ISSN: 2008-8000

Assessment Limitations of University Entrance Exam in Intra-individual Abilities

Mahdi Karvandi Renani — Ph.D. Student in Measurement and Assessment, Faculty of Psychology and Education, University of Tehran, Tehran, Iran. Email: mahdikarvandi@ut.ac.ir

Ebrahim Khodaie ❋ — *Corresponding Author:* Professor, Faculty of Psychology and Education, University of Tehran, Tehran, Iran. Email: khodaie@ut.ac.ir

Keyvan Salehi — Associate Professor, Faculty of Psychology and Education, University of Tehran, Tehran, Iran. Email: keyvansalehi@ut.ac.ir

Ali Moghaddam Zadeh — Associate Professor, Faculty of Psychology and Education, University of Tehran, Tehran, Iran. Email: amoghadamzadeh@ut.ac.ir

Abstract

Objective: The national university entrance exam (Konkur) has been the most important exam and assessment conducted in Iran in its 60-year history (from 1968 until today). Therefore, it is crucial to understand what the test actually evaluates. . In this article, our goal is to examine the ergodicity assumption by examining the condition of homogeneity of the phenomenon in question in society. This is important because if the ergodicity assumption does not exist in the factor structure of the Kankor, these factors cannot be interpreted as intra-individual causes.

Method: The present research was designed as a cross-sectional study to investigate the hypothesis of ergodicity in the national exam of mathematics and physics. The statistical population included all subjects who participated in the mathematics and physics entrance exam in the year 2021, from which 12,699 people were randomly selected as a sample. The variables used in this study were the grades (in percentage) of the candidates in mathematics and physics courses, including Persian literature, Arabic, Islamic studies, English language, mathematics, physics, and chemistry. In the process of data analysis, exploratory factor analysis was initially used to identify the underlying factors of the entrance exam and indicators related to each factor. Subsequently, the fit of the obtained model was examined using structural equation modeling. Following this, hierarchical cluster analysis was employed to create five clusters of candidates participating in the entrance exam based on the percentages of different courses. Finally, the invariance of the factor structure obtained in these clusters was tested.

Conclusion: The absence of measurement invariance indicates the heterogeneity of the phenomenon among the subjects, which means that the Ergodicity assumption is not met. Therefore, the intra-individual interpretation of factors is not possible. This challenges the traditional interpretation of factors obtained from latent variable models as individual causes. It suggests that the obtained factors may be emergent properties of the population, but this claim faces challenges related to empirical justification and alignment with the purpose of using the entrance exam. In summary, discussing the technical aspects of the entrance exam without considering its theoretical and conceptual dimensions is analogous to "milking a sheep whose gender has not yet been determined": it may yield results, but the type of the product cannot be guaranteed.

Keywords: Latent Variable Models, Causal Interpretation, Ergodicity, Measurement Invariance, Konkur, Intra-individual abilities

[Downloaded from ihej.ir on 2025-08-01]

محدودیت آزمون سراسری در ارزیابی توانایی‌های درون‌فردی

مهدی کروندی رنانی
دانشجوی دکتری سنجش و اندازه‌گیری، گروه علوم تربیتی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران.
رایانامه: mahdikarvandi@ut.ac.ir

ابراهیم خدایی ❋
نویسنده مسئول:
استاد، گروه علوم تربیتی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران.
رایانامه: khodaie@ut.ac.ir

کیوان صالحی
دانشیار، گروه علوم تربیتی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران.
رایانامه: keyvansalehi@ut.ac.ir

علی مقدم‌زاده
دانشیار، گروه علوم تربیتی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران.
رایانامه: amoghadamzadeh@ut.ac.ir

چکیده

هدف: آزمون سراسری ورود به دانشگاه (کنکور) مهم‌ترین آزمون و ارزیابی صورت گرفته در تاریخ ٦٠ ساله خود (از سال ١٣٤٧ هجری شمسی تا به امروز) در ایران بوده است. به همین دلیل، درک اینکه این آزمون چه چیزی را مورد سنجش قرار می‌دهد ضروری است. در این مقاله، هدف ما بررسی مفروضه ارگودیستی از طریق بررسی شرط همگنی پدیده مورد نظر در جامعه است. این امر بدان جهت از اهمیت برخوردار است که اگر مفروضه ارگودیسیتی در ساختار عاملی کنکور وجود نداشته باشد، نمی‌توان این عوامل را به صورت علّت‌های درون‌فردی تفسیر کرد.

روش: پژوهش حاضر از نوع تحقیقات بین‌مقطعی و به منظور بررسی تغییرناپذیری اندازه‌گیری در ساختار عاملی کنکور است. جامعه پژوهش کلیه داوطلبین آزمون سراسری در کنکور ریاضی سال ١٤٠٠ است که از این بین ١٢٦٩٩ نفر با روش نمونه‌گیری تصادفی ساده به‌عنوان نمونه انتخاب شدند. شاخص‌های (متغیرهای) بکار رفته در این مقاله را نمرات (بر حسب درصد) داوطلبین دروس آزمون سراسری گروه علوم ریاضی و فنی شامل ادبیات فارسی، عربی، معارف اسلامی، زبان انگلیسی، ریاضیات، فیزیک و شیمی تشکیل داده بود. یافته‌ها: نتایج تحلیل عاملی یک ساختار سه عاملی برای کنکور پیشنهاد داد. همچنین، با استفاده از تحلیل خوشه سلسله مراتبی، داوطلبین آزمون ریاضی و فنی به ٦ گروه تقسیم شدند. اما میان این ٦ خوشه مختلف آزمودنی‌ها در ساختار عاملی بدست آمده در مرحله اول، تغییرناپذیری اندازه‌گیری در هیچ سطحی (حتی سطح پیکربندی) برقرار نبود.

نتیجه‌گیری: با توجه به اینکه تغییرناپذیری اندازه‌گیری ساختار عاملی ما تایید نشد، بنابراین ساختار عاملی در کنکور سراسری از ویژگی همگنی جامعه برخوردار نیست. از آنجا که مفروضه ارگودیستی وابسته به شرط همگنی جامعه است، بنابراین، کنکور از مفروضه ارگودیستی برخوردار نیست. این بدان معنا است که نمی‌توان عامل‌های حاصله از مدل‌های متغیر مکنون در خصوص کنکور را به صورت علّت‌های درون‌فردی تفسیر کرد. این نتیجه نشان از عدم روایی و توجیه‌پذیری بسیاری از تفسیرات انجام گرفته بر اساس عامل‌های حاصل از آزمون سراسری ورود به دانشگاه دارد.

واژه های کلیدی: مدل‌های متغیر مکنون، تفسیر علّی، ارگودیسیتی، تغییرناپذیری اندازه‌گیری، کنکور، توانایی‌های درون‌فردی

[Downloaded from ihej.ir on 2025-08-01]

مقدمه و بیان مسئله

ارزیابی و آزمون برای قرن‌ها نقشی محوری در فرآیند انتخاب شغل و ورود به تحصیل داشته است. تاریخچهٔ ارزیابی و آزمون‌گیری مشابه یک قالیچه غنیِ بافته شده با روش‌ها، نظریه‌ها و عمل‌های مختلف است که در گذر تاریخ، تکامل یافته و با نیازهای متغیر و متنوع جامعه سازگار شده است. از اشکال اولیه امتحانات شفاهی در تمدن‌های باستان تا آزمون‌های استاندارد مورد استفاده در سیستم‌های آموزشی مدرن، سفر ارزیابی و آزمون‌گیری برای درک چگونگی ارزیابی افراد برای موقعیت‌ها و فرصت‌های آموزشی بسیار جذاب و ضروری است.

برای مثال، در چین باستان آزمون‌هایی به منظور انتخاب ادیبان برای استخدام در مشاغل حکومتی به صورت دوره‌ای برگزار می‌شدند. محور اصلی این آزمون‌ها را آموزه‌های کنفسیوس، فیلسوف و اندیشمند نامدار آن دوران، شکل داده بود. با این حال، با رونق گرفتن آزمون‌گیری و ورود این رشته به دانشگاه‌ها، مفاهیم مرتبط با سنجش به‌شکلی جدی‌تر مورد بحث و توسعه قرار گرفته است. از جملهٔ این مفاهیم می‌توان به «روایی» اشاره کرد. اگرچه تا به امروز، اجماع نظری در خصوص مفهوم روایی شکل نگرفته است اما معمولا آزمونی روا در نظر گرفته می‌شود که دقیقا توانایی‌های مورد نظر را مورد سنجش قرار دهد (Borsboom et al., 2004) یا زمانی که استفاده از نمرات آزمون و تفسیرهای حاصله از آن را بتوان توجیه کرد، نمرات آزمون روا در نظر گرفته می‌شوند (Kane, 2016). اهمیت مفهوم روایی در آزمون‌سازی سبب شده است تا محتوا و معنای آزمون به‌گونه‌ای جدی‌تر مورد کنکاش قرار گیرد. به همین دلیل، احتمالا با استانداردهای فعلی در خصوص آزمون‌سازی، دیگر آزمونی متشکل از آموزه‌های کنفسیوس به عنوان یک آزمون مناسب (روا) برای گزینش مشاغل حکومتی درنظر گرفته نمی‌شود.

