

بررسی روائی و پایائی آزمون OSCE/OSPE در ارزیابی مهارت‌های عملی آزمایشگاه بیوشیمی دانشجویان سال اول رشته پرستاری دانشکده پرستاری و مامائی دانشگاه علوم پزشکی شهید بهشتی

سعیده حسینی^۱، ژاکلین وارطانوسیان^۲، آغا فاطمه حسینی^۳، فرح ناز فرزین فرد^۴

۱. مربی و عضو هیئت علمی گروه علوم پایه، دانشکده پرستاری و مامائی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید بهشتی، تهران، ایران
 ۲. کارشناس آزمایشگاه گروه علوم پایه (دانشجوی کارشناسی ارشد آموزش پزشکی دانشگاه علوم پزشکی شهید بهشتی)، دانشکده پرستاری و مامائی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید بهشتی، تهران، ایران
 ۳. آغا فاطمه حسینی، مربی و عضو هیئت علمی گروه آمار ریاضی، دانشکده مدیریت و اطلاع رسانی پزشکی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی ایران، تهران، ایران
 ۴. فرح ناز فرزین فرد، مربی و عضو هیئت علمی گروه روان، دانشکده پرستاری و مامائی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید بهشتی، تهران، ایران
- نویسنده مسئول مکاتبات: ژاکلین وارطانوسیان، ایران، تهران، تقاطع خیابان ولی عصر و بزرگراه نیایش، دانشکده پرستاری و مامائی دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید بهشتی، آدرس پست الکترونیکی: jacqva@gmail.com

چکیده

زمینه و هدف: ارزیابی بالینی/عملی عینی ساختاریافته OSCE/OSPE یکی از شناخته‌شده‌ترین روش‌های ارزشیابی در علوم پزشکی است. کمبود پژوهش در زمینه ارزشیابی مهارت‌های بالینی/عملی علوم پایه موجب شد تا پژوهش حاضر جهت بررسی روائی و پایائی آزمون OSCE/OSPE در ارزیابی مهارت‌های عملی آزمایشگاه بیوشیمی دانشجویان سال اول رشته پرستاری به عمل درآید.

مواد و روش‌ها: پژوهش حاضر یک مطالعه کاربردی از نوع توصیفی است که بر روی ۳۱ نفر از دانشجویان ترم اول رشته پرستاری انجام شد. ابتدا با توجه به اهداف پژوهش و کسب نظرات و پیشنهادات مدرسین آزمایشگاه بیوشیمی، ده ایستگاه جهت آزمون OSCE/OSPE طراحی گردید. برای سنجش روائی محتوا، چک لیست‌ها (برگه‌های آزمون) و پرسش‌نامه‌های آن در اختیار ۴ نفر از متخصصین رشته قرار گرفت، که نظرات آنان در طراحی آزمون اعمال شد. برای تعیین روائی ملاکی، ارتباط نمرات کل آزمون OSCE/OSPE با نمرات کتبی درس بیوشیمی دانشجویان محاسبه شد. جهت بررسی همسانی درونی، همبستگی میان نمره کل آزمون OSCE/OSPE با میانگین نمره دانشجویان در هر یک از ایستگاه‌ها محاسبه شد. به منظور تعیین پایائی آزمون همبستگی بین نمرات گزارش شده توسط دو ارزیاب (پایائی ارزیابان) تعیین شد و نیز از روش بازآزمایی و تکرار آزمون (بعد از ۷۲ ساعت) استفاده گردید.

یافته‌ها: یافته‌ها نشان داد رابطه مستقیمی بین نمرات کل آزمون OSCE/OSPE با نمره کتبی درس بیوشیمی وجود دارد، ولی از نظر آماری معنی‌دار نمی‌باشد. همبستگی به ترتیب برای ارزیاب اول ($r=0/107$) و ($p=0/567$) و برای ارزیاب دوم ($r=0/201$) و ($p=0/227$) بدست آمد. نتایج حاصل از بررسی همسانی درونی، نشان داد که همبستگی معنی‌داری بین نمره کل آزمون با هشت ایستگاه وجود دارد ولی همبستگی معنی‌داری بین دو ایستگاه با نمره کل آزمون حاصل نشد. در تعیین پایائی آزمون نیز همبستگی معنی‌داری بین نتایج آزمون OSCE/OSPE حاصل توسط هر دو ارزیاب بدست آمد ($r=0/859$)

و $p < 0.001$) و نیز همبستگی معنی داری بین نمرات ارزیاب اول در روز آزمون و ۷۲ ساعت بعد (روز سوم) $r = 0.681$ و $p < 0.001$) و نیز همبستگی معنی داری بین نمرات ارزیاب دوم در روز آزمون و ۷۲ ساعت بعد (روز سوم) $r = 0.671$ و $p < 0.001$ بدست آمد. با توجه به ضریب آلفای کرونباخ که عدد 0.907 محاسبه شد پایایی آزمون مورد تأیید است.

نتیجه گیری: استفاده از آزمون OSCE/OSPE به عنوان روشی روا و پایا برای سنجش مهارت‌های عملی آزمایشگاه بیوشیمی توصیه می شود.

کلمات کلیدی: OSCE/OSPE، آزمایشگاه بیوشیمی، روانی، پایایی.

مقدمه

(clinical/practical Examination)

یکی از پسنندیده‌ترین ابزارهای ساختارمند از این نوع است که در مقایسه با آزمون‌های معلم ساخته عادلانه‌تر است (بواله‌ری و همکاران، ۲۰۱۱). در ابتدا آزمون ساختار یافته عینی بالینی (OSCE) و کمی بعد آزمون ساختار یافته عینی عملی (OSPE) بعنوان شیوه‌های ارزیابی مهارت‌های بالینی و عملی توسط دکتر Harden در سال ۱۹۷۹ معرفی شد (تقوا و همکاران، ۲۰۰۷؛ Harden & Glesson، ۱۹۷۹؛ رسولیان و همکاران، ۲۰۰۷). مدرسین مدرسه پرستاری دانشگاه مک مستر نیز در سال ۱۹۸۴ برای اولین بار از آن برای سنجش مهارت‌های پرستاری در مراقبت‌های اولیه دانشجویان سال سوم استفاده کردند (معطری و همکاران، ۲۰۰۷؛ Ross، ۱۹۸۸). این آزمون امکان ارزیابی مهارت‌های گوناگون آزمودنی را در موقعیت‌های متفاوت فراهم می‌کند (رسولیان و همکاران، ۲۰۰۷)، و نتایج حاصل از آن توانایی و صلاحیت فراگیر را به نحو مطلوبی مورد ارزیابی قرار می‌دهد (بواله‌ری و همکاران، ۲۰۱۱) و برخلاف آزمون شفاهی سنتی، گستره‌ای از مهارت‌های بالینی و عملی سنجیده می‌شود (تقوا و همکاران، ۲۰۰۷). امروزه، آزمون ساختار یافته عینی عملی (OSPE) بعنوان وسیله‌ای جهت ارزیابی مهارت‌های عملی قبل و هم راستا با حیطه بالین پذیرفته شده است. هر چند هیچ گونه دستور العمل دقیق و محدود کننده‌ای در نوع سناریوهای به کار گرفته شده در این آزمون وجود ندارد. در انگلستان، آمریکا، کانادا و بیشتر دانشکده‌های پزشکی معتبر این آزمون را شیوه‌ای استاندارد جهت ارزیابی صلاحیت، مهارت‌های بالینی در نظر می‌گیرند که در کنکاش مطالب به‌طور رضایت بخشی می‌تواند مکمل آزمون دانش‌شناختی در بررسی آزمون عینی و کتبی باشد (Abraham و همکاران، ۲۰۰۵). در این راستا گروه بیوشیمی بسیاری از دانشکده‌های پزشکی از این روش به منظور ارزیابی عملکرد دانشجویان در مهارت‌های آزمایشگاهی استفاده می‌کنند (Kunada و همکاران، ۲۰۱۳).

