روایی سازه و پایایی ابزارهای ارزشیابی امتحان پایان دوره تحصیلی دانشجویان مامایی دانشکده پرستاری و مامایی مشهد

ناهید جهانی شوراب*، ناهید گلمکانی، سیدرضا مظلوم، کبری میرزاخانی، صدیقه اظهری، افسانه نوابیان

چکیده

مقدمه: عملکرد نامطلوب مراقبین زایمانی یکی از علل مهم مرگ مادران است. فقدان یک ابزار عینی، امکان ارزیابی سیستماتیک و معتبر را در امتحانات بالینی پایان دوره مامایی کاهش می‌دهد. بنابراین پژوهش حاضر با هدف بررسی روایی و پایایی ابزارهای سنجش و ارزیابی در امتحان جامع پایان دوره تحصیلی دانشجویان مامایی انجام شد.

روش‌ها: در این مطالعه همبستگی مهارت‌های دانشجویان مامایی در ۴ حیطه امتحانی (زایشگاه، درمانگاه زنان، مراقبت‌های دوران بارداری، و بهداشت مادر و کودک) مطابق ابزارهای ارزشیابی امتحان جامع پایان دوره موجود در دانشکده پرستاری و مامایی مشهد در تیرماه ۸۹ بررسی شد. ۳۱ دانشجو (۲۲ نفر دانشجوی ترم ۸، و ۹ نفر دانشجوی ترم ۶ مامایی) در طی ۵ روز متوالی توسط دو ممتحن ارزیابی گردیدند. روایی محتوی با استفاده از نظرات کارشناسان تعیین شد. جهت بررسی روایی سازه از مقایسه میانگین نمرات کسب شده در هر حیطه امتحانی و میانگین نمره کل دانشجویان در ترم‌های بالا و پایین (ترم‌های ۸ و۶) محاسبه شد. برای اندازه گیری شاخص روایی از فرمول لاوشه استفاده شد. همچنین برای بررسی پایایی، ضریب آلفای کرونباخ تعیین شد.

نتایج: وجود تفاوت معنادار آماری بین میانگین نمره کل دانشجویان ترم ۸ با دانشجویان ترم۶ در امتحان جامع پایان دوره تحصیلی به میزان ۲۶ درصد، وجود روایی سازه در فرم‌های ارزشیابی دانشجویان مامایی را بر طبق تعریف کالج سلطنتی زنان و مامایی در این آزمون اثبات می‌کند (p=۰/۰۱). ضرایب آلفای کرونباخ متوسط تا بالا (۰/۵۱-۰/۸۳) در هر حیطه امتحانی بیانگر وجود همسانی درونی در ابزارهای ارزشیابی موجود است. روایی محتوی نیز با شاخص روایی بالا (۰/۸) محاسبه شد. کمترین و بیشترین پایایی ارزیابان به ترتیب ۰/۸۲ و ۷۳/۵ درصد با تعریف معیارهای ارزش هر نمره مطابق تعاریف دانشگاه سین سیناتی تعیین شد.

نتیجه‌گیری: ابزارهای ارزشیابی امتحان جامع پایان دوره کارشناسی مامایی در دانشکده پرستاری و مامایی مشهد به عنوان فرم‌های معتبر و پایا برای ارزیابی دانشجویان در زمان دانش‌آموختگی توصیه می‌شوند.

واژه‌های کلیدی: روایی، پایایی، ابزار ارزشیابی، ارزشیابی، دانش‌آموختگان، مهارت‌های بالینی

مجله ایرانی آموزش در علوم پزشکی/ آذر۱۳۹۰؛ ۱۱(۵):۴۸۵ تا ۴۹۵

مقدمه

بخش مهمی از فرایند آموزشی، ارزشیابی پیشرفت

* **نویسنده مسؤول:** ناهید جهانی شوراب (مربی)، کارشناس ارشد مامایی، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی مشهد، خراسان رضوی، ایران. jahanishn@mums.ac.ir
ناهید گلمکانی (مربی)، کارشناس ارشد مامایی، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی مشهد، خراسان رضوی، ایران. (golmakanin@mums.ac.ir)؛ دکتر سیدرضا مظلوم (استادیار)، دکترای پرستاری، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی مشهد، خراسان رضوی، ایران. (mazlomr@mums.ac.ir)؛ کبری میرزاخانی، کارشناس ارشد مامایی، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی مشهد، خراسان رضوی، ایران. (mirzakhanik@mums.ac.ir)؛ صدیقه اظهری، کارشناس ارشد مامایی، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی مشهد، خراسان رضوی، ایران. (azharis@mums.ac.ir)؛ افسانه نوابیان، کارشناس ارشد مامایی، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی مشهد، خراسان رضوی، ایران. (navaiyana@mums.ac.ir).

این مقاله در تاریخ ۸۹/۸/۳۰ به دفتر مجله رسیده، در تاریخ ۹۰/۵/۱۰ اصلاح شده و در تاریخ ۹۰/۵/۱۸ پذیرش گردیده است.

