

Validation of Managers' Promotional Assessment Center in Isfahan Gas Company

Mahdiehsadat Khoshouei.Ph.D

Psychology, University of Isfahan, Isfahan, Iran.

Hamidreza Oreyzi.Ph.D

Academic member, University of Isfahan, Isfahan, Iran.

Afshin Jahanbazi PhD student

Psychology, University of Isfahan, Isfahan, Iran.

اعتباریابی کانون ارزیابی ارتقایی مدیران در شرکت گاز اصفهان^۱

مهديه سادات خشوعی*

پسادکتری روان‌شناسی صنعتی و سازمانی، دانشگاه اصفهان، اصفهان، ایران.

حمیدرضا عریضی

عضو هیئت‌علمی دانشگاه اصفهان، اصفهان، ایران.

افشین جهان‌بازی

دانشجو دکتری دانشگاه اصفهان، اصفهان، ایران.

Abstract

The goal of this research was validation of manager's promotional assessment center in Isfahan Gas company. In order to achieve the goal, according to assessment center designed Isfahan Gas company (khoshouei, Oreyzi and Nouri, 2014) eight mid-level managers of company as participant and thirteen experts with masters' degree at industrial and organizational psychology as expert and assessor selected purposefully (homogeneous sampling). Then, experts were completed the training course effectiveness questionnaire and simulation exercise effectiveness questionnaire, assessors were completed the simulation exercise effectiveness questionnaire and behaviorally anchored rating scale and managers were completed the simulation exercise effectiveness questionnaire and managerial dimensional questionnaire (Khoshouei, 2012). Therefore, standardization (by survey of experts and managers), face validity (by survey of experts and assessors), content validity (by survey of managers), concurrent validity (by correlation of simulations ratings and managerial dimensional questionnaire) and construct validity (multi-trait multi-method analysis) were established. Reliability of assessment center were found to be confirmatory by inter-rater reliability, intra-rater reliability, reliability of same dimensions in different exercises, and behaviorally anchored rating scale reliability.

Keywords: Managers' promotional assessment center, Standardization, Validity, Reliability

چکیده

هدف این پژوهش، اعتباریابی کانون ارزیابی ارتقایی مدیران در شرکت گاز اصفهان بود. برای دستیابی به این هدف متناسب با کانون ارزیابی طراحی شده شرکت گاز اصفهان (خشوعی، عریضی و نوری، ۱۳۹۳)، تعداد ۸ نفر از مدیران میانی شرکت گاز اصفهان به عنوان شرکت کننده و تعداد ۱۳ نفر کارشناس ارشد روان‌شناسی صنعتی و سازمانی به عنوان کارشناس و ارزیاب به روش نمونه گیری هدفمند از نوع همگون انتخاب شدند. سپس کارشناسان پرسشنامه اثربخشی دوره آموزشی، پرسشنامه اثربخشی تمرین شبیه سازی، ارزیاب‌ها پرسشنامه اثربخشی تمرین شبیه سازی و مقیاس امتیازدهی رفتاری و مدیران پرسشنامه اثربخشی تمرین شبیه سازی و پرسشنامه ابعاد مدیریتی (خشوعی، ۱۳۹۱) را تکمیل نمودند. به این ترتیب از طریق نظرخواهی از کارشناسان و مدیران، استاندارد بودن کانون، از طریق نظرخواهی از کارشناسان و ارزیاب‌ها اعتبار صوری کانون و از طریق نظرخواهی از مدیران، اعتبار محتوای کانون تأیید گردید. ارتباط نمره‌های تمرین با پرسشنامه ابعاد مدیریتی و تحلیل چند صفت - چند روش نیز به ترتیب تأیید کننده اعتبار همزمان و اعتبار سازه کانون بود. بر اساس پایایی بین ارزیاب‌ها، پایایی درون ارزیاب‌ها، پایایی بین ابعاد یکسان در تمرین‌های مختلف و پایایی مقیاس امتیازدهی رفتاری نیز نتایج حاکی از پایایی مطلوب کانون بود.

واژه‌های کلیدی: کانون ارزیابی ارتقایی مدیران، استاندارد، اعتبار، پایایی

۱. این پژوهش با حمایت مالی شرکت ملی گاز انجام شده است.

* نویسنده مسئول: khoshouei.mahdieh@gmail.com

پذیرش: ۹۷/۰۶/۲۵

وصول: ۹۷/۰۳/۲۷

مقدمه

(ابوالعلائی و غفاری، ۱۳۸۶) و شرکت ملی پالایش و پخش فرآورده‌های نفتی (عبدی و دیانتی، ۱۳۸۷) به کار رفته است. به دلیل پیچیدگی کانون و متعدد بودن مؤلفه‌های آن، پژوهشگران و سازمان‌ها در خارج از کشور به بررسی و بهبود هر یک از مؤلفه‌های آن پرداخته‌اند. برای مثال می‌توان به چگونگی آموزش ارزیاب‌ها (اسچلیچر، دی، میس و ریجیو^{۱۱}، ۲۰۰۲) و رفع خطاهای ارزیابی (خشوعی، عریضی و نوری، ۲۰۱۲) اشاره نمود؛ اما یکی از مهم‌ترین موارد مرتبط با کانون، ویژگی‌های روان‌سنجی آن یعنی استانداردسازی، اعتبار و روایی است (خشوعی، ۱۳۹۳) که متأسفانه پژوهشگران ایرانی از آن غفلت نموده‌اند.

استانداردسازی، یعنی اجرای یکنواخت روش ارزیابی به‌طوری‌که دستورالعمل، محیط و شرایط محرک برای تمام شرکت‌کنندگان، یکسان باشد. همچنین، نمره‌گذاری باید استاندارد شده باشد به‌نحوی‌که رفتار معین، به نحو یکسانی ارزیابی شود (تورنتن و مولر-هانسون، ۲۰۰۴). در زمینه استانداردسازی، تحقیقات تجربی انجام نشده است، اما بر اساس دیدگاه تورنتن و مولر-هانسون (۲۰۰۴) طراحی دستورالعمل دقیق و واضح برای شرکت‌کننده، ارزیاب، ایفاگر نقش و شخص منبع، آموزش ارزیاب، ایفاگر نقش و شخص منبع، استفاده از یک ایفاگر نقش و شخص منبع ثابت برای تمام شرکت‌کنندگان و تصادفی قرار دادن شرکت‌کنندگان در تمرین بحث گروهی، باعث افزایش استاندارد می‌شود.

در بررسی اعتبار کانون نیز، چهار اعتبار مهم است: ۱- صوری یا وفاداری^{۱۲} اشاره به شباهت ظاهری تمرین‌ها و ابعاد موجود در کانون با شغل هدف دارد (تورنتن و مولر-هانسون، ۲۰۰۴). در رابطه با اعتبار صوری به اعتقاد اکثریت افراد حاضر در کانون، تمرین‌ها از اعتبار صوری بالایی برخوردار است (کودیچ^{۱۳} و همکاران، ۱۹۹۹). همچنین بر اساس نظر پژوهشگران نیز کانون، اعتبار صوری بالایی دارد (کاسیو و آگوینیس^{۱۴}، ۲۰۰۵). ۲- محتوا: به این معنا که محتوای تمرین‌ها و ابعاد موجود در آن با شغل هدف مرتبط

بعد از شناخته شدن شایستگی در موفقیت شغلی افراد، پژوهشگران از ابزارها و روش‌های مختلفی برای ارزیابی این شایستگی‌ها استفاده کردند (تورنتن و مولر-هانسون^۱، ۲۰۰۴)؛ اما یکی از این روش‌ها کانون ارزیابی (AC)^۲ نام دارد. روش کانون ارزیابی با عمر پنجاه‌ساله خود (بولر و وهر^۳، ۲۰۰۶) به‌عنوان یک روش ارزیابی (هافمن و ماد^۴، ۲۰۱۲) به سنجش شایستگی‌های شرکت‌کنندگان (اکثرأ مدیران) (دیلچرت و اونس^۵، ۲۰۰۹) توسط تیم ارزیاب‌ها با استفاده از تمرین‌های شبیه‌سازی دلالت دارد (تورنتن، ۱۳۸۶). این روش، حاوی چهار مؤلفه است: اول، شایستگی یا در اصطلاح کانون، ابعاد^۶ (چن^۷، ۲۰۰۶) نام دارد و منظور از آن مجموعه عوامل مرتبط با شغل است که باعث موفقیت فرد در شغلش می‌شود (بورنت و داچ^۸، ۲۰۰۶). دوم، تمرین شبیه‌سازی: نوعی آزمون که در آن یک موقعیت سازمانی یا شغلی، شبیه‌سازی شده و به‌وسیله آن می‌توان رفتار افراد را مورد ارزیابی قرار داد (تورنتن و مولر-هانسون، ۲۰۰۴). سوم، شرکت‌کننده: فردی است که شایستگی‌های او به‌وسیله تمرین‌های شبیه‌سازی مورد ارزیابی قرار می‌گیرد. چهارم، ارزیاب: فردی است که شایستگی‌های شرکت‌کننده را با استفاده از تمرین‌های شبیه‌سازی مورد سنجش قرار می‌دهد (بالانتاین و پوا، ۱۳۸۵). این روش به دلیل قابلیت ارزیابی شایستگی، کاربردهای متعددی برای سازمان‌های مختلف دارد که از آن جمله می‌توان به استفاده از این روش در استخدام، گزینش (یا انتخاب)، ارتقاء، تعیین شغل، ارزیابی عملکرد، انتقال، آموزش و پرورش، توسعه سازمانی، برنامه‌ریزی منابع انسانی و خاتمه خدمت اشاره نمود (خشوعی، ۱۳۹۳). به همین دلیل این روش در کشورهای مختلف و در سازمان‌های خصوصی و دولتی (ملچرز، کلیمن و پرینز^۹، ۲۰۱۰) متعددی مورد استفاده قرار گرفته (تورنتن و گیونز^{۱۰}، ۲۰۰۹) و در ایران نیز در طی ده سال اخیر در سازمان‌هایی مانند گسترش و نوسازی صنایع ایران

1. Thornton, Mueller-Hanson

3. Bowler, Woehr

5. Dilchert, Ones

7. Chen

9. Melchers, Kleinmann & Prinz

11. Schleicher, Day, Mayes & Riggio

13. Kudisch

2. Assessment Center (AC)

4. Hoffman, Meade

6. dimensions

8. Burnett, Dutsch

10. Gibbons

12. fidelity

14. Cascio & Aguinis

از این تحلیل، اعتبار همگرا موقعی حاصل می‌شود که ضرایب یک صفت-چند روش (MTHM)^۹ به اندازه کافی از نظر آماری در حد بالا باشد. اعتبار واگرا نیز موقعی حاصل می‌شود که ضرایب یک صفت-چند روش (MTHM) از ضرایب چند صفت-چند روش (HTHM)^{۱۱} بالاتر باشد و در معیاری سخت‌گیرانه‌تر وقتی که ضرایب چند صفت-یک روش (HTMM)^{۱۱} از ضرایب یک صفت-چند روش (MTHM) کمتر باشد (کولک، بون، وندر فیلر و اولمان^{۱۲}، ۲۰۰۲). در رابطه با اعتبار سازه نیز شواهد حاکی از وجود اعتبار سازه در کانون است (لیوینز و کنوی^{۱۳}، ۲۰۰۱).

