

مجله‌ی علوم تربیتی دانشگاه شهید چمران اهواز
بهار و تابستان ۱۳۹۷، دوره‌ی ششم، سال ۲۵
شماره‌ی ۱، صص: ۱۸۴-۱۶۵

تاریخ دریافت مقاله: ۱۳۹۵/۰۶/۲۷
تاریخ بررسی مقاله: ۱۳۹۵/۰۷/۰۸
تاریخ پذیرش مقاله: ۱۳۹۶/۰۱/۲۳

بررسی مقایسه‌ای آزمون‌های مهارتی مبتنی بر فناوری و آزمون‌های مبتنی بر کاغذ (مورد مطالعه: اداره کل آموزش فنی و حرفه‌ای البرز)

حسین باقرپور*

حسین محمودی سفیدکوهی**

چکیده

رشد سریع فناوری ارتباطات و اطلاعات در تدریس و یادگیری، رویکرد اصلی را از سیستم آزمون مبتنی بر قلم و کاغذ به آزمون مبتنی بر فناوری تغییر داده است. رایانه‌ها و فناوری‌های مرتبط ابزارهای قدرتمندی برای رفع چالش‌های ارزشیابی ارائه می‌دهند و ثبت مجموعه‌ی گسترده‌تری از مهارت‌های شناختی را آسان می‌کنند. این مطالعه با هدف بررسی مقایسه‌ای آزمون‌های مهارتی مبتنی بر فناوری و آزمون‌های مبتنی بر کاغذ از دیدگاه متقاضیان آزمون اداره کل آموزش فنی و حرفه‌ای استان البرز انجام شده است. جامعه آماری این تحقیق تعداد ۴۲۰ نفر از آزمون دهندگان سازمان آموزش فنی و حرفه‌ای بودند که تعداد ۲۰۱ نفر به صورت نمونه گیری خوشه‌ای با استفاده از فرمول کوکران، به عنوان حجم نمونه انتخاب شدند. این پژوهش از نظر هدف کاربردی و از نظر روش گردآوری داده‌ها، آزمایشی است. گردآوری داده‌ها شامل پرسشنامه محقق ساخته با طیف پنج درجه‌ای لیکرت و با پایایی ۰٫۸۹ می‌باشد. پرسشنامه در سه بعد عوامل فنی، محیطی و مدیریتی تهیه شد. روایی پرسشنامه از طریق نرم‌افزار AMOS و تحلیل عاملی بر اساس معادلات ساختاری، مطلوب بوده است. بررسی معناداری بارهای عاملی نشان از معنادار بودن همه بارهای عاملی در سطح ۰/۰۰۱ بوده است. آزمون شاخص‌های برازش نشان داد که مدل از برازش نسبتاً مطلوبی برخوردار است. پرسشنامه بر روی دو گروه آزمودنی مستقل، شامل گروه آزمودنی مبتنی بر رایانه و گروه آزمودنی مبتنی بر کاغذ اجرا شد. تجزیه و تحلیل داده‌ها در دو سطح آمار توصیفی و آمار استنباطی با استفاده از نرم افزار SPSS صورت گرفت و تحلیل یافته‌ها از طریق تحلیل واریانس و t نمونه‌های مستقل انجام شد، نتایج آزمون نرمالیتی نشان داد که توزیع داده‌ها به صورت نرمال بوده است. نتایج آزمون t نمونه‌های مستقل برای تمامی ابعاد در سطوح فنی، محیطی و مدیریتی نشان داد که اختلاف معناداری میان میانگین‌های کسب‌شده در دو گروه

* کارشناس ارشد تکنولوژی آموزشی، کارشناس دفتر ارزشیابی مهارت سازمان آموزش فنی و حرفه‌ای
(نویسنده‌ی مسئول)
hbagh1981@gmail.com

** دانشجوی دکتری مهندسی صنایع، مدیرکل دفتر ارزشیابی مهارت سازمان آموزش فنی و حرفه‌ای

آزمودنی‌ها (کاغذی و الکترونیکی) وجود داشته است. همچنین نتایج تحقیق، نشان دهنده بالاتر بودن سطح میانگین نظرات آزمون شوندگان الکترونیکی و اختلاف معنادار میانگین ارزشیابی از طریق آزمون‌های مبتنی بر فناوری در مقایسه با آزمون‌های سنتی یا کاغذی بوده است. به توجه به نتایج تحقیق به نظر می‌رسد بایستی برنامه‌ریزی برای پیاده‌سازی الگوهای نوین ارزشیابی و اجرای انواع آزمون‌های مبتنی بر فناوری برای انواع آزمون‌ها صورت پذیرد.

واژه‌های کلیدی: ارزشیابی، آزمون مبتنی بر فناوری، آزمون کاغذی، سازمان آموزش فنی و حرفه‌ای

مقدمه

آزمون‌ها میزان تحقق اهداف آموزشی و همچنین میزان برآورده شدن نیازهای مردم و جامعه از مؤسسات آموزشی را تعیین می‌کنند (Shah, 2002, cited in Jamil, Tariq & Zakariys, 2012, 371)، آزمون‌ها به ارزیابی اهداف آموزشی یا اجتماعی و نیازها محدود نمی‌شوند، بلکه در نظام آموزشی مشارکت و تعامل دارند (Havens, 2002, cited in Jamil et al., 2012, 371). آزمون‌ها در تعیین آنچه که در محیط آموزشی صورت می‌گیرد، از لحاظ موضوع و نحوه‌ی تدریس مریبان و یادگیری فراگیران نقش عمده‌ای ایفا می‌کنند و می‌توانند هم تدریس و هم یادگیری را تحت تأثیر قرار دهند (Rehmani, 2003, 9). پژوهش ارزشیابی اطلاعات مهمی را درباره‌ی هزینه و فایده و مسائل مربوط به برنامه‌های جایگزین فراهم می‌سازد. ارزشیابی‌ها همچنین به منظور کمک به مدیران در اتخاذ تصمیم‌گیری‌های مناسب در ارتباط با طرح نیروی انسانی و بودجه انجام می‌گیرند (Gal, Borg & Gal, 2005, cited in Ali imohammadi, 2011, 7). استانداردهای یک برنامه‌ی ارزشیابی مبین ضرورت ویژگی‌هایی چون سودمندی، امکان‌پذیری، تناسب و دقت در یک نظام ارزشیابی است که به نیازهای اطلاعاتی مورد نظر کاربران، واقع‌بینی، عقلایی بودن، حساسیت و اقتصادی بودن، مناسبت داشتن یک ارزشیابی بر حسب انطباقش با الزامات قانونی و اخلاقی و قابلیت اعتماد داشتن داده‌های ارزشیابی دلالت دارد (Lynch, Greer, Larsen, Cummings, Harriett, Dreyfus et al, 2003, cited in Azizi & Heydari, 2012, 192).

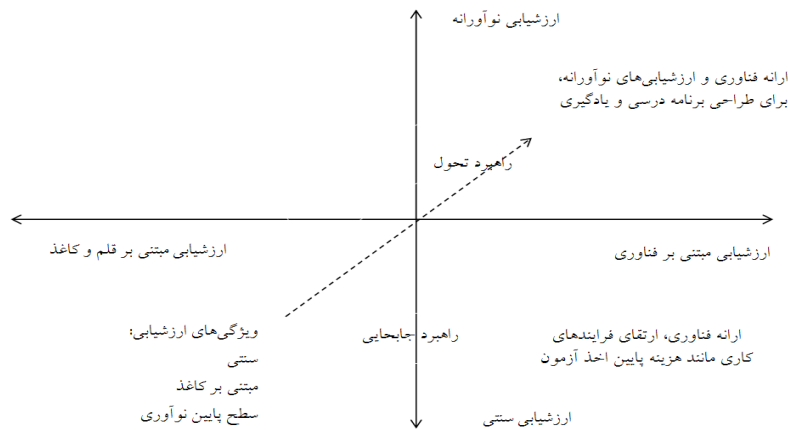
در مؤسسات آموزش عالی به منظور ارزیابی پیشرفت تحصیلی از روش‌های مختلف آزمون استفاده شده است. مثل آزمون‌های قلم- کاغذی، تکالیف، ارائه‌ها و غیره. Sim, Holifield and Brown (2004) بیش از پنجاه تکنیک مختلف را تشخیص دادند که در آموزش

