

آشنایی با الگوی ارزشیابی CIPP^۱

معرفی مقاله

تألیف: دکتر علیرضا کیامنش

نظرات و دیدگاهها پیرامون ارزشیابی در سه دهه اخیر همواره در حال تکوین و تکامل بوده است. در نتیجه عمل ارزشیابی از پدیده‌های آموزشی به صورتهای مختلف توصیف و برای اجرای آن نیز الگوهای مختلفی ارائه گردیده است. گروهی هدف ارزشیابی را بررسی میزان تحقق یافتن هدفها قلمداد نموده و گروه دیگر به خصوصیات درونی و عوامل تشکیل دهنده برنامه آموزشی توجه نموده‌اند. گروهی هدف ارزشیابی را بررسی اثرات حاصل از اجرای برنامه آموزشی دانسته و سرانجام گروه دیگر هدف ارزشیابی را فراهم آوردن اطلاعات لازم به منظور تسهیل در امر تصمیم‌گیری قلمداد نموده‌اند. و الگوهای تسهیل در تصمیم‌گیری را ارائه نموده‌اند. از میان الگوهای تسهیل در تصمیم‌گیری، الگوی سیپ (CIPP) که توسط استافل‌بیم و همکارانش تدوین و ارائه شده است، در این مقاله معرفی شده است.

در این الگو ارزشیابی عبارتست از: فرایند تعیین کردن، بدست آوردن، و فراهم ساختن اطلاعات توصیفی و قضائتی در مورد ارزش و مطلوبیت هدفها، طرح، اجراء و نتایج به منظور هدایت تصمیم‌گیری، خدمت به نیازهای پاسخگویی و درک بیشتر از پدیده‌های مورد بررسی.

این الگو تصمیمهای آموزشی را به چهار گروه:
۱- برنامه‌ریزی، ۲- ساختن، ۳- اجرایی، و ۴- دوباره‌سازی تقسیم می‌کند، و برای هر یک نوعی ارزشیابی زیر عنوانین: ارزشیابی

زمینه؛ ارزشیابی درون داد؛ ارزشیابی فراگرد؛ و بالاخره ارزشیابی
برون‌داد معرفی می‌کند.

این مقاله توسط آقای دکتر علیرضا کیامنش عضو هیئت علمی
دانشگاه تربیت معلم برای فصلنامه تهیه گردیده است. با ابراز تشکر و
قدردانی از توجه و علاقه ایشان نسبت به فصلنامه، امید است بتوانیم
از الگوی ارائه شده در ارزشیابیهای آموزشی و درسی بهره‌مند
گردیم.

«فصلنامه»



پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مقدمه:

در سه دهه اخیر جهت ارزشیابی آموزشی چندین الگوی مختلف توسط صاحب نظران تدوین و ارائه گردیده است. ریچارد ولف،^۲ در کتاب ارزشیابی آموزشی (1979) به بررسی تفاوت‌های موجود بین این مدلها پرداخته و در مجموع چهار تفاوت اساسی را بشرح زیر بیان نموده است:

(۱) تفاوت در هدفهای اصلی ارزشیابی: مثلاً: اصلاح برنامه‌های آموزشی (کرانباخ، 1963)،^۳ یا تعیین میزان موفقیت برنامه آموزشی در کسب هدفهای آموزشی مورد نظر (تایلر 1942)^۴

(۲) تفاوت در زمینه‌های آموزشی و تجربی صاحب نظران. مثلاً: آموزش و تجربه در حوزه برنامه‌ریزی درسی، اندازه‌گیری و سنجش، روشهای پژوهش، روان‌شناسی تربیتی و...

(۳) تفاوت در ماهیت موضوع درس یا مطلبی که مرکز توجه کوششهای ارزشیابی است. مثلاً: درس هنر یا درس ریاضی.

(۴) تفاوت در تأکیدی که بر مدل روی بخشی خاصی از فرایند ارزشیابی مبذول می‌دارد. مثلاً: تأکید بر مدیریت و تصمیم‌گیری، یا تأکید بر سنجش دقیق رفتار.

شرح مفصل این الگوها در کتاب‌های مختلف ارزشیابی منجمله پاپ‌هام (1975)^۵، و استاقیل بیم و شینک فیلد (1986)^۶ ارائه گردیده است.

از میان این الگوها تقریباً از سه الگو می‌توان نام برد که موضوع مورد توجه آنها ارائه خدمات لازم به تصمیم‌گیرندگان و مسئولین آموزش و پرورش جهت تصمیم‌گیری در مورد فعالیت‌های آموزشی است. این الگوها عبارتند از: الگوی ارائه شده توسط استاقیل بیم و دیگران^۷ (1971) که به آن الگوی ارزشیابی سیپ (CIPP) گفته می‌شود. الگوی ارائه شده توسط مرکز مطالعات ارزشیابی دانشگاه کالیفرنیا - لوس‌آنجلس یا الگوی ارزشیابی CSE^۸ (1974)، و الگوی تفاوت یا اختلاف^۹ که توسط پروفیسور پرووس^{۱۰} (1971) ارائه گردیده است. به این سه الگو در مجموع الگوهای تسهیل در تصمیم‌گیری^{۱۱} گفته می‌شود.

در الگوهای تسهیل در تصمیم‌گیری، ارزشیاب کمتر به ارزشیابی مطلوبیت^{۱۲} و ارزش^{۱۳}

پدیده‌های آموزشی توجه دارد. سعی ارزشیابی در این الگوها آنست که اطلاعات لازم و مناسب را فراهم آورده و آنها را جهت تعیین مطلوبیت و ارزش فعالیت‌های آموزشی در اختیار افسراد تصمیم گیرنده قرار دهد.

