

مقایسه مفاهیم و مفروضه‌های نظریه کلاسیک و نظریه جدید سؤال - پاسخ در ساخت آزمونهای روانی و تربیتی

دکتر حسین سپاسی*

چکیده

مطالعه مسائل اندازه‌گیری به منظور پیدا کردن راه‌حلهایی برای آنها در روانشناسی و تعلیم و تربیت نظریه‌ای به نام نظریه آزمون‌سازی را به وجود آورده است. هدف هر نظریه آزمون‌سازی کمک به دست‌اندرکاران این فن است تا آنها را از مدل‌های منطقی و ریاضی که زیربنای ساخت آزمونهای روانی و تربیتی را تشکیل می‌دهند آگاه سازد. در این مطالعه تحلیلی مفاهیم و مفروضه‌های دو نظریه مهم در ساخت آزمونهای روانی و تربیتی مورد بررسی قرار گرفت. نظریه کلاسیک آزمون که می‌توان پیدایش آن را به چارلز اسپیرمن روانشناس انگلیسی نسبت داد و امروز از آن به عنوان یک نظریه قدیمی در آزمون‌سازی یاد می‌شود، عمدتاً بر سه مفروضه اساسی نمره مشاهده شده، نمره واقعی، و نمره خطا استوار است. این نظریه تاکنون مبنای اندازه‌گیری ویژگیهای رفتاری مورد استفاده بسیاری از متخصصان آزمون‌سازی قرار گرفته است. محدودیتهایی که اخیراً در این نظریه مشاهده شد موجب گردید تا نظریه دیگری توسط روانسنجانی چون لردونویک (۱۹۶۸)، همبلتون و همکاران (۱۹۹۳)، وانگهارد (۱۹۹۴) ارائه گردد. نظریه سؤال-پاسخ تأثیر مهمی در ساخت آزمونهای روانی و تربیتی داشته است. بر پایه مبانی نظری، این نظریه نشان داده است که هم در پاسخگویی به سؤالات آزمون‌سازی و هم ارائه راهکارهای مناسب بر نظریه کلاسیک آزمون برتری دارد. در این مقاله مفروضه‌های دو نظریه قدیم و جدید و تفاوت‌های موجود در مدل‌های ریاضی آنها ارائه می‌شود. درباره شاخصهائی که این دو نظریه

*- عضو هیات علمی دانشکده علوم تربیتی و روانشناسی دانشگاه شهید چمران اهواز

برای برآورد پارامترهای سؤالهای آزمونهای روانی و تربیتی استفاده می‌کنند بحث خواهد شد. علاوه بر این، مزایا و محدودیتهایی که هر یک از دو نظریه در ساخت آزمونهای روانی و تربیتی دارند بررسی می‌شود. در پایان، به تفاوتها و تشابهات بین دو نظریه و شیوه‌هایی که آزمونها بر اساس این تفاوتها و تشابهات ساخته می‌شود اشاره خواهد گردید.

کلید واژگان: نظریه کلاسیک آزمون، نظریه سؤال- پاسخ، آزمونهای روانی و تربیتی

مقدمه

(۱۹۵۰)، لالی^۲ (۱۹۵۰)، لازارفیلد^۳

(۱۹۵۴) و لردونوویک^۴ (۱۹۶۸)، و غیره آن را به صورت آنچه که امروز مورد استفاده آزمون سازان، محققان و دانشجویان قرار می‌گیرد در آورده‌اند. برای مثال، گولیکسون مفاهیم نظریه کلاسیک آزمون را به صورت جامع‌تری ارائه داد. لردونوویک سازه‌های اصلی در نظریه کلاسیک را با استفاده از رویکردهای ریاضی و آماری مجدداً به شکل تازه‌ای فرمول بندی کردند. تلاش آنها در این جهت بود تا توجه را از نمره کل آزمون به سمت هر سؤال به صورت انفرادی و ویژگیهای آن، که بعداً نقش اساسی در ساختار نظریه آزمون سازی جدید ایفا کردند، جلب نمایند. لالی (۱۹۵۰) و لرد (۱۹۵۲) نشان دادند که می‌توان مفاهیم

در طول قرن بیستم دو رویکرد مهم به وجود آمد که عمده‌تاً اساس نظریه‌های مختلف را در اندازه‌گیری خصوصیات رفتاری تشکیل داد. با این که تفاوتی را می‌توان در این دو نظریه پیدا کرد، ولی هر دو نظریه سعی دارند با ارائه مدلهای متفاوتی ویژگیهای مختلف رفتاری مثل تواناییها، استعدادها، و نگرشها را به صورت کمیت در آورند و سپس به تفسیر آنها بپردازند.

نظریه‌ای که به طور منسجم به اندازه‌گیری خصوصیات رفتاری انسان پرداخت به نظریه کلاسیک آزمون مشهور است و پیدایش آن را عمده‌تاً می‌توان به کارهای چارلز اسپیرمن نسبت داد. این نظریه از بدو پیدایش تاکنون تغییراتی پیدا کرده و دانشمندانی چون گولیکسون^۱

² - Lawley

³ - Lazarfield

⁴ - Lord & Novick

¹ - Gullikson

زمان عکس‌العمل، تمیز شنوایی، برآورد وزن نسبی اشیاء، یا سایر متغیرهای حسی که در آزمایشگاه‌های ادراکی در آلمان مورد مطالعه قرار می‌گرفت گسترش یافته است، معهدا ویژگیهای رفتاری که متخصصان روان فیزیک در شرایط کنترل شده از طریق مشاهده، گردآوری کردند در به وجود آوردن نظریه خصیصه مکنون^۲ لازارفیلد (۱۹۵۴)، نظریه منحنی ویژگی سؤال^۳ لرد (۱۹۷۷)، و بالاخره نظریه سؤال-پاسخ^۴ لردونوویک (۱۹۶۸)، که وابستگی نظریه را به پاسخهای آزمودنیها به سؤالات آزمون نشان می‌دهد نقش اساسی داشته است.

نظریه‌ای که تاکنون در کشور ما مبنای کار سنجش و اندازه‌گیری ویژگیهای رفتاری انسان قرار گرفته عمدتاً بر نظریه کلاسیک آزمون مبتنی است. متخصصان آزمون‌سازی توانسته‌اند با استفاده از مبانی این نظریه مسأله ثبات ابزار مطالعاتی خود را به راههای مقتضی حل و فصل نمایند. آن چه از مطالعه نوشته‌ها و بررسی روش‌شناسی مطالعات و تحقیقاتی که در

اصلی در نظریه کلاسیک را به عنوان تابعی از پارامترهای منحنی ویژگی سؤالات آزمون بیان کرد. کار این دو موجب شد تا مفاهیم اصلی نظریه روان‌سنجی جدید به وجود آید که بعداً به نظریه سؤال-پاسخ شهرت یافت.

