

تاثیر خودآزمایی مهارت نگارش به‌روشنای سازنده و پایان‌محور بر عملکرد نگارش و دقت خودآزمایی زبان‌آموزان ایرانی

مسعود سیری*

استادیار آموزش زبان انگلیسی واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

ستاره قربان‌دایی**

کارشناسی ارشد آموزش زبان انگلیسی واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران

(تاریخ دریافت: ۹۵/۰۸/۱۲، تاریخ تصویب: ۹۶/۰۶/۱۴، تاریخ چاپ: مهر ۱۳۹۶)

چکیده

هدف از خودآزمایی به‌صورت سازنده و یا پایان‌محور می‌تواند در میزان دقت نمره‌دهی خودآزمایی تاثیر بگذارد، بنابراین این پژوهش به‌بررسی اینکه آیا خودآزمایی به‌روشنای سازنده و پایان‌محور تاثیرات متفاوتی بر نگارش و میزان دقت نمره‌دهی زبان‌آموزان می‌گذارد یا خیر می‌پردازد. بدین منظور، ۶۰ زبان‌آموز سطح میانی به دو گروه سازنده و پایان‌محور به‌صورت تصادفی تقسیم شدند، و یک پیش‌آزمون برای سنجش فعلی آن‌ها در نگارش پاراگراف اجرا شد. در گروه سازنده، نمره خودآزمایی زبان‌آموزان از میانگین نمره معیار دو مصحح تفریق می‌شد تا میزان خطا در خودآزمایی مشخص شود. در گروه پایان‌محور شرایط مشابهی وجود داشت اما میزان دقت در انجام خودآزمایی به‌عنوان بخشی از نمره پایان دوره در نظر گرفته شد. بعد از اجرای پس‌آزمون، نتایج بررسی کواریانس نشان داد که زبان‌آموزان هر دو گروه بهبود معنادار و برابری در کیفیت عملکرد نگارش داشتند، اما زبان‌آموزان گروه پایان‌محور افزایش بیشتری در میزان دقت خودآزمایی نشان دادند.

واژه‌های کلیدی: سنجش جایگزین، خودآزمایی، سنجش سازنده، سنجش پایان‌محور، عملکرد نگارش، دقت سنجش.

* E-mail: m.siyari@srbiau.ac.ir

** E-mail: setare.daei@yhaoo.com

مقدمه

سنجش آموزش محور^۱ بر خلاف سنجش‌های مبتنی بر اصول روان‌سنجی^۲ سعی بر آن دارد تا در کنار سنجش و ارزشیابی فرصت یادگیری و آموزش را نیز به‌یادگیرنده‌ها بدهد (گیپس، ۱۹۹۴؛ براون، ۱۹۹۸). برای دستیابی به این هدف، روش‌های سنجش جایگزین^۳ می‌توانند راهگشا باشند. این روش‌ها شامل استفاده از ویدیو و صوت، مشاهدات مستمر مدرسان از کلاس‌ها، یادداشت برداری و وقایع‌نگاری توسط یادگیرنده‌ها، پورتفولیو، خودآزمایی^۴ و هم‌آزمایی^۵ است (براون، ۱۹۹۸؛ براون و هادسون، ۱۹۹۸؛ مکی، ۲۰۰۶). از بین این روش‌ها، خودآزمایی و هم‌آزمایی در سال‌های اخیر توجه بسیاری را در زمینه آموزش و پرورش و به‌طور خاص زمینه آموزش زبان خارجه به‌سوی خود جلب کرده‌اند. این بدان دلیل است که امروزه روش‌های نوین آموزش و پرورش بر استقلال^۶ یادگیرنده تاکید بسیاری دارند. در کنار این انگیزه، خودآزمایی و هم‌آزمایی خود می‌توانند اثرات مثبتی بر یادگیری افراد داشته باشند. به‌عنوان مثال، براون و هادسون (۲۰۰۲) معتقدند که افزون بر صرفه‌جویی در وقت کلاس، با استفاده از خودآزمایی و هم‌آزمایی، یادگیرنده‌ها به‌طور فعال و معناداری درگیر روند آموزش و یادگیری می‌شوند، که این خود به‌افزایش انگیزه یادگیرنده‌ها و استقلال آن‌ها کمک می‌کند (اسکارسون، ۱۹۸۹). تاپینگ (۲۰۰۳) نیز خاطرنشان کرده‌است که خودآزمایی و هم‌آزمایی فعالیت‌هایی هستند که از لحاظ پردازش و شناختی پیچیده هستند، و این امر باعث می‌شود که یادگیرنده‌ها در مورد روند یادگیری از خود سوال پرسیده و به آن فکر کنند که این خود انگیزه‌ای است تا آن‌ها مسئولیت و مدیریت یادگیری خود را به‌عهده بگیرند و به‌خودبازدهی^۷ و فراشناخت^۸ خود کمک کنند.

نگاهی به ادبیات پشینۀ خودآزمایی و هم‌آزمایی هم در زمینه جامع آموزش و پرورش و هم در آموزش زبان خارجه، نشانگر آن است که پژوهش‌های عملی و آزمایشی بسیاری نیز از گذشته تاکنون به‌تاثیرات مثبت این دو روش بر روی یادگیری یادگیرنده‌ها و زبان‌آموزان اذعان

1. Educational assessment
2. Psychometric tests
3. Alternative assessment
4. Self-assessment
5. Peer-assessment
6. Autonomy
7. Self-efficacy
8. Metacognition

دارند (مثال: چنگ و وارن، ۲۰۰۵؛ ژیاو و لوکینگ، ۲۰۰۸؛ بیرجندی و سیری، ۲۰۱۰؛ بیرجندی و حدیدی تمجید، ۲۰۱۲). باوجود تمامی این حمایت‌های نظری و عملی از خودآزمایی و هم‌آزمایی، معمولاً این دو روش سنجش آموزش‌محور در محیط‌های آموزش به‌ویژه کلاس‌های آموزش زبان زیاد کاربرد زیادی ندارند. این مشکل شاید بیشتر به‌این دلیل باشد که معلمان به‌توانایی یادگیرنده‌ها در اجرای دقیق خودآزمایی و هم‌آزمایی زیاد اطمینان ندارند (اسکارسون، ۱۹۸۹). این نبود اطمینان تا حدی نیز قابل توجه‌است چرا که باور اینکه کسی بتواند به‌سنجش توانایی خود و یا هم‌تایان خود در مورد مهارت و یا توانایی پردازد در حالیکه خود هنوز در آن مهارت و یا توانایی در حال آموزش دیدن است دشوار است. افزون بر این، بعضی مطالعات عملی در ادبیات پیشینه در زمینه پایایی و روایی خودآزمایی و هم‌آزمایی نشانگر آنند که یادگیرنده‌ها در انجام خودآزمایی و هم‌آزمایی به‌طور نسبی از خود خطای سنجش نشان می‌دهند (پتری، ۲۰۰۲). اما شمار دیگری از مطالعات نیز نشانگر آنند که خودآزمایی و هم‌آزمایی می‌توانند به‌اندازه سنجش‌های انجام شده توسط خود معلمان، دارای پایایی و اعتبار باشند (بلانچ و مرینو، ۱۹۸۹؛ راس، ۱۹۹۸؛ دوچی و سگرز، ۱۹۹۹؛ فالچیکوف و گلدفینچ، ۲۰۰۰).