الگوی ارزشیابی CIPP

در سال ۱۹۶۵ میلادی دولت امریکا یک هزینه چندین میلیون دلاری جهت بهبود آموزش دانش‌آموزان محروم و همچنین بالا بردن سطح کلی سیستم آموزش در مدارس ابتدایی و متوسطه اختصاص داد. این برنامه^{۱۲} (ESEA) مسئولین آموزش و پرورش را موظف می‌ساخت که پروژه‌های آموزشی را مورد ارزشیابی قرار دهند. مناطق آموزش و پرورش فقط به شرطی می‌توانستند در بودجه اختصاص داده شده سهم بگیرند که همراه هر یک از برنامه‌های پیشنهادی خود طرح ارزشیابی از آن برنامه را نیز ارائه دهند. به علت عدم تجربه و توانایی مناطق آموزش و پرورش برای طرح‌ریزی و اجرای ارزشیابی، از دانشگاه‌ها و سازمانهای پژوهشی برای این منظور کمک گرفته شد. در دانشگاه ایالت اُهایو^{۱۵} نیز برای کمک به ارزشیابی برنامه‌های ESEA مرکزی بوجود آمد و استافل‌بیم به سرپرستی این مرکز انتخاب گردید.

استافل‌بیم و همکاران وی ابتدا با استفاده از الگوی ارزشیابی تحقق یافتن هدفها^{۱۴} که توسط تایلر ارائه گردیده بود، سعی نمودند میزان موفقیت برنامه‌های اجرا شده را در کسب هدفهای تعیین شده مشخص نمایند. برای این منظور مجبور بودند که هدفهای هر یک از برنامه‌های مورد مطالعه را مشخص نمایند، ابزارهای مناسب جهت اندازه‌گیری عملکرد فراگیران را انتخاب و یا تهیه نمایند، و پس از انجام آموزش (اجرای عمل آزمایشی) با استفاده از ابزارهای تعیین شده عملکرد فراگیران را مورد اندازه‌گیری قرار داده و نتایج را با هدفهای تعیین شده مورد مقایسه قرار دهند. ولی بزودی به غیر واقعی بودن این فرض که «مربیان قادر هستند رفتارهایی را که فراگیران در اثر شرکت در برنامه کسب خواهند نمود را معین نمایند» پی بردند. از طرف دیگر، استفاده از روش تایلر باعث می‌گردید که نتایج ارزشیابی در پایان اجرای پروژه فراهم گردد و از آنها برای تشخیص مشکلات برنامه و رفع آن مشکلات آموزشی توان استفاده نمود. همچنین از طریق مشاهده مستقیم فراگرد برنامه‌های ESEA این نتیجه حاصل گردید که بین آنچه که در عمل به اجرا در می‌آید و آنچه که در طرح برنامه‌ها پیش‌بینی شده تفاوت اساسی وجود داشته و چگونگی اجرای یک برنامه در مدارس و حتی کلاسهای مختلف با هم یکسان نمی‌باشد.

بر اساس تجارب بدست آمده از ارزشیابی در مدارس و تحلیل منطقی شیوه‌های کلاسیک ارزشیابی، استافل‌بیم و همکاران وی در کمیته ملی مطالعات ارزشیابی (PDK, 1969)^{۱۷} به این نتیجه رسیدند که ارزشیابی نباید با فعالیت‌هایی نظیر اندازه‌گیری^{۱۸}، قضاوت حرفه‌ای^{۱۹}، طرح آزمایشی^{۲۰} و یا تطابق نتایج و هدفها (تایلرین) یکسان قلمداد گردد. اگر چه هر یک از شیوه‌های بالا دارای امتیازاتی هستند ولی هیچکدام برای ارزشیابی پروژه‌های ESEA کافی نبودند. در نتیجه ضرورت تدوین یک مفهوم بهتر از ارزشیابی به منظور غلبه بر ضعفهای موجود مطرح گردید. در جدول شماره یک امتیازات و محدودیت‌های هر کدام از چهار روش بالا به نقل از کتاب استافل‌بیم و شینک فیلد

