

ارزشیابی نگارش در کلاس آموزش زبان: آشنایی با برخی مفاهیم و روش‌ها

نسیم قنبری

دکترای آموزش زبان انگلیسی، عضو هیئت‌علمی دانشگاه خلیج فارس بوشهر

پریساعبدالرضا پور

دکترای آموزش زبان انگلیسی، عضو هیئت‌علمی دانشگاه سلمان فارسی کازرون

چکیده

این مقاله به بررسی اجمالی ارزشیابی نگارش در کلاس آموزش زبان انگلیسی در ایران، به‌عنوان زبان خارجی (EFL) می‌پردازد. بدین‌منظور در ابتدا برخی مفاهیم کلیدی مرتبط تعریف می‌شوند و سپس معیار نمره‌دهی به‌عنوان یکی از مهم‌ترین عوامل تأثیرگذار در ارزشیابی متن، مورد کنکاش قرار می‌گیرد. توجه به ارزشیابی مهارت نگارش، که شواهد فراوانی اهمیت آن را نشان می‌دهد، نیاز مبرم معلمان و کادر آموزشی را به معیارهای عینی جهت آزمون هرچه دقیق‌تر این مهارت آشکار می‌کند. درواقع آنچه در اکثر محیط‌های ارزشیابی EFL در جریان است، ارزشیابی متن به شکل سلیقه‌ای است که تنها می‌تواند تصویر مبهمی از توانایی نوشتاری فرد به دست دهد. به علاوه، عدم شفافیت در کار و مقایسه‌ناپذیر بودن نتایج چنین ارزشیابی‌هایی در غیاب بازخورد دادن تشخیصی به فراگیر، همگی نیاز به آشنایی بیشتر با مفاهیم نظری ارزشیابی متن و نیز معیار نمره‌دهی و انواع آن را به‌خوبی آشکار می‌سازد.

کلید واژه‌ها

ارزشیابی نگارش، زبان انگلیسی به‌عنوان زبان خارجه (EFL)، نگارش EFL، معیار نمره‌دهی



ارزشیابی نگارش

واقعی تأکید می‌شد. در ایالات متحده نیز در دهه‌های پایانی سال ۱۸۰۰، دانشگاه هاروارد آزمون کتبی را جایگزین آزمون شفاهی کرد. به دنبال چنین تغییراتی در اروپا و آمریکا، نیاز بیشتری به استانداردسازی آزمون‌ها احساس شد که خود موجب توجه بیشتر به آمار و نظریه‌های اندازه‌گیری و نیز یافتن روش‌هایی جهت ارزشیابی توانایی واقعی فراگیران شد. ارزشیابی نگارش کلاس‌محور تا دهه ۱۹۵۰ روش رایج آزمون مهارت نگارش متن به شمار می‌رفت. به دنبال افزایش چشمگیر متقاضیان دانشگاه و نیز آزمون‌های بسندگی بین‌المللی، نیاز به آزمون‌هایی که از سطح بالایی از ویژگی‌های روان‌سنجی برخوردار باشند افزایش یافت. بنابراین، آزمون‌های غیرمستقیم نگارش (Grabe & Kaplan, 1996) که ویژگی‌های خاص لغوی و دستوری متن را مورد ارزیابی قرار می‌دادند، به وجود

مهارت نوشتن، که زمانی قلمرو نخبگان و افراد بسیار تحصیل کرده یک جامعه به‌شمار می‌رفت، امروز به ابزار اصلی ارتباطی افراد در تمامی شئون زندگی تبدیل شده است (Weigle, 2002). همگان اذعان دارند که نوشتن نه تنها اطلاعات را منتقل می‌کند، بلکه نقش بسزایی در تغییر ماهیت دانش و خلق دانش جدید دارد. از این منظر، نوشتن، اهمیت اساسی در فراگیری زبان در سراسر جهان دارد. بنابراین، هرگاه یادگیری مهارت زبانی خاصی، مهم جلوه کند، به همان میزان ارزشیابی آن نیز حائز اهمیت می‌شود و مهارت نگارش نیز از این قاعده مستثنی نیست. بر این اساس، همگام با اهمیت فزاینده نگارش در مبحث آموزش زبان‌های خارجی، نیاز جدی‌تری به روش‌های پایا و با روایی بالا جهت آزمون مهارت نگارش در کلاس آموزش زبان شکل می‌گیرد.

