

تاریخ دریافت مقاله: ۸۵/۰۱/۳۰

تاریخ بررسی مقاله: ۸۵/۳/۶

تاریخ پذیرش مقاله: ۸۵/۹/۱۲

مجله علوم تربیتی و روانشناسی
دانشگاه شهید چمران اهواز، زمستان ۱۳۸۵
دوره سوم، سال سیزدهم، شماره ۴
ص ص: ۵۷-۷۸

بررسی و تحلیل سطوح حیطة شناختی و شاخصهای روان سنجی امتحان نهایی دروس عربی، حسابان و زیست شناسی دانش آموزان دختر پایه سوم متوسطه در سه منطقه متفاوت اقتصادی - اجتماعی استان خوزستان

دکتر حسین سپاسی*

چکیده

هدف عمده این پژوهش بررسی سطوح شناختی و شاخصهای روان سنجی سؤالات امتحان نهایی دروس عربی، حسابان و زیست‌شناسی در مناطق برخوردار، نیمه برخوردار و غیربرخوردار و کل استان خوزستان بود. تعداد ۱۰۰۰ برگ امتحانی از هر جامعه به طور تصادفی نمونه‌های پژوهش بود. نتایج نشان داد که ویژگیهای روان سنجی سؤالات امتحانات دروس در مقایسه با ملاکهای پیشنهادی قابل قبول بود. ضرایب تمیز اکثر سؤالات امتحانی دروس مذکور بین دو گروه از دانش‌آموزان قوی و ضعیف در سطح احتمال $P < 0/05$ تفاوت گذاشتند. ضرایب پایایی نشان داد که امتحانات از دقت و ثبات کافی برخوردار بودند. اعتبار محتوایی امتحانات حاکی از آن بود که طراحان سؤالات توانسته بودند بین اهداف و محتوای دروس توازن مطلوبی برقرار کنند.

کلید واژگان: حیطة شناختی، ضریب دشواری، ضریب تمیز، پایایی و اعتبار

* عضو هیئت علمی دانشکده علوم تربیتی و روان شناسی دانشگاه شهید چمران اهواز

مقدمه

لهمان^۱، ۱۹۸۴؛ ایبل و فریزی^۲، ۱۹۹۱؛ بوم گارتنر و جکسون^۳، ۱۹۹۱؛ گیج^۴، ۱۹۹۲؛ پیچ و پیترسون^۵، ۲۰۰۳).

برای دسترسی به اهداف فوق، متخصصان آزمون سازی راههای متفاوتی پیشنهاد می‌کنند. آنچه اکثریت بر سر آن توافق دارند این است که این اهداف از طریق شرکت دانش‌آموزان در نوعی آزمون تحقق پذیر خواهد بود. بنابراین، برای این که آزمون بتواند به اهداف فوق جامعه عمل بپوشاند، سازنده سؤالات، معلم، باید از مهارت‌های خاصی برخوردار باشد که ایبل و فریزی (۱۹۹۱) برخی از معلمان را فاقد چنین مهارت‌هایی می‌دانند.

متأسفانه با این که شاهد اجرای آزمون‌های بسیاری در نظام‌های آموزشی و غیرآموزشی هستیم، ولی تاکنون از آزمون‌های مطلوب کمتر استفاده کرده‌ایم. بسیاری از معلمان بدون این که آموزش کافی در سنجش و اندازه‌گیری و شیوه ساخت آزمون‌ها کسب کنند آماده تدریس می‌شوند، و یا این که

آئین نامه‌های امتحانات دوره‌های مختلف تحصیلی به ویژه آئین نامه امتحانات دوره متوسطه برای برگزاری امتحانات دانش‌آموزان به اهداف چندی اشاره می‌کنند. از جمله این اهداف می‌توان به نقش ارزشیابی مستمر به منظور ایجاد انگیزه یادگیری دانش‌آموزان و تشخیص نقاط قوت و ضعف آنان اشاره کرد. همچنین، آگاه کردن دانش‌آموزان از میزان عملکرد تحصیلی و اعلام نتایج امتحانات به اولیاء از اهداف دیگر امتحانات ذکر شده است. علاوه بر این، پرورش استعدادها، پرورش روحیه تلاش، ابتکار و رقابت‌های سالم بین دانش‌آموزان، تشخیص استعدادها و علائق دانش‌آموزان به منظور راهنمایی آنان در فعالیتهای آینده و تحصیلات پیش دانشگاهی و دانشگاهی را می‌توان از اهداف دیگر برگزاری این امتحانات ذکر کرد. نهایتاً، حصول اطمینان از تحقق شرایط لازم به منظور ارتقاء دانش‌آموزان به پایه تحصیلی بالاتر از اهداف این امتحانات ذکر شده است (آئین نامه امتحانات دوره متوسطه نظری، ۱۳۶۸). بررسی سایر منابع نیز اهداف مشابهی برای برگزاری امتحانات مدارس ذکر می‌کنند (سیف، ۱۳۷۲؛ شریفی، ۱۳۷۴؛ مهرنرز و

- 1- Mehrens & Lehman
- 2- Ebel & Frisbie
- 3- Baumgartner & Jackson
- 4- Gage
- 5- Page & Peterson

اطلاعات آنان درباره ساخت آزمون و ویژگیهای آن به زمان دانشجویی و یا به آنچه از همکاران خود دیده یا شنیده‌اند و یا به تجربه‌های کوتاه مدت دوران تحصیلاتشان خلاصه می‌شود. وقتی از دانش‌آموزان درباره سؤالات امتحانی معلمانشان پرسیده شد پاسخ دادند سؤالات نامفهوم بود، سؤالات بسیار دشوار بود و یا سؤالات بسیار آسان و به پرسش مفاهیم ساده و پیش پا افتاده که فقط به حفظ طوطی‌وار مطالب نیاز پیدا می‌کرد، اختصاص داشت (مهرنزی و لهما، ۱۹۸۴؛ تاگوموری^۱، ۱۹۹۳).

پرداختن به عوامل گوناگون عدم موفقیت آزمونها در دسترسی به اهداف قصد شده، به مطالعات زیادی نیاز دارد که از حوصله یک پژوهش خارج است. از بین این عوامل این تحقیق به بررسی و تحلیل سطوح حیطة شناسی و شاخصهای روان سنجی سؤالات امتحانات نهایی دروس عربی، حسابان و زیست شناسی دانش‌آموزان دختر پایه سوم متوسطه در خرداد ۱۳۸۲ استان خوزستان پرداخته است.

اهمیت این موضوع از آنجا ناشی می‌گردد که متخصصان آزمون سازی معتقدند که نتایج

آزمونها ارزش بالقوه‌ای در تصمیم‌گیریهای مهم آموزشی از جمله ارتقای دانش‌آموزان به مدارج علمی بالاتر دارد (لرد و نوک^۲، ۱۹۶۸، تورندایک و هیگن^۳، ۱۹۸۲). بنابراین، چنانچه این آزمونها توسط معلمانی باتجربه ساخته شود تصمیماتی که بر اساس نتایج این آزمونها اتخاذ می‌گردد از صحت و اعتبار بالایی برخوردار خواهد بود. برعکس، چنانچه این آزمونها توسط معلمان کم تجربه، عدم آشنا با شیوه‌های ساخت و محاسبه شاخصهای روان سنجی سؤالات امتحان تهیه شود، روشن است که تصمیم‌گیری بر اساس نتایج چنین آزمونهایی از عدم صحت و اعتبار لازم برخوردار خواهد بود. بنابراین، به واسطه اهمیتی که امتحانات نهایی پایه سوم متوسطه در ارتقای دانش‌آموزان به دوره پیش دانشگاهی دارند و ضمناً یکی از اولویتهای پژوهشی سازمان آموزش و پرورش به شمار می‌رفت، ضرورت پیدا کرد تا سطوح حیطة شناسی سؤالات این آزمونها بررسی شود و تحلیلی روی شاخصهای روان‌سنجی سؤالات امتحانات نهایی دروس عربی، حسابان و زیست شناسی دانش‌آموزان پایه

2- Lord & Novick
3- Thorndike & Hagen

1- Tagomori

سوم متوسطه انجام پذیرد.

