

نقدی بر آزمون‌های کنکور کارشناسی ارشد رشته آموزش زبان انگلیسی

حسین فرهادی

دانشیار دانشگاه علم و صنعت ایران

چکیده

امروزه آزمون‌ها به عنوان یکی از قوی‌ترین و در عین حال خطرناک‌ترین وسیله اعمال قدرت در سطوح مختلف آموزش و پرورش مورد استفاده قرار می‌گیرند. موارد متعدد استفاده از آزمون‌ها در تصمیم‌گیری‌های مختلف از قبیل سنجش دانش و معلومات افراد، سنجش توانایی افراد در انجام وظایف محوله در آینده، اعطای گواهینامه‌های مختلف، انتخاب افراد صالح برای شغل‌های متفاوت و بالاخره صدور مجوز برای ادامه تحصیل دانش‌آموزان و دانشجویان، مسئله آزمون و آزمون‌سازی را به صورت یک پدیده پیچیده در آموزش و پرورش امروز در آورده است.

یکی از مهم‌ترین کاربردهای آزمون‌ها، گزینش دانشجو برای مقاطع مختلف تحصیلی است. این کار مسئولیت خطیری است که هر چه در سطوح بالاتر انجام گیرد، از اهمیت و پیچیدگی بیشتری هم برخوردار می‌شود. اهمیت ویژه این امر به دلیل سرمایه‌گذاری عظیم مالی و فکری بر روی افراد گزینش شده و پیچیدگی بیشتر آن، به دلیل گستردگی مطالب علمی مورد سنجش می‌باشد. به دلیل این گستردگی و پیچیدگی، این مقاله به یکی از صدها رشته آموزشی محدود شده است. هدف این مقاله، به طور اخص بررسی دقیق آزمون‌های زبان انگلیسی در مقطع کارشناسی ارشد و در محدوده سه سؤال زیر می‌باشد.

۱- محتوای آزمون‌های کارشناسی ارشد رشته زبان انگلیسی باید چگونه باشد؟
پاسخ به این سؤال از طریق تجزیه و تحلیل دقیق مطالب درسی مقاطع پایین‌تر
امکان پذیر است.

۲- محتوای آزمون‌های کارشناسی ارشد رشته زبان انگلیسی موجود چگونه است؟
پاسخ به این سؤال نیاز به تجزیه و تحلیل آزمون‌های چند سال اخیر دارد.
۳- چگونه باید بین آنچه که هست و آنچه که باید باشد، ارتباط منطقی ایجاد کرد؟
در پاسخ به این سؤال راهکارهای عملی جهت تهیه و تدوین آزمون‌های کارشناسی
ارشد که از اصول علم آزمون‌سازی بهره می‌گیرد، ارائه خواهد شد.

واژه‌های کلیدی

سنجش آموزش، آزمون‌های جزء به جزء، آزمون‌های کل، پایایی، روایی، اهداف آموزشی

مقدمه

آموزش و سنجش دو پدیده‌ای هستند که با یکدیگر در ارتباط تنگاتنگ بوده و بر همدیگر
اثرات مستقیمی دارند. تغییر در شیوه‌ها، مطالب، و معیارهای آموزشی بر نحوه سنجش
مطالب فرا گرفته شده اثر می‌گذارد. به همین ترتیب، تغییر در شیوه‌های سنجش نیز اثرات
مستقیمی بر فرآیند آموزش دارد. بنابراین، لازمه یک نظام آموزشی موفق که اهداف از
پیش تعیین شده‌ای را دنبال می‌کند، وجود یک نظام سنجش است که بتواند پاسخ‌گوی
فعالیت‌های آموزشی نظام بوده، نکات قوت و ضعف آن را با شیوه‌های علمی روشن کرده و
به متولیان مربوطه راهکارهای مناسب ارائه دهد.

به دلیل گستردگی موضوع، این مقاله به بحث در مورد کیفیت آزمون‌های رشته زبان
انگلیسی در سطح کارشناسی ارشد محدود شده است. البته علی‌رغم محدودیت دامنه بحث،
شیوه‌های پیشنهادی می‌تواند در رشته‌های دیگر نیز کاربرد داشته باشد. بنابراین محورهای
بحث این مقاله مشخصاً عبارتند از:

۱- بررسی شیوه‌های آموزش و سنجش زبان انگلیسی در جهان و مقایسه آن با
وضعیت موجود در کشور.

نقدی بر آزمون‌های کنکور کارشناسی ارشد... ۸۱

۲- نقدی بر آزمون‌های کارشناسی ارشد رشته آموزش زبان انگلیسی در کشور.

۳- آرایه راهکارهای مناسب جهت بهبود کیفیت سنجش زبان انگلیسی در این رشته.

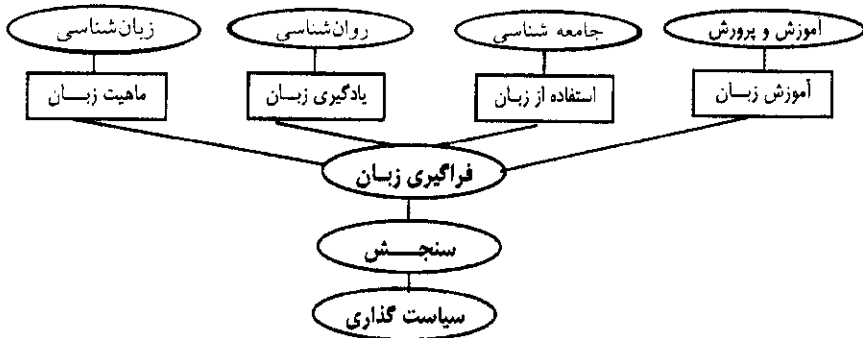
۱- شیوه‌های آموزش و سنجش زبان انگلیسی

آموزش زبان به صورت غیر علمی و سنتی حدود ۳۰۰۰ سال تاریخ مکتوب دارد (کلی^۱، ۱۹۷۵) طبیعی است که همراه با آموزش غیر علمی، شیوه‌های غیر علمی سنجش نیز متداول بوده است.

با آغاز حاکمیت شیوه‌های علمی بر آموزش به طور اعم و آموزش زبان به طور اخص، نحوه آموزش و سنجش زبان از اوایل قرن بیستم دستخوش تغییرات عمده‌ای شد. در این مدت در زمینه آموزش و سنجش زبان تئوری‌های مختلفی ارائه گردید، رشد کرد، و حتی برخی از آنها به فراموشی سپرده شد. از آنجا که آگاهی از این تحولات، وضعیت موجود را تا اندازه‌ای تبیین می‌کند، شمه‌ای از آن به اختصار بیان می‌گردد.

به طور کلی، آموزش و سنجش زبان متأثر از تئوری‌های آموزش و سنجش است. یعنی زبان از طریق شیوه‌ای مناسب بر اساس تئوری خاصی تدریس و یادگیری آن نیز از طریق آزمونی، که آن هم بر تئوری خاصی استوار است سنجیده می‌شود. این فرآیند به ظاهر ساده از پیچیدگی‌های خاصی بر خوردار است، چون شاخه‌های متفاوت علم از جمله زبان‌شناسی، روان‌شناسی، آموزش و پرورش، جامعه‌شناسی و سنجش و اندازه‌گیری هر کدام به نوعی در تعیین و تبیین چگونگی فرآیند آموزش و سنجش زبان تأثیر مستقیم یا غیر مستقیم دارند. نمودار شماره ۱ ارتباط زمینه‌های مختلف علمی را در آموزش و سنجش زبان نشان می‌دهد.