تاریخچه ارزشیابی نگارش

سابقه ارزشیابی نگارش به دوره امپراتوری چو (بین سال‌های ۱۱۱۱ تا ۷۷۱ قبل از میلاد مسیح) در چین برمی‌گردد (2001 Hamp-Lyons). در آن زمان چندین مصحح به کار گرفته می‌شد تا از پایایی شیوه انتخاب مقامات کشوری، که براساس آزمون‌های کتبی بود، اطمینان حاصل کنند. در اروپای استعماری نیز به دلیل نیاز به فرستادن افراد باسواد به کشورهای مختلف، بر ارزشیابی مستقیم و بلاواسطه نگارش متن براساس متن‌های

به دنبال گرایش نگرش‌های ارتباطی، در آموزش و ارزشیابی زبان، در اواخر دهه هفتاد و اوایل دهه هشتاد میلادی، ارزشیابی بلاواسطه و مستقیم نگارش، یا همان ارزشیابی عملکرد نگارش، در محیط‌های آموزشی زبان اول و دوم مورد توجه قرار گرفت



آمدند. به‌عنوان نمونه، آزمون کتبی انگلیسی استاندارد (Test of Standard Written English) در مؤسسه ETS

جهت فراگیران زبان اول تولید شد.

به دنبال گرایش نگرش‌های ارتباطی، در آموزش و ارزشیابی زبان، در اواخر دهه هفتاد و اوایل دهه هشتاد میلادی، ارزشیابی بلاواسطه و مستقیم نگارش، یا همان ارزشیابی عملکرد نگارش، در محیط‌های آموزشی زبان اول و دوم مورد توجه قرار گرفت. در این شیوه از فراگیران خواسته می‌شد که متنی را در مدت زمان ۳۰-۴۵ دقیقه بنویسند؛ سپس یک مصحح آن را با استفاده از یک معیار ارزیابی، به شکل جزئی یا کلی تصحیح می‌کرد (Cumming, 1997). علاوه بر آزمون‌های بلاواسطه نگارش در کلاس، آزمون‌های مهم و استاندارد هم‌چون آلتس و تافل نیز این شیوه ارزشیابی نگارش را اتخاذ کردند. اکنون در کنار این گونه آزمون‌های بسندگی بین‌المللی که توانایی کلی نگارش افراد را می‌سنجند، اهداف آموزشی مهم دیگری برای ارزشیابی مستقیم نگارش وجود دارد. برای مثال، آزمون‌های پایانی کلاس که پیشرفت فراگیر را در یک دوره تحصیلی می‌سنجند و یا آزمون‌های تشخیصی که به‌منظور شناسایی نقاط ضعف و قوت توانایی نگارشی به‌کار می‌روند، از این دسته‌اند.

در فضای ارزشیابی عملکرد، معیار نمره‌دهی به‌عنوان تعریف عملیاتی سازه مورد نظر، یک سری از سطوح سازمان یافته را فراهم می‌کند که براساس آن عملکرد فراگیران با یکدیگر مقایسه می‌شود

سنت آزمون-نمره که نشئت گرفته از مقاله ارزشمند اسپیرمن (Spearman, 1904) است، بر خطای اندازه‌گیری و تجزیه نمره عینی به دو عنصر نمره واقعی و مقادیر خطا، متمرکز است. این سنت تحت لوای نظریه کلاسیک آزمون، موجبات پیدایش برخی نظریه‌های مهم همچون نظریه تعمیم‌پذیری (Brennan, 1997; Cronbach, Gleser, Nanda, & Rajaratnam, 1972)، تحلیل عاملی و مدل معادلات ساختاری، گردید (Joreskog, 2007).

دیگر نظریه تأثیرگذار، نظریه معیار-محور، به واسطه سيطرة علم روان-جسمانی قرن نوزدهم پدیدار شد. تمرکز سنت معیار-محور، فراهم آوردن نقشه‌هایی متغیر است که جایگاه افراد و آیت‌ها را روی یک معیار متغیر نهانی، مشخص می‌کند. بر اساس گفته معروف ثوندایک که هر آنچه وجود دارد تا اندازه‌ای وجود دارد (Clifford, 1984) سنت معیار-محور، معیارهای ارزشیابی فراوانی را که ریشه در نظریه گویه-پاسخ دارد، عرضه کرده است.

علیرغم اقبال عمومی به ارزشیابی مستقیم نگارش، دغدغه‌هایی در باب اعتبار محتوایی و سازه‌ای این گونه ارزشیابی وجود دارد. درواقع، این اعتباریابی در صدر تحقیقات مرتبط در این زمینه بوده است. در این میان عوامل متعددی همچون مصحح، نوع فعالیت مدنظر نگارشی، معیار نمره‌دهی و ... وجود دارد. در ادامه، معیار نمره‌دهی به‌عنوان برجسته‌ترین تفاوت این گونه از ارزشیابی نگارش با دوره‌های پیشین مورد بررسی قرار خواهد گرفت.