نظریه دیگری که به بررسی و اندازه‌گیری ویژگیهای رفتاری پرداخت ابتدا به کارهای متخصصان روان فیزیک^۱ نسبت داده می‌شود. دانشمندان آلمانی چون ویلهلم ونت، ارنست وبر، گوستاو فخنر که بعداً به طرفداران روان فیزیک شهرت یافتند، ویژگی رفتاری انسان را از بعد دیگر می‌نگریستند و ابزار و وسایلی که برای اندازه‌گیری این ویژگیها آرایه دادند با نظریه‌های دیگر متفاوت بود. برای مثال، آنها به جای این که پاسخ آزمودنیها را به سؤالات محور مطالعاتی خود قرار دهند، شدت و ضعف پاسخی که آزمودنیها به محرکهای فیزیکی در شرایط کاملاً کنترل شده می‌دادند مورد مطالعه قرار دادند.

با وجود اینکه اندازه‌گیری تواناییهای ذهنی تا به امروز به مفاهیمی اساسی‌تر از

^۲ - latent trait theory

^۳ - item characteristic curve theory

^۴ - item response theory

^۱ - Psychophysicists

نظریه کلاسیک آزمون

در این مقاله به دلیل آشنایی خوانندگان علاقمند به آزمون سازی، از شرح مفصل نظریه کلاسیک آزمون خودداری می‌شود. ولی، به منظور یادآوری به مفاهیم و مفروضه‌های اساسی این نظریه به اختصار اشاره خواهد شد. بر اساس این نظریه نمره‌های آزمون سه مفهوم مختلف ارایه می‌دهند. نمره مشاهده شده، نمره واقعی، و نمره خطا. در چارچوب این نظریه، مدل‌های خطی ساده مختلف ساخته شده که نمره مشاهده شده (X) را به مجموع نمره‌های دو متغیر غیر قابل مشاهده، نمره واقعی (T)، و نمره خطا (E) ارتباط می‌دهد. چون در این معادله برای هر آزمودنی دو متغیر غیر قابل مشاهده وجود دارد، بنابراین آن را نمی‌توان حل کرد مگر این که فرضهای خاصی در آن به کار رود. این مفروضه‌ها بر اساس نظریه کلاسیک آزمون در جامعه‌ای از آزمودنیها عبارتند از:

۱. میانگین نمره‌های خطا برابر صفر است.
۲. همبستگی بین نمره‌های واقعی و نمره خطا صفر است.

کشورمان در زمینه علوم رفتاری استنباط می‌شود این است که مبانی نظریه جدید آزمون سازی کمتر مورد استفاده محققان، متخصصان آزمون سازی، و دانشجویان دوره‌های تحصیلات تکمیلی قرار گرفته است. علاوه بر این، بررسی مطالب مجلات علمی معتبر نیز که از طرف دانشگاهها یا مراکز آموزش عالی معتبر انتشار می‌یابند، بندرت می‌توان به مقاله‌ای اشاره کرد که از مبانی نظریه سؤال- پاسخ در ساختن آزمونهای روانی و تربیتی استفاده کرده باشند.

بنابراین، به دلیل اهمیت موضوع، در این مقاله، مفاهیم، مدلها، و روشهای به کار برده شده در نظریه کلاسیک آزمون و نظریه سؤال- پاسخ مورد مقایسه قرار خواهند گرفت. چون در ارایه مدل‌های این دو نظریه تفاوت‌های آماری و ریاضی وجود دارد. بنابراین، شیوه‌هایی که آزمون بر اساس این تفاوتها ساخته می‌شوند مورد بررسی قرار خواهد گرفت. در پایان هم به تشابهات و تفاوت‌های بین دو نظریه اشاره خواهد شد.

۳. همبستگی بین نمره‌های خطای آزمونهای همتا برابر صفر است. بر اساس مفروضه‌های فوق، نقش آزمونهای همتا در ساختن آزمونها بر اساس نظریه کلاسیک آزمون حیاتی است. آزمونهای همتا به آزمونهایی گفته می‌شود که به اندازه‌گیری محتوی یکسانی می‌پردازند و در آن آزمودنیها نمره واقعی یکسانی کسب می‌کنند. علاوه بر این، مقدار خطاهای اندازه‌گیری در فرمهای همتای آزمون یکسان است. بر اساس مبانی این نظریه، تعاریف، مفروضه‌ها، و مدلهای جدیدی در ساختن آزمونهای روانی و تربیتی به کار گرفته شده و تا به امروز بسیاری از آزمونهای مهم بر اساس مبانی این نظریه ساخته و پرداخته شده است.
- برای پیدا کردن مدلهای بهتری در چارچوب نظریه کلاسیک آزمون، متخصصان آزمون سازی تلاش زیادی انجام داده‌اند. گاهی با حذف برخی مفروضه‌ها و گاهی با تجدید نظر در آنها، و گاهی هم با اضافه کردن مفروضه‌های دیگری درباره توزیع نمره واقعی و نمره خطا توانسته‌اند تغییراتی در مبانی نظری آن بوجود آورند. برای مثال، برخی متخصصان توانسته‌اند اجزای جداگانه‌ای برای نمره خطا چون خطای نمره‌گذار، خطایی که از به کار بردن انواع مختلف سؤال ناشی می‌شود، و خطای اجرای آزمون پیدا کنند و سپس مطالعاتی برای ارزیابی این عوامل و اثراتی که می‌توانند بر روی مقدار واریانس و پایایی داشته باشند، مورد بررسی قرار دهند (شاولسون^۱ ۱۹۹۱ و لوید^۲، ۱۹۸۸).
- مدلهای نظریه کلاسیک آزمون عمدتاً بر روی نمره کل آزمون تأکید می‌گذارد. بدین معنا به جای این که نمره‌های هر سؤال را به نمره واقعی ارتباط دهد، نمره کل آزمون را به نمره واقعی مربوط می‌سازد. لذا، از شاخصهای آماری چون دشواری و قدرت تمیز سؤال و ارتباطی که این شاخصها در فرایند ساختن آزمون با میانگین، انحراف معیار، و پایایی پیدا می‌کنند استفاده می‌شود تا این که آزمونهایی با ویژگیهای آماری مناسبی تهیه گردد. (میل‌مان و گرین^۳، ۱۹۹۸). با این که نقش این آماره‌ها در ساختن آزمونها با ارزشند، ولی چون کاملاً به نمونه‌ای از
- ^۱ - Shavelson
^۲ - Loyed
^۳ - Millman & Green

آزمودنیها وابسته‌اند محدودیت‌هایی در استفاده آنها فراهم می‌آید. از این آمارها زمانی می‌توان استفاده کرد که نمونه‌های آزمودنیها به جامعه‌ای که آزمون روی آنها ساخته شده مربوط باشند. چنانچه نمونه به طرق نامعلومی با جامعه متفاوت باشد، که امکان وقوع چنین کاری در آزمون‌سازی بسیار زیاد است. استفاده مطلوب از شاخصهای آماری را کاهش می‌دهد. (میل‌مان، ۱۹۹۸)

فواید مدل‌های نظریه کلاسیک آزمون در این است که بر مفروضه‌هایی استوارند که محاسبه آنها از طریق نمره‌های آزمون آسان است و سابقه طولانی در ساختن آزمونهای روانی و تربیتی دارند. به بیان دیگر، هم پارامترهایی که به فرد مربوط می‌شوند (نمره‌های واقعی) و هم پارامترهای سؤال (دشواری و تمیز) به آزمون و نمونه‌ای که آزمودنی در آن قرار می‌گیرد بستگی دارد. ویلسون^۱ (۱۹۹۴) معتقد است که چنین وابستگی‌هایی کاربرد شاخصهای آماری آزمودنی و سؤال را در فرایند ساختن آزمون محدود و کار هرگونه تحلیلی را پیچیده‌تر سازد.

