



Construction of a Standardized Questionnaire to Measure the Culture of Evaluation in the Higher Education System: A Mixed Research of Instrument Development Model

Morteza Malmir¹ | Keyvan Salehi^{2*} | Ali Moghadam Zadeh³ | Marzieh Dehghani⁴

1. Department of Educational Measurement and Evaluation, University of Tehran. Email: Morteza.malmir@ut.ac.ir
2. Corresponding Author, Department of Curriculum Development & Instruction Methods, University of Tehran. Email: Keyvansalehi@ut.ac.ir
3. Department of Curriculum Development & Instruction Methods, University of Tehran. Email: amoghadamzadeh@ut.ac.ir
4. Department of Curriculum Development & Instruction Methods, University of Tehran. Email: dehghani_m33@ut.ac.ir

ARTICLE INFO

Article type:
Research Article

Received: 31 December 2023
Revised: 25 March 2024
Accepted: 20 April 2024
Published online: 24 February 2026

Keywords:
Evaluation of culture,
Quality culture,
Construction,
IRT,
Quality assurance.

ABSTRACT

Objective: In this research, the conceptualization and construction of a scaled tool for measuring quality of culture and evaluation was discussed. **Methods:** This research was based on the exploratory mixed method of the instrument development model. In the phenomenology stage, with the aim of analyzing the lived experiences of Beneficiaries and stakeholders, after interviewing 61 participants, data saturation was achieved. In the quantitative part, which was based on the correlation method, a total of 510 university activists, including faculty members, teaching and executive staff, and university students, participated in the study in an accessible way. In the construction of the tool, the psychometric characteristics were examined based on Classical Test Theory (CTT) and Item Response Theory (IRT). **Results:** At first, content and construct validity were examined and confirmed using factor analysis; In the following, reliability was obtained using McDonald's omega, yielding an overall value of 0.869 for the culture assessment and subscale omegas exceeding 0.70. In the analysis of the items based on the IRT, the findings showed that the partial score model (PCM) has a better fit with the data of the evaluation of cultural questionnaire. **Conclusion:** Based on the obtained results, this questionnaire can be used to measure the levels of evaluation of culture in the academic environment.

Cite this article: Malmir, M.; Salehi, K.; Moghadam Zadeh, A. & Dehghani, M. (2026). Construction of a Standardized Questionnaire to Measure the Culture of Evaluation in the Higher Education System: A Mixed Research of Instrument Development Model. *Organizational Culture Management*, 24 (1), 1-17. <http://doi.org/10.22059/jomc.2024.371651.1008636>



© Authors retain the copyright and full publishing rights.
DOI: <http://doi.org/10.22059/jomc.2024.371651.1008636>

Publisher: University of Tehran Press.



ساخت پرسشنامه‌ای میزان شده برای اندازه‌گیری فرهنگ ارزشیابی در آموزش عالی: یک پژوهش آمیخته، مدل توسعه ابزار

(ابزاری میزان شده برای سنجش فرهنگ ارزشیابی در سیستم آموزش عالی)*

مرتضی مالمیر^۱ | کیوان صالحی^{۲*} | علی مقدم‌زاده^۳ | مرضیه دهقانی^۴

۱. گروه سنجش و اندازه‌گیری، گروه روش‌ها و برنامه‌های درسی و آموزشی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. رایانامه: morteza.malmir@ut.ac.ir
۲. نویسنده مسئول، گروه سنجش و اندازه‌گیری، گروه روش‌ها و برنامه‌های درسی و آموزشی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. رایانامه: keyvansalehi@ut.ac.ir
۳. گروه سنجش و اندازه‌گیری، گروه روش‌ها و برنامه‌های درسی و آموزشی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. رایانامه: amoghadamzadeh@ut.ac.ir
۴. گروه برنامه‌ریزی درسی، گروه روش‌ها و برنامه‌های درسی و آموزشی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. رایانامه: dehghani_m33@ut.ac.ir

اطلاعات مقاله

چکیده

نوع مقاله:

مقاله پژوهشی

هدف: این پژوهش با هدف مفهوم‌پردازی و ساخت ابزاری میزان شده برای اندازه‌گیری فرهنگ ارزشیابی شکل گرفت.

تاریخ‌های مقاله:

تاریخ دریافت: ۱۴۰۲/۱۱/۱۰

تاریخ بازنگری: ۱۴۰۳/۰۱/۰۶

تاریخ پذیرش: ۱۴۰۳/۰۲/۰۱

تاریخ انتشار: ۱۴۰۴/۱۲/۰۵

روش پژوهش: این پژوهش مبتنی بر روش آمیخته اکتشافی مدل توسعه ابزار بود. در مرحله پدیدارشناسی، با هدف تحلیل تجارب زیسته یاران آموزشی، ادراک چهار طبقه شامل اعضای هیئت علمی، کارکنان، مدیران، و دانشجویان دانشکده‌های مختلف دانشگاه پس از مصاحبه با ۶۱ شرکت‌کننده اشباع داده حاصل شد. در بخش کمی مبتنی بر روش همبستگی انجام شد و ۵۱۰ نفر از کُشگران دانشگاهی شامل اعضای هیئت علمی و کادر آموزش و اجرایی و دانشجویان دانشگاه به شیوه در دسترس در مطالعه شرکت کردند. در ساخت ابزار، ویژگی‌های روان‌سنجی بر اساس نظریه کلاسیک آزمون (CTT) و نظریه سؤال-پاسخ (IRT) مورد بررسی قرار گرفت.

کلیدواژه:

تضمین کیفیت،

ساخت ابزار،

فرهنگ ارزشیابی،

فرهنگ کیفیت،

نظریه سؤال-پاسخ.

یافته‌ها: در ابتدا روایی محتوایی و سازه با استفاده از تحلیل عاملی مورد بررسی و تأیید قرار گرفت. در ادامه اعتبار با استفاده از روش آمگا مک دونالد در فرهنگ ارزشیابی به صورت کلی ۰/۸۶۹ و در همه خرده‌مقیاس‌ها بالای ۰/۷ به دست آمد. در تحلیل گویه‌ها مبتنی بر نظریه سؤال-پاسخ، یافته‌ها نشان داد مدل امتیاز پاره‌ای (PCM) با داده‌های پرسشنامه فرهنگ ارزشیابی برازش مناسب‌تری دارد.

نتیجه: بر اساس نتایج به‌دست‌آمده و ویژگی‌های مطلوب روان‌سنجی از ابزار ساخته‌شده، از این پرسشنامه می‌توان برای سنجش سطوح فرهنگ ارزشیابی در محیط دانشگاهی استفاده کرد.

استناد: مالمیر، مرتضی؛ کیوان، مقدم‌زاده، علی و دهقانی، مرضیه (۱۴۰۵). ساخت پرسشنامه‌ای میزان شده برای اندازه‌گیری فرهنگ ارزشیابی در آموزش عالی: یک پژوهش آمیخته، مدل توسعه ابزار (ابزاری میزان شده برای سنجش فرهنگ ارزشیابی در سیستم آموزش عالی). مدیریت فرهنگ سازمانی، ۲۴ (۱) ۱-۱۷. <http://doi.org/10.22059/jomc.2024.371651.1008636>

ناشر: مؤسسه انتشارات دانشگاه تهران.

© نویسندگان.

DOI: <http://doi.org/10.22059/jomc.2024.371651.1008636>



* این مقاله برگرفته از رساله دکتری رشته سنجش و اندازه‌گیری دانشگاه تهران است.

مقدمه

آموزش با رشد و توسعه ارتباط دارد و زیربنای پروژه‌های بزرگ برای پیشرفت جوامع است (Chen, 2023). بررسی‌ها در این زمینه گویای نقش کلیدی و مرتبط آموزش در حیطه‌های مختلف و توسعه پایدار و الگوی رشد است (Angiel & Pokojski, 2019; 8: 2016; Kalinowska et al.). بر این اساس آموزش عالی نقشی مهم در توسعه منابع انسانی، رشد اقتصادی، و توسعه کشورها دارد (Jalaliyoon & Taherdoost, 2012). دانشگاه مهم‌ترین دستگاه آموزش عالی است که فرهنگ و الگوی سیاستی اتخاذشده در آن بسیار بااهمیت است و ساختار حاکم بر آن تأثیری چندوجهی در زمینه‌های آموزشی، کاربردی، آینده‌نگرایانه، و ... دارد (Rospigliosi, 2019; Gandini, 2016). رشد و پیشرفت آموزش عالی در گرو یک برنامه انعطاف‌پذیر اجرایی و پذیرش سیستم برای تغییر متناسب با شرایط است و چنین هدفی با ارزشیابی و ارزیابی در نظام آموزش عالی و متغیرهای مرتبط با آن ممکن می‌شود (Brown, 2019). بر همین اساس، یکی از ارکان اساسی رشد و توسعه اقتصادی، اجتماعی، فرهنگی، و سیاسی کشورها آموزش است. همه صاحب‌نظران عقیده دارند که اگر اجرای نظام صحیح آموزش موفق باشد، بر بسیاری از مشکلات و معضلات فائق می‌آید و جامعه را به سوی پیشرفت و ترقی در بسیاری از زمینه‌ها سوق خواهد داد. اجرای نظام صحیح آموزش مستلزم برنامه‌ریزی‌های دقیق علمی و اجرای آن‌هاست. بی‌شک یکی از اقدامات مسئولان برای آگاه شدن از چگونگی میزان تحقق اهداف برنامه خود ارزشیابی است (بنی‌سی و همکاران، ۱۳۹۶). ارزشیابی عاملی مهم در ارتقای کیفی سیستم دانشگاهی است (Han, 2011) و بررسی وضعیت کارکنان، اعضای دانشگاه، و افراد درگیر در سیستم آموزشی بسیار مهم است (Sun, 2020). هوانگ^۱ (۲۰۲۲) از مدل ارزیابی جامع فازی ANP- برای ایجاد یک سیستم شاخص برای ساختن دانشگاه‌های سطح بالا استفاده کردند و آن را عامل مهم کیفیت‌افزایی در سیستم دانشگاهی در نظر گرفتند. بر این اساس باید دانشگاه را هدف ارزشیابی آموزشی قرار داد و از باتوم ارزشیابی حداکثر بهره را برد تا نشاط و توان دانشگاه‌ها به طور مستمر بهبود پیدا کند (Lee, 2023). بر این اساس گرچه یک نگاه واحد به اهمیت مسئله ارزشیابی سیستمی و جامع در بهبود کیفیت در هر سیستم آموزشی، صنعتی و تجاری وجود دارد، ابزارسازی تشخیصی دقیق در این زمینه با هدف شناخت فرهنگ ارزشیابی مسلط در نظام‌های سازمانی و آموزشی وجود ندارد و به همین سبب پژوهش حاضر بررسی در این مفهوم را بر اساس ساخت ابزار تشخیصی در این زمینه مد نظر دارد.

