



## Recognizing and prioritizing fair classroom assessment themes using Fuzzy Analytical Hierarchy Process (FAHP) method

Ali Baniasadi<sup>1</sup>, Keyvan Salehi<sup>2</sup>, Ebrahim Khodaie<sup>3</sup>, Khosro Bagheri<sup>4</sup>, Balal Izanloo<sup>5</sup>

1. Ph.D. Student, University of Tehran, Faculty of Psychology and Education, Tehran, Iran, [ali.baniasadi@ut.ac.ir](mailto:ali.baniasadi@ut.ac.ir)

2. Assistant Professor, University of Tehran, Faculty of Psychology and Education, Tehran, Iran. Corresponding Author: [keyvansalehi@ut.ac.ir](mailto:keyvansalehi@ut.ac.ir)

3. Associate Professor, University of Tehran, Faculty of Psychology and Education, Tehran, Iran, [Khodaie@ut.ac.ir](mailto:Khodaie@ut.ac.ir)

4. Professor, University of Tehran, Faculty of Psychology and Education, Tehran, Iran, [khabagheri@ut.ac.ir](mailto:khabagheri@ut.ac.ir)

5. Associate Professor, University of Kharazmi, Faculty of Psychology and Education, Tehran, Iran, [izan.b@khu.ac.ir](mailto:izan.b@khu.ac.ir)

### Article Info

### ABSTRACT

#### Article Type:

#### Research Article

#### Received

2021/04/13

#### Received in

revised form

2022/07/12

#### Accepted

2022/07/27

#### Published online

2022/08/14

**Objective:** The aim of this study was to identify and weigh the themes and sub- themes of fair classroom assessment.

**Methods:** To achieve this aim, an interpretive phenomenology study was carried out. Semi-structured qualitative interviews were used to collect data. Using criterion sampling, after conducting 29 individual interviews, one group interview and one focus group interview with students who had at least two semesters of experience in the University of Tehran, inductive thematic saturation was obtained. Fuzzy analytical hierarchy process method was used to weigh these themes. Data for this stage were collected by a pairwise comparison questionnaire that included 53 comparisons from a sample of 9 experts in the field of assessment.

**Results:** Students' point of view analysis was performed using the Ricoeur method. Extraction of meaning units led to the identification of 20 sub-themes in 5 main themes of "educational fairness", "procedural fairness", "Nature of assessment", "Interactional fairness", and "fairness in paper and pencil test". The results of fuzzy analytical hierarchy process showed that among the 5 main themes, "educational fairness" with a weight of 0.318 was in the first place, followed by the themes of "procedural fairness", "nature of assessment", "interactional fairness", and "fairness in paper and pencil test" respectively. The final weights of the sub- themes were obtained by multiplying the weight of the main themes by the relative weight of the sub- themes; Accordingly, the sub- theme of "teaching quality" with a weight of 0.22 in the first place and the sub- themes of "informational fairness" and "professionalism in teaching" with weights of 0.11 and 0.09 were in the next positions, respectively.

**Conclusion:** Students' perception of fairness in assessment is a combination of the concepts of measurement theories and organizational justice, and the high weight of the theme of "educational fairness" among the all themes and sub- theme of "teaching quality" of this theme among all sub- themes obtained showed that Violation of educational fairness in the dimension of teaching quality will lead to an unfair perception of assessment. Conversely, if people feel that the teacher's teaching is of good quality and that the procedures used to assign the results are fair, even if it is not desirable for them, they will accept the results and understand them more fairly.

**Keywords:** *classroom assessment; fairness; justice; phenomenology study*

**Cite this article:** Baniasadi, Ali, Salehi, and Keyvan. Khodaie, Ebrahim. Bagheri, Khosro. & Izanloo, Balal. (2022). Recognizing and prioritizing fair classroom assessment themes using Fuzzy Analytical Hierarchy Process (FAHP) method. *Educational Measurement and Evaluation Studies*, 12 (37): 1-19 pages. DOI: 10.22034/EMES.2022.527579.2241



© The Author(s).

Publisher: National Organization of Educational Testing (NOET)

# شناسایی و اولویت‌بندی مضمون‌های سنجش کلاسی عادلانه با استفاده از روش فرایند تحلیل سلسه مراتبی فازی

علی بنی‌اسدی<sup>۱</sup>، کیوان صالحی<sup>۲</sup>، ابراهیم خدایی<sup>۳</sup>، خسرو باقری<sup>۴</sup>، بلال ایزانلو<sup>۵</sup>

۱. دانشجوی دکتری سنجش و اندازه‌گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران، [ali.baniasadi@ut.ac.ir](mailto:ali.baniasadi@ut.ac.ir)

۲. استادیار گروه روش‌ها و برنامه‌های آموزشی، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران، نویسنده مسئول: [keyvansalehi@ut.ac.ir](mailto:keyvansalehi@ut.ac.ir)

۳. دانشیار گروه روش‌ها و برنامه‌های آموزشی، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران، [khodaie@ut.ac.ir](mailto:khodaie@ut.ac.ir)

۴. استاد گروه مبانی فلسفی و اجتماعی آموزش پرورش، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران، [khabagheri@ut.ac.ir](mailto:khabagheri@ut.ac.ir)

۵. استادیار گروه برنامه‌ریزی درسی، دانشکده روانشناسی و علوم تربیتی، دانشگاه خوارزمی، تهران، ایران، [izan.b@khu.ac.ir](mailto:izan.b@khu.ac.ir)

## اطلاعات مقاله چکیده

نوع مقاله:	هدف: پژوهش حاضر با هدف شناسایی و وزن‌دهی مضمون‌ها و زیرمضمون‌های سنجش کلاسی عادلانه انجام شد.
مقاله پژوهشی	<b>روش پژوهش:</b> برای دستیابی به این هدف، ابتدا یک مطالعه کیفی به روش پدیدارشناسی تفسیری انجام شد. برای جمع‌آوری داده‌ها از مصاحبه نیمه ساختاریافته استفاده شد. با استفاده از نمونه‌گیری ملاکی، پس از انجام ۲۹ مصاحبه انفرادی، یک مصاحبه گروهی و یک گروه کانونی با دانشجویانی که حداقل دو ترم تجربه مورد سنجش عملکرد واقع شدن در دانشگاه تهران داشتند، اشباع موضوعی استقرایی به‌دست آمد. برای وزن‌دهی به این مضمون‌ها از روش فرآیند تحلیل سلسله‌مراتبی فازی استفاده شد. داده‌های این مرحله با استفاده از پرسشنامه مقایسه زوجی که شامل ۵۳ مقایسه بود از یک نمونه نه نفری از متخصصان حوزه سنجش جمع‌آوری شد.
دریافت	۱۴۰۰/۰۱/۲۴
اصلاح	<b>یافته‌ها:</b> تحلیل ادراک دانشجویان با استفاده از روش ریکور و در سه مرحله ساده‌خوانی، تحلیل ساختاری و درک جامع انجام شد. استخراج واحدهای معنایی به شناسایی ۲۰ زیرمضمون منجر شد که در ۵ مضمون اصلی «عدالت آموزشی»، «عدالت رویه‌ای»، «ماهیت سنجش»، «عدالت تعاملی» و «عدالت در آزمون کتبی» دسته‌بندی شد. نتایج حاصل از تحلیل سلسله‌مراتبی فازی نشان داد که در بین ۵ مضمون اصلی، «عدالت آموزشی» با وزن ۰/۳۱۸ در جایگاه اول قرار داشت و بعد از آن به ترتیب مضمون‌های «عدالت رویه‌ای»، «ماهیت سنجش»، «عدالت تعاملی» و «عدالت در آزمون کتبی» قرار داشتند. وزن‌های نهایی زیرمضمون‌ها از ضرب وزن مضمون‌های اصلی در وزن نسبی زیرمضمون‌ها به‌دست‌آمد؛ بر این اساس زیرمضمون «کیفیت تدریس» با وزن ۰/۲۲ در جایگاه اول و زیرمضمون‌های «عدالت اطلاعاتی» و «حرفه‌ای‌گری در تدریس» به ترتیب با وزن‌های ۰/۱۱ و ۰/۰۹ در جایگاه‌های بعدی قرار گرفتند.
پذیرش	۱۴۰۱/۰۴/۲۱
انتشار	۱۴۰۱/۰۵/۰۵
نتیجه‌گیری:	ادراک دانشجویان از عدالت در سنجش، ترکیبی از مفاهیم نظریه‌های اندازه‌گیری و عدالت سازمانی است و وزن بالای مضمون «عدالت آموزشی» در بین مضمون‌ها و زیرمضمون‌ها کیفیت تدریس این مضمون در بین تمام زیرمضمون‌های به‌دست آمده نشان داد که نقض عدالت آموزشی در بُعد کیفیت تدریس به درک ناعادلانه سنجش منجر خواهد شد و در مقابل اگر افراد احساس کنند تدریس استاد از کیفیت مطلوبی برخوردار است و رویه‌های مورد استفاده برای تخصیص نتایج عادلانه است حتی اگر برای آن‌ها مطلوب نباشد، نتایج را می‌پذیرند و آن را عادلانه‌تر درک می‌کنند.
انتشار	۱۴۰۱/۰۵/۲۳

## واژه‌های کلیدی: انصاف، پدیدارشناسی، سنجش کلاسی، عدالت، عدالت‌دارک‌شده

**استناد:** بنی‌اسدی، علی؛ صالحی، کیوان؛ خدایی، ابراهیم؛ باقری، خسرو و ایزانلو، بلال. (۱۴۰۱). شناسایی و اولویت‌بندی مضمون‌های سنجش کلاسی عادلانه با استفاده از روش فرایند تحلیل سلسله‌مراتبی فازی. *مطالعات اندازه‌گیری و ارزشیابی آموزشی*، ۱۲ (شماره ۳۷)، ۱۹-صفحه.

DOI: 10.22034/EMES.2022.527579.2241



حق مؤلف © نویسندگان.

ناشر: سازمان سنجش آموزش کشور

## مقدمه

روایی، پایایی و عدالت کیفیت‌هایی هستند که در نظریه اندازه‌گیری برای سنجش آموزشی بهینه به آن‌ها اشاره شده است (تیرنی<sup>۱</sup>، ۲۰۱۳). لازم است این نقطه را به خاطر داشته باشیم که فرضیه‌های نظریه اندازه‌گیری با سنجش کلاسی به‌طور کامل برازش ندارند؛ چراکه سنجش کلاسی وابسته به محیط آموزشی است که در یک زمینه روانی اجتماعی انجام می‌شود (بروخارت<sup>۲</sup>، ۲۰۰۳). چنین محیط ارزیابی با توجه به رویکرد معلم در سنجش که کربی<sup>۳</sup> (۱۹۹۹) آن را «فرهنگ ارزشیابی کلاس‌ها»<sup>۴</sup> خواند، طیف وسیعی از محیط‌ها را از یادگیرنده-محور تا معلم-محور در برمی‌گیرد (به نقل از بروخارت، ۲۰۰۳)؛ بنابراین، سنجش کلاسی فراتر از کاربرد فنون قابل‌اندازه‌گیری بوده و شامل نگرانی‌های فردی و اجتماعی است (اللال<sup>۵</sup>، ۲۰۱۳) و تعامل معلم و یادگیرنده و تأثیرات اجتماعی ممکن است روایی اطلاعات سنجش را کاهش داده و در نتیجه عدالت نمره تعیین‌شده توسط استاد را تحت تأثیر قرار دهد (آلم و کلنرود<sup>۶</sup>، ۲۰۱۵). با در نظر گرفتن این تفاوت زمینه‌ای، استیگینز<sup>۷</sup> (۲۰۰۱) به‌جای روایی و پایایی از اصطلاح «سنجش کلاس با کیفیت بالا»<sup>۸</sup> استفاده می‌کند (به نقل از بروخارت، ۲۰۰۳).

