



ریاست جمهوری

سازمان برنامه و بودجه کشور

سازمان مدیریت و برنامه ریزی استان چهارمحال و بختیاری
مدیریت آموزش و پژوهش های توسعه و آینده نگری

برنامه آموزشی احراز صلاحیت تدریس و تربیت مدرسان دوره های آموزشی کارکنان دولت

عنوان دوره :

روش های سنجش و ارزشیابی آموزشی

مدت دوره :

۸ ساعت

مدرس:

محمدحسین حیدری

تمامی حقوق این اثر محفوظ و متعلق به سازمان مدیریت و برنامه ریزی
استان چهارمحال و بختیاری می باشد.

فهرست مطالب

۴	مقدمه
۴	تعریف اصطلاحات
۵	انواع مقیاس‌های اندازه‌گیری
۵	۱- مقیاس اسمی
۵	۲- مقیاس ترتیبی
۵	۳- مقیاس فاصله‌ای
۶	۴- مقیاس نسبتی
۶	رویکردهای ارزشیابی آموزشی
۷	الگوها (مدل‌ها) ارزشیابی
۷	۱- الگوی برخاسته از رویکرد مبتنی بر هدف
۸	۲- الگوی ارزشیابی تایلر
۸	۳- الگوی ارزشیابی برخاسته از رویکرد مبتنی بر مدیریت
۸	۴- الگوهای ارزشیابی سیپ
۹	انواع ارزشیابی از نظر زمان اجرا و هدف
۹	سنجش ورودی (pre-assessment)
۱۰	ارزشیابی تکوینی (formative evaluation)
۱۰	ارزشیابی تشخیصی (diagnostic evaluation)
۱۱	ارزشیابی بر اساس حیطه‌های ۶گانه شناختی
۱۳	انواع ارزشیابی
۱۴	اجرای آزمون
۱۵	انواع ابزارها و آزمون‌های ارزشیابی آموزشی و نحوه طراحی آنها
۱۵	آزمون‌های عینی
۱۵	آزمون صحیح-غلط (true-false)

۱۶	آزمون چندگزینه‌ای (multiple-choice)
۱۷	آزمون های جور کردنی
۱۷	آزمون های ذهنی (غیر عینی)
۱۸	آزمون های تشریحی (essay test)
۱۹	فرآیند تحلیل سوال های آزمون
۲۱	تحلیل و تفسیر اطلاعات ارزشیابی
۲۳	۱۰ نکته مهم در ارزشیابی آموزشی

مقدمه

سنجش و ارزشیابی را می‌توان بخشی از فرایند آموزش دانست که مدرسان برای مشاهده تغییرات به وجود آمده در یاد گیرنده از آن بهره می‌برند. ارزشیابی بخش جدایی ناپذیر در برگزاری موثر دوره‌های آموزشی نیز محسوب می‌شود. در واقع، اطمینان از نیل به آنچه به عنوان اهداف آموزشی در برگزاری یک دوره آموزشی انتخاب شده، از راه ارزشیابی حاصل می‌شود و نقش کلیدی در کنترل کیفیت چرخه آموزش دارد. در این نوشتار تلاش شده است رویکردهای ارزشیابی و روش‌ها و ابزارهای سنجش و ارزشیابی فعالیت‌های آموزشی معرفی گردند.

تعریف اصطلاحات

اندازه‌گیری (measurement): فرایندی دقیق و نظام‌دار که تعیین می‌کند یک شخص یا یک شیء چه مقدار از یک ویژگی برخوردار است که معمولاً به صورت عدد (نمره) مشخص می‌گردد. البته روش نسبت‌دادن اعداد باید روشن باشد تا امکان تفسیر یا سوگیری شخصی وجود نداشته باشد.

آزمون (test): ابزار یا وسیله‌ای که برای اندازه‌گیری یکی از ویژگیهای روانی یا آموزشی- تربیتی یک فرد یا گروهی از افراد مورد استفاده قرار می‌گیرد. «آزمون» نوعی از اندازه‌گیری است ولی هر نوع اندازه‌گیری (مثلاً مساحت اتاق) لزوماً آزمون نیست.

سنجش (assessment): سنجش مفهوم وسیع‌تری از آزمون دارد. زمانی که معلمان علاوه بر آزمون از راههای متنوعی چون مشاهده، مبادله کلامی، بررسی تکالیف و گزارش‌های کتبی و غیره، به مجموعه نسبتاً کاملی از اطلاعات درباره فراگیران دست می‌یابند، در حقیقت به سنجش وضعیت تحصیلی فراگیر اقدام کرده‌اند.

ارزشیابی (evaluation): فرایند نظام‌دار برای جمع‌آوری، تحلیل و تفسیر اطلاعات (بدست آمده از طریق آزمون یا سنجش) به منظور تعیین این که آیا اهداف مورد نظر محقق شده‌اند و به چه میزانی؟ ارزشیابی پیشرفت تحصیلی، شامل داوری ارزشی درباره عملکرد آموزشی فراگیران است که با اتکا به معیارهای معین، صورت می‌گیرد. پس هرچه اطلاعات جامع‌تر و دقیق‌تر باشد، ارزشیابی مطلوب‌تر خواهد بود به این دلیل که :

- چون ارزشیابی حالت فرایندی دارد، جریان آن متوقف در مرحله خاصی از فعالیت یا فعالیت‌ها نمی‌باشد.

- فعالیت ارزشیابی، نظامدار و هدفمند است و مراحل منظم و مستمر و منسجمی دارد که بر اساس آن، تحلیل یافته‌ها و مقایسه‌ها صورت می‌گیرد.
- نتیجه ارزشیابی، داوری و تصمیم‌گیری است، در غیر این صورت فاقد ارزش اساسی و توجیه علمی است.

انواع مقیاس‌های اندازه‌گیری

مقیاس‌های اندازه‌گیری در چهار دسته طبقه‌بندی می‌شوند: اسمی، ترتیبی، نسبی، فاصله‌ای

۱- مقیاس اسمی: اندازه‌گیری اسمی مستلزم جای دادن اشیاء یا افراد در طبقاتی است که از نظر کیفی با هم متفاوت هستند. یک مثال ساده در این مورد متغیر جنسیت است. در سطح اندازه‌گیری اسمی استفاده از اعداد تنها به منظور مشخص کردن طبقات می‌باشد و هیچ رابطه‌ای که با اعمال ریاضی مطابقت کند بین آن‌ها وجود ندارد. اعداد به کار رفته در یک مقیاس اسمی معرف مقدار مطلق یا نسبی هیچ ویژگی نیستند.

