

**Impact meta-evaluation: Introducing a checklist¹**Fatemeh Sabbaghi², Keyvan Salehi^{3*}

(Received: 2025.04.05 - Accepted: 2025.07.21)

1- This article is derived from independent research.

2- Ph. D. Student in Measurement, Faculty of Psychology and Educational Sciences, University of Tehran, Tehran, Iran

3- Associate Professor, Department of Research and Measurement, Faculty of Psychology and Educational Sciences, University of Tehran, Tehran, Iran

*- Corresponding Author: keyvansalehi@ut.ac.ir

Abstract

Impact evaluation is one of the most necessary steps in implementing plans and programs, and the most strategic way of judging the quality of the activities, especially in the information age. Based on the importance of this evaluation method, compiling and introducing a detailed checklist for evaluating the quality of impact evaluations (meta-impact evaluation) is an extreme necessity. In the first stage, studying, coding, and detailed analysis of 16 documents based on Altheide's strategy led to the identification of 35 items. Then, in the second stage, 18 initially eligible studies were evaluated using the compiled checklist. At this stage, the findings indicate that the studies did not reach the mastery level in 94.28% of the indicators, equivalent to 33 indicators out of 35 indicators (precisely, 100% of the evaluated studies did not reach the mastery level in 65.71% of the indicators, more than 50% of the evaluated studies did not reach the mastery level in 28.57% of the indicators, and 50% of the evaluated studies did not reach the mastery level in 5.71% of the indicators). Therefore, the analysis of the results indicates the extreme weakness of impact evaluation studies in observing the underlying principles of this type of evaluation, including presenting a theory of change, preparing an evaluation plan, developing an evaluation checklist, forming a trained assessment team, and other key steps, raises serious doubts about the quality of these studies. Also, the lack of attention to the principles of impact evaluation as the most important part of program implementation, interventions, and changes causes neglect of the natural, long-term, and sustainable effects of programs and attention to superficial, immediate, and short-term results. Therefore, the attention of managers, evaluators, and researchers in the field of impact evaluation to the basic principles of impact evaluation and essential items presented in the checklist of the current research and evaluation based on them is necessary.

Keywords: Impact evaluation, meta-evaluation, checklist, Quality of evaluation studies, Effectiveness and efficiency of plans, Quality of evaluation programs



فرا ارزشیابی تأثیر؛ معرفی یک فهرست‌وارسی^۱

فاطمه صباغی^۲، کیوان صالحی^{۳*}

(دریافت: ۱۴۰۴/۰۱/۱۶ - پذیرش: ۱۴۰۴/۰۴/۳۰)

چکیده

ارزشیابی تأثیر یکی از ضروری‌ترین گام‌ها در فرآیند اجرای طرح‌ها و برنامه‌ها و راهبردی‌ترین شیوه قضاوت درباره کیفیت فعالیت‌های انجام‌شده به خصوص در عصر اطلاعات است. با توجه به اهمیت و جایگاه این شیوه ارزشیابی، تدوین و معرفی یک فهرست‌وارسی دقیق برای ارزیابی کیفیت ارزشیابی‌های تأثیر (فرا ارزشیابی تأثیر) از ضرورت دوجندانی برخوردار است. در پژوهش حاضر تلاش شده‌است تا با استفاده از رویکرد آمیخته از نوع اکتشافی و با تلفیق دو روش آسنادی و مطالعات ارزشیابی، یک فهرست‌وارسی فرا ارزشیابی تأثیر، طراحی و تدوین شود. در مرحله نخست مطالعه، کدگذاری و تحلیل دقیق ۱۶ سند مبتنی بر راهبرد آلتاید، منجر به شناسایی ۳۵ گویه شد. در مرحله دوم با استفاده از فهرست‌وارسی تدوین‌شده، به ارزشیابی ۱۸ پژوهش واجد شرایط اولیه پرداخته‌شد. در این مرحله، یافته‌ها نشان‌دهنده عدم دستیابی پژوهش‌ها به سطح تسلط در ۹۴/۲۸ درصد از نشانگرها معادل ۳۳ نشانگر از ۳۵ نشانگر (به طور دقیق عدم دستیابی ۱۰۰ درصد پژوهش‌های ارزیابی‌شده به حد تسلط در ۶۵/۷۱ درصد از نشانگرها، عدم دستیابی بیش از ۵۰ درصد پژوهش‌های ارزیابی‌شده به حد تسلط در ۲۸/۵۷ درصد از نشانگرها، و عدم دستیابی ۵۰ درصد پژوهش‌های ارزیابی‌شده به حد تسلط در ۵/۷۱ درصد از نشانگرها) است. بنابراین تحلیل نتایج حاکی از ضعف مفرط مطالعات ارزشیابی تأثیر در رعایت اصول زیربنایی این نوع ارزشیابی از جمله ارائه یک نظریه تغییر، تهیه طرح ارزشیابی، تهیه فهرست‌وارسی ارزشیابی، تشکیل تیم ارزیابی آموزش‌دیده و موارد دیگر است که این مسأله کیفیت این مطالعات را با تردید جدی همراه می‌سازد. همچنین کم‌توجهی به اصول ارزشیابی تأثیر به‌عنوان مهم‌ترین بخش در اجرای برنامه‌ها، مداخلات و تغییرات، به غفلت از تأثیرات واقعی، بلندمدت و پایدار برنامه‌ها و توجه به نتایج سطحی، فوری و کوتاه مدت می‌انجامد، لذا توجه مدیران، ارزشیابان و پژوهشگران در حوزه ارزشیابی تأثیر به اصول اساسی ارزشیابی تأثیر و گویه‌های مهم ارائه‌شده در فهرست‌وارسی پژوهش حاضر و ارزشیابی بر اساس آن‌ها ضروری است.

واژگان کلیدی: ارزشیابی تأثیر، فرا ارزشیابی، فهرست‌وارسی، کیفیت مطالعات ارزشیابی، اثربخشی و کارایی طرح‌ها، کیفیت برنامه‌های ارزشیابی

۱- این مقاله مستخرج از پژوهش مستقل است.

۲- دانشجوی دکتری سنجش و اندازه‌گیری، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران

۳- دانشیار بخش تخصصی پژوهش و سنجش، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران

* نویسنده مسئول: keyvansalehi@ut.ac.ir

مقدمه

در سال‌های اخیر، با عمیق‌تر شدن بحران‌های مالی، اقتصادی، اجتماعی، زیست‌محیطی، سیاسی، فناوری، و بهداشتی جهانی (میشل و اشنايدر^۱، ۲۰۲۵، ۱)، علاقه به ارزشیابی تأثیر در حوزه‌های مختلفی از جمله سیاست‌گذاری، اقتصاد، بهداشت، آموزش، و علوم اجتماعی به شدت افزایش یافته‌است و از این‌رو تمرکز فزاینده‌ای بر ارزشیابی تأثیر در جوامع توسعه یافته‌است (فرولیچ و اسپرلیچ^۲، ۲۰۱۹، ۱؛ گیوگتی، کارلان، و ولش^۳، ۲۰۱۶؛ بیللیس، هانی روزس، برنر، کوبرا، ازین دبلاس، فرارو، ... و واندر^۴، ۲۰۱۶، ۵۸؛ و وایت^۵، ۲۰۱۰، ۱۵۳). اصطلاح ارزشیابی تأثیر به طور گسترده‌ای برای اشاره به هر بحثی در مورد نتایج و شاخص‌های تأثیر، و گاهی محدودتر از آن به مطالعاتی که به صراحت به دنبال نسبت دادن نتایج به یک مداخله مشخص هستند، استفاده می‌شود (کوپستیک^۶، ۲۰۱۴، ۴۱۶؛ وایت، ۲۰۱۰، ۱۵۴). بنابراین می‌توانیم بگوییم افراد مختلف، تعاریف مختلفی از «ارزشیابی تأثیر» دارند. از آنجایی که این یک موضوع صرفاً معنایی است، تعاریف هیچ‌یک از طرفین درست یا نادرست نیستند، فقط این تعاریف متفاوتند. در بین ارزیاب‌ها، «تأثیر» معمولاً به سطح نهایی زنجیره علی اشاره دارد، که با نتایج متفاوت است (وایت، ۲۰۱۰، ۱۵۴)، زیرا «تأثیر» به اثرات بلندمدت اشاره دارد (پیرزمن، گایت و پازن^۷، ۲۰۱۵، ۶؛ پیرزمن^۸، ۲۰۱۴، ۵؛ بمبرگر^۹، ۲۰۱۲، ۳؛ پرین^{۱۰}، ۲۰۱۲، ۱؛ راجرز^{۱۱}، ۲۰۱۲، ۲؛ و وایت، ۲۰۱۰، ۱۵۴). به عنوان مثال، تعریف ارائه‌شده توسط شبکه ارزشیابی کمیته کمک به توسعه^{۱۲} سازمان همکاری و توسعه اقتصادی^{۱۳} از تأثیر عبارت است از «اثرات مثبت و منفی، اولیه و ثانویه بلندمدتی است که توسط یک مداخله، به طور مستقیم یا غیر مستقیم، عمدی (خواسته) یا ناخواسته ایجاد می‌شود» (پیرزمن و همکاران، ۲۰۱۵، ۶؛ پیرزمن، ۲۰۱۴، ۱؛ بمبرگر، ۲۰۱۲، ۳؛ پرین، ۲۰۱۲، ۱؛ راجرز، ۲۰۱۲، ۲؛ و وایت، ۲۰۱۰، ۱۵۴). از این‌رو، بسیاری از مطالعات را می‌توان به عنوان ارزشیابی تأثیر در نظر گرفت، زیرا آنها نتایج و شاخص‌های تأثیر را مورد بحث قرار می‌دهند، در حالی که هیچ تلاشی برای نسبت دادن تغییرات در آن شاخص‌ها به مداخله انجام نمی‌دهند. در واقع چنین مطالعاتی اغلب به صراحت بیان می‌کنند که اسناد امکان‌پذیر نیست. در نتیجه چنین مطالعاتی برای کسی که اسناد علی برای او مشخصه تعیین‌کننده‌است، ارزشیابی تأثیر نیست. به ویژه آنهایی که از

1- Michel & Schneider

2- Frölich, & Sperlich

3- Gugerty, Karlan, & Welsh

4- Baylis, Honey-Rosés, Börner, Corbera, Ezzine-de-Blas, Ferraro, ... & Wunder

5- White

6- Copestake

7- Peersman, Guijt, & Pasanen

8- Peersman

9- Bamberger

10- Perrin

11- Rogers

12- The Development Assistance Committee (DAC)

13- Organisation for Economic Co-operation and Development (OECD)

روش‌های کمی برای اندازه‌گیری تأثیر استفاده می‌کنند (وایت، ۲۰۱۰، ۱۵۴). این نکته قابل توجه است که تأثیرات دیرتر از نتایج میانی و در نتیجه آنها رخ می‌دهد. تمایز بین نتایج و تأثیرات می‌تواند نسبی باشد و به اهداف بیان‌شده مداخله بستگی دارد (بمبرگر، ۲۰۱۲، ۳؛ پرین، ۲۰۱۲، ۱؛ و راجرز، ۲۰۱۲، ۲). در عمل، اغلب مفید است که یک ارزشیابی هم نتایج و هم تأثیرات را در برگیرد (راجرز، ۲۰۱۲، ۲)؛ بنا به این تعریف، هر ارزشیابی که به شاخص‌های تأثیر (یا حداقل پیامدها) اشاره می‌کند، یک ارزشیابی تأثیر است. از این رو، نظرات بر نتیجه نیز می‌تواند تحت عنوان ارزشیابی تأثیر قرار گیرد. در مقابل اما بسیاری از افرادی که روی ارزشیابی تأثیر کار می‌کنند، برای مثال در بانک جهانی، این تعریف را به اشتراک نمی‌گذارند و از نظر آنها تأثیر به عنوان تفاوت در متغیر موردنظر (Y)، با مداخله (Y_1) و بدون مداخله (Y_0) تعریف می‌شود. یعنی تأثیر $Y_1 - Y_0 =$ (وایت، ۲۰۱۰، ۱۵۴؛ راولیون^۱، ۲۰۰۸). بنابراین از این منظر می‌توان گفت منظور از ارزشیابی تأثیر، هر ارزشیابی‌ای است که به طور سیستماتیک و تجربی تأثیرات ایجادشده توسط یک مداخله را (که می‌تواند به آن نسبت داده‌شود) بررسی می‌کند (بمبرگر، ۲۰۱۲، ۳؛ پرین، ۲۰۱۲، ۱؛ و راجرز، ۲۰۱۲، ۲)؛ لذا، ارزشیابی تأثیر نه تنها تغییرات رخ داده را اندازه‌گیری یا توصیف می‌کند، بلکه به دنبال درک نقش یک مداخله خاص در ایجاد این تغییرات است. به عبارت دیگر، ارزشیابی تأثیر به دنبال اسناد علی است (گرتلر، مارتینز، پرماند، رولینگز و ورمش^۲، ۲۰۱۶، ۸؛ پرین، ۲۰۱۲، ۹؛ پیرزمن و همکاران، ۲۰۱۵، ۶؛ پیرزمن، ۲۰۱۴، ۵؛ راجرز، ۲۰۱۲، ۲). اسناد علی توسط کمیته کمک به توسعه چنین تعریف می‌شود: «ترسیم پیوند علی بین تغییرات مشاهده‌شده یا مورد انتظار با یک مداخله خاص» به گونه‌ای که سایر علل و عوامل زمینه‌ای را نیز در نظر گیرد (پیرزمن و همکاران، ۲۰۱۵، ۶). تعاریف یادشده، دو گونه متفاوت از تعریف ارزشیابی تأثیرند که هیچ‌یک به طور کامل درست یا نادرست نیست (وایت، ۲۰۱۰، ۱۵۴). وظیفه اصلی یک ارزیاب اندازه‌گیری تأثیر، بررسی تأثیرات یک مداخله به گونه‌ای است که افراد مرتبط بتوانند تصمیم بگیرند که آیا مداخله برنامه ارزش حمایت دارد و آیا برنامه باید ادامه یابد، گسترش یابد یا منحل شود (گرتلر و همکاران، ۲۰۱۶، ۲۶؛ و کندکر، کول‌وال، و صمد^۳، ۲۰۰۹).

ارزشیابی تأثیر نیز مانند پژوهش، تحت تأثیر سه رویکرد کلی کمی، کیفی و آمیخته است (پوزارنی و برینگتون^۴، ۲۰۱۶، ۱؛ کوپستیک، ۲۰۱۴، ۴۱۸؛ پیرزمن، ۲۰۱۴، ۴؛ بمبرگر، ۲۰۱۲، ۳ و ۴؛ کندکر و همکاران، ۲۰۰۹، ۱۹؛ گاربارینو و هالند^۵، ۲۰۰۹، ۱۱)، چرا در این زمینه اعداد و ارقام به تنهایی نمی‌توانند توضیح دهند که تأثیر چگونه است (میشل و اشنایدر، ۲۰۲۵، ۲). در این میان، رویکردهای کمی یا تأییدی به طور قابل توجهی در پیشینه ارزشیابی تأثیر برجسته‌ترند، در حالی که رویکردهای کیفی یا اکتشافی با تحقیقات

1- Ravallion

2- Gertler, Martinez, Premand, Rawlings, & Vermeersch

3- Khandker, Koolwal, & Samad

4- Pozarny & Barrington

5- Garbarino & Holland

تأثیر از نوع بازتر و طبیعی‌تر پیوند نزدیک‌تری دارند. رویکردهای کمی می‌توانند با اتخاذ یک نگاه فرضی - قیاسی مبتنی بر نظریه‌های از پیش تعیین‌شده برای موارد تغییر، سخت‌گیری بیشتری داشته باشند؛ در مقابل، رویکردهای کیفی به اجتناب از سوگیری و در نظر گرفتن پیامدهای ناخواسته، اهمیت بیشتری می‌دهند (کوپستیک، ۲۰۱۴، ۴۲۲ و ۴۲۳). همچنین داده‌های کیفی مانند درک بافت اجتماعی-فرهنگی و نهادی محلی، و همچنین جزئیات برنامه و شرکت‌کنندگان، برای یک ارزشیابی کمی مناسب ضروری است (کندکر و همکاران، ۲۰۰۹، ۱۸). این رویکرد موارد مذکور را با روش تولید شواهد تجربی متمایز از تأثیر بر اساس اظهارات علی‌روایتی ذی‌نفعان پروژه موردنظر بدون نیاز به مصاحبه با گروه کنترل پیگیری می‌کند. شواهد انتساب به جای تکیه بر استنتاج آماری مبتنی بر مواجهه متغیر با X ، از طریق توضیح خود پاسخ‌دهندگان از سازوکارهای علی مرتبط با X ، و از طریق توجه دقیق به جزئیات روش، تحلیل داده‌ها و شفافیت رویه‌ها انجام می‌شود (کوپستیک، ۲۰۱۴، ۴۱۸). با این وجود یک ارزشیابی کیفی به تنهایی نمی‌تواند نتایج را در مقابل جایگزین‌های مرتبط یا نتایج خلاف واقع ارزشیابی کند؛ یعنی واقعا نمی‌تواند نشان دهد که در غیاب برنامه چه اتفاقی می‌افتد (کندکر و همکاران، ۲۰۰۹، ۱۹). به طور کلی به ندرت یک روش ارزشیابی واحد وجود دارد که بتواند تمام پیچیدگی‌های نحوه عملکرد برنامه‌ها در دنیای واقعی را به طور کامل نشان دهد (بمبرگر، ۲۰۱۲، ۳). در نتیجه، ضروری است ارزشیابی از طیف وسیعی از روش‌های کمی و کیفی برای ارائه اطلاعات جامع‌تری در مورد آنچه اتفاق می‌افتد، چرایی و مناسب بودن یا نبودن آن و راهنمایی برای مسیرهای آتی استفاده کند (پوزارنی و برینگتون، ۲۰۱۶، iii؛ پرین، ۲۰۱۲، ۴). بنابراین ارزیابان باید راه‌های خلاقانه‌ای را برای ترکیب چارچوب‌ها، ابزارها یا فنون ارزشیابی مختلف ببابند. ویژگی منحصر به فرد رویکردهای آمیخته این است که آنها به دنبال ادغام موضوع مورد پژوهش یا ارزشیابی با رویکردهای عمدتاً کمی و عمدتاً کیفی در نظریه، جمع‌آوری و تحلیل و تفسیر داده‌ها هستند (بمبرگر، ۲۰۱۲، ۳ و ۴). اگرچه بسیاری از ارزیابان در حال حاضر به طور معمول از روش‌های مختلفی استفاده می‌کنند، «آنچه که ارزشیابی با رویکرد آمیخته را متمایز می‌کند، استفاده عمدی یا برنامه‌ریزی شده از روش‌های متنوع برای اهداف ترکیبی خاص با استفاده از طرح‌های آمیخته خاص است» (گرین، ۲۰۰۵). به بیانی دیگر آنچه «ارزشیابی با رویکردهای آمیخته» را متمایز می‌کند، ادغام نظام‌مند روش‌های کمی و کیفی در تمام مراحل ارزشیابی است (پیرزمن، ۲۰۱۴، ۴؛ کندکر و همکاران، ۲۰۰۹، ۱۹). یک دلیل کلیدی برای ترکیب روش‌ها این است که این ادغام به غلبه بر نقاط ضعف ذاتی هر روش در صورت استفاده مجزا (پیرزمن، ۲۰۱۴، ۴؛ بمبرگر، ۲۰۱۲، ۴) کمک می‌کند. همچنین اعتبار یافته‌های ارزشیابی را با نمایان کردن همگرایی داده‌ها از طریق چندسویه‌سازی افزایش می‌دهد (صالحی و گل‌افشانی، ۲۰۱۰، ۱۸۷)، و از این طریق می‌تواند درک برنامه/سیاست، اثرات و زمینه آن را عمیق‌تر کند (پیرزمن، ۲۰۱۴، ۴). بنابراین هدف

رویکرد آمیخته استفاده از نقاط قوت هر دو رویکرد کمی و کیفی و ادغام آنها برای غلبه بر نقاط ضعف آنهاست (بمبرگر، ۲۰۱۲، ۴؛ صالحی و گل‌افشانی، ۲۰۱۰، ۱۸۷).

