

## اثرات و نحوه ارزیابی پویا، ارزیابی پویای رایانه‌ای، ارزیابی ایستا و درک مطلب

مهدی یعقوبی\*

سید آیت‌الله رزمجو\*\*

### چکیده

هدف این مطالعه مقایسه اثرات ارزیابی پویا رایانه‌ای، ارزیابی پویا و ارزیابی ایستا بر عملکرد درک مطلب بود. همچنین نگرش زبان‌آموزان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای مورد بررسی قرار گرفت. روش پژوهش از نوع ترکیبی و راهبرد اکتشافی - متوالی بوده است. به این منظور، امتیازات آزمون به‌عنوان داده‌های کمی و محتوای به‌دست‌آمده از مصاحبه‌های نیمه ساختاریافته به‌عنوان داده‌های کیفی مورد استفاده قرار گرفتند. شرکت‌کنندگان در این مطالعه شامل ۶۰ زبان‌آموز ایرانی بودند که زبان انگلیسی را به‌عنوان یک زبان خارجی می‌آموختند و سطح مهارت آنان در زبان انگلیسی، متوسط بود. زبان‌آموزان به گروه‌های ارزیابی پویا و ارزیابی پویا رایانه‌ای تقسیم‌بندی شدند. به‌منظور اطمینان حاصل کردن از عدم وجود تفاوت معنادار و قابل‌توجه در درک مطلب هر یک از این گروه‌ها، عملکرد خوانش زبان‌آموزان قبل از انجام این مطالعه مورد بررسی و ارزیابی قرار گرفت. پس از آن زبان‌آموزان در پس‌آزمون خواندن و درک مطلب شرکت کردند. در پایان در خصوص بررسی نگرش آنان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای از شرکت‌کنندگان مصاحبه نیمه ساختاری به عمل آمد. نتایج تحلیل آماری نشان داد که هر دو ارزیابی پویا و پویا رایانه‌ای، تأثیرات قابل‌توجهی بر بهبود عملکرد درک مفاهیم شرکت‌کنندگان داشته‌اند. همچنین ارزیابی پویا رایانه‌ای نسبت به ارزیابی پویا در زمینه عملکرد درک مطلب مؤثرتر بوده است. تحلیل مصاحبه‌ها نشان داد که همه شرکت‌کنندگان، نگرش مثبتی نسبت به ارزیابی پویا رایانه‌ای در مقایسه با انواع قدیمی ارزیابی‌های پویا و ایستا داشتند.

**واژه‌های کلیدی:** ارزیابی ایستا، ارزیابی پویا، ارزیابی پویا رایانه‌ای، درک مطلب، نگرش

مقاله حاضر برگرفته از رساله دکتری رشته زبان انگلیسی دانشگاه آزاد اسلامی واحد قشم است.

\* دانشجوی دکتری گروه زبان انگلیسی، واحد قشم، دانشگاه آزاد اسلامی، قشم، ایران.

\*\* استاد دانشکده زبان‌های خارجی، دانشگاه شیراز، شیراز، ایران. (نویسنده مسئول) [razmjooshiraz@yahoo.com](mailto:razmjooshiraz@yahoo.com)

## مقدمه

دستیابی به سطح بالایی از توانایی در درک مطلب، هدف اغلب زبان آموزان است. لذا برای اطلاعات و لذت بردن از شغل و اهداف مطالعه تمایل به خواندن دارند. با توجه به اینکه متون نوشته شده، انواع مختلفی از اهداف آموزشی را ارائه می‌دهد، مواجه با متون قابل درک زبان‌شناختی می‌تواند فرایند یادگیری زبان را بهبود بخشد. بعلاوه متون قابل درک، الگوی مناسب برای نوشتن را ارائه می‌دهد و فرصت‌هایی را برای بحث و مطالعه زبان (واژگان، گرامر و اصطلاحات) فراهم می‌آوردند. از این رو خواندن و درک مطلب توسط افراد بسیار ارزشمند است (باخدا و شعبانی<sup>۱</sup>، ۲۰۱۷). در شرایطی که انگلیسی زبان خارجی است، احتمال ارتباط با زبان مادری کم یا نزدیک به صفر است. از سوی دیگر، زبان آموزان نیاز بیشتری به ورودی هر چیزی برای یادگیری زبان مقصد دارند. ارزیابی به‌عنوان راهی برای ارزیابی پتانسیل واقعی یادگیرندگان شناخته شده است که ماهیت تعاملی یادگیری را به روند ارزیابی گسترش می‌دهد.

بسیاری از مدرسان و معلمان، ارزیابی را به‌عنوان بخش جدایی‌ناپذیری از زندگی خود در نظر می‌گیرند، به‌عبارت‌دیگر ارزیابی بخش مهمی از وظایف کاری آنان است. همچنین ارزیابی به‌عنوان بخشی ضروری از آموزش و یادگیری در نظر گرفته می‌شود. استفاده از شواهد ارزیابی بخشی از فرآیند برنامه‌ریزی در کلاس درس است (لمبرتو و لاینز<sup>۲</sup>، ۲۰۰۰). طبق نتایج تحقیقات، ارزیابی کلاس درس می‌تواند منجر به موفقیت دانشجویان گردد، لازمی این نتیجه آن است که ارزیابی‌ها، بازخوردهای کافی و مؤثری را برای دانشجویان بر اساس عملکرد آنان ارائه کنند. همچنین این ارزیابی‌ها باید به‌طور پیوسته و بر اساس یک مجموعه از استانداردهای مشخص برای یادگیری اجرا شوند (کریک و یو<sup>۳</sup>، ۲۰۱۸). در واقع امکان استفاده از ارزیابی در نظارت بر پیشرفت دانشجویان وجود دارد و همچنین ارزیابی جهت اصلاح آموزش به‌منظور شناخت نیازهای زبان‌آموزان انجام می‌شود (شاشیو و یان<sup>۴</sup>، ۲۰۱۰).

