

بسمه تعالی

درسنامه مهارت طراحی سوالات امتحانی

تدوین

دکتر معصومه شوهانی

## مقدمه

براساس روانشناسی یادگیری و آموزشی، نظارت و ارزشیابی حلقه آخر فعالیتهای آموزشی و تکمیل کننده انجام آموزش محسوب می شود و عمدتاً با هدف اصلاح و بهبود فرایند آموزش صورت می گیرد. اساساً مفهوم "آموزش" به یک فعالیت، منوط به انجام ارزشیابی است. به این ترتیب نظارت و ارزشیابی آموزشی یکی از مهمترین وظایف هر آموزش دهنده است و جزء لاینفک از اجزاء یک برنامه آموزشی تلقی می شود. ابزار یا وسیله ای که برای سنجش پیشرفت تحصیلی در موقعیت های آموزشی و تدریس به کار می رود آزمون یا امتحان نام دارد. طرح سوالات آزمون، فرایندی است که نیاز به دقت و توجه خاص در رابطه با علم آموزان، کتاب و مواد درسی، منبع آزمون و شرایط و نوع امتحان دارد؛ زیرا موضوعات و سرفصل های درسی و آموزشی با هم متفاوتند و علم آموزان نیز از استعداد، درجه یادگیری، سطح دانش و اطلاعات و شرایط روحی یکسان برخوردار نیستند. بنابر این باید اصول و مقرراتی را در تنظیم سوال امتحانی رعایت نمود تا علم آموزان با استعداد و کم استعداد، در شرایط کاملاً مساوی قرار نگیرند و برای علم آموزان ضعیف نیز این فرصت را ایجاد کرد تا بتوانند به تعدادی از سوالات طرح شده متناسب با استعداد خویش جواب داده و نمره هایی که مانع از افت گردد، کسب نمایند که این خود باعث ایجاد حس اعتماد به نفس و افزایش علاقه مندی برای مطالعه و تلاش بیشتر در امتحانات بعدی او خواهد شد. توجه داشته باشید که اهداف آموزشی و آزمون باید با یکدیگر همخوانی داشته باشند و در طراحی آزمون باید بیان روشن و دقیقی از اهداف آموزش و تدریس در اختیار داشته باشیم.

## امتحان چیست ؟

امتحان یکی از عناصر فرآیند آموزش و پرورش است که اطلاعات اساسی را برای تعیین سطح کارایی و بهبود شرایط آموزشی فراهم می سازد. اما برای این که اطلاعات کسب شده مفید و موثر واقع شوند، لازم است که این اطلاعات با محتوای هدف های آموزشی، برنامه، روش، وسایل و ویژگی های دانش آموزان رابطه مستقیمی داشته باشند. زیرا بدون توجه به این عوامل، محتوای امتحان را نمی توان به درستی تعیین و مشخص نمود. علاوه بر آن نمی توان از طریق این اطلاعات به هدف های آن جامعه عمل پوشاند.

## هدف های امتحان

۱- امتحان به عنوان وسیله ای برای شناساندن هدف های تدریس به دانش آموزان

۲- امتحان به عنوان وسیله ای برای تصمیم گیری جهت شروع مراحل بعدی تدریس

۳- امتحان به عنوان وسیله ای برای یافتن نارسایی های آموزشی دانش آموزان و کشف نقاط ضعف آنان

۴- امتحان به عنوان وسیله ای برای رفع نارسایی های آموزشی دانش آموزان و کیفیت بخشیدن به یادگیری آنان

۵- امتحان به عنوان وسیله ای برای بهبود و اصلاح روش ها و رویکرد های تدریس

### اصول طرح ریزی امتحان

متأسفانه گاهی معلمان بدون داشتن هیچ گونه طرح قبلی و برنامه ریزی اصولی پرسش های امتحانی خود را تنها چند لحظه مانده به شروع امتحان طرح و تهیه می کنند که متأسفانه اینگونه پرسش ها اغلب با هدف های تدریس و محتوای برنامه درسی همخوانی نداشته و در تعارض هستند .

انجام امتحان نیز مانند هر کار عملی دیگر مستلزم تهیه طرح و نقشه ای پیش بینی شده و مبتنی بر اصول معین است . برای این که امتحان به عنوان وسیله موثری جهت هدایت یادگیری دانش آموزان و بهبود کیفیت آموزشی مورد استفاده قرار گیرد ، لازم است که در طرح ریزی آن به اصول زیر توجه نمود :

الف - سوالات امتحان باید تحقق هدف های آموزشی و تدریس را در هر یک از مواد درسی ارزیابی نماید .

ب - سوالات امتحان باید با محتوای مواد آموزشی سازگار باشند و مطالب مهم هر درس را که یادگیری آن ها در رابطه با دستیابی به اهداف درسی از اهمیت قابل قبولی برخوردار است ، شامل شوند .

ج - سوالات امتحانی باید توانایی و مهارت های دانش آموزان را در سطوح مختلف یادگیری و هدف های آموزشی مورد سنجش قرار دهند .

د - سوالات امتحانی باید از میان تمامی سوال های ممکن در رابطه با هدف های رفتاری و محتوای هر درس چنان انتخاب شوند که نمونه جامع یاز آموخته های دانش آموزان را مورد سنجش و اندازه گیری قرار دهند .

ه - سوالات امتحانی از نظر محتوایی باید برای دستیابی به هدف هایی که امتحان به منظور دستیابی به آن ها انجام می پذیرد ، مناسب باشند .

و - نوع و چگونگی هر یک از سوالات امتحانی می بایست با نوع رفتار مورد سنجش تناسب داشته باشد .

ز - امتحان باید به صورتی طرح ریزی شود که نتایج آن تا جایی که ممکن است از خطای اندازه گیری به دور باشند .

**سنجش یادگیری به کمک آزمون صحیح - غلط**

**هدف های یادگیری**

از خواننده انتظار می رود پس از مطالعه و یادگیری مطالب این فصل بتواند:

۱- آزمون صحیح- غلط را تعریف کند.

۲- ویژگی های سؤال های صحیح- غلط را توضیح دهد.

۳- انواع سؤال های صحیح- غلط را توضیح دهد و برای هر یک مثالی بزند.

۴- چگونگی تهیهی سؤال‌های صحیح- غلط را توضیح دهد.

۵- قواعد تهیهی سؤال‌های صحیح- غلط را شرح دهد.

۶- امتیازها و محدودیت‌های آزمون‌های صحیح- غلط را توضیح دهد.

۷- موارد استفاده‌ی آزمون‌های صحیح- غلط را توضیح دهد.

۸- برای هدف‌های یادگیری سطح پایین یک درس یک آزمون صحیح- غلط با رعایت کلیه‌ی پیشنهاد‌های داده شده در این فصل تهیه نماید.

یک دسته از روش‌های شناخته شده‌ی سنجش روش‌های عینی با گزینش پاسخ نام دارند. روش‌های عینی سنجش یادگیری از سه نوع صحیح- غلط، جورکردنی، و چندگزینه‌ای تشکیل می‌شوند.

### تعریف و انواع آزمون‌های صحیح- غلط

در آزمون‌هایی که صحیح- غلط<sup>۱</sup> نام دارند هر سؤال از یک جمله تشکیل می‌شود که آزمون شونده (دانش‌آموز یا دانشجو) آن را به یکی از صورت‌های صحیح- غلط، درست- نادرست، بله- نه، موافق- مخالف، واقعیت- عقیده، و مانند اینها پاسخ می‌دهد. اما از آنجا که این نوع سؤال‌ها غالباً به صورتی نوشته می‌شوند که آزمون‌شونده یا پاسخ‌دهنده به صورت صحیح یا غلط به آنها جواب می‌دهد، به سؤال‌های صحیح- غلط شهرت یافته‌اند. به نمونه‌های زیر توجه کنید:

نوع صحیح- غلط

صحيح غلط

✓

متن سؤال

ماده‌ی سبز موجود در گیاهان کلروفیل نام دارد.

نوع صحیح- غلط با اصلاح

راهنمایی: اگر گفته‌ی زیر درست است پاسخ "صحیح" را انتخاب کنید (دور آن را خط بکشید)، اگر غلط است "غلط" را انتخاب کنید و برای کلمه‌ای که زیر آن خط کشیده شده کلمه‌ی دیگری برگزینید که جمله را درست کند.

صحيح غلط

✓

متن سؤال

استفاده از نیروی بخار موجب انقلابی در حمل و نقل در قرن هفدهم شد.

<sup>۱</sup> -True- false

در واقع، این نوع سؤال ترکیبی از سؤال صحیح- غلط و سؤال کوتاه پاسخ است. قسمت اول که تصمیم‌گیری درباره‌ی درستی یا نادرستی سؤال است بخش صحیح- غلط و قسمت دوم که یافتن کلمه‌ی درست برای کلمه‌ی نادرست و نوشتن آن در جواب سؤال است بخش کوتاه پاسخ است.

نوع بله- نه

بله نه

متن سؤال

امکان دارد بدون یادگیری پیش‌نیازهای اساسی یک درس مطالب آن درس را یاد گرفت؟ ✓

نوع بله- نه با توضیح

یک نظرسنجی از ۵۰۰ نفر عضو حزب دموکرات و ۵۰۰ نفر عضو حزب جمهوری خواه به عمل آمده و در آن سؤال زیر پرسیده شده است، "حکومتی بهترین است که کمترین حکومت را می‌کند". نتیجه این نظرخواهی به شرح زیر اعلام شده است:

درصد	درصد	
مخالف	موافق	
٪۳۵	٪۱۲	مردان عضو حزب دموکرات
٪۱۴	٪۳	زنان حزب دموکرات
٪۱۲	٪۴۸	مردان حزب جمهوری خواه
٪۷	٪۲۸	زنان حزب جمهوری خواه
نه	بله	

✓ من معتقدم که این نظرسنجی ثابت می‌کند که اکثریت مردم طالب این

هستند که حکومت‌ها کمتر حکومت کنند. نظر من درست است؟

اگر جواب شما "نه" است، توضیح دهید چرا من اشتباه می‌کنم.

(نیتکو، ۲۰۰۱، ص ۱۳۶)

نوع موافق- مخالف

راهنمایی: با کشیدن دایره‌ای به دور "موافق" یا "مخالف" نظر خود را نسبت به گفته‌ی زیر مشخص کنید.

موافق مخالف

✓

متن سؤال

"هر گسی آن درود عاقبت کار که کشت".

نوع واقعیت - عقیده

بیان زیر یک حقیقت علمی است یا یک عقیده‌ی شخص؟

واقعیت علمی عقیده‌ی شخص

✓

متن سؤال

"در بعضی سیارگانی که به دور ستارگان دور دست می‌چرخند

نوعی زندگی هوشمندانه یافت می‌شود."

(لین و گرانلاند، ۲۰۰۰، ص ۱۷۹)

نوع چند صحیح - غلط (نوع خوشه‌ای)

کشور ایران هم مرز است با کشور:

صحیح غلط

✓

۱-عراق

✓

۲- قزاقستان

✓

۳- افغانستان

✓

۴- سوریه

سؤال‌های چند صحیح- غلط شبیه به سؤال‌های چند گزینه‌ای هستند، با این تفاوت که در سؤال‌های چندگزینه‌ای آزمون شونده از میان گزینه‌های پیشنهادی تنها یکی را برمی‌گزیند، اما در سؤال‌های چند صحیح- غلط لازم است به همه‌ی گزینه‌ها جواب بدهد. بنابراین، به هر یک سؤال چندگزینه‌ای، صرف‌نظر از تعداد گزینه‌های آن، تنها ۱ نمره تعلق می‌گیرد، اما برای هر سؤال چند صحیح- غلط، به تعداد انتخاب‌های آن، نمره منظور می‌شود. به همین دلیل، در تنظیم سؤال‌های یک آزمون به تنه یا صورت سؤال‌های چند صحیح- غلط شماره داده نمی‌شود، اما همه‌ی گزینه‌ها شماره‌گذاری می‌شوند.

نکته‌ی دیگری که در رابطه با سؤال‌های چند صحیح- غلط باید به یاد بسپارید این است که در هر مجموعه همه‌ی انتخاب‌های صحیح یا غلط باید به موضوع واحدی مربوط باشند. همچنین امکان دارد که همه‌ی انتخاب‌های یک مجموعه سؤال چند صحیح- غلط صحیح باشند یا غلط. به نمونه‌ی دیگری از این نوع سؤال‌ها توجه کنید.

مدافعان برده‌داری در کشور آمریکا پیش از جنگ داخلی با توسل به گفته‌های زیر از عقیده‌ی خود دفاع می‌کردند. هر یک از گفته‌ها را بررسی کنید و ببینید که با موازین دموکراسی همخوانی دارد یا نه؟

### صحیح غلط

- ✓ ۱- برده‌داری کار درستی است زیرا در تمام طول تاریخ وجود داشته است.
- ✓ ۲- برده‌داری کار درستی است زیرا مورد قبول کسانی بوده است که قانون اساسی آمریکا را نوشته‌اند.
- ✓ ۳- برده‌داری کار درستی است زیرا مورد قبول فیلسوف سرشناس یونان باستان، ارسطو، بوده است.

(ایبل، ۱۹۷۲، ص ۱۷۴)

### تهیه‌ی سؤال‌های صحیح- غلط

صاحب‌نظران سنجش در رابطه با تهیه‌ی سؤال‌های صحیح- غلط قواعد مختلفی را به دست داده‌اند. ما در اینجا قواعدی را که اگر رعایت شوند بر کیفیت سؤال‌های صحیح- غلط می‌افزایند مورد بحث قرار می‌دهیم. اما پیش از معرفی این قواعد به پیشنهادهای رابرت ایبل (۱۹۷۲) یکی از طرفداران آزمون‌های صحیح- غلط در این باره توجه کنید.

ایبل در رابطه با تهیه‌ی سؤال‌های صحیح- غلط گفته است ابتدا قسمتی از موضوع درسی را که قابل طرح در یک پاراگراف باشد انتخاب کنید. "این قسمت از موضوع یا این جزء دانش باید به سادگی قابل توصیف در یک پاراگراف واحد مانند آنچه در کتاب‌های درسی خوب یافت می‌شود باشد" (ص ۱۷۲). سپس درباره‌ی آن موضوع جمله یا گزاره‌ای با ویژگی‌های زیر بنویسید.

۱- حاوی اندیشه‌ی مهمی باشد.

#### پاراگراف

گرفت یا گرفتگی زمانی رخ می‌دهد که یک جرم نجومی مانع تابش بخشی از یا تمامی نور جرم نجومی دیگری می‌شود. در خورشید گرفتگی، ماه در فاصله‌ی بین خورشید و زمین قرار می‌گیرد و مانع رسیدن نور خورشید به زمین می‌شود. در ماه گرفتگی، سایه‌ی زمین بر روی ماه می‌افتد و مانع بازتابش نور خورشید بر سطح کره‌ی ماه می‌شود.

گزاره



گرفتگی (مثلاً خورشیدگرفتگی یا ماه‌گرفتگی) در منظومه‌ی شمسی بر اثر افتادن سایه‌ی یک جرم نجومی بر جرم نجومی دیگر حاصل می‌شود.

شکل ۱-۷ یک پاراگراف مطلب و گزاره‌ی استخراج شده از آن (ایبل، ۱۹۷۲، ص ۱۷۳)

۲- به وسیله‌ی اشخاص صاحب‌نظر قابل دفاع باشد.

۳- توسط کسانی که از موضوع مطرح شده اطلاع کافی ندارند قابل جواب دادن نباشد؛ یعنی براساس اطلاعات عمومی نتوان به آن پاسخ درست داد.

در شکل ۱-۷ به یک نمونه پاراگراف و جمله یا گزاره‌ی حاصل از آن توجه کنید.

ایبل (۱۹۷۲) می‌گوید سؤال نویسان باتجربه به ندرت گزاره‌ی را که سؤال از آن استخراج می‌شود عملاً می‌نویسند، بلکه به هنگام نوشتن سؤال آن را در ذهن خود دارند. با این حال، وی توصیه می‌کند، هنگامی که در نوشتن یک سؤال صحیح- غلط به مشکل برمی‌خورید از خود بپرسید که "این سؤال قرار است چه چیزی را بسنجد؟" یا "گزاره‌ی که براساس آن من می‌خواهم سؤال را بنویسم چیست؟"

گام بعدی تبدیل گزاره‌ی حاصل از پاراگراف به سؤال صحیح- غلط است. برای این منظور، ۵ پیشنهاد ارائه شده است. این پیشنهادها همراه با نمونه‌ی سؤال‌های حاصل از پاراگراف و گزاره‌ی شکل ۱-۷ در زیر توضیح داده می‌شوند.

**پیشنهاد ۱-** اندیشه‌ی اصلی را در قالب کلمات دیگری تغییر دهید و در صورت صحیح یا غلط سؤال را بنویسید.

#### صورت غلط

#### صورت صحیح

وقتی که یک جرم نجومی مانند ماه مانع تابش قسمتی وقتی که نور ستاره‌ی مانند خورشید از سطح سیاره‌ی از نور ستاره‌ی مانند خورشید بر سطح سیاره‌ی مانند زمین به سطح جرم دیگری مانند ماه منعکس زمین می‌شود، می‌گویند گرفتن رخ داده است. می‌شود، می‌گویند گرفتگی رخ داده است.

**پیشنهاد ۲-** تنها بخشی از اندیشه‌ی اصلی را به صورت صحیح یا غلط بنویسید نه تمام آن را.

#### صورت غلط

#### صورت صحیح

همه گرفتها از سایه تشکیل می‌شوند. اگر پرتوهای نور قابل خم شدن نباشند، گرفتها رخ نمی‌دهند.

**پیشنهاد ۳-** اندیشه‌ی اصلی را به اندیشه‌ی دیگری تغییر دهید.

## صورت صحیح

## صورت غلط

ماه گرفتگی تنها زمانی رخ می‌دهد که ماه قرص کامل ماه گرفتگی تنها زمانی رخ می‌دهد که ماه نو باشد. باشد.

پیشنهاد ۴- پیامدهای اندیشه‌ی اصلی را مشخص کنید.

## صورت صحیح

## صورت غلط

پیش‌بینی گرفته‌ها مستلزم داشتن اطلاع درباره‌ی حرکات پیش‌بینی گرفته‌ها مستلزم داشتن اطلاع درباره‌ی شیب مداری اجرام موردنظر است. محور زمین است. پیشنهاد ۵- اثر موقعیت‌های متفاوت (حتی غیرممکن) را استنتاج کنید.

## صورت صحیح

## صورت غلط

اگر جرم‌های نجومی درخشان مانند ماه وجود نداشتند در حضور ستارگان درخشان گرفته‌ها رخ نمی‌دهند. گرفته‌ها قابل مشاهده نبودند.

نمونه‌ی سؤال‌های بالا نشان می‌دهند که از یک اندیشه‌ی واحد می‌توان سؤال‌های صحیح- غلط متنوعی درست کرد. البته سؤال‌های بالا تنها سؤال‌هایی نیستند که از پاراگراف فوق درباره‌ی خسوف یا کسوف می‌توان نوشت. مسلماً سؤال‌های بیشتری نیز می‌توان درباره‌ی آن طرح کرد. چیزی که نباید فراموش شود این است که "ماده‌های صحیح- غلط صرفاً به آزمون یادآوری مطالب خلاصه نمی‌شوند، بلکه می‌توانند مسائل تازه را در اختیار آزمون‌شونده قرار دهند و او را وادار به تفکر و حل مسئله نمایند" (ایبل، ۱۹۷۲، ص ۱۷۵)

درباره‌ی کیفیت صورت درست و صورت غلط سؤال‌های صحیح - غلط، ایبل (۱۹۷۲) اظهار می‌دارد که آزمون‌ساز باید بتواند درباره‌ی صورت صحیح سؤال بگوید: "درستی این بیان بر همه کس آشکار نیست، اما من می‌توانم شواهدی دال بر درستی آن به دست دهم." همچنین درباره‌ی صورت غلط سؤال باید بتواند بگوید: "این بیان یک غلط آشکار است به نظر می‌رسد، اما من می‌توانم شواهدی دال بر اینکه آن واقعاً غلط است به دست دهم. اگر سؤال‌های شما این آزمون را با موفقیت بگذرانند احتمالاً سؤال‌های صحیح- غلط خوبی هستند" (ص ۱۷۵).

## قواعد تهیه‌ی سؤال‌های صحیح - غلط

افزون بر راهنمایی‌های بالا، برای تهیه سؤال‌های صحیح- غلط، از سوی متخصصان اندازه‌گیری و سنجش پیشرفت تحصیلی، قواعدی پیشنهاد شده است، ما در اینجا به توضیح آنها می‌پردازیم.

۱- موضوع‌های مهم و در ارتباط با هدف‌های آموزشی را در سؤال‌های خود بیاورید. سعی کنید هر یک از سؤال‌های شما یک موضوع مهم مربوط به یک هدف آموزشی را بسنجد. هر چند که این پیشنهاد مخصوص سؤال‌های صحیح-غلط نیست و به همه‌ی انواع سؤال‌های پیشرفت تحصیلی مربوط می‌شود، با این حال هنگام نوشتن سؤال‌های صحیح-غلط بیشتر به آنها توجه کنید، زیرا غالباً در این‌گونه سؤال‌هاست که نکات جزئی و کم‌اهمیت امکان طرح شدن می‌یابند. بهترین وسیله‌ی رعایت این پیشنهاد استفاده از جدول مشخصات است.

#### نمونه‌ی ضعیف:

صحیح غلط

✓

بعضی چیزها در بعضی چیزهای دیگر حل می‌شوند.

#### نمونه‌ی بهتر:

صحیح غلط

✓

در یک لیتر آب گرم نمک بیشتری حل می‌شود تا در یک لیتر آب سرد

(اپیل، ۱۹۷۲، ص ۱۷۶)

سؤال ضعیف بالا یک بیان کلی است و معرف هیچ نکته‌ی مهمی نیست. در عوض، نمونه‌ی سؤال بهتر فهم یک رابطه‌ی مهم را می‌سنجد.

۲- سؤال‌هایی طرح کنید که درک و فهم را بسنجد نه یادآوری کلمه به کلمه‌ی مطالب را. تکرار مطالب کتاب در سؤال‌های امتحانی تنها محفوظات دانش‌آموزان را می‌سنجد. برای سنجش درک و فهم باید مطالب را بازسازی کنید و آنها را در قالب‌های دیگر بجز آنچه در کتاب آمده است بنویسید.

#### نمونه‌ی ضعیف:

صحیح غلط

✓

در یک مثلث قائم‌الزاویه مربع وتر برابر است با مجموع مربع‌های دو ضلع دیگر

#### نمونه‌ی بهتر:

صحیح غلط

✓

اگر وتر یک مثلث قائم‌الزاویه‌ی متساوی‌الساقین ۷ سانتیمتر است، هر یک از دو ضلع

دیگر آن باید از ۵ سانتیمتر بزرگتر باشد.

(ایبل، ۱۹۷۲، ص ۱۷۷)

سؤال ضعیف بالا دربرگیرنده‌ی یک جمله‌ی آشنای کتاب‌های ریاضی است و لذا صرفاً اطلاعات حفظی دانش‌آموزان را می‌سنجد. در مقابل، نمونه‌ی سؤال بهتر که از راه بازنویسی همان موضوع ولی با بیان دیگری درست شده می‌تواند درک و فهم دانش‌آموز را از موضوع موردنظر بسنجد. در زیر به نمونه‌های دیگری از همین نوع سؤال‌ها توجه کنید:

نمونه‌ی ضعیف:

صحیح غلط

✓

برای هر عملی یک عکس‌العمل برابر ولی مخالف با آن وجود دارد.

نمونه‌ی بهتر:

صحیح غلط

✓

وقتی که دستی دری را با نیروی معینی می‌فشارد، در با همان نیرو به دست فشار وارد می‌آورد.

(ایبل، ۱۹۷۲، ص ۱۷۶)

۳- سؤال‌هایی طرح کنید که درست یا غلط بودن آنها قابل دفاع باشد. هر سؤال باید به گونه‌ای باشد که نوشته شود که کسانی که از موضوع آن اطلاع کافی دارند به طور آشکار آن را درست یا غلط بدانند. به سخن دیگر، سؤال باید برای پاسخ‌دهنده‌ی آگاه از موضوع سؤال کاملاً درست یا کاملاً غلط باشد.

نمونه‌ی ضعیف:

صحیح غلط

✓

ستارگان از خود نورهایی تولید می‌کنند که سوسو می‌زند.

نمونه‌ی بهتر:

صحیح غلط

✓

سوسو زدن نور ستارگان به علت حرکت جو زمین است.

(ایبل، ۱۹۷۲، ص ۱۷۷)

یک شخص مطلع از موضوع، می‌تواند سؤال ضعیف بالا را با توجه به دلایل زیر مورد تردید قرار دهد: نور منتشره از ستاره سوسو نمی‌زند. نور ستاره نسبتاً ثابت است، اما به علت تغییرات موجود در جو زمین نوری که از ستارگان به چشمان ما می‌رسد سوسو می‌زند. "سؤال ضعیف و غیرقابل دفاع بالا یا ناشی از دانش اندک نویسنده سؤال یا به سبب بی‌دقتی او در نوشتن سؤال است" (ایبل، ۱۹۷۲، ص ۱۷۸).

۴- از جمله‌های پیچیده و چند قسمتی استفاده نکنید. جمله‌های پیچیده و طولانی حواس آزمون‌شوندگان را پرت می‌کنند و باعث می‌شوند که توجه او به جای نکته‌ی اصلی سؤال به مسائل فرعی جلب شود.

#### نمونه‌ی ضعیف:

##### صحيح غلط

- ✓ جین آستین، داستان سرای آمریکایی، متولد ۱۷۹۰، یک نویسنده‌ی پرکار بود و شهرتش به خاطر داستان معروف او غرور و تعصب است که در سال ۱۸۲۰ انتشار یافت.

#### نمونه‌ی بهتر:

##### صحيح غلط

- ✓ شهرت جین آستین به خاطر داستان غرور و تعصب است.

(پین، ۲۰۰۳، ص ۱۹۴)

۵- در هر سؤال بیشتر از یک موضوع قرار ندهید. هر سؤال باید تنها یک موضوع مهم را شامل شود قرار دادن بیش از یک موضوع در یک سؤال، به ویژه اگر یکی درست و دیگری غلط باشد، وسیله‌ی مناسبی برای سنجش یادگیری نیست. پوفام (۲۰۰۲) در این باره گرفته است "حضور دو مفهوم در یک سؤال، حتی اگر هر دو صحیح یا هر دو غلط باشند، باعث سردرگمی دانش‌آموز می‌شود، و نتایجی به دست می‌دهد که منجر به استنباط‌های غلطی درباره‌ی دانش‌آموزان خواهد شد" (ص ۱۳۱). نول<sup>۱</sup> و اسکانل<sup>۲</sup> (۱۹۷۲، ص ۲۱۴)، برای این مورد مثال‌های زیر را آورده‌اند:

##### صحيح غلط

##### متن سؤال

آب و هوای انگلستان، به سبب نفوذ گلف استریم اقیانوس اطلس، از بیشتر

<sup>۱</sup> -Noll

<sup>۲</sup> - Scannell

کشورهای هم عرض و جغرافیایی آن سردتر است.

بیان فوق مبهم است، زیرا دو موضوع را دربرمی‌گیرد که یکی از آنها درست و دیگری غلط است. بهتر آن است آن را به دو سؤال مستقل مانند نمونه‌های زیر تبدیل کنیم:

صحیح    غلط

✓ ۱- آب و هوای انگلستان از بیشتر کشورهای هم عرض جغرافیایی آن سردتر است.

✓ ۲- آب و هوای انگلستان تحت تأثیر گلف استریم اقیانوس اطلس قرار دارد.

نیتکو (۲۰۰۱، ص ۱۴۱) نیز در ارتباط با این سفارش که در سؤال‌های صحیح، غلط بیشتر از یک موضوع فرار ندهید مثال‌های زیر را ذکر کرده است:

نمونه‌ی ضعیف:

صحیح    غلط

✓ رود مونونگاهلا به سوی شمال جریان می‌یابد و در کولومبوس به رود الگه‌نی می‌پیوندد و در آنجا با همدیگر رود اوهایو را به وجود می‌آورند.

نمونه‌ی بهتر:

صحیح    غلط

✓ رود مونونگاهلا و رود الگه‌نی به هم می‌پیوندند و رود اوهایو را به وجود می‌آورند.

نیتکو (۲۰۰۱) در رابطه با سؤال ضعیف بالا می‌گوید دانش‌آموز ممکن است به یکی از دلایل موجه زیر برای این سؤال جواب غلط را انتخاب کند: (۱) ممکن است (اشتباهاً) فکر کند که رود مونونگاهلا به سوی شمال جریان ندارد. (۲) ممکن است نداند که به هم پیوستن رودخانه‌ها در پیتسبورگ اتفاق می‌افتد. (۳) ممکن است چیزی درباره‌ی این سه رود نداند. بنابراین، دانش‌آموز به سؤال درست جواب می‌دهد بدون اینکه اطلاعات لازم برای درست جواب دادن را داشته باشد.

پس، بهتر است از گنجاندن بیشتر از یک موضوع در هر سؤال پرهیز شود. حتی اگر دو موضوع هر دو صحیح یا هر دو غلط باشند. نیز دشواری به وجود می‌آورد، زیرا اگر یک آزمون‌شونده از هیچ یک از دو موضوع آگاه نباشد و آزمون

شونده‌ی دیگری از یکی از آن دو مطلع باشد، نتیجه‌ی سؤال برای هر دو نفر یکی خواهد بود. تنها استثناء برای این قاعده زمانی است که بخواهید رابطه‌ی علت و معلولی بین دو موضوع را بسنجید.

۶- در صورت امکان از زبان کمی و دقیق استفاده کنید نه از زبان کیفی و غیردقیق. کلماتی مانند اندکی، بسیاری، کوچک، بزرگ، و مهم مبهم و قابل تفسیرند مگر اینکه با یک معیار مقایسه به کار بروند.

#### نمونه‌ی ضعیف:

در انتخابات ریاست جمهوری سال ۱۹۹۲، بسیاری از مردم به بیل کلینتون رأی دادند.

#### نمونه‌ی بهتر:

در انتخابات ریاست جمهوری سال ۱۹۹۲، بیشتر از ۶۰ درصد رأی‌دهندگان به بیل کلینتون رأی دادند. (پین، ۲۰۰۳، ص ۱۹۲)

۷- تا آنجا که ممکن است از طرح سؤال‌های منفی پرهیز کنید. جواب دادن به سؤال‌های منفی معمولاً دشوار است. پوفام (۲۰۰۲) می‌گوید فرض کنید در یک سؤال صحیح- غلط از شما خواسته شده که درباره‌ی درست یا غلط بودن بیان زیر تصمیم بگیرید:

#### صحیح غلط

✓

جامعه‌ی جهانی بلافاصله پس از پایان جنگ جهانی دوم تشکیل نشد.

جواب درست این سؤال صحیح است، زیرا جامعه‌ی ملل پیش از جنگ جهانی دوم تشکیل شده بود. اما وجود نشد در سؤال بعضی دانش‌آموزان را دچار سردرگمی می‌کند. حتی اگر بدانند که جامعه‌ی ملل پیش از شروع جنگ جهانی دوم وجود داشته، احتمالاً به سؤال جواب غلط خواهند داد. بهتر است سؤال را به یکی از صورت‌های زیر درآورید:

#### صحیح غلط

✓

۱- جامعه‌ی جهانی پیش از شروع جنگ جهانی دوم تشکیل شده بود.

✓

۲- جامعه‌ی جهانی بلافاصله پس از پایان جنگ جهانی دوم تشکیل شد.

با توجه به توضیحات بالا، تا آنجا که ممکن است در نوشتن سؤال‌های صحیح- غلط از کاربرد عبارت‌های منفی پرهیز کنید.

۸- از کلمات راهنما با اشاره‌کننده به جواب درست استفاده نکنید. بعضی وقت‌ها کاربرد کلمه‌هایی مانند ممکن است، همیشه، همه، هرگز و هیچ صحیح یا غلط بودن یک سؤال را کاملاً آشکار می‌سازد. به گونه‌ای که پاسخ‌دهنده‌ی ناآگاه از موضوع مورد آزمون به سؤال جواب درست می‌دهد. به نمونه‌های زیر نگاه کنید:

صحیح غلط

- ✓ ۱- همه‌ی شهرهای بزرگ با قطار به هم مرتبط‌اند.
- ✓ ۲- هیچ نظام آموزشی با کمک‌های محلی حمایت نمی‌شود.
- ✓ ۳- امکان دارد که هر زاویه‌ای را دونیم کرد.
- ✓ ۴- غیرممکن است که یک کیلومتر را در ۲ دقیقه دوید.
- ✓ ۵- انقلاب‌ها همیشه به هدف‌های مطلوب اجتماعی انجامیده‌اند.
- ✓ ۶- در یک دموکراسی جنگ هرگز توجیه نمی‌شود.

(ساکس، ۱۹۹۷، ص ۸۷)

استفاده از کلمه‌های همه، هیچ، غیرممکن، همیشه، و هرگز، در سؤال‌های ۱، ۲، ۴، ۵ و ۶ بالا غلط بودن آنها را آشکار می‌سازد. در مقابل، آزمون‌شونده سؤال ۳ را نیز به راحتی صحیح تشخیص می‌دهد. زیرا کمتر فعالیتی غیرممکن است. بنا به گفته‌ی ساکس (۱۹۹۷)، "در حالی که استفاده از کلمه‌های راهنما ممکن است دانش‌آموزان را تشویق کند تا به صورت آزمون‌دهندگان مجرب درآیند، اما کمک زیادی به سنجش هدف‌های درس نمی‌کند" (ص ۸۷). در زیر به نمونه‌های دیگر از این سؤال‌ها توجه کنید:

نمونه‌ی ضعیف:

صحیح غلط

- ✓ یک "برد مادر"<sup>۱</sup> معیوب ممکن است باعث خاموش شدن کامپیوتر شخصی بشود.

نمونه‌ی بهتر:

صحیح غلط

- ✓ یک "برد مادر" معیوب باعث خاموش شدن کامپیوتر شخصی می‌شود.

(پین، ۲۰۰۳، ص ۱۹۱)

---

<sup>۱</sup> - mother board



پین (۲۰۰۳) می‌گوید آزمون‌شوندگان مجرب با توجه به کلمه‌های "ممکن است" به راحتی پاسخ سؤال ضعیف بالا را پیدا خواهند کرد. البته اگر در طول آزمون به تساوی از این‌گونه کلمات در سؤال‌های صحیح و غلط استفاده شود احتمال حدس زدن جواب درست سؤال با توجه به آنها کاهش خواهد یافت.

۹- سعی کنید تعداد سؤال‌های غلط بیشتر از تعداد سؤال‌های درست باشد. ایبل (۱۹۷۲) در توجیه این پیشنهاد می‌گوید دانش‌آموزان وقتی که در انتخاب بین جواب صحیح و جواب غلط تردید می‌کنند ترجیح می‌دهند جواب صحیح را انتخاب کنند، چیزی که به آمایه‌ی توافق<sup>۱</sup> شهرت یافته است. "چندین پژوهشگر دریافته‌اند که بیانات غلط تا حدی بهتر از بیانات صحیح بین دانش‌آموزان با پیشرفت بالا و دانش‌آموزان با پیشرفت پایین تمیز قائل می‌شوند" (ص ۱۸۰). ایبل برای سؤال‌های غلط رقم ۶۷ درصد و برای سؤال‌های درست رقم ۳۳ درصد را پیشنهاد داده است.

تمایل به دادن پاسخ به سؤال‌های صحیح- غلط در یک جهت خاص (صحیح یا غلط) آمایه‌ی پاسخ<sup>۲</sup> نام دارد. بنا به تعریف، آمایه‌ی پاسخ به گرایش آزمون‌شونده در دادن پاسخ به سؤال‌های با توجه به شکل سؤال نه محتوای آن گفته می‌شود. منطق حکم می‌کند که اگر تعداد سؤال‌های صحیح و تعداد سؤال‌های غلط در یک آزمون صحیح- غلط مساوی باشند از مشکل آمایه‌ی پاسخ جلوگیری شود، "اما شواهدی موجودند که نشان می‌دهند سؤال‌های غلط قدرت تمیز بیشتری دارند و شاید نسبت ۶۰ به ۴۰ به نفع سؤال‌های غلط بهترین پیشنهاد در این باره باشد" (پین، ۲۰۰۳، ص ۱۹۱).

در مقابل پیشنهادهای بالا، رینالدز و لیوینگستون (۲۰۱۲) تعداد مساوی سؤال‌های صحیح و غلط را سفارش داده‌اند. آنان در این باره گفته‌اند:

برخی نویسندگان قبلی توصیه کرده بودند که سؤال‌های صحیح- غلط به گونه‌ای نوشته شوند که جواب‌های صحیح ۶۰٪ کل جواب‌ها باشند. این پیشنهاد به عنوان وسیله‌ای برای افزایش یادگیری داده شد زیرا اکثریت بیاناتی که در این سؤال‌ها آزمون‌شونده آنها را می‌خواند درست هستند. این عمل در موارد خیلی خاص مؤثر است و در آزمون‌های عملکرد نوعی که به جای صحیح و غلط از بله و نه استفاده می‌شود قابل کاربست نیست. ولی همان بهتر که از تعداد تقریباً مساوی پاسخ‌های صحیح و غلط استفاده شود. (ص ۲۰۴)

۱۰- طول سؤال‌های صحیح و سؤال‌های غلط را تقریباً هم‌اندازه انتخاب کنید. کسانی که در نوشتن سؤال‌های صحیح- غلط مهارت کافی ندارند، غالباً سؤال‌های صحیح را از سؤال‌های غلط طولانی‌تر می‌نویسند. این کار باعث

---

<sup>۱</sup> -acquiescence set

<sup>۲</sup> -response set

می‌شود که دانش‌آموزان متوجه مطلب بشوند و بدون اینکه از موضوع طرح شده در سؤال اطلاع داشته باشند و به آن درست پاسخ دهند. بنابراین، سعی کنید سؤال‌های صحیح و سؤال‌های غلط را هم اندازه بنویسید یا اگر به ناچار طول یکی از دیگری بیشتر می‌شود طول آنها را به تناوب و به یک نسبت کم و زیاد کنید.

### امتیازها و محدودیت‌های آزمون‌های صحیح- غلط

بر سر استفاده از آزمون‌های صحیح- غلط برای سنجش میزان یادگیری دانش‌آموزان و دانشجویان در میان متخصصان سنجش اختلاف نظر وجود دارد. برخی استفاده از آن را مفید می‌دانند. از جمله کسانی که این آزمون‌ها را مفید تشخیص داده‌اند، ایبل (۱۹۷۲) است. او، در دفاع از این آزمون‌ها، گفته است "دلیل عمده‌ی استفاده از سؤال‌های صحیح- غلط این است که آنها وسیله‌ی ساده و مستقیمی برای اندازه‌گیری بازدهی اساسی آموزش رسمی هستند" (ص ۱۶۴). پین (۲۰۰۳) نیز در این باره می‌گوید:

سؤال صحیح- غلط یک شیوه‌ی بالقوه ارزشمند برای جمع‌آوری اطلاعات است... استفاده از سؤال‌های صحیح- غلط دارای امتیازهایی است که مهم‌ترین آنها کارآمدی است. معلم می‌تواند تعداد زیادی از این سؤال‌ها را در یک جلسه‌ی امتحانی مورد استفاده قرار دهد. این کار امکان نمونه‌گیری از دامنه‌ی بزرگی از مطالب درسی را به منظور سنجش دانش یادگیرنده فراهم می‌آورد. (ص ۱۹۰).

به طور کلی، می‌توان امتیازها یا محاسن سؤال‌های صحیح- غلط را به صورت زیر خلاصه کرد:

۱-تهیه‌ی این نوع سؤال‌ها کار نسبتاً آسانی است.

۲- تصحیح جواب آزمون‌شوندگان به این نوع سؤال‌ها به سادگی و سرعت و با عینیت انجام‌پذیر است.

۳- به این سبب که می‌توان تعداد زیادی سؤال صحیح- غلط را در یک جلسه‌ی امتحان مورد استفاده قرار داد، بخش گسترده‌ای از محتوای درسی را می‌توان به این سؤال‌ها سنجش کرد.

مخالفان آزمون‌های صحیح- غلط نیز ایرادهایی به آن وارد آورده‌اند. از جمله هارپر و هارپر<sup>۱</sup> (۱۹۹۰) و لین و گرانلاند (۲۰۰۰) گفته‌اند، به‌خلاف تصور مدافعان آزمون‌های صحیح- غلط، تهیه‌ی این آزمون‌ها کار آسانی نیست. "ساختن سؤال‌های صحیح- غلط که بازده‌های مهم یادگیری را بسنجند نیازمند مهارت فراوان است" (لین و گرانلاند، ۲۰۰۰، ص ۱۸۱). هارپر و هارپر (۱۹۹۰) نیز گفته‌اند، "تجارب دردناک نشان داده‌اند که سؤال‌های صحیح- غلط به آن سادگی که به نظر می‌رسد نیستند و تهیه‌ی آنها غالباً به مشکل می‌انجامد" (ص ۲۱۳).

در رد امتیاز دیگر عنوان شده برای سؤال‌های صحیح- غلط، یعنی اینکه چون در یک جلسه‌ی امتحانی تعداد زیادی از آنها را می‌توان مورد استفاده قرار داد پس حجم بزرگی از محتوای درس را به وسیله‌ی این سؤال‌ها می‌توان در یک

---

<sup>۱</sup>-Harper

جلسه‌ی امتحانی سنجش کرد، نیز نظرهایی داده شده است. از جمله گفته شده که همه‌ی موضوع‌های درسی را نمی‌توان به صورت عبارات‌های صحیح یا غلط درآورد. بنا به گفته‌ی لین و گراندلاند (۲۰۰۰)، "در همه‌ی موضوع‌های درسی، از جمله علوم اجتماعی، زمینه‌هایی وجود دارند که درباره‌ی آنها نوشتن بیانات مطلقاً درست یا مطلقاً غلط که قابل تبدیل به سؤال‌های صحیح- غلط باشند امکان‌پذیر نیست. تنها موضوع‌های کم‌اهمیت را می‌توان به این‌گونه بیانات تبدیل کرد" (ص ۱۲۸).

ایراد دیگری که به سؤال‌های صحیح- غلط گرفته‌اند میزان بالای حدس‌پذیری آنهاست. به طور متوسط، نصف همه‌ی آزمون‌شوندگان می‌توانند تنها با حدس زدن به یک سؤال صحیح- غلط درست جواب دهند. با یک آزمون‌شونده صرفاً با حدس زدن می‌تواند ۵۰ درصد سؤال‌های صحیح- غلط یک آزمون را درست جواب دهد. حتی آزمون‌شوندگان بی‌اطلاع خوش اقبال ممکن است بیشتر از نصف سؤال‌ها را درست جواب دهند.

هارپرو هارپر (۱۹۹۰) در ارتباط با امکان حدس‌پذیری بالای آزمون‌های صحیح- غلط گفته‌اند آمار ساده‌ی نظریه‌ی احتمال نشان می‌دهد که از هر ۱۰۰ آزمون‌شونده یک نفر خوش‌اقبال می‌تواند به گونه‌ی زیر کسب نمره نماید:

از هر ۱۰ سؤال به ۸ سؤال درست جواب دهد. (۰.۸۰٪)

از هر ۵۰ سؤال به ۳۳ سؤال درست جواب دهد. (۰.۶۶٪)

از هر ۱۰۰ سؤال به ۶۲ سؤال درست جواب دهد. (۰.۶۳٪)

از هر ۲۰۰ سؤال به ۱۱۶ سؤال درست جواب دهد. (۰.۵۸٪)

افزون بر ایرادهای بالا، عیب بزرگ آزمون‌های صحیح- غلط این است که برای سنجش هدف‌های آموزشی و بازده‌های مهم یادگیری محدودیت دارند. لین و گراندلاند (۲۰۰۰) در این باره گفته‌اند:

سؤال‌های صحیح- غلط برای سنجش بازده‌های یادگیری فراتر از دانش چندان مفید نیستند. تنها استثناهای این قاعده تمیز داده بین واقعیت از عقیده و تشخیص رابطه‌ی علت و معلولی است. این دو بازده یادگیری احتمالاً مهم‌ترین بازده‌هایی هستند که با این نوع سؤال‌ها سنجش پذیرند. بسیاری از بازده‌های دیگر یادگیری را می‌توان با گونه‌های دیگر آزمون‌های انتخاب پاسخ، به ویژه نوع چندگزینه‌ای، به نحو مؤثرتری اندازه‌گیری کرد. (ص ۱۸۲)

پین (۲۰۰۳) نیز درباره‌ی استفاده از سؤال‌های صحیح- غلط گفته است استفاده از این سؤال‌ها در موقعیت‌هایی مناسب است که در آنها تنها دو پاسخ وجود دارند. وی به عنوان نمونه سؤال زیر را مثال زده است.

صحیح غلط

✓

درهای خروجی اضطراری مدرسه باید به سمت درون باز شوند.

(پین، ۲۰۰۳، ص ۱۹۰)

لین و گراند (۲۰۰۰) استفاده از سؤال‌های صحیح- غلط را به دو موقعیت محدود کرده‌اند. یکی موقعیتی که در آن تنها دو شق چپ- راست، بیشتر- کمتر، بالا- پایین، و مانند اینها وجود دارند. موقعیت دیگر زمانی است که آزمون‌شونده باید بین عقیده از واقعیت، علت از معلول، خرافه از علم، مربوط از نامربوط، معتبر از غیرمعتبر، و مانند اینها تمیز قائل شود.

با توجه به توضیحات بالا، لازم است کاربرد سؤال‌های صحیح- غلط به حداقل کاهش یابد، و عمدتاً برای سنجش هدف‌های یادگیری سطوح پایین از آنها استفاده شود. علاوه بر آزمون‌های صحیح- غلط، آزمون‌ها و ابزارهای متنوع دیگری وجود دارند که به نحو شایسته‌تری هدف‌های سطح بالای یادگیری را می‌سنجند.

### موارد استفاده از آزمون‌های صحیح- غلط

در این قسمت نمونه‌هایی از کاربرد آزمون‌های صحیح- غلط را که از سوی متخصصان برای سنجش هدف‌های گوناگون آموزشی تهیه شده‌اند معرفی می‌کنیم.

#### یادآوری دانش امور واقعی

صحیح غلط

✓

۱- میان بهترین شاخص‌گرایش مرکزی است.

جنگ جهانی اول:

صحیح غلط

✓

۲- به تأسیس جامعه‌ی ملل انجامید.

✓

۳- منجر به تقسیم آلمان به دو قسمت شد.

✓

۴- به تأسیس دولت مستقل اسرائیل انجامید.

✓

۵- تنها جنگی است که در آن از گازهای سمی به مقیاس وسیعی استفاده شد.

#### یادآوری دانش مفهومی

صحیح غلط

✓

۶- می‌توان به کمک تقویت منفی رفتار مطلوب یک دانش‌آموز را افزایش داد.

#### یادآوری دانش روندی

صحیح غلط

✓

۷- برای محاسبه‌ی ضریب همبستگی یک سؤال با کل آزمون از روش همبستگی دو

نقطه‌ای رشته‌ای استفاده می‌کنیم.

### یادآوری دانش فراشناختی

صحيح غلط

✓

۸- یکی از روش‌های موفق یادگیری کتاب‌های درسی علمی این است که ابتدا با یک مرور مقدماتی با موضوع‌های مختلف یک آشنایی کلی کسب کنیم و بعد برای یادگیری کامل آنها یک برنامه‌ی زمان‌بندی تهیه نماییم.

### فهمیدن دانش مفهومی

صحيح غلط

✓

۹- با افزودن تعدادی سؤال مشابه سؤال‌های یک آزمون به آن آزمون هم‌روایی و هم‌پایایی آزمون بالا می‌رود.

✓

۱۰- تفاوت بین نمره‌های خام منطبق با صدک‌های چهل و پنجم و پنجاه و پنجم از تفاوت بین نمره‌های خام منطبق با صدک‌های پنجم و پانزدهم کمتر است.

### فهمیدن دانش روندی

صحيح غلط

✓

۱۱- یکی از راه‌های خوب استخراج فرضیه‌های علمی مراجعه به نظریه‌های موجود در زمینه‌ی مورد پژوهش است.

✓

۱۲- در استفاده از ضریب همبستگی در تحلیل داده‌های پژوهشی، هم می‌توان رابطه‌ی زمانی بین متغیرها و هم رابطه‌ی علت و معلولی را تعیین کرد.

### کار بستن دانش امور واقعی

صحيح غلط

✓

۱۳- اگر در یک اتاق در بسته در یک یخچال روشن را باز بگذاریم درجه‌ی حرارت اتاق کاهش خواهد یافت.

### کار بستن دانش روندی

برای نمره‌های ۸، ۹، ۱۰، ۱۱ و ۱۲

صحيح غلط

- ✓ ۱۴- میانگین برابر است با ۱۰
- ✓ ۱۵- میانه برابر است با ۱۰
- ✓ ۱۶- نما برابر است با ۱۰
- ✓ ۱۷- دامنه برابر است با ۷
- ✓ ۱۸- با در دست داشتن فرمول  $\frac{PV}{T} = K$ ، اگر  $K = 2$ ،  $P = 5$  و  $T = 50 \cdot C$  باشد، آن‌گاه  $V = 20$ .
- ✓ ۱۹- یک آزمون صحیح- غلط ۳۰ سؤالی در دست داریم که ضریب پایایی آن  $0/60$  است. برای افزودن برپایایی این آزمون می‌خواهیم ۱۰ سؤال مشابه سؤال‌های موجود به آن اضافه کنیم. ضریب پایایی آزمون جدید ۴۰ سؤالی ما  $0/67$  خواهد بود.

### تحلیل دانش امور واقعی

میانگین تعداد روزهایی که دانش‌آموزان ۱۳ ساله‌ی کشورهای مختلف در طول یک سال تحصیلی به مدرسه می‌روند. (اطلاعات مربوط به سال‌های ۹۱-۱۹۹۰، نقل شده در لین و گراندلند، ۲۰۰۰، ص ۲۲۴)

میانگین تعداد روزهایی که دانش‌آموزان ۱۳ ساله‌ی کشورهای مختلف در طول یک سال تحصیلی به مدرسه می‌روند. (اطلاعات مربوط به سال‌های ۹۱-۱۹۹۰، نقل شده در لین و گراندلند، ۲۰۰۰، ص ۲۲۴)

راهنمایی: سؤال‌هایی که در دنبال می‌آیند به اطلاعات موجود در شکل بالا اشاره می‌کنند. هر یک از آنها را بخوانید و پاسخ خود را مطابق با دستورالعمل زیر علامت بزنید:

صحیح: اگر براساس اطلاعات موجود در شکل سؤال صحیح است.

غلط: اگر براساس اطلاعات موجود در شکل سؤال غلط است.

نامشخص: اگر اطلاعات موجود در شکل برای تشخیص دادن صحیح و غلط بودن سؤال کافی نیست.

### صحیح غلط نامشخص

- ✓ ۲۰- تعداد روزهای رفتن دانش‌آموزان به مدرسه در ایالات متحده کمتر از کشورهای دیگر است.
- ✓ ۲۱- تعداد روزهای رفتن دانش‌آموزان به مدرسه در سه کشور آسیایی بیشتر از بقیه‌ی کشورهاست.

✓

۲۲- دانش‌آموزان آمریکایی از دانش‌آموزان ژاپنی ساعت‌های کمتری را در مدرسه سپری می‌کنند.

### حوزه‌ی عاطفی

#### دریافت کردن (علاقه)

بله      نه

۲۳- دوست دارید در یک کلوپ ورزشی عضو شوید؟

موافق   مخالف

۲۴- اگر یکی از اعضای خانواده من موسیقی رادیو گوش دهد من ترجیح می‌دهم اتاق را ترک کنم.

#### پاسخ دادن

موافق   مخالف

۲۵- من تکالیف درسی‌ام را با میل و رغبت انجام می‌دهم.

بله      نه

۲۶- برای شما غیرممکن است که یک ساعت مطالعه کنید و حوصله‌تان سر نرود؟

#### ارزش‌گذاری (نگرش)

موافق   مخالف

۲۷- سازمان ملل یک نهاد بی‌خاصیت است.

#### ارزش‌گذاری (باور)

موافق   مخالف

۲۸- حکومت‌ها باید مواظب مردم باشند.

#### سازمان دادن

موافق   مخالف

۲۹- خواندن کتاب برای من به صورت یک سرگرمی تازه درآمده است.

## سنجش یادگیری به کمک آزمون جور کردنی

### هدف‌های یادگیری

از خواننده انتظار می‌رود پس از مطالعه و یادگیری مطالب این فصل بتواند:

- ۱- آزمون جور کردنی را تعریف کند.
  - ۲- ویژگی‌های سؤال‌های جور کردنی را توضیح دهد.
  - ۳- انواع سؤال‌های جور کردنی را شرح دهید و برای هر یک مثالی بزند.
  - ۴- قواعد تهیه سؤال‌های جور کردنی را توضیح دهد.
  - ۵- امتیازها و محدودیت‌های سؤال‌های جور کردنی را توضیح دهد.
  - ۶- موارد استفاده از آزمون‌های جور کردنی را توضیح دهد.
  - ۷- برای هدف‌های یادگیری یک درس که با آزمون‌های جور کردنی قابل سنجش‌اند یک آزمون جور کردنی تهیه نماید.
- در این فصل نوع دیگر این روش‌ها یعنی آزمون جور کردنی را معرفی می‌کنیم. آگاهی یافتن معلم از ویژگی‌ها و چگونگی تهیه سؤال‌های جور کردنی او را در استفاده‌ی بهینه از این سؤال‌ها به عنوان مکمل سؤال‌های صحیح-غلط یاری خواهد کرد.

### تعریف و انواع آزمون‌های جور کردنی

هر آزمون جور کردنی<sup>۱</sup> از تعدادی تمرین<sup>۲</sup> (سؤال) تشکیل می‌شود و هر تمرین نیز شامل سه بخش زیر است:

---

<sup>۱</sup> -matching test

<sup>۲</sup> -exercise



(۱) راهنمای انجام دادن تمرین؛ (۲) فهرست محرک‌ها یا پرسش‌ها؛ (۳) فهرست پاسخ‌ها. آزمون‌شونده، با توجه به توضیحات راهنما، فهرست پاسخ‌ها را با فهرست پرسش‌ها جور می‌کند. نگاه کنید به نمونه‌ی زیر:

**راهنمایی:** در جاهای خالی مقابل اکتشافات و اختراعات ستون سمت راست زیر، حروف اول مربوط به کاشفان و مخترعان ستون سمت چپ را بنویسید، هر پاسخ را بیشتر از یک بار به کار نبرید.

