

مجله علوم اجتماعی و انسانی دانشگاه شیراز
دوره هیجدهم، شماره دوم، پاییز ۱۳۸۱ (پیاپی ۳۶)
(ویژه نامه علوم تربیتی)

مقایسه تطبیقی روند تحولات ارزشیابی آموزشی در غرب و ایران

دکتر نعمت‌الله موسی‌پور*
دانشگاه شهید باهنر کرمان

چکیده

در مقاله حاضر، هدف آن بوده است که با اتکا بر دیدگاهی تطبیقی، پهنی برای این پرسش فراهم شود که ارزشیابی آموزشی، طی دو قرن گذشته، چه وضعی داشته و چگونه تکوین یافته است. گرچه ترسیم چگونگی تکوین ارزشیابی آموزشی در ایران بر پایه تحلیل منابع موجود به عمل آمده است؛ اما نظریه مدوس، استافل بیم و اسگریون (۱۹۸۸) در شناخت وضع ارزشیابی آموزشی در غرب، مبنای کار قرار گرفته است. تحلیل تاریخی مدوس، استافل بیم و اسگریون (۱۹۸۸) از تحول ارزشیابی آموزشی در غرب، آنان را به این جمع‌بندی رسانده است که ارزشیابی آموزشی، شش دوره را تجربه کرده است و آنان این دوره‌ها را با عناوین دوره‌های دگرگونی (۱۹۰۰-۱۸۰۰)، کارایی و آزمون (۱۹۳۰-۱۹۰۰)، تایلری (۱۹۴۵-۱۹۳۰)، معصومیت (۱۹۵۷-۱۹۴۶)، بسط (۱۹۷۲-۱۹۵۸) و حرفه‌ای شدن (۱۹۸۳-۱۹۷۳) طبقه‌بندی کرده‌اند.

در ایران نیز ارزشیابی آموزشی در تعامل با سایر جوامع بشری، طی همین دوره‌ها، تحول یافته؛ اما با آنچه در غرب گذشته، همسانی ندارد. تحلیل مدارک موجود نشان می‌دهد که ارزشیابی آموزشی در ایران از ابتدای قرن نوزدهم میلادی (۱۱۷۹) تاکنون، هفت دوره را گذرانده است و می‌توان آنها را با عناوین دوره‌های بی‌نیازی (۱۲۷۹-۱۱۷۹)، آگاهی (۱۳۰۹-۱۲۸۰)، تلاش رسمی (۱۳۲۴-۱۳۱۰)، دشواری (۱۳۳۶-۱۳۲۵)، علاقمندی (۱۳۵۱-۱۳۳۷)، اقدام (۱۳۷۰-۱۳۵۲) و امیدواری (۱۳۷۱ تاکنون) معرفی کرد. شایان ذکر است که این مرحله‌بندی به صورت تطبیقی صورت گرفته است و شک نیست که در یک نگاه تحلیلی، گستره و نام دوره‌ها متفاوت خواهد شد.

واژه‌های کلیدی: ۱. ارزشیابی آموزشی ۲. ایران ۳. طبقه‌بندی ۴. مطالعه تطبیقی ۵. غرب ۶. تحول ارزشیابی

۱. مقدمه

استفاده از واژه ارزشیابی^۱ در بین دست‌اندرکاران حوزه‌های مختلف امور انسانی و به خصوص آموزش و پرورش و حتی عامه مردم، چنان گسترده است که گویی معنای آن برای همه بدیهی است؛ حال آن که نه تنها عامه مردم فهم درستی از معنای آن در ذهن ندارند؛ بلکه پس از حدود یک صد سال بحث‌های تخصصی، هنوز تعریف مشخصی از ارزشیابی در بین صاحب‌نظران نیز رواج نیافته است و برای کسب آگاهی در باره آن باید به ناچار به نام صاحب‌نظران

متوسل شد. به عبارت دیگر هنوز نمی‌توان از «معنا و چیستی» ارزشیابی پرسش کرد؛ بلکه باید ارزشیابی و معنای آن را تنها بر اساس دیدگاه اشخاص دنبال کرد.

پرداختن به تعریف‌های مختلف ارزشیابی که توسط صاحب‌نظران این حوزه ارایه شده‌اند (کرنباخ،^۲ ۱۹۶۳؛ پایهام،^۳ ۱۹۷۱؛ استیک،^۴ ۱۹۶۷؛ بولا،^۵ ۱۳۶۲؛ هاوس،^۵ ۱۹۸۸؛ کیامنش،^۶ ۱۳۷۳؛ هومن،^۷ ۱۳۷۵؛ سیف،^۸ ۱۳۷۵ و اردبیلی،^۹ ۱۳۷۶) نیازمند تدوین مقاله‌های مستقلی است. بر همین اساس، در اینجا با پرهیز از هر نوع بررسی و بحث در باره تعریف‌های ارایه شده، ارزشیابی آموزشی فعالیت نظامداری تصور می‌شود که ضمن فراهم نمودن ملاک‌هایی برای ارزش‌گذاری، به تهیه اطلاعاتی در باره پدیده یا موضوع مورد نظر می‌پردازد و با مقایسه آن دو، از ارزش و سودمندی موضوع یا پدیده خبر می‌دهد. تأکید بر نظامدار بودن فعالیت ارزشیابی، بدان جهت است که در این فعالیت، اطلاعات با روش خاص و به طور دقیق و گسترده جمع‌آوری می‌شوند و با شیوه مناسب تحلیل می‌گردند و تفسیر آنها مبتنی بر اصول و موازین علمی است. در عین حال، در نظامدار بودن، این معنا نیز نهفته است که فعالیت‌های ارزشیابی آموزشی دارای نظم خاصی هستند. این نظم که در تعریف نیز بر آن تأکید شده است، می‌تواند بیانگر مراحل انجام ارزشیابی آموزشی باشد. در فعالیت ارزشیابی آموزشی، سه گام اساسی برداشته می‌شود: نخستین گام، فراهم کردن ملاک‌هایی برای ارزش‌گذاری است. این ملاک‌ها از ارزش‌های حاکم بر نظام آموزشی، به دست می‌آیند و تدارک آنها بر پایه اصول و موازین خاصی صورت می‌گیرد. دومین گام، شناخت موضوع یا پدیده مورد نظر است که با جمع‌آوری اطلاعات دقیق و همه‌جانبه در باره آن به دست می‌آید. سومین گام، مقایسه سیمای موضوع یا پدیده مورد بررسی (وضع موجود) با ملاک‌ها و معیارهای تدوین شده (وضع مطلوب) است. در این گام، پدیده یا موضوع مورد بررسی به محک زده می‌شود تا ارزش و عیار آن تعیین شود. بحث مفصل در باره این که هر یک از سه گام مذکور را چگونه می‌توان به پیش برد و چه کسانی صلاحیت انجام آن را دارند، به آینده سپرده می‌شود و در این مقاله تلاش می‌گردد تا مشخص شود که با این مفهوم، ارزشیابی آموزشی طی دو قرن گذشته چه تحولاتی را تجربه کرده است.

فعالیت‌های قابل‌تحسینی برای توصیف سیر تحول، معنا، مراحل و چگونگی ارزشیابی به عمل آمده است و در عین حال کوشش شده است تا فعالیت‌های مختلف ارزشیابی در قالب دسته‌هایی کم و بیش همانند (طبقه) ارایه گردند (پایهام، ۱۹۷۱؛ استیک، ۱۹۷۶؛ هاوس، ۱۹۸۸؛ کیامنش، ۱۳۷۳؛ هومن، ۱۳۷۵ و سیف، ۱۳۷۵). تلاش برای طبقه‌بندی تاریخ ارزشیابی نیز چندان همسو نبوده است و یکی از پیامدهای آن، طبقه‌بندی مدوس، استافل بیم و اسکریون^۶ (۱۹۸۸) است که از شش مرحله متفاوت یاد کرده است. اینان به بررسی تاریخی تحول ارزشیابی از ابتدای قرن نوزدهم پرداخته‌اند و معتقدند طی قریب به دو قرن گذشته، ارزشیابی در مفهوم و در عمل، تحولاتی را تجربه کرده است. آنها این تحولات را در قالب شش مرحله با عناوین ذیل قرار داده‌اند: ۱- دوره دگرگونی^۷ (۱۹۰۰-۱۸۰۰)، ۲- دوره کارایی و آزمون^۸ (۱۹۳۰-۱۹۰۰)، ۳- دوره تایلری^۹ (۱۹۴۵-۱۹۳۰)، ۴- دوره معصومیت^{۱۰} (پاکی) (۱۹۵۷-۱۹۴۶)، ۵- دوره بسط^{۱۱} (۱۹۷۲-۱۹۵۸)، و ۶- دوره حرفه‌ای شدن^{۱۲} (۱۹۸۳-۱۹۷۳).

در ایران نیز ارزشیابی تحولاتی را تجربه کرده است که به دو شیوه می‌توان به طبقه‌بندی آنها اقدام کرد. شیوه نخست، بررسی و طبقه‌بندی تحولات، بدون توجه به تغییرات ارزشیابی در سایر کشورهاست به طوری که می‌توان آن را شیوه تحلیلی نام نهاد. شیوه دوم، بررسی و طبقه‌بندی تحولات بر اساس تغییراتی است که ارزشیابی در سایر کشورها داشته است و می‌توان آن را شیوه تطبیقی نام داد. در شیوه دوم، البته نیازمند به تحلیل می‌باشد؛ اما در عین حال به مقایسه نیز مبادرت می‌شود. به همین دلیل، در اینجا شیوه دوم استفاده می‌شود و طبقه‌بندی مدوس، استافل بیم و اسکریون (۱۹۸۸) مبنای کار قرار می‌گیرد.

بررسی تحولات ارزشیابی آموزشی در ایران، نشان می‌دهد که دوره دگرگونی ارزشیابی در غرب، با دوره بی‌خبری از اهمیت و نقش ارزشیابی آموزشی در ایران همزمان بوده است و می‌توان آن را دوره بی‌نیازی (۱۲۷۹-۱۱۷۹) نام نهاد. برنامه‌ریزی برای آموزش و در عین حال تلاش رسمی برای ایجاد مدارس جدید در ابتدای قرن، با توجه به برخی جوانب ارزشیابی همراه شد که در عمل، دوره آگاهی از اهمیت ارزشیابی آموزشی (۱۳۰۹-۱۲۸۰) شناخته شده است. اولین کارها برای ارزشیابی برنامه‌های آموزشی توسط کسانی به عمل آمد که در دستگاه‌های دولتی مسؤول آموزش و

برنامه‌ریزی آموزشی بودند. این دوره را می‌توان دوره تلاش رسمی (۱۳۳۴-۱۳۱۰) نام داد. تلاش‌های اولیه برای ارزشیابی برنامه، با مشکلاتی مواجه بود. کمبود نیروهای متخصص و نوشته‌های تخصصی در حوزه ارزشیابی در این سال‌ها مشهود بود و دشواری‌هایی را در امر ارزشیابی پدید می‌آورد. این دوره به همین دلیل، دوره دشواری (۱۳۳۶-۱۳۲۵) نام گرفته است. کارهای اولیه برای تغییر در نظام آموزشی و در عین حال حساسیت به نظام تازه در دهه چهل، شرایطی را پدید آورد که ارزشیابی برنامه‌ها با استقبال دست‌اندرکاران و صاحب‌نظران همراه شد که می‌توان آن را دوره علاقمندی (۱۳۳۷-۱۳۵۱) نامید. تمایل جمع قابل توجهی از دست‌اندرکاران و صاحب‌نظران در دهه های پنجاه و شصت به انجام ارزشیابی آموزشی، به این امر رونق بخشید. این دوره را می‌توان دوره اقدام (۱۳۷۰-۱۳۵۱) نام داد. در دهه هفتاد، تغییر نظام آموزش متوسطه و در عین حال تلاش برای تغییر دوره‌های پیش از متوسطه، با کوشش برای طراحی الگوهایی برای ارزشیابی همراه شد و امید می‌رود که این تلاش‌ها منجر به نتایج با اهمیتی در آینده بشود. به همین دلیل این دوره را دوره امیدواری (۱۳۷۱) می‌توان نام داد. بر این اساس، به نظر می‌رسد ارزشیابی آموزشی در ایران، هفت مرحله را پشت سر گذاشته است که شش مرحله اول آن با مراحل مطرح شده مدوس، استافل بیم و اسکریون (۱۹۸۸) انطباق داده شده است (جدول ۱).

جدول ۱: تحول ارزشیابی برنامه آموزشی

تحول ارزشیابی در غرب (مدوس و همکاران)		تحول ارزشیابی در ایران (پیشنهادی)	
نام دوره	دوره زمانی	نام دوره	دوره زمانی
۱- دگرگونی	۱۸۰۰-۱۹۰۰	۱- بی نیازی	۱۱۷۹-۱۲۷۹
۲- کارایی و آزمون	۱۹۰۰-۱۹۳۰	۲- آگاهی	۱۲۸۰-۱۳۰۹
۳- تایلری	۱۹۳۰-۱۹۴۵	۳- تلاش رسمی	۱۳۱۰-۱۳۲۴
۴- معصومیت	۱۹۴۶-۱۹۵۷	۴- دشواری	۱۳۲۵-۱۳۳۶
۵- بسط	۱۹۵۸-۱۹۷۲	۵- علاقمندی	۱۳۳۷-۱۳۵۱
۶- حرفه‌ای شدن	۱۹۷۳ تاکنون	۶- اقدام	۱۳۵۲-۱۳۷۰
		۷- امیدواری	۱۳۷۱ تاکنون

۲. دوره دگرگونی در غرب (۱۸۰۰-۱۹۰۰ م)، دوره بی نیازی در ایران (۱۱۷۹-۱۲۷۹ ش)

این دوره، در تاریخ ارزشیابی با قرن نوزدهم آغاز می‌شود. در قرن نوزدهم، انقلاب صنعتی با ایجاد تغییرات در فن‌آوری و اقتصاد، بسیاری از ساخت‌های جامعه را دگرگون کرد و تغییرات عظیم اجتماعی، انقلاب‌ها و اصلاحات متعددی طی آن ظهور یافت (پینکر، ۱۹۷۱). تلاش برای اصلاح آموزش و پرورش، دادگاه‌ها، بیمارستان‌ها و یتیم خانه‌ها و بهداشت عمومی در سراسر قرن نوزدهم در انگلیس ادامه داشت و برای ارزشیابی این گونه مأموریت‌ها و عملکردهای اجتماعی، به صورت غیر رسمی و احساسی اقدام شد و اغلب گروه‌هایی انجام این کار را بر عهده داشتند که از سوی حکومت تعیین شده بودند. انجام ارزشیابی‌های سالیانه در هر یک از نظام‌های اجتماعی هنگام به اجرا گذاشتن برنامه‌ها در بریتانیا، امری غیر معمول نبود. برای مثال، در تعلیم و تربیت بازرسان مدارس، به صورت سالیانه به هر یک از مدارس سرکشی می‌کردند و در باره موقعیت مدرسه و وضع دانش‌آموزان آن، گزارش تهیه می‌کردند (کلاگان و مدوس، ۱۹۸۲). بین سال‌های ۱۸۸۷ تا ۱۸۹۸، جوزف رایس^{۱۳} اولین ارزشیابی رسمی برنامه تربیتی در آمریکا را طرح کرد. او اقدام به انجام یک مطالعه تطبیقی در باره ارزش تمرین در آموزش خواندن کرد و تعدادی از مدارس ایالتی را در این ارزشیابی، مورد بررسی قرار داد. رایس (۱۸۹۷) نمرات آزمون را به عنوان معیار اندازه‌گیری خود، در ارزشیابی آموزش خواندن بکار گرفت. او هیچ تفاوت معنی‌داری بین آن مدارس که ۲۰۰ دقیقه در یک هفته صرف مطالعه و خواندن کرده بودند و آنهایی که کمتر از ده دقیقه در هفته به چنین کاری مبادرت ورزیده بودند، نیافت. نتایج رایس مریبان را به تجدید نظر در رویکرد به تدریس خواندن، هدایت کرد. دیدگاه تاریخی او در باره ارزشیابی برنامه، مهمتر از نتایجی

است که او به دست آورده است. او عقیده داشت که «مربیان باید متفکران تجربه‌گرا (آزمایشی) و کمی باشند (ص: ۱۴۸) و طرح تحقیق تطبیقی خود او برای مطالعه پیشرفت تحصیلی دانش‌آموزان نمونه‌ای برای این کار بود (رایس، ۱۹۱۴). رایس از پیشروان (منادبان) رویکرد تجربی به ارزشیابی است. نخستین پیرو او لینکوئیست^{۱۴} (۱۹۵۳) بود، اما فعالیت‌های کمپبل^{۱۵} و دیگران در دهه‌های ۱۹۶۰ و ۱۹۷۰ دیدگاه‌های او را گسترش داد و کامل کرد.

