The University of Tehran Press

Journal of Foreign Language Research

Online ISSN: 2588-7521

Journal Homepage: https://jflr.ut.ac.ir/

Exploring the Effectiveness of Online Assessment Systems in Higher Education in Iran and Germany: A Qualitative Study

Sonya Anwar [1]✉ Mohammadreza Dousti Zadeh [2]. Stefan Piasecki [3]

1. Department of German Language and Literature, Faculty of Literature and Humanities Shahid Beheshti University, Tehran, Iran, Email: s_anvar@sbu.ac.ir

2- Department of German Language and Literature, Faculty of Foreign Languages, University of Tehran, Tehran, Iran. Email: dostizadeh@ut.ac.ir

3-Professor at the North Rhine-Westphalia Police and Public Administration University (HSPV NRW). Email: stefan.piasecki@hspv.nrw.de

Article Info

Article type:
Research Article

Article history:
Received: 25 August 2025
Received in revised form: 14 November 2025
Accepted: 15 November 2025
Available online : 22 November 2026

Keywords:
online assessment, higher education, qualitative research, lived experience, Iran and Germany, validity, e-learning

ABSTRACT

Objective: This study aimed to investigate the effectiveness of online assessments in higher education in Iran and Germany. The theoretical framework was grounded in interpretive and consequential validity as well as the concept of constructive alignment in instructional design.

Method: Conducted between 2024 and 2025, the research employed a qualitative phenomenological approach based on the lived experiences of university faculty members. Data were collected from 13 instructors through semi-structured face-to-face interviews in Iran and open-ended qualitative and audio questionnaires in Germany. The textual and audio data were analyzed using MAXQDA 2022 software, following Braun and Clarke's six-phase thematic analysis. To enhance trustworthiness, strategies such as data triangulation, participant review, and parallel coding by two researchers were applied.

Results: The findings revealed that, in addition to challenges such as exam security and cheating, reduced quality of feedback and educational interaction, technical difficulties, and unequal access, the participants also highlighted opportunities including temporal and spatial flexibility, cost efficiency, scalability of assessment, and functionality during crises. These themes were organized into three core components: learning-centeredness, equity, and feedback. A comparative analysis between Iran and Germany showed that although security and technical issues were common in both countries, their intensity was higher in Iran due to weaker technological infrastructure and lower digital readiness. Conversely, both countries experienced similar advantages such as flexibility, cost reduction, and the enhancement of digital skills. Based on these findings, several strategies were proposed, including redesigning assessments with an emphasis on open-ended and project-based questions, developing secure local platforms, training instructors in test design skills, and improving students' digital literacy.

Conclusions: These results can contribute to educational policymaking, enhancement of assessment quality, and strengthening of trust in online evaluation.

Cite this article: Anwar, S. Dousti Zadeh , M, & Piasecki, S. (2026). Exploring the Effectiveness of Online Assessment Systems in Higher Education in Iran and Germany: A Qualitative Study. *Journal of Foreign Language Research*, *16*(4), 459-476. https://doi.org/ 10.22059/jflr.2025.401256.1240

© Author(s) retain the copyright. **Publisher**: University of Tehran Press.
DOI: https://doi.org/ 10.22059/jflr.2025.401256.1240

انتشارات دانشگاه تهران

پژوهشهای زبانشناختی
در زبانهای خارجی

شاپا الکترونیکی: ۷۵۲۱-۲۵۸۸
Journal Homepage: https://jflr.ut.ac.ir/

واکاوی اثربخشی سیستم‌های سنجش آنلاین در آموزش عالی ایران و آلمان: یک مطالعهٔ کیفی

سونیا انور [۱] ✉ محمدرضا دوستی‌زاده [۲] اشتفان پیاسکی [۳]

۱. گروه زبان و ادبیات آلمانی دانشگاه شهید بهشتی ، دانشکده ادبیات و علوم انسانی دانشگاه شهید بهشتی ، تهران ، ایران. رایانامه: s_anvar@sbu.ac.ir
۲- گروه زبان و ادبیات آلمانی دانشگاه تهران ، دانشکده زبان ها و ادبیات خارجی دانشگاه تهران ، تهران ، ایران . رایانامه: dostizadeh@ut.ac.ir
۳-. استاد دانشگاه پلیس و مدیریت دولتی نوردراین-وستفالن ((دانشگاه پلیس و مدیریت عمومی نوردراین-وستفالن)) رایانامه: stefan.piasecki@hspv.nrw.de

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی
تاریخچه مقاله:
تاریخ دریافت: ۱۴۰۴/۰۶/۰۳
تاریخ بازنگری: ۱۴۰۴/۰۸/۲۳
تاریخ پذیرش: ۱۴۰۴/۰۸/۲۴
تاریخ انتشار: ۱۴۰۵/۰۹/۰۱

کلیدواژه‌ها:
سنجش آنلاین،
آموزش عالی،
پژوهش کیفی،
تجربهٔ زیسته،
ایران و آلمان،
اعتبار و روایی،
یادگیری الکترونیکی

چکیده

هدف: این پژوهش با هدف واکاوی اثربخشی سنجش‌های آنلاین در آموزش عالی ایران و آلمان انجام شد. چارچوب نظری مطالعه بر مبنای روایی استنباطی و پیامدی و مفهوم هم‌ترازی سازه‌ای در طراحی آموزشی استوار گردید.

روش: پژوهش در بازهٔ ۲۰۲۴ تا ۲۰۲۵ با رویکرد کیفی پدیدارشناسانه و بر اساس تجارب زیستهٔ اعضای هیئت علمی انجام گرفت. داده‌ها از طریق مصاحبه‌های نیمه‌ساختاریافتهٔ حضوری در ایران و پرسشنامه‌های کیفی بازمتن و صوتی در آلمان از ۱۳ استاد گردآوری شد. متون و فایل‌های صوتی با بهره‌گیری از نرم‌افزار MAXQDA 2022 و بر اساس روش تحلیل مضمون شش‌مرحله‌ای براون و کلارک بررسی گردید. برای ارتقای اعتمادپذیری، از راهبردهایی چون مثلث‌سازی داده‌ها، بازبینی مشارکت‌کنندگان و کدگذاری موازی توسط دو پژوهشگر استفاده شد.

یافته‌ها: یافته‌ها نشان داد که استادان علاوه بر چالش‌هایی نظیر امنیت آزمون و تقلب، کاهش کیفیت بازخورد و تعامل آموزشی، مشکلات فنی و نابرابری دسترسی، فرصت‌هایی همچون انعطاف‌پذیری زمانی و مکانی، صرفه‌جویی هزینه‌ای، مقیاس‌پذیری ارزیابی و کارکرد در شرایط بحران را برجسته کردند. این مضامین در سهٔ مؤلفهٔ اصلی سازمان‌دهی شدند: یادگیری محوری (عمق یادگیری و هم‌ترازی با اهداف آموزشی)، عدالت (شمول‌پذیری و برابری دسترسی) و بازخورد (کیفیت و فردی‌سازی). مقایسهٔ تطبیقی ایران و آلمان نشان داد که گرچه مسائل امنیتی و فنی در هر دو کشور مشترک‌اند، شدت آن‌ها در ایران به دلیل ضعف زیرساخت‌های فناوری و آمادگی دیجیتال بیشتر است. در مقابل، هر دو کشور مزایای مشابهی مانند انعطاف‌پذیری، کاهش هزینه‌ها و ارتقای مهارت‌های دیجیتال را تجربه کردند. بر اساس این تحلیل، راهکارهایی شامل بازطراحی آزمون‌ها با تأکید بر پرسش‌های بازپاسخ و پروژه‌محور، توسعهٔ پلتفرم‌های بومی امن، آموزش مهارت‌های طراحی آزمون به استادان و ارتقای سواد دیجیتال دانشجویان پیشنهاد شد.

نتیجه‌گیری: این یافته‌ها می‌تواند به سیاست‌گذاری آموزشی، ارتقای کیفیت ارزیابی و تقویت اعتماد به سنجش آنلاین یاری رساند.

استناد: انور، سونیا، دوستی‌زاده،محمدرضا و پیاسکی، اشتفان. (۱۴۰۵). واکاوی اثربخشی سیستم‌های سنجش آنلاین در آموزش عالی ایران و آلمان: یک مطالعه کیفی. *پژوهشهای زبانشناختی در زبانهای خارجی*، ۱۶ (۴)، ۴۵۹-۴۷۶. https://doi.org/ 10.22059/jflr.2025.401256.1240

ناشر: انتشارات دانشگاه تهران.

© نویسندگان.

