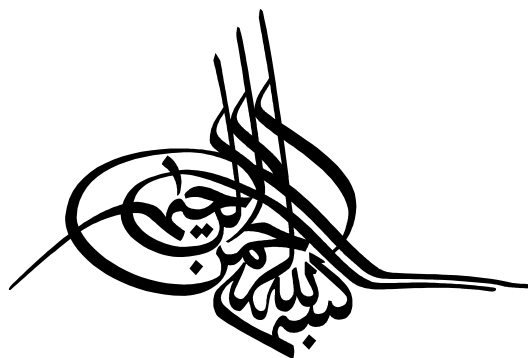




سنجش و اندازه‌گیری در تعلیم و تربیت

سری کتاب‌های کمک آموزشی کارشناسی ارشد

مجموعه علوم تربیتی ۱

مؤلف: گروه مولفان

ویراستار علمی: سید جمال بارخدا

سرشناسه	: گروه مولفان
عنوان	: سنجش و اندازه‌گیری در تعلیم و تربیت
مشخصات نشر	: تهران : مشاوران صعود ماهان، ۱۴۰۴
مشخصات ظاهری	: ۱۹۹ ص
فروست	: سری کتاب‌های کمک آموزشی کارشناسی ارشد
شابک	: ۹۷۸-۶۰۰-۳۸۹-۰۱۶-۹
وضعیت فهرست نویسی	: فیپای مختصر
یادداشت	: این مدرک در آدرس http://opac.nlai.ir قابل دسترسی است.
شناسه افزوده	: سیدجمال، بارخدا؛ ویراستار علمی
شماره کتابشناسی ملی	: ۳۷۶۹۴۵۶



نام کتاب: سنجش و اندازه‌گیری در تعلیم و تربیت
مولف:.....گروه مولفان
ویراستار علمی:سیدجمال بارخدا
ناشر:.....مشاوران صعود ماهان
نوبت و تاریخ چاپ:دوم/۱۴۰۴
تیراژ:.....۱۰۰۰ نسخه
قیمت:..... ۴/۸۳۰/۰۰۰ ریال
شابک: ISBN: ۹۷۸-۶۰۰-۳۸۹-۰۱۶-۹

انتشارات مشاوران صعود ماهان: خیابان ولیعصر، بالاتر از تقاطع مطهری،

روبروی قنادی هتل بزرگ تهران، 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.

بدین‌وسیله از مجموعه اساتید، مولفان و همکاران محترم خانواده بزرگ ماهان که در تولید و به‌روزرسانی تالیفات ماهان نقش موثری داشته‌اند، صمیمانه تقدیر و تشکر می‌نماییم.

دانشجویان عزیز و اساتید محترم می‌توانند 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

کتاب حاضر که در دست خوانندگان محترم است، نسخه جدید و کاملاً ویراست شده کتاب سنجش و اندازه‌گیری به‌عنوان یکی از دروس ضریب‌دار برای کنکور کارشناسی ارشد رشته‌های علوم انسانی و به‌ویژه روان‌شناسی ۱ و ۲، علوم تربیتی ۱ است. در انتخاب عناوین و سرفصل‌های انتخاب شده برای این کتاب تا جای ممکن سعی شده است که سرفصل‌های استراتژیک و مهم با در نظر گرفتن حجم لازم آن، که هر ساله تعداد بسیاری از سوالات را به خود اختصاص می‌دهد مورد توجه قرار بگیرد. برای مثال در دو فصل روایی و پایایی به‌عنوان دو فصل بسیار مهم، تمام سعی مولفان و ویراستار کتاب بر این امر بوده است که مطالب مهم و در حد اتکایی جمع‌آوری گردد. این کتاب شامل سرفصل‌های مهم کنکور می‌باشد که در هر فصل سوالات تألیفی و کنکور سراسری به‌همراه پاسخ تشریحی گنجانده شده است. در پایان با تشکر از تمامی دوستانی که در این امر ما را یاری داده‌اند، امیدواریم بتوانیم کمک هرچند کوچکی در زمینه مباحث این درس به داوطلبان کوشا کرده باشیم.

با تشکر

سید جمال بارخدا

فصل اول: کلیات سنجش و اندازه‌گیری	۷
تاریخچه روان‌سنجی (سنجش و اندازه‌گیری)	۸
مراحل اساسی اندازه‌گیری	۱۲
مقایسه آزمون، آزمودن، اندازه‌گیری، سنجش و ارزشیابی	۱۵
دسته‌بندی انواع آزمون‌ها و ارزشیابی آموزشی	۱۶
انواع ارزشیابی‌های آموزشی	۲۲
دسته‌بندی ارزشیابی‌های آموزشی	۲۲
دسته‌بندی ارزشیابی‌های آموزشی با توجه به موضوع ارزشیابی	۲۶
دسته‌بندی ارزشیابی‌های آموزشی با توجه به ارزشیابان	۲۷
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل اول	۳۰
فصل دوم: مقیاس‌های اندازه‌گیری	۳۱
اندازه‌گیری	۳۳
مقیاس‌های اندازه‌گیری	۳۳
مقیاس پیوسته و مقیاس گسسته	۳۶
فصل سوم: مراحل ارزشیابی آموزشی و طبقه‌بندی اهداف آموزشی	۳۷
مراحل ارزشیابی آموزشی	۳۸
تهیه طرح ارزشیابی پیشرفت تحصیلی	۴۱
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل سوم	۴۸
فصل چهارم: آزمون‌ها	۵۱
انواع آزمون‌های پیشرفت تحصیلی کتبی	۵۳
آزمون‌های عینی	۵۳
آزمون‌های تشریحی	۵۸
آزمون‌های کوتاه پاسخ	۵۹
آزمون‌های شفاهی	۶۱
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل چهارم	۶۳
فصل پنجم: روش‌های سنجش فرآیند و فرآورده‌های یادگیری غیرشناختی	۶۵
تقسیم‌بندی آزمون‌ها در حوزه‌های سه‌گانه بلوم	۶۶
آزمون‌های عملکردی	۶۸
روش‌های سنجش مشاهده‌ای	۷۱
روش‌های ویژه سنجش هدف‌های حوزه عاطفی	۷۴
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل پنجم	۷۶
فصل ششم: اجرا، نمره‌گذاری و تحلیل آزمون	۷۷
شرایط اجرای آزمون	۷۸
نمره‌گذاری	۷۹
تحلیل سوال‌های آزمون	۸۰
فصل هفتم: کاربرد روش‌های آماری در تفسیر نمره‌های آزمون	۸۷

تعریف داده‌های تجربی و آماری، متغیر، فراوانی و توزیع فراوانی.....	۹۱
تهیه توزیع فراوانی نمرات.....	۹۱
نمایش نمرات به صورت نمودار.....	۹۲
اندازه‌های گرایش مرکزی.....	۹۳
اندازه‌های تغییرپذیری.....	۹۳
همبستگی.....	۹۳
فصل هشتم: هنجارها و نیم‌رخ‌ها.....	۹۸
تعریف هنجار.....	۹۹
انواع هنجار.....	۹۹
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل هشتم.....	۱۰۳
فصل نهم: روایی آزمون.....	۱۰۴
تعریف و انواع روایی.....	۱۰۵
پایایی.....	۱۰۷
عوامل موثر بر روایی آزمون.....	۱۱۱
روایی در تحقیقات علوم انسانی.....	۱۱۱
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل نهم.....	۱۱۴
فصل دهم: پایایی آزمون.....	۱۱۵
تعریف پایایی.....	۱۱۷
روش‌های تعیین پایایی.....	۱۲۰
خطای معیار اندازه‌گیری.....	۱۲۲
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل دهم.....	۱۲۷
فصل یازدهم: ارزشیابی معلم.....	۱۲۸
تعریف ارزشیابی معلم.....	۱۳۰
هدف‌های ارزشیابی معلم.....	۱۳۰
انواع روش‌های ارزشیابی معلم.....	۱۳۰
سؤالات چهارگزینه‌ای و پاسخنامه تالیفی فصل یازدهم.....	۱۳۷
سوالات و پاسخ تشریحی کنکور سراسری ۹۵-۱۴۰۱.....	۱۳۸
منابع.....	۱۹۹

فصل اول

کلیات سنجش و اندازه گیری

- تاریخچه روان سنجی (سنجش و اندازه گیری)
- مراحل اساسی اندازه گیری
- مقایسه آزمون، آزمودن، اندازه گیری، سنجش و ارزشیابی
- دسته بندی انواع آزمون ها و ارزشیابی آموزشی

کلیات سنجش و اندازه‌گیری

تاریخچه روان‌سنجی (سنجش و اندازه‌گیری)

سنجش و اندازه‌گیری^۱ در زندگی انسان سابقه‌ای طولانی دارد و از قدمتی برابر با تاریخ خلقت انسان برخوردار می‌باشد. انسان‌ها از زمان‌های قدیم برای تهیه خوراک (مواد غذایی بهتر)، سرزمین‌ها و مراتع بهتر و همچنین برای مقابله با رقبای خود و یا دفاع و مبارزه در مقابل یورش‌های قبایل دیگر به ارزیابی قدرت خود و دشمن می‌پرداختند و پس از آن وارد عمل می‌شدند. به تدریج که اجتماعات بشری پیشرفت کردند افرادی پیدا شدند که برای تحلیل ویژگی‌های هموعان خود قدم‌های مؤثری برداشتند. این افراد، به کمک ابزار یا آزمون‌هایی که ساختند، توانستند انسان‌ها را به صورت کاملاً علمی مورد مطالعه قرار دهند، ویژگی‌های آن‌ها را حتی الامکان به طور عینی اندازه بگیرند و تفاوت‌های فردی را معلوم سازند.

افلاطون و ارسطو در آثار خود از تفاوت‌های فردی بین افراد و گروه‌ها سخن گفته‌اند. در امپراطوری چین قدیم در ۲۲۰۰ سال پیش از میلاد مسیح، کارکنان سازمان‌های کشور هر سه سال یکبار با استفاده از امتحانات شفاهی مورد آزمون قرار می‌گرفتند و نتایج امتحانات به عنوان ملاکی برای ارزشیابی عملکرد و ارتقاء آنان مورد استفاده قرار می‌گرفت. در قرن ۱۹ ابتدا کشورهای انگلیس، فرانسه، آلمان و بعدها دولت آمریکا امتحانات استخدامی را برای انتخاب کارکنان دولت به کار بستند.

بنابراین هر چند (ریشه) روان‌سنجی^۲ یا سنجش و اندازه‌گیری را می‌توان در دوران باستان جستجو کرد؛ اما رشد و توسعه آن و جنبه علمی پیدا کردن آن در قرن نوزدهم و از زمان جدایی روان‌شناسی از فلسفه صورت گرفته می‌توان گفت سنجش و اندازه‌گیری یا علم روان‌سنجی تحت تأثیر سه عامل عمده در قرن ۱۹ پا به عرصه وجود گذاشته است. این سه عامل عبارتند از: پیدایش روان‌شناسی تجربی یا آزمایشی^۳، مطالعه تفاوت‌های فردی^۴ و مطالعه عقب ماندگی‌های ذهنی^۵ یا ظهور روان‌شناسی کودکان استثنایی^۶.

پیدایش روان‌شناسی تجربی:

اولین قدم‌ها، برای تهیه آزمون‌های روانی در آزمایشگاه‌های روان‌شناسی برداشته شد، زیرا روان‌شناسان آزمایشی بودند که برای اولین بار سعی کردند زمان واکنش را اندازه بگیرند. آن‌ها به این امر توجه داشتند که آزمایش‌های روان‌شناسی نیاز مبرم به کنترل دقیق شرایط مشاهده دارند. تعداد زیادی از روان‌شناسان معتقدند که ویلهلم وونت روان‌شناس آلمانی، اولین کسی است که بیش از همه به روان‌سنجی علاقمند شده بود و در سال ۱۸۷۹، در شهر لایپزیک آلمان، اولین آزمایشگاه روان‌شناسی را دایر کرد. او ابتدا به

1- Assessment

2- Psycho metrics

3- Experimenttal psychology

4- Individual different

5- Mental Retardation

6- Exceptional children psychology

تحقیقات آزمایشگاهی در زمینه حواس مختلف انسان از جمله اندازه‌گیری قدرت شنوایی، بینایی، زمان واکنش^۱ و... پرداخت و سپس به مسائل دیگری نظیر اندازه‌گیری حوزه ادراک توجه نمود.

مطالعه تفاوت‌های فردی:

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

فرانسیس گالتون (۱۸۸۳) یکی از نخستین کسانی است که به مطالعه و اندازه‌گیری تفاوت‌های فردی پرداخت. گالتون عقیده داشت که بین توانایی ذهنی (هوش) و تمییز حسی رابطه وجود دارد. هرچه میزان هوش بالاتر باشد، سطح تمییز حسی نیز بالاتر است، بر این اساس بود که وی برای اندازه‌گیری تمییز حسی، آزمون‌های مختلفی را ابداع کرد. او معتقد بود که با این آزمون‌ها می‌توان هوش و توانایی‌های ذهنی را مورد اندازه‌گیری قرار داد. گالتون تحت تأثیر جان لاک فیلسوف انگلیسی (۱۶۳۲-۱۷۰۴) معتقد بود که تقویت حواس باعث تقویت هوش و به‌ویژه دقت می‌گردد. نمونه‌هایی از ابزارهای گالتون که امروزه اشکال مدرن‌تر آنها در آزمایشگاه‌های روان‌شناسی مورد استفاده قرار می‌گیرند به شرح زیر است:

- سوت گالتون: برای اندازه‌گیری قدرت شنوایی افراد در شنیدن بالاترین فرکانس صوتی که قابل شنیدن است، استفاده می‌شود.
- رنگ سنج^۲: برای اندازه‌گیری دقت آزمودن دو لکه رنگ استفاده می‌شود.
- وزنه‌های گالتون: که از نظر شکل، رنگ و حجم یکسان اما از نظر وزن به ترتیب سنگینی مرتب شده بودند و برای سنجش حساسیت جنبش^۳ از آنها استفاده می‌شد.
- میله‌هایی با اندازه‌های متفاوت برای آزمون تخمین دامنه دیداری.

