

تاریخ دریافت مقاله: ۸۳/۴/۸

بررسی مقاله: ۸۳/۸/۱۸

پذیرش مقاله: ۸۳/۶/۴

مجله علوم تربیتی و روانشناسی

دانشگاه شهید چمران اهواز، ۱۳۸۴

دوره سوم، سال دوازدهم، شماره ۲

صص: ۱-۲۰

بررسی سؤالات امتحانی اعضای هیأت علمی دانشگاه شهید چمران از طریق بررسی شاخصهای روان سنجی سؤالات امتحانی موجود

دکتر حسین سپاسی*

دکتر یوسفعلی عطاری*

چکیده

هدف از تحقیق حاضر در درجه نخست بررسی شاخصهای روان سنجی سؤالات امتحانی اعضای هیأت علمی دانشگاه شهید چمران بود. همچنین، بررسی استفاده از انواع سؤالات امتحانی و مقایسه شیوه‌های ارزشیابی اساتید دانشکده‌های مختلف هدف دیگر این تحقیق را تشکیل داد. ۱۰۹ نفر از اعضاء هیأت علمی دانشگاه داوطلبانه در این مطالعه شرکت و اوراق امتحانی پایان ترم یکی از دروس خود را جهت بررسی شاخصهای روان سنجی آن ارسال داشتند. سؤالات آزمون اساتید به طور انفرادی بررسی و علاوه بر محاسبه ضرایب دشواری، تمیز و پایایی ویژگیهای انفرادی سؤالات نیز تحلیل گردید. نتایج این ضرایب و تحلیلها محرمانه و مکتوب همراه با ارائه پیشنهاداتی به اطلاع اساتید داوطلب در تحقیق رسانده شد. بررسی نتایج تحقیق نشان داد اساتید با توجه به نوع درسی که ارائه می‌دهند، از شیوه‌های امتحانی یکسانی استفاده نمی‌کنند. بین ضرایب دشواری، تمیز و پایایی سؤالات در مقایسه با ملاکهای مطلوب این ضرایب تفاوت مشاهده گردید. در کل تحلیل اوراق امتحانی اساتید نشان داد که سؤالات از ضرایب دشواری بالا و ضرایب تمیز پائین برخوردارند و به این معنی که دانشجویان در پاسخ دادن درست به سؤالات آنها مشکل چندانی نداشتند. بر اساس نتایج این تحقیق پیشنهاد می‌شود دانشگاه یک مرکز آزمون سازی دایر سازد تا از طریق آن اساتید بتوانند نیازهای ارزشیابی خود را در ساختن سؤالات امتحانی استاندارد و محاسبه شاخصهای روان سنجی سؤالات برطرف نمایند.

کلید واژگان: شاخصهای روانسنجی، دشواری سؤال، تمیز سؤال، پایایی

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

بیان مسأله

از قدیم ایام گرفتن نوعی آزمون از شهروندان برای احراز مناصب دولتی مرسوم بوده است. شاید مراسمی که در آن قبایل بدوی عادات و رسوم قبیله‌ای و تحمل و آمادگی جوانان را برای اجازه ورود به مراتب بسزرگسالی آزمون می‌کردند، در زمره قدیمی‌ترین آزمونهایی باشد که بوسیله انسان به کار گرفته شده است. مقامات چینی در ۲۲۰۰ سال قبل از میلاد از آزمونهای کتبی برای گزینش افراد لایق برای احراز تصدی امور مملکتی استفاده می‌کردند (ابیل^۱، ۱۹۹۱). در ایران باستان در عهد شاپور پادشاه ساسانی در دانشگاه جندی شاپور هم جلساتی برای آزمون دانشجویان پزشکی بر پا می‌کردند و اعطای گواهی نامه پزشکی موقوف به موفقیت دانشجویان در این امتحانات بود (کیامنش، ۱۳۷۰). همچنین از مطالعه آثار و نوشته‌های بزرگان کشورمان نیز چنین می‌توان استنباط کرد که آنها هم برای ارزشیابی اندوخته‌ها علمی شاگردان و طلاب حوزه درسی خود از آزمونها به صورت شفاهی در محضر استاد و یا کتبی استفاده می‌کردند (عیسی صدیق، ۱۳۳۶).

به بیان دیگر، در آن ایام آزمونهای شفاهی یا کتبی در قالب سؤالات تشریحی و در حضور استاد از رونق زیادی برخوردار بودند.

شاید دلیل عمده استفاده از این نوع آزمونها را بتوان به تعداد کم شاگردان که در کلاس درس حضور پیدا می‌کردند نسبت داد. حضور قلبی در کلاس درس به استاد اجازه می‌داد تا در وقت مناسب مفاهیم عمده درس را به صورت شفاهی یا کتبی از دانشجویان سؤال کند. هم اکنون نیز استفاده از این نوع آزمونها در بین برخی از معلمان، دبیران و اساتید دانشگاه رایج است.

امروزه بیش از ۲۰ میلیون کودک، نوجوان، جوان در مقاطع مختلف به تحصیل اشتغال دارند که بیش از دو میلیون آنها در دانشگاهها و مؤسسات آموزش عالی کشور تحصیل می‌کنند. برای تصمیم‌گیری در مورد میزان پیشرفت و ارتقای آنان به مقاطع بالاتر، معلمان، دبیران و استادان دانشگاه به طور مرسوم از انواع مختلف آزمونها استفاده می‌کنند. به بیان دیگر، تنها وسیله‌ای که می‌توان اطمینان حاصل کرد که این فراگیران به یادگیری مطلوب دست یافته‌اند و ارزشیابی آنها از عینیت، اعتبار و ثبات لازم برخوردار می‌باشد، استفاده از آزمونهایی است که توسط معلمان، دبیران و اعضای هیأت علمی دانشگاهها و مؤسسات آموزش عالی

ساخته می شود.

ایفا می کند.

در نتیجه، در نظام آموزشی کنونی دگرگونی‌هایی در شیوه‌های ارزشیابی از آموخته‌های فراگیران صورت پذیرفته و به تدریج آزمونهای سنتی جای خود را به شیوه‌های جدیدتر اندازه‌گیری و ارزشیابی داده است. برخی از این شیوه‌های ارزشیابی به دلیل افزایش شمار دانشجویان و تعداد قلیل مدرسان کسارایی خود را از دست داده‌اند و شیوه‌های جدیدتری جایگزین آنها شده است. ارزشیابی از عملکرد تحصیلی دانشجویان یکی از اهداف عمده دانشگاهها و مؤسسات آموزش عالی کشور است که گاهی بحث‌انگیز و مسأله‌ساز نیز می شود. برای ارزشیابی می توان دو نقش عمده متصور شد: یکی تصمیم‌گیری در مورد این که آیا دانشجویان شرایط لازم را برای ارتقاء کسب کرده‌اند یا نه. نقش دیگر ارزشیابی به یادگیری در طول دوره تحصیلی مربوط می شود و هدف آن ایجاد سهولت، انگیزه، جهت دادن، تعیین نقاط قوت و ضعف و بالاخره هدایت یادگیری دانشجویان است (بروکر و اسمیت^۱، ۱۹۹۶). این دو نوع ارزشیابی اهداف متفاوتی را دنبال می کنند. ولی ارزشیابی که در پایان دوره تحصیلی از آموخته‌های دانشجویان انجام می شود به لحاظ دانشگاه نقش مهمتری

موقعی که ارزشیابی در جهت تأمین خواسته‌های دانشگاه به معنی گذراندن تعداد واحدهای درسی انجام پذیرد، فرایند ارزشیابی یک وظیفه آموزشی است که هدف آن تعیین صلاحیت دانشجوی برای ارتقاء و تفویض قدرت به استاد درس است. هدف چنین فرایندی ارزشیابی پایانی است که موجب افزایش کنترل و اعمال مقررات و قوانین آموزشی می شود.

