



The University of
Tehran Press

Journal of Foreign Language Research

Online ISSN: 2588-7521

Journal Homepage: <https://jflr.ut.ac.ir/>



EFL Teachers' Critical Pedagogy Perceptions and Assessment Practices: A Needs Analysis Approach to Iranian Secondary School ELT Curriculum

Samaneh Abbaszadeh¹ Saeed Ahmadi² Mohammad Exir³

1. Department of Foreign Languages, Bu.C, Islamic Azad University, Bushehr, Iran. Email: samaneh.abbaszadeh8501@iau.ir

2. Department of Foreign Languages, Bu.C, Islamic Azad University, Bushehr, Iran. Email: s.ahmadi20@iau.ac.ir

3- Department of Foreign Languages, Bu.C, Islamic Azad University, Bushehr, Iran. Email: mohammad.exir@iau.ac.ir

Article Info

Article type:
Research Article

Article history:

Received: 2 January 2026

Received in revised form: 23
April 2026

Accepted: 27 April 2026

Available online : 22 June
2026

Keywords:

*Dynamic assessment, group
dynamic assessment,
argumentative reading, EFL
learners, Iranian context*

ABSTRACT

Objective: Argumentative reading comprehension requires higher-order skills such as evaluation, reasoning, and critical thinking, yet these abilities remain underdeveloped among many English as a Foreign Language (EFL) learners in Iran, where static, grammar-focused testing dominates. While dynamic assessment (DA) has been shown to enhance language skills, few studies have directly compared different group-based DA models in the domain of argumentative reading. This study investigated the effectiveness of interventionist group dynamic assessment (IGDA) and cumulative group dynamic assessment (CGDA) in comparison with static assessment (SA).

Methodology: Ninety low-intermediate Iranian EFL learners (aged 18–25) participated in a six-week intervention and completed pretests, posttests, and interviews.

Findings: Quantitative analyses revealed that both IGDA and CGDA led to significantly greater gains in argumentative reading comprehension than SA. IGDA proved most effective for fostering immediate analytical strategies through teacher mediation, whereas CGDA promoted deeper, sustained comprehension by encouraging collaborative reflection across sessions. Qualitative findings highlighted that dynamic assessment reduced anxiety, strengthened metacognitive awareness, and helped learners internalize reasoning strategies, while SA offered little developmental value.

Conclusion: These results extend DA research by distinguishing the complementary strengths of IGDA and CGDA, particularly in addressing the challenges of argumentative reading. The study underscores the need to move beyond static testing in Iranian EFL contexts and suggests that incorporating DA can cultivate higher-order comprehension and critical literacy.

Cite this article: Abbaszadeh,S , Ahmadi,S and Exir,M . (2026). The Impact of Interventionist and Cumulative Group Dynamic Assessment versus Static Assessment on Iranian EFL Learners' Argumentative Reading Comprehension. *Journal of Foreign Language Research*, 16(2), 199-219.
<https://doi.org/10.22059/jflr.2026.413170.1286>



© Author(s) retain the copyright.

Publisher: University of Tehran Press.

DOI: <https://doi.org/10.22059/jflr.2026.413170.1286>



پژوهشهای زبانشناختی در زبانهای خارجی

شاپا الکترونیکی: ۷۵۲۱-۲۵۸۸

Journal Homepage: <https://jflr.ut.ac.ir/>



انتشارات دانشگاه تهران

تأثیر ارزشیابی پویای گروهی مداخله‌گر و تجمعی در مقایسه با ارزشیابی ایستا بر درک مطلب خواندن استدلالی زبان آموزان ایرانی زبان انگلیسی به‌عنوان زبان خارجی

سمانه عباس زاده^۱ سعید احمدی^۲ محمد اکسیر^۳

۱. گروه زبان‌های خارجی، دانشگاه آزاد اسلامی واحد بوشهر، بوشهر، ایران. رایانامه: samaneh.abbaszadeh8501@iau.ir

۲. گروه زبان‌های خارجی، دانشگاه آزاد اسلامی واحد بوشهر، بوشهر، ایران. رایانامه: s.ahmadi20@iau.ac.ir

۳- گروه زبان‌های خارجی، دانشگاه آزاد اسلامی واحد بوشهر، بوشهر، ایران. رایانامه: mohammad.exir@iau.ac.ir

چکیده

اطلاعات مقاله

هدف: درک مطلب خواندن استدلالی^۱ مستلزم مهارت‌های سطح بالایی مانند ارزیابی، استدلال و تفکر انتقادی است؛ با این حال، در میان بسیاری از زبان‌آموزان زبان انگلیسی در ایران، این توانایی‌ها به اندازه کافی رشد نکرده‌اند. دلیل اصلی این کاستی را باید در غلبه آزمون‌های ایستا و دستور-محور جست‌وجو کرد که در نظام آموزشی ایران حاکمند. این مطالعه به بررسی اثربخشی ارزشیابی پویای گروهی مداخله‌گر^۲ و ارزشیابی پویای گروهی تجمعی^۳ در مقایسه با ارزشیابی ایستا^۴ پرداخت.

روش شناسی: نود زبان‌آموز ایرانی زبان انگلیسی در سطح پایین‌تر از متوسط با دامنه سنی ۱۸ تا ۲۵ سال که در مؤسسه‌های خصوصی آموزش زبان در شهر بوشهر مشغول به تحصیل بودند، در یک آزمایش شش‌هفته‌ای شرکت کردند و در پیش‌آزمون‌ها، پس‌آزمون‌ها و مصاحبه‌ها حضور یافتند.

یافته‌ها: تحلیل‌های کمی با استفاده از تحلیل واریانس آمیخته نشان داد که هر دو نوع ارزشیابی پویای گروهی مداخله‌گر و تجمعی در مقایسه با ارزشیابی ایستا به پیشرفت معناداری در درک مطلب خواندن استدلالی منجر شدند. ارزشیابی پویای گروهی مداخله‌گر در تقویت راهبردهای تحلیلی فوری از طریق میانجی‌گری معلم مؤثرتر بود، در حالی که ارزشیابی پویای گروهی تجمعی با تشویق به تفکر مشارکتی، درک عمیق‌تر و پایدارتر را ارتقا داد. یافته‌های کیفی نشان داد که ارزشیابی پویا موجب کاهش اضطراب، تقویت آگاهی فراشناختی و کمک به درونی‌سازی راهبردهای استدلالی در زبان‌آموزان شد، در حالی که ارزشیابی ایستا به رشد یادگیری زبان‌آموزان کمک چندانی نکرد.

نتیجه‌گیری: این نتایج، با آشکار ساختن نقاط قوت مکمل ارزشیابی پویای گروهی مداخله‌گر و تجمعی، به‌ویژه در رویارویی با چالش‌های خواندن استدلالی، افق‌های تازه‌ای در پژوهش‌های ارزشیابی پویا می‌گشایند. این مطالعه پیشنهاد می‌کند به‌کارگیری ارزشیابی پویا ظرفیت آن را دارد که درک مطلب سطح بالا و سواد انتقادی را در زبان‌آموزان پرورش دهد.

استاد: باس زاده، سمانه، احمدی، سعید و اکسیر، محمد. (۱۴۰۵). تأثیر ارزشیابی پویای گروهی مداخله‌گر و تجمعی در مقایسه با ارزشیابی ایستا بر درک مطلب خواندن استدلالی زبان‌آموزان ایرانی زبان انگلیسی به‌عنوان زبان خارجی. *پژوهشهای زبانشناختی در زبانهای خارجی*، ۱۶ (۲)، ۱۹۹-۲۱۹.

<https://doi.org/10.22059/jflr.2026.413170.1286>



© نویسندگان.

ناشر: انتشارات دانشگاه تهران.

¹ Argumentative reading comprehension

² Interventionist Group Dynamic Assessment (IGDA)

³ Cumulative Group Dynamic Assessment (CGDA)

⁴ Static Assessment (SA)

⁵ Dynamic Assessment (DA)

۱- مقدمه

درک مطلب خواندن یکی از مهارت‌های بنیادین برای زبان‌آموزان انگلیسی است. این مهارت نه تنها پایه‌گذار موفقیت تحصیلی است، بلکه نقش مهمی در یادگیری مؤثر واژگان و ساختارهای دستوری، عملکرد بهتر در آزمون‌های استاندارد^۱، و پیشرفت در تحصیلات دانشگاهی دارد. همچنین، درک مطلب خواندن با برقراری ارتباط مؤثر مرتبط است، زیرا خواننده‌ای که متون را به خوبی درک می‌کند، توانایی بیشتری در تولید پیام‌های نوشتاری و گفتاری روشن و مستدل دارد. در این میان، متون استدلالی در مقایسه با متون روایی^۲ یا توصیفی^۳ پیچیده‌تر هستند، زیرا خواندن آن‌ها از زبان‌آموزان می‌خواهد تا ادعاها را شناسایی، شواهد را ارزیابی، و ساختارهای استدلالی را تحلیل کنند (کوهن، ۲۰۱۷؛ گریب و ژانگ، ۲۰۲۲). این فرایندها به تفکر سطح بالا و سواد انتقادی نیاز دارند؛ مهارت‌هایی که بسیاری از زبان‌آموزان به دلیل دانش زبانی محدود و مواجهه ناکافی با متون اصیل، در کسب آن‌ها با دشواری روبه‌رو هستند.

شیوه‌های ارزشیابی در آموزش و یادگیری درک مطلب خواندن نقش مهمی دارند. در بسیاری از نظام‌های آموزشی، از جمله ایران، ارزشیابی‌های ایستا مثل آزمون‌های چندگزینه‌ای هنوز هم رایج هستند. این نوع ارزشیابی‌ها فقط دانش فعلی زبان‌آموزان را می‌سنجند و مشکلات زیادی دارند از جمله اینکه نمی‌توانند تفکر انتقادی و مهارت‌های استدلالی را بسنجند، فقط بر حفظیات و اطلاعات سطحی تمرکز می‌کنند، فرایند یادگیری و رشد تدریجی زبان‌آموزان را نادیده می‌گیرند، و بازخورد مؤثر و رشدی ارائه نمی‌دهند (اسمیت و همکاران، ۲۰۲۱). چنین روش‌هایی فشار و اضطراب روانی ایجاد می‌کنند و مانع از پیشرفت زبان‌آموزان در جریان ارزشیابی می‌شوند (استیگینز، ۲۰۰۵). در ایران، ارزشیابی بیشتر بر دستور زبان و حفظ کردن مطالب استوار است (رحیمی‌نیا، ۲۰۱۶). به همین دلیل، فرصت‌های بسیار کمی برای پرورش مهارت‌های استدلالی لازم برای درک مطلب خواندن استدلالی در اختیار زبان‌آموزان قرار می‌گیرد (غفور و رحیمی، ۲۰۲۱).

ارزشیابی پویا به عنوان مکملی برای آزمون‌های ایستا معرفی شده است. این رویکرد ریشه در نظریه اجتماعی - فرهنگی^۴ و یگوتسکی دارد و بر این نکته تأکید می‌کند که یادگیری از طریق میانجی‌گری در منطقه تقریبی رشد^۵ اتفاق می‌افتد. در این چارچوب، ارزشیابی با آموزش تلفیق می‌شود تا حمایت و بازخورد در طول فرایند فراهم گردد (لنتوف و پوهنر، ۲۰۱۴). زبان‌آموزان برای حل تکالیفی هدایت می‌شوند که به تنهایی قادر به انجام آن‌ها نیستند. این رویکرد نه تنها توانایی کنونی آنان، بلکه ظرفیت بالقوه رشدشان را نیز آشکار می‌سازد (ورتش، ۲۰۰۸). میانجی‌گری می‌تواند از طریق راهنمایی‌ها، سرنخ‌ها یا تعامل سازمان‌یافته ارائه شود و به زبان‌آموزان کمک کند تا راهبردهای جدید را درونی‌سازی کنند. از این رو، ارزشیابی پویا فرایندی تلقی می‌شود که همزمان با سنجش عملکرد، یادگیری را نیز تقویت می‌کند.

