

مقایسه تأثیر ارزشیابی تکوینی و  
ارزشیابی پایانی بر پیشرفت تحصیلی دانشجویان  
دانشگاه شهید چمران  
دکتر حسین سپاسی \*

چکیده

مدت زمانی طولانی است که معلمان، علاقه‌مند به فهم مساعادت ارزشیابی تکوینی در یادگیری بوده‌اند. بسیاری اعتقاد دارند که ارزشیابی تکوینی پیشرفت بهتر دانش آموزان را فراهم می‌آورد و بر نگرش مثبت آنها نسبت به یادگیری تأکید می‌گذارد. این پژوهش برای آزمون این فرضیه که به کار بردن ارزشیابی تکوینی نتایج یادگیری را بهبود می‌بخشد طرح ریزی شد. تعداد ۱۱۲ دانشجوی رشته‌های دبیری از بین دانشجویانی که برای درس سنجش و اندازه‌گیری ثبت نام کرده بودند به عنوان افراد مورد مطالعه انتخاب شدند. آنها در آغاز پژوهش به طور تصادفی به دو گروه آزمایشی و گواه تقسیم شدند. هر دو گروه در یک پیش آزمون شرکت کردند. تفاوت معنی داری بین دو گروه در ابتدای پژوهش از لحاظ دانش، مهارت، و آگاهی‌شان به دست نیامد ( $p < 0/05$ ). گروه آزمایشی در ۵ آزمون تکوینی و یک آزمون نهائی (پس آزمون)، و گروه گواه فقط در یک آزمون نهائی شرکت داشتند. شیوه‌های تدریس به جز مداخله آزمایشی (ارزشیابی تکوینی) برای هر دو گروه یکسان بود. نتایج نشان داد که اختلاف معنی داری بین نمره‌های آزمون دو گروه وجود دارد. این نتیجه نشان می‌دهد که معلمان باید اهمیت بیشتری به نقش ارزشیابی تکوینی در پیشرفت تحصیلی دانش آموزان بدهند.

مقدمه

اندازه‌گیری و ارزشیابی در تعلیم و تربیت تاریخچه‌ای طولانی دارد. مراسمی که در آن قبایل بدوی عادات و رسوم قبیله‌ای و تحمل و آمادگی جوانان را برای اجازه ورود به مراتب بزرگسالی آزمون می‌کردند را می‌توانیم به آنچه امروزه به آن ارزشیابی تکوینی می‌گوئیم ربط دهیم. به عقیده رومبرگ (Romberg, ۱۹۹۲) شاید این قدیمی‌ترین نوع امتحانی است که به وسیله انسان به کار گرفته شده است. از نقطه نظر منابع تاریخی، می‌دانیم که مقامات چینی در ۲۲۰۰ سال

\* - دکتر حسین سپاسی، عضو هیأت علمی گروه علوم تربیتی دانشکده علوم تربیتی و روانشناسی دانشگاه شهید چمران اهواز.

پیش از نتایج آزمونها، با شیوه ارزشیابی پایانی، به عنوان روشی برای انتخاب مسئولان اداری استفاده می‌کردند (ابل ۱۹۹۱، Abel). در ایران باستان در عهد شاپور پادشاه ساسانی در جندی‌شاپور هم جلساتی برای آزمون دانشجویان پزشکی برپا می‌کردند و اعطاء گواهی نامه پزشکی موکول به موفقیت دانشجویان در این امتحان می‌گردید (کیامنش، ۱۳۷۰).

امروزه مقاصد ارزشیابیهای تعلیم و تربیت با گذشته تفاوت بسیار دارند و برای اهداف گوناگون به طور چشم‌گیری تغییر یافته‌اند. با گسترش سیستمهای آموزش و پرورش و افزایش تعداد دانش آموزان در پایه‌های مختلف تحصیلی شیوه‌های ارزشیابی برای تعیین میزان یادگیری دانش آموز از حالت فردی یا گروههای ناهمگن کوچک به آزمونهائی در مقیاسی وسیعتر و گروههای متجانستر تغییر پیدا کرده‌اند. معلمان مدارس در شرف تغییر شیوه‌های ارزشیابی خود از آزمون شفاهی که به صورت فردی انجام می‌شود به شیوه‌های مختلف آزمونهای کتبی می‌باشند.