• مجموعه‌ای از بطری‌های پر از مواد مختلف: که برای سنجیدن حس بویایی آزمودنی‌ها مورد استفاده قرار می‌گرفت.

خدمت مهم گالتون، ابداع روش‌های آماری برای کمی‌ساختن نتایج حاصل از اجرای آزمون‌ها برای مطالعه تفاوت‌های فردی و تجزیه و تحلیل این نتایج بود. روش‌های آماری مورد مطالعه وی بعدها توسط کارل پیرسون توسعه پیدا کرد که به تدوین روش همبستگی گشتاوری منجر شد. گالتون برای جمع‌آوری اطلاعات در زمینه تحقیقات خود برای اولین بار از روش پرسش‌نامه تداعی آزاد استفاده کرد. مک کین کتل روان‌شناس آمریکایی برای اولین بار در سال ۱۸۹۰ در مقاله‌ای اصطلاح «آزمون‌های روانی»^۴ را پیشنهاد کرد که با هدف سنجش تفاوت‌های فردی اشخاص نوشته شده بود. کتل شاگرد سابق وونت در لایپزیک اقدام به تأسیس آزمایشگاه روان‌شناسی نمود. او در رساله پایانی دکترای خود یک مجموعه آزمون جهت سنجش توانایی‌های مختلف افراد ابداع کرده بود و آنها را بر روی کودکان، دانشجویان و بزرگسالان آزمود. آزمون‌های کتل بیشتر به سنجش کارکرد حواس و واکنش‌های حرکتی می‌پرداختند و از این نظر تحت تأثیر کارهای گالتون قرار داشتند. کتل نیز مانند گالتون معتقد بود که از طریق آزمون‌های تن سنجی، زمان واکنش و افتراق حسی، می‌توان کنش‌های ذهنی از جمله هوش را اندازه‌گیری کرد.

معایب آزمون‌های کتل:

- ۱- آزمون‌های او پدیده‌های ساده فعالیت‌های روانی را اندازه‌گیری می‌کردند.
- ۲- فاقد فرم برای مقایسه افراد بودند.
- ۳- اعتبار و روایی آزمون‌های او پایین بود. به همین علت جای خود را به آزمون‌های دقیق‌تری دادند.

1- Reactiontime
2- Photometer
3- Kinesthetic
۴- Mental test

مطالعه عقب ماندگی‌های ذهنی:

در سال‌های اول قرن بیستم مسأله‌ای بسیار بنیادی در وزارت آموزش و پرورش کشور فرانسه مطرح شد و آن ناتوانی برخی کودکان مدارس در یادگیری متون پیشنهادی برنامه‌های رسمی بود. تمایل به همگانی کردن آموزش و پرورش و رعایت دموکراسی ایجاب می‌کرد که همه کودکان به تحصیل بپردازند، درحالی‌که برخی در نیمه راه متوقف می‌شدند. گفته می‌شد که اگر کودکان با توجه به میزان یادگیری‌شان به دسته‌های مختلف تقسیم گردند و در کلاس‌های موازی جای داده شوند، مشکل حل خواهد شد و با در نظر گرفتن ساعات تدریس بیشتر می‌توان به جبران ظرفیت محدود آن دسته از کودکانی که ضعیف هستند، پرداخت.

بنابراین نیاز به وجود مقیاسی که با استفاده از آن بتوان دانش‌آموزان را به طبقات ضعیف، قوی و متوسط تقسیم کرد، روز به روز بیشتر احساس می‌شد. در سال ۱۹۰۴ وزارت آموزش و پرورش فرانسه کمیته‌ای را برای مطالعه در مورد توانایی‌های یادگیری کودکان عقب‌مانده ذهنی مأمور کرد. بینه و یک روانپزشک به نام تئودور سیمون در این کمیته انواع تکالیف ذهنی را که کودکان در سنین مختلف می‌توانستند انجام دهند، بررسی کردند و به این ترتیب نخستین آزمون هوش را ساختند. این آزمون در سال ۱۹۰۵ انتشار یافت و ۳۰ سؤال داشت که به ترتیب از نظر سطح دشواری از آسان به مشکل تنظیم شده بود و سه کارکرد شناختی یعنی قضاوت، درک و فهم و استدلال را اندازه‌گیری می‌کرد. در سال ۱۹۰۸ این آزمون مورد تجدید نظر قرار گرفت و محتوای آن گسترش یافت و مفهوم سن عقلی در آن مطرح شد. با این شیوه محاسبه می‌توان تشخیص داد که یک کودک نسبت به سن تقویمی خود یک عقب‌مانده یا یک پیشرفته ذهنی است. چنانچه سن عقلی محاسبه شده کودک از سن تقویمی^۱ او بیشتر می‌شد، نتیجه می‌گرفتند که آن کودک پیشرفته ذهنی است و چنانچه بر عکس، کمتر از سن تقویمی او می‌شد معلوم می‌شد که آن کودک عقب‌ماندگی ذهنی دارد. در سال ۱۹۱۶ روان‌شناسی آمریکایی برای کودکان آمریکایی میزان و آزمونی به کاربرد که آزمون استنفورد - بینه نامیده شد. ترمن در این آزمون از اصطلاح هوشبهر^۲ (IQ) استفاده کرد که به‌صورت خارج قسمت سن عقلی بر سن تقویمی تعریف شده و برای حذف ارقام اعشاری در عدد ۱۰۰ ضرب می‌شود.

در جنگ جهانی اول، برای سنجش افرادی که وارد ارتش آمریکا می‌شدند ضرورت آزمون هوش گروهی احساس شد. رابرت یرکز رئیس انجمن روان‌شناسی آمریکا در رأس کمیته‌ای متشکل از ۴۰ روان‌شناس به تهیه چنین آزمونی اقدام کرد. این گروه پس از بررسی آزمون‌های مختلف، آزمون تدوین شده توسط آرتور اتیس را الگو قرار داد. اتیس در تدوین آزمون هوشی خود پرسش‌های چند گزینه‌ای را برای نخستین بار به‌کار برد. بدین‌سان آزمون ارتشی آلفا و بتا ساخته شد، آزمون آلفا که یک آزمون کلامی بود برای افراد با سواد که خواندن و نوشتن به زبان انگلیسی را می‌دانستند، به‌کار می‌رفت. آزمون بتا یک آزمون غیر کلامی بود که به‌جای راهنمایی کتبی یا شفاهی، روش نشان دادن از راه پانتومیم در آن به‌کار بسته می‌شد و برای افراد بی‌سواد یا کسانی که زبان انگلیسی را نمی‌دانستند، از این آزمون استفاده می‌شد.

علاوه بر آزمون‌های هوش، آزمون‌های شخصیت نیز در جنگ جهانی اول مورد توجه قرار گرفتند. پیش از آن امیل کرپلین روانپزشک آلمانی «آزمون تداعی آزاد اندیشه‌ها»^۳ را که قبلاً توسط گالتون ابداع شده بود، تدوین کرد و در مورد بیماران روانی به‌کار برد. بعدها کارل یونگ روش مشابهی را با عنوان «تداعی آزاد کلمات» برای تشخیص عقده‌های روانی بیماران به‌کار بست. در جنگ جهانی اول، ارتش آمریکا برای تشخیص افراد نظامی مبتلا به روان رنجوری به یک آزمون شخصیت نیاز پیدا کرد. رابرت وودورث در این رابطه به تهیه برگه اطلاعات فردی^۴ همت گماشت والتردیل اسکات یکی از شاگردان وونت، آزمون‌هایی را برای گزینش کارکنان موسسات صنعتی و تجاری تدوین کرد. به‌این ترتیب کاربرد آزمون‌های روانی به‌گونه گسترده‌ای در تمامی سازمان‌های دولتی و غیردولتی برای استخدام و گزینش رواج یافت.

1- Cronological Age
2- Intelligence Quotient
3- Free Association test
4- Personal Data sheet

۱- نیتکو (۲۰۰۱) اندازه‌گیری آموزشی را به‌صورت زیر تعریف کرده است:
"شیوه نسبت دادن یک عدد که معمولاً به آن نمره می‌گویند به صفت یا ویژگی معین یک شخص به‌گونه‌ای که آن عدد درجه‌ای از آن صفت یا ویژگی را که آن شخص دارا می‌باشد، نشان دهد."
۲- علی‌رغم کاربرد وسیع اصطلاح اندازه‌گیری ارائه تعریفی جامع از آن که مورد پذیرش همه افراد باشد، بسیار دشوار است، بنابراین هرچند در تعریف اندازه‌گیری از نظر لغوی میان روان‌شناسان و بقیه افراد اتفاق نظر وجود دارد، اما در تعریف کامل و علمی از اندازه‌گیری در میان متخصصان اختلاف نظر مشاهده می‌شود.
مفاهیم اساسی روان‌سنجی (سنجش و اندازه‌گیری) شامل: اندازه‌گیری^۱، آزمون^۲، آزمودن^۳، سنجش^۴، ارزشیابی^۵، ارزشیابی آموزشی^۶ و پژوهش آموزش^۷ می‌باشد. در اینجا هر کدام از این مفاهیم را توضیح می‌دهیم.

اندازه‌گیری:

در اندازه‌گیری، ویژگی‌ها یا صفات اشیا و افراد تعیین می‌گردد و مقدار آن ویژگی‌ها یا صفات به‌صورت عدد یا رقم گزارش می‌شود؛ بنابراین، می‌توان گفت: «اندازه‌گیری عبارت است از فرآیندی که تعیین می‌کند یک شخص یا یک چیز چه مقدار از یک ویژگی برخوردار است.» (گی، ۱۹۹۱)
اندازه‌گیری فرآیند یا شیوه‌ای دقیق و نظام‌دار است.

ناتالی (۱۹۷۸) در تعریفی که از اندازه‌گیری به‌دست داده بر کاربرد قواعد^۸ تأکید کرده است: «اندازه‌گیری دربرگیرنده قواعدی است برای نسبت دادن اعداد به اشیا (یا افراد) به‌گونه‌ای که صفاتی از آن‌ها را به‌صورت کمیت نشان دهد». در تعریف فوق سه خصیصه اساسی فرآیند اندازه‌گیری بیان شده است که وجود هر کدام از آن‌ها شرط کافی در تعریف اندازه‌گیری محسوب می‌شود. منظور از کاربرد قواعد در اندازه‌گیری این است که روش‌های نسبت دادن اعداد باید به‌طور روشن بیان شوند. مثلاً در نمره‌گذاری نتایج یک آزمون حساب، وقتی که جواب‌های درست دانش‌آموزان شمارش می‌شوند و به تعداد جواب‌های درست به آن‌ها نمره داده می‌شود، شاهد نمونه‌ای از کاربرد قاعده در اندازه‌گیری هستیم؛ بنابراین اندازه‌گیری «مشی‌هایی هستند برای نشان دادن مقدار (یا نوعی ویژگی دیگر) شیء موردنظر اندازه‌گیری» (کوهن و سورد لیک ۲۰۰۲، ص ۷۳).
فایده بیان قواعد اندازه‌گیری این است که کمک می‌کند تا افراد مختلفی که می‌خواهند وسیله اندازه‌گیری را مورد استفاده قرار دهند آن را یکسان به کار برند.

بیان قواعد اندازه‌گیری از مسایل مهم استاندارد کردن یا میزان کردن^۹ آزمون‌های روانی و تربیتی است. از مشخصات مهم آزمون‌های استاندارد یا میزان شده این است که افراد مختلفی که این وسائل اندازه‌گیری را به کار می‌بندند، نتایج یکسانی به‌دست می‌آورند.

ویژگی‌های خصایص افراد:

تاکید بر اصطلاح ویژگی یا خصایص افراد در تعریف اندازه‌گیری، به این منظور است که نشان داده شود در فرآیند اندازه‌گیری، هیچ‌گاه خود اشیا یا افراد اندازه‌گیری نمی‌شوند بلکه آنچه مورد اندازه‌گیری واقع می‌شود، ویژگی، خصیصه یا صفاتی از اشیا و یا افراد است، اصطلاح «متغیر» به کار رفته است که از لحاظ علمی و عملی یعنی صفتی که مقدار آن نزد اشیا با افراد مختلف، متفاوت است مثلاً وزن شاگردان برحسب کیلوگرم خصیصه‌ای است که مقدار آن برای همه آن‌ها یکسان است، به کار بردن این اصطلاح در تعریف اندازه‌گیری، دلالت بر این دارد که فرآیند اندازه‌گیری همواره ارتباط ویژگی خاصی از شیء یا فرد دارد.

-
- 1- Measurment
 - 2- Test
 - 3- Testiny
 - 4- Assessmenttal
 - 5- Evaluation
 - 6- Educational Evaluation
 - 7- Educational Research
 - 8- Rules
 - ۹- Standardization

کمی‌سازی

کاربرد اصطلاح کمیت یا کمی‌سازی خصیصه‌ها در تعریف اندازه‌گیری، به این معنی است که فرآیند اندازه‌گیری باید نشان دهد که چه مقدار از یک صفت در فرد با یک شیء موردنظر وجود دارد. به عبارت دیگر اندازه‌گیری صحیح آن است که نتیجه آن حتماً به صورت عدد بیان شود.

