



## Moving Beyond Mechanical Objectivity: Addressing the Challenges and Exploring Novel Solutions in Iranian University Admissions

Mahdi. Karvandi Renani<sup>1</sup>, Keyvan. Salehi<sup>2\*</sup>, Ebrahim. Khodaie<sup>3</sup>

<sup>1</sup> PhD student in Measurement and Assessment, University of Tehran, Tehran, Iran

<sup>2</sup> Associate Professor, Faculty of Psychology and Education, University of Tehran, Iran

<sup>3</sup> Professor, Faculty of Psychology and Education, University of Tehran, Tehran, Iran

\* Corresponding author email address: keyvansalehi@ut.ac.ir

### Article Info

#### Article type:

Review Article

#### How to cite this article:

Karvandi Renani, M., Salehi, K., & Khodaie, E. (2026). Moving Beyond Mechanical Objectivity: Addressing the Challenges and Exploring Novel Solutions in Iranian University Admissions. *Quarterly Journal of Research and Planning in Higher Education*, 32(2), 165-187.



© 2026 the authors. Published by Institute for Research and Planning in Higher Education (IRPHE), Tehran, Iran. This is an open access article under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

### ABSTRACT

Extensive criticisms have been raised against Iran's university admission system, prompting the addition of final examinations alongside the national entrance exam. Despite these changes, standardized testing remains central, driven by a commitment to objectivity and ethical impartiality. In this paper, we argue that the traditional conception of objectivity—specifically mechanical objectivity or the value-free ideal—assumes that test scores alone are sufficient for admission, creating serious challenges for the validity of entrance exams. Classical test theory, rooted in operationalism and aligned with mechanical objectivity, emphasizes reliability over validity, failing to distinguish between appropriate and inappropriate assessments. Consequently, mechanical objectivity alone cannot ensure justice or epistemic integrity. Drawing on Daston and Galison's work, we highlight the need for additional epistemic virtues such as truth-to-nature and trained judgment. Latent variable models, which aim to uncover the underlying essence of phenomena, are presented as alternatives to classical test theory. Furthermore, trained judgment leverages expert evaluation to offer a more holistic view of student abilities. Recent perspectives also frame objectivity as context-sensitive and goal-oriented. Therefore, redefining objectivity in university admissions requires identifying clear social and academic goals and designing evaluations accordingly. Ultimately, we propose a hybrid framework that integrates latent variable-based assessments with expert-driven processes, aiming to create a fairer, multidimensional, and more just admission model for Iran's universities.

**Keywords:** Mechanical Objectivity, Value-Free Ideal, National Entrance Examination, Epistemic Virtues, Validity



## فراسوی عینیت مکانیکی: چالش‌های عینیت و راهکارهای نوین در پذیرش دانشجو

مهدی کروندی رنانی<sup>۱</sup>، کیوان صالحی<sup>۲</sup>، ابراهیم خدائی<sup>۳</sup>

۱. دانشجوی دکتری، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران

۲. دانشیار، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران

۳. استاد، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران

\*ایمیل نویسنده مسئول: keyvansalehi@ut.ac.ir

### اطلاعات مقاله

### چکیده

### نوع مقاله

مروری

### نحوه استناد به این مقاله:

کروندی رنانی، مهدی، صالحی، کیوان، و خدائی، ابراهیم. (۱۴۰۵). فراسوی عینیت مکانیکی: چالش‌های عینیت و راهکارهای نوین در پذیرش دانشجو. *فصلنامه پژوهش و برنامه‌ریزی در آموزش عالی*، ۳۲(۲)، ۱۸۷-۱۶۵.



© ۱۴۰۵ تمامی حقوق انتشار این مقاله متعلق به نویسنده است. انتشار این مقاله به صورت دسترسی آزاد مطابق با گواهی (CC BY 4.0) صورت گرفته است.

در دهه‌های اخیر، نظام پذیرش دانشجو در کشور ایران با نقدهای گسترده‌ای مواجه شده است. اتکای صرف به آزمون‌های استاندارد شامل آزمون سراسری و آزمون‌های نهایی، که با منطق «عینیت مکانیکی» توجیه می‌شوند، موجب شده است تا عدالت آموزشی بیش از حد با استانداردهای سازشی و یکسان‌سازی گره بخورد. در این مطالعه نشان داده شد که عینیت مکانیکی، هرچند با نیت حفظ بی‌طرفی و پایایی معرفت‌شناختی به کار گرفته می‌شود، به دلیل تبدیل نمره آزمون به شرط کافی پذیرش، به یک چالش جدی در رواسازی آزمون‌ها تبدیل شده است. به‌ویژه نظریه کلاسیک آزمون که با عملیات‌گرایی هم‌ریشه است، قادر به تمایز میان آزمون‌های مناسب و نامناسب نیست، چراکه پایایی به‌عنوان اصلی‌ترین مفهوم در این نظریه مستقل از روایی آزمون است. در پاسخ به این محدودیت‌ها، در این مطالعه سه برهان اصلی مطرح شد. نخست، با تکیه بر دیدگاه «حقیقت‌طبیعی» نشان داده شد که مدل‌های متغیر مکنون، همچون نظریه سؤال‌پاسخ، با اتکا بر کشف ذات پدیده‌ها می‌توانند چارچوبی کارآمدتر برای آزمون‌سازی فراهم کنند و از محدودیت‌های نظریه کلاسیک آزمون فراتر روند. دوم، با ارجاع به «فضاوت‌تمرین‌شده» استدلال شد که تخصص انسانی و مهارت‌های تفسیری ارزیابان باید بخشی جدایی‌ناپذیر از فرایند پذیرش باشند، زیرا عوامل کیفی همچون انگیزه، زمینه اجتماعی و توانایی‌های بالقوه دانشجو را نمی‌توان صرفاً از طریق آزمون‌ها و رویه‌های استاندارد بازنمایی کرد. سوم، با پذیرش دیدگاه‌های نوین فلسفه علم، عینیت به‌عنوان یک مفهوم وابسته به اهداف و موقعیت بازتعریف شد؛ یعنی عدالت در پذیرش دانشجو فقط با کمی‌سازی معیارها محقق نمی‌شود، بلکه مستلزم تعریف روشن اهداف اجتماعی و دانشگاهی و تلاش برای تحقق آنهاست. در بخش انتهایی، یک چارچوب ترکیبی پیشنهاد شده است که آزمون‌های مبتنی بر نظریه متغیر مکنون را با فرایندهای متخصص‌محور ادغام می‌کند. چنین مدلی، ضمن حفظ حداقلی از استانداردهای سازشی، امکان ارزیابی کل‌نگر و عادلانه‌تر داوطلبان را فراهم می‌آورد. بدین ترتیب، نظام پذیرش دانشجو در کشور ایران می‌تواند از سلطه تک‌بعدی عینیت مکانیکی عبور و به سوی الگویی چندبعدی، معتبرتر و سازگارتر با نیازهای علمی و اجتماعی جامعه ایرانی حرکت کند.

کلیدواژه‌گان: عینیت مکانیکی، ایده‌آل آزاد از ارزش، آزمون سراسری، فضایل معرفت‌شناختی، روایی.



## مقدمه

عینیت<sup>۱</sup> از دیرباز یک مفهوم بنیادی برای قضاوت منصفانه و بی‌طرفانه، به‌ویژه در فرایند پذیرش دانشجو<sup>۲</sup> بوده است (Gillispie, Au, 2013). به شکل سنتی، عینیت با یک نگرش مکانیکی<sup>۳</sup> (Daston & Galison, 2007) یا آنچه آرمان آزاد از ارزش<sup>۴</sup> (Douglas, 2009) نامیده می‌شود، یکسان در نظر گرفته شده، که هدف آن حذف سوگیری‌ها و احساسات شخصی با تکیه بر ملاک‌های استاندارد و قابل مشاهده است. این درک از عینیت، اتکا به آزمون‌های استاندارد و ارزیابی‌های الگوریتمی را در فرایند پذیرش دانشجو تقویت کرده و به دنبال تضمین عدالت<sup>۵</sup> با به حداقل رساندن تأثیر ذهنیت‌گرایی<sup>۶</sup> بوده است (Woo et al., 2023). با این حال، همان‌طور که داستون و گالیسون (Daston & Galison, 2007) استدلال کرده‌اند، عینیت مفهومی یکپارچه و همیشه حاضر در تاریخ علم نبوده است، بلکه به‌عنوان شیوه‌ای از قضاوت علمی<sup>۷</sup> از نظر تاریخی از طریق فضایل معرفتی<sup>۸</sup> رقیب تکامل یافته است که هر کدام از این فضایل خود به شیوه‌ای متفاوت، قضاوت علمی و تصمیم‌گیری‌های مبتنی بر علم را تحت تأثیر قرار داده‌اند. در حالی که عینیت مکانیکی اثرگذار است، اما به تنهایی برای ایجاد یک فرایند پذیرش دانشجو کارآمد و عادلانه کافی نیست. سایر فضایل معرفتی در پرتو سایر شیوه‌های قضاوت علمی، مانند حقیقت‌طبیعت<sup>۹</sup> و قضاوت‌تمرین‌شده<sup>۱۰</sup>، به همان اندازه برای دستیابی به عدالت، دسترسی همگانی<sup>۱۱</sup> و ارزیابی کل‌نگر<sup>۱۲</sup> متقاضیان ورود به آموزش عالی ضروری هستند.

فضایل معرفتی مجموعه‌ای از ویژگی‌ها، رویه‌ها یا تعهدات روش‌شناختی هستند که اندیشمندان را در جست‌وجوی دانش قابل اعتماد راهنمایی می‌کنند و چگونگی مشاهده، تفسیر و بازنمایی دنیای طبیعی توسط آنها را شکل می‌دهند (Zagzebski, Daston & Galison, 2007). فضایل معرفتی ایده‌آل‌سازی و ذات‌گرایی که در شیوه قضاوت علمی حقیقت‌طبیعت که قبل از ظهور عینیت مکانیکی برجسته بود، بر کشف ویژگی‌ها و الگوها و رای آنچه مستقیماً قابل مشاهده است، تأکید دارد و بنابراین، از منظر این رویکرد معیارهای سفت و سخت یا قوانین الگوریتمی گاهی ممکن است مانع از کشف ذات<sup>۱۳</sup> اساسی طبیعت شوند. همچنین در زمینه پذیرش دانشجو، از منظر این رویکرد، عدالت صرفاً حذف ذهنیت‌گرایی نیست، بلکه عدالت درباره کشف دقیق پتانسیل فکری، آموزشی، شخصیتی و غیره متقاضیان است که با صرف تکیه بر ملاک‌های مکانیکی اغلب نمی‌توان آنها را به تصویر کشید. قضاوت تمرین‌شده یکی دیگر از شیوه‌های قضاوت علمی است که با تأکید بر فضیلت معرفتی دقت<sup>۱۴</sup> در پاسخ به عینیت مکانیکی ظهور کرد. این رویکرد با برجسته کردن نقش ارزیاب‌های خبره، تأکید دارد که درک زمینه‌ای<sup>۱۵</sup> و مهارت‌های تفسیری ارزیاب‌ها، تصمیم‌گیری دقیق‌تری را درخصوص پیروی از قوانین ثابت و خشک ممکن می‌سازد. در مجموع، این فضایل معرفتی و شیوه‌های قضاوت علمی رقیب، با به چالش کشیدن سلطه عینیت مکانیکی از یک چارچوب ارزیابی گسترده‌تر و همچنین حساس به زمینه دفاع می‌کنند.

<sup>1</sup> objectivity

<sup>2</sup> university admission

<sup>3</sup> mechanical

<sup>4</sup> value-free ideal

<sup>5</sup> fairness

<sup>6</sup> subjectivity

<sup>7</sup> scientific judgment

<sup>8</sup> epistemic virtues

<sup>9</sup> truth-to-nature

<sup>10</sup> trained judgment

<sup>11</sup> inclusivity

<sup>12</sup> holistic

<sup>13</sup> essence

<sup>14</sup> accuracy

<sup>15</sup> contextual



خود مفهوم عینیت نیز در سال‌های اخیر کانون بحث‌های فلسفی زیادی بوده است (Cartwright et al., 2022; Douglas, 2000; Longino, 1990; Lloyd, 1995). داستون و گالیسون (۲۰۰۷) بیان کرده‌اند که عینیت مکانیکی (یا آرمان آزاد از ارزش) از نظر تاریخی در پاسخ به نگرانی‌ها درباره سوگیری<sup>۱</sup> علمی پدیدار شد. با این حال، دیدگاه‌های جدیدتر عینیت را به صورت پاک کردن قضاوت‌های انسانی از فرایند علمی مفهوم‌سازی نمی‌کنند، بلکه بر این باورند که عینیت یک مفهوم وابسته به موقعیت است (Cartwright et al., 2022; Koskinen, 2021) و باید با توجه به اهداف خاص که معمولاً از طریق گفت‌وگو و اجماع بین ذینفعان مختلف اتخاذ می‌شود، تعریف شود (Longino, 1990). این درک در حال تکامل از عینیت پیامدهای بسیاری برای فرایند پذیرش دانشجو دارد و نیاز به رویکردی جامع‌تر و فراگیر را نشان می‌دهد که به زمینه و هدف‌ها به اندازه دقت عددی اهمیت دهد. درک دقیق‌تر از عینیت و ادغام این فضایل معرفتی رقیب می‌تواند نحوه پذیرش دانشجو در کشور ایران را متحول کند. عینیت مکانیکی با ترویج بیش از حد استانداردسازی و یکسان‌سازی، اغلب عوامل اجتماعی، فرهنگی و ساختاری گسترده‌تر را که بر فرصت‌ها و عملکرد متقاضیان اثرگذار است، نادیده می‌گیرد.