علاوه بر تقسیم ارزشیابی تأثیر بر اساس رویکرد، ارزشیابی تأثیر را می‌توان بر اساس **زمان** به دو دسته آینده‌نگر و گذشته‌نگر تقسیم کرد (گرتلر و همکاران، ۲۰۱۶، ۹؛ کندکر و همکاران، ۲۰۰۹، ۲۰). ارزشیابی‌های آینده‌نگر، همزمان با طراحی برنامه ایجاد می‌شوند و در اجرای برنامه گنجانده می‌شوند. داده‌های پایه قبل از اجرای برنامه برای هر دو گروه دریافت‌کننده مداخله و گروه مقایسه جمع‌آوری می‌شود (گرتلر و همکاران، ۲۰۱۶، ۹)؛ در این نوع ارزشیابی تلاش می‌شود تأثیرات موردنظر برنامه‌ها و سیاست‌های آینده را، با توجه به وضعیت فعلی یک منطقه هدف‌گذاری شده بالقوه، اندازه‌گیری کند، این ارزشیابی‌ها فواید فوری دارند و واقعیت را منعکس می‌کنند (کندکر و همکاران، ۲۰۰۹، ۲۰). به طور خلاصه به سه دلیل، ارزشیابی تأثیر آینده‌نگر به احتمال زیاد نتایج ارزشیابی قوی و معتبری را به همراه دارد: اول، داده‌های پایه را می‌توان برای تعیین ملاک‌های نتایج موردنظر قبل از شروع برنامه جمع‌آوری کرد. دوم، تعریف ملاک‌های موفقیت یک برنامه در مرحله برنامه‌ریزی، هم برنامه و هم ارزشیابی را بر نتایج موردنظر متمرکز می‌کند. سوم، در یک ارزشیابی آینده‌نگر، گروه‌های مداخله و مقایسه قبل از اجرای مداخله در حال ارزشیابی مشخص می‌شوند و بهترین شانس را برای ایجاد خلاف واقع معتبر دارند (گرتلر و همکاران، ۲۰۱۶، ۱۰). در مقابل ارزشیابی‌های گذشته‌نگر، تأثیر برنامه را پس از اجرای برنامه ارزشیابی می‌کنند (گرتلر و همکاران، ۲۰۱۶، ۹)، یعنی این نوع ارزشیابی، تأثیرات واقعی را که توسط ذی‌نفعان به مداخله برنامه نسبت داده می‌شود، اندازه‌گیری می‌کند (کندکر و همکاران، ۲۰۰۹، ۲۱)، و به طور کلی این نوع ارزشیابی با استفاده از داده‌های موجود برای ارزشیابی برنامه‌هایی که در گذشته ایجاد شده‌اند ضروری است (گرتلر و همکاران، ۲۰۱۶، ۱۰)؛ این ارزشیابی می‌تواند بسیار پرهزینه‌تر از ارزشیابی‌های آینده‌نگر باشد، زیرا به جمع‌آوری داده‌هایی در مورد نتایج واقعی، و همچنین سایر عوامل اجتماعی و اقتصادی همراهی که ممکن است مسیر مداخله را تعیین کرده باشند، نیاز دارد. یک هزینه اضافی در این زمینه نیز پس از اتمام، شکست مداخله است، که ممکن است از طریق تحلیل قبلی پیش‌بینی شده باشد (کندکر و همکاران، ۲۰۰۹، ۱۹). در این نوع ارزشیابی‌ها، تیمی که ارزشیابی را انجام می‌دهد، اغلب اطلاعات محدودی دارند و از این‌رو تحلیل اینکه آیا برنامه با موفقیت اجرا شده و آیا شرکت‌کنندگان واقعاً از آن سود برده‌اند یا خیر، مشکل است (گرتلر و همکاران، ۲۰۱۶، ۱۰)، و همچنین ممکن است هزینه‌های اضافی در پی داشته باشد. یکی از رویکردها ترکیب هر دو تحلیل و مقایسه تخمین‌های پس از آن با پیش‌بینی‌های قبلی است (راوالیون، ۲۰۰۸). این رویکرد می‌تواند به توضیح چگونگی ظهور مزایای برنامه کمک کند، به‌ویژه اگر برنامه در مراحل مختلف اجرا شود و انعطاف‌پذیر باشد. به طور کلی اگر محققین طراحی قبلی از ارزشیابی تأثیر داشته باشند، انجام

یک ارزشیابی پس از مداخله آسان‌تر است؛ یعنی اگر قبل از اجرای مداخله، طرحی برای ارزشیابی تأثیر برنامه‌ریزی شود، ارزشیابی دقیق‌تر و مؤثرتر است (کندکر و همکاران، ۲۰۰۹، ۱۹).

علاوه بر دسته‌بندی ارزشیابی تأثیر بر اساس رویکرد و زمان، ارزشیابی تأثیر بر اساس **ارزیابان** به سه دسته داخلی، خارجی و مشارکتی تقسیم می‌شود (گرتلر و همکاران، ۲۰۱۶، ۲۷؛ پیرزمن و همکاران، ۲۰۱۵، ۹؛ راجرز، ۲۰۱۴، ۳ و ۱۰؛ راجرز، ۲۰۱۲، ۳). منظور این است که ارزشیابی تأثیر را می‌توان توسط کارکنان سازمان انجام داد، یا می‌توان آن را با ارزیاب‌های خارجی قرارداد بست و یا با همکاری و مشارکت هر دو گروه ارزیابان داخلی و خارجی انجام داد؛ اما به طور کلی باید توجه داشت تیم در حالت ایده‌آل باید هم شامل ارزیاب و هم متخصص موضوع باشد (پیرزمن و همکاران، ۲۰۱۵، ۹؛ راجرز، ۲۰۱۲، ۳). برای مجریان پرمشغله برنامه، اختصاص یک تیم خارجی با مسئولیت طراحی و اجرای ارزشیابی تأثیر اغلب جذاب است و از طرفی نیز برون‌سپاری می‌تواند ارزشیابی تأثیر را مستقل‌تر کند. با این حال، برون‌سپاری کامل می‌تواند با خطرات قابل توجهی همراه باشد (گرتلر و همکاران، ۲۰۱۶، ۲۰۸). از یک طرف ارزیابان خارجی همیشه بی‌طرف تلقی نمی‌شوند (راجرز، ۲۰۱۲، ۳)، و از طرفی نیز انتخاب و نظارت بر تیم طرف قرارداد می‌تواند برای واحد اجرای برنامه، چالش‌برانگیز باشد. برای کاهش این خطرات، معمولاً ترجیح داده می‌شود که تیم خط‌مشی از قبل، یک طرح ارزشیابی تأثیر تهیه کنند و یا شیوه مشارکتی و به خصوص نظارت بر کار ارزیابان خارجی را در پیش گیرند (گرتلر و همکاران، ۲۰۱۶، ۲۷).

حال پس از شناسایی انواع رویکردهای ارزشیابی تأثیر ضروری است بدانیم چرا ارزشیابی تأثیر انجام می‌دهیم؟ برخی از دلایل رایج برای انجام ارزشیابی تأثیر شامل؛ اطلاع‌رسانی به سیاست‌گذاران در مورد طیف وسیعی از تصمیم‌ها (گرتلر و همکاران، ۲۰۱۶، ۲۰۸)، تصمیم‌گیری در مورد تأمین مالی یک مداخله (با ارزشیابی تأثیر آینده‌نگر و تخمین هزینه‌ها) (راجرز، ۲۰۱۲، ۳)، تصمیم‌گیری در مورد تکرار، ادامه، گسترش، و یا توقف مداخله (گیوگتی و همکاران، ۲۰۱۶، ۷؛ گرتلر و همکاران، ۲۰۱۶، ۲۴۰؛ پیرزمن و همکاران، ۲۰۱۵، ۸؛ پیرزمن، ۲۰۱۴، ۱؛ راجرز، ۲۰۱۴، ۲)، محدود کردن برنامه‌های ناکارآمد و افزایش مداخلات مؤثر (گرتلر و همکاران، ۲۰۱۶، ۲۱)، یادگیری نحوه تکرار یا افزایش مقیاس یک پایلوت (پیرزمن، ۲۰۱۴، ۱؛ راجرز، ۲۰۱۴، ۲)، یادگیری نحوه انطباق موفقیت‌آمیز یک مداخله موفق با زمینه دیگر، اطمینان بخشیدن به سرمایه‌گذاران (پاسخگویی به بالادستی‌ها)، و اطلاع‌رسانی به ذی‌نفعان و جوامع موردنظر (پاسخگویی به پایین‌دستی‌ها)، است (راجرز، ۲۰۱۲، ۳).

در این نوع ارزشیابی نیز مانند تمام پژوهش‌ها و ارزشیابی‌ها ضروری است به یکسری اصول اخلاقی توجه و از آنها پیروی کنیم (گرتلر و همکاران، ۲۰۱۶، ۲۳۲؛ پیرزمن و همکاران، ۲۰۱۵، ۸؛ راجرز، ۲۰۱۲، ۱۴). از جمله این مسائل اخلاقی می‌توان به انتخاب شاخص مناسب برای اولویت‌بندی ذی‌نفعان در دسترسی آنها به برنامه، ارائه نتایج بی‌طرفانه، قابل اعتماد و معتبر (گرتلر و همکاران، ۲۰۱۶، ۲۳۱)، پرهیز

از انتشار داده‌های شخصی و محرمانه (گرتلر و همکاران، ۲۰۱۶، ۲۳۷؛ راجرز، ۲۰۱۴، ۱۴؛ راجرز، ۲۰۱۲، ۱۴)، ارائه جزئیات برای حمایت از ادعای مطرح‌شده، شفافیت در روش کار و نتایج، جلب رضایت آگاهانه شرکت‌کنندگان، رعایت حقوق و رفاه افراد درگیر، اختصاص یک شماره شناسایی رمزگذاری‌شده به هر فرد مشارکت‌کننده (گرتلر و همکاران، ۲۰۱۶، ۲۳۷)، آسیب نرساندن به شرکت‌کنندگان (گرتلر و همکاران، ۲۰۱۶، ۲۳۲؛ راجرز، ۲۰۱۴، ۱۴)، و در نظر گرفتن منفعت عمومی جامعه (راجرز، ۲۰۱۲، ۱۵)، اشاره نمود.

همچنین ضروری است پیش از شروع کار و در مراحل طراحی و برنامه‌ریزی کار، تیم ارزشیابی، متناسب با اهداف و شرایط ارزشیابی تأثیر، یک طرح ارزشیابی که شامل جنبه‌های عملیاتی و تحقیقاتی باشد، تدوین نمایند (گیوگتی و همکاران، ۲۰۱۶، ۵؛ گرتلر و همکاران، ۲۰۱۶، ۲۷، ۲۳۶، ۲۱۰؛ پیرزمن، ۲۰۱۴، ۵؛ راجرز، ۲۰۱۴، ۵؛ راجرز، ۲۰۱۲، ۱۵ و ۱۷؛ کندکر و همکاران، ۲۰۰۹، ۱۹)؛ این طرح می‌تواند شامل مواردی همچون روش‌های گردآوری و تحلیل داده‌ها (گرتلر و همکاران، ۲۰۱۶، ۲۳۶؛ راجرز، ۲۰۱۴، ۸)، راهبرد شناسایی، شاخص‌های اصلی نتیجه، اندازه‌های نمونه تقریبی، مشخص نمودن نقش‌ها و مسئولیت‌های تیم، قوانین کلیدی عملیاتی، خطرات عملیاتی بالقوه (گرتلر و همکاران، ۲۰۱۶، ۲۱۰)، شناسایی کاربران موردنظر، سؤال‌های ارزشیابی کلیدی، شفاف‌سازی ارزش‌ها، توسعه نظریه تغییر، توصیف متغیرهای مهم (راجرز، ۲۰۱۲، ۱۷)، جزئیاتی در مورد چگونگی، مکان، زمان، جامعه، راهبرد نمونه‌گیری و نمونه، منابع مالی و تخمین بودجه، نوع مداخله و اسناد علی، گروه‌های مقایسه، سؤال اصلی ارزشیابی و روش‌های پاسخگویی به آن (راجرز، ۲۰۱۴، ۶)، باشد. از سوی دیگر پس از انجام کار و تکمیل ارزشیابی ضروری است تیم ارزشیابی بدون توجه به مثبت یا منفی بودن نتایج، یک گزارش ارزشیابی کامل، جامع، منطقی، منصفانه و شفاف از فرآیند و نتایج ارزشیابی با رعایت اصول اخلاقی ارزشیابی تدوین و منتشر نمایند (گرتلر و همکاران، ۲۰۱۶، ۲۰۴؛ پیرزمن و همکاران، ۲۰۱۵، ۱۲؛ راجرز، ۲۰۱۴، ۱۲؛ کندکر و همکاران، ۲۰۰۹، ۲۸). قالب گزارش ارزشیابی باید در زمانی که ارزشیابی تأثیر برنامه‌ریزی می‌شود مورد توافق قرار گیرد (راجرز، ۲۰۱۲، ۱۳). گزارش‌ها باید با طول متوسط، در محدوده ۳۰ تا ۵۰ صفحه، شامل چکیده‌ای کمتر از یک صفحه، و یک خلاصه اجرایی ۲ تا ۴ صفحه‌ای همراه با اهداف، نتایج، توصیه‌ها، و پیوست‌ها (شامل جزئیات فنی، مستندات مرتبط، ابزار گردآوری داده‌ها و همچنین تحلیل‌ها) نگارش شود (گرتلر و همکاران، ۲۰۱۶، ۲۵۴). کیفیت گزارش‌های ارزشیابی را می‌توان با مشارکت مناسب ذی‌نفعان افزایش داد (راجرز، ۲۰۱۲، ۱۳).

یکی از موارد مهم در ارزشیابی تأثیر، **نظریه تغییر**^۱ است در واقع می‌توان گفت یک ابزار ضروری در ارزشیابی تأثیر، یک نظریه تغییر است که به خوبی توسعه یافته‌است (میشل و اشنايدر، ۲۰۲۵، ۲؛ گیوگتی و همکاران، ۲۰۱۶، ۵؛ پیرزمن، ۲۰۱۴، ۱ و ۲؛ راجرز، ۲۰۱۴، ۴)، و موجب بهبود فرآیند ارزشیابی تأثیر می‌شود (پیرن، ۲۰۱۲، ۸). بنابراین اغلب مفید است که ارزشیابی تأثیر را بر اساس یک نظریه یا مدلی از

نحوه درک مداخله برای ایجاد تأثیرات موردنظر خود قرار دهیم (راجرز، ۲۰۱۲، ۶). نظریه تغییر، مدلی است که توضیح می‌دهد چگونه انتظار می‌رود که یک مداخله به اثرات موردنظر یا مشاهده‌شده منجر شود (پرین، ۲۰۱۲، ۸)؛ این نظریه که اغلب به عنوان نظریه برنامه، زنجیره نتایج، مدل منطقی برنامه یا منطق اسناد شناخته می‌شود (پرین، ۲۰۱۲، ۸؛ راجرز، ۲۰۱۲، ۶)، توضیح می‌دهد که برنامه یا خط‌مشی چگونه کار می‌کند یعنی یک مدل علی را به تصویر می‌کشد که روابط مفروض میان ورودی‌ها، فعالیت‌ها، خروجی‌ها، نتایج کوتاه‌مدت، نتایج بلندمدت و اثرات را به تصویر می‌کشد (پیرزمن، ۲۰۱۴، ۱ و ۲؛ پرین، ۲۰۱۲، ۸؛ راجرز، ۲۰۱۲، ۷؛ و وایت، ۲۰۱۰، ۱۵۳). نظریه تغییر به وضوح عناصر برنامه و نتایج موردنظر را مشخص می‌کند (گرتلر و همکاران، ۲۰۱۶، ۲۰۴؛ گیوگتی و همکاران، ۲۰۱۶، ۵)، و به دو دلیل کلیدی ضروری است. اول، یک نظریه واضح تغییر، داستانی را در مورد چگونگی تبدیل برنامه، منابع و فعالیت‌ها به خروجی‌ها و از نظر نظری، نتایج و تأثیرات بیان می‌کند. ترسیم این فرآیند برای شناسایی سؤال مناسب برای ارزشیابی ضروری است و به طراحی مطالعه اطلاع می‌دهد. دوم، هر ارزشیابی تأثیر جدید باید دانش را در مورد اثربخشی یک مداخله معین گسترش دهد (گیوگتی و همکاران، ۲۰۱۶، ۵). این نظریه توجه را بر چگونگی ایجاد تغییر متمرکز می‌کند و برای نمایش برنامه‌هایی که در آن فعالیت‌های مختلف در امتداد زنجیره علی رخ می‌دهند، مفید است (راجرز، ۲۰۱۲، ۷). این ابزار نه تنها در مرحله طراحی برنامه مفید است، بلکه به تمرکز ارزشیابی تأثیر بر آنچه که ذی‌نفعان باید در مورد برنامه یا خط‌مشی برای حمایت از تصمیم‌گیری بدانند کمک می‌کند (پیرزمن، ۲۰۱۴، ۲). فرآیند توسعه نظریه تغییر و توافق بر سر سؤالات برای ارزشیابی می‌تواند چالش‌برانگیز باشد (گیوگتی و همکاران، ۲۰۱۶، ۵). بهترین کار این است که نظریه تغییر را به عنوان بخشی از برنامه‌ریزی یک مداخله توسعه دهیم، و سپس در حین برنامه‌ریزی برای ارزشیابی تأثیر، آن را بررسی و در صورت لزوم بازبینی کنیم (راجرز، ۲۰۱۲، ۶). در این فرآیند توصیه می‌شود با اتخاذ رویکرد مشارکتی و مشارکت دادن کارکنان، سهام‌داران اصلی و سایر ذی‌نفعان کلیدی در توسعه نظریه تغییر، درک مشترکی از مداخله ایجاد کنید (میشل و اشنايدر، ۲۰۲۵، ۲؛ پرین، ۲۰۱۲، ۹). اگر یک نظریه تغییر فقط توسط یک متخصص با تکیه گسترده بر مستندات ایجاد شود، ممکن است مقبولیت و مالکیت مدل توسط سایر ذی‌نفعان وجود نداشته‌باشد، و همچنین ممکن است واقعیت آنچه در حال وقوع است را منعکس نکند (پرین، ۲۰۱۲، ۹).