بیان دیگر، سؤالات، ضرایب دشواری بالا و ضرایب تمیز پایینی داشتند و به این معنا بود که دانشجویان برای پیدا کردن پاسخ درست سؤالات امتحانی با مشکل چندانی مواجه نبودند.

پیشینه پژوهش

مطالعات زیادی در زمینه محاسبه شاخصهای روان سنجی سؤالات آزمونهای معلم ساخته انجام نشده است. بیشتر تلاشها عمدتاً حول محور مطالعه ضرایب اعتبار و پایایی آزمونهای روانی صورت پذیرفته است. شاید دلیل این کار را محرمانه بودن این قبیل سؤالات درسی است که قبل از اجرا نمی‌توان دشواری، تمیز و میزان پایایی و ثبات آنها را محاسبه کرد. با محاسبه شاخصهای روان سنجی آزمونهای معلم ساخته، می‌توان از اصلاحات انجام شده برای اجراهای بعدی آزمون و یا در تهیه بانکی از سؤالات استفاده کرد.

مطالبات مشابهی نیز در این زمینه توسط قدیمی (۱۳۸۱) و پولادی (۱۳۸۰) انجام گرفته است. قدیمی میزان اطلاعات اعضای هیئت علمی گروههای آموزشی دانشکده‌ها را درباره ساخت سؤالات امتحانی دانشجویان بر اساس شاخصهای روان‌سنجی آنها بررسی کرد. وی گزارش داد که اعضای هیأت علمی گروه علوم تربیتی و روانشناسی در مقایسه با اعضای هیأت علمی سایر گروههای آموزشی اطلاعات وسیع‌تری در ساخت و طراحی سؤالات امتحانی دارند. یافته مطالعات قدیمی نشان داد که اعضای هیئت علمی گروه علوم پایه در مقایسه با سایرین، کمترین اطلاعات را در رابطه با ساخت سؤالات امتحانی داشتند. پولادی (۱۳۸۰) نیز نشان داد که بین اطلاعات معلمان دیپلم و کارشناس در ساخت آزمونهای معلم- ساخته تفاوت معنی‌داری وجود دارد. شریفی (۱۳۷۴)، سلیمی‌زاده (۱۳۷۷)، ضمن این که امتحانات را بخش جداناپذیر در فرایند تدریس می‌دانند، معتقدند که سؤالات امتحانی باید

سپاسی (۱۳۸۴) شاخصهای روان‌سنجی سؤالات امتحانی اعضای هیئت علمی دانشگاه شهید چمران را مورد بررسی قرار داد. نتایج تحقیق وی نشان داد که اعضای هیئت علمی با توجه به نوع مطلبی که تدریس می‌کنند از سؤالات امتحانی یکسانی استفاده نمی‌کنند. نتایج تحلیل تحقیق وی نشان داد که بین ضرایب دشواری، تمیز و پایایی سؤالات اساتید در مقایسه با ملاکهای مطلوب این ضرایب تفاوت وجود دارد. به

چنان تهیه شوند که نمونه‌ای کافی از عملکرد دانش‌آموزان را آزمون کند. آنها اظهار می‌دارند که سؤالات امتحانی باید به گونه‌ای ساخته شود که از نتایج آنها بتوان احتمال موفقیت فراگیران را با دقت پیش‌بینی کرد. ولی، در پایان می‌افزایند سؤالات امتحانی باید دارای دو ویژگی مهم آزمون‌سازی یعنی پایایی و اعتبار را دارا باشند.

شرقی (۱۳۸۳) در تحقیق مشابهی با پژوهش حاضر به بررسی و تحلیل سطوح حیطة شناختی و شاخصهای روان‌سنجی

سؤالات امتحانات نهایی دروس عربی، حسابان و زیست‌شناسی دانش‌آموزان پسر پایه سوم متوسطه سه منطقه متفاوت اقتصادی-اجتماعی استان خوزستان پرداخت. نتایج یافته‌های وی نشان داد که ضرایب دشواری و تمیز و همسانی درونی سؤالات امتحانات نهایی دروس مذکور در مقایسه با ملاکهای معتبر از وضعیت مناسبی برخوردارند. ولی، بین عملکرد تحصیلی دانش‌آموزان در سه وضعیت آموزشی برخوردار، نیمه برخوردار و غیربرخوردار استان تفاوت معنی‌داری وجود دارد.

برخی از تحقیقات به مطالعه این موضوع پرداخته‌اند که کلاسهای بازآموزی و حین خدمت چقدر توانسته‌اند معلمان را یاری دهند قادر باشند مهارتهای لازم در ساخت سؤالات آزمون کسب کنند. برخی دیگر از مطالعات مهارتهای معلمان دیپلم، کارشناس و بالاتر و سالهای تجربه تدریس آنان را در ساخت آزمونهای معلم- ساخته با ثبات و معتبر مورد بررسی قرار داده‌اند. یافته‌های فراری^۳ (۱۹۹۵)، دانیل و کینگ^۴ (۱۹۹۸) نشان داد عملکرد معلمان که در دوره‌های

علاوه بر مطالعات داخل کشور، پیشینه پژوهش در خارج از کشور نیز بررسی شد. نتایج یافته‌های اکثر این مطالعات نشان

- 1- Marso & Big
- 2- Tagomori & Bishop
- 3- Feraree
- 4- Daneil & King

روش شناسی پژوهش

چون هدف تحقیق بررسی شاخصهای روان سنجی سؤالات امتحان نهایی دروس مختلف در سطح استان خوزستان بود، ابتدا براساس شاخصهای اقتصادی- اجتماعی سازمان مدیریت و برنامه ریزی، استان خوزستان به سه منطقه برخوردار، نیمه برخوردار و غیربرخوردار تقسیم گردید. این تقسیم بندی این امکان را در اختیار محقق قرار داد تا مقایسات گسترده ای با توجه به مناطق مختلف آموزش و پرورش در استان خوزستان انجام دهد.

در این تحقیق شهرهای بهبهان، شوشتر، اندیمشک، دزفول، خرمشهر همراه با نواحی چهارگانه آموزش و پرورش اهواز منطقه برخوردار را تشکیل دادند. ضمناً شهرهای رامهرمز، امیدیه، ماهشهر، بندر امام، ایذه، باغملک، آبادان، سوسنگرد، مسجدسلیمان و شادگان جزء منطقه نیمه برخوردار قرار گرفتند. در نهایت، شهرهایی مثل لالی، گتوند، هندیجان، هفتگل، اروند، بستان، دهدز، سردشت، اندیکا، شاهرور، زیدون، رامشیر و آغاچاری جزء منطقه غیر برخوردار دسته بندی شدند.

