

پایایی و روایی نخستین آزمون ساختاریافته عینی - بالینی (OSCE) روانپزشکی در ایران

دکتر ارسیا تقوا*، دکتر مریم رسولیان**، دکتر لیلی پناشی***، دکتر جعفر بوالهری****،
دکتر مهران ضرغامی*****، دکتر مهدی نصر اصفهانی****، دکتر برابان هاجز*****

چکیده

هدف: این پژوهش با هدف بررسی پایایی و روایی آزمون ساختاریافته عینی - بالینی (OSCE) در حوزه روانپزشکی و در میان دستیاران انجام شده است.

روش: پس از بررسی و بحث‌های فراوان میان استادان روانپزشکی کشور، نه ایستگاه ۱۲ دقیقه‌ای طراحی گردید. در هر ایستگاه دو ارزیاب ممتحن جداگانه به تکمیل چک‌لیست‌های (برگ آزمون‌های) از پیش طراحی شده پرداختند. در هشت ایستگاه از بیمار استاندارد شده استفاده شد. روایی ظاهری و محتوایی، روایی ساختاری، روایی هم‌زمان و پایایی بین دو ارزیاب، ثبات درونی و پایایی بین ایستگاه‌ها بررسی شد.

یافته‌ها: روایی ظاهری و محتوایی آزمون پس از برگزاری جلساتی با حضور استادان و دبیر دانش‌نامه تخصصی کشور و استخراج کلیه هدف‌های عملی و به بحث گذاشتن آن در نشست‌های گوناگون به دست آمد. روایی ساختاری با مقایسه نمرات OSCE دستیاران با نمره‌های استادان مستقیم ایشان و روایی هم‌زمان با بررسی همبستگی نمره‌های OSCE با آزمون دانش‌نامه کتبی، شفاهی، ارتقا و ورودی به رشته روانپزشکی بررسی گردید، که نشان‌دهنده روایی مناسب این آزمون بود. بررسی پایایی آزمون نیز به کمک پایایی بین دو ارزیاب، ثبات درونی و پایایی بین ایستگاه‌ها نشان داد که این آزمون از پایایی مناسب برخوردار است.
نتیجه‌گیری: نخستین آزمون OSCE روانپزشکی دارای پایایی و روایی چشم‌گیری بود.

کلیدواژه: آزمون ساختاریافته عینی - بالینی (OSCE)، روایی، پایایی، روانپزشکی

* روانپزشک، استادیار دانشگاه علوم پزشکی ارتش. تهران، بزرگراه افسریه، نرسیده به سه راه کلاهدوز، بیمارستان ۵۰۶ دورنگار: ۰۹۱-۳۳۲۱۷۰۲۱-۰۲۱ (نوبنده مستول).
E-mail: drarsiataghva@irimc.org

** روانپزشک، استادیار دانشگاه علوم پزشکی و خدمات بهداشتی - درمانی ایران. انستیتو روانپزشکی تهران و مرکز تحقیقات بهداشت روان.

*** متخصص پزشکی اجتماعی، پژوهشکده خانواده، دانشگاه شهید بهشتی.

**** روانپزشک، دانشیار دانشگاه علوم پزشکی و خدمات بهداشتی - درمانی ایران. انستیتو روانپزشکی تهران و مرکز تحقیقات بهداشت روان.

***** روانپزشک، استاد مرکز تحقیقات روانپزشکی و علوم رفتاری و گروه روانپزشکی دانشگاه علوم پزشکی و خدمات بهداشتی - درمانی مازندران.

***** روانپزشک، دانشیار و نایب‌رییس گروه روانپزشکی دانشگاه تورنتو و رییس مرکز پژوهش در آموزش ویلسون.

OSCE در ایران در رشته‌های گوناگون، مانند پزشکی اجتماعی، طب کودکان و پرتونگاری برگزار می‌گردد، اما تا کنون تجربه‌ای از OSCE روانپزشکی به صورت عملی در کشور ما گزارش نشده است. دبیرخانه شورای آموزش پزشکی و تخصصی کشور که متولی برگزاری امتحانات نهایی رشته‌های تخصصی است، همه رشته‌ها را واداشته است که امتحانات بالینی نهایی خود را از سال ۱۳۸۴ به شیوه OSCE برگزار نمایند. هر سال با پایان آزمون دانش‌نامه تخصصی اعتراض‌های زیادی در زمینه غیرعادلانه بودن آزمون به شیوه مرسوم وجود دارد و به نظر می‌رسد که کاربرد OSCE در زمینه روانپزشکی تا اندازه‌ای بتواند تأثیر شانس و نیز غیرعادلانه بودن را کاهش دهد. از این رو پژوهش حاضر با هدف بررسی پایایی و روایی OSCE روانپزشکی در ایران انجام شد.

