

ضرایب تعمیم‌پذیری و اتکاپذیری نمره‌های مصاحبه‌آزمون پذیرش مقطع دکتری روان‌شناسی

ابراهیم علیزاده*

چکیده

نمره‌های مصاحبه به شرطی می‌تواند مبنای تصمیم‌گیری قرار بگیرد که میزان خطای اندازه‌گیری آن کم باشد. هدف از اجرای این پژوهش، برآورد ضرایب تعمیم‌پذیری و اتکاپذیری نمره مصاحبه پذیرش مقطع دکتری روان‌شناسی دانشگاه شهید بهشتی بود؛ بنابراین، نمره‌های مصاحبه ۲۷ نفر از متقاضیان پذیرش در رشته روان‌شناسی دانشگاه شهید بهشتی با یک طرح دوجهی متقاطع تحلیل شد. واریانس واقعی که مربوط به مؤلفه هدف اندازه‌گیری بود، حدود یک‌سوم از واریانس نمره مشاهده شده را به خود اختصاص داده بود. خطای اندازه‌گیری نامنظم یک‌پنجم واریانس و خطاهای اندازه‌گیری منظم، حدود نیمی از واریانس نمره مشاهده شده را به خود اختصاص داده بودند. یافته‌ها نشان داد که ضرایب تعمیم‌پذیری و اتکاپذیری نمره‌های مصاحبه قابل قبول و نزدیک به ضرایبی است که برای رشته‌های پزشکی برآورد شده است. با وجود این تعامل بین مصاحبه‌کنندگان با متقاضیان بیشترین سهم را در تولید خطای اندازه‌گیری منظم دارد که می‌تواند اعتبار نمره‌های به‌دست‌آمده را کاهش دهد. بر پایه نتایج، پیشنهاد می‌شود کارگاه آموزشی برای گروه‌های آموزشی روان‌شناسی برگزار شود که آنها بتوانند تحلیل پایایی انجام دهند تا شناخت بیشتری از کیفیت نمره‌های مصاحبه به دست بیاورند و مشخص کنند که تغییر در کدام‌یک از وجه‌ها باعث کاهش خطای منظم و افزایش پایایی نمره مصاحبه می‌شود.

واژگان کلیدی: ضریب تعمیم‌پذیری، ضریب اتکاپذیری، مصاحبه پذیرش، نظریه تعمیم‌پذیری.

*استادیار گروه روان‌شناسی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه شهید بهشتی، تهران، ایران (نویسنده

مسئول: E_Alizadeh@sbu.ac.ir)

مقدمه

نخستین گام برای بهبود کیفیت آموزش در دانشگاه این است که مناسب‌ترین افراد برای تحصیل در دانشگاه انتخاب شوند. البته جامعه نیز انتظار دارد که چنین شرایطی به بهترین وجه ممکن فراهم شود؛ زیرا جامعه از خدمات افرادی استفاده می‌کند که توسط دانشگاه‌ها، دانش‌آموخته می‌شوند (ارانیه^۱، ۲۰۱۶). شواهد موجود (تیل، مای‌فورد و دوئل^۲، ۲۰۱۳؛ تایمر و کلاسون^۳، ۲۰۱۱) نشان می‌دهد که سنجش شایستگی‌های داوطلبان تحصیل در مقاطع تحصیلات تکمیلی، کار دشواری است و استفاده از روش‌های نامناسب برای سنجش شایستگی میزان خطاهای اندازه‌گیری را افزایش می‌دهد.

بیشتر دانشگاه‌های جهان برای پذیرش دانشجو در مقاطع مختلف تحصیلی از روش‌های خاصی استفاده می‌کنند. استفاده از آزمون‌های ارزیابی موقعیت^۴، مراکز گزینش^۵ دانشجو، آزمون‌های استعداد، سوابق تحصیلی، سنجش هوش هیجانی و شخصیت، توصیه‌نامه، رزومه دانشجوی داوطلب و مصاحبه‌های ساختاریافته یا بدون ساختار از آن جمله هستند. در کنار هر یک از روش‌های گزینش دانشجو معمولاً از مصاحبه نیز استفاده می‌شود (پترسون، نایت، دوئل، نیکلسون، کازانس و کلند^۶، ۲۰۱۶). بدیهی است که همه دانشگاه‌ها از همه این روش‌ها استفاده نمی‌کنند؛ برای نمونه، برای پذیرش دانشجویان رشته‌های پزشکی در کشور کانادا، استرالیا و آمریکا بیشتر از آزمون‌های استعداد استفاده می‌شود (مک‌اندرو، ایس و والتین^۷، ۲۰۱۷). با وجود این، روش مصاحبه یا به‌عنوان یک مرحله جداگانه برای پذیرش دانشجو یا به‌عنوان یک مرحله پایانی در کنار روش‌های پیش‌گفته استفاده می‌شود تا میزان تناسب شغل با شاغل را مشخص کنند (مک‌اندرو و همکاران، ۲۰۱۷؛ پترسون و همکاران، ۲۰۱۶). کارکرد اصلی مصاحبه‌های پذیرش دانشجو این است که شایستگی‌های خاصی مانند تعاملات بین فردی، تفکر انتقادی (ارسه-فرر و کاستیلو^۸، ۲۰۰۷)، صداقت،

1. Oranye

2. Till, Myford & Dowell

3. Timer & Clauson

4. Situational Judgment Tests

5. Selection centers

6. Patterson, Knight, Dowell, Nicholson, Cousins & Cleland

7. McAndrew, Ellis & Valentine

8. Arce-Ferrer & Castillo

مسئولیت‌پذیری، پایبندی به اصول اخلاقی، همکاری و مشارکت (لیمی، لاکیر، کولین و برونل^۱، ۲۰۰۷) را می‌توان ارزیابی کرد که این شایستگی‌ها یا به‌وسیله روش‌های دیگر قابل ارزیابی نیست یا با صرف هزینه زیادتری امکان‌پذیر می‌شود. روش مصاحبه پذیرش دانشجو، انواع مختلفی دارد که مصاحبه‌های ساختاریافته، غیر ساختاریافته و خرده مصاحبه‌های چندگانه^۲ از آن جمله هستند (رزنفلد، ریتر، ترین و او^۳، ۲۰۰۸).