به‌طور کلی، وقتی به‌روند مطالعات فوق نگاه عمیق‌تری می‌کنیم، این امر واضح و مبرهن است که این مطالعات هر یک در شرایط متفاوتی با عوامل و متغیرهای دخیل گوناگونی انجام گرفته‌اند، و اگر هر دسته از مطالعات، به‌نتایج متفاوتی رسیده‌اند بدان دلیل است که این متغیرها و عوامل به‌طور نسبی کنترل شده و یا نشده‌اند. به‌عنوان مثال عوامل و متغیرهایی از جمله پیشینه دانشی یادگیرنده‌ها، آموزش در زمینه خودآزمایی و هم‌آزمایی، انتظارات دوستان، والدین، و معلمان و یا خود یادگیرنده‌ها از خود و هم‌تایانشان، ابزار سنجش و غیره می‌توانند در میزان دقت سنجش، پایایی و روایی خودآزمایی و هم‌آزمایی تاثیر داشته باشند (فالچیکوف و گلدفینچ، ۲۰۰۰). بیرجندی و سیری (۲۰۱۶: ۱۱۸) با انجام مروری تقریباً جامع از مطالعات مختلف در طی دو سه دهه گذشته، فهرستی از این متغیرها ارائه کرده‌اند که در بیشتر مطالعات عملی مورد بررسی قرار گرفته‌اند و می‌توانند در میزان پایایی و روایی خودآزمایی و هم‌آزمایی تاثیر داشته باشند. این متغیرها و عوامل عبارتند از:

۱. سطح توانایی یادگیرنده‌ها از لحاظ توانایی مورد سنجش (دیویدسون و هنینگ، ۱۹۸۵؛ فالچیکوف و بود، ۱۹۸۹؛ بلانچ، ۱۹۸۸؛ یانسن ون دایتن، ۱۹۸۹؛ هیلنمن، ۱۹۹۰؛ جعفرپور و یمینی، ۱۹۹۵؛ تاپینگ، ۱۹۹۸؛ فالچیکوف و گلدفینچ، ۲۰۰۰)

۲. میزان وضوح معیار سنجش (براون، ۲۰۰۵؛ اسمیت و همکاران، ۲۰۰۲؛ تاپینگ و همکاران، ۲۰۰۰)
 ۳. موضوع مورد خودآزمایی و یا همتاآزمایی (فالچیکوف و بود، ۱۹۸۹)
 ۴. تعداد افراد حاضر در همتاآزمایی (استفانی، ۱۹۹۴)
 ۵. مکان و موقعیتی که در آن خودآزمایی و همتاآزمایی صورت می‌گیرد (فالچیکوف و گلدفینچ، ۲۰۰۰)
 ۶. میزان جزئی بودن معیار سنجش و روش نمره‌گذاری (کلی و جزئی) (لیک و وایویل، ۲۰۰۱، ۲۰۰۲؛ میلر، ۲۰۰۳؛ مک آلیپین، ۱۹۹۹)
 ۷. وجود بازخورد از جانب معلم (تاراس، ۲۰۰۱)
 ۸. عرف‌های فرهنگی (ماتسونو، ۲۰۰۹)
 ۹. جنسیت یادگیرنده‌ها (آندراده و همکاران، ۲۰۰۹؛ فالچیکوف و مگین، ۱۹۹۷؛ لنگن و همکاران، ۲۰۰۸؛ پوپ، ۲۰۰۵)
 ۱۰. سبک‌های یادگیری (کسیدی، ۲۰۰۷؛ ویکرمن، ۲۰۰۹)
 ۱۱. اضطراب، دیدگاه‌ها، عواطف و ویژگی‌های شخصیتی (الفلی، ۲۰۰۴؛ بلانچ، ۱۹۹۸؛ دلاسکا و کرکلر، ۲۰۰۸؛ پوپ، ۲۰۰۵)
- بیرجندی و سیری (۲۰۱۶) در ادامه توضیح می‌دهند که این متغیرها به دو دسته موقعیتی و فردی قابل تقسیم‌اند. متغیرهای موقعیتی آن‌هایی هستند که قابل کنترل و بهبود پذیرند و از شرایط بیرونی افراد سرچشمه می‌گیرند، مانند روش سنجش و معیار سنجش که قابل آموزش و کنترل هستند تا پایایی و روایی سنجش بالا رود. از سوی دیگر، متغیرهای فردی همچون سبک‌های یادگیری و نوع شخصیت قابل کنترل نیستند و می‌توانند در روایی و پایایی خودآزمایی و همتاآزمایی تاثیر منفی بگذارند. در تایید همین نکته، براون و هادسون (۲۰۰۲) نیز باور دارند که میزان دقت در انجام خودآزمایی و همتاآزمایی، می‌تواند با واضح سازی معیار و ابزار سنجش بالا رود. اسکارسون (۱۹۸۹) نیز تاکید بر این دارد که آموزش و تمرین دادن یادگیرنده‌ها برای انجام خودآزمایی و همتاآزمایی صحیح می‌تواند پایایی سنجش را بالا ببرد. در همین راستا، مطالعات بسیاری تاکنون در پژوهش‌های عملی بر اهمیت ضرورت آموزش و معیارهای واضح برای سنجش صحنه گذاشته‌اند (هریس، ۱۹۹۷؛ یانسن ون دایتن، ۱۹۸۹؛ تاپینگ و همکاران، ۲۰۰۰؛ سیری ۲۰۱۰).