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

تحصیلی و ارزشیابی میزان یادگیری دانشجویان است. مطابق تعریف، ارزشیابی پیشرفت تحصیلی عبارت است از سنجش عملکرد یادگیرندگان و مقایسه نتایج حاصل با هدف‌های آموزشی ازپیش تعیین شده، به منظور تصمیم‌گیری در این باره که آیا فعالیت‌های آموزشی به نتایج مطلوب رسیده‌اند؟ آیا هدف‌های موردنظر تحقق یافته‌اند؟ و به چه میزانی به این اهداف دست یافته‌ایم(۱). کرونباخ ارزشیابی را جمع‌آوری و کاربرد اطلاعات به منظور تصمیم‌گیری درباره یک برنامه آموزشی تعریف کرده است(۲). لذا باید از تناسب و سازگاری روش‌های ارزشیابی با اهداف ویژه یادگیری اطمینان حاصل شود. متأسفانه در حال حاضر بر نحوه برگزاری امتحانات معمول بالینی ایراداتی وارد است که از آن جمله می‌توان به تعداد زیاد شرکت‌کننده در هر امتحان،زمان کوتاه اختصاص یافته برای ارزشیابی، وغیر عینی بودن ارزشیابی‌ها اشاره کرد(۳). تحقیقی در سال ۱۳۸۰ عدم رضایت دانشجویان پزشکی را از نحوه ارزشیابی فعلی گزارش می‌کند؛ به طوری که دانشجویان در ٤٦/٦ درصد موارد عدم انطباق نمره پایان‌بخش را با توانایی‌های کسب شده گزارش نمودند(٤). حتی گاهی عدم رضایت از سنجش‌ها توسط مربیان بالینی نیز اعلام می‌شود، به طوری که در ارزشیابی مهارت‌های بالینی پرستاری در اندونزی ۲۰ درصد مربیان عدم رضایت خود را از روش بررسی اعلام نمودند و در ۲۱ درصد موارد مربیان گزارش کردند در صورت شکست دانشجو در امتحان فینال بر اثر قضاوت آنان احساس ناراحتی می‌کنند؛ و حتی ۵۲ درصد مربیان گزارش نموده‌اند در زمان شکست دانشجویان احساس مسؤولیت می‌نمایند(۵). از سوی دیگر اهمیت ارزشیابی مطلوب دانشجویان مامایی در زمان فارغ‌التحصیلی هنگامی آشکار می‌شود که به عواقب فقدان مهارت لازم در زمینه کنترل زنان و به خصوص زنان باردار توجه گردد. طبق گزارش سازمان جهانی بهداشت وقوع ۵۱۵ هزار مرگ مادر سالانه در جهان نیاز به آموزش و افزایش بیشتر مهارت‌های زایمانی را نشان می‌دهد(٦). با وجود آن که ۹۷ درصد زایمان‌ها در ایران توسط ماماها صورت می‌گیرد(۷)؛ در بررسی موردی علل مرگ مادران عمده‌ترین دلیل، عملکرد بد و ناصحیح در بخش زایمانی است(۸). توجه به این حقایق، حساسیت و تلاش مربیان بالینی مامایی را در بررسی مهارت‌های ضروری مامایی افزایش می‌دهد.

روش‌های متعددی در بررسی صلاحیت‌های بالینی مانند کارپوشه (پورت فولیو)، OSCE، و مشاهده مستقیم پیشنهاد شده است(۹). اما سنجش مهارت‌های بالینی گروه علوم پزشکی، عمدتاً با مشاهده مستقیم میسر می‌گردد(۱۰).

هم اکنون در دانشکده پرستاری و مامایی مشهد نیز برای ارزیابی بالینی دانشجویان مامایی از بین ابزارهای مشاهده‌ای، فهرست وارسی (مصوب شورای عالی برنامه‌ریزی و نظارت سال٨٤) در چهار بخش زایشگاه، زنان، بهداشت مادر و کودک و پره‌ناتال استفاده می‌شود. در روش فهرست وارسی تنها بود و یا نبود یا کمیت ویژگی موردنظر ثبت می‌شود. ویژگی‌های یک فهرست وارسی خوب شامل کوتاه بودن، روشن بودن، تأکید بر رفتار قابل مشاهده و ثبت رفتارهای مهم است(۱).

صرف‌نظر از شیوه سنجش صلاحیت‌های بالینی باید به معیارهای یک آزمون عادلانه که شامل روایی، پایایی و عینیت است توجه نمود. مطابق تعریف، آزمونی دارای روایی است که برای اندازه‌گیری آنچه که موردنظر است مناسب باشد و پایایی ابزار به این معناست که ابزار در زمان‌های متفاوت با شرایط یکسان نتایج یکسان را به دست دهد(۱۱).

کالج سلطنتی زنان و مامایی تأیید روایی سازه را برای هر ابزاری که جهت سنجش مهارت به کار می‌رود الزامی برشمرده است(۲۱). از طرفی ماهیت پیچیده روش‌های تحلیل عاملی جهت تعیین روایی سازه، باعث می‌شود که کمتر روایی سازه فرم‌های ارزشیابی مهارت‌های بالینی

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

تعیین گردد و این در حالی است که در بسیاری از مواقع، دانشجویان در امتحانات بالینی به غیرواقعی بودن نمرات اعتراض دارند(۱۰). یک راه عملی تعیین روایی سازه، در صورتی که معیار طلایی برای اندازه‌گیری ساختاری آزمون وجود نداشته باشد، این است که ببینیم آیا نمره افراد با سطوح مهارت‌های مختلف در این آزمون متفاوت است یا خیر؟(۱۳). شرط اثبات روایی سازه آزمون‌های مهارتی، توانایی آزمون در تشخیص تعلیم‌یافتگان از افراد آموزش‌ندیده و حتی تعیین سطوح آموزشی فراگیران نسبت به یکدیگر (تمایز سطوح آموزشی بالاتر از پایین‌تر) است. اختلاف میانگین نمرات عملکردی دانشجویان بین دو سال تحصیلی در حد ۲۵ درصد روایی سازه ابزار را حمایت می‌کند(۱۲).

فقدان اهداف عینی و نداشتن یک ابزار عینی، نه تنها امکان ارزیابی سیستماتیک و معتبر دانشجویان را در زمان فارغ‌التحصیلی کاهش می‌دهد بلکه طراحی برنامه‌های مداخله‌ای برای بهبود عملکرد دانشجویان را پس از فراغت از تحصیل غیرممکن می‌سازد. بنابراین مطالعه حاضر با هدف «بررسی روایی سازه و پایایی ابزارهای ارزشیابی امتحان پایان دوره تحصیلی دانشجویان مامایی دانشکده پرستاری و مامایی مشهد» انجام شد.

روش‌ها

این مطالعه همبستگی در بیمارستان‌ها و مراکز آموزشی دانشگاه علوم پزشکی مشهد (بیمارستان ام‌البنین و امام رضا، مراکز بهداشتی دانش‌آموز و نجفی) در تیرماه ۱۳۸۹ انجام‌گردید. جامعه پژوهش، دانشجویان مامایی مقطع کارشناسی دانشکده پرستاری و مامایی مشهد و واحدهای مورد پژوهش نیز دانشجویان ترم ٦ و ٨ مامایی در زمان انجام مطالعه بودند.