آزمون سراسری ورود به دانشگاه‌ها و موسسات آموزش عالی کشور (کنکور) مهم‌ترین آزمون و ارزیابی صورت گرفته در تاریخ ۶۰ ساله خود (از سال ۱۳۴۷ هجری شمسی تا به امروز) در ایران بوده است. این آزمون (حداقل به صورت نظری) برای ارزیابی ابعاد مختلف دانش، مهارت و استعدادهای مرتبط با انتخاب بهترین داوطلبان برای برخورداری از آموزشِ عالی طراحی شده است. درک معنا و ضرورت این ابعاد برای افزایش اثربخشی و عادلانه بودن فرآیند گزینش بسیار مهم است.

زمانی که بحث ابعاد و عوامل اساسی شکل‌دهندهٔ آزمون سراسری مطرح می‌شود، مدل‌های متغیر مکنون (Bartholomew et al., 2011; Borsboom, 2008a) به عنوان یک گزینهٔ در دسترس و قابل اتکا برای شناخت این ابعاد و عوامل در نظر گرفته می‌شوند(Izanloo, 2021). این مدل‌ها به دلیل ماهیت شهودی و توانایی ارائه بینش نسبت به مجموعهٔ متنوعی از رفتارها و خلاصه کردن این رفتارها تنها در چند عامل مکنون توجه بسیاری را به خود جلب کرده‌اند. با این حال، درک معنای عامل‌های مکنون حاصل از مدل‌های متغیر مکنون بدون چالش نیست.

[Downloaded from ihej.ir on 2025-08-01]

یکی از دلایل اصلی تردید در تفسیر عامل‌ها، وجود مشکل عدم قطعیت عامل است. مشکل عدم قطعیت به این مسئله اشاره دارد که با توجه به داده‌های در دسترس، می‌توان بینهایت پاسخ‌های عاملی متفاوت بدست آورد که از لحاظ قدرت تبیین داده‌ها با یکدیگر یکسان هستند و معمولا یک روش قطعی برای انتخاب درست‌ترین پاسخ وجود ندارد (Elffers, 1980; Elffers et al., 1978). برابری مدل‌ها مشکل دیگری است که بر تفسیر نتایج حاصل از مدل‌های متغیر مکنون تاثیرگذار است. برابری مدل به مفهومی اطلاق می‌شود که در آن مدل‌های مختلف، علیرغم داشتن قالب یا ساختارهای متفاوت، تبیین یا پیش‌بینی‌های معادلی را برای مجموعه یکسانی از داده‌ها ارائه می‌دهند (van Bork et al., 2021). برای مثال، داده‌های مربوط به یک آزمون هوشبهر را می‌توان همزمان به صورت عامل‌های پیوسته، گسسته (تایپ‌ها) و یا به عنوان یک شبکه روان‌شناختی در نظر گرفت (Van Der Maas et al., 2006). با توجه به عدم قطعیت عامل‌ها و برابری مدل‌ها، نمی‌توان معنای عامل‌های حاصل از مدل‌های متغیر مکنون را صرفا با استفاده از داده‌های تجربی به‌صورت قطعی و دقیق مشخص کرد (همچنین به Michell (۲۰۲۳) مراجعه کنید). به همین دلیل، برای درک معنای این عامل‌ها نیاز به اتخاذ پیش‌فرض‌هایی در خصوص این مدل‌ها است که در ادامه به آن خواهیم پرداخت.

مبانی نظری و پیشینه پژوهش

تفسیر مدل‌های متغیر مکنون معمولاً به دو صورت امکان‌پذیر است: ۱) تفسیر توصیفی و ۲) تفسیر علّی. در ادامه، در ابتدا دیدگاه Jonas and Markon (۲۰۱۶) و Markon and Jonas (۲۰۱۶) در خصوص تفسیر توصیفی مطرح می‌گردد. سپس، نظر Bork et al. (۲۰۱۷) در مورد تفسیر علّی (در مقابل تفسیر توصیفی) و مزیّت‌های دیدگاه علّی مطرح خواهد شد.

مارکون و یونس در مقالات خود با مبنا قرار دادن نظریهٔ اطلاعات[1] به دفاع از خوانش توصیفی از متغیرهای مکنون پرداخته‌اند. از دیدگاه پارادایم توصیفی، متغیرهای مکنون تنها به‌عنوان خلاصه‌ای از داده‌ها تفسیر می‌شوند و ملاک مورد استفاده برای مقایسهٔ مدل‌ها، اصل امساک[2] در خلاصه کردن داده‌ها است. همان‌طور که مشخص است، مبنای زیربنایی پارادایم توصیفی که نظریه اطلاعات است با رویکردهای احتمالاتی رایج، همچون نگاه فراوانی‌گرا[3] و بیزی[4] تفاوت‌هایی دارد. در نگاه فراوانی‌گرا و بیزی، یک مفروضه اساسی در خصوص توزیع داده‌ها وجود دارد. به بیان دیگر، در این دیدگاه‌ها ما به این مفروضه، که توزیع مشخصی وجود دارد که داده‌های موردنظر از آن نمونه‌گیری شده است، باور داریم. بنابراین، در این دیدگاه‌ها هدف اساسی ایجاد روش‌هایی برای تخمین هرچه دقیق‌تر از این حقیقت[5] است. یا به بیان Myung et al.

1. information theory
2. parsimony
3. Frequentist
4. Bayesian
5. truth

[Downloaded from ihej.ir on 2025-08-01]

(۲۰۰۶): «نتیجه این است که هدف از انتخاب مدل[1] یافتن مدلی است که به حقیقت نزدیکتر است (ص۱۶۷)». در مقابل، پارادایم توصیفی پیشنهاد می‌دهد که هدف اصلی باید یافتن قاعده‌مندی‌ها[2] در داده‌ها باشد، و از این قاعده‌مندی‌ها برای فشرده‌سازی مجموعهٔ داده‌ها استفاده شود تا بدین وسیله بتوانیم داده‌ها را با استفاده از نمادهای کمتری، نسبت به تعداد نمادهای مورد نیاز برای بازنمایی واقعی داده‌ها، توصیف کنیم.

انتقادهای وارده از سوی مارکون و یونس نسبت به روان‌شناسان در استفاده از مدل‌های مکنون به‌صورت علّی را می‌توان به دو بخش تقسیم کرد. بخش اول از انتقادها را می‌توان به معضل عدم‌تعیین و برابری مدل‌ها در خصوص مدل‌های متغیر مکنون دانست (Jonas & Markon, 2016). معضل عدم‌تعیین و برابری مدل‌ها، همانطور که پیش‌تر مطرح شد، به این موضوع اشاره دارد که برازش مدل با داده‌ها، به‌صورت قطعی نمی‌تواند بیان دارد که مدلِ مفروض، مکانیزم ایجاد کنندهٔ داده‌ها بوده است. به بیان دیگر، بسیاری از مکانیزم‌های تولید داده می‌توانند ساختاری مشابه با مدل فرضی را در داده‌ها ایجاد کنند (Borgstede & Eggert, 2023). بنابراین، مدل‌های متغیر مکنون می‌توانند با داده‌ها برازش مناسب داشته باشند بدون اینکه مکانیزم واقعی ایجادکنندهٔ داده‌ها بوده باشند.

با درنظرگرفتن این مشکل، یونس و مارکون استفاده از مدل‌های متغیر مکنون را به‌صورت توصیفی پیشنهاد داده‌اند؛ زیرا همان‌طور که بیان گشت، دیدگاه توصیفی نسبت به مکانیزم ایجادکننده داده‌ها، بی‌تفاوت است. در این دیدگاه، مزیّتِ استفاده از یک متغیر مکنون، قابلیت انتقال‌پذیری آن در موقعیت‌های گوناگون است و از این نظر امساک بیشتری در مقایسه با استفاده نکردن از هیچ مدلی دارد؛ چراکه در این صورت، نیاز به یک مدل جدید برای هر مجموعه از داده‌ها داریم. به بیان دیگر، ما با استفاده از متغیر مکنون به امساک و تعمیم‌پذیری دست پیدا می‌کنیم، اما کدگذاری داده‌ها به قیمت از دست رفتن بخشی از اطلاعات تمام می‌شود. این هزینه متناسب با خطای اندازه‌گیری است، بدین معنا که هرچه خطای اندازه‌گیری بیشتر باشد، هزینه افزایش خواهد یافت (Jonas & Markon, 2016).

