Research Article

Industrial and Organizational Psychology Studies

Ind Organ Psychol Stud
Spring & Summer 2026, 13(1), 97-110
https://doi.org/10.22055/jiops.2026.50680.1522

Received: 20 September 2025
Received in revised form: 07 April 2026
Accepted: 07 April 2026

The Cognitive Error of the Halo Effect in Classifying Leadership Competency Indicators in Assessment Centers: Comparing the Performance of Industrial and Organizational Psychologist Assessors with Managers

Hamid Reza Oreyzi Samani[1], Naghi Radi Afsouran[2*]
George. C. Thornton III[3], Fatemeh Jaefari Alamdari[4]

1. Professor. Department of Psychology, University of Isfahan, Isfahan, Iran
2. Assistant Professor, Department of Educational Sciences, University of Guilan, Rasht, Iran
3. Emeritus Professor, Department of Psychology, Colorado State University, Fort Collins, Denver. USA
4. Master of Art in Industrial and Organizational Psychology, Department of Psychology, University of Isfahan, Isfahan, Iran

Citation: Oreyzi Samani, H. R., Radi Afsouran, N., Thornton III, G. C. & Jaafari Alamdari, F. (2026). The Cognitive Error of the Halo Effect in Classifying Leadership Competency Indicators in Assessment Centers: Comparing the Performance of Industrial and Organizational Psychologist Assessors with Managers. (e21104). *Industrial and Organizational Psychology Studies*, *13*(1), 97-110 e21104 https://doi.org/10.22055/jiops.2026.50680.1522

Abstract

Introduction

It is clear that all managers and employees working in various organizations and companies are subject to performance appraisal. However, some may feel that their appraisal was inaccurate or biased. One possible cognitive error involved is the halo effect, where a person's other characteristics and performances are judged based on one outstanding characteristic. Since assessment centers are a primary method of performance appraisal in the workplace, this study aimed to compare the cognitive error of the halo effect among industrial and organizational psychology assessors and managers.

Method

The research employed a quantitative strategy with an experimental method. A total of 240 industrial and organizational psychologist assessors, trained in assessment centers courses, and 145 manager assessors with at least five years of experience participated in the study. Data analysis involved the use of the ROC curve and the area under the curve (AUC). Additionally, the odds ratio for correct classification of patient/non-patient and risk factor and non-risk factor indices—converted into a 2×2 psychologist/manager table, correct/incorrect classification, through an analogical transfer of related characteristics—was used.

Results

The study results indicated that assessments conducted by industrial and organizational psychologists are more likely to fall within the correct classification of leadership competence than those conducted by managers. Additionally, the odds ratio for correctly identifying leadership competence indicators among industrial and organizational psychologists is much closer to that of the criterion assessors group. Furthermore, the accuracy of industrial and organizational psychologists in the assessor role closely matches that of the criterion assessors. Finally, industrial and organizational psychologists outperformed managers in assessing leadership competency in assessment centers and exhibited fewer halo effect cognitive errors.

Discussion

The findings of this study can be utilized within the framework of accrediting assessment centers. They

* **Corresponding Author:** Naghi Radi Afsouran
Email: naghi.radi.afsouran@guilan.ac.ir

Authors retain the copyright and full publishing rights.
Published by Shahid Chamran University of Ahvaz. This article is an open access article licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0)

can help organizations design and implement the assessment centers process with greater confidence, maximize the validity of the results, and make more effective decisions with higher execution guarantees.

Keywords: Halo effect, assessment centers, industrial and organizational psychologists, managers, correct classification

مقاله پژوهشی

مطالعات روان‌شناسی صنعتی و سازمانی

بهار و تابستان ۱۴۰۵، ۱۳(۱)، ۱۱۰-۹۷
https://doi.org/10.22055/jiops.2026.50680.1522

تاریخ دریافت: ۱۴۰۴/۰۶/۲۹
تاریخ بازنگری:۱۴۰۵/۰۱/۱۸
تاریخ پذیرش: ۱۴۰۵/۰۱/۱۸

خطای شناختی اثر هاله‌ای در طبقه‌بندی نشانه‌های شایستگی رهبری در کانون‌های ارزیابی: مقایسه عملکرد ارزیابان روان‌شناسان صنعتی و سازمانی با مدیران باسابقه

حمیدرضا عریضی سامانی[۱]، نقی رعدی افسوران[۲*]، جورج. سی. ثورنتون[۳]، فاطمه جعفری علمداری[۴]

۱. استاد، گروه روان‌شناسی، دانشگاه اصفهان، اصفهان، ایران
۲. استادیار، گروه علوم تربیتی، دانشگاه گیلان، رشت، ایران
۳. استاد نمونه، گروه روان‌شناسی، دانشگاه ایالتی کلرادو، فورت کالینز، دنور، آمریکا
۴. کارشناس ارشد روان‌شناسی صنعتی و سازمانی، گروه روان‌شناسی، دانشگاه اصفهان، اصفهان، ایران

چکیده

هدف پژوهش حاضر مقایسه خطای ادراکی اثر هاله‌ای در بین ارزیابان روان‌شناسی صنعتی و سازمانی و مدیران باسابقه بود. در قسمت روش پژوهش، راهبرد پژوهش کمی و روش تجربی بود. مشارکت‌کنندگان در پژوهش، ۲۴۰ نفر ارزیاب روان‌شناس صنعتی و سازمانی آموزش‌دیده در دوره‌های کانون ارزیابی و ۱۴۵ نفر ارزیاب مدیر باسابقه حداقل پنج سال بودند. برای تجزیه و تحلیل داده‌ها از منحنی ROC و سطح زیر منحنی (AUC) استفاده شد. همچنین، نسبت شانس در طبقه‌بندی درست بیمار/غیر بیمار و شاخص‌های عوامل ریسک و عوامل بی‌خطر که با یک انتقال قیاسی ویژگی‌های مرتبط به جدول ۲*۲ روان‌شناس/مدیر، طبقه‌بندی درست/نادرست، تبدیل گردید، بکار گرفته شد. نتایج پژوهش نشان داد که روان‌شناسان صنعتی و سازمانی در مقایسه با مدیران، عملکرد مطلوب‌تری در ارزیابی شایستگی‌های رهبری در کانون‌های ارزیابی داشته و از خطای شناختی اثر هاله‌ای کمتری برخوردارند. یافته‌های این پژوهش را می‌توان در چارچوب اعتباریابی کانون‌های ارزیابی قلمداد کرد و آن را مورد بهره‌برداری قرارداد.

کلیدواژگان: اثر هاله‌ای، کانون‌های ارزیابی، روان‌شناسان صنعتی و سازمانی، مدیران، طبقه‌بندی درست

مقدمه

کاملاً واضح است که همه مدیران و کارکنان شاغل در سازمان‌ها و شرکت‌های مختلف در معرض ارزیابی عملکرد قرار می‌گیرند. اما احتمالاً تعدادی از آنها گلایه‌هایی دارند مبنی بر اینکه ارزیابی از آنها نادرست و سلیقه‌ای بوده است. دراین‌رابطه، یکی از سوءگیرهای شناختی احتمالی خطای اثر هاله‌ای است که در آن بر اساس یک ویژگی برجسته فرد، سایر ویژگی‌ها و عملکردهای او مورد قضاوت قرار می‌گیرد. به این صورت که یک ویژگی خاص فرد مانند هوش، ظاهر یا رفتار او (مثبت یا منفی) به‌عنوان پایه‌ای برای ارزیابی ویژگی‌های دیگرش در نظر گرفته می‌شود. این پدیده برای اولین‌بار توسط Thorndike (1920) به‌عنوان یکی از مهم‌ترین منابع خطای شناختی معرفی شده که می‌تواند به ارزیابی‌های نادرست و غیرمنصفانه در محیط

* **نویسنده مسئول:** نقی رعدی افسوران
رایانامه: naghi.radi.afsouran@guilan.ac.ir

Authors retain the copyright and full publishing rights.
Published by Shahid Chamran University of Ahvaz. This article is an open access article licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0)

سازمانی هم منجر شود. چنانچه، نتایج پژوهش Nasr esfahani & et al (2016) نشان داد که یکی از خطاهای بالقوه ارزیابی عملکرد منابع انسانی خطای شناختی اثر هاله‌ای است. بااین‌حال، ارزیابی عملکرد منابع انسانی در سازمان‌ها و شرکت‌ها گریزناپذیر و از الزامات اساسی اولیه است که روش‌های مختلفی برای آن وجود دارد. یکی از آن روش‌ها کانون‌های ارزیابی است.

