

# عوامل مؤثر بر نمره‌دهی آزمون بالینی کوتاه (Mini-CEX): مطالعه مروری

رقیه گندم‌کار، محمد جلیلی\*

## چکیده

**مقدمه:** در سال‌های اخیر استفاده از آزمون Mini-CEX به دلیل قابلیت اجرای آسان آن گسترش یافته است. علی‌رغم حمایت مطالعات اولیه از پایایی و روایی آزمون مذکور، سودمندی آن به دلیل عدم ثبات در قضاوت ارزیابان در نمره‌دهی آن مورد بحث جدی است. هدف از مطالعه حاضر مرور متون مرتبط با سودمندی آزمون Mini-CEX به منظور جمع‌بندی و ارائه عوامل مؤثر بر نمره‌دهی آن است.

**روش‌ها:** در این مطالعه مروری narrative برای دستیابی به مقالات مرتبط کلید واژه‌های Mini-CEX، validity، utility، reliability، score variability و کلید واژه‌های معادل فارسی آن در پایگاه‌های اطلاعاتی MEDLINE، SCOPUS، EMBASE، SID و Magiran جستجو شد. نتایج بازبازی شده به فاصله زمانی سال ۱۹۹۵ تا ۲۰۱۴ میلادی محدود شد. در نهایت نتایج ۱۹ مطالعه مرتبط با موضوع مقاله حاضر از بین ۵۱ مقاله مورد بررسی قرار گرفت.

**نتایج:** بر اساس نتایج مطالعات عوامل تأثیرگذار بر نمره‌دهی آزمون Mini-CEX به ترتیب اهمیت شامل عوامل مربوط به ارزیابان، عوامل مربوط به فرم‌های آزمون و مواجهه بالینی بود. تفاوت در میزان سخت‌گیری ارزیابان در نمره‌دهی و ارفاق و نیز ارتباط بالای بین سؤالات از مهم‌ترین عوامل تأثیرگذار بر نمره‌دهی آزمون بود. نتایج پژوهش‌ها در خصوص نحوه تأثیر برخی از عوامل مانند توانمندسازی ارزیابان بر نمره‌دهی آزمون Mini-CEX قطعی نبود.

**نتیجه‌گیری:** با در نظر گرفتن عدم ثبات در نمرات آزمون Mini-CEX و مشخص نبودن شرایط تأثیرگذار بر نمره‌دهی آن، توصیه می‌شود نمره آزمون به تنهایی مبنای ارزیابی پایانی به ویژه آزمون‌های پایانی هنجار محور به منظور رتبه‌بندی عملکرد دانشجویان قرار نگیرد.

واژه‌های کلیدی: آزمون بالینی کوتاه، سودمندی، روایی، ثبات نمره‌دهی، پایایی

مجله ایرانی آموزش در علوم پزشکی / دی ۱۳۹۳؛ ۱۴(۱۰): ۸۶۰ تا ۸۶۸

## مقدمه

معمول یکی از اعضای هیأت‌علمی بالینی عملکرد فراگیر را در مواجهه با بیمار و به مدت ۱۵ تا ۲۰ دقیقه ارزیابی می‌کند. هدف اصلی در این روش ارائه بازخورد بر اساس عملکرد مشاهده شده است. هر مواجهه با بیمار ۱۵ دقیقه طول می‌کشد و ۵ تا ۱۰ دقیقه نیز به ارائه بازخورد به فراگیر اختصاص می‌یابد (۱). در آزمون Mini-CEX انتظار می‌رود عملکرد فراگیر در طول یک سال، در چند مواجهه، در محیط‌های بالینی مختلف، توسط ارزیابان متفاوت و در برخورد با طیف وسیعی از مشکلات بالینی ارزیابی شود. توانمندی‌هایی که به طور معمول در

آزمون Mini-CEX یکی از انواع ارزیابی‌های مبتنی بر محل کار است که برای اولین بار در سال ۱۹۹۵ میلادی توسط بورد داخلی آمریکا در پاسخ به محدودیت‌های آزمون‌های شفاهی مطرح شد. در این آزمون به طور

\* نویسنده مسؤول: دکتر محمد جلیلی (دانشیار)، گروه طب اورژانس، گروه آموزش پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی تهران، mjalili@tums.ac.ir  
دکتر رقیه گندم‌کار (پزشک عمومی)، کارشناس ارشد آموزش پزشکی، دانشجوی دکتری آموزش پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی تهران، ایران.  
(rgandomkar@razi.tums.ac.ir)

تاریخ دریافت مقاله: ۹۳/۳/۳۱، تاریخ اصلاحیه: ۹۳/۷/۳۰، تاریخ پذیرش: ۹۳/۹/۱۶

Mini-CEX با دیگر آزمون‌های برگزار شده در محیط بالینی شامل آزمون شفاهی ساختارمند ( $r=0/73$ )، آزمون بالینی در کنار بستر بیمار ( $r=0/67$ )، ارزیابی درون‌بخش دوره کارآموزی ( $r=0/43$ ,  $p=0/001$ ) و ارزیابی در درمانگاه دوره کارآموزی ( $r=0/35$ ,  $p=0/001$ ) و در نتیجه روایی هم‌زمان این روش ارزیابی حمایت می‌کند (۹ تا ۸). علاوه بر این، آزمون Mini-CEX با تمایز قائل شدن بین سطوح متفاوت عملکرد فراگیران در سال‌های مختلف تحصیل و فراگیران سطوح مختلف عملکرد (ضعیف، مرزی و عالی) و همچنین افزایش نمرات آزمون فراگیران در طول سال تحصیلی از روایی سازه معقولی برخوردار است (۱۰ تا ۱۳).

