حسب وظيفة كل عنصر من العناصر العامة المتضمنة في تلك الأعمال (الجاسم، 2010) ويرفض فكرة عمومية الذكاء وتجانسه. فهو ليس نوعية موحدة ولا مركب من عاملين من عوامل سبيرمان أو مجموعة عوامل. يتمثل الذكاء لدى ثورندايك في:

- الذكاء المجرد: وهو القدرة على معالجة الألفاظ والرموز.

- الذكاء الميكانيكي (العملي) وهو القدرة على معالجة الأشياء والمواد العيانية.

- الذكاء الاجتماعي: وهو القدرة على التعامل بفاعلية مع الآخرين.

د- نظرية القدرات العقلية الأولية Primary Mental Abilities: يعتبر ثورستون L. L. Thurstone الأخصائي النفسي أكبر المساهمين في نظرية العوامل المتعددة فيرى أن الذكاء لا يمثل العامل العام كما نادي به سبيرمان، ولا بالترابطات المتعددة لقدرات جد خاصة كما يناهز بها ثورندايك، وإنما يمكن تقسيم الذكاء إلى عدد من القدرات الأولية. حيث كان افتراضه أن بعض العمليات العقلية تشترك في عامل أولي (وهو ليس بالضرورة "g") الذي يمنحها الوحدة النفسية والوظيفية والتي تميزها عن العمليات العقلية الأخرى. هذه العمليات العقلية تشكل في حد ذاتها مجموعة، والعامل الأولي أو القدرة العقلية التي تعمل من خلال تلك المجموعة أو مجموعة من تلك العمليات تسمى "عامل المجموعة" أو العامل الأولي الذي يكون مستقلاً نسبياً عن العوامل الأولية الأخرى التي تعمل في مجموعة أو مجموعات أخرى من العمليات العقلية. وتوصل ثورستون بعد إجرائه لهذه التحليلات إلى وجود سبعة عوامل أساسية أطلق عليها العوامل العقلية الأولية، وذلك بعد إجرائه مجموعة من البحوث بتطبيق 65 اختباراً على 240 طالباً في المرحلة الثانوية باستخدام أسلوب التحليل العاملي المتمثل في تدوير المحاور (صالح، 1992)، وهذه القدرات هي كالتالي:

- عامل العدد (N): القدرة على القيام بحسابات رقمية بسرعة ودقة.
- العامل اللفظي (V): القدرة التي تقاس عن طريق الاختبارات التي تتطوي الفهم القرائي. القدرة على فهم اللغة المسموعة والمقروءة
- عامل الفضاء (القدرة المكانية) (S): يقصد بها قدرة الفرد على تصور الأشياء الفراغية وإدراك العلاقات المكانية أو الهندسية، التصور البصري للعلاقات.
- عامل الطلاقة اللغوية (W): القدرة تتضمن الوتيرة السريعة عند التفكير في الكلمات المنعزلة. القدرة على التفكير في الكلمات بسرعة، كما هو الحال في حل الجناس أو التفكير في الكلمات التي قافية.
- عامل الاستدلال (R): قدرة الفرد على اكتشاف القاعدة ثم الوصول إلى تعميم من الخاص إلى العام، وتقاس باختبارات التناظر سواءً اللفظية أو العددية.
- عامل الحفظ والتذكر (M): القدرة على الحفظ بسرعة. القدرة على تذكر منبهات لفظية مثل أزواج الكلمات أو الجمل.
- عامل السرعة الإدراكية (P): وتتمثل في قدرة التعرف على الأشياء والموضوعات بسرعة. وأضاف عاملين آخرين فيما بعد ليصبح عدد العوامل تسعة، عامل المنطق، عامل الاستدلال الاستقرائي. وتأكيداً لنظريته وضع ثورستون بطارية القدرات العقلية الأولية (1942) مكونة من خمس بطاريات موجهة لمختلف المراحل العمرية والتي تبدأ من الحضانة إلى المرحلة الثانوية
- هـ- **نظرية التنظيم الهرمي**: بحلول نهاية الأربعينيات من القرن العشرين، أصبح من الواضح أنه لا يمكن لنظرية العامل سبيرمان ولا لنظرية العوامل الأولية لثورستون وصف الارتباطات بين اختبارات القدرات المعرفية بشكل كافٍ. اقترح علماء النفس في المملكة المتحدة (Burt, 1940; P. E. Vernon, 1950) الجمع بين مقاربات سبيرمان وثورستون في وصف هرمي واحد للقدرات البشرية. فالعامل العام الذي يقيس القدرة العقلية العامة ويقاس النشاط العقلي المعرفي هو المستوى الأول للقدرات، والعامل الخاص للقدرات يتحدد من خلال ما يميز كل اختبار وهو يُعد المستوى النوعي أو الخاص، أما مستوى القدرات الطائفية فهو الصفة المشتركة لمجموعة من القدرات الخاصة (الجاسم، 2010)
- **نموذج (Burt, 1940)**:

بتزايد عدد العوامل المكتشفة حديثاً، أصبحت هناك حاجة إلى ترتيبها وفق العلاقة المنطقية. كان سيريل بيرت، عالم النفس البريطاني، من أوائل الذين حاولوا ذلك. واقترح أن القدرات العقلية قد تكون منظمة بشكل هرمي. من خلال دراسته للتحليل العاملي، جادل بأن العامل العام الشامل يمكن أن يستخدم لتمثيل كل الأداء العقلي. ويمكن أيضاً تقسيم هذا العامل العام إلى مجموعة من العوامل المتعددة تمثل فئات واسعة ومختلفة من السلوك العقلي. يمكن تقسيم عوامل المجموعة الواسعة هذه وفقاً لتصوير بيرت إلى عوامل مجموعة أضيق، ثم إلى العديد من العوامل المحددة جداً التي تُصور التسلسل الهرمي المثالي داخل الثنائيات المتعاقبة، وحدد مستويات مختلفة من التشعب والعلاقات على أعلى مستوى. حيث يضع في المستوى الأدنى

للهرم عمليات الإحساس البسيط والنشاط الحركي، يلي ذلك العمليات الأكثر تعقيدا والتي تتعلق بالإدراك والحركات التآزرية، وفي المستوى الثالث يضع بيرت عوامل الذاكرة والعادات التي تم اكتشافها وتكوينها، أما في المستوى الرابع فنجد عمليات العلاقات، وتنقسم إلى الثنائية الكبيرة والمتمثلة في الفهم من جهة والاستخدام من جهة ثانية. والذكاء العام أو "الامكانية التكاملية للعقل" كما يسميه بيرت يظهر في كل مستوى من المستويات الأربعة السابقة، إلا إن ظهوره يختلف في كل مستوى من حيث الدرجة والنوع (مقدم، 2003).

- نموذج (Vernon, 1950): يشير فيرنون إلى أنه يمكننا تصوير العقل على أنه نوع من التسلسل الهرمي أو شجرة الأنساب، حيث يمثل "g" القدرة الذهنية الأبرز، بمعنى أنه يمثل أكبر نسبة من الاختلافات في القدرات. في الجزء العلوي من التسلسل الهرمي، يضع فيرنون عامل "g" لسبيرمان. في المستوى التالي هناك مجموعتين من العوامل الطائفية، المجموعة الأولى تقابلها القدرات التعليمية اللفظية (v: ed) والمجموعة الثانية تقابلها القدرات العملية الميكانيكية (k: m). هذه العوامل الرئيسية قد تكون مقسمة بشكل أكبر. فالعامل التعليمي اللفظي، على سبيل المثال، ينتج عوامل فرعية شفهية وعددية وغيرها. وبالمثل فإن العامل العملي الميكانيكي يمكن أن ينقسم إلى المعلومات العملية الميكانيكية والقدرات المكانية والحركية. ويمكن إجراء مزيد من التحليل لتحديد عوامل أضيق. إن هذه العوامل الطائفية الصغرى هي بدورها تنفرع إلى عوامل بسيطة وفي أي أداء عقلي يمكن وضعه على أنه يتضمن نسبا معينة من العامل العام والعوامل الطائفية ثم العوامل الخاصة، وذلك لتفسير النشاط الذي قام به الفرد (مقدم، 2003).

و- نظرية الذكاء السائل والذكاء المتبلور **Fluid and Crystallized Ability**: تم تطوير نظرية الذكاء السائل (fluid) والمتبلور (crystallized) في الأصل من قبل Raymond B. Cattell 1941، 1963، وتم التحقيق فيها فيما بعد بشكل أعمق مع متعاونه، John L. Horn (هورن، 1968، 1985؛ هورن وكاتيل، 1966، 1967) (Gardner, 2011) يرتبط الذكاء السائل بمهام مثل الاستدلال الاستقرائي، الاستدلال الاستنتاجي، فهم العلاقات بين المحفزات، استيعاب المعاني، واستنتاجات الرسم (Horn & Hofner، 1982) وهذا النوع من القدرات يكون متحررا نسبيا من العوامل الثقافية. وقد ربط (Horn & Cattell 1966) الذكاء السائل مع القدرة البيولوجية الأساسية للتعلم. من ناحية أخرى يرتبط الذكاء المتبلور بمهام مثل المفردات والمعارف الثقافية. فهو مرتبط بالخبرة الثقافية والتعرض للتعليم النظامي؛ بمعنى المعارف المكتسبة جاءت بفضل تجارب محيطه. الذكاء السائل والمتبلور غير مستقلين عن بعضهم البعض؛ فالارتباط بينهم يتراوح ما بين 0.4 إلى 0.5. إن التمييز بين الذكاء السائل والذكاء المتبلور هو أكثر من مجرد تمييز أكاديمي. كما أن الذكاء السائل عرضة للتراجع بسبب تلف الجهاز العصبي المركزي، في حين أن الذكاء المتبلور يبقى سليم نسبيا بعد هذا الضرر (Horn & Hofner، 1982). وعلاوة على هذا فإن الذكاءين يتعرضان لأنماط مختلفة من النمو والتراجع على مدى العمر الافتراضي. فيبلغ الذكاء السائل ذروته في وقت مبكر إلى منتصف سن العشرينيات وينخفض بعد ذلك. وتبلغ قدرة المتبلور ذروتها

في وقت متأخر (في أوائل الأربعينيات)، وفي العديد من الحالات تظل عالية حتى في مرحلة البلوغ المتأخرة (Horn & Hofner, 1982)

ي-نظرية القدرات المعرفية - الترتيب الهرمي الثلاثي - Three-Stratum Factor Analytic Theory of Cognitive Abilities: وقد اقترح John B. Carroll النظرية التحليلية للعوامل الثلاثة للقدرات العقلية (1993; 1997) هذا النموذج يشبه نموذج التنظيم الهرمي في أنه هرمي، ومع ذلك، فإنه يقترح ثلاثة مستويات فقط من التسلسل الهرمي. في الجزء العلوي من التسلسل الهرمي نجد العامل g. في وسط التسلسل الهرمي هناك ثماني مجموعات واسعة النطاق (Sattler, 2008): (1) الذكاء السائل؛ (2) الذكاء المتبلور؛ (3) الذاكرة العامة والتعلم (4) الإدراك البصري الواسع؛ (5) إدراك سمعي واسع؛ (6) قدرة استرجاع واسعة. (7) السرعة الإدراكية الواسعة؛ و (8) سرعة المعالجة. في الجزء السفلي من التسلسل الهرمي نجد 65 قدرة ضيقة. يشبه هذا النموذج نظرية ثورستون، نظرية القدرات العقلية الأولية (من حيث أنها تحتوي على مجموعة واسعة من العوامل) ونظرية هورن وكاتيل للقدرة السائلة والمتبلورة (من حيث أنها تتضمن هذه العوامل صراحةً كمجموعة) (Gardner, 2011).

ر-نموذج البنية العقلية "SI Structure of Intellect": اقترحت نظرية البنية العقلية من طرف النفساني الأمريكي (J. P. Guilford 1964; 1967; Guilford & Hoepfner, 1971). يرفض جيلفورد فكرة عامل الذكاء العام g، والمجموعات الواسعة للعوامل مثل القدرات الأولية لثورستون. نموذج "SI" هو محاولة للتعرف على طبيعة القدرات العقلية الخاصة عن طريق توليد نماذج من الذكاء تفترض العديد من المهارات العقلية. ويقترح أنه بما أن الناس يمتلكون بوضوح مجموعات غير متساوية من القدرات العقلية والنماذج النظرية، يجب أن تعكس مقاييس الذكاء هذه الاختلافات بطريقة منتظمة. يفترض نموذج SI 180 عنصرًا محددًا من الذكاء استنادًا إلى 3 فئات عامة. ووفقًا له، فإن كل عملية عقلية أو نشاط عقلي له ثلاثة أبعاد، ثلاثة معايير أساسية يمكن أن يحدث خلالها أي سلوك فكري محتمل، أي "العملية" و "المنتج" و "المحتوى"

2-3. النظريات المعرفية: في الوقت الذي تركز فيه النظريات السيكومترية على بنية الذكاء البشري، نجد النظريات المعرفية تركز على العمليات التي ينطوي عليها الذكاء البشري. ونستعرض البحوث والنظريات التي تتضمن العمليات المعرفية البسيطة إلى الأبحاث التي تتطوي على العمليات المعرفية المعقدة.

أ- القياس الحسي البسيط:

يعتقد المدافعون عن القياس الحسي البسيط أن الأفراد الأكثر ذكاءً يكونون أكثر قدرة على القيام بعمليات تمييز حسية جيدة، مثل التمييز بين الارتفاعات، أو الأوزان ذات القيمة المماثلة. وكان هذا واضحًا في العمل المبكر لسبيرمان (1904b). يعتقد آخرون أن الأفراد الأكثر ذكاءً سيكون لديهم طول القامة وقدرات جسدية فائقة (تيرمان، 1925). والاهتمام المبكر لمنهج القياس الحسي البسيط للذكاء بدأ بأعمال غالتون، متأثرًا بنظرية داروين في أهمية الفروق الفردية في التكيف والتكاثر البيئي، فجاء على إثرها تركيز غالتون على الفروق الفردية في القدرة الطبيعية، ويرى أن أولئك الذين يتمتعون بقدرة طبيعية أكبر سيكونون

بارزين في مجالات دراستهم، بينما أولئك الذين لديهم قدرة طبيعية أقل سيخفقون. ويعتقد بانتقال الذكاء عن طريق الوراثة واختلاف الأجناس ذكاءهم، وأن تحسين النسل أو التكاثر المخطط، يمكن استخدامه لتحسين "ذكاء" أمة ما.

في الولايات المتحدة، دافع جيمس ماكين كاتيل عن أفكار غالتون. حيث بدأ برنامجًا لأخذ قياسات بسيطة من الطلاب في الفصل الدراسي الأول من كل عام (Cattell & Farrand, 1896) بجامعة كولومبيا. وتتكون من 21 قياس، وشملت أشياء مثل قوة اليد، حدة البصر، وزمن رد الفعل. لكن طالب الدراسات العليا كلارك ويسلر، أجرى دراسة للتحقق من صحة تعلق درجات الطلاب في جامعة كولومبيا بالقياسات الحسية البسيطة. ولم تكون نتائج ويسلر (1901) مشجعة للغاية. حيث وجد أن درجات الطلاب تميل إلى الترابط مع بعضها البعض، ولكن ليس مع بيانات الاختبارات الحسية البسيطة التي تم جمعها. أنهت دراسة Wissler إلى حد كبير الاهتمام المبكر بالتدابير الحسية البسيطة كمؤشر للذكاء.