با توجه به موارد فوق می‌توان گفت در فرآیند اندازه‌گیری، ویژگی‌ها، صفات یا خصایص اشیاء یا افراد مشخص می‌شود و آنگاه مقدار آن ویژگی‌ها یا خصایص را به صورت عدد یا رقم گزارش می‌کنند. بنابراین فرآیند اندازه‌گیری باید دارای مراحل باشد که در زیر به آن‌ها اشاره می‌کنیم.

مراحل اساسی اندازه‌گیری

همان‌گونه که گفته شد هدف از اندازه‌گیری به دست آوردن اطلاعات دقیق در مورد ویژگی افراد یا اشیاء است. به منظور دستیابی به این هدف هومن (۱۳۶۸) سه مرحله اساسی را مشخص کرده است که به اختصار توضیح داده می‌شود.

الف- تشخیص و تعریف خصیصه (صفت) مورد اندازه‌گیری

نخستین گام در اندازه‌گیری هر خصیصه تعریف دقیق آن است. در مواردی که صفات مورد اندازه‌گیری از نوع کمیت‌های فیزیکی مانند طول، قد، وزن و مانند آن باشد، به آسانی قابل تعریف است؛ اما تعریف صفات کیفی نظیر هوش، علاقه و غیره بسیار دشوار است، در چنین مواردی صفات مورد اندازه‌گیری باید به صورت رفتارهای قابل مشاهده و قابل اندازه‌گیری تعریف شوند. مثلاً حافظه سمعی ممکن است به صورت زیر تعریف شود: «تعداد کلماتی که فرد از میان بیست کلمه بلافاصله بعد از یکبار شنیدن آن‌ها را به خاطر می‌آورد» این نوع تعریف‌ها را تعاریف عملیاتی می‌نامند.

ب- تهیه وسایل اندازه‌گیری (تعیین رشته عملیات لازم برای آشکارسازی و قابل درک کردن خصیصه مورد اندازه‌گیری)

دومین گام در اندازه‌گیری یک صفت تهیه وسیله‌ای است که به کمک آن بتوان صفات موردنظر را اندازه گرفت، مانند قدسنج برای اندازه‌گیری قد. برای اندازه‌گیری صفات کیفی مانند هوش، استعداد، معلومات و غیره این وسیله باید به طریقی ساخته شود که بروز رفتارهای مورد اندازه‌گیری را به صورتی که تعریف شده‌اند موجب گردد. به عنوان مثال هر گاه صفت مورد اندازه‌گیری به صورت «توانایی حل مسائل مربوط به چهار عمل اصلی در حساب» تعریف شده باشد، وسیله اندازه‌گیری این صفت آزمون خواهد بود که هر یک از پرسش‌های آن توانایی آزمایش‌شوندگان را در یکی از زمینه‌های مربوط به «حل چهار عمل اصلی در حساب» نشان می‌دهد.

ج- تهیه مقیاس کمی (کمی‌سازی خصیصه مورد مطالعه به صورت واحدهای درجه یا مقدار)

کمی‌سازی بر این امر اشاره دارد که چقدر یا چه میزان از یک ویژگی یا خصیصه در یک شیء یا فرد وجود دارد؛ بنابراین، پس از مشخص ساختن سلسله عملیات لازم جهت اندازه‌گیری خصیصه موردنظر برای بیان نتایج، لازم است ماحصل یا نتیجه کار به صورت کمیت (اعداد و ارقام) بیان شود، یعنی برای اینکه مقدار موردنظر ویژگی در فرد باشد حتماً باید به شکل اعداد یا ارقام بیان شود.

همچنین علاوه بر اینکه مقدار ویژگی موردنظر به صورت عدد یا رقم بیان می‌شود باید برای مقیاس موردنظر واحد یا درجه در نظر گرفته شود، زیرا برخلاف اکثر صفات فیزیکی که در ابزار اندازه‌گیری آن‌ها واحدهای مقیاس با هم برابرند، در روان‌شناسی و آموزش و پرورش صفاتی که مورد مطالعه هستند دارای واحدهای برابر نیستند و از قبل نیاز به تعریف عملیاتی متغیر و مقیاس‌های اندازه‌گیری مناسب دارند.^۱

۱- جهت مطالعه بیشتر به فصل انواع مقیاس‌ها رجوع کنید.

آزمون:

گفتیم که اندازه‌گیری یک فرآیند است. این فرآیند نیاز به وسیله‌ای دارد که به آن وسیله اندازه‌گیری می‌گویند. اندازه‌گیری ویژگی‌های مختلف انسان‌ها و افراد به دو شیوه زیر انجام می‌شود:

۱- مستقیم } ویژگی‌های فیزیکی
رفتارهای آشکار یا محصولات رفتاری

۲- غیرمستقیم ← ویژگی‌ها یا صفات روانی

- برای اندازه‌گیری ویژگی‌های فیزیکی مثل طول و وزن اشیاء یا قد و وزن افراد از متر و ترازو استفاده می‌شود.
- برای اندازه‌گیری رفتارهای آشکار یا محصولات رفتاری نیز از روش غیرمستقیم استفاده می‌شود. مثلاً تعداد دفعاتی که دانش‌آموز در یک جلسه درس صندلی‌اش را ترک می‌کند (رفتار آشکار) یا تعداد مسائل ریاضی که درست حل می‌کند (محصول رفتار) را می‌توان به‌طور غیرمستقیم اندازه گرفت.

- ویژگی‌ها یا صفات روانی مانند هوش، خلاقیت، انگیزش و... را باید به‌طور غیرمستقیم اندازه‌گیری کرد؛ زیرا هیچ کس به‌طور مستقیم به این ویژگی‌ها دسترسی ندارد. معمول‌ترین وسیله اندازه‌گیری ویژگی‌ها یا صفات روانی آزمون است. آزمون وسیله‌ای عینی و استاندارد شده است که برای اندازه‌گیری نمونه‌ای از رفتار یا خصایص آدمی به‌کار می‌رود. در این تعریف مراد از عینی آن است که روش اجرا، نمره دادن و تعبیر و تفسیر نتایج آزمون براساس قواعدی معین و مشخص صورت می‌گیرد و قضاوت و نظر شخصی در آنها بی‌تأثیر است. برای هر آزمون دستورالعمل خاصی در مورد نحوه اجرا، طول مدت آزمایش، دستورهای شفاهی و نحوه ارائه مثال‌ها و روش تفسیر نمره‌ها تهیه می‌شود.

اصطلاح استاندارد شده به این معنا است که آزمون قبلاً در مورد گروه نمونه‌ای از افراد موردنظر در بوته آزمایش گذاشته شده و نتایج پژوهش‌های مربوط به آن از راه روش‌های آماری مورد تجزیه و تحلیل قرار گرفته و روابی و اعتبار آن تعیین شده است، به‌علاوه دارای جداول نرم یا هنجارهایی است که نمره‌های خام آزمون بر مبنای آن‌ها تعبیر و تفسیر می‌شوند.

مقصود از رفتار یا خصایص آن است که در هر آزمون نمونه‌های کوچکی از رفتار یا خصایص فرد که به دقت انتخاب شده است مورد مشاهده و اندازه‌گیری قرار می‌گیرد. همچنانکه یک متخصص شیمی برای تعیین مواد محلول در آب قطره‌ای از آن را مورد آزمایش قرار می‌دهد روان‌شناس نیز مثلاً برای اندازه‌گیری استعداد در ریاضی یک فرد می‌تواند نمونه‌هایی از توانایی او را در حل مسائل، محاسبات کمی و با استفاده استدلال ریاضی مورد مطالعه قرار دهد.

معرف بودن محتوای آزمون و قدرت تشخیص و پیش‌بینی آن بسته به این است که ماهیت و تعداد پرسش با چه دقتی از میان پرسش‌های ممکن انتخاب شده باشد و تا چه اندازه نمونه‌هایی واقعی رفتار را مورد سنجش قرار می‌دهد؛ بنابراین باید مواد آزمون با رفتارهای مورد پیش‌بینی مشابهت داشته باشد.

علاوه بر موارد بالا باید به این نکته هم اشاره کنیم که نتایج هر آزمون تنها در مورد گروهی از افراد یا جامعه‌ای معتبر است که این آزمون قبلاً در مورد آن‌ها اجرا شده باشد یعنی افراد نمونه‌ای استاندارد شده که از نظر خصایص و متغیرهای بسیار مؤثر در نتایج آزمون‌های اجرا شده عامل فرهنگی مشترکند.

آزمون همچنین وسیله یا روش نظام‌دار (نظام‌مند) برای اندازه‌گیری نمونه‌ای از رفتار است. علاوه بر آزمون روش‌های دیگری برای اندازه‌گیری ویژگی‌های روانی وجود دارد. با وجود این، معروف‌ترین وسیله اندازه‌گیری در آموزش و پرورش و روان‌شناسی آزمون است که به‌صورت مجموعه‌های سؤال و غالباً به‌طور کتبی برای پاسخگویی در اختیار آزمون‌شوندگان قرار داده می‌شود.

واژه آزمون به معنی محدود و متداول آن یعنی مجموعه‌ای از سؤال‌های چند گزینه‌ای و در مورد معنای گسترده آن به مجموعه‌ای از سؤالات عینی چند گزینه‌ای، جور کردنی، صحیح - غلط، کوتاه پاسخ، سؤال‌های تشریحی و همچنین سایر شیوه‌های اندازه‌گیری روانی و تربیتی نظیر پرسشنامه به‌کار می‌رود. در کل برای اندازه‌گیری تمامی ویژگی‌های روانی و تربیتی عملاً هیچ‌گونه محدودیتی وجود ندارد.

دیوید شوfer (۱۹۹۴) آزمون‌های مورد استفاده در روان‌شناسی و آموزش و پرورش را در سه دسته کلی زیر قرار داده است.

- ۱- آزمون‌هایی که در آن‌ها آزمون‌شونده تکلیف بخصوصی را انجام می‌دهد، مانند نوشتن، جواب سؤال‌های تشریحی، انتخاب جواب سؤال‌های چند گزینه‌ای، یا به‌طور ذهنی چرخاندن تصاویری که بر روی صفحه کامپیوتر ظاهر می‌شوند.
- ۲- مشاهده رفتار آزمون‌شونده در یک موقعیت ویژه، مانند مشاهده رفتار یک فروشنده در برخورد با مشتریان بد خلق، به‌منظور تعیین توانایی او در شغل فروشنده‌گی.
- ۳- مقیاس‌های گزارش شخصی که در آن‌ها آزمون‌شونده حساب‌ها، نگرش‌ها و باورهای خود را بیان می‌کند، نظیر پرسش‌نامه‌ها، فرم‌های نظر سنجی و غیره.

آزمون سنجش:

وقتی برای اندازه‌گیری یکی از ویژگی‌های روانی یا تربیتی یک فرد یا گروهی از افراد از آزمون استفاده می‌شود به این عمل یا فعالیت آزمون می‌گویند. در تعریف آزمون گفتیم که آزمون نمونه‌ای از رفتارهای موردنظر اندازه‌گیری است، از این‌رو "اندازه‌گیری متغیرهای روانشناختی به‌وسیله تدابیر یا روش‌هایی که به منظور کسب نمونه‌هایی از رفتار طراحی شده‌اند، آزمون روانی یا روان آزمایی نام دارد" (کوهی و سورد کب ۲۰۲، ص ۴).

از آنجا که آزمون یک وسیله اندازه‌گیری است، هر وقت آزمون صورت می‌پذیرد، یعنی آزمونی به کار می‌رود، نوعی اندازه‌گیری می‌شود، اما هر نوع اندازه‌گیری الزاماً آزمون نیست.

چنانکه دیدیم، وقتی که اندازه‌گیری صرفاً با استفاده از آزمون انجام می‌شود، آزمون صورت می‌پذیرد؛ اما زمانی که در اندازه‌گیری به جای آزمون یا علاوه بر آن وسایل آزمون مختلفی را به کار می‌برند، به آن سنجش می‌گویند. از جمله آن‌ها می‌توان به آزمون پرسش‌نامه، مقیاس درجه‌بندی، فهرست‌وارسی و... اشاره کرد؛ بنابراین سنجش نیز نوعی اندازه‌گیری است که مفهوم وسیع‌تری از آزمون و اندازه‌گیری دارد. اصطلاح سنجش بیشتر در حوزه روان‌شناسی بالینی و آموزش و پرورش به کار می‌رود.

ارزشیابی:

اصطلاح (ارزشیابی یا ارزیابی) به‌طور ساده، به تعیین ارزش^۱ برای هر چیزی یا داوری ارزشی^۲ گفته می‌شود. ال.ج. کروبنانه (۱۹۶۴) ارزشیابی را به‌عنوان "گردآوری و کاربرد اطلاعات به منظور اخذ تصمیم درباره برنامه آموزشی تعریف می‌کند و سی.ای.بی.بی (۱۹۷۵) آن را به‌عنوان فرآیند نظام‌دار^۳ گردآوری و تفسیر و شواهدی که منجر به داوری ارزشی می‌شود و در نهایت به عمل می‌انجامد (به نقل هومن ۱۳۷۴) و گی (۱۹۹۷) "ارزشیابی به یک فرآیند نظام‌دار (سیستماتیک) برای جمع‌آوری، تحلیل و تفسیر اطلاعات گفته می‌شود، به این منظور که تعیین شود آیا هدف‌های موردنظر تحقق یافته‌اند یا در حال تحقق یافتن هستند و به چه میزانی" (به نقل از سیف ۱۳۸۲)

آنچه که از تعاریف بالا استنباط می‌شود این است که ارزشیابی، خود حوزه بسیار گسترده‌ای دارد و علاوه بر موارد بالا جنبه‌های کیفی رفتار را نیز برآورده می‌کند.

