



Identifying and prioritizing Equitable classroom assessment themes using Fuzzy Analytic Hierarchy Process (FAHP) method

Ali Baniasadi¹, Keyvan Salehi², Ebrahim Khodaie³, Khosro Bagheri⁴, Balal Izanloo⁵

1. Ph.D. Student of Assessment And Measurement, Faculty of Psychology and Educational sciences, University of Tehran, Tehran, Iran, ali.baniasadi@ut.ac.ir

2. Assistant Professor, Department of Methods and Curriculum Planning, Faculty of Psychology and Educational sciences, University of Tehran, Tehran, Iran. Corresponding Author: keyvansalehi@ut.ac.ir

3. Associate Professor, Department of Methods and Curriculum Planning, Faculty of Psychology and Educational sciences, University of Tehran, Tehran, Iran, Khodaie@ut.ac.ir

4. Professor, Department of Philosophical and social foundations of education, Faculty of Psychology and Educational sciences, University of Tehran, Tehran, Iran, khbagheri@ut.ac.ir

5. Assistant Professor, Department of Curriculum Planning, Faculty of Psychology and Educational Sciences, Kharazmi University, Tehran, Iran, izan.b@khu.ac.ir

Article Info

ABSTRACT

Article Type:

Research Article

Received

2021/09/30

Received in

revised form

2022/02/18

Accepted

2022/02/28

Published online

2022/03/26

Objective: The aim of this study was identifying and weighting the themes and sub- themes of equitable classroom assessment.

Methods: To achieve this goal, at first a qualitative study was conducted using interpretive phenomenology. A Semi-structured interview was used to collect data. Using criterion sampling, after conducting 29 individual interviews, a group interview and a focus group with students who had at least two and a half year of measurement performance at Tehran University, inductive thematic saturation was obtained. Fuzzy analytic hierarchy process method was used to weight these themes. The Data of this stage was collected by a pairwise comparison questionnaire that included 53 comparisons from a sample of 9 experts in the assessment field.

Results: The analysis of students' perception was done using the Ricoeur method and in three stages: simple reading, structural analysis and comprehensive understanding. Extraction of semantic units led to the identification of 20 sub-themes, which were categorized into 5 main themes: "educational justice", "procedural justice", "the nature of assessment", "interactive justice" and "fairness in the written test". The results of fuzzy analytic hierarchy process showed that among the 5 main themes, "educational fairness" with a weight of 0.318 was in the first place, after that the themes of "procedural fairness", "nature of assessment", "interactional fairness", and "fairness in paper and pencil assesment" placed 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 after that the sub- themes of "informational fairness" and "professionalism in teaching" with weights of 0.11 and 0.09 were placed, 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 "educational justice" among the themes and sub- theme of "teaching quality" of this theme among all of the obtained sub- themes showed that Violation of educational justice in terms of the quality of teaching will lead to an unfair understanding of assessment, and on the other hand, if people feel that the teacher's teaching has a good quality and the procedures used to allocate the results are fair, even if it is not favorable for them, they will accept the results and understand it as more fair.

Keywords: justice, phenomenology, class assessment, equity, perceived justice.

Cite this article: Baniasadi, Ali, Salehi, Keyvan. Khodaie, Ebrahim. Bagheri, Khosro. & Izanloo, Balal. (2022). Identifying and prioritizing Equitable classroom assessment themes using Fuzzy Analytic Hierarchy Process (FAHP) method. *Educational Measurement and Evaluation Studies*, 12 (37): 31-56 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 سلسله مراتبی فازی

علی بنی‌اسدی^۱، کیوان صالحی^۲، ابراهیم خدایی^۳، خسرو باقری^۴، بلال ایزانلو^۵

۱. دانشجوی دکتری سنجش و اندازه‌گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران، 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
۳. دانشیار گروه روش‌ها و برنامه‌های آموزشی، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران، khodaie@ut.ac.ir
۴. استاد گروه مبانی فلسفی و اجتماعی آموزش پرورش، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران، khbagheri@ut.ac.ir
۵. استادیار گروه برنامه‌ریزی درسی، دانشکده روانشناسی و علوم تربیتی، دانشگاه خوارزمی، تهران، ایران، izan.b@khu.ac.ir

اطلاعات مقاله چکیده

نوع مقاله:	چکیده
مقاله پژوهشی	هدف: پژوهش حاضر با هدف شناسایی و وزن‌دهی مضمون‌ها و زیرمضمون‌های سنجش کلاسی عادلانه انجام شد. روش پژوهش: برای دستیابی به این هدف، ابتدا مطالعه‌ای 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 نیم‌سال تجربه مورد سنجش عملکرد واقع شدن در دانشگاه تهران را داشتند، اشباع موضوعی استقرایی به دست آمد. برای وزن‌دهی به این مضمون‌ها از روش فرایند تحلیل سلسله‌مراتبی فازی استفاده شد. داده‌های این مرحله با استفاده از پرسشنامه مقایسه زوجی که شامل ۵۳ مقایسه بود، از نمونه نه نفری از متخصصان حوزه سنجش جمع‌آوری شد.
دریافت	یافته‌ها: 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 بیست زیرمضمون منجر شد که در پنج 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 پنج 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 به دست آمد. بر این اساس زیرمضمون «کیفیت تدریس» با وزن ۲۲/۰ 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.
اصلاح	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 و زیرمضمون کیفیت تدریس در بین تمام زیرمضمون‌های به دست 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 کیفیت مطلوبی دارد و رویه‌های مورد استفاده برای تخصیص نتایج عادلانه است، حتی اگر برای آنها مطلوب نباشد، نتایج را می‌پذیرند و آن را عادلانه‌تر درک می‌کنند.
پذیرش	واژه‌های کلیدی: انصاف، پدیدارشناسی، سنجش کلاسی، عدالت، عدالت‌آدارک‌کشده
انتشار	

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)، صفحه ۳۱-۵۶
DOI:10.22034/EMES.2022.527579.2241



ناشر: سازمان سنجش آموزش کشور © نویسنده‌گان.