ب- زمن المعاينة:

طرح Douglas Vickers زمن معاينة الفقرة (Inspection time-IT) نحو عام 1970 على أنه محدد أساسي للمعدل الذي يمكن على أساسه تجميع المعلومات الخارجية لاتخاذ القرار في مخازن حسية مؤقتة (ستيرنبرج و كوفمان، 2015). ويشير زمن المعاينة إلى مدة التعرض المطلوبة من شخص ما لتحديد تأثير بسيط. وهي أيضا مدة التحفيز المميزة التي يحتاجها الفرد من أجل اتخاذ قرار بمستوى معياري دقيق. وبالرغم من استخدام إجراء زمن المعاينة في سياقات بحثية أخرى، إلا أن الاستخدام الأكثر شيوعًا كان في مجال الذكاء، حيث تم استخدامه لاختبار إمكانية أن الفروق الفردية في الذكاء المقاس قد تتعلق بالفروق الفردية في زمن المعاينة (Deary & Stough , June 1996) والفكرة المتعلقة بسرعة تناول المعلومات البصرية قد تكون ذات صلة بالفروق في القدرات العقلية الأكثر عمومية التي حدثت لـ McKeen Cattell حين كان يعمل في مختبر Wundt ومع ذلك لم يتم متابعة هذه الفكرة في أي برنامج بحثي منتظم. ويعد Vickers المسؤول الأول عن صياغة نظرية زمن المعاينة، إلا أن Nettelbeck كان أول من قام بفحص العلاقة بين زمن المعاينة ونسبة الذكاء. حين كان طالب دكتوراه عند Vickers في جنوب استراليا (Deary & Stough , June 1996) وانطلاقًا من الاقتراح القائم على أن الفروق في القدرات العقلية يمكن أن تعزى إلى السرعة التي تشتغل بها هذه القدرات في بعض من آليات البحث والتحليل، قام كل من (Nettelbeck et Lally (1976) et Lally et Nettelbeck (1977)) بفحص العلاقة بين زمن المعاينة ونسبة الذكاء على مقياس وكسلر للراشدين، ووجدت قيمة الارتباط بين زمن المعاينة IT ونسبة الذكاء IQ تتراوح ما بين -0,8 و -0,9. لذلك تعد هذه الدراسة أول الأعمال التي أظهرت وجود علاقة بين الفروق الفردية في زمن المعاينة المقاس والذكاء النفسي. وفي مراجعة شبه كمية لـ Nettelbeck 1987 استنتج أن زمن المعاينة يمثل حوالي 25% من تباين نسبة الذكاء، ونقول بمعنى آخر أن الارتباط بين زمن المعاينة والذكاء النفسي بلغ نحو -0,5 لدى عموم الناس (Deary & Stough , June 1996).

ج- زمن الرجع البسيط/ الاختياري: بعض المؤلفين افترضوا بأن "السرعة العقلية" «vitesse mentale»، تعتبر خاصية أساسية تؤثر على جميع العمليات المعرفية، ويمكن لها أن تفسر الفروق في الأداء التي تكمن في عامل g. ولاختبار هذه الفرضية استخدموا نموذجين تجريبيين رئيسيين زمن المعاينة (تطرقنا إليه أعلاه) ونموذج زمن الرجع الذي كيّفه (Jensen, 1982 Lautiry, 2005). سيتم عرض هذا النموذج ومناقشته هنا. لقد دافع آرثر جنسن عن نموذج أكثر تعقيداً قليلاً للتحقيق في العلاقة بين العمليات المعرفية الأولية والذكاء: وقت رد الفعل البسيط / الاختياري. ويقصد بزمن الرجع بالزمن المنقضي بين بداية المنبه وبداية الاستجابة، هناك نوعان رئيسيان منه، الأول زمن الرجع البسيط وفيه يطلب من المفحوص الاستجابة بأسرع ما يمكن عند تقديم المنبه الضغط على زر معين عند ظهور الضوء. والثاني زمن الرجع الاختياري، حيث يعرض على المفحوص منبهان ويطلب منه الاستجابة بطريقة معينة للمنبه الأول وبطريقة أخرى للمنبه الثاني ويمكن أن تتزايد بدائل الاستجابة بحيث تكون ثلاثة أو أربعة (طه، 2006). ولخص إيزنك نتائج مناقشة جينسن للتراث بما يشمله من أعماله وأعمال زملائه في النقاط التالية (طه، 2006).

- أظهر زمن الرجع البسيط ارتباطاً منخفضاً بالذكاء يتراوح بين -0,2 إلى -0,3.

- أظهر زمن الرجع البسيط ارتباطاً بالذكاء يمكن إهماله.

- أظهر زمن الرجع الاختياري ارتباطاً أكثر جوهرية من زمن الرجع البسيط بالذكاء.

- يتزايد الارتباط بين زمن الرجع وبين نسبة الذكاء مع تزايد عدد الاختبارات المستخدمة.

- يرتبط تباين زمن الرجع ارتباطاً سلبياً بالذكاء.

- ترتبط مهمة زمن الفحص ارتباطاً دالاً سالباً بالذكاء.

د- الذاكرة العاملة: يذكر Baddeley & Hitch 1974 أن الذاكرة العاملة عبارة عن أنظمة تخزين خاصة وظيفتها تخزين المعلومات اللفظية، تسمى هذه الأنظمة (المكون اللفظي) بالإضافة إلى أنها تحتوي على أنظمة أخرى خاصة بمعالجة المعلومات تسمى المعالج المركزي، وفي عام 1992 أضاف "بادلي" مكوناً آخر أطلق عليه المكون (غير اللفظي) وظيفته معالجة الصور المكانية والبصرية وإدراك العلاقات المكانية، وتعمل هذه المكونات في تكامل وانسجام وتناسق (بلبل و حجازي، 2016) قياس الذاكرة العاملة يتم عادة بالمهام التي تتطلب التشفير والتخزين والتذكر والتعرف على المثبرات، ويتم تفعيل قدرة الذاكرة العاملة عن طريق عدد المفردات التي يستطيع الفرد حفظها في الذاكرة العاملة، ويقدر متوسط القدرة على الحفظ بأربع مفردات، حيث تتنوع في الناس ما بين مفردتين إلى ست مفردات (Chuderski, 2013).

و- نموذج الترابطات المعرفية: أفضل ما يميز مقارنة الترابطات المعرفية هي أعمال Earl Hunt وزملائه (Hunt, Frost et Lunneborg, 1973; Hunt, Lunneborg et Lewis, 1975; Hunt et al., 1978, 1976; Lansman, 1975). ويدخل هذا البحث في سياق نموذج عام للذاكرة يصور تدفق المعلومات من خلال سلسلة من المخازن الحسية إلى ذاكرة قصيرة المدى، ثم من خلال ذاكرة متوسطة المدى إلى ذاكرة طويلة المدى. هناك نظام تنفيذي يتحكم في تدفق المعلومات والوصول إلى مستويات

مختلفة من تخزين الذاكرة (هاننت، 1971، 1973). مع هذا النموذج كإطار مرجعي، طرح هاننت سلسلة من الأسئلة حول الاختلافات بين مجموعات القدرة اللفظية العالية والمنخفضة التي ترتبط بالهياكل والعمليات والمعايير الخاصة بنظام معالجة المعلومات.

كان الهدف من هذا العمل هو إظهار أن الذكاء اللفظي الذي يتم قياسه بالوسائل التقليدية مرتبط بمتغيرات أساسية تمت دراستها في النظرية المعرفية الحديثة، لا سيما معلمات النماذج التي تصف تحويل المعلومات في الذاكرة القصيرة والطويلة الأجل. بشكل عام، يستنتج هاننت وزملاؤه أن اختبارات الذكاء اللفظي تستفيد مباشرة من معرفة الشخص باللغة، مثل معاني الكلمات، القواعد النحوية، والعلاقات الدلالية بين المفاهيم التي تشير إليها الكلمات، وأن هذه الاختبارات أيضا تقيم بشكل غير مباشر قدرات معالجة المعلومات التي تعتبر عناصر أساسية في الدراسات التجريبية للذاكرة (Gardner, 2011).

ي-مدخل المكونات المعرفية: يستخدم مدخل المكونات المعرفية تحليل المهام كوسيلة لتحديد مكونات تناول وتجهيز المعلومات التي يتطلبها الأداء على مفردات الاختبارات العقلية المقننة، ثم يقيس مقدار الفروق الفردية في هذه المكونات حسب Sternberg (1977) (الشيخ، 2014) فالباحثون في هذا المدخل يهدفون إلى تحليل العمليات المتضمنة في النشاط العقلي للإنسان بشكل مباشر، بدل من البحث عن الارتباطات بين هذا النشاط والعمليات المعرفية. وقد ارتبط هذا المدخل باسم Sternberg، حيث أجرى في بحثه تحليلا معرفيا لعدد كبير من المهام السيكمترية وبصفة خاصة مهام الاستقراء والاستنباط. وفي هذا الإطار يميز ستيرنبرغ بين ثلاثة أنواع من العمليات المعرفية الرئيسية، ويعتبرها المصدر الأساسي للفروق الفردية في الذكاء وهذه الأنواع هي (الشيخ، 2014)

أ- ما وراء المكونات: هي عمليات تحكم "ضبط" ذات مستوى أعلى، وهي عمليات تنفيذ تُستخدم في تخطيط أداء الفرد في المهمة، والتهيؤ أو التوجيه له، وتقويمه. ويشير ستيرنبرج إلى عشرة عمليات مهمة من هذا النوع.

ب- مكونات الأداء: عبارة عن عمليات معرفية ذات مستوى أدنى، تستخدم في تنفيذ الاستراتيجيات المختلفة لأداء المهمة.

ج- مكونات اكتساب المعرفة: وهي عمليات متضمنة في تعلم معلومات جديدة وتخزينها في الذاكرة.
ر- سرعة المعالجة والشيخوخة:

يعرض الذكاء السائل والمتبلور مسارات مختلفة من التطور في مرحلة البلوغ. تزداد قدرة السائل خلال فترة المراهقة، ويصل إلى قمته في أوائل إلى منتصف العشرينات، ثم تنخفض بعد ذلك. في المقابل، تزداد القدرة المتبلورة حتى أوائل الأربعينيات، وغالبا ما تظل مرتفعة إلى وقت متأخر من مرحلة البلوغ. وبالنظر إلى أن القدرة المتبلورة تمثل منتجات التعليم والثقافة، فإن نموها البطيء يمكن فهمه على أنه التراكم التدريجي للمعرفة. ولكن ما الذي يمكن أن يفسر الانخفاض في الذكاء السائل التي تحدث خلال منتصف العمر وما بعده (Gardner, 2011).

يقترح (Timothy Salthouse 1985, 1993, 1996) أن الانخفاض في قدرة السائل، والوظائف العقلية المماثلة هو نتيجة لتباطؤ سرعة المعالجة للعمليات المعرفية مع الشيخوخة. ويربط هذا التباطؤ إلى آليتين أساسيتين لضعف الأداء: (1) الآلية الزمنية المحدودة، و (2) آلية التزامن. ووفقاً لآلية الزمن المحدودة فإن وقت إجراء العمليات اللاحقة مقيد إلى حد كبير عندما تشغل نسبة كبيرة من الوقت المتاح من خلال تنفيذ العمليات المبكرة. تنص آلية التزامن على أنه بسبب التباطؤ قد تفقد منتجات المعالجة المبكرة بحلول الوقت الذي تكتمل فيه المعالجة. نتيجة كل من هذه الآليات هو انخفاض في الأداء، ليس فقط في السرعة، ولكن أيضاً في الدقة مع تقدم السن (Gardner، 2011).

بالرغم من أن نظرية Salthouse لا تشرح الذكاء في حد ذاته، إلا أنها تقدم تفسيراً للانخفاض في قدرة السائل خلال الحياة. ويشير أيضاً إلى عامل مهم يجب مراعاته عند تصميم التعليم والتدريب لكبار السن: وهي سرعة المعالجة. فعند تقديم المعلومات إلى جمهور أكبر سناً، يجب على المرء أن يقوم بتسوية السرعة البطيئة التي يمكن تشفير وتحليل المنبهات بها، وإمكانية ألا يكون لدى الأفراد الأكبر سناً في وقت واحد إمكانية الوصول إلى معلومات مختلفة مثلما يفعل المتعلمون الأصغر سناً.

3-3. النظريات المعرفية-السياقية

تحاول النظريات المعرفية - السياقية شرح السلوك الذكي من حيث السياق الذي يتم عرضه فيه. رغم أن المعالجة المعرفية مهمة جداً في هذه النظريات، إلا أنها تميل إلى اتباع النهج البيولوجي بصفة أكثر: يتم تحديد وتطوير بعض المهارات (وتشجيعها) انطلاقاً من البيئة أو الثقافة. وبالتالي قد لا يكون هناك شيء اسمه الذكاء الشامل (Gardner، 2011). فالذكي أو الأذكى هم المتوافقين مع البيئة أو الثقافة. لهذه النظريات نظرة أكثر شمولية من عديد النظريات المعرفية التي تطرقنا لها سابقاً، فهي لا تميل إلى التركيز على العمليات المعرفية الفردية، وإنما على فئات العمليات أو أنواع الذكاء.

أ-نظرية التنظيم الثلاثي للذكاء (Robert Sternberg 1977a, 1977b):

ينظر Robert Sternberg إلى الذكاء على أنه مجموعة من المكونات، "مجموعة واسعة من المهارات المعرفية وغيرها من المهارات" وهي في نفس الوقت قوية يوحدّها ما يسميه العمليات التنفيذية، فهذه النظرية تسعى إلى فهم الذكاء من خلال ثلاث نظريات فرعية؛

أ. نظرية فرعية سياقية تربط الذكاء بالمحيط الخارجي للفرد، وتتضمن هذه النظرية التكيف الهادف، واختيار وتشكيل بيئات العالم الحقيقي المرتبطة بحياة الفرد، ويتجلى ذلك في ثلاثة أنواع من الذكاء هي: الذكاء الأكاديمي، والذكاء العملي، والذكاء الابتكاري.

ب. ونظرية المكونات حيث تربط الذكاء بالبيئة الداخلية للفرد، ويرى أن الذكاء يرتبط بالمكونات الداخلية للفرد، ويتألف من العمليات الأدائية والعمليات المعرفية الماورائية وعمليات اكتساب المعرفة

ت. ونظرية فرعية خبراتية (ذكاء الخبرة) تنطبق على كل من المحيط الداخلي والخارجي. ويقاس الذكاء

هنا:

. بمدى توفر القدرة على التعامل مع المهمات والمواقف الجديدة (القدرة الابتكارية).
 . القدرة على المعالجة الذاتية للمعلومات، وبأقل جهد ممكن (قدرة الاستبصار).
 يرى Sternberg أن الذكاء يمارس في ثلاث مجالات: الخارجي (أو السياق)، التجريبي والداخلي.
 الذكاء هو نشاط عقلي، لكن كل جزء في النظرية يعتبره على علاقة بمجال مختلف.
 السياق ببساطة هو البيئة الخارجية التي يشتغل فيها الذكاء، سواء داخل الصف الدراسي أو المكتب.. قد يستخدم نفس الشخص ذكائه بطريقة مختلفة في كل بيئة. أما التجربة أو الخبرة فهي المجال الذي يواجه فيه الأشخاص مواقف جديدة، وفيها البديهية والبصيرة والإبداع - وهي عمليات غير عقلانية لا تدخل ببساطة في الآليات العقلية المعتادة لمعالجة الصور، التي ترتبط بها الذكاء بالعالم الداخلي فرد، يتم جلبها على الذكاء من خلال الخبرة.

