

مفهوم‌شناسی ارزشیابی و تفاوت آن با سایر مفاهیم

دکتر علیرضا کیامنش^۱

مقدمه

قبل از پرداختن به مفهوم ارزشیابی لازم است که بین مفهوم عامیانه و روزمره که بیشتر فردی و تجربی است و بر اساس احساس و معیارهای شخصی صورت می‌گیرد با ارزشیابی حرفه‌ای که به صورت هدفمند صورت می‌گیرد تفاوت قائل شد. همچنین برای درک بهتر مفهوم ارزشیابی باید مشابهت‌ها و تفاوت‌های این مفهوم با سایر مفاهیم نظیر نظارت، ارزیابی، سنجش، پژوهش، پاسخگویی، آزمون و اندازه‌گیری روشن شود.

تعاریف ارزشیابی

ارزشیابی حرفه‌ای با هدف خاص انجام می‌گیرد و مستلزم جمع‌آوری و سنجش اطلاعات به صورتی روشمند و برنامه‌ریزی شده است. در فرهنگ‌های لغت، ارزشیابی به صورت عام «تعیین ارزش کردن، امتحان و قضاوت کردن» و به صورت خاص «تعیین میزان موفقیت یک برنامه، یک ...، در رسیدن به هدف‌های تعیین شده» تعریف شده است. تایلر^۲ (۱۹۵۰) ارزشیابی را فرایند تعیین میزان انطباق هدف‌های آموزشی و عملکرد برنامه می‌داند. ویلی (۱۹۷۰)، به نقل از استنهاوس^۳، (۱۹۷۵) از ترکیب تعریف هریس (۱۹۶۳) «ارزشیابی تلاش منظم برای جمع‌آوری شواهد در مورد تغییرات حاصل از تجارب آموزشی در رفتار دانش‌آموزان» و کرناخ (۱۹۶۳) «ارزشیابی جمع‌آوری و استفاده از اطلاعات برای تصمیم‌گیری در مورد برنامه‌های آموزشی» ارزشیابی را جمع‌آوری و استفاده از اطلاعات در مورد تغییرات بوجود آمده در رفتار دانش‌آموزان برای تصمیم‌گیری در مورد برنامه آموزشی می‌داند.

دانشنامه ارزشیابی (ماتیسون^۴، ۲۰۰۵)، ارزشیابی را فرایند جستجوی کاربردی برای جمع‌آوری و تحلیل شواهد در جهت رسیدن به نتیجه نهایی در مورد ارزش، شایستگی، معنی‌دار بودن، کیفیت برنامه، تولید، فرد، خط مشی، یا برنامه و نقشه (هر آنچه که موضوع ارزشیابی است) تعریف کرده است. از دید پاتن، ۲۰۰۸ (به نقل از مرتنز^۵، ۲۰۱۰) شایستگی بیانگر ارزش درونی یک برنامه (برآورده کردن نیاز مخاطبان برنامه) و ارزش به ارزش بیرونی برای افراد بیرون برنامه (جامعه به معنی وسیع کلمه) اشاره دارد. به عبارتی شایستگی بیانگر کیفیت مطلق یا نسبی موضوع ارزشیابی شده با توجه به معیاری خاص است و ارزش بیانگر بازده ارزشیابی است. این دو وابسته به یکدیگر نیستند. یک برنامه یا فعالیت ممکن است دارای شایستگی ولی کم ارزش و یا دارای ارزش ولی فاقد شایستگی باشد.

اسکریون به نقل از الکین^۶ (۲۰۱۳) هدف ارزشیابی را همیشه قضاوت می‌داند و بر این اساس ارزشیابی را «قضاوت در مورد ارزش^۷ یا شایستگی^۸ موضوع مورد ارزشیابی (هر چیزی که ارزشیابی می‌شود)» تعریف می‌کند. وی ارزشیابی را بدون ارزشگذاری، ارزشیابی نمی‌داند. بد، بد است و خوب و خوب.