اجرای آزمونها برای هدف دوم به ارزشیابیهای تکوینی (مرحله‌ای) و پایانی نیاز دارد که به اندازه‌گیری میزان تبهر و شایستگی دانشجو در طی دوره و پایان دوره در رسیدن به اهداف ویژه درس می پردازد (نایت^۲ و همکاران، ۱۹۹۶). این هدف، از طریق اجرای آزمونهایی که زیر نظر استاد صورت می گیرد قابل حصول است، تا بتوان از نتیجه تلاش و کوشش دانشجو در رسیدن به اهداف درس اطمینان حاصل کرد (براون و نایت^۳، ۱۹۹۴). اجرای این آزمونها وقت‌گیرند، باز خورد فراهم می کند و معمولاً عینیت دارند و متصفانه برگزار می شوند.

1- Brooker & Smith

2- Night

3- Brown & Knight

اولین مورد به این مسأله اشاره می‌کند که آزمونها به آنچه انتظار می‌رود که دانشجو پس از فراغت از تحصیل انجام دهد مربوط نمی‌شود. دومین مورد چنانچه سؤالات آزمون به اندازه‌گیری محفوظات بپردازد یادگیری سطحی را بین دانشجویان تشویق می‌کند. سومین موضوع به این مطلب اشاره دارد که چنانچه آزمونها تحت شرایط بسیار دشوار و اضطراب زیاد برگزار شود امکان دارد از توانایی دانشجو در پاسخ به سؤالات آزمون بکاهد و نتیجه مطلوب حاصل نگردد. توضیح چهارمین مورد این است که اغلب آزمونها اهداف مورد نظر و مطلوب واحد درسی را اندازه‌گیری نمی‌کنند. بنابراین، آزمونها در تعیین سرعت، سطح و نوع یادگیری و میزان تسلط دانشجویان با مشکل مواجه خواهد بود. این موضوع زمانی آشکار می‌شود که دانشجویان در مقابل پاسخ به سؤالات تشریحی قرار گیرند و از آنها خواسته می‌شود تا به تحلیل، و سازمان‌دهی افکار خود پرداخته و به صورت مستدل به سؤالات پاسخ دهند. در یک مطالعه، ۴۸ درصد از دانشجویان در پاسخ به سؤالات آزمون تشریحی با مشکل مواجه شده بودند (بالی و توهی^۱، ۱۹۹۷). در توضیح پنجمین دشواری باید به این نکته

ولی گاهی در اجرا و ساخت آزمونها مسائل و مشکلاتی پیدا می‌شود که به جای افزایش سطح و نوع یادگیری امکان دارد نتایج معکوسی بدنبال داشته باشد. بنابراین، لازم است جنبه‌های مثبت در ساخت و اجرای آزمونها را مدنظر داشت و از آن ویژگیها که امکان دارد به یادگیری مطلوب دانشجو لطمه بزند کاست. آیا امکان ساخت چنین آزمونهایی وجود دارد به گونه‌ای که استاد بتواند نتایج مطلوب یادگیری واحد درسی را ارزشیابی کند و در عین حال بسیاری از مشکلاتی که حاصل اجرای آزمونهای ضعیف است بکاهد؟

جنبه‌های مسأله‌ساز که از اهمیت یک آزمون می‌کاهد و باید از آنها اجتناب کرد عبارتند از:

۱. آزمونها شباهتی به فعالیتهای حرفه‌ای دانشجویان پس از فراغت تحصیل نداشته باشند.
۲. آزمونها فقط توانایی پاسخ به سؤالات را بسنجند و یادگیری سطحی و حفظی را ترغیب نمایند.
۳. نتایج آزمونها تحت فشار اضطراب آزمودنیها در جلسه امتحان لطمه ببیند.
۴. آزمونها بر اساس اهداف اصلی درس تنظیم نشده باشند.
۵. آزمونها امکان بررسی تفاوت بین سطوح مختلف یادگیری دانشجویان را اجازه ندهد.

ترکیب و ارزشیابی)، تعریف اهداف درس در قالب هدفهای رفتاری (دارا بودن ملاک عملکرد و شرایط عملکرد)، تهیه جدول مشخصات آزمون (تعیین ابعاد مختلف درس در قالب سطوح مختلف حیطه شناختی، تعیین اولویت اهداف درس)، آشنایی با انواع مختلف آزمونهای (تشریحی، شفاهی، کوتاه- پاسخ، صحیح/ غلط، چند گزینه‌ای) و آگاهی از این نوع آزمونها که برای اندازه‌گیری کدام یک از اهداف درس مناسبترند، لازم است. علاوه بر این، محاسبه ضرایب دشواری و تمیز، (ثبات و دقت و تمایز بین دانشجویان قوی و ضعیف)، هماهنگ کردن دشواری و قدرت تمیز سؤالات با تواناییهای دانشجویان، کاهش میزان دخالت خطاها (خطای ناشی از اجرای آزمون، خطای ناشی از تهیه سؤالات، خطای ناشی از خود آزمون)، که امکان دارد نتایج نمره پیشرفت دانشجویان را مخدوش کند از دیگر مهارتهای است که استاد باید نسبت به آنها آگاهی کامل داشته باشد. در نهایت، آگاهی از میزان اعتبار آزمون (اعتبار محتوایی) برای پاسخ به این که آیا سؤالات آزمون همان اهداف مورد نظر را اندازه می‌گیرد یا نه لازم است.

برای کاهش میزان خطا و افزایش اعتبار و

اشاره کرد که آزمونهایی که با اهداف محدودی ساخته می‌شوند قادر نیستند تفاوت بین دانشجویان قوی و ضعیف را اندازه‌گیری کنند. فقط دانشجویانی از این آزمونها سود می‌برند که به حفظ کردن مطالب جزئی درس می‌پردازند لذا پاداشی برای دانشجویان قوی فراهم نمی‌شود. این نوع آزمونها قادر نیستند تمایزی بین سطوح مختلف یادگیری دانشجویان فراهم کنند.

بنابراین، لازم است آن دسته از ویژگیهای آزمون که مشکل‌ساز هستند شناسایی و در جهت کاهش آنها قدمهای مؤثری برداشته شود. استاد باید با اطمینان و برنامه‌ریزی دقیق و اختصاص دادن بخش عمده‌ای از اوقات آموزشی خود به ساختن آزمونها بپردازد تا از بروز برخی از مشکلات که در بالا به آنها اشاره شد جلوگیری شود.

