
نظريه سوال پاسخ اصول و کاربردها

انتشارات موسسه آموزش عالی غياث الدين جمشيد کاشاني

ولف:

رونالد هميلتون - هاريهار، سر ميناتان

مترجمان:

مریم یار سايان

سپما نقی زاده (استاديار سازمان سنجش آموزش کشور)

حبيب نادری (استاديار دانشگاه سيستان و بلوچستان)



موسسه آموزش عالی
غياث الدين جمشيد کاشاني

سرشناسه	: هیملتون، رونالد کی.
عنوان و نام پدیدآور	: Hambleton, Ronald K. نظریه سوال پاسخ اصول و کاربردها / تألیف رونالد هیملتون، هاریهاران سوامیناتان؛ ترجمه مریم پارسایان، سیما نقی زاده، حبیب نادر.
مشخصات نشر	: آریک: موسسه آموزش عالی غیاث الدین جمشید کاشانی، ۱۳۹۵.
مشخصات ظاهری	: ج. ۴۵۲ ص: ۱ مصور، جدول، نمودار.
شابک	: ۹۷۸-۶۰۰-۹۶۵۳۷-۱-۶
وضعیت فهرست نویسی	: فیبا
یادداشت	: عنوان اصلی: Item response theory: principles and applications, C1۹۹۱.
یادداشت	: واژه نامه
یادداشت	: کتاب نامه ص: ۴۲۵ - ۴۵۰.
عنوان دیگر	: مبانی نظریه پرستی - پاسخ
موضوع	: نظریه سوال پاسخ
موضوع	: Item response theory
شناسه افزوده	: سوامیناتان، هاریهاران
شناسه افزوده	: Swaminathan, Hariharan
شناسه افزوده	: پارسایان، مریم، ۱۳۵۹ - مترجم
شناسه افزوده	: نقی زاده، سیما، مترجم
شناسه افزوده	: نادر، حبیب، ۱۳۵۸ - مترجم
رده بندی ده	: BF۱۷۷/۵۸۴۲ ۱۳۹۵
رده بندی ده	: ۲۸۷/۱۵۰
ردیف کتابخانه ملی	: ۴۳۷۴۹۷



موسسه آموزش عالی
غیاث الدین جمشید کاشانی

انتشارات موسسه آموزش عالی غیاث الدین جمشید کاشانی

نام کتاب	: نظریه سوال پاسخ اصول و کاربردها
ناشر	: انتشارات موسسه آموزش عالی غیاث الدین جمشید کاشانی
مؤلفان	: رونالد هیملتون، هاریهاران سوامیناتان
مترجمان	: مریم پارسایان، سیما نقی زاده (انتشارات سنجش آموزش کشور، حبیب نادر و غیاث الدین جمشید کاشانی)
طراح روی جلد	: ندا سهرابی مطلق فاضلی
نوبت چاپ	: اول
سال چاپ	: ۱۳۹۵
شمارگان	: ۲۰۰
شابک	: ۹۷۸-۶۰۰-۹۶۵۳۷-۱-۶
قیمت	: ۳۰۰۰۰۰ ریال

دفتر انتشارات: آریک، میدان مادر بلوار نیاپیش، موسسه غیاث الدین جمشید کاشانی

تلفن تماس: ۰۲۸۳۲۸۸۱۲۲۲

فکس: ۰۲۸۳۲۸۸۱۲۰۹

فهرست مطالب

ک	پیشگفتار مولف
س	پیشگفتار مترجمین
ق	فهرست نمودارها
خ	فهرست جداول
غ	فهرست علائم اختصاری
۱	فصل ۱. پیش‌زمینه‌هایی در مورد نظریه سوال پاسخ
۱	۱-۱. محدودیت‌های روش‌های استاندارد آزمون
۵	۲-۱. مروری بر ادبیات موضوع
۱۱	۳-۱. نظریه سوال پاسخ
۱۳	۴-۱. ویژگی‌های مدل‌های سوال پاسخ
۱۶	۵-۱. خلاصه
۱۹	فصل ۲. فرضیه‌های نظری سوال پاسخ
۱۹	۱-۲. مقلمه
۲۱	۲-۲. ابعاد فضای مکنون
۲۹	۳-۲. استقلال موضعی