نظریه سؤال - پاسخ

تقریباً ۵۰ سال پیش بود که لرد (۱۹۵۲) یکی از پیشگامان روان‌سنجی جدید این نظریه را مطرح کرد که نمره مشاهده شده و نمره واقعی آزمودنیها با نمره تواناییهای آنها مترادف نیستند. آنچه که لرد می‌خواهد بگوید این است که توانایی هر آزمودنی از نمره‌ای که او در آزمون کسب می‌کند مستقل است در حالی که نمره مشاهده شده و نمره واقعی فرد می‌تواند به نوع سؤالاتی که از حیطة مورد نظر انتخاب می‌گردد تفاوت پیدا کند. برای مثال، از قبل می‌دانیم موقعی که آزمودنیها با توانایی پایین در آزمون دشوار شرکت کنند نمره کمتر و چنانچه در آزمونهای آسان شرکت کنند نمره بالاتری خواهند گرفت. در حالی که توانایی این قبیل آزمودنیها در اندازه‌گیری خصیصه موردنظر معمولاً یکسان می‌باشد.

لرد معتقد است که تواناییها را می‌توان در اثر آموزش یا سایر عوامل تغییر داد ولی هنگامی که موقع ارزشیابی فرا می‌رسد، نمره توانایی هر آزمودنی در ارتباط با خصیصه مورد نظر قابل تعریف

¹ - Willson

و بررسی است و در پاسخ به نمونه‌هایی از سوالات که برای اندازه‌گیری حیطة مورد نظر ساخته شده بدون تغییر (مستقل) باقی می‌ماند. همبلتون و همکاران^۱ (۱۹۹۳). بر اساس مطالعه کارهای لرد نتیجه می‌گیرند که قاعداً مفهوم اساسی‌تری غیر از نمره واقعی آزمودنی در پاسخ دادن به سؤال آزمون دخالت دارد. آنها اضافه می‌کنند که این همان مفهوم اساسی‌تر است که اغلب باید در ارزشیابی مورد توجه قرار گیرد.

روان سنجان تلاش کرده‌اند پارامترهایی را برای سؤال توصیف کنند که نسبت به تغییراتی که در ترکیب نمونه‌ای از آزمودنیها به وجود می‌آید، و یا تحلیلی که بر روی آنها انجام می‌پذیرد، نامتغیر باقی بماند. دلیل این کار کاملاً روشن است. چنانچه پارامترهای سؤال نامتغیر باقی بمانند، آنها را می‌توان با استفاده از یک گروه از داده‌ها برآورد کرد و این برآوردها را می‌توان با اطمینان در مورد هر گروه دیگری از آزمودنیها، از جمله تمام جامعه آزمودنیها، به کار بست (انگلهارد^۲، ۱۹۹۴).

در حقیقت موضوع نامتغیر ماندن پارامترهای سؤال ابتدا توسط روان‌سنجان مشهوری چون تورندایک، تورستون و راش مورد مطالعه قرار گرفته است. هر یک از این روان‌سنجان اهمیت حذف اثر نمونه را در برآورد پارامترهای سؤال مورد تأکید قرار داده‌اند، ولی شیوه‌هایی که آنها برای حل این موضوع ارائه می‌کنند تا اندازه‌ای با هم تفاوت دارند. تورندایک و تورستون با قبول فرض نرمال بودن، نامتغیر ماندن پارامترهای سؤال را براساس توصیف توزیع توانایی گروهی و با استفاده محاسبه میانگین و انحراف معیار مورد تأکید قرار می‌دهند (انگلهارد، ۱۹۸۴).

راش موضوع نامتغیر ماندن پارامترهای سؤال (دشواری، تمیز) را مستقیماً بر تحلیل انفرادی سؤال قرار می‌دهد و بنابراین مدلهای پیشنهادی او نیاز به مفروضه‌های نرمال بودن توزیع ندارد. به عقیده راش، احتمال درست پاسخ دادن هر آزمودنی به سؤال آزمون به عنوان تابعی از تفاوت بین دشواری سؤال (P) و توانایی (θ)^۳ او بدست می‌آید و قاعداً باید برای

^۱ - Hambelton & et al.

^۲ - Engelhard

^۳ - teta

موضوع برزندگی داده‌ها با مدلی که راش آن را پیشنهاد می‌کند بیشتر مورد توجه روان سنجان قرار گرفته است و به نظر وایس و دیویدسون^۱ (۱۹۸۶)، یکی از مزایا مهم نظریه سؤال- پاسخ این است که اجازه می‌دهد تا دقت اندازه‌گیری با بررسی برزندگی پاسخ فرد با مدل مورد ارزشیابی قرار گیرد.

به عقیده کروکر و آلگینا^۲ (۱۹۸۶) از شاخصهای آماری که از تحلیل سؤال در نظریه کلاسیک آزمون بدست می‌آیند، مثل نسبت پاسخهای صحیح و همبستگی بین نمره هر سؤال با نمره کل آزمون، نمی‌توان انتظار داشت که از انتخاب افراد آزمودنی نامتغیر باقی بمانند. بر عکس، منابع زیادی را می‌توان پیدا کرد که گزارش می‌کنند پارامترهای سؤال در نظریه سؤال- پاسخ در انتخاب آزمودنیها نامتغیر باقی می‌ماند (گولیکسون، ۱۹۵۰، لرد، ۱۹۵۲، ۱۹۷۷، ۱۹۸۰، لرد ونویک، ۱۹۶۸، همبلتون و کوک^۳، ۱۹۷۷، انگلهارد، ۱۹۹۴).

به عبارت دیگر، برای روان سنجانی که برای نامتغیر ماندن شاخصهای آماری

هر نمونه محاسبه این برآوردها نامتغیر باقی بماند. به بیان دیگر، هدف راش ارایه مدلی بود که توسط آن بتواند اثر جامعه را حذف و برآوردهای نمونه را نامتغیر باقی نگاه دارد. همبلتون (۱۹۹۳) می‌گوید ترستون و تورندایک راههای خاصی برای مدل سازی پاسخ فرد به سؤال پیشنهاد نمی‌کنند و فقط نرمال بودن توزیع توانایی را برای مسأله نامتغیر ماندن برآوردها کافی می‌دانند.