۲- مقیاس ترتیبی: مقیاس‌های ترتیبی برای ما امکان رتبه‌بندی ترتیب سطوح متغیر مورد بررسی را فراهم می‌آورد. برای مثال چهارویژگی داشتن روابط عمومی قوی، تجربه کافی، تحصیلات مرتبط و کارکردن به صورت مستقل، را در نظر می‌گیریم و از پاسخگویان می‌خواهیم این صفات را برحسب درجه اهمیت آن‌ها از ۱ تا ۵ رتبه‌بندی کنند. به این ترتیب تفاوت ارجحیت هر مقوله از دید پاسخگو تعیین می‌شود. با این حال یکی از مشکلات مقیاس ترتیبی آن است که هرچند محقق می‌تواند تفاوت‌های مورد نظر را برحسب بزرگتر یا کوچکتر بودن، بهتر یا بدتر بودن و غیره رتبه‌بندی کند، ولی قادر نیست فاصله این تفاوت‌ها را بدست آورد. به عبارت دیگر مقادیر حاصل در مقیاس ترتیبی فاقد معنای کمی و عددی هستند. علت این امر آن است که در مقیاس ترتیبی هیچ مبدایی وجود ندارد تا ارزش‌ها نسبت به آن سنجیده شوند.

۳- مقیاس فاصله‌ای: ویژگی عمده مقیاس فاصله‌ای آن است که به ارزش‌های آن نمره‌های عددی تعلق می‌گیرد. این مقیاس علاوه بر آن که اطلاعات مربوط به مقیاس‌های طبقه‌ای و ترتیبی را در بر دارد، اطلاعات اضافی در مورد تفاوت‌ها به دست می‌دهد. برای مثال دو دمای ۳۰ و ۴۰ درجه در نظر بگیرید. این دو دما با یکدیگر تفاوت دارند، دمای ۴۰ بیشتر از ۳۰ است و همچنین دقیقاً ۱۰ درجه تفاوت وجود دارد. به این ترتیب در این مقیاس می‌توان روشهای آماری مختلف مانند محاسبه میانگین، واریانس، همبستگی و... را انجام داد. لازم به ذکر است که

این مقیاس از نقطه صفر مطلق شروع نمی شود ولی می توان برای آن صفر قراردادی برگزید. به طور مثال اگر سه نفر در یک آزمون روانشناسی نمرات ۳۰ و ۱۵ گرفته باشند نمی توان گفت نمره ۳۰ معرف دو برابر بودن دانش فرد نسبت به فرد با نمره ۱۵ است.

۴- مقیاس نسبتی: این مقیاس نه تنها دارای تمام ویژگی‌های مقیاس‌های ذکر شده است بلکه دارای مبدأ مطلق یا صفر حقیقی نیز می باشد و بنابراین امکان محاسبه نسبت‌های مختلف از لحاظ صفت مورد نظر در این مقیاس وجود دارد. متغیرهایی مانند سن، نرخ مرگ و میر، وزن، قد افراد، جمعیت و ... از این دست هستند. کلیه اعمال آماری و ریاضی را می توان بر روی این مقیاس انجام داد.

رویکردهای ارزشیابی آموزشی

صاحب نظران ارزشیابی آموزشی در دیدگاه‌هایشان نسبت به این که ارزشیابی چیست و چگونه باید آن را انجام داد اختلاف نظر دارند. این اختلاف نظرها منجر به رویکردهای ارزشیابی آموزشی مختلفی شده است که تقسیم‌بندی «ورتن و سندرز» از معروفترین آنهاست:

۱- **رویکرد مبتنی بر هدف:** در رویکرد مبتنی بر هدف، هدف‌های کلی و هدف‌های دقیق مورد تاکید قرار می گیرند و ارزشیابی آموزشی به دنبال این است که تعیین کند این هدف‌ها به چه میزان تحقق یافته اند.

۲- **رویکرد مبتنی بر مدیریت:** هدف تشخیص و تحقق بخشیدن به نیازهای اطلاعاتی مدیران تصمیم گیرنده است.

۳- **رویکرد مبتنی بر مصرف کننده:** برکسب اطلاعات مربوط به نتایج کلی آموزشی تاکید دارد که توسط مصرف کنندگان آموزشی از میان برنامه‌های درسی مختلف انتخاب و مورد استفاده قرار می گیرد .

۴- **رویکرد مبتنی بر نظر متخصصان:** در این رویکرد تاکید اصلی بر کاربرد مستقیم نظر متخصصان در قضاوت درباره کیفیت فعالیتهای آموزش و پرورش است .

۵- **رویکرد مبتنی بر اختلاف نظر متخصصان:** در این رویکرد اختلاف نظرهای ارزشیابان موافق و مخالف مورد تاکید است .

۶- **رویکرد طبیعت گرایانه و مبتنی بر مشارکت کنندگان:** در این رویکرد بررسی طبیعت گرایانه و کوشش مشارکت کنندگان در تعیین ارزش‌ها، ملاک‌ها، نیازها و داده‌های ارزشیابی مورد تاکید است.

این شش رویکرد «ورتن و سندرز» بر تقسیم بندی دوگانه «هاوس» از ارزشیابی آموزشی استوار است. هاوس انواع ارزشیابی آموزشی را به دو دسته تقسیم کرد:

۱- عین گرا در مقابل ذهن گرا

۲- فایده گرا در مقابل کثرت گرا - شهودگرا

در حقیقت این نوع تقسیم بندی پیوستاری می سازد که در یک کرانه آن عین گرا یا فایده گرا و در کرانه دیگر آن ذهن گرا یا شهودگرا - کثرت گرا قرار می گیرد.

رویکرد مبتنی بر هدف و رویکرد مبتنی بر مدیریت با کرانه عین گرایی یا فایده گرایی منطبق هستند و رویکرد طبیعت گرایانه و مبتنی بر مشارکت کنندگان بر کرانه ذهن گرایی یا شهودگرایی - کثرت گرایی انطباق دارد و بقیه طبقات در حد فاصل این دو طبقه واقع هستند.

در ارزشیابی فایده گرا، بیشترین خوبی آن است که به بیشترین افراد فایده برساند. فایده گرایی به این فکرند که شادمانی را در جامعه باید به حداکثر رساند و حداکثر خوبی را برای حداکثر افراد تعیین نمایند. رویکرد شهودگرا یا کثرت گرا به این فکرند که :

ارزش به تاثیر برنامه بر فرد فرد افراد وابسته است نه بر اکثریت آنان .

