



Investigating the Effectiveness of General Competencies Model of Professional Managers in Public-Sector

Milad Shams Zare*

PhD Student in Public Administration, Department of Public Administration, Faculty of Accounting and Management, Allameh Tabataba'i University, Tehran, Iran.

MirAli Seyyednaghavi

Professor, Public Administration Department, Allameh Tabataba'i University, Tehran, Iran.

Vajhollah Ghorbanizadeh

Professor, Public Administration Department, Allameh Tabataba'i University, Tehran, Iran.

Mohammad afkane

Assistant Professor, Public Administration and Business Management Department, Faculty of Humanities, Shahed University, Tehran, Iran.

Received: 22/10/2023

Accepted: 10/12/2023

Abstract: The purpose of this research is to investigate the effectiveness of general competencies model of professional managers included in the "Guidelines on how to evaluate and develop general competencies of professional managers" announced by the Administrative and Recruitment Affairs Organization of Iran. For this purpose, by designing and implementing the assessment center for 1616 qualified government employees in three public organizations and analyzing the scores of eight competencies based on the assessment center scores, competency scores, and job performance scores based on the evaluation of direct managers, the required data was collected. The results of internal construct validity, external construct validity, criterion validity of competencies, and the fit of the competency model were analyzed. The results indicate that the current competency model does not have sufficient validity and effectiveness. The competencies of "responsibility" and "organizational commitment and belonging" had the lowest levels of construct validity and criterion validity. "Supervision and control" competence had the highest external construct validity and "analytical thinking and problem solving" competence had the highest criterion validity. At the end, practical, policy and research proposals were presented.

Keywords: Assessment Center, General Competency Model of Professional Managers, Construct Validity, Criterion Validity.

Corresponding Author, Email: Milad.shz@gmail.com

Original Article

DOI: 10.22034/jipas.2023.415532.1647

Print ISSN: 2676-6256

Online ISSN: 2676-606X

بررسی اثربخشی مدل شایستگی‌های عمومی مدیران حرفه‌ای دولت

میلاذ شمس زارع *

دانشجوی دکتری رشته مدیریت دولتی، گرایش مدیریت منابع انسانی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.

میرعلی سیدنقوی

استاد، گروه مدیریت دولتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.

وجه الله قربانی زاده

استاد، گروه مدیریت دولتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.

محمد افکانه

استادیار، گروه مدیریت دولتی-بازرگانی، دانشکده علوم انسانی، دانشگاه شاهد، تهران، ایران.

پذیرش: ۱۴۰۲/۰۹/۱۹

دریافت: ۱۴۰۲/۰۷/۳۰

چکیده:

هدف این پژوهش بررسی اثربخشی مدل شایستگی‌های عمومی مدیران حرفه‌ای مندرج در «دستورالعمل نحوه ارزیابی و توسعه شایستگی‌های عمومی مدیران حرفه‌ای» ابلاغی سازمان اداری و استخدامی کشور است. به این منظور، با طراحی و اجرای کانون ارزیابی برای ۱۶۱۶ نفر از کارکنان دولتی واجد شرایط در سه سازمان دولتی و تحلیل نمرات هشت شایستگی براساس نمرات کانون ارزیابی، نمرات شایستگی‌ها و نمرات عملکرد شغلی بر مبنای ارزیابی مدیران مستقیم، داده‌های موردنیاز جمع‌آوری شد. نتایج بررسی روایی سازه داخلی، روایی سازه خارجی، روایی ملاکی شایستگی‌ها و همچنین برازش مدل شایستگی، تحلیل شد. نتایج حاکی از آن است که مدل شایستگی فعلی، دارای اعتبار و اثربخشی کافی نیست. شایستگی‌های «مسئولیت‌پذیری» و «تعهد و تعلق سازمانی» کمترین میزان روایی سازه و روایی ملاکی را داشتند. شایستگی «نظارت و کنترل» بیشترین میزان روایی سازه خارجی و شایستگی «تفکر تحلیلی و حل مسئله» بیشترین میزان روایی ملاکی را داشت. در انتها، پیشنهادها، کاربردی و سیاستی و پژوهشی ارائه شدند.

واژگان کلیدی: کانون ارزیابی، مدل شایستگی عمومی مدیران حرفه‌ای، روایی سازه، روایی ملاکی.

* نویسنده مسئول: Milad.shz@gmail.com

نوع مقاله: پژوهشی

DOI: 10.22034/jipas.2023.415532.1647

شاپا چاپی: ۲۶۷۶-۶۲۵۶

شاپا الکترونیک: ۶۰۶-۲۶۷۶X



مقدمه

سازمان‌ها به‌طور گسترده‌ای از مدل‌های شایستگی در حوزه‌های مختلف مدیریت منابع انسانی استفاده می‌کنند (Campion et al., 2011; Thornton & Rupp, 2006). کانون‌های ارزیابی از روش‌هایی هستند که مبتنی بر مدل‌های شایستگی، جهت مقاصد انتخابی یا توسعه‌ای افراد در سازمان‌ها استفاده می‌شود. کانون‌های ارزیابی معمولاً شامل چندین تمرین شبیه‌سازی (به‌عنوان مثال تمرین ایفای نقش، موردکاوی و بحث گروهی) هستند که این تمرین‌ها بایستی مرتبط با شغل هدف باشند. طی این تمرین‌ها، ارزیابی‌شوندگان براساس عملکردشان در هر شایستگی، در هر تمرین مورد ارزیابی و نمره‌دهی قرار می‌گیرند (Kleinmann & Ingold, 2019). براساس گزارش روندهای جهانی ارزیابی، حدود ۵۲ درصد شرکت‌ها در دنیا در حال استفاده از این روش در فرایندهای ارزیابی خود هستند یا قصد استفاده از آن در آینده نزدیک را دارند و حدود ۶۹ درصد از ۵۰۰ شرکت برتر فورچون^۱ از کانون ارزیابی استفاده می‌کنند (SHL, 2018).

پژوهش‌های مختلفی درباره کاربردها و نقاط ضعف کانون‌های ارزیابی صورت گرفته است. به‌طور مثال، بررسی‌های لوری^۲ (۱۹۹۶) نشان داد که مرکزهای ارزیابی بدمحور برای اهداف مختلفی که شامل انتخاب، استخدام و توسعه کارکنان است کاربرد دارد و اثر تمرین باعث می‌شود ابعاد مرتبط با شغل که مورد بررسی قرار می‌گیرند، به‌خوبی سنجش نشوند. این ضعف مرکز ارزیابی بدمحور در زمینه انتخاب، باعث عدم شناسایی فرد مناسب‌تر برای شغل و در زمینه توسعه باعث بازخورد غلط در رابطه با شایستگی‌ها به داوطلبین می‌شود (Fleenor, 1996). برای حل این مسئله در روایی سازه، راه‌های مختلفی پیشنهاد شد که چندان اثربخش نبوده است (Lance et al., 2004).

لنس^۳ (۲۰۰۸) نشان داد مشکلات مربوط با مرکزهای ارزیابی بدمحور پایدار بوده و از بین نمی‌روند. سکت^۴ و لیوینز^۵ (۲۰۰۸)، بیان کردند که اجماع تقریبی وجود دارد که تمارین مرکز ارزیابی، نسبت به ساختار این مرکزها، اهمیت بیشتری دارد (Meriac et al., 2008;)

۱. Fortune: یک مجله معتبر تجاری آمریکایی است که به‌طور مرتب لیست‌های رتبه‌بندی‌شده ۵۰۰ فرد ثروتمند و همچنین رتبه‌بندی شرکت‌ها براساس درآمد را معرفی می‌کند.

2. Lowry
3. Lance
4. Sackett
5. Lievens

(Arthur et al., 2008) و نباید ابعاد را به‌عنوان آنچه در مرکز ارزیابی سنجش می‌شود، در رأس قرار داد (Jackson et al., 2012).

در ایران نیز از سال ۱۳۹۶، سازمان اداری و استخدامی کشور با ابلاغ «دستورالعمل نحوه ارزیابی و توسعه شایستگی‌های عمومی مدیران حرفه‌ای»، انتخاب و انتصاب مدیران دولتی را مبتنی بر روش کانون ارزیابی و گزارش نتایج ارزیابی فرد بر اساس مدل شایستگی‌های عمومی مدیران، حرفه‌ای کرد.

با توجه به اهمیت موضوع ارزیابی و انتخاب مدیران در هر جامعه‌ای، توسعه روش‌های ارزیابی مدیران مبتنی بر ملاحظات زمینه‌ای هر کشور و بخش مربوطه، از پیش‌نیازهای رعایت عدالت و روایی در هر روشی است. با توجه به تأکید پژوهش‌ها بر انجام مطالعات مبتنی بر بستر و فرهنگ‌های جدید درباره کانون ارزیابی (Bobko & Roth, 2013; Meiring & Buckett, 2016) تاکنون هیچ بررسی‌ای روی اثربخشی نتایج دستورالعمل مذکور صورت نگرفته است. از این رو هدف پژوهش حاضر بررسی اثربخشی مدل شایستگی‌های عمومی مدیران حرفه‌ای از طریق بررسی روایی سازه درونی و بیرونی و روایی ملاکی شایستگی‌های مذکور است. بر اساس نتایج این پژوهش می‌توان الگوی بهینه ابعاد و تمرین‌های کانون ارزیابی را پیشنهاد کرده و علاوه بر افزایش دقت و صحت شایستگی‌ها و طراحی کانون، نرخ بازگشت سرمایه روش‌های ارزیابی و اقدامات توسعه‌ای را از طریق کاهش هزینه‌ها بهبود بخشید.