۳۲		۴-۲	منحنی های ویژگی سوال
۳۹		۵-۲	سرعت
۴۰		۶-۲	خلاصه

۴۱

فصل ۳. مدل های سوال پاسخ

۴۱		۱-۳	مقدمه
۴۱		۲-۲	طبیعت داده های آزمون
۴۲		۳-۱	رایج ترین مدل های سوال پاسخ
۴۴		۱-۲-۳	مدل اجازت برمال دو پارامتری
۴۵		۲-۳-۳	مدل نه با اثر رولونستیک
۴۶		۳-۳-۳	مدل سه پارامتری راجستیک
۵۶		۴-۳-۳	مدل یک پارامتری لوجستیک (مدل راش)
۵۹		۵-۳-۳	مدل چهار پارامتری لوجستیک
۶۰		۶-۳-۳	مدل های اجایونر مال یک، دو، سه و چهار پارامتری
۶۱		۷-۳-۳	مدل پاسخ اسمی
۶۲		۸-۳-۳	مدل پاسخ مدرج
۶۳		۹-۳-۳	مدل پاسخ پیوسته
۶۴		۴-۳	خلاصه

۶۷

فصل ۴. مقیاس های توانایی

۶۷		۱-۴	مقدمه
۶۸		۲-۴	تعریف توانایی و تبدیل مقیاس توانایی
۷۷		۳-۴	رابطه نمره های توانایی با نمره های دامنه
۸۰		۴-۴	رابطه بین توزیع توانایی و توزیع نمره دامنه

۸۳.....	رابطه بین نمره دامنه مشاهده شده و توانایی	۵-۴
۸۸.....	رابطه بین پیش‌بینی توزیع نمره مشاهده شده و توزیع توانایی	۶-۴
۸۹.....	نمره‌های کامل	۷-۴
۹۰.....	لزوم بررسی‌های اعتبار	۸-۴
۹۳.....	خلاصه	۹-۴

فصل ۵. برآورد توانایی

۹۷.....	مقدمه	۱-۵
۹۸.....	تایید درستی	۲-۵
۱۰۴.....	آورد تایی: روش ماکسیم درستی شرطی	۳-۵
۱۱۲.....	خواص برآوردگای ماکسیم درستی	۴-۵
۱۱۶.....	برآورد بیزی	۵-۵
۱۲۲.....	برآورد θ برای نمره‌ها: دامنه غیر	۶-۵
۱۲۳.....	خلاصه	۷-۵
۱۲۵.....	پیوست: مشتق تابع اطلاع	۸-۵

فصل ۶. تابع اطلاع و کاربردهای آن

۱۲۷.....	مقدمه	۱-۶
۱۲۸.....	تابع اطلاع نمره	۲-۶
۱۳۱.....	توابع اطلاع آزمون و سوال	۳-۶
۱۴۴.....	وزن‌های نمره‌دهی	۴-۶
۱۴۹.....	تاثیر متریک توانایی بر اطلاع	۵-۶
۱۵۲.....	دقت نسبی، کارایی نسبی و کارایی	۶-۶
۱۵۴.....	ارزیابی دقت اندازه‌گیری	۷-۶
۱۵۵.....	خلاصه	۸-۶

فصل ۷. برآورد پارامترهای سوال و توانایی

۱۵۷

۱۵۷.....	۱-۷. مقدمه.....
۱۵۷.....	۲-۷. شناسایی پارامترها.....
۱۵۹.....	۳-۷. پارامترهای تصادفی و ساختاری.....
۱۶۲.....	۴-۷. برآورد ماکسیمم درستمایی توام.....
۱۷۴.....	۵-۷. برآورد ماکسیمم درستمایی شرطی.....
۱۷۶.....	۶-۷. برآورد ماکسیمم درستمایی حاشیه‌ای.....
۱۷۹.....	۷-۷. برآورد بیزی.....
۱۸۳.....	۸-۷. روش‌های برآورد تقریبی.....
۱۸۶.....	۹-۷. روش‌های کامپیوتری.....
۱۸۸.....	۱۰-۷. خلاصه.....
۱۹۰.....	۱۱-۷. پیوست: ماتریس اطلاع برای برآوردهای پارامتر سوال.....