ب-نظرية الذكاءات المتعددة 1999, 1983 Howard Gardner:

يقدم "هاورد جاردر" في كتابه "إطارات العقل، نظرية الذكاء المتعدد" (1983) وجهة نظر جديدة ومختلفة عن الكفاءات العقلية البشرية. ويجادل بكل جرأة وقناعة بأننا جميعنا نولد مع إمكانية تطوير مجموعة من الذكاءات (Pal, Pal, & Tourani, 2005)

يرى جاردر أن البشر جميعا يمتلكون ثمانية ذكاءات متميزة على الأقل، وهذه توجد على هيئة خليط نسبي مميز خاص بكل فرد. ويعرّف الذكاء على أنه "القدرة على حل المشكلات، أو استحداث نتائج لها قيمة في موقف ثقافي أو آخر" (Gardner, 2011) وبذلك يرفض مفهوم الذكاء بوصفه قدرة عامة، تشدد هذه النظرية على مجالات الذكاء، وبدرجة أقل على العمليات العقلية. ولكي يكون جزءاً من نموذج الذكاء المتعدد، يتعين على الذكاء المرشح أن:

- أ. يكون قابلاً للعزل في حالة الإصابة الدماغية.
- ب. يمتلك القدرة على التطور.
- ت. يشمل مجموعة عمليات مركزية يمكن التعرف إليها.
- ث. يطوّع نظاماً من التمثيلات الرمزية.
- ج. يكون له تاريخ تطوري مع القدرة على الأداء الخبير.
- ح. يكون ظاهراً في أشخاص مميزين، مثل العلماء.
- خ. يكون هناك دليل عليه في علم النفس التجريبي.
- د. يكون مدعوماً بدليل من البحث السيكميترى. يتطور كل واحد من هذه الذكاءات من خلال التفاعلات بين نزعات الفرد البيولوجية والفرص التي تتيحها بيئة الفرد. كذلك يصف جاردر نوعين أساسيين من مظاهر الذكاء، هما: الافراد ذوو المستوى العالي في نوع أو نوعين من الذكاء يطلق عليهم (أصحاب الملامح الليزرية Laser profiles) في حين يوصف ذوو التوسيع الأوسع (ذوو ملامح كاشفات ضوئية Searchlight profiles).

الذكاءات الثمانية لجاردنر:

1. الذكاء اللغوي: أو الحساسية للغة الشفوية والمكتوبة، القدرة على تحليل المعلومات وإنشاء منتوجات تتضمن اللغة الشفوية والكتابية، مثل الخطابات والكتب والمذكرات.
2. الذكاء المنطقي الرياضي: القدرة على التحليل المنطقي للمشكلات، وإجراء العمليات الحسابية والتحقق في القضايا بطريقة علمية
3. الذكاء المكاني: القدرة على التعرف على أنماط المساحات المفتوحة والتعامل معها وإدارتها.. بالإضافة إلى المناطق المحصورة.

تشبه هذه الأنواع الثلاثة من الذكاء، القدرات التي يمكن قياسها بوساطة اختبارات الذكاء التقليدية (ستيرنبرج وكوفمان، 2015).

4. الذكاء الموسيقي: الإحساس بخصائص موسيقية متعددة، والقدرة على تقدير النغمات، والألحان، والإيقاعات، وإنتاجها، ومزجها
5. الذكاء الحس-حركي: مهارة الفرد في استخدام جسمه
6. الذكاء الشخصي: يشير إلى فهم الفرد لدوافعه ومشاعره، ونقاط قوته وضعفه.
7. الذكاء الاجتماعي: فهم دوافع الآخرين، وتصرفاتهم ومشاعرهم والاحساس بها.
8. الذكاء الطبيعي: التمييز الحاذق، وتصنيف الأنماط الطبيعية أو الأشياء المادية بحسب الفئات.

ج-نظرية مرحلة النمو لبياجيه Piaget: نظرية بياجيه هي نظرية نمائية تبحث في خصائص نمو الأطفال، في داخل هذه الخصائص يعالج موضوع الذكاء، فهو يناقش الذكاء في ضوء مكانه داخل مخطط النمو العقلي للطفل. وفي نفس الوقت هي نظرية معرفية لأنها تناقش عمليات التفكير، وسياقية لأنها تؤكد على دور البيئة في تحفيز النمو المعرفي (Gardner, 2011) نموذج بياجيه لنمو الذكاء هو نموذج نوعي، فالأطفال الأكبر سنا يعرفون أكثر من الأطفال الأقل سنا منهم، ويختلفون عنهم في معرفتهم. ويعرّف الذكاء بأنه القدرة على التكيف العقلي مع المستجدات الراهنة"

التصور الذي وضعه بياجيه حول النمو العقلي جاء بعد دراسة طويلة، وهو انعكاس لاهتمامه الأول بالبيولوجيا والابستمولوجيا، حيث يفترض أن البشر يرثون اتجاهين أساسيين التنظيم والتكيف (Gardner, 2011) التنظيم (ميل عام لتنظيم كل من العمليات الفيزيائية والنفسية في نظم متماسكة) وهذا الميل موجود سواء على المستوى البيولوجي أو النفسي، فالفرد في تفاعله مع العالم الخارجي يميل إلى أن تتكامل أبنيته النفسية في نظم متماسكة، فمثلا، يكون لدى الطفل الصغير القدرة على النظر للأشياء أو القدرة على القبض عليها، كل قدرة منهما مستقلة أو منفصلة عن الأخرى. ولكن بعد فترة من النمو، ينظم الطفل هذه الأبنية المنفصلة في بنية ذات مستوى أعلى، تمكنه من أن يقبض على شيء معين وينظر إليه في نفس الوقت.

فالتنظيم إذن هو ميل مشترك في كل أشكال الحياة، والاتجاه الثاني هو التكيف (ميل لتمثل وملاءمة البيئة)، وبما أن عملية الهضم البيولوجي تحول الطعام إلى شكل يمكن أن يستخدمه الجسم، يعتقد Piaget أن العمليات العقلية تحول الخبرات إلى شكل يمكن للطفل استخدامه في التعامل مع مواقف جديدة، فكل خبرة لدينا (أطفال مراهقين أو راشدين)، تؤخذ في العقل ويتم إعدادها بحيث تتسق أو تدخل في الخبرات السابقة الموجودة هناك، والخبرة الجديدة تحتاج لأن تعدل بدرجة ما، لكي يمكن إدخالها في البنية المعرفية القائمة. وبعض الخبرات لا يمكن تمثيلها، لأنها لا تلائم البنية (الشيخ، 2014) ومعنى هذا أن العقل يتمثل أو يستوعب الخبرات الجديدة عن طريق التغيير فيها. بحيث تلائم البنية التي تم تكوينها وسمّها بياجيه المخططات. وفي هذا يقول بياجيه "الذكاء هو تمثيل بالدرجة التي يستوعب فيها كل بيانات الخبرة المعينة في إطاره الخاص" (الشيخ، 2014) كذلك تتأثر الأبنية العقلية القائمة بالبيئة التي يشتغل فيها العقل. فعملية التمثيل تحدث أثناء استقبال خبرات تكون مرت بالفرد من قبل، بمعنى عملية التمثيل تحدث كلما استجاب الفرد في موقف جديد كما فعل في مواقف مشابهة سابقا، لكن يحدث أن يتعرض لمحفزات بيئية (خبرات جديدة)، لم يسبق له أن مرّ بها أو تعرض لها، وهذا يمكن أن يتسبب في ملاءمة الطفل للمؤثرات من خلال إنشاء مخططات معرفية جديدة أكثر تطورا تتناسب بشكل أفضل مع البيئة. فإذا كانت عملية التمثيل هو تفسير المؤثرات الجديدة وغير المعتادة في ضوء المخططات القديمة أو الحالية، فإن عملية الملاءمة تعني تعديلا في المخططات المعرفية عن العالم، حتى يمكن له استيعاب الخبرات الجديدة (الشيخ، 2014). يُنظر إلى الطفل على أنه يتطور من خلال مراحل معقدة متزايدة، المرحلة الحسية الحركية (0-2 سنوات) ، مرحلة ما قبل العمليات العقلية (2-7 سنوات)، مرحلة العمليات المحسوسة (7-11 سنة) مرحلة العمليات الشكلية أو الذكاء المجرد (11+ سنوات).

الظاهر أن نظرية بياجيه لا تعالج الذكاء من منظور القياس النفسي، فإنها تعالج بالتأكيد تطور الآليات المعرفية التي تكمن وراء الكثير من الاستدلال المجرد وحل المشكلات. وهي تنظر إلى الفرد على أنه شخص يتفاعل مع البيئة، ويسعى جاهداً لجعل أفكاره (أو مخططاته) تتماشى مع التجربة. وهو يجعل الادعاء القائل بأن جميع الأفراد يمرون خلال المراحل بنفس الترتيب قوياً، مع عدم تخطي أي "مرحلة" إلى مراحل أعلى. في حين انتقد العديد من الباحثين الأعمار التي حدّدها بياجيه لتحقيق مهارات خاصة، إلا أن القليل منهم انتقد قدرة نظريته على وصف اكتساب الأطفال التدريجي للفكر المعقد.

3-4. النظريات البيولوجية للذكاء:

إنّه من الواضح أن الذكاء البشري موجود داخل الدماغ. وقد أدى الاهتمام الأخير في علم النفس العصبي المعرفي إلى الأمل في أننا سنكتشف الآليات البيولوجية والفسولوجية الخاصة المسؤولة عن الذكاء (Gardner، 2011) وسنستعرض بعض النظريات البيولوجية المطروحة في هذا المجال، ويجدر بنا أن نشير إلى الحاجة لتذكر أن الدماغ هو عضو بالغ التعقيد، ولا تزال معرفتنا بعمله في مهدها.

أ. حجم الدماغ والذكاء:

أجريت العديد من الدراسات لتقييم العلاقة بين حجم المخ والذكاء. ففي دراسات مبكرة، تم قياس حجم الدماغ عبر بديل أكثر سهولة مثل محيط الرأس (يطلق عليه أحياناً "المحيط"). في مراجعة 35 دراسة سابقة تضم 56793 فرداً، ذكر Vernon, Wickett, Bazana, and Stelmack 2000 وجود ارتباط متوسط مرجح مقدّر بـ 0.191 بين حجم الرأس والذكاء. كما خلصت التحليلات الماورائية التي قام بها كل من Gignac, Vernon et Wickett 2003 لمجموعة من التحقيقات (14 دراسة)، $n = 858$ أجراها باحثون آخرون قبل عقد من الزمن، إلى وجود ارتباط بين حجم المخ ونسبة الذكاء Q_i وقدر المتوسط المرجح لهذا الارتباط في ست تحقيقات من أصل أربعة عشر تحقيقاً بـ 0.37 بين حجم الدماغ ونسبة الذكاء Q_i . كما أفاد Gignac وآخرون (2003) عن متوسط ارتباط مصحح غير مرجح يقدر بـ 0.43 بين حجم الدماغ ونسبة الذكاء (Gignac & Bates, 2017). كذلك قام Mc Daniel 2005 بإعادة فحص الارتباط بين الذكاء وحجم الدماغ من خلال إجراء تحليل ماورائي أكثر شمولاً من الذي قام به Gignac وآخرون سنة 2003، إذ حدّد Mc Daniel مجموعة من المعايير الصارمة في التحقيقات المعتمدة، تمثلت في اعتماد عينات سوية كلينيكية، قياس حجم الدماغ الكلي، ومقاييس ذكاء معدة بطريقة جيدة (مقياس وكسلر، مقياس رافن) واستناداً إلى العينات التي تستجيب لهذه المعايير حيث $K = 37$ و $n = 1530$ قدر الارتباط الملاحظ بـ 0.29 بين حجم الدماغ والذكاء العام، وارتباط مصحح قدره 0.33. وجددير أن ننوه هنا إلى أن هذه القيم لمعاملات الارتباط بين حجم الدماغ والذكاء التي أشار إليها Mc Daniel أقل من القيمة التي أعلنها Gignac وآخرون (Gignac & Bates, 2017).

ب- طول سلسلة EEG والذكاء:

عرض Hendrickson نموذجاً نظرياً يربط الذكاء بتعقيدات التخطيط الكهربائي للدماغ (EEG). في الأساس، يفترض هذا النموذج أن الأفراد الأكثر ذكاءً لديهم أخطاء أقل في عمليات الإرسال المشبكية، وأن هذا سيؤدي إلى نمط موجة EEG أكثر تعقيداً. يتم تحديد تعقيد نمط موجة EEG بطول "سلسلة" متراكبة فوق شكل الموجة؛ المزيد من التخطيط الدماغى المعقد يؤدي إلى سلسلة أطول (Gardner, 2011).

ج- نموذج معدل استقلاب (Métabolique) الجلوكوز والذكاء:

تشير دراسة (haier, siegel, tang, abel, & buchsbaum, 1992) للتصوير المقطعي للانبعثات البوزتروني (PET) لمعدل استقلاب سكر الجلوكوز (Glucose) في الدماغ لدي عينة من المتطوعين العاديين، إلى وجود ارتباطات عكسية (سلبية) بين هذا الأخير والدرجات المحصل عليها في مصفوفة Raven المتقدمة، والطلاقة اللفظية. فالأشخاص الذين حصلوا على أعلى الدرجات في مصفوفة RAPM أظهروا انخفاض أكثر في معدل استقلاب الجلوكوز في الدماغ العام، وفي مناطق خاصة أيضاً. كذلك استكشف هاري وزملائه هذه النظرية في مهمة التعليم، حيث أخضع ثمانية أفراد للتصوير المقطعي للانبعثات البوزتروني (PET) لمعدل استقلاب سكر الجلوكوز في الدماغ أثناء لعبهم للكمبيوتر (لعبة Tetris)، وبعد

مدة زمنية تراوحت بين أربعة إلى ثمانية أسابيع أجرى عليهم تصوير آخر وهم في حالة ممارسة للعبة، كشفت النتائج بين التصويرين ظهور انخفاض في معدّل الاستقلاب لصالح التصوير الثاني، وهذا نتيجة التحسينات التي اكتسبها المتعاونين خلال ممارستهم لهذه اللعبة. وقد تعكس هذه النتائج حسبهم تحسينات مهمة في المعالجة التلقائية عند ذوي القدرة العالية بعد التطبيق الثاني وبالتالي انخفاض أكثر في معدّل استقلاب الجلوكوز في الدماغ.

يفسر هاري هذه النتائج على أنّها ناتجة عن فعالية الدماغ، فالأفراد الأكثر ذكاءً يتطلبون نشاطاً عصبياً أقل (وبالتالي استقلاب أقل في الجلوكوز) لحل المشكلات العقلية مقارنة بالأفراد الأقل ذكاءً (HAIER, (SIEGEL, TANG , ABEL , & BUCHSBAUM , 1992))

د- نموذج جينسن (Jensen) للتذبذب العصبي:

اقترح Jensen 1982 "نموذج التذبذب العصبي" لشرح اكتشافه في أنّ زمن رد الفعل (TR) في مهمة زمن رد الفعل بسيط / اختياري يزيد خطياً مع أجزاء من المعلومات في عرض التحفيزات (Gardner, 2011) افتراض جنسن أن الجهاز العصبي يستخدم شبكة ثنائية هرمية لمعالجة المعلومات في مهمة وقت رد الفعل البسيطة / الاختيارية. لهذا يفترن الافتراض بأن كل عقدة في الشبكة كانت عرضة لدورة متذبذبة بين الفترات النشطة (عندما تتمكن العقدة من التشغيل ومن ثم نقل المعلومات) وفترات حرارية (عندما يتعذر إطلاق العقدة). افتراض أن الناس أكثر إشراقاً لديهم دورات التذبذب العصبي أقصر، وبالتالي كانوا قادرين على العودة إلى حالة "إطلاق النار" بسرعة أكبر من الناس أقل سطوعاً. أثبت جنسن أن مثل هذا النموذج يمكن أن يفسر النتائج الأساسية التي توصل إليها: (1) الزيادة الخطية في TR مع زيادة أجزاء المعلومات؛ و (2) زيادة الانحراف المعياري TR مع زيادة أجزاء المعلومات (Gardner, 2011)

نتائج النظريات البيولوجية للذكاء توفر لنا أدلة على العلاقة بين جوانب الدماغ والذكاء أو مع العمليات العقلية. لكن هذه النتائج مرتبطة فقط بهذه النظريات التي اقترحها هؤلاء الباحثين، وهذا راجع إلى كون علم الأعصاب المعرفي علم فتي نسبياً، حيث نجد الروابط السببية بين بنية الدماغ ووظيفته من ناحية والأداء العقلي من ناحية أخرى ليست واضحة بشكل كامل. وقد تصبح هذه الروابط السببية أكثر وضوحاً في المستقبل مع تقدم الأبحاث في هذا المجال.

