

## تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان در فعالیت‌های انتخابی و

### تولیدی درک مطلب خواندن

#### کبری توسلی\*

استادیار، گروه آموزش زبان انگلیسی، دانشگاه آزاد اسلامی، واحد کرج،  
کرج، ایران

#### فاطمه نیک‌مرد\*\*

دانشجوی دکتری، گروه آموزش زبان انگلیسی، دانشگاه آزاد اسلامی، واحد کرج،  
کرج، ایران

(تاریخ دریافت: ۹۷/۰۲/۱۶، تاریخ تصویب: ۹۸/۰۱/۲۱، تاریخ چاپ: تیر ۱۳۹۸)

#### چکیده

هدف این تحقیق، بررسی تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان در دو نوع فعالیت انتخابی و تولیدی درک مطلب خواندن بود. ۶۰ زبان‌آموز از میان ۷۸ نفر براساس عملکردشان در آزمون نلسون سطح مبتدی، انتخاب و به دو گروه ۳۰ نفره کنترل و تجربی تقسیم شدند. در هر دو گروه، از بخش درک مطلب خواندن آزمون KET که شامل ۲۸ سوال انتخابی و ۲۷ سوال تولیدی می‌شد، به عنوان پیش‌آزمون و پس‌آزمون استفاده شد. زبان‌آموزان گروه تجربی از ارزشیابی پویا مدل "آزمون-میانجی-آزمون مجدد" بهره بردند. در پایان، برای بررسی عملکرد دو گروه و پاسخ به سوالات، از تحلیل واریانس دوسویه سنجش مکرر و تحلیل واریانس چند متغیری استفاده شد. مهم‌ترین یافته این تحقیق این است که ارزشیابی پویا تأثیر قابل توجهی در بهبود عملکرد زبان‌آموزان در هر دو نوع فعالیت انتخابی و تولیدی درک مطلب خواندن دارد. نتایج بیشتر و جنبه‌های کاربردی آن توضیح داده می‌شود.

**واژه‌های کلیدی:** ارزشیابی، ارزشیابی پویا، درک مطلب خواندن، زبان‌آموز، فعالیت‌های انتخابی، فعالیت‌های تولیدی.

\* نویسنده مسئول: [kobra.tavassoli@kiaiu.ac.ir](mailto:kobra.tavassoli@kiaiu.ac.ir)

\*\* [E-mail: fatemehnikmard@gmail.com](mailto:fatemehnikmard@gmail.com)

۱- مقدمه

یکی از مهمترین مهارت‌هایی که هر زبان‌آموز باید بتواند به‌خوبی از عهده آن برآید، درک مطلب خواندن است. شوری و مختاری<sup>۱</sup> (۲۰۰۱) در تعریف درک مطلب خواندن گفته‌اند که درک مطلب خواندن هم کنشی میان متن، خواننده، و شرایطی است که خواندن در آن اتفاق می‌افتد. از طرف دیگر، کرن<sup>۲</sup> (۱۹۸۹) و هورنر<sup>۳</sup> (۱۹۹۸) معتقدند که درک مطلب خواندن تلاشی دشوار در همه زبان‌هاست، چراکه خواننده باید قادر به هماهنگی چندین فرایند از جمله توجه به متن، نگه‌داشتن متن در حافظه، فرایندهای بینشی و ادراکی، و غیره باشد. این فرایندها و نیز فرایندهای دیگر باید در کنار هم کار کنند تا بتوان معنی درست را درک کرد و انجام این مهم در زبان دوم کاری سخت‌تر است. در نتیجه، چگونگی آموزش و ارزشیابی این مهارت، وظیفه‌ای مهم برای معلمان شمرده می‌شود.

درباره درک مطلب خواندن، الدرسن، هاپاکانگاس، هوها، نیمینن، و اولاکونوجا<sup>۴</sup> (۲۰۱۵) گفته‌اند که بیشتر خوانندگان زبان دوم یا زبان خارجی، متون را به‌زبانی می‌خوانند که برایشان بطور کامل قابل فهم نیست و غالب مشکلاتشان مربوط به‌زبان است نه مربوط به‌درک مطلب خواندن. در نتیجه، برای پیش‌بینی مشکلات آن‌ها، معلم یا ممتحن بایستی مشکلات زبانی خوانندگان را بیابد. همچنین، ارزشیابی مهارت درک مطلب خواندن به‌همان اندازه آموزش آن حائز اهمیت است. یکی از انواع ارزشیابی‌هایی که در حال حاضر مرکز توجه قرار گرفته و با آموزش عجین شده است، ارزشیابی پویاست. ارزشیابی پویا، به‌عنوان مدل جدیدی از ارزشیابی، فرایند مداخله فعالانه معلم در یادگیری دانش‌آموزان و سنتجش دائم پاسخ‌های آن‌ها به‌این مداخلات است. در این مدل، رابطه‌ای مداوم بین معلم و دانش‌آموز وجود دارد که در آن معلم همیشه در حال ارائه راهکار، سوالات هدایت‌کننده، و دادن پاسخ‌های مناسب است (هی وود و لیدز،<sup>۵</sup> ۲۰۰۷).

به‌گفته پوهنر<sup>۶</sup> (۲۰۰۸)، ارزشیابی پویا که توسط وایگوتسکی، روانشناس به‌نام روسی معرفی شد، نوعی ارزشیابی فعل و انفعالی است که بیشتر در زمینه تعلیم و تربیت مورد استفاده

1- Sheorey & Mokhtari

2- Kern

3- Horner

4- Alderson, Haapakangas, Huhta, Nieminen, & Ullakonoja

5- Haywood & Lidz

6- Poehner

قرار می‌گیرد. تمرکز وایگوتسکی<sup>۱</sup> بر روی فعل و انفعالاتی بود که به تشریک مساعی و ارتباطات میان افراد بستگی دارند. این نوع نگرش به ارزشیابی در مطالعات مربوط به زبان دوم در ابتدا توسط پوهنر و لتالف<sup>۲</sup> (۲۰۰۵) مطرح شد. در ارزشیابی پویا، معلم همیشه در هم‌کنشی با شاگردان است و زبان‌آموزان را به سوی پیشرفت راهنمایی می‌کند (هاردینگ، الدرسن، و بران فات<sup>۳</sup>، ۲۰۱۵). معلم در جریان عملکرد زبان‌آموز، در موقع لزوم سخنان وی را قطع کرده، و برای عملکرد بهتر وی، سوالاتی می‌پرسد و نکاتی را گوشزد می‌کند. این فرایند "میانجی‌گری" نامیده می‌شود. هر چند ارزشیابی پویا، پتانسیل زیادی برای استفاده در تمرین‌های کلاسی دارد، اما تاکنون معلمان زیادی از آن در کلاس‌هایشان استفاده نکرده‌اند (وین، دابر، و اورس<sup>۴</sup>، ۲۰۱۶).

در این راستا، هدف تحقیق حاضر بررسی تأثیر ارزشیابی پویا بر عملکرد درک مطلب خواندن زبان‌آموزان در دو نوع فعالیت مهم انتخابی و تولیدی است. بدین منظور، در بخش بعدی، پیشینه تحقیق در زمینه درک مطلب خواندن، اهمیت ارزشیابی این مهارت، انواع فعالیت‌های درک مطلب خواندن از جمله انتخابی و تولیدی، و ارزشیابی پویا آمده است. پس از آن، روش تحقیق، تجزیه تحلیل داده‌ها، و نتایج بدست آمده به تفصیل توضیح داده شده است.

## ۲- پیشینه تحقیق

### ۲-۱- درک مطلب خواندن

مهارت درک مطلب خواندن، فرایند تشخیص و درک معنی لغات موجود در متن، پیدا کردن رابطه دستوری بین آن‌ها و دریافت یک معنی منسجم از کل متن است (اوسو-خوان و مارتینز-فلورز<sup>۵</sup>، ۲۰۰۶). کرن (۱۹۸۹) و هورنر (۱۹۹۸) معتقدند که درک مطلب خواندن، کار دشواری در همه زبان‌های دنیاست. خواننده بایست قادر به هماهنگی چندین کار از جمله توجه به مطلب ارائه شده، حفظ آن در حافظه، تجزیه تحلیل متن در ذهن، و درک خود متن باشد.

1- Vygotsky

2- Poehner & Lantolf

3- Harding, Alderson, & Brunfaut

4- Veen, Dobber, & Oers

5- Uso'-Juan & Martinez-Flor

بنابراین می‌توان گفت که خواندن به‌زبانی غیر از زبان مادری که افزون بر موارد فوق نیازمند دانستن لغات، دستور زبان و نیز سایر موارد زبانی است که کاری به‌مراتب سخت‌تر خواهد بود. در این راستا صفری و محقق‌متطری (۲۰۱۷) به‌دنبال انجام تحقیقی درباره‌ی تأثیر انواع ساده‌سازی‌های متنی بر روی عملکرد زبان‌آموزان در فعالیت‌های درک مطلب خواندن به‌این نتیجه رسیدند که ساده‌سازی‌های متنی، واژگانی، و دستور زبانی کمک شایانی به‌فهم زبان‌آموزان می‌کند.

به اعتقاد گرب و استولر<sup>۱</sup> (۲۰۰۱)، درک مطلب خواندن مهم‌ترین مهارت زبانی فرهنگی است چراکه امروزه عمده نیاز دانش‌آموزان و دانشجویان فهم مطالبی است که می‌خوانند. درک مطلب خواندن در حقیقت یک فرایند معنی‌سازی است که ارتباط دائم بین خواننده و متن را برای درک معنی متن می‌طلبد (گودمن<sup>۲</sup>، ۱۹۸۸). از فعالیت‌های ذهنی مورد استفاده خواننده برای درک بهتر متن، معمولاً با نام راهبردی درک مطلب خواندن یاد می‌شود (پاریس، واسیک، و ترنر<sup>۳</sup>، ۱۹۹۱). همچنین، نتایج پژوهش‌های نوین نشان داده است که انجام فعالیت‌های درک مطلب خواندن که خود زبان‌آموز در آن‌ها حضور فعال دارد و مهم‌ترین عضو حاضر در انجام آن فعالیت است، کمک قابل توجهی به‌بهبود عملکرد وی می‌کند (معماری هنجانی و لی<sup>۴</sup>، ۲۰۱۷).

هر متنی، ترکیبی از کلمات، جملات و پاراگراف‌هاست. برای فهمیدن یک متن، خواننده باید ابتدا بتواند کلمات آن را تشخیص دهد. یعنی می‌توان گفت که دانش واژگان امری مهم در فهمیدن یک متن است. با این وجود، نکته مهم این است که مهارت‌های دیگری هم وجود دارند که خواننده برای درک یک متن نیازمند دانستن آن‌هاست. یکی دیگر از این مهارت‌ها، دانش درک ساختار بکار رفته در متن است. برای فهم یک متن، درک جملات بکار رفته در آن نیز ضروری است. افزون بر این، نتایج تحقیقات نشان داده است که مهارت‌های سطح بالاتری از قبیل استنباط و نظارت بر درک هم در فرایند درک مطلب خواندن دخیل‌اند (اوک‌هیل، کین، مک‌کارتی، و نایتینگل<sup>۵</sup>، ۲۰۱۳).