متغیرهای بسیاری در بافت مصاحبه پذیرش دانشگاه وجود دارد که می‌تواند خطای اندازه‌گیری منظم تولید کند و پایایی^۴ نمره‌های به‌دست‌آمده از مصاحبه را کاهش دهد (ارسه-فرر و کاستیلو، ۲۰۰۷). ویژگی‌های جمعیت‌شناختی داوطلبان پذیرش دانشگاه مانند سن، جنسیت و نژاد (لامادرید و همکاران^۵، ۲۰۱۲؛ جانس و فورستر^۶، ۲۰۱۱)، دانشگاه قبلی محل تحصیل و وضعیت اقتصادی-اجتماعی داوطلبان ورود به مقاطع بالاتر (تیلور، گرین و اسپروس^۷، ۲۰۱۵) می‌تواند بر نتایج مصاحبه اثر منفی داشته باشد. همچنین برخی از داوطلبان ورود به مقطع تحصیلی تلاش می‌کنند تا نظر مصاحبه‌کنندگان را از طریق مدیریت برداشت تغییر دهند (ارسه-فرر و کاستیلو، ۲۰۰۷). مدیریت برداشت به این معنی است که متقاضیان با رفتارهای کلامی (استورات، داستین، باریک و دارنولد^۸، ۲۰۰۸) و رفتارهای غیرکلامی مانند ظاهر و نوع لباس، دست دادن به مصاحبه‌کنندگان و لبخند زدن (باریک، شافر و دیگراسی^۹، ۲۰۰۹) بر نظرات مصاحبه‌کنندگان تأثیر می‌گذارند.

افزون بر ویژگی‌های جمعیت‌شناختی داوطلبان، برخی از مصاحبه‌کنندگان بر اساس ویژگی‌های فردی مانند سطح خلق روز مصاحبه و تجربه مصاحبه (زیدی، اسوبودا، وانگ و منوال^{۱۰}، ۲۰۱۴)، نظرات متفاوتی با سایر مصاحبه‌کنندگان دارند. یا این امکان وجود دارد که مصاحبه‌کنندگان تحت تأثیر اثر هاله‌ای قرار بگیرند؛ یعنی در ابتدای مصاحبه، یک برداشت کلی از مصاحبه‌شونده به دست بیاورند و تحت تأثیر آن، سایر

1. Lemay, Lockyer, Collin & Brownell

2. Multiple mini-interviews

3. Rosenfeld, Reiter, Trinh & Eva

4. reliability

5. Lamadrid et al

6. Jones & Forister

7. Taylor, Green & Spruce

8. Stewart, Dustin, Barrick & Darnold

9. Barrick, Shaffer & DeGrassi

10. Zaidi, Swoboda, Wang & Manuel

ویژگی‌ها را نادیده بگیرند. موارد پیش‌گفته از سوی مصاحبه‌کنندگان بر نمره مصاحبه تأثیر منفی بر جای می‌گذارد و خطای اندازه‌گیری نمره مصاحبه را افزایش می‌دهد (اوا، ریتر، رزنفلد و نورمن^۱، ۲۰۰۴). در این رابطه، نتایج پژوهش‌های انجام گرفته نشان می‌دهد که واریانس تعامل بین مصاحبه‌کنندگان و داوطلبان پذیرش در دانشگاه که تولیدکننده خطای اندازه‌گیری منظم است، بین ۳۷ (هانسن، کولاسگارام، وودز، فچتیک و اندرسن^۲، ۲۰۱۲) تا ۴۵ درصد از واریانس نمره‌های مشاهده شده را به خود اختصاص داده است (اوا و همکاران، ۲۰۰۴).

پرسش‌های مورد استفاده در مصاحبه پذیرش دانشگاه نیز می‌توانند در تولید خطای اندازه‌گیری منظم نقش داشته باشند. گاهی، پرسش‌ها به صورت اختصاصی سهم اندکی در تولید واریانس خطای منظم دارند؛ اما زمانی که پرسش با عامل ارزیاب تعامل پیدا می‌کند، سهم آن در واریانس خطای منظم افزایش می‌یابد. برای نمونه، برخی پژوهشگران (زیدی و همکاران، ۲۰۱۴) به این نتیجه رسیده‌اند که پرسش‌های مصاحبه پذیرش کمتر از ۲ درصد واریانس نمره مشاهده شده را به خود اختصاص داده است؛ در مقابل، پژوهشگرانی نیز وجود دارند که به این نتیجه رسیده‌اند که پرسش‌های مصاحبه چنانچه به صورت درست طراحی نشود، می‌تواند سهم زیادی در واریانس خطای منظم داشته باشد و این سهم تا ۲۵ درصد نیز برآورد شده است (ارسه-فرر و کاستیلو، ۲۰۰۷).

روش مصاحبه، ارزش کمی برای پیش‌بینی عملکرد تحصیلی و پیشرفت تحصیلی دانشجویان متقاضی پذیرش در دانشگاه دارد؛ برای نمونه فراتحلیل انجام شده درباره میزان اعتبار^۳ پیش‌بین نمره‌های مصاحبه در پیش‌بینی موفقیت تحصیلی دانشجویان رشته‌های پزشکی نشان می‌دهد که روش مصاحبه توانسته بین ۰/۰۳ تا ۰/۰۶ عملکرد دانشجویان را پیش‌بینی کند (گوهو و بلاک من^۴، ۲۰۰۶). بنابراین اگر از مصاحبه‌های ساختاریافته به همراه مقیاس‌های نمره‌دهی استاندارد و مصاحبه‌کنندگان آموزش‌دیده استفاده شود، پایایی و اعتبار نمره‌های مصاحبه افزایش می‌یابد و با آن می‌توان عملکرد تحصیلی دانشجویان را پیش‌بینی کرد (پترسون و همکاران، ۲۰۱۶). با وجود این با مصاحبه نمی‌توان توانایی افراد توانمند را به خوبی اندازه‌گیری کرد؛ به بیان دیگر اگر

1. Eva, Reiter, Rosenfeld & Norman

2. Hanson, Kulasegaram, Woods, Fechtig & Anderson

3. validity

4. Goho & Blackman

سطح توانایی همه داوطلبان بالا باشد و بخواهیم از بین آنها چند نفر را انتخاب کنیم، کار بسیار دشوار خواهد شد (ارانیه، ۲۰۱۶).

مصاحبه‌های پذیرش دانشجو در همه دانشگاه‌های دنیا صورت می‌گیرد و شواهد پژوهشی منتشر شده نشان می‌دهند که ضرایب پایایی آنها که در چهارچوب نظریه تعمیم‌پذیری برآورد شده است، قابل قبول است؛ برای نمونه ضریب تعمیم‌پذیری در یکی از دانشگاه‌های علوم پزشکی کشور کره ۰/۸۸ (کیم، نام و وون^۱، ۲۰۱۷)، ضریب تعمیم‌پذیری در یکی از دانشگاه‌های علوم پزشکی کشور مکزیک بین ۰/۸۴ تا ۰/۸۷ در پنل‌های مختلف مصاحبه (ارسه-فرر و کاستیلو، ۲۰۰۷)، ضریب تعمیم‌پذیری برآورد شده در یکی از دانشگاه‌های علوم پزشکی کشور کانادا ۰/۷۵ (هانسن و همکاران، ۲۰۱۲)، ضریب تعمیم‌پذیری برآورد شده در یکی از دانشگاه‌های علوم پزشکی کشور آمریکا ۰/۷۸ (اوا و همکاران، ۲۰۰۴) برآورد شده است. بر پایه نتایج موجود می‌توان بیان کرد که دامنه ضرایب متغیر است و تغییرپذیری موجود ناشی از خطاهای اندازه‌گیری نامنظم است.