تاکنون مطالب زیادی در باره دو شیوه مهم و بالنسبه متفاوت ارزشیابی، به نامهای ارزشیابی تکوینی (Formative Evaluation) و ارزشیابی پایانی (Summative Evaluation) به رشته تحریر درآمده است. همچنین این فکر به طور گسترده‌ای مورد قبول واقع شده است که این دو روش ارزشیابی بر جنبه‌های متفاوت یادگیری اثر می‌گذارد. روشهای ارزشیابی تکوینی بر کسب دانش و اطلاعات روزانه یا هفتگی و فرایندهای مربوط به آنها تأثیر می‌گذارد، در حالی که ارزشیابی پایانی نمونه‌ای از حیطه دانش مورد نظر فراهم می‌آورد. ارزشیابی تکوینی با استفاده از داده‌های سودبخشی فرایند یادگیری را نشان می‌دهد. در مقابل ارزشیابی پایانی در پایان فرایند کاری انجام می‌پذیرد و پیشرفت را در انتهای واحد درسی تعیین می‌کند، (اسکریون Scriven ۱۹۷۶).

به طور کلی، ارزشیابی تکوینی موقعیت و وضعیت دانش آموزان را در حالی که هنوز در بعضی مراحل پیشرفت قرار دارند اندازه‌گیری می‌کند. مهمرنز و لهمان (Mehrens & Lehman ۱۹۸۴) می‌گویند منظور از ارزشیابی تکوینی گردآوری اطلاعات است در حالی که یادگیری جریان دارد تا بدین وسیله بتوانیم تجربیات یادگیری دانش آموز را افزایش دهیم. شووتز (Schutz ۱۹۷۰) اظهار می‌دارد می‌توان انتظار داشت که نتایج فعالیتهای ارزشیابی تکوینی بتواند برنامه‌های آموزشی را پیشرفت دهد و همچنین داده‌هایی برای اعتباریابی نتایج آزمونها فراهم آورد. متخصصان سنجش و اندازه‌گیری در تعلیم و تربیت، بیکر و کولیس (Baker & Collis ۱۹۸۲) یا ابل و فریزبی (Ebel & Frisbee ۱۹۹۱)، و بلوم (Bloom ۱۹۷۶) اعتقاد دارند که در فرایند پیشرفت دانش آموز باید توجه خاصی به اهمیت ارزشیابی تکوینی داده شود. آنها همچنین به اهمیت ارزشیابی تکوینی به عنوان وسیله‌ای برای ارزیابی قدرت و ضعف تجربیات یادگیری دانش آموز اشاره می‌کنند.

ارزشیابی تکوینی می‌تواند در نگرش مثبت دانش آموز نسبت به آنچه در کلاس درس آموزش داده می‌شود مؤثر باشد. بلوم (Bloom, ۱۹۷۶) نشان داد که بین نگرش و اندازه‌ای که پیشرفت تحصیلی را می‌سجد همبستگی معنی داری وجود دارد. وی تخمین می‌زند که می‌توانیم ۲۵ درصد واریانس نمره‌های پیشرفت تحصیلی را به نگرشهای مثبت نسبت دهیم. فونچ و دیگران (Funch. et al, ۱۹۸۵) عملکرد دانشجویان رشته‌های دبیری دو گروه آزمایشی و گواه را مورد بررسی قرار دادند. آنها در یافتند که عملکرد دانشجویان رشته‌های دبیری در گروه آزمایشی که به وسیله ارزشیابی تکوینی مورد سنجش قرار گرفته بودند از عملکرد دانشجویان گروهی که هیچ‌گونه بازخوردی دریافت نکرده بودند بهتر بود.