۱. استاد دانشگاه آزاد اسلامی واحد علوم و تحقیقات، تهران

1. Tyler
3. Stenhouse
4. Mathison
5. Mertens
6. Alkin
7. value
8. merit

لذا ارزشیاب وظیفه دارد که نشان دهد بد کدام است و خوب کدام. استافل بیم و شینک فیلد^۹ (۱۹۸۵) در یک تعریف جامع، ارزشیابی را «فرایند تعیین کردن، به دست آوردن، و فراهم ساختن اطلاعات توصیفی و قضاوتی در مورد ارزش و مطلوبیت هدفها، طرح، اجرا و نتایج به منظور هدایت تصمیم‌گیری خدمت به نیازهای پاسخگویی^{۱۰} و درک بیشتر از پدیده‌های مورد بررسی» تعریف کرده‌اند.

از دید فیتزپاتریک، سندرز و ورتن^{۱۱} (۲۰۰۴) کار ارزشیابی «مشخص کردن و روشن کردن معیارهای عملی و قابل دفاع، جمع‌آوری اطلاعات و تعیین ارزش موضوع مورد ارزشیابی در ارتباط با این معیارها» است. هولی، هارتی و نیو کامر^{۱۲} (۲۰۱۰) برای فعالیت‌های ارزشیابی دو دلیل ذکر کرده‌اند:

۱) کسب اطلاعات بیشتر برای پاسخگویی در قبال هزینه‌های عمومی و یا کمک‌های مردمی هزینه شده و ۲) کمک به سازمان یا مؤسسه برای بهبود کارآمدی برنامه‌های اجرا شده. از دید خاندکر، کولوال و صمد^{۱۳} (۲۰۱۱) ارزشیابی فعالیتی است که در آن بازده-های حاصل از اجرای یک برنامه به صورت هدفمند و عینی مورد سنجش قرار می‌گیرند. به عبارتی تلاش مجری ارزشیابی آن است که نشان دهد در تغییرات ایجاد شده در برنامه هدف فقط حاصل اجرای فعالیت انجام شده است. الکین (۲۰۱۱) تعاریف ارائه شده برای ارزشیابی «قضاوت در مورد ارزش یا شایستگی برنامه ارزشیابی شده» را عبارتی در مورد هدف ارزشیابی می‌داند که به دو جنبه قضاوت در مورد ارزش درونی و ارزش بیرونی ارزشیابی اشاره دارد. وی ارزشیابی را یک جستجوی نظام‌مند می‌داند و برای آن یک فرایند در نظر می‌گیرد.

در مجموع از عمل ارزشیابی انتظار می‌آورد علاوه بر تعیین ارزش و شایستگی موضوع مورد ارزشیابی، با فراهم آوردن اطلاعات توصیفی و قضاوتی در زمینه میزان کارآمد بودن برنامه در برآورده کردن نیازهای مصرف‌کنندگان برنامه و تحقق هدف‌های تعیین شده، تصمیم‌گیری را برای تصمیم‌گیران تسهیل نماید و به بهبود برنامه‌ها و بسط دانش در زمینه آنچه که ارزشیابی شده کمک کند. در حوزه آموزش از ارزشیابی برای تعیین میزان موفقیت دانش‌آموزان در رسیدن به هدف‌های تعیین شده (ارزشیابی دانش آموز) کارآمد بودن برنامه‌های مدرسه (ارزشیابی برنامه) و کارآمد بودن برنامه‌های درسی (ارزشیابی برنامه درسی) استفاده می‌شود.

