

## مقدمه‌ای بر کاربرد نظریه سوال-پاسخ در پژوهش‌های مدیریتی (مطالعه موردی، کیفیت خدمات شرکت‌های بیمه)

قاسم کشاورز گرامی\*، محمود خواجه‌پور\*\*

### چکیده

هدف اصلی از انجام پژوهش حاضر معرفی نظریه سوال-پاسخ و کاربرد آن در پژوهش‌های مدیریتی است که در این راستا کیفیت خدمات به عنوان یکی از متغیرهای ممکن انتخاب شد. به دلیل اهمیت کیفیت در صنایع خدماتی و آثار برجسته آن بر رضایتمندی مشتریان، همواره این سؤال مطرح می‌شود که چگونه می‌توان کیفیت خدمات را ارزیابی کرد؟ در این راستا سنجش کیفیت خدمات شرکت بیمه بر اساس نظریه سوال-پاسخ انجام پذیرفت. برای دستیابی به این هدف ۳۰۰ نفر از مشتریان یک شرکت بیمه باروش تصادفی ساده به عنوان گروه نمونه انتخاب شدند. برای سنجش کیفیت خدمات از پرسشنامه سروکوال استفاده شد. مقدار آلفای کرونباخ این پرسشنامه ۰,۹۴۱ و مقدار ضریب اعتبار حاصل از تئوری سوال-پاسخ با روش حاشیه‌ای ۰,۹۶۲ بدست آمد که حکایت از اعتبار بالای این تست دارد. تحلیل عاملی با روش محورهای اصلی نشان داد که عامل اول با مقدار ویژه ۱۰,۰۶ توانایی تبیین ۴۳,۷۲ درصد از کل نمرات پرسشنامه را دارد و با عامل دوم که مقدار ویژه آن ۱,۵۹ است فاصله زیادی دارد که نشان‌دهنده برقراری مفروضات نظریه سوال-پاسخ است. برای انجام محاسبات در این پژوهش از مدل پاسخ درجه‌بندی نظریه سوال-پاسخ در برنامه مولتی‌لوگ استفاده شده است. نتایج محاسبات ویژگی‌های روانسنجی نشان داد که ۲۲ سؤال پرسشنامه در سنجش کیفیت خدمات ویژگی‌های روانسنجی مناسبی دارند. بر اساس امکانات و توانایی‌های بالای سنجش مشاهده شده در

\* دانشجوی دکتری سنجش و اندازه‌گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه علامه طباطبائی

GH.KZ.59@Gmail.com.

\*\* دانشجوی دکتری مدیریت تولید و عملیات، دانشکده مدیریت و حسابداری، دانشگاه شهید بهشتی

نظریه سؤال- پاسخ پیشنهاد شد که در پژوهش‌های مدیریتی از این نظریه در سنجش صفات و ویژگی‌ها استفاده شود.

**واژگان کلیدی:** نظریه سؤال-پاسخ، پژوهش در مدیریت، کیفیت خدمات

**مقدمه**

رویکرد نظریه سوال-پاسخ<sup>۱</sup> در سال‌های اخیر مورد توجه روز افزون محققان قرار گرفته و یک روش جایگزین مدرن و برتر برای ارزیابی نظریه کلاسیک اندازه‌گیری محسوب می‌شود اما در بسیاری از رشته‌ها مانند مدیریت خبری از این جایگزینی نیست و آشنایی با این رویکرد اندازه‌گیری اتفاق نیافتاده است. از سوی دیگر، مصرف‌کنندگان دیگر حاضر به پذیرش هر نوع کالا یا خدمتی نیستند و از این رو سازمان‌هایی از دور رقابت خارج می‌شوند که نتوانند انتظارات مشتریان خود را تأمین کنند. در حال حاضر بخش خدمات، سهم عمده‌ای در بازار بسیاری از کشورها دارد. برای مثال، کاتلر (۲۰۰۳) بیان می‌کند که ۷۹ درصد از کل شاغلین، در بخش خدمات مشغول به کار هستند و ۷۶ درصد از تولید ناخالص ملی اقتصاد ایالات متحده از بخش خدمات به دست می‌آید (Kotler, 2003: 112). همچنین سازمان توسعه و همکاری اقتصادی<sup>۲</sup> اعلام کرده است که بیش از ۷۰ درصد از مردم سراسر جهان در بخش خدمات مشغول به فعالیت می‌باشند ( <http://ocde.P4.Siteinternet.Com>, 2005). بازتاب این تغییرات بر این نکته تأکید دارد که بازاریابی خدمات به یک زمینه مسلم برای بررسی علمی تبدیل شده است و امروزه یک پارادایم جایگزین برای بازاریابی کالاها ارائه می‌دهد (Lovelock & Gummesson, 2004). ماهیت پیچیده خدمات به علت ویژگی‌های آن یعنی: ناملموس بودن، فناپذیر بودن، درگیری زیاد مشتری، همزمانی تولید و مصرف و ناهمگون بودن می‌باشد (روستا و همکاران، ۱۳۸۷). این ویژگی‌ها به همراه افزایش سهم بخش خدمات، باعث افزایش نیاز به کیفیت بهتر خدمات برای شرکت‌هایی شده است که به دنبال بهبود عملکرد مالی و جذب مشتریان در محیط بسیار رقابتی می‌باشند (Wang, & Hui, 2003). همچنین محققین بر این موضوع اتفاق نظر دارند که یک مبحث مهم در بررسی خدمات، کیفیت خدمات است که یک عامل تعیین کننده حیاتی در عملکرد فعالیت تجاری و قابلیت سوددهی بلند مدت شرکت می‌باشد (Carrillat, 2007). دلیل این امر این است که کیفیت خدمات منجر به رضایتمندی مشتری می‌گردد که به نوبه خود، اثر مثبتی بر روی تبلیغات دهان به دهان، وفاداری نگرشی، و نیات خرید مشتری می‌گذارد (Gremier & Gwinner, 2000)

---

1 item-response theory

2 OECD

## ادبیات پژوهش

به دلیل تنوع خدمات، تعریف آنها همواره کاری دشوار بوده است. آنچه این امر را پیچیده‌تر می‌کند، این واقعیت است که به دلیل نامحسوس بودن اکثر داده‌ها و ستاده‌ها، غالباً درک و تشخیص راه‌های انجام و عرضه خدمات آسان نیست (لاولاک و رایت، ۱۳۸۲). کلمه خدمت وسعت و تنوع معانی وسیعی دارد و این ویژگی به ابهام زیادی در مفهوم آن در متون مدیریتی منتهی شده است (Johns, 1999). گرونروس<sup>۱</sup> (۲۰۰۱) خدمت را چنین تعریف می‌کند: "یک خدمت، فعالیت یا مجموعه‌ای از فعالیتهای کم و بیش ناملموس است که معمولاً و نه لزوماً در تعاملات بین مشتری و کارکنان خدمات و/ یا منابع فیزیکی یا کالاها و/ یا سیستمهای عرضه کننده خدمات صورت می‌گیرند، بلکه به عنوان راه‌حلهایی برای مشکلات مشتری ارائه می‌شوند" (Gronroos, 2001).

کیفیت از مشتری شروع می‌شود و هرگونه توجه به کالا یا خدمت، بدون توجه به نظر مشتری، الزاماً کیفیت را بدنبال ندارد (فیتز سیمونز<sup>۲</sup>، ۱۳۸۲). لذا معمولاً کیفیت را بر اساس تطابق‌پذیری محصول با ویژگی‌های مدنظر مشتری، و یا میزانی که محصول می‌تواند انتظارات قبلی مشتری را برآورده سازد، تعریف می‌کنند (Nilsson & Johnson, 2003).