عالی به‌منظور ارزیابی به‌کار رفتند که در این بین، آزمون‌ها بیشترین کاربرد را داشته‌اند. پیشرفت سریع فناوری ارتباطات و اطلاعات (ICT) در تدریس و یادگیری، رویکرد اصلی را از سیستم آزمون مبتنی بر قلم و کاغذ به آزمون مبتنی بر رایانه تغییر داده است که معمولاً به آنها «آزمون با کمک رایانه»، «سنجش رایانه‌ای»، «آزمون مبتنی بر رایانه (CBT)»، «سنجش به کمک رایانه^۳ (CAA)»، «سنجش مبتنی بر رایانه»، «سنجش آنلاین^۴»، «سنجش الکترونیکی^۵» و «سنجش مبتنی بر وب» می‌گویند. به عقیده Bull (1999) و McKenna (2001) و بسیاری از محققان دیگر، آزمون‌های مبتنی بر رایانه شکلی از سنجش هستند که در آنها رایانه یک بخش جدانشدنی در برگه‌های پرسش، ذخیره‌سازی پاسخ، علامت زدن پاسخ یا ثبت نتایج یک آزمون یا تمرین می‌باشد.

Clark (1994) معتقد است فرض بر این است که اگر محتوا و فعالیت‌های شناختی اجزای آزمون‌های مبتنی بر کاغذ و قلم آزمون‌های مبتنی بر رایانه یکسان باشند، بایستی نتایج مشابهی داشته باشند. با این حال، در بسیاری از مطالعات تفاوت‌هایی با توجه به زمینه‌ی اندازه‌گیری یافت شده است (Clariana & Wallace, 2002, 593). عوامل کلیدی که عملکرد آزمون دهندگان را تحت‌تأثیر قرار می‌دهد، عبارتند از: توانایی بررسی و تجدید نظر در پاسخ‌ها، ارائه‌ی گرافیک و متن در صفحه‌ی نمایش رایانه و تجربه‌ی قبلی کار با رایانه (Russil et al, 2003, 288). Martin (2008) به نقش نوع‌آوری و تحول در ارزشیابی و حرکت از مسیر آزمون‌های قلم و کاغذی به سمت آزمون‌های مبتنی بر فناوری تأکید کرده است. این مسیر در نمودار شماره ۱ نمایش داده شده است.

نمودار دارای چهار ربع است که سه قسمت آن به موضوع ارتباط مستقیم دارد. ربع پایین سمت چپ، نشان‌دهنده‌ی ارزشیابی سنتی است که معمولاً مبتنی بر قلم و کاغذ است، در طول سال هیچ تمایلی برای توسعه یا نوآوری نشان نمی‌دهد و بخش عمده‌ای از آزمون‌های مدارس و آموزش عالی را در بر می‌گیرد. حرکت از ربع پایین سمت چپ به ربع پایین سمت راست یک راهبرد جابه‌جایی را نشان می‌دهد. یک مثال می‌تواند انتقال از یک آزمون چندگزینه‌ای مبتنی بر

-
- 1- Computer Assisted Testing
 - 2- Computerized Assessment
 - 3- Computer Aided Assessment
 - 4- Online Assessment
 - 5- E-Assessment



نمودار ۱. ابعاد نوآوری ارزشیابی الکترونیکی

کاغذ به سمت آزمون مبتنی بر صفحه‌ی نمایش باشد. ربع تحتانی راست، نه تنها نشان‌دهنده‌ی جابه‌جایی آزمون‌های کاغذی است، بلکه شامل آزمون‌ها و ارزشیابی‌های توسعه‌یافته مبتنی بر رایانه است که در عین حال بسیاری از مشخصات آزمون‌های سنتی مبتنی بر کاغذ را انعکاس می‌دهد. راهبرد جابه‌جایی، مزایای بسیاری دارد و دلایل قانع‌کننده‌ای برای ارائه‌ی آزمون از طریق صفحه‌ی نمایش وجود دارد. این مزایا عبارتند از: کاهش هزینه‌ها، بهبود گزارشات آزمون و تجزیه و تحلیل عملکرد فراگیران، بالابردن قابلیت اطمینان، بالابردن سرعت علامت‌گذاری و چرخه‌ی گزارش‌دهی. با این حال، ویژگی مربوط به آزمون الکترونیکی در ربع تحتانی راست، این است که هیچ ابتکار مؤثری در برنامه‌ی درسی وجود ندارد. در مقابل، ربع فوقانی سمت راست یک راهبرد تحول با پشتوانه‌ی استفاده از فناوری در ارزشیابی را ارائه می‌کند. ویژگی تعیین‌کننده‌ی ارزشیابی‌های نوآورانه اشاعه‌ی نوآوری در طراحی برنامه‌ی درسی و یادگیری است. برای مثال، یک ارزشیابی مبتنی بر رایانه‌ی حل مسأله (که از فناوری برای نوآوری و یا طراحی ماهیت حل مسأله استفاده شده است) به دنبال ارائه‌ی ارزشیابی مهارت‌هایی است که به‌طور معمول از طریق آزمون‌های کاغذی قابل ارزیابی نیست (همان منبع).

مزیت‌ها و چالش‌های سنجش از طریق فناوری

رایانه‌ها و فناوری‌های مرتبط ابزارهای قدرتمندی برای رفع چالش‌های جدید طراحی و روش‌های اجرای ارزشیابی ارائه می‌دهند که از فعالیت‌های معمول فراتر می‌روند و ثبت

مجموعه‌ی گسترده‌تری از مهارت‌های شناختی را آسان می‌کنند.

طرفداران ارزشیابی الکترونیکی اغلب به مزیت‌هایی که می‌توان به آن دست یافت، اشاره می‌کنند. این مزایا ممکن است برای کاهش هزینه‌های تولید آزمون، توانایی برای استفاده‌ی مجدد، ایجاد آزمون‌های تطبیقی و قدرتمند یا ارتقای سیستم باشد. با این حال، طرفداران آزمون الکترونیکی، کمتر به نیروی بالقوه‌ی فناوری برای حمایت از نوآوری‌های آموزشی و ارتقای مهارت‌های قرن ۲۱ مانند حل مسأله، ارتباط، کارگروهی، خلاقیت و نوآوری توجه می‌کنند (Ripley, 2009, 4).

آزمون‌های مبتنی بر رایانه نسبت به آزمون‌های کاغذی یا کاغذ و قلمی چندین مزیت دارند؛ سنجش مبتنی بر فناوری، فرصت‌هایی برای ارزیابی شکل پیچیده‌ی دانش را فراهم می‌کند و برای اندازه‌گیری مواردی که امکان مواجهه و سنجش از طریق روش‌های قدیمی ممکن نیست، سودمند است (Bodmann & Robinson, 2004, 57)، وجود ارتباط بین مشاهده و تفسیر از طریق فناوری‌های مبتنی بر رایانه، امتیازدهی و تفسیر چندین جنبه از عملکرد فراگیر در گستره‌ی زیادی از وظایفی که برای ویژگی‌های شناختی انتخاب شده‌اند را ممکن می‌سازد و امکان مقایسه‌ی نتایج شاخصه‌هایی که ارزش تفسیری دارند را فراهم در سنجش آنلاین امکان دسترسی و مدیریت اطلاعات و مدیریت معرفی مهارت‌های ارتباطی وجود دارد که نمی‌توانند در آزمون‌های عادی مورد ارزیابی قرار گیرند. ابزارهای آموزش و آزمون مبتنی بر وب به‌طور عمده به این دلیل برای فراگیران کار گرفته شده‌اند که تجارب یادگیری جذاب‌تر و با تعامل بیشتری را فراهم کنند، از آموزش در کلاس درس حمایت کنند و آن را ارتقا دهند. همچنین می‌توانند به مربیان کمک کنند تا فعالیت‌های تدریس اثربخش انجام دهند. آن را در حمایت از محیط یادگیری، سازمان‌دهی دروس و حمایت از جریان پیشرفت دانش‌آموزان به‌کار گیرند و یک ابزار ارزشیابی و گزارش‌گیری فوری را فراهم آورند (Brown, Race & Bull, 1999; cited in Jamil et al, 2012, 371).