کانون‌های ارزیابی یک فرآیند ساختاریافته علمی است که برای تعیین شایستگی‌های فردی و حرفه‌ای در انجام انواع مشاغل، به‌ویژه مشاغل مدیریتی و فرماندهی بکار می‌رود (Cascio & Aguinis, 2019; Radi Afsouran et al., 2022). این فرآیند شامل اجرای تمرین‌های شبیه‌سازی مختلفی مانند مصاحبه، بحث گروهی، ارائه شفاهی، ایفای نقش و آزمون‌های روان‌سنجی است که بر اساس رفتار شرکت‌کنندگان در شرایط شبیه‌سازی‌شده، استعدادها، توانایی‌ها و حتی ویژگی‌های روانی آن‌ها را ارزیابی می‌کند (Schippmann et al., 2006; Radi Afsouran et al., 2023). در واقع، از کانون‌های ارزیابی به‌عنوان ابزاری برای تشخیص و طبقه‌بندی درست افراد در حیطه‌های مختلف مانند انتخاب داوطلبان جدید، ارتقاء مدیران و کارکنان داخلی و بهبود و توسعه شایستگی‌های آنها استفاده می‌شود (Lievens & Conway, 2013; Radi Afsouran et al., 2019; Salimi & Moshtaghian Abarghoei, 2025). بااین‌حال، یکی از چالش‌های مهم در کانون‌های ارزیابی وجود اثرات مخدوش‌کننده‌ای است که می‌توانند دقت ارزیابی را تحت‌تأثیر قرار داده و به تصمیم‌گیری‌های نادرست بیانجامند. یکی از این اثرات، اثر هاله‌ای است که به‌صورت سوءگیری شناختی عمل کرده و ارزیابان را در کانون‌های ارزیابی به تأثیرپذیری از ویژگی‌های خاص یک فرد در ارزیابی ویژگی‌های دیگرش می‌کشاند (Landy & Conte, 2016; Thorndike, 1920). این اثر می‌تواند به‌صورت مثبت یا منفی و به سه حالت تأثیر رزومه (Landy & Conte, 2016)، تأثیر اولیه (Asch, 1997) و شیمی شخصی (Harris & Schaubroeck, 1988) عمل نموده و تصمیمات ارزیابان را به‌سوی یک طرفه‌گرایی هدایت کند. دراین‌رابطه، پژوهش Vajir & Fayyazi (2015) نشان داده است که طراحی نادرست، اجرای ضعیف و عدم حمایت مدیریت ارشد از جمله آسیب‌های اصلی کانون‌های ارزیابی هستند. همچنین، یکی از مؤلفه‌های کاهش اعتبار کانون‌های ارزیابی خطای اثر هاله‌ای گزارش شده است. ازسوی‌دیگر، پژوهش‌های مختلفی به ارائه راهکار برای کاهش خطای اثرات هاله‌ای در کانون‌های ارزیابی هم پرداخته‌اند. به طور مثال، Roch (2024) در پژوهش خود که باهدف بررسی ادراک ارزیابان از تمرین‌های شبیه‌سازی و ابزارهای مختلف کانون‌های ارزیابی با روش آزمایشی روی ۲۸۶ ارزیاب شاغل انجام داد، به این نتیجه رسید که نوع تمرین و شیوه‌ی ارائه‌ی آن به‌شدت بر ادراک عادلانه‌بودن ارزیابی، انگیزه ارزیابان و سوگیری‌های شناختی از جمله اثر هاله‌ای تأثیر می‌گذارد. همچنین، پژوهش Salami et al (2021) داده است که فقدان طراحی علمی و دقیق تمرین‌های شبیه‌سازی، امکان بروز سوگیری‌های ادراکی مانند اثر هاله‌ای را افزایش می‌دهد. چنانچه از بررسی ادبیات تجربی پژوهش قابل برداشت است، برای کاهش خطای شناختی اثر هاله‌ای لازم است بر روی ارزیابان کانون‌های ارزیابی تمرکز ویژه‌ای صورت بگیرد. چون هر چه ارزیاب متخصص‌تر و آموزش دیده‌تر باشد احتمال خطای اثر هاله‌ای می‌تواند به‌مراتب کمتر باشد. ازآنجاکه کانون‌های ارزیابی یکی از موضوعات رشته روان‌شناسی صنعتی و سازمانی است، از نظر تاریخی و تدریس دانشگاهی به رشته روان‌شناسی صنعتی و سازمانی تعلق دارد و به‌صورت تخصصی و حرفه‌ای در این رشته تحصیلی مورد مطالعه و آموزش قرار می‌گیرد. بنابراین، به نظر می‌رسد که ارزیابان آن از عملکرد بالاتری در ارزیابی مشارکت‌کنندگان برخوردار بوده و کمتر دچار خطای اثر هاله‌ای شوند. دراین‌رابطه، Smith & Anderson (2021) نشان دادند که اثر هاله‌ای می‌تواند به طور معناداری بر دقت ارزیابی‌های مدیریتی در کانون ارزیابی تأثیر بگذارد. بعلاوه، روان‌شناسان صنعتی و سازمانی در مقایسه با مدیران، توانایی بالاتری در کاهش این اثرات دارند. Woehr & Arthur (2003) در یک مطالعه‌ی فراتحلیلی بر روی ۸۳ پژوهش مرتبط با کانون‌های ارزیابی نشان دادند که نوع ارزیاب نقش مهمی در کیفیت قضاوت‌ها ایفا می‌کند. به طور مشخص، ارزیابان روان‌شناس در تشخیص و تفکیک ابعاد رفتاری عملکرد به‌طور معناداری موفق‌تر از ارزیابان مدیر عمل کردند. روان‌شناسان به دلیل آموزش تخصصی خود در مشاهده و تحلیل رفتار، در مقایسه با مدیران اعتبار همگرای بالاتری داشتند و امتیازدهی آن‌ها سازگارتر و پایاتر بود. در مقابل، مدیران بیشتر به قضاوت‌های کلی و شهودی تمایل داشتند که این