۱. مقدمه

شتاب فزایندهٔ فناوری‌های آموزشی در سال‌های اخیر، نظام‌های سنجش آنلاین را از یک گزینهٔ جانبی به یکی از ارکان اصلی ارزیابی در آموزش عالی بدل کرده است. اگرچه این رویکرد مزایایی همچون انعطاف‌پذیری، دسترسی گسترده‌تر و کاهش هزینه‌ها را فراهم می‌آورد، همچنان پرسش‌های مهمی دربارهٔ اعتبار، اصالت پاسخ‌ها، عدالت آموزشی، حریم خصوصی و تجربهٔ کاربران مطرح است (باتلر–هندرسون و کرافورد[1]، ۲۰۲۰؛ هایل و ایفنتهالر[2]، ۲۰۲۳). پژوهش‌های کیفی اخیر نیز نشان داده‌اند که حتی در آموزش زبان‌های خارجی، اجرای برخط می‌تواند با بار شناختی مضاعف و ملاحظات ویژه‌ای در زمینهٔ روایی و کیفیت یادگیری همراه باشد (مونس قراملکی، آهنگری و حدیدی تمجید[3]، ۲۰۲۴) چنین یافته‌هایی نشان می‌دهد که صرف انتقال شیوه‌های سنتی آزمون به فضای دیجیتال کفایت نمی‌کند و طراحی و اعتباربخشی سنجش آنلاین باید با در نظر گرفتن تجربهٔ زیستهٔ اساتید و دانشجویان بازاندیشی شود. ضرورت این موضوع تنها از دل بحران‌هایی همچون پاندمی‌ها یا محدودیت‌های جغرافیایی ناشی نمی‌شود، بلکه در بسترهایی نظیر جنگ‌ها، تنش‌های سیاسی و اجتماعی و محدودیت‌های مالی نیز به‌عنوان تنها راه تداوم یادگیری مطرح است. بااین‌حال، مرور ادبیات نشان می‌دهد که بخش عمده‌ای از مطالعات موجود بیشتر بر جنبه‌های فناورانه یا امنیتی تمرکز داشته و کمتر به ابعاد کیفی و تجربه‌نگر سنجش آنلاین پرداخته‌اند (دشتستانی و کرمی[4]، ۲۰۱۹). این یافته‌ها در کنار مطالعات بین‌المللی، شکافی جدی را در زمینهٔ سنجش آنلاین از منظر تجربهٔ زیستهٔ اساتید دانشگاهی آشکار می‌سازد. از منظر نظری، اعتبار (روایی) آزمون نه ویژگی ابزار، بلکه ویژگی تفسیر و کاربرد نمره‌ها ست و نیازمند زنجیره‌ای از استدلال‌ها و شواهد معتبر است (مسیک[5]، ۱۹۹۵: ۷۴۱؛ کین[6]، ۲۰۱۳: ۲) و در بستر آزمون‌های آنلاین، این زنجیره باید علاوه بر انطباق با اهداف آموزشی و هم‌ترازی سازه‌ای (بیگز و تانگ[7]، ۲۰۱۱: ۴۵–۴۷)، شامل ملاحظاتی همچون یکپارچگی آکادمیک، عدالت و حفظ حریم خصوصی نیز باشد (دیویس و همکاران[8]، ۲۰۲۳: ۱۱۹؛ موتیموکوه و همکاران[9]، ۲۰۲۰: ۶۴–۶۵). با وجود این، مطالعاتی که به‌طور خاص به تجارب فراملی اساتید و مقایسهٔ ادراک آنان از اثربخشی سنجش آنلاین در بافت‌های فرهنگی گوناگون بپردازند، هنوز محدودند (هایل و ایفنتهالر، ۲۰۲۳: ۱۴–۱۵). بر این اساس، پژوهش حاضر با تمرکز بر دو بافت فرهنگی و نهادی متفاوت ایران و آلمان، در پی آن است که ابعاد اثربخشی، چالش‌ها و فرصت‌های سنجش آنلاین را از منظر اساتید دانشگاه واکاوی کند و راهکارهای کاربرمحور برای بهبود طراحی و اجرای آن ارائه دهد. در این پژوهش، تمرکز اصلی بر بررسی مؤلفه‌های ادراک شدهٔ اثربخشی سنجش آنلاین از دیدگاه اساتید دانشگاه در دو بافت فرهنگی ایران و آلمان است. در این راستا، مطالعه حاضر می‌کوشد ابعاد مختلف اثربخشی از جمله یادگیری‌محوری، بازخورد، عدالت و هم‌ترازی سازه‌ای را شناسایی و تبیین کند. همچنین، با واکاوی تجارب دانشگاهیان در دو کشور، تهدیدها و فرصت‌های مشترک و متمایز موجود در این دو بستر فرهنگی مورد بررسی قرار می‌گیرد. هدف دیگر پژوهش، استخراج راهبردهای کاربرمحور برای بهبود طراحی و اجرای سنجش‌های آنلاین است؛ به‌ویژه از منظر عواملی که می‌توانند اعتماد، عدالت و پذیرش دانشجویان را در فرآیند ارزیابی افزایش دهند. بر این اساس، این مطالعه در پی آن است که دریابد اساتید ایرانی و آلمانی اثربخشی سنجش آنلاین را چگونه معنا می‌کنند، چه چالش‌ها و عوامل کلیدی در هر کشور نقش پررنگ‌تری دارند و کدام وجوه اشتراک و افتراق در تجارب آنان برجسته‌تر است. در نهایت، پژوهش بر این گزاره استوار است که شفافیت معیارها و اصالت تکالیف می‌تواند به ارتقای اعتبار سنجش آنلاین بینجامد و همچنین، توجه به حریم خصوصی و عدالت شرط ضروری برای پذیرش سنجش‌های مراقبت‌شده به‌شمار می‌آید. در ادامه، ابتدا مبانی نظری و پژوهش‌های

---

1. Butler-Henderson & Crawford.
2. Heil & Ifenthaler.
3. Mounesi Qaramalki, Āhangarī & Haddīdi Tamjid.
4. Dashtestani & Karami.
5. Messick.
6. Kane.
7. Biggs & Tang.
8. Davis et al.
9. Mutimukwe et al.

پیشین مرور می‌شود. سپس روش تحقیق شامل شرکت‌کنندگان، ابزار گردآوری داده‌ها و شیوهٔ تحلیل توضیح داده می‌شود. بخش بعد به یافته‌ها اختصاص دارد که در قالب مضامین اصلی ارائه می‌شود. این یافته‌ها در بخش بحث با ادبیات مرتبط مقایسه و تفسیر می‌گردند. در پایان، نتیجه‌گیری شامل جمع‌بندی یافته‌ها، پیامدهای نظری و عملی، محدودیت‌ها و پیشنهادها مطرح خواهد شد.

۲. پیشینهٔ تحقیق

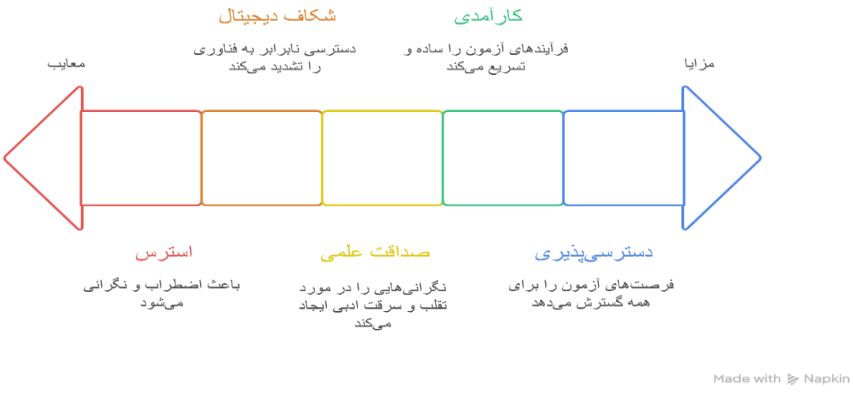
بحث سنجش آنلاین در آموزش عالی طی یک دههٔ اخیر به‌طور گسترده در ادبیات بین‌المللی بررسی شده است. چارچوب نظری این حوزه به‌ویژه بر سه بعد اساسی تأکید دارد: نخست، زیرساخت‌های فناورانه و سرمایه‌گذاری اولیه به‌عنوان پیش‌شرط کارآمدی؛ دوم، اعتبار و روایی[1] به‌عنوان بنیان تفسیر نمرات و نتایج (مسیک، ۱۹۹۵: ۷۴۱؛ کین، ۲۰۱۳: ۲) و سوم، تجربهٔ زیسته و نگرش ذی‌نفعان شامل دانشجویان و اساتید که مستقیماً بر پذیرش و اثربخشی سنجش آنلاین اثرگذار است (هایل و ایفنتهالر، ۲۰۲۳: ۱۴–۱۵). مرور مطالعات اخیر در ایران، آلمان و سایر کشورها نشان می‌دهد که با وجود نقاط اشتراک، هر بافت با چالش‌ها و فرصت‌های خاصی مواجه است. مطالعات بین‌المللی در فاصلهٔ سال‌های ۲۰۱۶ تا ۲۰۲۵ سه عامل کلیدی در اثربخشی سنجش‌های آنلاین را برجسته کرده‌اند: زیرساخت‌های فناورانهٔ قوی، آموزش و آمادگی کاربران و معیارهای شفاف طراحی سنجش. گاماگ و همکاران (گمج[2] و همکاران، ۲۰۲۲: ۱۴–۱۵) در مرور نظام‌مند کاربرد Moodle دریافتند که در صورت وجود ظرفیت‌های فنی و پشتیبانی مناسب، سنجش آنلاین می‌تواند موجب ارتقای مشارکت و عملکرد دانشجویان شود؛ هرچند هزینه‌های توسعه و مشکلات فنی همچنان مانع هستند. مظفر و همکاران[3] (مظفر و همکاران، ۲۰۲۰: ۷–۸) نیز نشان دادند که مسائل امنیتی و صداقت علمی، چالش‌هایی جهانی هستند که در بیشتر کشورها تکرار می‌شوند. مطالعات دیگری مانند لوریرو و گومز[4] (لوریرو و گومز، ۲۰۲۲: ۲۳۳–۲۳۴) و کریگ و کِی[5] (کریگ و کِی، ۲۰۲۵: ۴۹) بر تجربهٔ یادگیری و تعامل در بسترهای آنلاین متمرکز بوده و نقش همتایابی و ناشناس‌بودن را در ارتقای کیفیت بازخورد و انگیزش روشن ساخته‌اند. در حالی که تشومه[6] (تشومه، ۲۰۲۵: ۸۸–۸۹) و طاهر و همکاران[7] (۲۰۲۵: ۶۵–۶۶) نشان داده‌اند که آزمون‌های آنلاین در مرز میان کارآمدی و تهدید صداقت علمی قرار دارند و می‌توانند شکاف دیجیتال را تعمیق کنند. در ایران، پژوهش‌ها تصویر مکملی از چالش‌ها ارائه می‌دهند. عظیمی و همکاران[8] (۲۰۱۶: ۱۲۲–۱۲۳) در مطالعه‌ای تطبیقی میان ایران و روسیه نشان دادند که ضد سرقت علمی و پیش‌آزمون در ایران کمتر مورد توجه است، حال آنکه در روسیه این عناصر بنیان عدالت آموزشی محسوب می‌شوند. دشستانی (۲۰۲۰: ۸۵) نیز گزارش کرد که در حالی دانشجویان نگرش مثبتی به دوره‌های آنلاین دارند، مدرسان و مدیران دانشگاهی نسبت به اعتبار آزمون‌ها مردد هستند. عزیزی[9] (۲۰۲۲: ۴۱–۴۲) بر عدالت توزیعی، رویه‌ای و تعاملی به‌عنوان ارکان طراحی آزمون آنلاین در ایران تأکید کرده و ضعف زیرساخت اینترنتی و پایین‌بودن سواد دیجیتال را مانع جدی معرفی کرده است. رحیمی و همکاران[10] (۲۰۱۹: ۵۶) نیز به مشکلات فنی و قطع‌های مکرر اینترنت اشاره کرده‌اند که بر اعتبار آزمون‌ها سایه می‌اندازد. در آلمان، مطالعات سیاستی مانند پژوهش هایل و ایفنتالر (۲۰۲۳: ۱۶–۱۸) نشان داده‌اند که سرمایه‌گذاری در زیرساخت و تدوین معیارهای شفاف دو ضرورت اساسی برای تضمین کیفیت هستند. با این حال، داده‌های تجربی دانشگاهی محدود بوده است. برای رفع این شکاف، مطالعات میدانی و فنی در آلمان

---

1. validity & reliability.
2. Gamage et al.
3. Muzaffar et al.
4. Loureiro & Gomes.
5. Craig & Kay.
6. Teshome.
7. Tahir et al.
8. Azimi et al.
9. Azizi.
10. Rahimi et al.