(Noyes and Garland (2008) بر این باورند که به علت مزایای آزمون‌های مبتنی بر رایانه مانند سرعت، نتایج عینی و سهولت مقایسه‌ی آن با نتایج دیگران، این روش بسیار محبوب است (cited in solak, 2014, 203) علاوه بر این حرکت به سمت آزمون‌های استاندارد ریشه در مزیت‌هایی دارد که این نوع آزمون در مقایسه با نوع قلم- کاغذی فراهم کرده است، این مزیت‌ها عبارتند از: مدیریت هزینه- فایده، کنترل آسان‌تر، دقت بیشتر، بی‌واسطه بودن نمره‌دهی و گزارش‌دهی فوری، مکان و برنامه‌ریزی آزمون انعطاف‌پذیر. این مطالعه، همچنین نشان داد دانش‌آموزانی که

آشنا‌تر به رایانه هستند احساس راحتی بیشتری در استفاده از آن دارند. ایشان مزایای استفاده از رایانه را در ارزشیابی به پنج دسته‌ی اصلی تقسیم کردند. این موارد به شرح ذیل می‌باشند:

۱. وزن‌دهی به برقراری ارتباط؛ برای مثال استفاده از گرافیک امکان ارائه‌ی پویا برای محتوای آزمون را فراهم می‌کند. رایانه این قابلیت را دارد که در سرعت‌ها و سطوح مختلف، متناسب با نیازهای افراد ارائه شود. برخلاف فعالیت‌های قلم و کاغذی، رایانه می‌تواند استفاده از تبادل بازخورد دوطرفه را ایجاد کند.

۲. فراوانی کاربران؛ آزمون مبتنی بر رایانه از طریق اینترنت، این امکان را می‌دهد که نمونه‌های متنوع بیشتری در دسترس باشند، چرا که مردم برای دسترسی تنها به یک رایانه نیاز دارند. همچنین به افراد اجازه می‌دهد که بخشی از آزمون را در منازلشان انجام دهند و حتماً لازم نیست برای آزمون به علت مسائل مربوط به معلولیت یا غیره در یک موقعیت آزمایشگاهی حاضر شوند و در دسترس باشند.

۳. استانداردسازی محیط آزمون؛ به این معنی که آزمون در روش و شکل یکسان برای یک زمان مشخص ارائه می‌شود. بنابراین، خطاها در برنامه‌ریزی که ممکن است برنامه را به انحراف بکشاند، به حداقل می‌رسد.

۴. نمره‌گذاری برخط؛ موجب بازخورد سریع‌تر و دقت بهتر که منجر به کاهش خطای انسانی می‌شود. اطلاعات مرتبط با رفتارهای حین آزمون مثلاً اینکه چقدر زمان برای هر آزمون صرف شده به آسانی جمع‌آوری می‌شود. به‌طور کلی، این موضوع ثابت شده است که ارائه و نمره‌دهی در آزمون آنلاین منجر به صرفه‌جویی در هزینه‌های اقتصادی، به‌ویژه برای نمونه‌های بزرگ می‌شود.

۵. کیفیت و کمیت نوشتن؛ Goldberg et al. در یک تحلیل مطالعات مقایسه‌ای در مورد رایانه و کاغذ در طی یک دوره‌ی ۱۰ ساله دریافتند که رایانه‌ها در کمیت و کیفیت نوشتن مورد توجه بیشتری هستند. به عبارت دیگر، یک ادبیات نوشتاری متفاوتی در نوشتن، کیفیت و ارزشیابی روی خروجی نوشتاری رایانه و کاغذ وجود دارد که نباید نادیده گرفت (MacCann, Eastment & Pickering, 2002). البته آزمون‌های مبتنی بر رایانه تهدیدهایی را هم به همراه دارد که از جمله‌ی آن‌ها می‌توان به احتمال نفوذ امنیتی^۱ (Van Lent, Gerben, 2009) اشاره کرد. یکی از چالش برانگیزترین مسائل برای بدنه‌ی ارزشیابی، اطمینان یافتن از این است که همه‌ی

1- Likelihood of serious security breaches

آزمون‌دهندگان با شرایط یکسانی در آزمون مواجه می‌شوند. نیاز است که همه‌ی آزمون‌دهندگان به مجموعه‌ی از پیش تعیین شده‌ای از سخت‌افزار، نرم‌افزار و زیرساخت‌های مرتبط با آن در مرکز آزمون تأیید شده، دسترسی داشته باشند (Kikis-Papadakis & Kollias, 2009, 101).

مروری بر تحقیقات انجام‌شده

برای بررسی پیشینه‌ی تحقیقات انجام‌شده در زمینه‌ی مورد مطالعه، منابع داخلی مرتبط با موضوع یافت نشد، لیکن مطالعات متعددی در سطح بین‌المللی در خصوص انواع مدل‌های سنجش مهارت اعم از آزمون‌های قلم - کاغذی، آزمون‌های مبتنی بر رایانه، مبتنی بر وب و مبتنی بر ابزارهای الکترونیکی همراه و همچنین مقایسه‌ی آن‌ها انجام شد که خلاصه روش و یافته‌های چند مطالعه در ذیل قید می‌شود:

اداره‌ی ایالات متحده‌ی آمریکا، امکان ارزشیابی مبتنی بر فناوری (TBA)، در چارچوب یک پروژه توسط مرکز ملی آمار تحصیلی (NCES) مورد بررسی قرار گرفت. هدف اصلی از این مطالعه، آماده‌سازی برای انتقال از آزمون‌های مبتنی بر کاغذ به آزمون‌های مبتنی بر فناوری بود. این پروژه چهار پرسش اصلی را ارزیابی کرد: (۱) اندازه‌گیری، (۲) برابری، (۳) بهره‌وری و (۴) موارد عملیاتی. پروژه، سه حوزه را در بر می‌گرفت: ریاضیات برخط، نوشتن برخط و حل مسأله در محیط‌های آکنده از فناوری. نتایج نشان داد که به‌طور میانگین، آزمون مبتنی بر رایانه به خوبی عمل کرد اگر چه تأثیر رسانه تا حدی به شکل موضوع و آشنایی شرکت‌کنندگان در مطالعه به رایانه وابسته بود. در اروپا، برخی از کشورها در حال حاضر، برخی شکل‌های آزمون‌های رایانه‌ای را به کار برده‌اند و چند مؤسسه‌ی ملی دیگر در حال ارائه کردن آزمون‌های رایانه‌ای در مقیاس بزرگ هستند (مراجعه شود به Scheuermann & Pereira, 2008).

به نظر (Fluck, Pullen and Harper (2009) مریبان باید بدانند که کدام تکنیک‌های سنجش به فراگیران اجازه می‌دهند تا امکانات فناوری جدید را به کار بگیرند. این محققان یک آزمون الکترونیکی برای دانشجویان سال چهارم کارشناسی برنامه‌ی تحصیلی دانشگاه تاسمانیا^۱ انجام دادند. موفقیت تحصیلی تعداد ۲۷۰ نفر از دانشجویان از طریق دو فعالیت هم وزن شده مورد بررسی قرار گرفت. فعالیت اول یک تکلیف در خانه بود که در آن دانشجویان از طریق فناوری

1- Tasmania.

اطلاعات با محتوای آموزش آشنا شدند. فعالیت دوم ۲ ساعت آزمون متشکل از ۱۴ سؤال بر مبنای تمام موضوعات درسی آن واحد بود. در پایان آزمون، به دانشجویان یک برگه‌ی نظرسنجی با پنج سؤال داده شد. در نظرسنجی مشخص شد ۳۸ درصد از پاسخ‌دهندگان به نظرسنجی قبلاً یک آزمون مبتنی بر رایانه داشته‌اند. ۷۸ درصد قبل از آزمون از لوح فشرده‌ی تمرین استفاده کرده بودند و ۷۱ درصد معتقد بودند که آزمون بسیار مفید یا نسبتاً مفید بود و از تعداد ۲۳۰ پاسخ‌دهندگان، ۹۴٫۵ درصد آزمون‌های مبتنی بر رایانه را ترجیح دادند. عامل بسیار مهم در ترجیح دادن روش رایانه‌ای، ارتباط قبلی با آزمون‌های مبتنی بر رایانه بود.