علاوه بر این، برای حصول اطمینان از این که آزمونها ویژگی و شایستگیهای لازم برای اندازه‌گیری پیشرفت مطلوب دانشجویان را دارند، استاد لازم است مهارتهایی در زمینه ساختن آزمونهای درسی مورد نیاز کسب کند. آگاهی از اهداف اجرای آزمون، (آزمون برای چه اهدافی برگزار می‌شود)، آشنایی با سطوح حیطه شناختی (دانش، فهم، کاربرد، تحلیل،

نهایی این بررسی را تشکیل داد.

روش تحقیق

شمار اساتیدی که در سال تحصیلی ۸۱-۸۲ در دانشگاه شهید چمران به عنوان عضو هیأت علمی به فعالیت آموزشی و پژوهشی اشتغال داشتند طبق آمار مدیریت برنامه‌ریزی و ارزشیابی بالغ بر ۴۹۸ نفر بودند. از این تعداد ۲۰ نفر در مرتبه استادی، ۱۶۶ نفر در مرتبه دانشیاری، ۲۴۱ نفر در مرتبه استاد یاری و بقیه ۷۱ نفر با مدرک کارشناسی ارشد در مرتبه مربی به تدریس اشتغال داشتند.

ابتدا طی نامه‌ای به اطلاع اعضای هیأت علمی رسید که قرار است شاخصهای روان‌سنجی سؤالات امتحانی پایان ترم آنها در قالب یک طرح پژوهشی توسط محقق مورد بررسی و تحلیل قرار گیرد. لذا از آنها درخواست شد که به صورت داوطلب و به انتخاب خود اوراق امتحانی یکی از دروس خود را جهت تحلیل این شاخصها در اختیار این محقق قرار دهند. تعداد ۱۰۹ نفر از ده دانشکسده در وقت مقرر داوطلبانه اوراق امتحانی خود را جهت تحلیل ضرایب

انطباق سؤالات آزمون با سطوح مختلف تواناییهایی که پیشرفت تحصیلی آزمودنیها را به طور عینی نشان دهد، متخصصان آزمون سازی (بلاوم^۱، گران لاند^۲، ۱۹۸۶؛ کرونباخ، ۱۹۸۶؛ مهرنر و لمان^۳، ۱۹۸۴؛ ایبل، ۱۹۹۱) محاسبه شاخصهایی را برای سؤالات آزمون پیشنهاد کرده‌اند. این متخصصان عقیده دارند چنانچه آزمونهای روانی و تربیتی فاقد این ویژگیها باشند به تصمیم‌گیری در مورد نتایج این آزمونها نمی‌توان اطمینان حاصل کرد. آزمونهایی که توسط معلمان، دبیران و اعضای هیأت علمی دانشگاه برای اندازه‌گیری پیشرفت فراگیران ساخته می‌شود از این قاعده مستثنا نیستند.

هدف عمده این تحقیق بررسی سؤالات امتحانی اعضای هیأت علمی دانشگاه شهید چمران اهواز از طریق ارزشیابی شاخصهای روان‌سنجی سؤالات امتحانی آنها بود. محاسبه تناسب دشواری سؤالات امتحانی با سطوح توانایی دانشجویان و همچنین بررسی میزان قدرت افتراق سؤالات بین دانشجویان قوی و ضعیف (که هدف هر آزمون پیشرفت محسوب می‌شود) هدف دوم این تحقیق بشمار می‌رود. بالاخره، بررسی دقت و ثباتی که سؤالات امتحانی در اندازه‌گیری میزان عملکرد تحصیلی دانشجویان داشته‌اند هدف

1- Bloom

2- Gronlund

3- Mehrens & Lehman

دشسواری، تمیز و پایائی و بررسی سایر ویژگیهای سؤالات ارسال داشتند. بنابراین، اعضای هیأت علمی که داوطلبانه علاقمند به بررسی سؤالات امتحانی پایان ترم خود بودند نمونه آماری تحقیق حاضر را تشکیل داد.

نمونه‌گیری تصادفی از یک جامعه وسیع زمانی مقدور است که موضوعی کلی با آنها در میان گذاشته شود. برای مثال، روش نگرش سنجی که نمونه‌ها به صورت تصادفی انتخاب می‌شوند در آن افراد به سؤالات خاصی پاسخ می‌دهند و زمان زیادی برای پاسخگویی مورد نیاز نیست. ولی اگر تحقیقی طرح‌ریزی شود که طی آن محقق بخواهد مثلاً شیوه تدریس یا روش نمره‌گذاری و طراحی سؤالات اعضای هیأت علمی مؤسسه‌ای را به طور تصادفی مورد بررسی قرار دهد، علیرغم انتخاب تصادفی، تعداد اندکی به دلایلی حاضر می‌شوند در تحقیق شرکت نمایند. بنابراین، موقعی که عده‌ای افراد نمونه از شرکت در تحقیق امتناع ورزند، افراد باقی مانده دیگر یک نمونه تصادفی به شمار نخواهد آمد. زیرا احتمالاً ویژگیهای آن دسته از افرادی که اعلام موافقت می‌کنند با ویژگیهای آن عده از افراد که از موافقت سرباز می‌زنند متفاوت خواهد بود. علاوه بر این، به لحاظ مراعات برخی مسائل اخلاقی باید موافقت شرکت افراد قبل از شروع

تحقیق نیز جلب گردد.

نتایج مطالعات بیگز، کولیس و کوین^۱ (۱۹۸۲) نشان می‌دهد که ۱- افراد داوطلب موقعی که در معرض سؤالات حساس قرار می‌گیرند از افراد غیر داوطلب اعتماد به نفس بالاتری در پاسخ به سؤالات از خود نشان می‌دهند، ۲- افراد داوطلب علاقه بیشتری برای تولید دانش و معلومات از خود نشان می‌دهند تا افراد غیر داوطلب، ۳- افراد داوطلب کمتر از افراد غیر داوطلب خودخواه هستند و بیشتر حاضرند برای پیشبرد اهداف تحقیق با محقق همکاری نمایند. بنابراین، به دلیل اهمیت موضوع، شاخصهای روان‌سنجی سؤالات امتحانی آن دسته از اعضای هیأت علمی دانشکده‌ها که داوطلبانه حاضر شدند و ویژگی سؤالات امتحانی آنها تحلیل شوند نمونه‌گیری مطلوب برای این تحقیق بشمار آمد.

روش تحلیل شاخصهای روان‌سنجی سؤالات امتحانی

نمره‌ای که هر دانشجو در هر درس دریافت می‌کند از مجموع نمره سؤالات امتحانی آن درس تشکیل می‌شود. به عبارت دیگر، تمام سؤالات امتحانی به نحوی در نمره کل دانشجو

1- Biggs, Collis & Kevin

پس از دسته‌بندی و گردآوری اطلاعات، کار تجزیه و تحلیل و انجام محاسبات لازم روی آنها به طور جداگانه صورت گرفت. ابتدا، اسامی دانشکده‌ها و تعداد اعضای هیأت علمی، ویژگیهای سؤالات امتحان پایان ترم آنها با استفاده از آمار توصیفی بررسی شد. آنگاه، ضرایب دشواری، تمیز و پایایی سؤالات امتحانی محاسبه و با سلاکهای مطلوب این سؤالات مقایسه گردید. در پایان، تحلیل نهائی آزمون بر اساس شاخصهای محاسبه شده برای عضو هیأت علمی ارسال گردید.

