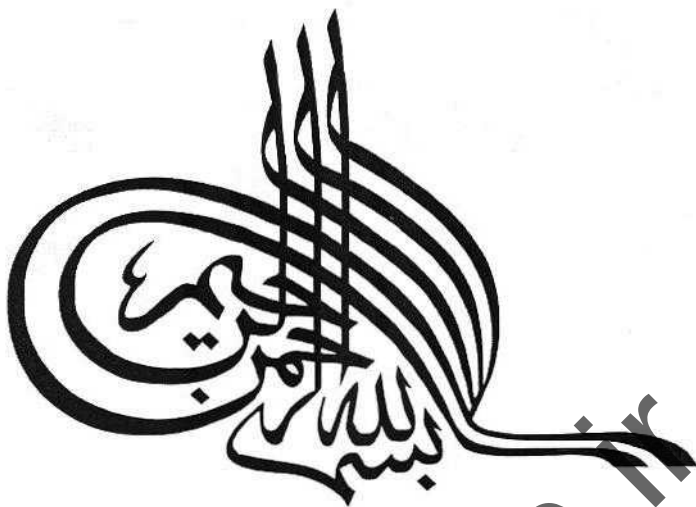




ساخت و اعتباریابی سوال‌های چندگزینه‌ای

توماس ام. هالادینا (دانشگاه ایالتی آریزونا)

ویرایش سوم

مترجمین:

سیده طیبه مطیعی لنگرودی (سازمان سنجش آموزش کشور)

دکتر داود برزآبادی فراهانی (دانشگاه تهران)

دکتر غلامرضا یادگارزاده (سازمان سنجش آموزش کشور)

علیرضا نظری انامق

سرشناسه: هالادینا، توماس ام. Haladyna, Thomas M.

عنوان و نام پدیدآور: ساخت و اعتباریابی سوال‌های چند گزینه‌ای

مشخصات نشر: تهران، چوک آشتیان، ۱۳۹۴.

مشخصات ظاهری: ۴۲۶ ص.

شابک: ۲۴۰۰۰۰ ریال: ۴-۳۹-۵۲۹۰-۶۰۰-۹۷۸.

وضعیت فهرست‌نویسی: فیای مختصر

یادداشت: فهرست‌نویسی کامل این اثر در نشانی: <http://opac.nlai.ir> قابل دسترسی

است.

یادداشت: عنوان اصلی: Developing and validating multiple-choice test items (third edition)

یادداشت: مترجمین سیده طیبہ مطیعی لنگرودی، داود برزآبادی فراهانی، غلامرضا یادگارزاده، علیرضا نظری انامق.

شناسه افزوده: مطیعی لنگرودی، سیده طیبہ، ۱۳۳۷، مترجم

شماره کتاب‌شناسی ملی: ۳۷۸۹۲۶

عنوان کتاب: ساخت و اعتباریابی سوال‌های چندگزینه‌ای

ناشر: انتشارات چوک آشتیان

مترجمین: سیده طیبہ مطیعی لنگرودی (سازمان سنجش آموزش کشور)؛ دکتر داود برزآبادی فراهانی (دانشگاه تهران)؛ دکتر غلامرضا یادگارزاده (سازمان سنجش آموزش کشور)؛ علیرضا نظری انامق

صفحه آرا: ژیلّا حسینی

ویراستار: سیده سارا نیک‌پناه

لیتوگرافی و چاپ: چاپ نشاط

نوبت چاپ: اول

تیراژ: ۱/۰۰۰ نسخه

قطع: وزیری

قیمت: ۲۴۰/۰۰۰ ریال

مرکز بخش: stm190@gmail.com

شابک: ۴-۳۹-۵۲۹۰-۶۰۰-۹۷۸

کلیه حقوق این اثر متعلق به نشر چوک آشتیان می‌باشد.

