

پدیدارشناسی، ساخت و اعتباربخشی انگاره‌روایی پسینی سنجه زبان در آزمون ورودی دانشگاه: کاربردهایی برای آزمون‌سازی در تحصیلات تکمیلی

محمود خسروانی^۱، مرتضی رستمیان^۲، حمید اشرف^۳، حسین خدابخش‌زاده^۴

چکیده

رویکردهای استاندارد به ساخت آزمون علی‌رغم مزایای گوناگون، به دلیل تفکیک آموزش از ارزشیابی در ساحت گوناگون، باعث ایجاد تأثیرات تعقیبی غیرسازنده فراوانی شده‌اند. در این میان، اهداف برنامه ملی درسی ایران که یک تلاش راهبردی برای ایجاد تغییرات مثبت در آموزش کشور بود، به دلیل همین اثرات تعقیبی، محقق نشده است. به دلیل خلأ تحقیقاتی در زمینه پدیدارشناسی این موضوع و نقایص روش‌شناختی و معرفت‌شناختی در پرداخت به آن، تحقیق حاضر تلاشی است تا با اتخاذ یک طرح چندمرحله‌ای - چندروشی، ضمن کشف مغایرت‌های سنجه زبان انگلیسی آزمون ورودی دانشگاه با برنامه ملی درسی در کتاب‌های ویزن، یک انگاره معتبر از مؤلفه‌ها و متغیرهای روایی پسینی این آزمون ارائه دهد و سپس با تحلیل ساختاری در ماتریس اثرات متقاطع، راهبردهایی را در اختیار سیاست‌گذاران آموزشی قرار دهد. بدین منظور ده نسخه در زمینه آموزش زبان انگلیسی و ۵۱۰ نفر از شرکت‌کنندگان در آزمون ورودی دانشگاه، به ترتیب، با نمونه‌گیری هدف‌دار و در دسترس برای مراحل تحقیق انتخاب شدند. در گام اول، پدیدارشناسی با روش مصاحبه نیمه‌ساختاریافته از نخبگان حاکی از مغایرت اهداف آزمون با برنامه ملی

۱. دانش‌آموخته دکترای تخصصی، رشته زبان انگلیسی، گروه زبان انگلیسی، واحد تربت حیدریه، دانشگاه آزاد اسلامی، تربت حیدریه، ایران.

۲. استادیار، رشته زبان انگلیسی، گروه زبان انگلیسی، واحد تربت حیدریه، دانشگاه آزاد اسلامی، تربت حیدریه، ایران (m.rostamian.edu@gmail.com) نویسنده مسئول

۳. استادیار، گروه زبان انگلیسی، دانشکده پزشکی، دانشگاه علوم پزشکی گناباد، ایران.

۴. استادیار، گروه زبان انگلیسی، واحد تربت حیدریه، دانشگاه آزاد اسلامی، تربت حیدریه، ایران.

(دریافت: ۱۴۰۱/۰۹/۲۲ - پذیرش: ۱۴۰۱/۱۱/۱۲)

درسی در پنج بُعد اساسی است. مدل‌سازی معادله ساختاری نیز نشان داد که ۲۶ متغیر در هفت مؤلفه (فاکتور) در انگاره روایی پسینی این آزمون دخالت دارند. نتایج تحلیل ساختاری نیز متغیرهای راهبردی، حیاتی، تأثیرگذار، دووجهی و تنظیم‌گر در این نظام تأثیرات را مشخص و اولویت‌بندی کرد و متناسب با آن کاربردهایی برای آزمون‌سازی در تحصیلات تکمیلی دانشگاه پیشنهاد شد. این نتایج دارای کاربردهای نظری، عملی، روش‌شناختی و راهبردی است.

واژه‌های کلیدی: اثر تعقیبی، انگاره‌سازی، برنامه ملی درسی، پدیدارشناسی، تحصیلات تکمیلی، زبان انگلیسی، ویژن.

Phenomenology , and validating a Consequential Validity Model for English Module of University Entrance Exam: Pedagogical Implications for Post-Graduate Testing in University

Mahmood Khosravani¹

Morteza Rostamian

Hamid Ashraf²

Hossein Khodabakhshzadeh³

Abstract

Standard approaches to test designing, despite their benefits, have brought about destructive effects due to the separation of education from assessment of different aspects. Iranian national curriculum was a strategic attempt to make positive changes but failed in its mission due to the test's consequential effects. Considering the gap regarding the phenomenography of this issue and methodological and epistemological deficiencies in research on test consequences, the present research is an attempt to employ a multi-phase/multi-method design and provide a consequential validity model for the English module of university entrance exam based on the Vision book series and then a structural analysis of the variables of the model to feed educational and pedagogical practitioners for implementing appropriate strategies. To this aim, ten elite participants in ELT and 510 recent examinees of the university entrance exam were selected through purposive and convenience sampling ,respectively. In the first phase, the phenomenography was conducted through semi-structured interviews with elites and revealed that the vision curriculum contradicts the national curriculum in five basic dimensions. Moreover, the results of structural equation modeling revealed that 26 variables subsumed under seven factors are involved in the developed consequential validity model. Finally, the findings of structural interpretation also determined and prioritized the strategic, critical, bi-dimensional, regulatory, and influential nature of the variables. These findings have theoretical, methodological, practical, and strategic implications for different practitioners

Keywords: consequential Effect, modeling, national curriculum, phenomenography, post-Graduate, english, vision

-
1. PhD Student, English Department, Torbat-e Heydariyeh Branch, Islamic Azad University, Torbat-e Heydariyeh, Iran
 2. Assistant Professor, English Department, Torbat-e Heydariyeh Branch, Islamic Azad University, Torbat-e Heydariyeh, Iran. Corresponding Author. (m.rostamian.edu@gmail.com)
 3. Assistant Professor, English Department, Torbat-e Heydariyeh Branch, Islamic Azad University, Torbat-e Heydariyeh, Iran

مقدمه

اغلب مطالعاتی که به بررسی پدیده‌های مرتبط با یادگیری پرداخته‌اند، نگاهی موزاییکی به حیطه‌های یادگیری اتخاذ کرده‌اند و در بسیاری از موارد، بخش‌های مختلف یادگیری، مانند آموزش و تدریس، ارزشیابی و آزمون و برنامه‌ریزی و سیاست‌گذاری آموزشی را به صورت مجزا بررسی کرده و از اتخاذ نگاهی کل‌نگر عاجز مانده‌اند. براساس شیزل^۱ (۲۰۱۸)، چنین رویکرد تحقیقاتی به دریافتی سطحی از پدیده پیچیده و چندوجهی یادگیری منجر شده است، زیرا برهم‌کنش تأثیرات وجود روابط علی بین بخش‌های آموزشی، رویه‌ای و مدیریتی در یادگیری، اجتناب‌ناپذیر است. در این میان، رویکردهای سنتی به آموزش، افزون‌بر نقایص فراوان معرفت‌شناختی (مانند اتخاذ نگاه محصول‌محور به یادگیری) و روش‌شناختی (مانند اعتقاد به خطی بودن و انتقال‌پذیر بودن یادگیری)، قائل به تفکیک دو حوزه مهم آموزش و ارزشیابی بوده‌اند که حتی در اصلاحات استاندارد روش‌های کلاسیک نیز به چشم می‌خورد (فرخی‌پور و همکاران، ۱۳۹۹). محققان زیادی معتقدند یکی از حیطه‌هایی که رابطه بین آموزش و ارزشیابی را مفروض دانسته و به بررسی ابعاد گوناگون آن پرداخته است، روایی پسینی^۲ آزمون است (پوفام^۳، ۱۹۸۷؛ شپارد^۴، ۱۹۹۳؛ چنگک^۵ و همکاران، ۲۰۰۷؛ ژی^۶، سان^۷، ۲۰۱۶). بررسی پیشینه ادبی مفهوم «روایی»^۸ آزمون نشان می‌دهد که این مفهوم از حیطه «روایی ساخت»^۹ به حیطه‌های گسترده‌تری توسعه یافته است و متغیرهایی مانند تأثیرات ثانویه آموزشی، نگرشی، اقتصادی و غیره، کاربردهای آزمون، عدالت و انصاف در آزمون، چگونگی تفسیر نمرات و چگونگی کاربری تفسیر نمرات را در خود جای داده است که به صورت خلاصه ذیل روایی پسینی آزمون جای گرفته‌اند (مزیک^{۱۰}، ۱۹۹۶؛ کین^{۱۱}، ۲۰۰۶). بدیهی است که با اتخاذ نگاه جزءنگر، امکان بررسی شواهد مرتبط با اثبات روایی پسینی یک آزمون وجود

1. Schiesel
2. consequential validity
3. Popham
4. Shepard
5. Cheng
6. Xie
7. Sun
8. validity
9. construct validity
10. Messick
11. Kane

ندارد و شواهد مبتنی بر روایی ساخت، روایی محتوایی و روایی شکلی آزمون‌ها، معیار مناسبی برای بررسی کارایی آن‌ها نیست (بکمن و پالمر^۱، ۱۹۹۶). در همین ارتباط، رویکردهای کلاسیک به آزمون‌سازی، علیرغم تغییرات بنیادین در شیوه‌ها و معرفی آزمون‌های استاندارد برای حل بخشی از مشکلات آموزشی زمان خود، از اوایل دهه هشتاد قرن بیستم، مورد انتقاد جدی کاربران قرار گرفتند، زیرا بالا بودن قدرت و تأثیرگذاری این نوع آزمون‌ها به ایجاد عواقب خواسته و ناخواسته آموزشی و غیر آموزشی فراوانی منجر می‌شد که از آن جمله می‌توان به تحدید برنامه‌های درسی، آموزش برای آزمون به جای آموزش برای یادگیری، آسیب‌های اجتماعی و روان‌شناختی اشاره کرد (چن^۲ و همکاران ۲۰۲۰). به گفته گرین^۳ (۲۰۲۰) برای بررسی این عواقب، اصطلاحات نوینی در حوزه آموزش و ارزشیابی ظهور یافت که از آن جمله می‌توان به واشبک^۴ (تأثیر معکوس آزمون بر آموزش) اشاره کرد که ذیل همین روایی پسینی قابل بحث است (شپل^۵، ۲۰۲۰). کیم^۶ (۲۰۱۷) از محققان پیش‌رو در زمینه بررسی روایی پسینی آزمون‌هاست که برای رسیدن به یک قالب کاری جامع حول روایی پسینی، نظریات زیادی در حوزه آزمون‌سازی را ترکیب کرد و برای نخستین بار مفاهیمی مانند معناداری آزمون^۷، جهت‌دهی آزمون، شفافیت آزمون و غیره را از مؤلفه‌های روایی پسینی عنوان کرد که برای همه آزمون‌های تأثیرگذار ملزم به تحقق آن‌ها هستند.