داده‌های توصیفی

اعضای هیأت علمی ۹ دانشکده و همراه با دانشکده‌های اقماری (شاخه شمالی، شاخه جنوبی، شاخه شوشتر) به عنوان یک دانشکده و در کل ده دانشکده در این تحقیق شرکت داشتند. تعداد اعضای هیأت علمی هر یک از دانشکده‌ها که به طور داوطلب در این مطالعه شرکت کردند متفاوت است. جدول ۱ و نمودار ۱ اسامی دانشکده‌ها و تعداد اعضای هیأت علمی که در تحقیق شرکت داشته‌اند را همراه با میزان درصد آنها نشان می‌دهد.

از تعداد ۱۰۹ عضو هیأت علمی که

سهیم هستند. بنابراین، شاخصهای روان‌سنجی بر روی تمام سؤالات به صورت انفرادی انجام می‌شود تا نقش مثبت یا منفی آن سؤال در نمره کل دانشجو نشان داده شود.

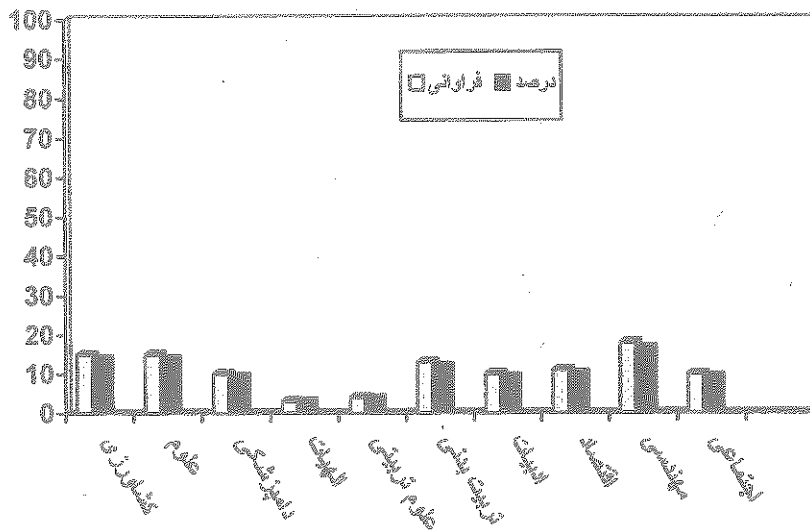
از این رو، در این تحقیق شاخصهای روان‌سنجی (ضرایب دشواری و تمیز) بررسی تمام سؤالات به صورت انفرادی انجام گردید تا نقش مثبت یا منفی هر سؤال در نمره کل دانشجو معلوم شود. برای محاسبه این شاخصها از "روش تفاوت" اختلاف بین نسبت پاسخهای دو گروه ملاک (قوی و ضعیف) استفاده شد (کروکر و آلگینا، ۱۹۸۶؛ گرانلاند، ۱۹۸۶؛ ویلسون، ۱۹۹۵). علاوه بر این، محاسبه پایایی آزمونهای تشریحی و عینی جداگانه بر اساس فرمولهای مربوطه (آلفای کرونباخ و کودر-ریچاردسون ۲۱) محاسبه گردید.

تجزیه و تحلیل داده‌ها

از اعضای هیأت علمی دانشگاه درخواست گردید به دلخواه اوراق امتحانی یکی از دروس خود را جهت تجزیه و تحلیل شاخصهای روان‌سنجی در اختیار محقق قرار دهند. تعداد ۱۰۹ نفر از ده دانشکده داوطلبانه در تحقیق شرکت کردند و اوراق امتحانی پایان ترم دانشجویان را در حالی که نمره‌گذاری شده بود برای محقق ارسال داشتند.

جدول ۱. آمار اعضای هیات علمی شرکت کننده در تحقیق

درصد فراوانی	درصد	فراوانی	دانشکده
۱۳/۸	۱۳/۸	۱۵	کشاورزی
۲۷/۵	۱۳/۸	۱۵	علوم
۳۶/۷	۹/۲	۱۰	دامپزشکی
۳۹/۴	۲/۸	۳	الهیات
۴۳/۱	۲/۷	۴	علوم تربیتی
۵۵/۰	۱۱/۹	۱۳	تربیت بدنی
۶۶/۴	۹/۲	۱۰	ادبیات
۷۴/۳	۱۰/۱	۱۱	اقتصاد
۹۰/۸	۱۶/۵	۱۸	مهندسی
۱۰۰/۰	۹/۲	۱۰	اجتماعی
	۱۰۰/۰	۱۰۹	کل



نمودار ۱. فراوانی دانشکده‌های مورد تحقیق

داوطلبانه در تحقیق شرکت داشتند تقریباً ۸۹ درصد آنها مرد و ۱۱ درصد زن بودند. احتمالاً نسبت هیأت علمی مرد به اعضای هیأت علمی زن در کل دانشگاه نیز با این درصدها مطابقت دارد.

دامنه میانگین تعداد دانشجویان کلاسهایی که اعضای هیأت علمی تدریس می‌کردند و از آنها امتحان به عمل آوردند از ۲۳ تا ۳۹ نفر متغیر بود. بالاترین میانگین تعداد دانشجویان به دانشکده الهیات و کمترین میانگین حجم کلاس به دانشکده‌های اقماری دانشگاه تعلق داشت.

اعضای هیأت علمی عمدتاً از سؤالات تشریحی، عینی (کوتاه پاسخ، کامل کردنی، صحیح غلط و عمدتاً چهار گزینه‌ای) به عنوان ابزاری برای اندازه‌گیری پیشرفت دانشجویان در امتحان پایان ترم استفاده کرده بودند. سؤالات تعداد معدودی از اساتید، به ویژه اساتید دانشکده مهندسی، از نوع مسأله تشکیل می‌شد. تقریباً ۶۴ درصد اعضای هیأت علمی از سؤالات تشریحی و ۳۶ درصد از انواع دیگر سؤالات استفاده کردند.

معمولاً سؤالات امتحانی همراه با راهنمایی در اختیار دانشجویان گذاشته می‌شود. این کار به آزمودنی کمک می‌کند تا آگاهیهای لازم نسبت به تعداد سؤالات آزمون،

بارم یا امتیازی که به سؤال یا اجزای آن تعلق می‌گیرد، مدت زمان آزمون، اعمال نمره منفی و جز اینها قبل از شروع امتحان کسب کنند. آگاهی از این موارد به آزمودنی کمک می‌کند تا با خیال آسوده افکار خود را برای پاسخ به سؤالات آزمون معطوف سازد. در تحقیق حاضر اوراق امتحانی که بین دانشجویان در پایان ترم توزیع گردیده بود تقریباً ۳۲ درصد دارای برگ راهت‌ما و ۶۸ درصد فاقد برگ راهت‌ما بود.