در بخش دوم، Markon and Jonas (۲۰۱۶) با معرفی مدل بلوک تصادفی[3] چندین مزیّت استفاده از این مدل‌ها به‌عنوان یک ابزار توصیفی در مقایسه با نگاه عاملی را برشمرده‌اند. از جمله مهم‌ترینِ ایرادات مطرح شده به نگاه عاملی در مقایسه با دیدگاه موردنظر مارکون و یونس را می‌توان مربوط به بحث شکل علّیت در سیستم‌های پیچیده[4] و همچنین ویژگی‌های برآینده[5] حاصل از این سیستم‌ها دانست. اما در نگاه ما، این بخش از انتقادات وارد شده را نمی‌توان به سود و در جهت حمایت از پارادایم توصیفی تفسیر کرد. در واقع، مدل بلوک تصادفی

[1]. model selection
[2]. regularities
[3]. stochastic block model
[4]. complex systems
[5]. emergent properties

[Downloaded from ihej.ir on 2025-08-01]

را می‌توان از جمله روش‌های مربوط به شبکه دانست. مدل‌های شبکه اگرچه از لحاظ مفهوم‌سازی علّی با دیدگاه رایج در مورد متغیرهای مکنون تفاوت دارند؛ اما بازهم با شکلی از علّیت سروکار دارند (به Borsboom (۲۰۱۷)؛ McNally et al. (۲۰۱۵) نگاه کنید). البته Markon and Jonas (۲۰۱٦) نیز به این نکته واقف بوده‌اند و در پایان مقالهٔ خود بیان داشته‌اند: «مهم است که تأکید کنیم که ما بحث نمی‌کنیم که امساک و علّیت باید به طور کامل در تحقیق ساختاری از هم جدا شوند، اگر اصلاً این امر امکان‌پذیر باشد. اگرچه مدل‌های ساختاری به مفروضات علّیت وابسته نیستند، تحلیل‌های ساختاری اغلب به مکانیزم‌های علّیِ احتمالی اشاره می‌کنند یا با طرح‌های علّی آگاهی‌بخش، تقویت می‌شوند. به‌علاوه، همان‌طور که در این قسمت مورد تأکید قرار گرفت، استقراهای علّیت و امساک محور با افزایش اطلاعات تجربی به یکدیگر میل می‌کنند، [بدین‌وسیله این نکته را] پیشنهاد می‌دهند که یک راه برای افزایش امساک، افزایش زایندگی علّی در مدل‌ها است (ص۱۱۵۵)». به همین دلیل در ادامه مقالهٔ فعلی به این بخش از انتقادات مارکون و یونس پرداخته نخواهد شد.

در پاسخ به انتقادات وارده از سوی Jonas and Markon (۲۰۱٦)، Bork et al. (۲۰۱۷) سه مزیّت را در خصوص استفاده از دیدگاه علّی در مقایسه با تفسیر توصیفی این مدل‌ها برشمرده‌اند:

۱. بسیاری از تحقیقات به دنبال تبیین همبستگی بین متغیرهای مورد نظر هستند تا خلاصه‌کردن آن، که این هدف با تفسیر علّی مطابقت بیشتری دارد.

۲. تفسیر علّی از مدل‌های متغیر مکنون، تمرکز عمده بر واریانس مشترکِ بین متغیرهای نشانگر[1] بجای توجه به واریانس اختصاصی را به‌صورت مشروع توجیه می‌کند.

۳. توجیه مفروضهٔ استقلال موضعی تنها با تفسیر علّی از این مدل‌ها به‌صورت مشروع قابل قبول است.

همان‌طور که در بخش قبل بیان شد، یکی از ایرادات وارده به دیدگاه علّی از سوی طرفداران دیدگاه توصیفی به این مسئله مربوط بود که با هر مجموعهٔ دلبخواهی از متغیرها، امکان ایجاد متغیرهای نهفته از واریانس مشترک متغیرهای نشانگر وجود دارد. اما این دیدگاه باهدف تحقیقات روان‌شناختی سازگاری ندارد. چراکه، متغیرهای نشانگر به‌صورت تصادفی انتخاب و وارد مدل نمی‌گردند. بلکه، از مدل‌های متغیر مکنون (به‌صورت عمده) برای امر اندازه‌گیری استفاده می‌شود که با انتخاب یک گروه همگن از متغیرهای نشانگر توسط محقق همراه است. همچنین، دیدگاه توصیفی به بسیاری از پدیده‌های مورد توجه در روان‌شناسی، مانند همبودی بین نشانگان آسیب‌شناسی پاسخی نمی‌دهد. در

[1]. indicator

[Downloaded from ihej.ir on 2025-08-01]

مقابل، نگاه علّی به مدل‌های متغیر مکنون، امکان تبیین این پدیده را فراهم می‌کند؛ بنابراین، نگاه علّی در مقایسه با رقیب توصیفی خود، پاسخ مناسب‌تری به مسئلهٔ تبیین، به‌عنوان یکی از مباحث اساسی در پیشرفت علم می‌دهد (Bork et al., 2017).

در مدل‌های متغیر مکنون بین واریانس مشترک و واریانس اختصاصی هر یک از نشانگرها تفاوت وجود دارد. اما از بین این دو، آنچه موردنظر محققین قرار دارد و در تحقیقات خود سعی در تبیین و تفسیر آن دارند، واریانس مشترک است. به بیان Bork et al. (۲۰۱۷)، این توجه به واریانس مشترک تنها زمانی قابل‌توجیه است که پژوهشگر به وجود یک عامل زیربنایی به‌عنوان علّت این واریانس مشترک باور داشته باشد: «[...] زمانی که مشخص شده است که سازه جایگاه علّت مشترک را ندارد [...] یا وضعیت سازه [از لحاظ علّتِ مشترک بودن] ناشناخته است [...]، عمل به تفسیر واریانس مشترک به‌تنهایی در میان مجموعه‌ای از سؤالات، مورد حمایت نیست (ص۵۹۳)».

استقلال موضعی یکی از مفروضات اساسی در مدل‌های متغیر مکنون به‌حساب می‌آید. استقلال موضعی بیان می‌دارد که مشروط بر متغیر(های) نهفته، متغیرهای نشانگر از یکدیگر مستقل هستند. با قبول یک تفسیر علّی از مدل‌های متغیر مکنون، می‌توان انتظار داشت که یک علّت مشترک در بین نشانگرها وجود دارد که سبب ایجاد همبستگی در بین این متغیرها گردیده و با کنترل آن، این همبستگی نیز از بین خواهد رفت. اما در سوی دیگر، دیدگاه توصیفی نمی‌تواند توجیه مناسبی برای استفاده از مفروضهٔ استقلال موضعی فراهم آورد. در واقع، می‌توان این ایراد را مطرح کرد که چرا باید از مدل‌های متغیر مکنون که بر مفروضه استقلال موضعی استوار هستند، استفاده شود و نه روش‌هایی چون تحلیل مؤلفه‌های اصلی (Bork et al., 2017).

همان‌طور که پیش‌تر بیان گشت، خانواده مدل‌های متغیر مکنون، تنها روش‌های آماری برای کاهش ابعاد داده‌ها نیستند. روش‌هایی چون تحلیل مولفه‌های اساسی، همبستگی کانونی، مقیاس‌بندی چند بُعدی و تحلیل خوشه نیز برای این منظور قابل‌استفاده هستند. این روش‌ها، در مقایسه با مدل‌های متغیر مکنون، صرفاً تبدیل‌ها و یا ویژگی‌های هندسی موجود در داده‌ها را نشان می‌دهند و نه عناصر موجود در یک مدل آماری را (Skrondal et al., 2004). برای نمونه، روش تحلیل مولفه‌های اصلی در مقایسه با تحلیل عاملی، تمامی واریانس (و نه واریانس مشترک) را در متغیر مکنون در نظر می‌گیرد و بدین ترتیب یک خلاصه با امساک بیشتر را از داده‌ها فراهم می‌آورد. برای نتیجه‌گیری باید گفت که یک نگاه توصیفی صرف، نمی‌تواند اصل استقلال موضعی را توجیه کند و بنابراین استفاده از دیدگاه علّی دراین‌خصوص مزیّت بیشتری دارد.