بنابراین اصطلاح ارزشیابی در معنای وسیع خود، عبارت است از عملی که با آن درباره یک واقعه، یک فرد یا یک شیء با توجه به یک معیار (یا چند معیار) قضاوت می‌شود. از طرف دیگر موضوع قضاوت و معیار ارزشیابی هرچه باشد در نفس عمل تغییر ایجاد می‌کند؛ بنابراین ارزشیابی حوزه بسیار گسترده‌ای دارد اما در معیار محدود، آن نوع ارزشیابی که در جهت قضاوت ارزشی در مورد تحقیق یافتن یا نیافتن اهداف آموزشی به کار می‌رود، ارزشیابی آموزشی نامیده می‌شود.

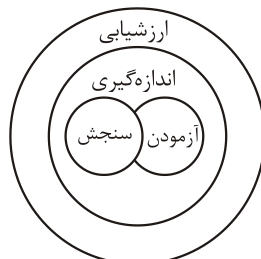
1- Value

2- Value judgement

3- Systematic

مقایسه آزمون، سنجش، اندازه‌گیری و ارزشیابی:

اینک می‌توانیم روابط مفاهیم اساسی روان‌سنجی یعنی آزمون، سنجش، اندازه‌گیری و ارزشیابی را با هم به شکل زیر نشان دهیم. در شکل ۱-۱ روابط بین این مفاهیم نشان داده شده‌اند. آزمون و سنجش بخشی از اندازه‌گیری هستند و اندازه‌گیری نیز جزو ارزشیابی است، اما سنجش هم مستقل از آزمون می‌باشد و هم می‌تواند علاوه بر آزمون شامل فعالیت‌های دیگر نیز باشد. محدودترین مفهوم روان‌سنجی آزمون است و گسترده‌ترین آن ارزیابی می‌باشد. مفاهیم اساسی روان‌سنجی به ترتیب از محدود به گسترده از لحاظ اندازه‌گیری، عبارتند از آزمون، سنجش و ارزشیابی.



شکل ۱-۱: مقایسه آزمون، سنجش، اندازه‌گیری و ارزشیابی

ارزشیابی آموزشی و پژوهش آموزشی:

هم پژوهش آموزشی و هم ارزشیابی آموزشی جزو فعالیت‌های علمی به حساب می‌آیند، زیرا هر دو تقریباً به روش و فنون بررسی تجربی متکی‌اند.

با وجود این، بین پژوهش آموزشی و ارزشیابی آموزشی تفاوت‌هایی وجود دارند که عمده‌ترین این تفاوت‌ها به قرار زیرند:

۱- گرچه هم در ارزشیابی آموزشی و هم در پژوهش آموزشی اطلاعات جمع‌آوری می‌شوند و دانش کسب می‌گردد، اما هدف ارزشیابی به‌طور عمده تصمیم‌گیری است، ولی هدف پژوهش استنتاج است. بعضی پژوهشگران عمدتاً علاقه‌مند به درک پدیده‌ها برای شناسایی بهتر آن‌ها هستند؛ اما متخصصان ارزشیابی یا ارزشیابان علاقه‌مند به فهم پدیده‌ها برای هدایت اعمال دیگران هستند.

۲- در پژوهش آموزشی، مانند هر نوع پژوهش دیگری، تعمیم‌پذیری نتایج پژوهشی از اهمیت خاصی برخوردار است؛ بنابراین، در فعالیت پژوهشی کوشش فراوان به عمل می‌آید تا یافته‌ها به موارد دیگر تعمیم یابند. در مقابل ارزشیابی آموزشی بر یک پدیده آموزشی ویژه متمرکز است و تعمیم‌پذیری در آن نقش ندارد.

۳- ارزشیابی آموزشی با ارزش سروکار دارد، درحالی‌که پژوهش آموزشی مانند سایر پژوهش‌ها خالی از ارزش است. چنان‌که قبلاً گفته شد، هدف اصلی هرگونه فعالیت ارزشیابی به ویژه ارزشیابی آموزشی تعیین قدر و ارزش پدیده مورد ارزشیابی است تا اینکه به افراد علاقه‌مند و مسئول کمک کنند تا درباره آن پدیده تصمیم‌های درستی اتخاذ نمایند. در مقابل، پژوهشگران به دنبال حقایق علمی هستند و هیچ‌گونه قدر و ارزش مثبت یا منفی برای حقایق به‌دست آورده خود قائل نمی‌شوند.

۴- به اعتقاد بسیاری از صاحب‌نظران از جمله اپین (۱۹۹۴) پژوهش وابسته به نظریه است، اما ارزشیابی الزاماً مبنای خطر ندارد. به اعتقاد این صاحب‌نظران، هدف اصلی پژوهش تولید دانش تازه و افزودن آن به نظریه‌های موجود است. درحالی‌که مطالعات ارزشیابی عموماً به منظور حل مشکلات عملی ویژه و کمک به تصمیم‌گیری صورت می‌پذیرد. (اپین ۱۹۹۴، ص ۱۲، سیف ۱۳۸۲).

مقایسه آزمون، آزمودن، اندازه‌گیری، سنجش و ارزشیابی

مفهوم آزمون از مفهوم اندازه‌گیری محدودتر است، چرا که آزمون یک وسیله اندازه‌گیری است. آزمون فقط برای عمل آزمودن استفاده می‌شود، اما در اندازه‌گیری از فنون و ابزار مختلف استفاده می‌شود.

آزمودن یک ابزار است و یک وسیله اندازه‌گیری می‌باشد، در آزمودن فقط از آزمون استفاده می‌شود. سنجش نوعی تحلیل بالینی به منظور پیش‌بینی عملکرد می‌باشد.

اندازه‌گیری و آزمودن عموماً با کمیت سروکار دارند، اما سنجش الزاماً به کمیت منحصر نمی‌شود یعنی نتیجه سنجش می‌تواند به صورت غیر کمی (توصیفی) گزارش شود.

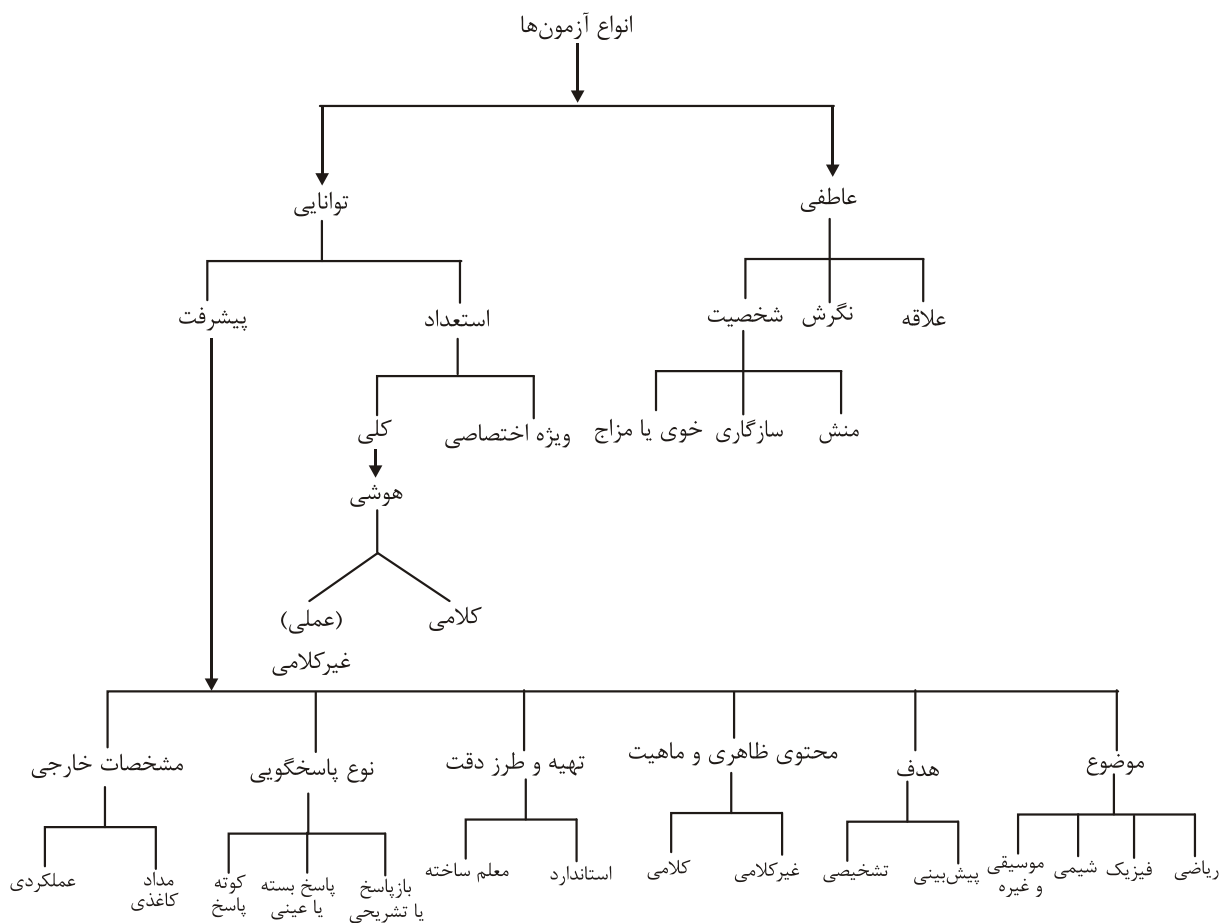
ارزشیابی جامع‌تر از اندازه‌گیری است. اندازه‌گیری همان‌طور که گفته شد صرفاً کمی است، اما ارزشیابی هم کمی و هم کیفی است. اندازه‌گیری برای مفاهیم تعدادی از افراد از لحاظ یک ویژگی مورد استفاده قرار می‌گیرد ولی ارزشیابی برای تعیین میزان کارآمدی یک برنامه یا دوره آموزشی انجام می‌گیرد، به عبارت دیگر در ارزشیابی داوری و قضاوت وجود دارد، ولی در اندازه‌گیری هیچ‌گونه داوری در مورد نتایج به دست آمده، انجام نمی‌شود. اندازه‌گیری درست پایه و اساس ارزشیابی می‌باشد.

بخش اول: دسته‌بندی انواع آزمون‌ها و ارزشیابی آموزشی

این فصل تقسیم‌بندی انواع آزمون‌ها و ارزشیابی آموزشی با توجه به ملاک‌های مختلف بیان می‌شود و سعی می‌شود نوعی الگوی تقسیم‌بندی ارائه شود که تقریباً به صورت جامع انواع آزمون‌ها را شامل می‌شود.

دسته‌بندی آزمون‌ها

همان‌طور که در فصل نخست بیان کردیم متداول‌ترین ابزار اندازه‌گیری ویژگی‌ها یا صفات روانی افراد آزمون است. هرچند در آموزش و پرورش و روان‌شناسی علاوه بر آزمون روش‌های دیگر اندازه‌گیری و سنجش مورد استفاده قرار می‌گیرند، با این حال آزمون یکی از ابزارهای متداول است.

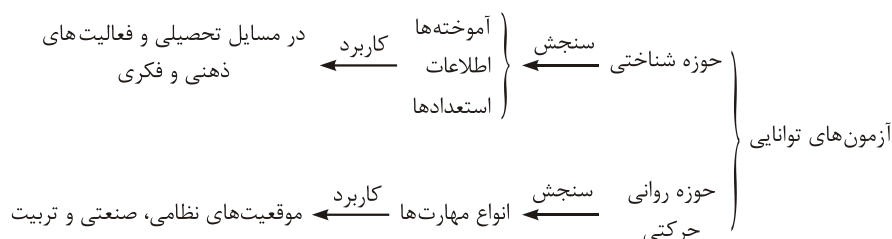


نمودار ۱-۱

آزمون‌های مختلف را می‌توان با توجه به ویژگی‌های گوناگونی که اندازه می‌گیرند ابتدا به دو دسته کلی آزمون‌های توانایی^۱ یا شناختی و روانی حرکتی و یا غیرعاطفی و آزمون‌های عاطفی^۲ تقسیم کرد و بعد هر یک از این دو دسته آزمون را نیز به اجزای کوچک‌تری به شرح زیر دسته‌بندی نمود.

آزمون‌های توانایی:

این آزمون‌ها دو مقوله مهم توانایی‌های شناختی^۳ و توانایی‌های حرکتی^۴ را اندازه می‌گیرند؛ بنابراین آزمون‌های توانایی، توانایی‌های شناختی و ذهنی مثل آموخته‌ها و اطلاعات و استعدادها نیز توانایی‌های ذهنی و عملی مثل انواع مهارت‌ها را مورد سنجش قرار می‌دهند. آزمون‌های شناختی اساساً در مسایل تحصیلی و فعالیت‌های ذهنی و فکری به کار می‌روند، اما آزمون‌های روانی - حرکتی در موقعیت‌های صنعتی، نظامی، حرفه‌ای و تربیت بدنی به کار می‌روند. ویژگی اصلی آزمون‌های توانایی (شناختی و روانی - حرکتی) این است که هنگام اجرای آن‌ها از آزمون شونده، خواسته می‌شود تا حداکثر سعی خود را به کار ببندند تا بهترین نمره را به دست آورد.



آزمون‌های توانایی به دو دسته آزمون‌های استعداد^۵ و پیشرفت تحصیلی^۶ تقسیم می‌شوند.

آزمون‌های استعداد

این آزمون‌ها مقدار توانایی یا آمادگی فرد را برای انجام دادن کارهایی که در پیش دارد و همچنین ظرفیت یادگیری‌های مختلف را می‌سنجند، به عبارت دیگر توانایی‌های بالقوه فرد را می‌سنجند، ناظر بر آینده هستند و برای پیش‌بینی موفقیت‌های آتی فرد به کار می‌روند. آزمون‌های هوش و استعدادهاى مختلف معرف آزمون‌های استعداد هستند.

