

دن داگلاس

درآمدی بر آزمون سازی زبان

مترجمان:

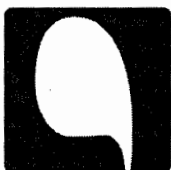
دکتر رضا مراد صحرایی

(عضو هیئت علمی دانشگاه علامه طباطبائی)

هدی سلیمی

فائزه مرصوص





درآمدی بر آزمون سازی زبان

نویسنده:

دن داگلاس

مترجمان:

دکتر رضامراد صحرایی

(عضو هیئت علمی دانشگاه علامه طباطبائی)

هدی سلیمی

فائزه مرصوص



سازمان اسناد و کتابخانه ملی
جمهوری اسلامی ایران

- موضوع: سوشلسه
عنوان و نام پدیدآور: دوآمدی بر آزمون‌سازی زبان / نویسنده دن داگلاس؛ ترجمه رضاغداد صحرایی، هدی سلیمی،
فائزه مرصوصی
- مشخصات نشر: تهران: نشر نویسه پارسی، ۱۳۹۶.
مشخصات ظاهری: ۲۳۸ ص: مصور، جدول، نمودار
شابک: ۹۷۸-۶۰۰-۷۰۳۰-۵۶-۱
- وضعیت فهرست نویسی: فیبا
یادداشت: کتابخانه
یادداشت: واژه‌نامه
یادداشت: عنوان اصلی: Understanding language testing, c 2010.
موضوع: زبان - آزمون‌ها
موضوع: Language and languages - Examinations
- شماره افزوده: صحرایی، رضاغداد، ۱۳۵۵ - مترجم
شماره افزوده: سلیمی، هدی، ۱۳۶۷ - مترجم
شماره افزوده: مرصوصی، فائزه، ۱۳۶۰ - مترجم
زده بندی کنگره: Pa۳/۴۵۳۵۴ ۱۳۹۶
زده بندی دیویی: ۴۰۷/۶
شماره کتابشناسی ملی: ۴۹۵۷۴۰۲

این کتاب برگردانی است از:

Understanding Language Testing

Dan Douglas

© 2010 Dan Douglas

نویسنده: دن داگلاس Dan Douglas

مترجمان: رضامراد صحرایی، فائزه مرصوص، هدی سلیمی

طراح جلد، گرافیک، صفحه آرایی و ناظر فنی چاپ: محمد محرابی

www.mehraabi.ir / 09125163545

ناشر: نشر نویسه پارسی

دفتر انتشارات: ۰۲۱-۷۷۰۵۳۲۴۶

فروشگاه: ۰۲۱-۶۶۹۵۷۱۳۲

سامانه پیام کوتاه: ۳۰۰۰۴۵۵۴۵۵۴۱۴۲

وبگاه: www.neveeseh.com

نوبت چاپ: اول، ۱۳۹۷

شمارگان: ۳۰۰ نسخه

شابک: ۹۷۸-۶۰۰-۷۰۳۰-۵۶-۱

چاپ و صحافی: روز

قیمت: ۲۵۰۰۰ تومان

کلیه حقوق محفوظ و متعلق به «نشر نویسه پارسی» است.
تکثیر و انتشار این اثر یا قسمتی از آن به هر شیوه، بدون مجوز قبلی و کتبی
ممنوع و مورد پیگیری قانونی قرار خواهد گرفت.



کاغذ بالک:

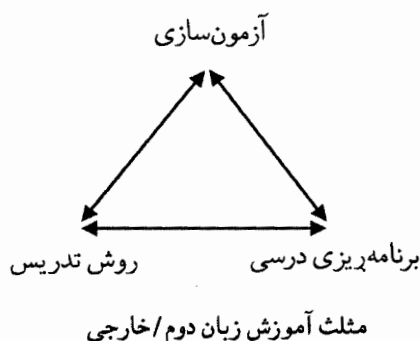


کاغذی است که وزن سبک، جرم حجمی کم و منخامت بالا دارد و رنگ آن برای مطالعه مناسب است. این کاغذ از الیاف چوب تهیه شده و کتاب‌های به وجود آمده از آن وزن بسیار کمتری نسبت به کتاب‌های ساخته شده از کاغذهای تحریر سبدرنگ (WoodFree) دارند و همین موضوع، مطالعه و جابه‌جایی آن را آسان می‌کند. این کاغذ در کشورهای توسعه‌یافته بزرگ‌ترین سهم را در انتشار کتاب دارد.

آزمون‌سازی زبان به عنوان هنری خاص و سری شهرت یافته که حوزه فعالیت متخصصان آموزش دیده است، این هنر مدت‌هاست که از کلاس‌های عادی زبان و حتی از دنیای واقعی کاربرد زبان نیز حذف شده است. اغلب، مسئولان بالادستی بدون اینکه تجربه و پیشینه ذهنی زبان‌آموز، روش‌های تدریس و موقعیت‌ها را در نظر بگیرند، آزمون‌های زبان را تهیه می‌کنند و در اختیار ما می‌گذارند. یکی از انگیزه‌های اصلی تألیف کتاب حاضر روشن کردن فرایند آزمون‌سازی زبان و چگونگی طراحی آزمون است که به اساتید رشته زبان و دانشجویان رشته آموزش زبان و زبان‌شناسی کاربردی کمک کند تا ماهیت آزمون‌سازی زبان به عنوان ابزاری برای اندازه‌گیری توانایی زبانی را درک کنند. این کتاب نخستین مرحله در آزمون‌سازی زبان است، و بستری را فراهم می‌کند تا دریابیم که اندازه‌گیری مشخصه ناملموسی همچون دانش زبانی، به چه معناست. هدفم از نگارش این کتاب متوجه ساختن اساتید و دانشجویان این رشته به اهمیت تجربه شخصی‌شان در آموزش و یادگیری زبان به عنوان یک عامل مهم، هم در تفسیر نتایج آزمون و هم در طراحی آزمون‌هایشان است.