نمودار شماره ۱: مؤلفه‌های آموزش زبان انگلیسی

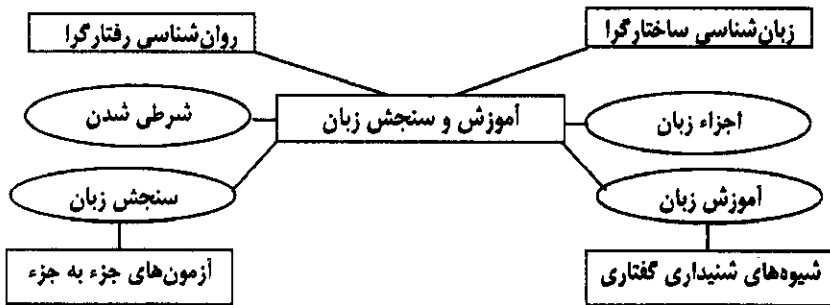


روند توسعه علمی در زمینه آموزش زبان و علم زبان‌شناسی نشانگر آن است که در دهه‌های ۱۹۴۰ و ۱۹۵۰ تئوری حاکم بر زبان‌شناسی، تئوری ساختارگرایی بود. بر اساس این نظریه، زبان مجموعه‌ای از الگوهای ساختاری است که از اصوات، کلمات، و جملات متفاوتی تشکیل می‌شود که تک به تک قابل بررسی و مطالعه است. مجموعه این اجزا به صورت مهارت‌های گفتاری و نوشتاری نمود پیدا می‌کند که شامل گوش کردن، گفتن، خواندن، و نوشتن است. تدریس این اجزاء، زبان‌آموز را به سوی کسب مهارت‌های مختلف سوق می‌دهد که در نهایت آن را توانایی زبان می‌نامیم.

از طرف دیگر، تبیین فرآیند یادگیری در دهه‌های فوق‌الذکر، تحت سلطه تئوری رفتارگرایی بود. این نگرش یادگیری را تغییر در رفتار انسان تعریف می‌کرد و این تغییر رفتار در اثر تقلید، تکرار و حفظ کردن صورت می‌گرفت. بر اساس تئوری‌های زبان‌شناسی و روان‌شناسی، متخصصان آموزش زبان جهت انتقال دانش زبان به فراگیرندگان روشی را اتخاذ کردند که معروف به شیوه گفتاری-شنیداری^۱ است. در این شیوه، اجزاء زبان می‌بایست به زبان آموزان ارائه شود تا آنها این اجزاء را تقلید و آن قدر تکرار کنند تا به صورت مجموعه‌ای از محفوظات در ذهنشان جای گیرد و در موارد ضروری به طور خودکار با استفاده از اجزاء صوتی، واژگانی، و جمله‌ای مقاصد معنی‌دار خود را بیان کنند.

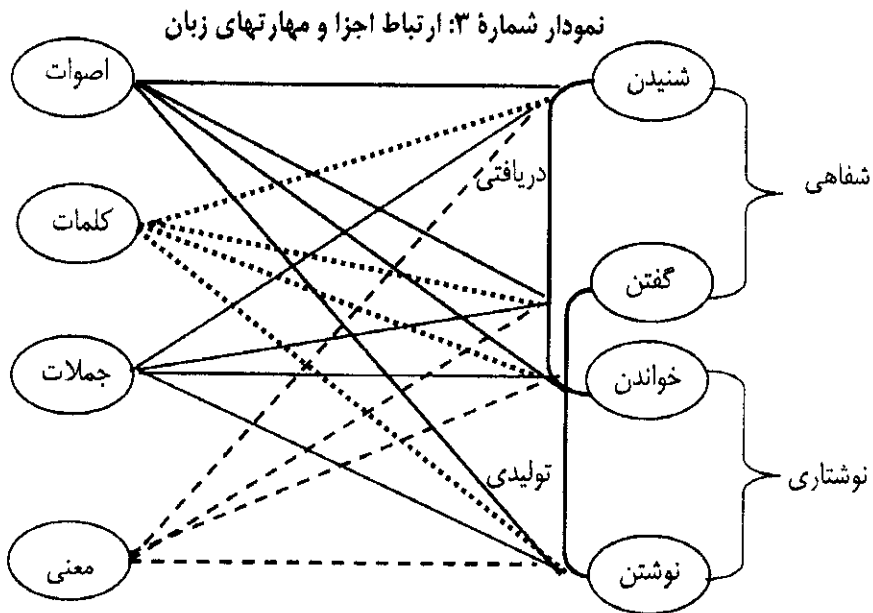
آموزش به شیوه علمی، سنجش به شیوه علمی را می‌طلبید. با توجه به اینکه زبان‌شناسان وقت، زبان را به صورت مجموعه‌ای از اجزاء مختلف توصیف کرده بودند، آزمون‌ها نیز بر اساس اجزاء تعیین شده که شامل اصوات، واژگان، و در نهایت جمله می‌بود طراحی شد. این شیوه آزمون‌سازی به نام آزمون تفکیکی^۱ معروف است. نمودار شماره ۲ چگونگی اولین تحول علمی را در آموزش و سنجش زبان نشان می‌دهد.

نمودار شماره ۲: اولین تحول علمی در آموزش و سنجش زبان



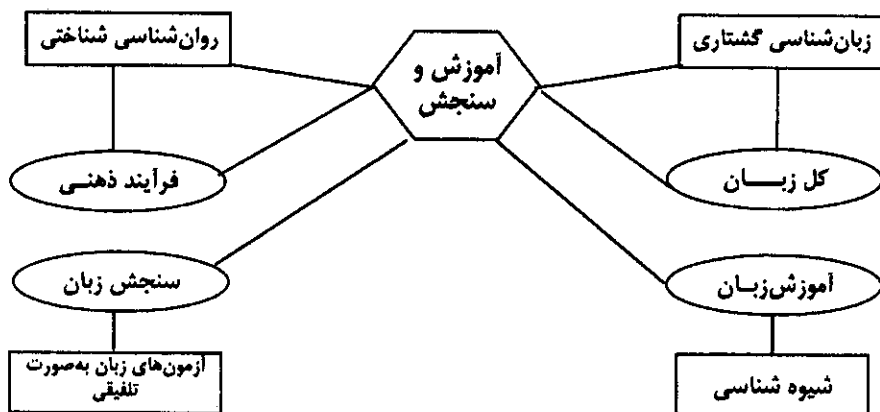
این شیوه آموزش و سنجش به اصطلاح علمی زبان، بیش از یک ربع قرن بر دنیا حکومت کرد، ولی با تغییراتی که در تئوری‌های زبان‌شناسی و روان‌شناسی به وجود آمد، آموزش زبان و پیامد آن سنجش زبان نیز دچار تغییر و تحول شد. زبان‌شناسی تحت تأثیر زبان‌شناسی گشتاری-زایشی، روان‌شناسی تحت تأثیر روان‌شناسی شناختی پایه‌های تئوری زبان‌شناسی ساختارگرا و روان‌شناسی رفتارگرا را متزلزل کردند. بر اساس این تئوری‌های جدید، زبان به جای مجموعه‌ای از اجزاء، به عنوان یک کل شناخته می‌شد که جدا کردن اجزای آن به آسانی ممکن نبود. از طرف دیگر، یادگیری نتیجه تقلید و تکرار و تمرین محسوب نمی‌شد، بلکه از آن به عنوان یک فرآیند ذهنی پیچیده با مؤلفه‌های مختلف یاد می‌شد. با این تغییرات، شیوه‌های آموزش زبان و سنجش زبان نیز دگرگون گشت.

این گروه از دانشمندان اعتقاد داشتند که به دلیل تبلور اجزای زبان در مهارت‌های آن، تسلط به مهارت‌ها به ناچار تسلط بر اجزاء را نیز در برمی‌گیرد. با توجه به ارتباط بسیار نزدیکی که در نمودار شماره ۳ مشخص شده است، جدا کردن اجزاء و حتی مهارت‌ها از یکدیگر نه تنها کار آسانی نبود، بلکه خلاف اصول علم زبان‌شناسی نیز به حساب می‌آمد.



ارتباطات بسیار نزدیک بین اجزاء و مهارت‌ها، نظریهٔ کل بودن زبان را بیشتر تأیید می‌کرد، در نتیجهٔ استفاده از شیوه‌های سنتی سنجش زبان به صورت تفکیکی نیز کم‌کم مورد سؤال قرار گرفت و به جای آن شیوهٔ سنجش تلفیقی (Integrative) مرسوم گردید. آزمون‌های بسته، دیکته، مصاحبه و نگارش نمونه‌هایی از این نوع آزمون‌ها که از بین آنها، مصاحبه و نگارش به دلیل مشکلات اجرایی به طور عام مرسوم نشد، ولی آزمون‌های بسته و دیکته جای خود را در عرصه سنجش زبان در سی سال گذشته به خوبی باز کرده و تحقیقات وسیعی نیز در این زمینه به عمل آمده است. نمودار شماره ۴ دومین تحول علمی را در زمینهٔ آموزش و سنجش زبان نشان می‌دهد.