آزمون‌های استعداد را به دو دسته آزمون‌های استعداد کلی و استعداد ویژه تقسیم می‌کنند.

۱- **آزمون‌های استعداد کلی**^۷: این آزمون که توانایی‌های ذهنی را که تقریباً در همه انواع فعالیت‌های فکری و شناختی ضروری‌اند، مورد سنجش قرار می‌دهد و به تعبیر دیگر هدف از کاربرد آزمون‌های هوش کلی پیش‌بینی موفقیت‌های همه جانبه افراد است.

۲- **آزمون استعداد ویژه یا جزئی**^۸: این آزمون استعدادهاى اختصاصی افراد را در زمینه‌های مختلف اندازه‌گیری می‌کند، هدف از کاربرد استعدادهاى ویژه پیش‌بینی موفقیت یا شایستگی افراد در زمینه‌های فکری و عملی مشخص است. آزمون‌های استعداد ویژه با توجه به زمینه‌هایی که استعداد افراد در آن سنجیده می‌شود، نام‌گذاری می‌شوند.

-
- 1- Ability
 - 2- Affectibve
 - 3- Cognitive
 - 4- Psychomotor
 - 5- Aptitude
 - 6- Academic(scholastic)Achirement
 - 7- Specific
 - 8- General

آزمون‌های پیشرفت تحصیلی

این آزمون‌ها دانش و مهارت‌هایی را که فرد تا لحظه اجرای آزمون کسب کرده است اندازه می‌گیرند، به عبارت دیگر توانایی‌های بالفعل را مورد سنجش قرار می‌دهند. آزمون‌های پیشرفت تحصیلی با موضوع یا محتوایی که اندازه می‌گیرند مشخص می‌گردند و در نتیجه با عنوان آن موضوع‌ها نام‌گذاری می‌شوند، مانند آزمون پیشرفت ریاضی، آزمون پیشرفت فیزیک و آزمون پیشرفت مهارت‌های فنی و غیره.

اجرای آزمون‌های پیشرفت تحصیلی در جریان آموزش مهم هستند، زیرا اولاً این آزمون‌ها کاربرد فراوان دارند و ابزار کار هر معلمی هستند و به جهت ارزشیابی در میزان تحقق هدف‌های تدریس به کار می‌روند. ثانیاً نتایج به دست آمده از این آزمون‌ها به عنوان شاخصی برای تعیین میزان آموخته‌های قبلی و درجه تبحر و شایستگی فرد در راهنمایی او برای موفقیت احتمالی آینده در انتخاب رشته تحصیلی و شغلی مربوط به آنان است.

طبقه‌بندی آزمون‌ها بر اساس موضوع

آزمون‌های پیشرفت تحصیلی در طبقه‌بندی بر حسب موضوع انواع زیادی دارند؛ مثلاً می‌توان از آزمون پیشرفت تحصیلی ریاضیات، پیشرفت تحصیلی فیزیک و سایر علوم نام برد.

طبقه‌بندی آزمون‌ها از نظر هدف

آزمون‌های پیشرفت تحصیلی از نظر هدف به آزمون‌های پیش‌بینی و تشخیصی تقسیم می‌شوند.

الف - آزمون‌های پیش‌بینی: در این آزمون‌ها می‌خواهیم بدانیم آزمودن دارای چه ویژگی‌هایی خواهد بود. پس این آزمون‌ها ناظر بر آینده می‌باشند و در برنامه‌ریزی برای آینده فرد یا موفقیت‌های آتی مورد استفاده قرار می‌گیرند.

ب - آزمون‌های تشخیصی: در این آزمون‌ها می‌خواهیم ویژگی‌های فعلی روانی آزمودنی (ویژگی‌های هوشی، تحصیلی و روانشناختی فرد) را بشناسیم. این آزمون‌ها همچنین ناظر بر حال می‌باشند.

طبقه‌بندی آزمون‌ها از نظر محتوای ظاهری و ماهیت

این آزمون‌ها به دو دسته آزمون‌های کلامی^۱ (غیرعملی) و غیر کلامی^۲ (عملی) تقسیم می‌شوند.

الف - آزمون‌های کلامی: در این آزمون‌ها سؤالات به صورت گفتاری یا نوشتاری مطرح می‌شوند و آزمودنی پاسخ خود را به صورت گفتار یا نوشتار ارائه می‌کند. آزمون‌های کلامی معمولاً وابسته به فرهنگ هستند، این آزمون‌ها به سه دسته تقسیم می‌شوند:

- ۱- **گفتاری - گفتاری:** سؤالات و پاسخ‌ها به صورت گفتار ارائه می‌شوند، مثل آزمون بینه و سیمون یا آزمون رنه زازو.
- ۲- **گفتاری - نوشتاری:** سؤالات به صورت گفتار مطرح و پاسخ‌های نیز به صورت نوشتار ارائه می‌شوند.
- ۳- **نوشتاری - نوشتاری:** سؤالات و پاسخ‌ها به صورت نوشتار ارائه می‌شوند.

در همه این آزمون‌ها زبان کلامی نقش اساسی دارد و کسانی که استعداد کلامی بالاتری دارند نمرات بهتری دریافت می‌کنند. برخی افراد آزمون‌های گفتاری - نوشتاری و نوشتاری - نوشتاری را آزمون‌های مداد - کاغذی می‌نامند.

ب - آزمون‌های غیر کلامی: آزمون‌هایی هستند که در آن‌ها نقش زبان به حداقل ممکن کاهش می‌یابد، این آزمون که ممکن است از نوع عملکرد باشد پاسخ به آن‌ها غالباً مستلزم انجام فعالیتی از طرف آزمودنی است. در این آزمون‌ها استعدادهای کلامی

1- Verbal

2- Non Verbal

کمترین نقش را دارند و در عوض هوش عملی یا مکانیکی فعال می‌شود. آزمون‌های عملی اکثراً استعدادهای ویژه را مورد سنجش قرار می‌دهند و بیشتر جزو آزمون‌های ناوابسته به فرهنگ^۱ می‌باشند، از جمله آزمون‌های غیرکلامی می‌توان از آزمون هوش ناوابسته به فرهنگ لایتر، آزمون مکعب‌های چوبی کس، آزمون قطعات چوبی وگلی، آزمون جعبه دکرولی نام برد. علاوه بر آزمون‌های کلامی، آزمون‌های دیگری وجود دارند که سؤالات آن‌ها برخی کلامی و برخی غیرکلامی می‌باشند. از جمله این آزمون‌ها به مقیاس‌های وکسلر و استنفورد - بینه می‌توان اشاره کرد. این آزمون‌ها را به علت تنوع و گسترش حوزه‌های سنجش «آزمون‌های جامع» نیز می‌نامند.

طبقه‌بندی آزمون‌ها از نظر دقت و طرز تهیه

آزمون‌های پیشرفت تحصیلی از نظر دقت و طرز تهیه به آزمون‌های استاندارد شده^۲ و آزمون‌های معلم ساخته^۳ تقسیم می‌شوند:

الف - آزمون‌های میزان شده (معیار شده و استاندارد شده): آزمون‌هایی هستند که با استفاده از نمونه‌های بزرگ آزمون‌شوندگان و غالباً به‌وسیله موسسات یا بنگاه‌های آزمون سازی دولتی یا خصوصی و یا مراکز آموزشی و علمی و برای اندازه‌گیری پیشرفت تحصیلی یا هوش و استعداد گروه‌های بزرگ دانش‌آموزان و با مقاصد ارزشیابی از دوره‌های تحصیلی تهیه و توزیع می‌شوند، در این نوع آزمون‌ها، ملاک نسبی اندازه‌گیری مورد استفاده قرار می‌گیرد.

کرون باخ (۱۹۹۵) "استاندارد بودن یا میزان بودن آزمون را به این صورت بیان می‌کند که شیوه‌های اجرا، مشاهده، تجهیزات، مواد و قواعد نمره‌گذاری تا چه اندازه‌ای تثبیت شده‌اند که تا حد امکان در شرایط و زمان‌های مختلف نتایج یکسانی به‌دست می‌آید". به‌طورکلی آزمون‌های میزان شده، براساس دستورالعمل‌های خاص و به شیوه یکسان اجرا و نمره‌گذاری و تفسیر می‌شوند. برای تفسیر نمره آزمونی در آزمون استاندارد شده، فرم‌ها یا هنجارهایی براساس عملکرد گروه نمونه وسیعی از آزمودنی‌ها فراهم شده است. تفسیر این آزمون‌ها براساس مقایسه نمره آزمودنی با گروه هنجار می‌باشد.

ب - آزمون‌های معلم ساخته یا غیراستاندارد: آزمون‌هایی هستند که معلمان برای سنجش پیشرفت تحصیلی دانش‌آموزان خود تهیه می‌کنند و در طول دوره‌های آموزشی و یا در پایان دوره به اجرا درمی‌آورند. این آزمون‌ها غالباً برای تعیین میزان توفیق یادگیرندگان در رسیدن به هدف‌های آموزشی طرح‌ریزی می‌شوند، از این جهت معمولاً در آن‌ها ملاک مطلق اندازه‌گیری به‌کار گرفته می‌شود. برتری این نوع آزمون‌ها نسبت به آزمون‌های شفاهی این است که اولاً شرایط آزمون برای همه دانش‌آموزان یکسان است و ثانیاً در تصحیح اوراق نظر تصحیح‌کننده دخالت می‌کند.

مقایسه آزمون‌های استاندارد شده با آزمون‌های غیر استاندارد (معلم ساخته)

تراندایک و هیگن (۱۹۷۹) سه نوع تفاوت را بیان می‌کنند.

- ۱- تهیه آزمون‌های استاندارد شده در مقایسه با آزمون‌های معلم ساخته، به زمان و تخصص بیشتری نیاز دارد.
- ۲- آزمون‌های میزان شده برای استفاده در مناطق گوناگون با یکدیگر تهیه می‌شوند. به این دلیل باید اهداف آموزشی مناطق مختلف را در این آزمون‌ها قرار داد؛ اما آزمون‌های معلم ساخته با توجه به محتوا و هدف‌های ویژه دروس به خصوصی تهیه می‌شوند؛ لذا چنین آزمون‌هایی برای مقایسه میزان پیشرفت موازن مناطق مختلف خیلی مفید نیستند.
- ۳- اصطلاح «آزمون استاندارد شده» به وجود داده‌های هنجاری اشاره می‌کند و ناشران این نوع آزمون‌ها، برای معنا دار کردن نمرات آزمون خود، هنجارهای لازم را برای آن‌ها تدارک می‌بینند. اما آزمون‌های معلم ساخته نیازی به هنجار یا هنجاریابی ندارند.

1- Free culture
2- Standardized
3- Teach made

ویژگی آزمون‌های معلم ساخته:

- ۱- این نوع آزمون‌ها همه هدف‌هایی را که به‌وسیله معلمان مشخص شده‌اند، دربرنمی‌گیرند. همچنین آزمون‌های معلم ساخته تنها می‌تواند یادآوری ساده را اندازه بگیرد همچنین سطح دیگر یادگیری از جمله فهمیدن، به‌کار بستن، ترکیب و ارزشیابی را اندازه‌گیری می‌کند.
- ۲- محتوای این‌گونه آزمون‌ها به مطالب جزئی و کم اهمیت گرایش دارد و قدرت خلاقیت و توانایی ترکیب جنبه‌های مهم مطالب را بررسی نمی‌کند.
- ۳- مواد آزمونی که معلم در کلاس به کار می‌برد مطابقت چندانی با ارزشیابی فرآیندهای ذهنی یا موضوعاتی که آزمون باید اندازه بگیرد، ندارند، همچنین محتوای آزمون به‌ندرت معرف همه مطالبی است که در یک دوره معین به‌وسیله معلم تدریس شده است.
- ۴- متن سؤال‌هایی که به‌وسیله معلمان تهیه می‌شوند، مبهم، سخت و مشکل هستند.

طبقه‌بندی آزمون‌ها از نظر مشخصات خارجی یا فرم و شکل

- آزمون‌های پیشرفت تحصیلی از نظر مشخصات خارجی به آزمون‌های مداد- کاغذی^۱ و عملکردی^۲ تقسیم می‌شوند:
- الف- آزمون‌های مداد- کاغذی: در این آزمون‌ها، آزمودنی به‌طورکتابی به سؤال‌هایی که از او پرسیده می‌شود، پاسخ می‌دهد.
- ب- آزمون‌های عملکردی: در این آزمون‌ها از آزمودنی خواسته می‌شود که یک سلسله کارهای دستی بر روی ابزارهایی که به او ارائه می‌شود، انجام دهد.

طبقه‌بندی آزمون‌ها از نظر نوع پاسخ‌دهی

- آزمون‌های پیشرفت تحصیلی نیز به سه دسته تشریحی یا باز پاسخ، عینی یا بسته پاسخ و کوتاه پاسخ تقسیم می‌شوند. از آنجا که این دسته از آزمون‌های پیشرفت تحصیلی موضوع فصول بعدی هستند، از توضیح بیشتر درباره آنها خودداری می‌کنیم.

آزمون‌های عاطفی:

برخلاف آزمون‌های غیرعاطفی (شناختی و روانی- حرکتی) که توانایی‌های فکری و عملی را می‌سنجند، این آزمون‌ها خصوصیات و ویژگی‌های شخصیتی افراد را ارزیابی می‌کنند، این آزمون‌ها، رفتار و ویژگی‌های آزمون نمونه را در موقعیت‌های طبیعی زندگی می‌سنجند و همچنین این آزمون‌ها بیشتر به امور مشاوره در حل مسائل روانی- تربیتی و حرفه‌ای کمک می‌کنند، در ضمن این آزمون‌ها به اندازه آزمون‌های توانایی، عینی و دقیق نیستند. آزمون‌های عاطفی دارای انواع زیادی هستند که مهم‌ترین آنها آزمون‌های علاقه، شخصیت و نگرش می‌باشند.