موقعیت‌های آموزشی، پیچیده است و چالش‌های متعددی را شامل می‌شود، یکی از آن‌ها مناقشه بین استاد و یادگیرنده است به‌طوری‌که رفتار استاد منبع مهم این مناقشه است و یک جنبه مهم این تعارض ادراک بی‌عدالتی است. این تعارض‌ها از طریق درک متقابل، حل مسئله، بهبود و تقویت روابط استاد و دانشجو و در نهایت رضایت از فرایندها و پیامدها حل می‌شود (جوالدین و ریسلایت<sup>۹</sup>، ۲۰۱۶). پیشرفت یادگیرندگان در یادگیری به کیفیت اساتید و معلمان بستگی دارد که در سه حوزه برنامه‌درسی، طراحی آموزشی و سنجش، تصمیم‌گیری می‌کنند و چنانچه هر یک از این حوزه‌ها نادیده انگاشته شود تصمیمات به‌اندازه کافی مؤثر نخواهد بود. چنانچه مک‌للان<sup>۱۰</sup> (۲۰۰۴) بیان می‌کند سواد سنجشی ناکافی اساتید در زمینه‌های روایی، پایایی و عدالت، پیامدهای منفی روی یادگیری دانشجویان خواهد داشت (به نقل از دلوکا<sup>۱۱</sup>، ۲۰۱۲) و با توجه به اینکه حدود ۵۰ درصد زمان تدریسشان را درگیر فعالیت‌های مرتبط با سنجش‌اند باید به‌طور پیوسته، توانایی سنجشی‌شان را افزایش دهند (یامتیم و ونگ وانگ<sup>۱۲</sup>، ۲۰۱۴).

برداشت از عدالت می‌تواند تحت تأثیر باورها و ارزش‌های فرهنگی باشد. تجربه‌ای که در یک کشور عادلانه تلقی می‌شود ممکن است در کشور دیگری آن‌گونه نباشد. در پژوهش تاتا<sup>۱۳</sup> (۲۰۰۵) نشان داده شد که فرهنگ (فردگرایی - جمع‌گرایی) بر دو بعد عدالت رویه‌ای، از جمله صدا (درواقع به معنای اظهارنظر دانشجویان در مورد سنجش و موضوعات سنجشی است) و عدالت بین فردی<sup>۱۵</sup> تأثیر می‌گذارد و برداشت دانشجویان چینی از عدالت بیشتر تحت تأثیر عدالت بین فردی است، درحالی‌که برداشت دانشجویان آمریکایی از عدالت بیشتر تحت تأثیر صدا قرار می‌گیرد. بافتی بودن موضوعات اخلاقی و عدالت در سنجش در پژوهش انجام‌شده توسط لیو<sup>۱۶</sup> و همکاران (۲۰۱۶) نیز نشان داده شده است به‌طوری‌که در این پژوهش که با عنوان «مطالعه مقایسه‌ای ادراک معلمان چینی و آمریکایی درباره موضوعات اخلاقی در کلاس درس» و با طراحی ۳۶ سناریو که عمل یا اندیشه‌ی معلم در ارتباط با شیوه‌های سنجش را نشان می‌دادند انجام شد نتایج حاصل از تحلیل‌های ۱۷۳ معلم آمریکایی و ۱۷۴ معلم چینی نشان داد که نظر معلمان دو کشور در ۲۲ سناریو از ۳۶ تا به‌طور معنی‌داری متفاوت بود (اندازه اثر بیشتر از ۰/۳) و از هشت سناریوی مربوط به عدالت/ سوگیری در ۳ سناریو تفاوت‌ها معنادار بودند.

- 
1. Tierney
  2. Brookhart
  3. Kirby
  4. the evaluative culture of classrooms
  5. Allal
  6. Alm & Colnerud
  7. Stiggins
  8. high-quality classroom assessment
  9. Čiuladienė & Račelytė
  10. MacLellan
  11. DeLuca
  12. Yamtim & Wongwanich
  13. Tata
  14. Voice
  15. interpersonal fairness
  16. Liu

تنها پژوهشی که در ایران در حوزه ادراک دانشجویان از سنجش کلاسی انجام شده است پژوهشی کمی است که مبتنی بر پرسشنامه ادراک فراگیران از سنجش که روی چهار عامل مختلف «بهبود»، «تأثیر»، «مرتبط نبودن» و «اسنادهای بیرونی» تمرکز می‌کند و توسط برون<sup>۱</sup> و همکاران (۲۰۱۴) با عنوان «ادراک دانشجویان دانشگاه‌های ایران درباره سنجش: کاربرد سنجش برای بهبود خود» انجام شده است. شرکت کنندگان ۷۶۰ دانشجو (۲۲۶ مرد و ۵۳۴ زن، ۶ تا ۳۰ ساله) با دامنه سنی ۱۸ تا ۳۸ سال بودند. همانند پژوهش سولیمونیدو و میچیلایدز<sup>۲</sup> (۲۰۱۴) و تلما<sup>۳</sup> و همکاران (۲۰۱۱)، دانشجویان سنجش را به‌عنوان عامل «بهبود یادگیری» بیش‌تر از عاملی برای «بهبود تدریس اساتید» تأیید کردند که نشان می‌دهد دانشجویان اصل مسئولیت‌پذیری و تلاش راهبردی برای یادگیری را در پاسخ به سنجش پذیرفته‌اند. اسنادهای بیرونی تأیید متوسطی را با تأکید بیشتر بر ارزیابی به‌عنوان شاخص کیفیت مدرسه تا به‌عنوان پیش‌بینی کننده آینده فردی به دست آوردند. «مرتبط نبودن» و «تأثیر»، میانگین نمره زیر ۳ به دست آوردند که تأیید ضعیفی را نشان می‌دهد. هرچند موافقت کمی با این دیدگاه که سنجش بد است (میانگین = ۳/۱۳، زیرمقیاسی از «مرتبط نبودن»، شامل آیتم‌های: سنجش، بی‌ارزش است، سنجش برای دانشجویان ناعادلانه است و نتایج سنجش از دقت برخوردار نیستند) داشتند. نکته‌ای که باید مدنظر قرار گیرد آن است که پرسشنامه ادراک دانش‌آموزان از سنجش مبتنی بر نسخه معلم برای دانش‌آموزان در دوره متوسطه نیوزند و نه برای دانشجویان تدوین یافته است.

سایر پژوهش‌ها، چالش‌های ارزشیابی دانشجویان را مورد بررسی قرار داده‌اند که می‌توانیم به پژوهش سراجی و همکاران (۱۳۹۲) اشاره کنیم. تحلیل داده‌های کیفی حاصل از مصاحبه با ۱۲ نفر از اساتید نشان داد که عواملی مانند دانش و مهارت ناکافی استادان درباره ارزشیابی، عدم امکان استفاده از شیوه‌های ارزشیابی نوین، پیامدهای منفی ناشی از تأکید بیش‌ازحد بر ارزشیابی پایانی، عدم توجه به ارزشیابی مستمر، عدم توجه به اهداف متفاوت و حیطه‌های یادگیری و غفلت از توجه به تفاوت‌های فردی، از چالش‌های مهم ارزشیابی از آموخته‌های دانشجویان به حساب می‌آیند. همچنین قنبری و همکاران (۱۳۹۴) در پژوهش خود با عنوان «تأثیر چالش‌های ارزشیابی آموخته‌های دانشجویان بر رویکرد مطالعه عمل فکورانه» که از پرسشنامه شناسایی چالش‌های ارزشیابی دانشجویان استفاده کردند نشان دادند که از بین شش چالش تأکید بیش‌ازحد بر ارزشیابی پایانی، عدم استفاده از ارزشیابی فرایندی و مستمر، عدم استفاده از ارزشیابی‌های غیر عینی، عدم توجه به تفاوت‌های فردی، عدم توجه به مجموعه هدف‌های یادگیری و عدم آشنایی اساتید با فلسفه و هدف ارزشیابی چالش مربوط به عدم توجه به تفاوت‌های فردی بیشترین و چالش مربوط به عدم استفاده از ارزشیابی‌های غیرعینی کمترین بار عاملی را داشتند. همچنین نتیجه پژوهش پورنادری و همکاران (۱۳۹۲) که با عنوان «بررسی استفاده از روش‌های مختلف سنجش و سنجش‌های مبتنی بر یادگیری پایدار» انجام شد و در آن از پرسشنامه محقق ساخته شامل سه بخش انواع سنجش، دقت و صحت سنجش و یادگیری‌های پایدار استفاده کردند نشان داد که روش‌های سنجش موجود غیرکاربردی، تراکمی، محدود و تا حد زیادی مبتنی بر امتحانات پایان‌ترم است و خودسنجی و هم‌تاسنجی در آن جایی ندارد. همچنین اقدامات سنجش در جهت افزایش یادگیری‌های دانشجویان طرح‌ریزی نمی‌شوند و برای اندازه‌گیری آموخته‌ها از دقت و صحت کافی برخوردار نیستند.

باور به سفید و سیاه‌ها، صفر و یک‌ها و به‌طور کلی نظام دو ارزشی به گذشته برمی‌گردد و به یونان قدیم و ارسطو می‌رسد. البته قبل از ارسطو نوعی ذهنیت فلسفی وجود داشت که به ایمان دودویی با شک و تردید می‌نگریست. بودا در هند، پنج قرن قبل از مسیح و تقریباً دو قرن قبل از ارسطو زندگی می‌کرد. اولین قدم در نظام اعتقادی او گریز از جهان سیاه‌وسفید و برداشتن این حجاب دو ارزشی بود؛ یعنی نگرستن به جهان به صورتی که هست. از دید بودا، جهان را باید سراسر تناقض دید، جهانی که چیزها و ناچیزها در آن وجود دارد. در آن گل‌های رز هم سرخ هستند و هم غیر سرخ (کاسکو، ۱۳۸۹، ص. ۲۵). در دنیای ما غالب چیزهایی که درست به نظر می‌رسند «نسبتاً» درست هستند و در باب صحت و سقم پدیده‌های واقعی همواره درجاتی از «عدم قطعیت» صدق می‌کند به عبارت دیگر پدیده‌های واقعی صرفاً سیاه یا سفید نیستند بلکه تاندازه‌های «خاکستری» هستند. منطق فازی جهان‌بینی جدیدی است که با وجود ریشه داشتن در فرهنگ مشرق زمین با نیازهای دنیای پیچیده امروز بسیار سازگارتر از منطق ارسطویی است. منطق فازی جهان را آن‌طور که هست به تصویر می‌کشد. رویکرد فازی با فراهم آوردن امکان رده‌بندی ویژگی<sup>۴</sup>، سنجش انطباق عناصر با مفاهیم متعدد را ممکن می‌سازد که کاربرد اصلی آن ساخت و سنجش مفاهیم ترکیبی بر اساس وجود درجاتی از صدق چند معنا در یک مصداق است. تمایز مهمی میان درجه انطباق مصادیق با یک مفهوم و درجه انطباق آن‌ها با چند مفهوم وجود دارد. در حالت اول مقایسه مصادیق در درون یک مجموعه معنایی و بر اساس اصل بیشتر بودن یا کمتر بودن انطباق عنصر x در مقایسه با y با مفهوم A صورت

1. Brown
2. Solomonidou & Michaelides
3. Tillema
4. Property ranking

می‌گیرد درحالی‌که رده‌بندی ویژگی بر اساس مقایسه بین مصادیق مشترک انجام می‌شود و اینکه عنصر  $x$  بیشتر با مفهوم  $A$  تطابق دارد یا  $B$  (ساروخانی و صادقی پور، ۱۳۹۲).