نمودار شماره ۴ : دومین تحول علمی در آموزش و سنجش زبان



اواخر دهه ۷۰ و اوایل دهه ۸۰ اوج استفاده از این شیوه سنجش بود، تا جایی که آزمون‌های جزء به جزء زبان تقریباً منسوخ شد و به جای آنها آزمون‌هایی که اجزاء متفاوت را یکجا و در بافت واقعی استفاده از زبان مورد سنجش قرار می‌داد، معمول گشت. این تحول، پیشرفت محسوسی را در زمینه سنجش به وجود آورد و با تأثیرپذیری از تحولات آموزش زبان برای ایجاد ارتباط علم سنجش زبان نیز عوامل ارتباطی را مورد توجه قرار داد و آزمون‌هایی طراحی شدند که زبان و مشخصه‌های ارتباطی را مورد سنجش قرار می‌داد. از آغاز دهه ۹۰ این نوع آزمون‌ها که هدفمندتر از آزمون‌های گذشته بود، به سرعت رشد کرد. از نمونه‌های مختلف این نوع آزمون‌ها در انگلستان، استرالیا و کانادا تست‌های

International English Language Testing Service	(IELTS)
of English for Speakers of Other Languages	(OTESOL)
Ontario Test	
Carleton Academic English Language Assessment	(CAELA)
Test of English for Educational Purposes	(TEEP)
Canadian Test for Science and Technology	(CanTEST)

را می‌توان نام برد. متولیان این تست‌ها معتقدند که نتایج حاصله از این آزمون‌ها، به مراتب بهتر از آزمون‌های سنتی نظیر (TOEFL) Test of English as a Foreign Language میزان توانایی واقعی آزمون‌دهندگان را مشخص می‌کند.

در نهایت در تعریف بعضی از مفاهیم آزمون نیز تغییرات بنیادی به وجود آمد. یکی از این تغییرات در مورد مفهوم روایی آزمون است که توسط مزیک (Messick, 1989, 1991) بیان شده است. به اعتقاد وی تعریف سنتی روایی (میزان سنجش تست از آنچه که باید بسنجد) ناقص و نارساست. وی دو نوع روایی را مطرح می‌کند، روایی استنادی (Evidential Validity) و روایی پیامدی (Consequential Validity). نوع اول روایی مجموعه روایی‌های سنتی شامل روایی صوری، روایی محتوایی، روایی معیاری، و روایی سازه‌ای را در برمی‌گیرد. ولی نوع دوم که به نوبه خود تحولی جدید در مفهوم روایی به حساب می‌آید، مربوط به آینده آزمون‌دهندگان می‌شود. این روایی برای اثراتی که آزمون و نتیجه آن در زندگی آینده افراد می‌گذارد، اهمیت خاصی قایل است. روایی پیامدی عبارت از یک سلسله ارزش‌ها است که نمرات آزمون می‌تواند به داوطلبان آزمون نسبت دهد. برای مثال در هر نظام آموزشی، اطلاق عناوینی چون قبولی، مردودی، ضعیف و قوی بودن به داوطلبان، اثرات روانی و اجتماعی متفاوتی بر روی آنان می‌گذارد. بنابراین، آزمون‌سازی باید در مسیری حرکت کند که نتایج حاصل از آزمون‌ها اثرات روحی و روانی منفی بر روی آزمون‌دهندگان نگذارد. با توجه به همین تحولات، بحث اخلاقیات Ethics در آزمون‌سازی به شکلی جدی مطرح شده که در نوشتاری جداگانه مورد بحث قرار گرفته است (فرهادی، ۱۳۷۹).

۲- بررسی آزمون‌های کارشناسی ارشد رشته آموزش زبان

آزمون‌های کارشناسی ارشد مورد بحث در این مقاله مربوط به رشته آموزش زبان انگلیسی است. این آزمون‌ها در دو بخش عمومی و اختصاصی برگزار می‌شود که در قسمت عمومی

با رشته زبان و ادبیات انگلیسی مشترک است. به دلیل ویژگی‌های متفاوت این دو بخش، هر یک به طور جداگانه مورد بحث قرار می‌گیرد.

۱-۲ بخش آزمون‌های عمومی

آزمون زبان عمومی شامل سه قسمت دستور زبان، واژگان و خواندن و درک متن است که در هر سه قسمت نوع پرسش‌ها چهار گزینه ایست. بررسی اجمالی آزمون‌های سال‌های ۷۰ تا ۷۶ نشان می‌دهد که به طور متوسط به قسمت‌های واژگان و دستور زبان انگلیسی هر کدام ۴۰ سؤال و به قسمت خواندن و درک متن ۲۰ سؤال اختصاص داده شده است. البته در آزمون بعضی از سال‌ها اختلافات جزئی در تعداد سؤال‌ها وجود دارد که تأثیر چندانی بر کل مجموعه نمی‌گذارد.

همان‌طور که قبلاً اشاره شد، از آنجا که نتیجه یک آزمون در روند زندگی افراد تأثیر بسزایی دارد، باید مطالبی که آزمون بر اساس آن تهیه می‌شود به دقت مورد بررسی قرار گرفته و نکات مهم مطالب تدریس شده به طور منطقی طبقه‌بندی و به نسبت در آزمون‌ها گنجانده شوند. در مورد تعیین محتوای آزمون‌های عمومی زبان دو شیوه معمول است. در شیوه اول تعیین محتوای آزمون‌ها براساس مطالب تدریس شده در دوره‌های پیشین آموزشی است. و اگر این شیوه در تعیین محتوای آزمون عمومی رشته آموزش زبان به کار گرفته شود باید محتوای آزمون بر اساس مطالب تدریس شده در دوره‌های کارشناسی تربیت دبیری، مترجمی، و زبان و ادبیات انگلیسی تهیه شود.

شیوه دیگر این است که محتوای آزمون‌های عمومی زبان بر اساس معیارهای از پیش تعیین شده برای سنجش توانایی زبان انگلیسی طراحی شود. این مقوله را می‌توان مثل آزمون TOEFL تعریف کرد که ارتباط مستقیمی با مطالب قبلی نداشته و به عنوان آزمون عمومی توانایی زبان تهیه و به کار برده می‌شود.

آزمون‌های عمومی کارشناسی ارشد موجود با هیچ‌یک از شیوه‌های فوق مطابقت ندارد و علی‌رغم زحمات مسئولان سازمان سنجش آموزش کشور نارسایی‌های زیادی در این

آزمون‌ها به چشم می‌خورد. البته دلیل بعضی از این نارسایی‌ها را می‌توان در مشکلات اجرایی و وسعت برگزاری آزمون‌ها جستجو کرد. ولی بعضی دیگر مربوط به عدم استفاده صحیح از یافته‌های علمی در زمینه‌های آزمون‌سازی است که با کمی توجه و دقت اگر کلاً قابل رفع نباشد، به طور محسوسی قابل کنترل است. در اینجا به اختصار چند مورد ذکر می‌شود.

بررسی دقیق سؤالات قسمت دستور زبان نشان می‌دهد که نه تنها برای تهیه این سؤالات مطالب دوره کارشناسی رشته‌های مربوطه مورد توجه قرار نگرفته است، بلکه بسیاری از سؤالات، مشابه سؤالات آزمون سراسری دانشگاه‌ها در مقطع کارشناسی است. به نمونه‌های زیر توجه کنید:

آزمون ورودی سال ۷۱-۷۰:

10. Old people remember the past and young people dream ----- the future.

- a. at b. of c. to d. with

این سؤال حرف اضافه of را می‌سنجد که بطور کلیشه‌ای با فعل dream به کار می‌رود.

40. Ali: You look tired today. Bahram: I feel tired too. I wish I ----- to bed earlier last night.

- a. had gone b. have gone c. did go d. were gone

این سؤال نیز عبارت کلیشه‌ای wish را می‌سنجد که جواب آن به عبارت معروف، فرمولی است.

آزمون ورودی سال ۷۳-۷۲

54. The examiner made us ----- our identification in order to be admitted to the test center.

- a. show b. showed c. showing d. to show

این سؤال نیز طرز استفاده از make را می‌سنجد که بعد از آن فعل به شکل مصدر بدون to می‌آید.

آزمون ورودی سال ۷۴-۷۳

۳۱. He never has time for his family because there are so many clubs that he belongs -----.

- a. for b. from c. of d. to

این سؤال حرف اضافه مربوط به فعل belong را می‌سنجد که مطلبی کاملاً ازبرکردنی است.