یک رویداد دیگر نیز در این دوره، از اهمیت برخوردار است. دیدگاه اعتبار بخشی^{۱۶} یا رویکرد قضاوت حرفه‌ای به ارزشیابی^{۱۷} که با تشکیل مؤسسه کالج‌های مرکزی شمال و مدارس متوسطه در اواخر دهه ۱۸۸۰، شکل گرفت؛ ریشه در این دوره دارد. حرکت اعتبار بخشی تا دهه ۱۹۳۰، موقعیت مناسبی بدست نیاورد و از این زمان به بعد، با اعتبار بخشی شش مؤسسه سنتی محلی که در سطح ملی (آمریکا) انجام شد، ارج و منزلتی پیدا کرد. از آن زمان حرکت اعتبار بخشی به عنوان یک ابزار عمده ارزشیابی شایستگی^{۱۸} مؤسسات آموزشی، مطرح شده است.

آموزش و پرورش ایران، طی این دوره (۱۲۷۹-۱۱۷۹) بیشتر مبتنی بر مکتب‌خانه‌ها بوده است (کمپفر، ۱۳۶۳؛ سلطان‌زاده، ۱۳۶۴؛ ضمیری، ۱۳۷۱). آموزش مکتب‌خانه‌ای در اساس با ارزشیابی همراه نبود. گرچه می‌توان نوعی مقبولیت عمومی را در مکتب‌خانه‌ها یافت که کم و بیش مبتنی بر ارزش‌گذاری بر کار مکتب‌خانه‌ها بوده است (شریعی، بی تا) و شاید بتوان با مسامحه آن را شکل اولیه ارزشیابی قلمداد کرد، اما این امر نیز دارای چند ویژگی اساسی بوده: نخست آنکه این نوع ارزشیابی‌ها غیر رسمی بوده است. ارزشیابی فعالیت مکتب‌خانه‌ها، کثرتی بوده است که توسط افراد علاقمند به تحصیل برای خود یا فرزندان انجام می‌شده است و بر پایه اطلاعاتی صورت می‌گرفته. که به شیوه غیر علنی بدست می‌آمده‌اند. دوم آنکه این ارزشیابی‌ها سازمان‌نیافته بوده است. شیوه خاصی بر کار ارزشیابی حاکم نبوده است و ملاک‌های مشخصی برای تعیین ارزش مورد استفاده قرار نمی‌گرفته است. سوم آنکه این ارزشیابی‌ها شخصی بوده است. ارزشیابی عمومی از مکتب‌خانه، کمتر متوجه برنامه‌های آموزشی و بیشتر متوجه مکتب‌دار بوده است. به عبارت دیگر، ارزشیابی، متوجه شخص اداره کننده مکتب بوده است نه متوجه برنامه‌ها و یا حتی وضعیت مکتب‌خانه.

لازم به ذکر است که در میانه این دوره مدرسه دارالفنون آغاز به کار کرده است (محبوبی اردکانی، ۱۳۷۰)، و در پی آن مدارس از قبیل مشیریه، رشیدیه، ادیب، سادات، اسلام، اقدسیه، علمیه، کمالیه، مروّت، سلطانی، شرف، مظفریه، دانش، مدیریه و خیریه تأسیس شده‌اند (صافی، ۱۲: ۱۳۷۲). این مدارس که با الگوگیری از مدارس اروپایی در ایران شروع به کار کردند؛ از نظم و ترتیب و مقرراتی تبعیت می‌کردند که در مدارس اروپایی رواج داشت (محبوبی اردکانی، ۱۳۷۰). مسأله آزمون از آموخته‌های دانش‌آموزان در این مدارس، معمول بود؛ اما ارزشیابی برنامه به معنایی که امروزه مورد نظر است، در این مدارس دیده نمی‌شد. بیشتر مواد درسی و برنامه‌ها از زبان‌های خارجی ترجمه شده بود و شاید این امر بر احساس عدم نیاز به ارزشیابی آنها تأثیر داشته است.

۳. دوره کارایی و آزمون (۱۹۳۰-۱۹۰۰)، دوره آگاهی در ایران (۱۳۰۹-۱۲۷۹)

طی بخش اول قرن بیستم، اندیشه مدیریت علمی در نظریه رهبری تربیتی، همانند بخش‌های صنعتی قوت یافت (بیدل و الن،^{۱۹} ۱۹۶۴؛ کالاهان، ۱۹۶۲؛ کریمین،^{۲۰} ۱۹۶۲). این حرکت بر نظم بخشی، معیار کردن و مهمتر از همه کارایی تأکید داشت. نمونه‌ای از این تأکید بر کارایی در عناوین چهاردهمین و پانزدهمین کتاب‌های سال انجمن ملی مطالعه تعلیم و تربیت^{۲۱} که با عناوین «روش‌هایی برای سنجیدن کارایی معلمان» و «استانداردها و آزمون‌هایی برای سنجش کارایی مدارس و نظام مدرسه» منتشر شدند؛ مشاهده می‌شود.

در بسیاری از نظام‌های بزرگ مدرسه‌ای، طی این دوره، با تأکید بر اثربخشی مدرسه و یا معلم و با استفاده از ملاک‌های متفاوتی (مثل مخارج، میزان افت تحصیلی دانش‌آموز، میزان پیشرفت و ...) بررسی‌هایی انجام شد. در سال ۱۹۱۵، سی تا چهل نظام مدرسه‌ای بزرگ از حیث همه مراحل زندگی آموزشی، تحت بررسی‌های جامع قرار گرفتند (کنندال، ۱۹۱۵؛ اسمیت و جود، ۱۹۱۴). در تعداد زیادی از این بررسی‌ها، به شیوه نوینی برای آزمون‌های ریاضیات، خواندن، املا و انشای انگلیسی، «هدف» تعیین شد تا از این طریق، کیفیت تدریس مورد سنجش قرار گیرد. این

آزمون‌ها، اغلب در ایالت‌های بزرگ آمریکا توسط یک هیأت یا دانشکده تهیه می‌شد تا برای بهبود کارایی ناحیه، مورد استفاده قرار گیرند. برای مثال، دانشکده مدیریت و پژوهش تربیتی در مدارس دولتی بوستون، تعدادی آزمون تهیه کرد که امروزه آزمون‌های "هدف محور"^{۲۲} نامیده می‌شوند (بالو، ۱۹۱۶). شاید آزمون‌هایی از قبیل آزمون‌های بوستون، دارای خصیصه "هنجار محوری"^{۲۳} بوده‌اند؛ زیرا درصد دانش‌آموزانی که موفق به گذراندن آنها می‌شدند، معیاری بوجود می‌آورد تا بر آن اساس، معلمان قضاوت کنند که آیا کلاس آنها پیشتر یا پس‌تر از معیارهای عمومی برای شهر بوده است (بالو، ۱۹۱۶). در کنار این آزمون‌ها، تعدادی آزمون نیز بوسیله پژوهشگرانی از قبیل کورتیس، آیرز، تورندایک و دیگران تهیه شد که مجموعه خیلی دقیقی از هدف‌های آموزشی را اندازه می‌گرفتند. این آزمون‌ها که به وسیله پژوهشگران مشهور روز تهیه شده بودند، داده‌های هنجاری برای مقایسه یک سیستم با خودش و با دیگر سیستم‌ها فراهم می‌آوردند (تیک و هانسوت، ۱۹۸۲).

بسیاری از مطالعات اوایل قرن بیستم، نمونه روش قدیمی از آشکار سازی مشکلات بودند. مسأله دیگری که با مطالعات اولیه مرتبط است، اما ممکن است برای ارزشیابان امروزی شناخته شده نباشد، استفاده از نتایج عینی برای تبلیغات بود.

با توسعه استاندارد پیشرفت تحصیلی بعد از جنگ جهانی اول، مدارس ایالتی از این آزمون‌ها برای مشخص کردن اثربخشی برنامه استفاده کردند (می، ۱۹۷۱). در سراسر تاریخ، ارزشیابی با آزمودن^{۲۷} همراه بوده است. داده‌های آزمون، اغلب منبع اصلی داده‌های ارزشیابی بوده‌اند و استفاده از آزمون‌ها یک موهبت چند وجهی در امر ارزشیابی محسوب می‌شود.

در اواخر دهه ۱۹۲۰ و اوایل دهه ۱۹۳۰؛ مؤسسات دانشگاهی ویژه رشته پژوهش (مطالعات) در آمریکا پدید آمدند و مطالعاتی را برای مدارس محلی انجام دادند. مشهورترین این مؤسسات موردی بود که بوسیله جورج استریر در کالج معلمان تأسیس شد (تیک و هانسوت، ۱۹۸۲). این مؤسسات، منادیان آن دست از مراکز دانشگاهی بودند که به رشد ارزشیابی در دهه‌های ۱۹۶۰ و ۱۹۷۰ کمک کردند.

نکته مهم آن است که مطالعات اثر بخشی و آزمودن، در بخش اعظم مدارس محلی در ایالات متحده، مورد قبول واقع شدند. در آمریکا، بر خلاف طرح‌های برنامه‌ریزی درسی ملی اواخر دهه ۱۹۵۰ و اوایل دهه ۱۹۶۰؛ برنامه‌ریزی درسی، تا قبل از دهه ۱۹۳۰ به طور وسیعی در دست معلمان یا کمیته‌های معلمان بود. به هر حال طبیعی بود که این دسته از مطالعات ارزشیابی در پی پاسخگویی به پرسش‌های محلی باشد. این تأکید و تمرکز بر پرسش‌های محلی ارزشیابی تا دهه ۱۹۶۰ ادامه پیدا کرد و این مسأله بر کاربرد نتایج ارزشیابی در دهه ۱۹۶۰ مؤثر بوده است. در دهه ۱۹۷۰، جهت‌گیری ارزشیابی تغییر کرد و مسأله عمومیت بخشی (تعمیم) برای مربیان و ارزشیابان مطرح شد و مورد بحث قرار گرفت.

این دوره (۱۳۰۹-۱۲۷۹) در ایران، با شروع فعالیت‌های رسمی برای سامان بخشی فعالیت‌های آموزش و پرورش همراه است. ارتباط و آشنایی با کشورهای اروپایی، که از طریق سفرهای سردمداران، اعزام دانشجویان، و پذیرش هیأت‌های نمایندگی صورت گرفت، زمینه را برای تشکیل مدارس جدید فراهم کرد. اولین مدارس توسط انجمن‌ها و هیأت‌های تبلیغ مسیحیت و بعضی از تحصیل کردگان ایرانی به صورت خصوصی تأسیس شدند (محبوبی اردکانی، ۱۳۷۰). گرچه در ابتدا با این مدارس، مخالفت‌هایی از سوی مردم و اصناف صورت می‌گرفت و حتی تحصیل علوم جدید از سوی عده‌ای ممنوع اعلام شد (شریعتی، بی تا) اما این مدارس، کم‌کم جای خود را در جامعه باز کردند.

پیروزی مشروطه خواهان و تدوین قانون اساسی مشروطه، شرایط تازه‌ای را برای آموزش و پرورش پدید آورد. طبق اصول ۱۸ و ۱۹ قانون اساسی مشروطه، آموزش، اجباری و همگانی اعلام شد و در اصل نوزدهم چنین بیان شد که "تأسیس مدارس به مخارج دولتی و ملتی و تحصیل اجباری باید مطابق قانون وزارت علوم و معارف مقرر شود" (ثواقب، ۶: ۱۳۶۹). بر همین اساس دولت به تأسیس وزارتخانه‌ای به نام "علوم" اقدام کرد تا مسؤلیت نظارت بر تمامی فعالیت‌های آموزشی را عهده‌دار شود. این وزارتخانه پس از دو سال به وزارت معارف تغییر نام داد، و اولین دستور تعلیمات شش ساله مدارس ابتدایی و سه ساله مدارس متوسطه را در سال ۱۲۸۶ به تصویب رساند (مجله تعلیم و

تربیت، ۱۳۱۳). قانون اساسی معارف، مصوبه دیگر این وزارتخانه است که در سال ۱۲۹۰، تصویب شده است و بر پایه آن، تعلیمات ابتدایی برای تمام ایرانیان، اجباری اعلام گردید و قید شده است که هر کس مکلف است فرزند خود را از سن هفت سالگی به تحصیل ابتدایی بگمارد (اداره کل نگارش، ۱۳۱۷). تأسیس وزارتخانه‌ای به منظور تنظیم آموزش در ایران نقطه عطفی در تاریخ تشکیل مدارس جدید و بخصوص تمرکزگرایی در آموزش و پرورش ایران محسوب می‌شود. این وزارتخانه در سال ۱۳۰۶، برای نخستین بار، برنامه ساعات هفتگی دروس در مدارس ابتدایی را با عنوان "دستور تعلیمات شش ساله ابتدایی مدارس" به تصویب رساند (مجله وزارت معارف و اوقاف و صنایع مستظرفه، ۱۳۰۶).

این دوره را می‌توان کم و بیش آغاز فعالیت‌های رسمی ارزشیابی مدارس قلمداد کرد اما کمتر می‌توان نشانه‌ای از ارزشیابی برنامه را در آن نظاره نمود. طبق ماده دوازدهم قانون اساسی معارف: "مفتشین دولتی در کل مدارس و مکاتب رسمی و غیررسمی حق ورود و تفتیش دارند و هیچ مدیری در هیچ موقعی حق مخالفت ندارد" (قانون اساسی معارف، ۱۲۹۰). محتوای این ماده وقتی که در کنار دستورات مطرح شده در سایر مواد (۲۸ ماده) قرار می‌گیرد، مؤید توجه به ارزشیابی مدرسه است. توجه به این مسأله که قانون اساسی معارف تهیه برنامه مدارس (پروگرام) را بر عهده وزارت معارف نهاده بود و وزارت معارف نیز در سال ۱۳۰۶ برای نخستین بار ساعات درسی مدارس ابتدایی را مشخص کرده است، شاهد دیگری بر مدعای مطرح شده است.

در این دوره بازرسان (یا مفتشین) دولتی برای بررسی مدارس اقدام می‌کردند و اطلاعاتی را در باره آنها جمع‌آوری و گزارش‌هایی را تهیه می‌کردند. اما روح حاکم بر این بررسی‌ها تأیید یا تعطیل کردن برخی مدارس با انگیزه‌های دیگری بوده است. به عبارت دیگر ارزشیابی مدارس بیش از آنکه به مسأله اصلاح مدارس نظر داشته باشد یک ابزار سیاسی شده بود.