تفاوت ارزشیابی با مفاهیم دیگر

اگر چه در منابع مختلف به صورت دقیق نمی‌توان بین بعضی از واژه‌ها یک خط و مرز مشخص و دقیق رسم نمود، ولی آشنایی با مفهوم هر واژه از دید صاحب‌نظران مختلف به درک بهتر هر یک از این مفاهیم می‌تواند کمک کند. در این نوشته به صورت مختصر تفاوت ارزشیابی با مفاهیم اندازه‌گیری^{۱۴}، آزمون^{۱۵}، سنجش^{۱۶}، ارزیابی^{۱۷}، نظارت^{۱۸}، پاسخگویی و پژوهش مورد بحث قرار گرفته است.

9. Stufflebeam & Shinkfield

10. Accountability

11. Fitzpatrick, Sanders & Worthen

12. Wholey, Hatry & Newcomer

13. Khandker, Koowal & Samad

14. Measurement

15. Testing

16. Assessment

17. Appraisal

18. Monitoring

اندازه‌گیری همانند ارزشیابی یک فرایند نظام‌دار است که با استفاده از قواعدی که صحت آن را می‌توان بررسی کرد به مقادیر مختلفی از یک صفت (مثلاً قد) که در یک فرد یا یک شیئی وجود دارد عدد (مثلاً ۱۶۵ سانتی متر) اختصاص داده می‌شود. عمل اندازه‌گیری امکان توصیف عینی‌تر پدیده‌های اندازه‌گیری شده و همچنین امکان مقایسه پدیده‌ها را فراهم می‌آورد. برای اندازه‌گیری بعضی از صفات (مانند هوش، استعداد، عملکرد تحصیلی) و بیان میزان این صفات به صورت کمی از **آزمون** استفاده می‌شود. حوزه و کاربرد آزمون از حوزه و مفهوم اندازه‌گیری محدودتر است. آزمون تنها یکی از ابزارهای اندازه‌گیری است. داده‌های کمی را می‌توان با توجه به صفت مورد اندازه‌گیری (مثلاً طول، سطح، حجم، رضایت شغلی و ...) با ابزارهای دیگر غیر از آزمون نیز به دست آورد. داده‌های کمی حاصل از آزمون و اندازه‌گیری غیر ارزشی بوده و در مورد این داده‌ها قضاوت و داوری صورت نمی‌گیرد. داشتن ابزار اندازه‌گیری شرط لازم برای به کمیت در آوردن پدیده‌ها و زیر بنای اصلی سنجش، ارزیابی، ارزشیابی، نظارت، پژوهش و پاسخگویی است. الکین (۲۰۱۱) آزمون را وسیله سنجش و ارزشیابی می‌داند.

مفاهیم سنجش، ارزیابی، ارزشیابی و نظارت در فرهنگ‌های لغت گاه مترادف و هم معنی بیان شده‌اند. در این چهار مفهوم، عناصر ارزش‌گذاری و قضاوت و برآوردهای کمی دیده می‌شود. ولف^{۱۹} (۱۹۸۴) با تفاوت قائل شدن بین هدف ارزشیابی و ارزیابی، معتقد است که ارزشیابی (قضاوت در باره ارزش یا موثر بودن یک برنامه) نباید جایگزین مسئولیت معلم در ارزیابی فراگیران (قضاوت در مورد هر فراگیر) شود. از دید وی نتایج ارزشیابی آموزشی که معمولاً از یک نمونه معرف دانش‌آموزی به دست می‌آید، به فرایند ارزیابی فرد فراگیران کمک نمی‌کند. ارزشیابی یک فعالیت بیرونی است که با قصد پی بردن به میزان موفقیت برنامه به اجرا در می‌آید، همانند ارزشیابی‌های در مقیاس بزرگ نظیر ارزشیابی آموزشی در سطح منطقه، ملی یا بین‌المللی. در این مطالعات از گروه نمونه برای جمع‌آوری اطلاعات استفاده می‌شود. درحالی‌که ارزیابی فعالیت‌های درونی است که در مدرسه برای ارزیابی عملکرد معلم و یا توسط معلم برای ارزیابی عملکرد آموزشی و رفتاری دانش‌آموزان انجام می‌گیرد (ارزیابی معلم در کلاس درس). برای کامل شدن کار ارزیابی سطح تبحر و توانایی فرد دانش‌آموزان اندازه‌گیری می‌شود.