۳۵. No matter how many people like it, I still do not care ----- ice-cream.

- a. for b. of c. to d. with

این سؤال عبارت کلیشه‌ای care for را می‌سنجد که کاملاً محاوره‌ای است.

علاوه بر اشکالات فنی مربوط به علم آزمون‌سازی که در این نمونه‌ها مشهود است، با بررسی نه چندان عمیقی معلوم می‌شود که مطالب آزمون ارتباط چندانی با مطالب تدریس شده در دوره چهار ساله کارشناسی زبان ندارد. تمام این موارد در سطح مطالب دبیرستان است. باید متذکر شد که چون اطلاعات آماری در مورد این سؤالات در دسترس نیست، تمام مطالب ذکر شده بر اساس تجربه و باورهای نگارنده است و چه بسا که این سؤالات در مجموعه آزمون بسیار خوب عمل کرده باشد، که اگر این چنین باشد، در آن صورت فاجعه عمیق‌تر از حد تصور است و ضمن تأسف خوردن به سطح دانش فارغ‌التحصیلان رشته زبان در سطح کارشناسی، باید فکر اساسی برای تغییرات بنیادی در برنامه‌های کارشناسی مربوط به رشته زبان به عمل آورد.

بحث در مورد کیفیت این سؤالات از جنبه میزان دشواری، پیچیدگی، و یا سادگی بسیار حایز اهمیت است. البته ممکن است عده‌ای ادعا کنند که چون آزمون به صورت مسابقه برگزار می‌شود بنابراین میزان سادگی یا دشواری سؤالات تأثیر چندانی بر رتبه داوطلبان نمی‌گذارد. لازم به توضیح است که چنین اعتقادی نه تنها صحیح نیست، بلکه در شرایط حساسی می‌تواند بسیار زیانبار هم باشد. آسان بودن بیش از حد سؤالات برای داوطلبان باسوادتر عادلانه نیست، چون داوطلبان کم‌سواد هم می‌توانند به این سؤالات پاسخ دهند. در نتیجه آزمون قدرت تشخیص داوطلبان با معلومات‌تر را از دست می‌دهد. در

پاسخ دهند. در نتیجه آزمون قدرت تشخیص داوطلبان با معلومات‌تر را از دست می‌دهد. در این صورت روایی و پایایی آزمون به عنوان مؤلفه‌های کیفیت آزمون به شکل جدی زیر سؤال می‌رود.

عده‌ای هم ممکن است ادعا کنند که سؤالات دستور زبان باید از نظر جمله‌بندی آسان و از نظر نکات دستوری روشن و واضح طراحی شود. این ادعا نیز صحت ندارد، چون در این صورت نباید فرقی بین آزمون دستور زبان کلاس دوم راهنمایی و آزمون دستور زبان دورهٔ دکتری زبان وجود داشته باشد. این باور در دهه‌های ۵۰ و ۶۰ یعنی حدود ۳۰ سال پیش درست بود، اما امروزه با تحولات علمی در آزمون‌سازی، این نظریه کاملاً مردود است. چرا باید تصور کنیم که دانشجویان دانشگاه‌ها پس از چهار سال تحصیل در رشتهٔ انگلیسی، خواندن کتاب‌های متفاوت علمی و ادبی و گذراندن حدود ۱۴۰ واحد درسی به زبان انگلیسی، هنوز هم مانند فارغ‌التحصیلان دبیرستان‌ها ایند! آیا نظام آموزش عالی کشور قادر نیست در مدت چهار سال تغییر قابل ملاحظه‌ای در میزان توانایی دانشجویان زبان ایجاد کند؟ اگر پاسخ مثبت است، پس چطور می‌توان ادعا کرد که باید نحوه آزمون دستور زبان در مقاطع مختلف تحصیلی از راهنمایی تا دکتری به بهانهٔ ساده بودن ساختار آزمون‌های دستور زبان یکسان باشد؟

بخش واژگان نیز دست کمی از بخش دستور زبان ندارد. همانقدر که در مورد دستور زبان با ساده اندیشی عمل می‌شود، در بخش واژگان نیز معیار خاصی که حاکم بر انتخاب لغات و گزینه‌ها باشد به چشم نمی‌خورد. از ظاهر سؤالات چنین برمی‌آید که کلمات به طور تصادفی انتخاب و در جمله‌ای که حداقل بافت ممکن زبان را دارد ارائه می‌شود و گزینه‌های درست و غلط هم از روی فرهنگ لغات انتخاب می‌شوند.

درست است که تعداد واژگان یک زبان تقریباً نامحدود است و هر نوع انتخابی تا اندازه‌ای با مسئله شانس، تصادف و سلیقهٔ تدوین‌کنندگان آزمون ارتباط دارد، اما در تمام مراکز آزمون‌سازی دنیا رسم بر این است که محدودهٔ واژگان را با توجه به سطح دانش

داوطلبان تعیین کنند. مثلاً اعلام می‌کنند که برای ورود به دوره کارشناسی ارشد رشته زبان، هر داوطلب باید گنجینه واژگانی معادل ده هزار کلمه داشته باشد. سپس از روی این ده هزار کلمه که لیست‌های متفاوت کامپیوتری آنها تهیه شده است، سؤالات واژگان را طراحی می‌کنند. این حداقل قضیه است. در شرایط مورد نظر در این مقاله باید لیست تمام لغاتی که دانشجویان در چهار سال دانشگاه در کلاس‌های مختلف مطالعه می‌کنند، تهیه شود و از بین آنها افعال، اسامی، صفات و قیود طبقه‌بندی و هر کدام براساس فراوانی خود در آزمون‌ها گنجانده شود. اگر به این مسئله ابتدایی توجه می‌شد، دیگر شاهد سؤالاتی مشابه آنچه که در زیر آمده است، نمی‌بودیم:

آزمون ورودی سال ۷۴-۷۳

5. It is ----- to see a policeman being given a ticket for speeding.

- a. apparent b. ironic c. objective d. subjective

6. He ate a ----- amount of the homemade bread; no matter he is sick.

- a. capricious b. precocious c. prodigious d. copious

مقایسه گزینه‌های سؤالات ۵ و ۶ به وضوح نشان می‌دهد که این گزینه‌ها از هماهنگی لازم برخوردار نیستند. گزینه‌های سؤال ۵ از لغات بسیار ساده و روزمره تشکیل شده، در صورتی که گزینه‌های سؤال ۶ از لغات بسیار مشکل و سطح بالا انتخاب شده‌اند.

آزمون ورودی سال ۷۵-۷۴

33. What ----- can be had by the young widow of the man who was killed in an avoidable car accident.

- a. haven b. hawser c. ocular d. solace

37. We might just as well have stayed at home ----- the enjoyment we had.

- a. as far as b. concerning c. for all d. on account of

در این دو سؤال نیز اختلاف سطح کلمات در گزینه‌ها کاملاً مشهود است. برای آزمایش، این چهار نمونه بین دانشجویان دوره دکتری و فارغ التحصیلان دوره کارشناسی ارشد توزیع شد. جالب اینکه اکثر آنان معانی لغات گزینه‌های سؤالات شماره ۶ و ۳۳ را نمی‌دانستند. پس چگونه از داوطلبان ورود به دوره کارشناسی ارشد باید انتظار داشت که این لغات را بدانند. در ضمن، چه تجانسی بین گزینه‌ها در سؤالات ۵ و ۳۳ از یک طرف و ۶ و ۳۷ از طرف دیگر وجود دارد. کدام یک برای سطح آزمون کارشناسی ارشد زبان مناسب است؟

در بخش مربوط به آزمون خواندن و درک متن نیز ضابطه‌ای برای انتخاب متون و سؤالات به چشم نمی‌خورد. بین درجه دشواری متون هماهنگی لازم وجود ندارد و سؤالات مربوط به آن نیز با دقت و با توجه به اصول آزمون سازی تهیه نشده است. دو سؤال زیر که از آزمون ورودی ۷۱-۷۲ استخراج شده‌اند مبین این مسئله است:

10. According to the passage, what is one main function of the root system?

- a. To absorb light for growth b. To create minerals for nutrition
c. To secure the plant in the ground d. To speed the reproduction process

11. The major food-producing part of a plant is the ----- .

- a. flower b. leaf c. root d. shoot

برای پاسخ دادن به این سؤال‌ها، نیازی به خواندن متن نیست، چون جواب صحیح به طور منطقی در هر دو سؤال گزینه C می‌باشد.