فهرست مطالب

مقدمه.....	۵
بخش اول: پایه و اساس سوال‌های چندگزینه‌ای.....	۹
فصل ۱: نقش سوال در روایی آزمون.....	۱۱
فصل ۲: محتوا و فرایندهای شناختی.....	۳۵
فصل ۳: قالب‌های سوال.....	۶۷
بخش دوم: ساخت سوال‌های چندگزینه‌ای.....	۱۰۷
فصل ۴: قالب‌های چندگزینه‌ای.....	۱۰۹
فصل ۵: راهبردهایی برای ساخت سوال‌های چندگزینه‌ای.....	۱۵۱
فصل ۶: ارائه مدل‌هایی از سوال‌های مناسب و قالب سوال‌های خلاقانه.....	۱۹۳
فصل ۷: ساخت سوال.....	۲۱۹
بخش سوم: به‌دست آوردن روایی از طریق فرایند ساخت سوال و اعتباریابی نظریه سوال - پاسخ.....	۲۶۳
فصل ۸: به‌دست آوردن روایی از طریق فرایند ساخت سوال.....	۲۶۵
فصل ۹: به‌دست آوردن روایی از طریق بررسی آماری نظریه سوال - پاسخ.....	۲۹۱
فصل ۱۰: استفاده از الگوهای پاسخ سوال در بررسی مسائل خاص.....	۳۲۹
بخش چهارم: آینده آزمون‌سازی و اعتباریابی پاسخ‌ها.....	۳۶۹
فصل ۱۱: رویه‌های جدید در نگارش سوال و اعتباریابی پاسخ سوال‌ها.....	۳۷۱
فهرست منابع.....	۳۹۵

مقدمه

سومین ویرایش کتاب حاضر، حاوی دستورالعمل‌های جامع برای ساخت سوال‌های چندگزینه‌ای است. این کتاب به منظور اصلاح و بهبود روند ساخت سوال‌های چندگزینه‌ای، به بررسی پاسخ سوال‌ها نیز می‌پردازد. امید است که تجدیدنظرها و اصلاحات صورت گرفته در این ویرایش، نقاط ضعف و اشکالات دو نسخه قبل را پوشش دهد.

دلایل انجام ویرایش سوم

ویرایش و اصلاح یک کتاب برای چندمین بار، نیاز به دلایل موثری دارد. در مورد این کتاب، چهار عامل را می‌توان دلیل انجام ویرایش سوم دانست:

دلیل اول به میزان علاقه خوانندگان به حیطه ساخت سوال‌های چندگزینه‌ای و اعتباریابی پاسخ سوال‌های آنها برمی‌گردد. این گونه مخاطبان راه‌های بی‌شماری را برای بهبود فرایند ساخت این گونه سوال‌ها به من نشان می‌دهند. برخی از آنها اشتباهات موجود در کتاب را به من گوشزد می‌کنند، ایده‌های خود را با من درمیان می‌گذارند، قالب‌های جدیدی از سوال‌های چندگزینه‌ای را پیشنهاد می‌کنند، سوال می‌پرسند و یا نظراتی ارائه می‌دهند که در بهبود سوال‌های چندگزینه‌ای مفید و مؤثر واقع می‌شوند.

دلیل دوم، این است که منابع علمی کافی و مبنای نظری قوی برای ساخت سوال وجود ندارد (کروناخ، ۱۹۷۰؛ هالادینا و داوینینگ، ۱۹۸۹a، ۱۹۸۹b؛ هالادینا، داوینینگ و رودریگوئز، ۲۰۰۲؛ نیتکو، ۱۹۸۵؛ روید و هالادینا، ۱۹۸۲). منتقدین به کمبود پژوهش در حیطه ساخت سوال اشاره کرده‌اند و می‌توان گفت کتاب حاضر، پاسخی به این انتقادات است.

با مطالعه سوابق قبلی مشخص می‌شود که پژوهش‌های مربوط به ساخت سوال‌های چندگزینه‌ای، در گذشته دارای غنای لازم نبوده‌اند و این مسئله را می‌توان به‌عنوان سومین دلیل ویرایش کتاب حاضر در نظر گرفت. پژوهش‌های صورت گرفته، به اوایل قرن بیستم یعنی زمانی که سوال‌های چندگزینه‌ای برای اولین بار معرفی شدند، بازمی‌گردد. از آن زمان به بعد، متخصصان و صاحب‌نظران بسیاری، نظرات، تجربیات و دست‌نوشته‌های خود را برای بهبود فرایند ساخت سوال‌های چندگزینه‌ای در اختیار یکدیگر قرار دادند. این کتاب نتیجه چنین همکاری‌هایی است.

چهارمین و آخرین دلیل، به تجربه کاری ۳۰ ساله من در زمینه طراحی، اجرا، ارزشیابی برنامه‌های سنجشی و تدریس در مقاطع تحصیلی مختلف برمی‌گردد. در اختیار داشتن چنین تجربه‌ای به من کمک نموده تا درک بهتری از فرایند ساخت سوال‌های چهارگزینه‌ای، مزایا و معایب این سوالات و اهمیت اعتباریابی پاسخ‌های آنها به‌دست آورم.