بر اساس مطالعه شاشیو و یان (۲۰۱۰) پیشرو نظری ارزیابی پویا ویگوتسکی است که مفهوم منطقه رشد مجاور<sup>۵</sup>، به‌عنوان یکی از مفاهیم اساسی مرتبط با ارزیابی پویا را ارائه کرد. از طریق ارزیابی پویا، مدرسان به‌عنوان مروج‌ها ایفای نقش می‌کنند و

1. Bakhoda, I., & Shabani, K.  
2. Lambert, D., & Lines, D.  
3. Crick, R. D., & Yu, G.

4. Xiaoxiao, L., & Yan, L.  
5. Zone Of Proximal Development (ZPD)

بازخوردهای سریعی را به رویه‌های یک کار ارائه می‌کنند. همچنین آنان بیان می‌کنند که ارزیابی پویا نوعی دستورالعمل برنامه‌ریزی‌شده و درون‌پویی است که به‌جای نتایج پیشرفت گذشته، بر پیشرفت آینده زبان‌آموزان تمرکز دارد.

آلیس<sup>۱</sup> (۲۰۰۸) ادعا می‌کند که اصطلاح ارزیابی پویا حالتی از ارزیابی است که هدفش اصلاح عملکرد زبان‌آموزان در طول ارزیابی است. بر اساس مطالعه پوهنر و لنتلف<sup>۲</sup> (۲۰۰۵)، مفهوم ارزیابی پویا در مقابل اندازه‌گیری ایستا از سطح مهارت زبان‌آموزان قرار دارد که هیچ‌گونه بازخورد یا مداخله‌ای را ارائه نمی‌کند.

امروزه پیشرفت‌های فن‌آوری اطلاعات، پارادایم (الگوواره) جدیدی را برای واحدهای انتقال دانش مرتبط با آموزش بزرگ‌سالان فراهم کرده است. در واقع این پارادایم، یادگیری زبان به کمک رایانه<sup>۳</sup> بوده و استفاده از رایانه در زمینه آموزشی نیز در حال افزایش است (سیمث و سارمنی-شولر<sup>۴</sup>، ۲۰۱۴). بر اساس مطالعه لوی<sup>۵</sup> (۱۹۹۷)، یادگیری زبان به کمک رایانه مربوط به جستجو برای برنامه‌های رایانه‌ای آموزش و یادگیری زبان است. به‌منظور استفاده از رایانه در پروسه‌های ارزشیابی، اخیراً ارزیابی پویا رایانه‌ای رایج گشته و موضوع بسیاری از مطالعات بوده است. داوودی و عطائی تبار<sup>۶</sup> (۲۰۱۵) به تأثیر ارزیابی پویای کامپیوتری نوشتن L2 در توسعه نوشتاری زبان‌آموزان زبان انگلیسی EFL پرداختند. عبادی و سعیدیان<sup>۷</sup> (۲۰۱۵) به بررسی اثرات ارزیابی پویای رایانه‌ای در ارتقای مهارت‌های خواندن دانش‌آموزان پیشرفته ایرانی EFL در معرض خطر پرداختند. در مطالعه‌ای دیگر عبادی و سعیدیان (۲۰۱۶) به برنامه‌ریزی برنامه‌های آموزشی آینده را از طریق ارزیابی پویای کامپیوتری L2 پرداختند.

ارزیابی پویا رایانه‌ای می‌بایست با استفاده از رایانه در اجرای ارزیابی پویا باشد (تو، ۲۰۱۴). با تمرکز بر مزایای کاربرد تکنولوژی در زمینه‌های آموزشی، دلایل قانع‌کننده‌ای برای استفاده از رایانه در مجموعه‌های آموزشی و شخصی وجود دارد. در این روش، امکان انجام فعالیت‌های یادگیری در خارج از کلاس نیز وجود دارد و اکثر زبان‌آموزان می‌توانند از رایانه به‌عنوان وسیله ارتباطی و یادگیری استفاده کنند (دونالدسون و هاگستروم<sup>۸</sup>، ۲۰۰۶).

1. Ellis, R.  
2. Poehner, M. E., & Lantolf, J. P.  
3. Computer Assisted Language Learning (CALL)  
4. Simuth, J., & Sarmany-Schuller, I.  
5. Levy, M.

6. Davoudi, M., & Ataie-Tabar, M.  
7. Ebadi, S., & Saedian, A.  
8. Donaldson, R. P., & Haggstrom, M. A.

همچنین رایانه فرصتی را در اختیار فرد قرار می‌دهد که محدودیت‌های زمانی و مکانی روش‌های یادگیری قدیمی وجود ندارند (اینان و لاوتر، ۲۰۰۷). انعطاف‌پذیری منطقی و ذاتی موجود در یادگیری زبان به کمک رایانه به‌منظور یادگیری و جستجوی آنلاین اطلاعات سبب شده است که یادگیری زبان به کمک رایانه، ابزاری مناسب برای یادگیری دانشجو محور باشد (ناوارو و لارا، ۲۰۱۷؛ هاگستروم، ۲۰۱۹).