شرنز، کلاس و توماس^{۲۰} (۲۰۰۳) برای استفاده از هر واژه به هدف یا موضوع بررسی توجه می‌کنند. از دید این افراد از واژه:

- **ارزشیابی** برای تمام موضوعات بخصوص ارزشیابی برنامه،
 - **ارزیابی** زمانی که موضوع بررسی معلم است،
 - **سنجش** زمانی که موفقیت تحصیلی هر یک از دانش‌آموزان مورد توجه است،
 - **نظارت** زمانی که فعالیت‌های روزمره نظام‌های آموزشی و سازمانها مورد توجه است،
- استفاده می‌شود.

الکین (۲۰۱۱) نیز معتقد است که در سه واژه سنجش، ارزیابی و ارزشیابی مفهوم ارزش‌گذاری (قضاوت در مورد ارزش و شایستگی) وجود دارد. وی از واژه ارزشیابی برای قضاوت در مورد برنامه، از واژه سنجش برای قضاوت در مورد مصرف کنندگان^{۲۱} برنامه نظیر سنجش ملی آموزش، یا سنجش دانش‌آموزان و از واژه ارزیابی برای قضاوت در مورد کارکنان برنامه استفاده می‌کند. به عبارتی برنامه‌ها ارزشیابی می‌شوند، دانش مصرف کنندگان سنجیده می‌شود و کارکنان برنامه ارزیابی می‌شوند. با بررسی منابع مختلف این نتیجه

19. Wolf

20. Scheerens, Glas & Thomas

21. Client

حاصل می‌شود که بین صاحب‌نظران مختلف در مورد مفهوم هر یک از واژه‌ها توافق نظر وجود دارد. این موضوع در منابع فارسی هم بوضوح دیده می‌شود.

پاسخگویی به عبارت ساده یعنی پاسخگو بودن مجریان برنامه در قبال کار انجام شده. در مفهوم وسیع‌تر پاسخگویی فراهم آوردن اطلاعات دقیق و مستند در مورد بازده‌های برنامه با هدف نشان دادن میزان و کیفیت بازده‌های حاصل از اجرای برنامه است. مثلاً معلم در قبال آموزشی که به دانش‌آموزان خود ارائه می‌دهد، با توجه به میزان کنترلی که بر عوامل مؤثر بر یادگیری و رفتار دانش‌آموزان دارد باید به اولیاء و مدرسه پاسخگو باشد. در تعریف استافیل بیم و شینگ فیلد (۱۹۸۵) یکی از هدف‌های ارزشیابی «خدمت به نیازهای پاسخگویی» ذکر شده است. این بدان معنی است که از بازده‌های ارزشیابی می‌توان در زمینه‌های مختلف از جمله عمل پاسخگویی بهره گرفت. به عبارتی ارزشیابی وسیله پاسخگویی است و از نتایج ارزشیابی در صورت نیاز می‌توان جهت پاسخگویی استفاده نمود.

اگر چه ارزشیابی و پژوهش در عمل از جمله فرایندهای منظم و هدفمند، جمع‌آوری اطلاعات، استفاده از نمونه، نمونه‌گیری، ابزارهای اندازه‌گیری، تحلیل داده‌های کمی و یا کیفی شباهت بسیار دارند، ولی این دو یکسان نیستند. یکی از تفاوت‌های عمده بین ارزشیابی و پژوهش در حوزه‌های مختلف را می‌توان در ماهیت سیاسی ارزشیابی جستجو کرد. در ارزشیابی علم و سیاست کاملاً در هم تنیده‌اند، به عبارتی یافته‌های ارزشیابی زمینه را برای تصمیم‌های سیاسی عمده در زمینه اصلاح، هدایت، تغییر و یا توقف فعالیت‌ها فراهم می‌آورد. در منابع مختلف از جمله ورتن ساندرز (۱۹۷۳)، پاپهام^{۲۲} (۱۹۷۵)، فیتزپاتریک، سندرز و ورتن (۲۰۰۴)، اسمیت و بران دون^{۲۳} (۲۰۰۸) و الکین (۲۰۱۱) به چند تفاوت عمده بین ارزشیابی و پژوهش اشاره شده است. تفاوت‌های عمده این دو مفهوم به صورت خلاصه عبارتند از:

۱) هدف پژوهش شناخت رابطه و در بسیاری موارد رابطه علت و معلولی بین متغیرها، بسط دانش و کمک به توسعه و رشد نظریه است. در مقابل هدف عمده ارزشیابی کمک به دست‌اندرکاران اجرای برنامه است تا با استفاده از اطلاعات فراهم آورده شده در مورد برنامه اجرا شده تصمیم‌گیری (ادامه، اصلاح و بازسازی یا توقف برنامه) کنند. در ارزشیابی ضمن توصیف پدیده مورد ارزشیابی، گاه با توجه به نیاز ذینفعان رابطه‌های علت و معلولی هم مورد توجه قرار می‌گیرند. به عبارت واضح‌تر پژوهش نتیجه-مدار و ارزشیابی تصمیم-مدار است.

۲) سؤال‌های پژوهش توسط خود پژوهشگر بیان و تدوین می‌شوند، در نتیجه برای فراهم آوردن پاسخ از نظر زمانی تحت فشار قرار نمی‌گیرد. در مقابل سؤال‌های ارزشیابی توسط سفارش‌دهندگان ارزشیابی طرح می‌شوند و مجری ارزشیابی باید به نیازها و خواسته‌های ذینفعان و سفارش‌دهندگان توجه کند. به همین دلیل مجری ارزشیابی باید در زمان تعیین شده، اطلاعات لازم را برای تصمیم‌گیری در مورد موضوع ارزشیابی فراهم آورد.

۳) کار پژوهشگر معمولاً با فراهم آوردن پاسخ سؤال یا سؤال‌ها و ارائه گزارش (بدون داوری در مورد خوب یا بد، مؤثر یا نامؤثر و با ارزش یا فاقد ارزش بودن یافته‌ها) پایان می‌یابد. روایی درونی و بیرونی یافته‌های پژوهشی و همچنین تصمیم‌پذیری یافته‌ها معیارهای قضاوت پژوهشگر می‌باشند. در مقابل مجری ارزشیابی با توجه به اطلاعات فراهم شده در مورد آنچه ارزشیابی شده قضاوت و داوری (خوب یا بد، مؤثر یا نامؤثر و با ارزش یا فاقد ارزش بودن یافته‌ها) می‌کند و در نهایت بهبود و اصلاح پدیده ارزشیابی شده را یکی از

22. Pop ham

23. Smith & Brandon

بخش‌های ارزشیابی می‌داند. معیارهای ارزشیابی عبارتند از درستی^{۲۴}، سودمندی^{۲۵}، عملی بودن^{۲۶} و اخلاقی و قانونی^{۲۷} بودن (کمیته مشترک تعیین معیارهای ارزشیابی، ۱۹۹۴).

۴) در پژوهش سؤال‌های (فرضیه‌های) مطرح شده مورد بررسی قرار می‌گیرند، در حالی که در ارزشیابی فرایند اجرای برنامه و میزان موفقیت برنامه در رسیدن به هدف‌های تعیین شده مورد سنجش، قضاوت و داوری قرار می‌گیرد. به عبارتی پژوهش در پی تأیید یا اثبات و ارزشیابی در پی بهبود و اصلاح است.