بر اساس تحقیقات انجام شده طراحی تمرین‌های خانگی به‌جای استفاده از تمرین‌های پیش‌ساخته، انجام تحلیل شایستگی، وظیفه، سازمان و استخراج شایستگی‌ها و محتوای تمرین‌ها بر اساس این تحلیل‌ها (تورنتن و راپ، ۲۰۰۶)، استفاده از روان‌شناس به‌عنوان ارزیاب (لیوینز، ۲۰۰۲)، استفاده از روش آموزشی چارچوب مرجع (FOR)^{۱۴} (روشی برای حذف معیارهای فردی و دستیابی به یک چارچوب مرجع مشترک) در آموزش ارزیاب‌ها (اسچلیچر و همکاران، ۲۰۰۲)، استفاده از مقیاس امتیازدهی مبتنی بر رفتار (BARS)^{۱۵} (کراس^{۱۶} و تورنتن، ۲۰۰۹) و کاهش تعداد ابعاد مورد سنجش برای هر ارزیاب (بوهر و وهلر، ۲۰۰۶) باعث افزایش اعتبار کانون ارزیابی می‌شود.

پایایی نیز معمولاً به معنای توافق بین ارزیاب‌های مختلف است که از توافق بین آن‌ها حاصل می‌شود (تورنتن و راپ، ۲۰۰۶). تحقیقات حاکی از وجود پایایی بین ارزیاب‌ها^{۱۷} در کانون است (جانسن، لیوینز و کلینمن^{۱۸}، ۲۰۱۱). به عبارت دقیق‌تر پژوهشگران در پژوهش خود به این نتیجه رسیدند که توافق ارزیاب‌ها در شایستگی‌های بین‌فردی مانند مدیریت تعارض و همدلی کمتر از شایستگی‌های عملکردی مانند تصمیم‌گیری و انعطاف‌پذیری است (دمیتز، مانزی کلینمن و سورین^{۱۹}، ۲۰۰۳). برخی نیز به بررسی توافق بین ارزیاب‌ها در هر بُعد در هر تمرین

باشد (تورنتن، ویلسون، جانسون و راجرز^۱، ۲۰۰۹). یکی از بهترین روش‌ها برای اندازه‌گیری اعتبار محتوا نظرخواهی از کارشناسان در زمینه مرتبط بودن محتوای تمرین و ابعاد با وظایف موجود در شغل هدف است (ریجو، میس و اسچلیچر، ۲۰۰۳). در رابطه با اعتبار محتوا نیز تحقیقات حاکی از مرتبط بودن محتوای تمرین‌ها با شرایط شغلی و مشابه بودن ابعاد مورد سنجش در کانون با عملکرد شغلی موفقیت‌آمیز در دنیای واقعی است (وهر و آرتور^۲، ۲۰۰۳). به همین علت نیز شواهد، حاکی از وجود اعتبار محتوا در کانون است (کاهون، بولر و بولر^۳، ۲۰۱۲). ۳- ملاکی: به معنی ارتباط بین امتیاز حاصل از کانون با سایر معیارهای عملکردی مرتبط است (کاهون و همکاران، ۲۰۱۲). این اعتبار شامل دو اعتبار همزمان و پیش‌بین است (تورنتن و راپ، ۲۰۰۶). در رابطه با اعتبار همزمان نتایج حاکی از وجود اعتبار همزمان کانون است (گاگلر، روزتال، تورنتن و بنسون^۴، ۱۹۸۷) و البته اعتبار همزمان کانون با امتیازدهی سرپرست از اعتبار همزمان مصاحبه رفتاری و امتیاز سرپرست بیشتر است (هارل، آردیتی-وگل و جانز^۵، ۲۰۰۳). در زمینه اعتبار پیش‌بین نیز تحقیقات، همگی حاکی از اعتبار پیش‌بین متوسط تا بالا برای کانون است (اسچلیچر و همکاران، ۲۰۰۲). ۴- سازه: یعنی آیا عناصری که کانون را می‌سازند، مناسب بوده و به روشنی تعریف شده‌اند (بالانتاین و پوا، ۱۳۸۵). این اعتبار در صورتی حاصل می‌شود که اعتبار همگرا و اعتبار واگرا (یا افتراقی) ایجاد شود. اعتبار واگرا به معنی همبستگی پایین بین شایستگی‌های مختلف در یک تمرین و اعتبار همگرا به معنی همبستگی بالا بین شایستگی‌های یکسان در تمرین‌های مختلف است (کوک و کریپس^۶، ۱۳۸۸). برای ارزیابی اعتبار سازه کانون نیز در اکثر پژوهش‌ها از تحلیل چند صفت-چند روش (MTMM)^۷ استفاده شده است. در این ماتریس، ابعاد به‌عنوان صفت و تمرین‌ها به‌عنوان روش در نظر گرفته می‌شوند (لانس، نیوبولت، گیتوود، فاستر، فرنچ و اسمیت^۸، ۲۰۰۰). با استفاده

1. Thornton, Wilson, Johnson & Rogers

3. Cahoon, Bowler & Bowler

5. Harel, Arditi-Vogel, Janz

7. Multi-Trait Multi-Method (MTMM) analysis

9. Mono-Trait Hetero-Method (MTHM)

11. Hetero-Trait Mono-Method (HTMM)

13. Conway

15. Behaviorally Anchored Rating Scales (BARS)

17. inter-rater reliability

19. Damitz, Manzey, Severin

2. Arthur

4. Gaugler, Rosenthal, Thornton & Bentson

6. Cook & Cripps

8. Lance, Newbolt, Gatewood, Foster, French, Smith

10. Hetero-Trait Hetero-Method (HTHM)

12. Kolk, Born, van der Flier & Olman

14. Frame Of Reference (FOR)

16. Krause

18. Jansen, Kleinmann

از نظرات آن‌ها برای سنجش استاندارد بودن کانون و اعتبار محتوای آن و از نمرات آن‌ها در تمرین‌ها برای سنجش اعتبار سازه و همزمان و پایایی‌ها استفاده شده است. از نظر حجم نمونه نیز با توجه به سه نکته، این تعداد، انتخاب شدند: اول، یک کانون ارزیابی ارتقایی معمولاً نیم روز یا یک روز به طول می‌انجامد (تورنتن و راپ، ۲۰۰۶). با توجه به برآورد زمانی، ارزیابی ۸ نفر، یک روز به طول می‌انجامد. دوم: با افزایش تعداد تمرین‌ها، مدت زمان اجرای کانون افزایش می‌یابد، با توجه به این‌که در این پژوهش در مقایسه با فراتحلیل‌های انجام شده در این زمینه (کراس و تورنتن، ۲۰۰۹) از تعداد تمرین بیشتری استفاده شده است، لذا انتخاب تعداد بیشتر شرکت‌کنندگان، اجرای کانون را طولانی نموده، بنابراین پژوهشگران از تعداد بیشتر صرف نظر نمود. سوم: در یکی از تمرین‌های این کانون به نام بحث گروهی لازم است تا شرکت‌کنندگان در قالب گروه دور هم جمع شده و در رابطه با موضوع ارائه شده در تمرین با یکدیگر به بحث پردازند. بر اساس مبانی نظری در این نوع تمرین، گروه‌های کمتر از چهار نفر و بیشتر از شش نفر مناسب نیست (بالانتاین و پوا، ۱۳۸۵). از آنجایی که اجرا و اداره کردن گروه شش نفری، دشوار است پژوهشگر به دو گروه چهارنفری (در مجموع ۸ نفر) قناعت نمود و افراد را به صورت تصادفی در گروه‌ها قرار داد.

۲- کلیه کارشناسان ارشد روان‌شناسی صنعتی و سازمانی که از این جامعه، تعداد ۱۳ نفر کارشناسان ارشد روان‌شناسی صنعتی و سازمانی به صورت داوطلب و نمونه‌گیری هدفمند از نوع همگون انتخاب شدند. لازم به ذکر است که از این کارشناسان قبل از آموزش به‌عنوان کارشناس در زمینه سنجش استاندارد بودن و اعتبار صوری کانون نظرخواهی شد. سپس همه آنان مورد آموزش قرار گرفته و ۴ نفر از آنان به‌عنوان ارزیاب ذخیره و ۹ نفر از آنان به‌عنوان ارزیاب اصلی انتخاب شدند و از نظرات ارزیاب‌های اصلی برای سنجش اعتبار صوری کانون و از امتیازدهی آنان برای سنجش اعتبار سازه و پایایی‌ها استفاده شد. این ۹ نفر به مدت تقریباً ۲۳ ساعت به روش آموزش چارچوب مرجع (FOR) و مبتنی بر راهنمای بین‌المللی گروه کاری کانون ارزیابی (۲۰۰۹) و

پرداخته‌اند که بر اساس بررسی این پژوهشگران، میانگین میزان پایایی، ۰/۸۴ است (لانسی، فاستر، نامت، جنتری و درولینگر^۱، ۲۰۰۷). برخی نیز به این نتیجه رسیدند که به ترتیب تمرین ارائه، تعاملی، بحث گروهی، کازیه و بازی تجاری از بیشترین توافق برخوردار است (تزینر و دولان^۲، ۱۹۸۲). بر اساس دیدگاه صاحب‌نظران، آموزش ارزیاب‌ها و استفاده از روش چارچوب مرجع در آموزش ارزیاب‌ها (اسچلیچر و همکاران، ۲۰۰۲) باعث افزایش پایایی کانون ارزیابی می‌شود.