ارزشیابی شایستگی بالینی دانشجویان و کارآیی آن‌ها یکی از مشکل‌ترین وظایف اعضای هیئت علمی و آموزش دهندگان برنامه‌های بهداشتی است (نوحی و همکاران، ۲۰۰۸؛ Wilkinson & Frampton، ۲۰۰۴). برنامه‌ریزی و ارزشیابی آموزش علوم پزشکی مبتنی بر مهارت‌ها است (Carraccio & Englander، ۲۰۰۰)، علی‌رغم در دسترس بودن روش‌های مختلف ارزشیابی بالینی، شواهد نشان می‌دهد که به‌طور معمول ارزیابی دانشجویان محدود به اطلاعات ذهنی بوده و به ارزیابی دقیق مهارت‌های بالینی آن‌ها توجه نشده است. این در حالی است که مهارت و کار عملی در آموزش پزشکی نقش اصلی را داشته و معلومات ذهنی از درجه دوم اهمیت برخوردار است (عزیزی، ۲۰۰۳؛ Karns & Nowotny، ۱۹۹۱؛ نوحی و همکاران، ۲۰۰۸). Lechner، Sisler & Harper و همکاران (Lechner، ۱۹۸۶)، دریافتند که عامل شانس، نوع بیمار و آزمونگر در نتیجه نهایی آزمون نقش اساسی دارد. این نقش به ویژه در نمره دانشجویانی که در مرز رد یا قبول هستند اهمیت می‌یابد. در عین حال افزایش دقت پایانی آزمون شفاهی تن‌ها با بهبود دقت آزمون گر دشوار است، به نظر می‌رسد که سنجش مهارت‌های گوناگون (و نه تن‌ها یک مهارت ویژه) و استفاده از چند ارزیاب باعث بالا رفتن روانی و پایایی آزمون شفاهی خواهد شد (تقوا و همکاران، ۲۰۰۷). یک آزمون خوب زمانی پذیرفتنی است که پایا، روا و عملی باشد (Tooth و همکاران، ۱۹۸۹).

ارزشیابی بالینی در پرستاری نیز، به عنوان یکی از ارکان مهم آموزش، باید شایستگی و توانمندی‌های پرستار را اندازه‌گیری کند، استفاده از روش‌های متفاوت ارزشیابی برای سنجش ابعاد مختلف عملکرد پرستاری کاملاً لازم است (زارعی، ۲۰۰۲؛ معطری و همکاران، ۲۰۰۷). آزمون‌های رایج بالینی از بعد روانی و پایایی محدودیت دارند، در این راستا وجود و گسترش روش‌های عملی ساختار یافته، برخی از این محدودیت‌ها را کاهش داده است. آزمون OSCE/OSPE (Objective structured

حیطه، دانش، نگرش و مهارت دانشجویان را مورد ارزشیابی قرار می‌دهد، تعدادی از چک لیست‌ها مربوط به انجام یک پروسیجر و تعدادی در ارتباط با ارزیابی دانش و اطلاعات دانشجویان بود، به‌طور مثال چک لیست‌های مربوط به تشخیص کیفی قندها، لیپیدها، تشکیل صابون و بیوشیمی ادرار از دسته اول و چک لیست‌های مربوط به تشخیص ابزار و تجهیزات آزمایشگاه و تهیه محلول‌های مولار (حل مسئله) و از گروه دوم است. چک لیست‌های تهیه شده توسط صاحب نظران بازبینی شد و نظرات اصلاحی و پیشنهادی آنان مد نظر قرار گرفت. چک لیست‌های مربوط به انجام یک پروسیجر مرحله به مرحله عملکرد دانشجویان را ارزیابی کرده و هر مرحله امتیاز مستقل داشت، زیرا هر عملکرد شامل چندین مهارت عملی بود و پرسش‌های چک لیست‌ها با گزینه درست (انجام داد / نداد) ارزشیابی شد. چک لیست‌های مربوط به حیطه شناخت و دانش نیز با گزینه (تشخیص داد / نداد) و یا درست (محاسبه کرد / نکرد) مورد ارزشیابی قرار گرفت که این چک لیست‌ها نیز می‌توانست چندین مهارت عملی را ارزیابی کند. به این ترتیب نیز تنظیم چک لیست‌ها و پرسش‌نامه‌ها چندین بار توسط اساتید صاحب نظر انجام شد. فهرست مهارت‌های مورد ارزیابی در ایستگاه‌ها در جدول شماره ۱ آورده شده است.