بیشتر پژوهش‌ها درباره پایایی مصاحبه‌های پذیرش در رشته پزشکی اجرا شده است (هانسن و همکاران، ۲۰۱۲؛ کیم و همکاران، ۲۰۱۷؛ ارسه-فرر و کاستیلو، ۲۰۰۷؛ اوا و همکاران، ۲۰۰۴)؛ اما پژوهش در حوزه روان‌شناسی بسیار محدود است (هلمز و پاچانا^۲، ۲۰۰۸؛ فابر^۳، ۲۰۰۶). جستجو در سامانه سازمان اسناد و مدارک علمی ایران نشان داد که تاکنون پژوهشی در کشور با هدف برآورد پایایی مصاحبه‌های پذیرش دکتری روان‌شناسی منتشر نشده است؛ بنابراین، مشخص نیست که در مصاحبه پذیرش مقطع دکتری، خطاهای منظم، مانند نظرات مصاحبه‌کنندگان و ناهمگنی پرسش‌های مصاحبه، چند درصد از پراکندگی نمره مشاهده شده را به خود اختصاص می‌دهند و اگر مصاحبه‌کنندگان بر اساس نمره‌های تخصیص داده شده، بخواهند داوطلبان ورود به مقطع دکتری را با یکدیگر مقایسه کنند، تصمیم آنان چقدر پایایی دارد، یا اینکه اگر کمیته مصاحبه‌کننده بخواهد نمره افراد مصاحبه شده را با یک نمره برش مقایسه کنند تا مشخص شود که کدام فرد واجد شرایط قبولی در مقطع دکتری خواهد بود، آشکار نیست که این تصمیم‌گیری چقدر قابل اتکا است.

1. Kim, Nam & Kwon

2. Helmes & Pachana

3. Fauber

به‌منظور پاسخ به پرسش‌های پیش‌گفته باید خطاهای منظم موجود در بافت مصاحبه پذیرش مقطع دکتری روان‌شناسی با استفاده از یک طرح اندازه‌گیری مبتنی بر نظریه تعمیم‌پذیری^۱، برآورد شود تا آشکار شود که نمره‌های مصاحبه در چه شرایطی پایا است (برینان^۲، ۲۰۰۱) و خطاهای منظم و نامنظم چه سهمی در نمره مشاهده شده دارند. پژوهش درباره ضریب پایایی نتایج مصاحبه‌های پذیرش مقطع دکتری در رشته‌های روان‌شناسی، اطلاعات مناسبی از وضعیت موجود کیفیت مصاحبه‌های پذیرش مقطع دکتری روان‌شناسی در کشور ایجاد خواهد کرد تا گروه‌های روان‌شناسی دانشگاه‌های کشور با آگاهی از وضعیت موجود، بتوانند با ایجاد تغییرات لازم، کیفیت مصاحبه‌های پذیرش مقطع دکتری را بهبود دهند؛ بنابراین هدف از نگارش این مقاله، شناسایی مؤلفه‌های خطاهای اندازه‌گیری منظم در نمره‌های مصاحبه‌های پذیرش مقطع دکتری با استفاده از نظریه تعمیم‌پذیری است. در این راستا پرسش‌هایی به این شرح وجود دارد:

الف- هدف اندازه‌گیری چند درصد از واریانس نمره مشاهده شده را در نمره‌های مصاحبه پذیرش مقطع دکتری به خود اختصاص داده است؟ ب- کدام یک از منابع واریانس خطای اندازه‌گیری منظم، بیشترین سهم را در میزان ضرایب تعمیم‌پذیری یا اتکالپذیری نمره‌های مصاحبه پذیرش مقطع دکتری دارند؟ ج- اگر داوطلبان ورود به مقطع دکتری بر اساس نمره مصاحبه رتبه‌بندی شوند، نمره‌های به‌دست‌آمده چقدر پایایی دارند؟ د- اگر داوطلبان ورود به مقطع دکتری بر اساس یک نمره برش قبول یا رد شوند، نتایج قبولی یا رد شده چقدر قابل اتکا هستند؟

روش و نمونه

این پژوهش، یک مطالعه توصیفی است که در چهارچوب نظریه تعمیم‌پذیری، سهم منابع خطای منظم نمره‌های مصاحبه آزمون پذیرش مقطع دکتری و اثرات آنها را بر پایایی نمره‌های مصاحبه، توصیف می‌کند. نظریه تعمیم‌پذیری با استفاده از طرح‌های تحلیل واریانس، مجموعه‌ای از طرح‌های اندازه‌گیری را معرفی می‌کند که با استفاده از آنها می‌توان خطای موجود در داده‌ها را به دو مؤلفه واریانس خطای تصادفی و خطای منظم تجزیه کرد. این نظریه با برآورد مؤلفه‌های واریانس^۳، شرایطی فراهم می‌کند تا

1. Generalizability Theory (GT)

2. Brennan

3. variance components

سهم نسبی عوامل تشکیل‌دهنده خطای منظم و خطای نامنظم در نمره مشاهده شده تعیین شود (شولسون، وب و رولی^۱، ۱۹۸۹).

در این پژوهش، نمره‌های مصاحبه‌پذیرش مقطع دکتری روان‌شناسی دانشگاه شهید بهشتی در سال تحصیلی ۹۷-۱۳۹۶ تحلیل شده است. نمره‌های مصاحبه برای ۲۷ داوطلب پذیرش رشته روان‌شناسی عمومی بودند که توسط پنج مصاحبه‌کننده مورد ارزیابی قرار گرفته بودند. همه مصاحبه‌کنندگان عضو هیئت علمی گروه آموزشی مورد نظر بودند. لازم به ذکر است که همه داوطلبان ورود به مقطع دکتری، مرد بودند.