اولین هدف در این مطالعه با تکیه بر بینش‌های داستون و گالیسون (Daston & Galison, 2007) و دیدگاه‌های کثرت‌گرایانه کارترایت و همکاران (Cartwright et al., 2022) و لونگینو (Longino, 1990) در مفهوم‌سازی مجدد از عینیت، ایجاد ارتباط بین پیشرفت‌ها و تغییرات ایجاد شده در حوزه فلسفه علم و تأثیرات بالقوه ناشی از پذیرش این دیدگاه‌ها در فرایند پذیرش دانشجو بود. دومین هدف، معرفی یک رویکرد پذیرش جایگزین با الهام و یکپارچه‌سازی نتایج به دست آمده در بخش اول (هدف اول) است. با توجه به اهداف بیان شده، در این مطالعه به دنبال پاسخگویی به سؤالات زیر بودیم:

۱. عینیت مکانیکی چگونه بر ساختار فعلی نظام پذیرش دانشجو در ایران، به ویژه بر اتکای آن به آزمون سراسری و امتحانات نهایی، تأثیر گذاشته است و این رویکرد چه چالش‌های اساسی‌ای را برای روایی آزمون‌ها ایجاد می‌کند؟
  ۲. فضایل معرفتی رقیب با عینیت مکانیکی (حقیقت طبیعت و قضاوت تمرین شده) چگونه می‌توانند بر فرایند پذیرش دانشجو تأثیرگذار باشند؟
  ۳. با توجه به تحولات اخیر در فلسفه علم، مفهوم عینیت در فرایند پذیرش دانشجو چگونه باید بازتعریف شود تا علاوه بر ملاک‌های کمی، اهداف مشخص اجتماعی و دانشگاهی را نیز در بر گیرد؟
  ۴. یک چارچوب ترکیبی و بهینه برای پذیرش دانشجو در ایران شامل چه مؤلفه‌هایی است و چگونه می‌توان آزمون‌های استاندارد را با فرایندهای ارزیابی تخصصی ادغام کرد تا به سیستمی عادلانه‌تر و معتبرتر دست یابیم؟
- برای رسیدن به این اهداف، ابتدا عینیت مکانیکی معرفی و چگونگی تأثیر این دیدگاه بر شیوه پذیرش دانشجو در کشور ایران توضیح داده و سپس درباره برخی از ایرادات ناشی از این نگرش مکانیکی بر روایی آزمون‌ها بحث شد. نگرش حقیقت‌طبیعت، دومین دیدگاهی است که معرفی و ارتباط آن با مدل‌های متغیر مکنون (که یک چارچوب مهم در آزمون‌سازی به شمار می‌رود) روشن شده است. در بخش بعد، در خصوص قضاوت تمرین شده و پیامدهای به کارگیری آن در فرایند پذیرش دانشجو بحث شده است. همان‌طور که بیان شد، مفهوم عینیت نیز در سال‌های اخیر تکامل یافته است. در این مطالعه سعی شد تا حدودی دلایل این تغییرات روشن و همچنین نگاه‌های نوین عینیت معرفی شود. سپس درباره پیامدهای اتخاذ این دیدگاه نوین از عینیت در پذیرش دانشجو بحث شده است. در نهایت، با بهره‌گیری و ترکیب این دیدگاه‌ها تلاش شد تا رهنمودهایی برای یک چارچوب سنجش نوین در فرایند پذیرش دانشجو ارائه شود.

<sup>1</sup> bias

## عینیت مکانیکی

عینیت مکانیکی یکی از شیوه‌های قضاوت علمی است که با محوریت قرار دادن بی‌طرفی<sup>۱</sup> به‌عنوان یک فضیلت معرفتی، در قرن نوزدهم میلادی به‌عنوان پاسخی به خطرهای درک شده ناشی از مداخله ذهنی در بازنمایی علمی پدیدار شد. این تغییر رویکرد صرفاً یک تحول روش‌شناختی نبود، بلکه یک ضرورت اخلاقی<sup>۲</sup> بود که دانشمندان را ملزم می‌کرد تا از تفسیر شخصی، آرایه‌های ادبی و ایده‌آل‌سازی نظری دست بردارند و در عوض، تصاویری دقیق و بی‌واسطه از طبیعت را بازنمایی کنند. به بیان داستون و گالیسون (Daston & Galison, 2007)، عینیت مکانیکی به دنبال «سرکوب مداخله عمدی<sup>۳</sup> نویسنده- هنرمند<sup>۴</sup>» از طریق استفاده از پروتکل‌های دقیق و دستگاه‌های مکانیکی بود تا از بی‌طرفانه‌ترین بازنمایی ممکن اطمینان حاصل شود.

تمایل به عینیت مکانیکی از این نگرانی فزاینده علمی ناشی می‌شد که ادراک و تفسیر انسان می‌تواند مشاهدات تجربی را مخدوش کند. خویش‌ن علمی که قبلاً مستقیماً درگیر ایده‌آل‌سازی طبیعت برای نمایش یک حقیقت کلی بود، اکنون ملزم به اعمال خویش‌ن‌داری است و به طبیعت اجازه می‌دهد تا «خودش [از طرف خودش] صحبت کند»<sup>۵</sup> (Daston & Galison, 2007)؛ به بیان دیگر، دانشمندان باید از ابزارها و روش‌هایی استفاده کنند که حتی در برابر اراده نادرست خود آنها نیز مقاوم باشد. هدف، حذف سوگیری‌های ذهنی و تحمیل نوعی نظم و خویش‌ن‌داری است که خلوص مشاهدات علمی را تضمین کند. اثر عینیت مکانیکی بر پذیرش دانشجو به دو شیوه مستقیم و غیر مستقیم قابل شناسایی است.

**شیوه مستقیم:** تطابق بین عینیت مکانیکی و نظام ورود به دانشگاه در کشور ایران را می‌توان از منظر دو اصل فلسفی توجیه کرد: پایایی معرفت‌شناختی<sup>۶</sup> و بی‌طرفی اخلاقی<sup>۷</sup>. توجیه اول از این فرض ناشی می‌شود که آزمون‌های استاندارد شده، پایاترین (مطمئن‌ترین) روش برای ارزیابی توانایی‌های فکری داوطلبان هستند، چراکه نتایج را به شکلی کمی و به دور از سوگیری‌های فردی بازتاب می‌دهند. همان‌طور که عینیت مکانیکی در علم به دنبال حذف تأثیر قضاوت انسانی در بازنمایی پدیده‌های طبیعی است، آزمون سراسری و امتحانات نهایی نیز به‌منظور حذف تأثیر عوامل ذهنی مانند پیشینه اقتصادی- اجتماعی، سرمایه اجتماعی<sup>۸</sup> یا روابط شخصی که ممکن است ارزیابی علمی را مخدوش کنند، طراحی شده است. این رویکرد، بازتاب تعهد مکتب اثبات‌گرایی به تأیید تجربی (وارسی‌گرایی<sup>۹</sup>) است. همان‌طور که ادعاهای علمی باید توسط داده‌های قابل مشاهده (یا قابل اندازه‌گیری) پشتیبانی شوند، شایستگی علمی نیز باید از طریق ملاک‌های کمی و قابل سنجش ارزیابی شود (Misak, 1995). ساختار پذیرش دانشجو در کشور ایران نیز با تکیه کامل بر آزمون‌ها (آزمون سراسری و آزمون‌های نهایی) با ملاک‌های دقیق نمره‌دهی و حذف ارزیابی اختیاری، نشان‌دهنده این رویکرد مکانیکی در قضاوت است؛ همان‌طور که اندازه‌گیری دما با دماسنج دیجیتال فقط به یک صورت ممکن است و بازتاب خواص فیزیکی و استانداردهای به‌کار گرفته شده در کالیبره کردن دستگاه است، نمرات تراز نهایی نیز فقط نشان‌دهنده توانایی‌های دانشجو فارغ از هرگونه دیدگاه‌های شخصی است.

<sup>1</sup> neutrality

<sup>2</sup> ethical

<sup>3</sup> willful

<sup>4</sup> artist-author

<sup>5</sup> speak for itself

<sup>6</sup> epistemic reliability

<sup>7</sup> ethical impartiality

<sup>8</sup> social capital

<sup>9</sup> verifiability

اصل دوم؛ یعنی بی‌طرفی اخلاقی بر این ایده استوار است که عدالت زمانی به بهترین شکل محقق می‌شود که همه داوطلبان در شرایط یکسان قضاوت شوند. مشابه با منطق تحقیقات آزمایشی، زمانی مقایسه تأثیر عمل آزمایشی در گروه‌ها ممکن است که شرایط برای همه گروه‌ها یکسان بوده باشد. در فرایند ورود به دانشگاه در ایران نیز معمولاً وسواس خاصی برای یکسان بودن شرایط از هر نظر وجود دارد. از نظر شرایط محیطی همه چیز از جمله تاریخ و زمان شروع و پایان آزمون یکسان است. محتوای مورد سنجش کاملاً برای تمام داوطلبان هر گروه آزمایشی یکسان و از پیش مشخص شده است. شیوه نمره‌دهی نیز در آزمون سراسری کاملاً مکانیکی است و در امتحانات نهایی با بیشترین دقت برای یکسان‌سازی شیوه تصحیح اوراق انجام می‌پذیرد.

در عمل، منطق به کارگیری عینیت مکانیکی (دو اصل پایایی معرفت‌شناختی و بی‌طرفی اخلاقی) درست مانند منطق آزمایش‌های علمی است که در آن شرایط کنترل‌شده و اندازه‌گیری دقیق به ما کمک می‌کنند تا مطمئن شویم نتایج فقط مربوط به پدیده‌هایی هستند که قصد مطالعه آنها را داشته‌ایم و نه عوامل دیگر (Michell, 2020).

**شیوه غیر مستقیم:** ارتباط دیگر عینیت مکانیکی با شیوه پذیرش دانشجو در کشور ایران مربوط به نحوه طراحی و ساخت آزمون سراسری است؛ به بیان دقیق‌تر، مدل‌های روانسنجی مختلفی برای توسعه آزمون‌های روانی و تربیتی وجود دارد که از این بین، نظریه کلاسیک آزمون<sup>۱</sup> برای تحلیل و توسعه سؤالات آزمون سراسری توسط سازمان سنجش آموزش کشور استفاده می‌شود. بهترین شیوه برای معرفی مدل کلاسیک، معرفی مدل صوری<sup>۲</sup> آن است:

$$x_p = \tau_p + e_p \quad (1)$$

در این معادله  $x_p$  نشان‌دهنده نمره مشاهده شده (کسب شده)،  $\tau_p$  نمره حقیقی و  $e_p$  نمره خطا برای شخص  $p$  است. ایده زیربنای این معادله بسیار ساده است. این معادله بیان می‌کند که نمره حقیقی فرد در یک آزمون، توسط خطای اندازه‌گیری مخدوش شده است؛ به بیان دیگر، نمره مشاهده شده شخص در آزمون از یک نمره حقیقی و مقداری خطای اندازه‌گیری تشکیل شده است. حال اگر ما بی‌نهایت بار برای شخص  $p$  آزمون مشخصی را تکرار کنیم، با فرض اینکه نتیجه آزمون‌های مختلف از یکدیگر مستقل باشند (برای مثال، با شست‌وشوی ذهنی شخص  $p$  بعد از هر آزمون) میانگین نمره مشاهده شده در این آزمون‌ها برابر با نمره حقیقی فرد در آن آزمون خواهد بود. به‌طور صوری، نمره حقیقی به‌عنوان امید ریاضی متغیر تصادفی<sup>۳</sup>  $X_p$  تعریف می‌شود ( $x_p$  یک تحقق<sup>۴</sup> از  $X_p$  محسوب می‌شود) که  $E()$  نشان‌دهنده امید ریاضی (میانگین) است:

$$\tau_p = E(X_p) \quad (2)$$

مهم‌ترین مفهوم در نظریه کلاسیک آزمون، نمره حقیقی است. اما این نمره حقیقی به چه نحوی قابل تفسیر است؟ رایج‌ترین تفسیر از نمره حقیقی در نظریه کلاسیک آزمون را لرد و ناویک (Lord & Novick, 1968) در کتاب *نظریه‌های آماری نمرات آزمون ذهنی*<sup>۵</sup> ارائه کرده‌اند.

<sup>1</sup> classical test theory

<sup>2</sup> formal

<sup>3</sup> random variable

<sup>4</sup> realization

<sup>5</sup> Statistical Theories of Mental Test Scores

این نویسندگان به تفسیر نمره حقیقی به صورت عملیاتی<sup>۱</sup> روی آورده‌اند. برای فهم بهتر اینکه منظور از تعریف نمره حقیقی به صورت عملیاتی چیست، باید به معادله (۲) توجه کرد. در این معادله نمره حقیقی به عنوان میانگین یا امید ریاضی نمرات مشاهده شده در یک آزمون مشخص تعریف شده است. به بیان لرد و نوویک (Lord & Novick, 1968): «... [تعریف عملیاتی از نمره حقیقی به مفهوم نمره واقعی به عنوان نمره (متوسط) است که یک فرد در آزمونی با طول بی‌نهایت به دست می‌آورد].»