در مجموع، به‌منظور بهره‌مندی از کاربردهای متعدد کانون ارزیابی باید ویژگی‌های روان‌سنجی آن مورد تأیید قرار گیرد و بنابراین اهمیت این موضوع و عدم بررسی این ویژگی در ایران، ضرورت انجام این پژوهش است. بنابراین، پرسش‌های این پژوهش عبارت است از:

۱. آیا کانون ارزیابی طراحی شده استاندارد است؟
۲. آیا کانون ارزیابی طراحی شده از اعتبار لازم برخوردار است؟
۳. آیا کانون ارزیابی طراحی شده از پایایی مطلوبی برخوردار است؟

روش پژوهش

روش پژوهش، جامعه آماری و نمونه: در این پژوهش از دو جامعه و به تبع آن دو نمونه استفاده شده است: ۱- اولین جامعه آماری شامل کلیه مدیرانی بود که در سال ۱۳۹۱ در شرکت گاز استان اصفهان مشغول به خدمت بودند. از بین این مدیران، به روش نمونه‌گیری هدفمند یا غیراحتمالی^۳ و از نوع نمونه‌گیری همگون^۴ تعداد ۸ نفر انتخاب شدند. در این روش، افرادی انتخاب می‌شوند که متعلق به یک زیرفرهنگ بوده یا خصوصیات مشترکی دارند (رنجبر و همکاران، ۱۳۹۱). در این پژوهش نیز همگی از مدیران میانی شرکت گاز استان اصفهان بودند. البته در انتخاب این افراد که به کمک واحد تحقیقات شرکت گاز انجام گرفت این مورد نیز لحاظ گردید که به دلیل حجم زیاد فعالیت مورد نیاز، آن‌ها خود داوطلب به همکاری باشند. لازم به ذکر است که این مدیران به‌عنوان شرکت‌کننده در این پژوهش محسوب شده و

1. Lance, Foster Nemeth, Gentry & Drollinger
3. purposive sampling

2. Tziner, Dolan
4. homogeneous sampling

تعیین الگوی شایستگی شرکت (خشوعی، عریضی، نوری و جهانبازی، ۱۳۹۲) حاوی ۹ بُعد یا شایستگی مرتبط با شغل (حل مشکل، برنامه‌ریزی، آگاهی بین‌فردی، انعطاف‌پذیری، نتیجه‌محوری، مربی‌گری، ارتباط شفاهی، ارتباط نوشتاری و ارتباط شنیداری) و ۶ تمرین شبیه‌سازی مرتبط با شغل (مطالعه موردی، ارائه شفاهی، بحث گروهی، کشف حقیقت، کازیه و تعاملی)، ۸ شرکت‌کننده (مدیران میانی شرکت گاز) و ۹ ارزیاب (کارشناسان ارشد روان‌شناسی صنعتی و سازمانی) بود که در آن شایستگی‌های شرکت‌کنندگان توسط ارزیاب‌ها به‌وسیله تمرین مورد ارزیابی و برای هر شرکت‌کننده امتیاز هر بُعد در هر تمرین شبیه‌سازی، امتیاز کلی بُعد، امتیاز هر تمرین و امتیاز کلی عملکرد تعیین می‌گردید (خشوعی، عریضی و نوری، ۱۳۹۳).

در این پژوهش از چهار ابزار به شرح زیر استفاده شد:

۱- پرسشنامه ابعاد مدیریتی (MDQ)^۱: این پرسشنامه توسط پژوهشگران در فرم دیگر ارزیابی (همکار و سرپرست) با ۵۱ سؤال و با هدف سنجش ۹ بُعد مورد سنجش در کانون طراحی گردید (خشوعی، ۱۳۹۱). فرم مذکور بر اساس ۳ پرسشنامه معتبر و پایا تدوین شده است: ۱- پرسشنامه پایا و معتبر هوش هیجانی (EQ-I)^۲ بار-آن^۳، برای ابعاد انعطاف‌پذیری و حل مسئله (گل‌پرور و خاکسار، ۱۳۸۹). ۲- پرسشنامه پایا و معتبر ارزیابی عملکرد ۳۶۰ درجه برای بُعد برنامه‌ریزی (نوری و همکاران، ۱۳۸۴). ۳- پرسشنامه پایا و معتبر وظایف مدیریتی (MPQ)^۴ برای ابعاد آگاهی بین‌فردی، مربی‌گری، ارتباط شفاهی، ارتباط نوشتاری و ارتباط شنیداری (نواک^۵، ۱۹۸۹). ضمناً برای استفاده از MPQ، نخست، ابعاد مذکور به فارسی ترجمه و مجدد توسط فرد مسلط به انگلیسی برگردانده شد. بدین ترتیب فرم فارسی برای ارزیابی در ایران آماده گردید. پایایی این ابعاد نیز از طریق اجرای آن بر روی ۲۰ نفر از مدیران شرکت گاز استان اصفهان با روش آلفای کرونباخ مطلوب برآورد شد (خشوعی، ۱۳۹۱). ضمناً به‌منظور طراحی این پرسشنامه، برحسب نیاز، برخی از عبارات‌های پرسشنامه‌ها تغییر یافت. همچنین در طیف لیکرتی هر یک از سؤالات

سایر منابع (تورنتن و مولر- هانسون، ۲۰۰۴؛ تورنتن و راپ، ۲۰۰۶؛ بالانتاین و پوا، ۱۳۸۵؛ تیلور^۱، ۲۰۰۷) تحت آموزش نظری و عملی (تورنتن و مولر- هانسون، ۲۰۰۴) قرار گرفتند. هدف این آموزش، آشنایی ارزیاب‌ها با وظایف خود یعنی، مشاهده رفتار، ثبت رفتار، طبقه‌بندی رفتار و امتیازدهی رفتار شرکت‌کنندگان یعنی امتیاز یک بُعد در یک تمرین (PEDR)^۲ امتیاز نهایی بُعد (ODR)^۳، امتیاز تمرین شبیه‌سازی (OER)^۴ و امتیاز کلی ارزیابی (OAR)^۵ (لانس و همکاران، ۲۰۰۷) در تمرین‌ها بود. از نظر حجم نمونه نیز با توجه به سه نکته، این تعداد، انتخاب شدند: اول، در این پژوهش با توجه به مزایای روش درون‌بُعدی (خشوعی، ۱۳۹۳) از این روش امتیازدهی استفاده شد، یعنی هر ارزیاب، بعد از پایان تمام تمرین‌های شبیه‌سازی و پس از بحث با دیگران اقدام به امتیازدهی می‌نماید. همچنین در هنگام مشاهده و ارزیابی نیز هر ارزیاب از بین ابعاد موجود در هر تمرین، با تأکید بر یک بُعد، به مشاهده و ارزیابی یک شرکت‌کننده (یا شرکت‌کنندگان) می‌پردازد. دوم، بر اساس پژوهش‌های موجود معمولاً نسبت ارزیاب به شرکت‌کنندگان ۱ به ۲ است (بالانتاین و پوا، ۱۳۸۵). این ملاک در صورتی حاکم است که در هر تمرین، یک ارزیاب برای هر شرکت‌کننده بیش از یک بُعد را مورد ارزیابی قرار دهد. در این پژوهش چون هر ارزیاب، در هر شرکت‌کننده یک بُعد را مورد مشاهده و ارزیابی قرار می‌دهد تعداد ارزیاب‌ها متناسب با ابعاد افزایش یافته است. سوم، یکی دیگر از عوامل مؤثر بر تعداد ارزیاب‌ها، تعداد ابعاد مورد ارزیابی در هر تمرین شبیه‌سازی است (تورنتن و راپ، ۲۰۰۶).

ابزار سنجش

قبل از ارائه ابزار سنجش لازم به ذکر است که سنجش ویژگی‌های روان‌سنجی زمانی امکان‌پذیر است که ابتدا کانون ارزیابی طراحی شود. بنابراین کانون مورد استفاده در پژوهش، کانون ارزیابی ارتقایی مدیران در شرکت گاز استان اصفهان بود. این کانون پس از بررسی شایستگی‌های مدیریتی (خشوعی، عریضی و نوری، ۲۰۱۳)، تحلیل شرکت گاز و

1. Taylor
3. Overall Dimension Ratings (ODR)
5. Overall Assessment Rating (OAR)
7. Emotional Quotient Inventory (EQ-I)
9. Management Practices Questionnaire (MPQ)

2. Postexercise Dimension Ratings (PEDR)
4. Overall Exercise Rating (OER)
6. Managerial Dimensional Questionnaire (MDQ)
8. Bar-On
10. Nowak

۳- پرسشنامه اثربخشی تمرین شبیه‌سازی: این پرسشنامه بر اساس مطالعه منابع مرتبط با کانون ارزیابی (مانند تورنتن و مولر- هانسون، ۲۰۰۴) توسط پژوهشگران در دو نسخه «الف» با عنوان «پرسشنامه اثربخشی از دیدگاه کارشناسان و ارزیاب‌ها» و «ب» با عنوان «پرسشنامه اثربخشی از دیدگاه شرکت‌کننده» طراحی شد. در نسخه «الف» این پرسشنامه برای تمرین‌های مطالعه موردی، کازیه، بحث گروهی، ارائه شفاهی، حاوی ۷ سؤال با طیف پنج‌درجه‌ای لیکرتی بسیار زیاد (۵)، زیاد (۴)، متوسط (۳) کم (۲) و بسیار کم (۱) و یک سؤال به‌صورت باز پاسخ (اظهارنظر و پیشنهاد) و برای تمرین‌های کشف حقیقت و تعاملی حاوی ۸ سؤال با طیف پنج‌درجه‌ای لیکرتی و یک سؤال به‌صورت باز پاسخ (اظهارنظر و پیشنهاد) بود. نسخه «ب» نیز برای تمرین‌های مطالعه موردی، کازیه، ارائه شفاهی، حاوی ۹ سؤال با طیف پنج‌درجه‌ای لیکرتی کاملاً موافقم (۵)، موافقم (۴)، نظری ندارم (۳) مخالفم (۲) و کاملاً مخالفم (۱) و یک سؤال به‌صورت باز پاسخ (اظهارنظر و پیشنهاد) و برای تمرین‌های کشف حقیقت، تعاملی و بحث گروهی حاوی ۱۰ سؤال با طیف پنج‌درجه‌ای لیکرتی و یک سؤال به‌صورت باز پاسخ است. اعتبار صوری این پرسشنامه با نظرخواهی از ۵ نفر از روان‌شناس صنعتی و سازمانی با درجه کارشناسی ارشد مورد تأیید قرار گرفت. اعتبار محتوای آن نیز با استفاده از نسبت اعتبار محتوا (CVR) و با تأیید پژوهشگران انجام گرفت که چون همه سؤالات در ۵۰ درصد موارد به‌عنوان سؤال ضروری تشخیص داده شد (بونیمی کامپوس و دوپرادو، ۲۰۱۲)، اعتبار محتوای پرسشنامه نیز حاصل شد. پایایی پرسشنامه نیز با استفاده از روش آزمون-بازآزمون برای دو نسخه الف و ب به ترتیب ۰/۸۴ و ۰/۷۸ به دست آمد که نتیجه مطلوبی است. یکی از سؤالات پرسشنامه نسخه «الف» و نسخه «ب» به ترتیب عبارت بود از: راهنمای شرکت‌کننده واضح بود، محتوای تمرین‌ها، امکان بروز شایستگی‌های من را فراهم آورد. لازم به ذکر است که هر یک از نمونه‌های پژوهش متناسب با اهداف پژوهش، به برخی از سؤال‌های این پرسشنامه پاسخ دادند.