با توجه به مطالب فوق و اینکه تاکنون در پژوهش‌های حوزه سنجش کلاسی در ایران پژوهشی جهت شناسایی مضمون‌های سنجش کلاسی عادلانه انجام نشده است و با توجه به اینکه عدالت مفهومی اخلاقی و اجتماعی است که تمام عواملی که در ایجاد اخلاق اجتماعی یک ملت مؤثرند (اعتقادات، نیازها، عوامل اقتصادی، اجتماعی، تاریخی و جغرافیایی) در به‌وجود آمدن عدالت تأثیر دارند (کاتوزیان، ۱۳۸۶) و از طرفی در باب مفاهیم اجتماعی با درجاتی از انتزاع و انضمام روبه‌رو هستیم که سبب‌ساز وجود نوعی ابهام در صدق معنا به مصداق می‌شود؛ به عبارت دیگر، به‌جای حالات دوگانه "صدق" / "عدم صدق" با حالات چندگانه "درجات صدق" مواجهیم که می‌توانند در موقعیت‌هایی بین حالات دوگانه مزبور قرار گیرند و رویکرد مبتنی بر منطق فازی در ساخت و سنجش این مفاهیم می‌تواند مؤثر واقع شود (ساروخانی و صادقی پور، ۱۳۹۲) در پژوهش حاضر به دنبال آن هستیم تا به سؤال‌های زیر پاسخ دهیم:

- ۱- مضمون‌های سنجش کلاسی عادلانه از نظر دانشجویان چیستند؟
- ۲- اولویت و وزن هر یک از مضمون‌ها و زیرمضمون‌های شناسایی شده براساس نظر متخصصان و با کاربرد رویکرد سلسه‌مراتبی تحلیلی فازی<sup>۱</sup> نسبت به یک‌دیگر چگونه است؟

### روش پژوهش

پژوهش حاضر در دو بخش زیر انجام شده است که در هر بخش روش، نمونه و ابزار گردآوری داده‌ها به تفکیک بیان شده است.

الف- برای به‌دست‌آوردن مضمون‌ها و زیرمضمون‌های سنجش کلاسی عادلانه و ترسیم درخت سلسه‌مراتبی، روش پدیدارشناسی مورد استفاده قرار گرفت؛ چراکه این روش با هدف پی بردن به معنای ضمنی یک پدیده از دید کسانی که آن پدیده را تجربه کرده‌اند، انجام می‌شود (بازرگان، ۱۳۹۱). جامعه پژوهش حاضر همه دانشجویان دانشگاه تهران بودند. از نمونه‌گیری ملاکی<sup>۲</sup> برای انتخاب دانشجویان استفاده شد؛ از این‌رو با توجه به اینکه ادراک دانشجویان از عدالت در سنجش باید موردبررسی قرار می‌گرفت، دانشجویان به‌گونه‌ای انتخاب شدند که تجربه حداقل دو ترم سنجش در دانشگاه تهران داشته باشند. از طرف دیگر، به‌دلیل طیف گسترده رشته‌ها و دانشکده‌های دانشگاه تهران، به‌منظور پوشش تجارب و دیدگاه‌های متنوع‌تر، دانشکده‌ها در ۵ حوزه علوم انسانی، علوم پایه، فنی و مهندسی، کشاورزی و هنر طبقه‌بندی شدند و با دانشجویان مصاحبه شد. روش استفاده‌شده برای جمع‌آوری داده‌ها مصاحبه‌های نیمه‌ساختاریافته<sup>۳</sup> بود. مصاحبه‌های نیمه ساختاری قالبی از مصاحبه است که به‌طور گسترده برای تحقیقات کیفی مورد استفاده قرار می‌گیرد و می‌تواند به‌صورت جداگانه یا گروهی انجام شود. مصاحبه‌های گروهی اغلب در مطالعات گروه‌های کانونی مورد استفاده قرار می‌گیرد که در آن چندین شرکت‌کننده، دانش و تجربیات خود را در مورد یک موضوع خاص به اشتراک می‌گذارند (دیکوبلوم و کرابتری<sup>۴</sup>، ۲۰۰۶ به نقل از بنی‌اسدی و صالحی، ۱۳۹۸). نتایج کامل این بخش در پژوهش بنی‌اسدی وهمکاران (۱۴۰۰) قابل دسترسی است.

در مورد تولید داده در روش پدیدارشناسی، اسپزیال<sup>۵</sup> و همکاران (۲۰۱۱) اشاره کرده‌اند که تولید یا جمع‌آوری داده‌ها تا زمانی که پژوهشگر به اشباع داده‌ها نرسد ادامه می‌یابد. در این پژوهش اشباع موضوعی استقرایی<sup>۶</sup> بعد از ۲۹ مصاحبه فردی، یک مصاحبه گروهی و یک گروه کانونی به‌دست آمد. ویژگی‌های توصیفی مشارکت‌کنندگان در جدول ۱ ارائه شده است. پس از حصول به اشباع نظری یافته‌ها، روایی محتوایی مضمون‌ها و زیرمضمون‌های به‌دست‌آمده توسط نه استاد متخصص در حوزه سنجش تأیید شد. همچنین مضمون‌ها و زیرمضمون‌های به‌دست‌آمده در قالب درخت سلسه‌مراتبی که شامل ۵ مضمون و ۲۰ زیرمضمون است ارائه شده‌اند (شکل ۱).

ب- در این بخش که هدف وزن‌دهی و اولویت‌بندی مضمون‌ها و زیرمضمون‌های به‌دست‌آمده از بخش الف است روش پژوهش، توصیفی- تحلیلی است. روش‌های وزن‌دهی به دو دسته روش‌های عینی و ذهنی تقسیم‌بندی می‌شوند. در فرایند سلسه‌مراتبی تحلیلی که نمونه‌ای از روش‌های ذهنی است و ساتی<sup>۷</sup> (۱۹۷۷) آن را پیشنهاد داده است (الدیان و تیلور<sup>۸</sup>، ۲۰۰۵) مشکلات مربوط به تصمیم‌گیری‌های پیچیده از طریق تعیین اهمیت

1. Fuzzy Analytical Hierarchy Process (FAHP)

2. criterion sampling

3. Semi- structured interviews

4. DiCicco-Bloom & Crabtree

5. Speziale

6. inductive thematic saturation

7. Saaty

8. Aldian & Taylor

نسبی مضمون‌های دخیل در یک تصمیم‌گیری حل می‌شوند و یک مسئله تصمیم‌گیری چندملاکی به مجموعه‌ای از تصمیمات مرتبط تجزیه می‌شود (وینود<sup>۱</sup> و همکاران، ۲۰۱۴).

جدول ۱. ویژگی‌های توصیفی مشارکت‌کنندگان (دانشجویان)

شماره	شماره ترم	رشته تحصیلی	جنسیت	زمان مصاحبه	شماره	شماره ترم	رشته تحصیلی	جنسیت	زمان مصاحبه
۱	۳ ارشد	اکویدرولوژی	مرد	۶۰ دقیقه	۱۶	۴ ارشد	ژئومورفولوژی	زن	۵۰ دقیقه
۲	۴ ارشد	آمار	زن	۵۵ دقیقه	۱۷	فارغ-التحصیل کارشناسی	مدیریت	مرد	کتابی
۳	۳ دکتری	هوافضا	مرد	۶۰ دقیقه	۱۸	۳ کارشناسی	شهرسازی	مرد	کتابی
۴	۷ کارشناسی	زیست‌شناسی سلولی مولکولی	زن	۴۰ دقیقه	۱۹	۳ ارشد	مدیریت شهری	مرد	۴۰ دقیقه
۵	۴ ارشد	شیمی عالی	زن	کتابی	۲۰	۳ ارشد	تصویرسازی	زن	۴۵ دقیقه
۶	۱۲ دکتری	روانشناسی	زن	کتابی	۲۱	فارغ-التحصیل کارشناسی	معماری	زن	کتابی
۷	فارغ-التحصیل ارشد	برنامه‌ریزی آموزشی	مرد	۵۵ دقیقه	۲۲	۱۲ دکتری	عمران سازه	مرد	۴۰ دقیقه
۸	۸ دکتری	بیوفیزیک	زن	۵۰ دقیقه	۲۳	۵ ارشد	مدیریت شهری	مرد	کتابی
۹	۳ ارشد	روانشناسی کودکان استثنایی	زن	کتابی	۲۴	۷ دکتری	حقوق	مرد	۳۰ دقیقه
۱۰	۳ کارشناسی	آمار	زن	۴۰ دقیقه	۲۵	۵ کارشناسی	زمین‌شناسی	مرد	۳۵ دقیقه
۱۱	۳ کارشناسی	انسان‌شناسی	زن	۴۵ دقیقه	۲۶	۹ کارشناسی	مردم‌شناسی	مرد	۴۰ دقیقه
۱۲	۳ کارشناسی	مهندسی مخابرات	زن	۵۰ دقیقه	۲۷	۵ کارشناسی	معماری	مرد	۵۵ دقیقه
۱۳	۷ کارشناسی	روانشناسی	زن	۴۵ دقیقه	۲۸	۵ کارشناسی	مکانیک	مرد	۴۰ دقیقه
۱۴	۷ کارشناسی	برق	مرد	۴۰ دقیقه	۲۹	۳ ارشد	صنایع سلولزی	زن	۴۰ دقیقه
۱۵	۷ دکتری	سنجش	مرد	۵۰ دقیقه					

تعیین وزن در روش‌های ذهنی بر اساس نظر متخصصانی است که در زمینه مورد نظر دانش و تجربه دارند؛ از این رو تحلیل‌گر برای به دست آوردن قضاوت‌های ذهنی معمولاً مجموعه‌ای از سؤال‌ها را در اختیار تصمیم‌گیرندگان قرار می‌دهد (آلدیان و تیلور، ۲۰۰۵) و پرسشنامه طراحی شده بر اساس مقایسه‌های زوجی از مؤثرترین و معتبرترین روش‌های جمع‌آوری داده‌ها در این روش است (آذر، ۱۳۸۸). بر این اساس در پژوهش حاضر پرسشنامه مقایسه‌ی زوجی که شامل ۵۳ مقایسه بود طراحی شد و برای ۱۳ نفر از متخصصین حوزه سنجش ارسال شد که در نهایت پاسخ‌های نه نفر از آن‌ها دریافت شد.

<sup>1</sup>. Vinodh

مدل AHP اولیه نقایصی دارد؛ به‌عنوان مثال عدم قطعیت مرتبط با فرایندها را دربر نمی‌گیرد. برای غلبه بر مشکلات این چینی نظریه فازی با AHP در قالب AHP فازی ادغام شده است (وینود و همکاران، ۲۰۱۴). روش AHP فازی این پژوهش برگرفته از روش میانگین هندسی باکلی<sup>۱</sup> می‌باشد (هاسی<sup>۲</sup> و همکاران، ۲۰۰۴). گام‌های این روش در زیر آورده شده است.

فرض کنید  $\tilde{P}_i$  مجموعه‌ای از ترجیحات تصمیم‌گیران در مورد یک شاخص نسبت به دیگر شاخص‌ها باشد. ماتریس مقایسه‌های زوجی به صورت زیر تشکیل می‌شود:

$$\tilde{A} = \begin{bmatrix} 1 & \tilde{P}_{12} & \tilde{P}_{1n} \\ \tilde{P}_{21} & 1 & \tilde{P}_{2n} \\ \tilde{P}_{n1} & \tilde{P}_{n2} & 1 \end{bmatrix}$$

که  $n$  تعداد عناصر مرتبط در هر سطر است. وزن‌های فازی هر شاخص ماتریس مقایسه‌های زوجی به وسیله روش میانگین هندسی باکلی به دست می‌آید (هاسی و همکاران، ۲۰۰۴). میانگین هندسی ارزش مقایسه‌های فازی شاخص  $i$  به هر شاخص از رابطه زیر به دست می‌آید.

$$\tilde{r}_i = \left( \prod_{j=1}^n \tilde{P}_{ij} \right)^{1/n} \quad i = 1, 2, 3, \dots, n \quad (1)$$

سپس وزن فازی  $i$  مین شاخص به وسیله یک عدد فازی مثلثی نشان داده می‌شود.

$$w_i = r_i \otimes (r_1 \oplus r_2 \oplus \dots \oplus r_m)^{-1} \quad (2)$$

بعد از محاسبه فاکتورهای وزن فازی، به‌وسیله فرمول زیر وزن‌ها را دیفازی کرده و سپس نرمال می‌کنیم.

$$w_{crisp} = \frac{l + 2m + u}{4} \quad (3)$$

در این پژوهش جهت محاسبه وزن در مقایسه‌های زوجی، از عبارات کلامی و اعداد فازی مثلثی مندرج در جدول استفاده شده است. تعیین وزن ملاک‌ها با کاربرد داده‌های کیفی می‌تواند تحت تأثیر اولویت تصمیم‌گیرنده باشد به همین دلیل ساتی (۱۹۷۷) مقیاس عددی ۹-۱ را برای تبدیل داده‌های کیفی به کمی که در آن ۱ اهمیت برابر و ۹ اهمیت فوق‌العاده را نشان می‌دهد پیشنهاد کرد.