بیان مسئله

در سال‌های اخیر و هم‌سو با برجسته شدن نقش آموزش عالی در توسعه و پیشرفت جوامع، ارزشیابی آموزش عالی مورد توجه بیشتری قرار گرفته است. آستین^۲ و آنتونیو^۳ (۲۰۱۲) سیستم شاخص ارزیابی سطح توسعه آموزش عالی را مطرح کردند و روش ارزیابی مقرون‌به‌صرفه بودن مؤسسات آموزش عالی بر اساس مدل تحلیل پوششی داده‌های مرکب با به‌کارگیری رویکردی نوین در این زمینه را مورد تأکید قرار دادند. پرودنته^۳ و همکارانش (۲۰۲۳) با ذکر اهمیت آموزش عالی ارزشیابی کارآمد را عاملی مهم در توسعه منابع انسانی آن، به‌ویژه اعضای هیئت علمی، دانستند. مالمیر و همکارانش (۱۴۰۰) ارزشیابی کارآمد را در گرو یک سیستم کیفی صحیح درون‌دانشگاهی و به‌کارگیری فرهنگ کیفیت ارزیابی کردند. بر همین اساس مطالعات گسترده‌ای از سهم انکارناپذیر فرهنگ کیفیت و نظام ارزشیابی کارآمد در تعالی دانشگاهی و تعالی جامعه حکایت دارد (Bognar & Bungić, 2014; Zhang & Guo, 2015; Balesh et al., 2014) و در این حیطه ارزشیابی نقشی چندوجهی در بهبود آموزش و اصلاح شرایط دارد (Huber & Skedsmo, 2017). ارزشیابی با هدف ارائه بازخورد و اصلاح شرایط بر مبنای نیازهای موجود است که همین مسئله ضرورت وجود و بررسی عوامل و شرایط اثرگذار در پویایی، بالندگی، و اثرگذاری آن را برجسته می‌سازد (Kim, 2015).

با توجه به آنچه تشریح شد ارزشیابی فرایندی فراگیر و ضروری برای همه سازمان‌ها است و برای مؤسسات آموزش عالی، که عدالت و تعالی را هدف خود قرار داده‌اند، ضرورتی فزون‌تر دارد. محوری‌ترین ارزش‌نهادهای دانشگاهی در درجه نخست بهبود کیفیت یاددهی و یادگیری است. از این‌رو، برنامه‌های ارزشیابی ضرورتاً بخش بااهمیتی از فعالیت‌های آنان را تشکیل می‌دهد.

1. Huang
2. Astin
3. Prudente

تضمین کیفیت یاددهی- یادگیری در نظام دانشگاهی مستلزم نگاه جدی به تمام عوامل درگیر با فرایند یاددهی- یادگیری است و ارزشیابی و به عبارت مشخص تر فرهنگ ارزشیابی نقشی بسیار مهم و اساسی در اجرای باکیفیت ارزشیابی دارد. بررسی ها گویای این مسئله است که اگر ساختاری می خواهد اصلاح شود در قدم اول باید فرهنگ تغییر و فرهنگ ارزشیابی و شناخت نقاط قوت و ضعف بر جامعه حاکم باشد و این فرهنگ زمینه اجرای یک ارزشیابی صحیح و دقیق را مشخص می کند (Mayne, 2008: 6). فرهنگ ارزشیابی ناکارآمد بسیاری از تلاش ها برای ایجاد یک ارزیابی مؤثر و الگوی اصلاحی بر اساس آن را تضعیف و بی کیفیت می سازد (مالمیر و همکاران، ۱۴۰۲) و بر همین اساس شناخت فرهنگ ارزشیابی و عناصر مربوط به آن امری مهم و اثرگذار در به کارگیری فرهنگ ارزشیابی کارآمد و مبتنی بر بهبود است.

بر اساس پژوهش ماین (۲۰۱۰) فرهنگ ارزشیابی اشاره به سیستم ارزشیابی درون سازمانی و همراهی افراد درون سازمان دارد و ویژگی های یک سازمان با فرهنگ ارزشیابی قوی را به شرح زیر خلاصه می کند: اولاً، سازمان درگیر خودبازتابی و خودآزمایی در آن انجام می پذیرد و از اطلاعات به دست آمده برای بهبود تصمیم ها استفاده می کند. ثانیاً، سازمان درگیر یادگیری مبتنی بر شواهد می شود، به این معنی که یادگیری را در اولویت قرار می دهد و درگیر دانش می شود. ثالثاً، سازمان آزمایش و تغییر را تشویق می کند؛ ریسک می کند و به دنبال راه های جدیدی برای بهبود در شرایط سازمانی است (Mayne, 2010). ویژگی های بسیار رایجی که فرهنگ ارزشیابی ناکارآمد را به تصویر می کشد شامل جمع آوری اطلاعات سازمانی اما استفاده از آن عمدتاً برای گزارش دهی خارجی است و در این شرایط ارزشیابی در جهت بهبود کیفیت در سازمان نیست (Mayne, 2020). بر این اساس فرهنگ ارزشیابی مسئله ای چندبعدی است و بر اساس یک الگوی درونی از سیستم باورها و ارزش های در ارتباط با مفهوم ارزشیابی شکل می گیرد و تمامی افراد سازمان با هر سمت و نقشی در میزان و الگوی حاکم بر آن نقش دارند (Chaytor, 2023).

آنچه تشریح شد بیانگر نقش بسیار مهم فرهنگ ارزشیابی کارآمد در بهبود عملکرد سیستم آموزشی و تضمین کیفیت است. فرهنگ ارزشیابی در هر محیط آموزشی و سیستم سازمانی وجود دارد و حاکم بودن هر الگوی فرهنگ ارزشیابی پیامدهایی دارد که عموماً شناختی نسبت به این تفاوت و اثر هر یک از این ابعاد فرهنگ نشده است. آنچه مسئله مهم پژوهش حاضر و تحلیلی در این زمینه است خلأ ابزاری برای سنجش فرهنگ ارزشیابی و نوین بودن مسئله فرهنگ ارزشیابی است، حال آنکه نظر به اهمیت بسیار بالای این مفهوم سنجش مناسب و کارآمدی در این زمینه در مطالعات داخلی و خارجی مشاهده نمی شود. بر همین اساس پژوهش حاضر با همین رویکرد و با هدف ساخت ابزاری برای سنجش فرهنگ ارزشیابی و الگوی حاکم بر آن در دانشگاه ها و سیستم آموزش عالی شکل گرفت.

پیشینه پژوهش

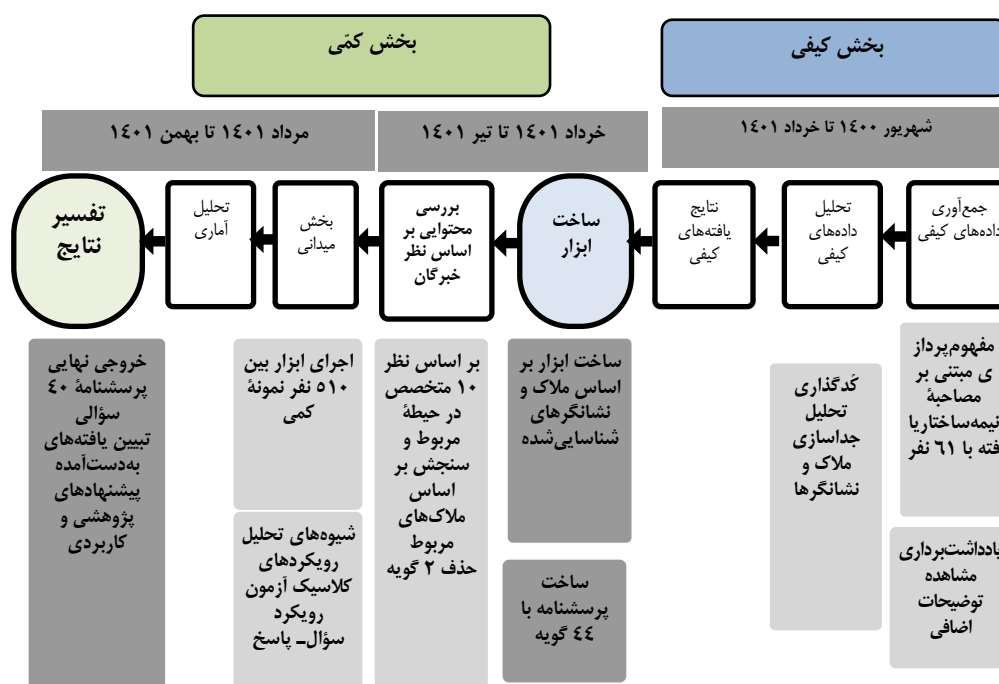
در زمینه فرهنگ ارزشیابی مطالعاتی انجام شده است و از آن جمله می توان به مطالعه مالمیر و همکارانش (۱۴۰۲ و ۱۴۰۲) اشاره کرد که در یک پژوهش فرهنگ ارزشیابی را زیرمجموعه ای از فرهنگ کیفیت برشمردند و در پژوهش دیگر فرهنگ ارزشیابی را مقوله ای چندبعدی و بر اساس سطوح مختلف در افراد ذی نفع و ذی ربط در دانشگاه ارزیابی کردند. سیمپر^۱ و همکارانش (۲۰۲۱) فرهنگ ارزشیابی در آموزش عالی را عامل تغییر و اهرم سیاست گذاری برشمردند. اما اشاره ای مشخص به ملاک ها و نشانگرهای مرتبط با آن نکردند. بامبر^۲ (۲۰۲۰) فرهنگ در دانشگاه را مبتنی بر فرهنگ کلی و فرهنگ دانشگاه ارزیابی کرد. ابوریزاه^۳ (۲۰۲۲) ساختار سازی چندبعدی را عامل مهم در شکل دهی یک کارآمد سازی اجرایی در به کارگیری فرهنگ ارزشیابی ارزیابی کرد. سیمپر^۴ و همکارانش (۲۰۲۱) تغییرات مبتنی بر سیاست گذاری اجرایی و سازمانی را مهم ترین حیطه در توسعه فرهنگ ارزشیابی تحلیل کرد. خادم محترم و سلیمی (۱۳۹۸) ارزشیابی و تضمین کیفیت در نظام آموزش عالی را منطبق بر انعطاف سیستمی و به کارگیری الگوی تغییرات در این زمینه با اهمیت شمرد. جوزه ماتوس گومز^۵ (۲۰۲۴) بیان کرد فرهنگ ارزشیابی مسئله ای است که همه

1. Simper
2. Bamber
3. Aburizaizah
4. Simper
5. José Matus Gómez

افراد ذی‌نفع و مرتبط در امر آموزش عالی باید در آن مشارکت فعال داشته باشند. ثانیاً، عناصر فرهنگ ارزشیابی از منظر دانشگاهی غیرتجزیه‌ای است و باید به صورت یک کلیت بررسی شود. در دیگر پژوهش‌ها اشاره شد برای توسعه بهتر سیستم آموزش عالی یک کشور یا منطقه مدیران باید وضعیت فعلی سیستم را کنترل کنند، که نیاز به ارزیابی منظم و واقع‌بینانه از کیفیت و پایداری آموزش عالی دارد و فرهنگ ارزشیابی ضرورت کارآمدسازی در این بخش است (Cao et al., 2023: 1). در یک جمع‌بندی از پیشینه‌های اشاره‌شده به صورت کلی مشخص است فرهنگ ارزشیابی کارآمد با بهبود توسعه کیفی در فرایند آموزشی و اجرایی همراه است. با این حال بررسی تحلیلی و عمیق با هدف مفهوم‌سازی و بالاخص ابزارسازی در این زمینه با شکاف‌های جدی در این بخش همراه است و محدود بررسی‌های انجام‌شده نیز بدون بررسی عمیق و تحلیلی و ارائه مدلی دقیق در این زمینه انجام شده است که ضرورت پژوهش در این حیطه را بیش از پیش بااهمیت می‌سازد.