۴- مقیاس‌های امتیازدهی رفتاری: پژوهشگران با مطالعه منابع مختلف (مانند تورنتن و مولر- هانسون، ۲۰۰۴؛ تورنتن و راپ، ۲۰۰۶؛ بالانتاین و پوا، ۱۳۸۵) متناسب با ابعاد

پرسشنامه‌ها به غیر از پرسشنامه ارزیابی عملکرد ۳۶۰ درجه و MPQ یک گزینه دیگر با عنوان «غیرقابل مشاهده» اضافه گردید (در این دو پرسشنامه، این طیف موجود بود). با توجه به این‌که در این پرسشنامه ۹ بُعد مورد سنجش قرار می‌گیرد یک نمونه سؤال متناسب با هر بُعد ارائه می‌شود: حل مشکل (هنگام تلاش برای حل یک مشکل، راه‌حل‌های ممکن را در نظر می‌آورد، سپس بهترین را انتخاب می‌کند)، برنامه‌ریزی (فعالیت‌های واحد خود را اولویت‌بندی می‌کند)، آگاهی بین‌فردی (به نیازهای دیگران توجه دارد)، انعطاف‌پذیری (به‌راحتی می‌تواند فکرش را در مورد مسائل تغییر دهد)، نتیجه‌محوری (کارمندان خود را از نظر دستیابی به اهداف سازمانی، تحت نظارت قرار می‌دهد)، مربی‌گری (کارمندان را تحت آموزش و حمایت مناسب، قرار می‌دهد)، ارتباط شفاهی (شیوا و شمرده صحبت می‌کند)، ارتباط نوشتاری (به شیوه‌ای واضح، صریح و مختصر می‌نویسد) و ارتباط شنیداری (در هنگام صحبت افراد بدون قطع صحبت آنان صبورانه به صحبت آنان گوش می‌کند)

۲- پرسشنامه اثربخشی دوره آموزشی: این پرسشنامه که بر اساس مطالعه منابع مرتبط با کانون ارزیابی (مانند تورنتن و مولر- هانسون، ۲۰۰۴؛ تیلور، ۲۰۰۷) توسط پژوهشگران طراحی شد، شامل ۱۴ است که ۱۲ سؤال با طیف پنج‌درجه‌ای لیکرتی، بسیار زیاد (۵)، زیاد (۴)، متوسط (۳) کم (۲) و بسیار کم (۱) و دو سؤال به‌صورت باز پاسخ (اظهارنظر و پیشنهاد) است. اعتبار صوری این پرسشنامه با نظرخواهی از ۵ نفر از روان‌شناس صنعتی و سازمانی با درجه کارشناسی ارشد مورد تأیید قرار گرفت. اعتبار محتوای آن نیز با استفاده از نسبت اعتبار محتوا (CVR)^۱ و با تأیید پژوهشگران انجام گرفت که چون همه سؤالات در ۵۰ درصد موارد به‌عنوان سؤال ضروری تشخیص داده شد (بونیمی کامپوس و دوپرادو، ۲۰۱۲)، اعتبار محتوای پرسشنامه نیز حاصل شد. پایایی پرسشنامه نیز با استفاده از روش آزمون-بازآزمون ۰/۹۱ به دست آمد که نتیجه مطلوبی است. یکی از سؤالات پرسشنامه عبارت بود از: برنامه آموزشی در آشنایی دقیق شما با تمرین‌های شبیه‌سازی چه میزان نقش داشته است. لازم به ذکر است که هر یک از نمونه‌های پژوهش متناسب با اهداف پژوهش، به برخی از سؤال‌های این پرسشنامه پاسخ دادند.

تجزیه و تحلیل داده‌ها

پس از جمع‌آوری داده‌ها از آزمون t تک نمونه‌ای (با رعایت نرمال بودن توزیع) برای سنجش استاندارد بودن کانون، اعتبار صوری و اعتبار محتوای آن و از ضریب همبستگی به‌منظور سنجش اعتبار همزمان، اعتبار سازه و پایایی بین ارزیاب‌ها، پایایی درون‌ارزیاب‌ها، پایایی بین ابعاد یکسان در تمرین‌های مختلف و پایایی مقیاس امتیازدهی رفتاری استفاده شد. لازم به ذکر است که با توجه به شرایط کانون ارزیابی، امکان گزینش تعداد زیادی ارزیاب یا شرکت‌کننده وجود ندارد (تورتن و راپ، ۲۰۰۶)، بنابراین از آنجایی که حجم نمونه در معنی‌دار شدن ضریب همبستگی تأثیرگذار است. لذا ممکن است این ضریب معنی‌دار نشود. در این مواقع، بر اساس دیدگاه صاحب‌نظران، معنی‌دار شدن ضریب چندان مهم نیست، بلکه جهت رابطه یعنی مثبت و منفی بودن آن مهم است (پالانت^۱، ۲۰۱۰) و لذا در این‌گونه موارد، تفسیر نیز بر اساس جهت و میزان رابطه انجام گرفته است.

یافته‌ها

استانداردسنجی: در سه مرحله نظرخواهی انجام گرفت: در مرحله اول، قبل از اجرای کانون از کارشناسان ارشد روان‌شناسی (۱۳ نفر)، در مرحله دوم، قبل از اجرای کانون از کارشناسان ارشد روان‌شناسی (۱۳ نفر) و در مرحله سوم، بعد از اجرای کانون از شرکت‌کنندگان سؤالاتی پرسیده شد (جدول ۱). بر اساس این جدول، در تمامی مراحل، میانگین نمره پاسخ‌دهندگان در هر سؤال از حد متوسط سؤالات (۳) بیشتر است.

بنابراین در پاسخ به سؤال اول پژوهش باید این‌گونه پاسخ داد که با توجه به نتایج دریافت شده (جدول ۱)، کانون ارزیابی طراحی شده در این پژوهش، استاندارد است.

اعتبارسنجی: چهار اعتبار به شرح زیر مورد سنجش قرار گرفت: اعتبار صوری: در دو مرحله نظرخواهی انجام گرفت: در مرحله اول، قبل از اجرای کانون از کارشناسان ارشد روان‌شناسی (۱۳ نفر) و در مرحله دوم بعد از اجرای کانون از ارزیاب‌ها (۹ نفر) سؤالاتی پرسیده شد (جدول ۲). بر اساس این جدول، در تمامی مراحل، میانگین نمره پاسخ‌دهندگان در هر سؤال از حد متوسط سؤالات (۳) بیشتر است. بنابراین بازخورد دریافت شده از هر دو مرحله گویای وجود این اعتبار است.

۹ گانه کانون و این‌که هر یک از ابعاد در دو تمرین شبیه‌سازی مورد ارزیابی قرار گرفتند، اقدام به طراحی ۱۸ مقیاس امتیازدهی رفتاری در طیف پنج‌درجه‌ای لیکرت یعنی ضعیف (۱)، پایین‌تر از انتظار (۲)، در حد انتظار (۳)، بالاتر از حد انتظار (۴) و مطلوب (۵) نمودند. اعتبار صوری این پرسشنامه با نظرخواهی از ۵ نفر از روان‌شناس صنعتی و سازمانی با درجه کارشناسی ارشد مورد تأیید قرار گرفت. اعتبار محتوای آن نیز با استفاده از نسبت اعتبار محتوا (CVR) و با تأیید پژوهشگران انجام گرفت که چون هر ۱۴ سؤال در ۵۰ درصد موارد به‌عنوان سؤال ضروری تشخیص داده شد (بونیمی کامپوس و دوپرادو، ۲۰۱۲) اعتبار محتوای پرسشنامه نیز حاصل شد. از جمله عبارت‌های رفتاری در ۱۸ مقیاس امتیازدهی رفتاری عبارت است از: ارائه بیش از یک راه‌حل (حل مشکل در ارائه شفاهی)، پرسش واضح از شخص منبع (حل مشکل در کشف حقیقت)، اولویت‌بندی کردن امور (برنامه‌ریزی در مطالعه موردی)، انجام تمرین پیش از وقت مقرر (برنامه‌ریزی در کازیه)، بیان احساسات، بدون کلام مانند اخم کردن، لبخند زدن و غیره (آگاهی بین‌فردی در بحث گروهی)، گفتار محترمانه مانند خواهش می‌کنم و لطفاً (آگاهی بین‌فردی در تعاملی)، تغییر گفته‌های خود بر اساس اطلاعات جدید (انعطاف‌پذیری در بحث گروهی)، پذیرش دیدگاه ایفاگر نقش با ارائه دلیل منطقی توسط ایفاگر نقش (انعطاف‌پذیری در تعاملی)، سؤال پرسیدن برای دستیابی به هدف (نتیجه‌محوری در بحث گروهی)، پرهیز از شتاب‌زدگی در دستیابی به نتیجه (نتیجه‌محوری در تعاملی)، ترغیب افراد برای بیان نظراتشان (مربی‌گری در بحث گروهی)، تحسین صحبت در صورت صحت آن (مربی‌گری در تعاملی)، تنظیم صحبت با مهلت زمانی مقرر (ارتباط شفاهی در ارائه شفاهی)، شمرده صحبت کردن (ارتباط شفاهی در کشف حقیقت)، خلاصه نوشتن متن (ارتباط نوشتاری در مطالعه موردی)، عدم وجود خط‌خوردگی (ارتباط نوشتاری در کازیه)، توازن بین سخن گفتن و گوش دادن (ارتباط شنیداری در کشف حقیقت)، خلاصه کردن سخنان ایفاگر نقش (ارتباط شنیداری در تعاملی).