ذهن گرایی : در این رویکرد، احساس ها و دریافتهای فردی ملاک خوبی و درستی به حساب می آیند.

کثرت گرایی : در این رویکرد ارزشیابی تاکید بر توزیع دستاوردها در میان همه افراد و خرده گروه ها است.

شهودگرایی : متعادل ساختن و وزن دادن به ارزشیابی ها و ملاکهای گوناگون در این رویکرد عمدتاً جنبه شهودی دارد.

الگوها (مدلها) ارزشیابی

۱- **الگوی برخاسته از رویکرد مبتنی بر هدف :** در رویکرد ارزشیابی مبتنی بر هدف ، ابتدا اهداف و مقاصد یک فعالیت آموزشی تعیین می شود و بعد از آن در ارزشیابی کوشش به عمل می آید تا نشان دهد که آن اهداف و مقاصد چه اندازه تحقق یافته اند. در آموزش و پرورش فعالیت ارزشیابی می تواند از یک روز آموزشی کلاسی تا تمامی نظام آموزشی را شامل شود. نتایج بدست آمده از ارزشیابی مبتنی برهدف را می توان برای اصلاح و تجدید نظر اهداف فعالیت های آموزشی ، خود فعالیت آموزشی و روش ها و تدابیر به کار رفته برای تحقق اهداف به کار برد.

۲- الگوی ارزشیابی تایلر: نخستین الگوی ارزشیابی آموزشی است که تاریخ پیدایش آن دهه ۱۹۳۰ میلادی است. مفهوم ارزشیابی آموزشی تایلر این است که تعیین کند هدف‌های آموزشی برنامه درسی تا چه میزانی تحقق یافته‌اند. تایلر در الگوی ارزشیابی مراحل زیر را پیشنهاد داده است:

۱- تعیین هدف‌های کلی و هدف‌های دقیق

۲- طبقه‌بندی هدف‌ها

۳- بیان هدف‌ها به صورت رفتاری

۴- یافتن موقعیت‌هایی که در آنها بتوان دستیابی به هدف‌ها را نشان داد

۵- تهیه روش‌ها و فنون اندازه‌گیری

۶- گردآوری داده‌های مربوط به عملکرد یادگیرندگان

۷- مقایسه داده‌های مربوط به عملکرد یا هدف‌های رفتاری

در نتیجه ارزشیابی، اگر معلوم شد که بین عملکرد و هدف‌ها اختلاف وجود دارد، باید نواقص برطرف شده و بار دیگر ارزشیابی انجام گیرد. بنا به گفته ورتن و سندرز (۱۹۸۷)، «الگوی ارزشیابی تایلر، از لحاظ منطقی درست، از نظر علمی قابل پذیرش و بوسیله متخصصان ارزشیابی به آسانی قابل استفاده است و حوزه برنامه درسی را شدیداً تحت تاثیر قرار داده است.»

۳- الگوی ارزشیابی برخاسته از رویکرد مبتنی بر مدیریت: دسته‌ای از الگوهای ارزشیابی، از رویکرد مبتنی بر مدیریت به قصد کمک به مدیران و تصمیم‌گیرندگان برخاسته‌اند. پدیدآورندگان این الگوها بر یک رویکرد نظام‌دار تاکید دارند که در آن، در مورد درون داده‌ها، فرایند و برون داده‌ها تصمیم‌گیری می‌شود، مخاطبین این الگوها تصمیم‌گیرندگان آموزش و پرورش هستند.

۴- الگوهای ارزشیابی سیپ: پدیدآورندگان این الگوی ارزشیابی استافل بیم و همکاران او هستند. عنوان سیپ (SIPP) از حروف اول کلمات بافت (structure)، درون داد (input)، فرآیند (process) و فرآورده (product) به دست آمده است. این الگو دارای چهارچوبی است که مدیران و تصمیم‌گیرندگان را در چهار نوع تصمیم‌گیری زیر کمک می‌کند:

۱- ارزشیابی از بافت یا زمینه: مهمترین نوع ارزشیابی از دیدگاه استافل بیم و همکاران او، ارزشیابی از بافت است. هدف آن فراهم آوردن یک زمینه منطقی برای تعیین هدفهای آموزشی است. روشهای ارزشیابی از بافت عمدتاً توصیفی و مقایسه‌ای هستند.

۲- ارزشیابی از درون‌داد: در ارزشیابی درون‌داد اطلاعات مورد نیاز درباره چگونگی استفاده از منابع به منظور دستیابی به هدفهای برنامه جمع‌آوری می‌شوند و به تصمیم‌گیرندگان کمک می‌کنند تا روش‌هایی را انتخاب و طراحی کنند که برای دستیابی به هدفها، ضروری تشخیص داده شده‌اند.

۳- ارزشیابی از فرایند: این فرایند بیشتر به دنبال تعیین پاسخ به پرسش‌هایی از قبیل، آیا برنامه به خوبی در حال پیاده شدن است؟ چه موانعی بر سر راه موفقیت آن قرار دارند؟ تغییرات ضروری کدام است که پاسخ این پرسش‌ها به کنترل و هدایت شیوه‌های اجرایی کمک می‌کند.

۴- ارزشیابی از فراورده: در این مرحله نتایج به دست آمده با نتایج برنامه مقایسه می‌گردد و رابطه بین انتظارات و نتایج واقعی مشخص می‌شود پس از بررسی کامل نتایج، اطلاعات در اختیار تصمیم‌گیرندگان قرار می‌گیرد تا تصمیم بگیرند که برنامه ادامه داشته باشد یا متوقف شود یا تجدیدنظر در آن صورت گیرد.

انواع ارزشیابی از نظر زمان اجرا و هدف

ارزشیابی فرایند آموزشی را با توجه به زمان ارزشیابی و هدفهای آن می‌توان به چهار دسته به شرح زیر تقسیم کرد و آنها را به منظورهای متفاوت در زمان و مقاطع مختلف آموزشی به کار برد.