کارآموزی شرکت کردند در مقایسه با آنانی که در این کلاسها شرکت نداشتند در ساخت سؤالات آزمونهای پایا و معتبر تفاوت معنی داری وجود دارد. تاگوموری و بی شاب (۱۹۹۴) و زانک و بری^۱ (۱۹۹۵) مطالعاتی در زمینه شیوه های ساخت سؤالات معلمان دوره های مختلف آموزش در ارتباط با مدرک تحصیلی، تجربه و سابقه تدریس آنان انجام دادند. نتایج یافته های آنان نشان داد که دبیران با مدرک کارشناسی و کارشناسی ارشد در ساخت آزمونهای معلم- ساخته از معلمان با مدرک دیپلم موفق تر عمل کرده اند.

سؤالات پژوهش

آیا در ساخت سؤالات امتحانی طراحان به اندازه گیری سطوح مختلف حیطه شناختی بر اساس جدول مشخصات آزمون توجه کرده اند؟

تا چه اندازه ساخت سؤالات امتحان نهایی بر پایه شاخصهای روان سنجی سؤالات استوار است؟

چگونه می توان دقت، ثبات و درستی سؤالات امتحانات نهایی را ارزیابی کرد؟

جامعه آماری

در این تحقیق سه جامعه آماری مختلف به شرح زیر مورد استفاده قرار گرفت: کلیه اوراق امتحان نهایی نمره‌گذاری شده درس عربی مربوط به دانش‌آموزان دختر پایه سوم متوسطه رشته ادبیات و علوم انسانی استان خوزستان در خرداد ۱۳۸۲.

کلیه اوراق امتحان نهایی نمره‌گذاری شده درس حسابان مربوط به دانش‌آموزان دختر پایه سوم متوسطه ریاضی- فیزیک استان خوزستان در خرداد ۱۳۸۲.

کلیه اوراق امتحان نهایی نمره‌گذاری شده درس زیست‌شناسی مربوط به دانش‌آموزان دختر پایه سوم متوسطه رشته علوم تجربی استان خوزستان در خرداد ۱۳۸۲.

روش نمونه‌گیری

با استفاده از روش نمونه‌گیری تصادفی مرحله‌ای، از هر جامعه آماری تعداد ۱۰۰۰ برگ اوراق امتحانی انتخاب گردید. سپس با توجه به تقسیم‌بندی استان به مناطق برخوردار، نیمه برخوردار و غیربرخوردار، نسبت به تعداد دانش‌آموزان دختر شرکت کننده در امتحانات هر درس در کل استان، بر اساس شماره‌های داوطلبی دانش‌آموزان که روی اوراق امتحانات نهایی نوشته شده بود، به طور تصادفی ساده اوراق مورد نظر

انتخاب گردید.

جدول ۱ تعداد اوراق نمونه مربوط به دانش‌آموزان را در دروس عربی، حسابان و زیست‌شناسی به تفکیک مناطق مختلف آموزش و پرورش استان نشان می‌دهد.

منابع داده‌های پژوهش

از سه آزمون عربی، حسابان و زیست‌شناسی به عنوان ابزارهای این تحقیق برای اندازه‌گیری عملکرد تحصیلی دانش‌آموزان دختر پایه سوم متوسطه استان در امتحانات نهایی دروس مربوطه استفاده شد. تمام سؤالات امتحانات نهایی توسط افرادی ذی صلاح و با نظم و رعایت نکات ایمنی طراحی، تکثیر و زیر نظر دایره امتحانات سازمان اجرا گردید. طراحان سؤالات علاوه بر تهیه و ساخت سؤالات، به منظور افزایش عینیت و سهولت در تصحیح اوراق امتحانی، ارزش هر سؤال (بارم) را با توجه به اهمیت نسبی آن نیز تعیین کرده بودند.

روش اجرای پژوهش

پس از برگزاری امتحانات نهایی دروس عربی، حسابان و زیست‌شناسی، اوراق امتحانی دروس مذکور توسط افرادی ذی صلاح نمره‌گذاری و نتایج برای اطلاع

جدول ۱ اوراق امتحانی نمونه شهرهای مناطق استان در دروس عربی، حسابان و زیست شناسی

جمع کل	غیربرخوردار				نیمه برخوردار				برخوردار			منطقه درس	
	جمع	هندیجان	آغاچاری	رامشیر	جمع	رامهرمز	مسجدسلیمان	آبادان	جمع	بهبهان	ناحیه ۴		ناحیه ۱
۱۰۰۰	۱۰۱	۵۸	۱۴	۲۹	۳۹۲	۱۲۲	۱۴۳	۱۲۷	۵۰۷	۱۵۹	۲۰۲	۱۲۶	عربی
۱۰۰۰	۴۰	۱۲	۱۵	۱۳	۲۶۴	۶۰	۸۴	۱۲۰	۶۹۶	۱۸۶	۳۵۷	۱۵۳	حسابان
۱۰۰۰	۸۷	۱۸	۲۱	۴۸	۳۸۵	۱۰۷	۱۵۲	۱۲۶	۵۲۸	۱۳۳	۲۰۰	۱۹۵	زیست شناسی
۳۰۰۰	۲۲۸	۸۸	۵۰	۹۰	۱۰۴۱	۲۸۹	۳۷۹	۳۷۳	۱۷۳۱	۴۷۸	۷۵۹	۴۷۴	کل

استان کسب کرده بودند در فرم مخصوص یادداشت می‌کردند. پس از گردآوری داده‌ها از برنامه SPSS برای تحلیل آنها استفاده شد. برای تحلیل داده‌هایی که از طریق برنامه SPSS مقدور نبود، برنامه تازه‌ای به کمک افراد ذی صلاح برای کامپیوتر نوشته شد.

یافته‌های پژوهش

مقادیر مربوط به شاخصهای مرکزی و پراکندگی نمره‌های امتحانات نهایی دروس عربی، حسابان و زیست شناسی پایه سوم متوسطه رشته‌های مختلف تحصیلی در خرداد ۱۳۸۲ استان خوزستان در جدول ۲ ارائه شده است.

دانش‌آموزان اعلام شده بود. اوراق امتحانی زیر نظر دایره امتحانات پلمپ و برای نگهداری در اختیار سازمان آموزش و پرورش استان قرار گرفته بود. با کسب مجوز از اداره کل سنجش و ارزشیابی تحصیلی وزارت آموزش و پرورش برای انتخاب تصادفی نمونه‌های مورد نظر از جامعه‌های آماری مربوطه، بسته‌های اوراق امتحانی فک پلمپ گردید.

برای دقت در استخراج نمره اوراق امتحانی دروس مختلف، منشیهای آموزش دیده برای این کار گمارده شدند. منشیها ابتدا اوراق امتحانی دروس مختلف را به طور انفرادی مشاهده و سپس نمره‌ای که دانش‌آموزان در پاسخ درست به سؤالات آن

جدول ۲ شاخصهای مرکزی و پراکندگی امتحانات نهایی دروس عربی، حسابان و زیست شناسی

زیست شناسی			حسابان			عربی			ناحیه شاخص
SD	Mdn	\bar{X}	SD	Mdn	\bar{X}	SD	Mdn	\bar{X}	
۴/۵۱	۱۲/۷۵	۱۲/۱۰	۵/۰۷	۱۲/۵	۱۲/۲۴	۳/۱۵	۸/۲۵	۸/۷۳	برخوردار
۴/۲۳	۱۲/۰۰	۱۱/۸۱	۴/۳۰	۱۰	۹/۸۴	۳/۴۴	۷/۲۵	۷/۹۷	نیمه برخوردار
۴/۰۳	۱۳/۵۰	۱۲/۸۶	۴/۸۳	۱۲	۱۲/۱۹	۲/۹۱	۱۱/۷۵	۱۱/۷۹	غیربرخوردار
۴/۳۵	۱۲/۲۵	۱۱/۹۷	۴/۹۳	۱۲/۲۵	۱۱/۴۰	۳/۴	۸/۲۵	۸/۸۲	کل استان