مقدمه

روایی، پایایی و عدالت کیفیت‌هایی هستند که در نظریه اندازه‌گیری برای سنجش آموزشی بهینه به آنها اشاره شده است (تیرنی^۱، ۲۰۱۳). لازم است این نکته را به خاطر داشته باشیم که فرضیه‌های نظریه اندازه‌گیری با سنجش کلاسی به طور کامل برازش ندارد چرا که سنجش کلاسی وابسته به محیط آموزشی است که در یک زمینه روانی اجتماعی انجام می‌شود (بروختارت^۲، ۲۰۰۳). چنین محیط ارزیابی با توجه به رویکرد معلم در سنجش که کربی^۳ (۱۹۹۹) آن را «فرهنگ ارزشیابی کلاس‌ها»^۴ خوانده است، طیف وسیعی از محیط‌ها را از یادگیرنده محور تا معلم محور در بر می‌گیرد (به نقل از بروختارت، ۲۰۰۳). بنابر این، سنجش کلاسی فراتر از کاربرد فنون قابل اندازه‌گیری و شامل نگرانی‌های فردی و اجتماعی است (الللال^۵، ۲۰۱۳). تعامل معلم و یادگیرنده و تأثیرات اجتماعی ممکن است با کاهش روایی اطلاعات سنجش، عدالت نمره تعیین شده توسط استاد را تحت تأثیر قرار دهد (آلم و کلنرود^۶، ۲۰۱۵). با در نظر گرفتن این تفاوت زمینه‌ای، استیگینز^۷ (۲۰۰۱) به جای روایی و پایایی از اصطلاح «سنجش کلاسی با کیفیت بالا»^۸ استفاده می‌کند (به نقل از بروختارت، ۲۰۰۳).

موقعیت‌های آموزشی پیچیده و شامل چالش‌های متعددی است. مناقشه میان استاد و یادگیرنده یکی از این چالش‌هاست که به طوری که رفتار استاد منبع مهم این مناقشه و جنبه مهم این تعارض ادراک بی‌عدالتی است. این تعارض‌ها با درک متقابل، حل مسئله، بهبود و تقویت روابط استاد و دانشجو و در نهایت رضایت از فرایندها و پیامدها حل می‌شود (جوالدین و ریسلایت^۹، ۲۰۱۶). پیشرفت یادگیرندگان در یادگیری به کیفیت اساتید و معلمان بستگی دارد که در سه حوزه برنامه درسی، طراحی آموزشی و سنجش، تصمیم‌گیری می‌کنند و چنانچه هر یک از این حوزه‌ها نادیده انگاشته شود تصمیمات به اندازه کافی مؤثر نخواهد بود. چنانچه مک‌للان^{۱۰} (۲۰۰۴) بیان می‌کند، سواد سنجشی ناکافی اساتید در زمینه‌های روایی، پایایی و عدالت، پیامدهای منفی بر یادگیری دانشجویان خواهد داشت (به نقل از دلوکا^{۱۱}، ۲۰۱۲). با توجه به اینکه اساتید حدود پنجاه درصد زمان تدریس درگیر فعالیت‌های مرتبط با سنجش‌اند باید به طور پیوسته، توانایی سنجشی‌شان را افزایش دهند (یامتیم و ونگ وانگ^{۱۲}، 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^{۱۳} (۲۰۰۵) نشان داده شد که فرهنگ (فردگرایی - جمع‌گرایی) بر دو بُعد عدالت رویه‌ای، از جمله صدا^{۱۴} (اظهار نظر دانشجویان در مورد سنجش و موضوعات سنجشی) و عدالت

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

بین فردی^۱ تأثیر می‌گذارد و برداشت دانشجویان چینی از عدالت بیشتر تحت تأثیر عدالت بین فردی است، در حالی که برداشت دانشجویان آمریکایی از عدالت بیشتر تحت تأثیر صدا قرار می‌گیرد. بافت یا زمینه‌ای بودن موضوعات اخلاقی و عدالت در سنجش در پژوهش انجام شده توسط لیو^۲ و همکاران (۲۰۱۶) نیز نشان داده شده است. در پژوهش لیو و همکاران که با عنوان «مطالعه مقایسه‌ای ادراک معلمان چینی و آمریکایی درباره موضوعات اخلاقی در کلاس درس» و با طراحی ۳۶ سناریو که عمل یا اندیشه معلم در باره شیوه‌های سنجش را نشان می‌داد انجام شد. نتایج حاصل از تحلیل‌های دو پاسخ‌های ۱۷۳ معلم آمریکایی و ۱۷۴ معلم چینی نشان داد که نظر معلمان دو کشور در ۲۲ سناریو از ۳۶ تا به طور معنی‌داری متفاوت بود (اندازه اثر بیشتر از ۰/۳) و از هشت سناریوی مربوط به عدالت / سوگیری در سه سناریو تفاوت‌ها معنی‌دار بود. 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 با تمرکز بر چهار عامل «بهبود»، «تأثیر»، «مرتبط نبودن» و «اسنادهای بیرونی» و توسط برون^۳ و همکاران (۲۰۱۴) با عنوان «ادراک دانشجویان دانشگاه‌های ایران درباره سنجش: کاربرد سنجش برای بهبود خود» انجام شده است. شرکت‌کنندگان ۷۶۰ دانشجو (۲۲۶ مرد و ۵۲۸ زن، ۶ تا گم شده) با دامنه سنی ۱۸ تا ۳۸ سال بودند. همانند پژوهش سولیمونیدو و میچیلایدز^۴ (۲۰۱۴) و تلمه^۵ و همکاران (۲۰۱۱)، دانشجویان، سنجش را به عنوان 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 به عنوان شاخص کیفیت مدرسه تا به عنوان پیش‌بینی کننده آینده فردی به دست آورد. «مرتبط نبودن» و «تأثیر»، میانگین نمره زیر سه به دست آوردند که تأیید ضعیفی را نشان می‌دهد. هر چند موافقت کمی با این دیدگاه که سنجش بد است (میانگین = ۳/۱۳، زیرمقیاسی از «مرتبط نبودن»، شامل مقوله‌های سنجش، بی‌ارزش است، سنجش برای دانشجویان ناعادلانه است و نتایج سنجش دقیق نیست) داشتند. نکته‌ای که باید مدنظر قرار گیرد آن است که پرسشنامه ادراک دانش‌آموزان از سنجش مبتنی بر نسخه معلمان برای دانش‌آموزان در دوره متوسطه نیوزیلند و نه برای دانشجویان تدوین شده است. سایر پژوهش‌ها، چالش‌های ارزشیابی دانشجویان را بررسی کرده است 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 بیش از حد 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 بیش از حد بر

1. interpersonal fairness

2. Liu

3. Brown

4. Solomonidou & Michaelides

5. Tillema

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 غیر عینی کمترین بار عاملی را داشت. 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. نتایج پژوهش نشان داد که روش‌های سنجش موجود غیر کاربردی، 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 نمی‌شود و برای اندازه‌گیری آموخته‌ها دقت و صحت کافی ندارد. باور به سفید و سیاه‌ها، صفر و یک‌ها و به طور کلی نظام دو ارزشی به گذشته 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 سیاه و سفید و برداشتن این حجاب دو ارزشی، یعنی نگرستن به جهان به صورتی که هست، بود. 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 تا اندازه‌ای «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 ویژگی‌ها، سنجش انطباق عناصر با مفاهیم متعدد را ممکن می‌سازد. کاربرد اصلی این رویکرد 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 درون هر مجموعه معنایی و بر اساس اصل انطباق بیشتر یا کمتر عنصر x در مقایسه با y با مفهوم A صورت می‌گیرد، در حالی که رده‌بندی ویژگی بر اساس مقایسه بین مصادیق مشترک انجام می‌شود و اینکه عنصر x با مفهوم A بیشتر 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 برای 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 به وجود آمدن عدالت تأثیر دارند (کاتوزیان، ۱۳۸۶) و از طرفی در باب مفاهیم اجتماعی با درجاتی از انتزاع و