معیارهای موردنظر برای ورود به این پژوهش عبارت بودند از دارا بودن شرایط ورود به کارآموزی در عرصه برای دانشجویان داوطلب ترم ٦ و دارا بودن شرایط آموزشی لازم برای امتحان جامع پایان دوره برای دانشجویان داوطلب ترم ٨، و نیز نداشتن سابقه شرکت در امتحان جامع پایان دوره. همچنین به منظور آشنا بودن تمامی شرکت‌کنندگان با محیط امتحان، دانشجویانی در مطالعه شرکت داده شدند که واحد کارآموزی مربوطه را قبلاً در محل برگزاری امتحان گذرانده باشند. عدم تمایل داوطلبان به ادامه شرکت در امتحانات بالینی به عنوان معیارهای خروج از مطالعه در نظر گرفته شد.

به دلیل محدود بودن جامعه پژوهش، نمونه‌گیری به روش سرشماری انجام شد. در ابتدا اهداف مطالعه برای دانشجویان ترم ٨ (٢٢ نفر) و ترم ٦ مامایی (٩ نفر) شرکت‌کننده توضیح داده شد. پس از کسب رضایت از شرکت‌کنندگان، برنامه‌ای شامل زمان و محیط‌های امتحانی همچنین فرم‌های ارزشیابی امتحانی در هر واحد بالینی جهت اطلاع از موضوعات آزمون در اختیار آنان گذاشته شد. ارزیابان خارج از دانشکده ٤ دانشجوی کارشناسی ارشد با سابقه مربی‌گری بالینی بودند. ممتحنین، مربیان بالینی دانشکده بودند که در امتحان فینال شرکت داشتند. قبل از شروع امتحان، توسط پژوهشگر و مسؤولین هر واحد امتحان بالینی توضیحات لازم جهت اهداف پژوهش، چگونگی ارزیابی و نحوه تکمیل فرم‌های ارزشیابی به ممتحنین و ارزیابان داده شد. لازم به توضیح است که در دانشکده پرستاری و مامایی مشهد دروس مامایی در ٤ حیطه بهداشت مادر و کودک، زنان، زایشگاه و پره‌ناتال طبقه‌بندی می‌شوند و هر واحد دارای یک مسؤول است. پس از اطمینان از آشنایی کامل ممتحنین و ارزیابان با شیوه تکمیل فرم‌ها این ابزارها، برای ٥ نفر از دانشجویان در حین کارآموزی در چهار حیطه مختلف توسط ارزیاب موردنظر با مسؤول واحد بالینی تکمیل و ضریب همبستگی نمرات ارزیاب و مسؤول واحد بالینی تعیین شد. در صورت اخذ ضریب همبستگی مورد تأیید ($r>0/7$)، فرد منتخب به عنوان ارزیاب در امتحان شرکت می‌نمود. امتیازدهی فرم‌های ارزشیابی

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

(فهرست‌های وارسی) به صورت لیکرت ۵ قسمتی بود (بسیار ضعیف، ضعیف، متوسط، خوب، بسیارخوب) که به ترتیب برحسب عملکرد دانشجو نمرات صفر، ۲۵، ۵۰، ۷۵، ۱۰۰ درصد در برابر هر گویه به وی اختصاص می‌یافت. جهت عینی‌تر شدن تعاریف عملکرد (بسیار ضعیف تا بسیار خوب) و اختصاص نمرات به روش دقیق‌تر، در مطالعه حاضر توضیح به روش دانشگاه سین سیناتی (۱٤) نیز داده شد و امتیازات به صورت زیر نیز تعریف گردید:

امتیاز صفر: عملی که دانشجو باید انجام می‌داده ولی انجام نداده است. (هیچ نمره‌ای از گویه به دانشجو تعلق نمی‌گرفت)

امتیاز ۱: عملکرد دانشجو ناقص و یا اشتباه بوده است (با راهنمایی هم دانشجو نمی‌توانست عملکرد قابل قبول داشته باشد و در این صورت ۲۵ درصد نمره گویه پرسشنامه به دانشجو تعلق می‌گرفت).

امتیاز ۲: دانشجو جهت انجام عمل نیازمند کمک و یا راهنمایی است (۵۰ درصد نمره گویه در پرسشنامه به دانشجو تعلق می‌گرفت)

امتیاز ۳: دانشجو برای انجام مهارت نیاز به حداقل کمک دارد (عملکرد صحیح وقابل قبول است که ۷۵ درصد نمره گویه پرسشنامه به دانشجو تعلق می‌گرفت).

امتیاز ٤: عملکرد دانشجو کاملاً با مهارت و بدون کمک انجام شده است (در این صورت۱۰۰ درصد نمره گویه پرسشنامه به دانشجو تعلق می‌گرفت).

به طوری که امتیاز صفر این سیستم معادل بسیارضعیف در مقیاس لیکرت و امتیاز ٤ برابر بسیار خوب در نظر گرفته شد.

در مطالعه مقدماتی امتیازدهی به هردو صورت انجام شد ولیکن به دلیل پایین‌تر بودن ضرایب همبستگی در روش امتیازبندی لیکرت، در مطالعه انجام شده از روش امتیازدهی سین سیناتی استفاده گردید.

امتحان در ۵ روز متوالی در هر ٤ واحد بالینی با حضور دو ممتحن برگزار گردید و فرم‌های ارزشیابی بالینی هر دانشجو توسط ممتحن و ارزیاب تکمیل شد. از تمامی ممتحنین خواسته شده بود در خصوص ترم تحصیلی، زمینه‌های آشنایی قبلی و یا نظرات شخصی خود درباره دانشجویان مورد ارزیابی تا پایان تکمیل فرم‌ها به هیچ وجه با یکدیگر بحث و گفتگویی نداشته باشند.

ابزارهای بکار گرفته شده در این پژوهش دو پرسشنامه (شامل مشخصات ممتحن و واحد پژوهش، و ٤ فهرست وارسی (فرم ارزشیابی بالینی دانشجویان در بخش‌های زنان، زایشگاه، پره‌ناتال و بهداشت مادر و کودک) بود. پرسشنامه مشخصات ممتحن شامل سؤالاتی در زمینه سن، سابقه آموزشی در واحد بالینی امتحانی، سطح تحصیلات بود و پرسشنامه واحد پژوهش دارای سؤالاتی در زمینه سن، ترم تحصیلی، وضعیت تأهل، محل سکونت، و سال ورود به دانشگاه بود.