پیش از ادامه بحث لازم است تا با ریشه‌ها و دلایل ایجاد تعریف عملیاتی<sup>۲</sup> در فلسفه علم آشنا شویم. مکتب و نگرش غالب فلسفه علم در ابتدای قرن بیستم میلادی مکتب اثبات‌گرایی<sup>۳</sup> بود. یکی از آرمان‌های مکتب اثبات‌گرایی حذف هستی‌شناسی<sup>۴</sup> یا به بیان دقیق‌تر، هر گزاره‌ای بود که نتوان به صدق آن از طریق داده‌های تجربی یا استدلال منطقی پی برد (آنچه به عنوان اصل وارسی‌گرایی شناخته می‌شد). این تعهد اثبات‌گرایان سبب شد تا به دنبال حذف هر اصطلاحی برآیند که نتوان معنای آن را با داده‌های تجربی مشخص کرد. یک گروه خاص از این اصطلاحات، اصطلاحات نظری<sup>۵</sup> بودند. در واقع، در نظریات علمی از اصطلاحاتی استفاده می‌شود که مستقیماً در جهان طبیعی قابل شناسایی نیستند. از دیدگاه اثبات‌گرایان هر اصطلاح نظری برای آنکه از نظر علمی قابل قبول در نظر گرفته شود، باید با اصطلاحات مشاهده‌پذیر رابطه تناظر<sup>۶</sup> داشته باشد:

(۳) اصطلاحات مشاهده‌پذیر  $\equiv$  اصطلاح نظری

برای مثال، مفهوم سرعت در نظریه مکانیک نیوتونی یک اصطلاح نظری است. از دیدگاه اثبات‌گرایان برای آنکه سرعت یک مفهوم نظری قابل قبول در نظر گرفته شود، باید بتوان این اصطلاح را به اصطلاحات قابل مشاهده تقلیل داد. سرعت در نظریه نیوتون از تقسیم مسافت بین دو نقطه بر بازه زمانی مشخص برای پیمودن آن مسافت ایجاد می‌شود و بنابراین، قابل تقلیل به اصطلاحات مشاهده‌پذیر است (باید دقت شود که زمان با وجود آنکه با چشم غیرمسلح قابل مشاهده نیست، ولی بدین دلیل که با دقت بالا قابل اندازه‌گیری است، یک اصطلاح مشاهده‌پذیر در نظر گرفته می‌شود). اما همیشه رابطه بین اصطلاحات مشاهده‌پذیر و نظری بدین صورت روشن نیست و به راه حلی برای ایجاد تناظر بین این اصطلاحات وجود نیاز دارد. یکی از نظریات بیان شده برای این منظور عملیات‌گرایی است. به زبان ساده، در عملیات‌گرایی اصطلاحات نظری مساوی با اعمالی هستند که در اندازه‌گیری آنها به کار گرفته شده است. برنهاده عملیات‌گرایی عموماً به فیزیکدان امریکایی، بریجمن، منتسب می‌شود. بریجمن (Bridgman, 1927) در کتاب معروف خود به نام منطق فیزیک مدرن<sup>۷</sup> این برنهاده را بدین شکل بیان کرده است: «منظور ما از هر مفهوم چیزی بیش از یک سری از عملیات‌ها نیست؛ مفهوم هم معنا با مجموعه‌ای از عملیات‌های متناظر است». همان‌طور که قابل مشاهده است، عملیات‌گرایی در اصل به عنوان یک برنهاده معنایی، در خصوص مفهوم اصطلاحات نظری پیشنهاد شده است. در عملیات‌گرایی اصطلاحات نظری با عملیات خاصی که برای مشاهده (اندازه‌گیری) آنها انجام می‌شود، به صورت کامل پوشش داده می‌شود.

<sup>1</sup> operationally  
<sup>2</sup> operational definition  
<sup>3</sup> positivism  
<sup>4</sup> ontology  
<sup>5</sup> theoretical terms  
<sup>6</sup> correspondence  
<sup>7</sup> the logic of modern physics

با توجه به مطالب بیان شده، می‌توان استدلال کرد که عینیت مکانیکی و عملیات‌گرایی دو روی یک سکه هستند. این ارتباط را عمدتاً از تعهد به رویه‌گرایی<sup>1</sup> و همچنین تعهد به مقابله با ذهنیت‌گرایی در تولید علم که اساس هر دو دیدگاه است، می‌توان استنباط کرد. همان‌طور که بیان شد، عملیات‌گرایی بر تعریف مفاهیم از طریق رویه‌هایی که در اندازه‌گیری آن به کار گرفته شده است، تأکید دارد. این شکل از تعریف رویه‌ای مانند یک دستورالعمل است که اگر رعایت شود، موجب درک ثابت و انتقال‌پذیر از مفهوم می‌شود. همین دستورالعمل‌ها اساس عینیت مکانیکی است که بر پایبندی به رویه‌ها و پروتکل‌های استاندارد برای ایجاد بازنمایی جهان تأکید دارد. علاوه بر این، هم عملیات‌گرایی و هم عینیت مکانیکی از یک سوء ظن عمیق نسبت به قضاوت ذهنی و پتانسیل سوگیری انسانی برای آلوده کردن دانش علمی ناشی می‌شوند. هر دو دیدگاه به دنبال پاکسازی و خالص کردن دانش از طریق خارج کردن عنصر انسانی (ذهنی) از فرایند تولید آن هستند. بنابراین، می‌توان نتیجه گرفت که عینیت مکانیکی به صورت غیرمستقیم و از طریق نظریه کلاسیک آزمون بر فرایند پذیرش دانشجو در دانشگاه‌های ایران اثرگذار بوده است.

تا اینجا درباره چگونگی اثرپذیری شیوه پذیرش دانشجو در ایران از دیدگاه‌های مرتبط با عینیت مکانیکی بحث شد. برای توجیه روند ادامه مقاله، ذکر برخی از ایرادات این نگرش اتخاذ شده ضروری است. اولین اثر جانبی ناشی از اتخاذ عینیت مکانیکی به روایی آزمون مربوط می‌شود. روایی یک مفهوم مورد مناقشه در حوزه آزمون‌سازی در نظر گرفته می‌شود، با این حال، بسیاری از متخصصان روایی را مربوط به توجیه‌پذیری تفاسیر مرتبط با نمرات آزمون در نظر می‌گیرند (Cronbach, 1971; Kane, 2006; Messick, 1989). یکی از این دیدگاه‌ها، چارچوب پیشنهادی کین (Kane, 2006, 2013, 2016) است. چارچوب مبتنی بر استدلال کین بر این نکته تأکید دارد که یک آزمون برای منظوری مشخص (یک تفسیر مشخص از نمرات) رواست، اگر شواهد (منطقی و تجربی) کافی برای تأیید (پذیرش در جمع متخصصان) این تفسیر وجود داشته باشد. اگرچه تا کنون توجه مناسبی به بحث رواسازی (آزمون‌های مرتبط با) شیوه پذیرش دانشجو در ایران صورت نگرفته، نگرش مکانیکی برخی ایرادات منطقی را برای روایی این آزمون‌ها ایجاد کرده است. همان‌طور که بیان شد، عینیت مکانیکی را می‌توان با منطق تحقیقات آزمایشی مقایسه کرد که تأکید بر کنترل تمام متغیرهای اثرگذار و همچنین استفاده از اندازه‌گیری‌های (کمی) دقیق است.

اگرچه این ویژگی‌ها (کنترل و کمی‌سازی) نقاط قوت این دیدگاه است، برخی ایرادات مبنایی را برای پذیرش دانشجو ایجاد کرده‌اند. در ابتدا برای محقق ساختن کنترل به ثابت نگه داشتن تمام شرایط برای داوطلبان نیاز است. این ثابت نگه داشتن فقط به شرایط زمینه‌ای آزمون (ها) مانند زمان برگزاری مربوط نیست، بلکه اساسی‌ترین ویژگی به محتوای آزمون در هر گروه آزمایشی مربوط می‌شود. برای مثال، در آزمون سراسری گروه علوم تجربی، مواد امتحانی ثابت و متشکل از زمین‌شناسی، زیست‌شناسی، ریاضی، فیزیک و شیمی است. همچنین مواد تشکیل‌دهنده امتحانات نهایی نیز برای تمام شرکت‌کنندگان ثابت و متشکل از دروس دینی، ریاضی، شیمی، سلامت و بهداشت، زیست‌شناسی، فیزیک، عربی، علوم اجتماعی، فارسی و زبان خارجی است. اما با توجه به رشته‌های متنوعی که از طریق آزمون سراسری گروه تجربی امکان پذیرش در آن وجود دارد (برای مثال، پزشکی، علوم تغذیه، بینایی‌سنجی، بهداشت محیط، شیمی محض و کاربردی، اقیانوس‌شناسی، مدیریت مالی، حسابداری و گردشگری)، این ایراد مطرح می‌شود که مواد امتحانی از نظر منطقی نمی‌توانند نشانگرهای مناسبی (روا) برای شایستگی پذیرش در تمام رشته‌های مذکور باشند. برای مثال، کسانی را که از طریق آزمون سراسری در رشته پزشکی پذیرفته می‌شوند، در نظر بگیرید. این افراد معمولاً در تمام دروس از جمله ریاضیات و فیزیک و شیمی درصدهای (نمرات) بالایی دارند (در غیر این صورت رتبه مناسب برای پذیرش در این رشته را کسب نمی‌کردند). این موضوع دقیقاً مشخص نیست که چرا فردی که می‌خواهد در رشته‌ای همچون پزشکی تحصیل

<sup>1</sup> proceduralism



کند، باید چنین وقت و تلاشی برای حوزه‌ای مانند فیزیک انجام دهد، چرا که این توانایی کاربرد زیادی برای بیشتر پذیرفته‌شدگان این رشته در آینده ندارد و منطقاً نمی‌تواند یک توانایی ضروری یا یک نشانگر مناسب برای موفقیت در رشته‌ای مانند پزشکی در نظر گرفته شود. یک راه حل احتمالی، تخصصی‌تر کردن نشانگرهای (مواد امتحانی) مربوط به هر گروه آزمایشی است. با این حال، این عمل نیز نمی‌تواند تمام مشکل را برطرف کند، زیرا این سؤال مطرح می‌شود که این تخصصی‌سازی تا چه میزان باید ادامه پیدا کند؟ آیا ایجاد یک آزمون با مواد مرتبط‌تر به رشته‌های پزشکی و پیرا پزشکی می‌تواند یک آزمون روا برای این گروه از رشته‌ها باشد یا اینکه هر کدام از این رشته‌ها نیاز به آزمون تخصصی خود خواهند داشت؟ یا حتی اینکه آیا یک آزمون تخصصی برای هر رشته می‌تواند تمام نشانگرهای مرتبط با یک رشته را در خود جای دهد؟ آنچه در اینجا قصد بیان آن را داریم، بسیار ساده است؛ فراهم کردن شواهد روایی برای آزمون‌هایی که شرط لازم و کافی برای ورود به یک رشته دانشگاهی (یا یک مجموعه رشته یا حتی یک گروه آزمایشی مانند علوم تجربی) در نظر گرفته می‌شود، (در بهترین حالت) بسیار دشوار است.

در پاسخ به این ایراد ممکن است این راه حل پیشنهاد شود تا رویه مبتنی بر آزمون‌ها برای پذیرش دانشجویان به‌عنوان شرط کافی در نظر گرفته نشود. اما درست در همین لحظه است که اثر جانبی دیگر نگرش عینیت مکانیکی خود را نشان می‌دهد. در واقع، به دلیل اهمیت کمی‌سازی در این نگرش این امکان وجود ندارد؛ به بیان دیگر، برای رسیدن به عینیت (مکانیکی) حداکثری به حذف تمام سوگیری‌ها و دخالت‌های انسانی نیاز است که یک جزء مهم در این میان، کمی‌سازی و مبنا قرار دادن این کمیته‌ها در فرایند تصمیم‌گیری است (مشابه آنچه در خصوص پایایی معرفت‌شناختی مطرح شد؛ به (Popper, 1959). بنابراین، اگر کمی‌سازی به معنای واقعی آن (Michell, 1990) هدف باشد، در نتیجه آزمون به شرط کافی برای پذیرش دانشجویان تبدیل می‌شود.

انتقاد آخر به مدل کلاسیک آزمون مرتبط است. یکی از ویژگی‌های هر نظریه آزمون‌سازی، توانایی ایجاد تمایز بین آزمون‌های مناسب و نامناسب است که در نظریه کلاسیک آزمون مفهوم پایایی این نقش را بر عهده دارد. در این چارچوب، پایایی به‌صورت مجذور همبستگی بین نمره حقیقی و مشاهده‌شده یک آزمون در جامعه هدف تعریف شده است. ابتدا نگاهی به همبستگی بین نمره مشاهده شده و حقیقی داشته باشیم:

$$\rho_{XT} = \frac{\sigma_{XT}}{\sigma_X \sigma_T} \quad (4)$$

در این معادله  $\sigma_{XT}$  نشان‌دهنده کوواریانس نمرات مشاهده شده و حقیقی،  $\sigma_X$  انحراف استاندارد نمرات مشاهده شده،  $\sigma_T$  انحراف استاندارد نمرات حقیقی و در نهایت،  $\rho_{XT}$  همبستگی بین نمرات حقیقی و مشاهده شده است. برای ساده‌تر کردن معادله (4)، کوواریانس بین نمره مشاهده شده و حقیقی را می‌توان بدین صورت محاسبه کرد:

$$\sigma_{XT} = \sigma(T + E, T) = \sigma_T^2 + \sigma(E, T) = \sigma_T^2 \quad (5)$$

با جایگزینی نتیجه معادله (5) در معادله (4) خواهیم داشت:



$$\rho_{XT} = \frac{\sigma_T}{\sigma_X} \quad (6)$$

حال، مجذور همبستگی بین نمره مشاهده شده و حقیقی برابر خواهد بود با:

$$\rho_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_E^2 + \sigma_T^2} \quad (7)$$

بنابراین، در چارچوب مدل کلاسیک آزمون، پایایی برابر با نسبتی از واریانس نمره کل (مشاهده شده) آزمون است که به نمره حقیقی تعلق دارد. اگرچه مفهوم پایایی آن گونه که در نظریه کلاسیک آزمون معرفی شد، می‌تواند نشان‌دهنده خوبی از ثبات نمرات آزمون باشد، با این حال، معیار دقیقی در تمایزگذاری بین آزمون مناسب و نامناسب نیست. برای درک بهتر موضوع، این سناریو خیالی را در نظر بگیرید (Borsboom, 2005):

شخصی قصد ساخت آزمونی به منظور سنجش مهارت‌های محاسباتی را دارد. این شخص برای سنجش این توانایی از سؤالات زیر بهره برده است:

۱. دوست دارم برای ادامه تحصیل به کشور دیگری مهاجرت کنم.

۲.  $(1234 - 456) * (12 + 12 - 24) = ?$

۳.  $123 - 12 + 44 = ?$

۴.  $\frac{(1+1+165)}{876*56-3} = ?$

۵. من مسافرت به کشورهای شرقی را به غربی ترجیح می‌دهم.

۶. منطقه خاورمیانه نسبت به اروپا از جاذبه‌های گردشگری بیشتری برخوردار است.

