

مبانی فلسفی پژوهش در سنجش زبان انگلیسی در ایران: بررسی مقالات چاپ شده در نشریات علمی پژوهشی از ۲۰۰۸ تاکنون

کیومرث رضوی پور*

استادیار زبان‌شناسی کاربردی، دانشگاه شهید چمران اهواز،

اهواز، ایران

(تاریخ دریافت: ۹۷/۰۵/۰۷، تاریخ تصویب: ۹۸/۰۲/۱۵، تاریخ چاپ: آبان ۱۳۹۸)

چکیده

سنجش زبان، مانند دیگر شاخه‌های زبان‌شناسی کاربردی، همیشه متأثر از پارادایم‌های فلسفی زمانه است که عدم آگاهی به آنها ممکن است کیفیت پژوهش را تحت تاثیر قرار دهد. مقاله‌ی حاضر به بررسی مبانی فلسفی پژوهش در سنجش زبان انگلیسی در ایران می‌پردازد. بدین منظور همه‌ی مقالات قابل دسترس برخط که در نشریات علمی پژوهشی کشور در مورد سنجش زبان چاپ شده بودند مورد بررسی قرار گرفتند. برای تشخیص مبانی فلسفی مقالات، عنوان، چکیده، روش پژوهش و چارچوب نظری آنها مورد تحلیل و بررسی قرار گرفت. در گام بعدی، مقالات به چهار مقوله‌ی فلسفی اثبات‌گرایی، واقع‌گرایی، برساخت‌گرایی و ابزارانگاری تقسیم شدند. نتایج نشان داد که حدوداً دو سوم پژوهش‌ها در سنجش زبان در ایران دارای مبانی فلسفی اثبات‌گرایی و واقع‌گرایی می‌باشند. کمتر از یک سوم مقالات مبتنی بر مکتب برساخت‌گرا هستند و سهم پژوهش با مبانی ابزارانگاری تقریباً ناچیز است.

واژه‌های کلیدی: سنجش زبان، هستی‌شناسی، معرفت‌شناسی، اثبات‌گرایی، واقع‌گرایی، برساخت‌گرایی، ابزارانگاری.

۱- مقدمه

فلسفه، به معنای عام، به پرسش‌های بنیادی در باره‌ی هستی و چیستی پدیده‌ها می‌پردازد. در باب فلسفه علم، پرسش‌های فلسفی به دو دسته‌ی هستی‌شناسی^۱ و معرفت‌شناسی^۲ تقسیم می‌شوند. هستی‌شناسی به پرسش در مورد ماهیت پدیده‌ها می‌پردازد. به‌طور خاص، در علوم اجتماعی، اینکه آیا پدیده‌های اجتماعی وجود عینی دارند یا ساخته‌ی ذهن و زبان پژوهشگران هستند پرسش جدی هستی‌شناسی است. بر همین مبنا، معرفت‌شناسی به‌امکان و چگونگی دانش انسان در مورد پدیده‌ها می‌پردازد. قضاوت در مورد کیفیت مکانیسم سنجش در علوم تربیتی و آموزشی مستلزم ورود به مباحث فلسفی هستی‌شناسی و معرفت‌شناسی است (ماس، جیرارد و هانیفورد،^۳ ۲۰۰۶). در سنجش زبان، قبل از اقدام به طراحی و تولید ابزار و کیفیت سنجش توانش‌های زبانی، پرسش هستی‌شناسی این است که آیا اساساً آن ماهیتی که سنجشگر به دنبال اندازه‌گیری آن است وجودی مستقل از ذهن و زبان پژوهشگر دارد؟ اگر چنین است، ماهیت وجودی آن چیست؟ به‌عنوان مثال، تعریف کلاسیک روایی^۴ آزمون این است که آیا آزمون ماهیت مد نظر را می‌سنجد یا نه (نیوتن و شاو،^۵ ۲۰۱۴). بعضی از صاحب‌نظران معتقدند که چنین تعریفی از روایی آزمون مبتنی بر مبانی سنجش فیزیک نیوتن است که به‌وجود ماهیت‌های مستقل از سنجش باور دارند در حالیکه امروزه فیزیک کوانتوم چنین باوری را زیر سوال برده است (کانتلی،^۶ ۲۰۱۶). در فلسفه‌ی سنجش زبان، "ماهیت مد نظر سنجش" بدیهی قلمداد نمی‌شود بلکه پرسشی بنیادین و دامنه‌دار است که پاسخ دادن به آن مستلزم توسل به نظریه‌های فلسفه علم است (فالچر،^۷ ۲۰۱۵).

معرفت‌شناسی در سنجش آموزشی به‌صورت عام و در سنجش زبان به‌صورت خاص، به ارتباط میان ماهیت مورد سنجش و ابزار سنجش می‌پردازد (برسبوم،^۸ ۲۰۰۵؛ فالچر، ۲۰۰۹).

- 1- Ontology
- 2- Epistemology
- 3- Moss, P. A., Girard, B. J., & Haniford, L. C.
- 4- Validity
- 5- Newton, P., & Shaw, S.
- 6- Cantley, I.
- 7- Fulcher, G.
- 8- Borsboom, D.

میچل^۱، ۱۹۹۰: هود^۲، ۲۰۰۹). پرسش بنیادی معرفت‌شناسی سنجش در علوم انسانی و اجتماعی این است که آیا اصولاً ویژگی‌های باطنی و اجتماعی همچون سازه‌های فیزیکی از قبیل فاصله و وزن، شمارش‌پذیر و قابل بیان در قالب عدد و ارقامند؟ بعضی از محققان معتقدند که سنجش در حوزه‌ی آموزش و روان‌سنجی کلاً به بیراهه رفته است بهاین دلیل که کارشناسان این رشته قبل از آنکه ماهیت کمی سازه‌های باطنی و اجتماعی را مشخص نمایند به‌اندازه‌گیری و کمیت‌سازی پرداخته (میچل، ۱۹۹۷) و تجزیه و تحلیل آماری را معادل اندازه‌گیری و سنجش دانسته‌اند (مارکوس و بورسبوم^۳ (۲۰۱۳). کانتلی (۲۰۱۶) معتقد است که امروزه سنجش در آموزش و روان‌شناسی همچنان مبتنی بر مبانی سنجش فیزیک نیوتنی است، که بنیادهای معرفت‌شناسی آن توسط فیزیک کوانتوم به‌چالش جدی کشیده شده است.

لذا پرسش این است که آیا سنجش ماهیتی به‌نام توانش زبانی اساساً امکان‌پذیر است یا توهمی است که چون نهادینه شده است کسی در مورد آن پرسشگری نمی‌کند. معرفت‌شناسی سنجش آموزشی و روان‌شناسی، نفس عمل سنجش و اندازه‌گیری را به‌چالش می‌کشد. آیا به‌کمیت درآوردن پدیده‌های اجتماعی و انسانی امکان‌پذیر است؟ اگر چنین است، آیا روش سنجش در علوم اجتماعی باید همانند شیوه‌های سنجش در علوم طبیعی باشد و یا اگر غیر از این است، سنجش در علوم اجتماعی به‌طور عام، و در سنجش زبان به‌طور خاص، چگونه باید باشد؟ اینها همه پرسش‌هایی فلسفی است که نوع پاسخ به‌آنها جهت‌گیری و روش‌های پژوهش در سنجش زبان را تحت تاثیر قرار می‌دهد. تا آنجا که نویسنده واقف است، تاکنون در باب مبانی فلسفی سنجش زبان انگلیسی در ایران پژوهشی انجام نشده است. لذا امید است این کوشش فروتنانه در باب مبانی فلسفی سنجش زبان در ایران مورد اقبال پژوهشگران این حوزه قرار بگیرد.

لازم به توضیح است که در این نوشتار، مراد ما از سنجش زبان هر گونه تلاش نظام‌مند برای اندازه‌گیری توانش زبانی زبان‌آموزان در زبان انگلیسی به‌عنوان زبان خارجی می‌باشد. در همین راستا، هر پژوهشی که به‌پرسشگری پیرامون چیستی، چگونگی، و چرایی سنجش زبان انگلیسی به‌عنوان زبان خارجی بپردازد در شمول پژوهش در سنجش زبان قرار می‌گیرد

1- Michell, J.

2- Hood, S. B.

3- Markus, K. A., & Borsboom, D.

(دیویس^۱، ۲۰۰۸). در مقاله‌ی حاضر، ابتدا به شرح مکاتب عمده‌ی هستی‌شناختی و معرفت‌شناختی شامل اثبات‌گرایی، برساخت‌گرایی، و ابزارانگاری می‌پردازیم. سپس دلالت‌های هر یک را برای پژوهش در سنجش زبان مورد بررسی قرار خواهیم داد. در ادامه، روش انجام پژوهش حاضر را شرح می‌دهیم و در انتها یافته‌ها را گزارش کرده و آنها را به بحث خواهیم گذاشت. لازم به ذکر است، در این مقاله مراد ما از سنجش زبان، صرفاً سنجش در زبان انگلیسی است.

۲- اثبات‌گرایی

محققان ریشه‌های اثبات‌گرایی را به آثار قدمای یونانی و ایرانی مانند ارسطو، ابن سینا و فارابی نسبت می‌دهند (مول، ایریبارا و ویلسون^۲، ۲۰۱۵). در میان متأخرین، هیوم و جان لاک تنها روش پیشرفت علم را تجربه و حس می‌دانستند و لذا از آنها به عنوان پیشقراولان اثبات‌گرایی^۳ مدرن یاد می‌شود (پکر^۴، ۲۰۱۷). در این مکتب، مشاهده‌ی مستقیم تنها روش کسب دانش قلمداد می‌شود (مال و همکاران، ۲۰۱۵). اتحاد مکتب اثبات‌گرایی و منطق دکارتی در حلقه‌ی معروف وین، منجر به ظهور پارادایمی شد که امروزه از آن به عنوان اثبات‌گرایی منطقی^۵ یاد می‌شود (همان). این مکتب مبتنی بر دو عنصر منطق و تجربه می‌باشند. منطق همان نظریه‌ی علمی است و منظور از تجربه اندازه‌گیری و سنجش می‌باشد. وازه‌ها و مفاهیم نظریه‌های علمی باید طوری بیان شوند که قابل اندازه‌گیری باشند و این نیازمند تعریف عملیاتی^۶ مفاهیم محوری نظریه می‌باشد.

در اثبات‌گرایی منطقی^۷ گزاره‌ها به سه دسته تقسیم می‌شوند: گزاره‌های بدیهی مانند قضایای ریاضی، گزاره‌های تجربی، که درست یا نادرست بودنشان از روش‌های تجربی قابل بررسی است و دسته‌ی سوم گزاره‌هایی هستند که حاوی ارجاعات متافیزیک هستند و از طریق روش علمی نه قابل رد شدن هستند و نه قابل پذیرش. اثبات‌گرایان منطقی گزاره‌های دسته سوم

1- Davies, A.