روش پژوهش

پژوهش حاضر با استفاده از طرح آمیخته اکتشافی مدل ساخت ابزار انجام شد. در طرح‌های آمیخته فرصت استفاده بهینه از هر دو دسته از روش‌های پژوهش کمی و کیفی برای پاسخگویی به سؤال‌های پژوهشی و کمینه کردن کاستی‌های رویکردهای تک‌روشی و بی‌پیشینه‌سازی دقت فراهم می‌شود (Salehi & Golafshani, 2010).



شکل ۱. فرایند اجرایی پژوهش حاضر

در این مطالعه و متناسب با دو بخش آن، نخست ملاک‌ها و نشانگرها بر اساس رویکرد کیفی مبتنی بر روش پدیدارشناسی تفسیری مورد شناسایی قرار گرفت و در بخش دوم، با روش همبستگی، به ساخت و بررسی ویژگی‌های روان‌سنجی ابزار مربوط پرداخته شد. در بخش کیفی برای جمع‌آوری داده‌ها از مصاحبه عمیق و نیمه‌ساختاریافته استفاده شد. پروتکل مصاحبه مبتنی بر اصول پیشنهادی بنی‌اسدی و صالحی (۱۳۹۸) تدوین شد. محور سؤال‌های مصاحبه واکاوی در ادراک کُنشگران به مسئله فرهنگ ارزشیابی در نظام دانشگاه بود. در این زمینه ابتدا یک سؤال کلی مبنی بر بررسی تجارب و ساختار مبتنی بر آن مطرح شد و سپس به منظور دستیابی به داده‌های غنی‌تر سؤال‌های پیگیرانه پرسیده شد. تحلیل داده‌ها با استفاده از روش هفت مرحله‌ای دیکلمن انجام شد. در بخش کمی مبتنی بر روش همبستگی و با استفاده از روش‌های کلاسیک (CTT) و نوین اندازه‌گیری (IRT) ویژگی‌های روان‌سنجی ابزار مورد بررسی قرار گرفت.

میدان پژوهش و جامعه آماری

در بخش کیفی، میدان پژوهش دانشگاه تهران بود و شرکت کنندگان دربرگیرنده چهار دسته از ذی نفعان شامل اعضای هیئت علمی، کارکنان، مدیران، و دانشجویان دانشکده‌های مختلف دانشگاه تهران در سال ۱۴۰۰ - ۱۴۰۱ بودند و نمونه‌گیری به روش غیرتصادفی و هدفمند تا اشباع نمونه (۶۱ نفر) ادامه یافت. به این معنا که اطلاعات قبلی از شرکت کنندگان تکرار شد و اطلاعات جدیدی به دست نیامد. برای ارتقای دقت انتخاب شرکت کنندگان از چند ملاک ورود استفاده شد که در زیرگروه اعضای هیئت علمی دارا بودن سابقه چهارساله عضویت در هیئت علمی گروه‌های آموزشی، در دانشجویان ضرورت حضور در دوره‌های تحصیلات تکمیلی، در کارکنان و مدیران سابقه فعالیت بیش از سه سال در دانشگاه بود. همچنین داشتن آگاهی به مفهوم و رضایت آگاهانه (در همه گروه‌ها) از دیگر ملاک‌های مشارکت در تحقیق بود. پس از انجام ۶۱ مصاحبه با شرکت کنندگان اشباع داده‌ها به دست آمد. مصاحبه‌ها با رضایت مشارکت کنندگان ضبط و مکتوب شدند. ملاحظات اخلاقی در تمام مراحل پژوهش به طور کامل اجرا شد. به مصاحبه کنندگان در مورد حفظ اصول رازداری توضیح و اطمینان داده شد و برای پنهان ماندن هویت شرکت کنندگان هیچ اسمی ذکر نشد و از شماره و عبارت‌هایی مانند مصاحبه‌شونده‌ای گفت برای نقل قول‌هایشان استفاده شد. جامعه آماری در بخش کمی تحقیق شامل اعضای هیئت علمی، کادر آموزش و اجرایی، و دانشجویان دانشگاه تهران بود و به منظور انتخاب نمونه از روش نمونه‌گیری در دسترس استفاده شد. حجم نمونه بر اساس مقادیر به دست آمده در اجرای مقدماتی تعیین شد و بر اساس ضرورت استفاده از نمونه بالای ۲۵۰ نفر در بررسی و ساخت ابزار ۵۱۰ پرسشنامه بدون مشکل جمع‌آوری و تحلیل آماری شد.

یافته‌ها

عوامل اثرگذار بر فرهنگ ارزشیابی

در این بخش بر اساس مصاحبه کیفی نیمه‌ساختاریافته ملاک‌ها و نشانگرهای شناسایی شده مرتبط با مفهوم فرهنگ ارزشیابی ارائه شده است.

جدول ۱. ملاک و نشانگرهای فرهنگ ارزشیابی

| ردیف | نمونه مصاحبه | ملاک | نشانگر | تکرارپذیری |
|------|---|-------------------|---|-----------------------------|
| ۱ | یعنی انتقادپذیری، یعنی عیبی که به حق باشد را بپذیریم و پذیرش اینکه انسان کاملی نیستیم و باید از ما انتقاد شود. | پذیرش نقد شدن | فرهنگ‌سازی پذیرش نقص | ۲۴ ۲۰ |
| ۲ | آیین‌نامه‌ها تغییر کنند و از نگاه اقتصادی به دانشگاه‌ها جلوگیری شود. الان دانشگاه‌ها را بنگاه اقتصادی می‌بینند و این بر کیفیت اثر می‌گذارد. | سیاست‌گذاری | ویژگی‌های ارزیابی منطق ارزیابی تدوین الگوی ارزیابی ملاحظات ارزیابی سیاست‌های تشویقی سیاست‌های نظارتی، کنترلی | ۸ ۶ ۱۰ ۷ ۸ ۹ |
| ۳ | اساس ایجاد فرهنگ کیفیت و ارزشیابی اینه؛ هوای هم را داشتن در راستای بهبود شرایط کنار گذاشته بشه. یعنی سعی کنیم شرایط رو به نفع تحقق اهداف بنا بذاریم. | کاهش مصلحت‌اندیشی | جدیت عملیاتی جدیت فکری عدالت‌محوری | ۱۰ ۹ ۱۲ |
| ۴ | پس فرهنگ ارزیابی در فرهنگ کیفیت مؤثره. چون همین ارزیابی باعث می‌شود بفهمیم چقدر یا این اهداف و چشم‌انداز فاصله داریم، به کدام یک از بخش‌ها نسبتاً نزدیک‌تر شدیم و به کدام یکی تقریباً رسیدیم. اگر ارزیابی نکنیم، اصلاً نمی‌تونیم بفهمیم که از لحاظ کیفیت در کجای کاریم. | نقش ارزیابی | اثرگذاری تیم ارزیابی حرفه‌ای شناخت همه گروه‌های دانشگاهی از وجود سیستم کمی‌سازی گزارش‌ها و میزان اثربخشی | ۱۹ ۱۰ ۸ ۶ ۴ |
| ۵ | مثلاً بودجه دانشگاه و برنامه‌ریزی‌هایی که خود وزارت علوم انجام می‌دهد بسیار در شکل‌گیری فرهنگ ارزشیابی مهمه. | بودجه‌بندی | تبلیغات معرفی حمایت مالی | ۱۲ ۸ ۱۱ |

ادامه جدول ۱.