با توجه به مباحث مطرح شده می‌توان گفت که استفاده از یک نگاه علّی به تفسیر و مشروعیت استفاده از مدل‌های متغیر مکنون کمک می‌کند. با این‌حال، نوع تفسیر علّی از متغیرهای مکنونِ حاصل از داده‌های مقطعی (داده‌های مربوط به آزمون سراسری نیز از نوع مقطعی

[Downloaded from ihej.ir on 2025-08-01]

است)، محل بحث و مناقشه بوده است (Borsboom, 2015; Weinberger, 2015). تفسیر علّی درون‌فردی در مقابل تفسیر علّی بین‌فردی، دو گونهٔ اساسی از تفسیرهای مرتبط با عامل‌های مکنون را شکل می‌دهند (Borsboom et al., 2003; Molenaar, 2004).

در تفسیر درون فردی، متغیرهای مکنون به عنوان علّت‌هایی که سبب بروز رفتارهای مشاهده شده در هر فرد گردیده است، مفهوم‌سازی می‌شوند. برای درک بهتر، بیماری دیابت نوع اول را در نظر بگیرید. افرادی که از دیابت نوع اول رنج می‌برند، همگی اختلالی در لوزالمعده خود دارند که سبب کاهش تولید و ترشح انسولین در بدن می‌شود. به بیان دیگر، کاهش تولید انسولین علّت ایجاد علائم مربوط به بیماری دیابت (برای مثال، کاهش وزن، تشنگی شدید، خستگی بیش از حد و غیره) در هر فرد است (Borsboom, 2008b). در این مثال، دیابت نوع اول معادل متغیر مکنون و علائم مربوط به این بیماری معادل نشانگرها در مدل‌های مکنون هستند. برای درک بهتر این مثال را درنظر بگیرید: در تفسیر درون فردی از هوش، می‌توان این عامل را دلیل بروز رفتارهای هوشمندانه در افراد دانست که هرچه فرد بیشتر از این عامل برخوردار باشد، احتمال بروز رفتارهای هوشمندانه در وی بیشتر خواهد بود.

در مقابل، تفسیر علّی بین‌فردی متغیرهای مکنون محل بحث و مناقشه بوده است (Borsboom, 2015; Weinberger, 2015). در تفسیر بین‌فردی عمدهٔ تمرکز بر درک تفاوت‌های بین‌فردی در افراد جامعه است تا تغیرات درون هر یک از افراد. به بیان دیگر، تفاوت‌های جامعه در موقعیت متغیر مکنون باعث تفاوت‌های جامعه در ارزش مورد انتظار (میانگین) پاسخ به سوالات می‌شود (Borsboom et al., 2003). برای مثال، اگر ما هوش را یک متغیر مکنون علّی بین‌فردی بدانیم، می‌توان نتیجه گرفت که تفاوت‌های جامعه در طول پیوستار هوش سبب تفاوت‌های آن جامعه در میانگین بروز رفتار هوشمندانه بوده است. در این صورت آنچه این متغیر به ما می‌گوید، جایگاه هر شخص در مقایسه با دیگر افراد جامعه است. با این‌حال، این متغیر در مورد اینکه این علّت مکنون بر رفتار هر یک از آزمودنی‌ها چه تاثیری دارد، هیچ اطلاعاتی در اختیار ما قرار نمی‌دهد.

ناگفته پیداست که محققین در بکارگیری مدل‌های متغیر مکنون مِیل به استفاده از تفسیر علّی درون‌فردی دارند. با این وجود، تنها تحت شرایط خاصی که از آن به عنوان مفروضه ارگودیسیتی[1] نام برده می‌شود، می‌توان با داده‌های مقطعی از تفسیر علّی درون‌فردی دفاع کرد. Molenaar (۲۰۰٤) در مقاله خود با عنوان «بیانیه‌ای در خصوص روان‌شناسی به عنوان علم ایدیوگرافیک: بازگرداندن شخص به روان‌شناسی علمی، این بار برای همیشه» شرایط لازم برای رعایت مفروضه ارگودیستی را مطرح کرده است. به بیان مولینار، ارگودیسیتی تنها زمانی حاصل می‌شود که توزیع نمرات بین فردی در یک جامعه و توزیع نمرات حاصل از نمونه‌گیری مکرر در اشخاص در بینهایت، با یکدیگر یکسان باشند. بطور مثال، توزیع نمرات پرتاب صد سکه در یک زمان، با توزیع نمرات صد مرتبه پرتاب یک سکه، یکسان است که نشان از رعایت

[1]. Ergodicity

[Downloaded from ihej.ir on 2025-08-01]

فرض ارگودیستی دارد. به شکل دقیق‌تر، برای رعایت مفروضه ارگودیسیتی این دو شرط مورد نیاز است: ۱) همگنی جامعه در ویژگی مورد نظر (تمام زیر جوامع دارای توزیع احتمالاتی یکسان باشند) و ۲) ویژگی ایستا[1] بودن (مکانیزم ایجاد کننده داده‌ها در طول زمان ثابت باشد). با توجه به مطالب بیان شده، به‌نظر می‌رسد که تعمیم ابعاد حاصل از داده‌های بین‌فردی به درون افراد بسیار دشوار و حتی نادر است.

بررسی مفروضه ارگودیسیتی بسیار دشوار است؛ با این‌حال، غیر ممکن نیست. به بیان دقیق‌تر، اینکه چه میزان از شواهد تجربی برای تایید پیش‌فرض ارگودیسیتی مورد نیاز است مشخص نیست. اما در مقابل، یک مثال نقض از تخطی از این پیش‌فرض (برای مثال، تخطی از شرط همگنی جامعه) می‌تواند شاهد محکمی بر رد پیش‌فرض ارگودیسیتی در نظر گرفته شود؛ همانطور که دیدن تعداد زیادی قو سپید نمی‌تواند شاهدی قطعی بر درستی این گزاره که تمام قوها سپید هستند باشد اما در مقابل، دیدن یک قو سیاه می‌تواند شاهدی قطعی بر این گزاره که تمام قوها سپید نیستند باشد (Popper, 1962).

در این مقاله، هدف اصلی ما بررسی پیش‌فرض ارگودیسیتی از طریق بررسی شرط همگنی جامعه در ساختار عاملی آزمون سراسری است. به این منظور، در ابتدا ساختار عاملی مربوط به آزمون سراسری در رشته ریاضی شناسایی می‌شود. سپس، آزمودنی‌ها به چند خوشه متفاوت تقسیم می‌شوند. در ادامه، تغییرناپذیری اندازه‌گیری[2] در این خوشه‌ها نسبت به عامل‌های حاصل شده در گام اول مورد بررسی قرار خواهد گرفت. در صورتی که تغییرناپذیری اندازه‌گیری بین خوشه‌ها حمایت نشود، می‌توان نتیجه گرفت که تفسیر درون‌فردی عامل‌های مربوطه ممکن نیست.

بدین ترتیب، سوالات تحقیق در این مقاله عبارت‌اند از:

- دروس آزمون سراسری در گروه علوم ریاضی و فنی از چه ساختار عاملی برخوردار است؟
- داوطلبین شرکت کننده در آزمون سراسری گروه علوم ریاضی و فنی چه خوشه‌هایی را تشکیل می‌دهند؟
- آیا در ساختار عاملی مربوط به دروس آزمون سراسری گروه علوم ریاضی و فنی با توجه به خوشه‌های متفاوت داوطلبین تغییرناپذیری اندازه‌گیری وجود دارد؟

روش پژوهش

[1]. stationary
[2]. measurement invariance

[Downloaded from ihej.ir on 2025-08-01]

طرح پژوهش حاضر از نوع طرح‌های بین مقطعی و با هدف بررسی مفروضه ارگودیسیتی در آزمون سراسری رشتهٔ ریاضی و فیزیک بوده است.

جامعه آماری شامل تمامی آزمودنی‌های شرکت کننده در آزمون سراسری گروه علوم ریاضی و فنی در سال ۱۴۰۰ می‌شود که از این میان ۱۲۶۹۹ نفر به روش تصادفی به عنوان نمونه انتخاب شدند.

شاخص‌های (متغیرهای) بکار رفته در این مقاله را نمرات (بر حسب درصد) داوطلبین دروس آزمون سراسری گروه علوم ریاضی و فنی شامل ادبیات فارسی، عربی، معارف اسلامی، زبان انگلیسی، ریاضیات، فیزیک و شیمی تشکیل داده است.