مخاطبین این کتاب

کتاب حاضر، برای آن دسته از افرادی در نظر گرفته شده که به‌طور جدی در زمینه طراحی و ساخت سوال مشغول فعالیت هستند. دانشجویانی که در دوره تحصیلات تکمیلی رشته سنجش آموزشی قرار دارند نیز می‌توانند از این کتاب، برای درک بهتر دو مفهوم مهم در طراحی آزمون، یعنی ساخت سوال و اعتباریابی پاسخ آن، استفاده نمایند. این کتاب برای افرادی که مستقیماً درگیر طراحی آزمون هستند نیز مناسب است.

محدودیت‌های کتاب حاضر

هرچند این کتاب، دستورالعمل‌های جامعی را برای ساخت سوال‌های چندگزینه‌ای و اعتباریابی پاسخ سوال‌ها ارائه می‌کند، اما دارای محدودیت‌هایی نیز هست.

نظریه‌های آماری بسیاری در زمینه اعتباریابی پاسخ سوال و ساخت مقیاس وجود دارد و پژوهش‌ها و فعالیت‌های بسیاری نیز در این حیطه صورت گرفته است. تکنیک این کتاب در نگارش سوال، بر هیچ‌یک از نظریه‌های کنونی استوار نیست. امید است پیشرفت‌های آتی در زمینه نظریه‌های مربوط به ساخت سوال و اعتباریابی پاسخ سوال‌ها، باعث ایجاد هماهنگی بیشتری میان مسائل ذکر شده و نظریه یادگیری شناختی شود. برای اینکه بتوانیم در حیطه‌هایی چون مفهوم‌سازی سازه مورد سنجش، ساخت سوال، استفاده از نتایج حاصله و ارزشیابی پاسخ‌ها مباحثی ارائه کنیم، باید ابتدا در نظریه‌ها پیشرفتی منطقی حاصل بیاید. نظرات و عقاید موجود درباره هوش انسان، دائماً در حال تغییر و تحول است. علاقه به سنجش توانایی‌های شناختی افراد، ما را بر آن داشته که از جدیدترین روش‌ها برای سنجش دوره‌های آموزشی و تحصیلی استفاده کنیم. نحوه نگرش ما به یادگیری انسان، وابسته به این است که وی چگونه می‌آموزد و چگونه مورد آزمون قرار می‌گیرد. سه دیدگاه معروف درباره یادگیری انسان عبارتند از: دیدگاه رفتاری، دیدگاه شناختی و دیدگاه ساختارگرایی اجتماعی. از آنجایی که نظریه یادگیری رفتاری، کاربرد بسیار زیادی در موضوعات آموزشی، یادگیری در حد تسلط [یادگیری در سطح بسیار بالا] و سنجش ملاک - مرجع دارد، رونق بسیار زیادی یافته است. اخیراً، دیدگاه ساختارگرایی اجتماعی در حیطه سنجش، مورد استقبال عام قرار گرفته است. توانایی‌های شناختی، در طول زندگی افراد به تدریج در حال رشد و گسترش هستند. روان‌شناسان شناختی و نظریه‌پردازان آماری، با همکاری یکدیگر شروع به سنجش چنین توانایی‌هایی کرده‌اند. در این کتاب، توانایی شناختی، پارادایم مفیدی برای تعریف یادگیری و سنجش آن ارائه می‌دهد. تا زمانی که مفهوم توانایی‌های شناختی مورد پذیرش عام قرار نگیرد و تبدیل به دیدگاهی عملیاتی و کاربردی نشود، تعارضاتی درباره نحوه و چگونگی اجرای بهترین نوع سنجش و ارائه آموزش وجود خواهد داشت. نظریه یادگیری رفتاری تقابلات شدیدی با دو دیدگاه یادگیری شناختی و ساختارگرایی اجتماعی دارد. به نظر می‌رسد که دو دیدگاه اخیر، نقاط اشتراک زیادی با یکدیگر دارند، درحالی که اختلاف بنیادینی میان آنها با دیدگاه رفتاری وجود دارد. در نتیجه، برای افرادی مانند ما که در کار تدریس هستیم، روش سنجش آموخته‌های دانشجویان باید تاحدودی تابع روشی باشد که آنها در یادگیری مورد استفاده قرار می‌دهند.

بعید به نظر می‌رسد به آسانی بتوان این‌گونه مشکلات را حل و فصل کرد. ساخت سوال و اعتباریابی پاسخ آن، حیطه‌ای نوظهور است و نیاز به یک نظریه یکپارچه و منسجم دارد. بدون شک طراحی سوال، همچنان به عنوان یک عمل خلاقانه قلمداد می‌شود.