2- Maul, A., Irribarra, D. T., & Wilson, M.

3- Positivism

4- Packer, M. J.

5- Logical positivism (empiricism)

6- Operational definition

7- Logical Positivism

را خارج از گستره علم دانسته و آنها را مردود می‌دانند (لیدیمن، ۲۰۰۷). از نظر حلقه‌ی وین، دلیل اینکه فیزیک نیوتن با آن همه شوکت و اقتدارش در مقابل نسبت انیشتین رنگ باخته بود استفاده‌ی نیوتن از مفاهیم متافیزیکی مانند جاذبه بود (پکر، ۲۰۱۷). لذا آنان چاره را در این می‌دیدند که زبان علم را از تعابیر متافیزیکی تنقیح نموده و منطق استقرا را محور کاوش‌های علمی قرار دهند. البته بعدها پوپر ضمن موافقت با کلیت پارادایم اثبات‌گرایی، روش قیاس را بر روش استقرا ترجیح داد (پوپر، ۲۰۱۴).

اگر چه خاستگاه اثبات‌گرایی علوم طبیعی بود، اما به حوزه‌های علوم اجتماعی از جمله زبان‌شناسی کاربردی نیز تسری یافت چنانکه امروزه بیشتر پژوهش‌های علوم اجتماعی در چارچوب همین پارادایم انجام می‌شوند. در ایران، لطف آبادی^۱ (۱۳۸۶) با بررسی پژوهش‌های رشته روانشناسی به این نتیجه رسید که اکثر قریب به اتفاق این مطالعات مبتنی بر مبانی اثبات‌گرایی هستند.

در مکتب اثبات‌گرایی منطقی، سازه‌های باطنی و درونی وجود خارجی ندارند. اثبات‌گرایان منطقی سازه^۲ را معادل روند منظم رفتاری می‌دانند. لذا آنها منکر این هستند که سازه‌ها به توانایی‌ها و ویژگی‌های واقعی غیرقابل مشاهده در باطن انسان دلالت دارند. بلکه سازه‌ها را عناوینی می‌دانند که توسط محقق برای مختصر سازی ساخته می‌شوند. افزون بر این، مفهوم علیت نیز در اثبات‌گرایی منطقی با معنای متعارف آن متفاوت است. اثبات‌گرایان منطقی علیت را به دو معنا تعبیر می‌کنند: یک تعبیر این است که علیت صرفاً یک رابطه همبستگی میان دو روند منظم است. به تعبیر دوم، علیت چیزی نیست غیر از ربطی که میان دو متغیر در ذهن پژوهشگر شکل می‌گیرد و ما به ازای خارجی واقعی ندارد (نوریس^۳، ۱۹۸۳). در مکتب اثبات‌گرایی منطقی، معیار درستی نظریه‌های علمی انسجام درونی آنهاست، نه اینکه تا چه حد معادل درستی برای واقعیت بیرونی هستند (همان). درستی نظریه‌های علمی از روش قیاس و فرضیه آزموده می‌شود. به بیان دیگر، بر اساس یک نظریه، مجموعه‌ای از پیش‌بینی‌ها انجام می‌شود و سپس هر کدام از این پیش‌بینی‌ها به روش تجربی آزموده می‌شود. هر چقدر که تعداد پیش‌بینی‌های صحیح مبتنی بر نظریه‌ای بیشتر باشد، اعتبار آن نظریه بیشتر است و بالعکس.

1- Lotfabadi, H.

2- Construct

3- Norris, S. P.

۳- اثبات‌گرایی و سنجش زبان

فلسفه‌ی اثبات‌گرایی راه خود را از طریق روانشناسی رفتارگرایی به‌علوم اجتماعی و زبانشناسی کاربردی باز کرد (بورسبوم، ۲۰۰۵). مفاهیم بسیار رایجی مانند تعریف عملیاتی که بعضاً ما هنوز هم به آنها مقید هستیم میراث فلسفه‌ی اثبات‌گرایی در روانشناسی هستند (همان). در واقع مراد از تعریف عملیاتی این است که پژوهشگران از کاربرد واژه‌های متافیزیک پرهیز کرده و واژه‌های علمی را به مصادیق قابل مشاهده آنها تقلیل دهند. اما در حوزه سنجش و بخصوص در مبحث روایی‌سنجی و اعتبار آزمون‌ها، بعضی از مبانی اثبات‌گرایی منطقی را می‌توان در مقاله معروف کروناخ و میل^۱ (۱۹۵۵) مشاهده نمود. آنان اعتبارسنجی ابزارهای سنجش را مبتنی بر شواهد تجربی و استدلال نظری می‌دانستند، که این دو همان دو عنصر اصلی اثبات‌گرایی منطقی هستند. در مقاله کروناخ و میل، از استدلال نظری به‌عنوان شبکه منطقی^۲ یاد می‌شود. اما تعبیر آنها از علیت با تعبیر رایج در اثبات‌گرایی منطقی متفاوت است. به عبارت دیگر، کروناخ و میل از دیدگاه هیومی علیت فاصله گرفته و به واقع‌گرایی گرایش پیدا می‌کنند (نوریس، ۱۹۸۳). همچنین تعبیر آنان از هستی‌شناسی سازه‌های باطنی نیز با تعبیر رایج در اثبات‌گرایی منطقی و روانشناسی رفتارگرایی متفاوت است به طوری که آنان عملکرد آزمون‌دهندگان را در آزمون، ناشی از فرایندهای ذهنی و باطنی آنان می‌دانند. چنانچه بیان شد، تجلی هستی‌شناسی اثبات‌گرایی در سنجش زبان را باید در آزمون‌هایی جست که طراحان آنها آبخور فکری رفتارگرایی داشته‌اند. لذا آزمون‌سازی مبتنی بر زبان‌شناسی ساختارگرایی، که متأثر از روان‌شناسی رفتارگرایی بود، را باید به لحاظ فلسفه علم، آزمون‌سازی مبتنی بر اثبات‌گرایی دانست. همچنین شیوع آزمون‌های گسسته و چندگزینه‌ای در دهه‌های اول قرن بیستم (ویر، ۲۰۰۵، ۳) را باید فرزند با واسطه اثبات‌گرایی دانست. از آنجا که در مکتب رفتارگرایی، یادگیری زبان به نشان دادن پاسخ زبانی مناسب در مقابل محرک زبانی خاص خلاصه می‌شد، در نتیجه سنجش یادگیری زبان نیز به‌ارزیابی گسسته و جداگانه هر یک از پاسخ‌های زبانی ختم می‌شد و بنابراین یکی از اصول غیرقابل انکار این مکتب آزمون‌سازی این است که هر پرسش فقط باید یک تکه جداگانه از رفتار زبانی را بسنجد (فرهادی، جعفرپور و بیرجندی،^۴ ۱۳۹۴). مکتب زبان‌شناسی ساختارگرا، که زبان را

1- Cronbach, L. J., & Meehl, P. E.

2- Nomological network

3- Weir, C. J.

4- Farhady, H., Jafarpur, A., & Birjandi, P.

قابل تجزیه به عناصر بسیار کوچک می‌دانست، پشتوانه نظری لازم را برای طراحی و تولید آزمون‌های گسترده فراهم می‌ساخت (ویر، ۲۰۰۵). افزون بر این، در آزمون‌سازی مبتنی بر فلسفه اثبات‌گرایی، دغدغه غالب سنجشگران زبان پایایی آزمون بود و نه روایی آن (همان). این اولویت‌بندی با مکتب کلاسیک سنجش نیز همخوانی داشت چرا که در این مکتب نیز روایی آزمون از اهمیت کمتری نسبت به پایایی آن برخوردار است (مارکوس و بورسبوم، ۲۰۱۳). همچنین پژوهش‌های مبتنی بر رویکرد سازه‌ها و روش‌های چندگانه^۱ نیز متأثر از مکتب اثبات‌گرایی قلمداد می‌شوند (مول و همکاران، ۲۰۱۶).

۴- واقع‌گرایی^۲

بر خلاف مکتب اثبات‌گرایی، که بالکل وجود مستقل سازه‌های غیرقابل مشاهده را انکار می‌کند، باور به وجود ماهیت‌های غیرقابل مشاهده از اصول محوری واقع‌گرایی است (جردن^۳، ۲۰۰۴). علاوه بر این، در واقع‌گرایی این فرض پذیرفته است که سازه‌های باطنی و اجتماعی وجودی مستقل از ذهن و زبان دارند (فالچر، ۲۰۱۴: جردن، ۲۰۰۴: مال و همکاران، ۲۰۱۵). بر این اساس، مفاهیم محوری مورد استفاده در نظریه‌های علمی صرفاً استعاره‌هایی تخیلی نبوده، بلکه دال بر ماهیت‌هایی موجود در دنیای واقعی هستند. در نتیجه، واقع‌گرایان به طبیعت‌گرایی^۴ در علوم اجتماعی باور دارند یعنی می‌پذیرند که به همان شیوه که طبیعت در علوم طبیعی مورد کاوش قرار می‌گیرد، می‌توان امور اجتماعی و انسانی را مورد مطالعه قرار داد (کارتر و نیو، ۲۰۰۴). آنان یکی دانستن نظریه‌های علمی با ماهیت پدیده‌های مورد مطالعه را مصداق مغالطه‌ی معرفت‌شناسی^۵ می‌دانند (همان).

۵- واقع‌گرایی و سنجش زبان

در زبان‌شناسی کاربردی، چامسکی اولین کسی بود که پس از دهه‌ها استیلاهی رفتارگرایی و عملیات‌گرایی^۶، باور به قابلیت‌های زبانی غیرقابل مشاهده را رواج داد (جردن، ۲۰۰۴). نتیجه

1- Multi-trait-multi-method

2- Realism

3- Jordan, G.