با وجود مطالعات حمایت‌کننده از پایایی و روایی آزمون Mini-CEX، سودمندی (utility) آن در بحث روایی به دلیل عدم ثبات در قضاوت ارزیابان مورد انتقاد است (۱۴ تا ۱۶). در مطالعه ولر (Weller) و همکاران، سخت‌گیری (stringency) ارزیاب در نمره‌دهی و ذهنی بودن قضاوت ارزیابان به ترتیب ۴۰ و ۱۵ درصد منابع تغییرپذیری نمرات آزمون را به خود اختصاص دادند (۱۷). مطالعات انجام شده در شرایط و محیط‌های دیگر نیز یافته‌های این مطالعه را تأیید می‌کند (۱۸ تا ۱۹). با وجود استفاده از فرم‌های ساختارمند در آزمون Mini-CEX و آموزش ارزیابان، هنوز تغییرپذیری نمره‌دهی به عنوان یکی از مهم‌ترین عوامل مؤثر بر روایی این آزمون مورد بحث و تردید است. مطالعات انجام شده در زمینه سودمندی آزمون Mini-CEX به بررسی منابع تغییرپذیری نمرات آزمون پرداخته‌اند، ولی هیچ یک به بررسی عوامل مؤثر بر نمره‌دهی آن به صورت جامع نپرداخته‌اند.

با توجه به به‌کارگیری رو به رشد آزمون Mini-CEX در محیط‌های بالینی در دانشگاه‌های علوم پزشکی کشور، آگاهی از عوامل دخیل در نمره‌دهی آزمون، می‌تواند به کنترل منابع خطا و در نتیجه برگزاری آزمون‌های معتبر

Mini-CEX مورد سنجش قرار می‌گیرند شامل مصاحبه، معاینه بالینی، تعهد حرفه‌ای، قضاوت بالینی، مشاوره، سازمان‌دهی و کارآیی و توانمندی کلی است. در انتها پس از مشاهده عملکرد فراگیر با استفاده از مقیاس درجه‌بندی (معمولاً مقیاس لیکرت با درجه‌بندی ۹ تایی یا ۶ تایی) هر یک از توانمندی‌ها نمره‌گذاری می‌شود (۲).

هر چند آزمون Mini-CEX اولین بار به منظور ارزیابی مهارت‌های بالینی در دوره دستیاری داخلی طراحی شد، ولی قابلیت اجرای آسان آن نسبت به دیگر روش‌های ارزیابی در محیط واقعی موجب گسترش استفاده از آن در دیگر برنامه‌های دستیاری، پزشکی عمومی و دیگر رشته‌های علوم پزشکی از جمله پرستاری و مامایی شده است (۳ تا ۴). علاوه بر این، با وجود این که هدف ذاتی آزمون مذکور ارائه بازخورد و اصلاح توانمندی‌های بالینی فراگیران است ولی گزارشات متعددی در خصوص استفاده از آن با هدف پایانی و تصمیم‌گیری‌های نهایی در مورد عملکرد فراگیران وجود دارد (۵). بنابراین، اطمینان از پایایی و روایی این آزمون به منظور انجام یک ارزیابی دقیق و عادلانه از توانمندی‌های بالینی اهمیت بسیار دارد.

در آزمون Mini-CEX سعی شده است با افزایش تعداد مشاهدات (مواجهه با بیمار)، استفاده از ارزیابان مختلف و استفاده از فرم‌های ساختارمند با تعداد سؤالات مناسب پایایی و روایی آزمون بهبود یابد. در این راستا، مطالعات اولیه حاکی از پایایی و روایی مناسب این آزمون است. به طوری که در ارزیابی دستیاران برای ۸ و ۱۴ مواجهه به ترتیب ضرایب تعمیم‌پذیری ۰/۷۱ و ۰/۸۱ و در ارزیابی دانش‌آموختگان پزشکی بین‌المللی در استرالیا برای ۸ مواجهه ضریب تعمیم‌پذیری ۰/۸۸ گزارش شده است (۶ تا ۷). در مطالعه دیگری در مقطع دستیاری، برای ۸ مواجهه ضریب تعمیم‌پذیری ۰/۷۷ به دست آمد که نشان‌دهنده پایایی مناسب این آزمون است (۸).