در واقع تفاوتی که به نظر انگلهارد (۱۹۸۴) می‌توان بین ترستون، تورندایک، و راش در مورد نامتغیر ماندن پارامترهای سؤال به آن اشاره کرد این است که ترستون و تورندایک می‌خواهند به این سؤال- پاسخ دهند که آیا مدل برزننده داده‌ها است؟ در حالی که سؤالی که راش به آن پاسخ می‌دهد این است که آیا داده‌ها با مدل برزندگی دارند یا خیر؟ راش معتقد است که ابتدا باید به ساختن سؤالات آزمون پرداخت آن گاه آنهایی را که با مدل برزندگی دارند انتخاب کرد تا این که به اهداف اندازه‌گیری دست یافت. ترستون و تورندایک تأکید نسبی بیشتری بر روی داده‌ها دارند در حالی که راش بیشتر بر روی مدل تأکید می‌گذارد.

^۱ - Weiss & Davidson

^۲ - Crocker & Algina

^۳ - Cook

سؤال و توانائی آزمودنی اهمیت قائلند راه حل قضیه را در مفاهیم و مدل‌های ارایه شده در نظریه سؤال- پاسخ جستجو می‌کنند. بنابراین، نظریه سؤال- پاسخ و مدل‌های مربوطه و کاربردهای آنها بیش از ۴۰ سال است که مورد مطالعه و بررسی قرار گرفته و در ربع آخر قرن بیستم به طور گسترده‌ای در ساخت آزمون‌ها تربیتی و روانی به کار برده شده است (همبلتون و جونز^۱، ۱۹۹۳).

چنانچه بخواهیم نظریه جدید را براساس مفاهیم و اجزاء تشکیل دهنده‌اش مورد بررسی قرار دهیم، نظریه سؤال- پاسخ یک نظریه آماری است که چگونگی عملکرد آزمودنی را به توانایی او که به وسیله سؤالات آزمون اندازه‌گیری می‌شود ارتباط می‌دهد (کروکر، آلگینا، ۱۹۸۶). در این نظریه، سازنده آزمون چنین فرض می‌کند که می‌توان احتمال پاسخهایی را که به سؤالات یک آزمون داده می‌شود روی پیوستار توانائی تعریف کرد. در این نظریه چنین فرض می‌شود که توانائی واحدی برای پاسخ به سؤالات در یک آزمون وجود دارد. به عقیده کروکر و آلگینا رکن اصلی را در این نظریه یک مدل ریاضی تشکیل می‌دهد که چگونگی پاسخ آزمودنیها با سطوح مختلف توانایی را به سؤالات آزمون بیان می‌کند.

اطلاعاتی که این مدل ریاضی در اختیار سازنده آزمون می‌گذارد این است که به او اجازه می‌دهد تا عملکرد آزمودنیهایی را که در آزمونهای مختلف شرکت کرده‌اند با هم مقایسه کند. این مدل هم چنین اجازه می‌دهد تا نتایج تحلیل سؤالی را در به سایر گروه‌ها با سطوح توانایی مختلف، به غیر از گروهی که مورد مطالعه قرار گرفته است، تعمیم داد.

به نظر لوید (۱۹۸۸)، میچل^۲ (۱۹۹۴) نظریه سؤال- پاسخ به دلیل محدودیتهایی که در نظریه کلاسیک آزمون وجود دارد بوجود آمد. بر عکس نظریه کلاسیک، در نظریه سؤال- پاسخ واحد مورد نظر، به جای کل آزمون، هر سؤال آزمون است. متخصصان آزمون سازی طرفدار نظریه سؤال- پاسخ از یک مدل ریاضی برای برقراری ارتباط بین احتمال درست پاسخ دادن به سؤال و خصیصه مکنون (توانائی) استفاده می‌کنند. لوید و میچل معتقدند که

² - Michell

¹ - Jones

ویژگیهای این نظریه بدین معنی است که در اندازه‌گیری توانایی به نمونه‌ای از افراد که به سؤالات آزمون ویژه‌ای پاسخ می‌دهند نیازی نیست. اطلاعاتی که از پاسخ نمونه‌ای از آزمودنیها به مجموعه‌ای از سؤالات بدست می‌آید می‌توان برای برآورد پارامترهای مدل ریاضی نظریه سؤال به کار برد. هم چنین پس از برآورد این مقادیر، چنین فرض می‌شود که مجموعه سؤالات و نمونه‌ای از آزمودنیهای اصلی محدودیت‌هایی در کار اندازه‌گیری فراهم نمی‌آوردند. در عوض، زیر مجموعه‌ای از سؤالات را می‌توان به گروه دیگری از آزمودنیها تعمیم داد و از پاسخها می‌توان برای تعیین محل هر آزمودنی روی مقیاس خصیصه مکنون استفاده کرد. این مفروضه چنین پیشنهاد می‌کند که ویژگیهای سؤال تنها در مورد نمونه‌ای از آزمودنیها یا در مورد مجموعه‌ای از سؤالات آزمون اصلی کاربرد ندارند، موضوعی که در نظریه کلاسیک آزمون صحت دارد.

ویژگی مدل‌های نظریه سؤال - پاسخ

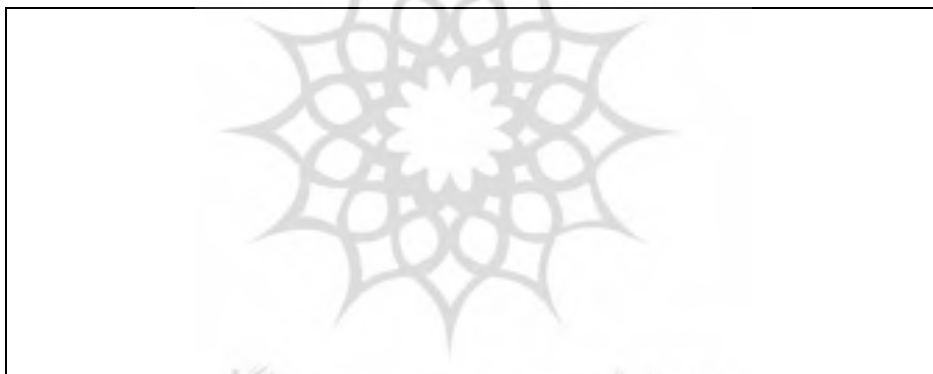
در حالی که تا اندازه‌ای با مفاهیم نظریه جدید آزمون سازی آشنایی پیدا کردیم،

حال برای آگاهی بیشتر به بررسی ویژگیها، مفروضه‌ها، و مدل‌های این نظریه می‌پردازیم. در بسیاری از موقعیتهای اندازه‌گیری روانی و تربیتی یک متغیر زیربنایی وجود دارد که مایل به اندازه‌گیری آن هستیم. برای مثال، از هوش، توانایی خواندن، توانایی حل مسأله و جزء اینها می‌توان به عنوان چنین متغیرهایی نام برد. روان سنجان به هر یک از این متغیرها، سازه، خصیصه مکنون، یا به طور کلی توانایی می‌گویند. هدف اصلی در اغلب اندازه‌گیریهای روانی و تربیتی این است تعیین کند که هر فرد چقدر از این توانایی دارا می‌باشد.