پاسخ شرکت‌کنندگان به این آزمون فرضی جمع‌آوری و با استفاده از روش‌های موجود، پایایی نمرات محاسبه شده است. پایایی حاصل شده برابر با ۰٫۹۵ بود (با فرض اینکه پایایی به صورت مستقیم قابل محاسبه است) که نشان‌دهنده پایایی عالی نمرات است. اما آیا این آزمون یک آزمون مناسب برای سنجش مهارت‌های محاسباتی است؟ احتمالاً شما نیز با ما هم نظر هستید که این گونه نیست. این مشکل از کجا ناشی می‌شود؟ همان‌طور که بیان شد، در نظریه کلاسیک آزمون مهم‌ترین مفهوم، نمره حقیقی است که به صورت ارزش مورد انتظار (میانگین) نمرات یک آزمودنی در آزمونی با طول بی‌نهایت تعریف می‌شود. اما در این نظریه هیچ رابطه‌ای بین هر یک از سؤالات و نمره حقیقی در نظر گرفته نشده است؛ به عبارت دیگر، این نظریه آزمون محور است و نه سؤال محور. بدین ترتیب، سازه مورد بررسی که با استفاده از این نظریه سنجیده می‌شود، با توجه به آزمون مورد استفاده در سنجش آن قابل تعریف است که این آزمون می‌تواند شامل هر سؤالی باشد. این ضعف به استفاده از تعریف عملیاتی در تعریف نمره حقیقی مربوط است. در تعریف عملیاتی، هر سازه (اصطلاح نظری) مساوی با اعمالی است که در اندازه‌گیری آن به کار برده شده است. به همین دلیل سازه فقط یک نام برای مجموعه سؤالاتی است که در آزمون استفاده شده‌اند و در این چارچوب، رد اینکه سازه مورد ادعا اندازه‌گیری نشده است، ناممکن است (ابطال‌پذیر نبودن) (Vessonen, 2021). بنابراین، در چارچوب نظریه کلاسیک آزمون، یک آزمون می‌تواند پایا باشد، بدون آنکه اطلاعاتی در خصوص روایی آن در اختیار ما قرار دهد که نقطه ضعفی بزرگ محسوب می‌شود.



## حقیقتِ طبیعت

حقیقتِ طبیعت شیوه‌ای از قضاوت علمی است که با برجسته کردن فضیلت معرفت‌شناختی ایده‌آل‌سازی، به دنبال گذر کردن از تفاوت‌های فردی در پدیده مورد بررسی است. از نظر داستون و گلیسن (Daston & Galison, 2007)، هدف نگرش حقیقتِ طبیعت رسیدن به ذات یک پدیده طبیعی است؛ به بیان دیگر، هدف این دیدگاه بازنمایی طبیعت به شکلی است که به یک مشاهده تجربی خاص محدود نشود، بلکه یک طرح زیربنایی را بازتاب دهد. دلیل اصلی توسعه نگرش حقیقتِ طبیعت مقابله با تغییرات بیش از حد در طبیعت بود که باعث ایجاد دشواری در بازنمایی می‌شد. برای مثال، طبیعت‌شناسان تمرکز بیش از حد بر ناهنجاری‌های طبیعی را تهدیدی برای ثبات طبقه‌بندی‌های علمی می‌دانستند؛ به بیان دیگر، طبیعت‌شناسان به جای فهرست تمام انحرافات در هر گونه جانوری، سعی داشتند تا ویژگی‌های اساسی این گونه‌ها را استخراج کنند. به همین دلیل دانشمندان در مقابل طبیعت منفعلانه عمل نمی‌کردند، بلکه دست به انتخاب، مقایسه و ترکیب فعالانه پدیده‌ها می‌زدند تا (حقیقتِ) طبیعت را همان‌گونه که هست بازنمایی کنند. بنابراین، دانشمندان به جای فقط ثبت آنچه دیده می‌شد، سعی در تمایز ویژگی‌های ضروری از ویژگی‌های غیرضروری داشتند.

حقیقتِ طبیعت پیش از استیلای عینیت مکانیکی بر نگرش علمی و در قرون ۱۸ و ۱۹ میلادی برجسته بود. با این حال، این بدان معنا نیست که عینیت مکانیکی به‌طور کلی، جایگزین این نگرش شده باشد. حقیقتِ طبیعت همچنان موضوعی مهم در فلسفه علم باقی مانده است، چرا که بر تعامل بین مشاهده تجربی و تفسیر نظری تأکید دارد. اما آنچه در این مطالعه حایز اهمیت است، ارتباط نگرش حقیقتِ طبیعت به آزمون‌سازی با نگاه ویژه به آزمون‌های میزان‌شده (استاندارد شده) است. آنچه حقیقتِ طبیعت و آزمون‌سازی را به یکدیگر مرتبط می‌کند، وجود رگه‌هایی از ذات‌گرایی<sup>۱</sup> در مدل‌های متغیر مکنون است که چارچوب مهمی در حوزه آزمون‌سازی محسوب می‌شود. در ادامه، چگونگی این ارتباط مورد بحث قرار گرفته است.

**ذات‌گرایی در مدل‌های متغیر مکنون:** مدل‌های متغیر مکنون به دسته‌ای از مدل‌های گوناگون اطلاق می‌شوند. ساخت و استفاده از اولین مدل متغیر مکنون را می‌توان به اسپیرمن (Spearman, 1904) نسبت داد. اسپیرمن در مقاله‌ی پر استناد خود به نام «هوش عمومی، عیناً مشخص و اندازه‌گیری شده» روشی را معرفی کرد که بسیار شبیه به روش‌های تحلیل عاملی امروزی بود. پس از اسپیرمن، توسعه مدل‌های متغیر مکنون در حوزه‌های مختلف و به‌صورت مستقل از یکدیگر ادامه یافت. از مدل‌های پرسش‌پاسخ در روان‌ناسی و علوم تربیتی (Lord & Novick, 1968; Rasch, 1960)، مدل‌های کلاس مکنون در جامعه‌شناسی و اقتصاد (Heinen, 1996; Lazarsfeld & Henry, 1968) و مدل‌های تحلیل عاملی (Joreskog, 1971) می‌توان به‌عنوان مهم‌ترین دسته از مدل‌های توسعه‌یافته متغیر مکنون نام برد. مدل‌های متغیر مکنون بیان می‌کنند که داده‌های مشاهده‌شده  $y_i$  از ساختار زیربنایی و غیرقابل مشاهده  $\eta$  ناشی می‌شوند. یک فرمول‌بندی رایج تحلیل عاملی به شرح زیر است:

$$y_i = \mu_i + \lambda_i \eta + \varepsilon_i \quad (1)$$

که در آن  $\mu_i$  عرض از مبدأ،  $\lambda_i$  بارهای عاملی و  $\varepsilon_i$  خطاهای تصادفی است که فرض می‌شود مستقل از  $\eta$  است. برای مدل‌هایی که با پاسخ‌های دو حالتی (مانند پاسخ‌های بلی/خیر یا درست/غلط) سروکار دارند، مانند نظریه سؤال پاسخ (IRT)، معمولاً این‌گونه فرمول‌بندی می‌شوند:

<sup>1</sup> essentialism



$$P(Y_{ij} = 1 | \theta_j) = \frac{\exp(-\alpha_i(\theta_j - b_i))}{1 - \exp(-\alpha_i(\theta_j - b_i))} \quad (2)$$

که  $\theta_j$  سطح صفت مکنون در آزمودنی  $j$  ام،  $\alpha_i$  پارامتر تمیز سؤال (متغیر مشاهده شده) و  $b_i$  پارامتر دشواری سؤال  $i$  ام هستند. در حوزه‌های مرتبط با سنجش روانشناختی و آموزشی،  $\eta$  یا  $\theta$  می‌تواند نشان‌دهنده ساختارهای مکنونی مانند هوش، افسردگی، مهارت ریاضی و غیره باشد، با این شرط که این ساختارهای مکنون علت (یا ذات) مشترک تغییرات بین متغیرهای مشاهده شده هستند. بنابراین، یک فرض اساسی در مدل‌های متغیر مکنون، استقلال موضعی<sup>۱</sup> است؛ به بیان دیگر، همبستگی بین متغیرهای مشاهده شده حاکی از ارتباط مستقیم بین آنها نیست، بلکه ناشی از تأثیرپذیرفتن از یک متغیر مشترک است. بنابراین، رابطه بین متغیرهای مشاهده شده مشروط به وجود یک عامل مکنون است. به زبان صوری:

$$P(Y_1, Y_2, \dots, Y_k) = P(Y_1 | \theta) P(Y_2 | \theta) \dots P(Y_k | \theta) \quad (3)$$

در این معادله  $Y_1, Y_2, \dots, Y_k$  متغیرهای تصادفی مشاهده شده،  $\theta$  بردار متغیرهای مکنون،  $P(Y_1, Y_2, \dots, Y_k)$  احتمال مشترک متغیرهای مشاهده شده و  $P(Y_1 | \theta) P(Y_2 | \theta) \dots P(Y_k | \theta)$  ضرب احتمالات شرطی است.

تفسیر متغیرهای مکنون در مدل‌های متغیر مکنون به دو شکل توصیفی و واقع‌گرایانه ممکن است. تفسیر واقع‌گرایانه از مدل‌های مکنون به دلیل مطلوبیت معرفت‌شناختی بیشتر بر تفسیر توصیفی ارجحیت دارد (به (Karvandi Renani et al., 2024) نگاه کنید). مدل‌های متغیر مکنون بدین دلیل که ادعای وجود داشتن ساختارهای زیربنایی و تغییرناپذیری را مطرح می‌کنند که دسترسی معرفت‌شناختی به آنها ممکن نیست، رویکردی ذات‌گرایانه دارند (Borsboom & Cramer, 2013; Borsboom, 2008; Borgstede & Eggert, 2023). برای مثال، صورت‌بندی ارائه‌شده از تحلیل عاملی در معادله (۸)، به صورت ضمنی نشان‌دهنده دیدگاه ذات‌گرایانه است، چرا که ادعا دارد تغییرات بین متغیرهای مشاهده شده  $Y_i$  در نهایت، به ذات مکنون  $\eta$  مربوط می‌شود که در موقعیت‌های متفاوت ثابت می‌ماند. برای مثال، ساختار هوش را در تاریخچه سنجش روانی در نظر بگیرید. در مدل‌های الهام گرفته از  $g$  اسپیرمن، متغیر مکنون  $g$  به‌عنوان یک ظرفیت شناختی بنیادی در نظر گرفته می‌شود که به‌طور علی عملکرد افراد در فعالیت‌های مختلف را تحت تأثیر قرار می‌دهد. در اینجا  $g$  به‌عنوان ذات زیربنایی توانایی‌های شناختی مختلف در نظر گرفته شده است، یک ویژگی پایدار و واقعی که همبستگی‌های مشاهده‌شده بین نمرات آزمون‌های مختلف را تبیین می‌کند، با این حال در هیچ‌یک از رفتارهای مشاهده شده به صورت کامل نمود پیدا نکرده است.

در خصوص شیوه پذیرش دانشجو این تفسیر ذات‌گرایانه از مدل‌های متغیر مکنون بدین دلیل اهمیت دارد که ویژگی‌های (های) مکنون زیربنایی در این مدل‌ها، تحت یک خوانش واقع‌گرایانه به‌عنوان ظرفیت واقعی تحصیلی یا توانایی شناختی در نظر گرفته می‌شود که در موقعیت‌های متفاوت آزمون‌گیری ثابت است. بنابراین، نمرات آزمون فرضاً بازتاب ذات واقعی توانایی‌های تحصیلی دانش‌آموزان هستند و نه نوسانات عملکردی گذرا یا خطای اندازه‌گیری. استفاده از مدل‌های متغیر مکنون، و به‌صورت خاص مدل‌های سؤال‌پاسخ، در ساخت و توسعه آزمون‌های سرنوشت‌ساز در سراسر جهان رایج است. برای مثال، در کشور آمریکا استفاده از مدل‌های سؤال‌پاسخ در ابتدا برای ساخت آزمون‌های سرنوشت‌سازی همچون

<sup>1</sup> local independence



SAT و GRE در دهه ۸۰ میلادی استفاده شد (Lord, 1980). در کشور انگلستان بخش عمده‌توسعه‌آزمون‌های مرتبط با گواهی عمومی آموزش متوسطه<sup>۱</sup> (GCSE) با استفاده از مدل‌های سؤال‌پاسخ انجام می‌گیرد. در کشور استرالیا نظریه سؤال‌پاسخ (IRT) به‌طور گسترده در برنامه ملی ارزیابی - خواندن و حساب (NAPLAN) به‌کار گرفته شده است (Griffin, 2018). همچنین این مدل‌ها در آزمون‌های ورودی دانشگاه در کشور ترکیه نیز نقشی کلیدی دارند (Bulut, 2016).

استفاده گسترده از مدل‌های سؤال‌پاسخ بی‌دلیل نبوده است. برای مثال، ایراداتی همچون وابسته به نمونه بودن<sup>۲</sup> در نظریه کلاسیک آزمون با توجه به اندازه‌گیری تغییرناپذیر<sup>۳</sup> در نظریه سؤال‌پاسخ به یک مزیت در این نظریه تبدیل شده است که از قابلیت مقایسه نمرات در اجراهای متفاوت آزمون پشتیبانی می‌کند (Hambleton et al., 1991). همچنین توانایی نظریه سؤال‌پاسخ در تحلیل تک‌تک سؤالات آزمون با در نظر گرفتن پارامترهای دشواری، تمیز و حدس زدن، به ارزیابی‌های هدفمندتر و کارآمدتر منجر شد (Baker & Kim, 2018). علاوه بر این، این نظریه امکان شناسایی عملکرد افتراقی سؤالات<sup>۴</sup> و در نتیجه، ایجاد آزمون‌هایی با سؤالاتی را که برای گروه‌های مختلف عملکرد مشابهی داشته باشند، فراهم کرده است (Ohiri et al., 2024). نظریه سؤال‌پاسخ امکان آزمون‌دهی تطبیقی رایانه‌ای<sup>۵</sup> (CAT) را فراهم می‌کند که امکان ارائه سؤالات بر اساس توانایی داوطلبان را ممکن می‌کند که موجب ارزیابی بهینه‌تر و صرفه‌جویی در زمان و هزینه می‌شود. بنابراین، استفاده از مدل‌های سؤال‌پاسخ می‌تواند موجب بهبود فرایند آزمون‌سازی در ایران باشد.