فرم ارزشیابی یا فهرست وارسی واحد زنان مرکب از ١٦ خرده مقیاس با عناوین ویژگی‌های فردی دانشجو، نحوه ارتباط با بیمار، اخذ شرح حال، معرفی بیمار، آمادگی جهت معاینه، مشاهده، معاینه ولو، اسپکولوم‌گذاری، بررسی واژن، سرویکس، رحم و آدنکس‌ها، اسمیربرداری، نسخه‌نویسی، تشخیص نهایی، آموزش به بیمار و پاسخ به سؤالات علمی با ۳۰ گویه بود.

فهرست وارسی واحد زایشگاه شامل ٦ خرده مقیاس (نحوه پذیرش مادر باردار، اداره مرحله اول، دوم، سوم، چهارم زایمان و پذیرش نوزاد) و ۳۷ گویه بود.

فرم ارزشیابی مراقبت‌های پره‌ناتال از ترکیب ۹ خرده مقیاس با عناوین چگونگی اخذشرح حال از بیمار، توجه و تفسیر آزمایش‌ها، سونوگرافی‌ها و ایمن‌سازی، معاینه و ارزیابی معمول مادر باردار، انجام مراقبت‌های ویژه، آموزش‌ها، توصیه‌ها و مراقبت بعد زایمان و پاسخ به سؤالات علمی با ٢٤ گویه تشکیل شده بود.

فرم ارزشیابی واحد بهداشت مادر و کودک شامل ٤ خرده مقیاس با عناوین مراقبت از کودک، بررسی

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

وضعیت رشد، واکسیناسیون، تنظیم خانواده و ۱۷ گویه بود. قابل ذکر است که درتمامی فرم‌ها ارزش هر گویه در مقابل آن درج شده بود (قبلاً بر اساس نظر کارشناسان امتیاز آن مهارت مشخص شده و در فرم‌ها ثبت بود) و هر دانشجو بر مبنای عملکرد خود از صفر تا صد در صد نمره گویه را بر اساس معیار لیکرت ۵ گانه به خود اختصاص می‌داد. بدین صورت که پس از علامت زدن نحوه عملکرد دانشجو در یکی از درجات بسیار ضعیف، ضعیف، متوسط، خوب و بسیار خوب به ترتیب صفر، ۵۰،۷۵، ۲۵ و ۱۰۰ درصد امتیاز آن گویه به عملکرد دانشجو تخصیص می‌یافت.

مجموع امتیازات کل در فرم زنان و زایشگاه به ترتیب ۸۰ و ۱۰۰ و در دو فرم دیگر هر کدام ٤٠ بود لذا برای محاسبه نمره دانشجو بر مبنای بیست، مجموع امتیاز کسب شده توسط دانشجو در حیطه زنان و زایشگاه به ترتیب بر ٤ و ٥ و در دو واحد دیگر بر ۲ تقسیم شد.

فرم‌های ارزشیابی (فهرست‌های وارسی) برای بررسی روایی محتوی در اختیار ۵ نفر از صاحب‌نظران این رشته گذاشته شد و از آنان خواسته شد مطابق اهداف ارزشیابی مندرج در راهنمای بالینی، در برابر هر سؤال موجود در ابزار سنجش یکی از گزینه‌های ضروری است، مفید است و نیازی نیست را انتخاب نمایند و بر اساس فرمول لاوشه(Lawshe)(۱۵) شاخص روایی محتوی (Content Validity Ratio: CVR) تعیین گشت.

$$CVR = \frac{ne - N/2}{N/2}$$

در این فرمول **ne** نشانگر تعداد سؤالاتی است که ضروری می‌باشند و N معرف تعداد کل سوالات است.

همچنین برای تعیین روایی صوری نیز ازآنان سؤال شد. برای تأیید روایی سازه از روش تمایز گروه‌های شناخته شده استفاده شد. بدین ترتیب که میانگین نمرات کسب شده دانشجویان شرکت نموده در امتحان (دانشجویان ترم ۸ و ٦) ارزیابی گردید. وجود تفاوت معنادارآماری بین میانگین نمرات کسب شده در هر حیطه و نمره کل آزمون بین دانشجویان با سطح بالاتر آموزشی (دانشجویان ترم ۸) و دانشجویان ترم پایین‌تر (ترم ٦) به عنوان شرط وجود روایی سازه ابزار تعیین گشت. برای تعیین پایایی هم‌ارز، از ضرایب همبستگی نمره دو ممتحن به طور هم زمان (با روش مشاهده مستقیم عملکرد) استفاده شد. به منظور تعیین پایایی همسانی درونی از روش آلفا کرونباخ استفاده گشت. از آن جا که آزمون پایان تحصیلی از ٤ خرده آزمون تشکیل شده است ضریب پایایی آلفای نمرات کل دانشجویان در امتحان جامع با نمرات دانشجویان در هر حیطه محاسبه گشت در تمامی موارد وجود شاخص روایی و پایایی بیشتر از ۰/۷ بالا، ۰/۷-۰/۵ متوسط ، و کمتر از ۰/۵ پایین تفسیر شد(۱و۱۵).

همچنین اطلاعات در برنامه SPSS-13 وارد و مورد تجزیه و تحلیل قرار گرفت. به کمک آمار توصیفی توزیع فراوانی، میانگین و انحراف معیار مشخصات واحدهای پژوهش تعیین شد و برای محاسبه شاخص‌های روایی از آزمون ضریب همبستگی پیرسون و برای بررسی اختلاف میانگین بین نمرات دانشجویان ترم‌های ٦ و ۸، به دلیل غیرنرمال بودن توزیع، از آزمون من‌ویتنی استفاده شد. آزمون کرونباخ نیز برای تعیین همبستگی بین نمرات خرده آزمون‌ها با نمره کل و بررسی پایایی در این مطالعه استفاده گشت.

نتایج

در این مطالعه ۷۰/۹ درصد (۲۲ نفر) شرکت‌کنندگان، دانشجویان ترم ۸ ورودی سال۸۵ دانشکده پرستاری و مامایی مشهد و ۲۹ درصد (۹ نفر) دانشجویان ترم ٦ داوطلب بودند. میانگین و انحراف معیار سن دانشجویان شرکت‌کننده ۰/٦±۲۲ سال (دامنه ۲۰ تا ۲۵ سال) محاسبه شد. ۵۳/۳ درصد متأهل و ٤٦/۷ درصد ساکن خوابگاه

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

بودند. متوسط مدت ارزیابی آنان در کل امتحان جامع ۰/۲±۳/۱ ساعت محاسبه شد.