در فرایند تحلیل داده در ابتدا از تحلیل عاملی اکتشافی به منظور شناسایی عامل‌های زیربنایی آزمون سراسری و شاخص‌های مرتبط با هر عامل استفاده شد. در ادامه، برازش مدل بدست آمده با استفاده از معادلات ساختاری مورد بررسی قرار گرفت. این مرحله به این دلیل از اهمیت برخوردار است که سنجش تغییرناپذیری اندازه‌گیری به برازش مدل بر داده‌های کلی وابسته است. سپس، با استفاده از تحلیل خوشه سلسله‌مراتبی، پنج خوشه از داوطلبین شرکت‌کننده در آزمون سراسری بر مبنای درصدهای دروس مختلف ایجاد گردید. در نهایت، تغییرناپذیری ساختار عاملی بدست آمده در این خوشه‌ها مورد آزمون قرار گرفت.

لازم به ذکر است که تمام تحلیل‌های آماری با استفاده از نرم‌افزار JASP (Team) نسخه 0.18.3 انجام پذیرفت.

یافته های پژوهش

نتایج تحلیل عاملی اکتشافی با استفاده از دروس آزمون سراسری گروه علوم ریاضی و فنی، یک ساختار سه عاملی را نشان داد (جدول ۱). برای شناسایی تعداد عامل‌ها از تحلیل موازی[1] استفاده شد. در تحلیل موازی، ارزش‌های ویژه داده‌های تجربی با ارزش ویژه داده‌های شبیه‌سازی شده مقایسه می‌شود و عواملی انتخاب می‌شوند که ارزش ویژهٔ بیشتری از داده‌های شبیه‌سازی شده داشته باشند (نمودار ارزش‌های ویژه تجربی و شبیه‌سازی شده در شکل ۱ قابل مشاهده است). شاخص‌های برازش، حاکی از برازش مناسب مدل با داده‌ها بود (= RMSEA 0.017, RMSEA 90% confidence = [0.009, 0.027], SRMR = 0.003, TLI = 0.998, CFI = 1, BIC = -13.885).

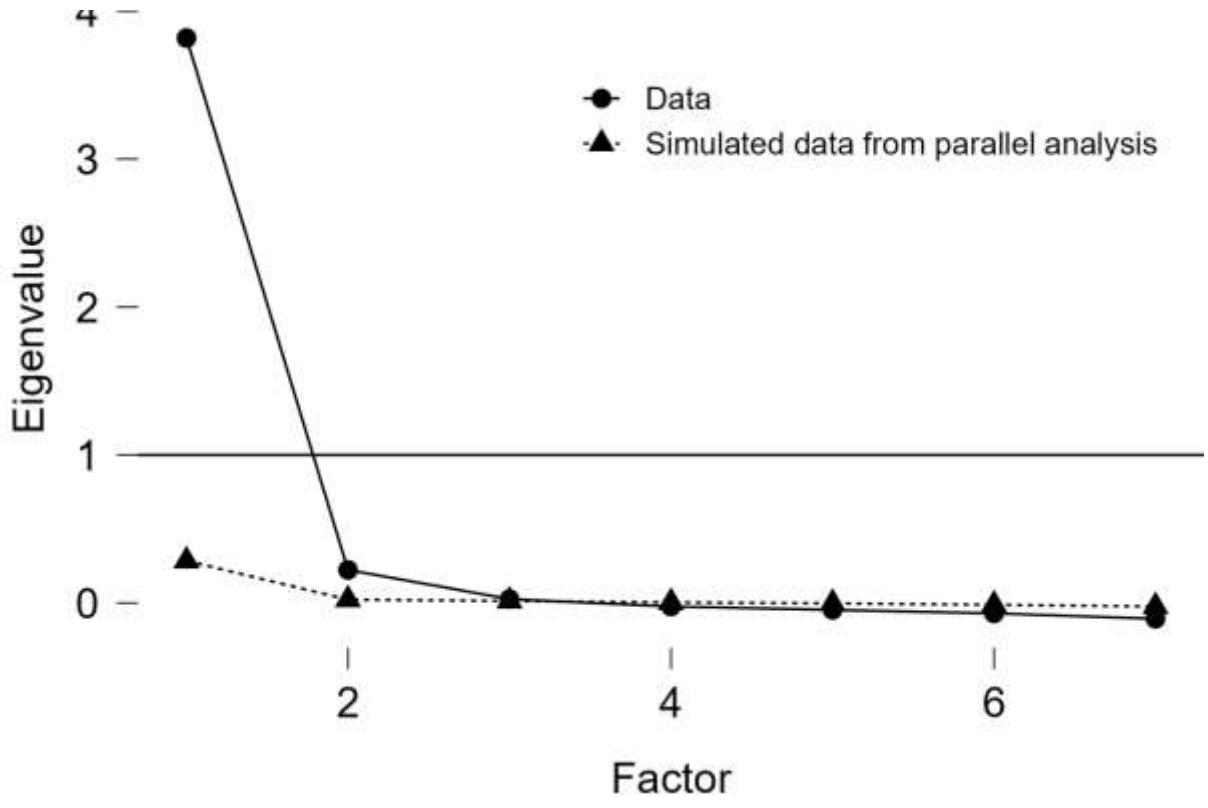
جدول ۱ مشخصات عامل‌ها در تحلیل عاملی اکتشافی پیش و پس از چرخش

	پاسخ بدون چرخش	پاسخ چرخش داده شده

[1]. parallel analysis

[Downloaded from ihej.ir on 2025-08-01]

	ارزش‌های ویژه	مجموع مجذورات بارها	سهم واریانس	سهم واریانس تراکمی	مجموع مجذورات بارها	سهم واریانس	سهم واریانس تراکمی
عامل اول	4.250	3.892	0.556	0.556	2.025	0.289	0.289
عامل دوم	0.696	0.322	0.046	0.602	1.709	0.244	0.533
عامل سوم	0.639	0.077	0.011	0.613	0.556	0.079	0.613



شکل ۱ نمودار مربوط به ارزش‌های ویژه داده‌های تجربی و شبیه‌سازی شده

پس از چرخش پرومکس، بارهای عاملی متناسب با هر درس به همراه ماترس ساختاری در جدول ۲ نمایش داده شده است. پس از چرخش، دروس ادبیات فارسی، عربی و معارف در عامل اول، دروس ریاضیات، فیزیک و شیمی در عامل دوم و زبان انگلیسی و ریاضیات در عامل سوم نقش پر رنگ‌تری ایفا کرده‌اند.

[Downloaded from ihej.ir on 2025-08-01]

جدول ۲ بارهای عاملی و ماتریس ساختاری مربوط به تحلیل عاملی اکتشافی بعد از چرخش پرومکس

		عامل اول	عامل دوم	عامل سوم	اختصاصی
بارهای عاملی	فارسی	0.641	0.007	0.154	0.410
	عربی	0.646	0.028	0.158	0.371
	معارف	0.873	0.057	-0.090	0.276
	انگلیسی	0.157	0.096	0.349	0.691
	ریاضیات	-0.065	0.544	0.338	0.398
	فیزیک	0.049	0.913	-0.047	0.163
	شیمی	0.311	0.489	0.027	0.402
ماتریس ساختاری	فارسی	0.761	0.604	0.636	-
	عربی	0.785	0.632	0.659	-
	معارف	0.849	0.646	0.601	-
	انگلیسی	0.488	0.474	0.537	-
	ریاضیات	0.594	0.747	0.695	-
	فیزیک	0.700	0.914	0.671	-
	شیمی	0.699	0.743	0.624	-

در ادامه، به منظور سنجش برازش ساختار عاملی بدست آمده از تحلیل عاملی اکتشافی، از مدل‌سازی معادلات ساختاری استفاده شد. این مرحله به این دلیل انجام شد که برای بررسی تغییرناپذیری اندازه‌گیری در این عامل‌ها، ابتدا باید این عامل‌ها به صورت کلی با داده‌ها از برازش مناسب برخوردار باشند. اگرچه آزمون برازش کلی مدل مجذور کای معنادار بود ($X^2 = 338.471, df = 7, p < 0.001$) که به

[Downloaded from ihej.ir on 2025-08-01]

دلیل حجم نمونهٔ تحلیل فعلی طبیعی است، با این‌حال دیگر شاخص‌های مربوط به برازش مدل حاکی از برازش مناسب این مدل با داده‌ها داشت (به جدول۳ نگاه کنید). نمودار مربوط به مدل ساختاری بکار برده شده، در شکل ۲ نمایش داده شده است.