همانطور که از مباحث بالا واضح است، پژوهش‌های زیادی از سال‌ها پیش تاکنون به‌صورت نظری و عملی متغیرهای فردی و موقعیتی دخیل در میزان پایایی و روایی خودآزمایی و هم‌تأزمایی را مورد بررسی قرار داده‌اند. اما به‌نظر می‌رسد همچنان متغیرهای دخیل دیگری وجود دارد که از نظر پژوهشگران دور مانده‌است. یکی از این متغیرها که از شرایط و موفقیت بیرون می‌تواند بر روایی و پایایی خودآزمایی و هم‌تأزمایی تأثیر بگذارد هدف از انجام خودآزمایی و هم‌تأزمایی می‌باشد که به‌اهداف پایان‌محور^۱ و سازنده^۲ تقسیم می‌شود. منظور از سنجش با هدف پایان‌محور، در نظر گرفتن نمره سنجش در نمره‌های پایان دوره، و منظور از خودآزمایی سازنده در نظر گرفتن نمره‌های خودآزمایی در طول دوره تنها به‌منظور ارائه بازخورد مناسب به‌مدرس و زبان‌آموز است. البته ناگفته نماند که محققین و صاحب نظران کمابیش به‌اهمیت هدف از انجام خودآزمایی و هم‌تأزمایی اشاره کرده‌اند و تأکید بر این داشته‌اند که هدف از استفاده از نمره‌های خودآزمایی و هم‌تأزمایی و تأکید بر دقت انجام خودآزمایی و هم‌تأزمایی به‌یادگیرنده‌ها می‌تواند بر میزان پایایی و روایی سنجش تأثیر بگذارد (بتنهاوزن و فدور، ۱۹۹۷؛ اودانل و تاپینگ، ۱۹۹۸ ارجاع داده شده در سیتو و فوجینا، ۲۰۰۴). سرانجام، باوجود حلاء پژوهشی که در ادبیات پیشینه وجود دارد، می‌توان اینگونه انتظار داشت که اگر خودآزمایی و هم‌تأزمایی در شرایطی صورت گیرد تا میزان دقت یادگیرنده‌ها در انجام سنجش، بخشی از معیار نمره‌دهی در آخر دوره آموزشی باشد (سنجش پایان‌محور)، این احتمال وجود دارد تا آن‌ها با دقت بیشتری سنجش و نمره‌دهی را در مقایسه با سنجش سازنده انجام دهند. همچنین می‌توان این احتمال را داد که با جدی گرفتن روند سنجش، از آنجایی که یادگیرنده‌ها بر روی معیار عملکرد مناسب و نمره‌دهی دقت بیشتری می‌کنند، این خود انگیزه پیشرفت بیشتر در عملکرد خودشان نسبت به‌مهارت و یا توانایی مورد سنجش می‌شود.

اهداف و سوال‌های پژوهش

با توجه به‌اشاره‌ای که به‌خلاء پژوهشی در زمینه تأثیر اهداف سازنده و پایان‌محور از انجام خودآزمایی و هم‌تأزمایی بر میزان دقت یادگیرنده‌ها در سنجش و پیشرفت در مهارت و یا

1. Summative
2. Formative

توانایی مورد سنجش صورت گرفت، پژوهش جاری با تمرکز فقط بر روی خودآزمایی نگارش و در شرایط کلاس‌های زبان انگلیسی مؤسسات آموزشی، به بررسی پاسخ پرسش‌های زیر پرداخت:

۱. آیا تفاوت معناداری بین تاثیر خودآزمایی سازنده و خودآزمایی پایان‌محور بر عملکرد نگارش زبان‌آموزان ایرانی وجود دارد؟
۲. آیا تفاوت معناداری بین تاثیر خودآزمایی سازنده و خودآزمایی پایان‌محور بر میزان دقت در نمره‌دهی خودآزمایی نگارش زبان‌آموزان ایرانی وجود دارد؟

روش پژوهش جامعه آماری

شرکت کنندگان این پژوهش متشکل از ۶۰ زبان‌آموز دختر و پسر در سطح میانی بودند که در آموزشگاه کاسپین واقع در پاکدشت از توابع استان تهران مشغول به گذراندن دوره‌های آموزش زبان بودند. سن زبان‌آموزان بین ۱۶ تا ۲۸ سال بود، و زبان‌آموزان کلاس‌های موجود به همان شکل چیدمان آموزشگاه^۱ مورد استفاده قرار گرفتند، اما کلاس‌ها به صورت تصادفی به دو گروه خودآزمایی سازنده و پایان‌محور تقسیم شدند.

ابزار پژوهش

معیار و مقیاس سنجش نگارش: به منظور سنجش و نمره‌دهی به نگارش زبان‌آموزان در این پژوهش، از مقیاس سنجش نگارش جیکابز و همکاران (۱۹۸۱)^۲ استفاده شد. بایسته یادآوری است که از این مقیاس در مطالعات مشابه دیگری در گذشته همچون پژوهش‌های فوجیتا (۲۰۰۴) و ماتسونو (۲۰۰۹) برای انجام سنجش توسط معلمان و زبان‌آموزان در خودآزمایی و هم‌آزمایی مورد استفاده قرار گرفته است، بنابراین در پژوهش جاری نیز از این مقیاس هم برای نمره‌دهی مصححان و هم برای خودآزمایی زبان‌آموزان استفاده شد. البته باید خاطر نشان کرد که این مقیاس در نسخه اولیه به ارزیابی بررسی^۳ نگارش می‌پردازد که در آن ۵ المان

1. Intact classes
2. The ESL Composition Profile
3. Analytic scoring

جداگانه در نظر گرفته می‌شود: محتوی، سازماندهی، واژگان، کاربرد ساختارهای زبان، و مسائل سازگاری غیر زبانی (علامتگذاری، حروف کوچک و بزرگ، املا، و اسلوب). افزون بر این، هر کدام از این المان‌ها به‌همراه پاره‌ای توضیحات^۱ مختصر در مورد طبیعت آن‌ها و چگونه در نظر گرفتنشان در سنجش و نمره‌دهی^۲ آورده شده‌اند و واژگان کلیدی در این توضیحات نیز به‌اختصار تعریف شده‌اند. از آنجایی که درک درست از معیار سنجش نقش بالایی در روایی و پایایی دارد (برنتمایر و ندرپلنک، ۲۰۰۸)، توضیحات این مقیاس در جزوه‌ای جداگانه با مثال‌ها و توضیحات واضح‌تر توسط سیری (۲۰۱۱) تهیه شده‌است که نسخه‌ای از همان در اختیار مصححان و زبان‌آموزان قرار گرفت. البته این توضیحات برای زبان‌آموزان به‌فارسی می‌باشد و بیان آن به‌زبان ساده تر است (رجوع شود به پیوست).

آزمون توانایی کلی زبانی: برای تعیین سطح کلی زبانی زبان‌آموزان دو گروه، از آزمون او پی تی (آلن، ۲۰۰۴) استفاده شد. این آزمون از دو قسمت مهارت شنیداری و دستور زبان تشکیل شده که هر قسمت با مهارت خواندن و واژگان ادغام شده‌است.