خلاصة:

يظهر من العرض المقدم حول مراجعة نظريات الذكاء الأربع: نظريات القياس النفسي للذكاء، النظريات المعرفية والنظريات المعرفية السياقية، والبيولوجية، أنّ نظريات القياس النفسي للذكاء ناضجة نسبياً. كما يُعد التمثيل الهرمي للقدرات أحسن تمثيل للعلاقات البنوية بين القدرات، مع وجود العامل العام g في أعلى الهرم، وعوامل المجموعة في المستويات الوسطى، وفي الجزء السفلي العوامل الخاصة نسبياً، والتفضيل بين النموذج الهرمي البريطاني ونظرية كارول قد يكون مسألة شخصية بالدرجة الأولى.

النظريات المعرفية للذكاء حاولت توضيح المعالجة المعرفية التي يقوم عليها السلوك الذكي، فالمدى الذي يوجد فيه إجماع هو أنّ السلوك الذكي مرتبط بالعمليات الواسعة المتدخلة في عدد كبير من المهام، ومن أمثلة هذه العمليات، الذاكرة العاملة، الانتباه، مكونات التفكير العام، انتقاء الاستراتيجية وتنفيذها (ما وراء المكونات)، سرعة معالجة المعلومة مهمة أيضاً، لكن لا يتم ابرازها إلا في حالات خاصة (أثناء الشيخوخة، مشكلات جديدة) والملاحظ أنّه لا توجد هناك نظرية معرفية واحدة متفق عليها.

النظريات المعرفية السياقية وسعت النظريات المعرفية من خلال إدماجها في السياق، سواء في السياق البيئي أو الثقافي، تؤكد هذه النظريات أنّه لا يمكن تطوير العمليات العقلية وتقييمها إلا في بيئة تختارها أو تدعمها أو داخل ثقافة، بهذه الطريقة فإنّ الذكاء هو تفاعل بين الفرد والبيئة. أما النظريات البيولوجية حاولت شرح الذكاء من حيث بنية ووظيفة الدماغ، فالعديد من الباحثين أشاروا إلى الارتباطات التي تجمع بين الدماغ والذكاء، لكن لا تزال هناك صعوبات في تحديد هذه الارتباطات، بسبب حداثة مجال علم الاعصاب المعرفي نسبياً.

محاضرة رقم 10: اختبارات الذكاء

4-اختبارات الذكاء :

تُسجل طبعة من سلسلة الاختبارات المطبوعة (tests in print TIP) ما مجموعه مئتي اختبار واختبارين في قسم (الذكاء والكفاية العامة)، من هذه الاختبارات نجد سبعة وعشرون اختباراً فقط يستخدم مصطلح الذكاء في العنوان (Urbina, 2011) فالاختبارات المنشورة في العقود الماضية القليلة تجنبت استخدام مصطلح الذكاء في العنوان، في حين الاختبارات الأقدم لا تزال تستخدم المصطلح حتى في طبعاتها الحديثة، لكي توفر استمرارية لها (Urbina, 2011) وتعد اختبارات الذكاء التقليدية -وكسلر، بنيه- الأكثر استخداماً ودراسة.

لعل غياب تعريف واحد مقبول للذكاء يسهم في الاختلاف في كيفية تقييمه، ومن دون الاتفاق على تعريفات الذكاء يبقى من الصعب التوصل لاتفاق على كيفية قياسه، فالذكاء مفهوم مجرد يستدل عليه فقط، ويكون مبنياً على البيانات المتوافرة لنا والمعايير الخاصة بنا. وفي تعريف لاختبارات الذكاء أشارت (Urbina, 2011) إلى أنّ اختبارات الذكاء تقيس شيئاً ما في الواقع. تنتج من هذه الاختبارات أعداد. يحصل عليها المُمْتَحَنون بحسب إجاباتهم، طبقاً لعينات السلوك التي تبني كل اختبار، وتُصنّف هذه الأعداد بحسب معايير، أو قواعد محددة.

1-4. أنواع اختبارات الذكاء: توجد أربع طرق رئيسية -على الأقل- يمكن بها تصنيف اختبارات الذكاء (Urbina, 2011) هي:

أ- عن طريق نظام الإدارة، أي اختبارات فردية مقابل اختبارات جماعية.

ب- عن طريق نوع المجتمعات المستهدفة من الاختبارات: مثل الاختبارات الموجهة للأطفال أو البالغين، أو أي مجموعات أخرى محددة.

ت- عن طريق نوع المحتوى: اختبارات لفظية وغير لفظية.

ث- عن طريق ما إذا كانت الاختبارات بطارية كاملة أو إصدارات مختصرة. مع أنّ هذا التصنيف قائم على الاختبارات التي تحمل مسمى الذكاء في العنوان، فإنه يمكن تطبيقه على تلك التي تستخدم أسماء مختلفة مثل اختبارات القدرة العامة، المعرفية.

ونستعرض هذه التصنيفات لاختبارات الذكاء بشيء من الشرح، مع ذكر أهم الاختبارات التي تنتمي إلى كل تصنيف، ويكون التركيز على أكثرها تداولاً واستخداماً من طرف المختصين في وقتنا الراهن. وقد يرد اختبار واحد في أكثر من تصنيف، بمعنى يكون اختباراً فردياً موجهاً للأطفال، ويحمل مهمات تتطلب مصطلحات لفظية وكتابية.

1-1-4. الاختبارات الفردية: هي تلك الاختبارات التي تُقدّم للمختبرين فردياً، حيث يختبر فيها الفاحص كل شخص على حدي. في مثل هذه الاختبارات يتطلب تدريب عالي للقائم على الاختبار، إلى جانب قياس الذكاء في مثل هذه الاختبارات نجدها توفر ثروة من المعلومات تفيد في التقييم الكلينيكي ودراسة الحالات، ومن أمثلة الاختبارات الفردية:

أ- **اختبار ستانفورد بنيه SB:** وهو من الاختبارات الفردية واللفظية، تم تطوير هذا الاختبار من طرف بنيه وسيمون 1905 وقد مر بثلاثة مراحل قبيل وفاة بنيه وهي 1905. 1908. 1911، وطوّر في الأصل لتحديد الأطفال الذين يعانون من صعوبات فكرية خطيرة - بحيث أنهم لن ينجحوا في نظام المدارس العامة والذين لا ينبغي وضعهم في نفس الصفوف مع التلاميذ الآخرين. هذا الاختبار يقيس الأشياء اللازمة للنجاح في المدرسة مثل فهم اللغة واستخدامها، والمهارات الحسابية، والذاكرة، والقدرة على اتباع التعليمات. وتكون إجابة المفحوص فردية على أربع مجالات: التفكير اللفظي. التفكير الكمي. التفكير المجرد -البصري. ذاكرة قصيرة المدى. في عام 1916 تمت مراجعته من طرف تيرمان بجامعة ستانفورد بالولايات المتحدة الأمريكية وأصبح الاختبار يحمل تسمية اختبار ستانفورد بنيه للذكاء. ويعتمد على نسبة الذكاء Qi، حيث قياس الذكاء هنا يأخذ في الاعتبار العمر العقلي والزمني للطفل.

$$100 \times CA / MA = \text{نسبة الذكاء.}$$

العمر العقلي (MA): مستوى الذكاء النموذجي الذي تم اكتشافه للأشخاص في عمر زمني معين.

العمر الزمني (CA): العمر الفعلي للطفل مع اختبار الذكاء

وفيما يلي أبرز التطورات لهذا المقياس (Roid, 2003).

1905 إصدارات Binet and Simon بفرنسا.

1916 مراجعة Stanford للاختبار من طرف Lewis M. Terman. من هنا أصبح يعرف باختبار ستانفورد بنيه.

1937 النسخة الثانية Lewis M. Terman, Maud A Merrill SB2 –

1960 النسخة الثالثة SB3 نموذج L-M Lewis M. Terman, Maud A Merrill

1972 دمج النموذجين L-M-Lewis M. Terman, Maud A. Merrill, Robert L. Thorndike

1986 اختبار الذكاء SB4 (Robert L. Thorndike, Elizabeth P. Hagen, Jerome M. Sattler)

2003 النسخة الخامسة والأحدث لاختبار الذكاء، (Stanford-Binet SB5) طورها Roid Gale

. وهناك نسخة مكيفة ومقننة في البيئة العربية، من إصدار المؤسسة العربية لإعداد وتقنين ونشر الاختبارات

النفسية. ويعتمد على حساب نسبة الذكاء Qi.

يمتاز مقياس ستانفورد بنيه للذكاء (الإصدار الخامس)، أنه مناسب لنطاق عمري واسع يتراوح بين عامين إلى 85 سنة. وكذا اعتماد التطورات الحديثة في نظرية القياس أثناء عملية التقنين، يستغرق تطبيقه 50 دقيقة (حوالي 5 دقائق لكل من الاختبارات الفرعية العشرة). يصلح لاستخدامات واسعة، مثل، تشخيص حالات العجز الارتقائي لدى الأطفال والمراهقين، وبالغين. التقييم الإكلينيكي والعصبي السيكولوجي، تقييم الطفولة المبكرة، التقديرات النفسية التربوية المتعلقة بالالتحاق ببرامج التربية الخاصة، التقديرات الخاصة بتعويضات العمال، تقديم معلومات عن التدخلات مثل الخطط العائلية الفردية للصغار، الخطط التربوية الفردية للأطفال في سن المدرسة التقييم المهني (التخطيط للانتقال من المدرسة إلى العمل) بالنسبة للمراهقين، التغيير المهني للراشدين وتصنيف وانقضاء الموظفين، تشخيص حالات الإعاقة العقلية (في كل الأعمار)، صعوبات التعلم التأخر المعرفي الارتقائي في الأطفال الصغار، إلحاق الطلاب ببرامج الموهوبين عقليا في المدارس، تقييم إصابات العمل وحجم الإعاقة الناتجة عنه (Roid, 2003).

كما يسمح بتقييم الأفراد ذوي اللغة المحدودة من خلال مقياس غير لفظي، فهو مفيد للأفراد الذين قد يتم التقليل من قدرتهم من خلال المهام اللفظية وحدها (الصم وضعاف السمع، ضعاف اللغة، الأفراد المصابين بالتوحد، إصابات الرأس، فقدان القدرة على الكلام، أو إصابات الدماغ). الملاحظ أيضا على هذه النسخة تضمنها لاختبار غير لفظي لكن الإجابة تكون لفظية وقليلة.

ب-مقياس وكسلر للذكاء the Wechsler Intelligence scale : نشر Davide Wechsler سلسلة من

اختبارات الذكاء بدأت سنة 1939 تحت مسمى Wechsler Bellevue واختار أن يبقى مصطلح نسبة الذكاء Qi ليشير إلى الدرجات الكلية في تلك المقاييس، لكن كانت نسبة الذكاء الانحرافي لوكسلر تختلف عن نسبة الذكاء عند بنيه. ونشير هنا إلى أحدث إصدارات سلسلة اختبارات الذكاء لوكسلر دون الإشارة إلى الطبقات السابقة، وقد تطرقنا في عنصر التطورات التاريخية لاختبارات الذكاء إلى هذا الموضوع.

-**اختبار وكسلر لذكاء الأطفال والمراهقين (WISC-5)** أحدث اصدار له في 2014 بالولايات المتحدة الأمريكية، يهدف هذا الاختبار إلى تقييم القدرات العقلية للأطفال والمراهقين الذين تتراوح أعمارهم بين 6 سنوات و16 سن و11 شهرا. يمكن استخدام هذا المقياس للتعرف على الصعوبات المعرفية، لتقدير نقاط الضعف والقوة في الموضوع مقارنة بالشخص نفسه وبالأخرين، لتحديد وتشخيص الصعوبات أو اضطرابات التعلم، لتقييم أثر تلف الدماغ على العمليات العقلية، تحديد الأطفال أو المراهقين ذوي الإمكانيات العالية أو الذين لديهم تأخر عقلي (Terriot & Ozenne, 2015).

-**اختبار وكسلر لذكاء الراشدين (WAIS-4)** أحدث اصدار له في 2008 بالولايات المتحدة الأمريكية، شمل عشر اختبارات فرعية درست العديد من جوانب العقل البشري، بما في ذلك الذاكرة. وهذا يجعل الاختبار أداة مفيدة في التشخيص المبكر لمرض الزهايمر، يصلح للفئة العمرية من 16 سنة إلى 90 سنة. كما يفيد في الدراسات الإكلينيكية لتشخيص الحالات، مثل، الإعاقة الذهنية (الخفيفة، ومعتدلة الشدة) الموهبة، اضطراب التوحد، صعوبات التعلم القرائية والرياضية، الزهايمر، الاكتئاب.

-**اختبار وكسلر لذكاء الأطفال ما قبل المدرسة والابتدائية (WPPSI 4)** أحدث اصدار له في 2012 في الولايات المتحدة الأمريكية، يصلح للفئة العمرية 2سنتين إلى 6سنوات.

ج-مقياس رينولدز للتقييم العقلي (RIAS؛ Reynolds & Kamphaus، 2003): هو اختبار ذكاء يصلح للاستخدام مع الأفراد الذين تتراوح أعمارهم بين 3 و94. صممه الباحثان بإدماج اختبارات فرعية تقيس القدرات السائلة والقدرات المتبلورة مع استبعاد الاختبارات الفرعية التي أثبت تاريخياً أنها غير متشعبة بمعامل g (على سبيل المثال، تلك التي تنطوي على السرعة الحس حركية)، ويتطلب تطبيقه وقت قصير، بذل مؤلفو هذا المقياس جهداً كبيراً لتطوير أداة عادلة ثقافياً، توفر تفسيراً موحداً لنتائج الاختبارات عبر المجموعات العرقية (Dombrowski & Mrazik, 2008). اعتمد الباحثان في تطوير **RIAS** على اعتبارات نظرية، تجريبية وتطبيقية، مستوحى جزئياً من النظرية الثلاثية للقدرات العقلية لـ (Carroll 1993) وأعمال (Horn et Cattell 1966)). كما ينسب اختبار **RIAS** إلى البنية الهرمية لـ g ، حيث يقيس العامل الأعلى g إلى جانب عوامل الترتيب الأدنى للذكاء المتبلور والسائل (Dombrowski & Mrazik, 2008)

يحتوي **RIAS** على اختبارات فرعية تتشابه مع اختبارات قياس الذكاء Qi التي لها تجربة طويلة في قياس الذكاء. يحتوي في مجمله على أربع اختبارات فرعية للذكاء، اثنتين لفظيين، واثنين غير لفظيين، فضلا عن اختبارين ثانويين للذكاء، هذا الأخير لا يساهم في نتيجة مؤشر الذكاء.

