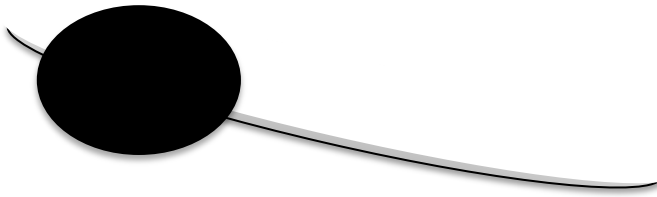


آزمون و آزمون‌سازی در
تحقیقات تربیتی و روان‌شناسی
(مبانی و روش‌های ساخت و انطباق‌یابی پرسشنامه‌ها)



فهرست

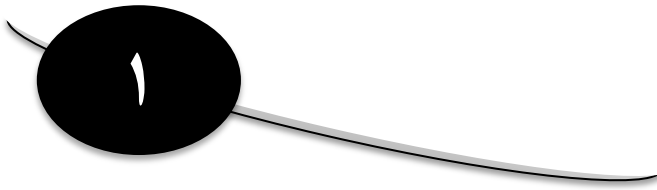
پیشگفتار.....	۷
فصل ۱. کلیات و مبانی.....	۹
فصل ۲. طبقه‌بندی آزمون‌ها.....	۲۱
فصل ۳. کاربرد و ساخت آزمون در پژوهش و عمل.....	۲۷
فصل ۴. انطباق‌یابی آزمون.....	۳۹
فصل ۵. اصول طراحی سؤالات آزمون (آیتم‌ها).....	۴۷
فصل ۶. روایی آزمون.....	۵۹
فصل ۷. تحلیل عاملی: اصول و روش‌ها.....	۶۹
فصل ۸. پایایی آزمون.....	۹۵
فصل ۹. تحلیل آیتم در ساخت و انطباق‌یابی آزمون.....	۱۱۵
فصل ۱۰. هنجاریابی آزمون.....	۱۳۵
فصل ۱۱. نمرات برش در اندازه‌گیری‌های روانی و تربیتی.....	۱۴۷
فصل ۱۲. سوگیری آیتم و آزمون.....	۱۵۹
فصل ۱۳. نظریه‌های اندازه‌گیری کلاسیک و سؤال - پاسخ.....	۱۶۷
فصل ۱۴. فرارزشیابی.....	۱۸۹
منابع.....	۲۰۳
پیوست.....	۲۰۷

سنجش و اندازه‌گیری پدیده‌ها یا متغیرها و ساخت پرسشنامه یا آزمون از جمله فعالیت‌های اساسی پژوهشگران و فعالان حوزه‌های علوم انسانی به‌خصوص روان‌شناسی و تعلیم و تربیت است. پیچیدگی پدیده‌ها در این حیطه سنجش و اندازه‌گیری را نیز پیچیده‌تر کرده است. لذا ساخت ابزارها در این حیطه نیازمند تسلط بر مبانی نظری مربوطه و نیز تسلط بر روش‌ها و الگوهای اندازه‌گیری مختلف و همچنین ذهن نقاد، خلاق و جستجوگر پژوهشگر است. کتاب‌ها و مقاله‌های مختلفی در این حیطه نوشته شده است و هر کدام از منظری به مهارت‌ها و روش‌های مربوطه پرداخته‌اند. با وجود چنین منابعی، جای خالی کتابی که بتواند راهنمای ساده ساخت و *انطباق‌یابی* آزمون‌های روانی و تربیتی باشد و بتواند روش‌ها و رویکردهای جدیدی هم برای پژوهشگران و عملگران این حیطه‌ها، ارائه دهد، احساس می‌شود. برجستگی این کتاب تهیه راهنمای ساده کاربرد، ساخت و به‌خصوص انطباق‌یابی آزمون است. ارائه روش‌های عینی‌تری برای احراز روایی محتوایی نیز از دیگر ویژگی‌های این کتاب است که در سایر آثار کمتر به چشم می‌خورد. علاوه بر این اختصاص بخش‌هایی از کتاب به تعیین نقاط برش، سوگیری آزمون و آیتم و فرارزشیابی هم از نقاط قوت کتاب حاضر است که در ساخت و انطباق‌یابی آزمون‌ها و پرسشنامه‌ها کاربرد گسترده‌ای باید داشته باشد. امید است اثر حاضر مورد استفاده پژوهشگران علاقه‌مند به ساخت آزمون یا پرسشنامه در حوزه‌های مختلف علوم انسانی قرار گیرد. علاوه بر این این کتاب می‌تواند به‌عنوان منبع درسی و کمک‌درسی دروس سنجش و اندازه‌گیری در علوم تربیتی و روان‌سنجی در دوره‌های کارشناسی و کارشناسی ارشد قرار گیرد. پیشنهادها و دیدگاه‌های اساتید و دانشجویان در بهبود کمی و کیفی کتاب مؤثر خواهد بود. پیشاپیش از همکاری و اشتراک دیدگاه‌های همکاران عزیزم سپاسگزارم.