امیدوارم که خوانندگان کتاب پس از مطالعه این بخش مقدماتی، علاقه و اراده لازم را پیدا کنند که هرچه بیشتر در حوزه آزمون‌سازی زبان به مطالعه بپردازند و دیگر کتاب‌ها و مقالات چاپ شده در این حوزه را نیز بخوانند و در نتیجه دید عمیق‌تری از این رشته پیدا کنند. در هر فصل از کتاب، منابع مختلف و تألیفات گوناگونی معرفی شده‌اند، علاوه بر آن، قسمتی با عنوان «برای مطالعه بیشتر» در ارتباط با وجوه خاص اندازه‌گیری توانایی زبانی پیش‌بینی شده است. همچنین، در انتهای هر فصل، یک قسمت با عنوان «تمرین» طراحی شده که هم تجربه‌ای از درک آزمون‌سازی زبان و هم تجربه‌ای در زمینه انجام آن برای خواننده فراهم می‌کند. آزمون‌سازی زبان، بخش مهمی از آموزش و یادگیری زبان است، چون نتایج آزمون‌های زبان می‌توانند تأثیر مهمی بر زندگی دانشجویان ما، تدریس خود ما، برنامه‌های کاری ما و در نهایت جوامعی که در آن زندگی می‌کنیم، داشته باشند. مسئولیت اخلاقی ما این است که نه تنها ماهیت آزمون‌سازی زبان را درک کنیم بلکه کارکرد آزمون‌ها و تفسیر کاربردی از عملکرد آزمون را نیز بطور مؤثری ارتقا دهیم.

آموزش زبان دوم / خارجی یک حرفه تخصصی است که سه بخش اصلی و کلیدی دارد: روش تدریس^۱، برنامه‌ریزی درسی^۲ و آزمون‌سازی^۳. این سه بخش با یکدیگر رابطه فرایندی و چرخه‌ای دارند و به‌مثابه سه ضلع یک مثلث هستند که آموزش زبان دوم / خارجی بر آنها استوار می‌گردد:



معلم زبان حرفه‌ای و موفق کسی است که علاوه بر آشنایی کافی با اصول و فنون تدریس و برنامه‌ریزی درسی، با اصول و مبانی آزمون‌سازی زبان و روش‌های تهیه آزمون نیز آشنا باشد. ارزشیابی و آزمون‌سازی جزء لاینفک حرفه معلمی زبان است، زیرا مبنای اصلی تصمیم‌گیری بهینه و عادلانه در نظام آموزش است. بر این اساس، معلم زبان باید با انواع آزمون و هدف هر یک از آنها آشنا باشد؛ اصول و فنون آزمون‌سازی را بداند؛ بر فرایند تهیه و اجرای آزمون مسلط باشد و شیوه‌های تفسیر و تحلیل نمرات آزمون را بشناسد.

آزمون هم برای معلم اهمیت دارد و هم برای زبان‌آموز. آزمون به معلم کمک می‌کند تا نقاط قوت و ضعف روش تدریس خود را کشف کند؛ سطح زبان‌آموز را تشخیص دهد و برای ارتقاء او به سطح بالاتر تصمیم بگیرد. آزمون برای زبان‌آموز نیز اهمیت کلیدی دارد. آزمون خوب به زبان‌آموزان انگیزه مطالعه و یادگیری می‌دهد؛ در آنها احساس پیشرفت ایجاد می‌کند و مقیاس قابل اعتمادی برای مقایسه آنها با خودشان و سایر همکلاسی‌هایشان است.

کتاب درآمدی بر آزمون‌سازی زبان، پس از تبیین اهمیت آزمون در فرایند کلان برنامه‌ریزی آموزشی^۴ و ملاحظات اخلاقی تهیه و اجرای آزمون، انواع آزمون‌های زبان و اهداف هر یک از آنها را مورد بحث قرار داده است. در این کتاب، انواع آزمون‌های آزمون و نحوه تحلیل آنها؛ نحوه سنجش هر یک از توانایی‌های زبانی؛ و ویژگی‌های آزمون‌های معتبر و پایا نیز با زبانی گویا و نسبتاً ساده مورد بحث

¹ Methodology

² Syllabus design

³ Testing

⁴ Curriculum Development

قرار گرفته است. یکی از مزیت‌های این کتاب، معرفی آزمون‌های رایانه‌ای زبان و ابزارهای آنها است که در بخش انتهایی فصل سوم به آنها اشاره شده است. دیگر مزیت ویژه این کتاب نسبت به کتاب‌های مشابه، توجه آن به تکنولوژی‌های نوین آزمون‌سازی زبان است. فصل ششم کتاب به این موضوع اختصاص یافته است.

کتاب درآمدی بر آزمون‌سازی زبان یک منبع ساده و قابل فهم برای هر کسی است که به کار آموزش زبان دوم/خارجی اشتغال دارد. در این بین، نیت اصلی مترجمان، ارائه یک منبع آموزشی مناسب برای آزمون‌سازی زبان فارسی بوده است. در برنامه درسی مقطع کارشناسی ارشد «آموزش زبان فارسی به غیرفارسی‌زبانان» یک درس دو واحدی با عنوان «آزمون‌سازی زبان» گنجانده شده است که کتاب حاضر می‌تواند منبع مناسبی برای آن درس باشد. امیدواریم این ترجمه گام مؤثری در توسعه دانش آزمون‌سازی معلمان زبان دوم/خارجی، خصوصاً معلمان ایرانی و غیرایرانی زبان فارسی، باشد.