شیوه پاسخ دانشجویان به سؤالات امتحانی در بین اعضای هیأت علمی با یکدیگر متفاوت است. برخی از آنان از دانشجویان خواسته بودند که به سؤالات در روی برگه امتحانی پاسخ دهند، برخی کاغذ جداگانه فراهم کرده بودند و بالاخره عده‌ای هم به دانشجویان پاسخنامه داده بودند. ۴۹ درصد از اعضای هیأت علمی از دانشجویان خواسته بودند که روی برگه‌های امتحانی به سؤالات پاسخ دهند، ۴۷ درصد از برگه‌های جداگانه برای پاسخ به سؤالات استفاده کرده بودند و بالاخره ۴ درصد نیز از دانشجویان خواسته بودند که از پاسخنامه برای علامت‌گذاری روی گزینه‌ها استفاده کنند.

آیا جایی که در برگه امتحانی برای پاسخ به سؤالات در نظر گرفته می‌شود چقدر باید

شخصه‌های روان‌سنجی برای سؤالات هر آزمون به طور انفرادی و برای هر عضو هیأت علمی اعم از این که سؤالات تشریحی باشند یا عینی یا دست‌محاسبه شدند. به دلیل عدم دسترسی به نرم‌افزار مناسب، این کار با صرف وقت زیاد برای اطمینان از نتایج و با حوصله انجام پذیرفت. انجام محاسبات با دست را هم می‌توان از مزایا و هم از محدودیتهای تحقیق حاضر به شمار آورد. جدول ۲ محاسبه میانگین، انحراف معیار، ضریب تمیز، ضریب دشواری، ضریب پایایی و تعداد دانشجویان را برای هر دانشکده را به طور انفرادی و جدول ۳ این شاخصها را برای کل اعضای هیأت علمی مورد تحقیق نشان می‌دهد. علاوه بر این، چون اعضای هیأت علمی دانشکده‌ها از آزمونهای متفاوتی در امتحان پایان ترم استفاده کرده بودند، بنابراین، ویژگیهای روان‌سنجی سؤالات هم به صورت دانشکده و هم به طور انفرادی برای هر یک از اعضای هیأت علمی محاسبه گردید. به مناسب محرمانه بودن اعلام این نتایج، از ذکر نام عضو هیأت علمی خودداری و به جای آن از شماره کد هر استاد استفاده گردید. شاخصهای روان‌سنجی سؤالات همراه با تحلیل نهائی نتایج طی فرمهای ویژه برای آگاهی هر یک از اعضای هیأت علمی ارسال گردید.

باشد؟ گاهی اوقات ملاحظه می‌شود که معلمان با این موضوع تنگ نظرانه برخورد می‌کنند و جای کافی برای ارایه پاسخ در اختیار آزمودنی قرار نمی‌دهند. بررسی آن دسته از اوراق امتحانی که از دانشجویان خواسته شده بود پاسخهای خود را روی برگه‌های امتحانی زیر هر سؤال و در جایی که برای آن در نظر گرفته شده بنویسند، نشان داد که ۹۴ درصد جاهای در نظر گرفته برای پاسخ دانشجویان کافی و ۶ درصد از جاها برای پاسخ دانشجویان به سؤالات کافی نبود.

اعضای هیأت علمی از شیوه‌های متفاوت برای ارایه سؤالات در امتحانات پایان ترم استفاده می‌کنند. ۲۸ درصد از اعضای هیأت علمی مورد تحقیق سؤالات را روی یک برگ کاغذ و با دست نوشته بودند. ۴۰ درصد از دو روی یک صفحه کاغذ استفاده کرده و سؤالات را با دست نوشته بودند. ۱۳ درصد سؤالات خود را روی یک برگ کاغذ تایپ کرده و ۱۸ درصد از دو روی یک برگ کاغذ استفاده کرده و سؤالات را تایپ شده در اختیار دانشجویان قرار داده بودند.

محاسبه شاخصهای روان‌سنجی سؤالات

طبق روشی که قبلاً به آن اشاره شد،

جدول ۲. شاخصهای روان سنجی سؤالات امتحانی اعضای هیات علمی دانشکده به تفکیک

دانشکده ها	شاخصهای روان سنجی		ضریب همبستگی
	میانگین	انحراف معیار	
کشاورزی	میانگین	۰/۲۲	۰/۵۸
	میانگین	۰/۲۲	۰/۵۸
	انحراف معیار	۰/۲۵	۰/۱۴
علوم پایه	میانگین	۰/۲۳	۰/۶۰
	میانگین	۰/۲۲	۰/۶۱
	انحراف معیار	۰/۱۸	۰/۱۷
دامپزشکی	میانگین	۰/۲۸	۰/۶۰
	میانگین	۰/۲۱	۰/۶۱
	انحراف معیار	۰/۲۱	۰/۱۳
الهیات	میانگین	۰/۲۰	۰/۷۴
	میانگین	۰/۲۰	۰/۷۴
	انحراف معیار	۰/۳۰	۰/۱۲
علوم تربیتی و روانشناسی	میانگین	۰/۱۸	۰/۶۵
	میانگین	۰/۱۹	۰/۶۶
	انحراف معیار	۰/۱۹	۰/۱۵
تربیت بدنی	میانگین	۰/۱۹	۰/۶۳
	میانگین	۰/۲۰	۰/۷۲
	انحراف معیار	۰/۱۷	۰/۱۰
ادبیات	میانگین	۰/۲۵	۰/۶۵
	میانگین	۰/۲۱	۰/۶۷
	انحراف معیار	۰/۱۴	۰/۱۳
اقتصاد	میانگین	۰/۲۸	۰/۶۷
	میانگین	۰/۲۲	۰/۷۳
	انحراف معیار	۰/۱۷	۰/۱۱
مهندسی	میانگین	۰/۲۶	۰/۵۳
	میانگین	۰/۲۵	۰/۵۲
	انحراف معیار	۰/۱۱	۰/۱۱
اقتصادی	میانگین	۰/۲۰	۰/۶۲
	میانگین	۰/۲۰	۰/۶۴
	انحراف معیار	۰/۱۰	۰/۱۲

جدول ۳. میان، میانگین، انحراف معیار و پایایی سؤالات امتحانی کل اعضای هیأت علمی مورد تحقیق

انحراف معیار	میانگین	میان	شاخصهای آماری شاخصهای روانسنجی
۰/۱۳	۰/۶۱	۰/۶۲	ضریب دشواری
۰/۱۳	۰/۲۳	۰/۲۱	ضریب تمیز
۰/۲۰	۰/۵۵	۰/۲۰	ضریب پایایی

بحث و بررسی

تشخیص داده بودند که استفاده از سؤالات تشریحی و حل مسأله بهتر می تواند آموخته های دانشجویان آنها را اندازه گیری کنند. لذا همه سؤالاتی که اعضای هیأت علمی این دانشکده برای امتحان پایان ترم انتخاب کرده بودند از این نوع سؤالات بود.

در بررسی نوع سؤالات امتحانی در بین سایر دانشکده ها چنین انسجامی دیده نمی شد. برای مثال تقریباً ۵۰ درصد اساتید دانشکده دامپزشکی مورد تحقیق از سؤالات تشریحی و ۵۰ درصد از سؤالات عینی و عمدتاً چهار گزینه ای استفاده کرده بودند. اعضای هیأت علمی دانشکده اقتصاد و علوم اجتماعی بیشتر سؤالات امتحانی خود را به استفاده از آزمونهای چهار گزینه ای اختصاص داده بودند.