با توجه به اینکه ارتقاء فرهنگی در سه عامل (ذاتی، درونی و بیرونی) و دو بعد (فرهنگی و آموزشی)، مؤلفه‌های فرهنگی مؤثر در آموزش زبان خارجی هستند (محمدخانی و همکاران، ۱۳۹۶) و روش آموزش زبان مبتنی بر محتوا به‌عنوان بهترین مدل آموزشی برای دانش‌آموزان است (حمادی و همکاران، ۱۳۹۷)؛ ارائه نوعی از ارزیابی به زبان‌آموزان ضروری است که به آنان در اتخاذ رویکرد تعاملی و بازسازی معنای صحیح کلی متن و همچنین آشنایی با روابط در سطوح کلی و جزئی متن کمک کند. ارزیابی‌های پویا و پویا رایانه‌ای دارای چنین ویژگی‌هایی هستند و لذا می‌توان از آن‌ها در درک مطلب استفاده کرد. مروری بر مطالعات پیشین درباره ارزیابی‌های پویا و پویا رایانه‌ای نشان می‌دهد که اگرچه برخی مطالعات اثرات ارزیابی‌های پویا و پویا رایانه‌ای بر درک مطلب را موردبررسی قرار داده‌اند، اما هیچ‌کدام از این مطالعات اثرات این دو نوع ارزیابی را نسبت به ارزیابی ایستا بررسی نکرده‌اند. بعلاوه، هیچ‌کدام از مطالعات پیشین، نگرش زبان‌آموزان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای را موردبررسی قرار نداده‌اند؛ بنابراین هدف این مطالعه، مقایسه اثرات ارزیابی پویا قدیمی و رایانه‌ای بر درک مطلب است. همچنین نگرش زبان‌آموزان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای نیز موردبررسی قرار گرفت. بر همین اساس، سؤالات تحقیق به‌صورت زیر تنظیم شدند:

- ارزیابی پویا (DA) چه تأثیری بر عملکرد درک مطلب دارد؟
- ارزیابی پویا رایانه‌ای (CDA) چه تأثیری بر عملکرد درک مطلب دارد؟
- آیا تفاوت چشمگیری میان اثرات ارزیابی پویا و ارزیابی پویا رایانه‌ای بر عملکرد درک مطلب وجود دارد؟
- نگرش زبان‌آموزان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای چیست؟

## روش

در این مطالعه از طراحی ترکیبی از طریق تحلیل کیفی و کمی داده‌ها استفاده شده است (کرسول، ۲۰۰۹). به این منظور، امتیازات آزمون به‌عنوان داده‌های کمی و محتوای به‌دست‌آمده از مصاحبه‌های نیمه ساختاریافته به‌عنوان داده‌های کیفی مورد استفاده قرار گرفتند. شرکت‌کنندگان در این مطالعه از میان زبان‌آموزان سطح متوسط EFL که در دوره زبان انگلیسی دانشگاه آزاد اسلامی واحد مریوان ثبت‌نام کرده بودند، انتخاب شدند. ۶ کلاس در این مطالعه شرکت کردند. بعد از اجرای آزمون مهارت که شامل آزمون زبان سطح متوسط (PET) بود، دانشجویانی که توانستند نمرات بین یک انحراف معیار بالاتر و پایین‌تر از میانگین کسب کنند، به‌عنوان شرکت‌کنندگان در این مطالعه انتخاب شدند. دانشجویان ۶ کلاس به‌صورت تصادفی به سه گروه تقسیم‌بندی شدند. در مجموع، نمونه انتخاب‌شده شامل ۷۵ زبان‌آموز EFL بود که شامل ۲۸ نفر آقا و ۴۷ نفر خانم بودند. دو کلاس به‌عنوان گروه آزمایشی که یکی از کلاس‌ها همراه با ارزیابی پویا بود، دو کلاس به‌عنوان گروه آزمایشی دیگر همراه با ارزیابی پویا رایانه‌ای و دو کلاس به‌عنوان گروه کنترل همراه با ارزیابی ایستا بودند. شرکت‌کنندگان از میان آقایان و خانم‌های زبان‌آموز که سن آنان بین ۲۰ تا ۳۰ سال قرار داشت، انتخاب شدند؛ زیرا پژوهشگران به گروه‌های با سنین دیگر دسترسی نداشتند و حضور این افراد در تحقیق به‌صورت اختیاری بوده است.

آزمون زبان سطح متوسط<sup>۲</sup> در این مطالعه از آزمون PET به‌منظور انتخاب زبان‌آموزان همگن با توجه به مهارت زبانی عمومی آن‌ها استفاده گردید. آزمون PET شامل سه بخش است: بخش اول مهارت‌های خوانش و نوشتاری که با هم موردسنجش قرار می‌گیرند، بخش دوم مهارت شنیداری و بخش سوم مهارت گفتاری (یک مصاحبه) مورد سنجش قرار می‌گیرد.

این آزمون بر روی ۳۰ نفر که دارای مشخصات مشابه با شرکت‌کنندگان هدف بودند نیز اجرا شد که هدف از آن، اطمینان حاصل کردن از مناسب بودن این آزمون برای شرکت‌کنندگان موردنظر بود. ضریب آلفای کرونباخ کل برای نمرات نهایی آزمون PET، ۰/۸۶ به دست آمد که میزان رضایت‌بخشی است. شاخص پایایی برای مهارت‌های گفتاری و نوشتاری آزمون PET به ترتیب ۰/۷۹ و ۰/۸۱ بودند که مقادیر قابل‌قبولی را دارا هستند.

1. Creswell, J. W.

2. PET

در این مطالعه، از نمرات مهارت خوانش مربوط به پیش‌آزمون - پس‌آزمون به‌منظور همگنی شرکت‌کنندگان استفاده شد. هر آزمون خوانش از ۵ قسمت تشکیل شده و در مجموع دارای ۳۵ سؤال بود.