روش

برای طراحی ایستگاه‌ها، چک‌لیست‌ها و سایر موارد، سه ماه کار پژوهشی، نه ماه کار اجرایی، سه کارگاه مفصل و چهار کارگاه فشرده برای استادان برگزار گردید و در نهایت سناریو و چک‌لیست هر ایستگاه توسط استادان روانپزشکی دانشگاه‌های علوم پزشکی مختلف در کل کشور تهیه شد.

برای اجرای پژوهش، همه دستیاران سال سوم رشته روانپزشکی که امتحان فارغ‌التحصیلی را پشت سر گذاشته بودند (پروبرد) در ۲۲ مرداد ۱۳۸۳ برای انجام این آزمون فراخوانده شدند. ۲۲ دستیار از هفت دانشگاه در مرکز مهارت‌های بالینی بیمارستان فیروزگر (محل آزمون) شرکت کردند.

در این آزمون نه ایستگاه وجود داشت. مدت زمان هر ایستگاه ۱۲ دقیقه بود. یک دقیقه پیش از پایان وقت در هر ایستگاه یک زنگ کوتاه نواخته می‌شد که به دستیار گوشزد می‌کرد که زمان رو به پایان است و پس از آن با شنیدن صدای زنگ پیوسته، دستیار ایستگاه را به قصد ایستگاه بعدی

آزمون ساختاریافته عینی-بالینی^۱ (OSCE) به منظور ارزیابی مهارت‌های بالینی به صورت عینی و ساختاریافته پدید آمده و در آموزش پزشکی نخستین بار توسط هاردن^۲ و گلیسون^۳ (۱۹۷۹) به کار برده شده است. گسترش OSCE شاید مرهون انتقادهایی باشد که به آزمون شفاهی سنتی وارد شده بود، یعنی آزمونی‌هایی که در آن آزمون‌شونده تنها با یک بیمار و تنها توسط یک ارزیاب مورد ارزیابی قرار می‌گیرد. آنچه که سبب زیر سؤال رفتن امتحانات شفاهی به شیوه مرسوم شد، نتایج بررسی‌های بوردهای آموزش پزشکی^۴ در ایالات متحده آمریکا بود که در آن ده هزار دانشجوی پزشکی در طی سه سال بررسی شدند و توافق دو ارزیاب در مورد یک بیمار برای یک دانشجوی کمتر از ۰/۲۵ گزارش گردید. چند سال بعد لیچنر^۵، سیسلر^۶ و هارپر^۷ (۱۹۸۶) دریافتند که عامل شانس، نوع بیمار و آزمون‌گر در نتیجه نهایی آزمون نقشی اساسی دارد (همان‌جا). این نقش به ویژه در نمره دانشجویانی که در مرز رد و قبول هستند اهمیت می‌یابد. از آنجا که افزایش دقت پایایی آزمون شفاهی تنها با بهبود دقت آزمون‌گر دشوار است، به نظر می‌رسد که سنجش مهارت‌های گوناگون (و نه تنها یک مهارت ویژه) و استفاده از چندین ارزیاب باعث بالارفتن روایی و پایایی آزمون شفاهی خواهد شد (همان‌جا).

هر آزمون OSCE در بردارنده چند ایستگاه است و در هر چرخه آزمون، گستره‌ای از مهارت‌های بالینی و عملی سنجیده می‌شود. در هر ایستگاه مهارت‌های بالینی توسط یک ارزیاب که معمولاً ارتباط و رویارویی مستقیمی با آزمون‌شونده ندارد، ارزیابی می‌شود. ارزیابی به کمک برنامه و چک‌لیست‌های از پیش طراحی شده انجام می‌شود (هاردن و گلیسون، ۱۹۷۹).

چک‌لیست‌ها شفاف و روشن هستند و هدف‌های عینی و عملی را ارزیابی می‌کنند. در OSCE معمولاً از بیماران واقعی بهره گرفته نمی‌شود، بلکه بیماران استاندارد شده^۸ (SP) یا شبیه‌سازی شده^۹ شرکت داده می‌شوند. بیمار استاندارد شده بایستی وظایف محول شده را به صورتی یکپارچه برای هر شرکت کننده اجرا نماید (باروز^{۱۰} و آبراهامسون^{۱۱}، ۱۹۶۴). کاربرد OSCE در روانپزشکی سال‌هاست که رواج دارد و پایایی و روایی آن نیز به طور گسترده بررسی و در سطح قابل قبولی گزارش شده است (هاجز^{۱۲}، رگهر^{۱۳}، هانسون^{۱۴}، ۱۹۹۷، ۱۹۹۸؛ حاجز، ترن‌بال^{۱۵} و کوهن^{۱۶}، ۱۹۹۶؛ حاجز، هانسون، مک‌ناتان^{۱۷} و رگهر، ۱۹۹۹).