استاد از چه نوع سؤالهایی برای امتحانات درسی خود استفاده کند حائز اهمیت زیادی است. هر نوع از سؤالات آزمون ویژگیهای خاص خود را دارد و برای اندازه گیری مفاهیم خاصی از درس مناسب است. برای مثال،

مطالعه و بررسی روی اوراق امتحانی هیأت علمی در یک مؤسسه آموزش عالی کار چندان آسانی نیست. شاید بدلائیل عدم تجربه کسافی در ارزیابی سؤالات امتحانی، روحیه استقلال طلبی و محرمانه بودن سؤالات، اعضای هیأت علمی اگره دارند که شاخصهای روان سنجی سؤالات مثل عینیت، دشواری، قدرت، تمیز، ثبات و دقت سؤالات امتحانی آنها از جانب یک محقق بی طرف مورد ارزیابی قرار گیرد. لذا با این که از همه اعضای هیأت علمی دعوت شده بود با ارسال اوراق امتحانی خود در این تحقیق مشارکت کنند، فقط ۱۰۹ نفر داوطلبانه در این تحقیق شرکت جستند.

بررسی نوع سؤالات نشان داد که اعضای هیأت علمی دانشکده ها با توجه به موضوعی که تدریس می کنند از آزمونهای یکسانی استفاده نمی کنند. برای مثال، اعضای هیأت علمی دانشکده مهندسی چنین

اینها از جمله مسائلی است که باید هنگام استفاده از این آزمونها مورد ملاحظه عضو هیأت علمی قرار گیرد. در کلام آخر، استاد لازم است مزایا و محدودیتهای این نوع سؤالات را بداند و هنگامی که قصد دارد مفاهیم ذهنی خود را در قالب سؤالات امتحانی به صورت سؤالات تشریحی روی کاغذ آورد دقت و توجه کافی به این ویژگیها مبذول نماید.

از جمله آزمونهای عینی می توان به آزمونها کوتاه- پاسخ، کامل- کردنی، جور- کردن، صحیح- غلط و چهار- گزینه ای اشاره کرد. از این نوع آزمونها عمدتاً می توان برای اندازه گیری مفاهیم حفظی درس استفاده کرد. دانش قواعد، دانش فرمولها، دانش اصول و جز اینها مطالبی هستند که به آسانی از طریق استفاده از آزمونها عینی به دقت قابل اندازه گیری است.

بررسی اوراق امتحانی نشان داد که اساتید به نحو شایسته ای از این نوع سؤالات برای مفاهیم فوق استفاده کرده بودند. تنها نوع آزمون عینی که اخیراً توجه خاصی را به خود معطوف داشته و به سبب محاسبه ویژگیهای روان سنجی آن مورد توجه متخصصان آزمون سازی قرار گرفته است آزمونهای عینی از نوع چهار-گزینه ای است. این آزمونها اصطلاحاً به آزمون تستی شهرت دارند. این نوع آزمونها چنانچه با دقت و مهارت کافی ساخته شوند

ویژگی سؤالات آزمونهای تشریحی در این است که تنها یک پاسخ صحیح ندارند. به بیان دیگر، دانشجو در ارابه پاسخ آزادی عمل دارد. این نوع سؤالات باید این فرصت را برای دانشجو فراهم کند تا با اندیشیدن پاسخ سؤالات را به شیوه منطقی تنظیم و مستدل رایه کند. علاوه بر این، چنانچه توانایی انتقال مفاهیم ذهنی به صورت نوشتاری از اهداف مهمی است که آموزش عالی دنبال می کند، این آزمونها تنها انواعی هستند که قادرند به این هدف مهم جامه عمل بپوشانند. به بیان دیگر، نوع سؤالات تشریحی برای اندازه گیری مفاهیم عالیتر ذهنی مناسب اند و تا جایی که مقدور باشد نباید از آنها برای اندازه گیری محفوظات درسی استفاده کرد.

برخی از پاسخهای آزمونهایی که سؤالات به صورت تشریحی ارابه شده بود فاقد ویژگیهای فوق بود. بررسیهای بیشتر نشان داد اغلب پاسخهایی که دانشجویان به یک سؤال تشریحی داده بودند یکسان بود. یکسان بودن تمام پاسخهای دانشجویان به سؤالات آزمون تشریحی به این معنی است که دانشجویان از محفوظات خود برای پاسخ به سؤال استفاده کرده اند که این عمدتاً از اهداف استفاده از سؤالات تشریحی به شمار نمی رود. مسأله عدم عینیت در نمره گذاری، ضعف نمونه گیری از تمام محتوای درس، عدم ثبات و دقت و جز

اختیار دانشجو قرار دهند. این کار سبب خواهد شد ویژگیها مورد نظر استاد را، دانشجو هنگام پاسخ به سؤالات مراعات کند. ۶۷ درصد از اوراق امتحانی راهنمایی در زمینه زمان آزمون، نمره گذاری سؤالات (بارم بندی)، تعداد سؤالات، و جز اینها در اختیار دانشجویان قرار نداده بودند. تنها ۳۳ درصد از اوراق امتحانی به این امر توجه کرده و اطلاعات لازم را قبلاً در اختیار دانشجویان قرار داده بودند.

در اجرای آزمون، سؤالات زیادی مطرح می شود که عضو هیأت علمی باید قبلاً پاسخهای منطقی برای آنها در نظر بگیرد. از جمله این سؤالات یکی این است که شیوه پاسخ دانشجویان به سؤالات چگونه باید باشد؟ آیا سؤالات به شیوه ارایه می شود که دانشجویان باید پاسخ آنها را در جایی که استاد برای این کار در نظر گرفته است بنویسند؟ آیا دانشجویان باید پاسخهای خود را در ورقه جداگانه وارد کنند؟ یا این که دانشجو باید از پاسخنامه استفاده کند. تقریباً ۵ درصد اوراقی که لازم بود دانشجو پاسخ خود را زیر هر سؤال و یا در برگ جداگانه بنویسد یکسان و به ترتیب ۴۸ و ۴۷ درصد بود. فقط ۵ درصد از پاسخنامه استفاده کرده بودند.

در بحث و بررسی مورد فوق می توان به این نکته اشاره کرد که آن دسته از اوراق امتحانی که از دانشجو خواسته شده بود پاسخ

برای اندازه گیری تمام اهداف درس در تمام سطوح فکری و در تمام رشته های تحصیلی مناسب است (مهرنژولهمان، ۱۹۸۶؛ ایبل، ۱۹۹۱). به بیان دیگر، اگر پایایی و اعتبار را ویژگی عمده سؤالات آزمون بدانیم آزمونهای چهار-گزینه ای از بوته آزمایش روان سنجان سر بلند بیرون آمده است. اکثر آزمونهای استاندارد که در سطح بین المللی از جمله Miller Anology Test, GRE, TOEFEL و جزء اینها برای گزینش دانشجو در سطوح مختلف تحصیلی انجام می پذیرد، از نوع چهار-گزینه ای هستند. (کیامنش، ۱۳۷۷) آزمونهای پیشرفت تحصیلی استاندارد زیادی نیز را می توان نام برد که از این نوع سؤالات برای ارایه مطالب درسی استفاده کرده اند.