هشتاد درصد از ممتحنین (۱۵ نفر) از مدرسین دانشکده که در برگزاری امتحان شرکت داشتند سن بالای ٤٠ سال و بالاتر داشتند و میانگین سابقه آموزشی آنان ۷/٤±۳/ ۱۲(دامنه ۲۳-۱) سال بود. میانگین سن ارزیابان ۱/۲±۲۵ سال با میانگین سابقه کارآموزشی ۰/٤±۲/۵ سال (۳-۱) بود. مدرک تحصیلی تمامی ممتحنین کارشناسی ارشد بود و ارزیابان دانشجویان ارشد مامایی با سابقه مربی‌گری بودند. اختلاف معنادارآماری بین سن و سابقه کار ممتحنین و ارزیابان وجود داشت (t=۲/۵۲، p=۰/۰۰۱).

پایایی کل مقیاس، با استفاده از آلفای کرونباخ ۰/۸ به دست آمد. جهت بررسی همسانی درونی از ضریب آلفای کرونباخ استفاده شد. این ضرایب برای بخش پره‌ناتال ۰/۷۷ و زایشگاه ۰/۸۳، در واحد زنان ۰/۵۰ و در حیطه بهداشت مادر و کودک ۰/٦۳ تعیین گشت.

با استفاده از ضرایب همبستگی نمرات ممتحنین (که یکی از آنان نسبت به سطح تحصیلی دانشجویان بی‌اطلاع بود) پایایی هم‌ارز نیز تعیین شد. ضرایب همبستگی ارزیابان در هر ٤ واحد امتحانی در حد مطلوبی وجود داشت. با بکارگیری معیار لیکرت بالاترین ضریب همبستگی پیرسون در بخش بهداشت مادر و کودک (p=۰/۰۰, r=۰/۸۲) و در واحدهای زایشگاه (p=۰/۰۷۵, r=۰/۰۰)، زنان (۰/۵۱ , p=۰/۰۳۱r=) و پره‌ناتال (p=۰/۰۰۱, r=۰/٦۰) محاسبه شد. با اعمال روش امتیازدهی سین سیناتی و محاسبه مجدد، ضرایب همبستگی پیرسون در سه واحد دیگر نیز بین ممتحنین زایشگاه،زنان و پره‌ناتال به ترتیب به (p=۰/۰۰۰, r=۰/۷٦) (p=۰/۰۰۰, r=۰/۷٤) و (r= ۰/۷۳ , p=۰/۰۰۱) ارتقا یافت، در بخش بهداشت ضریب تغییری نکرد. روایی صوری تمام فرم‌های ارزشیابی در بخش‌های مختلف (پره‌ناتال، زایشگاه، زنان و بهداشت مادر و کودک) توسط ۵ نفر از صاحب‌نظران گروه مامایی تأیید شد. روایی محتوای فرم‌های ارزیابی در امتحان جامع پایان دوره مامایی نیز با استفاده از قانون لاوشه در تمامی موارد بالا در حد ۰/۸ محاسبه شد به جز: در فرم ارزشیابی واحد زنان که تنها یک گزینه (شماره۱ -مرتبط با ویژگی‌های فردی) به دلیل پایین بودن ضریب روایی(۰/۲) حذف شد. فرم‌های ارزشیابی آزمون جامع پایان دوره. پس از اعمال اصلاحات موردنظر کارشناسان، تغییر یافت.

جدول۱: مقایسه میانگین نمرات حیطه‌های امتحانی درامتحان جامع پایان دوره بین دانشجویان ترم ٦ و۸

حیطه‌های امتحانی	میانگین نمره دانشجویان ترم ٦ (نمره کل ۲۰)	میانگین نمره دانشجویان ترم۸ (نمره کل ۲۰)	درصد تفاوت نمرات دانشجویان ترم ٦ و ۸	p	z
پره‌ناتال	۱٤/٤±۰/۹	۱۵±۱/۹	۰/۰٤	۰/۲٦۱	۰/۳۳
زایشگاه	۹±۲/۵	۱٤/۷±۱/۲	۰/٦۳	۰/۰۰۱	۳/۳۷
زنان	۱۲/۸±۲	۱۳/۷±۱/۸	۰/۰٦	۰/۲٦۱	۰/٤۱
بهداشت مادروکودک	۹/۵±۱/۲	۱٤/٦±۲	۰/۵۳	۰/۰۹۱	۲/۷۷
نمره کل فینال	٤۵/۷±۲/۱	۵۸±۱/۸	۰/۲٦	۰/۰۱۰	۳/۳۷

در این مطالعه تفاوت معنادار آماری بین نمره کل دو گروه شرکت‌کننده و در تمام موارد ارزشیابی در حیطه‌های زایشگاه و بهداشت مادر و کودک امتحان جامع پایان دوره وجود داشت (جدول ۱). اما در سایر حیطه‌ها

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

این گونه نبود، به طوری که در واحد پره‌ناتال، اختلاف آماری معنادار بین دو گروه تنها در مهارت انجام معاینه بیماران وجود داشت (p=۰/۰۳۱, z=۲/۸۲).

همچنین علی‌رغم بالاتر بودن میانگین نمرات دانشجویان ترم ۸ نسبت به دانشجویان ترم ششم (۱/۸±۱۳/۷ در برابر۲±۱۲/۸) در واحد زنان، بین میانگین نمرات کل دو گروه شرکت‌کننده اختلاف آماری معنادار مشاهده نشد. البته دانشجویان ترم ۸ کارشناسی مامایی مهارت بیشتری جهت آمادگی بیماران برای معاینه و بررسی واژن و رحم نسبت به دانشجویان ترم ٦ داشتند.