۲-۲ بخش آزمون اختصاصی

بخش آزمون اختصاصی نیز مانند بخش عمومی از سه قسمت تشکیل شده است. این سه قسمت شامل دروس آزمون‌سازی، روش تدریس، و زبان‌شناسی است. که برای هر سه قسمت از سؤالات چهار گزینه‌ای استفاده شده است. در مورد آزمون اختصاصی ذکر سه مورد ضروری به نظر می‌رسد.

نکته اول در مورد آزمون اختصاصی این است که در این آزمون توانایی زبان مد نظر نیست. بلکه معلومات تخصصی داوطلبان مورد سؤال قرار می‌گیرد. بنابراین باید محتوای تست مطابق مطالب تدریس شده در دوره کارشناسی باشد. در دوره کارشناسی حدود ۱۴۰ واحد درسی ارائه می‌گردد که از این ۱۴۰ واحد، حدود ۸۰ واحد عمومی و ۶۰ واحد اختصاصی است. از این ۶۰ واحد اختصاصی که شامل درس‌های متعددی است، آزمون کارشناسی‌ارشد فقط بر اساس ۱۰ واحد مربوط به سه درس آزمون‌سازی، زبان‌شناسی و روش تدریس طراحی و تدوین می‌گردد.

به نظر من، هیچ دلیل قانع‌کننده‌ای برای محدود کردن امتحان کارشناسی‌ارشد به این سه درس وجود ندارد. آزمون‌سازی، روش تدریس و زبان‌شناسی نه تنها مزیت خاصی نسبت به درس‌هایی چون ترجمه‌ها، روش تحقیق‌ها، ادبیات‌ها و غیره که در دوره کارشناسی تدریس می‌شوند، ندارند بلکه این شیوه تعیین مطالب آزمون، اهمیت درس‌های دیگر را در دوره کارشناسی کم رنگ می‌کند و دانشجویان برای درسهایی که جزء آزمون کارشناسی‌ارشد نیستند، اهمیت زیادی قایل نمی‌شوند و شاید هم عدم توجه به آنها در آزمون کارشناسی‌ارشد، باعث افت تحصیلی دانشجویان در این درس‌ها بشود.

به علاوه، یکی از مطالب بسیار مهم دنیای آموزش امروز اثر آزمون بر مطالب درسی و شیوه‌های تدریس است که به washback معروف است. بدین معنی که آموزش در هر مقطعی متأثر از محتوای آزمون است. بنابراین حذف بسیار آشکار حدود ۵۰ واحد از درس‌های اختصاصی دوره کارشناسی در آزمون کارشناسی‌ارشد اثرات نامطلوبی بر مطالب، شیوه آموزش، و میزان توجه دانشجویان به درس‌های مذکور می‌گذارد.

نکته دوم در مورد آزمون‌های کارشناسی‌ارشد، منابعی است که سؤالات از آنها استخراج می‌شود. طبق اصول آزمون‌سازی، داوطلبان باید به طور کامل از منابع و مآخذ سؤالات امتحانی اطلاع داشته باشند. معمولاً سؤالات یا از کتاب خاصی انتخاب می‌شوند که داوطلبان باید مطالب آن کتاب را به صورت بسیار ریز مطالعه و حفظ کنند و یا اینکه

سؤالات، مفاهیم رشته‌مربوطه را می‌سنجد که در این صورت با مطالب ریز یک کتاب خاص ارتباطی ندارد. در هر دو حالت داوطلبان باید از منابع و مآخذ سؤالات آگاهی کامل داشته باشند.

گاهی مشاهده می‌شود که برخی از سؤالات به دلایلی عیناً از کتاب خاصی استخراج می‌شود. اگر داوطلب کم‌شانسی به طور تصادفی آن کتاب خاص را مطالعه نکرده باشد، علی‌رغم داشتن اطلاعات کافی، از پاسخ دادن به آن سؤالات عاجز خواهد بود. نمونه‌های زیر این مورد را تأیید می‌کند.

آزمون ورودی سال ۷۲

117. An EFL approach which does not tolerate any learner errors is ----- oriented.

- a. cognitively b. structurally c. traditionally d. transformationally

(Yule, 1988:129)

119. The approach used by fluent readers is normally -----.

- a. data driven b. conceptually driven c. interactive d. linguistic

(Yule, 1988:92)

آزمون ورودی سال ۷۴

In the communicative approach, students try to state a sentence in different ways. This is because -----.

- a. one might not remember a certain structure at a certain time
b. one function can have many different linguistic forms
c. the students have to expand their vocabulary
d. the students need much practice in learning a structure

(Freeman, 1986:128)

آزمون ورودی سال ۷۵

113. The Community Language Learning takes its principles from ----- .

- a. Counselling learning b. Communication theory
c. Cognitive approach of learning d. Communicative language learning

(Freeman, 1986:89)

112. In bottom-up processing of information, ----- .

- a. the concepts we have in mind help us interpret more specific information
- b. thought moves from specific to more general information
- c. thought operates from a knowledge base to work on specific information
- d. the reader activates his prior knowledge to recreate the author's meaning

(Chastain, 1988:36)

117. In top-down processing of information, -----.

- a. thought moves from specific pieces of information to more general knowledge
- b. thought operates from a knowledge base to work on pieces of information
- c. the information we receive help us interpret the concepts we have in our mind
- d. the reader uses specific information to receive the author's meaning

(Chastain, 1988:36)

برای پاسخ دادن به این سؤالات فقط آشنایی با جملات و یا حفظ کردن آنها مرتبط با کتاب‌های مرجعی که برای هر سؤال آورده شده است، کفایت می‌کند. این سؤالات نشان می‌دهد که طراحان جمله‌ای را از کتابی به عنوان سؤال انتخاب کرده و کلمه‌ای از آن را حذف و گزینه‌ها را بر اساس آن کلمه تهیه نموده‌اند. این نوع سؤالات، داوطلبان را به سمت حفظ کردن مطالب کتاب سوق می‌دهد. اگر قرار است که کتاب خاصی مورد استفاده طراحان سؤال قرار گیرد و مطالب آن کتاب از اول تا آخر توسط داوطلبان حفظ شود باید همین شیوه به طور دقیق به اطلاع داوطلبان برسد، اما اگر هدف آزمون سنجش معلومات داوطلبان به صورت مفهومی است و کتاب خاصی مورد نظر نیست، در این صورت سؤالات باید قائم به مفهوم باشد، نه قائم به جمله‌ای از یک کتاب. حتی با کمی دقت معلوم می‌شود که سؤالات شماره ۱۱۲ و ۱۱۷ از یک صفحه کتاب مذکور انتخاب شده‌اند.

نکته سوم در مورد این آزمون‌ها نوع سؤالات است. با توجه به مطالبی که در مقدمه ذکر شد. در واقع نظام آزمون‌سازی کارشناسی ارشد هنوز از اصول دهه‌های ۶۰ و ۷۰ تبعیت می‌کند که سؤالات بر اساس اجزای ریز و حفظ کردنی استوار است. اکثر قریب به اتفاق

دانشمندان بر این باورند؛ سؤالاتی که اجزای بسیار کم اهمیت دانش و معلومات را می‌سنجند، برای آزمون‌های سطح بالا مناسب نیستند. با این نوع سؤال‌ها قدرت تفکر، خلاقیت، توانایی‌های شناختی و فراشناختی داوطلبان سنجیده نمی‌شود.

البته مشکلات و محدودیت‌های اجرایی یک آزمون سراسری را نباید نادیده گرفت و ادعای آرمان‌گرایی هم در شرایط فعلی صحیح نیست. ولی به اعتقاد من، با همین شرایط موجود، با همین توان علمی موجود در دانشگاه‌ها، با همین علاقه موجود بین مسئولان سازمان سنجش آموزش کشور، و با توجه به تعداد قابل توجه داوطلبان، نظام سنجش در کارشناسی‌ارشد می‌تواند به نحو چشم‌گیری بهبود یابد. در قسمت پایانی این مقاله راه کارهایی که به نظر عملی می‌آید، پیشنهاد می‌گردد.