۱- آزمون علاقه

علاقه، به احساس فرد نسبت به یک فعالیت اشاره می‌کند، به آزمون علاقه، آزمون رغبت^۳ نیز می‌گویند. آزمون‌های رغبت، برای اندازه‌گیری علایق آزمودنی‌ها درمورد فعالیت‌ها، رشته‌های تحصیلی و شغلی ساخته می‌شوند. علاقه و رغبت برای اندازه‌گیری علایق مثبت و منفی افراد نسبت به یک فعالیت اشاره می‌کنند. داشتن علاقه به چیزی صرفاً به این معنی است که شخص برای اندیشیدن درباره آن چیز وقت صرف می‌کند یا به آن واکنش عملی نشان می‌دهد. صرف نظر از اینکه این واکنش‌ها مثبت یا منفی باشند. رغبت‌سنج‌ها از جمله این نمونه‌های این نوع می‌باشند.

1- Pencil Paper

2- Performance

3- Interest

۲- آزمون نگرش^۱

نگرش را می‌توان برحسب نظریه‌های یادگیری و رویکردهای شناختی تعریف کرد. گوردون آلپورت (۱۹۳۵) نگرش را چنین تعریف کرده است: «نگرش یک حالت آمادگی ذهنی و عصبی است که از طریق تجربه سازمان پویا برجا می‌گذارد» (سیرز و دیگران، ۱۹۹۱). نگرش نسبت به هر شیء، اندیشه، شخص، گروه و یا موقعیت به‌عنوان جهت‌گیری پایدار نسبت به عناصر شناختی، عاطفی و رفتاری تلقی می‌شود. بخش شناختی نگرش، شامل همه افکار، حقایق، دانش‌ها و باور داشت‌هایی است که شخص در مورد موضوع نگرش دارد. بخش عاطفی یا ارزشی شامل آمادگی فرد برای پاسخ دادن و گرایش او به انجام عمل درمورد موضوع نگرش است؛ بنابراین نگرش از یک طرف با آموخته‌ها و تجارب قبلی ما و از سوی دیگر با برداشت‌های ذهنی زمان حال ما در مورد موضوع نگرش در ارتباط است و در هر صورت واکنش منفی یا مثبت ما را نسبت به اشیاء، موقعیت‌ها، سازمان‌ها، مفاهیم و افراد برمی‌انگیزد. از میان روش‌های مختلف سنجش نگرش، استفاده از پرسش‌نامه‌های سنجش و مقیاس‌های نگرش سنج مناسب‌ترین روش به‌شمار می‌روند. یکی از نخستین مقیاس‌های نگرش سنج، مقیاس فاصله اجتماعی است که در ۱۹۲۵ توسط بوگارد ساخته شد. پرسش‌نامه‌های نگرش دو قطب (بنیادگرایی - محافظه‌کاری و وطن‌پرستی - جهان‌وطنی) را مشخص می‌کند.

۳- آزمون شخصیت

واژه شخصیت^۲ معادل انگلیسی پرسونالیتی است که از ریشه پرسونا به معنای نقاب (ماسک) گرفته شده است و منظور از شخصیت نیز همان چهره یا نقابی است که افراد در اجتماع به خود زده‌اند. بیشتر رفتارهای اجتماعی افراد تعیین‌کننده شخصیت آنهاست. لارنس ای. پروین: در کتاب خود می‌نویسد: شخصیت آن دسته از ویژگی‌های افرادی است که الگوهای ثابت رفتاری آن‌ها را نشان می‌دهد. بعضی از روان‌شناسان علاوه بر رفتارهای آشکار، ویژگی‌های غیررفتاری از جمله هوش، قد و قامت، وضع سلامتی، کیفیت صدا و وضع ظاهر را به تعریف شخصیت اضافه می‌کنند. آلپورت بیش از ۵۰ تعریف از شخصیت ارائه می‌کند که جامع‌ترین تعریف از شخصیت عبارت است از "سازمان پویای نظام‌های روانی فرد که سازگاری‌های او با محیط را تعیین می‌کند."

آزمون‌های شخصیت به‌طور عمده به اندازه‌گیری متغیرهای انگیزش^۳، مزاج^۴، سازگاری^۵، منش^۶ و چیزهایی از این دست اختصاص یافته است. آزمون‌های شخصیت را براساس دیدگاه‌ها و نظریه‌های روان‌شناسان می‌توان به دو دسته تقسیم‌بندی کرد که عبارت از آزمون‌های ترکیبی و آزمون‌های تحلیلی می‌باشند.

۱- **آزمون‌های تحلیلی:** در این آزمون‌ها بر استقلال عناصر شخصیت تاکید می‌شود و ابزار سنجش آن عبارت از پرسش‌نامه‌ها و آزمون‌های عینی می‌باشند.

۲- **آزمون‌های ترکیبی:** در این آزمون‌ها بر اصل وحدت فرد توجه می‌کنند، این آزمون‌ها فرد را برحسب شیوه‌ای که در موقعیت‌های مختلف به کار می‌برد، تعریف می‌کنند.

ابزارهای اندازه‌گیری شخصیت را می‌توان به سه دسته عمده تقسیم نمود که به شرح زیر می‌باشند:

۱- پرسش‌نامه‌های گزارش^۷ شخصی: این ابزارها وسایلی هستند که در آنها فرد نظر شخصی‌اش را نسبت به خود بیان می‌کند.

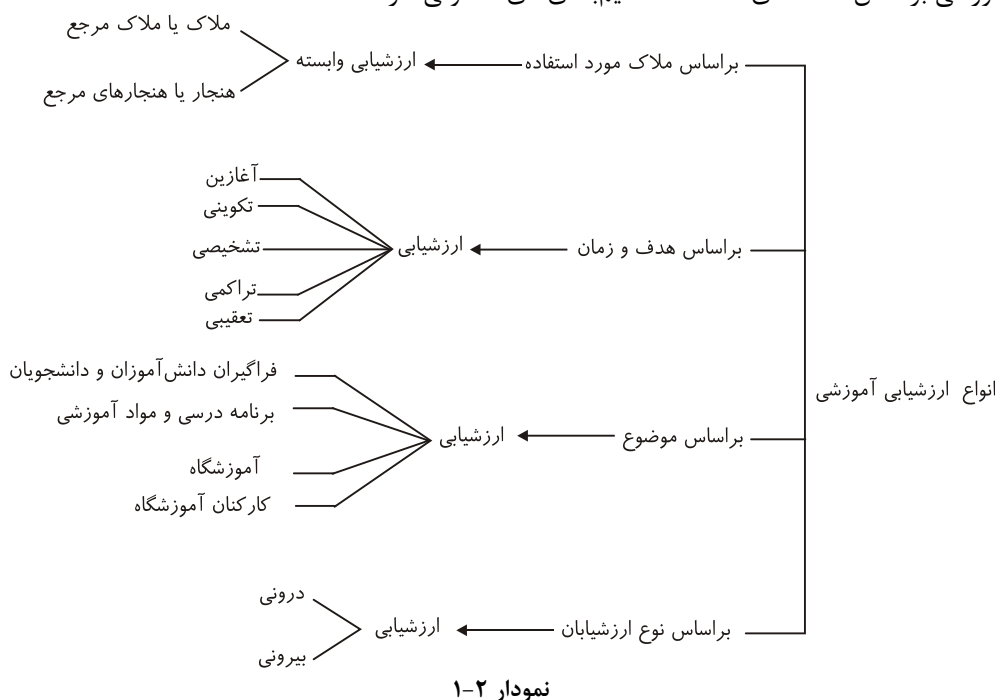
۲- پرسش‌نامه‌های جامعه سنجی: در این وسایل اندازه‌گیری، افراد دیگر نظرشان را نسبت به فرد مورد سنجش ابراز می‌کنند.

۳- فنون مشاهده‌ای: در این نوع اندازه‌گیری فرد در یک موقعیت به خصوص مورد مشاهده قرار می‌گیرد.

1- Attitude
2- Personality
3- Motivation
4- Temperament
5- Hdjastment
6- Charater
7- Questionnairs

بخش دوم: انواع ارزشیابی‌های آموزشی

ارزشیابی‌های آموزشی براساس ملاک‌های مختلف، تقسیم‌بندی‌های متفاوتی دارند.



دسته‌بندی ارزشیابی‌های آموزشی

الف- دسته‌بندی ارزشیابی‌های آموزشی با توجه به ملاک مورد استفاده

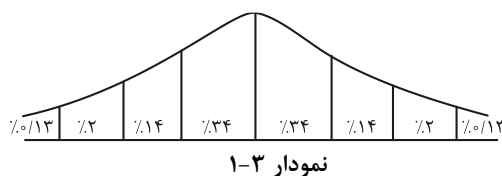
معروف‌ترین تقسیم‌بندی برای ارزشیابی‌های آموزشی تقسیم‌بندی آن‌ها به ارزشیابی وابسته به ملاک و ارزشیابی وابسته به هنجار است.

۱- ارزشیابی وابسته به ملاک

در ارزشیابی وابسته به ملاک یا ارزشیابی ملاکی^۱، معیار یا ملاک ارزشیابی از پیش تعیین می‌شود و عملکرد یادگیرنده در آزمون با توجه به آن ملاک مورد قضاوت قرار می‌گیرد. در این نوع ارزشیابی که بیشتر ویژه پیشرفت تحصیلی دانش‌آموزان و دانشجویان است، نوعی ملاک مطلق که همان هدف‌های آموزشی از پیش تعیین شده هستند، مورد استفاده قرار می‌گیرد. در این نوع ارزشیابی، از آزمون‌شدگان انتظار می‌رود که به هدف‌های آموزشی معینی دست یابند یا به سطح بخصوصی از عملکرد برسند. در استفاده از ملاک مطلق در ارزشیابی پیشرفت تحصیلی، میزان یادگیری یا پیشرفت یادگیرنده با توجه به آنچه آموزش داده شده است، ارزشیابی می‌شود.

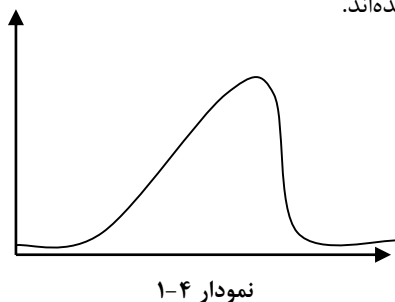
نکته:

۱- در توزیع، اکثر نمرات حول و حوش میانگین قرار دارند و تعداد بسیار کمی از آن‌ها بالاتر از میانگین و به همان نسبت هم تعداد بسیار کمی پایین‌تر از میانگین قرار دارند.



1- Criterion- Referenced Evaluation

۲- این نوع توزیع‌ها نشان‌دهنده بالا بودن سطح همه نمرات دانش‌آموزان است که نشان می‌دهد اکثریت بزرگی از آنان به اکثریت هدف‌های آموزشی موردنظر (ملاک مطلق) رسیده‌اند.



۲- ارزشیابی وابسته به هنجار^۱

ارزشیابی وابسته به هنجار یا ارزشیابی هنجاری: به آن نوع ارزشیابی‌هایی گفته می‌شود که به جای یک ملاک مطلق از پیش تعیین شده، بر نوعی ملاک نسبی وابسته است. در این نوع ارزشیابی اگر بخواهیم پیشرفت تحصیلی گروهی از دانش‌آموزان را ارزشیابی کنیم، عملکرد دانش‌آموزان را با یکدیگر مقایسه می‌کنیم (با یک ملاک مشخص و از پیش تعیین شده). علت نام‌گذاری این نوع ارزشیابی به ارزشیابی هنجاری این است که در آن ملاک یا معیار ارزشیابی نوعی هنجار گروهی است بنابراین در این نوع ارزشیابی می‌توان تعیین کرد که پیشرفت یک دانش‌آموز نسبت به دانش‌آموزان دیگر چه وضعی دارد؛ اما نمی‌توان مشخص کرد پیشرفت یک دانش‌آموز نسبت به هدف‌های آموزشی یا آنچه آموزش داده شده چگونه است.

۳- مقایسه ارزشیابی وابسته به ملاک با ارزشیابی وابسته به هنجار

ارزشیابی وابسته به ملاک (ملاک مرجع)	ارزشیابی وابسته به هنجار (هنجار مرجع)
۱- مبتنی بر ملاک مطلق است.	۱- مبتنی بر ملاک نسبی است.
۲- سنجش پیشرفت تحصیلی	۲- سنجش هوش و استعداد، سنجش شخصیت و سایر جنبه‌های عاطفی
۳- عملکرد آزمون‌شونده با ملاک مطلق مقایسه می‌شود مثل امتحان کلاسی	۳- عملکرد آزمون‌شونده، با عملکرد گروه هنجار مقایسه می‌شود
۴- شکل توزیع طبیعی می‌باشد.	۴- مثل کنکور سراسری، شکل توزیع کمی منفی است.

ب- دسته‌بندی ارزشیابی‌های آموزشی با توجه به زمان و هدف استفاده از آنها

۱- ارزشیابی آغازین یا سنجش آغازین

نخستین ارزشیابی معلم که پیش از انجام فعالیت‌های آموزشی او به اجرا در می‌آید، ارزشیابی آغازین یا سنجش آغازین^۲ است. این نوع ارزشیابی به دو منظور یعنی برای پاسخ دادن به دو پرسش زیر مورد استفاده قرار می‌گیرد:

۱- آیا یادگیرندگان بر دانش‌ها و مهارت‌های پیش نیاز درس تازه از پیش مسلط‌اند؟

۲- یادگیرندگان چه مقدار از هدف‌ها و محتوای درس تازه را قبلاً یاد گرفته‌اند؟

سؤال یک سنجش آغازین، به منظور اندازه‌گیری رفتارهای ورودی^۳ و دانش‌ها و مهارت‌های پیش‌نیاز یادگیری درس تازه به کار می‌رود و سؤال دوم به آزمون جابجایی^۴ یا پایه‌گزینی که به آن پیش آزمون^۱ گفته می‌شود، مربوط است.