پژوهشگران برای آزمون خوانش، چهار متن را از اینترنت بر مبنای داشتن موضوعات کلی و انتشار متون بعد از سال ۲۰۱۸ دانلود کرده و از آن‌ها به‌عنوان مواد آزمون خوانش در این مطالعه استفاده کردند. همچنین پژوهشگران میزان خوانایی این متن‌ها را محاسبه کردند تا از میزان سختی یا آسانی آن‌ها اطمینان حاصل کنند. بدین منظور از یک مقیاس خوانایی جهت تعیین شاخص خوانایی مواد آزمون استفاده گردید که توسط یک ابزار آنلاین رایگان به نام 'Textalyser' ارائه می‌شود. کاربرد Textalyser جهت تخمین خوانایی در این مطالعه با دستورالعمل‌های ارائه‌شده توسط گانینگ<sup>۲</sup> (۱۹۵۲) و همچنین ارائه‌شده در مطالعه کول و اشکالینیک<sup>۳</sup> (۲۰۰۸) سازگار و هماهنگ بود. از دیدگاه گانینگ، شاخص خوانایی حدود ۱۱ یا ۱۲ می‌تواند به‌عنوان سطح بسیار سخت در نظر گرفته شود در حالی که شاخص خوانایی مناسب برای زبان‌آموزان سطح متوسط حدوداً بین ۵ تا ۶ است. هر متن حدوداً شامل ۳۰۰ کلمه بود و پنج سؤال درک مطلب نیز در ادامه آن قرار داشت. متن‌ها و سؤالات درک مطلب در هر سه گروه دقیقاً مشابه بودند. اگرچه در گروه کنترل، متن‌ها به همراه ارزیابی ایستا، در گروه ارزیابی پویا نیز متن‌ها در کنار ارزیابی پویا و در گروه ارزیابی پویا رایانه‌ای نیز متن‌ها در کنار ارزیابی ارائه‌شده توسط نرم‌افزار مرتبط با این هدف ارائه می‌شدند.

پژوهشگران مجموعه سؤالات مصاحبه نیمه ساختاریافته را با توجه به نگرش افراد نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای آماده کردند. از یک تحلیل محتوای کیفی نیز جهت گزارش نتایج مصاحبه‌ها استفاده گردید.

آزمون PET برای ۷۵ زبان‌آموز اجرا گردید و تنها افرادی که امتیازشان در حدود  $\pm 1$  انحراف معیار از میانگین بود، به‌عنوان شرکت‌کنندگان اصلی در این مطالعه انتخاب شدند. بر همین اساس، ۶۰ زبان‌آموز انتخاب و به ۳ گروه بیست‌نفره تقسیم‌بندی شدند. امتیازات مهارت خوانش آزمون PET مورد استفاده برای همگنی مهارت زبانی افراد،

1. <http://textalyser.net>  
2. Gunning, R.

3. Kol, S., & Scholnik, M.

جهت اطمینان حاصل کردن از این که قبل از انجام این مطالعه سه گروه از نظر درک مطلب تفاوت چشمگیری نداشته باشند، مورد ارزیابی قرار گرفت.

به یکی از گروه‌ها، ارزیابی پویا برای خوانش متن بر اساس مطالعه پوهنر (۲۰۰۸) ارائه شد. بدین منظور زبان‌آموزان گروه ارزیابی پویا، هر جلسه یک متن را می‌خواندند و سپس به سؤالات درک مطلب پاسخ می‌دادند. مدرس از دانشجویان خواست که به سؤالات در حد یک سطر پاسخ دهند. اگر پاسخ یک زبان‌آموز اشتباه بود، مدرس میانجی‌گری را از طریق ارائه اشارات و راهنمایی‌های بسیاری ضمنی به منظور کمک به زبان‌آموز در رسیدن به پاسخ صحیح آغاز می‌کرد. در صورتی که همچنان زبان‌آموز قادر به یافتن پاسخ صحیح نبود، مدرس اشارات مستقیم و صریحی را ارائه می‌کرد تا دانشجو بتواند به پاسخ صحیح دست یابد.

در گروه ارزیابی پویا رایانه‌ای نیز مدرس یک نرم‌افزار را بر روی لپ‌تاپ‌های زبان‌آموزان نصب کرد و نحوه خواندن متن‌ها و پاسخ به سؤالات درک مطلب را برای آنان شرح داد. در این نرم‌افزار نیز اگر پاسخ‌های زبان‌آموز غلط می‌بودند، نرم‌افزار از طریق یک روش گام‌به‌گام، ضمنی‌ترین تا صریح‌ترین راهنمایی‌ها را در اختیار زبان‌آموز قرار می‌داد تا این که بتواند به پاسخ صحیح برسد (پوهنر، ۲۰۰۸).

در گروه کنترل نیز زبان‌آموزان متن‌های مشابه را خواندند و به سؤالات درک مطلب پاسخ دادند ولی هیچ‌گونه ارزیابی پویایی به آن‌ها ارائه نگردید. در این گروه، زبان‌آموزان صرفاً به سؤالات پاسخ دادند و مدرس فقط صحیح یا غلط بودن پاسخ آنان را مشخص می‌کرد.

این مطالعه به مدت ۴ جلسه به طول انجامید و سپس بلافاصله بعد از آن، زبان‌آموزان هر سه گروه در پس‌آزمون درک مطلب شرکت کردند. شایان‌ذکر است که از ۱۰ نفر از هر گروه، به منظور بررسی نگرش آنان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای مصاحبه به عمل آمد.

### یافته‌ها

به‌منظور اطمینان حاصل کردن از عدم وجود تفاوت بسیار زیاد از نظر درک مطلب هر سه گروه، نمرات درک مطلب مربوط به پیش‌آزمون آنان از طریق تحلیل واریانس یک‌طرفه<sup>۱</sup> تحلیل و بررسی گردید. جدول ۱، آمار توصیفی امتیازات درک مطلب مربوط به پیش‌آزمون هر سه گروه را نشان می‌دهد.

جدول ۱. آمار توصیفی امتیازات درک مطلب پیش‌آزمون

تعداد	میانگین	انحراف معیار	خطای استاندارد	حداقل	حداکثر
۲۰	۲۱/۹۰۰۰	۴/۸۶	۰/۴۱۶۱۲	۱۲	۲۵
۲۰	۲۱/۸۵۰۰	۴/۴۲	۰/۶۶۱۱۲	۱۳	۲۷
۲۰	۲۱/۵۵۰۰	۵/۱۶	۰/۴۸۳۸۲	۱۱	۲۶

قبل از انجام تحلیل واریانس یک‌طرفه، فرض همگنی واریانس‌ها نیز بررسی گردید  
 $F(2, 57) = 2/191, P > 0/05$  (جدول ۲).