بررسی اوراق امتحانی که از سؤالات چهار-گزینه ای استفاده کرده بودند نشان داد بسیاری از اعضا از این نوع سؤالات برای اندازه گیری مفاهیم درسی استفاده کرده بودند. در بحثهایی که درباره ضرایب دشواری، تمیز و پایایی خواهیم داشت به این مطلب که آیا ویژگیهای روان سنجی این سؤالات در مقایسه با ملاک مطابقت دارند یا خیر اشاره خواهد شد.

همان طور که قبلاً اشاره شد، استاد لازم است قبل از شروع امتحان، اطلاعات و راهنمایی درباره چگونگی اجرای امتحان در

اوراق امتحانی زیادی برای ارایه سؤالات استفاده کرد. نتایج بررسی اوراق امتحانی نشان داد که ۶۸ درصد سؤالات با دست و روی یک یا دو برگ کاغذ نگارش شده بود و ۳۲ درصد سؤالات به صورت تایپ شده در اختیار دانشجویان قرار گرفته بود. در نگارش سؤالاتی که تایپ شده در اختیار دانشجویان قرار داده شده بود غلطهای املائی و دستوری قابل توجهی به چشم می خورد.

آنچه لازم است در اینجا به آن اشاره شود، این است که چنانچه قرار است سؤالات را به صورت دست نویس (چه تشریحی و چه چهار-گزینه‌ای) در اختیار دانشجو قرار دهیم، حتماً باید خوش خط و خوانا نوشته شود. چنانچه از خط خوشی برخوردار نیستیم، به دلیل مشکلاتی که دانشجو در خواندن بعضی کلمات در جلسه امتحان با آن مواجه خواهد شد، حتماً سعی گردد سؤالات به صورت تایپ شده آماده توزیع بین دانشجویان باشد. ویراستاری سؤالاتی که توسط متشی گروه تایپ می شود، برای اطمینان از صحت و سقم آن، لازم است دقیقاً واریسی شوند.

تحلیل دشواری، تمیز و ضریب همسانی درونی سؤالات

جدول ۳ میانگین دشواری، تمیز و همسانی درونی (KR20) سؤالات امتحانی

سؤالات را در برگ جداگانه بنویسند و یا در پاسخ نامه وارد کنند مشکلی مشاهده نشد. ولی در مورد اوراق امتحانی که دانشجو لازم بود پاسخ به سؤالی را در زیر آن (جایی که برای آن پیش بینی شده‌اند) بنویسد، ناهماهنگی‌هایی مشاهده گردید. از یک سو به دلیل حجم زیاد پاسخ و از سوی دیگر جای ناکافی که برای پاسخ به سؤال در نظر گرفته شده بود، دانشجو مجبور شده بود از قسمتهای سفید یا حاشیه ورقه امتحانی استفاده کند و توجه استاد را با ايماء و اشاره به دنباله پاسخهای خود برای مثال، در پشت صفحه و یا زیر سؤال معطوف سازد. خوشبختانه درصد قابل توجهی (۹۴ درصد) از اوراق امتحانی، جای کافی برای پاسخ دانشجو به سؤالات آزمون فراهم کرده بودند. فقط در ۶ درصد از اوراق، دانشجو برای نوشتن تمامی پاسخهای خود با مشکل مواجه شده بودند.

ارایه سؤالات امتحانی می تواند به شیوه‌های مختلف انجام پذیرد. می توان سؤالات را با دست نوشت یا این که آنها را تایپ کرد. چنانچه از آزمونهای تشریحی استفاده می کنیم و به دلیل تعداد کم سؤالات، معمولاً می توان سؤالات را در یک برگ با دست نوشت. ولی چنانچه آزمونها عینی است و به ویژه از نوع چهارگزینه‌ای استفاده می شود، به دلیل تعداد زیاد سؤالات باید از

آزمون برای اندازه‌گیری پیشرفت دانشجو برای اطمینان از رسیدن به هدف مورد نظر وجود دارد.

چنانچه این ثبات درونی بین سؤالات مشاهده نشود، خواهیم گفت آزمون با مقداری خطا همراه است. برای آزمونهایی که هدف آنها اندازه‌گیری پیشرفت تحصیلی دانشجو است، متخصصان آزمون سازی ضریب پایایی برابر ۰/۶۷ را پیشنهاد می‌کنند (مهرنزلوهمان، ۱۹۸۴؛ اپیل، ۱۹۹۱؛ رامبرگ، ۱۹۹۰). هر چه این ضریب به عدد یک نزدیک شود، بین سؤالات آزمون همسانی درونی بالایی در اندازه‌گیری پیشرفت دانشجو وجود دارد و به کاهش میزان خطای اندازه‌گیری و ثبات بیشتر نمره‌ها می‌انجامد.

پیشنهادات

بر اساس نتایج تحقیق حاضر می‌توان موارد زیر را پیشنهاد کرد:

۱. دانشگاه باید فرصتهای مناسبی برای اعضای هیأت علمی فراهم کند تا بیشتر با نظریه‌ها و مدل‌های جدید آزمون سازی که می‌توانند در اندازه‌گیری تواناییهای عالی فکر دانشجویان در کلاس درس کاربرد داشته باشند ارتقاء دهند.
۲. کلاسهای آزمون‌سازی، به ویژه برای اعضای که به تازگی به جمع اعضای هیأت

دانشکده‌های مختلف را نشان می‌دهد. علاوه بر این، نتایج شاخصهای فوق برای هر یک از اوراق محاسبه همراه با پیشنهاداتی برای استاد مربوطه ارسال گردید. علاقه‌مندان می‌توانند برای مشاهده تحلیل ضرایب دشواری، تمیز و پایایی به صورت انفرادی به جدول مراجعه کنند. ضمناً، میانگین این ضرایب همراه با سایر اطلاعات آزمون در جدول به طور کامل ارائه شده است.

متخصصان آزمون سازی معتقدند که سؤالات امتحان نباید زیاد دشوار و یا زیاد آسان ساخته شوند. سؤالاتی که همه دانشجویان پاسخ آنها را می‌دانند سبب می‌شود نمره ثابتی به نمره هر یک از دانشجویان اضافه شود. چنانچه بتوان دشواری سؤالات را در حد متوسط قرار دهیم، می‌گوییم که امتحان از سطح دشواری مطلوبی (متوسطی) برخوردار است. این گونه سؤالات نیز قدرت بیشتری در تمایز بین دانشجویان قوی و ضعیف (ضریب تمیز) دارند.