جدول ۱- مهارت‌های ارزیابی شده در ایستگاه‌های مختلف

آزمون OSCE/OSPE آزمایشگاه بیوشیمی

شماره ایستگاه	مهارت مورد ارزیابی
اول	توانایی تشخیص ابزار و تجهیزات مورد نیاز آزمایشگاه بیوشیمی
دوم	محاسبه مکتوب کلرور سدیم لازم برای تهیه ۲۵۰ میلی لیتر محلول ۰/۱ مولار آن
سوم	توانایی کار با اسپکتروفوتومتر
چهارم	تشخیص میزان دانسیته ادرار (قرائت میزان)
پنجم	توانایی انجام مراحل کار برای تشخیص پلی ساکراید (واکنش ید)
ششم	توضیح مکتوب مکانیسم آزمایش بارفود برای تشخیص منوساکراید
هفتم	توانایی انجام مراحل کار برای تشخیص قندهای کتوز (آزمایش سلیوانف)
هشتم	ارزیابی قدرت حلالیت چربی‌ها در حلال‌های مختلف
نهم	توانایی انجام مراحل کار برای تشخیص قابلیت صابونی شدن چربی‌ها
دهم	توضیح مکتوب مکانیسم آزمایش ایستگاه نهم

در ایران مسئله آزمون بالینی و آزمایشگاهی هم چنان از مشکلات لاینحل است و نبود استانداردهای مشخص کار در محیط بالینی و نبود اهداف مشخص و مدونی که ویژگی‌های اختصاصی بودن، قابل دستیابی بودن و قابل اندازه‌گیری بودن، مناسب بودن و زمان بندی لازم را داشته باشد، روند یادگیری دانشجویان را تحت تأثیر قرار داده است (معطری و همکاران، ۱۳۹۲)، از طرف دیگر با توجه به گسترش برنامه‌های بازآموزی در دانشگاه‌های ایران، ضرورت ارزیابی دقیق این برنامه‌ها و پژوهش درباره ی آن‌ها به عنوان یک نیاز احساس می‌شود (بحری و همکاران، ۲۰۰۹). با علم به اینکه OSCE/OSPE در ایران در رشته‌های گوناگون مانند پزشکی اجتماعی، روان پزشکی، طب کودکان و پرستاری برگزار می‌گردد، و تاکنون تجربه‌ای از OSCE/OSPE در ارزیابی مهارت‌های بالینی/عملی در علوم پایه گزارش نشده و یا بسیار محدود است، اهمیت و لزوم برگزاری آزمون معتبر برای ارزشیابی مهارت‌های بالینی علوم پایه باعث شد تا پژوهش حاضر جهت بررسی روائی و پایائی آزمون OSCE/OSPE در ارزیابی مهارت‌های عملی آزمایشگاه بیوشیمی دانشجویان سال اول رشته پرستاری به عمل آید.

مواد و روش‌ها

این تحقیق یک مطالعه کاربردی از نوع توصیفی است. پس از کسب نظرات داوران محترم و پس از اتمام جلسات آزمایشگاه بیوشیمی، دانشجویان رشته پرستاری ورودی ۱۳۸۹ دانشکده پرستاری و مامائی شهید بهشتی مورد ارزشیابی قرار گرفتند. ابتدا فهرستی از کلیه مهارت‌های آزمایشگاه بیوشیمی تهیه شد و برای این کار از کتاب‌های مرجع و مرتبط با رئوس واحد آزمایشگاه بیوشیمی استفاده شد. برای طراحی ایستگاه‌ها ابتدا چند نشست با حضور تعدادی از مدرسین آزمایشگاه بیوشیمی در دانشگاه‌های مختلف برگزار شد. در طی این جلسات چند هدف عملی و کاملاً قابل اجرا در مدت زمان مشخص (ده دقیقه) تعیین شد. این اهداف همان هدف‌های کلی آموزش عملی واحد بیوشیمی بود که توسط دبیرخانه شورای تخصصی وزارت بهداشت درمان و آموزش پزشکی تهیه شده بود. سپس در نشست‌های متعدد امکانات عملی ایستگاه‌ها با هدف‌های تعیین شده با توجه به مکان آزمون سازگار شد. سرانجام آزمون OSCE/OSPE آزمایشگاه بیوشیمی در ده ایستگاه که زمان هر ایستگاه بطور میانگین ۸ دقیقه بود طراحی شد. آنگاه چک لیست‌های مهارت‌هایی که دانشجویان در طی نیم سال تحصیلی کسب کرده‌اند تهیه شد و از آن جایی که آزمون OSCE/OSPE در سه

Test re- test) بود که از مقایسه نتایج روز سوم (۷۲ ساعت پس از انجام آزمون) با دو روز گذشته بدست آمد. به‌طوری که همان آزمون و پرسشنامه‌ها که در روز اول و دوم به دانشجویان داده شده بود در روز سوم مجدداً به همان دانشجویان داده شد و آزمون به همان روش دو روز گذشته انجام شد و همبستگی بین نتایج روز سوم با دو روز گذشته از طریق محاسبه ضریب همبستگی پیرسون تعیین گردید. قابل ذکر است که چون در ابتدا آزمون کولموگروف-اسمیرنوف، توزیع داده‌ها را نرمال نشان داد، استفاده از ضریب همبستگی پیرسون اعمال شد.

برای تعیین پایایی بین دو ارزیاب (Examiner's reliability) و پایایی بازآزمایی آزمون نیز از ضریب همبستگی پیرسون استفاده شد.

برای تعیین روانی از دو روش روانی محتوا (Content validity) و روانی ملاکی (Criterion-related validity) استفاده شد. برای سنجش روانی محتوا، پرسشنامه آزمون، همراه با اهداف پژوهش در اختیار ۴ نفر از متخصصین رشته قرار گرفت و سپس نظرات متخصصین در طراحی آزمون نهایی اعمال شد. برای تعیین روانی ملاکی ارتباط نمرات آزمون OSCE/OSPE با نمرات آزمون کتبی درس بیوشیمی دانشجویان محاسبه شد. برای تعیین همسانی درونی، همبستگی نمرات کل آزمون OSCE/OSPE با نمرات حاصل از هر ایستگاه بدست آمد. برای تعیین روانی ملاکی و همسانی درونی از ضریب همبستگی پیرسون استفاده شد.

یافته‌ها

یافته‌های حاصل از تجزیه و تحلیل داده‌ها نشان داد که ۴۸/۴ (۱۵ نفر) درصد نمونه‌ها پسر و ۵۱/۶ (۱۶ نفر) درصد نمونه‌ها دختر بودند. دامنه سنی دانشجویان ۱۸ سال تا ۲۲ سال با میانگین ۱۹/۷۶ سال و انحراف معیار ۱/۰۹ سال بود. ۹۶/۲ (۲۵ نفر) درصد نمونه‌ها مجرد و ۳/۸ (یک نفر) درصد متأهل بودند. ۴۴/۴ (۸ نفر) درصد نمونه‌ها با سهمیه منطقه یک، ۳۸/۹ (۷ نفر) درصد منطقه دو، ۱۱/۱ (۲ نفر) درصد منطقه سه و ۵/۶ (یک نفر) درصد از سایر سهمیه‌های کنکور سراسری شرکت کرده بودند.

یافته‌ها نشان داد که رابطه مستقیمی بین نمرات کل آزمون OSCE/OSPE با نمره کتبی درس بیوشیمی وجود دارد، ولی از نظر آماری معنی‌دار نیست. ضریب همبستگی پیرسون به ترتیب برای ارزیاب اول ($r=0/107$ و $p=0/567$) و برای ارزیاب دوم ($r=0/201$ و $p=0/277$) بدست آمد.

