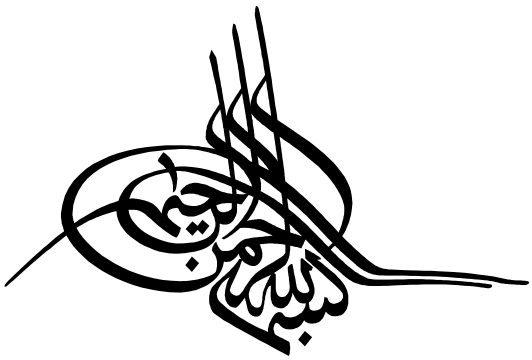

کاربرد آزمون‌های روانی در مشاوره

سری کتاب‌های کمک آموزشی کارشناسی ارشد

مجموعه روان‌شناسی ۲

مؤلفان: حسین کشاورز افشار، میلاد قربانی

سرشناسه : کشاورز افشار،حسین
عنوان : کاربرد آزمون‌های روانی در مشاوره
مشخصات نشر : تهران: مشاوران صعود ماهان، ۱۴۰۴
مشخصات ظاهری : ۳۱۷ ص
فروست : سری کتاب‌های کمک آموزشی کارشناسی ارشد
شابک : ۹-۰۲۹-۳۸۹-۶۰۰-۹۷۸
وضعیت فهرست نویسی : فیپای مختصر
یادداشت :این مدرک در آدرس http://opac.nlai.ir قابل دسترسی است.
شناسه افزوده : قربانی، میلاد
شماره کتابشناسی ملی : ۳۷۷۵۳۵۲

نام کتاب: ..... کاربرد آزمون‌های روانی در مشاوره
مولفان: ..... حسین کشاورز افشار، میلاد قربانی
مدیرتولید: ..... سمیه بیگی
ناشر: ..... مشاوران صعود ماهان
نوبت و تاریخ چاپ: ..... اول /۱۴۰۴
تیراژ: ..... ۱۰۰۰ نسخه

قیمت:
ISBN: ..... ۵/۱۹۰/۰۰۰ ریال
شابک: ۹-۰۲۹-۳۸۹-۶۰۰-۹۷۸

انتشـــارات مشاوران صعود ماهان: خیابان ولیعصر، بالاتر از تقـاطع مطهری،
روبروی قنادی هتل بزرگ تهـــران، جنب بانک ملی، پلاک ۲۰۵۰
تلفن: ۴- ۸۸۱۰۰۱۱۳

سخن ناشر

«ن والقلم و ما یسطرون»

کلمه نزد خدا بود و خدا آن را با قلم بر ما نازل کرد.

به پاس تشکر از چنین موهبت الهی، موسسه ماهان درصدد برآمده است تا در راستای انتقال دانش و مفاهیم با کمک اساتید مجرب و مجموعه کتب آموزشی خود برای شما داوطلبان ادامه تحصیل در مقطع کارشناسی ارشد گام موثری بردارد. امید است تلاش‌های خدمتگزاران شما در این موسسه پایه‌گذار گام‌های بلند فردای شما باشد.

مجموعه کتاب‌های کمک آموزشی ماهان به‌منظور استفاده داوطلبان کنکور کارشناسی ارشد سراسری و آزاد تالیف شده‌اند. در این کتاب‌ها سعی کرده‌ایم با بهره‌گیری از تجربه اساتید بزرگ و کتب معتبر داوطلبان را از مطالعه کتاب‌های متعدد در هر درس بی‌نیاز کنیم.

بدین وسیله از مجموعه اساتید، مولفان و همکاران محترم خانواده بزرگ ماهان که در تولید و به روزرسانی تالیفات ماهان نقش موثری داشته اند، صمیمانه تقدیر و تشکر می نماییم. دانشجویان عزیز و اساتید محترم می توانند هرگونه انتقاد و پیشنهاد درخصوص تالیفات ماهان را از طریق سایت ماهان به آدرس mahan.ac.ir با ما در میان بگذارند.

موسسه آموزش عالی آزاد ماهان

سخن مؤلف

انسان تنها موجودی است که «هستی» از او سؤال می‌پرسد و این روحیه پرسشگری اوست که از کودکی تا بودن در این جهان، همواره با تیغ ظریف سؤال‌هایش به جراحی باورها و عقاید خود، دنیا و آینده می‌پردازد لذا انسان و کنجکاوی‌هایش او را در جهت آزمودن واقعیت‌ها برای گذر از مراحل خاصی از زندگی سوق می‌دهد و هر زمان که با خویشتن‌نگری خویش به‌دنبال کسب آرزوها و اهدافش در تکاپوست، دست به دامن دوست دیرینه‌اش «کتاب» می‌برد.

با توجه به اهمیت روزافزون مشاوره در عصر کنونی، در سالیان اخیر شاهد اقبال زیاد دانشجویان به رشته مشاوره در ایران بوده‌ایم. اما با توجه به ظرفیت محدود پذیرش دانشگاه‌ها در این رشته متأسفانه داوطلبان زیادی پشت سد کنکور باقی می‌مانند. لذا حدّاکثر توان خود را به کار بسته‌ایم تا مناسب‌ترین منابع هر درس را جمع‌آوری و خلاصه آن را به‌صورت کتاب واحدی در اختیار داوطلبان قرار دهیم. کتاب حاضر خلاصه‌ای است از مطالب درس کاربرد آزمون‌های روانی در مشاوره که برای دانشجویان رشته راهنمایی‌ومشاوره جهت آمادگی در آزمون کارشناسی ارشد تدوین شده است. با توجه به منابع گسترده‌ای که این درس دارد، در تدوین این اثر سعی شده است که مطالب عمدتاً از منابع اصلی و مورد سوال جمع‌آوری گردد. با این وجود گردآورنده توصیه می‌کند که داوطلبان در کنار این کتاب سایر منابع در این زمینه را مطالعه کنند، زیرا در آزمون کارشناسی ارشد درصدهای جزیی تعیین‌کننده موفقیت هستند. کتاب حاضر در هشت فصل تدوین شده است که به ترتیب شامل فصل‌های تاریخچه آزمون‌های روانی، نظریه‌های هوش، اندازه‌گیری هوش، آزمون‌های هوش، آزمون‌های شخصیت،آزمون‌های فرافکن شخصیت، آزمون‌های تشخیصی و در نهایت سنجش رغبت‌ها، نگرش‌ها و ارزش‌ها می‌باشد. در پایان هر فصل به‌ترتیب نکات کلیدی مربوط به آن فصل، سوالات تالیفی، سوالات آزمون سراسری و در نهایت سوالات آزمون دانشگاه آزاد همراه با پاسخ تشریحی آورده شده است. در آزمون سراسری سال ۱۴۰۴ سوالاتی را شاهد بودهایم که در صورت مطالعهٔ این کتاب کاملا از قدرت پاسخگویی به آنها برخوردار خواهیم بود؛ برای نمونه:

سوال ۳ در فصل دو کتاب (بخش رویکردهای روان‌سنجی)، سوال ۷ در فصل چهار کتاب (بخش آزمون نابسته به فرهنگ کتل)، سوال ۸ و ۲۰ در فصل هشت کتاب (بخش نظریه هالند)، سوال ۱۰ در فصل یک کتاب (بخش ویژگی‌های آزمون‌ها)، سوال ۱۴ در فصل پنج کتاب (بخش مقیاس‌های بالینی MMPI)، سوال ۱۶ در فصل یک کتاب (بخش انواع نیمرخ‌ها)، سوال ۱۸ در فصل ۶ کتاب (بخش ویژگی‌های آزمون‌های فرافکن) آمده است.

بی‌شک مطالب ارایه شده خالی از نقص و اشکال نیست؛ بنابراین نظرات و انتقادات شما می‌تواند راهنمای ما در تدوین اثری مفیدتر و بهتر باشد. امید است مطالب ارایه شده برای شما خواننده گرامی مفید واقع شود.

فهرست

عنوان | صفحه

فصل اول: تاریخچه آزمون‌ها ۱۱
زمینه استفاده از آزمون‌های روانی ۱۴
پایایی ۱۷
روایی ۲۱
روایی محتوایی ۲۱
روایی ملاکی ۲۱
روایی سازه ۲۲
روایی در روش‌های بالینی ۲۳
ضریب دشواری ۲۵
ضریب تمییز ۲۵
ضرایب همبستگی ۲۵
اجرای آزمون‌ها ۲۵
انواع نرم‌ها ۲۷
امتیازات و محدودیت‌های ضریب هوشی ۲۹
هنجارهای سنی ۲۹
نقد آزمون‌ها ۳۱
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل اول ۳۷
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل اول ۴۰
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل اول ۵۸
فصل دوم: نظریه‌های هوش **۶۵**
ماهیت هوش ۶۶
چهار رویکرد مهم درباره هوش ۶۷
الگوی سلسله مراتبی هوش ۶۸
رویکردهای عصبی – زیستی ۶۹
نظریه‌های شناختی ۷۱
نظریه‌های معاصر هوش ۷۲
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل دوم ۷۷
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل دوم ۷۹
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل دوم ۸۷

فصل سوم: اندازه‌گیری هوش ۹۱
اندازه‌گیری هوش و انتقادهای واردشده ۹۲
آزمون بینه و سیمون ۹۳
مقیاس ۱۹۰۸ ..... ۹۳
آزمون ۱۹۱۱ ..... ۹۴
مقیاس جدید زازو ۹۴
آزمون استنفورد ـ بینه ۹۴
مفهوم سن‌تارک ۹۵
سقف آزمون ۹۵
آزمون هوش وکسلر ۹۵
خرده آزمون‌های وکسلر کودکان (WISC-R) ۹۶
خرده آزمون‌های وکسلر بزرگسالان (WAIS-R) ۹۶
خرده آزمون‌های وکسلر پیش‌دبستانی (WIPPSI-R) ۹۶
روایی پایایی ۹۷
اعتبار وکسلر بزرگسالان ۹۷
محاسن و محدودیت‌های آزمون وکسلر ۹۸
هوشبهر ۹۸
هوشبهر کلامی و عملی ۹۹
نمره‌های عامل ۹۹
تغییرپذیری خرده آزمون‌ها درون مقیاس‌ها ۹۹
تغییرپذیری درون خرده آزمون‌ها ۱۰۰
مقیاس‌های کلامی ۱۰۰
اطلاعات عمومی ۱۰۰
فراخنای ارقام ۱۰۰
گنجینه لغات ۱۰۱
محاسبه ۱۰۱
خرده آزمون درک و فهم ۱۰۱
خرده آزمون شباهت‌ها ۱۰۲
خرده آزمون‌های عملی ۱۰۲
انواع خرده آزمون‌ها (تکمیل تصاویر، تنظیم تصاویر، طراحی مکعب‌ها، الحاق قطعات، نماد ارقام، مازها) ۱۰۲
سنجش آسیب مغزی و خلاصه فرضیات مطرح شده درباره آسیب‌های مغزی ۱۰۴
ناتوانایی‌های یادگیری ۱۰۵
ناتوانایی هوش ۱۰۵
رویکرد ستز – ماجل و یودین ۱۰۶
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل سوم ۱۱۱
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل سوم ۱۱۴
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل سوم ۱۲۳

فصل چهارم: آزمون‌های هوش ۱۳۱
مازهای پرتئوس ۱۳۲
آزمون مکعب‌های کهس ۱۳۳
آزمون ماتریس‌های پیشرونده ریون ۱۳۳
مقیاس عملی لایتر ۱۳۴
عملکرد بورلی – اوله رون ۱۳۴
آزمون اندره – ری ۱۳۴
آزمون‌های شیرخوارگی ۱۳۵
مجموعه آزمون‌های چندسطحی ۱۳۶
آزمون‌های گروهی هوش ۱۳۷
آزمون‌های گروهی کلامی هوش ۱۳۸
آزمون‌های گروهی غیرکلامی هوش ۱۳۸
آزمون $\mathbf{B}_{۵۲}$ دکتر بورنادل ۱۳۹
آزمون آدمک گودیناف ۱۳۹
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل چهارم ۱۴۷
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل چهارم ۱۴۹
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل چهارم ۱۵۷
فصل پنجم: آزمون‌های شخصیت ۱۶۳
پرسش‌نامه اطلاعات شخصی وودورث ۱۶۵
پرسش‌نامه شخصیتی آیزنگ ۱۶۵
پرسش‌نامه شخصیتی چندجنبه‌ای برون رویتر ۱۶۶
پرسش‌نامه سنجش سازگاری بل (Bell) ۱۶۶
پرسش‌نامه رشد اجتماعی واینلند ۱۶۶
پرسش‌نامه علائم روانی (SCL ۹۰) ۱۶۷
پرسش‌نامه اضطراب کتل ۱۶۷
پرسش‌نامه افسردگی بک ۱۶۸
پرسش‌نامه ۱۶ عاملی شخصیت کتل ۱۶۸
پرسش‌نامه دانش‌آموزان دبیرستانی ۱۶۹
پرسش‌نامه شخصیتی چندوجهی مینه سوتا (MMPI) ۱۶۹
سه مقیاس روایی L، F، K ۱۶۹
مقیاس‌های بالینی MMPI ۱۷۰
مقیاس‌های فردی CPI ۱۷۳
محتوای آزمون ۱۸۰
مقیاس‌های هر شاخص ۱۸۰
پرسش‌نامه شخصیتی آیزنگ ۱۸۱
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل پنجم ۱۸۶
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل پنجم ۱۸۹
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل پنجم ۱۹۴