۴. دوره تایلری در غرب (۱۹۴۵-۱۹۳۰)، دوره تلاش رسمی در ایران (۱۳۲۴-۱۳۱۰)

رالف دبلیو تایلر بر تعلیم و تربیت به طور کلی، و بر ارزشیابی تربیتی به طور خاص آشکارا اثر داشته است. او را اغلب پدر ارزشیابی تربیتی معرفی می‌کنند. تایلر با ارایه یک دیدگاه گسترده در باره برنامه درسی و ارزشیابی، کار خود را آغاز کرد. در دیدگاه او، برنامه درسی مجموعه‌ای از تجربیات مدرسه‌ای برنامه‌ریزی شده تصور می‌شود که به منظور کمک به دانش‌آموزان، طراحی و اجرا می‌شود تا به بروندهای رفتاری مشخص دست یابند (تایلر، ۱۹۴۹). تایلر، واژه ارزشیابی تربیتی را به منظور سنجش مقدار هدف‌هایی که به عنوان بخشی از برنامه آموزشی تحقق یافته‌اند ابداع کرد. طی اوایل و نیمه دهه ۱۹۳۰ او برداشت خود از ارزشیابی را برای کمک به مریبان ایالت اوهایو بکار گرفت تا تدریس و آزمون توسط آنها را بهبود بخشد.

در آمریکا، طی دوره رکود بزرگ^{۲۸} مدارس مثل بسیاری دیگر از مؤسسات ملی از فقر منابع در رنج بودند. همزمان با کوشش‌های روزولت با برنامه‌های نیودیل^{۲۹} برای مشکل‌گشایی از اقتصاد، جان دیویی نیز کوشش می‌کرد تا تعلیم و تربیت را بازسازی^{۳۰} کند. بازسازی در تعلیم و تربیت که امروزه با عنوان نهضت آموزش و پرورش پیشرو^{۳۱} شناخته می‌شود، فلسفه پراگماتیسم^{۳۲} و ابزارهای مورد استفاده در روان‌شناسی رفتارگرا^{۳۳} را وارد عرصه آموزش و پرورش کرد. تایلر وقتی که مقدمات کار را برای انجام مطالعه مشهور هشت ساله (اسمیت و تایلر، ۱۹۴۲) فراهم می‌کرد بطور مستقیم به نهضت آموزش و پرورش پیشرو پیوست. مطالعه هشت ساله (۱۹۴۰-۱۹۳۲)، برای پاسخگویی به این پرسش دهه ۱۹۳۰ انجام شد که تجربیات مدارس عالی چه ارتباطی با پیشرفت تحصیلی در مدارس متوسطه دارند.

در مطالعه هشت ساله، دیدگاه تازه‌ای از ارزشیابی مطرح شد و با آنچه در اثر بخشی (کارایی) و آزمون مطرح بود، تفاوت داشت. ارزشیابی که به وسیله تایلر مفهوم پردازی شده بود به معنای مقایسه بین بروندهای خواسته شده^{۳۴} و بروندهای واقعی^{۳۵} در نظر گرفته شد. در ارزشیابی تایلری، مقایسه درونی بروندها با هدف‌ها^{۳۶} مطرح بود و بر این اساس، به مقایسه بین گروه‌های آزمایشی و گواه (کنترل)، چنانکه در رویکرد تجربی مبتنی بر نظر رایس مطرح بود، نیازی نبود. از آنجا که در این رویکرد، اندازه‌گیری هدف‌هایی که بطور رفتاری تعریف شده‌اند، مبنای کار ارزشیابی است؛ به جای تأکید بر درون داده‌های سازمانی و آموزش، بر بروندهای یادگیری تمرکز می‌شود. بدین ترتیب، تایلر با

کارها و نوشته‌های خود در دهه ۱۹۴۰، بنیادی را پدید آورد که بر فعالیت‌های آموزشی و تربیتی بطور عام و بر آزمون‌داری و ارزشیابی به طور خاص، طی ۲۵ سال بعد اثر داشته است.

این دوره در ایران (۱۳۲۴-۱۳۱۰)، با تدوین برنامه‌ها و قوانین متعددی برای آموزش و پرورش، همراه بوده است. برنامه‌های تهیه شده چندان دوامی نداشتند و با سرعت تغییر می‌کردند. برنامه‌تهیه شده برای آموزش ابتدایی که مصوب سال ۱۳۰۶ بود، در سال ۱۳۱۶ و ۱۳۱۹ دو بار تغییر کرد (اداره کل نگارش، ۱۳۱۷ و ۱۳۲۲). همچنین، در این دوره (سال ۱۳۲۲)، قانون آموزش و پرورش اجباری و مجانی تصویب شد و به دولت ده سال فرصت داد تا آموزش و پرورش ابتدایی را در تمام کشور عمومی و اجباری سازد (اداره کل نگارش، ۱۳۲۲).

طی این دوره، تلاش برای متمرکز کردن برنامه‌های درسی، توسط دولت افزایش یافت. تألیف کتاب‌های یک‌دست برای دبیرستان‌ها، که در سال ۱۳۱۷ آغاز شد و تا سال ۱۳۲۰، هشتاد جلد کتاب این دوره آموزشی را آماده کرد، (ریاحی، ۴: ۱۳۴۲) از جمله این فعالیت‌هاست. گرچه برنامه‌ریزی متمرکز می‌توانست زمینه مناسبی برای ارزشیابی برنامه‌های درسی فراهم کند؛ اما چنین نتیجه‌ای از آن حاصل نشد. تغییراتی که در برنامه‌ها و کتاب‌های درسی صورت می‌گرفت، فاقد پشتوانه پژوهشی و ارزشیابی بود و به همین دلیل است که در دهه بعد، وقتی تجربه کتاب‌های یک‌دست مورد بازنگری قرار می‌گیرد ضمن اذعان به برخی نقص‌هایی که در کتاب‌ها بوده است، ادعا می‌شود: «تردیدی نیست که اگر تدریس آن کتاب‌ها چند سال ادامه می‌یافت و دستگاه منظمی ضمن آزمایش با استفاده از انتقادات صاحب‌نظران در اصلاح و تهذیب و تکمیل آنها می‌کوشید، بطور قطع کتاب‌های مطلوبی فراهم می‌آمد» (ریاحی، ۵: ۱۳۴۲).

نکته حایز اهمیت در این دوره، کنترل محتوای برنامه‌های درسی توسط وزارت معارف (فرهنگ) است. شورای عالی فرهنگ، که مسؤلیت نظارت و کنترل و سامان بخشی فعالیت‌های آموزشی را بر عهده داشت، از سال ۱۳۱۶، برای تهیه برنامه‌های درسی مناسب‌تر به این فکر افتاد که سرفصل درس‌های هر یک از برنامه‌ها را مشخص کند و از مؤلفان برای تهیه کتاب درسی دعوت نماید و مجوز استفاده از کتاب در مدرسه را پس از تطبیق کتاب تهیه شده با سرفصل مورد نظر صادر کند (ریاحی، ۱۳۴۲ و ودیعی، ۱۳۵۳). این دوره، که در تحول برنامه‌ریزی درسی ایران دوره انتخاب مقید (موسی‌پور، ۱۳۷۴) نام گرفته است، تا اوایل دهه ۱۳۴۰ ادامه یافت. در این دوره، در واقع نوع خاصی از ارزشیابی برنامه درسی وجود دارد که توسط یک سازمان مرکزی و به شیوه رسمی انجام می‌شود؛ البته برنامه درسی در اینجا به معنای محدود آن یعنی کتاب درسی مورد نظر است.

گرچه در این دوره فعالیت‌های پراکنده‌ای برای بررسی برنامه‌های درسی توسط صاحب‌نظران و معلمان انجام می‌شد؛ اما مبنای اصلی این نوع ارزشیابی‌ها بر اطلاعاتی بود که حاصل نظرات شخصی بودند و از نظم خاصی برخوردار نبودند. در عین حال، بیشتر متوجه ظاهر برنامه‌های درسی بودند (ره‌آورد، ۱۳۲۷؛ ریاحی، ۱۳۴۲). حتی فعالیت‌های دفتر تطبیق کتاب‌های درسی نیز از نظم و قاعده خاصی برخوردار نبود.

در این دوره همچنین توجه به آزمون‌های روانی (هوش و معلومات) آغاز شد. برخی مریبان که با این آزمون‌ها در کشورهای غربی آشنایی پیدا کرده بودند، سعی کردند کاربرد آنها را در ایران رواج بخشند. فعالیت‌هایی برای معرفی و معیار قرار دادن این نوع آزمون‌ها در ایران انجام شد که در رأس آنها، انتشار کتاب «سنجش هوش یا روان‌شناسی عملی: تجارب روان‌شناسی در آزمایشگاه» توسط مرحوم دکتر محمدباقر هوشیار در سال (۱۳۱۷)، قرار داشت.

۵. دوره معصومیت در غرب (۱۹۵۷-۱۹۴۶)، دوره دشواری در ایران (۱۳۳۶-۱۳۲۵)

گرچه دوره (۱۹۵۷ تا ۱۹۴۶) دوره معصومیت نام گرفته است؛ اما تا همین حد مناسب خواهد بود اگر دوره گناهکاری نام گیرد (مدوس، استافل بیم و اسکریون، ۱۹۸۸). این دوره در اروپا با بازسازی خرابی‌های جنگ همراه بوده است. به منظور بازسازی و اصلاح نظام آموزش و پرورش، کارکنان و امکانات آموزش و پرورش گسترش یافت، ساختمان‌های جدید برپا شد و انواع جدید مؤسسات تربیتی از قبیل کالج‌های تجربی و کالج‌های عمومی پدید آمد. مدارس کوچک ایالتی با همدیگر ادغام شدند تا توانایی لازم برای فراهم کردن طیف وسیعی از خدمات تربیتی از قبیل خدمات بهداشت روانی و جسمی، تغذیه، آموزش موسیقی، برنامه‌های ویژه، آموزش فنی و بازرگانی و آموزش مداوم که

در نظام مدارس بزرگتر مرسوم بود را بدست آورند. اواخر دهه (۱۹۵۰-۱۹۴۰) در سراسر جامعه آمریکا و اروپا، زمان فراموشی جنگ، پشت بفر گذاشتن رکود، سازندگی، افزایش توانایی، دستیابی به منابع، و زمان مهندسی کردن و پدید آوردن یک زندگی خوب بود.

این حس عمومی در جامعه و در آموزش و پرورش آمریکا و اروپا، در ارزشیابی نیز انعکاس یافت. در حالیکه آموزش و پرورش به سرعت توسعه می یافت و به حد اعلاى خود نزدیک می شد و افزایش درآمدهای ناشی از مالیات، نگرانی برای منابع کشاورزی را تقلیل می داد؛ هیچ توجه خاصی صرف بخشی از جامعه که مسؤول تربیت بودند، نمی شد. تقاضا برای مدیریت مؤثر و افزایش کارایی کوشش هایی که مربیان به عمل می آوردند، اندک بود. مربیان در باره ارزشیابی حرف می زدند و مطالبی را می نوشتند؛ اما آنها معمولاً موارد خاصی از داده ها را جمع آوری می کردند که بدان نیاز داشتند؛ معمولاً برای توسعه یا تحدید یک برنامه جدید، قضاوت هایی انجام می شد. به هر حال مدارک اندکی وجود داشت و همین داده های اندک برای قضاوت و اصلاح کیفیت برنامه ها مورد استفاده قرار می گرفتند. طی این دوره، رشد قابل توجهی در برخی از جنبه های ارزشیابی حاصل شد که با توسعه مرسوم انواع دیگر تکنولوژی همراه بود. این رشد کم و بیش از رویکرد آزمون در ارزشیابی تأثیر پذیرفته بود؛ اما در عین حال در رویکرد آزمایشی و رویکرد تجانس بین هدف ها و برون داده ها نیز ریشه داشت. در میان این تحولات رشد نهضت استاندارد کردن آزمون با اهمیت (مهم) است. برخی، آزمون ها را خریداری نموده و بر روی هزاران نفر به اجرا در آوردند و با استفاده از ماشین به نمره گذاری آن اقدام کردند و با روش های جدید آماری به تحلیل آن پرداختند. نهضت آزمون سازی در سال ۱۹۴۷ وقتی که ای. اف لینکوئیست و رالف تایلر و دیگران به تأسیس مرکز آزمون سازی تربیتی^{۳۷} همت گماشتند به ترقی دیگری دست یافت. طی دهه ۱۹۵۰ عمل استاندارد کردن آزمونها گسترش قابل توجهی یافت و سازمان های حرفه ای مرتبط با آزمون سازی در صدد بر آمدند تا مجموعه ای از اقدامات طراحی شده را برای تنظیم فعالیت های مرتبط با آزمون سازی اعضای خود فراهم کنند. در سال ۱۹۵۴ یکی از کمیته های انجمن روان شناسی آمریکا توصیه های تکنیکی برای آزمون های روان شناختی و تکنیک های تشخیصی^{۳۸} (APA، ۱۹۵۴) را تهیه کرد و منتشر ساخت. پس از این اقدام فعالیت های مشابهی از سوی سازمان های مربوط به عمل آمد (مدوس، استافل بیم و اسکریون، ۱۹۸۸).

توسعه آزمون سازی فقط با تحول تکنیکی مرتبط با ارزشیابی برنامه طی این دوره همراه نبوده است. لینکوئیست (۱۹۵۳) اصول آماری طرح تجربی را تهیه و ارایه نمود. این امر در ابتدا مورد استقبال قرار گرفت ولی بعدها برخی ارزشیابان و مربیان دریافته اند که مسایل مربوط به کاربری طرح های تجربی در مدارس غلبه نیاافتنی هستند.

در طی دهه ۱۹۵۰ و دهه ۱۹۶۰، همچنین رشد تکنیکی قابل توجهی در ارتباط با دیدگاه تایلری ارزشیابی حاصل شد. در رویکرد تایلر به دلیل اینکه ارزشیابی نیازمند هدف های واضح بود، لازم بود تا مربیان و دیگر کسانی که وظیفه تعیین هدفها را بر عهده داشتند وظایف خود را به خوبی انجام دهند. بر همین اساس تکنیک هایی برای کمک به برنامه ریزی که در پی تعیین هدفها بودند تهیه شد که از میان آن می توان به طبقه بندی هدف های تربیتی^{۳۹} (بلوم و دیگران، ۱۹۵۶ و کراتول و دیگران، ۱۹۶۴) اشاره کرد که بدین منظور پدید آمده بود. مدل عقلانی تایلر طی این دوره به طور فزاینده ای برای تربیت معلم و برای تهیه آزمون ها مورد استفاده قرار گرفت. تا این زمان، هنوز ارزشیابی محدود به مأموریت های محلی بود و مأموریت های ایالتی و ملی، در فعالیت های ارزشیابی مورد نظر قرار نمی گرفتند (مدوس، استافل بیم و اسکریون، ۱۹۸۸).