این دانش‌آموزان در مقایسه با دانش‌آموزان سایر مناطق بهتر بود. دانش‌آموزان منطقه غیربرخوردار در درس دشوار حسابان عملکردی مشابه با دانش‌آموزان منطقه

بررسی جدول فوق نشان می‌دهد علیرغم سکونت در منطقه‌ای که به لحاظ وضعیت آموزشی غیربرخوردار بود، میانگین نمره‌های امتحانات نهایی دروس عربی و زیست‌شناسی

برخوردار داشتند. دانش‌آموزان هر دوی این مناطق (برخوردار و غیربرخوردار) عملکرد بهتری در درس حسابان نسبت به دانش‌آموزان منطقه نیمه برخوردار از خود نشان دادند. میانگین امتحان نهایی درس عربی، حسابان و زیست شناسی در کل استان برای دانش‌آموزان پایه سوم رشته‌های تحصیلی مختلف به ترتیب ۸/۸۲ و ۱۱/۴۰ و ۱۳۸۲ در نظر گرفته بود نشان می‌دهد. ۱۱/۹۷ گزارش گردید.

جدول ۳ درصد حیطه شناختی امتحان نهایی درس عربی، حسابان و زیست شناسی

ارزشیابی	ترکیب	تحلیل	کار بستن	فهمیدن	دانش	سطوح درس
-	-	۲۲/۵	۲۵	۴۵	۷/۵	عربی
-	-	۲۶/۲۵	۳۳/۷۵	۴۰	-	حسابان
-	-	۲۳/۷۵	۵	۱۵	۵۶/۲۵	زیست شناسی

بررسی جدول ۳ نشان می‌دهد طراحان به ساخت سؤالاتی پرداخته بودند که صرفاً به اندازه‌گیری قلمروهای دانش، فهم، کاربرد و تحلیل مطالب درسی اختصاص داشتند. طراحان امتحان نهایی درس حسابان، شاید به دلیل اهمیت مطالب این درس، سؤالاتی برای اندازه‌گیری سطح دانش در نظر نگرفته بودند.

ضریب دشواری و تمیز سؤالات امتحانات نهایی

در این بخش ضریب دشواری و تمیز

جالب توجه است که هیچ یک از طراحان درس مختلف سؤالاتی برای اندازه‌گیری در قلمروهای ترکیب و ارزشیابی حیطه شناختی

برای سؤالات امتحانات نهایی دروس عربی، حسابان و زیست شناسی ابتدا برای هر درس و سپس برای همه مناطق و کل استان بررسی شده است. دانش‌آموزان مناطق و استان در جدول ۵ مشاهده کرد.

ضریب دشواری و تمیز سؤالات درس حسابان

جدول ۶ حداقل (L)، حداکثر (H) و میانگین (\bar{X}) ضرایب دشواری و تمیز سؤالات امتحان نهایی درس حسابان دانش‌آموزان مناطق و کل استان را نشان می‌دهد.

از ۱۸ سؤال امتحان نهایی درس حسابان، تعداد اندکی از آنها برای دانش‌آموزان مناطق و کل استان بسیار آسان و یا بسیار دشوار بودند و یا این که نتوانسته بودند بین تواناییهای پاسخ درست دانش‌آموزان قوی و ضعیف تمیز بگذارند. جدول ۷

ضریب دشواری و تمیز سؤالات درس عربی

جدول ۴ میانگین ضرایب دشواری و تمیز سؤالات امتحان نهایی درس عربی دانش‌آموزان مناطق مختلف و کل استان را نشان می‌دهد.

از ۲۱ سؤال امتحان نهایی درس عربی، تعداد اندکی از سؤالات بودند که برای دانش‌آموزان مناطق و کل استان بسیار آسان، بسیار دشوار و یا این که نتوانسته بودند بین دانش‌آموزان با توانایی‌های مختلف تفاوت بگذارند. محاسبات ضرایب نشان داد که اکثر سؤالات در چارچوب ملاکهای قابل قبول پیشنهادی متخصصان آزمون سازی قرار

جدول ۴ میانگین ضرایب دشواری و تمیز سؤالات امتحان نهایی درس عربی پایه سوم متوسطه

منطقه	برخوردار	نیمه برخوردار	غیربرخوردار	کل استان
ضریب	\bar{X}	\bar{X}	\bar{X}	\bar{X}
دشواری	۰/۴۴	۰/۴۲	۰/۵۹	۰/۴۷
تمیز	۰/۲۳	۰/۲۷	۰/۲۳	۰/۴۹

جدول ۵ محاسبه ضرایب دشواری و تمیز سؤالات امتحان نهائی درس عربی کل استان

ردیف ۲	α	X^2	D	pq	q	p	P_L	p_u	ردیف ۱
۸	۰/۰۰۰	۱۷۵/۶۸	۰/۵۷	۰/۲۵	۰/۵۰	۰/۵۰	۰/۲۱	۰/۷۸	۱
۱۸	۰/۰۰۰	۱۶۳/۴۷	۰/۵۲	۰/۲۳	۰/۶۶	۰/۳۴	۰/۰۸	۰/۶۰	۲
۹	۰/۰۰۰	۱۵۴/۸۲	۰/۵۳	۰/۲۵	۰/۵۴	۰/۴۶	۰/۱۹	۰/۷۲	۳
۷	۰/۰۰۰	۱۶۴/۲۸	۰/۵۰	۰/۲۵	۰/۴۴	۰/۵۶	۰/۲۸	۰/۸۳	۴
۳	۰/۰۰۰	۷۲/۷۰	۰/۳۲	۰/۲۰	۰/۲۷	۰/۷۳	۰/۵۷	۰/۸۹	۵
۵	۰/۰۰۰	۸۸/۷۶	۰/۳۹	۰/۲۴	۰/۳۹	۰/۶۱	۰/۴۱	۰/۸۰	۶
۱	۰/۰۰۰	۴۶/۲۵	۰/۲۱	۰/۱۳	۰/۱۵	۰/۸۵	۰/۷۴	۰/۹۵	۷
۲	نیست	۱۳/۴۲	۰/۱۲	۰/۱۴	۰/۱۷	۰/۸۳	۰/۷۷	۰/۸۹	۸
۱۵	۰/۰۰۰	۸۵/۲۵	۰/۴۷	۰/۲۳	۰/۶۴	۰/۳۶	۰/۱۷	۰/۵۴	۹
۴	۰/۰۰۰	۵۳/۶۳	۰/۳۰	۰/۲۲	۰/۳۳	۰/۶۷	۰/۵۲	۰/۸۲	۱۰
۱۳	۰/۰۰۰	۱۱۶/۸۲	۰/۴۵	۰/۲۴	۰/۶۲	۰/۳۸	۰/۱۶	۰/۶۱	۱۱
۱۹	۰/۰۰۰	۹۴/۵۷	۰/۳۹	۰/۲۱	۰/۷۱	۰/۲۹	۰/۱۰	۰/۴۹	۱۲
۲۱	۰/۰۰۰	۵۳/۹۵	۰/۲۳	۰/۱۲	۰/۸۶	۰/۱۴	۰/۰۳	۰/۲۶	۱۳
۱۰	۰/۰۰۰	۱۶۵/۸۷	۰/۵۵	۰/۲۵	۰/۵۵	۰/۴۵	۰/۱۸	۰/۷۳	۱۴
۱۷	۰/۰۰۰	۹۶/۷۱	۰/۴۰	۰/۲۳	۰/۶۵	۰/۳۵	۰/۱۵	۰/۵۵	۱۵
۶	۰/۰۰۰	۷۳/۴۵	۰/۳۶	۰/۲۴	۰/۳۹	۰/۶۱	۰/۴۳	۰/۷۹	۱۶
۲۰	۰/۰۰۰	۵۳/۹۵	۰/۲۳	۰/۱۲	۰/۸۶	۰/۱۴	۰/۰۳	۰/۲۶	۱۷
۱۲	۰/۰۰۰	۱۳۹/۰۹	۰/۵۰	۰/۲۴	۰/۶۲	۰/۳۸	۰/۱۳	۰/۶۳	۱۸
۱۱	۰/۰۰۰	۱۹۳/۱۲	۰/۵۹	۰/۲۴	۰/۵۹	۰/۴۱	۰/۱۲	۰/۷۱	۱۹
۱۴	۰/۰۰۰	۹۶/۰۹	۰/۴۰	۰/۲۳	۰/۶۳	۰/۳۷	۰/۱۷	۰/۵۷	۲۰
۱۶	۰/۰۰۰	۱۲۳/۱۵	۰/۴۶	۰/۲۳	۰/۶۵	۰/۳۵	۰/۱۲	۰/۵۸	۲۱