رویکردهای متفاوتی برای ارزشیابی پویا ارائه شده‌اند. ارزشیابی پویای مداخله‌گر از میانجی‌گری ساختاریافته استفاده می‌کند؛ به این صورت که حمایت به شیوه‌ای سلسله‌مراتبی و از راهنمایی‌های ضمنی به کمک‌های صریح ارائه می‌شود (پوهنر و لنتوف، ۲۰۱۰). در مقابل، ارزشیابی پویای تجمعی، میانجی‌گری را به صورت مکرر در طول جلسات فراهم می‌کند تا زبان‌آموزان بتوانند به تدریج درک خود را تثبیت کنند (کرامش، ۲۰۱۵). پژوهش‌ها نشان داده‌اند که هر دو رویکرد برای مهارت‌هایی مانند واژگان و درک شنیداری مؤثر هستند (میرزایی و همکاران، ۲۰۱۷؛ رضایی و همکاران، ۲۰۲۳). با این حال، توجه اندکی به خواندن استدلالی معطوف شده است؛ مهارتی که به استدلال پیچیده‌تری نیاز دارد. همچنین پژوهش‌های بسیار کمی به مقایسه‌ی مستقیم ارزشیابی پویای مداخله‌گر و تجمعی در این حوزه پرداخته‌اند و اثربخشی نسبی این دو رویکرد همچنان روشن نیست (چنگ و همکاران، ۲۰۲۳).

¹ Standardized tests² Narrative³ Descriptive⁴ Sociocultural theory⁵ Zone of Proximal Development (ZPD)

این خلأ پژوهشی در بافت آموزشی ایران آشکارتر است. اگرچه ارزشیابی پویا در حوزه‌هایی مانند یادگیری واژگان و درک شنیداری به کار گرفته شده است (صفایی‌اصل و همکاران، ۲۰۲۱؛ رضایی و همکاران، ۲۰۲۳)، اما نقش آن در خواندن استدلالی کمتر بررسی شده است. شمار پژوهش‌هایی که به مقایسه‌ی مستقیم ارزشیابی پویای گروهی مداخله‌گر و جمعی پرداخته‌اند نیز بسیار اندک است. در نتیجه، این پرسش همچنان بی‌پاسخ مانده است که این رویکردها چگونه می‌توانند از زبان‌آموزان در پرورش درک مطلب خواندن استدلالی حمایت کنند.

پژوهش حاضر سه روش ارزشیابی را در کلاس‌های زبان انگلیسی ایران با هم مقایسه می‌کند: ارزشیابی پویای گروهی مداخله‌گر، ارزشیابی پویای گروهی جمعی، و ارزشیابی ایستا. اهمیت این پژوهش در دو چیز است: یکی از نظر نظری و دیگری از نظر آموزشی. از نظر نظری، این پژوهش ارزشیابی پویا را به حوزه‌ی خواندن استدلالی می‌برد؛ حوزه‌ای که به تفکر انتقادی و مهارت ارزیابی نیاز دارد. از نظر آموزشی، یافته‌ها می‌تواند نشان دهد که چگونه می‌توان در کلاس‌های ایران از ارزشیابی پویا بهره گرفت تا سطح درک مطلب زبان‌آموزان ارتقا یابد.

۲- پیشینه

۲-۱- درک مطلب خواندن در بافت‌های زبان انگلیسی به‌عنوان زبان خارجی

درک مطلب خواندن مهارتی پیچیده است که سه بخش را در بر می‌گیرد: رمزگشایی متن، یکپارچه‌سازی اطلاعات، و ارزیابی انتقادی. پژوهشگران معمولاً بین دو نوع درک مطلب تمایز می‌گذارند: درک لفظی که به یادآوری جزئیات آشکار متن اشاره دارد، و درک استنباطی که نیاز به تفسیر، پیش‌بینی و ارزیابی دارد (شاناهان و شاناهان، ۲۰۲۰؛ کینتس و راوسن، ۲۰۲۱). برای زبان‌آموزان انگلیسی، هر دو نوع دشوار است، زیرا آنها همزمان با دشواری‌های زبانی و نیز هنجارهای فرهنگی و بلاغی ناآشنا روبه‌رو می‌شوند (گریب و ژانگ، ۲۰۲۲).

دانش واژگانی همیشه یکی از عوامل مؤثر بر درک مطلب شناخته شده است. دایره واژگان محدود، توانایی زبان‌آموزان را در رمزگشایی متن و حدس زدن معنا کاهش می‌دهد (ژانگ و لیو، ۲۰۲۳). همچنین دانش پیشین نقش مهمی دارد، زیرا آشنایی با موضوع، فرایند استنباط و یکپارچه‌سازی اطلاعات را آسان‌تر می‌کند (گریب و ژانگ، ۲۰۲۲). افزون بر این، خواندن مؤثر نیازمند به کارگیری راهبردهایی مانند پیش‌بینی، خلاصه‌سازی و پرسش‌گری است. پژوهش‌ها نشان داده‌اند که آموزش مستقیم این راهبردها می‌تواند درک مطلب زبان‌آموزان انگلیسی را، به ویژه در متون پیچیده، بهبود بخشد (آلبریشت و پیرسون، ۲۰۲۰).

فراشناخت نیز در درک مطلب نقش اساسی دارد و به توانایی نظارت و تنظیم درک فرد از متن گفته می‌شود. راهبردهای فراشناختی، مانند برنامه‌ریزی برای خواندن متن، بررسی میزان فهم در حین خواندن، و ارزیابی درک پس از خواندن، برای یادگیری عمیق ضروری هستند (ژانگ و لیو، ۲۰۲۳). اما ارزشیابی‌های سنتی به ندرت این راهبردها را می‌سنجند یا به کارگیری آنها را تشویق می‌کنند. به همین دلیل، نیاز به رویکردهای جایگزین، مانند ارزشیابی پویا، برجسته می‌شود.

۲-۲- درک مطلب خواندن استدلالی

درک مطلب خواندن مهارتی پیچیده است که سه بخش را در بر می‌گیرد: رمزگشایی متن، یکپارچه‌سازی اطلاعات، و ارزیابی انتقادی. پژوهشگران معمولاً بین دو نوع درک مطلب تمایز می‌گذارند: درک لفظی که به یادآوری جزئیات آشکار متن اشاره دارد، و درک استنباطی که نیاز به تفسیر، پیش‌بینی و ارزیابی دارد (شاناهان و شاناهان، ۲۰۲۰؛ کینتس و راوسن، ۲۰۲۱). برای زبان‌آموزان انگلیسی، هر دو نوع دشوار است، زیرا آنها همزمان با دشواری‌های زبانی و نیز هنجارهای فرهنگی و بلاغی ناآشنا روبه‌رو می‌شوند (گریب و ژانگ، ۲۰۲۲).

دانش واژگانی همیشه یکی از عوامل مؤثر بر درک مطلب شناخته شده است. دایره واژگان محدود، توانایی زبان‌آموزان را در رمزگشایی متن و حدس زدن معنا کاهش می‌دهد (ژانگ و لیو، ۲۰۲۳). همچنین دانش پیشین نقش مهمی دارد، زیرا آشنایی با موضوع، فرایند استنباط و یکپارچه‌سازی اطلاعات را آسان‌تر می‌کند (گریب و ژانگ، ۲۰۲۲). افزون بر این، خواندن مؤثر نیازمند به

کارگیری راهبردهایی مانند پیش‌بینی، خلاصه‌سازی و پرسش‌گری است. پژوهش‌ها نشان داده‌اند که آموزش مستقیم این راهبردها می‌تواند درک مطلب زبان آموزان انگلیسی را، به ویژه در متون پیچیده، بهبود بخشد (آلبریشت و پیرسون، ۲۰۲۰).

فرانشاخت نیز در درک مطلب نقش اساسی دارد و به توانایی نظارت و تنظیم درک فرد از متن گفته می‌شود. راهبردهای فرانشاختی، مانند برنامه‌ریزی برای خواندن متن، بررسی میزان فهم در حین خواندن، و ارزیابی درک پس از خواندن، برای یادگیری عمیق ضروری هستند (ژانگ و لیو، ۲۰۲۳). اما ارزشیابی‌های سنتی به ندرت این راهبردها را می‌سنجند یا به کارگیری آنها را تشویق می‌کنند. به همین دلیل، نیاز به رویکردهای جایگزین، مانند ارزشیابی پویا، برجسته می‌شود.

۲-۳- ارزشیابی در آموزش زبان

ارزشیابی در آموزش زبان نقش محوری دارد؛ زیرا هم به معلمان اطلاعاتی درباره دانش زبان آموزان می‌دهد و هم تصمیم‌های آموزشی را جهت می‌دهد. در این حوزه، تمایز شناخته‌شده‌ای بین دو نوع ارزشیابی وجود دارد: ارزشیابی تکوینی^۱ و ارزشیابی پایانی^۲. ارزشیابی تکوینی به فعالیت‌هایی گفته می‌شود که در جریان یادگیری انجام می‌گیرند. هدف آن فراهم کردن بازخورد فوری و کمک به بهبود یادگیری است. نمونه‌هایی از این نوع ارزشیابی عبارتند از: پرسش و پاسخ کلاسی، آزمون‌های کوتاه، و ارزشیابی همتایان. این روش‌ها به معلمان امکان می‌دهند نقاط ضعف زبان آموزان را شناسایی کنند و در عین حال به خود زبان آموزان کمک می‌کنند تا یادگیری خود را تنظیم نمایند (بلک و ویلیام، ۱۹۹۸؛ ویلیام، ۲۰۱۱).

در مقابل، ارزشیابی پایانی در پایان یک دوره آموزشی انجام می‌شود تا پیشرفت زبان آموزان بر اساس معیارهایی از پیش تعیین‌شده (مانند آزمون‌های نهایی یا استاندارد) سنجیده شود (پوفام، ۲۰۰۱). هرچند ارزشیابی پایانی برای پاسخگویی آموزشی ضروری است، اما معمولاً تصویری ایستا از عملکرد ارائه می‌دهد و نشان نمی‌دهد که توانایی‌های زبان آموزان چگونه می‌تواند با دریافت حمایت رشد کند.

محدودیت‌های ارزشیابی پایانی در بافت‌های آموزش زبان انگلیسی به عنوان زبان خارجی آشکارتر است. در این بافت‌ها، آزمون‌های استاندارد معمولاً تأکید زیادی بر دستور زبان و واژگان دارند و در نتیجه توانایی ارتباطی واقعی زبان آموزان را به درستی منعکس نمی‌کنند. پژوهش‌ها نشان داده‌اند که چنین آزمون‌هایی توانایی‌های سطح بالاتر مانند استنباط، ارزیابی و استدلال را به درستی نمی‌سنجند (اسمیت و همکاران، ۲۰۲۱). علاوه بر این، ارزشیابی پایانی اغلب اضطراب را افزایش می‌دهد، به ویژه در زبان آموزانی که اعتماد به نفس کمتری در توانایی‌های زبانی خود دارند (استیگنیز، ۲۰۰۵). به همین دلیل، پژوهشگران بر نیاز به الگوهای ارزشیابی تأکید کرده‌اند که یادگیری را تقویت کنند و به جای اندازه‌گیری صرف پیشرفت ایستا، رشد تدریجی زبان آموزان را نشان دهند.

انواع ارزشیابی در آموزش زبان را نمی‌توان به صورت رویکردهایی کاملاً مستقل از یکدیگر در نظر گرفت، زیرا هر یک جنبه متفاوتی از فرایند یادگیری را آشکار می‌کنند. ارزشیابی تکوینی بر پایش و بهبود یادگیری در جریان آموزش تمرکز دارد، در حالی که ارزشیابی پایانی میزان دستیابی زبان آموزان به اهداف آموزشی را در پایان دوره نشان می‌دهد. با این حال، هر دو رویکرد عمده‌تاً بر عملکرد مشاهده‌شده زبان آموز تکیه دارند و اطلاعات محدودی درباره ظرفیت بالقوه یادگیری او فراهم می‌کنند. در پاسخ به این محدودیت، ارزشیابی پویا مطرح شد که با تلفیق سنجش و میانجیگری آموزشی، نه تنها عملکرد موجود، بلکه توانایی رشد و تحول یادگیرنده را نیز مورد توجه قرار می‌دهد. از این رو، ارزشیابی پویا را می‌توان گسترشی بر رویکردهای سنتی ارزشیابی دانست که بر فرایند رشد یادگیرنده تأکید بیشتری دارد.