جدول ۲. آزمون لون برای بررسی برابری واریانس‌های خطا مربوط به امتیازات پیش‌آزمون

F	df1	Df2	معناداری
۲/۱۳۱	۲	۵۷	۰/۱۰۱

جدول ۳ نتایج تحلیل واریانس یک‌طرفه انجام شده بر روی امتیازات پیش‌آزمون را نشان می‌دهد.

جدول ۳. نتایج تحلیل واریانس امتیازات پیش‌آزمون

مجموع مربعات	Df	میانگین مربعات	F	معناداری
۱/۲۳۳	۲	۰/۷۱۷	۰/۲۱۰	۰/۲۱۲
۲۳۰/۹۵۰	۵۷	۲/۰۵۲		
۲۳۲/۱۸۳	۵۹			

همان‌طور که در جدول ۳ نشان داده شده است، هیچ‌گونه تفاوت آماری معناداری میان سه گروه در رابطه با پیش‌آزمون وجود ندارد ( $F = 0/210, P < 0/05$ )؛ بنابراین از



این نتایج می‌توان دریافت که قبل از انجام این مطالعه، هر سه گروه از لحاظ درک مطلب همگن بوده‌اند.

به‌منظور بررسی تفاوت موجود میان سه گروه از لحاظ درک مطلب و پس از انجام این مطالعه، نمرات پس‌آزمون هر سه گروه به‌وسیله روش آماری تحلیل واریانس یک‌طرفه مورد تحلیل و بررسی قرار گرفت.

جدول ۴. آمار توصیفی نمرات بخش خوانش پیش‌آزمون

تعداد	میانگین	انحراف معیار	خطای استاندارد	حداقل	حداکثر
۲۰	۲۴/۵۲۰۰	۳/۳۶	۰/۹۲۱۴۲	۱۰	۳۴
۲۰	۲۸/۲۸۰۰	۳/۱۲	۱/۲۳۲۵	۱۲	۳۵
۲۰	۲۳/۸۷۰۰	۳/۹۸	۰/۸۷۵۸۴۵	۹	۲۶

همانند فرض تحلیل واریانس یک‌طرفه قبل از اعمال این روش آماری نیز همگنی واریانس‌ها برقرار بود. قبل از انجام آزمون‌های میانگین، برابری واریانس‌ها در گروه‌ها را توسط آزمون لون بررسی شده است. در جدول ۵ آزمون لون برای بررسی برابری واریانس‌های خطا مربوط به امتیازات پس‌آزمون نشان داده شده است ( $P > 0/05$ ،  $F(2, 57) = 2/121$ ).

جدول ۵. آزمون لون برای بررسی برابری واریانس‌های خطا مربوط به امتیازات پس‌آزمون

F	df1	df2	معناداری
۲/۱۲۱	۲	۵۷	۰/۱۸۵

در جدول ۶ نیز نتایج تحلیل واریانس یک‌طرفه نمرات پس‌آزمون هر سه گروه را نشان داده شده است.

جدول ۶. نتایج تحلیل واریانس یک‌طرفه امتیازات پیش‌آزمون

مجموع مربعات	Df	میانگین مربعات	F	معناداری
۳۳۲/۰۳۳	۲	۱۶۶/۰۱۷	۳۸/۲۹۱	۰/۰۰۲
۲۴۰/۱۵۰	۵۷	۴/۲۱۳		
۵۷۲/۱۸۳	۵۹			

نتایج تحلیل واریانس در جدول ۶ نشان داد که تفاوت معناداری میان سه گروه تحت مطالعه وجود داشت ( $F = 38/29, P < 0/05$ ). از وجود تفاوت معنادار در امتیازات بخش خوائش مربوط به پس‌آزمون سه گروه می‌توان برداشت کرد که حداقل یکی از گروه‌ها نسبت به دو گروه دیگر یا حداقل نسبت به یک گروه دیگر دارای عملکرد بهتری بوده است. به‌منظور یافتن منشأ اصلی تفاوت میان سه گروه، از آزمون تعقیبی توکی<sup>۱</sup> استفاده گردید. در جدول ۷ نتایج مقایسه‌های چندگانه انجام شده از طریق آزمون توکی برای نمرات بخش خوائش پس‌آزمون نشان داده شده است.

جدول ۷. نتایج مقایسه‌های چندگانه از طریق آزمون توکی

روش D)	روش J))	تفاضل میانگین (I-J)	خطای استاندارد		معناداری	
			خطای استاندارد	خطای استاندارد	معناداری	معناداری
ارزیابی پویا رایانه‌ای	ارزیابی پویا	*۷۶۰۰/۳	۶۴۹۰۹/۰	۰۰۲/۰	۳۱۲۰/۶-	۱۸۸۰/۳-
		*۶۵۰۰/۱۱	۶۴۹۰۹/۰	۰۰۵/۰	۱۱۲۰/۱-	۰۱۲۰/۲
ارزیابی پویا رایانه‌ای	ارزیابی پویا	*۷۶۰۰/۳-	۶۴۹۰۹/۰	۰۰۲/۰	۱۸۸۰/۳	۳۱۲۰/۶
		*۴۱۰۰/۱۵	۶۴۹۰۹/۰	۰۰۰/۰	۶۳۸۰/۳	۷۶۲۰/۶
ارزیابی پویا رایانه‌ای	ارزیابی پویا	*۶۵۰۰/۱۱-	۶۴۹۰۹/۰	۰۰۵/۰	۰۱۲۰/۲-	۱۱۲۰/۱
		*۴۱۰۰/۱۵-	۶۴۹۰۹/۰	۰۰۰/۰	۷۶۲۰/۶-	۶۳۸۰/۳-

بر اساس نتایج مقایسه‌های چندگانه (جدول ۷)، گروه ارزیابی پویا رایانه‌ای عملکرد بسیار بهتری نسبت به گروه کنترل داشته است ( $p = 0/00 < 0/05$ ). همچنین گروه ارزیابی پویا نیز عملکرد بسیار بهتری را نسبت به گروه کنترل از خود نشان داده است ( $p = 0/05 < 0/05$ ). بعلاوه، تفاوت معناداری در عملکرد گروه‌های ارزیابی پویا و پویا رایانه‌ای وجود دارد، به‌عبارت‌دیگر گروه ارزیابی پویا رایانه‌ای عملکرد بهتری نسبت به گروه ارزیابی پویا داشته است. پژوهشگران در این مطالعه، نگرش زبان‌آموزان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای را مورد بررسی قرار دادند.

