

اعتباریابی آزمون ساختاریافته عینی - بالینی روانپزشکی:

یک بررسی راهنما

دکتر عباس عطاری*، دکتر غلامرضا میرسپاسی**، دکتر ارسیا تقوایی***، دکتر جعفر بوالهروی****،
مهین امین‌الرعایا*****، دکتر اکبر حسن زاده*****

چکیده

هدف: آزمون ساختاریافته عینی - بالینی (OSCE) برای امتحانات جامع دستیاری و بورده تخصصی روانپزشکی به کار گرفته می‌شود. در یک بررسی راهنما که با همکاری دانشگاه‌های علوم پزشکی ایران، تهران و اصفهان در مرکز تحقیقات علوم رفتاری اصفهان انجام شد، ضمن آشناسدتن دستیاران و استادان رشته روانپزشکی با این آزمون، روایی و پایایی آن بررسی گردید.
روش: این بررسی برای تعیین روایی و پایایی آزمون OSCE در روانپزشکی در پاییز ۱۳۸۴ در مرکز تحقیقات علوم رفتاری اصفهان انجام شد. برای این منظور یک کارگاه سه روزه تشکیل گردید. ۲۲ نفر از استادان و ۲۲ نفر از دستیاران در این برنامه شرکت نمودند. شش ایستگاه ۱۰ دقیقه‌ای طراحی گردید. برای تعیین روایی، روش روایی صوری و برای تعیین پایایی روش‌های پایایی بین ایستگاهی، همسانی درونی، ضریب وابستگی و پایایی موازی به کار برده شد.
یافته‌ها: میانگین کل نمرات چک‌لیست دستیاران ۸۴/۲۸ (انحراف معیار ۱۱/۹۱) و ضریب وابستگی بین نمرات ارزیاب‌ها (در مجموع شش ایستگاه) برابر با ۰/۹۶ و ضریب همبستگی آن ۰/۸۲ به دست آمد. ضریب همبستگی پیرسون بین میانگین نمره‌های چک‌لیست و نمره کلی محاسبه و ضریب همبستگی ۰/۸۹ ($p < 0.001$) به دست آمد. روایی آزمون در حد بسیار خوب بود.
نتیجه‌گیری: آزمون شفاهی به روش OSCE برای دستیاران روانپزشکی انجام‌شدنی است و از روایی و پایایی مناسبی برخوردار است.

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

* روانپزشک، استاد مرکز تحقیقات علوم رفتاری اصفهان. اصفهان، خیابان استانداری، مرکز پزشکی نور (خورشید). دورنگار: ۲۲۲۲۱۳۵-۰۳۱۱ (نویسنده مسئول).

E-mail: bsre@mui.ac.ir

** روانپزشک، دانشیار بازنشسته دانشگاه علوم پزشکی و خدمات بهداشتی - درمانی تهران و نایب رییس انجمن علمی روانپزشکان ایران.

*** روانپزشک، استادیار دانشگاه علوم پزشکی و خدمات بهداشتی - درمانی ارتش.

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

***** کارشناسی ارشد مدیریت آموزشی، کارشناس مسئول پژوهشی مرکز تحقیقات علوم رفتاری اصفهان.

***** دکترای آمار حیاتی، استادیار دانشکده بهداشت دانشگاه علوم پزشکی اصفهان.

مقدمه

آزمون ساختاریافته عینی- بالینی^۱ (OSCE) را نخستین بار دکتر رونالد هاردن^۲ در دهه ۱۹۷۰ میلادی برای سنجش مهارت‌های بالینی دانشجویان به کار برد (هاردن و گلیسون^۳، ۱۹۷۹). در امتحانات شفاهی کنونی یک یا دو نفر به عنوان ممتحن حضور دارند و تنها یک بیمار (موقعیت بالینی) توسط دانشجو ارزیابی می‌شود، در حالی که در امتحان OSCE موقعیت‌های بالینی متفاوتی ارزیابی می‌شوند و دانشجو توسط چند نفر ناظر یا ممتحن مورد آزمون قرار می‌گیرد. به هر روی OSCE یک ابزار ارزیابی نوین است که مزیت‌های مهمی بر امتحان شفاهی دارد (هاجز^۴، هانسن^۵، مک‌ناگتون^۶ و رگهر^۷، ۲۰۰۲). در روش OSCE امکان استفاده از بیماران استاندارد شده^۸ (SP) فراهم شده است. کاربرد بیمار استاندارد شده نخستین بار توسط باروز^۹ و آبراهامسون^{۱۰} (۱۹۶۴) ارایه گردید. به کارگیری بیماران استاندارد شده سبب یکسان شدن ماهیت مشکل بالینی و سطح دشواری آزمون برای همه دانشجویان می‌شود.