- | | |
|--|--|
| 1- Objective Structured Clinical Examination | 3- Gleeson |
| 2- Harden | 4- National Board of Medical Examination |
| 5- Lechner | 6- Sisler |
| 7- Harper | 8- standardized patient |
| 9- simulated patient | 10- Barrows |
| 11- Abrahamson | 12- Hodges |
| 13- Regehr | 14- Hanson |
| 15- Turnbull | 16- Cohen |
| 17- McNaughtan | |

جدول ۱- مهارت‌های ارزیابی‌شده در ایستگاه‌های مختلف OSCE روانپزشکی

شماره	مهارت مورد ارزیابی	ایستگاه
۱	مصاحبه و برقراری ارتباط و جهت‌دهی مصاحبه متناسب با زمان (در مواجهه با بیمار مرد مبتلا به افسردگی اساسی)	
۲	معاینه وضعیت روانی ^۱ (MSE) به‌خصوص مشکلات ادراکی در زن مبتلا به اسکیزوفرنیا	
۳	ارزیابی مشکلات شناخت و سوگیری با استفاده از ^۱ (MMSE) در مرد مشکوک به دمانس	
۴	توانایی تشخیص تظاهر به جنون در مرد مجرم	
۵	توانایی ارزیابی و مدیریت درمان بیمار کنترل شده مبتلا به اختلال دوقطبی نوع دو	
۶	مشاهده و مصاحبه با خانواده بیمار مبتلا به اسکیزوفرنیا	
۷	ارزیابی و اداره رفتارهای کلامی و غیرکلامی بیمار با شخصیت مرزی	
۸	معاینه وضعیت روانی بیمار مبتلا به اختلال وسواسی-اجباری	
۹	طبقه‌بندی و فرمولاسیون مکتوب بیمار ایستگاه هشت	

1- Mental Status Examination; 2- Mini-Mental State Examination

۱۹

نمودند و در پاره‌ای از موارد وضعیت قرارگیری اشیاء را در اتاق تغییر دادند. سپس با برگزاری یک آزمون آزمایشی بدون حضور دانشجوی^۱، روش قرارگیری اشیاء و شیوه ایفای نقش بیمار استاندارد بررسی شد.

چون OSCE روانپزشکی برای نخستین بار در ایران اجرا می‌شد، بیمار استاندارد شده وجود نداشت. از این رو هشت نفر از کارکنان انستیتو روانپزشکی تهران (چهار نفر) و دانشجویان روانشناسی (چهار نفر)، چهار جلسه و نزدیک به هشت ساعت آموزش دیدند. روش آموزش ایشان نخست مطالعه و سپس تمرین‌های پی‌درپی به‌صورت ایفای نقش در حضور روانپزشکی که با اصول انجام این کار آشنا بود انجام شد. افزون بر این، روز پیش از برگزاری آزمون، بیماران استاندارد شده توسط چند ارزیاب از نظر ایفای مناسب نقش بررسی شدند و در چهارچوب فضای ایستگاه داوری گردیدند و تا حد ممکن خطای آنها تصحیح شد.

در هر ایستگاه دو ارزیاب حضور داشتند که به‌صورت مستقل دستیاران را با استفاده از چک‌لیست ارزیابی می‌نمودند. در همه ایستگاه‌ها هیچ ارتباطی میان ارزیاب‌ها و دستیاران وجود نداشت.

ترک می‌نمود. در هر ایستگاه یک مهارت خاص ارزیابی می‌شد. روی در ایستگاه و نیز روی صندلی آزمون‌شونده دستورالعملی مبتنی بر وظایف مورد انتظار آزمون‌شونده درج شده بود که شامل سه بخش بود: ۱- هدف از ایستگاه، ۲- فعالیت مورد انتظار و ۳- شیوه ارزشیابی.

پس از چهار ایستگاه، دستیار وارد ایستگاهی می‌شد که می‌توانست در آن استراحت نماید. به‌جز ایستگاه شماره ۸ که در آن آزمون‌شونده به جمع‌بندی نوشتاری ایستگاه شماره هشت می‌پرداخت، در هر ایستگاه یک بیمار استاندارد و دو ارزیاب که به‌صورت مستقل نمره می‌دادند حضور داشتند. جمع‌بندی نوشتاری ایستگاه شماره هشت در ایستگاه ۸ انجام می‌شد و دو ایستگاه هشت و ۹ روی هم یک نمره می‌گرفتند. فهرست مهارت‌های مورد ارزیابی در ایستگاه‌ها در جدول ۱ دیده می‌شود.