مطالعات مختلف از ارتباط قابل ملاحظه نمرات آزمون

مقایسه‌ها بود. بر اساس نتایج مطالعات بررسی شده عوامل تأثیرگذار بر نمره‌دهی آزمون Mini-CEX شامل عوامل مربوط به ارزیابان، فرم‌های نمره‌دهی و مواجهه بالینی مورد بررسی است. مهم‌ترین روش‌های مورد استفاده در مقالات به منظور تحلیل نتایج آزمون Mini-CEX شامل تئوری تعمیم‌پذیری (G theory)، آلفای کرونباخ، تحلیل عاملی و ضرایب همبستگی بود. در ۱۰ مطالعه از تئوری تعمیم‌پذیری استفاده شده بود و سهم منابع خطا مانند عوامل مربوط به ارزیابان، مواجهه بر دقت نمرات آزمون Mini-CEX تعیین شده بود. در مقالاتی که به تأثیر ساختار فرم‌های مورد استفاده در آزمون بر نمرات آن پرداختند از آلفای کرونباخ و تحلیل عاملی استفاده شده بود که در ۵ مورد از آلفای کرونباخ و در یک مورد از تحلیل عاملی استفاده شده بود. علاوه بر این، مطالعاتی (۵ مورد) که به تعیین ارتباط عوامل مربوط به مواجهه با نمرات آزمون پرداخته بودند ضرایب همبستگی مانند پیرسون، اسپیرمن را با توجه به جنس متغیرها مورد استفاده قرار داده بودند. در ۵ مطالعه که به مقایسه تأثیر آموزش ارزیابان بر قضاوت آن‌ها پرداخته بودند از آزمون‌های آماری مقایسه میانگین بین گروه‌ها استفاده شده بود. لازم به ذکر است در برخی از مقالات از چند روش به طور هم‌زمان استفاده شده بود. در ادامه عوامل تأثیرگذار بر نمره‌دهی آزمون Mini-CEX و نحوه تأثیر آن به تفصیل ارائه شده است.

#### ۱. عوامل مربوط به ارزیابان

یکی از مشکلات آزمون Mini-CEX ذهنی بودن قضاوت ارزیابان است. به این معنی که نظر یک ارزیاب در خصوص داوطلبان مختلف با عملکرد یکسان و نیز در مورد یک داوطلب یکسان در موقعیت‌های مختلف می‌تواند متغیر باشد. به نظر می‌رسد عوامل زیر از طرف ارزیابان بر خطا در نمره‌دهی آزمون مؤثر باشد:

- ویژگی‌های ارزیاب: یکی از ویژگی‌های ارزیابان که بر

به ویژه با اهداف پایانی کمک کند. هدف از مطالعه حاضر مرور متون مرتبط با سودمندی آزمون Mini-CEX به منظور جمع‌بندی و ارائه عوامل مؤثر بر نمره‌دهی آزمون مذکور است.

#### روش‌ها

این مطالعه از نوع مروری narrative است. جهت دستیابی به مقالات مرتبط کلید واژه‌های انگلیسی Mini-CEX ، reliability , validity , utility , score , variability و کلید واژه‌های معادل فارسی آن در پایگاه‌های اطلاعاتی MEDLINE, EMBASE, Magiran و SID SCOPUS جستجو شد. نتایج بازیابی شده به فاصله زمانی سال ۱۹۹۵ (زمان انتشار اولین مقاله در مورد آزمون Mini-CEX) تا ماه می (May 2014) محدود شد. بخش منابع مقالات به دست آمده نیز به منظور جستجوی مقالات بیشتر بررسی شد. بر اساس نتایج جستجو ۵۱ مقاله به دست آمد که ۴۹ مقاله به زبان انگلیسی و ۲ مقاله به زبان فارسی بود. در بررسی چکیده آن‌ها ۱۹ مقاله به طور مستقیم با موضوع مقاله حاضر مرتبط بود و به بررسی عوامل مؤثر بر نمره‌دهی آزمون Mini-CEX پرداخته بود و انتخاب شد. مقالاتی که تنها به بیان تجربه اجرای آزمون مذکور در یک مؤسسه پرداخته بودند، مقالاتی که در نمره‌دهی آزمون به جای فرم نمره‌دهی کلی از چک‌لیست استفاده کرده بودند یا به بررسی پیامدهای آزمون Mini-CEX پرداخته بودند، حذف شدند. در مواردی که اصل مقالات از کتابخانه دیجیتال دانشگاه علوم پزشکی تهران قابل بازیابی نبود، مقاله خریداری شد.

#### نتایج

متن کامل ۱۹ مقاله انتخاب شده مطالعه شد. کل مقالات انتخاب شده به زبان انگلیسی بودند. ۱۷ مقاله پژوهشی اصیل، یک مقاله مروری (narrative) و یک مقاله هم

تا ۷ در فرم‌های نمره‌گذاری ۹ تایی بودند. علاوه بر این، ارزیابان بسیاری از موارد با عملکرد نامطلوب را در طیف فراگیران با عملکرد مطلوب قرار می‌دهند و از مردود اعلام کردن داوطلبان به ویژه دستیاران اجتناب می‌ورزند. در حمایت از مسأله ارفاق در نمره‌دهی آزمون، پژوهش‌ها نشان داده‌اند، اگر برگزاری آزمون Mini-CEX بخشی از یک پروژه تحقیقاتی و نه با هدف ارزیابی باشد یا زمانی که ارزیابان به مواجهه‌های ضبط شده قبلی نمره می‌دهند، نمرات پایین‌تر و عموماً دقیق‌تری نسبت به زمانی که این آزمون در محیط واقعی برگزار می‌شود، می‌دهند (۱۹).