با این که ارزشیابهای تکوینی از پیشرفت دانش آموز در کلاس درس می‌تواند در فرایند آموزش و یادگیری بسیار با اهمیت باشد، بسیاری شیوه ساختن آزمونهای تکوینی را مورد انتقاد قرار داده‌اند. استیگینز (Stiggins, ۱۹۸۹) یا آوری می‌کند که در تمام پایه‌های تحصیلی آزمونهای مداد - کاغذی که به سئوالات حفظی اختصاص داده شده است تنها شکلهای ارزشیابی به شمار می‌روند. بندرت سئوالاتی برای اندازه‌گیری قدرت فهم و درک و یا مفاهیم عالی ذهنی پرسیده می‌شود. برخی معلمان اعتقاد دارند که ارزشیابی تکوینی موجب اتلاف وقت کلاس درس می‌شود و از زمان تدریس مطالب می‌کاهد (راس Ross, ۱۹۵۴). عده‌ای دیگر هم براین باورند که اجرای آزمونهای مکرر برای کودکان در کلاسهای ابتدائی یا برای دانش آموزان با توانائی ضعیف لازم و مناسب است (هایکینز Hopkins, ۱۹۸۷).

با وجودی که سودبخشی ارزشیابی تکوینی توسط بسیاری از متخصصان سنجش و اندازه‌گیری به عنوان وسیله، تسهیل کسب تجربیات یادگیری به دانش آموز مورد تأکید قرار گرفته است، ولی به نظر می‌رسد که برای بررسی و مقایسه کاربردهای فرایندهای مختلف یادگیری و شیوه‌های مختلف ارزشیابی در اندازه‌گیری پیشرفت تحصیلی دانش آموزان به پژوهشهای بیشتری نیاز داریم (پیج Paige, ۱۹۶۶، آلکین Alkin, ۱۹۶۹). همچنین با اینکه معلمان رشته‌های مختلف احتمالاً "مهارتهای فکری را تا حدودی به دانش آموزان می‌آموزند، ولی غالباً" آموزش کافی برای ارزشیابی چنین مهارتهائی ندارند. علاوه بر این، چون معلمان در مدارس از شیوه‌های متفاوتی برای اندازه‌گیری پیشرفت دانش آموزان استفاده می‌کنند، این میزان تأثیر ارزشیابی تکوینی بر نتایج پیشرفت تحصیلی نیاز به پژوهشهای بیشتری دارد.

تحقیقاتی که در زمینه یادگیری دانش آموزان انجام گرفته است نشان می‌دهد که آنها شرکت کنندگان غیر فعالی در کسب آنچه به آنها آموخته می‌شود نیستند. بلکه آنها با مقایسه اطلاعات جدید با دانش و تجربیاتی که از پیش آندوخته‌اند به تفسیرهای تازه‌ای از مفاهیم درسی می‌پردازند (ویلسون Wilson, ۱۹۹۲). دیدگاههای جدیدی که در مورد یادگیری ارائه شده است پیشنهادهایی برای نظریه‌ها و مدل‌های تازه اجرای آزمونها به عنوان جایگزینی برای آزمونهای

ملاکی که سنتاً روی رفتارهای معین دانش آموز تأکید می‌گذراند به همراه داشته است (بلوم و هستنیک Bloom & Hasting، ۱۹۷۱). این نظریه‌های جدید گرفتن آزمون، همچون ارزشیابی عملکرد (وان هیل Van Hiele، ۱۹۸۶) و طبقه بندی ساختار نتایج یادگیری (بیکر و کولیس Baker & Collis، ۱۹۸۲)، نیازمند آنند که معلمان شیوه‌های طبقه بندی پاسخهایی را که دانش آموزان به سئوالات آزمون می‌دهند به صورت ساختاری از دانش و معلومات بدانند و نوعی ارزشیابی کلاسی یا ارزشیابی تکوینی برای دسترسی به چنین اطلاعاتی به وجود آورند. با توجه به این که این گونه نظریه‌های ارزشیابی دانش، مهارتها، و آگاهیهای دانش آموزان از موضوع درس تازگی دارند، معلمان باید چگونگی تدریس این شیوه‌های جدید، راههای ارزشیابی معلومات دانش آموزان، انواع شیوه‌های ارزشیابی را که می‌شود در کلاس درس به کار برد بدانند. مؤسساتی که عهده دار تعلیم و تربیت معلمان هستند باید تلاش کنند تا با فراهم آوردن فرصتهای مناسب برای آموزش نظریه‌ها و مدل‌های جدید آزمون سازی که می‌توانند در اندازه گیری تواناییها و مهارتهای عالی فکری به کار برده شوند و نوع فعالیتهای کلاسی که به وسیله آنها می‌توان به ارزشیابی پیشرفت دانش آموز کمک کرد فراهم آورند.