کوشش شد که در هر ایستگاه ارزیاب از میان استادانی انتخاب شود که در فرآیند طراحی و اجرای اولیه ایستگاه نقشی نداشته باشند. هر چند در عمل به دلیل مشکلات اجرایی، در برخی از ایستگاه‌ها یکی از ارزیاب‌ها از طراحان سناریو یا چک‌لیست همان ایستگاه بود.

برای بررسی استاندارد بودن فضای فیزیکی ایستگاه‌ها، یک روز پیش از برگزاری آزمون، ارزیاب‌ها مکان را بررسی

توافق گردید. این ۱۵ هدف در یک کارگاه هشت ساعته با حضور استادان روانپزشکی برجسته کشور در قالب هشت ایستگاه برنامه‌ریزی شد. در کارگاه بعدی متخصصان روانپزشکی کلیه دانشگاه‌هایی که به آموزش روانپزشکی در سطح تخصصی اشتغال داشتند، پرسش‌ها را طراحی نموده، پرسش‌ها به تفکیک دانشگاه‌های مختلف بررسی و بازبینی اجرایی شدند. به این ترتیب کلیه پرسش‌های چک‌لیست چندین بار توسط استادان دانشگاه‌ها بررسی شدند.

سپس مکان مناسب برای انجام آزمون OSCE فراهم شد و استادان امکانات عملی ایستگاه‌ها را با هدف‌های تعیین‌شده سازگار نمودند. سناریوهای طراحی‌شده توسط هر گروه از استادان توسط گروه‌های دیگر نقد و بررسی گردید.

برای محاسبه روایی ساختاری پرسش‌نامه، کارنامه‌های دستیاران شرکت‌کننده در OSCE به تفکیک دانشگاه‌ها جدا شد و سپس رتبه هر کدام از ایشان در هر ایستگاه مشخص گردید. سپس کارنامه‌ها (بدون طبقه‌بندی معین‌شده در OSCE) با نمره ارزیابی مهارت‌های عملی (نمره درون‌بخشی سال سوم) که به صورت سنتی در بخش‌های آموزشی انجام شده بود مقایسه شد.

برای بررسی روایی هم‌زمان، همبستگی بین نمرات OSCE با نمرات آزمون دانش‌نامه کتبی، دانش‌نامه شفاهی و ارتقای سال اول، دوم و سوم محاسبه گردید.

یافته‌ها

یافته‌های مربوط به بررسی روایی هم‌زمان، همبستگی بین نمرات OSCE با نمرات آزمون کتبی دانش‌نامه، آزمون شفاهی دانش‌نامه و آزمون‌های ارتقای سال اول، دوم و سوم در **جدول ۲** نشان داده شده است. نتایج مربوط به ارزیابی پایایی بین ارزیاب‌ها^۱ با توجه به کمی بودن نمرات دستیاران، بر پایه ضریب همبستگی بین طبقه‌ای^۱ (ICC) در **جدول ۳** ارائه گردیده است. ضریب‌های آلفای کرونباخ به دست آمده در هر ایستگاه در **جدول ۴** ارائه شده است.

برای بررسی پایایی این ابزار، نمره‌های دستیاران در ایستگاه‌های مختلف، به کمک ضریب همبستگی پیرسون بررسی گردیدند (**جدول ۵**). هم‌چنین نمره کل هر دستیار با نمره وی در هر ایستگاه به کمک ضریب همبستگی پیرسون بررسی شد (**جدول ۶**).