### قضاوت تمرین‌شده

قضاوت تمرین‌شده به‌عنوان شیوه‌ای از قضاوت علمی تا حدودی موضعی مابین عینیت مکانیکی و حقیقت طبیعت دارد، اما اساساً با هر دو آنها متفاوت است. این نگرش به انکارناپذیر بودن نقش تخصص انسانی در تفسیر داده‌ها اذعان دارد و بنابراین، تولید دانش را صرفاً یک فرایند مکانیکی نمی‌داند، بلکه بینش حاصل از تجربه را جزء اساسی از این چرخه به‌شمار می‌آورد. با توجه به نظر داستون و گالیسون (Daston & Galison, 2007)، «دید علمی<sup>۶</sup> به یک هنر تجربی<sup>۷</sup> تبدیل شده بود» که ادعا دارد دیدن صرفاً عمل مشاهده نیست، بلکه کنشی از قضاوت تمرین‌شده است. این نگرش در نوشتارهای پوپر (Popper, 1959) به خوبی رخ نشان داده است:

«اما در واقع، این باور که ما می‌توانیم فقط با مشاهدات خالص و بدون هیچ چیزی در ماهیت نظریه شروع کنیم، پوچ است. [...] بیست و پنج سال پیش سعی کردم در شروع یک سخنرانی با دستورالعمل‌های زیر، همین نکته را برای گروهی از دانشجویان فیزیک در وین بیان کنم: «مداد و کاغذ بردارید؛ با دقت مشاهده کنید و آنچه را مشاهده کرده‌اید، بنویسید!» آنها البته، پرسیدند که من [از آنها] می‌خواهم چه چیزی را مشاهده کنند. به وضوح دستورالعمل «مشاهده کنید» پوچ است. (مشاهده کنید) حتی اصطلاحی نیست، مگر اینکه مفعول فعل متعدی را بتوان فهمید. مشاهده همیشه انتخابی است. [مشاهده کردن] به یک شیء انتخاب شده، به یک کار مشخص، یک علاقه، یک دیدگاه و یک مسئله نیاز دارد و وصف آن مستلزم زبانی توصیفی با واژه‌های خاص است؛ این امر مستلزم تشابه و طبقه‌بندی است که خود منافع، دیدگاه‌ها و مشکلات را پیشفرض می‌گیرد».

<sup>1</sup> General Certificate of Secondary Education

<sup>2</sup> sample dependency

<sup>3</sup> invariant measurement

<sup>4</sup> differential item functioning

<sup>5</sup> Computerized Adaptive Testing

<sup>6</sup> scientific sight

<sup>7</sup> empirical art



بنابراین، اصلی‌ترین خصیصه دیدگاه قضاوت تمرین‌شده، رد رویه‌های صرفاً مکانیکی و استفاده از بینش متخصصان در تولید و تشخیص علمی است. داستون و گالیسون (Daston & Galison, 2007) نشان دادند که رویکرد قضاوت تمرین‌شده عمیقاً در مدل کارآموزی<sup>۱</sup> ریشه دارد که در آن، تخصص به مرور زمان پرورش می‌یابد و فقط حاصل پیروی از رویه‌های خشک قانون‌مآبانه نیست. در بسیاری از علوم این باور پدیدار شده بود که شناخت الگوها در پدیده‌های مورد مطالعه صرفاً با به‌کارگیری قواعد از پیش تعیین شده ناممکن است و قضاوت افراد متخصص در این میان نقشی انکارناپذیر دارد. در حوزه‌هایی مانند الکتروانسفالوگرافی<sup>۲</sup> (مغزنگاری الکترونیک) و نجوم، توانایی شناسایی الگوهای مرتبط نه از مجموعه‌ای دقیق از ملاک‌های قابل اندازه‌گیری، بلکه از آشنایی با ساختارهای تکرار شونده و درک شهودی از تغییرات ناشی می‌شود. بنا بر دیدگاه داستون و گالیسون این توانایی را نمی‌توان به توانایی تشخیص ویژگی‌های ذاتی (آن‌گونه که در حقیقت طبیعت مطرح شد) ارتباط داد، بلکه نشان‌دهنده توانایی متخصصان در پیدا کردن شباهت‌های خانوادگی<sup>۳</sup> است. این مفهوم ویتگنشتاینی<sup>۴</sup> انحراف از شیوه‌های قضاوت علمی قدیمی‌تر همچون حقیقت طبیعت را که به دنبال ثبت ذات واحد یک نوع (یا گونه) بود و همچنین عینیت مکانیکی را، که بر بازتولید جزئیات بدون دخالت انسان تأکید داشت، برجسته می‌کند.

بعد اخلاقی قضاوت تمرین‌شده نیز قابل توجه است. برخلاف عینیت مکانیکی که بر جداسازی سختگیرانه بین مشاهده‌گر و شیء مورد بررسی تأکید دارد، قضاوت تمرین‌شده اذعان می‌کند که متخصص، مسئول شکل‌دهی دانش به‌گونه‌ای است که به پیشرفت پژوهش علمی خدمت کند. باید توجه داشت که این به معنای بازگشت به فردگرایی به شیوه صاحب‌نظران خبره در علم نیست، بلکه ادغامی از تخصص با تحلیل مبتنی بر ابزار است. این رویکرد ترکیبی به متخصصان اجازه داد تا به داده‌ها به‌عنوان ابزارهای کارکردی نگاه کنند و نه به‌عنوان یک قطعه بازتاب‌دهنده (یا یک پازل از) طبیعت.

تغییر دیدگاه به سمت قضاوت تمرین‌شده بازتاب تحولات تاریخی گسترده‌تر در حوزه معرفت‌شناسی است. این گذار علمی را می‌توان از تبدیل استعاره «عینیت نباید فدای دقت شود» به «دقت نباید فدای عینیت شود» متوجه شد (Niaz, 2017). عینیت مکانیکی به‌عنوان واکنشی در برابر ذهنیت‌گرایی ادراک‌شده در رویکرد حقیقت طبیعت پدید آمد، در حالی که قضاوت تمرین‌شده در مقابله با محدودیت‌های تولید دانش ماشینی‌شده ایجاد شد و بر اهمیت نقش متخصصان در درک زمینه برای تفسیر داده‌ها اشاره داشت. همان‌طور که داستون و گالیسون (Daston & Galison, 2007) استدلال می‌کنند، این تغییر به معنای رد کامل فضایل معرفتی قبلی نیست، بلکه نشان‌دهنده تغییر اولویت‌هاست. در ادامه پیامدهای به‌کارگیری رویکرد قضاوت تمرین‌شده بر رویه پذیرش دانشجو بیان شده است.

**قضاوت تمرین‌شده در پذیرش دانشجو:** قضاوت تمرین‌شده بر ضرورت استفاده از تخصص انسانی در تصمیم‌گیری تأکید دارد؛ این رویکرد بر این باور است که نمی‌شود همه چیز را با قواعد از پیش تعیین شده و یکسان پیش برد، چرا که تمام عوامل اثرگذار را نمی‌توان در قالب یک الگو و رویه مشخص خلاصه کرد و بنابراین، باید بر توانایی متخصصان در شناسایی الگوهای معنادار از الگوهای تصادفی تکیه کرد. این طرز تفکر با روش‌های پذیرش جدید که در بسیاری از کشورها به‌کار گرفته می‌شود، سازگار است. در این روش‌ها متخصصان آموزش دیده عوامل گوناگونی را برای سنجش داوطلبان بررسی می‌کنند و صرفاً بر نمره آزمون در پذیرش متکی نیستند. برای مثال، در فرایندهای پذیرش کل‌نگر<sup>۵</sup> در ایالات متحده آمریکا و اروپا از کمیته‌های آموزش‌دیده‌ای استفاده می‌شود که نمرات آزمون را در کنار مقالات، فعالیت‌های فوق برنامه و اظهارات شخصی برای پذیرش ارزیابی می‌کنند. به‌واقع توانایی شهودی این ارزیابان که ریشه در قضاوت تمرین‌شده آنها دارد، این امکان

<sup>1</sup> apprenticeship model

<sup>2</sup> electroencephalography

<sup>3</sup> family resemblances

<sup>4</sup> Wittgenstein

<sup>5</sup> holistic



را فراهم می‌سازد تا افراد مستعد را حتی زمانی که بهترین نمرات را ندارند، شناسایی کنند. اگرچه این قضاوت ممکن است کاملاً با تکیه بر قواعد از پیش تعیین شده نباشد (نقض عینیت مکانیکی)، با این حال، به دقت و حساسیت فرایند پذیرش کمک می‌کند. برای نمونه، موسسه فناوری ماساچوست<sup>۱</sup> (MIT) که در رتبه‌بندی دانشگاه‌های برتر QS<sup>۲</sup> در رتبه نخست قرار دارد<sup>۳</sup>، در نظام پذیرش خود (در دوره کارشناسی) بر ترکیبی از آزمون‌ها و مصاحبه متکی است. در ابتدا برای ورود به فرایند پذیرش نیاز است که هر داوطلب در یکی از آزمون‌های ACT یا SAT شرکت و حداقل نمره قابل قبول برای شرکت در فرایند پذیرش را کسب کند<sup>۴</sup>. در بخش «مقالات، فعالیت‌ها و مسائل آکادمیک» هر داوطلب به برخی سؤالات (عمدتاً در خصوص علایق و نحوه برخورد با مسائل) پاسخ می‌دهد و یک کاربرگ از سوابق تحصیلی (شبهه به رزومه ولی در قالب مشخص دانشگاه) فراهم می‌کند<sup>۵</sup>. همچنین هر داوطلب به ارائه یک رونوشت از سوی مشاور مدرسه نیاز دارد که در خصوص عملکرد وی در مدرسه اطلاعات لازم را فراهم کند<sup>۶</sup>. علاوه بر این، دو توصیه‌نامه نیز باید از معلم‌های هر داوطلب ارائه شود که بنا بر ترجیح دانشگاه، بهتر است یکی از توصیه‌نامه‌ها از معلم ریاضیات یا علوم و دیگری از معلم‌های انسانی، علوم اجتماعی و یا ادبیات باشد<sup>۷</sup>. دانشگاه MIT تأکید دارد که فقط به اسناد ارائه شده قناعت نکند و به همین دلیل، یک مصاحبه نیز در فرایند پذیرش گنجانده شده است<sup>۸</sup>. در نهایت، زمانی که فرایند تکمیل اطلاعات از سوی داوطلب پایان می‌یابد، ابتدا یکی از افسران ارشد پذیرش درخواست داوطلب را به صورت جامع و در بستر مناسب آن بررسی می‌کند. سپس درخواست‌های قوی توسط افسران پذیرش دیگر ارزیابی و برای کمیته پذیرش خلاصه می‌شوند. این خلاصه‌ها، همراه با کل اطلاعات داوطلب، به کمیته انتخاب ارسال می‌شوند، جایی که گروه‌های مختلفی از کارکنان پذیرش و اعضای هیئت علمی نظر خود را اعلام می‌کنند. حداقل دوازده نفر به‌طور جدی درباره یک درخواست بحث و بررسی می‌کنند، قبل از آنکه در فهرست پذیرش قرار گیرد<sup>۹</sup>. با پیاده‌سازی مدل قضاوت تمرین‌شده در فرایند پذیرش دانشجو انتظار می‌رود مسئولان پذیرش به متخصصانی بدل شوند که قادر به سنجش دقیق و همه‌جانبه توانایی‌های علمی متقاضیان باشند. بدین معنا که به جای رتبه‌بندی مکانیکی دانشجویان بر اساس نمرات آزمون، ارزیابان طیف کامل‌تری از دستاوردهای متقاضی را ارزیابی و عملکرد آنها را با در نظر گرفتن زمینه شخصی و اجتماعی - اقتصادی‌شان بررسی کنند. این روش موجب می‌شود که ارزیابی دقیق‌تر و منصفانه‌تری صورت گیرد و احتمال نادیده گرفته شدن دانشجویان با استعدادی که شاید در آزمون‌های استاندارد نتوانند بهترین عملکرد را داشته باشند، کمتر شود.

## عینیت در مسیر تکامل

عینیت در مقابل ذهنیت معنا پیدا می‌کند. منظور از ذهنیت در علم وارد شدن ارزش‌های شخصی، اجتماعی، اخلاقی و سیاسی در فرایند علمی است. این نگرش نسبت به عدم دخالت ارزش‌های شخصی، اجتماعی، اخلاقی و سیاسی در علم با نام «ایده‌آل آزاد از ارزش»<sup>۱۰</sup> شناخته می‌شود (Douglas, 2009). اما در کنار این ارزش‌های غیر معرفت‌شناختی<sup>۱۱</sup> می‌توان برخی از ارزش‌های معرفت‌شناختی را نیز شناسایی کرد. به بیان ساده، ارزش‌های معرفت‌شناختی ملاک‌هایی هستند که دانشمندان از آنها بهره می‌جویند تا بتوانند بین مدل‌ها، نظریه‌ها و دیدگاه‌های موجود

<sup>1</sup> Massachusetts Institute of Technology

<sup>2</sup> <https://www.topuniversities.com/world-university-rankings>

<sup>3</sup> جست‌وجوها در تاریخ ۱۷ اسفند ۱۴۰۳ ساعت ۳۰ دقیقه بامداد به وقت محلی تهران انجام شده است

<sup>4</sup> <https://mitadmissions.org/apply/firstyear/tests-scores/>

<sup>5</sup> <https://mitadmissions.org/apply/firstyear/essays-activities-academics/>

<sup>6</sup> <https://mitadmissions.org/apply/firstyear/transcripts/>

<sup>7</sup> <https://mitadmissions.org/apply/firstyear/letters-of-recommendation/>

<sup>8</sup> <https://mitadmissions.org/apply/firstyear/interview/>

<sup>9</sup> <https://mitadmissions.org/apply/process/selection/>

<sup>10</sup> value-free ideal

<sup>11</sup> non epistemic values





بهترین گزینه را مشخص کنند (McMullin, 1982). از جمله ارزش‌های معرفت‌شناختی می‌توان به دقت پیش‌بینی، توانایی یکپارچه‌سازی، ثبات بیرونی، انسجام درونی و بارورسازی اشاره کرد. باید دقت داشت که این فهرست از ارزش‌ها بسته نیستند و می‌توان شاخص‌های دیگری را نیز به آن وارد کرد. از سوی دیگر، توجه به این نکته مهم است که این ملاک‌ها دستورالعمل‌های ثابت و تغییرناپذیری را برای ترجیح یا عدم ترجیح یک نظریه بر نظریه دیگر فراهم نمی‌کنند، بلکه آن گونه که کوهن (Kuhn, 1981) استدلال می‌کند، با اینکه این ارزش‌ها تقریباً در تمام قضاوت‌های علمی دانشمندان دخیل هستند، ممکن است در یک حوزه، پیش‌بینی دقیق و در حوزه دیگر توانایی یکپارچه‌سازی در اولویت باشد. با این همه، آنچه طرفداران ایده‌آل آزاد از ارزش از آن دفاع می‌کنند، بیرون راندن ارزش‌های غیر معرفت‌شناختی از فرایند علمی است. اما آیا رسیدن به آرمان آزاد از ارزش ممکن و عقلانی است؟ مخالفان این دیدگاه انتقادهایی را به این رویکرد وارد کرده‌اند که در ادامه به برخی از آنها خواهیم پرداخت.