ابزار اندازه‌گیری

ابزار اندازه‌گیری مصاحبه‌کنندگان در جلسات مصاحبه، پرسشنامه‌ای با نام فرم شماره ۴-۱ بود که در سازمان سنجش برای مصاحبه‌پذیرش مقطع دکتری طراحی شده است. این پرسشنامه پنج سؤال به صورت عبارت‌های مصدری داشت که عبارت‌اند از: ۱- تسلط در تجزیه و تحلیل مسائل علمی و پاسخگویی به پرسش‌ها؛ ۲- وسعت نظر، نوآوری و کارآفرینی؛ ۳- شخصیت، متانت و نحوه تعامل؛ ۴- نگرش و اطلاعات فناورانه مرتبط با رشته تحصیلی؛ ۵- توانایی فن بیان و انتقال مطالب. در این پرسشنامه، عضو هیئت علمی گروه آموزشی میزان هر شایستگی را در یک مقیاس پنج‌درجه‌ای از یک تا پنج (خیلی زیاد تا خیلی کم) ارزیابی می‌کند.

شیوه اجرا

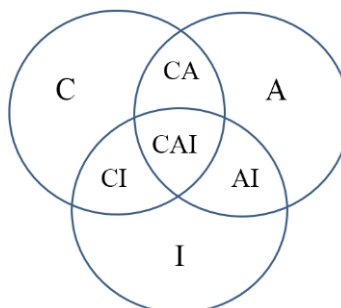
هر داوطلب به صورت جداگانه وارد اتاق مصاحبه شده و مصاحبه‌کنندگان سؤال‌هایی را از او پرسیده‌اند تا تعیین کنند که او هر یک از شایستگی‌های مورد نظر را چقدر دارد. روال کار به این شکل بوده که مدیر جلسه پیش از ورود داوطلب درباره رشته لیسانس و فوق لیسانس و محل اخذ مدرک متقاضی، اطلاعاتی را برای اعضای گروه می‌خواند. همچنین هر مصاحبه‌کننده، مسئول پرسش یکی از سؤال‌های پرسشنامه مصاحبه بود؛ بنابراین اگر یک مصاحبه‌کننده سؤالی را می‌پرسید، سایر مصاحبه‌کنندگان آن سؤال را نمی‌پرسیدند و بر اساس پاسخ‌های داوطلب به پرسش مورد نظر، سایر مصاحبه‌کنندگان

¹. Shavelson, Webb & Rowley

سطح شایستگی مورد نظر را ارزیابی می‌کردند. لازم به ذکر است که مصاحبه‌کنندگان به صورت مستقل از هم نظراتشان را در فرم ارزیابی مصاحبه ثبت کرده‌اند. مدت زمان مصاحبه برای هر نفر به طور متوسط ۱۵ دقیقه بوده است.

تحلیل داده‌ها

برای برآورد پایایی نمره مصاحبه از یک طرح دووجهی متقاطع استفاده شد. وجه‌های اندازه‌گیری این طرح (وجه مصاحبه‌کنندگان و وجه پرسش‌های مصاحبه) دربردارنده دو منبع تولیدکننده خطاهای اندازه‌گیری منظم بودند و فرض بر این بود که باعث کاهش پایایی نمره‌های مربوط به هدف اندازه‌گیری مصاحبه می‌شوند. وجه مصاحبه‌کنندگان و وجه پرسش‌ها به عنوان وجه تصادفی در نظر گرفته شدند؛ چون فرض شد که کمیت مصاحبه‌کننده می‌توانست از اعضای سایر گروه‌های روان‌شناسی نیز انتخاب شوند. همچنین، پرسش‌های مصاحبه از بین یک خزانه سؤال به صورت تصادفی انتخاب شده بود و به غیر از اینها، پرسش‌های دیگری نیز می‌توانست طراحی و انتخاب شود. در مصاحبه‌های صورت گرفته، ارتباط بین وجه‌های اندازه‌گیری به این صورت بود که وجه‌های اندازه‌گیری با هدف اندازه‌گیری متقاطع شده بودند؛ بنابراین طرح اندازه‌گیری را می‌توان به صورت زیر نشان داد:



نمودار (۱) طرح اندازه‌گیری و ارتباط وجه‌های اندازه‌گیری پژوهش

توصیف نمادهای اختصاری در جدول (۱) نوشته شده است. برای برآورد پایایی مبتنی بر طرح دووجهی متقاطع باید ابتدا ۷ مؤلفه واریانس به صورت هم‌زمان برآورد شود. همه مؤلفه‌ها در جدول زیر توصیف شده‌اند.

جدول (۱) توصیف مؤلفه‌های واریانس درگیر در نمره‌های مصاحبه پذیرش دکتری

توصیف	نماد	منابع واریانس
این وجه هدف اندازه‌گیری است و افزایش واریانس آن نشان‌دهنده تفاوت‌های فردی بین داوطلبان پذیرش مقطع دکتری است.	C	وجه متقاضیان پذیرش مقطع دکتری
برخی از مصاحبه‌کنندگان، بر اساس ویژگی‌های فردی مانند سطح خلاق روز مصاحبه و تجربه مصاحبه، نظرات متفاوتی با سایر مصاحبه‌کنندگان داشته‌اند.	I	وجه مصاحبه‌کنندگان
واریانس موجود در بین پرسش‌های مصاحبه نشان می‌دهد که محتوای پرسش‌ها با یکدیگر ناهمگن است.	A	وجه پرسش‌های مصاحبه
واریانس موجود نشان می‌دهد که برخی از مصاحبه‌کنندگان، به دلیل مدیریت برداشت نظر خاصی نسبت به برخی از داوطلبان پذیرش مقطع دکتری داشته‌اند؛ به همین دلیل، شایستگی‌های آنها را بیش برآورد یا کم برآورد کرده‌اند.	CI	وجه تعامل مصاحبه‌کنندگان با وجه متقاضیان پذیرش مقطع دکتری
واریانس موجود نشان می‌دهد که محتوای برخی از پرسش‌ها، به گونه‌ای بوده است که برخی از داوطلبان پذیرش مقطع دکتری صرف نظر از ویژگی واقعی، نمره بالاتر یا پایین‌تری به دست آورده‌اند.	CA	وجه تعامل متقاضیان پذیرش مقطع دکتری با وجه پرسش‌های مصاحبه
واریانس موجود نشان می‌دهد که برخی از مصاحبه‌کنندگان، تفسیر خاصی از برخی سپرسش‌های مصاحبه داشته‌اند.	AI	وجه تعامل مصاحبه‌کنندگان با وجه پرسش‌های مصاحبه
واریانس موجود نشان می‌دهد که وجه‌های اندازه‌گیری با یکدیگر و با هدف اندازه‌گیری تعامل مرتبه بالاتری را پیدا کرده‌اند که قابل تفسیر نیست. علاوه بر آن خطای اندازه‌گیری نامنظم نیز با این نوع تعاملات ترکیب شده است.	CAI	وجه تعامل‌های مرتبه بالا و خطای اندازه‌گیری نامنظم

برای برآورد مؤلفه‌های واریانس، از نسخه ۶/۱ نرم‌افزار EduG (کاردینت، جانسون و پینی^۱، ۲۰۱۰) استفاده شد. بزرگ بودن یک مؤلفه واریانس نشان می‌دهد که آن مؤلفه تا چه حد در خطای اندازه‌گیری نقش دارد. نتایج به‌دست‌آمده در جدول (۲) ارائه شده است.