بحث

این مطالعه با هدف بررسی اعتبار سازه و پایایی ابزارهای مورد استفاده در امتحان جامع پایان دوره مامایی صورت گرفت. در مطالعه حاضر برای تأیید روایی سازه، از اختلاف میانگین امتیازات عملکردی بین دانشجویان سال‌های سوم و چهارم مامایی استفاده شد. وجود تفاوت معنادار آماری بین میانگین نمره کل کسب شده توسط دانشجویان ترم‌های ٦ و ٨ مامایی به میزان ٢٦ درصد در امتحان جامع پایان دوره تحصیلی روایی سازه ابزارهای ارزشیابی دانشجویان مامایی را تأیید نمود. نتایج این بخش از یافته‌های ما از نظر تفاوت بین نمره کل کسب شده با سایر تحقیقات انجام شده در این باره همخوانی دارد.

تأیید روایی سازه به روش فوق در مطالعات ون بلریکوم، فردلیچ انجام شده است. ون بلریکوم و همکاران، اختلاف میانگین امتیازات کسب شده دو گروه رزیدنت سال اول ودوم (٢٤ رزیدنت سال اول از دانشگاه واشنگتن و٢٤ رزیدنت سال دوم از دانشگاه مادیگان ارتش) را، در ایستگاه‌های امتحانی OSCE محاسبه نمودند و به دلیل وجود تفاوت معنادار آماری بین مجموع امتیازات کسب شده (به هر دوروش چک‌لیست و رتبه‌بندی) رزیدنت‌های سال‌های اول و دوم تحصیلی روایی‌سازه ابزار ارزیابی را اثبات کردند(١٦).

محققان دیگری نیز در بررسی مهارت‌های جراحی ٣١ دستیار سال سوم و چهارم جراحی در تورنتو به روش OSCE، با دو ابزار چک‌لیست و رتبه‌بندی لیکرت، روایی سازه ابزارهای ارزیابی را تأیید کردند. وجود اختلاف معنادار آماری بین میانگین نمرات رزیدنت‌های سال سوم و چهارم به میزان ٢٥ درصد مشاهده شد(١٧).

همچنین ویفت و همکاران برای اثبات روایی سازه آزمون مهارت‌های تکنیکی رزیدنت‌های زنان سال‌های اول تا چهارم، در ١٠ ایستگاه امتحانی Observed Structured Assessment of Ttechnical Skills (OSATS) از مقایسه اختلاف میانگین امتیازات کسب شده، استفاده نمودند(١٨).

در این مطالعه تفاوت معنادار آماری بین نمره کل دوگروه شرکت‌کننده و در تمام موارد ارزشیابی در حیطه‌های زایشگاه و بهداشت مادر و کودک امتحان جامع پایان دوره وجود داشت اما عدم وجود تفاوت معنادار آماری در دو حیطه زنان و پره‌ناتال مشاهده شد. از آن جا که در مطالعه حاضر، دانشجویان ترم ٨ واحدهای بالینی مذکور را قبل از ورود به عرصه گذرانیده بودند، عدم ایجاد فرصت برای تمرین موقعیت‌های فوق توسط دانشجویان در ترم‌های بالاتر می‌تواند تداوم یادگیری را مختل کرده باشد. گوف و همکاران جهت ارزیابی دو روش آموزشی هیسترسکوپی با مدل استاندارد و روش شبیه‌سازی در هفت ایستگاه OSATS، میانگین امتیازات متخصصین و رزیدنت‌های زنان را با یکدیگر مقایسه نمودند که دامنه اختلاف امتیازات دو گروه افزایش داشت(١٩). تفاوت نتایج این تحقیق با تحقیق گوف می‌تواند به دلیل حجم کم نمونه و یا تفاوت گروه‌های هدف از نظر فراغت تحصیل باشد. بنابراین توصیه می‌شود که این مطالعه در سایر دانشکده‌های مامایی تکرار شود.

نتایج بررسی پایایی آزمون پایان دوره تحصیلی

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

دانشجویان مامایی دانشکده مشهد نشان داد که به دلیل بالابودن ضریب آلفا، همسانی درونی آزمون در حد مطلوبی قرار دارد. حقایق و همکاران نیز با استفاده روش آلفا کرونباخ پایایی مقیاس خودکارآمدی مدیریت بیماران دیابتی تیپ را تعیین کردند. ضریب همبستگی این مقیاس در ٤ بعد رژیم غذایی، فعالیت بدنی، مصرف دارو و آزمون قند خون به ترتیب ۰/۵۳، ۰/۸۷، ۰/۸٦ و ۰/۱۹ اعلام شد در این مطالعه ضریب همبستگی بین نمره کل با خرده مقیاس‌ها بیش از همبستگی بین خرده مقیاس‌ها با یکدیگر بود (بالاتر از ۰/۵). بنابراین پژوهشگران وجود ساختار و هماهنگی خوب در پرسشنامه را تأیید کردند(۲۰).

معطری و همکاران بالاترین ضریب همسانی در آزمون OSCE دانشجویان پرستاری را ۰/۷۷ اعلام داشتند و گزارش نمودند به دلیل ضرایب بالای همبستگی بین نمرات OSCE و نمرات ایستگاه‌ها (بجز یک ایستگاه) همه این ایستگاه‌ها یک حوزه رفتاری را اندازه‌گیری می‌کنند. تأیید روایی محتوای این آزمون پس از مطالعه کتب معتبر علمی و با استفاده از نظرات کارشناسان و سرپرستاران منتخب بیمارستان‌ها، و بدون ذکر شاخص روایی محتوا بود(۳). در مطالعه ما شاخص روایی محتوا مطلوب ارزیابی شد. تعیین شاخص روایی محتوا از نقاط قوت هر مطالعه می‌تواند محسوب شود، به طوری که پن و همکاران برای تأیید روایی و پایایی مقیاس کوتاه و چند بعدی افسردگی (MFSI-SF-C) پس از ترجمه و کاربرد آن برای ۱۰۷ بیمار مراجعه‌کننده به دو مرکز بیمارستانی تایوان، شاخص‌های روایی محتوا و ضریب آلفا را تعیین نمودند و به دلیل وجود ضریب آلفای بالا (۰/۸۳) و شاخص محتوای مطلوب (۰/۹۳) مقیاس را دارای روایی و پایایی دانستند(۲۱).