$$\bar{D} = ۰/۴۹$$

$$\bar{P} = ۰/۴۷$$

جدول ۶ حداقل، میانگین و حداکثر دشواری و تمیز سؤالات امتحان نهایی درس حسابان

منطقه	برخوردار			نیمه برخوردار			غیر برخوردار			کل استان
	H	\bar{X}	L	H	\bar{X}	L	H	\bar{X}	L	
دشواری	۰/۷۷	۰/۶۰	۰/۴۰	۰/۶۸	۰/۵۰	۰/۲۲	۰/۸۶	۰/۶۰	۰/۳۶	۰/۷۳
تمیز	۰/۶۲	۰/۴۰	۰/۲۰	۰/۷۰	۰/۳۶	۰/۱۸	۰/۵۷	۰/۳۸	۰/۱۴	۰/۸۲

جدول ۷ محاسبه ضرایب دشواری و تمیز سؤالات امتحان نهایی درس حسابان کل استان

ردیف ۱	p_u	P_L	p	q	pq	D	X^2	α	ردیف ۲
۱	۰/۸۸	۰/۳۰	۰/۵۹	۰/۴۱	۰/۲۴	۰/۵۸	۱۸۶/۶۷	۰/۰۰۰	۱۰
۲	۰/۹۱	۰/۳۷	۰/۶۴	۰/۳۶	۰/۲۳	۰/۵۴	۱۷۳/۴۵	۰/۰۰۰	۳
۳	۰/۹۶	۰/۳۸	۰/۶۷	۰/۳۳	۰/۲۲	۰/۵۸	۲۰۷/۱۶	۰/۰۰۰	۲
۴	۰/۸۹	۰/۳۸	۰/۶۴	۰/۳۶	۰/۲۳	۰/۵۱	۱۴۹/۹۹	۰/۰۰۰	۴
۵	۰/۸۸	۰/۳۲	۰/۶۰	۰/۴۰	۰/۲۴	۰/۵۶	۱۷۶/۲۱	۰/۰۰۰	۹
۶	۰/۸۰	۰/۲۷	۰/۵۴	۰/۴۶	۰/۲۵	۰/۵۳	۱۵۰/۱۹	۰/۰۰۰	۱۱
۷	۰/۸۲	۰/۳۲	۰/۶۲	۰/۳۸	۰/۲۴	۰/۶۰	۲۰۸/۹۲	۰/۰۰۰	۷
۸	۰/۷۹	۰/۲۷	۰/۵۳	۰/۴۷	۰/۲۵	۰/۵۳	۱۴۹/۸۹	۰/۰۰۰	۱۳
۹	۰/۸۹	۰/۱۶	۰/۵۳	۰/۴۷	۰/۲۵	۰/۷۳	۲۸۸/۳۶	۰/۰۰۰	۱۲
۱۰	۰/۹۲	۰/۳۳	۰/۶۳	۰/۳۷	۰/۲۳	۰/۵۹	۲۰۰/۳۵	۰/۰۰۰	۵
۱۱	۰/۸۸	۰/۲۲	۰/۵۵	۰/۴۵	۰/۲۵	۰/۶۶	۲۳۷/۲۵	۰/۰۰۰	۱۴
۱۲	۰/۷۲	۰/۰۵	۰/۳۹	۰/۶۱	۰/۲۴	۰/۶۷	۲۵۵/۷۳	۰/۰۰۰	۱۸
۱۳	۰/۹۱	۰/۰۹	۰/۵۰	۰/۵۰	۰/۲۵	۰/۸۲	۳۶۱/۷۹	۰/۰۰۰	۱۵
۱۴	۰/۹۱	۰/۳۳	۰/۶۲	۰/۳۸	۰/۲۴	۰/۵۹	۱۹۵/۹۳	۰/۰۰۰	۶
۱۵	۰/۸۰	۰/۱۳	۰/۴۶	۰/۵۴	۰/۲۵	۰/۶۶	۲۳۸/۵۲	۰/۰۰۰	۱۶
۱۶	۰/۹۰	۰/۵۶	۰/۷۳	۰/۲۷	۰/۲۰	۰/۳۴	۸۰/۸۴	۰/۰۰۰	۱
۱۷	۰/۷۸	۰/۰۵	۰/۴۱	۰/۵۹	۰/۲۴	۰/۷۳	۲۹۹/۰۸	۰/۰۰۰	۱۷
۱۸	۰/۹۰	۰/۳۰	۰/۶۰	۰/۴۰	۰/۲۴	۰/۶۰	۲۰۴/۶۹	۰/۰۰۰	۸

$$\bar{D} = ۰/۶۰$$

$$\bar{P} = ۰/۵۷$$

ضرایب دشواری و تمیز هر سؤال امتحان نشان می‌دهد. نهایی درس حسابان را برای کل دانش‌آموزان استان نشان می‌دهد. علاوه بر این، می‌توان معنی‌دار بودن ضرایب تمیز سؤالات را به طور انفرادی در سطح احتمال مربوطه برای افتراق بین دانش‌آموزان قوی و ضعیف در جدول ۷ مشاهده کرد. ستون آخر این جدول (ردیف دو) ترتیب دشواری سؤالات را با توجه به ضرایب دشواری محاسبه شده آنها نشان می‌دهد.