دکتر حسین کارشکی

دانشیار روان‌شناسی تربیتی دانشگاه فردوسی مشهد

kareshki@um.ac.ir



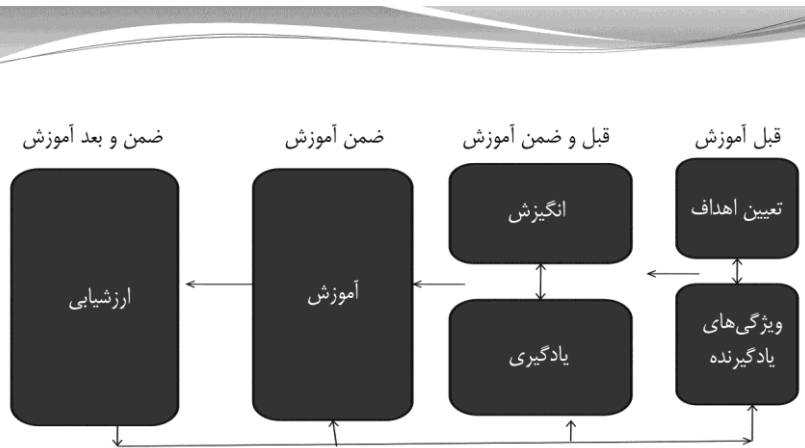
کلیات و مبانی

جایگاه سنجش و اندازه‌گیری در روان‌شناسی و تعلیم و تربیت

اندازه‌گیری از زمان‌های اولیه پیدایش بشر و پیدایش علوم مختلف مورد استفاده بوده و یکی از دغدغه‌های مردم و دانشمندان و پژوهشگران، در مبادلات گسترده آنها تعیین اندازه ویژگی‌های مورد مبادله بوده است. ضرورت توجه به اندازه‌های دقیق‌تر در حال حاضر بیشتر هم شده است. امروزه، در فیزیک، شیمی، زیست‌شناسی، علوم انسانی و سایر علوم و فنون، کاربرد گسترده‌ای دارد. ارائه گزارش پژوهش یا خدمات علمی بدون اندازه‌گیری مقدر نیست. به‌عنوان مثال، برای اندازه‌گیری دما، جرم، وزن، نم، طول و ضخامت مواد و پدیده‌های مختلف، از گذشته‌های دور، ابزارهای کمابیش دقیقی استفاده می‌شده که امروزه نسبت به گذشته خیلی تکامل یافته است.

به‌جز علوم فیزیکی، اندازه‌گیری، در روان‌شناسی و تعلیم و تربیت نیز کاربرد گسترده‌ای داشته و خواهد داشت. در موقعیت روانی و تربیتی، روان‌شناس یا معلم باید تشخیص دهد و تصمیم‌گیری کند و لازمه تشخیص و تصمیم‌گیری درست، اندازه‌گیری‌های دقیق پدیده‌هاست. در یک موقعیت آموزشی، فرایندها و فرآورده‌های زیادی وجود دارند که باید مورد ارزیابی قرار گیرند. برای تعیین اهمیت اندازه‌گیری، معرفی مدل عمومی آموزش کمک‌کننده است. شکل (۱-۱)، مدل عمومی آموزشی (گیج و برلاینر، ۱۹۸۶، ۵۷) را با اندک تغییراتی نشان می‌دهد.

در این مدل، قبل از آموزش هدف‌های آموزشی تعیین می‌شود و ویژگی‌های یادگیرنده مشخص می‌گردد، قبل و یا ضمن آموزش، فرایندهای انگیزشی و یادگیری مورد بررسی و مداخله قرار می‌گیرد و مطابق هدف‌ها، یادگیرندگان و فرایندهای انگیزشی و یادگیری، آموزش ارائه می‌گردد و نهایتاً پس از آموزش، نتیجه آموزش مورد ارزیابی قرار می‌گیرد و نتیجه اندازه‌گیری در ارزشیابی عملکرد یادگیرندگان و معلم مورد استفاده قرار می‌گیرد. نقص در هر مرحله مشکلاتی را در روند یادگیری و آموزش ایجاد می‌کند. تمام فرایندها و فرآورده‌های این چرخه یادگیری و آموزش باید مورد ارزشیابی قرار گیرند. لذا می‌توان گفت ارزشیابی در تمام مراحل آموزش حضور دارد و فقط در انتهای آموزش و



شکل ۱-۱. فرایندهای درگیر در آموزش برحسب مدل عمومی آموزش (گیج و برلایتر، ۱۹۸۶، ص ۵۷)

به‌عنوان فقط هدف نهایی محسوب نمی‌شود. لذا ارزشیابی باید در خدمت آموزش باشد. برخلاف تصور رایج که آموزش و یادگیری باید در خدمت ارزشیابی باشد، ارزشیابی باید در خدمت آموزش و یادگیری باشد. در موقعیت‌های مدیریتی و سازمانی نیز، نیاز به اندازه‌گیری و تعیین اثربخشی سازمانی، کارمندان و مدیران وجود دارد.