به نظر می‌رسد که باید بین تمام سؤالات یک آزمون یک هماهنگی و همسانی منطقی برقرار گردد. این موضوع به این دلیل مطرح می‌شود که معمولاً تمام سؤالات یک آزمون در جهت اندازه‌گیری اهداف واحدی ساخته می‌شود. لذا، ضریب همسانی بالا نشان می‌دهد که یک ارتباط منطقی بین سؤالات

- علمی پیوسته‌اند در طول سال تحصیلی برگزار گردد. شرکت در این کلاسها کمک می‌کند که اعضا با شیوه‌های مختلف آزمون‌سازی آشنایی پیدا کنند. خوشبختانه در سالهای اخیر مدیریت برنامه‌ریزی و ارزشیابی دانشگاه کلاسهایی در جهت آشنا ساختن اعضای هیأت علمی با روشهای تدریس و آزمون‌سازی تشکیل داده است. شرکت اعضای هیأت علمی در این کلاسها توانسته است به این پیشنهاد تحقیق جامعه عمل بپوشاند.
۴. پیشنهاد می‌شود مرکز آزمون‌سازی در حوزه مدیریت برنامه‌ریزی و ارزشیابی دانشگاه به طور دایم تشکیل گردد. اعضای هیأت علمی در دانشکده‌های مختلف لازم است اوراق امتحانی خود را پس از نمره‌گذاری به این مرکز ارسال دارند. این مرکز موظف است ضمن بررسی سؤالات امتحانی، اطلاعاتی در مورد ضرایب دشواری، تمیز، پایایی و سایر ویژگیهایی که در این تحقیق به آنها اشاره شد در اختیار عضو هیأت علمی قرار دهد. اگر بررسیهای این مرکز نشان داد که سؤالات امتحان از ویژگیهای مطلوب برای اندازه‌گیری پیشرفت تحصیلی برخوردارند، مراتب به اطلاع عضو هیأت رسانده شود.
- چنانچه بررسی ویژگیهای مختلف آزمون نشان داد که آزمون نیاز به تجدیدنظر کلی یا جزئی در برخی موارد دارد، این‌گونه اطلاعات نیز برای عضو هیأت علمی ارسال گردد. عضو هیأت علمی لازم است در مواردی که به عنوان نقایص آزمون به آنها اشاره شده است با راهنمایهای مرکز در سؤالات آزمون خود تجدیدنظر اساسی اعمال نماید.
۴. بانکی از سؤالات برای دروس عمومی که توسط چند استاد تدریس می‌شود تهیه گردد. ویژگیهای روان‌سنجی این قبیل سؤالات قبلاً محاسبه و با سئالکهای مورد نظر مقایسه شده‌اند. سؤالات امتحان نهایی این دروس می‌تواند از بین سؤالات این بانک به طور تصادفی انتخاب گردد.
۵. با این که این تحقیق بر روی افراد داوطلب انجام پذیرفت پیشنهاد می‌شود که تحقیق مشابهی بر روی کسب جامعه اعضای هیأت علمی دانشگاه و یا یک نمونه تصادفی بزرگ از آنها انجام پذیرد. برای انجام این کار، لازم است اوراق امتحانی که به طور تصادفی انتخاب می‌شوند برای بررسی و محاسبه شاخصهای روان‌سنجی در اختیار محقق قرار داده شود. به نظر می‌رسد در صورت تحقق چنین مطالعه‌ای نتایج به دست آمده تصویر روشنتری از شیوه‌های ارزشیابی اعضای هیأت علمی در سطح دانشکده‌های دانشگاه ارایه شود.
۶. معمولاً ارزشیابی دانشجویان از اساتید خود قبل از اجرای آزمون آنها انجام می‌شود. نتایج مطالعاتی که پس از اجرای آزمونها انجام

گرفته نشان می دهند. آزمودنیها رضایت چندانی از شیوه آرایه سؤالات، نمره گذاری، وقت آزمون و جز اینها نشان نداده اند (مهرنزلوهمان، همان و گلوک، ۱۹۹۷). بنابراین

پیشنهاد می شود که اوراق امتحانی پس از اجرای آزمون توسط افراد بی طرف مورد بررسی و تجزیه و تحلیل قرار گیرد.

منابع

فارسی

- بلوم، بنجامین (۱۹۸۲). ویژگیهای آماری و یادگیری آموزشی، ترجمه علی اکبر سیف (۱۳۲۳). تهران مرکز نشر دانش.
- صدیق، عیسی (۱۳۳۶). تاریخ فرهنگ ایران. انتشارات دانشگاه تهران. شماره ۴۲۴.
- کیامنش، علیرضا (۱۳۷۱). آزمون چند-گزینه‌ای و تأثیر آن بر یادگیری دانش‌آموزان، محنتی آموزش و روش تدریس. فصلنامه تعلیم و تربیت، شماره ۳۱.
- کیامنش، علیرضا (۱۳۷۳). ارزشیابی دروس ریاضی و هندسه نظام جدید آموزش متوسطه، فصلنامه تعلیم و تربیت، شماره‌های ۳ و ۴.

لانی

- Ahman, J. & Glock, M. (1971). *Evaluating Pupil Growth: Principles of the Tests and Measurements* (4th ed.), Boston: Allyn and Bacon.
- Baillie, C. & Toohey, S. (1997). 'The power test', *Assessment and Evaluation in Higher Education*, 22(1), pp 33-48.
- Brinbaum, M. & Tatuska, K.K. (1987). Open versus multiple choice response formats. *Journal of Applied Psychological Measurement*, 11, 385-95.
- Brown, S. & Knight, P. (1994). *Assessment of Learners in Higher Education*. London: Kogan Page.
- Brooker, R. & Smith, D. (1996). Assessing tertiary student in an education faculty. Rhetoric and Reality, *Higher Education Research & Development*, 5(2), pp 103-175.

- Baker, L. (1990). *Developing Comprehensive Assessment of Higher Order Thinking*. Englewood cliffs, New Jersey: Prentice Hall.
- Biggs, B., Collis, M. & Kevin, F. (1982). *Evaluating the value of quality learning. The Solo Taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press, Co.
- Crocker, L. & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winson, Inc.
- Cronbach, L.J. (1984). Further evidence on response sets and test design. *Educational and Psychological Measurement*, 10, 3-13.
- Cook, Averil (2002). Assessing the use of flexible assessment. *Assessment and Evaluation in Higher Education*, Vol. 26, No. 6, pp 540-549.
- Ebel, A.L. & Frisbie, A.D. (1991). *Essentials of Educational Measurement*. Englewood Cliffs, New Jersey: Prentice Hall.
- Gronlund, N.E. (1986). *Constructing Achievement Tests*. Englewood cliffs, New Jersey: Prentice Hall.
- Mehrens, A.W. & Lehman, J.I. (1984). *Measurement and Evaluation in Education and Psychology*. New York: Holt, Rinehart and Winston Inc.
- Night, P., Te Wiata, I., Toohey, S., Ryan, G., Hughes, C., & Maginn, D. (eds) (1996). *Assessing Learning in Universities*. Sydney: University of New South Wales Press.
- Ritter, Leonora (2000). The quest for an effective form of assessment: The evolution and evaluation of a controlled assessment procedure (CPA). *Assessment and Evaluation in Higher Education*, Vol. 25, No. 4 PP 307-320.
- Romberg, A.T. (1990). *A New World View of Assessment in Mathematics*. Washington DC: American Association for the Advancement of Science.
- Trub, R.E., & Hamblen, R.K. (1972). The effect of scoring instructions and degree of speediness on the validity and reliability of multiple-choice tests. *Educational and Psychological Measurement*, 32, 737-758.
- Wilson, M. (1995). *Investigating of Structured Problem Solving Items*. New York: Academic Press.