• آموزش ارزیابان: هر چند با توجه به نقش مهم تفاوت سخت‌گیری ارزیابان در نمره‌دهی، منطقی به نظر می‌رسد که آموزش این افراد موجب ثبات در قضاوت ارزیابان شود، نتایج پژوهش‌های انجام شده با هدف تأثیر آموزش ارزیابان بر ثبات قضاوت ارزیابان و دقت نمره‌دهی نشان داد، این آموزش‌ها تأثیر معناداری بر ثبات بین قضاوت ارزیابان نداشت. با حضور در برنامه‌های آموزشی، تنها اعتماد به نفس شرکت‌کنندگان (اساتید) در اجرای آزمون بهبود یافت، آن‌ها در مشاهده مستقیم عملکرد فراگیران احساس راحتی بیشتری داشتند و سخت‌گیرتر از همکاران گروه کنترل قضاوت کردند. همچنین برگزاری کارگاه آموزشی بر دانش ارزیابان در مورد Mini-CEX و افزایش احتمال ارائه بازخورد به داوطلبان مؤثر بود (۲۴ تا ۲۲).

۲. عوامل مربوط به فرم‌های آزمون:

یکی از مشکلات نمره‌دهی آزمون Mini-CEX ارتباط بالای بین سؤالات (آیتم‌ها) است. مطالعات مختلف سطح بالایی از ثبات درونی را با محاسبه ضریب آلفای کرونباخ برای فرم‌های نمره‌دهی آزمون مذکور گزارش کردند (۱۸ و ۱۷ تا ۲۹).

در مطالعه ولر، آلفای کرونباخ ۰/۹۵ برای فرم‌های Mini-

نتایج آزمون تأثیرگذار است موقعیت شغلی آن‌ها از نظر هیأت‌علمی یا دستیار بودن است. دیده شده است که استادان در آزمون Mini-CEX سخت‌گیرانه‌تر (more stringently) از دستیاران به دانشجویان پزشکی نمره می‌دهند (به ترتیب: ۴/۳۵ و ۴/۵۰ در فرم‌های با لیکرت ۶ تایی،  $P < ۰/۰۱$ ) (۷ تا ۷). ولی در خصوص دیگر ویژگی‌های ارزیابان پژوهش‌های بسیار کمی انجام شده است. هیچ ارتباطی بین ویژگی‌های ارزیابان شامل اطلاعات دموگرافیک و نیز میزان تجربه آن‌ها (سال‌های تدریس) به عنوان عضو هیأت‌علمی بالینی پیدا نشد (۲۰ تا ۲۱). در حالی که بین مهارت‌های بالینی استادان و نمره‌گذاری آن‌ها ارتباط منفی معناداری وجود داشت، به این صورت که استادانی که در هر یک از حیطه‌های مهارت‌های بالینی از توانمندی بیشتری برخوردار بودند به طور سخت‌گیرانه‌تری به حیطه مربوطه نمره دادند (مهارت مصاحبه:  $r = -۰/۴۱$ ,  $P < ۰/۰۱$ ; مهارت معاینه فیزیکی:  $r = -۰/۴۲$ ,  $P < ۰/۰۱$ ; مهارت سازمان‌دهی:  $r = -۰/۳۶$ ,  $p < ۰/۰۵$ ).

• میزان سخت‌گیری ارزیابان در نمره‌دهی: از جمله مواردی که از طرف ارزیابان بر نمره‌دهی این آزمون تأثیرگذار است تفاوت آن‌ها از نظر میزان سخت‌گیری (stringency) در نمره‌دهی است. برخی از ارزیابان سهل‌گیر بوده و معمولاً انتهای بالای طیف نمرات را انتخاب می‌کنند (Doves) و برخی دیگر سخت‌گیر بوده و انتهای پایین طیف را انتخاب می‌کنند (Hawks) (۲). در برخی از مطالعات، تا ۴۰ درصد واریانس نمرات آزمون Mini-CEX مربوط به این موضوع بوده است (۱۷).

• ارفاق (Leniency): یکی دیگر از عوامل مؤثر، سهل‌گیری کلی ارزیابان نسبت به آزمون دهندگان و ارفاق به ایشان در نمره‌گذاری است. دیده شده است که ارزیابان به طور کلی نمرات بالایی به عملکرد فراگیران می‌دهند، به طوری که بررسی فرم‌های آزمون Mini-CEX مورد داخلی آمریکا نشان داد اکثریت نمرات بین ۶

• پیچیدگی موارد: مواجهه‌های با موارد بیماری دشوارتر نمرات بالاتری دریافت کردند. {موارد بیمار با پیچیدگی کم: ۴/۱۷، موارد بیماری با پیچیدگی متوسط: ۴/۴۰ و موارد بیماری با پیچیدگی زیاد: ۴/۵۰} (در فرم‌های با لیکرت ۶ تایی)  $P < 0.01$  (۵).

• بخش تخصصی: معمولاً نمره‌گذاری ارزیابان در بخش‌های تخصصی مختلف متفاوت بود. به عنوان مثال در برخی از مطالعات، جراحی پایین‌ترین نمره و بخش بیماری‌های زنان و زایمان بالاترین نمره را دارا بودند (به ترتیب: ۴/۲۵ و ۴/۵۳ در فرم‌های با لیکرت ۶ تایی،  $P < 0.01$ ) (۵).

• تمرکز مورد: این احتمال وجود دارد که نوع تمرکز مواجهه نیز بر نمره آزمون تأثیرگذار باشد. به عنوان مثال دیده شده است که مواجهات با تمرکز بر تشخیص بیماری کم‌ترین نمره را به خود اختصاص داده‌است (۵). از دیگر عوامل مربوط به مواجهه بالینی، تأثیر سن بیمار و طول مواجهه مورد بررسی قرار گرفته است که هیچ ارتباطی بین نمرات آزمون و دو مورد مذکور مشاهده نشده است (۵ و ۶ تا ۱۰).