موضوع می‌تواند زمینه‌ساز اثر هاله‌ای و کاهش دقت ارزیابی شود. علاوه بر این، نتایج نشان داد که آموزش ساختاریافته‌ی ارزیابان و استفاده از دستورالعمل‌های استاندارد می‌تواند به شکل چشمگیری سوگیری‌های ادراکی را کاهش دهد و بهبود پایایی بین ارزیابان را به همراه داشته باشد. Sagie & Magnezy (1997) در پژوهشی نشان دادند که بین دانشجویان رشته روان‌شناسی و دانشجویان رشته مدیریت از نظر تشخیص ابعادی که مورد ارزیابی قرار می‌گیرد و میزان تمایز آنها تفاوت وجود دارد. دانشجویان رشته روان‌شناسی مهارت‌های نرم بیشتری دارند و می‌توانند این تفکیک را در سطح بالاتری نسبت به مدیران انجام دهند. همچنین، Coetzee & Schreuder (2016) به بررسی تأثیر آموزش چارچوب مرجع در کنار استفاده از مدل مشاهده، ثبت، طبقه‌بندی و ارزشیابی (ORCE) پرداختند. بر اساس یافته‌های این مطالعه، آموزش سیستماتیک به ارزیابان با بهره‌گیری از ORCE، به طور معناداری باعث افزایش دقت، اعتبار و انسجام ارزیابی‌ها در محیط‌های شبیه‌سازی‌شده سازمانی شد. ارزیابان پس از آموزش، توانستند رفتارهای کلیدی را دقیق‌تر مشاهده، ثبت و طبقه‌بندی کرده و در نهایت به‌صورت منسجم‌تری ارزیابی کنند. نویسندگان تأکید کرده‌اند که روش ORCE، به‌ویژه زمانی که با آموزش چارچوب مرجع همراه می‌شود، می‌تواند از بروز سوگیری‌های شناختی و قضاوت‌های شتاب‌زده در ارزیابی شایستگی‌ها جلوگیری کند. مطابق با نتایج تحقیقات، ارزیابانی که سابقه و تجربه کافی در زمینه‌های مرتبط با ارزیابی دارند، مهارت‌های تحلیلی بالاتری در مقایسه با ارزیابان کم‌تجربه و یا بدون تجربه دارند (Schippmann et al, 2006; Akhbari dorcheh & et al, 2020). بااین‌حال، روان‌شناسان صنعتی و سازمانی که به دلیل آموزش‌های تخصصی و تجربه‌هایشان در ارزیابی ویژگی‌های شخصیتی، مهارت‌های تحلیلی بالاتری دارند، می‌توانند به‌صورت دقیق‌تر در کانون ارزیابی عمل کنند (Lievens & Conway, 2013). در مقابل، مدیران که معمولاً بر اساس تجربه‌های عملی خود عمل می‌کنند، ممکن است در تمایز بین ویژگی‌های مختلف کم‌ترین موفقیت را نشان دهند. این موضوع می‌تواند ناشی از عدم داشتن آموزش‌های نظام‌مند و تخصصی در کانون‌های ارزیابی باشد (Anderson & Shackleton, 2017). همچنین، مدیران ممکن است به دلیل نداشتن پیشینه کافی تحصیل دانشگاهی در ارزیابی ویژگی‌های رفتاری، بیشتر به سوگیری‌های شناختی مثل اثر هاله‌ای دچار شوند (Borman & Motowidlo, 2015). طبق بررسی پژوهشگران این مطالعه، در ایران پژوهشی دررابطه‌با شناسایی خطای شناختی اثر هاله‌ای در ارزیابان کانون‌های ارزیابی صورت نگرفته است. درحالی‌که این موضوع برای بهبود کیفیت کانون‌های ارزیابی و ارتقای اعتبار آن دارای اهمیت دوچندان است. ازآنجاکه رشته روان‌شناسی صنعتی و سازمانی پایه‌گذار کانون‌های ارزیابی بوده است، به نظر می‌رسد که در این زمینه می‌تواند نقش‌آفرینی کرده و از اثربخشی و اعتبار کانون‌های ارزیابی به‌خوبی دفاع کند. اما آنچه در سیاست‌گذاری‌های سازمان‌ها و شرکت‌ها مورد مشاهده قرار می‌گیرد، این است که جایگاه روان‌شناسان صنعتی و سازمانی در فرایند مدیریت عملکرد منابع انسانی کم‌رنگ بوده و یا مورد غفلت قرار گرفته است. به طور مثال، در ویراست اول استاندارد کانون ارزیابی و توسعه شرکت ملی نفت ایران، رشته روان‌شناسی صنعتی و سازمانی، نه پیشرو، بلکه در ردیف دیگر رشته‌های مرتبط با کانون‌های ارزیابی آمده است (salami, 2024). این مسئله بستری برای طراحی پژوهش حاضر فراهم کرد که به مطالعه خطای اثر هاله‌ای در عملکرد ارزیابان روان‌شناسان صنعتی و سازمانی با مدیران باسابقه بپردازد. بنابراین، سؤال محوری در پژوهش حاضر این است که آیا روان‌شناسان صنعتی و سازمانی در مقایسه با مدیران باسابقه کمتر دچار خطای اثر هاله‌ای در کانون‌های ارزیابی می‌شود یا نه؟ نتایج این مطالعه می‌تواند به سازمان‌ها کمک کند تا با اطمینان بیشتری به طراحی و اجرای فرایند کانون‌های ارزیابی پرداخته، از اعتبار نتایج آن دفاع حداکثری داشته و به پشتوانه آن تصمیمات اثربخش‌تر و با ضمانت اجرایی بالاتری اتخاذ کنند.

روش‌شناسی پژوهش

پژوهش حاضر با استفاده از راهبرد کمی و روش تجربی انجام شده است. در این پژوهش از روش‌های سنجش و اندازه‌گیری و منحنی رک استفاده شده است، به همین دلیل این یک رویکرد کمی می‌باشد. همچنین، به دلیل آنکه در پژوهش دو گروه روان‌شناسان صنعتی و سازمانی و مدیران با یکدیگر مقایسه می‌گردند این پژوهش تجربی می‌باشد. بنابراین،

این پژوهش را در مجموع می‌توان جزء پژوهش‌های کمی - تجربی دانست. جامعه آماری پژوهش ارزیابان روان‌شناسی سازمانی (شامل دانشجویان و فارغ‌التحصیلان کارشناسی‌ارشد و دکتری دانشگاه اصفهان) و مدیران باسابقه (با حداقل ۵ سال سابقه مدیریتی در صنایع برق و ذوب آهن اصفهان) در سال ۱۴۰۴ بودند که قبلاً شرایط شرکت در کانون‌های ارزیابی را احراز کرده بودند. برای نمونه‌گیری، از روش نمونه‌گیری هدفمند و در دسترس استفاده شد و در مجموع ۳۸۵ نفر شامل ۲۴۰ ارزیاب روان‌شناسی صنعتی و سازمانی و ۱۴۵ ارزیاب مدیریت صنعتی انتخاب شدند. داده‌ها از طریق اجرای تمرین مصاحبه در کانون‌های ارزیابی و مشاهده فیلم آنها و تکمیل پرسش‌نامه سمت حرفه‌ای و مدیریتی گردآوری شدند. پرسش‌نامه سمت حرفه‌ای و مدیریتی (PMPQ) که توسط McCormick (1976) ساخته شده است، در ایران نخستین بار بر روی مدیران شرکت گاز در چهار استان اصفهان (Nouri et al., 2005)، یزد، چهارمحال و بختیاری و مازندران (Oreyzi, 2007) و همچنین مدیران شرکت سایپا (Oreyzi, 2006) اجرا گردید. این پرسش‌نامه در پژوهش‌های متعدد مورد استفاده قرار گرفته، پایایی و روایی آن نیز بررسی شده است. McCormick (1976) روایی همزمان این پرسش‌نامه را با نمره‌های شناختی در آزمون استعداد عمومی به دست آورده و ضرایب همبستگی بین دو آزمون را بین ۰/۶۱ تا ۰/۸۳ گزارش کرده است. Nouri et al. (2005) روایی آن را با روش‌های روایی ملاک و سازه مورد تحقیق قرار داده‌اند و نتایج مطلوبی به دست آوردند. این پرسش‌نامه با ۱۰۸ سؤال ویژگی‌های شخصیتی و شش بعد وظایف مدیریتی را می‌سنجد که عبارت است از: ۱) برنامه‌ریزی: شامل برنامه‌ریزی، فهرست‌بندی، بودجه‌بندی و سازمان‌دهی فعالیت‌های خود فرد یا دیگران، ۲) پردازش اطلاعات: شامل رونوشت برداری، تألیف یا رمزگذاری اطلاعات، پردازش کمی اطلاعات یا داده‌ها، تحلیل و ترکیب اطلاعات و ایده‌ها، ۳) قضاوت: شامل قضاوت در مورد افراد، عملیات‌ها، اهداف و منابع مالی، ۴) ارتباطات: شامل ارتباطات گفتاری، نوشتاری و استفاده از زبان، ۵) روابط بین‌فردی: شامل نظارت و هدایت، آموزش، هماهنگی، مصاحبه، مشاوره دادن، نمایندگی و مذاکره، و ۶) مهارت‌های فنی: شامل استفاده از تجهیزات و ابزارها، استفاده از روند، فنون و فرایندها. لازم به ذکر است که ابزار تحلیل شغل دارای چهار مقیاس پیچیدگی، بخشی از شغل، نقصان عملکرد و مسئولیت است که مشارکت‌کنندگان در پژوهش که قبلاً شغل موردنظر را مورد مشاهده قرار داده بودند آن را نمره‌گذاری می‌کنند. در پژوهش حاضر به دلیل آن که نمونه پژوهش، مدیران و روان‌شناسان صنعتی و سازمانی هستند، ۸۴ ویژگی با چهار مقیاس مدیریت (پیچیدگی، بخشی از شغل، مسئولیت و فقدان ویژگی) در نظر گرفته شد.