گرونروس (۲۰۰۱) کیفیت خدمات را «اندازه مغایرت بین ادراک مشتری از خدمت و انتظارات او» تعریف کرده است (Gronroos, 2001). زیتهمل<sup>۳</sup> (۱۹۹۶) کیفیت خدمات را «قضایوت همه‌جانبه مشتری درباره ماهیت برتر خدمت نسبت به خدمات مشابه با مزیت‌های برجسته آن» تعریف کرده است ((Zeithaml, 1996) عثمان و اون<sup>۴</sup> در سال ۲۰۰۲ کیفیت خدمات را به عنوان درجه‌ای از اختلاف بین ادراک و انتظارات مشتریان از خدمات تعریف می‌کنند<sup>۱</sup> (Othman & Owen, 2002). پاراسورامان<sup>۵</sup> (۱۹۸۵) کیفیت خدمت را «سازگاری پایدار با انتظارات مشتری و شناخت انتظارات مشتری از خدمت خاص» تعریف کرده است ((Parasuraman, 1985).

در بخش خدمات، ارزیابی کیفیت آن حین فرایند ارائه خدمات انجام می‌گیرد. هر تماس مشتری به عنوان لحظه‌ای برای ایجاد اعتماد و فرصتی برای راضی یا ناراضی کردن به شمار می‌رود (فیتز سیمونز،

1 Gronroos, Christian

2 Fitzsimmons, James A

3 Zeithaml, Valarie

4 Othman, Abdulqawi & Owen, Lynn

5 Parasuraman, A

۱۳۸۲). از زمانی که نقش خدمات در زندگی روزمره آشکار شد، مقوله "کیفیت خدمات" نیز به عنوان مشخصه اصلی رقابت بین سازمان‌ها مورد توجه قرار گرفت؛ به طوری که توجه به کیفیت خدمات، سازمان را از رقابش متمایز ساخته و سبب کسب مزیت رقابتی می‌شود. با توجه به اهمیت کیفیت خدمات و تاثیرگذاری آن بر رضایت مشتری و کسب مزیت رقابتی برای سازمان‌ها، پژوهش‌های گسترده‌تر و دقیق‌تر در جهت شناخت بیشتر مولفه‌های مؤثر بر کیفیت خدمات امری لازم و ضروری می‌نمایاند. پژوهش در این زمینه نیازمند ابزارها و پرسشنامه‌هایی روا<sup>۱</sup> و پایا<sup>۲</sup> است که با نظریه‌های قوی و کارآمد سنجش و اندازه‌گیری ساخت و هنجاریابی شده باشند. در کشور ما اغلب پرسشنامه‌های موجود، خصوصاً در رشته‌ی مدیریت، بر اساس تئوری کلاسیک اندازه‌گیری، تجزیه و تحلیل و هنجاریابی می‌شوند، در حالی که تئوری‌های جدید سنجش و اندازه‌گیری مانند مدل‌های نظریه سؤال- پاسخ به شکل و شیوه اساسی‌تری می‌توانند به کمی‌سازی صفت‌های مکنون و ویژگی‌ها پردازند.

نظریه کلاسیک آزمون‌سازی نخستین بار در سال ۱۹۰۴ توسط چارلز اسپیرمن<sup>۳</sup> شاگرد گالتون<sup>۴</sup> مطرح شد. اوج تکامل این نظریه در دو کتاب «مبانی نظری آزمون‌های روانی» اثر گالیکسن<sup>۵</sup> در سال ۱۹۰۵ و «نظریه‌های آماری آزمون‌های روانی» لرد و نویک<sup>۶</sup> (۱۹۶۸) انعکاس یافته است. از همان سال‌ها تاکنون در مورد این نظریه و میزان کارایی آن پژوهش‌های بسیاری انجام گرفته است. هرکدام از این پژوهش‌ها از جنبه‌های مختلف این نظریه را مورد کنکاش قرار داده‌اند. از این میان می‌توان به پژوهش‌های همبلتون<sup>۷</sup>، لیندن<sup>۸</sup>، لرد<sup>۹</sup>، آلن و ین<sup>۱۰</sup> و همبلتون و کوک<sup>۱۱</sup> اشاره کرد که توانستند این نظریه را نقد کرده و ایرادهای وارد شده بر آن را مورد بررسی قرار دهند (ستاری، ۱۳۸۲).

- 
- 1 Valid
  - 2 Reliabel
  - 3 Spearman
  - 4 Galton
  - 5 H.Gulliksen
  - 6 Lord & Novik
  - 7 Hambleton
  - 8 Linden
  - 9 LOrd
  - 10Allen & Yen
  - 11Cook

متغیر بودن پارامترهای سؤال برای گروه‌های متفاوت در نمونه‌گیری، وابستگی صفت سنجیده شده به آسانی و سختی سؤالات، غیرقابل آزمایش بودن مفاهیم و عدم اطمینان از اندازه‌گیری خالص سازه مورد نظر، از اصلی‌ترین نقاط ضعف نظریه کلاسیک آزمون می‌باشند (Hambleton, 1993; Lord, 1980). بر اساس چنین ایراداتی است که نظریه کلاسیک را «نظریه ضعیف اندازه‌گیری» نیز نامیده‌اند. این نقاط ضعف باعث می‌شود سنجش صفات با پرسشنامه‌های ساخت و اندازه‌گیری شده در چارچوب این نظریه، با خطای زیادی همراه باشد و نتایج بدست آمده از آن نیز با نظریه سؤال-پاسخ تفاوت‌های قابل توجهی داشته باشد (آلن و ین، ۱۹۷۹، ثراندایک<sup>۱</sup>، ۱۹۸۲، مگنوسون، ۱۹۶۶).

در نظریه سؤال-پاسخ به جای تاکید بر نمرات کل آزمون، بر پاسخ‌های آزمودنی‌ها به تک تک سوال-های آزمون تکیه می‌شود. در این نظریه با استفاده از مدل‌های ریاضی می‌توان احتمال پاسخ درست به یک سؤال آزمون را به عنوان تابعی از توانایی آزمون شونده به حساب آورد و همچنین برخی ویژگی-های سؤال را پیش‌بینی کرد. ویژگی‌هایی که برای سوال‌ها یا ماده‌های آزمون بدست می‌آیند به نوع مدل یا الگوی نظریه سؤال-پاسخ وابسته‌اند (سیف، ۱۳۸۹).

توابع اجایو نرمال<sup>۲</sup> (تراکمی) و تابع منطقی<sup>۳</sup> (لجستیک) جزء معروفترین مدل‌های ریاضی نظریه سوال-پاسخ محسوب می‌شوند که انتخاب هر کدام بستگی به سادگی و سهولت استفاده از مطالب ریاضی دارد که از به کار بردن آن حاصل می‌شود. هر دو مدل نتایج تقریباً مشابهی بدست می‌دهند (هومن، ۱۳۷۱؛ متسا مورنن، ۲۰۰۲).

مدل‌های سؤال-پاسخ بر اساس تعداد پارامترهای برآورد شده نیز قابل دسته‌بندی می‌باشند. در همه مدل‌های نظریه سؤال-پاسخ، پارامتر توانایی<sup>۴</sup> آزمودنی برآورد می‌شود، اما تعداد پارامترهایی که برای سؤالات برآورد می‌گردد متفاوت است. هر سؤال می‌تواند از سه پارامتر عمده و اساسی تشکیل شده باشد که عبارتند از: دشواری سؤال، پارامتر تشخیص سؤال و پارامتر حدس سؤال. پارامتر دشواری سؤال معرف دشواری و سطح صفت زیربنایی تتا است (هومن، ۱۳۷۱). ضریب تمیز<sup>۵</sup>، پارامتری است که شیب منحنی یا نسبت تغییر احتمال موفقیت را در امتداد پیوستار توانایی نشان می‌دهد. پارامتر سوم

1 - Thorndike

2- Normal ogive function

3- Logistic model

4- Difficulty parameter

5- Discrimination parameter

پارامتر حدس<sup>۱</sup> است که پایین‌ترین سطح مجانب در منحنی شاخص سؤال می‌باشد. در صورتی که احتمال پاسخ درست دادن به یک سؤال، به عنوان دانستن جواب آن، صفر نباشد گفته می‌شود عامل یا پارامتری در پاسخگویی دخالت دارد که چون از روی حدس و شانس است، پارامتر حدس یا شانس کاذب نامیده می‌شود (Suen, 1990).