الف- ماری کوری	(پ) ۱- پنی سیلین
ب- دی باکی	(ث) ۲- پرتو
پ- فلمینگ	(الف) ۳- رادیوم
ت- هاروی	(ج) ۴- واکسن فلج
ث- رونتگن	(ب) ۵- قلب مصنوعی
ج- سالک	(ت) ۶- گردش خون
چ- پاستور	

نمونه‌ی بالا متداول‌ترین الگوی سؤال‌های جور کردنی است، اما سؤال‌های جور کردنی به این الگو منحصر نمی‌شوند و الگوهای دیگری را نیز می‌توان مورد استفاده قرار داد. برای مثال، می‌توان نقشه‌ای را که در آن نقاط مختلفی با شماره یا حرف مشخص شده‌اند در اختیار دانش‌آموزان گذاشت و از آنان خواست تا شماره‌ها یا حرف‌ها را به نام شهرها، دریاچه‌ها، رودخانه‌ها، یا مواردی از این قبیل جور کنند. یا می‌توان شکل یک دستگاه برقی را که در آن قسمت‌های مختلف دستگاه شماره‌گذاری شده‌اند به دانش‌آموزان داد و از آنان خواست تا نام قسمت‌های مختلف دستگاه را به شماره‌های داخل شکل ارتباط دهند.

در تدوین سؤال‌های جور کردنی، رابطه‌ی پرسش‌ها و پاسخ‌ها را می‌توان به گونه‌های دیگری تعیین کرد. از جمله اینکه می‌توان از آزمون‌شوندگان خواست تا از هر پاسخ خطی به پرسش مربوط به آن بکشند. هر روشی را که مفید تشخیص می‌دهید، به کار بندید، اما همواره سعی کنید که آن را در راهنمای تمرین یا سؤال به روشنی توضیح دهید.

نوع معروفی از سؤال‌ها یا تمرین‌های جور کردنی، نوع رده‌بندی<sup>۱</sup> یا طبقه‌بندی است که به نوع ماده‌ی فهرست اصلی<sup>۲</sup> نیز شهرت دارد (نیتکو، ۲۰۰۱). در این نوع سؤال‌ها، یک طرح رده‌بندی (مثلاً قسمت‌های جمله، انواع

---

<sup>۱</sup> -Classification

<sup>۲</sup> - master list item

جانوران، دوره‌های تاریخی، انواع گیاهان، انواع واکنش‌های شیمیایی، و غیره) ارائه می‌شود و از دانش‌آموزان خواسته می‌شود تا هر سؤال را طبق این طرح رده‌بندی کنند.

در آزمون‌های نوع جور کردنی رده‌بندی یا ماده فهرست اصلی هر یک ماده (سؤال) یا تمرین دارای سه قسمت زیر است:

(۱) راهنمایی برای پاسخ دادن به سؤال

(۲) فهرست اصلی گزینه‌ها (که همان فهرست یا ستون پرسش‌هاست) و

(۳) فهرست پاسخ‌ها.

چنان‌که گفتیم، دانش‌آموز، برای پاسخ دادن به این‌گونه سؤال‌ها یا تمرین‌ها، باید هر یک از پرسش‌ها را که با شماره مشخص شده است بخواند، و بعد برای یافتن جواب آن در میان فهرست پاسخ‌هایی که با حروف نشان داده شده‌اند پاسخ متناسب را بیابد.

نول و اسکانل (۱۹۷۲) و مهرنر و لهما (۱۹۸۴) می‌گویند این نوع سؤال‌ها بیشتر برای سنجش توانایی انتقال یادگیری، توضیح، و درک مطالب مناسب‌اند. به مثال زیر توجه کنید:

راهنمایی: در ستون اول نام دستوری قسمت‌های مختلف یک جمله داده شده، در جای خالی مقابل هر جمله‌ی ستون دوم حرف مربوط به نام دستوری کلمه‌ای را که با حروف سیاه نوشته شده است بنویسید.

الف- صفت

ب- قید

پ- حرف اضافه

ت- حرف ربط

(پ) ۱- کسی در مدرسه نبود جز ناظم

(ت) ۲- تا درس نخوانی چیزی نمی‌آموزی.

(ت) ۳- دیروز به کتابفروشی رفتم و دو جلد کتاب خریدم.

(الف) ۴- همه‌ی درهای بسته باز خواهند شد.

(الف) ۵- پرویز از احمد دانش‌آموز کوشاتری است.

-

-

-

(ب) ۷- پروین برای جبران سهل‌انگاری گذشته‌ای سخت می‌کوشد.

سؤال‌های جور کردنی رده‌بندی از نوع بالا با سؤال‌های جور کردنی معمولی از لحاظ تعداد ماده‌های فهرست پرسش و پاسخ فرق دارند. از آنجا که در سؤال‌های رده‌بندی از یک پاسخ بیشتر از یک بار می‌توان استفاده کرد، در هر تمرین تعداد پرسش‌ها می‌تواند بیشتر از تعداد پاسخ‌ها باشد، در صورتی که در سؤال‌های جور کردنی معمولی تعداد پاسخ‌ها بیشتر از تعداد پرسش‌هاست.

یکی از امنیازهای سؤال‌های جور کردنی از نوع رده‌بندی با ماده‌ی فهرست اصلی این است که در تصحیح جواب‌ها می‌توان از پاسخنامه‌های جداگانه شبیه آنچه برای آزمون‌های چندگزینه‌ای به کار می‌رود استفاده کرد. به نمونه‌ی زیر توجه کنید:

راهنمایی: در پرسش‌های زیر شما باید تأثیر ورزش را بر فرآیندها و مواد مختلف بدن تعیین کنید. فرض کنید که در ارگانیسم هیچ تغییری بجز آنچه در نتیجه‌ی ورزش رخ می‌دهد اتفاق نمی‌افتد. برای هر پرسش مطابق با توضیحات زیر در پاسخنامه علامت بگذارید.

الف- اگر نتیجه‌ی ورزش قطعاً افزایش چیزی است که در پرسش آمده

ب- اگر نتیجه‌ی ورزش قطعاً کاهش چیزی است که در پرسش آمده

پ- اگر ورزش هیچ‌گونه تأثیر چشم‌گیری بر چیزی که در پرسش آمده ندارد یا تأثیر آن غیرقابل پیش‌بینی است.

الف	ب	ج
✓		
✓		
	✓	
	✓	
۱-ضربان قلب		
۲- فشارخون		
۳- مقدار قند موجود در خون		
۴- مقدار هوای اضافی در شش‌ها		
۵- و غیره		

(ایبل، ۱۹۷۲، ص ۱۰۵)

نکته‌ای را که لازم است در رابطه با سؤال‌های جور کردنی یادآور شویم این است که فهرست پرسش‌ها و فهرست پاسخ‌ها را هم می‌توان در دو ستون موازی در کنار هم قرار داد و هم می‌توان آن‌ها را پشت سر هم نوشت (مانند مثال بالا). آنجا که طول عبارت‌های فهرست پرسش و فهرست پاسخ کوتاه است و دو ستون در کنار هم جا می‌گیرند بهتر است آن‌ها را به طور موازی پهلوی هم قرار دهید، اما در جایی که مطالب فهرست پرسش یا فهرست پاسخ مفصل است و در کنار هم جا نمی‌گیرند آن‌ها را به دنبال یکدیگر بیاورید.

## قواعد تهیه‌ی سؤال‌های جور کردنی

صاحب‌نظران سنجش یادگیری در رابطه با چگونگی تهیه‌ی سؤال‌های جور کردنی قاعده‌هایی را پیشنهاد داده‌اند که ما در این قسمت به توضیح آنها می‌پردازیم. در بحث از سؤال‌های جور کردنی ما اصطلاح تکلیف را مورد استفاده قرار داده‌ایم. منظور ما از تکلیف مجموعه‌ای از پرسش‌ها و پاسخ‌های همگون است که آزمون‌شونده باید آنها را با هم جور کند. این مجموعه بر روی هم بیشتر از یک سؤال محسوب می‌شود. در واقع تعداد سؤال‌های هر تکلیف برابر تعداد پرسش‌های آن است.

### ۱- اطلاعات کامل درباره‌ی چگونگی جور کردن پرسش‌ها و پاسخ‌ها را در راهنمای سؤال بنویسید.

هرچند که در آزمون‌های جور کردنی نحوه‌ی عمل جور کردن پرسش‌ها با پاسخ‌ها نسبتاً آشکار است، اما باز هم لازم است در راهنمای سؤال‌ها چگونگی این کار را توضیح دهید. این توضیحات از سردرگمی دانش‌آموزان جلوگیری می‌کند. با این حال، از نوشتن توضیحات خیلی مفصل در راهنمای سؤال‌ها که موجب به هدر دادن مقداری از وقت آزمون‌شوندگان می‌شود پرهیز کنید.

به طور کلی، در راهنمای سؤال‌ها سه دسته اطلاع زیر را قرار دهید: (۱) اساس رابطه‌ی بین پرسش‌ها و پاسخ‌ها با مبنای جور کردن آنها؛ (۲) روش مشخص کردن پاسخ‌ها یا جور کردن پرسش‌ها و پاسخ‌ها؛ و (۳) آیا از یک پاسخ تنها یک بار استفاده بشود یا بیشتر از یک بار. در راهنمای سؤال‌هایی که تاکنون در این فصل مثال زدیم موارد سه‌گانه‌ی بالا دیده می‌شوند. در زیر به نمونه‌ی دیگری از راهنمای سؤال که هر سه مورد بالا را شامل است توجه کنید.

راهنمایی: در زیر در ستون با دو فهرست مطلب وجود دارند، ستون اول دربرگیرنده‌ی عناوین داستان‌های کوتاه و ستون دوم شامل نام‌های نویسندگان است. برای هر یک از داستان‌های ستون سمت راست نام یکی از نویسندگان ستون سمت چپ را برگزینید، حرف مقابل نام نویسنده را در جای خالی مقابل عنوان داستان بنویسید، تعداد نام‌های نویسندگان بیشتر از تعداد عنوان‌های داستان‌هاست. بنابراین، تعدادی نام نویسنده‌ی اضافی خواهید آورد. همچنین ممکن است نام یک نویسنده برای بیشتر از یک داستان انتخاب شود. (نول و اسکائل، ۱۹۷۲، ص ۸)

هنگام اجرای آزمون‌های جور کردنی، علاوه بر توضیحات کتبی، اگر لازم بود، توضیحات شفاهی هم بدهید. این عمل به ویژه در رابطه با دانش‌آموزان سن‌های پایین‌تر مفید است. برای این منظور، می‌توانید مثالی بر روی تابلو بنویسید. همچنین می‌توانید از دانش‌آموزان بخواهید تا، به جای نوشتن حرف‌های مربوط به پاسخ‌ها در مقابل پرسش‌ها، از هر پاسخ خطی به پرسش مربوط به آن وصل کنند.

۲- موضوع‌های مهم و در ارتباط با هدف‌های آموزشی را در سؤال‌ها قرار دهید. ما در ارتباط با تهیه‌ی سؤال‌های صحیح- غلط بر این نکته تأکید کردیم. در اینجا نیز آن را یادآور می‌شویم. هر تمرین جور کردنی باید یک موضوع مهم مربوط به یک هدف آموزشی را بسنجد. از قرار دادن مطالب جزئی، بی‌اهمیت، و نامربوط به محتوای درس و هدف‌های یادگیری بپرهیزید. در تدارک همه‌ی روش‌های سنجش این مطلب را به یاد داشته باشید.

۳- پرسش‌ها و پاسخ‌های هر تمرین را متجانس انتخاب کنید. منظور از تجانس این است که مجموعه‌ی پرسش‌ها و پاسخ‌های یک تمرین به یک مفهوم، طبقه یا زمینه‌ی واحد مربوط باشند. در زیر به نمونه‌ای از سؤال‌های نامتجانس توجه کنید.

راهنمایی: در جای خالی مقابل هر شماره در ستون سمت راست، حرف مربوط به مطلب ستون سمت چپ را که با آن جور درمی‌آید بنویسید. هر پاسخ را بیشتر از یک بار به کار نبرید.

(پ) ۱- پنی سیلین را کشف کرد. الف- آرمسترانگ

(ب) ۲- آمریکا را کشف کرد. ب- کریستف کلمب

(الف) ۳- اولین فضانوردی که بر ماه قدم نهاد. پ- فلمینگ

(ت) ۴- اولین کاشف که دماغه‌ی امیرنیک را دور زد. ت- گلن

ث- ماژلان

ج- نیوتن

تجانس یا همگونی یک امر نسبی است و به همین دلیل طراحان آزمون ممکن است در نوشتن سؤال‌های آزمون جور کردنی با مشکل مواجه شوند، زیرا آنچه برای یک گروه از آزمون‌شوندگان متجانس است ممکن است برای گروهی دیگر نامتجانس جلوه کند. لین و گرانلاند (۲۰۰۰، ص ۱۸۸) در این باره سؤال زیر را ذکر کرده‌اند.

راهنمایی: در جای خالی مقابل هر شماره در ستون سمت راست حرف مربوط به مطلب ستون سمت چپ را که با آن جور درمی‌آید بنویسید. هر یک از نام‌های ستون سمت چپ را می‌توانید یک بار یا بیشتر از یک بار به کار برید یا اصلاً آن را مورد استفاده قرار ندهید.

(الف) ۱- مخترع تلفن الف- الکساندر گراهام بل

(ب) ۲- کاشف آمریکا ب- کریستوفر کولومبوس

(پ) ۳- اولین فضانورد آمریکایی که زمین را دور زد. پ- جان گلن

(ت) ۴- اولین رئیس‌جمهور آمریکا ت- ابراهام لینکن

ث- فردیناند مگلان

## ج- جورج واشنگتن

### ج- الی ویتنی

تمرین بالا از آنجا که تنها انسان‌ها را شامل می‌شود متجانس است. این تمرین ممکن است برای دانش‌آموزان ابتدایی متجانس به نظر برسد، اما برای دانش‌آموزان سال‌های بالاتر که قدرت تشخیص بیشتری دارند خیلی نامتجانس به نظر خواهد رسید زیرا شامل مکتشفان، مخترعان، و رؤسای جمهور است. بنابراین، برای استفاده در کلاس‌های سطح بالاتر باید، به جای این تمرین، تمرین دیگری تهیه کرد که تنها شامل مخترعان، مکتشفان، یا رؤسای جمهور باشد. یعنی تمرین را می‌توان به سه تمرین متفاوت تبدیل کرد. بنابراین، به همان نسبتی که بخواهیم قدرت تمیز یا تشخیص سؤال‌های جور کردنی را بالا ببریم، باید بر تجانس یا همگونی عناصر پرسش و پاسخ آنها بیفزاییم.

**۴- طول فهرست پرسش‌ها و پاسخ‌ها را کوتاه بگیرید.** فهرست کوتاه مطالب هم برای معلم بهتر است هم برای دانش‌آموزان، معلم در تهیه‌ی مطالب کمتر بیشتر می‌تواند تجانس آنها را رعایت کند. از لحاظ مفید بودن برای دانش‌آموزان، خواندن فهرست‌های کوتاه‌تر مطالب دقت و تمرکز حواس را بیشتر امکان‌پذیر می‌سازد.

متخصصان سنجش و ارزشیابی پیشرفت تحصیلی رقم‌های ۴ تا ۱۰ ماده را برای فهرست پرسش‌ها و پاسخ‌ها پیشنهاد کرده‌اند. پوفام (۲۰۰۲) و نول و اسکانل (۱۹۷۲) بالاترین رقم را برای تعداد پرسش‌ها و پاسخ‌ها ۱۰ تا ۱۲ ذکر کرده‌اند. پوفام در رابطه با طولانی بودن فهرست پرسش‌ها و فهرست پاسخ‌ها می‌گوید "مشکل فهرست‌های طولانی این است که دانش‌آموزان آن قدر وقت صرف جدا کردن پاسخ مناسب برای یک پرسش می‌کنند که ممکن است یادشان برود که به دنبال چه چیزی می‌گردند" (ص ۱۴۲)

**۵- توضیحات مفصل را در فهرست پرسش‌ها قرار دهید و در فهرست پاسخ‌ها مطالب کمتری بنویسید.** قرار دادن مطالب مختصر در ستون پاسخ‌ها به کارآمدی سؤال کمک می‌کند. این امر باعث می‌شود که دانش‌آموزان ابتدا مطالب طولانی‌تر پرسش‌ها را به دقت مطالعه کنند، و آن‌گاه به سرعت پاسخ‌ها را بخوانند و تصمیم خود را در رابطه با انتخاب آنها بگیرند.

دانش‌آموز، هنگام جواب دادن به یک تمرین جور کردنی، ابتدا یک پرسش را می‌خواند، بعد از میان فهرست پاسخ‌ها به دنبال پاسخ مناسب برای آن پرسش می‌گردد. برای هر پرسش این کار را تکرار می‌کند. یعنی به تعداد پرسش‌ها تمامی فهرست پاسخ‌ها را می‌خواند. بنابراین، بهتر است مطالب مفصل در فهرست پرسش‌ها- که دانش‌آموز هر یک از آنها را تنها یک بار می‌خواند- بیاید، و مطالب کوتاه (کلمه‌ها، نمادها و غیره) در فهرست پاسخ‌ها- که بارها باید خوانده شود - قرار گیرد. به نمونه‌ی زیر توجه کنید.

راهنمایی: در جای خالی مقابل هر پرسش در ستون سمت راست حرف مربوط به ضریب پایایی یکی از موارد ستون چپ را بنویسید. هر یک از حرف‌های مربوط به ستون سمت چپ را می‌توانید یک بار یا چند بار به کار برید، یا هرگز از آن استفاده نکنید.

(ث) ۱-نمرات استانفورد- بینه با نمرات آزمون وکسلر همبستگی الف- ثبات دارند.

(ت) ۲- روشی که غالباً برای تخمین پایایی آزمون‌های معلم ب- تعادل ساخته به کار می‌رود.

(ت) ۳- معمولاً کمترین برآورد پایایی را به دست می‌دهد. پ- ثبات و تعادل

(ث) ۴- نتایج تحت تأثیر آزمون‌های سرعت قرار دارند. ت- دو نیمه کردن

(الف) ۵- آزمون استعداد تحصیلی یک بار به دانش‌آموزان سال اول ث- هیچ یک از موارد بالا دبیرستان و بار دوم به همان دانش‌آموزان در سال آخر دبیرستان داده می‌شود.

(مهرنر و لهما، ۱۹۷۳، ص ۲۶۰)

۶- از نوشتن تمرین‌ها یا سؤال‌هایی که پرسش‌ها و پاسخ‌های آنها کاملاً جور می‌شوند پرهیز کنید. منظور از این پیشنهاد آن است که تمرین‌هایی ننویسید که در آنها هر پاسخ تنها یک بار با یک پرسش جور می‌شود و تعداد پاسخ‌ها و پرسش‌های آن برابر است. در این‌گونه تمرین‌ها، اگر دانش‌آموز پاسخ همه‌ی پرسش‌ها بجز یکی از آنها را بداند، آخری را نیز خود به خود جواب خواهد داد. یا اگر در یکی از پاسخ‌ها اشتباهی صورت گیرد در آخرین پاسخ هم همین اشتباه تکرار خواهد شد. اگر تعداد پاسخ‌ها از تعداد پرسش‌ها بیشتر باشد این مشکل برطرف خواهد شد. همچنین سعی کنید پاسخ‌هایی بنویسید که بیش از یک بار مورد استفاده قرار گیرند.

بنابراین، در نوشتن سؤال‌های جور کردنی، لازم است تعداد پاسخ‌ها از تعداد پرسش‌ها بیشتر باشد یا اینکه بتوان هر پاسخ را برای یک پرسش مورد استفاده قرار داد (مانند سوال‌های رده‌بندی). همچنین می‌توان سؤال را به گونه‌ای طراحی کرد که در آن بعضی پاسخ‌ها یک بار و بعضی دیگر بیشتر از یک بار مورد استفاده قرار گیرند، و بعضی از آنها اصلاً قابل استفاده نباشند. این مطلب باید در راهنمای سؤال توضیح داده شود.

۷- هر یک از پاسخ‌ها را به گونه‌ای انتخاب کنید که برای تمام پرسش‌ها درست جلوه کند. سؤال را به گونه‌ای طرح نکنید که پاسخ‌های غلط کاملاً مشخص باشند، طوری که دانش‌آموز به راحتی بتواند آنها را تشخیص داده کنار بگذارد. چنان‌که قبلاً توضیح دادیم، هر پاسخی چه درست و چه غلط باید برای دست‌کم یک پرسش

درست جلوه کند. اگر قاعده‌ای را که درباره‌ی متجانس بودن پرسش‌ها و پاسخ‌ها قبلاً توضیح دادیم رعایت کنید این منظور برآورده می‌شود، زیرا معمولاً پاسخ‌های متجانس برای فهرستی از پرسش‌های متجانس درست به نظر می‌رسند.

۸- پرسش‌ها را با شماره و پاسخ‌ها را با حرف مشخص کنید. در تمرین‌های مربوط به آزمون‌های جور کردنی هر پرسش یک نمره‌ی جداگانه دریافت می‌کند. بنابراین، هر پرسش باید دارای یک شماره باشد و به عنوان یک سؤال مستقل منظور شود. اگر در یک آزمون از انواع مختلف سؤال‌ها استفاده می‌کنید، همان‌طور که برای هر یک از آنها یک شماره اختصاص می‌دهید، برای هر یک از پرسش‌های سؤال‌های جور کردنی نیز یک شماره اختصاص دهید. مثلاً، اگر ۱۵ سؤال اول آزمون شما صحیح، غلط است و به دنبال آنها یک تمرین جور کردنی دارای ۵ پرسش قرار می‌دهید، پرسش‌های تمرین جور کردنی را از ۱۶ تا ۲۰ شماره‌گذاری کنید. در این حالت ارزش یک تمرین جور کردنی ۵ پرسشی شما ۵ نمره خواهد بود.

۹- فهرست پرسش‌ها و فهرست پاسخ‌ها را به طور منطقی مرتب کنید. اگر فهرست پرسش‌ها و پاسخ‌های یک تمرین جور کردنی به گونه‌ای منطقی مرتب شده باشند، دانش‌آموزان مطلع از موضوع بدون صرف وقت پاسخ درست هر پرسش را پیدا خواهند کرد. در غیر این صورت، ممکن است وقت زیادی را صرف یافتن آن بکنند. علاوه‌براین، مرتب کردن فهرست پرسش‌ها و فهرست پاسخ‌ها موجب می‌شود که دانش‌آموزان نتوانند اشاره‌های احتمالی را که منجر به یافتن پاسخ‌های درست می‌شود کشف کنند.

برای مرتب کردن مطالب، نام‌ها را به ترتیب حروف الفبا، تاریخ‌ها را به ترتیب زمان، و اعداد را از بزرگ به کوچک یا به عکس مرتب کنید. به نمونه‌ی زیر توجه نمایید:

راهنمایی: در جاهای خالی مقابل آثار ستون سمت راست، حروف مربوط به نویسندگان آن آثار را که در ستون سمت چپ نوشته شده‌اند قرار دهید.

الف-بوعلی‌سینا	(پ) ۱- بوستان
ب- حافظ	(ت) ۲- چهار مقاله
پ- سعدی	(ج) ۳- شاهنامه
ت- عروضی	(الف) ۴- شفا
ث- غزالی	(ث) ۵- کیمیای سعادت
ج- فردوسی	



۱۰- در برگه‌های امتحانی، همه‌ی پرسش‌ها و پاسخ‌های یک سؤال را در یک صفحه قرار دهید. رعایت این سفارش اجرای آزمون‌های جور کردنی را آسان می‌کند. اگر نصب مطالب یک تمرین در یک صفحه و نصف دیگر آن در صفحه‌ای دیگر باشد، مراجعه‌ی مکرر از این صفحه به آن صفحه موجب حواس‌پرتی و صرف وقت زیاد از دانش‌آموزان می‌شود. در ضمن، ورق زدن برگه‌های امتحانی توسط دانش‌آموزان مختلف باعث تولید سر و صدای زیاد و ایجاد مزاحمت می‌شود. بنابراین، سؤال‌ها را طوری تنظیم کنید که تمام مطالب دو ستون پرسش و پاسخ یک تمرین در یک صفحه قرار بگیرند تا آزمون‌شوندگان مجبور نباشند برای یافتن پاسخ‌های موردنظر مرتباً برگه‌ها را ورق بزنند. همچنین قرار گرفتن بخشی از پاسخ‌ها در صفحه‌ی دیگر ممکن است موجب شود که دانش‌آموزان آنها را نادیده بگیرند.

۱۱- در نوشتن صورت سؤال‌ها ابتکاری عمل کنید. لازم نیست در نوشتن سؤال‌های جور کردنی (یا هر نوع سؤال دیگری) صرفاً از قالب‌های متداول پیروی نمایید. با حفظ اصل اساسی حاکم بر نوع سؤال (در آزمون‌های جور کردنی، عمدتاً سنجش رابطه) ابتکار به خرج دهید و سؤال‌های تازه و متنوع بسازید. در زیر به نمونه‌ای از سؤال‌های جور کردنی ابتکاری توجه کنید.

راهنمایی: شما درباره‌ی زمین و حرکات آن، آن‌گونه که واقعاً وجود دارد، اطلاعاتی کسب کرده‌اید، در این سؤال باید آثار شرایطی کاملاً فرضی را شناسایی کنید. در مقابل شماره هر سؤال در پاسخنامه، برحسب موارد زیر علامت بزنید.

الف- در صورتی که زمین نسبت به مدار خود مایل نبود سؤال درست بود.

ب- در صورتی که مدار به جای بیضی دایره بود سؤال درست بود.

پ- در صورتی که زمین به جای شرق به طرف غرب می‌چرخید سؤال درست بود.

ت- در صورتی که زمین با همین جرم نصف قطر فعلی‌اش را داشت سؤال درست بود.

ث- در صورتی که زمین ماه نداشت سؤال درست بود.

۱- طول همه‌ی روزهای شمسی مساوی می‌بودند.

۲- اشیاء روی زمین چهار برابر وزن فعلی خود را داشتند.

۳- خط استوای کره‌ی آسمانی و مدار خورشید یکی بود.

۴- خورشید از شرق غروب می‌کرد.

۵- ستاره‌ی شمالی دیگری باید انتخاب می‌شد.

۶- نیروی جاذبه چهار برابر بیشتر از حالا می‌بود.

۷- سرعت حرکت زمین بر روی مدارش در ضمن سال تغییر نمی‌کند.

۸- ما خیلی کمتر از حالا درباره‌ی ماهیت خورشید اطلاع می‌داشتیم.

۹- طول روز و شب در تمام طول سال در همه‌ی عرض‌های جغرافیایی برابر بود.

(از بلوم و همکاران، ۱۹۵۶، ترجمه‌ی سیف و علی‌آبادی، ۱۳۶۸، ص ۱۴۸)

### امتیازها و محدودیت‌های آزمون‌های جور کردنی

صاحب‌نظران سنجش و اندازه‌گیری، در رابطه با آزمون‌های جور کردنی، نظرهای موافق و مخالف ابراز داشته‌اند. به طور کلی، امتیازهای استفاده از سؤال‌های جور کردنی را می‌توان به صورت زیر خلاصه کرد:

۱- از آنجا که پاسخ دادن به این سؤال‌ها زمان زیادی لازم ندارد، تعداد زیادی از آنها را می‌توان در یک زمان معین مورد استفاده قرار داد. بنابراین، استفاده از این نوع سؤال‌ها، مانند سؤال‌های عینی دیگر، نمونه‌ی نسبتاً بزرگی از محتوای درس و هدف‌های آموزشی فراهم می‌آورد.

۲- تصحیح سؤال‌های جور کردنی به سرعت و سهولت امکان‌پذیر است و نیازی به مصحح متخصص ندارد. علاوه‌براین، تصحیح این سؤال‌ها به طور کاملاً عینی امکان‌پذیر است و در آن امکان دخالت نظر شخصی مصحح وجود ندارد.

۳- امکان حدس زدن جواب درست آزمون‌های جور کردنی کمتر از سایر آزمون‌های گزینش پاسخ، به ویژه آزمون‌های صحیح-غلط، است.

۴- از سؤال‌های جور کردنی برای سنجش هدف‌های یادگیری سطح بالاتری از سؤال‌های صحیح-غلط می‌توان استفاده کرد.

در مقابل فایده‌های بالا، برای آزمون‌های جور کردنی محدودیت‌هایی ذکر کرده‌اند، مهم‌ترین آنها به قرار زیر بودند:

۱- از آنجا که در یادگیری تداعی بین امور حفظ طوطی‌وار دخالت زیادی دارد و معلمان غالباً در تهیه‌ی این نوع سؤال‌ها بر روابط حفظی بین امور تأکید می‌کنند، این آزمون‌ها اگر به دقت تهیه نشوند ممکن است دانش‌آموزان را به یادگیری مطالب سطحی تشویق کنند.

۲- بعضی وقت‌ها پیدا کردن مجموعه‌ای پرسش که به اندازه‌ی کافی شبیه به هم باشند تا بتوان برای آنها مجموعه‌ی پاسخ‌های همگونی پیدا کرد کار دشواری است.

۳- علاوه‌بر معایب بالا، بعضی از سؤال‌های جور کردنی معمولی که به وسیله‌ی معلمان تهیه می‌شوند دارای عیب‌های زیر نیز هستند:

الف- مبهم بودن راهنمای سؤال

ب- طولانی بودن فهرست پرسش‌ها و پاسخ‌ها

پ- مبهم بودن بیان پرسش‌ها

### موارد استفاده‌ی آزمون‌های جور کردنی

آزمون‌های جور کردنی، در قیاس با آزمون‌های صحیح- غلط، برای سنجش دامنه‌ی وسیع‌تری از هدف‌های آموزشی قابل استفاده‌اند. عمده‌ترین مورد استفاده‌ی آزمون‌های جور کردنی سنجش هدف‌هایی است که روابط بین امور را شامل می‌شوند - به ویژه سنجش توانایی آزمون‌شوندگان در تشخیص روابط میان موضوع‌های نسبتاً مشابه یا مربوط به هم. در این سؤال‌ها، می‌توان از دانش‌آموز خواست تا تاریخ‌ها را به رویدادها، مؤلفان را به کتاب‌ها، ابزارها را به موارد استفاده‌ی آنها، و مانند اینها ربط داد. لین و گراندلند (۲۰۰۰) نمونه‌ی روابط مهم از لحاظ آموزش معلمان را در فهرست زیر آورده‌اند:

اشخاص.....دستاوردها

تاریخ‌ها.....رویدادهای تاریخی

اصطلاحات.....تعاریف

قواعد .....مثال‌ها

نمادها.....مفاهیم

نویسندگان .....عنوان‌های کتابها

کلمات خارجی .....معادل‌های زبان مادری

ماشین‌ها.....موارد استفاده

گیاهان یا جانوران .....طبقه‌بندی‌ها

اصل‌ها .....نمونه‌ها

اشیاء .....نام‌های اشیاء

اجزاء .....کارکردها

علاوه‌بر موارد بالا، از آزمون‌های جور کردنی می‌توان برای سنجش توانایی تشخیص روابط در شکل‌ها و نمودارها نیز استفاده کرد. برای نمونه می‌توان از دانش‌آموزان خواست تا شکل‌ها را با کلمات جور کنند یا موقعیت‌های مختلف را در نقشه‌ها، نمودارها، طرح‌ها، و مانند اینها تشخیص دهند. "صرف‌نظر از صورت ظاهر ارائه‌ی سؤال، وظیفه‌ی

دانش‌آموز این است که دو چیز را که براساس نوعی منطق با هم جور درمی‌آیند به هم ربط دهد" (لین و گراندلاند، ۲۰۰۰، ص ۱۸۷).

علاوه‌بر سنجش توانایی تشخیص روابط میان امور، از آزمون‌های جور کردنی نوع رده‌بندی یا ماده‌ی فهرست اصلی می‌توان برای سنجش درک مفاهیم نیز استفاده کرد. نیتکو (۲۰۰۱) در این باره گفته است "برای اینکه از این نوع سؤال‌ها (جور کردنی نوع فهرست اصلی) در سنجش درک مفهوم استفاده نمایید، مثال‌هایی را که به دانش‌آموزان می‌دهید تا دسته‌بندی کنند نه باید مثال‌هایی که در کلاس مورد استفاده قرار داده‌اید باشند نه مثال‌هایی که در کتاب درسی آمده‌اند" (ص ۱۷۳). ثراندیک، کانینگهام<sup>۱</sup>، ثرندایک، و هیگن (۱۹۹۱)، درباره‌ی سؤال‌های جور کردنی نوع ماده‌ی فهرست اصلی، گفته‌اند "این‌گونه سؤال‌ها را می‌توان با اطمینان کامل برای سنجش کار بستن، فهمیدن، و تفسیر کردن مورد استفاده قرار داد. این سؤال‌ها راه مناسبی برای کشف و تعیین دامنه‌ای از یادگیری مفاهیم مختلف فراهم می‌آورند" (ص ۲۳۸).

همچنین از آزمون‌های جور کردنی (بازهم از نوع ماده یا سؤال فهرست اصلی و نوع رده‌بندی) می‌تواند برای سنجش توانایی تحلیل استفاده کرد. "سؤال نوع فهرست اصلی راه مناسبی برای سنجش توانایی‌های زیر است:

(۱) تحلیل متن‌های مختصر، جدول‌ها، یا نمودارها و (۲) تشخیص تفسیر یا استنتاج مناسب از مطالب تفسیری" (نیتکو، ۲۰۰۱، ص ۱۷۴). در زیر به نمونه‌هایی از سؤال‌های جور کردنی برای سنجش سطوح مختلف توجه کنید:

### یادآوری دانش امور واقعی

راهنمایی: در جای خالی مقابل نام هر یک از نویسندگان ستون سمت راست، حرف مربوط به عنوان یکی از کتاب‌های ستون سمت چپ را قرار دهید. هر یک از حروف ستون چپ را می‌توانید یک بار یا بیش از یک بار مورد استفاده قرار دهید یا اصلاً از آن استفاده نکنید. مورد اول به عنوان نمونه جواب داده شده است.

#### ستون عنوان کتاب‌ها

الف- چرخها

ب- خیابان اصلی

پ- بادهای جنگ

ت- داستان عاشقانه

ث- هفت دقیقه

#### ستون نویسندگان

(الف) x- آرتور هیلی

(ت) ۱- اریک سیگال

(ث) ۲- آیرونینگ والاس

(ج) ۳- جوزف وامباگ

(پ) ۴- هرمان ووک

<sup>۱</sup> -Cunningham

- (ج) ۵- لئون اُرس  
 (خ) ۶- پرل باک  
 (ج) ۷- الکس هیلی  
 ج- کیوبی هفت  
 چ- پله‌های چرخان  
 ح- فرماندهان جدید  
 خ- خاک خوب  
 د- ریشه‌ها

(مهرنر و لهمان، ۱۹۸۴، ص ۱۴۰)

### یادآوری دانش مفهومی

راهنمایی: ستون سمت چپ دربرگیرنده‌ی اصطلاحات مختلف روانشناسی یادگیری و ستوت سمت راست شامل موقعیت‌های مختلف مُعرف آن اصطلاحات است. با گذاشتن حروف اول مقابل مطالب ستون سمت چپ در جاهای خالی مقابل شماره‌های مربوط به مطالب ستون سمت راست، آنها را با هم جور کنید. از هر یک از موارد الف و ب و پ می‌توانید یک بار یا بیشتر از یک بار استفاده کنید یا اصلاً از آن استفاده نکنید.

- |  |                 |
|--|-----------------|
| الف) ۸- ارائه تقویت‌کننده‌ی مثبت پس از رفتار | الف- تقویت مثبت |
| ب) ۹- حذف تقویت‌کننده‌ی منفی پس از رفتار     | ب- تقویت منفی   |
| پ) ۱۰- ارائه تقویت‌کننده‌ی منفی پس از رفتار  | پ- تنبیه        |
| ث) ۱۱- حذف تقویت‌کننده‌ی مثبت پس از رفتار    | ت- محروم کردن   |
|  | ث- جریمه کردن   |

### فهمیدن دانش مفهومی

راهنمایی: مطالب موجود در ستون سمت راست معرف انواع روایی‌های داده شده در ستون سمت چپ است. در جای خالی مقابل هر یک از پرسش‌های ستون سمت راست حرف مربوط به یکی از پاسخ‌های ستون سمت چپ را بنویسید. از هر یک از پاسخ‌ها می‌توانید یک یا چند بار استفاده کنید یا اصلاً آن را مورد استفاده قرار ندهید.

- الف) ۱۲- برای ساختن یک آزمون دیکته، معلم از کلماتی که الف- روایی  
 محتوایی که دانش‌آموزان در نوشتن انشاء غالباً آنها را غلط ب- روایی سازه  
 می‌نویسند استفاده می‌کند. پ- روایی ملاکی  
 ت- هیچ یک از موارد بالا

(ب) ۱۳- معلمی می‌خواهد مشخص کند که اشکالات لغتی دانش‌آموزان در دیکته همان اشکالات لغتی آنان در نوشتن انشاء است.






(پ) ۱۴- در دفترچه‌ی راهنمای یک آزمون هوش بین نمرات حاصل از آن آزمون و آزمون هوش استانفورد- بینه ضریب همبستگی  $+0.76$  گزارش شده است.

(پ) ۱۵- میزان تولید کارگران با درجه‌بندی سرپرست کارگاه از موفقیت کارگران مقایسه شده است.

(الف) معلم به تهیه‌ی جدول مشخصات اقدام کرده است.

(مهرنر و لهما، ۱۹۷۳، ص ۲۶۳)

راهنمایی: در ستون سمت راست پنج شکل هندسی و در ستون سمت چپ نام‌های شکل‌های هندسی مختلف داده شده‌اند. با قرار دادن حرف اول نام شکل‌های فهرست سمت چپ در مقابل شکل‌های فهرست سمت راست آنها را با هم جور کنید. از هر پاسخ تنها یک بار می‌توانید استفاده کنید یا هرگز استفاده نکنید.

الف- مربع		(پ) ۱۷-
ب- مستطیل		(الف) ۱۸-
پ- ذوزنقه		(ب) ۱۹-
ت- لوزی		(ج) ۲۰-
ث- استوانه		(ت) ۲۱-
ج- متوازی‌الاضلاع		
چ- مثلث		
ح- هرم		

### کار بستن دانش مفهومی

راهنمایی: در مقابل شماره هر جفت بیان زیر یکی از حرف‌های الف، ب، یا پ را مطابق با توضیحات داده شده قرار دهید:

الف- اگر افزایش در مورد اول همراه با افزایش در مورد دوم است.

ب- اگر افزایش در مورد اول همراه با کاهش در مورد دوم است.

پ- اگر افزایش در مورد اول اثر چندانی بر مورد دوم ندارد.

(پ) ۲۲- تعداد موش‌های صحرایی در یک منطقه از قطب شمال

تعداد گوزن‌ها در همان منطقه

(الف) ۲۳- تعداد گل‌سنگ‌ها در یک منطقه از قطب شمال

تعداد گوزن‌ها در همان منطقه

(الف) ۲۴- درجه‌ی حرارت محیط زندگی یک جانور

درجه‌ی حرارت بدن همان جانور

(ب) ۲۵- مقدار کربنات‌های محلول در آب یک رودخانه

تعداد صدف‌ها در همان رودخانه

(ب) ۲۶- فشردگی خاک در یک منطقه

مقدار جذب آب به وسیله‌ی خاک پس از یک بار باران سنگین

(الف) ۲۷- فراوانی آتش‌سوزی در یک جنگل کاج

تعداد درخت‌های سپیدار همان جنگل

(الف) ارتفاع محیط زندگی یک جانور

میزان تقسیم گلبول‌های سرخ خون آن جانور

(بلوم و همکاران، ۱۹۵۶، ترجمه سیف و علی‌آبادی، ۱۳۶۸، ص ۱۴۴)

بلوم و همکاران (۱۳۶۸) در معرفی سؤال بالا گفته‌اند "دانش‌آموز باید اصول بوم‌شناسی را به یاد آورد و اصول مناسب را در موقعیت‌هایی که در سؤال آمده به کار بندد... موقعیت‌ها واقعی هستند، اما احتمالاً دانش‌آموز با آنها آشنایی ندارد" (ص ۱۴۴).

### سنجش یادگیری به کمک آزمون چندگزینه‌ای

آزمون چندگزینه‌ای نوع سومی از روش‌های سنجش‌گزینه‌ای است. آزمون‌های چندگزینه‌ای از آزمون‌های صحیح-غلط و جور کردنی معروف‌تر و پراستفاده‌ترند. در دنباله‌ی مطالب به بحث مفصل این آزمون‌ها توجه کنید.

### تعریف و انواع آزمون‌های چندگزینه‌ای

هر آزمون چندگزینه‌ای<sup>۱</sup> از تعدادی سؤال درست می‌شود و هر سؤال چندگزینه‌ای از سه قسمت اصلی زیر تشکیل می‌یابد: (۱) تنه‌ی سؤال،<sup>۲</sup> (۲) گزینه‌ی کلید<sup>۳</sup>، و (۳) گزینه‌های انحرافی<sup>۴</sup>. تنه‌ی سؤال متن یا صورت سؤال را تشکیل می‌دهد و دربرگیرنده‌ی مسئله یا موضوعی است که سؤال باید آن را بسنجد. گزینه‌ی کلید همان پاسخ سؤال است که آزمون‌شونده‌ی آگاه از موضوع سؤال آن را برمی‌گزیند. گزینه‌های انحرافی پیشنهادهای غلطی هستند که نقش آنها منحرف کردن آزمون‌شوندگان ناآگاه از موضوع سؤال است. شکل ۱-۹ قسمت‌های اصلی یک سؤال چندگزینه‌ای را نشان می‌دهد.

اصطلاحات بالا متداول‌ترین اصطلاحات مورد استفاده در سؤال‌های چندگزینه‌ای هستند. با این حال، در متون مختلف از اصطلاحات دیگری نیز استفاده می‌شود که آشنایی با آنها مفید به نظر می‌رسد. برای گزینه‌ی کلید "پاسخ کلید"، "پاسخ درست" یا صرفاً "کلید"؛ برای پاسخ‌های انحرافی "پاسخ‌های پیشنهادی"، "انتخاب‌ها"؛ و برای گزینه‌های انحرافی "گزینه‌های ردگم کنی" استفاده می‌شوند.

---

<sup>۱</sup> -multiple – choice

<sup>۲</sup> -Stem

<sup>۳</sup> -Keyed option

<sup>۴</sup> - distractors



تعداد گزینه‌های سؤال‌های چندگزینه‌ای (گزینه‌ی درست و گزینه‌های انحرافی) از ۲ تا ۵ تغییر می‌کند. از لحاظ نظری، هر چه تعداد گزینه‌ها بیشتر باشد امکان حدس زدن کمتر است. اما به سبب اینکه پیدا کردن بیش از سه گزینه‌ی انحرافی کار دشواری است. آزمون‌های چهار گزینه‌ای که در آنها امکان حدس زدن جواب درست یک در چهار است شهرت بیشتری کسب کرده‌اند.

هر چند که همه‌ی سؤال‌های آزمون‌های چندگزینه‌ای از سه قسمت اصلی نشان داده در شکل ۱-۹ درست می‌شوند، اما انواع مختلفی دارند و معلم علاقمند به تهیه‌ی سؤال‌های ابتکاری می‌تواند از آنها بهره بگیرد. در زیر گونه‌های شناخته شده‌ی آزمون‌های چندگزینه‌ای را توضیح می‌دهیم.

### نوع پرسشی

در سؤال‌های نوع پرسشی آزمون‌های چندگزینه‌ای، تنه‌ی سؤال را یک جمله‌ی پرسشی تشکیل می‌دهد و گزینه‌ها (پیشنهادی و کلید) پاسخ‌های آن پرسش هستند.

تنه‌ی سؤال بیت زیر چه معنی می‌دهد؟ زمین شد شش و آسمان گشت هشت	گزینه‌ی کلید بر اثر گرد و خاک حاصل از سم چهارپایان پایان یک لایه از زمین به آسمان چسبید.
گزینه‌های انحرافی ب - دلیل آشفته‌گی ناشی از هرج و مرج میدان نبرد همه به اشتباه محاسبه گرفتار شدند . پ - نیروی خارج از تصویری که رزمندگان به نمایش گذاشتند حقارت زمین در برابر عظمت آسمان‌ها را نمایان‌تر کرد ت - برخورد سهمگین اسبان و اسب‌سواران با یکدیگر چشمان بینندگان را تیره و تار کرد	

شکل ۱-۹ قسمت‌های مختلف یک سؤال چندگزینه‌ای

### مثال

ژان‌والژان اولین بار که محکوم به اعمال شاقه شد چه چیزی دزدیده بود؟

الف- شمعدان‌های اسقف

(×) ب- یک قرص نان

پ- چند قطعه چوب

ت- گاو یک بیوه‌زن

ث- پوشش سکوی خطابه‌ی کلیسا

(بلوم و همکاران، ۱۹۵۶، ترجمه‌ی سیف و علی‌آبادی، ۱۳۶۸، ص ۸۸)

نوع ناتمام

در سؤال‌های نوع ناتمام آزمون‌های چندگزینه‌ای، تنه‌ی سؤال یک جمله ناتمام است و گزینه‌ها (پیشنهادی و کلید) تکمیل‌کننده‌ی آن هستند.

### مثال

مهم‌ترین ویژگی هدف‌های آموزشی رفتاری این است که .....

(×) الف- برحسب عملکرد قابل اندازه‌گیری نوشته می‌شوند.

ب- مهارت‌های سطح بالای یادگیری را شامل می‌شوند.

پ- مطالب جزئی محتوای درس را دربرمی‌گیرند.

ث- دانش‌آموزان به راحتی آنها را می‌آموزند.

نوع تنها گزینه‌ی درست

در این نوع سؤال‌ها، همه‌ی گزینه‌ها بجز گزینه‌ی کلید به طور قطع غلط هستند. بنابراین از این سؤال‌ها بیشتر در ریاضیات و علوم می‌توان استفاده کرد.

مثال:

چه کسی چرخ خیاطی را اختراع کرد؟

الف- فولتون

(×) ب- هوو

پ- سینگر

ت- وایت

ث- ویتنی (نیتکو، ۲۰۰۱، ص ۱۴۷)

نوع بهترین گزینه

در این نوع سؤال‌ها، همه‌ی گزینه‌ها به طریقی هستند، اما یکی از آنها از همه درست‌تر است. این نوع سؤال‌های چندگزینه‌ای در زمینه‌های علوم انسانی و علوم اجتماعی که برای سؤال‌ها معمولاً جواب قطعی یافت نمی‌شود مفیدند.

### مثال

بنا به تعریف، می‌توان گفت که یادگیری عبارت است از ..... .

الف- کسب اطلاعات تازه در نتیجه‌ی آموزش

(×) ب- ایجاد تغییرات نسبتاً پایدار در توان رفتار یادگیرنده

پ- ایجاد تغییرات مطلوب و موردنظر در رفتار یادگیرنده

ت- توانایی به خاطر سپردن و یادآوری تجارب پیشین

### نوع منفی

در این نوع سؤال‌ها، تنه‌ی سؤال به صورت منفی بیان می‌شود و همه‌ی گزینه‌ها بجز یکی از آنها درست هستند.

### مثال

کدام مورد زیر از قواعد تهیه‌ی آزمون‌های چندگزینه‌ای نیست؟

الف- مربوط بودن همه‌ی گزینه‌ها به موضوعی واحد

(×) ب- سنجش تنها یک هدف با هر سؤال

پ- قرار دادن حداکثر دو مطلب در هر سؤال

ت- استفاده نکردن از گزینه‌ی "همه‌ی موارد بالا"

یکی از مشکلات تهیه‌ی سؤال‌های چندگزینه‌ای این است که بعضی وقت‌ها تهیه‌ی تعدادی گزینه‌ی انحرافی غلط که به ظاهر درست جلوه کنند بسیار دشوار است. در این‌گونه موارد، چنانچه تهیه‌ی گزینه‌های درست آسان‌تر باشد می‌توان از سؤال‌های نوع منفی که در آنها همه‌ی گزینه‌ها بجز یکی درست‌اند استفاده کرد، مانند نمونه‌ی بالا.

### نوع جایگزینی

در سؤال‌های نوع جایگزینی، در متن یا تنه‌ی سؤال غلط‌هایی گنجانیده می‌شود و آزمون‌شونده باید آنها را با گزینه‌های پیشنهادی جایگزین کند.

### مثال

"به درستی که از نیروهای آموزش و پرورش باید به طور کامل استفاده شود تا جوانان از ماهیت واقعی خطراتی که دموکراسی را تهدید می‌کنند آگاه شوند زیرا"<sup>(۱)</sup> هیچ جای دیگری بجز آموزشگاه نمی‌تواند فرصت‌های خوب یا بهتری<sup>(۲)</sup> در اختیار جوانان قرار دهد تا به طور واقعی<sup>(۳)</sup> به کُنه مسائل پی ببرند."

مطلب بالا را بخوانید و با توجه به سؤال‌هایی که در زیر آمده برای قسمت‌هایی که با شماره مشخص شده‌اند جایگزین کنید.

۱-

(×) الف-، زیرا

ب-؛ زیرا

پ- - زیرا

ت- هیچ علامتی قبل از زیرا لازم نیست.

۲-

الف- فرصت‌های خوب یا خوب‌تری

ب- فرصت‌های بهتر یا بهتری

پ- فرصت‌های بهتر یا برتری

(×) ت- فرصت‌های بهتری

۳-

الف- به صورت منطقی

(×) ب- به طور منطقی

پ- به طور عقلانی

ت- به صورت علمی

(نیتکو، ۲۰۰۱، ص ۱۴۷)

### نوع ناتمام پیشنهادی

در این نوع سؤال‌ها، آزمون‌شونده با توضیحاتی که در متن سؤال می‌آید راهنمایی می‌شود و او باید با توجه به آن راهنمایی‌ها جواب درست سؤال را پیدا کند.

### مثال

به سیبی که مزه‌اش ترش و شیرین است چه نوع سیبی گفته می‌شود (جواب یک کلمه سه حرفی است، کدام گزینه‌ی زیر حرف اول آن کلمه است؟)

الف- ت

(×) ب- م

پ- ش

ت- گ

### نوع پاسخ ترکیبی

در این نوع سؤال‌ها، جمله‌های یک پاراگراف به هم ریخته‌اند و از آزمون‌شونده خواسته می‌شود تا آنها را پشت سرهم مرتب کند.

### مثال

جمله‌های زیر را به چه ترتیبی باید پشت سرهم قرار دهیم تا از مجموع آنها یک پاراگراف مطلب معنی‌دار به دست آید؟

... ۱	جمله‌ها
... ۲	
... ۳	
... ۴	
... ۵	

گزینه‌ها:

الف- ۱، ۲، ۳، ۴، ۵

ب- ۵، ۲، ۳، ۱، ۴

پ- ۴، ۲، ۳، ۱، ۵

ت- ۵، ۴، ۳، ۲، ۱

### قواعد تهیهی سؤال‌های چندگزینه‌ای

قواعدی که در اینجا برای تهیهی سؤال‌های چندگزینه‌ای پیشنهاد می‌شوند براساس این طرز فکر تدوین شده‌اند که یک سؤال چندگزینه‌ای باید آزمون‌شونده را با تکلیفی که هم مهم و هم کاملاً برای او قابل فهم است مواجه سازد. هر سؤال باید طوری نوشته شود که تنها به وسیلهی کسانی که هدف مورد اندازه‌گیری (جواب سؤال) را می‌دانند به درستی پاسخ داده شود، و کسانی که تسلط کامل بر مطلب ندارند نتوانند به آن پاسخ درست بدهند.

۱- هر سؤال را به گونه‌ای انتخاب کنید که یک موضوع مهم یا یک هدف آموزشی را اندازه بگیرد. در طرح هر سؤال، نویسندهی سؤال باید از خود بپرسد که "آیا دانستن و ندانستن پاسخ سؤال در صلاحیت آزمون‌شونده، از

لحاظ مطلب مورد آزمایش، تأثیر دارد یا نه؟" اگر جواب این سؤال مثبت بود آن گاه باید به گنجاندن موضوع در سؤال اقدام کند. بنابراین، هر سؤال باید یک هدف آموزشی را بسنجد و شامل بخش مهمی از محتوای درس باشد. نیتکو (۱۹۸۳) در این باره می گوید ابتدا یک هدف آموزشی یا یک مطلب مهم را انتخاب کنید و تصمیم بگیرید که برای آن چند سؤال تهیه خواهید کرد. بعد برای نوشتن هر سؤال اقدامات زیر را انجام دهید.

الف- با نوشتن یک جمله‌ی سؤال‌ی یا یک جمله‌ی ناتمام تنه‌ی سؤال را بنویسید.

ب- جواب سؤال یا بخش ناتمام جمله را به صورت گزینه‌ی کلید بنویسید.

پ- گزینه‌های انحرافی را به صورتی بنویسید که نظر دانش‌آموزان نامطلع را به خود جلب کنند.

به مطلب و هدف آموزشی زیر و سؤال نمونه‌ای که برای آن آماده شده است توجه کنید.

مطلب درسی: اگر زمان آزمون محدود باشد، یک آزمون چندگزینه‌ای از یک آزمون تشریحی نمونه‌ی بهتری از هدف‌ها و محتوا را به دست می‌دهد.

هدف آموزشی: دانشجو بتواند با ذکر دلیل توضیح دهد که، برای استفاده در یک زمان محدود، آزمون‌های چندگزینه‌ای از آزمون‌های تشریحی مناسب‌ترند.

#### سؤال نمونه:

زمانی که وقت آزمون محدود و مقدار مطلبی که باید سنجش شود گسترده است، سؤال‌های چندگزینه‌ای از سؤال‌های تشریحی بهترند، زیرا سؤال‌های چندگزینه‌ای ...

الف- راحت‌تر تصحیح می‌شوند.

ب- امکان کمی برای حدس زدن به آزمون‌شونده می‌دهند.

پ- نمونه‌ی بزرگتری از مطالب را شامل می‌شوند. (x)

ت- توانایی‌های سطح بالای یادگیرندگان را می‌سنجند.