فصل ۸. روش‌هایی برای بررسی برازش مدل به داده‌ها

۱۹۳.....	۸-۱. مقدمه.....
۱۹۴.....	۲-۸. آزمون‌های آماری معنی‌داری.....
۱۹۹.....	۳-۸. بررسی فرضیه‌های مدل.....
۲۰۴.....	۸-۳-۱. تک‌بعدی بودن.....
۲۰۶.....	۸-۳-۲. شاخص‌های برابری قدرت تمیز.....
۲۰۷.....	۸-۳-۳. حدس زدن.....
۲۰۷.....	۸-۳-۴. سرعت آزمون.....
۲۰۸.....	۸-۴. بررسی ویژگی‌های مدل.....
۲۱۱.....	۵-۸. بررسی پیش‌بینی‌های اضافی مدل.....
۲۱۶.....	۶-۸. خلاصه.....

فصل ۹. مثال‌هایی در رابطه با برازش مدل به داده‌ها

- ۲۱۹ ۱-۹. مقدمه.....
- ۲۱۹ ۲-۹. شرح تمرین ریاضیات NAEP.....
- ۲۲۰ ۳-۹. شرح داده‌ها.....
- ۲۲۱ ۴-۹. بررسی فرضیه‌های مدل.....
- ۲۲۱ ۴-۹. ۱- تک بعدی بودن
- ۲۲۲ ۴-۹. ۲- شاخص‌های تساوی قدرت تمیز سوال
- ۲۲۸ ۵-۹. ررسی ویژگی‌های مدل.....
- ۲۲۸ ۵-۹. ۱- تغییرپذیری پارامتر سوال
- ۲۳۴ ۶-۹. بررسی پیش‌بینی‌های تنافی مدل.....
- ۲۳۴ ۶-۹. ۱- تحلیل باقیمانده‌ها
- ۲۳۸ ۶-۹. ۲- بررسی‌های فرضیه تحقیق
- ۲۴۴ ۷-۹. خلاصه.....

فصل ۱۰. همترازسازی آزمون

- ۲۴۷ ۱-۱۰. مقدمه.....
- ۲۴۸ ۲-۱۰. طرح‌هایی برای همترازسازی.....
- ۲۴۹ ۳-۱۰. شرط‌هایی برای همترازسازی.....
- ۲۵۱ ۴-۱۰. روش‌های کلاسیک همترازسازی.....
- ۲۵۴ ۵-۱۰. همترازسازی از طریق نظریه سوال پاسخ.....
- ۲۵۹ ۶-۱۰. تعیین ثابت‌های همترازسازی.....
- ۲۵۹ ۱-۶-۱۰. روش‌های رگرسیونی
- ۲۶۰ ۱-۶-۲. روش میانگین و سیگما
- ۲۶۱ ۱-۶-۳. روش میانگین و سیگمای نیرومند

۲۶۴ ۱۰-۶-۴. روش منحنی ویژگی

۲۶۵ ۱۰-۷. روش های استفاده شده با طرح های همترازسازی

۲۶۶ ۱۰-۷-۱. طرح تک گروهی

۲۶۶ ۱۰-۷-۲. طرح گروهی معادل

۲۶۷ ۱۰-۷-۳. طرح آزمون لنگر

۲۶۸ ۱۰-۸. همترازسازی نمره واقعی

۲۷۰ ۱۰-۹. همترازسازی نمره مشاهده شده با استفاده از نظریه سوال پاسخ

۲۷۵ ۱۰-۱. گامی در همترازسازی با استفاده از نظریه سوال پاسخ

۲۷۸ ۱۰-۱۱. خ

۲۸۳ فصل ۱۱. ساخت آزمون

۲۸۳ ۱۱-۱. مقدمه

۲۸۴ ۱۱-۲. تعمیم آزمون های استفاده شده در مدل های سوال پاسخ

۲۸۵ ۱۱-۲-۱. حوزه آزمون سوال ها

۲۸۸ ۱۱-۲-۲. انتخاب سوال

۲۹۶ ۱۱-۲-۳. پایایی آزمون

۲۹۸ ۱۱-۳. طراحی مجدد آزمون ها

۳۰۱ ۱۱-۴. مقایسه پنج روش انتخاب سوال

۳۰۱ ۱۱-۴-۱. تولید منبع سوال

۳۰۴ ۱۱-۴-۲. روش های انتخاب سوال

۳۰۵ ۱۱-۴-۳. نتایج

۳۰۹ ۱۱-۵. انتخاب سوال های آزمون برای برازش منحنی های هدف

۳۰۹ ۱۱-۵-۱. روش

۳۱۳	۱۱-۵-۲. نتایج
۳۱۴	۱۱-۵-۳. خلاصه
۳۱۴	۱۱-۶. خلاصه

فصل ۱۲. بانکداری سوال ۳۱۵

۳۱۵	۱۲-۱. مقدمه
۳۱۶	۱۲-۲. مدل‌های سوال پاسخ و بانکداری سوال
۳۱۸	۱۲-۳. انتخاب سوال آزمون مرجع معیار
۳۲۵	۱۲-۴. کاربرد نظریه سوال پاسخ برای هنجار کردن
۳۳۴	۱۲-۵. ارزیابی سیستم پیش‌بینی یک نمره آزمون
۳۳۵	۱۲-۵-۱. روش
۳۳۸	۱۲-۵-۲. نتایج
۳۴۴	۱۲-۵-۳. نتیجه‌گیری‌های حاصل از مساعده