فرض قابل قبول این است هر آزمودنی که به سؤال آزمونی پاسخ می‌دهد مقداری از این توانایی برخوردار می‌باشد. بنابراین، می‌توان فرض کرد که هر آزمودنی یک ارزش عددی، یک نمره دارد که موقعیت او را در امتداد پیوستار توانایی معین می‌سازد. به این خصیصه مکنون یا توانایی به یونانی θ می‌گویند. در هر سطحی از توانایی احتمال معینی وجود دارد که هر آزمودنی با آن توانایی بتواند به سؤال - پاسخ درست بدهد. به این احتمال، $P(\theta)$ گفته می‌شود. در پاسخ به یک سؤال

آزمون، این احتمال برای آزمودنیها با توانایی پائین کم و برای آزمودنیها با توانایی بالا زیاد خواهد بود. حال اگر $P(\emptyset)$ را به عنوان تابعی از توانایی به صورت منحنی رسم کنیم، حاصل کار منحنی به شکل S در خواهد آمد. این

منحنی S شکل ارتباط بین احتمال پاسخ درست به سؤال و پیوستار توانایی را توصیف می‌کند و در نظریه سؤال-پاسخ به آن منحنی ویژگی سؤال می‌گویند. در آزمون هر سؤال منحنی ویژگی خاص خود را دارد. به علاوه، به چنین



شکل ۱. منحنی ویژگی سؤال

منحنی به دلیل تراکمی بودن آن منحنی اوجیو نرمال^۱ نیز گفته می‌شود. از این منحنی در آغاز پیدایش نظریه سؤال-پاسخ استفاده می‌شد (لالی، ۱۹۵۰؛ لرد، ۱۹۵۲).

شکل ۱ چند ویژگی مهم یک منحنی اوجیو نرمال را به شرح زیر نشان می‌دهد:

۱. از چپ به راست، منحنی قوس صعودی پیدا می‌کند.
۲. ضلع چپ منحنی به صفر نزدیک می‌شود ولی هرگز به آن نمی‌رسد. ضلع بالای منحنی به یک می‌رسد.
۳. منحنی اوجیو نرمال مستقیماً به منحنی نرمال استاندارد مربوط می‌شود. می‌توان از جدول نمره استاندارد Z برای محاسبه هر نمره برای تعیین مساحت در سمت چپ آن استفاده کرد. منحنی نرمال اوجیو استاندارد نیز می‌تواند از این ویژگیها استفاده کند.

¹ - ogive normal curve

مدلهای منطقی

با این که از منحنی اوجیو نرمال در اوایل آغاز کار نظریه سؤال- پاسخ بسیار استفاده می‌شد، ولی امروزه جای خود را به مدل‌های منطقی (سه مدل) که به محاسبات آسان‌تری نیاز دارند داده است. برای هر یک از این سه مدل، توزیع تابع منطقی تراکمی^۱ به عنوان اساسی برای توصیف منحنی ویژگی سؤال به کار برده می‌شود. فرمول این تابع به این صورت نوشته می‌شود:

$$p_g(\theta) = \frac{e^x}{1+e^x}$$

در این فرمول، e پایه سیستم لگاریتم طبیعی^۲ است x، یک نماد^۳ دلخواه است و به نمره مشاهده شده اطلاق می‌شود. در هر یک از سه مدل منطقی که در زیر می‌آیند، منحنی ویژگی سؤال^۴ بسته به تغییراتی که در فرمول بالا داده می‌شود توصیف می‌گردد. این سه مدل به واسطه

موقعی که از منحنی اوجیو نرمال به عنوان ویژگی خصیصه سؤال استفاده می‌کنیم اعداد روی محور افقی مقادیر ممکن را برای خصیصه مکنون، θ بیان می‌کنند. ارتفاع منحنی در بالای هر مقدار از θ نسبت آزمودنی‌هایی را که در آن سطح توانایی می‌توانند به سؤال- پاسخ درست دهند نشان می‌دهد. معادله منحنی اوجیو نرمال برای ویژگی خصیصه سؤال به صورت زیر نوشته می‌شود.

$$p_g(\theta) = \int_{-\infty}^w f(z) dz$$

در این معادله $P_g(\theta)$ نسبت آزمودنی‌ها با توانایی مکنون θ است که می‌توانند به سؤال g پاسخ درست بدهند است. مقادیر سمت راست، مقادیر تراکمی را برای اوجیو نرمال نشان می‌دهد و بدان معنی است که مساحت بین h - و w زیر منحنی اوجیو باید محاسبه شود. مقدار w یک عدد حقیقی است و برابر است با $w=a(\theta-b)$ ، در این معادله a و b به ترتیب پارامترهای تمیز و دشواری سؤال هستند. پارامترهای a و b با پارامترهای دشواری و تمیز در نظریه کلاسیک آزمون یکسان نیستند (داگلاس و همکاران، ۱۹۷۹).

¹ - cumulative logistic distribution function

² - natural logarithm system

³ - symbol

⁴ - item characteristic curve (ICO)

مدل یک - پارامتری

مدل یک - پارامتری منطقی را می‌توان حالت خاصی از مدل دو - پارامتری دانست که در آن تمام سؤالات فرض می‌شود پارامتر تمیز یکسانی دارند. چون تمام سؤالات پارامتر تمیز مشابهی دارند، به جای این که آن را به صورت متغیر g نشان داد با ثابت a معلوم می‌سازیم. معادله مدل یک - پارامتری منحنی ویژگی سؤال به شرح زیر است:

$$P_g(\theta) = \frac{e^{D_a(\theta - b_g)}}{1 + e^{D_a(\theta - b_g)}}$$

در این فرمول نسبت آزمودنیهایی که به سؤال - پاسخ درست می‌دهند تابعی است هم از توانایی آزمودنیها و هم از دشواری هر سؤال. مدل یک - پارامتری به مدل راش¹ شهرت دارد.

مدل سه - پارامتری

مشکلی که در کاربرد مدل‌های منطقی یک و دو - پارامتری و منحنی اوجیونرمال بروز می‌کند زمانی است که داده‌ها از سؤالات آزمون چهار - گزینه‌ای و صحیح

تعداد پارامترهایی که به سؤال مربوط می‌سازند از هم متمایز می‌گردند.

