



طراحی الگوی مفهومی برای سنجش تحصیلی دانشگاه باز و از دور

فهمیه‌السادات حقیقی *

مهران فرج‌اللهی **

چکیده

هدف این تحقیق، طراحی و اعتباریابی الگوی مفهومی برای آزمون‌های دانشگاه باز و از دور است. روش تحقیق از نوع آمیخته (کمی و کیفی) است. جامعه تحقیق، اعضای هیأت علمی دانشگاه پیام‌نور ایران در سال تحصیلی ۹۳-۱۳۹۲ بودند. حجم نمونه کمی طبق جدول کرجسی و مورگان، ۲۸۶ نفر تعیین شد که به طور تصادفی انتخاب شدند. نمونه کیفی شامل ۷ متخصص طراحی سؤال بود که به روش گلوله برفی انتخاب شدند. ابزارهای تحقیق، پرسش‌نامه محقق‌ساخته با ۵۸ سؤال با پایایی ۰/۹۶۴ و مصاحبه نیمه ساختارمند بود. عناصر پنج‌گانه فلسفه، اهداف، طراحی، اجرا و ارزیابی نتایج و هم‌چنین شاخص‌های سه‌گانه هر عنصر مورد تحلیل توصیفی و آزمون t تک‌گروهی قرار گرفتند. نتایج نشان داد که عنصر فلسفه از میانگینی در حد ملاک (۹) و سایر عناصر میانگینی بالاتر از ملاک داشتند و هر ۵ عنصر در سطح ۰/۹۹٪ معنی‌دار بودند. شاخص‌های چشم‌انداز توسعه، فلسفه آموزش عالی و دانشگاه باز و از دور عنصر فلسفه، میانگینی در حد ملاک (۳/۶) و ۰/۹۹٪ معنی‌داری؛ شاخص‌های اهداف دانشگاه، اهداف شاخه، رشته و گروه از عنصر اهداف، میانگینی بالاتر از ملاک و ۰/۹۹٪ معنی‌داری؛ شاخص‌های قوانین ارزشیابی و کیفیت آزمون از عنصر طراحی، میانگینی بالاتر و معنی‌داری ۰/۹۹٪ داشتند؛ اما شاخص استاندارد طراحی سؤال میانگینی در حد ملاک داشت که معنی‌دار نبود. شاخص‌های سیاست و شیوه اجرا از عنصر اجرا، میانگینی در حد ملاک و ۰/۹۹٪ معنی‌داری و شاخص شرایط اجرا میانگینی بالاتر از ملاک و ۰/۹۵٪ معنی‌داری داشت. دو شاخص تصحیح و نمره‌گذاری از عنصر ارزیابی، از میانگینی در حد ملاک و شاخص تحلیل و بازخورد میانگینی بالاتر از ملاک برخوردار و هر سه شاخص ۰/۹۹٪ معنی‌دار بودند. یافته‌ها نشان دادند که می‌توان از این الگو در طراحی آزمون‌های دانشگاه باز و از دور استفاده کرد.

واژگان کلیدی

آموزش عالی، دانشگاه باز و از دور، سنجش پیشرفت تحصیلی، الگو، آزمون

* دکتری برنامه‌ریزی آموزش از دور، دانشگاه پیام‌نور، تهران، ایران fahimehaghghi@yahoo.com

** دانشیار گروه علوم تربیتی، دانشگاه پیام‌نور، تهران، ایران farajollahim@yahoo.com

نویسنده مسؤول یا طرف مکاتبه: فهمیه‌السادات حقیقی

مقدمه

امروزه به دلایل مختلف آموزش باز و از دور^۱ به عنوان یکی از گزینه‌های دست‌یابی به اطلاعات مداوم و همیشگی در هر کجا و در هر زمان، نگرش جهان را به انسان از نگاه ابزاری به منبع ثروت و دارایی فکری^۲ (دوران مغز افزاری^۳) تغییر داده و فرآیند یادگیری مستقل را به خصوص در آموزش عالی به عنوان ابزار توانمندسازی او معرفی نموده است. فیگنباوم (Fegenbaum, 1994) و فریمن (Freeman, 1993) اعتقاد دارند که محیط‌های آموزشی و دانشگاه‌ها باید از قدرت تخیل، دانش، هوشمندی و قضاوت برای حل مشکلات و تضادهای روزافزون برخوردار باشند. بارنت، پاری و کوت (Barnett, Parry & Coate, 2001) معتقد هستند که آموزش عالی و به تبع آن دانشگاه‌ها در عصر کنونی دچار وضعیت ابر پیچیدگی^۴ می‌باشند. هدف نظام آموزش باز و از دور رسیدن به استانداردها و کیفیت مطلوب آموزشی است. اصلی‌ترین چالش فرا روی آموزش باز و از دور، اطمینان از تحقق، یادگیری و چگونگی بالا رفتن از هرم یادگیری از مرحله دانستن به مرحله کار بستن در عرصه کار و زندگی می‌باشد و این مهم بر عهده نظام ارزشیابی است.

در ادبیات و مباحث نظری نظام آموزشی، نسل اول ارزشیابی^۵ (اندازه‌گیری) در دهه ۱۹۳۰ به دنبال محاسبه نرخ حافظه و دانش و شناخت دانشجویان پا به عرصه حیات علمی گذاشت و نهضت اندازه‌گیری تا مدت‌ها رویه معمول ارزشیابی بود. تایلر (Tyler, 1969) نقش متفاوتی برای ارزشیابی در نظر گرفت و اهداف آموزشی را برجسته و نسل دوم ارزشیابی توصیفی را پایه‌گذاری کرد. نسل سوم ارزشیابی تکیه بر کارکرد مدیریتی و قضاوتی داشت. ورتن و سنדר تمامی رویکردهای ارزشیابی نظام سنتی را در شش الگو خلاصه نموده‌اند که در ادامه توصیف می‌شود (cited in Seif, 2005, 70):

الف. الگوی ارزشیابی مبتنی بر هدف: تایلر (Tyler, 1969)، پرووس (Provus, 1973)، پوفام (Popham, 1988)، تابا (Taba, 1962)، بلوم (Bloom, 1968).

ب. الگوی ارزشیابی مبتنی بر مدیریت: استافل بیم (Stufflebeam, 1971)

ج. الگوی ارزشیابی مبتنی بر نظر مصرف‌کنندگان: اسکریون (Scriven, 1967).

1. Open and Distance Education
2. Intellectual Property
3. Brain Age Software
4. Highest Complexity
5. Evaluation
6. Assessment

د. الگوی مبتنی بر نظر متخصصان و الگوی مبتنی بر اختلاف نظر متخصصان.

ه. الگوی مبتنی بر مشارکت کنندگان: استیک (Stake, 1967).

عناصر کلیدی این الگوها بر اساس عناوین شان شامل: اهداف، نتایج، فرآیند، اهداف آزاد و غیررسمی می باشد.

از دهه ۱۹۷۰ به بعد طرح سؤال‌هایی در زمینه کیفیت یادگیری در آموزش، باعث شد تا نهضت ارزشیابی کیفی^۱ و ارزشیابی فرآیندی^۲ (نسل چهارم) پا بگیرد. تالنت-رانلز (Tallent-Runnels, 2006) اعتقاد دارند که ارزشیابی کیفی کمک می کند تا متغیرها و فرضیه‌هایی برای تحقیقات بعدی فراهم شود. مریام (Merriam, 2002) هدف از ارزشیابی کیفی را ساخت تئوری بر مبنای شواهد تجربی می داند. ماک کالاج (McCulloch, 1997)، مک آلیستر (McAlister, 1998) و هندرسون و پات (Henderson & Putt, 1999, 25) اعتقاد دارند که کارکرد واقعی یادگیری در مقابل کارکرد مورد انتظار یادگیری به وسیله سنجش فرآیندی خود را نشان می دهد. کاراسلی^۳، تاشاکوری و تدلی^۴ ارزشیابی فرآیندمدار و نتیجه‌مدار را با هم توصیه می کنند (cited in Ruhe & Zumbo, 2009, 60).