مراحل گردآوری داده‌ها

در آغاز دوره، آزمون او پی تی به‌زبان‌آموزان هر دو گروه داده شد، و میانگین سطح توانایی عمومی زبان آن‌ها مشخص شد که اطلاعات آن در جدول شماره ۱ آورده شده‌است. با محاسبه نسبت اوج و چولگی (عدد اوج و چولگی تقسیم بر خطای استاندارد) مشخص شد این اعداد کمتر از $1/96+$ می‌باشد که نشان دهنده توزیع نسبتاً نرمال داده‌های سطح توانایی کلی زبانی زبان‌آموزان دوگروه می‌باشد. بنابراین با استفاده از آزمایش مستقل تی میانگین دو گروه مقایسه شد و تفاوت معناداری بین دو گروه از لحاظ سطح کلی زبانی پیدا نشد؛ $0/05 > \text{سطح معناداری}$ ، $0/274 = (29)$ تی.

1. Descriptors
2. Scoring rubrics

جدول ۱. آمار توصیفی دو گروه از لحاظ میزان سطح توانایی زبانی عمومی

نسبت چولگی	نسبت اوج	انحراف از معیار	میانگین	حداکثر	حداقل	تعداد
۱/۴۴	۱/۲۳	۲۳/۳۰	۱۱۳/۸۸	۱۴۷	۷۲	۳۰
۰/۹۰	۰/۸۷	۲۳/۱۰	۱۱۰/۰۲	۱۴۲	۶۸	۳۰

نکته: حداکثر نمره از ۲۰۰ بوده است

بعد از دریافت اطمینان دربارهٔ برابری دو گروه از لحاظ سطح توانایی عمومی زبان، یک پیش‌آزمون نگارش پاراگراف به‌زبان‌آموزان دو گروه داده شد و نمره‌های آن ثبت شد. یکی از پژوهشگران این پژوهش به‌همراه مدرسی دیگر که هر دو دارای مدرک کارشناسی ارشد در آموزش زبانند، بعد از اینکه طبق اصول آموزش مصححان توسط ویگیل (۲۰۰۲) و جیکابز و همکاران (۱۹۸۱) به‌هماهنگی لازم رسیدند، نگارش زبان‌آموزان را بعد از هر جلسه آموزشی از پیش‌آزمون تا پس‌آزمون مورد سنجش و نمره‌دهی قرار دادند. بعد از برگزاری پیش‌آزمون، دورهٔ آموزشی آغاز شد. در جلسه اول مطالب پایه‌ای پاراگراف‌نویسی همچون طریقه نوشتن جمله موضوع و جملات متن و نتیجه‌گیری به‌زبان‌آموزان آموزش داده شد. بیشتر آموزش‌های در زمینه نگارش از کتاب پاراگراف‌نویسی آرنادت و بری (۱۹۹۰) بود. در جلسه‌های دوم و سوم، مقیاس سنجش نگارش جیکابز و همکاران (۱۹۸۱) به‌زبان‌آموزان معرفی و پاراگراف‌های نوشته شده توسط آن‌ها در پیش‌آزمون برای انجام خودآزمایی به‌آن‌ها داده شد. سپس زبان‌آموزان نمره‌های خودآزمایی خود را با میانگین نمره‌های مصححان مقایسه کردند و هر گونه ابهامی در این زمینه توسط مدرسان در هر جلسه بحث و رفع شد. این رویه در تمامی جلسه‌های آموزشی بعد تکرار شد.

بعد از این جلسه‌های ابتدایی، به‌مدت در کل ۹ جلسه، در هر جلسه یک روش پاراگراف‌نویسی به‌همراه تمرینات مربوطه کتاب به‌زبان‌آموزان دو گروه ارائه شد. سپس دو موضوع برای پاراگراف‌نویسی به‌زبان‌آموزان معرفی می‌شد که از بین آن‌ها یک موضوع انتخاب

و یک پاراگراف توسط زبان‌آموزان نوشته می‌شد. بر اساس نوع گروه، به‌زبان‌آموزان گفته می‌شد که نمره‌های خودآزمایی برای اهداف سازنده و یا پایان‌محور قرار بود در نظر گرفته شود. همانطور که پیشتر توضیح داده شد هم نمره‌های خودآزمایی و هم نمره‌های مصححان بر اساس یک مقیاس مشترک بود، بنابراین زبان‌آموزان می‌توانستند مقایسه‌ای دقیق بین نمره‌های خود و میانگین نمره‌های مصححان داشته باشند. همانطور که قبلاً توضیح داده شد میانگین نمره‌های مصححان به‌عنوان نمره‌های معیار لحاظ شدند، بنابراین با تفریق این نمره معیار از نمره خودآزمایی هر زبان‌آموز میزان خطای زبان‌آموزان در خودآزمایی محاسبه شد. به‌طور طبیعی هر چه این نمره خطا پایینتر بود، میزان دقت در نمره خودآزمایی بالاتر بود. تا اینجا تمامی مراحل در هر دو گروه سنجش یکسان بود، اما در گروه پایان‌محور قرار بر این شد تا میزان دقت در خودآزمایی زبان‌آموزان و پیشرفت آن‌ها در این زمینه تا پایان دوره به‌عنوان ۱۵ درصد از نمره پایان دوره آن‌ها لحاظ شود ولی در گروه سازنده چنین چیزی لحاظ نشد. سرانجام بعد از جلسه نهم یک پس‌آزمون به‌زبان‌آموزان دوگروه داده شد تا میزان دگرگونی آن‌ها از پیش‌آزمون تا پس‌آزمون از لحاظ کیفیت عملکرد نگارش و میزان دقت در نمره‌دهی خودآزمایی مورد بررسی قرار گیرد.

بررسی داده‌ها

در این قسمت، هر سوال پژوهش در قسمت‌های فرعی جداگانه بررسی می‌شود. برای انجام بررسی‌های آماری، در ابتدا پیش‌فرض‌های آماری مربوطه شامل نرمال بودن داده‌ها، تساوی واریانس‌ها^۱، و تساوی شیب رگرسیون^۲ به‌ترتیب با در نظر گرفتن نسبت چولگی و اوج، آزمایش لوین، و هم‌کنشی کووریت و متغیر غیر وابسته بررسی شد.

بررسی سوال ۱: مقایسه نمره‌های نگارش در پس‌آزمون

برای پاسخ به‌سوال اول پژوهش، مقایسه میانگین نمره‌های زبان‌آموزان دو گروه در پس‌آزمون ضروری بود. اما از آنجایی که مشخص نبود که آیا دو گروه در شروع دوره دارای

1. Homogeneity of variances
2. Homogeneity of regression slopes

میانگین پیش‌آزمون برابر بودند یا نه، لازم بود که در مقایسه پس‌آزمون هر گونه تفاوت چشمگیری در پیش‌آزمون بین دو گروه (کووریت^۱) کنترل شود، که این کار با به‌کارگیری بررسی کوواریانس^۲ صورت گرفت.