گروه مترجمان

اسفندماه ۱۳۹۶

فهرست مطالب

۵	پیشگفتار نویسنده
۷	پیشگفتار مترجمان

| فصل اول: خط کش منعطف لاستیکی / ۱۳ |

۱۵	۱-۱. چرا یادگیری زبان را می‌سنجیم؟
۱۷	۱-۲. آزمون زبان چیست؟
۱۷	۱-۲-۱. خصوصیات ابزارهای اندازه‌گیری چیست؟
۱۸	۱-۳. خط کش منعطف لاستیکی
۲۰	۱-۴. آزمون، اندازه‌گیری و ارزشیابی
۲۰	۱-۴-۱. ارزشیابی بدون اندازه‌گیری
۲۰	۱-۴-۲. اندازه‌گیری بدون آزمون
۲۱	۱-۴-۳. آزمون‌ها
۲۵	۱-۵. کاربردهای اخلاقی آزمون‌های زبان
۲۶	۱-۵-۱. پایایی
۲۶	۱-۵-۲. روایی
۲۹	منابع پیشنهادی برای مطالعه بیشتر
۲۹	تمرین‌ها

| فصل دوم: سنجش توانایی زبانی و تصمیم‌گیری / ۳۵ |

۳۷	۲-۱. سنجش توانایی زبانی
۴۰	۲-۲. تصمیم‌گیری درباره‌ی زبان‌آموزان، مدرسان، برنامه‌ریزان آموزشی و سیاست‌های آموزشی
۴۱	۲-۳. محیط‌های کاربرد زبان
۴۳	۲-۳-۱. مؤلفه‌های بافتی
۴۳	۲-۳-۱-۱. موقعیت (S)
۴۳	۲-۳-۱-۲. مشارکان (P)
۴۴	۲-۳-۱-۳. اهداف (E)
۴۴	۲-۳-۱-۴. ترتیب کنش‌ها (A)
۴۴	۲-۳-۱-۵. کلیدهای راهنما (K)
۴۵	۲-۳-۱-۶. ابزارهای واسط (I)
۴۵	۲-۳-۱-۷. هنجارهای تعامل (N)
۴۵	۲-۳-۱-۸. ژانر (G)

۴۶	۲-۳-۲. اصالت
۴۹	۴-۲. تفسیر معتبر از عملکرد آزمون
۴۹	۲-۴-۱. ثبات اندازه‌گیری
۵۳	۲-۴-۲. روایی: شاهدی برای تفسیر عملکرد زبان‌آموزان در آزمون
۵۳	۲-۴-۱. چه چیزی در جعبه سیاه وجود دارد؟
۵۹	۲-۴-۲. شواهد روایی آزمون
۶۰	۵-۲. نتیجه‌گیری: سوگیری ایده‌آل
۶۲	برای مطالعه بیشتر
۶۲	تمرین‌ها

فصل سوم: طراحی آزمون / ۶۵

۶۷	۱-۳. چه مواردی را در یک آزمون بسنجیم؟ تجزیه و تحلیل نیازها
۶۹	۱-۱-۳. مشخص کردن اهداف آزمون
۶۹	۲-۱-۳. جمع‌آوری اطلاعات اولیه
۷۰	۳-۱-۳. جمع‌آوری داده‌های مقدماتی
۷۲	۳-۱-۴. جمع‌آوری داده‌های ثانویه
۷۴	۵-۱-۳. تحلیل پرسش‌هایی که کاربرد زبان هدف و ویژگی‌های زبانی را مورد بررسی قرار می‌دهند
۸۰	۲-۳. چگونه توانایی‌های زبانی را بسنجیم؟ تبدیل کاربردهای زبان هدف به پرسش‌های آزمون
۸۱	۳-۲-۱. طراحی تکلیف در آزمون
۸۲	۳-۲-۲. طراحی طرح کلی آزمون
۸۳	۳-۲-۲. گزینه‌هایی برای پرسش‌های آزمون
۸۳	۳-۲-۱. پاسخ انتخابی
۸۴	۳-۲-۲. پاسخ کوتاه
۸۵	۳-۲-۳. پاسخ بلند
۸۸	۳-۲-۴. خلاصه‌نویسی
۸۸	۳-۳. چگونه آزمون بدهیم؟ برگزاری آزمون
۸۹	۳-۳-۱. محیط آزمون
۸۹	۳-۳-۲. مسئولین برگزاری آزمون
۹۰	۳-۳-۳. شیوه‌های برگزاری آزمون
۹۰	۳-۳-۴. نمره‌دهی
۹۲	۴-۳. چگونه می‌توانیم در طراحی یک آزمون از رایانه استفاده کنیم؟
۹۳	۳-۴-۱. نرم‌افزار آزمون‌سازی هات پوتیتوز
۹۶	۳-۴-۲. مودل
۹۷	۳-۴-۳. وب‌سیتی

۱۰۱	۵-۳. نتیجه‌گیری
۱۰۲	منابع پیشنهادی برای مطالعه بیشتر

فصل چهارم: ارزیابی‌های نوین / ۱۰۷

۱۰۹	۱-۴. آزمون‌های هنجار- مرجع و آزمون‌های ملاک - مرجع
۱۱۲	۲-۴. آزمون‌های ارتباطی زبان
۱۱۳	۳-۴. آزمون‌های زبان با اهداف «عمومی» و «ویژه»
۱۱۴	۴-۴. آزمون‌های نکته- تفکیکی و آزمون‌های جامع
۱۱۶	۵-۴. ارزیابی تکوینی و ارزیابی تراکمی
۱۱۶	۶-۴. رویکردهای نوین ارزیابی
۱۱۷	۱-۶-۴. ارزیابی مبتنی بر کنفرانس
۱۱۸	۲-۶-۴. ارزیابی مبتنی بر کاربوشه
۱۲۰	۳-۶-۴. خودارزیابی و ارزیابی هم‌کلاسی‌ها
۱۲۱	۴-۶-۴. ارزیابی تکلیف-محور و ارزیابی عملکرد
۱۲۴	۵-۶-۴. ارزیابی پویا
۱۲۶	۶-۶-۴. خلاصه
۱۲۶	۷-۴. نتیجه‌گیری
۱۲۷	برای مطالعه بیشتر
۱۲۷	تمرین‌ها