د-بطارية كوفمان لتقييم الأطفال K-ABC: ظهرت لأول مرة بالولايات المتحدة الأمريكية سنة 1983، ويعتبر اصدار 2004 أحدث نسخة لهذه البطارية، وهي من اختبارات الذكاء، تدار بشكل فردي للأطفال الذين تتراوح أعمارهم بين 3 إلى 18. تم تصميم الاختبار القائم لاستخدامه في الحالات الإكلينيكية، والنفس تربوية، والحالات النفسية العصبية.

ونذكر أيضا من اختبارات الذكاء الفردية، اختبار وودكوك جوهانسون Woodcock-Johnson، ومقاييس القدرة المتميزة، واختبارات أخرى.

مزايا اختبارات الذكاء الفردية:

- توفر ثروة من المعلومات عن المختبر إلى جانب درجة الاختبار التي يحصل عليها.
- التعليمات وأساليب التطبيق تكون متطابقة قدر الإمكان، بحيث يخضع المختبرين لنفس ظروف التطبيق.
- تسمح بملاحظة الفروق في السلوك والمواقف بين الأفراد أثناء تطبيق الاختبار، ذلك أثناء الفشل أو النجاح في الإجابة على بند معين.
- بعدما يكتسب الفاحص خبرة في إدارة وتطبيق الاختبار الفردي، يمكن له ملاحظة الفروق في ردود الفعل بين المفحوصين تحت نفس الظروف، فالفاحصين من ذوي الخبرة بإمكانهم مع مرور الوقت تطوير معايير داخلية خاصة بهم، حيث يصبحون على فكرة عن كيفية تفاعل معظم المختبرين مع موقف أو مهمة معينة. وتكون لهم دراية بالتفاعلات أو ردود الفعل غير العادية بشكل أسهل.
- توفر الاختبارات الفردية فرصة للحصول على معلومات عن المختبر تتجاوز قيمتها تلك المعلومات التي يحصل عليها في المقابلة أو من الأصدقاء أو الأسرة، لأنهم نادرا ما يكونون موضوعيين وغير مدربين على ملاحظة السلوك البشري، كذلك تكون الظروف أثناء تطبيق الاختبار مقننة.

2-1-4. اختبارات الذكاء الجماعية:

تطبق على عدد كبير من الأفراد وفي نفس التوقيت الزمني، وتتضمن اختبارات الذكاء الجماعية، اختبارات لفظية في المقام الأول التي تتطلب استخدام اللغة للإجابة على بنود الاختبار، واختبارات أدائية لا تتطلب استخدام اللغة في الإجابة عن بنودها، واختبارات مركبة أي مزيج بين اللفظي وغير اللفظي. وبصفة عامة الاختبارات الجماعية هي اختبارات الورقة والقلم أو كتيب وقلم، فهي تتوفر على كتيب مطبوع فيه بنود، دليل الاختبار، مفتاح التصحيح، ورقة إجابة وقلم الرصاص، ومع التطورات التكنولوجية أصبحت الاختبارات المحوسبة les tests informatisés الأكثر طلبا للاستخدام، (Kaplan & Saccuzzo, 2009) كما أن معظم الاختبارات الجماعية متعددة الاختيارات، ولكن بعضها يتطلب استجابة حرة مثل إكمال جملة أو رسم، أو كتابة مقال. ونظراً لوجود عدد كافٍ من الاختبارات الجماعية المناسبة من الناحية السيكومترية بالنسبة لمعظم الأغراض، فمستخدم الاختبار هنا يحتاج فقط إلى اختبارات ذات موثوقية عالية وسليمة من ناحية الشروط السيكومترية. وسوف نتطرق إلى بعض الاختبارات المنتشرة بقليل من الشرح وتسمية البعض الآخر فقط.

أ-اختبارات الذكاء اللفظي للمجموعة: تلك الاختبارات التي تستلزم استخدام اللغة ويتم تطبيقها على مجموعة من الأفراد في وقت واحد، وتستخدم عادة لأغراض الفحص والتصنيف، وتندرج تحت هذه الفئة مجموعة من الاختبارات هي:

- اختبار ألفا للجيش (تم تطويره في الحرب العالمية)
- اختبار التصنيف العام للجيش (تم تطويره في الحرب العالمية الثانية)

- اختبار أوتيس لينون للقدرة المدرسية (الطبعة الثامنة) OLSAT Otis-Lennon School Ability :Test

تعتبر اختبارات OLSAT 8 الصيغة الأحدث من سلسلة الاختبارات التي أعدها كل من S. Otis and Roger T. Lennon 1918 ولقد تم تغيير اسم الاختبار من اختبار أوتيس - لينون للقدرة العقلية " OLMAT " إلى اختبار أوتيس - لينون للقدرة المدرسية (Otis & Lennon, 2005) OLSAT، تطبق بصورة جماعية وهي من اختبارات الورقة والقلم كما تطبق عن طريق الأنترنت، ومدة تطبيقها متباينة حسب المستويات وأقصى مدة يستغرقها المختبر 75 دقيقة. ويصلح تطبيقه على الفئات العمرية التي تمتد من التعليم التحضيري إلى غاية الصف الثاني عشر. تتكون من سبعة مستويات حيث تمتد من الروضة إلى غاية الصف الثاني عشر.

يقيس OLSAT القدرات المعرفية المتعلقة بقدرة التلميذ على التعلم في المدرسة. فهو يقيم القدرة على التفكير التجريدي والتفكير لدي التلميذ، كما يزود المعلمين بمعلومات تمكنهم من تحسين الفهم للاختبارات التحصيلية التقليدية المطبقة في الصف. وذلك عندما يتم تطبيقه مع سلسلة الاختبارات التحصيلية لستانفورد (الطبعة العاشرة)، يمكن أيضاً استخدام درجات OLSAT 8 لربط الإنجاز الفعلي للتلميذ وقدراته المدرسية. يتضمن مهام مثل اكتشاف التشابهات والاختلافات، وتذكر الكلمات والأرقام، وتحديد الكلمات، اتباع التعليمات، التصنيف، ووضع تسلسل، وحل المشكلات الحسابية، واستكمال التشبيهات ... (Otis & Lennon, 2005)

- اختبار كاليفورنيا للنضج العقلي: هو واحد من اختبارات الذكاء الحديثة التي جذبت الانتباه، كونه يحاول القيام بقياس تشخيصي للوظائف العقلية، وقد صمم الاختبار لقياس الأنواع التالية من القدرات العقلية: الذاكرة، العلاقات المكانية، والمفاهيم اللفظية، الاستدلال المنطقي. ويندرج تحت كل قدرة عدة اختبارات فرعية، تنتج هذه الاختبارات ثلاث معاملات للذكاء، معامل الذكاء (Qi) خاص بالعامل اللغوي، معامل الذكاء (Qi) خاص بالعامل غير اللغوي، وكذلك النوع المعتاد من معامل الذكاء القائم على الدرجة الكلية للاختبار.

- اختبار كولمان أندرسون (الطبعة الثامنة): يصلح تطبيقه على الفئات العمرية التي تمتد من رياض الأطفال إلى غاية الصف الدراسي الثاني عشر، وتتراوح مدة تطبيقه ما بين 50 و 75 دقيقة، يحتوي على ثمانية مستويات وكل مستوى تدرج تحته عديد من الاختبارات مع بنود متنوعة (Kaplan & Saccuzzo, 2009). بنود اختبار KAT في بدايتها غير لفظية (البنود التي تقع في المستويات المنخفضة غير لفظية وتكون لفظية في المستوى الأعلى)، وهو من الاختبارات المفضلة أيضاً لدى الفئات ذوي الحاجات الخاصة،

ويتميز أنه يعطي ثلاث درجات للذكاء: درجة لفظية ودرجة كمية وأخرى عامة، ويعبر عن درجاته بنسبة الذكاء الإنحرافية ومعايير المدى المئني لتقدير فترات الثقة في الرتب المئينية (Kaplan & Saccuzzo, 2009).

-**اختبار هينمون نيلسون (Henmon-Nelson Test (H-NT):** وهو اختبار جيد ومقنن واسع الاستخدام، مصمم بعناية ليقس القدرات العقلية لجميع مستويات الصفوف الدراسية. وتنتج عنه درجة عامة واحدة مرتبطة بالعامل العام عند سبيرمان، ويعتقد أنها تعكس الذكاء العام، فهو يتجاهل أو لا يفكر في الذكاءات المتعددة. يتوفر على مجموعتان من المعايير، المجموعة الأولى تستند على توزيعات للدرجات الخام حسب العمر، والأخرى على توزيعات للدرجات الخام حسب الصف الدراسي. ويمكن تحويل الدرجات الخام إلى معامل الذكاء الانحرافي وكذلك إلى نسب مئوية، فهو يمنح الحصول على قياس سريع نسبياً للذكاء العام (يستغرق حوالي 30 دقيقة لإكمال 90 عنصراً) (Kaplan & Saccuzzo, 2009).

باختصار اختبار H-NT أداة سليمة للغاية. حيث يمكن أن يساعد في التنبؤ بسرعة بالنجاح الأكاديمي في المستقبل، ويجب أن ننتبه إليه بعض القيود الهامة عند استخدامه كأداة واحدة للفحص لاختيار الموهوبين أو تحديد صعوبات التعلم في الأقليات، والأطفال المتنوعين ثقافياً والمحرومين اقتصادياً.

ب-اختبارات الذكاء غير اللفظية للمجموعة: لا تقتضي هذه الاختبارات استخدام اللغة وتطبق على مجموعة من الأفراد في وقت معين، إذ تتطلب اختبارات الأداء التلاعب بالأشياء أو المواد الملموسة الموردة في الاختبار حسب الموضوع. فالاستجابات ذات طابع ذاتي محض ونادراً ما تتطلب استخدام الورق والقلم من قبل الممتحن (باستثناء حالات مثل اختبار المتاهة وما إلى ذلك) هذه الاختبارات لا تحتوي على كلمات أو أرقام. إنها تحتوي على صور ورسوم بيانية وأشكال هندسية مطبوعة في كتيب. والمطلوب هنا القيام بأنشطة لملى بعض المساحات الفارغة، واستخلاص بعض الأشكال الهندسية البسيطة للإشارة إلى أوجه التشابه والاختلاف وما إلى ذلك. وعلى الرغم من أن المختبر يستخدم الورق والقلم، فإنه لا يحتاج إلى معرفة الكلمات أو الأرقام العددية. ما يفعله الفاحص هو شرح واضح من خلال تمثيل ظاهري، ومن الأمثلة على هذا النوع من الاختبارات نذكر ما يلي:

-**اختبار رسم الرجل لجود انف (Goodenough-Harris Drawing Test (G-HDT):** من إعداد الباحثة الأمريكية جود انف. يعد من الاختبارات غير اللفظية المتحررة من أثر الثقافة، يطبق بصورة جماعية كما يمكن تطبيقه فردياً، من مميزاتة سهولة التطبيق والتصحيح، اقتصاد في الجهد والوقت والتكاليف، كما يعتبر أداة جيدة في الدراسات المسحية التي تهدف إلى تصنيف أطفال الروضة والمرحلة الابتدائية بحسب مستوياتهم العقلية. بساطة العمل المطلوب في هذا الاختبار يجعل التلاميذ منجذبين إليه، ولا يشعرون بالخوف والتهديد، يتفرد أيضاً هذا الاختبار بالتقائية في رسومات الأطفال المختبرين والتي توفر بدورها أي الرسومات لغة غنية بالمعاني النفسية التي تتخطى عوائق التعبير اللفظي وتجعل هذا الاختبار أداة جيدة

في قياس النمو العقلي عند الأطفال العاديين وذوي الاحتياجات الخاصة. يستخدم (G-HDT) (G-HDT) على نحو ملائم مع مصادر المعلومات الأخرى (Kaplan & Saccuzzo, 2009).

- اختبار المصفوفات المتتابعة لرافن Raven Progressive Matrices RPM: هو واحد من أفضل الاختبارات الجماعية غير اللفظية المعروفة والأكثر شعبية. فعلى الرغم من استخداماته في البيئات التعليمية بالدرجة الأولى، إلا أنه اختبار مناسب لتقدير الذكاء العام للفرد. ويمكن تطبيق اختبار RPM بصورة جماعية وفردية أيضا، على الرغم من أن الكتاب السنوي للقياس العقلي لا يذكره كاختبار فردي (Kaplan & Saccuzzo, 2009) يناسب الفئات العمرية من خمس سنوات إلى غاية الرشد وتتراوح مدة تطبيقه بين 20 و45 دقيقة، وله تعليمات بسيطة. يستند اختبار RPM في قياسه للذكاء على نظرية سبيرمان للعامل العام.

يقدم فكرة عن قدرة شخص ما على الملاحظة وحل المشكلات والتعلم. يحتوي على 60 بندا معروضة في 5 مجموعات (A-E)، مع 12 بندا لكل مجموعة. يقلل من تأثير العامل الثقافي على الدرجة المحصل عليه.

- اختبارات كاتل المتحررة من أثر الثقافة.

- اختبار دافيز - ايلز els-Davis لقياس الذكاء العام.

- اختبار بيتا للجيش test Beta Army .

- اختبار متاهة بورتيوس test Maze Porteus PMT إعداد بورتيوس

مزايا الاختبارات الجماعية: تتفرد اختبارات الذكاء ذات التطبيق الجماعي بمزايا وخدمات متعددة وفريدة نذكرها فيما يلي:

- اختبارات الذكاء الجماعية فعالة من حيث التكلفة، فهي توفر الوقت في أثناء التطبيق والتقييم (تصحيح الاختبار)، تتضمن مواد أقل تكلفة ولا تتطلب تدريب وكفاءة عالية من الفاحصين عكس الاختبارات الفردية.

- التقييم في الاختبارات الجماعية أكثر موضوعية وبالتالي أكثر موثوقية من التقييم الذاتي في العديد من الاختبارات الفردية.

- تسمح بتطبيقها على أعداد كبيرة في نفس الفترة الزمنية.

- بما أن اختبارات الذكاء الفردية تطبق بصفة أكبر لأغراض التقييم وتشخيص المشكلات النفسية أو الصحية، نجد الاختبارات الجماعية تطبق بشكل أوسع في المدارس وعلى جميع الصفوف الدراسية، الجيش، الصناعة، وفي أغراض البحث.

- يمكن استخدام نتائج الاختبارات الجماعية لأغراض الفحص والكشف، الاختيار، لتقييم القدرات العقلية والمهنية الخاصة.

محاضرة رقم 11- قياس التحصيل الدراسي:

1-الاختبارات التحصيلية: من المهم الإشارة إلى معنى كلمة التحصيل قبل التطرق إلى اختبارات التحصيل، ترتبط كلمة التحصيل بشكل مباشر بالأداء الدراسي للمتعلمين، إذ تستخدم للإشارة إلى المدى الذي تحققت فيه الأهداف التعليمية. فالتحصيل "هو درجة الاكتساب التي يحققها الفرد، أو مستوى النجاح الذي يحرزه أو يصل إليه في مادة دراسية أو مجال تعليمي" (علام، 2000) وتعرف الاختبارات التحصيلية بأنها تلك الاختبارات التي تقيس المعارف والمهارات التي تعلمها الفرد في مجال محدد، بعد تعرضه لخبرة تعليمية رسمية أو غير رسمية، وتشمل اختبارات التحصيل الاختبارات التي يصممها المدرسون في الفصول الدراسية، والاختبارات المقننة التي تضعها دوائر متخصصة، مدارس محلية أو وطنية، أو منظمات دولية (Fernandez-Ballesteros, 2003).

2-أغراض اختبارات التحصيل: تستخدم الاختبارات التحصيلية للأغراض التالية:

-أغراض: مثل قياس تحصيل المتعلمين، تعيين الدرجات، لتحديد الانتقال إلى القسم الأعلى أو النجاح في الدورة، مقارنة نتائج التلاميذ في إطار نفس النظام التعليمي وبين مختلف النظم التعليمية العالمية. تقييم فعالية العملية التدريسية، البرامج والمحتويات، والمقاطعات، البلدان في إطار المساءلة التربوية.