آموزش و پرورش ایران طی این دوره (۱۳۲۶-۱۳۲۵) در تمام مقاطع شاهد تحولاتی بوده است. برنامه های آموزش ابتدایی دستخوش تغییر شد (اداره کل تعلیمات ابتدایی کشور، ۱۳۳۸)، طول دوره و رشته های دوره متوسطه دگرگونی یافت (گویا، ۱۳۷۸) و نه تنها دانشگاه ها و دانشکده های بیشتری شروع بکار کردند، بلکه رشته های فراوان تری نیز در آنها دایر گشت. طی این دوره، کار برنامه ریزی درسی به شیوه "انتخاب مقید" (موسی پور، ۱۳۷۴) انجام می شد و ارزشیابی برنامه های درسی کم و بیش روال دوره قبل را داشت. ارزشیابی برنامه های درسی به شکل رسمی و توسط ادارات وزارت فرهنگ انجام می شد و فعالیت های درون این ادارات به تطبیق محتوا با سر فصل تهیه شده محدود بود و در این ادارات نیز به ندرت کسی اقدام به ارزشیابی برنامه درسی می کرد. فقدان ملاک های مشخص برای ارزشیابی

برنامه‌های درسی در ادارات مسئول این کار، جنبه‌های مختلف نیازمند بررسی و تصمیم‌گیری، و در عین حال تعداد قابل توجه کتاب‌های متقاضی کسب مجوز کتاب درسی، شورای عالی فرهنگ را بز آن داشت تا در اواخر این دوره اساسنامه‌ای در باره رسیدگی به کتاب‌های درسی دبیرستانی و چگونگی چاپ و صحافی و تعیین بهای آنها به تصویب برساند (ریاحی، ۱۳۴۲). بر اساس این اساسنامه، که در سال ۱۳۳۵ به تصویب رسیده است مقرر شد، کتاب‌های درسی به منظور حصول اطمینان از صحت مطالب، هماهنگی با برنامه، رعایت نظم منطقی و هماهنگی اصطلاحات، سادگی و روانی عبارات و ملاحظات تربیتی در کمیسیون‌هایی مورد رسیدگی قرار گیرند و ناشران در چاپ کتاب‌ها نکات مورد نظر وزارت فرهنگ را مراعات کنند (نائینی، ۵: ۱۳۶۲). کمیسیون‌های رسیدگی زیر نظر اداره مطالعات و برنامه‌ها عمل می‌کردند که در سال ۱۳۳۴ تأسیس شده بود. هدف این اداره آن بود که فرصتی را برای مطالعه و تحقیق پیرامون هدف‌ها و تجدید نظر در برنامه‌ها و روش‌های تعلیم و تربیت ایجاد نماید (نائینی، ۵: ۱۳۶۲). برای تحقق بخشیدن به هدف مذکور بود که اداره مطالعات و برنامه‌ها از سال ۱۳۳۵ کار تألیف کتاب را به منظور ارایه کتاب‌هایی نمونه برای مؤلفان آزاد آغاز کرد (خواجه نصیر طوسی، ۱۸۸: ۱۳۵۱).

۶. دوره بسط در غرب (۱۹۷۲-۱۹۵۸)، دوره علاقمندی در ایران (۱۳۵۱-۱۳۳۷)

دوره معصومیت در ارزشیابی تا رسیدن به اواخر دهه ۱۹۵۰ و اوایل دهه ۱۹۶۰ به طول انجامید. در این زمان ارزشیابی برنامه‌های درسی گسترده - مقیاس با تکیه بر سرمایه‌های ملی مطرح شد. همین امر سبب پایان بخشیدن به دوره معصومیت در ارزشیابی شد و تغییراتی را در پی داشت که ارزشیابی را به عنوان یک صفت و یک حرفه مبتنی بر درآمدهای مالیاتی مطرح نمود.

موفقیت روس‌ها در پرتاب اسپوتنیک در سال ۱۹۵۷ سبب شد تا دولت آمریکا مسأله قانون دفاع ملی آموزش^{۳۸} را مطرح نماید و در سال ۱۹۵۸ آن را به صورت قانون به تصویب رساند. از جمله پیامدهای این قانون فراهم شدن زمینه برای اصلاح برنامه‌های آموزش ریاضیات، علوم و زبان‌های خارجی بود و خدمات مشاوره و راهنمایی و برنامه‌های آزمون را در مدارس محلی افزایش بخشید. پروژه‌های جدید برنامه‌ریزی برای برنامه درسی ملی به خصوص در حوزه علوم و ریاضیات مطرح شد و در عین حال مؤسساتی برای ارزشیابی این فعالیت‌های برنامه‌ریزی درسی تأسیس گردید. هر چهار رویکرد به ارزشیابی که مورد بحث بوده اند طی این دوره بسط یافته‌اند. نخست رویکرد تایلری است که برای کمک به تعریف کردن هدف‌های برنامه‌های درسی جدید و برای سنجش میزان تحقق یافتن هدف‌ها مورد استفاده قرار گرفت. دوم رویکرد آزمون است. در این دوره آزمون‌های جدیدی در سطح ملی خلق و استاندارد شدند تا هدف‌ها و محتوای برنامه درسی جدید را بهتر منعکس نمایند. سوم رویکرد قضاوت حرفه‌ای است که برای درجه‌بندی طرح‌ها و برای کنترل دوره‌ای در فعالیت‌های مقاطعه کاری بکار گرفته شد. چهارم، رویکرد آزمایشی است که توسط بسیاری از ارزشیابان، برای ارزشیابی فعالیت‌های برنامه‌ریزی درسی بکار گرفته شد.

با تأسیس اتحادیه حرفه‌ای^{۳۹} (P.D.K. 1971) یکی از کمیته‌های آن، مأموریت یافت تا یک مطالعه ملی در باره ارزشیابی به عمل آورد. این کمیته ارزشیابی تربیتی را مستولی شدن بر یک بیماری (شرارت) عظیم^{۴۰} تعریف کرد و اعلام داشت: توسعه نظریه‌های جدید و روش‌های نوین ارزشیابی به اندازه برنامه‌های جدید آموزش برای ارزشیابی ضروری است (کمیسیون فای دلنا کاپا برای ارزشیابی، ۱۹۷۱). در همین زمانها است که بعضی مفهوم پردازی‌های تازہ از ارزشیابی آغاز شد. پرووس، (۱۹۶۹ و ۱۹۷۱)؛ هاموند،^{۴۱} (۱۹۶۷)؛ آیزنر، (۱۹۶۷) و متفسل و مایکل،^{۴۲} (۱۹۶۷)، به منظور اصلاح الگوی تایلر اقدام کردند. گلاسر،^{۴۳} (۱۹۶۳)؛ تایلر، (۱۹۶۷)؛ و پاپهام، (۱۹۷۱) آزمون‌های (ملاک - مرجع)^{۴۴} را به عنوان رقیب برای آزمون‌های (هنجار - مرجع)^{۴۵} مطرح کردند. کوک (۱۹۶۶) به کار بستن رویکرد تحلیل سیستمی را برای ارزشیابی برنامه‌ها مطرح کرد. اسکریون، (۱۹۷۵)، استافل بیم، (۱۹۶۷ و ۱۹۷۱) با دیگران و استیک، (۱۹۷۶) الگوهای تازه‌ای برای ارزشیابی مطرح کردند که بطور افراطی از رویکردهای پیشین متفاوت بود. این مفهوم پردازی تازه از ارزشیابی هدفهای آموزشی، درون داد‌ها، کاربرد و نوع خدمات را همسنگ برون‌دادهای خواسته شده و ناخواسته برنامه قرار داد. در این نظریه‌ها همچنین بر نیاز به قضاوت در باره قدر و ارزش موضوع مورد ارزشیابی تأکید

می شد. اواخر دهه (۱۹۶۰) و اوایل دهه (۱۹۷۰)، با تعابیر، مباحثات و جدل هایی در باره "چگونه باید ارزشیابی کرد؛ همراه بود و جوی متشنج را پدید آورده بود. به هر حال این دوره در تاریخ ارزشیابی برنامه، با کاهش توجه نسبت به ارزشیابی پایان می یابد. برخی ارزشیابی های مهم به یافته های منفی منتهی شد که اولین آنها مطالعه مشهور کلمن بود. این مطالعه که با نام "برابری فرصت های آموزشی" (کلمن و همکاران، ۱۹۶۶) انجام می شد، این نتیجه را در پی داشت که "توفیق مدارس در تأثیر بر پیشرفت تحصیلی کودکان مستقل از سوابق و زمینه اجتماعی عمومی آنها اندک است". البته در پژوهش های دیگری (پیکاریلو، ۱۹۶۲؛ گلاس و دیگران، ۱۹۷۰؛ مؤسسه تربیتی ایالات متحده، ۱۹۷۰) اثر بخشی برنامه های مدارس نشان داده شده است.

آموزش و پرورش ایران در این دوره (۱۳۵۱-۱۳۳۷) شاهد تحولی در برنامه ریزی درسی است و با عنوان "دوره انتخاب محدود" (موسی پور، ۱۳۷۴) مطرح شده است. گرچه سال های اولیه این دوره، با همان سبک و سیاق دوره قبل سپری شده است؛ اما از سال ۱۳۴۲، که با تصویب هیأت وزیران، سازمانی برای تألیف و تدوین کتابهای درسی بوجود آمد (نائینی، ۱۳۶۲)؛ ارزشیابی برنامه درسی نیز شاهد تغییراتی بوده است. "سازمان کتابهای درسی ایران" وظیفه داشت به تألیف و تهیه کتابهای درسی، تدوین کتابهای تکمیلی، تهذیب و تنقیح و تکمیل کتاب هایی که از طرف مؤلفان آزاد، به عنوان کتاب درسی و تکمیلی تهیه می شد حق التألیف مؤلفان را پرداخت نماید" (نائینی، ۱۳۶۲). تغییر نظام آموزشی، که از سال ۱۳۴۵ آغاز شد، زمینه مناسبی را برای فعالیت های گسترده سازمان کتابهای درسی فراهم آورد و با تأیید افکار عمومی و حمایت های سیاسی، دوره "یک درس، یک کتاب" در آموزش و پرورش ایران آغاز شد.

در این دوره، اداره ای به منظور ارزشیابی برنامه های درسی با عنوان "اداره نظارت و ارزشیابی" در درون سازمان کتابهای درسی تأسیس شد (نائینی، ۱۳۶۲) که برنامه های درسی تهیه شده را مورد ارزشیابی قرار می داد. طی این دوره، گرچه کتابهای درسی کم و بیش بدون ارزشیابی تکوینی تهیه می شدند؛ اما پس از آنکه استفاده از آنها شروع می شد، مورد ارزشیابی قرار می گرفتند. روشهای مختلف برای ارزشیابی کتابهایی که تهیه شده بود، مورد استفاده قرار گرفته است. نظر خواهی از معلمان، اولیای دانش آموزان، گروه های آموزشی و در عین حال، استفاده و تحلیل نمرات درسی دانش آموزان از جمله این روشها هستند (اداره نظارت و ارزشیابی، ۱۳۴۹). شاید بتوان گفت که اولین فعالیت های سازمان یافته ارزشیابی در این دوره آغاز و انجام شده است و یا حداقل سندهای موجود مؤید چنین امری هستند. بیشترین فعالیت های ارزشیابی این دوره راه اداره نظارت و ارزشیابی انجام داده است و کمتر می توان مواردی را یافت که محققان آزاد به ارزشیابی برنامه ها اقدام کرده باشند. البته باید دانشجویان دوره های کارشناسی ارشد را جدا کرد. این گروه، طی این دوره فعالیت هایی را برای ارزشیابی برنامه های درسی انجام داده اند (برای مثال نیستی، ۱۳۴۰). روح حاکم بر ارزشیابی در این دوره، کمی نگری است. تمام سعی و تلاش ارزشیابی آن بوده است که نتایج خود را با عدد و رقم نشان دهند (مرکز تحقیقات و برنامه ریزی درسی، ۱۳۵۰). حتی در جایی که مشاهده گر، به مشاهده پرداخته است و یا گروه های آموزشی به بحث نشسته اند، در نهایت نتایج در قالب های کمی ارائه شده است (مرکز تحقیقات و برنامه ریزی درسی، ۱۳۵۳). تلاش برای انتشار کتاب ها و مقاله هایی که به افزایش آگاهی مردم از روش ها و اهمیت ارزشیابی کمک می کرد، در این دوره قابل توجه است. تهیه و انتشار کتاب های "هوش و خرد" (سیاسی، ۱۳۴۱)؛ "اصول علمی تهیه، اجرا و استاندارد کردن تست" (اردبیلی، ۱۳۴۵)؛ "آزمون های روانی" (هاشمی و بطحایی، ۱۳۴۵)؛ "روان شناسی شخصیت" (سیاسی، ۱۳۴۹)؛ و "روان شناسی بالینی" (شاملو، ۱۳۵۰)، در امتداد توجه مریبان به آزمون در آموزش و پرورش و انتشار مقاله ها و کتابهایی از قبیل "راهنمای تهیه پرسشنامه" (سپه پور، ۱۳۴۵)؛ "در باره بازرسی و ارزشیابی آموزشی" (نصفت، ۱۳۴۸)؛ "اصول ارزشیابی در آموزش و پرورش" (نصفت، ۱۳۴۸)؛ اندازه گیری و سنجش در روان شناسی و تعلیم و تربیت" (مهندسی، ۱۳۴۹)؛ "ارزشیابی و هدفهای آن" (فروغی، ۱۳۵۱)؛ "نقش نمره و معدل در ارزشیابی انسانها" (سلیمی، ۱۳۵۱)؛ "سؤالهای تستی" (حاجتی و شاهورانی، ۱۳۵۱)؛ ارزشیابی و اندازه گیری در آموزش و پرورش با استفاده از شیوه های آماری" (حقیقی، ۱۳۴۹)؛ "اصول ارزشیابی (مهندسی، ۱۳۵۱) و "یک مدل ارزشیابی از برنامه های درسی در تایلند" (تایلر، ۱۳۵۱)، و برگزاری نشست هایی از قبیل "کنفرانس ارزیابی انقلاب آموزشی (شیراز، ۱۳۴۸) که یکی از کمیسیون های آن به ارزیابی مسایل برنامه ریزی آموزشی

اختصاص یافته بود، سبب شد تا آگاهی عمومی از ارزشیابی وسعت و عمق بیشتری پیدا کند و آشنایی علاقمندان با روش‌های مختلف ارزشیابی بیشتر شود.

۷. دوره حرفه‌ای شدن در غرب (۱۹۷۳ تاکنون)، دوره اقدام در ایران (۱۳۷۰-۱۳۵۲)

حدود سال ۱۹۷۳، رشته ارزشیابی تبلور یافت و به عنوان یک حرفه مجزا، که سازمان یافته و دارای کیفیت مشخص است، سر برآورد. حدّ اعلای ارزشیابی، پژوهش و آزمون است. در حالیکه امروزه رشته ارزشیابی به عنوان یک حرفه پیشرفت قابل توجهی کرده است؛ اما ملاحظه تحول آن در دوره پیشین آموزنده است. در آن زمان ارزشیابها هویت نقاد داشتند. آنها دارای شغل خاصی نبودند و باید به عنوان پژوهشگر، آزمونگر، مدیر، معلم یا فیلسوف فعالیت می‌کردند. گرچه آنها می‌بایست اطلاعاتی را کسب می‌کردند؛ اما کیفیت خاصی مورد نظر نبود. هیچ سازمان حرفه‌ای متولی امر ارزشیابی نبود؛ هیچ نشریه خاصی منتشر نمی‌شد؛ گرچه هر یک از ارزشیابها باید برای کار خود اطلاعاتی را کسب می‌کردند. بطور کلی هیچ پیشینه‌ای در باره ارزشیابی برنامه بجز مقاله‌های منتشر شده‌ای که به طور اتفاقی نشر یافته بودند، وجود نداشت. البته فعالیت‌های محدودی توسط مجریان انجام می‌شد. فرصت‌های معدودی پیش از خدمت و در ضمن آن برای ارزشیابی وجود داشت. با فاصله به نسبت زیاد، معیارهایی تهیه می‌شد تا بر آن اساس، آزمون‌های تربیتی و روان‌شناختی ساخته شوند. رشته ارزشیابی هنوز شکل نایافته بود. برخی ارزشیابی‌ها توسط افراد فاقد آموزش انجام می‌شد و برخی را روش‌شناسانی انجام می‌دادند که سعی می‌کردند، روش‌هایشان را با ارزشیابی برنامه بطور موفقیت آمیزی انطباق دهند (گویا،^{۴۶} ۱۹۶۶). مطالعات ارزشیابی با سر درگمی، اضطراب و دشمنی همراه بود. ارزشیابی به عنوان یک رشته، منزلت اندکی داشت و فاقد اهمیت سیاسی بود.