4- Naturalism

5- Epistemic fallacy

6- Operationalism

منطقی باور به مبانی واقع‌گرایی در سنجش زبان این است که توانش‌های ارتباطی و زبانی مورد سنجش دارای ماهیت مستقل و پایا می‌باشند که وجودشان منوط به نظریه‌ی خاصی نیست (فالچر، ۲۰۱۵). واقع‌گرایان در سنجش معتقدند که میان واریانس موجود در نمرات مبتنی بر یک آزمون و تغییرات در مقادیر ماهیت مورد سنجش رابطه‌ی علی معلولی وجود دارد. با این وصف، مدل‌های توانش زبانی چامسکی، مدل توانش ارتباطی بکمن^۱ (۱۹۹۰) و مدل گالاکسی و تیلور^۲ (۲۰۱۸) همگی مبتنی بر مفروضات واقع‌گرایی هستند. امروزه عمده‌ی آزمون‌هایی که تحت عنوان آزمون انگلیسی برای اهداف خاص^۳ مانند آزمون زبان هوانوردی ساخته می‌شوند، متکی به گزاره‌های هستی‌شناسی واقع‌گرا هستند (فالچر، ۲۰۱۳). همچنین، مطالعات مبتنی بر تحلیل عاملی و معادلات ساختاری، آزمون‌های اختلالات ارتباطی و آزمون‌های تشخیصی همه بر این فرض مبتنی هستند که ماهیت مورد سنجش وجودی مستقل از ابزار سنجش دارد (فالچر، ۲۰۱۳؛ کانتلی، ۲۰۱۶؛ مارکوس و بورسیوم، ۲۰۱۳).

۶- برساخت‌گرایی^۴

کسانی که پیرو مکتب هستی‌شناختی برساخت‌گرایی هستند خود به دو نحله تقسیم می‌شوند. دسته‌ی اول هستی و وجود پدیده‌های اجتماعی را در عالم بیرون از ذهن محقق انکار می‌کنند. دسته‌ی دوم کسانی هستند که منکر وجود هستی و ذات ماهیت‌های مستقل از ذهنیات آدمی نیستند اما دسترسی و امکان کشف آنها را غیرممکن می‌دانند. به عبارت دیگر دسته‌ی دوم هستی‌شناسی واقع‌گرا را با معرفت‌شناسی برساخت‌گرایی ترکیب می‌نمایند. استعاره‌ی غار افلاطون و پیل در تاریکی مولانا هر دو استعاره‌هایی هستند مبنی بر اینکه دست بشر هرگز به اسرار موجود بر نخیل هستی نخواهد رسید. ویتگنشتاین اگر چه دانش را جلوه‌ایزبانی از واقعیت می‌دانست اما بعدها به این باور رسید که وقتی میان واژه‌ها و مصادیق‌شان در عالم واقع نسبتی واقعی وجود ندارد، بر همین مبنا، میان ترکیب واژه‌ها (گزاره‌های علمی)، و ترکیب مصادیق (پدیده‌ها) هم نسبتی نیست (کورسن^۵، ۱۹۹۷). فالچر (۲۰۱۳) معتقد است که هستی‌شناختی برساخت‌گرایی غالباً منجر به خوانشی بدبینانه از مسایل و شرایط اجتماعی

1- Bachman, L. F.

2- Galaczi, E., & Taylor, L.

3- English for Specific Purposes (ESP)

4- Constructivism

5- Corson, D.

می‌شود. همچنین، از آنجا که در مکتب برساخت‌گرایی، همه روایت‌ها و خوانش‌های علمی صرفاً زاده نگاه پژوهشگر قلمداد می‌شوند، برساخت‌گرایان معتقدند که شرایط موجود می‌شود غیر از آنی که هست باشد. این البته مرحله اول در هستی‌شناختی برساخت‌گرا قلمداد می‌شود. مراحل سه‌گانه برساخت‌گرایی به شرح زیر هستند (هکینگ^۱، ۱۹۹۹):

الف) وضعیت موجود غیرقابل اجتناب نیست. یعنی می‌شد که غیر از این باشد.

ب) وضعیت موجود وضعیتی نامطلوب است.

ج) وضعیت موجود باید تغییر کند.

بعضی از پیروان این مکتب در سطح اول می‌مانند و بسنده می‌کنند به اینکه واقعیت همیشه برآیند شرایط اجتماعی سیاسی است و ماهیتی واقعی ندارد. برساخت‌گرایان افراطی‌تر پا را فراتر گذاشته و همیشه وضعیت موجود را در قیاس با وضعیت مطلوب، نامطلوب می‌بینند. در آخرین مرتبه از نظریه برساخت‌گرایی، برساخت‌گرایان افراطی در صدد تغییر شرایط موجود برمی‌آیند. به همین دلیل گفته می‌شود که یکی از مزایای هستی‌شناسی مبتنی بر برساخت‌گرایی این است که آزادی‌بخش است و با توسل به آن همیشه می‌توان ساختارهای حاکم را مورد تشکیک قرار داد و نهایتاً برانداخت. البته همگان با این ظرفیت آزادی‌بخشی مکتب برساخت‌گرایی موافق نیستند. هکینگ (۱۹۹۹) معتقد است که این پتانسیل آزادی‌بخشی برساخت‌گرایی تنها برای کسانی مفید است که خود در آستانه آزادی قرار داشته باشند. همچنین، پوپر از معرفت‌شناسی برساخت‌گرا به‌عنوان معرفت‌شناسی بدبینانه^۲ یاد می‌کند و معتقد است که به لحاظ فلسفه سیاسی نه تنها به آزادی و دموکراسی نمی‌انجامد، بلکه خود منجر به زایش نظام‌های انحصارطلب و اقتدارگرا می‌شود (پوپر^۳، ۲۰۱۴).

در مکتب برساخت‌گرایی، از آنجا که گزاره‌های علمی در ارتباط تنگاتنگ و همیشگی با ایدئولوژی قدرت حاکم می‌باشند (شوهامی^۴، ۲۰۱۴)، رد پای قدرت حاکم در همه گزاره‌های علمی مشهود است و لذا تحت یک ایدئولوژی متفاوت، همه گزاره‌های علمی هم صورت و محتوایی کاملاً متفاوت می‌داشتند. بنابراین، در مکتب برساخت‌گرایی، گزاره‌های علمی را هرگز نمی‌توان گزاره‌هایی عینی و خالی از جانبداری دانست (نادینگز^۵، ۲۰۱۰).

1- Hacking, I.

2- Pessimistic epistemology

3- Popper, K.

4- Shohamy, E.

5- Noddings, N.

۷- برساخت‌گرایی و سنجش زبان

با این اوصاف، از نظرگاه هستی‌شناسی برساخت‌گرا، متغیرهایی که سنجشگران به دنبال سنجش آنها هستند، ماهیتی مستقل از اذهان پژوهشگران نداشته و صرفاً زائیده قراردادهای اجتماعی و زبانی هستند. در افقی نازل‌تر از برساخت‌گرایی، اگر حتی متغیرهای مذکور ماهیت مستقل داشته باشند، میان واژه‌ها و باطنیات فردی و اجتماعی تناظری وجود ندارد. در حالت اول، برساخت‌گرایان با کسانی هم‌پاله می‌شوند که در نگاه اول، شاید بیشترین فاصله را با آنان دارند: رفتارگرایان کلاسیک. همانند برساخت‌گرایان نوع اول، رفتارگرایان کلاسیک هم بالکل منکر وجود سازه‌های درونی، روحی و روانی بودند و همه‌چیز را در رفتار می‌دیدند. بر این اساس، ابزارهای اندازه‌گیری در علوم انسانی و اجتماعی، مصنوعاتی قراردادی‌اند که خود مبتنی بر مفاهیمی قراردادی هستند. همچنین، مفاهیم قراردادی یادشده برآمده از نظریه‌هایی هستند که نسبتی با ماهیت واقعی باطن انسان و اجتماع، که هستی آنها هم زیر سوال است، ندارند. از این رو، روایت‌هایی که در اثبات روایی و اعتبار آزمون‌ها ارایه می‌شوند فقط یک روایت از هزاران روایت ممکن هستند. بنابراین، اساساً این ادعا که ابزاری برای سنجش و کمیت‌بخشی به باطنیات، دانسته‌ها و نگرش آدم‌ها وجود دارد از نظرگاه معرفت‌شناسی برساخت‌گرایی باطل است. از آنجا که فی الجمله منکر وجود خارجی سازه‌هایی هستند که آزمون‌سازان ادعای اندازه‌گیری‌شان را دارند، پژوهشگران برساخت‌گرا مأموریت خود را ساخت آزمون‌هایی با روایی بهتر نمی‌دانند چرا که با اصل تعریف کلاسیک روایی، مبتنی بر اینکه آزمون روا آزمونی است که بسنجد آنچه را قرار است بسنجد، مخالف جدی دارند (فالچر، ۲۰۱۳). پس به جای ارایه‌ی آلت‌رناتیو و ساخت آزمون‌های باکیفیت بهتر، برساخت‌گرایان ترجیح می‌دهند منتقد همیشگی وضعیت موجود بمانند. به طور خاص، یکی از دغدغه‌هایشان پرده برداری و افشاکردن تبعیض‌هایی است که عمداً یا تصادفاً از طریق آزمون‌ها علیه اقلیت‌ها صورت می‌گیرد (مک نامارا و روور، ۲۰۰۶). برساخت‌گرایان معتقدند که آزمون‌ها به ابزارهایی سهل‌الوصول و مقرون‌به‌صرفه برای اعمال زور و فشار به لایه‌های فرودست جامعه تبدیل شده‌اند. همچنین، برساخت‌گرایان باور دارند که آزمون‌ها، اگرچه در ابتدا به نیت تحقق عدالت اجتماعی و شایسته‌سالاری ابداع شدند، اما خود به تدریج به مانعی سر راه تحقق دموکراسی و عدالت اجتماعی تبدیل شده‌اند، چرا که عموماً طراحی و ساخت آزمون‌ها

در دست دولت‌ها، نهادها و یا کمپانی‌های قدرتمند است و لذا محتوا و روش‌های سنجش و آزمون همیشه در راستای اولویت‌ها و منافع قدرتمندان است. برساخت‌گرایان معتقدند که اصحاب قدرت با اعمال سانسور در محتوای آزمون‌ها به یک نوع نگرش خاص مشروعیت بخشیده، از دیگر اندیشه‌ها مشروعیت‌زدایی کرده، و به صورت نامرئی آزادی بیان و اندیشه را محدود می‌نمایند (فوکو^۱، ۱۹۷۷؛ شوهای، ۲۰۱۴). بر همین اساس، برساخت‌گرایان آزمون را به زندان پناپتیکن^۲ شبیه می‌دانند که در آن زندانیان بدون نیاز به نگهبان به همه خواست‌های زندان‌بان پایبند می‌مانند (فالچر، ۲۰۱۰). با این اوصاف، برساخت‌گرایان خود را مکلف به افزایش سطح آگاهی عمومی جامعه در باره مقاصد مستتر در ورای آزمون‌ها و امتحان‌های گوناگون می‌دانند. آنان همگان را ترغیب می‌کنند به آزمون‌ها همواره به دیده تردید و بدبینی بنگرند و مقهور جلوه‌های آماری و عددی آنها نشوند (شوهای، ۲۰۱۴). از نظر برساخت‌گرایان، حتی آزمون‌های کلاسی هم مبرا از روابط قدرت نیستند. به دیگر بیان، حتی معلمان نیز از امتحان به عنوان اهرمی برای اعمال فشار، تهدید، و کنترل استفاده می‌کنند. میشل فوکو (۱۹۷۷) حتی از این هم فراتر رفته و معتقد است که کل تکنولوژی آزمون بخشی از زندانی است که توسط قدرتمندان برای سرکوب و محاصره انسان معاصر ساخته شده است.