-أغراض تكوينية: مثل تحديد وتشخيص نقاط القوة والضعف لدى المتعلمين، تحفيز التلاميذ والمدرسين والإدارة لتحقيق مستويات أعلى من الأداء. كما تعلمنا عن نجاح السياسة التعليمية من عدمها.

-أغراض التنسيب والتشخيص: مثل انتقاء التلاميذ وتوزيعهم على مستويات مقرر معين أو منحهم مكانا مقدما في مادة ما، كذلك تساعد في تشخيص اضطرابات التعلم، والموهبة وغيرها من الاحتياجات الخاصة.

3-تصنيف الاختبارات التحصيلية: يمكن تصنيف الاختبارات التحصيلية المقننة إلى عدة أنواع، منها ما يصنف بناءً على الغرض من الاختبار، طول الاختبار، طريقة تطبيق الاختبار، شكل بنود الاختبار، طريقة تفسير درجات الاختبار (Fernandez-Ballesteros, 2003).

1-3. الغرض من الاختبار: تصنف الاختبارات التحصيلية هنا إلى نوعين، اختبارات الفحص أو الكشف، والاختبارات التشخيصية الشاملة. تميل اختبارات الكشف لأن تكون موجزة نسبيا مع اختبار فرعي واحد يغطي كل نطاق محتوى، مصممة لمسح عدد كبير من التلاميذ بسرعة لتحديد الحالات التي تجب إحالتها لمزيد من التقييم المعمق. أما الاختبارات التشخيصية الشاملة عادة ما تتضمن أكثر من اختبار فرعي واحد لكل نطاق محتوى، وكل اختبار هنا بإمكانه أن يقدم كشفا أكثر عمقا.

2-3. طول الاختبار: هناك اختبارات ذات الموضوع الواحد أو المادة الواحدة تتضمن عدد محدد من الاختبارات الفرعية، تتراوح بنودها بين تقييم المستويات الأدنى للمهارات إلى المستويات الأعلى في مجال واحد على الأقل. واختبارات ذات المواد المتعددة تقيم على الأقل ثلاث مواد دراسية، مثل القراءة، الرياضيات

واللغة الكتابية، ونذكر منها اختبارات ايوا Iowa، سلسلة اختبارات الأداء ستانفورد، اختبارات التحصيل والاتقان.

3-3. طريقة التطبيق: تنقسم اختبارات التحصيل حسب طريقة التطبيق إلى قسمين، اختبارات جماعية تطبق على مجموعة من المختبرين في نفس التوقيت الزمني داخل الفصول الدراسية، تكون في الغالب متعددة المواد وتتضمن اختبارات فرعية تتناسب مع مختلف المستويات الدراسية، واختبارات فردية تطبق بشكل فردي، يمكن أن يكون لها موضوع واحد أو موضوعات متعددة، تستخدم في الحالات الإكلينيكية والتربوية.

3-4. شكل البنود: تصنف الاختبارات التحصيلية بناء على شكل البنود إلى اختبارات موضوعية ذات الإجابة الثابتة أو المحددة، واختبارات المقال ذات الإجابة الانشائية. تتميز الاختبارات الموضوعية بإمكانية تغطية محتوى المادة بشكل كامل، وسرعة التنفيذ والتصحيح، الصعوبة في الإعداد والموضوعية في التصحيح، شكل الأسئلة الأكثر شيوعا في هذا النوع من الاختبارات هو الاختيار من متعدد، الصواب والخطأ، المطابقة، التكملة أو ملأ الفراغات. غير انها تشوبها عيوب مثل اهمالها لمهارات التفكير العليا وحل المشكلات، كذلك تثبط التفكير الإبداعي.

أما الاختبارات المقالية ميزتها الأساسية أنها تتطلب بناء الإجابة من المختبر وليس تحديد الاختيار الصحيح من السؤال. وتقييم قدرة التلاميذ على التنظيم، وربط الأفكار وحل المشكلات، تقليل تأثير التخمين، وسهولة الإعداد. مساوي البنود ذات الاستجابة المبنية أنها تطرح أسئلة قليلة نسبياً وبالتالي تغطية كافية لنطاق المحتوى وكذا تدخل الذاتية في التصحيح.

3-5. تفسير النتائج: عندما تُفسر نتائج الاختبار التحصيلي في ضوء الجماعة المعيارية، فإن الاختبار يطلق عليه اختبار معياري المرجع، درجات التلاميذ في الاختبار المعياري المرجع تفسر في إطار درجات العمر المكافئة أو الصف الدراسي، تصمم أيضا للتمييز بين أداء التلاميذ دون أن تقدم معلومات عن كمية المعلومات المكتسبة. وعندما يتم تفسير درجات الاختبار من حيث مدى اتقان المتعلم لمعارف أو مهارات معينة دون الرجوع على المجموعة المعيارية، يطلق عليه اختبار محكي المرجع، وعادة ما يتم التعبير عن نتائج هذا النوع من الاختبارات بنسب مئوية أو بعبارات وصفية مثل متحكم/ غير متحكم. هناك بعض الاختبارات المرجعية يمكن تفسير نتائجها معياريا مثل اختبار BASIS.

محاضرة رقم 12: قياس الشخصية:

1- مفهوم الشخصية:

كثيرا ما نستخدم مفهوم الشخصية في حياتنا اليومية ونطلق أحكاما على الآخرين فنقول عن هذا أنه "شخصية قوية" وآخر "ليس له شخصية" وإلى غير ذلك من الأحكام التي نشكلها عن الآخرين سواء بوعي منا أو بغير وعي، هذه الأحكام غالبا ما تشير إلى الكفاءة الاجتماعية للشخص والقبول الاجتماعي له. فاستخدام مفهوم الشخصية من قبل جهات غير مختصة في علم النفس يشير إلى حكم ضمني شامل متكون من جميع الانطباعات والمشاعر التي أسسها أو كونها شخص ما أو مجموعة أشخاص عن الآخر، هذه التقييمات تبقى ذاتية تركز على الخصائص السطحية الأكثر وضوحا.

لكن نحن بحاجة إلى أن نكون علميين، وإلى استخدام أساليب تجريبية لتقديم مفهوم موضوعي وأعمق للشخصية، فالبحث في الشخصية يركز على الوصف الموضوعي والشامل والدقيق، ففهم شخص ما، يستدعي مزيدا من المعرفة عنه، مثلا حول النمط السلوكي له، وطاقفه العقلية، الدوافع والمعتقدات والقيم التي يحملها، وكذا طريقة تنظيمها. إن تجميع هذه الأبعاد يشكل لنا الشخصية الظاهرية لهذا الشخص. وقد أفرز البحث في مفهوم الشخصية عشرات التعريفات المقدمة لها، ففي دراسة مسحية مبكرة أجراها Gordon Allport 1937 أحصى فيها ما لا يقل عن خمسين تعريفا قدم للشخصية (علام ، 2000) هذا الاختلاف في تحديد مفهوم الشخصية يعكس تعقّد المفاهيم المتعلقة بالسلوك الإنساني من جهة، والاستخدامات المختلفة لمفهوم الشخصية من جهة أخرى.

لعل أبرز التعريفات التي قدمت للشخصية تعريف Gordon Allport 1937 "الشخصية تنظيم ديناميكي للنظم النفسية الجسدية داخل الفرد، والتي تحدد تكيفه الفريد مع البيئة" وفي عام 1961، قام بتعريفها مرة أخرى، "الشخصية تنظيم ديناميكي داخل الفرد للأنظمة النفسية الجسدية التي تحدد الأنماط المميزة لسلوكه وتفكيره" (Coaley , 2010).

وفي إطار تعريف الشخصية واختلاف الباحثين في تحديدها نجد (Mayer, 2007) صرح معتقدا أن هناك تعريفا مركزيا للشخصية يستخدم اليوم (وتاريخيا) داخل تخصصنا. على الرغم من صياغته من قبل علماء نفسيين مختلفين، فإن فكرته المركزية تبقى كما هي:

"الشخصية عبارة عن نظام من الأجزاء التي يتم تنظيمها وتطويرها والتعبير عنها في تصرفات الشخص" نظام الأجزاء يشمل: مكونات مثل الدوافع، العواطف، النماذج العقلية، والذات. واستدل على ذلك باستعراضه لخمسة تعريفات للشخصية من خمس كتب لباحثين مختلفين وهي:

- تشير الشخصية إلى الأنماط المميزة للفكر، للعاطفة والسلوك لدى الفرد، إلى جانب الآليات النفسية - الخفية أو التي تقف وراء تلك الأنماط (Funder، 2004).

- "الشخصية هي مجموعة من السمات والآليات النفسية داخل الفرد التي يتم تنظيمها واستمرارها نسبيا والتي تؤثر على تفاعلاته مع البيئات النفسية والفيزيائية والاجتماعية والتكيف معها" (Larsen & Buss، 2005).

- علم النفس الشخصية هو الدراسة العلمية للفرد من جميع النواحي... علم النفس يهتم بعدد الأمور: الإدراك والانتباه، المعرفة والذاكرة، الأعصاب ودوائر الدماغ... نحن نحاول أن نفهم الفرد البشري ككل مركب... [و] تقديم وصف علمي جامع ذا مصداقية حول الفرد البشري. (McAdams, 2006).

- الشخصية هي نظام مرتب وناشئ داخل الفرد، والذي يمثل العمل الجماعي للأنظمة الفرعية النفسية الرئيسية لهذا الفرد (Mayer, 2007)

الشخصية تشير إلى خصائص الشخص التي تمثل أنماطاً ثابتة من المشاعر والتفكير والسلوك (Pervin, Cervone & John, 2005).

التعريفات السابقة كثيراً ما تشترك في الرأي القائل: (أ) الشخصية هي نظام نفسي (ب) يتألف من مجموعة من الأجزاء (د) التي تتفاعل (هـ) وتتطور (و) لتؤثر على التعبير السلوكي للفرد. وذكر (Coaley, 2010) أيضاً أن العديد من المراجع الأكثر شيوعاً أجمعت على النقاط التالية في تعريف الشخصية:

- أسلوب أو نمط من السلوك.
- خصائص مستقرة نسبياً أو دائمة تمكن التنبؤ.
- التفرد، حيث تميز بين فرد وآخر.
- التكيف أو التفاعل مع البيئة.
- أنماط مميزة من السلوك والتفكير والمشاعر.

2- قياس الشخصية: قياس الشخصية هو أي إجراء منهجي ينسب الأرقام إلى الخصائص المميزة لأسلوب الشخص مقارنة بالأشخاص الآخرين، وفقاً لبعض القواعد الواضحة. يمكن بعد ذلك استخدام هذه الأرقام لعمل تنبؤات حول استجابات ذلك الشخص في ضوء مؤثرات معينة مستقبلاً (Hogan, Hogan, & Roberts, May 1996) كما أن القياس الجيد للشخصية لا بد أن يكون له خاصيتين على الأقل، الأولى أن تكون درجات القياس ثابتة عبر الزمن (بمعنى استقرار النتائج بمرور الوقت). والثانية، تتطلب أن تكون هناك أدلة ذات مصداقية تشير إلى أن الدرجات في القياس تتعلق بمؤثرات السلوك المختبر (أي أن الدرجات لا بد أن تتنبأ بالأداء الحقيقي) (Hogan, Hogan, & Roberts, May 1996). فالمشكلة المطروحة في مجال قياس الشخصية إذن ليست على مستوى ماهية هذا القياس، بل في جودة هذا القياس، فهناك مجموعة صغيرة فقط من اختبارات أو أدوات قياس الشخصية تستجيب للمعيارين المذكوران سابقاً.

3- اختبارات الشخصية: هناك مجموعة واسعة من اختبارات الشخصية، البعض منها يطلق عليها أدوات محددة النطاق، تسعى لقياس تكوين فرضي واحد، يعتقد أنها ذات قيمة تفسيرية وتنبؤية كبيرة، والأخرى يطلق عليها أدوات غير محددة النطاق أو غامضة الهدف. تسعى لقياس تكوينات فرضية متعددة (Krug, 2013). وهذه الأخيرة تستخدم محفزات غير منظمة نسبياً، كما ينتج عنها نطاق لا نهائي من الاستجابات. ونستعرض هذين القسمين العريضين من مقاييس الشخصية هما: المقاييس الموضوعية للتقرير الذاتي، والأساليب الإسقاطية.

1-3. الاختبارات الموضوعية: لسنوات عديدة يُستخدم مصطلح الموضوعية لتمييز مجموعة من الاختبارات عن المجموعة الأخرى، مثل اختبار رورشاخ واختبار تفهم الموضوع والتي يطلق عليها تسمية الاختبارات الإسقاطية، هذه الأخيرة تستخدم مثيرات غير محددة البنية نسبياً مثل بقع الحبر والرسومات التي تتطلب قدرة عالية لتسجيل الدرجات وتبقى إشكالية استقرار الدرجات عبر القياسات المختلفة أهم عيوب الاختبارات الإسقاطية. في حين الاختبارات الموضوعية تتطلب الإجابة على مجموعة معيارية من البنود، محصورة نطاق الإجابة مثل مفردات الصواب والخطأ، وموازنين التقدير الذاتي (مطلقاً، أحياناً، غالباً، دائماً) أو مفردات جبرية الاختيار، التي تعتمد على الاختيار بين عبارتين متساويتين في إمكانية قبولهما من حيث المرغوبة الاجتماعية. مثلاً: "أستمتع بقضاء أوقاتي في المذاكرة"، و "أستمتع بقضاء أوقاتي في مشاهدة الأفلام" وينبغي على المختبر هنا أن يختار العبارة التي تصفه أكثر أو بدرجة أفضل. ونستخدم هنا قواعد بسيطة للحصول على الدرجة النهائية لاستجابات المختبر، هذا الشكل من الاختبارات يوفر لنا ثبات جيد عبر الزمن، كما تتصف أنها اقتصادية.