جدول شماره ۱: امتیازات و محدودیت‌های شیوه‌های کلاسیک ارزشیابی

شبهه	امتیازات	محدودیت‌ها
۱ - اندازه‌گیری (هتچار مدار)	<ul style="list-style-type: none"> - براساس تئوری روان‌شناسی - استفاده از تکنیک‌های استاندارد کردن - اطمینان به اعتبار و پایایی نتایج - در دسترس بودن آزمون‌های منتشر شده و خدمات نمره‌گذاری - معیارها منشأ گرفته از هتچارها 	<ul style="list-style-type: none"> - تنها به ابزارهای موجود توجه می‌کند. - عدم انعطاف‌پذیری به علت زمان و هزینه لازم جهت تهیه ابزارهای جدید - از دست دادن اعتبار محتوی - در آراء پایایی نمرات تفاوت فردی - تأکید بر توانایی‌ها و دانشی که از طریق آزمون‌های قلم - کاغذی به‌سہولت اندازه‌گیری می‌شوند. - استفاده از هتچار برای بررسی همه فراگران
۲ - قضاوت حرفه‌ای	<ul style="list-style-type: none"> - اجرای ساده - امکان بررسی تمام متغیرها - استفاده از تجربه و تخصص - در زمان انتظار برای تحلیل داده‌ها وقت از بین نمی‌رود 	<ul style="list-style-type: none"> - صرفاً حاصل قضاوت شخصی است - پایایی و عینیت نتایج زیر سؤال است - داده‌ها و معیارها مبهم هستند - امکان بررسی اعتبار نتایج وجود ندارد - تمییز یافته‌ها خیلی مشکل است
۳ - طرح آزمایشی	<ul style="list-style-type: none"> - بالا بودن ارزش علمی - امکان بررسی رابطه علت و معلول - بالا بودن پایایی، اعتبار و عینیت یافته‌ها - یافته‌ها از طریق استنباط آماری مورد نائید قرار می‌گیرند. 	<ul style="list-style-type: none"> - شرایط کنترل متغیرها، در محیط‌های آموزشی قابل اجرا نیست و با اهداف ارزشیابی نیز مغایرت دارند. - مداخله در اجرای عادی پروژہ تحت مطالعه - نیاز به فرض‌های غیر قابل اجرا - محدود بودن مطالعه به چند متغیر - نتایج فقط در پایان آزمایش فراهم می‌شوند
۴ - تطابق بازده‌ها و اهداف (شیوۀ تایلرین)	<ul style="list-style-type: none"> - توافق کامل با طراحی آموزشی - جمع‌آوری اطلاعات از فراگیران و برنامه - امکان باز خورد دوره‌ای - مشخص بودن معیارهای موفقیت 	<ul style="list-style-type: none"> - به ارزشیابی یک نقش تکنیکی می‌دهد. - توجه به هدفها منحصر می‌گردد - معرفی رفتار بعنوان معیار نهایی هر فعالیت آموزشی - عدم ارزشیابی هدفها - توجه به ارزشیابی بعنوان یک فراگرد پایانی

کمیته PPK به پنج مسئله در مفاهیم کلاسیک ارزشیابی اشاره می‌کند که برای بسط مفهوم ارزشیابی باید مورد اصلاح قرار گیرند. این مسائل عبارتند از:

(۱) تعریف^{۲۵}: تعاریف موجود که ارزشیابی را با اندازه‌گیری، طرح آزمایشی و یا قضاوت حرفه‌ای یکسان می‌دانند، نارسا هستند. این تعاریف خیلی مبهم هستند و با توجه به فرض‌های مطرح شده در تعریف از اشکال برخوردار بوده و از نظر متدولوژی نیز ضعیف هستند.

(۲) تصمیم‌گیری^{۲۶}: ارزشیابی با تصمیم‌گیری ارتباط دارد و باید تلاش‌های مربوط به تغییر و نوآوری را هدایت نماید، ولی نوع تصمیمی که باید برای این منظور بخدمت گرفته شود مشخص نیست. جهت هدایت ارزشیابی در تعیین مسائل تصمیم‌گیری و جمع‌آوری اطلاعات مورد نیاز، هیچ فرایندی موجود نیست.

(۳) ارزش‌ها^{۲۷}: در حالیکه باید براساس معیارهای منعکس‌کننده سیستم‌های ارزشی از فعالیت‌های آموزشی اندازه‌گیری بعمل آید ولی جهت‌گیری بعضی از فعالیت‌های ارزشیابی انجام شده بخطا ارزش — آزاد بوده است. ارزشیاب‌ها فاقد ابزارهای لازم جهت تشخیص نوع سیستم‌های ارزشی هستند تا بتوانند از آنها بعنوان پایه‌ای جهت فعالیت‌های جمع‌آوری اطلاعات استفاده نمایند.

(۴) سطوح^{۲۸}: اطلاعات فراهم شده برای افراد مختلف در سطوح مختلف یک سازمان باید متفاوت باشد. یک طرح ارزشیابی که بمنظور پاسخگویی به نیازهای مثلاً کادر آموزشی یک مدرسه به اجرا در می‌آید، احتمالاً نمی‌تواند نیازهای مسئولین آموزش و پرورش منطقه را نیز پاسخگو باشد. در نتیجه ارزشیابی به فرایندی نیاز دارد که ارزشیاب‌ها را در شناخت گروه‌های مخاطب، سنجش خصوصیات و اطلاعات مورد نیاز گروه‌ها، طرح‌ریزی جهت جمع‌آوری اطلاعات و گزارش فعالیت‌ها بمنظور برآورده کردن نیازهای گروه‌های مخاطب یاری دهد.

(۵) طرح پژوهش^{۲۹}: پژوهش و ارزشیابی از نظر هدف با هم تفاوت دارند و در نتیجه از نظر شیوه‌های رسیدن به هدف نیز باید با هم متفاوت باشند. هدف فعالیت‌های پژوهشی بیشتر کشف حقایق و تعمیم نتایج است، در حالیکه هدف فعالیت‌های ارزشیابی فراهم آوردن اطلاعات و راهنمایی‌های لازم جهت تصمیم‌گیری مناسب و منطقی در مورد پدیده‌های آموزشی است.

استافل بیم، هر یک از مسائل پنجگانه بالا را مورد تجزیه و تحلیل قرار داد و مدل CIPP را جهت پاسخگویی به این مسائل ارائه نمود. کامل‌ترین تعریف برای ارزشیابی در مدل CIPP بشرح زیر است:

«ارزشیابی عبارتست از فرایند تعیین کردن^{۳۰}، بدست آوردن^{۳۱}، و فراهم ساختن^{۳۲} اطلاعات توصیفی^{۳۳} و قضای^{۳۴} در مورد ارزش و مطلوبیت هدفها، طرح، اجرا و نتایج بمنظور هدایت تصمیم‌گیری، خدمت به نیازهای پاسخگویی^{۳۵} و درک بیشتر از پدیده‌های مورد بررسی» (استافل بیم و شینک‌فیلد، ۱۹۸۶، صفحه ۱۵۹).