هر پرسش در چک‌لیست با نمره ای بین ۳-۰ (از ضعیف تا عالی) بر حسب عملکرد دستیار نمره‌گذاری می‌گردید. در حقیقت چک‌لیست، فهرستی از مهارت‌هایی بود که در آن ایستگاه خاص ارزیابی می‌شد. این پرسش‌ها در رابطه با شیوه عملکرد دستیاران بودند؛ آیا به نمونه، آیا به سابقه خانوادگی بیمار توجه شده است؟ آیا به خودکشی توجه کرده است؟ در پایان چک‌لیست در هر ایستگاه (به جز ایستگاه نُه) دو پرسش وجود داشت. پرسش نخست درباره نظر کلی استاد ارزیاب نسبت به دستیار و پرسش دیگر نیز درباره نظر کلی بیمار استناداردهنده نسبت به دستیار بود. هر دو پرسش یادشده بین ۳-۰ نمره‌گذاری شد. در نهایت نمره هر دستیار به صورت کلی و بدون در نظر گرفتن پرسش آخر محاسبه گردید. بالاترین امتیاز در نظر گرفته شده برای تمام ایستگاه‌ها ۲۰ بود؛ به غیر از ایستگاه شماره یک که ۳۰ امتیاز و ایستگاه شماره چهار که ۱۰ امتیاز داشت (ایستگاه‌های شماره هشت و نه روی هم ۲۰ امتیاز می‌گرفتند). در نمره ای که توسط استادان به عنوان نمره کلی در نظر گرفته شده، نظر بیماران استناداردهنده حساب نشده است و نمره دستیار تنها بر پایه ریز نمره‌های ماده‌ها حساب شد. چک‌لیست‌ها، نمره‌های هر ایستگاه و همه امتیازبندی‌ها چند بار توسط استادان بازبینی شد. یک هفته پس از برگزاری آزمون، برای هر دستیار کارنامه‌ای صادر شد که در آن نمره کلی دستیار، بالاترین نمره به دست آمده، میانگین و بالاترین نمره در هر ایستگاه نوشته شده بود. برای بررسی پایایی و روایی OSCE از روش‌های آلفای کرونباخ، کاپا و ضریب همبستگی پیرسون بهره گرفته شد. گفتنی است که چون آزمون کولموگروف - اسمیرنوف توزیع داده‌ها را نرمال نشان داد، بهره‌گیری از ضریب همبستگی پیرسون منطقی به نظر می‌رسید.

برای ساخت چک‌لیست و ایستگاه‌های OSCE کارهای زیر انجام شد. پس از برگزاری چند نشست با حضور دبیر هیئت‌بورد روانپزشکی و استادان دانشگاه‌های علوم پزشکی تهران، ایران، شهید بهشتی، مازندران و اصفهان، نخست هدف‌های کلی تعیین شد. این هدف‌ها بیشتر شامل همان هدف‌های کلی آموزش روانپزشکی بود که دبیرخانه شورای تخصصی وزارت بهداشت، درمان و آموزش پزشکی تهیه کرده بود. سپس مشخص شد که کدامیک از این هدف‌ها عملی و از نظر بالینی قابل اجرا است. این هدف‌ها با حضور استادان بیشتری از سایر دانشگاه‌های کشور بازبینی شد و در نهایت بر سر پانزده هدف

1- inter-rater reliability

2- intraclass correlation coefficient

جدول ۲- ضریب‌های همبستگی بین نمرات OSCE و امتحان ورودی دستیاری و ارتقای سال‌های مختلف و امتحان شفاهی و کتبی دانش‌نامه تخصصی

نمره OSCE	نمره ورودی دستیاری	ارتقای ۱ به ۲	ارتقای ۲ به ۳	نمره درون بخشی	کتبی دانش‌نامه	شفاهی دانش‌نامه
-	* ۰/۶۴	۰/۳۱	۰/۳۱	۰/۳۶	۰/۳۶	* ۰/۶۳
		۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵
* ۰/۶۴	-	۰/۱۰۶	۰/۱۰۹	۰/۲۸	۰/۵۰	۰/۵۰
۰/۱۰۵		۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵
۰/۳۱		-	* ۰/۷۵	* ۰/۴۴	* ۰/۴۶	۰/۲۴۱
۰/۱۰۵		۰/۱۰۵	۰/۱۰۱	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵
۰/۳۱		* ۰/۷۵	-	- ۰/۱۹	* ۰/۶۷	- ۰/۰۲
۰/۱۰۵		۰/۱۰۵	۰/۱۰۱	۰/۱۰۵	۰/۱۰۱	۰/۱۰۵
۰/۳۰		* ۰/۷۵	-	- ۰/۰۹	* ۰/۷۰	- ۰/۰۶
۰/۱۰۵		۰/۱۰۵	۰/۱۰۱	۰/۱۰۵	۰/۱۰۱	۰/۱۰۵
۰/۳۶		۰/۲۸	۰/۴۴	-	۰/۱۰	* ۰/۶۴
۰/۱۰۵		۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰
۰/۳۶		۰/۵۰	* ۰/۶۶	* ۰/۷۰	-	۰/۱۳
۰/۱۰۵		۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵	۰/۱۰۵
* ۰/۶۳		۰/۵۰	۰/۲۴	* ۰/۶۴	۰/۱۳	-
۰/۱۰۵		۰/۱۰۵	۰/۱۰۵	۰/۱۰۱	۰/۱۰۵	۰/۱۰۵