لونگینو (Longino, 1990) به تأثیر ارزش‌های شخصی و اجتماعی بر انتخاب زمینه‌های تحقیقاتی پرداخته است. بودجه تحقیقات علمی از نهادهای اجتماعی یا خصوصی تامین می‌شود و همواره اهداف و ارزش‌های این نهادها بر جهت‌گیری تحقیقات علمی اثرگذار است. با وجود اثرگذاری این ارزش‌ها در جهت‌گیری علمی، آیا ممکن است که همچنان به آرمان آزاد از ارزش پایبند ماند؟ در واقع، در همین نقطه است که ارتباط موضوع جداسازی زمینه کشف<sup>۱</sup> از زمینه توجیه<sup>۲</sup> در فرایند عینیت بخشی به علم رخ نشان می‌دهد (Popper, 1959; Reichenbach, 1938). طرفداران دیدگاه ایده‌آل آزاد از ارزش با جداسازی فرایند شکل‌گیری ایده‌ها از توجیه آنها معتقدند که فرایند داخلی و اساسی علم مربوط به زمینه توجیه است که از انتقاد وارده در امان است، چراکه فارغ از نحوه شکل‌گیری نظریه (فرضیه)، می‌توان (و باید) آن را تحلیل منطقی کرد. حتی اگر ارزش‌ها وارد فرایند توجیه شوند، نشان‌دهنده عمل علمی بد<sup>۳</sup> است که با توجه به قابلیت خود تصحیحی<sup>۴</sup> علم به مرور زمان خنثی می‌شود (Longino, 1990). لونگینو در ادامه بیان می‌کند که ممکن است برخی از ملاک‌های عقلانی برای ارزیابی نظریه‌ها، مستقل از ارزش‌ها وجود داشته باشند، اما برآورده کردن این ملاک‌ها لزوماً بدین معنا نیست که نظریه مد نظر آزاد از ارزش باشد. وی علت اصلی این پدیده را این‌گونه بیان کرده است: «حمایت شواهد از فرضیه‌ها یک موضوع نسبی است. در حالی که در زمینه یک مجموعه از باورها یا مفروضات، X شاهدهی برای h خواهد بود، در زمینه مجموعه دیگری [از باورها و مفروضات] X شاهدهی برای h نیست، بلکه برای برخی فرضیه‌های دیگر [مانند] h' است یا اصلاً برای هیچ فرضیه‌ای [شاهد] نیست» (Longino, 1990).

برای فهم بهتر موضوع، این مثال را در نظر بگیرید. فرض کنید یک دانشمند در حال بررسی تأثیر داروی X بر بهبودی یک بیماری خاص است. در سایه مجموعه‌ای از باورهای پیشین A (بر اساس پیشینه پژوهشی یا نظرهای متخصصان)، اگر دانشمند مشاهده کند بیمارانی که داروی X مصرف کرده‌اند در مقایسه با افرادی که دارو مصرف نکرده‌اند بهبودی در وضعیت خود نشان داده‌اند، ممکن است این موضوع را به‌عنوان شاهدهی از این فرضیه که داروی X بر درمان این بیماری مؤثر است، در نظر بگیرد. با این حال، در یک زمینه متفاوت یا با مجموعه‌ای از باورها و مفروضات متفاوت A'، مشاهده یکسان ممکن است لزوماً به‌عنوان شاهدهی در حمایت از این فرضیه که داروی X بر درمان بیمارانی اثرگذار است، در نظر گرفته نشود. برای مثال، اگر مطالعات قبلی با نتایج متناقضی وجود داشته باشند یا درباره قابلیت اطمینان داده‌های جمع‌آوری شده تردید وجود داشته باشد، مشاهده اثرهای بهبود در بیمارانی که داروی X را دریافت کرده‌اند، ممکن است به‌عنوان شاهدهی

<sup>1</sup> context of discovery

<sup>2</sup> context of justification

<sup>3</sup> bad scientific practice

<sup>4</sup> self-corrective



برای یک فرضیه متفاوت، مانند این فرضیه که اثرهای بالقوه دارونما در ایجاد نتایج اثرگذار بوده است، در نظر گرفته شود. بنابراین، اینکه چگونه مشاهدات و شواهد خود را تفسیر کنیم، وابسته به برخی پیشفرض‌هاست.

زمانی که این دیدگاه را در ذهن داشته باشیم، واضح است که گروه‌های مختلف می‌توانند به شکلی متفاوت شواهد مشابهی را تفسیر کنند، چرا که از پیشفرض‌های متفاوتی برخوردارند. با این همه، ممکن است دیدگاه تمام این گروه‌ها نیز عاقلانه در نظر گرفته شود (Kellert et al., 2006; Longino, 1990). حال این سؤال پیش می‌آید که اگر ارزش‌ها همواره در فعالیت علمی دخیل هستند، پس چگونه باید به عینیت دست یابیم؟ از نظر لونگینو (Longino, 1990) عینیت زمانی به دست می‌آید که تمام گروه‌ها بتوانند ارزش‌های خود را در فضایی دموکراتیک بیان کنند و این ارزش‌ها در دیگر گروه‌ها بحث و انتقاد شود.

چه میزان از شواهد برای تأیید یا رد یک فرضیه مورد نیاز است (John, 2021; Elliott, 2022; Douglas, 2009)؟ این سؤال که به برهان مخاطره‌استقرایی<sup>۱</sup> شهرت دارد، اساس دومین انتقاد مخالفان آرمان آزاد از ارزش را شکل می‌دهد. تصور کنید که یک پزشک این فرضیه را که یک شیوه جدید از جراحی بر بهبود بیماری Z اثرگذار است، آزمایش می‌کند. از سوی دیگر، یک شیمیدان این فرضیه را که ماده n باعث افزایش خاصیت چسبندگی چسب می‌شود، بررسی می‌کند. آیا میزان شواهدی که پزشک و شیمیدان برای مثبت ارزیابی کردن مداخله خود نیاز دارند یکسان است؟ احتمالاً شما نیز با ما هم نظر هستید که این گونه نیست. در واقع، پزشک به سبب عواقب بسیار خطرناک‌تر در صورت بروز خطای مثبت کاذب<sup>۲</sup>، شواهد بیشتری را برای رد فرض صفر خود نیاز دارد. نکته حایز اهمیت، نقش ارزش‌ها در تعیین این سطح از شواهد مورد نیاز برای رد یا تأیید یک فرضیه در هر حوزه علمی است. به بیان ساده‌تر، اینکه چه سطحی از سخت‌گیری برای تأیید هر فرضیه استفاده شود، به پیش‌بینی نتایج حاصل از رد یا تأیید آن فرضیه بستگی دارد که اساساً یک امر ارزشی است.

با در نظر گرفتن تمام مباحث مطرح شده، عینیت در علم چه جایگاهی دارد؟ برخی مانند هکینگ (Hacking, 2015) ادعا کرده‌اند که باید مفهوم عینیت را به‌طور کلی از علم کنار بگذاریم. از نظر وی آنچه به‌عنوان عینیت می‌شناسیم، صرفاً یک تز منفی است که ما را از برخی اعمال در انجام دادن فعالیت علمی باز می‌دارد. با این حال، این مفهوم هیچ‌گونه پیشنهادی در خصوص اینکه در یک فعالیت علمی خاص باید چگونه عمل کنیم، به ما ارائه نمی‌دهد. به بیان ساده، هکینگ معتقد است که به جای اینکه ادعا کنیم فعالیت انجام پذیرفته عینی است، بهتر است دقیقاً بیان کنیم چه ویژگی مطلوبی در خصوص آن وجود داشته است.

کارترایت و همکاران (Cartwright et al., 2022) معتقدند که عینیت را می‌توان به‌عنوان یک مفهوم با تعریف فراخ<sup>۳</sup> که می‌تواند بیش از یک تعریف داشته باشد، در نظر گرفت. از دلایل کارترایت در کنار نگذاشتن عینیت، وجود جنبه هنجاری<sup>۴</sup> در برخی مفاهیم علوم انسانی از جمله فقر، جنگ، دموکراسی و البته، عینیت است. وجود جنبه هنجاری یعنی اینکه این مفاهیم علاوه بر معنایی که حمل می‌کنند، هدایت‌کننده اعمال<sup>۵</sup> نیز هستند. به بیان دیگر، این تعاریف چیزهایی را توصیف می‌کنند که برای ما حایز اهمیت‌اند. از نظر کارترایت و همکاران (Cartwright et al., 2022)، از دانشمندان انتظار می‌رود تا علاوه بر رعایت عینیت، عینیت را در حوزه مطالعه خود تعریف کنند: «شما می‌خواهید که دانشمندان درگیر نه تنها عینی باشند، بلکه به این پی ببرند که عینی بودن در شرایط به چه معناست». بر این اساس، اهداف یک مؤلفه اساسی در تعریف عینیت است. برای درک بهتر این مبحث، اهداف به دو دسته اهداف ذاتی<sup>۶</sup> و بیرونی<sup>۷</sup> تقسیم می‌شود (Chang, 2022). برای

<sup>1</sup> inductive risk

<sup>2</sup> false positive

<sup>3</sup> loose

<sup>4</sup> normative

<sup>5</sup> action-guiding

<sup>6</sup> inherent

<sup>7</sup> external





یک ابزار موسیقی، هدف ذاتی ایجاد نوت‌های موسیقایی مربوط است. از سوی دیگر، اجرا در یک کنسرت، جلب توجه دوستان یا حتی ایجاد حس آرامش برای نوزاد برای خواباندن وی، همگی می‌توانند اهداف بیرونی مربوط به عمل استفاده از یک ابزار موسیقی باشند. با این تفسیر، هدف ذاتی امری مشخص و خنثی است، اما در مقابل، هدف‌های بیرونی می‌توانند کاملاً متفاوت و حتی متعارض باشند. در واقع، منظور کارترایت و همکاران (2022) از تعریف عینیت در هر موقعیت این است که محقق باید در هر زمینه اهداف بیرونی مناسب را شناسایی کند و بتواند طرح مناسب برای رسیدن به این اهداف را پیاده‌سازی کند (Cartwright et al., 2022). برای جمع‌بندی باید بیان داشت که مفهوم عینیت دیگر به معنای خالی کردن ارزش‌ها از فرایند علمی نیست، بلکه شناسایی ارزش‌ها و اهداف مطلوب در هر زمینه و تلاش برای رسیدن به آن را می‌توان پارادایم جدید عینیت دانست.

**عینیت نوین در پذیرش دانشجو:** برای پیاده‌سازی دیدگاه نوین عینیت، اساسی‌ترین گام مشخص کردن اهداف مطلوب است. اهداف اساسی در پذیرش دانشجو را می‌توان به دو بخش اهداف دانشگاه و اهداف اجتماع تقسیم کرد. اهداف دانشگاه عموماً در خطوط تحقیقاتی (و مأموریت‌ها) گروه‌های آموزشی قابل مشاهده هستند. خطوط تحقیقاتی به اهداف و کارکردهای نظری و کاربردی هر گروه آموزشی اشاره دارد. مشخص کردن خطوط تحقیقاتی کارکردهای متفاوتی دارد که از جمله آنها می‌توان به موارد زیر اشاره کرد:

- ۱) تبیین هدف: خطوط تحقیقاتی به فهم دلیل وجودی گروه آموزشی و مشارکت منحصر به فرد آن در سازمان کمک می‌کند.
- ۲) انسجام موضوعی: خطوط تحقیقاتی موجب ایجاد تمرکز بر یک موضوع، حوزه یا مسئله خاص می‌شوند و از این طریق انسجام را در مطالعات متعدد تضمین می‌کنند.
- ۳) جهت‌گیری بلندمدت: برخلاف پروژه‌های منفرد که دارای مدت زمان محدودی هستند، خطوط تحقیقاتی به‌طور مستمر توسعه پیدا می‌کنند.
- ۴) دانش تجمیعی: به‌کارگیری خطوط تحقیقاتی و انجام دادن مطالعات مطابق با آنها سبب شکل‌گیری تخصص و افزایش عمق فهم در طول چندین مطالعه می‌شود.
- ۵) چارچوب هدایت‌کننده تحقیق: خطوط تحقیقاتی به محققان در طراحی پروژه‌های تازه، جذب بودجه و همکاری کارآمد در اجرای پژوهش‌ها (یا پروژه‌ها) کمک می‌کنند.
- ۶) ایجاد انگیزه و تعامل در کارکنان: خطوط تحقیقاتی به اعضای تیم کمک می‌کند تا اهمیت کار خود را به وضوح درک کنند که این امر موجب تقویت انگیزه و همبستگی میان کارکنان می‌شود.

بنابراین، اساسی‌ترین گام در دستیابی به عینیت نوین (با توجه به اهداف دانشگاه)، ایجاد و توسعه خطوط تحقیقاتی مناسب در هر گروه آموزشی است. اما در خصوص اهداف اجتماعی به نظر نیاز به تعریف اهداف نیست، چراکه اهداف اجتماعی اساسی مانند عدالت<sup>۱</sup> در دسترسی به آموزش عالی (Ebrahimi et al., 2025) و شایسته‌سالاری<sup>۲</sup> (Amini Bagh & Salimi, 2025; Hosseini & Salimi, 2025) از قبل مشخص هستند. در گام دوم تحقق بخشی به عینیت نوین، تلاش برای رسیدن به اهداف قرار دارد. در خصوص اهداف دانشگاهی این تلاش مربوط به جذب داوطلبانی می‌شود که به بهترین شکل توانایی (یا ظرفیت) سازگاری با خطوط تحقیقاتی یا کمک به توسعه آنها را داشته باشند. در خصوص اهداف اجتماعی، تلاش برای رسیدن به اهدافی مانند شایسته‌سالاری و عدالت در اولویت قرار دارند، اما این اهداف گاهی ممکن است در تضاد با یکدیگر قرار گیرند. اگرچه تلاش برای جذب مناسب‌ترین داوطلبان (با توجه به خطوط تحقیقاتی) با هدف شایسته‌سالاری همگراست،

<sup>1</sup> fairness

<sup>2</sup> meritocracy



با این حال، این دو هدف گاهی ممکن است با هدف عدالت در دسترسی به آموزش عالی متعارض باشند. بنابراین، ایجاد تعادل و راهکاری بهینه در حل این تعارضات برای رسیدن به عینیت ضروری است.