جدول ۱. نتایج مربوط به استاندارد بودن کانون ارزیابی

مرحله	آموزش	سؤال	درصد			شاخص	
			بسیار زیاد	زیاد	متوسط		کم
مرحله اول		آشنایی با شرکت گاز اصفهان	۶۱/۱۵	۰	۳۸/۵	۰	۴/۲۳
		آشنایی با وظایف شرکت کننده	۳۵/۸	۳۸/۵	۱۵/۴	۷/۷	۴/۰۸
		آشنایی با کانون	۷۶/۹	۲۳/۱	۰	۰	۴/۷۷
		آشنایی ابعاد	۶۹/۲	۳۰/۸	۰	۰	۴/۶۹
		آشنایی با تمرین ها	۶۹/۲	۳۰/۸	۰	۰	۴/۶۹
		یادگیری مشاهده رفتار	۵۳/۸	۳۸/۵	۷/۷	۰	۴/۴۶
		یادگیری ثبت رفتار	۶۱/۵	۳۰/۸	۷/۷	۰	۴/۵۴
		یادگیری طبقه بندی رفتار	۴۶/۲	۴۶/۲	۷/۷	۰	۴/۳۸
		یادگیری ارزیابی رفتار	۵۳/۸	۴۶/۲	۰	۰	۴/۵۴
		یادگیری نگارش بازخورد	۵۳/۸	۱۵/۴	۱۵/۴	۷/۷	۴
مرحله دوم		یادگیری شخصی منبع	۷/۷	۶۱/۵	۲۳/۱	۰	۳/۶۱
		یادگیری ایفای نقش	۳۰/۸	۳۸/۵	۱۵/۴	۷/۷	۳/۷۷
		وضوح محتوای تمرین	۵۳/۸	۲۳/۱	۲۳/۱	۰	۴/۴۶
		وضوح راهنمای شرکت کننده	۵۳/۸	۳۰/۸	۱۵/۴	۰	۴/۶۱
		وضوح راهنمای ارزیاب	۴۶/۲	۴۶/۲	۷/۷	۰	۴/۷۷
		وضوح محتوای تمرین	۶۱/۵	۳۸/۵	۰	۰	۴/۶۱
		وضوح راهنمای شرکت کننده	۶۹/۲	۲۳/۱	۷/۷	۰	۴/۶۱
		وضوح راهنمای ارزیاب	۶۹/۲	۳۰/۸	۰	۰	۴/۶۹
		وضوح محتوای تمرین	۴۶/۲	۳۸/۵	۱۵/۴	۰	۴/۳۱
		وضوح راهنمای شرکت کننده	۲۳/۱	۵۳/۸	۲۳/۱	۰	۴
مرحله سوم		وضوح راهنمای ارزیاب	۳۸/۵	۵۳/۸	۷/۷	۰	۴/۳۱
		وضوح محتوای تمرین	۰	۳۸/۵	۴۶/۲	۱۵/۴	۳/۲۳
		وضوح راهنمای شرکت کننده	۴۶/۲	۳۸/۵	۱۵/۴	۰	۴/۳۱
		وضوح راهنمای ارزیاب	۶۹/۲	۱۵/۴	۱۵/۴	۰	۴/۵۴
		وضوح راهنمای شخصی منبع	۳۸/۵	۴۶/۲	۱۵/۴	۰	۴/۳۱
		وضوح محتوای تمرین	۶۱/۵	۱۵/۴	۲۳/۱	۰	۴/۴۶
		وضوح راهنمای شرکت کننده	۶۱/۵	۳۸/۵	۰	۰	۴/۶۱
		وضوح راهنمای ارزیاب	۷۶/۹	۲۳/۱	۰	۰	۴/۷۷
		وضوح محتوای تمرین	۰	۳۸/۵	۴۶/۲	۱۵/۴	۳/۲۳
		وضوح راهنمای شرکت کننده	۰	۶۱/۵	۳۰/۸	۷/۷	۳/۵۴
مرحله سوم		وضوح راهنمای ارزیاب	۲۳/۱	۵۳/۸	۲۳/۱	۰	۴
		وضوح راهنمای ایفاگر نقش	۰	۵۳/۸	۳۸/۵	۷/۷	۳/۴۶
		وضوح محتوای تمرین	۱۲/۵	۸۷/۵	۰	۰	۴/۱۲
		وضوح راهنمای تمرین	۱۲/۵	۸۷/۵	۰	۰	۴/۱۲
		وضوح محتوای تمرین	۵۰	۵۰	۰	۰	۴/۵۰
		وضوح راهنمای تمرین	۲۵	۷۵	۰	۰	۴/۲۵
		وضوح محتوای تمرین	۱۲/۵	۸۷/۵	۰	۰	۴/۱۲
		وضوح راهنمای تمرین	۲۵	۷۵	۰	۰	۴/۲۵
		وضوح محتوای تمرین	۱۲/۵	۷۵	۱۲/۵	۰	۴
		وضوح راهنمای تمرین	۲۵	۷۵	۰	۰	۴/۲۵
مرحله سوم		وضوح محتوای تمرین	۲۵	۶۲/۵	۱۲/۵	۰	۴/۱۲
		وضوح راهنمای تمرین	۲۵	۶۲/۵	۰	۱۲/۵	۴

جدول ۲. نتایج مربوط به اعتبار صوری کانون ارزیابی

مرحله	سؤال	نوع تمرین	درصد				
			بسیار زیاد	زیاد	متوسط	کم	بسیار کم
مرحله اول	ارتباط محتوای تمرین با شغل	مطالعه موردی	۶۱/۵	۱۵/۴	۲۳/۱	۰	۰
		ارائه	۵۳/۸	۴۶/۲	۰	۰	۰
		بحث گروهی	۳۸/۵	۴۶/۲	۱۵/۴	۰	۰
		کشف حقیقت	۳۰/۸	۶۱/۵	۷/۷	۰	۰
		کازیه	۶۹/۲	۲۳/۱	۷/۷	۰	۰
		تعاملی	۳۰/۸	۶۹/۲	۰	۰	۰
مرحله دوم	ارتباط ابعاد شغل با تمرین	کل تمرین‌ها	۳۳/۳	۴۴/۴	۲۲/۲	۰	۰
		کل تمرین‌ها	۵۵/۶	۳۳/۳	۱۱/۱	۰	۰

تمامی مراحل، میانگین نمره پاسخ‌دهندگان در هر سؤال از حد متوسط سؤالات (۳) بیشتر است. بنابراین بازخورد دریافت شده از هر دو مرحله گویای وجود این اعتبار است.

اعتبار محتوا: در دو مرحله نظرخواهی انجام گرفت: در مرحله اول، در زمان اجرای کانون از شرکت‌کنندگان (۸ نفر) و در مرحله دوم بعد از اجرای کانون از شرکت‌کنندگان سؤالاتی پرسیده شد (جدول ۳). بر اساس این جدول، در

جدول ۳. نتایج مربوط به اعتبار محتوای کانون ارزیابی

مرحله	سؤال	نوع تمرین	درصد				
			کاملاً موافقم	موافقم ندارم	نظری مخالفم	کاملاً مخالفم	میانگین
مرحله اول	ارتباط محتوای تمرین با شغل	مطالعه موردی	۳۷/۵	۳۷/۵	۱۲/۵	۰	۴
		ارائه	۰	۲۵	۶۲/۵	۱۲/۵	۴/۱۲
		بحث گروهی	۱۲/۵	۶۲/۵	۲۵	۰	۳/۸۷
		کشف حقیقت	۰	۶۲/۵	۳۷/۵	۰	۳/۶۲
		کازیه	۲۵	۷۵	۰	۰	۴/۲۵
		تعاملی	۲۵	۵۰	۱۲/۵	۱۲/۵	۳/۸۷
مرحله دوم	ارتباط ابعاد شغل با تمرین	کل تمرین‌ها	۵۰	۳۷/۵	۱۲/۵	۰	۴/۳۷
		کل تمرین‌ها	۷۵	۱۲/۵	۱۲/۵	۰	۴/۶۲

اعتبار همزمان: از طریق ارتباط بین امتیاز کلی بُعد با امتیاز حاصل از پرسشنامه ابعاد مدیریتی (ارزیابی همکاران و ارزیابی سرپرست) مورد سنجش قرار گرفت (جدول ۴). بر اساس این جدول، میانگین کل ضریب همبستگی یا رابطه بین ارزیابی همکار و ارزیابی سرپرستها برابر با ۰/۳۸ و میانگین کل رابطه بین ارزیابی سرپرست و ارزیابی ارزیابها برابر ۰/۵۲ است. لذا کانون از اعتبار همزمان متوسطی برخوردار است.

جدول ۴. نتایج مربوط به اعتبار همزمان کانون ارزیابی

ارزیابی سرپرست		ارزیابی همکار		نوع ارزیابی		
r	انحراف معیار	میانگین	r	انحراف معیار	میانگین	شاخص بُعد
۰/۲۵	۰/۷۷	۴/۴۰	۰/۴۱	۱/۰۹	۴/۷۵	حل مشکل
*۰/۷۲	۰/۴۷	۴/۰۷	۰/۲۶	۰/۲۵	۴/۴۶	برنامه‌ریزی
۰/۶۷	۰/۶۳	۴/۸۲	۰/۲۷	۰/۷۶	۵/۸۰	آگاهی بین فردی
۰/۵۴	۱/۲۸	۲۳/۳۱	۰/۵۲	۲/۶۶	۲۶/۶۲	انعطاف‌پذیری
**۰/۸۴	۰/۲۸	۴/۲۳	۰/۳۲	۰/۴۴	۴/۴۴	نتیجه‌محوری
۰/۳۶	۰/۷۰	۴/۵۹	۰/۵۵	۰/۸۱	۵/۸۷	مربی‌گری
۰/۴۹	۰/۴۴	۴/۲۲	۰/۴۶	۰/۵۹	۴	ارتباط شفاهی
۰/۵۶	۰/۷۸	۴/۲۰	۰/۵۹	۰/۵۳	۶/۲۵	ارتباط نوشتاری
۰/۲۷	۰/۴۳	۴/۲۰	۰/۱۲	۰/۴۸	۳/۲۲	ارتباط شنیداری

*P < ۰/۰۵ **P < ۰/۰۱

اعتبار سازه: در این پژوهش برای تعیین اعتبار سازه از تعیین اعتبار همگرا و اعتبار واگرا استفاده شد. به منظور ارزیابی اعتبار همگرا و واگرا نیز ماتریس چند صفت - چند روش (HTHM) مورد استفاده قرار گرفت (جدول ۵). در این جدول، منظور از صفت همان بُعد و روش همان تمرین است، روابط یک صفت - چند روش (MTHM) در داخل پراکنش نوشته شده است، روابط چند صفت - چند روش (HTHM) به صورت زیرخط دار نوشته شده است، روابط چند صفت - یک روش (HTMM) به صورت پررنگ نوشته شده است.