همانطور که پیشتر توضیح داده شد، پیش فرض‌های بررسی کوواریانس از جمله نرمال بودن داده‌ها، تساوی واریانس‌ها، و تساوی شیب رگرسیون قبل از اجرای تحلیل بررسی شد. جدول ۲ آمار توصیفی نمره‌های نگارش در پیش‌آزمون و پس‌آزمون را نشان می‌دهد. همانطور که مشخص است دو گروه هم در پیش‌آزمون و هم در پس‌آزمون دارای تفاوت‌هایی اند که معناداری این تفاوت‌ها با تحلیل کوواریانس بررسی شد.

جدول ۲. آمار توصیفی نمره‌های نگارش

جلسه	گروه							
	گروه خودآزمایی سازنده				گروه خودآزمایی پایان‌محور			
	میانگین	انحراف از معیار	نسبت اوج	نسبت چولگی	میانگین	انحراف از معیار	نسبت اوج	نسبت چولگی
پیش‌آزمون	۴۸/۶۷	۸/۴۷	۱/۰۱	۰/۷۷	۵۱/۱۱	۷/۵۸	۰/۲۹	۰/۴۵
پس‌آزمون	۶۰/۱۱	۵/۸۶	۱/۱۲	۱/۳۲	۶۱/۶۸	۵/۳۶	۱/۸۹	۰/۲۲

جدول ۳ ارائه‌کننده نتایج اصلی تحلیل کوواریانس است. ردیف اول نشان‌دهنده این است که دو گروه در پیش‌آزمون تفاوت معناداری داشتند ($0/05 < \text{سطح معناداری}$)، اما ردیف سوم نشان می‌دهد که تفاوت معناداری بین میانگین نمره‌های نگارش دو گروه در پس‌آزمون نیست ($0/05 > \text{سطح معناداری}$). به عبارت دیگر، فرضیه صفر مربوط به سوال اول پژوهش مبنی بر نبود وجود تفاوت معنادار بین تاثیر دو روش خودآزمایی با اهداف سازنده و پایان‌محور بر عملکرد نگارش تایید شد.

1. Covariate
2. Analysis of covariance (ANCOVA)

جدول ۳. نتایج تحلیل کوواریانس درباره مقایسه نمره‌های نگارش دو گروه در پس‌آزمون

ضریب تاثیر	سطح معناداری	اف	درجه آزادی	
	۰/۰۰	۳۸/۷۶	۱	مقایسه میانگین‌های پیش‌آزمون (کووریت)
	۰/۲۰	۱/۶۱	۱	تعامل کووریت و متغیر غیر وابسته
۰/۰۷	۰/۲۳	۴۸/۱۵	۱	مقایسه میانگین‌های پس‌آزمون (متغیر وابسته)
			۵۷	خطا

بررسی سوال ۲: مقایسه میزان خطا در خودآزمایی در پس‌آزمون

برای پاسخ به سوال دوم پژوهش، مقایسه میانگین میزان خطا در خودآزمایی زبان‌آموزان دو گروه در پس‌آزمون ضروری بود. اما از آنجایی که مشخص نبود که آیا دو گروه در شروع دوره دارای میانگین میزان خطا در خودآزمایی پیش‌آزمون برابر بودند یا نه، لازم بود که در مقایسه پس‌آزمون هر گونه تفاوت چشمگیری در پیش‌آزمون بین دو گروه (کووریت) کنترل شود، که این کار با به‌کارگیری تحلیل کوواریانس انجام گرفت.

همانطور که پیشتر توضیح داده شد پیش فرض‌های تحلیل کوواریانس از جمله نرمال بودن داده‌ها، تساوی واریانس‌ها، و تساوی شیب رگرسیون قبل از اجرای تحلیل بررسی شد. جدول ۴ آمار توصیفی میزان خطا در خودآزمایی در پیش‌آزمون و پس‌آزمون را نشان می‌دهد. همانطور که مشخص است دوگروه هم در پیش‌آزمون و هم در پس‌آزمون دارای تفاوت‌هایی‌اند که معناداری این تفاوت‌ها با تحلیل کوواریانس بررسی شد.

جدول ۴. آمار توصیفی بر میزان دقت در نمره‌دهی خودآزمایی

جلسه	گروه							
	گروه خودآزمایی سازنده				گروه خودآزمایی پایان‌محور			
	میانگین	انحراف از معیار	نسبت اوج	نسبت چولگی	میانگین	انحراف از معیار	نسبت اوج	نسبت چولگی
پیش‌آزمون	۱۱/۰۳	۵/۱۰	۱/۵۱	۰/۱۱	۱۲/۱۱	۶/۲۷	۰/۹۷	۰/۰۲
پس‌آزمون	۴/۵۳	۲/۷۴	۱/۰۱	۱/۹۰	۷/۳۴	۴/۸۲	۱/۰۲	۱/۷۸

جدول ۴ ارائه‌کننده نتایج اصلی تحلیل کوواریانس می‌باشد. ردیف اول نشان‌دهنده این است که دو گروه در پیش‌آزمون تفاوت معناداری داشتند ($0/05 < \text{سطح معناداری}$). همچنین ردیف سوم نشان می‌دهد که تفاوت معناداری بین میانگین میزان دقت در نمره‌دهی خودآزمایی دو گروه در پس‌آزمون وجود دارد ($0/05 < \text{سطح معناداری}$). به عبارت دیگر، فرضیه صفر مربوط به سوال دوم پژوهش مبنی بر نبود تفاوت معنادار بین تاثیر دو روش خودآزمایی با اهداف سازنده و پایان‌محور بر بر میزان دقت در نمره‌دهی خودآزمایی نگارش رد شد. با توجه به آمار توصیفی ارائه شده در جدول ۳ می‌توان نتیجه گرفت که زبان‌آموزان در گروه پایان‌محور در پس‌آزمون پیشرفت چشمگیرتری در میزان دقت در نمره‌دهی خودآزمایی از خود نشان دادند.