یافته‌های این مطالعه نشان می‌دهند که امکان تمرکز بهتر در سیستم آزمون‌های رایج وجود دارد. باین‌حال، علیرغم شواهد برخی شرکت‌کنندگان در این مطالعه همچنان معتقدند که «بازنگری سؤالات سخت‌تر در زمان دیگر» و مرور پاسخ‌هایشان در آزمون‌های رایج، برای آن‌ها بهتر خواهد بود. دیدگاه یکی از شرکت‌کنندگان در این

1. Tukey

مطالعه: «همه متن را در اختیار داشتم و امکان رد شدن از سؤالات سخت‌تر و پاسخ دادن به آن‌ها در زمان دیگری وجود داشت. گاهی اوقات پاسخ یک سؤال را در لحظه فراموش می‌کنید ولی ممکن است بعد از مدتی آن را با یاد بیاورید که در این صورت شانس برگشتن به آن سؤال و پاسخ به آن وجود دارد».

مهم‌ترین مانعی که شرکت‌کنندگان در آزمون ارزیابی ایستا با آن مواجه بودند، نبود میانجی‌گری<sup>۱</sup> و اشاراتی بود که در زمان تردید میان گزینه‌های مختلف وجود داشت. اکثر شرکت‌کنندگان اظهار داشتند که در زمان آزمون‌های رایج ارزیابی ایستا، هیچ چیزی یاد نگرفته‌اند، زیرا هیچ‌گونه اشاره یا مداخله‌ای در میانه اجرای آزمون وجود نداشت. یکی از مصاحبه‌شوندگان در این باره گفت: «گاهی احساس گیج شدن می‌کردم زیرا بین انتخاب دو گزینه تردید داشتم و زمانی که یکی از آن‌ها را انتخاب می‌کردم، صحیح یا غلط بودن آن گزینه را نمی‌دانستم». موضوع استفاده از اشارات و راهنمایی‌ها در زمان انجام آزمون به‌روشنی در مصاحبه‌ها مورد تأکید قرار گرفته است و اهمیت میانجی‌گری در آموزش مدرن و تشویق برای خودراهبری<sup>۲</sup> نشان داده شده است. این مسئله یک مانع بوده که اگر دانشجویان از نرم‌افزار طراحی شده و آزمون مدرن استفاده می‌کردند، قادر بودند که بر این مانع غلبه پیدا کنند.

«مدیریت زمان» مانع مهم نهایی بوده که در زمان آزمون‌های ارزیابی ایستا و پویا که می‌بایست بر آن غلبه کرد. اکثر شرکت‌کنندگان به موضوع عدم مدیریت مناسب زمان در زمان آزمون اشاره داشتند. یکی از شرکت‌کنندگان درباره نگرانی بابت کمبود زمان کافی جهت پاسخ به سؤالات گفت: «آزمون برای من استرس‌زا بود زیرا نمی‌دانستم چه میزان وقت برای هر بخش دارم». این مسئله توسط برخی دانشجویان دیگر نیز بیان شد و آن را به‌عنوان مشکل و مانعی در به دست آوردن نمره مناسب می‌دانستند.

### بحث و نتیجه‌گیری

هدف این مطالعه، مقایسه اثرات ارزیابی‌های پویا رایج و رایانه‌ای بر درک مطلب است. بعلاوه، نگرش زبان‌آموزان نسبت به ارزیابی‌های ایستا، پویا و پویا رایانه‌ای نیز مورد بررسی قرار گرفت. نتایج تحلیل‌های آماری نشان می‌دهند که ارزیابی‌های پویا و پویا رایانه‌ای، اثرات قابل توجهی بر بهبود درک مطلب شرکت‌کنندگان داشته‌اند. همچنین مشخص شد

که ارزیابی پویا رایانه‌ای در مقایسه با ارزیابی پویا، بر درک مطلب مؤثرتر بوده است. تحلیل محتوای مصاحبه‌ها نشان داد که همه شرکت‌کنندگان نگرش مثبتی نسبت به ارزیابی پویا رایانه‌ای در مقایسه با ارزیابی‌های رایج پویا و ایستا داشتند.

یافته‌های این مطالعه، نتایج مطالعات پیشین درباره تأثیرات مثبت ارزیابی پویا بر مهارت‌ها و بخش‌های مختلف زبانی را تأیید می‌کند (ذوقی و مالهییر، ۲۰۱۳؛ برزگر و آذری‌زاد، ۲۰۱۴؛ آنتون، ۲۰۰۳). مطالعه انجام شده توسط حسامی و قادری (۲۰۱۴)، درباره اثر ارزیابی پویا در یادگیری واژگان توسط زبان‌آموزان ایرانی EFL که با نتایج مطالعه حاضر سازگار و هماهنگ است، نشان می‌دهد که ارزیابی پویا اثر مثبتی بر دانش واژگان زبان‌آموزان ایرانی داشته است. همچنین مطالعه انجام شده توسط امینی (۲۰۱۵) که با نتایج مطالعه حاضر نیز هماهنگی دارد، نشان داد که ارزیابی پویا اثر قابل توجهی بر عملکرد درک مطلب دارد. بعلاوه، مطالعه انجام شده توسط فانی و رشتچی (۲۰۱۵) نشان داد که ارزیابی پویا تأثیر مثبتی بر درک مطلب دارد. بر اساس نتایج مطالعه حاضر می‌توان استنتاج کرد که اثربخشی ارزیابی‌های پویا و پویا رایانه‌ای، مرتبط با هسته مشترک این دو ارزیابی است که همان پویایی در ارائه آزمون به زبان‌آموزان است.