بر مجموعه‌ای از ابزارهای سنجش آنلاین تمرکز کرده‌اند. برای نمونه، ریس و همکاران[۱] (۲۰۰۹: ۳) و گراسیگر و شنبل[۲] (۲۰۲۴: ۵) نشان داده‌اند که خودسنجی‌های آنلاین[۳] در شناسایی مهارت‌های مرتبط با تحصیل و ارتقای خودبازاندیشی مؤثرند. پتری و کرستینگ[۴] (۲۰۲۴: ۸) نیز کارکرد حمایتی این ابزارها در انتخاب رشته را برجسته کرده‌اند. مطالعاتی چون لمان و همکاران[۵] (۲۰۱۵: ۱۲) و باوئر و همکاران[۶] (۲۰۰۹: ۷۷) بر ارزیابی همتایان[۷] تأکید کرده‌اند که به بهبود کیفیت بازخورد و افزایش موفقیت تحصیلی منجر می‌شود. در حوزهٔ آزمون‌های خودکار، لن‌هارد و همکاران[۸] (۲۰۰۷: ۱۵۵) با استفاده از تحلیل معنایی نهفته نشان دادند که این روش همبستگی بالایی با نمره‌گذاری انسانی دارد، اگرچه چالش پذیرش پایین و محدودیت تحلیل معنایی پابرجا است. دوهمن و همکاران[۹] (۲۰۲۱: ۹۲) و یرمیاس و زلمر[۱۰] (۲۰۱۹: ۱۰۱) تجربهٔ آزمون‌های الکترونیک با سامانه‌هایONYX ، OPALو MAXIMAرا بررسی کرده و ضمن تأیید کارایی فنی، به مسائلی چون توزیع منابع، تغییر بار کاری و دغدغه‌های اخلاقی اشاره کرده‌اند. همچنین، کولر و همکاران[۱۱] (۲۰۲۰: ۴۸) و استریکروت و همکاران[۱۲] (۲۰۱۱: ۶۳) بر ظرفیت آزمون‌های سازگار و تشخیصی برای افزایش دقت سنجش شایستگی‌ها و پشتیبانی یادگیری تأکید کرده‌اند. در یک جمع‌بندی، مطالعات آلمان نشان می‌دهند که سنجش آنلاین نه‌تنها از نظر فنی امکان‌پذیر است بلکه به افزایش خودبازاندیشی، یادگیری تعاملی و رضایت دانشجویان کمک می‌کند. با این حال، پذیرش پایین برخی فناوری‌ها، محدودیت‌های نهادی و دغدغهٔ عدالت آموزشی همچنان چالش‌برانگیزند (رایس و همکاران، ۲۰۰۹: ۳؛ لهمان و همکاران، ۲۰۱۵: ۱۲؛ لن‌هارد و همکاران، ۲۰۰۷: ۱۵۵؛ دومن و همکاران، ۲۰۲۱: ۹۲) بدین ترتیب، در هر دو کشور ایران و آلمان، مسئلهٔ بازخورد و کیفیت تعامل به‌عنوان چالش مشترک باقی می‌ماند. مرور مطالعات پیشین نشان می‌دهد که: (۱) روایی و عدالت دو مؤلفهٔ محوری در سنجش آنلاین هستند؛ (۲) تفاوت‌های نهادی و فرهنگی در ایران و آلمان بر تجربهٔ اساتید تأثیرگذار است و (۳) شکاف قابل‌توجهی در مطالعات مقایسه‌ای ایران و آلمان وجود دارد. پژوهش حاضر با رویکرد کیفی و اتکا بر مصاحبه‌های حضوری در ایران و پرسشنامه‌های باز در آلمان، می‌کوشد این خلأ را پر کرده و تصویری روشن از وجوه اشتراک و افتراق در تجربهٔ زیستهٔ اساتید دو کشور ارائه دهد.

آزمون‌های آنلاین بین مزایا و معایب نوسان می‌کنند.



[1]. Reiss et al.
[2]. Grassinger & Schnebel.
[3]. Online Self-Assessments.
[4]. Petri & Kersting.
[5]. Lehmann et al.
[6]. Bauer et al.
[7]. Peer Assessment.
[8]. Lenhard et al.
[9]. Dohmen et al.
[10]. Jeremias & Sellmer.
[11]. Köller et al.
[12]. Strickroth et al.

۳. روش تحقیق

این پژوهش با رویکرد کیفی و با هدف واکاوی اثربخشی نظام‌های سنجش آنلاین در آموزش عالی ایران و آلمان طراحی و اجرا شد. انتخاب رویکرد کیفی بر این اساس صورت گرفت که موضوع مورد مطالعه نیازمند درک عمیق تجربه‌های زیستهٔ اساتید و کشف ابعاد معنایی پدیده است و نه صرفاً اندازه‌گیری کمّی متغیرها (کِرسوِل و پاث[1]، ۲۰۱۸: ۴۲). ازاین‌رو داده‌های آماری که در این فصل ارائه می‌شوند صرفاً نقش توصیفی دارند و برای معرفی ویژگی‌های مشارکت‌کنندگان به‌کار رفته‌اند، نه برای تحلیل آماری. داده‌های ارائه‌شده در این فصل کاملاً محصول پژوهش میدانی حاضر هستند و از طریق مصاحبه‌های مستقیم با اساتید ایرانی و آلمانی گردآوری شده‌اند. جامعهٔ پژوهش شامل ۱۳ نفر از اعضای هیئت علمی دانشگاه‌های ایران و آلمان بود که تجربهٔ تدریس حضوری و آنلاین داشتند و در طراحی یا اجرای سنجش‌های آنلاین مشارکت کرده بودند. در ایران، مصاحبه‌ها به‌صورت حضوری و در اتاق اساتید دانشگاه‌های تهران، علامه طباطبایی و شهید بهشتی انجام شد؛ این گفت‌وگوها با رضایت مشارکت‌کنندگان ضبط، سپس واژه‌به‌واژه پیاده‌سازی شدند. در آلمان، با توجه به پراکندگی جغرافیایی و محدودیت‌های زمانی، مصاحبه‌های نیمه‌ساختاریافته از طریق ارسال پرسش‌ها انجام گرفت و اساتید پاسخ‌های خود را به‌صورت کتبی و گاهی فایل‌های صوتی از طریق واتساپ یا ایمیل ارائه کردند. تمامی پاسخ‌ها پس از گردآوری، به متن کامل تبدیل و در نرم‌افزار MAXQDA برای تحلیل مضمون وارد شدند. برای تقویت اصالت و باورپذیری داده‌ها، نمونه‌ای از پرسشنامهٔ مصاحبهٔ نیمه‌ساختاریافته پیاده شده و نیز تصویر مراحل کدگذاری مضامین در پیوست مقاله جهت تحلیل شفاف‌تر مسیر تحقیق ارائه شده است. ویژگی‌های جمعیت‌شناختی مشارکت‌کنندگان نشان داد که ۷ نفر (۵۳٫۸۵٪) از ایران و ۶ نفر (۴۶٫۱۵٪) از آلمان بودند. از میان این افراد، ۹ نفر مرد (۶۹٫۲۳٪) و ۴ نفر زن (۳۰٫۷۷٪) بودند. ترکیب جنسیتی نشان داد که در آلمان سه زن و سه مرد حضور داشتند، در حالی که در ایران یک زن و شش مرد مشارکت داشتند. از نظر تحصیلات، ۱۲ نفر دارای مدرک دکتری (۹۲٫۳۱٪) و یک نفر دارای کارشناسی ارشد (۷٫۶۹٪) بودند. از حیث مرتبهٔ علمی نیز ۶ نفر استاد، ۴ نفر استادیار، یک نفر دانشیار، یک نفر مربی و یک نفر استاد تمام بودند. این تنوع تحصیلی و مرتبهٔ علمی امکان دسترسی به طیف متنوعی از دیدگاه‌ها را فراهم آورد. برای تحلیل داده‌ها، از روش تحلیل مضمون براساس الگوی براون و کلارک (۲۰۰۶: ۷۸) استفاده شد. مراحل تحلیل شامل پنج گام اصلی بود: خوانش مکرر داده‌ها برای آشنایی کامل، تولید کدهای اولیه، گروه‌بندی کدها در مقوله‌های مفهومی، استخراج مضامین فرعی و نهایتاً شکل‌دهی به مضامین اصلی. این فرآیند امکان شناسایی الگوهای معنایی و ابعاد متنوع اثربخشی سنجش آنلاین را فراهم کرد. به‌منظور تضمین اعتبار و اعتمادپذیری یافته‌ها، راهبردهای پیشنهادی لینکلن و گوبا[2] (۱۹۸۵: ۲۸۹) به‌کار گرفته شد. اعتبار[3] از طریق بازبینی یافته‌ها توسط مشارکت‌کنندگان[4] و تحلیل موازی توسط دو پژوهشگر مستقل تضمین شد. انتقال‌پذیری[5] با توصیف غنی زمینه‌های نهادی و فرهنگی در هر دو کشور تقویت گردید. پایایی[6] از طریق مستندسازی کامل مراحل پژوهش در دفترچه میدانی و تأییدپذیری[7] با بهره‌گیری از یادداشت‌های تأملی پژوهشگر و بازبینی همتایان حاصل شد. انجام این مطالعه در دو بستر فرهنگی و نهادی متفاوت امکان مقایسهٔ تطبیقی تجربه‌های اساتید ایرانی و آلمانی را فراهم آورد. یافته‌ها نشان دادند که در کنار چالش‌های مشترکی همچون دغدغه‌های امنیت، عدالت و اعتبار آزمون‌های آنلاین، تفاوت‌های فرهنگی و ساختاری نیز برجسته هستند. این طراحی تطبیقی موجب شد نتایج پژوهش نه‌تنها به درک عمیق‌تر از سنجش آنلاین در بافت‌های خاص منجر شود، بلکه بتواند به بازاندیشی سیاست‌های کلان آموزشی و ارتقای کیفیت طراحی و اجرای آزمون‌های آنلاین در آموزش عالی نیز کمک کند.