جدول ۴. نتایج تحلیل کوواریانس در مورد مقایسه میزان خطا در خودآزمایی دو گروه در پس‌آزمون

ضریب تاثیر	سطح معناداری	اف	درجه آزادی
	۰/۰۵	۲/۷۱	۱
	۰/۰۱	۴/۱۶	۱
۰/۳۴	۰/۰۲	۰/۰۲۱	۱
			۵۷
خطا			

بحث و نتیجه‌گیری

تحلیل داده‌ها در این پژوهش حاکی از آن بود که در هر دو شرایط خودآزمایی سازنده و پایان‌محور زبان‌آموزان پیشرفت برابری را در عملکرد نگارش‌شان نشان دادند. اما از لحاظ پیشرفت در میزان دقت نمره‌دهی در خودآزمایی، زبان‌آموزان گروه سنجش پایان‌محور پیشرفت معناتری نسبت به گروه خودآزمایی سازنده از خود نشان دادند. به‌نظر می‌رسد این نتایج، مطالعات عملی پیشین موجود در زمینه تاثیرات مثبت خودآزمایی در بهبود مهارت مورد سنجش را تایید میکند (مثال: چنگ و وارن، ۲۰۰۵؛ براون، ۲۰۰۵؛ ژیاو و لوکینگ، ۲۰۰۸؛ بیرجندی و سیری، ۲۰۱۰؛ بیرجندی و حدیدی تمجید، ۲۰۱۲). دلیل چنین تاثیرات مثبت در انجام خودآزمایی را می‌توان در نوشته‌های گیپس (۱۹۹۴) یافت که می‌نویسد خودآزمایی به‌عنوان یکی از نمودهای ارزشیابی و سنجش آموزش‌محور دارای مزایای بسیاری است؛ از آن جمله می‌توان به‌مواردی همچون ارائه بازخورد مناسب، ارائه بازخورد مثبت، ارتقا خود‌مدیریتی^۲ در یادگیرنده‌ها، افزایش عزت نفس و انگیزه یادگیرنده‌ها و معلمان، ارتقا سنجش مبتنی بر معیارهای واضح، و تسریع پیشرفت و یادگیری اشاره کرد. در همین راستا، براون و هادسون (۲۰۰۲) نیز اذعان دارند که خودآزمایی زمان کمی برای اجرا در کلاس لازم دارد و در عین حال باعث درگیر کردن زبان‌آموزان در روند یادگیری می‌شود که این خود منجر به استقلال بیشتر و انگیزه بالا در آن‌ها میشود. اسکارسون (۱۹۸۹) نیز به‌مواردی همچون ارتقا یادگیری، افزایش آگاهی، بهبود هدف‌گذاری‌های یادگیری، ایجاد مسئولیت مشترک در سنجش و ارزشیابی با معلمان، و ایجاد تاثیرات مثبت فرا کلاسی به‌عنوان مزایای خودآزمایی اشاره میکند. به‌همین ترتیب، سیتو و فوجیتا (۲۰۰۴) اظهار می‌دارند که مزایایی که خودآزمایی دارد در دیگر اشکال سنجش و ارزشیابی دیده نمیشود. از آن جمله می‌توان به‌این نکته اشاره کرد که خودآزمایی زبان‌آموزان را وا می‌دارد که عملکرد خود را با معیارهای معلم مقایسه کنند که این کار منجر به افزایش آگاهی و تفکر نقادانه آن‌ها می‌شود. افزون‌براین، پذیرش مسئولیت در سنجش و ارزشیابی منجر به مسئولیت‌پذیری زبان‌آموزان در روند یادگیری نیز میشود. درباره افزایش چشمگیر در دقت نمره‌دهی خودآزمایی که در گروه سنجش پایان‌محور این مطالعه اتفاق افتاد، به‌نظر می‌رسد مطالعه مشابهی تاکنون دقیق روی هدف از انجام خودآزمایی

1. Washback

2. Self-monitoring

(به صورت پایان‌محور و سازنده) به عنوان متغیر تاثیر گزار تمرکز نکرده است. اما مروری بر ادبیات پیشینه نشانگر این است که عده‌ای به اهمیت این متغیر اشاره کرده‌اند. به عنوان مثال، سیتو و فوجیتا (۲۰۰۴) به اظهارات بنتهاوزن و فدور (۱۹۹۷) و اودانل و تاپینگ (۱۹۹۸) اشاره می‌کنند که در آن از هدف از انجام خودآزمایی به منظور ارزشیابی و یا تنها یادگیری به عنوان متغیری تاثیر گزار یاد شده است. بنابراین مطالعه جاری می‌تواند به پیشینه عملی در مورد خودآزمایی اضافه کند چرا که نتایج نشان داد که وقتی معلم در ازای دقت بالاتر در انجام خودآزمایی به آن‌ها نمره می‌دهد، زبان‌آموزان خودآزمایی را جدی تر می‌گیرند و به مرور زمان سطح دقتشان در نمره‌دهی به معلمان نزدیکتر می‌شود. قابل توجه اینکه خود مدرسان این مطالعه ادعان به شهود جدید بیشتر در بین زبان‌آموزان گروه پایان‌محور داشتند.

افزون بر نکته‌های بالا، یافته‌های این پژوهش می‌تواند مدرسان را برای به کار گیری خودآزمایی در بین یادگیرنده‌ها تشویق کند، چرا که یکی از نگرانی‌های معلمان بی‌دقتی یادگیرنده‌ها در انجام خودآزمایی است. در ضمن شاید بتوان نتیجه گرفت، با وجود تاثیرات منفی متغیرهای موقعیتی و فردی (بیرجندی و سیری، ۲۰۱۶)، جای امیدواری وجود دارد که دست‌کم از میزان خطا در خودآزمایی ناشی از تاثیرات متغیرهای موقعیتی و فردی می‌تواند کاست، و بنابر نتایج این پژوهش یکی از این راه‌ها در این راستا اجرای خودآزمایی به صورت نه فقط پایان‌محور محض بلکه حداقل با تلفیقی از خودآزمایی سازنده و پایان‌محور می‌تواند باشد. و در پایان، دور از دسترس نخواهد بود تا بتوان با اجرای خودآزمایی پایان‌محور به شکلی که در این مطالعه صورت گرفت، نوعی از فرهنگ اعمال جدید در خودآزمایی را بین یادگیرنده‌ها جا انداخت تا به مرور زمان بتوان از نتایج خودآزمایی و حتی هم‌تا‌آزمایی در تصمیم‌گیری‌های مهمتر استفاده کرد و از مسئولیت مدرسان و مصححان در این زمینه به نفع هم یادگیرنده و هم مدرسان و مسئولان کاست.

در پایان این نکته باید خاطر نشان شود که در بیشتر مطالعات صورت گرفته در زمینه خودآزمایی از جمله پژوهش حاضر، به طور معمول تاثیر یک متغیر بر روی میزان دقت نمره‌دهی در خودآزمایی مورد بررسی قرار گرفته است، اما چیزی که جا برای پژوهش بیشتر دارد بررسی هم‌کنش این متغیرها با یکدیگر است. به عنوان مثال بررسی هم‌کنش هر یک از متغیرهای بررسی شده در ادبیات پیشینه همچون اضطراب، دیدگاه‌ها، عواطف و ویژگی‌های شخصیتی، سطح توانایی یادگیرنده‌ها از لحاظ توانایی مورد سنجش، موضوع یا مهارت مورد خودآزمایی، مکان و موقعیت خودآزمایی، عرف‌های فرهنگی، جنسیت یادگیرنده‌ها، و

سبک‌های یادگیری با هدف از انجام خودآزمایی (سازنده و یا پایان‌محور) می‌تواند پتانسیل بالایی برای بررسی و پژوهش داشته باشد.