مدل دو - پارامتری

چون این مدل بسیار به منحنی اوجیونرمال شبیه است ابتدا به شرح آن می‌پردازیم. در مدل دو - پارامتری مقدار X را می‌توان از فرمول زیر محاسبه کرد.

$$X = D_{ag}(\theta - b_g)$$

در حالی که D یک مقدار ثابت و مقدار آن برابر $1/7$ تعیین گردیده است. چنانچه D برابر $1/7$ باشد، منحنی اوجیونرمال و مدل‌های منطقی برای هر نمره بیش از $0/01$ با هم تفاوت پیدا نخواهند کرد (لرد و نوویک، ۱۹۶۸). پارامترهای a_g و b_g نقش مشابهی در مدل‌های منطقی مثل منحنی اوجیونرمال دارند. با انجام جایگزینی برای x در فرمول اصلی مدل منطقی، به یک فرمول کلی‌تر برای مدل دو - پارامتری خواهیم رسید.

$$P_g(\theta) = \frac{e^{D_{ag}(\theta - b_g)}}{1 + e^{D_{ag}(\theta - b_g)}}$$

¹ - Rasch Model

همان طور که گفته شد، مدل‌های خصیصه مکنون ارتباط بین کمیت‌های قابل مشاهده و غیرقابل مشاهده را به عنوان یک تابع ریاضی توصیف می‌کنند. بنابراین، مدل‌های این نظریه عمدتاً مدل‌های ریاضی هستند.

در صفحات قبلی به دلیل آشنایی بیشتر خوانندگان درباره نظریه کلاسیک آزمون به اختصار بحث کردیم ولی مبانی، مفروضه‌ها و مدل‌های نظریه سؤال-پاسخ مفصل‌تر مورد بحث و بررسی قرار گرفت. اشاره گردید که نظریه سؤال-پاسخ از منحنی اوجیونرمال و مدل‌های منطقی برای تعیین احتمال پاسخ درست آزمودنی به هر سؤال آزمون و تعیین جایگاه وی روی پیوستار توانایی استفاده می‌کند. برای آگاهی بیشتر علاقه‌مندان به نظریه جدید، مطالعه همبلتون و همکاران (۱۹۷۸)، بیکر^۱ (۱۹۸۵)، همبلتون و جونز (۱۹۹۳) انگلهارت (۱۹۸۴)، رایت و استون^۲ (۱۹۸۰)، و لرد (۱۹۷۷) توصیه می‌شود.

حال در صفحات باقی مانده قصد داریم مقایسه‌ای بین نظریه کلاسیک آزمون و نظریه سؤال-پاسخ انجام دهیم و طی آن

غلط حاصل می‌گردد. دلیل بروز این مشکل این است که این نوع سؤالات آزمون به آزمودنی اجازه حدس زدن در پاسخ به سؤالات را می‌دهد. در مدل‌های یک و دو-پارامتری با کاهش θ مقدار $P_g(\theta)$ به صفر نزدیک می‌شود. ولی، برای آزمودنی‌هایی که در توانایی بسیار پایین هستند، می‌توان انتظار داشت که نسبت پاسخ درست بزرگتر از صفر باشد برای این که این قبیل آزمودنی‌ها می‌توانند پاسخ‌های صحیح را از راه شانس علامت بزنند. برای رفع این دشواری، از مدل منطقی سه-پارامتری می‌توان استفاده کرد. معادله برای منحنی ویژگی سؤال در این مدل به شرح زیر است:

$$P_g(\theta) = C_g + \frac{e(1-C_g)e^{D_{ag}(\theta-b_g)}}{1+e^{D_{ag}(\theta-b_g)}}$$

پارامتر C_g ، پارامتر حدس زدن نامیده می‌شود.

به طور خلاصه، خصیصه مکنون یا توانایی و منحنی ویژگی سؤال دو مفهوم اساسی در نظریه سؤال-پاسخ بشمار می‌روند که دانستن آنها برای علاقه‌مندان به مبانی نظریه جدید لازم و ضروری است.

^۱ - Baker

^۲ - Wright & Stone

سؤالات و نمره‌های خصیصه مکنون محاسبه می‌شود و پارامتر تمیز در نظریه سؤال- پاسخ تقریباً ارتباط نزدیکی وجود دارد. این ارتباط را لرد به شرح زیر نمایش می‌دهد:

$$a_i = \frac{r_i}{\sqrt{1-r_i^2}}$$

در این معادله a_i مقدار پارامتر تمیز برای سؤال i در نظریه سؤال- پاسخ و r_i ضریب همبستگی دو رشته‌ای سؤال در نظریه کلاسیک آزمون است.

به دلیل تفاوتی که در توزیع و اطلاق نمره‌ها در دو نظریه وجود دارد، این فرمول به جای یک ارتباط واقعی یک ارتباط تقریبی را نشان می‌دهد. فراوانی نمره‌های پاسخ درست (X) در نظریه کلاسیک آزمون و فراوانی نمره‌های توانایی در نظریه سؤال- پاسخ توزیع‌هایی با شکلهای متفاوت درست می‌کنند و بنابراین می‌توان گفت ارتباط بین X و \emptyset غیرخطی است. به علاوه، نمره کل آزمون X با خطاهای اندازه‌گیری همراه است در حالی که نمره‌های توانایی در نظریه سؤال- پاسخ با چنین خطاهائی همراه نیستند (لرد، ۱۹۸۰).

شایستگیها و محدودیتهای هر یک را در حل مسائل و مشکلاتی که در ساختن آزمونهای روانی و تربیتی وجود دارد مورد بررسی قرار دهیم. تاکنون مقالات زیادی مقایساتی بین دو نظریه انجام داده‌اند. آنچه در زیر می‌خوانید فشرده‌ای است از تفاوت‌های بین این دو نظریه آزمون سازی.

مقایسه نظری کلاسیک و نظریه

سؤال- پاسخ

در هر دو نظریه کلاسیک و سؤال- پاسخ از شاخصهای دشواری و تمیز برای برآورد پارامترهای سؤال استفاده می‌شود. با این که شیوه محاسبه این آماره‌ها متفاوت و هر یک مفاهیم جداگانه‌ای را در دو نظریه ارائه می‌دهند، ولی لرد (۱۹۸۰) ارتباط بین پارامترهای دشواری و تمیز در مدل نظریه کلاسیک آزمون و مدل منطقی دو- پارامتری در نظریه سؤال- پاسخ را مورد مطالعه قرار داد. وی نشان می‌دهد که تحت شرایط معینی (مثل موقعی که عملکرد تحت تأثیر حدس قرار نمی‌گیرد)، بین پارامتر تمیز در نظریه کلاسیک آزمون که از طریق محاسبه و یک همبستگی دو رشته‌ای بین نمره

آزمودنیهای همگن بدست خواهند داد. در زمینه شاخصهای دشواری، مقادیر بزرگتر از نمونه‌های آزمودنی با توانایی بالای متوسط و مقادیر کوچکتر از نمونه‌های آزمودنی با توانایی پایین‌تر از متوسط بدست خواهند آمد (همبلتون، ۱۹۸۹).

محدودیت دیگر نظریه کلاسیک آزمون این است که نمره‌هایی که از کاربرد این نظریه بدست می‌آیند کاملاً به آزمون بستگی دارند. نتیجتاً، دشواری سؤال مستقیماً بر نتایج نمره‌های آزمون تأثیر می‌گذارد. روان‌سنجان این را یک محدودیت مهم در ساخت آزمونهای روانی و تربیتی بشمار می‌آورند؛ برای این که در اندازه‌گیری اغلب لازم است عملکرد آزمودنیها از یک جامعه واحد را با استفاده از نتایجی که از سؤالات آزمونهاى ديگرى بدست می‌آید، مورد مقایسه قرار گیرند. در واقع، نظریه کلاسیک آزمون را می‌توان یک نظریه آزمون-محوری و نظریه سؤال-پاسخ را می‌توان یک نظریه سؤال-محوری دانست.