رو و زامبو (Ruhe & Zumbo, 2009) الگوهای ارزشیابی در نظام آموزش باز و از دور را بر اساس مدل مسیک (Messick, 1988, 1989) در دو بعد علمی (۹ مدل) و نتایج و برآیند (۸ مدل) تقسیم‌بندی کرده‌اند (جدول ۱).

جدول ۱. مبانی علمی و برآیندی مدل‌های ارزشیابی آموزش از دور

نام محققان	توضیحات	محتوی	ابعاد	زمینه
(Van Slyk, 1998) (Belange & Jordan, 2000) (Bates, 1995)	جمع آوری اطلاعات از طریق مصاحبه، مشاهده، نظر سنجی درون وبسایت و محتوای درسی	شواهد علمی	تفسیر	
(Scanlon, Jones, Barnard, Thompson & Calder, 2000) (Kirkpatrick, 2006) (Hughes & Attwel, 2002)	بین برنامه درسی و نیازهای جامعه، معنی داری برنامه درسی برای یادگیرنده، انتقال یادگیری به زمینه‌های اخلاقی و عملی	ارتباط	استفاده	مبانی علمی
(Lam & McNaught, 2005) (Zaharias, 2005)	دانشگاه، دانشجو	هزینه - منفعت		
(Gooler, 1979) (Rumbel, 1981) (Collis, 1993) (Clark, 1994) (Mann, 1998)	اهداف و محتوای برنامه، نقش مشارکت کننده‌ها، استانداردها و ...	آشکار سازی ارزش‌ها	تفسیر	برآیند و نتیجه
(Lorenzo & Moor, 2002) (Ruhe & Zumbo, 2009) (Baker & O'Neil, 2006)	ساختاری، اجتماعی، تحقق برنامه و ...	نتایج ناخواسته	استفاده	

تلاش‌ها و مطالعات گسترده‌ای که در زمینه سنجش در آموزش از دور صورت گرفته، بدون تصویری درست از یادگیری عمیق و شرایط یادگیری در محیط از دور و باز آموزشی انجام شده است. به همین دلیل سنجش در آموزش از دور با موانع متعدد روبرو شده است. بروک‌هارت (Brookhart, 2007) بیان می‌کند که تئوری‌های اندازه‌گیری و سنجش آموزش از دور به دلیل اختلاف محتوایی، یکپارچگی آموزش و سنجش با فن آوری و مقاصد متفاوت سنجش ورودی، تکوینی و پایانی تفاوت چشم‌گیری نسبت به تئوری‌های قدیمی سنجش دارند.

در زمینه پیشینه تحقیقات خارجی می‌توان به تحقیق اکهات، دیویس، مرچلسون و گودبرن (Eckhout, Davis, Merchelson & Goodburn, 2005) درباره الگویی برای آماده‌سازی فعالیت‌های سنجش معلمان مشغول تدریس و داوطلبان معلمی، اشاره کرد. نتایج نشان داد که در پیش‌آزمون، نمونه تحقیق در استفاده از روش‌های مختلف سنجش پیشرفت تحصیلی عملکرد بالایی داشتند، اما در زمینه شفاف‌سازی اهداف پیشرفت تحصیلی عملکرد ضعیف داشتند، پس آزمون نشان داد که در زمینه درک، تولید و استفاده از انواع آزمون عملکرد بالایی نشان دادند.

اما در شفاف‌سازی اهداف یادگیری ضعیف عمل کردند. پگی (Peggy, 2005) در تحقیقی با عنوان «میزان سواد داوطلبان معلمی در سنجش»، نشان داد که معلمانی که آموزش دیده‌اند و از سواد و ادبیات سنجش و اندازه‌گیری برخوردارند، در سنجش یادگیری بهتر از داوطلبانی که به این مسایل اشراف ندارند عمل می‌کنند. مرتلر (Mertler, 2004) در تحقیق دیگری با عنوان «سواد سنجش معلمان دبیرستان: آیا تجربیات کلاسی تغییری ایجاد می‌کند؟»، نشان داد که سواد سنجش معلمانی که مشغول تدریس هستند با معلمانی که در شرف تدریس هستند تفاوت دارد. معلمانی که تدریس می‌کردند و تجربیات کلاسی داشتند در اجرای سنجش، نمره‌گذاری و تفسیر نتایج بسیار قوی عمل کرده بودند؛ اما معلمانی که داوطلب تدریس بودند رویه‌های مناسب‌تری را برای سنجش به کار برده بودند. در کل، معلمانی که تجربه کلاسی داشتند امتیازهای بهتری در آزمون‌های تحقیق کسب کردند. ایمپرا، پلک و فاگر (Impara, Plak & Fager, 1993) در تحقیق خود با عنوان «پیش‌دانسته‌های معلمان از سنجش و اندازه‌گیری و تمایل آنها در آزمون‌سازی»، نشان دادند که ۷۰٪ معلمان آگاهی اندکی از سواد سنجش دارند و ۳۰٪ اصلاً دانشی از سنجش و اندازه‌گیری ندارند.

در خصوص پیشینه تحقیقاتی داخلی با مرور در ادبیات، تحقیق، نظریه و الگوی خاصی درباره آزمون‌های دانشگاه باز و از دور یافت نشد. اما پیرامون سنجش و اندازه‌گیری در نظام سنتی نتایجی به دست آمد که در ادامه بررسی می‌شود.

یونسی (Younesi, 2005) ویژگی‌های روان‌سنجی سؤال‌های آزمون‌های فراگیر رشته روانشناسی دانشگاه پیام‌نور را بررسی نمود. جهت تجزیه و تحلیل آزمون و سؤالات آن بر طبق نظریه کلاسیک اندازه‌گیری، ضریب دشواری، ضریب تمیز، روش لوب، ضریب پایایی آزمون، واریانس سؤالات و فراوانی گزینه‌های انحرافی محاسبه شد. نتایج نشان داد که کلیه آزمون‌ها تک بعدی هستند و تحلیل گزینه‌های انحرافی نشان داد که گزینه‌های انحرافی تمامی سؤالات در همه آزمون‌ها هم احتمال نیستند و عملکرد معیوبی داشته‌اند. درویشی (Darvishi, 2005) وضعیت ارزشیابی پیشرفت تحصیلی دانشجویان دانشگاه بیرجند را بررسی نمود. نتایج نشان داد که بین دیدگاه دانشجویان در خصوص ارزشیابی قبل، ضمن و پس از تدریس (آغازین، تکوینی و پایانی) تفاوت معناداری وجود دارد. هم‌چنین ۴۹٪ دانشجویان اظهار نمودند که اهداف ارزشیابی آغازین به اطلاع آنان رسیده نمی‌شود. ۶۷٪ دانشجویان ابراز نمودند که ارزشیابی تکوینی تا حد متوسط

انجام می‌شود و ۱۸٪ دانشجویان ابراز نمودند که اهداف ارزشیابی تکوینی به اطلاع آنها رسانده نمی‌شود. در ارزیابی پایانی اغلب دانشجویان و استادان آزمون تشریحی را تأیید کردند. حافظی کنشگری (Hafezi Koneshgari, 1996) پژوهشی با عنوان «بررسی دیدگاه‌های استادان و دانشجویان دانشگاه شیراز در زمینه شیوه‌های ارزشیابی از آموخته‌های دانشجویان» انجام داد. پرسش‌نامه حول محورهای عامل بازخورد به دانشجو، بازخورد به استاد، آگاه نمودن دانشجو از منبع امتحانی و هدف‌های آموزشی، و نحوه تنظیم سؤال‌های امتحانی، تنظیم شده بود. نتایج نشان داد که مدرسان در دانشکده‌های مختلف تنها در یک مورد (آگاه نمودن دانشجو از منابع و اهداف آموزشی) اختلاف معنی‌دار نشان داده‌اند و در بقیه موارد اختلافی مشاهده نشد.