فصل ششم: آزمون‌های فرافکن شخصیت ۱۹۹
ویژگی‌های اصلی آزمون‌های فرافکن از نظر پیرپیشو ۲۰۰
ویژگی‌های آزمون‌های فرافکن ۲۰۰
انواع روش‌های فرافکن ۲۰۱
آزمون رورشاخ ۲۰۱
تفسیر ۲۰۳
آزمون اندریافت موضوع (TAT) ۲۰۴
آزمون CAT ۲۰۵
آزمون اندریافت موضوع کودکان و نوجوانان P.N ۲۰۵
آزمون سایموندز ۲۰۵
آزمون GAT ۲۰۶
TAT کنگو ۲۰۶
تفسیر نتایج ۲۰۷
آزمون تداعی کلمات یونگ ۲۰۷
آزمون فرافکنی تکمیلی ۲۰۸
آزمون جمله‌های ناتمام ۲۰۸
آزمون داستان‌های ناتمام ۲۰۹
آزمون‌های فرافکن ترسیمی ۲۰۹
آزمون فرافکنی آدمک مک‌اور (H.F.D) ۲۱۰
آزمون فرافکنی ترسیمی مینکووسکا ۲۱۰
آزمون فرافکنی خانواده ۲۱۰
آزمون فرافکنی گفت و گو در موقعیت‌های ناکام‌ساز ۲۱۲
آزمون فرافکنی ناکامی روز نزوایگ برای کودکان ۲۱۲
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل ششم ۲۱۵
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل ششم ۲۱۷
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل ششم ۲۲۲
فصل هفتم: آزمون‌های تشخیصی ۲۲۷
آزمون‌های هوش بندر گشتالت ۲۲۸
دستورالعمل‌های اجرا ۲۲۸
ملاک‌های کیفی تفسیر آزمون ۲۳۰
شناخت ویژگی‌های رفتاری ۲۳۱
ملاک تشخیص ضایعات مغزی بندر گشتالت ۲۳۲
ملاک‌های تشخیص روان نژندی ۲۳۲
ملاک‌های تشخیص مشکلات جنسی ۲۳۳
ملاک‌های تشخیص حالت‌های اضطرابی ۲۳۳
ملاک‌های تشخیص هیستری ۲۳۳
ملاک‌های تشخیص افراد وسواسی ۲۳۳

ملاک‌های تشخیص افسردگی واکنشی ۲۳۳
شاخص‌های بندر گشتالت سایکوپاتیک ۲۳۳
ملاک‌های تشخیص الکلیسم و اعتیاد ۲۳۳
ملاک‌های تشخیص سایکوسوماتیک ۲۳۴
ملاک‌های تشخیص روان‌پریشی ۲۳۴
ملاک‌های تشخیص مانیک و دپرسیو ۲۳۴
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل هفتم ۲۳۸
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل هفتم ۲۴۰
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل هفتم ۲۴۳
فصل هشتم: سنجش رغبت‌ها، نگرش‌ها و ارزش‌ها ۲۴۷
نظریه هالند ۲۴۹
نظریه رو ۲۵۰
نظریه سوپر ۲۵۰
پرسش‌نامه‌های رغبت‌سنج ۲۵۰
رغبت‌سنج شغلی هالند ۲۵۱
رغبت‌سنج شغلی استرانگ (SVIB) ۲۵۱
رغبت‌سنج استرانگ – کمپبل (SCII) ۲۵۱
رغبت‌سنج شغلی کودر (KOIS) ۲۵۲
رغبت‌سنج شغلی مینه سوتا (MVII) ۲۵۲
رغبت‌سنج کالج‌های آمریکا ۲۵۲
مقیاس‌های نگرش‌سنج ۲۵۳
مقیاس لیکرت ۲۵۳
مقیاس ترستون ۲۵۳
مقیاس گاتمن ۲۵۳
مقیاس افتراق معنایی ۲۵۳
مقیاس فاصله اجتماعی ۲۵۴
پرسش‌نامه مطالعه ارزش‌ها ۲۵۴
پرسش‌نامه ارزش‌های شغلی ۲۵۴
ارزش‌سنج راکیچ ۲۵۵
مقیاس قضاوت اخلاقی کلبرگ ۲۵۵
آزمون‌های خلاقیت ۲۵۵
سوالات چهارگزینه‌ای تالیفی و پاسخنامه فصل هشتم ۲۵۷
سوالات چهارگزینه‌ای سراسری و پاسخنامه فصل هشتم ۲۵۹
سوالات چهارگزینه‌ای آزاد و پاسخنامه فصل هشتم ۲۵۹
سوالات کنکور از سال ۹۵ تا۱۴۰۴ ..... ۲۶۰

فصل اول

تاریخچه آزمون‌ها

- خدمات گالتون
- زمینه استفاده از آزمون‌های روانی
- زمینه‌های کاربرد آزمون‌های روانی از نظر کرانباخ
- ویژگی آزمون‌ها
- روش‌های اصلی تعیین روایی
- ضرایب همبستگی
- انواع نرم‌ها
- امتیازات ضریب هوشی
- پراکندگی روانی

تاریخچه آزمون‌ها

آزمون: اندازه عینی و استاندارد شده از چگونگی یک رفتار است (تعریف آناستازی). عینی یعنی اینکه اطلاعات ذهنی ما در آزمون دخالت ندارد و براساس ملاک‌هایی می‌خواهیم اندازه‌هایی را به‌دست آوریم. مفهوم آزمون از مفهوم اندازه‌گیری محدودتر است و معمولاً نمونه‌ای از رفتار را اندازه می‌گیرد.

سنجش: زمانی‌که در اندازه‌گیری به‌جای آزمون یا علاوه بر آن از وسایل دیگری (مانند مشاهده، مصاحبه و...) استفاده می‌کنیم، در واقع سنجش نوعی تحلیل بالینی است و به‌منظور پیش‌بینی عملکرد به‌کار می‌رود.

اندازه‌گیری: اختصاص دادن اعداد به افراد و اشیا به شیوه منظم برای نشان دادن ویژگی‌ها است. به‌عبارت دیگر، اندازه‌گیری عبارت است از تخصیص کمیّات به کیفیات مطابق اصول و قواعد معین.

روان‌سنجی: مجموعه شیوه‌هایی که با استفاده از آن حالات روان را از کیفی به کمّی درمی‌آورند.

ارزشیابی: کرانباخ ارزشیابی را به‌عنوان «گردآوری و کاربرد اطلاعات به‌منظور اخذ تصمیم درباره برنامه آموزشی» و سی.ای.بی.بای آن را به‌عنوان «فرآیند نظام‌دار گردآوری و تفسیر شواهدی که میخر به داوری ارزشی منجر می‌شود و درنهایت به عمل می‌انجامد» تعریف کرده است. ارزشیابی حوزه بسیار گسترده‌ای دارد و علاوه‌بر جنبه‌های کمّی، جنبه‌های کیفی را نیز برآورد می‌کند. درباره تفاوت اندازه‌گیری و ارزشیابی می‌توان گفت در اندازه‌گیری هیچ‌گونه داوری درباره نتایج به‌دست آمده انجام نمی‌شود، اما در ارزشیابی داوری و قضاوت وجود دارد و برای تعیین میزان کارآمدی یک برنامه یا دوره آموزشی انجام می‌گیرد.

اندیشه اولیه جهت انفرادی‌کردن

اولین بار جمجمه‌شناسان و قیافه‌شناسان در اوایل قرن نوزدهم برای ارزیابی عینی هوش تلاش کردند.

ریشه آزمون‌ها به‌معنای امروزی آن با روان‌شناسی آزمایشی شروع شد و در واقع، اولین قدم‌ها برای اندازه‌گیری هوش در آزمایشگاه‌ها برداشته شد و اولین روان‌شناسان علمی فیزیولوژیست‌ها و فیزیکدان‌ها بودند.

اولین آزمایشگاه روان‌شناسی در سال ۱۸۷۹ توسط وونت در لایپزیک آلمان پایه‌گذاری شد.

اولین مراحل تشکیل آزمون‌ها به‌منظور اندازه‌گیری صحیح و ایجاد فنون مناسب برای تحلیل واکنش افراد و تهیه هنجارها، استفاده از ریاضیات برای تحلیل نتایج آزمون‌هاست.

روان‌شناسان اولیه به مسأله تفاوت‌های فردی توجهی نشان نمی‌دادند؛ به‌نظر آنها تفاوت آزمودنی‌ها نتیجه خطاهایی بود که در اندازه‌گیری پدیده‌های روانی اتفاق می‌افتاد. آنها معتقد بودند پدیده‌های روانی بسیار ناپایدارتر از پدیده‌های فیزیکی هستند (هدفشان تنها اندازه‌گیری بود).

درحالی‌که، روان‌شناسی در نیمه دوم قرن نوزدهم به مرور در دامن فیزیولوژی پرورده می‌شد، روان‌سنجی نیز از سال ۱۸۸۵ به این‌سو از مباحث مربوط به تفاوت‌های فردی به‌وجود آمد.

اولین تلاش‌ها برای اندازه‌گیری به مطالعه فرآیندی سطح پایین (مثل سرعت واکنش افراد، زمینه‌های حسی و حرکتی افراد، زمان واکنش و...) اختصاص یافت.