۲- بیشتر از یک هدف یا یک مطلب در هر سؤال قرار ندهید. هر سؤال باید تنها به یک مطلب یا یک هدف آموزشی مربوط باشد. اگر بیش از یک مطلب یا یک هدف در سؤال مطرح شود، نه تنها سؤال را پیچیده می‌کند بلکه اگر آزمون‌شونده در پاسخگویی به سؤال موفق نباشد معلوم نیست که کدام یک از مطالب را ندانسته و علت عدم موفقیت او چه بوده است.

۳- سؤال‌ها را کاملاً روشن و در حد درک آزمون‌شوندگان بنویسید. استفاده از جمله‌ها و عبارات‌های مبهم در سؤال‌های چندگزینه‌ای ممکن است باعث شود که آزمون‌شوندگان با وجود اینکه جواب درست را می‌دانند به علت ابهام و دشواری متن سؤال نتوانند به آن جواب درست بدهند.

به‌ویژه سؤال‌هایی که برای دانش‌آموزان سطح‌های پایین‌تر تحصیلی نوشته می‌شوند باید با زبانی ساده و درخور درک آنان باشند. سؤال‌هایی که در آن‌ها کلمه‌ها و عبارتهای دشوار و غیرقابل درک دانش‌آموزان به کار می‌روند نمی‌توانند به خوبی توانایی موردنظر را سنجش کنند، زیرا اگر دانش‌آموزان در جواب دادن به این نوع سؤال‌ها موفق نباشند معلوم نمی‌شود که علت شکست آنان ندانستن جواب سؤال یا ناتوانی در درک سؤال است. میزان دشواری یک سؤال باید به محتوا و هدف آموزشی مورد سنجش مربوط باشد نه به کلمه‌ها و عبارتهایی که در آن به کار می‌روند. تنها در مواردی که سؤال برای اندازه‌گیری توانایی خواندن و فهمیدن یا توانایی کلامی آزمون‌شوندگان نوشته می‌شود استفاده از ساختمان پیچیده‌ی دستوری و لغات و عبارات دشوار مجاز است.

۴- تکرار مطالب در گزینه‌ها خودداری کنید. تا آنجا که ممکن است پاسخ‌ها را مختصر و کوتاه بنویسید. ضمناً سعی کنید مطالب را یک بار آن هم در متن یا تنه‌ی سؤال بیاورید و در صورت امکان از تکرار آن در پاسخ‌ها خودداری نمایید.

## مثال

### سؤال ضعیف:

در آزمون‌های عینی، اصطلاح عینی.....

الف- به روش مشخص کردن نتایج یادگیری مربوط است.

ب- به روش انتخاب محتوای آزمون مربوط است.

(×) پ- به روش تصحیح جواب‌ها مربوط است.

ت- به روش ارائه مطالب مربوط است.

### سؤال بهتر:

در آزمون‌های عینی، اصطلاح عینی به روش ..... مربوط است.

الف- مشخص کردن نتایج یادگیری

ب- انتخاب محتوای آزمون

(×) پ- تصحیح جواب‌ها

ت- ارائه مطالب

### سؤال باز هم بهتر:

در آزمون‌های عینی، اصطلاح عینی به چه چیزی مربوط است؟

الف- مشخص کردن نتایج یادگیری

ب- انتخاب محتوای آزمون

(×) پ- تصحیح جوابها

ت- ارائه مطالب

فایده‌ی رعایت قاعده‌ی بالا و جلوگیری از تکرار مطالب در گزینه‌ها صرفه‌جویی در وقت خواندن برای آزمون‌شوندگان و نیز صرفه‌جویی در وقت نوشتن و تایپ کردن برای طراح سؤال و تکثیرکننده‌ی آزمون است.

**۵- مطالب اصلی سؤال را به طور کامل در تنه‌ی سؤال بنویسید.** تنه‌ی سؤال باید شامل مطالب مهم سؤال باشد، به گونه‌ای که آزمون‌شونده، پس از خواندن و پیش از مراجعه به گزینه‌ها، اندیشه‌ی اصلی مورد سؤال را درک کند و بداند که سؤال به چه مطلبی مربوط است. مشکل سؤال‌هایی که این نکته را در نظر نمی‌گیرند این است که آزمون‌شونده، پیش از خواندن گزینه‌ها، متوجه نخواهد شد که سؤال درباره‌ی چه مطلبی نوشته شده است. یک راه کسب اطمینان از این امر آن است که روی گزینه‌ها را بپوشانید و تنه‌ی سؤال را به تنهایی بخوانید تا متوجه شوید که مسئله یا مطلب به طور کامل بیان شده است یا نه. اگر جواب شما منفی است در تنه‌ی سؤال تجدیدنظر کنید.

**مثال**

**سؤال ضعیف:**

جدول مشخصات.....

الف- نشان می‌دهد که چگونه می‌توان از آزمون برای بهبود امر آموزش و یادگیری استفاده کرد.

(×) ب- نمونه‌ی بهتری از محتوای درس و هدف‌های آموزش به دست می‌دهد.

پ- روش مورد استفاده در تصحیح پاسخ‌های آزمون را مشخص می‌کند.

ت- عینیت آزمون را افزایش می‌دهد.

**سؤال بهتر:**

مهم‌ترین فایده‌ی استفاده از جدول مشخصات در تهیه‌ی آزمون‌های پیشرفت تحصیلی آن است که.....

الف- آموزش و یادگیری را بهبود می‌بخشد.

(×) ب- نمونه‌ی بهتری از محتوای درس و هدف‌های آموزش به دست می‌دهد.

پ- روش مورد استفاده در تصحیح پاسخ‌های آزمون را مشخص می‌کند.

ت- عینیت آزمون را افزایش می‌دهد.

**۶- بکشید تا همه‌ی گزینه‌های یک سؤال متجانس و به موضوع واحدی مربوط باشند.** همه‌ی گزینه‌های

یک سؤال باید از لحاظ محتوا همگون یا متجانس باشند نه اینکه هر یک از آنها به مطلب جداگانه‌ای اشاره کند. در



چنین حالتی سؤال از صورت چندگزینه‌ای خارج شده و هر یک از گزینه‌ها به صورت یک سؤال صحیح- غلط درمی‌آید که از دیگر گزینه‌ها مستقل است. نمونه‌ی معمول این گونه سؤال‌ها سؤالی است که تنه‌ی آن این جمله را شامل می‌شود: "کدام یک از مطالب یا بیانات زیر درست است؟" و گزینه‌های آن هر یک به مطلب جداگانه‌ای اشاره می‌کند.

**مثال:**

### سؤال ضعیف

کدام یک از موارد زیر درست است؟

- (×) الف- سؤال‌های انشایی بهترین نوع سؤال برای ارزشیابی هدف‌های طبقه‌ی آفریدن هستند.
- ب- تهیه‌ی جدول مشخصات برای آزمون‌های انشایی مفید نیست.
- پ- سؤال‌های عینی را تنها در کلاس‌های دانشگاه می‌توان به اجرا درآورد.
- ت- هدف‌های رفتاری را دیرباز مورد استفاده معلمان قرار داشته‌اند.

### سؤال بهتر:

- در رابطه با استفاده از سؤال‌های انشایی و ارزشیابی هدف‌های مختلف آموزشی، کدام مورد زیر درست است؟
- (×) الف- بهترین نوع سؤال برای ارزشیابی هدف‌های طبقه‌ی آفریدن است.
- ب- برای ارزشیابی هدف‌های هیچ یک از طبقات حوزه‌ی شناختی، مفید نیستند.
- پ- بین سؤال‌های انشایی و عینی از لحاظ ارزشیابی هدف‌های گوناگون چندان تفاوتی وجود ندارد.
- ت- بهتر است از سؤال‌های انشایی به ندرت استفاده شود.

تجانس گزینه‌ها یکی از راه‌های کنترل سطح دشواری سؤال‌هاست. قاعده‌ی کلی این است که هر چه گزینه‌ها متجانس‌تر باشند به همان نسبت سؤال دشوارتر است. البته چنان‌که در مورد سؤال‌های جور کردنی نیز گفته شد تجانس یک امر نسبی است. گزینه‌هایی که برای یک گروه دانش‌آموز متجانس به نظر می‌رسند ممکن است برای گروه دیگر (معمولاً گروه سطح بالاتر) از تجانس کمتری برخوردار باشند.

۷- سؤال را طوری بنویسید که پاسخ درست تنها پاسخ درست یا قطعاً درست‌ترین پاسخ باشد. قبلاً گفتیم که گزینه‌ی درست هر سؤال یا باید تنها گزینه‌ی درست در میان سایر گزینه‌ها باشد یا از گزینه‌های انحرافی به طور قطع درست‌تر یا بهتر باشد.

**مثال**

### سؤال ضعیف:

یک پسر بچه‌ی کوررنگ، کوررنگی خود را از چه کسی به ارث می‌برد؟

الف- پدر

ب- مادر

پ- پدرِ مادر

ت- پدرِ پدر

ث- نیاکان دور

سؤال بهتر:

یک پسر بچه‌ی کوررنگ به احتمال بسیار زیاد کوررنگی خود را از چه کسی به ارث می‌برد؟

الف- پدر

(x) ب- مادر

پ- پدرِ پدر

ت- مادرِ مادر

ث- نیاکان دور

(ثرندایک و هیگن، ۱۹۶۹)

ثرندایک و هیگن درباره‌ی سؤال‌های بالا اظهار می‌دارند که گرچه در سؤال ضعیف گزینه‌ی ب درست است، گزینه‌های پ و ت و ث هم کاملاً غلط نیستند. در سؤال تجدیدنظر شده (سؤال بهتر) اضافه شدن کلمه‌ی احتمالاً گزینه‌ی ب را از دیگر گزینه‌ها درست‌تر می‌سازد.

۸- برای اندازه‌گیری فرآیندهای پیچیده‌ی ذهنی از موقعیت‌های تازه استفاده کنید. باید به خاطر داشت که مسائل و موقعیت‌هایی که در کلاس و کتاب برای آموزش فعالیت‌های ذهنی چون کار بستن، ترکیب، و از این قبیل به کار رفته‌اند دیگر نمی‌توانند برای سنجش این توانایی‌ها مورد استفاده قرار گیرند. اگر این مطالب در سؤال‌ها تکرار شوند، تنها دانش را می‌سنجند. بنابراین، در طرح سؤال‌هایی که به منظور سنجش توانایی‌های بالاتر از سطح دانش به کار می‌روند از موقعیت‌ها و مثال‌های تازه‌ای که با مثال‌های کتاب و موقعیت‌های به کار رفته در کلاس متفاوت‌اند استفاده کنید. پرن‌دایک و هیگن (۱۹۶۹) به عنوان نمونه‌ای برای روشن کردن این مورد، دو سؤال زیر را معرفی کرده‌اند. سؤال اول یک سؤال در سطح یادآوری است که محفوظات یادگیرنده را می‌سنجد، ولی سؤال دوم یک سؤال کاربردی است که در آن یک موقعیت تازه ارائه شده است.

سؤال سطح یادآوری:

کدام یک از موارد زیر وقتی که در بدن مصرف می‌شود بیشترین کالری را تولید می‌کند؟

(×) الف- یک گرم چربی

ب- یک گرم شکر

پ- یک گرم نشاسته

ت- یک گرم پروتئین

### سؤال کاربردی:

کدامیک از موارد زیر اگر از غذاهای روزانه حذف شود بیشترین مقدار کاهش کالری را به دنبال خواهد داشت؟

(×) الف- یک قاشق کره

ب- یک قاشق شکر

پ- یک قرص نان

ت- یک تخم‌مرغ آب‌پز

در زیر به سؤال دیگری که برای اندازه‌گیری توانایی کار بستن اصول نوشته شده و در آن یک موقعیت جدید مطرح گردیده است توجه کنید:

فرض کنید ۴ لیتر رنگ برای رنگ کردن مساحتی برابر ۴۰ مترمربع کافی باشد، چند لیتر رنگ برای رنگ کردن ۴ دیولر اتاقی با طول و عرض ۱۰×۳ متر لازم است؟

الف- ۶

ب- ۸

(×) پ- ۱۲

ت- ۱۸

۹- گزینه‌های انحرافی را طوری بنویسید که توجه آزمون‌شوندگان بی‌اطلاع از موضوع سؤال را به خود جلب کنند. هدف از گزینه‌های انحرافی در سؤال‌های چندگزینه‌ای این است که توجه به آزمون‌شوندگان بی‌اطلاع مورد سؤال را به خود جلب کند. این گزینه‌ها باید ظاهری منطقی داشته باشند و در یک بررسی سطحی درست جلوه کنند. یعنی نباید آشکارا غلط باشند، آن‌گونه که هر کسی بتواند غلط بودن آنها را تشخیص دهد.

### مثال

#### سؤال ضعیف:

امیرکبیر صدراعظم کدامیک از شاهان زیر بود؟

الف- شاه عباس

ب- سلطان محمود

(×) پ- ناصرالدین شاه

ت- نادرشاه

### سؤال بهتر:

امیرکبیر صدراعظم کدام یک از شاهان زیر بود؟

الف- مظفرالدین شاه

ب- محمدعلیشاه

(×) پ- ناصرالدین شاه

ت- احمدشاه

برای نوشتن گزینه‌های انحرافی پیشنهادهای زیر را مورد توجه قرار دهید:

- از غلطی‌های معمول دانش‌آموزان استفاده کنید.
  - از کلماتی که مبهم به نظر می‌رسند (مانند بااهمیت، معنی‌دار، و دقیق) و به تنه‌ی سؤال ربط پیدا می‌کنند استفاده کنید، اما در این کار افراط ننمایید.
  - از کلماتی که با تنه‌ی سؤال تداعی دارند استفاده کنید (مانند، سیاست‌مدار، سیاسی).
  - از زبان کتاب درسی و عبارتهای دیگری که ظاهری درست‌نما دارند استفاده کنید.
  - از پاسخ‌های غلطی که از کج‌فهمی‌ها و بی‌دقتی‌های دانش‌آموزان ناشی می‌شوند استفاده کنید (مثلاً اینکه فراموش می‌کنند که دسیمتر را به سانتیمتر تبدیل نمایند).
  - از گزینه‌های انحرافی که از لحاظ محتوا متجانس و شبیه به گزینه‌ی درست هستند استفاده کنید.
  - از گزینه‌های انحرافی که در شکل ظاهر با تنه‌ی سؤال موازی و از لحاظ دستوری با آن هماهنگی دارند استفاده کنید.
  - گزینه‌های انحرافی را از لحاظ طول، واژگان، ساختار جمله، و پیچیدگی موضوع شبیه به گزینه‌ی درست بنویسید.
- احتیاط: گزینه‌های انحرافی باید پاسخ‌دهندگان بی‌اطلاع از موضوع سؤال را منحرف کنند، اما نباید برای پاسخ‌دهندگان مسلط بر موضوع سؤال گمراه‌کننده باشند.

ازلین و گرانلاند (۲۰۰۰)

۱۰- گزینه‌های هر سؤال را طوری بنویسید که از نظر دستوری و جمله‌بندی به نحو درست مکمل متن سؤال باشند. یکی از معایب سؤال‌های چندگزینه‌ای افراد تازه‌کار این است که بین تنه‌ی سؤال و گزینه‌های پیشنهادی از لحاظ دستوری هماهنگی وجود ندارد، و این ناهماهنگی در رابطه با گزینه‌های غلط بیشتر به چشم می‌خورد. در نتیجه، این امر به صورت یک راهنما برای آزمون‌شونده درمی‌آید و متوجه می‌شود که گزینه‌های غلط آنهایی هستند که از نظر دستوری تنه‌ی سؤال را کامل نمی‌کنند. بنابراین، سعی کنید سؤال‌ها را طوری بنویسید که همه‌ی گزینه‌ها از لحاظ دستوری با تنه‌ی سؤال هماهنگی داشته باشند.

این نکته به ویژه در رابطه با سؤال‌هایی که متن آنها نیمه تمام و پاسخ‌ها مکمل آن هستند باید به دقت رعایت شود. همچنین در سؤال‌هایی که آخرین گزینه‌های آنها همه‌ی موارد بالا با هیچ یک از موارد بالا است غالباً این مشکل دیده می‌شود. به سؤال زیر که از این لحاظ اشکال دارد توجه کنید:

نزدیک به هفتاد درصد جمعیت مردم یک مملکت افرادی هستند که هوشبهر آنان.....

الف- بین ۱۰۰ و ۱۵۰ است.

ب- بین ۵۰ و ۱۰۰ است.

پ- بین ۸۵ و ۱۱۵ است. (x)

ت- هیچ‌کدام درست نیست.

دقت کنید که اگر عبارت "هیچ‌کدام درست نیست" را در دنباله تنه‌ی سؤال قرار دهیم، یک جمله‌ی درست و منطقی به دست نمی‌آید. بنابراین، یا کلاً باید سؤال بالا را از لحاظ شکل ظاهری تغییر دهیم یا گزینه‌ی ت را به گونه‌ای تغییر دهیم که از لحاظ دستوری مکمل تنه‌ی سؤال باشد.

سؤال زیر نمونه‌ی دیگری از این سؤالات ضعیف است.

چرا زندگی شهری و روستایی در ناحیه‌ی کنار خزر به هم وابسته است؟

الف- شهرها بازار فروش محصولات روستایی است. (x)

ب- شهرها بازار فروش محصولات روستایی است.

پ- وجود بازارهای هفتگی

ت- همه‌ی موارد بالا

پیداست که هیچ یک از گزینه‌های سؤال بالا، به ویژه گزینه‌ی ت، از نظر دستوری و جمله‌سازی جواب درست و منطقی سؤال نیست. برای اصلاح این سؤال باید شکل کلی سؤال را تغییر داد. باز هم به مثالی دیگر توجه کنید.

**سؤال ضعیف:**

یک مبدل برق (ترانسفورماتور) را می‌توان برای..... به کار برد.

الف- ذخیره‌سازی برق

(×) ب- افزایش ولتاژ یک جریان برق متناوب

پ- انرژی برقی را به انرژی مکانیکی تبدیل می‌کند.

ت- جریان متناوب را به جریان مستقیم تبدیل می‌کند.

**سؤال بهتر:**

یک مبدل (ترانسفورماتور) برقی را می‌توان ..... برای به کار برد.

الف- ذخیره‌سازی برق

(×) ب- افزایش ولتاژ یک جریان برق متناوب

پ- تبدیل انرژی برقی به انرژی مکانیکی

ت- تغییر جریان متناوب به جریان مستقیم

(لین و گراندلند، ۲۰۰۰)

۱۱- از به کار بردن اشاره‌های دستوری نامربوط خودداری کنید. اشاره‌هایی از قبیل زمان افعال، مفرد و جمع

بودن آنها، و یا حروف تعریف a و an در زبان انگلیسی را در سؤال‌ها آن‌گونه که آزمون‌شونده با توجه به این اشاره‌ها

گزینه‌ی درست را انتخاب کند به کار نبرید. در ضمن اگر تنه‌ی سؤال و گزینه‌ی صحیح از بعضی جهات مشابه باشند

ممکن است آزمون‌شونده جواب درست را با توجه به تشابه ظاهری پیدا کند، بدون اینکه واقعاً پاسخ سؤال را بداند. به

سؤال زیر که از این جهت دارای اشکال است توجه کنید:

اسناد مالکیت املاک مردم در کجا ثبت می‌شوند؟

(×) الف- اداره کل ثبت اسناد و املاک

ب- اداره کل ثبت احوال

پ- وزارت امور اقتصادی و دارایی

ت- وزارت علوم و آموزش عالی

به نمونه‌ی دیگری از این‌گونه سؤال‌های ضعیف توجه کنید.

پایتخت کشور مکزیک کجاست؟

الف- آکاپولکو

ب- مریاد

(×) پ- شهر مکزیک

ت- واشنگتن

۱۲- در سؤالات منفی کلمات منفی را برجسته جلوه دهید. برجسته جلوه دادن کلمات منفی به آزمون شونده کمک می کند تا متوجه منفی بودن سؤال باشد. اگر آزمون به صورت چاپی در اختیار آزمون شونده قرار می گیرد، می توان از طریق سیاه نوشتن کلمات منفی آنها را برجسته کرد، اما در سؤالات آزمون هایی که به صورت دستنویس یا تایپی در اختیار آزمون شونده قرار می گیرند می توان زیر کلمات منفی خط کشید. به مثال زیر توجه کنید.

در طرح سؤال های چندگزینه ای تمامی موارد زیر را باید رعایت کرد بجز:

(×) الف- کوتاه کردن متن اصلی سؤال یا طولانی کردن پاسخها

ب- نوشتن متن اصلی سؤال به صورت یک جمله ای استفهامی

پ- خط کشیدن زیر عبارت منفی یا برجسته کردن آن به طریقی دیگر

ت- استفاده از موقعیت های واقعی در سؤال های کاربردی

به عنوان یک قاعده ی کلی، بهتر است از کاربرد کلمات و عبارات منفی مانند بجز، هرگز، و غیره پرهیز شود یا به حداقل کاهش یابد. تنها در صورتی باید از آنها استفاده شود که این کار ضروری باشد. رینالدز و لیوینگستون (۲۰۱۲) گفته اند در بعضی موارد مهم است آزمون شونده بداند که چه کاری را نباید انجام دهد. مثلاً "اگر بوی گاز شنیدید چه کاری نباید انجام دهید؟"

۱۳- از نوشتن سؤال هایی که در آنها متن سؤال منفی و گزینه ها هم منفی هستند، یعنی منفی مضاعف، احتراز کنید. سؤال هایی که در آنها هم متن سؤال منفی است و هم گزینه ها با عبارات و کلمات منفی نوشته می شوند بسیار دشوار هستند، و خواندن آنها آزمون شونده را گیج و سردرگم می کند.

مثال:

سؤال ضعیف:

در رابطه با استفاده از سؤال های انشایی در ارزشیابی هدف های آموزشی، نمی توان گفت که...

(×) الف- سؤال های انشایی برای ارزشیابی هدف های طبقه ی ترکیب بهترین نوع نیستند.

ب- سؤال های انشایی برای ارزشیابی همه ی طبقات حوزه ی شناختی مفید نیستند.

پ- سؤال های انشایی و عینی تفاوت زیادی با یکدیگر ندارند.

ت- استفاده از سؤال های عینی برای هدف های دانش همیشه صلاح نیست.

شکل بهتر این سؤال در صفحه‌ی ۲۰۴ آمده است.

۱۴- از کاربرد عبارت‌هایی مانند "همه‌ی آنچه در بالا گفته شد"، "تمامی موارد بالا" و مانند اینها

خودداری کنید. سؤال‌هایی که در آنها این‌گونه عبارت‌ها به کار می‌روند، به ویژه اگر به جای گزینه‌ی درست سؤال به کار روند، بسیار آسان هستند زیرا اگر آزمون‌شونده بداند که حداقل دو گزینه از گزینه‌ها درست است، چه از دست و غلط بودن بقیه‌ی گزینه‌ها اطلاع داشته باشد چه نداشته باشد، گزینه‌ی صحیح را انتخاب خواهد کرد. به جای این‌گونه سؤال‌ها می‌توانید از سؤال‌های دیگری نظیر آنچه در زیر آمده است استفاده کنید.

مثال

در ترکیب هوا در اتاقی که هیچ‌گونه منفذی به بیرون ندارد و تنها تعدادی گیاه در آن قرار دارند چه تغییراتی رخ می‌دهد؟

الف- اکسیژن کاهش می‌یابد.

ب- دی‌اکسیدکربن افزایش می‌یابد.

پ- اکسیژن افزایش می‌یابد.

ت- هم دی‌اکسیدکربن و هم اکسیژن افزایش می‌یابد.

(×) ث- اکسیژن کاهش می‌یابد ولی دی‌اکسیدکربن افزایش می‌یابد.

(مهرنر و لهما، ۱۹۸۲)

اشکال دیگر این‌گونه سؤال‌ها آن است که ممکن است آزمون‌شونده گزینه‌ی الف را که خواند و متوجه شد که آن گزینه درست است همان را انتخاب کند و به سراغ گزینه‌های بعدی نرود.

۱۵- تا حد امکان از کاربرد گزینه‌ی "هیچ یک از موارد بالا" پرهیز کنید. صاحب‌نظران سنجش و اندازه‌گیری استفاده از گزینه "هیچ‌کدام" یا "هیچ‌یک از موارد بالا" را پیشنهاد نمی‌کنند. بهترین مورد استفاده این گزینه در ارتباط با سؤال‌های درسهای ریاضیات است؛ با این حال، بعضی از متخصصان استفاده از این گزینه را حتی با سؤال‌های ریاضی هم مفید نمی‌دانند. از جمله پین (۲۰۰۳) در رابطه با نگرانی خود مورد زیر را ذکر کرده است:

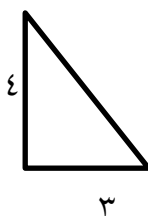
سؤال ضعیف:

دو ضلع مجاور زاویه‌ی قائمه یک مثلث قائم‌الزاویه ۳ و ۴ سانتیمتر است. مساحت این مثلث چند سانتیمترمربع

است؟

الف- ۷

ب- ۱۲





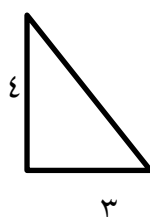
پ- ۲۵

(×) ت- هیچ کدام از موارد بالا

جواب سؤال بالا ۶ سانتیمتر است و دانش آموزی که جواب سؤال را می داند گزینه ی ت را انتخاب خواهد کرد. اما دانش آموز دیگری که مسئله را اشتباه حل می کند (مثلاً به جای مساحت مثلث وتر آن را که ۵ سانتیمتر است محاسبه می نماید و جواب ت را انتخاب می کند) از یک راه حل غلط به یک جواب درست می رسد. بنابراین، گزینه ی "هیچ یک از موارد بالا" فقط زمانی درست عمل می کند که از آن به عنوان گزینه ی غلط استفاده شود و گزینه ی درست در میان سایر گزینه ها باشد.

سؤال بهتر:

دو ضلع مجاور زاویه ی قائمه یک مثلث قائم الزاویه ۳ و ۴ سانتیمتر است، مساحت این مثلث چند سانتیمتر مربع است؟



(×) الف- ۶

ب- ۷

پ- ۱۲

ت- هیچ کدام از موارد بالا

در غیر از مسائل و محاسبات ریاضی، استفاده از گزینه ی "هیچ یک از موارد بالا" پیشنهاد نمی شود زیرا در سؤال هایی که "هیچ یک از موارد بالا" گزینه ی درست است و دانش آموز آن را انتخاب می کند و به ما نشان می دهد که هیچ کدام از گزینه های پیشنهادی درست نیست، باز هم نخواهیم فهمید که او جواب درست سؤال را می داند یا نه. به مثال زیر توجه کنید:

مرکز استان مازندران کدام یک از شهرهای زیر است؟

الف- چالوس

ب- آمل

پ- بابل

(×) ت- هیچ کدام

دانش آموزی که گزینه ی ت (گزینه ی درست) سؤال بالا را انتخاب می کند می داند که شهرهای چالوس، آمل و بابل مرکز استان مازندران نیستند، اما بر ما معلوم نمی شود که او جواب درست سؤال را که شهرسازی است می داند یا نه. نکته ی دیگری که موضوع را پیچیده تر می کند این است که دانش آموزان قوی بعضی وقتها در گزینه ی صحیح

اشکال می‌بینند و به همین دلیل گزینه‌ی "هیچ‌کدام" را انتخاب می‌کنند" (ریس<sup>۱</sup>، براون<sup>۲</sup>، و اسمیت<sup>۳</sup>، ۲۰۰۵، ص ۴۷).

۱۶- از کاربرد همیشه و هرگز در گزینه‌ها پرهیز کنید. در استفاده از همیشه و هرگز احتیاط کنید، زیرا این کلمات صرفاً در رابطه با ریاضیات قابل توجیه‌اند. آزمون‌شوندگان زیرک این را می‌دانند و از این اطلاعات برای کنار گذاشتن این گونه گزینه‌ها استفاده می‌کنند.

۱۷- سؤال‌ها را مستقل از یکدیگر بنویسید. پاسخ دادن به یک سؤال نباید شرط لازم برای پاسخ دادن به سؤال دیگری باشد، و جواب یک سؤال نباید به پاسخ دادن به سؤال دیگری کمک بکند. فرض کنید در یک آزمون ریاضی دو سؤال پشت سرهم آمده باشند که در سؤال اول از آزمون‌شونده خواسته شده باشد تا مساحت یک شکل را تعیین کند و در سؤال دوم از او خواسته شده باشد که هزینه‌ی نوعی کفپوش را برای مساحت تعیین شده به دست آورد. اگر جواب آزمون‌شونده به سؤال اول غلط باشد، جواب سؤال دوم نیز الزاماً غلط خواهد بود. اگر آزمون‌شونده نحوه‌ی محاسبه‌ی هزینه‌ی کفپوش را بداند اما، به علت یک اشتباه کوچک در محاسبه، سؤال اول را غلط جواب دهد، برای هیچ یک از دو سؤال نمره‌ای نخواهد گرفت.

۱۸- از طرح سؤال‌های گمراه‌کننده پرهیزید. استفاده از سؤال‌های گمراه‌کننده در آزمون‌های پیشرفت تحصیلی به هیچ وجه جایز نیست. اغلب اتفاق می‌افتد آزمون‌شوندگان مطلع‌تر که اطلاعات وسیع‌تری دارند به وسیله‌ی این‌گونه سؤال‌ها گمراه می‌شوند. این‌گونه سؤال‌ها با هدف اصلی آزمون‌های پیشرفت تحصیلی که اندازه‌گیری میزان فهم و دانش آزمون‌شوندگان است مغایرت دارد.

## مثال

### سؤال گمراه‌کننده:

با توجه به اینکه نیروی جاذبه‌ی کره‌ی زمین شش برابر نیروی جاذبه‌ی کره‌ی ماه است، اگر جرم جسمی در روی زمین ۶۰ کیلوگرم باشد، جرم این جسم در کره‌ی ماه چند کیلوگرم است؟

الف- ۰/۶

ب- ۶

پ- ۱۰

---

<sup>۱</sup>-Race

<sup>۲</sup>-Brown

<sup>۳</sup>-Smith

(x) ت - ۶۰

توضیح مقدمه تنه‌ی سؤال بالا ممکن است برای تعدادی از آزمون‌شوندگان گمراه‌کننده باشد. برای اصلاح این سؤال باید آن توضیح را به کلی حذف کرد.

### سؤال بهتر:

اگر جرم جسمی در روی زمین ۶۰ کیلوگرم باشد، جرم این جسم در کره‌ی ماه چند کیلوگرم است؟

الف-۰/۶

ب- ۶

پ-۱۰

(x) ت - ۶۰

۱۹- دو گزینه‌ی متضاد را که یکی از آنها درست است به کار نبرید. هر گاه در میان گزینه‌ها تنها یک زوج گزینه‌ی متضاد به کار رود که یکی از آنها گزینه‌ی درست سؤال باشد، احتمالاً آزمون‌شونده انتخاب خود را با همین دو گزینه محدود خواهد کرد، زیرا تصور می‌کند که ممکن نیست هر دو پاسخ غلط باشند. در این‌گونه موارد، نمی‌توان سؤال را چندگزینه‌ای دانست، زیرا آن سؤال مانند سؤال‌های صحیح- غلط عمل می‌کند و احتمال حدس زدن جواب صحیح نیز افزایش می‌یابد. بنابراین، بهتر است یا به طور کلی از کاربرد متضادها صرف‌نظر شود یا دو زوج متضاد برای سؤال‌های چهارگزینه‌ای مورد استفاده قرار گیرد.

### مثال

### سؤال ضعیف

وقتی که شباهت بین محرک غیرشرطی و محرک شرطی افزایش می‌یابد:

(x) الف-تعمیم بیشتر اتفاق می‌افتد، اما تمیز دشوارتر می‌شود.

ب- تمیز بیشتر اتفاق می‌افتد، اما تعمیم دشوارتر می‌شود.

پ-تعمیم و تمیز کمتر اتفاق می‌افتد.

ت- اتفاق تازه‌ای رخ نمی‌دهد.

### سؤال بهتر:

وقتی که شباهت بین محرک غیرشرطی و محرک شرطی افزایش می‌یابد:

(x) الف-تعمیم بیشتر اتفاق می‌افتد، اما تمیز دشوارتر می‌شود.

ب- تمیز بیشتر اتفاق می‌افتد، اما تعمیم دشوارتر می‌شود.

پ-تعمیم و تمیز کمتر اتفاق می‌افتند.

ت-تعمیم و تمیز بیشتر اتفاق می‌افتند.

۲۰- طول گزینه‌ی درست را در سؤال‌های مختلف تغییر دهید. گزینه‌ی درست در مقایسه با گزینه‌های انحرافی باید طوری نوشته شود که به هیچ وجه نمایان نباشد. گزینه‌ی درست نباید در چندین سؤال پشت سرهم از سایر گزینه‌ها بلندتر یا کوتاه‌تر باشد، زیرا ممکن است آزمون‌شونده را، صرف‌نظر از اینکه جواب درست سؤال را می‌داند یا نمی‌داند، در پیدا کردن گزینه‌ی درست کمک کند. بنابراین، سعی کنید گاهی گزینه‌ی درست را بلندتر از سایر گزینه‌ها، گاه کوتاه‌تر از آنها، و اغلب هم‌تراز با آنها بنویسید. اما مواظب باشید که در چند سؤال پشت سر هم گزینه‌ی درست کوتاه‌تر یا بلندتر از سایر گزینه‌ها نباشد.

۲۱- محل گزینه‌ی درست را در میان گزینه‌های انحرافی به طور تصادفی انتخاب کنید. اگر در تعیین محل گزینه‌های درست یک آزمون نوعی نظم و ترتیب حاکم باشد، یا اگر گزینه‌های صحیح در جای معینی قرار داده شوند، ممکن است آزمون‌شونده این مطلب را کشف کند و بدون اینکه جواب سؤال را بداند گزینه‌ی درست را انتخاب کند. ساده‌ترین راه برای انتخاب محل گزینه‌های درست استفاده از شماره‌ی صفحات یک کتاب است. کتابی را به طور دلخواه باز کنید و به صفحه‌ای که شماره‌ی آن فرد است نگاه کنید. اگر صفحه‌ای را که باز کرده‌اید شماره‌اش ۱ است، محل پاسخ درست را الف انتخاب کنید و اگر شماره‌اش ۳ است، ب، و همچنین برای شماره‌های ۵، پ، ۷، ت، و ۹، ث را منظور کنید. این روش را برای همه‌ی سؤال‌ها به کار برید.

محل گزینه‌ی درست	رقم پایانی شماره‌ی صفحه‌ی فرد
الف	۱
ب	۳
پ	۵
ت	۷
ث	۹

روش دیگر قرار دادن گزینه‌ها در دنبال تنه‌ی سؤال استفاده از نوعی نظم و ترتیب طبیعی است که این خود سبب می‌شود که گزینه‌ی انحرافی در سؤال‌های مختلف جای تصادفی داشته باشد. مثلاً می‌توان گزینه‌ها را به ترتیب حروف الفبا به دنبال هم قرار داد (اگر از یک کلمه تشکیل یافته‌اند)، یا به ترتیب بزرگی و کوچکی اعداد (اگر از اعداد تشکیل شده‌اند)، یا به ترتیب بلندی یا کوتاهی طول پاسخ‌ها، به مثال زیر توجه کنید:

مثال

مقدار  $X$  در معادله  $X^2 + 1 = 5$  چقدر است؟

الف- ۱

ب- ۲ (x)

پ- ۳

ت- ۴

۲۲- برای هر سؤال بین ۳ تا ۵ گزینه در نظر بگیرید. در سؤال‌های چندگزینه‌ای کاربرد ۳ تا ۵ گزینه تفاوت زیادی ندارد. البته از لحاظ نظری هر چه تعداد گزینه‌ها بیشتر باشد احتمال حدس زدن پاسخ سؤال کمتر است. اما از آنجا که نوشتن تعداد زیادی گزینه‌ی انحرافی دشوار است، گاهی اوقات تعداد گزینه‌های کمتر یعنی ۳ تا ۴ گزینه‌ی مناسب‌تر است. به ویژه خواندن تعداد کمتری گزینه وقت کمتری از آزمون‌شونده را می‌گیرد، و این امر سبب می‌شود که در یک جلسه‌ی امتحان تعداد بیشتری سؤال با گزینه‌های کمتری مورد استفاده قرار گیرد.

مهرنز و لهما (۱۹۸۴) پیشنهاد کرده‌اند که بهتر است برای اندازه‌گیری پیشرفت تحصیلی دانش‌آموزان سطوح پایین از سؤال‌های دارای گزینه‌های کمتر استفاده شود. آنان براساس تجربه‌های خود نظر داده‌اند که دانش‌آموزان کلاس دوم از عهده‌ی پاسخ دادن به سؤال‌های سه گزینه‌ای برمی‌آیند، دانش‌آموزان کلاس سوم و چهارم می‌توانند سؤال‌های ۳ تا ۴ گزینه‌ای را جواب دهند، و سؤال‌های ۴ تا ۵ گزینه‌ای بیشتر برای دانش‌آموزان کلاس ششم و بالاتر مناسب‌اند.

۲۳- در آزمون‌های تکوینی از گزینه‌ی "نمی‌دانم" استفاده کنید. هر چند که درباره‌ی کاربرد این گزینه میان صاحب‌نظران اختلاف عقیده وجود دارد، اما می‌توان از آن در سؤال‌های آزمون‌های تکوینی که برای شناسایی کم و کسر یادگیری و آموزش به کار می‌روند استفاده کرد. چنان‌که قبلاً در این کتاب (فصل ۴) توضیح داده شد، آزمون‌های تکوینی به عنوان کمکی برای آموزش و به منظور تشخیص کم و کسرهای یادگیری دانش‌آموزان و آموزش معلمان مورد استفاده قرار می‌گیرند و هدف از آنها نمره دادن به دانش‌آموزان نیست. در چنین حالتی به نفع یادگیرنده است که اگر نمی‌تواند به سؤالی جواب دهد، به جای حدس زدن کورکورانه، صادقانه عمل کند و گزینه‌ی "نمی‌دانم" را انتخاب نماید. زیرا اگر حدس بزند و حدس او درست باشد، بی‌آنکه جواب سؤال را بداند، در مرحله‌ی بالاتری از آموزش قرار داده خواهد شد که ممکن است برای او مفید نباشد. بنابراین، در آزمون‌هایی که به منظور تشخیص نواقص آموزش و یادگیری به کار می‌روند، نه برای تعیین نمره، گنجاندن گزینه‌ی "نمی‌دانم" هم به نفع دانش‌آموز است و هم به نفع معلم. البته باید به دانش‌آموزان گفته شود که هر گاه جواب سؤالی را که نمی‌دانند به جای اینکه حدس بزنند گزینه‌ی "نمی‌دانم" را انتخاب کنند.

مثال:

در معادله  $\frac{X}{3} - \sqrt{9} = 1$  چقدر است؟

الف-۳

ب-۴

پ-۸ (×)

ت-۲

ث- نمی دانم

۲۴- در سؤال‌هایی که تنه‌ی آنها جمله‌ی ناتمام است، جای خالی را در قسمت آخر جمله قرار دهید. موضوع اصلی هر سؤال باید در تنه‌ی سؤال باشد، به گونه‌ای که آزمون‌شونده بعد از مطالعه‌ی آن و پیش از مراجعه به گزینه‌ها منظور اصلی سؤال را بفهمد. یکی از سفارش‌های صاحب‌نظران در این باره قرار دادن جای خالی در انتهای تنه‌ی سؤال است نه در اوایل یا اواسط آن. به نمونه‌های زیر توجه کنید:

مثال

سؤال ضعیف:

طبق نظر دیباکی..... عامل اصلی بیماری قلبی است.

(×) الف-کلسترول

ب- فریپی

پ-سیگار کشیدن

ت- استرس

سؤال بهتر:

طبق نظر دیباکی عامل اصلی بیماری قلبی ..... است.

(×) الف-کلسترول

ب- فریپی

پ-سیگار کشیدن

ت- استرس

(مهرنز و لهما، ۱۹۸۴)

چنان‌که دیده می‌شود، در سؤال ضعیف بالا جای خالی وسط سؤال مانع از ادامه‌ی مطالعه‌ی آزمون‌شونده می‌شود و موجب سردرگمی او می‌گردد. در سؤال بهتر جای خالی سؤال در آخر تنه‌ی سؤال آمده است و آزمون‌شونده پس از خواندن سؤال و پیش از پرداختن به گزینه‌ها از هدف سؤال به خوبی آگاه می‌شود.

۲۵- تا آنجا که ممکن است در تنه‌ی سؤال به جای جمله‌ی ناتمام از جمله‌ی پرسشی استفاده کنید. متخصصان سنجش و اندازه‌گیری پیشرفت تحصیلی پیشنهاد کرده‌اند که آزمون‌ساز (به‌ویژه آزمون‌ساز تازه‌کار) کوشش کند تا تنه‌ی سؤال‌های خود را در قالب جمله‌های سؤالی یا پرسشی طرح کند و از عبارتهای ناتمام پرهیز نماید. جمله‌ی پرسشی هدف سؤال را به طور آشکارتری بیان می‌کند تا جمله‌ی ناتمام. پین (۲۰۰۳) در این‌باره گفته است "یکی از مشکلات استفاده از جمله‌های ناتمام این است که جمله‌ی ناتمام موجب می‌شود که آزمون‌شونده سؤالی را که در ذهن آزمون‌ساز بوده است بازسازی کند، که خود این ممکن است باعث به هدر رفتن وقت و نیروی آزمون‌شونده گردد" (۱۹۶).

به رغم نگرانی‌های فوق، چنان‌که طرح سؤال در قالب جمله‌ی پرسشی خیلی طولانی‌تری از طرح آن به صورت عبارت ناتمام باشد، بهتر است از همان شکل جمله‌ی ناتمام استفاده کنید. اما سعی نمایید که تمام گزینه‌های پیشنهادی تکمیل‌کننده‌ی جمله ناتمام باشند.

### آزمون چندگزینه‌ای رایانه‌ای

یکی از کاربردهای ابتکاری سؤال‌های چندگزینه‌ای استفاده از رایانه (کامپیوتر) است. برای این منظور مجموعه‌ی متنوعی از سؤال‌های چندگزینه‌ای در یک زمینه‌ی علمی شامل سؤال‌های دارای درجه‌های دشواری از ساده تا دشوار و ضریب‌های تمیز خوب (سؤال‌هایی که گروه‌های قوی و ضعیف آزمون‌شونده را از هم جدا می‌کنند) انتخاب و به دستگاه داده می‌شود. در جریان اجرای آزمون، ابتدا آزمون‌شونده سؤالی از رایانه دریافت می‌کند. اگر توانست به سؤال پاسخ دهد، سؤال نسبتاً دشوارتری دریافت می‌کند، و همچنان این جریان ادامه می‌یابد. اگر آزمون‌شونده به سؤال اول جواب غلط بدهد، سؤال بعدی او یک سؤال ساده‌تر خواهد بود و این جریان نیز تا دادن جواب درست ادامه می‌یابد. رایانه برنامه‌ریزی شده است که سطح توانایی آزمون‌شونده را پیدا کند و تا رسیدن به این سطح به سؤال پرسیدن ادامه می‌دهد. "این‌گونه آزمون‌تاکنون در حوزه‌ی آموزش پزشکی بیشترین مورد استفاده را داشته و معلوم گشته وقتی که با آزمون‌های سنتی به کار می‌رود روش بسیار مفید و معتبری است؟ (ریس، براون، و اسمیت، ۲۰۰۵، ص ۵۴).

## امتیازها و محدودیت‌های آزمون‌های چندگزینه‌ای

آزمون‌های چندگزینه‌ای تمام محاسن آزمون‌های صحیح- غلط و جورکردنی را دارند. علاوه بر آن، بنا به گفته‌ی ایبل (۱۹۷۹)،

این آزمون‌ها قادر به اندازه‌گیری غالب بازده‌های یادگیری از دانش گرفته تا فهمیدن، قضاوت کردن، حل مسئله، ارائه پیشنهاد‌های عملی، و پیش‌بینی امور هستند. تقریباً هرگونه درک و فهم یا توانایی را که بتوان با آزمون‌های دیگر- کوتاه پاسخ، کامل کردنی، صحیح، غلط، جور کردنی، یا تشریحی- اندازه گرفت می‌توان با آزمون‌های چندگزینه‌ای سنجش کرد. (ص ۱۳۵)

به طور کلی امتیازهای آزمون‌های چندگزینه‌ای را می‌توان به شرح زیر بیان کرد:

۱- این آزمون‌ها از سایر آزمون‌های عینی انعطاف‌پذیرترند. آنها علاوه بر دانش، توانایی استدلال، قضاوت و بسیاری بازده‌های مهم دیگر یادگیری را می‌سنجند.

۲- با استفاده از سؤال‌های چندگزینه‌ای می‌توان در یک زمان محدود تعداد زیادی از هدف‌های آموزشی و بخش مهمی از محتوای درس را اندازه‌گیری کرد.

۳- آزمون‌های چندگزینه‌ای نسبت به آزمون‌های صحیح- غلط کمتر امکان حدس زدن کورکورانه را به آزمون‌شونده می‌دهند.

۴- پاسخ‌های این سؤال‌ها به سادگی و یا عینیت کامل قابل تصحیح هستند.

۵- دانش‌آموزان و معلمان این نوع آزمون‌ها را به سایر آزمون‌های عینی ترجیح می‌دهند.

۶- اگر گزینه‌های انحرافی این آزمون‌ها با توجه به اشتباهات و کج‌فهمی‌های متداول دانش‌آموزان تهیه شوند، منبع بسیار مناسبی برای تشخیص مشکلات یادگیری دانش‌آموزان خواهند بود.

آزمون‌های چندگزینه‌ای از چند نظر مورد انتقاد قرار گرفته‌اند. اما در اینجا این انتقادها را به عنوان محدودیت‌های آزمون‌های چندگزینه‌ای برمی‌شماریم.

۱- ساختن این آزمون‌ها بسیار دشوار است. از جمله اینکه، معلمان غالباً نمی‌توانند تعدادی گزینه‌ی انحرافی خوب برای سؤال‌های چندگزینه‌ای انتخاب کنند.

۲- معلمان اغلب سؤال‌های چندگزینه‌ای را در حد سنجش اطلاعات جزئی و کم اهمیت می‌نویسند. (البته این ایرادی است که به همه‌ی آزمون‌های عینی از جمله به آزمون‌های چندگزینه‌ای می‌گیرند.)

۳- در مقایسه با بعضی آزمون‌های دیگر، خواندن این آزمون‌ها و پیدا کردن گزینه‌ی درست مستلزم صرف وقت نسبتاً زیادتری است، به ویژه اگر گزینه‌ها خیلی شبیه بهم ساخته شوند.



۴- زمانی که به پاسخ‌های غلط نمره‌ی منفی داده می‌شود دانش‌آموزانی که خطر می‌کنند از سایر دانش‌آموزان نمره‌های بهتری می‌گیرند. همچنین دانش‌آموزانی که در پاسخ دادن به این‌گونه سؤال‌ها خیره و بانجربه یا به اصطلاح آزمون آشنا<sup>۱</sup> هستند از دانش‌آموزان تازه‌کار نمره‌های بهتری به دست می‌آورند.

۵- دانش‌آموزان قوی بیش از دانش‌آموزان معمولی قادر به پیدا کردن اشکالات، پیچیدگی‌ها، و نظرگاه‌های واگرا (غیرمعمولی) در سؤال‌ها هستند، اما از آنجا که تنها یک پاسخ صحیح در هر سؤال وجود دارد، بابت این تیزبینی‌های خود نه‌تنها تشویق نمی‌شوند بلکه با دادن جوابی که موردنظر طراح سؤال نیست ممکن است تنبیه هم بشوند.

### موارد استفاده‌ی آزمون‌های چندگزینه‌ای

آزمون‌های چندگزینه‌ای بهترین ابزارهای سنجش‌گزینه‌ی پاسخ یا بهترین نوع آزمون‌های بسته پاسخ هستند. سؤال‌های این آزمون‌ها می‌توانند هدف‌های آموزشی مختلفی را اندازه بگیرند و با محتوای غالب موضوع‌های درسی سازگاری دارند (لین و گراندلاند، ۲۰۰۰). از آنجا که سؤال‌های چندگزینه‌ای بسیار انعطاف‌پذیرند و قابلیت کاربرد متنوعی دارند، تاکنون اکثریت آزمون‌های استاندارد شده از آنها استفاده کرده‌اند. سؤال‌های چندگزینه‌ای عموماً برای سنجش هدف‌های یادآوری، فهمیدن، و کار بستن مفیدند. با این حال، اگر در تهیه‌ی آنها دقت و ابتکار به خرج دهید می‌توانید یادگیری‌های پیچیده‌تر را نیز بسنجید.

### سنجش یادگیری به کمک آزمون تشریحی

شناخته شده‌ترین روش سنجش مبتنی بر ساختن پاسخ سنجش تشریحی یا آزمون انشایی است که موضوع بحث فصل حاضر است.

### تعریف و انواع آزمون‌های تشریحی

هر چند که آزمون تشریحی (انشایی)<sup>۲</sup> نزد همه‌ی ما به خوبی شناخته شده است، بازهم چند ویژگی مهم آن را یادآور می‌شویم: (۱) هیچ پاسخ واحدی برای سؤال‌های این نوع سنجش نمی‌توان در نظر گرفت که صددرصد کامل و به طور قطع درست باشد، (۲) پاسخ‌ها از لحاظ درجه‌ی کیفیت یا درستی با هم فرق دارند، و (۳) در این نوع سنجش

---

<sup>۱</sup> -test – wise

<sup>۲</sup> -essay test

(به ویژه آزمون‌های تشریحی گسترده پاسخ) به آزمون‌شونده آزادی عمل زیادی داده می‌شود (استالناکر<sup>۱</sup>، ۱۹۵۱، به نقل مهرنز و لهما، ۱۹۷۳، ۲۰۶).

آزمون‌های تشریحی را، با توجه به آزادی عمل آزمون‌شونده در پاسخ دادن به سؤال‌های آزمون، به دو دسته تقسیم می‌کنند. ویژگی‌های این دو نوع آزمون را در زیر توضیح می‌دهیم.

### آزمون تشریحی گسترده پاسخ

در آزمون‌های تشریحی گسترده پاسخ<sup>۲</sup>، هیچ‌گونه محدودیتی برای آزمون‌شونده منظور نمی‌شود، بلکه او عملاً آزاد است تا هر طور که مایل باشد پاسخ خود را بپروراند و سازمان دهد. آزمون‌شونده در پاسخ دادن به سؤال‌های این نوع آزمون، هم از لحاظ زمان پاسخدهی و هم از جهت مقدار پاسخ، آزادی کامل دارد. در زیر به نمونه‌هایی از سؤال‌های گسترده پاسخ توجه کنید.

- آنچه را که به نظر شما باید در یک برنامه‌ی امتحانی یک آموزشگاه گنجانیده شود بنویسید، نوع آزمون مناسب برای این برنامه را با ذکر دلایل منطقی خود توضیح دهید.

(مهرنز و لهما، ۱۹۸۴، ص ۹۶)

- شرطی شدن کلاسیک پاولفی و شرطی شدن کنشگر اسکینری را با هم مقایسه کنید. در توضیحات خود از کاربرد هر یک از این دو نوع شرطی شدن در زمینه‌های آموزشی و بالینی مثال‌هایی بیاورید.

(کوهن و سوردلیک، ۲۰۰۵، ص ۲۰۵)

- تأثیر قانون مندل را بر تحول زیست‌شناسی به عنوان یک علم شرح دهید.

(لین و گرانلاند، ۲۰۰۰، ص ۲۳۸)

آزمون‌شونده در پاسخ دادن به این سؤال‌ها آزاد است تا مطالبی را ارائه دهد که به نظر او منطقی و مفید هستند و مطالب خود را به گونه‌ای که می‌خواهد سازمان دهد. همچنین آزمون‌شونده باید اطلاعات و دانش‌های خود را ارزیابی کند و اندیشه‌های خود را به نحوی منطقی، منسجم، و بدیع روی کاغذ آورد. بنابراین، سؤال‌های آزمون‌های تشریحی گسترده پاسخ برای سنجش هدف‌های تحلیل و آفرینندگی (بالاترین طبقه‌های طبقه‌بندی هدف‌های آموزشی، حوزه‌ی شناختی) مناسب‌ترین سؤال‌ها هستند.

### آزمون تشریحی محدود پاسخ

---

<sup>۱</sup> -Stalnaker

<sup>۲</sup> -Extended response

در آزمون‌های تشریحی محدود پاسخ<sup>۱</sup>، آزمون‌شونده در دادن پاسخ به سؤال‌ها آزادی کامل ندارد، بلکه صورت سؤال او را ملزم می‌سازد تا پاسخ خود را در چهارچوب شرایطی خاص محدود کند. همچنین این سؤال‌ها برای پاسخ آزمون‌شوند، هم از لحاظ زمان پاسخدهی و هم از نظر مقدار پاسخ، محدودیت‌هایی قائل می‌شوند. به نمونه‌های زیر از این نوع سؤال‌ها توجه کنید:

-پاولف در آزمایش‌های خود مشاهده کرد سگی که شرطی شده بود تا با شنیدن صدای زنگ ترشح بزاق کند بعضی وقت‌ها از این کار سر باز می‌زد. علت این امر چیست؟ جواب خود را حداکثر در نصف صفحه بنویسید.

(اقتباس براساس مهرنز و لهما، ۱۹۸۴، ص ۹۶)

-توضیح دهید چرا فشارسنج (بارومتر) یکی از ابزارهای مفید برای پیش‌بینی هواست، جواب خود را در یک پاراگراف بنویسید.

(لین و گرانلاند، ۲۰۰۰، ص ۲۳۷)

این سؤال‌ها آزمون‌شونده را محدود به جواب دادن درباره‌ی یک موضوع خاص می‌کنند، و از این لحاظ به آنها محدود پاسخ می‌گویند. افزون‌براین، پاسخ آزمون‌شونده به نصف صفحه و یک پاراگراف محدود شده است.

ما در آزمون‌های محدود پاسخ، با محدود کردن پاسخ آزمون‌شوندگان به یک موضوع مشخص، تصحیح پاسخ‌های آنان را هماهنگ‌تر و آسان‌تر می‌کنیم. اما این محدودیت سبب می‌شود که آزمون‌شونده نتواند توانایی خود را در ترکیب کردن اندیشه‌ها و بیان آنها به صورتی منسجم و منطقی ابراز کند. بنابراین، سؤال‌های آزمون محدود پاسخ برای اندازه‌گیری توانایی یادگیرندگان در سطوح فهمیدن، کار بستن، و تحلیل مناسب‌اند، ولی برای سنجش توانایی ترکیب (خلاقیت) و ارزشیابی چندان مفید نیستند.