معرفی کانون ارزیابی و مؤلفه‌های آن

روش کانون ارزیابی، مشخصه‌های منحصر به فردی نسبت به دیگر روش‌های ارزیابی دارد که این مشخصه‌ها در تعریفی که در استاندارد بین‌المللی کانون ارزیابی توسط برترین پژوهشگران این حوزه تدوین شده است به شکل زیر تعریف شده است:

یک ارزشیابی استاندارد از رفتارها براساس ورودی‌های متعدد. هر کانون ارزیابی از مؤلفه‌های متعددی تشکیل شده است که شامل تمارین شبیه‌سازی رفتاری هستند که در آن‌ها چندین ارزیاب آموزش‌دیده به مشاهده و ثبت رفتارها، طبقه‌بندی آن‌ها بر اساس سازه‌های رفتاری مدنظر کانون و (به‌صورت انفرادی یا جمعی) به نمره‌دهی (تک‌تک یا مجموع) رفتارها می‌پردازند. سپس براساس جلسه اجماع یا تجمیع آماری، نمرات ارزیابی مشخص می‌شوند که این نمرات نشان‌دهنده جایگاه ارزیابی‌شونده در سازه رفتاری و یا نمره‌ی نهایی تجمیعی ارزیابی هستند.

هر کانون ارزیابی هشت مؤلفه ضروری دارد:

- یک تحلیل از شغل موردنظر و زمینه آن، برای تعیین سازه‌های مرکزی (یعنی ابعاد، شایستگی‌ها، وظایف) مورد ارزیابی ایجاد می‌شود.
 - فنون ارزیابی متعددی استفاده می‌شود.
 - شبیه‌سازی‌های مربوط به شغل (تمرین‌های شبیه‌سازی) برای استخراج رفتار مربوط به سازه‌های مرکزی استفاده می‌شود.
 - در حین ارزیابی، رفتارهای آشکار واقعی داوطلبان مشاهده، طبقه‌بندی و ارزیابی می‌شود.
 - از ارزیابی‌های متعدد استفاده می‌شود.
 - ارزیاب‌ها آموزش می‌بینند.
 - یک روش نظام‌مند برای ضبط و ثبت کردن رفتار استفاده می‌شود.
 - داده‌ها از طریق ارزیاب‌ها و شبیه‌سازی‌ها و احتمالاً از طریق سازه‌های مرکزی که برای ایجاد امتیازهای کلی ارزیابی می‌شوند، یکپارچه‌سازی می‌شوند.
- درحالی‌که هر یک از این ۸ مؤلفه می‌تواند مؤلفه دیگر روش‌های ارزیابی نیز باشد، ویژگی متمایز روش کانون ارزیابی این است که تمام این ویژگی‌ها را داراست (Guidelines and Ethical Considerations for Assessment Center Operations, 2015)

انواع کانون‌های ارزیابی و کاربردهای آن

انواع مختلفی از کانون وجود دارد. از کانون‌های سنتی^۱ تا کانون‌های یادگیری^۲، کانون‌های مشارکتی^۳ و حتی کانون‌های کارویژه‌ای^۴. به‌رحال، تمایز روشی بین کانونی که برای انتخاب افراد به کار گرفته می‌شود (که به آن کانون ارزیابی می‌گوییم) با کانونی که برای مقاصد توسعه‌ای به کار گرفته می‌شود (که به آن کانون توسعه گفته می‌شود) وجود دارد. علاوه بر این نوع دیگری از کانون به نام کانون‌های تشخیصی^۵ نیز وجود دارد.

کارکردهای مختلفی برای کانون‌های ارزیابی برشمرده می‌شوند که مهم‌ترین آن‌ها عبارت‌اند از: انتخاب و استخدام، جانشین‌پروری، شناسایی نقاط قوت و ضعف کارکنان، آموزش و توسعه کارکنان (Thornton & Rupp, 2015)

1. Traditional ACs
2. Learning ACs
3. Collaborative ACs
4. Functional ACs
5. Diagnostic ACs

ابعاد، شایستگی‌ها و سازه‌های رفتاری در کانون‌های ارزیابی

در ادبیات کانون ارزیابی با دو اصطلاح که به‌طور جایگزین به‌عنوان سازه‌های مرکزی جهت ارزیابی مورد استفاده قرار می‌گیرند روبرو هستیم: ابعاد رفتاری و شایستگی‌های رفتاری.

ابعاد رفتاری

بایهام^۱ (۱۹۷۰) یک بعد رفتاری را به‌عنوان یک دسته از رفتارهای قابل مشاهده تعریف کرده است که منطقاً همگن هستند. مشخصه تعریف یک بعد این است که آن را می‌توان به‌صورت رفتاری تعریف کرد. این مورد در زمینه کانون ارزیابی، کلیدی است، زیرا این رفتارهای واقعی هستند که حین اجرای تمرین‌ها مشاهده، طبقه‌بندی و به‌عنوان یک دسته امتیازبندی می‌شوند. سازمان‌ها در به‌کارگیری راهبرد مدیریت استعداد هدفمند ممکن است از تعاریف اختصاصی^۲ برای ابعاد استفاده کنند. با وجود این تنوع، ابعاد مشترکی وجود دارد که در بسیاری از کانون‌های ارزیابی استفاده می‌شود، به‌ویژه در سطوح مدیریتی که در بین شغل‌ها، سازمان‌ها و فرهنگ‌های مختلف، مجموعه‌ای از ویژگی‌های پایدارتر مورد نیاز است. به‌عنوان مثال آرتور و همکاران (۲۰۰۳) از ۳۴ پژوهش تجربی کانون ارزیابی، مجموعه‌ای از ۱۶۸ بعد را شناسایی کردند. این ابعاد به‌صورت نظام‌مند به هفت بعد تقسیم می‌شوند، که در جدول زیر فهرست شده‌اند. عناوین مرسوم‌تر مورد استفاده برای ابعاد در پرانتز آورده شده است. ووهر و همکاران (۲۰۱۴)، یک طبقه‌بندی آشیانه‌ای از ابعاد را تأیید کردند که در آن هفت بعدی را که آرتور و همکاران مشخص کرده بودند، به سه دسته کاهش یافت: ابعاد اجرایی^۳، انگیزشی^۴، و ارتباطی^۵ که از پیشینه پژوهش رهبری مشتق شده‌اند. ارتباط این سه بعد گسترده توسط هافمن و همکاران (۲۰۱۱) تأیید شده بودند.

جدول ۱: ابعاد کانون ارزیابی: سه طبقه، هفت بعد با تعاریف، عناوین جایگزین رایج و درصد استفاده در یک نظرسنجی بین‌المللی

اجرائی

سازماندهی و برنامه‌ریزی: میزانی که یک فرد به‌طور نظام‌مند، کار و منابع خود و دیگران را برای انجام اثربخش وظیفه، سازماندهی می‌کند؛ و میزانی که فرد اقتضات آینده را لحاظ می‌کند و برای آن آماده می‌شود (کنترل، تفویض اختیار، توسعه زیردستان، برنامه‌ریزی و سازماندهی) (۵۴٪)

1. Byham
2. firm-specific
3. administrative
4. drive
5. relational

حل مسئله: حدی که فرد به گردآوری اطلاعات، درک اطلاعات فنی مرتبط، تحلیل مؤثر داده‌ها و اطلاعات، خلق راهکارهای مؤثر و انتخاب اقدامات مؤثر برای مسائل می‌پردازد. (تحلیل، خلاقیت، قاطعیت، قضاوت، یادگیری عملی، محدوده علایق، دانش فنی / حرفه‌ای) (۷۷٪)

انگیزه

انگیزه: میزانی که فرد سطح بالایی از فعالیت را از خود نشان می‌دهد و حفظ می‌کند، استانداردهای عملکردی بالایی را تعیین می‌کند و در دستیابی به موفقیت خود پافشاری می‌کند، و تمایل به پیشرفت به سطوح بالاتر را نشان می‌دهد (جاه‌طلبی شغلی، انرژی، ابتکار، انگیزه شغلی، سرسختی، استانداردهای شغلی) (۶۷٪)

ارتباطی

ارتباط: میزانی که فرد اطلاعات کتبی و شفاهی را انتقال می‌دهد و به سؤالات و چالش‌ها پاسخ می‌دهد (ارتباط شفاهی، ارائه شفاهی) (۷۰٪)

توجه / آگاهی: میزانی که اقدامات فرد، ملاحظه افراد و نیازهایشان را می‌کند و همچنین آگاهی از تأثیر و پیامدهای تصمیم‌گیری‌هایش بر داخل و خارج سازمان (آگاهی برون‌سازمانی، حساسیت برون‌سازمانی، آگاهی سازمانی، حساسیت سازمانی، شناخت نیازهای ایمنی، حساسیت) (۴۲٪)

تأثیر بر دیگران: میزانی که فرد دیگران را متقاعد می‌کند که برای رسیدن به نتایج مطلوب، کاری را انجام دهند یا با یک دیدگاه موافقت کنند، و اقداماتی انجام می‌دهد که عمدتاً تحت تأثیر نظرات و عقاید قاطع خودش است، نه تأثیر نظرات دیگران. (استقلال، صداقت، رهبری) (۷۰٪)

تحمل استرس / عدم اطمینان: میزانی که فرد در موقعیت‌های گوناگون و انواع فشارها، مخالفت‌ها و ناامیدی، اثربخشی خود را حفظ می‌کند (سازگاری، تاب‌آوری، ریسک‌پذیری، تحمل استرس) (۳۹٪)

جدول بالا تعاریف هفت بعد ارائه‌شده توسط آرتور و همکاران (۲۰۰۳) را نیز ارائه می‌دهد. پیمایش‌های تجارب کانون ارزیابی طی سال‌های اخیر نشان داده است که برنامه‌های کانون ارزیابی عموماً از ابعاد به‌عنوان سازه‌های مرکزی خود استفاده می‌کنند، اگر چه اغلب در سازمان‌ها به‌عنوان شایستگی شناخته می‌شوند. برای مثال هیوز و همکاران (۲۰۱۲) پاسخ‌های ۵۴۳ نماینده کانون ارزیابی از ۵۳ کشور درباره تعداد زیادی از ویژگی‌های کانون ارزیابی را بررسی کردند. اگر چه این نظرسنجی، سازه‌های مرکزی را شایستگی نامید، در واقع یک فهرست اصلاح‌شده از ابعاد آرتور و همکاران (۲۰۰۳) بود. اعداد داخل پرانتز در جدول ۱، درصد پاسخ‌دهندگانی است که نشان دادند در برنامه آن‌ها یک یا چند بعد از آن طبقه ارزیابی شده بود. رایج‌ترین ابعاد، ارتباط با افراد و حل مسئله بود. میانگین تعداد ابعاد ارزیابی شده ۶/۷ بود، اما ۵۳ درصد گزارش دادند که در مجموع، هشت بعد یا بیشتر را ارزیابی کردند. این تعداد احتمالاً بیش از حد بزرگ است، زیرا مانع تمایز دقیق بین ابعاد می‌شود. از سوی دیگر، اکثر پاسخ‌دهندگان گزارش می‌کنند که در هر تمرین تنها سه تا شش بعد را ارزیابی می‌کنند.