خدمات گالتون

۱- گالتون اعتقاد داشت تفاوت‌های فردی ذاتاً منافعی دارند. وی به توارث در زمینه صفات جسمی و روانی اعتقاد داشت. او در کتاب مشهورش به‌نام نبوغ ارثی سعی کرد نبوغ را از طریق انحرافی که افراد از میانگین گروه متعلق به خود پیدا می‌کنند، اندازه بگیرد.
۲- از دیگر خدمات گالتون منحنی اُجیو است. ۳- تاسیس آزمایشگاه مردم‌شناسی است که در آن هرکس می‌توانست از جنبه‌های مختلف بدنی مورد آزمایش قرار گیرد.۴- استفاده از منحنی طبیعی و استفاده از روش‌های آماری برای طبقه‌بندی و تفسیر اطلاعات است. ۵- استفاده از پرسش‌نامه و مقیاس‌های درجه‌بندی مفهوم همبستگی بین توانایی‌های ذهنی، ۶- همبستگی بین توانایی‌های ذهنی و تمییز حسی ۷- مفهوم تغییرپذیری.
اصطلاح آزمون روانی را اولین بار شاگرد وونت به‌نام جیمز مَکین کَتِل به‌کار برد. کتل این اصطلاح را در مقاله‌ای که با هدف سنجش تفاوت‌های فردی اشخاص نوشته بود، پیشنهاد داد.
خلاصه ماهیت آزمون‌های کتل: کتل روان‌سنجی را از چهار دیواری آزمایشگاه بیرون آورد تا به خدمت اجتماع درآورد.
مهم‌ترین موارد مورد سنجش در آزمون‌های کتل عبارتند از:
۱- اندازه‌گیری قدرت ماهیچه‌ای یا نیروسنجی، ۲- اندازه‌گیری سرعت حرکات، ۳- اندازه‌گیری فشار لازم برای ایجاد درد، ۴- اندازه‌گیری زمان واکنش، ۵- اندازه‌گیری آستانه اختلافی حس وزن، ۶- اندازه‌گیری سرعت تشخیص چهار رنگ مطلوب، ۷- اندازه‌گیری دقت در تقسیم یک خط ۵۰ سانتی‌متری، ۸- اندازه‌گیری حافظه کوتاه مدت، ۹- اندازه‌گیری آستانه حس لامسه ۱۰- اندازه‌گیری دقت در تولید یک فاصله زمانی ۱۰ ثانیه‌ای.
معایب آزمون‌ها: بینه و هانری از آزمون‌های کتل به‌علت اینکه توانایی‌های سطح پایین را می‌سنجند، انتقاد می‌کنند.
آزمون‌های کتل فاقد هنجار هستند و هدف نهایی آنها اندازه‌گیری توانایی ذهنی افراد نیست، بلکه این آزمون‌ها می‌خواهند ثبات فرآیندهای ذهنی یا تغییر آنها را برحسب موقعیت‌های زمانی تعیین کنند و روابط موجود بین کنش‌ها را به‌دست آورند.
شروع واقعی کار آزمون‌های روانی با بینه و هانری است، آنها مقاله‌ای را منتشر کردند که در این مقاله به آزمون‌هایی که سطوح پایین را می‌سنجند، انتقاد می‌کنند. به نظر آنها تفاوت افراد در زمینه واکنش‌های عالی ذهنی است و این آزمون‌ها از دقت و صراحت فوق‌العاده‌ای برخوردار هستند. اندازه‌گیری هوش ازطریق جلوه‌های هوش است.
در آزمون‌های هوش به‌دنبال جلوه‌های خارجی و رایج آن هستیم و مهم‌ترین هدف آزمون‌های بینه و سیمون نیز همین بوده است.
آزمون‌های هوشی برای اولین بار به‌منظور شناسایی کودکان کم‌توان ذهنی تهیه شدند.
آزمون بینه و سیمون اولین بار در سال ۱۹۰۵ شروع شد. بینه پدر آزمون‌های روانی است؛ زیرا مفهومی که او از اندازه‌گیری هوش ارائه می‌کرد، کاملاً تازه بود و برای حصول به اهداف کاملاً عملی به‌کار می‌رفت.
بینه معتقد بود هوش تا ۱۶ سالگی رشد می‌کند و بعد از ۱۶ سالگی هوش دیگر رشد نمی‌کند.
آزمون بینه و سیمون ۳۰ سوال داشت که به‌ترتیب سطح دشواری از آسان به مشکل تنظیم شده بودند که سه کارکرد شناختی یعنی قضاوت، درک و فهم و استدلال را اندازه‌گیری می‌کردند. در تجدیدنظر سال ۱۹۰۸ آزمون بینه و سیمون، مفهوم سن عقلی در آن مطرح شد. آزمون‌های بینه بیشتر درمورد کودکان ۳ تا ۱۳ سال کاربرد داشتند.
آزمون بینه و سیمون توسط تِرمَن تجدیدنظر شد که به‌نام استنفورد - بینه معروف شد. در این تجدیدنظر که در سال ۱۹۱۶ صورت گرفت، ترمن اصطلاح هوشبهر را برای اولین بار به‌کار برد. این آزمون در سال ۱۹۳۷ دوباره توسط ترمن و مریل تجدیدنظر شد و به‌صورت دو فرم موازی L و M ارائه گردید.
علت انتقاد از آزمون بینه: این آزمون فقط یک عدد به‌دست می‌دهد و سطح هوش را بدون توجه به چگونگی فعالیت ذهنی تعیین می‌کند.
آزمون تفکر مفهومی: فرد از پاسخ‌های غلط برای نحوه ساخت روان استفاده می‌کند.

آزمون تفکر مفهومی نحوه تکوین و به‌وجود آمدن مفاهیم را در ذهن نشان می‌دهد. این آزمون ساخت روان را در هر مرحله از رشد نشان می‌دهد. شکل‌گیری مفاهیم در ذهن یک نفر را تفکر مفهومی گویند (مثلاً مفهوم حجم در ذهن کودکان).
اسپیرمن با استفاده از ریاضیات به نظریه دو عاملی G و S دست پیدا کرد.
بونارد در فرانسه مطمئن‌ترین راه را برای راهنمایی و انتخاب شغل، روش تحلیل عاملی دانسته است.
تحلیل عاملی (Factor analysis) از تعدادی سوال تشکیل شده که با همدیگر همبستگی بالایی دارند.
تحلیل عاملی تنها روشی است که می‌تواند روان‌سنجی را از روش‌های غیرعلمی و سنتی برهاند.
آخرین مرحله تلاش‌های روان‌سنجی آزمون‌های شخصیت است.
هدف آزمون‌های هوش تعیین کارآیی است. در این آزمون‌ها به‌دنبال عناصر شناختی رفتار هستیم (جنبه‌های کمی).
در آزمون‌های شخصیت به‌دنبال عناصر غیرشناختی رفتار هستیم (جنبه‌های کیفی، ناپایدار رفتار و عاطفی).
شخصیت از عناصر متعددی تشکیل شده که این عناصر مستقل از هم هستند و می‌توان آنها را اندازه گرفت.
علت گسترش آزمون‌های شخصیت در قرن اخیر استقبال زیاد روان‌پزشکان و روان درمانگران است.

زمینه استفاده از آزمون‌های روانی

۱- زمینه تحصیلی: آزمون‌های تحصیلی قبل از آزمون‌های روانی به‌وجود آمدند، هرچند کندتر از آزمون‌های روانی گسترش پیدا کردند. زمینه‌های تحصیلی، اجرای اولیه آزمون‌ها و کاربرد مفید آنها در آموزش و پرورش است. (بینه و سیمون کودکان عقب‌مانده و عادی را جدا کردند و تحت آموزش خاص قرار دادند.)
بلوغ آموزشی یعنی تجلی توانایی‌های ذهنی لازم برای شروع بعضی از یادگیری‌ها به‌خصوص خواندن، نوشتن، حساب کردن و...
وظایف روان‌شناس تحصیلی پیدا کردن راه‌حل برای دشواری‌های عمومی و اختصاصی دانش‌آموزان است.
۲- زمینه‌های شغلی: کاربرد روان‌شناسی در زمینه سازمان‌دهی کار آنقدر متنوع شده که امروزه کاربرد عبارت سایکوتکنیک به زمینه شغلی اختصاص یافته است.
سایکوتکنیک: یعنی مجموعه آزمون‌هایی که با آنها استعدادهای افراد را اندازه‌گیری می‌کنند.
سایکوتکنسین: کسی که در اجرا و تعبیر و تفسیر آزمون‌ها تخصص دارد.
سازمان‌دهی شغلی: شامل راهنمایی، انتخاب، انطباق و آموزش شغلی است.
تایلوریسم: کاربرد روان‌شناسی برای مطالعه اثر عوامل انسانی در افزایش تولید کارخانه.
۳- زمینه نظامی: از جنگ جهانی اول - آزمون و تربیت جاسوسان و فرماندهان ارتش. آزمون‌های ارتش که در جریان جنگ جهانی اول مورد استفاده قرار گرفت، نقطه شروع آزمون‌های گروهی بود. رابرت یرکس آزمون‌های آلفا و بتا را ابداع کرد.
۴-کودکان نابهنجار و بزهکار: به‌منظور کمک رساندن به کودکان و نوجوانانی که دشواری‌هایی دارند، تاسیس شده است.
۵- آزمون‌ها و کلینیک‌های پزشکی: استفاده از آزمون‌های روانی به‌خصوص آزمون‌های شخصیت توسط روان‌پزشکان در سی سال اخیر رو به افزایش است.
کلینیک مُسَلَح: در آن از انواع آزمون‌ها استفاده می‌شود تا تشخیص بهتر و سریع‌تری انجام پذیرد.
سایر زمینه‌های کاربرد آزمون‌ها: بررسی جنایی و قضایی، پلیس قضایی، دروغ‌سنج‌ها، مسابقات ورزشی، تبلیغات.
همیشه اندیشه پی‌بردن به تفاوت‌های فردی، انفرادی‌کردن کارها و نتایج عملی حاصل از آنها پایه کاربرد آزمون‌ها بوده است.

جامعه‌سنجی (سوسیومتری) مورنو ۱۹۳۴

- مورنو پدر جامعه سنجی (سوسیودرام) است.
- شاخص موقعیت اجتماعی یا پذیرش اجتماعی شاخصی است که با شمردن تعداد دفعاتی که یک نفر انتخاب می‌شود، به‌دست می‌آید.
- نمایش تصاویر نتایج حاصل از اجرای آزمون را جامعه‌نما (سوسیوگرام) گویند.

• شاخص گیرندگی اجتماعی نشان می‌دهد که یک فرد معین از طرف چند نفر مورد پذیرش قرار گرفته است.
• شاخص گسترش عاطفی تعداد کسانی را نشان می‌دهد که توسط یک نفر معین انتخاب شده‌اند.

زمینه‌های کاربرد آزمون‌های روانی از نظر کرانباخ

کرانباخ در کتاب خود چهار زمینه را برای آزمون‌ها ذکر کرده است:
۱- پیش‌بینی: آزمون‌ها برای اندازه‌گیری ویژگی‌ها و توانایی‌های فردی‌اند که پایه و اساس تصمیم‌گیری‌اند.
۲- گزینش: مثل آزمون استخدامی.
۳- طبقه‌بندی: برای این به‌کار می‌رود که فرد را در چه گروهی قرار دهیم و براساس چه روشی او را مورد آموزش یا درمان قرار دهیم.
۴- ارزشیابی: برنامه‌ها و روش‌ها و تدابیر درمانی مورد قضاوت واقع می‌شوند.

اهداف اصلی آزمون‌ها

۱- حل مسائل نظری یا پیدا کردن راه‌حل منطقی برای دشواری‌های موجود (تحقیقات بنیادی)
۲- حل مشکل موردنظر (تحقیقات کاربردی)
آزمون‌ها را می‌توان برحسب ملاک‌های متفاوتی طبقه‌بندی کرد:

۱- فرم و شکل و مشخصات خارجی:
۱-۱- آزمون‌های مدادی - کاغذی (کتبی)
۱-۲- آزمون‌های عملکردی (دستی)

۲- محتوای ظاهری:
۲-۱- کلامی (از زبان استفاده می‌شود).
۲-۲- غیرکلامی (وابسته به فرهنگ نیستند).

۳- اجرا:
۳-۱- فردی ← مانند: آزمون هوشی استنفرد - بینه، مقیاس هوشی وکسلر، آزمون رورشاخ و TAT
۳-۲- گروهی ← مانند: آزمون هوشی ریون، آزمون استعداد تحصیلی SAT و پرسش‌نامه شخصیتی MMPI
برای بررسی عمیق افراد و تشخیص عقب‌ماندگی‌ها (بررسی همه جانبه فرد) از آزمون‌های فردی استفاده می‌کنیم.

محاسن آزمون‌های فردی:
۱- دسترسی به اطلاعات کیفی
۲- اطمینان از درک دستور العمل
۳- درمورد خردسالان قابل اجراست.
۴- درمورد افراد کم‌توان ذهنی و بیماران روانی قابل اجراست.
۵- در صورت ناتوانی آزمودنی، همه سوال‌ها از او پرسیده نمی‌شود.
نکته: سن پایه به سنی گفته می‌شود که افراد می‌توانند به تمام سوالات آن سن پاسخ دهند.

معایب آزمون‌های فردی:
۱- نیازمند تخصص است.
۲- یکسان نبودن شرایط
۳- به زمان زیاد برای اجرا نیاز دارد.
۴- هزینه زیاد

محاسن آزمون‌های گروهی:

۱- شرایط یکسان برای همه

۲- صرفه‌جویی در هزینه

۳- صرفه‌جویی در زمان

۴- تخصص زیاد نیاز ندارد.

۵- نمره‌گذاری‌اش سریع و آسان و عینی است.

معایب آزمون‌های گروهی:

۱- امکان دسترسی به اطلاعات کیفی وجود ندارد.

۲- درمورد خردسالان قابل اجرا نیست.

۲-۱- به‌خاطر عدم درک توضیحات آزمونگر.

۲-۲- مشکلات مربوط به جمع‌آوری آن‌ها در یک محل خاص برای اجرای آزمون.

۳- بر روی بیماران روانی و عقب‌مانده‌ها قابل اجرا نیست.

۴- امکان دسترسی به آزمودنی‌ها پس از اجرای آزمون وجود ندارد.