1- Grabe & Stoller

2- Goodman

3- Paris, Wasik, & Turner

4- Memari Hanjani & Li

5- Oakhill, Cain, McCarthy, & Nightingale

۲-۱-۱- اهمیت ارزشیابی درک مطلب خواندن

رایج‌ترین سوال مربوط به ارزشیابی، درباره لزوم استفاده از آن است، سوالی که ALTE (۱۹۹۸) پاسخی واضح برای آن دارد، ارزشیابی جزئی ضروری از فرایند یادگیری است که نمی‌توان آن را از یادگیری جدا کرد، بدین معنی که بدون ارزشیابی، معلم قادر به سنجیدن سطح سواد دانش‌آموزان و یا تشخیص نقاط ضعف و قوت آنان نیست (ALTE، ۱۹۹۸). بهترین راه تصمیم‌گیری درباره دانستن وضعیت فعلی دانش‌آموزان، ارزشیابی است. البته ارزشیابی مزایای دیگری هم دارد که از آن جمله می‌توان به این مورد اشاره کرد که معلم می‌تواند به وسیله ارزشیابی به روش‌های تدریس خود بیندیشد تا مطمئن شود که آیا موثر بوده‌اند یا خیر. با وجود این، نوع ارزشیابی بسته به نوع فعالیتی که در حال ارزشیابی آنیم متفاوت خواهد بود. به ادعای براون و ابی ویکراما<sup>۱</sup> (۲۰۱۰)، مشکلی اساسی در راه ارزشیابی مهارت خواندن وجود دارد چراکه فرایند درک مطلب خواندن را نمی‌توان مشاهده کرد و تنها چیزی را که می‌توان از شخصی که در حال خواندن است، دید حرکت چشم‌هایش، ورق زدن صفحات، و در بهترین حالت دنبال کردن خطوط متن با مداد یا وسیله‌ای شبیه آن است. در حقیقت، راهی وجود ندارد که قادر به کنترل آنچه در ذهن خواننده می‌گذرد بود. این امر فرایند ارزشیابی مهارت خواندن را بسیار مشکل می‌کند، چرا که تنها راه فهمیدن ذهن خواننده، استنباط خواهد بود. بنابراین، همانطور که گرین<sup>۲</sup> (۲۰۱۴) می‌گوید این حقیقت که درک مطلب خواندن یک موضوع قابل مشاهده نیست، اصلی‌ترین چالش ارزشیابی درک مطلب خواندن است. در واقع، هدف اصلی ارزشیابی درک مطلب خواندن اینست که به زبان‌آموزان کمک کند تا خوانندگانی متفکر، منتقد، مسئول، و موثر باشند (آلورمان و کمیراس<sup>۳</sup>، ۱۹۹۸).

مطالعات بسیاری در زمینه سنجش درک مطلب خواندن انجام شده است. بطور مثال، وثوقی (۲۰۱۳) طی تحقیقی درباره آزمون‌های درک مطلب انتخابی به این نتیجه رسید که طراحان سوال معمولاً سعی می‌کنند متون درک مطلب را طوری انتخاب کنند که طولانی‌تر از یک پاراگراف نباشند، اما مشکل اینجاست که این کار تأثیر بسزایی در کاهش درک مطلب زبان‌آموزان دارد. حسینی (۲۰۱۵) نیز بر روی مهارت درک مطلب خواندن تحقیق کرد و به این نتیجه رسید که زبان‌آموزان وقتی به شکل دو نفری بر روی تمرین‌های درک مطلب فعالیت

1- Brown & Abeywickrama

2- Green

3- Alvermann & Commeyras

می‌کنند، با راحتی بیشتری روش‌های لازم را فرا می‌گیرند و قادر به فعال کردن دانش قبلی خود برای فهم بهتر متن خواهند بود.

برای طراحی یک آزمون درک مطلب خواندن مناسب، شناختن انواع مختلف فعالیت‌های مربوطه شامل فعالیت‌های انتخابی، تولیدی، استنباطی، و غیره ضروری است. فعالیت‌های انتخابی، فعالیت‌هایی‌اند که با استفاده از آن‌ها توانایی خواننده در تشخیص واژگان، دستور زبان، یا فهم متن، بدون نیاز به تولید زبان، سنجیده می‌شود. در این نوع فعالیت‌ها، تمرکز بر جنبه‌های دارای قاعده زبان مثل واژگان، دستور زبان و برخی جنبه‌های گفتاری است. از جمله سوال‌هایی که می‌توان در این نوع فعالیت استفاده کرد، می‌توان از غلط/درست، چندگزینه‌ای، وصل کردنی‌ها، و فعالیت‌های دارای عکس نام برد (براون و ابی ویکراما، ۲۰۱۰).

از طرف دیگر، سوروف و هگل هیمر<sup>۱</sup> (۲۰۱۴)، فعالیت‌های تولیدی را نوعی فعالیت توصیف می‌کنند که در آن‌ها از خواننده انتظار می‌رود با متن ارتباط برقرار کرده، معنی آن را بفهمد، و براساس فهم متن، چیزی تولید کند. در واقع، این نوع فعالیت‌ها تا حدی تولید دربردارند. متن‌های موجود در این نوع فعالیت‌ها بطور معمول از متن‌های موجود در فعالیت‌های انتخابی طولانی‌ترند. از جمله سوال‌هایی که می‌توان برای این نوع فعالیت استفاده کرد، می‌توان به پرکردن جای خالی با کلمه مناسب که خود خواننده باید آن را تشخیص دهد، پاسخ دادن به سوالات با استفاده از تعداد محدودی کلمات، و یافتن کلمه مناسب برای تعریف ارائه شده، نام برد (سوروف و هگل هیمر، ۲۰۱۴).

در نهایت، براون و ابی ویکراما (۲۰۱۰)، فعالیت‌های استنباطی را نوعی فعالیت درک مطلب خواندن توصیف کرده‌اند که در آن خواننده باید قادر باشد وارد سطوح بالاتری از متن شده و آنچه را نیاز دارد از کل متن استخراج کند. بایسته یادآوری است که تمرکز تحقیق حاضر بر دو نوع از این فعالیت‌ها، یعنی فعالیت‌های انتخابی و تولیدی متمرکز بوده است.

## ۲-۲- ارزشیابی پویا

ارزشیابی پویا، نوعی ارزشیابی است که در آن معلم از روش‌های مختلفی استفاده می‌کند تا پتانسیل یادگیری هر دانش‌آموز را بسنجد و هدف آن نیز نشان دادن حداکثر میزان عملکرد دانش‌آموزان در شرایطی است که کمک دریافت می‌کنند (پوهنر و لتالف، ۲۰۰۵). بنا به گفته

1- Suvorov & Hegelheimer

پوهنر و لتالف (۲۰۰۵)، اگر ارزشیاب‌ها بخواهند از میزان توانایی دانش‌آموزان بطور کامل آگاه شوند، افزون بر آنچه دربارهٔ آزمون‌های استاندارد سنتی می‌دانند باید در دو زمینه دیگر هم اطلاعات داشته باشند: عملکرد دانش‌آموز در صورت کمک یک میانجی و میزان فایده‌ای که دانش‌آموز از این کمک می‌برد.

در فرایند ارزشیابی پویا، فرد میانجی دانش فعلی زبان‌آموز را با استفاده از یک پیش‌آزمون می‌سنجد، مداخلات کمکی را به‌وی عرض می‌کند، و در آخر هم با استفاده از یک آزمون نهایی مشابه پیش‌آزمون، میزان بهبود عملکرد زبان‌آموز را می‌سنجد. مهم‌ترین فایدهٔ ارزشیابی پویا اینست که بازخوردی که در جریان ارزشیابی به‌زبان‌آموز ارائه می‌شود، باعث کاهش بار فکری وی شده و کمک شایانی به پیشرفت فعالیت‌های زبانی او می‌کند (جنگ و واگنر<sup>۱</sup>، ۲۰۱۴).

پایه و اساس ارزشیابی پویا بر انگارهٔ فرهنگی - اجتماعی گسترش آگاهی کودکان است که توسط وایگوتسکی معرفی شد (پوهنر، ۲۰۰۸). بر اساس این انگاره، هیچ کودکی بدون اینکه در کنار دیگران باشد، قادر به پیشرفت نیست. هدف این انگاره در زمینهٔ تحصیل، یافتن راهی است که از طریق آن آموزش بر پیشرفت دانش‌آموز اثر گذارد. مفهوم کلیدی در انگاره وایگوتسکی منطقهٔ نزدیک به پیشرفت (ZPD) است که خود وایگوتسکی آن را شکاف بین سطح فعلی توانایی هر شخص در حل مسئله و توانایی وی در حل مسئله با راهنمایی یک میانجی توصیف می‌کند، یعنی میزان عملکرد یک فرد در صورت کمک‌گرفتن از میانجی که بدون این کمک قادر به انجام آن کار نیست. سطح واقعی عملکرد زبان‌آموزان همانی است که به‌تنهایی، مستقل و بدون کمک هیچ شخص دیگری قادر به انجام آنند در حالی که منطقهٔ نزدیک به پیشرفت، توانایی عملکردی را مشخص می‌کند که زبان‌آموزان با کمک یک میانجی می‌توانند انجامش دهند (الجفرا و لتالف<sup>۲</sup>، ۱۹۹۴؛ بیرجندی و نجفی سارم، ۲۰۱۲؛ پوهنر و لتالف، ۲۰۰۵؛ شعبانی، خطیب، و عبادی، ۲۰۱۰؛ کلارک، ترولاو، هولم، و اسنولینگ<sup>۳</sup>، ۲۰۱۴).

لتالف و پوهنر (۲۰۰۴) و پوهنر (۲۰۰۸) ادعا کرده‌اند که دو نوع کلی میانجی‌گری در ارزشیابی پویا وجود دارد: کنشگر و مداخله‌جویانه. تمرکز ارزشیابی پویای کنشگر بر نوعی میانجی‌گری است که بر مبنای هم‌کنشی بین زبان‌آموز و ارزشیاب و در همان زمان هم‌کنشی

1- Jang & Wagner

2- Aljaafreh & Lantolf

3- Clarke, Truelove, Hulme, & Snowling

تعیین می‌شود، یعنی هیچ نقطه پایانی از پیش تعیین شده‌ای در این شیوه وجود ندارد. از طرف دیگر، ارزشیابی پویای مداخله‌جویانه، مراحل استاندارد از پیش تعیین شده‌ایند، یعنی میزان کمکی که هر زبان‌آموز دریافت خواهد کرد و نیز شیوه دریافت این کمک و راهبردهای مورد استفاده در دریافت آن از پیش تعیین شده است (پوهنر، ۲۰۰۸؛ شعبانی، ۲۰۱۱).

برای رسیدن به اهداف این تحقیق، مدل میانجی‌گری مداخله‌جویانه انتخاب شد. بر این اساس و طبق نظر پوهنر (۲۰۰۸)، معلم نباید به مشکلات زبان‌آموزان به‌هر شکلی که مایل است، پاسخ دهد، بلکه موظف است از یک برنامه دقیقاً طراحی شده‌ای که راهنمایی‌ها، اشارات و سوالات راهنما را دربردارد، استفاده کند. بایسته یادآوری است که تمامی مراحل این راهنمایی به‌شکل سلسله‌مراتبی از ضمنی و تلویحی تا آشکار و واضح مرتب شده‌اند. این نوع میانجی‌گری به‌طور معمول در ارائه پیشنهادات برای تعیین سطح و آموزش‌های متعاقب آن بسیار مفیدند. بطور خلاصه می‌توان گفت که ارزشیابی پویای مداخله‌جویانه نیازمند استفاده از فرایندهای اجرایی استاندارد و روش‌های آماری سنتی برای تحلیل داده‌ها و تفسیر آن‌هاست.

بنابه‌گفته گرب و جیانگ<sup>۱</sup> (۲۰۱۴)، دریافت بازخورد و هم‌کنشی میان معلم و دانش‌آموز، عامل بسیار مهمی در یادگیری است که از طریق آن معلم قادر به ارائه مداوم کمک و توضیح بیشتر درباره راهکار در هر بخشی از فرایند یادگیری است، بخصوص در مواردی که به‌نظر رسد، زبان‌آموز درست متوجه مطلب نشده است. هدف اصلی ارائه بازخورد دادن جواب درست به دانش‌آموز نیست، بلکه هدفش افزایش یادگیری، آگاه‌سازی زبان‌آموزان از راهبردهای یادگیری، و تشویق آن‌ها به استفاده از این راهبردها در جهت پیشرفت بیشتر است.

پوهنر و لتالف (۲۰۰۵) با بررسی تعدادی تحقیق انجام شده در زمینه ارزشیابی پویا به این نتیجه رسیدند که این نوع ارزشیابی تأثیر قابل توجهی بر چگونگی عملکرد زبان‌آموزان می‌گذارد. کامپرنول و ژانگ<sup>۲</sup> (۲۰۱۴) در تحقیق خود تفاوت چشمگیری بین نمرات معمول زبان‌آموزان و نمرات آن‌ها به هنگام کمک‌گرفتن از میانجی مشاهده کردند که نشان می‌داد میانجی‌گری، به‌خصوص میانجی‌گری مداخله‌جویانه، تأثیر بسزایی در فراگیری زبان‌آموزان داشت.