### قواعد تهیه‌ی سؤال‌های تشریحی

امتیاز مهم آزمون‌های تشریحی این است که آزمون‌شوندگان را وامی‌دارند تا اندیشه‌های خود را به طور منطقی، منسجم، و سازمان یافته بیان کنند و در این کار خلاقیت خود را نشان دهند. نویسنده‌ی سؤال‌های تشریحی باید همواره این نکته‌ی مهم را در نظر داشته باشد. علاوه‌براین، سؤال‌های تشریحی را باید به گونه‌ای نوشت که تا حد امکان از معایب آنها که بعداً توضیح داده خواهند شد کاسته شوند قواعد زیر به منظور تحقق این منظورها پیشنهاد شده‌اند.

۱- دقت کنید که سؤال‌ها به طور مستقیم به هدف‌های آموزشی مربوط شوند. از آنجا که در آزمون‌های تشریحی تعداد سؤال‌ها کمتر از آزمون‌های عینی (مثلاً صحیح- غلط یا چندگزینه‌ای) است، دقت عمل در نوشتن

<sup>۱</sup> Restricted response

این نوع سؤال‌ها بیشتر ضروری است، زیرا بد بودن یک سؤال در یک آزمون تشریحی ۵ سؤالی خیلی بیشتر از بد بودن یک سؤال در یک آزمون عینی ۵۰ سؤالی به نتیجه‌ی ارزشیابی لطمه می‌زند. برای افزودن بر دقت خود، سعی کنید هر سؤال را با توجه به یک هدف دقیق آموزشی بنویسید. بنابراین، پیش از هر اقدام دیگر، جدول مشخصات موضوعی را که می‌خواهید برای آن آزمون تهیه کنید آماده سازید و هدف‌های آموزشی را در آن وارد کنید. سطوح یادگیری هدف‌ها یعنی طبقه‌بندی‌هایی را که این هدف‌ها تعلق دارند نیز مشخص کنید.

۲- استفاده از آزمون‌های تشریحی را تنها به اندازه‌گیری هدف‌هایی محدود کنید که با سایر انواع آزمون‌ها به خوبی قابل اندازه‌گیری نیستند. از سؤال‌های تشریحی برای اندازه‌گیری هدف‌های دانش استفاده نکنید. این‌گونه سؤال‌ها را به هدف‌های طبقات بالای طبقه‌بندی هدف‌های آموزشی اختصاص دهید. سؤال‌های تشریحی گسترده پاسخ بهترین سؤال‌ها برای طبقه‌ی تحلیل، ترکیب (آفرینندگی) و ارزشیابی هستند، اما از سؤال‌های محدود پاسخ می‌توان برای طبقات پایین‌تر مانند کار بستن و فهمیدن استفاده کرد.

۳- صورت سؤال‌های تشریحی را با عبارات و کلمات واضح و روشن بنویسید و از کلی‌گویی و ابهام در بیان بپرهیزید. کلمات و عبارات روشن منظور سؤال را بهتر نشان می‌دهند. کلمات و عبارات "تعریف کنید"، "حل کنید"، "خلاصه کنید" و "طبقه‌بندی کنید" از کلمات "بحث کنید" و "بررسی کنید" روشن‌تر هستند. اگر لازم بود که از "بحث یا بررسی کنید" استفاده شود، باید به طور دقیق بنویسید که چه نکاتی باید مورد بحث و بررسی قرار گیرند. همچنین در کاربرد کلمات "توضیح دهید" و "مقایسه کنید" بهتر است منظور خود را به صراحت بیان نمایید، در غیر این صورت ممکن است آزمون‌شوندگان برداشت‌های متفاوتی از این کلمات داشته باشند و به گونه‌های مختلف به سؤال‌ها پاسخ دهد.

۴- از کاربرد کلمات "چه کسی"، "چه وقت"، "کجا"، و جز اینها بپرهیزید. این کلمات تنها یادآوری معلومات را اندازه می‌گیرند. به جای این کلمات، بهتر است از کلمات "چرا"، "چگونه"، "به چه دلیل" و از این قبیل استفاده کنید. این کلمات در مقایسه با کلمات قبلی هدف‌های سطح بالاتری را می‌سنجند.

۵- تا حد امکان از سؤال‌های تازه و موقعیت‌های جدید استفاده کنید. سؤال‌هایی که در کتاب یا کلاس توضیح داده شده‌اند صرفاً معلومات و حافظه‌ی آزمون‌شونده را می‌سنجند، اما استفاده از موقعیت‌های جدید در طرح سؤال‌ها قدرت کاربرد و درک و فهم را اندازه می‌گیرند. در طرح سؤال برای سنجش کار بستن و خلاقیت، رعایت این نکته ضروری است.

۶- سؤال‌های مربوط به موضوعات و عقاید بحث‌برانگیز را طوری طرح کنید که از آزمون‌شونده بخواهند تا شواهد لازم برای مستند کردن عقیده‌ی انتخابی را بیان کند، نه اینکه از او بخواهید تا صرفاً عقاید

شخصی خودش را شرح دهد. این نوع سؤالها باید به طور روشن از پاسخ‌دهنده بخواهند تا موضوع انتخابی خود را با دلایل علمی و منطقی مستند سازد. پاسخ‌دهنده باید بداند که پاسخ داده شده تنها به این خاطر که او نظر خاصی را می‌پسندد نمی‌تواند مورد پذیرش قرار گیرد.

۷- به آزمون‌شوندگان حق انتخاب چند سؤال از میان تعدادی سؤال را ندهید. از همه‌ی آزمون‌شوندگان بخواهید تا به تمامی سؤالها پاسخ دهند. این امر به شما اجازه می‌دهد تا با کاربرد ملاکی یکنواخت به تصحیح اوراق بپردازید و پاسخ‌ها را بهتر یا یکدیگر مقایسه کنید. اگر آزمون‌شوندگان به سؤالهای مختلف جواب دهند امکان مقایسه‌ی آنها با یکدیگر به طور کامل میسر نخواهد بود.

از آنجا که یکی از معایب عمده‌ی سؤالهای تشریحی عدم نمونه‌گیری درست از آموخته‌های آزمون‌شوندگان است، و از آنجا که هنگام دادن حق انتخاب به آزمون‌شوندگان آنان سؤالهایی را جواب می‌دهند که آمادگی بیشتری برای آنها دارند، دادن حق انتخاب به آزمون‌شوندگان مشکل نمونه‌گیری را حادث می‌کند. مهرنز و لهما (۱۹۸۴) مشکلات دادن حق انتخاب چند سؤال از میان تعدادی سؤال را در سه مورد زیر خلاصه کرده‌اند:

الف) تهیه‌ی سؤالهایی که از نظر درجه‌ی دشواری همسنگ باشند کار آسانی نیست.

ب) آزمون‌شوندگان توانایی انتخاب سؤالهایی را که جواب آنها را بهتر می‌دانند ندارند.

پ) از آنجا که آزمون‌شوندگان قوی‌تر بیشتر به سراغ سؤالهای دشوار می‌روند، ممکن است داشتن حق انتخاب در نهایت به ضرر این گونه افراد تمام شود.

با این حال، در مواردی که هدف سؤالها اندازه‌گیری میزان یادگیری مطالب نیست، بلکه منظور از آزمون صرفاً اندازه‌گیری مهارت پاسخ‌دهندگان در نوشتن و پروراندن مطالب است، دادن حق انتخاب تعدادی سؤال از میان سؤالهای تشریحی ممکن است قابل دفاع باشد.

۸- برای پاسخ دادن به سؤالها، زمان کافی در نظر بگیرید و زمان هر سؤال را نیز به طور جداگانه مشخص کنید. وقتی را که بدین منظور تعیین می‌کنید باید هم شامل دادن جواب باشد و هم فرصتی برای فکر کردن آزمون‌شونده در نظر بگیرید. از آنجا که دادن وقت کافی ممکن است سبب شود که تعداد کمی سؤال در هر آزمون به کار رود، بهتر است برای اندازه‌گیری هدفهای سطح بالای آموزشی در هر دفعه امتحان تعداد سؤالهای کمتری به کار برید، و مشکل نمونه‌گیری را با انجام آزمونهای متعدد و مکرر برطرف کنید.

۹- با نوشتن سؤالهایی که به جواب کوتاه نیاز دارند، تعداد سؤالها را افزایش دهید تا از مشکل ضعف نمونه‌گیری آزمونهای تشریحی کم کنید. هر چه تعداد سؤالهای یک آزمون بیشتر باشد، نمونه بهتری از محتوا

و هدف‌های درس را شامل خواهد شد و نتیجه‌ی بهتری به دست خواهد داد. در ضمن سؤال‌های دارای پاسخ کوتاه روشن‌تر هستند و جواب دادن به آنها از سوی آزمون‌شوندگان مختلف یکنواخت‌تر انجام می‌گیرد.

#### ۱۰- عواملی را که در ارزشیابی آزمون‌های تشریحی دخالت خواهید داد از پیش تعیین کنید و آنها را به

**اطلاع آزمون‌شوندگان برسانید.** ما این موضوع را در بحث پیرامون قواعد تصحیح جواب سؤال‌های تشریحی نیز توضیح خواهیم داد. هدف از بیان آن در اینجا این است که آزمون‌شوندگان پیش از جواب دادن به یک سؤال بدانند که چه چیزی را باید در جواب‌های خود منظور کنند که نظر طراح آزمون را تأمین کند. برای مثال، مشخص کنید که آیا عواملی چون املائی درست کلمات، انشای درست مطالب، نکات دستوری، خوش‌خطی و روشنی بیان در ارزشیابی جواب سؤالات مؤثر خواهند بود یا نه. به یاد داشته باشید که بجز مواردی که توانایی نوشتن، املاء و انشاء را امتحان می‌کنید، از دخالت دادن این عوامل در تصحیح اوراق امتحانات تشریحی بپرهیزید.

#### ۱۱- از روش کارپوشه برای تکمیل آزمون‌های تشریحی استفاده کنید. یکی از اشکالات آزمونهای تشریحی

نمونه‌گیری محدود آنها از محتوا و هدف‌های درس است. برای غلبه بر این مشکل در زمینه‌ی انشاء یا مقاله‌نویسی، از دانش‌آموزان بخواهید تا در طول یک مدت نسبتاً طولانی، مثلاً در طول یک ترم یا سال تحصیلی، تعدادی انشاء بنویسند. چند انشاء مختلف از یک انشاء واحد که در یک جلسه‌ی امتحانی نوشته می‌شود بهتر می‌تواند توانایی دانش‌آموزان را بسنجد. شما می‌توانید این مجموعه انشاء را که در طول ترم تحصیلی جمع‌آوری شده‌اند ارزشیابی کنید.

### قواعد تصحیح پاسخ‌های سؤال‌های تشریحی

برخلاف آزمون‌های عینی که در آنها اندازه‌گیری توانایی آزمون‌شوندگان مستقل از نظر شخصی تصحیح‌کنندگان ورقه‌ها صورت می‌گیرد، در آزمون‌های تشریحی بخش اعظم اندازه‌گیری توانایی مورد سنجش به شخصی که پاسخ‌های سؤال‌ها را می‌خواند و به آنها نمره می‌دهد وابسته است. بنابراین، دقت روش اندازه‌گیری با آزمون‌های تشریحی به مقدار زیاد به نحوه‌ی تصحیح برگه‌های امتحانی و دقت عمل مصححان مربوط است. دلیل عمده‌ی مخالفت برخی صاحب‌نظران با کاربرد این آزمون‌ها نیز از بی‌ثباتی نمره‌گذاری آنها ناشی می‌شود. لذا، رعایت نکات زیر که به منظور بهبود کیفیت نمره‌گذاری این نوع آزمون‌ها پیشنهاد شده از عوامل بسیار مؤثر است.

#### ۱- پاسخ‌های سؤال‌های تشریحی را تنها براساس هدفی که در سؤال گنجانیده شده است تصحیح کنید. از

دخالت دادن نکات دستوری، املائی، انشایی، خوش‌نویسی، و جز اینها خودداری کنید. این نکات را تنها وقتی که امتحان مربوط به دستور زبان فارسی، املاء، و خوش‌نویسی است منظور کنید. زمانی که علاوه بر محتوا و مطالب

درسی، اندازه‌گیری نکات دستوری فوق نیز موردنظر است، برای هر مورد نمره‌ی جداگانه‌ای در نظر بگیرید. این مطالب باید هنگام امتحان به اطلاع آزمون‌شوندگان برسد. در هر صورت معلم باید دقت کند که عوامل فوق ارزشیابی او را از محتوا تحت تأثیر قرار ندهند.

**۲- پاسخ‌ها را سؤال به سؤال تصحیح کنید نه برگه به برگه.** یعنی ابتدا پاسخ سؤال اول همه‌ی برگه‌ها یا ورقه‌ها را تصحیح کنید، بعد پاسخ‌های سؤال دوم همه‌ی برگه‌ها را، و همین‌طور تا آخر. این روش امکان مقایسه‌ی پاسخ‌های آزمون‌شوندگان مختلف به یک سؤال را فراهم می‌آورد، کار نمره‌گذاری به آن سؤال را آسان می‌سازد، و امکان کاربرد یک ملاک واحد را به صورت هماهنگ برای کلیه‌ی افراد میسر می‌کند.

**۳- هنگام تصحیح برگه‌های امتحانی از شناسایی نام صاحبان آنها خودداری کنید.** اگر معلم در هنگام تصحیح برگه‌های امتحانی صاحبان آنها را نشناسد از دخالت نظر شخصی در نمره دادن که یکی از معایب بزرگ آزمون‌های تشریحی است کاسته خواهد شد. برای این منظور از سربرگ‌های ورقه‌ها که مشخصات پاسخ‌دهندگان بر روی آنها نوشته می‌شود و در وقت تصحیح از آنها جدا می‌گردد استفاده کنید. همچنین می‌توان از دانش‌آموزان خواست تا نام خود را در پشت ورقه‌ها بنویسند یا از نوعی رمز عددی استفاده کنند. حتی پیشنهاد می‌شود که پیش از شروع به تصحیح ورقه‌ها آنها را خوب به هم بزنید و مخلوط کنید تا ترتیب تقدم و تأخر آنها به هم بخورد و بر شما معلوم نباشد که برگه‌های رو از آن دانش‌آموزان قوی است یا ضعیف.

**۴- در صورت امکان، از یکی دو نفر از همکارانتان بخواهید تا سؤال‌هایی را که شما تصحیح کرده‌اید تصحیح کنند.** این روش تنها راه بررسی عینیت و دقت ارزشیابی شما از پاسخ‌های دانش‌آموزان است. این کار باید بدون آگاهی طرفین از نمرات یکدیگر و بدون مشورت با هم صورت پذیرد. این روش نمره‌گذاری در شرایطی که قرار است نتایج ارزشیابی برای تصمیم‌های مهم و غیرقابل بازگشت به کار رود، مانند انتخاب افراد برای ادامه‌ی تحصیل یا برای دریافت سایر امتیازهای تحصیلی، بسیار بااهمیت است. در این مورد، میانگین نمرات مصححان مختلف از نمرات هر یک از آنان دقیق‌تر خواهد بود. ایبل (۱۹۷۹) پیشنهاد می‌کند که اگر تصحیح همه‌ی برگه‌ها توسط سایر اشخاص ممکن نیست؛ دست کم نمونه‌ای از آنها را برای تصحیح به کس دیگری واگذار کنید.

**۵- تمام پاسخ‌های آزمون‌شوندگان به یک سؤال را در یک نشست و بدون وقفه‌ی زمانی تصحیح کنید.** انسان‌ها از روزی به روز دیگر و حتی از ساعتی به ساعتی دیگر تغییر می‌کنند. اگر پاسخ‌های تعدادی از آزمون‌شوندگان را امروز و تعدادی دیگر را روز بعد تصحیح کنید، ممکن است عوامل مختلف در نحوه‌ی ارزشیابی شما تأثیر بگذارند. مشکلات خانوادگی، حالات جسمانی، و اتفاقات غیرمنتظره در این کار دخالت می‌کنند. البته این بدان معنی نیست که معلم نباید به منظور تنفس و رفع خستگی برای مدت کوتاهی از کار دست بکشد. منظور این

است که در شرایط مختلف به تصحیح پاسخ‌های آزمون‌شوندگان به یک سؤال واحد نپردازد. وقتی که تصحیح پاسخ یک سؤال را آغاز می‌کنید تمام پاسخ‌های آزمون‌شوندگان به آن سؤال را تصحیح کنید و اگر خواستید می‌توانید بقیه‌ی سؤال‌ها را برای وقت دیگری بگذارید. به هر حال، اگر ناچار شدید تصحیح تعدادی از پاسخ‌های سؤال را که مقداری از آنها را تصحیح کرده‌اید به وقت دیگری واگذارید، در آغاز شروع مجدد به ارزشیابی، تعدادی از برگه‌هایی را که قبلاً تصحیح کرده‌اید بازخوانی کنید تا با نحوه‌ی نمره‌گذاری و شیوه‌ی ارزشیابی قبلی خود آشنا شوید و برگه‌های باقیمانده را نیز به همان سبک ارزشیابی کنید.

۶- به نمرات سؤال‌هایی که قبلاً تصحیح کرده‌اید نگاه نکنید. نمرات سؤال‌هایی را که قبلاً تصحیح کرده‌اید در جایی قرار دهید که هنگام تصحیح سؤال‌های بعدی دور از دید شما باشند. عدم دسترسی به نمرات سؤال‌های قبلی موجب می‌شود که معلم هر سؤال تازه را به طور مستقل و بدون تأثیر گرفتن از نمرات سؤال‌های قبلی تصحیح کند. یک راه انجام این کار آن است که نمرات را بر روی یک صفحه‌ی مجزا بنویسید و پس از جواب دادن به تمام پاسخ‌های یک سؤال برگه‌های آزمون را به طور کامل به هم بزنید. اگر پاسخ سؤال در حد یک صفحه باشد، می‌توان هر سؤال را بر روی یک صفحه نوشت و از پاسخ‌دهندگان خواست که جواب هر سؤال را در همان صفحه بنویسند. این روش برای آزمون‌های ورودی که در آنها طراحان سؤال‌ها و مصححان متفاوت هستند بسیار مفید است.

۷- بر روی برگه‌های آزمون اشتباهات دانش‌آموزان را تصحیح کنید و اظهارنظرهای خود را بنویسید. اظهارنظرها و توضیحات کتبی معلمان بر روی برگه‌های امتحانی دانش‌آموزان منجر به بهبود عملکرد آنان می‌شود (پیچ<sup>۱</sup>، ۱۹۵۸؛ وولفلک، ۱۹۸۷). یکی از هدف‌های آزمون کردن و ارزشیابی پیشرفت تحصیلی دانش‌آموزان فراهم آوردن اطلاعات یا بازخورد درباره‌ی پیشرفت آنان به منظور ایجاد انگیزش و آموزش غیرمستقیم است. این کار همچنین به معلم کمک می‌کند تا الگوی اشتباهات دانش‌آموزان را کشف کند و به نقائص روش آموزش خود و روش یادگیری دانش‌آموزان پی ببرد.

نکته‌ی دیگری که درباره‌ی دادن بازخورد از نتیجه‌ی آزمون به دانش‌آموزان پیشنهاد شده، مرور کردن سؤال‌ها و جواب‌ها با فرد فرد دانش‌آموزان است. چند دقیقه صحبت کردن با هر دانش‌آموز درباره‌ی چگونگی پاسخ‌ها و نقاط قوت و ضعف او بسیار مفید و آموزنده است. این صحبت شفاهی از هر نوع اظهارنظر کتبی مفیدتر است.

۸- با نوشتن یک نمونه یا الگو برای هر سؤال به عنوان کلید، از دخالت عوامل نامربوط جلوگیری کنید. در امتحانات تشریحی محدود پاسخ تهیه‌ی یک کلید پیش ساخته عملی‌تر از انجام این کار در امتحانات گسترده پاسخ است. با این حال، سعی کنید برای سؤال‌های گسترده پاسخ نیز یک الگوی پاسخ درست کنید که

---

<sup>۱</sup> -Page



دست کم در آن نکات مهمی را که می‌خواهید به آن امتیاز بدهید معلوم نمایید. این کلید در یکنواختی نمره گذاری کمک زیادی به مصحح می‌کند. در تهیه‌ی پاسخ‌های نمونه‌ی خود سعی کنید پاسخ هر سؤال را به چند قسمت تقسیم نمایید و برای هر قسمت امتیاز جداگانه‌ای منظور کنید و به پاسخ‌های دانش‌آموزان با توجه به این امتیازها نمره بدهید. دقت کنید که در تقسیم نمره‌ی کل هر پاسخ به اجزای آن کلیه‌ی عوامل مهم را در نظر بگیرید. مثلاً برای سازمان دادن پاسخ، مثال‌های داده شده، روابط بین اجزاء، دلایل ذکر شده، و از این قبیل امتیازهای لازم را منظور کنید.

**۹- از راهنمای نمره‌گذاری استفاده کنید.** برای نمره‌گذاری و تصحیح پاسخ سؤال‌های تشریحی راهنمای نمره‌گذاری<sup>۱</sup> بسیار مفید است. "راهنمای نمره‌گذاری به یک مجموعه قواعد منسجم گفته می‌شود که برای سنجش کیفیت پاسخ یا عملکرد دانش‌آموزان به کار می‌رود. ما معمولاً برای سنجش یک هدف یادگیری یا بخشی از عملکرد از یک راهنما استفاده می‌کنیم، اما عملکردهای پیچیده از چند قسمت یا چند هدف تشکیل می‌شوند؛ لذا برای این منظور چند راهنمای نمره‌گذاری لازم‌اند" (نیتکو، ۲۰۰۱، صص ۲۴۳-۲۴۲). ما در زیر مهم‌ترین راهنماهای نمره‌گذاری را توضیح می‌دهیم.

### روش تحلیلی

در روش تحلیلی<sup>۲</sup> نمره‌گذاری که به آن روش امتیازبندی نیز گفته می‌شود پاسخ نمونه‌ی خود را به اجزای کوچک‌تری تقسیم می‌کند و برای هر جزء مشخص نمره یا امتیاز جداگانه‌ای در نظر بگیرید. همچنین در این روش برای عواملی چون "قدرت بیان"، "ساختمان منطقی پاسخ"، و "ذکر دلایل" امتیازهای جداگانه‌ای منظور کنید.

### روش کلی

در روش کلی یا سراسری<sup>۳</sup> نمره‌گذاری که به آن روش درجه‌بندی نیز گفته می‌شود، پاسخ نمونه به اجزاء و قسمت‌های کوچک تقسیم نمی‌شود بلکه تنها به صورت یک معیار به کار می‌رود. در این روش، معلم تمامی پاسخ هر فرد به یک سؤال را می‌خواند و یک برداشت کلی از آن کسب می‌کند و بعد این برداشت کلی را به یک نمره تبدیل می‌نماید. در این روش به هیچ عامل واحدی امتیاز خاصی داده نمی‌شود، بلکه همه‌ی عوامل موردنظر قرار می‌گیرند و کل پاسخ یکباره مورد قضاوت واقع می‌شود.

---

<sup>۱</sup> -Scoring rubric

<sup>۲</sup> -analytical

<sup>۳</sup> -global

در روش سراسری می‌توان شیوه‌های مختلفی را برای نمره‌گذاری پیشنهاد کرد: می‌توان یک نمره دوگانه مورد استفاده قرار داد (قبول- مردود)، یا مقیاس را تا ۵ واحد به شرح زیر افزایش داد.

۵	خیلی خوب
۴	خوب
۳	متوسط
۲	ضعیف
۱	خیلی ضعیف

ایبل (۱۹۷۹) درباره‌ی نمره‌گذاری با روش کلی گفته است که در مجموع این روش نمره‌گذاری آسان‌تر و سریع‌تر از روش تحلیلی است، اما دلایل توجیهی روشنی برای دادن نمرات به دست نمی‌دهد و برای دانش‌آموزان نیز تصویر روشنی از چگونگی کامل نبودن جوابشان فراهم نمی‌آورد، از این بابت روش تحلیلی نمره‌گذاری بهتر است.

گی (۱۹۹۱) گفته است در حالی که هر دو روش تحلیلی و کلی نمره‌گذاری را می‌توان هم با آزمون‌های وابسته به ملاک به کار بست و هم با آزمون‌های وابسته به هنجار، اما روش تحلیلی برای آزمون‌های وابسته به ملاک مناسب‌تر است، زیرا نمرات بیشتر بر پاسخ‌های واقعی آزمون‌شوندگان مبتنی هستند، و روش کلی بیشتر مناسب آزمون‌های وابسته به هنجار است. او همچنین می‌گوید، زمانی که می‌خواهید روش نمره‌گذاری را با آزمون‌های وابسته به هنجار به کار بندید، ابتدا نمونه‌ای از پاسخ‌ها را مطالعه کنید تا نسبت به کیفیت کل سؤال‌ها یک برداشت حاصل نمایید، بعد هر پاسخ را به عنوان خیلی خوب، متوسط، ضعیف، یا خیلی ضعیف مورد داوری قرار دهید.

ویرسما<sup>۱</sup> و جورس<sup>۲</sup> (۱۹۹۰) بهترین شیوه‌ی نمره‌گذاری را ترکیب دو روش فوق دانسته‌اند. این صاحب‌نظران گفته‌اند ابتدا همه‌ی پاسخ‌ها را با روش کلی تصحیح کنید تا یک احساس کلی نسبت به کیفیت و گستره‌ی آنها کسب نمایید. با این کار می‌توانید پاسخ‌ها را در طبقات نسبتاً کلی دسته‌بندی کنید و سپس پاسخ‌ها را به طور تحلیلی تصحیح نمایید و این بار بر نکات یا اجزای دقیق‌تر تأکید کنید. به باور ویرسما و جورس، این شیوه‌ی نمره‌گذاری ترکیبی سبب جلوگیری از معایب هر دو روش می‌شود و روش نمره‌گذاری معلم را قوت می‌بخشد.

### روش ویژگی‌های اصلی

---

<sup>۱</sup> -Wiersma

<sup>۲</sup> - Jurs

روش دیگر نمره‌گذاری سؤال‌های آزمون‌های تشریحی روش ویژگی‌های اصلی<sup>۱</sup> نام دارد. در این روش، معلم یا مصحح ویژگی‌های اصلی پاسخ دانش‌آموز یا دانشجو به هر سؤال را سنجش می‌کند و برای هر یک از آنها نمره‌های ۴ تا ۰ را که معرف عالی تا غیرقابل قبول‌اند منظور می‌نماید. ساکس (۱۹۹۷)، به عنوان نمونه، گفته است اگر در سؤال تشریحی معلم از دانش‌آموز خواسته است تا نامه‌ای دوستانه بنویسد، هنگام نمره‌گذاری، باید به دنبال ویژگی‌های زیر باشد: بذله‌گویی، توصیف دقیق، و دوستانه بود. مطالب مربوط به راهنمایی استفاده از یک برنامه‌ی کامپیوتری را می‌توان براساس روشنی بیان، سازمان مطالب، و مفید بودن مثال‌ها ارزشیابی کرد. همچنین یک نامه‌ی متقاعدکننده یا مجاب‌گر باید بتواند خواننده را متقاعد کند یا نگرش او را تغییر دهد.

### امتیازها و محدودیت‌های آزمون‌های تشریحی

صاحب‌نظران و متخصصان نظرهای موافق و مخالف زیادی نسبت به کاربرد آزمون‌های تشریحی ابراز داشته‌اند. مدافعان این نوع آزمون‌ها امتیازهای آنها را به شرح زیر برشمرده‌اند:

- ۱- تهیه‌ی آزمون‌های تشریحی از تهیه‌ی آزمون‌های عینی به‌ویژه آزمون‌های چندگزینه‌ای، آسان‌تر است.
- ۲- این نوع آزمون بهترین وسیله‌ی موجود برای سنجش توانایی آزمون‌شونده در پروراندن جواب سؤال‌ها و سازمان دادن مطالب و بیان آنهاست.
- ۳- این آزمون‌ها توانایی تولید پاسخ‌ها را می‌سنجند نه توانایی انتخاب پاسخ‌ها را.
- ۴- آزمون‌های تشریحی از آزمون‌های عینی موقعیت‌های واقعی‌تری را در اختیار آزمون‌شوندگان می‌گذارند.
- ۵- آزمون‌های تشریحی بر روش مطالعه‌ی یادگیرندگان تأثیر مثبت به جای می‌گذارند. اگر یادگیرندگان از پیش بدانند که با استفاده از آزمون‌های تشریحی یادگیری آنان سنجش خواهد شد، با احتمال بیشتری در میان مطالب به جستجوی روابط خواهند پرداخت و اطلاعات را به راه‌های معنی‌دار سازمان خواهند داد (اگن و کاوچاک، ۲۰۰۱، ص ۶۰۷). به عنوان نمونه، در پژوهشی که به منظور بررسی تأثیر روش‌های متفاوت سنجش بر چگونگی مطالعه و یادگیری دانشجویان دانشگاه انجام گرفت (فتح‌آبادی، ۱۳۸۵)، معلوم شد که استفاده از امتحانات تشریحی دانشجویان را به سمت رویکرد عمقی مطالعه سوق می‌دهد.
- ۶- استفاده از آزمون‌های تشریحی عادت‌های مطالعه‌ی بهتری را در دانش‌آموزان و دانشجویان تشویق می‌کند. در مقابل امتیازهای بالا، برای آزمون‌های تشریحی محدودیت‌هایی را نیز ذکر کرده‌اند. مهم‌ترین آنها به قرار زیر بودند:

- ۱- یک آزمون تشریحی نمونه‌ی کوچکی از محتوای درس و هدف‌های آموزشی را اندازه می‌گیرد.

---

<sup>۱</sup> -primary traits

۲- تصحیح پاسخ‌های تشریحی نمی‌تواند با دقت و عینیت انجام گیرد.

۳- تصحیح پاسخ‌های تشریحی بسیار وقت‌گیر است.

معایب ذکر شده‌ی بالا به مقدار زیادی از کارایی آزمون‌های تشریحی می‌کاهند. مشکل نمونه‌گیری محدود برای این نوع آزمون‌ها یک مشکل جدی است. از آنجا که در یک زمان محدود امتحانی نمی‌توان تعداد زیادی سؤال تشریحی به آزمون‌شوندگان داد، لذا همیشه تعداد سؤال‌های یک آزمون تشریحی، به ویژه آزمون‌های تشریحی گسترده پاسخ، بسیار محدود است. در نتیجه آزمون‌های تشریحی معمولاً نمونه‌ی ضعیفی از محتوا و هدف‌های درس را اندازه می‌گیرند. بنابراین، لازم است به هنگام کاربرد این آزمون‌ها، تا آنجا که ممکن است از سؤال‌های محدود پاسخ استفاده کرد که تعداد بیشتری از آنها را در یک جلسه‌ی امتحان می‌توان مورد استفاده قرار داد؛ و به جای یک نوبت چندین نوبت امتحان به عمل آورد، تا نمونه‌ی بزرگتری از محتوا و هدف‌های آموزشی سنجش شوند.

مشکل دیگر آزمون‌های تشریحی ذهنی بودن نمره‌گذاری مصححان و یکنواخت نبودن ارزشیابی آنان است. عوامل بسیار زیادی در این زمینه دخالت می‌کنند که باعث می‌شوند نمره‌گذاری مصححان مختلف به یک برگه‌ی امتحانی یا حتی نمره‌گذاری یک مصحح به یک برگه در زمان‌های متفاوت یکسان نباشد. برای کاستن از این مشکل بهتر است هنگام نوشتن صورت سؤال‌های این آزمون‌ها دقت لازم را به کار بندید تا تمام آزمون‌شوندگان منظور سؤال را درک کنند و در نوشتن پاسخ‌های خود یکنواخت عمل نمایند. همچنین در تصحیح پاسخ‌های سؤال‌های این آزمون‌ها لازم است از یک شیوه‌ی نمره‌گذاری یکنواخت که از قبل تعیین می‌شود استفاده کنید.

سومین عیب آزمون‌های تشریحی، یعنی وقت‌گیر بودن تصحیح پاسخ سؤال‌ها، یک مشکل جدی دیگر است. برخلاف پاسخ‌های آزمون‌های عینی که به وسیله‌ی ماشین یا افراد غیرمتخصص قابل تصحیح است. تصحیح برگه‌های امتحانی آزمون‌های تشریحی نیاز به افراد متخصص دارد. بنابراین، زمانی که هدف‌های آموزشی معلم با آزمون‌های عینی قابل اندازه‌گیری است، بهتر است از این آزمون‌ها استفاده کند و آزمون‌های تشریحی را صرفاً به هدف‌های سطح بالای یادگیری اختصاص دهد.

به‌رغم این مشکلات، آزمون‌های تشریحی از ابزارهای مهم سنجش یادگیری دانش‌آموزان و دانشجویان هستند. چنان‌که در صفحات قبلی این کتاب توضیح دادیم، بسیاری از هدف‌های آموزشی را نمی‌توان با سؤال‌های گزینش پاسخ یا پاسخ‌گزين سنجش کرد و روش‌های سنجش تولید پاسخ با پاسخ‌ساز همواره مورد نیاز معلم‌اند.

تنها به کمک روش‌های سنجش مبتنی بر تولید پاسخ که آزمون‌های تشریحی از انواع خوب آنها هستند می‌توان توانایی سازمان‌دهی، پروراندن مطالب، خلاقیت، و قدرت بیان یادگیرندگان را سنجش کرد. حتی با هدف‌هایی که ممکن است با روش‌های سنجش مبتنی بر گزینش پاسخ قابل اندازه‌گیری باشند، مانند یادآوری، فهمیدن، حل مسئله

و تحلیل، بازم این روش‌ها صرفاً توانایی بازشناسی یادگیرندگان را می‌سنجند نه بازخوانی را با توجه به اینکه بازخوانی جنبه‌ی مهم یادگیری و یادآوری یادگیرندگان است، معلمان در فعالیت‌های سنجش و اندازه‌گیری خود ناگزیر از کاربرد روش‌های سنجش وابسته به تولید پاسخ از جمله روش‌های سنجش تشریحی هستند.

### موارد استفاده‌ی آزمون‌های تشریحی

چنان که گفتیم، از آزمون‌های تشریحی می‌توان برای سنجش انواع مختلف فرآیندها و فرآورده‌های یادگیری دانش‌آموزان و دانشجویان سود برد. جدول ۱-۱۰ یادگیری‌هایی را که می‌توان با این آزمون‌های سنجش کرد نشان می‌دهد. این جدول الگوی کلی سؤال‌های تشریحی را که برای مقاصد گوناگون سنجش قابل استفاده‌اند نشان می‌دهد. در زیر به چند نمونه‌ی واقعی از سؤال‌های تشریحی که برای سنجش هدف‌های مختلف نوشته شده‌اند توجه کنید.

### فهمیدن دانش امور واقعی

۱- توضیح دهید چرا اندکی پس از آنکه ظرفی را بر روی شمع می‌دهیم آن شمع خاموش می‌شود.

(لین و گرانلاند، ۲۰۰۰، ص ۲۴۴)

۲- تأثیر نمره‌های خیلی بالا و خیلی پایین را بر میانه و میانگین با هم مقایسه کنید.

### فهمیدن دانش مفهومی

۳- دلایلی را که موجب شده بود تا آمریکاییان اولیه در دهه‌ی ۱۷۸۰ به غرب سفر کنند توضیح دهید.

۴- در حداکثر ۵ سطر امتیازها و محدودیت‌های آزمون‌های تشریحی را خلاصه کنید.

۵- معنی گفته‌ی زیر را توضیح دهید:

"اگر نمی‌توانید حرارت را تحمل کنید، از آشپزخانه خارج شوید."

(لین و گرانلاند، ۲۰۰۰، ص ۲۴۴)

جدول ۱-۱۰ انواع یادگیری‌هایی که به وسیله‌ی آزمون‌های تشریحی محدود پاسخ و گسترده پاسخ سنجش می‌شوند

(اقتباس براساس لین و گرانلاند، ۲۰۰۰، ص ۲۴۰)

نوع سؤال تشریحی	یادگیری‌هایی که می‌توان سنجش کرد
سؤال تشریحی محدود پاسخ	توانایی: تبیین روابط علت و معلولی توصیف چگونگی کاربرد اصول

<p>ارائه دلایل لازم</p> <p>تدوین فرضیه‌های قابل دفاع</p> <p>تدوین نتیجه‌گیری‌های معتبر</p> <p>بیان فرصت‌های لازم</p> <p>توصیف محدودیت داده‌ها</p> <p>تبیین روش‌ها و شیوه‌ها</p> <p>(و انجام کارهای مشابه براساس توانایی دانش‌آموز)</p>	
<p>توانایی:</p> <p>تولید، سازمان دادن، و بیان اندیشه‌ها</p> <p>ترکیب کردن یادگیری‌های حاصل از زمینه‌های متفاوت</p> <p>آفریدن شکل‌های نو یا اصیل (مانند طراحی یک آزمایش تازه)</p> <p>خلاصه کردن (مانند خلاصه کردن یک داستان)</p> <p>ساختن داستان‌های ابتکاری</p> <p>تبیین مفاهیم با اصول (مثلاً نوشتن یک مقاله‌ی تشریحی)</p> <p>متقاعد کردن خواننده (مثلاً نوشتن یک مقاله‌ی متقاعدکننده)</p> <p>و انجام کارهای مشابه براساس توانایی دانش‌آموز برای نوشتن یک مقاله به منظور هدفی خاص</p>	<p>سؤال تشریحی گسترده پاسخ</p>

### کار بستن دانش روندی

۶- فرض کنید تولدتان را جشن گرفته‌اید و ۹ نفر از دوستان خود را دعوت کرده‌اید. دو پیتزای بزرگ سفارش داده‌اید اما هر یک از آنها به چهار قسمت تقسیم شده است. مشکل شما چیست؟ چه تدبیری می‌اندیشید تا هر نفر یک تکه پیتزا داشته باشد؟

(کوبیسزین<sup>۱</sup> و بوریچ<sup>۲</sup>، ۲۰۰۳، ص ۱۳۱)

<sup>۱</sup> -Kubiszyn

<sup>۲</sup> -Borich

۷- وزن شما در سطح کره‌ی ماه و در سطح خورشید بیشتر یا کمتر از وزن شما در سطح زمین خواهد بود؟ دلایل خود را بنویسید.

(پین، ۲۰۰۳، ص ۲۳۵)

۸- طرحی تهیه کنید که به وسیله‌ی آن بتوانید تعیین نمایید که آیا طرفداران دو حزب معروف سیاسی به طور مساوی در تمام شهر پراکنده‌اند یا طرفداران هر یک از این دو حزب در محله‌ی خاصی متمرکزند.

(اقتباس براساس نیتکو، ۲۰۰۱، ص ۱۸۷)

### تحلیل دانش مفهومی

۹- مقاله‌ی پیوست را بخوانید. بیاناتی را که نشان‌دهنده‌ی عقیده‌ی نویسنده‌اند مشخص کنید. دلایل خود را ذکر نمایید.

(نیتکو، ۲۰۰۱، ص ۱۹۲)

۱۰- رنگ‌های آبی و خاکستری به آب و هوای سرد مربوط‌اند. چه رنگ‌های دیگری به آن ربط دارند؟ این رنگ‌ها بر شکلی که می‌خواهید بکشید چه تأثیری خواهند گذاشت؟

(کوبیسزین و بوریچ، ۲۰۰۳، ص ۱۳۱)

### ارزشیابی دانش مفهومی

۱۱- در تعیین رفتار انسان‌ها طبیعت بیشتر نقش دارد یا تربیت؟ نظر خود را با ذکر دلیل توضیح دهید.

(پین، ۲۰۰۳، ص ۲۳۵)

### آفریدن دانش مفهومی

۱۲- نامه‌ای به مدیر مدرسه‌تان بنویسید و موافقت او را برای بردن دانش‌آموزان به یک گردش علمی جلب کنید.

(لین و گرانلاند، ۲۰۰۰، ص ۲۴۴)

۱۳- با مشاهده‌ی ابرهای موجود در آسمان درباره‌ی وضع طوفانی که در راه است چه پیش‌بینی‌ای می‌کنید؟ (چه فرضیه‌ای می‌سازید؟) چه چیزی در این پیش‌بینی به شما کمک می‌کند؟

(کوبیسزین و بوریچ، ۲۰۰۳، ص ۱۳۱)

۱۴- برای اجرا با بی‌سوادها، روانشناسان آزمون‌هایی ساخته‌اند که به صورت ایما و اشاره قابل استفاده‌اند، و در آنها تصاویر به جای نوشته به کار رفته است. شما می‌توانید برای اجرا با دانش‌آموزانی که قادر به خواندن نیستند روش دیگری ابداع کنید؟ روش خود را به طور کامل و همراه با مثال توضیح دهید.

(مهرنر و لهمان، ۱۹۷۳، ص ۲۰۹)

## پرسش شفاهی

یکی از تدابیر متداول و مؤثر معلمان در ارزشیابی‌های غیررسمی از یادگیری دانش‌آموزان، به منظور دادن بازخورد به آنان، تشویق و ترغیب آنان در یادگیری، و رفع مشکلات یادگیری آنان استفاده از پرسش شفاهی<sup>۱</sup> است. معمولاً پرسش شفاهی را نمی‌توان نوعی آزمون به حساب آورد که از نتایج آن بتوان برای نمره‌گذاری میزان آموخته‌های دانش‌آموزان استفاده کرد. اما از این نوع پرسش برای بهبود روش‌های یادگیری دانش‌آموزان و شیوه‌های آموزشی معلم می‌توان به خوبی استفاده کرد، زیرا بازخوردهای فوری برای معلم و دانش‌آموز فراهم می‌آورد.

می‌توان پرسش شفاهی را نوعی سؤال تشریحی دانست که به صورت غیررسمی طرح و جواب آن هم به صورت شفاهی در کلاس ارائه می‌شود. البته امتحانات رسمی شفاهی را می‌توان به گونه‌ای ترتیب داد که از نتایج آنها برای نمره‌گذاری دانش‌آموزان استفاده شود. اما از آنجا که این نوع امتحانات دارای همان مشکلات امتحانات تشریحی هستند، و علاوه بر آن بسیار وقت‌گیر و غیردقیق‌تر از امتحانات تشریحی کتبی‌اند، در شرایط معمولی کاربرد آنها به عنوان وسیله‌ای برای ارزشیابی و نمره‌گذاری توصیه نمی‌شود. بهتر است پرسش‌های شفاهی را در همان محدوده‌ی کلاس درس و به عنوان وسیله‌ای برای واری نمره‌ی پیشرفت یادگیرندگان و تدبیری برای بازخورد دادن به آنان به کار برد.

مهرنز و لهما (۱۹۸۴) کاربرد پرسش‌های شفاهی را برای مقاصد تشخیصی پیشنهاد می‌کنند. از آنجا که در گفت‌وگوشنود رودر رو امکان هدایت گفتگو به مسیرهای دلخواهی میسر است. معلم می‌تواند از این طریق دانش‌آموزان را به بیان مشکلات خود و یافتن اشکالات یادگیری هدایت نماید. "یک پرسشگر ماهر می‌تواند پاسخ‌هایی را در دانش‌آموزان برانگیزاند که نشان‌دهنده‌ی فرآیندهای فکری او باشند" (ص ۱۲۱).

از امتیازهای دیگر پرسش‌های شفاهی این است که دانش‌آموز می‌تواند از معلم بخواهد تا نکات مبهم سؤال را برای او کاملاً روشن کند. قبلاً گفتیم که یکی از اشکالات بعضی از آزمون‌های تشریحی کتبی مبهم بودن متن سؤال‌هاست. بسیاری از اوقات دانش‌آموزان نه به خاطر بی‌اطلاعی بلکه به سبب ابهام در صورت سؤال‌ها به آنها جواب غلط می‌دهند. در پرسش‌های شفاهی نیز ابهام ممکن است وجود داشته باشد، اما چون پرسشگر و پاسخ‌دهنده رودرروی هم قرار دارند امکان برطرف شدن ابهام موجود است.

یکی دیگر از موارد استفاده‌ی پرسش شفاهی یا امتحان شفاهی کاربرد آن با دانش‌آموزان معلول جسمی است. برخی از دانش‌آموزان به خاطر معلولیت دست یا چشم قادر به دادن جواب‌های کتبی نیستند، لذا در چنین شرایطی بهترین روش آزمون شفاهی است.

---

<sup>۱</sup> -oral question



استفاده از پرسش‌های شفاهی در موقعیت‌های خاص، مانند انجام مصاحبه با داوطلبان شغل معلمی، یا داوطلبان ورود به دوره‌های دکتری، یا علاقمندان به استخدام در هیأت‌های علمی دانشگاه‌ها، می‌تواند در شناخت توانایی افراد، به ویژه نگرش‌ها و واکنش‌های رفتاری آنان، کمک کند. به عنوان نمونه، در جلسه‌ی پرسش و پاسخ شفاهی، اگر داوطلب در پاسخ دادن به سؤال‌های تردید کند، برای یافتن کلمه‌های مناسب مین و مین کند، یا علائم تنش از خود نشان دهد بر سنجش شایستگی یا اعتماد به نفس او تأثیر منفی به جای می‌گذارد. داوطلب شغل معلمی در مصاحبه و گفتگوی شفاهی با هیأت داوران باید بتواند هنگام صحبت کردن از دانش خود استفاده نماید، درست و معقول صحبت کند، و بر اعصاب خود مسلط باشد.

به طور کلی، می‌توان گفت امتحان شفاهی بخش مکمل امتحان کتبی است. زمانی که علاوه بر دانش و مهارت‌هایی که آزمون‌شونده در برگه‌ی کتبی امتحانی خود نشان می‌دهد، سنجش مهارت‌های گفتاری و توانایی‌های ایجاد واکنش مناسب در جمع و نیز کنترل هیجان ضروری است، می‌توان از پرسش شفاهی به عنوان یک وسیله‌ی مؤثر ارزشیابی سود جست.

### آزمون بیرون‌بردنی

نوع دیگر آزمون تشریحی آزمون بیرون‌بردنی<sup>۱</sup> نام دارد. در این امتحان یا آزمون، دانش‌آموز یا دانشجو در ساعت معینی (مثلاً ۸ صبح) سؤال‌ها را دریافت می‌کند و با خود به کتابخانه، منزل، اتاق کامپیوتر، یا هر کجا که بخواهد می‌برد. طبعاً می‌تواند از هر منبعی کمک بگیرد و با هر کس که بخواهد مشورت کند. در ساعت معینی که از قبل اعلام شده، مثلاً ۴ بعدازظهر، جواب سؤال‌ها را تحویل می‌دهد. ریس، براون، و اسمیت (۲۰۰۵) می‌گویند این نوع آزمون از لحاظ روایی بسیار خوب هستند، زیرا بر آنچه در زندگی واقعی اتفاق می‌افتد منطبق‌اند. "البته مشکل اعتماد کردن درباره‌ی آنها وجود دارد- اما این هم با آنچه در زندگی واقعی رخ می‌دهد منطبق است" (ص ۵۴). البته زمان امتحان بیرون‌بردنی را می‌توان به بیشتر از یک روز، مثلاً به تعطیلات آخر هفته، گسترش داد.

### سنجش یادگیری به کمک آزمون کوتاه پاسخ

حد وسط بین آزمون‌های تشریحی و آزمون عینی آزمون کوتاه پاسخ است. شباهت این نوع آزمون با آزمون تشریحی آن است که در هر دو آنها صورت سؤال به وسیله‌ی طراح سؤال تهیه می‌شود و جواب آنها را آزمون‌شونده تهیه می‌کند، یعنی هر دو از نوع تولید پاسخ هستند. با وجود این، جواب سؤال‌های آزمون‌های کوتاه پاسخ آن قدر مختصر

---

<sup>۱</sup> -take- away exam

است که این آزمون‌ها را به کلی از آزمون‌های تشریحی مجزا می‌سازد و به آزمون‌های عینی نزدیک می‌کند. افزون بر این، چون تصحیح پاسخ سؤال‌های این آزمون‌ها با دقت و عینیت بیشتری انجام می‌پذیرد، گروهی از متخصصان اندازه‌گیری و ارزشیابی آنها را در ردیف آزمون‌های عینی قرار داده‌اند.

### تعریف و انواع آزمون‌های کوتاه پاسخ

منظور از آزمون کوتاه پاسخ مجموعه‌ای سؤال مختصر است که غالباً برای سنجش هدف‌های آموزشی سطح پایین طرح می‌شوند. سؤال‌های آزمون کوتاه پاسخ از آزمون‌شونده می‌خواهند تا کلمه، عبارت، جمله، عدد، یا علامتی را در پاسخ به یک سؤال یا تکمیل یک جمله بنویسند.

آزمون‌های کوتاه پاسخ را به سه دسته پرسشی، کامل‌کردنی، و تشخیصی یا تداعی تقسیم می‌کنند. به نمونه سؤال‌هایی برای هر یک از این سه دسته آزمون توجه کنید:

#### نوع پرسشی

- مرکز استان مازندران کجاست؟ (ساری)
- یک میلی‌متر چند میکرون است؟ (۱۰۰۰)

#### نوع کامل‌کردنی

بزرگترین دریاچه دنیا دریاچه ..... است. (خزر)

$$۴+۸ = ..... (۶)$$

#### نوع تشخیصی یا تداعی

بعد از نام هر استان، نام مرکز آن استان را بنویسید.

خوزستان..... (اهواز)

خراسان..... (مشهد)

مرکزی..... (اراک)

لرستان.....(خرم‌آباد)

-بعد از عناصر شیمیایی زیر، علامت شیمیایی آنها را بنویسید.

باریوم..... (Ba)

کلسیم.....(Ca)

پتاسیم.....(K)

روی.....(Zn)

چنان که می بینید، در آزمون های کوتاه پاسخ نوع پرسشی، سؤال به صورت یک جمله ای استفاده می شود و آزمون شونده پاسخ کوتاهی به آن می دهد. در آزمون های کوتاه پاسخ نوع کامل کردنی، سؤال به صورت یک جمله ای ناقص نوشته می شود که در آن یک جای خالی وجود دارد و آزمون شونده با پر کردن جای خالی به سؤال جواب می دهد. و در آزمون های کوتاه پاسخ تشخیصی با تداعی، سؤال به صورت یک موضوع یا مطلب ارائه می شود و آزمون شونده باید جواب آن را تداعی کند.

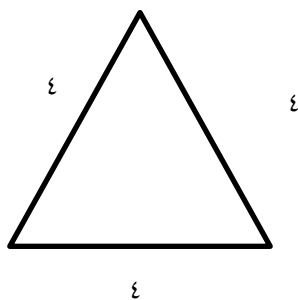
### آزمون های کوتاه پاسخ تشکیل یافته از مسائل عددی ساده

ایبل (۱۹۷۹) آزمون هایی را که از سؤال های مربوط به محاسبات عددی تشکیل می شوند نوع ویژه ای از آزمون های کوتاه پاسخ می داند و به آنها نام آزمون های مسائل عددی<sup>۱</sup> داده است. او می گوید "در حالی که مسائل عددی را می توان به صورت سؤال های چندگزینه ای نوشت، اما اغلب به صورت سؤال های کوتاه پاسخ نوشته می شوند" (ص ۱۷۱). این گونه سؤال ها در ریاضیات، علوم، حسابداری و زمینه های دیگری که با کاربرد اعداد سروکار دارند قابل استفاده اند. امتیاز مهم آزمون های مسائل عددی این است که می توان آنها را در موثعیت های تازه به کار بست و توانایی فهم و کار بستن آزمون شوندگان را مورد سنجش قرار داد. به نمونه هایی از این نوع سؤال ها توجه کنید:

- در تساوی  $4x + 6 = 30$  مقدار  $x$  چقدر است؟

- یک کودک ۷ سال و ۷ ماهه، براساس آزمون هوش استفورد-بینه دارای سن هوشی ۸ است. هوشبهر او را حساب کنید.

- در مثلث زیر اندازه ی ضلع ها معلوم است، اندازه ی هر یک از زاویه های داخلی را برحسب درجه حساب کنید.



### قواعد تهیه ی سؤال های کوتاه پاسخ

۱- هر سؤال را با توجه به یک موضوع مهم بنویسید. از قرار دادن موضوع ها و مطالب بی اهمیت و خیلی جزئی در سؤال ها بپرهیزید. ثرندایک، کانینگهام، ثرندایک، و هیگن (۱۹۹۱) دو سؤال زیر را به عنوان نمونه ی سؤال هایی که در آنها موضوع های بی اهمیت مطرح شده مثال زده اند:

<sup>۱</sup> -numerical problem

-در سال ۱۹۵۰، در کشور آمریکا چند مورد بیماری آبله مشاهده شده است؟

- در سال ۱۹۵۰، تنها یک نفر از ..... نفر مردم آمریکا از دندان‌های خود مراظبت کامل به عمل آورده‌اند. (ص ۲۵۴)

۲- صورت سؤال را کاملاً روشن بنویسید به گونه‌ای که به پاسخ مشخص و معینی نیاز داشته باشد. از مبهم و چندپهلوی نوشتن صورت سؤال بپرهیزید. ایبل (۱۹۷۹) می‌گوید یک مشکل معمولی سؤال‌های کوتاه‌پاسخ این است که سؤالی که به نظر طراح آن پاسخ مشخصی را می‌طلبد از سوی آزمون‌شوندگان پاسخ‌های درست دیگری دریافت می‌کند. برای مثال، سؤال "ذغال‌سنگ چیست؟"

که پاسخ آن از نظر طراح سؤال "نوعی سوخت" است، ممکن است از سوی آزمون‌شوندگان مختلف جواب‌های زیر را دریافت کند که همه درست هستند:

یک ماده‌ی قابل اشتعال

کربن ناخالص

بقایای سفت شده‌ی مواد گیاهی

برای رفع ابهام این سؤال و روشن بیان کردن موضوع موردنظر و نیز برای جلوگیری از دشواری‌های بعدی تصحیح پاسخ‌ها، سؤال را باید طوری نوشت که پاسخ واحد و مشخصی را از همه‌ی آزمون‌شوندگان بخواهد. برای مثال به نمونه‌ی سؤال‌های زیر توجه کنید.

-ذغال‌سنگ غالباً به چه منظوری به کار می‌رود؟

- ذغال‌سنگ از چه موادی درست شده است؟

- ذغال‌سنگ از چه عنصر شیمیایی درست می‌شود؟

به مثال‌های دیگری در زیر توجه کنید:

**سؤال‌های ضعیف:**

برگ تنباکو حاوی چه چیزی است؟ (نیکوتین)

برگ تنباکو حاوی (نیکوتین) ایت.