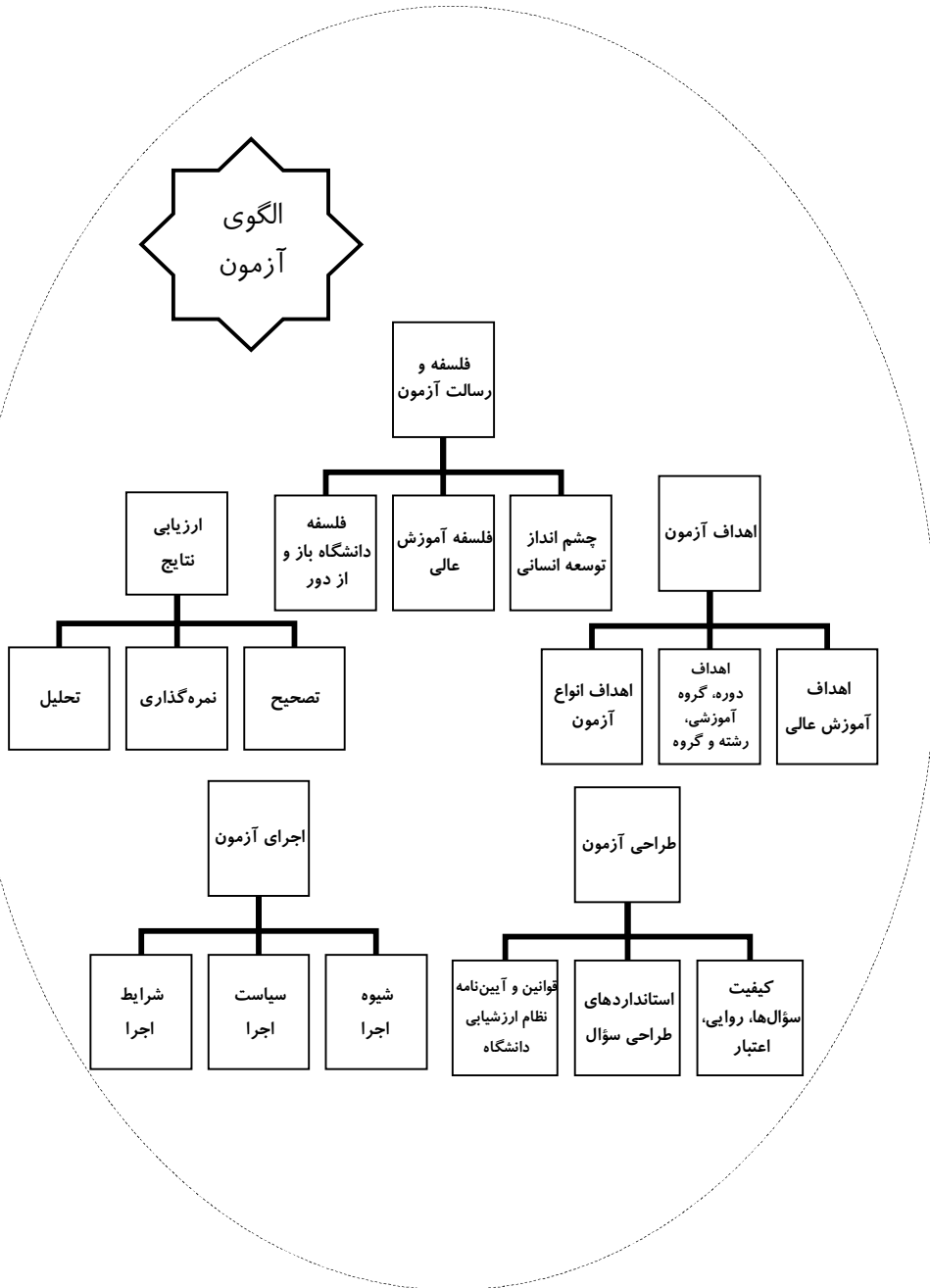
رضایی (Rezaei, 2001) روند ارزشیابی دانشجویان دانشگاه علوم پزشکی زاهدان را توسط استادان و اعضای هیأت علمی علوم پایه بررسی و به این نتایج دست یافت: ۲۷ درصد از استادان از ارزشیابی تکوینی و ۷۲ درصد از ارزشیابی مرحله‌ای استفاده می‌کردند. ایزدی فیروزآبادی (Eizadi Firouzabadi, 2003) پژوهشی با عنوان «تهیه نرم‌افزاری جهت ارزشیابی کلینیکی دانشجویان در بخش ارتدنسی دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید صدوقی یزد» انجام داد و نرم‌افزار کامپیوتری طراحی کرد که امکان خودارزیابی و گروه‌ارزیابی (کلاس‌ارزیابی) با توانایی ضریب دشواری و تمیز و اعلام نتیجه و توانایی تعیین حیطه‌های شناختی هر سؤال آزمون ارتدنسی را داشت. این نرم‌افزار روی ۳۸ دانشجویی که با روش نمونه‌گیری آسان انتخاب شدند، برای آزمون پایان ترم در کنار آزمون کاغذی به اجرا درآمد و پس از آزمون پرسش‌نامه‌ای توسط تمامی دانشجویان تکمیل شد. میانگین نمرات محاسبه شد و نمرات شیوه کامپیوتری بیشتر از کاغذی بود. نظرسنجی نشان داد که پرسش‌نامه با نرخ ۰/۶۶ پایا بوده و دانشجویان از روش کامپیوتری راضی بودند. اسدزاده دهرایی (Asadzadeh Dahraei, 1993) تحقیقی با عنوان «بررسی دیدگاه دانشجویان درباره چگونگی ارزشیابی استادان از آموخته‌های آنان» انجام داد. نتایج تحقیق در خصوص میزان استفاده استادان از انواع ارزشیابی نشان داد که آزمون‌های تشریحی و چندگزینه‌ای معمولی‌ترین ابزار ارزشیابی هستند. بررسی مطابقت سؤال‌ها با آموخته‌ها، نشان داد که ملاک مطابقت سؤال با آموخته‌ها طرح سؤال در سطح دانش، درک و تجزیه و تحلیل می‌باشد. در قسمت دیگر نتایج نشان داد که استادان از نتایج آزمون برای تشخیص و رفع مشکلات یادگیری دانشجویان استفاده نمی‌کنند و آزمون‌ها برای نمره دادن به دانشجویان

کاربرد دارد. در ضمن سؤال‌های تشریحی نمی‌تواند تمامی اهداف آموزشی را پوشش دهد. دانشگاه شیراز (Shiraz University, 1991) در طرح پژوهشی با عنوان «ارزش امتحانات تشریحی در پیش‌بینی موفقیت دانشجویی» نشان داد که آزمون‌های تشریحی، همانند آزمون‌های چهار گزینه‌ای، قدرت پیش‌بینی چندان‌ی ندارند، ولی عملکرد متوسطه در مقایسه با هریک از دو نوع آزمون از قدرت پیش‌بینی بالایی برخوردار می‌باشد.

با توجه به مسایل فوق، این سؤال مطرح می‌گردد که چه الگویی برای آزمون‌های دانشگاه باز و از دور مناسب می‌باشد؟ بر اساس چارچوب نظری مطالعه شده، الگوی انتخابی ترکیبی از عناصر الگوهای سنتی و آموزش باز و از دور را در نظر گرفته و عنصر جدید فلسفه را به عناصر الگوهای قبلی اضافه نموده است.