همچنین، تحقیقاتی چند در زمینه ارزشیابی پویا در ایران انجام شده است که در اینجا به‌چند مورد از آن‌ها اشاره می‌شود. نفرزاده (۲۰۱۶) از این نوع ارزشیابی در تدریس و افزایش

1- Grabe & Jiang

2- Compernelle & Zhang



یادگیری مهارت شنیداری زبان‌آموزان استفاده کرد و به این نتیجه رسید که ارزشیابی پویا تأثیر چشمگیری بر روی فراگیری مهارت شنیداری زبان‌آموزان دارد، چراکه در تحقیق وی، زبان‌آموزان در پس‌آزمون بسیار بهتر از پیش‌آزمون عمل کردند. همسو با این تحقیق، زندی (۲۰۱۸) نیز به بررسی تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان در فعالیت‌های انتخابی و تولیدی مهارت شنیداری پرداخت و دریافت که این نوع ارزشیابی تأثیر قابل توجهی بر هر دو نوع فعالیت انتخابی و تولیدی مهارت شنیداری، به خصوص نوع تولیدی این فعالیت‌ها دارد. در تحقیقی دیگر، آردین (۲۰۱۸) تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان در فعالیت‌های توصیفی و روایتی مهارت نوشتاری را بررسی کرد. نتایج بدست آمده این تحقیق هم نشان از تأثیر چشمگیر ارزشیابی پویا بر عملکرد زبان‌آموزان در هر دو نوع فعالیت داشت. کاظمی (۲۰۱۸) نیز با انجام تحقیقی در زمینه تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان در مهارت گفتاری به نتیجه‌ای مشابه سایر موارد رسید. به عبارت دیگر، در تحقیق وی ارزشیابی پویا تأثیر قابل توجهی بر بهبود مهارت گفتاری زبان‌آموزان داشت. سرانجام، شعبانی (۲۰۱۸) از نوع گروهی ارزشیابی پویا برای کمک به بهبود عملکرد نوشتاری زبان‌آموزان استفاده کرد و به این نتیجه رسید که استفاده از این نوع ارزشیابی می‌تواند کمک شایانی به تشخیص دلایل اصلی مشکلات زبان‌آموزان کند و همچنین نقاط قوت آن‌ها را هم به معلمان و هم به خود زبان‌آموزان بشناساند. از جمله سایر نتایج بدست آمده این تحقیق، این بود که ارزشیابی پویای گروهی، کمک بزرگی به ایجاد رابطه دوستانه بین معلمان و دانش‌آموزان می‌کند.

### ۳- روش تحقیق

همانطور که در بالا ذکر شد، هدف این تحقیق، بررسی تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان در فعالیت‌های انتخابی و تولیدی درک مطلب خواندن بود. بدین منظور سوالات زیر برای انجام این تحقیق مطرح شدند.

- ۱- آیا ارزشیابی پویا تأثیر معناداری بر عملکرد زبان‌آموزان در فعالیت‌های انتخابی مهارت درک مطلب خواندن دارد؟
- ۲- آیا ارزشیابی پویا تأثیر معناداری بر عملکرد زبان‌آموزان در فعالیت‌های تولیدی مهارت درک مطلب خواندن دارد؟
- ۳- آیا ارزشیابی پویا تفاوت معناداری در عملکرد زبان‌آموزان در فعالیت‌های انتخابی و تولیدی مهارت درک مطلب خواندن ایجاد می‌کند؟

برای پاسخگویی به این سوالات، در این بخش، شرکت‌کنندگان در تحقیق، ابزار استفاده شده در تحقیق، و مراحل انجام تحقیق توضیح داده می‌شود.

### ۳-۱- شرکت‌کنندگان در تحقیق

برای انجام این تحقیق، از ۶۰ زبان‌آموز خانم به‌عنوان شرکت‌کنندگان اصلی تحقیق استفاده شد که این افراد از بین ۷۸ زبان‌آموز سطح مبتدی یک آموزشگاه زبان که در ۴ کلاس دست‌نخورده حضور داشتند، انتخاب شدند. در این تحقیق، از زبان‌آموزانی استفاده شد که محققان به‌آنها دسترسی داشتند و به‌همین دلیل می‌توان گفت مدل نمونه‌گزینی غیر تصادفی یا آنچه بست و کان<sup>۱</sup> (۲۰۰۶) آن را نمونه‌گزینی در دسترس نامیده‌اند، مورد استفاده قرار گرفت. این زبان‌آموزان بین ۱۸ تا ۴۰ سال سن داشتند. بایسته یادآوری است که ۶۰ نفر زبان‌آموز مربوطه، از بین ۷۸ نفر اولیه بر مبنای عملکردشان در آزمون نلسون سطح مبتدی برگزیده شدند. بدین ترتیب که زبان‌آموزانی که نمره‌شان بین یک انحراف معیار بالاتر و یا پایینتر از میانگین آزمون نلسون بود، انتخاب شدند و بقیه زبان‌آموزان گرچه در تمامی مراحل تحقیق در کلاس‌هایشان حضور داشتند، ولی از آنجایی که نمره مناسب در آزمون نلسون را بدست نیاوردند، از اطلاعات و داده‌های آنها در مراحل آماری تحقیق استفاده نشد و در پاسخگویی به سوالات تحقیق شرکت داده نشدند. این ۴ کلاس بطور تصادفی به دو گروه تجربی ارزشیابی پویا و گروه کنترل تقسیم شدند. به دلیل تعداد کم زبان‌آموزان در هر کلاس، هر گروه متشکل از ۲ کلاس دست‌نخورده بود.

### ۳-۲- ابزار تحقیق

از چند ابزار مختلف در این تحقیق استفاده شد که هر کدام در ذیل توصیف شده‌اند:

### ۳-۲-۱- آزمون نلسون (Nelson, ۱۹۷۶)

از این آزمون برای تعیین سطح دانش زبان انگلیسی شرکت‌کنندگان این تحقیق برای یافتن زبان‌آموزان هم سطح استفاده شد. به‌گفته‌ی وود و لیدز (۲۰۰۷)، داشتن گروه‌های هم‌سطح در انجام تحقیق بسیار مهم است، چرا که این امر باعث کاهش تفاوت فاحش عملکرد

شرکت‌کنندگان شده و نتایج تحقیق را نیز قابل اطمینان‌تر می‌کند. از نسخه ۱۵۰ این آزمون که برای چک‌کردن سطح دانش زبان‌آموزان مبتدی طراحی شده استفاده شد. ۱۵۰ در این آزمون، به معنی این است که زبان‌آموزان، ۱۵۰ ساعت آموزش زبان انگلیسی دیده‌اند. بایسته یادآوری است که این آزمون از ۵۰ سوال چهار گزینه‌ای تشکیل شده است.

### ۳-۲-۲- آزمون KET

در تحقیق حاضر از دو آزمون KET به‌عنوان پیش‌آزمون (۲۰۱۵) و پس‌آزمون (۲۰۱۴) استفاده شد. به‌گفته‌ی الدرسن<sup>۱</sup> (۲۰۰۰) آزمون KET، آزمونی مناسب برای زبان‌آموزان سطح مبتدی است که از آن‌ها انتظار می‌رود بتوانند متون ساده‌ی مورد استفاده در زندگی روزمره یا در حین مسافرت به‌خارج از کشور را بفهمند. موضوعات مورد استفاده در این متون معمولاً درباره‌ی سفارش غذا در رستوران‌ها، خرید، شروع یک مکالمه و ادامه‌ی آن، و موضوعاتی مشابه است. بنابراین متون مورد استفاده اغلب به شکل آگهی‌ها، فرم‌ها، روزنامه‌ها و از این قبیل‌اند. هر آزمون KET از ۵۵ سوال از انواع مختلفی مثل چندگزینه‌ای، جای خالی، و وصل‌کردنی تشکیل شده است. سوالات هر کدام از دو آزمون مورد استفاده طوری ویرایش شدند که در نهایت این آزمون دارای ۲۸ سوال انتخابی و ۲۷ سوال تولیدی باشد. این دو آزمون ویرایش شده قبل از استفاده مورد مطالعه آزمایشی قرار گرفتند و پایایی آن‌ها بررسی شد. به‌عنوان نمونه‌ای از فعالیت‌های انتخابی مورد استفاده در آزمون KET استفاده شده در این تحقیق می‌توان به سوال زیر اشاره کرد. در این فعالیت، دانش‌آموزان باید جمله‌ای را که مربوط به مسافرت رفتن است، خوانده و گزینه مناسب برای پر کردن جای خالی را از بین پاسخ‌های ارائه شده انتخاب کنند.

They wanted to ..... somewhere near the sea.

- a. stand                      b. put                      c. stay

همچنین، به‌عنوان نمونه‌ای از فعالیت‌های تولیدی استفاده شده در آزمون KET این تحقیق می‌توان به سوال زیر اشاره کرد. در این فعالیت، دانش‌آموزان باید کلمه‌ای را از متن بیابند که معادل تعریف ارائه شده باشد.

I will repair your car for you. ....

### ۳-۲-۳- آزمون‌های KET

**تمرینی.** در طول ترم، از چهار آزمون KET تمرینی با موضوعات و چهارچوبی دقیقاً مشابه با پیش‌آزمون و پس‌آزمون در گروه تجربی ارزشیابی پویا استفاده شد. سوالات موجود در این آزمون‌ها هم دقیقاً شبیه به سوالات موجود در پیش‌آزمون و پس‌آزمون به شکل فعالیت‌های انتخابی و تولیدی بودند.

### ۳-۳- مراحل انجام تحقیق

هفتاد و هشت زبان آموز خانم سطح مبتدی یک موسسه آموزش زبان انگلیسی از ۴ کلاس دست‌نخورده که بین ۱۸ تا ۴۰ سال سن داشتند به صورت غیر تصادفی و نمونه‌گزینی در دسترس انتخاب و در آزمون نلسونی که متناسب با تعداد ساعت‌هایی بود که آموزش زبان دیده بودند شرکت کردند. سپس، پاسخ‌های آن‌ها تصحیح شد و بر مبنای نمرات به دست آمده، شرکت‌کنندگان نهایی در این تحقیق مشخص شدند. شرکت‌کنندگان نهایی ۶۰ نفری بودند که نمره آن‌ها در محدوده یک انحراف معیار بالاتر و پایین‌تر از نمره میانگین کل شرکت‌کنندگان اولیه در آزمون نلسون بود. سپس، این ۴ کلاس بطور تصادفی به دو گروه تجربی ارزشیابی پویا و گروه کنترل تقسیم شدند که هر گروه متشکل از ۲ کلاس ۳۰ نفره بود.

از آنجایی که مبنای ارزشیابی پویا، آزمون- میانجی‌گری- آزمون مجدد است، از افراد حاضر در گروه تجربی علاوه بر پیش‌آزمون و پس‌آزمون، چهار سری آزمون تمرینی KET هر کدام دوبار، یک بار قبل از میانجی‌گری و یک بار پس از میانجی‌گری، گرفته شد. به عبارت دیگر، در این گروه پس از انجام آزمون نلسون و انتخاب شرکت‌کنندگان اصلی، یک پیش‌آزمون (KET, ۲۰۱۵) به عمل آمد. در جلسه بعد کلاس، اولین آزمون تمرینی KET از زبان‌آموزان گرفته شد و پس از تصحیح برگه‌ها نقاط ضعف زبان‌آموزان مشخص شد و معلم با استفاده از راهبردهای از پیش تعیین شده به زبان‌آموزان کمک کرد تا مشکلات خود را بشناسند، پس از این کمک، دوباره همان آزمون تمرینی از زبان‌آموزان به عمل آمد. در اینجا، برای مشخص شدن چگونگی انجام میانجی‌گری در این گروه، مثالی ذکر میشود. در تعدادی از سوالات یکی از آزمون‌ها، از دانش‌آموزان خواسته شده بود تا کلمه‌ای مناسب برای پرکردن هر جای خالی بیابند. معلم پس از تصحیح برگه‌ها و یافتن مشکلات زبان‌آموزان در این قسمت، برای دانش‌آموزان توضیح داد که یکی از مشکلات عمده آن‌ها در این قسمت از آزمون بوده که نیاز به یافتن کلمه‌ای مناسب برای پرکردن جای خالی داشته‌اند. وی به بیان توضیحاتی چند در خصوص رفع این مشکل زبان‌آموزان پرداخت، بطور مثال، معلم بدون ارائه پاسخ درست، تنها

به دانش‌آموزان یادآوری می‌کرد که هر جای خالی نیاز به کلمه‌ای از یک گروه خاص واژگانی مثل اسم، فعل، صفت، و غیره دارد. با دانستن این موضوع، زبان‌آموزان می‌توانستند کلمه مناسب را راحت‌تر بیابند. این فرایند چهار بار تکرار شد. در آخر دوره هم پس‌آزمونی مشابه پیش‌آزمون (۲۰۱۴، KET) از زبان‌آموزان گرفته شد تا با مقایسه نتایج بدست آمده از این آزمون و پیش‌آزمون در مورد میزان تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان در فعالیتهای انتخابی و تولیدی درک مطلب خواندن نتیجه‌گیری کرد. گروه کنترل هم روال معمول فراگیری مهارت درک مطلب خواندن در کلاس‌های زبان را دنبال کردند و تنها در آزمون‌های نلسون، پیش‌آزمون و پس‌آزمون اصلی شرکت کردند. در طول دوره، تمرکز این گروه بر روی فعالیتهای مهارت درک مطلب خواندن موجود در کتاب درسیشان (۲۰۱۴، Touchstone) بود. در این گروه، در پایان هر متن درک مطلب خواندن، از انواع فعالیتهای انتخابی و تولیدی جهت فهم بهتر متن استفاده می‌شد تا زبان‌آموزان گروه کنترل نیز بطور کامل با این نوع فعالیتهای آشنا شوند. در انتها، برای نشان دادن میزان اثربخشی ارزشیابی پویا بر عملکرد زبان‌آموزان در فعالیتهای انتخابی و تولیدی درک مطلب خواندن، نتایج به دست آمده در پس‌آزمون این گروه با گروه تجربی مقایسه شد.