| ردیف | نمونه مصاحبه | ملاک | نشانگر | تکرارپذیری |
|------|--|--------------------------------------|--|-----------------------------|
| ۶ | مؤلفه اول توی خود دانشگاه تمام درون‌داد این سیستمه؛ معلمش، دانشجوی، امکاناتش، همین‌طور نحوه تعامل بین معلم‌ها و استاد و دانشجویها و دانشجویها با هم. | تعامل درون‌داد و برون‌داد | روابط سیستمی نوع الگوی ارتباطی درون‌دانشگاهی روابط میکرو و ماکروی سیستم | ۱۰ ۷ ۵ |
| ۷ | اگر بخواهیم فرهنگ ارزشیابی در دانشگاه اجرا شود نباید به دنبال منافع شخصی باشیم. باید به دنبال منافع جمعی باشیم. اگر این نگاه رو نداشته باشیم هیچ‌وقت فرهنگ ارزشیابی نمی‌تواند جا بیفتد. | فرهنگ عمومی | ارزش‌های عمومی باورهای عمومی نگاه اجتماع به مسئله ارزشیابی تحمل‌پذیری جامعه ارزش جامعه برای مقوله ارزشیابی | ۲۱ ۲۶ ۲۰ ۷ ۱۰ |
| ۸ | اگر ما روحیه انتقادپذیری نداشته باشیم این اتفاق در دانشگاه نمی‌افتد. یعنی اول باید این نوع نگاه از دوره پیش‌دبستان و دبستان تربیت شده باشه. | روحیه نقادی | گشودگی به تجربه انعطاف‌پذیری چالش‌پذیری | ۱۱ ۱۲ ۷ |
| ۹ | اون قدر سعه صدر داشته باشیم که بشنویم و پرخاش نکنیم و تامل کنیم. | وجود سعه صدر | تحمل‌پذیری اصلاحگری تامل‌گری و به کار بستن | ۲۹ ۱۱ ۸ |
| ۱۰ | بعد اخلاقی‌ش رو رصد کنه تا بعد علمی‌ش را وقتی لایه‌هایی صورت می‌گیره که باعث می‌شود افراد سطح‌پایین وارد دانشگاه شوند. وقتی گزینش دانشجو درست نیست همه این‌ها دخیله. | نظارت | نظارت بر مفاد نظارت بر اجرا نظارت بر تیم اجرا کنترل ارزشیابی موقعیتی ارزشیابی برای همه گروه‌ها | ۸ ۱۰ ۶ ۴ ۵ |
| ۱۱ | مهم‌ترین راهکار برای گسترش این فرهنگ اینه که تک به تک ما پذیرا باشیم و قبول کنیم اگر می‌خواهیم این فرهنگ در جامعه و دانشگاه اجرا شود. | آموزش فرهنگی | آموزش پایه به ارزشیابی معرفی ارزشیابی و ضرورت آن اصل‌سازی ضرورت ارزشیابی در هر سطح | ۹ ۷ ۱۰ |
| ۱۲ | یکی از مهم‌ترین ساخت‌های شکل‌دهنده برای فرهنگ ارزشیابی اینه که شفاف بفهمیم چی می‌خواد این فرهنگ و بعد گسترش بدیم. بنابراین شفاف‌سازی توی این مسئله خیلی مهمه. | شفاف‌سازی مفهوم | تعریف ارزشیابی تعریف ابعاد ارزشیابی مشخص ساختن نحوه ارزشیابی مشخص ساختن تاثیر ارزشیابی تبیین اهداف ارزشیابی | ۱۳ ۹ ۱۴ ۱۴ ۱۰ |
| ۱۳ | مگر از طریق جلسات هم‌فکری و اصلاح آیین‌نامه‌های ارتباطی یا حتی نحوه جذب دانشجو باید نظرسنجی بشوند. | هم‌فکری | تیم‌سازی برای تبیین‌گذاری در جریان قرار گرفتن تمامی افراد مشارکت همگانی سیستم | ۶ ۶ ۸ |
| ۱۴ | هم اینک جامعه نمی‌طلبه هم اینک استادان به دلیل بار اقتصادی که بر دوشون هست ذهنیت باز را ندارند که بتونن برن دنبال ارتقای کیفیت. | بهبود شرایط افراد ذی‌نفع و ذی‌ربط | ارزشیابی در جهت اصلاح ارزشیابی در جهت بهبود ارزشیابی در جهت شناخت ارزشیابی در جهت پاسخگویی ارزشیابی در جهت شفاف‌سازی | ۸ ۷ ۱۰ ۷ ۵ |
| ۱۵ | در نظر گرفتن هزینه مالی برای متولیان این امر در ایجاد انگیزه برای اجرای کار به بهترین شکل می‌تواند نقش اساسی داشته باشد. | انگیزه‌افزایی | انگیزه مشارکت انگیزه تعامل انگیزه همکاری انگیزه معنوی | ۱۱ ۱۴ ۱۳ ۷ |
| ۱۶ | باید به صورت پیوستاری و منطبق بر قانون تدوین شود. | قانون‌سازی ارزشیابی | اجباری‌سازی اجرائی‌سازی ابلاغ به همه بخش‌ها | ۷ ۱۷ ۶ |
| ۱۷ | بله، تا وقتی که داریم ارزشیابی می‌کنیم یعنی داریم بررسی می‌کنیم که آیا اون ملاک‌ها درست‌اند در حد خودشان یا خیر و آیا چیزهایی که بر اساس اهداف مشخص کردیم حاصل شده یا خیر. پس ارزیابی به ما کمک می‌کنه که بفهمیم به چه میزان به اهدافمون رسیدیم. | ارزشیابی کارآمد | ارزشیابی نتیجه‌گرا ارزشیابی قانونمند ارزشیابی عادلانه مبتنی بر اسناد بالادستی شفافیت ملاک‌های ارزشیابی ارزشیابی فرایندی | ۸ ۱۰ ۹ ۶ ۵ ۶ |
| ۱۸ | مهم‌ترین عامل برای اجرایی و عملی شدن یک کار رعایت تمام و کمال پارامترهای آن توسط کسانی است که به عنوان متولیان این کار شناخته می‌شوند. | مسئولیت‌پذیری فردی | تعهد فردی تعهد سازمانی تعهد کاری | ۸ ۱۲ ۱۹ |
| ۱۹ | کار گروهی که هر ساله با اهتمام تمام وظیفه تحلیل نتایج این کار را بر عهده داشته باشد. در واقع با وجود متخصصین در این کار گروه هر ساله دستاوردها، بهبودها، یا بدتر شدن شرایط به لحاظ کیفی و کمی ارزیابی شود. | کارگروه تخصصی | تیم متخصص تیم حرفه‌ای تیم آموزش‌دیده تیم مورد اطمینان | ۱۸ ۱۶ ۷ ۱۰ |

بررسی در زمینه ملاک‌ها و نشانگرهای فرهنگ ارزشیابی منتج به ۱۹ ملاک و ۷۹ نشانگر شد. پذیرش نقد شدن، سیاست‌گذاری، کاهش مصلحت‌اندیشی، نقش ارزیابی، بودجه‌بندی، تعامل درون‌داد و برون‌داد، فرهنگ عمومی، روحیه نقادی، وجود سعه صدر، نظارت، آموزش فرهنگی، شفاف‌سازی مفهوم، هم‌فکری، بهبود شرایط افراد ذی‌نفع و ذی‌ربط، انگیزه‌افزایی، قانون‌سازی ارزشیابی، ارزشیابی کارآمد، ذیری فردی، کارگروه تخصصی مهم‌ترین ملاک‌های مرتبط با اشتراک همه افراد در فرهنگ ارزشیابی است.

ساخت ابزار سنجش فرهنگ ارزشیابی

بر اساس الگوی برآمده از اشتراک ادراک یاران آموزشی در مصاحبه‌های کیفی، پرسشنامه فرهنگ ارزشیابی ۴۴ گویه ساخته شد. این پرسشنامه دارای طیف ۶ درجه‌ای بود که ملاک وضعیت موجود بود و بر این اساس به هر گویه از ۱ تا ۶ امتیاز داده شد؛ به گونه‌ای که امتیاز ۶ اشاره به وجود آن ویژگی در حداکثر وضعیت موجود دانشکده و نمره ۱ اشاره به حداقل داشت. گفتنی است به سبب جلوگیری از پاسخ‌های تکراری در امتیاز به صورت عددی توسط پاسخ‌دهندگان بود. نمونه مشارکت‌کننده در بخش کمی را ۴۷ درصد مردان و ۵۳ درصد زنان تشکیل دادند. ۸۷ درصد نمونه تحقیق در بخش کمی دانشجوی، ۲/۴ درصد مدیر، ۵/۹ درصد هیئت علمی، و ۴/۷ درصد کارمند بودند. همچنین، از نظر سن، جامعه آماری در بخش کمی ۹۰/۶ درصد زیر ۴۰ سال بودند.

برای ساخت پرسشنامه تحقیق برگرفته از همه نشانگرهای اصلی پرسشنامه‌ای ۶۰ گویه‌ای آماده شد که پس از بررسی اولیه توسط تیم پژوهش و نیز مطالعه مقدماتی ۲۰ گویه حذف و ۴ گویه جدید برگرفته از مفاهیم ضروری حذف شده با پیشینه پژوهش و ساختار متفاوت وارد پرسشنامه شد و پرسشنامه با ۴۴ گویه در اختیار خبرگان برای بررسی روایی محتوایی و ویژگی‌های روان‌سنجی قرار گرفت. برای بررسی روایی^۱ پرسشنامه محقق ساخته از سه نوع از روایی استفاده شد.

روایی صوری: این روایی ملاکی ذهنی و مطالعه‌ای دارد و برای تأمین روایی در این بخش ۳ نفر از افراد همکار و متخصص، با مطالعه، پرسشنامه را مورد بررسی دقیق قرار دادند و پس از چند مرحله ویرایش مورد تأیید قرار گرفت.

روایی محتوایی: برای تعیین روایی محتوایی از روش CVI و CVR استفاده شد. بر اساس بررسی ۴۴ گویه پذیرفته شده در دور اول بررسی ۴۲ گویه بر اساس ملاک‌های روایی محتوایی و ملاکی تأیید نهایی شد و در گروه نهایی مشارکت‌کنندگان در تحقیق استفاده شد. در بررسی روایی سازه و تحلیل عاملی میزان KMO برای پرسشنامه فرهنگ ارزشیابی $[KMO=0/872]$ بود که در سطح عالی قرار داشت و در ادامه از تحلیل عاملی اکتشافی استفاده شد.

تحلیل عاملی اکتشافی به روش تحلیل محورها اصلی انجام شد و از چرخش واریماکس به منظور رسیدن به عامل‌های به‌هم‌پیوسته استفاده شد. در این بخش پس از ۶ چرخش و حذف ۲ گویه و اصلاح عامل‌های ناهمگون همه عوامل در بعد خویش قرار گرفتند. بر اساس نتایج ذکر شده مقادیر مرتبط با میزان اشتراکات استخراجی کلیه سؤالات با هدف تحلیل عاملی بالاتر از ۰/۴ قرار داشتند که نشان‌دهنده همبستگی قابل قبول بین تک‌تک مؤلفه‌ها با کل آزمون و مناسب بودن برای تحلیل عاملی بود. برای ارزیابی برازش مدل از شاخص‌های ریشه دوم برآورد واریانس خطای تقریب (RMSEA)، ریشه دوم میانگین مجزورات پس‌مانده‌های استاندارد شده (RMSR)، شاخص تاکر-لویز (TLI) استفاده شد. مقدار شاخص‌های RMSR، RMSEA، و TLI در فرهنگ ارزشیابی ۰/۰۵۸، ۰/۰۶۹، و ۰/۹۰ بود. فاصله اطمینان ۹۰ درصد برای شاخص RMSEA ۰/۰۶۷ بود. مقدار تمام این شاخص‌ها با ملاک‌ها و قواعد سرانگشتی که در تفسیر آن‌ها به کار می‌رود (کلاين، ۲۰۰۵) مطابقت داشت. بارهای عاملی در جدول ۲ ارائه شده است.

جدول ۲. میزان اشتراک مواد ابزار پژوهش در فرهنگ ارزشیابی

| میزان اشتراک | گویه | میزان اشتراک | گویه | میزان اشتراک | گویه | میزان اشتراک | گویه | میزان اشتراک | گویه |
|--------------|------|--------------|------|--------------|------|--------------|------|--------------|------|
| ۰/۵۸۲ | ۵ | ۰/۵۹۸ | ۴ | ۰/۷۱۱ | ۳ | ۰/۶۰۳ | ۲ | ۰/۴۸۱ | ۱ |
| ۰/۴۹۳ | ۱۰ | ۰/۵۷۱ | ۹ | ۰/۴۹۸ | ۸ | ۰/۶۴۱ | ۷ | ۰/۵۸۸ | ۶ |
| ۰/۴۷۱ | ۱۵ | ۰/۵۹۳ | ۱۴ | ۰/۵۴۷ | ۱۳ | ۰/۵۹۸ | ۱۲ | ۰/۵۵۵ | ۱۱ |
| ۰/۵۰۳ | ۲۰ | ۰/۷۱۱ | ۱۹ | ۰/۵۸۲ | ۱۸ | ۰/۵۸۸ | ۱۷ | ۰/۵۶۴ | ۱۶ |
| ۰/۶۴۴ | ۲۵ | ۰/۵۴۷ | ۲۴ | ۰/۴۸۱ | ۲۳ | ۰/۶۷۴ | ۲۲ | ۰/۴۲۳ | ۲۱ |
| ۰/۴۷۹ | ۳۰ | ۰/۶۲۸ | ۲۹ | ۰/۶۲۲ | ۲۸ | ۰/۵۶۴ | ۲۷ | ۰/۴۴۶ | ۲۶ |
| ۰/۵۵۹ | ۳۵ | ۰/۵۸۱ | ۳۴ | ۰/۵۵۱ | ۳۳ | ۰/۵۷۷ | ۳۲ | ۰/۷۲۷ | ۳۱ |
| ۰/۵۸۳ | ۴۰ | ۰/۶۰۵ | ۳۹ | ۰/۵۷۴ | ۳۸ | ۰/۵۹۵ | ۳۷ | ۰/۶۰۲ | ۳۶ |
| | | | | | | ۰/۶۱۹ | ۴۲ | ۰/۷۰۱ | ۴۱ |

در جدول ۲ مشاهده می‌شود که تمامی مقادیر مرتبط با همبستگی گویه‌ها با کل بالاتر از ۰/۴ قرار دارند که نشان‌دهنده همبستگی بالا بین تک‌تک گویه‌ها با کل آزمون و مناسب بودن آن جهت تحلیل عاملی است. در چرخش‌های نهایی که شش بار انجام شد دو گویه از فرهنگ ارزشیابی حذف و بر این اساس پرسشنامه نهایی بر اساس الگوی کلاسیک دارای ۴۰ سؤال در بخش فرهنگ ارزشیابی شد که نتایج بر اساس دسته‌های انجام گرفته و تقسیم شده در ادامه گزارش می‌شود.