۳- راهکارهای عملی برای بهبود کیفیت آزمون کارشناسی‌ارشد رشته آموزش زبان انگلیسی با توجه به آنچه که در دو بخش گذشته ذکر شد، می‌توان نتیجه گرفت که کیفیت آزمون‌های کارشناسی‌ارشد رشته آموزش زبان در حد مطلوب نیست. نارسایی‌های موجود را می‌توان به ترتیب زیر فهرست کرد:

۱- در آزمون عمومی و در آزمون اختصاصی، رابطه‌ای بین محتوای دروس دوره کارشناسی و محتوای آزمون وجود ندارد.

۲- هم در آزمون عمومی و هم در آزمون اختصاصی از شیوه سنتی، یعنی سنجش اجزاء کوچک و بعضاً کم اهمیت، استفاده می‌شود.

۳- صنعت آزمون‌سازی از پیشرفت‌های علمی بهره‌چندانی نبرده است.

۴- آزمون‌ها توانایی‌های پیچیده و سطح بالا را نمی‌توانند بسنجند.

۵- سؤالات آزمون‌های عمومی و اختصاصی بسیار کوتاه و در بعضی موارد به گونه‌ای طراحی شده‌اند که فقط می‌تواند مسئله مورد نظر را بسنجد.

۶- محتوای سؤال‌ها در آزمون اختصاصی بیشتر در جهت سنجش محفوظات است، تا معلومات.

۷- در آزمون‌های اختصاصی بعضی از سؤالات قائم به یک کتاب خاص می باشند.
۸- دو قسمت از آزمون عمومی، یعنی دستور زبان و واژگان، خارج از بافت واقعی زبان
سنجیده می‌شود و در قسمت سوم که مربوط به آزمون خواندن و درک متن است، به
سنجش مهارت‌های خواندن توجه نمی‌شود.

۹- در آزمون عمومی برای انتخاب واژگان از یک روش نظام مند استفاده نمی‌شود.

۱۰- به مهارت‌های متفاوت زبان در امتحان عمومی توجهی نمی‌شود.

برای مرتفع ساختن بعضی از این نارسایی‌ها، در چندین جهت می‌توان اقداماتی به
عمل آورد و چون سنجش فرآیندی طولانی دارد، خیلی از اقدامات نمی‌تواند نتایج فوری در
بر داشته باشد. بنابراین می‌توان راهکارها را به دو دسته کوتاه مدت و بلند مدت
تقسیم‌بندی کرد. راهکارهای کوتاه مدت را بدون فوت وقت می‌شود انجام داد، ولی برای
راهکارهای بلند مدت باید برنامه‌ریزی‌های پژوهشی لازم به عمل آید.

البته باید اعتراف کرد که در چنین شرایطی، مسئولیت سازمان سنجش آموزش کشور
بسیار سنگین بوده و انجام آزمون سراسری مطلوب کاری بسیار دشوار است. باید از همین
جا به مسئولان سنجش آموزش کشور تبریک گفت که این وظیفه خطیر را با توجه به
امکانات و شرایط موجود به نحو احسن انجام می‌دهند. پیشنهادهایی که در این مقاله ارائه
می‌شود، فقط بر اساس دانش و علم سنجش و اندازه‌گیری است و باید توجه داشت که در
آزمون‌سازی فاصله زیادی بین تجمل‌گرایی نظری و استیصال عملی وجود دارد و هر گونه
افراط و تفریط، حاصلی جز ضرر و زیان نخواهد داشت. بنابراین انتظار می‌رود که
پیشنهادها تا آنجایی که ممکن است در قالب امکانات از قوه به فعل درآید. پیشنهادها را بر
اساس نارسایی‌هایی که قبلاً ذکر شد، می‌توان بر سه محور خلاصه کرد:

۱- تعیین محتوای آزمون‌ها

۲- تعیین نوع سؤالات

۳- تعیین شیوه‌های تفسیر نمرات داوطلبان

اگر این مقاله بتواند در ارایهٔ راهکارهای عملی جهت بهبود آزمون‌ها در این سه محدوده موفقیتی کسب کند، هدف آن تا اندازه‌ای بر آورده شده است.

۱-۳- تعیین محتوای آزمون‌ها

یکی از ارکان آزمون‌سازی و آزمون‌شناسی، همان‌طور که قبلاً نیز اشاره شد، روایی آزمون‌ها است، که یکی از ارکان روایی آزمون محتوای آن است. بدین معنی که اولاً محتوای آزمون باید با محتوای آنچه که باید سنجیده شود رابطهٔ تنگاتنگ و مستقیمی داشته باشد. ثانیاً محتوای آزمون باید آیینۀ تمام‌نمای اهداف آموزشی مقطع مورد نظر باشد. ثالثاً، منابع و مآخذ محتوای آزمون باید در کمال صراحت و صداقت به اطلاع کلیه داوطلبان برسد. جهت روشن شدن مطلب برای هر مورد توضیح مختصری ارائه می‌شود.

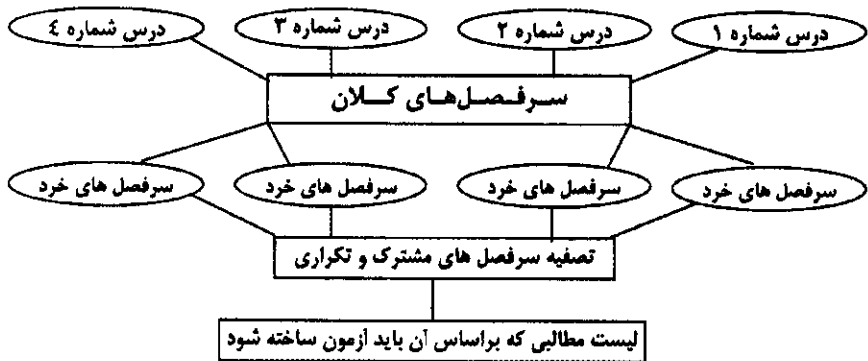
بدون شک نتایج آزمون کارشناسی‌ارشد اثرات غیر قابل‌انکاری بر نظام آموزشی آیندهٔ کشور دارد، چرا که اکثر فارغ‌التحصیلان این دوره به کادر مدرسان دانشگاه‌های کشور می‌پیوندند. عدهٔ دیگر نیز وارد دوره‌های دکتری داخل یا خارج کشور می‌شوند. بنابراین دقت در تهیهٔ این آزمون‌ها هر چقدر هم که هزینه‌بر و وقت‌گیر باشد، ارزشمند است. پس لازم است حداقل اصول علمی در تهیهٔ این آزمون‌ها رعایت شود.

اولین قدم در تعیین محتوای آزمون کارشناسی‌ارشد رشتهٔ آموزش زبان انگلیسی، بررسی دقیق مطالب مصوب و تدریس شده در دوره‌های کارشناسی آموزش زبان، ادبیات، مترجمی، و تربیت دبیری است. مطالب این دوره‌ها هم محتوای آزمون عمومی و هم محتوای آزمون اختصاصی را تا حد قابل‌قبولی تعیین می‌کند.

با بررسی این مطالب، سرفصل‌های کلان و خرد برای هر واحد درسی در دوره‌های فوق تعیین می‌شود. نتیجهٔ این بررسی ایجاد منبعی غنی از مطالب تدریس شده در این دوره‌ها است. با توجه به وجوه اشتراک و مطالب احتمالاً تکراری در درس‌های مختلف، می‌توان همهٔ سرفصل‌ها را از یک صافی گذراند و فهرستی از سرفصل‌های تصفیه شده را تهیه نمود. نمودار شماره ۵ فرآیند این بررسی را نشان می‌دهد.

پس از تعیین مطالب، مسلماً تعداد عناوین به دست آمده به مراتب بیشتر از تعداد سؤالاتی خواهد بود که امکانات اجرایی یک آزمون سراسری اجازه می‌دهد. بنابراین مسئله انتخاب عناوین یا آنچه که در آزمون‌سازی به تهیه جدول بودجه‌بندی (Table of Specifications) معروف است پیش می‌آید. با توجه به اهمیت آزمون کارشناسی ارشد، باید در تهیه جدول به مفهوم مشمولیت (Inclusiveness) توجه شود. بدین معنی که اگر مثلاً برای چهار مورد از مطالب تعیین شده، فقط یک سؤال می‌تواند در آزمون ارائه گردد، باید موردی را انتخاب کرد که دانستن آن لازمه‌اش دانستن موارد دیگر نیز باشد. مثلاً در آزمون عمومی، اختصاص دادن یک سؤال برای سنجیدن توانایی داوطلبان در مورد فقط یک حرف اضافه به نظر منطقی نمی‌آید، بلکه می‌توان همان حرف اضافه را در بافت پیچیده زبان سنجید. یا اینکه موردی را در سؤال گنجانند که دانستن آن توانایی استفاده از حرف اضافه را تضمین بکند. در سؤالات اختصاصی هم به جای محدود کردن یک سؤال به یک مفهوم، که تعریف یک خطی و حفظی دارد، می‌توان مطالبی را عنوان کرد که لازمه‌اش دانستن آن تعریف نیز باشد. البته سؤالات نمونه در قسمت بعدی ارائه خواهد شد.