فصل ۱۳. کاربردهای گوناگون ۳۴۷

۳۴۷	۱۳-۱. مقدمه
۳۴۷	۱۳-۲. ارزیابی سوال
۳۶۵	۱۳-۳. آزمون انطباقی
۳۷۱	۱۳-۴. وزن‌دهی افتراقی گزینه‌های پاسخ
۳۷۳	۱۳-۵. برآورد نمره‌های توان
۳۷۵	۱۳-۶. خلاصه

فصل ۱۴. نکاتی عملی در استفاده از مدل‌های IRT ۳۷۷

۳۷۷	۱۴-۱. مقدمه
۳۷۸	۱۴-۲. قابلیت اجرای نظریه سوال پاسخ

- ۳۸۰.....انتخاب مدل ۳-۱۴
- ۳۸۲.....گزارش نمره‌ها ۴-۱۴
- ۳۸۳.....نتیجه‌گیری ۵-۱۴
- ۳۸۴.....پیوست A: مقادیر $ex/1 + ex$ برای ۱۴-۶.

- ۳۸۵ نام‌نامه
- ۳۹۵ واژه‌نامه انگلیسی به فارسی
- ۴۰۷ واژه‌نامه فارسی به انگلیسی
- ۴۱۹ منابع

پیشگفتار مولف

در دهه ۱۹۷۰، نظریه سوال پاسخ جذابی برای مطالعه توسط متخصصین اندازه گیری شد. اما پیدایش نظریه سوال پاسخ به اواسط ۳۰ سالگی و اوایل ۴۰ سالگی برمی گردد. درواقع عبارت "منحنی ویژگی سوال" که یکی از مفاهیم اصلی نظریه سوال پاسخ است را می توان به لیدیارد توکر^۱ در ۱۹۴۲ نسبت داد. با وجود این در اوایل تلاش های تحقیقاتی، علاقه به کار در زمینه نظریه سوال پاسخ تا دهه ۱۹۶۰ طول کشید و تعمیم نظریه نمره واقعی نیرومند پدید آمد. با وجود اینکه نظریه نمره واقعی به سرعت تعمیم یافت و توجه متخصصین روانسنجی را به خود جلب کرد ولی مشکلات و ضعف های ذاتی در تهیه آن باعث به وجود آمدن نگرانی هایی شد. مشکلاتی از قبیل عدم تغییرناپذیری پارامترهای سوال در همه گروه های آزمودنی و نارسایی روش های آزمون کلاسیک برای شناسایی اریبی سوال یا تهیه یک مبنای صوتی برای اندازه گیری در "آزمون مناسب" باعث شدند که نظریه سوال پاسخ مورد توجه واقع شود.

همان طور که اکنون می دانیم انگیزه برای تعمیم نظریه سوال پاسخ توسط لرد از طریق کارهای پیشگام وی فراهم آمد (لرد، ۱۹۵۲؛ ۱۹۵۳a؛ ۱۹۵۳b). در ادامه نگاه به دلیل پیچیدگی ریاضیاتی موضوع و عدم دسترسی به برنامه های کامپیوتری، پیشرفتی به آندی انجام شد و کار پیرزحمتهی بود. کار دیگری که توسط لرد (۱۹۶۸) انجام شد، نوشتن پنج فصل در مورد نظریه سوال پاسخ در کتاب نظریه های آماری نمره های آزمون روانی^۲ در ۱۹۶۸ بود که این کتاب توسط لرد و ناویک نوشته شده است و کار بنجامین رایب و دارل بوک و دانشجویهای آنها در دانشگاه شیکاگو بیانگر علاقه دوباره ای در این زمینه هستند که منجر به سیل عظیمی از مقاله های تحقیقاتی و کاربردهای نظریه شد. موضوع های ویژه مجله اندازه گیری آموزشی (۱۹۷۷) و اندازه گیری کاربردی روانشناسی (۱۹۸۲) راجع به نظریه سوال پاسخ و کاربردها اختصاص داده