در این تعریف برای ارزشیابی سه هدف تعیین گردیده است: هدایت تصمیم‌گیری؛ فراهم کردن اطلاعات برای پاسخگویی و درک بیشتر از پدیده‌های مورد بررسی. استافل بیم همانند اسکریون^{۳۶} برای ارزشیابی دو نقش^{۳۷} قائل می‌گردد: نقش تصمیم‌گیری یا نقش تکوینی^{۳۸} و نقش پاسخگویی یا نقش مجموعی^{۳۹} (پایانی).

این تعریف همچنین ارزشیابی را یک فعالیت مستمر می‌داند که شامل سه مرحله تعیین کردن، بدست آوردن و فراهم ساختن اطلاعات است. مرحله تعیین کردن شامل فعالیت‌هایی نظیر مشخص کردن، تعریف کردن و روشن ساختن می‌باشد که تصمیم‌گیرندگان به آنها نیاز دارند. مرحله بدست آوردن شامل جمع‌آوری، سازمان دادن و تجزیه و تحلیل است که با استفاده از روشهای تکنیکی نظیر اندازه‌گیری و آمار صورت می‌گیرد. مرحله فراهم ساختن شامل ترکیب اطلاعات مختلف بدست آمده بصورتی که در اهداف ارزشیابی مورد استفاده قرار گیرند از مراحل سه‌گانه، مراحل تعیین کردن و فراهم کردن اطلاعات با همکاری مشترک ارزشیاب و تصمیم‌گیرنده انجام می‌پذیرد؛ ولی مرحله بدست آوردن اطلاعات چون یک فعالیت کاملاً تکنیکی است توسط ارزشیاب یا گروه ارزشیابی انجام می‌گیرد.

اطلاعات توصیفی و قضایوی هر دو برای سنجش و کمک به بهبود برنامه مورد نظر مناسب هستند. با توجه به تعریف ارائه شده، جنبه‌های کلیدی برنامه که باید مورد سنجش قرار گیرند عبارتند از: تعیین هدفهای برنامه؛ تهیه طرح مناسب برای رسیدن به اهداف؛ اجرای طرح تهیه شده و بررسی محصول یا بازده برنامه که به ترتیب توسط ارزشیابی زمینه، ارزشیابی درون داد، ارزشیابی فراگرد و ارزشیابی برون داد مورد ارزشیابی قرار می‌گیرند. معیارهای اساسی برای هر چهار نوع ارزشیابی عبارتند از: ارزش و مطلوبیت.

الگوی CIPP تصمیم‌های آموزشی را به چهار گروه تقسیم می‌کند:

- ۱ - تصمیم‌های برنامه‌ریزی^{۲۰} جهت تعیین اهداف برنامه.
 - ۲ - تصمیم‌های ساختی^{۲۱} جهت طراحی شیوه‌های آموزش مناسب بمنظور رسیدن به اهداف برنامه.
 - ۳ - تصمیم‌های اجرایی^{۲۲} جهت استفاده، هدایت و بهبود شیوه‌های آموزشی.
 - ۴ - تصمیم‌های دوباره‌سازی^{۲۳} بمنظور قضاوت و عکس‌العمل در مورد نتایج بدست آمده.
- و برای هر یک از تصمیم‌های بالا یک نوع ارزشیابی معرفی می‌کند.

پرتال جامع علوم انسانی

ارزشیابی زمینه

ارزشیابی زمینه با استفاده از روش‌های توصیفی و مقایسه‌ای تلاش می‌کند که مسائل و مشکلات آموزشی، جنبه‌های مثبت و منفی (قوت و ضعف) برنامه‌ها، نیازهای برآورده نشده و امکانات آموزشی بخدمت گرفته نشده (بطور کلی محیط آموزش بمعنی عام کلمه) را از طریق مصاحبه با افراد ذیصلاح و دست‌اندرکار مسائل آموزشی؛ مشاهده مستقیم جریان آموزش؛ اجرای آزمونهای تشخیصی جهت تعیین میزان عملکرد فراگیران و ... مورد شناسایی قرار دهد. بررسی دقیق و همه‌جانبه محیط آموزشی به شناخت وضعیت اهداف موجود آموزشی و برنامه‌های آموزشی منجر می‌گردد. همچنین از طریق مصاحبه با افراد ذیصلاح و دست‌اندرکار مسائل آموزشی و بررسی نیازهای جامعه و فراگیران، اهداف آموزشی که برای سیستم آموزش مطلوب و ضروری بنظر می‌رسند مشخص می‌گردند. اهداف وضعیت موجود با اهداف وضعیت مطلوب مورد مقایسه قرار گرفته و تفاوت‌های بین دو وضعیت معین می‌گردد. به هر هدفی که در وضعیت مطلوب وجود دارد ولی در وضعیت موجود جای آن خالی است،

یک نیاز آموزشی^{۴۴} گفته می‌شود. از میان نیازهای تعیین شده با توجه به شرایط و امکانات، نیازهای مهمتر و ضروری‌تر انتخاب شده تا در جهت برآورده شدن آنها برنامه‌ریزی صورت گیرد. به هر نیاز آموزشی که بمنظور برآورده شدن در برنامه آموزشی گنجانیده می‌شود یک مسئله آموزشی^{۴۵} گفته می‌شود. بدین ترتیب حاصل ارزشیابی زمینه تعیین آن دسته از اهداف آموزشی یا مسائل آموزشی است که برای نیل به آنها باید یک برنامه آموزش طراحی و تدوین گردد. ضمناً ارزشیابی زمینه اساسی جهت قضاوت در مورد برون دادهای برنامه فراهم می‌سازد.