۱. عنصر فلسفه با سه شاخص: چشم‌اندازهای توسعه انسانی، فلسفه آموزش و دانشگاه باز و از دور، فلسفه تربیتی و آموزش عالی.
 ۲. عنصر هدف با سه شاخص: الف. اهداف دانشگاه باز و از دور؛ ب. اهداف دوره، گروه آموزشی، رشته، گرایش، درس؛ ج. اهداف آزمون.
 ۳. عنصر طراحی با سه شاخص: الف. قوانین، مقررات و آیین‌نامه سیستم ارزشیابی دانشگاه؛ ب. استانداردهای طراحی سؤال‌های آزمون (حیطه و سطوح مختلف اهداف و انتظارات، انواع سؤال‌های عینی، غیر عینی و عملکردی، بودجه‌بندی، سطح ذهنی مخاطبان، فن‌آوری طراحی سؤال، جدول مشخصات آزمون، ملاک مطلق و هنجاری، راهنما و دستورالعمل پاسخ‌گویی، تناسب درصد نمرات میان ترم و پایان ترم؛ و ج. کیفیت سؤال‌های آزمون.
 ۴. عنصر اجرا با سه شاخص: الف. سیاست اجرا (متمرکز یا غیرمتمرکز)؛ ب. شیوه اجرا (حضور، غیرحضور، الکترونیکی، مکتوب)؛ و ج. شرایط و سلامت آزمون.
 ۵. عنصر ارزیابی (تصحیح، تحلیل و بازخورد) با سه شاخص: الف. تصحیح (الکترونیکی، دستی)؛ ب. نمره‌گذاری (کلی، تحلیلی)؛ و ج. تحلیل و بازخورد می‌باشد.
- نمای گرافیکی الگو در شکل ۱، ارائه گردیده است. بنابراین، این تحقیق درصدد است تا به سؤال زیر پاسخ دهد:

به چه میزان عناصر و شاخص‌های الگوی پیشنهادی سنجش تحصیلی دانشگاه باز و از دور از اعتبار برخوردار می‌باشند؟



شکل ۱. عناصر اصلی و شاخص‌های سه‌گانه الگوی پیشنهادی آزمون دانشگاه باز و از دور

روش

تحقیق از نظر هدف کاربردی و از نظر روش آمیخته (کمی و کیفی) است. تکنیک جمع‌آوری اطلاعات جهت تعیین درصد توافق افراد پاسخ‌گو با عناصر الگوی طراحی شده توصیفی است. جامعه آماری تحقیق را ۱۲۹۹ عضو هیأت علمی دانشگاه پیام‌نور تشکیل می‌دهند که در سال ۹۳-۱۳۹۲ در طراحی سؤال با مرکز آزمون دانشگاه همکاری داشتند. نمونه با روش تصادفی انتخاب و حجم آن طبق جدول کرجسی و مورگان در قسمت کمی ۲۸۶ نفر می‌باشند. نمونه بر حسب جنسیت شامل ۸۸ زن و ۱۹۲ مرد (۶ نفر از نظر جنسیت نامعلوم) و از لحاظ مدرک تحصیلی ۱۶۲ نفر دکتری و ۱۱۳ نفر کارشناسی ارشد (۹ نفر نامعلوم) بود. در قسمت کیفی نیز با روش گلوله برفی، هفت نفر متخصص طراحی سؤال دانشگاه پیام‌نور (تماماً مرد با مدرک دکتری علوم تربیتی و سنجش و اندازه‌گیری و سابقه کار بیش از ۲۰ سال) انتخاب شدند.

ابزار جمع‌آوری اطلاعات در بخش کمی پرسش‌نامه محقق ساخته بود که در چهار بخش تنظیم شد. بخش اول درخواست همکاری محقق و ارایه هدف تحقیق (۱ صفحه)، بخش دوم مطالب علمی برای رفع ابهام و یکسان‌سازی درک گویه‌ها (۱۴ صفحه)، بخش سوم اطلاعات جمعیت شناختی شامل جنسیت، مدرک، سابقه کار و بخش چهارم گویه‌های پرسش‌نامه در دو ستون (وضعیت موجود و مطلوب) تنظیم گردید. پرسش‌نامه از نوع بسته پاسخ با مقیاس چند درجه‌ای لیکرت از خیلی کم (۱ امتیاز) تا خیلی زیاد (۵ امتیاز) می‌باشد که در ابتدا بر اساس ادبیات ارزشیابی و آموزش از دور از ۶۰ گویه برخوردار بود و پس از بررسی استادان و تنی چند از صاحب‌نظران، به ۵۸ گویه تقلیل یافت و روایی محتوایی آن تأیید شد. ضریب پایایی پرسش‌نامه (ضریب آلفای کراباخ) معادل ۰/۹۶۴ به دست آمد. شایان ذکر است که میانگین ملاک برای هر شاخص، در دامنه ۱ الی ۵ معادل ۳/۶ (دهک ششم یا ۶۰ درصد) در نظر گرفته شده است. طریق محاسبه آن $۳/۶ = ۶۰ \times ۰/۶$ ، $۱+۵=۶$ می‌باشد. میانگین نظری برای هر مؤلفه (مجموعه سه شاخص) در دامنه ۳ الی ۱۵ معادل ۹ (دهک پنجم یا ۵۰ درصد) در نظر گرفته شد که به همان طریق محاسبه گردیده است.

در بخش کیفی از ابزار مصاحبه نیمه ساختارمند استفاده شد. در ابتدا چهارچوبی برای سؤال‌های مصاحبه تعیین گردید تا انعطاف‌پذیری در سؤال‌ها حفظ گردد و بعد از هر پاسخ، از مصاحبه‌شونده چرایی پاسخ برای توضیح بیشتر درخواست گردید.

داده‌های جمع آوری شده با نرم افزار SPSS مورد تحلیل توصیفی (فراوانی، فراوانی تجمعی و شاخص‌های مرکزی مانند میانگین و انحراف معیار و واریانس) و استنباطی (آزمون t تک گروهی) قرار گرفتند.

یافته‌ها

سؤال تحقیق: به چه میزان عناصر و شاخص‌های الگوی پیشنهادی سنجش تحصیلی باز و از دور از اعتبار برخوردار می‌باشند؟
برای تحلیل سؤال تحقیق، تمامی پنج عنصر و زیرشاخه‌های آنها به تفکیک بررسی توصیفی و استنباطی شدند.

الف. شاخص‌های عنصر فلسفه: فراوانی شاخص‌های عنصر فلسفه در سه سطح به همراه میانگین، انحراف معیار و نتایج آزمون t در جدول ۲ درج گردیده است.

جدول ۲. فراوانی، میانگین، انحراف معیار و نتایج آزمون t شاخص‌های عنصر فلسفه

شاخص‌های فلسفه	فراوانی (درصد) سطوح			میانگین	انحراف استاندارد	t	درجه آزادی	سطح معنی‌داری دوطرفه	تفاوت میانگین‌ها
	زیاد و خیلی زیاد	متوسط	خیلی کم و کم						
چشم‌انداز توسعه نیروی انسانی	۲۵۲ (۸۸/۱)	۲۳ (۸)	۱۱ (۳/۸)	۴/۵۶	۰/۰۴۵۸	-۲۷/۹۳	۲۸۵	۰/۰۰۰**	۰/۹۶۰۰
فلسفه آموزش عالی	۲۵۰ (۸۷/۱)	۳۰ (۱۰/۵)	۶ (۲/۱)	۴/۳۰	۰/۴۱۰۳	-۳۰/۷۵	۲۸۵	۰/۰۰۰**	۰/۷۴۷۲
فلسفه دانشگاه باز و از دور	۲۴۱ (۸۴/۳)	۳۹ (۱۳/۶)	۶ (۲/۱)	۴/۶۳	۰/۴۳۴۹	-۳۰/۲۶	۲۸۵	۰/۰۰۰**	۱/۰۳۰۰

* معناداری در سطح ۰/۰۵ ** معناداری در سطح ۰/۰۱

داده‌های جدول ۲ نشان می‌دهد که سه شاخص عنصر فلسفه از فراوانی‌های سطوح خیلی زیاد و زیاد بالاتر از ۸۴ درصد برخوردارند و میانگین آنها بالاتر از میانگین ملاک (۳/۶) می‌باشد. برای اطمینان از معنی‌دار بودن میانگین‌ها آزمون t تک گروهی انجام شد و نتایج نشان می‌دهد که خطای

آزمون از ۰/۰۵ کمتر می‌باشد. از دیدگاه اعضای هیأت علمی دانشگاه پیام‌نور سه شاخص فلسفه الگوی آزمون دانشگاه باز و از دور با ۹۹ درصد از اعتبار معنی‌دار در حد بالاتر از ملاک برخوردارند.

ب. شاخص‌های عنصر هدف: فراوانی شاخص‌های عنصر هدف در سه سطح به همراه میانگین، انحراف معیار و نتایج آزمون t در جدول ۳ درج گردیده است.