- ورودی یا آغازین (pre-assessment)
- تکوینی (formative evaluation)
- تشخیصی (diagnostic evaluation)
- تراکمی یا پایانی (summative evaluation)

سنجش ورودی (pre-assessment): این نوع ارزشیابی همان‌گونه که از نام آن نیز مشخص است، نخستین ارزشیابی مدرس است که پیش از فعالیتهای آموزشی او به اجرا در می‌آید. مدرس در این نوع ارزشیابی به آموخته‌های پیشین فراگیران واقف می‌گردد. زمان اجرای آن در بدو ورود به دوره می‌باشد و با دو هدف انجام می‌گردد:

- آمادگی: آیا فراگیران پیش نیازهای درس جدید را آموخته‌اند؟

- جایابی: تا چه حدی فراگیران مطالب درس جدید را از قبل می‌دانند؟

سؤالات این نوع آزمون باید دو دسته باشند: یک دسته از محتوای دروس قبلی و دسته دوم از محتوای درس جدید.

ارزشیابی تکوینی (formative evaluation): ارزشیابی تکوینی زمانی به اجرا در می‌آید که فعالیت‌های آموزشی هنوز جریان دارد و یادگیری فراگیران در حال تکوین یا شکل‌گیری است. معمولاً تحقق اهداف آموزشی بتدریج و به مرور زمان امکان‌پذیر می‌شود، به این دلیل، مدرس ناگزیر است تحقق هدف‌های آموزشی هر بخش را در فاصله زمانی معینی، متناسب با توانایی و امکانات فراگیران انتظار داشته و برای آن که از چگونگی تحقق هدف‌های آموزشی هر بخش از مطالب تدریس شده آگاه شود، لازم است در پایان هر بخش، آموخته‌های شاگردان را مورد ارزشیابی قرار دهد. بنابراین این ارزشیابی با دو هدف انجام می‌گردد:

۱- همراه کردن فراگیران با فرایند تدریس و کاهش بار محتوایی و روانی آزمون پایانی.

۲- اطلاع کلی مدرس از نقاط قوت و ضعف فراگیران، روش تدریس، محتوا و ... به منظور اصلاح فرایند.

ارزشیابی تشخیصی (diagnostic evaluation): زمانی که مدرس می‌خواهد اطمینان حاصل کند که فراگیران مهارت‌های لازم برای ورود به مطالب جدید در دوره را دارند و همچنین آنها چه میزان از مطالبی را که قرار است تدریس شوند را می‌دانند، از این روش ارزشیابی استفاده می‌کند. سؤالات مطرح شده در ارزشیابی باید به گونه‌ای تنظیم شود که رفتار ورودی و هدف‌های اجرایی درس جدید را در بر بگیرد و مشخص کند که آیا شاگردان پیش‌نیاز درس جدید را بطور مطلوب آموخته‌اند یا نه و یا از مفاهیم درس جدید چقدر آگاهی دارند. ارزشیابی تشخیصی علاوه بر اینکه معیار سنجش رفتار ورودی شاگرد است، وسیله‌ای است که به کمک آن می‌توان اولین گام آموزشی را بدقت معین کرد. زمان اجرای این ارزشیابی در طول دوره (بازخورد ارزشیابی تکوینی) و هدف عمده اجرای آن تشخیص مشکلات جدی و مکرر برخی از فراگیران است. نتایج این نوع ارزشیابی به مدرس کمک خواهد کرد تا علاوه بر طراحی آموزش‌های ویژه و تکمیلی (ترمیمی) از وجود یک روان‌شناس یا مشاور آموزشی مطلع و آگاه برای ریشه‌یابی ضعف برخی فراگیران استفاده کند.

ارزشیابی تراکمی (summative evaluation): زمان انجام ارزشیابی تراکمی در پایان یک دوره آموزشی (از یادگیری متراکم یا تجمیع شده فراگیران) می‌باشد. اهداف انجام این ارزشیابی به شرح زیر می‌باشد:

- ۱- امکان ارزشیابی همزمان کل محتوای آموزشی و محک زدن قدرت تشخیص فراگیران (ممکن است فراگیران در ارزشیابی‌های مستمر و تکوینی، از عهده‌ی پاسخ به سوالات هر مبحث یا فصل کتاب برآیند، اما آیا زمانی که با کل محتوا مواجه می‌شوند، از عهده تشخیص تفاوتها و شباهتها برمی‌آیند؟)
- ۲- ارتقاء فراگیران به پایه بالاتر.
- ۳- قضاوت درباره کل برنامه آموزشی.

ارزشیابی بر اساس حیطه‌های ۶گانه شناختی

همان گونه که تدریس مباحث نظری توسط یک مدرس، در حیطه‌های مختلفی قابل انجام است، ارزشیابی نیز باید منطبق بر هر یک از این سطوح باشد. در یک ارزشیابی استاندارد، مدرس تلاش می‌کند، سوالات متنوعی از حیطه‌های مختلف شناختی طرح نماید و پرسش‌های آزمون را به یک حیطه خاص، محدود نکند. حیطه‌های ۶گانه شناختی عبارتند از:

- ۱- دانش: تکرار پاسخ‌هایی که قبلاً در موقعیت یادگیری تمرین و آموخته شده‌اند: شامل: یادآوری، بازشناسی و تشخیص.
- ۲- فهمیدن یا فراگیری: توانایی درک منظور یا مقصود در یک مطلب شامل:
 - ترجمه کردن یا برگردان: ترجمه در واقع بازگویی مطلب به زبان خوداست. مهارت برگرداندن یا تغییر مطالب از شکلی به شکل دیگر بدون این که معنی و محتوای آن‌ها دگرگون شود، ترجمه نامیده می‌شود.
 - تفسیر کردن (بسط دادن و ارائه توضیحات بیشتر): تفسیر شامل تفکر درباره اهمیت نسبی اندیشه‌هایی است که ممکن است فهمیدن آنها مستلزم تنظیم مجدد اندیشه‌ها، به صورت ترکیبی تازه، در ذهن فرد باشد. در واقع تفسیر درک تدابیر گوناگون به کار رفته در انتقال مفاهیم و یک نوع بازیابی مفاهیم در ذهن است.
 - خلاصه کردن
 - مقایسه کردن
- ۳- کار بستن: توانایی استفاده از امور انتزاعی (قواعد و قوانین، اصول، اندیشه‌ها و روش‌ها در موقعیت‌های عینی و عملی) در موقعیت حل مساله است. در حقیقت، به کار بستن، فراتر از فهمیدن است. در این سطح، وقتی به فراگیر مساله تازه‌ای داده می‌شود، او مفاهیم انتزاعی مناسب آن را به کار می‌گیرد و مساله را حل می‌کند، بدون این که لازم باشد به او گفته شود از کدام مطلب انتزاعی استفاده کند

انواع ارزشیابی

در نگاهی کلی، عبارت «ارزشیابی آموزشی» در دو معنی بکار می رود:

۱- ارزشیابی آموزشی: این نوع ارزشیابی عموماً توسط برنامه ریزان و کارشناسان انجام می شود تا نتایج حاصل به عنوان بازخورد به نظام آموزشی بازگردانده شود. مثلاً ممکن است در یک دانشگاه، پرسش اساسی این باشد که معاونت آموزشی و زیرمجموعه های آن، در کل عملکردی مطلوبی دارند یا نه. آنگاه طی یک برنامه مدون، بررسی عملکرد معاونت آموزشی آن دانشگاه مورد ارزیابی قرار می گیرد.