ارزشیابی زمینه، با فراهم آوردن اطلاعاتی در رابطه با اهداف انتخاب شده و انتخاب نشده و ارتباط این دو نوع هدف با فرصت‌ها، نیازها و مسائل آموزشی به امر پاسخگویی نیز کمک می‌کند.

ارزشیابی درون داد

هدف اصلی ارزشیابی درون داد کمک به تدوین برنامه‌ای است که برای ایجاد تغییرات آموزشی و حصول به اهداف تعیین شده، در مرحله ارزشیابی زمینه طراحی می‌گردد. پس از مطالعه و شناخت محیطی که برنامه جدید آموزشی در آن به اجرا درمی‌آید، بسایند تمام عوامل تشکیل دهنده برنامه‌های آموزشی موجود بمنظور تصمیم‌گیری در مورد چگونگی طراحی و ساخت برنامه تازه ارزشیابی شود. امکانات مالی و انسانی، خط‌مشی‌ها، راهبردهای آموزشی، موانع و محدودیت‌های موجود در سیستم آموزش مورد بررسی قرار گرفته و اطلاعات لازم در زمینه نوع و میزان استفاده از سیستم آموزشی جهت رسیدن به اهداف تعیین شده فراهم می‌گردد. بطور کلی خصوصیات کلیه برنامه‌های آموزشی موجود که به نحوی می‌توانند به پیشبرد اهداف تعیین شده کمک نمایند، مورد شناسایی و ارزشیابی قرار می‌گیرد. تصمیم‌گیرندگان با استفاده از اطلاعات فراهم شده در این مرحله و با توجه به اهداف تعیین شده در مرحله قبلی (ارزشیابی زمینه) به انتخاب، انطباق و یا در صورت ضرورت طراحی یک برنامه جدید اقدام می‌کنند.

ارزشیابی درون داد، از طریق تعیین و تعریف واضح و روشن فعالیت‌های لازم جهت تدوین برنامه، بدست آوردن اطلاعات مناسب در زمینه آن فعالیت‌ها، و فراهم کردن اطلاعات بصورت یک گزارش کامل و ارائه به تصمیم‌گیرندگان به آنان کمک می‌کند تا در مورد نوع و چگونگی استفاده از منابع مختلف آموزشی برای رسیدن به هدفهای تعیین شده، اتخاذ تصمیم نمایند. مثلاً اگر هدف یک برنامه آموزشی بشرح زیر بیان شده باشد: «بالا بردن توانایی ریاضی دانش‌آموزان کلاس پنجم دبستان» وظیفه ارزشیابی درون داد آن خواهد بود که عوامل مختلف آموزشی از قبیل منابع و راهبردهای آموزشی، بودجه و زمان لازم جهت اجرای برنامه، موانع اجرایی و آموزشی، کمبودها و همچنین چگونگی بکارگیری هر یک از عوامل بالا را معین ساخته و ضمن بدست آوردن اطلاعات کافی در مورد هر یک، آنها را به صورت منسجم در اختیار تصمیم‌گیرندگان قرار دهد. تصمیم‌گیرندگان با استفاده از اطلاعات فراهم شده از میان برنامه‌های موجود برنامه‌ای را که قادر باشد اهداف تعیین شده را برآورده نماید جهت اجرا انتخاب می‌کند. در صورتیکه برنامه‌های موجود فاقد چنین توانایی باشند، تصمیم‌گیرندگان با ایجاد تغییرات و اصلاحات لازم در یکی از مناسب‌ترین برنامه‌های موجود آن را جهت اجرا آماده می‌کند و یا به طراحی و تدوین یک برنامه کاملاً جدید مبادرت می‌ورزد. در این

برنامه کلیه عوامل از قبیل راهبردهای آموزشی، منابع و امکانات کمک آموزشی، زمان اجرای هر یک از قسمت‌ها و فعالیت‌های برنامه، منابع مالی و انسانی پیش‌بینی و مشخص می‌گردد.

ارزشیابی درون داد، علاوه بر کمک به جریان تصمیم‌گیری، با فراهم آوردن اطلاعات دقیق در مورد علت انتخاب و یا طراحی برنامه جدید و رد دیگر برنامه‌های آموزشی و همچنین نشان دادن ضعف‌ها و امتیازات هر کدام از برنامه‌ها، به امر پاسخگویی نیز کمک می‌کند.