۲-۴- ارزشیابی پویا: مبانی نظری

ارزشیابی پویا ریشه در نظریه اجتماعی-فرهنگی ویگوتسکی دارد که بر نقش میانجی‌گری در یادگیری تأکید می‌کند (ویگوتسکی، ۱۹۷۸). ارزشیابی پویا این چارچوب نظری را به حوزه سنجش می‌گستراند و آموزش و آزمون‌سازی را با یکدیگر تلفیق می‌کند

^۱ Formative assessment

^۲ Summative assessment

(لنتولف و پوهنر، ۲۰۱۴). در دو دهه گذشته، ارزشیابی پویا به طور گسترده در یادگیری زبان دوم به کار رفته است، به ویژه در زمینه‌هایی که میانجی‌گری و داربست‌زنی برای رشد زبان‌آموزان ضروری تلقی می‌شوند. اصل اساسی در این رویکرد این است که ارزشیابی نباید تنها آنچه را که زبان‌آموزان به تنهایی می‌توانند انجام دهند، اندازه بگیرد؛ بلکه باید ظرفیت بالقوه آنان برای رشد را در صورت دریافت میانجی‌گری آشکار سازد (لنتولف، ۲۰۰۶). این اصل بر مفهوم منطقه تقریبی رشد استوار است؛ فاصله‌ای میان آنچه زبان‌آموزان به تنهایی می‌توانند انجام دهند و آنچه با راهنمایی دیگران قادر به انجام آن هستند (ورتش، ۲۰۰۸). در این چارچوب، یادگیری ماهیتی ذاتاً اجتماعی دارد و پیشرفت از طریق تعامل با معلمان یا همتایان حاصل می‌شود.

در ارزشیابی پویا، میانجی‌گری با فرایند سنجش ادغام می‌شود تا زبان‌آموزان را در حل تکالیفی هدایت کند که در غیر این صورت فراتر از توانایی مستقل آنان است. این میانجی‌گری می‌تواند به شکل راهنمایی، سرنخ، توضیح، یا گفت‌وگوی مشارکتی ارائه شود. این فرایند نه تنها میزان پاسخ‌دهی زبان‌آموزان به حمایت را ارزیابی می‌کند، بلکه شایستگی آنان را نیز گسترش می‌دهد. پژوهشگران استدلال کرده‌اند که همین ویژگی، ارزشیابی پویا را از نظر آموزشی معنادارتر از آزمون‌های ایستا می‌سازد؛ آزمون‌هایی که فقط عملکرد موجود را ثبت می‌کنند (پوهنر و لنتولف، ۲۰۱۰). نکته مهم این است که ارزشیابی پویا با ترکیب سنجش و آموزش، جدایی سنتی میان این دو حوزه را به چالش می‌کشد.

جذابیت ارزشیابی پویا در کارکرد دوگانه آن نهفته است. این رویکرد هم‌زمان تشخیص‌دهنده و رشد‌دهنده است. ارزشیابی پویا از طریق داربست‌زنی در جریان سنجش، تصویری پویا از توانایی زبان‌آموز ارائه می‌دهد و به او کمک می‌کند تا راهبردهایی را درونی‌سازی کند که قابلیت انتقال به تکالیف آینده را دارند (پوهنر، ۲۰۰۸). به همین دلیل، پژوهشگران بر این باورند که ارزشیابی پویا نشان‌دهنده تغییری در سنجش آموزشی است؛ تغییری که تمرکز را از نتایج نهایی به فرایندهای یادگیری معطوف می‌کند.

۲-۵- انواع ارزشیابی پویا

پژوهش‌ها دو رویکرد اصلی را برای ارزشیابی پویا شناسایی کرده‌اند: رویکرد مداخله‌گر که بر راهنمایی‌های سازمان‌یافته و رویه‌های استاندارد متکی است (لیدز و الیوت، ۲۰۰۰)، و رویکرد تعامل‌گرا که بر میانجی‌گری گفت‌وگومحور و حمایت انعطاف‌پذیر معلم تأکید دارد (پوهنر، ۲۰۰۸). مطالعات جدید نشان داده‌اند که هر دو رویکرد می‌توانند یادگیری را ارتقا دهند، هرچند از نظر میزان کنترل و انعطاف‌پذیری با هم تفاوت دارند.

در رویکرد مداخله‌گر، معلمان یا پژوهشگران تنها در صورت نیاز حمایت ارائه می‌کنند و سطح کمک ثبت می‌شود تا ظرفیت بالقوه زبان‌آموز مشخص گردد. این ویژگی امکان استانداردسازی آسان‌تر و مقایسه زبان‌آموزان را فراهم می‌کند (پوهنر و لنتولف، ۲۰۱۰). این رویکرد به ویژه در زمینه‌هایی که بازخورد فوری به زبان‌آموزان کمک می‌کند تا دشواری‌ها را سریع‌تر برطرف کنند، اثربخش است (چنگ و همکاران، ۲۰۲۳). با این حال، ساختار نسبتاً خشک آن ممکن است تعامل آزاد و باز را محدود کند.

اگرچه ارزشیابی پویا اغلب در سطح فردی به کار می‌رود، پژوهشگران چگونگی عملکرد آن در زمینه‌های گروهی را نیز بررسی کرده‌اند. پوهنر (۲۰۰۹) ارزشیابی پویای گروهی را معرفی کرد که در آن میانجی‌گری در قالب فعالیت‌های مشارکتی صورت می‌گیرد. مطالعات جدیدتر نشان می‌دهند که ارزشیابی پویای گروهی می‌تواند از طریق تعامل همتایان، هم رشد زبانی و هم آگاهی فراشناختی را تقویت کند (ابلیوا و لنتولف، ۲۰۱۱؛ رسایی، ۲۰۲۳).

ارزشیابی پویای جمعی بر میانجی‌گری پیاپی در جلسات گوناگون تأکید دارد. در این روش، حمایت به تدریج و در طول زمان ارائه می‌شود تا زبان‌آموزان بتوانند درک خود را گام به گام تثبیت کنند. به باور کرامش (۲۰۱۵)، این رویکرد تصویر واقع‌بینانه‌تری از فرایند رشد درازمدت به دست می‌دهد؛ برخلاف یک جلسه میانجی‌گری کوتاه که تأثیر محدودی دارد. نقطه قوت ارزشیابی جمعی در ایجاد یادگیری عمیق‌تر و ماندگارتر است، هرچند زمان بیشتری از کلاس را می‌گیرد و استاندارد کردن آن دشوارتر است. پژوهش‌ها نشان داده‌اند که این نوع ارزشیابی می‌تواند به زبان‌آموزان کمک کند از راهبردهای مؤثرتر استفاده کنند و خودتنظیمی را در آنان تقویت کند (سنائی‌فر و دیوکلایی، ۲۰۱۹).

مطالعات مقایسه‌ای در این حوزه محدود هستند. بعضی از آن‌ها نشان داده‌اند که ارزشیابی پویای تجمعی برای یادگیری واژگان و هماندها (هم‌نشین‌های زبانی) مؤثرتر است (میرزایی و همکاران، ۲۰۱۷؛ فراهانی و مقدم، ۲۰۲۰). در مقابل، ارزشیابی پویای مداخله‌گر شاید برای کارهایی که نیاز به درک سریع یا حل مسئله دارند مناسب‌تر باشد (چنگ و همکاران، ۲۰۲۳). این نتایج یک پرسش اساسی را پیش می‌کشند: آیا نوع خاصی از ارزشیابی پویا برای خواندن متون پیچیده، مثلاً درک مطلب خواندن استدلالی، کارآمدتر است؟

ارزشیابی پویای گروهی مداخله‌گر به ویژه برای آن دسته از تکالیفی مناسب است که نیاز به درک سریع و حل فوری مسئله دارند؛ مانند تمرین‌های خواندن انتخابی و تولیدی. شواهد جدید نشان می‌دهد که دانش، آموزش و هویت حرفه‌ای معلم نقش بسیار مهمی در موفقیت این نوع ارزشیابی دارد (عسگری، ۲۰۲۴). پژوهش‌هایی که در ایران انجام شده تأیید می‌کند که استفاده از ارزشیابی پویای مداخله‌گر می‌تواند عملکرد زبان‌آموزان را به طور معناداری بهبود بخشد زیرا این زبان‌آموزان بازخوردی سازمان‌یافته و فوری دریافت می‌کنند که دقیقاً دشواری‌های خواندن را هدف می‌گیرد. به همین دلیل، این روش برای کمک به زبان‌آموزان در پرورش مهارت‌های تحلیلی مورد نیاز برای درک متون پیچیده بسیار مناسب است. نکته مهم این است که داربست‌سازی (حمایت سازمان‌یافته) از سوی معلم یا همکلاسی، که یکی از ویژگی‌های اصلی ارزشیابی پویاست، به رشد درک خواندن در درون منطقه تقریبی رشد زبان‌آموزان کمک می‌کند (پسندی و همکاران، ۲۰۲۳). این تعامل داربست-محور نه تنها به فهم فوری یاری می‌رساند، بلکه زبان‌آموزان را تشویق می‌کند راهبردهایی را درونی کنند که پایه و اساس خواندن در سطح بالاست.

۲-۶- مطالعات تجربی درباره ارزشیابی پویا و خواندن

در سال‌های اخیر، فناوری نقش مهمی در تقویت ارزشیابی پویا داشته است. برنامه‌های تلفن همراه، سکویهای برخط و ابزارهای هوش مصنوعی به کار گرفته شده‌اند تا میانجی‌گری را به شیوه‌های نوآورانه ارائه دهند. برای نمونه، مطالعه‌ای در ایران نشان داد که داربست‌زنی مبتنی بر هوش مصنوعی، هم درک مطلب خواندن و هم درگیری انتقادی با متون را بهبود می‌بخشد (محمدی و رضایی، ۲۰۲۵). سایر پژوهش‌ها نیز نقش فناوری‌های دیجیتال را در ارائه بازخورد فوری و فردی در تکالیف خواندن و نوشتن تأیید کرده‌اند (وانگ و شو، ۲۰۲۴). این یافته‌ها نشان می‌دهند که ارزشیابی پویا به سمت ادغام با نظام‌های هوشمند پیش می‌رود. شمار فزاینده‌ای از پژوهش‌ها به بررسی ارزشیابی پویا در ارتباط با خواندن و مهارت‌های وابسته به آن پرداخته‌اند. فانی و رشتچی (۲۰۱۵) نشان دادند که هر دو نوع ارزشیابی پویای فردی و گروهی، درک مطلب خواندن فراگیران ایرانی را بهبود می‌بخشد، هرچند تفاوت معناداری میان این دو قالب مشاهده نشد. میری و همکاران (۲۰۱۷) دریافتند که هر دو نوع ارزشیابی پویای گروهی (همزمان و تجمعی) یادگیری مقالات را تقویت می‌کنند، اما ارزشیابی همزمان اندکی اثربخش‌تر بود. رضایی و همکاران (۲۰۲۳) نیز نشان دادند که ارزشیابی پویای گروهی، آگاهی فراشناختی و درک شنیداری را بهبود می‌بخشد. این نتایج حاکی از آن است که فواید ارزشیابی پویا به مهارت‌های شناختی سطح بالا نیز گسترش می‌یابد.

در حوزه تخصصی خواندن استدلالی، چنگ و همکاران (۲۰۲۳) نشان دادند که ارزشیابی پویای مداخله‌گر، توانایی فراگیران را در تحلیل و ارزیابی استدلال‌ها تقویت می‌کند. این یافته‌ها نشان می‌دهند که هر دو شکل ارزشیابی پویا سودمندند، اما شیوه اثرگذاری آن‌ها در طول زمان ممکن است متفاوت باشد. با این حال، بیشتر این مطالعات دوره‌های آزمایش کوتاه و نمونه‌های نسبتاً کوچکی داشته‌اند؛ عاملی که تعمیم‌پذیری یافته‌ها را کاهش می‌دهد.