1. Property ranking

انضمام روبه‌رو هستیم که سبب‌ساز وجود نوعی ابهام در صدق معنا به مصداق می‌شود؛ به عبارت دیگر، به جای حالات دوگانه «صدق»/«عدم صدق» با حالات چندگانه «درجات صدق» مواجهیم که می‌تواند در موقعیت‌هایی بین حالات دوگانه مزبور قرار گیرد 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 چیست؟
- ۲- اولویت و وزن هر یک از مضمون‌ها و زیرمضمون‌های شناسایی شده بر اساس نظر متخصصان و با کاربرد رویکرد سلسه‌مراتبی تحلیلی فازی^۱ 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 استفاده شده است چرا که این روش با هدف پی بردن به معنای ضمنی 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^۲ برای انتخاب دانشجویان استفاده شد. از این رو با توجه به اینکه ادراک دانشجویان از عدالت در سنجش باید بررسی می‌شد، دانشجویان به گونه‌ای انتخاب شدند که تجربه حداقل دو نیمسال سنجش در دانشگاه تهران داشته باشند. از طرف دیگر، به دلیل طیف گسترده رشته‌ها و دانشکده‌های دانشگاه تهران، به منظور پوشش تجارب و دیدگاه‌های متنوع‌تر، دانشکده‌ها در پنج حوزه علوم انسانی، علوم پایه، فنی و مهندسی، کشاورزی و هنر طبقه‌بندی شدند و با دانشجویان مصاحبه شد. روش استفاده شده برای جمع‌آوری داده‌ها مصاحبه‌های نیمه‌ساختاریافته^۳ بود. مصاحبه‌های نیمه ساختاری قالبی از مصاحبه است که به طور گسترده برای تحقیقات کیفی استفاده می‌شود و می‌تواند به صورت جداگانه یا گروهی انجام شود. مصاحبه‌های گروهی اغلب در مطالعات گروه‌های کانونی استفاده می‌شود که در آن چندین شرکت‌کننده، دانش و تجربیات خود را در مورد یک موضوع خاص به اشتراک می‌گذارند (دیکوبلوم و کرابتری^۴، ۲۰۰۶ به نقل از بنی‌اسدی و صالحی، ۱۳۹۸). نتایج کامل این بخش در پژوهش بنی‌اسدی و همکاران (۱۴۰۰) قابل دسترسی است.

در مورد تولید داده در روش پدیدارشناسی، اسپزیال^۵ 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^۶ بعد از ۲۹ مصاحبه فردی، یک مصاحبه گروهی و یک گروه کانونی به دست آمد. ویژگی‌های توصیفی مشارکت‌کنندگان در جدول (۱) ارائه شده است. پس از رسیدن به اشباع نظری یافته‌ها، روایی محتوایی مضمون‌ها و زیرمضمون‌های به دست آمده توسط نه استاد متخصص

1. Fuzzy Analytical Hierarchy Process (FAHP)
2. criterion sampling
3. Semi- structured interviews
4. DiCicco-Bloom & Crabtree
5. Speziale
6. inductive thematic saturation

در حوزه سنجش تأیید شد. همچنین مضمون‌ها و زیرمضمون‌های به‌دست‌آمده در قالب درخت سلسله مراتبی که شامل پنج مضمون و بیست 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 سلسله مراتبی تحلیلی که نمونه‌ای از روش‌های ذهنی است و ساتی^۱ (۱۹۷۷) آن را پیشنهاد داده است (آلدیان و تیلور^۲، ۲۰۰۵) مشکلات مربوط به تصمیم‌گیری‌های پیچیده با تعیین اهمیت نسبی مضمون‌های دخیل در یک تصمیم‌گیری حل می‌شود و مسئله تصمیم‌گیری چند ملاکی به مجموعه‌ای از تصمیمات مرتبط تجزیه می‌شود (وینود^۳ و همکاران، ۲۰۱۴).

جدول ۱. ویژگی‌های توصیفی مشارکت‌کنندگان (دانشجویان)

ردیف	سابقه تحصیلی	رشته تحصیلی	جنسیت	زمان مصاحبه	ردیف	سابقه تحصیلی	رشته تحصیلی	جنسیت	زمان مصاحبه
۱	۳ ارشد	اکوهیدرولوژی	مرد	۶۰ دقیقه	۱۶	۴ ارشد	ژئومورفولوژی	زن	۵۰ دقیقه
۲	۴ ارشد	آمار	زن	۵۵ دقیقه	۱۷	فارغ‌التحصیل کارشناسی	مدیریت	مرد	کتبی
۳	۳ دکتری	هوافضا	مرد	۶۰ دقیقه	۱۸	۳ کارشناسی	شهرسازی	مرد	کتبی
۴	۷ کارشناسی	زیست‌شناسی سلولی مولکولی	زن	۴۰ دقیقه	۱۹	۳ ارشد	مدیریت شهری	مرد	۴۰ دقیقه
۵	۴ ارشد	شیمی عالی	زن	کتبی	۲۰	۳ ارشد	تصویرسازی	زن	۴۵ دقیقه
۶	۱۲ دکتری	روانشناسی	زن		۲۱	فارغ‌التحصیل کارشناسی	معماری	زن	
۷	فارغ‌التحصیل ارشد	برنامه‌ریزی آموزشی	مرد	۵۵ دقیقه	۲۲	۱۲ دکتری	عمران‌سازه	مرد	۴۰ دقیقه
۸	۸ دکتری	بیوفیزیک	زن	۵۰ دقیقه	۲۳	۵ ارشد	مدیریت شهری	مرد	کتبی
۹	۳ ارشد	روانشناسی کودکان استثنایی	زن	کتبی	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

1. Saaty
2. Aldian & Taylor
3. Vinodh

ردیف	سابقه تحصیلی	رشته تحصیلی	جنسیت	زمان مصاحبه	رتبه	سابقه تحصیلی	رشته تحصیلی	جنسیت	زمان مصاحبه
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 تحلیل گر 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 مقایسه زوجی که شامل ۵۳ مقایسه بود طراحی شد و برای ۱۳ نفر از متخصصان حوزه سنجش فرستاده شد که در نهایت پاسخ‌های نه نفر از آنها دریافت شد. الگوی اولیه فرایند تحلیل سلسله مراتبی (AHP) نقایصی دارد. به عنوان مثال عدم قطعیت مرتبط با فرایندها را در بر نمی‌گیرد. برای غلبه بر چنین مشکلاتی نظریه فازی با فرایند تحلیل سلسله مراتبی (AHP) در قالب فرایند تحلیل سلسله مراتبی فازی (FAHP) ادغام شده است (وینود و همکاران، ۲۰۱۴). روش مورد استفاده در این پژوهش برگرفته از روش میانگین هندسی باکلی^۱ است (هاسی^۲ و همکاران، ۲۰۰۴). گام‌های این روش در زیر آورده شده است.