هم‌اکنون در بسیاری از دانشکده‌های پزشکی سراسر دنیا، OSCE به عنوان شیوه انتخابی برای ارزیابی مهارت‌های بالینی دانشجویان پزشکی به کار برده می‌شود (بلاسکوویچ^{۱۱}، پارک^{۱۲}، چینسال^{۱۳} و پاول^{۱۴}، ۲۰۰۴). در روش امتحان به صورت OSCE ایستگاه‌های چندی طراحی می‌گردد. در هر ایستگاه دانشجو باید طی مدت معینی یک بیمار (SP) را معاینه و ارزیابی کند و هر ایستگاه هدف خاصی را دنبال می‌کند. هم‌چنین کوشش می‌شود در یک آزمون، مطالب مهم آموزشی عملی مانند مصاحبه و معاینه بالینی، کسب اطلاعات و تشخیص و درمان گنجانده شود (بلاسکوویچ و همکاران، ۲۰۰۴). معمولاً دانشجویان با دو روش ارزیابی می‌شوند. یکی به صورت چک‌لیست‌های دوگزینه‌ای بلی و خیر که بیشتر برای ارزیابی محتوا^{۱۵} مانند گرفتن تاریخچه، روش برقراری ارتباط، تشخیص افتراقی و غیره به کار می‌رود و دوم نمره‌دهی کلی^{۱۶} (GR) که برای ارزیابی فرآیندی^{۱۷} هم‌چون مهارت‌های میان‌فردی مناسب است (باروز و آبراهامسون، ۱۹۶۴؛ بلاسکوویچ و همکاران، ۲۰۰۴). اگر چه OSCE در پزشکی عمومی و تخصصی به خوبی تبیین گردیده است ولی کاربرد آن در روانپزشکی به تازگی آغاز گردیده و تنها در سال‌های اخیر OSCE در روانپزشکی به بحث و چالش گذاشته شده است. طراحی امتحان OSCE در روانپزشکی از هر رشته تخصصی دیگر دشوارتر بوده، نیاز به وقت و هزینه بیشتری دارد (والترز^{۱۸}، اسپورن^{۱۹} و راون^{۲۰}، ۲۰۰۵).

در زمینه تعیین روایی و پایایی این آزمون در روانپزشکی بررسی‌های اندکی انجام شده است (هاجز، ۲۰۰۳؛ فاماویوا^{۲۱}، زکریا^{۲۲} و ایلچو کوو^{۲۳}، ۱۹۹۱؛ هاجز، رگهر، هانسون و مک‌ناگتون، ۱۹۹۷؛ هاجز، رگهر، هانسون و مک‌ناگتون، ۱۹۹۸؛ هاجز، رگهر، مک‌ناگتون، تیبریوس^{۲۴} و هانسون، ۱۹۹۹؛ پارک^{۲۵} و همکاران، ۲۰۰۴). در یک تجربه شش ساله در نیجریه نشان داده شد که OSCE برای ارزیابی دانشجویان دوره عمومی روانپزشکی انجام‌شدنی است و از نظر آماری از روایی و پایایی خوبی برخوردار است (فاماویوا و همکاران، ۱۹۹۱). در دانشکده پزشکی تورنتوی کانادا در سال ۱۹۹۵ روایی و پایایی آزمون OSCE در روانپزشکی بررسی شد و در آن دو ردیف ایستگاه‌های ده‌گانه موازی طراحی گردید. نتایج نشان داد که این آزمون برای ارزیابی مهارت‌های پیچیده روانپزشکی قابل اجرا است و از روایی و پایایی قابل قبولی برخوردار است (هاجز و همکاران، ۱۹۹۷). در بررسی دیگری، ۳۳ نفر از دانشجویان دوره بالینی عمومی با ۱۷ دستیار روانپزشکی در دانشگاه تورنتو در یک امتحان OSCE شرکت نمودند. این بررسی نشان داد که میانگین نمرات ارزیابی کلی (GR) در دستیاران بالاتر از دانشجویان است، ولی با نمره‌های چک‌لیست، تفاوت چندانی نداشت. بررسی روایی (همزمان^{۲۶} و سازه^{۲۷}) نشان داد که OSCE در روانپزشکی دارای روایی قابل قبولی برای ارزیابی دانشجویان دوره بالینی است (هاجز و همکاران، ۱۹۹۸). در پژوهش دیگری ۱۴ نفر دانشجوی بالینی با ۱۴ نفر دستیار روانپزشکی و ۱۴ نفر پزشک خانواده در یک آزمون OSCE مقایسه شدند؛ میانگین نمرات نمره‌دهی کلی (GR) پزشکان خانواده بالاتر از دستیاران و دانشجویان و میانگین نمره‌های برگ آزمون^{۲۸} دستیاران بالاتر از پزشکان خانواده و دانشجویان بود. یافته‌های این پژوهش نشان داد که چک‌لیست‌های دوگزینه‌ای برای سنجش میزان شایستگی بالینی برای مقاطع پایین‌تر تحصیلی معتبر می‌باشد (هاجز و همکاران، ۱۹۹۹).

- | | |
|--|-----------------|
| 1- Objective Structured Clinical Examination | 3- Gleeson |
| 2- Ronald Harden | 5- Hanson |
| 4- Hodges | 7- Regehr |
| 6- McNaughton | 9- Barrows |
| 8- standardized patients | 11- Blaskiewicz |
| 10- Abrahamson | 13- Chibnall |
| 12- Park | 15- content |
| 14- Powell | 17- process |
| 16- global rating | 19- Osborn |
| 18- Walters | 21- Famuyiwa |
| 20- Raven | 23- Ilechukwu |
| 22- Zachariah | 25- Park |
| 24- Tiberius | 27- construct |
| 26- concurrent | |
| 28- checklist | |