۲- ارزشیابی پیشرفت تحصیلی: این نوع ارزشیابی، زیر مجموعه ای از ارزشیابی آموزشی است و هدف آن بررسی میزان تحقق اهداف آموزشی در کلاس است. یکی از تخصص های مورد نیاز دست اندرکاران فعالیتهای آموزشی در سطوح مختلف سازمانی و اجرائی، برخوردار بودن از دانش و مهارت سنجش و ارزشیابی پیشرفت تحصیلی می باشد که فقدان آن در عرصه عمل و اجرا، ناکارآمدی فرایند یاددهی - یادگیری را به دنبال خواهد داشت. برای انجام ارزشیابی پیشرفت تحصیلی، مدرس باید چشم انداز درستی نسبت به انواع آزمونها داشته باشد. در شکل زیر، انواع کلی آزمونها به تفکیک بیان شده است:



نکاتی درباره انواع آزمون

- آزمونهای شناختی بیشتر قابلیت‌های فکری و ذهنی را در زمینه‌های تحصیلی می‌سنجند. درحالی که آزمونهای روانی - حرکتی مهارت‌های عملی افراد را مورد آزمون قرار می‌دهد.
- آزمونهای استعداد ناظر به آینده اند ولی آزمون‌های پیشرفت تحصیلی ناظر به گذشته‌اند. البته تفکیک این دو گاه به روشنی امکان پذیر نیست.
- لازم است که هنگام شروع یک دوره آموزشی، مدرس، علاقه و انگیزه‌های فراگیران را در موضوعاتی مانند رشته تحصیلی، موضوع درس، وضعیت کلاس و مدرس و ... را از طریق آزمونهای عاطفی بررسی نماید.

اجرای آزمون

شرایط اجرای آزمون‌های پیشرفت تحصیلی باید به گونه ای باشد که از هر لحاظ آسایش جسمی و فکری آزمون شونده‌گان را تأمین نماید. علاوه بر شرایط فیزیکی، شرایط درونی آزمون شونده‌گان نیز بر عملکرد آنان تأثیر دارد. مهمترین این عوامل، اضطراب زیاد است. تحقیقات نشان می‌دهد اندکی اضطراب افزایش دقت و کوشش آزمون شونده می‌شود، اما اضطراب زاد بر عملکرد او تأثیر منفی دارد. توجه به موارد زیر در اجرای آزمون‌ها حائز اهمیت است:

- حفظ نگرش مثبت امتحان دهندگان
- به حداکثر رساندن انگیزه پیشرفت و تلاش امتحان دهندگان برای عملکرد بهتر در آزمون
- توجه یکسان و برابر به فراگیران زرنگ و تنبل
- از قبل در مورد زمان امتحان اطلاع رسانی نماید تا آنها غافلگیر نشوند.
- روشن بودن قوانین
- توزیع چرخشی برگه های آزمون بین افراد
- اعلام تعداد برگه ها و سوالات امتحان قبل از پاسخ دادن
- نظارت بر دانش آموزان در زمان برگزاری آزمون
- به حداقل رساندن عوامل حواس پرتی افراد

- هشدار در مورد میزان زمان باقی مانده ۵ - ۱۰ - ۱۵ (دقیقه به پایان امتحان)
- جمع آوری یکنواخت برگه آزمون (و اینکه آیا برگه های سوال را می خواهید یا فقط پاسخ نامه ها را)

انواع ابزارها و آزمون های ارزشیابی آموزشی و نحوه طراحی آنها

اولین گام در اجرای آزمون های پیشرفت تحصیلی، آماده سازی آزمون ها و شرایط اجرای آنهاست. تهیه پرسشنامه ها و پاسخنامه ها و ترتیب پشت سر هم قرار گرفتن سؤال های مختلف باید به گونه ای باشد که به آزمون شونده گان در ارائه بهترین عملکرد کمک نماید. دسته بندی سؤالات مختلف با توجه به نوع آنها به آزمون شونده نظم فکری می دهد و پاسخ دهی به سؤالات را برای او آسان می سازد. آزمون خوب آزمونی است که بتواند به بهترین شکل منعکس کننده تمامی اهداف و محتوای درسی ارائه شده باشد. آزمون های پیشرفت تحصیلی بطور کلی به دو دسته آزمون های عینی و آزمون های ذهنی (غیر عینی) تقسیم می شوند:

آزمون های عینی: آزمون هایی هستند که سوال و جواب در اختیار فراگیر قرار می گیرد و فراگیر درباره جواب آن

تصمیم می گیرد این آزمون ها به سه دسته تقسیم می شوند: صحیح و غلط، چند گزینه ای، جور کردنی

آزمون صحیح-غلط (true-false): این آزمون ها به چهار شکل ممکن است طراحی شوند، نوع صحیح غلط یعنی سوال به شکل یک جمله خبری طرح و صحیح غلط آن خواسته می شود. نوع دوم بلی یا خیراست که در این حالت سوال به شکل یک جمله استفهامی طرح می گردد. نوع سوم اصلاحی است بدین صورت که صحیح و غلط بودن جمله داده شده، خواسته می شود و نوع چهارم به شکل خوشه ای یا سوال چند صحیح و غلط است، بدین صورت که اصل سوال به صورت یک جمله نا تمام نوشته می شود و به دنبال آن پاسخ هایی ارائه می گردد که لازم است صحیح و غلط بودن آنها مورد قضاوت قرار بگیرد.

محاسن :

۱- عینی بودن تصحیح جوابها؛

۲- طرح مجدد موضوعات مهم درسی در قالب آزمون؛

۳- امکان طرح پرسش در حیطه ی شناختی با بالاتر از دانش (مثلاً درک و فهم)

نکات مهم در طراحی:

۱- «محورهای اصلی» در قالب عبارات جدید یا بیان متفاوت طرح شود.