معیار نمره‌دهی و ارزشیابی نگارش

بررسی تاریخی نظریه‌های سنجش مرتبط با ارزشیابی نگارش در قرن بیستم، دو سنت غالب را نشان می‌دهند: سنت آزمون-نمره و سنت معیار-محور (Behizadeh & Engelhard, 2011).



نگارش، برخی از مهم‌ترین انواع این معیارها در ذیل آورده می‌شود:

انواع معیارهای نمره‌دهی

فارغ از نوع خاص معیار نمره‌دهی، رایج‌ترین انواع این معیارها عبارت‌اند از معیارهای کلی‌نگر، معیارهای جزئی‌نگر، معیارهای ویژگی اصلی و معیارهای چندویژگی (Hamp-Lyons, 1991; Hyland, 2003 and Weigle, 2002). همگی این معیارها در محیط‌های EFL/ESL جهت شناسایی سطوح بسندگی فراگیران به منظور اهداف مختلف به کار گرفته شده‌اند. براساس (Weigle, 2002) دو ویژگی عمده این چهار معیار نمره‌دهی را از هم متمایز می‌کند: نخست اینکه آیا معیار مختص یک فعالیت نگارشی خاص است و یا قابل تعمیم به گروهی از فعالیت‌هاست، و دوم اینکه آیا یک نمره واحد و یا چند نمره به هر نوشته اختصاص می‌یابد (جدول ۲، ۴). این چهار شیوه نمره‌دهی در ادامه به تفصیل مورد بحث قرار می‌گیرند.

در فضای ارزشیابی عملکرد، معیار نمره‌دهی به‌عنوان تعریف عملیاتی سازه مورد نظر، یک سری از سطوح سازمان‌یافته را فراهم می‌کند که براساس آن عملکرد فراگیران با یکدیگر مقایسه می‌شود. دیویس و همکاران (Davies et al, 1999) ساختار معیار نمره‌دهی را این‌گونه شرح می‌دهند: «غالباً این‌گونه معیارها نمایانگر تسلط صفر تا غایت تسلط (هم‌سطح فرد بومی‌زبان) است. سطوح تعریف‌شده در این معیارها براساس آنچه افراد می‌توانند با زبان انجام دهند (مانند فعالیت‌ها و عملکردهای قابل انجام) و نیز تسلط آن‌ها بر ویژگی‌های زبانی (مانند لغات، دستور زبان، روانی و انسجام کلام) است. معیارهای نمره‌دهی دربرگیرنده پاره‌ای از رفتارهای غالباً رایج هستند. آن‌ها به خودی خود، آزمون به شمار نمی‌آیند بلکه می‌بایست به همراه آزمون‌های هدفمند به کار گرفته شوند. مصححین معمولاً جهت استفاده از این معیارها، آموزش می‌بینند و از این نظر پایایی معیار تأمین می‌شود.» (P. 153) نظر به اهمیت فراوان معیارهای نمره‌دهی در ارزشیابی

جدول شماره ۱: انواع معیارهای نمره‌دهی در ارزشیابی نگارش (برگرفته از Weigle, 2002)

مربوط با یک گونه فعالیت نگارشی خاص	قابل تعمیم به دسته‌ای از فعالیت‌های نگارشی	نمره واحد
ویژگی اصلی	کلی‌نگر	
چندین ویژگی	جزیی‌نگر	چندین نمره

نمره‌دهی کلی‌نگر

در رابطه با معایب نمره‌دهی کلی‌نگر، چهار نقد عمده مطرح شده است: نخست اینکه این شیوه نمره‌دهی، پروفایلی از اطلاعات نگارشی فراگیران فراهم نمی‌آورد. به عبارت دیگر، یک نمره واحد، اطلاعات تشخیصی مفیدی در مورد جنبه‌های مختلف نگارش مانند تسلط بر دستور زبان، میزان آشنایی با واژگان، سازماندهی متن و غیره را به دست نمی‌دهد. ویگل (2002) نیز بر این نکته تأکید می‌کند که تصویر غیرواقعی که نمره‌دهی کلی‌نگر فراهم می‌آورد، خصوصاً برای فراگیران زبان دوم دردساز است، زیرا نویسندگان مختلف با سرعت‌های متفاوتی در بخش‌های مختلف نگارش پیشرفت می‌کنند. برای مثال، برخی نویسندگان، مهارت بالایی در محتوا و سازماندهی متن دارند اما ممکن است تسلط آن‌ها بر دستور زبان به همان اندازه نباشد. دسته‌ای دیگر از نویسندگان به خوبی با ساختار جمله آشنایی دارند اما ممکن است در سازماندهی منطقی متن ضعف داشته باشند.