¹. Cardinet, Johnson & Pini

جدول (۲) مؤلفه‌های واریانس برآورد شده برای طرح اندازه‌گیری دوجبهی متقاطع

منابع واریانس	نماد	درجه آزادی	میانگین مجذورات	مؤلفه‌های واریانس برآورد شده	سهم مؤلفه در واریانس نمره کل (به درصد)	خطای استاندارد برآورد
وجه داوطلبان	C	۲۰	۱۱/۰۱	۰/۳۷	۳۲/۸	۰/۱۳
وجه مصاحبه‌کنندگان	I	۴	۹/۳۱	۰/۰۶	۵/۵	۰/۰۵
وجه پرسش‌های مصاحبه	A	۴	۱۲/۲۵	۰/۱	۸/۹	۰/۰۷
وجه تعامل مصاحبه‌کنندگان با وجه داوطلبان	CI	۸۰	۱/۵۶	۰/۲۶	۲۳	۰/۰۵
وجه تعامل داوطلبان با وجه پرسش‌های مصاحبه	CA	۸۰	۰/۳۸	۰/۰۳	۲/۲	۰/۰۱
وجه تعامل مصاحبه‌کنندگان با وجه پرسش‌های مصاحبه	AI	۱۶	۱/۴۷	۰/۰۶	۵/۱	۰/۰۲
وجه تعامل‌های مرتبه بالا و خطای اندازه‌گیری نامنظم	CAI	۳۲۰	۰/۲۶	۰/۲۶	۲۲/۵	۰/۰۲

مؤلفه واریانس وجه داوطلبان پذیرش مقطع دکتری که به‌عنوان هدف اندازه‌گیری به شمار آورده می‌شود، حدود ۳۰ درصد واریانس نمره مشاهده‌شده را به خود اختصاص داده است که نشان می‌دهد، همه کسانی که در مصاحبه شرکت کرده‌اند، از نظر مصاحبه‌کنندگان، تفاوت‌های فردی زیادی را در شایستگی‌های مورد ارزیابی نداشته‌اند. مؤلفه‌های واریانس تعامل‌های مرتبه دوم بین وجه داوطلبان پذیرش مقطع دکتری، مصاحبه‌کنندگان و پرسش‌های مصاحبه و خطاهای نامنظم که به‌عنوان واریانس باقی‌مانده به شمار می‌آید و همچنین حدود ۲۰ درصد واریانس نمره مشاهده شده ناشی از وجه تعامل مصاحبه‌کنندگان با وجه متقاضیان پذیرش بوده است. علاوه بر این، خطاهای استاندارد برآورد شده نشان دادند که برآوردهای صورت گرفته برای مؤلفه‌های

واریانس و سهم اختصاصی آنها تحت تأثیر حجم نمونه و تغییرپذیری نمونه‌گیری قرار نداشته‌اند.

پس از برآورد مؤلفه‌های واریانس، پایایی نمره‌های مصاحبه برآورد شد. در نظریه تعمیم‌پذیری، دو نوع ضریب پایایی قابل برآورد است که با نام ضرایب تعمیم‌پذیری و اتکاپذیری شناخته می‌شوند. برای برآورد ضریب تعمیم‌پذیری همه مؤلفه‌هایی که با هدف اندازه‌گیری در تعامل است به‌عنوان خطای اندازه‌گیری به شمار می‌آید. همچنین برای برآورد ضریب اتکاپذیری به‌جز مؤلفه هدف اندازه‌گیری، همه مؤلفه‌های موجود در طرح اندازه‌گیری به‌عنوان مؤلفه خطا در نظر گرفته می‌شود (برینان، ۲۰۰۱)؛ بر این اساس، فرمول‌ها برای برآورد ضرایب مورد نظر به شرح جدول (۳) خواهد بود.

جدول (۳) ضریب تعمیم‌پذیری و ضریب اتکاپذیری نمره‌های مصاحبه پذیرش مقطع دکتری

$E\rho^2 = \frac{\sigma^2 C}{[\sigma^2 C] + \left[\frac{\sigma^2 I}{nI} + \frac{\sigma^2 A}{nA} + \frac{\sigma^2 CI}{nCI} + \frac{\sigma^2 CA}{nCA} + \frac{\sigma^2 IA}{nIA} + \frac{\sigma^2 CIA}{nC, nI, nA} \right]}$	ضریب تعمیم‌پذیری
$\phi = \frac{\sigma^2 C}{[\sigma^2 C] + \left[\frac{\sigma^2 I}{nI} + \frac{\sigma^2 A}{nA} + \frac{\sigma^2 CI}{nCI} + \frac{\sigma^2 CA}{nCA} + \frac{\sigma^2 IA}{nIA} + \frac{\sigma^2 CIA}{nC, nI, nA} \right]}$	ضریب اتکاپذیری

ضرایب تعمیم‌پذیری و اتکاپذیری به ترتیب برابر با ۰/۸۵ و ۰/۷۸ به دست آمد که نتایج به‌دست‌آمده نشان می‌دهد که نمره‌های مصاحبه، پایایی قابل قبولی داشته‌اند، چون بالای ۰/۷ هست. در چهارچوب نظریه تعمیم‌پذیری، پایایی، نسبت واریانس نمره مرجع^۱ (واریانس هدف اندازه‌گیری) به واریانس نمره مشاهده مورد انتظار^۲ است. در چهارچوب نظریه تعمیم‌پذیری می‌توان دو نوع ضریب پایایی با عنوان ضریب تعمیم‌پذیری^۳ و ضریب اتکاپذیری^۴ برآورد کرد (کاردینت و همکاران، ۲۰۱۰؛ شولسون و همکاران، ۱۹۹۱). بر اساس ضریب تعمیم‌پذیری می‌توان مشخص کرد که اگر بنا بود که نمره‌های مصاحبه‌شوندگان به ترتیب از بیشترین به کمترین نمره مرتب شود و سپس

1. universe score

2. expected observed score

3. generalizability coefficient

4. dependability coefficient

افرادی که بالاترین نمره‌ها را به دست آورده‌اند، در دانشگاه پذیرفته شوند، این تصمیم تا چه اندازه پایا بود. ضریب اتکاپذیری نیز مشخص می‌کند که اگر بنا بود فقط داوطلبانی انتخاب شوند که همه نمره‌های آنها بر پایه مقیاس‌های پاسخ‌دهی پرسشنامه در حد زیاد بود (نمره ۴ از ۵) تا چه اندازه تصمیم‌گیری مصاحبه‌کنندگان باثبات می‌بود.