در مقابل پیشرفت‌هایی که توسط ارزشیاب‌های تربیتی، برای حرفه‌ای شدن این رشته، حاصل شد؛ طی دهه ۱۹۷۰، کاملاً قابل توجه است. تعدادی نشریه شامل ارزشیابی تربیتی و تحلیل سیاست، مطالعات ارزشیابی، مرور ارزشیابی، توصیه‌های نوین برای ارزشیابی برنامه، ارزشیابی و برنامه‌ریزی برنامه، و خبرهای ارزشیابی برای نشر اطلاعات در باره واقعیت‌های ارزشیابی برنامه منتشر گردید. بیش از سال‌های پیش کتاب‌ها و تکنوگاری‌هایی که بطور عمیق در باره ارزشیابی بحث می‌کردند، نشر می‌یافت. در این زمان، دیگر مسأله روز، فقدان پیشینه برای ارزشیابی نبود؛ بلکه مسأله جمع کردن پیشینه بود (مدوس، استافل بیم و اسکرین، ۱۹۸۸).

برخی دانشگاه‌ها به برگزاری دوره‌هایی، حداقل یک دوره، در باره روش‌شناسی ارزشیابی، به عنوان جزئی از روش‌شناختی تحقیق اقدام کردند، تعدادی از دانشگاه‌ها، از قبیل ایلینویز، استانفورد، کالج بوستون، UCLA، مینه سوتا و میشیگان غربی در ایالات متحده، برنامه‌های تکمیلی^{۴۷} در ارزشیابی را طراحی کردند. دانشگاه نوآدا برای اولین بار، یک دوره ارزشیابی را در مقطع دکتری ارائه کرد. وزیر آموزش و پرورش آمریکا مسؤلیت یک برنامه ملی آموزش ضمن خدمت هفت ساله در ارزشیابی را، برای مریبان ویژه تقبل کرد (برینکر هاف، ۱۹۸۸) و چندین سازمان حرفه‌ای، کارگاه‌ها و دوره‌هایی را در باره عناوین متنوع ارزشیابی عرضه کردند. در این دوره، مراکزی برای تحقیق و توسعه در خصوص ارزشیابی تأسیس شد که از آن جمله می‌توان به واحد ارزشیابی آزمایشگاه آموزش ناحیه شمال غربی، مرکزی برای مطالعه ارزشیابی در UCLA، کنسرسیون (مرکز) ارزشیابی استانفورد، مرکزی برای تحقیق آموزشی و ارزشیابی برنامه در دانشگاه ایلینویز، مرکز ارزشیابی دانشگاه میشیگان غربی، مرکزی برای مطالعه سنجش، ارزشیابی و سیاست در کالج بوستون اشاره کرد. مؤسسه ارزشیابی دانشگاه سان فرانسیسکو بطور خلاصه ارزشیابی بیرون از حوزه محصول و کارکنان را توسعه داد. ایالت لوئیزیانا، سیاست و برنامه‌ای را برای تعیین ارزشیابی‌ها (پیک،^{۴۸} ۱۹۸۱) تعیین کرد و ایالت ماساچوست بطور همزمان کار بر روی برنامه‌های مشابهی برای ارزشیابی همزمان را آغاز کرد. جانسون (۱۹۸۰)، برای اولین بار طرح تهیه راهنمایی برای ارزشیابی و نمایندگی‌های ارزشیابی را مطرح نمود. بطور فزاینده‌ای رشته ارزشیابی قابلیت فرا ارزشیابی^{۴۹} را بدست آورد (اسکرین، ۱۹۷۵؛ استافل بیم، ۱۹۷۸) تا به عنوان وسیله‌ای برای کنترل کیفیت ارزشیابی بکار رود.

طی این دوره ارزشیابها بطور فزاینده‌ای متوجه این امر شدند که روش‌های ارزشیابی باید نتایجی را بدست دهد که در درجه اول، پاسخگوی نیازمندی‌ها و انتظارات عمومی از ارزشیابی باشد؛ و این در حالی بود که رشته ارزشیابی، هنوز در حال تحول بود و روش‌شناسی عملکرد آن برای پاسخگویی به همه نیازمندی‌ها، تکمیل نشده بود. برخی از خواسته‌ها را الگوهای طراحی شده جواب می‌گفتند. این الگوها شامل ارزشیابی هدف آزاد (اسکریون، ۱۹۷۵ و ایورس،^{۵۰} ۱۹۸۰) الگوی مبارزه جویی - مدافعه‌ای (استیک و جرد، ۱۹۷۴) مدافعه‌ای (رینهارد،^{۵۱} ۱۹۷۲)، فراتحلیلی (کلاس، ۱۹۷۶)، ارزشیابی پاسخ دهنده (استیک، ۱۹۷۵) و ارزشیابی طبیعت‌گرا (گوبا و لینکلن،^{۵۲} ۱۹۸۱) بودند. آیزنر (۱۹۷۵) و دانشجوینش روش‌هایی را برای بکارگیری روش‌های مورد استفاده در نقد و ارزشیابی مواد هنری طراحی کردند. ویستر (۱۹۷۵) و دانشجوینش الگوی سیپ (CIPP) را که توسط استافل بیم مطرح شده بود؛ برای ارزشیابی مدارس محلی بکار بستند. استیک (۱۹۷۶)، روش مطالعه موردی را برای استفاده در ارزشیابی آماده کرد. روث^{۵۳} (۱۹۷۷ و ۱۹۷۸)؛ سوارز، (۱۹۸۰)؛ اسکریون و روث (۱۹۷۷)؛ استافل بیم و دیگران (۱۹۷۱) فعالیت‌هایی را در هر دو حوزه مفهوم پردازی و عمل کردن آغاز کردند.

این توسعه حرفه‌ای قابل توجه در ارزشیابی، نتایج پیچیده‌ای را به بار آورد. نخست آنکه سازمان یافتگی رشته را بیش از پیش در پی داشت و بلوغ آن را موجب شد (کرنباخ، ۱۹۸۰). دوم آنکه امکان آموزش ارزشیابها را فراهم آورد و آنها را به حرفه خود مطمئن ساخت و موجب اطمینان به مؤسسه‌هایی شد که خدماتی را برای افزایش کیفیت افراد ارزشیاب ارایه می‌کردند (استیک، ۱۹۸۱). سوم آنکه بین مؤسسه‌های حرفه‌ای درگیر در ارزشیابی آموزش، همکاری پدید آورد و آنها را به رعایت برخی معیارهای عمومی فرا خواند. در نهایت آنکه اطمینان عمومی به مؤسسه‌های حرفه‌ای جدید که برای انجام مأموریت‌های ارزشیابی پدید آمده بودند، افزایش یافت.

دل‌نگرانی عمده در این دوره، تعارض دیدگاه‌ها در ارزشیابی بود. گرچه ارتباط چشمگیر و رو به افزایش رویکردهای مثبت‌گرا - کمی و روش‌شناسی - کیفی وجود دارد؛ اما در عین حال نباید از خطر تفکیک بین این دیدگاه‌ها غافل ماند. ریشه این تضاد را نمی‌توان در روش‌شناسی یافت؛ بلکه باید در تفاوت‌های عقیدتی جستجو کرد. مدوس و مک‌دونگ (۱۹۸۲)، خطر این تضاد را بدین شرح توصیف کرده‌اند:

«در دو صورت، ممکن است، ارزشیابها به یک طبقه روحانی بدل شوند و هشدار دهند و توصیه کنند اما هشدار و توصیه نپذیرند. یکی اینکه آنها تحت نام علم شروع به سخنرانی و موعظه نمایند و دیگر اینکه آنها تحت نام شخصیت‌های برجسته بدین کار مبادرت ورزند» (ص: ۳۶).

بنابراین لازم است در کنار افزایش تلاش برای جستجوی روش‌های مناسب ارزشیابی به افزایش ارتباط و انسجام و درک متقابل از طریق هدایت و رهبری روش‌شناسانه و تحول روش‌های جدید اقدام شود. در عین حال نباید فراموش کرد که تربیت ارزشیابهایی که روش‌های جدید را مورد استفاده قرار دهند و روش‌های سنتی را تحول بخشند؛ نیازمند تلاش چشمگیر و سعی وافر است. البته لازم است، در همه این کوشش‌ها بر تهته یک روش‌شناسی که با نیازهای اجتماعی، نیازهای مؤسسه‌ها و نیازهای شهروندان انطباق داشته باشد، تأکید شود.

در این دوره (۱۳۷۰-۱۳۵۲) آموزش و پرورش ایران، شاهد رشد قابل توجه پژوهش‌های ارزشیابی بوده است. تغییر نظام آموزشی در انتهای دهه ۱۳۴۰، زمینه را برای انجام ارزشیابی‌های رسمی و غیررسمی در دهه ۱۳۵۰ فراهم آورد. موافقان و مخالفان طرح جدید آموزش و پرورش، با انگیزه‌های متفاوت برای ارزشیابی اقدام کردند و از روش‌های متفاوتی برای این کار سود جستند. ویژگی‌های اساسی پژوهش‌های ارزشیابی در این دوره را می‌توان بدین شرح برشمرد:

۱.۷. افزایش کمی

شاید بتوان ادعا کرد که پژوهش‌های ارزشیابی در نظام آموزش و پرورش ایران، با تغییر نظام آموزشی در سال ۱۳۴۴ پیوند دارند. از این زمان به بعد فعالیت‌های مختلفی برای ارزشیابی از برنامه‌های آموزشی به عمل آمده است که کم و بیش دارای شرایط یک مطالعه ارزشیابی می‌باشند. اسناد موجود نشان می‌دهند که دویست و هفت مورد پژوهش ارزشیابی تا نیمه سال ۱۳۷۵، در باره موضوعات مختلف آموزشی انجام شده است و بیش از نود و پنج درصد آنها، بعد از

سال ۱۳۵۰، به اتمام رسیده‌اند (بانک اطلاعاتی علوم تربیتی). این رشد کمی تا حدی ناشی از آن بوده است که برای انجام صحیح‌تر فعالیت‌های آموزشی در نظام جدید، سرمایه‌گذاری شده بود و تا حدی نیز ریشه در حضور صاحب‌نظران تعلیم و تربیت در عرصه فعالیت‌های آموزش و پرورش داشته است.

۲.۷. جهت‌گیری کمی (کمی‌نگری)

پژوهش‌های ارزشیابی در این دوره، بطور عمده مبتنی بر رویکرد (تجربی - کمی) بوده‌اند. در اکثر ارزشیابی‌های انجام شده، سعی شده است با اعداد و ارقام، تبیینی از مسأله ارائه شود (دفتر تحقیقات و برنامه‌ریزی درسی، ۱۳۵۱، ۱۳۵۲ الف؛ ۱۳۵۲، ب؛ ۱۳۵۴؛ مرکز برنامه‌ریزی آموزشی، ۱۳۵۳؛ صدرالسادات، ۱۳۵۴؛ میرجلیلی، ۱۳۵۹؛ نصف، ۱۳۶۵؛ و سحرخیز، ۱۳۷۰) و حتی گاه پژوهش‌گران سعی کرده‌اند اطلاعاتی را که با روش‌های کیفی به دست آورده‌اند، به نحوی به کمیت تبدیل نمایند و با اعداد و ارقام گزارش دهند (دفتر تحقیقات و برنامه‌ریزی درسی، ۱۳۵۲ الف و ب و ج).

۳.۷. یکسان‌انگاری پژوهش و ارزشیابی

پژوهش‌ها معمولاً توسط کسانی انجام شده است که در ارزشیابی تخصصی نداشته‌اند و ارزشیابی را از زاویه عمومی تحقیق، می‌دیده‌اند. به همین علت است که بسیاری از ارزشیابی‌های انجام شده، طریق معمول در پژوهش‌های تربیتی را پیموده‌اند و به این مسأله که بین ارزشیابی و تحقیق، به معنای عام آن، تفاوتی وجود دارد؛ توجه نداشته‌اند (برای نمونه به گزارش‌های ارزشیابی که توسط دفتر تحقیقات و برنامه‌ریزی درسی در سال‌های ۱۳۵۱، ۱۳۵۲، ۱۳۵۳ و ۱۳۵۴ منتشر شده است، نظر کنید). شاید همین درهم آمیختگی است که سبب شده است در گزارش‌های ارزشیابی موجود، به ارابه مجموعه‌ای از اعداد و ارقام بسنده شود.

پیامد دیگر عدم تفکیک بین ارزشیابی و پژوهش، بی‌توجهی به الگوهای ارزشیابی بوده است. در گزارش‌های موجود بندرت می‌توان الگوی مورد استفاده برای ارزشیابی را مشخص کرد. حتی در مواردی که پژوهشگر بیان کرده است که از الگوی خاصی استفاده می‌کند (میرجلیلی، ۱۳۵۹؛ بطحایی و توحیدی‌فر، ۱۳۷۴) چنین ادعایی به خوبی دنبال نشده است. برای مثال، بطحایی و توحیدی‌فر (۱۳۷۴) بیان کرده‌اند که از ارزشیابی مبتنی بر الگوی قضاوت (ص: ۷) استفاده می‌کنند؛ در حالیکه پس از گردآوری اطلاعات، با استفاده از پرسشنامه، نتایج را با استفاده از آزمون‌های آماری تحلیل کرده و به صورت کمی گزارش نموده‌اند.

۴.۷. تمرکز بر الگوی هدف - محور

بیشتر ارزشیابی‌های انجام شده طی این دوره، چنان که گفته شد، فاقد الگوی مشخص بوده‌اند و ویژگی‌های یک پژوهش ارزشیابی را در آنها نمی‌توان مشاهده کرد. در معدود مواردی که می‌توان آنها را پژوهش ارزشیابی نام داد؛ معمولاً هدف‌ها محور ارزشیابی بوده‌اند و تلاش شده است از میزان انطباق فعالیت‌ها با هدف‌ها و توفیق برنامه در تحقق بخشیدن هدف‌ها خبر دهند. تقریباً تمامی گزارش‌هایی که توسط دفتر تحقیقات و برنامه‌ریزی درسی طی دهه ۱۳۵۰ منتشر شده‌اند، دارای این ویژگی هستند.

۵.۷. توجه به زمینه‌های نظری و مفهوم‌پردازی

بحث در باره ارزشیابی به شکل جدی در این دوره و بخصوص در دهه ۱۳۶۰ و ۱۳۷۰، آغاز شده است. ترجمه و تألیف کتاب‌هایی در باره ارزشیابی و انتشار مقاله‌هایی در نشریه‌های تخصصی در این دوره رواج بیشتری یافت. بخصوص حضور برخی از تحصیل‌کردگان رشته ارزشیابی در این دوره زمینه را برای تشریح و تبیین نظریه‌های ارزشیابی فراهم آورد. در عین حال، ورود درس‌های مربوط به ارزشیابی در دوره‌های دانشگاهی، آشنایی دانشجویان با ارزشیابی و نظریه‌های آن را گسترش بخشید. گرچه پس از حدود دو دهه فعالیت‌های جدی در حوزه ارزشیابی آموزشی، هنوز تعداد کتاب‌هایی که در زبان فارسی وجود دارند، از تعداد انگشتان دست تجاوز نمی‌کند و اکثر این نوشته‌ها نیز به تبیین دیدگاه‌های صاحب‌نظران خارجی اختصاص یافته است و کمتر می‌توان حاصلی از اندیشه و ابداع نویسندگان ایرانی را در آنها جستجو کرد.