در بافت آموزشی ایران، برنامه درسی ملی (۱۳۹۰) تلاشی بود تا برخی از نارسایی‌های آموزشی کشور را به ویژه در حوزه عدم تطابق اهداف آموزشی با اهداف آزمون‌سازی، مرتفع سازد. اگرچه این برنامه درسی در حوزه طراحی آموزشی، برخی از نارسایی‌های پیشین را پوشش داد، اما در زمینه همگام‌سازی اهداف آموزش با آزمون، نه تنها موجب همگرایی شد، بلکه بر شکاف ایجادشده افزود تا روایی پسینی آزمون‌های استاندارد ورود به دانشگاه، حداقل از لحاظ نظری، به کم‌ترین مقدار ممکن تنزل پیدا کند (فتحی^۸، ۲۰۲۱). شواهد آموزشی و تجربی نشان داده است که در دهه گذشته نیز، سنجه

1. Bachman & Palmer
2. Chen
3. Green
4. Washback
5. Chapelle
6. Kim
7. test meaningfulness
8. Fathi

زبان انگلیسی آزمون‌های ورودی دانشگاه، به‌عنوان بخشی از یک آزمون تأثیرگذار، عواقب گوناگون آموزشی و فرابافتی داشته است که نه‌تنها مورد غفلت سیاست‌گذاران آموزشی قرار گرفته است بلکه حیطه شمول این عواقب از فضای آموزشی فراتر رفته است و ساحات اجتماعی، خانواده، اقتصاد و حتی سیاست را در بر گرفته است (خسروانی^۱ و همکاران، ۲۰۲۲). براساس فتحی^۲ (۲۰۲۱، ص ۱۱۲)، برنامه ملی درسی ایران در بحث آموزش زبان با توسعه و جایگزینی کتاب‌های سری «ویژن»، اقدام مؤثری برای پوشش نارسایی‌های آموزشی در زمینه بومی‌سازی محتوا، تعدیل جنبه‌های فرهنگی، روش تدریس، تأکید بر مهارت‌های شفاهی آموزش و تمرینات درسی معنادار و خلاقیت آفرین را رقم زد، به‌طوری که براساس جعفری‌گهر و جمالوندی (۱۳۹۷) با ویژگی‌های کتاب مطلوب آموزش زبان انگلیسی تطابق داشت، اما واگرایی اهداف آموزشی با اهداف آزمون‌سازی در کنکور وردی دانشگاه را به حداکثر رساند و باعث شد بخش زیادی از اهداف آموزشی این برنامه محقق نشود، زیرا محتوای سنجه آزمون زبان انگلیسی با محتوای کتاب‌های ویژه از حیث شکل سؤالات و ارجحیت مهارت‌های یادگیری زبان دوم، تفاوت رادیکال داشت. علی‌رغم این مشکلات بنیادین در روایی پسینی سنجه زبان انگلیسی آزمون ورودی دانشگاه، تاکنون تحقیق جامعی در مورد ابعاد گوناگون آن صورت نگرفته است. همچنین در برخی تحقیقات که به بررسی تأثیر معکوس آزمون بر برنامه درسی ایران انجام شده، از الگوهای نظری فرابافتی استفاده شده است که حداقل در ابعاد فراآموزشی، با زیست‌بوم آموزشی ایران تطابق ندارند. همچنین، براساس علوی و مسجدلو^۳ (۲۰۱۷)، برخی نتایج حاصل شده از تحقیقات بخشی در این زمینه، فاقد اعتبار بخشی لازم و در اغلب مواقع از کمینه بازنمایی روایی^۴ برخوردار هستند. بر همین اساس، تحقیق حاضر تلاشی است تا با اتخاذ طرح تحقیقی چندمرحله‌ای-چندروشی^۵ تصویر جامعی از این پدیده را ارائه کند و ضمن کشف دلایل واگرایی کارکردی برنامه درسی ملی ایران با اهداف سنجه زبان انگلیسی آزمون ورودی دانشگاه از طریق پدیدارشناسی، با کاربرست مدل‌سازی معادله ساختاری، ابعاد و مؤلفه‌های عواقب پسینی این آزمون را بررسی و تحلیل

1. Khosravani
2. Fathi
3. Alavai & Masjedlu
4. construct underrepresentation
5. multi-phase multi-method design

کند و در نهایت با استخدام تحلیل ساختاری در ماتریس اثرات متقاطع، راهبردهایی را در این زمینه ارائه دهد. بر همین اساس، تحقیق حاصل معطوف به پاسخ‌گویی به سؤالات زیر خواهد بود:

۱. نقاط واگرایی اهداف سنج‌ه زبان انگلیسی آزمون ورودی دانشگاه و برنامه ملی درسی ایران کدام‌اند؟
۲. مؤلفه‌ها و متغیرهای عواقب پسینی سنج‌ه زبان انگلیسی آزمون ورودی دانشگاه کدام‌اند؟
۳. انگاره معتبر روایی پسینی این سنج‌ه کدام است و شاخص‌های برآزش آن کدام‌اند؟
۴. ماهیت و رفتار هر کدام از متغیرها در نظام انگاره حاصل شده چگونه است و راهبرد آموزشی متناسب با هر متغیر کدام است؟

مبانی نظری و پیشینه پژوهش

چارچوب نظری تحقیق

بر اساس مک‌نامارا^۱ (۲۰۰۶)، تأثیرگذارترین نوآوری نظری در زمینه روایی پسینی آزمون را مزیک^۲ (۱۹۸۹) ارائه کرده است. یکی از رویکردهای کاربردی برای بررسی مفهوم روایی پسینی که مستند به براهین مطرح شده توسط مزیک است و بیش‌ترین قرابت نظری را با این رویکرد دارد، «برهان کاربردی آزمون»^۳ بک من^۴ (۲۰۰۵) است. بنابراین در تحقیق حاضر، برای بررسی ابعاد گوناگون روایی پسینی سنج‌ه زبان انگلیسی آزمون ورودی دانشگاه‌ها از این چارچوب نظری استفاده شده است. این موضع نظری که بعدتر توسط بک‌من و پالمر (۲۰۱۰) به بلوغ رسید، روایی آزمون را به‌عنوان مفهومی یکپارچه مطرح می‌کند که مبتنی بر تفاسیر و استنباط‌های حاصل شده از نمرات آزمون و کاربرد این تفاسیر برای اهداف آموزشی یا انتخابی (استخدامی) است. بر این اساس، تصمیم‌های اتخاذ شده بر مبنای تفاسیر حاصل شده از نمرات آزمون، دارای عواقب ناخواسته و ناشناخته بر روی افراد، سازمان‌ها و برنامه‌های

1. McNamara
2. Messick
3. Assessment Use Argument
4. Bachman

آموزشی است. براساس بک‌من و پالم (۲۰۱۰، ص ۲۲)، این چارچوب نظری دارای گزاره‌هایی مرتبط با نتایج آزمون است که مواردی مانند تفاسیر نمرات، تصمیمات ناشی از آن و عواقب منتج از آن را به شکل یکپارچه و در تعامل با دیگر گزاره‌ها بررسی می‌کند و براساس تحلیل جامع این گزاره‌ها، شواهدی را برای اثبات روایی ارائه می‌دهد. این چارچوب نظری از چند حیث برای تحقیق حاضر اهمیت دارد؛ این چارچوب ساختاری صریح و با وضوح نظری از مفهوم روایی ارائه می‌دهد که موجب هدایت ادراکی محقق برای بررسی این پدیده پیچیده می‌شود. همچنین، برخلاف بسیاری از چارچوب‌های نظری مطرح شده در این حوزه که به بعد آموزشی روایی آزمون‌ها پرداخته‌اند، چارچوب نظری حاضر امکان بررسی روایی پسینی در قالب تأثیرات آموزشی و فراآموزشی را فراهم می‌کند.

مرور پیشینه تحقیق و ادبیات مرتبط

پیشینه تحقیقات انجام شده در زمینه روایی پسینی آزمون‌ها از دو منظر تحقیقات نظری و تحقیقات عملی (تجربی) قابل بررسی است. از منظر اول، مرور ادبیات تحقیق نشان می‌دهد که بحث روایی آزمون‌ها حتی در روش‌های کلاسیک تدریس و آزمون‌سازی، مورد توجه بوده است و به مرور زمان در روش‌های نوین آزمون‌سازی به بلوغ نظری رسیده است. کوریتون^۱ (۱۹۵۱) از اولین نظریه پردازان در این حوزه، روایی را صرفاً به ارتباط آزمون با اهداف مورد نظر خود تعبیر می‌کند. بعد از آن کرونباخ^۲ (۱۹۸۰) روایی را فراتر از صحت و تطابق کارکرد - هدف آزمون می‌داند و حوزه‌هایی مانند تأثیر آزمون بر زندگی فراگیران را نیز مورد توجه قرار می‌دهد و تصریح می‌کند که این تأثیرات، نیاز به اعتبارسنجی نظری و آماری دارد. در همین سال، مزیک^۳ (۱۹۸۹، ص ۱۲) موضع نظری تأثیرگذاری را در مورد روایی آزمون‌ها مطرح می‌کند و روایی شکلی، محتوایی و ساخت را شرط لازم ولی ناکافی برای اعتبار آزمون عنوان می‌کند و بحث معناداری همه‌جانبه تفاسیر و استنباط‌های اتخاذ شده از نمرات آزمون را شرط اصلی برای تحقق روایی آزمون می‌داند. این موضع نظری، با کمک مکاتب فلسفی، معرفت‌شناختی و هرمنوتیکی و پشتیبانی‌های آماری، بعدها توسط کرونباخ (۱۹۸۸)، موس^۴ (۱۹۹۸)، بک من (۲۰۰۵)، مک‌نامارا^۴ (۲۰۰۶) و در نهایت بک‌من و

1. Cureton
2. Cronbach
3. Moss
4. McNamara

پالمر (۲۰۱۰) توسعه یافت که در چارچوب نظری مقاله حاضر، به صورت خلاصه ارائه شده است. این در شرایطی است که از منظر عملی نیز، تحقیقات زیادی پیرامون این تغییرات نظری انجام شده است. ورنون^۱ (۱۹۵۶) از اولین تحقیقات تجربی بود که بررسی تأثیر آزمون بر آموزش را بررسی کرد و دریافت که معلمان بررسی شده در این تحقیق، محتوای آموزشی را به دلیل تمرکز بر محتوای آزمون، در کلاس‌ها حذف می‌کنند. دیویس^۲ (۱۹۶۸) نیز در تحقیقات خود تأثیر آزمون را موجب تحدید سیلابس درسی در لایه‌های طراحی و آموزش عنوان کرده است. سوین^۳ (۱۹۸۵)، هوگس^۴ (۱۹۸۹) و بایلی^۵ (۱۹۹۹) نیز در تحقیقاتی که در بافت‌های آموزش زبان خارجی در اروپا و آسیا انجام دادند، به تأثیر مخرب آزمون‌های تأثیرگذار بر ابعاد گوناگون آموزش، مریان و فرایندهای مدیریتی، اشاره کرده‌اند. در زمینه بسط نظری مفهوم روایی پسینی آزمون، تحقیق برجسته‌ای در بافت آموزش زبان انگلیسی ایران انجام نشده است. اما مرور پیشینه تحقیق نشان می‌دهد که بحث روایی پسینی آزمون در دهه گذشته نیز مورد توجه محققان داخلی و خارجی فراوانی قرار گرفته است که به فراخور تحقیق حاضر، به برخی از مهم‌ترین و مرتبط‌ترین آن‌ها، اشاره شده است. به دلیل کم بودن مطالعات مرتبط انجام شده در کشور، ادبیات مرتبط تحقیق به صورت ترکیبی با مطالعات خارجی و صرفاً براساس توالی تاریخی، بررسی شده است.