۴- زمان اجرا:

۴-۱- محدود: مکعب‌های کهس - موزائیک ژیل - وکسلر - آزمون هوشی کتل و ریون

۴-۲- نامحدود: مازهای پرتئوس - بندرگشتالت

۵- از نظر موضوع:

۵-۱- توانایی: توانایی‌های نظری و ذاتی را می‌سنجند، مثل وکسلر - مکعب‌های کهس - ریون

۵-۲- مهارت و پیشرفت: معلومات و مهارت‌های کسب شده را می‌سنجند. مثل آزمون مهارت ماشین نویسی.

۵-۳- شخصیت: ویژگی‌های شخصیتی را می‌سنجند، مثل: نئو - MMPI - مازها، عزت نفس کوپر اسمیت.

۵-۴- رغبت: مانند جان هالند - استرانگ - میشل گوکلن

۶- دقت:

۶-۱- معلم ساخته: برای تحقق اهداف آموزشی است.

۶-۲- تراز شده (میزان شده): آزمون‌هایی که هنجار و فرم دارند، مانند کنکور.

۷- شناخت:

۷-۱- شناختی: آزمون‌هایی که فعالیت‌های فکری و ذهنی آزمودنی را اندازه‌گیری می‌کنند، مانند آزمون‌های هوش و استعداد.

۷-۲- غیرشناختی: آزمون‌هایی‌که ویژگی‌های شخصیتی و عاطفی آزمودنی را اندازه‌گیری می‌کنند؛ آزمون‌های شخصیت، نگرش‌سنج‌ها و رغبت‌سنج‌ها نمونه‌هایی از این نوع آزمون‌ها هستند.

منظور از آزمون غیرشناختی آزمونی است که آزمودنی‌ها با استفاده از قسمت‌های هیپوتالاموس به آنها پاسخ می‌دهند. در آزمون‌های شناختی فرد با استفاده از کورتکس مغز به آنها پاسخ می‌دهد، مثل سخن گفتن، حل کردن، دلیل آوردن، حساب کردن، اما هیپوتالاموس مانند خشم، نفرت و... است.

آزمون‌های شناختی: در این نوع آزمون فعالیت‌های فکری و ذهنی آزمودنی‌ها دخالت دارند، اما در آزمون‌های غیرشناختی ویژگی‌های شخصیتی و عاطفی فرد دخالت دارد.

اهداف:

پیش‌بینی: آزمون‌های هوش، کنکور

تشخیص: به‌منظور شناسایی ویژگی‌های هوشی به‌کار می‌رود.

از نظر مفروضه‌ها:

۱- آزمون‌های فرافکنی (ترکیبی) مفروضه‌ها:

۱-۱- شخصیت پدیده‌ای پویاست.

۱-۲- شخصیت به‌صورت یک کل درنظر گرفته می‌شود.

۲- آزمون‌های تحلیلی (عینی) مفروضه‌ها:

۲-۱- شخصیت از عناصر متعددی تشکیل شده است.

۲-۲- این عناصر از یکدیگر نسبتاً مستقل هستند.

ویژگی‌های آزمون‌ها

پایایی، روایی، ضریب تمییز، ضریب دشواری، اجرای آزمون‌ها، نقد آزمون‌ها، بیان نتایج آزمون‌ها.

پـایـایی

اشاره به درجه ثبات نمرات یک آزمون دارد و در واقع قابلیت پیش‌بینی و دقت یک آزمون است. پایایی یعنی برآورد توانایی‌های واقعی فرد (حدس و شانس و... را پایین می‌آورد.) هدف پایایی برآورد تغییرات ناشی از خطاست.

$$x = T + e$$

داده آزمون = نمره واقعی + خطا

ما، در پایایی به دنبال این هستیم که مقدار خطا کم شود. هرچقدر خطا کمتر باشد، پایایی بیشتر می‌شود.

خطای معیار اندازه‌گیری: پایایی ۱ باشد، خطای معیار اندازه‌گیری صفر است.

$$\text{خطای معیار اندازه‌گیری} = S_t\sqrt{1-r_{11}}$$

زیربنای مفهوم پایایی خطای اندازه‌گیری است.

پایایی شرط لازم روایی است. اگر یک آزمون پایا باشد حتماً رواست، اما ممکن است روا باشد، اما پایا نباشد.

به‌صورت آرمانی متخصصان بالینی درمورد آزمون‌هایی که برای تصمیم‌گیری درباره افراد به‌کار می‌روند، انتظار پایایی ۹۰٪ یا بالاتر را می‌دهند. برای مقاصد بالینی پایایی ۹۰٪ و برای مقاصد پژوهشی انتظار ۷۰٪ را داریم.

روش‌های برآورد پایایی

روش‌های برآورد پایایی چهار مورد است:

۱- بازآزمایی: (آزمونی را ۲ بار اجرا می‌کنیم و ضریب همبستگی آنها را به‌دست می‌آوریم.)

فاصله اجرای دو بار آزمون را باید به‌گونه‌ای تعیین کنیم که نخست، اثر اجرای بار اول بر روی آزمودنی تا حد قابل‌قبولی از بین رفته باشد. دوم آنکه، تغییرات جسمی و تحولات روانی مربوط به رشد وارد صحنه نشوند. البته رعایت این دو مورد از آزمونی به آزمون دیگر و در سنین مختلف فرق می‌کند؛ برای مثال در آزمون‌های مربوط به استدلال و ابتکار فاصله زمانی باید بیشتر از آزمون‌های دیگر باشد؛ زیرا اثر یادگیری مدت زمان زیادی در ذهن می‌ماند. همچنین در آزمون‌های مربوط به کودکان اگر فاصله زمانی زیاد باشد، تحولات مربوط به رشد اثرگذار خواهند بود، درصورتی‌که در بزرگسالان فاصله زمانی بین چند هفته تا چند ماه نیز مجاز است. روش بازآزمایی فقط درمورد آزمون‌هایی مناسب است که علاوه‌بر شرایط یکسان اجرا، تکرار آزمون اثر قابل توجهی در نتایج آن نداشته باشد.

قاعده کلی: به‌طور کلی بازآزمایی درمورد اندازه‌گیری متغیرهایی مفید است که از ثبات نسبی برخوردارند، مثل هوش.

۲- فرم‌های همتا: زیربنای منطقی فرم‌های همتا این است که اگر صفت یک شخص با استفاده از فرم‌های همتای یک آزمون چندین بار اندازه‌گیری شود نتایج حاصل از این اندازه‌گیری‌ها باید یکسان باشد.

فرم‌های همتا از نظر شکل ظاهری متفاوتند، اما میانگین، واریانس، ضریب تشخیص و ضریب دشواری یکی می‌باشد.

نخستین مشکل فرم‌های همتا این است که آیا این دو فرم واقعاً با یکدیگر معادلند یا نه؟
دومین مشکل موجود در این زمینه این است که بین دو اجرای آزمون زیاد وقفه نباشد، این وقفه ممکن است موجب شود عملکرد آزمودنی‌ها در دو آزمون متفاوت شود.

استفاده از روش فرم‌های همتا یا هم‌ارز درمواردی‌که امکان تهیه سوال‌های زیاد و هم‌ارز وجود دارد (مانند آزمون‌های هوش و پیشرفت تحصیلی) امکان‌پذیر است، اما در آزمون‌هایی مثل آزمون شخصیت که تهیه سوال‌های کاملاً هم‌ارز در آن دشوار و حتی غیر ممکن است، امکان استفاده از این روش وجود ندارد. امتیاز فرم‌های همتا به روش بازآزمایی این است که مشکل یکسانی محتوا در فرم‌های هم ارز حل شده است.

۳- روش‌های همسانی درونی: در دو روش بازآزمایی و فرم‌های هم‌ارز، به ثبات زمانی نمره‌ها توجه می‌شود؛ درصورتی‌که در روش همسانی درونی، ثبات زمانی نمره‌ها مطرح نیست؛ زیرا آزمون فقط یک‌بار اجرا می‌شود. در روش همسانی درونی، نوعی همسانی در محتوای سوال‌ها مطرح است. در روش همسانی درونی آزمون یک بار اجرا می‌شود و سپس با یکی از روش‌های زیر به تعیین ضریب پایایی می‌پردازند:

الف- روش دو نیمه‌کردن: در این روش آزمون یک‌بار بر روی گروه نمونه اجرا می‌شود و سپس آزمون به دو نیمه تقسیم می‌شود. بهترین روش برای تقسیم آزمون به دو قسمت این است که سوال‌های زوج را در یک بخش و سوال‌های فرد را در بخش دیگر قرار بدهیم. در این روش به شرط آنکه ماده‌های آزمون به‌ترتیب دشواری تنظیم شده باشند، نمره‌های دو بخش تقریباً هم‌ارز خواهند بود.
از آنجا که آزمون در این شیوه سنجش به دو نیمه تقسیم می‌شود، ضریب همبستگی حاصله، ضریب اعتبار نیمی از آزمون تلقی می‌گردد. برای محاسبه ضریب اعتبار کل آزمون فرمول پیشنهادی اسپیرمن - براون به‌شرح زیر به‌کار گرفته می‌شود:
در این فرمول علائم استفاده شده عبارتند از:

$$r_{۱۱} = \frac{۲r_{\frac{۱}{۲}\frac{۱}{۲}}}{۱ + r_{\frac{۱}{۲}\frac{۱}{۲}}}$$

$r_{۱۱}$ = ضریب اعتبار کل آزمون

$r_{\frac{۱}{۲}\frac{۱}{۲}}$: ضریب همبستگی بین دو نیمه آزمون

در این روش ضریب اعتبار کل آزمون همیشه بیشتر از ضریب همبستگی بین دو نیمه آزمون به‌دست می‌آید.
شرط به‌کارگیری ضریب اعتبار اسپیرمن - براون یکسان بودن واریانس نمره‌های دو نیمه آزمون است.
درصورتی‌که واریانس نمره‌های دو نیمه آزمون یکسان نباشد، از ضریب آلفای کرانباخ استفاده می‌کنیم.
فرمول آلفای کرانباخ به شرح زیر است:

$$\alpha = \frac{۲\left[s_t^۲ - \left(s_x^۲ + s_y^۲\right)\right]}{s_t^۲}$$

عناصر فرمول عبارتند از:

α = ضریب پایایی آزمون از راه دو نیمه کردن

$s_t^۲$ = واریانس نمره‌های کل آزمون

$s_x^۲$ = واریانس نمره‌های یکی از دو نیمه آزمون

$s_y^۲$ = واریانس نمره‌های نیمه دیگر آزمون

باید توجه داشت که ضریب آلفای کرانباخ حد پایین ضریب اعتبار آزمون را برآورد می‌کند و احتمال دارد که ضریب اعتبار واقعی آزمون از آن بیشتر باشد.

نکته: روش دو نیمه کردن بهترین روش برای تعیین پایایی صفتی است که نوسان‌های زیادی دارد.

نکته: روش دو نیمه‌سازی برای آزمون‌هایی کارایی دارد که سوالات آنها از ماده‌های همگن درست شده‌اند.

ب- روش کودر - ریچاردسون: این روش زمانی به‌کار می‌رود که:

۱- سوالات باهم همگن نباشند؛ یعنی سوالات از گروه‌های مختلف باشند (تشریحی، تستی، جورکردنی، کامل کردنی و...)

۲- نمره‌گذاری سوال متفاوت باشد.

۳- روش کودر - ریچاردسون هم مثل روش دونیمه سازی یک‌بار اجرا می‌شود.

نکته: کودر - ریچاردسون تجانس درونی سوالات و یا همسالانی درونی را می‌سنجد.

نکته: ضریب همبستگی کودر - ریچاردسون بین ۰ و ۱ است.

نکته: یکی از مقاصد کودر - ریچاردسون صرف وقت کمتر برای محاسبه است.

نکته: در کودر - ریچاردسون نمره‌گذاری سوالات به‌صورت صفر و یک می‌باشد.

فرمول کودر – ریچاردسون ۲۰:

عناصر فرمول:

$$KR_{20} = r_{11} = \frac{n}{n-1}\left(\frac{s^2 - \sum pq}{S^2}\right)$$

r_{11}: ضریب پایایی آزمون

n : تعداد سوالات آزمون

s^2: واریانس نمره‌های خام آزمون

p: نسبت دانش‌آموزانی که به سوال پاسخ درست داده‌اند.

$q(1-p)$: نسبت دانش‌آموزانی که به سوال پاسخ غلط داده‌اند.

Σpq : مجموع واریانس‌های تمام سوال‌ها.