اجرای پژوهش به این صورت بود که یک کارشناس خبره به‌عنوان ارزیاب ملاک، دو ارزیاب روان‌شناسی صنعتی و سازمانی و دو ارزیاب مدیریت که اجراهای متعدد کانون‌های ارزیابی را بر عهده داشته‌اند، فیلم اجرای کانون‌های ارزیابی در مصاحبه آن را مشاهده کردند تا در مورد هر ویژگی وضعیت درست ویژگی در هر طبقه را مشخص کنند (جای‌گذاری درست/جای‌گذاری نادرست). ازآنجاکه این پژوهش یک پژوهش آزمایشی بود، متغیر مداخله‌ای فیلم توسط پژوهشگر با ۵ کارشناس خبره (شامل ۱ نفر ارزیاب ملاک، ۲ نفر ارزیاب روان‌شناسی صنعتی و سازمانی و ۲ نفر مدیر باسابقه) اجرا شد. مبنای داستان این فیلم‌ها ۲۱ شایستگی می‌باشد که در همین بخش به‌عنوان نمونه سوم ابعاد ذکر شده است. بر اساس مقیاس تحلیل شغلی PAQ (McCormick et al, 1972) شش بعد اصلی شغلی (برنامه‌ریزی، پردازش اطلاعات، قضاوت، ارتباطات، روابط بین‌فردی، مهارت‌های فنی) به ۲۱ شایستگی/بعد تفکیک شدند تا تجزیه و تحلیل رفتاری در کانون‌های ارزیابی قابلیت مناسب‌تری داشته باشد. مهم‌ترین شرط ورود افراد خبره (روان‌شناسی صنعتی و سازمانی/مدیریت) آن بود که با PMPQ هیچ آشنایی قبلی نداشته باشند. هر یک از این ۲۱ بعد، بر اساس ۴ معیار مدیریتی (پیچیدگی، بخشی از شغل، مسئولیت و فقدان ویژگی) ارزیابی می‌شوند، که این ترکیب (۴*۲۱) باعث تشکیل ۸۴ سؤال اصلی در ابزار پژوهش شده است. در مورد نمونه دوم که تمرین‌ها می‌باشد نیز مبنای ساخت این تمرین‌ها ابزار تحلیل شغل PAQ می‌باشد. این فیلم‌ها بعضی دارای شبیه‌سازی بالا در تمرین و برخی دارای شبیه‌سازی پایین در تمرین می‌باشند که بر اساس ارزیابی داوران که تا چه اندازه معرف آن بعد می‌باشند، دسته‌بندی گردید. از هر ارزیاب خواسته شد که در جلسه اتمام پروژه مشخص کنند که از کدام نشانه پی به ارجاع آن به طبقه‌بندی نشانه‌های رهبری گرفته‌اند. برای ویژگی‌های موردنظر، ۸۴ ویژگی با چهار مقیاس

مدیریت (پیچیدگی، بخشی از شغل، مسئولیت و فقدان ویژگی) در نظر گرفته شد که چون ۲۱ سؤال مربوط به این چهار بعد در پرسش‌نامه PMPQ است (McCormick et al, 1972)، تعداد ویژگی‌های موردنظر ۸۴ (۲۱ *۴) می‌باشد. برای بررسی فرضیه‌های پژوهش از روش نظریه پیگردی علامت (SDT) و سطح زیر منحنی رک (AUC) استفاده شد. هانلی و مکنیل Hanley & McNeil (1982) توضیح می‌دهند که سطح زیر منحنی رک که به اختصار AUC نامیده می‌شود، بیانگر احتمال رتبه‌بندی صحیح «یک نمونه دارای نشانه» نسبت به «یک نمونه فاقد نشانه» است. به‌عبارت دیگر، اگر دو نمونه به‌صورت تصادفی انتخاب شوند، AUC نمایانگر احتمال این است که ارزیاب، نمونه دارای نشانه را با امتیاز بزرگتری رتبه‌بندی کند. همچنین، از شاخص OR برای بررسی نسبت شانس قرار دادن نشانه درست شایستگی رهبری استفاده شد و با بررسی نمودار ROC و محاسبه AUC دقت ارزیابان مورد سنجش قرار گرفت و داده‌ها با استفاده از نرم‌افزارهای MedCalc و XLSTAT و PROXSCAL تجزیه و تحلیل شدند. لازم به ذکر است که نسبت بخت (OR) همانند بسیاری از مطالعات اپیدمیولوژیک در علوم پزشکی، در این پژوهش نیز به‌عنوان نمایه‌ای از احتمال بروز خطا در نظر گرفته شده است.

دلیل استفاده از باز نمونه‌گیری در محاسبه سطح زیر خم رک این است که بیشتر نرم‌افزارهای آماری مانند SAS، STATA و SPSS رویه‌هایی را برای تولید خم رک پیش‌بینی کردند و نیز محاسبه سطح زیر خم رک را برای داده‌های مقطعی پیشنهاد می‌کنند (Liu et al., 2005). در اینجا یک مشکل آماری پیش می‌آید، به دلیل اینکه در مدل رگرسیونی، ضرایب مشترکی برای محاسبه مقادیر پیش‌بینی شده به کار می‌روند، برآوردهای بر مبنای مدل همواره دارای احتمالات پیش‌بینی شده غیرمستقل می‌باشند. نقض شرط استقلال تبعات جدی خواهد داشت. به همین دلیل برآورد احتمالات پیش‌بینی شده از طریق مدل سنجش‌های تکرار شده، نمی‌تواند صورت گیرد چون استقلال نقض می‌شود. از طرف دیگر ضرایب درون آزمودنی، بر شدت وابستگی احتمالات پیش‌بینی شده می‌افزایند و این اعتبار محاسبه سطح زیر خم (AUC) را دچار اشکال می‌کند. نمونه‌ی بوتستراپی برای محاسبه‌ی سطح زیر خم رک را می‌توان از طریق الگوریتم ویلکاکسون ناپارامتری برآورد کرد (Liu et al., 2005).

یافته‌ها

فرضیه اول: ارزیابی روان‌شناسان سازمانی بیشتر از مدیران باسابقه در ناحیه طبقه‌بندی درست شایستگی رهبری قرار می‌گیرد.

برای آزمون این فرضیه، از شاخص‌های OR استفاده شده است. شاخص OR برای هر سه گروه ارزیاب ملاک، روان‌شناسان صنعتی و سازمانی و مدیران محاسبه شده که مقادیر آن به همراه نتایج در جدول ۱ قابل مشاهده می‌باشد.

Table 1: Odds ratio index results for the three groups

Groups	Sign	Class of signs of leadership competence	Class of signs' lack of leadership competence	Odds ratio	Confidence interval
Managers	Present in class	41	43	2.24	1.96-2.34
	Absent in class	28	56		
Psychologists	Present in class	28	57	0.56	0.49-0.61
	Absent in class	39	45		
Criterion assessor	Present in class	22	62	0.26	0.23-0.28
	Absent in class	48	36		

همان‌طور که نتایج نشان می‌دهد، نسبت شانس در گروه ارزیاب ملاک معادل ۰/۲۶، در گروه روان‌شناسان صنعتی و سازمانی ۰/۵۶ و در گروه مدیران ۲/۲۶ می‌باشد. مقدار نسبت شانس در گروه روان‌شناسان صنعتی و سازمانی به نسبت شانس گروه ارزیاب ملاک بسیار نزدیک‌تر می‌باشد. بنابراین، فرضیه اول پژوهش حاضر تأیید می‌شود. مقدار OR روان‌شناسان

صنعتی و سازمانی برابر ۰.۵۶ و مدیران برابر ۲.۲۴ بوده است. باتوجه‌به اینکه در این پژوهش، OR به‌عنوان نمایه‌ای از خطای ارزیابی تفسیر شده، OR پایین‌تر بیانگر دقت بالاتر است. بنابراین، روان‌شناسان سازمانی در طبقه‌بندی نشانه‌های شایستگی رهبری خطای کمتری نسبت به مدیران داشته‌اند. این موضوع نشان می‌دهد که ارزیابی‌های روان‌شناسان بیشتر در ناحیه‌ی طبقه‌بندی درست قرار گرفته است. در نتیجه، فرضیه اول تأیید می‌شود.

فرضِیه دوم: بین ارزیابان روان‌شناسان صنعتی و سازمانی همبستگی معناداری در ارزیابی نشانه‌های شایستگی رهبری وجود دارد.