وضعیت کنونی آزمون‌های چندگزینه‌ای

شاید در گذشته، استفاده بیش از حد از سوال‌های چندگزینه‌ای باعث شد تا منتقدانی مانند فردریکسن (۱۹۸۴) و شیپارد (۱۹۹۳ و ۲۰۰۰) ادعا کنند که آزمون‌های چندگزینه‌ای، نوع خاصی از آموزش را می‌طلبند. فردریکسن خاطرنشان نمود که به دلیل تصور رایج در مورد قالب سوالات چندگزینه‌ای (اینکه استفاده از این قالب، برای سنجش میزان یادگیری دانش‌آموزان

نسبت به روش‌های دشوارتر سنجش، کاری آسان است)، آسیب‌های ناخواسته‌ای به افراد وارد می‌آید. شیبپارد نیز معتقد است تأکید بیش از اندازه بر روی حافظه در سنجش افراد، ممکن است باعث شود که مسائل مهم‌تر، مانند کاربرد دانش و مهارت‌های لازم در موقعیت‌های واقعی زندگی نادیده گرفته شوند.

قطعاً، بررسی و تحلیل آنچه که معلم یاد می‌دهد و می‌سنجد، این ایده را تقویت می‌کند که بخش اعظمی از محتوای آموزشی تنها حافظه سطحی افراد را مورد سنجش قرار می‌دهد. چنین ایرادی را نمی‌توان بر گردن سوال‌های چندگزینه‌ای انداخت. همان‌طور که در این کتاب بارها تأکید شده است، برنامه تحصیلی، روش تدریس و یادگیری افراد باید بر نتایجی تکیه کنند که از لحاظ شناختی پیچیده باشند و انعکاسی از توانایی‌های افراد را در زمینه خواندن، نوشتن، صحبت کردن، گوش دادن، ریاضی، حل مسائل علمی، تفکر انتقادی و امور اخلاقانه نشان دهند. سوال‌های چندگزینه‌ای نقش مهمی را در این میان برعهده دارند.

علی‌رغم حملات انتقادی صورت گرفته به سوال‌های چندگزینه‌ای، این سوال‌ها در سال‌های اخیر غنی‌تر شده‌اند. آگاهی دادن به سیاست‌گذاران و ارزشیابی‌کنندگان برای حمایت از برنامه‌های سنجشی بسیار ضروری است. شکی نیست که سنجش، سرمایه‌گذاری بزرگی است که به‌صورت مستقیم یا غیرمستقیم تمامی افراد موجود در ایالات متحده را تحت تأثیر خود قرار می‌دهد؛ پدیده‌ای که در ایالات متحده و دیگر مناطق جهان در حال گسترش است (فیلیپس، ۱۹۹۸ و ۲۰۰۰). از آزمون‌های چندگزینه‌ای به دلایل مختلفی استفاده می‌شود: تعیین جایگزین، انتخاب افراد، اعطای گواهی، اعطای مجوز، اخذ تأییدیه مهارت، اعطای رتبه و درجه، تشخیص میزان یادگیری افراد و ...

براساس این کتاب، می‌توان آزمون‌های چندگزینه‌ای را در مقیاس کلاسی، مقیاس‌های بزرگتر (برای سنجش یادگیری دانش‌آموزان) و آزمون‌های تعیین شایستگی در مشاغل مختلف به‌کار برد. بسیاری از طراحان و کاربران آزمون نیازمند آن هستند که از استانداردهای آموزشی و روان‌شناختی سنجش مطلع شوند (انجمن پژوهشی آموزشی آمریکا، انجمن روان‌شناسی آمریکا و مؤسسه بین‌المللی سنجش در آموزش) و همچنین از دستورالعمل‌های آزمون‌های دارای اهمیت بالا که انجمن پژوهشی آموزشی آمریکا منتشر کرده استفاده کنند. در ویرایش سوم این کتاب، این استانداردها و دستورالعمل‌ها در کنار مثال‌هایی که برای ساخت سوال و اعتباریابی پاسخ سوال‌ها مطرح شده، آمده‌اند. زمانی که کاربران آزمون درباره‌ها و نیایدهای استفاده از نتایج آزمون (به این معنی که کجا و چگونه می‌توانند از آنها استفاده کنند)، اطلاعاتی کسب نمایند، آنگاه می‌توانیم کیفیت آزمون‌ها را با استفاده از روش‌های دقیق و کاربردی ساخت سوال و استفاده از روش‌های اعتباریابی مناسب افزایش دهیم.

ساختار کتاب

در ویرایش سوم، تغییرات بیشتری نسبت به ویرایش دوم دیده می‌شود. برای حفظ رویه قبلی، این کتاب نیز در چهار بخش ارائه شده است.