فرض کنید P_{ij} مجموعه‌ای از ترجیحات تصمیم‌گیران در مورد یک شاخص نسبت به دیگر شاخص‌ها باشد. ماتریس مقایسه‌های زوجی به صورت زیر تشکیل می‌شود:

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

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

1. buckley
2. Hsieh

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

سپس وزن فازی هر (i) شاخص با 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، با رابطه زیر وزن‌ها را دیفازی کرده و سپس نرمال می‌کنیم.

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

در این پژوهش برای محاسبه وزن در مقایسه‌های زوجی، از عبارات کلامی و اعداد فازی مثلثی مندرج در جدول (۲) 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 عدد یک اهمیت برابر و عدد نه اهمیت فوق‌العاده را نشان می‌دهد، پیشنهاد کرد.

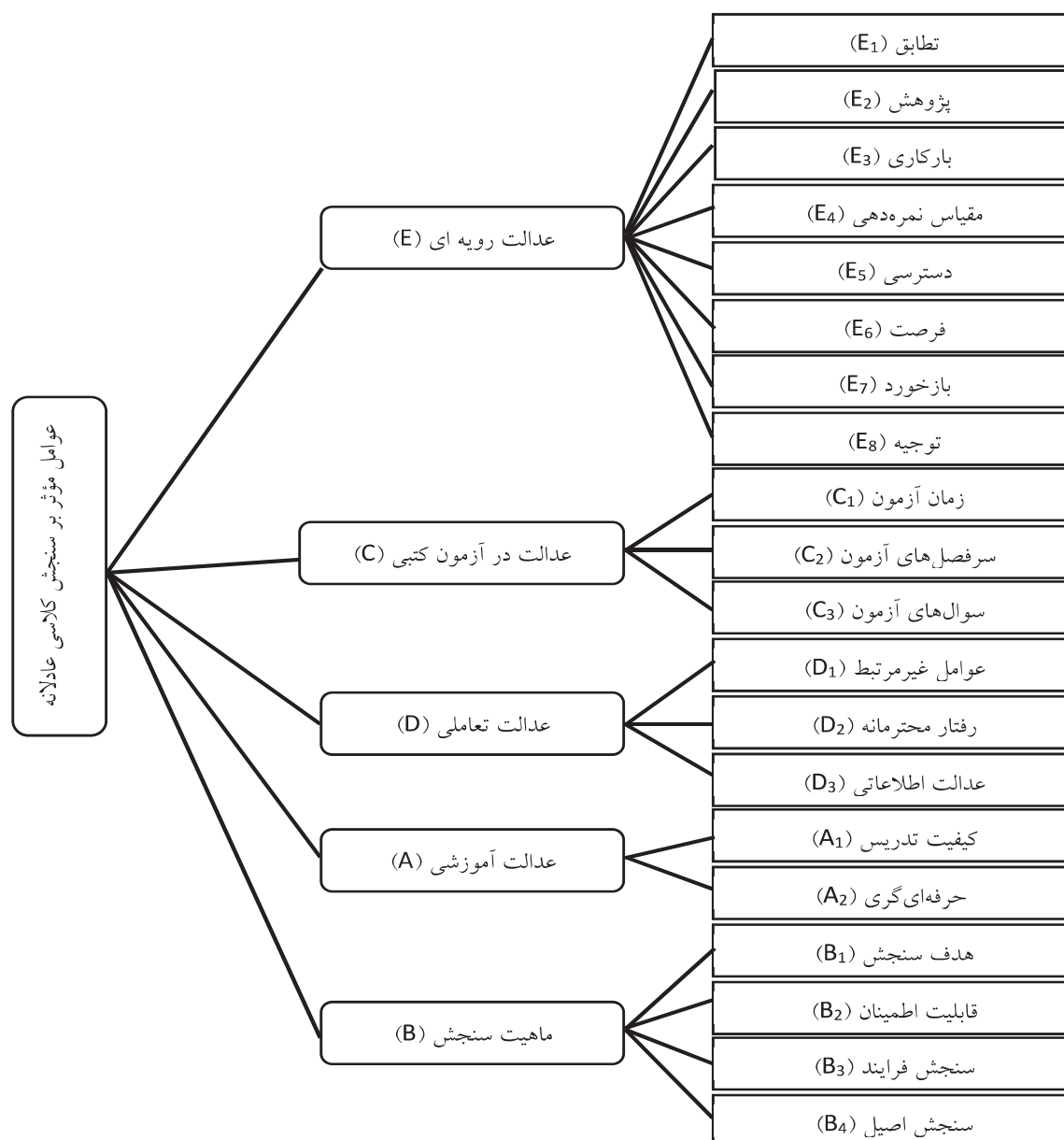
جدول ۲. عبارات کلامی و اعداد فازی برای وزن‌دهی به مضمون‌ها

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

یافته‌ها

مضمون‌ها و زیرمضمون‌ها

متن‌های پیاده‌سازی شده مصاحبه‌ها با استفاده از روش پدیدارشناسی، اقتباس شده از ریکور^۱ (۱۹۷۶) که شامل ساده خواندن، تحلیل ساختاری و درک جامع است، تحلیل شد. نتایج این بخش یعنی مضمون‌ها و زیرمضمون‌های سنجش کلاسی عادلانه، در درخت سلسه مراتبی ارائه شده است (شکل ۱).