«تک بعدی بودن»<sup>۲</sup> و «استقلال موضعی»<sup>۳</sup> مهم‌ترین مفروضه‌های اساسی نظریه سؤال-پاسخ به حساب می‌آیند. تک بعدی بودن به این معنا است که همه سؤالات آزمون فقط یک حوزه توانایی یا دانش را اندازه بگیرد. مفروضه استقلال موضعی بیان می‌کند که پاسخ‌های آزمودنی‌ها به سؤالات آزمون از لحاظ آماری مستقل از یکدیگرند اگر و فقط اگر سطح توانایی آزمودنی به عنوان عامل اصلی به حساب آید. اگر این پیش‌فرض برقرار باشد، عملکرد آزمودنی نباید تحت تأثیر پاسخ‌های خوب یا بد او در دیگر سؤالات باشد (سیجت سما و مولن آیر<sup>۴</sup>، ۲۰۰۲).

مدل‌های نظریه سوال-پاسخ می‌توانند نمره‌های دو مقوله‌ای و چند مقوله‌ای را تحلیل کنند به شکلی که طبقه‌بندی نمره‌ها می‌تواند منظم یا نامنظم باشند ((Ronald. and Russell, 2008). این امکان وجود دارد که نظریه سؤال-پاسخ برای تجزیه و تحلیل نمره‌های اندازه‌گیر شده بوسیله سؤالات چند وجهی بکار رود. با این مدل‌ها حتی می‌توان گزینه‌های انحرافی و اشتباه‌های آزمودنی‌ها را تحلیل نمود. مدل پاسخ رتبه‌ای (GRM) که در سال (۱۹۶۹) توسط سامی جما<sup>۵</sup> ارائه گردید، توانایی تجزیه و تحلیل پاسخ‌های طبقه‌ای در طول یک پیوستار ترتیبی مانند طیف لیکرت را داراست.

(Weiss & yoes, 1994)

نظریه کلاسیک آزمون و نظریه سؤال-پاسخ چارچوب‌های متفاوتی برای سنجش ارائه می‌دهند. اگرچه در گذشته نظریه کلاسیک بیشتر از نظریه سؤال-پاسخ استفاده می‌شد، اما در دهه‌های اخیر نظریه سؤال-پاسخ جایگاه خود را بدست آورده است. نظریه سؤال-پاسخ برای تحلیل سوال‌ها، مقایسه آزمون‌ها، سنجش توانایی‌ها، نگرش‌ها و سازه‌های دیگر کاربرد دارد (آلن و ین، ۱۹۷۹). بررسی سوگیری سوال‌ها و برازش سوال‌ها، از دیگر امکانات نظریه سؤال-پاسخ می‌باشند. بحث نحوه

- 
- 1- Guess parameter
  - 2- Unidimensionality
  - 3- Local independence
  - 4- Sijtsma.& Molenaar
  - 5- Samejima

برخورد با ارزیابی پایایی و روایی امری ضروری در نظریه می‌باشد. پایایی، توانایی فرد یا گروه را به شکل باثبات اندازه‌گیری می‌کند و روایی می‌گوید آزمون باید آن چیزی را بسنجد که برای اندازه‌گیری آن ساخته شده است. نظریه کلاسیک اندازه‌گیری روش‌های به نسبت ساده‌ای را برای تعیین پایایی و روایی آزمون داراست. در تحلیل کلاسیک، از طریق بررسی نمرات کل آزمون و روش‌های آماری ساده‌ای روایی و اعتبار تعیین می‌شود. اما در نظریه سؤال- پاسخ روش‌های پیچیده‌تری برای تعیین روایی و پایایی آزمون ارائه شده است. اگر نظریه کلاسیک اندازه‌گیری بر کل آزمون تمرکز کرده، نظریه سؤال- پاسخ بر هر ماده و هر فرد آزمودنی تمرکز می‌کند (Cole & Moss, 1993).

بنابراین در این پژوهش به دنبال این هستیم که توانایی نظریه سؤال- پاسخ را در سنجش کیفیت خدمات بررسی کنیم و ببینیم پارامترهای سؤالات برآورد شده سؤالات تست کیفیت خدمات در نظریه سؤال- پاسخ چگونه است، خصوصیات روانسنجی تست (روایی و پایایی) برآورد شده تست سنجش کیفیت خدمات در نظریه سؤال- پاسخ چگونه است، هنجارهای برآورد شده تست کیفیت خدمات در نظریه سؤال- پاسخ چگونه است و در نهایت اینکه کدامیک از نظریه‌های کلاسیک اندازه‌گیری و سؤال- پاسخ، اطلاعات بیشتری در رابطه با توانایی سنجش تست‌ها ارائه می‌دهند. به منظور کاهش خطاها، در انجام محاسبات از نرم افزارهای SPSS16 و مولتی لوگ ۷ استفاده شده است.

### روش پژوهش

این پژوهش توصیفی در دسته تحقیقات کمی و از نوع پیمایشی است که با استفاده از ابزار پرسشنامه اطلاعات را جمع‌آوری کرده است، به لحاظ پیامد از نوع کاربردی و بر حسب زمان از نوع مقطعی می‌باشد (علی احمدی و نهائی، ۱۳۸۶). جامعه آماری پژوهش حاضر از کل مشتریان یکی از شرکت‌های بیمه منطقه ۳ تهران که در اردیبهشت ماه سال ۱۳۹۵ به آن مراجعه کرده‌اند، تشکیل شده است. تعداد گروه نمونه ۳۰۰ نفر زن و مرد بودند که با روش نمونه‌گیری تصادفی انتخاب شدند. گروه نمونه از نظر متغیرهای جمعیت شناختی گوناگونی مورد بررسی قرار گرفت که خلاصه‌ای از توصیف متغیرهای جنسیت، سن و تحصیلات گروه نمونه با هدف آشنایی بیشتر در جدول ۱ آمده است.



جدول ۱- توصیف ویژگی‌های گروه نمونه

متغیرها	مرد		زن		کل	
	فراوانی	درصد	فراوانی	درصد	فراوانی	درصد
جنسیت	۲۰۸	۶۹,۳۳	۹۲	۳۰,۶۷	۳۰۰	۱۰۰
سن:						
۲۱ تا ۳۰ سال	۳۳	۱۱	۱۴	۴,۶۷	۴۷	۱۵,۶۷
۳۱ تا ۴۰ سال	۵۳	۱۷,۶۷	۴۷	۱۵,۶۷	۱۰۰	۳۳,۳۳
۴۱ تا ۵۰ سال	۷۶	۲۵,۳۳	۲۷	۹	۱۰۳	۳۴,۳۳
بالای ۵۰ سال	۴۶	۱۵,۳۳	۴	۱,۳۳	۵۰	۱۶,۶۷
تحصیلات:						
دیپلم	۸۶	۲۸,۶۷	۳۷	۱۲,۳۳	۱۲۳	۴۱
کاردانی	۷	۲,۳۳	۳	۱	۱۰	۳,۳۳
کارشناسی	۱۰۴	۳۴,۶۷	۴۷	۱۵,۶۷	۱۵۱	۵۰,۳۳
کارشناسی ارشد	۱۱	۳,۶۷	۵	۱,۶۷	۱۶	۵,۳۳

در جدول ۱ مشاهده می‌شود که ۲۰۸ نفر از حجم نمونه مرد و ۹۲ نفر از آن‌ها زن هستند. در بررسی سن گروه نمونه مشاهده می‌شود که ۴۱ تا ۵۰ ساله‌ها با فراوانی ۱۰۳ نفر و ۳۴٪ بیشترین مقدار حضور را داشته‌اند. از نظر تحصیلی نیز بیشترین فراوانی با مقدار ۱۵۱ نفر و ۵۰٪ به مدرک تحصیلی کارشناسی تعلق گرفته است.