در یک مطالعه شبه تجربی در بافت آموزش زبان ایران، احمدی صفا و گودرزی^۶ (۲۰۱۴) به بررسی تأثیر معکوس ارزشیابی وظیفه‌محور^۷ بر آموزش زبان پرداختند. در این مطالعه که بر روی ۱۱۰ فراگیر سطح متوسطه انجام شد، نشان داد که رویکردهای وظیفه‌محور به ارزشیابی برخلاف رویکردهای سنتی، در تقارن بیشتری با اهداف آموزش قرار دارند و موجب تقویت آموزش می‌شوند. از نقاط ضعف این تحقیق این بود که رویکردهای مذکور در حیطه آزمون‌های پیشرفت^۸ قرار می‌گیرند و مانند آزمون‌های ورودی دانشگاه، در زمره آزمون‌های تأثیرگذار قرار نمی‌گیرند. چو^۹ (۲۰۱۵) به بررسی

-
1. Vernon
 2. Davies
 3. Swain
 4. Hughes
 5. Bailey
 6. Ahmadi Safa & Goodarzi
 7. task-based assessment
 8. achievement tests
 9. Cho

بخش شنیداری آزمون تأثیرگذار ورودی دانشگاه‌های تایوان را بر انتظارات فراگیران و ابعاد آموزش زبان در دبیرستان‌های این کشور بررسی کرد. نتایج حاصل از این تحقیق نشان داد که این آزمون به دلیل تغییر در ادراک فراگیران از آموزش، باعث ایجاد تأثیرات منفی بر آموزش شده است. تأثیرات آزمون بسندگی فارغ‌التحصیلی دانشگاه‌های همین کشور نیز در تحقیقی دیگر توسط پن و روور^۱ (۲۰۱۶) بررسی شد و نتایج آن نشان داد که به دلیل تأثیر تعقیبی زیاد این آزمون بر فراگیران، عواقب منفی اجتماعی زیادی را برای آن‌ها ایجاد کرده است. سان^۲ (۲۰۱۶) از مدل‌سازی معادله ساختاری برای تدوین و اعتباربخشی انگاره تأثیرات منفی یک آزمون زبان تأثیرگذار در چین استفاده کرد. انگاره حاصل شده نشان داد که در میان تأثیرات منفی آزمون، تأثیرات آموزشی، بافتی و اجتماعی به چشم می‌خورد. دی‌جنارو^۳ (۲۰۱۷) در تحقیقی مشابه در بافت آموزش زبان کره جنوبی دریافت که آزمون تأثیرگذار جیمیت^۴ در این کشور متناسب با سطح سواد، سن و طبقه اجتماعی افراد، تأثیرات تعقیبی گوناگونی را بر آن‌ها اعمال می‌کند. در تحقیقی متفاوت در بافت ایران، رضوی پور^۵ و همکاران (۲۰۱۸) تأثیر مشخصات سنجه ادبیات انگلیسی آزمون کارشناسی ارشد را بر القاء تأثیر معکوس منفی بر آموزش بررسی کردند. نتایج تحقیق نشان داد که این سنجه، به دلیل برخی ویژگی‌های ماهوی، بر نگرش و یادگیری فراگیران ادبیات انگلیسی، تأثیر می‌گذارد. وو و ژانگ^۶ (۲۰۱۹) در یک نوآوری تحقیقاتی، به بررسی کنش‌گری فرهنگی-اجتماعی آزمون تأثیرگذار تائیک^۷ در کره جنوبی پرداختند و دریافتند که آزمون مذکور علاوه بر نقض برخی از شواهد روایی آزمون‌های استاندارد، دارای برخی اثرات تعقیبی مخرب است که نیاز به بررسی بیشتر دارد. در تحقیقی کیفی که در دبیرستان‌های سوئد انجام گرفت، لارسونا و اولین-شلر^۸ (۲۰۲۰) دریافت که تأثیر منفی آزمون زبان در این کشور بر روی برخی مهارت‌های زبان مانند مهارت نوشتار و گفتار بیشتر از سایر مهارت‌ها است. و درنهایت، در تحقیقی که توسط جمالی‌فر^۹ و همکاران

1. Pan & Roever
2. Sun
3. Di-Genaro
4. GMATE
5. Razavipour
6. Wu & Zhang
7. TOEIC
8. Larssona & Olin-Scheller
9. Jamalifar

(۲۰۲۱) در بافت تحصیلات تکمیلی ایران انجام گرفت، به کمک توزیع پرسش‌نامه، تأثیرات معکوس آزمون توانش زبان دانشگاه آزاد اسلامی موسوم به ای‌پی‌تی^۱ به کمک تحلیل مضمون یافته‌ها مورد بررسی قرار گرفت. نتایج این تحقیق نشان داد که این آزمون، تأثیرات تعقیبی فراوانی را بر روی شرکت‌کنندگان بر جای می‌گذارد.

تمایزات و نوآوری تحقیق

تحقیق حاضر از سه حیث معرفت‌شناسی، روش‌شناسی و کارکردشناسی، از تحقیقات انجام‌شده در این زمینه متمایز است و دارای نوآوری است. از لحاظ معرفت‌شناختی، تحقیق حاضر ریشه در آن دسته از دیدگاه‌های نظری دارد که نگاهی هرمنوتیکی به روایی پسینی آزمون دارند و با ترکیب مواضع نظری گوناگون، جنبه‌های مختلف آموزشی و فراآموزشی را به صورت جامع در یک قالب نظری قرار داده است. به همین لحاظ، تحقیق حاضر هم تأثیرات معکوس آزمون را بر آموزش می‌سنجد و هم تأثیرات فرابافتی تعقیبی بر مشارکان گوناگون را مورد بررسی قرار می‌دهد. از لحاظ روش‌شناختی نیز در سطوح گوناگون طرح تحقیق، گام‌های تحقیق، تعداد و نوع شرکت‌کنندگان در تحقیق، تلفیق روشی، کاربست نرم‌افزارهای تحلیل کمی و کیفی، و حصول نتایج نظری و عملی، اعتباربخشی یافته‌ها و تعمیم‌پذیری آن‌ها در نوع خود بی‌نظیر است که به تفصیل به آن پرداخته شده است. از لحاظ کاربردشناسی نیز، تحقیق حاضر اولین تحقیق در نوع خود است که مستند به نتایج کمی و کیفی گام‌های پیشین خود، تغذیه معتبری را در اختیار راهبردهای آموزش، آینده‌پژوهان این حیطه و مدیران آموزشی قرار می‌دهد و منطبق با نیازهای واقعی فضای آموزشی کشور است.

روش

در تحقیق حاضر به تناسب اهداف ذکرشده، از طرح تحقیقی چندفازی-چند روشی^۲ استفاده شد. در برخی از فازهای مختلف این طرح نیز به دلیل الزامات روش تحقیق استفاده شده، برای به دست آوردن اطلاعات لازم از تلفیقی از روش‌های کمی و کیفی استفاده شده است تا تصویری کامل از پدیده مورد مطالعه به دست آید. کرسول (۲۰۰۸) معتقد است استفاده هم‌زمان از روش‌ها و رویکردهای گوناگون کمی و کیفی به تحقیق باعث غنای

1. EPT

2. multi-phase multi-method design

یافته‌ها و افزایش تعمیم‌پذیری آن‌ها می‌شود. در این تحقیق، به تناسب سؤالات مطرح‌شده از سه فاز اصلی پدیدارشناسی کیفی، مدل‌سازی معادله ساختاری و تحلیل ساختاری در ماتریس اثرات متقاطع استفاده شد که در ادامه بحث شده‌اند:

الف) پدیدارشناسی^۱

برای پاسخ‌گویی به سؤال اول تحقیق و به‌دست آوردن نقاط واگرایی اهداف سنجۀ زبان انگلیسی آزمون ورودی دانشگاه و برنامه ملی درسی از روش کیفی پدیدارشناسی استفاده شد. براساس مارتون و پونگک (۲۰۰۵) برخلاف روش‌های سنتی کیفی که به بررسی رخدادهای خارج از حس و ادراک انسان می‌پردازند، پدیدارشناسی یک روش تحقیق کیفی است که به نگاشت ادراک و تجربه انسان از یک رخداد منجر می‌شود. برای این منظور در ابتدای شش نفر و درنهایت ده نفر با تخصص آزمون‌سازی زبان انگلیسی (دارای سابقۀ پژوهش، تألیف و یا تدریس دانشگاهی در زمینه آزمون‌سازی که یکی از زیرشاخه‌های آموزش زبان انگلیسی در دانشگاه است) با گرایش‌های آموزش زبان انگلیسی و زبان‌شناسی کاربردی از دانشگاه‌های آزاد و دولتی و آموزش و پرورش بیرجند، مشهد و تهران به‌صورت هدف‌دار انتخاب شدند. دلیل افزایش شرکت‌کنندگان از شش به ده، بررسی روند اشباع نظری یافته‌ها براساس مضمون‌های ظهور یافته بود که با شش نفر محقق نشد. ابزار اصلی گردآوری اطلاعات مصاحبه نیمه‌ساختار^۲ یافته بود که براساس دستورالعمل مطرح‌شده در تایت^۳ (۲۰۱۵) برنامه‌ریزی شد. برای جلوگیری از ریزش معنایی در ترجمۀ مفاهیم، مصاحبه‌ها به زبان فارسی برگزار شد. پس از رونویسی مصاحبه‌ها به شیوۀ نگاشت مفهومی^۴، فرایندهای تحلیل محتوا و کدگذاری انجام شد و محورهای اصلی واگرایی اهداف آزمون از برنامه درسی ملی احصا شد. برای تضمین پایایی و روایی یافته‌ها از راهبردهای تخمین بین-مصاحبه‌شوندگان^۵ (اسجوستورم و دالگرین^۶، ۲۰۰۲) و کنترل عضو^۷ (رضایی^۸ و همکاران، ۲۰۲۱) استفاده شد.