۸- ابزارانگاری^۳

در مکتب فلسفی ابزارانگاری^۴، پرسش مربوط به هستی‌شناختی پدیده‌ها بی‌پاسخ می‌ماند چرا که ابزارگرایان معتقدند اصولاً اینکه واژه‌ها و مفاهیم استفاده‌شده در نظریه‌های علمی به ماهیت‌های واقعی خارج از ذهن دلالت دارند یا نه دغدغه‌ای مهم نیست (مول، و همکاران، ۲۰۱۶). فیلسوفانی مانند پیرس، جیمز، رورتی، پوتنم و تولمین را پیشگامان این مکتب فلسفی می‌دانند (همان). در واقع، در نگاه ابزارانگاران به علم، رسالت علم نه توصیف جهان هستی بلکه تغییر دادن آن در راستای مطلوب است. بنابراین، اگر نظریه‌ای علی‌رغم اینکه توصیف نادرستی از پدیده مورد نظر است اما پیروی از آن منجر به تغییرات مطلوبی شود، نظریه سودمندی قلمداد شده و برتر از نظریه‌ای است که از صحت هستی‌شناختی برخوردار است اما

1- Foucault, M.

2- Panopticon Prison

3- Pragmatism

4- Instrumentalism

پیامدهای سودمندی بر آن مترتب نیست. گفته می‌شود همین نگاه پراگماتیسم به هستی‌شناختی و شناخت‌شناسی است که باعث شده بعد از دهه‌ها جدال میان دو پارادایم پژوهشی کیفی و کمی، امروزه روش‌های پژوهشی کمی و کیفی به موازات هم در پژوهش‌های مختلط مورد استفاده قرار بگیرند (مکسول و میتاپالی^۱، ۲۰۱۰).

۹- ابزارانگاری و سنجش زبان

چنانکه گفته شد، ابزارانگاری فلسفی نسبت به این پرسش هستی‌شناسی که آیا سازه‌هایی مانند توانش ارتباطی^۲، و یا روانی زبانی^۳ ماهیت واقعی مستقل از گفتمان پژوهشگران دارند یا نه بی‌طرف می‌مانند، یعنی نه تایید می‌کنند و نه انکار. نمود مکتب فلسفی ابزارانگاری در ادبیات روان‌سنجی، سنجش آموزش و سنجش زبان را می‌توان در رویکرد روایی اتخاذ شده در منابعی همچون مسیک^۴ (۱۹۸۹)، کین^۵ (۲۰۱۳)، و بکمن (۲۰۱۰) یافت. در ماتریس معروف روایی مسیک، عواقب و تبعات تصمیمات مبتنی بر آزمون، بخش مهمی از شواهد لازم برای روایی یک آزمون به شمار می‌آیند. بر این اساس، روایی نه خاصیت خود آزمون، بلکه خاصیت کاربردها و اقداماتی است که بر اساس آزمون در پیش گرفته می‌شوند. رویکرد استدلال‌محور کین به روایی آزمون‌ها، که مبتنی بر ساختار استدلالی تولمین می‌باشد، مظهر تمام عیار فلسفه‌ی پراگماتیسم در روان‌سنجی و سنجش آموزش است. در سنجش زبان، برای قضاوت در مورد کیفیت آزمون‌های زبان، بکمن و پالمر^۶ (۱۹۹۶) چارچوبی تحت عنوان سودمندی آزمون^۷ ارائه دادند که باز جلوه‌های پراگماتیسم آن را در انتخاب صفت سودمندی می‌توان دید. بر اساس نظریه‌ی استدلال‌گرایی روایی، روایی تفاسیر و کاربردهای مبتنی بر آزمون میزان باورپذیری استدلالی است که سازندگان یا کاربران برای تفاسیر یا کاربردهای مذکور ارائه می‌دهند (بکمن، ۲۰۰۵: چپله^۸، ۲۰۱۲). ملهم از فلسفه‌ی استدلال تولمین^۹ (۲۰۰۲)، در روایی استدلال‌گرا،

1- Maxwell, J. A., & Mittapalli, K.

2- Communicative competence

3- Fluency

4- Messick, S.

5- Kane, M. T.

6- Bachman, L. F., & Palmer, A. S.

7- Test Usefulness

8- Chappelle, C. A.

9- Toulmin, S. E. (2002).

روایی از زنجیره‌ای از استدلال‌ها تشکیل می‌شود که هر حلقه از زنجیره‌ی مذکور به عنوان داده‌ی حلقه‌ی بعدی استدلال تلقی می‌شود و لذا ضعف هر حلقه منجر به شکست و پارگی کل زنجیره‌ی استدلال می‌شود (زیامونگ^۱، ۲۰۱۰). امروزه بسیاری از پژوهشگران سنجش زبان، به رویکرد استدلال‌گرایی روی آوردند. چپله، انریخت و جمیسون^۲ (۲۰۰۸) یکی از مفصل‌ترین پژوهش‌های پیشگام استدلال‌گرا را در مورد نسل جدید آزمون تافل انجام دادند. طی دو دهه‌ی اخیر تعداد مطالعات با رویکرد استدلال‌گرایی رو به افزایش بوده است (ناچ و چپله^۳، ۲۰۱۸؛ جانسون و ریاضی^۴، ۲۰۱۵؛ لافلر و استاپلز^۵، ۲۰۱۷).

۱۰- پژوهش در مبانی فلسفی سنجش زبان

اگرچه در دل هر پژوهشی در سنجش زبان نگاهی فلسفی مستتر است، اما پژوهش‌هایی که به بررسی مبانی فلسفی پژوهش در سنجش زبان پرداخته باشند زیاد نیستند. به نظر می‌رسد این فقدان توجه به مبانی فلسفی سنجش در مورد سنجش آموزش به صورت عام و روان‌سنجی نیز مصداق دارد (مول و همکاران، ۲۰۱۶). مک‌نامارا و روور (۲۰۰۶) معتقدند که در قیاس با سایر شاخه‌های زبان شناسی کاربردی، اقبال به مبانی اثبات‌گرایی در سنجش زبان بیشتر است. آنان این گرایش فلسفی به اثبات‌گرایی را ناشی از اتکای تاریخی سنجش زبان به رشته‌ی روان‌سنجی^۶ می‌دانند. فالچر و دیویدسون^۷ (۲۰۰۷) اما معتقدند که بعد از دهه‌ی هشتاد میلادی، پژوهش در مورد روایی آزمون‌های زبان از مبانی اثبات‌گرایی فاصله گرفته است. لازم به ذکر است که نه ادعای مک‌نامارا و روور و نه نتیجه‌گیری فالچر و دیویدسون هیچکدام مبتنی بر داده‌های میدانی نیستند. بنابراین، می‌توان ادعا کرد که پیرامون پارادیم‌های فلسفی پژوهش در سنجش زبان دامنه‌ی کارهای انجام شده چندان وسیع نیست. یکی از این محدود آثار پژوهشی کاری بود که چنگ و فاکس^۸ در مورد رساله‌های دکتری نوشته شده در دانشگاه‌های کانادا انجام دادند. بررسی روش

1- Xiaoming, X.

2- Chapelle, C. A., Enright, M. K., & Jamieson, J. M.

3- Knoch, U., & Chapelle, C. A.

4- Johnson, R. C., & Riazi, A. M.

5- LaFlair, G. T., & Staples, S.

6- Psychometrics

7- Fulcher, G., & Davidson, F.

8- Cheng, L., & Fox, J.

تحقیق تعداد ۲۴ رساله‌ی دکتری سنجش زبان در کانادا توسط چنگ و فاکس (۲۰۱۳) نشان داد که رویکرد غالب این رساله‌ها مکتب ابزارانگاری بوده و روش تحقیق مختلط کمی و کیفی در آنها رواج چشمگیری داشته است. اما در باب مقالات پژوهشی، همپلاینز و لینچ^۱ (۱۹۹۸) چکیده‌های مقالات کنفرانس معروف ال تی آرسی را بررسی کردند. آنان با تقسیم‌بندی چکیده‌ها به سه دسته‌ی شدیداً اثبات‌گرا، دسته‌ی نسبتاً برساخت‌گرا، و دسته‌ی مبهم به این نتیجه رسیدند که عمده‌ی تحقیقات پیرامون روایی آزمون‌های زبان در پارادایم اثبات‌گرایی انجام شده است. بیش از دو دهه از تاریخ چاپ مقاله‌ی همپلاینز و لینچ و سه چهار دهه از تاریخ چاپ چکیده‌های تحلیل شده در آن می‌گذرد و طی این مدت مبانی معرفت‌شناسی علوم اجتماعی دچار دگرگونی‌هایی شده که قاعدتاً سنجش زبان هم از آن مستثنی نبوده است. افزون بر این، از آنجا که اکثریت غالب چکیده‌های تحلیل شده در کار مذکور توسط پژوهشگران اروپایی و آمریکایی نوشته شده بودند و نظر به اینکه مبانی فکری و فلسفی تحت تاثیر جغرافیای طبیعی و فرهنگی می‌باشند (نیسبت، ۲۰۰۴)، لذا مطالعه‌ی مبانی فلسفی پژوهش در سنجش زبان در ایران می‌تواند تصویر کلی راهگشایی در مورد بنیان‌های فکری پژوهش در سنجش زبان به دست دهد.

هدف پژوهش حاضر تعیین میزان رواج هر یک از مکاتب فلسفی یادشده در مطالعات مربوط به سنجش زبان انگلیسی در ایران می‌باشد. به عبارت دیگر، پرسش این است که تا چه اندازه پژوهش در حوزه‌ی آزمون‌سازی و سنجش زبان متأثر از مکاتب فلسفی اثبات‌گرایی، واقع‌گرایی، برساخت‌گرایی، و ابزارگرایی است. لازم به ذکر و تاکید است که هدف این مطالعه، قضاوت در مورد مبانی فلسفی پژوهشگران نیست چرا که اولاً، مبانی فلسفی یک پژوهشگر در ظرف زمان و مکان متغیرند (نیسبت، ۲۰۰۴) و دوم آنکه، انجام کار پژوهشی در یک پارادایم فلسفی مانع کار در دیگر پارادایم‌ها نیست.