### فرضیه و موضوع تحقیق

پس از بازنگری مختصر پیشینه تحقیق، این موضوع برای تحقیق برگزیده شد که آیا تأثیر ارزشیابی تکوینی بر پیشرفت تحصیلی دانشجویان از ارزشیابی پایانی بیشتر است؟ فرضیه‌ای که در تحقیق حاضر مورد آزمون قرار گرفت عبارت بود از: میانگین نمره‌های امتحان نهائی دانشجویانی که در برنامه ارزشیابی تکوینی شرکت می‌کنند از میانگین نمره‌های دانشجویانی که تنها در پایان نیمسال از آنان یک امتحان نهائی به عمل می‌آید بیشتر است.

### روش نمونه گیری

یکصد و دوازده دانشجو به عنوان نمونه تحقیق انتخاب گردید. این افراد را دانشجویان رشته‌های دبیری دانشکده‌های ادبیات و الهیات و معارف اسلامی تشکیل می‌دادند که از بین دانشکده‌هایی که در نیمسال دوم سال تحصیلی ۷۱ - ۱۳۷۰ درس سنجش و اندازه گیری را در دانشگاه شهید چمران اهواز انتخاب کرده بودند به طور تصادفی انتخاب شدند.

## ابزار تحقیق

از دو فرم همتای آزمون پیشرفت تحصیلی معلم - ساخته به طور تصادفی به عنوان پیش آزمون و پس آزمون استفاده شد. ضریب پایانی فرمهای همتای آزمون قبلاً در رابطه با دانشجویان مشابهی که برای این درس ثبت نام کرده بودند برابر با ۰/۹۱ محاسبه گردیده بود.

## مراحل اجرای تحقیق

برای اینکه امکان تقسیم تصادفی دانشجویان به دو گروه آزمایشی و گواه حاصل شود و تداخلی در انتخاب زمان ارائه این درس با درسهای دیگر فراهم نشود، با هماهنگی آموزش دانشکده‌های مزبور مدت زمان کلاسهای این واحد ۴ ساعت در نظر گرفته شد و هنگام انتخاب واحد به آنها یادآوری گردید. پس از انتخاب واحد، دانشجویان به طور تصادفی به دو گروه تجربی و گواه تقسیم شدند. زمان تشکیل کلاسها در روز شنبه و یکشنبه هر هفته از ساعت ۱۴ تا ۱۶ به طور تصادفی به گروه آزمایشی و از ساعت ۱۶ تا ۱۸ به گروه گواه اختصاص یافت. در اولین جلسه تشکیل کلاسها سرفصل درس، هدفهای کلی درس، شیوه نمره گذاری و مطالبی از این قبیل به طور کامل و به صورت مکتوب به اطلاع دو گروه آزمایشی و گواه رسید. آزمودنیهای گروه آزمایشی در شروع کلاس درس در یک پیش آزمون و متعاقب آن مطابق با برنامه تنظیمی پس از تدریس بخشهایی از درس در فواصل زمانی معین در ۵ آزمون تکوینی در طول ترم و بالاخره در یک آزمون پایانی (پس آزمون) که زمان برگزاری آن هنگام انتخاب واحد توسط آموزش دانشکده‌ها به اطلاع رسیده بود شرکت کردند. آزمونهای تکوینی گروه آزمایشی هر کدام به ارزش ده درصد و امتحان آخر ترم (پس آزمون) نیز معادل ۵۰ درصد شیوه نمره گذاری دانشجویان گروه آزمایشی را تا پایان ترم تشکیل می داد. زمان آزمونهای تکوینی و تعداد صفحاتی که آزمون از آنها گرفته می شد، جزوه و منابعی که می باید مورد مطالعه آنها قرار می گرفت قبلاً به اطلاع آنها رسیده بود. آزمونهای تکوینی در جلساتی که به همین منظور در نظر گرفته شده بودند به مدت ۱۵ تا ۲۰ دقیقه اجراء گردید و در اختیار آنان قرار گرفت. به علاوه، بازخوردها همراه با اطلاعات آماری از قبیل میانگین نمره‌های کلاس، نمره خام دانشجو، تبدیل نمره خام آنها به صدک همراه با برگه امتحانی به اطلاع آزمودنیهای گروه آزمایشی رسید. در مقابل آزمودنیهای گروه گواه غیر از پیش آزمون تنها در یک آزمون آخر ترم (پس آزمون) شرکت کردند و نمره درسی آنها فقط براساس نمره پایان ترم آنها معادل صد در صد محاسبه شد. بجز عامل آزمایشی (آزمونهای تکوینی) که در مورد آزمودنیهای گروه آزمایشی به اجرا در آمد، روش