در یک پژوهش کنترل‌شده، فراگیرانی که میانجی‌گری مداخله‌گر را تجربه کرده بودند، عملکرد بهتری نسبت به همتایان خود در گروه ارزشیابی ایستا داشتند. به همین ترتیب، پسندی و همکاران (۲۰۲۳) با تحلیل ریزپردازشی نشان دادند که داربست‌زنی معلم و همسالان در عمل چگونه رخ می‌دهد و جزئیات راه‌های میانجی‌گری در آموزش خواندن را آشکار ساختند. پژوهش آن‌ها تأیید کرد که هر دو شکل داربست‌سازی مؤثر هستند، هرچند از سازوکارهای تعاملی متمایزی عمل می‌کنند. همچنین فرخی‌پور و همکاران (۲۰۲۰) با طراحی و اعتباریابی یک الگوی بومی ارزشیابی پویا نشان دادند که می‌توان از این رویکرد برای تشخیص و

رفع دشواری‌های یادگیری در مهارت‌های تولیدی دانش‌آموزان دبیرستانی بهره گرفت. در مجموع، این پژوهش‌ها حاکی از آن است که میانجی‌گری سازمان‌یافته در چارچوب ارزشیابی پویا، عملکرد خواندن زبان دوم و توانش راهبردی را بهبود می‌بخشد. علیرغم شواهد قابل توجه در خصوص اثربخشی ارزشیابی پویا، کاستی‌های مهمی در پیشینه پژوهش دیده می‌شود. نخست، بیشتر مطالعات بر مهارت‌های پایه‌تر مانند واژگان و درک شنیداری متمرکز بوده‌اند، در حالی که خواندن استدلالی، که به تلفیق دانش زبانی با توانایی‌های فراشناختی، تحلیل انتقادی و ارزیابی منطقی نیاز دارد، کمتر مورد توجه قرار گرفته است. دوم، حتی در پژوهش‌هایی که به ارزشیابی پویا در خواندن پرداخته‌اند، قالب غالب به کارگیری ارزشیابی فردی بوده و ارزشیابی گروهی به عنوان یک مداخله اجتماعی - فرهنگی کمتر واکاوی شده است. از میان مطالعات معدودی که ارزشیابی پویای گروهی را به کار گرفته‌اند، اکثراً به مقایسه آن با ارزشیابی ایستا یا فردی بسنده کرده‌اند و مقایسه مستقیم و کنترل‌شده میان دو رویکرد اصلی ارزشیابی پویای گروهی، یعنی مداخله‌گر و تجمعی، به ندرت در یک طرح پژوهشی واحد انجام شده است. سوم، در حوزه خاص خواندن استدلالی، هیچ مطالعه‌ای تأثیر همزمان دو رویکرد ارزشیابی پویای گروهی را در کنار ارزشیابی ایستا بررسی نکرده است.

به بیان روشن‌تر، شکاف اصلی پژوهشی در سه محور زیر قابل بررسی است:

- مشخص نیست کدام یک از دو رویکرد ارزشیابی پویای گروهی (مداخله‌گر یا تجمعی) تأثیر بیشتری بر بهبود درک مطلب خواندن استدلالی دارد.
- روشن نیست که آیا الگوهای زمانی اثرگذاری این دو رویکرد متفاوت است؛ به گونه‌ای که رویکرد مداخله‌گر تأثیر کوتاه‌مدت قوی‌تر و رویکرد تجمعی تأثیر پایدارتر و عمیق‌تری بر یادگیری داشته باشد.
- با توجه به غلبه ارزشیابی‌های ایستا و دستورمحرور در نظام آموزشی ایران، هیچ پژوهش نظام‌مندی این سه روش ارزشیابی (مداخله‌گر گروهی، تجمعی گروهی و ایستا) را در کنار یکدیگر و در حوزه خواندن استدلالی مقایسه نکرده است.

علاوه بر این، ضعف‌های روش‌شناختی پژوهش‌های پیشین، از جمله حجم نمونه کوچک، دوره‌های تحقیق کوتاه‌مدت (اغلب کمتر از یک ترم تحصیلی)، و تنوع محدود بافت‌های پژوهشی، تعمیم‌پذیری یافته‌های موجود را به شدت کاهش می‌دهد. در نتیجه، نیاز به پژوهش‌هایی با طرح منسجم، مقایسه‌ای نظام‌مند، و در برگیرنده هر سه متغیر اصلی (ارزشیابی پویای گروهی مداخله‌گر، ارزشیابی پویای گروهی تجمعی و ارزشیابی ایستا) در حوزه خواندن استدلالی کاملاً مشهود است.

برای پر کردن این شکاف پژوهشی، مطالعه حاضر به طور نظام‌مند ارزشیابی پویای گروهی مداخله‌گر و ارزشیابی پویای گروهی تجمعی را با ارزشیابی ایستا در حوزه درک مطلب خواندن استدلالی در میان فراگیران ایرانی زبان انگلیسی مقایسه می‌کند. این پژوهش از دیدگاه نظری به گسترش دامنه تحقیقات ارزشیابی پویا به سوی نوع متنی کمتر بررسی‌شده (خواندن استدلالی) و مقایسه مستقیم دو رویکرد گروهی آن کمک می‌کند. از نظر آموزشی نیز نشان می‌دهد که چگونه می‌توان شیوه‌های ارزشیابی در کلاس‌های زبان انگلیسی ایران را بازطراحی کرد تا درک انتقادی تقویت شود و اتکا به حفظ مطالب کاهش یابد.

بر این اساس، سوالات تحقیق زیر مطرح شدند:

۱. آیا بین سه روش ارزشیابی (ارزشیابی پویای گروهی مداخله‌گر، ارزشیابی پویای گروهی تجمعی و ارزشیابی ایستا) از نظر تأثیر بر درک مطلب خواندن استدلالی زبان‌آموزان ایرانی تفاوت معناداری وجود دارد؟ و آیا میزان پیشرفت از پیش‌آزمون به پس‌آزمون در این سه گروه متفاوت است؟
۲. دیدگاه زبان‌آموزان انگلیسی درباره مؤثر بودن سه روش ارزشیابی در بهبود درک مطلب خواندن استدلالی چیست؟

۳- روش پژوهش

۳-۱- طرح پژوهش

در این مطالعه از طرح پژوهشی تلفیقی^۱ بر اساس الگوی تبیینی متوالی^۲ استفاده شد. این طرح دو مرحله داشت. در مرحله نخست، از یک طرح شبه‌آزمایشی کمی شامل پیش‌آزمون، پس‌آزمون و گروه کنترل استفاده شد تا تأثیر روش‌های مختلف ارزشیابی بر درک مطلب خواندن استدلالی زبان‌آموزان سنجیده شود. سه گروه با یکدیگر مقایسه شدند: ارزشیابی پویای گروهی مداخله‌گر، ارزشیابی پویای گروهی تجمعی و ارزشیابی ایستا. در مرحله دوم، داده‌های کیفی از طریق مصاحبه‌های نیمه‌ساختاریافته با بخشی از شرکت‌کنندگان گردآوری شد تا تجربه‌ها و دیدگاه‌های آنان درباره روش‌های ارزشیابی بررسی شود.

دلیل انتخاب طرح تبیینی متوالی این بود که داده‌های کمی به تنهایی نمی‌توانستند تعامل پیچیده زبان‌آموزان با رویکردهای ارزشیابی را نشان دهند. بخش کمی پژوهش به پژوهشگران امکان داد تا تفاوت‌های معنادار میان گروه‌ها را شناسایی کرده و مشخص کنند که آیا روش‌های ارزشیابی پویا اثرات قابل اندازه‌گیری دارند یا نه. بخش کیفی نیز بینش‌هایی درباره نگرش زبان‌آموزان به بازخورد، همکاری و راهبردهای یادگیری فراهم کرد. ادغام این دو بخش به پژوهش اجازه داد تا به هر دو سؤال تحقیق پاسخ دهد: سؤال نخست به شواهد آماری درباره تفاوت‌ها نیاز داشت و سؤال دوم مستلزم آن بود که شرکت‌کنندگان توضیح دهند چگونه تجربه‌های ارزشیابی بر درک مطلب آنان تأثیر گذاشته است. این تلفیق با پیوند دادن نتایج عملکردی به تجربه‌های واقعی زبان‌آموزان، اعتبار یافته‌ها را افزایش داد.

۱-۳- شرکت‌کنندگان

این مطالعه با مشارکت ۹۰ زبان‌آموز ایرانی زبان انگلیسی انجام شد که در مؤسسه‌های خصوصی آموزش زبان در شهر بوشهر ثبت‌نام کرده بودند. انتخاب مؤسسه‌های خصوصی به عنوان بافت پژوهش از آن رو حائز اهمیت است که بسیاری از زبان‌آموزان ایرانی برای آمادگی در آزمون‌های بین‌المللی (نظیر آیلتس و تافل) یا تقویت مهارت‌های تحصیلی خود به این مراکز مراجعه می‌کنند؛ از این رو، بررسی درک مطلب خواندن استدلالی در چنین بافت‌هایی از ضرورت آموزشی بالایی برخوردار است. شرکت‌کنندگان ابتدا با استفاده از آزمون تعیین سطح مورد ارزیابی قرار گرفتند تا از همگنی آنان در سطح پایین‌تر از متوسط اطمینان حاصل شود. سپس، با به‌کارگیری روش نمونه‌گیری طبقه‌ای تصادفی^۳ بر اساس جنسیت (دو طبقه: زن و مرد)، شرکت‌کنندگان در هر طبقه به طور تصادفی در سه گروه مساوی (هر گروه ۳۰ نفر) تقسیم شدند: گروه ارزشیابی پویای گروهی مداخله‌گر (IGDA)، گروه ارزشیابی پویای گروهی تجمعی (CGDA) و گروه ارزشیابی ایستا (SA).

دامنه سنی شرکت‌کنندگان بین ۱۸ تا ۲۴ سال بود (میانگین = ۲۰/۷، انحراف معیار = ۱/۸). از میان آنان، ۴۸ نفر زن (۵۳/۳ درصد) و ۴۲ نفر مرد (۴۶/۷ درصد) بودند. بیشتر شرکت‌کنندگان از خانواده‌هایی با وضعیت اقتصادی - اجتماعی متوسط برخوردار بودند، در حالی که ۲۲ زبان‌آموز (۲۴/۴ درصد) پیشینه اقتصادی - اجتماعی پایین‌تری را گزارش کردند. همه شرکت‌کنندگان پیش از ورود به مؤسسه، دست‌کم چهار سال زبان انگلیسی را در نظام آموزش عمومی کشور آموخته بودند. هیچ‌یک از آنان سابقه زندگی یا تحصیل در کشورهای انگلیسی‌زبان نداشتند.

۳-۳- ابزارهای پژوهش

در این مطالعه از چندین ابزار کمی و کیفی استفاده شد: آزمون درک مطلب خواندن استدلالی، آزمون تعیین سطح (PET)، و مصاحبه نیمه‌ساختاریافته.

۱-۳-۳- آزمون درک مطلب خواندن استدلالی

برای سنجش درک مطلب خواندن استدلالی زبان‌آموزان، یک آزمون محقق‌ساخته شامل ۲۰ پرسش چهارگزینه‌ای طراحی گردید. متن پایه آزمون برگرفته از موضوع «بازیافت اجباری»^۴ بود که به دلیل ماهیت استدلالی (شامل ادعا، شواهد و ضد استدلال) و نیز ارتباط با تجربه روزمره زبان‌آموزان ایرانی انتخاب شد. دلیل استفاده از متن مستقل و عدم استفاده از متون کتاب درسی مؤسسه این

^۱ Mixed-methods

^۲ Sequential explanatory

^۳ Stratified random sampling

^۴ Mandatory recycling

است که متون کتاب درسی قبلاً در کلاس تدریس شده‌اند و استفاده از آن‌ها در پیش‌آزمون و پس‌آزمون می‌توانست باعث شود دانش پیشین نابرابر یا اثر حافظه بر نتایج تأثیر بگذارد. مدت زمان پاسخگویی به این آزمون ۳۰ دقیقه در نظر گرفته شد. متن از نظر شاخص‌های خوانایی^۱ (مانند شاخص سهولت خوانایی فلش^۲ و شاخص فاگ^۳) در سطحی معادل متون سطح B1 ارزیابی شد. افزون بر این، واژگان و ساختارهای دستوری متن با پیکره متون سطح B1 تطابق داده شد. برای حصول اطمینان از دشواری مناسب و جلوگیری از خستگی و ناامیدی زبان‌آموزان، متن در یک مطالعه مقدماتی بر روی ۲۰ زبان‌آموز با سطح زبانی یکسان (که در مطالعه اصلی شرکت نداشتند) اجرا گردید.