ده روز قبل از آزمون OSCE/OSPE جلسه توجیهی برای کلیه دانشجویان برگزار شد و راهنمای آزمون در اختیار تک تک آن‌ها قرار گرفت. در این راهنما مواردی مانند تعداد ایستگاه‌ها، زمان امتحان، خاموش کردن و تحویل تلفن همراه و ... قید شده بود. این آزمون توسط سه نفر ارزیاب (ممتحن) و یک پرسنل خدماتی در سه روز متوالی انجام شد. با توجه به این که زمان انجام آزمون با این تعداد دانشجو در ده ایستگاه و توسط دو ارزیاب در یک روز به درازا می‌کشید، بنابراین، روز اول آزمون پنج ایستگاه، روز دوم نیز پنج ایستگاه طراحی شد. روز سوم (۷۲ ساعت بعد) جهت بررسی پایایی آزمون، بازآزمایی مهارت‌های روزهای قبل (در ده ایستگاه) انجام شد. بدیهی است که دانشجویان از تکرار مهارت‌های عملی در روز سوم بی اطلاع بودند، زیرا در جلسه ی توجیهی فقط از برگزاری آزمون در سه روز آگاه می‌شدند. در ورودی هر ایستگاه مراحل کار به‌طور دقیق چسبانده شده بود و دانشجو فرصت کافی برای مطالعه داشت. برای تعیین زمان در هر ایستگاه از ساعت‌های زنگ‌دار آزمایشگاه استفاده می‌شد و دانشجویان با شنیدن صدای زنگ متوجه می‌شدند که باید به ایستگاه بعدی بروند. در روز آزمون پس از اطمینان از حضور کلیه دانشجویان، تلفن همراه شان در پاکت‌هایی که به تعداد دانشجویان با شماره و بر چسب مهیا شده بود، جمع‌آوری و تا آخر آزمون به آن‌ها تحویل داده نمی‌شد. آنگاه دانشجویان به‌طور داوطلب وارد آزمایشگاه بیوشیمی شده و از ایستگاه شماره یک شروع می‌نمودند. با ورود دانشجو به هر ایستگاه دو ارزیاب به‌طور همزمان عملکرد دانشجو را به‌طور جداگانه و توسط چک لیست‌ها گام به گام ارزشیابی کرده و با زدن علامت، چک لیست هر ایستگاه را کامل می‌کردند. ارزیاب سوم (ممتحن سوم) ورود و خروج دانشجویان را کاملاً کنترل می‌کرد تا دانشجویانی که آزمون را داده‌اند هنگام خروج هیچ گونه تماسی با دانشجویان دیگر نداشته باشند (دانشجویان کاملاً قرنطینه بودند). ارزیاب سوم در خاتمه آزمون در هر روز نمرات دو ارزیاب را جمع، کنترل و مجموع آن‌ها را در هر چک لیست یادداشت کرده و در روز سوم پس از اتمام آزمون فرم نظر سنجی را به دانشجویان داده و از درست پر شدن آن توسط آنان اطمینان حاصل کرد. فرم نظرسنجی شامل اطلاعات دموگرافیک، میزان رضایت دانشجو از نحوه آزمون، میزان استرس و بود. در روز سوم به علت طولانی شدن امتحان، پذیرایی مختصری از دانشجویان به عمل آمد.

برای تعیین پایایی آزمون با تاکید بر همسانی درونی از ضریب آلفای کرونباخ استفاده شد. به همین ترتیب دومین روشی که برای تعیین پایایی آزمون به کار گرفته شد روش آزمون مجدد

جهت بررسی همسانی درونی (جدول شماره ۲)، از ضریب همبستگی پیرسون بین نمرات حاصل از ایستگاه‌های مختلف با نمره کل آزمون OSCE/OSPE استفاده شد که به ترتیب با ایستگاه اول ($r=0/416$ و $p=0/020$)، و ایستگاه دوم ($r=0/479$ و $p=0/006$)، با ایستگاه سوم ($r=0/450$ و $p=0/011$)، با ایستگاه چهارم ($r=0/501$ و $p=0/004$)، با ایستگاه پنجم ($r=0/384$ و $p=0/033$)، با ایستگاه ششم ($r=0/732$ و $p<0/001$)، با ایستگاه هفتم ($r=0/370$ و $p=0/040$) و با ایستگاه نهم ($r=0/522$ و $p=0/003$) بدست آمد. ولی با ایستگاه‌های هشتم ($r=0/080$) و دهم ($p=0/667$ و $r=0/305$) معنی‌دار نبود که این دو ایستگاه حوزه روان حرکتی دانشجویان را اندازه‌گیری نمی‌کرد.

جدول ۲- ماتریس ضرایب همبستگی بین نمرات بدست آمده در ایستگاه‌های مختلف
با نمره کل آزمون OSCE/OSPE