جدول ۲. عبارات کلامی و اعداد فازی جهت وزندهی به مضمونها

کد	اولویت‌ها	معادل فازی اولویت‌ها		
		حد پایین (L)	حد متوسط (m)	حد بالا (u)
۱	اهمیت یکسان	۱	۱	۱
۲	یکسان تا نسبتاً مهم‌تر	۱	۲	۳
۳	نسبتاً مهم‌تر	۲	۳	۴
۴	نسبتاً مهم‌تر تا اهمیت زیاد	۳	۴	۵
۵	اهمیت زیاد	۴	۵	۶
۶	اهمیت زیاد تا بسیار زیاد	۵	۶	۷
۷	اهمیت بسیار زیاد	۶	۷	۸
۸	بسیار زیاد تا کاملاً مهم‌تر	۷	۸	۹
۹	کاملاً مهم‌تر	۸	۹	۱۰

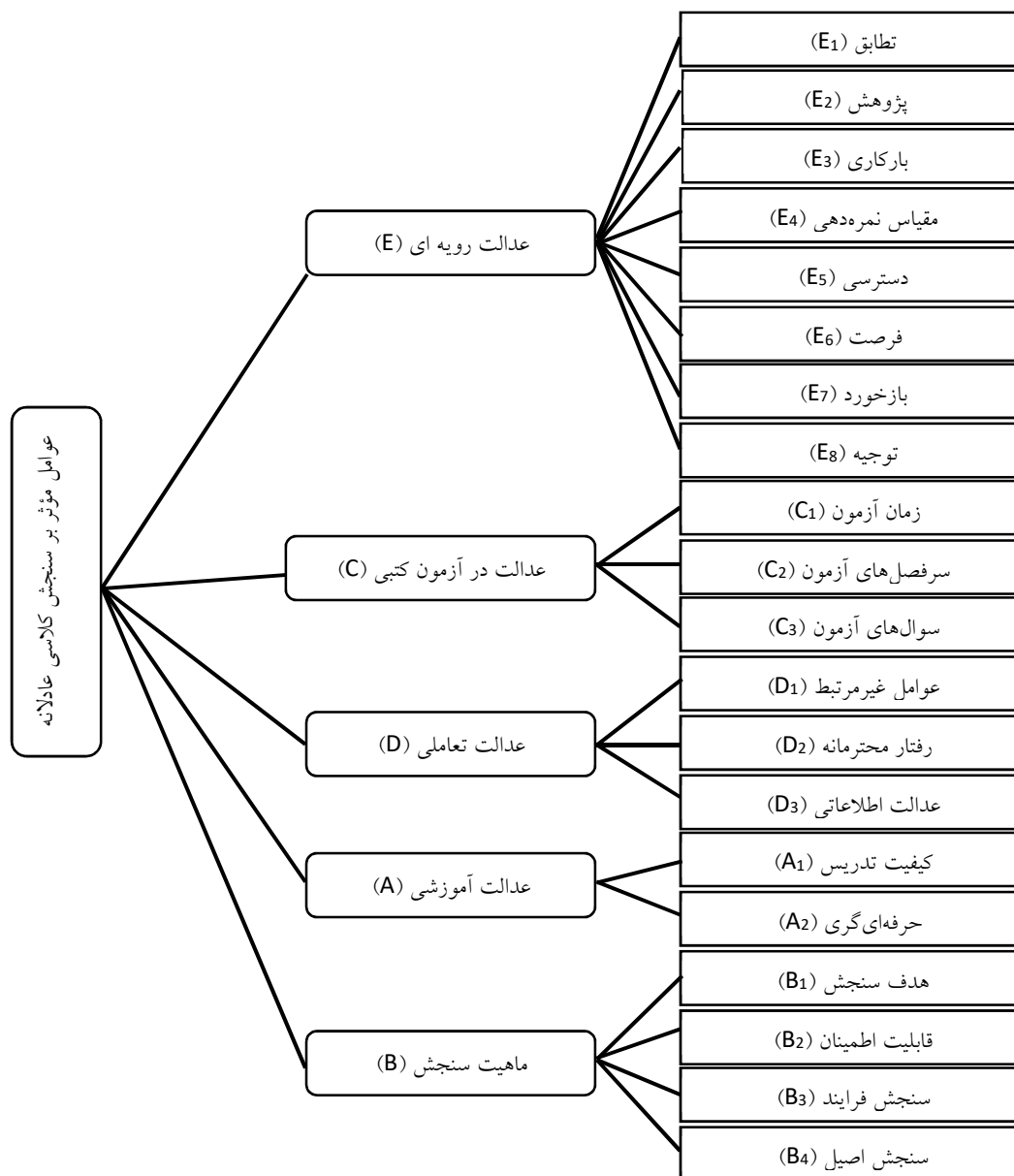
1. buckley

2. Hsieh

## یافته‌ها

## مضمون‌ها و زیرمضمون‌ها

متن‌های پیاده‌سازی شدهٔ مصاحبه‌ها با استفاده از روش پدیدارشناسی، اقتباس شده از ریکور<sup>۱</sup> (۱۹۷۶) که شامل ساده خواندن، تحلیل ساختاری و درک جامع است؛ تحلیل شدند. نتایج این بخش یعنی مضمون‌ها و زیرمضمون‌های سنجش کلاسی عادلانه در درخت سلسه‌مراتبی ارائه شده است (شکل ۱ را ببینید).



شکل ۱. درخت تحلیل سلسه‌مراتبی مضمون‌ها و زیرمضمون‌های سنجش کلاسی عادلانه

<sup>1</sup>. Ricoeur



همچنین برای محاسبه‌ی روایی محتوایی از گروهی از متخصصان خواسته می‌شود تا هر مضمون را از نظر ارتباط با سازه‌ی زیربنایی درجه‌بندی کنند. با توجه به اینکه روایی محتوایی شاخصی از توافق ارزیابان و متخصصان است و این توافق با پاسخ تصادفی و شانسی متورم می‌شود لین<sup>۱</sup> (۱۹۸۶) حداقل سه متخصص را توصیه می‌کند اما بیش از ۱۰ نفر را ضروری نمی‌داند (ذکر شده در پالت و بک<sup>۲</sup>، ۲۰۰۶). بر این اساس متخصصانی که ما برای رواسازی مضمون‌های به‌دست‌آمده استفاده کردیم شامل نه استاد دانشگاه بودند. ملاک انتخاب این افراد تخصص در حوزه‌ی سنجش و تجربه تدریس حداقل ۵ سال در دانشگاه بود. ما ابتدا از متخصصان خواستیم تا میزان ارتباط هر مضمون با سازه‌ی "سنجش کلاسی عادلانه" را در یک طیف لیکرت چهارنقطه‌ای، "کاملاً مرتبط"، "بسیار مرتبط"، "تا حدودی مرتبط" و "مرتبط نیست" درجه‌بندی کنند. انتخاب مقیاس رتبه‌ای چهار نقطه‌ای به‌جای مقیاس درجه‌بندی ۳ یا ۵ نقطه‌ای برای اجتناب از داشتن نقطه‌ی میانی خنثی و دوسویه بود. برای هر مضمون شاخص روایی محتوایی<sup>۳</sup> با تقسیم تعداد متخصصانی که به رتبه ۳ یا ۴ داده‌اند بر تعداد کل متخصصان محاسبه می‌شود. برای اینکه خطای استاندارد نسبت در محاسبات وارد شود لین (۱۹۸۶) پیشنهاد می‌کند با داشتن ۵ متخصص یا کمتر همه باید در مورد روایی محتوا به توافق برسند به‌عبارت‌دیگر شاخص روایی محتوایی آیتم باید برابر ۱ باشد و ۶ متخصص یا بیش‌تر شاخص روایی محتوایی کمتر از ۰/۷۸ را توصیه نمی‌کند (ذکر شده در پالت و بک<sup>۴</sup>، ۲۰۰۶). در این پژوهش شاخص روایی محتوایی کلی به توصیه‌ی پالت و بک (۲۰۰۶) میانگین شاخص‌های روایی محتوایی زیرمضمون‌ها در نظر گرفته شد؛ چراکه به‌جای میانگین عملکرد کارشناسان بر میانگین کیفیت زیرمضمون‌ها متمرکز است. بسیاری از پژوهشگران مقدار قابل قبول روایی محتوایی کلی را ۰/۸۰ و بالاتر از آن می‌دانند.

هر نه کارشناس زیرمضمون‌های بازخورد، فرصت، دسترسی، پژوهش‌های تخصیص داده‌شده به دانشجویان، بارکاری، تطابق، کیفیت تدریس، سنجش فرایند، قابلیت اطمینان، عوامل غیر مرتبط، عدالت اطلاعاتی، رفتار محترمانه، سرفصل‌های آزمون، سؤال‌های آزمون و زمان آزمون را کاملاً مرتبط یا بسیار مرتبط با مفهوم سنجش کلاسی عادلانه دانستند. با توجه به تعریف شاخص روایی محتوایی هرکدام از این زیرمضمون‌ها برابر ۱ (۹÷۹) است. همچنین هر یک از زیرمضمون‌های توجیه، مقیاس نمره‌دهی، حرفه‌ای‌گری در تدریس، سنجش اصیل و هدف سنجش توسط هشت کارشناس مرتبط یا بسیار مرتبط تشخیص داده شد و شاخص روایی محتوایی هرکدام از این زیرمضمون‌ها برابر ۰/۸۸ (۸ ÷ ۹) است؛ بنابراین روایی محتوایی هر از زیرمضمون‌ها بالاتر از مقدار معیار ۰/۷۸ است. روایی محتوایی کلی نیز بنا به تعریف برابر ۰/۹۴ است [۲۰ ÷ (۸۸ × ۰/۸۰ + ۱ × ۱۵)]؛ که بالاتر از مقدار معیار ۰/۸۰ است.

### نتایج تحلیل سلسه‌مراتبی فازی

در مرحله‌ی قبل عوامل پژوهش شناسایی و تأیید شدند. در این گام ابتدا مقایسه‌های زوجی مضمون‌ها تشکیل شد و در اختیار نه پاسخ‌دهنده قرار داده شد. بعد از پاسخگویی به مقایسه‌های زوجی، نرخ ناسازگاری جداول محاسبه شد. همگی از ۰/۱ کوچک‌تر بود که نشان‌دهنده‌ی این است که ثبات و قابلیت اطمینان مقایسه‌های زوجی در حد قابل قبولی است. سپس با استفاده از روش میانگین هندسی پاسخ‌ها ادغام شدند. مقایسه‌های زوجی ادغام‌شده در ادامه آورده شده است. وزن‌های مقایسه‌های زوجی نیز با استفاده از روش میانگین هندسی باکلی محاسبه شد.

#### الف- تشکیل مقایسه‌های زوجی

در این بخش مقایسه‌های زوجی مضمون‌های اصلی و سپس زیرمضمون‌ها تشکیل شد این مقایسه‌های زوجی بر اساس طیف ۱ تا ۹ فازی انجام گرفته است. نتایج در جدول‌های ۳ تا ۸ ارائه شده است.