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

1. Ricoeur

همچنین برای محاسبه روایی محتوایی از گروهی از متخصصان خواسته می‌شود تا هر مضمون را از نظر ارتباط با سازه 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^۱ (۱۹۸۶) حداقل سه متخصص را توصیه می‌کند اما بیش از ده نفر را ضروری نمی‌داند (ذکر شده در پالت و بک^۲، ۲۰۰۶). بر این اساس متخصصانی که ما برای رواسازی مضمون‌های به دست آمده استفاده کردیم شامل نه استاد دانشگاه بود. ملاک انتخاب این افراد، تخصص در حوزه سنجش و تجربه تدریس حداقل پنج سال در دانشگاه بود. ما ابتدا از متخصصان خواستیم تا میزان ارتباط هر مضمون با سازه «سنجش کلاسی عادلانه» را در یک طیف لیکرت چهار نقطه‌ای، «کاملاً مرتبط»، «بسیار مرتبط»، «تا حدودی مرتبط» و «مرتبط نیست» درجه‌بندی کنند. انتخاب مقیاس رتبه‌ای چهار نقطه‌ای به جای مقیاس درجه‌بندی سه یا پنج نقطه‌ای برای اجتناب از داشتن نقطه میانی خنثی و دوسویه بود. برای هر مضمون شاخص روایی محتوایی^۳ با تقسیم تعداد متخصصانی که به رتبه سه یا چهار داده‌اند بر تعداد کل متخصصان محاسبه می‌شود. برای اینکه خطای استاندارد نسبت در محاسبات وارد شود، لین (۱۹۸۶) پیشنهاد می‌کند با داشتن پنج متخصص یا کمتر همه باید در مورد روایی محتوا به توافق برسند. به عبارت دیگر شاخص روایی محتوایی آیتم باید برابر یک باشد و با شش متخصص یا بیشتر شاخص روایی محتوایی کمتر از $0/78$ را توصیه نمی‌کند (ذکر شده در پالت و بک^۴، ۲۰۰۶). در این پژوهش شاخص روایی محتوایی کلی به توصیه پالت و بک (۲۰۰۶) میانگین شاخص‌های روایی محتوایی زیرمضمون‌ها در نظر گرفته شد، چراکه به جای میانگین عملکرد کارشناسان، بر میانگین کیفیت زیرمضمون‌ها متمرکز است. بسیاری از پژوهشگران مقدار قابل قبول روایی محتوایی کلی را $0/80$ و بالاتر از آن می‌دانند.

هر نه کارشناس، زیرمضمون‌های بازخورد، فرصت، دسترسی، پژوهش‌های تخصیص داده شده به دانشجویان، بار کاری، 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 هر کدام از این زیرمضمون‌ها برابر $1 (9 \div 9)$ 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 هر کدام از این زیرمضمون‌ها برابر $0/88 (8 \div 9)$ است. بنابر این روایی محتوایی هر یک از زیرمضمون‌ها بالاتر از مقدار معیار $0/78$ است. روایی محتوایی کلی نیز بنا به تعریف برابر $0/94$ است $[20 \div (15 \times 1 + 5 \times 0/88)]$ ، که بالاتر از مقدار معیار $0/80$ است.

1. Lynn
2. Polit & Beck
3. Content Validity Index (CVI)
4. Polit & Beck

نتایج تحلیل سلسه‌مراتبی فازی

در مرحله قبل، عوامل پژوهش شناسایی و تأیید شدند. در این گام ابتدا مقایسه‌های زوجی مضمون‌ها تشکیل و در اختیار نُه پاسخ‌دهنده قرار داده شد. بعد از پاسخگویی به مقایسه‌های زوجی، نرخ ناسازگاری جدول‌ها محاسبه شد. همگی از ۰/۱ کوچک‌تر بود که نشان‌دهنده این است که ثبات و قابلیت اطمینان مقایسه‌های زوجی در حد قابل قبولی است. سپس با استفاده از روش میانگین هندسی، پاسخ‌ها ادغام شد. 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)	(۱/۴۵۶, ۱/۹۵۶, ۲/۴۵۴)	(۲/۶۳۱, ۳/۷۶۹, ۴/۸۴)	(۱/۶۴, ۲/۱۱۳, ۲/۶۷۷)	(۱/۲۰۴, ۱/۵۱۳, ۱/۹۰۱)
B	(۰/۴۰۸, ۰/۵۱۱, ۰/۶۸۷)	(۱,۱,۱)	(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)	(۱,۱,۱)	(۰/۱۹۷, ۰/۲۳۲, ۰/۲۸۴)	(۰/۲۴۴, ۰/۳۰۳, ۰/۴۱۱)
D	(۰/۳۷۴, ۰/۴۷۳, ۰/۶۱)	(۱/۱۲, ۱/۳۳۹, ۱/۵۴۱)	(۳/۵۲, ۴/۳۱۱, ۵/۰۷)	(۱,۱,۱)	(۳/۲۹۷, ۴/۳۵۸, ۵/۳۹۵)
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, ۳/۶۸۷, 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. مقایسه‌های زوجی زیرمضمون‌های عدالت رویه‌ای (نرخ ناسازگاری: ۰/۰۳)

	E1	E2	E3	E4	E5	E6	E7	E8
E1	(۱,۱,۱)	(۰/۲۱, ۰/۲۷۲) (-۰/۱۷۳)	(۰/۵۲۱, ۰/۶۷۲) (-۰/۴۲۴)	(۰/۴۹۲, ۰/۶۲) (-۰/۳۸۹)	(۰/۲۷, ۰/۳۲۲) (-۰/۲۳۵)	(۰/۲۷, ۰/۳۶۲) (-۰/۲۲۱)	(۰/۳۹۸, ۰/۳۹) (-۰/۲۲۷)	(۰/۲۳۲, ۰/۲۸۸) (-۰/۱۹۷)
E2	(۴/۷۵۲, ۵/۷۹۵) (۳/۶۸۲)	(۱,۱,۱)	(۱/۲۳۹, ۱/۴۹۶) (۱)	(۰/۶۷۱, ۰/۸۱۹) (-۰/۵۲۹)	(۰/۶۷, ۰/۷۷۴) (-۰/۵۷۳)	(۰/۶۹۴, ۰/۷۹۹) (-۰/۶۰۸)	(۱/۴۵۱, ۱/۸۹۵) (۱/۱۰۷)	(۰/۵۹۹, ۰/۷۹۹) (-۰/۴۵۸)
E3	(۱/۹۱۹, ۲/۳۵۶) (۱/۴۸۹)	(۰/۸۰۷, ۱) (-۰/۶۶۹)	(۱,۱,۱)	(۰/۸۸۳, ۱/۱۰۴) (-۰/۶۸۹)	(۰/۵۲۴, ۰/۶۱۹) (-۰/۴۴۵)	(۰/۵۰۱, ۰/۶۳۴) (-۰/۴۱۲)	(۰/۶۷۷, ۰/۸۷۵) (-۰/۵۰۷)	(۰/۶۶۳, ۰/۹۳۱) (-۰/۴۸۶)
E4	(۲/۰۳۵, ۲/۵۶۸) (۱/۶۱۲)	(۱/۴۹, ۱/۸۵۴) (۱/۲۲)	(۱/۱۳۲, ۱/۴۵۱) (-۰/۹۰۶)	(۱,۱,۱)	(۰/۶۴۲, ۰/۸۳۵) (-۰/۵۱۷)	(۰/۸۶, ۱/۱۲۶) (-۰/۶۷۲)	(۱/۷۴۱, ۲/۰۳۵) (۱/۴۶۴)	(۰/۴۶۸, ۰/۶۵۷) (-۰/۳۶)
E5	(۲/۰۳۶, ۲/۲۵۹) (۳/۱۱)	(۱/۴۹۳, ۱/۷۴۴) (۱/۲۹۳)	(۱/۹۰۹, ۲/۲۴۹) (۱/۶۱۶)	(۱/۵۵۸, ۱/۹۳۴) (۱/۱۹۸)	(۱,۱,۱)	(۰/۴۵۲, ۰/۵۵۳) (-۰/۳۸۵)	(۱/۲۴۲, ۱/۶۰۸) (-۰/۹۷۴)	(۰/۸۵۳, ۱/۰۲) (-۰/۶۷۴)
E6	(۲/۶۹۷, ۴/۵۲۴) (۲/۷۶۵)	(۱/۴۴۲, ۱/۶۴۴) (۱/۲۵۱)	(۱/۹۹۷, ۲/۴۲۹) (۱/۵۷۶)	(۱/۱۶۳, ۱/۴۸۹) (-۰/۸۸۸)	(۲/۲۱, ۲/۵۹۷) (۱/۸۰۷)	(۱,۱,۱)	(۱/۸۵۲, ۲/۲۴) (۱/۴۳۸)	(۰/۴۱۶, ۰/۶۳) (-۰/۳۰۷)
E7	(۲/۳۵۶, ۴/۲۱۹) (۲/۵۶۵)	(۰/۶۸۹, ۰/۹۰۳) (-۰/۵۲۸)	(۱/۴۷۶, ۱/۹۷۳) (۱/۱۴۳)	(۰/۵۷۴, ۰/۶۸۳) (-۰/۴۹۲)	(۰/۸۰۵, ۱/۰۳۷) (-۰/۶۲۲)	(۰/۵۴, ۰/۶۹۶) (-۰/۴۴۷)	(۱,۱,۱)	(۰/۳۲۶, ۰/۴۵۷) (-۰/۲۵۲)
E8	(۴/۲۹۴, ۵/۰۷) (۳/۴۷۴)	(۱/۶۷, ۲/۱۸۲) (۱/۲۵۱)	(۱/۵۰۹, ۲/۰۴۶) (۱/۰۷۴)	(۲/۱۳۵, ۲/۷۸) (۱/۵۲۲)	(۱/۱۷۳, ۱/۴۸۴) (-۰/۹۸)	(۲/۴۰۶, ۲/۲۵۹) (۱/۵۸۷)	(۲/۰۶۴, ۳/۹۷۱) (۲/۱۸۹)	(۱,۱,۱)