تدریس، استفاده از جزوه‌های درسی و منابع کتابخانه‌ای، همدلی با دانشجو، و جز اینها برای هر دو گروه یکسان بود. به بیان دیگر، سعی گردید تا حد ممکن از عواملی که موجب خطای اندازه گیری می‌شدند کاسته گردد تا چنانچه تفاوتی بین نمره‌های پایانی ترم آزمودنیهای دو گروه مشاهده شود بتوان آن را به تأثیر و مداخله عامل آزمایشی نسبت داد.

### روش تحلیل داده‌ها

با توجه به صورت‌بندی که در مورد آزمودنیهای دو گروه انجام پذیرفته بود، برای تحلیل داده‌ها از یک طرح آزمایشی به شرح زیر استفاده شد.

	پیش آزمون	عامل آزمایشی	پس آزمون
R گروه آزمایشی	01	X	02
R گروه گواه	01	-	02

X = آزمون تکوینی

02 = آزمون پایانی

آزمونهای آماری متفاوتی بین نمره‌های 1- پیش آزمونهاى دو گروه آزمایشی و گواه، 2- پیش آزمون و پس آزمون گروه آزمایشی، 3- پیش آزمون و پس آزمون گروه گواه، 4- پس آزمونهاى دو گروه آزمایشی و گواه، به منظور بررسی تأثیر عامل آزمایشی انجام گرفت. لازم به توضیح است که بر اساس طرح آزمایشی فوق، اختلاف معنی دار میانگین نمره‌های پس آزمونهاى گروه‌های آزمایشی و گواه و عدم تفاوت میانگینهای پیش آزمون این دو گروه که با آزمون t گروه‌های مستقل انجام شد موجبات آزمون فرضیه صفر این تحقیق را فراهم آورد.

### یافته‌های تحقیق

جدول 1 نتایج تحلیلهای آماری دو گروه آزمایشی و گواه را در پیش آزمون و پس آزمون برای گروه‌های مستقل نشان می‌دهد.

جدول ۱ - میانگینها و نتایج آزمونهای t مربوط به مقایسه نمره‌های پیش آزمون و نمره‌های پس آزمون گروههای آزمایشی و گواه.

df	t	گروه گواه $\bar{x}$	گروه آزمایشی $\bar{x}$	گروهها نوع آزمون
۱۱۸	۱/۱۷	۱۵/۱۲	۱۴/۴۴	پیش آزمون
۱۱۸	۲/۰۷	۲۰/۴۱	۴۰/۴۲	پس آزمون

نتایج محاسبات نشان داد که مقدار t محاسبه شده بین نمره‌های امتحان نهایی گروههای آزمایشی و گواه از مقدار بحرانی برابر ۱/۶۵۸ در سطح احتمال ۵ درصد بزرگتر بود. یعنی اینکه بین دو گروه آزمایشی و گواه در امتحان نهایی تفاوت معنی داری وجود دارد. به بیان دیگر، دلیل معنی دار بودن اختلاف بین نمره‌های افراد دو گروه در امتحان نهایی را می‌توان نه به دلیل عامل شانس بلکه به مداخله آزمایشی (ارزشیابی تکوینی در مقابل ارزشیابی پایانی) نسبت داد. بنابراین، فرضیه‌ای که این تحقیق به منظور آن به اجرا در آمد مورد تأیید واقع گردید.