---

[1] Creswell & Poth
[2] Lincoln & Guba
[3] credibility
[4] member checking
[5] transferability
[6] dependability
[7] confirmability

جدول ۱. ویژگی‌های جمعیت‌شناختی مشارکت‌کنندگان، (منبع: داده‌های جامعهٔ پژوهش، ۲۰۲۴–۲۰۲۵)

کشور	ردیف	شرکت‌کننده کد	جنسیت	دانشگاه	تحصیلات	موقعیت شغلی / مرتبهٔ علمی
آلمان	1	T.N1	زن	Fachhochschule Polizei Sachsen-Anhalt	PhD	پروفسور، استاد دانشگاه
	2	T.N2	مرد	Uni Duisburg / Uni Leipzig / Dr. phil / Uni Kassel		
	3	T.N3		CVJM-Hochschule Kassel / YMCA-University of Applied Sciences Kassel		
	4	T.N4	زن	Hochschule für Polizei und öffentliche Verwaltung NRW		
	5	T.N5	مرد	Universität Paderborn		
	6	T.N6	زن	CVJM-Hochschule Kassel / Universität Kassel		
ایران	7	T.N7	مرد	دانشگاه تهران	PhD	هیئت علمی، استادیار
	8	T.N8	مرد		PhD	هیئت علمی، دانشیار
	9	T.N9			Master	هیئت علمی، مربی
	10	T.N10		دانشگاه علامه طباطبایی	PhD	هیئت علمی، استاد تمام
	11	T.N11				هیئت علمی، استادیار
	12	T.N12	زن	دانشگاه شهید بهشتی		
	13	T.N13	مرد			

جدول ۲. آمار توصیفی مشارکت‌کنندگان، (منبع: داده‌های جامعهٔ پژوهش، ۲۰۲۴–۲۰۲۵)

متغیر	گروه	تعداد	(%) درصد
کشور	ایران	7	53.85
	آلمان	6	46.15
جنسیت	مرد	9	69.23
	زن	4	30.77
کشور × جنسیت	آلمان – زن	3	23.08
	آلمان – مرد	3	23.08
	ایران – زن	1	7.69
	ایران – مرد	6	46.15
تحصیلات	PhD	12	92.31
	Master	1	7.69
موقعیت شغلی	Professor	6	46.15
	Assistant Prof.	4	30.77
	Associate Prof.	1	7.69

	Lecturer	1	7.69
	Full Prof.	1	7.69
کل	—	13	100.00

۴. نتایج و بحث و بررسی

یافته‌های این پژوهش نشان داد که تجربهٔ سنجش آنلاین در دو بستر فرهنگی و آموزشی متفاوت، یعنی ایران و آلمان، با مجموعه‌ای از چالش‌ها، فرصت‌ها و راهکارهای پیشنهادی همراه است. اگرچه بسیاری از این مسائل ماهیت جهانی دارند، شدت و جلوهٔ آن‌ها در هر کشور بسته به سطح توسعهٔ زیرساخت‌های دیجیتال، فرهنگ نهادی و سیاست‌های آموزشی متفاوت است.

چالش‌های سنجش آنلاین

نخستین مضمون اصلی به چالش‌های سنجش آنلاین اختصاص داشت. یکی از پررنگ‌ترین دغدغه‌ها، موضوع امنیت آزمون و تقلب بود. مشارکت‌کنندگان بارها به این نگرانی اشاره داشتند که آیا فردی که در آزمون آنلاین شرکت می‌کند واقعاً همان دانشجوی ثبت‌نام‌شده است یا شخص دیگری جای او پاسخ می‌دهد. این نگرانی در ایران به دلیل ضعف در زیرساخت‌های نظارت تصویری و محدودیت سامانه‌های کنترل از راه دور برجسته‌تر مطرح شد. چنان‌که یکی از مشارکت‌کنندگان ایرانی تأکید کرد: «مسئلهٔ تقلب در ایران یک دغدغهٔ جدی است؛ به‌خصوص وقتی زیرساخت‌های نظارت تصویری کافی وجود ندارد» (T.N7، ایران). در آلمان نیز هرچند شرایط نظارتی بهتر است؛ اما همچنان تردید نسبت به اصالت پاسخ‌ها باقی است (T.N2، آلمان). چالش دیگر، افت کیفیت تعامل و بازخورد بود. استادان در هر دو کشور معتقد بودند که فضای آنلاین امکان ارائه بازخورد فردی و عمیق را محدود می‌سازد و کیفیت تعامل میان استاد و دانشجو در آزمون‌های آنلاین به‌ویژه در بخش‌های تشریحی پایین‌تر از کلاس حضوری است. یکی از استادان آلمانی در این زمینه اشاره کرد: «ما نمی‌توانیم همان‌طور که در کلاس حضوری بازخورد عمیق و شخصی بدهیم، در بستر آنلاین این کار را انجام دهیم» (T.N1، آلمان). افزون بر این، مشکلات فنی و محدودیت زیرساختی از دیگر موانع اساسی بود. در ایران مشکلاتی مانند قطع اینترنت، کمبود تجهیزات در منازل دانشجویان و ناسازگاری نرم‌افزارها بیشتر گزارش شد (T.N9، ایران). در آلمان نیز با وجود امکانات گسترده‌تر، خطر از کار افتادن پلتفرم‌ها در آزمون‌های بزرگ وجود داشت که این امر به‌عنوان یک ریسک جدی مطرح شد (T.N5، آلمان).

فرصت‌های سنجش آنلاین

با وجود چالش‌های یادشده، سنجش آنلاین فرصت‌های قابل توجهی را نیز در اختیار نظام‌های آموزشی قرار داد. مهم‌ترین مزیت، انعطاف‌پذیری زمانی و مکانی بود. دانشجویان می‌توانستند از هر مکان و در شرایط مختلف در آزمون‌ها شرکت کنند، امکانی که به‌ویژه برای دانشجویان شاغل یا کسانی که در شهرهای دیگر سکونت دارند بسیار ارزشمند تلقی شد (T.N6، آلمان؛ T.N11، ایران). دومین فرصت، کاهش هزینه‌ها و منابع بود. برگزاری آزمون به‌صورت آنلاین موجب صرفه‌جویی قابل توجه در زمان و هزینه‌های دانشگاه‌ها شد، چرا که نیازی به تخصیص مکان‌های فیزیکی و نیروی انسانی گسترده برای نظارت وجود نداشت. این مزیت نه‌تنها برای دانشگاه‌ها، بلکه برای دانشجویان نیز اهمیت داشت؛ زیرا هزینهٔ رفت‌وآمد و اقامت حذف می‌شد (T.N3، آلمان؛ T.N12، ایران). فرصت دیگر به امکان ارزیابی گسترده‌تر و مقیاس بزرگ‌تر بازمی‌گشت. دانشگاه‌ها توانستند هم‌زمان صدها دانشجو را ارزیابی کنند؛ امری که در شرایط حضوری با محدودیت‌های فیزیکی و لجستیکی عملاً غیرممکن بود (T.N4، آلمان؛ T.N10، ایران). در ایران، به‌ویژه در دوران همه‌گیری کرونا، این ویژگی به‌عنوان یک مزیت کلیدی مورد تأکید قرار گرفت.

راهکارهای ارتقای اثربخشی سنجش آنلاین

مضمون سوم به راهکارهای پیشنهادی برای ارتقای اثربخشی مربوط بود. مشارکت‌کنندگان معتقد بودند که طراحی آزمون‌ها باید به سمت کاربرمحوری و الگوهای ترکیبی[1] حرکت کند. بازطراحی آزمون‌ها بر مبنای مسائل باز و کاربردی، در کنار استفاده از

[1] Hybrid Design

ترکیب حضوری و آنلاین، می‌تواند امنیت و اعتبار بیشتری به آزمون‌ها بدهد (T.N2، آلمان؛ T.N13، ایران). راهکار دوم، تقویت زیرساخت فناورانه و امنیتی بود. به‌کارگیری نرم‌افزارهای معتبر بین‌المللی، سرمایه‌گذاری در سرورهای قدرتمند و توسعهٔ پلتفرم‌های بومی امن، به‌ویژه در ایران، از الزامات دانسته شد (T.N5، آلمان؛ T.N7، ایران). راهکار سوم به توانمندسازی استادان و دانشجویان اختصاص داشت. آموزش مهارت‌های فناورانه، فرهنگ‌سازی و ارتقای سواد دیجیتال، شرایطی را فراهم می‌کند که نگرانی‌ها کاهش یابد و استادان بتوانند طراحی آزمون‌های متناسب با فضای آنلاین را بیاموزند. همان‌گونه که یک مشارکت‌کنندهٔ ایرانی اشاره کرد: «اساتید باید یاد بگیرند چگونه آزمون آنلاین طراحی کنند و فقط انتقال آزمون حضوری به آنلاین کافی نیست» (T.N8، ایران).

تفسیر یافته‌ها

تحلیل تطبیقی نتایج نشان داد که اگرچه چالش‌های امنیتی، تعامل آموزشی و محدودیت‌های فنی در هر دو کشور مطرح هستند، شدت آن‌ها در ایران بیشتر احساس می‌شود. این امر تا حد زیادی ناشی از ضعف زیرساخت‌های فناوری و سطح پایین‌تر آمادگی دیجیتال است. در مقابل، مزایایی چون انعطاف‌پذیری، کاهش هزینه‌ها و امکان ارزیابی در مقیاس وسیع‌تر به‌طور مشترک در هر دو کشور تجربه شد. این یافته‌ها با پژوهش‌های پیشین نیز همخوانی دارند. به‌عنوان نمونه، کریسپ و وارد[1] (۲۰۲۲) و جاگین[2] (۲۰۱۰) بر اهمیت طراحی نوآورانهٔ آزمون‌ها در ارتقای اعتبار و کاهش آسیب‌پذیری تأکید کرده‌اند. همچنین، بر اساس چارچوب نظری لینکلن و گوبا (۱۹۸۵)، اعتمادپذیری داده‌ها از طریق مثلث‌سازی منابع و بازبینی مشارکت‌کنندگان تقویت شد. نوآوری این پژوهش در مقایسهٔ بین‌فرهنگی ایران و آلمان نهفته است که نشان می‌دهد فرهنگ نهادی، سیاست‌های آموزشی و سطح توسعهٔ زیرساخت دیجیتال نقشی تعیین‌کننده در تجربهٔ سنجش آنلاین ایفا می‌کنند.