1. phenomenography
2. semi-structured interview
3. Tight
4. concept mapping
5. inter-rater estimation
6. Sjostrom, & Dahlgren
7. member-checking
8. Rezaei

ب) مدل‌سازی معادله ساختاری

برای پاسخ‌گویی به سؤالات دوم و سوم تحقیق از مدل‌سازی معادله ساختاری استفاده شد. این مدل‌سازی در علوم انسانی روشی ترکیبی از رویکردهای کیفی و کمی در نرم‌افزارهای تحلیل داده اس.پی.اس.اس. و ایموس است که به بررسی و اعتباربخشی روابط علی بین متغیرهای ممکن در یک ساخت ذهنی می‌پردازد. این روش در علوم انسانی شامل سه مرحله اصلی ساخت پرسش‌نامه^۱، تحلیل عاملی اکتشافی^۲ و تحلیل عاملی تأییدی^۳ است. در مطالعه حاضر برای ساخت پرسش‌نامه و رسیدن به انگاره اولیه مفروض از سه منبع مصاحبه نخبگانی، سندپژوهی، و مرور ادبیات تحقیق استفاده شد. بر این اساس، از پنج نفر از شرکت‌کنندگان مرحله قبل استفاده شد و در مصاحبه‌هایی بازپاسخ، عواقب پسینی سنجه زبان انگلیسی آزمون وردی دانشگاه بر روی جنبه‌های مختلف آموزشی و فراآموزشی مورد سؤال قرار گرفت. به صورت هم‌زمان ادبیات تحقیق در پایگاه‌های معتبر داده مورد بررسی قرار گرفت. علاوه بر آن، به صورت موازی و متناسب با یافته‌ها، سندپژوهی در اسناد آموزش و پرورش انجام گرفت. در یک فرایند چرخشی^۴ از این سه روش جمع‌آوری داده، و تا زمان حصول اشباع نظری استفاده شد. منظور از چرخشی بودن فرایند این است که داده‌های به‌دست آمده در هر بخش باعث هدایت نظری یا تکمیل یافته‌های حاصل از بخش دیگر می‌شد. براساس یافته‌های حاصل شده، و با استفاده از دستورالعمل مطرح شده توسط و دورنیه و تاگوچی^۵ (۲۰۱۰)، پرسش‌نامه اولیه ساخته شد و در یک مطالعه آزمونه (پایلوت) بر روی شصت دانشجوی سال اول کارشناسی رشته زبان انگلیسی که در بازه زمانی یک سال گذشته در آزمون ورودی دانشگاه شرکت کرده بودند، مورد بررسی قرار گرفت و پایایی آن (با ضریب آلفای کرونباخ ۰,۷۸) مورد تأیید قرار گرفت. این دانشجویان به صورت نمونه‌برداری در دسترس از دانشگاه‌های بیرجند انتخاب شدند. سپس برای کشف عامل‌ها و متغیرهای ذیل آن از تحلیل عاملی اکتشافی استفاده شد. در این مرحله ۲۷۰ شرکت‌کننده از دانش‌آموزان پشت‌کنکوری و دانشجویان سال اول دانشگاه، مدیران آموزش پرورش و

1. questionnaire development
2. exploratory factor analysis
3. confirmatory factor analysis
4. iterative
5. Dornyei & Taguchi

معلمان زبان به صورت در دسترس از مدارس و دانشگاه‌های بیرجند، تربت حیدریه و مشهد انتخاب شدند و براساس پاسخ آن‌ها به پرسش‌نامه، بار عاملی به دست آمده، تعداد فاکتورها و متغیرهای ذیل آن‌ها کشف شدند و انگاره اولیه تحقیق شکل گرفت. برخلاف مطالعات انجام شده در ادبیات تحقیق که اغلب معطوف به یک گروه از ذینفعان آزمون است، در مطالعه حاضر، دلیل استفاده از شرکت کنندگان گوناگون، جمع‌آوری اطلاعات از همه ذینفعان آزمون سراسری و ارائه تصویری کامل تر از عواقب این آزمون است. سپس برای تعیین شاخص‌های برآزش انگاره حاصل شده و انطباق آن با واقعیت‌های فضای آموزشی کشور، از تحلیل عاملی تأییدی در افزونه ایموس و نرم‌افزار اس.پی.اس.اس. (نسخه ۲۳) استفاده شد. در این مرحله نیز ۲۴۰ شرکت کننده متفاوت به صورت در دسترس از جمعیت بالا مشارکت داشتند. دلیل تغییر تعداد و ماهیت شرکت کنندگان در گام‌های مختلف تحقیق، الزامات روش مدل‌سازی معادله ساختاری و رعایت پیش‌فرض‌های مورد نیاز برای تحلیل‌های عاملی اکتشافی و تأییدی است.

ج) تحلیل ساختاری میک‌مک

برای پاسخ‌گویی به سؤال چهارم تحقیق، از میان ده نفر استاد شرکت کننده در مطالعه، چهار نفر به صورت هدف‌دار انتخاب شدند و ب اساس متغیرهای حاصل شده، تأثیر هر متغیر را در نظام انگاره کشف شده، در ماتریس اثرات متقاطع، ارزشیابی شدند. سپس گراف رفتاری مدل و گراف ضرایب تأثیرات مستقیم احصا شد و براساس آن نوع متغیرها، ماهیت آن‌ها و رفتار آن‌ها در نظام روایی پسینی آزمون مذکور، احصا شد و متناسب با برخی راهبردهای آموزشی پیشنهاد شد.

یافته‌ها

برای پاسخ‌گویی به سؤال اول تحقیق مبنی بر بررسی نقاط واگرایی اهداف برنامه ملی درسی ایران و سنجه زبان انگلیسی آزمون ورودی دانشگاه، از پدیدارشناسی و از طریق ابزار مصاحبه از نخبگان استفاده شد. نتایج حاصل از این بخش که در جدول شماره ۱، نشان داده شده است، حاکی از آن است که این دو مهم در پنج بُعد اصلی دچار تضاد اساسی هستند که حداقل در شانزده محور با بیشترین بسامد، نمود یافته است.

جدول ۱ ابعاد اساسی واکرایی اهداف برنامه ملی درسی ایران و سنجه زبان انگلیسی آزمون ورودی دانشگاه

محور اثرگذاری	اثر تعقیبی آزمون	انتظارات برنامه ملی درسی
فرایند آموزش	۱. آموزش زبان محور ۲. نگاه محصول‌گرا به آموزش	۱. آموزش یادگیری محور ۲. نگاه فرایند‌گرا به آموزش
	۳. برنامه درسی ساختارگرا ۴. تأکید بر شکل زبان	۳. برنامه درسی کارکردگرا ۴. تأکید بر ارتباط زبانی
	۵. آزمون تجمعی ۶. اعتقاد به انتقال دانش	۵. آزمون تکوینی ۶. اعتقاد به ایجاد زمینه خلق دانش
	۷. تسلط فراگیر بر ساختار زبان	۷. تأکید بر خودبستگی فراگیر در کاربست زبان
ادراک و نگرش	۱. انتظارات صرف آموزشی ۲. ارزشیابی خروجی محور	۱. انتظارات آموزشی، فرهنگی و اجتماعی ۲. ارزشیابی چندجانبه
سیاست آموزشی	۱. تأکید بر مهارت‌های آزمون زنی ۲. برنامه درسی تجویزی	۱. تأکید بر مهارت‌های ارتباطی ۲. برنامه درسی نیازمحور
	۳. مدیریت آموزشی	۳. مدیریت آموزشی، عقیدتی و رفتاری
عاطفه و حس	۱. انگیزش بیرونی ۲. نادیده‌گیری عاطفه	۱. انگیزش درونی ۲. تأکید بر نیازهای عاطفی
اجتماعی و فرهنگی	۱. نجات اجتماعی فراگیر ۲. خلأ فرهنگی	۱. شکوفایی اجتماعی فراگیر ۲. جهانی-بومی سازی فرهنگ

برای پاسخ‌گویی به سؤالات دوم و سوم تحقیق و رسیدن به انگاره معتبر از روایی پسینی سنجه زبان انگلیسی آزمون ورودی دانشگاه، از مدل‌سازی معادله ساختاری استفاده شد. ساخت پرسش‌نامه و تخمین پایایی آن براساس چارچوب مطرح‌شده توسط دورنیه و تاگوچی (۲۰۱۰) انجام گرفت که در بالا به آن اشاره شد. سپس برای انجام تحلیل عاملی اکتشافی، عاملیت‌پذیری داده‌ها مورد بررسی قرار گرفت که مقادیر مربوط به آن در جدول ۲، ذکر شده است:

جدول ۲ تخمین عاملیت‌پذیری داده‌ها برای انجام تحلیل عاملی اکتشافی

مقیاس کی‌ام‌او برای کفایت نمونه ^۱	۰,۶۲۷
آزمون کرویت بارتلت ^۲	۳۶۱۶,۶۴۵
دی اف ^۳	۳۵۱
معناداری	۰,۰۰۰

1. KMO Measure of Sample Adequacy
2. Bartlett's test of sphericity
3. Df

نتایج حاصل از جدول ۲ نشان می‌دهد که هر سه مفروض انجام تحلیل عاملی اکتشافی شامل کفایت نمونه، عاملیت‌پذیری داده‌ها و آزمون کرویت از مقادیر قابل قبولی برخوردارند. براساس پالانت (۲۰۰۷)، مقادیر کی‌ام‌او بیشتر از ۰٫۵ قابل قبول است. همچنین براساس پالانت (۲۰۱۳)، معناداری آزمون کرویت باید کم‌تر از ۰٫۰۵ باشد که در مطالعه حاضر محقق شده است (ارزش پی^۱ برابر با ۰٫۰۰۰). همچنین، پلات پراکندگی^۲ و نتایج حاصل از بار عاملی آیت‌ها (متغیرها) نشان داد که ۲۶ متغیر ذیل هفت مؤلفه (فاکتور) دسته‌بندی شده‌اند که در جدول ۳ نشان داده شده است.