(Karadeniz (2009) به مطالعه‌ی تأثیر سنجش مبتنی بر کاغذ، سنجش مبتنی بر وب و سنجش مبتنی بر تلفن همراه روی موفقیت دانشجویان پرداخت. یک گروه از ۳۸ دانشجو به مدت ۳ هفته آزمایش شدند. تفاوت‌های معنی‌دار بین نمرات دانشجویان در هفته‌ی دوم یافت شد. محققان معتقد بودند که دانشجویان در خصوص سنجش مبتنی بر وب و موبایل، به دلیل سهولت کاربرد، بازخورد جامع و فوری آن، نظر مثبتی داشتند. به علاوه، مطلوب‌ترین و مناسب‌ترین آزمون‌ها، آزمون‌های مبتنی بر وب بودند و کمترین علاقه معطوف به آزمون‌های مبتنی بر کاغذ بود.

دانشگاه ملی سنگاپور در سال ۲۰۰۴ آزمون مبتنی بر رایانه را ارائه کرد. Lim, Ong, Wilder-Smith & Seet (2006) نظر دانشجویان پزشکی در مورد آزمون رایانه‌ای در مقایسه با آزمون کاغذی را بررسی کردند. ۲۷۳ نفر (۲۳٫۵٪) از دانشجویان سال آخر رشته‌ی پزشکی عمومی از طریق یک سنجش آنلاین مورد آزمون قرار گرفتند که از میان آنها ۹۱ نفر (۷۹٫۸٪) آزمون مبتنی بر رایانه و ۱۱ نفر (۹٫۶٪) شکل قلم و کاغذی را ترجیح دادند. همچنین ۱۲ نفر (۱۰٫۵٪) در انتخاب یک روش خاص، تردید داشتند. محققان بیان کردند که ۴۲ نفر آزمون‌های مبتنی بر رایانه را به دلیل کیفیت خوب تصاویر و عدم وابستگی به موقعیت‌های در نظر گرفته شده برای نشستن جهت آزمون، دوست داشتند. ۲۲ نفر به این دلیل آن را پسندیدند که هر طور دوست دارند آن را انجام می‌دهند، ۱ نفر عنوان کرد که آزمون‌های مبتنی بر رایانه جالب و مفرح بود. ۴ نفر از راحتی آزمون مبتنی بر رایانه لذت بردند و ۶ نفر به «برابری» به عنوان دلیلی که آزمون مبتنی بر رایانه را به آزمون قلم و کاغذی ترجیح می‌دادند، اشاره کردند.

(Bodmann & Robinson (2004) برای مقایسه‌ی تفاوت‌های عملکرد و سرعت در میان آزمون‌های مبتنی بر رایانه و آزمون‌های قلم و کاغذی، یک مطالعه‌ی آزمایشی انجام دادند. ۵۵ دانشجوی دوره‌ی کارشناسی رشته‌ی روان‌شناسی تربیتی، در آزمایش شرکت کردند که قبلاً با

آزمون‌های رایانه‌ای آشنا بودند. در هر دو روش آزمون‌های مبتنی بر کاغذ و مبتنی بر رایانه، ۳۰ سؤال وجود داشت که ۳۵ دقیقه زمان داشتند. تقریباً نیمی از کلاس (۲۸ دانشجو)، ابتدا آزمون را روی رایانه انجام دادند و بقیه ترجیح دادند آزمون اول را به صورت کاغذی انجام دهند. شیوه‌ها در آزمون‌های دوم تغییر کردند، به طوری که گروه اول آزمون مبتنی بر کاغذ و گروه دوم آزمون مبتنی بر رایانه را با یک فاصله‌ی دو هفته‌ای دریافت کردند. نتیجه این بود که دانشجویان، آزمون مبتنی بر رایانه را سریع‌تر از آزمون مبتنی بر کاغذ به پایان رساندند و هیچ تفاوتی در نمرات نبود.

Koppel & Hollister برای بررسی تأثیر یک سنجش مبتنی بر رایانه بر عملکرد فراگیر در مقایسه با یک روش آزمون قدیمی، مطالعه‌ای انجام دادند. نظرات و نمرات آزمون دانشجویان در هر دو روش سنجش مبتنی بر کاغذ و مبتنی بر رایانه بررسی شدند. از مجموع ۱۲۳ دانشجو، ۹۱ نفر هیچ تجربه‌ی قبلی از آزمون مبتنی بر رایانه نداشتند و ۴۲ نفر باقیمانده در دوره‌های قبلی آن را تجربه کرده بودند، آزمون مبتنی بر رایانه، شامل ۲۵ سؤال بود. نمرات دانشجویان در آزمون مبتنی بر رایانه به طور خودکار ثبت شده بودند که شامل امتیازات کلی اکتسابی و بر اساس توانایی دانشجویان برای تکمیل یک فعالیت مهارت‌محور ویژه در کاربرد اکسل^۱ بود. آزمون مبتنی بر رایانه در یک دوره‌ی کلاسی انجام شد و در پایان، دانش‌آموزان نتایج خود را مرور کردند. سنجش مبتنی بر کاغذ به طور دستی با دیدن پرینت و فایل اکسل انجام شد و نتایج آزمون مبتنی بر کاغذ به صورت دستی با یادداشت‌برداری اطلاعات مورد نیاز در یک صفحه‌ی اکسل ثبت شده بودند، انجام شد. نمره‌گذاری مربوط به سنجش مبتنی بر کاغذ بر اساس محصول نهایی ارسالی توسط هر دانشجو انجام شد. در مورد پرسش سهولت استفاده، اکثر دانشجویان (۵۹ درصد) معتقد بودند کاربری نرم‌افزار آسان است، ۲۹ درصد اعتقاد داشتند متوسط و ۱۲ درصد معتقد بودند تا حدی دشوار است. در مورد گزارش میزان مهارت‌ها، ۷۶ درصد پاسخ دادند که سنجش مبتنی بر رایانه مؤثرتر بود. در ارزیابی مشکل پاسخگویی روی کاغذ سؤال، ۶۵ درصد نظر متوسط داشتند. ۳۴ درصد از دانشجویان حس کردند سیستم نمره‌دهی اتوماتیک عادلانه بود، در حالی که ۳۹ درصد نمره‌بندی آن را عادلانه نمی‌دانستند. تنها ۱۹ درصد احساس کردند که سنجش مبتنی بر رایانه بر عملکرد آنها تأثیر منفی می‌گذارد. سنجش مبتنی بر رایانه فقط از نظر ۱۴ درصد دانشجویان آسان‌تر بود در حالی که ۴۹ درصد آن

1- Excel

را دشوارتر می‌دانستند. در مجموع، ۵۸ درصد سنجش مبتنی بر رایانه و ۴۲ درصد نیز سنجش مبتنی بر کاغذ را ترجیح دادند. به خاطر زمان مورد نیاز کمتر در آزمون، کاهش زمان نمره‌دهی، شیوه‌ی آسان ثبت نمرات و افزایش اعتبار آزمون از طریق تحلیل آماری پایان آزمون، در خصوص سنجش مبتنی بر رایانه نظر مثبت‌تری داشتند (cited in jamil et al 2012, 372).

(2002) Calarina & Wallace در مطالعه‌ی نشان دادند که ارائه‌ی آزمون‌های مبتنی بر رایانه در مقایسه با آزمون‌های کاغذی، بر امتیازات فراگیران، تأثیر مثبتی دارد و عملکرد گروه آزمون رایانه‌ای بهتر از آزمون عملکرد گروه آزمون کاغذی بود.

روش‌شناسی

این پژوهش از نظر هدف کاربردی و از نظر روش گردآوری داده‌ها، پیمایشی است. حجم جامعه‌ی آماری تحقیق، شامل ۴۲۰ نفر از متقاضیان آزمون‌های مهارتی اداره‌ی کل آموزش فنی و حرفه‌ای استان البرز بوده‌اند که تعداد ۲۰۱ نفر به صورت نمونه‌گیری خوشه‌ای و با استفاده از فرمول کوکران در چهار حرفه‌ی مهارتی، به شرح جداول شماره‌ی ۱ و ۲ به عنوان جامعه‌ی نمونه انتخاب شده‌اند.