فصل پنجم: همراه با اعداد: دوره‌ای کوتاه از بررسی آمار / ۱۳۱

۱۳۳	۱-۵. مقدمه
۱۳۶	۲-۵. توزیع نرمال
۱۳۸	۳-۵. معدل یا میانگین
۱۳۹	۴-۵. انحراف معیار
۱۴۲	۱-۴-۵. انحراف معیار به عنوان واحد اندازه‌گیری
۱۴۴	۵-۵. همبستگی
۱۴۸	۶-۵. احتمال و معنی‌داری آماری
۱۵۱	۷-۵. آزمون تی در اختلاف بین دو میانگین
۱۵۴	۸-۵. تحلیل واریانس
۱۵۷	۹-۵. پایایی
۱۵۸	۱-۹-۵. روش جداسازی دو نیمه

۱۶۰	۲-۹-۵. روش ثبات درونی
۱۶۲	۳-۹-۵. خطای استاندارد اندازه‌گیری
۱۶۴	۱۰-۵. پایایی نمره‌دهی ارزیاب
۱۶۶	۱۱-۵. نتیجه‌گیری
۱۶۷	برای مطالعه بیشتر
۱۶۸	تمرین‌ها

الفصل ششم: فناوری و آزمون‌سازی زبان / ۱۷۱

۱۷۳	۱-۶. مقدمه
۱۷۴	۲-۶. مسائل مطرح در فناوری و آزمون‌سازی زبان
۱۷۵	۱-۲-۶. فناوری و نگرش آزمون‌دهنده‌ها
۱۷۶	۲-۲-۶. عملکرد زبانی و وسایل ارتباط جمعی گوناگون
۱۷۷	۳-۲-۶. فناوری و ساختی که قراراست اندازه‌گیری شود
۱۷۷	۴-۲-۶. فناوری و تکالیف ارزیابی
۱۷۸	۵-۲-۶. محدودیت‌های نمره‌دهی خودکار
۱۷۹	۳-۶. فناوری و انواع تکالیف زبانی
۱۷۹	۱-۳-۶. تکلیف‌های شنیداری
۱۸۲	۲-۳-۶. گوش کردن تلفیقی و تکلیف‌های گفتاری
۱۸۶	۳-۳-۶. تکلیف‌های نوشتاری
۱۸۸	۴-۳-۶. تکلیف‌های خواننداری
۱۹۰	۴-۶. فرصت‌ها و تهدیدهای نمره‌دهی خودکار
۱۹۰	۱-۴-۶. نمونه‌هایی از برنامه‌های رایج نمره‌دهی خودکار
۱۹۳	۲-۴-۶. نگرانی‌هایی در مورد نمره‌دهی خودکار
۱۹۴	۵-۶. بازخورد آزمون و اعلام نتایج آزمون
۱۹۸	۶-۶. منابع آماری برخط و رایانه‌محور
۱۹۸	۱-۶-۶. مایکروسافت اکسل
۲۰۰	۲-۶-۶. منابع برخط
۲۰۳	۷-۶. جمع‌بندی
۲۰۴	برای مطالعه بیشتر
۲۰۴	تمرین‌ها
۲۰۹	فهرست منابع
۲۱۷	واژه‌نامه
۲۳۳	نمایه

فصل اول

مقدمه: خطکش منعطف لاستیکی

مدرسان زبان، هم در کلاس و به هنگام آموزش و یادگیری و هم در بیرون از کلاس و طی ساعات اداری و سفرهای علمی با زبان‌آموزان خود هستند در نتیجه، زمان زیادی را با آنها می‌گذرانند. مدرسان زبان، عملکرد زبان‌آموزان را در یادگیری زبان دوم، مثلاً در تلفظ، تعداد واژه‌هایی که می‌دانند و درستی استفاده از آنها، همچنین نحوه صحیح استفاده از قوانین نحوی، به طور پیوسته مورد بررسی قرار می‌دهند. برای مثال، اگر از آنها پرسیده شود که عملکرد ریچارد در کلاس زبان فرانسه چگونه است، یا اینکه پائولا چقدر در کلاس زبان انگلیسی پیشرفت می‌کند، می‌توانند ارزیابی خود از پیشرفت زبان‌آموزانشان را اعلام کنند، مثلاً به این صورت: «حُب، پائولا واژه‌ها و جملات را خوب تلفظ می‌کند، او فقط کمی لهجه دارد، اطلاعاتش در زمینه دستور به‌ویژه زمان گذشته خیلی بالا نیست و فهم درستی از جمله‌واره‌های موصولی ندارد. دامنه لغات او برای کسی که سال دوم یادگیری زبان خارجی را می‌گذراند، عالی است.»

اگر مدرسان زبان بتوانند پیشرفت زبان‌آموزان را با این روش، ارزیابی کنند؛ که اکثر آنها بین دو تا سه هفته به این توانایی دست پیدا می‌کنند؛ پس در این صورت چه نیازی به آزمون‌های زبان دارند؟ شاید مهم‌ترین دلیل، رعایت عدالت باشد. آنها می‌خواهند اطمینان پیدا کنند که با همه زبان‌آموزان به یک شیوه رفتار شده و به همه به یک اندازه فرصت داده می‌شود که نشان دهند چه چیزهایی آموخته‌اند و چه کارهایی می‌توانند با آنچه یاد گرفته‌اند، انجام دهند. آزمون‌ها این فرصت را فراهم می‌کنند که معلم با یک دستورالعمل مشابه و در شرایط مشابهی، اطلاعات یکسانی را در اختیار زبان‌آموزان قرار دهد. او همچنین می‌تواند پیشرفت زبان‌آموزان را از زوایای مختلف بسنجد و از آن اطمینان پیدا کند. معلم از طریق آزمون می‌تواند نقاط ضعف زبان‌آموز را تشخیص دهد و راه‌حل‌های گوناگونی برای برطرف نمودن آن پیشنهاد نماید. او همچنین به تعدادی استاندارد دست می‌یابد تا به وسیله آنها عملکرد و پیشرفت زبان‌آموزان و کیفیت عملکردشان را بسنجد. درنهایت، آزمون به معلم کمک می‌کند که پیشرفت زبان‌آموزان را هر بار با شیوه مشابهی قضاوت کند و ارزیابی دقیق و پایانی داشته باشد.