به‌جز موقعیت‌های آموزشی که فرایندهای روانی و شناختی مختلفی اندازه‌گیری می‌شود، روان‌شناسان نیز برای انجام فعالیت‌های حرفه‌ای خود نیازمند ارزیابی جامع و دقیق مراجعان خود هستند تا تشخیص و در نتیجه تصمیم‌گیری‌ها و مداخلات صحیحی را انجام دهند. بخش قابل توجهی از فعالیت‌های روان‌شناس در جلسات ابتدایی و انتهایی و حتی میانی درمان، صرف ارزیابی مراجع با استفاده از فنون مختلف سنجش می‌شود. این امر با استفاده از مصاحبه، مشاهده، آزمودن، پرسشنامه، آزمون و سایر ابزارهای اندازه‌گیری صورت می‌گیرد. با توجه به اهمیت سنجش^۱ و اندازه‌گیری^۲ در موقعیت‌های روانی و تربیتی، آشنایی با مفاهیم و روش‌های آن ضروری است. نقطه شروع این آشنایی با مفاهیم و سازه‌های نزدیک به آن خواهد بود.

مفاهیم مورد استفاده در اندازه‌گیری‌های روانی و تربیتی

قبل از ورود به سنجش و اندازه‌گیری در روان‌شناسی و تعلیم و تربیت لازم است، تفاوت‌ها و شباهت‌های سنجش و اندازه‌گیری با هم و با سایر اصطلاحات مشابه مورد بررسی قرار گیرد. آزمون،

1- assessment
2- measurement

آزمودن، ارزشیابی و پژوهش، مهم‌ترین واژه‌های مشابه سنجش و اندازه‌گیری هستند که در بخش‌های بعدی به آنها پرداخته خواهد شد.

کلی‌ترین و فراگیرترین واژه این حوزه، اندازه‌گیری است. اندازه‌گیری یعنی تعیین مقدار ویژگی‌ها یا صفات اشیا و افراد. با این تصور از این واژه، می‌توان گفت در سایر حیطه‌های علمی و غیرعلمی نیز کم‌وبیش، این عمل، مورد استفاده قرار می‌گیرد. برای مثال نمونه‌هایی از اندازه‌گیری عبارت‌اند از: "قد علی، ۱۷۲ سانتیمتر است، وزن زهرا، ۵۵ کیلوگرم است، نمره افسردگی علی، ۱۷ است، نمره ریاضی هاشم، ۱۸/۷۵ است، طبقه اجتماعی فاطمه بالاست و بالاخره، مدرک تحصیلی آرزو، کارشناسی ارشد است. همه اینها با ابزارهای خاص خود انجام شده و معرف کمیت یک ویژگی است. گروترو و وان-در منبع^۱ (۲۰۰۸، ص ۲) تعریفی از استیونس در مورد اندازه‌گیری نقل می‌کنند که عدد و قاعده در آن برجسته است "اختصاص اعداد به اشیا یا رویدادها بر طبق قواعدی!" همچنین، ورثن و ساندرز (۱۹۸۷، ص ۲۳) معتقدند اندازه‌گیری، وصف کمی رفتار، اشیا و وقایع و تعیین کمیت یک پدیده است. داشتن قواعد به یکسانی اجرا در موارد متعدد کمک می‌کند و از جنبه‌های مهم استاندارد کردن یا میزان کردن است. با مشاهده مثال‌های یادشده، می‌توان گفت برخی از اندازه‌گیری‌ها، دقیق‌تر و برخی کلی‌تر هستند. برخی اندازه‌گیری‌ها ساده‌اند و برخی پیچیده، برخی میزان اندازه به دست آمده آنها بدون بحث و اختلاف نظر میزان آن مشخص می‌گردد، ولی در برخی موارد، میزان آنها قابل بحث‌اند؛ برخی مربوط به ویژگی‌های مادی و جسمانی و برخی مربوط به ویژگی‌های روانی و تربیتی هستند. اندازه‌گیری قد و وزن، اندازه‌گیری‌های مادی هستند، ولی اندازه‌گیری‌های هوش و شخصیت، اندازه‌گیری‌های روانی و تربیتی محسوب می‌گردند. دسته دوم یعنی اندازه‌گیری‌های روانی و تربیتی مدنظر و مورد تأکید ماست. لذا تأکید بر واژه‌های آزمودن و سنجش، محور این کتاب است که نوع خاصی از اندازه‌گیری است. اندازه‌گیری قد و وزن دقیق است و مستقیم و با ابزارهایی همچون متر و ترازو صورت می‌گیرد. اما اندازه‌گیری‌های روانی و تربیتی غیرمستقیم و با ابزارهای گوناگون صورت می‌گیرد که گاهی مکمل یا حتی متناقض همدیگر هستند. ممکن است نتایج اجرای آزمون هوش روی کودک با نتایج مشاهده عملکرد او توسط دیگران متناقض یا مکمل همدیگر باشند.