ابزار استفاده شده در این پژوهش پرسشنامه سنجش کیفیت خدمات سروکوآل<sup>۱</sup> است. این تست استاندارد از ۲۲ عبارت تشکیل شده است که براساس طیف لیکرت ۵ مقوله‌ای نمره گذاری می‌شود (اوپنهایم، ۱۹۸۴). این پرسشنامه دارای دویخش کلی است. در بخش اول وضعیت موجود مولفه‌های کیفیت خدمات و در بخش دوم کیفیت خدمات در وضعیت مطلوب مورد بررسی قرار می‌گیرد. در این

تست سؤالات ۱، ۶، ۱۱ و ۱۶ مربوط به شواهد فیزیکی است. سؤالات ۲، ۷، ۱۲، ۱۷ و ۲۱ مربوط به قابلیت اطمینان است. سؤالات ۴، ۹، ۱۴ و ۱۹ مربوط به مؤلفه پاسخگویی است. سؤالات ۳، ۸، ۱۳ و ۱۸ مربوط به مؤلفه تضمین است و سؤالات ۵، ۲۰، ۱۵، ۱۰ و ۲۲ مربوط به مؤلفه همدلی است. برای این پرسشنامه مقدار آلفای کرونباخ ۰،۹۴۱ بدست آمد. مقدار ضریب اعتبار تست سنجش کیفیت خدمات در نظریه سؤال - پاسخ ۰،۹۶۲ برآورد شد. در مطالعات روایی سازه با کمک تحلیل عاملی اکتشافی (ضریب کفایت نمونه گیری ۰/۹۱۲ و ضریب بار تلت  $p < ۰/۰۰۰$ ) سازه‌ای یک عاملی شناسایی شد.

### یافته‌های پژوهش

در اولین قدم پیش فرض‌های مهم نظریه سؤال- پاسخ یعنی تک بعدی بودن و استقلال موضعی مورد بررسی قرار گرفت. برای بررسی تک بعدی بودن با توجه به برنامه‌های موجود از تحلیل عاملی خطی استفاده شد. برای دستیابی به این منظور بعد از استخراج ماتریس ضرایب همبستگی پلی کوریک توسط برنامه لیزرل، اطلاعات توسط اس. پی. اس. اس و با روش محورهای اصلی تحلیل عاملی شد. نمودار سنگریزه<sup>۱</sup> به همراه جدول مقادیر ویژه استخراج شده عوامل حاصل از تحلیل عاملی در ضمیمه ۱ گزارش شده است.

ریکاس<sup>۲</sup> عنوان می‌کند در صورتی که عامل اول حداقل ۲۰٪ از کل واریانس تست را تبیین کند و یا مقدار ویژه عامل اول حداقل ۳ برابر عامل بعد از خود باشد، پیش فرض تک بعدی بودن در حد قابل قبولی برقرار است. در پیوست ۱ مشاهده می‌شود که مقدار ویژه عامل اول ۱۰،۰۶ بدست آمده است. این مقدار فاصله زیادی با عامل بعد از خود با مقدار ۱،۵۹ دارد بطوریکه عامل اول ۶،۳۲ برابر عامل دوم می‌باشد. همچنین عامل اول به تنهایی ۴۳،۷۲ در صد از واریانس کل تست را تبیین می‌کند. با توجه به مطالب بالا تست کیفیت خدمات را می‌توان تک بعدی نامید. در نمودار اسکری نیز می‌توان فاصله مناسب عامل اول با عامل دوم را مشاهده نمود ((Reckase, 2006). همبلتون (۱۹۸۹) عنوان می‌کند در صورتی که فرض تک بعدی بودن برقرار باشد، فرض استقلال موضعی نیز برقرار خواهد

1-Screplot

2-Reckase

بود. در واقع با استخراج یک عامل غالب کوواریانس بین متغیرها صفر و یا نزدیک به صفر می‌شود و رابطه بین سوال‌ها با برداشته شدن عامل غالب از بین خواهد رفت. در ادامه پژوهش، محاسبات خصوصیات روانسنجی سؤالات بر اساس مدل درجه‌بندی شده نظریه سوال-پاسخ محاسبه شد که خلاصه‌ای از نتایج آن در جدول ۲ گزارش شده است. در این جدول ۶ سؤال از ۲۲ سؤال تست سنجش کیفیت خدمات گزارش شده است.

جدول ۲ - خصوصیات روانسنجی سؤالات تست سنجش کیفیت خدمات

سؤال	ضریب تمییز	آستانه ۱	آستانه ۲	آستانه ۳	آستانه ۴	آستانه ۵	آستانه ۶
سؤال ۱	۱,۰۶	-۱,۵۵	-۰,۷۴	-۰,۱	۰,۸۸	۱,۵۲	۲,۳۷
سؤال ۲	۱	-۲,۳۴	-۱,۴۴	-۰,۴۵	۰,۵۸	۱,۸۴	۲,۶۶
سؤال ۱۳	۲,۷	-۱,۶۱	-۰,۹۴	-۰,۴۳	۰,۱۱	۰,۴۸	۱,۱۲
سؤال ۱۴	۲,۷۴	-۱,۴۲	-۰,۸۹	-۰,۴۸	-۰,۰۷	۰,۵۶	۰,۹۳
سؤال ۱۵	۳,۱۷	-۱,۱۹	-۰,۷۹	-۰,۳۷	۰,۰۸	۰,۴۱	۱
سؤال ۲۰	۰,۹۹	-۲,۳۴	-۱,۴۱	-۱,۰۶	-۰,۳۳	۰,۷۱	۱,۶۳

در جدول ۲ مشاء هده می‌کنیم که سؤالات ۲۰، ۲ و ۱ با مقادیر ۰,۹۹، ۱ و ۱,۰۶ دارای کمترین ضرایب تمییز در بین سؤالات تست سنجش کیفیت خدمات هستند. سؤالات ۱۳، ۱۴ و ۱۵ با مقادیر ۲,۷، ۲,۷۴ و ۳,۱۷ دارای بیشترین ضرایب تمییز در بین سؤالات هستند. همچنین سؤالات ۲۰، ۲ و ۱ دارای نامناسبترین ضرایب آستانه‌ای و سؤالات ۱۳، ۱۴ و ۱۵ دارای بهترین ضرایب آستانه‌ای در بین سؤالات هستند. کمترین مقدار آستانه در آستانه‌اول سؤالات ۲۰ و ۲ با مقدار -۲,۳۴- بدست آمده است. بیشترین مقدار آستانه نیز در آستانه ششم سؤال ۲ با مقدار ۲,۶۶ بدست آمده است.