جدول ۱ شاخص‌های خوانایی متن مورد استفاده در پژوهش

شاخص خوانایی	مقدار به‌دست‌آمده	دامنه تفسیر	سطح متن
شاخص سهولت خوانایی فلش	۶۴/۵	۶۰-۷۰	B1
شاخص فاگ	۸/۲	۸-۱۰	B1
میانگین طول جمله	۱۴/۳ واژه	۱۲-۲۰ واژه	B1
میانگین هجا در هر کلمه	۱/۴۵	۱/۷-۱/۳	B1

مقادیر به‌دست‌آمده برای شاخص‌های خوانایی در دامنه استاندارد متون سطح B1 قرار دارند. شاخص سهولت خوانایی فلش (۶۴/۵) نشان‌دهنده متن نسبتاً روان، شاخص فاگ (۸/۲) بیانگر نیاز به حدود ۸ سال تحصیل رسمی، میانگین طول جمله (۱۴/۳ واژه) حاکی از ساختار نحوی مناسب، و میانگین هجا در هر کلمه (۱/۴۵) نشان‌دهنده غالب بودن واژگان کوتاه و رایج است. در مجموع، متن «بازیافت اجباری» از نظر دشواری برای زبان‌آموزان سطح پایین‌تر از متوسط (B1) مناسب ارزیابی می‌شود.

نتایج بررسی پایایی آزمون مقدماتی نشان داد که آزمون از پایایی مناسب (آلفای کرونباخ بالاتر از ۰/۸۰) برخوردار است. بنابراین، این آزمون به عنوان ابزاری معتبر برای پیش‌آزمون و پس‌آزمون مطالعه اصلی در نظر گرفته شد. برای سنجیدن روایی محتوایی، نسخه اولیه مصاحبه توسط سه متخصص آموزش زبان انگلیسی (هر سه دارای مدرک کارشناسی ارشد در رشته آموزش زبان انگلیسی، با تخصص در ارزشیابی پویا و مهارت‌های خواندن، و با بیش از ۸ سال سابقه تدریس در مؤسسه‌های زبان) بررسی شد و اصلاحاتی در جهت شفافیت و تناسب با زبان‌آموزان هدف اعمال گردید.

۳-۳-۲ آزمون تعیین سطح (PET)

برای اطمینان از همگنی شرکت‌کنندگان در سطح پایین‌تر از متوسط، از آزمون تمرینی B1 Preliminary (PET) نسخه ۲۰۲۴ استفاده گردید. این آزمون توسط [ارزشیابی زبان کمبریج](#)^۴ منتشر شده و شامل چهار بخش خواندن، نوشتن، شنیدن و صحبت کردن است. در این پژوهش، تنها بخش خواندن (شامل ۳۲ سؤال) به کار گرفته شد و مدت زمان اجرای آن ۴۵ دقیقه بود.

۳-۳-۳ مصاحبه نیمه‌ساختاریافته

برای بررسی ادراک و تجربه زبان‌آموزان از روش‌های ارزشیابی، از مصاحبه نیمه‌ساختاریافته به صورت انفرادی استفاده شد. دلیل انتخاب مصاحبه انفرادی به جای گروهی، جلوگیری از تأثیرپذیری پاسخ‌ها از نظرات دیگران و افزایش عمق و صحت داده‌ها بود. مصاحبه بر اساس مرور پژوهش‌های پیشین (پوهنر، ۲۰۰۸؛ رسایی، ۲۰۲۳) و با تمرکز بر سه حوزه طراحی شد: (الف) تجربه زبان‌آموزان از فرایند ارزشیابی، (ب) تأثیر روش‌ها بر درک متون استدلالی، و (ج) نگرش آنان نسبت به بازخورد و میانجی‌گری. برای سنجیدن روایی محتوایی، نسخه اولیه مصاحبه توسط سه متخصص آموزش زبان انگلیسی بررسی شد و اصلاحاتی در جهت شفافیت و تناسب با زبان‌آموزان هدف اعمال گردید. سپس، مصاحبه با ۵ نفر از همان گروه ۲۰ نفره مطالعه مقدماتی متن خواندن (که در مطالعه اصلی شرکت نداشتند) به صورت آزمایشی اجرا شد.

^۱ Readability indices

^۲ Flesch Reading Ease Index

^۳ Fog Index

^۴ Cambridge Assessment English

۴-۳- مراحل اجرا

پژوهش در سه مرحله پیش‌آزمون، آزمایش و پس‌آزمون انجام شد. برای اطمینان از همگنی سطح توانش زبانی شرکت‌کنندگان، آزمون PET اجرا شد. این آزمون چهار مهارت اصلی خواندن، نوشتن، شنیدن و صحبت کردن را می‌سنجد. نمره هر مهارت از ۲۵ محاسبه می‌شود و نمره کل آزمون از ۱۰۰ است. بر اساس چارچوب مشترک اروپایی مرجع برای زبان‌ها (CEFR)، سطح B1 معمولاً با نمره کل بین ۷۰ تا ۸۴ مشخص می‌شود. با این حال، در این پژوهش شرکت‌کنندگانی انتخاب شدند که نمره آنان در بازه ۶۰ تا ۷۴ قرار داشت. این بازه منطبق با بخش پایین‌تر سطح B1 است و سطح هدف مطالعه را نشان می‌دهد. هدف از این معیار، اطمینان از قرار داشتن همه شرکت‌کنندگان در سطح پایین‌تر از متوسط و کاهش تفاوت‌های احتمالی در توانش زبانی اولیه بود. سپس پیش‌آزمون درک مطلب خواندن استدلالی گرفته شد تا سطح پایه عملکرد مشخص شود. پس از این مرحله، شرکت‌کنندگان به طور تصادفی در سه گروه مساوی (هر گروه ۳۰ نفر) تقسیم شدند: ارزشیابی پویای گروهی مداخله‌گر (IGDA)، ارزشیابی پویای گروهی تجمعی (CGDA) و ارزشیابی ایستا (SA) که به عنوان گروه کنترل در نظر گرفته شد.

آزمایش به مدت شش هفته و در دوازده جلسه (هفته‌ای دو جلسه) برگزار شد. هر جلسه ۹۰ دقیقه به طول انجامید و تمام این مدت به فعالیت‌های مرتبط با پژوهش، از جمله خواندن متون استدلالی، ارائه بازخورد و ارزشیابی گروهی اختصاص داشت. در گروه IGDA، زبان‌آموزان تکالیف خواندن را به صورت فردی انجام می‌دادند. هنگامی که زبان‌آموزی با دشواری مواجه می‌شد، معلم از راهنمایی سلسله‌مراتبی استفاده می‌کرد: ابتدا یک سرخ‌ضمنی ارائه می‌داد، سپس کمک صریح‌تر می‌کرد و در نهایت اگر نیاز بود، پاسخ را مستقیماً توضیح می‌داد. این بازخورد فوری و فردی بود و سطح کمک ثبت می‌شد. در گروه CGDA، زبان‌آموزان در گروه‌های کوچک سه تا چهار نفره به صورت مشارکتی کار می‌کردند. معلم نقش تسهیل‌گر داشت. هنگامی که گروهی درخواست کمک می‌کرد، معلم پاسخ مستقیم نمی‌داد، بلکه سؤال را به کل گروه ارجاع می‌کرد، بحث را با پرسش‌های جهت‌دهنده هدایت می‌نمود و به خلاصه‌سازی نکات کلیدی کمک می‌کرد. این بازخورد در طول جلسات تکرار می‌شد تا زبان‌آموزان به تدریج درک خود را تثبیت کنند. در گروه SA، زبان‌آموزان همان تکالیف خواندن را به صورت فردی و بدون هیچ کمک یا بازخوردی از سوی معلم انجام می‌دادند. معلم تنها برگه‌های آزمون را توزیع و جمع‌آوری می‌کرد. این گروه به عنوان گروه کنترل در نظر گرفته شد و در طول آزمایش داده‌های کیفی ارائه نکرد.

پس از پایان آزمایش، همه شرکت‌کنندگان پس‌آزمون درک مطلب خواندن استدلالی (هم‌ارز با پیش‌آزمون) را تکمیل کردند. سپس، با یک زیرنمونه از گروه‌های IGDA و CGDA (در مجموع ۱۲ نفر) مصاحبه‌های نیمه‌ساختاریافته انجام شد تا ادراک و تجربه آنان از روش‌های ارزشیابی بررسی گردد.

پیش از آغاز پژوهش، از همه شرکت‌کنندگان رضایت آگاهانه کتبی دریافت شد. به آنان اطمینان داده شد که اطلاعاتشان محرمانه خواهد ماند، نمراتشان در ارزشیابی‌های کلاسی تأثیری ندارد و می‌توانند در هر مرحله از پژوهش انصراف دهند. برای حفظ ناشناس بودن، نام شرکت‌کنندگان با کدهای عددی جایگزین شد. پژوهش با تأیید کمیته اخلاق پژوهش مؤسسه انجام شده است.

۵-۳- تحلیل داده‌ها

برای تحلیل داده‌های کمی، از تحلیل واریانس آمیخته^۱ با دو عامل نوع ارزشیابی (سه سطح: مداخله‌گر، تجمعی و ایستا) به عنوان عامل بین-گروهی و زمان (دو سطح: پیش‌آزمون و پس‌آزمون) به عنوان عامل درون-گروهی استفاده شد. نمرات درک مطلب خواندن استدلالی به عنوان متغیر وابسته در نظر گرفته شد. اثرات اصلی زمان و نوع ارزشیابی و نیز اثر تعاملی آن‌ها بررسی گردید. به دلیل تمرکز پژوهش بر مقایسه مستقل هر رویکرد پویا با گروه کنترل، تحلیل‌های اصلی به صورت دوتایی انجام شد. اندازه اثر به صورت اتای جزئی^۲ گزارش گردید. پیش از تحلیل اصلی، مفروضه‌های نرمال بودن توزیع داده‌ها و همگنی واریانس‌ها بررسی شد. داده‌های کیفی حاصل از مصاحبه‌ها با استفاده از تحلیل مضمونی بررسی شدند. مضامین استخراج‌شده شامل ادراک از روش‌های

^۱ Mixed Analysis of Variance

^۲ Partial eta squared

ارزشیابی، نقش بازخورد، و چالش‌های مربوط به درک متن بود. در نهایت، یافته‌های کیفی با نتایج کمی تلفیق شد تا تبیین روشن‌تری از نحوه تأثیر روش‌های مختلف ارزشیابی بر درک مطلب خواندن ارائه شود.

۴- نتایج

ابتدا آزمون تعیین سطح (PET) روی شرکت‌کنندگان اجرا شد و زبان‌آموزانی که نمره ۶۰ تا ۷۴ (از ۱۰۰) را کسب کردند، به عنوان افراد دارای سطح پایین‌تر از متوسط برای ادامه پژوهش انتخاب شدند.

جدول ۲ آمار توصیفی نمره‌های آزمون PET

مهارت	نمره کل (از ۱۰۰)	انحراف معیار	میانگین نمره (از ۲۵)
خواندن	۶۵/۲	۱/۲	۱۶/۳
نوشتن	۶۲/۴	۱/۳	۱۵/۶
شنیدن	۶۰/۸	۱/۲	۱۵/۲
صحبت کردن	۶۳/۵	۱/۳	۱۵/۹

نتایج آزمون PET نشان داد که همه شرکت‌کنندگان در چهار مهارت زبانی از توانش مشابهی برخوردار بودند. میانگین نمره‌ها در هر مهارت بین ۱۵/۲ تا ۱۶/۳ از ۲۵ بود. میانگین کل نمرات ۶۲/۵ از ۱۰۰ به دست آمد که در بازه ۶۰ تا ۷۴ قرار می‌گیرد. این بازه با سطح پایین‌تر از متوسط همخوانی دارد و تأیید می‌کند که همه شرکت‌کنندگان در سطح مورد نظر بودند. همگنی سطح زبانی برای اعتبار یافته‌های پژوهش ضروری بود، زیرا از تأثیر تفاوت‌های فردی در توانش زبانی بر نتایج جلوگیری می‌کرد.

پیش از اجرای مطالعه اصلی، پیش‌آزمون و پس‌آزمون درک مطلب استدلالی با ۲۰ زبان‌آموز از همان جامعه (که در مطالعه اصلی شرکت نداشتند) اجرا شد. هدف از این مطالعه مقدماتی، بررسی پایایی و روایی ابزارهای سنجش بود.