### بحث

بر اساس نتایج مطالعات بررسی شده عوامل تأثیرگذار بر نمره‌دهی آزمون Mini-CEX شامل عوامل مربوط به ارزیابان، فرم‌های ارزیابی و عوامل مربوط به مواجهه بودند.

نتایج مطالعات همبستگی نشان داد که از جمله ویژگی‌های ارزیابان که در نمره‌دهی آزمون Mini-CEX مؤثر بود می‌توان به موقعیت این افراد به عنوان هیأت‌علمی یا دستیار و همچنین میزان تبحر ایشان در مهارت‌های بالینی اشاره کرد (۲۰ تا ۲۱). اهمیت موضوع اول در این است که اگر بخواهیم از Mini-CEX به عنوان ارزیابی پایانی استفاده کنیم باید منابع و افراد نمره‌دهنده بین دانشجویان متوازن شوند. در ارتباط با عامل دوم به نظر می‌رسد پزشکان بالینی طبابت و مهارت‌های خود را مبنای

CEX گزارش شد و پژوهشگران نشان دادند همه سؤالات سهم نسبتاً مساوی در نمره یا ارزیابی عملکرد دارند. علاوه بر این، دو مطالعه دیگر فرم‌های آزمون Mini-CEX را با استفاده از تحلیل عاملی در مقاطع دستیاری و پزشکی عمومی بررسی کردند و نتیجه گرفتند تنها یک عامل تحت عنوان توانمندی بالینی کلی (overall clinical competence) مسؤول واریانس نمرات Mini-CEX است (۲۹ و ۳۰). بنابراین، به نظر می‌رسد Mini-CEX تنها یک حیطة کلی از مهارت‌های بالینی را ارزیابی می‌کند.

در مورد دیگر ویژگی‌های فرم‌ها و تأثیر آن بر نمره‌دهی آزمون، تنها دو مطالعه به بررسی این موضوع پرداخته است (۲۵ و ۳۰). نتایج دو مطالعه مذکور نشان داد پایایی بین ارزیابان با تغییر تعداد گزینه‌های فرم‌ها و دیگر تغییرات مانند کاهش تعداد توانمندی‌های مورد ارزیابی مقیاس تغییری نکرد و فقط مقیاس نمره‌دهی ۹ تایی نمرات دقیق‌تری فراهم می‌کرد؛ به نحوی که فرم‌های نمره‌دهی ۹ تایی با دقت بیشتری توانست دستیاران را به داوطلبان با عملکرد غیرمطلوب یا عالی تقسیم نماید. بنابراین با وجود این که ارتباط بین دو نوع فرم زیاد بود ولی فرم‌های نمره‌دهی ۹ تایی بهتر توانست دستیارانی را که به برنامه‌های جبرانی نیاز دارند مشخص کند (۲۵). بنابراین، با توجه به این که این نتایج تنها از دو مطالعه به دست آمده است، نمی‌توان آن را به عنوان عوامل تأثیرگذار بر نمره‌دهی آزمون Mini-CEX مطرح نمود و نیازمند تحقیقات بیشتری می‌باشد.

### ۳. عوامل مربوط به مواجهه بالینی

بر اساس مرور متون یکی از عوامل دیگری که می‌تواند بر نمره‌دهی آزمون Mini-CEX تأثیرگذار باشد عوامل مربوط به مواجهه بالینی است. مهم‌ترین جنبه‌های مربوط به مواجهه که در متون به آن اشاره شده است عبارت است از:

کنند تا قضاوت ردی دستیاران را حمایت کند. در نتیجه آن‌ها صبر می‌کردند تا دستیاران توانمندی لازم جهت قبولی را کسب کنند و سپس عملکرد آنها را مورد ارزیابی قرار می‌دادند تا مجبور نباشند آنها را رد کنند. همچنین در مواردی که داوطلب مرزی بود به گونه‌ای رفتار می‌کردند که به نفع دانشجو باشد و در نتیجه در نهایت نمرات بیش از آن چیزی که باید باشد، گزارش می‌شده است (۳۲).

از دیگر مسائلی مربوط به نمره‌دهی نوع فرم‌های مورد استفاده و سؤالات آن است که به منظور تحلیل تأثیر این عوامل از آلفای کرونباخ و تحلیل عاملی استفاده می‌شود (۲۹). نتایج تحلیل فرم‌های نمره‌دهی آزمون Mini-CEX با استفاده از تحلیل عاملی و آلفای کرونباخ نشان داد که یکی از مشکلات نمره‌دهی آزمون ارتباط بالای بین سؤالات است. علل این مسأله می‌تواند عدم توانایی ارزیابان در تمایز قائل شدن بین ابعاد یا حیطه‌های مختلف مورد ارزیابی، ارتباط بسیار نزدیک حیطه‌ها با یکدیگر یا اثر هاله (halo effect) باشد. این احتمال وجود دارد که بخشی از مورد بیماری که برای مشاهده متمرکز استفاده می‌شود نمی‌تواند در عملکرد فراگیر در حیطه‌های مورد ارزیابی تمایز قائل شود یا این که اکثر فراگیران نقاط قوت و ضعف خیلی مشخصی در یک توانمندی خاص را نشان نمی‌دهند. حتی ممکن است شاخص‌های نمره‌دهی استفاده شده در فرم نمره‌دهی برای سؤالات مختلف تا حدودی مشابه باشد و با هم هم‌پوشانی داشته باشد. به عنوان مثال، توجه به راحتی بیمار هم در معاینه بالینی و هم در رفتار حرفه‌ای وجود دارد (۱۸). علت ارتباط زیاد بین سؤالات در حیطه‌های توانمندی مختلف هر چه باشد پیشنهادکننده این موضوع است که نمره دانشجو در یک حیطه می‌تواند به منظور ارائه بازخورد و ارزشیابی سازنده که هدف اصلی آزمون Mini-CEX است به کار رود، ولی در صورتی که آزمون Mini-CEX به عنوان آزمون پایانی استفاده