جدول ۳ ماتریس مؤلفه‌ها (فاکتورها) براساس چرخش وریمکس^۳

مؤلفه							
۷	۶	۵	۴	۳	۲	۱	
						۰٫۸۹۲	آیت‌م ۱
						۰٫۹۱۱	آیت‌م ۲
						۰٫۷۶۳	آیت‌م ۳
						۰٫۸۳۱	آیت‌م ۴
						۰٫۸۵۷	آیت‌م ۵
						۰٫۹۴۰	آیت‌م ۶
					۰٫۹۲۳		آیت‌م ۷
					۰٫۷۹۱		آیت‌م ۸
					۰٫۸۷۰		آیت‌م ۹
				۰٫۹۴۳			آیت‌م ۱۱
				۰٫۸۵۱			آیت‌م ۱۲
				۰٫۸۷۲			آیت‌م ۱۳
				۰٫۹۳۱			آیت‌م ۱۴
				۰٫۸۵۳			آیت‌م ۱۵
			۰٫۸۵۴				آیت‌م ۱۶
			۰٫۷۳۲				آیت‌م ۱۷
			۰٫۷۵۰				آیت‌م ۱۸
		۰٫۸۹۰					آیت‌م ۱۹
		۰٫۹۳۱					آیت‌م ۲۰

1. P Value
2. scree plot
3. Varimax rotation

ادامه جدول ۳

مؤلفه						
۷	۶	۵	۴	۳	۲	۱
		۰,۶۹۵				آیتم ۲۱
	۰,۷۶۳					آیتم ۲۲
	۰,۸۲۱					آیتم ۲۳
	۰,۷۱۰					آیتم ۲۴
۰,۹۱۳						آیتم ۲۵
	۰,۸۲۲					آیتم ۲۶
	۰,۷۲۷					آیتم ۲۷

همان‌طور که در جدول ۳ نشان داده شده است، ۲۶ متغیر با بار عاملی قابل قبول در هفت فاکتور دسته‌بندی شده‌اند که مؤلفه‌ها و متغیرهای مدل فرضی اولیه این مطالعه را شکل داده‌اند. این متغیرها که به شکل آیتم در جدول ۳ حضور دارند به همراه شرح آن‌ها در جدول ۴ ارائه شده‌اند:

جدول ۴ شرح مؤلفه‌ها و متغیرهای ذیل هر مؤلفه در انگاره‌فرضی روایی پسینی سنجه زبان انگلیسی آزمون ورودی دانشگاه

مؤلفه	متغیر	شرح متغیر
آموزشی ^۱	آیتم ۶	زبان مادری (فارسی) در کلاس آموزش زبان، جایگزین زبان انگلیسی می‌شود.
	آیتم ۲	روش تدریس متناسب با آزمون ورودی دانشگاه و متناقض با کتاب‌های ویزن است.
ارزشیابی	آیتم ۱	ارزشیابی کلاسی و بازخوردها متناسب با آزمون ورودی دانشگاه و متناقض با کتاب‌های ویزن است.
	آیتم ۵	زبان معلمان و تعاملات کلاسی معطوف به لغات و دستور زبان است.
مهارت‌های ارتباطی و گفتاری-شنیداری	آیتم ۴	مهارت‌های ارتباطی و گفتاری-شنیداری در کلاس نادیده گرفته می‌شوند.
	آیتم ۳	محتوای درسی و محتوای آموزشی متناسب با آزمون ورودی دانشگاه و متناقض با کتاب‌های ویزن است.
اجرائی ^۲	آیتم ۷	اهداف برنامه درسی کتاب‌های ویزن با اهداف سنجه زبان انگلیسی آزمون ورودی دانشگاه متناقض است.
	آیتم ۹	سیاست آموزشی آزمون ورودی، خلق تأثیرات و تغییرات مثبت در نظام آموزشی نیست.
آیتم ۸	تغییرات آزمون ورودی دانشگاه متناسب با تغییرات برنامه درسی کتاب‌های ویزن نیست.	

1. pedagogical
2. administrative

ادامهٔ جدول ۴

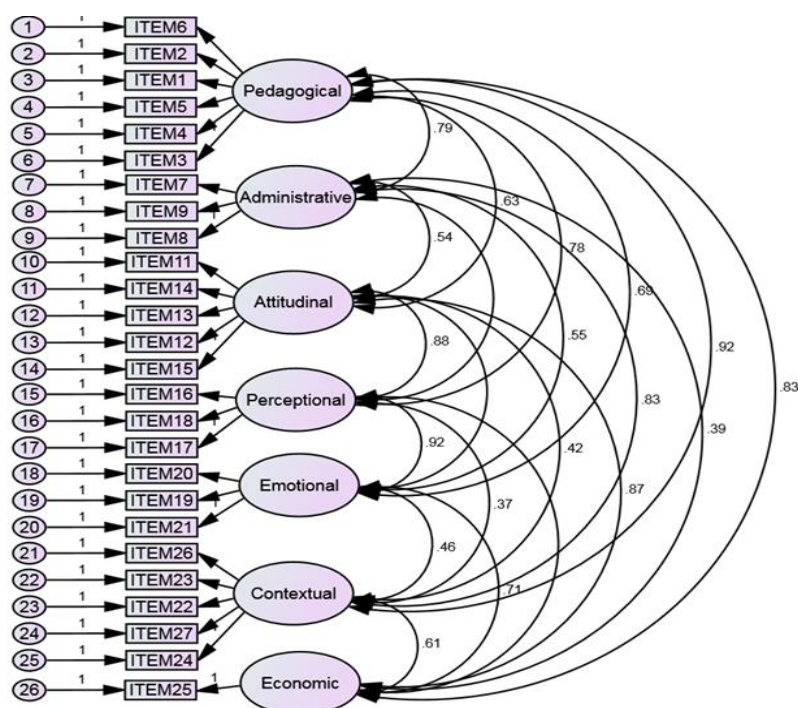
مؤلفه	متغیر	شرح متغیر
نگرشی ^۱	آیتم ۱۱	در ذهن فراگیران، کسب توانش آزمون ورودی دانشگاه، بر کسب توانش زبانی ارجحیت دارد.
	آیتم ۱۴	انگیزش فراگیران برای یادگیری مهارت آزمون‌زنی فراتر از انگیزش آن‌ها برای کسب مهارت زبانی است.
	آیتم ۱۳	از دید مدارس، موفقیت در آزمون ورودی دانشگاه بر موفقیت در یادگیری زبان دوم ارجحیت دارد.
	آیتم ۱۲	در ذهن فراگیران، ارزش تعقیبی آزمون ورودی دانشگاه، بسیار فراتر از ارزش یادگیری زبان دوم است.
	آیتم ۱۵	در ذهن معلمان، تدریس برای آزمون بر تدریس برای اهداف ارتباطی کتاب‌های ویژن، ارجحیت دارد.
ادراکی ^۲	آیتم ۱۶	مشارکان (معلمان، فراگیران، ...) دریافته‌اند که موفقیت براساس عملکرد آزمون ورودی قضاوت می‌شود.
	آیتم ۱۸	مشارکان (معلمان، فراگیران، ...) دریافته‌اند که اولویت یادگیری، عملکرد در آزمون ورودی دانشگاه است.
	آیتم ۱۷	مشارکان (معلمان، فراگیران، ...) معتقدند که یادگیری توانش زبانی یک موضوع عارضی (شانسی) است.
عاطفی ^۳	آیتم ۲۰	آزمون ورودی دانشگاه در مقایسه با برنامهٔ درسی ویژن، فشار روانی و اضطراب بیشتری بر فراگیران دارد.
	آیتم ۱۹	اجبار و تهییج فضای آموزشی کشور به سمت آزمون ورودی دانشگاه است تا برنامهٔ درسی مصوب.
	آیتم ۲۱	فشار آزمون ورودی، خلاقیت زبانی و تفکر انتقادی را در فراگیران کتاب‌های ویژن، از بین می‌برد.
بافتی ^۴	آیتم ۲۶	آزمون ورودی دانشگاه تنها مسیر ممکن برای کسب موقعیت اجتماعی است.
	آیتم ۲۳	آزمون ورودی دانشگاه تنها مسیر ممکن برای انتخاب رشتهٔ تحصیلی مورد علاقه است.
	آیتم ۲۲	اهداف آزمون ورودی دانشگاه، با نیازهای فراگیر زبان دوم تناقض رادیکال دارد.
	آیتم ۲۷	آزمون ورودی دانشگاه کم‌ترین سازگاری را با علایق آموزشی و زندگی فراگیران ندارد.
	آیتم ۲۴	آزمون ورودی دانشگاه تنها مسیر ممکن برای یافتن شغل است.
اقتصادی ^۵	آیتم ۲۵	در مقایسه با برنامهٔ درسی، آزمون دانشگاه فشار مالی فراوانی بر فراگیران و خانواده‌ها دارد.

1. attitudinal
2. perceptual
3. emotional
4. contextual
5. economic

در گام سوم مدل‌سازی معادله ساختاری، و مشخصاً در پاسخ به سؤال سوم تحقیق، به منظور تعیین تطابق انگاره استخراج شده با واقعیت‌های محیطی، از تحلیل عاملی تأییدی استفاده شد و ضمن تعیین شاخص‌های برازش مدل، گراف انگاره از طریق افزونه ایموس، ارائه شد که در جدول ۵ و شکل ۱، ارائه شده است.