جدول ۳. مقایسه‌های زوجی مضمون‌ها (نرخ ناسازگاری: ۰/۰۸)

	A	B	C	D	E
A	(۱،۱)	(۱/۴۵۶، ۱/۹۵۶، ۲/۴۵۴)	(۲/۶۳۱، ۳/۷۶۹، ۴/۸۸۴)	(۱/۶۴، ۲/۱۱۳، ۲/۶۷۷)	(۱/۲۰۴، ۱/۵۱۳، ۱/۹۰۱)
B	(۰/۴۰۸، ۰/۵۱۱، ۰/۶۸۷)	(۱،۱)	(۳/۱۷۱، ۳/۹۶۴، ۴/۷۲۶)	(۰/۶۴۹، ۰/۷۵۲، ۰/۸۹۳)	(۲/۱۱۸، ۲/۶۸۳، ۳/۲۱۸)
C	(۰/۲۰۷، ۰/۲۶۵، ۰/۳۸)	(۰/۲۱۲، ۰/۲۵۲، ۰/۳۱۵)	(۱،۱)	(۰/۱۹۷، ۰/۲۳۲، ۰/۲۸۴)	(۰/۲۴۴، ۰/۳۰۳، ۰/۴۱۱)
D	(۰/۳۷۴، ۰/۴۷۳، ۰/۶۱)	(۱/۱۲، ۱/۳۲۹، ۱/۵۴۱)	(۳/۵۲، ۴/۳۱۱، ۵/۰۷)	(۱،۱)	(۳/۲۹۷، ۴/۳۵۸، ۵/۳۹۵)

1. Lynn

2. Polit & Beck

3. Content Validity Index (CVI)

4. Polit & Beck

E (۰/۵۲۶، ۰/۶۶۱، ۰/۸۳) (۰/۳۱۱، ۰/۳۷۳، ۰/۴۷۲) (۲/۴۳۴، ۳/۳۰۲، ۴/۱۰۳) (۰/۱۸۵، ۰/۲۲۹، ۰/۳۰۳) (۱،۱،۱)

جدول ۴. مقایسه‌های زوجی زیرمضمون‌های عدالت آموزشی (نرخ ناسازگاری: ۰/۰۰۰)

	A1	A2
A1	(۱،۱،۱)	(۱/۹۳۷، ۲/۳۷، ۲/۸۰۵)
A2	(۰/۳۵۶، ۰/۴۲۲، ۰/۵۱۶)	(۱،۱،۱)

جدول ۵. مقایسه‌های زوجی زیرمضمون‌های ماهیت سنجش (نرخ ناسازگاری: ۰/۰۶)

	B1	B2	B3	B4
B1	(۱،۱،۱)	(۰/۳۲۸، ۰/۴۰۳، ۰/۵۲۷)	(۰/۷۱۴، ۰/۸۴۶، ۱/۰۱۲)	(۰/۳۱۲، ۰/۳۸۵، ۰/۵۱۴)
B2	(۱/۸۹۸، ۲/۴۸۴، ۳/۰۴۸)	(۱،۱،۱)	(۰/۹۸۷، ۱/۱۴۹، ۱/۳۷۹)	(۰/۷۵۹، ۱/۰۵۴، ۱/۴۱۳)
B3	(۰/۹۸۸، ۱/۱۸۲، ۱/۴۰۱)	(۰/۷۲۵، ۰/۸۷، ۱/۰۱۳)	(۱،۱،۱)	(۰/۲۱۱، ۰/۲۷۱، ۰/۳۵۴)
B4	(۱/۹۴۶، ۲/۶۰۱، ۳/۲۰۱)	(۰/۷۰۸، ۰/۹۴۹، ۱/۳۱۸)	(۲/۸۲۷، ۳/۳۸۷، ۴/۷۴۱)	(۱،۱،۱)

جدول ۶. مقایسه‌های زوجی زیرمضمون‌های عدالت در آزمون کتبی (نرخ ناسازگاری: ۰/۰۳)

	C1	C2	C3
C1	(۱،۱،۱)	(۰/۲۳، ۰/۲۷۹، ۰/۳۶۱)	(۰/۱۸۸، ۰/۲۳۴، ۰/۳۰۹)
C2	(۲/۷۷، ۳/۵۸۲، ۴/۳۵۶)	(۱،۱،۱)	(۱/۱۶۷، ۱/۳۸۵، ۱/۶۶۴)
C3	(۳/۲۳۶، ۴/۲۸۲، ۵/۳۱۱)	(۰/۶۰۱، ۰/۷۲۲، ۰/۸۵۷)	(۱،۱،۱)

جدول ۷. مقایسه‌های زوجی زیرمضمون‌های عدالت تعاملی (نرخ ناسازگاری: ۰/۰۶)

	D1	D2	D3
D1	(۱،۱،۱)	(۰/۵۰۷، ۰/۶۶۴، ۰/۸۵۷)	(۰/۶۸۳، ۰/۸۶، ۱/۱۵)
D2	(۱/۱۶۷، ۱/۵۰۷، ۱/۹۷۲)	(۱،۱،۱)	(۰/۵۲۸، ۰/۶۱۵، ۰/۷۱۳)
D3	(۰/۸۷، ۱/۱۶۳، ۱/۴۶۴)	(۱/۴۰۳، ۱/۶۲۷، ۱/۸۹۵)	(۱،۱،۱)

جدول ۸. مقایسه های زوجی زیرمضمون های عدالت رویدای (نرخ ناسازگاری: ۰/۰۳)

E8	E7	E6	E5	E4	E3	E2	E1
(۰/۱۹۷، ۰/۳۳۳، ۰/۷۸۸)	(۰/۳۲۷، ۰/۳۹۸، ۰/۴۹)	(۰/۲۲۱، ۰/۲۷، ۰/۳۶۲)	(۰/۳۲۵، ۰/۲۷، ۰/۳۳۲)	(۰/۷۸۹، ۰/۴۹۲، ۰/۶۲)	(۰/۴۴۴، ۰/۵۲۱، ۰/۶۷۳)	(۰/۱۷۳، ۰/۲۱، ۰/۳۷۳)	(۱، ۱، ۱)
(۰/۴۵۸، ۰/۵۹۹، ۰/۷۹۹)	(۱/۱۰۷، ۱/۴۵۱، ۱/۸۹۵)	(۰/۶۰۸، ۰/۶۹۴، ۰/۷۹۹)	(۰/۵۲۳، ۰/۶۷، ۰/۷۷۴)	(۰/۵۳۹، ۰/۶۷۱، ۰/۸۱۹)	(۱، ۱/۳۳۹، ۱/۴۹۶)	(۱، ۱، ۱)	(۳/۶۸۲، ۴/۷۵۲، ۵/۷۹۵)
(۰/۴۱۶، ۰/۶۶۳، ۰/۹۳۱)	(۰/۵۰۷، ۰/۶۷۷، ۰/۸۷۵)	(۰/۴۱۲، ۰/۵۰۱، ۰/۶۳۴)	(۰/۴۴۵، ۰/۵۲۴، ۰/۶۱۹)	(۰/۶۸۹، ۰/۸۸۳، ۱/۱۰۴)	(۱، ۱، ۱)	(۰/۶۶۹، ۰/۸۰۷، ۰/۹۰۱)	(۱/۴۸۹، ۱/۹۱۹، ۲/۳۵۶)
(۰/۳۶، ۰/۴۶۸، ۰/۶۵۷)	(۱/۴۶۴، ۱/۷۴۱، ۲/۰۳۵)	(۰/۶۷۲، ۰/۸۶، ۱/۱۲۶)	(۰/۵۱۷، ۰/۶۴۲، ۰/۸۳۵)	(۱، ۱، ۱)	(۰، ۰، ۰)	(۱/۲۲، ۱/۴۹، ۱/۸۵۴)	(۱/۶۱۲، ۲/۰۳۵، ۲/۵۶۸)
(۰/۶۷۴، ۰/۸۵۳، ۱/۰۲)	(۰/۹۷۴، ۱/۳۳۲، ۱/۶۰۸)	(۰/۳۸۵، ۰/۴۵۲، ۰/۵۵۳)	(۱، ۱، ۱)	(۱/۱۹۸، ۱/۵۵۸، ۱/۹۲۴)	(۱/۶۱۶، ۱/۹۰۹، ۲/۳۴۹)	(۱/۳۹۳، ۱/۴۹۳، ۱/۷۴۴)	(۳/۱۱، ۳/۷۰۶، ۴/۷۵۹)
(۰/۳۰۷، ۰/۴۱۶، ۰/۶۳)	(۱/۴۳۸، ۱/۸۵۲، ۲/۳۴)	(۱، ۱، ۱)	(۱/۸۰۷، ۲/۲۱، ۲/۵۹۷)	(۰/۸۸۸، ۱/۱۶۲، ۱/۴۸۹)	(۱/۵۷۶، ۱/۹۹۷، ۲/۳۴۹)	(۱/۲۵۱، ۱/۴۴۲، ۱/۶۴۴)	(۲/۷۶۵، ۳/۶۹۷، ۴/۵۲۴)
(۰/۲۵۲، ۰/۳۲۶، ۰/۴۵۷)	(۱، ۱، ۱)	(۰/۴۴۷، ۰/۵۴، ۰/۶۹۶)	(۰/۶۲۲، ۰/۸۰۵، ۱/۰۲۷)	(۰/۴۹۲، ۰/۵۷۴، ۰/۶۸۳)	(۱/۱۴۴، ۱/۴۷۶، ۱/۹۷۳)	(۰/۵۲۸، ۰/۶۸۹، ۰/۹۰۳)	(۳/۵۶۵، ۳/۳۵۶، ۴/۲۱۹)
(۱، ۱، ۱)	(۲/۱۸۹، ۳/۰۶۴، ۳/۹۷۱)	(۱/۵۸۷، ۲/۴۰۶، ۳/۲۵۹)	(۰/۹۸، ۱/۱۷۳، ۱/۴۸۴)	(۱/۵۲۲، ۲/۱۳۵، ۳/۷۸)	(۱/۰۷۴، ۱/۵۰۹، ۲/۰۶۶)	(۱/۲۵۱، ۱/۶۷، ۲/۱۸۲)	(۳/۴۷۴، ۴/۲۹۴، ۵/۰۷)

## ب- محاسبه وزن‌های فازی و نرمال

در این گام بر اساس رابطه (۱) و (۲) ابتدا میانگین هندسی اعداد فازی هر سطر جدول‌های ۳ الی ۸ را محاسبه می‌کنیم و سپس هر میانگین هندسی حاصل را بر مجموع میانگین‌های هندسی تقسیم می‌کنیم تا وزن فازی حاصل شود سپس هر وزن فازی را با استفاده از رابطه  $(l + 2m + u)/4$  غیرفازی می‌کنیم و برای نرمال‌سازی هر وزن غیرفازی کافی است آن وزن را بر مجموع وزن‌های غیرفازی تقسیم کنیم؛ به‌عنوان مثال برای مضمون A در جدول ۳ محاسبات به صورت زیر است:

$$\begin{aligned} \text{ابتدا میانگین هندسی درایه‌های سطرهای جدول ۳ را محاسبه می‌کنیم که به صورت زیر می‌شود.} \\ \text{میانگین هندسی سطر اول} = [(1,1,1) \times (1.456,1.956,2.454) \times (2.631,3.769,4.84) \times (1.64,2.113,2.677) \\ \times (1.204,1.513,1.901)]^{1/5} = (1.499,1.882,2.271) \end{aligned}$$

به طریق مشابه برای سطرهای دیگر نیز این محاسبات صورت می‌گیرد که نتایج در ستون دوم جدول ۹ برای همه سطرهای آورده شده است. سپس مجموع تمامی این میانگین‌های هندسی را به‌دست می‌آوریم که برابر با  $(۴/۸۷۷, ۵/۹۰۳, ۷/۰۴)$  می‌شود، آن‌گاه وزن فازی هر مضمون از تقسیم میانگین هندسی سطر آن مضمون بر مجموع میانگین‌های هندسی به‌دست می‌آید. به‌عنوان مثال برای مضمون A وزن فازی به صورت زیر محاسبه می‌شود:

$$\text{وزن فازی A} = \frac{(1.499,1.882,2.271)}{(4.877,5.903,7.04)} = (0.213,0.319,0.466)$$