در مدل نمره واقعی که نظریه کلاسیک آزمون بر آن مبتنی است، امکان بررسی الگوهای پاسخ آزمودنیها به سؤالات

همان طور که قبلاً نیز در نظریه کلاسیک آزمون اشاره شد، شاید مهمترین تفاوت بین نظریه‌های کلاسیک آزمون و سؤال-پاسخ ویژگی نامتغیر بودن پارامترهای سؤال و پارامترهای توانایی در نظریه سؤال-پاسخ است. از این ویژگی دو نتیجه می‌توان استنباط کرد. ۱. پارامترهایی که مربوط به آزمودنی است از سؤالات آزمون مستقل هستند و ۲. پارامترهایی که ویژگی هر سؤال را معین می‌کند از توزیع توانایی مجموعه‌ای از آزمودنیها مستقل هستند (همبلتون، سوامی ناتان و راجرز^۱، ۱۹۹۳).

به علاوه نظریه کلاسیک آزمون در مقایسه با نظریه سؤال-پاسخ محدودیتهای دیگری نیز دارد. برای مثال، دو شاخص آماری (دشواری و تمیز سؤال) که در تحلیل سؤال در نظریه کلاسیک آزمون جایگاه خاصی دارند، کاملاً به نمونه‌یی از آزمودنیها که این مقادیر از آنها استخراج شده بستگی دارند. در زمینه شاخصهای تمیز، مقادیر بزرگتر در نمونه‌ای از آزمودنیهای ناهمگن و مقادیر کوچکتر، در نمونه‌ای از

¹ - Hambelton, Sawminatan, Rogers

آزمون به طور انفرادی وجود ندارد. در نتیجه، اساسی برای پیش بینی این که چگونه گروهی از آزمودنیها به سؤال خاص آزمونی پاسخ می‌دهند وجود ندارد. در مقابل، نظریه سؤال- پاسخ به متخصصان آزمون سازی انعطاف بیشتری می‌دهد تا تفسیرهای گسترده‌تری در مورد سطح هر سؤال انجام دهند و احتمال پاسخ درست هر آزمودنی را به هر سؤالی تعیین کنند. بنابراین، چنانچه سازنده آزمون بخواهد آزمونی بسازد که تنها با ویژگیهای جامعه آزمودنی مورد مطالعه ارتباط نداشته باشد (یعنی جامعه- آزاد باشد)، این کار تنها از طریق نظریه سؤال- پاسخ مقدور است (همبلتون و همکاران، ۱۹۸۹).

برای مثال، چنانچه بخواهیم آزمونی بسازیم تا دانش‌آموزان با توانایی ضعیف را برای شرکت در یک برنامه اصلاحی شناسایی کند و یا اگر قصدمان تمیز بین دانشجویان برتر به منظور اعطای بررسی تحصیلی است، نیاز به ساختن چنین آزمونهایی امری عادی است. در این موارد می‌توان از این ویژگی نظریه سؤال- پاسخ در ساختن آزمونهای انطباقی نیز استفاده کرد.

مدلهای نظریه سؤال- پاسخ نیز با محدودیتهایی نیز همراه هستند. به لحاظ تکنیکی، مدل‌های نظریه سؤال- پاسخ بسیار پیچیده‌اند و در عمل مشکلاتی برآورد پارامترها توسط مدل بروز می‌کند. برازندگی داده‌ها با مدل نیز می‌تواند مشکل آفرین باشد. هنوز کاملاً روشن نیست که چگونه مسأله برازندگی مدل را می‌توان حل کرد به ویژه مشکلاتی که در ارتباط با پارامترهای آزمون (دشواری، تمیز و حدس) مطرح می‌شود. به علاوه، به لحاظ عملی چنین به نظر می‌رسد برآورده کردن مفروضه‌های نظریه سؤال- پاسخ در ساخت آزمونهای روانی و تربیتی از پیچیدگی بیشتری نسبت به برآورد کردن مفروضه نظریه کلاسیک آزمون برخوردار باشد. برای مثال، تأمین نیازهای مدل یک- پارامتری در نظریه سؤال- پاسخ مطمئناً از سایر مدل‌های این نظریه آسان‌تر است.

آگاهی از محدودیتهای نظریه کلاسیک آزمون و مزایای بالقوه نظریه سؤال- پاسخ موجب شده است بسیاری از متخصصان آزمون سازی کارهای خود را در چارچوب نظریه سؤال- پاسخ انجام دهند. دلیل این تغییر روش به وسیله روان

- سنجان و متخصصان آزمون سازی از مدل‌های نظریه کلاسیک آزمون به مدل‌های نظریه سؤال- پاسخ به واسطه کسب نتایج مهمی است که در اثر به کار بستن مدل‌های نظریه سؤال- پاسخ در حل مسائل اندازه‌گیری حاصل شده است. در زیر به برخی از مزایای دو نظریه به طور جداگانه اشاره می‌شود.
- مزایای استفاده نظریه سؤال- پاسخ به شرح زیرند.
۱. شاخصهای آماری از گروهی که این برآوردها گرفته شده مستقل است.
 ۲. نمره‌های پیشرفت آزمودنی به دشواری کل آزمون بستگی ندارد.
۳. مدل‌های آزمون، اساسی برای جور کردن سؤالات با سطوح مختلف توانایی فراهم می‌آورند.
۴. مدل‌های آزمون نیازی به آزمونهای هم‌تا برای محاسبه پایایی ندارند.
- مزایایی که به واسطه کار بستن مدل‌های نظریه کلاسیک آزمون برای رفع مشکلات اندازه‌گیری بدست می‌آید عبارتند از:
۱. تحلیل سؤالات به نمونه‌های کوچکتری نیاز دارد.
 ۲. تحلیل‌های ریاضی با نظریه سؤال- پاسخ یکسان است.