جدول ۱. توصیف جمعیت‌شناختی جامعه‌ی آماری به تفکیک نوع آزمون و جنسیت

درصد	جمع	جنسیت		نوع آزمون		عنوان حرفه
		زن	مرد	الکترونیکی	کاغذی	
۲۵,۴٪	۵۱	۲۷	۱۴	۲۶	۲۵	رایانه کار ICDL درجه‌ی دو
۲۴,۸٪	۵۰	۲۳	۲۷	۲۵	۲۵	نقشه‌کشی معماری
۲۴,۴٪	۴۹	۲۴	۲۵	۲۵	۲۴	نقشه‌کشی عمومی ساختمان
۲۵,۴٪	۵۱	۲۶	۲۵	۲۹	۲۲	مدیر آموزشگاه فنی و حرفه‌ای
	۲۰۱	۱۰۰	۱۰۱	۱۰۵	۹۶	جمع کل
		۴۹,۸٪	۵۰,۲٪	۵۲,۲٪	۴۷,۸٪	درصد

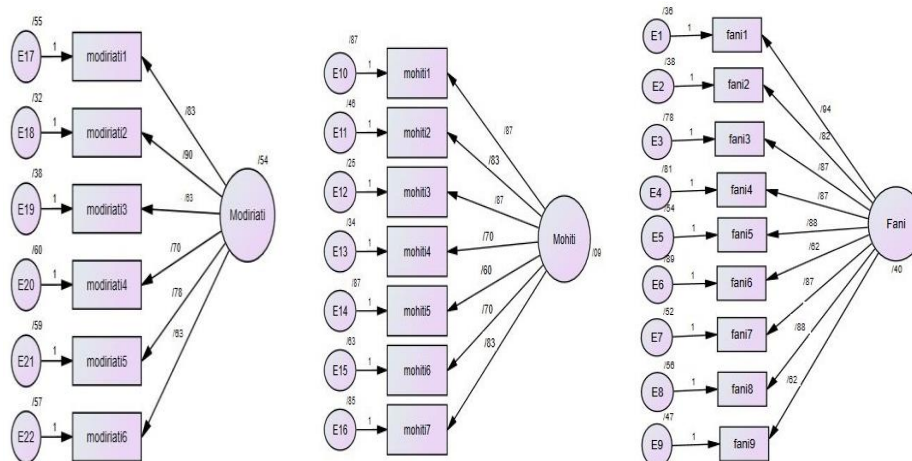
ابزار گردآوری داده‌ها در این تحقیق، پرسشنامه‌ی میزان مطلوبیت آزمون با طیف لیکرت پنج سطحی (از خیلی مخالف = ۱ تا خیلی موافق = ۵) می‌باشد. پرسشنامه بر روی دو گروه آزمودنی مستقل، شامل گروه آزمودنی مبتنی بر رایانه (۱۰۵ نفر) و گروه آزمودنی مبتنی بر کاغذ (۹۶ نفر) اجرا شد.

برای تعیین روایی تحقیق از روش روایی محتوایی و روایی سازه استفاده شد. روایی

جدول ۲. توصیف جمعیت‌شناختی جامعه‌ی آماری به تفکیک تحصیلات و وضعیت شغلی

حرفه	عنوان	تحصیلات							وضعیت شغلی		
		پایین‌تر از دیپلم	دیپلم	فوق دیپلم	لیسانس	بالا تر از لیسانس	مشغول تحصیل	بیکار	شغل آزاد	کارمند	جمع
رایانه کار ICDL درجه‌ی دو		۱۵	۱۹	۱۰	۴	۳	۲۱	۱۸	۴	۸	۵۱
نقشه‌کشی معماری		۸	۲۰	۱۴	۴	۴	۱۶	۲۱	۶	۸	۵۰
نقشه‌کشی ساختمان		۱۳	۱۷	۹	۶	۴	۱۶	۱۸	۸	۸	۴۹
مدیر آموزشگاه		۲	۴	۱۰	۲۱	۱۴	۱۴	۱۹	۸	۱۰	۵۱
جمع کل		۳۸	۶۰	۴۳	۳۵	۲۵	۶۷	۷۶	۲۶	۳۲	۲۰۱
درصد		٪۱۸	٪۲۹	٪۲۱	٪۱۷	٪۱۲	٪۳۳	٪۳۷	٪۱۲	٪۱۵	

محتوایی پرسشنامه با نظرسنجی از متخصصان در زمینه‌ی موضوع مورد پژوهش و پس از انجام اصلاح‌های ضروری، تأیید شد. در روایی سازه، تحلیل عاملی تأییدی از طریق نرم‌افزار AMOS استفاده شد تا با بررسی افتراقی مشخص شود گویه‌های پرسشنامه تا چه حدی، سازه‌های مورد نظر را می‌سنجند. سازه‌های پرسشنامه در این تحقیق عبارتند از: عوامل فنی شامل ۹ سؤال، عوامل محیطی شامل ۷ سؤال و عوامل مدیریتی شامل ۶ سؤال. بارهای عاملی و واریانس تبیین‌شده‌ی سازه‌های پرسشنامه به صورت شکل شماره‌ی ۱ می‌باشد:



شکل ۱. بارهای عاملی و واریانس تبیین‌شده‌ی سازه‌های تحقیق

بر اساس شکل ۱، سؤال اول سازه‌ی فنی، سؤال اول و سوم سازه‌ی عوامل محیطی و سؤال دوم سازه‌ی مدیریتی، بیشترین بار عاملی را روی سازه‌های خود دارند. نتایج آزمون‌های همانندی و شاخص‌های برازش در جدول شماره‌ی ۳ قید شده است: با توجه به مثبت بودن درجه‌ی آزادی مدل‌ها، نتیجه می‌شود که این مدل‌ها همانند هستند.

جدول ۳. نتایج آزمون‌های همانندی و شاخص‌های برازش متغیرهای تحقیق

عوامل مدیریتی	عوامل محیطی	عوامل فنی
Notes for Model (Default model)	Notes for Model (Default model)	Notes for Model (Default model)
Computation of degrees of freedom (Default model)	Computation of degrees of freedom (Default model)	Computation of degrees of freedom (Default model)
Number fo distinct sample moments: 21 Number fo distinct parameters to be estimated: 12 Degrees of freedom (21-12): 9	Number fo distinct sample moments: 28 Number fo distinct parameters to be estimated: 14 Degrees of freedom (28-14): 14	Number fo distinct sample moments: 45 Number fo distinct parameters to be estimated: 18 Degrees of freedom (45-18): 27
Result (Default model)	Result (Default model)	Result (Default model)
Minimum was achieved Chi-square= 41/818 Degrees of freedom= 9 Probability level= /000	Minimum was achieved Chi-square= 71/368 Degrees of freedom= 14 Probability level= /000	Minimum was achieved Chi-square= 236/391 Degrees of freedom= 27 Probability level= /000

بررسی معناداری بارهای عاملی از طریق آزمون معناداری نشان از معنادار بودن همه بارهای عاملی در سطح ۰,۰۰۱ بوده است. شاخص‌های برازش مدل در جدول شماره‌ی ۴ قید شده است که نشان می‌دهد که مدل از برازش نسبتاً مطلوبی برخوردار است.

جدول ۴. شاخص‌های برازش مدل ساختاری تحقیق

عوامل مدیریتی	عوامل محیطی	عوامل فنی	شاخص‌ها
۰,۰۰۰	۰,۰۰۰	۰,۰۰۰	P-Value
۰,۵۸	۰,۵۵	۰,۶	CFI
۰,۰۵	۰,۰۹	۰,۱	RMR
۰,۹۳	۰,۹۱	۰,۹۰	GFI

به منظور بررسی پایایی ابزار تحقیق نیز از پایایی تحقیق از روش آلفای کرونباخ و نرم افزار SPSS استفاده شد. میزان پایایی هر یک از ابعاد پرسشنامه در جدول ذیل قید شده است. با توجه به اینکه میزان پایایی در همه‌ی ابعاد بالای ۰,۷ بوده است، پایایی پرسشنامه، مطلوب ارزیابی می‌شود.