ایستگاه	ایستگاه اول	ایستگاه دوم	ایستگاه سوم	ایستگاه چهارم	ایستگاه پنجم	ایستگاه ششم	ایستگاه هفتم	ایستگاه هشتم	ایستگاه نهم	ایستگاه دهم	نمره کل آزمون OSCE/OSPE
ضریب همبستگی	۱	۰/۰۷۴	۰/۱۸۰	۰/۳۱۲	۰/۲۳۹	۰/۲۴۹	۰/۱۸۸	۰/۱۹۱	۰/۰۴۹	۰/۱۳۴	۰/۴۱۶ *
سطح معنی داری	-----	۰/۶۹۴	۰/۳۳۲	۰/۸۸	۰/۱۹۵	۰/۱۷۷	۰/۳۱۲	۰/۳۰۴	۰/۷۹۴	۰/۴۷۱	۰/۰۲۰
تعداد (ایستگاه اول)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۰۷۴	۱	۰/۰۲۳	۰/۳۱۶	۰/۱۱۰	۰/۱۴۲	۰/۴۰۱ *	۰/۰۴۸	۰/۱۳۷	۰/۱۰۸	۰/۴۷۹ **
سطح معنی داری	۰/۶۹۴	-----	۰/۹۰۲	۰/۰۸۳	۰/۵۵۶	۰/۴۴۵	۰/۰۲۵	۰/۸۰۰	۰/۴۶۳	۰/۵۶۴	۰/۰۰۶
تعداد (ایستگاه دوم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۱۴۴	۰/۲۸۲	۱	۰/۱۰۳	۰/۲۳۸	۰/۳۰۵	۰/۱۳۰	۰/۰۷۲	۰/۰۶۷	۰/۲۸۳	۰/۴۵۰ *
سطح معنی داری	۰/۴۳۹	۰/۱۲۴	-----	۰/۵۸۰	۰/۱۹۷	۰/۰۹۶	۰/۴۸۶	۰/۷۰۰	۰/۷۲۲	۰/۱۲۳	۰/۰۱۱
تعداد (ایستگاه سوم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۳۱۲	۰/۰۸۳	۰/۰۱۳	۱	۰/۵۱۹ **	۰/۴۱۶ *	۰/۱۹۵	۰/۰۱۵	۰/۰۸۵	۰/۴۸۳ **	۰/۵۰۱ **
سطح معنی داری	۰/۰۸۸	۰/۰۸۳	۰/۹۴۵	-----	۰/۰۰۳	۰/۰۲۰	۰/۲۹۳	۰/۹۳۷	۰/۶۵۰	۰/۰۰۶	۰/۰۰۴
تعداد (ایستگاه چهارم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۲۳۹	۰/۱۹۵	۰/۱۴۰	۰/۵۱۹ **	۱	۰/۱۸۴	۰/۱۶۲	۰/۰۳۷	۰/۱۶۵	۰/۰۸۴	۰/۳۸۴ *
سطح معنی داری	۰/۱۹۵	۰/۵۵۶	۰/۴۵۲	۰/۰۰۳	-----	۰/۳۲۲	۰/۳۸۳	۰/۸۴۲	۰/۳۷۵	۰/۶۵۲	۰/۰۳۳
تعداد (ایستگاه پنجم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۲۴۹	۰/۱۴۲	۰/۱۴۸	۰/۴۱۶ *	۰/۱۸۴	۱	۰/۰۱۵	۰/۰۳۵	۰/۱۶۳	۰/۴۱۶ *	۰/۷۳۲ **
سطح معنی داری	۰/۱۷۷	۰/۴۴۵	۰/۴۲۸	۰/۰۲۰	۰/۳۲۲	-----	۰/۹۳۸	۰/۸۵۰	۰/۳۸۲	۰/۰۲۰	۰/۰۰۰
تعداد (ایستگاه ششم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۱۸۸	۰/۴۰۱ *	۰/۲۵۴	۰/۱۹۵	۰/۱۶۲	۰/۰۱۵	۱	۰/۲۴۷	۰/۳۱۰	۰/۱۸۳	۰/۳۷۰ *
سطح معنی داری	۰/۳۱۲	۰/۰۲۵	۰/۱۶۹	۰/۲۹۳	۰/۹۳۸	۰/۹۳۸	-----	۰/۱۸۰	۰/۰۹۰	۰/۳۲۵	۰/۰۴۰
تعداد (ایستگاه هفتم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۱۹۱	۰/۰۴۸	۰/۱۰۱	۰/۰۱۵	۰/۰۳۷	۰/۰۳۵	۰/۲۴۷	۱	۰/۰۷۵	۰/۰۱۵	۰/۰۸۰
سطح معنی داری	۰/۳۰۴	۰/۸۰۰	۰/۵۸۸	۰/۹۳۷	۰/۸۴۲	۰/۸۴۲	۰/۱۸۰	-----	۰/۶۸۹	۰/۹۳۷	۰/۶۶۷
تعداد (ایستگاه هشتم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۰۸۹	۰/۰۴۵	۰/۰۶۷	۰/۵۲۲ **	۰/۲۳۹	۰/۲۵۸	۰/۰۹۵	۰/۲۹۳	۱	۰/۳۹۲ *	۰/۵۲۲ **
سطح معنی داری	۰/۶۳۵	۰/۸۱۲	۰/۷۷۲	۰/۰۰۳	۰/۱۹۵	۰/۱۶۱	۰/۶۱۳	۰/۱۰۹	-----	۰/۰۲۹	۰/۰۰۳
تعداد (ایستگاه نهم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱
ضریب همبستگی	۰/۱۳۴	۰/۰۸	۰/۴۱۷ *	۰/۴۸۳ **	۰/۰۸۴	۰/۴۱۶ *	۰/۱۸۳	۰/۰۱۵	۰/۳۳۹	۱	۰/۳۰۵
سطح معنی داری	۰/۴۷۱	۰/۵۶۴	۰/۰۲۰	۰/۰۰۶	۰/۶۵۲	۰/۰۲۰	۰/۳۲۵	۰/۹۳۷	۰/۰۶۲	-----	۰/۰۹۵
تعداد (ایستگاه دهم)	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱	۳۱

*در سطح معنی‌داری ۰/۰۵ است. **در سطح معنی‌داری ۰/۰۱ است.

بحث و نتیجه گیری

براساس نتایج پژوهش حاضر، آزمون OSCE/OSPE جهت ارزشیابی مهارت‌های عملی آزمایشگاه بیوشیمی از روانی و پایایی بالایی برخوردار است که با سایر مطالعات در حیطه پزشکی (Brown و همکاران، ۱۹۹۹؛ تقوا و همکاران، ۲۰۰۷؛ معطری و همکاران، ۲۰۰۷؛ Raj و همکاران، ۲۰۰۷) هم خوانی دارد. برای بررسی روانی ملاکی آزمون OSCE/OSPE نتایج نشان داد که رابطه مستقیمی بین نمرات کل آزمون OSCE/OSPE با نمره کتبی درس بیوشیمی وجود دارد ولی از نظر آماری معنی‌دار نیست. ضریب همبستگی پیرسون برای ارزیابی اول $r=0/107$ و $p=0/517$ و برای ارزیابی دوم $r=0/201$ و $p=0/277$ حاصل شد. Raj و همکاران نیز در سال ۲۰۰۷ از دو آزمون OSCE روماتولوژی برای ارزشیابی مهارت‌های عملی دست و زانو استفاده کردند. برای تعیین روانی همزمان، نمرات ۵۰ نفر از دانشجویان سال پنجم را با نمرات امتحان نهایی روماتولوژی سنجیدند. نتایج حاصل از ایستگاه‌های دست با نمرات کل OSCE/OSPE روماتولوژی همبستگی متوسطی داشت ($r=0/44$ و $p<0/01$) ولی همبستگی بین نمره کتبی روماتولوژی در قسمت دست و نمره نهایی OSCE/OSPE مشاهده نشد ($r=0/01$ و $p=0/93$). در بررسی مهارت ایستگاه‌های زانو نیز همبستگی متوسطی بین نتایج کل OSCE/OSPE سال پنجم با نمره نهایی امتحان OSCE/OSPE روماتولوژی مشاهده شد ($r=0/46$ و $p<0/01$) ولی بین نمرات OSCE/OSPE با نمرات کتبی کل آزمون روماتولوژی همبستگی معنی‌داری پیدا نکردند ($r=-0/04$ و $p=0/79$). پژوهش تقوا و همکاران در سال ۲۰۰۷ در روان پزشکی نیز نشان داد که نتایج OSCE/OSPE با نمره کتبی دانشنامه دستیاران سال سوم معنی‌دار نیست ولی با نمره شفاهی دانشنامه آنان که مهارت دستیار را می‌سنجد همبستگی معنی‌داری وجود دارد، پژوهش معطری و همکاران (۲۰۰۷) نیز در پرستاری حاکی از آن است که بین نمرات آزمون OSCE/OSPE مهارت‌های پرستاری با نمرات دروس نظری پرستاری ارتباط مثبت و ضعیفی وجود دارد. در حالی که بین نمرات OSCE/OSPE مهارت‌های پرستاری با نمرات دروس بالینی پرستاری ارتباط مثبت و معنی‌داری وجود دارد. Gerrow و همکاران (۲۰۰۳) به منظور بررسی روانی همزمان آزمون OSCE/OSPE دندان پزشکی در کانادا، ارتباط مهارت‌های عملی دندان پزشکی را در ۲۳۱۷ دانشجوی سال آخر دندان پزشکی را با نمرات کتبی مورد آزمون دندان پزشکی ملی کانادا به مدت ۵