جدول ۳- میزان آگاهی دهندگی سؤالات و تست سنجش کیفیت خدمات بر روی نقاط هفت گانه پیوستار تتا

سؤال	θ-3	θ-2	θ-1	θ0	θ1	θ2	θ3
سؤال ۱	۰,۱۶۵	۰,۲۸۳	۰,۳۵	۰,۳۶۲	۰,۳۶۱	۰,۳۴۲	۰,۲۶۳
سؤال ۲	۰,۲۳۴	۰,۳۰۱	۰,۳۱۹	۰,۳۲	۰,۳۱۷	۰,۳۰۹	۰,۲۶۲
سؤال ۱۳	۰,۱۶۲	۱,۴۲۱	۲,۲۳۴	۲,۳	۲,۱۱۴	۰,۵۶۳	۰,۰۴۴
سؤال ۱۴	۰,۰۹۶	۱,۰۶۴	۲,۳۴۶	۲,۳۳۵	۲,۰۶۱	۰,۳۶۴	۰,۰۲۶
سؤال ۱۵	۰,۰۳۲	۰,۶۶۷	۳,۰۷	۳,۱۷	۲,۷۶۳	۰,۳۸۳	۰,۰۱۷
سؤال ۲۰	۰,۲۳۲	۰,۲۹۸	۰,۳۱۷	۰,۳۱۶	۰,۳۰۵	۰,۲۵۶	۰,۱۶۲
کل	۶,۹۲۹	۱۹,۹۵۸	۲۹,۸۶	۳۰,۴۰۸	۲۸,۴۶	۱۳,۲۲۴	۳,۸۹۳
خطا استاندارد	۰,۳۸	۰,۲۲۴	۰,۱۸۳	۰,۱۸۱	۰,۱۸۷	۰,۲۷۵	۰,۵۰۷

در ادامه محاسبات سؤال- پاسخ، مقادیر آگاهی دهندگی سؤالات محاسبه شد که خلاصه‌ای از نتایج آن در جدول ۳ گزارش شده است. در این جدول میزان آگاهی دهندگی سؤالات تست سنجش کیفیت خدمات روی هفت نقطه از پیوستار تتا گزارش شده است. مشاهده می‌شود که سؤالات ۱۳، ۱۴ و ۱۵، بیشترین میزان آگاهی دهندگی را داشته است. سؤالات ۲۰، ۲ و ۱ نیز در بین تمامی سؤالات تست، نامناسب‌ترین آگاهی دهندگی را نشان داده است. سؤالات دیگر از نظر آگاهی دهندگی بین این سؤالات جای گرفته‌اند.

در دو سطر آخر جدول ۳ میزان آگاهی دهندگی و خطای اندازه‌گیری کل تست سنجش کیفیت خدمات روی هفت نقطه از پیوستار تتا گزارش شده است. همانگونه که مشاهده می‌شود، بیشینه آگاهی دهندگی تست سنجش کیفیت خدمات بین نقاط ۱- تا صفر تتا است. مقدار آگاهی دهندگی در ۱- تتا برابر با ۲۸,۸۶ و مقدار آگاهی دهندگی در نقطه صفر تتا ۳۰,۴ است. همچنین کمینه خطای استاندارد تست سنجش کیفیت خدمات بین نقاط صفر تا ۱- تتا است. مقدار خطای استاندارد در ۱- تتا برابر با ۰,۱۸۳

و خطای استاندارد در نقطه صفر تا ۰,۱۸۱ است. مقدار ضریب اعتبار حاصل از تئوری سؤال- پاسخ با روش حاشیه‌ای ۰,۹۶۲ بدست آمد. در جدول ۴ مشاهده می‌شود میانگین نمرات تا برای تست سنجش کیفیت خدمات ۰,۲۶۸ بدست آمد. انحراف معیار نمرات این مقیاس ۱,۰۷ است. دامنه تغییرات نمرات تا ۶,۷۲ بدست آمد که کمترین تنای مشاهده شده ۳,۲۵- و بیشترین تنای مشاهده شده ۳,۴۷ است.

جدول ۴- توصیف نمرات توزیع فراوانی (تا)

مقادیر	شاخص‌ها
۰,۲۶۸	میانگین
۱,۰۷	انحراف معیار
۶,۷۲	دامنه تغییرات
-۳,۲۵	کمترین
۳,۴۷	بیشترین
۱۸۳	حجم نمونه

در آخرین مرحله از تحلیل رتبه‌های افراد بر اساس نمرات هنجار  $Z$  که معمولاً در محاسبات نظریه کلاسیک استفاده می‌شود، و نمرات توانایی تا محاسبه شد. بخشی از نتایج آن که مربوط به ۴۰ آزمودنی می‌باشد در جدول ۵ گزارش شده است. این تحلیل به منظور مقایسه تفاوت رتبه‌بندی افراد در دو نظریه سؤال- پاسخ و کلاسیک اندازه‌گیری در سنجش سازه کیفیت خدمات صورت گرفت. در جدول ۵ مشاهده می‌کنیم رتبه آزمودنی ۲۰ در نمره هنجار کلاسیک ۵۶ است و در تنای نظریه سؤال- پاسخ ۸۵ است. بین دو رتبه ۲۹ نفر اختلاف وجود دارد. شرکت‌کننده‌های ۱، ۲، ۳ و ۱۹ نیز دارای تفاوت‌های قابل توجهی هستند. در بخش بحث و نتیجه‌گیری به بررسی دلایل وجود این تفاوت‌ها پرداخته می‌شود. در ادامه به بحث و نتیجه‌گیری یافته‌های پژوهش می‌پردازیم.

جدول ۵- مقایسه رتبه افراد بر اساس نمرات هنجار کلاسیک و نمرات تتا

اختلاف رتبه	رتبه در $\theta$	رتبه در $Z$	شرکت کننده	اختلا			شرکت کننده
				ف	رتبه در $\theta$	رتبه در $Z$	
۵	۱۸۲	۱۷۷	شرکت کننده ۱۱	۳۳	۲۴۸	۲۸۱	شرکت کننده ۱
۶	۹۶	۹۰	شرکت کننده ۱۲	۲۰	۵	۲۵	شرکت کننده ۲
۷	۸۵	۹۲	شرکت کننده ۱۳	۱۹	۱۶	۳۵	شرکت کننده ۳
۵	۶۶	۷۱	شرکت کننده ۱۴	۱۰	۸۶	۹۶	شرکت کننده ۴
۴	۶	۱۰	شرکت کننده ۱۵	۹	۲۶۷	۲۷۶	شرکت کننده ۵
۷	۱۴۵	۱۳۸	شرکت کننده ۱۶	۸	۳۴	۴۲	شرکت کننده ۶
۸	۴۳	۳۱	شرکت کننده ۱۷	۸	۵۶	۶۴	شرکت کننده ۷
۱۸	۶۹	۵۱	شرکت کننده ۱۸	۷	۷۱	۷۸	شرکت کننده ۸
۱۹	۲۵۲	۲۳۳	شرکت کننده ۱۹	۴	۲۰۳	۲۰۷	شرکت کننده ۹
۲۹	۸۵	۵۶	شرکت کننده ۲۰	۴	۱۳۷	۱۴۱	شرکت کننده ۱۰

## بحث و نتیجه گیری

هدف اصلی از انجام پژوهش حاضر معرفی نظریه سوال-پاسخ و مقدمه‌ای بر کاربرد آن در سنجش‌های مدیریتی بود که در این راستا کیفیت خدمات به عنوان یکی از متغیرهای ممکن انتخاب شد. به عبارت دیگر بر اساس ضرورت‌های پژوهشی تبیین شده کیفیت خدمات به عنوان یکی از سازه‌های قابل سنجش در رشته‌ی مدیریت انتخاب شد تا در خلال سنجش آن توانایی‌ها و امکانات نظریه سؤال- پاسخ مطرح و معرفی شود.

تست سنجش کیفیت خدمات براساس طیف لیکرت ۵ مقوله‌ای نمره گذاری می‌شود. در بین مدل‌های نظریه سؤال- پاسخ مدل پاسخ رتبه‌ای به گونه‌ای ساخته شده است که می‌تواند تست‌های ساخته شده بر اساس طیف لیکرت را تحلیل کند ((Thissen, 1991).