قضاوت عملکرد فراگیران قرار می‌دهند (۱۶ و ۳۱). این یافته‌ها اگر با پژوهش‌های دیگر تأیید شوند از این جهت اهمیت دارند که مهارت‌های بالینی برخلاف ویژگی‌های دموگرافیک قابل تغییر هستند و در نتیجه توانمندسازی ارزیابان می‌تواند علاوه بر مهارت‌های مشاهده و ارائه بازخورد شامل تقویت مهارت‌های بالینی استادان نیز باشد.

یکی از مهم‌ترین روش‌های تحلیل نتایج G theory است که به منظور تعیین واریانس یا تعیین میزان خطای مربوط به ارزیابان و مواجهه استفاده می‌شود. در اکثر مطالعاتی که در این مقاله مروری استفاده شد در تحلیل نتایج آزمون از این روش استفاده شده است.

مسأله ارفاق در قضاوت ارزیابان با استفاده از تحلیل نتایج آزمون و با استفاده از تئوری تعمیم‌پذیری نشان می‌دهد که عوامل دیگری غیر از توانمندی فراگیر و خصوصیات و ویژگی‌های ارزیاب بر قضاوت ارزیابان تأثیرگذار است. یکی از دلایل این ارفاق کاربرد نتایج آزمون برای دانشجو است. پژوهش‌ها نشان می‌دهد زمانی که نتایج آزمون برای دانشجو کاربردهای مهم‌تری دارد و منجر به تصمیم‌گیری‌های مهم در مورد وی مانند ارتقا به مرحله بعد می‌شود ارزیابان با ارفاق بیشتری عمل می‌کنند و نمرات بالاتری به آن‌ها می‌دهند (۱۹).

یکی دیگر از مسائلی مطرح در ارفاق در حین نمره‌دهی آزمون Mini-CEX این است که ارزیابان، بسیاری از موارد عملکرد نامطلوب را در طیف فراگیران با عملکرد مطلوب قرار می‌دهند. مشخص نیست علت این موضوع عدم توانایی ارزیابان در تمایز عملکرد این دو گروه از داوطلبان یا به دلیل اجتناب ارزیابان از مردود نمودن داوطلبان است. اعضای هیأت علمی در مطالعه کیفی الوس دلیمیا (Alves de Lima) و همکاران اظهار داشتند که آنها با اطمینان می‌توانستند دستیاران با عملکرد ضعیف را تشخیص دهند ولی اعلام ردی برای آنان مشکل بود. به این دلیل که نمی‌دانستند چه نوع رفتاری را باید مستند



خود می‌تواند بر روایی آزمون تأثیرگذار باشد. مرور عوامل مؤثر بر تغییرپذیری نمرات آزمون مذکور از این جهت اهمیت دارد که طراحان و اجراکنندگان آزمون با در نظر داشتن این عوامل می‌توانند روایی آزمون‌های Mini-CEX برگزار شده را ارتقا دهند. همان‌طور که در این مقاله بیان شد در مورد برخی از این عوامل نتایج پژوهش‌ها قوی بوده و در برخی موارد هنوز شواهد ناکافی است. به عنوان نمونه، به دلیل دخالت ویژگی‌های ارزیاب در نمره‌دهی، لازم است طراحان آزمون از ارزیابان مختلف برای ارزیابی یک فراگیر مشخص استفاده نمایند. این در حالی است که با توجه به ناکافی بودن پژوهش‌های انجام شده در زمینه توانمندسازی ارزیابان و یکسان نبودن نتایج آن نمی‌توان استنباط قطعی در زمینه اثربخشی دوره‌های آموزش ارزیابان داشت. به هر صورت پژوهش‌های بیشتری در این زمینه به منظور رسیدن به نتایج متقن مورد نیاز است. در مجموع بر اساس نتایج حاصل از مرور متون توصیه می‌شود نمره آزمون Mini-CEX به تنهایی مبنای ارزیابی پایانی قرار نگیرد. علاوه بر این، از نتایج آن در آزمون‌های پایانی هنجار محور به منظور رتبه‌بندی عملکرد دانشجویان استفاده نشود.