#### ۴- تجزیه تحلیل داده‌ها

در این قسمت، روش‌های آماری توصیفی و استنباطی بکار رفته جهت پاسخگویی به سوالات تحقیق توضیح داده می‌شود. نکته‌ای که در شروع انجام هر گونه تجزیه تحلیل آماری باید مورد توجه قرار گیرد، بررسی نرمال بودن عملکرد گروه‌ها در آزمون‌های مختلف مورد استفاده در تحقیق است. جدول ۱ و ۲ تجزیه تحلیل آماری مربوطه را نشان می‌دهد.

جدول ۱: آزمون کولموگروف- اسمیرنوف آزمون نلسون گروه‌های کنترل و ارزشیابی پویا

آزمون نلسون گروه پویا	آزمون نلسون گروه کنترل	
۳۰	۳۰	تعداد نفرات
۳۶,۹۰	۳۵,۵۳	میانگین
۳,۶۱	۴,۳۸	انحراف معیار
.۱۳	.۲۵	معناداری (دوسویه)

جدول ۱، میزان نرمال بودن آزمون نلسون را با استفاده از آزمون کولموگروف - اسمیرنوف برای هر دو گروه کنترل و ارزشیابی پویا نشان می‌دهد. از آنجا که اعداد معنادار دو گروه کنترل و ارزشیابی پویا به ترتیب ۲۵ و ۱۳. گزارش شده است، می‌توان این گونه نتیجه گرفت که نمرات نرمال‌اند زیرا این اعداد بزرگتر از میزان معناداری استاندارد ۰۰۵ می‌باشند. در نتیجه هرگونه تحلیل آماری درباره آزمون نلسون از طریق آزمونهای پارامتری قابل انجام است. بایسته یادآوری است که آزمونهای پارامتری دقیقتر از آزمونهای غیر پارامتری‌اند و نتایج بدست آمده از آنها قابل اطمینان تراند (پلنت، ۲۰۰۵).

جدول ۲: آزمون کولموگروف - اسمیرنوف پیش آزمون و پس آزمون گروه‌های کنترل و ارزشیابی پویا

پس آزمون گروه پویا	پیش آزمون گروه پویا	پس آزمون گروه کنترل	پیش آزمون گروه کنترل	
۳۰	۳۰	۳۰	۳۰	تعداد نفرات
۴۷,۴۷	۳۸,۷۷	۴۲,۶۰	۴۱,۸۰	میانگین
۳,۸۳	۵,۸۴	۴,۷۴	۴,۵۳	انحراف معیار
.۵۵	.۰۴*	.۲۶	.۰۶	معناداری (دوسویه)

از سوی دیگر، در جدول ۲ هم می‌توان میزان نرمال بودن عملکرد دو گروه کنترل و پویا را در پیش‌آزمون و پس‌آزمون دید. با نگاهی به این جدول، این نتیجه به‌دست می‌آید که بجز پیش‌آزمون گروه پویا که ۰۴ است، سایر نمرات همگی نرمال بوده‌اند، چرا که اعداد معناداری تمامی آنها بزرگتر از میزان معناداری استاندارد ۰۰۵ است. از طرفی عدد معناداری پیش‌آزمون گروه پویا نیز بسیار به‌میزان معناداری استاندارد نزدیک است. به این دلیل و از آنجایی که آزمونهای پارامتری دقیقتر از آزمونهای غیر پارامتری‌اند، مناسبترین نوع آزمون‌ها برای استفاده در این تحقیق، آزمونهای پارامتری در نظر گرفته شد.

پس از بررسی نرمال بودن عملکرد گروه‌ها در آزمونهای استفاده شده در تحقیق، پیش از انجام هر گونه تجزیه تحلیل آماری باید مطمئن شد که آیا شرکت‌کنندگان در گروه‌های مختلف تحقیق در ابتدای کار، سطح دانش زبانی مشابهی داشته‌اند یا خیر. از آنجایی که براساس

جدول ۱، عملکرد شرکت‌کنندگان در آزمون نلسون نرمال بود، برای مقایسه عملکرد دو گروه کنترل و ارزشیابی پویا در این آزمون، از یک آزمون تی مستقل استفاده شد که نتایج مربوطه به آن در جدول‌های ۳ و ۴ آمده‌اند. در ابتدا، جدول ۳ آمار توصیفی عملکرد دو گروه کنترل و ارزشیابی پویا در آزمون نلسون را نشان می‌دهد.

جدول ۳: آمار توصیفی عملکرد دو گروه کنترل و ارزشیابی پویا در آزمون نلسون

انحراف معیار	میانگین	تعداد دانش‌آموزان	
۴,۳۸	۳۵,۵۳	۳۰	گروه کنترل
۳,۶۱	۳۶,۹۰	۳۰	گروه ارزشیابی پویا

در جدول ۳، نمرات آزمون نلسون شرکت‌کنندگان حاضر در دو گروه نشان داده شده است. همانطور که مشخص است تفاوت کمی بین میانگین نمرات این گروه‌ها وجود دارد (میانگین گروه کنترل ۳۵,۵۳ و میانگین گروه ارزشیابی پویا ۳۶,۹۰) و به نظر می‌رسد که دو گروه از نظر زبانی تفاوت چشمگیری در ابتدای تحقیق نداشتند. همچنین، انحراف معیار دو گروه در آزمون نلسون کم و نزدیک به هم است (گروه کنترل ۴,۳۸ و گروه ارزشیابی پویا ۳,۶۱) که نشان می‌دهد تفاوت چشمگیری بین عملکرد شرکت‌کنندگان در هر گروه هم وجود نداشته است. با این وجود و برای بررسی دقیق وجود تفاوت معنادار بین این دو گروه، میزان این تفاوت با استفاده از آزمون تی مستقل سنجیده شد. نتایج این تجزیه تحلیل آماری در جدول ۴ آمده است.

جدول ۴: آزمون تی مستقل آزمون نلسون دو گروه کنترل و ارزشیابی پویا

آزمون تی مستقل برای تساوی میانگین‌ها			آزمون لوین برای تساوی واریانس‌ها		آزمون نلسون
معناداری (دوسویه)	درجه آزادی	ارزش t	معناداری	ارزش F	
۱۹.	۵۸	۳۱.۱-	۲۵.	۳۳.۱	واریانس‌های مساوی فرضی

با بررسی جدول ۴ می‌توان این گونه نتیجه گرفت که تفاوت چشمگیری بین سطح دانش زبانی شرکت‌کنندگان حاضر در دو گروه وجود نداشته است، چرا که عدد معناداری آزمون لوین برای تساوی واریانس‌های دو گروه ۲۵. گزارش شده است که بزرگتر از میزان معناداری استاندارد ۰۵. بوده و نشانه یکدستی عملکرد افراد در دو گروه کنترل و ارزشیابی پویا است. به‌همین دلیل، عدد معناداری آزمون تی مستقل مربوط به واریانس‌های مساوی فرضی جهت بررسی عملکرد دو گروه مورد استناد قرار گرفت که آن هم ۱۹. و بزرگتر از عدد استاندارد ۰۵. است و نشان می‌دهد که در ابتدای تحقیق تفاوت قابل ملاحظه‌ای بین عملکرد زبانی دو گروه وجود نداشته است. در نتیجه، می‌توان گفت که دو گروه کنترل و ارزشیابی پویا در ابتدای تحقیق سطح زبانی یکسانی داشته‌اند.

یکی دیگر از مهمترین نکاتی که در تحقیقات کمی باید بررسی شود، پایایی آزمون‌هایی است که در تحقیق مورد استفاده قرار گرفته‌اند. پایایی از طرق مختلف قابل اندازه‌گیری است ولی یکی از رایجترین انواع آن پایایی کرانیک آلفاست (براون<sup>۱</sup>، ۲۰۰۵). نتایج بدست آمده از این پایایی، برای آزمون‌های مورد استفاده این تحقیق در جدول ۵ ارائه شده است.

جدول ۵: پایایی پیش‌آزمون، پس‌آزمون و چهار آزمون درک مطلب خواندن KET

پس آزمون	آزمون ۲، بار دوم	آزمون ۴، بار اول	آزمون ۳، بار دوم	آزمون ۳، بار اول	آزمون ۲، بار دوم	آزمون ۲، بار اول	آزمون ۱، بار دوم	آزمون ۱، بار اول	پیش آزمون	
۵۸	۵۸	۵۸	۵۸	۵۸	۵۸	۵۸	۵۸	۵۸	۵۸	تعداد سوالات
۰.۷۶	-	-	-	-	-	-	-	-	۰.۷۶	گروه کنترل
۰.۷۹	۰.۷۴	۰.۷۵	۰.۷۶	۰.۷۶	۰.۷۵	۰.۷۷	۰.۷۶	۰.۷۷	۰.۷۸	گروه ارزشیابی پویا

همانطور که در جدول ۵ مشخص شده است، مقدار کرانیک آلفای بدست آمده برای تمامی آزمون‌ها، یعنی پیش‌آزمون، پس‌آزمون، و چهار آزمون تمرینی درک مطلب خواندن KET.

1- Brown



همگی بالاتر از عدد استاندارد ۷. است که این موضوع نشان‌دهنده پایایی و انسجام درونی قابل قبول همه این آزمون‌هاست. بنابراین، نتایج این تحقیق تا حد معقولی قابل اطمینان است و می‌توان نتایج آن را به جمعیت مورد نظر تعمیم داد.

نکته‌ای که باید درباره جدول ۵ توضیح داد مربوط به بعضی خانه‌های خالی در این جدول است. این خانه‌ها به این دلیل خالی‌اند که در گروه کنترل فقط پیش‌آزمون و پس‌آزمون از شرکت‌کنندگان گرفته شد در صورتی که در گروه تجربی ارزشیابی پویا، علاوه بر پیش‌آزمون و پس‌آزمون، چهار آزمون تمرینی درک مطلب خواندن KET هم مورد استفاده قرار گرفت که هر کدام دو بار از زبان‌آموزان گرفته شد، یکبار قبل از میانجی‌گری و یکبار پس از آن.

سپس، برای پاسخگویی به سوالات اول و دوم تحقیق و یافتن این نکته که آیا ارزشیابی پویا تأثیر معناداری بر میزان درک مطلب خواندن زبان‌آموزان در فعالیت‌های انتخابی و تولیدی دارد یا خیر و با توجه به نرمال بودن عملکرد شرکت‌کنندگان در پیش‌آزمون و پس‌آزمون که در جدول ۲ ذکر شد، از آزمون پارامتری تحلیل واریانس دوسویه سنجش مکرر یکبار برای مقایسه عملکرد زبان‌آموزان در پیش‌آزمون و پس‌آزمون فعالیت‌های انتخابی و یکبار برای مقایسه همین عملکرد در فعالیتهای تولیدی استفاده شد. جدول‌های ۶، ۷، و ۸ به‌علاوه شکل‌های ۱ و ۲ که در ادامه آمده است نتایج به‌دست آمده از این تجزیه تحلیل آماری را نشان می‌دهد.