^۱ Ledyard Tucker

^۲ Statistical Theories of Mental Test Scores

شد. کتاب راهنمایی در مورد تحلیل مدل راش، بهترین طرح آزمون توسط رایت و استون (۱۹۷۹) چاپ شد. مهم‌ترین دستاورد این تلاش‌ها، کتابی است که توسط لرد در سال ۱۹۸۰ در این رابطه چاپ شد. این کتاب که در واقع اظهارات شخصی اثرگذارترین فرد در زمینه نظریه سوال پاسخ است را باید به عنوان کتابی که به خوبی مطلب را می‌رساند در نظر گرفت.

با توجه به وجود این منابع، آیا نیازی به کتاب دیگری در این زمینه هست؟ لرد (۱۹۸۰a، صفحه xii) اشاره می‌کند که "داوران لزوم کتابی در زمینه نظریه سوال پاسخ که در آن نیازی به درک ریاضی نیست را پیشنهاد خواهند کرد [در کتاب من]." کتاب نوشته شده توسط لرد (۱۹۸۰a) به بلوغ ریاضیاتی قابل توجهی نیاز دارد و بر روی مدل سه پارامتری لوجستیک متمرکز می‌شود. رایت و استون (۱۹۷۹) کار خود را به وسیله متخصص ارزیابی کرده‌اند و به همان سطح پیچیدگی ریاضیاتی نیاز ندارند. اما توجه آنها بر روی مدل راش و کاربردهای آن می‌باشد. نیاز به کتابی است که شرح حلی تکنیکی از موضوع هدف را ارائه نمی‌دهد و اینکه به طور همزمان شامل طرز عمل بسیاری از مدل‌های سوال پاسخ و کاربردهای آنها می‌باشد که واضح به نظر برسند.

این کتاب اهداف سه‌گانه‌ای دارد. ما تلاش کرده‌ایم که ارائه غیرتکنیکی از موضوع هدف، طرز رفتار فراگیر مدل‌ها و کاربردهای آنها، گام‌های خاصی برای استفاده از مدل‌ها در انتخاب کاربردهای امیدبخش ارائه دهیم. گرچه هدف ما، این طرز رفتار غیرتکنیکی از هدف، رفتار ریاضیاتی و آماری است ولی گاه‌گاهی با توجه به اهمیت نظریه، این امر اجتناب‌ناپذیر است. اما تلاش کرده‌ایم که سطح ریاضیات و آمار را در حد مینیمم نگه داریم به طوری که خواننده‌ای که مقاطع مبتدی را در آمار می‌گذراند قادر خواهد شد که اکثر مطالب مطرح شده را دنبال کند. تحلیل‌های ریاضیاتی موجود در بخش‌های مربوط به فصل‌های ۱، ۲ و ۳ بدون اینکه مطلبی درک نشود می‌توان حذف کرد. با این وجود آشنایی با نظریه آزمون تدرسیک و اصول اندازه‌گیری مطلوب است.

این کتاب در چهار بخش سازمان یافته است:

- معرفی نظریه سوال پاسخ، مدل‌ها و فرضیه‌های نظریه سوال پاسخ؛
- مقیاس‌های توانایی، برآورد توانایی، توابع اطلاع و واسنجی آزمون‌ها؛
- بررسی‌های برازش مدل به داده‌ها؛
- همترازسازی، تعمیم آزمون و چندین کاربرد مهم دیگر.

بیشترین عملیات ریاضیاتی مربوط به بخش دوم کتاب است. این مورد را لازم و معتبر در نظر می‌گیریم. جایی که احساس شود که نتوان یک مثال عقلانی برای شرح اصول ارائه داد، توضیح‌هایی از کاربردها آورده شده است.

مساله‌ای که با آن روبه‌رو می‌شویم مربوط به نمادهاست. در یک توسعه سریع زمینه‌ای از قبیل نظریه سوال پاسخ، نمی‌توان همسانی نمادها را انتظار داشت. به جای اینکه از وفور نمادها استفاده شود طرح نمادی به کار گرفته شده در این کتاب تنها براساس لرد (۱۹۸۰a) است. امید داریم تا حدودی مساله تنوع نمادی در آینده کاهش یابد.