برای بررسی این فرضیه، از شاخص پایایی بین ارزیابان (Inter-rater Reliability) استفاده شد که نتایج آن در جدول ۲ گزارش شده است.

Table 2: inter-rater reliability index of industrial and organizational psychology

Group	Dimensions	Two-raters reliability		Reliability with criterion	
		Inter raters	Inside raters	First rater	Second rater
Assessors of Psychologist	Processing	0.79	0.82	0.72	0.74
	Judgment	0.77	0.81	0.74	0.71
	Communication	0.82	0.83	0.72	0.74
	Interpersonal Communication	0.85	0.87	0.74	0.79
	Technical	0.52	0.86	0.75	0.78

Note: p<0.001 is for all Correlation coefficients.

در جدول ۲، ضرایب پایایی بین ارزیابان برای روان‌شناسان صنعتی و سازمانی، در تمامی ابعاد، مقادیر بالا و قابل‌قبولی (بالاتر از ۰.۷۵ در اکثر موارد) را نشان داد. این همبستگی بالا بیانگر آن است که بین ارزیابان روان‌شناس صنعتی و سازمانی در تشخیص نشانه‌های رهبری، دیدگاه نسبتاً مشابه و همگنی وجود داشته است. بنابراین، فرضیه دوم تأیید می‌شود.

فرضیه سوم: بین ارزیابان مدیران باسابقه همبستگی معناداری در ارزیابی نشانه‌های شایستگی رهبری وجود دارد. برای بررسی این فرضیه، از شاخص پایایی بین ارزیابان استفاده شد که نتایج آن در جدول ۳ گزارش شده است.

Table 3: inter-rater reliability index of managers

Group	Dimensions	Two-raters reliability		Reliability with criterion	
		Inter raters	Inside raters	First rater	Second rater
Assessors of managers	Processing	0.53	0.72	0.58	0.54
	Judgment	0.48	0.74	0.54	0.52
	Communication	0.52	0.76	0.51	0.38
	Interpersonal Communication	0.61	0.68	0.43	0.41
	Technical	0.53	0.71	0.62	0.53

Note: p<0.001 is for all Correlation coefficients.

چنانچه در جدول ۳ قابل مشاهده است، در بررسی پایایی بین ارزیاب برای مدیران باسابقه، مقدار ضرایب در مقایسه با روان‌شناسان صنعتی و سازمانی پایین‌تر گزارش شد. بااین‌حال، در برخی ابعاد مانند ارتباطات، ضریب پایایی بین ارزیابان در حد متوسط به بالا (مثلاً ۰.۷۴) بوده است. اگرچه این ضرایب نسبت به روان‌شناسان صنعتی و سازمانی کمتر هستند، اما همچنان می‌توان آن‌ها را در دامنه‌ی همبستگی معنادار متوسط به بالا در نظر گرفت. در نتیجه، باتوجه‌به داده‌ها، فرضیه سوم نیز تأیید می‌شود، اما با همبستگی نسبتاً ضعیف‌تر در مقایسه با روان‌شناسان صنعتی و سازمانی (فرضیه دوم).

فرضیه چهارم: بین نسبت شانس تشخیص صحیح نشانه‌های شایستگی رهبری توسط ارزیابان روان‌شناسی صنعتی و سازمانی و ارزیابان مدیر تفاوت معناداری وجود دارد.

مطابق با داده‌هایی که در جدول ۱ مشاهده می‌شود، نسبت شانس در گروه ارزیاب ملاک معادل ۰/۲۶، در گروه روان‌شناسان صنعتی و سازمانی ۰/۵۶ و در گروه مدیران ۲/۲۶ می‌باشد. مقدار نسبت شانس در گروه روان‌شناسان صنعتی و سازمانی به نسبت شانس گروه ارزیاب ملاک بسیار نزدیک‌تر می‌باشد. بنابراین، فرضیه چهارم پژوهش حاضر تأیید می‌شود.

فرضیه پنجم: سطح زیر منحنی عملکرد (AUC) ارزیابان روان‌شناس صنعتی و سازمانی با ارزیابان مدیر تفاوت معناداری دارد. برای آزمون این فرضیه از شاخص سطح زیر منحنی عملکرد (AUC) و تحلیل نمودار (ROC) استفاده شد. در جدول ۴، فاصله اطمینان با ۰/۹۵ اطمینان (α=۰/۰۵) با باز نمونه‌گیری از صد نمونه بوتستراپ و در جدول ۵ با باز نمونه‌گیری از سیصد نمونه بوتستراپ گزارش شده است. همچنین، نمودارهای ۱ و ۲ به ترتیب متناظر با داده‌های جدول‌های ۴ و ۵ ترسیم شده است.

Table 4: analyzing the mean of AUC with 100 re-sampling (bootstrapping)

Groups	Mean of AUC	Confidence interval without bootstrap		Confidence interval with bootstrap		Result
Managers	0.74	0.867	0.612	0.854	0.629	Good
Criterion assessors	0.82	0.947	0.692	0.934	0.717	Excellent
Psychologists	0.81	0.937	0.682	0.926	0.709	Excellent

Table 5: analyzing the mean of AUC with 300 re-sampling (bootstrapping)

Groups	Mean of AUC	Confidence interval without bootstrap		Confidence interval with bootstrap		Result
Managers	0.72	0.825	0.614	0.809	0.630	Good
Criterion assessors	0.80	0.905	0.694	0.879	0.701	Excellent
Psychologists	0.78	0.885	0.674	0.864	0.690	Excellent

همان‌طور که یافته‌های مربوط به سطح زیر منحنی عملکرد (AUC) در هر دو جدول بالا نشان می‌دهد، دقت روان‌شناسان صنعتی و سازمانی نزدیک به ارزیاب ملاک و در سطح عالی است. اما میانگین AUC مدیران با فاصله‌ای بیشتر از ارزیاب ملاک و در سطح خوب قرار دارد. به این معنا که عملکرد ارزیابان روان‌شناسان صنعتی و سازمانی بهتر از عملکرد ارزیابان مدیران باسابقه است و این تفاوت در سطح اطمینان ۰/۹۵ معنادار است. بنابراین، فرضیه پنجم تأیید می‌شود.

فرضیه ششم: اثر هاله‌ای در ارزیابی نشانه‌های شایستگی رهبری هم در بین روان‌شناسان صنعتی و سازمانی و هم در بین مدیران باسابقه مشاهده می‌شود. برای آزمون این فرضیه از شاخص‌های پایایی بین ارزیابان، درون ارزیاب و پایایی با ملاک استفاده شد. اثر هاله‌ای زمانی قابل‌انتظار است که همبستگی بین ابعاد مختلف ارزیابی بالا بوده و نشان‌دهنده تعمیم نمره‌های ارزیاب به‌تمامی ابعاد، صرف‌نظر از تفاوت‌های واقعی باشد. مطابق با جدول‌های ۲ و ۳، در بین ارزیابان روان‌شناس صنعتی و سازمانی، فاصله پایایی بین ارزیابان و درون ارزیابان حد کفایت اندک است (حداکثر ۰.۰۴). درحالی‌که، پایایی فاصله بین ارزیابان و درون ارزیابان مدیران در حد قابل توجهی بیشتر است (حداقل ۰.۰۷ تا حداکثر ۰.۲۶). این به وضوح نشان‌دهنده اثر هاله‌ای می‌باشد. پایایی با ملاک هم در بین گروه ارزیابان روان‌شناسی صنعتی و سازمانی بسیار بالاتر است. همه همبستگی در حد $P<0.001$ معنادار می‌باشد. چنانچه، در گروه ارزیابان روان‌شناس صنعتی و سازمانی، ضرایب پایایی درون ارزیاب در همه ابعاد بسیار بالا (بالاتر از ۰.۸۰) و ضرایب پایایی بین ارزیاب نیز در اغلب موارد بالاتر از ۰.۷۵ بوده است. به طور خاص، پایایی درون ارزیاب در بُعد «ارتباطات بین‌فردی» برابر با ۰.۸۷ و در بُعد «فنی» برابر با ۰.۸۶ گزارش شده که نشان از ثبات بسیار بالا در ارزیابی همه ابعاد توسط یک ارزیاب دارد.