شود نمره تک تک سؤال‌ها ارزش افزوده‌ای برای آزمون بیش‌تر از نمره توانمندی کلی ندارد. هر چند به هر حال بررسی علت ارتباط بالای بین سؤالات فرم‌های نمره‌دهی آزمون و ارائه راه‌کار در مورد آن می‌تواند موضوع پژوهش‌های بیش‌تر باشد. در مورد عوامل مربوط به مواجهه بالینی ارتباط معنادار، ولی کوچکی بین نمره‌دهی ارزیاب و دشواری مورد بیماری مشاهده شد که نشان‌دهنده این موضوع است که ارزیابان سعی می‌کنند تا حدی دشواری مورد بیماری را با ارفاق جبران کنند. در نتیجه آموزش و توجیه ارزیابان در صورت دشواری موارد می‌تواند در تعدیل این مورد تأثیرگذار باشد. به منظور حذف اثر ارتباط نوع بخش و نمرات آزمون، باید دانشجویان به ویژه در مقطع پزشکی عمومی که در طول یک سال در بخش‌های مختلف چرخش دارند در همه بخش‌ها مورد ارزیابی قرار گیرند(۵). هر چند موارد مربوط به مواجهه بالینی در مطالعات مربوط به آزمون Mini-CEX کمابیش دیده می‌شود ولی به طور قطعی به اثبات نرسیده است.

### نتیجه‌گیری

همان‌طور که نشان داده شد یکی از نقاط ضعف آزمون Mini-CEX تغییرپذیری نمره‌دهی آن است که به نوبه

### منابع

1. Swanwick T. Understanding Medical Education: Evidence, Theory and Practice. West Sussex: John Wiley & Sons; 2010.
2. Holmboe ES, Hawkins RE. Practical guide to the evaluation of clinical competence. Philadelphia: Mosby/Elsevier; 2008.
3. Habibi H, Khaghanizade M, Mahmoodi H, Ebadi A, Seyedmazhari M. [Comparison of the Effects of Modern Assessment Methods (DOPS and Mini-CEX) with traditional method on Nursing Students' Clinical Skills: A Randomized Trial]. Iranian Journal of Medical Education. 2013; 13(5): 364-372. [Persian]
4. Imanipour M, Jalili M, Mirzazadeh A, Dehghan Nayeri N, Haghani H. [Viewpoints of Nursing Students and Faculties about Clinical Performance Assessment Using Programmatic Approach]. Iranian Journal of Medical Education. 2013; 12(10): 743-755. [Persian]
5. Hill F, Kendall K, Galbraith K, Crossley J. Implementing the undergraduate Mini-CEX: A tailored approach at Southampton University. Med Educ. 2009; 43(4): 326-34.
6. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The Mini-CEX (clinical evaluation exercise): A preliminary investigation. Ann Intern Med. 1995; 123(10): 795-9.
7. Nair BR, Alexander HG, McGrath BP, Parvathy MS, Kilsby EC, Wenzel J, et al. The mini clinical