آزمون به برنامه‌ریزان آموزشی، والدین، مسئولین اجرایی و همچنین مدرسان آینده اطمینان می‌دهد که زبان‌آموزان براساس استانداردهایی که عموماً پذیرفته شده‌اند یادگیری زبان را پیش می‌گیرند یا به سطح کفایت از توانش زبان دوم رسیده‌اند. آزمون‌های پیشرفت^۱ براساس مواد آموزشی که زبان‌آموزان در کلاس می‌خوانند یا براساس آنچه باید در آخر کلاس یاد گرفته باشند طراحی می‌شوند؛ این آزمون‌ها اطلاعاتی در اختیار عوامل اجرایی قرار می‌دهند؛ مثلاً می‌گویند که زبان‌آموزان پیشرفت لازم را داشته‌اند و مدرسان کار خود را به درستی انجام داده‌اند. از طرفی، آزمون‌های مهارتی^۲ کمکی برای مدرسان

¹ Achievement tests

² Proficiency tests

هستند که پیش‌بینی کنند زبان‌آموزان چه کارهایی می‌توانند در موقعیت‌های ارتباطی خارج از کلاس با زبان انجام دهند. برای مثال، از آزمون مهارت دانشگاهی زبان آلمانی زمانی مورد استفاده قرار می‌گیرد تا ببینند، آیا دانش‌زبانی فرد برای تحصیل در رشته تاریخ در دانشگاه برلین کافی است یا نه. آزمون زبان برای تجارت هم به کارفرمایان کمک می‌کند که مشخص کنند آیا یک متقاضی کار می‌تواند مذاکرات تجاری در آمریکا انجام دهد. این موارد، نکات مهم آموزشی و اجتماعی هستند که یک آزمون می‌تواند در اختیار ما بگذارد.

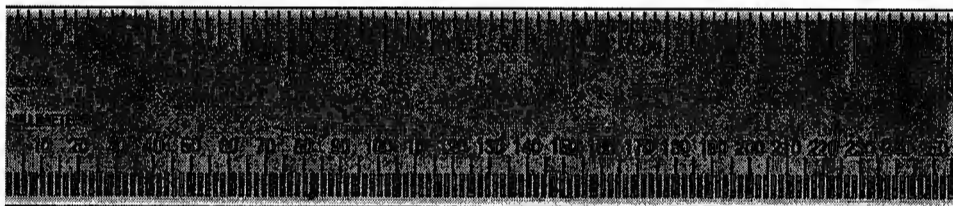
آزمون‌ها همچنین ابزار مهمی برای سیاست عمومی هستند. برای مثال، آزمون‌های ملی، برای اطمینان پیدا کردن از اینکه آیا استانداردهای مشابهی در سراسر کشور برای آموزش زبان‌آموزان در تمامی مؤسسه‌های آموزشی، به کار گرفته می‌شود. این آزمون‌ها با هدف «دروازه‌بانی» هم مورد استفاده قرار می‌گیرند؛ این آزمون‌ها به ما اطمینان می‌دهند که تنها زبان‌آموزانی که بهترین عملکرد را داشته‌اند به سطح بعدی راه پیدا می‌کنند؛ به‌ویژه، در کشورهایی که تعداد متقاضیان تحصیل بیشتر از توانایی دولت برای فراهم کردن امکانات مورد نیاز است. آزمون‌ها همچنین گاهی اوقات به‌عنوان ابزارهای مستقیم سیاسی مورد استفاده قرار می‌گیرند، برای مثال، برای کنترل تعداد مهاجران و کارگران مهاجر استفاده می‌شوند. چنین استفاده‌های سیاسی از آزمون‌ها نیازمند ملاحظات دقیق اخلاقی است، که در ادامه این فصل مورد بحث قرار خواهد گرفت.

یک آزمون باید کیفیت بسیار بالایی داشته باشد تا بتواند همه این مزایا و عملکردها را در خود جای دهد، در این صورت می‌توانیم به وسیله آزمون، اهدافی که در ذهن داریم را به نتیجه برسانیم و همچنین تفسیر درستی از این نتایج ارائه دهیم. درواقع پیام اصلی این کتاب، دستیابی به درکی درست از کاربردهای مختلف آزمون‌های زبان و فرایند آزمون‌سازی، نمره‌دهی به عملکرد زبان‌آموزان، تحلیل و تفسیر نتایج آزمون، و از همه مهم‌تر، استفاده از آزمون است که تا حد امکان، به‌صورت منصفانه و اخلاقی استفاده شود و در نتیجه، فرصت‌هایی به آزمون‌دهنده‌ها داده شود که بهترین عملکرد خود را نشان دهند، و تا بیشترین حد ممکن، یاد بگیرند و احساس مثبتی به یادگیری زبان داشته باشند.

اگر بخواهیم مفهوم آزمون زبان را به درستی توصیف کنیم، باید بگوییم که آزمون، یک ابزار اندازه‌گیری است و درواقع تفاوت زیادی با یک خط‌کش منعطف لاستیکی یا یک دماسنج ندارد. آزمون زبان، ابزاری برای سنجش توانایی زبانی است، حتی می‌توانیم ابعاد کمی آزمون را هم بررسی کنیم: یک فرد چه میزان اطلاعات زبانی دارد؟ اما وقتی که می‌گوییم می‌خواهیم توانایی زبانی یا میزان اطلاعات زبانی فرد را بسنجیم، چه معنایی دارد؟ چگونه می‌توانیم مفهومی انتزاعی مانند توانایی زبانی^۱ را اندازه‌گیری کنیم؟ برای پاسخ به این پرسش‌ها، ابتدا باید خصوصیات ابزارهای اندازه‌گیری معمول^۱ بیشتری را به‌طور خلاصه مورد بررسی قرار دهیم.