مواردی از قبیل حجم و هزینه کتاب، عواملی برای انتخاب محتوا و عمق پوشش بودند. بنابراین در کل نتاد قادر نبودیم که موضوع‌های مطرح شده، همه مطالب را بیاوریم. لزوماً بیشترین توجه محققین و ساینده‌های آزمون، مدل‌های آزمون یک، دو و سه پارامتری لوجستیک هستند. از آنجایی که در برد این سه مدل لوجستیک اطلاعات بیشتری موجود است و این مدل‌ها مورد استفاده اساسی هستند در آثار ما، به آنها اهمیت ویژه‌ای داده شده است. اما هیچ قید و شرطی نداریم در اینکه ممکن است رأیند مدل‌های سوال پاسخ دیگری به جای مدل‌های یک، دو و سه پارامتری لوجستیک تولید شوند از جمله کاربردهای مهم مدل‌های سوال پاسخ، همترازسازی آزمون، تعمیم آزمون و آریبی سران است که جزئیات توصیف شده‌اند ولی در مورد آزمون انطباقی، اندازه‌گیری مرجع-معیار و عدم تناسب اندازه‌گیری در این کتاب مطالب کافی آورده نشده است. خوانندگان می‌توانند برای اطلاعات بیشتر به مقاله ما با این موضوعات به هارنیش و تاتسوکا (۱۹۸۳)، لوین و روبین (۱۹۷۹)، لرد (۱۹۸۰a) و وایز (۱۹۸۳) مراجعه کنند.

وابستگی ما به آثار لرد کاملاً آشکار است. در حال حاضر با هدایت از اندام گرفته‌ایم و از بحث با او بسیار بهره‌مند شده‌ایم. اما مربیان ما یعنی راس تراب و رودریک مک‌دونالد بر این مقدمه اثر گذاشته‌اند. این افراد کسانی بودند که با نگرانی خود برای آموزش و پرورش و ارتقای روش‌ها و اوراق نمونه بورس تحصیلی خود، تفکر ما را تحت تأثیر قرار دادند. ما سعی کردیم که به گفته رودریک مک‌دونالد که موضوع مهمی را به ما یاد داد وفادار بمانیم که کار ما نباید "خطری برای خود و مزاحمتی برای دیگران باشد". بعدها چندین دانشجوی فارغ‌التحصیل پیشین ما از جمله کوک، دانیل ایگنر، گیفرد، هاتن، میلز، موری و سایمون با مشارکت در چندین پروژه تحقیقاتی، کمک شایانی در درک ما از نظریه سوال پاسخ کردند. گیفرد و موری از طریق انجام بسیاری از محاسبات استفاده شده در مثال‌ها و بازبینی نسخه خطی خدمات ارزشمندی را به ما ارائه دادند. مخصوصاً به برناردت مک‌دونالد و سیندی فیشر مدیون هستیم که با صبر و زحمت فراوان، تجدیدنظرهای

متعددی از متن، فرمول‌ها و جدول‌ها انجام دادند. اما متأسفانه اشتباهاتی که باقی می‌ماند مربوط به خود ما می‌باشد.

اکثر تعمیم‌ها و بررسی‌های گزارش شده در اینجا به ما نسبت داده شده است که با حمایت پرسنل و شعبه آموزشی دفتر تحقیقاتی ناول ممکن شده است (شماره قرارداد: N-۷۹-۰۰۱۴-۰۰۳۹)، منابع آزمایشگاهی (شماره قرارداد: FQ۷۶۲۴-۷۹-۰۰۱۴) و موسسه ملی آموزش (شماره قرارداد: ۲۰۳۱۹-۸۱-۰۲). ما از این سازمان‌ها و مخصوصاً از چارلز داویس، مال کلمبری، برایان واترز، فیل دلو، روج پنال به خاطر حمایت‌ها و تشویق‌هایشان از صمیم قلب سپاسگزاریم.