يستخدم مصطلح الموضوعية في بعض الأوساط للدلالة عن دقة النتائج التي نتحصل عليها من نوع معين من الأدوات، وعدم تحيزها واتصافها بالعلمية في مقابل أدوات أخرى تعتبر غير دقيقة، ومتحيزة وغير علمية، ولتجنب مثل هذه الأحكام القيمة ودلالاتها اقترح العديد من الباحثين استخدام مصطلحات محددة البنية وغير محددة البنية (Wiggins, 1973). في حين يستخدم (Cattell 1957) مصطلح الموضوعية في الحكم على المعلومة التي تم جمعها عن المفحوص ومدى بعدها عن التشويه والتحيز سواء عن طريق التقرير الذاتي أو عن طريق تقييمات الملاحظين. فالاختبارات الموضوعية إذن تكون محددة الهدف وواضحة البنية

-المدائل الرئيسية لبناء المقاييس الموضوعية للتقرير الذاتي:

لبناء وتطوير اختبارات التقرير الذاتي للشخصية، يجب التوقف عند هذه الأسئلة ومحاولة الإجابة عنها كونها الطريقة الأنسب لوضع مثل هذه الاختبارات (Cohen, Montagu, Nathanson, & Swerdlik, 1988)

- ما الهدف من اختبار الشخصية الذي طوره؟ أي تحديد الغرض من تصميم الاختبار.
- هل يتم استخدامها لقياس السمات أو الأنواع أو الحالات أو المزج بين هذه الأبعاد؟
- هل يتم استخدامها لقياس القوة النسبية لسمات مختلفة؟ إذا كان الأمر كذلك، ما هي السمات التي سيتم قياسها؟
- هل يتم استخدامها لتمييز الأفراد على أساس صحة شخصيتهم؟ هل يمكن استخدامها لتمييز الأفراد على أساس ملاءمة شخصياتهم لنوع معين من العمل؟ هل يجب استخدامه في البحث العام عن الشخصية؟

- ما هي أنواع البنود التي سيحتويها الاختبار؟ كيف يمكنك أن تقر محتوى وصياغة هذه البنود؟ هل يمكنك، على سبيل المثال، الاعتماد على نظرية معينة للشخصية في صياغة هذه البنود؟ أم أنك لن تعتمد على نظرية معينة، بل على تجاربك الخاصة؟
- عند كتابة بنود الاختبار الخاصة بك، هل استخدمت صيغة صواب / خطأ أو تنسيق آخر؟ هل سيتم تجميع بنود الاختبار على أساس ترتيب معين؟
- كيف يمكنك أن تثبت بصورة مقنعة أن اختبارك يقيس ما يعتزم قياسه؟

هذه الأسئلة أوجدت صعوبة في الإجابة عليها لدى مطوري اختبارات الشخصية، وهناك من اعتمد على نظريات الشخصية في وضع البنود، واستخدم آخرون الأساليب التجريبية. كما ابتكروا أيضا تصميمات تهدف إلى تقديم توصيف "جرد" عام للشخصية، بينما ابتكر آخرون أشكالاً أخرى لقياس جوانب معينة، مثل قوة سمة معينة. فمقاييس التقرير الذاتي للشخصية تختلف فيما يتعلق بالأساس المنطقي لنموذج القياس الذي يكمن وراء بناء الاختبار. وقد تم تصنيف المداخل أو الاستراتيجيات المختلفة لبناء الاختبار في عدد من الطرق أو المداخل المختلفة، وسوف ننظم مناقشتنا حول المداخل الرئيسية لبناء المقاييس في أربعة طرق رئيسية وهي التقسيمات الأكثر انتشارا في مختلف مراجع القياس النفسي (Reynolds & Cohen, Montagu, Nathanson, & Swerdlik, 1988) هي:

(1) مدخل منطقية المحتوى (2) مدخل المحك الإمبريقي (3) التحليل العاملي (4) المدخل النظري.
مدخل منطقية المحتوى: يعتمد هذا المدخل في كتابة مفردات المقاييس الموضوعية للشخصية على مدى مواعمتها الظاهرية للتكوين الفرضي المراد قياسه.

وأول الجهود الرسمية لقياس الشخصية استخدمت المدخل المنطقي لبناء الاختبار وتعد قائمة البيانات الشخصية لـ (Woodworth, 1917) أول اختبار طور لقياس الشخصية، صممت للمعاونة في جمع البيانات الشخصية المتعلقة بالمجندين في الجيش ابان الحرب العالمية الأولى، فقررت الاختبار تهدف إلى جمع معلومات عن شخصية المفحوصين عن طريق التقرير الذاتي، هذه المعلومات تتعلق بالانحرافات العصائية (Cohen, Montagu, Nathanson, & Swerdlik, 1988).

مدخل المحك الإمبريقي: يمكن تلخيص بناء اختبار الشخصية من خلال مدخل المحك الإمبريقي فيما يلي:
 - صياغة عدد معين من البنود لتشكل الاختبار الذي يفترض أن يقيس سمة واحدة أو أكثر.
 - تطبيق بنود الاختبار على مجموعتين من الأفراد كأقل حد: "مجموعة معيارية" تتألف من أشخاص تعرف أنهم يمتلكون السمة التي يتم قياسها، ومجموعة أخرى ضابطة يفترض أن أفرادها لا يمتلكون السمة المراد قياسها أو يمثلون عامة الناس.

- فحص المفردات إحصائيا والإبقاء على البنود التي تميز بين المجموعتين (المعيارية والضابطة).
 ويشار إلى هذه الطريقة في بناء الاختبار على أنها "تجريبية" (إمبريقيية)، لأنه يتم الاحتفاظ فقط بالبنود التي تثبت وجود علاقة فعلية (تجريبية) بين بنود الاختبار والسمة المراد قياسها (Cohen, Montagu,)

(Nathanson, & Swerdlik, 1988). وقد تمت الإشارة أيضاً إلى هذا الأسلوب في بناء الاختبار على أنه طريقة "المجموعات المتناقضة". ويعد اختبار مينيسوتا للشخصية متعدد الأوجه Minnesota Multiphasic Personality Inventory (MMPI) 1943 وقائمة كاليفورنيا النفسية California Psychological Inventory (CPI) من الاختبارات التي تم بناؤها باستخدام المحك الإمبريقي.

مدخل التحليل العاملي: التحليل العاملي أحد المداخل الإحصائية الذي يسمح لنا بتقويم وجود تكوينات فرضية بين مجموعة من المتغيرات، وله دور مهم في تحديد بنية الشخصية بنماذج يتراوح ما تشتمل عليه بين عدد قليل من العوامل وعوامل متعددة (Reynolds & Livingston, 2013). ويعد ريموند كاتل أول من نشر استبيان عوامل الشخصية والتي عدده 16 عام 1949. والطبعة الأحدث لهذا الاستبيان نشرت 1993. وهو مقياس شائع الاستخدام لوظائف الشخصية السوية وبخاصة في الإرشاد، والمواقف المدرسية. وفي السنوات الأخيرة نال نموذج العوامل الخمسة لل شخصية NEO قبولاً واستخداماً واسعاً.

المدخل النظري: بعض اختبارات الشخصية ترتبط ارتباطاً وثيقاً بنظرية معينة للشخصية، حيث تكون بنود هذا النوع من الاختبارات كلها مصممة لقياس سمات أو حالات يفترض أنها موجودة على أساس تلك النظرية، على سبيل المثال، قد يكون اختبار الشخصية الذي تم بناؤه في إطار نظرية التحليل النفسي يحتوي على بنود مصممة لتقييم الهوية، والأنا، والأنا الأعلى. وكمثال لبعض اختبارات الشخصية التي اعتمد في بنائها المدخل أو الاستراتيجية النظرية، نجد مؤشر النمط مايرز - بريجز Myers-Briggs 1985 والذي استند إلى تصنيف الشخصية الذي وضعه كارل يونغ، نموذج بحث الشخصية (استناداً إلى عمل هنري موراي)، نموذج بحث الشخصية (جاكسون 1984) الذي كان مستنداً إلى عمل هنري موراي، وجدول ادواردز التفضيل الشخصي (EPPS)، الذي وصفنا بيلو Belo.

النزعة المعاصرة في بناء مقاييس الشخصية تعتمد على مداخل واستراتيجيات متعددة عملية التصميم والبناء.

1-3. الاختبارات الإسقاطية للشخصية: تعد الاختبارات الإسقاطية اتجاهاً آخر في قياس الشخصية تختلف عن الاختبارات الموضوعية. وتتميز هذه الاختبارات بمهام غير محددة ومواد غامضة، حيث يتم تشجيع الأفراد الذين يتم اختبارهم على التعبير عن أنفسهم بطريقة تكشف عن البنية الأساسية وديناميكيات شخصيتهم. وذكرت في هذا الصدد Anne Anastasi، "أن التقنيات الإسقاطية تتميز بالمقاربة الشمولية لتقدير الشخصية. وتركيز الاهتمام على الصورة المركبة للشخصية بأكملها، وليس على قياس السمات المنفصلة. عادة ما تعتبر الأساطير تقنيات إسقاطية فعالة بشكل خاص في الكشف عن الجوانب السرية أو الكامنة أو اللاواعية للشخصية" وأضافت أيضاً "أن الاختبار الإسقاطي مهمة غير محددة البنية نسبياً، تسمح باستجابات غير محدودة تقريباً ومتنوعة. من أجل السماح باللعب الحر في خيال الفرد، لا تقدم سوى تعليمات عامة وموجزة" (Anastasi & Urbina, 1996) فهذه الاختبارات تعد بمثابة شاشة بيضاء يُسقط عليها المختبرون أفكارهم، ورغباتهم، ومخاوفهم، وحاجاتهم، وصراعاتهم أكثر تعلقاً بهم (Anastasi &

(Urbina, 1996) في ضوء هذه التعريفات المقدمة للاختبارات الإسقاطية نوجز أهم مميزاتها في النقاط التالية:

-الموقف المثير الذي يستجيب له الفرد غير متشكل وناقص التحديد والانتظام.
-يستجيب الفرد للمادة غير المتشكلة التي تعرض عليه دون أن تكون لديه معرفة عن كيفية تقدير الاستجابات.

-توفر للمختبر حرية ليعبر عن أفكاره ومشاعره وانفعالاته ورغباته في تشكيل المادة غير المتشكلة نسبياً.
-لا تقيس نواحي جزئية أو وحدات مستقلة تتألف منها الشخصية، بل تحاول أن ترسم صورة عن الشخصية ككل ودراسة مكوناتها وما بينهما من علاقات ديناميكية.

* **تقسيمات الاختبارات الإسقاطية:** قدمت تقسيمات متعددة للاختبارات الإسقاطية وأشهر هذه التقسيمات، التقسيم الذي قدمه لورانس فرانك، والأساس الذي اعتمده في هذا التقسيم هو نوع الاستجابة التي نحصل عليها من الفرد والهدف من الاختبار. ويشتمل على خمس تقسيمات هي (عباس، 2001):

-**الطرق التكوينية:** يطلب من المختبر أن يفرض على المادة المعروضة عليه نوعاً من التكوين والتنظيم، هذه المادة تكون في أساسها غامضة أو قريبة من الغموض وغير متشكلة وغير منتظمة، وبعد اختبار يقع الحبر لرورشاخ Rorschach مثال لهذا النوع من الاختبارات.

-**الطرق البنائية أو الإنشائية:** تتطلب من المختبر تشكيل مادة متشكلة متكونة ذات معنى محدد وخاص ومتميز ويعتبر اختبار لوينفلد lowenfeld الفيسفاسي مثال على هذه الاختبارات، حيث يتطلب من المفحوص ترتيب الأجزاء المختلفة الألوان والأشكال في صورة نماذج، تسمح لنا هذه الطرق بالحصول على مادة إسقاطية حين يكون المختبر وخاصة الطفل منهمكا أو مستغرقاً في نشاط اللعب أو الرسم أو التلوين.
-**الطرق التفسيرية:** هذه الطرق تقدم للمفحوص موقفاً أو عملاً أو صورة يستجيب إليه عن طريق القيام بنشاط مبدع يعبر فيه عن أفكاره ومشاعره وآماله. وتعد اختبارات تفهم الموضوع (T.A.T) أكثر الاختبارات استخداماً في الطرق التفسيرية. حيث نطلب من المختبر بعد أن نريه الصورة أو الرسم تأليف قصة عليها، كما يعد اختبار الإدراك الواعي للأطفال الذي أعده روبرتس Roberts-2 الأكثر حداثة وهو مصمم للأطفال والمراهقين الذين تتراوح أعمارهم بين 6 و18 عاماً.

-**الطرق التفرغية أو التطهيرية:** هذا النوع يساعد على التخفيف والتخلص من الانفعالات، وكثير من أنواع اللعب العلاجي للأطفال يشتمل على ناحيتين التخلص من الانفعالات والتعبير عنها.

-**الطرق التحريفية:** طريقة استخدام المادة أو المثير سواء كانت لفظية أو غير لفظية تمدنا بوسيلة للكشف عن شخصية الفرد (عباس، 2001).

تشارك اختبارات التقرير الذاتي والاختبارات الإسقاطية في عديد من النقاط:

- الضعف النسبي لمعامل الثبات والصدق
- توفير بعض المعلومات اللازمة للحصول على فهم أفضل للفرد.

- _ كلا النوعين تعترها جوانب قصور .
- يمكن أن تطبقها بشكل فردي أو في مجموعات.
- يتطلب تفسير كل من اختبارات الشخصية المحددة البنية وغير المحددة البنية من قبل أشخاص مؤهلين.
- لأن عملية التطبيق والتسجيل، والتفسير معقدة.
- هذا العرض المقدم للشخصية من تعاريف وطرق القياس والذي يتشارك فيه معظم الباحثين قد يجرنا إلى مشكلة أخرى وهي كيف نميز بين الشخصية (personnalité) والسمات الأخرى (traits)، مثل الميول والاتجاهات والدوافع.

عنوان المحاضرة: قياس الميول:**1-تعريف الميول Interest:**

يعرفه Guilford 1948 «الميل نزعة سلوكية عامة لدى الفرد للانجذاب نحو نوع معين من الأنشطة» (كراجة، 1997)

وعرفها أيضا Gardner and Tamir, 1989 على أنها "تفضيل الفرد المشاركة في أنشطة معينة أكثر من غيرها، وهو نوع محدد من الاتجاهات فعندما نهتم بظارة أو نشاط مخصص سنميل إلى الالتزام بها وبذل الوقت فيها" (خياطة، 2015)

تعريف Mattoo, 2011 مجموعة الخيارات التي يقوم بها الشخص من بين الأنشطة التي ينخرط فيها، حيث يفضل بعضها وينفر من الأخرى. (خياطة، 2015)

وتتميز الميول بمجموعة من الخصائص لخصها سترونج في النقاط التالية:

- الميول مظهر من مظاهر الشخصية المتعددة، وليست مسألة نفسية منفصلة عن غيرها.
- الميول تعبير عن الرضا ولكنه ليس بالضرورة دليلا على الكفاية، فالميل إلى دراسة العلوم لا يدل على المهارة في أدائها.
- على الرغم من إشارة الأبحاث إلى وجود علاقة بين الميول والقدرات، إلا أن قياس القدرات ينبغي ان يتم بطريقة مباشرة، وليس عن طريق اختبارات الميول.
- لتقدير الميول بدقة يجب أن يتم ذلك على يدي مختص مدرب، وتكون أكثر دقة عندما تقوم على أساس التجليل الإحصائي لمئات الحالات.
- معرفة ما يفضله الفرد وما ينفر منه أو يكرهه يمكّننا من أساس جيد لتقدير ما حدث في الماضي وما يمكن أن يحدث في المستقبل، لكنه لا يدل دلالة كبيرة عليهما (العنزي، 2011)

2-أنواع الميول: يميز (super & crites , 1962) بين أربع أنواع رئيسية لمصطلح "الميول " مرتبطة بأربع طرق تسمح لنا بالحصول على بيانات حول ميول الأفراد:

1-2. الميول المعلنة أو المعبر عنها لفظيا: يقصد بها التعبير اللفظي عن الميل نحو نشاط أو مهنة معينة، مثلا يعبر التلميذ عن حبه وميله لنشاط تعليمي معين. تختلف أهمية ودلالة مثل هذه التعبيرات عن الميول مع نضج الفرد وخبرته. في بعض الحالات، تمثل الميول المعبر عنها نزوات أو تخيلات مؤقتة.

2-2. الميول الظاهرة: يتم الاستدلال عليها من خلال مشاركة الفرد في نشاط أو مهنة. مثلاً يعبر الشخص الذي ينشط في نادي ثقافي عن ميوله من خلال الإقبال والمشاركة الفعلية في النشاط. الميول الظاهرة تتجه إلى أن تكون أكثر ثباتاً من الميول التي يتم التعبير عنها لفظياً كونها تستند إلى الخبرة الفعلية. هذه المقاربة في تحديد الميول هي الأخرى تعترضها قيود عدة. فاستظهار الميول قد يكون مقيد باعتبارات مالية أو عوامل بيئة أخرى، هذا النوع من الميول يوفر أدلة عن الأهداف التعليمية والمهنية الممكنة تسطيحها . غياب ظهور ميل خاص أو محدد لدى فرد معين قد يكون نتيجة الافتقار لفرصة تطويره في البيئة التي يعيش فيها.