- evaluation exercise (Mini-CEX) for assessing clinical performance of international medical graduates. *Med J Aust.* 2008; 189(3): 159-61.
8. Kogan JR, Bellini LM, Shea JA. Feasibility, Reliability, and Validity of the Mini-Clinical Evaluation Exercise (mCEX) in a Medicine Core Clerkship. *Acad Med.* 2003; 78(10 suppl): S33-35.
  9. Hatala R, Ainslie M, Kassen BO, Mackie I, Roberts JM. Assessing the mini-clinical evaluation exercise in comparison to a national specialty examination. *Med Educ.* 2006; 40(10): 950-6.
  10. Norcini JJ, Blank LL, Duffy FD, Fortna GS. The Mini-CEX : A method for assessing clinical skills. *Ann Intern Med.* 2003; 138(6): 476-81.
  11. Holmboe ES, Huot S, Chung J, Norcini J, Hawkins RE. Construct validity of the miniclinical evaluation exercise (miniCEX). *Acad Med.* 2003; 78(8): 826-30.
  12. Alves de Lima A, Barrero C, Baratta S, Castillo Costa Y, Bortman G, Carabajales J, et al. Validity, reliability, feasibility and satisfaction of the Mini-Clinical Evaluation Exercise (Mini-CEX) for cardiology residency training. *Med Teach.* 2007; 29(8): 785-90.
  13. Al Ansari A, Ali SK, Donnon T. The Construct and Criterion Validity of the Mini-CEX : A Meta-Analysis of the Published Research. *Academic Medicine.* 2013; 88(3): 413-20.
  14. Hawkins RE, Margolis MJ, Durning SJ, Norcini JJ. Constructing a validity argument for the mini-clinical evaluation exercise: A review of the research. *Acad Med.* 2010; 85(9): 1453-61.
  15. Pelgrim EA, Kramer AW, Mokkink HG, van den Elsen L, Grol RP, van der Vleuten CP. In-training assessment using direct observation of single-patient encounters: A literature review. *Adv Health Sci Educ Theory Pract.* 2011; 16(1): 189-199.
  16. Yeates P, O'Neill P, Mann K, W Eva K. 'You're certainly relatively competent': assessor bias due to recent experiences. *Med Educ.* 2013; 47(9): 910-22.
  17. Weller JM, Jolly B, Misur MP, Merry AF, Jones A, Crossley JG, et al. Mini-clinical evaluation exercise in anaesthesia training. *Br J Anaesth.* 2009; 102(5): 633-41.
  18. Margolis MJ, Clauser BE, Cuddy MM, Ciccone A, Mee J, Harik P, et al. Use of the mini-clinical evaluation exercise to rate examinee performance on a multiple-station clinical skills examination: A validity study. *Acad Med.* 2006; 81(10 suppl): S56-60.
  19. Alves de Lima A, Conde D, Costabel J, Corso J, Van der Vleuten C. A laboratory study on the reliability estimations of the Mini-CEX. *Adv in Health Sci Educ Theory Pract.* 2013; 18(1): 5-13.
  20. Kogan JR, Hess BJ, Conforti LN, Holmboe ES. What Drives Faculty Ratings of Residents' Clinical Skills? The Impact of Faculty's Own Clinical Skills. *Acad Med.* 2010; 85(10 Suppl): S25-28.
  21. Yeates P, O'Neill P, Mann K, Eva KW. Effect of exposure to good vs poor medical trainee performance on attending physician ratings of subsequent performances. *JAMA.* 2012; 308 (21): 2226-32.
  22. Holmboe ES, Hawkins RE, Huot SJ. Effects of Training in Direct Observation of Medical Residents' Clinical Competence. *Ann Intern Med.* 2004; 140(11): 874-81.
  23. Liao KC, Pul SJ, Liu MS, Yang CW, Kuo HP. Development and implementation of a mini-Clinical Evaluation Exercise (Mini-CEX) program to assess the clinical competencies of internal medicine residents: from faculty development to curriculum evaluation. *BMC Med Educ.* 2013; 13: 31.
  24. Cook DA, Dupras DM, Beckman TJ, Thomas KG, Pankratz VS. Effect of Rater Training on Reliability and Accuracy of Mini-CEX Scores: A Randomized, Controlled Trial. *J Gen Intern Med.* 2008; 24(1): 74-9.
  25. Cook DA, Beckman TJ. Does scale length matter? A comparison of nine- versus five-point rating scales for the mini-CE. *Adv in Health Sci Educ Theory Pract.* 2009; 14(5): 655-64.
  26. Durning SJ, Cation LJ, Markert RJ, Pangaro LN. Assessing the reliability and validity of the mini-clinical evaluation exercise for internal medicine residency training. *Acad Med.* 2002; 77(9): 900-4.
  27. Norcini JJ, Blank LL, Arnold GK, Kimball HR. Examiner differences in the Mini-CEX. *Adv Health Sci Educ Theory Pract.* 1997; 2(1): 27-33.
  28. Sahebalzamani M, Farahani H, Pouryani A. Assessing the Reliability and Validity of the Mini-Clinical Evaluation Exercise (Mini-CEX) for Nursing Student. *J Appl Environ Biol Sci.* 2013; 3(9): 78-82.
  29. Cook DA, Beckman TJ, Mandrekar JN, Pankratz VS. Internal structure of Mini-CEX scores for internal medicine residents: Factor analysis and generalizability. *Adv Health Sci Educ Theory Pract.* 2010; 15(5): 633-645.
  30. Donato AA, Pangaro L, Smith C, Rencic J, Diaz Y, Mensinger J, et al. Evaluation of a novel assessment form for observing medical residents: A randomized, controlled trial. *Med Educ.* 2008; 42(12): 1234-42.



31. Kogan JR, Conforti L, Bernabeo E, Iobst W, Holmboe E. Opening the black box of clinical skills assessment via observation: A conceptual model. *Med Educ.* 2011; 45(10): 1048–1060.
32. Alves de Lima A, Conde D, Aldunate L, Van der Vleuten CPM. Teachers' experiences of the role and function of the mini clinical evaluation exercise in postgraduate training. *Int J Med Educ.* 2010; 1: 68-73.

Archive of SID

# Factors Influencing Mini-Clinical Evaluation Exercise Scores: A Review Article

Roghayeh Gandomkar<sup>1</sup>, Mohammad Jalili<sup>2</sup>

## Abstract

**Introduction:** *The Mini-Clinical Evaluation Exercise (Mini-CEX ) is widely used to assess clinical performance because of its practicality. Although initial studies have supported its reliability and validity, the utility of Mini-CEX is under serious debate due to inconsistencies in rater judgments. This study aimed to review the relevant literature on the utility of Mini-CEX in order to find out the influencing factors on Mini-CEX scores.*

**Methods:** *In this narrative review study, the MEDLINE, EMBASE, SCOPUS, SID & Magiran were searched to retrieve the relevant articles published from 1995 to 2014, using the key words such as Mini-CEX , 'utility', 'validity', 'reliability', 'score variability'. Subsequently, nineteen relevant studies out of 51 articles were selected to be reviewed.*

**Results:** *Factors influencing Mini-CEX scores were related to the raters, rating forms and clinical exposures, in order of importance. Variability in raters' stringency and leniency and high intercorrelation between items were the most influencing factors affecting test scores. The results of the studies regarding factors such as rater training on Mini-CEX scores were not conclusive.*

**Conclusion:** *Considering the inconsistencies in Mini-CEX exam scores and uncertainty about the influencing circumstances on scoring, it is recommended that Mini-CEX scores should not be used for summative assessment purposes, especially in order to rank the performance of the students.*

**Keywords:** Mini-CEX , reliability, score consistency, utility, validity.

## Addresses:

<sup>1</sup>MD, MSc, PhD Candidate in Medical Education, Medical Education Department, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran. E-mail: rgandomkar@razi.tums.ac.ir

<sup>2</sup>(✉)Associate Professor, Emergency Medicine Department, Medical Education Department, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran. E-mail: mjalili@tums.ac.ir