ب- محاسبه وزن‌های فازی و نرمال

در این گام بر اساس رابطه (۱) و (۲) ابتدا میانگین هندسی اعداد فازی هر سطر از جدول‌های ۳ تا 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. به عنوان مثال برای مضمون A در جدول (۳) محاسبات به صورت زیر است:

میانگین هندسی سطر اول

$$= [(1,1,1) \times (1.456,1.956,2.454) \times (2.631,3.769,4.84) \times (1.64,2.113,2.677) \times (1.204,1.513,1.901)]^{\frac{1}{5}} = (1.499,1.882,2.271)$$

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

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

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

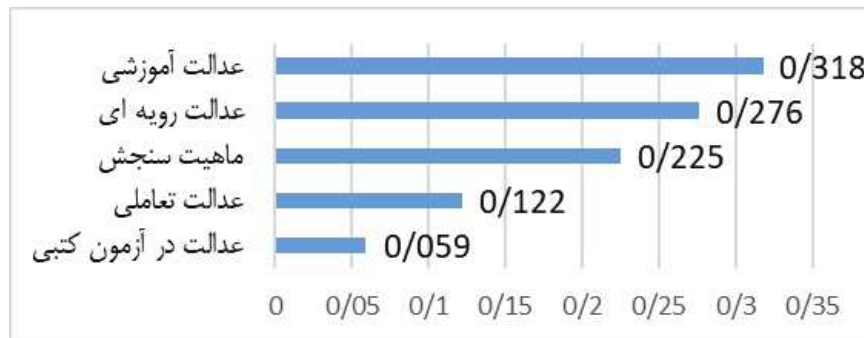
$$A \text{ وزن غیر فازی} = (0.213,0.319,0.466) \Rightarrow A \text{ وزن غیر فازی} = \frac{0.213 + 2 \times 0.319 + 0.416}{4} = 0.329$$

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 شیوه زیر عمل می‌کنیم:

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

جدول ۹. وزن فازی و غیرفازی مضمون‌های اصلی

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \tilde{P}_{ij})^{1/n})$	وزن فازی (\tilde{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)	۰/۲۲۲	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
$\sum (\prod_{j=1}^n \tilde{P}_{ij})^{1/n}$	(۴/۸۷۷، ۵/۹۰۳، ۷/۰۴)			

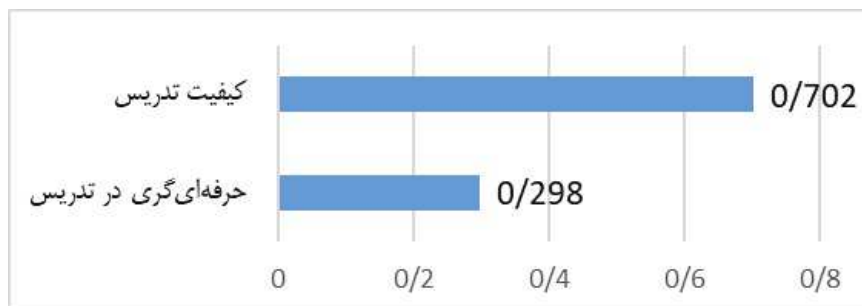


شکل ۲. وزن‌های مضمون‌های اصلی

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. این محاسبات برای دیگر مقایسه‌های زوجی (زیرمضمون‌ها) به روش مشابه صورت می‌گیرد که در ادامه آورده شده است.

جدول ۱۰. وزن فازی و غیرفازی زیرمضمون‌های عدالت آموزشی

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \tilde{P}_{ij})^{1/n})$	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
$\sum (\prod_{j=1}^n \tilde{P}_{ij})^{1/n}$	(۱/۹۸۹، ۲/۱۸۹، ۲/۳۹۳)			

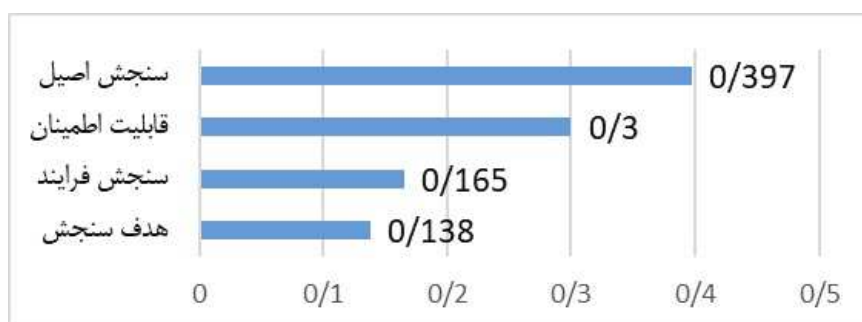


شکل ۳. وزن‌های زیرمضمون‌های عدالت آموزشی

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 $(\prod_{j=1}^n \beta_{ij})^{1/n}$	وزن فازی (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
$\sum \left(\prod_{j=1}^n \beta_{ij} \right)^{1/n}$	(۳/۶۴، ۴/۳۸۲، ۵/۲۴۱)			