در آزمون معنی داری که بین میانگین نمره‌های پیش آزمون و پس آزمون گروه آزمایشی انجام پذیرفت مقدار t محاسبه شده برای گروههای وابسته برابر ۲۳/۷۶ محاسبه گردید. مقایسه این مقدار با مقدار بحرانی آن ( $t = ۱/۶۷۱$ ) نشان داد که در پیشرفت تحصیلی گروه آزمایشی اختلاف معنی داری از آغاز تا پایان دوره آزمایشی به وقوع پیوست.

به علاوه، هنگامی که میانگین نمره‌های پیش آزمون و پس آزمون گروه گواه مقایسه شد، نتایجی مشابه گروه آزمایشی به دست آمد. میانگین پیش آزمون و پس آزمون گروه گواه به ترتیب برابر ۱۵/۱۲ و ۳۶/۱۶ محاسبه شد. مقدار t محاسبه شده برابر ۲۰/۴۱ از مقدار بحرانی برابر ۱/۶۷۱ در سطح احتمال آلفای مساوی ۵ درصد فاصله زیادی را نشان داد. این گروه نیز در طول دوره آزمایشی پیشرفت قابل ملاحظه‌ای از خود نشان دادند. جدول ۲- نتایج مقایسه‌های فوق را نشان می‌دهد.

جدول ۲- میانگینها و نتایج آزمونهای t مربوط به مقایسه نمره‌های پیش آزمون با پس آزمون گروه‌های آزمایشی و گواه.

df	t	پس آزمون $\bar{x}$	پیش آزمون $\bar{x}$	نوع آزمون گروهها
۵۹	۲۳/۷۶	۱۵/۱۲	۱۴/۴۶	گروه آزمایشی
۵۹	۲۰/۴۱	۳۶/۱۴	۴۰/۴۶	گروه گواه

### بحث و نتیجه گیری

بررسی دفتر حضور و غیاب دانشجویان در کلاسهای درس نشان داد که افراد گروه آزمایشی در تمام جلسات کلاس درس حضور داشتند و حتی به ندرت در آمدن به کلاس تأخیر کردند. پیشرفت تدریجی که در درصد پاسخهای صحیح به آزمونها برای هر دانشجو در گروه آزمایشی از متوسط ۸۰ درصد در اولین آزمون تا ۹۵ درصد در آخرین آزمون مشاهده گردید، بازخوردهائی از تلاش و کوشش آنان در یادگیری مطالب درسی فراهم آورد. از این نتایج امکان دارد چنین استنباط کنیم که بازخورد مداومی که گروه آزمایشی در باره عملکردشان از نمره هر آزمون تکوینی دریافت می‌داشتند آنها را قادر ساخته بود تا بتوانند مقایسه‌های درون فردی و درون گروهی انجام دهند. به علاوه، آنها قادر بودند عملکردشان را در آزمون تکوینی ارزشیابی کنند و پیشرفتشان را با تلاشی که برای یادگیری مطالب درسی انجام می‌دادند مقایسه کنند. این پیشرفت در یادگیری و همچنین حضور مرتب آنها در کلاس درس را می‌توان به عنوان عاملی بر نگرش آنان نسبت به کلاس و پیشرفتشان در دسترسی به هدفهای واحد درسی قلمداد کنیم.

در رابطه با گروه گواه، اجرای هیچ آزمونی در طول نیمسال برنامه ریزی نشده بود و ارزشیابی دانشجویان در این گروه فقط براساس یک آزمون در پایان نیمسال تحصیلی استوار بود. بررسی دفتر حضور و غیاب نشان داد که در مقایسه با حضور مرتب گروه آزمایشی عده‌ای از افراد گروه گواه غیبت داشتند و یا با تأخیر در کلاس حاضر می‌شدند. این مطلب را شاید بتوان به عنوان دلیلی بر اثر مثبت آزمونهای تکوینی در نگرش اعضا، گروه آزمایشی نسبت به روش و اندازه گیری این واحد درسی تلقی کنیم.