پیشگفتار مترجمین

در نظام‌های آموزشی با مفاهیم اندازه‌گیری، ارزشیابی و سنجش بسیار کار می‌شود. واژه سنجش به دو تعبیر به کار برده می‌شود: (۱) اندازه‌گیری (مانند دماسنج، فشارسنج و ...) و (۲) ارزشیابی (به معنی اندازه‌گیری و سپس قضاوت کردن مانند سنجیده سخن گفتن و ...). اندازه‌گیری عبارت از فرایندی است که تعیین می‌کند یک شخص یا یک شیء چه مقدار از یک ویژگی مانند اندازه‌گیری وزن، سرعت، هوش، پیشرفت تحصیلی و غیره برخوردار است. بنابراین اندازه‌گیری در برگزیده قواعدی است که اعداد را به اشیاء یا افراد تخصیص می‌دهد به‌طوری‌که خصیصه‌هایی از افراد یا اشیاء را برای قضاوت و مقایسه به صورت کمیت نشان دهد. روشن است تعبیر مورد استفاده در این کتاب، تعبیر دوم سنجش است از طرفی ارزشیابی، به یک فرایند نظام‌داری گفته می‌شود که براساس آن اطلاعات به منظور تعیین آنک، اسباب مورد نظر تحقق یافته‌اند یا در حال تحقق یافتن هستند و یا به چه میزانی تحقق یافته‌اند جمع‌آوری و تحلیل و تفسیر می‌شوند.

ارزشیابی بر دو نوع ارزشیابی از کار معلم و ارزشیابی پیشرفت تحصیلی یادگیرندگان است. ارزشیابی از کار معلم یکی از ارزشیابی‌های مهم آموزشی به شمار می‌آید که در بهبود کیفیت آموزشی نقش بسزایی دارد. سنجش عملکرد یادگیرندگان و مقایسه نتایج حاصل با اهداف آموزش از پیش تعیین شده، به منظور تصمیم‌گیری در این باره که آیا فعالیت‌های آموزشی معلم و کوشش‌های یادگیری یادگیرندگان به نتایج مطلوب انجامیده‌اند و یا این فعالیت‌ها چقدر در یادگیری یادگیرندگان تاثیر داشته است. در ارزشیابی پیشرفت تحصیلی دو اقدام (۱) تعیین اهداف آموزش و (۲) سنجش عملکرد یادگیرندگان ضروری است.

در علوم انسانی به خصوص روانشناسی و علوم تربیتی و حتی مدیریت، آزمون‌ها مهمترین و پرکاربردترین ابزار سنجش و اندازه‌گیری به حساب می‌آیند. آزمون وسیله‌ای عینی و استاندارد شده است که برای اندازه‌گیری نمونه‌ای از رفتار یا خصایص آزمودنی به کار می‌رود. در این تعریف مراد از عینی بودن آزمون آن است که روش اجرای آزمون، نمره دادن و تعبیر و تفسیر نتایج آزمون براساس قواعد معین و مشخص صورت می‌گیرد و قضاوت و نظر شخصی در آنها بی‌تاثیر

است و اصطلاح استاندارد شده بدین معناست که آزمون از قبل در مورد گروه نمونه‌ای از افراد موردنظر در بوته آزمایش گذاشته شده و نتایج پژوهش‌های مربوط به آن از طریق روش‌های آماری مورد تجزیه و تحلیل قرار گرفته و پایایی و اعتبار آن تعیین شده است. به طور کلی آزمون ابزاری برای به دست آوردن نمونه‌ای از رفتار فرد است. کاربرد آزمون‌ها چنان در مدارس، کلینیک‌ها، صنعت و سازمان‌های دولتی کشورهای جهان اول گسترش یافته است که مشکل بتوان فردی را یافت که صدها آزمون برای او اجرا نشده باشد. آزمون‌ها برای نیل به هدف‌های گوناگون طرح‌ریزی شده‌اند. آنها را می‌توان برای گزینش به کار برد؛ مانند هنگامی که از بین گروهی داوطلب عده‌ای برای استخدام انتخاب می‌شوند. از آزمون‌ها می‌توان برای طبقه‌بندی نیز استفاده نمود برای مثال می‌توان آزمون‌ها را در یکی از گروه‌های فعال، نیمه فعال و یا تنبل قرار داد. این طبقه‌بندی را می‌توان در ایگزینی نیز اعمال کرد مانند زمانی که دانش‌آموزی در یک کلاس پیشرفته زبان انگلیسی جایابی می‌شود. آزمون‌ها را می‌توان برای ارزشیابی نیز استفاده نمود مانند زمانی که نمره کلاسی به دانش‌آموزان داده می‌شود. به طور کلی هر آزمون سودمند، یک ویژگی یا رفتار را با دقت اندازه‌گیری می‌کند.