آزمون و آزمودن. اندازه‌گیری‌های فیزیکی با ابزاری چون ترازو و متر، جدا از دقتی دارند، مستقیماً صورت می‌گیرد. علاوه بر اینها برخی ویژگی رفتاری همچون تعداد جابجایی کودک در کلاس یا تعداد دست بلند کردن او مستقیماً قابل اندازه‌گیری است. متفاوت یا حتی برخلاف اندازه‌گیری‌های فیزیکی و اندازه‌گیری برخی ویژگی رفتاری که آشکارند، ویژگی‌ها یا صفات روانی و تربیتی مانند اضطراب، افسردگی، هوش، خلاقیت و انگیزش باید به‌طور غیرمستقیم اندازه‌گیری شوند. معمول‌ترین وسیله

اندازه‌گیری ویژگی‌ها یا صفات روانی و تربیتی آزمون یا پرسشنامه است. از آنجاکه فرد مستقیماً نمی‌تواند ویژگی‌ها یا صفات روانی و تربیتی مورد نظر را اندازه‌گیری کند، ناچار است به رفتارها یا نشانگرهای معرف آن ویژگی مراجعه کند و با توجه به آنها اندازه‌گیری کند. نشانگرهای معرف رفتارها مورد نظر زیادند، لذا در یک آزمون نمی‌گنجند و باید نمونه‌گیری شوند و نمونه معرف آن رفتارها انتخاب گردد. به همین دلیل است که گرانلاند و لین (۱۹۹۰)، به نقل از سیف، (۱۳۷۶، ص ۲۹) معتقدند *آزمون* وسیله‌ای نظام‌دار برای اندازه‌گیری نمونه‌ای از رفتار است. همچنین گروتر و وان در کمپ (۲۰۰۸، ص ۱) در تعریفی کامل‌تر بیان می‌کنند که آزمون، روندی^۱ برای نمونه‌گیری از رفتار و توصیف آن با طبقات و نمرات است. این تعریف به مفهوم آزمودن هم نزدیک است. هرگاه از آزمون برای اندازه‌گیری ویژگی‌های روانی و تربیتی استفاده شود، این نوع اندازه‌گیری، *آزمودن* نام دارد. هرگاه برای اندازه‌گیری هوش از آزمون هوش استفاده شود و نمره کلی هوش یا نمرات خرده مؤلفه‌های هوش به دست آید، آزمودن صورت گرفته است. با توجه به آنچه گفته شد می‌توان گفت اندازه‌گیری، گسترده‌تر از آزمودن است و در واقع، آزمودن نوع خاصی از اندازه‌گیری است که برای اندازه‌گیری غیرمستقیم رفتار انسانی مورد استفاده قرار می‌گیرد.

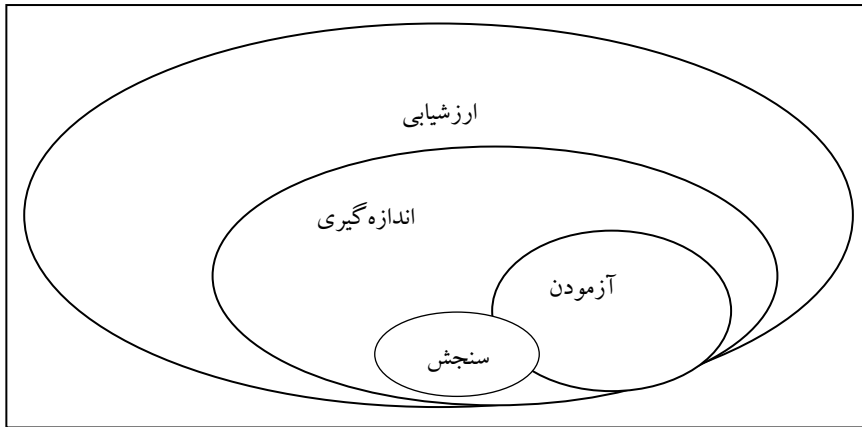
سنجش. اندازه‌گیری حاصل از آزمون گرچه مهم و قابل توجه است، اما همیشه دقیق و باکیفیت نیست. نمرات آزمون هوش، همیشه هوشمندی فرد را نشان نمی‌دهند و یا نمرات امتحانی دانش‌آموزان همیشه معرف توان و دانش فرد نیستند. به همین خاطر است که نمرات آنها همیشه مورد اعتراض دانش‌آموزان و یادگیرندگان و یا حتی خود معلم است. لذا به روش‌های اندازه‌گیری جامع‌تری در اندازه‌گیری‌های روانی و تربیتی نیازمندیم. هرگاه علاوه بر آزمون- یا بدون آزمون، از وسایل دیگری برای اندازه‌گیری همه‌جانبه استفاده گردد، **سنجش** صورت گرفته است. سنجش به اعتقاد ویگینز (۱۹۹۳)، به نقل از لطف‌آبادی، (۱۳۷۴، ص ۳۸) یک تحلیل جامع و چندوجهی از عملکرد است. در سنجش ممکن است از آزمودن هم استفاده شود؛ ولی هر آزمودنی، سنجش نیست. در مورد تفاوت و شباهت سنجش و اندازه‌گیری باید گفت سنجش نوعی اندازه‌گیری است که به اندازه‌گیری عمیق و همه‌جانبه فرد می‌پردازد. سنجش در موقعیت‌های روانی و تربیتی کاربرد گسترده‌ای دارد و در تشخیص و تصمیم‌گیری کمک قابل توجهی می‌کند. گزینش در سازمان‌ها که شامل امتحان کتبی، مصاحبه، بررسی‌های میدانی و اسنادی است، نوعی سنجش محسوب می‌گردد. چراکه به بررسی عمیق و همه‌جانبه فرد برای اختصاص جایگاه در سازمان پرداخته می‌شود.