۱۱- روش پژوهش

از آنجا که هدف این پژوهش، بررسی مبانی فلسفی پژوهش در سنجش زبان در ایران بود، پیکره‌ی لازم برای انجام این پژوهش از مجلات علمی-پژوهشی دانشگاه‌های دولتی ایران

1- Hamp-Lyons, L., & Lynch, B. K.

2- Nisbett, R.

گردآوری شد. بدین منظور همه‌ی نشریات علمی- پژوهشی مربوط به آموزش زبان و زبان‌شناسی کاربردی مورد بررسی قرار گرفتند (جدول ۱). لازم به ذکر است که به دلیل دشواری دسترسی به نسخه‌های کاغذی مجلات، صرفاً شماره‌های آنلاین مجلات مذکور در شمول پیکره‌ی فعلی قرار گرفتند. مقالات بررسی شده در بازه‌ی زمانی ۲۰۰۸ تا ژوئیه ۲۰۱۸ بودند و لذا مقالاتی که پس از این تاریخ در وبسایت نشریات فوق بارگزاری شده‌اند در شمول پیکره‌ی این مطالعه قرار نگرفته‌اند. در راستای رعایت موازین اخلاقی پژوهش، از عنوان مقالات یا نویسندگان آنها نامی برده نمی‌شود.

پس از انتخاب نشریات مورد نظر، فهرست مطالب هر یک از شماره‌های نشریات یادشده مورد مطالعه قرار گرفت و مقالات مربوط به سنجش و ارزیابی سنجش زبان شناسایی و استخراج شدند. استفاده از آزمون‌های زبان در پژوهش‌های زبان‌شناسی بسیار رایج است اما صرف استفاده از یک آزمون زبان کافی نیست تا پژوهشی را در شمول رشته‌ی آزمون‌سازی و سنجش زبان قلمداد کرد. در این پژوهش، تنها مقالاتی که پرسش اصلی شان ساخت و یا ارزیابی ویژگی‌های آزمون‌های زبان بودند مورد واکاوی قرار گرفتند. بنابراین مقالاتی که از آزمون‌های زبان صرفاً به عنوان ابزار گردآوری داده استفاده شده بود و آزمون سوژه‌ی اصلی پژوهش نبود در پیکره‌ی مد نظر قرار نگرفتند. همچنین در مواردی که در مورد مقاله‌ایتريد وجود داشت با رجوع به تقسیم بندی بنیاد بین‌المللی سنجش زبان^۱ تصمیم نهایی گرفته می‌شد. بنابراین، مقالاتی که در مورد ساخت یا روایی‌سازی ابزارهایی برای سنجش ویژگی‌هایی غیر از توانش زبانی مانند انگیزه، استرس و مانند آن بودند در شمول پیکره‌ی این پژوهش قرار نگرفتند. همچنین، مقالاتی که در مورد سنجش زبان‌هایی غیر از زبان انگلیسی بودند نیز از فهرست مقالات مورد نظر خارج شدند. افزون بر این، شماره‌های اختصاص داده شده به موضوعات خاص نیز، به دلیل روند داوری متفاوت شان (لازارتان^۲، ۲۰۰۵) در شمول این پژوهش قرار نگرفتند. از آنجا که موضوع حاضر محدود به سنجش زبان در ایران می‌باشد، یک مقاله که توسط پژوهشگری از کره‌ی جنوبی در مورد سنجش زبان در کره نوشته شده بود نیز از زمره‌ی مقالات مورد بررسی خارج شد. در نهایت، همه‌ی مقالات انتخاب شده با فرمت پی دی اف برای بررسی‌های بعدی ذخیره شدند.

1- International Language Testing Association (ILTA)

2- Lazaratan, A.

جدول شماره ۱ عنوان مجلات و تعداد مقالات مستخرج از هر یک را نشان می‌دهد. در مجموع طی یک‌دهه‌ی اخیر، تعداد مقالات چاپ شده در این نه نشریه ۱۰۹۵ مقاله است که تعداد ۷۸ مقاله مربوط به حوزه‌ی سنجش زبان هستند. به عبارت دیگر، ۷٪ کل مقالات چاپ شده در نشریات دانشگاهی ایران به موضوع سنجش زبان اختصاص دارد. در این میان، نشریه‌ی دانشگاه خوارزمی با اختصاص ۱۶٪ مقالات به سنجش زبان بیشترین سهم و نشریه‌ی دانشگاه تهران با اختصاص ۳٪ مقالات به سنجش زبان، کمترین سهم را در چاپ مقالات سنجش زبان داشته است. لازم به ذکر است که در میان نه مجله‌ی بررسی شده، نشریه‌ی دانشگاه تهران تنها نشریه‌ای است که به زبان فارسی چاپ می‌شود و شاید هم به همین دلیل باشد که بسامد نشر مقالات مربوط به سنجش زبان در آن پایین است.

یکی از دشواری‌های پژوهش‌های فلسفه‌ی علم این است که غالباً پژوهشگران از مبانی و مفروضات فلسفی استفاده‌شده در آثار خود سخنی به صراحت به میان نمی‌آورند که شاید به این دلیل باشد که خود پژوهشگران نیز الزاماً به این مبانی خودآگاهی کامل ندارند. لذا چاره‌ای نیست جز اینکه محقق، این مبانی را در روش‌شناسی پژوهش و انتخاب چارچوب نظری آنها جستجو نماید. در پژوهش حاضر، برای تعیین مبانی فلسفی مقالات، از تقسیم‌بندی فالچر (۲۰۱۳) بهره جستیم. دلیل این انتخاب این بود که در واقع، تقسیم‌بندی فالچر تنها مدل موجود در مورد فلسفه‌ی سنجش زبان می‌باشد. برای تعیین رویکرد فلسفی مقالات اما به موضوعاتی برخوردیم که در تقسیم‌بندی فالچر به آنها تصریحی نشده بود. به دیگر بیان، مدل فالچر از جامعیت لازم برخوردار نبود و لذا مقوله‌های جدیدی به آن افزودیم.

طبق جدول شماره ۲، که مبتنی بر تقسیم‌بندی فالچر (۲۰۱۳) است، پژوهش‌های سنجش زبان را می‌توان به سه دسته تقسیم کرد: پژوهش‌های مبتنی بر واقع‌گرایی، پژوهش‌های دارای مبانی برساخت‌گرا، و پژوهش‌های مبتنی بر ابزارانگاری. بر این اساس، پژوهش‌های مرتبط با ساخت آزمون برای اهداف خاص، آزمون‌های تشخیصی، مطالعات مبتنی بر تحلیل عاملی اکتشافی و تاییدی، و مقالات مربوط به شیوه‌های نمره دهی دارای مفروضات و مبانی فلسفی واقع‌گرا هستند. از طرف دیگر، پژوهش‌های دارای مبانی برساخت‌گرا عبارتند از مطالعات مربوط به نقش مخاطب در آزمون‌های شفاهی، سنجش منظورشناسی، انصاف و تبعیض در آزمون‌های زبان، کارکردهای پنهان آزمون‌های زبان، و ارزیابی پویا. همچنین طبق جدول شماره ۲، پژوهش‌هایی که شرح فرایند آزمون‌سازی را به عنوان شواهد روایی آزمون تلقی کرده و همچنین مطالعاتی که روایی را مبتنی بر استدلال می‌دانند دارای مبانی ابزارانگاری هستند.

جدول شماره ۱. نام مجلات، محل نشر، و تعداد مقالات چاپ شده در حوزه‌ی سنجش زبان

ردیف	عنوان مجله	ناشر	تعداد مقالات چاپ شده	تعداد مقالات مربوط به سنجش زبان	نسبت مقالات به سنجش زبان به کل مقالات
۱	Journal of Teaching Language Skills (JTLS)	دانشگاه شیراز	۲۰۴	۱۱	۵٪
۲	Journal of English Language Teaching and Learning (JELTL)	دانشگاه تبریز	۱۲۴	۱۰	۸٪
۳	Applied Research in English (ARE)	دانشگاه اصفهان	۹۰	۱۰	۱۱٪
۴	Foreign Language Research Journal (FLRJ)	دانشگاه تهران	۱۲۷	۴	۳٪
۵	Research in Applied Linguistics (RALs)	دانشگاه شهید چمران اهواز	۱۲۱	۷	۶٪
۶	Iranian Journal of Applied Language Studies (IJALS)	دانشگاه سیستان و بوچستان	۱۴۷	۱۱	۷٪
۷	Iranian Journal of Applied Linguistics (IJAL)	دانشگاه خوارزمی	۱۳۰	۱۶	۱۲٪
۸	Iranian Journal of Language Teaching Research (IJLTR)	دانشگاه ارومیه	۸۵	۵	۶٪
۹	Issues in Language Teaching (ILT)	دانشگاه علامه طباطبایی	۶۷	۴	۶٪
	مجموع		۱۰۹۵	۸۷	۷٪

برای تعیین مبانی فلسفی مقالات، ابتدا با توجه به عنوان هر مقاله و با توجه به اینکه در کدام یک از سه مقوله‌ی ذکر شده در جدول ۲ می‌گنجد، مقالات توسط نویسنده کدگذاری شدند. سپس چکیده‌ی مقاله را مورد بررسی قرار می‌دادیم. در صورتی که خوانش چکیده مویید کد تعیین شده بود، کد اولیه تایید می‌شد. اما در صورتی که میان دو کد داده شده، یکی بر اساس عنوان و دیگری با خواندن چکیده، همخوانی وجود نداشت، کل مقاله و بخصوص بخش چارچوب نظری مقاله و متدلوژی آن مورد بررسی قرار می‌گرفت. در این مرحله، متوجه شدیم که نقدی که بر مطالعات مشابه قبلی (لینچ و همپلایز، ۱۹۹۸، لطف آبادی، ۱۳۸۶) از جمله بر تقسیم بندی فالچر، وارد است این است که آنان اثبات‌گرایی و واقع‌گرایی را با هم مترادف گرفته‌اند. به عبارت دیگر، هم لطف‌آبادی و هم لینچ و همپلایز همه‌ی مطالعات کمی را اثبات‌گرا قلمداد کرده‌اند. همچنین، فالچر همه‌ی مطالعات صرفاً کمی را دارای مبانی واقع‌گرا می‌داند. مترادف دانستن اثبات‌گرایی و واقع‌گرایی به نظر می‌رسد برداشتی اشتباه و در عین حال بسیار رایج در میان پژوهشگران علوم اجتماعی است (مکسول و میتاپالی، ۲۰۱۰). در این میان، اما بورسیوم (۲۰۰۵) این دو مکتب را از هم متمایز می‌داند. به‌مانند بورسیوم، برای تحلیل بهتر مقالات، مقوله‌ی اثبات‌گرایی را نیز به تقسیم‌بندی فالچر اضافه کردیم (جدول ۲). در این راستا تقسیم بندی مطالعات صرفاً کمی بر اساس تقسیم بندی موضوعی فالچر امکان‌پذیر نبود. بنابراین، ناگزیر شدیم برای اطمینان از نتایج تقسیم‌بندی، معیارهای دیگری را نیز در نظر بگیریم.