این امکان وجود دارد که نوع اطلاعاتی که اعضای گروه آزمایشی در باره عملکردشان از هر نوع آزمون به دست می‌آوردند، آنها را قادر ساخته بود تا تجربیات یادگیری خود را پیشرفت



دهند. به بیان دیگر، افراد گروه گواه که بازخوردی در باره پیشرفت‌شان در دسترسی به هدفهای واحد درسی دریافت نمی‌کردند عملکردی در سطحی پائینتر نشان دادند، گرچه آنها در آزمون نهائی نسبت به پیش آزمون به طور معنی داری پیشرفت کرده بودند. احتمالاً چنین به نظر می‌رسد که ارزشیابی تکوینی که در مراحل مختلف یادگیری واقع گردیده بود فرصتهایی هم در اختیار معلم و هم در اختیار آزمودنیها گذاشته بود تا تغییراتی در اساس تدریس و پیشرفت یادگیری فراهم آورند.

از آنجا که ظاهراً آزمونهای تکوینی داده شده در طول واحد درسی موجبات اختلاف معنی داری با ارزشیابی پایانی در آزمون نهائی پایان نیمسال تحصیلی فراهم آورد، یافته‌های این تحقیق گفته‌های رومبرگ (Romberg, ۱۹۹۰) را که "پیشرفت به واسطه ایجاد تبحر نسبت به مراحل ساده درسی، به وجود آوردن سلسله مراتب در یادگیری، روشهای ساده و روشن دسترسی به هدفهای درس، طرح درسهای روزانه، آزمونهای مکرر، اجرای آزمونهای عینی در کوتاهترین مراحل یادگیری" فرایند یادگیری "دانش آموز را سهولت می‌بخشد، مورد تأیید قرار می‌دهد.

با توجه به اینکه به نظر می‌رسد، در پیشینه تحقیق نقش ارزشیابی تکوینی به نحو گسترده‌ای در افزایش یادگیری مورد قبول واقع شده است (استانلی و یوفام Stanlee & Popham, ۱۹۶۱)، لازم است برنامه‌هایی برای آموزش معلمان تهیه کنیم و این مطالب را به آنها منتقل سازیم. معلمان باید بدانند بالاترین پیشرفت تحصیلی به عواملی چون آگاهی از پیشرفت (روشی که به وسیله آن نتایج ارزشیابی تکوینی مورد بررسی قرار می‌گیرند)، انگیزش درونی (دانش آموزان در طول دوره تحصیلی جدیت و تلاش می‌کنند برای اینکه می‌خواهند نمره‌های خوبی بگیرند)، آگاهی از نتایج عملکرد آزمونهای تکوینی فرصتهای مغتنمی در اختیار دانش آموزان قرار می‌دهد تا به نقاط قوت و ضعف خود در موضوع درس مورد نظر پی ببرند، و شرکت فعال در کلاس (که از ثبت در دفتر حضور و غیاب معلوم می‌شود) بستگی دارد.

این تحقیق همچنین به نیاز معلمان آینده جهت آگاهی از اجرای آزمونهای مختلف و شیوه‌های ارزشیابی و ارزش آزمونهای تکوینی به عنوان روشهای مؤثر و مفید تدریس برای بهبود پیشرفت دانش آموز تأکید می‌گذارد.