**سؤال‌های بهتر:**

نام ماده‌ی سمی برگ تنباکو چیست؟ (نیکوتین)

نام ماده‌ی سمی برگ تنباکو (نیکوتین) است. (ثرندایک و همکاران، ۱۹۹۱، ص ۲۵۵)

بنابراین، بکوشید تا صورت سؤال را کاملاً آشکار و خالی از ابهام بنویسید. پوفام (۲۰۰۲) توصیه می‌کند که همواره سعی کنید خودتان را به طور ذهنی، به جای دانش‌آموزانتان قرار دهید و آن‌گاه ببینید که آنها سؤال شما را چگونه تفسیر خواهند کرد.

۳- صورت سؤال را عیناً از روی مطالب کتاب ننویسید. نقل مطالب کتاب موجب تشویق دانش‌آموزان به حفظ کردن مطالب و مانع درک و فهم آنان می‌شود. پیشنهاد نیتکو (۲۰۰۱) برای نوشتن صورت سؤال‌های کوتاه‌پاسخ این است که "ابتدا به پاسخ موردنظرتان فکر کنید و بعد سؤالی طرح نمایید که آن پاسخ تنها پاسخ درست آن باشد" (ص ۱۳۴).

۴- در سؤال‌هایی که پاسخ‌های آنها اعداد هستند، واحد مقیاس و میزان دقتی را که در محاسبات باید رعایت شوند مشخص کنید. سؤال‌ها باید واحدهای عددی و میزان دقت لازم را مشخص کنند. این کار تکلیف آزمون‌شوندگان را مشخص می‌کند و موجب می‌شود تا در وقت کسانی که در محاسبات خود بالاتر از دقت مورد انتظار معلم پیش می‌روند، صرفه‌جویی شود.

**سؤال ضعیف:**

مقدار عدد  $\pi$  چقدر است؟ (۳/۱، ۳/۱۴، ۳/۱۴۱)

**سؤال بهتر:**

مقدار عدد  $\pi$  تا سه رقم اعشار چقدر است؟ (۳/۱۴۱)

برای رعایت قاعده‌ی بالا، راهنمایی‌های لازم را فراهم آورید. به نمونه‌ی زیر که نوعی راهنمایی برای یک آزمون است توجه کنید:

راهنمایی:

این آزمون شامل ۶۰ سؤال است که بعضی از آنها به انجام دادن محاسبات نیاز دارند. برای سؤال‌هایی که احتیاج به محاسبه دارند باید، (۱) جواب‌های خود را تا دو رقم اعشاری نشان دهید، و (۲) آنها را به واحدهای مناسب چون پوند، گرم، یا ولت بنویسید.

(مهرنر و لهما، ۱۹۸۴، ص ۱۳۴)

۵- در سؤال‌های کوتاه‌پاسخ کامل کردنی، تنها کلمه‌ها و عبارت‌های مهم را حذف کنید. چنان‌که دیدیم،

سؤال کامل کردنی نوع خاصی از سؤال کوتاه‌پاسخ است. بنا به گفته‌ی کوهن و سوردلیک (۲۰۰۵)، "نام دیگر سؤال کامل کردنی سؤال کوتاه‌پاسخ است" (ص ۲۰۴). در نوشتن این نوع سؤال‌ها، از آزمون‌شوندگان بخواهید تا به

جنبه‌های مهم آنچه خواننده‌اند پاسخ دهند نه به جنبه‌های بی‌اهمیت و خیلی جزئی. در سؤال‌های کامل‌کردنی، جای خالی باید مطالب مهم را شامل شود، یعنی مطالبی که دانستن آنها از سوی آزمون‌شوندگان مهم است. جای خالی نباید به اجزای فرعی جمله مربوط باشد.

#### سؤال ضعیف:

کبد گلوکز اضافی را به صورت گلیکوژن در خود ..... می‌کند. (ذخیره)

#### سؤال بهتر:

در سوخت و ساز طبیعی بدن، گلوکز اضافی خون به صورت ..... ذخیره می‌شود. (گلیکوژن)

(ثرن‌دایک و همکاران، ۱۹۹۱، ص ۲۵۶)

۶- در سؤال‌های کامل‌کردنی که برای ارزشیابی پیشرفت تحصیلی به کار می‌روند تعداد زیادی جای خالی منظور نکنید. وقتی که در سؤالی بیش از یک جالی خالی قرار می‌گیرد، سؤال مبهم و ناخوانا می‌شود. این‌گونه سؤال‌ها ممکن است برای آزمون‌های هوش و خلاقیت مناسب باشند، اما برای سنجش یادگیری کارساز نیستند. ثرن‌دایک و هیگن (۱۹۷۹) می‌گویند "این نوع سؤال‌ها ممکن است برای اندازه‌گیری آفرینندگی یا اصالت اندیشه‌ی آزمون‌شوندگان در مورد کلمات مفید باشند" (ص ۲۶).

در سؤال‌های کامل‌کردنی گاه به عوض یک جای خالی می‌توان از دو جای خالی استفاده کرد. آن هم در شرایطی که صورت سؤال طوری طرح شود که هیچ‌گونه ابهامی در آن نباشد. به نمونه‌های زیر توجه کنید.

#### سؤال ضعیف:

..... و ..... روش‌های تصحیح آزمون‌های تشریحی هستند. (تحلیلی) و (کلی)

#### سؤال بهتر:

دو روش معمول تصحیح آزمون‌های تشریحی روش ..... و روش ..... هستند. (تحلیلی) و (کلی)

۷- جای خالی سؤال‌های کامل‌کردنی را تا آنجا که ممکن است در قسمت پایانی سؤال قرار دهید. اگر جای خالی در قسمت پایانی سؤال باشد، آزمون‌شونده می‌تواند، پیش از رسیدن به جای خالی، از موضوع مورد سؤال اطلاع کامل حاصل کند. در حالی که اگر جای خالی در اوایل یا اواسط صورت سؤال قرار داشته باشد، تا زمانی که آزمون‌شونده تمام صورت سؤال را می‌خواند ممکن است مطلب مهم و اساسی سؤال را فراموش کند. این مورد به ویژه درباره مطالب پیچیده و نیز در رابطه با کودکان خردسال که دامنه‌ی توجه و حافظه‌ی اندکی دارند صادق است. به نمونه‌های زیر توجه کنید.

### سؤال ضعیف:

در آزمایشات شرطی کردن پاولفی ..... به پاسخی گفته می‌شود که حیوان از آن طریق به صدای زنگ واکنش می‌کند. (پاسخ شرطی)

### سؤال بهتر:

در آزمایش‌های شرطی کردن پاولفی واکنش حیوان به صدای زنگ را ..... می‌گویند. (پاسخ شرطی)

۸- جای خالی سؤال‌های کامل کردنی را همواره یک اندازه تعیین کنید. اگر جای خالی هر سؤال کامل کردنی متناسب با طول پاسخ مورد انتظار باشد، دانش‌آموزان ممکن است به این مطلب پی ببرند و این امر موجب کمک به آنها در حدس زدن جواب درست باشد. بنابراین، بهتر است جای خالی در همه‌ی سؤال‌ها یک اندازه تعیین شود.

۹- در سؤال‌های کامل کردنی، از کاربرد اشاره‌های دستوری و موارد دیگری که جواب سؤال را مشخص می‌کنند خودداری کنید. اشاره‌هایی از قبیل زمان افعال، مفرد و جمع بودن آنها، یا در زبان انگلیسی حرف‌های a و an، و از این قبیل را به کار نبرید. ایبل (۱۹۷۹، ص ۱۶۸) سؤال زیر را به عنوان مثالی که در آن اشاره‌ای در صورت سؤال منجر به حدس زدن پاسخ سؤال می‌شود ذکر کرده است.

-موتور کشتی بخار با نیروی..... کار می‌کند. (بخار)

۱۰- تا حد امکان، به جای سؤال‌های کامل کردنی، از سؤال‌های پرسشی استفاده کنید. به طور کلی، سؤال‌های کوتاه‌پاسخ نوع پرسشی از سؤال‌های نوع کامل کردنی بهترند، زیرا: (۱) تهیه‌ی سؤال‌های نوع پرسشی آسان‌تر است و این سؤال‌ها دارای ابهام کم تری هستند و (۲) پاسخ دادن به سؤال‌های پرسشی برای دانش‌آموزان آسان‌تر است، زیرا آنان غالباً به پرسش‌های معلمان پاسخ می‌دهند نه به جمله‌های ناتمام. به سخن دیگر، آشنایی دانش‌آموزان با جمله‌های پرسشی خیلی بیشتر از جمله‌های ناتمام است. به ویژه وقتی که تغییر دادن جای خالی سؤال‌های کامل کردنی از اوایل صورت سؤال به قسمت آخر آن غیرعملی است، بهتر است صورت سؤال را از نوع کامل کردنی به نوعی پرسشی تغییر دهید. انجام دادن این کار به راحتی صورت می‌گیرد، زیرا سؤال کامل کردنی در واقع از نظر ماهیت همان سؤال پرسشی تغییر شکل یافته است. به نمونه زیر توجه کنید.

### سؤال کامل کردنی:

..... پایتخت کشور انگلستان است. (لندن)

### سؤال پرسشی:

نام پایتخت کشور انگلستان چیست؟ (لندن)

افزون بر آنچه گفته شد، معنی سؤال‌هایی که به صورت پرسشی تهیه می‌شوند روشن‌تر از زمانی است که به صورت کامل‌کردنی هستند. به سؤال‌های زیر توجه کنید:

### سؤال کامل‌کردنی:

نویسنده‌ی داستان آلیس در سرزمین عجایب..... است. (لوئیس کارول)

### سؤال پرسشی:

نام نویسنده‌ی داستان آلیس در سرزمین عجایب چیست؟ (لوئیس کارول)

نیتکو (۲۰۰۱) می‌گوید به سؤال کامل‌کردنی بالا چندین جواب می‌توان داد، از جمله:

"داستان‌سرا"، "ریاضی‌دادن"، "انگلیسی" و مانند اینها، اما جواب سؤال پرسشی فوق تنها همان یک جواب داده شده است.

### قواعد تهیه‌ی سؤال‌های کوتاه‌پاسخ ویژه مسائل عددی

علاوه بر قواعد کلی تهیه‌ی سؤال‌های آزمون‌های کوتاه‌پاسخ که قبلاً معرفی شدند، برای تهیه‌ی سؤال‌های مسائل عددی نیز قواعد زیر پیشنهاد شده‌اند (ایبل، ۱۹۷۹).

۱- تا حد امکان از اعداد ساده استفاده کنید. هدف سؤال‌های مسائل عددی آزمون کردن فهم دانش‌آموز از جریان یا فرایند محاسبات است نه دقت در محاسبات عددی. اگر می‌خواهید دقت محاسبات دانش‌آموز را سنجش کنید سؤال‌های مجزایی به آن اختصاص دهید و در این سؤال‌ها اعداد پیچیده را آن‌طور که می‌خواهید مورد استفاده قرار دهید.

۲- تا حد امکان مسائلی را طرح کنید که دارای جواب صحیح هستند. این کار نیاز به محاسبات پیچیده‌ی اعداد اعشاری و عدم اطمینان آزمون‌شونده را از اینکه محاسبات خود را تا چند عدد اعشاری ادامه دهد از بین می‌برد.

۳- میزان دقت مورد انتظار را تعیین کنید. اگر دانش‌آموزان درباره‌ی آنچه از آنها خواسته شده است نامطمئن باشند و اگر در این باره حدس غلط بزنند، اندازه‌گیری شما از میزان توانایی آنان غیردقیق خواهد بود.

۴- اگر لازم است آزمون‌شوندگان در پاسخ خود واحد اندازه‌گیری را نشان دهند، این مطلب را به اطلاع آنان برسانید. اگر از ذکر این نکته خودداری کنید، ممکن است سبب شود که بعضی از آزمون‌شوندگان به آن توجه نکنند و بی‌جهت از نمره‌ی آنان کاسته شود.



۵- در صورت امکان، مسائل پیچیده و چندمرحله‌ای را به تعدادی مسئله یک مرحله‌ای ساده تقسیم کنید. ایبل (۱۹۷۹، ص ۱۷۱) می‌گوید "این تصور اشتباهی است که هر چه یک مسئله پیچیده‌تر باشد توانایی آزمون‌شونده را بهتر می‌سنجد؛ معمولاً عکس این درست است." هر مسئله‌ی پیچیده‌ای شامل چند مرحله و تعدادی محاسبه است. هر یک از اینها را می‌توان به صورت یک مسئله‌ی مجزا درآورد. حل کردن مسئله به حل کردن این اجزاء وابسته است.

۶- تا آنجا که امکان دارد مسائل عددی را به زبانی ساده و به طور مختصر بیان کنید. این کار مستلزم دادن اطلاعات کافی در صورت مسئله و توضیحات روشن و دقیق است. همچنین از توضیحات اضافی و تکرار کلمات و مطالب خودداری کنید.

### امتیازها و محدودیت‌های آزمون‌های کوتاه پاسخ

یکی از امتیازهای آزمون‌های کوتاه پاسخ سهولت تهیه‌ی سؤال‌ها و تصحیح پاسخ‌های آنهاست. مهرنز و لهمان (۱۹۸۴) می‌گویند زمانی که دانش تعاریف یا اصطلاحات فنی دانش‌آموزان را می‌سنجید بهتر است از آنان بخواهید تا جواب سؤال را خود تهیه کنند نه آنکه جواب‌ها را از میان تعدادی جواب پیشنهاد شده انتخاب نمایند. آزمون‌های کوتاه پاسخ از این امتیاز برخوردارند. علاوه‌براین، آزمون‌های کوتاه پاسخ، در مقایسه با آزمون‌های عینی، نظیر آزمون‌های صحیح-غلط یا چندگزینه‌ای، دارای مزایای دیگری به شرح زیرند:

۱- تهیه و اجرای آنها آسان است.

۲- تقلب را کاهش می‌دهند.

۳- اطلاعات تشخیصی بیشتری را در اختیار معلمان می‌گذارند.

امتیاز دیگر آزمون‌های کوتاه پاسخ این است که در آنها حدس کورکورانه که از معایب عمده‌ی آزمون‌های عینی است وجود ندارد. چنان‌که پیش‌تر گفتیم، دانش‌آموزی که به یک آزمون صحیح-غلط جواب می‌دهد ۵۰ درصد شانس دارد که حدس بزند، و در آزمون‌های چهارگزینه‌ای این احتمال ۲۵ درصد است، اما در آزمون‌های کوتاه پاسخ هیچ‌گونه احتمال حدس زدن وجود ندارد.

در مقابل امتیازهای بالا، آزمون‌های کوتاه پاسخ دارای مشکلاتی هستند که کاربرد آنها را محدود می‌کند. از آنجا که پاسخ سؤال‌های این نوع آزمون‌ها بسیار مختصرند، نمی‌توان از آنها برای سنجش هدف‌های سطح بالای یادگیری استفاده کرد، و کاربرد آنها به اندازه‌گیری هدف‌های سطح دانش و حداکثر فهمیدن و کار بستن (آن هم بیشتر در رابطه با مسائل عددی) محدود می‌شود. مشکل دیگر این است که استفاده‌ی زیاد از این گونه آزمون‌ها سبب تشویق

یادگیرندگان به حفظ کردن اطلاعات جزئی و کم اهمیت درس خواهد شد، زیرا در این نوع آزمون‌ها بیشتر این گونه اطلاعات سنجیده می‌شوند.

همچنین تصحیح و نمره‌گذاری پاسخ‌های این آزمون‌ها به دقت و سرعت آزمون‌های عینی امکان‌پذیر نیست. آزمون‌شوندگان در پاسخ دادن به سؤال‌های آزمون‌های کوتاه‌پاسخ ممکن است یکنواخت عمل نکنند و در جواب سؤال‌ها پاسخ‌های مختلفی بدهند. در نتیجه مصحح باید تصمیم بگیرد که جواب‌ها درست‌اند، غلط‌اند، یا نیمه درست‌اند. این امر ممکن است به عدم دقت در تصحیح برگه‌های امتحانی و کاستن از میزان روایی و پایایی آزمون منجر شود. ساکس (۱۹۹۷) برای روشن کردن مطلب سؤال زیر را مثال زده است.

-نخستین رئیس‌جمهور ایالات متحده آمریکا چه کسی بود؟

ظاهراً معلم انتظار دارد که دانش‌آموز به سؤال بالا جواب جورج واشنگتن را بدهد. اما جواب‌های دیگری مانند یک ژنرال که در ویرجینیا متولد شد جواب‌های بالقوه درستی هستند. بنابراین، "ساختن سؤال‌های کوتاه‌پاسخ یا کامل کردنی که تنها یک پاسخ درست و غیرمبهم داشته باشند دشوار است" (ساکس، ۱۹۹۷، ص ۱۱۸)

### موارد استفاده‌ی آزمون‌های کوتاه‌پاسخ

آزمون‌های کوتاه‌پاسخ برای اندازه‌گیری هدف‌های شناختی سطوح پایین، به ویژه یادآوری دانش، مفیدند. مهرنز و لهما (۱۹۷۳) بهترین مورد استفاده را برای آزمون‌های کوتاه‌پاسخ سنجش دانش تعاریف و اصطلاحات فنی دانسته‌اند. بنا به گفته‌ی این متخصصان، "وقتی که ما علاقه‌مند به سنجش دانش تعاریف یا اصطلاحات فنی دانش‌آموزان هستیم، اگر از آنان بخواهیم تا جواب خود را تولید کنند بیشتر از کارمان مطمئن می‌شویم تا زمانی که از آنان می‌خواهیم تا جواب خود را انتخاب نمایند" (ص ۲۵۶).

البته در برخی مواقع از سؤال‌های کوتاه‌پاسخ برای اندازه‌گیری هدف‌های آموزشی سطح بالاتر نیز می‌توان استفاده کرد. گرانلاند و لین (۱۹۹۰) موارد زیر را ذکر کرده‌اند:

۱- توانایی تفسیر ساده‌ی اطلاعات و کاربرد قواعد، مثلاً شمارش تعداد هجاهای یک کلمه؛ نشان دادن دانش ارزش

مکانی اعداد؛ تشخیص دادن اندام‌های یک موجود زنده از روی عکس؛ کاربرد تعریف یک مثلث متساوی‌الساقین

۲- توانایی حل کردن مسائل عددی در ریاضیات و علوم

۳- توانایی دستکاری نمادهای ریاضی و حل کردن معادله‌های ریاضی و شیمیایی.

ایبل (۱۹۷۹) معتقد است که سؤال‌های آزمون‌های کوتاه‌پاسخ به طور عمده با کلمات و اعداد سروکار دارند. این

سؤال‌ها نام‌های اشخاص، مکان‌ها، اشیاء، فرآیندها، رنگها، و غیره را دربرمی‌گیرند. سؤال‌های کوتاه‌پاسخ همچنین

تعداد معرف تاریخ‌ها، فاصله‌ها، هزینه‌ها، و جمعیت‌ها را شامل می‌شوند. اگر آنها یک عبارت را شامل شوند، معمولاً

عبارتی کوتاه است. "سؤال‌هایی که پاسخ‌های طولانی‌تری را در برمی‌گیرند، مانند "سه دلیل برای ..... ذکر کنید" یا "ویژگی‌های ..... را نام ببرید" بهتر است در طبقه‌ی آزمون‌های تشریحی محدود پاسخ طبقه‌بندی شوند" (ص ۱۶۶). بنابراین، از نظر ایبل همه‌ی آزمون‌های کوتاه‌پاسخ برای اندازه‌گیری اطلاعات واقعی به کار می‌روند. با وجود این، اگر سؤال‌های آزمون‌های کوتاه‌پاسخ به گونه‌ای نوشته شوند که مسائل مختصر ریاضی و تبدیل و تفسیر مطالب را شامل شوند، می‌توان سطح آنها را از دانش به فهمیدن و کار بستن افزایش داد. برای مثال به سؤال کوتاه‌پاسخ زیر توجه کنید.

-چه عددی است که دو برابر آن به اضافه‌ی ۵ می‌شود ۳۹؟ (۱۷)

برای حل کردن این مسئله، دانش‌آموز ابتدا باید عدد موردنظر را مجهول فرض کند. بعد باید مسئله را از صورت کلامی به صورت معادله‌ای مانند  $2X + 5 = 39$  تبدیل نماید (تفسیر کردن). آن‌گاه با کاربرد فرمولی که برای حل کردن این‌گونه مسائل آموخته است به حل مسئله بپردازد (کار بستن). ما در دنباله‌ی مطالب این فصل تعدادی سؤال کوتاه‌پاسخ واقعی را که برای سنجش هدف‌های یادگیری مختلف تهیه شده‌اند می‌آوریم. خواندن دقیق آنها می‌تواند به شما در شناخت بهتر این‌گونه سؤال‌ها کمک کند.

### یادآوری دانش امور واقعی

۱- نمایندگان مجلس شورای اسلامی برای یک دوره‌ی ..... انتخاب می‌شوند.

۲- مرکز استان همدان کجاست؟

### یادآوری دانش مفهومی

۳- در طبقه‌بندی تجدیدنظر شده‌ی حوزه‌ی شناختی بالاترین سطح دانش چه نام دارد؟

۴- در نظریه‌ی ثرندایک، بنا به اصل گسترش اثر، تقویت یک پاسخ هم موجب نیرومندی آن پاسخ و هم نیرومندی پاسخ‌های مجاور آن می‌شود. در نظریه‌ی شرطی‌سازی کلاسیک پاولفی این اصل چه نام دارد؟

### یادآوری دانش روندی

۵- برای تعیین بار الکتریکی مثبت و منفی از چه وسیله‌ای استفاده می‌کنند؟

۶- وقتی که رتبه‌های دو درس دانش‌آموزان یک کلاس را در اختیار داریم، از چه روشی می‌توانیم همبستگی بین آنها را حساب کنیم؟

### فهمیدن دانش مفهومی

۷- اگر دمای گازی ثابت نگه داشته شود در حالی که فشار وارده بر آن افزایش می‌یابد، حجم آن چه تغییری می‌کند؟

۸- در عدد ۵۲۱ رقم ۵ چه ارزشی را نشان می‌دهد؟

### کار بستن دانش روندی

۹- در معادله  $X^2 + X - 1 = 20$ ، مقدار  $X$  چقدر است؟

۱۰- قیمت یک بسته پنیر یک کیلویی ۱۲۰۰۰ تومان است. قیمت بسته‌ی ربع کیلویی همان پنیر ۴۰۰۰ تومان

است. اگر به جای چهار بسته ی ربع کیلویی یک بسته‌ی یک کیلویی پنیر بخرید چند تومان به نفع شماست؟

### سنجش عملکردی

تا کنون آزمون‌های کتبی یا به اصطلاح آزمون‌های مداد و کاغذی را معرفی کردیم. این آزمون‌ها عمدتاً برای اندازه‌گیری هدف‌های آموزشی و بازده‌های یادگیری حوزه‌ی شناختی مناسب‌اند. از آنجا که هدف‌های حوزه‌ی شناختی بخش مهمی از موضوع‌های درسی همه‌ی دوره‌ی تحصیلی را شامل می‌شوند، آزمون‌های کتبی شهرت زیادی کسب کرده‌اند. اما استفاده از این آزمون‌ها برای اندازه‌گیری هدف‌های سایر حوزه‌های یادگیری، یعنی حوزه‌ی روانی- حرکتی و حوزه‌ی عاطفی و حتی برای اندازه‌گیری تمامی هدف‌های حوزه‌ی شناختی، به تنهایی، کافی نیست، و لازم است از تدابیر دیگری نیز سود جست. گرانلاند و لین (۱۹۹۰) تعدادی یادگیری را که با آزمون‌های کتبی با مداد و کاغذی به خوبی قابل سنجش نیستند و برای سنجش آنها روش‌ها و فنون دیگری به غیر از این آزمون‌ها مورد نیازند در جدول ۱-۱۲ خلاصه کرده‌اند.

جدول ۱-۱۲ یادگیری‌هایی که به خوبی با آزمون‌های متداول کتبی قابل سنجش نیستند (گرانلاند و لین،

۱۹۹۰، ص ۳۷۶)

انواع یادگیری	رفتارهای معرف
مهارت‌ها	سخن گفتن، نوشتن، گوش دادن، خواندن شفاهی، انجام آزمایش در آزمایشگاه، رسم کردن، نواختن آلات موسیقی، ژیمناستیک، مهارت‌های شغلی، مهارت‌های مطالعه کردن و مهارت‌های اجتماعی
عادت‌های کاری	توانایی در برنامه‌ریزی، استفاده‌ی درست از زمان، استفاده از وسائل، استفاده از منابع، نشان دادن ویژگی‌هایی چون ابتکار، خلاقیت، پشتکار، و قابل اعتماد بود.
نگرش‌های اجتماعی	علاقمندی به رفاه دیگران، احترام گذاشتن به قانون، احترام گذاشتن به اموال

دیگران، حساس بودن نسبت به مسائل اجتماعی، احترام گذاشتن به نهادهای اجتماعی، میل به کوشیدن برای پیشبرد اجتماعی	
اظهار علاقه نسبت به فعالیت‌های آموزشی، مکانیکی، هنری، علمی، اجتماعی، تفریحی، شغلی	علاقه‌ها
ابراز رضایت و شادمانی نسبت به طبیعت، موسیقی، هنر، ادبیات، مهارت‌های جسمانی، و خدمات‌های برجسته‌ی اجتماعی	ارج‌گذاری یا قدردانی
رابطه با همسالان، واکنش به ستایش و سرزنش، واکنش به مرجع قدرت، ثبات هیجانی، سازگاری اجتماعی	سازگاری

چنان‌که دیده می‌شود، هیچ یک از آزمون‌هایی که قبلاً معرفی شدند به‌تنهایی نمی‌تواند این یادگیری‌ها را به خوبی سنجش کند. بنابراین، لازم است روش‌ها و فنون دیگری را معرفی کنیم که به کمک آنها بهتر بتوان فرآیندها و فرآورده‌های این یادگیری‌ها را سنجش کرد. روش‌هایی که به کمک آنها هم می‌توان یادگیری‌های دانش‌آموزان و دانشجویان را در حوزه‌های روانی- حرکتی و عاطفی سنجش کرد و هم فرآیندها و فرآورده‌های فکری سطح بالای شناختی را که با آزمون‌های سنتی مداد و کاغذی به‌تنهایی و به سهولت قابل اندازه‌گیری نیستند مورد ارزشیابی قرار داد. از آنجا که به کمک این روش‌های سنجش غالباً می‌توان در موقعیت‌های طبیعی و شرایط واقعی به سنجش یادگیری اقدام کرد به آنها نام سنجش واقعی (اصیل)<sup>۱</sup> را داده‌اند. بنا به گفته‌ی پوفام (۲۰۰۲)، "بعضی وقت‌ها ممکن است ببینید که صاحب‌نظران آموزش و سنجش برای توصیف سنجش عملکردی از اصطلاح‌های دیگری استفاده کنند. مثلاً آنها ممکن است اصطلاح سنجش اصیل یا واقعی را به کار برند زیرا این سنجش با تکالیف زندگی واقعی و غیرآموزشگاهی بیشتر سروکار دارد" (ص ۱۷۵). نام دیگر این روش‌ها سنجش جایگزین است.

### تعریف و اهمیت سنجش عملکردی

یک دسته از روش‌های سنجش واقعی یا اصیل سنجش عملکردی<sup>۲</sup> نام دارد. در سنجش عملکردی یا آزمون‌های عملکردی<sup>۳</sup> فرآیندها و فرآورده‌های یادگیری دانش‌آموزان و دانشجویان به طور مستقیم سنجش می‌شوند. کوبیسزین

<sup>۱</sup> -authentic assessment

<sup>۲</sup> -Performance assessment

<sup>۳</sup> -Performance tests

و بوریچ (۲۰۰۳) در این باره گفته‌اند "معلم‌ان می‌توانند از آزمون‌های عملکردی برای سنجش یادگیری‌های شناختی پیچیده، نگرش‌ها، و مهارت‌های اجتماعی در موضوع‌های تحصیلی مختلف، مانند علوم، مطالعات اجتماعی، و ریاضیات استفاده کنند. برای این منظور، معلم‌ان باید موقعیت‌هایی را تدارک ببینند که در آن رفتار یادگیرندگان را به طور مستقیم مشاهده و سنجش نمایند" (ص ۱۵۴).

در سال‌های اخیر بر استفاده‌ی از سنجش عملکردی تأکید زیادی شده است. بنا به گفته‌ی پوفام (۲۰۰۲)، "در ضمن سال‌های اول دهه‌ی ۱۹۹۰ میلادی تعداد قابل توجهی از سیاست‌گذاران آموزش و پرورش شیفته‌ی سنجش عملکردی شدند (ص ۱۷۳). یکی از دلایل عمده‌ی تدکید و توجه به این روش سنجش پیشرفت‌های نسبتاً تازه در روانشناسی شناختی و تأثیر چشم‌گیر آن بر جریان‌های آموزش و پرورش بوده است. بنا به گفته‌ی ثندایک، کانینگهام، ثندایک، و هیگمن (۱۹۹۱)، "علاقه‌ی روزافزون به استفاده از اصول روانشناسی شناختی در کلاس درس نیاز به تشویق و پرورش فرآیندهای عالی فکری و مهارت‌های تفکر انتقادی را سبب شده است. باور عمومی این است که این فرآیندها و مهارت‌های عقلی را با فنون ارزشیابی عملکردی و محصول عملکردی بهتر می‌توان سنجش کرد" (ص ۲۷۱).

همچنین اسپرینتهال<sup>۱</sup>، اسپرینتهال، و آجا<sup>۲</sup> (۱۹۹۴) بیان داشته‌اند که علاقه‌مندی تازه‌ی متخصصان آموزشی و روانشناسان پرورشی به سنجش عملکردی یادگیرندگان امکان سنجش مهارت‌های شناختی سطح بالا و تکالیف مهم مانند توانایی نوشتن، ایجاد ارتباط‌های کلامی، و حل مسئله را فراهم آورده است. دمبو<sup>۳</sup> (۱۹۹۴)، درباره‌ی اینکه چرا در سال‌های اخیر آزمون‌های عملکردی گسترش یافته‌اند و از شهرت آزمون‌های عینی کاسته شده دلایل زیر را ذکر کرده است:

۱- نظریه‌های شناختی یادگیری بر آموزش تأثیر گذاشته‌اند و لذا مهارت‌های فکری پیچیده مورد تأکید قرار گرفته‌اند. در گذشته نظریه‌های رفتاری یادگیری افکار معلم‌ان را نسبت به آموزش تحت تأثیر قرار می‌دادند، و از آنجا که در نظریه‌های رفتاری اعتقاد بر این است که یادگیری در مراحل کوچک انجام می‌شود، لذا نظام ارزشیابی گذشته بر آزمون‌های عینی که دانش‌های اساسی و خُرد را می‌سنجند مبتنی بود.

۲- نظریه‌های شناختی جدید همچنین بر جنبه‌های فکری و خود نظم‌دهی یادگیرنده تأکید می‌کنند. لذا امروزه بیشتر توجه‌ها معطوف به این است که چگونه یادگیرندگان دانش را تفسیر می‌کنند و به کار می‌گیرند تا مسائل پیچیده را حل کنند. این گونه مهارت‌ها با آزمون‌های عینی، مثلاً آزمون‌های چندگزینه‌ای، قابل سنجش نیستند.

---

<sup>۱</sup> -Sprinthall

<sup>۲</sup> - Oja

<sup>۳</sup> - Dembo

۳- پژوهش‌های تازه نشان می‌دهند که یادگیری و انگیزش برهم تأثیر زیادی دارند؛ یادگیرنده ممکن است دانش فراوانی داشته باشد، اما نخواهد از آن استفاده کند. بنابراین، پژوهشگران بر این باورند که در زمینه‌ی نوشتن باید یادگیرندگان را وادارکنند تا نسبت به کار خود بیندیشند و کار خود را ارزشیابی نماید (ارزشیابی شخصی) این کار به آنان کمک خواهد کرد تا معیارهای سطح بالایی برای خود برگزینند.

علت دیگر تأکید بر آزمون‌های عملکردی با سنجش واقعی در سال‌های اخیر اعتراض‌های شدیدی بوده است که نسبت به آزمون‌های سنتی عینی (به ویژه آزمون‌های چندگزینه‌ای) ابراز شده است. بنا به گفته‌ی منتقدین، آزمون‌های سنتی مهارت‌هایی را می‌سنجند که در دنیای واقعی مانند ندارند (وولفلک، ۲۰۰۴).

از دانش‌آموزان خواسته می‌شود تا سؤال‌هایی را جواب دهند یا مسائلی را حل کنند که هرگز دوباره با آنها روبه‌رو نخواهند شد. از آنان می‌خواهند تا این کارها را تنها انجام دهند- بدون اتکاء به هیچ نوع ابزار یا منبعی و در شرایطی کاملاً محدود زمانی. زندگی واقعی اصلاً به اینها شبیه نیست. حل مسائل مهم به وقت کافی نیاز دارد و غالباً نیازمند استفاده از منابع، مشورت با دیگران و ادغام مهارت‌های اساسی با خلاقیت و تفکر سطح بالاست.

در پاسخ به این انتقادات، نهضت سنجش واقعی متولد شد. هدف این اقدام سنجش یادگیری‌های واقعی، مهم، و پیچیده بوده است. به این رویکرد نام‌های دیگر سنجش مستقیم یا سنجش جایگزین نیز داده شده است. این اصطلاحات نشان‌دهنده‌ی شیوه‌هایی هستند که نوع دیگری به غیر از آزمون‌های عینی چهار گزینه‌ای را دربرمی‌گیرند و عملکرد دانش‌آموزان را به طور مستقیم و در تکالیف زندگی واقعی می‌سنجند. (ص ۵۳۸).

ساکس (۱۹۹۷) در رابطه با این آزمون‌ها می‌گوید "آزمون‌های عملکردی از آزمون‌شونده می‌خواهند تا به انجام نوعی رفتار اقدام کند نه اینکه صرفاً سؤال‌هایی را جواب دهد؟ (ص ۱۶۷). لین و گراندلاند (۲۰۰۰) آزمون عملکردی را به عنوان "تکالیفی که از دانش‌آموزان می‌خواهند تا مسائلی را که در خارج از محدوده‌ی کلاسی با اهمیت هستند حل کنند یا به گونه‌هایی عمل نمایند که به خودی خود دارای اهمیت‌اند" (ص ۴۰) تعریف کرده‌اند.

### سنجش فرآیند و فرآورده‌ی یادگیری

آزمون‌های عملکردی با فرآیند<sup>۱</sup> یا شیوه‌ی اجرا، فرآورده<sup>۲</sup> یا محصول کار، و یا ترکیبی از این دو سروکار دارند. ماهیت عملکرد مورد سنجش مشخص می‌کند که تأکید بر چه قسمتی باید باشد. بعضی از عملکردها به محصولات یا فرآورده‌های ملموس و عینی نمی‌انجامند؛ مانند کار کردن با وسایل آزمایشگاهی، سخنرانی کردن، نواختن آلات

<sup>۱</sup> -Process

<sup>۲</sup> -product

موسیقی، و انجام دادن فعالیت‌های ورزشی نظیر شنا کردن یا پرتاب کردن توپ. این گونه فعالیت‌ها را باید ضمن انجام دادن سنجش کنیم و لذا با شیوه‌ی اجرا یا فرآیند عمل سروکار داریم.

در بعضی زمینه‌های عملکردی دیگر، محصول یا فرآورده مورد تأکید است و توجه چندانی به فرآیند یا شیوه‌ی اجرا نمی‌شود. در ارزشیابی از مقاله‌ی تحقیقی، داستان، انشاء، شعر، و نقاشی دانش‌آموزان و دانشجویان بیشتر فرآورده مورد توجه و ارزشیابی قرار می‌گیرد. در این مثال‌ها و مثال‌های مانند آن، با شیوه‌های اجرا و روش‌های مختلفی می‌توان به نتیجه‌ی واحدی دست یافت و لذا آنچه تولید می‌شود بیشتر اهمیت دارد تا فرآیند تولید آن. برای نمونه، معلم معمولاً به شعری که یادگیرنده سروده است بیشتر علاقه‌مند است تا به فرآیند یا جریان سرودن آن شعر.

فرآورده چیزی است که یادگیرنده تولید می‌کند و فرآیند جریانی است که او مورد استفاده قرار می‌دهد تا به فرآورده دست یابد. در بسیاری از موارد، هم فرآیند و هم فرآورده هر دو به عنوان جنبه‌های مهم عملکرد سنجش می‌شوند. برای مثال، مهارت در عیب‌یابی و تعمیر یک تلویزیون یا اتومبیل شامل استفاده از یک شیوه‌ی اجرا و تولید یک تلویزیون یا اتومبیل سالم است. معمولاً در آغاز دوره‌ی یادگیری بیشتر شیوه‌ی اجرا یا فرآیند کار مورد تأکید قرار می‌گیرد، و در مراحل پیشرفته‌تر یادگیری بر فرآورده یا محصول کار تأکید می‌شود. برای مثال، در سنجش مهارت ماشین‌نویسی (تایپ کردن) در آغاز دوره‌ی آموزش، درست لمس کردن دکمه‌ها ارزشیابی می‌شود، اما در مراحل پیشرفته‌تر آموزش، تمیزی، درستی و سرعت کار تولیدی مورد ارزشیابی قرار می‌گیرد. بنابراین، آنجا که هم شیوه‌ی اجرا یعنی فرآیند کار و هم فرآورده یا محصول کار مورد ارزشیابی قرار می‌گیرند میزان تأکیدی که به هر قسمت اختصاص می‌یابد هم به مهارت مورد سنجش و هم به زمان سنجش در طول دوره‌ی آموزش وابسته است.

سنجش عملکردی یا آزمون عملکردی با مهارت<sup>۱</sup> سروکار دارد- مهارت در استفاده از فرآیندها و شیوه‌های اجرایی، و نیز مهارت در تولید فرآورده‌ها (گرانلاند، ۱۹۹۸). برای مثال درس‌های علوم یا مهارت‌های آزمایشگاهی، درس‌های ریاضی با مهارت‌های حل مسئله، درس‌های فارسی و زبان‌های خارجی با مهارت‌های برقرار کردن ارتباط، و درس‌های علوم اجتماعی با مهارت‌هایی چون استفاده از نقشه و نمودار و کار در گروه سروکار دارند. اینها نمونه‌هایی از مهارت در استفاده از فرآیندها و شیوه‌های اجرایی هستند. علاوه بر اینها، مهارت در تولید فرآورده یا محصول در درس‌های هنر و موسیقی، آموزش صنعتی و تجاری، آموزش کشاورزی، تربیت بدنی، و مانند اینها نیز از هدف‌های مهم آموزش و یادگیری به حساب می‌آیند. بنابراین، آزمون‌های عملکردی وسایلی هستند که به عنوان مکمل آزمون‌های کتبی می‌توانند معلمان را در بهتر سنجیدن میزان توفیق یادگیرندگان در رسیدن به هدف‌های متنوع تحصیلی یاری دهند.

---

<sup>۱</sup>-Skill



کوبیسزین و بوریچ (۲۰۰۳) گفته‌اند آزمون‌های عملکردی را هم برای سنجش فرآیندهای یادگیری و هم برای سنجش فرآورده‌های آن می‌توان مورد استفاده قرار داد. آنان در این باره مثال‌های زیر را ذکر کرده‌اند: در یکی از مدارس آمریکا معلمان فرآیند خواندن هر یک از دانش‌آموزان را به وسیله‌ی تعیین درصد کلمه‌هایی که درست تلفظ می‌کند، تعداد جمله‌های معنی‌داری که در ضمن بیان یک داستان به زبان می‌آورد، و درصد عناصر داستانی که بعد از خواندن یک داستان می‌تواند درباره‌ی آنها صحبت کند سنجش می‌کنند. همچنین در مدرسه‌ی دیگری، دانش‌آموزان کلاس چهارم دبستان کارپوشه‌ای از کارهای خود فراهم می‌آورند و برای ارزشیابی آن را در اختیار معلم می‌گذارند. این کارپوشه‌ها دربرگیرنده‌ی تولیدات آنان درباره‌ی شعر، مقاله، زندگی‌نامه، و تفکرات دانش‌آموزان است. بازهم در مدرسه‌ای دیگر فرآیندها و فرآورده‌های یادگیری در درس مطالعات اجتماعی این‌گونه سنجش می‌شود که از دانش‌آموزان خواسته می‌شود تا درباره‌ی سوال زیر پروژه‌های گوناگونی را انجام دهند:

"براساس مطالعاتی که درباره‌ی تاریخ ایالت خود انجام داده‌اید، چه مسائلی در رابطه با این ایالت مطرح‌اند که نیاز به بحث و بررسی دارند، نظر شما درباره‌ی راه‌حل این مسائل چیست، و شما برای رسیدن به راه‌حل این مشکلات چه کمک‌هایی می‌توانید بکنید؟" دانش‌آموزان به صورت فردی و گروهی از طریق نوشتن گزارش‌های مختلف، ارائه شفاهی و تدارک نمایشگاه‌ها، به سؤال موردنظر جواب می‌دهند.

به طور کلی، هدف‌های آموزش و یادگیری را از لحاظ سنجش می‌توان به دو بخش زیر تقسیم کرد:

(۱) هدف‌هایی که به دانستن درباره‌ی امور مربوط می‌شوند (مثلاً اینکه میکروسکوپ برای مطالعه‌ی جانوران کوچکی که با چشم غیرمسلح قابل مشاهده نیستند به کار می‌رود)؛ (۲) هدف‌هایی که به دانستن نحوه‌ی انجام دادن عمل یا کاری مربوط می‌شوند (مثلاً کار کردن با میکروسکوپ). "آن دسته از وسایل اندازه‌گیری که می‌کوشند تا بسنجند که چگونه کسی می‌تواند کاری انجام دهد غالباً آزمون‌های عملکردی نام دارند، و آن دسته از وسایل اندازه‌گیری که می‌کوشند تا دامنه و صحت دانش شخص را بسنجند معمولاً آزمون‌های کتبی یا مداد و کاغذی نامیده می‌شوند" (بایلر<sup>۱</sup> و اسنومن<sup>۲</sup>، ۱۹۹۳، ص ۵۶۵).

## انواع سنجش‌های عملکردی

---

<sup>۱</sup> -Bieler

<sup>۲</sup> -Snowman

چنان‌که قبلاً اشاره کردیم، به سنجش‌های عملکردی یا آزمون‌های عملکردی آزمون‌های واقعی یا اصیل نیز گفته می‌شود. صاحب‌نظران، براساس پژوهش‌هایی که درباره‌ی سنجش عملکردی انجام داده‌اند، ویژگی‌های آنها را در چهار دسته‌ی زیر خلاصه کرده‌اند (نیتکو، ۲۰۰۱).

- ۱- تأکید بر کار بست، یعنی سنجش اینکه آیا دانش‌آموز، علاوه بر دانستن دانش، می‌تواند آن را به کار بندد.
- ۲- تأکید بر سنجش مستقیم، یعنی سنجش هدف هدف آموزشی به صورت مستقیم، به عوض غیرمستقیم.
- ۳- استفاده از مسائل واقعی، یعنی استفاده از مسائل و موقعیت‌هایی که در زندگی واقعی یافت می‌شوند یا شبیه به آنها هستند.
- ۴- ترغیب و تشویق تفکر باز، یعنی هدایت دانش‌آموزان به پیدا کردن راه‌حل‌های مختلف برای مسائل و تشویق آنان به درگیر شدن با مسائل به طور گروهی و کار کردن بر روی آنها برای مدتی طولانی (روزها، هفته‌ها، ماه‌ها). آزمون‌های عملکردی یا سنجش‌های عملکردی بسیار متنوع‌اند. ما در اینجا انواع مهم آنها را توضیح می‌دهیم.

### دسته‌بندی گرانلاند از آزمون‌های عملکردی

گرانلاند (۱۹۸۸) انواع روش‌های سنجش یا آزمون‌های عملکردی را به چهار دسته‌ی زیر تقسیم کرده است:

- ۱- آزمون‌های کتبی عملکردی
  - ۲- آزمون شناسایی
  - ۳- انجام عملکرد در موقعیت‌های شبیه‌سازی شده
  - ۴- نمونه کار
- در این تقسیم‌بندی، آزمون نوع اول، یعنی آزمون کتبی عملکردی، بیشترین فاصله را از عملکرد در زندگی واقعی دارد، ولی آخرین نوع آزمون، یعنی نمونه‌ی کار، کمترین فاصله را از عملکرد در زندگی واقعی دارد. روش‌های چهارگانه‌ی بالا را هم می‌توان جداگانه به کار بست و هم با یکدیگر مورد استفاده قرار داد. در بعضی موارد تنها یک روش کفایت می‌کند، اما در بیشتر مواقع کاربرد ترکیب دو یا چند روش ضروری به نظر می‌رسد، ما در زیر هر یک از این روش‌ها را توضیح می‌دهیم.

### آزمون کتبی عملکردی

هر چند در آغاز این فصل گفتیم که آزمون‌های کتبی یا مداد و کاغذی برای اندازه‌گیری بازده‌های یادگیری هدف‌های حوزه‌ی شناختی مناسب‌ترند، با این حال، در سنجش عملکرد، استفاده از آزمون‌های کتبی نیز مفید است.

اما باید دقت کرد که بین آزمون‌های کتبی مورد استفاده برای سنجش هدف‌های حوزه‌ی شناختی و آزمون‌های کتبی مورد استفاده برای سنجش عملکردی تفاوت وجود دارد. تفاوت عمده بین آزمون کتبی عملکردی<sup>۱</sup> و آزمون‌های کتبی که در فصل‌های قبلی این کتاب معرفی شدند این است که در آزمون کتبی عملکردی عمدتاً بر کاربست دانش و مهارت در موقعیت‌های عملی یا شبیه‌سازی شده یا موقعیت‌های عملی تأکید می‌شود. در این گونه آزمون‌های عملکردی، یا بازده‌های پایانی یادگیری سنجش می‌شوند یا مراحل میانی عملکردی که برای رسیدن به بازده‌های مطلوب پایانی ضروری هستند، مانند استفاده‌ی درست از ابزارها و دستگاهها.

به عنوان نمونه‌هایی از سؤال‌های آزمون کتبی عملکردی که در هر یک از آنها یک بازده مهم یادگیری سنجش می‌شود به موارد زیر توجه کنید.

دانشجوی درس اندازه‌گیری و ارزشیابی آموزشی در توانایی‌های زیر سنجش می‌شود:

- ساختن یک جدول مشخصات درس برای یک واحد آموزشی

- ساختن تعدادی سؤال آزمون متناسب با جدول مشخصات

ساختن یک فهرست واری برای ارزشیابی از یک آزمون پیشرفت تحصیلی

(گرانلند، ۱۹۸۸، ص ۸۷)

همچنین از دانش‌آموزان و دانشجویان می‌تواند خواست تا یک نقشه‌ی آب و هوایی، یک نمودار ستونی، طرح یک مدار برقی، طرح یک لباس، یک داستان کوتاه، یا نقشه‌ی یک آزمایش علمی را بسازند. در این مثال‌ها، محصول یادگیری که به صورت کتبی ارائه می‌شود، همه نتیجه‌ی دانش فرد است و هم نتیجه‌ی مهارت اول، و یک مقیاس عملکردی به دست می‌دهد که به خودی خود ارزشمند است و، برخلاف آزمون‌های غیرعملکردی، به عنوان وسیله‌ای برای قضاوت درباره‌ی یک صفت یا توانایی روانی غیرقابل مشاهده به کار نمی‌رود.

علاوه بر موارد بالا که در آنها محصول پایانی یا نهایی یادگیری سنجش می‌شوند، در بعضی مواقع می‌توان از آزمون‌های کتبی عملکردی برای سنجش مهارت‌ها یا فعالیت‌های واسطه‌ای نیز سود جست. برای مثال، پیش از استفاده از یک دستگاه یا یک ابزار خاص، مانند ریزسنج<sup>۲</sup> ممکن است ضروری باشد که یادگیرندگان، از روی عکس‌های مقیاس، موقعیت‌های مختلف را مطالعه کنند. اگرچه توانایی خواندن مقیاس شرط کافی برای سنجش دقیق نیست، اما یک شرط لازم است. در اینجا آزمون کتبی برای آزمون یادگیرندگان به صورت گروهی روش مناسبی است. علاوه بر این، آنجا که عملکرد پیچیده و دستگاه خیلی گران‌قیمت است، نشان دادن شایستگی انجام

---

<sup>۱</sup> -paper – and- pencil performance test

<sup>۲</sup> -micrometer

عمل در موقعیت‌های آزمون کتبی از حوادثی که ممکن است دستگاه را ناقص کند جلوگیری به عمل می‌آورد. همچنین از آزمون‌های کتبی در علوم بهداشت و پزشکی نیز می‌توان سود جست. مثلاً مهارت تشخیص و تجویز را در رابطه با بیماران فرضی می‌توان در دانشجویان پزشکی سنجش کرد تا اطمینان حاصل شود که بعدها در برخورد با بیماران واقعی با اشکال روبه‌رو نمی‌شوند. کوتاه سخن اینکه، آزمون عملکردی کتبی را می‌توان به عنوان وسیله‌ی سنجش اطلاعات مقدماتی فرد درباره‌ی کاری که انجام خواهد داد به کار بست. مثلاً اطلاعات فرد از ساختار ابزاری که مورد استفاده قرار خواهد داد، پیش از اینکه عملاً شروع به کار کردن با دستگاه بکند، اندازه‌گیری می‌شود. این نوع اطلاعات به صورت کتبی و در صورت لزوم به صورت گروهی قابل سنجش است.

## آزمون شناسایی

منظور از آزمون شناسایی<sup>۱</sup> روشی است برای سنجش توانایی یادگیرنده در تشخیص ویژگی‌ها، محاسن، معایب و موارد استفاده امور مختلف.

آزمون‌های شناسایی انواع و کاربردهای مختلفی دارند. در بعضی وقتها می‌توان از یادگیرنده خواست تا یک ابزار یا دستگاه را شناسایی نماید یا کار آن را توصیف کند. در موقعیت‌های دیگر می‌توان یادگیرنده را با مشکلی مواجه ساخت، مانند یک اتصال در یک دستگاه برقی، و از او خواست تا ابزارها، وسایل، و شیوه‌های عملی رفع مشکل را شناسایی کند. مورد پیچیده‌تری از این نوع آزمون گوش دادن به یک دستگاه معیوب، مثلاً موتور یک اتومبیل، و بعد از روی صدا شناسایی کردن علت یا علت‌های اشکال دستگاه است.

آزمون‌های شناسایی هم در آموزش صنعتی کاربرد دارند، هم در علوم زیستی، و هم در سایر رشته‌های علمی، برای نمونه، از دانش‌آموزان و دانشجویان درس زیست‌شناسی می‌توان خواست تا نمونه‌های جانوران مختلفی را که در اتاق آزمایش یا کلاس درس چیده شده‌اند شناسایی کنند، یا از دانش‌آموزان و دانشجویان درس شیمی می‌توان خواست از وسایل و مراحل انجام یک آزمایش شیمیایی را مشخص کنند. روش درست حل یک مسئله در درس ریاضی نیز مشمول آزمون‌های شناسایی هستند. علاوه بر اینها، شناسایی بهترین شیوه‌های عملی در درس‌های هنر، موسیقی، تربیت بدنی، و فعالیت‌های حرفه‌ای مانند کشاورزی، نجاری، مکانیکی، و غیره ضروری‌اند.

از آزمون‌های شناسایی به عنوان وسایل سنجش غیرمستقیم عملکرد نیز می‌توان استفاده کرد. برای مثال، انتظار می‌رود که یک لوله‌کش ماهر از یک لوله‌کش تازه‌کار اطلاعات بیشتری درباره‌ی ابزارها و وسایل لوله‌کشی داشته باشد. بنابراین، می‌توان از یک آزمون شناسایی برای سرنده کردن گروه بزرگی از داوطلبان این شغل استفاده کرد، و

---

<sup>۱</sup> -identification test

در مرحله‌ی اول انتخاب افراد ناوارد به کار لوله‌کشی را کنار گذاشت تا اینکه در مرحله‌ی بعد انتخاب ماهرترین افراد به وسیله‌ی آزمون‌های مستقیم سنجش عملکرد سهل‌تر گردد. گرانلاند (۱۹۸۸) در رابطه با استفاده از آزمون‌های شناسایی گفته است، "معمول‌تر آن است که از آزمون‌های شناسایی به عنوان یک تدبیر آموزشی برای آماده کردن دانش‌آموزان و دانشجویان در اجرای عمل در موقعیت‌های واقعی یا شبیه‌سازی شده استفاده شود" (ص ۸۸).

آزمون بالینی ساختارمند عینی (آسکی)<sup>۱</sup> یکی از موارد استفاده‌ی آزمون شناسایی که در آموزش پزشکی کاربرد زیادی دارد آزمون بالینی ساختارمند عینی است که به طور خلاصه آسکی خوانده می‌شود. از این نوع آزمون هم به منظور سنجش و هم با هدف آموزش استفاده می‌شود. در اتاق مخصوص سنجش ایستگاه‌های مختلفی وجود دارند که آزمون‌شونده باید به همه‌ی آنها سر بزند. در یکی از آن ایستگاه‌ها مثلاً عکس‌های بالینی وجود دارند که باید تفسیر شوند، در ایستگاه بعدی مدارک دیگر بیماری موجودند که باید براساس آنها بیماری تشخیص داده شود، در ایستگاه دیگر ممکن است یک بیمار واقعی (یا کسی که نقش یک بیمار را بازی می‌کند) حاضر باشد که باید معاینه و مصاحبه شود. بنا به گفته‌ی ریس، براون، و اسمیت (۲۰۰۵) این آزمون‌ها دارای روایی سطح بالایی هستند، زیرا چیزی را می‌سنجند که پزشکان باید قادر به انجام آنها باشند. "معمولاً نمره‌گذاری این آزمون‌ها آسان‌تر از امتحان‌های کتبی است (نه تنها به این دلیل که دست‌خط پزشکان از بدی زبان زد همگان است). اما برنامه‌ریزی این‌گونه آزمون‌ها وقت زیادی می‌برد" (ص ۵۳).