جدول ۵ شاخص‌های برازش برای انگاره استخراج شده

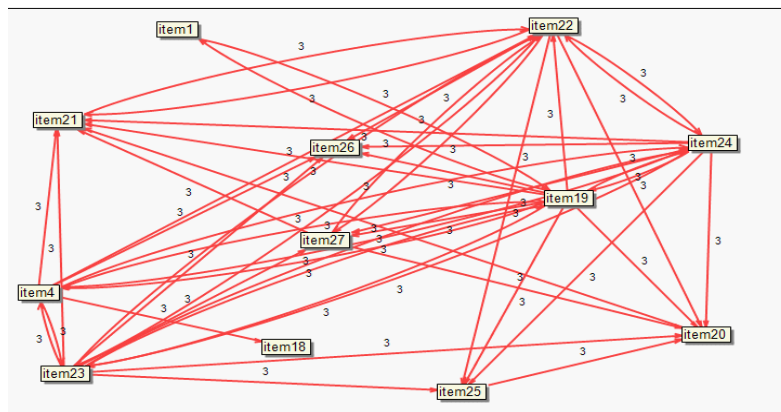
شاخص	سطح کنونی	سطح قابل قبول	نتیجه گیری
مربع خی (کای) ^۱	۰,۱۰۰	کم‌تر از ۳	قابل قبول
شاخص نرم‌شدگی برازش ^۲	۰,۹۱	بیشتر یا مساوی ۰,۹۰	قابل قبول
شاخص برازش تطبیقی ^۳	۰,۹۳	بیشتر یا مساوی ۰,۹۰	قابل قبول
ریشه میانگین مربعات خطای برآورد	۰,۰۱	کم‌تر از ۰,۰۵	قابل قبول



شکل ۱ گراف نهایی انگاره‌ی روایی پسینی سنجۀ زبان انگلیسی آزمون ورودی دانشگاه

1. Chi-Square (X²)/df
2. Normed Reference Index
3. Comparative Reference Index

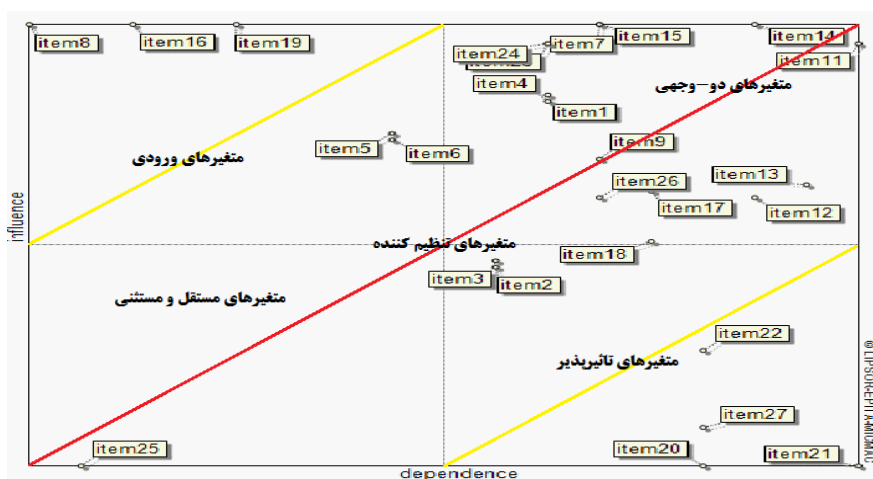
همان‌طور که در جدول ۵ نشان داده شده است، شاخص‌های برازش مدل مطابق با مقادیر قابل پذیرش (کوستا و سارمنتو^۱، ۲۰۰۷) است و انگاره استخراج شده از شاخص‌های تطابق قابل قبولی برخوردار است. گراف نهایی این انگاره در شکل ۱ نشان داده شده است. در گراف بالا، ۲۶ متغیر انگاره روایی پسینی سنجۀ زبان انگلیسی آزمون ورودی دانشگاه در هفت مؤلفه اصلی به همراه رابطه و تعامل آن‌ها در قالب ضرایب هم‌بستگی نشان داده شده است. این در حالی است که داده‌های به دست آمده، تنها تحلیلی از متغیرهای مؤثر در ساخت اثرات تعقیبی این آزمون تأثیرگذار را نشان می‌دهد. اما در مقاله حاضر به تحلیل وضع موجود^۲ بسنده نشده است و برای این که بستری برای تدوین راهبردهای آموزشی و مدیریتی در زمینه حذف یا تعدیل اثرات تعقیبی این آزمون فراهم شود، با استفاده از داده‌های اعتباربخشی شده موجود در انگاره فوق، به تحلیل ساختاری این متغیرها در ماتریس اثرات متقاطع پرداخته‌ایم. برای این منظور، پنج نفر از شرکت کنندگان نخبگانی حاضر در بخش اول تحقیق، در این تحلیل حضور داشتند و برای تعیین تأثیر این متغیرها در ماتریس میک‌مک، مقادیر لازم را در این ماتریس وارد کردند. بر این اساس، برای تأثیرات قوی، متوسط، ضعیف، بدون تأثیر و تأثیر بالقوه به ترتیب مقادیر یک، دو، سه، صفر و پی^۳، در نظر گرفته شد. نتایج حاصل از این تحلیل در ادامه آورده شده است.



شکل ۲ نگاشت متغیرهای دارای بیشترین تأثیر متقابل در نظام روایی پسینی سنجۀ زبان انگلیسی آزمون ورودی دانشگاه

1. Costa & Sarmento
2. Status Quo
3. Potential (P)

در نگاشت فوق نشان داده شده است که متغیرهای شماره ۱، ۲۲، ۲۱، ۲۶، ۱۹، ۲۴، ۲۷، ۴، ۱۸، ۲۵، ۲۰ یازده متغیری هستند که دارای بیشتر تأثیرگذاری در شکل‌گیری و تقویت و یا ایجاد هم‌افزایی با دیگر متغیرها هستند. بنابراین، این متغیرها در مؤلفه‌های خود دارای تأثیرات حیاتی هستند و برای کنترل نظام تأثیرات و یا تعدیل و تغییر این تأثیرات پسینی سنجۀ زبان انگلیسی آزمون ورودی دانشگاه، باید این متغیرها را کنترل و تغییر داد. این در شرایطی است برای بررسی رفتار این سیستم و اتخاذ راهبردهای آینده‌نگرانه، گراف تأثیر مستقیم متغیرها در ماتریس تأثیرات متقاطع، اطلاعات بیشتری را از ماهیت و شیوۀ رفتار این متغیرها در دستگاه فراهم می‌کند. این اطلاعات در شکل شماره ۳، نشان داده شده است.



شکل ۳ نگاشت رفتار و ماهیت متغیرها در سیستم روایی پسینی سنجۀ زبان انگلیسی آزمون ورودی دانشگاه

یافته‌های حاصل از این بخش که مهم‌ترین خروجی نرم‌افزار میک‌مک در تحلیل ساختاری است اطلاعات گوناگونی را برای اتخاذ راهبردهای آموزشی در اختیار کنش‌گران این بخش قرار می‌دهد که در ادامه به مهم‌ترین این یافته‌ها اشاره شده است و متناسب با آن راهبردهای مدیریتی ارائه شده است:

الف) پایداری سیستم روایی پسینی آزمون: براساس نگاشت تأثیرات مستقیم متغیرها در ماتریس اثرات متقاطع، توزیع متغیرها در نمودار نشان‌دهنده وضعیت پایداری/ناپایداری است. بر این اساس، اگر توزیع متغیرها در نمودار به شکل L باشد، نشان‌دهنده پایداری بودن

سیستم است و بدان معناست که این سیستم با تغییرات اندک و یا تعدیل برخی متغیرها به راحتی دچار تغییر وضعیت نخواهد شد. اما توزیع متغیرها در نگاشت فوق نشان‌دهنده پراکنده بودن متغیرها حول محورهای گوناگون است. این بدان معناست که برای تغییر سیستم روایی پسینی سنجۀ زبان انگلیسی، نیاز به تغییرات گسترده‌ای در مؤلفه‌های فراوانی است و با تغییرات بخشی در یک بُعد یا یک یا چند متغیر، نمی‌توان تأثیرات تعقیبی ناشی از این آزمون سرنوشت‌ساز را کنترل کرد، لذا، برای ایجاد تغییرات مثبت، نیاز به تغییرات زیادی در اغلب متغیرهاست. اولویت این تغییرات در ادامه (یافته‌های زیر) مشخص شده است.

ب) متغیرهای ورودی (متغیرهای حیاتی^۱/متغیرهای محیطی^۲): این متغیرها که در شمال غربی نگاشت وابستگی/تأثیر (شکل ۳) قرار دارند متغیرهای ورودی سیستم هستند و این بدان معناست که دست‌کاری و تغییر آن‌ها موجب بیشترین تغییرات در کل سیستم می‌شود. این متغیرها به دو دسته حیاتی و محیطی تقسیم می‌شوند. متغیرهای حیاتی در بالای خط زرد قطری بخش شمال غربی نگاشت قرار دارند و بیشترین تأثیرگذاری را در نظام اثرات تعقیبی این آزمون دارند. این متغیرهای حیاتی آیت‌های ۸، ۱۶، و ۱۹ هستند. همچنین در بخش پایینی خط قطری همین ناحیه، متغیرهای محیطی قرار دارند که علی‌رغم تأثیرگذاری زیاد آن‌ها بر سیستم، به دلیل محیطی بودن و وابسته به بافت بودن، تغییرات آن‌ها بسیار دشوار است. اما در صورت ایجاد زمینه تغییر آن‌ها، تأثیرات بسیار بنیادینی بر کاهش تأثیرات تعقیبی منفی آزمون زبان انگلیسی برجای می‌گذارند. متغیرهای ۵ و ۶ در این دسته قرار می‌گیرند.

ج) متغیرهای دووجهی^۳ (متغیرهای ریسک^۴/متغیرهای هدف^۵): این متغیرها که در قسمت شمال شرقی نگاشت تأثیرات قرار گرفته‌اند دارای ماهیتی بسیار ناپایدارند و براساس شرایط حاکم بر سیستم، می‌توانند در نقش متغیرهای ورودی یا متغیرهای خروجی سیستم ظاهر شوند. این متغیرها به دو دسته متغیرهای ریسک و هدف تقسیم‌بندی می‌شوند. متغیرهای ریسک نقش کلیدی‌تری در سیستم بازی می‌کنند و تمایل بیشتری برای تبدیل شدن به متغیر ورودی دارند. این متغیرها در بالای خط قطری قرمز این بخش قرار دارند و

1. Critical
2. Environmental
3. bi-dimensional
4. risk
5. target

بر این اساس متغیرهای ۱، ۴، ۷، ۱۴، ۹، ۱۵ و ۲۴ در این دسته جای می‌گیرند. متغیرهای هدف، که در زیر خط قطری این ناحیه قرار دارند، بیشتر تمایل دارند که تحت تأثیر سایر متغیرها به‌عنوان متغیر خروجی سیستم باشند. بنابراین، سرمایه‌گذاری بر روی تغییر و یا تعدیل آن‌ها، از لحاظ راهبردی، تأثیر کم‌تری بر ایجاد تأثیرات مثبت با کاهش تبعات منفی آزمون خواهد داشت. متغیرهای ۱۱، ۱۲، ۱۳، ۱۷، ۱۸ و ۲۶ در این دسته قرار دارند.

(د) متغیرهای خروجی: این متغیرها که در بخش جنوب شرقی نگاشت قرار دارند، بیش‌ترین حساسیت را نسبت به تغییر در سایر متغیرها از خود نشان می‌دهند و دچار تغییر می‌شوند. از این رو به این متغیرها وابسته یا متغیر خروجی گفته می‌شود. متغیرهای ۲، ۳، ۲۰، ۲۱، ۲۲ و ۲۷ در این دسته قرار می‌گیرند.