۲- درست یا غلط بودن هر پرسش منطقاً قابل دفاع باشد.

۳- هر پرسش حاوی یک موضوع باشد.

۴- از کلمات راهنما و تسهیل کننده استفاده نشود (انقلاب‌ها همیشه به اهداف مطلوب اجتماعی انجامیده‌اند)

۵- طول پرسش‌های صحیح و غلط تقریباً هم اندازه انتخاب شود.

آزمون چندگزینه‌ای (multiple-choice): این آزمون شامل سوالاتی است که هر سوال از یک قسمت اصلی و تعدادی

گزینه تشکیل می‌شود. هر سوال چند گزینه ای شامل سه قسمت می باشد، قسمت اصلی سوال، گزینه درست

یا پاسخ سوال و سه گزینه انحرافی

محاسن :

۱- عینی بودن تصحیح جواب‌ها؛

۲- کاهش امکان حدس پذیری (نسبت به صحیح-غلط)

۳- امکان طرح پرسش در حیطه‌های بالاتر شناختی

۴- امکان تشخیص بدفهمی‌های متداول فراگیران به کمک گزینه‌های انحرافی

نکات مهم در طراحی آزمونهای چند گزینه‌ای:

۱- از تکرار مطالب در گزینه‌ها خودداری شود.

۲- مطالب اصلی به طور کامل در تنه سوال بیاید.

۳- گزینه‌های انحرافی با پاسخ صحیح متجانس باشند تا توجه فراگیران بی اطلاع را به خود جلب کنند.

۴- از نوشتن جملات منفی در منفی خودداری شود.

۵- طول گزینه‌های صحیح و غلط تقریباً هم اندازه انتخاب شود.

۶- تا حد ممکن از عبارات «همه موارد» یا «هیچکدام» استفاده نشود.

۷- محل گزینه درست به صورت تصادفی انتخاب شود.

۸- در آزمون‌های تکوینی، گاه می توان ۳ تا ۵ گزینه در نظر گرفت.

۹- در آزمونهای تکوینی، گاه می توان از گزینه «نمی‌دانم» استفاده کرد.

۱۰- بهتر است در تنه سوال از جمله پرسشی استفاده شود تا جمله ناتمام.

۱۱- در سوال‌هایی که تنه جمله ناتمام است، بهتر است جای خالی در آخر جمله باشد.

آزمون های جور کردنی: از تعدادی سوال تشکیل می شوند که هر یک شامل دو ستون است. یک ستون پرسش ها و ستون های دیگر پاسخ ها قرار گرفته و آزمون شونده، پرسش ها را باید با پاسخ ها جور کند. این آزمون برای آزمون دانش اصطلاحات، تعاریف، تاریخ و رویداد ها مناسب است و همزمان امکان آزمون قسمت زیادی از محتوای درس و اهداف آموزشی را فراهم می کند، همچنین تصحیح آن به سهولت انجام می شود و نیازی به داشتن تخصص در تصحیح لازم نیست. اما این نوع از آزمون ها ممکن است فراگیر را به یادگیری سطحی تشویق کند و از طرف دیگر در بعضی مواقع پیدا کردن مجموعه ای از پرسش ها و پاسخ های همگون مشکل است

قواعد تهیه سوال های جور کردنی:

- ۱- پرسش ها و پاسخ ها متجانس و همگون انتخاب کنید.
- ۲- طول و فهرست پرسش و پاسخ را کوتاه انتخاب کنید.
- ۳- در راهنمای سوال اطلاعات لازم را درباره نحوه و مقایسه و جور کردن پاسخ ها در اختیار فراگیر قرار دهید.
- ۴- همه پرسش ها و پاسخ های یک سوال را در یک صفحه قرار دهید.
- ۵- در صورت امکان فهرست پاسخ ها را به طور منطقی مرتب کنید.
- ۶- هر یک از پاسخ ها باید برای تمام پرسش های یک سوال درست جلوه کند.
- ۷- پرسش ها را با شماره و پاسخ ها را با حروف مشخص کنید.
- ۸- در فهرست سوالهای جور کردنی از کاربرد جملات نیمه تمام خودداری گردد.

آزمون های ذهنی (غیر عینی): آزمون هایی هستند که سوال در اختیار آزمون شونده قرار می گیرد و او جواب

سوال ها را آماده می کند. می توان گفت این آزمون همان آزمون تشریحی می باشد.

نوعی دیگری از آزمون پیشرفت تحصیلی وجود دارد که حد وسط دو نوع آزمون عینی و ذهنی قرار دارند. در این آزمون ها که به آزمون های کوتاه پاسخ شهرت دارند سئوال ها بسیار شبیه به آزمون های تشریحی هستند، زیرا پاسخها را خود آزمون شونده باید تهیه کند، اما سئوال ها بسیار دقیق و مشخص هستند. جواب سئوال های کوتاه پاسخ معمولاً به یک یا چند کلمه حداکثر یک جمله) یا یک عدد و یا یک علامت خلاصه می شود.

آزمون‌های تشریحی (essay test)

- ۱- گسترده پاسخ - انشایی : در این نوع آزمون هیچ گونه محدودیتی از نظر زمان پاسخ دهی و مقدار پاسخ وجود ندارد. (این نوع آزمون‌ها برای سنجش هدف‌های طبقه ترکیب و ارزشیابی مناسب است).
 - ۲- محدود پاسخ - کنترل شده : آزمون شونده در پاسخ‌دادن آزادی کامل ندارد بلکه صورت سؤال او را ملزم می‌سازد تا پاسخ خود را در چهارچوب شرایطی خاص محدود کند .
- محاسن:

- ۱- تهیه آزمون‌های تشریحی از تهیه آزمون‌های عینی آسان‌تر است.
- ۲- در سنجش توانایی آزمون‌شوندگان، در پروراندن جواب سؤال‌ها و نحوه بیان آنها مناسب است.
- ۳- این آزمون‌ها توانایی پاسخ دادن به سؤال‌ها را می‌سنجند نه توانایی انتخاب پاسخ‌ها را.
- ۴- بر روش مطالعه یادگیرندگان تأثیر مثبت به‌جای می‌گذارند. فراگیر را مجبور به سازماندهی ذهنی و ترسیم روابط میان اجزاء بحث می‌نماید.
- ۵- مجال برای بروز خلاقیت فراگیران ایجاد می‌کند.
- ۶- توانایی نگارش فراگیران را تقویت می‌کند.