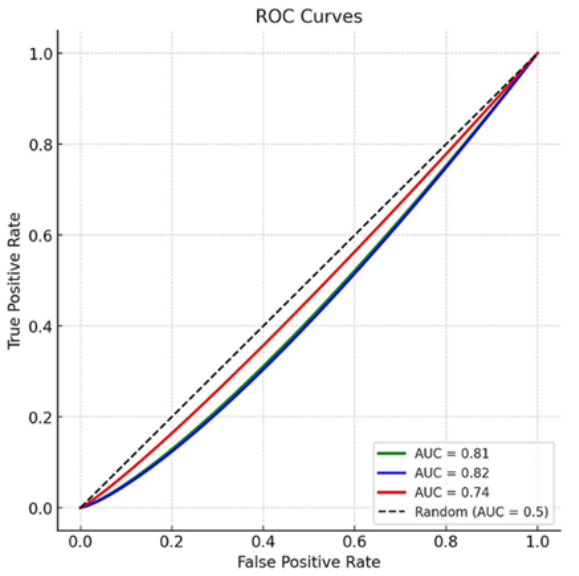

Chart 1. AUC of psychologist and manager based on the data of Table 4

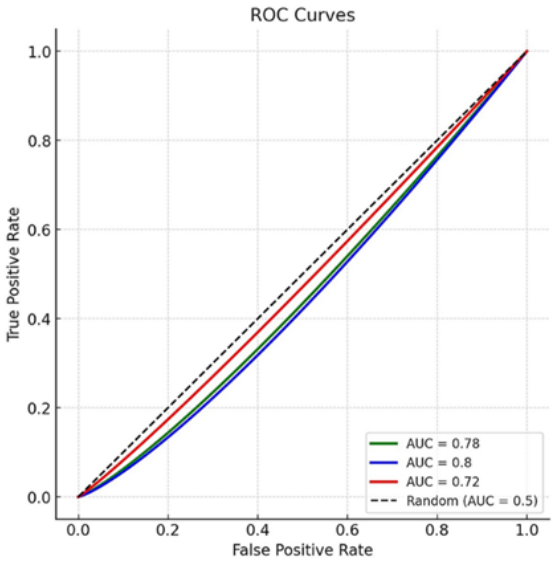

Chart 2. AUC of psychologist and manager based on the data of Table 5

همچنین، در گروه مدیران، ضرایب پایایی درون ارزیاب نیز به‌صورت معناداری بالا گزارش شده‌اند (بیشتر از ۰.۷۰ در اکثر موارد)؛ اما میزان پایایی بین ارزیابان مدیر، در حد متوسط و در برخی ابعاد کمتر از روان‌شناسان صنعتی و سازمانی بوده است. بالا بودن همبستگی بین ارزیابی‌ها درونی هر ارزیاب (در همه ابعاد) بدون توجه به تفاوت‌های واقعی هر بُعد، می‌تواند نشانه‌ی از اثر هاله‌ای باشد؛ به این معنا که ارزیاب بر اساس برداشت کلی مثبت یا منفی از یک شرکت‌کننده، به سایر شایستگی‌ها نیز نمرات مشابه داده است. این مسأله زمانی تقویت می‌شود که همبستگی‌ها در همه ابعاد بالا باشد، ولی داده‌های رفتاری تفاوت واقعی نشان دهند. بر اساس این نتایج، فرضیه ششم پژوهش مبنی بر وجود اثر هاله‌ای در بین هر دو گروه ارزیاب (روان‌شناسان صنعتی و سازمانی و مدیران باسابقه) تأیید می‌شود. بااین‌حال، اثر هاله‌ای در گروه مدیران به میزان بیشتری وجود داشته است. چرا که پایایی با ملاک و درون ارزیاب به‌صورت معناداری بالاتر گزارش شده است.

بحث و نتیجه‌گیری

پژوهش حاضر باهدف مقایسه خطای شناختی اثر هاله‌ای ارزیابان روان‌شناسان صنعتی و سازمانی و مدیران باسابقه در طبقه‌بندی نشانه‌های شایستگی رهبری در کانون‌های ارزیابی انجام شد. نتایج پژوهش نشان داد که: ارزیابی روان‌شناسان صنعتی و سازمانی بیشتر از مدیران در ناحیه طبقه‌بندی درست شایستگی رهبری قرار می‌گیرد. این نتیجه همسو با یافته‌های Salami et al (2021) ، Woehr & Arthur (2003) و Sagie & Magnezy (1997) است که نشان داده‌اند روان‌شناسان به دلیل آموزش تخصصی و مهارت در تحلیل رفتار، دقت بالاتری نسبت به مدیران دارند. بنابراین، روان‌شناسان صنعتی و سازمانی در طبقه‌بندی نشانه‌های رهبری خطای کمتری داشته و در ناحیه طبقه‌بندی درست قرار گرفته‌اند. همچنین، بین ارزیابان روان‌شناسی صنعتی و سازمانی همبستگی معناداری در ارزیابی نشانه‌های شایستگی رهبری وجود دارد. این نتیجه حاکی از آن است که روان‌شناسان صنعتی و سازمانی توانسته‌اند نشانه‌های شایستگی رهبری را باثبات و به طور هماهنگ ارزیابی کنند. این نتیجه با یافته‌های پژوهش Roch (2024) و Woehr & Arthur (2003) همسو است که نشان می‌دهند آموزش ساختاریافته و تخصص روان‌شناسان صنعتی و سازمانی باعث افزایش دقت ارزیابی می‌شود. بعلاوه، بین ارزیابان مدیریت همبستگی معناداری در ارزیابی نشانه‌های شایستگی رهبری وجود دارد. این نتیجه نشان می‌دهد که مدیران توانسته‌اند تا حدی

ارزیابی هماهنگ ارائه دهند، اما هنوز خطای ارزیابی و سوگیری‌های ذهنی بیشتری نسبت به روان‌شناسان دارند. همچنین، یافته‌های پژوهش حاضر نشان داد که بین نسبت شانس تشخیص صحیح نشانه‌های رهبری توسط ارزیابان روان‌شناس صنعتی و سازمانی و مدیران تفاوت معناداری وجود دارد. بعلاوه، سطح زیر منحنی عملکرد ارزیابان روان‌شناس صنعتی و سازمانی با ارزیابان مدیریت تفاوت معناداری دارد. این نتیجه نشان می‌دهد که روان‌شناسان قادرند به طور بهینه‌تر نسبت به مدیران، عملکرد واقعی افراد را پیش‌بینی کنند. این نتایج با مطالعات Salami et al (2021) و Sagie & Magnezy (1997) همسو است و بر اهمیت شاخص AUC به‌عنوان معیار دقت طبقه‌بندی تأکید می‌کند و در نهایت، اثر هاله‌ای در ارزیابی نشانه‌های شایستگی رهبری هم در بین روان‌شناسان صنعتی و سازمانی و هم در بین مدیران باسابقه مشاهده می‌شود. بااین‌وجود، روان‌شناسان صنعتی و سازمانی توانایی بیشتری در تمایز بین ابعاد رفتاری و عملکردی داشته و نهایتاً دقت بیشتری در جای‌گذاری نشانه‌های رفتاری شایستگی رهبری ارائه می‌دهند. این نتیجه با نتایج پژوهش Nasr esfahani et al (2016) ، Smith & Anderson (2021) ،Woehr & Arthur (2003)، Sagie & Magnezy (1997)، Vajir & Fayyazi (2015) همخوانی دارد.