در حجم قابل توجهی از پایان‌نامه‌ها و پژوهش‌های حوزه مدیریت از تست‌ها و پرسشنامه‌ها استفاده می‌شود. تمامی این پرسشنامه‌های مورد استفاده بر اساس نظریه کلاسیک نمره‌گذاری می‌شود. در واقع رتبه‌بندی افراد در سطح صفت مورد نظر بر اساس نمراتی انجام می‌شود که بیشینه دقت را ندارند. هر رشته علمی عدم توانایی در سنجش و اندازه‌گیری دقیق و معتبر متغیرها نقطه ضعف و محدودیتی بزرگ محسوب می‌شود. در جهت افزایش اعتبار محاسبات در اولین مرحله از تحلیل‌ها پیش فرض‌های اصلی نظریه سؤال- پاسخ مورد بررسی قرار گرفتند. در بخش یافته‌ها برقراری پیش فرض‌های مورد نیاز برای تحلیل داده‌های تست سنجش کیفیت خدمات بر اساس نظریه سؤال- پاسخ به اثبات رسید.

پارامتر شیب سؤال، ضرایب آستانه‌ای، تابع آگاهی سؤال و خطای اندازه‌گیری سؤال از جمله خصوصیات روانسنجی هستند که نظریه سؤال- پاسخ برای بررسی قدرت سنجش هر سوال ارائه می‌دهد. پارامتر شیب سؤال معادل ضریب تمیز در نظریه کلاسیک اندازه‌گیری می‌باشد. پارامترهای آستانه تعریف ساده‌ای دارند، بدین معنا که اندازه آنها معرف سطح صفتی است که شخص باید داشته باشد تا پاسخ او با احتمال ۰/۵ بالای آستانه آن مقوله قرار بگیرد (سیف، ۱۳۸۹).

در جدول ۲ مشاهده شد که سؤالات ۲۰، ۲ و ۱ بر اساس پارامتر شیب سؤال به عنوان ضعیف‌ترین سؤالات تست سنجش کیفیت خدمات هستند. سؤال ۱۳، سؤال ۱۴ و سؤال ۱۵ در پارامتر شیب سؤال نظریه سؤال- پاسخ اولین، دومین و سومین رتبه سؤال قوی را به خود اختصاص داده‌اند. یکی از مزایایی که مدل رتبه‌ای نظریه سؤال- پاسخ و نرم‌افزار آن مولتی‌لوگ (Thissen, 1991) دارد امکان

تحلیل هر مقوله می‌باشد. بخشی از این امکان با بررسی ضرایب آستانه‌ای و بخش دیگری از آن توسط منحنی‌های طبقه پاسخ<sup>۱</sup> (CRC) است. در پیوست ۲ و ۳ نمونه‌هایی از منحنی‌های طبقه پاسخ (CRC) و تابع آگاهی سؤال برای یک نمونه سؤال ضعیف و یک نمونه سؤال قوی گزارش شده است. پارامترهای سؤال شکل و جایگاه منحنی‌های طبقه پاسخ را تعیین می‌کنند. هر چه پارامتر شیب سؤال بیشتر باشد منحنی‌های طبقه پاسخ به هم فشرده‌تر و ارتفاع آنها بیشتر می‌شود مانند ضمیمه ۳ که به سؤال ۱۵ پرسشنامه (ارباب رجوعان در ارتباطشان با این سازمان احساس امنیت می‌کنند) اختصاص دارد. و هر چه شیب سؤال کمتر باشد نمودارها در هم تنیده‌تر و ارتفاع آنها کمتر می‌شود مانند ضمیمه ۲ که مربوط به سؤال ۲۰ پرسشنامه (این سازمان کارکنانی دارد که توجه شخصی به ارباب رجوعان نشان می‌دهند) است.

علاوه بر این مطالب می‌توانیم کارکرد هر طبقه پاسخ را با طبقه دیگر مقایسه کنیم. همچنین براساس میزان برازش طبقه سؤال می‌توانیم تعداد طبقه مورد نیاز برای هر سؤال تصمیم‌گیری کنیم. در ضمیمه ۲ مشاهده می‌کنیم طبقه ۳ بصورت کامل تحت پوشش طبقات ۲، ۵ و ۴ قرار گرفته و کارکرد مناسبی از خود نشان نداده است. طبقه پاسخ ۶ سؤال ۲۰ نیز در زیر طبقه ۷ قرار گرفته که آن نیز حکایت از ناتوانی این سؤال در تفکیک افراد به طبقات با توانایی‌های متفاوت دارد. به همین دلایل است که سؤال ۲۰ جزء سؤالات ضعیف قرار گرفته است. از طرف دیگر در سؤال ۱۵ (ضمیمه ۳) مشاهده می‌شود به غیر از طبقه ۵ هر طبقه شکل توزیع خود را حفظ کرده و دامنه خود را از طبقه پاسخ‌های دیگر مجزا ساخته است و در نتیجه این سؤال توانسته است افراد با توانایی‌های مختلف را در طبقات مجزا جای دهد. وضعیت طبقه پنج در این سؤال نشان می‌دهد که این سؤال می‌تواند با ۶ مقوله نیز کارکرد مناسبی داشته باشد.

در نظریه سؤال- پاسخ به جای خطای معیار اندازه‌گیری، از مفهوم شاخص دقت اندازه‌گیری<sup>۲</sup> استفاده می‌شود. تفاوت مهم بین این دو مفهوم این است که شاخص دقت اندازه‌گیری در طول مقیاس نمرات دارای اندازه‌های مختلف است، در حالی که خطای معیار اندازه‌گیری در طول مقیاس نمرات همواره دارای اندازه‌های یکسان می‌باشد. از دیگر امکانات مدل‌های نظریه- سؤال پاسخ تابع آگاهی<sup>۳</sup> است که

1-Category response curves

2-Precision of measurement index

3-Information function



با شاخص دقت اندازه‌گیری رابطه تنگاتنگی دارد. تابع آگاهی برای تک تک سؤالات قابل محاسبه است و از جمع آگاهی دهندگی سؤالات در هر نقطه از پیوستار توانایی تابع آگاهی تست به دست می‌آید. میزان دقت برآورد در یک نقطه از طریق واریانس آن نقطه به دست می‌آید و مقدار آگاهی در آن نقطه برابر با عکس واریانس در آن نقطه است. اگر مقدار آگاهی سؤال را در محور عمودی و سطوح توانایی را در محور افقی یک دستگاه مختصات رسم کنیم نموداری بدست می‌آید که به آن تابع آگاهی سؤال می‌گویند که نمونه‌ای از آن در ضمیمه‌های ۲ و ۳ گزارش شده است (نمودارهای سمت راست). بالا بودن مقدار آگاهی دهندگی به معنای آن است که می‌توان توانایی حقیقی فرد را در آن نقطه به دقت برآورد کرد. در نظریه سؤال- پاسخ با کمک آگاهی مبتنی بر یک سؤال نیز می‌توان سطوح توانایی افراد را محاسبه کرد و به همین دلیل نظریه سؤال- پاسخ را نظریه تک سوالی نیز نامیده‌اند. اما این کار کمتر صورت می‌گیرد زیرا آگاهی حاصل از یک سؤال اندک است. در هر نقطه که آگاهی دهندگی کم است خطای استاندارد برآورد ۱ افزایش می‌یابد. بهترین تابع آگاهی برای یک پرسشنامه تابعی است که مقدار بیشینه آن در نمره برش ۲ پرسشنامه قرار گرفته باشد (Baker, 1992) در بررسی توابع آگاهی سؤالات تست سنجش کیفیت خدمات دیدیم که سؤالات ۱۵، ۱۴ و ۱۳ با بیشینه آگاهی دهندگی ۳، ۱۷، ۲، ۳۴ و ۲، ۳ جزء قویترین سؤالات این مقیاس هستند. در بررسی تابع آگاهی تست سنجش کیفیت خدمات بیشینه آگاهی دهندگی در نقطه صفر تا ۴، ۳۰ و همچنین کمینه خطای استاندارد تست سنجش کیفیت خدمات نقطه صفر تا ۱۸۱، ۰ است.