شایستگی‌های رفتاری

شایستگی یک اصطلاح عمومی است که برای اشاره به انواع مختلفی از سازه‌های مرکزی استفاده شده است (Putka, 2013). مرجع استاندارد ریشه اصطلاح "شایستگی" مک‌کلند^۱ (۱۹۷۳) است، اما این مرجع تعریفی از این اصطلاح ارائه نمی‌کند. او پیشنهاد می‌کند که یک آزمون باید مشخصات عمومی لازم برای انواع برون‌داد زندگی مانند مهارت‌های ارتباطی، صبر و توسعه خود را ارزیابی کند. در زمینه‌های دیگر، شایستگی‌ها با ابعاد رفتاری متفاوت نیستند. تنها این مسئله وجود دارد که سازمان تصمیم می‌گیرد از برجسب شایستگی استفاده کند تا سازه‌های مرکزی برای افراد، کاربرپسندتر به نظر برسند و این که مدیران به چارچوب‌های شایستگی علاقه بیشتری دارند. در مواردی، اصطلاح "شایستگی" نشان‌دهنده ترکیبی از دانش، مهارت و استعدادها (KSA) یا ابعاد است. به‌عنوان مثال، شایستگی "رهبری" ممکن است ترکیبی از ابعاد رفتاری مانند هدف‌گذاری، انگیزش و متقاعدسازی باشد خاص‌تر (Thornton & Rupp, 2015).

بارترام^۲ (۲۰۰۵) شواهدی ارائه کرده است که هشت شایستگی فراگیر را نشان می‌دهد، که «هشت شایستگی برتر^۳» نامیده می‌شود. براساس بررسی فراتحلیلی، او ۱۱۲ شایستگی را ابتدا در ۲۰ دسته میانی و سپس در ۸ دسته اصلی تقسیم کرده است.

اثربخشی کانون‌های ارزیابی بر مبنای روایی ملاکی

مطالعات بی‌شماری در زمینه روایی ملاکی نمرات نهایی ارزیابی (OAR)^۴ وجود دارد. جدول ۲ خلاصه‌ای از نتایج فراتحلیل‌های موجود که این موضوع را بررسی می‌کنند، نشان می‌دهد.

در مطالعات روایی، از متغیرهای ملاک متفاوتی مانند رضایت شغلی، ارزیابی مدیر، ارزیابی عملکرد توسط همکاران، حقوق و دستمزد، مرتبه سازمانی و ... استفاده شده است و نتایج این مطالعات سازگاری مناسبی باهم دارد (Zysberg, 2012). همچنین بین نمرات مرکز ارزیابی و احتمال ارتقای افراد به سمت‌های بالاتر در آینده، رابطه به‌دست آمده است (Eisenburg, 2001).

گاگلر و همکاران (۱۹۸۷)، اولین بررسی در مقیاس بزرگ و فراتحلیل نمرات نهایی کانون‌ها (OAR) را ارائه دادند. نتایج بررسی آن‌ها نشان داد که میان ۲۹ نمونه، ارتباط پیش‌بینی‌شده بین OAR و عملکرد ۰/۳۷ بود (همبستگی تصحیح‌نشده ۰/۲۵ بود). اگرچه گاگلر و همکاران (۱۹۸۷)

1. McClelland
2. Bartram
3. the great eight
4. Overall Assessment Rating

دریافتند که OAR در بعضی زمینه‌ها و بسته به طراحی مرکز ارزیابی، ارتباط قوی‌تری با عملکرد دارد، اما هرملین و همکاران (۲۰۰۷) چنین اثراتی را تصدیق نکردند.

جدول ۲: خلاصه روایی ملاکی کانون ارزیابی از بررسی‌های فراتحلیل

مطالعه	نوع امتیاز کانون ارزیابی	تخمین اعتبار
گاگلر و همکاران (1987)	OAR	۰,۳۷
هرملین و همکاران (2007)	OAR	۰,۲۷
آرتور و همکاران (2003)	اجعه	۰,۴۷
مناهان و همکاران (2013)	نتنینه	۰,۲۷
سکت و همکاران (2022)	OAR	۰,۲۹

همچنین برخی مطالعات، روایی ملاکی مربوط به نمرات هر شایستگی در کانون‌های ارزیابی را بررسی کردند. آرتور و همکاران (۲۰۰۳) یک فراتحلیل از رابطه بین هفت بعد مرکز ارزیابی و نمرات سرپرست از عملکرد شغلی را بر مبنای ۳۴ مطالعه ارائه دادند. ابتدا نتایج آن‌ها نشان داد که روابط با عملکرد شغلی، به ابعاد مرکز ارزیابی وابسته است. آن‌ها یک تحلیل چند رگرسیونی برای تعیین این‌که آیا ابعاد انفرادی، اثرات منحصربه‌فردی را در متغیرهای معیار بیان می‌کنند، راهبری کردند. نتایج آن‌ها نشان داد که ابعاد، همبستگی چندگانه برابر با ۰/۴۵ را حاصل کردند (به این معنا که ابعاد ۲۱ درصد از واریانس را شرح می‌دهند). به‌طور قابل‌توجهی، زمانی که ابعاد متعددی به‌عنوان ورودی استفاده شدند، روایی ملاکی مرکزهای ارزیابی بیشتر از زمان به‌کارگیری یک OAR مختصر از همبستگی تصحیح‌شده برابر با ۰/۳۷ بود (یا ۱۴ درصد از واریانس).

براساس فراتحلیل‌ها و مطالعات تجمیعی انجام‌شده، نتایج چهار پژوهش معتبر بین‌المللی در جدول ۳، میزان اثربخشی و اعتبار پیش‌بین هر شایستگی را در کانون‌های ارزیابی نشان می‌دهد.

جدول ۳: اعتبار پیش‌بین هر شایستگی بر اساس فراتحلیل‌ها

شایستگی‌ها / تحقیقات	آرتور و همکاران، ۲۰۰۳	مریاک و همکاران، ۲۰۰۸	کانلی و همکاران، ۲۰۰۸ ۳۱۰۰ = N	لیونز و همکاران، ۲۰۰۹
ارتباطات	۰,۳۳ / ۰,۲۶	۰,۲۷ / ۰,۲۵	۰,۱۰	X
توجه / آگاهی	۰,۲۵ / ۰,۲۰	۰,۲۴ / ۰,۲۲	۰,۱۶	۰,۲۱
انگیزه	۰,۳۱ / ۰,۲۴	۰,۱۶ / ۰,۱۵	۰,۲۲	X
تأثیر بر دیگران	۰,۳۸ / ۰,۳۰	۰,۳۱ / ۰,۲۹	۰,۲۴	۰,۰۸

سازماندهی و برنامه‌ریزی	۳۷/۲۹.	۳۵/۳۳.	X	۰,۱۵
حل مسئله	۳۹/۳۰.	۳۲/۳۱.	۰,۲۵	۰,۱۸
تحمل استرس	X	۱۸/۱۶.	X	X

اثربخشی کانون‌های ارزیابی بر مبنای روایی سازه

مشابه انجام تحلیل عاملی در ساخت آزمون‌های روان‌سنجی برای تعیین تعداد و ماهیت سازه‌های آزمون، پژوهشگران کانون ارزیابی، رویکردهای مشابهی را برای بررسی تعداد و ماهیت سازه‌های سنجیده‌شده توسط کانون‌های ارزیابی به کار برده‌اند. چنین تحلیل‌هایی مبنای طراحی و ارزیابی مقیاس و آزمون هستند، زیرا آن‌ها نوعی از روایی سازه هستند و به پژوهشگران این اعتماد را می‌دهند که گویه‌ها سازه‌هایی را می‌سنجند که دقیقاً برای سنجش همان سازه در نظر گرفته شده‌اند.

طی ۴۰ سال گذشته پژوهش‌های کانون ارزیابی از تحلیل‌های چندصفت- چندروش (MTMM) به‌عنوان یک چارچوب برای تحلیل سازه درونی نمرات کانون ارزیابی بسیار استفاده کرده است. این چارچوب، زمانی که سازه‌های متعددی با استفاده از روش‌های مختلف سنجیده می‌شوند (همان‌طور که در مورد کانون‌های ارزیابی است) سودمند است. مطابق با این چارچوب، زمانی که یک سازه مشابه با استفاده از روش‌های مختلف سنجیده می‌شود، انتظار می‌رود که امتیازات سازه بسیار با یکدیگر همبستگی داشته باشند یا همگرا باشند. این موضوع به‌عنوان اعتبار یا روایی همگرا شناخته می‌شود. از سوی دیگر زمانی که سازه‌های مختلف با استفاده از روش‌های مختلف سنجیده می‌شوند، امتیازات نباید همپوشانی داشته باشند. تمایز بین سازه‌های مختلف به‌عنوان اعتبار یا روایی واگرا شناخته می‌شود (Thornton & Rupp, 2015).

نتایج مطالعات مشخصاً حاکی از روایی سازه پایین مرکزهای ارزیابی براساس توانایی بازتاب دادن شایستگی‌هایی که برای ارزیابی آن‌ها طراحی شده‌اند، است (Lance et al., 2004). به‌صورت کلی، در این پژوهش‌ها، روایی همگرا و واگرایی اندکی برای این نوع از مرکزها گزارش شده که در این حالت، نتایج اندازه‌گیری صفات، غیرمطمئن است. درواقع، نمرات ارزیابان در این مرکز ارزیابی اغلب، واریانس ناشی از تمرین‌ها را به‌جای واریانس ناشی از ابعاد، بازتاب می‌دهد (این مسئله، اثر تمرین نام دارد) (Jackson et al., 2005) و عامل تمرین، منبع اصلی واریانس

نمرات است (Jackson et al., 2012; Putka & Hoffman, 2013). به عبارت دیگر، در این مرکزها اثر تمرین باعث می‌شود ابعاد متفاوت در یک تمرین، همبستگی بالاتری از همبستگی ابعاد یکسان در تمرین‌های مختلف داشته باشند (Lance, 2008).