جدول ۳ شاخص‌های برازش مربوط به مدل تاییدی و تغییر ناپذیری اندازه‌گیری در سطوح پیکربندی، ضعیف و قوی

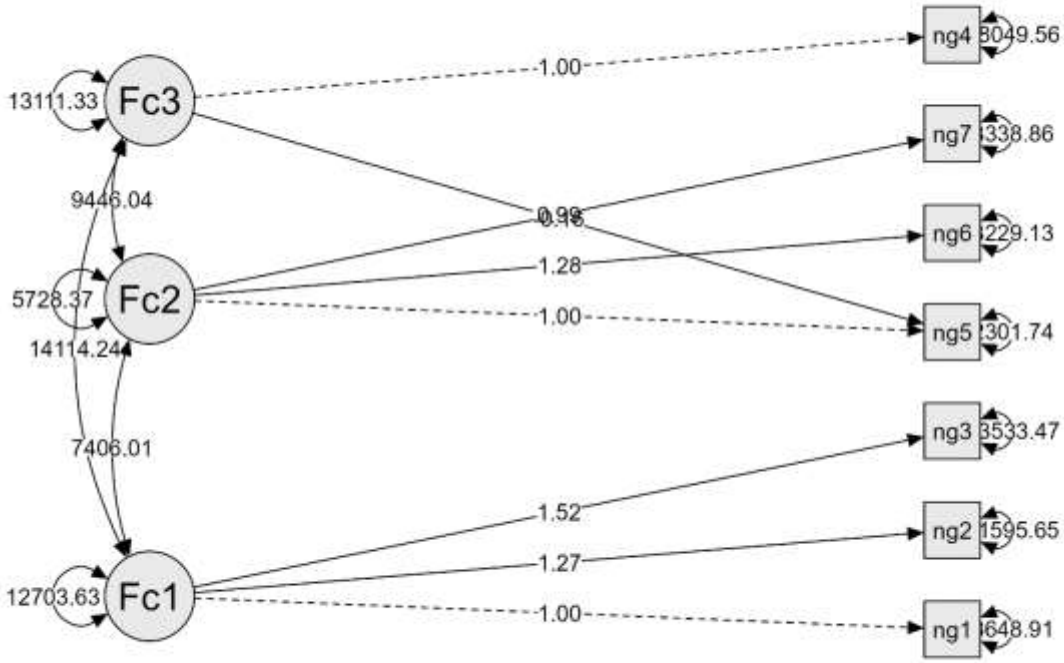
شاخص	مدل تاییدی با داده‌های کل داوطلبین	تغییرناپذیری پیکربندی	تغییرناپذیری ضعیف	تغییرناپذیری قوی
Comparative Fit Index (CFI)	0.993	0.821	0.586	0.495
T-size CFI	0.991	0.791	0.533	0.434
Tucker-Lewis Index (TLI)	0.978	0.659	0.351	0.441
Bentler-Bonett Non-normed Fit Index (NNFI)	0.978	0.659	0.351	0.441
Bentler-Bonett Normed Fit Index (NFI)	0.993	0.816	0.582	0.489
Parsimony Normed Fit Index (PNFI)	0.331	0.428	0.372	0.442
Bollen's Relative Fit Index (RFI)	0.978	0.649	0.345	0.435
Bollen's Incremental Fit Index (IFI)	0.993	0.823	0.588	0.495
Relative Noncentrality Index (RNI)	0.993	0.821	0.586	0.495
Log-likelihood	-541398.838	-509870.646	-510708.608	-511043.161
Number of free parameters	28.000	120.000	108.000	80.000
Akaike (AIC)	1.083e+6	1.020e+6	1.022e+6	1.022e+6
Bayesian (BIC)	1.083e+6	1.021e+6	1.022e+6	1.023e+6
Sample-size adjusted Bayesian (SSABIC)	1.083e+6	1.020e+6	1.022e+6	1.023e+6
Root mean square error of approximation (RMSEA)	0.061	0.095	0.131	0.122
RMSEA 90% CI lower bound	0.056	0.091	0.127	0.118
RMSEA 90% CI upper bound	0.067	0.100	0.135	0.125
RMSEA p-value	4.394e-4	1.221e-15	0.000	0.000
T-size RMSEA	0.067	0.045	0.060	0.056
Standardized root mean square residual (SRMR)	0.012	0.055	0.089	0.096
Hoelter's critical N (α = .05)	527.534	707.618	370.047	412.119
Hoelter's critical N (α = .01)	692.532	794.178	411.226	450.966

[Downloaded from ihej.ir on 2025-08-01]

شاخص	مدل تاییدی با داده‌های کل داوطلبین	تغییرناپذیری پیکربندی	تغییرناپذیری ضعیف	تغییرناپذیری قوی
Goodness of fit index (GFI)	0.995	0.989	0.979	0.973
McDonald fit index (MFI)	0.987	0.952	0.891	0.869
Expected cross validation index (ECVI)	0.031	0.123	0.253	0.301

مدل	$\chi 2$ ($\Delta\chi 2$)	Df (ΔDf)	CFI (ΔCFI)
پیکربندی	1314	55	0.821
ضعیف	2990 (1676)	67 (12)	0.586 (0.235)
قوی	3659 (669)	95 (28)	0.495 (0.091)

[Downloaded from ihej.ir on 2025-08-01]



شکل ۲ دیاگرام ساختار عاملی مربوط به مدل تاییدی

نکته. ng1 تا ng7 به ترتیب نشان‌دهندهٔ دروس زبان فارسی، عربی، معارف، انگلیسی، ریاضیات، فیزیک و شیمی است

در گام بعد، از تحلیل خوشه سلسله‌مراتبی با استفاده از فاصله اقلیدسی و اتصال‌دهندهٔ[1] والد[2] برای شناسایی خوشه‌های عمدهٔ شرکت کنندگان در آزمون سراسری گروه علوم ریاضی و فنی استفاده شد. انتخاب تعداد خوشه‌ها بر اساس میزان واریانس تبیین شده توسط خوشه‌ها در دروس آزمون سراسری و همچنین، تفسیرپذیری این خوشه‌ها انجام گرفت که در نهایت پاسخ پنج خوشه‌ای به دیگر پاسخ‌ها ترجیح داده شد. مشخصات مربوط به دقت تحلیل سلسله مراتبی در جدول ٤ نمایش داده شده است. همچنین، میانگین هر یک از دروس در هر خوشه در جدول ۵ قابل مشاهده است.

[1]. linkage
[2]. Wald

[Downloaded from ihej.ir on 2025-08-01]

جدول ٤ مشخصات عملکردی تحلیل خوشه سلسله مراتبی

خوشه‌ها	تعداد آزمودنی	R^2	AIC	BIC	Silhouette
5	12669	0.609	911143906.630	911144167.280	0.190

به منظور سنجش تغییرناپذیری اندازه‌گیری در ساختار عاملی مطلوب بدست آمده برای آزمون سراسری گروه علوم ریاضی و فنی، از متغیر مربوط به گروه‌بندی خوشه‌های آزمودنی‌ها استفاده شد. بدین ترتیب، تغییرناپذیری اندازه‌گیری در پنج سطح پیکربندی، متریک (ضعیف)، مقیاس (قوی)، اکید[1] و ساختاری[2] برای این شش گروه در دستور اجرا قرار گرفت.

جدول ٥ میانگین دروس داوطلبین در هر خوشه

	تعداد	فارسی	عربی	معارف	انگلیسی	ریاضیات	فیزیک	شیمی
خوشه اول	6076	50.217	80.576	61.305	3.144	17.907	3.975	9.374
خوشه دوم	1392	239.673	316.678	424.017	52.292	59.221	89.262	95.310
خوشه سوم	2645	78.915	179.148	150.258	122.977	24.625	8.615	23.509
خوشه چهارم	1817	165.668	236.981	277.059	531.001	57.275	58.718	68.451
خوشه پنجم	739	437.037	578.019	646.191	621.899	204.494	343.218	267.848

نکته. اعداد ارائه شده در جدول، درصدهای آزمودنی‌ها ضرب در عدد ۱۰ هستند. برای مثال، عدد ٤۳۷ نشان دهندهٔ ٤۳/۷ درصد است

اولین سطح که با عنوان تغییرناپذیری پیکربندی[3] شناخته می‌شود، نشان می‌دهد که عامل‌های معادل در مدل‌ها از نظر مفهومی در بین گروه‌ها یکسان هستند. این ابتدایی ترین سطح تغییر ناپذیری است، که نشان می دهد ساختار عامل‌ها در بین گروه‌ها مشابه است. شاخص‌های برازش مدل در بین گروه‌ها در سطح پیکربندی در جدول ۳ نمایش داده شده است. نتایج حاصله حاکی از عدم رعایت تغییرناپذیری

1. strict
2. structural
3. configural

[Downloaded from ihej.ir on 2025-08-01]

پیکربندی داشت ($CFI < 0.90$). با توجه به اینکه سطوح تغییرناپذیری اندازه‌گیری به سطح‌های پیشین وابسته هستند، بنابراین باید نتیجه بگیریم که تغییرناپذیری اندازه‌گیری در هیچ سطحی در بین خوشه‌های ما رعایت نشده بود (همانطور که در جدول ۳ قابل مشاهده است). با این حال، شاخص‌های برازش مربوط به سطوح متریک[1] و مقیاس[2] نیز در جدول ۳ نمایش داده شده است. لازم به ذکر است که سطوح اکید و ساختاری به دلیل مشکلات در برآورد مدل انجام نپذیرفت.