منابع

- Akin, M.G. (1969). *Evaluation Theory Development*. Belmont, CA: Wordsworth Publishing Co.
- Anastasi, A. (1988). *Psychological Testing*. (6th ed.) New York: Macmillan, Inc.
- Baker, F.B. (1985). *The Basics of Item Response Theory*, Portsmouth, N.H.: Heinemann Educational Books.
- Biggs, J.B. & Collis, K.F (1982). *Evaluating the Quality of Learning*. (SOLO Taxonomy, structure of the Observed Learning Outcomes). San Diego, CA.: Academic Press, Inc.
- Bloom, S.B. (1956). *Taxonomy of Educational Objectives handbook*. The cognitive Domain, New York: David Mckay Co.
- Bloom, S.B. (1968). *Learning for Mastery*. Evaluation. Comment, 1. UCLA: Center for the study of Evaluation.
- Bloom, S.B. (1976). *Human Characteristics and School Learning*. New York: McGraw Hill Book Company.
- Bloom, S.B. & Hastings (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw Hill Book Inc.
- Chansarkar, B.A. & Routary, U. (1981). *How Relevant is Continuous Assessment?* Assessment in Higher Education. 6, (No.1) 49 - 56.
- Ebel, L.E. & Frisbie, D.A. (1991). *Essentials of Educational Measurement*. Cliffs N.J: Englewood, Prentice Hall.

Feldhusen, J.E. (1964). *Student perceptions of Frequent Quizzes and post - Mortem Discussions of the Tests*. Journal of Educational psychology. 51, (No.1) 51 - 54.

Funchs, L.R. & Others (1985). *Effects of Curriculum-Based Measurement and Consultation on Teacher planning and student Achievement in Mathematics Operations*. Educational Research Journal 28, 617 - 641.

Hopking, C.D. (1987). *Classroom Measuement and Evaluation Itasca*. Ill: F.E. Peacock Rublisher.

Howsome, R.B. (1960). *Who's a Good Teacher? Problems and progress in Teacher Evaluation*. CA: Bulingame. A Special project prepared for and published by joint committee on personnel procedures.

Johnson, A.P. (1988). *Note on a Suggestion Index of Item Validity*. The U-L Index. Journal of Educational psychology. 62, 499 - 504.

Kulik, J.A. & Kulik, C.C. (1990). *Timing of Feedbak and Verbal Learning*. *Review of Educational Research*. 58, 79 - 97.

Kulm, G. editor (1990). *Assessing Higher Order Thinking in Mathematics*. Washington. DC: American Association for Advancement of Science.

Layman, H.B. (1986). *Test scores and what they mean*. Engelwood Cliff, NJ: Prentice Hal.

McMullen, P.M. & Weimer, M.G. (1987). *Student Examination Accentuating the postive*. Education Digest 52, 14 - 17.

Mehrens, A.W. & Lehman, J.I. (1984). *Measurement and Evaluation in Education and psychology*. Holt: Rinehart and Winston Inc.

Paige, D.D. (1966). *Learning while Testing*. The Journal of Educational Research. 59, 276 - 277.

Romberg, R.A., Wilson, L. & Khaketa, M. (1990). *An Examination of state and Foreign tests*. Madison, WI: Center for educational Research.

Romberg, R.A. (1992). *Mathematics Assessment and Evaluation, Imperative for Mathematics*. Albany, NY: state University of New York press.

Ross, C.C. (1974). *Measurement in todays' schools*. Engelwood, cliffs, NJ: Prentice Hall, Inc.

Schutz, R.E. (1970). *The Nature of Educational Development*. *Journal of Research Development in Education*. 3, 39 - 64.

Scriven, M. (1976). *The Methodology of evaluation*. AERA Monograph series on curriculum Evaluation, No.1. Chicago, Ill: Rand McNally.

Stiggins, R.J., Grisowld, M.M. & Wikelund, K.R. (1989). *Measuring Thinking skills through classroom Assessment*. *Journal of Educational Measurement*. 26, 233 - 246.

Stuffelbeam, D. L. & Webster, W.J. (1980). *An analysis of an Alternative Approach to Education*. *Educational Evaluation and Policy Analysis* 2, 5-20.

Van Hiele, P.M. (1986). *Structure and insight: A theory of Mathematic Education*. Orlando, FL: Academic press.

Wilson, M. (1992). *Measuring level of Mathematical Understanding in R.A. Romberg*. Ed. *Mathematics Assessment & Evaluation*. Imperative for Mathematics (213 - 247) Albany, NY: state University of New York press.

کیامنش، علیرضا. ارزشیابی چه هست و چه نیست. نشریه علوم تربیتی دانشگاه تهران، ۱۳۷۰.