نکرده بودند. با توجه به شمار شرکت کنندگان (استادان و دستیاران) شش ایستگاه ۱۰ دقیقه‌ای طراحی گردید. گفتنی است که ایستگاه‌های با زمان کوتاه، اجازه ارزیابی کامل را به دستیار نمی‌دهد و ایستگاه‌های با زمان بیش از ۱۲ دقیقه محدود بودن شمار ایستگاه‌ها را در پی دارد (هاجز و همکاران، ۲۰۰۲). برای طراحی ایستگاه‌ها شش نفر از استادان گروه‌های روانپزشکی کشور که اعضای امتحانات بودند و پره‌بورد در روانپزشکی کشور بودند، همکاری نمودند. دبیر بورد روانپزشکی کشور نیز در گروه طراحی پرسش‌ها حضور داشت. نخست هدف‌های آموزشی متناسب با هدف‌های کارگاه طی کار گروهی فهرست شد و پس از ادغام هدف‌های فرعی، شش هدف کلی استخراج گردید و هر ایستگاه بر پایه یکی از این هدف‌های کلی تدوین شد (جدول ۱). در طی چندین جلسه کارگاهی، برای هر ایستگاه سناریویی تدوین گردید. هم‌چنین فهرست‌هایی بر پایه نمره‌دهی به روش لیکرت تنظیم شد. فهرست‌ها بر پایه نمره کلی هر ایستگاه و هدف‌های فرعی مرتبط با آن ایستگاه طراحی گردید. با توجه به این که نمره کلی امتحان شفاهی بورد تخصصی روانپزشکی ۱۵۰ می‌باشد، جمع کل نمره‌های هر فهرست در شش ایستگاه، ۱۵۰ تعیین گردید (جدول ۱). محتوای فهرست‌ها برای ارزیابی مهارت‌های دستیار در انجام فعالیت‌هایی است که در هر ایستگاه از وی خواسته می‌شود.

برای هر ایستگاه دو نفر ارزیاب از استادان انتخاب شدند که بایستی به‌طور مستقل چک‌لیست‌ها را برای هر دستیار پُر می‌کردند. افزون بر نمره چک‌لیست یک نمره کلی ارزیابی از دستیار هم در نظر گرفته شد که برای ارزیابی کلی از فرآیند عملکرد دستیار بود. این نمره بین صفر تا سه تعیین شد. فردی که وظیفه داشت نقش بیمار را بازی کند (SP) بر پایه سناریوی تنظیم شده برای هر ایستگاه می‌بایست در جلسه‌های گروهی تمرین می‌کرد. چون کارگاه آزمایشی بود، SP‌ها از میان استادان شرکت‌کننده در کارگاه انتخاب شدند.

از SP‌ها در هر ایستگاه خواسته شد یک نمره ارزیابی کلی از عملکرد دستیاران را بین صفر تا سه در نظر بگیرند. زمان هر ایستگاه ده دقیقه بود؛ از این رو زمان امتحان برای هر دستیار ۶۰ دقیقه محاسبه گردید. روز سوم کارگاه، نخست دستیاران در یک جلسه توجیهی یک ساعته شرکت کردند و با کلیات شیوه برگزاری کارگاه آشنا شدند. دستیار پیش از ورود به هر

با توجه به این که در ایران آزمون‌های شفاهی رشته روانپزشکی هنوز به صورت سنتی انجام می‌شود، آزمون OSCE با وجود توصیه‌های معاونت آموزشی تخصصی، برای امتحانات جامع دستیاری و ارتقاء و دانش‌نامه تخصصی، بورد تخصصی و دانشجویی و اینترنتی به کار گرفته نشده است. نخستین کارگاه را دانشگاه علوم پزشکی ایران، در سال ۱۳۸۳ پس از نه ماه جلسات مداوم و با برگزاری آزمون در نه ایستگاه برگزار نمود (تقوا و همکاران، ۱۳۸۶). دومین کارگاه در پاییز ۱۳۸۴ توسط مرکز تحقیقات علوم رفتاری اصفهان با شش ایستگاه و تمهیدات زمانی در پنج روز برگزار گردید که مقاله حاضر دستاورد تحلیل آماری این کارگاه و آزمون انجام شده است. یافته‌های کیفی این کارگاه که بر پایه بحث گروهی^۱ انجام شده، طی مقاله‌ای جداگانه ارایه می‌گردد. از این رو این آزمون با همکاری دانشگاه‌های علوم پزشکی ایران، تهران و اصفهان در مرکز تحقیقات علوم رفتاری اصفهان تدوین شد که ضمن آشناسدن دستیاران و استادان رشته روانپزشکی با این آزمون، روایی و پایایی آن نیز بررسی گردید.

روش

این بررسی به منظور تعیین روایی و پایایی آزمون OSCE در روانپزشکی در پاییز ۱۳۸۴ در مرکز تحقیقات علوم رفتاری اصفهان و با همکاری گروه‌های روانپزشکی دانشگاه‌های علوم پزشکی تهران، ایران و اصفهان انجام شده است. بدین منظور یک کارگاه سه روزه برگزار گردید و بر پایه بازبینی پژوهش‌های پیشین و تجربه انجام یک آزمون آزمایشی OSCE در سال گذشته در دانشگاه علوم پزشکی و خدمات بهداشتی - درمانی ایران، این کارگاه تدوین گردید. پیش از برگزاری کارگاه شماری از گروه‌های روانپزشکی تهران و نیمه مرکزی و جنوبی ایران، مدیران گروه‌ها و شماری از استادان و دستیاران روانپزشکی سال سوم برای شرکت در کارگاه آزمون فراخوانده شدند و با توجه به برخی محدودیت‌ها، برای حضور شرکت‌کنندگان در کارگاه، برای هر دانشگاه سهمیه‌ای در نظر گرفته شد. این سهمیه به تناسب شمار دستیاران هر دانشگاه بود و افراد به‌طور داوطلب در این آزمون شرکت کردند. به این ترتیب ۲۲ نفر از استادان و ۲۲ نفر از دستیاران از دانشگاه‌های علوم پزشکی اصفهان، ایران، اهواز، تهران و کرمان در این برنامه شرکت نمودند. هیچ‌یک از این دستیاران پیشتر در آزمون به روش OSCE شرکت