از چالش‌های ارزشیابی تأثیر می‌توان به ایجاد گروه مقایسه مناسب (گیوگتی و همکاران، ۲۰۱۶، ۵؛ گرتلر و همکاران، ۲۰۱۶، ۸؛ بیلیس و همکاران، ۲۰۱۶، ۵۸؛ کندکر و همکاران، ۲۰۰۹، ۲۲)، غلبه بر مقاومت داخلی و خارجی در برابر ارزشیابی تأثیر، همسو نمودن انگیزه و اهداف کارکنان، تفکیک نحوه اجرای برنامه و محتوای برنامه در ارزشیابی، توافق روی سؤال درست ارزشیابی (گرتلر و همکاران، ۲۰۱۶، ۲۷؛ گیوگتی و همکاران، ۲۰۱۶، ۵؛ پیرزمن و همکاران، ۲۰۱۵، ۹)، پاسخ به سؤالات راهبردی سازمان، زمان‌سنجی و

انتخاب زمان مناسب ارزشیابی (گیوگتی و همکاران، ۲۰۱۶، ۶؛ گرتلر و همکاران، ۲۰۱۶، ۲۰۴؛ پیرزمن و همکاران، ۲۰۱۵، ۸ و ۱۰؛ پرین، ۲۰۱۲، ۷)، مقایسه مزایا و هزینه‌های گردآوری داده‌ها و ارزشیابی (گیوگتی و همکاران، ۲۰۱۶، ۳؛ گرتلر و همکاران، ۲۰۱۶، ۲۷ و ۲۵۳؛ راجرز، ۲۰۱۴، ۲؛ راجرز، ۲۰۱۲، ۲)، ایجاد سرمایه داخلی و جذب سرمایه خارجی، اشاره کرد (گیوگتی و همکاران، ۲۰۱۶، ۴). همانگونه که ذکر شد، زمان ارزشیابی یک مسأله چالش برانگیز و مهم است. اگر ارزشیابی خیلی زود انجام شود، ممکن است منجر به عدم دقت شود (پیرزمن و همکاران، ۲۰۱۵، ۸)، شواهد کافی مبنی بر وقوع یا تداوم اثرات وجود نداشته باشد (راجرز، ۲۰۱۲، ۱۴)، خطر یافتن تأثیرات جزئی یا بدون تأثیر وجود داشته باشد (گرتلر و همکاران، ۲۰۱۶، ۲۱۴)، و همچنین تصویری نادرست از تأثیرات ارائه شود. یعنی در برخی موارد، تأثیرات کمتر از حد اعلام می‌شود، زیرا برنامه، زمان کافی برای رشد نداشته است و در برخی موارد ممکن است تأثیرات اغراق آمیز باشد و در طول زمان کاهش یابد (راجرز، ۲۰۱۴، ۳؛ پرین، ۲۰۱۲، ۵ و ۹). همچنین زمانی که ارزشیابی تأثیر با تأخیر انجام می‌شود، این خطر وجود دارد که برنامه ممکن است حمایت مالی و عمومی را از دست بدهد یا یک برنامه بد طراحی شده ممکن است گسترش یابد (گرتلر و همکاران، ۲۰۱۶، ۲۱۴). یعنی اطلاعات مورد نیاز برای تصمیم‌گیری را خیلی دیر ارائه می‌کند و این یک مخاطره جدی است (راجرز، ۲۰۱۴، ۳؛ راجرز، ۲۰۱۲، ۱۴). به طور کلی ضروری است در زمان‌بندی گردآوری داده‌های پیگیری مشخص کنیم که پس از اجرای برنامه چقدر زمان برای آشکارشدن نتایج لازم است (گرتلر و همکاران، ۲۰۱۶، ۲۱۴؛ پیرزمن و همکاران، ۲۰۱۵، ۸؛ راجرز، ۲۰۱۴، ۳؛ پرین، ۲۰۱۲، ۹؛ راجرز، ۲۰۱۲، ۱۴ و ۱۵). به طور کلی بهترین زمان برای ارزشیابی تأثیر، پیش از گسترش برنامه و زمانی است که برنامه به اندازه کافی بالغ و پایدار شده باشد (گرتلر و همکاران، ۲۰۱۶، ۲۱۴). همچنین زنجیره نتایج برنامه به شناسایی شاخص‌های نتیجه و زمان مناسب برای اندازه‌گیری آنها کمک می‌کند (گرتلر و همکاران، ۲۰۱۶، ۲۱۴؛ راجرز، ۲۰۱۲، ۶). هدف برخی از برنامه‌ها ارائه مزایای کوتاه‌مدت است، در حالی که برخی دیگر به دنبال منافع بلندمدت هستند. همچنین برخی نتایج مانند نرخ باروری ناشی از اصلاحات بهداشتی، به دلیل ماهیت‌شان به زمان بیشتری برای ظهور نیاز دارند (گرتلر و همکاران، ۲۰۱۶، ۲۱۴؛ پرین، ۲۰۱۲، ۹).

برای روشن‌شدن مباحث پژوهش حاضر نیز ضروری است بدانیم منظور از فرا ارزشیابی، فرآیند طراحی، جمع‌آوری و کاربرد داده‌ها به منظور قضاوت در مورد صلاحیت ارزشیابی است و عموماً نقاط قوت و ضعف را گزارش می‌کند (سبزواری، ۱۳۸۴، ۱۱۷). به بیانی دیگر فرا ارزشیابی عبارت است از فرایند گردآوری، تهیه و کاربرد اطلاعات توصیفی و قضاوت در مورد سودمندی، عملی‌بودن، درستی، و دقت، به منظور ارزشیابی و گزارش نقاط قوت و نقاط قابل بهبود ارزشیابی‌هاست. در واقع فرا ارزشیابی عبارت از ارزشیابی از سیستم ارزشیابی، ابزار ارزشیابی، فرآیند ارزشیابی یا برنامه ارزشیابی است (رسولی و ذوالفقاری، ۱۳۹۶، ۱). پژوهش حاضر نیز با هدف ارزشیابی از مطالعات ارزشیابی‌های تأثیر انجام شده است. با توجه به نکات

ذکر شده و اهمیت نقش ارزشیابی تأثیر در حوزه‌های گوناگون از یک‌سو، و همچنین نبود ابزاری برای سنجش کیفیت مطالعات ارزشیابی تأثیر و در نتیجه فقدان شناخت درباره کیفیت مطالعات ارزشیابی تأثیر انجام‌شده از دیگرسو، ضروری است در نخستین گام پژوهشگران و ارزیابان علاقه‌مند به حوزه ارزشیابی تأثیر، از یک فهرست‌وارسی^۱ برای فرا ارزشیابی تأثیر مطالعات انجام‌شده استفاده نمایند؛ اما مرور پیشینه در زمینه ارزشیابی تأثیر نشان داد که به طور کلی در میان پژوهش‌های داخلی، هیچ مطالعه‌ای به ارائه فهرست‌وارسی برای ارزشیابی کیفیت مطالعات ارزشیابی تأثیر پرداخته‌است. همچنین در میان پژوهش‌های بین‌المللی، به‌رغم توجه جامعه علمی جهانی از سال‌های پیشین تا کنون و انجام مطالعات ارزشمندی در این زمینه، با فقر مطالعات به‌روز و جدید روبرو هستیم. از این‌رو پژوهشگران در این مطالعه با هدف «تدوین فهرست‌وارسی برای ارزشیابی کیفیت مطالعات ارزشیابی تأثیر»، تلاش نموده‌اند تا با تحلیل اسناد و مشورت با متخصصان حوزه روش تحقیق، به سه سؤال زیرینایی «اصلی‌ترین گویه‌ها در فرا ارزشیابی تأثیر چیست؟»، «مطالعات ارزشیابی تأثیر داخلی موجود از چه کیفیتی برخوردارند؟» و «آیا وضعیت تحقق گویه‌های تدوین‌شده برای ارزشیابی کیفیت مطالعات ارزشیابی تأثیر در حد تسلط است؟» پاسخ دهند.

پیشینه پژوهش

نخستین کار مهم در هر پژوهشی، روی آوردن به کارهای نظری و تجربی پیشین در حوزه موردنظر است. بنابراین با جستجو در منابع داخلی و بین‌المللی، مهم‌ترین پژوهش‌های انجام‌شده را گردآوری و تدوین نمودیم؛ در حوزه تهیه فهرست‌وارسی فرا ارزشیابی تأثیر، میشل و اشنایدر (۲۰۲۵) در پژوهش «بهام‌زدایی از ارزشیابی تأثیر: یک چارچوب ارزشیابی تأثیر» به ارائه چارچوب ارزشیابی تأثیر با هفت اصل مشخص پرداخته‌است (میشل و اشنایدر، ۲۰۲۵). این پژوهش جدیدترین و مرتبط‌ترین پژوهش در حوزه ارزشیابی تأثیر است. همچنین رید، فره، مارتین اورتگا، بلانچ، لافورد-رولف، دالیمر و هولدن^۲ (۲۰۲۱)، در مقاله «ارزیابی تأثیر پژوهش: یک چارچوب روش‌شناختی» به ارائه تعریف، گونه‌شناسی، و یک چارچوب روش‌شناختی ارزشیابی تأثیر پژوهش‌ها پرداخته‌اند (رید و همکاران، ۲۰۲۱). کراتو و پارولو^۳ (۲۰۱۹)، در کتاب خود با عنوان «ارزشیابی تأثیر سیاست مبتنی بر داده» به مباحثی در مورد ارزشیابی تأثیر به خصوص با تمرکز بر سیاست و اقدامات سیاستی پرداخته‌اند (کراتو و پارولو، ۲۰۱۹). همچنین فرولیچ و اسپرلیچ (۲۰۱۹)، در کتاب خود با عنوان «ارزشیابی تأثیر» به مباحثی در مورد ارزشیابی تأثیر پرداخته‌اند و به طور خاص به ارائه یک رویکرد جامع و به‌روز برای تحلیل ارزشیابی تأثیر کمی که به عنوان استنتاج علی یا تحلیل اثر درمان نیز شناخته می‌شود، پرداخته‌اند (فرولیچ و اسپرلیچ، ۲۰۱۹). پژوهش گیوگتی و همکاران

1- Checklist

2- Reed, Ferré, Martin-Ortega, Blanche, Lawford-Rolfe, Dallimer, & Holden

3- Crato & Paruolo

(۲۰۱۶)، با عنوان «چالش‌های سازمانی ارزشیابی تأثیر» به بررسی چهار چالش عمده ارزشیابی تأثیر و ارائه راهکارهایی برای مقابله با این چالش‌ها می‌پردازد (گیوگتی و همکاران، ۲۰۱۶). گرتلر و همکاران (۲۰۱۶)، نیز در کتابی با عنوان «ارزشیابی تأثیر در عمل» به تعریف ارزشیابی تأثیر، انواع روش‌های ارزشیابی تأثیر بر اساس زمان، تصمیم‌گیری در مورد مداخله و انجام ارزشیابی تأثیر، انواع برآوردهای تأثیر و مباحثی ارزشمندی در زمینه ارزشیابی تأثیر پرداخته‌اند (گرتلر و همکاران، ۲۰۱۶). پوزارنی و برینگتون (۲۰۱۶)، در یک فصل از کتاب «روش‌های کیفی در ارزشیابی تأثیر برنامه‌های انتقال وجه نقد در پروژه انتقال در جنوب صحرای آفریقا» به ضرورت و مزایای استفاده از روش‌های کیفی در ارزشیابی تأثیر پرداخته‌اند (پوزارنی و برینگتون، ۲۰۱۶). بیلپس و همکاران (۲۰۱۶)، در پژوهش خود با عنوان «جریان اصلی ارزشیابی تأثیر در حفاظت از طبیعت» به مباحثی از جمله مقایسه منابع و هزینه‌ها، مقرون به صرفه بودن ارزشیابی، مشارکت متخصصان و مجریان برنامه ارزشیابی پرداخته‌اند (بیلپس و همکاران، ۲۰۱۶). پیرزمن و همکاران (۲۰۱۵)، در کتابی با عنوان «سنجش قابلیت ارزشیابی برای ارزشیابی تأثیر؛ راهنمایی، فهرست‌وارسی و پشتیبانی تصمیم» به ویژگی‌های خاص ارزشیابی تأثیر بر اساس سه سؤال اصلی، موارد مهم در ارزشیابی تأثیر و تأکید بر تهیه فهرست‌وارسی در ارزشیابی تأثیر، پرداخته‌اند (پیرزمن و همکاران، ۲۰۱۵). ساموئل و دریک^۱ (۲۰۱۵) در پژوهشی با عنوان «ارزشیابی تأثیر اجتماعی: بررسی ادراکات ارزیاب از توصیف تأثیر تحت REF2014» به بیان نکاتی از ارزشیابی تأثیر همچون امتیازدهی، معرفی مقیاس ستاره‌ای، توجه به باورهای ارزیابان در فرآیند ارزشیابی تأثیر در قالب ارائه نتایج ارزشیابی رسمی ارزیابان و یا همتایان پرداخته‌اند (ساموئل و دریک، ۲۰۱۵). پیرزمن (۲۰۱۴) در پژوهشی با عنوان «بررسی اجمالی: روش‌های جمع‌آوری و تحلیل داده‌ها در ارزشیابی تأثیر» به مباحث گردآوری و تحلیل داده‌ها، نکات مهم در ارزشیابی تأثیر، معرفی و توضیح نظریه تغییر، تصمیم‌گیری در مورد انتخاب روش گردآوری داده، مدیریت داده‌ها و مسائلی از این قبیل پرداخته‌است (پیرزمن، ۲۰۱۴). کوپستیک (۲۰۱۴) در پژوهش خود با عنوان «ارزشیابی تأثیر معتبر در زمینه‌های پیچیده: رویکردهای تاییدی و اکتشافی» به ارائه تعریفی مختصر از ارزشیابی تأثیر و همچنین مقایسه رویکردهای کمی و کیفی (تاییدی و اکتشافی) در ارزشیابی تأثیر پرداخته‌است (کوپستیک، ۲۰۱۴). بمبرگر (۲۰۱۲) در پژوهش خود با عنوان «مقدمه‌ای بر روش‌های ترکیبی در ارزشیابی تأثیر» به بیان نقاط ضعف استفاده منحصر از روش‌های کمی و کیفی و مزایای ترکیب این دو روش در ارزشیابی تأثیر پرداخته‌است (بمبرگر، ۲۰۱۲). راجرز (۲۰۱۲)، در پژوهشی با عنوان «مقدمه‌ای بر ارزشیابی تأثیر» به مباحثی در مورد دلایل انجام ارزشیابی تأثیر، رویکرد مشارکتی در ارزشیابی تأثیر، سؤالات کلیدی در ارزشیابی تأثیر و مواردی از این قبیل پرداخته‌است (راجرز، ۲۰۱۲). پرین (۲۰۱۲) در پژوهش خود با عنوان «پیوند نظارت و ارزشیابی به ارزشیابی تأثیر» به بیان تعاریف، تمایزات و ویژگی‌های خاص هر کدام اشاره می‌کند (پرین،

1- Samuel, & Derrick

۲۰۱۲). وایت (۲۰۱۰)، در پژوهش خود با عنوان «سهمی در بحث‌های جاری در ارزشیابی تأثیر» به بیان دو تعریف مشابه و در عین حال متمایز از ارزشیابی تأثیر و بحث در مورد آن پرداخته‌است (وایت، ۲۰۱۰). کندکر و همکاران (۲۰۰۹) در کتاب خود با عنوان «ارزشیابی تأثیر؛ روش‌های کمی و عملیات‌ها» به نقش ارزشیابی تأثیر، اهمیت روش‌های کیفی در کنار روش‌های کمی، انواع ارزشیابی تأثیر به لحاظ زمانی و مزایا و معایب هر کدام پرداخته‌اند (کندکر و همکاران، ۲۰۰۹). گارباینو و هالند (۲۰۰۹)، در پژوهش خود «روش‌های کمی و کیفی در ارزشیابی تأثیر و نتایج اندازه‌گیری» در مورد رویکردهای کمی، کیفی و ترکیبی بحث نموده‌است (گارباینو و هالند، ۲۰۰۹). همچنین در میان پژوهش‌های داخلی، پژوهش علی‌میرزایی، پیرحیاتی، و صالحی (۱۳۹۴) با عنوان «ارزشیابی تأثیر در نظام آموزش عالی: چپستی، چرای و چگونگی» به مبانی اساسی ارزشیابی تأثیر پرداخته‌است و در میان منابع فارسی بهترین منبع موجود در زمینه ارزشیابی تأثیر است (علی‌میرزایی و همکاران، ۱۳۹۴).