نکات مهم در طراحی آزمون‌های تشریحی

- ۱- از سوالات تشریحی برای سنجش طبقات بالای حیطه‌های شناختی استفاده شود.
- ۲- در صورت سوال از عبارات روشن و دقیق استفاده شود و از کلی‌گویی پرهیز شود.
- ۳- از کلمات پرسشی استفاده شود که راه را برای تحلیل و توضیح هموار می‌کند (چگونه، به چه دلیل و ...)
- ۴- از افعالی استفاده شود که راه را برای تحلیل و توضیح هموار می‌کند.
- ۵- باید از فراگیر خواسته شود که اظهارات خود را با دلایل علمی و منطقی مستند سازد.
- ۶- حق انتخاب چند سوال از میان تعدادی سوال منطقی نیست. چرا؟

نکات مهم در تصحیح آزمون‌های تشریحی

- ۱- پاسخ را بر اساس اهداف آموزشی تصحیح کنید.
- ۲- هنگام طرح سوال، کلید پاسخ‌ها را تهیه کنید.
- ۳- پاسخ‌ها را سوال به سوال تصحیح کنید.

۴- تا جای ممکن، از شناسایی نام صاحب برگه اجتناب کنید.

۵- از ایجاد وقفه های مکرر در تصحیح سوالات اجتناب کنید.

۶- روی برگه اشتباهات را مشخص کنید و نظر اصلاحی خود را بنویسید.

فرآیند تحلیل سوال های آزمون

هدف از تحلیل سوال های آزمون، بررسی تک تک سوال ها و تعیین میزان دقت و نارسای های آنهاست. در تحلیل سوال های آزمون، نقاط قوت و ضعف آزمون و کیفیت سوال های آن تعیین می شود. لازم است مدرسین پس از اجرای هر آزمون سوال های آن را تحلیل کنند و با استفاده از نتایج حاصل به تجدید نظر در آزمون و بهبود کیفیت سوال ها برای استفاده های بعدی اقدام نمایند. بنابراین توصیه می شود که مدرسان، پیش و پس از اجرای هر آزمون، سوال های آن را تحلیل کنند و با استفاده از نتایج حاصله به تجدید نظر در آزمون و بهبود کیفیت سوال ها برای استفاده های بعدی اقدام نمایند.

مراحل تحلیل سؤال :

۱- تعیین نمره هر یک از دانشجویان

۲- رتبه بندی دانشجویان بر اساس شایستگی (از بالا به پایین)

۳- تعیین گروه های بالا و پایین

۴- محاسبه شاخص دشواری هر یک از سوالات

۵- محاسبه شاخص تمایز هر یک از سوالات

۶- ارزیابی انتقادی سوالات به منظور قبول، تجدید نظر و یا حذف هر کدام از آنها

اطلاعات مورد نیاز برای تحلیل سوال های یک آزمون، پاسخ هایی هستند که آزمون شونده گان به هر سؤال داده اند. بنابراین باید تعیین شود که در هر سؤال چند نفر گزینه درست را انتخاب کرده اند، هریک از گزینه های انحرافی چند نفر را به خود جلب کرده است و چند نفر آن را بی جواب گذاشته اند.

محاسبه ضریب دشواری سؤال

درصد کل آزمون شونده گانی که به یک سؤال جواب درست می دهند ضریب دشواری آن سؤال است که با حرف P نشان داده می شود. اگر در تحلیل یک سؤال کلیه افراد یا کلیه برگه های امتحانی دخالت داشته باشند برای محاسبه ضریب دشواری سؤال آن کافی است که تعداد کل افرادی را که به آن سؤال جواب درست داده اند بر

تعداد کل آزمون شوندگان تقسیم کنیم و نتیجه را در ۱۰۰ ضرب نماییم رقم حاصل ضریب دشواری سؤال است مطابق فرمول:

$$P = \left(\frac{R}{T} \right) 100$$

در این فرمول R معرف تعداد کسانی است که به سؤال جواب درست داده و T نشان دهنده تعداد کل آزمون شوندگان است. به عنوان مثال اگر تعداد پاسخ دهندگان به یک سؤال ۸۰ نفر باشند و از این تعداد ۶۵ نفر به آن سؤال جواب درست داده باشند ضریب دشواری این سؤال برابر خواهد بود با

$$P = 100 \times 65 \div 80 = 81$$

در مواردی که تعداد آزمون شوندگان یا همان تعداد برگه‌های آزمون زیاد است و اطلاعات ما به نحوه پاسخ‌دهی افراد گروه بالا و گروه پایین محدود می‌شود الزم است از فرمول زیر استفاده کنیم .

$$P = \frac{\text{انتخابهای درست گروه پایین} + \text{انتخابهای درست گروه بالا}}{\text{تعداد افراد گروه بالا} + \text{تعداد افراد گروه پایین}} \times 100$$

برای تعیین گروه‌های بالا و پایین، متخصصان آزمون سازی از جمله ویتنی و سیبرز (۱۹۷۰) و نیتکو (۱۹۸۳) پیشنهاد کرده‌اند که:

اگر تعداد داوطلبان ۲۰ نفر یا کمتر است، برگه‌های آزمون را به دو نیمه بالا و پایین تقسیم کنید. اگر تعداد کل داوطلبانی که در آزمون شرکت کرده‌اند (یعنی تعداد کل برگه‌های آزمون) بیشتر از ۲۰ و تا ۴۰ نفر هستند، ۱۰ برگه بالا و ۱۰ برگه پایین را انتخاب کنید و کاری به بقیه برگه‌ها نداشته باشید. در شرایطی که تعداد کل آزمون شوندگان بیش از ۴۰ نفر است بهترین رقم برای گروه بالا و پایین ۲۷٪ از کل برگه‌هاست. این رقم برای شرایطی که نمرات دارای توزیع بهنجار باشند پیشنهاد شده است، اما برای آزمون‌های کلاسی هر یک از ارقام ۲۵ تا ۳۳ درصد به عنوان تعداد برگه‌های گروه بالا یا پایین مناسب است.