در برخی از مطالعات نیز بیان شده که اثر تمرین روی نمرات ابعاد رفتاری، به جای تمرکز بر روی عدم اثر ابعاد رفتاری بر رفتارهای داوطلب، بیشتر به اثر مصنوعی تمرین نسبت داده شود (Monahan et al., 2013; Putka & Hoffman, 2013). بنابراین برای خیلی از پژوهشگران، چشم‌انداز سنجش ابعاد ناامیدکننده نیست و تمرکز آن‌ها بر نحوه نمره‌دهی به ابعاد رفتاری است. دو رویکرد در این جا مطرح می‌شود. رویکرد تعاملی در توسعه تمرین‌های مرکز ارزیابی و رویکرد تجمیع نمره ابعاد در کل تمرین‌ها. پژوهشگران بیان کرده‌اند که علت عدم همخوانی مرکزهای ارزیابی با منطق MTMM این است که تمرین‌ها به شکل همسان و موازی ابعاد رفتاری را نمی‌سنجند (Lee et al., 2017). گستره وسیع‌تر موقعیت‌ها در تمرین، روایی پیش‌بین بالاتر مرکزهای ارزیابی را به همراه دارد. ولی این مسئله سبب کاهش روایی همگرایی بین تمرین می‌شود (Speer et al, 2014). تمرین مختلف، براساس نظریه فعال‌سازی صفات، نشانه‌های رفتاری خاصی را از ابعاد موردنظر در تمرین متفاوت استخراج می‌کند. در نتیجه ابعاد رفتاری در هر تمرین به شکل متفاوتی ظاهر می‌گردند. مثلاً ممکن است شایستگی تفکر تحلیلی در ایفای نقش با همین شایستگی در تمرین کارتابل (یک نوع تمرین نوشتاری) به شکل معناداری متفاوت باشد. در این صورت، بهتر است نمره ابعاد رفتاری به‌عنوان یک مهارت خاص تمرین، سنجش و تفسیر شوند. با اتخاذ این دیدگاه، برای مثال، به جای تمرکز بر نمره کلی ابعاد در داوطلبان (مثلاً ارتباط مؤثر)، رویکرد "اگر-آن‌گاه" مینا می‌شود و کیفیت تغییرات ابعاد رفتاری در تمرین‌ها بررسی شود (Thornton & Rupp, 2015).

رویکرد دوم، رویکرد تجمیع نمره ابعاد رفتاری در کل تمرین‌ها است. بر این اساس بیان می‌شود که نتایج عدم روایی سازه، ممکن است آن‌چنان مشکل‌ساز نباشد، زیرا بیشتر پژوهش‌ها در این رابطه، بر نمره هر بعد رفتاری در هر تمرین متمرکز است اما از نظر عملیاتی، بیشتر مرکزهای ارزیابی از نمره نهایی بعد رفتاری استفاده می‌کنند. لذا پیشنهاد می‌شود از نمره کلی بعد رفتاری برای غلبه بر اثر تمرین، استفاده شود. در این روش تجمع نمره‌های بعد رفتاری، منجر به غلبه بر واریانس متأثر از تمرین‌ها شده و واریانس بعد رفتاری را برجسته می‌کند. اتخاذ این رویکرد، تفاوت‌های عمده در ساختار داخلی مرکزهای ارزیابی نسبت به رویکرد تعاملی ایجاد خواهد کرد. مثلاً در رویکرد تعاملی بر پراکندگی و تغییرات نمرات ابعاد بین تمرین تمرکز می‌شود ولی در این رویکرد به حذف این تفاوت‌ها توجه می‌شود. مضافاً، رویکرد تعاملی گرایش دارد بعد رفتاری را در

تعداد محدودی تمرین، مورد بررسی قرار دهد اما از نظر دیدگاه تجمعی بهتر است ابعاد رفتاری در تعداد زیادی تمرین بررسی گردد. همچنین برای ارائه یک بازخورد مناسب به افراد، ظاهراً رویکرد تعاملی گزینه بهتری است. اما رویکرد تجمعی برای مواقعی که ابعاد رفتاری موردنظر، دارای معنای یکسان و موازی در تمامی تمارین هستند بهتر است (Lee et al., 2017). در این رابطه لی و همکاران (۲۰۱۷)، با مطالعه یک نمونه بزرگ اذعان کردند که هرچند بعضی تمارین در مرکز ارزیابی ممکن است موجب تحریک بروز ابعاد رفتاری خاصی شوند ولی یافته‌های آن‌ها نشان داد که ابعاد رفتاری ثابت خوبی در تمارین مختلف مرکزهای ارزیابی دارند.

علاوه بر مسئله روایی سازه، انتقاد دیگری که بر این مرکزهای ارزیابی وارد است این است که باوجود گذشت بیش از ۶۰ سال پژوهش روی مرکز ارزیابی بعدمحور، هیچ نظریه‌ای از ابعاد مرکز ارزیابی پشتیبانی نمی‌کند. درعوض، معمولاً ابعاد رفتاری مبتنی بر اهداف سازمانی و با تجزیه و تحلیل شغل یا مدل‌سازی شایستگی بدون توجه کافی به رشد مداوم نظریه‌ها در زمینه ابزار و دسته‌بندی مهارت‌های رفتاری رهبری، انجام می‌شود. این مسئله باعث شده است که اجرای مرکز ارزیابی در مشاغل مختلف از نظر قابلیت تعمیم با مشکل مواجه گردد. به عبارت دیگر، تنوع در ابعاد رفتاری مختلف در مرکزهای ارزیابی مختلف، قابلیت مقایسه این مرکزها را کاهش می‌دهد. از طرف دیگر، پراکندگی در این ابعاد، از علل ابهام در مورد روایی سازه این مرکزها است. در مرکزهای ارزیابی، ابعاد رفتاری براساس محتوای شغل استخراج می‌گردند. این مسئله از دو جهت مشکل ساز شده است. اول آن‌که، ابعاد رفتاری و شایستگی‌های متعددی برای مشاغل مدیریتی مختلف شناسایی و احصا شده است که قابلیت تعمیم نتایج را دشوار می‌کند (Meriac et al., 2014). دوم آن‌که، توجه زیاد به اهداف عملکردی و تقاضاها در جهت تدوین شایستگی‌ها، بدون در نظر گرفتن نظریه‌ها و سنجش عملکرد، همپوشانی مفهومی محتوا را با مشکل مواجه کرده است. مقدار زیادی از این همپوشانی، رفتارهای مربوط به ابعاد رفتاری در مرکز ارزیابی است. لذا بین ابعاد یک مرکز ارزیابی، همپوشانی محتوا وجود دارد. با توجه به این همپوشانی، عدم تأیید روایی سازه توسط تحلیل ابعاد و روش MTMM، بعید نیست (Meriac et al., 2014).

جمع‌بندی بخش ادبیات نشان می‌دهد که عمده تمرکز پیشینه این حوزه جهت بررسی اثربخشی کانون‌های ارزیابی، بر بررسی روایی ملاکی کانون‌های ارزیابی با متغیرهای خارجی مانند نتایج ارزیابی عملکرد افراد، و همچنین بررسی روایی سازه شایستگی‌های کانون ارزیابی جهت تعیین میزان بهینه بودن ساختار داخلی کانون‌های ارزیابی، است.

روش‌شناسی پژوهش

پژوهش حاضر از نوع کمی و طرح پژوهش، طرح همبستگی مبتنی بر تحلیل عاملی تأییدی است.

نمونه و مشارکت‌کنندگان

داده‌ها براساس نتایج اجرای کانون ارزیابی و نمونه‌گیری در دسترس، در سه سازمان مختلف دولتی گردآوری شده است. جامعه آماری شامل تمام افراد واجد شرایط براساس دستورالعمل نحوه ارزیابی و توسعه شایستگی‌های عمومی مدیران حرفه‌ای است. حجم نمونه برابر با ۱۶۱۶ نفر از کارمندانی است که بین سال‌های ۱۳۹۸ تا ۱۴۰۲ در کانون ارزیابی این سه سازمان شرکت کرده‌اند.

روش گردآوری داده‌ها

در مرحله اول پس از اعلام نیاز سازمان‌ها به طراحی و اجرای کانون ارزیابی، مبتنی بر شایستگی‌های دستورالعمل، براساس اصول طراحی و اصول استاندارد کانون ارزیابی انجمن روانشناسی کانون بریتانیا (BPS) طراحی شده (BPS, 2015). شرح مشاغل هدف و الزامات سازمانی با خبرگان واحد منابع انسانی سازمان‌ها مورد بررسی قرار گرفت. بر این اساس، تمرین‌های شبیه‌سازی به‌عنوان ابزار اصلی گردآوری داده‌ها از معتبرترین منبع بین‌المللی (A PSI & DC) تهیه گردید. تمرین‌ها با نظر خبرگان واحد منابع انسانی سازمان‌ها، خبرگان شغلی از هر خانواده شغلی و پژوهشگران به‌عنوان طراحان کانون مورد بررسی قرار گرفت و براساس اصول استاندارد و رعایت میزان شبیه‌سازی (فیدلیتی)^۱ مناسب، از نظر روایی صوری و محتوایی مورد تأیید قرار گرفت. همچنین براساس نظریه فعال‌سازی صفت^۲ که در طراحی تمرین مورد استفاده قرار می‌گیرد، هر تمرین از نظر ترکیب متعادل محرک‌ها و شدت موقعیت و میزان دشواری مورد بررسی و تأیید قرار گرفت. در نهایت ماتریس تمرین-شایستگی تدوین شد. تمرین‌های شبیه‌سازی و شایستگی‌های مورد ارزیابی در هر تمرین طبق جدول زیر تنظیم شدند.