منابع

- Allan, D. (2004). *Oxford placement test 1*. Oxford: Oxford University Press.
- Andrade, H., Wang, X., D, Y., & Akawi, R. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research*, 102, 287-302.
- Arnaudet, M. L., & Barrett M. E. (1990). *Paragraph development. A guide for students of English (2nd ed.)*. Englewood Cliffs, NJ: Prentice Hall.
- Birjandi, P. & Hadidi Tamjid, N. (2012). The role of self-, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education*, 37(5), 513-533.
- Birjandi, P. & Siyyari, M. (2010). Self-assessment and Peer-assessment: A Comparative Study of Their Effect on Writing Performance and Rating Accuracy. *Iranian Journal of Applied Linguistics*, 13(1). 23-45.
- Birjandi, P. & Siyyari, M. (2016). Agreeableness and Conscientiousness as Predictors of University Students' Self/Peer-assessment Rating Error. *Irish Educational Studies*. 35(1), (117-135).
- Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal*, 19(1), 75-96.
- Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teacher and researchers. *Language Learning*, 39(3), 313-340.
- Brantmeier, C., & Vanderplank, R. (2008). Descriptive and criterion-referenced self-assessment with L2 readers. *System*, 36, 456-477.
- Brown, J. D. (ed.) (1998). *New ways of classroom assessment*. Alexandria, VA: TESOL Inc.
- Brown, A. (2005). Self-assessment of writing in independent language learning programs: The value of annotated samples. *Assessing Writing*, 10(3), 174-191.
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge University Press.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93-121.
- Davidson, F., & Henning, G. (1985). A self-rating scale of English proficiency: Rasch scalar analysis of items and rating categories. *Language Testing*, 2(2), 164-79.

- Dochy, F., & Segers, M., (1999). The Use of Self-, Peer and Co-assessment in Higher Education: A review. *Studies in Higher Education*, 24(3), 331-350.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395-430.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-322.
- Falchikov, N., Magin, D. (1997). Detecting gender bias in peer marking of students' group progress work. *Assessment & Evaluation in Higher Education*, 22, 385-396.
- Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. London: The Falmer Press.
- Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal*, 51, 12-20.
- Heilenmann, K. L. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing*, 7(2), 174-201.
- Jacobs, H. J., Zingraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Massachusetts: Newbury House.
- Jafarpur, A., & Yamini, M. (1995). Do Self-Assessment and Peer-Rating Improve with Training? *RELC Journal*, 26(1), 63-85.
- Janssen-van Dielen, A. (1989). The development of a test of Dutch as a second language: The validity of self-assessments by inexperienced subjects. *Language Testing*, 6(1), 30-46.
- Langan, A. M., Shuker, D. M., Cullen, W. R., Penney, D., Preziosi, R. F. & Wheeler, C. P. (2008). Relationships between student characteristics and self, peer and tutor evaluations of oral presentations, *Assessment Evaluation in Higher Education*, 3, 179-190.
- Lejk, M., & Wyvill, M. (2001). Peer assessment of contributions to a group project: A comparison of holistic and category-based approaches. *Assessment and Evaluation in Higher Education*, 26, 62-72.
- Lejk, M., & Wyvill, M. (2002). Peer Assessment of Contributions to a group project: Student attitudes to holistic and category-based approaches. *Assessment & Evaluation in Higher Education*, 27(6), 569-577.
- Macalpine, J. M. K. (1999). Improving and encouraging peer assessment of student Presentations, *Assessment & Evaluation in Higher Education*, 24, 15-25.
- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 75-100.
- McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.

- Miller, P. J. (2003). The effect of scoring criteria specificity on peer and self-assessment. *Assessment & Evaluation in Higher Education*, 28, 383-394.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and implications. *Language Testing*, 6(1), 1-13.
- Patri, M. (2002). The influence of peer feedback on self-and peer assessment of oral skills. *Language Testing*, 19(2), 109-131.
- Pope, N. (2005). The impact of stress in self- and peer-assessment. *Assessment & Evaluation in Higher Education*, 30, 51-63.
- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing* 15, 1-20.
- Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching Research*, 8(1), 31-54.
- Stefani, L. A. J. (1994). Peer, self and tutor assessment: relative reliabilities. *Studies in Higher Education*, 19, 69-75.
- Smith, H., Cooper, A., & Lancaster, L. (2002). Improving the quality of undergraduate peer assessment: A case for student and staff development. *Innovations in Education and Teaching International*, 39, 71-81.
- Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: Towards transparency for students and for tutors. *Assessment & Evaluation in Higher Education*, 26, 605-614.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249-276.
- Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in Higher Education*, 25, 149-169.
- Topping, K. (2003). *Self and peer assessment in school and university: Reliability, validity and utility*. In M. Segers, F. Dochy, & E. Cascallar (Eds.) *Optimizing new modes of assessment: In search of qualities and standards* (pp. 55-87). Dordrecht: Kluwer Academic Publishers.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Xiao, Y., & Lucking, R. (2008). *The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment*. *Internet and Higher Education*, 11, 186-193.

ضمیمه

مقیاس سنجش نگارش جیکابز و همکاران (۱۹۸۱)، ترجمه از سیری (۲۰۱۱)

Jacobs et al.'s (1981) ESL composition scoring profile

Content	محتوی
<p><u>30-27</u> EXCELLENT TO VERY GOOD</p> <p>knowledgeable*substantive*thorough development of <i>thesis (topic sentence)</i>* relevant to assigned topic</p>	<p>۲۷-۳۰ عالی تا خیلی خوب</p> <p>معلومات کافی از موضوع - ارایه جزییات کافی درباره موضوع - بسط جمله موضوع - بدون انحراف از موضوع اصلی</p>
<p><u>26-22</u> GOOD TO AVERAGE</p> <p>some knowledge of subject* adequate range* limited development of <i>thesis</i>* mostly relevant to topic, but lacks detail</p>	<p>۲۲-۲۶ خوب تا متوسط</p> <p>معلومات کم و بیش و کافی در مورد موضوع - بسط موضوع محدود اما مربوط به موضوع می‌باشد ولی جزئیات ارایه شده کافی نیست</p>
<p><u>21-17</u> FAIR TO POOR</p> <p>limited knowledge of subject* little substance* inadequate development of topic</p>	<p>۱۷-۲۱ متوسط ضعیف تا ضعیف</p> <p>معلومات محدود در مورد موضوع - ارایه جزییات کم - بسط نا کافی موضوع</p>
<p><u>16-13</u> VERY POOR</p> <p>does not show knowledge of subject* non-substantive* not pertinent * OR not enough to evaluate</p>	<p>۱۳-۱۶ بسیار ضعیف</p> <p>نبود وجود معلومات در مورد موضوع - جزییات وجود ندارد و یا مربوط نیست و یا برای ارزشیابی کافی نیست</p>