ارزشیابی. واژه دیگر مرتبط با اندازه‌گیری، ارزشیابی است. ارزشیابی یعنی تعیین ارزش چیزی. پس ابتدا باید مقدار چیزی مشخص گردد (اندازه‌گیری شود)، بعد در مورد این مقدار قضاوت ارزشی گردد. وقتی این امر در مورد پدیده‌های آموزشی صورت گیرد نام آن ارزشیابی آموزشی خواهد بود. این مفهوم به نظر بازرگان (۱۳۸۰، ص ۱۷) در منابع متعددی با قضاوت درباره شایستگی پدیده‌های آموزشی مترادف بوده است. به نحو قابل قیاسی، وقتی که ارزشیابی در مورد پدیده‌های روانی همچون شادکامی، افسردگی یا سلامت روانی باشد، نام آن ارزشیابی روانی یا روان‌شناسی خواهد بود.

نمره دانش‌آموز در درس ریاضی ابتدا باید دقیق اندازه‌گیری شود و بعد در این مورد قضاوت شود که آیا عملکرد او خوب بوده یا ضعیف عمل کرده است. پس مقدمه ارزشیابی‌های صحیح، اندازه‌گیری‌های دقیق است. این اندازه‌گیری مبنای ممکن است از طریق آزمودن، سنجش یا هر نوع اندازه‌گیری دیگر باشد. مقایسه وزن و قد کودک با منحنی طبیعی رشد یا مقایسه عملکرد ریاضی دانش‌آموز یا دانشجو با ملاک ده (۱۰)، نوعی ارزشیابی محسوب می‌گردد. پس می‌توان گفت ارزشیابی علاوه بر اندازه‌گیری، معیار یا ملاکی برای تصمیم‌گیری و قضاوت نیز نیاز دارد.

مقایسه واژه‌ها و مصادیق اندازه‌گیری، آزمودن، سنجش و ارزشیابی، اطلاعات ارزشمندی برای تفکیک اینها از همدیگر در اختیار قرار می‌دهد. سیف (۱۳۷۶، ص ۳۴) رابطه اندازه‌گیری، آزمودن، سنجش و ارزشیابی را با استفاده از یک نمودار نشان داده است (شکل ۲-۱). می‌توان گفت که اندازه‌گیری واژه کلی‌تری است که هر نوع کمی‌سازی و ویژگی‌های فیزیکی و رفتاری را شامل می‌شود و در واقع آزمودن و سنجش نوعی اندازه‌گیری محسوب می‌گردند. پس رابطه اندازه‌گیری با آزمودن و سنجش نوعی رابطه کل و جز است. اما در مورد رابطه ارزشیابی با بقیه مفاهیم (اندازه‌گیری، آزمودن و سنجش)، می‌توان گفت مبنای ارزشیابی درست، اندازه‌گیری‌های دقیق است؛ حال این اندازه‌گیری مستقیم، آزمودن یا سنجش باشد. ارزشیابی به‌جز اندازه‌گیری، ملاک ارزشیابی هم دارد که ارزشیاب بر اساس آن تصمیم می‌گیرد. دادن نمره دقیق به دانش‌آموز نوعی اندازه‌گیری است، ولی اینکه این فرد قبول است یا رد، به ملاک قبولی مربوط است که در ارزشیابی باید از قبل مشخص باشد.