1. fidelity

2. Trait activation theory

جدول ۲: ماتریس تمرین - شایستگی در کانون ارزیابی طراحی شده

	ایفای نقش	بحث گروهی بدون تخصیص نقش	بحث گروهی با تخصیص نقش	مطالعه موردی	کارتابل
۱				*	*
۲				*	*
۳	*	*	*		
۴				*	*
۵	*	*			
۶	*			*	
۷	*				*
۸			*	*	

ارزیابان نیز براساس اصول تربیت ارزیاب، آموزش چارچوب مرجع (FOR) به‌عنوان اثربخش‌ترین روش آموزش ارزیابان (Roch et al., 2012) طی یک کارگاه ۲۴ ساعته آموزش دیده و دارای سابقه ارزیابی حداقل ۲۰۰ نفر بودند. همچنین ارزیاب‌ها به‌خوبی بر تمرین‌ها مسلط بوده و هماهنگ شده بودند و مسلط بر کل مراحل چهارگانه ارزیابی یعنی مشاهده، ثبت، طبقه‌بندی و ارزش‌یابی بودند، و با تعاریف شایستگی‌ها و تمرین‌ها نیز آشنا بوده و مدارک تحصیلی آن‌ها نیز دکترای مدیریت یا روانشناسی بوده است.

فرایند اجرای کانون ارزیابی به این صورت بوده است: (۱) ارزیابی‌شوندگان در ابتدای فرایند در یک جلسه توجیهی شرکت می‌کنند. (۲) ارزیابی‌شوندگان طی کانون یک‌و نیم روزه، در هر پنج تمرین شرکت می‌کنند و توسط ارزیابان مختلف مورد ارزیابی قرار می‌گیرند به‌طوری‌که یک ارزیاب مشخص نمی‌تواند یک ارزیابی‌شونده مشخص را در بیش از یک تمرین ارزیابی کند تا دچار سوگیری نشود. علاوه بر این، ارزیابان در حین اجرای تمرین، تمام صحبت‌ها و عملکرد ارزیابی‌شونده را با دقت کامل یادداشت می‌کنند (۳) براساس ماتریس شایستگی - تمرین، نمرات هر ارزیابی‌شونده در هر تمرین بر اساس مقیاس ۵ نمره‌ای (۱= کاملاً پایین‌تر از حد انتظار، ۵= کاملاً بالاتر از حد انتظار) داده می‌شود. به این صورت که هر تمرین دارای تعدادی شایستگی است و هر شایستگی به‌ازای هر تمرین دارای شاخص‌های رفتاری (مصادیق، نشانگر) مشخصی است.

ضمناً ارزیابان به‌ازای تیک زدن هر شاخص رفتاری بایستی یک شاهد رفتاری در قالب یک بند دقیقاً از عملکرد فرد بنویسند تا دلیل قابل دفاعی برای تخصیص هر شاخص داشته باشد. (۴) ارزیابان در جلسه یکپارچه‌سازی (جمع‌بندی) حاضر می‌شوند و نمره نهایی هر فرد در هر شایستگی پس از اجماع تعیین می‌گردد.

جهت بررسی روایی سازه بیرونی و روایی ملاکی نمرات کانون ارزیابی، از مدیران مستقیم افراد ارزیابی‌شونده درخواست شد تا دو پرسشنامه شامل پرسشنامه شایستگی‌های فرد و پرسشنامه عملکرد شغلی را درباره این افراد تکمیل کنند. در پرسشنامه شایستگی‌ها، به‌ازای هر شایستگی که در کانون مورد ارزیابی قرار گرفته است؛ یک گویه کلی درباره وضعیت فرد در آن شایستگی و حدوداً ۵ گویه دیگر که از جنس شاخص‌های رفتاری در آن شایستگی و متناظر با شاخص‌های مشابه در کانون ارزیابی بودند مطرح گردید (ضریب آلفای کرونباخ بین ۰,۷۵ تا ۰,۹۰). پرسشنامه عملکرد شغلی شامل ۲۵ سؤال بود که براساس پرسشنامه استاندارد بات و همکاران^۱ (۲۰۰۳) شامل دو سازه عملکرد وظیفه‌ای و عملکرد زمینه‌ای (رفتار شهروندی سازمانی) در یک مقیاس ۵ نمره‌ای تنظیم شده بود. پایایی پرسشنامه معادل ۰,۹۱ بود که میزان قابل قبولی است. روش تحلیل داده‌ها

جهت بررسی روایی سازه شایستگی‌ها و برازش آن‌ها از روش تحلیل عاملی تأییدی با عامل صفت‌وروش (CFA-TM^۲) که ترکیبی از روش کلاسیک چندصفت- چندروش (MTMM)^۳ و تحلیل عاملی تأییدی است استفاده شد (Bong & Hong, 2010).

در روش MTMM، شایستگی‌ها به‌عنوان صفت و تمرین‌های کانون ارزیابی به‌عنوان روش در نظر گرفته می‌شوند و جزو روش‌های پرکاربرد در مطالعات روایی سازه کانون‌های ارزیابی است (Lance et al., 2000). در روش MTMM، روایی همگرا از طریق میزان همبستگی میان نمرات یک صفت در روش‌های مختلف مشخص می‌گردد که از آن با عنوان همبستگی تک‌صفت- چند روش^۴ یاد می‌شود. همچنین برای حصول روایی واگرا شروط زیر لازم است: همبستگی‌های میان نمرات یک صفت (شایستگی) در روش‌های (تمرین‌های) مختلف باید بزرگ‌تر از همبستگی‌های متناظر چندصفت- چندروش باشد و به‌عبارتی با شایستگی‌های دیگر همبستگی معنی‌دار نداشته باشد. همچنین شرط دوم که سخت‌گیرانه‌تر است این‌که همبستگی‌های

1 .Bott, Svyantek, Goodman, and Bernal

2. confirmatory factor analysis model with trait and method factors

3 nMulti -Trait Multi-Method

i .the monotrait-heteromethod correlations; MTHM

تک‌صفت- چندروش باید از همبستگی‌های چندصفت- تک‌روش^۱ بزرگ‌تر باشد (Campbell & Fiske, 1959).

جهت محاسبه و بررسی برآورده شدن شروط روایی سازه در روش MTMM، می‌توان از شاخص‌های برازش روش تحلیل عاملی استفاده کرد. یکی از اهداف اولیه انجام مدل اندازه‌گیری تحلیل عاملی تأییدی (CFA)، ارزیابی روایی سازه است. از این‌رو، روایی همگرا از طریق بررسی معناداری بار عاملی نشانگرها و روایی واگرا از طریق بررسی میانگین واریانس استخراجی که با عنوان میانگین واریانس استخراج‌شده (AVE) برچسب‌گذاری می‌شود، مشخص می‌شوند (Hair et al., 2010; Jr et al., 2018).

جهت بررسی روایی ملاکی کانون ارزیابی، از همبستگی میان نمرات نهایی هر شایستگی با نمرات ارزیابی عملکرد افراد که توسط مدیران آن‌ها تکمیل شده است استفاده شد.

یافته‌ها

در جدول زیر براساس نمرات نهایی هر شایستگی که براساس میانگین نمره هر شایستگی در هر تمرین به دست آمده است؛ ابتدا میانگین، انحراف معیار و همبستگی نمرات شایستگی‌ها در کانون ارزیابی، ذکر شده است. در ادامه، جهت بررسی روایی سازه خارجی، براساس نمرات مدیران مستقیم به پرسشنامه، میانگین، انحراف معیار و همبستگی نمرات شایستگی‌ها ذکر شده است و همچنین همبستگی این نمرات با نمرات در کانون که نشان‌دهنده روایی سازه خارجی مدل شایستگی است آورده شده است. در نهایت، در ردیف هفدهم، همبستگی نمرات شایستگی‌ها در کانون ارزیابی با نمره پرسشنامه ارزیابی عملکرد که توسط مدیران تکمیل شده است قید شده است که نشان‌دهنده میزان روایی ملاکی هر شایستگی است.

نتایج نشان می‌دهد در بخش نمرات کانون ارزیابی، میانگین نمرات شایستگی‌ها معادل ۲٫۶۵ است (از ۵ نمره). همچنین به‌طور میانگین، همبستگی بین شایستگی‌ها در کانون ۰٫۵۶ است. بیشترین میزان همبستگی در این قسمت بین شایستگی‌های هدف‌گرایی و هدایت عملکرد و شایستگی «نظارت و کنترل» با میزان ۰٫۷۵، و کمترین میزان همبستگی، بین شایستگی «آینده‌نگری» با شایستگی «تیم و شبکه‌سازی» با میزان ۰٫۳۵ است. میزان بالای همبستگی‌ها در این بخش نشان می‌دهد که مدل شایستگی از ساختار داخلی مناسبی برخوردار نیست و همپوشانی بالایی بین شایستگی‌ها دیده می‌شود.

1. heterotrait-monomethod correlations; HTMM

در بخش نمرات ارزیابی مدیر مستقیم، میانگین نمرات شایستگی‌ها معادل ۳,۵۷ است (از ۵ نمره). همچنین به‌طور میانگین، همبستگی بین شایستگی‌ها در کانون با نمرات شایستگی براساس ارزیابی مدیر مستقیم ۰,۰۸ است. این میزان پایین از همبستگی نشان‌دهنده عدم تحقق روایی سازه خارجی در مدل شایستگی است. بیشترین میزان روایی سازه خارجی در این قسمت متعلق به شایستگی «نظارت و کنترل» با میزان ۰,۲۳، کمترین میزان آن در این قسمت متعلق به شایستگی «تعهد و تعلق سازمانی» با میزان ۰,۰۳ است.