در این شیوه نمره‌دهی، توجه به ویژگی‌های زبانی، کلامی و محتوایی متن کنار گذاشته می‌شود. همچنین این گونه معیار نمره‌دهی عمدتاً در ارزشیابی‌های کلان مورد استفاده قرار می‌گیرد (Cooper, 1997). روال رایج نمره‌دهی کلی‌نگر به این شکل است که متن به سرعت خوانده می‌شود و سپس با معیار نمره‌دهی که دربرگیرنده شاخص‌های نمره‌دهی است، مورد مقایسه قرار می‌گیرد. معمولاً معیار مورد نظر با چند متن نمونه که نمایانگر معیارهای ارزشیابی مورد نظر در هر سطح هستند، همراه است. همچنین مصححین نیز به خوبی آموزش می‌بینند که چگونه در هنگام تصحیح متن از این معیارها استفاده کنند. به‌عنوان یکی از معروف‌ترین نمونه‌های معیار نمره‌دهی کلی‌نگر در فضای ESL می‌توان از معیار آزمون نگارش تافل نام برد. این معیار در برگیرنده توصیف‌گرهای دستوری و کلامی در شش سطح بسندگی نگارش است.

درواقع، مهم‌ترین برتری نگارش کلی‌نگر که به خوبی کاربرد فراوان آن در فضاهای مختلف را توجیه می‌کند، عملی بودن آن است

همچنین از آنجا که مصححین معیار/ معیارهای واحدی را جهت نمره‌دهی به کار نمی‌برند، تفسیر نمرات در این شیوه نمره‌دهی، مشکل است. برای مثال، یک مصحح ممکن است براساس جنبه‌های کلامی (محتوا، سازماندهی، بسط) به متنی نمره ۴ بدهد، و مصحح دیگر براساس جنبه‌های زبانی متن (تسلط بر دستور زبان و واژگان) همان نمره ۴ را بدهد. از دیگر نقدهای وارد شده به این شیوه نمره‌دهی همبستگی بالای آن با ویژگی‌های ظاهری متن مانند دستخط نویسنده و اندازه متن است (Weigle, 2002). همچنین تأکید بیش از حد بر همبستگی بالای بین مصححین که به بهای نادیده انگاشته شدن روایی متن است، از دیگر مشکلات در این زمینه است. به‌عنوان نمونه‌ای از یک معیار نمره‌دهی کلی‌نگر، معیار آزمون انگلیسی کتبی تافل در ذیل آورده می‌شود.

درواقع، مهم‌ترین برتری نگارش کلی‌نگر که به خوبی کاربرد فراوان آن در فضاهای مختلف را توجیه می‌کند، عملی بودن آن است. (White, 1984, 1985) که خود از پیشروان استفاده از نمره‌دهی کلی‌نگر است، مزایای متعددی را برای آن برمی‌شمرد: از دید او، یک بار خواندن متن و نمره‌دهی آن سریع‌تر از چند بار خواندن و در هر بار بخش متفاوتی را نمره دادن است. همچنین از آنجا که نمره‌دهی کلی‌نگر توجه خواننده متن را به نقاط قوت متن و نه ضعف‌های آن جلب می‌کند، نویسندگان را در سطح بالاتری از عملکرد ارزیابی می‌کند. به علاوه این روش نمره‌دهی می‌تواند به گونه‌ای طراحی شود که توجه خواننده را به بخش‌های خاصی از متن - که در یک فضای خاص مهم هستند - معطوف کند و از این طریق اطلاعات مهمی را در مورد آن‌ها به صورت کارآمدی ارائه کند. همچنین (White, 1984, 1985) ادعای بحث‌برانگیزی را مطرح می‌کند. او بر این باور است که «نمره‌دهی کلی‌نگر از نمره‌دهی جزئی‌نگر اعتبار بیشتری دارد زیرا بازتاب واکنش شخصی و واقعی فرد خواننده به متن است. از دید او، توجه بیش از حد به اجزای متن در نمره‌دهی جزئی‌نگر، معنای کلی را منسوخ می‌کند. (White, 1984, p. 409)

جدول شماره ۲: معیار نمره‌دهی کلی نگر آزمون تافل (برگرفته از Weigle, 2002)