نمودار شماره ۵: فرآیند بررسی مطالب درسی



آخرین مورد در تعیین مطالب آزمون، رعایت حق داوطلبان است، طبق اصول بسیار مهم آزمون‌سازی، داوطلبان باید بطور جدی در مسیر مطالب و عناوین و مراجع آزمون قرار گیرند. این حق مسلم داوطلبان است که بدانند:

۱- از چه منابعی و به چه نسبتی در آزمون گنجانده می‌شود.

- ۲- نوع سؤالات چند گزینه‌ای است یا تشریحی، مفهومی است.
- ۳- جزئیات بخش‌های آزمون چیست.
- ۴- سؤالات چگونه، در چه شرایطی و با چه ضوابطی تهیه می‌شوند.
- ۵- دانشگاه‌ها از چه موقعیتی برخوردارند و از کدام استاد سود می‌برند تا بتوانند دانشگاه مورد علاقه خود را انتخاب کنند.
- ۶- در شرایط مطلوب، چه از نظر روحی و چه از نظر فیزیکی و چه از نظر مکانی، به رقابت بپردازند.

تمام این حقوق را که به نظر می‌آید کلیه داوطلبان در شرایط فعلی از آن محرومند، می‌توان در یک دفترچه راهنما بطور مرتب و منظم نوشته و به هنگام ثبت نام در اختیار داوطلبان قرار داد. البته باید اذعان داشت که در چند سال اخیر سازمان سنجش آموزش کشور در این زمینه اقداماتی جدی به عمل آورده و اطلاعات نسبتاً مطلوبی در اختیار دانشجویان قرار داده است.

۲-۳- تعیین نوع سؤالات

مرحله دوم در تهیه آزمون‌های کارشناسی ارشد، دقت در نوع سؤال‌هاست. همانطور که قبلاً ذکر شد، سؤالات آزمون کارشناسی ارشد فعلی بر اساس فرضیه‌های ۳۰ الی ۴۰ سال پیش تهیه می‌شود و تغییر محسوسی هم در هفت سال گذشته در نوع آنها به عمل نیامده است.

در آزمون عمومی، بهتر است سؤالات دستور زبان و واژگان که به شکل کاملاً سنتی برگزار می‌شود، حذف گردد. چون سنجش این دو بخش بطور جداگانه با تحولات جدید آزمون‌سازی سازگاری ندارد. پس آزمون‌ها را باید به طرف سنجیدن توانایی داوطلبان در استفاده از زبان در بافت ارتباطی واقعی سوق داد. برای مثال، در دوره کارشناسی ارشد، دانشجویان معمولاً مقالات علمی می‌خوانند، آنها را خلاصه می‌کنند، از آنها یادداشت برداری می‌کنند و این یادداشت‌ها را به طور منسجم بصورت یک مقاله تحقیقی به استادان

خود تحویل می‌دهند. انجام این فعالیت‌ها، نیاز به دانستن چگونگی ارائه منابع و مراجع مقالات، نحوه انتخاب مقالات مربوط به موضوع، طرز استفاده از منابع موجود در کتابخانه، توانایی خواندن راهنمای کتاب‌ها، فهرست مطالب، و خیلی عناوین دیگر دارد.

همه این فعالیت‌ها که دانشجویان به طور واقعی در دوران تحصیل با آن مواجه خواهند شد، می‌تواند در آزمون‌های عمومی کارشناسی ارشد، به عنوان بافت زبان مورد استفاده قرار گیرد. اکثر مواردی که در سؤالات فعلی گنجانده شده، شاید لحظه‌ای بعد از امتحان در واقع به درد داوطلبان نخورد. بنابراین بهتر است که آزمون عمومی به سوی واقعیت‌های استفاده از زبان سوق داده شود. مثلاً دانشجویان کارشناسی ارشد به سخنرانی‌ها و اخبارهایی که کلاً به زبان انگلیسی پخش می‌شود گوش دهند تا مهارت گوش کردن و توانایی یادداشت برداری از سخنرانی‌ها را بیاموزند. این مهارت می‌تواند در آزمون کارشناسی ارشد مورد سنجش قرار گیرد. اختصاص دادن قسمتی از امتحان حتی در مرحله دوم به این نوع آزمون‌ها باعث می‌شود انتخاب اصلح، با توجه به نیاز واقعی داوطلبان، در محیط علمی آینده آنها به عمل آید.

بنابراین، پیشنهاد می‌شود که، آزمون عمومی کارشناسی ارشد مهارت‌های شنیدن و خواندن را در بافت ارتباطی زبان بسنجند نه اجزاء زبان را به شکل مواد خارج از متن. مثال‌های زیر هر کدام می‌تواند در این نوع آزمون‌ها به کار رود:

الف- داوطلبان به یک سخنرانی علمی در رشته خودشان گوش دهند و به سؤالات مربوط به آن پاسخ دهند. آزمون‌هایی که در مقدمه این مقاله ذکر شد، نمونه‌های قابل قبولی هستند و می‌توانند برای سنجش توانایی شنیداری داوطلبان در بافت واقعی زبان به کار گرفته شوند.

ب- داوطلبانی که متون واقعی را بخوانند و میزان درک متن آنان سنجیده شود. در زمینه خواندن، شاید صدها مقاله تحقیقی در ابعاد مختلف آن به چاپ رسیده است. مؤلفه‌های خواندن و درک متن بطور تقریباً دقیق تعیین شده‌اند. بنابراین بهتر است توانایی

خواندن دانشجویان با توجه به مؤلفه‌های خواندن و استفاده از متون واقعی سنجیده شود. منظور از متون واقعی، متونی است که داوطلبان در دنیای واقعی دانشگاهی با آنها برخورد می‌کنند؛ مثلاً مقالات علمی، مطالب روزنامه‌ها، مطالب مجلات علمی یا حتی مطالب متون اختصاصی می‌توانند در آزمون‌ها به کار گرفته شوند.

ج- مجموعه‌ای از خواندن و نوشتن که مهارت‌های بسیار ضروری دانشجویان کارشناسی‌ارشد است به صورت آزمون‌های بسته سنجیده شود. آزمون‌های بسته نیز انواع متفاوت دارند. این آزمون‌ها اغلب بر محور حذف کلمات ساخته می‌شود، ولی در سطح کارشناسی‌ارشد می‌توان این آزمون‌ها را بر محور حذف عبارت و جمله تهیه کرد.

د- در نهایت نوشتن خلاق که از مهارت‌های بسیار مهم تحصیلات تکمیلی است، باید به نحوی مورد سنجش قرار گیرد. پیشنهاد نوشتن انشاء در این مرحله بسیار ناپخته است، اما آزمون‌هایی هستند که بصورت غیر مستقیم توانایی نوشتاری داوطلبان را می‌سنجند.

در مورد آزمون‌های اختصاصی پیشنهاد می‌شود که سؤالات از نوع کاربردی باشد، بدین معنی که به جای اینکه از داوطلب سؤال شود که تعریف یک واژه فنی چیست، آن واژه را در بافت یک فرآیند تدریس، یا بافت زبان‌شناختی یا آزمون‌سازی قرار دهند و در مورد نحوه برخورد داوطلب با موارد کاربردی از آنان سؤال شود. در اینجا فقط به ذکر دو مثال اکتفا می‌کنیم.

Testing :

There are 200 incoming undergraduate students at your department. The students have different abilities regarding their English. As an administrator, you want to give them a test to place them in appropriate classes. What sources of materials would be most appropriate to prepare the test?

- a. What they have studied at high school
- b. Materials from existing standard tests

- c. Materials they will study at university
- d. What they have learned before university

Methodology

What would be the most efficient way to check the students' assignments in a class of 60 students?