یکی از معیارهایی که با استفاده از آن می‌توان گرایش و مبانی فلسفی کارهای پژوهشی مرتبط با سنجش زبان را تعیین کرد، دقت در معناشناسی روایی است (هتکورث^۱، ۲۰۱۳). بنابراین، تعریفی که پژوهشگران از روایی انتخاب می‌کنند می‌تواند تا حدودی نشان دهنده‌ی مبانی و مفروضات هستی‌شناسی و معرفت‌شناسی آنان در باره سنجش زبان باشد. چنانچه در دیدگاه‌های دارای سمت و سوی واقع‌گرایی، روایی صفت آزمون است اما در رویکرد ابزارانگاری، روایی صفتی است برای تفسیر و یا کاربرد آزمون (همان). جهت بررسی معناشناختی، واژه‌ی روایی و مشتقات آن را با استفاده از گزینه‌ی جستجو در نرم‌افزار اکروبات‌ریدر جستجو کردیم.

یک دسته از مقالات که تعیین رویکرد فلسفی‌شان دشوار می‌نمود، مقالاتی بودند که ما از آنها به عنوان پژوهش‌های پیرامون آزمون زدگی^۲ یاد می‌کنیم. این مطالعات به‌پیامدهای

1- Hathcoat, J. D.

2- Washback

آزمون‌ها برای آموزش زبان، زبان آموزان و معلمان زبان می‌پردازند. بنا بر تقسیم‌بندی فالچر، پژوهش‌های این‌چنینی با مبانی برساخت‌گرا سنخیت نظری بیشتری دارند. اما تحلیل روش پژوهشی آنها نشان داد که عمدتاً پیرو روش‌های کمی هستند و نه در روش پژوهش و نه در مبانی نظریشان نسبت چندانی با مکتب برساخت‌گرایی ندارند. از طرف دیگر، از نظر برسبوم (۲۰۰۶) به عنوان یک واقع‌گرای تمام‌عیار، روایی یک آزمون نسبتی با پیامدهای آن ندارد. گزینه‌ی بعد این بود که چنین مقالاتی را در زمره‌ی پژوهش‌های مبتنی بر ابزارانگاری قرار دهیم، چرا که از نظر مسیک پیامدهای آزمون را باید بخشی از شواهد روایی آزمون دانست. اما مسیک (۱۹۹۶) تنها پیامدهایی را به عنوان شواهد روایی یا عدم روایی آزمون می‌پذیرد که اگر آزمون مربوطه وجود نداشت، آن پیامدها اتفاق نمی‌افتادند. به همین دلیل است که توصیه می‌شود پژوهشگران علاقمند به آزمون‌زدگی همیشه باید در طرح پژوهشی خود یک عنصر مقایسه‌ای بگنجانند تا بتوانند تاثیر سایر متغیرهای دخیل در فرایند آزمون‌زدگی را کنترل نمایند (واتانابی^۱، ۲۰۰۴). با این توصیف، مقالات مربوطه را نمی‌شد به هیچ‌یک از چهارمقوله‌ی تعریف شده نسبت داد. از این رو، ما آنها را در مقوله‌ی "نامشخص" قرار دادیم. تنها در یک مورد، یکی از این مقالات را، که رویکردی انتقادی داشت، در مقوله‌ی برساخت‌گرا قرار دادیم. همچنین مقالاتی را که در مورد ارزیابی‌های جایگزین^۲ بودند به‌سادگی نمیشد به‌یکی از مکاتب فلسفی یادشده متعلق دانست. لذا این گروه مقالات را بسته به مبانی فلسفی‌شان به‌دو گروه برساخت‌گرا یا واقع‌گرا تقسیم کردیم. از نظر براون و هودسن^۳ (۱۹۹۸)، سنجش‌های جایگزین باید با معیارهای روایی و پایایی مبتنی بر واقع‌گرایی مورد ارزیابی قرار گیرند. لذا مقالاتی را که چنین نگاهی به سنجش‌های جایگزین داشتند در مقوله‌ی واقع‌گرا قرار داده شدند. اما دیگر صاحب‌نظران مانند لینچ و شاو^۴ (۲۰۰۵) معتقدند که روایی سنجش‌های جایگزین را نباید با معیارهای سنتی بررسی نمود. بلکه این سنجش‌ها مبتنی بر فلسفه‌ی برساخت‌گرا هستند. با این توجیه، مقالاتی که چنین نگاهی به سنجش‌های جایگزین داشتند در دسته‌ی فلسفه‌ی برساخت‌گرا قرار داده شدند. با این همه، تردیدی نیست که تقسیم مقالات به مکاتب فلسفی یادشده لاجرم میرا از سوگیری تحلیلی‌گر نیست اما برای کاهش میزان این سوگیری

1- Watanabe, Y.

2- Alternative Assessment

3- Brown, J. D., & Hudson, T.

4- Lynch, B., & Shaw, P.

تعداد پانزده مقاله (حدود بیست درصد از کل پیکره) مورد تحلیل مجدد قرار گرفت و اختلاف مشهودی میان کدبندی اولیه و ثانویه مشاهده نشد.

۱۲- یافته‌ها

از جنبه‌ی معنانشناسی واژه‌ی روایی، جستجو در فایل‌های پی‌دی‌اف مقالات نشان داد که در مجموع، واژه‌ی روایی و مشتقات آن ۴۱۲ بار و واژه‌ی رواسازی ۶۶ بار در مقالات بکار رفته بودند. با وجود این بسامد بالا، فقط چهار مقاله روایی و رواسازی را تعریف کرده بودند، که این نشان می‌دهد در این مقالات عمدتاً معنا و تعریف روایی بدیهی و بی‌نیاز از تعریف کردن تلقی شده است. این در حالی است که بیش از صد و پنجاه تعریف متفاوت از روایی وجود دارد (نیوتن و شاول، ۲۰۱۴). این نگرش به مفهوم روایی را شاید بتوان ناشی از استیلای نگاه‌های اثبات‌گرایی و واقع‌گرایی در پژوهش در سنجش زبان دانست (فالچر، ۲۰۱۳). جدول شماره ۳ بسامد مقالات مربوط به هر یک از مکاتب فلسفی را نشان می‌دهد. می‌بینیم که تعداد ۲۸ مقاله، مبتنی بر نظریه‌ی کلاسیک آزمون می‌باشند. این مقالات عموماً در مورد ساخت آزمون‌های زبان، مطالعه‌ی روایی و پایایی آزمون‌ها با روش‌های مبتنی بر نظریه‌ی کلاسیک اندازه‌گیری و یا بررسی دیگر ویژگی‌های آزمون‌ها که قابل مطالعه در نظریه‌ی کلاسیک سنجش^۳ هستند، می‌باشند. به عنوان نمونه، تعداد چهار مقاله به بررسی روایی آزمون کلوز^۴ و همین تعداد مقاله به جنبه‌های مختلف آزمون‌های چند گزینه‌ای پرداخته بودند. در این قبیل مطالعات، دو واژه روایی و آزمون به‌صراحت و متواتر در ترکیب صفت و موصوف استفاده شده بودند که نشان از اتخاذ نگرش واقع‌گرایی نسبت به روایی می‌باشد.

1- definition

2- Newton, P., & Shaw, S.

3- Classical Test Theory

4- Cloze Test

جدول ۲. دیدگاه‌های فلسفی و حوزه‌های پژوهشی سنجش زبان

تعداد مقالات	حیطه‌های مربوطه	مبنای فلسفی
۲۵	ساخت یا بررسی روایی و پایایی آزمون با استفاده از نظریه کلاسیک سنجش	اثبات‌گرایی
۲	آزمون‌های تشخیصی	واقع‌گرایی
۷	پژوهش‌های مبتنی بر تحلیل عاملی و آی آر تی	
۵	طراحی مدل‌های نمره‌دهی	
۶	سنجش‌های جایگزین	
۲۰	مجموع	
۱	تبیین نقش مخاطب در آزمون‌های شفاهی	برساخت‌گرایی
۶	سنجش منظورشناسی	
۵	انصاف و تبعیض در آزمون‌ها	
۰	افشا و برملا کردن کارکردهای پنهان آزمون‌ها	
۱۰	ارزیابی پویا	
۰	سنجش‌های متفاوت	
۲	آزمون‌زدگی با رویکرد انتقادی و سیاسی	
۲۴	مجموع	
۰	شرح و توصیف فرایند آزمون‌سازی	ابزارانگاری
۲	اعتبارسنجی مبتنی بر استدلال	
۲	مجموع	
۷	آزمون‌زدگی در مقیاس آموزش	نامشخص

همچنین جدول ۲ نشان می‌دهد که در مقوله‌ی واقع‌گرا، دو مقاله در مورد آزمون‌های تشخیصی شناختی^۱، تعداد هفت مقاله بر مبنای تحلیل عاملی و نظریه‌ی آیتم-پاسخ^۲، پنج مقاله پیرامون طراحی شیوه‌های نمره‌دهی مهارت‌های گفتاری و نوشتاری، و شش مقاله در مورد سنجش‌های جایگزین هستند. در مقوله‌ی مطالعات دارای خاستگاه برساخت‌گرا، یک مقاله به نقش مخاطب در آزمون‌های شفاهی، شش مقاله در مورد سنجش منظورشناسی، پنج مقاله پیرامون انصاف و تبعیض در آزمون‌های زبان، دو مقاله در مورد روابط قدرت و آزمون، و ده مقاله در مورد ارزیابی پویا بودند. فقط یک مقاله به طور مشخص دارای مبانی فلسفی ابزارگرایی بود و هفت مقاله نیز در دسته‌ی مقالات مبهم قرار گرفتند که با معیارهای اتخاذ شده تعیین مبانی هستی‌شناختی و معرفت‌شناسی‌شان دشوار می‌نمود.

جدول ۳ فراوانی مقالات سنجش زبان را در نشریات مختلف به تفکیک مبانی فلسفی شان نشان می‌دهد. واضح است که مکاتب اثبات‌گرایی و واقع‌گرایی همچنان رایج‌تر از دو نگاه دیگر هستند. پنجاه و هفت درصد مقالات بررسی شده به سوژه‌ی پژوهشی خود نگاه اثبات‌گرا یا واقع‌گرا داشتند. یک سوم مقالات (تعداد ۲۴ مقاله) دارای گرایش برساخت‌گرا بودند. حدود دو درصد پیرو فلسفه‌ی ابزارگرایی و حدودا در ده درصد موارد هم تعیین مبانی فلسفی مقالات دشوار بود و لذا در مقوله‌ی مبهم قرار گرفتند. چنانچه قبلا ذکر شد این مقالات عموماً در مورد آزمون‌زدگی بودند.