هرچه تعداد سوالات بالا باشد، ضریب پایایی بالاتر خواهد بود (ارتباط مستقیم با تعداد سوالات دارد.)

شرط استفاده از کودر ریچاردسون۲۱:

۱- نمره‌گذاری به‌صورت ۰ و ۱

۲- یکسان بودن درجه دشواری سوالات

فرمول کودر - ریچاردسون۲۱:

عناصر فرمول:

$$KR_{21} = r_{11} = \frac{n}{n-1}\left(1 - \frac{\bar{x}(n-\bar{x})}{n \times S^2}\right)$$

r_{11}: ضریب پایایی آزمون

n: تعداد سوال‌های آزمون

$\bar{x}$: میانگین نمره‌های گروه

s^2: واریانس نمره‌های خام آزمون

ج- روش ضریب آلفا: از روش ضریب آلفا زمانی استفاده می‌شود که آزمودنی نظر خود را بر روی یک پیوستار مشخص می‌کند و ما نمی‌توانیم سوالات را به‌صورت صفر و یک نمره‌گذاری کنیم. (مثلاً: کاملاً مخالفم، مخالفم، نظری ندارم، موافقم، کاملاً موافقم)

ضریب آلفا برای آزمون‌های شخصیت و نگرش‌سنج‌ها استفاده می‌شود:

$$\alpha = \left(\frac{n}{n-1}\right)\frac{S_t^2 - \Sigma S_i^2}{S_t^2}$$

عناصر فرمول:

s_t^2: واریانس کل آزمون

s_i^2: واریانس سوال i

n: تعداد سوال‌ها یا اظهارنظرها (در مقیاس نگرش‌سنج)

اگر آزمونی دارای بخش‌ها یا خرده آزمون‌هایی باشد و بخواهیم از خرده آزمون‌ها برای محاسبه پایایی کل آزمون بهره بگیریم، باید فرمول زیر را به‌کار ببریم:

$$r_\alpha = \frac{j}{j-1}\left(1-\frac{\Sigma\sigma_j^2}{\sigma_t^2}\right)$$

r_α: ضریب پایایی کل آزمون

j: تعداد بخش‌ها یا خرده آزمون‌ها

σ_j^2: واریانس نمره‌های هر بخش آزمون

σ^2: واریانس نمره‌های کل آزمون

۴- روش قضاوت نمره‌گذاران: مقایسه قضاوت نمره‌گذاران برای بررسی اعتبار آزمون‌هایی به‌کار برده می‌شود که ملاک‌های عینی همه پسندی برای ارزشیابی پاسخ به آنها ابداع نشده است. به‌طور مثال در آزمون‌های خلاقیت، انشایی پیشرفت تحصیلی، برخی از آزمون‌های فرافکن یا عملی حرفه‌ای، آزمونگران در مواردی که آزمون دستورالعمل صحیح را پیش‌بینی نکرده است خود را مجاز به ابتکار یا خلق ضوابط فردی می‌کنند. در چنین مواردی بهترین شیوه بررسی اعتبار آزمون مراجعه به قضاوت خبرگانی است که در کاربرد آن سابقه‌ی طولانی دارند.

پایایی آزمون‌های سرعت: در آزمون‌سازی و تفسیر نمره‌های آزمون، باید بین اندازه‌گیری آزمون‌های سرعت و قدرت (توانایی) تمایز قایل شویم. آزمون قدرت به آزمونی گفته می‌شود که ماده‌های آن به‌تدریج دشوارتر می‌شوند و حتی ممکن است ماده‌هایی وجود داشته باشد که هیچ‌یک از آزمودنی‌ها نتوانند به آن پاسخ درست بدهند، برعکس در آزمون‌های سرعت توانایی پاسخ‌دهی مطرح نیست، بلکه تفاوت‌های فردی ناشی از سرعت کارکرد مطرح است. برای پایایی آزمون‌های سرعت از ضرایب پایایی حاصل از یک‌بار اجرا از قبیل روش‌های فرد و زوج - کودر ریچاردسون و آلفای کرونباخ نمی‌توان استفاده کرد؛ زیرا در این نوع آزمون‌ها، آزمودنی‌ها توانایی پاسخگویی به کلیه سوالات را دارند، اما وقت کافی برای پاسخ‌گویی را ندارند؛ بنابراین اگر آزمون را به دو نیمه اول و دوم تقسیم کنیم و این دو نیمه را با هم همبسته کنیم ضریب پایایی محاسبه شده خیلی ضعیف خواهد بود.

نکته: بهترین روش برای محاسبه ضریب پایایی آزمون‌های سرعت روش بازآزمایی است.

البته با توجه به اینکه امکان تاثیر اجرای اول بر روی اجرای دوم وجود دارد، بهتر است دو فرم موازی از آزمون سرعت تهیه شود که میانگین و انحراف استاندارد آنها برابر باشند و این دو فرم بلافاصله بعد از یکدیگر روی آزمودنی‌های اجرا شود.

عوامل موثر در پایایی:

۱- افزایش تعداد سوالات

۲- افزودن تعداد سوال با ضریب تمییز بالا

۳- افزایش تعدادی سوال با ضریب دشواری متوسط

۴- انتخاب سوالات همگون (متجانس)

۵- انتخاب گروه ناهمگون

خطای معیار اندازه‌گیری: اگر آزمونی را به دفعات روی فردی اجرا کنیم، انحراف استاندارد نمراتی که از این آزمون‌ها به‌دست می‌آید، خطای معیار اندازه‌گیری نام دارد. منظور از خطای معیار اندازه‌گیری، انحراف معیار توزیع نمرات خطا در اندازه‌گیری‌های مکرر است. خطای معیار اندازه‌گیری را با استفاده از ضریب پایایی می‌توان محاسبه کرد:

در این فرمول:

$$SEM = s_t\sqrt{1-r_{11}}$$

SEM: خطای معیار اندازه‌گیری برحسب نمره‌های خام

s_t: انحراف معیار نمره‌های خام

r_{11}: ضریب پایایی آزمون

فرمول بالا نشانگر آن است که همیشه رابطه ضریب اعتبار و مقدار خطای معیار اندازه‌گیری معکوس خواهد بود. وقتی ضریب پایایی آزمون برابر یک باشد، خطای معیار اندازه‌گیری برابر صفر خواهد بود.

روایی

در روایی آزمون بررسی می‌شود که آیا آزمون همان چیزی را که باید بسنجد، می‌سنجد. بررسی و اثبات روایی آزمون بسیار دشوار است، به‌دلیل آنکه متغیرهای روانی مفاهیم انتزاعی‌اند، مثل اضطراب، هوش، خلاقیت. اینها واقعیت ملموس ندارند؛ بنابراین باید اینها را استنباط کنیم. مهم‌ترین عاملی که در ارزشیابی تست باید بدان توجه شود روایی آزمون است و مقصود از روایی آزمون مناسب بودن، با معنا بودن و مفید بودن استنباط‌های خاصی است که از روی نمره‌های آزمون به‌عمل می‌آید.

روش‌های اصلی تعیین روایی:

۱- روایی وابسته به محتوا

۲- روایی وابسته به ملاک

۳- روایی وابسته به سازه

۱ – روایی محتوایی

در نخستین مرحله ساختن آزمون باید روایی وابسته به محتوا درنظر گرفته شود، این امر به محتوای موضوعی ماده‌های آزمون برمی‌گردد و با انتخاب مناسب ماده‌های آزمون تحقق می‌یابد. تأکید زیاد بر قضاوت شخصی سازندگان آزمون عامل کلیدی روایی محتوا می‌باشد و به‌شدت ذهنی است. روایی محتوایی در اندازه‌گیری‌های پیشرفت تحصیلی از اهمیت خاصی برخوردار است.

نکته: در نخستین مرحله انتخاب ماده‌ها سازندگان آزمون باید زمینه مهارت‌ها و دانش مربوط به متغیری را که می‌خواهند اندازه بگیرند، به دقت مورد بررسی قرار دهند و انتخاب ماده‌ها براساس این تعاریف و مفاهیم تهیه می‌شوند.

روش تعیین روایی محتوایی

در این روش سوال‌های آزمون در اختیار متخصصان یا برخی از آزمودنی‌ها گذاشته می‌شوند و از آنها می‌خواهند که مشخص کنند آیا سوال‌های آزمون صفت موردنظر را اندازه‌گیری می‌کند یا خیر. درصورتی‌که بین افراد مختلف در زمینه روایی آزمون توافق وجود داشته باشد، آن آزمون دارای روایی محتوایی است. روایی محتوایی دو نوع است: روایی صوری و روایی منطقی.

روایی صوری: به قضاوت آزمایش‌شوندگان بستگی دارد. مسأله اصلی در روایی صوری درک آزمودنی از مفهوم آزمون است.

در واقع روایی صوری، جالب و جذاب بودن آزمون است. یک آزمون باید برانگیزاننده و رغبت‌انگیز باشد. (صفحه‌آرایی، نحوه چیدن سوالات و...)

روایی منطقی (نمونه‌گیری): شکل پیچیده‌تری از روایی صوری است. در این روش صفت مورد اندازه‌گیری به‌طور دقیق تعریف و اجزای تشکیل‌دهنده آن مشخص می‌شود؛ سپس برای هریک از اجزای سوال‌های منطقی طرح می‌کنند، به‌طوری‌که کلیه بخش‌های حیطه مورد اندازه‌گیری را بپوشاند.

تعیین روایی محتوایی اولین قدم در بررسی تمام آزمون‌هاست. استفاده از روایی محتوایی لازم، اما کافی نیست و بهتر است که غیر از روایی محتوایی، روایی ملاکی یا سازه نیز روی آزمون انجام شود.

۲- روایی ملاکی

روایی ملاکی از راه مقایسه نمره آزمون‌ها با گونه‌ای از عملکرد در یک مقیاس یا معیار خارجی تعیین می‌شود.

مثال: قبولی در کنکور و فارغ‌التحصیل شدن با معدل بالا در دانشگاه، کسانی‌که نمره‌های خوبی داشته‌اند در کنکور، نمره‌های خوبی هم در دانشگاه داشته‌اند.

این معیار خارجی باید با متغیری که قرار است به‌وسیله آزمون مورد اندازه‌گیری قرار گیرد، رابطه نظریه‌ای داشته باشد. بنابراین در روایی ملاکی از نمرات یک آزمون برای پیش‌بینی عملکرد افراد در آزمون دوم استفاده می‌شود. آزمون دوم را که عملکرد فرد در آن پیش‌بینی می‌شود، ملاک می‌گویند. روایی ملاکی به‌صورت ضریب همبستگی بیان می‌شود.

روایی وابسته به ملاک را می‌توان به روایی هم‌زمان و روایی پیش‌بینی تقسیم کرد.

روایی هم‌زمان به اندازه‌گیری‌هایی گفته می‌شود که هم‌زمان با اجرای آزمون یا تقریباً هم‌زمان با اجرای آزمون انجام می‌گیرد. برای نمونه، یک آزمون هوشی ممکن است هم‌زمان با سنجش پیشرفت تحصیلی برروی گروهی از افراد اجرا شود.

روایی پیش‌بین

به اندازه‌گیری‌هایی که مدتی پس از اجرای نمره‌های آزمون به‌دست می‌آید، روایی پیش‌بین می‌گویند.

کنکور ۸۸- ناخالصی ملاک (سرایت معیار) هنگامی رخ می‌دهد که آگاهی از نتایج آزمون قبلی در پیشرفت بعدی یا درجه‌بندی آن تأثیر می‌گذارد. برای نمونه، دانش‌آموزی که در ریاضیات قوی است معلم نمره ورزش بالایی به او می‌دهد.

از روایی پیش‌بین به‌دلیل اینکه محاسبه آن مدت‌ها پس از اجرای آزمون و طی تجربیات خاص ممکن می‌گردد، با عنوان روایی تجربی نیز نام برده‌اند.

۳- روایی سازه

پیچیده‌ترین و قوی‌ترین رویکرد آزمون‌سازی، روایی سازه است.

سازه برچسبی است که ما بر توانایی‌ها و ویژگی‌های انسان‌ها می‌زنیم مثل هوش، اضطراب (فرد مضطرب کسی است که تپش قلب دارد، بی‌خواب است، قادر به تمرکز نیست و...)