متفاوت از ایده مبنایی شکل (شکل ۲-۱) برای مقایسه، می‌توان گفت رابطه ارزشیابی با اندازه‌گیری، آزمودن و سنجش، برخلاف رابطه اندازه‌گیری با آزمودن و سنجش رابطه کل جز نیست، بلکه در داشتن ملاک تصمیم‌گیری است؛ ارزشیابی دارای ملاک ارزشیابی و تصمیم‌گیری است، ولی سایر واژه‌های اشاره شده و به‌طورکلی، اندازه‌گیری، فاقد آن است. بنابراین می‌توان شکل (۲-۱) را به نحو زیر اصلاح کرد (شکل ۳-۱). بخشی از اندازه‌گیری، ارزشیابی محسوب نمی‌گردد (بخشی که با هدف‌های دیگری جز ارزشیابی یا مثلاً برای تحقیق اندازه‌گیری شده است). بخشی از ارزشیابی هم اندازه‌گیری محسوب نمی‌شود (ملاک ارزشیابی). اما ارزشیابی همیشه مبتنی بر اندازه‌گیری‌های دقیق



شکل ۱-۲. رابطه آزمون، سنجش، اندازه‌گیری و ارزشیابی (سیف، ۱۳۷۶، ص ۳۴)

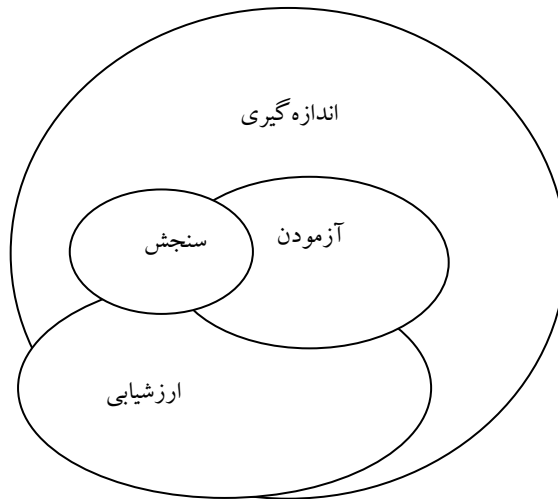
است که بخش مشترک اندازه‌گیری و ارزشیابی، همین اندازه‌گیری‌های مبنای ارزشیابی است. بخشی از ارزشیابی هم مبتنی بر سنجش و آزمون نیست، بلکه مبتنی بر اندازه‌گیری‌های فیزیکی است.

اندازه‌گیری و پژوهش. پژوهش‌گر با متغیر سروکار دارد و به اعتقاد کرینجر (۱۹۸۶، ترجمه شریفی، ۱۳۷۴، ص ۴۶۱) کار پژوهشگر در طرح پژوهش پیشنهادی‌اش، کنترل واریانس متغیرهاست. اگر چنین نقشی را برای پژوهش‌گر قائل باشیم، کنترل و بررسی متغیرها بدون اندازه‌گیری امکان ندارد. لذا می‌توان گفت اساس و مبنای پژوهش به‌خصوص پژوهش‌های کمی، اندازه‌گیری‌های دقیق متغیرهاست. همان نقشی که در ارزشیابی برای اندازه‌گیری مطرح کردیم، برای پژوهش نیز قائل هستیم. اگر هدف محقق بررسی "**رابطه شادکامی و هدفمندی در زندگی**" است، لازمه انجام چنین پژوهشی اندازه‌گیری دقیق میزان **شادکامی و هدفمندی** شرکت‌کنندگان است. اگر این اندازه‌گیری دقیق نباشد، بهترین طرح تحقیق و بهترین روش آماری هم نمی‌تواند توصیف دقیقی از **رابطه شادکامی و هدفمندی** فراهم کند. لذا می‌توان گفت اندازه‌گیری‌های دقیق، اساس پژوهش‌های معتبر است.

مقیاس‌های اندازه‌گیری

در اندازه‌گیری یا ارزشیابی، ارزشیاب، پژوهشگر یا به عبارتی کلی‌تر اندازه‌گیر، با کمی سازی متغیرها سروکار دارد و به ویژگی‌ها، عدد اختصاص می‌دهد. در این فرایند مفهوم مقیاس اندازه‌گیری^۱ یا سطح اندازه‌گیری حائز اهمیت است. سطح اندازه‌گیری بیان می‌کند که معنای عدد در مورد این ویژگی به چه

1 - measurement scale



شکل ۳-۱. رابطه آزمون، سنجش، اندازه‌گیری و ارزشیابی

معناست؛ یا به عبارتی دیگر مشخص می‌کند عددهای به کار رفته در مورد آن ویژگی تا چه حد دقیق هستند. عددی که برای اندازه‌گیری قد به کار می‌رود (اعداد ۱۵۰، ۱۷۰/۶۵ و ۱۹۵/۲۱) معنای متفاوتی نسبت به اندازه‌گیری جنسیت (۰ برای زنان و ۱ برای مردان) یا رتبه کلاسی (اول، دوم و بیستم) دارد. در مباحث سنجش و اندازه‌گیری و بخش‌های مقدماتی کتب معروف آماری، چهار سطح یا مقیاس اندازه‌گیری پیشنهاد شده است: *طبقه‌ای (اسمی یا مرتب)*، *رتبه‌ای*، *فاصله‌ای* و *نسبی*. سطح اندازه‌گیری متغیر در انتخاب آزمون و روش آماری نقش مهمی دارد. جدول (۱-۱)، ویژگی‌های تطبیقی مقیاس‌های اندازه‌گیری را نشان می‌دهد.