جدول ۳. تحلیل مضمون کیفی

مضامین اصلی (Themes)	مضامین فرعی (Sub-Themes)	کدهای اولیه (Initial Codes)	دادهٔ خام نقل‌قول / (Raw Data (مشارکت‌کنندگان
چالش‌های سنجش آنلاین (Challenges)	امنیت آزمون (Exam Security)	تقلب، اصالت پاسخ (Cheating, Authenticity)	همیشه نگران هستیم که آیا دانشجو خودش امتحان می‌دهد یا فرد دیگری (T.N2 – آلمان)
	امنیت آزمون	فقدان زیرساخت پروکتورینگ (Lack of Proctoring Infrastructure)	«در ایران زیرساخت کافی برای نظارت تصویری وجود ندارد» (T.N7 – ایران)
	افت کیفیت بازخورد (Feedback Decline)	بازخورد ناکافی (Low-Quality Feedback)	«کیفیت بازخورد پایین‌تر است» (T.N8 – ایران)
	کاهش تعامل آموزشی (Reduced Pedagogical Interaction)	تعامل کم (Low Interaction)	«ارتباط استاد و دانشجو ضعیف‌تر است» (T.N1 – آلمان)
	مشکلات فنی (Technical Difficulties)	قطع اینترنت، ناپایداری (Connectivity Issues)	«اینترنت در میانهٔ آزمون قطع شد» (T.N9 – ایران)
	مشکلات پلتفرمی (Platform Instability)	سقوط سرور (Server Crash)	«پلتفرم‌ها در آزمون‌های بزرگ از کار می‌افتند»

[1] Crisp & Ward
[2] Joughin

			(آلمان – T.N5)
	نبود کنترل منابع (Resource Control Issues)	دسترسی به منابع غیرمجاز (Unauthorized Materials)	دانشجویان به منابع غیرمجاز دسترسی دارند» «(ایران – T.N11)
	ناکافی بودن نظارت (Inadequate Supervision)	نظارت ناکافی (Insufficient Monitoring)	«امکان نظارت تصویری محدود است» (ایران – T.N12)
	آسیب‌پذیری محتوا (Content Vulnerability)	اشتراک‌گذاری سؤالات (Question Sharing)	سؤالات تستی به‌راحتی قابل اشتراک‌گذاری‌اند» «(آلمان – T.N3)
	امنیت محتوا (Content Security)	سرقت محتوای آزمون (Exam Content Theft)	دانشجوها اسکرین‌شات می‌گیرند و پخش می‌کنند» «(ایران – T.N13)
	عدالت آموزشی (Educational Justice)	نابرابری دسترسی (Unequal Access)	در ایران عدالت آموزشی زیر سؤال می‌رود چون همه دسترسی یکسان ندارند» «(ایران – T.N10)
	استرس و اضطراب دانشجویان (Student Stress)	فشار روانی ناشی از فناوری (Tech Anxiety)	دانشجویان با اینترنت ضعیف دچار استرس می‌شوند» «(ایران – T.N7)
	اعتمادپذیری پایین (Low Reliability)	بی‌اعتمادی به سیستم (Distrust in System)	وقتی سرور می‌خوابد، اعتماد دانشجو از بین می‌رود» «(آلمان – T.N4)
	اعتبار آزمون (Exam Credibility)	اعتبار اجتماعی پایین (Low Social Credibility)	دانشجویان می‌گویند آزمون آنلاین جدی گرفته نمی‌شود» «(آلمان – T.N6)
	توانمندی پایین استادان (Low Faculty Competence)	کمبود مهارت طراحی آزمون (Lack of Test Design Skills)	اساتید ابزار کافی برای طراحی آزمون ندارند» «(ایران – T.N8)
فرصت‌های سنجش آنلاین (Opportunities)	انعطاف زمانی و مکانی (Flexibility of Time & Place)	دسترسی از راه دور (Remote Access)	دانشجویان می‌توانند از هر کجا در آزمون شرکت کنند» «(آلمان – T.N6)
	دسترسی گسترده‌تر (Greater Accessibility)	دسترسی برای مناطق محروم (Access for Remote Areas)	این انعطاف برای دانشجویان شهرستانی در ایران خیلی ارزشمند است» «(ایران – T.N11)
	صرفه‌جویی هزینه‌ای (Cost Efficiency)	کاهش هزینهٔ نهادی (Institutional Cost Reduction)	دانشگاه‌ها هزینه‌های آزمون حضوری را کاهش می‌دهند» «(آلمان – T.N3)
	صرفه‌جویی فردی (Personal Cost Reduction)	کاهش هزینهٔ فردی (Reduced Student Expenses)	دانشجو هم دیگر هزینهٔ رفت‌وآمد نمی‌پردازد» «(ایران – T.N12)
	ارزیابی در مقیاس وسیع (Large-Scale Assessment)	مقیاس‌پذیری (Scalability)	توانستیم چند صد دانشجو را همزمان ارزیابی کنیم» «(آلمان – T.N4)

	کارکرد در بحران‌ها (Relevance in Crisis Contexts)	ضرورت در بحران (Crisis Necessity)	« این امکان در دوران کرونا در ایران خیلی مهم بود» (ایران – T.N10)
	تنوع فرهنگی (Cultural Diversity)	مقایسهٔ تطبیقی (Comparative Perspective)	«امکان مقایسهٔ بین‌فرهنگی فراهم شد» (آلمان – T.N2)
	تحلیل سریع‌تر (Faster Evaluation)	تسهیل تحلیل داده (Ease of Data Analysis)	« اساتید راحت‌تر می‌توانند نتایج را تحلیل کنند» (آلمان – T.N5)
	مدیریت زمان بهتر (Time Management)	کاهش زمان (Time Efficiency)	« امتحانات آنلاین باعث صرفه‌جویی در زمان می‌شود» (ایران – T.N9)
	عدالت دسترسی (Accessibility & Equity)	شمول‌پذیری (Inclusion)	« دانشجویان با مشکلات حرکتی هم می‌توانند شرکت کنند» (ایران – T.N7)
	امکان استفادهٔ ثانویه (Secondary Use)	ذخیره‌سازی داده‌ها (Data Archiving)	« داده‌ها برای پژوهش‌های بعدی ذخیره می‌شوند» (ایران – T.N13)
	رشد مهارت‌های دیجیتال (Digital Competence)	ارتقای سواد دیجیتال (Digital Literacy)	« دانشجوها با محیط دیجیتال بیشتر آشنا می‌شوند» (ایران – T.N8)
	آمادگی برای آینده (Future Readiness)	پیوند با آموزش ترکیبی (Blended Learning)	« این تجربه می‌تواند بخشی از آموزش ترکیبی آینده باشد» (آلمان – T.N1)
راهکارهای بهبود (Strategies)	طراحی نوآورانهٔ آزمون (Innovative Assessment Design)	نوآوری در طراحی (Innovative Test Design)	« اگر طراحی آزمون تغییر کند، امنیت هم بیشتر خواهد شد» (ایران – T.N13)
	مدل ترکیبی (Hybrid/Blended Models)	ترکیب حضوری و آنلاین (Hybrid Assessment)	«باید بخشی از آزمون‌ها حضوری بمانند» (آلمان – T.N2)
	بومی‌سازی فناوری (Localization)	توسعهٔ پلتفرم بومی (Local Platform Development)	« بدون پلتفرم بومی قوی، تقلب همیشه هست» (ایران – T.N7)
	تقویت زیرساخت (Infrastructure Development)	سرمایه‌گذاری دیجیتال (Digital Infrastructure Investment)	« سرمایه‌گذاری در زیرساخت ضروری است» (آلمان – T.N5)
	توانمندسازی استادان (Faculty Empowerment)	ارتقای مهارت استادان (Faculty Training)	« اساتید باید مهارت آزمون‌سازی بیاموزند» (ایران – T.N8)
	توانمندسازی دانشجویان (Student Empowerment)	آموزش دانشجویان (Student Training)	« برای دانشجویان کارگاه آموزشی برگزار کردیم» (آلمان – T.N6)
	تنوع در ارزیابی (Diversification of Assessment)	آزمون بازپاسخ (Open-Ended Exams)	« استفاده از آزمون‌های بازپاسخ تقلب را سخت‌تر می‌کند» (آلمان – T.N3)
	یادگیری عمیق‌تر (Deeper Learning)	آزمون مبتنی بر پروژه	«آزمون‌های پروژه‌محور مفیدتر هستند» (ایران – T.N9)

	(Project-Based Assessment)		
ارزیابی مستمر بهتر از یک آزمون نهایی است» «(ایران – T.N12)	ارزیابی مستمر (Continuous Assessment)	فرآیندمحوری (Process Orientation)	
«باید استانداردهای اخلاقی تعریف شود» (ایران – T.N11)	دستورالعمل اخلاقی (Ethical Guidelines)	اخلاق حرفه‌ای (Academic Integrity)	
نظارت تصویری هوشمند کمک می‌کند» «(آلمان – T.N4)	هوش مصنوعی در نظارت (AI Proctoring)	فناوری پیشرفته (Advanced Technology)	
دانشگاه‌ها باید سیاست‌های هماهنگ داشته باشند» «(ایران – T.N10)	سیاست‌گذاری کلان (Policy Alignment)	هماهنگی نهادی (Institutional Consistency)	
بازبینی نتایج توسط استاد و نرم‌افزار همزمان باشد» «(آلمان – T.N1)	بازبینی ترکیبی (Dual Evaluation)	چندلایه بودن ارزیابی (Multi-Layered Evaluation)	
تجربه‌های بین‌المللی باید بومی‌سازی شوند» «(آلمان – T.N2)	بومی‌سازی تجارب جهانی (Localization of Global Practices)	تطبیق فرهنگی (Cultural Adaptation)	
باید انجمن مشترک استادان برای طراحی آزمون تشکیل شود» «(ایران – T.N13)	همکاری جمعی (Collaborative Design)	هم‌افزایی علمی (Collaborative Improvement)	

۵. نتیجه‌گیری

نوآوری این پژوهش در رویکرد مقایسه‌ای و تجربه‌نگر آن نهفته است؛ بدین معنا که با اتکا به روایت استادان دانشگاه‌های ایران و آلمان، توانسته است تصویری چندوجهی از چالش‌ها، فرصت‌ها و راهکارهای ارتقای سنجش آنلاین در آموزش عالی ترسیم کند. یافته‌ها نشان داد که چالش‌های اصلی عمدتاً در سهٔ حوزه متمرکز هستند: امنیت و اصالت پاسخ‌ها، نگرانی از تقلب، ناکافی بودن نظارت و آسیب‌پذیری محتوای آزمون‌ها؛ (T.N2–آلمان، T.N7–ایران)، زیرساخت و پایداری فنی (قطع اینترنت، سقوط سرورها؛ T.N4 –آلمان، T.N9–ایران) و تعامل و بازخورد، کاهش کیفیت بازخورد فردی و ضعف تعامل آموزشی؛ (T.N1– آلمان، T.N8– ایران) این یافته‌ها همسو با مطالعات بین‌المللی است که بر ضرورت طراحی نوآورانهٔ آزمون‌ها برای ارتقای اعتبار و کاهش تقلب تأکید کرده‌اند (کریسپ و وارد، ۲۰۲۲؛ جوگین، ۲۰۱۰) در برابر این چالش‌ها، فرصت‌های ارزشمندی نیز شناسایی شد. مهم‌ترین آن‌ها انعطاف‌پذیری زمانی و مکانی، صرفه‌جویی در هزینه‌ها، مقیاس‌پذیری آزمون‌ها و کارکرد در بحران‌ها بود (T.N3– آلمان؛ – T.N10 ایران). همچنین، پیامدهای مثبتی همچون ارتقای سواد دیجیتال، شمول‌پذیری برای دانشجویان دارای محدودیت حرکتی و امکان ذخیره‌سازی داده‌ها برای تحلیل‌های بعدی مطرح شد (T.N7–ایران؛ T.N13 –ایران). این نتایج با پژوهش‌های آلمانی در حوزهٔ خودسنجی‌های آنلاین و آزمون‌های الکترونیک نیز همخوانی دارد (رایس و همکاران، ۲۰۰۹: ۳؛ لهمان و همکاران، ۲۰۱۵: ۱۲؛ گراسیگر و اشنبل، ۲۰۲۴: ۵؛ دومن و همکاران، ۲۰۲۱: ۹۲) از منظر عدالت آموزشی، یافته‌ها نشان دادند که ضعف زیرساخت اینترنتی و نابرابری دسترسی به تجهیزات در ایران، فشار روانی و بی‌اعتمادی به سامانه‌ها را تشدید می‌کند (ایران – T.N10؛ T.N7– ایران) در مقابل، انعطاف‌پذیری زمانی و مکانی، کاهش هزینه‌ها و امکان مشارکت دانشجویان مناطق محروم یا دارای محدودیت‌های جسمانی، ظرفیت‌هایی برای افزایش شمول فراهم می‌سازد (آلمان – T.N6؛ T.N11–ایران). این دوگانه

نشان می‌دهد که اعتبار پیامدی سنجش آنلاین (مسیک، ۱۹۹۵: ۷۴۱؛ کین، ۲۰۱۳: ۲) به شدت به سیاست‌های حمایتی و سرمایه‌گذاری در زیرساخت‌های ICT وابسته است.