رویکرد اساسی روایی سازه سنجش این مساله است که آزمون تا چه حد سازه نظری یا صفتی را اندازه می‌گیرد. این سنجش سه مرحله دارد: ۱- تجزیه و تحلیل صفت مورد اندازه‌گیری، ۲- چگونگی ارتباط صفات با متغیرهای دیگر، ۳- معلوم کردن اینکه آیا این روابط فرضی واقعاً وجود دارند یا نه.

برای تعیین روایی سازه از روش‌های زیر می‌توان استفاده کرد:

۱- تغییرهای تحولی یا تمایز سنی: به‌عنوان مثال، اگر در یک آزمون هوش، میانگین نمرات کودکان شش ساله بیشتر از نمرات پنج ساله‌ها و کمتر از هفت ساله‌ها باشد، آزمون مذکور روایی سازه خواهد داشت؛ زیرا سازه هوش با افزایش سن رشد می‌کند.

۲- تفاوت‌های گروهی: روش دیگر تعیین روایی سازه محاسبه تفاوت معناداری بین نمرات دو گروه در آزمون می‌باشد که معمولاً به‌صورت فرضیه از طرف پژوهشگر تدوین می‌شود.

۳- همبستگی با سایر آزمون‌ها (روایی همگرا و واگرا): راه دیگر تعیین روایی سازه محاسبه همبستگی بین نمره‌های آزمون موردنظر با آزمون دیگری است که دارای روایی است و برای آن سازه ساخته شده است. در روایی همگرا محقق سعی می‌کند ثابت کند که آزمون جدید از نظر سنجش سازه، همسو و هماهنگ با آزمون‌های معتبری است که همان سازه را مورد سنجش قرار می‌دهند. در روایی افتراقی محقق تلاش می‌کند نشان دهد که آزمون نخست، غیر از سازه‌ای که برای آن ساخته شده سازه دیگری را نمی‌سنجد؛ دوم آنکه، وجوهی از همان سازه را مورد سنجش قرار می‌دهد که آزمون‌های دیگر به آن نپرداخته‌اند.

۴- همسانی درونی: در بررسی همسانی درونی نمره کل آزمون ملاک قضاوت قرار می‌گیرد. برای بررسی همسانی درونی به سه طریق می‌توان عمل کرد:

۱) محاسبه همبستگی هریک از سوالات آزمون یا هریک از خرده آزمون‌ها با نمره کل آزمون؛ چنانچه همبستگی هر سوال یا خرده آزمون با کل آزمون معنادار نبود، باید آن را از مجموعه آزمون حذف کرد.

۲) در شیوه دوم نمره‌های آزمودنی‌ها را از بالاترین تا پایین‌ترین نمره ردیف می‌کنند و نمره‌های ۲۷ درصد بالا و پایین را به‌عنوان گروه قوی و ضعیف انتخاب می‌کنند؛ سپس عملکرد گروه قوی و ضعیف را در هر سوال با هم مقایسه می‌کنند. سوال‌ها باید به‌گونه‌ای باشند که نسبت پاسخ‌دهی گروه قوی به‌طور معناداری بیش از گروه ضعیف باشد.
۳) در شیوه سوم، پس از محاسبه همبستگی سوال‌های آزمون با یکدیگر، میانگین ضرایب همبستگی بین سوال‌ها را محاسبه کرده و با داشتن این میانگین و تعداد سوال‌ها، همسانی درونی آزمون را محاسبه می‌کنند.
۵- روایی عاملی: روایی عاملی شکلی از روایی سازه است که با استفاده از تحلیل عاملی انجام می‌شود. تحلیل عاملی تعداد و ماهیت متغیرهایی را که یک آزمون اندازه می‌گیرد، مشخص می‌کند و برای ایجاد سهولت، متغیرها را به چند عامل کاهش می‌دهد. سپس این عوامل با توجه به آنچه به‌وسیله سوال‌های آزمون اندازه‌گیری می‌شوند، نام‌گذاری می‌شوند. پس از شناسایی عوامل از آنها برای توصیف ترکیب عاملی آزمون استفاده می‌کنند. هر آزمون را می‌توان برحسب عامل‌های تشکیل‌دهنده آن آزمون و همچنین ضریب یا وزن عاملی توصیف کرد. همبستگی آزمون با هریک از عوامل را روایی عاملی آزمون می‌نامند.

روایی در روش‌های بالینی

روایی در روش‌های بالینی بستگی به این دارد که تا چه اندازه آزمون‌ها با یکدیگر همخوانی داشته و به دقت یکدیگر می‌افزایند. بعضی از آزمون‌ها روایی افزایشی را نشان می‌دهند، زیرا با افزایش تعداد منابع داده‌هایی که به‌کار بسته می‌شوند دقت آنها افزایش می‌یابد؛ بنابراین روایی به توانایی آزمون در فراهم کردن اطلاعاتی فراتر از آنچه که قبلاً شناخته شده، اطلاق می‌شود. با افزودن تعداد منابعی که از داده‌ها استخراج می‌شود دقت آزمون را بالا می‌بریم.
روایی مفهومی: توانایی متخصص بالینی در ساختن فرضیه‌ها، آزمودن فرضیه‌ها و ترکیب داده‌های حاصل از آزمون فرضیه‌ها برای ترسیم تصویری یکپارچه و جامع از آزمودنی است.
ملونی و وارد (۱۹۷۶) این رویکرد را رویکرد روایی مفهومی می‌نامند، زیرا مستلزم توصیف یکپارچه و مفهومی از شخص است.
برخلاف روایی سازه که از سازه‌های از پیش تعیین شده تشکیل شده است، روایی مفهومی سازه‌ها را به‌عنوان نتایج نهایی به‌وجود می‌آورد.
هدف روایی مفهومی این است که برای این سازه‌ها منابع اطلاعاتی روا فراهم کند. برای تعیین روایی مفهومی کار با افرادی آغاز می‌شود که هیچ سازه‌ای برایشان تنظیم نشده است.
مرحله دوم شامل مشاهده، جمع‌آوری اطلاعات و ساختن فرضیه است. اگر این فرضیه‌ها به‌وسیله مشاهده رفتار، شرح حال و یا سایر منابع داده‌ها تعیین شوند در این صورت فرضیه‌های مذکور به‌عنوان سازه‌های روا درمورد شخص تلقی می‌شوند.

عوامل موثر درونی بر روایی آزمون

۱- کیفیت سوالات
۲- ترتیب قرار گرفتن سوالات (همه سوال‌های سخت یا آسان نباید به دنبال هم باشند).
۳- تعداد سوال‌ها (با افزایش تعداد سوالات پایایی و روایی افزایش می‌یابد).

عوامل موثر بیرونی در روایی آزمون

۱- اجرای نامناسب (مثل دستورالعمل غلط، کمبود زمان، مکان نامناسب)
۲- مشکلات موجود در تصحیح پاسخ‌های آزمون‌شوندگان و نمره‌گذاری آنها.
۳- ویژگی‌های روانی آزمون شوندگان (مثل اضطراب بالا، علاقه‌مندی یا بی‌علاقگی فرد، سطح انگیزش بالا و پایین).
۴- ترکیب آزمودنی‌ها از لحاظ توانایی مورد اندازه‌گیری (هرچه از این نظر نامتجانس‌تر باشند، روایی آزمون بیشتر است).

اثر عوامل ذهنی

عواملی که دقت نتایج آزمون‌ها را تحت‌تأثیر قرار می‌دهند:
۱- ستایش یا نکوهش (تشویق و تنبیه): حتی اگر پنهانی باشد نتایج آزمون را به‌طور چشمگیری تحت‌تأثیر قرار می‌دهد.
۲- نگرش و رفتار آزماینده، بر انگیزه پاسخگویی به آزمون اثر دارد.

۳- اثر عوامل عاطفی: اگر آزمودنی از نظر عاطفی ناسازگار باشد، به خانواده سطح پایین تعلق داشته باشد و یا خردسال باشد و... اثر عوامل عاطفی بیشتر خواهد بود.
۴- اثر تمرین یا سازگاری ناآگاهانه با سطح کاری که آزمون ارائه می‌شود (اثر عادت و تمرین)؛ بنابراین کسانی‌که با آزمون آشنایی دارند بی‌شک برکسی که با آن بیگانه است، برتری دارد.

درجه‌بندی آزمون‌ها

درجه‌بندی آزمون به‌معنی تعیین درجات موفقیت واحدهای رفتار به‌منظور مقایسه عینی نتایج افراد مختلف است. این واحدها و این هنجارها امکان مقایسه افراد را فراهم می‌آورند و برحسب ماهیت متغیر مورد اندازه‌گیری متفاوت خواهند بود. به‌طور مثال، آزمون‌های رشد ذهنی برحسب سن درجه‌بندی می‌شوند. مثلاً پیش‌رس‌ها، بهنجارها و عقب‌مانده‌ها.

برخی از مولفان از ظرافت، حساسیت و قدرت تشخیص آزمون‌ها حرف می‌زنند که همگی اینها به درجه‌بندی آزمون‌ها مربوط می‌شوند. هر اندازه یک آزمون امکان اندازه‌گیری تفاوت‌های بسیار کوچک را بیشتر فراهم آورد، به همان اندازه حساسیت بیشتری خواهد داشت.

هرچقدر، در مقیاس کنش‌های روانی بالاتر می‌رویم (تفکر، استدلال و...) دشواری تهیه وسایل دقیق و حساس افزایش می‌یابد؛ زیرا هرچند تفاوت‌های افراد در اعمال عالی ذهنی بسیار آشکار است، این تفاوت‌ها به‌طور نامحسوس، از کمیّت محض به کمیّت آمیخته با کیفیت تبدیل می‌شوند (مثل خلاقیت، نوآوری، تجزیه و تحلیل و...).

استاندارد بودن

پس از تهیه یک آزمون برای اینکه مقایسه افراد با یکدیگر منطقی باشد، باید در اجرا و نمره‌گذاری سوالات روش یکسانی را به کار برد.
یکی از موارد استاندارد بودن آزمون‌ها یکسانی در اجرا و یکسانی در نمره‌گذاری است.
بسیاری از انتقادهای ناروا علیه آزمون‌ها بر اثر نتایج گمراه‌کننده‌ای است که به‌خاطر استاندارد نبودن آزمون (یا بی‌توجه به استاندارد بودن آزمون‌ها) و اجرای نامناسب، به‌وجود آمده است.
نمره‌گذاری نتایج باید طبق قواعد یکسان انجام گیرد. برای تفسیر و طبقه‌بندی پاسخ‌های آزمودنی، فقط یک راه وجود دارد، راهی که سازندگان آزمون آن را پیش‌بینی کرده‌اند و از آن استفاده کرده‌اند.
شخصیت آزمایشگر مهم‌ترین عاملی است که می‌تواند ایجاد یک فضای اطمینان بخش را تضمین کند.
قاعده: برخورد آزماینده در اجرای آزمون بزرگسالان باید خشک و جدّی و رسمی و مقرراتی باشد، اما در اجرای آزمون از خردسالان باید همراه با عطوفت و مهربانی باشد.

ویژگی‌های فرعی آزمون‌ها

۱- **مساله اقتصادی:** منظور از اقتصادی بودن آزمون، با صرفه بودن از نظر هزینه‌های تهیه آزمون یا اجرا و تصحیح نتایج آن است.
۲- **سرعت اجرا:** در بسیاری از موارد باید بتوان یک آزمون را بدون آنکه به صحت و دقت آن لطمه بخورد، به‌طور سریع اجرا کرد. آزمون‌های گروهی از نظر صرفه‌جویی در زمان بهترین نمونه هستند.
۳- **سادگی اجرا:** برای اجرای بعضی از آزمون‌ها داشتن تخصص خاصی لازم نیست. این نوع آزمون‌ها بیشتر در زمینه تحصیل‌اند.
۴- **رغبت:** رغبت جز ویژگی‌های ذاتی آزمون‌هاست و هر آزمونی باید رغبت آزمودنی‌ها را ترغیب کند.
نکته: در خردسالان برای تحریک رغبت می‌توان از بازی استفاده کرد.

اگر آزمودنی‌ها بزرگسال باشند نباید آزمون‌هایی به‌کار برد که محتوای آنها آشکارا از زندگی واقعی آنها دور باشد. می‌توان رغبت‌برانگیز بودن را با عنوان روایی صوری آزمون نیز مطالعه کرد.