سطح اندازه‌گیری طبقه‌ای. ساده‌ترین روش برای اندازه‌گیری متغیرها، تعیین طبقات آنها است. در این روش، پدیده‌ها، افراد و اشیاء در طبقات مجزا، قرار می‌گیرد و هر طبقه با یک کد که ارزش ریاضیاتی ندارد، مشخص می‌گردد. اختصاص شماره به بازیکنان فوتبال، اختصاص عدد به جنسیت، یا طبقه‌بندی انواع لوبیا و اختصاص عدد به هر طبقه نمونه‌ای از اندازه‌گیری‌های ساده و ابتدایی است. در این نوع اندازه‌گیری‌ها، اعداد ماهیت ریاضیاتی خود را ندارند و فقط برای طبقه‌بندی استفاده می‌شوند. در این سطح از اندازه‌گیری که اندازه‌گیر، فقط طبقه‌بندی می‌کند، گاهی طبقات نسبت به همدیگر ارجحیت یا برتری ندارند (مثل طبقات جنسیت یا شهرهای محل سکونت آزمودنی‌ها) و گاهی طبقات نسبت به همدیگر برتر دارند و سلسله مراتبی از کم به زیاد یا برعکس در طبقات قابل مشاهده است (مانند طبقات هوشی یا طبقات اقتصادی بالا، متوسط و پایین). دسته اول *طبقه‌ای/اسمی* نام دارد و دسته دوم

که سلسله مراتبی از طبقات از کم به زیاد یا برعکس است، *طبقه‌ای مرتب نام دارد*. یکی از ویژگی‌های سطوح اندازه‌گیری، قابلیت آنها برای انجام محاسبات آماری و ریاضی است. با توجه به قابلیت‌های ریاضی پایین این سطح، عملیات ریاضی و آماری محدودی با آن انجام می‌شود. عملیات آماری قابل استفاده در این مقیاس، شامل شمارش، نسبت، درصد و نماست. هیچ عملیات ریاضی در این مقیاس قابل اعمال و محاسبه نیست. نمی‌توان گفت اگر شهر مشهد با تهران جمع شود، می‌شود شهر X. یا مربی بازیکن شماره ۱۱ را از بازی خارج می‌کند و به جای آن بازیکن شماره ۶ و ۵ را با هم وارد بازی می‌کند.

سطح اندازه‌گیری رتبه‌ای. گرچه سطح اندازه‌گیری طبقه‌ای یگانه روش برای اندازه‌گیری برخی متغیرهاست، اما برای اندازه‌گیری برخی متغیرها، سطح دقیق‌تری هم امکان‌پذیر است که باید از آن استفاده کرد. اشکال سطح اندازه‌گیری طبقه‌ای این است که برتری افراد داخل یک طبقه نسبت به هم را نشان نمی‌دهد و متأثر از *خطای طبقه‌بندی* است. در یک مجموعه ۱۰ نفری، راه ساده‌تر این است که آنها را دو دسته بلندقد و کوتاه‌قد تقسیم کنیم؛ اما اشکال این است که علیرغم تفاوت‌هایی که افراد هر گروه نسبت به هم گروهی‌های خود دارند، بلندقدها، در اندازه قد، همه مانند هم و کوتاه‌قدها مانند هم تصور شده‌اند. روش دقیق‌تر اندازه‌گیری این است که افراد را برحسب قد مرتب کنیم و به هرکسی یک رتبه از اول تا دهم، اختصاص دهیم (تعداد افراد و رتبه‌ها باید با هم برابر باشند، حتی اگر قد یکسان هم داشته باشیم که در این صورت باید رتبه یکسان به آنها بدهیم ولی در هر صورت، تعداد رتبه‌ها باید ۱۰ تا باشد). این سطح، سطح اندازه‌گیری *رتبه‌ای یا ترتیبی* نام دارد. برتری این روش نسبت به سطح طبقه‌ای این است که امکان بررسی تفاوت‌های فردی کلیه اعضا نسبت به هم و تعیین جایگاه افراد در گروه، فراهم می‌شود. باین وجود، می‌توان گفت اشکالاتی هم دارد. فاصله رتبه‌ها با هم برابر فرض شده است، اما عملاً ممکن است برابر نباشد. در یک کلاس فاصله بین رتبه‌ها با هم یکسان فرض می‌شود، ولی در عمل ممکن است بیشتر یا کمتر از یک واحد اختلاف داشته باشند. فاصله رتبه اول با دوم و فاصله رتبه سیزدهم تا چهاردهم یکی فرض می‌شود، در صورتی که ممکن است یکی ۱۰ واحد با دیگری اختلاف داشته باشد و دیگری ۰/۰۰۱ واحد. این سطح با سطح اندازه‌گیری طبقه‌ای مرتب متفاوت است، چراکه اختلاف تک تک افراد با هم مورد توجه قرار می‌گیرد و نه برتری طبقاتی یک طبقه با طبقه دیگر.