جدول ۵. شاخص‌های برازش مدل ساختاری تحقیق

ابعاد	تعداد سؤالات	شماره‌ی سؤالات	میزان پایایی
عوامل فنی	۹	۱ تا ۹	۰,۸۳۷
عوامل محیطی	۷	۱۰ تا ۱۶	۰,۷۱۸
عوامل مدیریتی	۶	۱۷ تا ۲۲	۰,۸۳۱

یافته‌های تحقیق

شاخصه‌های مرکزی و پراکندگی مربوط به هریک از ابعاد ابزار تحقیق در جدول شماره‌ی ۶ نشان داده شده است.

جدول ۶. شاخصه‌های مرکزی و پراکندگی سازه‌های تحقیق

ابعاد	نوع آزمون	تعداد نمونه	میانگین	انحراف معیار	خطای استاندارد میانگین
عوامل فنی	کاغذی	۹۶	۳,۵۵	۰,۷۰۰	۰,۰۷۱
	الکترونیکی	۱۰۵	۴,۱۱	۰,۴۳۸	۰,۰۴۳
عوامل محیطی	کاغذی	۹۶	۳,۳۸	۰,۶۲۰	۰,۰۶۳
	الکترونیکی	۱۰۵	۴,۰۹	۰,۳۱۹	۰,۰۳۹
عوامل مدیریتی	کاغذی	۹۶	۳,۱۲	۰,۵۷۷	۰,۰۳۸
	الکترونیکی	۱۰۵	۴,۱۷	۰,۳۸۰	۰,۰۳۷

با توجه به جدول شماره‌ی ۶، بین میانگین نظرات آزمون‌دهندگان در همه‌ی ابعاد برای گروه‌های آزمون‌ی کاغذی و الکترونیکی تفاوت وجود دارد. این تفاوت در حرفه‌های مختلف برای میانگین به‌دست آمده‌ی کل سؤالات (سطح رضایت‌مندی) برای این دو گروه نیز مشاهده می‌شود که نتایج مقایسه در جدول شماره‌ی ۷ نشان داده شده است. در جدول میزان این اختلاف برای همه‌ی حرفه‌ها در دو نوع آزمون تقریباً برابر است.

مقدار چولگی و کشیدگی داده‌ها بررسی شد که این مقدار برای تمامی گویه‌ها در بازه‌ی (۲- و ۲-) قرار دارد و این نشان‌دهنده‌ی نرمال بودن توزیع داده‌هاست. مقادیر میانگین، چولگی و کشیدگی گویه‌ها در جدول شماره‌ی ۷ نشان داده شده است:

همچنین، نرمال بودن داده‌ها مطابق با جدول شماره‌ی ۹، از طریق آزمون کولموگروف - اسمیرنوف بررسی شد که با توجه به سطح معناداری بزرگ‌تر از ۰,۰۵ این نتیجه حاصل شد که داده‌ها نرمال می‌باشند.

جدول ۷. مقایسه‌ی شاخصه‌های مرکزی و پراکندگی کلیه‌ی سؤالات به تفکیک حرفه

عنوان حرفه‌ها	نوع آزمون	تعداد نمونه	میانگین	انحراف معیار	خطای استاندارد میانگین
رایانه کار ICDL درجه‌ی دو	کاغذی	۲۵	۳,۳۵	۰,۵۴۰	۰,۱۰۸
	الکترونیکی	۲۶	۴,۱۴	۰,۲۷۳	۰,۰۳۶
نقشه‌کشی معماری	کاغذی	۲۵	۳,۴۰	۰,۵۵۰	۰,۱۱۰
	الکترونیکی	۲۵	۴,۱۳	۰,۲۵۰	۰,۰۵۰
نقشه‌کشی عمومی ساختمان	کاغذی	۲۴	۳,۳۸	۰,۵۴۳	۰,۱۱۱
	الکترونیکی	۲۵	۴,۱۱	۰,۲۸۸	۰,۰۵۷۶
مدیر آموزشگاه فنی و حرفه‌ای	کاغذی	۲۲	۳,۴۰	۰,۵۷۷	۰,۱۲۳
	الکترونیکی	۲۹	۴,۱۰	۰,۳۲۴	۰,۰۶۰

جدول ۸. مقایسه‌ی میانگین، چولگی و کشیدگی کلیه‌ی سؤالات

گویه‌ها	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
میانگین	۳,۹۵	۴,۰۹	۳,۸۸	۳,۶۰	۳,۸۵	۳,۷۵	۳,۶۵	۳,۸۹	۳,۹۶	۳,۴۷	۳,۸۶
چولگی	-۰,۱۴۲	-۰,۸۵۶	-۰,۴۴۱	-۰,۲۹۷	-۰,۷۸۴	-۰,۳۱۸	-۰,۰۷۵	-۰,۰۸۸	-۰,۹۵۱	-۰,۱۳۹	-۰,۵۲۶
کشیدگی	-۰,۸۴۵	۰,۶۱۱	-۰,۸۲۶	-۰,۷۲۱	-۰,۴۲۲	-۰,۹۱۵	-۱,۲۵۶	۰,۶۸۳	۰,۱۷۲	-۰,۹۵۶	-۰,۳۷۸
گویه‌ها	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲
میانگین	۳,۸۴	۳,۹۹	۳,۶۶	۳,۸۲	۳,۶۳	۳,۶۵	۳,۵۶	۳,۵۹	۳,۶۹	۳,۸۹	۳,۶۴
چولگی	-۰,۳۷۴	-۱,۱۲۳	-۰,۵۷۹	-۰,۱۷۶	-۰,۴۷۸	-۰,۲۸۸	-۰,۵۰۸	-۰,۳۵۰	-۰,۰۸۸	-۰,۹۶۷	-۰,۲۷۵
کشیدگی	-۰,۳۶۰	-۱,۴۲۸	-۰,۹۱۸	-۰,۹۴۲	-۰,۳۶۱	-۰,۸۴۹	-۰,۵۴۵	-۰,۹۴۵	-۰,۹۲۵	-۰,۸۰۴	-۰,۷۴۰

جدول ۹. نتایج آزمون کلموگروف-اسمیرنوف

عوامل فنی	عوامل مدیریتی	عوامل محیطی	همه‌ی ابعاد
۱,۵۴۴	۱,۲۴۲	۱,۷۵۵	آماره‌ی کلموگروف اسمیرنوف
۰,۰۶۲	۰,۰۹۱	۰,۰۷۲	سطح معناداری

به‌منظور بررسی مقایسه‌ای آزمون‌های کاغذی و مبتنی بر رایانه، از آزمون t نمونه‌های مستقل استفاده شد. در این آزمون فرض بر این است که تفاوت معناداری میان میانگین دو

نمونه‌ی آماری وجود ندارد؛

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

آزمون فرضیه در دو حالت برابری واریانس و عدم برابری واریانس برای متغیر سطح رضایت‌مندی (میانگین تمامی گویه‌ها و سازه‌ها) مطابق با جدول شماره ۱۰ صورت گرفت. آزمون اول برابری واریانس‌ها، با مقدار $0,35,102$ برای آماره‌ی F با معیار تصمیم $0,000$ نشان می‌دهد که فرض برابر واریانس‌ها رد می‌شود. یعنی واریانس دو جامعه با هم برابر نیست. بنابراین، از فرض عدم برابری واریانس‌ها استفاده می‌شود.

جدول ۱۰. آزمون آماری t نمونه‌های مستقل برای تمامی سؤالات ابزار تحقیق

آزمون t نمونه‌های مستقل برای برابری میانگین		آزمون لون برای برابری واریانس		آماره‌ی t	سطح معناداری	آماره‌ی F	
ضریب اطمینان $0,95$	خطای میانگین	درجه‌ی آزادی	آماره‌ی t				
حد بالا	حد پایین						
$-0,62192$	$-0,86107$	$0,6041$	199	$-12,267$	$0,000$	$35,102$	فرض برابری واریانس
$-0,61848$	$-0,86361$	$0,6199$	$139,983$	$-11,954$			فرض عدم برابری واریانس

آزمون دوم در سطر دوم آزمون مقایسه‌ی میانگین‌هاست که با فرض عدم برابر واریانس‌ها، مقدار آماره‌ی t با معیار تصمیم $0,000$ دارد که کوچک‌تر از $0,05$ در ضریب اطمینان $0,95$ می‌باشد. فرض صفر رد می‌شود و می‌توان نتیجه گرفت که میان میانگین دو گروه آزمون‌های کاغذی و الکترونیکی اختلاف معناداری وجود دارد. همچنین با توجه به منفی بودن مقادیر حد بالا و حد پایین نتیجه می‌شود که تفاوت میانگین دو گروه کمتر از صفر است و میانگین گروه اول (آزمون‌های کاغذی) از گروه دوم (آزمون‌های الکترونیکی) کوچک‌تر است.