اما در حوزه مطالعات تجربی ارزشیابی تأثیر، بر اساس جستجوهای انجام‌شده در پایگاه داده‌های داخلی حدود ۳۰ پژوهش (تا حدودی مرتبط) وجود دارد که از آن میان می‌توان به پژوهش زارع، شجری، رحمانی، محدث‌حسینی، سالم، و انویه‌تکیه (۱۴۰۳) با عنوان «ارزشیابی تأثیر طرح سربازان سازندگی در بخش کشاورزی»، پژوهش بحرانی، فروتنی، و بوجلالی (۱۴۰۱) با عنوان «سنجش تأثیر سواد رسانه‌ای در استفاده از فضای مجازی بر اخلاق اجتماعی کاربران (مورد مطالعه: دانشجویان کارشناسی ارشد دانشگاه آزاد اسلامی واحد تهران جنوب)»، پژوهش رضایی و مرادی‌رمقانی (۱۴۰۱) با عنوان «سنجش تأثیر عملکرد مؤلفه‌های امنیت مقصدهای گردشگری بر رضایت‌مندی گردشگران خارجی در راستای توسعه پایدار صنعت گردشگری کلان‌شهر شیراز»، پژوهش دکامینی و احسانی‌فر (۱۴۰۰) با عنوان «سنجش تأثیر بلوغ مدیریت دانش بر کیفیت خدمات در بانک پارسیان»، پژوهش صادقی‌مزیدی، و ابراهیمیان (۱۴۰۰) با عنوان «سنجش تأثیر آموزش بر خط (آنلاین) بر میزان یادگیری درس زبان انگلیسی مدارس شرق استان فارس»، پژوهش صیفوری (۱۳۹۹) با عنوان «سنجش تأثیر معماری اطلاعات کتابخانه‌های دیجیتال بر یادگیری الکترونیکی (نمونه پژوهشی: سامانه منبع‌یاب وزارت بهداشت، درمان، و آموزش پزشکی)»، پژوهش شهیدی، بشیریه، صابری، و عباسی (۱۳۹۹) با عنوان «سنجش تأثیر اعتیاد و وضعیت سرپرستی بر کودک‌آزاری جنسی با تمرکز بر نمونه‌های به‌دست‌آمده از مراکز اورژانس اجتماعی تهران»، پژوهش قنبری، مؤدب، جواد، حسین‌زاده، موسوی، و براری (۱۳۹۸) با عنوان «ارزشیابی تأثیر برگزاری کارگاه ایمنی بیمار بر آگاهی و عملکرد دانشجویان پرستاری: چالش‌ها و راهکارها»، پژوهش بابایی، شکیبازاده، شجاعی‌زاده، یاسری، و محمدزاده (۱۳۹۹) با عنوان «ارزشیابی تأثیر مداخله آموزشی مبتنی بر الگوی اعتقاد بهداشتی بر رفتارهای ارتقاءدهنده سبک زندگی سالم در افراد مستعد بیماری قلبی-عروقی»، پژوهش آقا‌کاردان، و آرانی (۱۳۹۶) با عنوان «سنجش میزان تأثیر بازی گونه‌سازی آزمون‌های الکترونیکی در یادگیری دانش‌آموزان»، پژوهش

بارانی‌پسیان، نامی، و توکلی‌نعمه (۱۳۹۶) با عنوان «بررسی و سنجش تأثیر عوامل اجتماعی-اقتصادی بر سلامت شهری (شهر مقدس قم)»، پژوهش شهرام‌نیا، ابراهیمی‌پور، محموداوغلی، و ملکان (۱۳۹۶) با عنوان «سنجش تأثیر شبکه‌های اجتماعی بر مشارکت سیاسی: مطالعه موردی دانشجویان دانشگاه مازندران»، پژوهش هنری، میری، و معاشری (۱۳۹۵) با عنوان «ارزشیابی تأثیر آموزش بر اساس تئوری رفتار برنامه‌ریزی‌شده بر انجام مراقبت‌های دوران بارداری زنان باردار معتاد»، پژوهش رجبی‌فرجاد، و پوره‌اشمی‌خشک‌ناب (۱۳۹۵)، با عنوان «سنجش تأثیر فناوری اطلاعات بر بهره‌وری کارکنان»، پژوهش نجارنهادوندی و آبخوش (۱۳۹۴)، با عنوان «سنجش تأثیر آموزش شهروندی بر توانمندسازی زنان (مطالعه موردی: منطقه ۶ شهرداری تهران)»، پژوهش نوابخش و ثابتی (۱۳۹۴)، با عنوان «طراحی و ارائه الگو جهت سنجش تأثیر سرمایه فرهنگی بر ابعاد توسعه پایدار شهری با تأکید بر انرژی‌های پاک»، پژوهش عباسی‌شوازی و خواجه‌صالحی (۱۳۹۲)، با عنوان «سنجش تأثیر استقلال، مشارکت اجتماعی، و تحصیلات زنان بر تمایل به فرزندآوری (مطالعه موردی شهر سیرجان)»، و پژوهش محمدی‌زیدی، پاکپورحاجی‌آقا، و محمدی‌زیدی (۱۳۹۱) با عنوان «ارزشیابی تأثیر برنامه آموزشی بر اساس تئوری رفتار برنامه‌ریزی‌شده بر عملکرد ایمن کارگران» اشاره نمود.

روش پژوهش

با مطالعه پیشینه پژوهش و با توجه به هدف و سؤالات پژوهشی و همچنین فقر داده‌ها در زمینه ارزشیابی تأثیر و شاخص‌های فرا ارزشیابی تأثیر، ضرورت داشت تا بر اساس پارادایم عمل‌گرایی^۱ (صالحی و گل‌افشانی، ۲۰۱۰، ۱۸۷) و با بهره‌گیری از رویکرد آمیخته از نوع اکتشافی در دو مرحله پژوهش حاضر انجام شود. در مطالعه اول فهرست‌وارسی فرا ارزشیابی تأثیر (با روش کیفی) تدوین شد، و در مطالعه دوم، با استفاده از ابزار ساخته‌شده (با روش کمی) به ارزیابی مطالعات ارزشیابی تأثیر داخلی پرداخته شده‌است. در مرحله نخست برای پاسخگویی به سؤال اول پژوهش یعنی «گویه‌های مهم در فرا ارزشیابی تأثیر چیست؟» از روش اسنادی و شیوه تحلیل آلتاید^۲ استفاده شد؛ این روش برای تحلیل نظام‌مند اسناد و استخراج مفاهیم موردنظر از آنهاست (واخ و وارد^۳؛ به نقل از حقوقی و صالحی، ۱۳۹۸، ۴۰). مراحل این تحلیل به این ترتیب است که پژوهشگر پس از انتخاب مطالعات بر اساس ملاک‌هایی مانند سودمندی، اعتبار، ارتباط و موارد دیگر و ارزیابی اولیه آن مطالعات، ابتدا اسناد انتخاب‌شده را به ترتیب ارتباط با مسئله پژوهش، مطالعه کرده و در حین مطالعه، قسمت‌های مرتبط با مسئله پژوهش را کدگذاری و همه کدها را در یک فایل جمع‌بندی می‌کند. سپس با مرور مجدد و بررسی کدها، کدهای تکراری را حذف کرده و

1- Pragmatism

2- Altheide's "Process of Qualitative Document Analysis"

3- Wach & Ward

کدهای مشابه را با هم تجمیع و تحت عنوان یک ملاک یا گاهی یک گویه درمی آیند (حقوقی و صالحی، ۱۳۹۸، ۴۱). بنابراین در این بخش از پژوهش حاضر با استفاده از راهبرد آلتاید در گام نخست با توجه به فقر داده‌های داخلی، با جستجوی واژه «ارزشیابی تأثیر^۱» در پایگاه داده‌های بین‌المللی مانند گوگل اسکولار^۲، ریسرچ‌گیت^۳، و وب‌آف‌ساینس^۴ از میان اسناد موجود با ملاک‌های ورودی «سودمندی»، «ارتباط» و «معتبر بودن» به ۳۲ سند کاملاً سودمند، معتبر و مرتبط با موضوع ارزشیابی تأثیر دست‌یافتند که پس از بررسی دقیق متون این ۳۲ سند، تنها ۱۶ سند محتوای ارزشمندی در زمینه اصول و موارد مهم در ارزشیابی تأثیر داشتند. با تحلیل دقیق این اسناد و کدگذاری آنها، و سپس مرور مجدد کدها، کدهای تکراری و همچنین کدهایی با اهمیت و اولویت پایین حذف شدند و سپس با تجمیع کدهای مشابه، ۳۵ گویه فهرست‌وارسی شکل گرفتند و در نهایت اعتبار آنها با استفاده از روش مرور همتا^۵ با نظر یک متخصص بررسی و تأیید شد.

در مرحله دوم پژوهشگران برای پاسخگویی به سؤال «مطالعات ارزشیابی تأثیر داخلی موجود از چه کیفیتی برخوردارند؟» از روش مطالعه ارزشیابی^۶ استفاده نمودند. هر زمان که متخصص ارزشیابی بخواند یا نیاز داشته باشد همزمان با انجام یک پروژه ارزشیابی، پژوهشی نیز در همان راستا انجام دهد، یک مطالعه ارزشیابی انجام شده است (سیف، ۱۳۹۶). در این مرحله پژوهشگران با استفاده از گویه‌های تهیه‌شده به مطالعه و ارزشیابی مطالعات قابل قبول در حیطه ارزشیابی تأثیر پرداختند و از این طریق یک مطالعه فرا ارزشیابی تأثیر (یا ارزشیابی ارزشیابی‌های تأثیر) انجام دادند. در مطالعه نخست پژوهشگران با جستجو در پایگاه داده‌های داخلی نورمگز^۷ و پرتال جامع علوم انسانی^۸ با جستجوی دو عبارت «ارزشیابی تأثیر»، و «سنجش تأثیر» در ابتدا به حدود ۱۴۱ مقاله علمی دست یافتند. سپس بر اساس مرتبط بودن کامل عنوان، حدوداً ۶۲ مقاله باقی ماند که پس از مطالعه چکیده، این مقالات به ۳۰ مقاله کاهش یافتند. در نهایت با مطالعه کل پژوهش‌ها تنها تعداد ۱۸ مقاله در حیطه مطالعات ارزشیابی تأثیر جای‌گرفته و ارزشیابی شدند. در مرحله نهایی و به منظور پاسخگویی به سؤال سوم پژوهش یعنی «آیا وضعیت تحقق گویه‌های تدوین‌شده برای ارزشیابی کیفیت مطالعات ارزشیابی تأثیر در حد تسلط است؟» با استفاده از نرم‌افزار SPSS و با اجرای آزمون دوجمله‌ای^۹ (نسبت) به این سؤال پاسخ کامل داده‌شد. لازم به ذکر است که به منظور

1- Impact Evaluation

2- Google Scholar

3- Researchgate

4- WOS (Web of Science)

5- Peer Review

6- Evaluative research

7- Noormags

8- Ensani.ir

9- Binomial Test

رعایت حقوق مالکیت فکری نویسندگان مقالات، و ناشناس باقی ماندن وضعیت هر یک از مقالات، ترتیب مقاله‌های ارزیابی شده در جداول ۴-۵-۶ فایل پیوست تصادفی چیدمان شده است.

تحلیل یافته‌ها

در این بخش بر اساس سه سؤال اصلی پژوهش یافته‌ها ارائه می‌شوند که در ادامه به این سؤالات پاسخ داده می‌شود.

۱. گویه‌های مهم در فرا ارزشیابی تأثیر چیست؟

تهیه فهرست‌وارسی فرا ارزشیابی تأثیر

نخستین هدف مهم در ارزشیابی، شفاف‌سازی هدف (اهداف) ارزشیابی است. این کار انتظارات خاص سرمایه‌گذاران، و سایر ذی‌نفعان ارزشیابی را مشخص خواهد نمود (پیرزمن و همکاران، ۲۰۱۵)، و علاوه بر پاسخگویی به انتظارات آنها، ارزشیابی ما از شفافیت رویه‌ای برخوردار می‌شود (گرتر و همکاران، ۲۰۱۶؛ کوپستیک، ۲۰۱۴؛ راجرز، ۲۰۱۲). این کار افزون بر امکان تحلیل و تفسیر دقیق داده‌ها (کندر و همکاران، ۲۰۰۹)، و کمک به پوشش جامع موضوعات مرتبط و حائز اهمیت (پیرزمن و همکاران، ۲۰۱۵)، قابلیت پایایی (تکرار) ارزشیابی را افزایش می‌دهد، و به طور کلی اطمینان حاصل می‌شود که نتایج ارزشیابی بی‌طرفانه، قابل اعتماد و معتبر است و به دانش گسترده‌تری کمک می‌کند (گرتر و همکاران، ۲۰۱۶). از این‌رو برای مشخص شدن دقیق اهداف ارزشیابی و شفافیت آن به سراغ تهیه یک فهرست‌وارسی ارزشیابی می‌رویم. یک نکته قابل توجه اینکه یک فهرست‌وارسی یکسان برای هر موقعیتی مناسب نیست (پیرزمن و همکاران، ۲۰۱۵). گفتنی است که در این فهرست‌وارسی، از مقیاس ستاره‌ای سیستم ارزشیابی تأثیر^۱ استفاده شده است. این مقیاس در یک مطالعه ارزشیابی تأثیر، ارزشیابی‌کنندگان را ملزم می‌کند «تأثیر» را از صفر تا چهار ستاره ارزشیابی کنند (ساموئل و دریک، ۲۰۱۵). در اینجا نیز با الهام از این مقیاس ما کیفیت مطالعات ارزشیابی تأثیر را در هر گویه از صفر تا چهار ستاره ارزشیابی می‌کنیم.

جدول ۱. فهرست‌وارسی تدوین شده برای ارزشیابی کیفیت مطالعات ارزشیابی تأثیر

Table 1

Checklist developed for assessing the quality of impact evaluation studies

سطوح ارزش‌گذاری (Evaluation level)				گویه‌ها (Items)	شماره No.
****	***	**	*	در مطالعه منتخب، به هر یک از گویه‌های زیر تا چه اندازه توجه شده است؟ How much attention has been given to each of the following ?in the studies items	
یک مداخله خاص، واضح و مرتبط (تمرکز بر علیت) در مطالعه وجود دارد.					1

	There is a specific, clear, and relevant intervention (focus on causality) in the study	
2	مطالعه به تأثیرات بلندمدت توجه نموده‌است. The study focused on long-term effects.	
3	مطالعه به تأثیرات پایدار توجه نموده‌است. The study focused on lasting effects.	
4	مطالعه به تأثیرات ناخواسته نیز توجه نموده‌است. The study also looked at unintended effects.	
5	در مطالعه اهداف ارزشیابی به طور واضح بیان شده‌است. The evaluation objectives are clearly stated in the study.	
6	ارزشیابی انجام‌شده با اهداف تناسب دارد. The evaluation conducted is consistent with the objectives.	
7	کاربرد ارزشیابی و کاربران آن مشخص شده‌اند. The use of the evaluation and its users are identified.	
8	ارزشیابی به یک سؤال راهبردی و خاص پاسخ می‌دهد. The evaluation answers a specific strategic question.	
9	در انتخاب سؤالات کلیدی ارزشیابی، توافق نظر وجود دارد. There is consensus on the selection of key evaluation questions.	
10	به هزینه‌ها و بودجه‌بندی توجه شده‌است. Costs and budgeting have been considered.	
11	به زمان‌بندی ارزشیابی توجه شده‌است. The evaluation schedule has been considered.	
12	به قابلیت دسترسی به منابع ارزشیابی توجه شده‌است. The availability of evaluation resources has been considered.	
13	به قابلیت دسترسی ذی‌نفعان به برنامه توجه شده‌است. The accessibility of the program by stakeholders has been considered.	
14	به شفافیت روند و جزئیات کار توجه شده‌است. The transparency of the process and the details of the work have been considered.	
15	به شفافیت در مورد محدودیت‌ها توجه شده‌است. The transparency of the limitations has been considered.	
16	به معقول بودن انتظار تأثیر توجه شده‌است. The reasonableness of the expected impact has been considered.	
17	به سودمندی ارزشیابی تأثیر توجه شده‌است. The usefulness of the impact evaluation has been considered.	
18	به امکان‌سنجی ارزشیابی تأثیر توجه شده‌است. The feasibility of the impact evaluation has been considered.	
19	یک فهرست‌واری ارزشیابی واضح در مطالعه وجود دارد. There is a clear evaluation checklist in the study.	
20	یک طرح ارزشیابی در برنامه وجود دارد. There is an evaluation plan in the program.	
21	یک نظریه تغییر واضح در مطالعه ارائه شده‌است. A clear theory of change is presented in the study.	
22	به چندسویه‌سازی روش و منبع گردآوری داده‌ها توجه شده‌است.	

	The triangulation in the method and source of data collection has been considered.	
23	در مورد اجرای صحیح و پیگیری مسئولانه برنامه اطلاعاتی ارائه شده است. Information is provided on the proper implementation and responsible follow-up of the program.	
24	به نوآوری و به‌روز بودن ارزشیابی توجه شده است. The innovation and up-to-date nature of the evaluation have been considered.	
25	ارزشیابی در پر کردن شکاف‌های دانش نقشی ایفا نموده است. Evaluation has played a role in filling knowledge gaps.	
26	یک سیستم نظارتی قوی برای برنامه وجود داشته است. A strong monitoring system for the program is in place.	
27	رویکرد موردنظر (کمی، کیفی، یا ترکیبی) مشخص شده است. The intended approach (quantitative, qualitative, or mixed) is specified.	
28	نوع ارزشیابی بر اساس زمان (گذشته‌نگر، یا آینده‌نگر) مشخص شده است. The type of evaluation is specified based on time (retrospective or prospective).	
29	نوع ارزشیابی بر اساس داخلی، خارجی یا مشارکتی بودن مشخص شده است. The type of evaluation is specified based on whether it is internal, external, or participatory.	
30	تیم ارزشیابی تشکیل شده است و اعضای تیم آموزش دیده‌اند. The evaluation team is formed and team members are trained.	
31	یک روش ارزشیابی منطقی و قابل دفاع وجود دارد. The evaluation team is formed and team members are trained.	
32	به رعایت موارد اخلاقی ارزشیابی اشاره شده است. The ethical considerations of evaluation are addressed.	
33	مطالعه میان نتایج و تأثیرات تمایز قائل شده است. The study distinguishes between results, and impacts.	
34	ارزشیابی از روایی برخوردار است. The evaluation is valid.	
35	ارزشیابی از اعتبار برخوردار است. The evaluation is reliable.	