نتیجه‌گیری

نتایج تحلیل عاملی اکتشافی و در ادامه تحلیل عاملی تاییدی حاکی از برازش یک ساختار سه عاملی با داده‌ها بود. عامل اول از دروس زبان فارسی، عربی و معارف، عامل دوم از دروس ریاضیات، فیزیک و شیمی، و در نهایت، عامل سوم از ریاضیات و انگلیسی تشکیل شده بودند.

همانطور که در این مقاله بیان شد، تفسیر عامل‌های حاصله از مدل‌های متغیر مکنون به صورت درون‌فردی نیازمند پیش‌فرض ارگودیستی است. همچنین، بیان شد که این پیش‌فرض تنها زمانی تحقق می‌یابد که آزمودنی‌ها در پدیدهٔ مورد بررسی همگن بوده باشند. به همین دلیل، تغییرناپذیری ساختار عاملی کنکور در خوشه‌هایی از داوطلبین گروه علوم ریاضی و فنی مورد بررسی قرار گرفت. نتیجه حاکی از عدم وجود تغییرناپذیری اندازه‌گیری و در نتیجه ناهمگنی پدیده در بین آزمودنی‌ها بود. بنابراین، پیش‌فرض ارگودیستی رعایت نشده و امکان تفسیر درون‌فردی عامل‌ها میسر نیست.

حال این سوال مطرح می‌شود که عدم امکان تفسیر درون فردی عامل‌های آزمون سراسری به چه معنا است؟ فرض کنید نتیجه تحلیل عاملی نشان داده است که درس ریاضیات یک عامل مستقل را شکل داده است. حال اگر دو آزمودنی در در درس ریاضیات به ترتیب درصدهای ۴۰ و ۷۰ را کسب کرده باشند، بطور معمول نتایج تحلیل عاملی اینگونه تفسیر می‌شوند که سطح توانایی بالاتر آزمودنی دوم در عامل مکنونِ مربوط به ریاضیات، نسبت به آزمودنی اول، علّت کسب درصد بالاتر وی شده است. اما باید گفت که این تفسیر بدون داشتن شواهدی در خصوص رعایت پیش‌فرض ارگودیسیتی، صحیح به‌نظر نمی‌رسد. چراکه ممکن است علّت‌های کاملا متفاوتی در بدست آمدن نتیجهٔ حاصله دخیل بوده باشند. برای مثال، ممکن است که آزمودنی اول خانمی بوده باشد که در دوره عادت ماهانه خود به سر می‌برده است یا داوطلبی بوده که به دلیل استرس و اضطراب بیمارگونه نتوانسته است عملکرد مناسبی از خود نشان دهد. بنابراین، نباید عامل‌های حاصل شده از مدل‌های متغیر مکنون در خصوص آزمون سراسری را به صورت علّت‌های درون‌فردی تفسیر کرد.

[1]. metric
[2]. scalar

[Downloaded from ihej.ir on 2025-08-01]

اما اگر نمی‌توان عامل‌های بدست آمده از مدل‌های مکنون در کنکور را به صورت علّت‌های درون‌فردی تفسیر کرد، پس عامل‌های بدست آمده به چه معنایی هستند؟Borsboom et al. (۲۰۰۳) پیشنهاد می‌دهند که این عوامل را می‌توان به‌عنوان ویژگی‌های برآیندهٔ[1] جامعه درنظر گرفت. ویژگی برآینده اشاره به خواصی دارد که در اثر تعامل عناصر پایه‌ای یک سیستم ایجاد می‌شود، اما به‌طور کیفی با خواص تک تک این عناصر متفاوت است و به صورت کامل نمی‌توان آن را به ویژگی‌های عناصر تقلیل داد. برای نمونه، خاصیت آهن‌ربایی در اثر چینش خاص قطب‌های مغناطیسی عناصر آهن ایجاد می‌شود ولی قابل توضیح و تقلیل به ویژگی‌های یک عنصر آهن نیست. طبق این ادعا، عامل‌های حاصل شده در تحلیل ما نشان دهندهٔ ویژگی‌های برآینده در جامعهٔ آزمودنی‌ها است که قابل تقلیل به تک تک افراد نیست (Barcel'o et al., 2014).

پذیرش این ادعا که عامل‌ها نشان‌دهندهٔ ویژگی‌های برآیندهٔ جامعه است، با دو چالش رو به رو است: ۱) بدست آوردن شواهد مورد نیاز برای توجیه این ادعا بسیار دشوار است، و ۲) این ادعا با دلایل استفاده از آزمون سراسری همسو نیست. Mõttus (۲۰۱۶) شرط باور به وجود ویژگی‌های برآینده را نقش علّی این ویژگی‌ها دانسته است. به بیان دیگر، اگر این عامل‌ها هیچ نقش علّی مفیدی را در سیستم علم بازی نمی‌کنند، درنتیجه باور به وجود آنها عاقلانه نیست (ممکن است این عامل‌ها صرفا مصنوعات آماری باشند). اما بررسی نقش این عامل‌ها چندان ساده نیست. یک راه می‌تواند استفاده از نمرات عامل(های) بدست آمده و بررسی روابط این نمرات با مجموعه‌ای از بازده‌های آموزشی باشد. اما چندین مشکل در این میان وجود دارد. در ابتدا، آزمون سراسری یک آزمون خنثی نیست. به بیان دیگر، نتایج حاصل از آزمون سراسری به منظور گزینش دانشگاهی مورد استفاده قرار می‌گیرید و به همین دلیل، نتایج روابط مورد نظر ما را مقشوش می‌کند. حتی اگر بتوانیم بازده‌هایی را شناسایی کنیم که تحت تاثیر این ویژگی هنجاری آزمون سراسری قرار نگیرند، باید توجه داشته باشیم که سهم نمرات حاصل از عامل(ها) در پیش‌بینی این بازده‌ها باید چیزی بیش از اجزاء تشکیل‌دهندهٔ آن باشد تا بتوان استفاده از آن را توجیه کرد (Mõttus, 2016). علاوه بر این، در خصوص اینکه آیا متغیرهای بین‌فردی را نیز می‌توان به عنوان علّت در نظر گرفت تردید وجود دارد (Weinberger, 2015).

فارغ از مسئلهٔ قبل، این سوال مطرح می‌شود که از اساس به چه دلیلی از مدل‌های متغیر مکنون در تحلیل داده‌های آزمون سراسری استفاده می‌شود؟ به‌نظر می‌رسد دلیل اصلی استفاده از مدل‌های متغیر مکنون، ارتباط این مدل‌ها با مفهوم اندازه‌گیری باشد. به بیان دیگر، می‌توان از نتایج این مدل‌ها در توجیه روایی آزمون انجام پذیرفته استفاده کرد. برای مثال در نتایج تحلیل ما سه عامل برای کنکور شناسایی شد که عامل دوم مربوط به دروس تخصصی در گروه علوم ریاضی و فنی بود. بدین ترتیب، کسی می‌تواند ادعا کند که یک توانایی زیربنایی برای پاسخ

[1]. emergent properties

[Downloaded from ihej.ir on 2025-08-01]

دادن به دروس تخصصی در گروه ریاضی وجود دارد که آزمون سراسری آن را اندازه‌گیری می‌کند. با این تفاسیر، حتی در صورت پیدا کردن یک نقش علّی برای عامل‌های مکنون بین‌فردی، نمی‌توان از آن برای توجیه اندازه‌گیری توانایی‌های افراد استفاده کرد که نتیجهٔ آن عدم اشتیاق به این نوع از تحلیل و نتایج حاصل از آن است.