* p < ۰/۰۵

جدول ۳- ضریب‌های همبستگی بین دو ارزیابی در هر ایستگاه OSCE

شاخص آماری	ایستگاه	اول	دوم	سوم	چهارم	پنجم	ششم	هفتم	هشتم
ضریب همبستگی بین طبقه‌ای		۰/۷۵۱۰	۰/۶۶۰۳	۰/۹۰۱۷	۰/۷۷۳۱	۰/۹۱۱۵	۰/۸۰۰۸	۰/۸۶۷۸	۰/۷۱۵۲
سطح معنی داری		۰/۰۰۱	۰/۰۱	۰/۰۰۱	۰/۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۱	۰/۰۱

جدول ۴- ضرایب آلفای کرونباخ به دست آمده در هر ایستگاه برای بررسی پایایی آن

شماره ایستگاه	α کرونباخ
ایستگاه اول	۰/۸
ایستگاه دوم	۰/۷۱
ایستگاه سوم	۰/۷۸
ایستگاه چهارم	۰/۴۲
ایستگاه پنجم	۰/۶۳
ایستگاه ششم	۰/۵۹
ایستگاه هفتم	۰/۷۶
ایستگاه هشتم	۰/۷

جدول ۵- ضریب‌های همبستگی بین نمره‌های به‌دست آمده در ایستگاه‌های مختلف OSCE

ایستگاه	ایستگاه	ایستگاه	ایستگاه	ایستگاه	ایستگاه	ایستگاه	ایستگاه	ایستگاه	ایستگاه
اول	دوم	سوم	چهارم	پنجم	ششم	هفتم	هشتم	ایستگاه	ایستگاه
-	۰/۱۳	۰/۰۲	۰/۱۷	۰/۰۸	۰/۰۶	۰/۴۸*	۰/۰۴	ایستگاه	ضریب همبستگی
۰/۱۲	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	اول	سطح معنی داری
-	۰/۱۲	۰/۳۰	۰/۱۳	۰/۰۵	۰/۰۵	۰/۳۵	۰/۱۵۲	ایستگاه	ضریب همبستگی
۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	دوم	سطح معنی داری
۰/۰۲	۰/۳۰	-	۰/۱۱۳	۰/۰۴	۰/۲۸	۰/۱۱	۰/۰۳	ایستگاه	ضریب همبستگی
۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	سوم	سطح معنی داری
۰/۱۷	۰/۱۳	۰/۱۳	-	۰/۲۰	۰/۵۳*	۰/۲۳	۰/۴۲	ایستگاه	ضریب همبستگی
۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	چهارم	سطح معنی داری
۰/۰۸	۰/۱۵	۰/۰۴	۰/۲۰	-	۰/۵۰*	۰/۲۲	۰/۳۰	ایستگاه	ضریب همبستگی
۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	پنجم	سطح معنی داری
۰/۰۶	۰/۰۶	۰/۲۸	۰/۵۳*	۰/۵۱*	-	۰/۱۵۱	۰/۲۴	ایستگاه	ضریب همبستگی
۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	ششم	سطح معنی داری
۰/۴۸*	۰/۳۵۶	۰/۱۱	۰/۲۳	۰/۲۲	۰/۱۵	-	۰/۲۵	ایستگاه	ضریب همبستگی
۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	هفتم	سطح معنی داری
۰/۰۴	۰/۱۵	۰/۰۳	۰/۴۲	۰/۳۰	۰/۲۴	۰/۲۵	-	ایستگاه	ضریب همبستگی
۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	۰/۰۵	هشتم	سطح معنی داری

* p < ۰/۰۵

جدول ۶- همبستگی بین نمره‌های هر ایستگاه OSCE با نمره کل

این آزمون که نخستین تجربه OSCE روانپزشکی در ایران بود، پایایی و روایی چشم‌گیری را هم‌چون سایر بررسی‌ها در حوزه پزشکی (اسلوان^۱، دانلی^۲ و شوارتز^۳، ۱۹۹۵) و روانپزشکی (والترز^۴، اسبورن^۵ و ریون^۶، ۲۰۰۵) نشان داد. پایایی میان ارزیاب‌ها در هر ایستگاه معنی‌دار بود، هر چند که در برخی موارد مقادیر کاپا متوسط بود. همسانی درونی در هشت ایستگاه تقریباً در حد مناسبی بود. پایین‌ترین میزان همسانی درونی در ایستگاه شماره چهار (شخصیت ضداجتماعی با تقلید علایم سایکوز) دیده شد. ایستگاهی که از آغاز طرح بر سر سناریو و شیوه آماده‌سازی چک‌لیست آن بحث‌های زیادی وجود داشت و به نظر می‌رسید که به دلیل این اختلاف سلیقه‌ها، این ایستگاه از یکپارچگی درونی کافی برخوردار نبوده است. از همین رو نمره کمتری برای این ایستگاه در نظر گرفته شد.