پس از برآورد ضریب تعمیم‌پذیری و ضریب اتکاپذیری تصمیم گرفته شد تا برای بررسی تأثیرگذاری هر یک از منابع خطای اندازه‌گیری منظم و نامنظم بر ضریب تعمیم‌پذیری و ضریب اتکاپذیری، سهم واریانس مشاهده شده مربوط به هر یک از مؤلفه‌های خطا در یک مقیاس درصدی بررسی شد. بر این اساس، منابعی از خطای اندازه‌گیری منظم شناسایی شد که تأثیر منفی بر پایایی نمره‌های مصاحبه‌های صورت گرفته دارند. نتایج در جدول (۴) ارائه شده است.

جدول (۴) تفکیک مؤلفه‌های واریانس خطای منظم و نامنظم هنگام تصمیم‌گیری نسبی و مطلق

منابع واریانس	نماد	واریانس خطای نسبی	سهم واریانس خطای نسبی (به درصد)	واریانس خطای مطلق	سهم واریانس خطای مطلق (به درصد)
وجه مصاحبه‌کنندگان	I	۰/۰۱	۱۲/۱
وجه پرسش‌های مصاحبه	Q	۰/۰۲	۱۹/۸
وجه تعامل مصاحبه‌کنندگان با وجه داوطلبان	CI	۰/۰۵	۷۷/۳	۰/۰۵	۵۰/۹
وجه تعامل داوطلبان با وجه پرسش‌های مصاحبه	CQ	۰/۰۱	۷/۵	۰/۰۱	۵
وجه تعامل مصاحبه‌کنندگان با وجه پرسش‌های مصاحبه	IQ	۰/۰۰۱	۲/۳
وجه تعامل‌های مرتبه بالا و خطای اندازه‌گیری نامنظم	CIQ	۰/۰۱	۱۵/۱	۰/۰۱	۱۰
مجموع		۰/۰۷	۱۰۰	۰/۱	۱۰۰

تحلیل منابع خطای اندازه‌گیری نشان داد که مصاحبه‌کنندگان هنگام ارزیابی داوطلبان پذیرش مقطع دکتری، نظر خاصی نسبت به برخی از آنها داشته‌اند؛ به همین دلیل، شایستگی‌های آنها را بیش برآورد کرده‌اند. این نتیجه از روی مؤلفه تعامل بین مصاحبه‌کنندگان با داوطلبان پذیرش مقطع دکتری قابل بررسی است. بر این اساس، در تصمیم‌گیری نسبی و مطلق به ترتیب نزدیک ۸۰ درصد و ۵۰ درصد واریانس خطای منظم ناشی از تعامل بین مصاحبه‌کنندگان با داوطلبان پذیرش بوده است.

نتایج موجود نشان داد که خطای اندازه‌گیری نامنظم کمترین سهم را در واریانس نمره‌های مشاهده شده داشته است. سهم خطای اندازه‌گیری نامنظم از روی مؤلفه تعاملات مرتبه بالاتر بین وجه‌های اندازه‌گیری قابل بررسی است، به گونه‌ای که این مؤلفه به ترتیب فقط حدود ۱۵ درصد و ۱۰ درصد واریانس خطای نامنظم را به خود اختصاص داده‌اند.

بحث و نتیجه‌گیری

هدف از مصاحبه پذیرش مقطع دکتری این است که از بین افراد داوطلب تحصیل، افرادی انتخاب شوند که توانایی تحلیل مسائل علمی را داشته باشند، بتوانند با نوآوری و خلاقیت، راه حل‌های مناسبی را برای مسائل بیابند (ارسه-فرر و کاستیلو، ۲۰۰۷) و نسبت به فناوری‌های جدید مرتبط با رشته تحصیلی خود آگاهی داشته باشند؛ علاوه بر این، تعاملات بین فردی مناسبی داشته باشند تا بتوانند در گفتگوهای خود، مطالب را به نحو مناسبی به طرف مقابل انتقال دهند.

شناخت شایستگی‌های علمی داوطلبان پذیرش با استفاده از روش مصاحبه، کار بسیاری دشواری است (ارسه-فرر و کاستیلو، ۲۰۰۷)؛ زیرا منابع خطای بسیاری وجود دارند که می‌تواند پایایی نمره‌ها را کاهش دهد؛ بنابراین نیاز است که در مرحله نخست، پایایی جلسات مصاحبه برآورد شود تا آشکار شود که آیا ضریب به‌دست‌آمده قابل قبول است و کدام‌یک از منابع خطا، خطای منظم تولید می‌کنند. این پژوهش، نخستین مطالعه برای برآورد ضرایب تعمیم‌پذیری و اتکاپذیری مصاحبه پذیرش رشته روان‌شناسی در کشور ایران به شمار می‌آید.

یافته‌های به‌دست‌آمده نشان داد که هدف اندازه‌گیری، حدود ۳۰ درصد در واریانس نمره مشاهده شده نقش دارد و این نتیجه با پژوهش‌های انجام گرفته پیشین (هانسن و همکاران، ۲۰۱۲؛ ارسه-فرر و کاستیلو، ۲۰۰۷؛ او، ریتز، رزنفلد و نورمن، ۲۰۰۴) که

ضرایب پایایی مناسبی را در مصاحبه پذیرش رشته‌های پزشکی گزارش کرده‌اند، همسو است. در آن پژوهش‌ها، واریانس هدف اندازه‌گیری با سهم‌های متفاوتی به‌دست‌آمده است و در دامنه بین ۱۳ تا ۵۰ درصد (ارسه-فرر و کاستیلو، ۲۰۰۷)، ۱۴ درصد (اوا و همکاران، ۲۰۰۴)، ۴۰ درصد (هانسن و همکاران، ۲۰۱۲) متغیر است. همچنین، مشخص شد که خطای اندازه‌گیری نامنظم، حدود ۲۲/۵ درصد در واریانس نمره مشاهده شده نقش دارد. زیدی، اسوبودا، وانگ و منوال (۲۰۱۴) و هانسن، کولاسگارام، وودز، فچتیک و اندرسن (۲۰۱۲) نیز نتایج مشابهی به دست آورده‌اند. در آن پژوهش‌ها، واریانس خطای اندازه‌گیری نامنظم در دامنه بین ۱۵ درصد (زیدی و همکاران، ۲۰۱۴) تا ۲۱ درصد (هانسن، کولاسگارام، وودز، فچتیک و اندرسن، ۲۰۱۲) برآورد شده است.