جدول ۲. منابع استنادی گویه‌های ارزشیابی کیفیت فرا ارزشیابی تأثیر

Table 2

References for quality evaluation items of impact meta-evaluation

شماره No.	گویه‌ها (Items)	منابع (References)
1	یک مداخله خاص، واضح و مرتبط (تمرکز بر علیت) در مطالعه وجود دارد. There is a specific, clear, and relevant intervention (focus on causality) in the study.	(پیرزمن و همکاران، 2015، 6؛ راجرز، 2014، 7؛ بمبرگر، 2012، 3؛ پرین، 2012، 1؛ راجرز، 2012، 2) (Peersman et al., 2015, 6; Rogers, 2014, 7; Bamberger, 2012, 3; Perrin, 2012, 1; Rogers, 2012, 2)
2	مطالعه به تأثیرات بلندمدت توجه نموده است. The study focused on long-term effects.	(پیرزمن و همکاران، 2015، 6؛ پیرزمن، 2014، 1؛ راجرز، 2014، 8؛ بمبرگر، 2012، 3؛ پرین، 2012، 1؛ راجرز، 2012، 2؛ وایت، 2010، 154)

(Peersman et al., 2015, 6; Peersman, 2014, 1; Rogers, 2014, 8; Bamberger, 2012, 3; Perrin, 2012, 1; Rogers, 2012, 2; White, 2010, 154)	
(ساموئل و دریک، 2015، 234؛ راجرز، 2014، 8؛ راجرز، 2012، 2؛ پرین، 2012، 1) (Samuel & Drake, 2015, 234; Rogers, 2014, 8; Rogers, 2012, 2; Perrin, 2012, 1)	3 مطالعه به تأثیرات پایدار توجه نموده‌است. The study focused on lasting effects.
(پیرزمن و همکاران، 2015، 6؛ پیرزمن، 2014، 2؛ راجرز، 2014، 8؛ بمبرگر، 2012، 3؛ پرین، 2012، 1؛ راجرز، 2012، 2؛ وایت، 2010، 154) (Peersman et al., 2015, 6; Peersman, 2014, 2; Rogers, 2014, 8; Bamberger, 2012, 3; Perrin, 2012, 1; White, 2010, 154)	4 مطالعه به تأثیرات ناخواسته نیز توجه نموده‌است. The study also looked at unintended effects.
(گرتلر و همکاران، 2016، 236؛ راجرز، 2014، 7؛ راجرز، 2012، 17؛ کندکر و همکاران، 2009، 28) (Gertler et al., 2016, 236; Rogers, 2014, 7; Rogers, 2012, 17; Khandker et al., 2009, 28)	5 در مطالعه اهداف ارزشیابی به طور واضح بیان شده‌است. The evaluation objectives are clearly stated in the study.
(پیرزمن و همکاران، 2015، 9؛ راجرز، 2014، 7؛ کوپستیک، 2014، 416؛ پرین، 2012، 3؛ کندکر و همکاران، 2009، 28) (Peersman et al., 2015, 9; Rogers, 2014, 7; Copestake, 2014, 416; Perrin, 2012, 3; Khandker et al., 2009, 28)	6 ارزشیابی انجام شده با اهداف تناسب دارد. The evaluation conducted is consistent with the objectives.
(پیرزمن و همکاران، 2015، 10؛ راجرز، 2014، 2؛ پرین، 2012، 10) (Peersman et al., 2015, 10; Rogers, 2014, 2; Perrin, 2012, 10)	7 کاربرد ارزشیابی و کاربران آن مشخص شده‌اند. The use of the evaluation and its users are identified.
(گیوگتی و همکاران، 2016، 5؛ گرتلر و همکاران، 2016، 204؛ پیرزمن، 2014، 1؛ راجرز، 2012، 15) (Gugerty et al., 2016, 5; Gertler et al., 2016, 204; Peersman, 2014, 1; Rogers, 2012, 15)	8 ارزشیابی به یک سؤال راهبردی و خاص پاسخ می‌دهد. The evaluation answers a specific strategic question.
(گیوگتی و همکاران، 2016، 7؛ پیرزمن و همکاران، 2015، 7) (Gugerty et al., 2016, 7; Peersman et al., 2015, 7)	9 در انتخاب سؤالات کلیدی ارزشیابی، توافق نظر وجود دارد. There is consensus on the selection of key evaluation questions.
(گیوگتی و همکاران، 2016، 3 و 7؛ گرتلر و همکاران، 2016، 26؛ پیرزمن و همکاران، 2015، 8؛ راجرز، 2014، 4؛ راجرز، 2012، 2؛ کوپستیک، 2014، 416) (Gugerty et al., 2016, 3 & 7; Gertler et al., 2016, 26; Peersman et al., 2015, 8; Rogers, 2014, 4; Rogers, 2012, 2; Copestake, 2014, 416)	10 به هزینه‌ها و بودجه‌بندی توجه شده‌است. Costs and budgeting have been considered.
(گیوگتی و همکاران، 2016، 6؛ گرتلر و همکاران، 2016، 214؛ پیرزمن و همکاران، 2015، 10؛ راجرز، 2014، 3؛ راجرز، 2012، 14 و 15؛ پرین، 2012، 5 و 9؛ کندکر و همکاران، 2009، 28) (Gugerty et al., 2016, 6; Peersman et al., 2015, 10; Rogers, 2014, 3; Rogers, 2012, 14 & 15; Perrin, 2012, 5 & 9; Khandker et al., 2009, 28)	11 به زمان‌بندی ارزشیابی توجه شده‌است. The evaluation schedule has been considered.

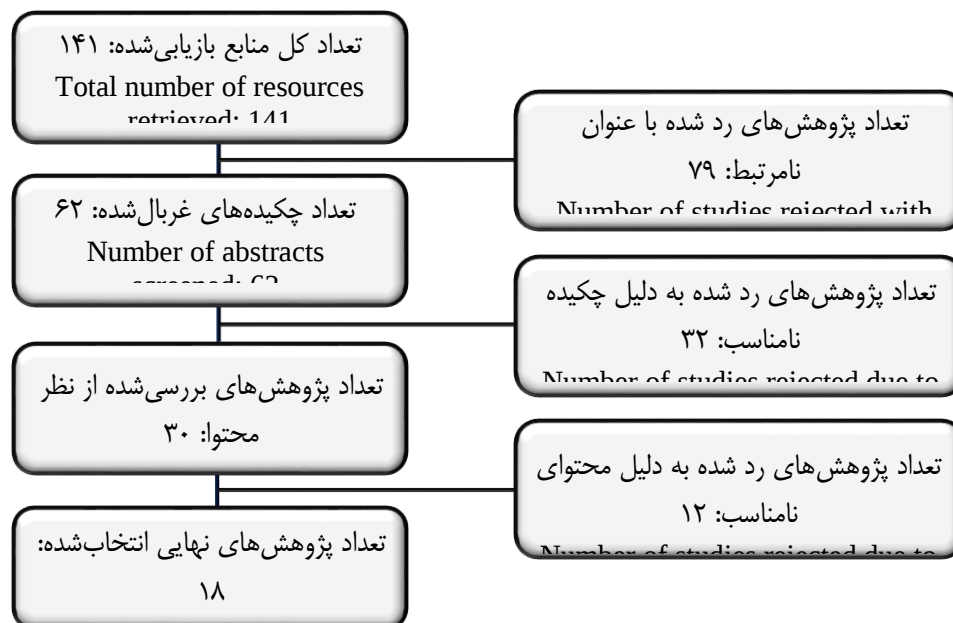
(Gugerty et al., 2016, 6; Gertler et al., 2016, 214; Peersman et al., 2015, 10; Rogers, 2014, 3; Rogers, 2012, 14 & 15; Perrin, 2012, 5 and 9; Khandker et al., 2009, 28)		
12	به قابلیت دسترسی به منابع ارزشیابی توجه شده است. The availability of evaluation resources has been considered.	(گرتلر و همکاران، 2016، 247؛ پیرزمن و همکاران، 2015، 10؛ راجرز، 2014، 4؛ راجرز، 2012، 15) (Gertler et al., 2016, 247; Peersman et al., 2015, 10; Rogers, 2014, 4; Rogers, 2012, 15)
13	به قابلیت دسترسی ذی‌نفعان به برنامه توجه شده است. The accessibility of the program by stakeholders has been considered.	(پیرزمن و همکاران، 2015، 8؛ راجرز، 2014، 13؛ راجرز، 2012، 15؛ کندکر و همکاران، 2009، 28) (Peersman et al., 2015, 8; Rogers, 2014, 13; Rogers, 2012, 15; Khandker et al., 2009, 28)
14	به شفافیت روند و جزئیات کار توجه شده است. The transparency of the process and the details of the work have been considered.	(گرتلر و همکاران، 2016، 189؛ پیرزمن، 2014، 5؛ کوپستیک، 2014، 418) (Gertler et al., 2016, 189; Peersman, 2014, 5; Copestake, 2014, 418)
15	به شفافیت در مورد محدودیت‌ها توجه شده است. The transparency of the limitations has been considered.	(پیرزمن و همکاران، 2015، 10؛ راجرز، 2012، 15) (Peersman et al., 2015, 10; Rogers, 2012, 15)
16	به معقول بودن انتظار تأثیر توجه شده است. The reasonableness of the expected impact has been considered.	(پیرزمن و همکاران، 2015، 12؛ پرین، 2012، 9؛ راجرز، 2012، 12) (Peersman et al., 2015, 12; Perrin, 2012, 9; Rogers, 2012, 12)
17	به سودمندی ارزشیابی تأثیر توجه شده است. The usefulness of the impact evaluation has been considered.	(رید و همکاران، 2021، 4؛ پیرزمن و همکاران، 2015، 12؛ راجرز، 2012، 15؛ پرین، 2012، 5) (Reed et al., 2021, 4; Peersman et al., 2015, 12; Rogers, 2012, 15; Perrin, 2012, 5)
18	به امکان‌سنجی ارزشیابی تأثیر توجه شده است. The feasibility of the impact evaluation has been considered.	(گرتلر و همکاران، 2016، 204؛ پیرزمن و همکاران، 2015، 12؛ پیرزمن، 2014، 5؛ راجرز، 2014، 14؛ راجرز، 2012، 6) (Gertler et al., 2016, 204; Peersman et al., 2015, 12; Peersman, 2014, 5; Rogers, 2014, 14; Rogers, 2012, 6)
19	یک فهرست‌واری ارزشیابی واضح در مطالعه وجود دارد. There is a clear evaluation checklist in the study.	(گرتلر و همکاران، 2016، 243؛ پیرزمن و همکاران، 2015، 9؛ پرین، 2012، 10) (Gertler et al., 2016, 243; Peersman et al., 2015, 9; Perrin, 2012, 10)
20	یک طرح ارزشیابی در برنامه وجود دارد. There is an evaluation plan in the program.	(رید و همکاران، 2021، 4؛ گیوگنی و همکاران، 2016، 5؛ گرتلر و همکاران، 2016، 27، 210 و 236؛ پیرزمن، 2014، 5؛ راجرز، 2014، 5؛ راجرز، 2012، 15؛ کندکر و همکاران، 2009، 19) (Gugerty et al., 2016, 5; (Reed et al., 2021, 4; Gertler et al., 2016, 27, 210 & 236; Peersman, 2014, 5; Rogers, 2014, 5; Rogers, 2012, 15; Khandker et al., 2009, 19)

21	یک نظریه تغییر واضح در مطالعه ارائه شده است. A clear theory of change is presented in the study.	(میشل و اشنایدر، 2025، 1؛ گیوگتی و همکاران، 2016، 5؛ گرتلر و همکاران، 2016، 204؛ پیرزمن، 2014، 1 و 2؛ راجرز، 2014، 4؛ راجرز، 2012، 6؛ پرین، 2012، 8) (Michel & Schneider, 2025, 1; Gugerty et al., 2016, 5; Gertler et al., 2016, 204; Peersman, 2014, 1 & 2; Rogers, 2014, 4; Rogers, 2012, 6; Perrin, 2012, 8)
22	به چندسویه سازی روش و منبع گردآوری داده ها توجه شده است. The triangulation in the method and source of data collection has been considered.	(میشل و اشنایدر، 2025، 2؛ پیرزمن و همکاران، 2015، 9؛ پیرزمن، 2014، 1؛ کوپستیک، 2014، 418؛ بمبرگر، 2012، 7؛ راجرز، 2012، 5) (Michel & Schneider, 2025, 2; Peersman et al., 2015, 9; Peersman, 2014, 1; Copestake, 2014, 418; Bamberger, 2012, 7; Rogers, 2012, 5)
23	در مورد اجرای صحیح و پیگیری مسئولانه برنامه اطلاعاتی ارائه شده است. Information is provided on the proper implementation and responsible follow-up of the program.	(گیوگتی و همکاران، 2016، 2؛ گرتلر و همکاران، 2016، 7؛ پیرزمن، 2014، 2 و 3؛ راجرز، 2014، 4؛ راجرز، 2012، 7؛ پرین، 2012، 7) (Gugerty et al., 2016, 2; Gertler et al., 2016, 7; Peersman, 2014, 2 & 3; Rogers, 2014, 4; Rogers, 2012, 7; Perrin, 2012, 7)
24	به نوآوری و به روز بودن ارزشیابی توجه شده است. The innovation and up-to-date nature of the evaluation have been considered.	(گیوگتی و همکاران، 2016، 5؛ پیرزمن، 2014، 6) (Gugerty et al., 2016, 5; Peersman, 2014, 6)
25	ارزشیابی در پر کردن شکاف های دانش نقشی ایفا نموده است. Evaluation has played a role in filling knowledge gaps.	(بیلیس و همکاران، 2016، 60؛ گیوگتی و همکاران، 2016، 7؛ پیرزمن و همکاران، 2015، 10؛ پیرزمن، 2014، 1 و 2؛ راجرز، 2014، 2؛ راجرز، 2012، 12 و 17؛ پرین، 2012، 10؛ گاربارینو و هالند، 2009، 5) (Baylis et al., 2016, 60; Gugerty et al., 2016, 7; Peersman et al., 2015, 10; Peersman, 2014, 1 and 2; Rogers, 2014, 2; Rogers, 2012, 12 and 17; Perrin, 2012, 10; Garbarino and Holland, 2009, 5)
26	یک سیستم نظارتی قوی برای برنامه وجود داشته است. A strong monitoring system for the program is in place.	(میشل و اشنایدر، 2025، 4؛ گیوگتی و همکاران، 2016، 3 و 7؛ گرتلر و همکاران، 2016، 7؛ پیرزمن، 2014، 2؛ راجرز، 2014، 3؛ پرین، 2012، 10) (Michel & Schneider, 2025, 4; Gugerty et al., 2016, 3 & 7; Gertler et al., 2016, 7; Peersman, 2014, 2; Rogers, 2014, 3; Perrin, 2012, 10)
27	رویکرد مورد نظر (کمی، کیفی، یا ترکیبی) مشخص شده است. The intended approach (quantitative, qualitative, or mixed) is specified.	(پوزارنی و برینگتون، 2016، 1؛ کوپستیک، 2014، 414؛ پیرزمن، 2014، 4؛ بمبرگر، 2012، 4؛ کندکر و همکاران، 2009، 19؛ گاربارینو و هالند، 2009، 11) (Pozarny & Barrington, 2016, 1; Copestake, 2014, 414; Peersman, 2014, 4; Bamberger, 2012, 4; Khandker et al., 2009, 19; Garbarino & Holland, 2009, 11)
28	نوع ارزشیابی بر اساس زمان (گذشته نگر، یا آینده نگر) مشخص شده است. The type of evaluation is specified based on time (retrospective or prospective).	(گرتلر و همکاران، 2016، 9؛ کندکر و همکاران، 2009، 20) (Gertler et al., 2016, 9; Khandker et al., 2009, 20)

29	نوع ارزشیابی بر اساس داخلی، خارجی یا مشارکتی بودن مشخص شده است. The type of evaluation is specified based on whether it is internal, external, or participatory.	(گرتلر و همکاران، 2016، 27؛ راجرز، 2014، 3 و 10؛ راجرز، 2012، 5) (Gertler et al., 2016, 27; Rogers, 2014, 3 & 10; Rogers, 2012, 5)
30	تیم ارزشیابی تشکیل شده است و اعضای تیم آموزش دیده اند. The evaluation team is formed and team members are trained.	(گرتلر و همکاران، 2016، 27؛ راجرز، 2014، 3؛ کندکر و همکاران، 2009، 28) (Gertler et al., 2016, 27; Rogers, 2014, 3; Khandker et al., 2009, 28)
31	یک روش ارزشیابی منطقی و قابل دفاع وجود دارد. The evaluation team is formed and team members are trained.	(گرتلر و همکاران، 2016، 8؛ گاربارینو و هالند، 2009، 4) (Gertler et al., 2016, 8; Garbarino & Holland, 2009, 4)
32	به رعایت موارد اخلاقی ارزشیابی اشاره شده است. The ethical considerations of evaluation are addressed.	(گرتلر و همکاران، 2016، 232؛ پیرزمن و همکاران، 2015، 10؛ راجرز، 2012، 15) (Gertler et al., 2016, 232; Peersman et al., 2015, 10; Rogers, 2012, 15)
33	مطالعه میان نتایج و تأثیرات تمایز قائل شده است. The study distinguishes between results, and impacts.	(پیرزمن، 2014، 5؛ راجرز، 2014، 3؛ راجرز، 2012، 2؛ پرین، 2012، 1) (Peersman, 2014, 5; Rogers, 2014, 3; Rogers, 2012, 2; Perrin, 2012, 1)
34	ارزشیابی از روایی برخوردار است. The evaluation is valid.	(گرتلر و همکاران، 2016، 254؛ پیرزمن، 2014، 1 و 6؛ راجرز، 2014، 14؛ بمرگر، 2012، 4) (Gertler et al., 2016, 254; Peersman, 2014, 1 & 6; Rogers, 2014, 14; Bamberger, 2012, 4)
35	ارزشیابی از اعتبار برخوردار است. The evaluation is reliable.	(گرتلر و همکاران، 2016، 242؛ پیرزمن، 2014، 6؛ راجرز، 2014، 14) (Gertler et al., 2016, 242; Peersman, 2014, 6; Rogers, 2014, 14)

۱. مطالعات ارزشیابی تأثیر داخلی موجود از چه کیفیتی برخوردارند؟

پژوهشگران پس از جستجوی پژوهش‌هایی با کلیدواژگان «ارزشیابی تأثیر»، و «سنجش تأثیر» در پایگاه داده‌های داخلی نورمگز و پرتال جامع علوم انسانی، و بازایی حدوداً ۱۴۱ پژوهش، با مطالعه اجمالی و بررسی اولیه پژوهش‌ها (از نظر عناوین کاملاً مرتبط)، ۶۲ پژوهش را انتخاب نموده و پس از مطالعه اجمالی چکیده و روش پژوهش، ۳۰ پژوهش باقی ماندند که با مطالعه کامل متن این پژوهش‌ها تنها ۱۸ پژوهش قابلیت ارزشیابی را داشت و این ۱۸ پژوهش به شرح جدول شماره ۳ (در فایل پیوست) وارد فرآیند ارزشیابی شدند.