بر اساس جدول ۶ نیز میانگین ضرایب همبستگی یک صفت - چند روش (MTHM) به اندازه کافی از نظر آماری در حد بالا است (r = ۰/۵۴)، بنابراین کانون از اعتبار همگرا برخوردار است. همچنین میانگین ضرایب یک صفت - چند روش (HTHM) بالاتر است بنابراین کانون از اعتبار واگرا نیز برخوردار است.

بنابراین در پاسخ به سؤال دوم پژوهش باید به این گونه پاسخ داد که با توجه به نتایج دریافت شده، کانون ارزیابی طراحی شده در این پژوهش، از اعتبار صوری (جدول ۲)، اعتبار محتوا (جدول ۲)، اعتبار همزمان (جدول ۳) و اعتبار سازه یعنی اعتبار همگرا و واگرا (جدول ۴ و ۵) برخوردار است.

بر اساس جدول ۶ نیز میانگین ضرایب همبستگی یک صفت - چند روش (MTHM) به اندازه کافی از نظر آماری در حد بالا است (r = ۰/۵۴)، بنابراین کانون از اعتبار همگرا

جدول ۵. نتایج اولیه مربوط به اعتبار سازه کانون ارزیابی

تمرین - ابعاد	مطالعه موردی		ارائه شفاهی		بحث گروهی		کشف حقیقت		کازیه		تعاملی								
	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	
مطالعه موردی																			
۱. برنامه‌ریزی																			
۲. ارتباط نوشتاری		۰/۵۲																	
ارائه شفاهی																			
۳. حل مشکل			۰/۳۲	-۰/۲۲															
۴. ارتباط شفاهی			-۰/۲۳	-۰/۲۹	۰/۳۴														
بحث گروهی																			
۵. انعطاف‌پذیری			-۰/۰۸	-۰/۰۶	-۰/۷۱	۰/۰۴													
۶. آگاهی بین‌فردی			۰/۲۳	۰/۰۱	۰/۱۳	۰/۶۴	۰/۴۴												
۷. نتیجه‌محوری			-۰/۴۴	-۰/۶۷	-۰/۲۱	۰/۴۶	۰/۰۸												
۸. مربی‌گری			۰/۲۷	-۰/۰۱	۰/۳۶	۰/۸۲	۰/۲۱												
کشف حقیقت																			
۹. حل مشکل			۰/۳۸	۰/۳۵	(۰/۵۵)	۰/۵۴	-۰/۴۵	۰/۲۰	-۰/۰۸	۰/۵۸									
۱۰. ارتباط شنیداری			-۰/۱۵	-۰/۴۴	۰	۰/۲۹	۰/۵۴	۰/۲۷	۰/۷۹	۰/۳۲	-۰/۲۶								
۱۱. ارتباط شفاهی			۰/۲۰	۰/۱۴	۰	(۰/۶۸)	۰/۳۰	۰/۴۹	۰/۶۶	۰/۷۳	۰/۶۶	۰/۲۳							
کازیه																			
۱۲. برنامه‌ریزی			(۰/۴۰)	۰/۱۶	۰/۲۱	-۰/۳۹	-۰/۱۳	۰/۲۲	-۰/۳۳	-۰/۱۵	-۰/۲۸	-۰/۵۸							
۱۳. ارتباط نوشتاری			۰/۷۷	(۰/۷۰)	-۰/۲۸	-۰/۵۳	۰/۳۳	-۰/۰۷	-۰/۱۱	۰/۰۸	-۰/۲۲	۰/۱۳							
تعاملی																			
۱۴. انعطاف‌پذیری			۰/۲۳	۰/۱۶	-۰/۵۳	-۰/۵۲	(۰/۶۲)	-۰/۰۵	-۰/۴۲	-۰/۲۵	-۰/۴۲	۰/۳۱	-۰/۰۴	۰/۰۹	۰/۵۸				
۱۵. آگاهی بین‌فردی			۰/۲۰	۰/۱۴	۰	-۰/۰۹	۰/۳۰	(۰/۳۷)	۰/۰۸	۰/۲۷	-۰/۱۳	۰/۲۳	-۰/۱۴	-۰/۰۱	-۰/۲۳	۰/۶۸	۰/۳۸		
۱۶. نتیجه‌محوری			۰/۵۷	۰/۱۲	-۰/۰۸	۰/۰۷	۰/۶۰	۰/۴۹	(۰/۲۵)	۰/۴۸	۰/۱۳	۰/۵۲	۰/۵۲	-۰/۰۱	۰/۵۷	۰/۳۳	-۰/۲۹	۰/۳۰	
۱۷. مربی‌گری			۰/۴۱	-۰/۲۲	۰/۶۹	۰/۶۰	-۰/۲۰	۰/۵۶	۰/۵۴	(۰/۶۶)	۰/۵۴	۰/۲۲	۰/۴۸	۰/۱۵	-۰/۱۲	-۰/۳۳	-۰/۲۹	۰/۳۰	
۱۸. ارتباط شنیداری			-۰/۲۱	-۰/۲۷	۰/۲۸	۰/۴۷	-۰/۱۵	-۰/۲۰	۰/۴۵	۰/۲۳	۰/۴۵	(۰/۴۷)	۰/۵۲	-۰/۸۰	-۰/۲۵	-۰/۱۹	-۰/۲۸	۰/۳۰	

جدول ۶. نتایج نهایی مربوط به اعتبار سازه کانون ارزیابی

میانگین r	ضرایب یک صفت چند روش	میانگین r	ضرایب یک صفت چند روش
۰/۵۲	مطالعه موردی	۰/۵۵	حل مشکل
۰/۳۴	ارائه شفاهی	۰/۶۸	برنامه‌ریزی
۰/۳۸	بحث گروهی	۰/۲۷	آگاهی بین‌فردی
۰/۲۱	کشف حقیقت	۰/۶۲	انعطاف‌پذیری
۰/۱۳	کازیه	۰/۲۵	نتیجه‌محوری
۰/۰۹	تعاملی	۰/۶۶	مربی‌گری
۰/۲۷	میانگین کل ضرایب تمرین‌ها	۰/۶۸	ارتباط شفاهی
میانگین r	ضرایب چند صفت چند روش	۰/۷۰	ارتباط نوشتاری
۰/۱۰	میانگین کل ضرایب ابعاد	۰/۴۷	ارتباط شنیداری
		۰/۵۴	میانگین کل ضرایب ابعاد

پایایی سنجی

به ذکر است که برای محاسبه پایایی در ابعاد به تفکیک هر تمرین و پایایی در هر تمرین، در تمرین‌های نوشتاری یعنی مطالعه موردی و کازیه از پایایی بین ارزیاب‌ها استفاده شد؛ اما در بحث گروهی، ارائه شفاهی، تعاملی و کشف حقیقت از پایایی بین ارزیاب با ارزیاب ملاک (پژوهشگر اول) استفاده شد زیرا به دلیل اجرای همزمان تمرین‌ها امکان ارزیابی تمام تمرین‌ها برای ارزیاب ملاک فراهم نبود.

پایایی از طریق پنج روش یعنی پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در ابعاد به تفکیک هر تمرین، پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در هر تمرین، پایایی درون ارزیاب‌ها، پایایی بین ابعاد یکسان در تمرین‌های مختلف و پایایی مقیاس امتیازدهی رفتاری مورد سنجش قرار گرفت (جدول ۷). لازم

بر اساس جدول ۷ نتایج پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در ابعاد به تفکیک هر تمرین نشان داد که ضرایب همبستگی از بالاترین رابطه یعنی ۰/۹۸ در بُعد ارتباط شنیداری تا کمترین رابطه یعنی ۰/۳۳ در بُعد ارتباط شفاهی در تمرین کشف حقیقت متغیر است. در مجموع، در تمامی ابعاد به غیر از بُعد ارتباط شفاهی در تمرین کشف حقیقت، این ضریب متوسط و بالا است. نتایج پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در هر تمرین نیز نشان داد که ضرایب همبستگی از بالاترین رابطه یعنی ۱ در تمرین ارائه شفاهی تا ۰/۵۵ در تمرین کازیه متغیر است. نتایج پایایی درون ارزیاب‌ها نیز نشان داد میزان پایایی برای تمرین مطالعه موردی، ۰/۸۵ و

برای تمرین کازیه ۰/۸۸ است. این نتایج حاکی از پایایی درون ارزیاب بالایی برای این دو تمرین است. پایایی مقیاس‌های امتیازدهی رفتاری نیز تأییدی دیگر برای پایایی کانون است که نتایج آن حاکی از بالاترین رابطه یعنی ۰/۹۶ در بُعد ارتباط شنیداری در تمرین کشف حقیقت تا کمترین رابطه یعنی ۰/۴۲ در بُعد آگاهی بین فردی در تمرین تعاملی است. در مجموع، در تمامی مقیاس‌ها این ضریب متوسط و بالا است. نتایج پایایی بین ابعاد یکسان نیز حاکی از بالاترین رابطه یعنی ۰/۷۰ در ارتباط نوشتاری تا کمترین رابطه یعنی ۰/۲۷ در آگاهی بین فردی متغیر است. این ضریب، در تمامی ابعاد به غیر از آگاهی بین فردی، متوسط و بالا است.