1 - Norm- Refereneed Evaluation

2- Entry Behavior

3- Entry Behavior

4- Place Ment Test

رفتار ورودی برخی اوقات هدف ورودی نامیده می‌شود. هدف ورودی یا رفتار ورودی به آمادگی فرد برای یادگیری هدف‌های آموزشی (نهایی و بین راه) اشاره می‌کند. از طریق سنجش رفتارهای ورودی معلوم می‌شود که یادگیرندگان برای یادگیری هدف‌ها و محتوای درس تازه، قبلاً آمادگی‌های لازم را کسب کرده‌اند یا نه.^۱ برای مثال، یک آزمون مربوط به مهارت‌های محاسباتی را در آغاز درس جبر یا یک آزمون دستور زبان انگلیسی را برای یادگیرندگان انگلیسی زبان در آغاز درس زبان آلمانی مورد استفاده قرار داد.^۲ (گرانلاند، ۱۹۸۸ ص ۲ به نقل از سیف ۱۳۸۲)

سؤال دوم که به آزمون‌های جابه‌جایی یا پایه پیش آزمون مربوط می‌شود عمدتاً همان آزمون نهایی است که شامل هدف‌های یادگیری یا هدف‌های آموزشی درس یا ترجیحاً فرم موازی آن است. هدف از اجرای این آزمون تعیین میزان اطلاعات یادگیرندگان از مطالبی است که قرار است به آنان آموزش داده شود. متخصصان پیشنهاد داده‌اند که بهتر است معلم در ارزشیابی آغازین خود ترکیبی از دو آزمون بالا یعنی آزمون رفتارهای ورودی و پیش آزمون را با هم به اجرا درآورد.

۲- ارزشیابی تکوینی^۲

آنچه عمدتاً به منظور کمک به اصلاح موضوع مورد ارزشیابی یعنی برنامه یا روش آموزشی، مورد استفاده قرار می‌گیرد، ارزشیابی تکوینی نام دارد.

هدف از اجرای ارزشیابی تکوینی در مورد این گونه برنامه‌ها یا مواد آموزشی، آگاه ساختن تولید کنندگان برنامه از نواقص برنامه خود و کمک به اصلاح آنهاست. هدف از کاربرد ارزشیابی تکوینی در رابطه با پیشرفت تحصیلی دانش‌آموزان و دانشجویان آگاهی از میزان و نحوه یادگیری آنان برای تعیین نقاط قوت و ضعف یادگیری و نیز تشخیص مشکلات روش آموزشی معلم در رابطه با هدف‌های آموزشی است. این ارزشیابی در طول دوره آموزش، یعنی زمانی که فعالیت آموزشی معلم هنوز در جریان است و یادگیری دانش‌آموزان در حال تکوین یا شکل‌گیری است، انجام می‌شود؛ به همین دلیل به آن نام ارزشیابی تکوینی داده‌اند. در ارزشیابی تکوینی در پایان هر واحد درسی، معلم یک آزمون دقیق که حاوی هدف‌های آموزشی آن واحد است، اجرا می‌کند و براساس نتایج حاصل از آن می‌تواند در زمانی که یادگیری هنوز در جریان است و امکان رفع مشکلات یادگیری دانش‌آموزان و برطرف کردن نواقص روش آموزشی خود او میسر است، به انجام چنین کاری اقدام کند.

لازم به یادآوری است که آزمون‌های مورد استفاده در ارزشیابی تکوینی، آزمون‌های ملاکی یا ملاک مرجع می‌باشند، زیرا این آزمون‌ها برای اندازه‌گیری هدف‌های دقیق آموزشی از پیش تعیین شده که همان هدف‌های واحد درسی یا واحد آموزشی هستند، تدوین می‌شوند. در ارزشیابی تکوینی هدف تعیین میزان توفیق یادگیرنده در رسیدن به تک تک هدف‌های واحد درسی است و هیچ گونه مقایسه میان افراد مختلف موردنظر نیست.

۳- ارزشیابی تشخیصی^۳

نوع دیگر ارزشیابی که مانند ارزشیابی تکوینی در جریان آموزش انجام می‌شود، ارزشیابی تشخیصی نام دارد. علت این نام‌گذاری آن است که این ارزشیابی با هدف تشخیص مشکلات یادگیری دانش‌آموزان در یک موضوع درسی به کار می‌رود.

ارزشیابی تشخیصی، زمانی مورد استفاده قرار می‌گیرد که معلم با مشکلات مبهم در یک یا چند دانش‌آموز روبه‌رو می‌شود که با روش‌های اصلاحی معمول ارزشیابی تکوینی قابل رفع شدن نیستند، به عنوان مثال: اگر دانش‌آموزی در درس ریاضیات یا درسی

1- Per Test
2- Formative Evaluation
3- Diagnostic Evaluation

دیگر مرتباً شکست می‌خورد و معلم از طریق تفسیر روش آموزشی خود با تدابیر جبرانی دیگری از عهده حل مشکل برنمی‌آید، آنگاه اجرای یک آزمون تشخیصی و یافتن ریشه‌های مشکل ضروری است.

لین و گرانلاند (۲۰۰۱) در مقایسه ارزشیابی تشخیصی با ارزشیابی تکوینی گفته‌اند "اگر بخواهیم از مقیاس پزشکی استفاده کنیم، سنجش تکوینی (ارزشیابی تکوینی) برای مشکلات ساده یادگیری جنبه کمک‌های اولیه را دارد؛ ولی سنجش تشخیصی به دنبال کشف علت‌های زیربنایی مشکلاتی است که به کمک‌های اولیه پاسخ نمی‌دهند."

آزمون‌های تشخیصی در مقایسه با آزمون‌های پیشرفت تحصیلی که پیشرفت یا توفیق دانش‌آموزان را در درس‌های مختلف می‌سنجند، توصیف کاملی از نقاط قوت و ضعف یادگیرندگان در زمینه‌های مهارتی خاص به‌دست می‌دهد. (اگن و کلوپاک، ۲۰۰۱، ص ۶۴)

ابزارهای مورد استفاده سنجش تشخیصی وابسته به هنجارهای ویژه هستند که برای همین منظور تهیه شده‌اند. علاوه بر این‌ها، از آزمون‌های ملاک مرجع یا فنون مشاهده نیز استفاده می‌شود. به نظر سیفرت (۱۹۹۱) برای تشخیص مشکلات یادگیری دانش‌آموزان، آزمون‌های تشخیصی وابسته به ملاک کارسازتر از آزمون‌های تشخیصی هنجار مرجع هستند؛ زیرا نتایج آزمون‌های تشخیصی «باید نشان دهند که دانش‌آموزان چه می‌دانند [یا چه نمی‌دانند] نه اینکه در مقایسه با دانش‌آموزان دیگر چه وضعی دارند».

۴- ارزشیابی تراکمی^۱

در ارزشیابی تراکمی، تمامی آموخته‌های دانش‌آموزان در طول یک دوره آموزشی تعیین می‌شوند و هدف، نمره دادن به دانش‌آموزان و قضاوت درباره اثر بخشی کار معلم و برنامه درسی یا مقایسه برنامه‌های مختلف درسی با یکدیگر است.

از آنجا به این نوع ارزشیابی، تراکمی یا مجموعه می‌گویند که به‌وسیله آن می‌توان یادگیری دانش‌آموزان یک دوره آموزشی را اندازه‌گیری کرد، این نوع ارزشیابی معمولاً در پایان دوره آموزشی به عمل می‌آید، به همین سبب به آن ارزشیابی پایانی نیز می‌گویند. با این حال از آن می‌توان به دفعات بیشتر در دوره آموزشی استفاده کرد. آزمون‌های پایانی که در ارزشیابی تراکمی به کار می‌روند، بسیار جامع و مفصل می‌باشند و نمونه کاملی از محتوای درس و هدف‌های آموزشی را دربرمی‌گیرند.

هدف ارزشیابی تراکمی هم نمره دادن و اتخاذ تصمیم‌های مختلف (ارتقای دانش‌آموزان به کلاس‌های بالاتر، دادن امتیازهای تحصیلی، اعزام به دوره‌های عالی‌تر تحصیلی، اعطای بورس تحصیلی و غیره) و هم تعیین اثر بخشی کار معلم و میزان توفیق او در پیاده کردن مراحل مختلف طرح آموزشی است.

بنابراین آزمون‌های مورد استفاده در ارزشیابی تراکمی هم آزمون‌های هنجاری یا وابسته به ملاک نسبی هستند و هم آزمون‌های ملاکی یا وابسته به ملاک.

۵- مقایسه ارزشیابی تراکمی و ارزشیابی تکوینی

مخاطبان	}	در ارزشیابی تکوینی	← طراحان تولیدکنندگان برنامه هستند.
		در ارزشیابی تراکمی	← قرار خواهند داد یا از آن استفاده می‌کنند.
مورد استفاده	}	در ارزشیابی تکوینی	← کمک به تصمیمات مربوط به توسعه و تولید در برنامه
		در ارزشیابی تراکمی	← یاری دادن تصمیمات مربوط به ادامه مقطع با گسترش برنامه

نمودار ۵-۱

ج- دسته‌بندی ارزشیابی‌های آموزشی با توجه به موضوع ارزشیابی

۱- ارزشیابی فراگیران از دانش‌آموزان و دانشجویان

منظور از ارزشیابی دانش‌آموزان^۱ عمدتاً ارزشیابی پیشرفت تحصیلی^۲ یا ارزشیابی میزان یادگیری^۳ آنان است؛ با این وجود در یک برنامه ارزشیابی از دانشجویی یا دانش‌آموزی، متغیرهای دیگری چون استعداد، هوش، شخصیت، نگرش و علاقه در نظر گرفته می‌شوند. اهداف مهم کاربرد برنامه ارزشیابی دانشجویی و دانش‌آموزی عبارتند از:

۱- قضاوت درباره عملکرد یادگیرندگان

۲- داوری درباره اثربخشی کیفیت آموزشی

۳- داوری درباره روش‌ها و مواد آموزشی

عمده‌ترین ملاک قضاوت در ارزشیابی آموزشی، هدف‌های آموزشی از پیش تعیین شده می‌باشد؛ بنابراین مهم‌ترین نوع ارزشیابی آموزشی که ارزشیابی پیشرفت‌های تحصیلی می‌باشد، تعیین می‌نماید که دانش‌آموزان یا دانشجویان تا چه میزان به هدف‌های آموزشی از پیش تعیین شده رسیده‌اند.

۲- ارزشیابی از برنامه‌های درسی و مواد آموزشی

ارزشیابی آموزشی گاهی به منظور تعیین اثربخشی برنامه درسی^۴ یا مواد آموزشی^۵ انجام می‌شود و شامل ارزشیابی از عواملی چون روش‌های آموزشی، کتاب‌های درسی، مواد دیداری، شنیداری و تدارکات فیزیکی و سازمانی است. در ارزشیابی از برنامه‌های درسی و مواد آموزشی دو ملاک عمده را باید در نظر گرفت:

۱- پیشرفت تحصیلی یا میزان یادگیری دانش‌آموزان و دانشجویان

۲- هزینه

۳- ارزشیابی از پروژه‌ها و برنامه‌های آموزشی

پروژه‌ها^۶ و برنامه‌های^۷ آموزشی نیز گونه‌ای دیگر از برنامه‌هایی هستند که مورد ارزشیابی قرار می‌گیرند. منظور از پروژه‌ها و برنامه‌های آموزشی ویژه کوشش‌هایی هستند که عبارتند از:

۱- موقعیتی که به‌عنوان بخشی از برنامه‌های معمول آموزشگاه‌ها به حساب نمی‌آیند.

۲- ماهیت نوآورانه دارند.

۳- طول مدت حضورشان به موفقیتشان وابسته است، مفاهیم پروژه، برنامه و برنامه درسی از هم متمایزند. (پین ۱۹۹۴، ص ۶)

۴- پروژه یک کوشش مجزا و احتمالاً یکباره است که با استفاده از یک نوآوری می‌کوشد تا تغییری به‌وجود آورد.

• برنامه از پروژه گسترده‌تر و پیچیده‌تر است. "یک پروژه خاص ممکن است یکی از چند نوآوری‌هایی باشد که به منظور کمک به رفع مشکل دانش‌آموزان ابتدایی تدارک دیده شده است."

• برنامه درسی از پروژه و برنامه‌های مفید است، پین (۱۹۹۴) برنامه درسی را به‌عنوان مجموع‌های از برنامه‌ها تعریف کرده است.

1- Student Evaluation
2- Academic Achirement
3- Learning
4- Curriculum
5- Instructional Materials
6- Programs
7- Progecl

۴- ارزشیابی آموزشی آموزشگاه

منظور از ارزشیابی آموزشی آموزشگاه^۱ این است که تعیین شود هدف‌های از پیش تعیین شده برای آموزشگاه به چه میزانی تحقیق یافته‌اند و نقاط قوت و ضعف کدامند. ارزشیابی آموزشگاه دربرگیرنده ارزشیابی از کل برنامه و مستلزم گردآوری اطلاعاتی لازم درباره همه جنبه‌های آموزشگاه است.