در این دوره، کتاب‌ها و مقاله‌هایی تألیف و یا ترجمه شد که بطور خاص به موضوع ارزشیابی و حوزه‌های وابسته به آن اختصاص داشتند و دانش تخصصی افراد علاقمند را بهبود بخشیدند. در میان این آثار، انتشار کتاب‌های «سنجش و اندازه‌گیری در آموزش و پرورش روان‌شناسی» (شریفی، ۱۳۵۲)؛ «تهیه آزمون‌های پیشرفت تحصیلی» (شوئر، ۱۳۵۳)؛ «مبانی نظری آزمون‌های روانی» (مگنسون، ۱۳۷۰)؛ «ارزیابی آموزشی و کاربرد آن در سواد آموزی تابعی» (بولا، ۱۳۶۲)؛ «اندازه‌گیری و ارزشیابی پیشرفت تحصیلی» (سیف، ۱۳۶۷)؛ «روان‌سنجی کاربردی» (ثراندیک، ۱۳۶۶)؛ «روش‌های ارزیابی آموزشی» (آذر، ۱۳۶۹) زمینه را برای آشنایی علاقمندان به ارزشیابی و مبانی نظری آن فراهم کرد و برخی از الگوهای ارزشیابی را برای بکار بستن در اختیار آنها نهاد.

۸. دوره امیدواری در ایران (۱۳۷۰ تاکنون)

از ابتدای دهه هفتاد شمسی در ایران، علاوه بر تلاش‌های نسبتاً موفقیت‌آمیز صاحب‌نظران برای شناساندن مبانی نظری و اهمیت عملی ارزشیابی در حوزه‌های آموزشی و در عین حال، اقدام به ارزشیابی و تربیت نیروی انسانی از سوی دست‌اندرکاران آموزش، کوشش‌هایی از جانب برخی علاقمندان و متخصصان ارزشیابی آموزشی در ایران به عمل آمده است تا الگوهای برای ارزشیابی برنامه آموزشی طراحی شود که با شرایط اجتماعی و فرهنگی ایران همخوان و منطبق باشد. چنین وضعیتی کم و بیش شرایط را برای اندیشه‌ورزی فراهم کرده است و بارقه‌هایی از امید به ابداع و ابتکار در حوزه ارزشیابی آموزشی را برافروخته است. ویژگی‌های اساسی این دوره را می‌توان بدین شرح خلاصه کرد:

۱. ۸. تلاش برای فراهم کردن دانش سازمان یافته و تخصصی در حوزه ارزشیابی

انتشار گزارش‌های ارزشیابی در سطح عمومی و تهیه و انتشار کتاب‌هایی از قبیل «ارزشیابی آموزشی؛ مبانی سنجش توانایی و بررسی برنامه» (ولف، ۱۳۷۱)؛ «اندازه‌گیری و آزمون در تعلیم و تربیت» (واپرزما، ۱۳۷۲)؛ «روش‌های ارزشیابی آموزشی» (کیامنش، ۱۳۷۳)؛ «زمینه ارزشیابی برنامه‌های آموزشی» (هومن، ۱۳۷۵)؛ «روش‌های اندازه‌گیری و ارزشیابی آموزشی» (سیف، ۱۳۷۵) و «ارزشیابی طرح‌ها و برنامه‌های آموزشی برای توسعه» (بولا، ۱۳۷۵) توانسته است سطح دانش ارزشیابی را در بین علاقمندان اعتلا بخشد. طی این مدت همچنین دهها مقاله در باره ارزشیابی آموزشی در نشریه‌ها و مجموعه مقالات همایش‌های علوم تربیتی (برای مثال سیف، ۱۳۷۶؛ بازرگان، ۱۳۷۴؛ احمدی، ۱۳۷۴؛ ابیلی، ۱۳۷۶) منتشر شده است و پایان‌نامه‌های متعددی (برای مثال مشفق آرانی و ایبسی، ۱۳۷۲؛ شفیع‌ی، ۱۳۷۴؛ اعظمی، ۱۳۷۴؛ موسی‌پور، ۱۳۷۶)، به پژوهش‌های ارزشیابی اختصاص یافته است. همه این فعالیت‌ها در جهت افزایش معرفت نظری در حوزه ارزشیابی مؤثر بوده‌اند. بر همین اساس می‌توان با مدوس، استافل بیم و اسکریون (۱۹۸۸) هم صدا شد که مسأله اساسی در این دوره، کمبود پیشینه برای ارزشیابی نیست؛ بلکه مسأله فرصت یافتن برای استفاده از آنهاست.

۲. ۸. توجه به آموزش‌های تخصصی

در این دوره ضرورت تربیت کسانی که دارای دانش تخصصی در حوزه ارزشیابی باشند؛ بیشتر احساس شد و پذیرش دانشجو در مقطع کارشناسی ارشد در رشته سنجش و اندازه‌گیری از سر گرفته شد (دانشگاه علامه طباطبایی در سال ۱۳۷۱ تعداد ۹ نفر دانشجو را برای تحصیل در این رشته پذیرفت که این امر طی سال‌های بعد استمرار داشته است) و در عین حال، برخی مؤسسه‌ها که نیاز به ارزشیابی برنامه‌های خود داشتند، دوره‌های کوتاه مدت برای تربیت ارزشیاب تدارک دیدند. تشکیل کارگاه‌های آموزش ارزشیابی در دانشگاه‌ها نیز در همین قالب قابل طرح است.

۳. ۸. توجه به ارزشیابی برنامه‌های آموزشی

طی این دوره بسیاری از مؤسسات آموزشی ارزشیابی برنامه را لازم تشخیص دادند و سعی کردند به نحوی به آن مبادرت ورزند. علاوه بر آنکه در سازمان پژوهش و برنامه‌ریزی درسی وزارت آموزش و پرورش، دفتری کار نظارت و ارزشیابی را بر عهده گرفت و در وزارت فرهنگ و آموزش عالی نیز دفتر مشابهی شروع به کار کرد (وزارت فرهنگ و آموزش عالی، ۱۳۷۲)، تمامی دانشگاه‌های کشور نیز به تأسیس دفتری برای نظارت و ارزیابی مبادرت ورزیدند. یکی از وظایف عمده این دفاتر، ارزشیابی مستمر برنامه‌های درسی بوده است؛ هر چند که به دلایلی، این کار در عمل توفیق

چندانی کسب نکرده است. نمونه قابل تأمل در اینجا تلاش برای ارزشیابی نظام جدید آموزش متوسطه جمهوری اسلامی (کیامنش و همکاران، ۱۳۷۲ الف؛ ۱۳۷۲ ب؛ ۱۳۷۳؛ ۱۳۷۴ و ۱۳۷۵ الف و ب) است که با تأکید و همکاری مقامات رسمی وزارت آموزش و پرورش انجام یافته است.

۴.۸. تلاش برای تهیه الگوهای ارزشیابی

طی این دوره از سوی برخی متخصصان و دست‌اندرکاران ارزشیابی تلاش شده است تا الگوهایی برای ارزشیابی برنامه‌های آموزشی و درسی طراحی و عرضه شود (احمدی و همکاران، ۱۳۷۴؛ بقایی، ۱۳۷۴؛ حسینی نسب، ۱۳۷۴؛ بازرگان، ۱۳۷۶؛ سیف، ۱۳۷۶؛ موسی پور، ۱۳۷۶) و یا در اقدام به ارزشیابی، الگوهای خاص بکار گرفته شود (کیامنش و همکاران، ۱۳۷۲؛ ۱۳۷۳؛ ۱۳۷۴ و ۱۳۷۵). این تلاش‌ها گرچه با موفقیت کامل همراه نبوده است، اما از دو جهت حایز اهمیت هستند. نخست آنکه ضرورت تهیه الگوهای متناسب با فرهنگ ایرانی را نشان می‌دهند و ضرورت بومی کردن الگوهای ارزشیابی را با خود دارند و دوم آنکه از نوعی خود باوری در عرصه تولید دانش خبر می‌دهند و تاریخ علم نشان می‌دهد که اعتلای علمی حاصل جرأت در تولید دانش است.

۹. جمع بندی و پیشنهاد

بررسی اسناد و مدارک نشان می‌دهد که طی دو قرن گذشته، ارزشیابی آموزشی در ایران، هفت مرحله را تجربه کرده و جریان حرکت آن اعتلایی و رو به بهبود بوده است که عناوین دوره‌های مذکور، دوره زمانی و خصایص عمده هر یک از آنها در جدول زیر ارائه می‌شود:

جدول ۲: تحول ارزشیابی آموزشی در ایران

نام دوره	گستره زمانی	خصایص اصلی
دوره بی نیازی	۱۱۷۹-۱۲۷۹	مقبولیت مبتنی بر برداشت شخصی و غیررسمی و نظام نایافته
دوره آگاهی	۱۲۸۰-۱۳۰۹	تحدید امور آموزش به دولت مرکزی و بررسی مراکز آموزشی در قالب مقررات مصوب
دوره تلاش رسمی	۱۳۱۰-۱۳۲۴	عنایت دولت به محتوای آموزش مدارس و تلاش در جهت بررسی و اطلاع از کم و کیف آن و احتمالاً اعمال نظر
دوره دشواری	۱۳۲۵-۱۳۳۶	تلاش دولت برای اعمال نظر در محتوای آموزش مدارس و حفظ ظاهر منطقی آن بدون وجود ملاک و معیار و در عین حال توجه به جنبه‌های مختلف آموزش
دوره علاقمندی	۱۳۳۷-۱۳۵۱	عنایت صاحب‌نظران و نهادهای ذیربط دولتی به بررسی و اظهار نظر درباره نظام آموزشی و برنامه‌های آن، افزایش آگاهی از مفهوم ارزشیابی آموزشی، انتشار مطالبی در باره ارزشیابی آموزشی
دوره اقدام	۱۳۵۲-۱۳۷۰	افزایش کمی، جهت‌گیری کمی در ارزشیابی، یکسان‌انگاری پژوهش و ارزشیابی، تمرکز بر الگوی هدف - محور در ارزشیابی، توجه به زمینه‌های نظری و مفهوم پردازی
دوره امیدواری	۱۳۷۰ تاکنون	تلاش برای فراهم کردن دانش سازمان یافته و تخصصی، توجه به آموزش‌های تخصصی، توجه به ارزشیابی برنامه‌های آموزشی، تلاش برای تهیه الگوهای ارزشیابی آموزشی

بررسی به عمل آمده همچنین نشان می‌دهد که در ارزشیابی آموزشی ایران، چند موضوع اصلی قابل تأمل

هستند:

اول آنکه آگاهی و پرداختن به ارزشیابی آموزشی، در قالب مفهوم جدید آن، عمر اندکی دارد و به واقع از دهه چهل قرن حاضر هجری شمسی آغاز شده است. گرچه نشانه‌هایی می‌توان یافت که آگاهی از ارزشیابی آموزشی را دارای سابقه بلندتری نشان می‌دهد، اما آگاهی توأم با عمل با تغییر نظام آموزشی در دهه چهل آمیخته است. پیامد عمده این امر، کمبود دانش تولید شده در این حوزه است و چنین کمبودی ارزشیابان آموزشی را با مشکل مواجه می‌سازد. حتی می‌توان گفت که عمر کوتاه ارزشیابی آموزشی فرصت لازم برای انتقال دانش تخصصی را نیز فراهم نیاورده است.

کتاب‌های معدود، مقاله‌های نادر و فقدان تشکیلات برای ایجاد ارتباط بین متخصصان و عمل‌کنندگان این حوزه، از واقعیت‌های حوزه ارزشیابی آموزشی در ایران هستند که پیامدهای ناخوشایندی را به دنبال داشته است. بر این اساس پیشنهاد می‌شود که متخصصان این حوزه از معارف بشری، تلاش نمایند، ابتدا به انتقال تجارب و یافته‌های علمی در این زمینه بپردازند و سپس همت خود را صرف تولید دانش مبتنی بر نیازهای محلی و فرهنگی ایران نمایند و به آرایه‌های گوناگونی برای ارزشیابی آموزشی در این مرز و بوم اقدام کنند. البته برای مهیا ساختن زمینه‌های تولید دانش، تشکیل انجمن ارزشیابان آموزشی (یا ارزشیابی آموزشی) و انتشار نشریات تخصصی در این زمینه ضرورت دارد.

دوم آنکه فعالیت ارزشیابی آموزشی در ایران با اقدام دولت همراه بوده است. نخستین گام‌های ارزشیابی آموزشی توسط وزارت فرهنگ و به منظور بررسی کتاب‌های درسی برداشته شده است. این امر از سویی این تصور را پدید آورد که گویی دولت تنها مسؤول چنین امری است و جایی برای دخالت و اقدام هیچ فرد یا گروهی وجود ندارد و از سوی دیگر، نگرش خاصی به ارزشیابی را به دنبال داشته است که گزارش‌های ارزشیابی به نوعی توجیه‌گر عملکرد مجریان قلمداد گردند. در عین حال، پیامد نامطلوب دیگر چنین اتفاقی، پدیدآیی رویکرد خاصی به ارزشیابی آموزشی است که مانع از بروز خلاق نظریات و الگوها شده است.

بر این اساس، انتظار می‌رود زمینه‌های مشارکت مردم و مراکز علمی در ارزشیابی آموزشی فراهم آید و ضمن ایجاد حساسیت در مردم، نوعی مسؤولیت در مراکز علمی برای اقدام به ارزشیابی پدید آید. این حساسیت و مسؤولیت با دو مسأله آگاهی از حقوق (برای مردم) و تسهیل انجام عمل (برای محققان) فراهم می‌شود.

سوم آنکه، فعالیت‌های ارزشیابی آموزشی جدا از یکدیگر انجام شده‌اند. از تجارب گذشته کمتر در اقدام‌های بعدی بهره گرفته شده است و در عین حال حتی بین تحقیقاتی که در یک دوره زمانی به عمل آمده‌اند، ارتباط لازم دیده نمی‌شود. بازنگری تجربه‌های گذشته علاوه بر اینکه زمینه را برای پیشگیری از تکرار خطاهای قبلی فراهم می‌کند، در عین حال، امکان خلاقیت را افزون می‌سازد. قریب به یک قرن است که آموزش نوین در ایران در حال قانون‌مند شدن است و طی این دوره، حداقل نیمی از آن با بحث ارزشیابی آموزشی توأم بوده است. لازم است که تجربه گذشته شناسایی و شناسانده شود، تا هم از پیشرفت کار آگاهی حاصل شود و هم زمینه برای اثر بخش ساختن فعالیت‌های آینده فراهم آید.

چهارم اینکه، ارزشیابی آموزشی حرکت رو به بهبود طی کرده است. از نیمه دوم دهه چهل هجری شمسی و با انتشار کتاب "اصول ارزشیابی در آموزش و پرورش" (نصفت، ۱۳۴۸)؛ نه تنها تاکنون چندین کتاب ترجمه و تألیف شده و مقاله‌هایی به جامعه آرایه گردیده و تحقیقاتی به عمل آمده و افرادی در این رشته تحصیل کرده‌اند؛ بلکه تلاش شده است تا الگوهایی برای ارزشیابی آموزشی در ایران طراحی گردد که متناسب با شرایط اجتماعی و فرهنگی ایران باشد. این حرکت هم از آن جهت که نوعی خودباوری را نشان می‌دهد و هم از آن جهت که واقع‌بینی را در شالوده کار خود دارد، قابل تحسین است. پیشنهاد می‌شود حمایت جدی از این گونه افراد به عمل آید تا هم توجه و توان آنان، و هم سایر افراد توانا، بدین امر معطوف گردد و از این طریق نیازهای اساسی جامعه پاسخ گفته شود.