جدول ۳. فراوانی، میانگین، انحراف معیار و نتایج آزمون t شاخص‌های عنصر هدف

تفاوت میانگین‌ها	درجه معنی‌داری دوطرفه	درجه آزادی	t	انحراف استاندارد	میانگین	فراوانی (درصد) سطوح مختلف شاخص‌ها			شاخص‌های هدف
						خیلی کم و کم	متوسط	زیاد و خیلی زیاد	
۰/۶۹۷۲	۰/۰۰۰ ^{°°}	۲۸۵	۱۳/۸۶	۰/۸۵۰۸	۴/۲۹	۱۷ (۵/۹)	۳۰ (۱۰/۵)	۲۳۹ (۸۳/۶)	آموزش عالی و دانشگاه
۰/۶۸۸۴	۰/۰۰۰ ^{°°}	۲۸۵	۱۱/۱۱	۱/۰۴۷۸	۴/۲۸	۱۸ (۶/۱)	۵۱ (۱۷/۷)	۲۱۷ (۷۵/۸)	دوره، گروه، رشته گرایش و درس
۰/۷۸۴۶	۰/۰۰۰ ^{°°}	۲۸۵	۱۷/۹۵	۰/۷۳۹۱	۴/۳۸	۴ (۱/۳)	۱۵ (۵/۲)	۲۶۷ (۹۳/۴)	آزمون

^{°°}معناداری در سطح ۰/۰۵ [°]معناداری در سطح ۰/۰۱

داده‌های جدول ۳ نشان می‌دهد که سه شاخص عنصر هدف از فراوانی‌های سطوح خیلی زیاد و زیاد بالاتر از ۷۵ درصد برخوردارند و میانگین آنها بالاتر از میانگین ملاک (۳/۶) می‌باشد. برای اطمینان از معنی‌دار بودن این میانگین‌ها، آزمون t تک گروهی انجام شد و نتایج نشان داد که خطای آزمون از ۰/۰۵ کمتر است. از دیدگاه اعضای هیأت علمی دانشگاه پیام‌نور سه شاخص هدف الگوی آزمون دانشگاه باز و از دور با ۹۹ درصد اطمینان از اعتبار معنی‌دار بالاتر از ملاک برخوردارند.

ج. شاخص‌های عنصر طراحی: فراوانی شاخص‌های عنصر طراحی در سه سطح به همراه میانگین، انحراف معیار و نتایج آزمون t در جدول ۴ درج گردیده است.

جدول ۴. فراوانی، میانگین، انحراف معیار و نتایج آزمون t شاخص‌های عنصر طراحی

تفاوت میانگین‌ها	درجه معنی‌داری دوطرفه	درجه آزادی	t	انحراف استاندارد	میانگین	فراوانی (درصد) سطوح مختلف شاخص‌ها			شاخص‌های طراحی
						زیاد و خیلی زیاد	متوسط	خیلی کم و کم	
۰/۴۶۰۶	۰/۰۰۰ ^{**}	۲۸۵	۹/۵۷۷	۰/۸۱۳۲	۴/۰۶	۱۹۷ (۶۸/۹)	۶۴ (۲۲/۳)	۲۵ (۸/۶)	قوانین نظام ارزشیابی
۰/۵۵۹۹	۰/۰۰۰ ^{**}	۲۸۵	۱۲/۰۵۵	۰/۷۸۵۵	۴/۱۶	۲۱۳ (۷۴/۵)	۶۶ (۲۳)	۷ (۲/۴)	کیفیت، روایی و اعتبار آزمون
۰/۰۰۲۵	۰/۹۴۸	۲۸۵	۰/۰۶۵	۰/۶۶۱۵	۳/۶۰	۷۹ (۲۷/۶)	۱۶۷ (۵۸/۱)	۴۰ (۱۳/۷)	استانداردهای طراحی سؤال

* معناداری در سطح ۰/۰۵ ** معناداری در سطح ۰/۰۱

داده‌های جدول ۴ نشان می‌دهد دو شاخص عنصر طراحی از فراوانی‌های سطوح خیلی زیاد و زیاد بالاتر از ۶۸/۹ درصد و یک شاخص از فراوانی بالاتر از ۲۷/۶ در همان سطح برخوردارند و میانگین آنها بالاتر از میانگین ملاک (۳/۶) می‌باشد. برای اطمینان از معنی‌دار بودن این میانگین‌ها آزمون t تک‌گروهی انجام شد. نتایج نشان می‌دهد که در دو شاخص اول خطای آزمون از ۰/۰۵ کمتر می‌باشد؛ اما در شاخص سوم معادل ۰/۹۴۸ می‌باشد. از دیدگاه اعضای هیأت علمی دانشگاه پیام‌نور دو شاخص طراحی الگوی آزمون دانشگاه باز و از دور با ۹۹ درصد اطمینان از اعتبار معنی‌دار بالاتر از ملاک برخوردارند و شاخص استانداردهای طراحی سؤال با میانگین در حد ملاک معنی‌دار نمی‌باشد.

د. شاخص‌های عنصر اجرا: فراوانی شاخص‌های عنصر اجرا در سه سطح به همراه میانگین، انحراف معیار و نتایج آزمون t در جدول ۵ درج گردیده است.

جدول ۵. فراوانی، میانگین، انحراف معیار و نتایج آزمون t شاخص‌های عنصر اجرا

تفاوت میانگین‌ها	درجه معنی‌داری دوطرفه	درجه آزادی	t	انحراف استاندارد	میانگین	فراوانی (درصد) سطوح مختلف شاخص‌ها			شاخص‌های اجرا
						زیاد و خیلی زیاد	متوسط	خیلی کم و کم	
-۰/۵۰۲	۰/۰۰۰ [°]	۲۸۵	-۱۰/۰۹۳	۰/۸۴۱۲	۳/۰۹	۴۱ (۱۴/۳)	۱۴۱ (۴۹/۲)	۱۰۴ (۳۶/۲)	سیاست اجرا
-۰/۸۶۵۱	۰/۰۱۲ [°]	۲۸۵	-۲/۵۱۷	۰/۵۸۱۱	۳/۵۱	۵۸ (۲۰/۲)	۱۹۲ (۶۷)	۳۶ (۱۲/۵)	شیوه اجرا
۰/۸۹۴۷	۰/۰۰۰ [°]	۲۸۵	۱۷/۶۹۵	۰/۸۵۶۸	۴/۴۹	۲۵۶ (۸۹/۴)	۱۰ (۳/۴)	۲۰ (۶/۹)	شرایط اجرا

* معناداری در سطح ۰/۰۵ ** معناداری در سطح ۰/۰۱

داده‌های جدول ۵ نشان می‌دهد که سه شاخص عنصر اجرا از فراوانی‌های متفاوت در سطوح مختلف برخوردارند. برای اطمینان از معنی‌دار بودن تفاوت این میانگین‌ها با میانگین ملاک، آزمون t تک‌گروهی انجام شد و نتایج نشان می‌دهد که خطای آزمون از ۰/۰۵ کمتر می‌باشد. دو شاخص سیاست و شیوه آزمون دارای میانگین بالاتر از ۳ است. در شاخص شرایط اجرا میانگین بالاتر از ۴ است. از دیدگاه اعضای هیأت علمی دانشگاه پیام‌نور شاخص اول و سوم عنصر اجرای الگوی آزمون دانشگاه باز و از دور با ۹۹ درصد اطمینان از اعتبار معنی‌دار به ترتیب کمتر و بیشتر از ملاک و شاخص دوم از اعتبار معنی‌دار در سطح ۰/۰۵ و کمتر از ملاک برخوردار است.

هـ شاخص‌های عنصر ارزیابی (تصحیح، تحلیل و بازخورد): فراوانی شاخص‌های عنصر ارزیابی (تصحیح، تحلیل و بازخورد) در سه سطح به همراه میانگین، انحراف معیار و نتایج آزمون t در جدول ۶ درج گردیده است.