برای همه مضمون‌ها محاسبات مشابه صورت می‌گیرد. وزن‌های فازی در ستون سوم جدول ۹ آورده شده است. سپس غیرفازی کردن هر وزن فازی به طریق زیر انجام می‌شود:

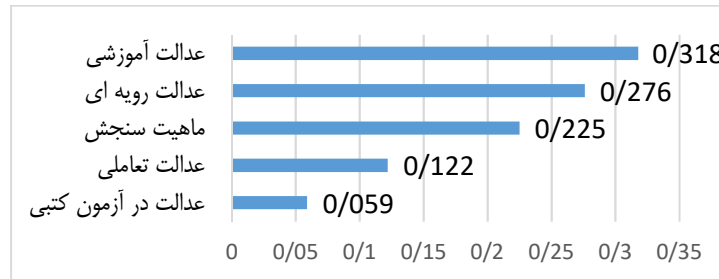
$$\text{وزن غیر فازی A} = \frac{0.213 + 2 \times 0.319 + 0.416}{4} = 0.329 \Rightarrow \text{وزن غیر فازی A} = (0.213,0.319,0.466)$$

برای همه مضمون‌ها نیز این فرایند صورت می‌گیرد که نتایج در ستون چهارم جدول ۹ آورده شده است. سپس برای نرمال‌سازی هر وزن غیر فازی به طریق زیر عمل می‌کنیم:

$$\text{وزن نرمال A} = \frac{0.329}{0.329 + 0.232 + 0.061 + 0.286 + 0.126} = 0.318 \Rightarrow \text{وزن نرمال A} = 0.329$$

جدول ۹. وزن فازی و غیرفازی مضمون‌های اصلی

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \bar{p}_{ij})^{1/n})$	وزن فازی $(\bar{W})$	وزن غیرفازی	وزن نرمال
A	(۱/۴۹۹, ۱/۸۸۲, ۲/۲۷۱)	(۰/۲۱۳, ۰/۳۱۹, ۰/۴۶۶)	۰/۳۲۹	۰/۳۱۸
B	(۱/۱۲۲, ۱/۳۲۵, ۱/۵۶۳)	(۰/۱۵۹, ۰/۲۲۵, ۰/۳۲)	۰/۲۳۲	۰/۲۲۵
C	(۰/۲۹۱, ۰/۳۴۲, ۰/۴۲۶)	(۰/۰۴۱, ۰/۰۵۸, ۰/۰۸۷)	۰/۰۶۱	۰/۰۵۹
D	(۱/۳۷۲, ۱/۶۳۹, ۱/۹۱۴)	(۰/۱۹۵, ۰/۲۷۸, ۰/۳۹۲)	۰/۲۸۶	۰/۲۷۶
E	(۰/۵۹۴, ۰/۷۱۵, ۰/۸۶۶)	(۰/۰۸۴, ۰/۱۲۱, ۰/۱۷۸)	۰/۱۲۶	۰/۱۲۲
	$\sum \left( \prod_{j=1}^n \bar{p}_{ij} \right)^{1/n}$ (۴/۸۷۷, ۵/۹۰۳, ۷/۰۴)			

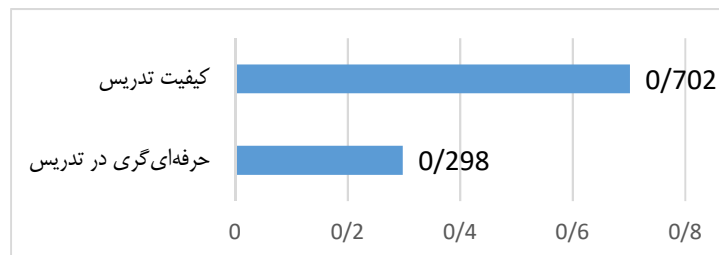


شکل ۲. وزن‌های مضمون‌های اصلی

با توجه به شکل ۲، مضمون عدالت آموزشی با وزن ۰/۳۱۸ رتبه اول را کسب کرده است. عدالت رویه‌ای با وزن ۰/۲۷۶ رتبه دوم و ماهیت سنجش با وزن ۰/۲۲۵ رتبه سوم را کسب کرده است. به طریق مشابه برای دیگر مقایسه‌های زوجی (زیرمضمون‌ها) این محاسبات صورت می‌گیرد که در ادامه آورده شده است.

جدول ۱۰. وزن فازی و غیرفازی زیرمضمون‌های عدالت آموزشی

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \bar{P}_{ij})^{1/n})$	وزن فازی $(\tilde{W})$	وزن غیرفازی	وزن نرمال
A1	(۱/۳۹۲، ۱/۵۴، ۱/۶۷۵)	(۰/۵۸۲، ۰/۷۰۳، ۰/۸۴۲)	۰/۷۰۸	۰/۷۰۲
A2	(۰/۵۹۷، ۰/۶۵، ۰/۷۱۸)	(۰/۲۴۹، ۰/۲۹۷، ۰/۳۶۱)	۰/۳۰۱	۰/۲۹۸
	$\sum \left( \prod_{j=1}^n \bar{P}_{ij} \right)^{1/n}$ (۱/۹۸۹، ۲/۱۸۹، ۲/۳۹۳)			

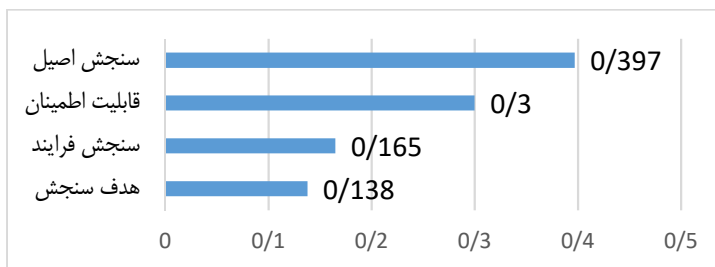


شکل ۳. وزن‌های زیرمضمون‌های عدالت آموزشی

با توجه به شکل ۳، در بین زیرمضمون‌های عدالت آموزشی، کیفیت تدریس با وزن ۰/۷۰۲ رتبه اول و حرفه‌ای‌گری در تدریس با وزن ۰/۲۹۸ رتبه دوم را کسب کرده است.

جدول ۱۱. وزن فازی و غیرفازی زیرمضمون‌های ماهیت سنجش

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \bar{P}_{ij})^{1/n})$	وزن فازی $(\tilde{W})$	وزن غیرفازی	وزن نرمال
B1	(۰/۵۲، ۰/۶۰۲، ۰/۷۲۳)	(۰/۰۹۹، ۰/۱۳۷، ۰/۱۹۹)	۰/۱۴۳	۰/۱۳۸
B2	(۱/۰۹۲، ۱/۳۱۷، ۱/۵۶۱)	(۰/۲۰۸، ۰/۳۰۱، ۰/۴۲۹)	۰/۳۱۰	۰/۳۰۰
B3	(۰/۶۲۴، ۰/۷۲۷، ۰/۸۴۲)	(۰/۱۱۹، ۰/۱۶۶، ۰/۲۳۱)	۰/۱۷۰	۰/۱۶۵
B4	(۱/۴۰۵، ۱/۷۳۷، ۲/۱۱۵)	(۰/۲۶۸، ۰/۳۹۶، ۰/۵۸۱)	۰/۴۱۰	۰/۳۹۷
	$\sum \left( \prod_{j=1}^n \bar{P}_{ij} \right)^{1/n}$ (۳/۶۴، ۴/۳۸۲، ۵/۲۴۱)			

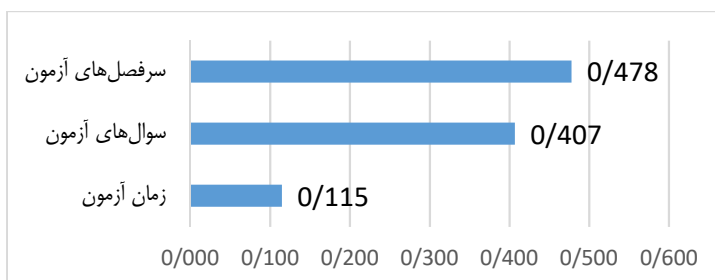


شکل ۴. وزن‌های زیرمضمون‌های ماهیت سنجش

با توجه به شکل ۴، در بین زیرمضمون‌های ماهیت سنجش، سنجش اصیل با وزن ۰/۳۹۷ رتبه اول را کسب کرده است. قابلیت اطمینان با وزن ۰/۳ رتبه دوم و سنجش فرایند با وزن ۰/۱۶۵ رتبه سوم را کسب کرده است.

جدول ۱۲. وزن فازی و غیرفازی زیرمضمون‌های عدالت در آزمون کتبی

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \tilde{P}_{ij})^{1/n})$	وزن فازی $(\tilde{W})$	وزن غیرفازی	وزن نرمال
C1	(۰/۳۵۱، ۰/۴۰۲، ۰/۴۸۱)	(۰/۰۸۶، ۰/۱۱۳، ۰/۱۵۶)	۰/۱۱۷	۰/۱۱۵
C2	(۱/۴۷۸، ۱/۷۰۶، ۱/۹۳۵)	(۰/۳۶۳، ۰/۴۷۸، ۰/۶۲۹)	۰/۴۸۷	۰/۴۷۸
C3	(۱/۲۴۸، ۱/۴۵۷، ۱/۶۵۷)	(۰/۳۰۶، ۰/۴۰۹، ۰/۵۳۹)	۰/۴۱۶	۰/۴۰۷
	$\sum \left( \prod_{j=1}^n \tilde{P}_{ij} \right)^{1/n}$ (۳/۰۷۷، ۳/۵۶۵، ۴/۰۷۴)			

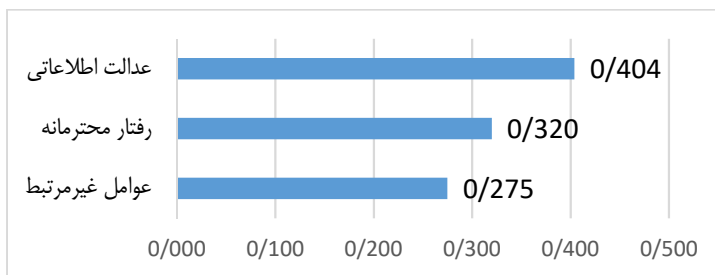


شکل ۵. وزن‌های زیرمضمون‌های عدالت در آزمون کتبی

با توجه به شکل ۵، در بین زیرمضمون‌های عدالت در آزمون کتبی، سرفصل‌های آزمون با وزن ۰/۴۷۸ رتبه اول را کسب کرده است. سؤال‌های آزمون با وزن ۰/۴۰۷ رتبه دوم و زمان آزمون با وزن ۰/۱۱۵ رتبه سوم را کسب کرده است.