(ه) متغیرهای مستقل (اهرمی/مستثنی): این متغیرها که در جنوب غربی نگاشت قرار دارند، ضعیف‌ترین رابطه را با سیستم تأثیرات دارند و به‌ندرت تأثیرگذار و تأثیرپذیر می‌شوند اما با توجه به کاربست آن‌ها در شرایط گوناگون، می‌توان از آن‌ها استفاده کرد. این متغیرها در دو دسته اصلی جای گرفته‌اند. متغیرهای اهرمی، که در بالای خط قطری قرمز در این ناحیه قرار دارند، دارای اهمیت راهبردی ثانویه هستند و در صورت ناکارآمد بودن تغییرات ایجادشده در متغیرهای حیاتی و محیطی، از آن‌ها به‌عنوان متغیر اهرمی و کمکی برای ایجاد تغییرات در سیستم استفاده می‌شود. هیچ متغیری از سیستم حاضر در این دسته جای نگرفته است. اما متغیر شماره ۲۵ در زیر خط قطری و در حیطه متغیرهای مستثنی قرار گرفته است که کم‌ترین رابطه را با سیستم آموزشی فوق دارد. صرف‌نظر از مکان قرارگیری متغیرهای فوق، هر چه این متغیرها به خط قطری قرمز نزدیک‌تر باشند، در مقایسه با هم‌دسته‌های خود، تأثیر راهبردی تری دارند و بر این اساس اولویت‌بندی می‌شوند.

بحث و نتیجه‌گیری

در یک روش تحقیق نوآورانه و با هدف پوشش دادن برخی نقص‌های روش‌شناختی و معرفت‌شناختی در مطالعه اثرات تعقیبی آزمون، تحقیق حاضر ناظر به سه وجه اصلی پدیده‌ی روایی پسینی در سنجه‌ی زبان انگلیسی آزمونی تأثیرگذار و سرنوشت‌ساز انجام گرفت. در وجه اول، با استخدام روش کیفی پدیدارشناسی و تحلیل نظرات نخبگانی، ریشه‌ی ناسازگاری اهداف این آزمون با اهداف نوین آموزش زبان انگلیسی که در برنامه ملی درسی عنوان شده است، کشف و گزارش شد. در وجه دوم، با استفاده از مدل‌سازی معادله

ساختاری، تحلیلی جامع از محورهای اثرات تعقیبی این آزمون و متغیرهای آن، به همراه روابط و چگونگی تعامل این متغیرها در قالب یک انگاره معتبرسازی شده، ارائه شد. در وجه سوم نیز با کاربست تحلیل ساختاری، ماهیت و رفتار این متغیرها در نظام تأثیرات تعقیبی آزمون کشف و متناسب با آن پیشنهادهایی در قالب چگونگی مهندسی و تغییر و تعدیل این مؤلفه‌ها براساس اولویت، ارائه شد. نتایج حاصل از این بخش، از لحاظ کمیت و کیفیت، جامع‌تر از بسیاری از تحقیقات پراستنادی است که در این حیطه انجام گرفته است. برای مثال، در بُعد متغیرهای بررسی شده بسیار جامع‌تر از هوانگ و هوانگ (۲۰۱۹) است. بُعد تحلیل ساختاری تحقیق حاضر نیز کاملاً منحصربه‌فرد و دارای کاربردهای فراوانی است. تحلیل یافته‌ها در بخش پدیدارشناسی که حاکی از عدم تطابق اساسی بین اهداف آزمون و اهداف برنامه درسی کتاب‌های ویزن در دوره دبیرستان است، با نتیجه‌گیری لی (۲۰۱۸) سازگاری دارد. نتایج پدیدارشناسی نشان داد که علی‌رغم کوشش‌های عملی، نظری و مدیریتی و حتی سیاست‌گذاری آموزشی برای تغییر برنامه درسی به سری کتاب‌های ویزن، بسیاری از اهداف آموزشی محقق نشده است و جریان اثر معکوس آزمون سرنوشت‌ساز ورودی دانشگاه ایران، بخش بسیار زیادی از این تلاش‌ها را خنثی کرده است. پیش‌تر نیز دیمووا (۲۰۱۷) دریافته بود که آزمون‌های تأثیرگذار، اثر برنامه آموزشی را در ابعاد گوناگون از بین می‌برند. براساس لی (۲۰۱۸، ص ۵۲) در کشورهای دنیا، تغییر اهداف آزمون قبل از تغییر اهداف آموزش صورت می‌گیرد و اساساً این آزمون است که موجب ایجاد تأثیرات مثبت و منفی در آموزش می‌شود و این در حالی است که تغییر برنامه آموزشی بدون تغییر اهداف آزمون فقط به اثرات تعقیبی منفی و غیرسازنده منجر می‌شود. بر همین اساس، پن و روور (۲۰۱۶) تصریح کرده‌اند از آنجایی که آزمون مبتنی بر نتیجه آموزش است، اهرم تأثیر قدرتمندی است که از آن می‌توان برای القای تغییرات مثبت در برنامه درسی استفاده کرد. بر همین اساس، به نظر می‌رسد تا زمانی که مفاد و اهداف آزمون ورودی دانشگاه (موسوم به کنکور) در ایران تغییر نکنند، هرگونه تغییر در برنامه درسی محکوم به شکست و یا نتیجه حداقلی است. در بُعد ساخت انگاره از روایی پسینی آزمون نیز، یافته‌های مطالعه حاضر با یافته‌های بک‌من و پالمر (۲۰۰۰) و قربانی (۲۰۰۸) تطابق دارد که نشان داده‌اند که آزمون‌های تأثیرگذار علاوه بر جنبه‌های آموزشی بر جنبه‌های غیرآموزشی و سیاسی و اجتماعی و اقتصادی نیز تأثیر می‌گذارند. البته در این مطالعه به دلیل اتخاذ رویه چرخشی در گردآوری داده‌ها تا مرحله اشباع نظری،

تعداد متغیرها و محورهای کشف شده در مقایسه با مطالعات داخلی و خارجی، بسیار گسترده‌تر است. در بُعد سوم تحقیق که معطوف به تحلیل ساختاری است نیز خلأ تحقیقاتی موجود در ادبیات تحقیق پوشش داده شده است و با استفاده از تحلیل ماتریس اثرات متقاطع، راهبردهایی آموزشی ارائه داده شده است که در نوع خود از لحاظ روش‌شناسی و کاربردشناسی منحصر به فرد است. نتایج این گام برای راهبردها سازان و آینده‌پژوهان منبع تغذیه‌ای بسیار معتبر به‌شمار می‌رود. همچنین، نتایج این تحقیقات برای معلمان زبان انگلیسی، مدیران آموزشی کشور و سیاست‌گذاران کلان حوزه آموزش، کاربرد بنیادین دارد. اما مهم‌ترین کاربردهای پژوهش حاضر، استفاده از ظرفیت تبیینی و پدیدارشناختی نتایج برای جلوگیری از عواقب سوء آزمون در آزمون‌های ورودی تحصیلات تکمیلی (کارشناسی ارشد و دکتری) است. اگرچه براساس مرور پیشینه تحقیق، برخی از گزارش‌ها ناظر به تغییر آزمون به‌عنوان متغیر ورودی به‌منظور ایجاد تغییر در برنامه آموزشی است، اما تحلیل ساختاری آزمون ورودی دانشگاه در ایران نشان داد که پدیده عواقب پسینی این آزمون دارای پیچیدگی‌های ساختاری و عنصری است. بر همین اساس، با تجربه‌نگاری حاصل از یافته‌های حاضر، پیشنهاداتی برای ارتقای روایی پسینی آزمون‌های تحصیلات تکمیلی دانشگاه براساس یک قالب تعاملی بین نظام آزمون‌سازی، نظام نیازسنجی آموزشی و نظام برنامه درسی ارائه شده است:

الف) آزمون‌سازی اقتضایی^۱: اگرچه به‌دلیل مشخص نبودن کانال آموزشی آزمون‌دهندگان در دانشگاه، سنجه زبان انگلیسی آزمون ورودی دانشگاه به‌صورت یک‌نواخت انجام می‌گیرد، اما در آزمون‌های تحصیلات تکمیلی، امکان آزمون‌سازی با اهداف ویژه وجود دارد که باعث انطباق اهداف آزمون با اهداف برنامه درسی و کاهش جریان معکوس منفی^۲ می‌شود. توجه به نیازهای زبانی رشته برای رسیدن به دانش تخصصی مطلوب و کاربست در فضای عملی، احصای نیازهای سازمان آموزشی یا حرفه‌ای یا ارگان ذینفع آزمون، احصای نیازهای ایدئال، استانداردها و شاخص‌های جهانی و نیازهای بافت آموزشی و حرفه‌ای بومی نیز از جمله این اقتضائات هستند که باعث کاهش عواقب پسینی و ارتقای روایی آزمون‌های تحصیلات تکمیلی می‌شوند.

ب) ترمیم گسست دوگانه آزمون-آموزش: آزمون‌های ورودی دانشگاه که در

1. adaptive testing
2. negative washback

شامل آزمون‌های نرم‌محور^۱ قرار می‌گیرد، نگاه محصول‌محور به آموزش دارند. این در حالی است که در نظریات جدید آموزش زبان، یادگیری به‌عنوان فرایند، و آزمون به‌عنوان ابزاری برای سنجش محصول به‌شمار نمی‌رود. یکی از راهبردهای فرایند دانستن یادگیری، عدم نگاه ابزاری انفکاک‌ی به آزمون و شامل آن در فرایند یادگیری است. از آنجایی که در آزمون‌های ورودی تحصیلات تکمیلی، رشته‌ها مجزا هستند، شرایط برای ترکیب کردن فرایندهای آزمون‌سازی و یادگیری مهیاتر است. این اتفاق باعث می‌شود که اهداف آزمون و یادگیری به‌صورت واحد دنبال شوند و اساتید دانشگاه نیز از فرایند آزمون‌های استاندارد حذف نشوند. انجام سازوکاری برای رسیدن به بسندگی زبان تخصصی در حین آموزش (برای مثال انجام ارزشیابی تکوینی زبان تخصصی هم‌زمان با آموزش و در طول تحصیل به جای آزمون تجمعی در آزمون ورودی و ارائه کفایت بسندگی) یکی از راه‌حل‌های ترمیمی در این حوزه به‌نظر می‌رسد.