جدول ۶. فراوانی، میانگین، انحراف معیار و نتایج آزمون t شاخص‌های عنصر ارزیابی

تفاوت میانگین‌ها	درجه معنی‌داری دو طرفه	درجه آزادی	t	انحراف استاندارد	میانگین	فراوانی (درصد) سطوح			شاخص‌های ارزیابی
						مختلف شاخص‌ها			
						زیاد و خیلی زیاد	متوسط	خیلی کم و کم	
۰/۴۵۸۹	۰/۰۰۰ ^{oo}	۲۸۵	-۹/۶۳۱	۰/۸۰۵۸	۳/۱۴	۴۷ (۱۶/۴)	۱۵۱ (۵۲/۸)	۸۸ (۳۰/۷)	تصحیح
۰/۳۹۶۰	۰/۰۰۰ ^{oo}	۲۸۵	-۷/۱۵۱	۰/۹۳۶۵	۳/۲۰	۶۹ (۲۴/۱)	۱۲۹ (۴۵/۱)	۸۸ (۳۰/۷)	نمره‌گذاری
۰/۷۴۶۱	۰/۰۰۰ ^{oo}	۲۸۵	۱۶/۹۸۸	۰/۷۴۲۷	۴/۳۴	۲۳۸ (۸۳/۲)	۴۴ (۱۵/۳)	۴ (۱/۳)	تحلیل و بازخورد
						* معناداری در سطح ۰/۰۵			
						** معناداری در سطح ۰/۰۱			

داده‌های جدول ۶ نشان می‌دهد که سه شاخص عنصر ارزیابی نتایج از فراوانی‌های سطوح مختلف برخوردارند. برای اطمینان از معنی‌دار بودن تفاوت این میانگین‌ها با میانگین ملاک، آزمون t تک‌گروهی انجام شد و نتایج نشان می‌دهد که خطای آزمون از ۰/۰۵ کمتر می‌باشد. دو شاخص تصحیح و نمره‌گذاری دارای میانگین کمی بیشتر از ۳ و شاخص تحلیل و بازخورد دارای میانگین بالاتر از ۴ است. از دیدگاه اعضای هیأت علمی دانشگاه پیام‌نور دو شاخص تصحیح و نمره‌گذاری الگوی آزمون دانشگاه باز و از دور با ۹۹ درصد اطمینان از اعتبار معنی‌دار نزدیک به ملاک و شاخص تحلیل و بازخورد از اعتبار معنی‌دار بالاتر از ملاک برخوردارند.

و. اعتبار عناصر الگوی آزمون دانشگاه باز و از دور: جدول ۷ آماره‌های توصیفی عناصر پنج‌گانه الگوی آزمون دانشگاه باز و از دور را در قیاس با ملاک ۹ معادل ۶۰٪ نشان می‌دهد.

جدول ۷. آماره‌های توصیفی و استنباطی محورهای پنج‌گانه آزمون دانشگاه‌های باز و از دور

عناصر الگوی آزمون	میانگین	انحراف استاندارد	t	درجه آزادی	سطح معنی‌داری	تفاوت میانگین‌ها
فلسفه آزمون	۸/۵۱	۰/۹۸۴۳	-۸/۳۰۴۳	۲۸۵	۰/۰۰۰ ^{***}	-۰/۴۸۴۲
اهداف آزمون	۱۲/۹۷	۲/۱۸۳۵	۳۰/۷۴۸۰	۲۸۵	۰/۰۰۰ ^{***}	۳/۹۷۰۲
طراحی آزمون	۱۱/۸۲	۱/۹۲۰۷	۲۴/۸۵۶۸	۲۸۵	۰/۰۰۰ ^{***}	۲/۸۲۳۱
اجرای آزمون	۱۱/۱۰	۱/۷۶۷۸	۲۰/۱۴۷۶	۲۸۵	۰/۰۰۰ ^{***}	۲/۱۰۶۱
ارزیابی	۱۰/۶۹	۱/۸۳۰۶	۱۵/۶۲۲۸	۲۸۵	۰/۰۰۰ ^{***}	۱/۶۹۱۱

* معناداری در سطح ۰/۰۵ ** معناداری در سطح ۰/۰۱

داده‌های جدول ۷ نشان می‌دهد که عنصر فلسفه دارای میانگینی نزدیک به ملاک نظری و عناصر هدف، طراحی، اجرا، تصحیح و تحلیل دارای میانگینی بالاتر از ۹ و ملاک ۶۰٪ می‌باشند. آزمون t تک‌گروهی انجام شد. از دیدگاه اعضای هیأت علمی دانشگاه پیام‌نور، عناصر فلسفه، اهداف، طراحی، اجرا، تصحیح و تحلیل آزمون‌های دانشگاه باز و از دور از اعتبار معنی‌داری در سطح ۹۹٪ برخوردارند.

برای اطمینان از نتایج و تأیید اعتبار الگو، مصاحبه نیمه‌ساختارمندی با هفت نفر متخصص سنجش و اندازه‌گیری و طراح سؤالات آزمون‌های دانشگاه پیام‌نور انجام شد. گزارش این مصاحبه نشان می‌دهد که اولاً کلیه عناصر و شاخص‌های الگوی پیشنهادی آزمون‌های دانشگاه باز و از دور مورد تأیید است و ثانیاً دلایل نسبتاً یکسانی برای طراحی آزمون‌های دانشگاه به شیوه سنتی و بر خلاف ماهیت آموزش از دور جمع‌آوری شد که طبقه‌بندی آنها در جدول ۸ ارائه گردیده است.

جدول ۸. نتایج مصاحبه نیمه ساختارمند

فراوانی	نکات مورد اشاره
۵	کاربرد اصول، قواعد و نگرش نظام سنتی آزمون‌سازی در نظام نوین آموزش از دور
۷	عدم توجه به فلسفه تربیتی و چشم‌انداز توسعه انسانی برای ترسیم اهداف کلی آموزش
۵	عدم توجه به فلسفه آموزش از دور و ویژگی‌های متقاضیان بزرگسال
۷	عدم توجه به ویژگی‌های منحصر به فرد آموزش باز و از دور
۷	شناخت ناکافی مدرسان به نقش خود در فرآیند آزمون‌سازی تکوینی
۷	برقراری قوانین سنتی در فرآیند طراحی سؤال‌های آزمون
۴	متمرکز بودن شیوه اجرای آزمون
۷	خودآموز نبودن کتب آموزشی
۳	شیوه‌های مختلف آموزش
۲	وجود معیارهای ملی کمی، عددی و نسبی در فرآیند ارزیابی
۵	ناهماهنگی نمرات میان‌ترم و پایان ترم
۲	عدم مهارت کادر اجرایی در اجرای آزمون‌های پایانی
۵	عدم رعایت استانداردها در طراحی سؤال

بحث و نتیجه‌گیری

برای اعتباریابی عناصر الگوی آزمون‌های دانشگاه باز و از دور، از اعضای هیأت علمی دانشگاه پیام‌نور که با مرکز آزمون همکاری داشتند، نظرخواهی شد.

عنصر اول «فلسفه» می‌باشد. پاسخ‌دهنده‌ها وجود عنصر فلسفه را در الگوی آزمون تأیید و اعتبار معنی‌داری را اظهار می‌کنند. در مصاحبه کیفی نیز متخصصان فن اشاره به بنیادی بودن این عنصر در الگوی آزمون نموده‌اند و علل کمی میانگین مشاهده شده نسبت به میانگین نظری را عدم شناخت کافی محیط آموزش باز و از دور و احساس بی‌تأثیری فعالیت مدرسان در یادگیری و ارزشیابی پایانی اظهار نموده‌اند. می‌توان نتیجه گرفت که با توجه به اظهار نظر آگاهانه ولی محافظه‌کارانه پاسخ‌دهنده‌ها و نکات طرح شده در مصاحبه‌ها، عنصر فلسفه جایگاه خود را در الگوی آزمون دانشگاه باز و از دور در مقام اول و پایه حفظ نموده است. متأسفانه در پیشینه و ادبیات پژوهشی هم‌سو با این نتیجه، تحقیقی یافت نشده است.