تعداد عوامل

یکی از روش‌های با پایه علمی قوی در تعیین تعداد عوامل هر پرسشنامه تحلیل موازی هرن (۱۹۶۵) است. «بر اساس ایده هرن، برآورد کران پایین ریشه پنهان یک گاتمن برای تعیین رتبه یک ماتریس همبستگی به عنوان کران بالای روان‌سنجی در تعیین تعداد عامل‌ها پذیرفته شده است. برای تعیین رتبه یک ماتریس نمونه باید ریشه‌های پنهانی که به خطای نمونه‌گیری قابل اختصاص هستند را از تعداد مؤلفه‌ها کم کرد.» (ایزائلو، ۱۴۰۱: ۸۶).

جدول ۳. شاخص‌های آماری اولیه در بخش ملاک‌های فرهنگ ارزشیابی

| عامل | ارزش ویژه | تحلیل موازی HORN ارزش ویژه‌های تصادفی | درصد تبیین‌کنندگی واریانس | درصد تجمعی |
|------|-----------|---------------------------------------|---------------------------|------------|
| ۱ | ۷/۹۳ | ۱/۶۴۹ | ۲۰/۸۳ | ۲۰/۸۳ |
| ۲ | ۳/۷۶ | ۱/۵۲۹ | ۹/۱۱ | ۲۹/۹۴ |
| ۳ | ۳/۰۲ | ۱/۵۲۳ | ۷/۹۳ | ۳۷/۸۷ |
| ۴ | ۲/۴۸ | ۱/۲۹۴ | ۵/۲۲ | ۴۳/۰۹ |
| ۵ | ۱/۸۹ | ۱/۲۱۵ | ۳/۴۸ | ۴۶/۵۷ |
| ۶ | ۱/۶۳ | ۱/۲۰۴ | ۳/۱۱ | ۴۹/۶۸ |
| ۷ | ۱/۲۴۴ | ۱/۱۰۲ | ۲/۶۹ | ۵۲/۳۷ |
| ۸ | ۰/۹۸۹ | ۱/۰۵۳ | | ۲ |

در جدول ۳ فرهنگ ارزشیابی یافته‌ها بر اساس روش موازی هورن حاکی از آن است که تعداد عوامل بهینه برای استخراج هفت عامل است و پس از عامل هفتم یعنی در عامل هشتم ارزش ویژه تصادفی (۱/۰۵۳) از ارزش ویژه تحلیلی عاملی (۰/۹۸۹) بیشتر است. در تحلیل حاضر هفت جزء با ارزش ویژه بزرگ‌تر از یک شناسایی شد که از ارزش‌های ویژه تصادفی بیشتر بود و ۵۲/۳۷ درصد از واریانس کل را برآورد می‌کند.

پس از مشخص شدن تعداد عوامل بر اساس تحلیل عاملی اکتشافی^۱ پس از چهار دور چرخش ۴۰ گویه در هفت سطح قرار گرفت که با بررسی تحلیلی و محتوایی فرهنگ ارزشیابی هر سطح نام گذاری شد.

تحلیل عاملی تأییدی

پس از انجام دادن تحلیل عاملی اکتشافی و حذف سؤال‌های نامناسب و مشخص شدن عامل‌ها، مدلی به دست آمد که به عنوان مدل مفروض برای انجام دادن تحلیل عاملی تأییدی مورد استفاده قرار گرفت. جدول ۴ شاخص‌های برازش را نشان می‌دهد.

جدول ۴. آماره‌های برازش در تحلیل عاملی تأییدی در مدل ۷ عاملی

| مقادیر | شاخص‌های آماری | χ^2/DF | AGFI | GFI | TLI | CFI | RMSEA |
|----------------|----------------|-------------|------|------|------|------|-------|
| فرهنگ ارزشیابی | مقدار برازش | ۲,۸۲۳ | ۰/۹۱ | ۰/۸۷ | ۰/۹۰ | ۰/۹۱ | ۰/۰۸۰ |

متداول‌ترین روش برای برآورد بهترین برازندگی روش بیشینه احتمال نامیده می‌شود. از مجموعه آماره‌های برازش مربوط به مدل ۵ عامل χ^2 ، AGFI، GFI، CFI، RMSEA در این مدل اندازه‌گیری شد. دو شاخص AGFI و GFI هر چه به یک نزدیک باشند برازش کامل مدل را نشان می‌دهند و کم بودن شاخص RMSEA به منزله مطلوبیت برازش مدل است. شاخص هر چه به یک نزدیک‌تر باشد مطلوبیت (CFI) برازندگی تطبیقی بیشتر مدل را نشان می‌دهد.

همچنین در تفسیر شاخص‌های برازندگی قواعد پیشنهادی زیر به کار برده شد. براون و کادک (۱۹۹۳) معتقدند که مقادیر بالاتر از ۰/۹۰ برای شاخص‌های TLI حاکی از برازش خوب مدل است. در مورد شاخص RMSEA نیز از ملاک پیشنهادی هیو و بنتلر (۱۹۹۹) استفاده شد. آن‌ها معتقدند اگر مقدار این شاخص کمتر از ۰/۰۸ باشد، می‌توان نتیجه گرفت که برازش مدل با داده‌ها خوب است (کلاين، ۲۰۰۵).

مقدار شاخص TLI برای فرهنگ ارزشیابی برابر با ۰/۹۰۴ است که نشان از مناسب بودن میزان به‌دست‌آمده در این بخش دارد. همچنین با توجه به اینکه $AGFI=0/91$ ، $GFI=0/87$ ، $CFI=0/91$ است و در این شاخص‌های برازش هر چه میزان به‌دست‌آمده نزدیک‌تر به یک باشد نشان‌دهنده برازش مناسب‌تر است و با توجه به اینکه میزان RMSEA برابر با ۰/۰۸۰ است که هر چه به صفر نزدیک‌تر باشد نشان‌دهنده برازش مناسب‌تر است این مدل دارای برازش بسیار مناسبی است و تمام ساختارها معنادار و منطبق بر مدل است. جهت بررسی اعتبار آزمون از ضریب مک دونالد استفاده شد.

با توجه به میزان بالای ۰/۷ در همه بخش‌های فرهنگ ارزشیابی، ابزار مورد نظر از اعتبار مناسب برخوردار است. همچنین برای بررسی ثبات ابزار تحقیق یک هفته پس از اجرای اصلی روی بخشی از نمونه در دسترس مجدد ارائه شد و یافته‌ها دلالت بر ثبات مناسب نتایج پرسشنامه داشت.

بررسی گویه با نظریه سؤال-پاسخ (IRT)

تحلیل سؤالات پرسشنامه اندازه‌گیری فرهنگ ارزشیابی برازش مدل‌های برآورد پارامتر گویه‌های پرسشنامه اندازه‌گیری فرهنگ کیفیت توسط نرم‌افزار R نسخه ۴,۲,۲ با استفاده از پکیج‌های ltm و eRm انجام شد و پس از مشخص شدن بهترین مدل منحنی ویژگی سؤالات (ICC) منحنی آگاهی سؤالات (IIC) شد. در این مرحله با استفاده از نرم‌افزار R و پکیج eRm پارامترهای شخص (تتا یا توانایی) و پارامترهای سؤال بر اساس مدل PCM برآورده شده. سپس پارامترهای آستانه‌ای بین طبقات، آماره کای دو، درجه آزادی، و سطح معناداری هر سؤال گزارش شده است.

۱. در فایل پیوست جدول مربوط ارائه شده است.

- adjusted goodness of fit index
- goodness of fit index
- comparative fit index
- root mean square error of approximation