مهم‌ترین ویژگی نظریه سؤال- پاسخ که در این پژوهش قصد تاکید بر آن را داشتیم نحوه رتبه بندی افراد مورد سنجش بر اساس نمرات اکتسابی آنها از تست‌ها است. هوگان (۲۰۰۷) عنوان می‌کند برخلاف روش کلاسیک که نمره آزمون شونده از مجموع پاسخ‌های درست او به سؤال‌های آزمون به دست می‌آید در روش سؤال- پاسخ نمره سؤال شونده برابر با مجموع پاسخ‌های درست او نیست، نمره آزمودنی که با تتا نمایش داده می‌شود هم تابع تعداد پاسخ‌های آزمون شونده و هم تابع ویژگی‌های سؤال است (Hogan, 2007, Steinberg & Thissen, 1995). مورد دیگر اختلاف بین این دو تئوری نا متغیر بودن ویژگی‌های روانسنجی سؤالات نسبت به نمونه است.

---

1-Standard error

2-Cut off score

در جدول ۵ دیدیم که بین رتبه‌های افراد بر اساس نمرات هنجار دو نظریه تفاوت‌های قابل توجه‌ای وجود دارد. بطور مثال آزمودنی ۲۰ بر اساس نمرات کلاسیک ۵۶ مین نفر است و بر اساس نمرات سؤال- پاسخ هشتاد و پنجمین (۸۵) نفر است که نشان‌دهنده اختلاف رتبه بیست و نه (۲۹) نفری بین ۳۰۰ نفر است. همچنین شرکت کننده ۱ بر اساس نمرات کلاسیک ۲۸۱ نفر و بر اساس نمرات سؤال- پاسخ دویست و چهل و هشتمین (۲۴۸) نفر است که نشان‌دهنده اختلاف رتبه سی و سه (۳۳) نفری بین ۳۰۰ نفر است. در گزینش‌ها و تصمیم‌گیری‌های چنین تفاوت رتبه‌ای بسیار اهمیت دارد. به عبارت دیگر فردی با صفت کمتر جای فردی با صفت بیشتر را می‌گیرد و نتایج را مختوش می‌کند. ریشه این تفاوت در نوع نمره‌گذاری دو نظریه است که در پاراگراف بالا به آن اشاره شد. با نگاهی به نمره‌گذاری در دو نظریه شکی باقی نمی‌ماند که نمره‌گذاری در نظریه سؤال- پاسخ بسیار معتبرتر از نمرات نظریه کلاسیک است. در نظریه سؤال- پاسخ هر سوالی که ضریب تمیز بیشتری دارد و می‌تواند سهم بیشتری در سنجش صفت مکنون داشته باشد وزن بیشتری بر نمره توانایی اعمال می‌کند. در حالی که در نظریه کلاسیک در حالت کلی تفاوت تأثیر سؤالات قوی و ضعیف در نمره برآورد شده فرد یکسان و در واقع هیچ است (Baker, 1992).

نتایج سنجش کیفیت خدمات با نظریه سؤال- پاسخ نشان داد که این نظریه پارامترها و ملاک‌های بیشتر و معتبرتری برای خصوصیات روانسنجی سؤالات و خصوصیات روانسنجی آزمون ارائه می‌دهد. همچنین نمرات توانایی تا که در نظریه سؤال- پاسخ برای سنجش صفت مکنون ارائه می‌شود بر اساس پایه‌های اساسی‌تری نسبت به نظریه کلاسیک محاسبه می‌گردد. یادآوری این مطلب ضروری است که این پژوهش فقط در ارتباط با تست سنجش خدمات انجام شده است لذا در تعمیم نتایج به سایر سازه‌های دیگر دقت لازم مبذول شود.

در این پژوهش سعی شد تا بخشی از توانایی‌های نظریه سؤال- پاسخ مانند شیب منحنی سؤال، ضرایب آستانه‌ای، توابع آگاهی سؤال و آزمون، نمرات تنا و انواع نمودارهای گرافیکی در سنجش صفات مکنون در رشته مدیریت معرفی شود. با توجه به مزایایی که در نظریه سؤال- پاسخ در سنجش کیفیت خدمات مشاهده شد پیشنهاد می‌شود در سنجش ویژگی‌ها و سازه‌های دیگر علوم مدیریت از مدل‌های نظریه سؤال- پاسخ استفاده شود. امید است این پژوهش توانسته باشد گام مفیدی در شناساندن قابلیت‌ها و کاربردهای نظریه سؤال- پاسخ برداشته باشد.

## منابع

- آلن، مری. یین، وندی‌ام. (۱۹۷۹). مقدمه‌ای بر نظریه‌های اندازه‌گیری، (ترجمه دکتر علی دلاور، ۱۳۷۴). تهران: انتشارات سمت.
- اوپنهایم، آبراهام. (۱۹۸۴). طرح پرسشنامه و سنجش نگرشها، (ترجمه مرضیه کرم‌نیا، ۱۳۷۵). چاپ دوم، مشهد: آستان قدس رضوی.
- ثراندایک، رابرت. (۱۹۸۲). روانسنجی کاربردی، (ترجمه حیدر علی هومن، ۱۳۷۵). تهران: انتشارات دانشگاه تهران.
- روستا، احمد. وتوس، داور و ابراهیمی، عبدالحمید، "مدیریت بازاریابی"، چاپ دوازدهم، تهران: سمت، ۱۳۸۷
- سیجت سما، کلاس؛ مولن آیر، ایو دبلیو. (۲۰۰۱). تئوری ناپارامتریک سؤال پاسخ، (ترجمه سلیمان ذولفقار نسب، ۱۳۸۵). مشهد: انتشارات کتابخانه رایانه‌ای.
- ستاری، بهزاد. (۱۳۸۲). روانشناسی پیشرفته کاربردی. مشهد: انتشارات به نشر.
- سیف، علی اکبر. (۱۳۸۹). اندازه‌گیری، سنجش و ارزشیابی آموزشی: ویرایش پنجم. تهران: نشر دوران.
- سپاسی، حسین. (۱۳۷۵). نظریه سؤال پاسخ در آزمونهای روانی و تربیتی. مجله علوم تربیتی و روانشناسی (دانشگاه شهید چمران اهواز).
- علی احمدی، علیرضا و نهائی، سعید (۱۳۸۶)، "توصیفی جامع از روش‌های تحقیق". تهران: انتشارات تولید دانش.
- فیتز سیمونز، جیمزای (۱۳۸۲)، "مدیریت خدمات؛ استراتژی، عملیات و تکنولوژی اطلاعات"، ترجمه دکتر سیدمحمد اعرابی و داود ایزدی، چاپ اول، تهران: دفتر پژوهشهای فرهنگی.
- لاولاک، کریستوفر و لارن. رایت (۱۳۸۲)، "اصول بازاریابی و خدمات"، ترجمه دکتر ابوالفضل (اردشیر) تاجزاده نمین، چاپ اول، تهران: سمت.
- لامعی، ابولفتح، «مدیریت جامع کیفیت؛ اصول، کاربرد و درس‌هایی از یک تجربه»، طب نوین، چاپ اول، ۱۳۸۲

متسا مورنن، جری. (۲۰۰۲). مبانی نظری آزمون و آزمون سازی، (ترجمه شیده کامکار، ۱۳۸۵). تهران: انتشارات بهینه.

هومن، حیدر علی. (۱۳۷۱). اندازه گیری روانی تربیتی و فن تهیه تست. تهران: بیک فرهنگ.

Baker, Frank B (1992). Item response theory: Parameter Estimating Trchnique theory. Sage.

Birnbaum, A. (1968). some latent trait model and their use in inferring an examinees ability. Ma:Addison Wesley.