دومین عنصر «هدف» می‌باشد. بر پاسخ‌دهنده‌ها آشکار بوده است که آزمون‌های دانشگاه باز و از دور باید از اهداف آموزش عالی و اهداف دانشگاه پیام‌نور تبعیت کند و ویژگی‌های ماهوی نظام دانشگاهی و نظام آموزش از دور را در نظر بگیرد. در فرآیند آزمون‌سازی نگاه به شاخه تحصیلی، گرایش تحصیلی، رشته تحصیلی و در نهایت دروس تحصیلی در آزمون لازم است. صرفاً اهداف آموزشی کلی و دقیق رفتاری نباید به عنوان اهداف اصلی در نظر گرفته شوند، بلکه، آن اهدافی از محتوای دروس که گرایش و رشته تحصیلی را هم پوشش دهند ملاک طراحی آزمون باید قرار بگیرد. در مصاحبه با استادان، ارزشیابی دانشگاه پیام‌نور نیز این عنصر با شاخص‌های مورد نظر تأیید گردید. از این رو، عنصر هدف الگوی آزمون‌های دانشگاه باز و از دور به عنوان دومین عنصر در الگو از جایگاه ویژه خود برخوردار است. یافته‌های این قسمت از تحقیق با نتایج تحقیق درویشی (Darvishi, 2005)، حافظی کاشگری (Hafezi Koneshgari, 1996)، و اکهات و همکاران (Eckhout et al., 2005) هم‌سو می‌باشد.

عنصر سوم الگو «طراحی» می‌باشد. میانگین شاخص استانداردهای طراحی معنی‌دار نبوده است. در بررسی علت این امر، به نظر می‌رسد پاسخ‌دهنده‌ها استفاده از انواع سؤال‌های عینی (جورکردنی، صحیح و غلط، چهار گزینه‌ای) و سؤال‌های غیرعینی (کوتاه پاسخ، بلند پاسخ) و آزمون‌های عملکردی، فهرست واریسی را مطلوب اعلام کرده‌اند. به عبارتی، برای طراحی سؤال، انواع عینی و غیرعینی را مناسب می‌دانند. این نوع نگاه به آزمون، نگرش سنتی را تداعی می‌کند که بیانگر فاصله شناختی با ماهیت آموزش از دور و یا سواد آزمون‌سازی است. اگر ماهیت دانشگاه باز و از دور پیام‌نور و سنجش این نظام به خوبی درک می‌شد، شاید این خطا اتفاق نمی‌افتاد. اگر سیاست‌گذاران نمره را ملاک و نتیجه یادگیری در نظر نمی‌گرفتند و یادگیری در حد تسلط به صورت کیفی و نه کمی را ملاک قرار می‌دادند، شاید این پراکنندگی نظر به دست نمی‌آمد. اغلب پاسخ‌دهنده‌ها اعتقاد دارند که آزمون هم ملاکی و هم رقابتی برگزار گردد. ادبیات علوم تربیتی ماهیت آموزش از دور را رقابتی نمی‌داند. ماهیت یادگیری از دور نیز ملاکی و در حد تسلط گزارش شده است. چطور می‌توان برای این ماهیت آزمون‌های هنجاری طرح کرد؟ نتایج تحقیق پگی (Peggy, 2005) نشان داد که معلمان که دوره آزمون‌سازی و سنجش را گذرانده‌اند در آزمون‌سازی و سنجش فعالیت‌های یادگیری موفق‌تر عمل کرده‌اند. نتایج تحقیق ملی ایمپرا و همکاران (Impara et al., 1993) نشان داده است که معلمان دانش‌اندکی از سنجش دارند و از

بهره‌برداری از آن، دانش بسیار کمتری دارند. شپارد (Shepard, 2001) ادعا می‌کند که سنجش در میان ترم باید یادگیری را تقویت کند و مشکل وقتی مشاهده می‌شود که طراحان آزمون از سواد و دانش کافی نسبت به سنجش برخوردار نباشند و مهارت کافی در آزمون‌سازی استاندارد نداشته باشند. مک میلان (McMillan, 2001) عامل شکست سنجش و هدف‌مند نبودن آن را در عدم درک و فهم طراحان سؤال از سنجش یادگیری می‌داند. اگر ماهیت سنجش و آزمون در آموزش باز و از دور شناخته می‌شود، شاید این خطا مشاهده نمی‌شد. اگر الگوی مشخصی برای آزمون‌سازی در اختیار سازندگان و طراحان سؤال قرار می‌گرفت این تمایز در پاسخ‌دهی اتفاق نمی‌افتاد و ناخودآگاه نگاه سنتی بر پاسخ‌ها حاکم نمی‌گشت. در مصاحبه با استادان ارزشیابی نیز مشکلات ذکر شده، مورد تأیید قرار گرفت. در نهایت، عنصر طراحی با شاخص‌های آن مورد تأیید نهایی (کمی و کیفی) قرار گرفته است و جایگاه خود را در الگوی آزمون حفظ نموده است. نتایج این تحقیق با قسمتی از یافته‌های تحقیق رضایی (Rezaei, 2001)، یونسی (Younesi, 2005)، طرح پژوهشی دانشگاه شیراز (Shiraz University, 1991) و اسدزاده دهراعی (Asadzadeh Dahraei, 1993) هم‌سو می‌باشد.

عنصر چهارم الگو «اجرا» می‌باشد. اجرا به عنوان فعالیت عملی هر الگویی بعد از طراحی شکل می‌گیرد. دانشگاه باز و از دور پیام‌نور با وسعت متقاضیان و پراکندگی جغرافیایی زیاد از مراکز متعددی برخوردار است که رعایت عدالت در آموزش و سنجش آن نیازمند برگزاری متمرکز آزمون می‌باشد. پاسخ‌دهنده‌ها تا حدودی خواهان سیاست اجرای نیمه‌متمرکز و غیرمتمرکز می‌باشند. وجود دروس عملی، تنوع سلیقه و مهارت در آموزش دروس یکسان، عدم آگاهی از ماهیت نظام آموزش از دور، عدم سنجش و طراحی استاندارد سؤال، عدم تناسب بین نمرات میان ترم و پایان ترم، عدم سرویس‌دهی یکسان ملی در زمان اجرای آموزش و آزمون، شرایط مختلف حاکم بر فضای فیزیکی متفاوت در زمان اجرا، عدم دسترسی به موقع به آیین‌نامه‌های نظام ارزشیابی، عدم دسترسی یکسان زمانی به کتب درسی، نگرش سنتی نسبت به ارزشیابی و آزمون‌سازی و آموزش همگی در اظهارنظر پاسخ‌دهنده‌ها مؤثر می‌باشند. در خصوص شاخص شیوه اجرا هر دو نوع اجرای آزمون به صورت حضوری و غیرحضوری در یک سطح انتخاب شده‌اند. پاسخ‌دهنده‌ها توانسته‌اند از نگاه سنتی فرار کنند و در عین حال اجرای الکترونیکی آزمون را نیز خواهان می‌باشند و موضع محافظه‌کارانه اتخاذ می‌کنند. در مصاحبه با صاحب‌نظران آموزش

باز و از دور نیز عنصر اجرای آزمون با شاخص‌های مورد نظر صریحاً مورد قبول و تأیید قرار گرفت. بنابراین، عنصر اجرا با شاخص‌های سه‌گانه، جایگاه خود را در الگوی آزمون حفظ نموده است. این نتیجه با یافته‌های تحقیق ایزدی فیروزآبادی (Eizadi Firouzabadi, 2003) در خصوص اجرای الکترونیکی و اسدزاده دهرایی (Asadzadeh Dahraei, 1993) در خصوص استفاده از سؤال تشریحی و چندگزینه‌ای در آزمون تکوینی و پایانی هم‌سو می‌باشد.