جدول ۱- هدفها، موضوع و مدت اجرا در ایستگاههای آزمون OSCE

شماره ایستگاه	هدف	فعالیت دستیار	مدت	دامنه نمره چک لیست
۱	درمان دارویی و غیردارویی اختلال‌های اضطرابی	ضمن مصاحبه با بیمار درمان دارویی و غیر دارویی بیمار مرد ۴۵ ساله مبتلا به OCD ^۱ را توضیح دهد	۱۰ دقیقه	۲۰-۰
۲	مشاوره خانواده	آموزش رابطه والد نوجوان به پدر ۴۰ ساله که فرزندش بیماری روانی دارد	۱۰ دقیقه	۲۰-۰
۳	معاینه وضعیت روانی ^۲	معاینه محتوای فکر مرد ۴۵ ساله مبتلا به سایکوز	۱۰ دقیقه	۳۰-۰
۴	تشخیص افتراقی ارگانیک و غیر ارگانیک	در مصاحبه با مرد ۵۵ ساله افسردگی را از دمانس تفکیک کند	۱۰ دقیقه	۳۰-۰
۵	آموزش به بیمار	به خانم ۳۰ ساله مبتلا به اختلال دو قطبی که قرار است از بخش ترخیص گردد آموزش‌های لازم را ارائه نماید	۱۰ دقیقه	۲۰-۰
۶	اداره بیمار ^۳	اداره بیمار ۲۵ ساله مبتلا به اختلال شخصیت مرزی	۱۰ دقیقه	۳۰-۰

1- obsessive-compulsive disorder, 2- mental status examination, 3- management

ایستگاه هوظیفه دستیار در آن ایستگاه را که پشت در نصب شده بود به مدت یک دقیقه مطالعه می‌کرد و سپس وارد اتاق می‌شد و مصاحبه و معاینه خود را با SP آغاز می‌کرد. یک دقیقه پیش از پایان زمان ده دقیقه‌ای، زنگ کوتاهی به صدا درمی‌آمد و پس از پایان زمان، زنگ ممتد به صدا در می‌آمد که باید دستیار بی‌درنگ اتاق را ترک کند. معمولاً در طی مصاحبه بین دستیار و استادان ارزیاب، پرسش و پاسخی رد و بدل نمی‌گردید و ارزیاب‌ها تنها بر پایه چک‌لیست‌ها و نمره کلی، نمره دهی می‌کردند و نمره SP هم منظور می‌گردید. در پایان کارگاه، جلسه بحث گروهی برای دستیاران برگزار گردید و فرم‌های نظر سنجی بین استادان و دستیاران توزیع شد. فرم‌های نظر سنجی در ارتباط با کل برنامه‌های کارگاه و هم چنین مربوط به هر ایستگاه به‌طور جداگانه بود.

برای تعیین روایی، روش روایی صوری^۱ به کار برده شد. در این رابطه طی کار گروهی، محتوا و سناریوها و چک‌لیست‌های هر کارگاه بررسی می‌شد و سپس در جمع استادان، شرکت‌کنندگان در کارگاه دوباره بحث و تبادل نظر می‌نمودند. برای تعیین پایایی^۲ از روش‌های پایایی بین ایستگاهی^۳، همسانی درونی^۴، ضریب وابستگی^۵ و پایایی موازی^۶ (بین ناظرین) بهره گرفته شد. برای این منظور ضریب همبستگی پیرسون به کار برده شد و همسانی درونی ایستگاه‌ها محاسبه گردید (جدول ۲).

یافته‌ها

میانگین کل نمره‌های چک‌لیست دستیاران $84/28 \pm 11/91$ و دامنه نمرات بین $56/25$ و $104/25$ بود (جدول ۲). ضریب وابستگی بین نمره‌های ارزیاب‌ها (در مجموع ۶ ایستگاه) برابر با $0/96$ و ضریب همبستگی آن $0/82$ به دست آمد.

جدول ۳ میانگین نمره‌های دستیاران را در ایستگاه‌های شش‌گانه نشان می‌دهد. ضریب همبستگی پیرسون بین میانگین نمره‌های چک‌لیست و نمره‌دهی کلی (GR) محاسبه گردید ($r=0/89$, $p<0/001$). به بیان دیگر بین نمره‌های چک‌لیست‌های دستیاران و نمره کلی ارزیاب همبستگی خطی مستقیم وجود داشت.

جدول ۴ پایایی درون ایستگاهی^۷ را برای هر ایستگاه به صورت جداگانه نشان می‌دهد. همان‌گونه که جدول نشان می‌دهد بین نمرات چک‌لیست هر یک از ارزیاب‌ها با نمره‌دهی کلی همان ارزیاب تقریباً در تمام ایستگاه‌ها رابطه خطی مستقیم وجود داشته است (به جز ارزیاب ۱ در ایستگاه ۱۰). همبستگی بین نمره‌های نمره‌دهی کلی (GR) دو ارزیاب، به جز ایستگاه پنج در سایر ایستگاه‌ها وجود داشته است.

برای بررسی ضریب وابستگی بین نمره‌های چک‌لیست‌ها و ارزیاب‌ها در هر ایستگاه، از ضریب‌های همبستگی درونی^۸ بهره گرفته شد. این بررسی نشان داد که غیر از ایستگاه شماره یک (به دلیل ضعیف بودن) و ایستگاه شماره چهار (به دلیل

1- face validity 2- reliability
3- inter-station reliability 4- interal consistency
5- contingency coefficient 6- parallel (interrater)
7- parallel reliability
8- intra-class correlation coefficient

جدول ۲- میانگین و انحراف معیار نمرات دستیاران* بر حسب نمره چک لیست و نمره کلی ارزیابی و ارزیاب‌ها