- 6 An essay at this level
- effectively addresses the writing task.
 - is well organized and well developed.
 - uses clearly appropriate details to support a thesis or illustrate ideas.
 - displays consistent facility in use of language.
 - demonstrates syntactic variety and appropriate word choice though it may have occasional errors.
- 5 An essay at this level
- may address some parts of the task more effectively than others.
 - is generally well organized and developed.
 - uses details to support a thesis or illustrate ideas.
 - displays facility in use of language.
 - demonstrates syntactic variety and appropriate word choice though it may have occasional errors
- 4 An essay at this level
- addresses the writing topic adequately but may slight parts of the task.
 - is adequately organized and developed.
 - uses some details to support a thesis or illustrate ideas.
 - demonstrates adequate but possibly inconsistent facility with syntax and usage.
 - may contain some errors that occasionally obscure meaning.
- 3 An essay at this level is seriously flawed by one or more of the following weaknesses:
- Inadequate organization or development
 - Inappropriate or insufficient details to support or illustrate generalizations
 - a noticeably inappropriate choice of words or word forms
 - an accumulation of errors in sentence structure and /or usage
- 2 An essay at this level is seriously flawed by one or more of the following weaknesses:
- Serious disorganization
 - Little or no detail, or irrelevant specifics
 - Serious and frequent errors in sentence structure or usage
 - Serious problems with focus
- 1 An essay at this level
- may be incoherent.
 - may be undeveloped.
 - may contain severe and persistent writing errors.
- 0 A paper is rated 0 if it contains no response, merely copies the topic, is off-topic, is written in a foreign language, or consists of only keystroke characters.

نمره‌دهی جزئی نگر

و نیز روایی سازه‌ای بیشتری برخوردار است (Weigle, 2002). یکی از نمونه‌های نمره‌دهی جزئی نگر، معیار (ESL Composition Profile, Jacobs, et al., 1981) است. در این معیار خاص، متن براساس پنج ویژگی نگارشی

در این نوع از نمره‌دهی، بخش‌های مختلفی از متن مانند محتوا، سازماندهی، مکانیک و غیره در نظر گرفته می‌شود. به همین دلیل نمره‌دهی جزئی نگر از پایایی

مورد قضاوت قرار می‌گیرد: محتوا، سازماندهی، واژگان، کاربرد زبان و مکانیک. این پنج جنبه، وزن نمره‌ای متفاوتی دارند به‌گونه‌ای که محتوا با ۳۰ نمره بالاترین تأکید را داراست و سپس کاربرد زبان (۲۵ نمره) سازماندهی و واژگان (هر کدام ۲۰ نمره) و مکانیک (با ۵ نمره) در سطوح مختلف این معیار جای دارند. این معیار به‌طور گسترده‌ای در برنامه‌های آموزشی مرتبط با نگارش، مورد اقبال واقع شده است. به‌عنوان نمونه دیگری از نمره‌دهی جزئی‌نگر، می‌توان از معیار نمره‌دهی (Hamp-Lyons Michigan Writing Assessment Scoring Guide, 1990; 1991) نام برد. این معیار جهت ارزشیابی آزمون نگارش ورودی دانشگاه به کار می‌رود و دربرگیرنده سه نوع معیار نمره‌دهی است: بحث و نظرات، ویژگی‌های کلامی و تسلط بر زبان. ویژگی برجسته این معیار این است که از درون کلاس‌های نگارش دانشگاه و با همراهی و مشورت با آن‌ها به وجود آمده است و متن خوب را از منظر چنین نگاهی، تعریف می‌کند. در نتیجه، این معیار مورد استقبال فراگیران، مشاوران و دانشگاهیان و دیگر افراد قرار گرفته است.

مزیت اصلی نمره‌دهی جزئی‌نگر، فراهم آوردن اطلاعات دقیق از توانایی فراگیران در رابطه با ویژگی‌های مختلف متن است و به همین منظور، این معیار مناسب اهداف تشخیصی است (Knoch, 2009). همچنین، برخی تحقیقات (Fancis & Adams, 1977, 1981) نشان داده‌اند که معیار نمره‌دهی جزئی‌نگر از آنجا که موجب می‌شود مصححین تازه‌کار به‌راحتی معیارهای نمره‌دهی را درک کنند و به کار ببرند، فایده بیشتری دارند. علاوه بر این، نمره‌دهی جزئی‌نگر برای فراگیران

زبان دوم که توانایی‌هایی متفاوتی در بخش‌های مختلف متن دارند، سود بیشتری دارد. و نهایتاً نمره‌دهی جزئی‌نگر پایایی بیشتری نسبت به نمره‌دهی کلی‌نگر دارد. همان‌طور که با افزایش تعداد گویه‌ها به آزمون، پایایی آن افزایش می‌یابد، معیار نمره‌دهی که چند نوع نمره را به یک متن اختصاص می‌دهد نیز موجب افزایش پایایی آن می‌شود.