نتایج پژوهش حاضر نشان داد که وجه پرسش‌های مصاحبه در صورت تصمیم‌گیری مطلق، ۲۰ درصد کل خطای اندازه‌گیری منظم را به خود اختصاص داده است؛ اما تعامل بین وجه تعامل متقاضیان پذیرش با وجه پرسش‌های مصاحبه، سهم بسیار ناچیزی در خطای اندازه‌گیری منظم داشته است؛ بنابراین می‌توان نتیجه گرفت که پرسش‌های پرسشنامه مصاحبه، یک ویژگی واحد را اندازه می‌گیرند. شبیه چنین نتیجه‌ای در سایر پژوهش‌های مشابه (زیدی و همکاران، ۲۰۱۴؛ هانسن و همکاران، ۲۰۱۲؛ ارسه-فرر و کاستیلو، ۲۰۰۷؛ اوا و همکاران، ۲۰۰۴) نیز به‌دست‌آمده است. هرچند پرسش‌های مصاحبه این پژوهش با پژوهش‌های یاد شده تفاوت داشته است؛ اما آنها نیز نشان داده‌اند که پرسش‌های مصاحبه، سهم چندانی در تولید خطای منظم ندارند.

نتایج پژوهش حاضر نشان داد که وجه تعامل مصاحبه‌کنندگان با وجه داوطلبان پذیرش، بیشترین سهم واریانس خطای منظم را در تصمیم‌گیری نسبی و مطلق داشته است و بنابراین برای افزایش پایایی نمره‌های مصاحبه باید این منبع خطا، کاهش یابد. برخی از پژوهش‌ها نیز به نتیجه مشابهی رسیده‌اند (مانند پژوهش هانسن و همکاران، ۲۰۱۲؛ اوا و همکاران، ۲۰۰۴) و نشان داده‌اند که نقش تعامل بین مصاحبه‌کنندگان با داوطلبان پذیرش، بیشترین نقش را در تولید خطای اندازه‌گیری منظم دارد.

این پژوهش شواهدی فراهم کرد، مبنی بر اینکه خطای نامنظم، کمترین سهم را در کاهش پایایی نمره‌های مصاحبه پذیرش دکتری دارد و این نتیجه دور از انتظار نبود؛ زیرا اعضای گروه آموزشی با هدف انتخاب بهترین داوطلبان، مصاحبه را انجام می‌دادند. در برخی از پژوهش‌ها نیز نتیجه مشابهی به‌دست‌آمده است (زیدی و همکاران، ۲۰۱۴؛ هانسن و همکاران، ۲۰۱۲)؛ هرچند پژوهش‌های یاد شده وجود دارند که حدود ۴۰ درصد

خطای اندازه‌گیری تصادفی داشته‌اند (سبوک، لو و کلینگر^۱، ۲۰۱۴؛ ارسه-فرر و کاستیلو، ۲۰۰۷).

نتایج مصاحبه پذیرش مقطع دکتری همواره تحت تأثیر خطاهای بالقوه اندازه‌گیری قرار دارد که می‌توانند خطای منظم تولید کنند (تیلور و همکاران، ۲۰۱۵؛ زیدی و همکاران، ۲۰۱۴؛ هانسن و همکاران، ۲۰۱۲؛ لامادرید و همکاران، ۲۰۱۲؛ جانس و فورستر، ۲۰۱۱؛ ارسه-فرر و کاستیلو، ۲۰۰۷؛ اوا و همکاران، ۲۰۰۴). به‌منظور شناسایی سهم هر یک از منابع خطا در واریانس نمره مشاهده شده مصاحبه پذیرش مقطع دکتری روان‌شناسی از طریق تفسیر مؤلفه‌های واریانس نشان دادیم که هر وجه اندازه‌گیری، چند درصد در واریانس نمره مشاهده شده اثر دارد و چند درصد واریانس نمره مشاهده شده مربوط به هدف اندازه‌گیری است.

یافته به‌دست‌آمده نشان داد که ضریب تعمیم‌پذیری به‌دست آمده تقریباً نزدیک به ضریبی است که در پژوهش‌های انجام گرفته پیشین (هانسن و همکاران، ۲۰۱۲؛ ارسه-فرر و کاستیلو، ۲۰۰۷؛ اوا و همکاران، ۲۰۰۴) به‌دست آمده است؛ اما ضریب اتکاپذیری پایین‌تر از آنها برآورد شده است. بنابراین اگر برای پذیرش دانشجویان، نمره مصاحبه آنها با یک نمره برش مقایسه شود، این تصمیم‌گیری از پایایی بالایی برخوردار نخواهد بود. علت اصلی آن این است که خطای اندازه‌گیری منظم ناشی از تعامل مصاحبه‌کنندگان با داوطلبان پذیرش، سهم زیادی در واریانس نمره مشاهده شده، داشته است.

نتایج این پژوهش نشان داد که نمره‌های مصاحبه‌های پذیرش مقطع دکتری، ضریب پایایی قابل قبولی دارند؛ با وجود این، خطاهای اندازه‌گیری منظم به‌ویژه تعامل بین مصاحبه‌کنندگان با داوطلبان پذیرش وجود دارد و سهم آن را در تولید واریانس خطا نمی‌توان نادیده گرفت؛ بنابراین برای افزایش ضرایب تعمیم‌پذیری و اتکاپذیری نمره‌های مصاحبه پذیرش مقطع دکتری در رشته‌های روان‌شناسی دانشگاه‌های کشور، چند راهکار پیشنهاد می‌شود. نخست آنکه نمره‌های مصاحبه از لحاظ منابع تشکیل‌دهنده خطای اندازه‌گیری منظم تحلیل شود و بازخوردهای لازم به گروه آموزشی مورد نظر داده شود تا آنها شناخت بیشتری درباره نحوه ارزیابی‌شان به دست آورند. علاوه بر این، هر دانشگاه با کمک متخصصان سنجش و اندازه‌گیری، کارگاه‌های آموزشی ویژه‌ای را

^۱. Sebok, Luu & Klinger

برای اعضای گروه‌های آموزشی روان‌شناسی برگزار کنند که آنها بتوانند خودشان تحلیل‌های پایایی را انجام دهند و در نتیجه مدیران گروه‌های آموزشی بتوانند خودشان در پایان مرحله پذیرش دانشجویان دکتری به این پرسش پاسخ دهند که روش مصاحبه تا چه اندازه توانسته است که متقاضیان توانمند را از بین سایرین جدا کند. همچنین مشخص کنند که تغییر در کدام‌یک از وجه‌ها باعث کاهش خطای منظم و افزایش پایایی نتایج مصاحبه می‌شود.