جدول ۱. تفاوت‌های بین مدل‌های نظریه‌های کلاسیک آزمون و سؤال- پاسخ

موضوع	نظریه کلاسیک	نظریه سؤال- پاسخ
مدل	خطی	غیرخطی
سطح	آزمون	سؤال
مفروضه‌ها	ضعیف (آسان محاسبه می‌شوند)	قوی (دشواری محاسبه می‌شوند)
رابطه سؤال با توانایی	روشن نیست	از طریق منحنیهای ویژگی سؤال حاصل می‌شود
توانایی	نمره‌های آزمون یا نمره‌های واقعی	نمره‌های توانایی روی پیوستار
مقیاس	روی مقیاس نمره آزمون یا هر مقیاس استاندارد گزارش می‌شود	+ تا یا هر مقیاس استاندارد گزارش می‌شود
نامتغیر بودن آماره‌ها	خیر- پارامترهای سؤال و فرد	بلی - پارامترهای سؤال و فرد
سؤال و فرد	به نمونه بستگی دارند	از نمونه مستقلند اگر مدل با داده‌ها برازندگی پیدا کند
آماره‌های سؤال و فرد	p, r	c, a, b برای مدل سه- پارامتری
حجم نمونه برای برآورد پارامتر	200 تا 500	به مدل بستگی دارد، ولی به نمونه‌های بزرگتری نیاز دارد

۳. برآورد پارامترهای مدل به لحاظ مفهومی بسیار آسان است. همان طور که اشاره شد، مهمترین رویکردی که تاکنون مورد استفاده سازندگان آزمون در کشور قرار گرفته بر اساس نظریه کلاسیک آزمون و عمدتاً بر مفهوم پایائی آزمون استوار بوده است. بر اساس این نظریه، سازنده آزمون بهترین مجموعه همگن از سؤالات را که به اندازه‌گیری یک سازه زیربنای رفتار موردنظر می‌پردازد تهیه می‌کند. برای

۴. برای اطمینان از برازندگی داده‌ها، تحلیل سؤالات به مدل نیاز ندارد. جدول ۱ تفاوت‌های بین مدل‌های کلاسیک آزمون و سؤال-پاسخ را نشان می‌دهد.

خلاصه

هدف از این مقاله بررسی شکل‌گیری نظریه جدید آزمون سازی و نشان دادن مفروضه‌ها، مدل‌ها، و کاربرد نظریه سؤال-پاسخ و نقشی که آنها در پیدا کردن راه‌حلهایی برای مسائل و مشکلاتی که بر سر راه ساختن آزمونهای روانی و تربیتی وجود دارد می‌باشد. علاوه بر این، مقایساتی بین دو نظریه کلاسیک آزمون و نظریه سؤال-پاسخ انجام گرفت و به نظر متخصصان آزمون‌سازی جدید، به برتری نظریه سؤال-پاسخ نسبت به نظریه کلاسیک آزمون در حل مسائل و مشکلاتی که آزمون سازی با آنها روبرو است اشاره گردید.

ساختن آزمون یا مقیاسی با این ویژگی، از بین مجموعه‌ای از سؤالات آنهایی را که ثبات درونی بیشتری با همدیگر دارند انتخاب می‌گردد. در عمل، این قبیل سؤالات از طریق محاسبه ضریب همبستگی سؤالات، نمره کل آزمون و آن گاه حذف سؤالاتی که ضرایب همبستگی معنی‌داری ندارند انجام می‌پذیرد. به طور نظری و بر اساس مدل کلاسیک آزمون سؤالاتی که باقی ماند بهترین مقیاس همگن را برای اندازه‌گیری ویژگیهای رفتاری تشکیل می‌دهد.

نظریه دیگری که در نیمه دوم قرن بیستم به دلیل محدودیتهایی که در نظریه

توانایی‌شان جور است، روشن می‌سازد. به علاوه، می‌توان توانایی آزمودنیها را مستقل از یکسان کردن نمره‌های آزمونهاى مختلف، که توانایی مشابهی را اندازه می‌گیرند، ممکن ساخت. بنابراین، مشکل ساختن فرمهای همتایی برای آزمونها از میان برداشته می‌شود.

ویژگی دیگری که در نظریه سؤال-پاسخ وجود دارد و در نظریه کلاسیک آزمون مشاهده نمی‌شود این است این امکان وجود دارد تا دقت برآوردهای توانایی را در هر نقطه روی پیوستار اندازه‌گیری کرد. بنابراین، به جای محاسبه خطای معیار اندازه‌گیری که در مورد همه آزمودنیها یکسان به کار برده می‌شود، بدون توجه به نمره‌های آزمون، امکان برآوردهای جداگانه‌ای برای هر آزمودنی یا در هر سطحی از توانایی روی پیوستار با استفاده از مدل‌های نظریه سؤال-پاسخ وجود دارد.

علاوه بر این، پارامترهای سؤال بر اساس مدل‌های نظریه سؤال-پاسخ در مورد تمام زیر مجموعه‌هایی از آزمودنیها نامتغیر باقی می‌ماند و لذا، امکان مقایسات را روی جامعه‌های متفاوت مورد علاقه مقدور می‌سازد و بنابراین از کاربرد بالائی

قدیمی وجود داشت شکل گرفت و توسط روان سنجان و متخصصان آزمون‌سازی به عنوان جانشینی برای نظریه کلاسیک آزمون ارایه گردید نظریه سؤال-پاسخ نام گرفت. همان طور که اشاره شد، مفروضه‌ها و مدل‌های نظریه سؤال-پاسخ نسبت به مفروضه‌ها و مدل‌های کلاسیک آزمون مزایا بیشتری دارند. شاید مهمترین آنها این باشد که می‌توان توانایی آزمودنیها را روی پیوستار توانایی مشابهی که از زیر مجموعه‌ای از سوالات ساخته شده، در صورتی که با مدل برازندگی پیدا کنند، برآورد کرد. این بدان معنی است که توانایی هر آزمودنی را می‌توان مستقل از هر آزمون یا هر تعداد سؤال برآورد کنیم و در نتیجه گشایش قابل توجهی در حیطه اندازه‌گیریهای روانی و تربیتی به وجود آمده است. آنچه از این ویژگی نظریه سؤال-پاسخ استنباط می‌شود این است که آزمودنیها را می‌توان باهم مقایسه کرد حتی اگر آنها به زیر مجموعه متفاوتی از سوالات آزمون پاسخ داده باشند. این ویژگی کاربرد مدل‌های نظریه سؤال-پاسخ را در ساختن آزمونهاى انطباقی جایی که آزمودنیها سوالاتی دریافت می‌کنند که با سطح

برای سازندگان آزمونهای روانی و تربیتی برخوردار است. نامتغیر ماندن پارامترهای سؤال بویژه در تهیه بانکی از سؤالات و جلوگیری از سوگیری از اهمیت مهمی برخوردار است.

نهایتاً، اشاره گردید که نظریه سؤال-پاسخ راه‌های دقیق و منطقی برای ساختن آزمونهای روانی و تربیتی پیشنهاد می‌کند. عوامل مهمی که تاکنون استفاده از این نظریه را محدود ساخته است عدم آشنایی سازندگان آزمون کشور با مبانی نظری، مفروضه‌ها، و مدلها، و برنامه‌های

نرم‌افزاری آن می‌باشد. در سالهای اخیر این مشکل تا اندازه زیادی با انتشار مقالات در مجلات علمی و چاپ کتب برطرف شده است. بنابراین، می‌توان انتظار داشت تا مبانی نظری، مفروضه‌ها و مدلهای نظریه سؤال-پاسخ به عنوان روشهایی جدید در آینده در ساختن آزمونهای روانی و تربیتی مورد استفاده سازندگان آزمون در کشورمان قرار گیرد.