نمره کلی	دستیار	ایستگاه‌های OSCE
۰/۵۲۹	۰/۰۱	ایستگاه اول
۰/۵۵۶	۰/۰۱	ایستگاه دوم
۰/۳۲۸	۰/۰۵	ایستگاه سوم
۰/۴۵۶	۰/۰۵	ایستگاه چهارم
۰/۵۹۵	۰/۰۱	ایستگاه پنجم
۰/۳۹۶	۰/۰۵	ایستگاه ششم
۰/۷۱۴	۰/۰۱	ایستگاه هفتم
۰/۵۱۷	۰/۰۱	ایستگاه هشتم

- 1- Sloan
2- Donnelly
3- Schwartz
4- Walters
5- Osborn
6- Raven

برای بررسی روایی هم‌زمان OSCE با سایر آزمون‌ها، ضریب همبستگی پیرسون به کار برده شد. همان‌گونه که انتظار می‌رفت نتایج OSCE با نمره کتبی دانش‌نامه، همبستگی معنی‌داری نداشت ولی با نمره شفاهی دانش‌نامه که آن نیز مهارت دستیار را می‌سنجد همبستگی معنی‌داری داشت. نمره‌های ارتقای هر سال با سال‌های دیگر و نمره کتبی دانش‌نامه همبستگی معنی‌داری داشت. این یافته با توجه به ماهیت همه امتحان‌ها که اولویت آن اندازه‌گیری میزان دانش دستیار است هم خوانی دارد. نکته در خور توجه این است که نمره درون‌بخشی با هیچ‌یک از موارد بالا همبستگی معنی‌داری نشان نداد. همبستگی نمره درون‌بخشی با نمره شفاهی دانش‌نامه نیز قابل توجه است؛ چرا که نمره‌های درون‌بخشی نیز افزون بر دانش نظری دستیار به مهارت‌های کلی او تکیه می‌کنند. هر چند که نظر شخصی استادان در نمره‌های درون‌بخشی تأثیر چشم‌گیری دارد.

این آزمون به منظور یک ارزیابی نهایی برای دریافت دانش‌نامه تخصصی (پورد) طراحی شده بود. بررسی فامویوا^۳، زکریا^۴ و ایچوکو^۵ (۱۹۹۱) نشان داد که استفاده از استادان به‌عنوان ارزیاب نقش برجسته‌ای در تکمیل یا پوشش دادن ضعف‌های نمره‌دهی OSCE دارد. تجربه‌ای که به نظر می‌رسد در ارزیابی‌های آینده نیز تا پیدا شدن راه کاری بهتر ادامه خواهد داشت. در بررسی هاجز و همکاران (۱۹۹۹)، دستیاران OSCE را روشی برای ارزیابی نهایی نمی‌دانستند، اما به نظر می‌رسد دستیارانی که در آن بررسی حضور غیرفعال داشته‌اند، مورد مناسبی برای ارزیابی نباشند. افزون بر این بررسی هاجز و همکاران (همان‌جا) با دانشجویان پزشکی انجام شد و نه دستیاران.

به‌طور کلی این بررسی روایی (صوری، محتوایی، ساختاری و هم‌زمان) و پایایی (بین دو ارزیاب، ثبات درونی و پایایی بین ایستگاه‌ها) خوبی برای اجرا به‌عنوان یک امتحان نهایی^۶ که توانایی‌های عملی دستیاران سال آخر روانپزشکی را بسنجد، نشان داد. این تجربه نخستین آزمون آزمایشی کشوری OSCE روانپزشکی، از نظر امکانات و هزینه‌ها با مشکلات بی‌شماری روبرو گردید. تأمین اعتبار مناسب، افزایش شمار داوطلبان شرکت‌کننده، همراه با تجربه‌اندوزی استادان به اعتبار بیشتر آن کمک می‌کند.

ضریب همبستگی میان بیشتر ایستگاه‌ها ارتباط معنی‌داری را با یکدیگر نشان ندادند و در مواردی که این همبستگی معنی‌دار بود، موضوع‌های ایستگاه‌ها با یکدیگر هم‌خوانی نداشتند. برای نمونه ایستگاه شماره یک گرفتن شرح حال روانی از یک بیمار افسرده را می‌سنجد و در ایستگاه شماره هفت آزمون‌شونده قرار بود که مدارا با رفتارهای هیجانی یک بیمار با اختلال شخصیت مرزی را ارزیابی کند. ایستگاه شماره شش که در آن مشاوره با مادر بیمار مبتلا به اسکیزوفرنیا مطرح شده بود، با دو ایستگاه چهار (اختلال شخصیت ضداجتماعی با تقلید علائم سایکوز) و پنج (مشاوره با خانمی با شرح حال اختلال دو قطبی تیپ II) ارتباط معنی‌داری را نشان داد. به‌نظر می‌رسد که در این آزمون به دلیل گوناگون‌بودن وظایف مورد سنجش در ایستگاه‌های مختلف، نباید انتظار داشت که ارتباط میان ایستگاه‌ها معنی‌دار باشد.