### ترکیب دیدگاه‌ها

درباره دیدگاه‌های متفاوت عینیت و تأثیر این دیدگاه‌ها بر پذیرش دانشجویان بحث شد. با این حال، صرف بیان جزیره‌ای هر یک این دیدگاه‌ها نمی‌تواند به نتیجه قابل قبولی در خصوص چگونگی پیاده‌سازی و هماهنگ‌سازی آنها منجر شود. بنابراین، در این بخش به دنبال بهینه و هماهنگ‌سازی نتایج به دست آمده در بخش‌های قبلی در قالب یک سیستم جامع پذیرش هستیم. برای رسیدن به این هدف، ابتدا سیستم پیشنهادی معرفی و در خصوص مزایا و معایب احتمالی آن بحث شده است.

نخستین مؤلفه این سیستم، آزمون‌های استاندارد است که در ارزیابی توانایی‌های شناختی داوطلبان نقش کلیدی دارند. این آزمون‌ها از چندین منظر با آزمون‌های پیشین تفاوت دارند. نخست اینکه این آزمون‌ها به‌عنوان تنها ملاک پذیرش در نظر گرفته نمی‌شوند؛ در این سیستم آزمون‌ها به‌عنوان یکی از چندین ابزار ارزیابی به کار گرفته می‌شوند. دومین تفاوت مربوط به نحوه استفاده از نتایج آزمون‌هاست. برخلاف آنچه در حال حاضر انجام می‌پذیرد که نتایج آزمون به شکل هنجار مرجع استفاده می‌شود، نتایج آزمون‌ها به صورت ملاک مرجع در نظر گرفته خواهد شد؛ یعنی هر رشته محل (یک رشته در یک دانشگاه خاص) برای ورود به چرخه پذیرش، حداقلی را در آزمون یا حتی آزمون‌های مد نظر، پیش‌نیاز در نظر می‌گیرد. علاوه بر این، در تحلیل و ساخت این آزمون‌ها از نظریه صفت‌مکون استفاده خواهد شد که نسبت به نظریه کلاسیک آزمون که در حال حاضر استفاده می‌شود، امکان ساخت آزمون‌هایی با ویژگی‌های فنی بهتری را فراهم می‌کند (بخش ذات‌گرایی در مدل‌های متغیر مکون). این تفاوت‌ها با سیستم قبلی در کنار یکدیگر چندین مزیت را، نسبت به آنچه در حال انجام است، دارند. ابتدا، نتایج آزمون‌ها در این سیستم شرط کافی برای پذیرش دانشجویان نیستند و همین موضوع کار را برای رواسازی (و دفاع از روایی) آزمون آسان‌تر می‌کند. این موضوع بر ساخت آزمون‌ها نیز اثرگذار است. مدت‌هاست که انطباق سؤالات آزمون سراسری با محتوای کتاب‌های درسی به یک الزام قانونی تبدیل شده است (Majles, 2021). با وجود این، هر ساله برخی سؤالات از خارج از محتوای کتاب‌های درسی طرح می‌شوند که علت اصلی آن نبود تمایزگذاری سؤالات مبتنی بر کتاب‌های درسی میان رتبه‌های برتر است. اما در این سیستم جدید، اساساً نیازی به تمایزگذاری دقیق بین این افراد نیست، چراکه آزمون تنها ملاک پذیرش دانشجویان نیست و بیشتر شبیه به یک غربال اولیه عمل می‌کند. از سوی دیگر، این امکان وجود دارد تا به جای تنها یک آزمون یا یک سلسله از آزمون‌های از پیش تعیین شده که همه باید در آنها شرکت کنند (مانند آزمون‌های نهایی و کنکور)، آزمون‌هایی را به صورت سلسله‌مراتبی داشته باشیم که در هر مرحله نسبت به مرحله قبل با توجه به نیازهای هر رشته تحصیلی (رشته‌های دانشگاهی و نه رشته‌های دوره متوسطه) تخصصی‌تر می‌شوند. علاوه بر این، این تغییرات، هم‌ترازسازی<sup>۱</sup> نمرات آزمون را که به چالش اساسی برای سازمان سنجش آموزش کشور تبدیل شده است، ساده‌تر می‌کنند. روش‌های هم‌ترازسازی مبتنی بر نظریه صفت‌مکون به دلیل ویژگی تغییرناپذیری پارامترها، نسبت به نظریه کلاسیک آزمون مزیت بیشتری دارند (Fitrah et al., 2023). همچنین با کم شدن حساسیت و اهمیت این آزمون‌ها، این امکان فراهم می‌شود تا برخی از سؤالات در آزمون‌ها تکرار شوند که یک نیاز اساسی برای برابری است (Kolen & Brennan, 2014).

مؤلفه دوم این سیستم پذیرش شامل بخش‌های متخصص‌محور می‌شود. منظور از بخش‌های متخصص‌محور، هرگونه فعالیتی در جریان پذیرش است که تفسیر آن به اعمال نظر متخصصان بستگی دارد. این بخش‌ها می‌توانند شامل شواهد مختلفی از جمله مصاحبه، توصیه‌نامه،

<sup>1</sup> equating





آزمون عملی شخصی سازی شده، رزومه و غیره باشد که در تفسیر آن نیاز به اعمال نظر تیم پذیرش دانشجویست. با این حال، در این مؤلفه از سیستم پذیرش یک عامل محدودکننده وجود دارد. تنها شواهدی که به نوعی با خطوط تحقیقاتی رشته محل مد نظر مرتبط هستند، به‌عنوان شواهد قابل قبول در فرایند پذیرش در نظر گرفته می‌شوند. این عامل محدودکننده به نوعی یک ملاک مشخص را برای بررسی شواهد هر داوطلب فراهم و تا حدی از اعمال نظرهای خارج از عَرَف جلوگیری می‌کند.

علاوه بر این، مشخص بودن خطوط تحقیقاتی هر رشته محل باعث می‌شود تا داوطلبان قبل از ورود به فرایند پذیرش دانشجوی بتوانند مناسب‌ترین گزینه را در اولویت خود قرار دهند و از آن مهم‌تر، این امکان برای داوطلبان فراهم می‌شود تا بتوانند تلاش‌های خود را معطوف به این خطوط تحقیقاتی کنند که این امر در سیستم فعلی ممکن نیست. در واقع، در سیستم فعلی (سیستم مبتنی بر کنکور)، امکان مطالعه تخصصی برای رشته محل مورد علاقه شخص وجود ندارد، چراکه این فعالیت‌ها و مطالعات تخصصی در پذیرش دانشجوی هیچ تأثیری ندارند. شاید حتی بتوان بیان داشت که سیستم پذیرش فعلی دانش‌آموزان را از انتخاب یک رشته به‌عنوان رشته مورد علاقه باز می‌دارد، چراکه پیش از کنکور مشخص نیست که بعد از اعلام نتایج چه رشته‌ای در انتظار فرد است و عملاً انتخاب رشته به فرایندی پس از آزمون سراسری تبدیل شده است. محدودیت دیگری که در این مؤلفه از سیستم پذیرش پیشنهادی مد نظر است، سهمیه‌هایی است که برای ایجاد عدالت در دسترسی به آموزش عالی وجود دارند. عدالت در دسترسی به آموزش عالی یک هدف اجتماعی است که نباید از آن چشم‌پوشی کرد و تیم پذیرش دانشجوی نیز باید در انتخاب‌های خود این مهم را مد نظر قرار دهد. با این حال، معمولاً بحث سهمیه‌ها در کشور ایران، یک بحث مناقشه‌انگیز بوده است. یکی از دلایل این مناقشات به ماهیت سهمیه‌ها مربوط می‌شود که موضوع بحث این مقاله نیست. اما یکی دیگر از دلایل نزاع، سطح علمی پایین داوطلبانی است که با سهمیه وارد رشته محل‌های پر طرفدار می‌شوند. این معضل در سیستم پیشنهادی تا حدی کمتر شده است، چراکه برای ورود به مؤلفه دوم پذیرش (برای مثال شرکت در مصاحبه)، نیاز به گذر کردن از مؤلفه اول است که در آن داوطلب باید حداقل نمره‌ای را کسب کند و به همین دلیل نیز یک غربال اولیه در خصوص توانایی‌های آموزشی این داوطلبان انجام شده است.

بر اساس مباحث پیشین، می‌توان یک مدل مفهومی تلفیقی برای نظام پذیرش دانشجوی در ایران طراحی کرد که به‌صورت نظام‌مند عناصر اصلی سه رویکرد معرفت‌شناختی (عینیت مکانیکی، حقیقت طبیعت و قضاوت تمرین‌شده) را در چارچوب عینیت نوین گرد هم می‌آورد. این مدل از سه لایه اساسی زیر تشکیل می‌شود:

**۱. لایه سنجش استاندارد (غربالگری اولیه):** در این مرحله آزمون‌های مبتنی بر نظریه متغیر مکنون، به‌ویژه مدل‌های سؤال پاسخ، به‌عنوان ابزار اولیه برای سنجش توانایی‌های شناختی و تحصیلی داوطلبان به‌کار گرفته می‌شوند. تمایز این لایه با نظام موجود آن است که آزمون‌ها دیگر شرط کافی برای پذیرش محسوب نمی‌شوند، بلکه کارکرد اصلی آن‌ها غربالگری اولیه و تأمین حداقل‌های لازم برای ورود به فرایند پذیرش است. افزون بر این، اتکای آزمون‌ها به چارچوب‌های روانسنجی پیشرفته، امکان ارتقای ویژگی‌های فنی و دفاع‌پذیری بیشتر از روایی آنها را فراهم می‌سازد.

**۲. لایه ارزیابی مبتنی بر قضاوت تمرین‌شده:** داوطلبانی که از حداقل‌های تعیین‌شده عبور می‌کنند، وارد مرحله دوم می‌شوند که در آن قضاوت متخصصان آموزش‌دیده نقش محوری دارد. این مرحله می‌تواند شامل مصاحبه‌های ساختاریافته، بررسی توصیه‌نامه‌ها، تحلیل سوابق تحصیلی و فعالیت‌های فوق‌برنامه باشد. ویژگی اصلی این لایه آن است که بر اساس بینش حرفه‌ای و تجربه انباشته متخصصان، ابعاد غیرقابل سنجش از طریق آزمون‌های استاندارد (مانند انگیزه، خلاقیت یا ظرفیت رهبری) ارزیابی می‌شود. بدین ترتیب، این لایه کارکرد مکمل آزمون‌ها را ایفا می‌کند و از حذف داوطلبانی که استعدادها بالقوه آنان در قالب نمرات عددی منعکس نمی‌شود، جلوگیری می‌کند.



۳. **لایه اهداف دانشگاهی و اجتماعی (چارچوب هنجاری):** نتایج حاصل از دو لایه پیشین در پرتو اهداف کلان دانشگاهی و اجتماعی تفسیر و نهایی می‌شوند. اهداف دانشگاهی در قالب خطوط تحقیقاتی، مأموریت‌های آموزشی و نیازهای علمی هر گروه آموزشی بازنمایی می‌شوند. در کنار آن، اهداف اجتماعی همچون عدالت در دسترسی به آموزش عالی، تنوع و شایسته‌سالاری معیارهای نهایی برای اعتباربخشی به کل فرایند هستند. این لایه در واقع، چارچوب هنجاری‌ای را فراهم می‌آورد که ضمن جهت‌دهی به کل نظام پذیرش، توازن میان کارآمدی علمی و عدالت اجتماعی را تضمین می‌کند.

### نتیجه‌گیری

در این مطالعه مفهوم عینیت مکانیکی و تأثیر آن بر نظام پذیرش دانشجو در کشور ایران بررسی شد. نشان داده شد که تکیه بر آزمون‌های استاندارد، نظریه کلاسیک آزمون و به‌کارگیری رویه‌های ثابت، ریشه در درک مکانیکی از عینیت دارد که پایایی معرفت‌شناختی و بی‌طرفی اخلاقی را به‌عنوان اصول بنیادی خود در نظر می‌گیرد. با این حال، این نگرش چالش‌هایی را در زمینه روایی آزمون‌ها، انعطاف‌پذیری در سنجش قابلیت‌های متقاضیان و ناتوانی در تشخیص ویژگی‌های ضروری برای موفقیت در رشته‌های مختلف ایجاد کرده است. برای غلبه بر این محدودیت‌ها، دو شیوه جابگزین قضاوت علمی؛ یعنی حقیقت‌طبیعت و قضاوت تمرین‌شده، معرفی و تحلیل شدند. دیدگاه حقیقت‌طبیعت که ریشه فکری مدل‌های متغیر مکنون هستند، امکان تحلیل عمیق‌تر روابط بین ویژگی‌های آزمون و توانایی‌های زیربنایی داوطلبان و در نتیجه، دقت بیشتری را در تفسیر نتایج آزمون‌ها فراهم می‌کند. از سوی دیگر، رویکرد قضاوت تمرین‌شده بر نقش متخصصان در فرایند ارزیابی تأکید دارد و این امکان را فراهم می‌کند تا داوطلبان از زوایای گوناگون و با در نظر گرفتن عوامل زمینه‌ای ارزیابی شوند.

علاوه بر این، بررسی تحولات اخیر در فلسفه علم نشان داد که مفهوم عینیت از یک ایده‌آل آزاد از ارزش به سوی عینیت وابسته به موقعیت تغییر یافته است. این تحول بر لزوم تعریف اهداف مشخص برای فرایند پذیرش دانشجو تأکید دارد که باید در تعامل با ذینفعان مختلف شکل بگیرد. بر این اساس، پذیرش دانشجو باید به‌گونه‌ای طراحی شود که نه تنها اهداف دانشگاه‌ها را در جهت خطوط تحقیقاتی‌شان تأمین کند، بلکه با اهداف اجتماعی نظیر عدالت آموزشی و شایسته‌سالاری نیز همخوان باشد. در نهایت، یک چارچوب ترکیبی برای پذیرش دانشجو پیشنهاد شد که شامل دو مؤلفه اصلی است: آزمون‌های استاندارد مبتنی بر نظریه متغیر مکنون که به‌عنوان یک غربال اولیه عمل می‌کنند و فرایندهای متخصص‌محور مانند مصاحبه و بررسی توصیه‌نامه‌ها که نقش تکمیلی را در سنجش جامع‌تر داوطلبان دارند. در این چارچوب پیشنهادی ضمن حفظ ملاک‌های کمی و قابل سنجش، امکان ارزیابی چندبعدی‌تر و عادلانه‌تر متقاضیان فراهم می‌شود و می‌تواند به بهبود فرایند پذیرش دانشجو در ایران منجر شود.