جدول ۷. نتایج مربوط به پایایی کانون ارزیابی

پایایی بین ابعاد یکسان در تمرین‌های مختلف	مقیاس امتیازدهی رفتاری	پایایی درون ارزیاب‌ها	پایایی بین ارزیاب‌ها در هر تمرین	پایایی بین ارزیاب‌ها در ابعاد به تفکیک هر تمرین	بُعد	تمرین
r	r	r	r	r	r	r
۰/۵۵	۰/۵۴	۰/۸۵	۰/۶۷	۰/۴۸	برنامه‌ریزی	مطالعه
۰/۴۰	***۰/۸۴			***۰/۷۳	ارتباط نوشتاری	موردی**
۰/۲۷	۰/۸۱	-	***۱	۰/۸۷	حل مشکل	ارائه شفاهی*
۰/۶۲	۰/۹۴			***۰/۹۷	ارتباط شفاهی	
۰/۲۵	۰/۹۰			۰/۸۷	انعطاف‌پذیری	
۰/۶۶	۰/۵۷	-	۰/۹۱	۰/۵۷	آگاهی بین فردی	بحث
۰/۶۸	۰/۹۴			۰/۸۷	نتیجه‌محوری	گروهی*
۰/۷۰	۰/۹۳			۰/۹۴	مربی‌گری	
۰/۴۷	۰/۹۳			۰/۶۹	حل مشکل	کشف حقیقت*
	***۰/۹۶	-	۰/۸۵	۰/۸۹	ارتباط شنیداری	
	۰/۸۱			۰/۳۳	ارتباط شفاهی	
	***۰/۷۱	۰/۸۸	۰/۵۵	***۰/۸۲	برنامه‌ریزی	کازیه**
	۰/۵۸			***۰/۸۰	ارتباط نوشتاری	
	۰/۹۰			۰/۹۰	انعطاف‌پذیری	
	۰/۴۲			۰/۸۷	آگاهی بین فردی	
	۰/۹۴	-	***۰/۹۸	۰/۹۴	نتیجه‌محوری	تعاملی*
	***۰/۹۵			۰/۵۷	مربی‌گری	
	۰/۸۶			***۰/۹۸	ارتباط شنیداری	

یادداشت: * پایایی از طریق پایایی بین ارزیاب‌ها با ارزیاب ملاک مورد سنجش قرار گرفته است.

** پایایی از طریق پایایی بین ارزیاب‌ها مورد سنجش قرار گرفته است.

***P < ۰/۰۵

****P < ۰/۰۱

این کارشناسان، با مشاغل مورد نظر مرتبط بود. این یافته با نتایج پژوهش‌های قبلی (کاهون و همکاران، ۲۰۱۲؛ وهر و آرتور، ۲۰۰۳) هماهنگ است.

در ارتباط با اعتبار همزمان نیز، یافته‌ها نشان داد که اعتبار همزمان کانون با امتیازدهی سرپرست بیشتر از اعتبار همزمان کانون با امتیازدهی همکار است، اما در مجموع اعتبار همزمان این روش، تقریباً مطلوب است. این یافته با پژوهش هارل و همکاران (۲۰۰۳) مبنی بر این‌که اعتبار همزمان کانون با امتیازدهی سرپرست از اعتبار همزمان مصاحبه رفتاری با امتیازدهی سرپرست بیشتر است، هماهنگ است.

برای سنجش اعتبار سازه نیز مطابق با سایر پژوهش‌ها (کولک و همکاران، ۲۰۰۲) از تعیین اعتبار همگرا و اعتبار واگرا و تحلیل چند صفت - چند روش استفاده شد. نتایج نشان داد که میانگین ضرایب یک صفت - چند روش به اندازه کافی از نظر آماری در حد بالا است، بنابراین کانون از اعتبار همگرا برخوردار بود. همچنین میانگین ضرایب یک صفت - چند روش از ضرایب چند صفت - چند روش یا چند صفت - یک روش بالاتر بود، بنابراین کانون از اعتبار واگرا نیز برخوردار بود. این یافته با پژوهش‌های قبلی در این زمینه نیز هماهنگ است (لیوینز و کنوی، ۲۰۰۱).

در رابطه با پایایی نیز پایایی کانون از روش پایایی ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در ابعاد به تفکیک هر تمرین، پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در هر تمرین، پایایی درون ارزیاب‌ها، پایایی بین ابعاد یکسان در تمرین‌های مختلف و پایایی مقیاس امتیازدهی رفتاری مورد سنجش قرار گرفت. نتایج پایایی کانون از روش پایایی ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در ابعاد به تفکیک هر تمرین، متوسط و بالا بود. این نتیجه با پژوهش لانس و همکاران (۲۰۰۷) هماهنگ است. در مورد پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب‌ها با ارزیاب ملاک) در هر تمرین، به ترتیب، تمرین ارائه شفاهی، تعاملی، بحث گروهی، کشف حقیقت، مطالعه موردی و کازیه از پایایی بالاتری برخوردار بودند. ترتیب این تمرین‌ها با پژوهش تیزر و دولان (۱۹۸۲) مبنی بر بالا بودن پایایی به ترتیب در تمرین ارائه شفاهی، تعاملی، بحث گروهی و کازیه هماهنگ است. در مورد پایایی درون ارزیاب‌ها نیز، یافته‌ها حاکی از پایایی متوسط و بالا در

بنابراین در پاسخ به سؤال سوم پژوهش باید به این‌گونه پاسخ داد که با توجه به نتایج دریافت شده، کانون ارزیابی طراحی شده در این پژوهش از پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در ابعاد به تفکیک هر تمرین، پایایی بین ارزیاب‌ها (و یا پایایی بین ارزیاب با ارزیاب ملاک) در هر تمرین، پایایی درون ارزیاب‌ها، پایایی بین ابعاد یکسان در تمرین‌های مختلف و پایایی مقیاس امتیازدهی رفتاری (جدول ۶) برخوردار است.

بحث و نتیجه‌گیری

این پژوهش با هدف بررسی ویژگی‌های روان‌سنجی کانون ارزیابی ارتقایی مدیران انجام و نتایج زیر حاصل شد:

در رابطه با استاندارد بودن کانون، بازخورد حاصل از کارشناسان ارشد روان‌شناسی و شرکت‌کنندگان گویای وجود استاندارد بودن کانون بود. در خصوص علت این مسئله می‌توان به دستورالعمل واضح برای شرکت‌کننده، ارزیاب، ایفاگر نقش و شخص منبع، آموزش ارزیاب، ایفاگر نقش و شخص منبع، استفاده از یک ایفاگر نقش یکسان برای تمام شرکت‌کنندگان و تصادفی قرار دادن شرکت‌کنندگان در تمرین بحث گروهی اشاره نمود. این یافته با دیدگاه تورنتن و مولر - هانسون (۲۰۰۴) هماهنگ است؛ اما با توجه به عدم وجود پژوهش تجربی در این زمینه، مقایسه دقیق‌تر امکان‌پذیر نیست.

در رابطه با اعتبار صوری کانون، بازخورد حاصل از کارشناسان ارشد روان‌شناسی صنعتی و سازمانی و ارزیاب‌ها، گویای وجود اعتبار صوری کانون بود. بدین معنی که هر یک از نه بُعد موجود در کانون و هر یک از شش تمرین از دیدگاه این کارشناسان، ظاهراً با مشاغل مورد نظر مرتبط بود. این نتایج با سایر پژوهش‌های انجام شده در این زمینه (کودیچ و همکاران، ۱۹۹۹؛ کاسیوو آگوینیس، ۲۰۰۵) هماهنگ است.

برای سنجش اعتبار محتوا نیز مطابق با سایر پژوهش‌ها (ریجو و همکاران، ۲۰۰۳) از کارشناسان نظرخواهی به عمل آمد. در رابطه با این نوع اعتبار، بازخورد حاصل از شرکت‌کنندگان، گویای وجود اعتبار محتوای کانون بود. به عبارت دقیق‌تر، در مجموع، ابعاد و تمرین‌ها از اعتبار محتوا برخوردار بود. بدین معنی که همه ابعاد و تمرین‌ها از دیدگاه

منابع

- ابوالعلائی، ب. و غفاری، ع. (۱۳۸۶). *مدیران آینده: مبانی نظری و تجارب عملی برنامه‌های استعدادیابی و جانشین‌پروری مدیران*. تهران: انتشارات سازمان مدیریت صنعتی.
- بالانتاین، ی. و پوا، ن. (۱۳۸۵). *کانون‌های ارزیابی و توسعه مدیران*. ترجمه مسعود سلطانی. تهران: انتشارات آسیا.
- تورتنن، ج. سی. (۱۳۸۶). *کانون‌های ارزیابی در مدیریت منابع انسانی*. ترجمه سعید جعفری مقدم، مجید سلیمی و سعید شهباز مرادی. تهران: موسسه انتشارات و چاپ دانشگاه تهران.
- خشوعی، م. س. (۱۳۹۱). *طراحی کانون ارزیابی ارتقایی مدیران و بررسی ویژگی‌های روان‌سنجی و فرایندهای آن*. پایان‌نامه دکتری روان‌شناسی دانشگاه اصفهان.
- خشوعی، م. س. (۱۳۹۳). *شبیه‌سازی سازمانی در مدیریت منابع انسانی: از نظریه تا عمل (راهنمای پژوهشگران و سازمان‌ها)*. اصفهان: انتشارات دانشگاه اصفهان.
- خشوعی، م. س.، عریضی، ح. ر. و نوری، ا. (۱۳۹۳). *طراحی و اجرای کانون ارزیابی ارتقایی مدیران (مطالعه موردی: شرکت گاز استان اصفهان)*. *مجله مدیریت منابع انسانی دانشگاه امام حسین (ع)*. ۳، ۴۷-۲۷.
- خشوعی، م. س.، عریضی، ح. ر. و نوری، ا. و جهانبازی، ا. (۱۳۹۲). *طراحی الگوی شایستگی مدیریتی. مطالعات مدیریت راهبردی*، شماره ۱۴، ۱۶۵-۱۴۵.
- رنجبر، ه.، حقدوست، ع. ا.، صلصالی، م.، خوشدل، ع. ر.، سلیمانی، م. ع. و بهرامی، ن. (۱۳۹۱). *نمونه‌گیری در پژوهش‌های کیفی: راهنمایی برای شروع*. *مجله دانشگاه علوم پزشکی ارتش جمهوری اسلامی ایران*، ۱۰ (۳)، ۲۵۰-۲۳۸.
- عبدی، ا. و دیبانتی، م. (۱۳۸۷). *مراکز ارزیابی، روش نوین جانشین‌پروری. تدبیر*، شماره ۳۱، ۱۹۵-۲۶.
- کوک، م. و کریس، ب. (۱۳۸۸). *سنجش و ارزیابی روان‌شناختی در محیط کار (راهنمای مدیران)*. ترجمه جهانبخش صادقی. تهران: انتشارات سمت.
- گل‌پرور، م. و خاکسار، ف. (۱۳۸۹). *رابطه بین هوش هیجانی با عملکرد شغلی کارکنان کارخانجات صنعتی شهر اصفهان*. *دانشور رفتار (روان‌شناسی بالینی و شخصیت)*، ۴۰، ۳۴-۱۹.
- نوری، ا.، عریضی، ح. ر.، مولوی، ح.، کلانتری، م.، نشاط دوست، ح. ط.، شفتی، ع. و موسویان، ح. (۱۳۸۴). *مهارت‌های روان‌شناختی مورد نیاز در مشاغل مدیریتی: حوزه مدیریت شرکت گاز استان اصفهان*. گزارش پژوهش برون‌دانشگاهی بین دانشگاه اصفهان و شرکت ملی گاز ایران.
- Bonini Campos, J. A. D., & Do Prado C, D. d. (2012). *Cross-cultural adaptation of the Portuguese version*