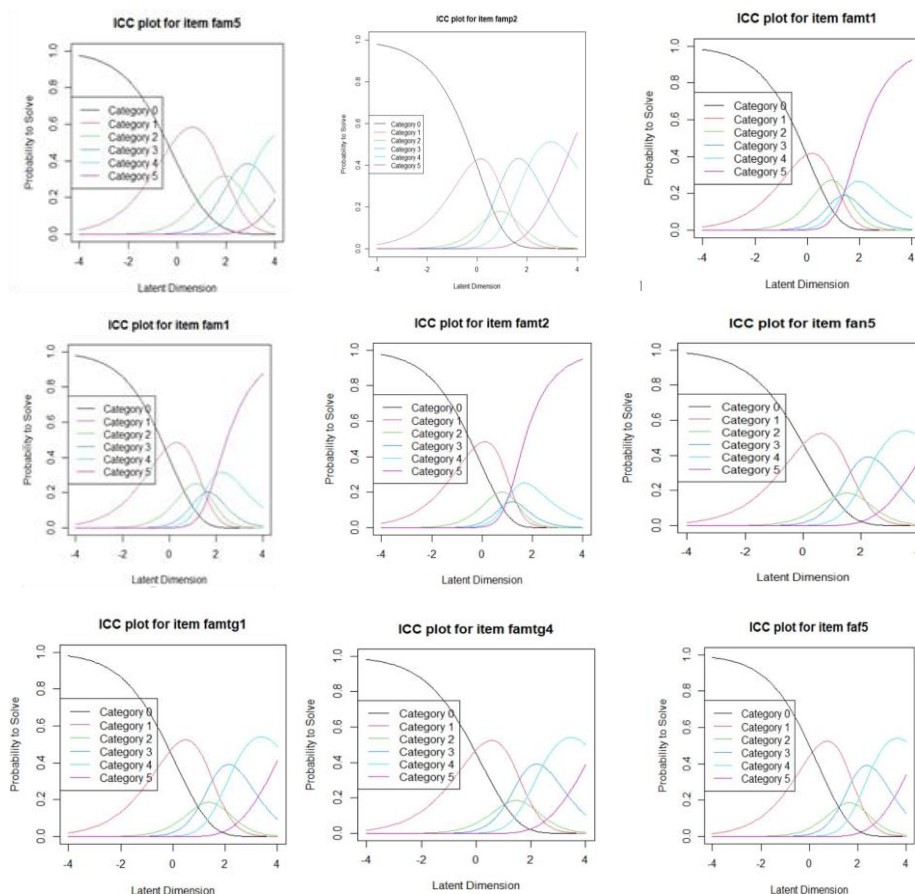
جدول ۵. بررسی اعتبار پژوهش

| خرده‌مقیاس | تعریف عملیاتی | تعداد گویه | آلفای کرونباخ | مک دونالد | همبستگی |
|-------------------------------------|---|------------|---------------|-----------|---------|
| فرهنگ ارزشیابی معیار | بر اساس نشانگرهای شناسایی شده، به به‌کارگیری ارزشیابی با توجه به آمار و گزارش‌ها با هدف بازخواست و سیستم تشویق و تنبیه نسبت به افراد و بخش‌ها اشاره دارد. | ۶ | ۰/۹۰۳ | ۰/۹۱۵ | ۰/۹۴ |
| فرهنگ ارزشیابی مسئولیت‌پذیری | بر اساس نشانگرهای شناسایی شده، به فرهنگ مبتنی بر به‌کارگیری ارزشیابی در جهت افزایش مسئولیت‌پذیری و پاسخگویی افراد و بخش‌های سازمان اشاره دارد. | ۶ | ۰/۸۹۸ | ۰/۹۰۶ | ۰/۷۸ |
| فرهنگ ارزشیابی مستند | بر اساس نشانگرهای شناسایی شده، به به‌کارگیری ارزشیابی در جهت مستندسازی و گزارش‌های تحلیلی در بخش‌های مربوط اشاره دارد. | ۶ | ۰/۸۸۹ | ۰/۸۹۴ | ۰/۸۵ |
| فرهنگ ارزشیابی نتیجه‌گرا | بر اساس نشانگرهای شناسایی شده، به به‌کارگیری ارزشیابی با ساختار عملیاتی و در جهت دستیابی به نتیجه کلی سازمانی اشاره دارد. | ۶ | ۰/۸۹۵ | ۰/۹۰۱ | ۰/۸۳ |
| فرهنگ ارزشیابی مبتنی بر توسعه افراد | بر اساس نشانگرهای شناسایی شده، به به‌کارگیری ارزشیابی در جهت توسعه فردی و دانشی و مهارتی اعضای درون سازمان در ابعاد مختلف اشاره دارد. | ۶ | ۰/۸۸۸ | ۰/۸۸۹ | ۰/۸۶ |
| فرهنگ ارزشیابی مبتنی بر تظاهرگرایی | بر اساس نشانگرهای شناسایی شده، به به‌کارگیری ارزشیابی در جهت ایجاد گزارش‌هایی با هدف خوب جلوه دادن بدون پشتوانه عملیاتی و واقعی بودن اشاره دارد. | ۵ | ۰/۸۶۰ | ۰/۸۶۳ | ۰/۹۴ |
| فرهنگ ارزشیابی فردگرا | بر اساس نشانگرهای شناسایی شده، به به‌کارگیری ارزشیابی با هدف پذیرش نقد شدن فردی و بالا رفتن ظرفیت‌های مورد بررسی قرار گرفتن اشاره دارد. | ۵ | ۰/۸۴۴ | ۰/۸۵۶ | ۰/۸۹ |
| فرهنگ ارزشیابی | به مجموعه گویه‌ها در همه فرهنگ‌های ارزشیابی موجود اشاره دارد. | ۴۰ | ۰/۸۵۵ | ۰/۸۶۹ | ۰/۸۵۴ |

جدول ۶. پارامترهای برآورد شده برای مدل امتیاز پاره‌ای به تفکیک خرده‌مقیاس‌ها برای پرسشنامه فرهنگ ارزشیابی

| خرده‌مقیاس | پرسش | δ1 | ۲δ | ۳δ | δ4 | δ5 | X ² | DF | P-VALUE |
|-------------------------------------|------|-------|-------|-------|-------|-------|----------------|-----|---------|
| فرهنگ ارزشیابی معیار | ۱ | - | -1.96 | 1.51 | -1.92 | -2.15 | 598.08 | 509 | 0.004 |
| | ۲ | -1.08 | -0.36 | -0.34 | 1.06 | 5.11 | 483.26 | 509 | 0.788 |
| | ۳ | -1.32 | -0.58 | -0.97 | 0.28 | 3.89 | 443.50 | 509 | 0.983 |
| | ۴ | -1.17 | 0.97 | 0.25 | 1.55 | 5.25 | 394.21 | 509 | 1.000 |
| | ۵ | 1.62 | -0.93 | 1.37 | -0.34 | 0.41 | 375.40 | 509 | 1.000 |
| | ۶ | -1.75 | -1.15 | -1.93 | -1.29 | 3.56 | 424.29 | 509 | 0.997 |
| فرهنگ ارزشیابی مسئولیت‌پذیری | ۷ | - | -0.83 | -0.67 | 1.03 | 5.19 | 376.90 | 509 | 1.000 |
| | ۸ | -2.03 | -1.69 | -2.35 | -1.36 | -2.89 | 531.26 | 509 | 0.239 |
| | ۹ | -1.92 | -1.52 | -2.24 | -2.88 | -3.43 | 455.95 | 509 | 0.956 |
| | ۱۰ | 1.85 | 1.91 | 1.98 | 0.62 | 4.05 | 409.19 | 509 | 1.000 |
| | ۱۱ | -1.32 | -0.18 | 0.11 | 1.82 | 5.44 | 373.66 | 509 | 1.000 |
| | ۱۲ | -1.70 | -1.01 | 0.19 | 0.21 | 0.31 | 561.24 | 509 | 0.054 |
| فرهنگ ارزشیابی مستند | ۱۳ | - | -0.74 | -0.41 | 1.05 | 4.85 | 475.23 | 509 | 0.856 |
| | ۱۴ | -1.21 | 0.27 | 0.86 | 1.03 | 4.69 | 500.76 | 509 | 0.594 |
| | ۱۵ | -3.40 | -2.95 | -2.02 | -1.69 | 2.03 | 470.42 | 509 | 0.889 |
| | ۱۶ | -1.30 | 0.07 | 0.59 | 0.60 | 2.81 | 443.50 | 509 | 0.983 |
| | ۱۷ | -1.77 | -0.12 | 0.63 | 0.69 | 3.02 | 436.01 | 509 | 0.991 |
| | ۱۸ | -1.45 | -0.77 | -0.57 | -0.44 | -0.15 | 475.23 | 509 | 0.856 |
| فرهنگ ارزشیابی نتیجه‌گرا | ۱۹ | - | -1.36 | -1.34 | -0.21 | 1.42 | 399.73 | 499 | 1.000 |
| | ۲۰ | -1.16 | 0.02 | 0.22 | 1.47 | 4.84 | 391.00 | 499 | 1.000 |
| | ۲۱ | -1.41 | -0.54 | -0.22 | 1.06 | 4.19 | 386.39 | 499 | 1.000 |
| | ۲۲ | -1.02 | 0.95 | 2.11 | 0.84 | 0.43 | 380.09 | 499 | 1.000 |
| | ۲۳ | -0.67 | -0.59 | -1.28 | 0.16 | 2.27 | 616.11 | 499 | 0.000 |
| | ۲۴ | -1.56 | -1.16 | -2.28 | 2.81 | 2.17 | 502.70 | 499 | 0.445 |
| فرهنگ ارزشیابی مبتنی بر توسعه افراد | ۲۵ | - | -1.83 | -2.34 | -1.30 | 3.27 | 414.12 | 499 | 0.998 |
| | ۲۶ | -1.23 | -1.22 | -0.13 | 1.04 | -1.18 | 569.25 | 499 | 0.016 |
| | ۲۷ | -1.06 | -0.19 | 0.00 | 1.26 | 4.60 | 476.43 | 499 | 0.759 |
| | ۲۸ | -1.51 | -0.62 | -0.47 | -0.04 | 2.79 | 425.36 | 499 | 0.993 |
| | ۲۹ | -1.13 | 0.28 | 0.67 | 1.95 | 5.58 | 387.22 | 499 | 1.000 |
| | ۳۰ | -1.58 | -0.65 | 0.7 | 1.11 | 2.15 | 390.95 | 499 | 1.000 |
| فرهنگ ارزشیابی مبتنی بر تظاهرگرایی | ۳۱ | - | -1.16 | -1.37 | -0.54 | 3.74 | 427.85 | 499 | 0.991 |
| | ۳۲ | -1.61 | -0.89 | -0.32 | 1.647 | 4.91 | 402.43 | 499 | 0.999 |
| | ۳۳ | 0.86 | 1.51 | 1.99 | 2.09 | 3.11 | 400.41 | 499 | 1.000 |
| | ۳۴ | 1.18 | -1.00 | -1.15 | 0.29 | -3.87 | 397.33 | 499 | 1.000 |
| | ۳۵ | 1.75 | 1.89 | 1.94 | 2.02 | 4.08 | 485.72 | 499 | 0.656 |
| | ۳۶ | - | -0.52 | -0.41 | 1.02 | 3.11 | 390.22 | 509 | 1.000 |
| فرهنگ ارزشیابی مبتنی بر فردگرایی | ۳۷ | -1.67 | -1.00 | 0.17 | 0.28 | 0.37 | 431.11 | 509 | 0.995 |
| | ۳۸ | -1.30 | 0.87 | 0.80 | 0.83 | 4.73 | 371.52 | 509 | 1.000 |
| | ۳۹ | -0.95 | 0.01 | 0.07 | 1.56 | 3.16 | 493.50 | 509 | 0.681 |
| | ۴۰ | -0.06 | -2.86 | 2.51 | -1.40 | 2.47 | 467.64 | 509 | 0.905 |

نتایج حاصل از جدول ۶ نشان می‌دهد نظر به ضریب کای دو سوالات ۲، ۳ و ۲۶ به طور معناداری با مدل PCM در خرده‌مقیاس خود برآزش نداشته‌اند. با توجه به اینکه از مدل امتیاز پاره‌ای برای تحلیل پاسخ‌های مقیاس نگرش که در آن آزمودنی‌ها در یک مقیاس چندنقطه‌ای به سوالات پاسخ می‌دهند استفاده می‌شود، برخلاف مدل پاسخ مدرج، پارامترهای تقاطع لزوماً مرتب نیستند. وقتی پارامترهای تقاطع مرتب نباشند، به این پدیده وارونگی گفته می‌شود. در پرسشنامه حاضر، سؤال‌های ۱، ۵، ۸، ۲۳، ۲۵، ۲۶، ۳۱، ۳۴ و ۴۰ دارای وارونگی هستند. اگر وارونگی وجود داشته باشد، با اطمینان می‌توان گفت دست‌کم یک طبقه وجود دارد که هرگز محتمل‌ترین گزینه مشروط به سطح صفت نیست (امبرتسون و رایس، ۲۰۰۰: ۱۳۶، ۱۴۸ و ۱۴۹). این حالت در خصوص پرسش‌هایی با ماهیت پرسشنامه فعلی کاملاً قابل توجیه است.