یافته‌ها همچنین بر اهمیت بازخورد و اعتمادپذیری تأکید داشتند. افت کیفیت بازخورد و سطحی‌شدن تعاملات آموزشی در هر دو کشور گزارش شد؛ با این حال، پیشنهادهایی همچون بازبینی چندلایه (استاد + نرم‌افزار)، آموزش مهارت‌های آزمون‌سازی به استادان و استفاده از داده‌های ذخیره‌شده برای تحلیل‌های تکمیلی می‌تواند به ترمیم چرخهٔ بازخورد و افزایش اعتماد نهادی کمک کند (T.N1– آلمان؛ T.N8 – ایران؛ T.N13 – ایران). این موضوع با یافته‌های شین[1] (۲۰۲۲: ۷۵) و مارتین و همکاران[2] (۲۰۲۰: ۱۱۲–۱۱۴) دربارهٔ کمبود بازخورد فردی و اهمیت عدالت دیجیتال مطابقت دارد. در سطح نظری، نتایج این پژوهش تأیید می‌کند که زنجیرهٔ روایی مسیک (۱۹۹۵) و چارچوب استنباطی کین (۲۰۱۳) در بستر آزمون‌های آنلاین با چالش‌های تازه‌ای مواجه می‌شوند؛ به‌ویژه در ارتباط با روایی استنباط[3] و روایی پیامدی[4]. به‌علاوه، تحلیل داده‌ها بر اساس چارچوب تحلیل مضمون براون و کلارک[5] (۲۰۰۶: ۷۸–۷۹) و معیارهای اعتمادپذیری لینکلن و گوبا (۱۹۸۵: ۲۸۹–۳۳۱) انجام شد که اعتبار نتایج را تقویت می‌کند.

راهکارهای پیشنهادی این پژوهش در سه سطح دسته‌بندی شدند:

- سیاست‌گذاری کلان: توسعهٔ پلتفرم‌های بومی، سرمایه‌گذاری در زیرساخت‌های فناورانه و هماهنگی نهادی (T.N5– آلمان؛ T.N7 – ایران؛ T.N10– ایران)
- توانمندسازی منابع انسانی: ارتقای مهارت‌های فناورانه و طراحی آزمون برای استادان و آموزش مهارت‌های دیجیتال به دانشجویان (T.N6–آلمان؛ T.N8– ایران)
- نوآوری در طراحی سنجش: آزمون‌های بازپاسخ و پروژه‌محور، الگوهای ترکیبی حضوری–آنلاین، ارزیابی مستمر و بازبینی چندلایه (T.N2–آلمان؛ T.N3– آلمان؛ T.N9– ایران)

به‌طور کلی، پژوهش حاضر نشان داد که تجربهٔ سنجش آنلاین در ایران و آلمان هم وجوه اشتراک دارد و هم تفاوت؛ چالش‌های امنیتی و زیرساختی در ایران پررنگ‌تر است، در حالی که هر دو کشور مزایای مشترکی مانند انعطاف‌پذیری و صرفه‌جویی هزینه‌ای را تجربه می‌کنند. بر این اساس می‌توان نتیجه گرفت که تنها از رهگذر طراحی نوآورانه، تقویت زیرساخت‌ها و توانمندسازی استادان و دانشجویان می‌توان به سمت تثبیت سنجش آنلاین به‌عنوان الگویی معتبر و پایدار در آموزش عالی حرکت کرد. از دل این یافته‌ها، مجموعه‌ای از راهکارهای کاربردی قابل استخراج است که می‌تواند به ارتقای اعتبار، عدالت و کارآمدی سنجش آنلاین در آموزش عالی کمک کند. در سطح کلان، توسعهٔ زیرساخت‌های فناورانه و سرمایه‌گذاری در پلتفرم‌های بومی امن، همان‌گونه که در آلمان از طریق سامانه‌هایی چون ONYX و OPAL تجربه شده است (دومن و همکاران، ۲۰۲۱: ۹۲)، ضرورتی اجتناب‌ناپذیر به شمار می‌رود. همچنین هماهنگی نهادی و تدوین قوانین حمایتی برای تضمین صداقت علمی و ارتقای اعتماد عمومی به آزمون‌ها، مشابه دستورالعمل‌هایی که در برخی دانشگاه‌های آلمان در حوزهٔ خودسنجی‌ها و آزمون‌های الکترونیک تدوین شده (رایس و همکاران، ۲۰۰۹: ۳؛ گراسیگر و اشنبل، ۲۰۲۴: ۵) می‌تواند به تقویت روایی پیامدی کمک کند. در سطح نهادی و سازمانی، توجه به امنیت آزمون‌ها از طریق بهره‌گیری از فناوری‌های نوین مانند پروکتورینگ هوش مصنوعی و بازبینی چندلایه، زمینه‌ساز کاهش تقلب و افزایش اعتماد دانشجویان خواهد بود (لن‌هارد و همکاران، ۲۰۰۷: ۱۵۵). تجربهٔ آلمان در تشکیل شبکه‌های همکاری علمی میان استادان نیز نشان می‌دهد که ایجاد کارگروه‌های مشترک می‌تواند به بازطراحی آزمون‌ها و تبادل تجارب میان‌رشته‌ای یاری رساند (هایل و ایفنتهالر، ۲۰۲۳: ۱۶–۱۸) در کنار این، طراحی پروتکل‌های مشخص برای برگزاری آزمون در شرایط بحران، چه در ایران و چه در آلمان، تضمین‌کنندهٔ تداوم آموزشی در موقعیت‌هایی همچون همه‌گیری کرونا یا اختلالات اینترنتی است (T.N10–

[1] Shin
[2] Martin et al.
[3] Validity of Inferences
[4] Consequential basis
[5] Braun & Clarke

ایران؛ T.N4 – آلمان). در سطح آموزشی و اجرایی، توانمندسازی استادان برای طراحی آزمون‌های نوآورانه و دیجیتالی اهمیت ویژه دارد. شواهد نشان می‌دهد که تنها انتقال آزمون حضوری به بستر آنلاین کافی نیست، بلکه استادان نیازمند مهارت‌های تازه‌ای برای طراحی آزمون‌های پروژه‌محور، بازپاسخ و ترکیبی هستند (T.N8–ایران؛ T.N2–آلمان). از سوی دیگر، توانمندسازی دانشجویان از طریق ارتقای سواد دیجیتال و فرهنگ‌سازی پیرامون صداقت علمی، به کاهش نابرابری‌های آموزشی و افزایش شمول‌پذیری کمک می‌کند (مارتین و همکاران[1]، ۲۰۲۰: ۱۱۲–۱۱۴). افزون بر این، تنوع‌بخشی به الگوهای سنجش و استفاده از روش‌هایی چون ارزیابی مستمر، آزمون‌های پروژه‌محور و مدل‌های ترکیبی حضوری–آنلاین، همان‌گونه که در ادبیات بین‌المللی توصیه شده است (کریسپ و وارد، ۲۰۲۲؛ جوگین، ۲۰۱۰)، می‌تواند زمینه‌ساز یادگیری عمیق‌تر و بازخورد اثربخش‌تر باشد. بر این اساس، می‌توان نتیجه گرفت که حرکت به سوی الگویی پایدار از سنجش آنلاین تنها زمانی امکان‌پذیر است که سه سطح سیاست‌گذاری، نهادی و آموزشی به‌طور همزمان تقویت شوند. چنین رویکردی نه‌تنها به رفع چالش‌های فعلی کمک خواهد کرد، بلکه بستری برای تبدیل سنجش آنلاین از یک ضرورت اضطراری به یک ظرفیت پایدار و تحول‌آفرین در آموزش عالی فراهم می‌سازد.

تعارض منافع

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تعارض منافعی در خصوص اجرای پژوهش، تحلیل داده‌ها و تدوین نتایج گزارش‌شده وجود ندارد.

منافع مالی

نویسندگان اعلام می‌دارند که این پژوهش فاقد هرگونه حمایت مالی یا پشتیبانی سازمانی خاص بوده و هیچ منفعت مالی در قبال اجرای آن برای پژوهشگران یا مؤسسهٔ محل اجرا حاصل نشده است.

مشارکت نویسندگان

نویسندهٔ اول: نگارش پایان‌نامه، تهیه و آماده‌سازی ابزارها و مواد پژوهش، اجرای مداخله، گردآوری داده‌ها، انجام محاسبات، تجزیه و تحلیل آماری داده‌ها، تحلیل و تفسیر نتایج، تهیهٔ پیش‌نویس مقاله.

نویسندهٔ دوم: استاد راهنمای اول، طراحی پژوهش، نظارت بر مراحل انجام پژوهش، بررسی و کنترل نتایج، اصلاح، بازبینی و نهایی‌سازی مقاله.

نویسندهٔ سوم: استاد راهنمای دوم، مشارکت در طراحی پژوهش، نظارت بر پژوهش، مطالعه و بازبینی مقاله.

همچنین، نویسندگان به‌طور کلی در مراحل طراحی پژوهش، اجرای مداخله، گردآوری داده‌ها، تحلیل و تفسیر یافته‌ها و بازبینی و نهایی‌سازی مقاله مشارکت داشته‌اند.

اعلامیهٔ هوش مصنوعی مولد و فناوری‌های مبتنی بر هوش مصنوعی در فرایند نگارش

در طول آماده‌سازی این اثر، نویسندگان از ابزار هوش مصنوعی ChatGPT به منظور ویرایش زبانی و بهبود انسجام نگارشی متن استفاده کردند. پس از استفاده از این ابزار، نویسندگان تمامی بخش‌های علمی را خود به‌طور کامل بررسی، اصلاح و نهایی کرده و مسئولیت کامل محتوای مقاله را بر عهده می‌گیرند.