مجموع دو ارزیاب	ارزیاب ۲	ارزیاب ۱	
۸۴/۲۸ ± ۱۱/۹۱	۸۵/۳ ± ۱۲/۹۴	۸۳/۸ ± ۱۱/۹۶	نمره چک لیست
۱۰/۹۹ ± ۲/۳۰	۱۱/۱۸ ± ۲/۵۵	۱۰/۸۴ ± ۲/۲۲	نمره کلی ارزیابی (GA)

* ضریب وابستگی ۰/۹۶، ضریب همبستگی پیرسون ۰/۸۹، $r = ۰/۸۹$ ، $p < ۰/۰۰۱$

مقایسه گردید. ضریب همبستگی نمرات کلی SPها با GR ارزیاب ۱ در مجموع شش ایستگاه برابر با $۰/۷۷$ ($p < ۰/۰۰۱$) و با GR ارزیاب ۲ برابر با $۰/۸۰$ ($p < ۰/۰۰۱$) بود که نشانه ارتباط مستقیم خطی بین ارزیابی‌های کلی آنان می‌باشد. همسانی درونی ایستگاه‌های یک تا شش به کمک آلفای کرونباخ به ترتیب عبارت از $۰/۵۱$ ، $۰/۷۹$ ، $۰/۵۹$ ، $۰/۴۶$ ، $۰/۵۴$ و $۰/۷۶$ محاسبه شد. میانگین نمره SP در مجموع شش ایستگاه در ارزیابی از عملکرد دستیاران $۱۱/۳۹$ (انحراف معیار $۲/۴$) بود. روایی آزمون به صورت روایی صوری در حد بسیار خوب بود. هدف‌های آموزشی، سناریوها و چک لیست‌ها به وسیله کار گروهی بازیابی شد و در هر گروه یک‌بار سناریو و چک لیست آزمون گردید. در پایان کارگاه دوباره از استادان شرکت کننده نظرخواهی شد که محتوای چک لیست‌ها را بسیار خوب ارزیابی کردند.

ناکافی بودن داده‌ها) در سایر ایستگاه‌ها ضریب وابستگی خوب بوده است. ضریب همبستگی پیرسون بین نمره‌های چک لیست و نمره‌های نمره‌دهی کلی (GR) برای ارزیاب ۱ در شش ایستگاه برابر با $۰/۸۵$ و برای ارزیاب ۲ برابر با $۰/۸۹$ به دست آمد. به بیان دیگر از بین نمره‌هایی که ارزیاب‌ها به چک لیست‌ها داده‌اند و ارزیابی کلی که از دستیار داشته‌اند، همبستگی خطی مستقیم دیده شد. همبستگی نمره‌های چک لیست‌ها بین دو ارزیاب در مجموع شش ایستگاه برابر با $۰/۸۲$ و ضریب وابستگی برابر با $۰/۹۶$ به دست آمد. ضریب همبستگی بین GR ارزیاب ۱ و GR ارزیاب ۲ برابر با $۰/۸۶$ بود. هر کدام از SPها هم به‌طور جداگانه به عنوان ارزیابی کلی از دستیار بین صفر تا سه نمره داده بودند. نمره‌های SPها با نمره‌های ارزیابی کلی ارزیاب‌ها (GR)

جدول ۳- میانگین و انحراف معیار نمرات چک لیست و نمره‌دهی کلی (GR)، و میانگین نمرات چک لیست و نمره‌دهی کلی دو ارزیاب و بیشترین نمره ممکن دستیاران در هر ایستگاه

ایستگاه ۱	ایستگاه ۲	ایستگاه ۳	ایستگاه ۴	ایستگاه ۵	ایستگاه ۶	
۱۴/۲۵ ± ۳/۲۵	۱۲/۰۹ ± ۳/۱۹	۱۵/۵ ± ۵/۴۲	۸/۳۰ ± ۳/۶۸	۱۳/۶۴ ± ۳/۰۳	۲۰ ± ۴/۸۰	نمره چک لیست
۱۲/۶۸ ± ۲/۴۲	۱۲/۶۴ ± ۳/۳۰	۱۶/۹۲ ± ۵/۳۵	۱۱/۵۴ ± ۴/۰۸	۱۱/۷۷ ± ۱/۷۷	۱۹/۶۸ ± ۵/۶۱	ارزیاب ۲
۲/۳۲ ± ۰/۵۷	۱/۸۶ ± ۰/۶۴	۱/۵۹ ± ۰/۷۲	۱/۲۹ ± ۰/۴۸	۲/۰۴ ± ۰/۷۸	۱/۷۳ ± ۰/۷۷	ارزیاب ۱
۲/۲۷ ± ۰/۷۰	۱/۸۶ ± ۰/۷۱	۱/۸۲ ± ۰/۵۷	۱/۷۷ ± ۰/۶۸	۱/۵۹ ± ۰/۶۷	۱/۸۶ ± ۰/۶۴	ارزیاب ۲
۱۳/۴۶ ± ۲/۳۵	۱۲/۳۶ ± ۳/۰۵	۱۶/۲۱ ± ۵/۰۳	۹/۹۲ ± ۳/۴۵	۱۲/۴۸ ± ۲/۳۲	۱۹/۸۴ ± ۴/۹۶	میانگین نمره چک لیست دو ارزیاب
۲/۲۷ ± ۰/۵۹	۱/۸۶ ± ۰/۶۲	۱/۷۰ ± ۰/۶۰	۱/۵۳ ± ۰/۵۲	۱/۸۲ ± ۰/۵۹	۱/۷۹ ± ۰/۶۳	میانگین نمره GR دو ارزیاب
۲۰	۲۰	۳۰	۳۰	۲۰	۳۰	حداکثر نمره ممکن چک لیست دو هر ایستگاه

جدول ۴- ضرایب همبستگی پیوسون برای نمره‌های چک‌لیست هر یک از ارزیاب‌ها و ارزیابی‌های درون ایستگاهی