ارزشیابی فراگرد

طراحی برنامه، باید چگونگی اجرای برنامه مورد ارزشیابی قرار گیرد. در این مرحله باید مشخص گردد که آیا برنامه طبق آنچه که پیش‌بینی شده است به اجرا درمی‌آید یا خیر؟ نقش ارزشیاب فراگرد آنست که چگونگی اجرای هر یک از بخش‌های برنامه را با آنچه که در برنامه پیش‌بینی شده است مقایسه نموده و کمبودها و نارسایی‌های احتمالی را مشخص نماید. مثلاً آیا کلیه امکانات کمک آموزشی پیش‌بینی شده در برنامه، در زمان تعیین شده فراهم و بصورت کامل به کار گرفته شده است؟ آیا بودجه پیش‌بینی شده برای اجرای برنامه، طبق برنامه فراهم و به مصرف رسیده است؟ آیا معلمین برای انجام فعالیت‌های تعیین شده، آموزش‌های لازم را فرا گرفته‌اند؟ آیا شیوه‌های آموزشی در نظر گرفته شده برای تدریس، توسط معلمین در کلاس بکار گرفته شده است؟ آیا مراحل مختلف برنامه طبق جدول زمان بندی شده به اجرا درآمده است؟ عبارت روشن‌تر، ارزشیاب فراگرد تلاش می‌کند که برای دو سوال عمده زیر اطلاعات لازم را جمع‌آوری و فراهم سازد:

۱) آیا بین زمان شروع و توالی قسمت‌های مختلف برنامه در حال اجرا با آنچه که در طرح پیش‌بینی شده مطابقت وجود دارد؟

۲) آیا قسمت‌های مختلف برنامه بهمان صورتی که در طرح پیش‌بینی شده است به اجرا درمی‌آید؟

این اطلاعات معمولاً از طریق مشاهده مستقیم و مداوم جریان اجرای برنامه و مصاحبه با نیروهای انسانی مجری برنامه جمع‌آوری می‌گردند. ارزشیاب می‌تواند از همکاری مدیران، معلمین، راهنمایان تعلیماتی، مشاورین راهنمایی، دانش‌آموزان و حتی اولیاء دانش‌آموزان در این زمینه بهره‌مند گردد.

استفاده از یادداشت‌ها و دفترچه خاطرات افراد مجری برنامه، طرح درس‌ها، دفتر حضور و غیاب کلاس، صورت جلسات شورای معلمان و... می‌تواند به نیازهای اطلاعاتی در این حوزه کمک نماید. شکست در ارزشیابی و پاسخگویی به سوالات بالا، فعالیت‌های ارزشیابی را با انتقاد جدی روبرو می‌سازد.

اهمیت اصلی ارزشیابی فراگرد در فراهم آوردن بازخوردهای مناسب برای مجریان و مسئولین برنامه است تا برنامه را طبق آنچه که در طرح پیش‌بینی شده به اجرا در آورده و بر ضعف‌ها و نارسایی‌های طرح و مشکلات اجرایی آن در جریان اجرا غلبه نمایند. از طریق نظارت مداوم و دقیق بر اجرای برنامه و در نتیجه شناخت موفقیت‌ها و عدم موفقیت‌های برنامه، شرایط لازم جهت تصمیم‌گیریهای معقولانه در مورد اجرای برنامه فراهم می‌گردد.

از اطلاعات فراهم شده در جریان ارزشیابی فراگرد علاوه بر تصمیم‌گیری می‌توان در زمینه پاسخگویی نیز استفاده نمود. معمولاً سازمان‌های برنامه‌ریزی، سیاسی و آموزشی مملکت و حتی مردم عادی و اولیاء دانش‌آموزان خواستار آن هستند که از کم و کیف اجرای طرح‌های آموزشی مطلع گردند. از نتایج فراهم شده در مرحله ارزشیابی فراگرد برای پاسخگویی به این خواست‌ها می‌توان استفاده نمود. ارزشیابی فراگرد همچنین می‌تواند به افرادی که قصد دارند برنامه مورد ارزشیابی را در موقعیت دیگری مورد استفاده قرار دهند، کمک نماید. علاوه بر آن نتایج حاصل از ارزشیابی فراگرد، یک منبع اطلاعاتی بسیار مهم برای تفسیر نتایج ارزشیابی برون داد بشمار می‌آید، زیرا برای بررسی نتایج بدست آمده از یک برنامه باید چگونگی به اجرا در آمدن آن برنامه مشخص شده باشد.

ارزشیابی برون داد

هدف از ارزشیابی برون داد اندازه‌گیری، تفسیر و قضاوت در مورد نتایج حاصل از برنامه است. سنجش و ارزشیابی برون دادهای برنامه معمولاً در سه مرحله زمانی متفاوت انجام می‌گیرد: ۱) در جریان اجرای برنامه و در پایان هر یک از مراحل برنامه، بمنظور پی‌بردن به چگونگی اجرای برنامه و موفقیت آن در نيل به هدفهای پیش‌بینی شده (ارزشیابی تکوینی)؛ ۲) در پایان اجرای کامل برنامه، بمنظور پی‌بردن به میزان موفقیت کلی برنامه در رسیدن به هدفهای تعیین شده در مرحله ارزشیابی زمینه (ارزشیابی مجموعی یا پایانی)؛ و ۳) مدت زمانی پس از اجرای کامل برنامه (مطالعات پی‌گیری)^{۴۴}، بمنظور پی‌بردن به میزان پایداری و اثربخشی نتایج حاصل از برنامه^{۴۷}. سعی ارزشیاب برون داد بر آنست که میزان موفقیت برنامه در برآورده کردن نیاز گروههایی که برنامه جهت خدمت به آنان طراحی و به اجرا گذاشته شده است را مشخص سازد. در جریان این تلاش‌ها، علاوه بر ارزشیابی برون دادهای پیش‌بینی شده^{۴۸} در برنامه (ارزشیابی هدفهای تعیین شده)، اثرات جانبی مثبت و یا منفی برنامه، یا به عبارت دیگر، اثرات پیش‌بینی نشده^{۵۰} در برنامه نیز مورد سنجش و قضاوت قرار می‌گیرد. کرانباخ (1968) معتقد است که:

«یک ارزشیابی ایده‌آل تنها به انتخاب و اندازه‌گیری بازده‌هایی که مورد توجه برنامه هستند نمی‌پردازد، بلکه تمام شایستگی‌هایی را که منطقاً از برنامه انتظار می‌رود اندازه‌گیری می‌کند. اگر قصد فقط بررسی میزان موفقیت برنامه در رسیدن به هدفهای تعیین شده باشد، آزمون باید متناسب با محتوی درس باشد. ولی اگر قصد پی‌بردن به میزان موفقیت برنامه در برآورده کردن نیازهای جامعه باشد، در آن صورت باید تمام نتایجی که رسیدن به آنها با ارزش است مورد اندازه‌گیری قرار گیرد». (به نقل از ولف 1979، صفحات ۲۹ - ۲۸).

دست اندرکاران آموزش و پرورش به این نتیجه رسیده‌اند که برنامه‌های آموزشی علاوه بر تأثیرات مورد انتظار، احتمالاً تأثیرات جانبی نیز بدنبال خواهند داشت، لذا ارزشیاب باید نسبت به آنچه که برای فراگیران در جریان شرکت در برنامه آموزشی به وقوع می‌پیوندد حساس بوده و آنها را بطور کامل مورد توجه و مطالعه قرار دهد.

در جریان ارزشیابی درون‌داد باید قضاوت‌های بعمل آمده توسط افراد و گروههای ذینفع در مورد اجرای برنامه و موفقیت‌های آن جمع‌آوری و مورد تحلیل قرار گیرد. استفاده از عقاید و

عکس‌العملهای افراد و گروه‌های ذینفع و در ارتباط با برنامه که از طریق پرسشنامه و مصاحبه جمع‌آوری می‌گردد می‌تواند به ارزشیابی برون‌دادهای برنامه و تصمیم‌گیری در مورد برنامه کمک نماید. مثلاً این احتمال وجود دارد که یک برنامه آموزشی با وجود موفقیت در برآورده کردن نیازهای مشخص شده و رسیدن به هدفهای تعیین شده، مخالفت و عکس‌العملهای منفی گروه‌های مختلف جامعه را به دنبال داشته باشد. عدم اطلاع از این مخالفت‌ها و عکس‌العملها و عدم تصمیم‌گیری صحیح، جهت برطرف کردن و یا کاهش آنها ممکن است به توقف و یا حذف یک برنامه موفق منجر گردد.

ارزشیابی برون‌دادهای برنامه از طریق سنجش عملکرد فراگیران شرکت‌کننده در برنامه در زمینه‌های مختلف آموزشی و پرورشی (شناختی^{۵۱}، عاطفی^{۵۲} و روانی-حرکتی^{۵۳}) صورت می‌گیرد. بمنظور سنجش و ارزشیابی اثرات برنامه آموزشی، معمولاً نتایج حاصل از آزمون‌های عملکرد فراگیران شرکت‌کننده، در برنامه آموزشی با معیارهایی نظیر: عملکرد آنان در پیش‌آزمون^{۵۴}، نیازهای از قبل تعیین شده، هنجارهای^{۵۵} انتخاب شده و یا عملکرد گروه گواه، مورد مقایسه قرار می‌گیرد. جهت سنجش عملکردهای فراگیران در زمینه‌های مورد نیاز، از آزمون‌های عینی بخصوص آزمون‌های عینی ملاک مطلق^{۵۶}، آزمون‌های عملی و مهارتی و یا مشاهده و مقیاس‌های درجه‌بندی^{۵۷} رفتار می‌توان استفاده نمود. همچنین از طریق انتخاب دقیق یک گروه نمونه تصادفی از افراد شرکت‌کننده در برنامه و انجام یک مطالعه موردی^{۵۸} و یا از طریق انجام یک مطالعه پیمایشی^{۵۹} می‌توان به اطلاعات لازم در زمینه تأثیرات همه‌جانبه برنامه دسترسی پیدا نمود.

کاربرد اساسی ارزشیابی برون‌داد آنست که ارزش برنامه به اجرا گذاشته شده را مشخص نموده و اطلاعات لازم جهت تصمیم‌گیری در مورد ادامه، توقف، تبدیل و تکرار برنامه را فراهم سازد. از نتایج حاصل از ارزشیابی برون‌داد جهت پاسخگویی در مورد اثرات برنامه و تهیه گزارش لازم‌تر می‌توان استفاده نمود. اگر نتایج ارزشیابی برون‌داد، موفقیت برنامه را مورد تأیید قرار دهد، زمینه لازم برای حمایت مادی و معنوی برنامه از طرف مقامات و دست‌اندرکاران مملکتی فراهم می‌گردد، بالعکس اعلام عدم موفقیت برنامه باعث می‌گردد که جریان سرمایه‌گذاری بیهوده متوقف گردد.

خلاصه

مدل CIPP از چهار نوع ارزشیابی (زمینه، درون‌داد، فراگرد و بیرون‌داد)، دو نقش (تصمیم‌گیری، و پاسخگویی) و سه مرحله (تعیین کردن، بدست آوردن و فراهم ساختن) تشکیل شده است. این مدل از دو معیار اساسی (ارزش و مطلوبیت) استفاده می‌کند. از چهار نوع ارزشیابی بالا می‌توان جهت تصمیم‌گیری، پاسخگویی و یا تصمیم‌گیری و پاسخگویی استفاده نمود.