برخی از مفاهیم و نتایج آموزشی از این مقاله وجود دارد که در زیر قابل ذکر است. اولین نتیجه، در نظر گرفتن افراد کم‌توان و یا ناتوان از لحاظ یادگیری است. برخلاف اشکال ارزیابی سنتی، روش ارزیابی پویا به‌طور کلی و ارزیابی پویا رایانه‌ای به‌طور خاص زبان‌آموزان را به‌واسطه اصول منطقه رشد مجاور (ZPD)، به موفقیت می‌رساند. اگرچه برخی از موضوعات مانند عدالت در آموزش، شیوه دسترسی به آن و غیره ممکن است برای ارزیابی پویا در یک حوزه تحقیقاتی خاص مزاحمت ایجاد کند، اما در اینجا «زبان‌آموزان کم‌توان در یادگیری» در یک شهر مرزی دورافتاده به شیوه منصفانه مورد بررسی قرار گرفتند و زبان‌آموزان نوعی از دستورالعمل را متناسب با اصول منطقه رشد مجاور آقای ویگوتسکی دریافت کنند.

شایان‌ذکر است که ارزیابی پویا و تمرکز آن بر تعامل، ممکن است بر مشارکت زبان‌آموزان زبان دوم و افزایش علاقه آنان به دروس مؤثر بوده باشد. در همین باره، بررسی انجام شده توسط ذوقی و مالهییر (۲۰۱۳) نشان می‌دهد که ارزیابی پویا اثر مثبتی بر انگیزه زبان‌آموزان داشته است. علاوه بر مطالعات تجربی که نتایج این مطالعه را تأیید می‌کنند، مطالعات تئوری نیز جهت تشریح نتایج مناسب هستند. به‌عنوان مثال، ارزیابی

پویا بر اساس اصول منطقه رشد مجاور است که موجب تشویق زبان‌آموزان و پیشرفت آنان نسبت به وضعیت توسعه قبلی آنان می‌گردد. همچنین این پیشرفت‌ها با کمک افراد با دانش مانند مدرسان افزایش می‌یابد. همچنین تعامل نیز عامل مهم دیگری در ارزیابی پویا است که در این مطالعه نیز مورد تأکید قرار گرفته است. همه مطالعات تئوری مبتنی بر منطقه رشد مجاور ویگوتسکی و نظریه اجتماعی او، در جهت تصدیق صحت و درستی نتایج این مطالعه می‌باشند.

از محدودیت‌های این مطالعه می‌توان به کوچک بودن مقیاس مطالعه، کمبود جلسات کمکی در ارزیابی رایانه‌ای، عدم امکان تشخیص اینکه در چه زمانی و در چه مواردی جواب صحیح حدس زده شده است، اشاره کرد. پیشنهاد می‌گردد که در تحقیقات آتی، تعداد شرکت‌کنندگان با تغییر اهداف مورد مطالعه و با استفاده از فن‌آوری‌های جدیدی ممکن است به میزان بیشتری افزایش یابد، همچنین عوامل دیگری همانند چگونگی تعامل استاد با دانشجو و دانشجو با همسالان و سایر ابزارهای انسانی یا فیزیکی را می‌توان کنترل کرد تا تفسیر بهتری از نتایج حاصل شود و در نهایت اثربخشی ارزیابی پویا به‌طور کلی و ارزیابی پویا رایانه‌ای به‌طور خاص بر روی سایر مهارت‌ها (صحبت کردن، گوش دادن و نوشتن) انجام گیرد.

### سیاسگزاری

از کلیه اساتید دانشجویان، زبان‌آموزان و همکاران عزیز و سایر کسانی که در انجام این پژوهش، ما را همراهی کردند، قدردانی می‌شود.

### منابع

حمادی، ح.، فقیهی، ع.، سیفی، م.، و ناطقی، ف. (۱۳۹۷). بررسی و تحلیل کیفی انواع مدل‌های آموزش دانش‌آموزان دوزبانه. *نشریه پژوهش در نظام‌های آموزشی*، ۱۲ (ویژه‌نامه)، ۸۳۹-۸۵۲.  
 محمدخانی، س.، قیومی، ع.، محمد داودی، الف.، امیرارجمندی، ن. (۱۳۹۶). عوامل مؤثر بر ارتقاء مؤلفه‌های فرهنگی آموزش زبان خارجی و ارائه یک الگوی مناسب (مورد مطالعه: زبان انگلیسی). *نشریه پژوهش در نظام‌های آموزشی*، ۱۱ (۳۹)، ۲۰۷-۲۱۰.

### References

- Anton, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals*, 42(3), 576-598.  
 Auerbach, C. F., & Silverstein, L. B. (2003). *Qualitative data: An introduction to coding and analysis*. New York, NY: New York University Press.