شماره ایستگاه‌ها	ضرایب همبستگی درون ایستگاهی		همبستگی GR بین دو ارزیاب		همبستگی بین نمره چک‌لیست و GR ارزیاب ۲		ICC
	p	r	p	r	p	r	
ایستگاه ۱	N.S.	۰/۱۶	۰/۰۰۱	۰/۸۴	۰/۰۰۱	۰/۷۵	۰/۰۰۱
ایستگاه ۲	۰/۰۰۱	۰/۹۳	۰/۰۰۱	۰/۶۹	۰/۰۰۱	۰/۸۱	۰/۰۰۱
ایستگاه ۳	۰/۰۰۱	۰/۸۳	۰/۰۰۱	۰/۷۷	۰/۰۰۱	۰/۹۰	۰/۰۰۱
ایستگاه ۴	داده‌ها برای ارزیابی کافی نمی‌باشد		۰/۰۰۱	۰/۵۸	۰/۰۱	۰/۵۸	۰/۰۱
ایستگاه ۵	۰/۰۵	۰/۷۳	N.S.	۰/۳۱	۰/۰۰۱	۰/۶۰	۰/۰۱
ایستگاه ۶	۰/۰۰۱	۰/۷۷	۰/۰۱	۰/۶۰	۰/۰۰۱	۰/۹۰	۰/۰۰۱
مجموع شش ایستگاه			۰/۰۰۱	۰/۸۶	۰/۰۰۱	۰/۸۹	۰/۰۰۱

بحث

ارزیابی‌ها در این آزمون بسیار خوب بود. از سوی دیگر نمراتی که ارزیاب‌ها برای هر دستیار از طریق چک‌لیست منظور می‌کردند با نمره‌دهی کلی (GR) که آنان منظور کرده بودند همبستگی خطی مستقیم داشتند (به جز در ایستگاه درمان دارویی و غیردارویی و سواس - ارزیاب یک) به بیان دیگر اگر ارزیابی کلی استادان در مورد یک دستیار ضعیف برآورد شده بود، این وضعیت با نمرات ضعیف در چک‌لیست مربوط به دستیار مطابقت داشت. از میان شش ایستگاه، تنها در دو ایستگاه اول و پنجم همبستگی نمرات چک‌لیست‌ها بین دو ارزیاب ضعیف بود و توافق قابل قبولی وجود نداشت. هم‌چنین در ایستگاه پنجم (آموزش خانواده) همبستگی بین نمره‌دهی کلی (GA) میان دو ارزیاب دیده نشد. به نظر می‌رسد که طراحی ایستگاه‌های اول و پنجم از نظر محتوایی و چک‌لیست‌ها و SPها نیاز به بازنگری داشته‌اند. در مجموع ضریب وابستگی این آزمون برابر ۰/۹۶ و ضریب همبستگی ۰/۸۹ به دست آمد. یافته‌های این بررسی با یافته‌های والترز و همکاران (۲۰۰۵) که پایایی آزمون OSCE در روانپزشکی را در حد متوسط تا خوب برآورد نمودند (۰/۶۸ - ۰/۵۵) هم‌سویی دارد. در چند بررسی انجام‌شده در نیجریه و کانادا میزان روایی و پایایی آزمون OSCE در روانپزشکی در حد قابل قبول گزارش شده است (فامایوا و همکاران، ۱۹۹۱؛ هاجز و همکاران، ۱۹۹۷؛ هاجز و همکاران، ۱۹۹۸). فامایوا و همکاران (۱۹۹۱) پایایی آزمون OSCE روانپزشکی را برای دانشجویان دوره عمومی در نیجریه برابر ۰/۷۶ گزارش نمودند. در بررسی هاجز و همکاران (۱۹۹۷) در تورنتو

در حال حاضر آزمون مورد شفاهی در ایران در بیشتر رشته‌های بالینی به صورت OSCE برگزار می‌گردد. از آنجا که در روانپزشکی تجربه کافی برای استادان و دستیاران در زمینه OSCE وجود نداشت و از سویی طراحی ایستگاه‌ها، آموزش و آماده‌سازی SP نیاز به وقت بیشتری داشت، امتحان شفاهی به همان روش سنتی برگزار می‌شد. از این رو، هیأت بودرد تخصصی روانپزشکی تصمیم گرفت که برای آشنایی استادان و دستیاران با روش آزمون OSCE، هم‌چنین ارزیابی روایی، پایایی و میزان عملی بودن آن، چند کارگاه برگزار نماید.

میانگین نمره کل چک‌لیست‌ها (۸۴/۲۸) از میانگین نمره دستیاران سال آخر در امتحان شفاهی بودرد روانپزشکی سال ۱۳۸۴ (۱۰۸) پایین‌تر است. از این رو بر پایه تجربه کارگاه اول (دانشگاه علوم پزشکی ایران - ۱۳۸۳) و بازبینی بررسی‌های پیشین، نسبت ۶۰٪ برای قبولی و نمره برش ۷۶ برای حد قبولی منظور گردید؛ به بیان دیگر ۶۰٪ افراد شرکت‌کننده نمره بالای ۷۶ داشتند. بنابراین لازم است در آزمون‌های OSCE در روانپزشکی، نمرات کل با در نظر گرفتن GA، روی منحنی برده شود و حد نصاب ۱۰۵ که در حال حاضر در آزمون شفاهی بودرد در نظر گرفته می‌شود تعدیل گردد (هاجز و همکاران، ۲۰۰۲).