به‌طورکلی، یافته‌های این پژوهش در دو بعد قابل تبیین و تفسیر است: اول، پایایی بالاتر درون‌فردی و بین‌فردی در گروه روان‌شناسان صنعتی و سازمانی و دوم، دقت بیشتر در طبقه‌بندی صحیح ویژگی‌های رهبری در تحلیل ROC. این یافته‌ها با دیدگاه Thornton (1992) و Thornton et al (2015) هم‌راستا است که تأکید دارند روان‌سنجی دقیق، مبنای ارزیابی‌های معتبر در کانون‌های ارزیابی است. همچنین، استفاده از منحنی ROC که در علوم پزشکی توسعه‌یافته، نوآوری خاص این پژوهش به حساب می‌آید. به‌کارگیری این ابزار برای سنجش اثر هاله‌ای در حیطه روان‌شناسی صنعتی و سازمانی و مدیریت نشان‌دهنده قابلیت انتقال میان‌رشته‌ای ابزارهای تحلیلی است. Thornton (1992) یکی از بزرگترین نظریه پردازان و پژوهشگران کانون‌های ارزیابی، فرایند ارزیابی را به ترتیب مشاهده، ثبت، طبقه‌بندی و ارزیابی رفتاری می‌داند. او در کتاب خود دو الگوی مدل ذهنی و مدل رفتاری را در مقابل هم قرار داد. در مدل ذهنی این ادعا مطرح می‌شد که ارزیابی‌ها در هر یک از این چهار مرحله شامل خطاها و سوگیری‌هایی می باشد که به ترتیب در مورد هر طبقه با مفاهیم مربوط به آن طبقه آن‌ها را شرح داد. اما در مدل رفتاری این دیدگاه وجود داشت که می‌توان بدون سوگیری و خطا، ارزیابی را با دقت کامل انجام داد. ثورنتون بین این دو الگو یک نوع ارزیابی تعریف کرد که آن را ارزیابی در کانون‌های ارزیابی در عمل دانست و دو الگوی رفتاری و ذهنی را به ترتیب الگوی خوشبینانه و بدبینانه دانست. در الگوی رفتاری هیچ خطایی رخ نمی‌دهد و ارزیابان می‌توانند بدون سوگیری و خطا شایستگی‌های افراد در کانون‌های ارزیابی را به‌دقت نشان دهند. در مدل ذهنی، برعکس، انواع خطاها و سوگیری‌ها رخ می‌دهد. در مورد طبقه‌بندی، دیدگاه او آن بود که انسان به لحاظ شناختی قادر به طبقه‌بندی رفتار نمی‌باشد، زیرا رفتار بد تعریف شده و طبیعی می‌باشد. منظور از طبیعی بودن طبقه آن است که با ویژگی‌های بسیار متمایز قابل تعریف نیست به همین دلیل خطاها در آن رخ می‌دهد. یکی از خطاهایی که در ادبیات پژوهشی بسیار مشهور شده است، خطای اثر هاله‌ای است. اثر هاله‌ای وقتی رخ می‌دهد که تعدادی از افراد که به ارزیابی می‌پردازند، نتایج‌شان با یکدیگر همبستگی بالایی داشته باشد اما در نوبت‌های متوالی که آن‌ها به ارزیابی همان رفتارها می‌پردازند، ضریب همبستگی بین آن ارزیاب‌ها پایین باشد. یا برعکس هنگامی که روابط بین ویژگی‌ها بین ارزیابان مختلف در سطح پایین ولی رابطه درون ارزیابی‌کنندگان (چند بار ارزیابی توسط یک فرد) در سطح بالایی باشد. این نشان می‌دهد که آن ویژگی در نظر افراد کم‌اهمیت یا بسیار دارای اهمیت است و این اهمیت را ارزیابان دیگر ندارند. به همین دلیل در پژوهش حاضر برای سنجش پایایی بین ارزیابان و پایایی درون ارزیابان تعدادی فیلم توسط ارزیابان مشاهده شد که در آن باید این نوع پایایی‌ها بر مبنای درجه‌بندی که به شش نوع رفتار مدیر/ رهبر انجام می‌دهند در سطح بالایی باشد. پایایی بین ارزیابی‌کنندگان و پایایی درون ارزیابی‌کنندگان برای اثر هاله‌ای هم مورد مطالعه قرار گرفت. همچنین پایایی ملاک که عبارت از فردی است که به لحاظ دانش و مهارت در سطح نسبتاً بالاتری از طبقه افراد روان‌شناسی صنعتی و سازمانی یا مدیران دارند نیز مورد بررسی قرار گرفت. این مفهوم اول اثر هاله‌ای در این پژوهش می‌باشد. برای مفهوم دوم از یک فن شناخته شده در روان‌سنجی که به ترسیم یک خم (منحنی) به اسم منحنی ROC می‌انجامد، استفاده شد. این ابزار دارای یک سطح زیر منحنی است که شاخصی برای محاسبه اثر هاله‌ای می‌باشد. در اینجا

لازم به تذکر است که این فن در علوم پزشکی و علوم زیستی و حیاتی ساخته شده است، به‌خصوص برای فیلم‌هایی که در رادیولوژی مورد استفاده قرار می‌گیرد، کاربرد دارد. در این روش طبقه‌بندی در یک جدول ۲*۲، متغیرهای بیمار / غیر بیمار و عوامل خطرپذیر برای ایجاد آن بیماری و عوامل غیر خطرپذیر مشاهده می‌گردد. ازآنجاکه این فن مشاهده‌ای برای ارزیابی اهمیت بسیار داشت از این الگو به‌صورت انتقال قیاسی استفاده گردید و مفاهیم موردنظر در الگوی بیماری به الگوی شناخت ویژگی‌های رهبر/مدیر تبدیل گردید. به این صورت که آیا ارزیابان طبقه‌بندی را درست یا نادرست انجام می‌دهند و دو گروهی که به ارزیابی می‌پردازند یعنی روان‌شناسان صنعتی و سازمانی و مدیران باسابقه برای این کار مورد مقایسه قرار گرفتند. در این پژوهش که تنها پژوهش شناخته شده از این نظر در ایران می‌باشد، نتایج به طور آشکار از برتری روان‌شناسان صنعتی و سازمانی نسبت به مدیران حمایت کرد. به این صورت که روان‌شناسان صنعتی و سازمانی دارای خطای کمتری در ارزیابی می‌باشند. بنابراین شایستگی‌های بیشتری برای قضاوت در مورد ارزیابی شایستگی‌های رهبری دارند به این ترتیب در این پژوهش برای اولین‌بار از مفاهیم روان‌سنجی برای اعتباریابی مصاحبه در کانون‌های ارزیابی در دو محیط مختلف فیلم و نوشتار استفاده گردیده است. این پژوهش دارای محدودیت‌هایی است که مهم‌ترین آن‌ها نمونه انتخاب شده می‌باشد. به این ترتیب که نمونه انتخاب شده فقط از بین روان‌شناسان صنعتی و سازمانی و مدیران باسابقه انتخاب شده است و برای تعمیم آن به گروه‌های دیگر ارزیابی به‌خصوص روانسنج‌ها (به‌خصوص به این دلیل که آن‌ها با مفاهیم روان‌سنجی و سوگیری آشنایی بیشتری دارند) انجام نشده است و در آینده احتمالاً پژوهش‌هایی که از گروه‌های متنوع‌تری از ارزیابان استفاده کنند قابل توصیه می‌باشد. همچنین، در این پژوهش فقط از دو صنعت برق و ذوب آهن استفاده گردیده است. در صورتی که صنایع بزرگی در ایران وجود دارند که در تعمیم نتایج نمونه حاضر به آن نمونه‌های کلی، نیاز به پژوهش و تطبیق بیشتری می‌باشد. بر مبنای یافته‌های این پژوهش پیشنهاد می‌شود که پیشنهاد می‌شود در طراحی و اجرای کانون‌های ارزیابی، از روان‌شناسان صنعتی و سازمانی به‌عنوان ارزیاب اصلی یا ناظر ارزیابی‌ها استفاده شود. همچنین، در تدوین استانداردهای شایستگی ارزیابان، ویژگی‌هایی مانند آشنایی با سوگیری‌های شناختی، توانایی در تحلیل رفتار و تجربه در آزمون‌های روان‌سنجی لحاظ شود و برای کاهش اثر هاله‌ای، آموزش رسمی در حوزه شناخت خطاهای ارزیابی برای مدیران و ارزیابان غیرروانشناس در نظر گرفته شود. پژوهشگران آینده می‌توانند نتایج این پژوهش را که در سطح روان‌سنجی و اعتباریابی مصاحبه در کانون‌های ارزیابی بوده است، به لحاظ آماری و با شیوه‌های آماری مورد آزمون قرار دهند. بعلاوه، این پژوهش فقط با تمرین مصاحبه انجام شده است و تمرین‌های دیگر از قبیل تمرین ارائه شفاهی، بحث‌های گروهی بدون رهبر، بحث‌های گروهی با رهبر و بازی‌های تجاری تمرین‌های مناسبی می‌باشند که می‌توان نتایج این پژوهش را با آن‌ها نیز بررسی کرد.