به دلیل دسترسی نداشتن به اطلاعات مصاحبه دیگر سازمان‌ها و به‌ویژه سایر دانشگاه‌ها نتایج این پژوهش نتوانست اطلاعات فراگیرتری بسته به وضعیت موجود پایایی نمره‌های مصاحبه پذیرش مقطع دکتری فراهم آورد؛ بنابراین نتایج این پژوهش، باید با احتیاط به مصاحبه‌های دانشگاه‌های دیگر، تعمیم داده شود؛ زیرا نتایج فقط از داده‌های یک دانشگاه به دست آمده است. در همین رابطه، به پژوهشگران علاقه‌مند پیشنهاد می‌شود که در پژوهش‌های آینده، داده‌های بیشتری را از دانشگاه‌های دولتی و آزاد به دست آورند و پایایی مصاحبه‌های پذیرش آنها را در چهارچوب نظریه تعمیم‌پذیری برآورد کنند.

منابع

- Arce-Ferrer, A. J. & Castillo, I. B. (2007). Investigating postgraduate college admission interviews: generalizability theory reliability and incremental predictive validity. *Journal of Hispanic Higher Education*, 6 (2), 118-134. Doi: 10.1177/1538192707299287.
- Barrick, M. R.; Shaffer, J. & DeGrassi, S. D. (2009). What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology*, 94 (6), 1394-1412. Doi: 10.1037/a0016532.
- Brennan, R. L. (2001). *Generalizability Theory*. New York: Springer-Verlag.
- Cardinet, J.; Johnson, S. & Pini, G. (2010). *Applying generalizability theory using Edu G*. Routledge.
- Eva, K. W.; Reiter, H. I.; Rosenfeld, J. & Norman, G. R. (2004). The relationship between interviewers' characteristics and ratings assigned during a multiple mini-interview. *Academic Medicine*, 79 (6), 602-609. Doi: 10.1097/00001888-200406000-00021.
- Fauber, R. L. (2006). Graduate admissions in clinical psychology: Observations on the present and thoughts on the future. *Clinical Psychology: Science and Practice*, 13, 227-234. Doi: 10.1111/j.1468-2850.2006.00029.x.
- Goho, J. & Blackman, A. (2006). The effectiveness of academic admission interviews: an exploratory meta-analysis. *Medical teacher*, 28 (4), 335-340. Doi: 10.1080/01421590600603418.
- Hanson, M. D. Kulasegaram, K. M. Woods, N. N. Fechtig, L. & Anderson, G. (2012). Modified personal interviews: resurrecting reliable personal interviews for admissions? *Academic Medicine*, 87 (10), 1330-1334. Doi: 10.1097/ACM.0b013e318267630f.
- Helmes, E. & Pachana, N. A. (2008). Value of interviews for admission to clinical training programs: Perspective of program directors. *Australian Psychologist*, 43 (4), 249-256. Doi: 10.1080/00050060802413362.
- Jones, P. E. & Forister, J. G. (2011). A comparison of behavioral and multiple mini-interview formats in physician assistant program admissions. *Journal of Physician Assistant Education*, 22 (1), 36-40. Doi: 10.1097/01367895-201122010-00006.
- Kim, K. J.; Nam, K. S. & Kwon, B. S. (2017). The utility of multiple mini-interviews: experience of a medical school. *Korean Journal of Medical Education*, 29 (1), 7-14. Doi: 10.3946/kjme.2017.48.

- Lamadrid-Figueroa, H.; Castillo-Castillo, L.; Fritz-Hernández, J. & Magaña-Valladares, L. (2012). Admissions criteria as predictors of students' academic success in master's degree programs at the national institute of public health of Mexico. *Public Health Reports*, 127 (6), 605-611. Doi: 10.1177/003335491212700612.
- Lemay, J. F.; Lockyer, J. M.; Collin, V. T. & Brownell, A. K. W. (2007). Assessment of non-cognitive traits through the admissions multiple mini-interview. *Medical Education*, 41 (6), 573-579. Doi:10.1111/j.1365-2923.2007.02767.x.
- McAndrew, R.; Ellis, J. & Valentine, R. A. (2017). Does a selection interview predict year 1 performance in dental school? *European Journal of Dental Education*, 21 (2), 108-112. Doi: 10.1111/eje.12188.
- Oranye, N. O. (2016). The Validity of Standardized Interviews Used for University Admission into Health Professional Programs: A Rasch Analysis. *SAGE Open*, 6 (3), 1-10. Doi: 10.1177/2158244016659119
- Patterson, F.; Knight, A.; Dowell, J.; Nicholson, S.; Cousans, F. & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education*, 50 (1), 36-60. Doi: 10.1111/medu.12817.
- Rosenfeld, J. M.; Reiter, H. I.; Trinh, K. & Eva, K. W. (2008). A cost efficiency comparison between the multiple mini-interview and traditional admissions interviews. *Advances in Health Sciences Education*, 13 (1), 43-58. Doi: 10.1007/s10459-006-9029-z.
- Sebok, S. S.; Luu, K. & Klinger, D. A. (2014). Psychometric properties of the multiple mini-interview used for medical admissions: findings from generalizability and Rasch analyses. *Advances in Health Sciences Education*, 19 (1), 71-84. Doi: 10.1007/s10459-013-9463-7.
- Shavelson, R.; Webb, N., & Rowley, G. (1989, June). Generalizability theory. *American Psychologist*, 44 (6), 922-932.
- Stewart, G. L.; Dustin, S. L.; Barrick, M. R. & Darnold, T. C. (2008). Exploring the handshake in employment interviews. *Journal of Applied Psychology*, 93 (5), 1139-1146. Doi: 10.1037/0021-9010.93.5.1139.
- Taylor, C. A.; Green, K. E. & Spruce, A. (2015). Evaluation of the effect of socio-economic status on performance in a multiple mini interview for admission to medical school. *Medical Teacher*, 37 (1), 59-63. Doi: 10.3109/0142159X.2014.923562.

- Till, H.; Myford, C. & Dowell, J. (2013). Improving student selection using multiple mini-interviews with multifaceted Rasch modeling. *Academic Medicine*, 88 (2), 216-223. Doi: 10.1097/ACM.0b013e31827c0c5d.
- Timer, J. E. & Clauson, M. I. (2011). The use of selective admissions tools to predict students' success in an advanced standing baccalaureate nursing program. *Nurse Education Today*, 31 (6), 601-606. Doi: 10.1016/j.
- Zaidi, N. B.; Swoboda, C.; Wang, L. L. & Manuel, R. S. (2014). Variance in attributes assessed by the multiple mini-interview. *Medical teacher*, 36 (9), 794-798. Doi: 10.3109/0142159X.2014.909587.