### انجام عملکرد در موقعیت‌های شبیه‌سازی شده

در آزمون شبیه‌سازی یا انجام عملکرد در موقعیت‌های شبیه‌سازی شده از یادگیرنده خواسته می‌شود تا در یک موقعیت شبیه‌سازی شده یا مصنوعی یا خیالی همان اعمالی را که انجام دهد که در موقعیت‌های واقعی ضروری هستند. برای نمونه، در تربیت بدنی، ضربه زدن به یک توپ خیالی، مشت‌بازی کردن به طور خیالی، یا انجام حرکات مختلف شنا در بیرون از آب نمونه‌هایی از عملکرد در موقعیت‌های شبیه‌سازی شده هستند. در درس‌های علوم و حرفه‌وفن کار در آزمایشگاه یا کارگاه به صورت عملکرد در شغل واقعی شبیه‌سازی شده است؛ اما در بعضی موارد، دستگاه‌ها و وسایل بخصوصی برای آموزش و آزمون از راه انجام عملکرد شبیه‌سازی شده فراهم می‌آیند. مثلاً در آموزش رانندگی، پیش از آنکه یادگیرنده را عهده‌دار راندن یک اتومبیل واقعی بکنند، او را در یک اتومبیل که مانند اتومبیل‌های واقعی شبیه‌سازی شده است و همه‌ی دستگاه‌های رانندگی از جمله فرمان، آینه، پدال‌ها و غیره را دارد قرار می‌دهند، و از او می‌خواهند تا اعمال رانندگی را در این خودرو مصنوعی تمرین کند. به همین منوال، برای

---

<sup>۱</sup> -Objective structured clinical exam (OSCE)

آموزش خلبانی نیز از دستگاه‌هایی که با هواپیماهای واقعی شبیه‌سازی شده‌اند استفاده می‌کنند. این‌گونه امکانات شبیه‌سازی شده از صدمه رسیدن به یادگیرنده و دیگران در مراحل اولیه‌ی یادگیری جلوگیری به عمل می‌آورد. در نمونه‌های بالا، انجام عملکرد در موقعیت‌های شبیه‌سازی شده بیشتر برای آموزش مهارت‌های مختلف مانند رانندگی اتومبیل مورد استفاده قرار می‌گیرد. با این حال، از این روش برای ارزشیابی از توانایی‌ها و مهارت‌های یادگیرندگان در انجام کارهای مختلف (فرآیند یادگیری) و تولید بازده‌های مطلوب (فرآورده‌های یادگیری) نیز می‌توان سود جست. برای مثال، آنجا که عملکرد دانش‌آموز یا دانشجو در انجام آزمایش‌های شیمی در آزمایشگاه سنجش می‌شود، این روش برای ارزشیابی از توانایی آموخته شده‌ی یادگیرنده مورد استفاده قرار می‌گیرد.

بعضی از صاحب‌نظران سنجش یادگیری (مانند هینتزه، استونر، و بول، ۲۰۰۰) به جای انجام عملکرد در موقعیت‌های شبیه‌سازی شده اصطلاح سنجش قیاس‌پذیر را به کار برده‌اند. منظور آنان از سنجش قیاس‌پذیر مشاهده و ارزشیابی عملکرد یادگیرنده در موقعیت‌هایی است که شبیه به موقعیت‌های واقعی است. هدف کسانی که سنجش قیاس‌پذیر را مورد استفاده قرار می‌دهند پیش‌بینی وقوع رفتارهای مورد سنجش در محیط زندگی واقعی است.

برای آشنایی بیشتر با آزمون‌های عملکردی از نوع انجام عملکرد در موقعیت‌های شبیه‌سازی شده، یک نمونه‌ی واقعی را توضیح می‌دهیم. این نمونه را از کتاب سنجش کلاسی تألیف جیمز پوفام (۲۰۰۲) نقل می‌کنیم. از این آزمون عملکردی به منظور سنجش توانایی ایجاد ارتباط کلامی در دانش‌آموزان و دانشجویان استفاده می‌شود.

### موقعیت سنجش: یک حالت اضطراری (اوژانس)

فرض کنید که تنها در منزل هستید و بوی دود می‌شنوید، شما [آزمون‌شونده یا دانش‌آموز] به اداره‌ی آتش‌نشانی تلفن می‌زنید و من [سنجشگر یا معلم] به تلفن شما جواب می‌دهم. با من صحبت کنید، آن‌گونه که در تلفن صحبت می‌کنید، هر چیزی را که لازم می‌دانید، برای من توضیح بدهید تا به کمک شما بیایم. (مستقیماً با من صحبت کنید، با گفتن آلو شروع کنید).

(پوفام، ۲۰۰۲، ص ۱۸۵)

نمونه‌ی کار

در روش نمونه‌ی کار<sup>۱</sup> از یادگیرنده خواسته می‌شود تا اعمالی را انجام دهد که معرف عملکرد واقعی مورد سنجش هستند. بدین لحاظ، این روش نزدیک‌ترین روش سنجش به عملکرد واقعی یادگیرنده در محیط‌های طبیعی است. در روش نمونه‌ی کار، نمونه‌ی اعمالی که از یادگیرنده می‌خواهیم تا انجام دهد باید شامل عناصر مهم عملکرد کلی باشد که در شرایط کنترل شده اجرا می‌شود. برای مثال، در آزمون رانندگی اتومبیل، از یادگیرنده خواسته می‌شود تا طول فاصله‌ای را که دربرگیرنده‌ی موقعیت‌های رانندگی معمولی هستند و فرد در رانندگی واقعی با آنها روبه‌رو می‌شود اتومبیل‌رانی کند. براساس عملکرد فرد در راندن اتومبیل در طول این فاصله‌ی معین، نسبت به توانایی کلی رانندگی او قضاوت صورت می‌گیرد.

مثالی دیگر از کاربرد روش نمونه‌ی کار در ارزشیابی از عملکرد یادگیرنده آموزش دانشجویی<sup>۲</sup> است. در آموزش دانشجویی از دانشجوی تربیت معلم خواسته می‌شود تا به عنوان نمونه، با استفاده از روش خاصی که در کلاس درس روانشناسی تربیتی یا روش‌های آموزشی یادگرفته است، درسی را به گروهی دانش‌آموز آموزش بدهد.

در سایر زمینه‌های آموزشی نیز می‌توان از روش نمونه‌ی کار به خوبی استفاده کرد. مثلاً در درس ترجمه می‌توان از دانشجویان خواست تا متنی را از زبانی به زبان دیگر برگردانند؛ در درس ماشین‌نویسی می‌توان از کارآموزان خواست تا یک نامه را ماشین‌کنند؛ یا در درس‌های حرفه‌وفن می‌توان از دانش‌آموزان خواست تا وسیله‌ای را تعمیر نمایند. روش نمونه‌ی کار برای سنجش توانایی افراد در انجام کارهای مختلف از دیرباز در صنعت و مشاغل نیازمند فعالیت‌های عملکردی متداول بوده است؛ می‌توان از آنها در محیط‌های آموزشی نیز حداکثر استفاده را به عمل آورد.

### دسته‌بندی نیتکو از روش‌های سنجش عملکردی

نیتکو (۲۰۰۱) از روش‌ها و فنون سنجش عملکردی یک تقسیم‌بندی چهارقسمتی به صورت زیر به دست داده است:

۱- سنجش ساختارمند یا کنترل شده

۲- سنجش در موقعیت‌های طبیعی

۳- پروژه‌های طولانی‌مدت

۴- کارپوشه

هر یک از روش‌های چهارگانه‌ی بالا شامل تعدادی فن سنجش است. در زیر به توضیح تقسیم‌بندی نیتکو از روش‌ها و فنون سنجش عملکردی توجه کنید.

---

<sup>۱</sup> -Work sample

<sup>۲</sup> -Student teaching

## سنجش ساختارمند

در سنجش ساختارمند<sup>۱</sup> یا کنترل شده معلم یا سنجش‌گر هم بر تکالیف سنجش و هم بر شرایطی که در آن دانش‌آموز باید پاسخ دهد کنترل اعمال می‌کند. به سخن دیگر، معلم تصمیم می‌گیرد که از چه مواردی استفاده شود و در چه زمانی آن موارد مورد استفاده قرار گیرند. همچنین به عهده‌ی معلم است که راهنمای عملکرد دانش‌آموز و نوع بازده‌های یادگیری او را تعیین کند. از این نظر، به تکالیف مورد استفاده در این روش تکالیف کنترل شده نیز گفته می‌شود.

فنون سنجش ساختارمند با توجه به تکالیف مورد استفاده دوتیپ (۱) تکالیف مداد و کاغذی و (۲) تکالیفی که به منابع و تجهیزات فراتر از مداد و کاغذ نیازمندند. تکالیف مداد و کاغذی در واقع همان آزمون‌های تشریحی یا انشایی هستند که انواع و چگونگی تهیه و تصحیح آنها را به تفصیل در فصل ۱۰ توضیح دادیم. در این نوع سنجش عملکردی از آزمون‌شونده خواسته می‌شود که نه تنها پاسخ خود را به صورت کتبی بنویسد بلکه درباره‌ی آن توضیح دهد و دلیل بیاورد. با این روش هم فرآیند اندیشه سنجش می‌شود و هم فرآورده‌ی آن. به نمونه‌های زیر از این‌گونه تکالیف سنجش توجه کنید:

- این مسئله داستان‌گونه را حل کنید و راه‌حل خود را توضیح دهید.

- پس از مطالعه‌ی نمودار زیر که نشان می‌دهد یک دانش‌آموز چگونه وقت خود را صرف می‌کند، درباره‌ی یک روز زندگی آن دانش‌آموز داستانی بنویسید.

در بسیاری از تکالیف سنجش، مداد و کاغذ به تنهایی کفایت نمی‌کند بلکه وسایل، ابزارها و تجهیزات دیگری مورد نیازمند. در این تکالیف دانش‌آموز هم با استفاده از وسایل کاری انجام می‌دهد و هم گاهی درباره‌ی کاری که انجام داده چیزی می‌نویسد. نمونه‌های این نوع تکالیف در زیر آمده‌اند:

- از یک مجموعه مثلث چهارتایی هر چند شکل هندسی که می‌توانید درست کنید.

- از طریق تلفت درباره‌ی یک شغل کسب اطلاع کنید و برای دریافت آن شغل تلفنی تقاضا نمایید.

- چگونگی مخلوط کردن اسید و آب را به طور عملی نشان دهید.

## سنجش در موقعیت‌های طبیعی

---

<sup>۱</sup>- Structured assessment



در روش سنجش در موقعیت‌های طبیعی<sup>۱</sup> به جای اینکه تکالیفی برای دانش‌آموز تعیین شود و موقعیت سنجش زیر کنترل قرار گیرد، معلم منتظر می‌ماند تا عملکرد در شرایط طبیعی رخ دهد و آن‌گاه به سنجش آن اقدام می‌کند. از آنجا که در این روش رفتارهای معمولی یا نوعی دانش‌آموز سنجیده می‌شوند به آن روش تکالیف عملکرد نوعی<sup>۲</sup> نیز گفته می‌شود. نمونه‌هایی از این‌گونه تکالیف به قرار زیرند:

- جمع‌آوری مطالبی که دانش‌آموز در زمینه‌های مختلف نوشته و تصحیح آنها از لحاظ غلط‌های دستوری، املائی، و انشایی.

- مشاهده‌ی یک دانش‌آموز در زمین بازی از لحاظ رفتاری که در ارتباط با درگیری بین دوستانش از خود نشان می‌دهد.

### پروژه‌های طولانی‌مدت

آنچه به پروژه‌های طولانی‌مدت<sup>۳</sup> شهرت یافته شامل دو قسمت است: (۱) پروژه‌های فردی<sup>۴</sup> و (۲) پروژه‌های گروهی<sup>۵</sup>. پروژه فردی به یک فعالیت طولانی‌مدت دانش‌آموزی گفته می‌شود که به نوعی فرآورده یا محصول مانند یک الگو، یک گزارش، یک وسیله، یا یک مجموعه (کلکسیون) می‌انجامد. ویژگی‌های پروژه‌ی دانش‌آموزی این است که در آن چند فعالیت مختلف با هم ترکیب می‌شوند تا یک فرآورده به وجود آید. برای نمونه، انجام یک تحقیق کتابخانه‌ای به عنوان یک پروژه‌ی دانش‌آموزی مستلزم انجام فعالیت‌های زیر است: استفاده از منابع تحقیقی موجود در کتابخانه، خلاصه‌برداری از مطالب، تهیه‌ی طرح یک گزارش، تدوین مطالب مختلف جمع‌آوری شده، تفسیر نتایج، سازمان‌بندی مطالب، و تهیه‌ی گزارش نهایی. در زیر نمونه‌های دیگری از پروژه‌های انفرادی داده شده‌اند:

- آگهی‌های روزنامه‌ها و مجله‌ها را جمع‌آوری و طبقه‌بندی کنید.

- الگوی یک دوربین عکاسی بسازید.

علاوه بر پروژه‌های فردی، معلم می‌تواند عملکرد دانش‌آموزان را در پروژه‌های گروهی نیز سنجش کند. در این پروژه‌ها دو یا چند تن از دانش‌آموزان همکاری می‌کنند. هدف اصلی پروژه‌ی گروهی، به عنوان یک وسیله‌ی سنجش، این

---

<sup>۱</sup> -naturally occurring

<sup>۲</sup> -Typical performance

<sup>۳</sup> Long term projects

<sup>۴</sup> -individual projects

<sup>۵</sup> - group projects

است که تعیین شود آیا دانش‌آموزان می‌توانند به صورت مشارکتی با هم کار کنند و یک محصول با ارزش تولید نمایند. نمونه‌هایی از پروژه‌های گروهی از قرار زیرند:

-تهیه‌ی یک روزنامه دیواری

-نوشتن یک مقاله درباره‌ی یک موضوع مورد علاقه‌ی دانش‌آموز

-تهیه‌ی یک گزارش درباره‌ی اینکه چگونه رودخانه‌ها شکل می‌گیرند.

## کارپوشه

منظور از کارپوشه<sup>۱</sup> مجموعه‌ی محدودی از کارهای دانش‌آموز است که برای سنجش عملکرد او مورد استفاده قرار می‌گیرد. کارپوشه هدف‌های مهم و متنوعی را برآورده می‌سازد که مهم‌ترین آنها یکی معرفی بهترین کارهای دانش‌آموز در طول یک مدت معین مثل یک ترم یا سال تحصیلی و دیگری نشان دادن چگونگی رشد تحصیلی او در یک مدت معین است.

## قواعد تهیه‌ی آزمون‌های عملکردی

در تهیه‌ی آزمون‌های عملکردی، همان مراحل و تدابیر مورد استفاده در ساختن سایر آزمون‌های معرفی شده مورد استفاده قرار می‌گیرند. با این حال، در تهیه‌ی آزمون‌های عملکردی رعایت پاره‌ای ملاحظات دیگر نیز ضروری‌اند که در این قسمت به آنها می‌پردازیم. یکی از تفاوت‌های عمده‌ی آزمون‌های عملکردی یا سایر آزمون‌هایی که در فصل‌های قبلی این کتاب معرفی شدند این است که تهیه و اجرای آزمون‌های عملکردی وقت بیشتری لازم دارد و نمره‌گذاری آنها دشوارتر است. قاعده‌هایی که در زیر مورد بحث قرار می‌گیرند، برای کاهش دادن دشواری تهیه، اجرا، و نمره‌گذاری این آزمون‌ها از سوی صاحب‌نظران پیشنهاد شده‌اند.

۱- **بازده‌های موردنظر عملکرد را مشخص کنید.** از میان هدف‌های آموزشی درس که قبلاً تهیه شده و در بُعد هدف جدول مشخصات قرار داده‌اید، آنهایی را برگزینید که به سنجش عملکردی نیاز دارند؛ هدف‌های قابل اندازه‌گیری با آزمون‌های مرسوم را به عهده‌ی همان آزمون‌ها بگذارید. دقت کنید آن دسته از هدف‌های آموزشی که برحسب عملکرد قابل اندازه‌گیری نوشته شده‌اند و هم شرایط عملکرد و هم ملاک عملکرد دارند با آزمون‌های عملکردی به راحتی قابل سنجش‌اند. بنابراین، سعی کنید، هنگام تهیه‌ی جدول مشخصات درس و نوشتن هدف‌های

---

<sup>۱</sup> -portfolio

آموزشی، هدف‌های عملکردی را در قالب عملکردهای یادگیرندگان بنویسید. ملاک یا معیار عملکرد را در هدف‌های عملکردی می‌توان با توجه به یکی از موارد زیر مشخص کرد.

الف- با توجه به دقت عملکرد، مثلاً:

با استفاده از یک دماسنج، دما را تا دهم درجه اندازه‌گیری کند.

ب- با توجه به سرعت عملکرد، مثلاً:

با در اختیار داشتن وسایل لازم، در پانزده دقیقه عیب یک رادیو را تشخیص دهد.

پ- با توجه به توالی درست مراحل، مثلاً:

ت- با توجه به مهارت عملکرد، مثلاً:

با در دست داشتن اره و چوب، یک قطعه چوب را با مهارت اره کند.

ث- با توجه به رعایت ایمنی، مثلاً:

اتومبیلی را بدون زیر پا گذاشتن هیچ یک از مقرراتی ایمنی رانندگی در دو چهارراه براند.

ملاک‌های عملکرد بالا را می‌توان به طور انفرادی یا به صورت ترکیبی مورد استفاده قرار داد. برای نمونه، در هدف زیر دو نوع ملاک عملکرد به کار رفته‌اند:

-یادگیرنده بتواند بر روی یک ماشین تحریر آی بی ام برقی، با سرعت پنجاه کلمه در دقیقه و حداکثر دو غلط در هر صفحه مطالبی را که به او می‌دهند ماشین کند.

در هدف بالا، هم از ملاک سرعت عملکرد (۵۰ کلمه در دقیقه) استفاده شده و هم از ملاک دقت عملکرد (حداکثر دو غلط در هر صفحه).

اینکه کدام یک از جنبه‌های عملکرد باید مورد تأکید قرار گیرد هم به ماهیت عملکرد و هم به مرحله‌ای از آموزش که عملکرد در آن باید سنجیده شود وابسته است. برای مثال، در سنجش مهارت، اندازه‌گیری در آزمایشگاه، در مراحل اولیه‌ی آموزش دقت مورد تأکید قرار می‌گیرد، ولی در مراحل آخر آموزش سرعت نیز اهمیت پیدا می‌کند. موقعیتی که در آن عملکرد باید انجام شود نیز باید مورد توجه قرار گیرد. مثلاً در سنجش مهارت ماشین‌نویسی در تایپ نامه‌های معمولی اداری سرعت عمل مهم است، در حالی که در ماشین کردن جدول‌های آماری یا گزارش‌های اقتصادی دقت عمل ضروری است.

گفتیم که از میان هدف‌های درس‌آنهايي را برای ارزشیابی با روش‌های سنجش عملکردی انتخاب کنید که با آزمون‌های مداد و کاغذی به خوبی قابل سنجش نیستند. کویبسیزین و بوریچ (۲۰۰۳) هدف‌های زیر را به عنوان نمونه‌های هدف‌های آموزشی معلم که بهتر است به کمک روش‌های سنجش عملکردی سنجش شوند ذکر کرده‌اند:

دانش آموز بتواند:

- برای سرزمین آمریکای شمالی، از حفظ، یک نقشه‌ی جغرافیایی رسم کند و بر روی آن محل ۱۰ شهر را علامت بزند.

- با استفاده از سیم، کلید، لامپ و باتری، یک مدار برقی درست کند.

- یک برنامه‌ی کامپیوتری برای حل معادله‌های یک مجهولی بنویسد.

۲- از تکالیف معنی‌دار استفاده کنید. تکالیفی که در آن یادگیرنده باید عملکرد یا توانایی خود را به منظور سنجش آن توسط معلم نشان دهد. باید تکالیفی معنی‌دار باشند. استفاده از تکالیف معنی‌دار سبب می‌شود که یادگیرندگان در انجام عملکرد یا حل مسائل شخصاً درگیر شوند. نیتکو (۲۰۰۱) برای خلق تکالیف معنی‌دار پیشنهادهای زیر را از بارون<sup>۱</sup> (۱۹۹۱) نقل کرده است:

الف- تکالیفی را برگزینید که برای دانش‌آموزان معنی‌دار باشند.

ب- موقعیت‌های تازه و آشنا را با هم ترکیب کنید به گونه‌ای که تکالیف سنجش برای دانش‌آموزان چالش‌انگیز باشند. یعنی تکالیف بسیار دشوار یا بیگانه را که موجب سرخوردگی دانش‌آموزان می‌شوند به کار نبرید و از تکالیف خیلی ساده هم پرهیز کنید.

پ- موقعیت‌هایی را برگزینید که بر تجارب دنیای واقعی دانش‌آموزان استوار باشند.

ت- از تکالیفی استفاده کنید که توانایی کار بستن دانش‌ها و مهارت‌های آموخته شده‌ی دانش‌آموز از درون کلاس به بیرون از کلاس را نشان دهد.

۳- موقعیت‌آزمون را واقع‌بینانه برگزینید. منظور از این پیشنهاد آن است که موقعیت‌ها و شرایطی را که می‌خواهید در آن عملکرد مورد سنجش را بیازمایید بین گونه‌ای انتخاب کنید که همه‌ی جوانب امر از جمله ملاحظات اقتصادی، زمانی، انسانی، عملی، و غیره را منظور نمایید. برای مثال، در آزمون مهارت‌های کمک‌های اولیه، استفاده از بیماران واقعی دارای استخوان‌های شکسته، زخم‌های شدید، و سایر مشکلات جدی نه میسر است و نه مطلوب. همچنین اگر استفاده از یک دستگاه گران‌قیمت و کمیاب برای آزمون مهارت آزمون‌شوندگان در راه‌اندازی آنها ممکن است موجب ایجاد خسارت به دستگاه بشود، بهتر است برای این منظور از یک دستگاه شبیه‌سازی شده یا طرح دستگاه واقعی استفاده شود.

۴- از راهنمایی‌ها و دستورالعمل‌هایی که به روشنی موقعیت‌آزمون و چگونگی انجام عملکرد را مشخص می‌کنند استفاده نمایید. منظور از این پیشنهاد آن است که راهنمایی‌هایی باید برای آزمون فراهم شوند که در آنها عملکرد

---

<sup>۱</sup> -Barone

موردنظر و شرایطی که در آن قرار است عملکرد نشان داده شود مشخص گردند. گرانلاند (۱۹۸۸) موارد زیر را به برای منظور کردن در راهنمایی یا دستورالعمل آزمون نمونه‌ی کار پیشنهاد داده است:

الف- مقصود از آزمون

ب- مواد و تجهیزات

پ- جریان آزمون:

- شرایط تجهیزات

- عملکرد مورد انتظار

- محدوده‌ی زمانی (در صورتی که محدودیت زمانی لازم باشد)

ت- روش نمره‌گذاری

دستورالعمل‌ها باید به صورت کتبی باشند تا اینکه همه‌ی افراد تکالیف را یکسان دریافت کنند. در بعضی مواقع، آزمون‌شوندگان دستورالعمل‌ها را خود می‌خوانند ولی در مواقع دیگر برای آنان خوانده می‌شوند. در هر حال، دستورالعمل‌ها یا راهنمایی‌های چگونگی انجام عملکرد باید کاملاً واضح باشند و برای همه‌ی دانش‌آموزان به طور یکنواخت به کار بروند.

۵- در ارزشیابی از عملکرد یادگیرندگان از راهنمای نمره‌گذاری استفاده کنید. منظور از راهنمایی ارزشیابی یا راهنمای نمره‌گذاری نوعی قاعده برای سنجش عملکرد یا داوری درباره‌ی پاسخ یادگیرنده است. عموماً برای ارزشیابی از عملکرد در روش‌های سنجش عملکردی از سه دسته‌ی راهنما با عنوان تحلیلی، کلی (یا سراسری)، و کمی، استفاده می‌شود.

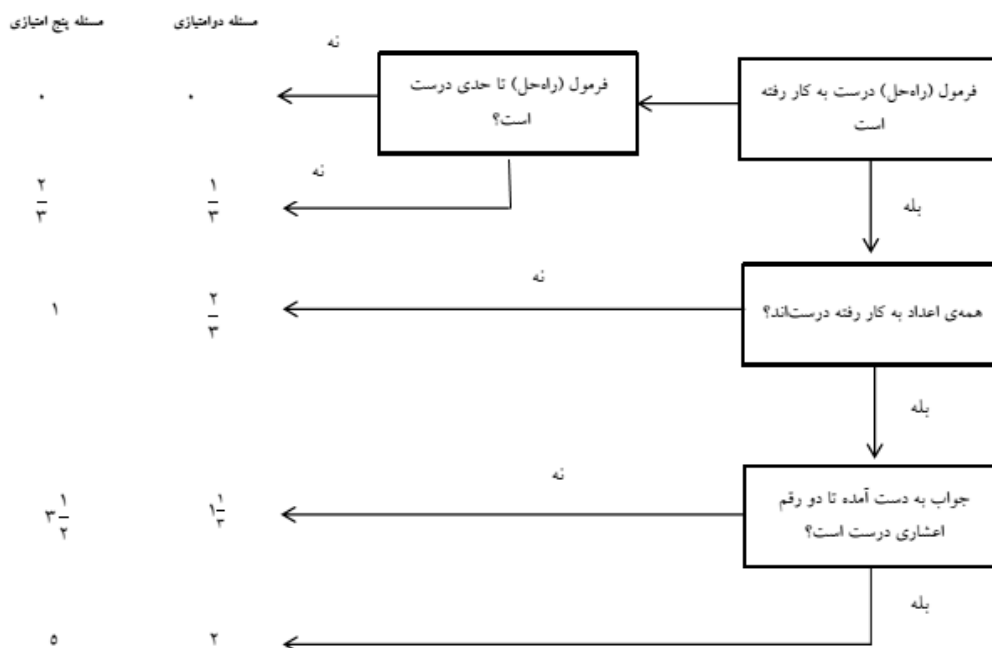
در روش تحلیلی، ابتدا ابعاد یا ویژگی‌های پاسخ مورد سنجش را مشخص می‌کنیم و سپس برای هر یک از آنها ارزش یا نمره‌ای در نظر می‌گیریم. به عنوان نمونه، یک گزارش کتبی درباره‌ی نتایج یک آزمایش علمی را می‌توان با توجه به صحت اطلاعات، کیفیت تحلیل نتایج، و چگونگی توجیه نتایج به دست آمده مورد ارزشیابی قرار داد (لین و گرانلاند، ۲۰۰۰).

در روش تحلیلی نمره‌گذاری، پس از آنکه اجزای پاسخ نمونه (فرآیند یا فرآورده) نمره‌گذاری شدند، این نمره‌ها یا امتیازهای جداگانه با هم جمع می‌شوند و نمره‌ی کل دانش‌آموز در سؤال تعیین می‌گردد. شکل ۱-۱۲ یک نمونه پاسخ تصحیح شده با روش تحلیلی نمره‌گذاری را نشان می‌دهد. در این مثال، اجرای راه‌حل برای یک مسئله ریاضی با دو نوع امتیازبندی (۲ و ۵) نمره‌گذاری شده‌اند. دقت کنید که در این مثال، چون نمره‌ی کل از ۲ به ۵ تغییر کرده، امتیازهای قسمت‌های مختلف راه‌حل مسئله نیز متفاوت شده است.

علاوه بر مثال‌های شکل ۱-۱۲، در جدول ۲-۱۲ نیز چند مورد راهنمای نمره‌گذاری به روش تحلیلی نشان داده شده. در این مثال‌ها، جزئیات پاسخ موردنظر به تفکیک مشخص شده‌اند. معلم در نمره‌گذاری خود می‌تواند برای هر قسمت امتیازهای مختلفی منظور کند.

شکل ۱-۱۲ روش تحلیلی نمره دادن به اجزای راه حل یک مسئله ریاضی (با دو نوع امتیازبندی) (نیتکو،

۲۰۰۱، ص ۱۹۵)



در روش نمره گذاری کلی یا سراسری یک داوری کلی درباره ی پاسخ یادگیرنده به عمل می آید. به این منظور، تعدادی طبقه از قبل مشخص می شوند و قاعده ی نمره گذاری کلی در رابطه با آن طبقه ها و با توجه به میزان شایستگی پاسخها تعریف می شود: مثلاً، خیلی خوب، خوب، متوسط و غیره. لین و گرانلاند (۲۰۰۰)، در مقایسه ی روش کلی ارزشیابی با روش تحلیلی، گفته اند "روش کلی با داوری های کلی مورد نیاز برای نمره دادن به یادگیرندگان سازگار است، اما بازخوردهای لازم را درباره ی نقاط قوت و ضعف یادگیرندگان، آن گونه که در روش تحلیلی امکان پذیر است، در اختیار آنان قرار نمی دهد" (ص ۲۷۱).

در روش کلی یا سراسری نمره گذاری، کل فرآیند یا فرآورده یکبار مورد داوری قرار می گیرد و به اجزاء یا قسمت های مختلف آن نمره یا امتیاز جداگانه داده نمی شود. جدول ۳-۱۲ نمونه ای از یک دستورالعمل یا راهنمای تصحیح با روش کلی را نشان می دهد. در این نمونه مهارت یادگیرنده در بیان شفاهی یک زبان خارجی ارزشیابی می شود.

در تعیین راهنما یا دستورالعمل برای نمره گذاری پاسخ های آزمون شوندگان، همچنین مشخص کنید که به آنان به صورت فردی نمره خواهید داد یا به صورت گروهی. استفاده از سنجش گروهی برای ارزشیابی از عملکرد دانش آموزان در موقعیت های یادگیری مشارکتی مناسب است. در این ارزشیابی هم می توان سهم هر یک از اعضای گروه را تعیین کرد و هم فرآورده یا محصول کار گروهی را سنجید.

جدول ۲-۱۲ مثال‌هایی از راهنمای نمره‌گذاری با روش تحلیلی (پین، ۲۰۰۳، ص ۲۹۳)

ریاضیات	نوشتن	هنر
حل مسئله	هدفها	عناصر صوری
• راهبردها:	- شرح دادن	- ساختار
- تشخیص مسئله	- ترغیب کردن	- ترکیب
- به کار بستن راهبردها، استفاده از مفاهیم، روش‌ها، ابزارها	- اطلاع دادن	عناصر فنی
	جنبه‌ها	- تکنیک‌ها
	- ترکیب	- مواد
بازنمایی:	- تأکید	
- چارت	- تفصیل	عناصر حسی
- نمودار	- سازمان	• بیانی:
	- امور قراردادی	- حالت
استدلال		- کیفیت هیجانی
- تفسیر		- کیفیت انرژی
- تعمیم		
ارتباط		عناصر هویتی
- آشکار		• انسجام عناصر
- سازمان یافته		• تأثیر عناصر
- کامل		
- مفصل		
- اصطلاحات، نمادها و علائم		
زبان ریاضی		
محتوا		



		- مفاهیم عددی - نسبت / درصد - مفاهیم جبری - مفاهیم هنری
--	--	--

جدول ۳-۱۲ مثالی از روش کلی سنجش عملکرد (سنجش مهارت شفاهی در زبان دوم)

(نیتکو، ۲۰۰۱، ص ۲۵۹)

سطح	وصف
عالی	<p>سطح عالی با توانایی یادگیرنده (گوینده) در ویژگی‌های زیر تعیین می‌شود:</p> <ul style="list-style-type: none"> <li>• شرکت مؤثر در اکثریت مکالمه‌های رسمی و غیررسمی درباره‌ی موضوع‌های عملی، اجتماعی، حرفه‌ای، و انتزاعی</li> <li>• جانب‌داری از عقاید و فرضیه‌ها با استفاده از راهبردهای گفتنی</li> </ul>
پیشرفته	<p>سطح پیشرفته با توانایی یادگیرنده (گوینده) در ویژگی‌های زیر تعیین می‌شود:</p> <ul style="list-style-type: none"> <li>• انجام گفتگو به شکلی روشن و مشارکتی</li> <li>• آغاز کردن، نگهداری، و به نتیجه رسانیدن انواع مختلف تکالیف ارتباطی، از جمله آنهایی که نیازمند توانایی سطح بالایی در انتقال معنی به کمک راهبردهای متنوع زبانی هستند.</li> <li>• پاسخگویی به خواسته‌های مدرسه</li> <li>• بیان و توصیف مطالب در قالب پاراگراف‌های منسجم</li> </ul>
متوسط	<p>سطح متوسط با توانایی یادگیرنده (گوینده) در ویژگی‌های زیر تعیین می‌شود:</p> <ul style="list-style-type: none"> <li>• استفاده از عناصر زبانی آموخته شده در ترکیب‌های تازه که عمدتاً حالت واکنشی داشته باشند.</li> <li>• آغاز کردن، تا حدی نگهداری کردن، و به نتیجه رساندن تکالیف ساده‌ی ارتباطی</li> <li>• پرسیدن سؤال و پاسخ دادن به سؤال‌ها</li> </ul>
ابتدایی	<p>سطح ابتدایی با توانایی یادگیرنده (گوینده) در ایجاد حداقلی از ارتباط با استفاده از مطالبی که آموخته است مشخص می‌شود.</p>

نکته‌ی دیگری که در اینجا لازم است به آن اشاره کنیم این است که از بعضی فنون سنجش عملکردی به این منظور که یادگیرندگان در مورد عملکرد خودشان داوری نمایند نیز می‌توان استفاده کرد. یعنی اینکه سنجش فعالیت‌های یادگیرندگان به صورت خودسنجی<sup>۱</sup> صورت پذیرد. ساتروک (۲۰۰۴) در این باره گفته است، "می‌توان از دانش‌آموزان و دانشجویان خواست تا درباره‌ی کیفیت گزارش شفاهی خودشان یا بازی‌های نمایشی‌شان داوری کنند" (ص ۵۳۸).

۶- از روش‌ها و فنون گوناگون مشاهده استفاده نمایید. چنان‌که قبلاً در آغاز این فصل توضیح دادیم، ارزشیابی از عملکرد بر فرآیند، فرآورده، یا ترکیبی از این دو تأکید می‌کند. فرآیندها و فرآورده‌های عملکرد معمولاً به وسیله‌ی روش‌ها و فنون مختلف مشاهده، نظیر فهرست واری، مقیاس درجه‌بندی، و نظایر آنها سنجیده می‌شوند. بنابراین، در سنجش عملکرد یادگیرندگان از این وسایل اندازه‌گیری حداکثر استفاده را بنمایید.

### یک نمونه آزمون عملکردی همراه با راهنمای انجام عملکرد و سنجش پاسخ

در شکل ۲-۱۲ یک تکلیف مربوط به سنجش عملکردی در علوم داده شده، در این تکلیف دانش‌آموزان باید یک آزمایش انجام دهند و اشیاء مورد آزمایش را طبقه‌بندی نمایند. آنها همچنین باید برای طبقه‌بندی خود دلیل بیاورند. اما تکلیف استفاده از آهنربا برای آزمایش کردن اشیاء، طبقه‌بندی آنها به دو دسته‌ی مختلف، و بیان دلایل تفاوت بین اشیاء در دو طبقه‌ی موردنظر توضیح داده شده است.

امتیازها و محدودیت‌های روش‌های سنجش عملکردی

یکی از امتیازهای سنجش عملکردی یا آزمون عملکردی این است که به وسیله‌ی آن می‌توان یادگیری‌هایی را سنجید که به کمک آزمون‌های کتبی یا مداد و ماغذی مرسوم میسر نیست. چنان‌که قبلاً اشاره کردیم، دانستن چگونه انجام دادن یک کار با توانایی انجام دادن آن فرق می‌کند. به عنوان مثال، "یک آزمون کتبی که می‌تواند دانش یادگیرنده را درباره‌ی ویژگی‌های مؤثر ایراد سخنرانی در حضور جمع بسنجد درباره‌ی اینکه آیا او می‌تواند در حضور جمع سخنرانی کند هیچ‌گونه اطلاعاتی به ما نمی‌دهد" (لین و گراندلاند، ۲۰۰۰، ص ۱۱).

دومین امتیاز سنجش عملکردی که آن هم قبلاً مورد اشاره واقع شد این است که در این روش هم فرآورده‌ی عمل یادگیرنده و هم فرآیند یا جریان کار او سنجیده می‌شود. به عنوان نمونه، وقتی که دانش‌آموزان مشغول انجام دادن یک آزمایش علمی هستند، به کمک سنجش عملکردی، هم می‌توان نقاط قوت و ضعفشان را در استفاده از ابزارها و فنون آزمایش سنجش کرد (سنجش فرآیند) و هم می‌توان توفیقشان را در کامل کردن آزمایش و مستدل ساختن نتیجه‌گیری‌های خود مورد ارزشیابی قرار داد (سنجش فرآورده).

<sup>۱</sup> -Self- assessment

سومین امتیازی که می‌توان برای سنجش عملکردی ذکر این است که در این نوع سنجش بر نظریه‌های جدید یادگیری تأکید می‌شود. بنا به نظریه‌های جدید یادگیری، از جمله نظریه‌های سازندگی یا سازنده‌گرایی<sup>۱</sup>، نه تنها یادگیرندگان دریافت‌کنندگان محض اطلاعات به حساب نمی‌آیند، بلکه چنین فرض می‌شود که آنان در ساختن و کشف دانش نقش مهمی برعهده دارند و فعالیت‌شان در جریان یادگیری تضمین‌کننده‌ی توفیق یادگیری آنان است (سیف، ۱۳۸۰). طبق این نظریه‌ها، اطلاعات جدید باید با اطلاعات قبلاً آموخته شده‌ی یادگیرندگان درهم آمیزد. "سنجش عملکردی دارای کیفیت سطح بالا دانش قبلی را مورد توجه قرار می‌دهد و یادگیرنده را در ساختن فعال معنی درگیر می‌سازد" (لین و گرانلاند، ۲۰۰۰، ص ۲۶۷).

در مقابل امتیازهای بالا، سنجش عملکردی دارای محدودیت‌هایی است که به آنها نیز اشاره می‌کنیم. مهم‌ترین محدودیتی که برای این روش گفته شد همان مشکل مربوط به آزمون‌های تشریحی است، یعنی کمبود پایایی یا دقت نمره‌گذاری. بنابراین، رعایت پیشنهادهایی که در در رابطه با آزمون‌های تشریحی گفتیم می‌تواند دقت روش‌های سنجش عملکردی را نیز افزایش دهد. به این منظور، معلمان باید بکوشند تا بازده‌های مورد سنجش را به روشنی تعریف کنند، تکالیفی را که در آن توانایی یادگیرنده قرار است سنجش شود انتخاب نمایند، و دستورالعمل یا راهنمای نمره‌گذاری را به طور کامل مشخص کنند و به اجرا درآورند.

محدودیت دیگر روش‌های سنجش عملکردی صرف وقت زیاد برای آنهاست. از آنجا که در سنجش عملکردی برای هر هدف یادگیری یک سنجش مستقل باید صورت پذیرد و هر نوبت سنجش به وقت قابل توجهی نیاز دارد، سنجش کامل عملکردهای یادگیرنده در ارتباط با یک موضوع درسی مستلزم صرف وقت زیادی است. برای رفع این مشکل، لازم است در طول ترم تحصیلی سنجش عملکردهای دانش‌آموزان به دفعات مختلف انجام شود، و کوشش به عمل آید تا هر تکلیف سنجش همراه یادگیری صورت پذیرد. یعنی معلم، ضمن سنجش، آموزش هم بدهد تا از وقت کلاس به نحو بهتری استفاده نماید.

### موارد استفاده‌ی روش‌های سنجش عملکردی

چنان‌که دیدیم، موارد استفاده‌ی روش‌های سنجش عملکردی بسیارند، و تقریباً همه‌ی هدف‌های حوزه‌های شناختی، عاطفی و روانی - حرکتی را دربرمی‌گیرند. در موقعیت‌هایی که روش‌های سنتی مداد و کاغذی کارساز نیستند، روش‌ها و فنون سنجش عملکردی قابل استفاده‌اند - به‌ویژه در شرایطی که هم فرآورده‌های رفتار و عملکرد یادگیرنده موردنظر است و هم فرآیندهای آن.

---

<sup>۱</sup> -Constructivism

پیش از معرفی نمونه‌های سنجش عملکردی، لازم است یادآوری کنیم که استفاده از روش‌ها و آزمون‌های عملکردی برای سنجش هدف‌های سطح پایین، یعنی یادآوری دانش و محفوظات، مفید نیست و به این منظور باید از آزمون‌ها و روش‌های نوع دیگری استفاده شود. علاوه بر این، هر چند که روش‌ها و فنون سنجش عملکردی برای سنجش و ارزشیابی یادگیری دانش‌آموزان و دانشجویان در کلاس درس، به ویژه به عنوان یک روش سنجش همراه با یادگیری، بسیار مفید است، اما استفاده از آن برای مقاصد هنجاری توصیه نشده. به عنوان نمونه، پوفام (۲۰۰۲) گفته است استفاده‌ی معلمان از روش سنجش عملکردی در سطح کلاس درس بسیار مؤثر است و موجب بهبود کیفیت آموزش معلمان می‌شود، اما استفاده از آن برای مقاصد پاسخگویی یا مسئولیت‌پذیری<sup>۱</sup> - یعنی تعیین اینکه معلمان یک آموزشگاه یا ناحیه یا استان در کار آموزشی خود تا چه اندازه موفق‌اند - چندان جایز نیست.

مثالی از سنجش عملکردی: سنجش هوش‌های چندگانه

هوارد گاردنر<sup>۲</sup> (۱۹۸۳، ۱۹۹۸)، یکی از نظریه پردازان هوش، نظریه‌ای وضع کرده شامل هشت نوع هوش متفاوت و مستقل از یکدیگر با نام‌های زیر: زبانی<sup>۳</sup>، منطقی- ریاضی<sup>۴</sup>، فضایی<sup>۵</sup>، بدنی- جنبشی<sup>۶</sup>، موسیقایی<sup>۷</sup>، میان‌فردی<sup>۸</sup>، درون فردی<sup>۹</sup>، و طبیعت‌گرایانه<sup>۱۰</sup>، توصیف مختصر هر یک از این هوش‌ها در جدول ۴-۱۲ آمده است.

جدول ۴-۱۲ توصیف مختصر هوش‌های هشت‌گانه‌ی گاردنر (اقتباس براساس نیتکو، ۲۰۰۱، ص ۲۴۵)

نوع هوش	توصیف
هوش زبانی	توانایی استفاده از زبان خودتان و دیگر زبان‌های برای فهماندن مقاصد خود به دیگران و درک مقاصد آنان
هوش منطقی - ریاضی	توانایی درک اصول علت و معلولی به گونه‌ای شبیه به درک دانشمندان یا منطقدانان، با استفاده از استدلال کمی و ریاضی به گونه‌ای شبیه به

<sup>۱</sup> - این مفهوم در فصل اول توضیح داده شد.

<sup>۲</sup> - Howard Gardner

<sup>۳</sup> - Linguistic

<sup>۴</sup> Logical- mathematical

<sup>۵</sup> - spatial

<sup>۶</sup> - bodily- kinesthetic

<sup>۷</sup> - musical

<sup>۸</sup> Interpersonal

<sup>۹</sup> - intrapersonal

<sup>۱۰</sup> - naturalistic

استفاده‌ی ریاضی‌دانان	
توانایی بازنمایی جنبه‌های فضایی دنیا در ذهن خود، مانند خلبانان، بازیکنان شطرنج، نقاشان، مجسمه‌سازان، یا آرشیتکت‌ها	هوش فضایی
توانایی استفاده از قسمتی یا تمامی بدن خود برای حل یک مسئله، تولید یک فرآورده، یا عمل کردن شبیه به ورزشکاران یا بازیگران	هوش بدنی - جنبشی
توانایی پردازش ذهنی موسیقی به گونه‌ای که موجب شناسایی الگوها، به یاد سپردن الگوها، و دستکاری موسیقی برای حل مسائل یا بیان درک و فهم بشود.	هوش موسیقایی
توانایی درک افراد دیگر و برقرار کردن رابطه‌ی معنی‌دار با آنان. این کار از راه درک اشخاص دیگر امکان‌پذیر است. یعنی چه کاری می‌توانند انجام دهند، چگونه با دنیا و دیگران رابطه برقرار می‌کنند، واکنش احتمالی آنان چیست، چه چیزی را دوست دارند، از چه چیزی اجتناب می‌کنند، و چه احساسی دارند.	هوش میان‌فردی یا بین افراد
توانایی درک خودتان، دانستن اینکه چه کسی هستید، توانایی‌ها و ناتوانایی‌های شما در چیست، آرزوها و هدف‌هایتان کدام‌اند، چه احساسی دارید، از چه چیزی اجتناب می‌کنید، و در شرایط مختلف چگونه واکنش می‌کنید.	هوش درون فردی
توانایی درک طبیعت و دنیای مدرن از راه تمیز قائل شدن بین موجودات طبیعی زنده و غیرزنده و اشیاء ساخته شده‌ی دست انسان و دسته‌بندی کردن آنها	هوش طبیعت‌گرایانه

به باور گاردنر، همه‌ی افراد از این هشت نوع هوش بهره‌مندند، اما در بعضی از آنان بعضی از این هوش‌ها قوی‌تر و بعضی ضعیف‌ترند. گاردنر در توجیه نظریه‌ی خود و وجود این هوش‌ها دلایل زیر را ذکر کرده است: (۱) این هوش‌ها برای بقای ما ضروری‌اند؛ (۲) اینها دارای ویژگی‌های تکاملی هستند (تا حدودی در جانداران غیرانسان نیز یافت می‌شوند)؛ (۳) به وسیله‌ی فعالیت‌های مغزی ویژه قابل شناسایی‌اند.

نیتکو (۲۰۰۱) گفته است در بعضی مدارس آمریکا معلمان فعالیت‌های آموزشی خود را براساس هوش‌های هشت‌گانه‌ی گاردنر سازمان می‌دهند. وی همچنین می‌گوید این نظریه الهام‌بخش یک فلسفه‌ی تربیتی و فعالیت‌های کلاسی متعدد بوده است که هدف آنها آموزش دادن موارد زیر به دانش‌آموزان است:

۱- آگاه باشند که در راه‌های مختلف می‌توانند هوشمند شوند.

۲- هوش‌های مختلف را در خود و دیگران ارج بگذارند.

۳- هوش‌های مختلف خود را به کار بندند، تمرین دهند، و تا حد توان خود در پرورش آنها بکوشند.

۴- پیشرفت خود را در هدف‌های آموزشی یا چند تا از هوش‌های خود نشان دهند.

### سنجش هوش‌های چندگانه

مدافعان نظریه‌ی هوش‌های هشت‌گانه‌ی گاردنر کوشیده‌اند تا روش سنجش عملکردی را برای اندازه‌گیری آنها به کار بندند. به باور این افراد، روش‌های سنجش عملکردی می‌تواند امکاناتی فراهم آورد تا پیشرفت دانش‌آموزان به چندین راه مختلف مورد شناسایی و ارزشیابی قرار گیرد. گاردنر در رابطه با سنجش عملکردی هوش‌های مختلف گفته است: تأکید جاری بر سنجش عملکردی با نظریه‌ی هوش‌های چندگانه به خوبی حمایت می‌شود... نخست اینکه، اجازه بدهید از صافی آزمون کوتاه پاسخ به چیزها نگاه نکنیم. در عوض، اجازه بدهید تا به عملکردی که برایش ارزش قائل هستیم نگاه کنیم- چه آن عملکرد زبانی باشد، چه منطقی، زیبایی‌شناختی، یا اجتماعی. دوم اینکه، هرگز اجازه ندهید که سنجش خود را از درک و فهم دانش‌آموز صرفاً بر یک اندازه‌ی خاص متمرکز کنیم، بلکه اجازه بدهید تا به دانش‌آموزان فرصت دهیم تا درک و فهم خود را در چند راه مختلف نشان دهند. (گاردنر، به نقل نیتکو، ۲۰۰۱، ص ۲۴۶).

بنا به گفته‌ی نیتکو (۲۰۰۱) استفاده از روش‌های سنجش عملکردی برای اندازه‌گیری هوش‌های چندگانه مستلزم این است که استانداردهای سطح بالایی برای توفیق یادگیرنده در هدف‌های یادگیری منظور شود و از او خواسته شود تا آن موفقیت را به راه‌های مختلفی نشان دهد. در جدول ۵-۱۲، برای هر یک از هوش‌های گاردنر، تعدادی عملکرد نمونه و مثال‌هایی از سنجش عملکرد آمده است.

جدول ۵-۱۲ سنجش عملکردی براساس نظریه‌ی هوش‌های هشت‌گانه‌ی گاردنر (اقتباس براساس نیتکو، ۲۰۰۱، ص ۲۴۶) به دانش‌آموزان اجازه داده می‌شود تا با انتخاب یک الگوی سنجش از میان یکی یا چند تا از طبقه‌های هوشی پیشرفت خود را نشان دهند.

هوش موسیقیایی	هوش زبانی
---------------	-----------

<ul style="list-style-type: none"> <li>• آهنگی درباره‌ی ..... بخوانید.</li> <li>• توضیح دهید که چگونه موسیقی یک آهنگ شبیه... است.</li> <li>• یک ابزار موسیقی بسازید که ..... را نشان دهد.</li> </ul> <p style="text-align: center;"><b>هوش بین فردی</b></p> <ul style="list-style-type: none"> <li>• به کسی درباره‌ی ..... آموزش دهید.</li> <li>• درباره‌ی موضوع ..... یک سخنرانی ایراد کنید.</li> <li>• از تکنولوژی به منظور ..... استفاده کنید.</li> </ul> <p style="text-align: center;"><b>هوش درون فردی</b></p> <ul style="list-style-type: none"> <li>• درباره‌ی ..... برای خود هدفی برگزینید و برای رسیدن به آن بکوشید.</li> <li>• یکی از ارزش‌های شخصی خود را درباره‌ی ..... توضیح دهید.</li> <li>• فعالیت خود را در رابطه با ..... بسنجید.</li> </ul> <p style="text-align: center;"><b>هوش طبیعت‌گرایانه</b></p> <ul style="list-style-type: none"> <li>• تغییرات ایجاد شده در محیط زندگی خود را توضیح دهید.</li> <li>• توضیح دهید که از حیوانات خانگی، حیوان‌های وحشی، باغچه، چگونه مواظبت به عمل می‌آورید؟</li> <li>• اشیاء طبیعی را نقاشی کنید یا از آنها عکس بگیرید.</li> </ul>	<ul style="list-style-type: none"> <li>• یک گفتگوی رادیویی درباره ..... تولید کنید.</li> <li>• درباره‌ی ..... بحث کنید.</li> <li>• درباره‌ی ..... یک شعر، قصه، یا مقاله بنویسید.</li> </ul> <p style="text-align: center;"><b>هوش منطقی - ریاضی</b></p> <ul style="list-style-type: none"> <li>• مطلب ..... را به یک فرمول ریاضی تبدیل کنید.</li> <li>• یک آزمایش علمی درباره‌ی ..... طرح و اجرا کنید.</li> <li>• با استدلال منطقی ..... را توضیح دهید.</li> </ul> <p style="text-align: center;"><b>هوش بدنی - جنبشی</b></p> <ul style="list-style-type: none"> <li>• با یک حرکت یا تعدادی حرکت ..... را نشان دهید.</li> <li>• یک ..... بسازید.</li> <li>• یک برنامه‌ی گردش علمی به منظور ..... طراحی کنید و در آن شرکت نمایید.</li> </ul> <p style="text-align: center;"><b>هوش فضایی (دیداری)</b></p> <ul style="list-style-type: none"> <li>• برای ..... یک نمودار، نقشه، یا طرح بکشید.</li> <li>• یک اثر هنری تولید کنید که ..... را نشان دهد.</li> <li>• یک بازی با چوب یا ورق ..... خلق کنید.</li> </ul>
---	---

#### مثالی از سنجش عملکردی: حل مسائل زندگی واقعی

در تکلیف زیر از دانش آموزان خواسته می‌شود تا درباره‌ی یکی از موقعیت‌های واقعی زندگی به جمع‌آوری اطلاعات لازم بپردازند و آنچه را که به صورت یک مسئله به آنان ارائه شده است حل نمایند. این سؤال عملکردی از مجموعه آزمون‌های مبانی یادگیری عمومی ایالت کانکتیکت آمریکا برگزیده شده (به نقل از وولفلک، ۲۰۰۴، ص ۵۵۵).

سؤال: بسیاری از فروشگاه‌های محلی ادعا می‌کنند که کالاهای خود را ارزان می‌فروشند. اما واقعاً این ادعا چه معنی می‌دهد؟ آیا معنی ادعای آنان این است که همه‌ی اجناس فروشگاه‌هایشان دارای قیمت پایین‌تر است با فقط بعضی از آنها؟ چگونه می‌توانید تعیین کنید که خرید از کدام فروشگاه بیشتر به نفع شماست.

تکلیف شما این است که پژوهشی را طراحی و اجرا کنید که به سؤال‌های بالا جواب دهد. چه اجناس و قیمت‌هایی را با هم مقایسه خواهید کرد و چرا؟ نمونه‌ای را که انتخاب خواهید کرد چگونه توجیه می‌کنید و اعتبار آن را چگونه می‌سنجید؟

اجرا، نمره گذاری و تحلیل آزمون



در این فصل روش اجرای آزمونهای کتبی، طرز نمره گذاری پاسخها، و چگونگی تحلیل سؤالهای این آزمونها را یاد خواهیم گرفت. همچنین در این فصل خواهیم آموخت که با استفاده از نتایج حاصل از آزمون خود چگونه به تجدید نظر در آن و رفع نواقص و اشکالات آن اقدام کنیم.

### شرایط اجرای آزمونهای کتبی

گام نخست در اجرای آزمونهای پیشرفت تحصیلی آماده سازی و فراهم آوردن شرایط مناسب برای اجرای آنهاست. تهیه دفترچه آزمون و پاسخنامه، ترتیب پشت سر هم قرار دادن سؤالهای مختلف، و در نظر گرفتن زمان و مکان مناسب برای اجرای آزمون نیاز به دقت فراوان دارند. مجموعه این عوامل باید به گونه‌ای فراهم شوند که برای هریک از آزمون شوندگان حداکثر امکانات را جهت نشان دادن بهترین عملکرد او فراهم آورند. نکاتی که در زیر توضیح میدهم ناظر بر این هدف هستند.

### ترتیب قرار دادن سؤالهای آزمون به دنبال یکدیگر

قبلاً در رابطه با چگونگی تهیه سؤالهای مختلف آزمونهای پیشرفت تحصیلی گفتیم که سؤالهایی باید تهیه شوند که تکالیفی را که در اختیار آزمون شوندگان قرار می‌دهند تا حد امکان برای آنان آشکار و قابل فهم باشند. یکی از راههای کمک به این امر آن است که سؤالهای مختلف یک آزمون را با رعایت اصول صحیح به دنبال هم قرار دهیم. این کار به آزمون شونده نظم فکری می‌دهد و به او کمک می‌کند تا راحت تر به سؤالها پاسخ بدهد.