۱۳- بحث و نتیجه‌گیری

هدف از انجام این پژوهش بررسی بستر فلسفی پژوهش در سنجش زبان در ایران با تاکید بر هستی‌شناسی و معرفت‌شناسی بود. در یک نگاه کلی به‌یافته‌های این پژوهش، شاید بتوان ادعا کرد که پژوهش در سنجش زبان ایران عمدتاً در بستر مکاتب اثبات‌گرا، واقع‌گرا و تا حدودی برساخت‌گرا انجام می‌شود و مکتب ابزارانگاری، علی‌رغم اقبال جهانی آن، هنوز نفوذ و رواج چندانی در آثار پژوهشی داخلی نداشته است.

1- Cognitive diagnostic assessment

2- Item Response Theory

جدول ۳. فراوانی مقالات با مبانی فلسفی مختلف به تفکیک نشریه

نشریه	اثبات‌گرا	واقع‌گرا	برساخت‌گرا	ابزارانگار	نامشخص	مجموع
FLRJ	۱	۲	۱	۰	۰	۴
JELTL	۶	۲	۱	۰	۱	۱۰
ARE	۳	۲	۲	۰	۳	۹
JTLS	۴	۲	۴	۱	۰	۱۱
RALS	۲	۲	۱	۰	۲	۷
IJALS	۵	۱	۴	۰	۱	۱۱
IJAL	۴	۹	۳	۰	۰	۱۶
IJLTR	۰	۰	۴	۱	۰	۵
ILT	۰	۰	۴	۰	۰	۴
مجموع	۲۵ (۳۲٪)	۲۰ (۲۵٪)	۴۲ (۳۰٪)	۲ (۲.۵٪)	۷ (۹٪)	۷۸

تا آنجا که به رواج مکتب اثبات‌گرایی مربوط است، می‌توان گفت که در قیاس با رویکرد کلی رشته‌ی زبان‌شناسی کاربردی، تصویر منتج از نتایج حاضر موید روند غالب در زبان‌شناسی کاربردی می‌باشد. ویدمن^۱ (۲۰۱۳) سیر تاریخی پژوهش در زبان‌شناسی کاربردی را در هفت مرحله خلاصه می‌کند که چهار مرحله‌ی اول متکی به اثبات‌گرایی و سه مرحله‌ی بعدی یعنی برساخت‌گرایی، پست‌مدرنیسم و نظریه‌ی سیستم‌های پیچیده از مبانی اثبات‌گرایی فاصله گرفته‌اند. ویدمن اما اذعان دارد که اثبات‌گرایی همچنان مکتب فلسفی غالب بسیاری از پژوهش‌های حوزه‌ی زبان‌شناسی کاربردی است.

اما قضاوت در مورد میزان تناسخ نتایج یک پژوهش با کارهای مشابه پیشین گاهی دشوار می‌نماید (ریان و آیکنهد^۲، ۱۹۹۲)، بخصوص که دریافت پژوهشگران رشته‌ی زبان‌شناسی کاربردی از حدود و ثغور پارادایم‌های فلسفه‌ی علم با هم متفاوت بوده و بر همین مبنا هم برداشت آنها از میزان رواج مکاتب فلسفی مختلف در سنجش زبان متفاوت و بعضاً متناقض

1- Weideman, A.

2- Ryan, A. G., & Aikenhead, G. S.

است. به عنوان نمونه همپلاینز و لینچ (۱۹۹۸) و لطف آبادی (۱۳۸۶) همه‌ی پژوهش‌های محاسباتی را در مکتب اثبات‌گرا گنجانده‌اند. دو مقاله‌ی مروری دیگری که در مورد مقالات سنجش زبان نوشته شده‌اند (چلوبدویل و دوپیل^۱، ۲۰۰۸، لوملی و براون^۲، ۲۰۰۵) نیز کم و بیش با همین تلقی به بررسی روش پژوهش در سنجش زبان پرداخته‌اند. چلوبدویل و دوپیل اما تصریح می‌کنند که اگرچه بیش از نود درصد مقالات نشریه‌ی سنجش زبان مبتنی بر مبانی روان‌سنجی هستند، در دیگر نشریه‌ی تخصصی این رشته، فصلنامه ارزشیابی زبان^۳، فقط پنجاه درصد مقالات در بستر روان‌سنجی انجام شده‌اند. اگر به مانند این محققین، دایره‌ی اثبات‌گرایی را این‌گونه وسیع قلمداد کنیم، نتیجه این خواهد بود که پژوهش در سنجش زبان ایران نیز عمدتاً اثبات‌گرا مانده است. در این صورت باید گفت، ما هنوز در عصر روان‌سنجی-ساختارگرایی^۴ (ویر، ۲۰۰۵) سنجش زبان مانده‌ایم.

از طرف دیگر، بعضی از محققان اثبات‌گرایی را مکتبی مرده می‌دانند که امروزه هیچ پیروی ندارد (فالچر، ۲۰۱۳: میسلوی^۵، ۲۰۱۸). همچنین از نظر متخصصان حوزه‌ی روان‌سنجی و فیلسوفان اجتماعی، عملیات‌گرایی و اثبات‌گرایی را نباید با واقع‌گرایی یکی دانست چرا که اثبات‌گرایی مبتنی بر نفی ماهیت توانش‌های باطنی است اما واقع‌گرایی دقیقاً نقطه‌ی عکس آن است (بورسبوم، ۲۰۰۵: مارکوس و بورسبوم، ۲۰۱۳). لذا بر اساس نتایج پژوهش حاضر، که مبتنی بر تمایز میان اثبات‌گرایی و واقع‌گرایی است، تحقیق در سنجش زبان در ایران در طیفی از مکاتب فلسفی اثبات‌گرایی، واقع‌گرایی و برساخت‌گرایی انجام می‌شود.

قضاوت در مورد اینکه آیا این وضعیت ایده آل است یا نامطلوب بستگی به آبشخور فلسفی شخص قضاوت‌گر دارد. از نظر بعضی فیلسوفان سنجش، این وضعیت کاملاً مطلوب بوده و اصولاً غیر از این نباید باشد (بورسبوم، ۲۰۰۵). چرا که وقتی سخن از سنجش به میان می‌آید، حتماً سخن از سنجش ماهیتی است و اگر آن ماهیت وجود خارجی نداشته باشد، وجودش محصول قراردادهای اجتماعی باشد و یا در وجود آن تردید داشته باشیم، اساساً سخن از اندازه‌گیری چنین ماهیتی بیهوده است. با همین استدلال، میچل (۱۹۹۰) کل تاریخ سنجش و

1- Chalhoub - Deville, M., & Deville, C.

2- Lumley, T., & Brown, A.

3- Language Assessment Quarterly

4- Psychometric-structuralist

5- Mislevy, R. J.

اندازه‌گیری در روانشناسی و دیگر علوم اجتماعی را زیر سوال می‌برد. او معتقد است که در علوم اجتماعی و روانشناسی تاکنون هیچ سنجش معتبری انجام نشده است. ایشان با مقایسه سنجش در روانشناسی و سنجش در فیزیک می‌گویند که در سنجش‌های فیزیکی و بخصوص اندازه‌گیری فاصله، ابزار اندازه‌گیری خود کسری از سازه‌ی فاصله است در حالی که در روانشناسی و علوم اجتماعی میان ابزارهای اندازه‌گیری و سوره‌های سنجش هیچ رابطه‌ی آشکاری دیده نمی‌شود. او حتی سنجش سازه‌های ثانویه فیزیکی مانند چگالی را نیز حاصل نسبت میان دو سازه‌ی جرم و حجم می‌داند که واحد اندازه‌گیری هر کدام کسری از سازه‌ی مورد نظر است.

از طرف دیگر، بسیاری از صاحب‌نظران پایبندی به فلسفه‌ی اثبات‌گرایی و واقع‌گرایی را در سنجش زبان نقطه ضعف آن می‌دانند (سیرسی، ۲۰۰۸، نقل در چپله ۲۰۱۲؛ مک‌نامارا و روور، ۲۰۰۶). هم سیرسی و هم مک‌نامارا و روور معتقدند که رویکرد واقع‌گرایانی مانند بورسبوم باعث می‌شود سنجش در آموزش هشتاد سال پسرفت داشته باشد، به این دلیل که در فلسفه‌ی واقع‌گرایی، نگاه به روایی یک نگاه روانشناسی صرف است در حالی که برساخت‌گرایان و ابزارانگاران سنجش را پدیده‌ای فرهنگی، اجتماعی و سیاسی می‌دانند که برای مطالعه‌ی روش‌های کمی ناکافی می‌باشند. بنابراین اگر از این منظر قضاوت کنیم، باید بگوییم استیلای مبانی اثبات‌گرا و واقع‌گرا در سنجش زبان ایران را نباید مطلوب قلمداد کرد.

دیگر یافته‌ی پژوهش فعلی این است که استفاده از فلسفه‌ی ابزارانگاری و نظریه‌ی استدلال محور روایی در میان پژوهش‌های انجام شده بسیار محدود (تنها دو مقاله) می‌باشد. این در حالی است که صاحب‌نظران فلسفه‌ی ابزارانگاری و مکتب استدلال‌گرایی در روایی را نقشه‌ی راه آینده‌ی علم سنجش می‌دانند (بکمن، ۲۰۰۵؛ بکمن و پالمر، ۲۰۱۰؛ چپله، ۲۰۱۲؛ فاکس و چنگ، ۲۰۱۳؛ فالچر، ۲۰۱۴؛ فالچر و دیویدسون، ۲۰۰۷). همچنین، این وضعیت با روند فعلی اقبال به روش تحقیق مختلط در سایر حوزه‌های علوم اجتماعی و انسانی (جانسون و انوگبوزی^۱، ۲۰۰۴) در تنافر می‌باشد. بعضی از محققان نظام‌های آموزشی را در این راستا قابل سرزنش می‌دانند از این جهت که به‌دانش‌آموزان چنین القا می‌کنند که علم کشف تدریجی رازهای مکنون عالم هستی است و دانش تطابق میان نظریه‌ها و واقعیات عالم هستی می‌باشد (ریان و آیکنهد، ۱۹۹۲؛ ساتون^۲، ۱۹۹۶). لذا این حقیقت که دانش همیشه از غربال فهم

1- Johnson, R. B., & Onwuegbuzie, A. J.

2- Sutton, C.