۱-۲-۱. خصوصیات ابزارهای اندازه‌گیری چیست؟

هدف اصلی ما از سنجش این است که بتوانیم، میزان و اندازه یک ویژگی یا خصوصیت را براساس قوانین و مقیاس‌های خاصی نمره‌دهی کنیم. یکی از واضح‌ترین و قابل استفاده‌ترین مقیاس‌ها، مقیاس شمارش ساده است: تعداد واژه‌های موجود در جمله قبل ۲۴ عدد است. البته متخصصان زبان بیشتر از ابزارهایی استفاده می‌کنند که براساس آن کیفیت و کمیت را با هم اندازه‌گیری می‌کنند، مانند دماسنج، خط‌کش، ترازو، سرعت‌سنج، ساعت، هواسنج و غیره. تمامی این ابزارها یک واحد اندازه‌گیری مانند میلی‌متر، اینچ، لیتر، دقیقه و دیگر اندازه‌ها دارند.



شکل ۱-۱. خط‌کش

وقتی که وزنمان زیاد می‌شود، مقیاس را مقصر می‌دانیم، درواقع استدلال ما این نیست که یک پوند، یک پوند است و یک کیلوگرم، یک کیلوگرم همانطور که باید باشد؛ ما واحدهای اندازه‌گیری را در زندگی روزانه‌مان به عنوان استانداردهایی که همه در مورد آن تفاهم داریم، می‌پذیریم. بعضی اوقات، تصور می‌کنیم که مقیاس ما درست اندازه‌گیری نمی‌کند، اما به راحتی می‌توانیم این مورد را بررسی کنیم، می‌توانیم خودمان را با مقیاس دیگری وزن کنیم و امید به نتایج بهتری داشته باشیم!

یکی از ویژگی‌های جالب بسیاری از ابزارهای اندازه‌گیری این است که واحدهای اندازه‌گیری آنها یکپارچه و منطبق برهم هستند. در خط‌کشی که در شکل بالا نشان داده شده است، برای مثال، می‌توانیم یک شیء را در هر نقطه‌ای از مقیاس، اندازه‌گیری کنیم و به نتیجه یکسانی برسیم. همچنین،

^۱ Language ability

این خط‌کش دارای نقطه صفر واقعی به معنی عدم وجود طول است؛ این ویژگی امکانی را فراهم می‌کند که دو شیء را مستقیماً براساس طولشان مقایسه کنیم؛ یک شیء ۱۰۰ میلی‌متر است و دو برابر یک شیء ۵۰ میلی‌متری بلند است. علاوه بر این، اندازه خط‌کش تغییر نمی‌کند، بنابراین اگر امروز اندازه چیزی ۱۰۰ میلی‌متر باشد، فردا هم همین خواهد بود، و ما همچنان منتظر نتایج مشابهی خواهیم بود مگر اینکه خود شی تغییر اندازه داده باشد؛ مثلاً، قد فرزندانمان را هر ماه اندازه‌گیری می‌کنیم تا ببینیم چقدر رشد کرده‌اند.

اگر بخواهیم خلاصه بگوییم، اکثر ابزارهای اندازه‌گیری این امکان را برای ما فراهم می‌کنند که با نمره‌دهی به آن چیزی که اندازه‌گیری می‌کنیم یا کیفیت مورد نظر، آنها را از هم متمایز و سپس رتبه‌بندی کنیم. فاصله‌های بین واحدهای اندازه‌گیری در یک مقیاس، مساوی و یکسان هستند و یک نقطه صفر واقعی دارند که به ما اجازه می‌دهند، چیزهایی که اندازه‌گیری می‌کنیم را با هم مقایسه کنیم. در نهایت، ابزارهای اندازه‌گیری که ما به‌طور معمول استفاده می‌کنیم، ابزارهای قابل اطمینانی هستند؛ ما به نتایج اطمینان داریم و انتظار داریم که هر بار نتایج مشابهی بگیریم.

۱-۳. خط‌کش منعطف لاستیکی

آزمون زبان چه نوع ابزار اندازه‌گیری است؟ اگر بخواهیم آزمون زبان را با یک خط‌کش مقایسه کنیم، احتمالاً می‌گوییم که آزمون زبان شبیه یک خط‌کش لاستیکی است^۱. تصور کنید که یک خط‌کش هم‌منظور که با آن اشیانی را اندازه‌گیری می‌کنید، کشیده و کوتاه و بلند می‌شود. مثلاً، وقتی قد دختر دو ساله خود را برای اولین بار اندازه می‌گیرید، رشد ۳۵ اینچ را نشان می‌دهد، بار دوم ۴۳ اینچ و بار سوم ۲۷ اینچ! شما نمی‌توانید از چنین مقیاسی برای اندازه‌گیری میزان رشد فرزندان استفاده کنید، و همچنین نمی‌توانید به‌طور معنی‌داری قد او را با برادر ۵ ساله‌اش مقایسه کنید. یک خط‌کش منعطف نمی‌تواند ابزار خیلی مفیدی باشد، با این حال، به عقیده من یک آزمون زبان تعداد زیادی از ویژگی‌های یک ابزار اندازه‌گیری مفید را دارد. نکته اول اینکه بحث‌های مختلفی درباره طبیعت واحدهای اندازه‌گیری در یک آزمون زبان و معانی آنها وجود دارد. برای مثال، ما اغلب سطح زبانی زبان‌آموزان را در سطوح مقدماتی، میانی یا پیشرفته می‌دانیم، اما هر کدام از ما نظر متفاوتی درباره معانی این اصطلاحات داریم. حتی اگر در یک آزمون بگوییم که نمره ۵۰ نشان‌دهنده سطح مقدماتی، ۷۰ سطح میانی، و ۹۰ به بالا سطح پیشرفته است، به‌سختی می‌توانیم افراد دیگر را با خود هم نظر کنیم. نکته دوم اینکه، فاصله بین واحدها در یک مقیاس همیشه مساوی نیستند. تعدادی از سازمان‌های آمریکا از

^۱ یکی از اساتید من در دانشگاه ادینبورگ، در دوره الیزابت انگرام آخر، در سال ۱۹۷۳، این واژه را به من آموخت که از آن زمان تاکنون از آن استفاده می‌کنم.