پنجمین عنصر «ارزیابی نتایج» است. اگر دقت آماری لحاظ گردد، تعدادی پاسخ مشترک در تصحیح دستی و الکترونیکی وجود خواهد داشت. به عبارتی، پاسخ‌دهنده هم نگاه سنتی را حفظ نموده و هم خود را با نظام آموزش از دور همانندسازی می‌کند. ۴۷/۹٪ پاسخ‌دهنده‌ها خواهان نمره‌گذاری به روش کلی در سؤال‌های آزمون می‌باشند. نمره‌گذاری به روش کلی معمول دروس تشریحی می‌باشد که به هیچ وجه عینیت کافی در نمره‌گذاری را نشان نمی‌دهد. پاسخ‌دهنده‌ها نقش خود را در آزمون، با این گونه پاسخ‌دادن آشکار می‌بینند که مغایر با نظام باز و از دور می‌باشد. در این نظام، مدرس فقط نقش راهنما و حل مشکلات یادگیری را دارد و ارزشیابی پایان ترم بر عهده او نمی‌باشد. ۹۸٪ پاسخ‌دهنده‌ها انتظار دارند که نتایج آزمون تحلیل گردد و ۹۸/۷٪ خواهان ارایه بازخورد می‌باشند. تحلیل و بازخورد نتایج نقطه پایانی فعالیت‌های آزمون‌سازی به قصد بازسازی مجدد فعالیت یاددهی - یادگیری نظام دانشگاهی است که به فراخور مخاطب خود صورت می‌گیرد. بنابراین، عنصر ارزیابی نتایج آزمون در الگوی آزمون دانشگاه باز و از دور ماهیت وجودی خود را حفظ می‌کند. نتایج این تحقیق با یافته‌های درویشی (Darvishi, 2005) هم‌سو می‌باشد.

بحث‌های کتابخانه‌ای ضرورت وجود الگوی آزمون دانشگاه باز و از دور را برجسته ساخت و بحث‌های میدانی، نقش و جایگاه الگوی آزمون‌سازی این نظام را برجسته ساخته است. عناصر این الگو در پنج شاخه اصلی: فلسفه آزمون، هدف آزمون، طراحی آزمون، اجرای آزمون و بازخورد نتایج آزمون خلاصه شده است و هر عنصر نیز سه شاخص را در بر گرفت که رهنمودی برای پیشنهادها و تحقیق می‌باشد.

- تربیت نیروی انسانی مورد نیاز برای تمامی زیرمجموعه‌های نظام دانشگاه باز و از دور.
- برگزاری کارگاه‌های آموزشی برای کارکنان، اعضای هیأت علمی و مدرسان مدعو در زمینه فلسفه نظام آموزش باز و از دور، نظام ارزشیابی و فرآیند یاددهی و یادگیری.

- بازنگری در شیوه‌های آزمون‌سازی، یاددهی، یادگیری، تدوین محتوی و کتب خودآموز.
- استفاده از ملاک کیفی در نمره‌گذاری.
- تهیه بسته‌های آموزشی برای طراحی سؤال.
- برگزاری کارگاه‌های آموزشی استفاده از نرم‌افزارهای آزمون‌سازی.
- جذب استادان با توجه به دانش و آگاهی آنان از نظام آموزش باز و از دور.
- ایجاد بستر مناسب مادی و معنوی برای انجام پژوهش‌هایی در زمینه ارزشیابی.
- اجرای الگوی آزمون دانشگاه باز و از دور به منظور اعتباریابی اجرا.
- استفاده از تجربه شخصی مدرسان در زمینه ارزشیابی پیشرفت تحصیلی.
- اعتباریابی الگوی ارائه شده در سطح وسیع تر.
- حمایت‌های پشتیبانی و مشاوره‌ای به گروه‌های آموزشی برای همکاری در فرآیند آزمون‌سازی.
- ارزیابی جامع از اجرا و پیاده‌سازی الگو و تحلیل نتایج به منظور مهندسی مجدد الگو.

References

1. Asadzadeh Dahraei, H. (1993). *Evaluation of students' views about the manner of educational evaluation of teachers of what they have learned*. Master's Thesis, Allameh Tabatabaiee University, Faculty of Psychology and Educational Sciences. (in Persian).
2. Barnett, R., Parry, G., & Coate, K. (2001). Conceptualizing curriculum change. *Teaching in Higher Education*, 6(4), 435-449.
3. Bloom, B. (1968). *Taxonomy of educational Objectives*. New York: Holt.
4. Brookhart, S. M. (2007). Expanding views about formative classroom assessment: A review of the literature. In J. McMillan (Ed.), *Formative classroom assessment: Theory in to practice* (pp. 29-42). New York: Teachers College.
5. Darvishi, M. (2005). *Evaluation of students' progress Birjand University*, M.A. Thesis, Faculty of Humanities, Birjnd University. (in Persian).
6. Eckhout, T., Davis, S., Mickelson, K., & Goodburn, A. (2005). A method for providing assessment training to in-service and pre-service teachers. *Paper presented at the Annual Meeting of the Southwestern Educational Research Association*, New Orleans, LA.
7. Eizadi Firouzabadi, F. (2004). *Preparing a software for clinical evaluation of students in the department of orthodontics*. Ph.D. Thesis, Shahid Sadoughi University of Medical Sciences and Health Services, Dentistry Faculty. (in Persian).

8. Fegenbaum, A. V. P. (1994). Quality education and mericus competitiveness. *Quality Progress*, 27(9), 83-84.
9. Freeman, R. (1993). *Quality assurance in training and education. How to apply BS 5750 (ISO 9000) Standards*. Kogan Page Limited, London.
10. Hafezi Koneshgari, E. (1996). *Examine the views of Professors and students of Shiraz University about students learned assessment practices, Shiraz University*. Master's Dissertation. Shiraz University, Department of Humanities (in Persian).
11. Henderson, L., & Putt, I. (1999). Evaluating audio-conferencing as an effective learning tool in corsss-cultural contexts. *Open Learning*, 14(1), 25-37.
12. Impara, J. C., Plake, B. S., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, 12(4), 10-12, 39.
13. McAlister, S. (1998). Crtedible or tentative? A model of open university students with 'low' educational qualifications. *Open Learning*, 13(3), 33-42.
14. McCulloch, K. H. (1997). Participatory evaluation in distance learning. *Open Learning*, 12(1), 24-30.
15. McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
16. Merriam, S. B. (2002). *Qualitative research in practice: Examples for discussion and analysis*. San Francisco: Jossey-Bass.
17. Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33(1), 49-64.
18. Messick, S. (1988). The once and future issue of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 33-45). Hillsdale, NJ: Lawrence Erlbaum Associates.
19. Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 13-103). New York: Macmillan.
20. Peggy, C. (2005). Teacher candidates' literacy in assessment. *Academic Exchange Quarterly*, 9(3).
21. Popham, W. J. (1988). *Educational evaluation*. Englewood Cliffs, NJ: Prentic Hall.
22. Provus, M. (1973). *Discrepancy for education program improvement and assessment*. Berkeley: MacCutchan Publishing Corporation.
23. Rezaei, G. (2001). *The evaluation of Zahedan University of Medical Sciences*. Ph.D. Dissertation Zahedan University. (in Persian).
24. Ruhe, V., & Zumbo, B. D. (2009). *Evaluation in distance education and e-learning*. Guilford Press. 27 Spring Street, New York, NY10012. www.guilford.com
25. Scriven, M. (1967). The methodology of Education. In curriculum evaluation. Edited by R. E. Stake. Aeramonograph Series on evaluation, No, I. Checago: Rand McNally.

26. Seif, A. A. (2005). *Educational assessment, measurement, and evaluation* (Third Edition). Douran Publishing Company. (in Persian).
27. Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *Handbook of research on teaching* (4th Ed.) (pp. 1066-1101). Washington, DC: American Educational Research Association.
28. Shiraz University. (1991). *Anatomical examinations value in predicting academic success*. Research Paper.
29. Stake, R. E. (1967). The counterance of educational evaluation. *Teachers College Record*, 68, 523-540.
30. Stufflebeam, D. L. (1971). *Educational evaluation and decision making*. Itasca. IL; Peacock.
31. Taba, H. (1962). *Curriculum development*. New York: Harcourt, Brace & World.
32. Tallent-Runnels, M. K. (2006). Teaching course online: A review of the research. *Review of Educational Research*, 76(1), 93-135.
33. Tyler, R. W. (1969). The objectives and plans for a national assessment of educational progress. *Journal of Educational Measurment*, 3(1), 1-4.
34. Younesi, J. (2002). *Psychometric properties of the test questions in Psychology course PNU*. Master's Thesis, Allameh Tabatabaei University. (in Persian).