در بخش روایی ملاکی که براساس میزان همبستگی شایستگی‌ها با نمره «عملکرد شغلی» مورد بررسی قرار می‌گیرد؛ متوسط همبستگی مدل شایستگی در کانون با متغیر ملاک معادل ۰,۱۸ است. بیشترین روایی ملاکی متعلق به شایستگی «تفکر تحلیلی و حل مسئله» با مقدار ۰,۲۷، و کمترین آن متعلق به شایستگی «مسئولیت‌پذیری» با مقدار ۰,۱۱ است

جدول ۳: میانگین، انحراف معیار و همبستگی میان شایستگی‌ها

۱۵	۱۴	۱۳	۱۲	۱۱	۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	SD	M	منبع داده / شایستگی‌ها
																	کانون ارزیابی
															۰,۴	۳,۱	۱. تعهد و تعلق سازمانی
															۰,۹	۲	
															۰,۴	۰,۷	۲. هدف‌گرایی و هدایت عملکرد
															۰,۲	۲	
															۰,۶	۰,۵	۳. مهارت‌های ارتباطی
															۱	۳	
															۰,۶	۰,۵	۴. تفکر تحلیلی و حل مسئله
															۰	۳	
															۰,۴	۰,۷	۵. تیم و شبکه‌سازی
															۰,۷	۳,۲	
															۰,۷	۵	
															۰,۵	۰,۵	۶. مسئولیت‌پذیری
															۰	۷	
															۰,۶	۰,۴	۷. نظارت و کنترل
															۰,۹	۶	
															۰,۵	۰,۳	۸. آینده‌نگری و تفکر راهبردی
															۱	۱	
																	ارزیابی مدیر مستقیم
																	۹. تعهد و تعلق سازمانی
															۰,۰۱	۰,۱۰	
															-	۰,۰۲	
															۰,۰۵	۰,۰۸	
															۰,۰۱	۰,۰۱	
															۰,۱۰	۰,۰۳	
															۰,۴۵	۴,۲۰	

منبع داده / شایستگی‌ها	M	SD	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
۱۰. هدف‌گرایی و هدایت عملکرد	۳,۱۱	۰,۶۲	۰,۰۵	۰,۱۱	۰,۰	۰,۰۸	۰,۰۱	۰,۱۸	۰,۲۵	۰,۲۰	۰,۲						
۱۱. مهارت‌های ارتباطی	۳,۲۰	۰,۶۱	۰,۱۱	۰,۱۱	۰,۲۰	۰,۱۳	۰,۰۶	۰,۰۹	۰,۱۸	۰,۱۶	۰,۳	۰,۵					
۱۲. تفکر تحلیلی و حل مسئله	۳,۰۱	۰,۶۵	۰,۰۶	۰,۱۵	۰,۰۳	۰,۱۹	۰,۰۲	۰,۲۱	۰,۱۳	۰,۱۸	۰,۲	۰,۶	۰,۳				
۱۳. تیم و شبکه‌سازی	۳,۵۰	۰,۶۶	۰,۱۰	۰,۰۸	۰,۱۲	۰,۰۵	۰,۱۱	۰,۰۱	۰,۱۳	۰,۰۱	۰,۲	۰,۶	۰,۷	۰,۴			
۱۴. مسئولیت‌پذیری	۳,۹۰	۰,۵۰	۰,۱۵	۰,۰۹	۰,۰۱	۰,۰۲	۰,۰۲	۰,۱۵	۰,۱۰	-	۰,۲	۰,۶	۰,۴	۰,۳	۰,۴		
۱۵. نظارت و کنترل	۴,۰۲	۰,۶۷	۰,۰۱	۰,۰۶	۰,۰۲	۰,۰۹	۰,۰۱	۰,۰۴	۰,۲۳	۰,۰۲	۰,۲	۰,۷	۰,۴	۰,۳	۰,۳	۰,۳	
۱۶. آینده‌نگری و تفکر راهبردی	۳,۶۴	۰,۶۶	۰,۰۱	۰,۱۰	۰,۰۲	۰,۱۲	۰,۰۱	۰,۰۵	۰,۲۴	۰,۱۸	۰,۲	۰,۵	۰,۵	۰,۳	۰,۳	۰,۳	
۱۷. عملکرد شغلی	۳,۵۸	۰,۶۰	۰,۱۳	۰,۲۴	۰,۲۰	۰,۲۷	۰,۲۱	۰,۱۱	۰,۱۵	۰,۱۷							

میزان معناداری تمام همبستگی‌ها در سطح ۰,۰۱ است.

جهت بررسی برازش مدل شایستگی‌ها از نرم‌افزار Amos استفاده شد. با توجه به این که هر شایستگی در حداقل ۲ تمرین مورد سنجش قرار گرفته است؛ برای ترسیم مدل در نرم‌افزار، هر شایستگی به‌عنوان یک متغیر مکنون در نظر گرفته شده است که نمرات نهایی هر شایستگی، داده‌های آن هستند و نمره هر شایستگی در هر تمرین به‌عنوان متغیرهای آشکار در نظر گرفته می‌شوند. همان‌طور که در بخش روش‌شناسی ذکر شد، براساس منطق MTMM، هر شایستگی یک صفت و هر تمرین یک روش در نظر گرفته می‌شود. نمودار مسیر تحلیل عاملی در پیوست قابل مشاهده است. بر این اساس، هر شایستگی با کد D و هر تمرین با کد E برچسب گذاری شده است.

برای حصول برازش مدل بایستی شروط زیر مهیا باشد: شاخص RMSEA کوچک‌تر یا مساوی ۰,۰۶، شاخص SRMR کوچک‌تر مساوی ۰,۰۸، و برای CFI و TLI بزرگ‌تر مساوی ۰,۹۵ باشد. همچنین شاخص نسبت کای اسکوئر به درجه آزادی نیز کمتر از ۳ باشد. (Hu and Bentler, 1999) نتایج جدول زیر نشان می‌دهد که هیچ‌یک از شرایط حصول برازش برآورده نشده است.

جدول ۶: آماره‌های برازش مدل شایستگی‌های عمومی مدیران حرفه‌ای

شاخص	برآورد
Df	۲۳۲
χ^2	۹۷۴,۴
Df/χ^2	۴,۲
RMSEA	۰,۰۸۱
SRMR	۰,۰۷۷
TLI	۰,۴۸۵
CFI	۰,۶۳۸

بررسی روایی سازه مبتنی بر تحلیل عاملی تأییدی

جهت بررسی روایی سازه در تحلیل عاملی تأییدی، از شاخص‌های CR^1 ، AVE و MSV استفاده می‌شود. شاخص پایایی مرکب (CR)، معیاری برای سازگاری درونی در مدل است و تعیین می‌کند که آیا آزمون به‌خوبی آن‌چه را که قرار است اندازه‌گیری کند، اندازه‌گیری می‌کند یا خیر. به‌عبارت‌دیگر، آیا آزمون به‌گونه‌ای ساخته‌شده است که آن‌چه را که ادعا می‌کند با موفقیت آزمایش کند؟

روایی سازه شامل دو روایی همگرا^۲ و واگرا^۳ است. روایی واگرا به میزانی اشاره می‌کند که یک سازه متمایز از سازه دیگر است؛ بنابراین، روایی واگرا، یکتایی سازه را نشان می‌دهد (Hair et al., 2010; Jr et al., 2018).

شاخص AVE یا میانگین واریانس استخراج‌شده، اندازه‌گیری مقدار واریانس است که توسط یک سازه در رابطه با مقدار واریانس ناشی از خطای اندازه‌گیری گرفته می‌شود. شاخص MSV، مربع بالاترین ضریب همبستگی بین سازه‌های نهفته است (Hair et al., 2010; Jr et al., 2018).

برای تأیید روایی سازه شاخص CR باید بالای ۰,۷ باشد و همان‌طور که از نتایج جدول زیر نیز مشخص است، شاخص CR برای همه شایستگی‌ها کمتر از ۰,۷ است و این یعنی متغیرهای سازه اصلی نتوانسته‌اند در اندازه‌گیری آن‌چه برای آن طراحی شده‌اند، موفق عمل کنند.

1. Composite Reliability
2. Convergent validity
3. Discriminant validity

برای تأیید روایی همگرا باید مقدار AVE که نشان دهنده میانگین واریانس استخراجی است، برای همه متغیرها بالای ۰,۵ باشد، در غیر این صورت، روایی همگرا تأیید نخواهد شد (Hair et al., 2018; Jr et al., 2010). همان‌طور که از نتایج زیر نیز مشخص است، همه AVE تا کمتر از ۰,۵ بوده و همگرایی مدل تأیید نمی‌شود، یعنی متغیرهای درونی هر سازه با یکدیگر همبستگی بالایی نداشته است. به‌طور مثال، نمرات شایستگی مسئولیت‌پذیری در تمرین اول و چهارم با یکدیگر همبستگی نداشته است.

برای تأیید روایی واگرا شاخص AVE با شاخص MSV مقایسه شده و در صورتی که مقدار MSV کمتر از AVE باشد می‌توان گفت روایی واگرا نیز تأیید می‌شود (Hair et al., 2018; Jr et al., 2010). اما همان‌طور که از نتایج نیز مشخص است، روایی واگرای هیچ‌یک از شایستگی‌ها تأیید نمی‌شود و این یعنی سازه‌ها که همان ۸ شایستگی هستند با یکدیگر همپوشانی بالایی داشته و می‌توان تعداد شایستگی‌ها را به عدد بهینه‌تری کاهش داد.

جدول ۷: شاخص‌های روایی سازه شایستگی‌ها براساس خروجی Amos

MSV	AVE	CR	شایستگی
۷,۶۱۲	۰,۰۳۶	۰,۰۲	تعهد و تعلق سازمانی
۱,۶۵۹	۰,۲۳۴	۰,۳۶	هدف‌گرایی و هدایت عملکرد
۱,۶۵۹	۰,۲۵۱	۰,۴۰	مهارت‌های ارتباطی
۱,۰۷۵	۰,۲۱۹	۰,۳۳	تفکر تحلیلی و حل مسئله
۱,۲۹۵	۰,۲۶۹	۰,۳۵	تیم و شبکه‌سازی
۱۰,۳۹۴	۰,۰۵۹	۰,۱۲	مسئولیت‌پذیری
۹,۸۷۴	۰,۰۶۵	۰,۱۳	نظارت و کنترل
۹,۴۲۵	۰,۱۰۴	۰,۱۸	آینده‌نگری و تفکر راهبردی

بحث و نتیجه‌گیری

هدف از این پژوهش، بررسی اثربخشی مدل شایستگی عمومی مدیران حرفه‌ای در بخشنامه ابلاغی سازمان اداری و استخدامی کشور بود. جهت انجام این پژوهش، با طراحی و اجرای کانون ارزیابی و سپس بررسی نتایج ۱۶۱۶ نفر از ارزیابی‌شوندگان؛ اثربخشی مدل شایستگی از طریق تحلیل روایی سازه درونی و بیرونی، برازش مدل و همچنین روایی ملاکی مدل شایستگی مذکور مورد بررسی قرار گرفت.