در کنار مزایای قبل، این شیوه نمره‌دهی از جنبه‌هایی هم مورد نقد واقع شده است. نخست اینکه شیوه نمره‌دهی جزئی‌نگر نسبت به نمره‌دهی کلی‌نگر بسیار زمان‌بر است، زیرا مصححین می‌بایست بیش از یک بار به هر متن نمره دهند و زمان بیشتری را صرف نمره‌دهی کنند. نقد دیگر این است که اگر نمرات بخش‌های مختلف با هم جمع شوند و یک نمره کل را به‌وجود بیاورند، تا میزان زیادی اطلاعاتی که نمره‌دهی جزئی‌نگر به‌دست می‌دهد، از بین می‌رود.

علاوه بر این، نمره‌دهی جزئی‌نگر نیازمند این مهم است که معیارهای خاصی ایجاد و نوشته شوند و مصححین نیز آموزش ببینند و در جلسات آموزشی فراوان شرکت کنند. چنین جلساتی به دنبال کاستن از اختلاف نمره‌دهی بین مصححین و در نتیجه افزایش روایی است. همچنین از آنجا که مصححین به دنبال جنبه‌های خاصی از متن هستند، سطح نمرات نسبت به نمرات مشابه در نمره‌دهی کلی‌نگر پایین‌تر است. آزمون‌گرها همواره هشدار می‌دهند که معیارهای نمره‌دهی جزئی‌نگر، ماهیت یکپارچگی مهارت‌ها را در ارزشیابی نگارش نادیده می‌گیرند. از سوی دیگر، معلمین نیز بر این باورند که نگارش بزرگ‌تر از مجموع اجزای آن است (Coombe & Wiens, 1999).



جدول شماره ۳: معیار نمره‌دهی جزئی‌نگر (Jacobs et al., 1981) ESL Composition Profile

ESL Composition Profile			
Student	Date	Topic	
Content	30-27	Excellent to very good: knowledgeable, substantive, thorough development of thesis, relevant to assigned topic	
	26-22	Good to average: some knowledge of subject, adequate range, limited development of thesis, mostly relevant to topic but lacks detail	
	21-17	Fair to poor: limited knowledge of subject, little substance, inadequate development of ideas.	
	16-13	Very poor: does not show knowledge of subject, non-substantive, not pertinent, or not enough to evaluate	
Organization	20-18	Excellent to very good: fluent expression, ideas clearly stated/supported, succinct, well organized, logical sequencing, cohesive.	
	17-14	Good to average: somewhat choppy, loosely organized but main ideas stand out, limited support, logical but incomplete sequencing	
	13-10	Fair to poor: non-fluent, ideas confused or disconnected, lacks logical sequencing and development	
	9-7	Very poor: does not communicate, no organization, or not enough to evaluate	
Vocabulary	20-18	Excellent to very good: sophisticated range, effective word/idiom choice and usage, word form mastery, appropriate register	
	17-14	Good to average: adequate range, occasional errors of word/idiom form, choice, usage but meaning not obscured	
	13-10	Fair to poor: limited range, frequent errors of word/idiom form, choice, usage, meaning confused or obscured	
	9-7	Very poor: essentially translation, little knowledge of English vocabulary, idioms, word form, or not enough to evaluate	
Language use	25-22	Excellent to very good: effective complex constructions, few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions	
	21-18	Good to average: effective but simple constructions, minor problems in complex constructions, several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions but meaning seldom obscured	
	17-11	Fair to poor: major problems in simple/complex constructions, frequent error of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, deletions, meaning confused or obscured	
	10-5	Very poor: virtually no mastery of sentence construction rules, dominated by errors, does not communicate, or not enough to evaluate	
Mechanics	5	Excellent to very good: demonstrates mastery of conventions, few errors in spelling, punctuation, capitalization and paragraphing	
	4	Good to average: occasional errors of spelling, punctuation, capitalization, paragraphing but meaning not obscured	
	3	Fair to poor: frequent errors of spelling, punctuation, capitalization, paragraphing, poor handwriting, meaning confused and obscured	
	2	Very poor: no mastery of conventions, dominated by errors of spelling, punctuation, capitalization, paragraphing, handwriting illegible, or not enough to evaluate	
Total score	Reader	Comments	

نمره‌دهی ویژگی اصلی

مانند گونه توضیحی یا ترغیبی) بنویسند. در نمره‌دهی ویژگی اصلی، معیار نمره‌دهی بر اساس نوع فعالیت خاص نگارشی تعریف می‌شود و متن فراگیران بر پایه میزان موفقیت آنان در انجام آن، مورد قضاوت قرار می‌گیرد. در واقع، در این شیوه، برای هر فعالیت نگارشی، معیاری تعریف می‌شود که دربرگیرنده این معیارهاست:

این شیوه نمره‌دهی با نام لوید جونز (1977) پیوند خورده است. جونز این شیوه را برای ارزشیابی ملی پیشرفت آموزشی (NAEP) که دربرگیرنده آزمون‌های مدارس ایالات متحده است، ابداع کرد. این گونه نمره‌دهی بر این فلسفه استوار است که فراگیران نهایتاً باید بتوانند در یک گونه خاص گفتگامانی



- فعالیت نگارشی
- بیان ویژگی کلامی اصلی مورد نظر فعالیت (برای مثال، متن انگیزشی، نامه تبریک)
- فرضیه در مورد عملکرد مورد انتظار
- بیان ارتباط بین فعالیت و ویژگی اصلی
- معیار نمره‌دهی که سطوح عملکرد را نشان می‌دهد
- نمونه متن در هر سطح
- توضیح در باب چگونگی تصحیح هر کدام از متن‌ها

ناگفته پیداست که چون برای هر فعالیت نگارشی می‌بایست راهنمای نمره‌دهی تعریف و ایجاد شود، نمره‌دهی ویژگی اصلی بسیار وقت‌گیر و خسته‌کننده است. به‌علاوه این روش چندان در برنامه‌های آموزشی یادگیری زبان دوم مورد استفاده واقع نشده است که این خود از رغبت به این روش می‌کاهد. با وجود این، روش نمره‌دهی ویژگی اصلی به‌طور خاص مناسب فراگیران زبان دوم در فضای مدرسه است، زیرا والدین که خود تسلط کافی به زبان مدرسه ندارند، می‌توانند از توصیف آنچه فرزندان‌شان به واسطه زبان انجام می‌دهند، بهره ببرند.

جدول شماره ۴: نمونه‌ای از یک معیار نمره‌دهی ویژگی اصلی (برگرفته از Tedick, 2002)

Prompt: persuading an audience

Write a persuasive letter to the editor of school newspaper...

Primary Trait scale:

0. The writer fails to persuade the audience.
1. The writer attempts to persuade but does not provide sufficient support.
2. The writer presents a somewhat persuasive argument but without consistent development and support

The writer develops a persuasive argument that is well developed and supported

نمره‌دهی چندویژگی

معیارهای جزئی‌نگر معمولاً به جنبه‌های سنتی و کلی‌تر متن می‌پردازند، اما معیارهای چندویژگی به ویژگی‌های خاص نگارش که مورد نیاز نگارش یک متن خاص هستند، توجه می‌کنند.

برخی از مزایای این شیوه خاص نمره‌دهی عبارتند از هماهنگی و ارتباط بیشتر معیارهای نمره‌دهی با فعالیت‌ها و برنامه درسی که این خود موجب افزایش روایی محتوایی و سازه‌ای ارزشیابی معیار-محور می‌گردد (McNamara, 1996). همچنین به دلیل ماهیت خاص فعالیت مورد نظر،

روش نمره‌دهی چندویژگی که بر اساس مفاهیم نمره‌دهی ویژگی اصلی طراحی شده است، به دنبال فراهم آوردن بازخورد تشخیصی به فراگیران و دیگر افراد مرتبط در مورد عملکرد نویسندگان بر مبنای معیارهای متناسب با فضا و فعالیت مورد نظر جهت یک متن خاص است (Hamp-Lyons, 1991).

معیارهای چندویژگی از آنجا که به بررسی عملکرد در چند دسته می‌پردازند، شباهت زیادی به معیارهای جزئی‌نگر دارند و گاهی این دو به جای یکدیگر به کار می‌روند. با وجود این،

ناگفته پیداست که چون برای هر فعالیت نگارشی می‌بایست راهنمای نمره‌دهی تعریف و ایجاد شود، نمره‌دهی ویژگی اصلی بسیار وقت‌گیر و خسته‌کننده است

هم معلم و هم فراگیر فعالانه در فرایند بازخورد حضور دارند. از معایب اصلی این روش نمره‌دهی (مشابه روش نمره‌دهی ویژگی اصلی) عدم تعمیم عملکرد نگارشی فراگیران به دیگر فعالیت‌ها است.