جدول ۶: آمار توصیفی فعالیت‌های انتخابی و تولیدی درک مطلب خواندن پیش‌آزمون و پس‌آزمون دو گروه کنترل و ارزشیابی پویا

فعالیت‌های تولیدی پس‌آزمون	فعالیت‌های تولیدی پیش‌آزمون	فعالیت‌های انتخابی پس‌آزمون	فعالیت‌های انتخابی پیش‌آزمون		
۳۰	۳۰	۳۰	۳۰	تعداد شرکت‌کنندگان	
۶۰.۱۹	۷۰.۱۸	۰۰.۲۳	۹۷.۲۲	میانگین	گروه کنترل
۵۷.۳	۲۳.۳	۴۴.۲	۸۴.۲	انحراف معیار	
۸۳.۲۲	۰۷.۱۶	۶۳.۲۴	۷۰.۲۲	میانگین	گروه ارزشیابی پویا
۹۸.۱	۰۸.۴	۹۳.۱	۹۲.۲	انحراف معیار	

بنابر آنچه در جدول ۶ دیده می‌شود، میانگین عملکرد گروه کنترل در فعالیت‌های انتخابی پیش‌آزمون ۲۲،۹۷ و میانگین عملکرد گروه پویا در این فعالیت‌ها ۲۲،۷۰ است که مشابه و بسیار نزدیک به هم است. همچنین با بررسی میانگین عملکرد دو گروه در فعالیت‌های تولیدی در پیش‌آزمون که به ترتیب ۱۸،۷۰ و ۱۶،۰۷ برای دو گروه کنترل و پویاست این نتیجه به دست می‌آید که گروه کنترل در فعالیت‌های تولیدی پیش‌آزمون عملکرد بهتری داشته است. حال آنکه با بررسی عملکرد دو گروه در فعالیت‌های انتخابی و تولیدی پس‌آزمون و مقایسه آن‌ها با پیش‌آزمون، می‌توان گفت که عملکرد همه شرکت‌کنندگان در این تحقیق از پیش‌آزمون تا پس‌آزمون پیشرفت داشته است. با بررسی جزئیات این جدول می‌توان نتیجه گرفت که پیشرفت شرکت‌کنندگان حاضر در گروه کنترل در هر دو نوع فعالیت انتخابی و تولیدی در مقایسه با گروه تجربی ارزشیابی پویا کمتر بوده است، چرا که میانگین پیشرفت آن‌ها در فعالیت‌های انتخابی از پیش‌آزمون تا پس‌آزمون، تنها تغییر کمی از عدد ۲۲،۹۷ به ۲۳،۰۰ داشته است و عدد میانگینی که پیشرفت آن‌ها را در فعالیت‌های تولیدی نشان می‌دهد، تغییر کمی از ۱۸،۷۰ به ۱۹،۶۰ داشته است. عدد گزارش شده انحراف معیار فعالیت‌های انتخابی گروه کنترل هم تغییر کمی از ۲،۸۴ به ۲،۴۴ و عدد مربوطه به فعالیت‌های تولیدی این گروه تغییر کمی از ۳،۲۳ به ۳،۵۷ را نشان می‌دهد که نشان دهنده تفاوت کم در بین عملکرد زبان‌آموزان گروه کنترل از پیش‌آزمون به پس‌آزمون است. این در حالی است که تفاوت عملکرد شرکت‌کنندگان گروه تجربی ارزشیابی پویا، بسیار قابل توجه‌تر بوده است. بدین معنی که میانگین عملکرد زبان‌آموزان این گروه در فعالیت‌های انتخابی از پیش‌آزمون تا پس‌آزمون از ۲۲،۷۰ به ۲۴،۶۳ و عملکرد همین شرکت‌کنندگان در فعالیت‌های تولیدی از پیش‌آزمون تا پس‌آزمون از ۱۶،۰۷ به ۲۲،۸۳ تغییر کرده است. نیز، اعداد انحراف معیار دو نوع فعالیت‌های انتخابی و تولیدی گروه ارزشیابی پویا هم به ترتیب از ۲،۹۲ به ۱،۹۳ و از ۴،۰۸ به ۱،۹۸ تغییر پیدا کرده است که مویده تغییر عملکرد در بین شرکت‌کنندگان این گروه است. بعبارت دیگر، عملکرد زبان‌آموزان گروه ارزشیابی پویا در پس‌آزمون به مراتب به یکدیگر نزدیکتر بوده است که می‌تواند نشانگر تأثیر ارزشیابی پویا بر یکسان‌سازی درک مطلب خواندن آن‌ها باشد. در آخر هم با مقایسه میانگین عملکرد گروه کنترل و ارزشیابی پویا در فعالیت‌های انتخابی پس‌آزمون که به ترتیب ۲۳،۰۰ و ۲۴،۶۳ است، و با مقایسه میانگین عملکرد دو گروه در فعالیت‌های تولیدی در پس‌آزمون که به ترتیب ۱۹،۶۰ و ۲۲،۸۳ برای دو گروه کنترل و ارزشیابی پویاست این نتیجه به دست آمد که دو گروه در پس‌آزمون عملکردی متفاوت نسبت به یکدیگر در فعالیت‌های انتخابی و تولیدی

داشته‌اند. به عبارت دیگر گروه ارزشیابی پویا، عملکرد بهتری در انجام هر دو نوع فعالیت، در پس‌آزمون داشته است.

با این وجود، نکته مهم این است که آیا پیشرفت زبان‌آموزان دو گروه از پیش‌آزمون تا پس‌آزمون، پیشرفت قابل توجه و معناداری داشته یا خیر. پاسخ به این سوال از طریق تحلیل واریانس دوسویه سنجش مکرر قابل انجام است. دلیل انتخاب این روش آماری، لزوم مقایسه درون‌گروهی نمرات از پیش‌آزمون به پس‌آزمون در هر گروه و مقایسه نمرات بین دو گروه ارزشیابی پویا و کنترل است (پیچ، براور، و مک کینان<sup>۱</sup>، ۲۰۰۳). نتایج مربوط به این روش آماری، در جداول ۷ و ۸ گزارش شده است.

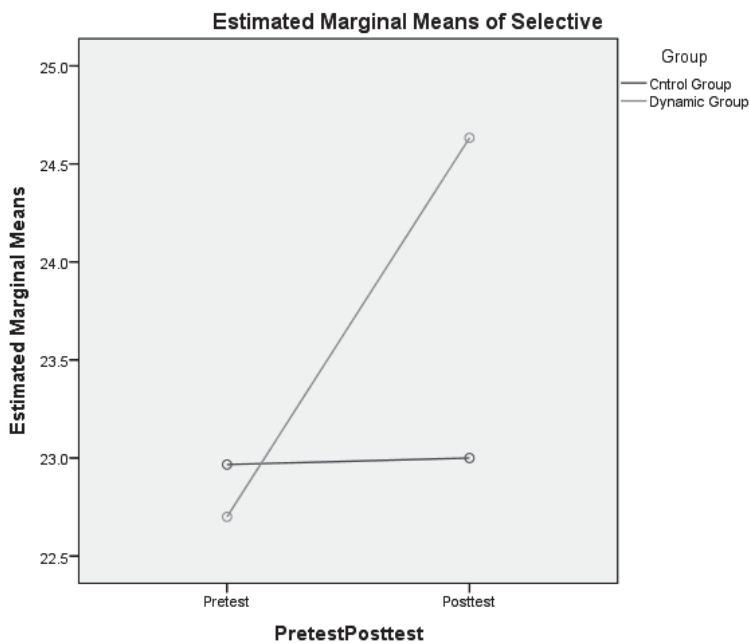
جدول ۷: تحلیل واریانس دوسویه سنجش مکرر فعالیت‌های انتخابی درک مطلب خواندن پیش‌آزمون و پس‌آزمون دو گروه کنترل و ارزشیابی پویا

تأثیر	مجموع مربع‌ها	درجه آزادی	میانگین مربع	ارزش F	معناداری	مجذور اتا جزئی
فعالیت‌های انتخابی	۲۹,۰۰	۱	۲۹,۰۰	۵۲.۸	*.۰۰	۱۲.
گروه	۱۴,۰۰	۱	۱۴,۰۰	۴۲.۱	۲۳.	۰۲.
فعالیت‌های انتخابی * گروه	۲۷,۰۷	۱	۲۷,۰۷	۹۵.۷	*.۰۰	۱۲.

در جدول ۷، عدد معناداری فعالیت‌های انتخابی ۰.۰۰ گزارش شده است که عددی کوچکتر از میزان معناداری استاندارد ۰.۰۵ است. در نتیجه می‌توان گفت که تفاوت قابل ملاحظه‌ای در عملکرد شرکت‌کنندگان در فعالیت‌های انتخابی از پیش‌آزمون تا پس‌آزمون دیده می‌شود. این بدان معناست که روش ارائه شده، تأثیر زیادی بر عملکرد شرکت‌کنندگان داشته است. افزون بر این، مجذور اتا جزئی گزارش شده هم ۱۲. است که نشان دهنده تأثیر متوسط روش ارائه شده بر عملکرد شرکت‌کنندگان در پس‌آزمون نسبت به پیش‌آزمون است. بایسته یادآوری است

که مجذور انا جزیی کمتر از ۰۱. نشانگر تأثیر کم، حدود ۰۶. تأثیر متوسط، و بالاتر از ۱۴. نشانگر تأثیر بالاست (پلنت، ۲۰۰۵). در ردیف دوم این جدول، عدد معناداری گزارش شده که برای گروه‌ها ۲۳. است، این عدد بزرگتر از میزان معناداری استاندارد ۰۵. است و این بدان معنی است که توانایی شرکت‌کنندگان دو گروه در انجام فعالیت‌های انتخابی در پس‌آزمون تفاوت قابل ملاحظه‌ای با یکدیگر نداشته است. یعنی، روش ارائه شده در گروه ارزشیابی پویا چندان متفاوت از روش مورد استفاده در گروه کنترل نبوده است. با این وجود، مهمتر از دو عدد معناداری گزارش شده قبلی، میزان معناداری بدست آمده برای هم‌کنشی فعالیت‌های انتخابی و گروه است که در ردیف سوم گزارش شده است. این عدد ۰۰. و کمتر از میزان معناداری استاندارد ۰۵. است. بنابراین، می‌توان نتیجه گرفت که تفاوت قابل ملاحظه‌ای در عملکرد شرکت‌کنندگان در فعالیت‌های انتخابی از پیش‌آزمون تا پس‌آزمون در دو گروه به‌وجود آمده است، اما این تفاوت در دو گروه به یک اندازه نبوده است. همانطور که در قسمت مجذور انا جزیی دیده می‌شود، عدد ۱۲. نشان‌دهنده تأثیر متوسط روش استفاده شده بر عملکرد شرکت‌کنندگان دو گروه در پس‌آزمون نسبت به پیش‌آزمون است.

افزون بر این، شکل ۱ هم که تفاوت بین پیش‌آزمون و پس‌آزمون فعالیت‌های انتخابی دو گروه کنترل و ارزشیابی پویا را نشان می‌دهد، نتایج ارائه شده بالا را تایید می‌کند. همانطور که در شکل ۱ دیده می‌شود، اعضای شرکت‌کننده در دو گروه این تحقیق، در پیش‌آزمون عملکرد تقریباً مشابهی در فعالیت‌های انتخابی داشته‌اند. با این وجود، در پس‌آزمون میزان پیشرفت آن‌ها به شکل قابل توجهی متفاوت از یکدیگر بوده است. به عبارت دیگر، عملکرد گروه ارزشیابی پویا پیشرفت بیشتری در پس‌آزمون نسبت به عملکرد گروه کنترل داشته است. با توجه به نتایج دیده شده در این شکل و آمار توصیفی ارائه شده در جدول ۶، می‌توان این گونه نتیجه گرفت که گروه تجربی حاضر در این تحقیق، یعنی گروه ارزشیابی پویا، پیشرفت عملکرد قابل توجهی در فعالیت‌های انتخابی درک مطلب خواندن نسبت به گروه کنترل داشته است.