- Bakhoda, I., & Shabani, K. (2019). Enhancing L2 learners' ZPD modification through computerized-group dynamic assessment of reading comprehension. *Innovation in Language Learning and Teaching*, 13(1), 31-44.
- Barzegar, R., & Azarizad, R. (2014). Using dynamic assessment to improve L2 learners' knowledge of grammar: evidence from the tenses. In *Classroom-oriented Research* (pp. 219-227). Springer, Heidelberg.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Los Angeles, CA: Sage.
- Crick, R. D., & Yu, G. (2008). Assessing learning dispositions: Is the effective lifelong learning inventory valid and reliable as a measurement tool? *Educational Research*, 50(4), 387-402.
- Gunning, R. (1952). *The technique of clear writing*. New York: McGraw-Hill.
- Davin, K. J. (2013). Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom. *Language Teaching Research*, 17(3), 303-322.
- Davoudi, M., & Ataie-Tabar, M. (2015). The effect of computerized dynamic assessment of L2 writing on Iranian EFL learners' writing development. *International Journal of Linguistics and Communication*, 3(2), 176-186.
- Donaldson, R. P., & Haggstrom, M. A. (Eds.). (2006). *Changing language education through CALL*. New York, NY: Routledge.
- Ellis, R. (2008). *The study of second language acquisition*. England: Oxford University Press.
- Ebadi, S., & Saeedian, A. (2015). The effects of computerized dynamic assessment on promoting at-risk advanced Iranian EFL students' reading skills. *Issues in Language Teaching (ILT)*, 4(2), 1-26.
- Ebadi, S., & Saeedian, A. (2016). Planning future instructional programs through computerized L2 dynamic assessment. *Teaching English with Technology*, 16(4), 12-32.
- Guthke, J., & Beckmann, J. F. (2000). The learning test concept and its applications in practice. In C. S. Lidz & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 17-69). Amsterdam: Elsevier.
- Hamadi, H., Faghihi, A., Seifi, M., & Nateghi, F. (1397). Qualitative analysis of various models of bilingual students' education. *Journal of Research in Educational Systems*, 12 (Special Issue), 839-852. [In Persian]
- Hessamya, G., & Ghaderi, E. (2014). The role of dynamic assessment in the vocabulary learning of Iranian EFL learners. *Procedia - Social and Behavioral Sciences*, 98, 645 - 652.
- Inan, F. A., & Lowther, D. L. (2007). A comparison analysis of computer-supported learning models and guidelines. In F. M. M. Neto, & F. U. Brasilerio (Eds.), *Advances in computer-supported learning*. Hershey, PA: Information Science Publishing.
- Kern, R., & Warschauer, M. (2000). Theory and practice of network-based language teaching. In M. Warschauer, & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp.1-19). England: Cambridge University Press.
- Kol, S., & Schcolnik, M. (2008). Asynchronous forums in EAP: Assessment issues. *Language Learning & Technology*, 12 (2), 49-70.
- Lambert, D., & Lines, D. (2000). *Understanding assessment: Purposes, perceptions, practice*. London: RoutledgeFalmer.
- Levy, M. (1997). *Computer assisted language learning: Context and conceptualization*. Oxford, England: Clarendon Press.

- MehriKamrood, A., Davoudi, M., Amiria, S. M. R., & Ghaniabadi, S. (2018). Transcendence of learning in an online computerized dynamic test of English listening. *CALL-EJ*, 19(1), 23-42.
- Mohammadkhani, S., Qayyumi, A., Mohammad Davoodi, A., & Amir Arjomandi, N. (1396). Factors affecting the promotion of cultural components of foreign language teaching and providing a suitable model (read: English language). *Journal of Research in Educational Systems*, 11 (39), 207-210. [In Persian]
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: TheMIT Press.
- Naeini, J., & Duvall, E. (2012). Dynamic assessment and the impact on English language learners' reading comprehension performance. *Language Testing in Asia*, 2 (2), 22-41.
- Navarro, J. J., Mourgues-Codern, C., Guzmán, E., Rodríguez-Ortiz, I. R., Conejo, R., Sánchez-Gutiérrez, C., De la Fuente, J., Martella, D., & Saracostti, M. (2018). Integrating curriculum-based dynamic assessment in computerized adaptive testing: Development and predictive validity of the EDPL-BAI battery on reading competence. *Frontiers in Psychology*, 9(1), 1-21.
- Nirmalakhandan, N. (2007). Computerized dynamic assessment. *International Conference on Engineering Education*, 1-6.
- Nirmalakhandan, N. (2009). Use of computerized dynamic assessment to improve student achievement: Case study. *Journal of Professional Issues in Engineering Education and Practice*, 135(2), 75-80.
- Pishghadam, R., & Barabadi, E. (2012). Constructing and validating computerized dynamic assessment of L2 reading comprehension. *Iranian Journal of Applied Linguistics (IJALI)*, 15(1), 73-95.
- Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9(3), 233-265.
- Ragan, L. C. (1998). Good teaching is good teaching: An emerging set of guiding principles and practices for the design and development of distance education. *DEOSNEWS*, 8 (12).
- Shabani, K. (2012). Dynamic assessment of L2 learners' reading comprehension processes: A Vygotskian perspective. *Procedia - Social and Behavioral Sciences*, 32, 321-328.
- Silverman, D. (2019). *Doing qualitative research: A practical handbook* (4<sup>th</sup>ed). London: SAGE.
- Simuth, J., & Sarmany-Schuller, I. (2013). Cognitive style variable in e-learning. *5th World Conference on Educational Sciences*, Italy, Rome.
- Simuth, J., & Sarmany-Schuller, I. (2014). Cognitive style variable in e-learning. *Procedia-Social and Behavioral Sciences*, 116, 1464-1467.
- Swanson, H. L., & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research*, 71(2), 321 - 363.
- Tzuriel, D., & Shamir, A. (2002). The effects of mediation in computer assisted dynamic assessment. *Journal of Computer Assisted Learning*, 18(1), 21-32.
- Vygotsky, L. S. (1986). *Thought and language*. (A. Kozulin, trans.). Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1994). The socialist relation of man. In R. de Veer, & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 175-184). Oxford, England: Blackwell.
- Xiaoxiao, L., & Yan, L. (2010). A case study of dynamic assessment in EFL process writing. *Chinese Journal of Applied Linguistics*, 33(1), 24-40.
- Yildirim, A. G. O. (2008). Vygotsky's sociocultural theory and dynamic assessment in language learning. *Anadolu University Journal of Social Sciences*, 8(1), 301-308.