یکی از جنبه‌های مهم ارزشیابی برای آموزشگاه، برنامه آزمون آموزشگاه است. در برنامه ارزشیابی آموزشگاهی مهم‌ترین بخش برنامه آزمون و منبع کسب اطلاعات، پیشرفت تحصیلی دانش‌آموزان و دانشجوین است.

۵- ارزشیابی از کارکنان آموزشگاه^۲

مقصود از کارکنان آموزشگاه، ارزشیابی از همه کسانی است که به‌طور مستقیم یا غیرمستقیم مسئول بازدهی آموزش موسسه آموزشی هستند؛ هر چند که این ارزشیابی بیشتر به منظور تعیین میزان اثربخشی یا کارآمدی هیأت آموزشی معلمان به‌کار رود. پیچیده‌ترین نوع ارزشیابی کارکنان، ارزشیابی کار معلمان است. علت این پیچیدگی، اعتبار و دقت پایین ابزار سنجش این نوع ارزشیابی است.

د- دسته‌بندی ارزشیابی‌های آموزشی با توجه به ارزشیابان

۱- ارزشیابی درونی^۳: منظور از ارزشیابی درونی، ارزشیابی توسط کارکنان و دست‌اندرکاران تهیه و اجرای برنامه، بعضی افراد درون برنامه به منظور داوری درباره کیفیت با ارزش مواد آموزشی، روش‌های آموزشی، برنامه‌های درسی یا برنامه‌های آموزشی در ضمن طراحی و اجرای آنها است.

۲- ارزشیابی بیرونی^۴: منظور از ارزشیابی بیرونی، ارزشیابی به‌وسیله اشخاصی خارج از برنامه به منظور داوری درباره کیفیت با ارزش مواد آموزشی، روش‌های آموزشی و نظام‌های آموزشی است، به‌عبارت دیگر ارزشیابی بیرونی یعنی بازبینی و واری فعالیت‌های آموزشگاه و گزارش نتایج آن به‌وسیله کسانی که اعضای سازمان آموزشگاه به حساب نمی‌آیند.

ویژگی‌های ارزشیابی درونی:

- ۱- ارزشیابی درونی معادل ارزشیابی تکوینی می‌باشد.
- ۲- هدف از ارزشیابی درونی کمک به شکل‌گیری، تجدید نظر، اصلاح و بهبود محصول آموزشی است.
- ۳- ارزشیاب، در ارزشیابی درونی شخص طراح یا تولیدکننده برنامه می‌باشد.
- ۴- ارزشیابی درونی را بیشتر به منظور نشان دادن اینکه یادگیرندگان به هدف‌های برنامه رسیده‌اند یا نه، به کار می‌برند.

ویژگی‌های ارزشیابی بیرونی:

- ۱- ارزشیابی بیرونی معادل ارزشیابی تراکمی یا مجموعه‌ای به معنی وسیع کلمه می‌باشد.
- ۲- هدف از ارزشیابی بیرونی این است که معلوم گردد دانش‌آموزانی که با روش خاصی آموزش دیده‌اند، نسبت به دانش‌آموزانی که برنامه دیگری را گذرانده‌اند، چه وضعی دارند.
- ۳- ارزشیاب، در ارزشیابی بیرونی از بیرون برای این منظور گمارده می‌شود و به برنامه مورد ارزشیابی وابسته نیست.

1- School Eveal Uation
2- School Personnal
3- Internal evaluation
4- External evaluation

۴- ارزشیابی بیرونی (تراکمی به معنی وسیع کلمه) به منظور نشان دادن نقاط قوت و ضعف برنامه و تعیین میزان توفیق برنامه، روش یا نظام در رسیدن به هدف‌هایش به کار می‌رود. این ارزشیابی به مصرف‌کننده یک محصول آموزشی می‌گوید که آن محصول آموزشی در چه شرایطی و تا چه میزانی مفید و اثر بخش است.

نکات کلیدی فصل اول

- افلاطون و ارسطو در آثار خود از تفاوت‌های فردی افراد و گروه‌ها سخن گفته‌اند. سه عامل مؤثر در پیدایش سنجش و اندازه‌گیری عبارت است از پیدایش روان‌شناسی تجربی، مطالعه تفاوت‌های فردی و مطالعه عقب‌ماندگی‌های ذهنی.
- فخنر (۱۸۷۷-۱۸۰۱) فیزیکدانی بود که به مسائل روان‌شناسی توجه داشت و در زمینه روان‌شناسی علمی گام‌های مؤثری برداشت.
- وونت اولین آزمایشگاه روان‌شناسی را در لایپزیک آلمان تاسیس و در زمینه حواس مطالعه کرد.
- در ظهور علم تفاوت‌های فردی، برخی ارایه (معادل شخصی) توسط بسل در اثر مطالعه ماجرایی رصدخانه گرینیچ را عامل اصلی می‌دانستند، اما اکثر روان‌شناسان تلاش‌های گالتون در راه‌اندازی آزمایشگاه مردم‌شناسی و کتل در تاسیس آزمایشگاه آزمون‌ها را عامل اصلی ظهور علم تفاوت‌های فردی می‌دانستند.
- گالتون در اثر مشهور خود به نام (نبوغ ارثی) سعی کرده است تا نبوغ را از طریق انحرافی که افراد از میانگین گروه متعلق به خود پیدا می‌کنند، اندازه بگیرد.
- اصطلاح آزمون‌های روانی را کتل مطرح کرد.
- شروع واقعی آزمون‌های روانی با تلاش بینه و سیمون در ساختن آزمون هوشی با هدف تشخیص دانش‌آموزان ناتوان در یادگیری می‌باشد. او مفهوم سن عقلی را بیان کرد. نرمی آزمون هوشی بینه و سیمون را تجدید نظر کرد و اصطلاح بهره هوشی (IQ) را بیان نمود.
- "رابرت یرکز" آزمون ارتشی آلفا و بتا، "کریپلین" آزمون تداعی آزاد اندیشه‌ها، "کارل یونگ" تداعی آزاد کلمات "رابرت وودورت" برگه اصلاحات فردی "والتر دیل اسکات" آزمونی جهت گزینش کارکنان موسسات صنعتی و تجاری را تدوین نمودند.
- اندازه‌گیری، قواعدی است برای اختصاص دادن اعداد به اشیاء یا افراد، به منظور کمی ساختن ویژگی‌های آن‌ها. مراحل اساسی اندازه‌گیری به‌عنوان یک فرآیند عبارت است از: تشخیص و تعریف خصیصه مورد اندازه‌گیری، تعیین رشته عملیات لازم برای آشکارسازی خصیصه مورد اندازه‌گیری و کمی سازی خصیصه به‌صورت واحدهای درجه یا مقدار.
- فرآیند اندازه‌گیری نیاز به وسایلی برای اندازه‌گیری ویژگی‌های فیزیکی و غیرفیزیکی دارد که برای اندازه‌گیری ویژگی‌های فیزیکی از وسایل اندازه‌گیری مستقیم مثل متر، ترازو و برای اندازه‌گیری ویژگی‌های روانی از وسایل اندازه‌گیری غیرمستقیم که مشهورترین آن‌ها آزمون می‌باشد، استفاده می‌شود.
- هر موقع برای اندازه‌گیری ویژگی‌های روانی و تربیتی از آزمون استفاده شود، به آن عمل آزمودن می‌گویند. برخی اوقات برای اندازه‌گیری ویژگی‌های روانی علاوه بر تست یا آزمون یا به‌جای تست از روش‌های دیگری نیز استفاده می‌شود که به این عمل سنجش می‌گویند؛ بنابراین سنجش نوعی اندازه‌گیری است که از آزمون گسترده‌تر است.
- ارزشیابی یک فرآیند نظام‌دار برای جمع‌آوری، تحلیل و تفسیر اطلاعات و قضاوت و داوری ارزشی می‌باشد.
- مفاهیم اساسی حوزه روان‌سنجی از محدودترین مفهوم تا گسترده‌ترین آن عبارتند از:
آزمون > سنجش > اندازه‌گیری > ارزشیابی
- آزمون‌ها و ارزشیابی‌ها را با توجه به انواع ملاک‌ها می‌توان به طبقه‌های متعددی تقسیم نمود.
- آزمون‌ها به‌طور کلی به دو شاخه آزمون‌های توانایی و عاطفی تقسیم می‌شوند. آزمون‌های توانایی (حوزه‌شناختی و روانی - حرکتی) آن‌هایی هستند که در اجرای آن از افراد می‌خواهند تا حداکثر تلاش خود را انجام دهند. آزمون‌های عاطفی وضعیت طبیعی فرد را مورد ارزیابی قرار می‌دهند.

• آزمون‌های توانایی به آزمون‌های استعداد و پیشرفت تحصیلی تقسیم می‌شود. اطلاعات جدول زیر آزمون استعداد در مقایسه با آزمون پیشرفت تحصیلی را نشان می‌دهد.

آزمون استعداد	آزمون پیشرفت تحصیلی
سنجش توانایی‌های بالقوه ناظر بر آینده هستند برای پیش‌بینی موفقیت‌های آتی فرد به کار می‌رود.	سنجش توانایی‌های ناظر به گذشته می‌باشد. برای تعیین آموخته‌های قبلی فرد مورد استفاده قرار می‌گیرد.

• آزمون استعداد به دو دسته، ۱- آزمون استعدادهای ویژه مانند آزمون استعداد موسیقی و ۲- آزمون استعدادهای کلی مانند آزمون هوشی تقسیم می‌شود.

• آزمون پیشرفت تحصیلی را براساس ملاک‌های معینی می‌توان به طبقات مختلفی تقسیم نمود. آزمون‌ها از نظر موضوع به طبقات متعددی چون آزمون ریاضی، آزمون علم و غیره تقسیم می‌شوند. این آزمون‌ها از نظر هدف به دو دسته آزمون‌های پیش‌بینی و تشخیصی، از نظر ماهیت و محتوای ظاهری به آزمون‌های کلامی و غیر کلامی، از نظر دقت و طرز تهیه به آزمون‌های استاندارد شده و معلم ساخته، از نظر مشخصات خارجی به مداد - کاغذی و عملکردی و از نظر نوع به سه دسته آزمون‌های تشریحی، عینی و کوتاه پاسخ تقسیم می‌شوند.

• آزمون‌های عاطفی خصوصیات و ویژگی‌های شخصیتی را ارزیابی می‌کنند و به سه دسته تقسیم می‌شوند: ۱- آزمون علاقه یا رغبت که به علاقه یا احساس فرد نسبت به یک فعالیت اشاره می‌کند. ۲- آزمون نگرشی که نگرش‌سنج نیز نامیده می‌شود که طرز فکر افراد را نسبت به گروه‌ها و زمان‌های اجتماعی اندازه می‌گیرد. ۳- آزمون شخصیت که شخصیت (مجموعه رفتارهای فرد در موقعیت‌های اجتماعی) را می‌سنجد و چهار نوع اساسی دارد: انگیزش، سازگاری، منش و مزاج یا خوی، ارزشیابی‌های آموزشی را نیز با توجه به چهار ملاک تقسیم‌بندی می‌کنند.

• ارزشیابی‌های آموزشی با توجه به ملاک مورد استفاده به دو دسته ارزشیابی وابسته به ملاک و ارزشیابی وابسته به هنجار تقسیم می‌شوند. نوع ملاک مورد سنجش در ارزشیابی وابسته به ملاک، ملاک مطلق می‌باشد و در ارزشیابی وابسته به هنجار، ملاک نسبی می‌باشد.

• ارزشیابی‌های آموزشی را با توجه به زمان و هدف استفاده از آن‌ها به چهار دسته تقسیم می‌کنیم:

۱- ارزشیابی آغازین شامل رفتار ورودی و پیش‌آزمون می‌باشد. رفتار ورودی به آمادگی فرد برای یادگیری هدف‌های آموزشی مربوط است. پیش‌آزمون همان آزمون جایابی یا همان آزمون نهایی درس می‌باشد که قبل از شروع درس برگزار می‌شود تا میزان اطلاعات یادگیرندگان از مطالب درس جدید، قبل از تدریس مشخص شود.

۲- ارزشیابی تکوینی در جریان آموزش اجرا می‌شود و هدف اصلی آن اصلاح روش تدریس معلم و تعیین نقاط قوت و ضعف یادگیری دانش‌آموزان است.

۳- ارزشیابی تشخیصی که از نوع تکوینی است عمدتاً در جریان آموزش برگزار می‌شود و هدف اصلی آن کشف مشکلات یادگیری فراگیران است.

۴- ارزشیابی تراکمی (پایانی) در پایان فرآیند یا دوره آموزشی اجرا می‌شود و هدف از آن نمره دادن به فراگیران و قضاوت درباره عملکرد آن‌ها و نیز قضاوت درباره اثر بخشی آموزشی معلم است.

• در انواع ارزشیابی‌های آموزشی مواد و سوالات امتحانی تقریباً یکسان هستند و آنچه که آن‌ها را از هم متمایز می‌سازد، هدف از اجرای آزمون است.

• ارزشیابی‌های آموزشی نیز براساس موضوع شامل پنج دسته ارزشیابی از فراگیران، ارزشیابی از برنامه درسی و مواد آموزشی و ارزشیابی از پروژه‌ها و برنامه‌های آموزشی، ارزشیابی از آموزشگاه و ارزشیابی از کارکنان آموزشگاه می‌باشد.

• ارزشیابی‌های آموزشی نیز با توجه به ارزشیابان به ارزشیابی بیرونی و درونی تقسیم می‌شوند. در ارزشیابی درونی طراحان و مجریان به ارزشیابی از برنامه‌های خود اقدام می‌کنند، اما در ارزشیابی بیرونی افراد خارج از برنامه به ارزشیابی‌های برنامه‌ها اقدام می‌کنند.