شکل ۱. تفاوت بین پیش‌آزمون و پس‌آزمون فعالیت‌های انتخابی دو گروه کنترل و ارزشیابی پویا

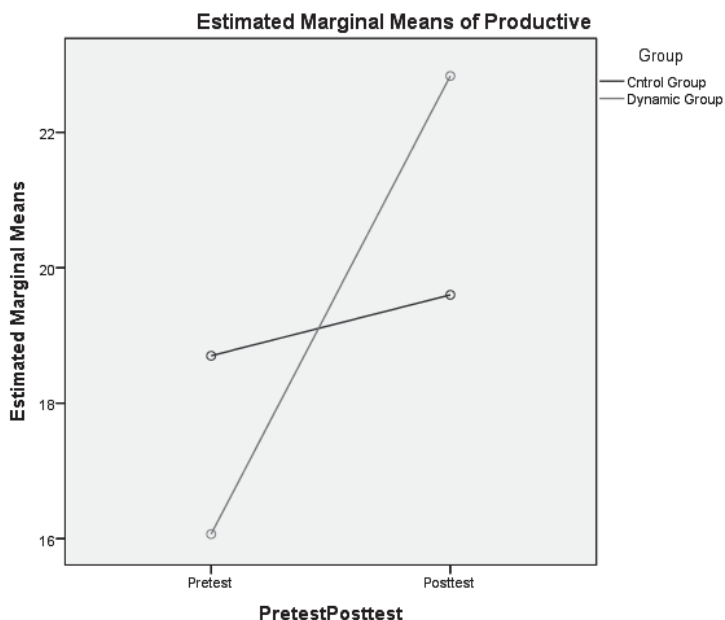
در همین راستا، جدول ۸ و شکل ۲ برای بررسی میزان تأثیرگذاری ارزشیابی پویا بر عملکرد زبان‌آموزان در فعالیت‌های تولیدی درک مطلب خواندن ارائه شده است.

جدول ۸: تحلیل واریانس دوسویه سنجش مکرر فعالیت‌های تولیدی درک مطلب خواندن پیش‌آزمون و پس‌آزمون دو گروه کنترل و ارزشیابی پویا

تأثیر	مجموع مربع‌ها	درجه آزادی	میانگین مربع	ارزش F	معناداری	مجذور اتا جزئی
فعالیت‌های تولیدی	۴۴۰٫۸۳	۱	۴۴۰٫۸۳	۱۵٫۸۸	*۰۰	۶۰
گروه	۲٫۷۰	۱	۲٫۷۰	۱۵	۶۹	۰۰
فعالیت‌های تولیدی * گروه	۲۵۸٫۱۳	۱	۲۵۸٫۱۳	۶۲٫۵۱	*۰۰	۴۷

در جدول ۸، عدد معناداری گزارش شده، برای فعالیت‌های تولیدی ۰۰ است که عددی کوچکتر از میزان معناداری استاندارد ۰۵ است، و این بدان معناست که تفاوت قابل ملاحظه‌ای در عملکرد کل شرکت‌کنندگان در فعالیت‌های تولیدی از پیش‌آزمون تا پس‌آزمون دیده شده است. بنابراین، روش بکار رفته در گروه‌ها تأثیر زیادی بر عملکرد آن‌ها در پس‌آزمون داشته است. افزون بر این، مجذور اتا جزئی گزارش شده مربوطه ۶۰ است که نشان دهنده تأثیر زیاد روش ارائه شده بر عملکرد شرکت‌کنندگان در پس‌آزمون نسبت به پیش‌آزمون است. در ردیف دوم، عدد معناداری گزارش شده برای گروه‌ها ۶۹ است که بزرگتر از عدد معناداری استاندارد ۰۵ است. در نتیجه می‌توان گفت که عملکرد شرکت‌کنندگان دو گروه در انجام فعالیت‌های تولیدی در پس‌آزمون، تفاوت قابل ملاحظه‌ای با یکدیگر نداشته است. در این جدول نیز اطلاعات موجود در ردیف سوم حائز اهمیت بیشتری است. میزان معناداری گزارش شده برای هم‌کنشی فعالیت‌های تولیدی و گروه ۰۰ است که باز هم کمتر از عدد معناداری استاندارد ۰۵ است. باز هم نتیجه بدست آمده این است که تفاوت قابل ملاحظه‌ای در عملکرد شرکت‌کنندگان دو گروه کنترل و ارزشیابی پویا در فعالیت‌های تولیدی از پیش‌آزمون تا پس‌آزمون به وجود آمده است، اما این تفاوت در دو گروه به یک اندازه نبوده است. همانطور که در قسمت مجذور اتا جزئی دیده می‌شود، عدد ۴۷ هم نشان‌دهنده تأثیر زیاد روش استفاده شده بر عملکرد شرکت‌کنندگان در فعالیت‌های تولیدی درک مطلب خواندن در پس‌آزمون نسبت به پیش‌آزمون است.

افزون بر این، شکل ۲ هم نتایج ارائه شده در جدول ۸ را تأیید می‌کند. همان‌طور که در این شکل می‌توان دید، در فعالیت‌های تولیدی پیش‌آزمون، عملکرد گروه کنترل بهتر از گروه ارزشیابی پویا بوده است، در حالی که در پس‌آزمون عکس این مورد دیده می‌شود و عملکرد گروه تجربی بالاتر از گروه کنترل است. در واقع، عملکرد شرکت‌کنندگان گروه ارزشیابی پویا در فعالیت‌های تولیدی پیشرفت بسیار قابل توجهی در پس‌آزمون در مقایسه با پیش‌آزمون داشته است، در حالی که گروه کنترل پیشرفت چندانی نداشته است. در نتیجه، با توجه به شکل ۲ و آمار توصیفی ارائه شده در جدول ۶ می‌توان گفت که گروه تجربی ارزشیابی پویا پیشرفت عملکرد قابل ملاحظه‌ای در فعالیت‌های تولیدی نسبت به گروه کنترل از پیش‌آزمون تا پس‌آزمون داشته است.



شکل ۲. تفاوت بین پیش‌آزمون و پس‌آزمون فعالیت‌های تولیدی دو گروه کنترل و ارزشیابی پویا

اولین نکته‌ای که پس از انجام پیش‌آزمون در مورد همگی شرکت‌کنندگان حاضر در این تحقیق روشن شد، این بود که تقریباً همه آن‌ها مشکلات بسیاری در انجام فعالیت‌های انتخابی و به‌خصوص تولیدی درک مطلب خواندن داشتند. به‌هر حال، نتایج این تحقیق نشان داد که ارزشیابی پویا می‌تواند عملکرد زبان‌آموزان را در فعالیت‌های انتخابی و به‌خصوص تولیدی درک مطلب خواندن را بطور چشمگیری بهبود بخشد. نکته قابل‌تأمل ضعف عمده زبان‌آموزان در انجام فعالیت‌های تولیدی در ابتدای تحقیق بود و مشخص شد که زبان‌آموزان بیشترین نیاز به‌کمک را در این نوع فعالیت‌ها دارند.

در جمع‌بندی این نتایج می‌توان گفت هر چند همه شرکت‌کنندگان حاضر در این تحقیق، در پس‌آزمون عملکرد بهتری نسبت به پیش‌آزمون داشته‌اند، اما این تفاوت در گروه کنترل چندان قابل‌ملاحظه نبوده است، در حالی‌که بهبود عملکرد زبان‌آموزان در گروه تجربی ارزشیابی پویا، قابل‌ملاحظه بوده است. این مسئله را می‌توان نتیجه بازخوردهایی دانست که شرکت‌کنندگان این گروه بطور متناوب دریافت می‌کردند، در حالی‌که زبان‌آموزان حاضر در گروه کنترل از اینگونه بازخوردهای پیاپی بی‌بهره بودند.

با توجه به نتایج بدست آمده در جدول‌های ۶ و ۷ و نیز شکل ۱ می‌توان به سوال اول تحقیق، بدین‌گونه پاسخ داد که "ارزشیابی پویا تأثیر معناداری روی عملکرد زبان‌آموزان در فعالیت‌های انتخابی مهارت درک مطلب خواندن دارد". همچنین با توجه به اطلاعات موجود در جدول‌های ۶ و ۸ و شکل ۲ می‌توان به سوال دوم تحقیق اینگونه جواب داد که "ارزشیابی پویا، تأثیر معناداری روی عملکرد زبان‌آموزان در فعالیت‌های تولیدی مهارت درک مطلب خواندن دارد".

در پایان، برای پاسخگویی به سوال سوم تحقیق، از تحلیل واریانس چند متغیری برای بررسی عملکرد زبان‌آموزان گروه ارزشیابی پویا در مقایسه با گروه کنترل از دو نوع فعالیت انتخابی و تولیدی در پیش‌آزمون و پس‌آزمون استفاده شد. دلیل انتخاب این تحلیل داده انجام مقایسه درون‌گروهی فعالیت‌های انتخابی و تولیدی از پیش‌آزمون به پس‌آزمون در هر گروه و مقایسه عملکرد دو گروه ارزشیابی پویا و کنترل با یکدیگر بود، چراکه برای پیدا کردن رابطه بین متغیرهای وابسته بصورت ترکیبی (در اینجا انتخابی و تولیدی) با توجه به متغیر(های) مستقل، نیاز به تحلیل واریانس چند متغیری است (هیتون، براون لو، مک مورای، و کوزنز<sup>۱</sup>، ۲۰۰۸). با استفاده از این نوع تحلیل، می‌توان میزان بهبود عملکرد شرکت‌کنندگان هر گروه و نیز بهبود عملکردشان نسبت به یکدیگر را از پیش‌آزمون تا پس‌آزمون سنجید. نتایج این تحلیل در جدول‌های ۹ و ۱۰ گزارش شده است. نخست، جدول ۹ تحلیل واریانس چند متغیری نتایج پیش‌آزمون و پس‌آزمون را بصورت کلی نشان می‌دهد.

جدول ۹: تحلیل واریانس چندمتغیری پیش‌آزمون و پس‌آزمون دو گروه کنترل و ارزشیابی پویا

مجدور اتا جزئی	معناداری	ارزش F	تأثیر	
.۰۱	.۳۵	۱.۰۵	گروه	آزمون و بیلکس لامبدا
.۲۵	.۰۰*	۱۹.۹۱	زمان (پیش‌آزمون و پس‌آزمون)	
.۱۶	.۰۰*	۱۱.۷۰	گروه * زمان	

1- Hinton, Brownlow, McMurray, & Cozens



با بررسی میزان معناداری گزارش شده در جدول ۹ برای گروه‌هایی که عدد ۳۵ و بزرگتر از میزان معناداری استاندارد ۰۰۵ است، می‌توان این گونه نتیجه گرفت که تفاوت چشمگیری بین عملکرد زبان‌آموزان دو گروه دیده نشده است. مجذور اتا جزئی گزارش شده برای گروه ۰۱ است که نشان‌دهنده تأثیر کم شیوه‌های بکار رفته در گروه‌هاست. از طرف دیگر، میزان معناداری زمان (از پیش‌آزمون به پس‌آزمون) ۰۰ است که کوچکتر از میزان معناداری استاندارد ۰۰۵ است و نشان می‌دهد عملکرد زبان‌آموزان دو گروه در پس‌آزمون، تفاوت قابل ملاحظه‌ای نسبت به پیش‌آزمون داشته است، میزان این تفاوت، زیاد هم بوده است، چرا که مجذور اتا جزئی گزارش شده از ۲۵ است. در این جدول نیز اطلاعات موجود در ردیف سوم حائز اهمیت بیشتری است. میزان معناداری گزارش شده، برای هم‌کنشی گروه و زمان ۰۰ است که کمتر از عدد معناداری استاندارد ۰۰۵ است و این نتیجه بدست می‌آید که تفاوت قابل ملاحظه‌ای در عملکرد شرکت‌کنندگان دو گروه کنترل و ارزشیابی پویا از پیش‌آزمون تا پس‌آزمون وجود دارد، اما این تفاوت در دو گروه به یک اندازه نبوده است. مجذور اتا جزئی این مورد هم که عدد ۱۶ است، نشان می‌دهد که این تأثیر زیاد بوده است. برای بررسی دقیقتر بهبود عملکرد شرکت‌کنندگان دو گروه در فعالیت‌های انتخابی و تولیدی از پیش‌آزمون به پس‌آزمون جدول ۱۰ ارائه شده است.