Organization		سازمان دهی	
20-18 EXCELLENT TO VERY GOOD	fluent expression* ideas clearly stated/supported* succinct*well-organized*logical sequencing*cohesive	بیان روان- بیان و تاکید واضح نظرات- موجز و بدون انحراف از موضوع- سازمان دهی قوی- ترتیب منطقی نکات- انسجام	۱۸-۲۰ عالی تا خیلی خوب
17-14 GOOD TO AVERAGE	GOOD TO AVERAGE: somewhat choppy*loosely organized but main ideas stand out*limited support* logical but incomplete sequencing	بیان تا حدی روان نیست- سازماندهی کمی ضعیف اما نکات اصلی متمایزند- تاکید محدود- بیان منطقی ولی ترتیب نکات نقص دارد	۱۶-۱۷ خوب تا متوسط
13-10 FAIR TO POOR	non-fluent* ideas confused or disconnected* lacks logical sequencing and development	بیان روان نیست- نکات و نظرات نا منظم و نا مربوط است- بسط و ترتیب نکات منطقی نیست	۱۰-۱۳ متوسط ضعیف تا ضعیف
9-7 VERY POOR	does not communicate* no organization*OR not enough to evaluate	پیامی نمی‌رساند- سازماندهی وجود ندارد- مطلب برای ارزشیابی کافی نیست	۷-۹ بسیار ضعیف

Vocabulary		واژگان	
20-18 EXCELLENT TO VERY GOOD	sophisticated range* effective word/idiom choice and usage* word form mastery * appropriate register	دایره واژگان گسترده- انتخاب و کاربرد موثر واژگان و اصطلاحات- تسلط بر اشکال کلمات- ساختارهای مناسب رسمی و غیر رسمی	۱۸-۲۰ عالی تا خیلی خوب
17-14 GOOD TO AVERAGE	adequate range* occasional errors of word/idiom form, choice, usage but meaning not obscured	دایره کافی واژگان- هر از گاهی اشتباهاتی در انتخاب و کاربرد واژگان و اصطلاحات و اشکال کلمات وجود دارد اما معنا مبهم نمی‌شود	۱۴-۱۷ خوب تا متوسط
13-10 FAIR TO POOR	limited range* frequent errors of word/idiom form, choice, usage* meaning confused or obscured	دایره محدود واژگان- اشتباهات مکرر در انتخاب و کاربرد واژگان و اصطلاحات و اشکال کلمات- معنا مبهم و نامشخص	۱۰-۱۳ متوسط تا ضعیف
9-7 VERY POOR	essentially translation* little knowledge of English vocabulary, idioms, word form* OR not enough to evaluate	متن در اصل یک ترجمه است- معلومات ناچیز واژگان، اصطلاحات و اشکال واژگان انگلیسی- و یا مطلب برای ارزشیابی کافی نیست	۷-۹ بسیار ضعیف

Language Use		کاربرد ساختارها	
25-22 EXCELLENT TO VERY GOOD	effective complex constructions* few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions	استفاده موثر از ساختارهای ترکیبی - اشتباهات بسیار کم در تطابق، زمان، تعداد، کاربرد و ترتیب واژگان، حروف تعریف، ضمائر و حروف اضافه	۲۲-۲۵ عالی تا خیلی خوب
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	ساختارها ساده ولی موثر هستند - اشکالات جزئی در ساختارهای ترکیبی - چند اشتباه در تطابق، زمان، تعداد، کاربرد و ترتیب واژگان، حروف تعریف، ضمائر و حروف اضافه وجود دارد ولی معنا به‌ندرت مبهم می‌شود	۱۸-۲۱ خوب تا متوسط
17-11 FAIR TO POOR	major problems in simple/complex constructions* frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, deletions * meaning confused or obscured	اشتباهات بزرگ در ساختارهای ترکیبی و ساده - اشتباهات مکرر در ساختارهای منفی، تطابق، زمان، تعداد، کاربرد و ترتیب واژگان، حروف تعریف، ضمائر و حروف اضافه، قرینه‌سازی، جملات ناقص و یا طولانی مبهم - معنا مبهم و نامشخص است	۱۱-۱۷ متوسط تا ضعیف
10-5 VERY POOR	virtually no mastery of sentence construction rules* dominated by errors* does not communicate* OR not enough to evaluate	تقریباً تسلطی بر قواعد جمله‌سازی وجود ندارد - مملو از اشتباه - پیامی نمی‌رساند - مطلب برای ارزشیابی کافی نیست	۵-۱۰ بسیار ضعیف

Mechanics		علامت‌گذاری، حروف کوچک و بزرگ، املا، و اسلوب	
5 EXCELLENT TO VERY GOOD	demonstrates mastery of conventions* few errors of spelling, punctuation, capitalization, paragraphing	تسلط بر قواعد نوشتار- اشتباهات بسیار کم در املا، علامت‌گذاری، حروف کوچک و بزرگ، حاشیه‌گذاری و پاراگراف بندی	۵ عالی تا خیلی خوب
4 GOOD TO AVERAGE	occasional errors of spelling, punctuation, capitalization, paragraphing but meaning not obscured	هر از گاهی اشتباهاتی در املا، علامت‌گذاری، حروف کوچک و بزرگ، حاشیه‌گذاری و پاراگراف بندی وجود دارد ولی معنا مبهم نیست	۴ خوب تا متوسط
3 FAIR TO POOR	frequent errors of spelling, punctuation, capitalization, paragraphing * poor handwriting* meaning confused or obscured	اشتباهات مکرر در املا، علامت‌گذاری، حروف کوچک و بزرگ، حاشیه‌گذاری و پاراگراف بندی- دست خط بد و معنا مبهم و نا مشخص می‌شود	۳ متوسط ضعیف تا ضعیف
2 VERY POOR	no mastery of conventions* dominated by errors of spelling, punctuation, capitalization, paragraphing* handwriting illegible* OR not enough to evaluate	نبود تسلط بر قواعد نوشتار- مملو از اشتباهات در املا، علامت‌گذاری، حروف کوچک و بزرگ، حاشیه‌گذاری و پاراگراف بندی- دست خط کاملاً ناخوانا- مطلب برای ارزشیابی کافی نیست	۲ بسیار ضعیف