بررسی این آزمون برای هر یک از ابعاد پرسشنامه (۱. عوامل فنی، ۲. عوامل مدیریتی و ۳. عوامل محیطی)، نتایج مشابهی را در بر داشته است که نتایج آن در جدول شماره ۱۱ نشان داده شده است. بنابراین، می‌توان نتیجه گرفت که در هر یک از ابعاد پرسشنامه، میان میانگین دو گروه آزمون‌های کاغذی و الکترونیکی اختلاف معناداری وجود داشته است. همچنین بررسی مقدار میانگین‌های هر یک از ابعاد پرسشنامه و میانگین کلی دو گروه آزمون‌های کاغذی و الکترونیکی نشان از بالاتر بودن سطح میانگین‌ها در گروه آزمون‌های الکترونیکی در مقایسه با آزمون‌های کاغذی بوده است.

جدول ۱۱. آزمون آماری t نمونه‌های مستقل برای سؤالات هریک از ابعاد ابزار تحقیق

	آزمون t نمونه‌های مستقل برای برابری میانگین		آزمون لون برای برابری واریانس						
			ضریب اطمینان ۰,۹۵	سطح معناداری (دوطرفه)	درجه‌ی آزادی	آماره‌ی t		سطح معناداری	
								حد بالا	حد پایین
۱	۰,۶۲۱۹۲	-۰,۷۲۸۸۲	۰,۰۰۰	۱۹۹	-۶,۹۵۳	۰,۰۰۰	۱۶,۴۰۲	فرض برابری واریانس	
	۰,۶۱۸۴۸	-۰,۷۳۲۳۰	۰,۰۰۰	۱۵۶,۷۲۱	-۶,۸۱۷			فرض عدم برابری واریانس	
۲	۰,۵۶۹۰۴	-۰,۸۴۰۲۳	۰,۰۰۰	۱۹۹	-۱۰,۲۴۸	۰,۰۰۰	۱۶,۴۰۲	فرض برابری واریانس	
	۰,۵۶۵۰۸	-۰,۸۴۴۱۹	۰,۰۰۰	۱۳۹,۰۶۰	-۹,۹۸۳			فرض عدم برابری واریانس	
۳	۰,۹۰۸۵۴	-۱,۱۷۸۲۷	۰,۰۰۰	۱۹۹	-۱۵,۲۵۶	۰,۰۰۰	۱۸,۰۱۹	فرض برابری واریانس	
	۰,۹۰۵۹۲	-۱,۱۸۰۸۹	۰,۰۰۰	۱۶۱,۸۶۷	-۱۴,۹۸۷			فرض عدم برابری واریانس	

بحث و نتیجه‌گیری

نتایج این تحقیق با نتایج تحقیق (2002) Bennett, Persky, Weiss و Calarina and Wallace and Jenkins (2007) که نشان از عملکرد بهتر گروه آزمودنی‌ها با رایانه داشت، همخوانی دارد و همسو با نتایج (2009) Fluck, Pullen and Harper که نشان از ترجیح دادن آزمون با رایانه نسبت به آزمون کاغذی داشت، می‌باشد. همچنین، یافته‌های تحقیق حاضر با نتایج تحقیق کارادنیز (2009) Karadeniz و (2006) Lim et al، که اظهار داشتند دانشجویان در تحقیق وی درخصوص آزمون‌های مبتنی بر فناوری نظر مثبت‌تری نسبت به آزمون‌های کاغذی داشتند، همسویی و همخوانی دارد.

نتایج توصیفی درباره‌ی وضعیت جمعیت‌شناختی آزمون‌دهندگان در دو گروه کاغذی و الکترونیکی بیانگر این نکات است که تعداد آزمون‌دهندگان مرد و زن در این تحقیق تقریباً برابر بوده است (۵۰ درصد) و همچنین درصد تعداد آزمون‌دهندگان در چهار حرفه‌ی مهارتی تقریباً برابر (حدود ۲۵ درصد) بوده است. بیشترین درصد آزمون‌دهندگان (۲۹ درصد)، در مجموع چهار حرفه‌آزمون، دارای مقطع تحصیلی دیپلم بوده‌اند و جمعیت بالاتر از لیسانس، کمترین درصد نمونه‌های آماری (۱۲ درصد) را تشکیل دادند. همچنین جمعیت بیکار، بیشترین

درصد آزمون‌دهندگان (۳۷ درصد) و جمعیت دارای شغل آزاد کمترین تعداد نمونه‌ی آماری (۱۲ درصد) را تشکیل دادند. این موضوع نشان‌دهنده‌ی نقش بیشتر افراد دارای مقطع تحصیلی دیپلم و بیکار در نمونه‌های آماری است.

بر اساس یافته‌های تحقیق، میانگین نمرات کسب‌شده‌ی گروه آزمون‌دهندگان الکترونیکی در تمامی ابعاد بالاتر از میانگین نمرات گروه آزمون کاغذی بوده است. در گروه آزمون‌های کاغذی، عوامل مدیریتی و فنی و در گروه آزمون‌های الکترونیکی، عوامل محیطی و مدیریتی به ترتیب دارای کمترین و بیشترین مقدار میانگین بوده‌اند.

نتایج آزمون t نمونه‌های مستقل برای تمامی ابعاد در سطوح فنی، محیطی و مدیریتی نشان داد که اختلاف معناداری میان میانگین‌های کسب‌شده در دو گروه آزمودنی‌ها (کاغذی و الکترونیکی) وجود داشته است. تفاوت میانگین دو گروه کمتر از صفر است و میانگین گروه دوم (آزمون‌های الکترونیکی) از گروه دوم (آزمون‌های کاغذی) بیشتر است. همچنین، بررسی میانگین نمرات کسب‌شده‌ی دو گروه در هر یک از گویه‌های پرسشنامه نشان داد که این تفاوت در حرفه‌های مختلف برای میانگین به‌دست آمده‌ی کل سؤالات (سطح رضایت‌مندی) نیز وجود دارد و میزان این اختلاف برای همه‌ی حرفه‌ها در دو نوع آزمون تقریباً برابر است.

با توجه ادبیات و یافته‌های تحقیق، به‌طور کلی می‌توان گفت مسیر تحول در ارزشیابی آموزشی در حال گذر از مرحله‌ی آزمون‌های قلم و کاغذی و به‌سوی آزمون‌های مبتنی بر فناوری است که مزایای زیادی در بر دارد. از مهم‌ترین مزایای آن می‌توان به کاهش هزینه‌های آزمون و سرعت اخذ آزمون و استانداردسازی محیط آزمون نام برد.

در حال حاضر، نظام ارزشیابی کشور ایران در مسیر نوآوری در روش‌های سنجش و ارزشیابی است و پیش‌بینی می‌شود در سال‌های آینده سرعت آن در دنیای مبتنی بر فناوری بیشتر شود. یافته‌های این تحقیق که از نمونه‌های آماری اداره‌ی کل آموزش فنی و حرفه‌ای استان البرز به‌دست آمد، نشان می‌دهد آزمون‌های الکترونیکی در حال به‌دست آوردن سهم واقعی خود در ارزشیابی مهارت متقاضیان انواع آزمون‌هاست. این نتایج می‌تواند معیار مناسبی برای سیاست‌گذاران و برنامه‌ریزان در سطوح کلان نظام آموزش و ارزشیابی و مبنایی برای آسیب‌شناسی، کیفیت‌بخشی و توسعه‌ی بیشتر این نوع سنجش مهارت باشد. بنابراین، لازم است مدیران نظام آموزشی درک مناسبی از فرایندها، مزیت‌ها و چالش‌های آزمون‌های مبتنی بر فناوری داشته باشند و بر اهمیت حرکت در مسیر نوآوری و ارتقای کیفیت واقف باشند.