ضمناً در هر یک از انواع ارزشیابی با هر نقشی که مورد استفاده قرار گیرد، مراحل سه‌گانه مورد بحث بکار گرفته می‌شود. ارزشیابی و تصمیم‌گیرندگان، نوع ارزشیابی را متناسب با نیاز برنامه مشخص می‌کنند. ممکن است تمام مراحل یک برنامه آموزشی به ارزشیابی نیاز داشته باشد، در نتیجه هر چهار نوع ارزشیابی در مورد برنامه به اجرا در می‌آید. گاه ممکن است فقط یک یا چند قسمت برنامه به ارزشیابی نیاز داشته باشد، در آنصورت فقط آن یک یا چند قسمت مورد ارزشیابی قرار می‌گیرد.

ارتباط چهار نوع ارزشیابی با دو نقش و سه مرحله بصورت خلاصه در جدول شماره دو به نقل از مقاله استافلیم در کتاب پاپهام (1974) ص ۱۲۳ در زیر ارائه می‌گردد.

جدول شماره ۲: چهارچوب طراحی مطالعات ارزشیابی مدل CIPP

انواع ارزشیابی

زمینه	درون‌داد	فراگرد	برون‌داد
نقش	نقش	نقش	نقش
تصمیم‌گیری	تصمیم‌گیری	تصمیم‌گیری	تصمیم‌گیری
پاسخگویی	پاسخگویی	پاسخگویی	پاسخگویی
تعیین کردن	چه نوع سؤالاتی مورد توجه قرار می‌گیرد؟		
بدست آوردن	چگونه اطلاعات موردنیاز بدست می‌آید؟		
فراهم کردن	چگونه اطلاعات بدست آمده گزارش خواهد شد؟		

مراحل ارزشیابی



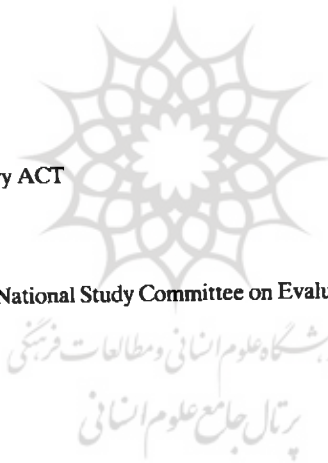
پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی



زیر نویسها:

۱) واژه سیپ (CIPP) از ترکیب حروف اول کلمات Context (زمینه)، Input (درون داد)، Process (فراگرد)، و Product (برون داد) تشکیل شده است.

- ۲) Richard, M, Wolf, Evaluation in Education
- ۳) Cronbach
- ۴) Tyler
- ۵) Popham
- ۶) Stufflebeam and Shinkfield
- ۷) Stufflebeam et al.
- ۸) Center for the study of Evaluation, university of California , Los Angeles.
- ۹) Discrepancy
- ۱۰) Provus
- ۱۱) Decision - Facilitation
- ۱۲) merit
- ۱۳) Worth
- ۱۴) Elementary and secondary ACT
- ۱۵) Ohio state university
- ۱۶) Goal - attainment
- ۱۷) Phi Delta Kappa (PDK) National Study Committee on Evaluation
- ۱۸) Measurement
- ۱۹) Professional judgment
- ۲۰) Experimental design
- ۲۱) Validity
- ۲۲) Reliability
- ۲۳) Content validity
- ۲۴) Norm
- ۲۵) Definition
- ۲۶) Decision making
- ۲۷) Values
- ۲۸) levels
- ۲۹) Research Design
- ۳۰) Delineating
- ۳۱) obtaining
- ۳۲) providing



- ۳۳) Descriptive
- ۳۴) Judgmental
- ۳۵) Accountability
- ۳۶) Scriven
- ۳۷) Role
- ۳۸) Formative
- ۳۹) Summative
- ۴۰) Planning
- ۴۱) Structuring
- ۴۲) Implementing
- ۴۳) Recycling
- ۴۴) Educational need
- ۴۵) Educational Problem
- ۴۶) Follow – up

(۴۷) در شماره یک و سه ارزشیابی چندین بار می‌تواند صورت گیرد.

- ۴۸) Intended Effect
- ۴۹) Side Effect
- ۵۰) Unintended Effect
- ۵۱) Cognitive
- ۵۲) Affective
- ۵۳) Psychomotor
- ۵۴) Pretest
- ۵۵) Norm
- ۵۶) Criterion – referenced test
- ۵۷) Rating Scales
- ۵۸) Case study
- ۵۹) Survey Study



پژوهشگاه علوم انسانی و مطالعات فرهنگی
رتال جامع علوم انسانی

منابع فارسی

- ۱ – بازرگان، عباس. مقدمه‌ای بر ارزیابی آموزش و الگوهای آن، همدان دانشگاه بوعلی، ۱۳۵۷
- ۲ – یولا، ه. س. ارزیابی و کاربرد آن در سوادآموزی تایمی، ترجمه دکتر بازرگان، عباس، تهران، مرکز نشر دانشگاهی، ۱۳۶۲

منابع انگلیسی

- 1 – Popham, W. J. **Evaluation in Education**, American Publishing Corporation, 1974.
- 2 – Popham, W. J. **Educational Evaluation**, Englewood Cliffs, N. J. prentice Hall, 1975
- 3 – Stufflebeam and skinkfield, **Systematic Evaluation**, Kluwer: Nijhoff publishing, 1986
- 4 – Wolf, R. M. **Evaluation in Education**, Foundation of competency Assessment and program Review, praeger publishers, 1979