جدول ۱۳. وزن فازی و غیرفازی زیرمضمون‌های عدالت تعاملی

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \tilde{P}_{ij})^{1/n})$	وزن فازی $(\tilde{W})$	وزن غیرفازی	وزن نرمال
D1	(۰/۷۰۲، ۰/۸۲۹، ۰/۹۹۵)	(۰/۲، ۰/۲۷۳، ۰/۳۸)	۰/۲۸۱	۰/۲۷۵
D2	(۰/۸۵۱، ۰/۹۷۵، ۱/۱۲)	(۰/۲۴۲، ۰/۳۲، ۰/۴۲۷)	۰/۳۲۷	۰/۳۲۰
D3	(۱/۰۶۸، ۱/۲۳۷، ۱/۴۰۵)	(۰/۳۰۳، ۰/۴۰۷، ۰/۵۳۶)	۰/۴۱۳	۰/۴۰۴
	$\sum \left( \prod_{j=1}^n \tilde{P}_{ij} \right)^{1/n}$ (۲/۶۲۲، ۳/۰۴۱، ۳/۵۲)			

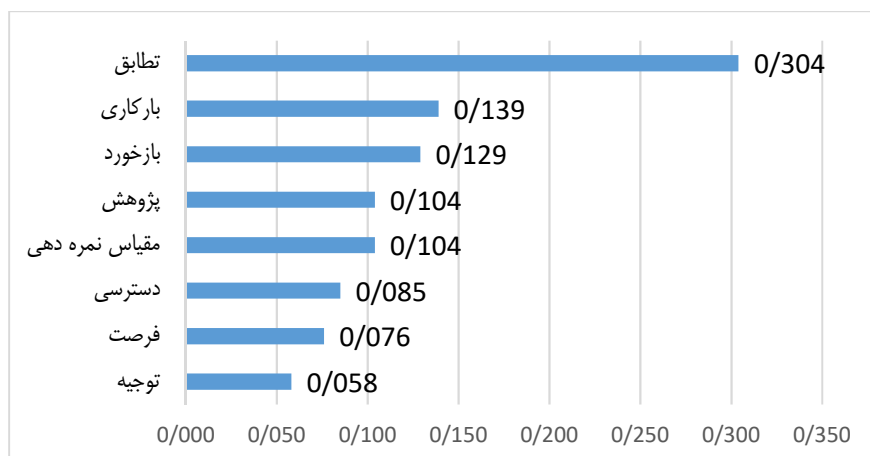


شکل ۶. وزن‌های زیرمضمون‌های عدالت تعاملی

با توجه به شکل ۶، در بین زیرمضمون‌های عدالت تعاملی، عدالت اطلاعاتی با وزن ۰/۴۰۴ رتبه اول را کسب کرده است. رفتار محترمانه با وزن ۰/۳۲۰ رتبه دوم و عوامل غیرمرتبط با وزن ۰/۲۷۵ رتبه سوم را کسب کرده است.

جدول ۱۴. وزن فازی و غیرفازی زیرمضمون‌های عدالت رویه‌ای

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \tilde{P}_{ij})^{1/n})$	وزن فازی $(\tilde{W})$	وزن غیرفازی	وزن نرمال
E1	(۲/۲۵۹ ، ۲/۷۸۹ ، ۳/۳۱۱)	(۰/۲۰۵ ، ۰/۳۰۷ ، ۰/۴۴۷)	۰/۳۱۷	۰/۳۰۴
E2	(۰/۷۸۹ ، ۰/۹۴۳ ، ۱/۱۴)	(۰/۰۷۲ ، ۰/۱۰۴ ، ۰/۱۵۴)	۰/۱۰۸	۰/۱۰۴
E3	(۱/۰۲۳ ، ۱/۲۵۶ ، ۱/۵۴۳)	(۰/۰۹۳ ، ۰/۱۳۸ ، ۰/۲۰۸)	۰/۱۴۴	۰/۱۳۹
E4	(۰/۷۶۳ ، ۰/۹۴۷ ، ۱/۱۵)	(۰/۰۶۹ ، ۰/۱۰۴ ، ۰/۱۵۵)	۰/۱۰۸	۰/۱۰۴
E5	(۰/۶۵۶ ، ۰/۷۷۳ ، ۰/۹۱۹)	(۰/۰۵۹ ، ۰/۰۸۵ ، ۰/۱۲۴)	۰/۰۸۸	۰/۰۸۵
E6	(۰/۵۶۳ ، ۰/۶۸۳ ، ۰/۸۴۵)	(۰/۰۵۱ ، ۰/۰۷۵ ، ۰/۱۱۴)	۰/۰۷۹	۰/۰۷۶
E7	(۰/۹۳۸ ، ۱/۱۷۳ ، ۱/۴۴۳)	(۰/۰۸۵ ، ۰/۱۲۹ ، ۰/۱۹۵)	۰/۱۳۵	۰/۱۲۹
E8	(۰/۴۱۲ ، ۰/۵۱۶ ، ۰/۶۷۲)	(۰/۰۳۷ ، ۰/۰۵۷ ، ۰/۰۹۱)	۰/۰۶۰	۰/۰۵۸
$\sum \left( \prod_{j=1}^n \tilde{P}_{ij} \right)^{1/n}$	(۷/۴۰۲ ، ۹/۰۸۱ ، ۱۱/۰۲۳)			



شکل ۷. وزن‌های زیرمضمون‌های عدالت رویه‌ای

با توجه به شکل ۷، در بین زیرمضمون‌های عدالت رویه‌ای، تطابق با وزن ۰/۳۰۴ رتبه اول را کسب کرده است. بارکاری با وزن ۰/۱۳۹ رتبه دوم و بازخورد با وزن ۰/۱۲۹ رتبه سوم را کسب کرده است.

## ج- وزن‌های نهایی زیرمضمون‌ها

وزن‌های نهایی زیرمضمون‌ها از ضرب وزن مضمون‌های اصلی در وزن نسبی زیرمضمون‌ها حاصل می‌شود که در جدول ۱۵ آورده شده است. بر این اساس زیرمضمون کیفیت تدریس رتبه اول را در بین تمامی زیرمضمون‌ها کسب کرده است.

جدول ۱۵. وزن نهایی زیرمضمون‌ها

رتبه	وزن نهایی زیرمضمون	وزن نسبی زیرمضمون	زیرمضمون	وزن مضمون	مضمون
۱	۰/۲۲۳۲	۰/۷۰۲	کیفیت تدریس	۰/۳۱۸	عوامل عدالت آموزشی
۳	۰/۰۹۴۸	۰/۲۹۸	حرفه‌ای‌گری در تدریس		
۱۰	۰/۰۳۱۱	۰/۱۳۸	هدف‌سنجش		
۷	۰/۰۶۷۵	۰/۳۰۰	قابلیت اطمینان	۰/۲۲۵	عوامل ماهیت سنجش
۸	۰/۰۳۷۱	۰/۱۶۵	سنجش فرایند		
۴	۰/۰۸۹۳	۰/۳۹۷	سنجش اصیل		
۲۰	۰/۰۰۶۸	۰/۱۱۵	زمان آزمون		عوامل عدالت در آزمون کتبی
۱۱	۰/۰۲۸۲	۰/۴۷۸	سرفصل‌های آزمون	۰/۰۵۹	
۱۲	۰/۰۲۴۰	۰/۴۰۷	سوال‌های آزمون		
۶	۰/۰۷۵۹	۰/۲۷۵	عوامل غیرمرتبط		عوامل عدالت رویه ای
۵	۰/۰۸۸۳	۰/۳۲۰	رفتار محترمانه	۰/۲۷۶	
۲	۰/۱۱۱۵	۰/۴۰۴	عدالت اطلاعاتی		
۹	۰/۰۳۷۱	۰/۳۰۴	تطابق		عدالت تعاملی
۱۵	۰/۰۱۲۷	۰/۱۰۴	مقیاس نمره دهی		
۱۳	۰/۰۱۷۰	۰/۱۳۹	بارکاری		
۱۵	۰/۰۱۲۷	۰/۱۰۴	پژوهش	۰/۱۲۲	
۱۷	۰/۰۱۰۴	۰/۰۸۵	دسترسی		
۱۸	۰/۰۰۹۳	۰/۰۷۶	فرصت		
۱۴	۰/۰۱۵۷	۰/۱۲۹	بازخورد		
۱۹	۰/۰۰۷۱	۰/۰۵۸	توجیه		

## بحث و نتیجه‌گیری

در این پژوهش مضمون‌ها و زیرمضمون‌های سنجش کلاسی عادلانه شناسایی شده و در نهایت با تحلیل سلسله‌مراتبی فازی اولویت‌بندی شدند. با تحلیل‌های صورت‌گرفته مشخص شد که در بین مضمون‌های پنج‌گانه اصلی، عدالت آموزشی بیش‌ترین وزن را دارد و زیرمضمون کیفیت تدریس بیش‌ترین وزن را در این مضمون و در مقایسه با تمام زیرمضمون‌های شناسایی شده به خود اختصاص داده است. عدالت آموزشی در این پژوهش به عدم موفقیت استاد در ایفای نقش خود به‌عنوان مربی با تمرکز بر فعالیت‌های تدریس اشاره می‌کند و چنانچه بمپچت<sup>۱</sup> و همکاران (۲۰۱۳) بیان می‌کنند آموزش بی‌اثر انگیزه دانشجویان را کاهش می‌دهد و باعث می‌شود آن‌ها فکر کنند تلاش‌هایشان به نتایج مورد انتظار منجر نمی‌شود. اهمیت کیفیت تدریس استاد و برداشت عادلانه از سنجش او در پیوند با زیرمضمون «تطابق» که یکی از نشانگرهای آن تطابق سطح سنجش با کیفیت تدریس استاد است و بالاترین وزن را در زیرمضمون‌های عدالت رویه‌ای دارد بیش‌تر نمایان می‌شود؛ چنانچه مک‌میلان<sup>۲</sup> (۲۰۱۷) نیز اشاره می‌کند نقض عدالت آموزشی در بُعد کیفیت تدریس به درک ناعادلانه سنجش منجر خواهد شد و در مقابل اگر افراد احساس کنند تدریس استاد از کیفیت مطلوبی برخوردار است و رویه‌های مورد استفاده برای تخصیص نتایج عادلانه است حتی اگر برای آن‌ها مطلوب نباشد،

1. Bempechat

2. McMillan



نتایج را می‌پذیرند و آن را عادلانه‌تر درک می‌کنند. همان‌طور که روداباو<sup>۱</sup> (۱۹۹۴) نشان می‌دهد، دانشجویان بیش از دریافت نمرات بالاتر، نگران عدالت در کلاس هستند؛ بنابراین، استادی که نمرات کمتری می‌دهد اما از رویه‌های عادلانه استفاده می‌کند، نسبت به کسی که نمرات بالاتری می‌دهد اما از رویه‌های ناعادلانه استفاده می‌کند، عادلانه‌تر ادراک می‌شود.

در مضمون «ماهیت سنجش» به‌عنوان مضمونی که در جایگاه سوم از نظر وزن‌های تخصیص داده شده قرار می‌گیرد زیرمضمون «سنجش اصیل»<sup>۲</sup> بالاترین وزن را به خود اختصاص داده است. در دهه گذشته، برخی از روندهای سنجش کلاسی برای سازگاری بهتر با دانش، مهارت‌ها و شرایط قرن ۲۱ ایجاد شده است که تغییر از فرهنگ آزمون<sup>۳</sup> به فرهنگ یادگیری<sup>۴</sup> را موجب شده است. این روندها باعث کاربرد سنجش‌های جایگزین، از جمله سنجش اصیل، سنجش عملکرد<sup>۵</sup>، نمونه کارها، نمایشگاه‌ها، نمایش‌ها، پژوهش‌ها و موارد پیشرفته فناوری شده است (مک میلان، ۲۰۱۷؛ تیرنی، ۲۰۱۴). سنجشی که صرفاً در سطح دانش باشد، عادلانه نیست امروزه بر سنجش کاربردی، مسئله محور، آینده‌نگر و در سطح تحلیل و کاربرد تأکید می‌شود. چنین معنایی از سنجش، مفهوم تکالیف اصیل را به ذهن متبادر می‌کند چراکه بر موقعیت‌های زندگی واقعی و ایجاد توانمندی‌های لازم برای مواجهه با آنچه دانشجویان بعد از تحصیل با آن روبرو خواهند شد، تأکید می‌کنند. سنجش اصیل باعث بهبود یادگیری و افزایش انگیزه دانشجویان می‌شود. علاوه بر این، توانایی بهبود عملکرد دانشجویان و فراهم آوردن فرصت‌های یادگیری بیشتر برای آن‌ها را دارد. استفاده از سنجش چندگانه به‌جای سنجش واحد، اضطراب دانش‌آموز را کاهش می‌دهد و در نتیجه ریسک‌پذیری، قدرت خلاقیت و اکتشاف وی را افزایش می‌دهد (موریلو و هیدالگو<sup>۶</sup>، ۲۰۱۷؛ تیرنی، ۲۰۱۳).