یادداشتها

- | | |
|---|------------------------------------|
| 1. Evaluation | 9. The Tylerian Age |
| 2. Cronbach, L.J. | 10. The Age of Innocence |
| 3. Popham, W.Y. | 11. The Age of Expansion |
| 4. Stake, R.E. | 12. The Age of Professionalization |
| 5. House, E.R. | 13. Joseph Rice |
| 6- Madaus, G.F & Stufflebeam, D.& Scriven, M.S. | 14. Lindquist |
| 7. The Age of Reform | 15. Campbell |
| 8. The Age of Efficiency and Testing | 16. Accreditation |

- | | |
|--|------------------------------------|
| 17. Professional - Judgment Approach | 35. Actual Outcomes |
| 18. Adequacy | 36. Objectives |
| 19. Biddle, B.J. and Ellena, W.Y. | 37. Educational Testing Service |
| 20. Cremin, L.A. | 38. National Defense Education Act |
| 21. National Society For The Study of Education (NSSE) | 39. Phi Delta Kappa |
| 22. Objective-Referenced | 40. Seized With Great Illness |
| 23. Ballou, F.A. | 41. Hammond |
| 24. Norm-Referenced | 42. Metfessel and Michael |
| 25. Tyack, D. and Hansot, E. | 43. Glaser, R. |
| 26. May | 44. Criterion Referenced |
| 27. Testing | 45. Norm-Referenced |
| 28. Great Depression | 46. Guba, E.G. |
| 29. New Deal | 47. Graduate |
| 30. Renew | 48. Pick |
| 31. Progressive Education Movement | 49. Meta Evaluation |
| 32. Pragmatism | 50. Evers, J.W. |
| 33. Behavioristic Psychology | 51. Reinhard, D.L. |
| 34. Intended Outcomes | 52. Lincoln, Y.S. |
| | 53. Roth, Y. |

منابع

الف: فارسی

- ابیلی، خدایار. (۱۳۷۶). *ارزشیابی ماهیت نظارت و ارزیابی جاری در آموزش عالی کشور، مجموعه مقالات نخستین سمینار آموزش عالی در ایران*، ج: ۲، تهران: دانشگاه علامه طباطبایی، چاپ اول، ۳۴۵-۳۳۱.
- اداره کل تعلیمات ابتدایی کشور. (۱۳۳۸). *اصول کلی برنامه تحصیلات شش ساله ابتدایی*، تهران: وزارت فرهنگ، چاپ پدرام، اسفند.
- اداره کل نگارش. (۱۳۱۷). *برنامه تحصیلات ابتدایی، سالنامه و آمار سال‌های ۱۳۱۹ تا ۱۳۲۲*، تهران: وزارت فرهنگ، ۴۴۴-۴۲۰.
- اداره کل نگارش. (۱۳۲۲). *برنامه تحصیلات ابتدایی، سالنامه و آمار سال‌های ۱۳۱۹ تا ۱۳۲۲*، تهران: وزارت فرهنگ.
- اداره نظارت و ارزشیابی. (۱۳۴۹). *ارزشیابی برنامه های درسی سال چهارم ابتدایی*، تهران: مرکز تحقیقات و برنامه ریزی آموزشی.
- آذر، رسول. (۱۳۶۹). *روشهای ارزیابی آموزشی*، تبریز: نیما.
- اردبیلی، یوسف. (۱۳۴۵). *اصول علمی تهیه، اجرا و استاندارد کردن تست*، تهران: سازمان تربیت معلم و تحقیقات تربیتی.
- اردبیلی، یوسف. (۱۳۷۶). *روشهای ارزشیابی در بخش های دولتی و غیر دولتی ایران*، تهران: بعثت.
- احمدی، غلامعلی و دیگران. (۱۳۷۴). *ارایه یک مدل ارزشیابی کاربردی*، تهران: دبیرخانه شورای تحقیقات آموزشی وزارت آموزش و پرورش.
- احمدی، غلامعلی. (۱۳۷۴). *جایگاه ارزشیابی در فرآیند برنامه ریزی درسی، همایش علمی-کاربردی بهبود کیفیت آموزش عمومی*، تهران: دبیرخانه شورای تحقیقات آموزشی وزارت آموزش و پرورش.

اعظمی، موسی. (۱۳۷۴). ارزشیابی مقایسه KASA (دانش، نگرش، مهارت و آرزوها) کشاورزان تحت پوشش طرح محوری گندم و کشاورزان خارج از طرح شهرستان سنقر و کلباتی در سال زراعی ۷۴-۱۳۷۳، رساله کارشناسی ارشد، تهران: دانشگاه تربیت مدرس.

بازرگان، عباس. (۱۳۷۴). ارزیابی درون دانشگاهی و کاربرد آن در بهبود مستمر کیفیت آموزش عالی، فصلنامه پژوهش و برنامه‌ریزی در آموزش عالی، ۱۱ و ۱۲، ۷۰-۴۹.

بازرگان، عباس. (۱۳۷۶). کیفیت و ارزیابی آن در آموزش عالی: نگاهی به تجربه‌های ملی و بین‌المللی، مجموعه مقالات نخستین سمینار آموزش عالی در ایران، ج: ۱، تهران: دانشگاه علامه طباطبایی، چاپ اول، ۲۷-۳.

بطحایی، سید محمد و توحیدی فر، احمد. (۱۳۷۴). ارزشیابی از آزمایشگاههای درس شیمی واحدهای آموزش دوره متوسطه نظری (نظام فعلی)، تهران: شورای تحقیقات آموزش و پرورش.

بقایی شیوا، نعمت. (۱۳۷۴). بررسی و طراحی الگوی ارزیابی مناسب برای اعتبار بخشی دبیرستان‌های نظام جدید، پایان‌نامه کارشناسی ارشد، تهران: دانشگاه تهران.

بولا، ه. س. (۱۳۶۲). ارزیابی آموزشی و کاربرد آن در سوادآموزی تابعی، ترجمه عباس بازرگان، تهران: نشر دانشگاهی، چاپ اول.

بولا، ه. س. (۱۳۷۵). ارزشیابی طرح‌ها و برنامه‌های آموزشی برای توسعه، ترجمه خدیار ایلی، تهران: مؤسسه بین‌المللی روش‌های آموزش بزرگسالان، چاپ اول.

تاریخچه معارف ایران، (۱۳۱۳). مجله تعلیم و تربیت، تهران: سال چهارم، شماره ۹، ۵۳۷-۵۳۰.

تایلر، رالف. (۱۳۵۱). اصول اساسی برنامه‌ریزی درسی و آموزش، ترجمه ابراهیم کظیمی، تهران: مرکز تحقیقات و برنامه‌ریزی درسی، چاپ اول.

ثرن‌دایک، رابرت، ال. (۱۳۶۶). روان‌سنجی کاربردی، ترجمه حیدرعلی هومن. تهران: دانشگاه تهران، چاپ اول.

ثواقب، جهانبخش. (۱۳۶۹). مروری بر تاریخچه مدارس جدید در ایران، روزنامه اطلاعات، آبان ماه، شماره‌های ۱۹۱۷۰ تا ۱۹۱۸۰.

حاجتی، نعیم؛ شاهورانی، کوروش. (۱۳۵۱). سؤالیهای تستی، ماهنامه آموزش و پرورش، ۴۱، ۴.

حسینی نسب، سیدداود. (۱۳۷۴). عوامل مهم و مؤثر در ساختن یک نظام ارزشیابی آموزشی یا در طراحی یک برنامه جامع ارزشیابی، همایش علمی-کاربردی بهبود کیفیت آموزش عمومی، تهران: دبیرخانه شورای تحقیقات آموزشی وزارت آموزش و پرورش.

حقیقی، حسن. (۱۳۴۹). ارزشیابی و اندازه‌گیری در آموزش و پرورش با استفاده از شیوه‌های آماری، تهران: وزارت آموزش و پرورش.

خواجه نصیرالدین طوسی، احمد. (۱۳۵۱). کتاب‌های درسی ایران، ماهنامه آموزش و پرورش، ۴۲، ۳.

دفتر تحقیقات و برنامه‌ریزی درسی. (۱۳۵۱). گزارش مراحل اجرایی فعالیت‌ها و برنامه‌های درسی سال اول مدارس راهنمایی تحصیلی، تهران.

دفتر تحقیقات و برنامه‌ریزی درسی. (۱۳۵۲ الف). ارزشیابی مراحل اجرایی برنامه‌های درسی سال اول راهنمایی تحصیلی، بررسی نتیجه آزمون‌های اردیبهشت ماه، تهران.

- دفتر تحقیقات و برنامه‌ریزی درسی. (۱۳۵۲ ب). ارزشیابی مراحل اجرایی برنامه‌های درسی سال اول راهنمایی، تهران.
- دفتر تحقیقات و برنامه‌ریزی درسی. (۱۳۵۲ ج). گزارش ارزشیابی برنامه‌های درسی سال دوم دوره راهنمایی تحصیلی و مراحل اجرایی آنها در سال تحصیلی ۵۱-۵۲، تهران.
- دفتر تحقیقات و برنامه‌ریزی درسی. (۱۳۵۳). گزارش ارزشیابی برنامه و کتاب‌های درسی پایه سوم دوره پنج ساله ابتدایی، تهران.
- دفتر تحقیقات و برنامه‌ریزی درسی. (۱۳۵۴). گزارش فعالیت‌های سال سوم مدارس راهنمایی تحصیلی، تهران.
- ره آورد، حسن. (۱۳۲۷). *تاریخچه فرهنگ ایران، سالنامه و آمار سال‌های تحصیلی ۱۳۲۷-۱۳۲۲*، تهران: اداره مطالعات و آمار وزارت فرهنگ.
- ریاحی (دکتر). (۱۳۴۲). *داستانی به نام کتاب درسی، ماهنامه آموزش و پرورش*، ۳۳، ۴.
- سپه پور، افسانه. (۱۳۴۵). *راهنمایی تهیه پرسشنامه*، تهران: مؤسسه مطالعات و تحقیقات تربیتی.
- سرخیز پور شیرازی، آمنه. (۱۳۷۰). *ارزشیابی محتوای کتاب ریاضی سال یک راهنمایی بر اساس عملکرد دانش‌آموزان و نظر دبیران در منطقه ۸ تهران در سال تحصیلی ۷۰-۶۹*، رساله کارشناسی ارشد، تهران: دانشگاه تربیت معلم.
- سلطان زاده، حسین. (۱۳۶۴). *تاریخ مدارس ایران از عهد باستان تا تأسیس دارالفنون*، تهران: انتشارات آگاه.
- سلیمی، عباس. (۱۳۵۱). *نقش نمره و معدل در ارزشیابی انسان ها، ماهنامه آموزش و پرورش*، ۴۱، ۱۱.
- سیاسی، علی اکبر. (۱۳۴۱). *هوش و خرد*، تهران: بی نا، چاپ اول.
- سیاسی، علی اکبر. (۱۳۴۹). *روان شناسی شخصیت*، تهران: سیمرغ.
- سیف، علی اکبر. (۱۳۶۷). *اندازه گیری و ارزشیابی پیشرفت تحصیلی*، تهران: نشر آگاه، چاپ اول.
- سیف، علی اکبر. (۱۳۷۵). *روش‌های اندازه‌گیری و ارزشیابی آموزشی*، تهران: نشر دوران، چاپ اول.
- سیف، علی اکبر. (۱۳۷۶). *روش‌های ارزشیابی کیفیت فعالیت‌های علمی - آموزشی اعضای هیأت علمی دانشگاه‌ها، مجموعه مقالات نخستین سمینار آموزش عالی در ایران*، ج: ۲، تهران: دانشگاه علامه طباطبایی، چاپ اول، ۳۸۸-۴۰۶.
- شاملو، سعید. (۱۳۵۰). *روانشناسی بالینی*، تهران: نشر چهر.
- شریعتی، علی. (بی تا). *مکتب: تعلیم و تربیت اسلامی*، تهران: انتشارات امام، چاپ اول.
- شریفی، حسن پاشا. (۱۳۵۲). *سنجش و اندازه‌گیری در آموزش و پرورش و روان‌شناسی*، تهران: بی نا، چاپ اول.
- شفیعی، ابوالقاسم. (۱۳۷۴). *ارزیابی تحلیلی از برنامه‌های آموزش نیروی انسانی و تطبیق آن با خط مشی‌های برنامه اول توسعه ج.ا. در شهرستان گناباد*، رساله کارشناسی ارشد، تهران: دانشگاه تربیت مدرس.
- شوئر، ل. ا. (۱۳۵۳). *تهیه آزمونهای پیشرفت تحصیلی*، ترجمه حمزه گنجی، تهران: بی نا.
- صافی، احمد. (۱۳۷۲). *تربیت معلم در ایران، هند و پاکستان*، مدرسه، تهران: بی نا، چاپ اول.

- صدرالسادات، جلال و دیگران. (۱۳۵۴). *ارزیابی نحوه هدایت تحصیلی دانش‌آموزان در اولین دوره سه ساله راهنمایی با اشاراتی به گذشته راهنمایی، رساله کارشناسی ارشد، تهران: دانشگاه تهران.*
- ضمیری، محمدعلی. (۱۳۷۱). *تاریخ آموزش و پرورش ایران و اسلام، شیراز: راهگشا.*
- فروغی، دانش. (۱۳۵۱). *ارزشیابی و هدف‌های آن، نشریه علوم تربیتی دانشگاه تهران، شماره ۱.*
- قانون اساسی معارف. (۱۲۸۹).
- کمپفر، انگلبرت. (۱۳۶۳). *سفرنامه کمپفر، ترجمه کیکاس جهاننداری، تهران: خوارزمی، چاپ سوم.*
- کیامنش، علیرضا؛ منصورنیا، حمید. (۱۳۷۲ الف). *اولین ارزشیابی از نظام جدید آموزش متوسطه، تهران: امور اجرایی نظام جدید، اردیبهشت.*
- کیامنش، علیرضا؛ موسی‌پور، نعمت‌الله. (۱۳۷۲ ب). *دومین ارزشیابی جامع از نظام جدید آموزش متوسطه، تهران: امور اجرایی نظام جدید، آذر.*
- کیامنش، علیرضا. (۱۳۷۳). *روش‌های ارزشیابی آموزشی، تهران: پیام نور، چاپ اول.*
- کیامنش، علیرضا؛ موسی‌پور، نعمت‌الله. (۱۳۷۳). *سومین ارزشیابی جامع از نظام جدید آموزش متوسطه، تهران: امور اجرایی نظام جدید.*
- کیامنش، علیرضا؛ موسی‌پور، نعمت‌الله؛ احمدی، غلامعلی. (۱۳۷۴). *چهارمین ارزشیابی جامع از نظام جدید آموزش متوسطه؛ (قسمت اول)، تهران: امور اجرایی نظام جدید.*
- کیامنش، علیرضا؛ احمدی، غلامعلی و موسی‌پور، نعمت‌الله. (۱۳۷۵ الف). *چهارمین ارزشیابی جامع از نظام جدید آموزشی متوسطه؛ (قسمت دوم)، تهران: امور اجرایی نظام جدید.*
- کیامنش، علیرضا؛ موسی‌پور، نعمت‌الله و حمزه‌لویی، احمد. (۱۳۷۵ ب). *چهارمین ارزشیابی جامع از نظام جدید آموزش متوسطه؛ (قسمت سوم)، تهران: امور اجرایی نظام جدید.*
- گزارش کنفرانس ارزیابی انقلاب آموزشی، (۱۳۴۸). شیراز.
- گویا، زهرا. (۱۳۷۸). *سیر تحول و شکل‌گیری برنامه درسی آموزشی متوسطه در ایران، فصلنامه تعلیم و تربیت (ویژه‌نامه برنامه درسی)، ۱۵، ۱.*
- مجله وزارت معارف و اوقاف. (۱۳۰۶). تهران: صنایع مستظرفه.
- محبوبی اردکانی، حسن. (۱۳۷۰). *تاریخ مؤسسات تمدنی جدید در ایران (۳ جلد). تهران: دانشگاه تهران.*
- مرکز برنامه‌ریزی آموزشی. (۱۳۵۳). *ارزشیابی کیفیت آموزش علوم در کلاس پنجم دبستان‌ها، تهران: موسسه تحقیقات و برنامه‌ریزی علمی و آموزشی.*
- مرکز تحقیقات و برنامه‌ریزی درسی. (۱۳۵۰). *ارزشیابی برنامه‌های درسی سال پنجم ابتدایی، تهران.*
- مرکز تحقیقات و برنامه‌ریزی درسی. (۱۳۵۱). *ارزشیابی برنامه‌های درسی سال اول راهنمایی، تهران.*
- مرکز تحقیقات و برنامه‌ریزی درسی. (۱۳۵۲). *ارزشیابی برنامه‌های درسی سال اول راهنمایی، تهران.*