یک مقیاس شناخته شده مهارت زبانی استفاده می‌کنند که دارای پنج سطح "مقدماتی" تا "فوق پیشرفته" است و براین اساس توانایی زبانی افراد را می‌سنجند.^۱ پیشرفت از سطح "مقدماتی" به سطح "پیش میانی" (که مهارت زبانی محدودی دارد) بسیار آسان‌تر از پیشرفت از سطح "پیشرفته" به "فوق پیشرفته" است؛ به همین دلیل است که می‌بینیم فاصله‌های موجود بین واحدها در یک مقیاس برابر نیستند. علاوه بر این، در مقیاس مهارت زبانی، نقطه صفر واقعی وجود ندارد، بنابراین امکان ندارد اذعان کنیم که چون نادیا دو برابر ماریو نمره گرفته است، پس دو برابر او زبان فرانسه می‌داند. در نهایت، اگر فردا دوباره همین آزمون را بگیریم، تقریباً می‌توانیم تضمین کنیم که همه زبان‌آموزان نمره متفاوتی خواهند گرفت! از نکات فوق می‌توانیم به این نتیجه برسیم که آزمون زبان، ابزار اندازه‌گیری چندان مناسبی نیست. البته نباید کلاً چنین ابزاری را کنار بگذاریم؛ با برداشتن چند گام به سوی جلو می‌بینیم که اگر تنها ابزار ما برای اندازه‌گیری، یک خط‌کش منعطف لاستیکی بود، این ابزار هم می‌تواند برای اندازه‌گیری مفید باشد.

ابتدا باید ببینیم که یک خط‌کش منعطف لاستیکی تا کجا کشیده می‌شود، و متوجه می‌شویم که میزان تغییرات در اندازه‌گیری ما پایان ندارد، بنابراین بهتر است محدودیتی برای میزان کشش خط‌کش و همچنین محدودیتی برای کوتاه شدن آن مشخص کنیم؛ اگر این محدودیت‌ها را بدانیم، حداقل می‌توانیم متوجه شویم که اندازه‌گیری ما چقدر می‌تواند نادرست باشد، یا برعکس درباره درستی اندازه‌گیری مان و اینکه چقدر به آن اطمینان داریم اطلاعاتی به دست آوریم. دوم اینکه، اگر همان‌طور که در بالا دیدیم، سه بار قد دخترتان را اندازه بگیرید و میانگین این سه بار را به دست آورید که مثلاً نتیجه آن عدد ۳۵ می‌شود، این عدد احتمالاً خیلی نزدیک‌تر به واقعیت از زمانی است که تنها یک بار قد او را اندازه‌گیری کرده باشید، حالا اگر ۱۰ بار این کار را انجام داده باشید، میانگین آن بسیار نزدیک‌تر می‌شود، بنابراین هرچه بیشتر اندازه‌گیری کنید، حتی با ابزاری که نواقص زیادی دارد، نتیجه تخمین شما از آنچه می‌خواهید اندازه‌گیری کنید، دقیق‌تر خواهد بود.

همین مسئله در مورد آزمون‌های زبان هم وجود دارد، همه ما می‌دانیم که بعضی از آزمون‌های زبان از تعداد زیادی پرسش تشکیل شده‌اند؛ چرا که زبان پیچیده است و مؤلفه‌های زیادی برای اندازه‌گیری دارد، همچنین، اینکه هرچه بیشتر به آزمون‌دهنده‌ها فرصت دهیم تا آنچه می‌دانند را نشان دهند، اندازه‌گیری درست‌تر و عادلانه‌تری خواهیم داشت. از سوی دیگر، راه‌هایی وجود دارد که ببینیم چقدر تغییر یا "کشش" در اندازه‌گیری‌های ما در آزمون‌ها وجود دارد و در این صورت، متوجه می‌شویم که چقدر می‌توانیم به نتایجی که به دست آورده‌ایم، اطمینان داشته باشیم. این بحث و مسائل مرتبط با آن را در فصل ۳ دنبال خواهیم کرد، اما اکنون، نکته‌ای که باید به آن توجه کنیم این است که پرفایده‌ترین آزمون‌های زبان، آن‌هایی هستند که کمترین میزان "کشش" را دارند؛ اگرچه همه آزمون‌ها مانند یک

خط‌کش لاستیکی هستند، باز هم شیوه‌های طراحی آزمونی وجود دارد که کمک می‌کنند تا اطمینان پیدا کنیم، آزمون‌های ما به سطح مشخصی از درستی رسیده‌اند؛ همچنین شیوه‌های تحلیلی وجود دارد که کمک می‌کنند نتایج آزمون را به درستی و با رعایت عدالت تفسیر کنیم.

۴-۱. آزمون، اندازه‌گیری و ارزشیابی

شاید، پیش از اینکه بیشتر در این کتاب جلو برویم، ایده خوبی باشد که نگاهی به بعضی مثال‌ها از انواع قضاوت‌های مدرسان دربارهٔ زبان‌آموزانشان و یادگیری آنها بیاندازیم، ما این کار را در بافتی انجام می‌دهیم که در آن سه مفهوم آزمون^۱، سنجش^۲ و ارزشیابی^۳ را مورد بحث قرار می‌دهیم، چون درک این مفاهیم به ما کمک می‌کند متوجه شویم چرا آزمون زبان می‌گیریم و چطور به تفسیر عادلانه و درستی از عملکرد آزمون‌دهنده‌ها در آن می‌رسیم^۴.