همچنین در این مقاله، ما از تحلیل خوشه استفاده کردیم که توجه به نتایج آن خالی از لطف نیست. در خوشهٔ اول که نیمی از آزمودنی‌های ما را در خود جای داده بود، نمرات افراد در تمامی دروس نزدیک به صفر بود. این نتیجه حاکی از آن است که در حدود نیمی از داوطلبین شرکت کننده در آزمون سراسری در اصل برای این آزمون تلاشی نکرده‌اند و صرفا نقش سیاهی لشکر را در این آزمون ایفا کرده‌اند. خوشهٔ دوم را افرادی تشکیل داده بودند که صرفا در دروس ادبیات فارسی، عربی و معارف درصدهای خوبی داشتند و در دروس تخصصی و زبان انگلیسی درصدهای پایینی (کمی بیشتر از صفر) را از خود به نمایش گذاشته‌اند. این نتیجه حاکی از آن است که گروهی از داوطلبان صرفا دروس عمومی را برای مطالعه انتخاب می‌کنند. در خوشهٔ سوم، افراد در همه دروس درصدهای ضعیفی از خود به نمایش گذاشته‌اند (بین ۱۰ تا ۲۰ درصد). این گروه از آزمودنی‌ها احتمالا برای آزمون سراسری آمادگی چندانی نداشته‌اند، با این حال شانس خود را در پاسخ دادن به سوالات با اطلاعات عمومی خود امتحان کرده‌اند. در خوشهٔ چهارم میانگین درصدها بسیار شبیه به خوشه دوم بوده است، به این معنا که داوطلبین در دروس عمومی موفق‌تر عمل کرده‌اند. اما تفاوت اساسی این خوشه با خوشهٔ دوم در درس زبان انگلیسی است که در خوشه چهارم شاهد افرادی هستیم که درصدهای عالی را از خود در این درس نشان داده‌اند. در نهایت در خوشهٔ پنجم شاهد افرادی هستیم که در تمام دروس (به صورت خاص، دروس تخصصی) عملکرد متوسط رو به بالایی داشته‌اند و در واقع داوطلبین اصلی کنکور بوده‌اند.

یکی از دلایل ما در استفاده از تحلیل خوشه، این بود که نشان دهیم دیگر روش‌های کاهش بعد و همچنین توجه به آزمودنی‌ها بجای متغیرها (درس‌ها) گاه می‌تواند اطلاعات ارزشمندتری در اختیار ما قرار دهد. باید توجه داشت که روش‌های دیگر کاهش ابعاد معمولا از مفروضهٔ سخت‌گیرانه‌ای به‌نام استقلال موضعی برخوردار نیستند که همین امر باعث باز بودن دست پژوهشگر در تفسیر اطلاعات حاصله از این تحلیل‌ها می‌شود (Henning, 1989). اما ذکر این نکته ضروری است که عدم پایبندی به مفروضه استقلال موضعی سبب می‌شود که نتوان با قطعیت نتایج حاصل شده از این تحلیل‌ها را با یک نگاه واقعیت‌گرایانه و در نتیجه علّی تفسیر کرد. لذا باید در انتخاب مدل به هدف استفاده از آن دقت کرد. در نهایت باید گفت که بحث در خصوص ویژگی های فنی کنکور بدون توجه به ابعاد نظری و مفهومی آن مانند دوشیدن گاوی است که هنوز جنسیت آن مشخص نشده باشد؛ ممکن است که شما بتوانید آن را خوب بدوشید، ولی نمی‌توانید از فرآورده آن اطمینان داشته باشید.

[Downloaded from ihej.ir on 2025-08-01]

تشکر و قدردانی

بدین وسیله از تمامی دوستانی که به نحوی در پیشبرد این مقاله به ما یاری رساندند تشکر می‌کنیم.

تعارض منافع

نویسندگان هیچ‌گونه تعارض منافعی در رابطه با نتایج پژوهش ندارند.

Mahdi Karvandi Renani https://orcid.org/0000-0002-8337-0749

Ebrahim Khodaie https://orcid.org/0000-0002-5634-5003

Keyvan Salehi https://orcid.org/0000-0001-8673-4248

Ali Moghaddam Zadeh https://orcid.org/0000-0001-7094-2923

منابع

Barcel'o, C., Carballo-Rubio, R., Garay, L. J., & Jannes, G. (2014). Electromagnetism as an emergent phenomenon: a step-by-step guide. *New Journal of Physics*, *16*.

Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). Latent Variable Models and Factor Analysis: A Unified Approach.

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

Bork, R. V., Wijsen, L. D., & Rhemtulla, M. (2017). Toward a Causal Interpretation of the Common Factor Model. *Disputatio*, *9*(47), 581-601. https://doi.org/doi:10.1515/disp-2017-0019

Borsboom, D. (2008a). Latent Variable Theory. *Measurement: Interdisciplinary Research and Perspectives*, *6*(1-2), 25-53. https://doi.org/10.1080/15366360802035497

Borsboom, D. (2008b). Psychometric perspectives on diagnostic systems. *Journal of clinical psychology*, *64 9*, 1089-1108.

Borsboom, D. (2015). What is causal about individual differences? : A comment on Weinberger. *Theory & Psychology*, *25*(3), 362-368. https://doi.org/10.1177/0959354315587784

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203-219. https://doi.org/10.1037/0033-295X.110.2.203

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*(4), 1061-1071. https://doi.org/10.1037/0033-295X.111.4.1061

Elffers, H. (1980). On Uninterpretability of Factor Analysis Results. *Transactions of the Institute of British Geographers*, *5*(3), 318-329. https://doi.org/10.2307/621845

Elffers, H., Bethlehem, J., & Gill, R. (1978). Indeterminacy problems and the interpretation of factor analysis results. *Statistica Neerlandica*, *32*(4), 181-199. https://doi.org/https://doi.org/10.1111/j.1467-9574.1978.tb01398.x

Henning, G. (1989). Meanings and implications of the principle of local independence. *Language Testing*, *6*(1), 95-108. https://doi.org/10.1177/026553228900600108

Izanloo, B. (2021). Comparing Methods of Determining Test Factor Structure Using Empirical data: The Case of National Entrance Exam in 2016. *Educational Measurement and Evaluation Studies*, *11*(33), 27-58. https://doi.org/10.22034/emes.2021.247577

Jonas, K. G., & Markon, K. E. (2016). A descriptivist approach to trait conceptualization and inference. *Psychological Review*, *123*(1), 90-96. https://doi.org/10.1037/a0039542

[Downloaded from ihej.ir on 2025-08-01]

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

Markon, K. E., & Jonas, K. G. (2016). Structure as cause and representation: Implications of descriptivist inference for structural modeling across multiple levels of analysis. *Journal of Abnormal Psychology*, *125*(8), 1146-1157. https://doi.org/10.1037/abn0000206

McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science*, *3*(6), 836-849. https://doi.org/10.1177/2167702614553230

Michell, J. (2023). "Professor Spearman has drawn over-hasty conclusions": Unravelling psychometrics' "Copernican Revolution". *Theory & Psychology*, *33*(5), 661-680. https://doi.org/10.1177/09593543231179446

Molenaar, P. C. M. (2004). A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*(4), 201-218. https://doi.org/10.1207/s15366359mea0204_1

Mõttus, R. (2016). Towards more rigorous personality trait–outcome research. *European Journal of Personality*, *30*(4), 292-303. https://doi.org/10.1002/per.2041

Myung, J. I., Navarro, D. J., & Pitt, M. A. (2006). Model selection by normalized maximum likelihood. *Journal of Mathematical Psychology*, *50*(2), 167-179. https://doi.org/10.1016/j.jmp.2005.06.008

Popper, K. R. (1962). *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge.

Skrondal, A., Rabe-Hesketh, S., & ProQuest. (2004). *Generalized latent variable modeling : multilevel, longitudinal, and structural equation models*. Chapman & Hall/CRC. http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=012824991&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA

Team, J. *JASP (Version 0.18.3)[Computer software]*. https://jasp-stats.org/

van Bork, R., Rhemtulla, M., Waldorp, L. J., Kruis, J., Rezvanifar, S., & Borsboom, D. (2021). Latent variable models and networks: Statistical equivalence and testability. *Multivariate Behavioral Research*, *56*(2), 175-198. https://doi.org/10.1080/00273171.2019.1672515

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842-861. https://doi.org/10.1037/0033-295X.113.4.842

Weinberger, N. (2015). If intelligence is a cause, it is a within-subjects cause. *Theory & Psychology*, *25*(3), 346-361. https://doi.org/10.1177/0959354315569832

[Downloaded from ihej.ir on 2025-08-01]