گرانلاند ولین (۱۹۹۰) در رابطه با قرار دادن سؤالهای آزمون به دنبال یکدیگر چند پیشنهاد زیر را ارائه داده اند:

۱. سؤالها را با توجه به نوع آنها دسته بندی کنید و پشت سر هم قرار دهید. در این باره می‌توان ترتیب زیر را برای ردیف کردن سؤالهای یک آزمون به دنبال هم مورد استفاده قرار دادن

الف - صحیح - غلط

ب - جورکردنی

پ - چندگزینه‌ای

ت - کوتاه پاسخ

ث - تشریحی

علاوه بر دسته بندی سؤالها برحسب نوع، می‌توان سؤالهای مختلف هر نوع را نیز برحسب طبقه‌بندی هدفهای آموزشی دسته بندی کرد. برای نمونه، سؤالهای مربوط به بخش چندگزینه‌ای با میتوان به ترتیب (۱) به یاد آوردن

دانش، (۲) فهمیدن دانش، (۳) به کار بستن دانش، (۴) تحلیل دانش و مانند اینها مرتب کرد. "کنار هم قرار دادن سؤلهایی که بازده های یادگیری مشابهی را می‌سنجند در تعیین نوع بازده های یادگیری که برای یادگیرندگان مشکل‌زا هستند معلم را یاری می‌دهند (گرانلاند ولین، ۱۹۹۰، ص ۲۳۴).

**۲. سؤلهای را از ساده به دشوار مرتب کنید.** اگر مرتب کردن سؤلهای یک آزمون با توجه به بازده های یادگیری امکان پذیر نباشد، می‌توانید آنها را، با توجه به ساده و دشوار بودن، پشت سر هم قرار دهید. یعنی ابتدا سؤلهای ساده را قرار دهید و بعد به ترتیب سؤلهایی را که از لحاظ درجه دشواری در سطح بالاتری قرار دارند به دنبال هم بیاورید.

وقتی که یک آزمون با سؤلهای ساده آغاز می‌شود، آزمون شوندگان با پاسخ دادن به این سؤلهای احساس اعتماد به نفس می‌کنند و این امر کارآیی آنان را در عملکرد روی سؤلهای پیچیده بعدی نیز افزایش میدهد. در این حالت، حتی ضعیف‌ترین دانش‌آموزان نیز خواهند کوشید تا حداکثر توان خود را در پاسخدهی به سؤلهای به کارگیرند. دقت کنید که در این روش مرتب کردن سؤلهای نیز ابتدا سؤلهای برحسب نوع دسته بندی می‌شوند؛ آنگاه، در درون هر نوع، سؤلهای به ترتیب از ساده به دشوار چیده می‌شوند. در واقع تنها فرق این روش با روش قبلی در این است که در روش قبلی سؤلهای هردسته یا هر نوع به ترتیب طبقات طبقه بندی هدفهای آموزشی، یعنی دانش، فهمیدن، کار بستن، و غیره مرتب می‌شوند، اما در روش دوم سؤلهای داخل هر نوع به ترتیب از ساده به دشوار مرتب می‌شوند.

**۳. سؤلهای را به ترتیب سازمان اصلی مطالب به دنبال هم مرتب کنید.** منظور این است که ترتیب سؤلهای باید با ترتیب مطالب درس هماهنگ باشد، یعنی به ترتیب فصلهای کتاب یا بخشهای درس که آموزش داده شده اند مرتب شوند. رعایت این نکته میتواند به ایجاد نظم فکری در آزمون شوندگان کمک کند. با این حال، گرانلاند ولین (۱۹۹۰) در این باره اظهار داشته اند که، "در ساختن آزمونهای پیشرفت تحصیلی کلاسی، مرتب کردن سؤلهای با توجه به ساختمان محتوای درس کمک چندانی نمی‌کند. اگر چنین کاری مطلوب جلوه کرد، مثلاً موقعی که می‌خواهید دوره های تاریخی را از هم جدا کنید، بکوشید تا این قسمت بندی را به حداقل برسانید" (ص ۲۳۴).

علاوه بر نکات بالا، سعی کنید سؤلهای را خوش خط و خوانا بنویسید و در بین سؤلهای فاصله‌های لازم را برای سهولت مطالعه رعایت کنید. آزمونهایی که سؤلهای آنها درهم اند و خوانا نیستند مقدار زیادی از وقت و نیروی آزمون‌شونده را به هدر می‌دهند و باعث دلسردی او می‌شوند.

در ضمن اگر سؤالی دارای شکل یا نمودار است آن را در بالای تنه سؤال قرار دهید نه در پایین گزینه‌ها. اگر شکلها و نمودارها، یا هر چیز دیگری که متن سؤال به آن اشاره دارد، پس از سؤال قرار داده شود، آزمون‌شونده ناچار خواهد بود بعد از خواندن هر گزینه به آن مراجعه نماید و این کار موجب اتلاف وقت او خواهد شد. این امر در رابطه با

قاعده ای که در فصل ۹ گفتیم، یعنی اینکه تمام مطالب مربوط به یک سؤال باید در تنه سؤال قرار داده شوند نه درگزینه ها، توضیح داده شد. همچنین به یاد داشته باشید که جای گزینه های درست در میان دیگر گزینه‌ها باید به طور تصادفی انتخاب شود و از هیچ الگوی مشخصی پیروی نکنند.

### نوشتن دستورعمل یا راهنمای آزمون

راهنمای آزمون باید مختصر و مفید باشد و به طور آشکار به آزمون شونده بگوید که چه کاری را انجام دهد. مهرنرولهمان (۱۹۸۴) گفته اند که راهنمای آزمون باید اطلاعات زیر را در اختیار آزمون شونده‌گان قرار دهد:

۱. زمان لازم برای هر قسمت

۲. ارزش هر سؤال

۳. مجازی غیرمجاز بودن حدس زدن

بجز برای دانش آموزان خردسال، راهنمای آزمون باید آن قدر آشکار باشد که آزمون شونده با خواندن آن و بدون نیاز به هیچگونه توضیح شفاهی کار خود را آغاز کند. اگر آزمون شونده‌گان نیاز به توضیح شفاهی داشته باشند، معلوم می شود که راهنمای آزمون به حد کافی روشن نبوده است و باید بازنویسی شود. اما هنگام اجرای آزمون با کودکان سالهای پایین ممکن است، علاوه بر توضیحات کتبی، راهنمایی‌های شفاهی هم ضرورت داشته باشند.

افزون بر یک راهنمایی کلی برای تمامی آزمون، برای هر بخش آزمون نیز یک راهنمایی جداگانه بنویسید (نگاه کنید به شکل ۱-۱۷). برای مسائل عددی، میزان دقتی را که انتظار دارید مشخص کنید (مثلاً چند عدد اعشاری باید محاسبه شود). همچنین واحدهای محاسبه چون ولت، اهم سانتی متر، یا میلی متر را مشخص کنید. اگر در تصحیح پاسخها به مواردی چون خوش خطی، املاي کلمه‌ها و جمله بندی نمره می‌دهید، این موارد را نیز در راهنمای سؤال بنویسید. نحوه نمره گذاری پاسخها را به اطلاع آزمون شونده‌گان برسانید. مثلاً در آزمون مربوط به مسائل ریاضی توضیح دهید که تنها جواب مسئله را می خواهید با مراحل رسیدن به جواب نیز ضروری هستند. همچنین مشخص کنید که به هر بخش چه مقدار نمره تعلق می‌گیرد.

در راهنمای آزمونهای عینی نحوه نمره گذاری پاسخها را در رابطه با حدس زدن یا حدس نزدن جوابها مشخص کنید. اگر برای حدس زدن از نمره دانش آموزان کسر میکنید، مقدار دقیق آن را معین کنید.

راهنمایی کلی: پیش از آنکه دفترچه آزمون را باز کنید، مطالب زیر را به دقت بخوانید. از آنجا که پاسخهای شما با ماشین تصحیح می شوند، لازم است محل پاسخهای خود را با مداد کاملاً سیاه کنیا- تا نمره ای را از دست ندهید.

**راهنماییهای اختصاصی:** این آزمون شامل ۷۵ سؤال چهارگزینه ای است که ارزش هر یک از آنها یک نمره است.

هر سؤال را به دقت بخوانید و ببینید کدام گزینه بهترین جواب سؤال یا کامل کننده جمله سؤال است.

۱. در پاسخنامه با خط خوانا نام خود، نام آزمون، و تاریخ آزمون را بنویسید.

۲. دقت کنید که پاسخ خود را برای هر سؤال در محل مربوط به آن در پاسخنامه علامت بزنید.

۳. از کاغذ اضافی که به شما داده شده است به عنوان پیش نویس استفاده کنید.

۴. برای هر سؤال یک پاسخ را علامت بزنید، اگر بیش از یک پاسخ برای هر سؤال انتخاب کنید به آن نمره ای تعلق نخواهد گرفت.

۵. نمره کل شما در این آزمون برابر با تعداد کل پاسخهای درست شما خواهد بود. سعی کنید به همه سؤالاها جواب بدهید، اما برای یک سؤال وقت زیادی صرف نکنید.

۶. اگر جواب سؤالی را نمیدانید حدس بزنید، نمره منفی وجود ندارد.

۷. حال می توانید شروع کنید. موفق باشید.

این آزمون شامل ۷۵ سؤال در ۸ صفحه است. دقت کنید سؤالی از نظر نیفتد.

شکل ۱-۱۷ راهنمایی یک آزمون چندگزینه ای ( اقتباس از اثر مهرنز و لهما، ۱۹۸۴)

(مثلاً توضیح دهید که در آزمونهای چهارگزینه ای برای هر سه پاسخ غلط ۱ نمره کم می‌کنید). روش نمره گذاری برای جبران حدس زدن در دنباله مطالب این فصل معرفی خواهد شد. مهرنز و الهمان (۱۹۸۴، ص ۱۸۲) می‌گویند: "ما براساس شواهد پژوهشی موجود متقاعد شده ایم که [در آزمونهای پیشرفت تحصیلی و استعداد تحصیلی باید به آزمون شوندگان گفته شود که همه سؤالاها را جواب بدهند و در تصحیح پاسخها هیچگونه نمره ای از آنها برای حدس زدن نباید کسر بشود."

### اجرای آزمون

شرایط اجرای آزمونهای پیشرفت تحصیلی باید به گونه ای باشد که از هر لحاظ آسایش جسمی و فکری آزمون شوندگان را تأمین نماید. این نکته را در رابطه با آزمون شوندگان خردسال با دقت بیشتری رعایت کنید. علاوه بر شرایط فیزیکی، شرایط روانی آزمون شوندگان نیز بر عملکرد آنان تأثیر دارد. مهم‌ترین این عوامل اضطراب زیاد است. پژوهشهای انجام شده نشان داده اند که اندکی اضطراب سبب افزایش دقت و کوشش آزمون شوندگان می شود، اما اضطراب زیاد بر عملکرد آنان تأثیر منفی دارد. بنابراین، سعی کنید که در جلسه امتحان از عوامل تنش زا و اضطراب برانگیز جلوگیری به عمل آورید.

### نمره گذاری

پس از تهیه و اجرائی آزمونهای پیشرفت تحصیلی، نوبت به تصحیح برگه های آزمون و نمره گذاری آنها میرسد. در فصل ۱۰ روش تصحیح آزمونهای تشریحی را به تفصیل توضیح دادیم، در این فصل روش صحیح آزمونهای عینی را شرح می دهیم.

تصحیح برگه های آزمون های عینی، به ویژه اگر پاسخنامه های جداگانه ای که در آن پاسخ سوالها به وسیله آزمون شوندگان علامت گذاری می شود به کار بروند، به سادگی امکانپذیر است. برای تصحیح این نوع پاسخنامه ها، ابتدا مصحح می تواند یکی از آنها را خود علامت گذاری کند و با سوراخ کردن محل پاسخهای درست پاسخنامه کلید را بسازد. بعد با قراردادن این کلید بر روی پاسخنامه های دانش آموزان و شمارش تعداد علامتهای درست، آنها را نمره گذاری کند. پیش از قراردادن پاسخنامه کلید بر روی پاسخنامه ای که باید نمره گذاری شود، دقت کنید که برای هر سؤال بیش از یک علامت وجود نداشته باشد. هر سؤالی که بیش از یک علامت برای آن منظور شده است باید حذف شود.

در نمره گذاری آزمونهای عینی دو روش عمده وجود دارد: (۱) محاسبه کلیه پاسخهای درست بدون کسر نمره برای حدس زدن؛ (۲) کسر مقداری از نمره برای جبران حدس زدن. روش اول روش متداول تر و ساده تری است. اما اگر معلم بخواهد روش دوم را مورد استفاده قرار دهد باید، هنگام اجرای آزمون، آزمون شوندگان را آگاه سازد که در ازای پاسخهای غلطی که می دهند مقداری از نمره آنان کم خواهد شد. برای تعیین این مقدار چندین روش وجود دارد که معروف ترین آنها استفاده از فرمول زیر است.

$$R = - \frac{W}{N - 1} = \text{نمره اصلاح شده}$$

در این فرمول:

R: تعداد پاسخهای درست:

w: تعداد پاسخهای غلط:

N: تعداد گزینه های هر سؤال:

فرض کنید که در یک آزمون چهار گزینه ای (N=۴) که دارای ۵۰ سؤال است، دانش آموزی به ۵ سؤال پاسخ داده است و از این تعداد ۶ سؤال غلط (W=۶) و ۳۹ سؤال درست است (R=۳۹).

نمره اصلاح شده این دانش آموز برابر است با:

$$R - \frac{W}{N - 1} = \text{نمره اصلاح شده}$$

$$\begin{aligned}
 \text{نمره اصلاح شده} &= 39 - \frac{6}{4-1} \\
 &= 39 - \frac{6}{3} \\
 &= 39 - 2 = 37
 \end{aligned}$$

### تحلیل سؤالیهای آزمون

هدف از تحلیل سؤالیهای آزمون واریسی تک تک سؤالیها و تعیین میزان دقت و نارساییهای آنهاست. در تحلیل سؤالیهای آزمون، نقاط قوت و ضعف یک آزمون و کیفیت همه سؤالیهای آن تعیین می شود. بنابراین، لازم است که معلمان، پس از اجرای هر آزمون، سؤالیهای آن را تحلیل کنند و با استفاده از نتایج حاصل به تجدید نظر در آزمون و بهبود کیفیت سؤالیها برای استفاده های بعدی اقدام نمایند.

برای تهیه یک آزمون خوب، علاوه بر رعایت پیشنهادهایی که در فصلهای پیش داده شد، تحلیل پاسخهای آزمون شوندگان و بررسی تجربی آنها ضروری است. تحلیل پاسخهای آزمون شوندگان همچنین اطلاعات تشخیصی لازم را برای بررسی کیفیت یادگیری دانش آموزان و مشکلات آموزشی معلمان فراهم می آورد.

### مراحل تحلیل سؤال

اطلاعات موردنیاز برای تحلیل سؤالیهای یک آزمون پاسخهایی هستند که آزمون شوندگان به هر سؤال داده اند. بنابراین، باید تعیین شود که در هر سؤال چند نفر گزینه درست را انتخاب کرده اند، هر یک از گزینه های انحرافی چند نفر را به خود جلب کرده است، و چند نفر آن را بی جواب گذاشته اند. برای این منظور، باید هر سؤال را در کارتی مانند شکل ۲-۱۷ وارد کنید. این کارت را می توان برای هر سؤال در روی یک صفحه کاغذ یا مقوا درست کرد و نتایج تحلیل آن سؤال را در یک طرف و صورت اصلاح شده سؤال را در طرف دیگر کارت نوشت.

اطلاعات موجود در شکل ۲-۱۷ که برای تحلیل یکی از سؤالیهای آزمون و محاسبه شاخصهای آماری آن سؤال ضروری هستند، از طریق تعیین دوگرو کرانی قوی (بالا) و ضعیف (پایین) آزمون شوندگان و تعیین نحوه پاسخدهی آنان به گزینه های مختلف سؤال به دست آمده اند. گروه بالا از ۲۵٪ تا ۳۳٪ آزمون شوندگان دارای نمره های بالا و گروه پایین از ۲۵٪ تا ۳۳٪ آزمون شوندگان دارای نمره های پایین تشکیل می یابد (کلاین<sup>۱</sup>، ۲۰۰۵). کسان دیگری از جمله رینالدزو البوینگستون، ۲۰۱۲) رقم ۲۷٪ را برای انتخاب گروه های کرانی بالا و پایین پیشنهاد داده اند. با این حال ۲۵٪ متداول ترین رقم مورد استفاده است. این اطلاعات برای تک تک سؤالیهای هر آزمون ضروری

<sup>۱</sup> . Kline

هستند. برای تعیین گروههای بالا و پایین، برگه های آزمون را به ترتیب نمره‌ای که گرفته اند از کوچک به بزرگ مرتب کنید. بعد از بالاترین نمره شروع کنید و تعداد برگه هایی را که می خواهید در گروه بالا قرار دهید انتخاب کنید. و به همین ترتیب از پایین ترین نمره شروع کنید و تعداد برگه هایی را که می خواهید در گروه پایین قرار دهید نیز برگزینید.

### عنوان آزمون: سنجش و اندازه گیری

#### موضوع سوال: پایانی

هدف یادگیری: دانشجو بتواند موارد استفاده پایایی های مختلف را تشخیص دهد.

#### سوال:

یک آزمون ۱۰۰ سوالی پیشرفت تحصیلی از ۵ خرده آزمون ۲۰ سوالی مربوط به موضوعهای درسی مختلف تشکیل شده است. اگر بخواهیم ضریب پایایی این آزمون را حساب کنیم بهترین روش کدام است؟

الف) اسپیرمن - براون

\* ب - آلفای کرونباخ

ج - KR۲۰

د - KR۲۱

گروه‌ها	الف	ب	پ	ت	بدون پاسخ
۲۵٪ بالا	۰	۵	۳	۰	۲
۲۵٪ پایین	۵	۲	۳	۰	۰
ضریب دشواری:	$P = \frac{5+2}{10+10} \times 100 = \frac{7}{20} \times 100 = 35$				
سوال نسبتاً دشوار است.					
ضریب تمیز یا تشخیص:	$P = \frac{5-2}{10} = \frac{3}{10} = 0.3$				
ضریب تمیز قدری کوچک اما قابل قبول است.					
گزینه انحرافی الف خیلی خوب عمل کرده است. گزینه ب نیاز به تجدیدنظر دارد.					
گزینه ت باید عوض شود، زیرا سوال عملاً به صورت سه گزینه‌ای درآمده است.					

شکل ۲-۱۷ کارت تحلیل سوال آزمون چند گزینه‌ای

متخصصان آزمون سازی (از جمله ویتنی<sup>۱</sup> و سیبرز<sup>۲</sup>، ۱۹۷۰؛ ونیتکو، ۱۹۸۳) پیشنهاد کرده‌اند که اگر تعداد کل دانش آموزانی که در آزمون شرکت کرده‌اند (یعنی تعداد کل برگه‌های آزمون) بیشتر از ۲۰ و ۴۰ نفر هستند، ۱۰ برگه بالا و ۱۰ برگه پایین را انتخاب کنید و کاری به بقیه برگه‌ها نداشته باشید. اگر تعداد دانش آموزان ۲۰ نفر یا کمتر است، برگه‌های آزمون را کلاً به دو دسته بالا و پایین تقسیم کنید. در شرایطی که تعداد کل آزمون شونده‌ها بیش از ۴۰ نفر است، بهترین رقم برای گروه بالا و پایین ۲۷٪ 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 (نیتکو، ۱۹۸۳). بعد از تعیین گروه‌های بالا و پایین، لازم است برای هر سؤال اطلاعات زیر را تهیه کنید.

۱. تعداد افراد گروه بالا که هر یک از گزینه‌های سؤال را انتخاب کرده یا آن را بی‌جواب گذاشته‌اند.

۲. تعداد افراد گروه پایین که هر یک از گزینه‌های سؤال را انتخاب کرده یا آن را بی‌جواب گذاشته‌اند.

پس از انجام مراحل بالا، اطلاعات موردنیاز برای تعیین شاخصهای آماری و تحلیل سؤالاها در دست هستند. این شاخصها به شرح زیر محاسبه می‌شوند.

### محاسبه ضریب دشواری سؤال

بنا به تعریف، درصد کل آزمون شونده‌گانی که به یک سوال جواب درست می‌دهند ضریب دشواری<sup>۳</sup> آن سؤال است که با حرف P نشان داده می‌شود. اگر در تحلیل یک سؤال کلیه افراد یا کلیه برگه‌های امتحانی دخالت داشته باشند برای محاسبه ضریب دشواری آن سؤال کافی است که تعداد کل افرادی را که به آن سؤال جواب درست داده‌اند بر تعداد کل آزمون شونده‌گان تقسیم کنیم و نتیجه را در ۱۰۰ ضریب نماییم. رقم حاصل ضریب دشواری سؤال است. مطابق فرمول زیر:

$$P = 100 \frac{R}{T}$$

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

$$P = 100 \frac{65}{80} = 81$$

<sup>۱</sup> . Whitney  
<sup>۲</sup> . Sabers  
<sup>۳</sup> . difficulty index



در مواردی که تعداد آزمون شوندگان (تعداد برگه‌های آزمون) زیاد است و اطلاعات ما به نحوه پاسخدهی افراد گروه بالا و گروه پایین محدود می‌شود، لازم است از فرمول زیر استفاده کنیم.

$$P = \frac{\text{انتخاب های درست} + \text{انتخاب های درست}}{\text{گروه پایین} + \text{گروه بالا}} \times 100$$

تعداد افراد گروه بالا + تعداد افراد گروه پایین  
برای محاسبه ضریب دشواری سؤال شکل ۲-۱۷ به گونه زیر عمل می‌کنیم:

$$P = \frac{5 + 2}{10 + 10} \times 100$$

$$= \frac{7}{20} \times 100$$

$$= 35$$

رقم ۳۵ نشان می‌دهد که درصد نسبتاً کمی از دو گروه بالا و پایین پاسخ درست سؤال را برگزیده‌اند. براساس اطلاعات مربوط به سؤال می‌توان نتیجه گرفت که این سوال نسبتاً دشوار است، زیرا حتی نصف گروه انتخابی (بالا و پایین) نتوانسته است به آن جواب بدهد (نگاه کنید به شکل ۲-۱۷).

براساس توضیحات بالا، هر اندازه ضریب دشواری یک سوال بزرگ تر (به ۱۰۰ نزدیک تر) باشد، آن سؤال آسان تر است و هر اندازه که این ضریب کوچک تر (به صفر نزدیک تر) باشد سؤال دشوارتر است. بنابراین، به جای ضریب دشواری می‌توان از ضریب آسانی یا سهولت<sup>۱</sup> نام برد. معمول این است که به آن ضریب دشواری بگویند.

چنان که ملاحظه شد، فرمول بالا ضریب دشواری را برحسب یک عدد صحیح دو رقمی نشان می‌دهد. برخی از متخصصان اندازه گیری و ارزشیابی ترجیح می‌دهند که این ضریب، مانند ضریب تمیز، که در قسمت بعد توضیح داده خواهد شد نیز برحسب اعداد اعشاری باشد که تفسیر آن با تفسیر ضریب تمیز همانند صورت گیرد. در این حالت، فرمول محاسبه ضریب دشواری به صورت زیر است:

$$P = \frac{\text{نسبت افراد گروه پایین که} + \text{نسبت افراد گروه بالا که}}{\text{به سؤال جواب درست داده‌اند.} + \text{به سؤال جواب درست داده‌اند.}}$$

برای محاسبه ضریب دشواری سؤال شکل ۲-۱۷ با این فرمول، به گونه زیر عمل می‌کنیم:

<sup>۱</sup> . facility index

$$P = \frac{0/5 + 0/2}{2}$$

$$= \frac{0/7}{2}$$

$$= 0/35$$

ملاحظه می‌کنید که ارقام حاصل از دو فرمول برابرند، اما در فرمول اول ضریب حاصل به صورت عدد صحیح نشان داده می‌شود در حالی که در فرمول دوم به صورت عدد اعشاری است.

### تفسیر ضریب دشواری

اگر در تفسیر ضریب دشواری یا ضریب سهولت سؤال با آزمونهای وابسته به هنجار سروکار داشته باشیم، می‌توانیم این ضریب را از دیدگاه آماری مورد بررسی قرار دهیم. یکی از انتظارات ما از آزمونهای وابسته به هنجار این است که برای آزمون شوندگان مختلف نمراتی به دست بدهد که در طول یک پیوستار پراکنده باشند، و هرچه این پراکندگی بیشتر باشد بهتر است. به سخن دیگر، هرچه واریانس نمرات حاصل از یک آزمون وابسته به هنجار بزرگ‌تر باشد آن آزمون، آزمون بهتری است. بنابراین، یکی از راههای قضاوت دربارهٔ مفید بودن سؤالهای یک آزمون آن است که بینیم 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 و (۲) همبستگیهای بین سؤالها. بنابراین، هر سؤالی که دارای واریانس بزرگ‌تری است و با سایر سؤالهای آزمون همبستگی بیشتری دارد به واریانس کل آزمون کمک بیشتری می‌کند. در مقابل، سؤالی که با سایر سؤالها همبستگی ندارد و دارای واریانس کوچکی است کمک زیادی به واریانس کل آزمون نمی‌کند. بنابراین، اگر در تحلیل آزمون به انتخاب سوالهایی پردازیم که همبستگی زیادی با یکدیگر دارند و در ضمن دارای واریانسهای بزرگ‌تری هستند، در مجموع آزمون بهتری را تدارک دیده‌ایم. اگر همهٔ سؤالهای یک آزمون با یکدیگر همبستگی کامل داشته باشند و ضریب دشواری همهٔ آنها ۰/۵ باشد، آنگاه نیمی از آزمون شوندگان در آزمون نمرهٔ صفر و نیم دیگر نمرهٔ کامل خواهند گرفت.

برای محاسبهٔ واریانس یک سؤال از فرمول زیر استفاده می‌شود:

$$P \times (1-P)$$

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 ۰ یا ۱ باشد، یعنی هیچ یک از آزمون شوندگان به آن سؤال جواب درست نداده باشد یا اینکه همهٔ آزمون شوندگان به آن جواب درست داده باشند.

$$\text{واریانس سؤال} = P \times (1-P)$$

$$= 0 \times (1-0) = 0$$

$$= 1 \times (1-1) = 0$$

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

$$\text{واریانس سؤال} = P \times (1-P)$$

$$= 0/5 \times (1-0/5)$$

$$= 0/5 \times 0/5 = 0/25$$

در نتیجه، از لحاظ انتخاب برای گنجاندن در فرم نهایی آزمون، سؤالهایی بهترند که ضریب دشواری آنها از ۱ کمتر و از صفر بیشتر و به  $0/5$  نزدیک باشد. البته انتخاب سوالهایی با ضرایب دشواری مناسب به نوع سوالهای آزمون مربوط است. در آزمونهای صحیح - غلط ضریب دشواری  $P=0/5$  زمانی به دست می آید که همهٔ آزمودنیها با حدس زدن به سؤال جواب بدهند. آلن<sup>۱</sup> و ین<sup>۲</sup> (۱۹۷۹) گفته‌اند برای آزمونهای چند گزینه‌ای سطح بهینهٔ دشواری اندکی کمتر از وسط فاصلهٔ بین ۱ و سطح موفقیت با حدس زدن است. منظور از سطح موفقیت با حدس زدن عبارت است از نسبت پاسخهای درستی که با حدس زدن به دست می‌آید. بنابراین، برای سوالهای چهار گزینه‌ای سطح موفقیت با حدس زدن حدود  $0/25$  است و لذا سطح بهینهٔ دشواری در وسط  $0/25$  و ۱ یعنی در حدود  $0/625$  خواهد بود.

برای تعیین رقم دقیق مُعرف ضریب دشواری بهینهٔ سؤال، ابتدا سطح موفقیت  $100\%$  را از سطح عملکرد با حدس زدن (مثلاً  $0/25$ ) کم می‌کنیم و نتیجه را به ۲ تقسیم می‌نماییم و نقطهٔ میانی را مشخص می‌کنیم.

$$\text{نقطهٔ میانی} = \frac{1/00 - 0/25}{2} = \frac{0/75}{2} = 0/375$$

پس از آن، رقم حاصل از محاسبات بالا (یعنی نقطهٔ میانی) را با سطح مُعرف عملکرد از طریق حدس زدن جمع می‌کنیم.

نقطه میانی	عملکرد با حدس زدن	ضریب دشواری بهینه
$0/375$	+	$0/25$
		= $0/625$

راه ساده تر برای به دست آوردن همان نتیجه این است که عدد ۱ را با عملکرد مبتنی بر حدس زدن جمع کنیم و نتیجه را بر ۲ تقسیم نماییم.

<sup>۱</sup> . Allen  
<sup>۲</sup> . Yen

$$\frac{1+0/25}{2} = 0/625$$

به طور کلی، ضریبهای دشواری بین ۰/۳ تا ۰/۷ حداکثر اطلاع را درباره تفاوت بین آزمودنیها به دست می‌دهند. با این حال، در شرایط ویژه، ضریب دشواری فرق می‌کند. مثلاً آزمونی که به صورت آزمون ورودی فوق لیسانس (کارشناسی ارشد) یا دکتری برای انتخاب بهترین داوطلبان به کار می‌رود و حدوداً ۱۰٪ بالای داوطلبان را برمی‌گزینند باید شامل سؤالهای بسیار دشوار باشد. در مقابل، آزمونی که برای انتخاب کودکان برای یک دوره آموزش ترمیمی یا اصلاحی مورد استفاده قرار می‌گیرد باید شامل سؤالهای بسیار آسان باشد. حتی در آزمونهای دشوار نیز آزمون ساز ممکن است ترجیح دهد که در آغاز آزمون چند سؤال آسان برای تشویق آزمون شوندگان قرار دهد.

### محاسبه ضریب تمیز سؤال

برخلاف ضریب دشواری که میزان آسان بودن یا دشوار بودن یک سؤال را برای گروه آزمون شوندگان نشان می‌دهد، ضریب تمیز<sup>۱</sup> که با D نشان داده می‌شود قدرت سؤال را در تمایزگذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان مشخص می‌کند؛ یعنی معلوم می‌نماید که سوال تا چه اندازه می‌تواند گروه قوی را از گروه ضعیف جدا سازد. برای محاسبه ضریب تمیز یک سؤال از فرمول زیر استفاده می‌شود:

$$D = \frac{\text{انتخابهای درست گروه پایین} - \text{انتخابهای درست گروه بالا}}{\text{ضریب تمیز سؤال}} =$$

تعداد افراد یک گروه (بالا یا پایین)

ضریب تمیز سؤال شکل ۲-۱۷ به شرح زیر محاسبه می‌شود:

$$D = \frac{5-2}{10} = \frac{3}{10} = 0/3$$

به جای فرمول بالا، می‌توان از فرمول زیر نیز برای محاسبه ضریب تمیز سؤال استفاده کرد:

$$D = \frac{\text{درصد کسانی از گروه پایین که به سؤال جواب درست داده‌اند} - \text{درصد کسانی از گروه پایین که به سؤال جواب درست نداده‌اند}}{2}$$

<sup>۱</sup> . discriminative index

طبق این فرمول ضریب تمیز سؤال شکل ۲-۱۷ برابر خواهد بود با:

$$D = 0.5 - 0.2 = 0.3$$

رقم ۰/۳ نشان می‌دهد که سؤال تا اندازه‌ای توانسته است گروه بالا و گروه پایین را از یکدیگر متمایز کند. یعنی از ۱۰ نفر گروه بالا ۵ نفر به سؤال جواب درست داده‌اند، درحالی که تنها ۲ نفر از ۱۰ نفر گروه پایین به این سؤال جواب درست داده‌اند.

### 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 سؤالی ۰/۹۰ باشد، آن سؤال آزمون شوندگان قوی و ضعیف را خیلی خوب از هم جدا خواهد کرد؛ اما اگر ضریب تمیز سؤالی ۰/۱۰ باشد، آن سوال از عهده جداسازی دانش آموزان قوی و ضعیف به خوبی برنخواهد آمد. ضریب تمیز صفر حاکی از این است که آن سؤال به هیچ وقت نتوانسته بین گروه قوی و ضعیف تمایز قائل شود.

گاه اتفاق می‌افتد که ضریب تمیز سؤالی منفی است. این نوع ضریب تمیز نشان می‌دهد که در آن سوال گروه قوی بدتر از گروه ضعیف عمل کرده است. این گونه سوالها 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 می‌تواند این باشد که افراد گروه بالا با آن سؤال را به طور کامل یاد نگرفته اند یا آن را به غلط آموخته اند. بنابراین، تمامی ضریبهای تمیز سوالهای آزمون باید مثبت باشند.

رینالدز و لیوینگستون (۲۰۱۲) درجات مختلف ضرایب تمیز را به شرح زیر ارزشیابی کرده و گفته اند "به عنوان یک قاعده کلی، ما سؤال های با ضریب تمیز ۰/۳۰ به بالا را پیشنهاد می‌کنیم (هرچه بزرگ تر بهتر)، و سؤال های دارای ضریب تمیز زیر ۰/۳۰ باید مورد واریسی دقیق قرار گیرند و اصلاح یا حذف شوند."

### ضریب تمیز

۰/۴۰ یا بزرگ تر	عالی
۰/۳۰-۰/۳۹	خوب
۰/۱۱-۰/۲۹	متوسط
۰/۰۰-۰/۱۰	ضعیف
منفی	معیوب

لازم به گفتن است که وجود ضریب تمیز منفی در آزمونهای توانایی نامطلوب است، اما "در ابزارهای سنجش شخصیت، علاقه، و نگرش ضریب تمیز منفی مشکل زا نیست. در این گونه آزمونها، تمیز دادن بین نسخه‌ها یا گروههای مختلف کاری مطلوب است و سؤلهای دارای ضریب تمیز بالا چه مثبت و چه منفی) در تمایز قائل شدن بین آنها کمک می‌کنند" (کلاین، ۲۰۰۵، ص ۹۸).

### همبستگی سؤال با کل آزمون

براساس توضیحات بالا، می‌توان دید که ضریب تمیز سؤال نوعی ضریب توافق سؤال با کل آزمون است. سؤالی را در نظر بگیرید که در آن نمره‌های افرادی که در کل آزمون نمره‌های بالایی گرفته‌اند بالا است و کسانی که در کل آزمون نمره‌های پایینی گرفته‌اند در آن سؤال نمره‌های شان پایین است. این سال با کل آزمون همخوانی دارد و ضریب تمیز آن مثبت و بالاست. در مقابل، سؤال دیگری را در نظر بگیرید که در آن افراد قوی (یعنی کسانی که در کل آزمون نمرات بالا گرفته‌اند) نمرات ضعیفی کسب می‌داند و افراد ضعیف (یعنی کسانی که در کل آزمون نمره‌های بالایی نگرفته‌اند) نمرات‌شان بالاست. این سؤال دارای ضریب تمیز منفی است. بنابر آنچه گفته شد، سؤلهای دارای قوه تمیز خوب با کل آزمون همبسته‌اند، اما سؤلهای دارای قوه تمیز صفر یا منفی یا باکل آزمون همبسته نیستند یا با آن همبستگی منفی دارند.

پس یکی دیگر از راههای محاسبه ضریب تمیز سؤلهای یک آزمون محاسبه ضریب همبستگی بین نمرات آزمون شوندگان در آن سؤالا و نمرات آنان در کل آزمون است. این ضریب همبستگی که معمولاً به صورت همبستگی دورشته‌ای نقطه‌ای<sup>۱</sup> محاسبه می‌شود در واقع ضریب توافق سؤال با کل آزمون است.<sup>۲</sup> هرچه این ضریب همبستگی بیشتر باشد، بدان معنی است که سؤال با مجموعه سؤلهای دیگر همخوانی بیشتری دارد، و لذا بین افراد قوی و ضعیف بیشتر تمیز قائل می‌شود؛ و هرچه این ضریب کمتر باشد مقدار این همخوانی کمتر است، و لذا آن سؤال بین افراد قوی و ضعیف کمتر تمیز قائل می‌شود. اگر ضریب همبستگی سؤالی باکل آزمون منفی باشد، بدان معنی است که سؤال باکل آزمون ناهمخوانی دارد، یعنی چیزی را می‌سنجد که با آنچه بقیه سؤلهای آزمون می‌سنجد متفاوت است، و لذا بین افراد قوی و ضعیف به طور منفی تمیز قائل می‌شود. و بالاخره اگر ضریب همبستگی سؤالی با کل آزمون صفر باشد، بدان معنی است که این سؤال نمی‌تواند بین افراد قوی و ضعیف تمیز قائل شود.

از امتیازهای استفاده از ضریب همبستگی بین سؤال و کل آزمون به عنوان ضریب تمیز، یکی این است که میتوان این ضریب را از لحاظ معنی داری آماری آزمون کرد. دوم اینکه، از لحاظ عملی می‌توان معنی‌داری این ضریب را مورد قضاوت قرار داد. برای مثال، اگر ضریب همبستگی بین یک سؤال و کل آزمون  $+0/40$  باشد، می‌دانیم که این سوال

<sup>۱</sup> point biserial correlation

<sup>۲</sup> برای چگونگی محاسبه این ضریب همبستگی، نگاه کنید به فصل ۱۸ کتاب.

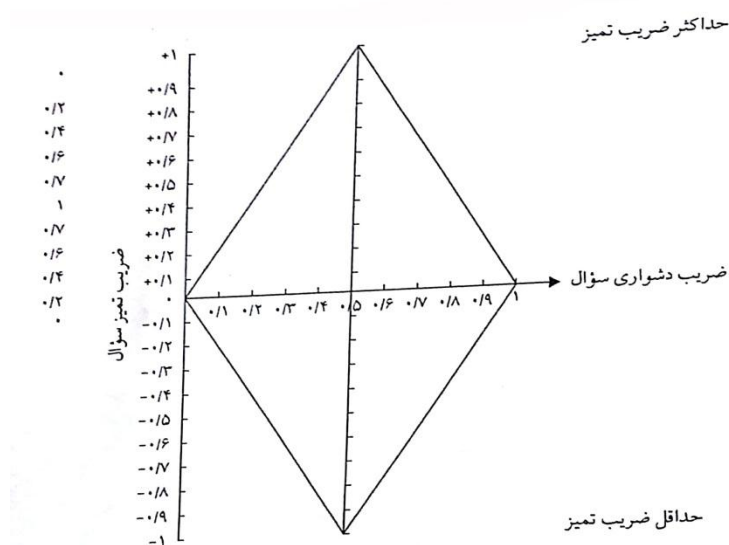
۱۶ درصد ( $4^2=16$ ) واریانس نمرات را به خود اختصاص می دهد. در حالی که ضریب تمیز به دست آمده از فرمول محاسبه ضریب تمیز با توجه به گروه های بالا و پایین را نمی توان این گونه تفسیر کرد. از اینها گذشته، چنانکه در فصل ۲۱ خواهیم دید، همبستگی بین سؤال و آزمون مستقیماً به پایایی آزمون مربوط است.

### رابطه بین ضریب دشواری و ضریب تمیز

شکل ۳-۱۷ رابطه بین ضریب دشواری و ضریب تمیز سؤال را نشان می دهد. حدود چهارضلعی این شکل حداکثر و حداقل ضرایب دشواری و تمیز را نشان می دهد. برای مثال، اگر ضریب دشواری برابر  $0/8$  باشد، ضریب تمیز بین  $0/4+$  و  $0/4-$  خواهد بود. دقت کنید که تنها در صورتی ضریب تمیز به حداکثر یا  $+1$  می رسد که ضریب دشواری برابر  $0/5$  یعنی در حد وسط باشد. در نتیجه سؤالهای خوب یک آزمون سؤالهایی هستند با ضریب دشواری متوسط و ضریب تمیز بزرگ.

### تحلیل گزینه های انحرافی

در تحلیل سؤالهای آزمون، علاوه بر تعیین ضریبهای دشواری و تمیز برای هر سؤال، بررسی نحوه پراکندگی پاسخهای مربوط به گزینه های انحرافی هر سؤال نیز ضروری است. پیشتر در این کتاب گفتیم که هدف از قراردادن گزینه های انحرافی در سوالها منحرف کردن آزمون شوندگانی است که جواب درست سوال را نمی دانند. بنابراین گزینه های انحرافی خوب باید عملاً از عهده این کار برآیند. برای تعیین این مطلب باید به گزینه هایی که توسط آزمون شوندگان انتخاب شده اند مراجعه کنیم.



شکل ۳-۱۷: رابطه ضریب تمیز با ضریب دشواری

به منظور بررسی گزینه‌های انحرافی سوال شکل ۲-۱۷، با مراجعه به اطلاعات داده شده معلوم می‌شود که گزینه انحرافی الف گزینه انحرافی بسیار خوبی است، زیرا از گروه قوی هیچ کس آن را انتخاب نکرده است اما از گروه ضعیف ۵ نفر آن را برگزیده‌اند. گزینه‌های انحرافی پ و ت کارشان را به خوبی انجام نداده‌اند. گزینه ت به کلی بی‌فایده است و بود و نبود آن فرقی نمی‌کند، زیرا هیچ یک از افراد دو گروه را به خود جلب نکرده است. گزینه پ نیز برای هر دو گروه یکسان عمل کرده است. در تجدیدنظر و اصلاح این سوال به منظور استفاده‌های بعدی، لازم است گزینه ت با گزینه تازه‌ای عوض شود و گزینه پ نیز اصلاح گردد.

در تحلیل گزینه‌های انحرافی قاعده کلی به شرح زیر است: هر گزینه انحرافی دست کم باید یک نفر از افراد گروه ضعیف را به خود جلب کند و اگر هر دو گروه قوی و ضعیف را به خود جلب می‌کنند تعداد افراد گروه ضعیف باید بیشتر از تعداد افراد گروه قوی باشد. دلیل این امر را این گونه می‌توان توضیح داد که افراد گروه ضعیف در مجموع اطلاعات کمتری از مطالب آزمون دارند. بنابراین صورتی یک سوال به خوبی عمل می‌کند که افراد ضعیف بیشتر از افراد گروه قوی گزینه‌های انحرافی آن سوال را انتخاب نمایند.

### تحلیل سوالهای آزمونهای تشریحی و آزمونهای عملکردی

هرچند که تحلیل سؤال بیشتر برای آزمونهای عینی متداول است، با این حال سؤالهای غیرعینی را نیز پس از اجرا برای مورد تجزیه و تحلیل قرار داد و اشکالهای آنها را برطرف کرد. ما در اینجا دو روش معروف تحلیل آزمونهای غیرعینی را توضیح می‌دهیم.

یکی از روشهای تحلیل سؤالهای آزمون تشریحی و عملکردی روش ویتنی و سیبرز (۱۹۷۰) است. این روش را همراه با یک مثال در زیر توضیح می‌دهیم.

۱. دو گروه ۲۵ درصدی بالا و پایین از کل دانش آموزان را به شرحی که قبلاً گذشت تعیین کنید.

۲. برای هر سؤال مجموع نمرات گروه بالا و گروه پایین را حساب کنید.

۳. برای محاسبه ضریبهای دشواری و تمیز نتایج را در فرمولهای زیر قرار دهید.

$$P = \frac{\sum_H + \sum_L - (2N \text{ Score}_{\min})}{2N(\text{Score}_{\max} - \text{Score}_{\min})}$$

ضریب دشواری =

$$D = \frac{\sum_H + \sum_L}{N(\text{Score}_{\max} - \text{Score}_{\min})}$$

ضریب تمیز =

در این فرمولها:

$\sum_H$  = مجموع نمرات های افراد یا برگه های گروه بالا (۲۵ درصد)



$\sum_L$  = مجموع نمره های افراد یا برگه های گروه پایین (۲۵ درصد)

$N$  = تعداد افراد یا برگه های ۲۵ درصد گروه موردنظر

$Score_{mix}$  = بالاترین نمره ممکن در سوال

$Score_{min}$  = پایین ترین نمره ممکن در سوال

مثال:

برای یکی از سوالهای یک آزمون تشریحی اطلاعات جدول ۱-۱۷ به دست آمده است.

با استفاده از اطلاعات بالا می توان ضریبهای دشواری و تمیز سؤال فوق را به روش زیر محاسبه کرد.

$$P = \text{ضریب دشواری} = \frac{93 + 81 - (2 \times 30 \times 0)}{2 \times 30 (5 - 0)} = \frac{174}{30} = 0/58$$

$$D = \text{ضریب تمیز} = \frac{93 - 81}{30(5 - 0)} = \frac{12}{150} = 0/08$$

جدول ۱-۱۷ اطلاعات موردنیاز برای تحلیل سؤال تشریحی

گروه پایین		گروه بالا		نمره سوال
fs	f	fs	f	s
۱۵	۳	۴۰	۸	۵
۲۸	۷	۲۰	۵	۴
۲۴	۸	۱۸	۶	۳
۱۲	۶	۸	۴	۲
۲	۲	۷	۷	۱
۰	۴	۰	۰	۰
۸۱	۳۰	۹۳	۳۰	

علاوه بر روش بالا، نیتکو (۲۰۰۱) نیز، برای تحلیل سؤال آزمونهای تشریحی و عملکردی، روش دیگری را معرفی

کرده که ما در اینجا به توضیح آن می پردازیم. در این روش، ضریب دشواری یک سؤال تشریحی یا عملکردی به

صورت نمره میانگین گروه آزمون شوندگان برای آن بخش بر دامنه ممکن نمرات سوال تعریف شده است، طبق فرمول زیر:

$$P = \text{ضریب دشواری} =$$

برای آشنا شدن با چگونگی محاسبه ضریب دشواری سوالهای تشریحی و عملکردی با روش نیتکو (۲۰۰۱) فرض کنید سوال شماره ۱ یک آزمون دارای ارزش از ۱ تا ۶ نمره است. اگر نمره میانگین این سؤال ۴/۲ باشد، آنگاه ضریب دشواری این سؤال برابر است با:

$$P = \frac{4/2}{6-1} = \frac{4/2}{5} = 0/84$$

در تفسیر ضریب دشواری  $P = 0/84$ ، می‌گوییم که به طور متوسط پاسخ دهندگان ۸۴٪ بالاترین نمره ممکن را در سوال به دست آورده‌اند (که البته حاکی از یک سوال نسبتاً ساده است).

برای محاسبه ضریب تمیز سوالهای آزمونهای عملکردی و تشریحی با روش نیتکو (۲۰۰۱)، تفاوت بین نمره میانگین گروه بالا و نمره میانگین گروه پایین را حساب می‌کنیم و نتیجه را بر دامنه ممکن نمرات سوال تقسیم می‌کنیم، طبق فرمول زیر:

$$D =$$

نمره میانگین گروه پایین - نمره میانگین گروه بالا  
 گروه بالا و گروه پایین همان گروه‌هایی هستند که قبلاً معرفی شدند، یعنی کسانی که در آزمون بیشترین نمرات (گروه بالا) و کسانی که کمترین نمرات (گروه پایین) را به دست آورده‌اند.  
 برای آشنا شدن با چگونگی محاسبه ضریب تمیز سؤال، فرض کنید میانگین نمرات گروه بالا برای یک سؤال ۵/۳ و میانگین نمرات گروه پایین برای همین سؤال ۲/۸ است. همچنین فرض کنید که دامنه نمرات ممکن برای این سؤال از ۱ تا ۶ است، ضریب تمیز این سؤال برابر است با:

$$D = \frac{5/3 - 2/8}{6-1} = \frac{2/5}{5} = 0/5$$

از آنجا که تفاوت بین نمرات میانگین گروه‌های بالا و پایین را به دامنه ممکن نمرات سؤال تقسیم می‌کنیم، می‌توانیم نتیجه حاصل را این گونه تفسیر نماییم که تفاوت بین نمرات میانگین گروه‌های بالا و پایین برای سوال فوق ۵۰٪ دامنه سوال است. این رقم ضریب تمیز خوبی است.

**تجدیدنظر و اصلاح آزمون**

پس از تعیین ضریبهای دشواری و تمیز همهٔ سوالهای یک آزمون و بررسی گزینه‌های هر سؤال، باید به اصلاح سؤالها و بازنویسی آنهايي که نیاز به تغییر دارند اقدام کرد. سؤالهایی که ضریب دشواری آنها بسیار بالا یا بسیار پایین است یا ضریب تمیز آنها خیلی کم است باید مورد تجدیدنظر قرار گیرند. همچنین گزینه‌های انحرافی معیوب نیز باید اصلاح یا عوض شوند.

### تحلیل سؤال‌های آزمونهای وابسته به ملاک

لازم به ذکر است که استفاده از ضریبهای دشواری و تمیز به شرحی که گذشت صرفاً برای آزمونهای وابسته به هنجار مناسب است، اما در آزمونهای وابسته به ملاک این ضریبها ورد استفاده‌ای ندارند. چنان که قبلاً در این کتاب گفته شد، نتیجهٔ آرمانی آزمونهای وابسته به ملاک این است که در آنها اکثر آزمون شوندگان اکثر سوالها را درست جواب دهند. در چنین حالتی ضریب دشواری سوالها بسیار بزرگ و ضریب تمیز آنها بسیار کوچک خواهد بود، یعنی برای گروه آزمودنی موردنظر اکثر سوالها آسان و قدرت تمیز آنها بسیار کم خواهد بود.

### تمیز بین پیش و پس از آموزش

بعضی از متخصصان ارزشیابی پیشرفت تحصیلی پیشنهاد می‌کنند که آزمونهای وابسته به ملاک مورد استفاده در سنجش یادگیری در حد تسلط تنها باید شامل سؤالهایی باشند که پس از آموزش مطالب درسی آسان‌تر از پیش از آموزش آن مطالب جلوه کنند. یعنی اینکه نشان دهنده سوالهای امتحان تا چه اندازه تاثیر آموزش را می‌سنجند. برای تحقق این امر انجام مراحل زیر پیشنهاد شده است.

۱. اجرای آزمون پیش از آموزش و پس از آن

۲. محاسبهٔ ضریب دشواری هر سال برای پیش آزمون و پس آزمون

۳. محاسبهٔ تفاوت ضرایب دشواری سوالهای پیش آزمون و پس آزمون

این ضریب تمیز را ضریب تمیز پیش و پس از آموزش<sup>۱</sup> می‌نامند و آن را یا  $D$  نشان می‌دهند (ویرسما و جورس، ۱۹۹۰). نام دیگر این ضریب، ضریب حساسیت نسبت به آثار آموزش<sup>۲</sup> است و آن را با  $S$  معرفی می‌کنند. این ضریب با فرمول زیر محاسبه می‌شود:

$$S \text{ یا } D = \frac{R_A - R_B}{T}$$

در این فرمول:

$R_A = RA$  تعداد کسانی که پس از آموزش به سوال جواب درست داده‌اند.

<sup>۱</sup> . pre and post discrimination

<sup>۲</sup> . sensitivity to instructional effects

تعداد کسانی که پیش از آموزش به سوال جواب درست داده‌اند.  $R_B = RB$

تعداد کسانی که هم پیش و هم پس از آموزش به سوال جواب داده‌اند،  $T =$

یعنی تعداد کل افراد

برای اینکه با چگونگی استفاده از این فرمول و محاسبه ضریب D یا S آشنا شوید، از داده‌های جدول ۲-۱۷ استفاده می‌کنیم. داده‌های این جدول نمرات ۶ دانش آموز در یک آزمون ۵ سؤالی هم پیش از آموزش (پیش آزمون) و هم پس از آموزش (پس آزمون) است. عدد ۱ نشان دهنده این است که دانش آموز سوال را درست جواب داده است و عدد صفر معرف این است که دانش آموز سوال را درست جواب نداده است. حرف B معرف نمره پیش آزمون (پیش از آموزش) و حرف A نشان دهنده نمره پس از آزمون (پس از آموزش) است.

جدول ۲-۱۷ عملکرد ۶ دانش آموز در یک آزمون هم پیش از آموزش (B) و هم پس از آموزش (A) (گرانلاند و لین، ۱۹۹۰)

سؤالاها		۱		۲		۳		۴		۵	
پیش از آزمون (B)											
پیش از آزمون (A)		A	B	A	B	A	B	A	B	A	B
دانش آموزان	الف	۰	۱	۱	۰	۰	۱	۱	۰	۱	۱
	ب	۰	۱	۱	۰	۰	۱	۱	۰	۱	۱
	پ	۰	۱	۱	۰	۰	۱	۱	۰	۱	۱
	ت	۰	۱	۱	۰	۰	۱	۱	۰	۱	۱
	ث	۰	۱	۱	۰	۰	۱	۱	۰	۱	۱
	ج	۰	۱	۱	۰	۰	۱	۱	۰	۱	۱

با استفاده از داده‌های جداول ۲-۱۷ و قراردادن آنها در فرمول ضریب پیش و پس از آموزش، برای سوالهای این جدول ضرایب D یا S به نحو زیر به دست می‌آیند.

$$d = \frac{6-0}{6} = 1 \text{ یا } S \text{ سؤال } 1$$

$$D = \frac{6-6}{6} = 0 \text{ یا } S \text{ سؤال } 2$$

$$D = \frac{0-0}{6} = 0 \text{ یا } S \text{ سؤال } 3$$

$$D = \frac{0-6}{6} = 1 \text{ یا } S \text{ سؤال } 4$$

$$D = \frac{5-2}{6} = 0.5 \text{ یا } S \text{ سؤال } 5$$

با توجه به داده‌های جدول ۲-۱۷ و ضرایب تمیز محاسبه شده برای سوالها، می‌توان هر سوال را در ارتباط با اثربخشی آن به عنوان وسیله‌ای برای قضاوت درباره آثار آموزش به نحو زیر مورد بحث قرار داد (گرانلاند و لین، ۱۹۹۰).

سؤال ۱: سؤال بسیار خوبی برای یک آزمون وابسته به ملاک مربوط به یادگیری در حد تسلط است، زیرا همه دانش آموزان پیش از آموزش آن را غلط جواب دادند اما پس از آموزش به آن جواب درست دادند.

سؤال ۲. این سوال بسیار آسان است و نتوانسته آثار آموزش را بسنجد، زیرا همه دانش آموزان پیش و پس از آموزش به آن جواب درست داده‌اند.

سؤال ۳. این سوال خیلی دشوار است و نتوانسته است آثار آموزش را اندازه بگیرد، یا اینکه آموزش در رابطه با موضوع این سوال ناکافی بوده است.

سؤال ۴. این یک سوال بسیار معیوب است یا یک سوال آسان است که با آموزش غلط دنبال شده است.

سؤال ۵. این یک سوال کارآمد است با یک الگوی نسبتاً طبیعی، زیرا بعضی از دانش آموزان پیش از آموزش به آن جواب درست داده‌اند، اما تعداد بیشتری از دانش آموزان پس از آموزش به آن جواب درست داده‌اند.