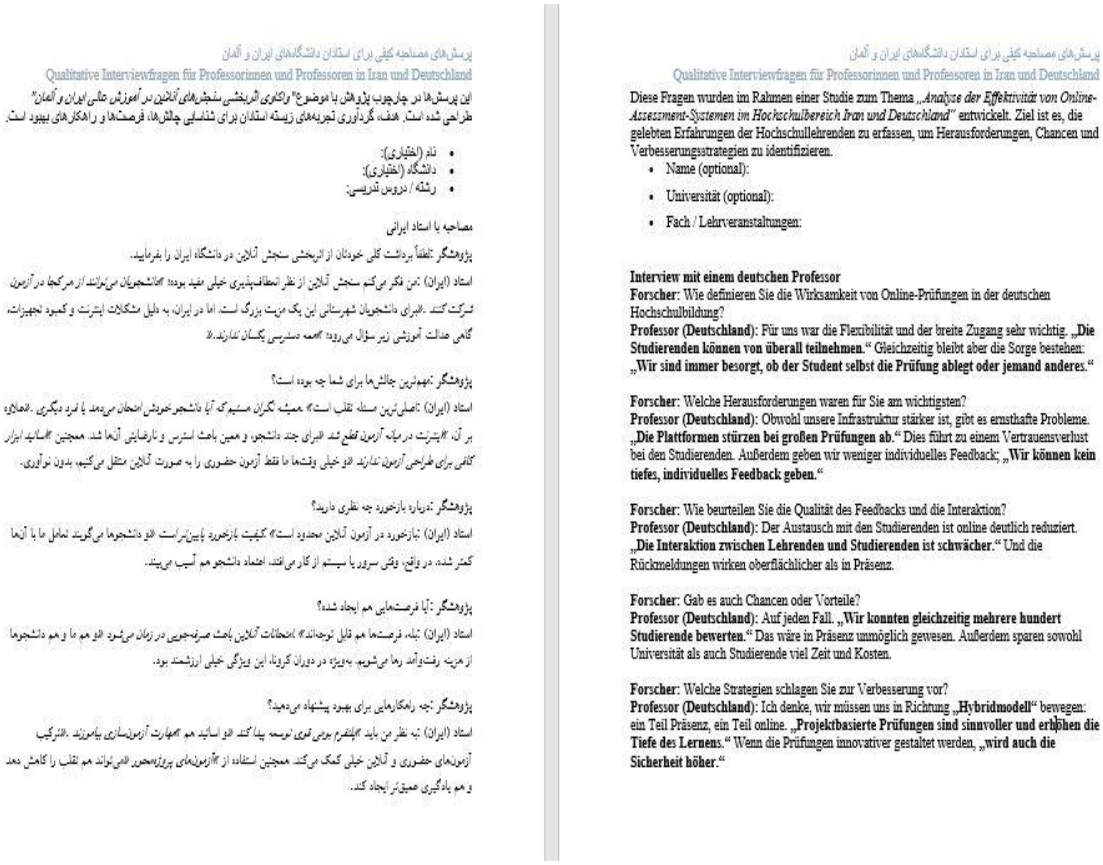
بیانیهٔ دسترسی به داده‌ها

داده‌های پژوهش حاضر در صورت درخواست معقول از نویسندهٔ مسئول قابل دسترسی است.

تشکر و قدردانی

[1] Martin et al.

نویسندگان بر خود لازم می‌دانند از تمامی مسئولان محترم مؤسسهٔ آموزشی «معراج اندیشه» که بستر مناسب برای اجرای مداخله را فراهم نمودند و همچنین از تمامی فراگیران نوجوان عزیز و خانواده‌های گرامی آنان که با همکاری صمیمانهٔ خود در مراحل مختلف این پژوهش مشارکت داشتند، صمیمانه قدردانی و تشکر نمایند.



پرسش‌های مصاحبه کیفی برای استادان دانشگاه‌های ایران و آلمان

Qualitative Interviewfragen für Professorinnen und Professoren in Iran und Deutschland

این پرسش‌ها در چارچوب پژوهش با موضوع «واکاوی اثربخشی سنجش‌های آنلاین در آموزش عالی ایران و آلمان» طراحی شده است. هدف، گردآوری تجربه‌های زیسته استادان برای شناسایی چالش‌ها، فرصت‌ها و راهکارهای بهبود است.

- نام (اختیاری):
- دانشگاه (اختیاری):
- رشته / دروس تدریسی:

مصاحبه با استاد ایرانی

پژوهشگر :لطفاً برداشت کلی خودتان از اثربخشی سنجش آنلاین در دانشگاه ایران را بفرمایید.

استاد (ایران) :من فکر می‌کنم سنجش آنلاین از نظر انعطاف‌پذیری خیلی مفید بوده؛ «دانشجویان می‌توانند از هرکجا در آزمون شرکت کنند .»اجرای دانشجویان شهرستانی این یک مزیت بزرگ است. اما در ایران، به دلیل مشکلات اینترنت و کمبود تجهیزات، گاهی عدالت آموزشی زیر سؤال می‌رود؛ «همه دسترسی یکسان ندارند.»

پژوهشگر :مهم‌ترین چالش‌ها برای شما چه بوده است؟

استاد (ایران) :اصلی‌ترین مسئله تقلب است؛» همیشه نگران هستم که آیا دانشجو خودش امتحان می‌دهد یا فرد دیگری .»اصلاً بر آن، «اینترنت در میانه آزمون قطع شد .»اجرای چند دانشجو، و همین باعث استرس و نارضایتی آنها شد. همچنین «اساتید ابزار کافی برای طراحی آزمون ندارند .»او خیلی وقت‌ها ما فقط آزمون حضوری را به صورت آنلاین منتقل می‌کنیم، بدون نوآوری.

پژوهشگر :درباره بازخورد چه نظری دارید؟

استاد (ایران) :بازخورد در آزمون آنلاین محدود است؛» کیفیت بازخورد پایین‌تر است .»او دانشجوها می‌گویند تعامل ما با آنها کمتر شده. در واقع، وقتی سرور یا سیستم از کار می‌افتد، اعتماد دانشجو هم آسیب می‌بیند.

پژوهشگر :آیا فرصت‌هایی هم ایجاد شده؟

استاد (ایران) :بله، فرصت‌ها هم قابل توجه‌اند؛» امتحانات آنلاین باعث صرفه‌جویی در زمان می‌شود .»او هم ما و هم دانشجوها از هزینه، رفت‌وآمد رها می‌شویم. به‌ویژه در دوران کرونا، این ویژگی خیلی ارزشمند بود.

پژوهشگر :چه راهکارهایی برای بهبود پیشنهاد می‌دهید؟

استاد (ایران) :به نظر من باید «پلتفرم بومی قوی توسعه پیدا کند .»او اساتید هم «مهارت آزمون‌سازی بیاموزند .»ترکیب آزمون‌های حضوری و آنلاین خیلی کمک می‌کند. همچنین استفاده از «آزمون‌های پروژه‌محور »می‌تواند هم تقلب را کاهش دهد و هم یادگیری عمیق‌تر ایجاد کند.

Diese Fragen wurden im Rahmen einer Studie zum Thema „*Analyse der Effektivität von Online-Assessment-Systemen im Hochschulbereich Iran und Deutschland*" entwickelt. Ziel ist es, die gelebten Erfahrungen der Hochschullehrenden zu erfassen, um Herausforderungen, Chancen und Verbesserungsstrategien zu identifizieren.

- Name (optional):
- Universität (optional):
- Fach / Lehrveranstaltungen:

Interview mit einem deutschen Professor

Forscher: Wie definieren Sie die Wirksamkeit von Online-Prüfungen in der deutschen Hochschulbildung?

Professor (Deutschland): Für uns war die Flexibilität und der breite Zugang sehr wichtig. **„Die Studierenden können von überall teilnehmen.“** Gleichzeitig bleibt aber die Sorge bestehen: **„Wir sind immer besorgt, ob der Student selbst die Prüfung ablegt oder jemand anderes.“**

Forscher: Welche Herausforderungen waren für Sie am wichtigsten?

Professor (Deutschland): Obwohl unsere Infrastruktur stärker ist, gibt es ernsthafte Probleme. **„Die Plattformen stürzen bei großen Prüfungen ab.“** Dies führt zu einem Vertrauensverlust bei den Studierenden. Außerdem geben wir weniger individuelles Feedback; **„Wir können kein tiefes, individuelles Feedback geben.“**

Forscher: Wie beurteilen Sie die Qualität des Feedbacks und die Interaktion?

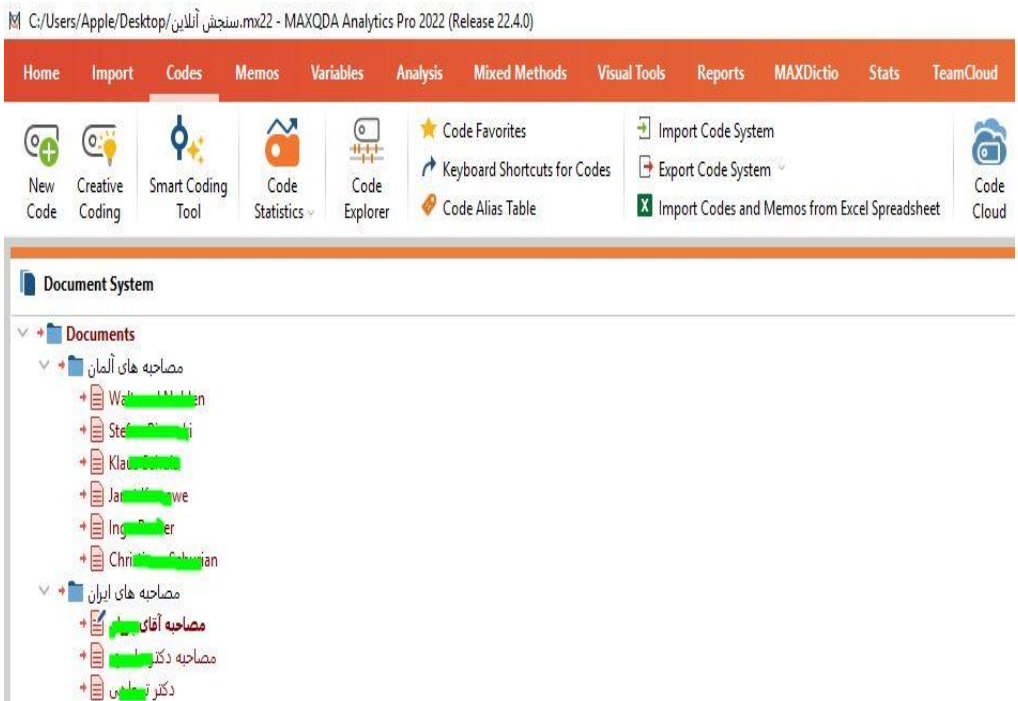
Professor (Deutschland): Der Austausch mit den Studierenden ist online deutlich reduziert. **„Die Interaktion zwischen Lehrenden und Studierenden ist schwächer.“** Und die Rückmeldungen wirken oberflächlicher als in Präsenz.