نتایج این بررسی نشان داد که اثربخشی مدل شایستگی عمومی مدیران حرفه‌ای وضعیت مناسبی ندارد و جهت برآورده کردن هدفی که برای آن طراحی شده است با تردیدهایی روبرو است. از نظر روایی سازه داخلی و برازش مدل، هیچ یک از شاخص‌های برازش مدل وضعیت مناسبی را نشان ندادند. هیچ یک از هشت شایستگی موردبررسی، روایی سازه داخلی مناسبی نداشتند و همبستگی بالایی بین نمرات شایستگی‌ها وجود دارد. نتایج این بخش با نتایج پژوهش‌های پیشین نیز همخوانی دارد و روایی سازه داخلی و برازش مدل در مدل‌های شایستگی در کانون‌های ارزیابی یک مسئله جدی در کانون‌های ارزیابی بوده است (Arthur et al., 2008; Hoffman et al. 2015; Reilly et al., 1990; Rupp et al., 2008) از نظر روایی سازه خارجی، به‌طور میانگین، همبستگی بین شایستگی‌ها در کانون ارزیابی با نمرات شایستگی براساس ارزیابی مدیر مستقیم ۰,۰۸ است که حکایت از عدم تحقق روایی سازه خارجی دارد. نتایج این بخش می‌تواند نتایج پژوهش‌های دیگر را که ارتباط سازه‌های دیگری را مانند شخصیت و هوش با نتایج مدل شایستگی کانون‌های ارزیابی بررسی کردند تکمیل کند (مانند Merkulova et al. 2016). همچنین از نظر روایی ملاکی و همبستگی نمرات مدل شایستگی در کانون با نمرات عملکرد شغلی افراد، صرفاً چهار شایستگی همبستگی بالای ۰,۲ داشتند که این موارد با نتایج برخی فراتحلیل‌ها در این زمینه مانند هرملین و همکاران (۲۰۰۷) و مناهان و همکاران (۲۰۱۳) همخوانی دارد که روایی ملاکی ۰,۲۷ را برای نمره نهایی کانون ارزیابی گزارش کردند. البته شایستگی‌های «تعهد و تعلق سازمانی» و «مسئولیت‌پذیری» که کمترین روایی ملاکی و سازه را داشتند اساساً در فراتحلیل‌های مدل شایستگی و کانون ارزیابی (مانند Arthur et al., 2003) نیز وجود نداشتند و با توجه به ماهیتشان از جنس دانش، توانایی و مهارت نیستند. شایستگی «تفکر تحلیلی و حل مسئله» با روایی ملاکی ۰,۲۷، مشابه با یافته‌های برخی فراتحلیل‌ها درباره این شایستگی است (Connelly et al, 2008).

به‌لحاظ نظری، نتایج این پژوهش می‌تواند سهم‌افزایی مناسبی در موضوع مورد بررسی داشته باشد. این پژوهش جزو محدود پژوهش‌هایی است که روی یک حجم نمونه گسترده و در فضای عملیاتی، اثربخشی مدل شایستگی را در کانون‌های ارزیابی با طراحی مبتنی بر استاندارد بررسی کرده است. علاوه بر این، دامنه پژوهش را منحصر به داده‌های کانون ارزیابی نکرده و با گردآوری داده‌های مناسب به‌ازای هر شایستگی، امکان بررسی دقیق روایی سازه خارجی و روایی ملاکی را فراهم کرده است. همچنین نتایج این پژوهش می‌تواند نکات مهمی را برای پژوهش‌های آتی حوزه کانون ارزیابی داشته باشد چراکه عدم تحقق روایی سازه داخلی و خارجی مدل شایستگی در کانون ارزیابی، هشدار جدی برای اتکا به نتایج مدل‌های شایستگی و تفسیر آن‌ها براساس بررسی این

داده‌ها برای اولین بار در بستر ایران و همچنین روایی سازه خارجی شایستگی‌ها برای مقاصد تحقیقاتی و اجرایی دیگر به همراه دارد.

به لحاظ کاربردی، این پژوهش برای اولین بار در کشور ایران به بررسی نتایج کمی و داده‌های مدل شایستگی و کانون ارزیابی در فضای دولتی در حجم نمونه گسترده کرده است. همچنین نه تنها برای اولین بار اثربخشی «دستورالعمل نحوه ارزیابی و توسعه شایستگی‌های عمومی مدیران حرفه‌ای» را بعد از ۵ سال از ابلاغ (و اجرای آن روی بیش از ۳۰ هزار داوطلب مدیریتی بخش دولتی) آن مبتنی بر داده‌های کمی بررسی کرده است؛ بلکه اولین تحلیل کمی و داده‌محور از نوع خود در بررسی دستورالعمل‌های حوزه منابع انسانی در بخش دولتی کشور ایران است و می‌تواند به عنوان مبنایی جهت ارتقا و بهبود دستورالعمل مذکور جهت انتخاب مدیران دولتی قرار گیرد.

پیشنادهای کاربردی و سیاستی

با توجه به این که مدل شایستگی عمومی مدیران حرفه‌ای در دستورالعمل مربوطه، مهم‌ترین عنصر ارزیابی و توسعه مدیران بخش دولتی به حساب می‌آید؛ توصیه می‌شود در طراحی این مدل، بازنگری اساسی صورت گیرد. همچنین بایستی در نحوه قبول یا رد داوطلبین براساس نمره هر شایستگی و تعیین حد نصاب ۶۰ درصد برای هر شایستگی نیز تردید کرد، چراکه براساس نتایج این پژوهش اساساً مشخص نیست که آیا عملکرد فرد در شایستگی مدنظر سنجیده شده است یا خیر. در نتیجه بایستی در الزام بر طراحی دوره‌های آموزشی و گذراندن آن‌ها به‌ازای هر شایستگی برای ارزیابی شوندگان تردید کرد.

جهت افزایش اثربخشی نتایج کانون‌های ارزیابی توصیه می‌شود به‌جای تأکید بر مدل شایستگی، کانون تمرکز، متوجه طراحی و محتوای تمرین‌های مورد استفاده در کانون‌های ارزیابی و حصول اطمینان از روایی صوری و محتوایی تمرین‌ها و ارتباط آن‌ها با مشاغل هدف شود. این موضوع براساس متن دستورالعمل مورد غفلت کامل قرار گرفته است. همچنین بررسی امکان استفاده از کانون‌های وظیفه‌محور^۱ نیز می‌تواند توصیه شود (Jackson, Lance, & Hoffman, 2012). در حقیقت با توجه به اولویت و سهم بالای تمرین‌ها در اثربخشی کانون ارزیابی (Putka & Hoffman, 2013)، نقش نظارتی سازمان اداری و استخدامی کشور بایستی به‌جای تمرکز بر تربیت ارزیاب، بر ممیزی محتوای کانون‌های ارزیابی باشد.

1. task-based AC

در آخر نیز توصیه می‌شود جهت کاهش ریسک تمرکز صرف بر نتایج کانون ارزیابی و مدل شایستگی و افزایش جامعیت در ارزیابی افراد، از منابع متعدد مانند داده‌های عملکرد شغلی افراد نیز در تحلیل افراد استفاده شود تا از طریق تجمیع چند روش، خطای روش‌ها کاهش پیدا کند.

محدودیت‌های پژوهش

در تعمیم نتایج این پژوهش به بافت‌های دیگر مانند سازمان‌های خصوصی باید با احتیاط عمل کرد، زیرا ارزیابی‌شوندگان در این پژوهش افراد واجد شرایط طبق دستورالعمل بودند و از این جهت با محدودیت دامنه در انتخاب ارزیابی‌شوندگان روبرو بودند. همچنین اثربخشی این شایستگی‌ها صرفاً در فضای اجرای کانون ارزیابی بررسی شده است نه لزوماً روش‌های دیگر. امکان برگزاری آموزش‌های چارچوب مرجع برای مدیران مستقیم افراد جهت تکمیل پرسشنامه‌های شایستگی و عملکرد شغلی وجود نداشت و از این جهت ممکن است سوگیری‌هایی در نتایج این پرسشنامه‌ها و وقوع اثر ناخالصی متغیر ملاک^۱ وجود داشته باشد.

پیشنهاد‌های پژوهشی

با توجه به قلمرو و محدودیت‌های این پژوهش، پیشنهادهایی برای پژوهش‌های آتی در این حیطه ارائه می‌گردد:

پیشنهاد می‌گردد نتایج مدل شایستگی عمومی مدیران حرفه‌ای با عامل سازه‌های دیگر مانند استفاده از ابعاد گسترده (Hoffman et al., 2011; Meriac et al., 2014) مورد بررسی و با نتایج حاضر مقایسه شود و در صورت حصول نتیجه بهتر، به‌عنوان مبنایی برای بهینه‌سازی مدل شایستگی فعلی مطرح گردد.

پیشنهاد می‌گردد نتایج مدل شایستگی فعلی با نتایج روش‌های غیر شبیه‌سازی مانند مصاحبه شایستگی محور ساختاریافته، آزمون‌های شناختی و آزمون‌های رفتاری و شخصیت مقایسه گردد و میزان روایی سازه و ملاکی و روایی فزاینده^۲ بررسی شود.

در نهایت پیشنهاد می‌گردد میزان اثربخشی مدل شایستگی فعلی از طریق بررسی داده‌های عملکرد شغلی آتی افراد و گزارش روایی پیش‌بین مورد بررسی و تحلیل قرار گیرد.