تحلیل و تفسیر اطلاعات ارزشیابی

تفسیر ضریب دشواری: یکی از راههای قضاوت درباره مفید بودن سؤال‌های یک آزمون آن است که ببینیم سؤال تا چه اندازه به پراکندگی یا واریانس نمرات کمک می‌کند. واریانس نمرات گروهی پاسخ دهنده به یک آزمون از دو قسمت تشکیل می‌یابد:

۱- واریانس سؤالها ۲- همبستگی‌های بین سؤالها

هر سؤالی که دارای واریانس بزرگتری است و با سایر سؤال‌های آزمون همبستگی بیشتری دارد به واریانس کل آزمون کمک بیشتری می‌کند و در مقابل سؤالی که با سایر سؤالها همبستگی ندارد و دارای واریانس کوچکی است کمک زیادی به واریانس کل آزمون نمی‌کند. اگر در تحلیل آزمون به انتخاب سؤال‌هایی پردازیم که همبستگی زیادی با یکدیگر دارند و در ضمن دارای واریانس‌های بزرگتری هستند در مجموع آزمون بهتری تدارک دیده ایم .

$$P = (p-1) * \text{واریانس سؤال}$$

در فرمول بالا ، P برابر است با درصد کسانی که به سؤال جواب درست داده‌اند یعنی همان ضریب دشواری ، یک سؤال زمانی دارای حداقل واریانس است که ضریب دشواری آن $+1$ یا 1 باشد یعنی هیچ یک از آزمون شوندگان به آن سؤال جواب درست نداده باشد یا اینکه همه آزمون شوندگان به آن جواب درست داده باشند.

$$P = (P-1) * \text{واریانس سؤال}$$

$$0 = 0 * (0-1) = 0$$

$$1 = 1 * (1-1) = 0$$

اما زمانی که ضریب دشواری سؤال $P=0,5$ باشد واریانس سؤال حداکثر خواهد بود

$$P = (P-1) * \text{واریانس سؤال}$$

$$0,5 = 0,5 * (0,5-1) = 0$$

$$0,25 = 0,5 * 0,5 = 0,25$$

بهتر است سؤال‌هایی برای آزمون انتخاب شوند که ضریب دشواری آنها از 1 کمتر ، و از صفر بیشتر و به $0,5$ نزدیک باشد. البته انتخاب سؤال‌هایی با ضرایب دشواری مناسب به نوع سؤال‌های آزمون مربوط است. در

آزمونهای صحیح - غلط ضریب دشواری مناسب به نوع سؤالهای آزمون مربوط است، در آزمونهای صحیح - غلط ضریب دشواری $p = +/5$ زمانی بدست می آید که همه آزمودنی ها را با حدس زدن به سؤال جواب بدهند .
 آن وین گفته اند برای آزمونهای چند گزینه ای سطح بهینه دشواری اندکی کمتر از وسط فاصله بین ۱ و سطح موفقیت با حدس زدن است . منظور از سطح موفقیت با حدس زدن عبارت است از نسبت پاسخ های درستی که با حدس زدن به دست می آید بنابراین برای سؤالهای چهارگزینه ای سطح موفقیت با حدس زدن حدود $+/25$ است و لذا سطح بهینه دشواری در وسط $+/25$ و 1 یعنی در حدود $+/6$ خواهد بود.

محاسبه ضریب تمیز سؤال

ضریب تمیز را با d نشان داده می دهند . قدرت سؤال را در تمایز گذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان مشخص می کند و یعنی معلوم می نمایند که سؤال تا چه اندازه می تواند گروه قوی را از گروه ضعیف جدا سازد . برای محاسبه ضریب تمیز یک سؤال از فرمول زیر استفاده می شود .

$$d = \frac{\text{تعداد افراد یک گروه (بالا یا پایین) - انتخابهای درست گروه بالا}}{\text{ضریب تمیز سوال}}$$

تفسیر ضریب تمیز

هر قدر ضریب تمیز بزرگتر باشد قوه تمیز آن سؤال بیشتر و هر قدر این ضریب کوچکتر باشد قوه تمیز آن کمتر است؛ مثلاً اگر قوه تمیز سؤال $+/90$ باشد آن سؤال آزمون شوندگان قوی و ضعیف را خیلی خوب از هم جدا خواهد کرد. اما اگر ضریب تمیز سؤال $+/10$ باشد آن سؤال از عهده جداسازی دانش آموزان قوی و ضعیف به خوبی بر نخواهد آمد. ضریب تمیز صفر حاکی از این است که آن سؤال به هیچ وجه نتوانسته بین گروه قوی ، ضعیف تمایز قائل شود گاه اتفاق می افتد که ضریب تمیز سؤالی منفی است این نوع ضریب تمیز نشان می دهد که در آن سؤال گروه قوی بدتر از گروه ضعیف عمل کرده است . این گونه سؤالها دارای معایب اساسی هستند که یا باید به کلی کنار گذاشته شوند یا در آنها تجدید نظر اساسی صورت پذیرد . اگر سؤال دارای اشکال فنی نباشد دلیل چنین ضریبی می تواند این باشد که افراد گروه بالا یا آن سؤال را به طور کامل یاد نگرفته اند یا آن را به غلط آموخته اند.

تفسیر نتایج حاصل از محاسبه ضریب تمیز:

- ۱- ضریب تمایز ۳۵٪ و بیشتر : عالی
- ۲- ضریب تمایز ۲۵ تا ۳۴٪ : خوب
- ۳- ضریب تمایز ۱۵ تا ۲۳٪ : بینابینی (باید سوال مورد تجدید نظر قرار گیرد)
- ۴- کمتر از ۱۵٪ : ضعیف (بهتر است سوال حذف شود)

۱۰ نکته مهم در ارزشیابی آموزشی :

- ۱- مهمترین هدف ارزشیابی، «اطمینان از روند مطلوب یادگیری است» نه صرفاً ابزاری برای گذر از یک دوره آموزشی.
- ۲- ارزشیابی، خود «میدانگاهی» برای یادگیری است.
- ۳- چون یادگیری فرایندی مستمر است، ارزشیابی نیز باید مستمر باشد.
- ۴- ارزشیابی باید تمام اهداف اصلی دوره را پوشش دهد.
- ۵- ارزشیابی باید معطوف به حیطه‌های یادگیری باشد.
- ۶- برای بهره‌گیری از محاسن انواع آزمون‌ها، بهتر است از آزمون‌های متنوع استفاده شود.
- ۷- ارزشیابی باید اطمینان بخش باشد نه تنش‌زا ...
- ۸- ارزشیابی باید تا حد ممکن «عینی» باشد.
- ۹- روند ارزشیابی باید از ساده به مشکل تنظیم شود.
- ۱۰- در ارزشیابی باید مجالی برای بروز خلاقیت وجود داشته باشد.