با توجه به ادبیات و یافته‌های این تحقیق پیشنهاد می‌شود:

۱. با بهره‌گیری از دیدگاه‌های کارشناسی و دقیق صاحب‌نظران و خبرگان سیستم ارزشیابی و فناوری اطلاعات، ضمن تدوین نقشه‌ی راه و تدوین راهبردهای توسعه‌ی فناوری در ارزشیابی، چالش‌ها و موانع توسعه‌ی آن شناسایی و راهکارهای عملیاتی برای رفع موانع اندیشیده شود.
۲. برنامه‌ریزی برای پیاده‌سازی الگوهای نوین ارزشیابی و اجرای انواع آزمون‌های مبتنی بر فناوری (رایانه، وب، تلفن همراه و...) برای تمامی انواع آزمون‌های مهارتی کتبی به‌ویژه آزمون‌های هماهنگ، ادواری، صنعت ساختمان و تفاهم‌نامه‌های سازمان آموزش فنی و حرفه‌ای، در تمامی شهرستان‌های کشور صورت پذیرد.
۳. با توجه به تجربه‌ی مناسب آزمون‌های مبتنی بر فناوری در بهبود عملکرد فراگیران، مطالعات لازم راجع به شبیه‌سازی آزمون‌های عملی در محیط واقعی و با استفاده از فناوری برای اجرای آزمون‌های عملی بسیاری از حرفه‌های مهارتی با اولویت حرفه‌های رشته‌ی فناوری اطلاعات و ارتباطات صورت پذیرد.
۴. برای رفع چالش‌های مربوط به ضریب امنیتی و حفاظتی آزمون‌های مبتنی بر رایانه و فناوری در سنجش مهارت متقاضیان، با بهره‌گیری از آخرین دستاوردهای علمی و فناوری، علاوه بر نصب دوربین‌های پیشرفته‌ی مداربسته و سیستم تشخیص چهره‌ی آزمون‌دهنده، نسبت به استقرار ابزارهای نرم‌افزاری و سخت‌افزاری مرتبط اعم از سیستم‌های ثبت وقایع، ابزارهای ارزیابی آسیب‌پذیری، نرم‌افزارهای ضد ویروس، دیواره‌ی آتش و سایر مکانیزم‌های امنیتی فناوری اطلاعات و ارتباطات اقدام، تمهیدات لازم اندیشیده شود.
۵. تعاملات سازمان‌های آموزشی به‌ویژه سازمان آموزش فنی و حرفه‌ای کشور با سازمان‌های پیشرو از جمله وزارت اطلاعات و فناوری ارتباطات با هدف ارتقای سطح فنی و تقویت زیرساخت‌های فناوری سازمان برای اجرای برنامه‌های نوآوری آموزشی و ارزشیابی افزایش یابد.
۶. مطالعات مشابه در مقیاس بزرگ‌تر با هدف شناسایی آسیب‌ها و سطح رضایت‌مندی آزمون‌شوندگان در سیستم سنتی و نوین، برای اخذ تصمیم‌های متناسب با وضعیت هر منطقه، صورت پذیرد.

References

- Ali Mohammadi, T., Mehralizadeh, Y., & Shahi, S. (2011). Evaluation of the third base of the theoretical girl's school of Ahvaz based on the theoretical model of the student from the viewpoints of students and graduates, *Shahid Chamran University of Ahvaz*, 6 (2), 189-208. (Persian).
- Azizi, N., & Heidari, Sh. (2010). The Attitude of Primary Teachers of Sanandaj City toward Descriptive Evaluation, *Journal of Educational Sciences, Shahid Chamran University, Ahvaz*, 2004, 6 (16), 167- 188. (Persian).
- Bennett, R. E., Persky, H., Weiss, A. R., & Jenkins, F. (2007). Problem solving in technology-rich environments: A report from the NAEP Technology-Based Assessment Project. Research and Development Series (NCES 2007-466). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Bodmann, S. M. & Robinson, D. H. (2004). Speed and Performance Differences among Computer-Based and Paper-Pencil Tests. *Journal of Educational Computing Research*, 31 (1), 51 – 60.
- Bull, J. (1999). Computer – Assisted Assessment: Impact on Higher Education Institutions. *Educational Technology & Society*, 2 (3). Retrieved December 05, 2004 from <http://ifets.ieee.org/periodicals>.
- Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: Key factors associated with the test mode effect. *British Journal of Educational Technology*, 33 (5), 593-602.
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42 (2), 21-29.
- Fluck, A., Pullen, D., & Harper, C. (2009). Case study of a computer based examination system. *Australasian Journal of Educational Technology*, 25 (4), 509- 523.
- Havens, A. (2002). Examinations and Learning: An Activity – Theoretical Analysis of the Relationship between Assessment and Learning. Retrieved December 03, 2010 from <http://www.leeds.ac.uk/educol/documents/00002238.htm>.
- Jamil, M., Tariq, R. H., Shami, P. A., & Zakariys, B. (2012). Computer-based vs paper-based examinations: Perceptions of university teachers. Tojet: *The Turkish Online Journal of Educational Technology*, 11 (4). 371-381.
- Karadeniz, S. (2009). The impacts of paper, web and mobile based assessment on students' achievement and perceptions. *Scientific Research and Essay*, 4 (10), 984 – 991. Retrieved May 15, 2011 from <http://www.academicjournals.org/sre>.
- Kikis-Papadakis, K., & Andreas, K. (2009). "Reflections on paper-and-pencil tests to eAssessments: Narrow and broadband paths to 21st century challenges." *The Transition to Computer-Based Assessment* 99.
- Lim, E., CH., Ong, B., KC., Wilder-Smith, E., PV., & Seet, R., CS. (2006).

- Computer-based Versus Pen-and-paper Testing: Students' Perception. *Ann Acad Med Singapore*, 35 (9), 599-603.
- MacCann, R., Eastment, B., & Pickering, S. (2002). Responding to free response examination questions: Computer versus pen and paper. *British Journal of Educational Technology*, 33, 173-188.
- Martin, R. (2008). New possibilities and challenges for assessment through the use of technology, in F. Scheuermann & A. Guimaraes Pereira (Eds) Towards a research agenda in computer-based assessment: Challenges and needs for European Educational Measurement (6-9).
- McKenna, C. (2001). Introducing Computers into Assessment Process: What is the Impact Upon Academic Practice? Paper Presented at Higher Education Close Up Conference 2, Lancaster University, 16 - 18 July. Retrieved November 4, 2004 from <http://leeds.ac.uk/educol/documents/00001805.html>
- Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*. 51 (9), 1352-1375.
- Rehmani, A. (2003). Impact of Public Examination System on Teaching and Learning in Pakistan. Retrieved December 24, 2010 from <http://www.aku.edu/AKUEB/pdfs/pubexam.pdf>.
- Ripley, M. (2008) Technology in the service of 21st century learning and assessment, in F. Scheuermann & A. Guimaraes Pereira (Eds) Towards a research agenda in computer-based assessment: Challenges and needs for European Educational Measurement (22-29).
- Russell, Michael & Amie Goldberg, & Kathleen. (2003). Computer-Based Testing and Validity: A Look Back and Into the Future, Technology and Assessment Study Collaborative, Boston College.
- Scheuermann, F., & Pereira, A. G. (2008). Towards a research agenda on Computer-based Assessment. *Challenges and needs for European Educational Measurement. Luxembourg*.
- Shah, J. H. (2002). Validity and Credibility of Public Examinations in Pakistan. An Unpublished Thesis Submitted for the Degree of Ph. D., in the Department of Education, Islamia University Bahawalpur, Pakistan.
- Sim, G., Holifield, P., & Brown, M. (2004). Implementation of Computer Assisted Assessment: Lessons from the iterature. ALT-J, *Research in Learning Technology*, 12 (3), 217 - 233.
- Solak, E. (2014). Computer versus Paper-Based Reading: A Case Study in English Language Teaching Context. *Online Submission*, 4 (1), 202-211.
- Van Lent, G., & Global, E. T. S. (2009). Risks and benefits of CBT versus PBT in high-stakes testing. *The Transition to Computer-Based Assessment*, 83.
- Van Lent, Gerben, & E. T. S. Global. (2009). "Risks and benefits of CBT versus PBT in high-stakes testing." *The Transition to Computer-Based Assessment*: 83.