- مشفق آرائی، بهمن و ابیضی، حجت اله. (۱۳۷۲). ارزشیابی برنامه، محتوا و امکانات آموزشی - اجرایی درس تولید و کاربرد مواد آموزشی مراکز تربیت معلم، تهران: سازمان پژوهش و برنامه‌ریزی درسی.
- مگنسون، دیوید. (۱۳۷۰). مبانی نظری آزمون‌های روانی، ترجمه محمدتقی براهنی، تهران: دانشگاه تهران، چاپ دوم.
- موسی‌پور، نعمت الله. (۱۳۷۰). ارزشیابی شیوه آرایه محتوای کتب فرهنگ اسلامی و تعلیمات دینی مقطع ابتدایی، پایان نامه کارشناسی ارشد، تهران: دانشگاه تربیت معلم.
- موسی‌پور، نعمت الله. (۱۳۷۴). روند تحول برنامه‌های درسی آموزش ابتدایی ایران: تحول روش‌ها، مجموعه مقالات اولین سمپوزیوم جایگاه و نقش مطلوب آموزش ابتدایی در جامعه، اصفهان: اداره کل آموزش و پرورش، چاپ اول.
- موسی‌پور، نعمت الله و شعبانی، بختیار. (۱۳۷۶). ارزشیابی برنامه درسی مقطع دکتری رشته های علوم تربیتی در ایران، مجموعه مقالات اولین سمینار آموزش عالی در ایران، تهران: دانشگاه علامه طباطبائی.
- موسی‌پور، نعمت الله. (۱۳۷۶). طراحی الگویی برای ارزشیابی برنامه درسی و بکارگیری آن در ارزشیابی درسی و روش‌ها و فنون تدریس دوره‌های تربیت دبیر دانشگاه‌ها، رساله دکتری، تهران: دانشگاه تربیت مدرس.
- مهندسی، خسرو. (۱۳۵۱). اصول ارزشیابی، تهران: وزارت آموزش و پرورش.
- مهندسی، خسرو. (۱۳۴۹). اندازه‌گیری و سنجش در روانشناسی و تعلیم و تربیت، تهران: نشر اندیشه.
- میرجلیلی، غضنفر. (۱۳۵۹). ارزیابی کتب درسی (فیزیک)، رساله کارشناسی ارشد، تهران: دانشگاه تهران.
- نائینی، محمدکاظم. (۱۳۶۲). تاریخچه سازمان پژوهش و برنامه‌ریزی آموزشی، نشریه شماره ۱۰۰، تهران: سازمان پژوهش و برنامه‌ریزی آموزشی.
- نصفت، مرتضی. (۱۳۴۸). در باره بازرسی و ارزشیابی آموزشی، تهران: بی نا، چاپ اول.
- نصفت، مرتضی. (۱۳۶۵). ارزیابی برنامه آزمایشی ۴۲ ساعت هفتگی برای مدارس تیزهوشان، تهران: مؤسسه روان‌شناسی.
- نیتی، محمدعلی. (۱۳۴۰). ارزیابی تعلیمات ابتدایی در ایران، پایان‌نامه کارشناسی ارشد، تهران: دانشگاه تهران.
- واحد تحقیقات و طرح و برنامه صنایع آموزشی. (۱۳۷۱). ارزشیابی مجموعه‌های آموزشی علوم و ریاضی دوره ابتدایی، تهران: شرکت صنایع آموزشی.
- وایرژما، ویلیام و جی، استفن. (۱۳۷۲). اندازه‌گیری و آزمون در تعلیم و تربیت، ترجمه غلامرضا خوبی نژاد. مشهد: آستان قدس رضوی، چاپ اول.
- ودیعی، کاظم. (۱۳۵۳). در باره کتابهای درسی در سمینار مشورتی کتابهای درسی، ماهنامه آموزش و پرورش، ۴۳، ۸.
- وزارت فرهنگ و آموزش عالی. (۱۳۷۲). آشنایی با دفتر نظارت فرهنگ و آموزش عالی، تهران.
- ولف، ریچارد. (۱۳۷۱). ارزشیابی آموزشی (مبانی سنجش توانایی و بررسی برنامه‌ها)، ترجمه علیرضا کیامنش، تهران: نشر دانشگاهی، چاپ اول.
- هاشمی، ابراهیم و بطحایی، حسن. (۱۳۴۵). آزمونهای روانی، تهران: نشر چهر.

هوشیار، محمد باقر. (۱۳۱۷). *سنجش هوش یا روان‌شناسی عملی: تجارب روان‌شناسی در آزمایشگاه*، تهران: بی‌نا، چاپ اول.

هومن، حیدرعلی. (۱۳۵۴). *اندازه‌گیری روانی و فن تهیه تست*، تهران: بی‌نا چاپ اول.

هومن، حیدرعلی. (۱۳۷۵). *زمینه ارزشیابی برنامه‌های آموزشی*، تهران: نشر پارسا، چاپ اول.

(۱۳۵۱). *یک مدل ارزشیابی از برنامه‌های درسی در تایلند، ماهنامه آموزش و پرورش*، ترجمه ابراهیم کظیمی، ۴۲،

۳.

ب: انگلیسی

Ballou, F.A.(1916). *Work of The Department of Educational Investigation and Measurment*, In: G.M. Whipple (Ed.), *Standards and Tests for the Measurment of Efficiency of School Systems. National Society for Study of Education, Fifteenth Yearbook, Part J*. Chicago: University of Chicago Press.

Biddle, B. J. and Ellena, W. J. (Ed.). (1964). *Contemporary Research on Teacher Effectiveness*, New York: Rinehart and Winston.

Bloom, B. S. and Others. (1956). *Taxonomy of Educational Objectives Handbook 1: The Cognitive Domian*, New York: David Mckay Co.

Brinkerhoff, R. (1988). *Evaluation Technical Heritage, Science*, 4433, 207, 831-36.

Callahan, R. E. (1962). *Education and the Cult of Efficiency*, Chicago: University of Chicago Press.

Coleman, J. S. and Others. (1966). *Equality of Educational Opportunity*, Washington, D.C: Office of Education, U.S. Department of Health, Education, and Welfare.

Cook, D. L. (1966). *Program Evaluation and Review Technique Applications in Education*, Washington D.C: Government Printing Office.

Cremin, L. J. (1962). *Toward Reform of Program Evaluation*, San Francisco: Jossey - Bass Publishers.

Cronbach, L. J. (1963). *Course Improvement Through Evaluation, Teachers College Record*, 64, 672-83.

Cronbach, L. J. (1980). *Toward Reform of Program Evaluation*, San Francisco: Jossey-Bass Publishers.

Cronbach, L. J. (1988). *Evaluation Models*, Boston: Kluwer - Nijhoff Publishing.

Eisner, E. W. (1967). *Educational Objectives: Help or Hindrance, The School Review*, 75, 250-60.

Eisner, E. W. (1975). *The Perceptive Eye: Toward the Reformation of Educational Evaluation*, California: Stanford Evaluation Consortium.

Evers, J.W. (1980). *A Field Study of Goal-Based and Goal-Free Evaluation Techniques*, Unpublished, **Doctoral Dissertation**, Westren Michigan University.

Gephart, William. J. and Ayers, Jerry. B. (1988). *Teacher Education Evaluation*, Boston: Kluwer Academic Publishers.

Glaser, R. (1963). *Instrucational Technology and The Measurment of Learning Outcomes: Some Questions, American Psychologist*, 18, 519-20.

- Glass, G. V. & et al (1970). **Data Analysis of the 1968-69 Survey of Compensatory Education, Title I, Final Report of Grant No. OEG 8-8-9618604003-(058)**, Washington, D.C: Office of Education.
- Glass, G. V. (1976). *Primary, Secondary, and Meta Analysis of Research, Educational Researcher*, 10, S.
- Guba, E. G. (1966). **A Study of Title III Activities: Report on Evaluation**, Indiana: Indiana University, National Institute for the Study of Educational Change.
- Guba, E. G. and Lincoln, Y.S. (1981). **Effective Evaluation**, San Francisco: Jossey-Bass Publishers.
- Hammond, R. L. (1967). **Evaluation at the Local Level**, Address to the Miller Committee for the National Study of ESSA Title III.
- House, Ernest. R. (1988). *Assumptions Underlying Evaluation Models*. in Madaus, G.F. Scriven, and D. L. Stufflebeam. (1988). **Evaluation Models**, Boston: Kluwer - Nijhoff Publishing.
- House, Ernest. R. (1993). **Professional Evaluation (Social Impact and Political Consequences)**, London: SAGE Publications.
- Johnson, R. (1980). **Directory of Evaluators and Evaluation Agencies**, New York: Exxon Corporation.
- Kellaghan, T. and Madaus. G. F. (1982). *Trends in Educational Standards in Great Britain and Ireland*, in G.R. Austin and H. Garder (Eds.), **The Rise and Fall of National Test Scores**, New York: Academic Press.
- Kendall, C. N. (1915). *Efficiency of School and School Systems*, In **Proceedings and Addresses of the Fifty - Third Annual Meeting of National Education Association**, New York: pp. 386-95.
- Krathwohl, D. R. and Others. (1964). **Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook II: Affective Domain**, New York: David Mckay Co.
- Lewy, Arieh (Ed.). (1991). **The International Encyclopedia of Curriculum**, New York: Pergamon Press.
- Lindquist, E. F. (1953). **Design and Analysis of Experiments in Psychology and Education**, Boston: Houghton-Mifflin.
- Madaus, G. F. and McDonagh, J.T. (1982). *ASI Roved Out: Folksong Collectioning as a Metaphor for Evaluation*, In N.L. Smith (Ed.), **Communicating in Evaluation: Alternative Forms of Representation**, Beverly Hills, California: Sage Publications.
- Madaus, G. F. Stufflebeam, D. L. and Scriven, F. D. (1988). *Program Evaluation: A Historical Overview*, In Madaus, G. F., Scriven. and Stuffleveam. D. L. (1988). **Evaluation Models**, Boston: Kluwer - Nijhoff Publishing.
- May, P. (1971). *Standardized Testing in Philadelphia, 1916-1938*, Unpublished Manuscript.
- Metfessel, N. S. and Michael, W.B. (1967). *A Paradigm Involving Multiple Criterion Measures for the Evaluation of the Effectiveness of School Programs*, **Educational and Psychological Measurement**, 27, 931-43.
- Peck, H. (1981). *Report on the Certification of Evaluators in Louisiana, Paper Presented at the Meating of the Southern Educational Research Association*, Kentucky: Lexington, Fall.

- Picariello, H. (1962). **Evaluation of Title I**, Washington, D.C: American Institute for the Advancement of Science.
- Pinker, R. (1971). **Social Theory and Social Policy**, London: Heinemann Educational Books.
- Phi Delta Kappa Commission on Evaluation. (1971). **Educational Evaluation and Decision Making**, Itasca, Illinois: J publishers.
- Popham, W. J. (1971). **Criterion - Referenced Measurement**, Englewood Cliffs, New Jersey: Educational Teschnology Publications.
- Provus, M. (1969). **Discrepancy Evaluation Model**, Pittsburgh, Pennsylvania: Pittsburgh Public School.
- Provus, M. (1971). **Discrepancy Evaluation**, Berkeley, California: Mc Cutcheon Publishing Co.
- Reinhard, D. (1972). **Methodolgy Development for Input Evaluation Using Advocate and Design Teams**, Unpublished Doctoral Dissertation, Ohio State University.
- Rice, J. M. (1914). **Scientific Management in Education**, New York: Hinds, Noble and Eldredge.
- Rice, J. M. (1897). *The Futility of the Spelling Grind*, **The Forum**, 23, 163-172.
- Roth, J. (1977). *Needs and the Needs Assessment Process*, **Evaluation News**, 5, 15-17
- Roth, J. E. (1978). **Theory and Practice of Needs Assessment with Special Application to Institutions of Higher Learning**, Unpublished Doctoral Dissertation, Berkeley: University of California.
- Scriven, M. S. (1975). **Evaluation Bias and its Control**, Occasional Paper Series no. 4. Western Michigan University Evaluation Center.
- Scriven, M. and Roth. J. E. (1977). *Need Assessment*, **Evaluation News**, 2, 25-28.
- Smith, H. L. and Judd, C. H. (1914). *Plans for Organizing School Surveys*, **National Society for the Study of Education**, Thirteenth Yearbook, Part II, Bloomington Illinois: Public School Publishing Co.
- Smith E. R. and Tyler, R.W. (1942). **Appraising and Recording Student Progress**, New York: Harper.
- Stake, R. E. (1967). *The Countenance of Educational Evaluations*, **Teachers College Record**, 523-40.
- Stake, R. E. and Gjerde. C. (1974). *An Evaluation of T-City, the Twin City Institute for Talented Youth*, **AERA Monograph Series in Curriculum Evaluation**, 7, Chicago: Rand. McNally.
- Stake, R. E. (1975). **Some Alternative Presumptions**, Urbana, Illinois: Center for Instructional Research an Curriculum Evaluation, Octobed.
- Stake, R. E. (1976). **Evaluation Educational Programmes**, Organization for Economic Co-Operational Development.
- Stake. R. E. (1981). *Setting Standards for Educational Evaluators*, **Evaluator News**, 2, 2, 148-52.
- Stufflebeam, D. L. (1967). *The Use and Abuse of Evaluation in Title III, Theory into Practice*, 6, 126-133.
- Stufflebeam, D. L. & et al. (1971). **Educational Evaluation and Decision- Making**, Ithaca, Illinois: Peacock Publishers.

- Stufflebeam, D.L.(1978). *Meta Evaluation: An Overview, Evaluation and the Professions*, 2, 1.
- Suarez, T. (1980). *Needs Assessments for Technical Assistance, A Conceptual Overview and Comparison of Three Strategies*, Unpublished Doctoral Dissertation, Western Michigan University.
- Tyack, D. and Hansot, E. (1967). *Changing Concepts of Educational Evaluation*, In: R,E. Stake (ed), *Perspectives of Curriculum Evaluation*, Vol. 1, NewYork: Rand McNally.
- Tyack, D. and Hansot, E. (1982). *Managers of Virtue*, New York: Basic Books, Inc.
- Tyler, R. W. (1949). *Basic Principles of Curriculum and Instruction*, Chicago: University of Chicago Press.
- Tyler, R. W. (1967). *Changing Concepts of Educational Evaluation*, In: R.E. Stake (Ed.), *Perspectiees of Curriculum Evaluation*, Vol. 1, NewYork: Rand McNally.
- Webster, W. J. (1975). *The Organization and Functions of Research and Evaluation in Large Urban School Districts*, Paper Presented at the Annual Meeting of the American Educational Research Association, Washington, D.C., March.