جدول ۸ حداقل، میانگین و حداکثر دشواری و تمیز سؤالات امتحان نهایی درس زیست شناسی

منطقه	برخوردار			نیمه برخوردار			غیر برخوردار			کل استان		
	H	\bar{X}	L	H	\bar{X}	L	H	\bar{X}	L			
دشواری	۰/۸۷	۰/۵۸	۰/۳۲	۰/۸۴	۰/۵۷	۰/۲۴	۰/۹۱	۰/۶۲	۰/۳۱	۰/۸۴	۰/۵۷	۰/۳۱
تمیز	۰/۵۶	۰/۳۴	۰/۳۰	۰/۴۵	۰/۳۵	۰/۱۳	۰/۶۹	۰/۳۲	۰/۶	۰/۶۶	۰/۵۵	۰/۲۳

از ۲۳ سؤالی که برای امتحان نهایی درس زیست شناسی دانش‌آموزان پایه سوم متوسطه رشته علوم تجربی در نظر گرفته شده بود، تعداد اندکی از آنها برای دانش‌آموزان مناطق و دانش‌آموزان در کل استان بسیار آسان یا بسیار دشوار بودند. افزون بر این، سؤالات نیز از قدرت تمیز خوبی برخوردار بودند و اکثر آنها توانستند، دانش‌آموزان با تواناییهای مختلف را از همدیگر متمایز سازند. جدول ۹ ضرایب دشواری و تمیز هر سؤال امتحان نهایی درس زیست شناسی را برای کل دانش‌آموزان استان نشان می‌دهد. همچنین، در این جدول می‌توان ضرایب دشواری و معنی‌دار بودن ضرایب تمیز سؤالات را

جدول ۹ محاسبه ضرایب دشواری و تمیز سؤالات امتحان نهایی درس زیست شناسی کل استان

ردیف ۲	α	X^2	D	pq	q	p	P_L	P_u	ردیف ۱
۲	۰/۰۰۰	۱۱۸/۰۶	۰/۴۲	۰/۲۰	۰/۲۹	۰/۷۱	۰/۵۰	۰/۹۳	۱
۱۰	۰/۰۰۰	۱۶۰/۶۱	۰/۵۴	۰/۲۴	۰/۴۱	۰/۵۹	۰/۳۲	۰/۸۶	۲
۲۰	۰/۰۰۰	۱۹۶/۹۴	۰/۶۰	۰/۲۵	۰/۵۱	۰/۴۹	۰/۱۹	۰/۷۹	۳
۴	۰/۰۰۰	۱۷۶/۷۱	۰/۵۴	۰/۲۲	۰/۳۳	۰/۶۷	۰/۴۰	۰/۹۴	۴
۱۴	۰/۰۰۰	۱۸۸/۶۶	۰/۵۹	۰/۲۵	۰/۴۳	۰/۵۷	۰/۲۸	۰/۸۶	۵
۳	۰/۰۰۰	۸۶/۷۴	۰/۳۷	۰/۲۱	۰/۳۰	۰/۷۰	۰/۵۲	۰/۸۹	۶
۲۱	۰/۰۰۰	۲۰۸/۳۷	۰/۶۲	۰/۲۵	۰/۵۵	۰/۴۵	۰/۱۴	۰/۷۶	۷
۸	۰/۰۰۰	۱۹۶/۶۶	۰/۵۹	۰/۲۴	۰/۳۹	۰/۶۱	۰/۳۱	۰/۹۰	۸
۱۳	۰/۰۰۰	۲۲۰/۷۵	۰/۶۳	۰/۲۵	۰/۴۳	۰/۵۷	۰/۲۵	۰/۸۹	۹
۱۵	۰/۰۰۰	۲۳۵/۵۹	۰/۶۶	۰/۲۵	۰/۴۴	۰/۵۶	۰/۲۳	۰/۸۹	۱۰
۱	۰/۰۰۰	۵۱/۹۵	۰/۲۳	۰/۱۳	۰/۱۶	۰/۸۴	۰/۷۳	۰/۹۶	۱۱
۲۳	۰/۰۰۰	۱۱۵/۴۲	۰/۴۳	۰/۲۱	۰/۶۹	۰/۳۱	۰/۰۹	۰/۵۲	۱۲
۱۹	۰/۰۰۰	۲۰۹/۰۰	۰/۶۲	۰/۲۵	۰/۵۵	۰/۴۵	۰/۱۴	۰/۷۶	۱۳
۱۶	۰/۰۰۰	۱۸۴/۹۶	۰/۵۹	۰/۲۵	۰/۴۹	۰/۵۱	۰/۲۱	۰/۸۰	۱۴
۱۱	۰/۰۰۰	۱۱۳/۵۵	۰/۴۵	۰/۲۴	۰/۴۱	۰/۵۹	۰/۳۶	۰/۸۱	۱۵
۶	۰/۰۰۰	۱۶۴/۶۷	۰/۵۳	۰/۲۳	۰/۳۷	۰/۶۳	۰/۳۶	۰/۹۰	۱۶
۱۷	۰/۰۰۰	۱۵۰/۰۳	۰/۵۳	۰/۲۵	۰/۵۳	۰/۴۷	۰/۲۰	۰/۷۳	۱۷
۹	۰/۰۰۰	۱۶۰/۰۰	۰/۵۳	۰/۲۴	۰/۴۰	۰/۶۰	۰/۳۳	۰/۸۷	۱۸
۲۲	۰/۰۰۰	۱۸۹/۵۰	۰/۵۹	۰/۲۴	۰/۵۸	۰/۴۲	۰/۱۳	۰/۷۱	۱۹
۱۸	۰/۰۰۰	۱۶۳/۳۴	۰/۵۵	۰/۲۵	۰/۵۴	۰/۴۶	۰/۱۹	۰/۷۳	۲۰
۷	۰/۰۰۰	۱۶۹/۲۸	۰/۵۴	۰/۲۴	۰/۳۸	۰/۶۲	۰/۳۴	۰/۸۹	۲۱
۵	۰/۰۰۰	۱۶۳/۷۴	۰/۵۳	۰/۲۳	۰/۳۵	۰/۶۵	۰/۳۹	۰/۹۱	۲۲
۱۴	۰/۰۰۰	۱۷۶/۱۵	۰/۵۷	۰/۲۵	۰/۴۷	۰/۵۳	۰/۲۴	۰/۸۱	۲۳

$$\bar{D} = ۰/۵۳$$

$$\bar{P} = ۰/۵۶$$

به طور انفرادی که می‌توانند در سطح احتمال مربوطه بین دانش‌آموزان قوی و ضعیف تمیز بگذارند مشاهده کرد. علاوه بر این، ستون آخر این جدول ترتیب دشواری سؤالات را با توجه به ضرایب دشواری محاسبه شده آنها نشان می‌دهد.

ضرایب پایایی امتحانات نهایی عربی، حسابان و زیست شناسی

جدول ۱۰ مجموعه واریانس سؤالات ($\sum pq$)، ضریب همسانی درونی (r_{tt})، خطای معیار اندازه‌گیری (SE)، ضریب اندازه سنجی (MM)، و ضریب پایایی کودر- ریچاردسون (KR-20) امتحانات نهایی دروس عربی، حسابان و زیست شناسی را نشان می‌دهد.

همان طور که در جدول فوق مشاهده می‌شود تمام ضرایب پایایی در سطح احتمال

$P < 0.001$ معنی دارند. سؤالات امتحانی طوری ساخته شده بودند تا از عواملی که موجب بروز خطاهای اندازه‌گیری می‌شوند جلوگیری نمایند. به بیان دیگر، آزمونها از ثبات و دقت خوبی برخوردار بودند یعنی این که نمره‌های مشاهده شده دانش‌آموزان در دروس مختلف با احتساب خطای معیار اندازه‌گیری به نمره‌های واقعی آنها نزدیک بودند.