تقدیر و تشکر

از همه مشارکت‌کنندگان که بدون همکاری آنها انجام این پژوهش ممکن نبود، تقدیر و تشکر می‌کنیم.

ملاحظات اخلاقی

اصول اخلاقی تماماً در این مقاله رعایت شده است.

حمایت مالی

مقاله حاضر مورد حمایت مالی هیچ نهاد و یا سازمانی نبوده است.

تضاد منافع

هیچ‌گونه تعارض منافع توسط نویسندگان گزارش نشده است.

References

Anderson, N., & Shackleton, V. J. (2017). Assessment centers in personnel selection and development. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology* (Vol. 2, pp. 283-307). Sage Publications.

Akhbari Dorcheh, M. S., Oreyzi Samani, H. R., Eskandari, H., & Bidram, H. (2020). The Phenomenology of the Experience of Experts and Novice Managers in the Oil Industry to Identify Their Cognitive Differences. *Strategic Studies in Petroleum and energy Industry*, *11*(44), 183-210. [in Persian]

Asch, S. E. (1997). Forming impressions of personality. In C. E. Hughes, D. Ariely, & D. A. Eckerman (Eds.), *Background readings for the joy of experimental psychology* (3rd ed., pp. 48–72). Kendall/Hunt Publishing Company. (Reprinted from "Journal of Abnormal and Social Psychology," 41, 1946, pp. 258–290)

Borman, W. C., & Motowidlo, S. J. (2015). Task performance and contextual performance: The meaning and measurement of job performance. *Journal of Applied Psychology*, 100 (3), 990-1000. https://doi.org/10.1037/a0038131

Cascio, W. F., & Aguinis, H. (2019). *Applied psychology in human resource management* (9th ed.). Pearson Education.

Coetzee, M., & Schreuder, D. (2016). The use of frame-of-reference training to improve observer accuracy in behavioural assessments. *SA Journal of Human Resource Management*, 14(1), 1–9.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29–36. https:// DOI: 10.1148/radiology.143.1.7063747

Harris, M., & Schaubroeck, J. (1988). The effects of halo error on performance appraisal validity. *Journal of Applied Psychology*, 73 (1), 132-139. https://doi.org/10.1037/0021-9010.73.1.132

Landy, F. J., & Conte, J. M. (2016). *Work in the 21st Century: An Introduction to Industrial and Organizational Psychology* (5th ed.). Wiley

Lievens, F., & Conway, J. M. (2013). Assessment centers as a selection tool: Current status and future directions. *Annual Review of Psychology*, 64 , 491-517. https://doi.org/10.1146/annurev-psych-113011-143812

Lievens, F., & Sackett, P. R. (2020). The validity of assessment centers for the prediction of practical intelligence. *Journal of Applied Psychology, 105* (4), 412–431.

Liu, H., Li, G., Cumberland, W. G., & Wu, T. (2005). Testing statistical significance of the area under a receiving operating characteristics curve for repeated measures design with bootstrapping. *Journal of Data Science*, 3(3), 257–278. https://doi.org/10.6339/JDS.2005.03(3).206

McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology*, 56(4), 347–368. https://doi.org/10.1037/h0033099

McCormick, E. J. (1976). Ergonomics – Future Perspectives. *SA Journal of Industrial Psychology*. Article 271. https://doi.org/10.4102/sajip.v0i0.271

Nouri, A. Oreyzi, H.R., Razi, H. Molavi, H. Kalantari, M. Neshat-Doost, H.T. Shafti, A. & Moussavian, H. (2005). Psychological skills required in managerial jobs: management of Isfahan Province Gas Company. *Research report*. University of Isfahan and National Iranian Gas Company. [In Persian]

Nasr Esfahani, A, Gerami, F, & Momen, Z. (2016). Investigating the effect of performance appraisal errors on organizational commitment. Case study: Isfahan Electricity Generation Management Employees. First National Conference on Economics, Management, and Accounting, Ahvaz. https://civilica.com/doc/620561 [in Persian]

Oreyzi, H.R. (2007). Management skills in gas companies of Yazd, Chahar Mahal Bakhtiari, and Mazandaran provinces. *Research project*. Vice-Chancellor of Research, University of Isfahan [In Persian]

Oreyzi, H.R. (2006). Analysis of jobs in Saipa Automobile Manufacturing Company. *Research report*. Iran Industrial Renovation Company and Saipa Automobile Manufacturing Company [In Persian]

Roch, S. G. (2024). Perceptions of assessment center exercises: Between-exercise differences and interventions. *Industrial and Organizational Psychology*, 17(1), 124-140. https://doi.org/10.1017/iop.2024.8

Radi Afsouran, N., C Thornton III, G., & Charkhabi, M. (2022). Leadership Development Assessment Center: A Review on Advantages and Disadvantages for Developing Leadership Behavioral Competencies. *International Journal of Organizational Leadership*, *11*(1), 71-88. https://doi.org/10.33844/ijol.2022.60620

Radi Afsouran, N. Siadat, A., Hoveida, R., Orizi, H. R., & Thornton III, G. (2019). Training and Developing Leadership Competencies Using Leader Development Assessment Center: Empirical Study of the Competencies of Transformation, Team Building, and Strategic Thinking. *Journal of Educational Sciences*. 26 (1), 97-118. https://doi.org/10.22055/edus.2019.25363.2493 [In Persian]

Radi Afsouran, N. Oreyzi Samani, H.R. Zareai, M. (2023). Comparing the effect of parallel simulation exercises, repeated simulation exercises, and simulation exercises with no description in developing the employees' relational energy competency. Industrial and organizational psychology studies. Vol, 10. No, 1. Pp, 1-14. https://doi.org/10.22055/jiops.2023.42912.1331 [In Persian]

Sagie, A., & Magnezy, R. (1997). Assessor type, number of distinguishable dimension categories, and assessment centre construct validity. *Journal of Occupational and Organizational Psychology*, 70(1), 103–108. https://doi.org/10.1111/j.2044-8325.1997.tb00634.x

Schippmann, J. S., Prien, E. P., & Ohlsson, S. (2006). *The practice of competency modeling*. *Personnel Psychology*, 59 (2), 233-266. https://doi.org/10.1111/j.1744-6570.2006.00048.x

Salimi M, Shokrkon H, Montakhab Yeganeh M. A. (2021). Critical Approach to the Construct -Related Validity of Assessment Centers: A Study of the Assessment Center of the National Iranian Oil Company. *International journal of psychology* .14 (2) :218-247. http://ijpb.ir/article-1-305-fa.html

Salimi, M. (2024). *Assessment and Development Centers Standard*. Tehran. National Iranian Oil Company. [In Persian]

Salimi, M., & Moshtaghian Abarghoei, H. (2025). Estimating the Reliability & Validity of Assessment Center Ratings Using Generalizability Theory. *Industrial and Organizational Psychology Studies*, *12*(2), 1-20. https://doi.org/10.22055/jiops.2025.49140.1476 [In Persian]

Smith, J., & Anderson, L. (2021). The impact of halo effect on managerial assessments in assessment centers. *Journal of Business Psychology*, 36(4), 587-602. https://doi.org/10.1007/s10869-020-09712-8

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4 (1), 25-29. https://doi.org/10.1037/h0068635

Thornton, A., Axinn, W. G., & Hill, D. H. (1992). Reciprocal effects of religiosity, cohabitation, and marriage. *American journal of sociology*, 98(3), 628-651.

Thornton III, G. C., Rupp, D. E., & Hoffman, B. J. (2015). *Assessment center perspectives for talent management strategies*. Routledge.

Vajir, L., & Fayyazi, B, B. M. (2015). The Pathology of the Assessment and Developmental Centers of Iran: An improvement Suggestion. *Iranian journal of management sciences*. 10 (37). 103-126 [in Persian]

Woehr, D. J., & Arthur, W. Jr. (2003). An examination of the construct validity of assessment center ratings: A meta-analysis. *Journal of Applied Psychology*, 88(3), 475–486. https://doi.org/10.1037/0021-9010.88.3.475