انسانی-اجتماعی عبور می‌کند، نادیده گرفته می‌شود (همان). همچنین محتوای کتب درسی سنجش زبان، که در دوره‌های تحصیلات تکمیلی تدریس می‌شوند، هم در تقویت این انگاره‌های معرفت‌شناسی بی‌تاثیر نیستند. در میان کتب درسی رایج، تنها بکمن (۲۰۱۰) و فالچر و دیویدسون (۲۰۰۷) به صورت نسبتاً مبسوط به مکتب استدلال‌گرایی روایی پرداخته‌اند. امروزه اما غالب صاحب‌نظران چندصدایی در مورد فلسفه‌ی پژوهش سنجش زبان را به سود این حوزه می‌دانند. فالچر (۲۰۱۵) معتقد است که مواضع افراطی چه واقع‌گرا باشند، چه برساخت‌گرا و یا ابزارگرا نه قابل دفاع هستند و نه به مصلحت پیشرفت علم سنجش زبان می‌باشند. لوملی و براون (۲۰۰۵) هم کمابیش چنین موضعی دارند.

چنانچه قبلاً نیز اشاره شد یکی از دشواری‌های پژوهش پیرامون فلسفه‌ی علم در سنجش زبان این است که معمولاً پژوهشگران از مبانی هستی‌شناسی و معرفت‌شناسی خود سخنی به‌میان نمی‌آورند. در بخش نتایج دیدیم که غالب مقالات به چارچوب نظری خود هیچ تصریحی نکرده بودند. شاید بهتر باشد که استادان و پژوهشگران زبان‌شناسی کاربردی و بخصوص سنجش زبان در دوره‌های روش تحقیق و راهنمایی پایان‌نامه‌ها بر اهمیت مبانی فلسفی چارچوب‌های نظری و روش‌های پژوهشی تاکید بیشتری داشته باشند چرا که این باعث می‌شود پژوهشگر در مبانی فلسفی خود به دقت بیشتری بنگرد و جایگاه کار پژوهشی خود را در میان نحله‌های فکری و فلسفی متعدد بهتر بشناسد، که این خودآگاهی باعث می‌شود پژوهشگران بتوانند نتایج کارهای پژوهشی خود را از افقی بالاتر مورد بحث و مذاقه قرار دهند. لطف آبادی (۱۳۹۶) حتی معتقد است که "پژوهش نوآورانه در علوم انسانی بدون دانش وسیع در باره بنیادهای معرفت‌شناسی پژوهش امکان‌پذیر نیست" (ص. ۱۰).

در آخر باید اذعان داشت که دشوار می‌توان یافت پژوهشی در گستره‌ی علوم اجتماعی که منزله از سوگیری‌ها و گرایش‌های خواسته و ناخواسته‌ی پژوهشگر باشد. لذا پژوهش حاضر هم لاجرم سهم خود را از کاستی‌ها دارد. اولاً، تعریف و مرزبندی داده شده از مکاتب فلسفی، بمانند دیگر مرزبندی‌های علمی، ناچاراً قراردادی است چرا که مرز کاملاً محرز و مشخصی میان مکاتب مذکور وجود ندارد. لذا پژوهشگران آینده می‌توانند با بازتعریف مکاتب یادشده به پژوهش در این حوزه کمک کنند. همچنین، در صورت امکان، چنانچه کار تقسیم‌بندی و کدگذاری مقالات توسط چند نفر انجام شود، حتماً به جبران کاستی‌های این پژوهش کمک شایانی خواهد کرد (این مهم در وسع نویسنده‌ی این مقاله نبود). سخن آخر اینکه، راه یافتن به مفروضات فلسفی مضمور در پژوهش، رسالتی در مرز ناممکن است.

بی‌گمان، پژوهش‌هایی که مبتنی بر شواهد متنعدد و از منابع متنوع همچون مصاحبه با پژوهشگران و تحلیل کلامی عمیق‌تر باشند خواهند توانست رهگشا و مقرون به واقعیت باشند.

۱۴- منابع

- لطف‌آبادی، ح. (۲۰۰۷). کاستی‌های معرفت‌شناسی و روش‌شناسی در مقالات پژوهشی روانشناسی در ایران. فصلنامه حوزه و دانشگاه روش‌شناسی علوم انسانی، ۱۳(۵۱)، ۹-۴۴.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2. doi:10.1207/s15434311laq0201_1
- Bachman, L. F., & Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*: Cambridge University Press.
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Cantley, I. (2017). A quantum measurement paradigm for Educational Predicates: Implications for validity in educational measurement. *Educational Philosophy and Theory*, 49(4), 405-421.
- Chalhoub - Deville, M., & Deville, C. (2008). Utilizing psychometric methods in assessment. In N. Hornberger & E. Shohamy (Eds.), *Encyclopedia of language and education* (pp. 2349-2361). NY: Springer.
- Chappelle, C. A., Enright, M. K., & Jamieson, J. M. (2011). *Building a validity argument for the Test of English as a Foreign Language™*: Routledge.
- Chappelle, C. A. (2012). Conceptions of validity. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 21-33). London: Routledge.
- Cheng, L., & Fox, J. (2013). Review of doctoral research in language assessment in Canada (2006–2011). *Language Teaching*, 46(4), 518-544.
- Corson, D. (1997). Critical realism: An emancipatory philosophy for applied linguistics? *Applied Linguistics*, 18(2), 166-188.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281-302.

- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347.
- Farhady, H., Jafarpur, A., & Birjandi, P. (1994). Testing language skills: From theory to practice. *Tehran: SAMT Publications*.
- Foucault, M. (1977). Discipline and punishment: The birth of the prison. *London: Allen Lane*.
- Fulcher, G. (2010). *Practical language testing*: Routledge.
- Fulcher, G. (2013). Philosophy and language testing. In A. J. Kunnan (Ed.), *The companion to language assessment* (Vol. 3, pp. 1431-1451). Chichester: John Wiley & Sons.
- Fulcher, G. (2015). *Reexamining language testing: A philosophical and social inquiry*.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. New York: Routledge
- Galaczi, E., & Taylor, L. (2018). Interactional Competence: Conceptualisations, Operationalisations, and Outstanding Questions. *Language Assessment Quarterly*, 15(3), 219-236. doi:10.1080/15434303.2018.1453816
- Hacking, I. (1999). *The social construction of what?* Cambridge, Mass.: Harvard university press.
- Hamp-Lyons, L., & Lynch, B. K. (1998). Perspectives on validity: A historical analysis of language testing conference abstracts. In A. J. Kunnan (Ed.), *Issues in Language Testing Research: Conventional Validity and Beyond* (pp. 253-277). NJ: Lawrence Erlbaum, Mahwah.
- Hathcoat, J. D. (2013). Validity semantics in educational and psychological assessment. *Practical assessment, research & evaluation*, 18(9), 1-14.
- Hood, S. B. (2009). Validity in psychological testing and scientific realism. *Theory & Psychology*, 19(4), 451-473.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14-26. doi:10.3102/0013189x033007014
- Johnson, R. C., & Riazi, A. M. (2015). Accuplacer Companion in a foreign language context: An argument-based validation of both test score meaning and impact. *Papers in Language Testing and Assessment*, 4(1), 31-58.
- Jordan, G. (2004). *Theory construction in second language acquisition*. Amsterdam: John Benjamins Publishing.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73.
- Knoch, U., & Chapelle, C. A. (2018). Validation of rating processes within an argument-based framework. *Language Testing*, 35(4), 477-499.

- LaFlair, G. T., & Staples, S. (2017). Using corpus linguistics to examine the extrapolation inference in the validity argument for a high-stakes speaking assessment. *Language Testing*, 34(4), 451-475. doi:10.1177/0265532217713951
- Lazaratan, A. (2005). Quantitative research methods. In E. Hinkle (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 209-244). NJ: Lawrence Erlbaum Associates, Publishers, Mahwah.
- Lumley, T., & Brown, A. (2005). Research methods in language testing. In E. Hinkle (Ed.), *Handbook of research in second language teaching and learning* (pp. 857-880). NJ: Lawrence Erlbaum Associates, Publishers, Mahwah.
- Lynch, B., & Shaw, P. (2005). Portfolios, Power, and Ethics. *TESOL Quarterly*, 39(2), 263-297. doi:10.2307/3588311
- Markus, K. A., & Borsboom, D. (2013). *Frontiers of test validity theory: Measurement, causation, and meaning*. New York: Routledge.
- Maul, A., Irribarra, D. T., & Wilson, M. (2016). On the philosophical foundations of psychological measurement. *Measurement*, 79, 311-320.
- Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In A. Tashakkori & C. Teddie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 145-168). California: Sage.
- McNamara, T., & Roever, K. (2006). *Language testing: The social dimension*. Oxford: Blackwell.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement*. New York: Macmillan.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256. doi:10.1177/026553229601300302
- Michell, J. (1990). *An introduction to the logic of psychological measurement*. Hillsdale, NJ: Erlbaum: Psychology Press.
- Mislevy, R. J. (2018). *Sociocognitive foundations of educational measurement*. New York: Routledge.
- Moss, P. A., Girard, B. J., & Haniford, L. C. (2006). Chapter 4: Validity in educational assessment. *Review of research in education*, 30(1), 109-162.
- Newton, P., & Shaw, S. (2014). *Validity in educational and psychological assessment*. California: Sage.
- Nisbett, R. (2004). *The geography of thought: How Asians and Westerners think differently... and why*. New York: Free Press.
- Noddings, N. (2018). *Philosophy of education*. New York: Routledge.
- Norris, S. P. (1983). The inconsistencies at the foundation of construct validation theory. *New directions for program evaluation*, 19, 53-74.

- Packer, M. J. (2017). *The science of qualitative research*. Cambridge: Cambridge University Press.
- Popper, K. (2014). *Conjectures and refutations: The growth of scientific knowledge*. New York: Routledge.
- Ryan, A. G., & Aikenhead, G. S. (1992). Students' preconceptions about the epistemology of science. *Science education*, 76(6), 559-580.
- Shohamy, E. (2014). *The power of tests: A critical perspective on the uses of language tests*. New York: Routledge.
- Sutton, C. (1996). Beliefs about science and beliefs about language. *International Journal of Science Education*, 18(1), 1-18.
- Toulmin, S. E. (2002). *The uses of argument*. Cambridge: Cambridge university press.
- Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing* (pp. 41-58). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Weideman, A. (2013). Applied linguistics beyond postmodernism. *Acta Academica*, 45(4), 236-255.
- Weir, C. J. (2005). *Language testing and validation*: Hampshire: Palgrave McMillan.
- Xiaoming, X. (2010). How do we go about investigating test fairness? *Language Testing*, 27(2), 147-170. doi:10.1177/0265532209349465