۱-۴-۱. ارزشیابی بدون اندازه‌گیری

همانطور که در ابتدای این فصل گفتم، ممکن است یک معلم، رویه پیشرفت زبان‌آموز خود در کلاس را به شیوه زیر ارزیابی کند:

پائولا واژه‌ها و جمله‌ها را خوب تلفظ می‌کند، او فقط کمی لهجه دارد، اما اطلاعاتش در زمینه دستور به ویژه زمان گذشته خیلی بالا نیست و فهم درستی از جمله‌واره‌های موصولی ندارد. دامنه لغات او برای کسی که سال دوم یادگیری زبان خارجی را می‌گذراند، عالی است.

واضح است که مدرس کیفیت یادگیری پائولا را ارزیابی می‌کند، و همچنین واضح است که هیچ آزمونی در کار نیست، و او هیچ نمره‌ای به این ارزشیابی نمی‌دهد، پس درواقع هیچ سنجش و اندازه‌گیری هم صورت نگرفته است. ارزشیابی درواقع قضاوت‌های کیفی ما دربارهٔ افراد یا وقایع است، ما از سنجش و اندازه‌گیری بهره می‌گیریم تا ما را در ارزشیابی کمک کنند، اما انجام این کار لزومی ندارد.

۲-۴-۱. اندازه‌گیری بدون آزمون

مدرسان طی سال تحصیلی، در بسیاری از موضوع‌های آموزشی، به زبان‌آموزانشان نمره می‌دهند تا روند پیشرفت آنها را به خودشان، والدینشان و مسئولان مدرسه گزارش دهند. برای مثال، مدرس با رجوع به عملکرد زبان‌آموز در کلاس، تکالیف او در خانه و پروژه‌هایی که خارج از کلاس انجام داده،

^۱ Test

^۲ Measurement

^۳ Evaluation

^۴ این بحث براساس کتاب بکمن (۱۹۹۵: ۲۴-۱۸) است.

به‌عنوان منابع اطلاعاتی رجوع می‌کند و بر این اساس به او نمره می‌دهد و در مقایسه با هم‌کلاسی‌هایش رتبه‌بندی می‌کند. این نمره‌ها می‌توانند به صورت الف، ب، پ و ...، یا اعدادی مانند ۵۰ درصد، ۷۰ از ۱۰۰ و غیره باشند، اگرچه رتبه‌بندی زبان‌آموزان در یک مقیاس شامل یک سنجش مبتنی بر ارزشیابی است، بدون اینکه آزمونی وجود داشته باشد.

در این دو مثال، مدرس از تجربه، اطلاعاتی که دربارهٔ زبان‌آموزان در کلاسش پیدا کرده و اطلاعاتی که برای نتیجه‌گیری دربارهٔ فرایند یادگیری زبان‌آموزان جمع‌آوری کرده، استفاده می‌کند تا عملکرد آنها را ارزشیابی و آنها را در مقایسه با هم‌کلاسی‌هایشان یا براساس معیارهای موجود در یک برنامهٔ درسی یا کتاب درسی رتبه‌بندی کند. همانطور که در ابتدای فصل بیان شد، مدرسان باید از این سطح ارزشیابی فراتر بروند تا مطمئن شوند که انصاف و عدالت در برابر همهٔ زبان‌آموزان رعایت می‌شود، قضاوت‌های موجود از روند پیشرفت آنها را تأیید کنند، یا اینکه روند پیشرفت آنها را با معیارهای بیرونی و در طول زمان مقایسه نمایند. برای انجام این کارها به آزمون زبان^۱ نیاز داریم.

۱-۴-۳. آزمون‌ها

به آزمون فرضی درک مطلب خواندن در شکل (۱-۲) نگاه کنید. اگرچه این نمونه آزمون کوتاه است، با این وجود تعدادی از اصول طراحی آزمون زبان را نشان می‌دهد. این آزمون برای ارزیابی درک مطلب خواندن در دانش‌آموزان سال هشتم در انگلستان طراحی شده بود. متن حاضر، یکی از پنج متن آزمون خواندن است، و ۴۰ سوال در کل آزمون وجود دارد. موضوع آن دربارهٔ یک تاجر جوان در شمال انگلستان است که برای زبان‌آموزان ۱۲ ساله مناسب است. زبان مورد استفاده در آن کمی برای این سطح تحصیلی چالش‌برانگیز است: اگرچه بیشتر واژه‌هایی که در آن استفاده شده بین ۲۰۰۰ واژهٔ اول رایج در زبان انگلیسی است، اما تعدادی واژهٔ خارج از این دایره هم دارد، مانند سرمایه^۲، روستایی^۳، پس‌زمینه^۴، پیمانکار^۵، و درآمد دولت^۶.

نزدیک به ۴ درصد از جمله‌های استفاده شده در این متن مجهول هستند و نمرهٔ آن در آزمون خوانش‌پذیری متن^۷ ۸۰۳ است که نشان می‌دهد سطح آن برای آزمون دهنندگان موردنظر، مناسب است. به نظر می‌آید، یکی از دلایل انتخاب متن مذکور، نشان دادن یکی از انواع متون مورد استفاده زبان‌آموزان در کلاس‌های زبان انگلیسی است. درواقع، دیگر متون خواندن در آزمون شامل شعر،

^۱ باید به این نکته اشاره کنم که بعضی اوقات، تمایزی بین واژه‌های آزمون و ارزیابی وجود دارد، آزمون زیرمجموعه‌ای از ارزیابی است.

^۲ asset
^۳ rural
^۴ backdrop
^۵ contractor
^۶ revenue
^۷ Flesch-Kincaid Grade Level