سال (۱۹۹۵ تا ۲۰۰۰) سنجیدند. یافته‌ها دلالت بر آن داشت که ارتباط مثبت و معنی‌داری بین نمرات کتبی مورد آزمون دندان پزشکی با نتایج آزمون OSCE/OSPE وجود دارد ($r=0/54$ و $p<0/001$). به نظر می‌رسد اختلاف بین نتایج پژوهش حاضر و تحقیق Gerrow، اولاً از تعداد زیاد نمونه‌ها و نیز طول مدت تحقیق (۵ سال) ناشی شود، ثانیاً تعداد و نوع سؤالات امتحان کتبی مورد باشد. Gerrow و همکاران در امتحان کتبی حداقل ۳۰۰ سؤال از انواع مختلف تک جوابی، چهار جوابی و مقایسه ای طرح کرده بودند. به هر حال شاید برای بررسی روانی ملاکی یا همزمان OSCE/OSPE با سایر آزمون‌ها، اگر از یک آزمون شفاهی و یا بالینی به عنوان معیار استفاده شود، نتایج بین دو آزمون، ارتباط مثبت و معنی‌دار قوی تری را نشان دهد. وجود شکاف بین دروس نظری و عملی همواره به عنوان چالشی فراروی صاحب نظران امر آموزش است. با در نظر گرفتن این نکته که آزمون OSCE/OSPE باید هر سه حیطه دانش، مهارت و نگرش را مورد ارزشیابی قرار دهد، پیدا کردن معیاری که هر سه حیطه به ویژه حیطه نگرشی را به‌طور مستقیم پوشش دهد عملاً دشوار است (معطری و همکاران، ۲۰۰۷). به این ترتیب با توجه به تعدادی از مطالعات که آن‌ها نیز در طراحی ایستگاه‌ها از حل مسئله، فرمولاسیون، طبقه بندی و استفاده کرده بودند (تقوا و همکاران، ۲۰۰۷؛ معطری و همکاران، ۲۰۰۷؛ Kunada و همکاران، ۲۰۱۳)، در پژوهش حاضر نیز سه ایستگاه به تفسیر اطلاعات، توضیح مکانیسم و اختصاص داده شد.

جهت بررسی همسانی درونی، همبستگی بین نمره کل آزمون OSCE/OSPE با میانگین نمره دانشجو در هر یک از ایستگاه‌ها بدست آمد که از ده ایستگاه ضریب همبستگی دو ایستگاه (ایستگاه‌های ۸ و ۱۰) معنی‌دار نبود. مهارت‌های عملی این دو ایستگاه حوزه روان حرکتی دانشجو را اندازه‌گیری نمی‌کرد. معنی‌دار بودن اکثر ایستگاه‌ها با نمره کل آزمون OSCE/OSPE می‌تواند نشان دهنده این باشد که دانشجو در طی سال مهارت‌های لازم را در آزمایشگاه کسب کرده و در همه جنبه‌های تعیین شده رشد کرده است. بیشترین همبستگی در ایستگاه ششم ($r=1/732$) و کمترین همبستگی در ایستگاه هفتم ($r=0/416$) بدست آمد (جدول شماره ۲). این نتایج می‌تواند به چگونگی طراحی ایستگاه‌ها ارتباط داشته باشد. مثلاً از آن جایی که ایستگاه ششم توضیح مکانیسم یک آزمایش می‌باشد، این نتیجه می‌تواند به تأکید مدرس در جلسه توجیهی

ایستگاه‌های زانو پایانی ارزیاب اول $r=0/91$ و $p<0/001$ و برای ارزیاب دوم $r=0/69$ و $p<0/03$ بدست آمد. Brown و همکاران (۱۹۹۹) آزمون OSCE/OSPE دندان پزشکی دانشجویان سال چهارم را به منظور بررسی بازآزمایی در نوبت صبح انجام و در نوبت عصر تکرار کردند که نتایج نشان دهنده پایانی بالایی بود. Roberts و همکاران (۱۹۹۰) نیز پایانی بازآزمایی آزمون OSCE/OSPE مهارت‌های عملی پرستاری را در ۷۱ دانشجوی سال اول در ده ایستگاه انجام دادند. برای بررسی پایانی بازآزمایی امتحان مجدد بلافاصله پس از اتمام آزمون انجام شد که پایانی آزمون خوب گزارش شد ($r=0/66-0/86$). بدین ترتیب ضریب آلفای کرونباخ آزمون OSCE/OSPE پژوهش حاضر عدد $0/907$ بدست آمد که مؤند پایانی آزمون است. مطالعات فوق نشان می‌دهد که برای تعیین بازآزمایی، با توجه به نوع آزمون و امکانات موجود می‌توان آزمون را بلافاصله، روز بعد، دو ماه بعد و یا مثل پژوهش حاضر ۷۲ ساعت پس از آزمون تکرار کرد. به هر حال چون ضرایب پایانی میزان خطای تصادفی اندازه گیری‌ها را در ارزیابی‌ها نشان می‌دهد با تعیین پایانی ارزیابی، آزمون می‌تواند روند بهبود و پیشرفت را طی کند (Feldt و همکاران، ۱۹۸۹).