شکل ۱. خلاصه فرآیند بازیابی و انتخاب منابع

Figure 1

Summary of the resource retrieval and selection process

نتایج ارزشیابی این پژوهش‌ها، حاکی از آن است که بیشتر مطالعات انجام‌شده در حوزه ارزشیابی تأثیر، تنها عنوان ارزشیابی یا سنجش تأثیر را به دوش می‌کشند، و تنها یک مطالعه پیمایشی، همبستگی و یا در بهترین حالت یک مطالعه نیمه‌آزمایشی ساده هستند. استدلال پژوهشگران در دفاع از این واقعیت این است که این مطالعات تا حد زیادی از اصول و قواعد ارزشیابی تأثیر که در اسناد بین‌المللی مرتبط با این حوزه به طور مکرر تأکید شده‌است، تبعیت نکرده و دچار انحراف چشم‌گیری شده‌اند و این به معنی آن است که پژوهشگران تنها با تعاریف ذهنی خود از سنجش و یا ارزشیابی تأثیر، در این حوزه قدم گذارده‌اند. بنابراین مطالعات بررسی شده در حقیقت هیچکدام در زمره پژوهش‌های ارزشیابی و یا سنجش تأثیر واقعی محسوب نمی‌شوند.

شیوه ارزشیابی هر پژوهش نیز به این صورت بوده‌است که یک ستاره برای (عدم رعایت گویه مورد نظر)، ۲ ستاره برای (رعایت ۱ تا ۳۳ درصدی گویه و عدم اشاره به آن)، ۳ ستاره برای (رعایت تقریبی ۳۳ تا ۶۶ درصدی گویه مورد نظر و اشاره ضمنی به آن)، و ۴ ستاره برای (رعایت تقریبی ۶۶ تا ۱۰۰ درصدی

گویه موردنظر و اشاره واضح به آن)، اختصاص داده شد. نتایج ارزشیابی این مقالات در حوزه ارزشیابی تأثیر در قالب (جداول ۴ و ۵ و ۶) در فایل پیوست ارائه خواهد شد.

پس از ارزشیابی کامل مقالات، پژوهشگران به این نتیجه دست یافتند که در میان پژوهش‌های ارزشیابی‌شده از بسیاری از گویه‌های مهم مانند «به هزینه‌ها و بودجه‌بندی توجه شده‌است»، «یک فهرست‌وارسی ارزشیابی واضح در مطالعه وجود دارد»، «یک طرح ارزشیابی در برنامه وجود دارد»، «یک نظریه تغییر واضح در مطالعه ارائه شده‌است»، «یک سیستم نظارتی قوی برای برنامه وجود داشته‌است»، «نوع ارزشیابی بر اساس داخلی، خارجی یا مشارکتی بودن ارزشیابی مشخص شده‌است»، «تیم ارزشیابی تشکیل شده‌است و اعضای تیم آموزش دیده‌اند»، «یک روش ارزشیابی منطقی و قابل دفاع وجود دارد»، «به رعایت موارد اخلاقی ارزشیابی اشاره شده‌است»، «مطالعه میان نتایج، پیامدها و تأثیرات تمایز قائل شده‌است»، غفلت شده و به آن‌ها توجه محدودی شده‌است.

در میان پژوهش‌ها به ترتیب پژوهش «سنجش تأثیر معماری اطلاعات کتابخانه‌های دیجیتال بر یادگیری الکترونیکی (نمونه پژوهشی: سامانه منبع‌یاب وزارت بهداشت، درمان و آموزش پزشکی)» با رعایت ۴۶/۶۶ درصد گویه‌ها، پژوهش «سنجش تأثیر شبکه‌های اجتماعی بر مشارکت سیاسی: مطالعه موردی دانشجویان دانشگاه مازندران»، پژوهش «سنجش تأثیر آموزش شهروندی بر توانمندسازی زنان (مطالعه موردی: منطقه ۶ شهرداری تهران)» و پژوهش «ارزشیابی تأثیر طرح سربازان سازندگی در بخش کشاورزی» با رعایت ۳۸/۰۹ درصد گویه‌ها، پژوهش «ارزشیابی تأثیر برنامه آموزشی بر اساس تئوری رفتار برنامه‌ریزی‌شده بر عملکرد ایمن کارگران» و پژوهش «ارزشیابی تأثیر برگزاری کارگاه ایمنی بیمار بر آگاهی و عملکرد دانشجویان پرستاری: چالش‌ها و راهکارها» با رعایت ۳۷/۱۴ درصد گویه‌ها، پژوهش «سنجش تأثیر بلوغ مدیریت دانش بر کیفیت خدمات در بانک پارسیان»، و پژوهش «ارزشیابی تأثیر مداخله آموزشی مبتنی بر الگوی اعتقاد بهداشتی بر رفتارهای ارتقاءدهنده سبک زندگی سالم در افراد مستعد بیماری قلبی-عروقی» با رعایت ۳۶/۱۹ درصد گویه‌ها، پژوهش «سنجش تأثیر استقلال، مشارکت اجتماعی، و تحصیلات زنان بر تمایل به فرزندآوری (مطالعه موردی شهر سیرجان)» با رعایت ۳۱/۴۲ درصد گویه‌ها، و پژوهش «سنجش تأثیر سواد رسانه‌ای در استفاده از فضای مجازی بر اخلاق اجتماعی کاربران (مورد مطالعه: دانشجویان کارشناسی ارشد دانشگاه آزاد اسلامی واحد تهران جنوب)» با رعایت ۳۰/۴۷ درصد گویه‌ها، به ترتیب ۱۰ پژوهش قابل قبول‌تر در میان ۱۸ پژوهش بررسی شده بودند. درصد رعایت گویه‌ها در مابقی پژوهش‌ها زیر ۳۰ درصد بود که این سطح رعایت و توجه به گویه‌ها بسیار پایین و ضعیف تلقی می‌گردد.

۱. آیا وضعیت تحقق گویه‌های تدوین‌شده برای ارزشیابی کیفیت مطالعات ارزشیابی

تأثیر در حد تسلط است؟

برای پاسخ به این سؤال، پژوهشگران مناسب دیدند از آزمون دوجمله‌ای استفاده کنند. بر اساس داده‌های ذکر شده در فایل جداول پیوست، در این ارزشیابی بر اساس سطوح ارزشیابی (۴ سطح) و با توجه به ملاک تسلط^۱ یعنی ۸۵ یا ۹۰ درصد (سیف، ۱۳۹۴، ص. ۵۲۱) و میانگین این ملاک یعنی ۸۷/۵ درصد (نه سهل‌گیرانه و نه سخت‌گیرانه)، نقطه برش برابر با ۳/۵ (۸۷/۵ درصد از ۴ سطح) و نسبت مفروض نیز برابر با ۰/۸۷۵ در نظر گرفته شد. بنابراین **فرض صفر (H_0)** به این صورت است که در گویه‌های ارزشیابی تأثیر، پژوهش‌ها به حد تسلط (۸۷/۵ درصد یا نسبت ۰/۸۷۵ هزارم) نرسیده‌اند؛ و **فرض مقابل یا خلاف (H_1)** به این صورت است که در گویه‌های ارزشیابی تأثیر پژوهش‌ها به حد تسلط (۸۷/۵ درصد یا نسبت ۰/۸۷۵ هزارم) رسیده‌اند.

جدول ۶ نتایج آزمون دوجمله‌ای را نشان می‌دهد. این آزمون علاوه بر مشخص نمودن تفاوت/عدم تفاوت بین نسبت مشاهده‌شده و نسبت مفروض (حبیب‌پورگتایی و صفری شالی، ۱۴۰۱)، می‌تواند مشخص کند که در صورت وجود تفاوت ($p\text{-value} < 0.05$)، رعایت گویه‌ها در هر یک از پژوهش‌ها در چه سطحی (بالا تر و یا پایین تر از حد تسلط) قرار دارند.

جدول ۶ نتایج آزمون دوجمله‌ای

Table 6

Binomial test results

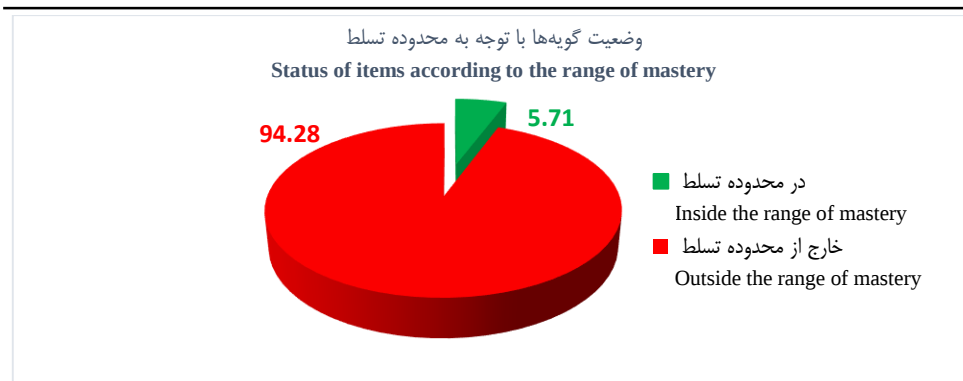
میزان تحقق کلی گویه‌ها (Overall fulfillment rate of the items)	معناداری آماری (Statistical significance)	Exact Sig (1-tailed)	نسبت مشاهده‌شده (Observed Prop)	مقدار پژوهش‌ها (N)	نقطه برش (Cut point)	گروه‌ها (Category)	گویه‌ها (Items)
%38.8	-	0.180	0.778	14	≤ 3.5	گروه ۱ (Group 1)	1
			0.222	4	> 3.5	گروه ۲ (Group 2)	
%55.5	-	0.605	0.889	16	≤ 3.5	گروه ۱ (Group 1)	2
			0.111	2	> 3.5	گروه ۲ (Group 2)	
%57.4	-	0.605	0.889	16	≤ 3.5	گروه ۱ (Group 1)	3
			0.111	2	> 3.5	گروه ۲ (Group 2)	
%24.07	-	0.090	1	18	≤ 3.5	گروه ۱ (Group 1)	4
			0	0	> 3.5	گروه ۲ (Group 2)	
%64.08	-	0.065	0.722	13	≤ 3.5	گروه ۱ (Group 1)	5
			0.278	5	> 3.5	گروه ۲ (Group 2)	
%79.06	*	0.000	0.5	9	≤ 3.5	گروه ۱ (Group 1)	6
			0.5	9	> 3.5	گروه ۲ (Group 2)	
%33.3	-	0.090	1	18	≤ 3.5	گروه ۱ (Group 1)	7

1- Mastery criteria

			0	0	>3.5	گروه ۲ (Group 2)	
			0.889	16	<=3.5	گروه ۱ (Group 1)	8
%53.7	-	0.605	0.111	2	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	9
%33.3	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	10
%1.85	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	11
%16.6	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	12
%27.7	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	13
%25.9	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	14
%48.14	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			0.944	17	<=3.5	گروه ۱ (Group 1)	15
%27.7	-	0.323	0.056	1	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	16
%38.8	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			0.833	15	<=3.5	گروه ۱ (Group 1)	17
%48.14	-	0.395	0.167	3	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	18
%53.7	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	19
%5.5	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	20
%1.85	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	21
%3.7	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	22
%7.4	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	23
%7.4	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			1	18	<=3.5	گروه ۱ (Group 1)	24
%53.7	-	0.090	0	0	>3.5	گروه ۲ (Group 2)	
			0.722	13	<=3.5	گروه ۱ (Group 1)	25
%51.8	-	0.065	0.278	5	>3.5	گروه ۲ (Group 2)	

26	گروه ۱ (Group 1)	≤ 3.5	18	1	0.090	-	1.85%
			0	0			
27	گروه ۱ (Group 1)	≤ 3.5	9	0.5	0.000	*	55.5%
			9	0.5			
28	گروه ۱ (Group 1)	≤ 3.5	16	0.889	0.605	-	44.4%
			2	0.111			
29	گروه ۱ (Group 1)	≤ 3.5	18	1	0.090	-	3.7%
			0	0			
30	گروه ۱ (Group 1)	≤ 3.5	18	1	0.090	-	0
			0	0			
31	گروه ۱ (Group 1)	≤ 3.5	18	1	0.090	-	3.7%
			0	0			
32	گروه ۱ (Group 1)	≤ 3.5	17	0.944	0.323	-	5.5%
			1	0.056			
33	گروه ۱ (Group 1)	≤ 3.5	18	1	0.090	-	1.85%
			0	0			
34	گروه ۱ (Group 1)	≤ 3.5	18	1	0.090	-	33.3%
			0	0			
35	گروه ۱ (Group 1)	≤ 3.5	18	1	0.090	-	35.1%
			0	0			

پس از انجام آزمون دوجمله‌ای و با توجه به نتایج جدول ۶ پژوهشگران به این نتیجه دست یافتند که دو گویه از مجموع ۳۵ گویه (معادل ۵/۷۱ درصد از کل گویه‌ها)، نسبت مفروض با نسبت مشاهده‌شده متفاوت است. مقایسه دو سطح نسبت مشاهده‌شده و مفروض نشان‌دهنده این است که هر دو گویه شماره ۶ «ارزشیابی انجام شده با اهداف تناسب دارد»، و شماره ۲۷ «رویکرد موردنظر (کمی، کیفی، یا ترکیبی) مشخص شده‌است» در سطح ۰/۰۵ معنادار بوده، و این تفاوت معنادار در محدوده تسلط بوده‌است و بنابراین در این دو مورد، فرض صفر (H_0) رد و فرض خلاف (H_1) پذیرفته می‌شود. این به معنی آن است که از میان ۳۵ گویه تنها در دو گویه شماره ۶ «ارزشیابی انجام شده با اهداف تناسب دارد»، و شماره ۲۷ «رویکرد موردنظر (کمی، کیفی، یا ترکیبی) مشخص شده‌است» پژوهش‌ها به حد تسلط (یعنی نمره ۱۷/۵ از ۲۰) رسیده‌اند، اما وضعیت مابقی گویه‌ها اینچنین نیست.



شکل ۲. نمودار دایره‌ای وضعیت گویه‌ها با توجه به محدوده تسلط

Figure 2

Pie chart of the status of items according to the mastery range

همچنین با توجه به سطح معناداری (بالتر از ۰/۰۵)، ۳۳ گویه (معادل ۹۴/۲۸ درصد از کل گویه‌ها) تفاوت معناداری بین نسبت دو سطح مشاهده‌شده و مفروض وجود ندارد. در ۲۳ گویه (معادل ۶۵/۷۱ درصد از کل گویه‌ها و ۶۹/۶۹ درصد از گویه‌های بدون تفاوت معنادار) شامل «مطالعه به تأثیرات ناخواسته نیز توجه نموده‌است»، «کاربرد ارزشیابی و کاربران آن مشخص شده‌اند»، «در انتخاب سؤالات کلیدی ارزشیابی، توافق نظر وجود دارد»، «به هزینه‌ها و بودجه‌بندی توجه شده‌است»، «به زمان‌بندی ارزشیابی توجه شده‌است»، «به قابلیت دسترسی به منابع ارزشیابی توجه شده‌است»، «به قابلیت دسترسی ذی‌نفعان به برنامه توجه شده‌است»، «به شفافیت روند و جزئیات کار توجه شده‌است»، «به معقول بودن انتظار تأثیر توجه شده‌است»، «به امکان‌سنجی ارزشیابی تأثیر توجه شده‌است»، «یک فهرست‌واری ارزشیابی واضح در مطالعه وجود دارد»، «یک طرح ارزشیابی در برنامه وجود دارد»، «یک نظریه تغییر واضح در مطالعه ارائه شده‌است»، «به چندسویه‌سازی داده‌ها توجه شده‌است»، «در مورد اجرای صحیح و پیگیری مسئولانه برنامه اطلاعاتی ارائه شده‌است»، «به نوآوری و به‌روز بودن ارزشیابی توجه شده‌است»، «یک سیستم نظارتی قوی برای برنامه وجود داشته‌است»، «نوع ارزشیابی بر اساس داخلی، خارجی یا مشارکتی بودن ارزشیابی مشخص شده‌است»، «تیم ارزشیابی تشکیل شده‌است و اعضای تیم آموزش دیده‌اند»، «یک روش ارزشیابی قوی و قابل دفاع وجود دارد»، «مطالعه میان نتایج، پیامدها و تأثیرات تمایز قائل شده‌است»، «ارزشیابی از روایی برخوردار است»، «ارزشیابی از اعتبار برخوردار است»، هیچ پژوهشی به حد تسلط (یعنی نمره ۱۷/۵ از ۲۰) دست نیافته‌است.

۱۰ گویه باقی‌مانده از ۳۳ گویه بدون تفاوت معنادار (معادل ۳۰/۳۰ درصد از گویه‌های فاقد تفاوت معنادار و ۲۸/۵۷ درصد از کل گویه‌ها) هرکدام به مقدار درصد متفاوتی به حد تسلط رسیده‌اند که این میزان کمتر از ۵۰ درصد از پژوهش‌ها است. این گویه‌ها شامل گویه شماره ۱ «یک مداخله خاص، واضح و مرتبط (تمرکز بر علیت) در مطالعه وجود دارد» در ۷۷/۷۷ درصد از پژوهش‌ها، گویه شماره ۲ «مطالعه به تأثیرات

بلندمدت توجه نموده‌است» در ۸۸/۸۸ درصد از پژوهش‌ها، گویه شماره ۳ «مطالعه به تأثیرات پایدار در طول زمان توجه نموده‌است» در ۸۸/۸۸ درصد از پژوهش‌ها، گویه شماره ۵ «در مطالعه اهداف ارزشیابی به طور واضح بیان شده‌است» در ۷۲/۲۲ درصد از پژوهش‌ها، گویه شماره ۸ «ارزشیابی به یک سؤال راهبردی و خاص پاسخ می‌دهد» در ۸۸/۸۸ درصد از پژوهش‌ها، گویه شماره ۱۵ «به شفافیت در مورد محدودیت‌ها توجه شده‌است» در ۹۴/۴۴ درصد از پژوهش‌ها، گویه شماره ۱۷ «به سودمندی ارزشیابی تأثیر توجه شده‌است» در ۸۳/۳۳ درصد از پژوهش‌ها، گویه شماره ۲۵ «ارزشیابی در پر کردن شکاف‌های دانش نقشی ایفا نموده‌است» در ۷۲/۲۲ درصد از پژوهش‌ها، گویه شماره ۲۸ «نوع ارزشیابی بر اساس زمان گذشته‌نگر، یا آینده‌نگر) مشخص شده‌است» در ۸۸/۸۸ درصد از پژوهش‌ها، و گویه شماره ۳۲ «به رعایت موارد اخلاقی ارزشیابی اشاره شده‌است» به ۹۴/۴۴ درصد از پژوهش‌ها به حد تسلط دست نیافته‌اند. در پایان اگر بخواهیم میزان تحقق کلی گویه‌ها را در سه سطح (حد تسلط، نزدیک به تسلط، و بدون تسلط) نشان دهیم به این صورت است که ۲۳ گویه (معادل ۶۵/۷۱ درصد از کل گویه‌ها) در محدوده بدون تسلط، ۱۰ گویه (معادل ۲۸/۵۷ درصد از کل گویه‌ها) در محدوده نزدیک به تسلط، و تنها دو گویه (معادل ۵/۷۱ درصد از کل گویه‌ها)، در محدوده تسلط قرار دارد.