جدول شماره ۵: نمونه‌ای از یک معیار نمره‌دهی چند ویژگی (برگرفته از Hyland, 2003)

Score	Content	Structure	Language
4	<ul style="list-style-type: none"> Events explicitly stated Clearly documents events Evaluates their significance Personal comment on events 	<ul style="list-style-type: none"> Orientation gives all essential information All necessary background provided Account in chronological/other order Reorientation 'rounds off' sequence 	<ul style="list-style-type: none"> Excellent control of language Excellent use of vocabulary Excellent control of grammar Appropriate tone and style
....			
1	<ul style="list-style-type: none"> Event not stated Not recognizable events No or confused evaluation No or weak personal comment 	<ul style="list-style-type: none"> Missing or weak instruction No background provided Haphazard and incoherent sequencing No reorientation or includes new matter 	<ul style="list-style-type: none"> Little language control Reader seriously distracted by grammar errors Poor vocabulary and tone

ESL Tedick Hyland

نتیجه گیری

اهمیت روزافزون نگارش به زبان دوم از سویی و فضای مبهم موجود در زمینه ارزشیابی دقیق و معتبر این مهارت از سوی دیگر، نیازمند آشنا ساختن هر چه بیشتر معلمان زبان در این زمینه است. مطالب ذکر شده در این مقاله را نیز باید تلاشی در این راستا قلمداد کرد. بی تردید اقدامات دیگری همچون برگزاری کارگاه‌های آموزشی که معلمان را به صورت عملی درگیر مقوله ارزشیابی خواهد کرد و اصلاح سیاست‌های آموزشی با تأکید بر اهمیت گونه کتبی ارتباطات در فضای مدرسه و دانشگاه نیز ضرورت دارد. در سایه چنین اقداماتی می‌توان امیدوار بود که مهارت به غفلت سپرده شده نگارش جایگاه خود را در برنامه‌های درسی آموزش و یادگیری زبان بیابد.

References

- Academic Publishers.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of Language Testing*. Cambridge: Cambridge University Press.
- Fancis, J. C. (1977). *Impression and analytic marking methods*, unpublished paper, Aldershot: Associated Examining Board.
- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing*. New York: Longman.
- Hamp-Lyons, L. (1990). Second language writing: Assessment issues. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom*. New York: Cambridge University Press.
- Hamp-Lyons, L. (1991). Reconstructing "academic writing proficiency". In: L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 127-154). Norwood NJ: Ablex.
- Hamp-Lyons, L. (2001). *Ethics, fairness(es) and developments in language testing*. Studies in Language Testing 9. Cambridge: Cambridge University Press, pp. 30-34.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Jacobs, H. L., Zinkgraf, S.A., Wormouth, D.R., Hartfiel, V. F. and Hughey, J. B. (1981). *Testing ESL composition: a practical approach*. Rowley, MA: Newbury House.
- Joreskog, K. G. (2007). Factor analysis and its extensions. In: R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions*. Mahwah, NJ: Erlbaum.
- Knoch, U. (2009). *Diagnostic assessment of writing: A comparison of two rating scales*. Language Testing, 26 (2), 275-304.
- Lloyd-Jones, R. (1977). Primary trait scoring. In C. R. Cooper & L. Odell (Eds.), *Evaluating writing*. New York: National Council of Teachers of English.
- McNamara, T. (1996). *Measuring second language performance*. Harlow: Longman.
- Spearman, C. (1904). "General intelligence," objectively determined and measured. In Behizadeh, N. & Engelhard, G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment in the United States. *Assessing writing*, 16 (3), 189-211.
- Tedick, D.J. (2002). Proficiency-oriented language instruction and assessment: Standards, philosophies, and considerations for assessment. In Minnesota Articulation Project, D. J. Tedick (Ed.), *Proficiency-oriented language instruction and assessment: A curriculum handbook for teachers* (Rev Ed.). CARLA Working Paper Series. Minneapolis, MN: University of Minnesota, The Center for Advanced Research on Language Acquisition.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge, UK: Cambridge University Press.
- White, E. M. (1985). *Teaching and assessing writing*. San Francisco: Jossey-Bass Inc.
- Adams, R. (1981). The reliability of marking of five June 1980 examinations, mimeographed paper, Guildford: Associated Examining Board.
- Behizadeh, N. & Engelhard, G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment in the United States. *Assessing writing*, 16 (3), 189-211.
- Brennan, R. L. (1997). A perspective on the history of generalizability. *Educational Measurement: Issues and Practice*, 16 (10), 14-20.
- Clifford, G. J. (1984). Edward L. Thorndike: The sane positivist. In Behizadeh, N. & Engelhard, G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment in the United States. *Assessing writing*, 16(3), 189-211.
- Coombe, C. & Wiens, B. (1999). Writing assessment: which marking scheme is right for you? Paper presented at the first Annual Egypt TESOL Conference, November 1999. Retrieved from:
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). The dependability of behavioral measurements: Theory of generalizability of scores and profiles. In Behizadeh, N. & Engelhard, G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment in the United States. *Assessing writing*, 16(3), 189-211.
- Cumming, A. (1997). The testing of writing in a second language. In C. Clapham & D. Corson (Eds.), *Encyclopedia of Language and Education* (Vol. 7: Language Testing and Assessment). Dordrecht: Kluwer