جدول ۳ تحلیل پایایی پیش‌آزمون و پس‌آزمون

نوع آزمون	آلفای کرونباخ
پیش‌آزمون استدلالی	۰/۸۵
پس‌آزمون استدلالی	۰/۸۷

مقادیر آلفای کرونباخ برای هر دو آزمون پیش‌آزمون و پس‌آزمون بالاتر از ۰/۸۰ بود که نشان‌دهنده پایایی بالای آنهاست. این نتایج تأیید کردند که آزمون‌ها از همسانی درونی مناسبی برخوردارند و برای سنجش سازه‌های مورد نظر مناسب هستند.

برای پاسخ به سؤال اول پژوهش، ابتدا تأثیر هر یک از دو رویکرد ارزشیابی پویا (مداخله‌گر و تجمعی) به طور جداگانه در مقایسه با گروه کنترل (ارزشیابی ایستا) بررسی شد. سپس میزان پیشرفت از پیش‌آزمون به پس‌آزمون در این مقایسه‌ها تحلیل گردید. در ادامه، نتایج هر جدول به همراه توجیه نقش آن در پاسخ به سؤال اول ارائه می‌شود. جدول ۴ نتایج تحلیل واریانس آمیخته را برای بررسی اثرات بین‌گروهی (IGDA و SA) بر نمره‌های درک مطلب استدلالی در پیش‌آزمون و پس‌آزمون نشان می‌دهد. این جدول به بخش نخست سؤال اول پاسخ می‌دهد که آیا تفاوت کلی و بدون در نظر گرفتن عامل زمان، بین گروه IGDA و گروه کنترل (SA) وجود دارد یا خیر.

جدول ۴ آزمون اثرات بین‌گروهی IGDA و SA

منبع	میانگین مجزورات	درجه آزادی	مجموع مجزورات نوع سوم	F	سطح معناداری	اتا اسکوتر جزئی
عرض از مبدأ	۷۶۳۱/۰۲۲	۱	۷۶۳۱/۰۲۲	۳۳۷۵/۰۷۰۷	۰/۰۰۰	۰/۹۷۵
گروه	۰/۵۵	۱	۰/۵۵	۰/۱۰	۰/۷۵	۰/۰۰۲
خطا	۵/۵۲	۵۸	۳۲۰/۱۰			

نتایج جدول ۴ نشان داد که اثر اصلی گروه معنادار نیست (p بزرگتر از ۰/۰۵). به عبارت دیگر، زمانی که عامل زمان در نظر گرفته نشود، تفاوت آماری معناداری بین دو گروه IGDA و SA در نمرات درک مطلب استدلالی وجود ندارد. این یافته نشان می‌دهد که صرفاً عضو بودن در گروه آزمایش یا کنترل، بدون در نظر گرفتن تغییرات ناشی از آزمایش، نمی‌تواند مبنای قضاوت درباره اثربخشی قرار گیرد.

جدول ۵ نتایج تحلیل واریانس آمیخته را برای بررسی اثرات بین‌گروهی (SA و CGDA) بر نمره‌های درک مطلب استدلالی در پیش‌آزمون و پس‌آزمون نشان می‌دهد. این جدول نیز به بخش نخست سؤال اول برای مقایسه گروه تجمعی با گروه کنترل پاسخ می‌دهد.

جدول ۵ آزمون اثرات بین‌گروهی SA و CGDA

منبع	میانگین مجزورات	درجه آزادی	مجموع مجزورات نوع سوم	F	سطح معناداری	اتا اسکوتر جزئی
عرض از مبدأ	۱۵۲۲۰/۰۰	۱	۱۵۲۲۰/۰۰	۲۷۰۹/۰۰	۰/۰۰	۰/۹۸
گروه	۰/۳۰	۱	۰/۳۰	۰/۰۵	۰/۸۳	۰/۰۰۱
خطا	۵/۶۲	۵۸	۳۲۵/۸۰			

نتایج جدول ۵ نشان داد که اثر اصلی گروه معنادار نیست (p بزرگتر از ۰/۰۵). مشابه مقایسه قبلی، بدون در نظر گرفتن عامل زمان، تفاوت آماری معناداری بین گروه SA و CGDA وجود ندارد. این یافته مجدداً ضرورت بررسی اثر تعاملی زمان و گروه را نشان می‌دهد.

جدول ۶ به بخش دوم سؤال اول پاسخ می‌دهد که آیا میزان پیشرفت از پیش‌آزمون به پس‌آزمون در دو گروه IGDA و SA متفاوت است یا خیر.

جدول ۶ آزمون اثرات درون‌گروهی SA و IGDA

منبع	میانگین مجزورات	درجه آزادی	مجموع مجزورات نوع سوم	F	سطح معناداری	اتا اسکوتر جزئی
زمان	۵۵۰/۲۰	۱	۵۵۰/۲۰	۱۵۰/۰۰	۰/۰۰	۰/۷۲
زمان × گروه	۲۶۱/۹۰	۱	۲۶۱/۹۰	۷۱/۰۰	۰/۰۰	۰/۵۵
خطا (زمان)	۳/۶۷	۵۸	۲۱۲/۷۵			

نتایج جدول ۶ نشان داد که اثر اصلی زمان معنادار است (p کوچکتر از ۰/۰۵). یعنی نمرات درک مطلب استدلالی از پیش‌آزمون به پس‌آزمون در هر دو گروه افزایش معناداری داشته است. همچنین اثر تعاملی زمان × گروه نیز معنادار است (p کوچکتر از ۰/۰۵). این یافته به طور مستقیم به بخش دوم سؤال اول پاسخ می‌دهد و نشان می‌دهد که میزان پیشرفت در طول زمان در گروه IGDA به طور معناداری بیشتر از گروه SA بوده است. بنابراین، ارزشیابی پویای مداخله‌گر نسبت به ارزشیابی ایستا تأثیر بیشتری بر بهبود درک مطلب استدلالی داشته است.

جدول ۷ نتایج تحلیل واریانس آمیخته را برای بررسی اثرات درون‌گروهی زمان (پیش‌آزمون و پس‌آزمون) و نیز اثر تعاملی زمان و گروه (SA و CGDA) نشان می‌دهد. این جدول نیز به بخش دوم سؤال اول برای مقایسه گروه تجمعی با گروه کنترل پاسخ می‌دهد.

جدول ۷ آزمون اثرات درون‌گروهی

منبع	میانگین مجدورات	درجه آزادی	مجموع مجدورات نوع سوم	F	سطح معناداری	اتا اسکوتر جزئی
زمان	۵۹۰/۰۰	۱	۵۹۰/۰۰	۱۶۵/۰۰	۰/۰۰	۰/۷۴
زمان × گروه	۳۳۱/۰۰	۱	۳۳۱/۰۰	۹۲/۴۰	۰/۰۰	۰/۶۱
خطا (زمان)	۳/۵۸	۵۸	۲۰۷/۵۰			

نتایج جدول ۷ نشان داد که اثر اصلی زمان معنادار است (p کوچکتر از ۰/۰۵). همچنین اثر تعاملی زمان × گروه نیز معنادار است (p کوچکتر از ۰/۰۵). این یافته بیانگر آن است که گروه CGDA افزایش بیشتری در نمرات درک مطلب استدلالی نسبت به گروه SA داشته است. بنابراین، ارزشیابی پویای تجمعی نیز نسبت به ارزشیابی ایستا تأثیر معنادارتری بر بهبود درک مطلب دارد. به طور خلاصه، یافته‌های پژوهش نشان داد که هر دو روش ارزشیابی پویا (مداخله‌گر و تجمعی) در مقایسه با ارزشیابی ایستا، پیشرفت بیشتری در درک مطلب خواندن استدلالی زبان‌آموزان ایجاد کردند. اگرچه پیش از شروع مداخله تفاوت چندانی میان گروه‌ها وجود نداشت، اما در طول زمان، زبان‌آموزانی که در گروه‌های ارزشیابی پویا قرار داشتند، افزایش بیشتری در نمرات خود نشان دادند. این نتیجه حاکی از آن است که به کارگیری ارزشیابی پویا، صرف نظر از نوع مداخله‌گر یا تجمعی بودن آن، نسبت به ارزشیابی سنتی و ایستا تأثیر مثبت و معناداری بر بهبود توانایی خواندن استدلالی دارد.

برای پاسخ به سؤال دوم پژوهش، مصاحبه‌های نیمه‌ساختاریافته با ۱۲ نفر از شرکت‌کنندگان (۴ نفر از هر گروه) انجام شد. در این مصاحبه‌ها، ادراک آنان از شیوه‌های ارزشیابی، نقش بازخورد، و چالش‌های تجربه‌شده بررسی گردید. برای حفظ محرمانگی، رونوشت مصاحبه‌ها با یک نظام کدگذاری حرفی - عددی (مانند IGDA1، CGDA3، SA2) مشخص شد. تحلیل داده‌ها پنج مضمون اصلی را آشکار ساخت که نشان می‌دهد چگونه شیوه‌های ارزشیابی پویا در مقایسه با ارزشیابی ایستا، تجربه زبان‌آموزان را شکل داده است.

۴-۱- ارزش ادراک‌شده شیوه‌های ارزشیابی

شرکت‌کنندگان به طور کلی رویکردهای پویا را به آزمون‌های ایستا ترجیح دادند. زبان‌آموزان گروه IGDA بیان کردند که راهنمایی معلم در حین انجام تکلیف به آنان کمک کرد تا راهبردهای مستقل را توسعه دهند. برای نمونه، IGDA2 اظهار داشت: «یاد گرفتم چگونه درباره مسئله فکر کنم، نه اینکه فقط منتظر پاسخ بمانم.» زبان‌آموزان گروه CGDA تعامل مشارکتی را ارزشمند دانستند. CGDA3 بیان کرد: «بحث‌های گروهی نشان داد که می‌توان یک متن را به شیوه‌های مختلف دید.» در مقابل، شرکت‌کنندگان گروه SA اغلب احساس ناکامی داشتند. SA1 اظهار کرد: «آزمون‌ها فقط نشان می‌داد چه چیزهایی را نمی‌دانم، بدون اینکه کمک کنند بهتر شوم.» با این حال، همه زبان‌آموزان مداخله را ترجیح ندادند. یکی از شرکت‌کنندگان گروه IGDA اشاره کرد که «راهنمایی بیش از حد باعث کندی کارم می‌شد» که نشان می‌دهد دیدگاه نسبت به حمایت یکسان نبود.

۴-۲- ماهیت و تأثیر بازخورد

زمان‌بندی و کیفیت بازخورد، تمایز مهمی میان گروه‌ها ایجاد کرد. زبان‌آموزان IGDA بازخورد فوری را مفید توصیف کردند. IGDA4 بیان کرد: «وقتی در تحلیل یک استدلال مشکل داشتم، پرسش‌های معلم نشان داد کجای استدلالم اشتباه است.» در گروه CGDA، بازخورد به صورت تجمعی در طول جلسات شکل گرفت. CGDA1 اظهار داشت: «هر بحث بر بحث قبلی بنا می‌شد، بنابراین درکم به تدریج عمیق‌تر می‌شد.» در مقابل، زبان‌آموزان SA بازخورد محدودی دریافت کردند. SA3 بیان کرد: «فقط نمره می‌گرفتیم، بدون توضیح.» با این حال، برخی از زبان‌آموزان CGDA نیز به چالش‌هایی اشاره کردند. یکی از آنان گفت: «گاهی بازخورد گروهی گیج‌کننده بود، چون نظرهای متفاوت بود.»

۴-۳- تجربه‌های شناختی و هیجانی

ارزشیابی پویا شیوه مواجهه زبان‌آموزان با متون دشوار را تغییر داد. زبان‌آموزان IGDA راهبردهای آنی را توسعه دادند IGDA3. بیان کرد: «یاد گرفتم جمله‌های پیچیده را تجزیه کنم، به جای اینکه مضطرب شوم». شرکت‌کنندگان CGDA از حمایت همسالان اعتماد به نفس بیشتری کسب کردند. CGDA2 گفت: «شنیدن تفسیر دیگران باعث شد بفهمم راه‌های مختلفی برای درک وجود دارد». در مقابل، آزمون‌های ایستا موجب اضطراب می‌شد. SA2 اظهار داشت: «آزمون‌ها مرا مضطرب می‌کرد، چون شبیه قضاوت بودند». با این حال، یکی از شرکت‌کنندگان SA اشاره کرد که «کار کردن به صورت فردی باعث شد بیشتر تلاش کنم»، که نشان می‌دهد همه جنبه‌های ارزشیابی ایستا منفی تلقی نشد.