- a. To allocate a full session for checking the assignments
- b. To ask one of the students to check the assignments
- c. To check a few assignments randomly each session
- d. To have students peer check one another's assignments

۳-۳- تعیین نحوه تفسیر نمرات

مورد سوم برای بهبود کیفیت آزمون‌های کارشناسی ارشد، استفاده از فناوری برای تجزیه و تحلیل داده‌هاست. برای مدتی طولانی تئوری آزمون‌سازی کلاسیک مورد استفاده دانشمندان بوده است. در این تئوری، نمره داوطلب بر اساس دو مؤلفه توانایی واقعی و خطای اندازه‌گیری تعیین می‌شود و سؤالات تک تک مورد تجزیه و تحلیل قرار می‌گیرند. تئوری کلاسیک دو محدودیت جدی دارد. اولاً عوامل مختلفی که می‌توانند بر روی نمره داوطلب اثر بگذارند، در تئوری لحاظ نمی‌شوند. ثانیاً ویژگی‌های آماری سؤالات، تابعی از توانایی داوطلبان است، یعنی اگر سؤالی برای یک گروه مشکل به نظر می‌رسد، همان سؤال می‌تواند برای گروه دیگر آسان جلوه کند.

نارسایی‌های فرضیه کلاسیک باعث شد که دانشمندان فرضیهٔ بهتری را که به تئوری G (Generalizability) معروف است ارائه دهند. در این تئوری، نمره داوطلبان بر اساس مؤلفه‌های مختلف تجزیه و تحلیل می‌شود، مثلاً نوع سؤالات، تقسیم‌بندی سؤالات و تقدم و تأخر آنها، زمان و مکان آزمون، همه این‌ها می‌توانند متغیرهایی باشند که در نمره داوطلب اثر مثبت یا منفی بگذارند. به وسیلهٔ فرضیه G می‌توان عواملی را که ارتباط

مستقیم با توانایی داوطلب ندارند ولی بر نمره او تأثیر می‌گذارند از نمرات آنان زدود و به نمره واقعی داوطلب نزدیک‌تر شد.

و بالاخره در دو دهه گذشته، فرضیه Item Response یا مدل راش تقریباً جایگزین فرضیه‌های دیگر شده است. از طریق این فرضیه می‌توان سؤالاتی را تهیه کرد که بدون توجه به متغیرهای مربوط به داوطلب، دشواری یا آسانی آنها از قبل تعیین شده و توانایی داوطلب، بدون توجه به سؤال خاصی تعیین کرد. به علاوه، می‌توان سؤالات را چنان منظم کرد که فقط برای داوطلبانی مناسب باشد که از توانایی بالایی برخوردارند. و سرانجام می‌توان از یک مجموعه سؤال، چندین آزمون کاملاً مشابه تهیه کرد که در فرم‌های مختلف از نظر ویژگی‌های آماری یکسان باشد، این مورد می‌تواند به امنیت سؤالات کمک قابل ملاحظه‌ای بکند، چون تمام آزمون محدود به یک فرم نمی‌شود، بلکه فرم‌های متفاوت در یک اجرا می‌توانند مورد استفاده قرار گیرند.

نتیجه‌گیری

سخن آخر اینکه واژه رنگ باخته پژوهش در آزمون‌سازی باید در مهد مرکز سنجش آموزش کشور رنگ بگیرد، و اعتبار از دست رفته خود را در این مرکز باز یابد. باید اعتراف کرد که آزمون‌های امروزی در سطوح مختلف پایه و اساس پژوهشی ندارند. چون با توجه به شیوه آزمون و امنیت حفظ سؤالات که بعضاً در فاصله زمانی کوتاهی بعد از امتحان به طور شاید نامعقول منتشر می‌شود و پس از مدتی توسط خود سازمان سنجش آموزش کشور به صورت مجموعه‌هایی منتشر می‌شود، نمی‌توان از آنها برای اجراهای بعدی استفاده کرد. به اعتقاد من، در تمام آزمون‌ها باید بخشی را به عنوان بخش تحقیقی اضافه کرد تا بتوان بدینوسیله از شمار سؤالاتی که به دلایلی خوب عمل نمی‌کنند، کاست و آنها را با سؤالات جدیدتر جایگزین کرد. به علاوه با شیوه‌های موجود نمی‌توان از آزمون‌ها در اجراهای بعدی استفاده کرد و به این دلیل هر سال باید سرمایه‌گذاری عظیمی برای تهیه

آزمون جدید به عمل آید، در صورتی که با استفاده از روش‌های پژوهشی و فناوری موجود، می‌توان چندین آزمون هم سطح تهیه کرد و به کار برد و از نشت سؤالات و حتی پیش‌بینی سؤالات توسط داوطلبان جلوگیری به عمل آورد.

توجه داشته باشیم سازمان آزمون‌های آموزشی (ETS) فقط در مورد آزمون TOEFL که مربوط به یکی از بخش‌های این سازمان است تا کنون بیش از ۵۰ جلد گزارش تحقیقی منتشر کرده است. در این سازمان با قرار دادن امکانات در اختیار متخصصان، با هر قسمت از آزمون به صورت یک پروژه تحقیقاتی برخورد می‌شود و نتایج تحقیقات در اختیار مسئولان گذاشته می‌شود تا در بهبود کیفیت آزمون مورد استفاده قرار گیرد.

مسئله مهم دیگر تعداد داوطلبان است که در هیچ کجای دنیا به اندازه ایران نیست. تعداد شرکت‌کنندگان در آزمون توفل که در سراسر جهان ۱۲ بار در سال اجرا می‌شود به زحمت به مرز یک میلیون می‌رسد، ولی در ایران فقط در دوره کارشناسی، و در یک اجرا بیش از یک میلیون نفر در آزمون شرکت می‌کنند. پس باید از این فرصت طلایی استفاده و برای تهیه آزمون‌های موازی و همگن که با اصول آزمون‌سازی و فن‌آوری آموزشی هماهنگ است استفاده کرد.

مطمئناً سازمان آموزش کشور با توجه به حضور متخصصان امر، به تمام این موارد توجه کرده و از بهترین شیوه ممکن در آزمون‌سازی استفاده خواهد کرد. امید است که مطالب ارائه شده در این مقاله، بعضی از نکات تأزیک آزمون‌سازی را در ایران روشن کرده باشد. در نهایت امید است که پیشنهادهای ارائه شده حداقل مورد بررسی قرار گرفته و با به کارگیری آنها شاهد بهبود کیفیت آزمون‌ها در آینده نزدیک باشیم.

منابع

- سیف، علی اکبر، اندازه‌گیری و ارزشیابی پیشرفت تحصیلی، تهران: آگاه، ۱۳۶۴.
لانگمن. ر؛ پلت، ج؛ و وبر، ه، فرهنگ توضیحی زبان‌شناسی کاربردی، وثوقی، حسین و میرحسینی، سید اکبر (مترجمین). تهران: مرکز ترجمه و نشر کتاب، ۱۳۷۲.
هومن، حیدر علی، اندازه‌گیری های روانی و تربیتی و فن تهیه تست، تهران: دیبا، ۱۳۶۸.

- Alderson, C. and Beretta, A. (1992). *Evaluating second language education*. CUP.
- Farhady, H., Jafarpour, A., and Birjandi, P. (1994). *Language testing Skills. from theory to practice*. Tehran: SAMT.
- Farhady, H. (2000). *Ethical Consideration in Language Testing*. Moddares.
- Fox, J. (Ed.). (1996). *The Carleton Academic English Language (CAEL) Assessment: Test Manual*. Ottawa: Carleton University.
- Heaton, G. B. (1984). *Writing English language tests*. Longman.
- Howatt, A.P.R. (1984). *The history of Foreign Language Teaching*. OUP.
- Kelly, L.G. (1975). *25 centuries of language teaching*. Rowley, Mass.: Newbury House.
- Messick, S. (1989). 'Validity', in Robert Linn (ed.) *Educational Measurement (3rd ed.)*, Washington: Macmillan
- McNamara, T.F. (1996). *Measuring Second Language Performance*. Addison Wesley, Longman Limited.
- Rea-Dickins, P. (1997). So, why do we need relationships with stakeholders in language testing? A view from the UK. *Language Testing*, 14 (3).
- Shohamy, E. (1993). The power of tests: the impact of language tests on teaching and learning. *National Foreign Language Center (NFLC) Occasional Papers*, June.
- Shohamy, E. (1997). Testing methods, testing consequences: Are they ethical? Are they fair? *Language Testing*, 14 (3).