اگر چه OSCE/OSPE یک روش مؤثر و کارآمد برای ارزیابی شایستگی بالینی/عملی دانشجویان است اما برای سنجش همه جنبه‌های بالینی/عملی مناسب نیست و عناصری مانند نیروی انسانی مجرب، منابع، امکانات و زمان بر بودن به عنوان مشکلات اجرایی آن مطرح می‌باشد. همچنین موفقیت آن به برنامه‌ریزی دقیق، مکان و ابزار اندازه‌گیری مناسب و امکانات برای ارزیابی دانشجو وابسته است (Wilkinson و همکاران، ۲۰۰۴)، و تأمین اعتبار مناسب، افزایش شمار داوطلبان شرکت کننده، همراه با تجربه اندوزی مدرسين به اعتبار بیشتر آن کمک خواهد کرد.

با توجه به یافته‌های پژوهش حاضر، می‌توان مدعی شد به علت حذف شانس، امکان بازخورد و عادلانه‌تر بودن نتایج OSCE/OSPE، علی‌رغم مشکلاتی مثل فضای فیزیکی مناسب برای تشکیل ایستگاه‌ها و قرنطینه شدن دانشجویان و از همه مهم‌تر لزوم حضور و مشارکت فعال ارزیابان در طول آزمون، روشی مطمئن برای ارزشیابی مهارت‌های عملی آزمایشگاه بیوشیمی برای دانشجویان رشته پرستاری است و پژوهشگران استفاده از آن را برای ارزشیابی کلیه آزمایشگاه‌های علوم پایه توصیه می‌نمایند.

در پایان لازم به یادآوری است که پس از اتمام پژوهش، برای ارزیابی نتایج متوالی آزمون OSCE/OSPE ارزشیابی نهایی

آزمون OSCE/OSPE به گنجانیدن نکات کتبی در آزمون مربوط باشد. نتیجه حاصل از ایستگاه هفتم نیز می‌تواند از دشواری دستورالعمل طراحی شده در ایستگاه ناشی شود و یا احیاناً استفاده از ابزارهای خاصی (مثل پیپتور آکسفورد و غیره در این ایستگاه) باشد. در تجارب بعدی آزمون OSCE/OSPE با در نظر گرفتن این نتایج سعی خواهد شد امتیازها در این ایستگاه‌ها تعدیل شوند تا تمام ایستگاه‌ها از یکپارچگی درونی کافی برخوردار شوند. پژوهش معطری و همکاران (۲۰۰۷) نیز که مهارت‌های پرستاری را آزمودند از ده ایستگاه فقط دو ایستگاه از همسانی درونی مطلوبی برخوردار نبود. نتایج پژوهش تقوا و همکاران (۲۰۰۷) که مهارت‌های روان پزشکی را در نه ایستگاه مورد آزمون قرار دادند، نشان داد که فقط در یک ایستگاه بین میانگین نمره کل هر دستیار با نمره وی در هر ایستگاه ارتباط معنی‌داری وجود ندارد.

جهت تعیین پایانی بین ارزیابان، همبستگی نمرات ارزیابان محاسبه شد که همبستگی معنی‌داری بین نتایج آزمون OSCE/OSPE توسط دو ارزیاب بدست آمد ($r=0/859$ و $p<0/001$). این نتیجه با نتایج Raj و همکاران (۲۰۰۷) مطابقت دارد، آن‌ها نیز از دو نفر ارزیاب در آزمون روماتولوژی کمک گرفته بودند که برای ایستگاه‌های دست ($r=0/86$ و $P<0/001$) و برای ایستگاه‌های زانو ($r=0/90$ و $p<0/001$) حاصل شد. Roberts و همکاران (۱۹۹۰) آزمون OSCE/OSPE دندان پزشکی را در ده ایستگاه بوسیله دو ارزیاب انجام دادند که پایانی بین ارزیابان وجود داشت، و به صورت $r=0/8-0/99$ گزارش شده است. Brown و همکاران (۱۹۹۹) آزمون OSCE/OSPE دندان پزشکی را در ۱۷ ایستگاه بوسیله دو ارزیاب انجام دادند که پایانی بین ارزیاب‌ها نسبتاً بالا بدست آمد.

در پژوهش حاضر برای بررسی پایانی آزمون از بازآزمایی نیز استفاده شد که همبستگی معنی‌داری برای ارزیاب اول در روز آزمون و ۷۲ ساعت بعد (روز سوم) و همبستگی معنی‌داری بین نمرات ارزیاب دوم در روز آزمون و ۷۲ ساعت بعد (روز سوم) بدست آمد. به ترتیب برای ارزیاب اول ($r=0/681$ و $p<0/001$) و برای ارزیاب دوم ($r=0/671$ و $p<0/001$). این نتایج با پژوهش‌های Roberts (۱۹۹۰)، Raj (۲۰۰۷) و Brown (۱۹۹۹) همخوانی دارد. Raj و همکاران (۲۰۰۷) برای بررسی پایانی بازآزمایی، فیلم دو امتحان OSCE روماتولوژی (دست و زانو) را پس از دو ماه توسط ارزیابان مجدداً امتیاز دادند که پایانی برای ارزیاب اول در ایستگاه‌های دست $r=0/94$ و $p<0/001$ و برای ارزیاب دوم $r=0/73$ و $p<0/001$ حاصل شد.

آزمایشگاه بیوشیمی دانشجویان پرستاری دانشگاه علوم پزشکی شهید بهشتی در هر نیمسال با این روش برگزار می‌گردد که نتایج حاصل مورد توجه قرار خواهد گرفت.

تشکر و قدردانی

این مقاله حاصل طرح پژوهشی به شماره ۷۸۲۵-۸۶-۱-۹۰ و

تاریخ ۸۹/۳/۲۱ دانشکده پرستاری و مامایی دانشگاه علوم پزشکی شهید بهشتی بوده است. پژوهشگران بدین وسیله مراتب سپاس و قدردانی خود را از معاونت محترم پژوهش سرکار خانم دکتر معصومه سیمبر، مدیر محترم گروه علوم پایه جناب آقای دکتر هومان منوچهری و کلیه دانشجویان ترم اول رشته پرستاری که در انجام این پژوهش مساعدت نمودند اعلام می‌دارند.