جهت بررسی پایایی هم‌ارز در مطالعه حاضر اکثریت (۷۵ درصد) ارزیابان گزارش کردند با استفاده از تعاریف دانشگاه سین سیناتی (نمرات ٤-۰) به جای مقیاس لیکرت مندرج در فرم‌های ارزشیابی توانایی بهتری برای تخصیص نمرات دانشجویان پس از انجام مهارت می‌یابند. بنابراین از تمامی ممتحنین و ارزیابان خواسته شد که بر مبنای سیستم ارزیابی سین سیناتی عمل نمایند. نتایج حاکی از افزایش پایایی هم‌ارز بود به طوری که در تمام حیطه‌های امتحانی، ضرایب بیشتر از ۰/۷ گزارش شد. گوف نیز در مطالعه خود ضریب پیرسون را با دو روش چک‌لیست و رتبه‌بندی بررسی نمود و نتیجه گرفت که با اجرای سیستم درجه‌بندی لیکرت ضریب همبستگی ارزیابان افزایش می‌یابد(۱۹). یکی از مشکلات درخصوص روش درجه‌بندی، دشواری در به دست آوردن خوب ارزش‌ها دانسته شده است(۲۲). منطقی است که بپذیریم هرچه تعاریف عینی‌تر باشد ارزش گذاری و قضاوت درباره آن آسانتر می‌شود و این از ویژگی‌های سیستم امتیازبندی سین سیناتی است.

لازم به ذکر است که همانند سایر پژوهش‌ها، محدودیت‌هایی دراین مطالعه وجود دارد از جمله میزان استرس درک شده توسط واحدهای پژوهش، تعداد کم دانشجویان مامایی به ویژه در ترم ٦، و نیز حضور تنها دانشجویان داوطلب در مطالعه، که به جهت رعایت مسائل اخلاقی ناگزیر از انجام آن بودیم؛ هرچند که آزمون تی بین میانگین معدل بالینی دانشجویان داوطلب و غیر داوطلب تفاوت معناداری را نشان نداد. اما پیشنهاد می‌شود اعتبار سازه این ابزار در سایر دانشکده‌ها و با حجم بالاتر بررسی شود. از طرفی روایی سازه در این ابزار تنها با استفاده از تمایز گروه‌ها انجام شد؛ توصیه می‌شود روش‌های عملی دیگر برای تأیید روایی سازه ابزارهای ارزشیابی بالینی مورد بررسی قرارگیرد همان طور که مولیناز و همکاران در بررسی روایی سازه مدل جدید لاپاراسکوپی از دو روش ترسیم منحنی یادگیری و نیز امتیازات کسب شده در ایستگاه‌های لاپاراسکوپی استفاده کردند(۲۳).

در نهایت ارائه روش‌های آسان و عملی روایی سازه می‌تواند با تسهیل روند ارزشیابی، امکان بررسی ابزارهای ارزشیابی را افزایش دهد. همچنین گنجاندن

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

تعاریف عینی‌تر مانند تعاریف سیستم سین سیناتی در سیستم نمره دهی میتواند در افزایش پایایی نمرات موثر واقع گردد، لذا استفاده از آن در سایر دانشکده‌ها برای ارزیابی مهارت‌های بالینی پیشنهاد می شود.

نتیجه‌گیری

در مجموع شواهد حاکی از وجود روایی سازه و کفایت ابزارهای ارزیابی بالینی دانشجویان مامایی در مقطع کارشناسی پیوسته بود.

قدردانی

نویسندگان بر خود لازم می‌دانند تا از معاونت پژوهشی دانشگاه علوم پزشکی مشهد که هزینه مالی این طرح تحقیقاتی را تقبل نموده‌اند و همچنین شورای پژوهشی دانشکده پرستاری و مامایی، کلیه مربیان و دانشجویانی که در انجام این طرح همکاری نموده‌اند سپاس‌گزاری نمایند.

منابع

1. Safe.A. [Andazegiri Sanjesh va arzeshyabi amozeshi]..4th ed.Tehran.Nashre Doran Publisher. 2006.[Persian]
2. Wolf .MR.[Arzeshyabi.Amozeshi :mabani sanjesh tavanai va baresi barnameh].tarjomeh Kyamnsh. Alireza.Tehran University Publication Centre.2003 [Persian]
3-Moatari M, Abdullah zargar SH, Mousavinasab M,Zare N, BeygyMarvdast P. [Barasi ravai & payai azmoon osce dar arzyabiye maharathaye balini daneshjouyane sal 4 parastari daneshkadeh parastari & mamaei Hazrate Fatemeh Shiraz]. Iran journal research in medicine 2008; 55-59. [Persian]
4. Bahreini Toosi MH, Modabber Azizi MJ, Kaveh Tabatbaie MS, Ebrahimzadeh Saeid, Bahreini Toosi Vahideh, Bahreini Toosi Kaveh, Medical Students' Viewpoints about the Evaluation Methods at Internship Stage (Mashad University 2001) Iranian Journal of Medical Education, 2002; 2 (0) :19-19. [Persian]
5. Wanda D. An investigation of clinical Assessment processes of student nurses in Jakarta Indonesia [dissertation]. Australian Catholic university. 2007.
6. Harvey SA, Blandón YC, McCaw-Binns A, Sandino I, Urbina L, Rodríguez C, & et al. Are skilled birth attendants really skilled? A measurement method, some disturbing results and a potential way forward. Bull World Health Organ. 2007 Oct;85(10):783-90.
7. Vesarat Behdasht va Darman Va Amozesh pezeshki. [Pishgiri as marg 420 madar 1390]. [Cited 2011 Oct 16]. Available from: http://WWW.pezeshkan.org/p=23316 [Persian]
8. Vesarat Behdasht, Darman va Amozesh pezeshki. [Rahnmay keshvari eraeh khadamat mamai va zayeman dar bimarestanhaye dostdar madar].dafter salamat khanevadeh va jamiat ,Edare . salamat .Tehran..2006. [Persian]
9. Casey PM, Goepfert AR, Espey EL, Hammoud MM, Kaczmarczyk JM, Katz NT & et al. To the point: reviews in medical education--the Objective Structured Clinical Examination. Am J Obstet Gynecol. 2009 Jan;200(1):25-34.
10. Jalili M, Mirzazadeh A, Azarpira A. A survey of medical students' perceptions of the quality of their medical education upon graduation. Ann Acad Med Singapore. 2008 Dec; 37(12): 1012-8
11. Khoinejad Gh, [Ravesh Tahghigh dar Oloume Tarbiati] 1th ed. Tehran .Meher Publisher; 2002. [Persian]
12. Bisson DI, Hyde J, Mearrs JE. Assessing practical skills in obstetrics and gynecology: Education issues and practical implications. Obstrician and Gynecologist 2006; 8(2): 107-112
13. Lodico MG, Spulding DT, Vogegtle KH. Methods in educational research:From theory to practice. 2th ed.San Francisco:Jon willy &sons. 2010
14. The University of Cincinnati College of Nursing. Clinical Skills Checklist-NMW II [citd2008Sep10]. Available from: http://nursingonline.uc.edu/docs/Clinical_Skills_Checklist_NMW_II.pdf
15. Haynes N, Richard D, Kubany E . Content validity in psychological assessment: A functional Approach to Concepts and Methods. American Psychological Assessment. 1995; 7(3): 238-247