با اندازه‌متغیر دارای سطح اندازه‌گیری رتبه‌ای، می‌توان عملیات ریاضی محدودی چون کوچک‌تر، مساوی و بزرگ‌تر را انجام داد. عمل جمع، تفریق، ضرب و تقسیم را نمی‌توان با رتبه‌ها انجام داد. نمی‌توان گفت رتبه اول کلاس دو برابر رتبه دوم می‌داند یا برعکس. با متغیر دارای این سطح اندازه‌گیری، عملیات آماری محدودی مانند ضریب همبستگی رتبه‌ای اسپیرمن می‌توان انجام داد.

سطح اندازه‌گیری فاصله‌ای. این سطح اندازه‌گیری دقیق‌تر از دو سطح قبلی است. ضمن اینکه مزایا و قابلیت‌های دو سطح قبلی را دارد، قابلیت‌های بیشتری هم نسبت به آن دو سطح دارد. اشکال اول سطح اندازه‌گیری رتبه‌ای این است که علیرغم اینکه فاصله واحدها برابر نیست، ولی برابر فرض می‌شود. اشکال دیگر آن این است که نقطه صفر ندارد که معیاری برای مقایسه با آن فراهم شود و مقایسه‌های بین گروهی و بین فردی امکان‌پذیر گردد. بنابراین متغیر دارای سطح اندازه‌گیری فاصله‌ای دارای صفر قراردادی است و فاصله بین واحدهای آن هم یکسان است. نمونه بارز سطح اندازه‌گیری فاصله‌ای، دمای هوا برحسب سانتی‌گراد است. حرکت جیوه در داخل لوله شیشه‌ای در شرایط انجماد آب (یخ زدن) تا نقطه بخار آب، به صد واحد برابر تقسیم شده است. نقطه یخ زدن، صفر قراردادی است و نقطه بخار عدد ۱۰۰ قراردادی است. متغیرهای روان‌شناختی و تعلیم و تربیتی که با نمره نشان داده می‌شوند، دارای سطح اندازه‌گیری فاصله‌ای هستند. نمره صفر در درس ریاضی، یک صفر مطلق نیست؛ بلکه صفری قراردادی است. صفر گرفتن در یک امتحان ریاضی به معنای فقدان توانایی ریاضی نیست. بلکه این آزمون نتوانسته برخی ویژگی‌های ریاضی را اندازه‌گیری کند. با توجه به قابلیت‌های سطح اندازه‌گیری فاصله‌ای، تمام عملیات آماری و عملیات ریاضی جمع و تفریق را می‌توان با اندازه‌های متغیرهای مربوطه انجام داد. اما عمل ضرب و تقسیم را نمی‌توان با متغیرهای فاصله‌ای انجام داد. نمی‌توان گفت دمای تهران (۱۵)، سه برابر گرم‌تر از دمای تبریز (۵) است. چون وجود صفر قراردادی امکان چنین مقایسه‌ای را از بین برده است. لذا می‌توان گفت عمده‌ترین مشکل سطح اندازه‌گیری فاصله‌ای [وجود صفر قراردادی و] فقدان صفر مطلق است.

سطح اندازه‌گیری نسبی. دقیق‌ترین سطح اندازه‌گیری، سطح نسبی است. در این سطح، صفر مطلق وجود دارد و صفر به معنای عدم است و فقط قراردادی نیست. متغیرهای فیزیکی و کمیت‌های تربیت‌بدنی مانند قد، وزن، موجودی حساب، تعداد بارفیکس و دراز نشستی که فرد می‌تواند انجام دهد، دارای سطح اندازه‌گیری نسبی هستند. این سطح اندازه‌گیری، تمام مزایای سایر سطوح قبلی اندازه‌گیری را دارد. علاوه بر این دارای صفر مطلق است. به همین دلیل تمام عملیاتی ریاضی و آماری را می‌توان با چنین متغیرهایی انجام داد.

تبدیل مقیاس‌ها. تبدیل مقیاس‌ها یا سطوح اندازه‌گیری، از سطوح دقیق‌تر (مثل نسبی) به غیردقیق‌تر (مثل رتبه‌ای) امکان‌پذیر است. اما از سطوح غیردقیق به دقیق امکان‌پذیر نیست. می‌توان برحسب اعداد قد که سطح نسبی است، رتبه قد و یا طبقه قد را به دست آورد، اما با داشتن طبقه قد، نمی‌توان رتبه قد افراد یا میزان دقیق قد آنها را به دست آورد. جدول (۱-۱)، امکان تطبیق مقیاس‌های اندازه‌گیری به همدیگر را نشان می‌دهد.