اعتبار امتحانات دروس عربی، حسابان و زیست شناسی

چون سؤالات این امتحانات قبلاً توسط طراحان بر اساس جدول مشخصات و با رعایت تناسب بین اهداف و محتوای دروس مختلف پایه سوم متوسطه بودجه بندی شده بود، چنین به نظر می‌رسد که امتحانات از اعتبار محتوایی مناسبی برخوردار بودند. با

جدول ۱۰ شاخصهای روان سنجی امتحان نهایی دروس عربی، حسابان و زیست شناسی

KR-20	MM	SE	r_{tt}	$\sum pq$	منطقه ضریب
۰/۶۲	۳/۸۸	۱/۸۹	۰/۶۹	۴/۴۷	عربی
۰/۸۲	۵/۴۰	۱/۸۴	۰/۸۲	۴/۲۸	حسابان
۰/۸۲	۵/۴۶	۲/۱۱	۰/۸۲	۵/۳۶	زیست شناسی

$P < 0.001$

وجود این، برای حصول از صحت و درستی بودجه بندی و رعایت تناسب بین اهداف و محتوای دروس مختلف، نظرات سه نفر از کارشناسان ذی صلاح در هر درس اخذ گردید. همه این کارشناسان به اتفاق اعتبار محتوایی آزمونهای مذکور را تأیید کردند.

بحث و نتیجه گیری

معلمان در تمام سطوح تحصیلی همیشه خود را در حال ساخت سؤالاتی برای امتحانی که پیش رو دارند می بینند و دانش آموزان نیز دائماً خود را برای شرکت در امتحانات آماده می کنند. یکی از نگرانیهای موسسه های آموزشی از جمله سازمانهای آموزش و پرورش (استانها) این است که آیا سؤالات امتحانی ساخته شده توسط معلمان با آئین نامه های مربوط که اهداف برگزاری چنین امتحاناتی را تعیین کرده است توافق دارد یا خیر. به بیان دیگر، آیا معلمان در تهیه سؤالات امتحانی با ویژگیهای آزمون سازی آشنایی دارند؟ بنابراین، همان طور که ملاحظه شد ویژگیهای روان سنجی سؤالات امتحانات نهایی دانش آموزان پایه سوم متوسط در سه درس عربی، حسابان و زیست شناسی بررسی و تحلیل شد. این تحقیق می خواست نقاط قوت و ضعف این سؤالات

را طی یک مطالعه سراسری دقیق و منسجم بررسی کند و بر اساس یافته ها، پیشنهاداتی در اختیار سازمان آموزش و پرورش استان خوزستان قرار دهد تا این که سازمان بتواند نقاط قوت آزمونها را تقویت و در جهت اصلاح نقاط ضعف ساخت سؤالات امتحانی چاره اندیشی کند. اجرای تحقیق حاضر در راستای رسیدن به این اهداف انجام گردید.

در این بخش ابتدا به تحلیل شکل ظاهری سؤالات پرداخته خواهد شد. سپس، قلمروهای حیطة شناختی که سؤالات این امتحانات عمدتاً برای اندازه گیری آنها ساخته شده بودند بررسی می شود. آنگاه، درباره سطح دشواری و قدرت تمیز سؤالات، پایایی و اعتبار آزمونها بحث خواهیم کرد. در پایان هم به تفسیر نتایج فرضیه های تحقیق و ارایه پیشنهادات می پردازیم.

شکل ظاهری سؤالات: سازندگان

آزمونها حداکثر استفاده را از کمترین صفحات برای ارائه سؤالات کرده بودند. علاوه بر طرح سؤالات، برگه امتحان حاوی اطلاعاتی بود که دانش آموزان برای پاسخ به سؤالات نیاز داشتند. طراحان از انواع سؤالات رایج در آزمونهای معلم- ساخته برای اندازه گیری قلمروهای حیطة شناختی استفاده کرده بودند.

در این بخش ابتدا به تحلیل شکل ظاهری سؤالات پرداخته خواهد شد. سپس، قلمروهای حیطة شناختی که سؤالات این امتحانات عمدتاً برای اندازه گیری آنها ساخته شده بودند بررسی می شود. آنگاه، درباره سطح دشواری و قدرت تمیز سؤالات، پایایی و اعتبار آزمونها بحث خواهیم کرد. در پایان هم به تفسیر نتایج فرضیه های تحقیق و ارایه پیشنهادات می پردازیم.

تعداد کم سؤالات و محدود کردن مجموع امتیازات در امتحان حداکثر به نمره بیست، طراحان سؤالات را مجبور ساخته بود تا سؤالات را طوری وزن گذاری کنند که در مجموع به نمره بیست ختم شود. محدود کردن مجموع بارمهای یک امتحان به نمره بیست چند اشکال اساسی می‌تواند به دنبال داشته باشد. اول این که طرح سؤالات بیشتری را برای اندازه‌گیری قلمروهای گسترده حیطه شناختی محدود می‌سازد. متخصصان آزمون سازی معتقدند که بین تعداد سؤالات آزمون و ثبات و دقت اندازه‌گیری آن رابطه مثبت وجود دارد. هر چه بر تعداد سؤالات امتحانی افزوده شود از خطای اندازه‌گیری آن کاسته و نمره مشاهده شده دانش‌آموز به نمره واقعی آن نزدیکتر می‌شود. علاوه بر این، کار تصحیح اوراق امتحانی را در اختصاص دادن بخش کوچکی از یک نمره (مثلاً ۰/۵ یا ۰/۲۵) نمره، به بخشی که دانش‌آموز به سؤال درست پاسخ داده است دشوار می‌سازد. بنابراین، سازنده آزمون می‌تواند سقف بارم سؤالات امتحان را افزایش دهد. این کار، وی را قادر می‌سازد تا هم بر تعداد سؤالات آزمون برای اندازه‌گیری سطوح گسترده حیطه شناختی درس مورد نظر بیفزاید و هم کار نمره گذاری پاسخهای درست و غلط را سهولت بیشتری به بخشد.

تحلیل حیطه شناختی سؤالات امتحانات: طراحان سؤالات برای هر یک از سطوح حیطه شناختی اهمیت ویژه‌ای برای اندازه‌گیری پیشرفت آموذنیها قائل شده بودند. سؤالات عمدتاً به اندازه‌گیری سطوح دانش، فهم، کاربرد و تجزیه و تحلیل مطالب دروس مختلف پرداخته بودند.

بررسی دشواری و قدرت تمیز: متخصصان آزمون سازی (مهرنزلوهمان ۱۹۸۴، ایبل و فریزی، ۱۹۹۱، تورندایک و هیگن، ۱۹۸۶) ضرایب دشواری بین ۰/۷۰ تا ۰/۳۰ را برای آزمونهای معلم- ساخته پیشنهاد می‌کنند. آنها به تجربه دریافتند سؤالاتی که دشواری بالاتر از ۰/۷۰ و پایین‌تر از ۰/۳۰ دارند برای دانش‌آموزان به ترتیب زیاد آسان و زیاد دشوارند.

اکثر ضرایب دشواری محاسبه شده برای سؤالات امتحانات عربی، حسابان و زیست شناسی در این دامنه از دشواری قرار داشتند. تعداد سؤالاتی که با این ملاک مطابقت نداشت اندک بودند. علاوه بر این، میانگین ضرایب دشواری برای دروس عربی، حسابان و زیست شناسی به ترتیب، ۰/۴۶، ۰/۵۷ و ۰/۵۶ بود که این ضرایب برای سؤالات

آزمونهای معلم- ساخته بسیار مطلوب ارزشیابی می‌شوند. متخصصان آزمون‌سازی معتقدند هر چه ضریب تمیز سؤالات آزمون بیشتر باشد، قدرت افتراق آنها بین دانش‌آموزان قوی و ضعیف بیشتر است. این متخصصان ضریب ۰/۲۰ را برای سؤالات آزمونهای معلم- ساخته پیشنهاد می‌کنند. در مقایسه با این ملاک، تحلیل و بررسی ضرایب تمیز سؤالات عربی، زیست‌شناسی در امتحانات نهایی این دروس در سطح استان نشان داد که اکثر سؤالات با این ملاک هم‌خوانی دارند. فقط تعداد محدودی از سؤالات ضرایب تمیزی داشتند که با این ملاک تفاوت داشت. به عبارت دیگر، بر اساس تحلیل ضرایب دشواری و تمیز سؤالات، نتایج تحقیق نشان داد که طراحان در ساخت سؤالاتی که با سطوح تواناییهای دانش‌آموزان هماهنگ باشند و بتوانند بین دانش‌آموزان قوی و ضعیف تفاوت بگذارند آگاهی کافی داشتند.