ج) اصلاح سیاست‌گذاری آموزشی: یکی از مشکلات اصلی تشریح‌شده در بخش اول تحقیق حاضر، عدم انطباق اساسی سیاست‌های آموزشی با واقعیت‌های بافتی است. نتایج پژوهش ذیل سؤال اول تحقیق نشان داد که در زمینه برنامه درسی ملی و واقعیت‌های سیلابس آموزشی زبان (ویژن) و آزمون ورودی دانشگاه تفاوت رادیکال وجود دارد که باعث ایجاد جریان منفی معکوس در فرایند آموزش شده و ابعاد گوناگونی از ذینفعان را متأثر کرده است. تلفیق و سازگارسازی برنامه درسی انتزاعی-ایدئال^۲ (احصاشده از طریق نیازسنجی‌های جامع و استاندارد)، برنامه درسی ابلاغی^۳ (احصاشده از طریق سیاست‌گذاری آموزشی)، و برنامه درسی کاربردی^۴ (انجام‌شده مبتنی بر واقعیت‌های بافتی در محیط‌های آموزشی) در سطح سیاست‌گذاری، باعث همسانی اهداف آزمون و آموزش و کاهش عواقب پسینی سوء می‌شود. ویژگی‌های بافتی دانشگاه، به‌دلیل امکان کنترل بسیاری از متغیرها، شرایط لازم برای این همسانی را دارد که مورد بهره‌برداری قرار نگرفته است.

د) چالش‌انگاری عواقب آزمون: واقعیت آن است که آزمون‌های سرنوشت‌ساز مانند آزمون‌های ورودی دانشگاه، به‌دلیل نقض الزامات روایی پسینی، از حیث استاندارد ساقط شده‌اند. علاوه بر عواقب ارائه‌شده در نتایج تحقیق حاضر، عدول از اخلاق، انصاف و

1. norm-referenced
2. Abstract-Ideal Curriculum
3. Curriculum as Artifact
4. Curriculum in use

مساوات که جزو شاخص‌های روایی هستند، در آزمون‌های سرنوشت‌ساز به چشم می‌خورد. اگرچه برخی از این نقایص به دلیل الزامات محیطی و استفاده این آزمون‌ها به عنوان ابزار انتخابی برای پالایش عده‌ی معدودی از داوطلبان ورود به دانشگاه است، ولی نظر به کم‌تر بودن تعداد داوطلبان، امکان احیای برخی از این شاخص‌ها در آزمون‌های تحصیلات تکمیلی وجود دارد. بسط تحقیق در این زمینه، مدل‌سازی عواقب آزمون‌ها قبل از اجرا (مشابه تحقیق حاضر) که تصویر معتبری از پدیده‌ی آزمون را نشان می‌دهد، تدوین حلقه‌های علی^۱ برای تبیین بازخوردهای علی و سناریونگاری سیاست‌های آموزشی قبل از اجرا و انتخاب پیش‌برنده‌ترین آن‌ها، نه تنها اطلاعات معتبری را در اختیار سیاست‌گذاران، و آینده‌پژوهان این حوزه قرار می‌دهد، بلکه صرفه‌جویی‌های زمانی و هزینه‌ای مطلوبی را نیز در پی دارد.

منابع

- برنامه ملی درسی ایران (۱۳۹۰). وزارت آموزش و پرورش.
- جعفری گهر، م.، و جمالوندی، ب. (۱۳۹۷). معیارهای کتاب تخصصی مطلوب در رشته زبان انگلیسی. *پژوهش و نگارش کتب دانشگاهی*، ۲۲ (۴۳)، ۹۰-۱۱۵.
- فرخی‌پور، س.، خوش‌سیما، ه.، سارانی، ع.، و گنجی، م. (۱۳۹۹). تبیین و بررسی تأثیر انگاره بومی ارزشیابی پویا در شناخت و رفع مشکلات یادگیری مهارت‌های زبانی مولد در دانش‌آموزان متوسطه. *پژوهش‌های زبان‌شناختی در زبان‌های خارجی*، ۱۰ (۱)، ۱۲۰-۱۳۴.
- Ahmadi Safaa, M., & Goodarzi, S. (2014). The Washback Effects of Task-based Assessment on the Iranian EFL Learners' Grammar Development. *Procedia - Social and Behavioral Sciences*, 98, 90 - 99.
- Alavi, S.M., & Masjedlu, A. (2017). Construct under-representation and construct irrelevant variances on IELTS academic writing task: Is There Any Threat to Validity? *Theory and Practice in Language Studies*, 7(11), 1097-1107.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1-34.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2000). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
- Bailey, K. (1999). *Washback in Language Testing*. Princeton, New Jersey: ETS.
- Chapelle, C. A. (2020). An introduction to language testing's first virtual special issue:

- Investigating consequences of language test use. *Language Testing*, 37(4), 638–645.
- Chen, Q., Hao, C., & Xiao, Y. (2020). When testing stakes are no longer high: Impact on the Chinese college English learners and their learning. *Language Testing in Asia*, 10(1), 10–25.
- Cheng, L., Klinger, D., & Zheng, Y. (2007). The challenges of the Ontario Secondary School Literacy Test for second language students. *Language Testing*, 24(2), 185–208.
- Chou, M.U. (2015). Impacts of the Test of English Listening Comprehension on students' English learning expectations in Taiwan. *Language, Culture and Curriculum*, 28 (2), 191–208.
- Creswell, J. W. (2008). *Educational Research*. Los Angeles, London: Sage.
- Cronbach, L. J. (1980). *Validity on parole: How can we go straight? New directions for testing and measurement: Measuring achievement over a decade*. In Proceedings of the 1979 ETS Invitational Conference (pp. 99-108). San Francisco: Jossey- Bass.
- Cureton, E. E. (1951). Validity. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 621-694). Washington, DC: American Council on Education.
- Davies, A. (1968). *Language testing symposium: A psycholinguistic approach*. Oxford: Oxford University.
- Di-Gennaro, J.A. (2017). *The Washback effects of an English exit exam on teachers and learners in a Korean university English program*. Unpublished PhD dissertation, University of Exeter, United Kingdom.
- Dimova, S. (2017). Life after oral English certification: The consequences of the Test of Oral English Proficiency for Academic Staff for EMI lecturers. *English for Specific Purposes*, 46, 45–58.
- Dornyei, Z., & Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing*. Routledge, London.
- Fathi, F. (2021). *Developing a Native Model for the University Entrance Exam Washback Effect on English Language Teaching, Testing and Learning in Iran's National Curriculum: A Socio-cultural Perspective*. Unpublished PhD Dissertation. Islamic Azad University, Tehran, Iran.
- Ghorbani, M. R. (2008). ELT in Iranian high schools in Iran, Malaysia, and Japan: Reflections on how tests influence the use of prescribed textbooks. *Reflections on English Language Teaching*, 8, 131–139.
- Green, A. (2020). Washback in language assessment. In C. A. Chapelle (Ed.), *concise encyclopedia of applied linguistics*, (pp. 1159–1164). Wiley-Blackwell.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Hung, S. T. A., & Huang, H. T. D. (2019). Standardized proficiency tests in a campus-wide English curriculum: A washback study. *Language Testing in Asia*, 9(21), 1–17.
- Jamalifar, G., Salehi, H., Tabatabaei, O., & Jafarigohar, M. (2021). Washback effect of the English proficiency test (EPT) on PhD candidates' language learning strategies. *Journal of Language and Translation*, 11(3), 179–191.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (pp.17–64). Westport, CT: American Council on Education.
- Khosravani, M., Rostamian, M., Ashraf, H. et al. (2022). A structural equation modeling of English tests' social and educational consequences: exploring target, leverage, risk, and critical variables. *Language Testing in Asia*, 12(46), 1-23.
- Kim, S.J. (2017). *Exploring a Framework for Consequential Validity for Performance-Based Assessments*. Unpublished PhD Dissertation. University of Minnesota. U.S.A.
- Larssona, M & Olin-Scheller, Ch. (2020). Adaptation and resistance: washback effects of the national test on upper secondary Swedish teaching. *The Curriculum journal*, 31 (4), 687-703.
- Li, X. (2018). Self-assessment as 'assessment as learning' in translator and interpreter education: Validity and washback. *The Interpreter and Translator Trainer*, 12(1), 48–67.

- Marton, F., & Pong, W. Y. (2005). On the unit of description in phenomenography. *Higher Education Research & Development*, 24, 335-348.
- McNamara, T. (2006). Validity in language testing: The challenge of Sam Messick's legacy. *Language Assessment Quarterly*, 3(1), 31-51.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement*, (pp. 13-103). American Council on Education/Macmillan.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- Moss, P. A., Pullin, D., Gee, J. P., & Haertel, E. H. (2005). The idea of testing: Psychometric and sociocultural perspectives. *Measurement: Interdisciplinary Research & Perspective*, 3(2), 63-83.
- Pallant, J. (2007). *SPSS survival manual: A step-by-step guide to data analysis using SPSS for Windows*. Open University Press & McGraw-Hill Education.
- Pallant, J. (2013). *SPSS Survival Manual. A step by step guide to data analysis using SPSS* (4th Ed.). Australia: Allen & Unwin.
- Pan, Y.C., & Roeber, C. (2016). Consequences of test use: a case study of employers' voice on the social impact of English certification exit requirements in Taiwan. *Language testing in Asia*, 6(6), 1-21.
- Popham, W. J. (1987). The Merits of Measurement-Driven Instruction. *The Phi Delta Kappan*, 68(9), 679-682.
- Razavipour, K., Moosavinia, S. R., & Atayi, S. (2018). Construct Ambiguity and Test Difficulty Generate Negative Washback: The Case of Admission Test of English Literature to Graduate Programs in Iran. *International Journal of Instruction*, 11(4), 717-732.
- Rezai, A., Alibakhshi, G., Farokhipour, S. et al. (2021). A phenomenographic study on language assessment literacy: hearing from Iranian university teachers. *Language Testing in Asia*, 11(26), 1-25.
- Shepard, L. A. (1993). Evaluating test validity. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 19, pp. 405-450). Washington, DC: American Educational Research Association.
- Sjostrom, B., & Dahlgren, L. O. (2002). Nursing theory and concept development or analysis: Applying phenomenography in nursing research. *Journal of Advanced Nursing*, 40(3), 339-345.
- Sun, Y. (2016). *Context, construct, and consequences: Washback of the College English Test in China*. Unpublished PhD dissertation, Queens University.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (Eds.), *Input and second language acquisition*, (pp. 235-253). Newbury House.
- Tight, M. (2015). Phenomenography: the development and application of an innovative research design in higher education research. *International Journal of Social Research Methodology*, 19(3), 319-338.
- Vernon, P. E. (1956). *The measurement of abilities*. University of London Press.
- Wu, X., & Zhang, L.J. (2019). The Sociocultural Activity of High Stakes Standardised Language Testing: TOEIC Washback in a South Korean Context. *Language Assessment Quarterly*, 16(3), 369-372
- Xie, Q. (2010). *Test design and use, preparation, and performance: A structural equation modeling study of consequential validity*. Unpublished doctoral dissertation, the University of Hong Kong, Hong Kong.