2-3. الميول التي تقيسها الاختبارات الموضوعية: هذا النوع من الميول تقاس باختبارات موضوعية ذات مفردات لغوية أو غيرها من المعلومات. يغلب على هذا النوع من الميول صفة الثبات، فحينما يظهر التلميذ ميل نحو مجال ما أو مادة دراسية معينة من خلال اختبار موضوعي نجده يطور معارفه ومهاراته مع مرور الوقت.

2-4. الميول التي تقيسها الاستفتاءات (الحصرية): تتيح طريقة الاستبيانات المقننة تحديد ميول الفرد، من خلال قياسها بعرض قائمة من الأنشطة أو المهن في استبيان، يطلب فيه من المستجيب أن يختار النشاط الذي يفضل من بين هذه الأنشطة والمواضيع المعروضة عليه، يمكن تفسير نتائج كل الفرد على أنها تعكس نمطاً من الميول العالية أو المنخفضة نسبياً في مختلف المجالات.

يمكن أن تكون استبيانات الميول والاهتمامات أدوات مساعدة قيّمة في التوجيه المهني. بيد أن الأدلة المستقاة من المصادر الثلاثة الأولى مفيدة في دراسة صحة الميول المقاسة بالاستبيانات وفي استكمال نتائجها لتقديم التوجيه الصحيح للفرد.

4- طرق قياس الميول والاهتمامات:

ميول الفرد (حب وكراهية. تفضيل ونفور) يمكن تحديدها وقياسها عبر عدة طرق، يقترح كل من (super & crites , 1962) أربع مقاربات يمكن استخدامها في تحديد ميول الفرد:

أ- الأسئلة المباشرة.

ب- الملاحظة المباشرة.

ج- الاختبارات الموضوعية.

د- استبيانات الميول : وتعتبر هذه الأخيرة أفضل أنواع المقاربات المذكورة في قياس الميول نظراً لنجاحاتها.

4-محكات بناء مقاييس الميول: نميز بين ثلاثة محكات أو أساليب يمكن اتباعها في تطوير وبناء مقاييس الميول (اسماعيل، 2004):

(أ) المحك التجريبي أو الواقعي: يتم هنا بناء المقاييس بناء على ميول مجموعات مهنية مرجعية، يعمل أفرادها في هذه المهن أو يقومون بهذه المهمة، يكونون راضيين ومتوافقين معها، ومن أمثلة ذلك اختبار "سترونج" للميول المهنية.

(ب) محك التجانس أو المضمون: يعتمد البناء هنا على مدى ارتباط بنود المقياس ببعضها البعض وكذلك صدق مضمونها، من الاختبارات التي طورت وفق هذا المحك استبيان "كودر" لقياس الميول والاهتمامات.

(ج) المحك المنطقي: يتم اختيار البنود من خلال التعرف على مكونات الميول المهنية المعينة وكذلك خصائص الأفراد الذين ينجحون في هذه المهن. حيث تنتقى مفردات الاستبيان استنادا إلى التحليل النظري أو المنطقي للسمة أو الخاصية المراد قياسها. واستخدام Throp هذا الأسلوب في بناء استبيان الميول المهنية.

5- نماذج لاستبيانات مقننة لقياس الميول:

1-5. استبيان سترونج-كامبل للميول: يعد هذا الاستبيان من أكثر أدوات قياس الميول انتشارا واستخداما، إذ يعد سترونج رائدا لحركة قياس الميول، يهدف هذا الاستبيان على قياس درجة اتفاق ميول الفرد مع ميول الأشخاص الناجحين في مهنة محددة، وتقدر ميول الفرد بمقارنة أجوبة الناجحين في تلك المهنة. يفترض هذا القياس أنه كلما ازداد التشابه بين ميول الفرد وميول الناجحين في مهنة معينة زاد احتمال نجاحه في تلك المهنة، ويبقى هذا النجاح مرتبط بالقدرة التي يمتلكها الفرد في المهنة أو التخصص الدراسي، بنود الاستبيان تعكس أوجه النشاط في عدد من المهن والمواد الدراسية والألعاب الرياضية والهوايات وأنماط التسلية ونماذج من النشاط العقلي، ويؤشر المفحوص على كل سؤال إذا كان يحب ذلك النشاط أو يكرهه. تستغرق مدة تطبيق الاختبار حوالي (40) دقيقة (القاسم، 2010)

2-5. استبيانات كيودر Kuder للميول: يقيس الميول في المستويات التعليمية المختلفة من المرحلة الابتدائية حت مستوى الراشدين، وهي تتناول ثلاثة جوانب أساسية (ميول تخصصية، ميول وظيفية مهنية، ميول شخصية وأنماط سلوكية) (منسي، 1988).

قائمة المراجع:

- Deary , I., & Stough , C. (June 1996). Intelligence and Inspection Time Achievements, Prospects, and Problems. *American Psychologist*, 5 L, 599-608.
- Fernandez-Ballesteros, R. (2003). *Encyclopedia of Psychological Assessment*. SAGE Publications.
- Gardner, M. (2011). Theories of Intelligence. In M. A. Bray, T. J. Kehle, & P. E. Nathan (Ed.), *The Oxford Handbook of School Psychology* (pp. 1-43). Oxford Library of Psychology.
- Goodwin, L., & Leech, N. (2003). The Meaning of Validity in the New Standards for Educational and Psychological Testing. *Measurement and Evaluation in Counseling and Development*, 36, 181-191.
- Julius, S., & Howard, W. A. (2010). Meet an assessment professional, in. In Cohen–Swerdlik, *Psychological Testing and Assessment: An Introduction to Tests and Measurement*. Edition, McGraw–Hill Primis.
- Legendre, R. (1998). *Dictionnaire actuel de l'éducation*. Larousse.
- Legg, S., & Hutter, M. (2007). A Collection of Definitions of Intelligence. *Frontiers in Artificial Intelligence and Applications, Vol.157 (2007) 17-24*, 17-24. Retrieved from www.vetta.org/.../A-Collection-of-Definitions-of-Intelligence.pdf
- Linn, R. L., & Gronlund, N. E. (1995). *Measurement and Assessment in Teaching* (7 ed.). Englewood Cliffs, N.J.
- Murphy, K., & Davidshofer, C. (2005). *Psychological Testing: Principles and Applications* (6 ed.). Pearson.
- Quaigrain, K., & Ato, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education*, 4, 5.
- Richard, J., & Sheila, C. (1999). *Item Analysis for Criterion Referenced Tests*. New York: The Center for Development of Human Services, University of New York.
- Soane, C., & Hawke, S. (2008). *Compact Oxford English Dictionary for University and College Students*.
- Terriot, K., & Ozenne, R. (2015). *L'échelle d'intelligence de Wechsler pour enfants, 5e édition (WISC-V)*. . ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant.
- Tracey, T., & Glidden-Tracey, C. (1999). Integration of Theory, Research Design, Measurement, and Analysis: Toward a Reasoned Argument. *The Counseling Psychologist*, 27, 299-324.
- AERA, APA, & NCIME. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Anastasi, A., & Urbina, S. (1997). *psychological testing* (ed. دار الفكر ناشرون وموزعون). (ص. محمود علام, Trans.) publishing by prentice hall.
- Anastasi, A., & Urbina, S. (1996). *Psychological testing*. New York: Macmillan.
- Brown, F. (1984). *Principles of educational and psychological testing*. N.Y; Holt, Rinechar+and winston.

- Cattell, M., & Farrand, L. (1896). Physical and mental measurements of the students of Columbia University. *Psychological Review*, 3, 618-648.
- Chuderski, A. (2013). When are fluid intelligence and memory isomorphic and when are they not? *intelligence*, 41, 244-262.
- Coaley, K. (2010). *An Introduction to Psychological Assessment and Psychometric*. Los Angeles: sage.
- Cohen, R. J., & Swerdlik, M. E. (2002). *Psychological testing and assessment: An introduction to tests and measurement* (5 ed.). McGraw-Hill Companies, Inc.
- Cohen, R., Montagu, P., Nathanson, L., & Swerdlik, M. (1988). *Psychological Testing: An Introduction to Tests and Measurements*. Mountain View, CA: Mayfield Publishing Co.
- Crocker, L., & Aigina, J. (2008). *Introduction to Classical and Modern Test Theory*. Ohio, USA: Cengage Learning.
- Dombrowski, S., & Mrazik, M. (2008). Test Review: Reynolds, C. R., & Kamphaus, R. W. (2003). RIAS: Reynolds Intellectual Assessment Scales. Lutz, FL: Reynolds Intellectual Assessment Scales. Lutz, FL: *Canadian Journal of School Psychology*, 23, 223-230.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of Educational Measurement*. New Delhi: Prentice, Hall of India private Limited.
- Gignaca, G., & Bates, T. (2017). Brain volume and intelligence: The moderating role of intelligence measurement quality. *Intelligence*, 64, 18-29.
- HAIER, R., SIEGEL, B., TANG, C., ABEL, L., & BUCHSBAUM, M. (1992). Intelligence and Changes in Regional Cerebral Glucose Metabolic Rate Following Learning. *Intelligence*, 16, 415-426.
- Henning, G. (1987). *A Guide to Language testing: Development, Evaluation and Research*. Heinle, Heinle /Thomson Learning Asia.
- Hogan, R., Hogan, R., & Roberts, B. (May 1996). Personality Measurement and Employment Decisions-Questions and Answers -. *American Psychologist*, 466-477.
- Holmbeck, G., & Devine, K. (2009). Editorial: An Author's Checklist for Measure Development and Validation Manuscripts. *Journal of Pediatric Psychology*, 34, 691-696.
- Kaplan, R., & Saccuzzo, D. (2009). *Psychological Testing* (7 ed.). USA: Wadsworth Cengage Learning.
- Knapp, T. (2009). *THE RELIABILITY OF MEASURING INSTRUMENTS*.
- Krug, (2013). APA Handbook of Testing and Assessment in Psychology. In Kurt F. Geisinger, *Objective Personality Testing* (Vol. 1, pp. 315-328). Washington: American Psychological Association.
- Laurent, J. (2010). A Statistics Refresher. In C. Swerdlik, *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (7 ed., p. 73). McGraw-Hill Primis.
- Mayer, J. D. (2007). *Asserting the Definition of Personality*. Retrieved from The Online Newsletter for Personality Science.

- Otis, A., & Lennon, R. (2005). *School Ability Test* (8 ed.). are trademarks of Harcourt Assessment Inc. registered in the U.S. and/or other jurisdictions.
- Pal, H. R., Pal, A., & Tourani, P. (2005). Theories of Intelligence. *Everyman's Science*, 39, 181-192.
- Pettersen, N. (2002). *Évaluation du potentiel humain dans les organisations*. Canada : Presses de l'Université du québec.
- Reynolds, C. r., & Livingston, R. b. (2013). (Ed.) *تقن القياس النفسي الحديث*. (ص. محمود علام، دار الفكر موزعون وناشرون عمان).
- Roid, G. (2003). *Introducing the Fifth Edition of the Stanford-Binet Intelligence Scales, Fifth Edition*. Riverside Publishin.
- Santrok, g. w. (2014). *A topical approach to life-span devlopment* (7 ed.). hill education.
- Shinn, M. R., Marston, D., & Ysseldyke, J. (June/jul1980). Review of Woodcock-Johnson Psycho-Educational Battery. *School Psychology International*, 1, 20-22.
- Singh, K., Junnarkar, M., & Kaur, J. (2016). *Measures of positive psychology*. india: springer india.
- Stevens, S. S. (1946, Jun). On the Theory of Scales of Measurement. *Science, New Series*, 103(2684), 677-680. Retrieved from <http://links.jstor.org/sici?sici=0036-8075%2819460607%293%3A103%3A2684%3C677%3AOTTOSO%3E2.0.CO%3B2-G>
- super, d., & crites, j. (1962). *Appraising vocational. fitness*, New York, Harper and Row.
- Thorndike, R. M., & others. (1991). *Measurement and evaluation in psychology and education*. (5. t. ed, Ed.) N.Y; Mcmillan publishing company.
- Urbina, S. (2011). أختبارات الذكاء. In R. STERNBERG, S. KAUFMAN, R. STERNBERG, & S. KAUFMAN (Eds.), *دليل جامعة كيمريدج للذكاء* (د. القرنة, & ع. عبد اللاه, Trans., pp. 20-37). العبيكان.
- Whiston, S. (2009). *Principles and applications of assessment in counseling*. Belmont, CA,
- Willis, J. O., Dumont, R., & Kaufman, A. S. (2015). Intelligence Testing. In R. L. Cautin, & S. O. Lilienfeld, *the Encyclopedia of Clinical Psychology* (pp. 1491-1507). Wiley-Blackwell.
- عبد الحفيظ مقدم. (2003). *الإحصاء والقياس النفسي والتربوي*, ط2. الجزائر: ديوان المطبوعات الجامعية.
- محمود عبد الحليم منسي. (1988). عمل الأم والسلوك الاجتماعي للأبناء من تلاميذ المرحلة الابتدائية بالمدينة المنورة. *مجلة العلوم الاجتماعية*.
- هبة الله خياطة. (2015). *الميول المهنية ومستوى الطموح في ضوء بعض المتغيرات*. رسالة ماجستير غير منشورة.
- السيد، ف. ا. (1989). *علم النفس الإحصائي وقياس العقل البشري*. مصر: دار الفكر العربي.
- الشيخ، س. ا. (2014). *سيكولوجية الفروق الفردية في الذكاء* (5 ed.). عمان: دار المسيرة للنشر والتوزيع.
- العنزي، ع. (2011). *الميول الميئية وعالقتها بالقيم الشخصية والتحصيل الدراسي لدى طلبة الصف الأول الثانوي في إدارة التربية والتعميم بمنطقة تبوك*. الأردن: رسالة ماجستير غير منشورة، جامعة مؤتة،
- القاسم، م. ب. (2010). *علم النفس المهني بين النظرية والتطبيق*. مؤسسة الوراق.

- أحمد تيغزة. (2008). نظرية الصدق الحديثة ومتضمناتها التطويرية لواقع القياس. ندوة علم النفس (الصفحات 1-35). الرياض: قسم علم النفس، جامعة الملك سعود.
- بشرى اسماعيل. (2004). المرجع في القياس النفسي. القاهرة: مكتبة الانجلو مصرية.
- صلاح الدين محمود علام. (2000). القياس والتقويم التربوي والنفسي (الإصدار 1). القاهرة: دار الفكر العربي.
- طارق عبد الرؤوف محمد عامر، وإيهاب عيسى المصري. (2018). الذكاء العاطفي والذكاء الاجتماعي. المجموعة العربية للنشر والتوزيع.
- عبد القادر كراجة. (1997). القياس والتقويم في عمم النفس (رؤية جديدة)، عمان: دار اليازوري العممية،
- فاطمة أحمد الجاسم. (2010). الذكاء الناجح والقدرات التحليلية الابداعية (الإصدار 1). الأردن: ديونو للطباعة والنشر والتوزيع.
- فيصل عباس. (2001). الاختبارات النفسية -تقنياتها وإجراءاتها- (الإصدار 1). دار الفكر العربي بيروت.
- محمد ربيع شحاتة. (2007). قياس الشخصية. عمان: دار المسيرة للنشر والتوزيع.
- محمد صلاح الدين أبو ناهية. (1994). القياس التربوي. القاهرة: مكتبة الأنجلو مصرية.
- محمد طه. (2006). الذكاء الانساني، اتجاهات معاصرة وقضايا نقدية. الكويت: عالم المعرفة.
- محمد عويضة كامل. (1996). القدرات العقلية في علم النفس. بيروت -لبنان: دار الكتب العلمية.