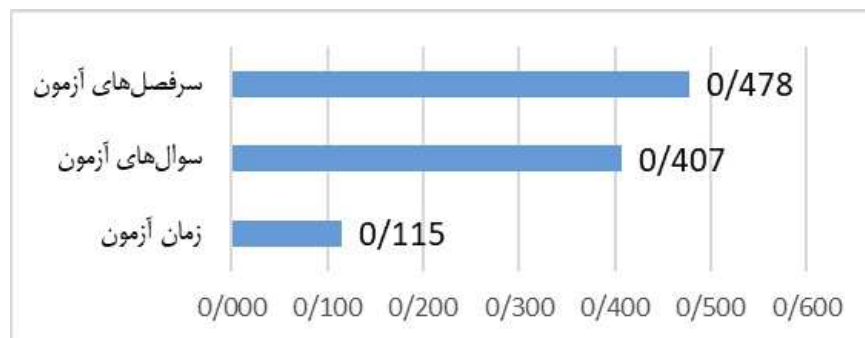


شکل ۴. وزن‌های زیرمضمون‌های ماهیت سنجش

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

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

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \beta_{ij})^{1/n})$	وزن فازی (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
$\sum \left(\prod_{j=1}^n \beta_{ij} \right)^{1/n}$	(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 ۴۷۸/۰ رتبه اول را کسب کرده است. سؤال‌های آزمون با وزن ۴۰۷/۰ رتبه دوم و زمان آزمون با وزن ۱۱۵/۰ رتبه سوم را کسب کرده است.

جدول ۱۳. وزن فازی و غیرفازی زیرمضمون‌های عدالت تعاملی

نام مضمون	میانگین هندسی $((\prod_{j=1}^n \beta_{ij})^{1/n})$	وزن فازی (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
$\sum \left(\prod_{j=1}^n \beta_{ij} \right)^{1/n}$	(۲/۶۲۲، ۳/۰۴۱، ۳/۵۲)			

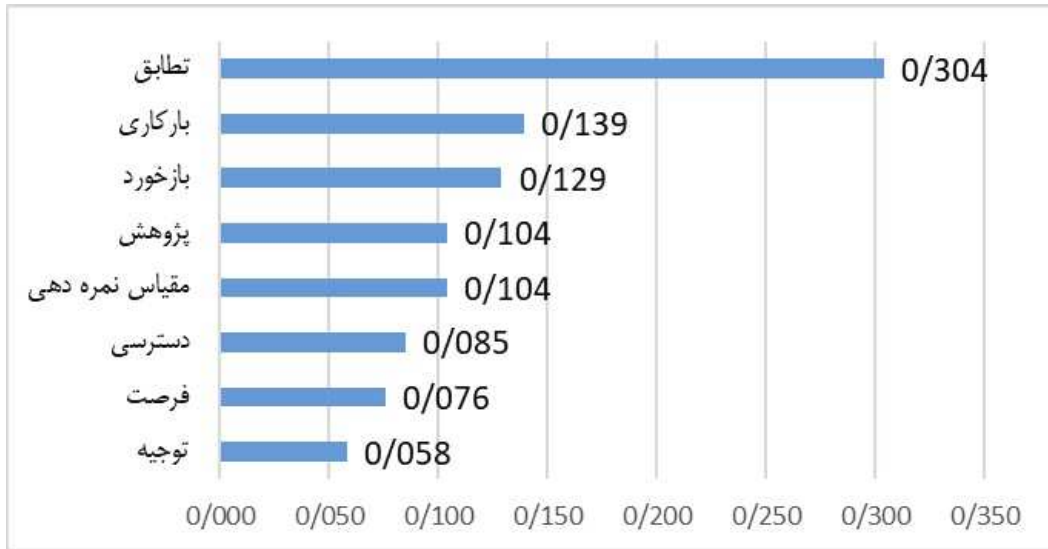


شکل ۶. وزن‌های زیرمضمون‌های عدالت تعاملی

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

جدول ۱۴. وزن فازی و غیرفازی زیرمضمون‌های عدالت رویه‌ای

نام مضمون	میانگین هندسی $(\prod_{j=1}^n P_{ij})^{1/n}$	وزن فازی (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
$\sum \left(\prod_{j=1}^n P_{ij} \right)^{1/n}$	(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 ۳۰۴/۰ رتبه اول را کسب کرده است. بارکاری با وزن ۱۳۹/۰ رتبه دوم و بازخورد با وزن ۱۲۹/۰ 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	هدف سنجش	۰/۲۲۵	عوامل ماهیت سنجش
۷	۰/۰۶۷۵	۰/۳۰۰	قابلیت اطمینان		
۸	۰/۰۳۷۱	۰/۱۶۵	سنجش فرایند		
۴	۰/۰۸۹۳	۰/۳۹۷	سنجش اصیل		

رتبه	وزن نهایی زیرمضمون	وزن نسبی زیرمضمون	زیرمضمون	وزن مضمون	مضمون
۲۰	۰/۰۰۶۸	۰/۱۱۵	زمان آزمون	۰/۰۵۹	عوامل عدالت در آزمون کتبی
۱۱	۰/۰۲۸۲	۰/۴۷۸	سرفصل‌های آزمون		
۱۲	۰/۰۲۴۰	۰/۴۰۷	سوال‌های آزمون		
۶	۰/۰۷۵۹	۰/۲۷۵	عوامل غیرمرتبط	۰/۲۷۶	عوامل عدالت رویه ای
۵	۰/۰۸۸۳	۰/۳۲۰	رفتار محترمانه		
۲	۰/۱۱۱۵	۰/۴۰۴	عدالت اطلاعاتی		
۹	۰/۰۳۷۱	۰/۳۰۴	تطابق	۰/۱۲۲	عدالت تعاملی
۱۵	۰/۰۱۲۷	۰/۱۰۴	مقیاس نمره دهی		
۱۳	۰/۰۱۷۰	۰/۱۳۹	بارکاری		
۱۵	۰/۰۱۲۷	۰/۱۰۴	پژوهش		
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 و در نهایت با تحلیل سلسله‌مراتبی فازی اولویت‌بندی شد. با تحلیل‌های صورت گرفته مشخص شد که در بین مضمون‌های پنج‌گانه اصلی، عدالت آموزشی بیشترین وزن را دارد و زیرمضمون کیفیت تدریس، در این مضمون و همچنین در مقایسه با تمام زیرمضمون‌های شناسایی شده، بیشترین وزن به خود اختصاص داده است. عدالت آموزشی در این پژوهش به عدم موفقیت استاد در ایفای نقش خود به عنوان مربی با تمرکز بر فعالیت‌های تدریس اشاره می‌کند و چنانچه بمپچت^۱ 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^۲ (۲۰۱۷) نیز اشاره می‌کند، نقض عدالت آموزشی در بُعد کیفیت تدریس، به