REFERENCES

- Abraham RR et al (2005). Student perspectives of assessment by TEMM model in physiolog . *Advances in Physiology Education*. 29 (2) 94-97.
- Azizi F (2003). [Medical Education, Mission, Vision and Challenges]. 1st ed. Tehran, Educational chancellor of ministry of health and medical education. (Persian)
- Bahri N et al (2009). [Participants attitude of on a job training in regards to the evaluation of the course with OSCE]. *Iranian Journal of Medical Education*. 9 (2) 181-182. (Persian)
- Bolhari J et al (2011). [OSCE Instructional Guideline in Psychiatry]. 1st ed. Tehran, Arjomand publication. (Persian)
- Brown Q Manogue M Martin M (1999). The validity and reliability of an OSCE in dentistry. *European Journal of Dental Education*. 3 (3) 117-25.
- Carraccio C Englander R (2000). The objective structured clinical examination. *Archives of Pediatrics and Adolescent Medicine*. 154 (7) 736-741.
- Feldt LS et al (1989). *Reliability. Educational Measurement*. 3rd edition. New York, American Council on Education.
- Gerrow JD et al (2003). Concurrent validity of written and OSCE components of the Canadian dental certification examinations. *Journal of Dental Education*. 67 (8) 896-901.
- Harden RM Gleeson FA (1979). Assessment of clinical competence using an objective structured clinical examination (OSCE). *Medical Education*. 13 (1) 41-54.
- Kundu D et al (2013). Objective structured practical examination in biochemistry: An experience in Medical College, Kolkata. *Journal of Natural Science, Biology and Medicine*. 4 (1) 103-107.
- Karns P Nowotny M (1991). Clinical structure and evaluation in baccalaureate schools of nursing. *Journal of Nursing Education*. 30 (5) 207-11.
- Leichner P Sisler GC Harper D (1986). The clinical oral examination in psychiatry: The patient variable. *Annals of the Royal College of Physician and Surgeons of Canada*. 19 283-284.
- Moattari M et al (2007). [Reliability and validity of OSCE in evaluating clinical skills of nursing students. *Pajouhesh*]. 31 (1) 55-59. (Persian)
- Noohi E Motsadi M Haghdost A (2008). [Clinical teachers viewpoints towards objective structured clinical examination in Kerman university of medical science]. *Iranian Journal of Medical Education*. 8 (1) 113-119. (Persian)
- Raj N et al (2007). Design and validation of 2 objective structured clinical examination stations to assess core undergraduate examination skills of the hand and knee. *The Journal of Rheumatology*. 34 (2) 421-4.
- Rasoulia M et al (2007). [Qualitative assessment of the first objective structured clinical examination (OSCE) in psychiatry in Iran]. *Iranian Journal of Psychiatry and Clinical Psychology*. 13 (1) 12-16. (Persian)
- Roberts J et al (1990). Reliability and learning from the objective structured clinical examination. *Medical Education*. 24 (3) 219-223.

- Ross M et al (1988). Using the OSCE to measure clinical skills performance in nursing. *J Advanc Nurs*. 13 (1) 45-56.
- Taghva A et al (2007). [Validity and reliability of the first objective structured clinical examination (OSCE) in psychiatry in Iran]. *Iranian Journal of Psychiatry and Clinical Psychology*. 13 (1) 17-24. (Persian)
- Tooth D Tonge K McManus (1989). Anxiety and study methods in preclinical students: causal relation to examination performance. *Medical Education*. 23 (5) 416-421.
- Wilkinson TJ Frampton CM (2004). Comprehensive undergraduate medical assessments improve prediction of clinical performance. *Medical Education*. 38 (10) 1111-6.
- Zareei Dehbaghi F (2002). [Assessment of OSCE method in evaluation clinical education of nursing students in the neonatal ward of nursing and midwifery faculty, Shahid Beheshti Medical Sciences University]. Thesis for MSc degree in nursing education. (Persian)

Validity & reliability of OSCE/OSPE in assessing biochemistry laboratory skills of freshman nursing students of Shahid Beheshti University of Medical Sciences

Hosseini S¹, Vartanoosian J², Hosseini F³, Farzin Fard F⁴

1. Faculty Member, Department of Basic Sciences, Faculty of Nursing and Midwifery, Shahid Beheshti University of Medical Sciences and Health Services, Tehran, Iran
 2. Medical Laboratory Scientist and Master student of Medical Education, Department of Basic Sciences, Faculty of Nursing and Midwifery, Shahid Beheshti University of Medical Sciences and Health Services, Tehran, Iran
 3. Faculty Member, Department of Statistic and Mathematics, School of Health Management and Information Science, Iran University of Medical Sciences, Tehran, Iran
 4. Faculty Member, Department of Psychology, Faculty of Nursing and Midwifery, Shahid Beheshti University of Medical Sciences and Health Services, Tehran, Iran
- Corresponding author: Vartanoosian Jacqueline, Medical Laboratory Scientist and Master student of Medical Education, Department of Basic Sciences, Faculty of Nursing and Midwifery, Shahid Beheshti University of Medical Sciences and Health Services, Tehran, Iran
Email: jacqva@gmail.com

Abstract

Background and aim: Objective Structured Clinical/Practical Examination (OSCE/OSPE) is a well known evaluation method in medical education. Lack of research in the evaluation of clinical/practical skills in basic sciences was the justification for the present study to investigate the reliability and validity of OSCE/OSPE in the evaluation of biochemical lab skills of the first year nursing students.

Materials and methods: This is an applied descriptive study. This study was carried out on 31 first year nursing students. At first the checklists were reviewed by the panel of experts, to implement their revisions and recommendations. According to research objectives 10 stations were designed. In each station two examiners separately completed the checklists during a three days period (after 72 hours). The first day, the tests were given in 5 stations, the second day also in 5 other stations, and the third day the skills of the previous days were repeated in 10 stations. To check the content validity the checklists along with research objectives were given to 4 experts in the field. Criterion related validity, internal consistency, examiners reliability, and test-retest were checked.

Findings: The findings showed a direct correlation between the total score gained in OSCE/OSPE and the scores from biochemical written test but it was not significant in terms of statistics, for the first observer calculated correlation was $r=0.107$, $p=0.567$ and for the second observer $r=0.201$, $p=0.277$. To find out the internal consistency, the correlation between the total score in OSCE/OSPE and the students' average score in each station were examined. There was a significant correlation among total score and scores of significant in 8 stations, but there was no significant correlation between the scores of two stations and the total score of the test in two other stations. A significant correlation between the results of the OSCE/OSPE by two examiners was obtained ($r=0.859$, $p<0.001$). In test-retest reliability, there was a significant correlation for the first examiner on the first and third day ($r=0.681$, $p<0.001$) also a significant correlation was found between the scores of the second examiner on the first and third day ($r=0.671$, $p<0.001$). In regards to Cronbach's Alpha coefficient ($=0.907$) the reliability of the designed research test was verified.

Conclusion: Using the OSCE/OSPE as a valid and reliable method is recommended in assessing biochemistry laboratory skills.

Keywords: OSCE/OSPE, Biochemistry Laboratory, Validity, Reliability.