۴-۴- توسعه راهبردهای یادگیری

ارزشیابی پویا آگاهی فراشناختی را تقویت کرد. زبان‌آموزان IGDA گزارش دادند که شیوه پرسشگری معلم را درونی کرده‌اند. IGDA1 توضیح داد: «اکنون همان پرسش‌هایی را که معلم می‌پرسید از خودم می‌پرسم». زبان‌آموزان CGDA نیز راهبردهایی را از همسالان آموختند. CGDA4 بیان کرد: «با گوش دادن به اعضای گروه یاد گرفتم چگونه استدلال‌های ضعیف را تشخیص دهم». در مقابل، زبان‌آموزان SA رشد اندکی در راهبردها گزارش کردند. SA4 اظهار داشت: «پیش از این مطالعه هرگز به شیوه خواندنم فکر نکرده بودم». با این حال، یکی از آنان به نوعی خودتأکیدی اشاره کرد و گفت: «انجام تکلیف بدون کمک مرا مجبور کرد پاسخ‌هایم را با دقت بررسی کنم».

۴-۵- ابعاد اجتماعی یادگیری

بعد اجتماعی در CGDA نقش برجسته‌ای داشت. شرکت‌کنندگان بیان کردند که همکاری موجب بهبود درک شد CGDA2. اظهار داشت: «توضیح دادن ایده‌هایم برای دیگران کمک کرد تفکرم روشن‌تر شود». حتی زبان‌آموزانی که در ابتدا مردد بودند، به تدریج اعتماد بیشتری پیدا کردند CGDA4. بیان کرد: «در ابتدا خجالتی بودم، اما بحث‌های ساختاریافته صحبت کردن را آسان‌تر کرد». این حمایت اجتماعی در IGDA و SA وجود نداشت. با این حال، برخی از شرکت‌کنندگان CGDA تعامل همسالان را گاه حواس‌پرت‌کن توصیف کردند CGDA3. اظهار داشت: «گاهی گروه از موضوع خارج می‌شد و تمرکز دشوار بود».

۴-۶- جمع‌بندی مضمون‌ها

مصاحبه‌ها نشان داد که ارزشیابی پویا عموماً حمایت‌کننده‌تر از آزمون‌های ایستا ادراک شد، هرچند تجربه‌ها یکسان نبود. مداخله فوری در IGDA و مداخله تجمعی مبتنی بر همسالان در CGDA به تقویت راهبردها و درگیری هیجانی کمک کرد. در مقابل، آزمون‌های ایستا راهنمایی محدودی ارائه دادند و اغلب موجب ناکامی شدند. با این حال، حتی در درون هر گروه نیز نگرانی‌هایی درباره سرعت اجرا، وضوح، یا حواس‌پرتی مطرح شد. این دیدگاه‌ها نشان می‌دهد که اثربخشی شیوه‌های ارزشیابی نه تنها به ساختار رسمی، بلکه به نحوه واکنش زبان‌آموزان به بافت‌های شناختی و اجتماعی ایجادشده در جریان ارزشیابی وابسته است.

۵- بحث

یافته‌های این پژوهش بر اساس نظریه اجتماعی - فرهنگی [ویگوتسکی \(۱۹۷۸\)](#) تبیین شد. در این نظریه، یادگیری به عنوان فرایندی اجتماعی و میانجی‌گری‌شده در چارچوب منطقه تقریبی رشد در نظر گرفته می‌شود. در این چارچوب، ارزشیابی فعالیتی جدا از آموزش نیست، بلکه ابزاری است که در صورت ادغام میانجی‌گری می‌تواند یادگیری را تقویت کند. پیشرفت معنادار گروه‌های IGDA و CGDA در درک مطلب استدلالی نشان داد که تعامل هدایت‌شده و حمایت داربست‌گونه نسبت به سنجش ایستا اثربخش‌تر است. این نتایج با مباحث بخش پیشینه پژوهشی هماهنگ است که بر نقش میانجی‌گری در رشد یادگیری تأکید داشت. یافته‌های کمی نشان داد که هر دو گروه ارزشیابی پویا عملکرد بهتری نسبت به گروه ارزشیابی ایستا داشتند. اثر قوی‌تر ارزشیابی پویای گروهی مداخله‌گر در کوتاه‌مدت را می‌توان با تأکید این رویکرد بر میانجی‌گری بی‌درنگ و فردی‌شده تبیین کرد. پرسش‌های راهنمای معلم، زبان‌آموزان را به تحلیل استدلال‌ها، بازشناسی شواهد و ارزیابی ساختارهای بلاغی ترغیب می‌کرد و همین امر راهبردهای فراشناختی آنان را پرورش داد. این نتایج با پژوهش‌هایی همسو است که بر اهمیت هویت و آموزش حرفه‌ای معلم در اجرای اثربخش ارزشیابی پویا تأکید می‌کنند ([عسگری، ۲۰۲۴](#)). افزون بر این، فرایند داربست‌سازی معلم که در کانون این

میانجی‌گری قرار دارد، نشان داده است که زبان‌آموزان را به خوبی در درون منطقه تقریبی رشد هدایت می‌کند و به درگیری عمیق‌تر آنان با مواد خواندن می‌انجامد ([پسندی و همکاران، ۲۰۲۳](#)). همچنین این یافته با [نظر آلبرشت و پیرسون \(۲۰۲۰\)](#) هماهنگ است که بر اهمیت آموزش راهبردی برای خواندن استدلالی تأکید ورزیدند. این نتیجه تأییدی بر پژوهش‌های پیشین [آلیوا \(۲۰۰۸\)](#) و [پوهنر \(۲۰۰۹\)](#) است که مزایای میانجی‌گری فردی‌شده را در تکالیف درک مطلب مستند ساختند. پژوهش حاضر این خط پژوهشی را با تمرکز ویژه بر متون استدلالی که نیازمند ارزیابی ادعاها و ادعاهای متقابل هستند، گسترش داد.

گروه CGDA نیز بهبود معناداری نشان داد. قالب مشارکتی فرصت‌هایی برای الگوگیری از همسالان و بحث مکرر درباره متن فراهم کرد. زبان‌آموزان بیان کردند که تعامل گروهی به آنان کمک کرد الگوهای سازمانی مانند مسئله-راه‌حل و علت-معلول را بهتر درک کنند. این نتایج با یافته‌های [رضایی و همکاران \(۲۰۲۳\)](#) و [میرزایی و همکاران \(۲۰۱۷\)](#) همخوان است که نشان دادند میانجی‌گری جمعی آگاهی فراشناختی و یادگیری یاری‌گرفته از همسالان را تقویت می‌کند. پژوهش حاضر شواهد جدیدی ارائه می‌دهد که میانجی‌گری جمعی می‌تواند درک مطلب استدلالی را نیز تقویت کند؛ نوعی متن که در پژوهش‌های پیشین ارزشیابی پویا کمتر بررسی شده است. همچنین یافته مربوط به اثربخشی بیشتر IGDA در یک برنامه شش هفته‌ای به مطالعات قبلی می‌افزاید. [میری و همکاران \(۲۰۱۷\)](#) و [صفایی‌اصل و همکاران \(۲۰۲۱\)](#) مزایای بلندمدت قوی‌تری برای CGDA گزارش کردند که نشان می‌دهد طول آزمایش در شکل‌گیری نتایج نقش دارد.

این نتیجه با پیشینه گسترده‌تر ارزشیابی پویا همسو است که بر ارزش داربست‌سازی همکلاسی در ارتقای یادگیری تأکید می‌کند. پژوهش‌ها نشان داده‌اند که داربست‌سازی همکلاسی می‌تواند با میانجی‌گری مبتنی بر معلم در تقویت دستاوردهای درک خواندن رقابت کند، زیرا همکلاسی‌ها دیدگاه‌های مکملی ارائه می‌دهند و راهبردهای یادگیری را به اشتراک می‌گذارند ([پسندی و همکاران، ۲۰۲۳](#)). این بعد مشارکتی، نقطه قوت متمایز الگوی ارزشیابی پویای گروهی جمعی به‌شمار می‌رود.

عملکرد ضعیف‌تر گروه SA محدودیت‌های ارزشیابی ایستا را تأیید می‌کند. نمره‌های پایین‌تر و ادراک‌های منفی آنان با نقدهای [پوپهام \(۲۰۰۱\)](#) و [اسمیت و همکاران \(۲۰۲۱\)](#) درباره آزمون‌های سنتی همسو است. گزارش زبان‌آموزان مبنی بر افزایش اضطراب و نبود راهنمایی در آزمون‌های ایستا نشان می‌دهد که چنین ارزشیابی‌هایی در ارتقای یادگیری ناتوان هستند. این یافته با بحث بخش پیشینه پژوهشی مرتبط است که جدایی ارزشیابی از آموزش را به عنوان یک مسئله پایدار در بافت‌های EFL مطرح کرد.

یافته‌های کیفی سازوکارهای زیربنایی نتایج کمی را روشن ساخت. زبان‌آموزان IGDA بیان کردند که میانجی‌گری معلم آنان را به تأمل و بازاندیشی سوق داد. این فرایند با مفهوم درونی‌سازی که [ورتش \(۲۰۰۸\)](#) توصیف کرده است همخوانی دارد. زبان‌آموزان CGDA گزارش کردند که همکاری اعتمادبه‌نفس آنان را افزایش داد که از ادعای [شاناهان و شاناهان \(۲۰۲۰\)](#) حمایت می‌کند مبنی بر اینکه متون پیچیده نیازمند راهبردهای فردی و جمعی هستند. در مقابل، زبان‌آموزان SA از ناکامی سخن گفتند که دیدگاه [استیگینز \(۲۰۰۵\)](#) را درباره تأثیر قالب ارزشیابی بر پیامدهای عاطفی تأیید می‌کند. ارتباط میان رشد راهبردها، کاهش اضطراب و پیشرفت درک مطلب در گروه‌های DA نشان می‌دهد که میانجی‌گری ابعاد شناختی و هیجانی یادگیری را به طور همزمان در بر می‌گیرد.

این یافته‌ها اجماع نوظهوری را در پژوهش‌های حوزه ارزشیابی تقویت می‌کنند: آزمون‌های ایستا و فاقد میانجی‌گری برای تسخیر یا پرورش دامنه کامل توان بالقوه رشدی زبان‌آموزان ناکافی هستند، به‌ویژه هنگامی که با مزایای تشخیصی و آموزشی ارائه‌شده از سوی ارزشیابی پویا مقایسه می‌شوند ([فرخی‌پور و همکاران، ۲۰۲۰](#)؛ [عسگری، ۲۰۲۴](#)؛ [پسندی و همکاران، ۲۰۲۳](#)).

یافته‌ها همچنین به عوامل فرهنگی در بافت ایرانی اشاره دارند. شرکت‌کنندگان CGDA بیان کردند که کار گروهی با شیوه‌های آموزشی سنتی شباهت داشت. این مشاهده با دیدگاه [رحیم‌نیا \(۲۰۱۶\)](#) همسو است. این هم‌راستایی فرهنگی ممکن است مقاومت در برابر فعالیت‌های مشارکتی را کاهش داده و مشارکت را افزایش داده باشد. [گریب و ژانگ \(۲۰۲۲\)](#) استدلال کردند که درگیری

عاطفی در رشد خواندن زبان دوم نقش اساسی دارد. یافته‌های حاضر این دیدگاه را تقویت می‌کند و نشان می‌دهد که الگوهای تعاملی آشنا از نظر فرهنگی می‌توانند انگیزش را افزایش دهند.

در مجموع، این پژوهش ادعاهای نظری پیشین درباره ارزشیابی پویا را تأیید کرد، مطالعات تجربی اخیر را با تمرکز بر درک مطلب استدلالی گسترش داد، و قابلیت انطباق DA را با بافت‌های فرهنگی مختلف نشان داد. همچنین ادغام یافته‌های کمی و کیفی نشان داد که ارزشیابی پویا نه تنها ابزار سنجش، بلکه رویکردی آموزشی است که فرایند یادگیری را شکل می‌دهد.