ضریب دشواری

ضریب دشواری عبارت است از نسبت آزمودنی‌هایی که به سوال پاسخ درست داده‌اند.

$$p = \frac{p_u + p_l}{2}$$

$$\Rightarrow p = \frac{\text{نسبت افراد گروه پایین که به سوال جواب درست داده‌اند} + \text{نسبت افراد گروه بالا که به سوال جواب درست داده‌اند.}}{2}$$

ضریب تمییز

یعنی توانایی آزمون در جدا کردن افراد قوی از ضعیف.

$$D = \frac{Nu - Nl}{\%27N}$$

$$\Rightarrow D = \frac{\text{انتخاب‌های درست گروه ضعیف} - \text{انتخاب‌های درست گروه قوی}}{\%27N}$$

نکته: ضریب تمییز سوال برآوردی از ضریب همبستگی سوال با نمره کل آزمون است.

نکته: هرچقدر نمره کل آزمون با سوال همبستگی بالاتری داشته باشد ضریب تمییز آن نیز بالاتر است.

نکته: زمانی‌که سوال را با نمره کل آزمون همبسته می‌کنیم، ضریب تمییز مساوی پایایی یا همان پایایی می‌باشد.

ضرایب همبستگی

زمانی‌که یک متغیر پیوسته و دیگری دو ارزشی باشد از ضریب همبستگی دو رشته‌ای نقطه‌ای استفاده می‌کنیم.

زمانی‌که دو متغیر رتبه‌ای باشند از اسپیرمن استفاده می‌کنیم.

زمانی‌که دو متغیر پیوسته باشند از پیرسون استفاده می‌کنیم.

وقتی همبستگی بین دو سوال را که به‌صورت صحیح و غلط است، بررسی می‌کنیم از ضریب همبستگی فی (φ) استفاده می‌کنیم:

$$\varphi = \frac{p_u - p_l}{2\sqrt{pq}}$$

نمرات T بین ۲۰ و ۸۰ می‌باشد.

$$T = 10z + 50$$

هر ۱۵ انحراف معیار یک نمره Z است.

$$Z = \frac{x - \bar{x}}{S}$$

نمره ۹ بخشی $= 2z + 5$

نمره تافل $= 100Z + 500$

انحراف معیار آزمون وکسلر $15z$

خرده مقیاس‌های وکسلر $3Z + 10$

هوش آزمایش وکسلر $15Z + 100$

اجرای آزمون‌ها

برای اجرای آزمون‌ها، آمارگیری‌های متعددی را لازم داریم و حتماً باید متخصصان آزمون‌ها را تفسیر کنند؛ در واقع مسئولیت تعبیر و تفسیر آزمون‌ها برعهده متخصصانی است که در آزمون‌ها تبحر دارند. داشتن صلاحیت و آمادگی شرایط لازم برای اجرای آزمون‌هاست. موارد و عواملی که در اجرای آزمون لازم است:

۱) **زمان اجرای آزمون:** باید در زمانی کاملاً مساعد انجام گیرد، زیرا هدف این است که آزمودنی بتواند از حداکثر توانایی‌های خویش استفاده کند تا بهترین عملکرد خویش را نشان دهد.

نکته: آزمودنی مفهوم آزمون را بداند و بخواهد از تمام توانایی‌های خویش استفاده کند.

۲) **مدت اجرای آزمون:** هر اندازه آزمودنی جوان‌تر (کم سن‌تر) باشد باید بیشتر مراقب بود تا وی دچار یکنواختی و کسالت نشود. چون تمرکز آنها پایین و رغبت آنها متنوع‌تر است.

نکته: خطرناک‌ترین کار درمورد کودکان خردسال این است که با اصرار از آنها خواسته شود تا کار خود را به پایان برسانند.

درباره مناسب‌ترین مدت برای اجرای آزمون نمی‌توان قاعده کلی وضع کرد؛ زیرا این امر با سن آزمودنی، شخصیت او، نگرش وی نسبت به آزمون و نیز با شیوه اجرای آزمون ارتباط پیدا می‌کند. حواس‌پرتی، افت انگیزش، افزایش پاسخ‌های «نمی‌دانم»، تحرک زیاد یا عصبانیت عواملی هستند که می‌توانند کافی بودن مدت اجرای آزمون را نشان دهند.

۳) **محل اجرای آزمون:** درمورد کودکان خردسال بهتر است آزمون در محل شناخته شده انجام بگیرد و اشیای زیادی که احتمال می‌رود دقت آزمودنی را مختل می‌کنند وجود نداشته باشد.

۴) **نگرش آزماینده:** مهم‌ترین عامل در اجرای آزمون‌هاست که نگرش آزماینده چه مثبت چه منفی اثرگذار است.

آزماینده باید رفتار خود را با آزمودنی سازگار کند، توصیه می‌شود رفتار با افراد بزرگسال خشک، عینی و رسمی و درمورد کودکان مضطرب بهتر است مدت زیادی صرف ایجاد رابطه و آزمایش شود تا بدین‌ترتیب کودک با محیط خو بگیرد و خاصیت آزمایش که ایجاد اضطراب است از بین برود.

نکته: بهتر است به آزمودنی‌هایی که بی‌اعتماد هستند گفته شود که این آزمون علیه آنها استفاده نخواهد شد و آزمون به زمینه مورد علاقه آنها مربوط می‌شود. وقتی آزمون محدودیت زمانی دارد می‌توان به آزمودنی اجازه داد تا مسائل دشوار را جا بیندازد. بهتر است گاهی به آزمودنی‌ها فهمانده شود که مرتکب خطا می‌شوند. در هر صورت آزمودنی باید بفهمد که دشواری سوال‌ها به‌تدریج افزایش خواهد یافت، تا جایی که نتواند به هیچ سوالی پاسخ دهد.

باید از نصیحت کردن آزمودنی پرهیز کرد. در جریان اجرای آزمون هرگز نباید آموزش داد. به همین دلیل وقتی آزمودنی به دشواری برمی‌خورد نباید به او کمک کرد. در برخی موارد، به‌ویژه وقتی آزمودنی به خاطر شکست خود واقعاً رنج می‌برد، می‌توان به وی کمک کرد، اما نباید پاسخ را در اختیار وی گذاشت. همچنین نباید در پایان اجرا معنای توضیحات و شیوه تصحیح پاسخ‌ها را در اختیار آزمودنی گذاشت. درمورد خردسالان بهتر است که آزمون به‌صورت بازی شروع شود. در سنین کمتر از شش سال، اگر آزماینده برای کودک آشنا باشد نتایج خوبی به‌دست نخواهد آمد.

بهتر است کودکان خردسال در جلسات مختلف تست شوند، زیرا عکس‌العمل آنها در روزهای مختلف فرق خواهد کرد.

۵) **آزمودنی‌ها:** بهتر است آزمون با سوالات جالب شروع شود تا آزمودنی برای پاسخگویی به خود سوالات کار بکند و موقعیت آزمون را فراموش کند (فراموش کردن آزمون و غرق شدن در آزمون). مهم این است که از حداکثر توانایی‌های خودش استفاده کند.

۶) **دستورالعمل‌ها:** هر آزمونی دستورالعمل‌هایی را دارد که چگونگی ارائه سوالات را مشخص می‌سازد.

تنها عاملی که باید در اجرای آزمون رعایت شود: پیروی بی‌قیدوشرط از دستورالعمل آزمون‌هاست یا رعایت مواردی که سازندگان آزمون آن را پیشنهاد دادند.

۷) **گزارش نتایج آزمون‌ها:** در تهیه گزارش باید به حریم زندگی خصوصی افراد توجه شود. گزارش نتایج آزمون‌های روانی فقط در اختیار کسانی قرار می‌گیرد که در حل مشکل آزمودنی و بهبود زندگی وی نقش دارند. معمولاً نتایج آزمون‌ها به افراد زیر گزارش می‌شود:

۱- متخصصان: در تهیه گزارش برای متخصصان، مانند روان‌شناسان و مشاوران، گزارش به‌صورت فنی و مختصر تهیه می‌شود و به‌طور معمول به‌صورت نیمرخ یا تفسیر کلی نتایج آزمون ارائه می‌شود.

۲- والدین: گزارش مخصوص والدین باید بدون استفاده از اصطلاحات فنی و با بیانی روشن و قابل فهم نوشته شود و بهتر است سطح عملکرد آزمودنی در رابطه با خصایص مورد اندازه‌گیری توضیح داده شود. بدون اجازه آزمودنی و والدین او نتایج آزمون نباید در اختیار فرد دیگری قرار بگیرد.

۳- آزمودنی: نتایج آزمون باید با بیانی روشن و ساده در یک جلسه مشاوره‌ای در اختیار آزمودنی گذاشته شود. برای اینکه آزمودنی واکنش هیجانی شدیدی از خود نشان ندهد، بهتر است آزمونگر ابتدا جنبه‌های مثبت نتایج آزمون را بیان کند.

بیان نتایج آزمون‌ها:

کسانی‌که از روان‌سنجی حرف می‌زنند، فرض می‌کنند که برای مقایسه افراد با یکدیگر یا مقایسه یک فرد با خودش در زمان‌های مختلف واحدهای معناداری وجود دارد. برای ترکیب نتایج به‌دست آمده، از واحدها و نرم‌ها که برحسب ماهیت متغیر مورد مطالعه و برحسب هدف آزمایش متفاوت خواهند بود استفاده می‌کنند.

نتایج یک آزمون زمانی معنادار خواهد بود که با یک هنجار یا نرم مقایسه شود.

هنجار و نرم:

به متوسط یا میانگین حالات، معلومات یا خصوصیات افرادی اطلاق می‌شود که در آزمون شرکت کرده‌اند و حالات، معلومات یا خصایص افراد دیگر را می‌توان با آن سنجید تا مشخص شود فرد در مقایسه با افراد مشابه با خود در چه وضعیتی قرار دارد. به عبارت ساده‌تر، میانگین نمرات گروه مرجع را نرم می‌گویند.

در استاندارد کردن آزمون‌های هنجار - مرجع ← تهیه نرم یکی از مراحل اصلی هنجاریابی است. بر مبنای مقایسه افراد باهم مثل رقابت درکنکور؛ اما در آزمون‌های ملاک - مرجع ← مثال نمره ۱۲ معیار قبولی است، براساس ملاک و معیار قبلی افراد را با آن مقایسه می‌کند.

انواع نرم‌ها

۱- نرم‌های سنی: متوسط ویژگی‌های افرادی که باهم، هم سن هستند؛ بنابراین برای هر ویژگی جسمانی (قد، وزن و...) و هر ویژگی روانی که با سن تغییر می‌کند، می‌توانیم نرم تهیه کنیم؛ به‌طور مثال: متوسط وزن و قد نوزادان متولد شده ۳/۱۰۰ وزن، ۵۰ قد.

نکته: استفاده از نرم‌های سنی تنها در سنین رشد مناسب است و بعد از آن چندان مناسب نیست.

معایب نرم‌های سنی: میزان رشد در طول یک‌سال در تمام سنین از همسانی برخوردار نیست؛ مثال: میزان رشد بین ۳ تا ۴ سالگی خیلی شدید و سریع است و با میزان رشد در سنین ۱۳ تا ۱۴ سالگی تفاوت دارد.

عیب دیگر این نرم این است که پیشرفت یا تفاوت با حد متوسط هر سن برابر با پیشرفت یا تفاوت با حد متوسط سن دیگر نیست و این موضوع بسته به سنی است که تفاوت در آن وجود دارد.

۲- نرم‌های کلاسی: متوسط ویژگی‌های افرادی که با هم در یک کلاس تحصیل می‌کنند. نرم‌های کلاسی را به‌صورت دو رقمی نشان می‌دهند که رقم اول نشان‌دهنده پایه تحصیلی و رقم دوم نشان‌دهنده ماه است و از آنجاکه سال تحصیلی ۹ ماه است، این شماره از ۰ تا ۹ در تغییر است.

نرم کلاسی بیشتر برای تحصیل پیشرفت تحصیلی در کلاس‌های ابتدایی مناسب است و همان معایب نرم سنی را دارد.

۳- نرم‌های سنی هنجاری: متوسط افرادی که هم پایه تحصیلی هم‌سن آنها برابر باشد.

۴- نرم سن عقلی: فاصله طی شده رشد هوشی از کودکی تا بزرگسالی را سن عقلی گویند که اولین بار توسط آلفرد بینه مطرح شد. محاسبه سن عقلی شبیه به محاسبه نرم‌های سنی است؛ بنابراین اگر نمره خام یک دانش‌آموز در یک آزمون هوش برابر میانگین نمرات خام کودکان ۸ ساله باشد، سن عقلی او ۸ ساله خواهد بود.