1. criterion contamination

2. incremental validity

مآخذ

- Arthur Jr, W., Woehr, D. J., & Maldegen, R. (2000). *Convergent and discriminant validity of assessment center dimensions: A conceptual and empirical reexamination of the assessment center construct-related validity paradox*. *Journal of Management*, 26(4), 813-835.
- Arthur Jr, W., & Villado, A. J. (2008). *The importance of distinguishing between constructs and methods when comparing predictors in personnel selection research and practice*. *Journal of Applied Psychology*, 93(2), 435.
- Arthur, W., Day, E. A., & Woehr, D. J. (2008). *Mend it, don't end it: An alternate view of assessment center construct-related validity evidence*. *Industrial and Organizational Psychology*, 1(1), 105-111.
- Arthur Jr, W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). *A meta analysis of the criterion related validity of assessment center dimensions*. *Personnel Psychology*, 56(1), 125-153.
- Barrick, M. R., Day, D. V., Lord, R. G., & Alexander, R. A. (1991). *Assessing the utility of executive leadership*. *The Leadership Quarterly*, 2(1), 9-22. [https://doi.org/10.1016/1048-9843\(91\)90004-L](https://doi.org/10.1016/1048-9843(91)90004-L)
- Bartram, D. (2005). *The Great Eight competencies: a criterion-centric approach to validation*. *Journal of applied psychology*, 90(6), 1185.
- Bong, M.; & Hong, S. (2010). "Multitrait-Multimethod Designs", In P. Peterson, E. Baker, & B. McGaw (eds.), (P. Peterson, E. Baker, & B. McGaw, eds.), *International Encyclopedia of Education* (Third Edition), (pp.111-119), Oxford:Elsevier. Retrieved from <https://www.sciencedirect.com/science/article/pii/B9780080448947002888>
- Bott, J. P., Svyantek, D. J., Goodman, S. A., & Bernal, D. S. (2003). *Expanding the performance domain: who says nice guys finish last?*. *The International Journal of Organizational Analysis*, 11(2), 137-152.
- BPS. (2015). *The design and delivery of Assessment Centres*. *The British Psychological Society*. <https://www.bps.org.uk/sites/bps.org.uk/files/Member Networks/Divisions/DOP/DOP The Design and Delivery of Assessment Centres.pdf>
- Byham, W. C. (1970). *Assessment center for spotting future managers*. *Harvard Business Review*, 48, 150-160, plus appendix.
- Campbell, D. T., & Fiske, D. W. (1959). *Convergent and discriminant validation by the multitrait-multimethod matrix*. *Psychological Bulletin*, 56(2), 81-105.
- Campion, M. A., Fink, A. A., Ruggeberg, B. J., Carr, L., Phillips, G. M., & Odman, R. B. (2011). *Doing competencies well: Best practices in competency modeling*. *Personnel Psychology*, 64, 225-262. doi:10.1111/j.1744- 6570.2010.01207.x

- Connelly, B. S., Ones, D. S., Ramesh, A., & Goff, M. (2008). *A pragmatic view of assessment center exercises and dimensions. Industrial and Organizational Psychology*, 1, 121–124.
- Fleener, J. W. (1996). *Constructs and developmental assessment centers: Further troubling empirical findings. Journal of Business and Psychology*, 10 (3), 319–333.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). *Meta-analysis of assessment center validity. Journal of applied psychology*, 72, 493–511.
- Hair, J. et al. (2010) '*Multivariate Data Analysis.pdf*', Australia : Cengage, p. 758.
- Hermelin, E., Lievens, F., & Robertson, I. T. (2007). *The validity of assessment centres for the prediction of supervisory performance ratings: A meta analysis. International Journal of Selection and Assessment*, 15(4), 405-411.
- Hoffman, B. J., Kennedy, C. L., LoPilato, A. C., Monahan, E. L., & Lance, C. E. (2015). *A review of the content, criterion-related, and construct-related validity of assessment center exercises. Journal of Applied Psychology*, 100, 1143–1168.
- Hoffman, B. J., Melchers, K. G., Blair, C. A., Kleinmann, M., & Ladd, R. T. (2011). *Exercises and dimensions are the currency of assessment centers. Personnel Psychology*, 64, 351–395.
- Hu, L. T., & Bentler, P. M. (1999). *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling*, 6.
- Hughes, D., Riley, P., Shalfröoshan, A., Gibbons, A., & Thornton, G. C. III (2012). *A global survey of assessment centre practices. Surrey, UK: Assessment & Development Consultants. www.adc.uk.com/resources/library/?resources_page=1&cat[]=16* Accessed August 25, 2023.
- International Taskforce on Assessment Center Guidelines. (2015). "*Guidelines and Ethical Considerations for Assessment Center Operations*", *Journal of Management*, Vol. 41. No. 4. PP. 1244-1273.
- Jackson, D. J. R., Lance, C. E., & Hoffman, B. J. (2012). *The psychology of assessment centers. New York, NY: Routledge.*
- Jr, J.F.H. et al. (2018) *Multivariate Data Analysis. doi:10.1002/9781119409137.ch4.*
- Kleinmann, M., & Ingold, P. V. (2019). *Toward a Better Understanding of Assessment Centers: A Conceptual Review. Annual Review of Organizational Psychology and Organizational Behavior*, 6(1), 349–372.
- Lance Jackson, D. J. R., Stillman, J. A., & Atkins, S. G. (2005). *Rating tasks versus dimensions in assessment centers: A psychometric comparison. Human Performance*, 18 (3), 213-241.

- Lance, C. E. (2008). *Why Assessment Centers Do Not Work the Way They Are Supposed To*. *Industrial and Organizational Psychology*, 1(1), 84–97.
- Lance, C. E., Lambert, T. A., Gewin, A. G., Lievens, F., & Conway, J. M. (2004). *Revised estimates of dimension and exercise variance components in assessment center post exercise dimension ratings*. *Journal of Applied Psychology*, 89(2), 377-385.
- Lance, Charles E.; Newbolt, William H.; Gatewood, Robert D.; Foster, Mark R.; French, Nita R.; & Smith, David E. (2000). "Assessment Center Exercise Factors Represent Cross-Situational Specificity, Not Method Bias", *Human Performance*, Vol. 13. No. 4. PP. 323-353.
- Lievens, F. (2002). *Trying to understand the different pieces of the construct validity puzzle of assessment centers: An examination of assessor and assessee effects*. *Journal of Applied Psychology*, 87(4), 675–686.
- Lievens, F. (2008). *What Does Exercise-Based Assessment Really Mean?* *Industrial and Organizational Psychology*, 1(1), 112–115. <https://doi.org/10.1111/j.1754-9434.2007.00020.x>
- Lievens, F., & Keer, E. (2001). *The construct validity of a Belgian assessment centre: A comparison of different models*. *Journal of Occupational and Organizational Psychology*, 74(3), 373–378.
- Meriac, J. P., Hoffman, B. J., & Woehr, D. J. (2014). *A conceptual and empirical review of the structure of assessment center dimensions*. *Journal of Management*, 40, 1269–1296.
- Meriac, J. P., Hoffman, B. J., Woehr, D. J., & Fleisher, M. S. (2008). *Further evidence for the validity of assessment center dimensions: A meta-analysis of the incremental criterion-related validity of dimension ratings*. *Journal of Applied Psychology*, 93(5), 1042- 1052.
- Merkulova, N., Melchers, K. G., Kleinmann, M., Annen, H., & Szvircsev Tresch, T. (2016). *A test of the generalizability of a recently suggested conceptual model for assessment center ratings*. *Human Performance*, 29(3), 226-250.
- Messick, S. (1980). *Test validity and the ethics of assessment*. *American Psychological Association*.
- Monahan Monahan, E. L., Hoffman, B. J., Lance, C. E., Jackson, D. J., & Foster, M. R. (2013). *Now you see them, now you do not: The influence of indicator-factor ratio on support for assessment center dimensions*. *Personnel Psychology*, 66, 1009–1047.
- Nako, Z. C., & ARTS, M. in. (2011). *Construct validity of a managerial assessment centre [, Citeseer]*. *EndNote Tagged Import Format*.
- Putka, D. J., & Hoffman, B. J. (2013). *Clarifying the contribution of assessee-, dimension-, exercise-, and assessor- related effects to reliable and unreliable variance in assessment center ratings*. *Journal of Applied Psychology*, 98, 114–133. doi:10.1037/a0030887

- Roch, S. G., Woehr, D. J., Mishra, V., & Kieszczynska, U. (2012). *Rater training revisited: An updated meta-analytic review of frame-of-reference training*. *Journal of Occupational and Organizational Psychology*, 85, 370-395.
- Rupp, D. E., Hoffman, B. J., Bischof, D., Byham, W., Collins, L., Gibbons, A., Hirose, S., Kleinmann, M., Kudisch, J. D., Lanik, M., Jackson, D. J. R., Kim, M., Lievens, F., Meiring, D., Melchers, K. G., Pendit, V. G., Putka, D. J., Povah, N., Reynolds, D.,... Thornton, G. (2015). *Guidelines and Ethical Considerations for Assessment Center Operations*. *Journal of Management*, 41(4), 1244–1273.
- Sackett, P. R., & Dreher, G. F. (1982). *Constructs and assessment center dimensions: Some troubling empirical findings*. Reprint: Nr. 228.
- Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). *Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range*. *Journal of Applied Psychology*, 107(11), 2040.
- SHL (2018). *2018 global assessment trends report*.
- Shore, T. H., Thornton, G. C., & Shore, L. M. (1990). *Construct validity of two categories of assessment center dimension ratings*. *Personnel Psychology*, 43(1), 101–114.
- Speer, A. B., Christiansen, N. D., Goffin, R. D., & Goff, M. (2014). *Situational bandwidth and the criterion-related validity of assessment center ratings: Is cross-exercise convergence always desirable?* *Journal of Applied Psychology*, 99 (2), 282-295.
- Thornton III, G. C., Rupp, D. E., & Hoffman, B. J. (2014). *Assessment center perspectives for talent management strategies*. Routledge.
- Thornton, G. C., & Rupp, D. E. (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis, and development*. Mahwah, NJ: Erlbaum.
- Thornton, G. C., Rupp, D. E., & Hoffman, B. J. (2015). *Assessment center perspectives for talent management strategies (Second edition)*. Routledge.
- Thornton, George C.; Mueller-Hanson, Rose A.; & Rupp, Deborah E. (2017). *“Developing organizational simulations: a guide for practitioners, students, and researchers”*, (Second Edition), New York London: Routledge, Taylor & Francis Group.
- Woehr, D. J., & Arthur Jr, W. (2003). *The construct-related validity of assessment center ratings: A review and meta-analysis of the role of methodological factors*. *Journal of Management*, 29(2), 231-258.
- Zysberg, L. (2012). *Assessment center dimensions predict performance-based bonus in business management settings*. *Psychology*, 3, 500–503.

پیوست

نمودار مسیر تحلیل عاملی