کانون بود. در مورد پایایی بین ابعاد یکسان در تمرین‌های مختلف نیز در تمامی ابعاد به غیر از بُعد آگاهی بین‌فردی، این ضریب متوسط و بالا بود. این نتیجه با پژوهش دمیتز و همکاران (۲۰۰۳) مبنی بر پایین بودن توافق در ابعاد بین‌فردی در مقایسه با سایر ابعاد هماهنگ است. پایایی مقیاس‌های امتیازدهی رفتاری نیز مطلوب بود. بنابراین در مجموع، کانون در این پژوهش از پایایی مطلوبی برخوردار بود. این یافته با پژوهش (جانسن و همکاران، ۲۰۱۱) هماهنگ است.

علاوه بر این، پژوهشگران با بکار بردن تمهیدات سایر پژوهشگران امکان بروز شرایط مذکور را فراهم نمودند. این تمهیدات، عبارت بود از: طراحی تمرین‌های خانگی به‌جای استفاده از تمرین‌های پیش‌ساخته، انجام تحلیل شایستگی، وظیفه، سازمان و استخراج شایستگی‌ها و محتوای تمرین بر اساس این تحلیل‌ها (تورتنن و راپ، ۲۰۰۶)، استفاده از روان‌شناس به‌عنوان ارزیاب (لیوینز، ۲۰۰۲)، استفاده از روش چارچوب مرجع در آموزش (اسچلیچر و همکاران، ۲۰۰۲)، استفاده از مقیاس امتیازدهی رفتاری (کراس و تورتنن، ۲۰۰۹)، کاهش تعداد ابعاد مورد سنجش برای هر ارزیاب (بوهر و وهلر، ۲۰۰۶) و آموزش ارزیاب‌ها (اسچلیچر و همکاران، ۲۰۰۲).

در مجموع، این پژوهش به‌عنوان اولین پژوهش در زمینه اعتباریابی کانون در ایران است علاوه بر این، از نظر پوشش تمامی ویژگی‌های روان‌سنجی کانون در مقایسه با سایر پژوهش‌های انجام شده در خارج از کشور نیز در نوع خود بی‌نظیر است؛ اما این پژوهش با محدودیت‌هایی نیز روبه‌رو بود از جمله این‌که اعتبار پیش‌بین کانون در این پژوهش مورد بررسی قرار نگرفت زیرا کانون طراحی شده در این پژوهش به‌عنوان یک کانون آزمایشی طراحی و اجرا شده بود و لذا امکان انتصاب فرد شایسته وجود نداشت. بنابراین، به سایر پژوهشگران پیشنهاد می‌شود با طراحی کانون و اجرای عملی آن، اعتبار پیش‌بین کانون را مورد بررسی قرار دهند. علاوه بر این، تعداد کم حجم نمونه به دلیل شرایط اجرای کانون و اجرای آن در جامعه مدیران از جمله محدودیت‌های دیگر این پژوهش محسوب می‌شود. بنابراین، به سایر پژوهشگران پیشنهاد می‌شود موضوع این پژوهش را در نمونه‌ای با حجم بیشتر و جامعه‌ای متفاوت یعنی سطوح غیرمدیریتی نیز مورد بررسی قرار دهند.

- of the Patient-Generated Subjective Global Assessment. *Nutrición Hospitalaria*, 27(2): 583-594.
- Bowler, M.C., & Woehr, D. J. (2006). A meta-analytic evaluation of the impact of dimension and exercise factors on assessment center ratings. *Journal of Applied Psychology*, 91(5), 1114-1124.
- Burnett, M., & Dutsch, J. D. (2006). Competency-based training and assessment center: Strategies, technology, process, and issues. *Advances in Developing Human Resources*, 8 (2), 141-143.
- Cahoon, M. V., Bowler, M. C., & Bowler, J. L. (2012). A reevaluation of assessment center construct-related validity. *International Journal of Business and Management*, 7 (9), 3- 19.
- Cascio, W. F., & Aguinis, H. (2005). *Applied Psychology in Human Resource Management*. New Jersey: Prentice Hall.
- Chen, H. C. (2006). Assessment center: A critical mechanism for assessing HRD effectiveness and accountability. *Advances in Developing Human Resources*, 8, 247-264.
- Damitz, M., Manzey, D., Kleinmann, M., & Severin, K. (2003). Assessment Center for Pilot Selection: Construct and Criterion Validity and the Impact of Assessor Type. *Applied Psychology: An International Review*, 52 (2), 193-212.
- Dilchert, S., & Ones, D. S. (2009). Assessment center dimensions: Individual differences correlates and meta-analytic incremental validity. *International Journal of Selection and Assessment*, 17(3), 254-270.
- Gaugler, B. B., Rosenthal, D.B., Thornton, G.C. III., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72, 493-511.
- Harel, G.H., Arditi-Vogel, A., & Janz, T. (2003). Comparing the validity and utility of behavior description interview versus assessment center ratings. *Journal of Managerial Psychology*, 18 (2), 94-104.
- Hoffman, B. J., & Meade, A. (2012). Alternate approaches to understanding the psychometric properties of assessment centers: An analysis of the structure and equivalence of exercise ratings. *International Journal of Selection and Assessment*, 20 (1), 82- 97.
- International Task Force on Assessment Center Guidelines. (2009). Guidelines and ethical considerations for assessment center operations. *International Journal of Selection and Assessment*, 17(3), 243-253.
- Jansen, A., Lievens, F., & Kleinmann, M. (2011). Do individual differences in perceiving situational demands moderate the relationship between personality and assessment center dimension ratings?. *Human Performance*, 24, 231-250.
- Khoshouei, M.S. Oreyzi, H. R. & Nouri, A. (2012 April). *Assessor Errors and Biases: Challenges in Assessment Centers*. Paper presented at the International Conference on Industrial and Organizational Psychology: Trends, Challenges, and Applications, 19.
- Khoshouei, M.S., Oreyzi, H.R., & Nouri, A. (2013). The Eight Managerial Competencies: Essential Competencies for Twenty First Century Managers. *Iranian Journal of Management Studies (IJMS)* .6 (2), 131- 152.
- Krause, D, E., & Thornton III, G.C (2009). A Cross-cultural look at assessment center practices: Survey results from Western Europe and North America. *Applied Psychology: An International Review*, 58(4), 557-585.
- Kolk, N. J., Born, M.P., van der Flier, H., & Olman, J. M. (2002). Assessment center procedures: cognitive load during the observation phase. *International Journal of Selection and Assessment*, 10 (4), 271- 278.
- Kudisch, J. D., Avis, J. M., Fallen, J. D., Thibodeaux, H. E, III, Roberts, F. E., Rollier, T. J., & Rotolo, C. T. (1999, June). *Benchmarking for success: A look at today's assessment center practices worldwide*. Paper presented at the 27th annual meeting of the International Congress on Assessment Center Methods, Orlando, FL.
- Lance, C. E., Foster, M. R., Nemeth, Y. M., Gentry, W. A., & Drollinger, S. (2007). Extending the nomological network of assessment center construct validity: Prediction of cross-situationally consistent and specific aspects of assessment center performance. *Human Performance*, 20, 345-362.
- Lance, C. E., Newbolt, W. H., Gatewood, R. D., Foster, M. R., French, N.R., & Smith, D. E. (2000). Assessment center exercise factors represent cross-situational specificity, not method bias. *Human Performance*, 13, 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, 675-686.
- Lievens, F., & Conway, J. M. (2001). Dimension and exercise variance in assessment center scores: A largescale evaluation of multitrait-multimethod studies. *Journal of Applied Psychology*, 86, 1202-1222.
- Melchers, K.G., Kleinmann, M., & Prinz, M. A. (2010). Do assessors have too much on their plates? The effects of simultaneously rating multiple assessment center candidates on rating quality. *International Journal of Selection and Assessment*, 18 (3), 329-341.
- Nowack, K. (1989). *Administration and software manual for the management practices questionnaire*. Woodland Hills, CA: Organizational Performance Dimensions.
- Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS* (4rd edition). New York: McGraw-Hill.
- Riggio, R. E., Mayes, B. T., & Schleicher, D. J. (2003). Using assessment center methods for measuring undergraduate business student outcomes. *Journal of Management Inquiry*, 12, 68-78.
- Schleicher, D. J., Day, D. V., Mayes, B. T., & Riggio, R. E. (2002). A new frame for frame-of-reference training: Enhancing the construct validity of

- assessment centers. *Journal of Applied Psychology*, 87, 735-746.
- Taylor, I. (2007). *A Practical Guide to Assessment Centres and Selection Method: Measuring Competency for Recruitment and Development*. Kogan Page.
- Thornton, G.C., & Gibbons, A. M. (2009). Validity of assessment centers for personnel selection. *Human Resource Management Review*, 19, 169- 187.
- Thornton, G. C., III., & Mueller-Hanson, R. A. (2004). *Developing organizational simulations: A guide for practitioners and students*. Lawrence Erlbaum Associates.
- Thornton, G. C., III, & Rupp, D. E. (2006). *Assessment centers in human resource management: strategies for prediction, diagnosis and development*. Lawrence Erlbaum Associates.
- Thornton III, G. C., Wilson, C. L., Johnson, R. M., Rogers, D. A. (2009). Managing Assessment Center Practices in the Context of Employment Discrimination Litigation. *The Psychologist-Manager Journal*, 12, 175-186.
- Tziner, A., & Dolan, S. (1982). Validity of an assessment center for identifying future female officers in the military. *Journal of Applied Psychology*, 67, 728-736.
- Woehr, D. J., & Arthur, 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, 231-258.