Carrillat, F. A., Jaramillo, F. and Mulki, J.P. "The Validity of the SERVQUAL and SERVPERF Scales", International Journal of service Industry Management, Vol.18, No.5, pp.472-790, 2007.

Cole, N. S., & Moss, P. A. (1993). Bias in test use. In R. L. Linn (Ed.), Educational Measurement(3rd ed.) (pp. 201-219). Phoenix, AZ: Oryx Press.

Everitt, B.S (1984). An introduction to latent variable models. Chapman and Hall, London.

Embretson, S.E. and, S.P. (2004). Item Reaponse Theory for Psychologists.

Gremler, D.D. and Gwinner, K.P. "Customer-Employee Rapport in Service Relationships",Journal of Service Research, Vol. 3, No. 1, pp.82-104, 2000.

-Gronroos, Christian, Service Management & Marketing, Second Edition, Wiley, 2001.

Hambleton, R. K. (1989)Principles and selected Application of item response theory. Educational measurement, New york: Mac Millan.

Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. Educational Measurement: Issues and Practice, 12(3), 535-556.

Hogan, T.D. (2007). Psychological testing: A practical introduction. John wily.

Johnson, M. D., and Nilsson, L., "The Importance of Reliability and Customization from Goods to Services", Quality Management Journal, 10, (1-15), 2003

-Johns, N., "What s this Thing Called Service?", European Journal of Marketing, 33(9/10), 958-973.1999

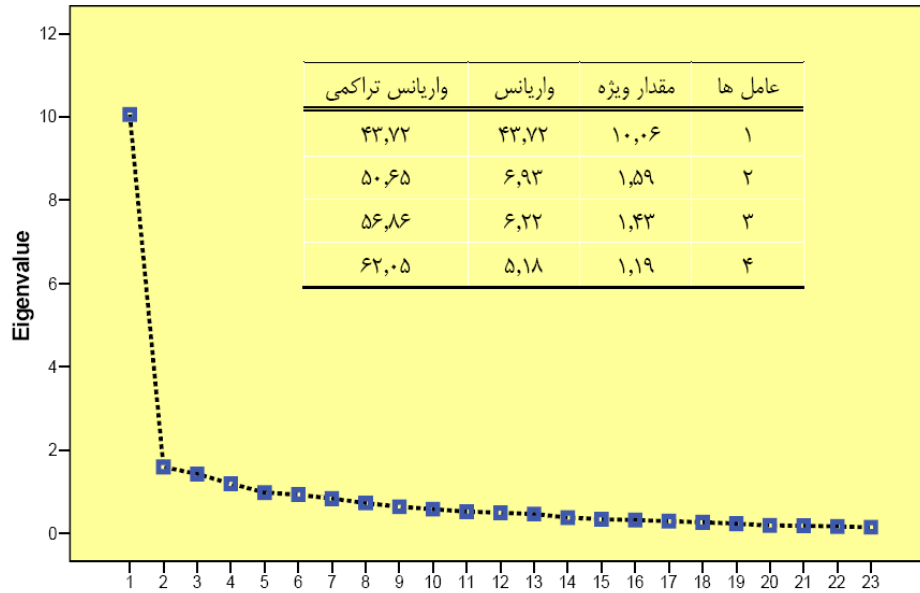
- Kotler, P. *Marketing Management*, 11th ed. Prentice-Hall, Englewood Cliffs, NJ, 2003.
- Lovelock, C.H. and Gummesson, E. "Whither Services Marketing? In Search of a New Paradigm and Fresh Perspectives", *Journal of Service Research*, Vol. 7, No. 1, pp. 20-41, 2004.
- Lord, F.M., & Novick, M. (1968). *Statistical theories of mental test score*. Reading, MA: Addison Wesley.
- Lord, F.M. (1980). *Applications of item response theory to practical testing problem*. NJ: Lawrence Erlbaum.
- Lord, F.M. (1977). *Practical applications of item characteristic curve theory*. *Journal of Educational measurement*, 14, 117-138.
- Mc Donald, R. P. (1981). *The dimensionality of tests and item*. *British Journal of Mathematical and Statistical Psychology*, 34, 100-117.
- OECD, available at: <http://ocde.P4.Siteinternet.Com/publications/doi/files/012005061t009.Xls>, 2005
- Othman, Abdulqawi & Owen, Lynn, "Adopting & Measuring Customer Service Quality (sq) in Islamic Banks", *International Journal of Islamic Financial Services*. Vol.3, No. 1, 2002.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. *A Conceptual Model of Service Quality and Its Implications for Future Research*", *Journal of Marketing*, Vol. 49, No. 4, pp. 41-50, 1985.
- Ronald K. Hambleton and Russell W. Jones (2008). *Instructional Topics in Educational measurement*. pp. 253- 261.
- Reckase, M (2006). *Multidimensional item response theory*. *Handbook of statistics*, vol 26.
- Reeve, B. B., & Fayers, P. (2005). *Applying item response theory modeling for evaluating questionnaire item and scale properties*. In P. Fayers and R. D. Hays (Eds.), *Assessing Quality of Life in Clinical Trials: Methods of Practice*. 2nd Edition. Oxford University Press. 55-73.
- Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*, *Psychometric monograph*.
- Sijtsma, K. & Molenaar, I.W. (1987). *Reliability of test scores* In *NIRT. psychometrica*. 52.
- Steinberg, L., & Thissen, D. (1995). *Item response theory in personality research*. Hillsdale, Nj: Lawrence Erlbaum Associates.
- Thissen, D. (1991). *MULTILOG users guide: Multiple categorical item analysis in Item response theory*: Chicago: scientific software Int.

Wang, Y., Lo, H. and Hui, Y.V. “The Antecedents of Service Quality and Product Quality and their Influences on Bank Reputation: Evidence from Banking Industry in China”, *Managing Service Quality*, Vol. 13, No. 1, pp. 72-83, 2003

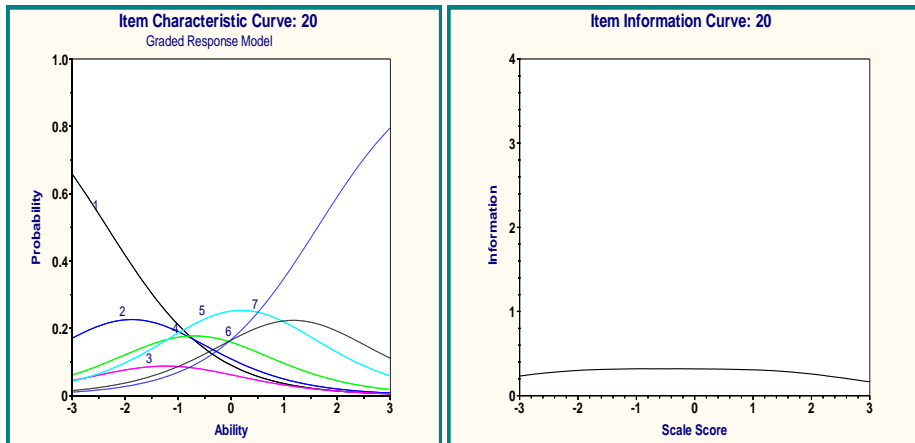
Weiss D.J. & yoes M.E. (1994).Improving measurement quality and Item Response theory. *Applied psychological measurement* 6,473-93.

Zeithaml, Valarie A., Bitner, Mary Jo, *Service Marketing*, McGraw Hill, Singapore, 1996.

ضمیمه ۱: نمودار سنگریزه به همراه جدول مقادیر ویژه عامل‌ها



ضمیمه ۲: منحنی‌های طبقه پاسخ و منحنی تابع آگاهی یک سؤال ضعیف (سؤال ۲۰)



ضمیمه ۳: منحنی‌های طبقه پاسخ و منحنی تابع آگاهی یک سؤال قوی (سؤال ۱۵)