ضریب وابستگی بین دو ارزیاب در مجموع شش ایستگاه ۰/۹۶ و ضریب همبستگی نمرات چک‌لیست میان دو استاد ارزیابی‌کننده ۰/۸۲ به دست آمد. از این رو پایایی بین

امتحان OSCE دستیارانی که در سمینار شرکت کرده بودند، به‌طور معنی‌داری بیشتر از نمرات دستیارانی بود که در سمینار شرکت نکرده بودند. به کمک آزمون OSCE می‌توان به‌عنوان یک آزمون معتبر، عملکرد دانشجویان را در سال‌های پس از آموزش نیز پیش‌بینی کرد (مارتین^۷ و جولی^۸، ۲۰۰۲).

یکی از مشکلاتی که در آزمون OSCE در روانپزشکی مطرح است، طراحی ایستگاه برای روانپزشکی کودکان است. انتخاب SP برای دوره نوجوانی با محدودیت‌هایی همراه است ولی به هر روی دست‌کم در دو بررسی در کانادا این امر تحقق یافته است (هانسون و همکاران ۲۰۰۲؛ هانسون، هاجز، مک‌ناگتون و رگهر، ۱۹۹۸). در ضمن طراحی برخی از ایستگاه‌ها می‌تواند به کمک ویدئو یا رایانه برای ارزیابی مهارت‌های ارتباطی تنظیم گردد (هامفریز^۹ و کانی^{۱۰}، ۲۰۰۰). فلی‌گل^{۱۱}، فروهنا^{۱۲} و منگروولکار^{۱۳} (۲۰۰۲) یک الگوی رایانه‌ای آزمون OSCE برای دانشجویان پزشکی (دوره عمومی) طراحی کردند. در این الگو به جای SP از الگوهای رایانه‌ای بهره گرفته شد و پرسش‌هایی برای هر ایستگاه طراحی گردید.

به‌طور کلی آزمون شفاهی به روش OSCE در روانپزشکی قابل اجرا بوده و از روایی و پایایی مناسبی برخوردار است. اجرایی شدن آن نیاز به کار گروهی و همکاری استادان با تجربه و علاقه‌مند در این زمینه دارد.

از جمله محدودیت‌های این طرح، محدود بودن جامعه آماری بود. پیشنهاد می‌شود در آزمون شفاهی مورد تخصصی روانپزشکی که در آن بیش از ۸۰ نفر شرکت کننده وجود دارد، این محدودیت از میان برداشته شده، تحلیل کامل آماری انجام شود. از نکات قابل توجه در این بررسی، به کارگیری استادان برای SP بود. در امتحان رسمی OSCE لازم است از افراد ناشناخته و آموزش دیده برای SP بهره گرفته شود.

سپاسگزاری

از استادان گرانقدر دانشگاه علوم پزشکی ایران به‌ویژه جناب آقای دکتر میرفرهاد قلعه‌بندی و جناب آقای دکتر مهدی نصر اصفهانی، هم‌چنین از استادان گرامی گروه

کانادا پایایی OSCE روانپزشکی برای دو گروه ۹۴ و ۹۸ نفره (در دو رشته ایستگاه موازی) به ترتیب برابر ۷۰/۴۷ و ۶۷/۶۶ به‌دست آمد. در بررسی دیگری هاجز و همکاران (۱۹۹۸) OSCE در روانپزشکی را برای ارزیابی مهارت‌های بالینی دانشجویان روانپزشکی، معتبر گزارش نمودند.

روایی صورتی این آزمون هم‌چون بررسی دیگر OSCE روانپزشکی در ایران (تقوا و همکاران، ۱۳۸۶) در حد بسیار خوب ارزیابی گردید. پارک و همکاران (۲۰۰۴) روایی آزمون OSCE در روانپزشکی را بررسی نموده، روایی محتوای چک‌لیست‌ها را بسیار بیشتر از نمره کلی برآورد نمودند و به کارگیری چک‌لیست همراه با ارزیابی کلی را در OSCE روانپزشکی سودمند دانستند.

در زمینه شمار ایستگاه‌ها و زمان هر ایستگاه نظرات متفاوتی مطرح شده است. در بررسی‌های گوناگون شمار ایستگاه‌ها بین چهار (پارک و همکاران، ۲۰۰۴) تا ۱۸ ایستگاه (والترز و همکاران، ۲۰۰۵) متفاوت گزارش شده است. نعیم^۱، راتر فورد^۲ و کن^۳ (۲۰۰۴) پنج تا شش ایستگاه ۷ دقیقه‌ای را پیشنهاد داده‌اند؛ در حالی که هانسون^۴ و همکاران (۲۰۰۲) هشت تا ده ایستگاه ۱۵-۱۲ دقیقه‌ای را بسیار مناسب دانسته‌اند. هر چه شمار ایستگاه‌ها بیشتر باشد، پایایی آزمون بیشتر است و از نمره مبتنی بر مورد^۵ جلوگیری می‌شود (والترز و همکاران، ۲۰۰۵). ولی باید در نظر داشت که بهتر است آزمون در یک روز انجام گردد تا دانشجویانی که امتحان داده‌اند با کسانی که هنوز امتحان را آغاز نکرده‌اند، در ارتباط قرار نگیرند. از این رو اگر شمار دستیاران زیاد باشد به ناچار باید شمار ایستگاه‌ها را کمتر نمود. از بازمینی بررسی‌های پیشین چنین بر می‌آید که کمترین ایستگاه‌های مورد نیاز در آزمون OSCE پنج تا شش ایستگاه است (هاجز و همکاران، ۲۰۰۲؛ نعیم و همکاران، ۲۰۰۴). با توجه به این که در بررسی دیگر OSCE در ایران (تقوا و همکاران، ۱۳۸۶) شمار ایستگاه‌ها هشت و در بررسی حاضر شش ایستگاه بود پیش‌بینی می‌شد که پایایی آزمون کاهش یابد؛ ولی با وجود کاهش شمار ایستگاه‌ها، پایایی در حد قابل قبول بود. با وجود این پیشنهاد می‌شود که برای افزایش پایایی، شمار ایستگاه‌ها به هشت افزایش یابد.