باید توجه داشت که شیوه پذیرش پیشنهادی ما به‌طور کامل نوآورانه نیست، اما اجرای آن در کشور ایران با چالش‌ها و مخالفت‌های جدی مواجه بوده است. هدف اصلی از نگارش این مقاله، پرداختن به یکی از مهم‌ترین دلایل این مخالفت‌ها؛ یعنی ادعای نبود عینیت در این نوع پذیرش بوده است. از این رو، در این مطالعه با بازاندیشی در مفهوم عینیت و نقد تلقی رایج از آن، تلاش شد تا این استدلال به چالش کشیده و از اعتبار رویکرد ترکیبی پیشنهادی دفاع شود. افزون بر این، تمرکز اصلی در این مطالعه بر فرم نظام پذیرش بوده است و نه بر محتوای ابزارها و معیارهای مورد استفاده در آن؛ به‌عبارت دیگر، پژوهش نشان داد که چگونه استفاده از آزمون‌های استاندارد و استانداردسازی فرایند پذیرش با مفهوم عینیت مکانیکی پیوند خورده است و این فرم حاکم چگونه می‌تواند از منظر دیدگاه‌های بدیل نقد و بازنگری شود. باید تأکید کرد که در این نوشتار نسبت به محتوای آزمون‌ها یا سایر ملاک‌های جایگزین پذیرش موضع‌گیری خاصی وجود نداشته است، چه این محتوا شامل نمرات مدرسه و چه آزمون‌های پایانی مانند کنکور باشد، مادامی که در راستای اهداف دانشگاهی و اجتماعی تعریف شوند. از این‌رو، نقدهای





ناظر بر محتوای ابزارهای سنجش (برای مثال، نپرداختن به فعالیت‌های مربوط به مدرسه)، هرچند قابل تأمل و ارزشمندند، خارج از حوزه تمرکز این مطالعه قرار دارند.

## تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافی وجود ندارد.

## References

- Amini Bagh, A., & Salimi, J. (2025). A qualitative analysis of faculty members' perceptions of the creative university: A pathway toward educational innovation. *Research and Planning in Higher Education Quarterly*, 31(2), 45-66. <https://doi.org/10.61838/kman.Irphe.31.2.3>
- Au, W. (2013). Hiding behind High-Stakes Testing: Meritocracy, Objectivity and Inequality in U.S. Education. *The International Education Journal: Comparative Perspectives*, 12, 7-20. <https://openjournals.library.sydney.edu.au/IEJ/article/view/7453>
- Baker, F. B., & Kim, S. H. (2018). *The Basics of Item Response Theory Using R*. Springer Publishing Company, Incorporated. <https://link.springer.com/content/pdf/10.1007/978-3-319-54205-8.pdf>
- Borgstede, M., & Eggert, F. (2023). Squaring the circle: From latent variables to theory-based measurement. *Theory & Psychology*, 33(1), 118-137. <https://doi.org/10.1177/09593543221127985>
- Borsboom, D. (2005). *Measuring the mind : conceptual issues in contemporary psychometrics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511490026>
- Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, 64(9), 1089-1108. <https://doi.org/10.1002/jclp.20503>
- Borsboom, D., & Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91-121. <https://doi.org/10.1146/annurev-clinpsy-050212-185608>
- Bridgman, P. W. (1927). *The logic of modern physics*. Macmillan. <https://marcellocosta.au/wp/wp-content/uploads/2024/04/Bridgmans-The-Logic-of-Modern-Physics.pdf>
- Bulut, O. (2016). Applying Item Response Theory Models to Entrance Examination for Graduate Studies: Practical Issues and Insights. *Journal of Measurement and Evaluation in Education and Psychology*, 6(2), 0-0. <https://doi.org/10.21031/epod.17523>
- Cartwright, N., Hardie, J., Montuschi, E., Soleiman, M., & Thresher, A. C. (2022). *The Tangle of Science: Reliability Beyond Method, Rigour, and Objectivity*. Oxford University Press. <https://doi.org/10.1093/oso/9780198866343.001.0001>
- Chang, H. (2022). *Realism for Realistic People: A New Pragmatist Philosophy of Science*. Cambridge University Press. <https://doi.org/10.1017/9781108635738>
- Cronbach, L. J. (1971). Test validity. In *Educational Measurement*. American Council on Education. <https://www.scirp.org/reference/referencespapers?referenceid=1457760>
- Daston, L., & Galison, P. (2007). *Objectivity*. Zone Books.
- Douglas, H. (2000). Inductive Risk and Values in Science. *Philosophy of Science*, 67(4), 559-579. <https://doi.org/10.1086/392855>
- Douglas, H. (2009). *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press. <https://doi.org/10.2307/j.ctt6wrc78>
- Ebrahimi, Y., Chavoushi, Z., & Roshan, A. (2025). An investigation of the socio-economic status of students in public universities of Tehran. *Journal of Research and Planning in Higher Education*, 31(2), 67-82. <https://doi.org/10.61838/kman.Irphe.31.2.4>
- Elliott, K. C. (2022). *Values in Science*. Cambridge University Press. <https://doi.org/10.1017/9781009052597>
- Fitrah, M., Ilyas, Akbarini, N., Istiyono, E., & Widhiastuti, W. (2023). Equating Principles in Assessment: A Literature Review in the Context of Education and Assessment. *IJECA (International Journal of Education and Curriculum Application)*, 6, 199-209. <https://doi.org/10.31764/ijeca.v6i3.19381>
- Gillispie, C. C. (1960). *The Edge of Objectivity: An Essay in the History of Scientific Ideas*. Princeton University Press. <https://doi.org/10.2307/j.ctvc779t6>
- Griffin, P. (2018). *Assessment for teaching*. Cambridge University Press. [https://books.google.com/books?hl=en&lr=&id=4i42DwAAQBAJ&oi=fnd&pg=PR7&dq=Griffin,+P.++\(2018\).+Assessment+for+teaching.+United+Kingdom,+Cambridge+University+Press.+%09&ots=XU3PeIS7Vu&sig=pl79J2K5xjEQ5rupCnLyPtXBS\\_o](https://books.google.com/books?hl=en&lr=&id=4i42DwAAQBAJ&oi=fnd&pg=PR7&dq=Griffin,+P.++(2018).+Assessment+for+teaching.+United+Kingdom,+Cambridge+University+Press.+%09&ots=XU3PeIS7Vu&sig=pl79J2K5xjEQ5rupCnLyPtXBS_o)
- Hacking, I. (2015). Let's Not Talk About Objectivity. In *Objectivity in Science: New Perspectives From Science and Technology Studies* (pp. 19-33). Springer. [https://doi.org/10.1007/978-3-319-14349-1\\_2](https://doi.org/10.1007/978-3-319-14349-1_2)
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage Publications, Inc. [https://books.google.com/books?hl=en&lr=&id=gW05DQAAQBAJ&oi=fnd&pg=PP1&dq=Hambleton,+R.+K.,+et+al.++\(1991\).+Fundamentals+of+item+response+theory,+Sage+Publications,+Inc.+%09&ots=h\\_K\\_KC0UAP&sig=25WhX6lrqnY9RkS8MRFI3rQ68zw](https://books.google.com/books?hl=en&lr=&id=gW05DQAAQBAJ&oi=fnd&pg=PP1&dq=Hambleton,+R.+K.,+et+al.++(1991).+Fundamentals+of+item+response+theory,+Sage+Publications,+Inc.+%09&ots=h_K_KC0UAP&sig=25WhX6lrqnY9RkS8MRFI3rQ68zw)
- Heinen, T. (1996). *Latent class and discrete latent trait models: similarities and differences*. Sage Publications. <https://psycnet.apa.org/record/1996-09133-000>
- Hosseini, N., & Salimi, G. (2025). Identification and analysis of the competencies and responsibilities of PhD students at Shiraz University regarding their supervisor's research guidance. *Research and Planning in Higher Education Quarterly*, 31(2), 21-44. <https://doi.org/10.61838/kman.Irphe.31.2.2>
- John, S. (2021). *Objectivity in Science*. Cambridge University Press. <https://doi.org/10.1017/9781009063647>
- Joreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36(2), 109-133. <https://doi.org/10.1007/BF02291393>



- Kane, M. T. (2006). Validation. In R. L. Brennan & E. Fourth (Eds.), *Educational Measurement* (pp. 17-64). <https://www.scirp.org/reference/referencespapers?referenceid=2372989>
- Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, 50(1), 1-73. <https://doi.org/10.1111/jedm.12000>
- Kane, M. T. (2016). Validity as the evaluation of the claims based on test scores. *Assessment in Education: Principles, Policy & Practice*, 23(2), 309-311. <https://doi.org/10.1080/0969594X.2016.1156645>
- Karvandi Renani, M., Khodaie, E., Salehi, K., & Moghaddam Zadeh, A. (2024). Assessment Limitations of University Entrance Exam in Intra-individual Abilities. *Iranian Higher Education*, 16(1), 110-130. <http://ihej.ir/article-1-2049-fa.html>
- Kellert, S. H., Longino, H., & Waters, C. K. (2006). Introduction: The Pluralist Stance. In *Scientific Pluralism* (pp. vii-xxix). University of Minnesota Press. [https://books.google.com/books?hl=en&lr=&id=B5UHqo7V1AsC&oi=fnd&pg=PR7&dq=Kellert,+S.+H.,+et+al.+\(2006\).+Introduction:+The+Pluralist+Stance.+Scientific+Pluralism,+University+of+Minnesota+Press:+vii-xxix.+%09&ots=9xyp5LrayM&sig=\\_r55GYb0MvhVwZjZ1AqQrgdKS8A](https://books.google.com/books?hl=en&lr=&id=B5UHqo7V1AsC&oi=fnd&pg=PR7&dq=Kellert,+S.+H.,+et+al.+(2006).+Introduction:+The+Pluralist+Stance.+Scientific+Pluralism,+University+of+Minnesota+Press:+vii-xxix.+%09&ots=9xyp5LrayM&sig=_r55GYb0MvhVwZjZ1AqQrgdKS8A)
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices*. Springer Science + Business Media. <https://doi.org/10.1007/978-1-4939-0317-7>
- Koskinen, I. (2021). Objectivity in contexts: withholding epistemic judgement as a strategy for mitigating collective bias. *Synthese*, 199(1), 211-225. <https://doi.org/10.1007/s11229-020-02645-9>
- Kuhn, T. S. (1981). Objectivity, value judgment, and theory choice. In *The Essential Tension: Selected Studies in Scientific Tradition and Change* (pp. 320-339). Duke University Press. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780203718087-6/objectivity-value-judgment-theory-choice-thomas-kuhn>
- Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent Structure Analysis*. Houghton Mifflin.
- Lloyd, E. A. (1995). Objectivity and the double standard for feminist epistemologies. *Synthese*, 104(3), 351-381. <https://doi.org/10.1007/BF01064505>
- Longino, H. E. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry* (Vol. 25). Princeton University Press. <https://doi.org/10.1515/9780691209753>
- Lord, F. M. (1980). *Applications of Item Response Theory To Practical Testing Problems*. Routledge. <https://doi.org/10.4324/9780203056615>
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Adison-Wesley.
- Majles, I. R. o. I. P. (2021). Expert opinion on the 'Bill mandating the National Organization of Educational Testing to design university entrance exam questions solely based on official school textbooks'. *Research Center of the Islamic Consultative Assembly, Expert Reports*, 129(29), 853-855. <https://www.noormags.ir/view/fa/articlepage/1786749>
- McMullin, E. (1982). Values in Science. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1982(4), 3-28. <https://doi.org/10.1086/psaprocbienmeetp.1982.2.192409>
- Messick, S. (1989). Validity. In R. Linn & e. rd (Eds.), *Educational Measurement* (pp. 13-100). American Council on Education. <https://www.scirp.org/reference/ReferencesPapers?ReferenceID=590541>
- Michell, J. (1990). *An introduction to the logic of psychological measurement*. Lawrence Erlbaum Associates, Inc. <https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.4324/9781315807614&type=googlepdf>
- Michell, J. (2020). The fashionable scientific fraud: Collingwood's critique of psychometrics. *History of the Human Sciences*, 33(2), 3-21. <https://doi.org/10.1177/0952695119872638>
- Misak, C. J. (1995). *Verificationism: Its History and Prospects*. Routledge. <https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.4324/9780203980248&type=googlepdf>
- Niaz, M. (2017). *Evolving Nature of Objectivity in the History of Science and its Implications for Science Education*. <https://doi.org/10.1007/978-3-319-67726-2>
- Ohiri, S. C., Christopher, M., & Benedict, I. (2024). Differential Item Functioning Detection Methods: An Overview. *International Journal of Research Publication and Reviews*, 5, 1555-1564. <https://doi.org/10.55248/gengpi.5.0224.0505>
- Popper, K. R. (1959). *The logic of scientific discovery*. Basic Books. <https://doi.org/10.1063/1.3060577>
- Rasch, G. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests*. Nielsen & Lydiche. <https://psycnet.apa.org/record/1962-07791-000>
- Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. University of Chicago Press. <https://doi.org/10.1037/11656-000>
- Spearman, C. (1904). 'General intelligence,' objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-293. <https://doi.org/10.2307/1412107>
- Vessonen, E. (2021). Respectful operationalism. *Theory & Psychology*, 31(1), 84-105. <https://doi.org/10.1177/0959354320945036>
- Woo, S. E., LeBreton, J. M., Keith, M. G., & Tay, L. (2023). Bias, Fairness, and Validity in Graduate-School Admissions: A Psychometric Perspective. *Perspectives on Psychological Science*, 18(1), 3-31. <https://doi.org/10.1177/17456916211055374>
- Zagzebski, L. T. (1996). *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139174763>