۶- نتیجه‌گیری

این پژوهش اثربخشی رویکردهای ارزشیابی پویا را در مقایسه با آزمون‌های ایستای سنتی برای بهبود درک مطلب خواندن در میان زبان‌آموزان ایرانی بررسی کرد. هر دو روش ارزشیابی پویای گروهی (مداخله‌گر و تجمعی) نتایج قوی‌تری نسبت به ارزشیابی ایستا نشان دادند. قالب مداخله‌گر به‌ویژه در حمایت از تحلیل متون استدلالی مؤثر بود، در حالی که قالب تجمعی در درک مشارکتی ساختارهای استدلالی مزیت داشت.

یافته‌ها از نظریه اجتماعی - فرهنگی [ویگوتسکی \(۱۹۷۸\)](#) حمایت می‌کنند که بیان می‌دارد میانجی‌گری هدایت‌شده در چارچوب منطقه تقریبی رشد موجب تحول می‌شود. زبان‌آموزان در گروه‌های ارزشیابی پویا نمره‌های بالاتری کسب کردند و سطح بیشتری از اعتمادبه‌نفس و درگیری در یادگیری را گزارش دادند. این نتایج با پژوهش‌های [پوهنر و لانتولف \(۲۰۱۰\)](#) و مطالعات جدیدتر مانند [ژانگ و لیو \(۲۰۲۳\)](#) و [چنگ و همکاران \(۲۰۲۳\)](#) همسو است که نشان دادند یادگیری میانجی‌گری‌شده، درک متون پیچیده را بهبود می‌بخشد. همچنین بافت آموزشی ایران ممکن است در پذیرش مثبت رویکردهای مشارکتی نقش داشته باشد، همان‌گونه که [رحیم‌نیا \(۲۰۱۶\)](#) اشاره کرده است.

نتایج این پژوهش دلالت‌های مهمی برای عمل آموزشی دارد. ارزشیابی می‌تواند با آموزش ادغام شود تا از کاربرد راهبردها و انگیزش یادگیری حمایت کند و صرفاً نقش ابزار سنجش را ایفا نکند. معلمان به آموزش تخصصی برای اجرای مؤثر روش‌های ارزشیابی پویا نیاز دارند و نهادهای آموزشی باید زمان و منابع کافی برای پشتیبانی از این رویکردها اختصاص دهند. چنین تغییراتی ارزشیابی را به فرایند یادگیری نزدیک‌تر می‌کند؛ پیشنهادی که با دیدگاه [استیگنیز \(۲۰۰۵\)](#) و [ویلیام \(۲۰۱۱\)](#) همسو است.

با وجود این دستاوردها، محدودیت‌هایی نیز باید در نظر گرفته شود. شرکت‌کنندگان پژوهش زبان‌آموزان ایرانی در سطح پایین‌تر از متوسط بودند و این امر تعمیم‌پذیری نتایج را محدود می‌کند. مدت آزمایش شش هفته بود که ممکن است برای آشکار شدن ظرفیت کامل میانجی‌گری تجمعی کافی نبوده باشد. تکالیف خواندن نیز به متون استدلالی محدود شد و دامنه کاربرد را کاهش داد. این محدودیت‌ها در پژوهش‌های مبتنی بر کلاس درس رایج هستند، اما مسیرهایی را برای پژوهش‌های آینده فراهم می‌کنند. پیشنهاد می‌شود مطالعات آینده سطوح مهارتی متفاوت و بافت‌های متنوعی را در بر گیرند، مدت آزمایش را افزایش دهند، و ارزشیابی پویا را در گونه‌های مختلف متنی به کار گیرند. پژوهش‌های اخیر [پوهنر و یو \(۲۰۲۲\)](#) و [رسایی \(۲۰۲۳\)](#) نشان داده‌اند که ابزارهای میانجی‌گری دیجیتال می‌توانند فرصت‌های اجرای ارزشیابی پویا را در محیط‌های آموزشی گسترش دهند.

این پژوهش به گسترش پیشینه مربوط به ارزشیابی جایگزین در آموزش زبان کمک کرد. نتایج نشان داد که ارزشیابی پویا می‌تواند هم عملکرد و هم نگرش نسبت به یادگیری را تقویت کند و مزایای مشخصی برای قالب‌های مداخله‌گر و تجمعی در درک مطلب استدلالی فراهم آورد. این یافته‌ها مباحث ارائه‌شده در بخش مرور پیشینه را گسترش می‌دهد و جهت‌گیری‌هایی برای بهبود عمل آموزشی و پژوهش‌های آینده در حوزه ارزشیابی ارائه می‌کند.

References

- Ableeva, R. (2008). The effects of dynamic assessment on L2 listening comprehension. *Language and Sociocultural Theory*, 1(1), 53–82. <https://doi.org/10.1558/lst.v1i1.53>
- Ableeva, R., & Lantolf, J. P. (2011). Mediated dialogue and the microgenesis of second language listening comprehension. *Assessment in Education: Principles, Policy & Practice*, 18(2), 133–

149.
<https://doi.org/10.1080/0969594X.2011.565107>
- Albrecht, J., & Pearson, P. D. (2020). Effective strategies for comprehending argumentative texts. *Journal of Educational Psychology*, 112(3), 589–601. <https://doi.org/10.1037/edu0000387>
- Asgari, L. (2024). The impact of teacher professional identity, knowledge and education on dynamic assessment processes in English language learning. *Journal of Foreign Language Research*. Advance online publication. 10.22059/jflr.2024.373917.1113
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148. <https://doi.org/10.1177/003172171009200119>
- Cambridge Assessment English. (2024). *B1 Preliminary (PET) practice tests*. Cambridge University Press.
- Cheng, L., Wu, S., Chen, M. C., & Lin, C. Y. (2023). Exploring the impact of argumentative text features on EFL learners' reading comprehension. *International Journal of English Language Teaching*, 11(12), 251–264. <https://doi.org/10.5539/ijel.v11n12p251>
- Fani, T., & Rashtchi, M. (2015). Dynamic assessment of reading comprehension ability: Group or individualized. *Education Journal*, 4(6), 325–331.
- Farahani, F., & Moghadam, F. S. A. (2020). The impact of the cumulative group dynamic assessment on the learning of congruent and non-congruent collocations among Iranian intermediate EFL learners. *Journal of Applied Linguistics and Language Research*, 7(4), 21–36.
- Farokhipour, S., Khoshsima, H., Sarani, A. and Ganji, M. (2020). Presenting and investigating the effect of a local model of dynamic assessment in diagnosing and removing learning difficulties of high school students in productive skills. *Journal of Foreign Language Research*, 10(1), 120-134. 10.22059/jflr.2019.273144.593
- Ghafour, Z., & Rahimi, M. (2021). The lived experiences of Iranian EFL learners: A narrative inquiry. *Journal of English Language Teaching and Research*, 12(1), 121–133. <https://doi.org/10.5539/ijel.v12n1p121>
- Grabe, W., & Zhang, C. (2022). Teaching reading to EFL learners: Text structures, strategies, and comprehension. *TESOL Quarterly*, 56(3), 487–512. <https://doi.org/10.1002/tesq.3077>
- Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. Cambridge University Press.
- Kintsch, W., & Rawson, K. A. (2021). Comprehension processes in reading. *Annual Review of Psychology*, 72, 405–432. <https://doi.org/10.1146/annurev-psych-010419-050901>
- Kramsch, C. (2015). *The applied linguist: A journey through theory and practice*. Routledge. <https://doi.org/10.4324/9781315746216>

- Kuhn, D. (2017). *The skills of argument*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511571350>
- Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, 28(1), 67–109. <https://doi.org/10.1017/S0272263106060037>
- Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, 1(1), 49–74. <https://doi.org/10.1558/japl.1.1.49.55872>
- Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge.
<https://doi.org/10.4324/9780203813850>
- Lidz, C. S., & Elliott, J. G. (2000). *Dynamic assessment: Prevailing models and applications*. Elsevier Science.
- Miri, M., Alibakhshi, G., Kushki, A., & Bavarsad, P. S. (2017). Going beyond one-to-one mediation in zone of proximal development (ZPD): Concurrent and cumulative group dynamic assessment. *Eurasian Journal of Applied Linguistics*, 3(1), 1–24.
- Mirzaei, A., Shakibaei, L., & Jafarpour, A. (2017). ZPD- based dynamic assessment and collaborative L2 vocabulary learning. *The Journal of Asia TEFL*, 14(1), 114–129.
- Mohammadi, M., & Rezaei, P. (2025). AI-based scaffolding and dynamic assessment in Iranian EFL classrooms. *Journal of English for Academic Purposes*, 68, 101334.
<https://doi.org/10.1016/j.jeap.2025.101334>
- Pasandi, S., Bagherzadeh Kasmani, M. & Ahangari, S. (2023). Microgenetic Analysis in an EFL Context: The Effects of Teacher and Peer Scaffolding on Reading Comprehension Pedagogy. *Journal of Foreign Language Research*, 13(1), 59-78. doi: 10.22059/jflr.2022.338520.935
- Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Springer.
- Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, 43(3), 471–491. <https://doi.org/10.1002/j.1545-7249.2009.tb00245.x>
- Poehner, M. E., & Lantolf, J. P. (2010). *Dynamic assessment in second language acquisition: Constructing a sociocultural approach to assessment*. John Benjamins.
<https://doi.org/10.1075/llt.25>
- Poehner, M. E., & Yu, L. (2022). Dynamic assessment of L2 writing: Exploring the potential of rubrics as mediation in diagnosing learner emerging abilities. *TESOL Quarterly*, 56(4), 1191–1217.
<https://doi.org/10.1002/tesq.3098>
- Popham, W. J. (2001). *The truth about testing: An educator's call to action*. Association for Supervision and Curriculum Development.

- Rahimnia, R. (2016). A critical analysis of the English language teaching curriculum in Iran: A sociocultural perspective. *International Journal of Applied Linguistics & English Literature*, 5(10), 241–248. <https://doi.org/10.7575/aiac.ijalel.v.5n.10p.241>
- Rassaei, E. (2023). Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners. *Computer Assisted Language Learning*, 36(3), 257–287. <https://doi.org/10.1080/09588221.2021.1924901>
- Rezai, M., Jafari, Z., & Rostami, Z. (2023). The effects of group dynamic assessment and process-based instruction on EFL learners' metacognitive awareness and listening comprehension. *Language Testing in Asia*, 14(2), 1–16. <https://doi.org/10.1186/s40468-023-00229-1>
- Safaei Asl, N., Ahangari, S., & Saeedi, M. (2021). The effect of concurrent and cumulative group dynamic assessment on homogeneous and heterogeneous EFL learners' auditory memory in listening tasks. *Journal of English Language Pedagogy and Practice*, 14(28), 67–86.
- Sanaeifar, S. H., & Divcolaii, M. S. (2019). Actualizing language learners' potentials through mediation: Cumulative vs concurrent group dynamic assessment and students' self-management of learning tasks. *Theory and Practice in Language Studies*, 9(7), 802–809. <https://doi.org/10.17507/tpls.0907.08>
- Shanahan, T., & Shanahan, C. (2020). Disciplinary literacy: A focus on comprehension of complex texts. *Journal of Educational Psychology*, 112(6), 1007–1021. <https://doi.org/10.1037/edu0000422>
- Smith, J., Johnson, K., & Williams, L. (2021). Limitations of static assessments in capturing dynamic reading comprehension. *Educational Assessment*, 26(2), 335–350. <https://doi.org/10.1080/10627197.2021.1888692>
- Stiggins, R. J. (2005). *Student-involved assessment for learning* (4th ed.). Prentice Hall.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, H., & Xu, Z. (2024). Technology-enhanced mediation in dynamic assessment of argumentative reading. *ReCALL*, 36(2), 245–262. <https://doi.org/10.1017/S0958344024000123>
- Wertsch, J. V. (2008). *Mind as action*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195117530.001.0001>
- Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.
- Zhang, J., & Liu, Y. (2023). Investigating the role of dynamic assessment in enhancing EFL learners' reading comprehension. *Journal of Language and Literacy Education*, 19(2), 68–85. <https://doi.org/10.1016/j.linged.2023.100123>