با توجه به عدد معناداری فعالیت‌های انتخابی مربوط به گروه در جدول ۱۰ که ۱۴ است، می‌توان نتیجه گرفت که تفاوت چشمگیری در عملکرد زبان‌آموزان دو گروه در فعالیت‌های انتخابی از پیش‌آزمون تا پس‌آزمون وجود نداشته است. بررسی مجذور اتا جزئی هم که عدد ۰۱ است، نشان می‌دهد که این تفاوت بسیار جزئی بوده است. همچنین، عدد معناداری گروه در فعالیت‌های تولیدی از پیش‌آزمون تا پس‌آزمون ۶۲ است که عددی بزرگتر از میزان معناداری استاندارد ۰۰۵ است که باز هم نشانگر آن است که تفاوت چشمگیری بین عملکرد زبان‌آموزان دو گروه در فعالیت‌های تولیدی از پیش‌آزمون تا پس‌آزمون نیز وجود نداشته است. عدد مجذور اتا جزئی این مورد هم ۰۰ است که موید تفاوت بسیار کم گروه‌هاست.

جدول ۱۰: تحلیل واریانس چندمتغیری فعالیت‌های انتخابی و تولیدی درک مطلب خواندن پیش‌آزمون و پس‌آزمون دو گروه کنترل و ارزشیابی پویا

منبع	متغیر وابسته	مجموع مربع‌ها	درجه آزادی	میانگین مربع	ارزش F	معناداری	مجذور اتا جزئی
گروه	فعالیت‌های انتخابی	۱۴,۰۰	۱	۱۴,۰۰	۲,۱۲	.۱۴	.۰۱
	فعالیت‌های تولیدی	۲,۷۰	۱	۲,۷۰	.۲۴	.۶۲	.۰۰
زمان (پیش‌آزمون و پس‌آزمون)	فعالیت‌های انتخابی	۲۹,۰۰	۱	۲۹,۰۰	۴,۳۹	.۰۳*	.۰۳
	فعالیت‌های تولیدی	۴۴۰,۸۳	۱	۴۴۰,۸۳	۴۰,۱۵	.۰۰*	.۲۵
گروه* زمان	فعالیت‌های انتخابی	۲۷,۰۷	۱	۲۷,۰۷	۴,۰۹	.۰۴*	.۰۳
	فعالیت‌های تولیدی	۲۵۸,۱۳	۱	۲۵۸,۱۳	۲۳,۵۱	.۰۰*	.۱۶

از آنجا که عدد معناداری فعالیت‌های انتخابی و تولیدی مربوط به زمان (یعنی از پیش‌آزمون به پس‌آزمون) به ترتیب اعداد ۰.۰۳ و ۰.۰۰ و هر دو کوچکتر از میزان معناداری استاندارد ۰.۰۵ اند، این نتیجه به دست می‌آید که عملکرد هر دو گروه در هر دو نوع فعالیت انتخابی و تولیدی در طول زمان پیشرفت چشمگیری داشته است. مجذور اتا جزئی مربوط به فعالیت‌های انتخابی مربوط به زمان عدد ۰.۰۳ است که نشانگر تفاوت کم در عملکرد زبان‌آموزان از پیش‌آزمون به پس‌آزمون در این نوع فعالیت است. با این وجود، مجذور اتا جزئی مربوط به فعالیت‌های تولیدی مربوط به زمان ۰.۲۵ است که نشانگر تفاوت زیاد در عملکرد زبان‌آموزان از پیش‌آزمون به پس‌آزمون در این نوع فعالیت است.

در این جدول نیز، مهمتر از نتایج بدست آمده درباره گروه و زمان بصورت جداگانه، میزان معناداری بدست آمده که مربوط به هم‌کنشی گروه و زمان است که در ردیف سوم گزارش شده است. این عدد برای فعالیت‌های انتخابی ۰۴. و برای فعالیت‌های تولیدی ۰۰. گزارش شده که هر دو کمتر از میزان معناداری استاندارد ۰۵. اند. بنابراین، می‌توان نتیجه گرفت که تفاوت قابل ملاحظه‌ای در عملکرد شرکت‌کنندگان در فعالیت‌های انتخابی و تولیدی از پیش‌آزمون تا پس‌آزمون در هر دو گروه به‌وجود آمده است که این تفاوت، نه تنها در دو گروه بلکه در دو نوع فعالیت، هم یکسان نبوده است. همانطور که در بخش مجذور ا تا جزیی دیده می‌شود، عدد ۰۳. نشان‌دهنده تأثیر کم روش استفاده شده بر عملکرد شرکت‌کنندگان دو گروه در پس‌آزمون فعالیت‌های انتخابی نسبت به پیش‌آزمون است. از طرف دیگر، مجذور ا تا جزیی گزارش شده برای فعالیت‌های تولیدی ۱۶. است که تأثیر زیاد روش بکار رفته بر عملکرد شرکت‌کنندگان را در این نوع فعالیت نشان می‌دهد.

با توجه به نتایج بدست آمده در جدول‌های ۹ و ۱۰ و بررسی مجدد آمار توصیفی گروه‌ها در هر دو نوع فعالیت انتخابی و تولیدی در پیش‌آزمون و پس‌آزمون ارائه شده در جدول ۶، می‌توان به‌سوال سوم تحقیق اینگونه پاسخ داد که "ارزشیابی پویا تفاوت معناداری در عملکرد زبان‌آموزان در فعالیت‌های انتخابی و تولیدی مهارت درک مطلب خواندن ایجاد می‌کند" بطوریکه تأثیر مثبت این ارزشیابی بر عملکرد زبان‌آموزان در فعالیت‌های تولیدی بیشتر است.

نتایج تحقیق حاضر همسو با نتایج بدست آمده از تحقیقات مشابه در زمینه ارزشیابی پویاست. بطور مثال لتالف و پوهنر (۲۰۰۴) تحقیقی در زمینه مدل دارای میانجی ارزشیابی پویا انجام دادند و به این نتیجه رسیدند که این نوع ارزشیابی تأثیر قابل ملاحظه‌ای بر روی عملکرد زبان‌آموزان داشته است، چرا که تغییر نمره محسوسی از پیش‌آزمون به پس‌آزمون دیده شد. آن‌ها همچنین به این نتیجه رسیدند که این نوع ارزشیابی معلم را قادر می‌سازد تا بطور دقیق زبان‌آموزان را به‌سه‌گروه ضعیف، متوسط، و قوی تقسیم کند، تا بتواند به هر گروه آموزش لازم را بدهد. آنتون<sup>۱</sup> (۲۰۰۳) در بخشی از تحقیق خود از ارزشیابی پویا استفاده کرد و به این نتیجه دست یافت که ارزشیابی پویا به معلم کمک می‌کند توانایی‌هایی از زبان‌آموزان را بشناسد که دستیابی به آن‌ها از هیچ طریق دیگری ممکن نیست. وین و همکارانش (۲۰۱۶) هم با استفاده از ارزشیابی پویا سعی در بهبود دامنه لغات وابسته به موضوع زبان‌آموزان کردند و در

1- Antón

پایان کار به این نتیجه رسیدند که با استفاده از این نوع ارزشیابی می‌توان تأثیر قابل توجهی بر بهبود عملکرد زبان‌آموزان در رابطه با یادگیری لغات موضوعی داشت. در تحقیقات جدید در زمینه ارزشیابی پویا، آردین (۲۰۱۸)، زندی (۲۰۱۸)، و کاظمی (۲۰۱۸) نیز همگی به این نتیجه رسیدند که ارزشیابی پویا، به ترتیب مهارت‌های نوشتاری، شنیداری، و گفتاری زبان‌آموزان را بطور چشمگیری افزایش می‌دهد. همچنین، دفتری فرد و بیرجندی (۲۰۱۷) با استفاده از دو نوع رویکرد نسبت به میانجی‌گری در ارزشیابی پویا، با نام‌های میانجی‌گری مربوط به پاسخ و میانجی‌گری مربوط به فعالیت، بر روی درک مطلب خواندن زبان‌آموزان ایرانی تحقیق کردند و همچون سایر محققان در گستره ارزشیابی پویا به این نتیجه رسیدند که شرکت‌کنندگان حاضر در هر دو گروه، پیشرفت قابل ملاحظه‌ای در پس‌آزمون داشتند. این در حالی است که رویکرد میانجی‌گری مربوط به فعالیت تأثیر فزونی بر عملکرد زبان‌آموزان داشت.

نتایج این تحقیق نشان داد که مدل آزمون - میانجی‌گری - آزمون مجدد که رایج‌ترین نوع ارزشیابی پویاست، می‌تواند تأثیر بسزایی در بهبود عملکرد زبان‌آموزان در فعالیت‌های انتخابی و تولیدی مهارت درک مطلب خواندن داشته باشد که البته این پیشرفت در فعالیت‌های تولیدی قابل ملاحظه‌تر بوده است. از آنجایی که پس از خواندن هر متنی و برای نشان دادن درک آن متن معمولاً نیاز است که به‌سوالاتی در مورد آن متن پاسخ داده شود، توانایی در پاسخگویی به انواع سوال‌ها و فعالیت‌ها برای نشان دادن دانش زبانی زبان‌آموزان ضروری است. همانطور که پیشتر گفته شد، فعالیت‌های متنوعی وجود دارد که برای بررسی درک یک متن می‌توان از آن‌ها استفاده کرد؛ اهم آن‌ها فعالیت‌های انتخابی و تولیدی است که معمولاً در مقایسه با یگدیگر بکار می‌روند. در نتیجه، توانایی زبان‌آموزان در پاسخگویی به این دو نوع فعالیت می‌تواند نشانگر افزایش سطح دانش زبانی آن‌ها حداقل در مهارت خواندن باشد. در این راستا، نتایج بدست آمده از این تحقیق، اهمیت استفاده از این نوع فعالیت‌ها، بخصوص فعالیت‌های تولیدی، در کلاس‌های زبان انگلیسی را نشان می‌دهد. یادگیری زبان، چه زبان اول و چه زبان دوم، شامل دو قسمت فهم زبان و تولید زبان است. عملکرد خوب زبان‌آموزان در فعالیت‌های انتخابی نشانگر توانایی آن‌ها در فهم زبان و عملکرد خوب آن‌ها در فعالیت‌های تولیدی نشانگر توانایی آن‌ها در تولید زبان است. عملکرد خوب زبان‌آموزان در هر کدام از این فعالیت‌ها، نشانگر پیشرفت دانش زبانی آن‌هاست. از آنجایی که نتایج این تحقیق نشان داد که استفاده از ارزشیابی پویا، تأثیر مثبت بسزایی هم در فهم زبان و هم در تولید زبان دانش‌آموزان دارد، می‌توان گفت که استفاده از ارزشیابی پویا در کلاس‌های زبان، بسیار حائز اهمیت و ضروری است.

از بارزترین مشخصات ارزشیابی پویا، نقش پررنگ معلم در کمک به پیشرفت زبان‌آموزان است. معلم در این نوع ارزشیابی تمام اهتمام خویش را در شناسایی و اصلاح نقاط ضعف دانش‌آموزان با استفاده از روش‌های متناسب انجام می‌دهد (هاردینگ و همکارانش ۲۰۱۵). بنا بر آنچه تاکنون گفته شد و با توجه به تأثیر چشمگیر ارزشیابی پویا بر عملکرد زبان‌آموزان در پژوهش‌های مختلف، می‌توان به این نوع ارزشیابی به‌عنوان مدلی مناسب برای آموزش به زبان‌آموزان نگاه کرد. در صورتی که بتوان استفاده از ارزشیابی پویا را در زبان‌آموزان و معلمان نهادینه کرد، کم‌کم دانش‌آموزان قادر به استفاده مستقل از این روش خواهند بود، بطوری‌که خودشان به‌سادگی نقاط ضعف خود را شناسایی کرده و در جهت رفع مشکلاتشان بر آیند. سرانجام، می‌توان گفت استفاده از این نوع ارزشیابی و آموزش آن به دانش‌آموزان در طولانی مدت، منجر به وجود آمدن زبان‌آموزان مستقل و کارآمد و کاهش فشار کاری معلمان می‌شود.

#### ۵- نتیجه‌گیری

این تحقیق به‌منظور بررسی تأثیر ارزشیابی پویا بر عملکرد زبان‌آموزان انگلیسی، به‌عنوان زبان دوم بر روی فعالیت‌های انتخابی و تولیدی مهارت درک مطلب خواندن طراحی شد و برای انجام آن از طرح پیش‌آزمون/ پس‌آزمون نیمه آزمایشی و مدل نمونه‌گزینی در دسترس استفاده شد.