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

16. VanBlaricom AL, Goff BA, Chinn M, Icasiano MM, Nielsen P, Mandel L. A new curriculum for hysteroscopy training as demonstrated by an objective structured assessment of technical skills (OSATS). Am J Obstet Gynecol. 2005 Nov;193(5):1856-65.
17. Friedlich M, Wood T, Regehr G, Hurst C, Shamji F. Structured assessment of minor surgical skills (SAMSS) for clinical clerks. Acad Med. 2002 Oct;77(10 Suppl):S39-41.
18. Swift SE, Carter JF. Institution and validation of an observed structured assessment of technical skills (OSATS) for obstetrics and gynecology residents and faculty. Am J Obstet Gynecol. 2006 Aug;195(2):617-21; discussion 621-3.
19. Goff BA, VanBlaricom A, Mandel L, Chinn M, Nielsen P. Comparison of objective, structured assessment of technical skills with a virtual reality hysteroscopy trainer and standard latex hysteroscopy model. J Reprod Med. 2007 May;52(5):407-12.
20. Haghayegh SA,Ghasemi N, Neshatdoost H T, Kajbaf M B, Khanbany M.[Ravai and payaie meghyas khodkaramadi modiriat bimaran diabet type2]. Iranian Journal of Endocrinology and Metabolism .2010; 12(2): 111-115 [Persian]
21. Pien LC, Chu H, Chen WC, Chang YS, Liao YM, Chen CH, Chou KR. Reliability and validity of a Chinese version of the Multidimensional Fatigue Symptom Inventory-Short Form (MFSI-SF-C). J Clin Nurs. 2011 Aug;20(15-16):2224-32. doi: 10.1111/j.1365-2702.2010.03691.x. Epub 2011 May 25.
22. Burns N. Grove Sk. The practice of Nursing Research. Appraisal ,Synthesis and Generation of Evidence.6th ed .USA: Saunders. 2009
23. Molinas CR, De Win G, Ritter O, Keckstein J, Miserez M, Campo R. Feasibility and construct validity of a novel laparoscopic skills testing and training model. Gynecol Surg (2008) 5:281–290.

[Downloaded from ijme.mui.ac.ir on 2026-11-12]

Construct Validity and Reliablity of Evaluation Tools for Midwifery Students at the Final Examination in Mashhad School of Nursing and Midwifery

Nahid Jahani Shoorab[1], Nahid Golmakani[2], Seyed Reza Mazluom[3], Kobra Mirzakhani[4], Sedighe Azhari[5], Afsaneh Navaiyan[6]

Abstract

Introduction: *Inappropriate performance of delivery care team is one of the most important causes of maternal mortality. Lack of objective tools, reduces the opportunity for a systematic valid evaluation in clinical examinations. Therefore, this study investigated the reliability and validity of assessment tools in comprehensive final examination of midwifery students at the end of their education (at the time of graduation).*

Methods: *In this correlational study, founded on existing evaluation tools for final comprehensive exam in Mashhad School of Nursing and Midwifery, students' skills in 4 areas of examination (delivery room, gynecological unit, pregnancy care, and maternal and child's health) were assessed (July 2010). Thirty one students (22 students of 8th and 9 students of 6th semester of midwifery BSc program) were evaluated within 5 consecutive days by 2 examiners. Content validity was determined by field experts. Mean scores achieved in each area of examination were compared in order to appraise construct validity as well as total mean score of students of senior and junior semesters (8 and 6). Alpha Cronbach coefficient was used for reliability analysis.*

Results: *A significant difference of 26 percent between total mean score of students studying at 8th semester with that of students studying at 6th semester in comprehensive final examination, proved the construct validity in evaluation forms of midwifery students based on Royal College of Gynecology and Midwifery Definition in this examination (p=0.01). There was an average to high Cronbach's Alpha coefficient (0.83 - 0.51) in each area of examination which reveals an internal consistency within the evaluation tools. Content validity was also high (0.8). The lowest and highest reliability of evaluators determined to be 0.82% and 73.5% respectively, according to definitions characterized by Cincinnati University.*

Conclusion: *The tools used for evaluating midwifery students at the final examination in Mashhad School of Nursing and Midwifery met reliability and construct validity criteria and are recommended to be used for evaluationg midwifery students for graduation.*

Keywords: Validity, reliability, evaluation tools, evaluation, undergraduate, clinical skills.

Addresses:
[1] (✉) Instructor, Department of Midwifery, School of Nursing and Midwifery, Mashhad University of Medical Sciences. Mashhad, Iran. E-mail: jahanishn@mums.ac.ir
[2] Instructor, Department of Midwifery, School of Nursing and Midwifery, Mashhad University of Medical Sciences. Mashhad, Iran. E-mail: golmakanin@mums.ac.ir
[3] Assistant Professor, Nursing, School of Nursing & Midwifery, Mashhad University of Medical Sciences, Mashhad, Iran. E-mail: mazlomr@mums.ac.ir
[4, 5 & 6] Staff, School of Nursing and Midwifery, Mashhad University of Medical Sciences. Mashhad, Iran. E-mails: mirzakhanik@mums.ac.ir, azharis@mums.ac.ir, navaiyana@mums.ac.ir

[Downloaded from ijme.mui.ac.ir on 2026-11-12]