نتیجه گیری

«ارزشیابی تأثیر» به معنای ارزشیابی «اثرات مثبت و منفی، اولیه و ثانویه بلندمدتی است که توسط یک مداخله، به طور مستقیم یا غیرمستقیم، عمدی (خواسته) یا ناخواسته ایجاد می‌شود» (پیرزمن و همکاران، ۲۰۱۵، ۶؛ پیرزمن، ۲۰۱۴، ۱؛ بمبرگر، ۲۰۱۲، ۳؛ پرین، ۲۰۱۲، ۱؛ راجرز، ۲۰۱۲، ۲؛ و وایت، ۲۰۱۰، ۱۵۴)، لذا در صورت وجود ارزشیابی به معنای واقعی کلمه، بسیاری از مطالعات مدعی در این حوزه را می‌توان به عنوان ارزشیابی تأثیر در نظر گرفت، اما این نکته قابل توجه است که تأثیرات دیرتر از نتایج میانی و در نتیجه آنها رخ می‌دهد. به بیان دیگر میان نتایج یک رابطه، مداخله و یا به طور کلی مطالعه، و تأثیرات آنها تفاوت‌ها و تمایزاتی وجود دارد. چرا که همان‌گونه که بحث شد، «تأثیر» معمولاً به سطح نهایی زنجیره علی که همان به اثرات بلندمدت است، اشاره دارد (پیرزمن، گایت و پازن، ۲۰۱۵، ۶؛ پیرزمن، ۲۰۱۴، ۵؛ بمبرگر، ۲۰۱۲، ۳؛ پرین، ۲۰۱۲، ۱؛ راجرز، ۲۰۱۲، ۲؛ و وایت، ۲۰۱۰، ۱۵۴). پژوهش حاضر ضمن در نظر گرفتن این تعریف و با هدف بازنمایی اصول و مبانی اساسی ارزشیابی تأثیر، به شناسایی نشانگرهای سنجش کیفیت مطالعات ارزشیابی تأثیر و تدوین فهرست‌واری ارزشیابی مطالعات ارزشیابی تأثیر، و بررسی کیفیت مطالعات ارزشیابی تأثیر داخلی و بررسی وضعیت تسلط مطالعات ارزشیابی‌شده در گویه‌های ارزشیابی تأثیر، انجام شده‌است.

در **گام نخست**، این پژوهش به بیان مسائل مهم پیرامون تعاریف ارزشیابی تأثیر، دسته‌بندی ارزشیابی تأثیر بر حسب رویکرد (شامل رویکرد کمی، کیفی و ترکیبی)، دسته‌بندی ارزشیابی تأثیر بر حسب زمان (گذشته‌نگر و آینده‌نگر)، دسته‌بندی ارزشیابی تأثیر بر حسب ارزیابان (داخلی، خارجی و مشارکتی)، دلایل انجام ارزشیابی تأثیر، توجه به اصول اخلاقی ارزشیابی تأثیر، ضرورت نگارش طرح ارزشیابی تأثیر، ضرورت ارائه نظریه تغییر در ارزشیابی تأثیر، و در نهایت چالش‌های ارزشیابی تأثیر پرداخته شده‌است. در **گام دوم**، با استفاده از راهبرد آلتاید و با مطالعه دقیق، کدگذاری، و مرور مجدد و مکرر ۱۶ سند بین‌المللی به تدوین ۳۵ گویه مهم در ارزشیابی تأثیر پرداخته‌است و در **گام سوم**، کیفیت ۱۸ مطالعه ارزشیابی بر اساس این ۳۵ گویه بررسی شده‌است. در نهایت در **گام چهارم**، با انجام آزمون دوجمله‌ای وضعیت دستیابی مطالعات در هر گویه به حد تسلط مشخص شده‌است.

در نتیجه تحلیل یافته‌های پژوهش حاضر مشخص شد که در ۶۵/۷۱ درصد از کل گویه‌ها (یعنی ۲۳ گویه) هیچ پژوهشی به حد تسلط دست نیافته‌است و تنها در ۵/۷۱ درصد از کل گویه‌ها (یعنی تنها ۲ گویه)، نیمی از پژوهش‌ها (۹ پژوهش معادل ۵۰ درصد از پژوهش‌ها) به حد تسلط دست یافته‌اند. در هر کدام از ۱۰ گویه دیگر، پژوهش‌ها به میزان مختلف به حد تسلط دست یافته‌بودند که این میزان کمتر از ۵۰ درصد از پژوهش‌ها است. کمترین میزان دستیابی پژوهش‌ها به حد تسلط در گویه شماره ۱۵ «به شفافیت در مورد محدودیت‌ها توجه شده‌است» و گویه شماره ۳۲ «به رعایت موارد اخلاقی ارزشیابی اشاره شده‌است» با ۵/۵۵ درصد و بیشترین میزان آن در گویه شماره ۵ «در مطالعه اهداف ارزشیابی به طور واضح بیان شده‌است» و گویه شماره ۲۵ «ارزشیابی در پر کردن شکاف‌های دانش نقشی ایفا نموده‌است» با ۲۷/۷۷ درصد بوده‌است.

بدیهی است غفلت از هر کدام از گویه‌های فهرست‌واری ارزشیابی موجود (که بر اساس بررسی دقیق منابع مهم ارزشیابی تأثیر ایجاد شده‌است)، پژوهش ارزشیابی تأثیر را با مشکلات متعددی روبرو می‌کند و هویت اصیل ارزشیابی تأثیر را مخدوش می‌نماید. به عنوان مثال غفلت از مشخص نمودن روش ارزشیابی، اهداف ارزشیابی، کاربرد ارزشیابی، کاربران و ذی‌نفعان ارزشیابی، سؤالات کلیدی ارزشیابی، هزینه‌ها و بودجه‌بندی ارزشیابی، زمان‌بندی ارزشیابی، فرایند دقیق و جزئیات ارزشیابی و همچنین فقدان طرح ارزشیابی، ابزار ارزشیابی (مانند فهرست‌واری)، و تیم ارزیابی آموزش دیده، صحت ارزشیابی را دچار خدشه می‌کند و جایگاه ارزشیابی را به یک بررسی عامیانه تقلیل می‌دهد. همچنین اگر مطالعه میان نتایج و تأثیرات تمایزی قائل نشود مشخص است منظور مطالعه از تأثیر همان نتایج حاصل از وجود یک رابطه میان دو یا چند متغیر است و به «تأثیر» آن گونه که بیان شد توجهی نداشته‌است. از سوی دیگر در صورت بی‌توجهی به دو مورد بسیار حساس «روایی» و «اعتبار»، یعنی قابلیت تعمیم و تکرار نتایج ارزشیابی در شرایط مشابه به خطر می‌افتد.

به طور کلی نتایج حاصل از چهار بخش این مطالعه حاکی از آن است که آگاهی و درک پژوهشگران از مفهوم اصلی «ارزشیابی تأثیر» ناقص، ناکافی و مبهم است که منشأ این اتفاق در فقدان آموزش صحیح روش‌های پژوهش در آموزش عالی و در نتیجه ناآگاهی و سرعت کند پژوهشگران در انطباق آنها با پیشرفت سریع روش‌های پژوهش و جزئیات و تمایزات آنهاست. پژوهش حاضر با بررسی عنوان و مطالعه اجمالی ۱۴۱ مقاله، بررسی کامل چکیده ۶۲ مقاله، بررسی کامل محتوای ۳۰ مقاله، و در نهایت ارزشیابی کامل ۱۸ مقاله بر اساس فهرست‌وارسی تهیه‌شده نتیجه نهایی این پژوهش را به این صورت ارائه می‌کند که منظور حداقل ۹۰ درصد از پژوهش‌هایی با عنوان «ارزشیابی تأثیر» و یا «سنجش تأثیر» با مفهوم اصلی «ارزشیابی تأثیر» کاملاً متفاوت است چرا که اولاً منظور از ارزشیابی و سنجش در این پژوهش‌ها یک بررسی اولیه بدون مشخص نمودن دقیق روش، اهداف، ابزار، فرآیند، تیم ارزیابی و موارد دیگر مؤثر در ارزیابی است، و ثانیاً منظور پژوهشگران از تأثیر آنچنان که منظور پژوهشگرانی مانند پیرزمن و همکاران (۲۰۱۵)، پیرزمن (۲۰۱۴)، بمبرگر (۲۰۱۲)، پرین (۲۰۱۲)، راجرز، (۲۰۱۲)، و وایت، (۲۰۱۰) است، یعنی «تأثیر» به اثرات بلندمدت اشاره دارد و معمولاً به سطح نهایی زنجیره علی گفته می‌شود، نیست و منظور آنها تنها بررسی نتایج یک رابطه میان دو یا چند متغیر است که آن هم نه با استفاده از روش آزمایشی بلکه با استفاده از روش پیمایشی و نهایتاً نیمه‌آزمایشی است.

پیشنهادهای پژوهشی و کاربردی

- ۱- با توجه به کمبود داده‌ها به ویژه در منابع داخلی در حوزه ارزشیابی تأثیر پیشنهاد می‌شود پژوهش‌هایی با رویکرد کیفی با تمرکز بر تجارب زیسته حرفه‌ای پژوهشگران در زمینه چپستی و چگونگی ارزشیابی تأثیر انجام شود.
- ۲- با توجه به محدودیت‌های ذکرشده در پژوهش حاضر، پیشنهاد می‌شود با استفاده از نظام نشانگرهای ارائه‌شده در این پژوهش (و ارتقاء کمی و کیفی آن در صورت لزوم)، این ارزیابی بر روی پژوهش‌های ارزشیابی تأثیر بیشتری در حوزه داخلی و بین‌المللی انجام شود تا ضمن افزایش دقت و اعتبار پژوهش حاضر، امکان مقایسه‌ای میان پژوهش‌های داخلی و بین‌المللی در حوزه ارزشیابی تأثیر فراهم شود.
- ۳- پیشنهاد می‌شود از طریق واحدهای روش تحقیق و یا کارگاه‌های آموزشی روش تحقیق، معنا و مفهوم، اهمیت، اصول، و فرآیند انجام پژوهش ارزشیابی تأثیر در دوره‌های تحصیلات تکمیلی آموزش داده شود.
- ۴- به سبب توجه بسیار پایین مقالات به اصول اساسی ارزشیابی تأثیر و درصد رعایت بسیار پایین گویه‌های تدوین شده که نشانگر اطلاع ناکافی و بسیار اندک پژوهشگران از فرآیند ارزشیابی تأثیر و در نتیجه ارائه کارهای علمی به دور از اصول و قواعد اساسی آن است، پژوهش حاضر توصیه به رعایت اصول مهم ارزشیابی تأثیر در پژوهش‌های این حوزه می‌پردازد تا هم پژوهشگران ضمن توجه به این اصول و رعایت گویه‌های مذکور در مطالعات خود، و هم داوران با توجه به این اصول در تدوین کاربرگ‌های

داوری مقالاتی با موضوع «ارزشیابی یا سنجش تأثیر» تحول شگرفی در زمینه ارزشیابی تأثیر ایجاد نمایند و در سال‌های آتی شاهد ارائه مقالات علمی با کیفیت بسیار بالا در زمینه ارزشیابی تأثیر در حوزه‌های مهم و تأثیرگذار علمی باشیم. بر همین اساس اصول مهم ارزشیابی تأثیر به شرح زیر شناسایی و معرفی می‌شود:

- روشن نمودن اهداف ارزشیابی تأثیر و توجه به انطباق ارزشیابی با اهداف.
- ارائه یک نظریه تغییر واضح و صریح در مطالعه ارزشیابی تأثیر.
- تهیه طرح ارزشیابی پیش از شروع ارزشیابی و انجام ارزشیابی بر اساس آن.
- مشخص نمودن ابزار ارزشیابی (در صورت وجود) و یا تهیه ابزار (به عنوان مثال یک فهرست‌وارسی) ارزشیابی تأثیر با گویه‌های واضح و وزن‌دهی شده.
- توجه به چندسویه‌سازی منابع گردآوری داده‌ها و روش ارزشیابی.
- توجه به تأثیرات بلندمدت و پایدار در طول زمان و تفکیک میان نتایج و تأثیرات.
- توجه به تأثیرات ناخواسته ارزشیابی علاوه بر تأثیرات خواسته و مورد انتظار.
- توجه به شفافیت روند و فرآیند انجام کار و فرصت‌ها و محدودیت‌های ارزشیابی.
- توجه به زمان مناسب ارزشیابی پس از اجرای طرح.
- تشکیل تیم ارزشیابی و نظارت و آموزش اعضای هر دو تیم.
- بررسی روایی و پایایی ارزشیابی تأثیر به منظور اطمینان از تعمیم‌پذیری نتایج و تکرارپذیری نتایج آن.

References

منابع

- آقاکاردان، احمد، و کاظمی‌آرانی، علی. (۱۳۹۶). سنجش میزان تأثیر بازی‌گونه‌سازی آزمون‌های الکترونیکی در یادگیری دانش‌آموزان، پژوهش در یادگیری آموزشگاهی و مجازی، ۵، ۶۴-۵۵.
<https://ensani.ir/fa/article/372654/>
- بابایی، سهراب، شکیبازاده، الهام، شجاعی‌زاده، داوود، یاسری، مهدی، و محمدزاده، علیرضا. (۱۳۹۹). ارزشیابی تأثیر مداخله آموزشی مبتنی بر الگوی اعتقاد بهداشتی بر رفتارهای ارتقاءدهنده سبک زندگی سالم در افراد مستعد بیماری قلبی-عروقی، آموزش بهداشت و ارتقاء سلامت ایران، ۸(۳)، ۲۳۹-۲۲۴.
 DOI: 10.29252/ijhehp.8.3.224
- بارانی‌پسیان، وحید، نامی، محمدحسن، و توکلی‌نغمه، مصطفی. (۱۳۹۶). بررسی و سنجش تأثیر عوامل اجتماعی-اقتصادی بر سلامت شهری (شهر مقدس قم). نگرش‌های نو در جغرافیای انسانی (جغرافیای انسانی)، ۹(۳)، ۴۱-۵۵.
 URL: <https://sid.ir/paper/505225/fa>

- بحرانی، عطیه، فروتنی، زهرا، و بوجالالی، زینب. (۱۴۰۱). سنجش تأثیر سواد رسانه‌ای در استفاده از فضای مجازی بر اخلاق اجتماعی کاربران (مورد مطالعه: دانشجویان کارشناسی ارشد دانشگاه آزاد اسلامی واحد تهران جنوب)، فصلنامه فرهنگ در دانشگاه اسلامی، ۱۱(۴)، ۶۷۶-۶۵۳.
URL: https://ciu.nahad.ir/article_983.html
- حبیب‌پورگتابی، کرم، و صفری‌شالی، رضا. (۱۴۰۱). راهنمای جامع کاربرد SPSS در تحقیقات پیمایشی (تحلیل داده‌های کمی)، لویه متفکران. URL: <https://www.gisoom.com/book/11169336/>
- حقوقی، سمیرا، و صالحی، کیوان. (۱۳۹۸). ظرفیت‌های روش پژوهش روایتی در حوزه مطالعات آموزش عالی، نامه آموزش عالی، ۱۲(۴۷)، ۳۳-۵۱.
URL: https://journal.sanjesh.org/article_37138.html?lang=en
- دکامینی، فاطمه، و احسانی‌فر، محمد. (۱۴۰۰). سنجش تأثیر بلوغ مدیریت دانش بر کیفیت خدمات در بانک پارسیان، نشریه بازیابی دانش و نظام‌های معنایی، ۸(۲۸)، ۷۱-۱۱۱.
DOI: 10.22054/jks.2021.60865.1438
- رجبی‌فرجاد، حاجیه، و پوره‌اشمی خشک‌ناب، سیدحسین. (۱۳۹۵). سنجش تأثیر فناوری اطلاعات بر بهره‌وری کارکنان. مدیریت منابع در نیروی انتظامی، ۴(۱)، ۱۹۳-۱۶۹.
URL: <https://sid.ir/paper/519300/fa>
- رسولی، سیده‌عصمت، و ذوالفقاری، فریبا. (۱۳۹۶). تأثیر فرا ارزشیابی در برنامه‌ریزی درسی، همایش بین‌المللی روان‌شناسی و مطالعات اجتماعی، تهران. URL: <https://civilica.com/doc/747531>
- رضایی، محمدرضا؛ مرادی‌رمقانی، رضا. (۱۴۰۱). سنجش تأثیر عملکرد مؤلفه‌های امنیت مقصدهای گردشگری بر رضایت‌مندی گردشگران خارجی در راستای توسعه پایدار صنعت گردشگری کلان‌شهر شیراز. جغرافیا و پایداری محیط، ۱۲(۴)، ۲۲-۱۰.
URL: <https://sid.ir/paper/1007548/fa>
- زارع، ابراهیم، شجری، شاهرخ، رحمانی، صفت‌اله، محدث‌حسینی، سیداحمد، سالم، جلال، و انویه‌تکیه، لورنس. (۱۴۰۳). ارزشیابی تأثیر طرح سربازان سازندگی در بخش کشاورزی. آموزش و ترویج نوین کشاورزی، ۲(۱)، ۶۳-۷۸.
URL: https://ameej.areeo.ac.ir/article_132376.html
- سبزواری، سکینه. (۱۳۸۴). فرا ارزشیابی و معیارهای آن. گام‌های توسعه در آموزش پزشکی، ۲(۲)، ۱۱۶-۱۲۴.
URL: <https://sid.ir/paper/116973/fa>
- سیف، علی‌اکبر. (۱۳۹۶). اندازه‌گیری، سنجش و ارزشیابی آموزشی، دوران. URL: <https://www.adinehbook.com/gp/product/6006208113>
- سیف، علی‌اکبر. (۱۳۹۴). روان‌شناسی پرورشی نوین (روان‌شناسی یادگیری و آموزش)، دوران. URL: <https://www.adinehbook.com/gp/product/6006208213>
- شهرام‌نیا، امیرمسعود، ابراهیمی‌پور، حوا، محموداوغلی، رضا، و ملکان، مجید. (۱۳۹۶). سنجش تأثیر شبکه‌های اجتماعی بر مشارکت سیاسی: مطالعه موردی دانشجویان دانشگاه مازندران. جامعه‌شناسی