نمودار ۱. منحنی ویژگی پرسش‌های دارای وارونگی در مدل امتیاز پاره‌ای

نمودارهای به‌دست‌آمده نیز حاکی از آگاهی مطلوب خرده‌مقیاس‌ها در اندازه‌گیری صفت مکنون در سطوح آگاهی متوسط است.^۱

نتیجه

توسعه پایدار و پیشرفت درون‌زاد هر کشوری تابعی از تصمیمات و اقدامات هوشمندانه برای ارتقای سطوح عالی فرهنگ کیفیت در آموزش عالی از طریق طراحی و اجرایی‌سازی نظام ارزشیابی یکپارچه و پویا با مشارکت تمامی یاران آموزشی (ذی‌نفع، ذی‌ربط، ذی‌نفع، ذی‌علاقه) است (مالمیر و همکاران، ۱۴۰۲). بهبود کیفیت درون‌زای دانشگاهی از نظام ارزشیابی و عوامل مربوط به آن تأثیر می‌پذیرد و در این حالت باید شرایطی فراهم شود که دانشگاه‌ها خود برای الگوی ارزشیابی و ساختار مبتنی بر آن پیش‌قدم شوند و اهداف ارزشیابی به صورت درونی و بر اساس ضرورت‌های خاص مبتنی بر آن ایجاد کنند (فراست‌خواه و همکاران، ۱۳۸۶). بررسی فرهنگ ارزشیابی استادان، سازمان و عملکرد اداری به عنوان یکی از اصلی‌ترین عوامل هدایت‌کننده

۱. این نمودارها در فایل پیوست ارسال شده است

کیفیت یک مؤسسه آموزش عالی (Braga et al., 2014) موضوعی است که به‌رغم وجود برخی تلاش‌ها برای نظارت و به‌کارگیری نظام‌های ارزشیابی محدود، چون ارزشیابی استادان و کادر، هنوز به عنوان یک واقعیت با کاربرد در هدایت مفاهیم آموزش مورد توجه جدی قرار نگرفته است. ریشه این کم‌توجهی از یک سو به ناکارآمدی نظام مورد اجرا و ابهامات عملیاتی آن و از سوی دیگر به ضعف‌ها و خلأهای بررسی تحلیلی در این زمینه برمی‌گردد. به‌کارگیری فرهنگ ارزشیابی مطلوب در سیستم آموزش اولین مقوله برای اصلاح و بازسازی است و وجود و پذیرا بودن ارزشیابی و فرهنگ موجود در آن اصلی است که با اصلاح و بازخورد به متولیان می‌تواند در بهبود شرایط اثرگذار باشد (Scheerens, 2004). اولین قدم در شناخت و ارزیابی مسئله فرهنگ ارزشیابی ساخت ابزاری به منظور سنجش الگوی فرهنگ حاکم در سیستم است که پژوهش حاضر آن را به صورت عملیاتی مورد بررسی قرار داد. بر اساس مبانی نظری و بر اساس الگوی مصاحبه کیفی پرسشنامه‌ای ۴۰ گویه‌ای طراحی شد. این پرسشنامه دارای طیف لیکرتی ۶ درجه‌ای بود که ملاک سنجش وضعیت موجود بود و بر این اساس به هر گویه از ۱ تا ۶ نمره داده شد؛ به گونه‌ای که نمره ۶ اشاره به وجود آن ویژگی در حداکثر وضعیت موجود دانشکده و نمره ۱ اشاره به حداقل داشت. در این پرسشنامه در فرهنگ ارزشیابی ۷ جزء با ارزش ویژه بزرگ‌تر از ۱ باقی مانده‌اند که ۵۲/۳۷ درصد از واریانس کل را برآورد می‌کردند. در تبیینی از یافته‌های به‌دست‌آمده در گام اول اعتبار مناسب ابزار تحقیق ریشه در همسانی درونی ماده‌های آزمون بر اساس الگوی صحیح محتوایی و اجرایی دارد. تدوین گویه‌ها به نحوی که به‌درستی عینیت مفاهیم هدف را مورد سنجش در هر بعد قرار دهد زمینه‌ای مشخص در ایجاد روایی محتوایی و سازه و اتکای درونی ماده‌های آزمون با بالا رفتن میزان اعتبار بوده است. همچنین مدل تحقیق حاضر بر اساس تحلیل عاملی تأییدی مدل تفکیک‌شده هر خرده‌مقیاس را مورد تأیید قرار داد که حاکی از توانایی تمیز هر خرده‌مقیاس به نسبت مقیاس دیگر بوده است. در بررسی گویه‌های تحقیق بر اساس نظریه سؤال-پاسخ نتایج حاکی از آن بود که مدل PCM با داده‌های پرسشنامه برازش بهتری دارد. در این مرحله با استفاده از نرم‌افزار R و پکیج eRm پارامترهای شخص (تتا یا توانایی) و پارامترهای سؤال بر اساس مدل PCM برآورده شد. سپس پارامترهای آستانه‌ای بین طبقات، آماره کای دو، درجه آزادی و سطح معناداری هر سؤال گزارش شد. نمودارهای آگاهی حاکی از آگاهی مطلوب خرده‌مقیاس‌ها در اندازه‌گیری صفت مکنون در سطوح آگاهی متوسط بود. در بررسی مطالعاتی انجام‌گرفته در این زمینه، پژوهشی در زمینه ساخت ابزار در مسئله فرهنگ ارزشیابی موجود نیست. اما، در فرهنگ کیفیت، سیدکلان و همکارانش (۱۳۹۸) در مطالعه‌ای اکتشافی به ساخت ابزاری در زمینه فرهنگ کیفیت اقدام کردند که دارای ده عامل «عناصر ساختاری/مدیریتی»، «عناصر فرهنگی/ روان‌شناختی»، «رهبری»، «ارتباطات»، «اعتماد»، «تعهد و مسئولیت‌پذیری»، «مشارکت»، «توانمندسازها»، «اصلاح زیرساخت‌ها»، «نتایج و برآیندهای فرهنگ کیفیت» بود. این مطالعه از یک سو ملاک‌هایی از فرهنگ کیفیت را بررسی نموده است که جامعیت کاملی در تمامی ابعاد ندارد و از سوی دیگر منحصر به نمونه دو گروه از یاران آموزشی، یعنی اعضای هیئت علمی و کارکنان، بوده است و گروه‌های دیگر مورد بی‌توجهی قرار گرفته است. از سوی دیگر در این مطالعه ویژگی‌های آزمون فقط بر اساس نظریه کلاسیک آزمون بررسی شده است و برازش و اعتبار مطلوب به سبب حجم زیاد سؤالات را نیاز به بررسی مجدد می‌گرداند. از سوی دیگر مطالعه حاضر با تحقیقات رهنوما (۲۰۲۰) هم‌سویی در یافته‌های کلی به‌دست‌آمده دارد و تبیین‌کننده این مسئله است که ابزارسازی در زمینه فرهنگ کیفیت و ارزشیابی باید با اتکا به ابعاد فردی، سازمانی و کلان باشد و این پژوهش ضمن پیروی از این ابعاد لایه‌ها و سطوح اثرگذار در هر یک از ابعاد را نیز مورد بررسی قرار داده است و با تکیه بر این مسئله توانسته است توان تبیینی و سنجشی در این زمینه را افزایش دهد. در یک جمع‌بندی ابزار مورد بررسی به سنجش مسئله فرهنگ ارزشیابی به نسبت سطوح موجود در فرهنگ ارزشیابی در بخش‌های مختلف پرداخته است. فرهنگ ارزشیابی در هر سازمان و محیط آموزشی وجود دارد. اما وجود هر یک از این فرهنگ‌ها نتایجی دارد که شناخت الگوی حاکم و فرهنگ ارزشیابی موجود زمینه‌ای روشنگرایانه و تحلیلی برای مقایسه فرهنگ ارزشیابی موجود و ایده‌آل است و به سیاست‌گذاران حوزه اولین قدم برای اصلاح را ارائه می‌کند و گامی مهم برای برطرف‌سازی مشکلات و خلأهای موجود در این زمینه است. با توجه به نتایج در بخش پژوهشی پیشنهاد می‌شود پژوهش حاضر در دیگر دانشگاه‌های کشور انجام پذیرد و زمینه مقایسه دانشگاه‌های رده «الف» ایرانی با دیگر دانشگاه‌ها در وزارت علوم و نیز مقایسه دانشگاه‌های دولتی با آزاد و پیام نور و غیرانتفاعی فراهم شود تا

زمینه‌ای مهم از روشنگری عملیاتی در مفهوم نوین فرهنگ ارزشیابی در ایران فراهم آید. همچنین با توجه به نتایج به دست آمده و با توجه به نقش فرهنگ عمومی در فرهنگ کیفیت و ارزشیابی پیشنهاد می‌شود جهت حرکت به سوی بهبود فرهنگ کیفیت و ارزشیابی، سیاست‌گذاری نوین و فرهنگ‌سازی در زمینه افزایش نقدپذیری و فرهنگ رک بودن عمل شود تا زمینه به‌سازی در این زمینه انجام پذیرد. همچنین با توجه به سند توسعه ملی و چشم‌اندازهای مربوط به آن که بخش زیادی از آن مبتنی بر دستیابی به توسعه با ملاک‌های کمی محور است و با توجه به ضرورت وجود انعطاف در ملاک‌های کیفیت و ارزشیابی بر اساس شرایط مختلف، همچون همه‌گیری کرونا و ویروس، پیشنهاد می‌شود ملاک‌های مربوط به این سند مبتنی بر افزایش انعطاف و قرار دادن معیارهای هر بند بر اساس طیف‌های چنددرجه‌ای ایجاد شود تا هم زمینه شفافیت قوانین فراهم آید و هم از سوی دیگر روند مبتنی بر حرکت به سوی کیفیت با ساختار مشخص انجام پذیرد. پرواضح است برای ایجاد فرهنگ ارزشیابی مطلوب باید الگوی بستر ساز در این زمینه مورد نظر قرار گیرد و ضروری است سازمان‌های مربوط با حمایت‌های مالی و معنوی زمینه ایجاد تیم ارزشیاب حرفه‌ای در سازمان را ایجاد کنند.