بررسی پایایی و اعتبار آزمونها: در این تحقیق ضرایب پایایی آزمونهای عربی، حسابان و زیست‌شناسی در امتحانات نهایی خرداد ۱۳۸۲ کل استان به ترتیب ۰/۶۲، ۰/۸۲ و ۰/۸۲ محاسبه شد. همه این ضرایب در سطح احتمال $P < 0/05$ معنی‌دار بودند.

ضرایب محاسبه شده نشان داد که آزمونها برای اندازه‌گیری پیشرفت دانش‌آموزان از ثبات و دقت کافی برخوردار بودند. به بیان دیگر، چنین به نظر می‌رسد نمره‌هایی که دانش‌آموزان در امتحانات نهایی دروس مختلف کسب کرده بودند، به واسطه کاهش خطای اندازه‌گیری، فاصله چندانی با نمره واقعی آنها نداشت. محاسبه ضریب پایایی برای اندازه‌گیری دقت و ثبات آزمون لازم است ولی کافی نیست. برای تکمیل ویژگیهای روان‌شناختی آزمونها، محاسبه و بررسی اعتبار آنها نیز لازم است. بنابراین، چنانکه آزمونها از اعتبار مطلوبی برخوردار باشند، ثبات و دقت هم خواهند داشت. عکس این مطلب صادق نیست.

طراحان دروس مختلف برای این که نشان دهند سؤالات در واقع هدفهای مورد نظر را می‌سنجد، از اعتبار محتوایی استفاده کردند. بررسی این اعتبار نشان داد که سؤالات براساس جدول مشخصات و با رعایت اهداف رفتاری، محتوای دروس، سطوح حیطة شناختی و با در نظر گرفتن اهمیت نسبی اهداف هر یک از آزمونهای دروس عربی، حسابان و زیست‌شناسی ساخته شده‌بودند. بنابراین، معلمان توانسته بودند با تهیه جدول مشخصات آزمونها

هماهنگی لازم بین مطالب را فراهم آوردند و آموزش و پرورش قرار دهد. علاوه بر این، بر اعتبار محتوایی امتحانات خود بیفزایند. پیشنهاد می‌شود بررسی ساخت آزمون‌ها سایر نتایج این تحقیق می‌تواند اطلاعات با دروس در دستور کار سازمان آموزش و ارزشی در ارتباط با ساخت آزمون‌های پرورش استان نیز قرار گیرد. امتحانات نهایی در اختیار مسئولین سازمان

منابع

فارسی

- پولادی، حسین (۱۳۸۰). *بررسی میزان اطلاعات تخصصی معلمان مقطع ابتدایی شهرستان قوچان در زمینه طرح سؤالات امتحانی و ارزیابی آموخته‌های دانش‌آموزان*. پایان نامه کارشناسی ارشد دانشگاه آزاد اسلامی واحد تهران. مرکز
- سپاسی، حسین (۱۳۸۴). *بررسی سؤالات امتحانی اعضای هیئت علمی دانشگاه شهید چمران از طریق بررسی شاخصهای روان‌سنجی سؤالات امتحانی موجود، مجله علوم تربیتی و روان‌شناسی (ویژه نامه علوم تربیتی)*، دوره سوم، سال دوازدهم، زمستان ۸۴، شماره ۴، ص ۱-۲۰.
- سلیمی‌زاده، محمدکاظم (۱۳۷۷). *آشنائی با روشهای تجزیه و تحلیل سؤالات و آزمونهای مورد استفاده در سازمان سنجش آموزش کشور*. انتشارات سازمان سنجش و آموزشی کشور ص ۸۰-۶۵.
- سیف، علی‌اکبر (۱۳۷۵). *روشهای اندازه‌گیری و ارزشیابی آموزشی*. تهران: نشر دوران.
- شرقی، حمید (۱۳۸۳). *بررسی و تحلیل سطوح حیطة‌شناختی و شاخصهای روان‌سنجی امتحان نهایی خرداد ماه ۱۳۸۲ دروس عربی، حسابان و زیست شناسی در پایه سوم*

متوسطه در سه منطقه متفاوت اقتصادی- اجتماعی استان خوزستان. شورای تحقیقات سازمان آموزش و پرورش استان خوزستان.

شریفی، حسن پاشا (۱۳۷۷). *آزمونهای باز پاسخ و روشهای عینی نمره‌گذاری آنها*. سازمان سنجش آموزشی کشور ص ۹۷-۱۰۷.

قدیمی مقدم، ملک محمد (۱۳۸۱). *بررسی میزان اطلاعات تخصصی اساتید دانشگاه آزاد اسلامی منطقه ۹*. پایان نامه کارشناسی ارشد دانشگاه آزاد اسلامی واحد تهران منطقه ۹.

ملکی، بهرام (۱۳۸۰). *بررسی میزان آشنائی و مهارت دبیران متوسطه استان کردستان با اصول سنجش و اندازه‌گیری در سال ۸۰-۷۹*. شورای تحقیقات سازمان آموزش و پرورش استان کردستان.

لاتین

- Baumgartner, T.A.J. & Jackson, A.S. (1991). *Measurement for evaluation in physical education and ducation and Exercise science* (4th ed.). Wis: Brown Publisher.
- Daneil, L. & King, D. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary school teachers. *Journal of Educational Research*, 91, 333-344.
- Ebel, L.E. & Frisbie, D.A. (1991). *Essentials of educational measurement*. Cliff, NJ: Engelwood, Printic Hall.
- Frary, R.B. (1993). Testing and grading practices and opinions of secondary school teachers of academic subjects. Implications for instruction opinions measurement. *Educational Measurement*. Issues and Practice. Vol 72 no. 3.
- Gage, L.R. (1992). *Educational evaluation and measurement*. (second ed.) Maxwell, Macmillan, International.
- Lord, F.M., & Novick, M.R. (1968). *Statistical theory of mental test scores*. Reading, Mass; Addison-Wesly.
- Marso, R.N. & Pigge, F.L. (1991). An analysis of teacher tests. Items, types, cognitive domains. *Journal of Contemporary Educational Psychology*, 16: 201-209.

-
- Mehrens, W.A. & Lehman, I.J. (1984). *Measurement and evaluation*. (third ed.). New York: Holt, Rinehart and Winston.
- Page, B.E. & Peterson, N.S. (1995). *The computer moves into essay grading*, Delta Kappa, 79, 7-15.
- Tagomori, H.T. & Bishop, L.A. (1994). *Content analysis of evaluation instruments used for student evaluation of classroom teaching for performance higher education*. Paper presented in annual meeting of the American Educational Research Association.
- Thorndike, R.L. & Hagen, E. (1986). *Examiner's manual-cognitive ability tests*. Form 4. Chicago: Riverside Publishing Company.
- Zhang, L. & Burry, S.J. (1995). *A multivariate analysis of teachers perceived assessment competency as a function of measurement training and years of teaching*. Paper presented at the annual meeting of the Mid-South Educational Research Association.