سن عقلی: یعنی هوش کلی که برحسب سال بیان می‌شود. در دوره‌های رشد که سرعت رشد بالاست به‌جای سن عقلی باید از (واحدهای زمانی مثل: ماه، هفته) ماه عقلی استفاده کنیم.

مثال: ماشینی که از تهران به قزوین رفته مسافت: سن عقلی، ضریب هوشی: سرعت حرکت کامیون، سرعت بیشتر مساوی مسافت بیشتر.

۵- بهره هوشی: یکی از روش‌های قدیمی تهیه نرم استفاده از بهره هوشی است که برای اولین بار در سال ۱۹۱۲ توسط اشترن مطرح شد:

$$\text{بهره هوشی یا ضریب هوشی} = \frac{MA}{CA} = \frac{\text{سن تعلیمی}}{\text{سن تقویمی}} \times ۱۰۰$$

هرچقدر افراد بتوانند فاصله بیشتری را در زمان کمتر طی کنند، بهره‌هوشی آنها بیشتر است (باهوش‌ترند). ضریب هوشی از ثبات نسبی برخوردار است و برای پیش‌بینی می‌توان از آن استفاده کرد. البته قبل از ۶ سالگی ضریب هوشی ثبات چندانی ندارد.

۶- نرم‌های درصدی: در نرم‌های درصدی نمره خام هر فرد را نمره درصدی و درصد افرادی را که زیر آن نمره قرار دارند، رتبه درصدی آن فرد گویند.

عناصر فرمول زیر عبارتند از:

PR : رتبه درصدی نمره‌ای که می‌خواهیم محاسبه کنیم.

CF : فراوانی تراکمی نمره پایین‌تر از نمره موردنظر

Fi : فراوانی مطلق طبقه‌ای که قصد محاسبه رتبه درصدی آن را داریم.

N : تعداد کل نمره‌ها در توزیع فراوانی

$$PR = \frac{CF_b + \frac{F_i}{۲}}{N} \times ۱۰۰$$

$$CEB = ۱۰۰۰Z + ۵۰۰۰$$

نکته: نتایج یک آزمون زمانی معنادار خواهد بود که با یک هنجار یا نرم مقایسه شود. روان‌سنجی حتماً باید هنجار و نرم داشته باشد. نرم‌های درصدی در راهنمایی شغلی و کنکور سراسری که در آنها آگاهی از موقعیت فرد نسبت به کلیه شرکت‌کنندگان، در ویژگی موردنظر ضروری است، مورد استفاده قرار می‌گیرد. همچنین از نرم‌های درصدی برای مقایسه نتایج آزمون‌های مختلف استفاده می‌شود.

نکته: بهتر است افراد نسبت به میانگین گروهی که به آن تعلق دارند، مقایسه شوند.

نکته: در ارزیابی تعداد زیادی از ویژگی‌های انسان باید عامل زمان و تحول را به حساب آورد. ممکن است یک عملکرد معین برای یک پیرمرد ۷۰ ساله بهنجار، ولی برای یک فرد جوان نابهنجار باشد.

نکته: تنها مرجع زمانی‌که به‌طور مطلق ثابت می‌ماند سن تقویمی است که به آن سن واقعی یا شناسنامه‌ای نیز گفته می‌شود. سن تقویمی یعنی مدت زمانی که فرد زندگی کرده است و از روز تولد به حساب می‌آید. هر سال زندگی مقداری رشد به همراه دارد که می‌توان آن را - که به تدریج کاهش می‌یابد - تعیین کرد.

نکته: ارزش کار بینه این بود که می‌خواست هوش کودک را برحسب سال‌های زندگی‌اش اندازه‌گیری کند.

اولین نتیجه ارزشیابی هوش برحسب سن این است که عقب‌مانده‌ها را به‌طور عینی طبقه‌بندی می‌کنیم. این طبقه‌بندی براساس نتایج آزمون‌ها انجام می‌گیرد و امتیاز آن در این است که با وجود قراردادی بودن همه، آن را به یک شکل می‌فهمند (یعنی همه آن را درک می‌کنند) مثلاً به نظر بینه کودک نابهنجار تحصیلی کودکی است که تا ۹ سالگی ۲ سال عقب‌ماندگی داشته باشد و بعد از ۹ سالگی ۳ سال عقب‌ماندگی ذهنی را نشان دهد. به فردی که به پایان رشد ذهنی رسیده؛ یعنی طبق نظر بینه ۱۶ ساله شده است، اما هوش کودک ۹ ساله را دارد، آموزش‌پذیر می‌گویند.

فردی که ۱۶ ساله است، اما سن عقلی ۶-۵/۵ ساله را دارد، کالیو (تربیت‌پذیر) ۱۶ ساله سن عقلی ۳-۲/۵ سال دارد ← کانا (حمایتی).

امتیازات ضریب هوشی

۱- مقایسه دو کودک

۲- پیش‌بینی درباره افراد

۳- ثبات ضریب هوشی نسبی است، بدین معنا که همیشه تغییرات و خطاهای جزیی وجود دارد. امکان دارد ضریب هوشی کودک بار اول ۱۱۰ باشد و بار دوم ۹۵ یا ۱۱۵ باشد.

محدودیت‌های ضریب هوشی

در ثبات ضریب هوشی محدودیتی وجود دارد، ضرایبی که پیش از تحصیلات رسمی به‌دست می‌آیند از ثبات بسیار ضعیفی برخوردار هستند (قبل از سن ۶ یا تحصیلات رسمی اندازه‌گیری‌ها از ثبات ضعیفی برخوردارند) برعکس در سن ۱۲-۷ سالگی از ثبات بسیار بالایی برخوردارند (از نوسان کمتری برخوردارند).

آزمون‌های سنین پایین فقط درجه پختگی یا رشد اعضا را نشان می‌دهند.

نسبی بودن ثبات ضریب هوشی این را نشان می‌دهد که ابزارهای اندازه‌گیری آزمون‌های سن عقلی کامل نیستند.

ارزیابی هوش بزرگسالان

۱- آزمون‌هایی را باید انتخاب کرد که محتوای متفاوت داشته باشد.

۲- باید برای هر سن آزمون‌ها را درجه‌بندی کنیم و میانگین‌های سوالی به‌دست بیاوریم.

پس‌روی بزرگسالان در آزمون‌های عملی هوش بیشتر است تا کلامی. پس برای بزرگسالان باید آزمون‌های خاص خودشان را به‌وجود آورد. نتایج را به‌صورت سن عقلی بیان نکرد؛ زیرا سن تقویمی افزایش می‌یابد، اما سن عقلی ثابت یا کاهش می‌یابد؛ بلکه باید به‌صورت نمرات استاندارد یا Z آنها را بیان کرد.

نقد ضریب هوش:

۱- معنای آن برحسب سن تغییر می‌کند.

۲- ضریب هوش برحسب آزمونی که برای اندازه‌گیری آن به‌کار می‌رود، فرق دارد.

۳- زازو ضریب هوش را مورد انتقاد قرار داد، او به‌جای اینکه هوش را تنها با یک عدد نشان دهد آن‌را در یک محدوده قرار می‌دهد.

هنجارهای سنی

سن جسمی: سن جسمی همان رشد جسمی است؛ به عبارت دیگر رشد قد و وزن که برحسب سال بیان می‌شود.

سن آناتومی یا رشد نظام استخوانی: که برحسب سال بیان می‌شود.

سن حرکتی: رشد تحریک‌پذیری را بیان می‌کند (چه حرکتی را می‌تواند با دستش انجام دهد؛ مثلاً می‌تواند توپی را شوت کند یا می‌تواند اسکیت سواری کند و...)

سن فیزیولوژیکی: رشد اندام‌ها به‌خصوص درجه بلوغ جنسی که برحسب سال بیان می‌شود.

سن تعلیمی: سطح آموزش که برحسب سال تقویمی بیان می‌شود. برای نمونه، یک دانش‌آموز زمانی سن تعلیمی ۷ سال را خواهد داشت که بتواند در پایان سال اول ابتدایی آنچه را که بیشتر دانش‌آموزان هم‌سال او یاد گرفته‌اند (آنچه که در برنامه تعیین شده است) یاد بگیرد.

$$\text{ضریب تعلیمی} = \frac{\text{سن تعلیمی}}{\text{سن تقویمی}}$$

وقتی در مقابل صفات، واکنش‌ها و رفتارهایی قرار می‌گیریم که نمی‌توانیم درمورد آنها از بازده و نتایج کمّی حرف بزنیم، بلکه بیشتر از کیفیت و شیوه‌های کودک صحبت می‌کنیم (مثلاً افسرده است، مضطرب است و..) استفاده از انواع شیوه‌های نمره‌گذاری کمّی منطقی نخواهد بود؛ مثلاً طبق شیوه‌ای که آزمودنی رفتار کرده است می‌توان او را در گروه درونگرایان، برونگرایان یا احساساتی‌ها قرار دارد.

در زمینه تکالیف تحصیلی با مشروعیت بیشتری می‌توان جنبه‌های کیفی را به جنبه‌های کمّی تبدیل کرد. این کار درمورد خط و انشاء تا اندازه‌ای تحقق می‌یابد.

آزمون تفکر مفهومی: نحوه تشکیل ساخت روان، آشکار شدن تیپ روان، مرحله تکوین، استنباط ساخت عملکرد روان (مثل آزمون‌های شیرر، بندرگشتالت، گلدشتاین، هافمن).

تفسیر ضریب هوشی

تفسیر یک نتیجه، زمانی درست خواهد بود که براساس جمعیت مرجع واقعی انجام می‌گیرد. نتایج حاصل از یک آزمون ابتدا به‌صورت مقایسه آنها با نتیجه جمعیت کل بیان می‌شود که در رابطه با کل جمعیت است.

نتیجه یک فرد همیشه باید با نتایج گروهی که فرد خود جزء آنهاست مقایسه شود. مثلاً نتیجه یک دانش‌آموز ۱۰ ساله شهرستانی باید با نتایج دانش‌آموزان ۱۰ ساله همان شهرستان مقایسه شود.

نیم‌رخ‌های روانی (پروفایل)

نیم‌رخ‌های روانی یعنی بیان تصویری نتایج آزمون‌ها.

نیم‌رخ‌های روانی ما را قادر می‌سازند تا افراد را نسبت به یک مبنای ثابت مقایسه کنیم؛ بنابراین نیم‌رخ‌های روانی نقاط مرجعی مطابق با هنجارها نشان می‌دهند. برای ترسیم نیم‌رخ باید نتایج خام را به هنجارهای سنی (مثل درصد، رتبه، نمره استاندارد، هنجارهای Z) تبدیل کنیم.

نیم‌رخ وسیله‌ای است که واقعیت ملموس تفاوت‌های روانی، توزیع و تغییرات آنها را به‌صورت کاملاً روشن نشان می‌دهند. در واقع تا بیان تصویری نتایج حد وسط بین بیان کلامی و بیان عددی یا ریاضی است.

نیم‌رخ‌های فردی

اگر اندازه‌های واقعی در دست باشند مثل اندازه‌های جسمانی، فیزیولوژیکی و نتایج تحصیلی، این نیم‌رخ‌های فردی یا بیان تصویری، قیافه درونی فرد را در یک یا چند زمینه نشان می‌دهند و بی‌نظمی‌ها و نقطه‌قوت‌ها و ضعف‌های او را نیز نشان می‌دهند.

با استفاده از نیم‌رخ‌ها می‌توانیم فرد معینی را در لحظه‌های مختلف زندگی با خودش مقایسه کنیم (مثلاً سرعت رشد جسمانی، سرعت رشد هوشی و...)

پراکندگی روانی

منظور از پراکندگی روانی که به‌عنوان شاخص عدم تجانس از آن یاد می‌شود، این است که پراکندگی فرد در مجموعه‌ای از آزمون‌ها را بررسی کنیم. هنگامی‌که شاخص پراکندگی بالا باشد، پیش‌بینی درمورد فرد دشوار خواهد بود.

امتیاز نیم‌رخ روانی در این است که نقاط قوت و ضعف فرد را ترسیم می‌کند و حتی پراکندگی روانی فرد را نیز نشان می‌دهد.

با استفاده از نیم‌رخ‌های روانی می‌توان افراد را با یکدیگر یا یک فرد معین را در زمان‌های مختلف با خودش مقایسه کرد.