زمان ۱۰ دقیقه برای هر ایستگاه را بیشتر دستیاران و استادان شرکت‌کننده مناسب تشخیص دادند. آموزش به دستیاران برای آشنایی با شیوه OSCE ضروری است. در یک سمینار روانپزشکی که برای آموزش OSCE در تورنتو کانادا برگزار گردید (ماس^۶، ۱۹۹۰)، نشان داده شد که نمرات

- | | |
|---------------------|---------------|
| 1- Naeem | 2- Rutherford |
| 3- Kenn | 4- Hanson |
| 5- case specificity | 6- Moss |
| 7- Martin | 8- Jolly |
| 9- Humphris | 10- Kaney |
| 11- Fliegel | 12- Frohna |
| 13- Mangrulkar | |

Harden, R. M., & Gleason, F. A. (1979). Assessment of clinical competence using an observed structured clinical examination. *Medical Education*, 13, 41-47.

Hodges, B., Hanson, M., McNaughton, N., & Regehr, G. (2002). Creating, monitoring and improving a psychiatry OSCE. *Academic Psychiatry*, 26, 134-161.

Hodges, B., Regehr, G., Hanson, M., & McNaughton, N. (1998). Validation of an objective structured clinical examination in psychiatry. *Academic Medicine*, 73, 910-912.

Hodges, B., Regehr, G., Hanson, M., & McNaughton, N. (1997). An objective structured clinical examination for evaluating psychiatric clinical clerks. *Academic Medicine*, 72, 715-721.

Hodges, B., Regehr, G., McNaughton, N., Tiberius, R., & Hanson, M. (1999). OSCE checklists do not capture increasing levels of expertise. *Academic Medicine*, 74, 1129-1139.

Hodges, B. (2003). Validity and the OSCE. *Medical Teacher*, 25, 250-254.

Humphris, G. M., & Kaney, S. (2000). The objective structured video examination for assessment of communication skills. *Medical Education*, 34, 939-995.

Martin, I. G., & Jolly, B. (2002). Predictive validity and estimated cut score of an objective structured clinical examination (OSCE) used as an assessment of clinical skills at the end of the first clinical year. *Medical Education*, 36, 418-425.

Moss, J. H. (1990). Evaluating a seminar designed to improve psychiatry skills of family medicine residents. *Academic Medicine*, 65, 658-660.

Naeem, A., Rutherford, J., & Kenn, C. (2004). The MRC Psych OSCE workshop: A new game to play? *Psychiatric Bulletin*, 28, 62-65.

Park, R. S., Chibnall, J. T., Blaskiewicz, R. J., 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.

Walters, K., Osborn, D., & Raven, P. (2005). The development, validity and reliability of a multimodality objective structured clinical examination in psychiatry. *Medical Education*, 39, 292-298.

روانپزشکی دانشگاه‌های علوم پزشکی اصفهان، اهواز، کرمان و ایران و دانشگاه علوم بهزیستی و توانبخشی، از آقای حسین شاره و کلیه اعضای مرکز تحقیقات علوم رفتاری اصفهان قدردانی می‌گردد.

از همکاری انستیتو روانپزشکی تهران، مرکز تحقیقات بهداشت روان و دانشگاه علوم پزشکی اصفهان که با حمایت‌های مالی و کمک‌های صمیمانه ما را در انجام این پژوهش یاری دادند، صمیمانه سپاسگزاری می‌شود.

دریافت مقاله: ۱۳۸۵/۲/۲۳؛ دریافت نسخه نهایی: ۱۳۸۵/۴/۱۴؛

پذیرش مقاله: ۱۳۸۵/۵/۴

منابع

تفوا، ارسیا؛ رسولیان، مریم؛ پناغی، لیلی؛ بوالهروی، جعفر؛ ضرغامی، مهران؛ نصرافصفا، مهدی؛ حاجز، برابان (۱۳۸۶). پایایی و روایی نخستین آزمون ساختاریافته عینی - بالینی (OSCE) روانپزشکی در ایران. *مجله روانپزشکی و روانشناسی بالینی ایران*، سال سیزدهم، شماره ۱، ۲۴-۱۷.

Barrows, H. S., & Abrahamson, S. (1964). The programmed patient: A technique for appraising student performance in clinical neurology. *Medical Education*, 39, 802-805.

Blaskiewicz, R. J., Park, R. S., Chibnall, J. T., & Powell, J. K. (2004). The influence of testing context and clinical rotation order on student's OSCE performance. *Academic Medicine*, 79, 597-601.

Famuyiwa, O. O., Zachariah, M. P., & Ilchukwu, S. T. (1991). The objective structured clinical examination in undergraduate psychiatry. *Medical Education*, 25, 45-50.

Fliegel, J. E., Frohna, J. G., & Mangrulkar, R. S. (2002). A computer-based OSCE station to measure competence in Evidence-based Medicine Skills in medical students. *Academic Medicine*, 77, 1157-1158.

Hanson, M., Hodges, B., McNaughton, N., & Regehr, G. (1998). The integration of child psychiatry into a psychiatry clerkship OSCE. *Canadian Journal of Psychiatry*, 93, 614-618.

Hanson, M., Tiberius, R., Hodges, B., Mackay, S., McNaughton, N., & Dickens, S. (2002). Adolescent standardized patients: Method of selection and assessment of benefits and risks. *Teaching and Learning in Medicine*, 14, 104-113.