1. Bempechat
2. McMillan

درک ناعادلانهٔ 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. همان‌طور که روداباو^۱ (۱۹۹۴) نشان می‌دهد، دانشجویان بیش از دریافت نمره‌های بالاتر، نگران عدالت در کلاس هستند؛ بنابر این، استادی که نمره‌های کمتری می‌دهد اما از رویه‌های عادلانه استفاده می‌کند، نسبت به کسی که نمره‌های بالاتری می‌دهد اما از رویه‌های ناعادلانه استفاده می‌کند، عادلانه‌تر ادراک می‌شود.

در مضمون «ماهیت سنجش» به عنوان مضمونی که از نظر وزن‌های تخصیص داده شده در جایگاه سوم قرار می‌گیرد، زیرمضمون «سنجش اصیل»^۲ بالاترین وزن را به خود اختصاص داده است. در دههٔ گذشته، برخی از روندهای سنجش کلاسی برای سازگاری بهتر با دانش، مهارت‌ها و شرایط قرن بیست و یکم ایجاد شده است که تغییر از فرهنگ آزمودن^۳ به فرهنگ یادگیری^۴ را موجب شده است. این روندها باعث کاربرد سنجش‌های جایگزین، از جمله سنجش اصیل، سنجش عملکرد^۵، نمونه کارها، نمایشگاه‌ها، نمایش‌ها، پژوهش‌ها و موارد پیشرفته فن‌آوری 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، چرا که بر موقعیت‌های زندگی واقعی و ایجاد توانمندی‌های لازم برای رویارویی دانشجویان با آنچه بعد از تحصیل با آن روبه‌رو خواهند شد، تأکید می‌کند. سنجش اصیل، باعث بهبود یادگیری و افزایش انگیزهٔ دانشجویان می‌شود. علاوه بر این، توانایی بهبود عملکرد دانشجویان و فراهم آوردن فرصت‌های یادگیری بیشتر برای آنها را دارد. استفاده از سنجش چندگانه به جای سنجش واحد، اضطراب دانش‌آموز را کاهش می‌دهد و در نتیجه خطرپذیری، قدرت خلاقیت و اکتشاف وی را افزایش می‌دهد (موریلو و هیدالگو^۶، ۲۰۱۷؛ تیرنی، ۲۰۱۳).

در مضمون «عدالت تعاملی» با تأکید بر ماهیت رابطهٔ استاد و دانشجو و با تمرکز بر کیفیت رفتار استاد و دانشجو هنگام اجرای رویه‌ها یا تخصیص نتایج (بمپیچت و همکاران، ۲۰۱۳)، زیرمضمون «عدالت اطلاعاتی» بیشترین وزن را دارد. عدالت اطلاعاتی دربرگیرنده نشانگرهای سنجش، انتظارات استاد از دانشجویان و عمل بر اساس ملاک‌های توافق شده توسط استاد است. ما این واحدهای معنایی را زیر عنوان «عدالت اطلاعاتی» قرار داده‌ایم. عدم توضیح استاد دربارهٔ روش‌ها و شیوه‌های نمره‌دهی، تعریف واضح و روشن انتظارات و سیاست‌های دوره در شروع آن توسط استاد، عدم ارائه راهنمای مشخص برای اختصاص نمره‌ها به‌ویژه برای حضور و مشارکت در کارگروهی، مشخص کردن موادی که برای آماده شدن هر 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 برنامهٔ

1. Rodabaugh
2. authentic assessment
3. Testing culture
4. Learning culture
5. performance assessment
6. Murillo & Hidalgo

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 خود در یک دوره مطلع شوند، برای رسیدن به اهداف، فرصت خودسنجی 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^۱، ۱۹۹۹). البته توجه به این مطلب ضروری است که تصریح ملاک‌ها با توجه به تأثیر نمره‌ها در آینده شغلی و تحصیلی یادگیرندگان نیز قابل بررسی است. چنان که تونگ^۲ (۲۰۰۹) معتقد است، دانشجویان عصر جدید، انتظار دارند مدرسان آنچه را آنها برای کسب نمره‌های بالا باید انجام دهند به صراحت و به طور دقیق 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، بیشتر وزن بر فرایندها، ماهیت سنجش و رویه‌ها قرار می‌گیرد. چنانچه جانسون و اسوینگی^۳ (۲۰۰۷) نیز تأکید می‌کنند فرهنگ جدید سنجش بر ارزیابی فرایندها و شایستگی‌های تفکر در سطح بالاتر به جای 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 بیش از حد بر ارزشیابی پایانی، استفاده نکردن از ارزشیابی فرایندی و مستمر، استفاده نکردن از ارزشیابی‌های غیر عینی، بی توجهی به تفاوت‌های فردی، بی توجهی به مجموعه هدف‌های یادگیری و ناآشنایی اساتید با فلسفه و هدف ارزشیابی، چالش بی توجهی به تفاوت‌های فردی بیشترین و چالش استفاده نکردن از ارزشیابی‌های غیر عینی کمترین بار عاملی را داشت.

با عنایت به اینکه بر اساس نقشه جامع علمی کشور (۱۳۹۰) تربیت انسان‌های مؤمن، توانا، خردمند، سالم، خلاق، آزاده، دارای فضائل اخلاقی و اعتماد به نفس، تقویت ارزش‌های معنوی و فضائل اخلاقی، ترویج فرهنگ آزاداندیشی، نقادی و نقدپذیری از وظایف نهاد علم است؛ ملاک‌ها و نشانگرهای به دست آمده از این پژوهش می‌تواند به عنوان یکی از منابع تدوین شیوه‌نامه ملی سنجش کلاسی عادلانه برای کمک به تحقق عدالت آموزشی استفاده شود 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. با این هدف استفاده از نشانگرها و ملاک‌های سنجش کلاسی عادلانه در ارزشیابی اعضای هیأت علمی می‌تواند استفاده شود. 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 اعضای هیأت علمی پیشنهاد می‌شود.

1. Houston & Bettencourt
2. Twenge
3. Jonsson & Svingby

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 Principles and Process of Construction and Validation of the Interview Protocol. *Higher Education Letter*, 12(46), 177-203. (Persian)
- Baniasadi, A., Salehi, K., Khodaie, E., 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, M. R., Karimi, I. (2015). Effect of the Challenges of Student Learning 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.