منابع

- بنی‌سی، پریناز؛ دلفان آذری، قنبرعلی و رشاد، فائقه (۱۳۹۶). نگاهی نو به برنامه‌ریزی آموزشی و درسی در هزاره سوم. تهران: آوای نور.
- بنی‌اسدی، علی و صالحی، کیوان (۱۳۹۸). مقدمه‌ای بر اصول و فرایندهای ساخت و رواسازی پروتکل مصاحبه. *نامه آموزش عالی*، ۱۲(۴۶)، ۱۱-۲۰۳. https://journal.sanjesh.org/article_36128.html
- خادم‌محترم، خلیل و سلیمی، مهتاب (۱۳۹۸). ارزشیابی و تضمین کیفیت در نظام آموزش عالی. *مطالعات کاربردی در علوم مدیریت و توسعه*، ۱۵(۱۵)، ۳۴-۴۷.
- سیدکلان، سید محمد؛ زاهدبابلان، عادل؛ معینی‌کیا، مهدی و رضایی‌شریف، علی (۱۳۹۷). فرهنگ کیفیت در دانشگاه فرهنگیان؛ ارائه مدلی بر پایه نظریه داده‌بنیاد. *راهبرد فرهنگ*، ۱۱(۴۴)، ۱۴۴-۱۷۰.
- فراست‌خواه، مقصود (۱۳۸۷). آینده‌اندیشی درباره کیفیت آموزش عالی ایران؛ مدلی برآمده از نظریه مبنایی (GT). *پژوهش و برنامه‌ریزی در آموزش عالی*، ۵۰، ۶۷-۹۷.
- فراست‌خواه، مقصود؛ بازرگان، عباس و لوکس، کارو (۱۳۸۶). بررسی رابطه نظام‌های تضمین کیفیت آموزش عالی با زمینه‌های فرهنگی و اجتماعی (نقشه شناختی با تأکید بر مورد ایران). *علوم اجتماعی*، ۳۱، ۱-۲۰.
- مال‌میر، مرتضی؛ صالحی، کیوان؛ مقدم‌زاده، علی و دهقانی، مرضیه (۱۴۰۰). ابعاد و سطوح فرهنگ کیفیت در نظام آموزش عالی: مروری منظم. *مطالعات اندازه‌گیری و ارزشیابی آموزشی*، ۱۱(۳۶)، ۲۹-۵۳. Doi: 10.22034/emes.2021.251556
- مال‌میر، مرتضی؛ صالحی، کیوان؛ مقدم‌زاده، علی و دهقانی، مرضیه (۱۴۰۲). معنا و مفهوم فرهنگ کیفیت از منظر ذی‌نفعان در دانشگاه تهران. *مدیریت فرهنگ سازمانی*، ۲۱. Doi: 10.22059/jomc.2023.364285.1008585
- _____ (۱۴۰۲). اشتراک و افتراق مفهوم فرهنگ ارزشیابی در نظام‌های دانشگاهی به منظور توسعه سرمایه انسانی- اجتماعی. *مدیریت سرمایه اجتماعی*. Doi: 10.22059/jscm.2023.364330.2441
- Aburizaizah, S.J. (2022). The role of quality assurance in Saudi higher education institutions. *International Journal of Educational Research Open*, 3. <https://doi.org/10.1016/j.ijedro.2022.100127>.
- Angiel, J. & Pokojski, W. (2019). Education for sustainable development – from students and geography teachers knowledge to educational activities. *Miscellanea Geographica*, 23(1), 47-52. Doi: <https://doi.org/10.2478/mgrsd-2018-0026>.
- Astin, A.W. & Antonio, A.L. (2012). *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*. Rowman & Littlefield Publishers: Lanham, MD, USA.
- Balesh, O., Popov, M., Rizhov, N., Ryaskov, Y., Shaposhnikov, S., & Shestopalov, M. (2015). Research on University Education Quality Assurance: Methodology and Results of Stakeholders' Satisfaction Monitoring. *Social and Behavioral Sciences*, 214, 344-358.
- Banisi, P., Delfan azari, Q.A., & Rashad, F. (2019). *A new approach to instructional pedagogical planning in the 3rd millennium*. Tehran: Avay Noor. (in Persian).
- Bamber, C.J. & Elezi, E. (2020). What culture is your university? Have universities any right to teach entrepreneurialism?. *Higher Education Evaluation and Development*, 14, 19-32. <https://doi.org/10.1108/HEED-09-2018-0021>.
- 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. (in Persian)
- Bognar, B. & Bungić, M. (2014). Evaluation in Higher Education. *Život i škola*, 31(60), 139-159.
- Braga, M., Paccagnella, M., & Pellizari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, 41, 77-88.
- Brown, GTL (2019). Is Assessment for Learning Really Assessment?. *Front. Educ.* 4, 64. Doi: 10.3389/educ.2019.00064.
- Chaytor, K. (2023). Building Evaluation Culture—The Missing Link.
- Chen, H. (2023). Evaluation of Schools of University and Analysis of Obstacle Factors under the Background of High-Quality Development. *Sustainability*, 15, 3754. <https://doi.org/10.3390/su15043754>.
- Cao, C., Wei, T., Xu, S., et al. (2023). Comprehensive evaluation of higher education systems using indicators: PCA and EWM methods. *Humanit Soc Sci Commun*, 10, 432. <https://doi.org/10.1057/s41599-023-01938-x>
- Farastkhah, M. (2008). Future thinking about the quality of higher education in Iran; A model derived from grounded theory (GT). *Journal of Research and Planning in Higher Education*, 50, 67-97. (in Persian)
- Farstakhah, M., Bazargan, A., & Lux, K. (2006). Investigating the relationship between higher education quality assurance systems and cultural and social contexts, a cognitive map with an emphasis on the case of Iran. *Social Sciences*, 31, 1-20. (in Persian)

- Gandini, A. (2016). *The reputation economy: Understanding knowledge work in digital society*. London: Springer.
- Han, H.B. (2011). Application research on universities efficiency assessment based on the combination model of AHP and DEA. *Technol. Innov. Manag*, 32, 399–402.
- Huang, S.M., Wang, Y.H., & Chen, T.Q. (2022). A fuzzy comprehensive evaluation method on the construction of high-level universities based on data from Jiangsu Province. *Sci. Program.*, 9343954.
- Huber, S.G. & Skedsmo, G. (2017). Standardization and assessment practices. *Educ Asse Eval Acc* 29, 1–3. <https://doi.org/10.1007/s11092-017-9257-1>
- Jalaliyoon, N. & Taherdoost, H. (2012). Performance Evaluation of Higher Education; A Necessity. *Procedia-Social and Behavioral Sciences*, 46, 5682-5686.
- José Matus Gómez, O. (2024). Culture of evaluation: A proposal from the role of university teaching. *National Autonomous University of Nicaragua, Managua*, 13(36). <https://doi.org/10.5377/rtu.v13i36.17634>
- Kalinowska, A., Szkop, Z., & Wiśniewski, R. (2016). 'A comparative study of environmental awareness among students pursuing Bachelor's studies in selected academic fields at the University of Warsaw Poland'. *Environmental & Socioeconomic Studies*, 4(2), 17–25.
- Kalinowska, A. (2010). 'Rola studiów ochrony środowiska w upowszechnianiu międzynarodowych inicjatyw edukacyjnych ONZ [The role of environmental protection studies in dissemination of international UN educational initiatives] in Ochrona środowiska w programach studiów przyrodniczych [Environmental protection in natural science programs] eds B Gabryś & G Gabryś Uniwersytet Zielonogórski, 7–17.
- Kim, S. (2015). *An Analysis Of Teacher Question Types In Inquiry-Based Classroom and Traditional Classroom Settings*. Thesis for the Doctor of Philosophy degree in Teaching and Learning (Science Education) in the Graduate College of The University of Iowa.
- Khadam Mohtaram, K. & Salimi, M. (2018). Evaluation and quality assurance in the higher education system. *Bimonthly Scientific-Specialist Journal of Applied Studies in Management and Development Sciences*, 4(15), 34-47. (in Persian)
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. (2nd edition). New York, NY: Guilford Press.
- Malmir, M., Salehi, K., Moghadam Zadeh, A., & Dehghani, M. (2023). Meaning and Concept of Quality Culture Based on The View of Stakeholders at Tehran University. *Organizational Culture Management*, 4. Doi: 10.22059/jomc.2023.364285.1008585. (in Persian).
- Malmir, M., Salehi, K., Moghadamzadeh, A., & Dehghani, M. (2021). Dimensions and Levels Of Quality Culture in the Higher Education System: A Systematic Review. *Educational Measurement and Evaluation Studies*, 11(36), 29-53. Doi: 10.22034/emes.2021.251556. (in Persian)
- Malmir, M., Salehi, K., Moghadam Zadeh, A., & Dehghani, M. (2023). Commonality and differentiation of the concept of evaluation culture in university systems in order to develop human-social capital. *Social Capital Management*. Doi: 10.22059/jscm.2023.364330.2441. (in Persian)
- Mayne, J. (2008). *Building an Evaluative Culture for Effective Evaluation and Results Management*. Rome, Italy. Institutional Learning and Change (ILAC).
- Mayne, J. (2010). Building an evaluative culture: The key to effective evaluation and result management. *Canadian Journal of Program Evaluation*, 24(2), 1–30.
- Mayne, J. (2020). Building evaluative culture in community services: caring for evidence. *Evaluation and Program Planning*, 80, 101450. <https://doi.org/10.1016/j.evalprogplan.2017.05.011>
- Prudente, J. & Yunaningsih, A. (2023). Improving university competitiveness and quality education through Human Resource Development in the digital era. *Bedan Research Journal*, 8(1), 173–201. <https://doi.org/10.58870/berj.v8i1.51>
- Rahnuma, N. (2020). Evolution of quality culture in an HEI: critical insights from university staff in Bangladesh. *Educ Asse Eval Acc*, 32, 53–81.
- Rospigliosi, P.A. (2019). The role of social media as a learning environment in the fully functioning university: preparing for Generation Z. *Journal Interactive Learning Environments*, 27(4), 429-431.
- Salehi, K. & Golafshani, N. (2010). COMMENTARY: Using mixed methods in research studies: An opportunity with its challenges. *International journal of multiple research approaches*, 4(3), 186-191. <https://doi.org/10.5172/mra.2010.4.3.186>
- Scheerens, J. (2004). The Evaluation Culture. *Studies in Educational Evaluation*, 30, 105-124.
- Seyedkalan, S., Zahed-babolan, A., Moenikia, M., & Rezaeisharif, A. (2019). Quality Culture in Farhangian University: Presenting a Model Based on Grounded Theory. *Et*, 11(44), 147-170. Doi: 10.22034/jsfc.2019.92147. (in Persian)
- Simper, N., Martensson, S., Berry, A., & Maynard, N. (2021). Assessment cultures in higher education: reducing barriers and enabling change. *Assessment & Evaluation in Higher Education*, 4. <https://doi.org/10.1080/02602938.2021.1983770>

- Sun, J. (2020). Research on faculty and staff for constructing the “double first-class” universities based on Grey-AHP comprehensive evaluation model. *Grey Syst. Theory Appl*, 10, 467–478.
- Zhang, F. & Guo, S. (2014). The Role of Higher Education Evaluation System in the Senior Human Resources Development. *International Conference on Education, Language, Art and Intercultural Communication (ICELAIC 2014)*.