مانند ضریب تمیز آموزنهای وابسته به هنجار، در اینجا نیز ضریب تمیز از -۱ تا +۱ در تغییر است. سؤالهای دارای ضریبهای تمیز مثبت به کیفیت خوب کل آزمون کمک میکنند. وقتی که ضریب تمیز سؤالی صفر یا نزدیک به صفر است نشان میدهد که سؤال به قدرت تمیز کل آزمون هیچ کمکی نمی کند و آن سؤالی است که همه آزمون شوندگان در پیش آزمون و پس آزمون به آن جواب درست یا غلط داده اند.

شرایط آرمانی برای هر سؤال آن است که تفاوت ضریبهای دشواری قبل و بعد از آموزش آن (ضریب حساسیت نسبت به آثار آموزش) +۱ باشد، یعنی ضریب دشواری سؤال در پیش آزمون صفر و ضریب دشواری آن در پس آزمون +۱ باشد، مانند سؤال شماره ۱ جدول ۲-۱۷. این بدان معنی است که پیش از آموزش هیچ یک از دانش آموزان جواب سؤال را نمی دانسته است، اما در نتیجه آموزش همه دانش آموزان جواب آن سؤال را یاد گرفته اند. بنابراین،

برای استفاده های بعدی سؤلهایی که دارای تفاوت ضریبهای دشواری قبلی و بعدی زیاد یا نزدیک به رقم ۱+ هستند باید انتخاب گردند.

با وجود این، نیتکو (۱۹۸۳) معتقد است که وقتی هدف از تهیه یک آزمون تفسیر پیشرفت مطلق آزمون شوندگان است، استفاده از شاخص بالا (تفاوت بین ضریبهای دشواری قبل و بعد از آموزش) کار مفیدی نیست. به اعتقاد او، از آنجا که محتوای سؤلهای در حوزه مورد نظر مهم است، انتخاب تعدادی از سؤلهای و کنار گذاشتن تعدادی دیگر مجموعه سؤلهایی را که معرف محتوای مورد نظر باشد به دست نمی دهد. همچنین تنها انتخاب سؤلهایی که از نظر تفاوت ضریب دشواری قبل و بعد از آموزش مثبت هستند سبب انتخاب آن دسته از سؤلهایی می شود که آموزش برای آنها اثربخش بوده است. در چنین حالتی، هیچ راهی برای ما وجود ندارد که تعیین کنیم: (۱) چه بخشی از محتوای درسی را دانش آموزان نیاموخته اند (یعنی موضوع سؤلهایی که تفاوت ضرایب دشواری قبلی و بعدی آنها صفر است)، یا (۲) آموزش چه بخشی از محتوای درس را به حالتی درآورده است که دانش آموزان نتوانند به سؤلهای آن جواب درست بدهند (یعنی موضوع سؤلهایی که تفاوت ضرایب دشواری قبل و بعد از آموزش آنها کمتر از صفر است).

#### تمیز بین حد تسلط و غیر حد تسلط

از آنجا که آزمونهای وابسته به ملاک دارای یک نمره مرزی برای حد تسلط هستند، مثلاً نمره ۸۵ از ۱۰۰، بنابراین لازم است معلوم شود که آیا سؤلهای آزمون در آن نقطه بین آزمون شوندگان قرعه ضیف تمیز قائل می شوند یا نه. برنان<sup>۱</sup> (۱۹۷۲) یک ضریب حد تسلط به غیر حد تسلط به دست داده که برای منظور فعلی ما مفید است. برای محاسبه این ضریب یک شکل چهارخانه ای شبیه شکل ۴-۱۷ درست می شود که در خانه های عمودی آن حد تسلط و غیر حد تسلط و در خانه های افقی آن درست و غلط قرار داده می شوند.

**آزمون**

حد تسلط	غیر حد تسلط	
۶	۴	
۱	۶	درست
۳	۰	سؤال غلط

شکل ۵-۱۷ اطلاعات مورد نیاز برای محاسبه ضریب تمیز حد تسلط - غیر حد تسلط

<sup>۱</sup> . brennan

در شکل ۴-۱۷، حرف a معرف تعداد آزمون شوندگانی است که سوال را درست جواب داده‌اند اما در کل آزمون نمره‌ای کمتر از حد تسلط گرفته‌اند. همچنین b تعداد آزمون شوندگانی است که سوال را درست جواب داده‌اند و در نمره کل آزمون نیز به حد تسلط رسیده‌اند. حرف C نشان دهنده تعداد آزمون شوندگانی است که سوال را غلط جواب داده‌اند و در کل به حد تسلط نرسیده‌اند. و بالاخره حرف d معرف تعداد کسانی است که سوال موردنظر را درست جواب نداده‌اند اما در نمره کل آزمون به حد تسلط رسیده‌اند. تعداد آزمون شوندگانی که به حد تسلط رسیده‌اند برابر است با b+d و تعداد کسانی که به حد تسلط نرسیده‌اند مساوی است با a+c.

ضریب تمیز، یعنی D، برابر است با تفاوت بین سطوح دشواری سوال برای کسانی که به حد تسلط رسیده و آنهایی که به حد تسلط نرسیده‌اند، مطابق با فرمول زیر:

$$D = \frac{b}{b+d} = \frac{a}{a+c}$$

تمیز بین حد تسلط و غیر حد تسلط

این ضریب تمیز از -۱ تا +۱ متغیر است. مقدار مثبت D نشان می‌دهد که سوال بین کسانی که به حد تسلط رسیده و کسانی که به حد تسلط نرسیده‌اند تمیز قائل شده است. وقتی که ضریب تمیز سوالی منفی است، نشان دهنده آن است که سوال در جهت خلاف سوالهای دیگر تمیز قائل می‌شود. چنین سوالی نیاز به بازبینی دارد. ممکن است معلم، پس از بازبینی، آن را به کلی کنار بگذارد. ممکن است لازم باشد سوال را با افراد کلاس در میان بگذارد تا درباره آن بحث کنند و علت مشکل را بیابند. ممکن است کلید تصحیح اشتباه بوده باشد. در هر صورت، ضریب تمیز منفی نشان دهنده مشکلی در سوال است که باید مورد توجه و رسیدگی قرار گیرد.

### مثالی برای تحلیل سؤال یک آزمون وابسته به ملاک

برای روشن تر ساختن نکات بالا، مثال کاملی برای تحلیل یک آزمون ملاکی توضیح می‌دهیم، فرض کنید اطلاعات جدول ۳-۱۷ نمرات حاصل از یک آزمون ۸ سوالی وابسته به ملاک است که با ۱۰ نفر دانش آموز اجرا شده است. در این جدول، نمره ۱ نشان دهنده این است که دانش آموز سوال را درست جواب داده است و نمره صفر معرف جواب غلط دانش آموز است. نمره کل دانش آموز برابر است با مجموع جوابهای درست.

در پایین جدول ۳-۱۷ ضریبهای دشواری و تمیز همه سوالها محاسبه شده‌اند. چنان که قبلاً گفتیم، ضریب دشواری عبارت است از درصد کل دانش آموزانی که سوال را درست جواب داده‌اند. در مثال بالا، همه دانش آموزان سوال شماره ۱ را درست جواب داده‌اند، بنابراین ضریب P آن ۱ است. از ۱۰ نفر دانش آموز، ۹ نفر سؤال شماره ۲ را درست جواب داده‌اند؛ بنابراین P آن برابر است با ۰/۹، بقیه ضریبهای دشواری نیز به همین ترتیب حساب شده‌اند. دقت کنید که ضریب دشواری سوال شماره ۶ بسیار کوچک تر از بقیه سوالهاست. این سوال باید مورد بازبینی قرار

گیرد تا علت دشواری بیش از اندازه آن معلوم گردد. شاید محتوای آن با محتوای بقیه سواها قدری متفاوت است یا شاید اشکال دیگری دارد. می توان از دانش آموزان سوال کرد که چرا با آن سوال بیشتر از بقیه مشکل داشته اید. حال فرض کنید که این آزمون دارای یک حد تسلط ۹۰ از صد بوده است، یعنی از پیش به دانش آموزان گفته شده است که باید در آزمون دست کم ۹۰ نمره از ۱۰۰ نمره را دریافت کنند یا ۹۰ درصد ۸ سوال را درست جواب دهند. بنابراین مرز حد تسلط، دانش آموزانی که دارای نمره ۷ یا بیشتر بوده اند به حد تسلط رسیده اند. براساس اطلاعات موجود در جدول ۳-۱۷، از ۱۰ نفر دانش آموز ۶ نفر به این حد از تسلط دست یافته اند.

جدول ۴-۱۷ نمرات مربوط به سواهای یک آزمون وابسته به ملاک (ویرسما و جورس، ۱۹۹۰، ص ۲۴۷)

سواها									
نمره	۸	۷	۶	۵	۴	۳	۲	۱	دانش آموزان
کل									
۷	۱	۱	۰	۱	۱	۱	۱	۱	الف
۴	۰	۰	۰	۱	۱	۰	۱	۱	ب
۸	۱	۱	۱	۱	۱	۱	۱	۱	پ
۷	۱	۱	۱	۰	۱	۱	۱	۱	ت
۴	۰	۰	۰	۱	۱	۱	۱	۱	ث
۵	۱	۱	۰	۰	۱	۰	۱	۱	ج
۸	۱	۱	۱	۱	۱	۱	۱	۱	چ
۳	۰	۱	۰	۱	۰	۰	۱	۱	ح
۸	۱	۰	۱	۱	۱	۱	۱	۱	خ
۷	۱	۱	۰	۱	۱	۱	۱	۱	د
	۰/۷	۰/۷	۰/۴	۰/۸	۰/۹	۰/۷	۰/۹	۱/۰	P= ضریب دشواری
	۰/۷۵	۰/۷۵	۰/۶۷	۰/۰۸	۰/۲۵	۰/۷۵	۰/۲۵	۰/۰	D= ضریب تمیز

حال به منظور اینکه ضریب تمیز بین حد تسلط و غیرحد تسلط را برای سواها محاسبه کنیم، اطلاعات مورد نیاز برای این محاسبه را طبق شکل ۴-۱۷ برای هر یک از سواها در فرمول زیر قرار می دهیم.



$$D = \frac{b}{b+d} = \frac{a}{a+c}$$

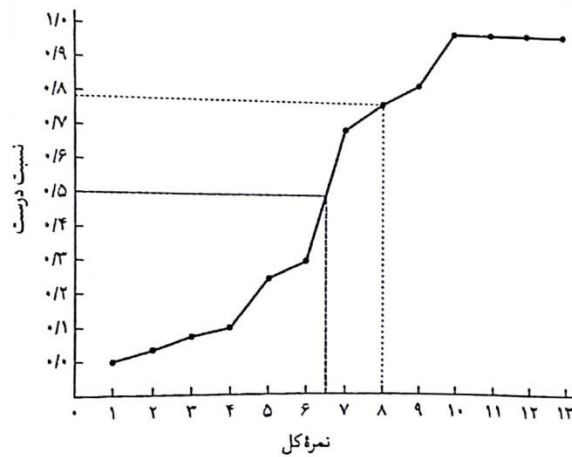
برای نمونه ضریب تمیز مربوط به سوال شماره ۳ را حساب می‌کنیم، اطلاعات موردنیاز این سوال در شکل ۵-۱۷ داده شده است.

$$\begin{aligned} D &= \frac{6}{6+0} = \frac{1}{1+3} \\ &= 1-0/25 \\ &= 0/75 \end{aligned}$$

چنان که از شکل ۵-۱۷ پیداست، از ۶ نفر دانش‌آموزی که به حد تسلط رسیده‌اند، هر ۶ نفر سوال را درست جواب داده و هیچ کس آن را غلط جواب نداده است. همچنین از ۴ نفر دانش‌آموزی که به حد تسلط نرسیده‌اند ۱ نفر سوال را درست و ۳ نفر آن را غلط جواب داده‌اند. ضریب محاسبه شده برای قدرت تمیز سوال ۳ برابر است با ۰/۷۵+ که نشان می‌دهد این سوال به خوبی و به صورت مثبت بین گروهی که به حد تسلط رسیده و گروهی که به این حد نرسیده‌اند تمیز قائل شده است. یعنی در جداسازی دانش‌آموزان دو گروه موفق بوده است.

برای محاسبه ضریبهای تمیز سایر سوالهای آزمون نیز می‌توان از همان روش مورد استفاده با سوال شماره ۳ سود جست. این ضریبها برای همه سوالها در پایین جدول ۳-۱۷ وارد شده‌اند.

بازبینی ضریبهای تمیز سوالهای جدول ۳-۱۷ نشان می‌دهد که همه آنها بجز سوال ۱ بالاتر از صفرند. سوالهای ۳، ۷ و ۸ دارای حداکثر ضریب تمیز هستند. ضریب تمیز سوال ۱ به این دلیل صفر است که همه دانش‌آموزان، چه آنهایی که به حد تسلط رسیده و چه آنهایی که به حد تسلط نرسیده‌اند، آن را درست جواب داده‌اند. این موضوع نباید تهیه‌کننده سؤال را نگران سازد، زیرا هدف آزمون این بوده است که همه دانش‌آموزان به آن جواب صحیح بدهند. با این حال، از لحاظ تمایزگذاری بین گروه حد تسلط و گروه غیرحد تسلط، این سوال درست کار نکرده است. ضریب تمیز سؤال ۵ نیز نزدیک به صفر است و درباره این سؤال نیز مانند سؤال شماره ۱ می‌توان داوری کرد.



شکل ۶-۱۷ اطلاعات موردنیاز برای محاسبه ضریب تمیز حد تسلط - غیر حد تسلط سوال شماره ۳

### جدول ۳-۱۷

#### تحلیل آزمون به کمک نظریه سؤال - پاسخ

روشهایی که در قسمت پیشین این فصل در رابطه با تحلیل آزمون توضیح داده شدند روشهای کلاسیک تحلیل آزمون<sup>۱</sup> نام دارند. روشهای کلاسیک یا قدیمی که به روشهای ساده تحلیل آزمون نیز شهرت دارند روشهای کمابیش سراسری هستند که به وسیله معلمان قابل استفاده اند. روشهای پیچیده تری برای این منظور به وجود آمده اند که با دست قابل استفاده نیستند و کاربرد آنها نیاز به کامپیوتر دارد. این روشهای نسبتاً پیچیده تحلیل آزمون به نظریه سؤال - پاسخ<sup>۲</sup> (آی آر تی) یا نظریه نوین آزمون سازی<sup>۳</sup> وابسته اند. ما در اینجا خوانندگان را با اصول کلی این روشها آشنا می کنیم، اما برای کسب اطلاعات جامع درباره آنها لازم است به منابع و مراجع تخصصی که در ضمن مباحث این قسمت معرفی می شوند مراجعه نمایند.

نظریه سؤال - پاسخ بر نظریه صفت مکنون<sup>۴</sup> استوار است (راست<sup>۵</sup> و کولومبوک<sup>۶</sup>، ۱۹۹۲). نظریه صفت مکنون این فرض را می پذیرد که یک صفت یا ویژگی زیربنایی وجود دارد که به شخص امکان میدهد تا در یک تکلیف معین موفقیت کسب نماید. از جمله اینگونه صفتها میتوان دانش معنی لغات، استدلال ریاضی<sup>۷</sup>، یا "تجسم فضایی" را نام برد (ثرندایک، کانینگهام، ثرندایک، و هیگن، ۱۹۹۱).

<sup>۱</sup> . classical test analysis  
<sup>۲</sup> . item - response theory (IRT)  
<sup>۳</sup> . modern test theory  
<sup>۴</sup> . latent trait theory  
<sup>۵</sup> . Rust  
<sup>۶</sup> . Golombok

اندیشه زیربنای آرتی بسیار روشن و منطقی است. این نظریه همان منطق زیربنای اندیشه آلفرد بینه<sup>۱</sup> بنیانگذار آزمونهای هوشی را به کار می‌گیرد (گرانلاند ولین، ۱۹۹۰). بینه می‌گفت نسبت کودکانی که پاسخ یک سؤال را می‌دانند با سن افزایش می‌یابد. در نظریه سؤال - پاسخ، به جای سن، توانایی به کار می‌رود و گفته میشود که احتمال پاسخ درست به سؤال با افزایش توانایی افزایش می‌یابد. به سخن دیگر، در نظریه سؤال - پاسخ فرض بر این است که مثلاً احتمال پاسخ درست دادن به یک سؤال جبر با افزایش دانش جبر آزمون شونده افزایش می‌یابد، و این بالقوه یک بیان منطقی است.

در نظریه سؤال - پاسخ (آرتی) یا نظریه صفت مکنون، میزان احتمال پاسخ درست دادن یک فرد (آزمون شونده) به یک سوال آزمون با سطح معینی از توانایی مورد اندازه‌گیری مشخص می‌شود. به سخن دیگر، آرتی مجموعه ای الگوی احتمالاتی به دست می‌دهد که به کمک آنها می‌توان رابطه بین سطح توانایی یک آزمون شونده [که معمولاً با  $\theta$  (تتا) نشان داده می‌شود] و احتمال پاسخ درست دادن او به هر سؤال آزمون را تعیین کرد.

شولتز و ویتنی (۲۰۰۵) میگویند الگوهای اولیه آرتی (در دهه های ۱۹۷۰-۱۹۴۰ میلادی)

شربا داده های دوارزشی (صفر برای جواب غلط و ۱ برای جواب درست) پدید آمدند که بیشتر پالرمونهای تواناییهای ذهنی قابل استفاده بودند، اما پژوهشگران بعدها متوجه شدند که آن الگوها را می‌توان با داده های حاصل از ابزارهای سنجش شخصیت و نگرش که در آنها پاسخ سؤالاها به صورت موافق - مخالف یا بله - نه تصحیح می‌شوند نیز مورد استفاده قرار داد. بعدها (تا دهه ۱۹۸۰) الگوهای ابداع شدند که با داده های چند ارزشی، مانند داده های حاصل از مقیاسهای ۱ تا ۵ ارزشی (کاملاً موافق تا کاملاً مخالف)، قابل کاربست‌اند.

### مقایسه نظریه کلاسیک با نظریه سؤال - پاسخ آزمون سازی

چنانکه دیدیم، در روش کلاسیک تحلیل آزمون، سؤالاها یا ماده های یک آزمون با توجه به شاخصهای دشواری (P) و تمیز (D) حفظ یا حذف می‌شوند. در این روش، مقدار P از طریق محاسبه درصد پاسخ دهندگانی که به یک سؤال پاسخ درست میدهند و مقدار D از راه محاسبه ضریب همبستگی بین سؤال وکل آزمون تعیین می‌گردند. افزون بر این، نمره آزمون شونده درکل آزمون (یا سطح توانایی او در موضوع مورد آزمون) از مجموع سؤالاها می‌گردد که درست جواب داده است - صرف نظر از اینکه کدام سؤالاها را درست جواب داده - به دست می‌آید.

به خلاف روش کلاسیک که در آن نمره آزمون شونده از مجموع پاسخهای درست او به سؤالاها آزمون به دست می‌آید، در روش آرتی نمره آزمون شونده برابر با مجموع پاسخهای درست او نیست. نمره آرتی که معمولاً با  $\theta$  نشان داده می‌شود هم تابع تعداد پاسخهای درست آزمون شونده وهو تابع ویژگیهای سؤالاها آزمون است (هوغان،

---

<sup>۱</sup> . Alfred Binet

۲۰۰۷). به عنوان مثال، اگر در یک آزمون دو آزمون شونده به تعداد مساوی سؤال درست جواب داده باشند اما یکی از آنها به سؤالهای آسان ولی دیگری به همان تعداد سؤال دشوار پاسخ داده باشد نمره آن دو مساوی نخواهد بود. به سخن دیگر، هر چند که این دو آزمون شونده به تعداد مساوی سؤال جواب داده اند، اما آنکه به سؤالهای سخت تر جواب داده است نمره  $\theta$  بیشتری خواهد گرفت.

مورد دیگر اختلاف بین این دو روش در آن است که در نظریه کلاسیک اندازه گیری اطلاعات مربوط به پارامترها یا شاخصهای سؤال، مانند ضریبهای دشواری و تمیز، در نمونه های مختلف آزمون شوندگان متفاوت به دست می آیند. به عنوان مثال، سؤالی که از یک آزمون ریاضیات دانشگاهی سطح کارشناسی) انتخاب شده باشد برای دانش آموزان دبیرستانی سؤالی سخت، برای دانشجویان دوره کارشناسی سؤالی معمولی، و برای دانشجویان کارشناسی ارشد سؤالی آسان به نظر خواهد رسید.

به خلاف نظریه کلاسیک، در نظریه سؤال - پاسخ، اطلاعات به دست آمده درباره پارامترهای دشواری و تمیز سؤال از نمونه ای از آزمون شوندگان به نمونه ای دیگر فرق نمیکند. دلیل این امر آن است که الگوی ریاضی مورد استفاده برای محاسبه پارامترهای سؤالها (از جمله الگوی لوجستیک<sup>۱</sup> یا الگوی بهنجار آجیو<sup>۲</sup>) براساس صفت مکنون به دست می آید نه از روی نمره کل آزمون شونده" (شولتز و ویتنی، ۲۰۰۵، ص ۳۲۹). بنابراین، در استفاده از آی آر تی، اطلاعات حاصل از یک نمونه از آزمون شوندگان برابر خواهد بود با اطلاعات حاصل از نمونه دیگری از آزمون شوندگان - صرف نظر از میانگین توانایی گروههای آزمون شونده.

نظریه کلاسیک آزمون سازی در طول سالها ابزار بسیار مفیدی در دست آزمون سازان برای ساختن آزمونهای مختلف بوده است، و هنوز هم روشهای وابسته به این نظریه برای تهیه و بازبینی آزمونهای مختصر روشهای قابل استفاده ای هستند و به طور گسترده مورد استفاده قرار می گیرند (شولتز و ویتنی، ۲۰۰۵). روشهای برخاسته از نظریه صفت مکنون یا نظریه سؤال - پاسخ روشهای نوتر و پیشرفته تری هستند که اطلاعات جامع تر و تعمیم پذیرتری به دست می دهند. با این حال، هر چند که این روشهای نو از روشهای کلاسیک پیشرفته ترند اما بیشتر برای استفاده با آزمونهای جامع و استاندارد شده ملی مفیدند، ولی برای آزمونهای مختصر معلم ساخته همان روش کلاسیک بیشتر قابل استفاده است. کوهن و سوردلیک (۲۰۰۵) علت کم مورد استفاده قرار گرفتن روشهای جدید برای آزمونهای معلم ساخته را به "مسائل فنی و پیچیده ای که باید حل و فصل شوند تا معلوم گردد که داده های به دست آمده با ریاضیات الگوی مورد نظر سازگارند یا نه" (ص ۲۱۹) نسبت داده اند.

---

<sup>۱</sup> . logistic

<sup>۲</sup> . normal ogive

نکته ای را که در اینجا باید یادآور شویم این است که روشهای کلاسیک و آی آر تی ضد هم نیستند، بلکه هردو را می توان برای اکثریت آزمونها به کار بست. با این حال، بنا به گفته کوهن (۲۰۰۷).

روشهای آی آر تی به نمونه های بزرگ نیاز دارند؛ در غیر این صورت، روشهای ریاضی آنها درست از کار در نمی آیند. از این رو، روشهای آی آر تی برای آزمونهای معلم ساخته کارایی ندارند. برای آزمونهای کلاسی معلم ساخته روشهای کلاسیک روانسنجی قابل استفاده ترند. در عوض، روشهای آی آر تی برای آزمونهای جامع و ملی استاندارد شده مفیدند. (ص ۱۴)

### منحنی ویژگی سؤال و استفاده از آن برای تحلیل سؤال آزمون

یکی از روشهای تحلیل آزمون برخاسته از نظریه سؤال - پاسخ استفاده از منحنی ویژگی سؤال (آی سی سی)<sup>۱</sup> یا تابع سؤال - پاسخ (آی آراف)<sup>۲</sup> است. منحنی ویژگی سؤال یک بازنمایی نموداری از رابطه بین احتمال پاسخ درست دادن به یک سؤال و موقعیت آزمون شونده در صفت مورد اندازه گیری است. در رابطه با سنجش توانایی، این منحنی راهی برای تعیین اینکه شخصی با توانایی X بتواند عملکردی در سطح Y داشته باشد به دست می دهد. در رابطه با سنجش شخصیت، این احتمال را که شخصی با X مقدار از یک ویژگی شخصیتی در ابزاری که برای سنجش آن ویژگی درست شده است عملکردی در سطح Y دارد تعیین می کند (کوهن و سوردلیک، ۲۰۰۵).

تهیه منحنی ویژگی سؤال و تعیین مقدار شاخصهای دشواری و تمیز عموماً با استفاده از کامپیوتر عملی است. با این حال، برای نشان دادن چگونگی رسم کردن این منحنی می توان تعداد اندکی نمره و محاسبه با دست را مورد استفاده قرار داد. ما در زیر، به منظور آگاه ساختن خوانندگان از چگونگی رسم این منحنی و نحوه استفاده از آن، مثالهای ساده ای می آوریم.

برای رسم منحنی ویژگی سؤال، ابتدا نسبتها یا درصدهای آزمون شوندگانی را که در آزمون نمرات مختلف گرفته اند و به آن سؤال جواب درست داده اند تعیین و این نسبتها را بر روی محور عمودی محور مختصات مشخص میکنیم. این نسبتها از یک تابع ریاضی به نام معادله لوجستیک<sup>۳</sup> استخراج می شوند (نگاه کنید به بیکر<sup>۴</sup>، ترجمه هومن و عسگری، ۱۳۸۱). پس از آن، بر روی محور افقی محور مختصات نمره های کل آزمون در قالب نمره های خام یا نمره های استاندارد به صورت  $\theta$  (تتا) را مشخص میکنیم. آنگاه، نقاط تقاطع خطهای عمودی برخاسته از هریک از نمره های آزمون

<sup>۱</sup> . item characteristic curve (ICC)

<sup>۲</sup> . item characteristic function (IRF)

<sup>۳</sup> . Logistic function

<sup>۴</sup> . Baker

برروی محور افقی) و خطهای عمودی برخاسته از هریک از نسبتها (بر روی محور عمودی) را تعیین میکنیم. سرانجام، از طریق وصل این نقاط به یکدیگر منحنی ویژگی سؤال را رسم میکنیم. هرچند انجام محاسبات موردنیاز و رسم منحنی ویژگی سؤال به کمک کامپیوتر انجام می گیرد، با این حال ما در اینجا از اطلاعات جدول ۴-۱۷ و محاسبات دستی استفاده می کنیم.)

جدول ۴-۱۷ نتایج یک آزمون ۱۳ سؤالی را نشان میدهد. در این آزمون ارزش هر سؤال ۱ نمره است، و نمره کل هر آزمون شونده از مجموع نمرات درست او به دست آمده است. در جدول ۴-۱۷ داده های لازم برای رسم منحنی های ویژگی سؤالهای ۱، ۲ و ۳ و ۴ این آزمون داده شده اند. به عنوان نمونه، در این جدول ۷۷ درصد کسانی که در آزمون نمره ۸ گرفته اند به سؤال ۱ جواب درست داده اند. با توجه به بقیه نسبتهای درست سؤال ۱ منحنی ویژگی آن در شکل ۶-۱۷ رسم شده است.

دقت کنید که در منحنی ویژگی سؤال ۱ (شکل ۶-۱۷) نسبتهای درست برای این سؤال همراه با افزایش نمره کل آزمون افزایش می یابد. این مطلب نشان می دهد که هرچه نمره کل آزمون یک فرد بزرگتر باشد با احتمال بیشتری به این سؤال پاسخ درست خواهد داد. یعنی، براساس منطق آی آر تی، هرچه توان آزمون شوندگان در ارتباط با ویژگی مورد اندازه گیری بیشتر بوده در این آزمون نمره بهتری به دست آورده است.

جدول ۴-۱۷ نسبت یا درصد درست پاسخهای ۴ سوال از یک آزمون ۱۳ سؤالی

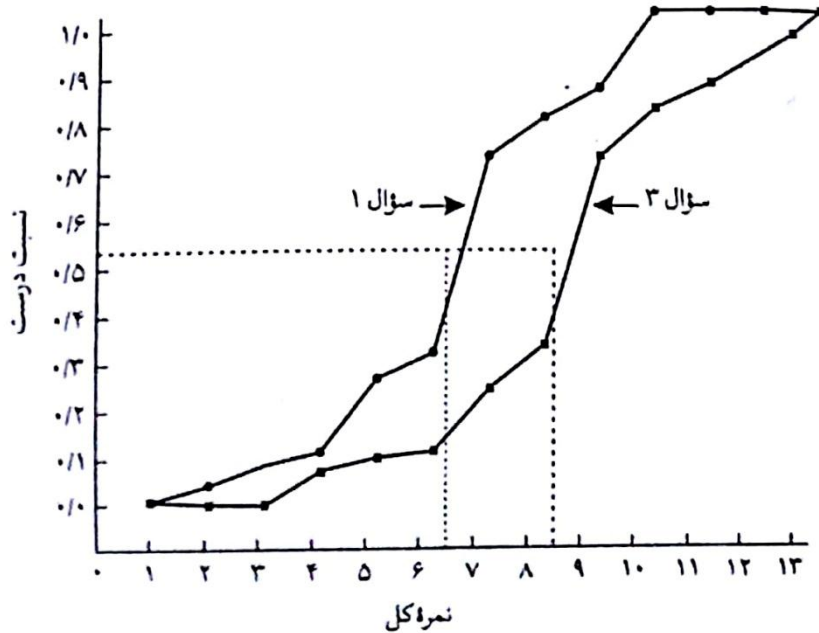
نمره کل	شماره سؤال			
	۱	۲	۳	۴
۰	.۰۰	.۰۰	.۰۰	.۰۰
۱	.۰۰	.۵۰	.۰۰	.۱۷
۲	.۰۳	.۶۱	.۰۰	.۱۸
۳	.۰۶	.۱۹	.۰۰	.۲۲
۴	.۰۸	.۳۰	.۰۳	.۳۱
۵	.۲۳	.۸۹	.۰۶	.۴۰
۶	.۳۱	.۳۱	.۰۸	.۴۳
۷	.۷۰	.۷۹	.۲۳	.۵۷
۸	.۷۷	.۲۰	.۳۱	.۶۱
۹	.۸۲	.۵۲	.۷۰	.۷۲

۱۰	.۹۹	.۵۹	.۷۷	.۸۲
۱۱	۱.۰۰	.۳۱	.۸۲	.۸۳
۱۲	۱.۰۰	.۷۰	.۹۹	.۸۵
۱۳	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰

از روی منحنی ویژگی سؤال می‌توان ضریبهای دشواری و تمیز سؤال را تعیین کرد. ضریب دشواری یا پارامتر دشواری با علامت  $b$  نشان داده می‌شود و معرف نقطه عطف منحنی ویژگی سؤال است. در سوالهایی که عامل حدس در پاسخگویی به سوالها تاثیر ندارد، نقطه عطف منحنی ویژگی سؤال ۵۰٪ (یا نسبت ۰/۵) است. یعنی ضریب دشواری یا سطح دشواری سؤال عبارت است از نمره معیاری که در آن ۵۰ درصد آزمون شوندگان به سؤال جواب درست داده‌اند. به سخن دیگر، "سطح دشواری سؤال نمره ای است که در آن پنجاه درصد آزمون شوندگان جواب درست می‌دهند" (کوالی<sup>۱</sup>، ۲۰۱۰، ص ۴۲). (از آنجا که در مثالهای فرضی ما عامل حدس منظور نشده است، ضریبهای دشواری سوالها از روی همان ۵۰ درصد محاسبه شده‌اند.) دقت کنید که محور نمرات کل آزمون یا اندازه‌های توانایی و سطح دشواری یکی است؛ از این رو، می‌توان دشواری سؤال را مستقیماً به سطح توانایی مربوط به آن تبدیل کرد. در منحنی ویژگی سؤال شکل ۶-۱۷ ضریب دشواری سؤال با نمره ۶/۵ نشان داده شده است. این ضریب نشان می‌دهد کسانی که در آزمون فرضی ۱۳ نمره ای ما نمره ۶/۵ و بالاتر گرفته باشند می‌توانند به این سؤال پاسخ درست بدهند.

اگر همه آزمون شوندگان به یک سؤال جواب درست بدهند، منحنی ویژگی آن سؤال یک خط افقی در بالای محور عمودی خواهد بود که از نسبت ۱ می‌گذرد. در این صورت ضریب دشواری سؤال  $b = 1$  است. به عکس، اگر همه آزمون شوندگان به یک سؤال جواب غلط بدهند، منحنی ویژگی آن سؤال نیز یک خط افقی خواهد بود که این بار از پایین محور عمودی و از نسبت صفر می‌گذرد. ضریب دشواری این سؤال  $b = 0$  است. با توجه به این توضیحات، می‌توان ضریب دشواری سؤال را مستقیماً به سطح زیر منحنی ویژگی سؤال ربط داد؛ هرچه سطح زیر منحنی ویژگی سؤال بیشتر باشد، ضریب دشواری سؤال بزرگ‌تر است (راجرز<sup>۲</sup>، ۱۹۹۵). شکل ۷-۱۷ منحنی‌های ویژگی سوالهای ۱ و ۳ جدول ۴-۱۷ و ضریبهای دشواری آنها را نشان می‌دهد. ضریب دشواری سؤال ۱ برابر با ۶/۵ و ضریب دشواری سؤال ۳ برابر با ۸/۵ است.

<sup>۱</sup> . coaly  
<sup>۲</sup> . Ragers



شکل ۷-۱۷ منحنیهای ویژگی سوالهای ۱ و ۳ جدول ۴-۱۷ با ضریبهای دشواری متفاوت و ضرایب تمیز برابر

نقطه تقاطع خط افقی برآمده از نسبت ۰/۵ و منحنی ویژگی سوال را که نقطه عطف منحنی است آستانه<sup>۱</sup> می‌گویند، زیرا در این نقطه جهت منحنی تغییر می‌کند. افرادی که نمره کل آزمون آنها پایین این نقطه است احتمال جواب درست دادن به سوال برای آنها در حد تصادف است، اما کسانی که نمره کل آزمون آنها بالاتر از این نقطه است با احتمال زیاد به سوال جواب درست خواهند داد. به این سبب گفته می‌شود که عملکرد آزمون شونده در سوال به توانایی او وابسته است.

ضریب تمیز با پارامتر قدرت تشخیص سوال با علامت  $a$  نشان داده می‌شود و برابر با ضریب زاویه شیب منحنی ویژگی سوال در آزمونهای استاندارد شده جامع که در آنها عامل حدس بی تأثیر است پارامتر تشخیص یا قدرت تمیز یا همبستگی دو رشته‌ای ارتباط دارد به گونه‌ای که می‌توان آن را با استفاده از ضریب همبستگی دو رشته‌ای با فرمول زیر تعیین کرد (لرد<sup>۲</sup>، ۱۹۷۵).

$$a = \frac{rbis}{\sqrt{1 - r^2 bis}}$$

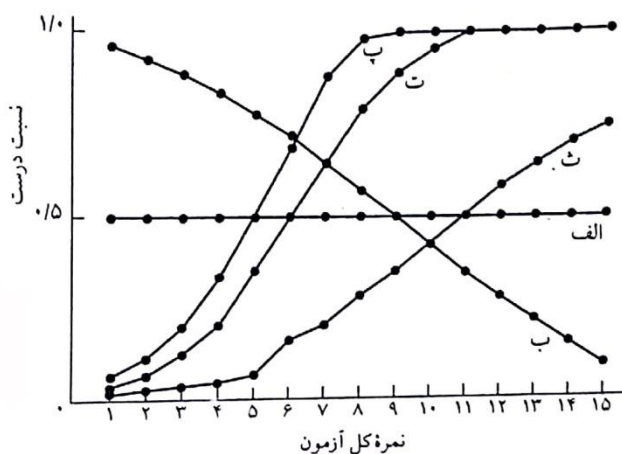
<sup>۱</sup> . threshold  
<sup>۲</sup> . Lord



در استفاده از فرمول بالا برای محاسبه ضریب تمیز سوال (a)، علاوه بر اینکه توانایی موردنظر  $\theta$  باید دارای توزیع نرمال (طبیعی) باشد، لازم است چند شرط آماری دیگر نیز برآورده شوند. برای اطلاعات بیشتر، به منابع تخصصی روانسنجی، از جمله لرد (۱۹۷۵)، لرد و نوک<sup>۱</sup> (۱۹۶۸)، و امبرتسون<sup>۲</sup> و رایز<sup>۳</sup> (۲۰۰۰) مراجعه کنید.

در شکل ۸-۱۷ منحنیهای ویژگی ۵ سوال با هم مقایسه شده‌اند. چنان که دیده می‌شود، سوال الف یک سوال ضعیف است، زیرا بین آزمون شوندگان دارای نمرات متفاوت در آزمون (یا آزمون شوندگان قوی و ضعیف) تمیز قائل نمی‌شود؛ احتمال جواب درست دادن به این سوال در تمام سطح نمرات کل آزمون یکساله است. سوال ب نیز یک سوال معیوب است، زیرا این سوال با کل آزمون رابطه معکوس دارد. به سخن دیگر، این سوال دارای ضریب تمیز منفی است. سوالهای پ و ت سوالهای نسبتاً خوبی هستند. منحنی ویژگی سوال ث که از منحنی های ویژگی سوالهای پ و ت خوابیده‌تر است حاکی از این است که به خوبی دو سوال قبلی بین گروههای قوی و ضعیف آزمون شوندگان تمیز قائل نمی‌شود.

حال به منحنی ویژگی سوال شکل ۹-۱۷ نگاه کنید. این منحنی در فاصله بین نمرات ۷ و ۸ دارای شیب بسیار تند است. سوالی که این منحنی معرف آن است آزمون شوندگان دارای نمرات پایین ۸ و آزمون شوندگان دارای نمرات بالای ۸ را به خوبی از یکدیگر متمایز می‌کند. یعنی اینکه، آزمون شوندگانی که نمره کل آنها در آزمون پایین تر از ۸ است به این سوال پاسخ درست نداده‌اند، ولی آزمون شوندگان که در آزمون نمره ۸ یا بالاتر گرفته‌اند به سوال پاسخ درست داده‌اند. هرچه منحنی ویژگی یک سوال به شکل ۹-۱۷ نزدیک تر باشد، سوال از لحاظ ضریب تمیز بهتر است، یعنی بهتر می‌تواند نمره کل آزمون شونده را پیش بینی کند. به سخن دیگر، هرچه منحنی ویژگی یک سوال حالت پلکانی بیشتری داشته باشد همبستگی بین آن سوال و کل آزمون بیشتر است.

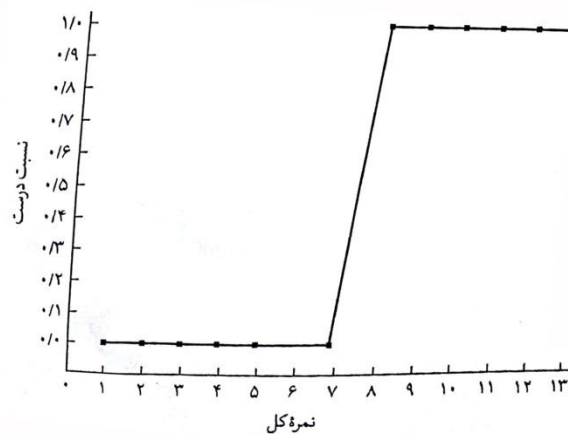


<sup>۱</sup> . Novic  
<sup>۲</sup> . embretson  
<sup>۳</sup> . Reise

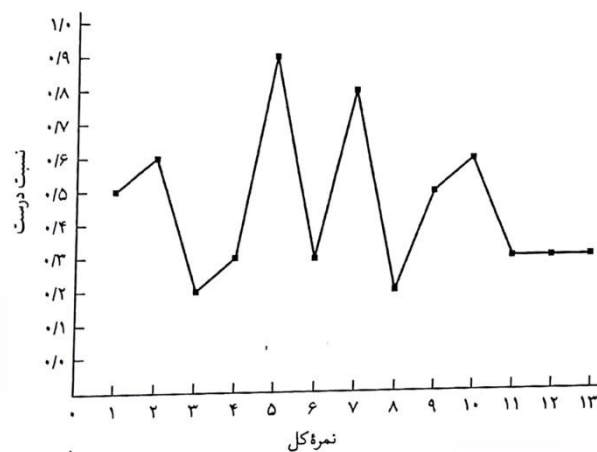
### شکل ۸-۱۷ مقایسه منحنیهای ویژگی ۵ سوال مختلف

شکل ۱۷-۱۰ منحنی ویژگی سوال ۲ مربوط به جدول ۴-۱۷ را نشان می دهد. این سوال به هیچ وجه بین آزمون شوندگان مختلف تمیز قائل نمی شود. این سوال هیچ گونه رابطه منحنی بین نسبتهای درست و نمرات کل آزمون نشان نمی دهد و لذا به خوبی عمل نمی نمایند.

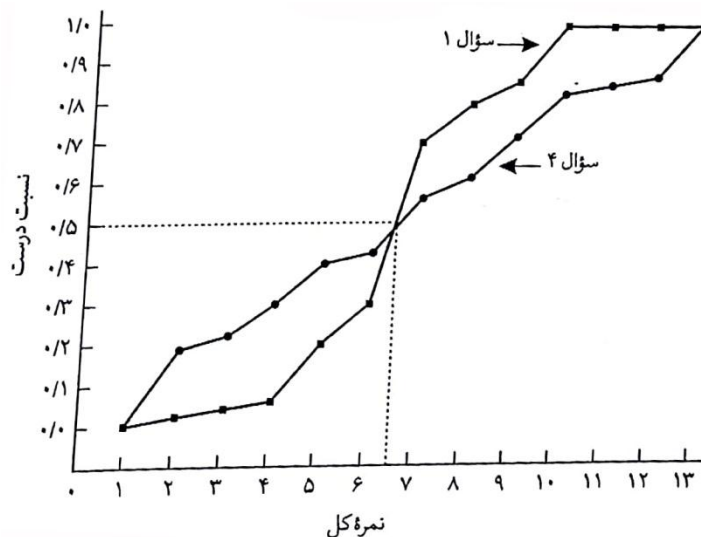
ویژگی های دشواری و تمیزسوالهای یک آزمون گرچه به هم وابسته اند، اما تا حدودی از یکدیگر مستقل اند. برای روشن شدن مطلب به منحنیهای ویژگی سوالهای ۱ و ۴ مربوط به جدول ۴-۱۷ در شکل ۱۱-۱۷ نگاه کنید. در این شکل نشان داده شده است که سوال ۱ دارای قدرت تمیز بیشتری از سوال ۴ است، زیرا شیب آن تندتر به حالت پلکانی نزدیک تر است. این در حالی است که آستانه یا ضریب دشواری هر دو سوال برابر است.



### شکل ۹-۱۷ منحنی ویژگی سوال با قدرت تمیز بسیار خوب



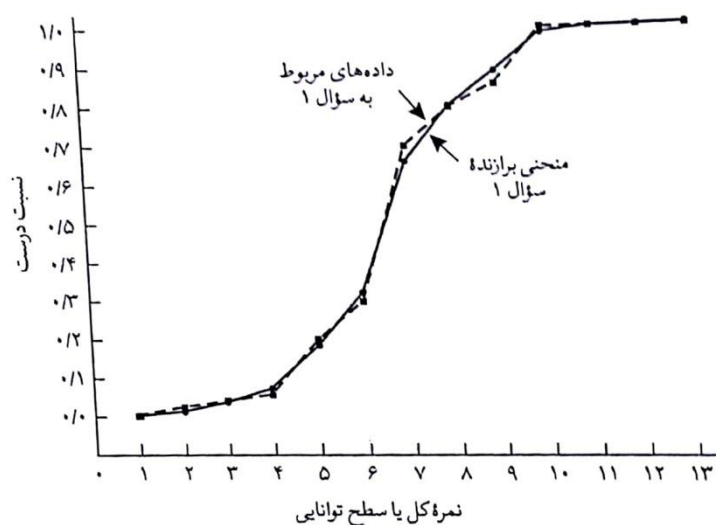
### شکل ۱۰-۱۷ منحنی ویژگی یک سوال ضعیف



شکل ۱۱-۱۷ منحنیهای ویژگی سوالهای ۱ و ۴ و جدول ۴-۱۷ با ضریبهای دشواری برابر و ضریبهای تمیز متفاوت

### برازنده کردن منحنی ویژگی سوال

برای استفاده بیشتر از منحنی ویژگی سوال، کوشش به عمل می‌آید تا مناسب‌ترین یا برازنده‌ترین<sup>۱</sup> منحنی ریاضی آن سوال رسم شود. معمولاً برازنده‌ترین منحنی ریاضی یک سوال به صورت منحنی‌ای که توان دوم فواصل بین آن و داده‌های مشاهده شده مربوط به سوال حداقل باشد تعریف می‌شود (راجرز، ۱۹۹۵). شکل ۱۲-۱۷ برازنده‌ترین منحنی ویژگی سوال ۱ جدول ۴-۱۷ را نشان می‌دهد.



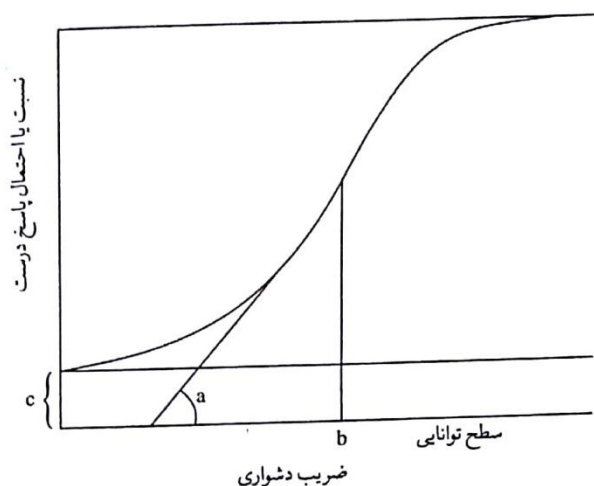
شکل ۱۲-۱۷ برازنده‌ترین منحنی ریاضی سوال شماره ۱ جدول ۳-۱۷

<sup>۱</sup> . item characteristic function

پس از آنکه منحنی ویژگی سوال به صورت برازنده رسم شد، از روی آن می‌توان چند پارامتر مهم سوال را تعیین کرد. سه پارامتر یا سه ارزش لازم‌اند تا به کک آنها منحنی ویژگی سوال توصیف شود. این پارامترها به قرار زیرند:

- پارامتر  $a$  که نشان دهنده شکل یا شیب منحنی ویژگی سوال است. چنان که قبلاً دیدیم، شیب صعودی یا منحنی دارای شکل پلکانی نشان دهنده قدرت تمیز زیاد سوال است.
- پارامتر  $b$  معرف آستانه یا همان ضریب دشواری سوال است. ضریب دشواری سوال برابر است با نمره‌ای از آزمون که مقابل نقطه تغییر یا نقطه عطف در منحنی ویژگی سوال است.
- پارامتر  $c$  یا پارامتر تصادف میزان حدس پذیری سوال را مشخص می‌کند. هرچه امکان کسب نمره از راه حدس زدن در یک سوال بیشتر باشد مقدار  $c$  برای آن سوال نیز بزرگ‌تر است.

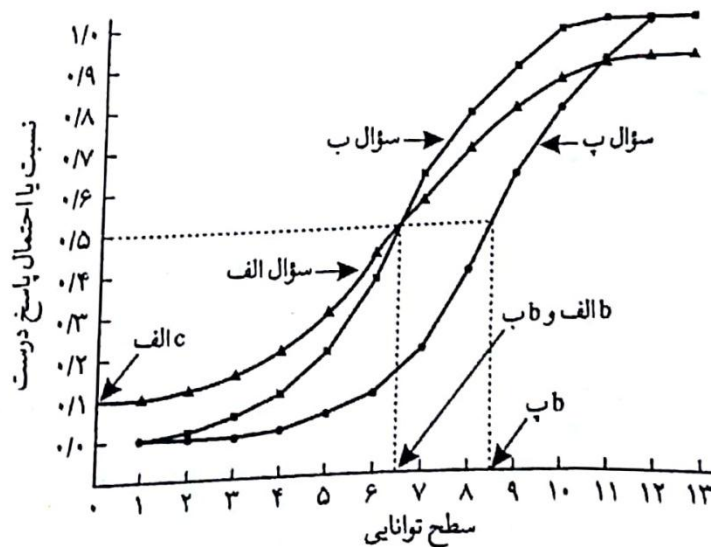
مجموع پارامترهای فوق را در منحنی ویژگی سوال برازنده شده یا هموار شده که تابع ویژگی سوال<sup>۱</sup> نیز نام دارد می‌توان به صورت شکل ۱۳-۱۷ نشان داد. دقت کنید که شکل این منحنی یا این تابع به صورت تابع تراکمی توزیع طبیعی اجابو است.



شکل ۱۳-۱۷ شکل کلی منحنی یا تابع ویژگی سوال  $a = \tan z$  = ضریب تمیز

پارامتر  $c$  یا میزان حدس پذیری در آزمونهایی مانند آزمونهای چندگزینه‌ای که در آنها امکان درست جواب دادن یک سوال از راه حدس زدن کورکورانه یا هشیارانه وجود دارد مطرح است. هرچه این امکان بیشتر باشد مقدار  $c$  نیز بیشتر است. اگر عامل حدس در سوالی بالا باشد، تعداد بیشتری از آزمون شوندگان دارای توانایی کم، احتمالاً، به آن جواب درست خواهند داد، و در این صورت منحنی از محلی در بالای نقطه صفر در روی محور عمودی شروع خواهد شد. حال به منحنی شکل ۱۱-۱۷ (ص ۴۵۶) نگاه کنید. می‌بینید که سوال ۴ دارای عامل حدس بالاست، اما در سوال ۱ عامل حدس صفر است؛ و این در حالی است که هر دو سوال دارای ضریب دشواری متوسط‌اند.

چنان که دیدیم،  $a$ ،  $b$ ،  $c$  پارامترهای ریاضی حاصل از منحنی ویژگی سوال برازنده شده هستند. شکل ۱۴-۱۷ منحنیهای هموار شده یا برازنده شده سه سوال مختلف را نشان می‌دهند.



شکل ۱۴-۱۷ منحنیهای ویژگی سه سوال دارای پارامترهای مختلف

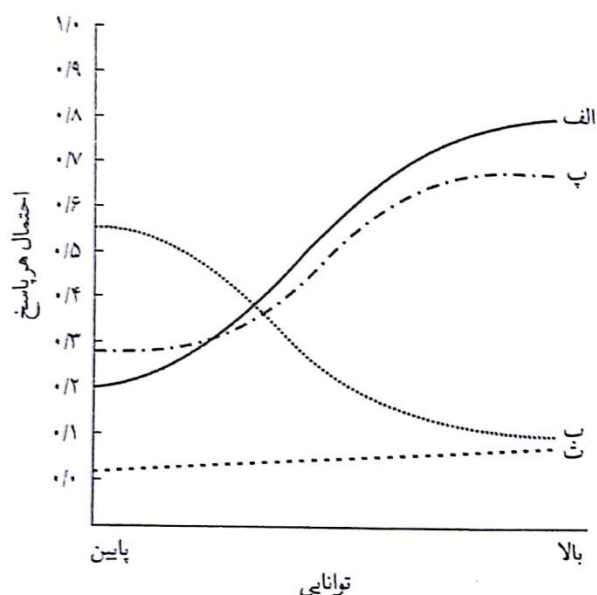
سؤالهای ب و پ از لحاظ پارامتر شکل یا قدرت تمیز مساوی‌اند، اما از لحاظ پارامتر آستانه یا ضریب دشواری با هم فرق دارند. سؤالهای الف و ب از لحاظ پارامتر ضریب دشواری برابرند، اما از لحاظ پارامتر ضریب تمیز با هم فرق دارند. سوال الف دارای ضریب حدس پذیری است، زیرا نسبت کوچکی از آزمون شوندگان دارای توانایی صفر سوال را به طور صحیح جواب داده‌اند.

#### تحلیل گزینه‌های انحرافی به کمک منحنی ویژگی سوال

در نظریه سؤال - پاسخ هر یک از گزینه‌های یک سوال چند گزینه‌ای (چه گزینه درست و چه گزینه‌های انحرافی) می‌تواند اطلاعاتی درباره دانش یا مهارت پاسخ دهنده به دست دهد. برای این منظور، لازم است، علاوه بر منحنی ویژگی گزینه درست، برای گزینه‌های غلط نیز منحنی ویژگی رسم کنیم. به عنوان نمونه، شکل ۱۵-۱۷ هم منحنی ویژگی گزینه درست (الف) و هم منحنیهای انحرافی (ب، پ، ت) را نشان می‌دهد.

چنان که در شکل ۱۵-۱۷ دیده می‌شود، گزینه درست یا گزینه الف دارای قدرت تمیز مثبت و بالاست. گزینه انحراف پ نیز دارای قدرت تمیز مثبت و نسبتاً بالاست. اما گزینه ت را کمتر کسی انتخاب کرده و قدرت تمیز ندارد. گزینه ب دارای قدرت تمیز منفی است، یعنی پاسخ دهندگان ضعیف بیشتر از پاسخ دهندگان قوی آن را انتخاب کرده‌اند.

در نمره گذاری سوال به روش کلاسیک، پاسخ به یک سوال یا درست ارزیابی می‌شود یا غلط. اگر پاسخ دهنده گزینه درست را انتخاب کند نمره ۱ می‌گیرد و اگر هر یک از گزینه‌های انحرافی را انتخاب کند نمره صفر می‌گیرد. اما در نظریه سؤال - پاسخ انتخاب گزینه‌های انحرافی از سوی پاسخ دهنده یکسان تلقی نمی‌شود، زیرا در سوالهای چند گزینه‌ای بعضی گزینه‌های غلط به گزینه درست نزدیک‌تر و بعضی از آنها از گزینه درست بسیار فاصله دارند. به عنوان مثال، در سوال شکل ۱۷-۱۵ گزینه پ بسیار نزدیک به گزینه الف است، اما گزینه‌های ب و ت با گزینه الف بسیار متفاوت‌اند. بنابراین، در نمره گذاری پاسخ دهندگان به این سوال، می‌توانیم به کسانی که گزینه الف را انتخاب کرده‌اند نمره کامل بدهیم. همچنین برای کسانی که گزینه پ را برگزیده‌اند نمره نزدیک به نمره کامل منظور کنیم. ولی کمترین نمره را برای گزینه‌های ت و ب در نظر بگیریم.



شکل ۱۷-۱۵ منحنی ویژگی گزینه‌های یک سوال چند گزینه‌ای