- کاربردی (مجله پژوهشی علوم انسانی دانشگاه اصفهان)، ۲۸ (۱) (پیاپی ۶۵)، ۱۹-۳۲. DOI: 10.22108/jas.2017.21248
- شهیدی، سپیده، بشیریه، تهمورث، صابری، سیدمهدی، و عباسی، اصغر. (۱۳۹۹). سنجش تأثیر اعتیاد و وضعیت سرپرستی بر کودک آزاری جنسی با تمرکز بر نمونه‌های به دست آمده از مراکز اورژانس اجتماعی تهران. *پژوهش‌های حقوق جزا و جرم‌شناسی*، ۸ (۱۶)، ۲۸۷-۲۶۷. DOI: 10.22034/jclc.2021.125314
- صادقی‌مزی‌دی، حسین، و ابراهیمیان، صدیقه. (۱۴۰۰). سنجش تأثیر آموزش برخط (آنلاین) بر میزان یادگیری درس زبان انگلیسی مدارس شرق استان فارس، *پیشرفت‌های نوین در روانشناسی*، ۴ (۴۵)، ۸۳-۹۱. URL: <https://ensani.ir/fa/article/498664/>
- صیفوری، ویدا. (۱۳۹۹). سنجش تأثیر معماری اطلاعات کتابخانه‌های دیجیتال بر یادگیری الکترونیکی (نمونه پژوهشی: سامانه منبع‌یاب وزارت بهداشت، درمان و آموزش پزشکی). *مدیریت اطلاعات*، ۶ (۱)، ۱۵۴-۱۳۵. DOI: 10.22034/aimj.2020.124930
- عباسی شوازی، محمدجلال، و خواجه صالحی، زهره. (۱۳۹۲). سنجش تأثیر استقلال، مشارکت اجتماعی، و تحصیلات زنان بر تمایل به فرزندآوری (مطالعه موردی شهر سیرجان). *زن در توسعه و سیاست (پژوهش زنان)*، ۱۱ (۱)، ۶۴-۴۵. DOI: 10.22059/jwdp.2013.35477
- علی‌میرزایی، مولود، پیرحیاتی، سارا، و صالحی، کیوان. (۱۳۹۴). ارزشیابی تأثیر در نظام آموزش عالی: چپستی، چرایی و چگونگی، چهارمین کنفرانس بین‌المللی روان‌شناسی و علوم اجتماعی، تهران. URL: <https://civilica.com/doc/492280>
- قنبری، عاطفه، مودب، فاطمه، جوادی، نازیلا، حسین زاده، طوبی، موسوی، ساقی، و براری، فاطمه. (۱۳۹۸). ارزشیابی تأثیر برگزاری کارگاه ایمنی بیمار بر آگاهی و عملکرد دانشجویان پرستاری: چالش‌ها و راهکارها. *توسعه‌ی آموزش جندی شاپور*، ۱۰ (۲)، ۹۷-۱۰۳. DOI: 10.22118/edc.2019.90844
- محمدی‌زیدی، عیسی، پاک‌پورحاجی‌آقا، امیر، و محمدی‌زیدی، بنفشه. (۱۳۹۱). ارزشیابی تأثیر برنامه آموزشی بر اساس تئوری رفتار برنامه‌ریزی‌شده بر عملکرد ایمن کارگران. *مجله دانشگاه علوم پزشکی مازندران (نامه دانشگاه)*، ۲۲ (۹۷)، ۱۷۷-۱۶۶. URL: <http://jmums.mazums.ac.ir/article-1-1775-fa.html>
- نجان‌نهادی، مریم، و آبخوش، معصومه. (۱۳۹۴). سنجش تأثیر آموزش شهروندی بر توانمندسازی زنان (مطالعه موردی: منطقه ۶ شهرداری تهران). *زن در توسعه و سیاست (پژوهش زنان)*، ۱۳ (۱)، ۱۳۱-۱۴۹. DOI: 10.22059/jwdp.2015.55210
- نوابخش، مهرداد، و ثابتی، مریم. (۱۳۹۴). طراحی و ارائه الگو جهت سنجش تأثیر سرمایه فرهنگی بر ابعاد توسعه پایدار شهری با تأکید بر انرژی‌های پاک. *مطالعات توسعه اجتماعی ایران*، ۸ (۱)، ۵۰-۳۷. DOI: <https://sid.ir/paper/231844/en>

- هنری، فاطمه، میری، محمدرضا، معاشری، بی بی نرگس. (۱۳۹۵). ارزشیابی تأثیر آموزش بر اساس تئوری رفتار برنامه‌ریزی شده بر انجام مراقبت‌های دوران بارداری زنان باردار معتاد. *مجله علمی دانشگاه علوم پزشکی بیرجند*، ۲۳ (۱): ۶۷-۷۷. URL: <http://journal.bums.ac.ir/article-1-1997-fa.html>
- Abbasi Shavazi, M. J., & Khajeh salehi, Z. (2013). An Assessment on the Impact of Women's Autonomy, Education and Social Participation on Childbearing Intention in Sirjan City. *Woman in Development & Politics*, 11(1), 45-64. [In Persian] DOI: 10.22059/jwdp.2013.35477
- A. Kardan, A., & Kazemi-Arani. A. (2017). Estimating the Impact of Electronic Exams Gamification on Students' Learning. *Quarterly Journal of Research in School and Virtual Learning*, 5, 55-64. [In Persian] URL: <https://ensani.ir/fa/article/372654/>
- Ali Mirzaei, M., & Pirhayati, S., & Salehi, K. (2015). *Impact Evaluation in the Higher Education System: What, Why and How*, Fourth International Conference on Psychology and Social Sciences, Tehran, [In Persian] URL: <https://civilica.com/doc/492280>
- Babaei, S., Shakibazadeh, E., Shojaeizadeh, D., Yaseri, M., & Mohamadzadeh, A. (2020). Effectiveness the Theory-Based Intervention Based on Health Belief Model on Health Promotion Lifestyle in Individuals Susceptible to Cardiovascular Diseases. *Iran J Health Educ Health Promot*, 8 (3), 224-239. [In Persian] DOI: 10.29252/ijhehp.8.3.224
- Bahrani, A., Forutani, Z., & Boujalali, Z. (1400). Measuring the impact of media literacy in the use of cyberspace on users' social ethics (Case study: Master's students of Islamic Azad University, South Tehran Branch). *Culture in Islamic University*, 11 (41), 653-676. [In Persian] URL: https://ciu.nahad.ir/article_983.html
- Bamberger, M. (2012). Introduction to mixed methods in impact evaluation. *Impact Evaluation Notes*, 3(3), 1-38. URL: <https://www.interaction.org/wp-content/uploads/2019/03/Mixed-Methods-in-Impact-Evaluation-English.pdf>
- Barani Pesian, V., Nami, M. H., & Tavakoli Naghmeh, M. (2017). Study and measurement of the impact of socio-economic factors on urban health (Holy city of Qom). *New Perspectives in Human Geography (Human Geography)*, 9(3), 41-55. [In Persian] URL: <https://sid.ir/paper/505225/fa>
- Baylis, K., Honey- Rosés, J., Börner, J., Corbera, E., Ezzine- de- Blas, D., Ferraro, P. J., ... & Wunder, S. (2016). Mainstreaming impact evaluation in nature conservation. *Conservation Letters*, 9(1), 58-64. DOI: 10.1111/conl.12180
- Copestake, J. 2014. Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches. *Evaluation*, 20(4), 412-427. DOI: 10.1177/1356389014550559
- Crato, N., & Paruolo, P. (2019). *Data-driven policy impact evaluation: How access to microdata is transforming policy design* (p. 346). Springer Nature. URL: <https://link.springer.com/book/10.1007/978-3-319-78461-8>

- Dekamini, F. & Ehsanifar, M. (2021). Evaluating the Effect of Knowledge Management Maturity on Service Quality in Parsian Bank. *Knowledge Retrieval and Semantic Systems*, 8(28), 71-111. [In Persian] DOI: 10.22054/jks.2021.60865.1438
- Frölich, M., & Sperlich, S. (2019). *Impact evaluation: Treatment Effects and Causal Analysis*. Cambridge University Press. DOI: 10.1017/9781107337008
- Garbarino, S., & Holland, J. (2009). *Quantitative and qualitative methods in impact evaluation and measuring results*. URL: <http://epapers.bham.ac.uk/646/1/eirs4.pdf>
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). *Impact evaluation in practice*. World Bank Publications. URL: <https://documents1.worldbank.org/curated/en/823791468325239704/pdf/Impact-evaluation-in-practice.pdf>
- Ghanbari, A., Mudaveb, F., Javadi, N., Hosseinzadeh, T., Mousavi, S., & Barari, F. (2019). Evaluating the impact of holding a patient safety workshop on nursing students' awareness and performance: Challenges and solutions. *Jundishapour Education Development*, 10(2), 97-103. [In Persian] DOI: 10.22118/edc.2019.90844
- Greene, J. C. (2007). *Mixed methods in social inquiry*. John Wiley & Sons. URL: <https://www.wiley.com/en-us/Mixed+Methods+in+Social+Inquiry-p-9780787983826>
- Gugerty, M. K., Karlan, D., & Welsh, D. (2016). Organizational Challenges of Impact Evaluation. GOLDILOCKS DEEP DIVE. IPA Communications Team. URL: https://poverty-action.org/sites/default/files/publications/Goldilocks-Deep-Dive-Organizational-Challenges-of-Impact-Evaluation_0.pdf
- Habibpourgtabi, K., & Safarishali, R. (2022). *A comprehensive guide to using SPSS in survey research (quantitative data analysis)*. Loya. [In Persian] URL: <https://www.gisoom.com/book/11169336/>
- Hoghoughi, S. & Salehi, K. (2019). Capacities of the narrative research method in the Field of Higher Education Studies. *Higher Education Letter*, 12(47), 33-51. [In Persian] URL: https://journal.sanjesh.org/article_37138.html?lang=en
- Honari, F., Miri, M., & Moasheri, B. N. (2016). Evaluating the effect of educational intervention based on Theory of Planned Behavior on prenatal care of addicted pregnant women. *Journals of Birjand University of Medical Sciences*, 23 (1), 67-77. [In Persian] URL: <http://journal.bums.ac.ir/article-1-1997-fa.html>
- Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2009). *Handbook on impact evaluation: quantitative methods and practices*. World Bank Publications. URL: <https://documents1.worldbank.org/curated/en/650951468335456749/pdf/520990PUB0EPI1101Official0Use0Only1.pdf>
- Michel, J., & Schneider, K. (2025). Demystifying impact evaluation: an impact evaluation framework. *Frontiers in epidemiology*, 5, 1460997. DOI: 10.3389/fepid.2025.1460997.

- Mohammadi Zaidi, I., Pakpourhaji Agha, A., & Mohammadi Zaidi, B. (2012). Evaluating the impact of a training program based on the theory of planned behavior on the safe performance of workers. *Journal of Mazandaran University of Medical Sciences (University Letter)*, 22(97), 166-177. [In Persian] URL: <http://jmums.mazums.ac.ir/article-1-1775-fa.html>
- Najjar Nahavandi, M. & Abkhosh, M. (2015). An assessment of the effect of citizenship education on empowering women (Case study: Tehran 6th district). *Woman in Development & Politics*, 13(1), 131-149. [In Persian] DOI: 10.22059/jwdp.2015.55210
- Navabakhsh, M., & Sabeti, M. (2016). Design and providing a model to assess the effects of cultural capital on aspects of sustainable urban development with an emphasis on clean energy. *Journal of Iranian Social Development Studies (JISDS)*, 8(1 (29)), 37-50. [In Persian] DOI: <https://sid.ir/paper/231844/en>
- Peersman, G., Guijt, I., & Pasanen, T. (2015). Evaluability assessment for impact evaluation: Guidance, checklists and decision support, A Methods Lab Publication. URL: <https://media.odi.org/documents/9802.pdf>
- Peersman, G. (2014). Overview: Data collection and analysis methods in impact evaluation. UNICEF Office of Research-Innocenti. URL: https://www.betterevaluation.org/sites/default/files/Data_collection_and_Analysis_ENG.pdf
- Perrin, B. (2012). Linking monitoring and evaluation to impact evaluation. *Impact Evaluation Notes*, 2, 1-22. URL: <https://www.interaction.org/wp-content/uploads/2019/03/Linking-Monitoring-and-Evaluation-to-Impact-Evaluation.pdf>
- Pozarny, P., & Barrington, C. (2016). *Qualitative Methods in Impact Evaluations of Cash Transfer Programmes in the Transfer Project in Sub-Saharan Africa*. (Davis, Benjamin; Handa, Sudhanshu; Hypher, Nicola; Winder Rossi, Natalia; Winters, Paul; & Yablonski, Jennifer (Eds.). pp. 71-93. Oxford, England: FAO. URL: <http://dx.doi.org/10.1093/acprof:oso/9780198769446.003.0004>
- Rajabi Farjad, H., & Pourhashmi Khoshknab, S. H. (2016). Measuring the impact of information technology on employee productivity. *Resource Management in the Police Force*, 4(1), 169-193. [In Persian] URL: <https://sid.ir/paper/519300/fa>
- Rasouli, S. E., & Zolfaghari, F. (2017). *The Effect of Meta-Assessment in Curriculum Planning*, International Conference on Psychology and Social Studies, Tehran. [In Persian] URL: <https://civilica.com/doc/747531>
- Ravallion, Martin. (2008). "Evaluating Anti-Poverty Programs." In *Handbook of Development Economics*, vol. 4, ed. T. Paul Schultz and John Strauss, 3787-846. Amsterdam: North-Holland. URL: <https://documents1.worldbank.org/curated/en/104761468315569641/pdf/wps3625.pdf>
- Reed, M. S., Ferré, M., Martin-Ortega, J., Blanche, R., Lawford-Rolfe, R., Dallimer, M., & Holden, J. (2021). Evaluating impact from research: A methodological

- framework. *Research Policy*, 50(4), 1-14. URL: www.elsevier.com/locate/respol
- Rezaei, M. R., & Moradi Ramghani, R. (2022). Measuring the impact of the performance of security components of tourism destinations on the satisfaction of foreign tourists in line with the sustainable development of the tourism industry in Shiraz metropolis. *Geography and Environmental Sustainability (Geographical Research Journal)*, 12(45), 1-22. [In Persian] URL: <https://sid.ir/paper/1007548/fa>
- Rogers, P. J. (2014). Overview of impact evaluation. *UNICEF Office of Research-Innocenti*. URL: <https://alnap.cdn.ngo/media/documents/brief-1-overview-eng.pdf>
- Rogers, P. J. (2012). Introduction to impact evaluation. *Impact Evaluation Notes*, 1, 1-18. URL: <https://www.interaction.org/wp-content/uploads/2019/03/1-Introduction-to-Impact-Evaluation.pdf>
- Sabzevari, S. (2005). Meta-evaluation and its criteria. *STRIDES IN DEVELOPMENT OF MEDICAL EDUCATION*, 2(2), 116-124. [In Persian] URL: <https://sid.ir/paper/116973/fa>
- Sadeghi Mazedi, H., & Ebrahimian, S. (2021). Measuring the impact of online education on the level of English language learning in schools in eastern Fars Province, *New Advances in Psychology, Educational Sciences and Education*, 4(45), 83-91. [In Persian] URL: <https://ensani.ir/fa/article/498664/>
- Saif, A.A., (2017). *Educational measurement, assessment and evaluation book*. Douran. [In Persian] URL: <https://www.adinehbook.com/gp/product/6006208113>
- Saif, A.A., (2015). *The book of modern educational psychology: psychology of learning and education*. Douran. [In Persian] URL: <https://www.adinehbook.com/gp/product/6006208213>
- Salehi, K., & Golafshani, N. (2010). Commentary: Using mixed methods in research studies: An opportunity with its challenges. *International journal of multiple research approaches*, 4(3), 186-191. DOI: 10.5172/mra.2010.4.3.186
- Samuel, G. N., & Derrick, G. E. (2015). Societal impact evaluation: Exploring evaluator perceptions of the characterization of impact under the REF2014. *Research evaluation*, 24(3), 229-241. DOI: 10.1093/reseval/rvv007
- Seifouri, V. (2020). Assessing the Impact of Digital Library Information Architecture on E-learning (Case Study: System of Resource Finder the Ministry of Health and Medical Education). *Iranian Journal of Information Management*, 6(1), 135-154. [In Persian] DOI: 10.22034/aimj.2020.124930
- Shahidi, S., Bashirieh, T., Saberi, S. M., & Abbasi, A. (2019). Measuring the impact of addiction and guardianship status on child sexual abuse with a focus on samples obtained from Tehran social emergency centers. *Criminal Law and*

-
- Criminology Research*, 8 (16), 267-287. [In Persian] DOI: 10.22034/jclc.2021.125314
- Shahramnia, A. M., Ebrahimipour, H., Mahmoudovgholi, R., & Malekan, M. (2017). Measuring the impact of social networks on political participation: A case study of students at the University of Mazandaran. *Applied Sociology*, 28 (1), 19-32. [In Persian] DOI: 10.22108/jas.2017.21248
- White, H. (2010). A contribution to current debates in impact evaluation. *Evaluation*, 16(2), 153-164. DOI: 10.1177/1356389010361562
- Zare, E., Shajari, Sh., Rahmani, S., Mohaddes Hosseini, S. A., Salem, J., & Anouyeh Tekey, L. (1403). Evaluating the impact of the construction soldiers project in the agricultural sector. *Modern Agricultural Education and Extension*, 2(1), 63-78. [In Persian] URL: https://ameej.areeo.ac.ir/article_132376.html