بخش اول، شامل سه فصل است که در مورد پایه و اساس نگارش سوال‌های چندگزینه‌ای صحبت می‌کنند. فصل یک، شامل مبحث بسیار مهم و اساسی در همه آزمون‌ها است: روایی. رابطه مستقیمی میان روایی نمره آزمون و روایی پاسخ‌ها وجود دارد. فرایندی که در اعتباریابی نمره آزمون به‌کار می‌بریم، در اعتباریابی پاسخ‌ها نیز کاربرد دارد و این به آن جهت است که

پاسخ سوال، اساسی‌ترین واحد موجود در سنجش و شکل‌گیری نمره آزمون است. فرایند ساخت سوال، در کنار شواهد مربوط به روایی، نشان‌دهنده کیفیت سوال است. فصل دو به موضوع محتوا و فرایندهای شناختی سوال‌های آزمون می‌پردازد. در این فصل نشان داده می‌شود که دانش، مهارت‌ها و توانایی‌ها به هم پیوسته هستند. این فصل، از روان‌شناسی شناختی اقتباس شده است. برنامه‌های سنجشی موجود، توصیه‌هایی را درباره آنچه که در سنجش نیازمند دانستن آنها هستیم، ارائه می‌دهند. فصل سوم، یک طبقه‌بندی از قالب سوال‌های آزمون ارائه

می‌دهد که حاوی سوال‌های ساخته - پاسخ^۱ (CR) و انواع مختلف سوال‌های چندگزینه‌ای است. همچنین در فصل سه، مباحثی ارائه می‌شود که نشان می‌دهد سوال‌های ساخته - پاسخ و چندگزینه‌ای، قابلیت سنجش محتوا و فرایندهای شناختی مختلفی را دارند. با معرفی چنین زیربنایی در بخش اول، نگرش سوال‌های چندگزینه‌ای در بخش دوم عملی می‌شود.

بخش دوم، به ساخت سوال‌های چندگزینه‌ای اختصاص یافته است. فصل چهارم به ارائه قالب‌های مختلفی از سوال‌های چندگزینه‌ای می‌پردازد. همچنین در این فصل، درباره نوع محتوا و فرایندهای شناختی که هر یک از این قالب‌ها قادر به سنجش آنها هستند، مباحثی مطرح می‌شود. فصل پنجم، دستورالعمل‌های ساخت سوال‌ها در قالب‌های مختلف را ارائه می‌کند. فصل ششم، شامل مثال‌هایی از سوال‌های چندگزینه‌ای خلاقانه است و دلایل و توجیهات مربوط به انتخاب هر یک از آنها را بیان می‌کند. فصل هفتم، راهنمایی‌های تازه‌ای را در حیطه تکنیک‌های ساخت سوال بیان می‌دارد.

بخش سوم، در مورد اعتباریابی پاسخ سوال‌ها است. فصل هشتم، شامل گزارشی است از روش‌های موجود برای بهبود سوال‌های آزمون. فصل نهم، با تحلیل سوال جهت ارزشیابی سروکار دارد. در این فصل، دیدگاه‌های نظریه‌ای مطرح شده‌اند که هر یک یا در حیطه نظریه کلاسیک آزمون قرار می‌گیرند یا در حیطه نظریه سوال - پاسخ. فصل دهم، اطلاعاتی درباره اینکه چگونه بررسی و مطالعه پاسخ سوال‌ها می‌تواند در مطالعه مسائل خاص موجود در سنجش مورد استفاده قرار گیرد، مباحثی را مطرح می‌کند.

بخش چهارم، حاوی فصل یازدهم است که با روش‌های موجود در نگارش و اعتباریابی سروکار دارد. برخلاف آنچه که ما امروزه تجربه می‌کنیم، نظریه‌پردازان شناختی بر روی روش‌هایی کار می‌کنند تا آنچه ما آموزش می‌دهیم و می‌آزماییم را تعریف کنند و نظریه‌پردازان آزمون نیز در حال بررسی نظریه‌های نگارش سوال و مدل‌های پاسخ سوال هستند که برای سنجش رفتارهای پیچیده‌تر استفاده می‌شوند. نتیجه بسیاری از این نظریه‌ها، آموزش و سنجش را عمیقاً دچار دگرگونی خواهد نمود.

در پایان، مسئله ساخت سوال‌های آزمون و اعتباریابی پاسخ‌های آنها، باز هم به عنوان دو گام انتقادی مهم و مرتبط با هم در طراحی آزمون باقی می‌ماند. این کتاب در نظر دارد به خوانندگان کمک کند تا مفاهیم، اصول و روش‌های ساخت سوال‌های چندگزینه‌ای که دارای روایی هستند را بهتر درک نمایند.

توماس هالادینا