در مضمون «عدالت تعاملی» که بر ماهیت رابطه استاد و دانشجو تأکید دارد و بر کیفیت رفتار استاد و دانشجو هنگام اجرای رویه‌ها یا تخصیص نتایج تمرکز دارد (بمپچت و همکاران، ۲۰۱۳) زیرمضمون «عدالت اطلاعاتی» بیش‌ترین وزن را دارد. عدالت اطلاعاتی دربرگیرنده تعیین نشانگرهای سنجش، انتظارات استاد از دانشجویان و عمل بر اساس ملاک‌های توافق شده توسط استاد است. ما این واحدهای معنایی را تحت عنوان «عدالت اطلاعاتی» قرار داده‌ایم. عدم توضیح استاد درباره روش‌ها و شیوه‌های نمره‌دهی، تعریف واضح و روشن انتظارات و سیاست‌های دوره در شروع آن توسط استاد، عدم ارائه راهنمای مشخص برای اختصاص نمرات به‌ویژه برای حضور و مشارکت در کارگروهی، مشخص کردن موادی که جهت آماده شدن یک امتحان کلاسی مهم هستند توسط مدرس، مشخص نکردن آنچه مورد سنجش قرار خواهد گرفت، راهنما برای آمادگی آزمون (راهنمای مطالعه، اهداف یادگیری و موارد دیگر)، مشخص بودن مضمون‌های نمره‌دهی مشارکت کلاسی در برنامه درسی، ارائه اطلاعات توسط استاد درباره تعداد آزمون‌ها، تاریخ برگزاری، قالب، محتوا، مدت زمان، درصد تأثیر در نمره نهایی هر یک و نظایر آن، نشانگرهایی هستند که در پژوهش‌ها برای «عدالت اطلاعاتی» آمده است. وقتی دانشجویان از ملاک‌های ارزیابی خودشان در یک دوره مطلع شوند فرصت خودسنجی در جهت رسیدن به اهداف را خواهند داشت بر این اساس، هنگامی که اساتید انتظارات و سیاست‌های خود را در ابتدای دوره تعیین می‌کنند (مادامی که رفتارهای بعدی استاد با آن انتظارات و سیاست‌ها سازگار باشد)، دانشجویان عملکرد اساتید خود را عادلانه می‌دانند (هستون و بتنکورت<sup>۷</sup>، ۱۹۹۹). البته توجه به این مطلب ضروری است که تصریح ملاک‌ها با توجه به تأثیر نمرات در آینده شغلی و تحصیلی یادگیرندگان نیز قابل بررسی است چنانچه تونگ<sup>۸</sup> (۲۰۰۹) معتقد است دانشجویان عصر جدید، انتظار دارند مدرسان آنچه را آن‌ها برای کسب نمرات بالا باید انجام دهند به صراحت و دقیقاً به آن‌ها بگویند.

«عدالت در آزمون کتبی» کمترین وزن را در بین مضمون‌های سنجش کلاسی عادلانه دارد. این موضوع نشان می‌دهد که در سنجش عادلانه وزن بیشتر بر فرایندها، ماهیت سنجش و رویه‌ها قرار می‌گیرد چنانچه جانسون و اسوینگی<sup>۹</sup> (۲۰۰۷) نیز تأکید می‌کنند فرهنگ سنجش جدید بر ارزیابی فرایندها و شایستگی‌های تفکر سطح بالاتر به‌جای دانش واقعی و مهارت‌های شناختی سطح پایین، تأکید دارد. همچنین قنبری و همکاران (۱۳۹۴) در پژوهش خود با عنوان "تأثیر چالش‌های ارزشیابی آموخته‌های دانشجویان بر رویکرد مطالعه عمل فکورانه" که از پرسشنامه شناسایی چالش‌های ارزشیابی دانشجویان استفاده کردند نشان دادند که از بین شش چالش تأکید بیش‌ازحد بر ارزشیابی پایانی، عدم استفاده از ارزشیابی فرایندی و مستمر، عدم استفاده از ارزشیابی‌های غیر عینی، عدم توجه به تفاوت‌های فردی، عدم توجه به مجموعه هدف‌های یادگیری و عدم

1. Rodabaugh

2. authentic assessment

3. Testing culture

4. Learning culture

5. performance assessment

6. Murillo & Hidalgo

7. Houston & Bettencourt

8. Twenge

9. Jonsson & Svingby

آشنایی اساتید با فلسفه و هدف ارزشیابی چالش مربوط به عدم توجه به تفاوت‌های فردی بیشترین و چالش مربوط به عدم استفاده از ارزشیابی‌های غیرعینی کمترین بار عاملی را داشتند.

با عنایت به اینکه بر اساس نقشه جامع علمی کشور (۱۳۹۰) تربیت انسان‌های مؤمن، توانا، خردمند، سالم، خلاق، آزاده، دارای فضائل اخلاقی و اعتمادبه‌نفس، تقویت ارزش‌های معنوی و فضائل اخلاقی، ترویج فرهنگ آزاداندیشی، نقادی و نقدپذیری از وظایف نهاد علم است؛ ملاک‌ها و نشانگرهای به‌دست‌آمده از این پژوهش می‌تواند به‌عنوان یکی از منابع تدوین دستورالعمل ملی سنجش کلاسی عادلانه در جهت کمک به تحقق عدالت آموزشی مورد استفاده قرار گیرد و با توجه به اهمیت کیفیت تدریس اعضای هیأت علمی در ادراک دانشجویان از عدالت در سنجش، پایش عملکرد آموزشی آن‌ها ضروری به نظر می‌رسد؛ در این راستا استفاده از نشانگرها و ملاک‌های سنجش کلاسی عادلانه در ارزشیابی اعضای هیأت علمی می‌تواند مورد استفاده قرار گیرد. همچنین با عنایت به اینکه آگاهی از مؤلفه‌های سنجش کلاسی عادلانه به افزایش سواد سنجشی اعضای هیأت علمی و ارتقاء کیفیت یادگیری دانشجویان و بهزیستی روان‌شناختی آنها کمک می‌کند اجرای دوره‌های آموزشی در این زمینه برای اعضاء هیأت علمی پیشنهاد می‌شود.

## References

- Aldian, A., Taylor, M. A. (2005). A consistent method to determine flexible criteria weights for multicriteria transport project evaluation in developing countries. *Journal of the Eastern Asia Society for Transportation Studies*, 6, 3948-3963.
- Azar, A. (2010). *Applied decision making (MADM approach)*. Tehran: Negah Danesh. (Persian)
- Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice*, 20(1), 20-34.
- Alm, F., & Colnerud, G. (2015). Teachers' experiences of unfair grading. *Educational Assessment*, 20(2), 132-150.
- Baniasadi. A.. Salehi. K. (2019). Introduction on the Priniciples and Process of Construction and Validation of the Interview Protocol. *Higher Education Letter*, 12(46), 177-203. (Persian)
- Baniasadi. A.. Salehi. K.. Khodaie. F.. Bagheri Noaparast. K.. IZANLOO. B. (2021). Students' perceptions of fair classroom assessment: A qualitative study. *Educational Measurement and Evaluation Studies*, 11(35), 68-88. (Persian)
- Bazargan, A. (2013). *An Introduction to Qualitative and Mixed Methods Research*. Tehran: Didar. (Persian)
- Bempechat, J., Ronfard, S., Li, J., Mirny, A., & Holloway, S. D. (2013). "SHE ALWAYS GIVES GRADES LOWER THAN ONE DESERVES:" A QUALITATIVE STUDY OF RUSSIAN ADOLESCENTS' PERCEPTIONS OF FAIRNESS IN THE CLASSROOM. *Journal of Ethnographic & Qualitative Research*, 7(4), 168-187.
- Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement: Issues and Practice*, 22(4), 5-12.
- Brown, G. T. L., Pishghadam, R., & Sadafian, S. S. (2014). Iranian university students' conceptions of assessment: Using assessment to self-improve. *Assessment Matters*, 6, 5-33.
- Čiuladienė, G., & Račelytė, D. (2016). Perceived unfairness in teacher-student conflict situations: students' point of view. *Polish Journal of Applied Psychology*, 14(1), 49-66.
- DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education*, 34(5-6), 576-591.
- Ghanbari, S., Ardalan, MR., Karimi, I. (2015). Effect of the Challenges of Student Learnings Evaluation on Deliberate Practice Study Approach. *Education Strategies Medical Sciences*, 8(2), 105-113. (Persian)
- Houston, M. B., & Bettencourt, L. A. (1999). But that's not fair! An exploratory study of student perceptions of instructor fairness. *Journal of Marketing Education*, 21(2), 84-96.
- Hsieh, T.Y., Lu, S.T. and Tzeng, G.H., 2004. Fuzzy MCDM approach for planning and design tenders selection in public office buildings. *International journal of project management*, 22(7), 573-584.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational

- consequences. *Educational research review*, 2(2), 130–144.
- Kosko, B. (1994). *Fuzzy Thinking: The New Science of Fuzzy Logic* (A. Ghaffari, Trans.). Tehran: K. N. Toosi University. (Persian)
- Liu, J., Johnson, R., & Fan, X. (2016). A comparative study of Chinese and United States pre-service teachers' perceptions about ethical issues in classroom assessment. *Studies in Educational Evaluation*, 48, 57–66.
- McMillan, J. H. (2017). *Classroom assessment: Principles and practice that enhance student learning and motivation*. Pearson.
- Murillo, F. J., & Hidalgo, N. (2017). Students' conceptions about a fair assessment of their learning. *Studies in Educational Evaluation*, 53, 10–16.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497.
- Pornaderi, p., Nasr, A, R., Nili, M, R., & Bagheri, Z. (2013). Evaluation of the Implementing Different Methods of Measurement and Sustainable Learning Based Measurements. *Technology of Education Journal*, 8 (1), 45-54. (Persian)
- Ricoeur, P. (1976). *Interpretation theory: Discourse and the surplus of meaning*. TCU press.
- Rodabaugh, R. C. (1994). College students' perceptions of unfairness in the classroom. *To Improve the Academy*, 13(1), 269–282.
- Seraji, F., Maroofi, Y., & Razeqi, T. (2013). Identifying the challenges of evaluating what students have learned in the Iranian higher education system. *Educational Measurement and Evaluation Studies*, 4 (5), 33-54. (Persian)
- Solomonidou, G., & Michaelides, M. (2017). Students' conceptions of assessment purposes in a low stakes secondary-school context: A mixed methodology approach. *Studies in Educational Evaluation*, 52, 35-41.
- Speziale, H. S., Streubert, H. J., & Carpenter, D. R. (2011). *Qualitative research in nursing: Advancing the humanistic imperative*. Lippincott Williams & Wilkins.
- Supreme Council for Cultural Revolution. (2012). *Comprehensive scientific map of the country*. Tehran: Secretariat of the Supreme Council of the Cultural Revolution. (Persian)
- Tata, J. (2005). The influence of national culture on the perceived fairness of grading procedures: A comparison of the United States and China. *The Journal of Psychology*, 139(5), 401–412.
- Tierney, R. D. (2013). Fairness in classroom assessment. *SAGE Handbook of Research on Classroom Assessment*, November, 125–144. <https://doi.org/10.4135/9781452218649.n8>
- Tierney, R. D. (2014). Fairness as a multifaceted quality in classroom assessment. *Studies in Educational Evaluation*, 43, 55–69.
- Tillema, H., Leenknech, M., & Segers, M. (2011). Assessing assessment quality: Criteria for quality assurance in design of (peer) assessment for learning – A review of research studies. *Studies in Educational Evaluation*, 37, 25-34.
- Twenge, J. M. (2009). Generational changes and their impact in the classroom: teaching Generation Me. *Medical Education*, 43(5), 398–405.
- Vinodh, S., Prasanna, M., & Prakash, N. H. (2014). Integrated Fuzzy AHP–TOPSIS for selecting the best plastic recycling method: A case study. *Applied Mathematical Modelling*, 38(19–20), 4662–4672.
- Yamtim, V., & Wongwanich, S. (2014). A study of classroom assessment literacy of primary school teachers. *Procedia-Social and Behavioral Sciences*, 116, 2998–3004.