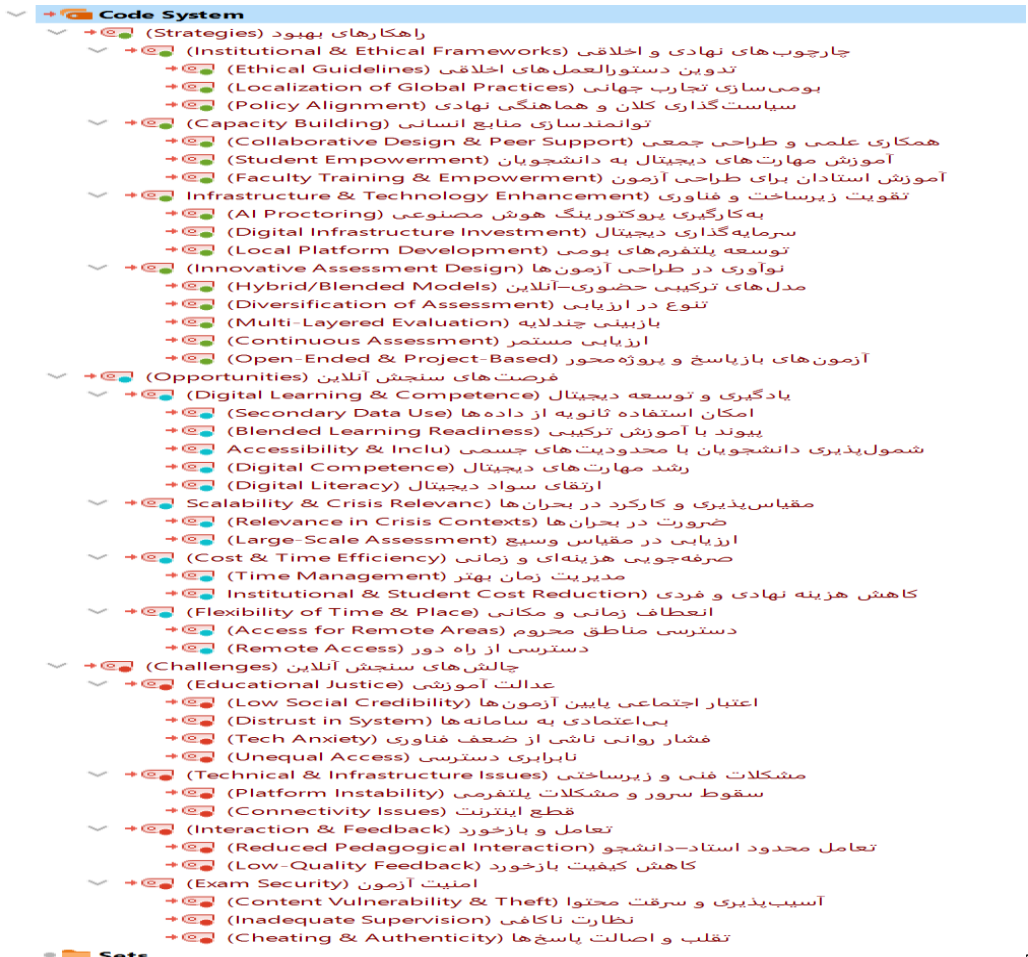
Forscher: Gab es auch Chancen oder Vorteile?

Professor (Deutschland): Auf jeden Fall. **„Wir konnten gleichzeitig mehrere hundert Studierende bewerten.“** Das wäre in Präsenz unmöglich gewesen. Außerdem sparen sowohl Universität als auch Studierende viel Zeit und Kosten.

Forscher: Welche Strategien schlagen Sie zur Verbesserung vor?

Professor (Deutschland): Ich denke, wir müssen uns in Richtung **„Hybridmodell“** bewegen: ein Teil Präsenz, ein Teil online. **„Projektbasierte Prüfungen sind sinnvoller und erhöhen die Tiefe des Lernens.“** Wenn die Prüfungen innovativer gestaltet werden, **„wird auch die Sicherheit höher.“**





Referen

Azimi, F., [illegible] arative study on the role of plagiarism detection and pre-tests. *Journal of Academic Ethics*, *14*(2), 121–134. https://doi.org/10.1007/s10805-016-9251-4

Azizi, M. (2022). Online assessment and distributive justice in Iranian higher education: Challenges and opportunities. *Journal of Educational Technology Development*, *18*(1), 39–54. https://doi.org/10.22054/edtech.2022.34567

Bauer, J., Figl, K., & Nistor, N. (2009). Peer assessment in higher education: Benefits and challenges. *Assessment & Evaluation in Higher Education, 34*(1), 75–89. https://doi.org/10.1080/02602930801895736

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). Open University Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education, 159,* 104024. https://doi.org/10.1016/j.compedu.2020.104024

Craig, E., & Kay, R. (2025). Enhancing student engagement in online assessments through peer review. *Journal of Online Learning Research, 11*(1), 45–55.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Crisp, G., & Ward, C. (2022). Re-imagining assessment in higher education: Digital, authentic, and inclusive practices. *Assessment & Evaluation in Higher Education, 47*(4), 565–580. https://doi.org/10.1080/02602938.2021.1887812

Dashtestani, R. (2020). Online language learning in Iran: Students' and instructors' perceptions. *Education and Information Technologies, 25*(1), 81–99. https://doi.org/10.1007/s10639-019-09945-0

Dashtestani, S. R. & Karami, H. (2019). An analysis of Iranian online EFL teachers' technological, pedagogical, and evaluation skills. *Journal of Foreign Language Research, 9*(3), 815–830. https://doi.org/10.22059/jflr.2019.261193.528

Davis, V. W., Smith, L. T., & Chan, M. (2023). Privacy, equity, and validity in online testing: A framework for higher education. *Assessment in Education: Principles, Policy & Practice, 30*(1), 118–134. https://doi.org/10.1080/0969594X.2022.2094567

Dohmen, D., Kröner, S., & Hartig, J. (2021). Digital examinations in German higher education: Experiences with ONYX, OPAL, and MAXIMA. *Zeitschrift für Hochschulentwicklung, 16*(2), 91–107. https://doi.org/10.3217/zfhe-16-02/05

Gamage, D., Ayres, J., & Behrend, M. (2022). A systematic review on the use of Moodle for online assessment in higher education. *Education and Information Technologies, 27*(1), 13–29. https://doi.org/10.1007/s10639-021-10683-3

Grassinger, R., & Schnebel, S. (2024). Online self-assessments in German higher education: Fostering reflection and learning outcomes. *Journal of Educational Measurement, 61*(1), 1–19. https://doi.org/10.1111/jedm.12345

Heil, C., & Ifenthaler, D. (2023). Quality and acceptance of online assessment in German universities: Policies and practices. *British Journal of Educational Technology, 54*(1), 13–29. https://doi.org/10.1111/bjet.13224

Jeremias, F., & Sellmer, J. (2019). Implementing e-assessments in higher education: Lessons learned from German universities. *Zeitschrift für Hochschulentwicklung*, *14*(3), 97–110.

Joughin, G. (2010). The hidden curriculum revisited: A critical review of research into the influence of summative assessment on learning. *Assessment & Evaluation in Higher Education, 35*(3), 335–345. https://doi.org/10.1080/02602930903221493

Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Köller, O., Watermann, R., & Wigfield, A. (2020). Adaptive testing and diagnostic assessment in German higher education. *Studies in Educational Evaluation, 66,* 47–58. https://doi.org/10.1016/j.stueduc.2020.100876

Lehmann, T., Hahn, M., & Halle, T. (2015). Peer assessment and its impact on higher education learning outcomes. *Assessment & Evaluation in Higher Education, 40*(1), 11–22. https://doi.org/10.1080/02602938.2014.976111

Lenhard, W., Baier, H., Hoffmann, J., & Schneider, W. (2007). Computer-based latent semantic analysis in online assessments. *Educational Psychology, 27*(2), 155–176. https://doi.org/10.1080/01443410601066742

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Loureiro, A., & Gomes, C. (2022). Online learning and peer interaction: The role of anonymity in feedback quality. *Computers in Human Behavior, 136,* 233–241. https://doi.org/10.1016/j.chb.2022.107382

Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. Computers & Education, 159, 104009. https://doi.org/10.1016/j.compedu.2020.104009

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Mounesi Qaramalki, A., Āhangarī, S., & Haddīdi Tamjid, N. (2024). Online foreign language testing: Validity and cognitive load considerations. *Iranian Journal of Applied Linguistics*, *27*(2), 115–134. https://doi.org/10.22054/ijals.2024.34567

Muzaffar, M., Ali, S., & Raza, T. (2020). Security challenges in online assessments: A systematic review. *International Journal of Emerging Technologies in Learning, 15*(10), 6–12. https://doi.org/10.3991/ijet.v15i10.13795

Mutimukwe, C., Lubega, J., & Muyinda, P. (2020). Online assessment in higher education: Equity and privacy concerns. *International Journal of Education and Development using ICT, 16*(2), 63–74.

Petri, K., & Kersting, M. (2024). Online self-assessments as guidance tools in German higher education. *Frontiers in Education, 9,* 1234567. https://doi.org/10.3389/feduc.2024.1234567

Rahimi, M., Aghajani, M., & Khodabakhsh, S. (2019). Technical challenges of online exams in Iranian universities. *Educational Technology Research, 11*(2), 53–61. https://doi.org/10.22054/edtech.2019.27890

Reiss, K., Sälzer, C., & Schiepe-Tiska, A. (2009). Online self-assessment in German higher education: Potentials and limitations. *Zeitschrift für Pädagogik, 55*(1), 1–15.

Shin, H. (2022). Redefining feedback in online higher education: From individualization to dialogic engagement. *Assessment & Evaluation in Higher Education, 47*(7), 1095–1110. https://doi.org/10.1080/02602938.2021.1938123

Strickroth, S., Hesse, F. W., & Fischer, F. (2011). Adaptive online testing in higher education: Potentials and challenges. *Computers & Education, 57*(1), 62–74. https://doi.org/10.1016/j.compedu.2010.12.004

Tahir, R., Yousaf, Z., & Khan, H. (2025). Online exams and the digital divide: A case study from South Asia. *International Review of Research in Open and Distributed Learning, 26*(1), 61–70. https://doi.org/10.19173/irrodl.v26i1.7654

Teshome, A. (2025). Online assessment in African universities: Opportunities and risks. *African Educational Review, 17*(2), 87–99. https://doi.org/10.1080/18146627.2025.123456