در یک آزمون روش‌های مختلفی برای طرح سوال امکان‌پذیر است. سوال‌ها ممکن است به صورت پاسخ آزاد^۱ تهیه شوند که در آن از آزمودنی‌ها خواسته می‌شود با استفاده از مفاهیم و درکی که از مطالب دارند پاسخ سوال را به صورت تشریحی بنویسند. استفاده از چنین سوال‌هایی باعث می‌شود تا آزماینده برداشت دقیقی از آنچه آزمون می‌داند یا برداشت می‌کند به دست آورد. با این حال تنظیم کردن سوال‌هایی با پاسخ آزاد و تشریحی غالباً دشوار است و هر نوع روش نمره‌گذاری که تدوین و تنظیم می‌شود نسبتاً پرهزینه است.

سوال‌ها می‌توانند به صورت چندگزینه‌ای نیز باشند. این نمره‌گذاری^۲ به آزمودنی امکان می‌دهد تا پاسخ خود را از بین دو یا چند گزینه انتخاب کند. نمره‌گذاری و تنظیم کردن سوال‌های چندگزینه‌ای آسان است و بیشتر با ماشین نمره‌گذاری می‌شوند که نتیجه‌اش صرفه‌جویی در مقدار قابل توجهی هزینه و کار انسانی است. برخی از متقدنان احساس می‌کنند که سوال‌های چندگزینه‌ای اساساً سطحی هستند و یا کاربرد آنها بیشتر به ضرر آزمودنی‌های خلاق است. با این وجود چنانچه سوال‌های چندگزینه‌ای با دقت طرح شوند قادر هستند فرایندهای پیچیده تفکر، درک ظریف معانی و خلاقیت آزمودنی را اندازه‌گیری کنند (مک دونالد، ۱۹۹۹).

هدف از هر نظریه آزمون، شرح استنباط‌هایی از چگونگی پاسخ به سوال‌ها و یا نمره‌های آزمونی است که می‌توان در مورد ویژگی‌های غیرقابل مشاهده آزمودنی نظر داد و یا اندازه‌گیری خصیصه‌هایی که توسط یک آزمون می‌باشد. از جمله نظریه‌های اندازه‌گیری رایج، نظریه آزمون کلاسیک و نظریه سوال‌پاسخ هستند. نظریه سوال‌پاسخ به دلیل محدودیت‌هایی که در نظریه آزمون کلاسیک وجود داشته پایه‌ریزی شده است. مطالعات نشان می‌دهد علیرغم اینکه بسیاری از آزمون‌های پیشرفت تحصیلی، استعداد و شخصیت با به کارگیری مدل‌ها و رویکردهای کلاسیک ساخته شده‌اند ارزشیابی آزمون‌های روانی و تربیتی که معمولاً براساس مدل‌های کلاسیک بنا نهاده شده‌اند محدودیت‌های زیادی دارند. به این دلیل متخصص‌های روانسنجی سیستم اندازه‌گیری جدیدی، همان نظریه سوال‌پاسخ است را برای مقابله با این محدودیت‌ها ابداع کردند. هر دو نظریه در ساخت و تحلیل آزمون‌ها و پرسشنامه‌ها کاربرد زیادی دارند. گرچه نظریه سوال‌پاسخ، نقطه‌ضعف‌های ادعای نظریه کلاسیک را جبران می‌نماید ولی دلایل قطعی برای حذف کامل این نظریه وجود ندارد. این نظریه همچنان در حوزه‌های متعدد علمی مورد استفاده قرار می‌گیرد. از سویی دیگر، برخی معتقد هستند که نظریه سوال‌پاسخ رقیب نظریه کلاسیک نیست بلکه تعمیم نظریه کلاسیک. چ‌چوب مدل‌های پیچیده‌تر و کامل‌تری می‌باشد. کتاب حاضر در راستای معرفی و ترویج تحریک سوال‌پاسخ انتخاب و ترجمه شده است. این کتاب برای آموزش مطالب اساسی نظریه سوال‌پاسخ بسیار مناسب است. جدول‌ها و شکل‌های موجود در این کتاب شامل تحلیل‌های واقعی نظریه سوال‌پاسخ هستند. تلاش شده است که مطالب کتاب به زبانی ساده و روان بیان شوند. درخاتمه برخورد لازم می‌دانیم که از خانم شهناز پارسایان که در ترجمه و خانم مریم مهدوی که در نحوه ویرایش و تایپ، کتاب کمک شایانی کردند تشکر و قدردانی نماییم. مطمئناً این ترجمه بدون نقص و اشتباه نیست لذا از خوانندگان محترم خواهشمندیم از هرگونه نظر و انتقادی ما را بی‌بهره نگذارند.