با توجه به نتایج به‌دست آمده در این تحقیق، می‌توان گفت ارزشیابی پویا می‌تواند به زبان‌آموزان، معلمان، طراحان آزمون، و نویسندگان کتاب‌های درسی کمک قابل توجهی کند. معلمان و دانش‌آموزان می‌توانند با استفاده از این نوع ارزشیابی، مشکلات فراگیری زبان دوم را یافته و با تمرین روش‌های مشخص شده منحصر به این نوع ارزشیابی، به‌حل این مشکلات بپردازند. آن‌ها می‌توانند پس از چندین بار تمرین، چگونگی یافتن مشکلات، با استفاده از ارزشیابی پویا، کاملاً بر مشکلات غلبه کنند و در همه آزمون‌ها موفق عمل کنند. این نوع ارزشیابی می‌تواند به طراحان آزمون نیز کمک کند تا آزمون‌ها را به‌گونه‌ای طراحی کنند که بر اساس دیدگاه‌های نوین آموزش و ارزشیابی باشد بطوریکه این آزمون‌ها نه فقط در جهت مشخص کردن سطح دانش زبان‌آموزان مورد استفاده قرار گیرند، بلکه بتوان از آن‌ها برای پیدا کردن اصلی‌ترین مشکلات زبان‌آموزان و کمک به آن‌ها در رفع این مشکلات در هر سطح استفاده کرد. به‌عبارت دیگر، اگر طراحان آزمون با هدف یافتن مشکلات زبانی زبان‌آموزان سطوح مختلف به‌طراحی سوال بپردازند، پس از مدت زمان کوتاهی، آزمون تبدیل به ابزاری نه

فقط برای سنجیدن میزان یادگیری زبان‌آموزان بلکه به‌ابزاری برای یافتن چرایی و چگونگی مشکلات رایج تبدیل خواهد شد. همچنین، طراحان کتاب‌های درسی هم می‌توانند با بهره‌گیری از این نوع ارزشیابی و راهبردهای مختص آن درجه سختی هر فعالیت را برای زبان‌آموزان سطوح مختلف مشخص کرده و روال آموزشی را طوری تعیین کنند که برای زبان‌آموز ساده‌تر و قابل فهم‌تر باشد. افزون بر این، آن‌ها می‌توانند در کتاب‌های راهنمای معلمان، نکاتی را بگنجانند که اگر زبان‌آموز مشکل خاصی در یادگیری زبان دوم یا انجام یک فعالیت خاص داشت، معلم چگونه و با استفاده از چه راهبردی به او کمک کند. بایسته یادآوری است که هر چند محققان تمام تلاش خود را انجام دادند که این تحقیق به‌بهترین شکل ممکن انجام شود، اما به‌هر حال در این راه با محدودیت‌هایی هم مواجه شدند. از جمله این محدودیت‌ها می‌توان به موارد زیر اشاره کرد:

- از آنجا که دستیابی به شرکت‌کنندگان به‌صورت کاملاً تصادفی کار بسیار دشواری است، محققان مجبور به استفاده از مدل نمونه‌گیری در دسترس شدند.
  - این تحقیق با توجه به‌زمان کوتاهی که محققان در اختیار داشتند، تنها بر روی یک گروه زبانی، سطح مبتدی، انجام شد. همچنین به‌دلیل دسترسی محققان به‌زبان‌آموزان خانم، فقط این گروه در تحقیق حاضر شرکت کردند.
- با توجه به‌نتایج به‌دست آمده در این تحقیق که نشان داد ارزشیابی پویا روشی مناسب برای کمک به‌بهبود عملکرد زبان‌آموزان است، می‌توان تحقیقات دیگری هم در این زمینه انجام داد. از جمله پیشنهادهایی برای انجام تحقیقات آتی همسو با این تحقیق می‌توان موارد ذیر را یادآوری کرد:
- می‌توان تأثیر ارزشیابی پویا بر انواع دیگر فعالیت‌های درک مطلب خواندن و یا انواع فعالیت‌های مربوط به‌مهارت‌های دیگر از جمله مهارت‌های شنیداری، نوشتاری، و گفتاری را سنجید.
  - از آنجایی که ارزشیابی پویا مدل دیگری به‌نام مدل کنشگر دارد، می‌توان از این مدل به‌جای مدل میانجی در تحقیقات بعدی استفاده کرد. همچنین می‌توان تحقیقی انجام داد که در آن از هر دو مدل ارزشیابی پویا، مدل میانجی و کنشگر، در دو گروه مجزا استفاده کرد و میزان تأثیر هر مدل بر یادگیری زبان‌آموزان را سنجید.
- در مجموع، همسو با نتایج تحقیقات پیشین، نتایج این تحقیق نیز تأثیر بسزای ارزشیابی پویا در یادگیری زبان‌آموزان را نشان داد. این بدان معنی است که می‌بایستی از این نوع ارزشیابی

بیشتر در کلاس‌های زبان استفاده شود، تا ضمن فراهم آوردن محیطی مناسب برای یادگیری، زبان‌آموزان پیشرفت خود را نیز بطور واضح مشاهده نمایند.

#### ۶- منابع

- Alderson, J.C. (2000). *Assessing reading*. New York: Cambridge University Press.
- Alderson, J.C., Haapakangas, E.L., Huhta, A., Nieminen, L., and Ullakonoja, R. (2015). *The diagnosis of reading in a second or foreign language*. New York, NY 10017.
- Aljaafreh, A., and Lantolf, J.P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78(4), 465-483.
- ALTE. (1998). *Handbook of language examinations and examination systems: The code of fair testing practices in education*. Washington D.C.: Joint Committee on Testing Practices.
- Alvermann, D.E., and Commeyras, M. (1998). Feminist poststructuralist perspectives on the language of reading assessment: Authenticity and performance. In C. Harrison, & T. Salinger (Eds.), *Assessing reading: Theory and practice international perspectives on reading assessment* (pp.50-60). New York, NY 10001.
- Antón, M. (2003). *Dynamic assessment of advanced foreign language learners*. Paper presented at the American Association of Applied Linguistics (March). Washington, D.C., USA.
- Ardin, M. (2018). *The effect of diagnostic assessment vs. dynamic assessment on EFL learners' descriptive and narrative writing*. Unpublished MA Thesis, Islamic Azad University, Karaj, Iran.
- Best, J.W., and Kahn, J.V. (2006). *Research in education (10th ed.)*. Pearson Education Inc.
- Birjandi, P., & Najafi Sarem, S. (2012). Dynamic assessment (DA): An evolution of the current trends in language testing and assessment. *Theory and Practice in Language Studies*, 2(4), 747-753.
- Brown, H.D., and Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices (2nd ed.)*. Longman.
- Brown, J.D. (2005). *Testing in language programs (2nd ed.)*. Prentice Hall.
- Clarke, P.J., Truelove, E., Hulme, C., & Snowling, M.J. (2014). *Developing reading comprehension*. John Wiley & Sons, Ltd.
- Compernelle, R.A.V., and Zhang, H.S. (2014). Dynamic assessment of elicited imitation: A case analysis of an advanced L2 English speaker. *Language Testing*, 31(4), 395-412.

- Daftarifard, P., and Birjandi, P. (2017). Impact of mediation types on Iranian EFL learners' reading comprehension strategies. *Chamran Journal of Research in Applied Linguistics*, 8(1), 22-45.
- Goodman, K. (1988). The reading process. In P.L. Carrell, J. Devine, & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp.11-21). New York: Cambridge University Press.
- Grabe, W., and Jiang, X. (2014). Assessing reading. In A.J. Kunnan (Ed.), *The companion to language assessment. Volume I: Abilities, contexts, and learners* (pp.185-200). John Wiley & Sons, Inc.
- Grabe, W., and Stoller, S.L. (2001). *Teaching and researching reading*. Harlow: Longman.
- Green, A. (2014). Adapting or developing source material for listening and reading tests. In A.J. Kunnan (Ed.), *The companion to language assessment. Volume II: Approaches and development* (pp.292-308). John Wiley & Sons, Inc.
- Harding, L., Alderson, J.C., and Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing*, 32(3), 317-336.
- Haywood, H.C., and Lidz, C.S. (2007). *Dynamic assessment in practice clinical and educational applications*. Cambridge: Cambridge University Press.
- Hinton, P.R., Brownlow, C., McMurray, I., & Cozens, B. (2008). *SPSS explained*. New York, NY 10016.
- Horner, S. (1998). Assessing reading in the English national curriculum. In C. Harrison, & T. Salinger (Eds.), *Assessing reading: Theory and practice international perspectives on reading assessment* (pp.84-95). New York, NY 10001.
- Hosseini, S.F. (2015). *The effect of collaborative strategic reading on EFL learners' receptive skills*. Unpublished MA Thesis, Islamic Azad University, Karaj, Iran.
- Jang, E.E., and Wagner, M. (2014). Diagnostic feedback in the classroom. In A.J. Kunnan (Ed.), *The companion to language assessment. Volume II: Approaches and development* (pp.157-175). John Wiley & Sons, Inc.
- Kazemi, N. (2018). *The effect of dynamic vs. diagnostic assessment on EFL learners' speaking ability*. Unpublished MA Thesis, Islamic Azad University, Karaj, Iran.
- Kern, R.G. (1989). Second language reading strategy instruction: Its effects on comprehension and word inference ability. *The Modern Language Journal*, 73(2), 135-149.
- Lantolf, J.P., and Poehner, M.E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, 1(1), 49-72.
- Memari Hanjani, A. and Li, L. (2017). Cooperative learning pedagogy: A response to an urgent need in the Iranian EFL reading comprehension context. *Journal of Teaching Language Skills (JTLS)*, 36(3), 33-58.



- Nafarzadeh Nafari, F. (2016). *The effect of self-regulated feedback in summative, formative, and dynamic assessments on listening comprehension of intermediate EFL learners*. Unpublished MA Thesis, Islamic Azad University, Karaj, Iran.
- Oakhill, J., Cain, K., McCarthy, D., and Nightingale, Z. (2013). Making the link between the vocabulary knowledge and comprehension skill. In M.A. Britt, S.R. Goldman, & J.F. Rouet (Eds.), *Reading from words to multiple text* (pp.101-114). New York, NY 10017.
- Page, M.C., Braver, S.L., and Mackinnon, D.P. (2003). *Levine's guide to SPSS for analysis of variance (2nd ed.)*. Mahwah, New Jersey 07430.
- Pallant, J. (2005). *SPSS survival manual: A step by step guide to data analysis using SPSS for windows (Version 12)*. Crows Nest NSW 2065 Australia.
- Paris, S.G., Wasik, B.A., and Turner, J.C. (1991). The development of strategic readers. In R. Barr, M.L. Kamil, P.B. Mosenthal, & P.D. Pearson (Eds.), *Handbook of reading research* (pp.609-640). New Jersey: Lawrence Erlbaum Associates.
- Poehner, M.E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. University Park, PA.
- Poehner, M.E., and Lantolf, J.P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9, 233-265.
- Safari, M., and Mohaghegh Montazeri, M. (2017). The effect of reducing lexical and syntactic complexity of texts on reading comprehension. *Journal of Teaching Language Skills (JTLS)*, 36(3), 59-83.
- Shabani, K. (2011). Dynamic assessment of L2 learners' reading comprehension processes: A Vygotskian perspective. *Social and Behavioral Sciences*, 32, 321-328.
- Shabani, K. (2018). Group dynamic assessment of L2 learners' writing abilities. *Iranian Journal of Language Teaching Research*, 6(1), 129-149.
- Shabani, K., Khatib, M., and Ebadi, S. (2010). Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. *English Language Teaching*, 3(4), 237-248.
- Sheorey, R., and Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System*, 29, 431-49.
- Suvorov, R., and Hegelheimer, V. (2014). Computer-assisted language testing. In A.J. Kunnan (Ed.), *The companion to language assessment. Volume II: Approaches and development* (pp.61-80). John Wiley & Sons, Inc.
- Uso'-Juan, E., and Martinez-Flor, A. (2006). *Current trends in the development and teaching of the four language skills*. Wiley & Sons, Inc.
- Veen, C.V.D., Dobber, M., and Oers, B.V. (2016). Implementing dynamic assessment of vocabulary development as a triological learning process: A practice of teacher support in primary education schools. *Language Assessment Quarterly*, 13(4), 329-340.

Vosooghi, M. (2013). To examine authenticity factors in reading comprehension achievement tests in Iranian high schools through Coh-matrix2. *Roshd*, 28(2), 4-9.

Zandi, G. (2018). *The effect of diagnostic assessment vs. dynamic assessment on EFL learners' performance on selective and productive listening comprehension tasks*. Unpublished MA Thesis, Islamic Azad University, Karaj, Iran.