افزون بر آن میانگین کل نمره هر دستیار در OSCE با نمره وی در هر ایستگاه ارتباط معنی‌داری را نشان داد. تنها استثناء در این مورد ایستگاه شماره شش بود؛ تنها ایستگاهی که SP در آن بازیگر نقش خود بیمار نبود. این ارتباط معنی‌دار نشان‌دهنده این است که اگر دستیار در طی سه سال دستیاری مهارت‌های لازم را کسب کرده باشد، در همه جنبه‌های تعیین‌شده رشد خواهد داشت و در غیر این صورت ممکن است مهارت‌های او در زمینه‌های گوناگون رشد یکسانی نداشته باشد.

از آن‌جا که از مدت‌ها پیش بر سر هدف‌ها، سناریوها و چک‌لیست‌های آن بحث‌های زیادی به‌صورت گروهی میان استادان باتجربه و صاحب‌نظر روانپزشکی انجام شده و بسیاری از هدف‌های عملی و قابل سنجش در زمینه روانپزشکی در آن گنجانده شده بود، آزمون‌شوندگان از واقعی به‌نظر رسیدن ایستگاه‌ها و بیمار استاددار شده رضایت داشتند.

بررسی روایی ساختاری^۱ نشان داد که OSCE به راستی مهارت دستیار را می‌سنجد؛ زیرا استادانی که با دستیاران از نزدیک آشنایی و بر فعالیت آنها نظارت داشتند، دستیاران را بر پایه توانایی‌های عملی همان‌گونه رتبه‌بندی نمودند که OSCE رتبه‌بندی کرده بود. این آزمون هم‌چون بررسی پارک^۲ و همکاران (۲۰۰۴) که آزمودنی‌های خود را از میان دانشجویان پزشکی انتخاب نموده بودند از روایی ساختاری خوبی برخوردار بود.

سیاسگزاری

از داوران مجله روانپزشکی و روانشناسی بالینی ایران که سبب پربارتر شدن مقاله شده‌اند قدردانی می‌شود.

دریافت مقاله: ۱۳۸۵/۲/۱۶؛ دریافت نسخه نهایی: ۱۳۸۵/۴/۱۲؛ پذیرش مقاله: ۱۳۸۵/۵/۴

منابع

- clinical exam for evaluating psychiatric clinical clerks. *Academic Medicine*, 72, 715-721.
- Hodges, B., Regehr, G., & Hanson, M. (1998). Validation of an objective structured clinical examination in psychiatry. *Academic Medicine*, 73, 910-912.
- Hodges, B., Turnbull, J., & Cohen, R. (1996). Evaluating communication skills in the OSCE format. *Medical Education*, 30, 38-43.
- Leichner, P., Sisler, G. C., & 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.
- Park, R. S., Chibnall, J. T., Blaskiwicz, R. Z., Furman, G. E., Powell, J. K., & Mohr, C. J. (2004). Construct validity of an objective structured clinical examination in psychiatry: Association with the clinical skills examination and other indicators. *Academic Psychiatry*, 28, 122-128.
- Sloan, D. A., Donnelly, M. B., & Schwartz, R. W. (1995). The OSCE the new gold standard for evaluation of post graduate clinical performance. *Annals of Surgery*, 222, 735-742.
- Walters, K., Osborn, D., & Raven, P. (2005). The development validity and reliability of a multimodality OSCE in Psychiatry. *Medical Education*, 39, 292-298.
- Barrows, H. S., & Abrahamson, S. (1964). The programmed patient: A technique for appraising student performance in clinical neurology. *Journal of Medical Education*, 39, 802-805.
- Famuyiwa, O. O., Zachariah, M. P., & Iiechukwu, S. T. (1991). The objective structured clinical exam in psychiatry. *Medical Education*, 25, 45-50.
- Harden, R. M., & Gleeson, F. A. (1979). Assessment of clinical competence using an observed clinical examination. *Medical Education*, 13, 41-47.
- Hodges, B., Hanson, M., McNaughtan, N., & Regehr, G. (1999). What do psychiatry residents think of an OSCE. *Academic Psychiatry*, 3, 198-204.
- Hodges, B., Regehr, G., & Hanson, M. (1997). Structured