جمع بندی

تعاریف ارائه شده برای ارزشیابی به صورت مستقیم و یا غیر مستقیم در چند ویژگی با هم مشترک هستند، از جمله: جمع آوری منظم و هدفمند اطلاعات، قضاوت و داوری و در نهایت تصمیم‌گیری و اقدام عملی برای بهبود آنچه که ارزشیابی شده است. مجری ارزشیابی علاوه بر جمع آوری هدفمند اطلاعات در مورد هدف‌های موضوع مورد ارزشیابی، به معیار یا معیارهایی برای مقایسه و داوری نیاز دارد. قضاوت در مورد موضوع ارزشیابی، باید با تصمیم‌گیری همراه گردد. بدین ترتیب ارزشیابی فرایندی است منظم و هدفمند که مجری ارزشیابی با توجه به هدف‌های مورد نظر شواهد دقیق و معتبر فراهم شده را از طریق مقایسه با معیار یا معیارهای پذیرفته شده تفسیر می‌کند و در باره آنچه که ارزشیابی شده قضاوت می‌کند. عمل ارزشیابی با قضاوت ارزشی و تصمیم‌گیری یا اقدام عملی در مورد آنچه که ارزشیابی شده، تکمیل می‌گردد. با توجه به تصمیم‌گیری و اقدام عملی در مورد آنچه که ارزشیابی شده (ادامه، اصلاح و بهبود یا قطع و توقف) ارزشیابی را فعالیتی «تصمیم‌مدار»^{۲۸} می‌نامند. در ارزشیابی از آزمون و اندازه‌گیری استفاده می‌شود. این مفهوم با مفاهیم سنجش، نظارت، پاسخگویی و پژوهش شباهت‌ها و تفاوت‌هایی دارد. توجه به این تفاوت‌ها در نوشتن و صحبت کردن، برای انتقال و درک بهتر پیام ضروری است.

فهرست منابع

- Alkin, C. Marvin. (2011). *Evaluation essentials from A to Z*. New York. The Guilford Press.
- Cronbach, I.J. (1963). *Course improvement through evaluation*. Teacher College Record 64: 672-683.
- Fitzpatrick, L.J., Sanders, R.J., & Worthen, R.B. (third Ed, 2004). *Program evaluation: Alternative approaches and practical guidelines*. Pearson Education, Inc.
- Joint Committee on Standards for Educational Evaluation. (1944). *The program evaluation standards*. Thousand Oaks, CA: Crowing Press.
- Mathison, S. (Ed). (2005). *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Publications, Inc.
- Mertens, M. D. (2010). *Research and evaluation in education and psychology*. (3rd edition). Sage Publications, Inc.
- Popham, W. (1975). *Educational evaluation*. Englewood Cliffs, NJ: Prentice Hall.
- Scheerens, J., Glas, C., & Thomas, M. S. (2003). *Educational evaluation, Assessment and Monitoring*. Taylor & Francis e-Library.

14. accuracy
15. utility
16. feasibility
17. propriety
18. decision-oriented

Shahidur R.Khandker, Gayatri B. Koolwal, Hussain A. Samad. (2010). *Handbook on impact evaluation: Quantitative methods and practices*. The World Bank. Washington DC.

Smith, L.N, & Brandon, R.P.(2008) . *Fundamental issues in evaluation*. The Guilford Prees. New York.

Stenhouse , Lawrence. (1975). *An introduction to curriculum research and development*. Heinemann Educational Books Ltd, London.

Stufflebeam,D.L., & Shinkfield,A.J. (1985). *Systematic evaluation*. Boston, MA: Kluwer- Nijhoff.

Tyler, R. W. (1950). *Basic principles of curriculum and instruction*. Chicago, Ill.: University of Chicago Press.

Wholey, S. J, Hatry,P. H, & Newcomer, E.N. (2010). *Planning and designing useful evaluations*. In Wholey, S. J, Hatry,P. H, & Newcomer, E.N (eds.), *Handbook of practical program evaluation*. John Wiley & Sons.

Wolf, R.M. (1984). *Evaluation in education: foundations for competency assessment and program review*. Praeger Publication, New York.