

## شناسایی عوامل مؤثر بر تفاوت نمره‌های ارزشیابی مستمر و نهایی هماهنگ پایه یازدهم از دیدگاه دبیران و دانش‌آموزان: مطالعه‌ای به روش پژوهش آمیخته

سمیه غلامی\*

ساره احمدی خانمینی\*\*

ناصر احمدیان\*\*\*

### چکیده

پژوهش حاضر با هدف شناسایی عوامل مؤثر بر تفاوت نمره آزمون‌های مستمر و نهایی در دانش‌آموزان سال یازدهم متوسطه دوم (سوم دبیرستان) از دیدگاه دبیران و دانش‌آموزان انجام گرفت. برای نیل به این هدف از رویکرد پژوهش آمیخته متوالی (کمی سپس کیفی) استفاده شد. جامعه پژوهش در مرحله کمی، دانش‌آموزان سوم دبیرستان شهرستان بستک و توابع و روش نمونه‌گیری در این مرحله سرشماری بود. در مرحله کیفی، جامعه آماری همه دبیران و دانش‌آموزان سال سوم متوسطه دوم بود و ۳۵ دانش‌آموز و ۲۰ دبیر با استفاده از روش نمونه‌گیری هدفمند انتخاب شدند. ابزار پژوهش در بخش کمی، نمره‌های پیشرفت تحصیلی و در بخش کیفی مصاحبه نیمه ساختار یافته بود. تفاوت در حجم دو آزمون، تفاوت در میزان دشواری دو آزمون، تفاوت در شاخص‌های روان‌سنجی دو آزمون، تفاوت در جو روانی اجتماعی دو آزمون، عدم تناسب کیفیت یاددهی-یادگیری، انگیزه و نگرش (عوامل مربوط به دانش‌آموزان) و دستکاری نمره‌ها توسط دبیران از جمله عوامل مؤثر بر تفاوت مشهود نمره‌های مستمر و پایانی بود.

واژگان کلیدی: آزمون مستمر، آزمون پایانی، دبیران، دانش‌آموزان

\* استادیار گروه روان‌شناسی، مجتمع آموزش عالی لارستان (نویسنده مسئول: somayeg@yahoo.com)

\*\* دکتری تخصصی برنامه‌ریزی درسی، دانشگاه علامه طباطبائی

\*\*\* کارشناسی ارشد تحقیقات آموزشی، دانشگاه آزاد اسلامی، واحد بندر لنگه

## مقدمه

ارزشیابی، پدیده‌ای قدمت‌دار و بخشی ذاتی از زندگی روزمره انسان است و ازجمله ملاک‌ها و معیارهایی است که در توسعه آموزش و پرورش نقش مهمی دارد. در عصر کنونی که نظام تعلیم و تربیت و معلمان زیر فشار زیادی برای به موفقیت رساندن دانش‌آموزان هستند، ارزشیابی به‌عنوان ابزاری برای استفاده معلمان در سنجش موفقیت و پیشرفت تحصیلی مورد استفاده قرار می‌گیرد. از ابزارهایی که توسط معلمان یا نظام تعلیم و تربیت برای ارزشیابی مواد درسی و دانش جدید استفاده می‌شود، ارزشیابی پایانی<sup>۱</sup> و ارزشیابی مستمر<sup>۲</sup> است (نهاد پژوهش آموزشی آمریکا،<sup>۳</sup> ۲۰۱۴).

ارزشیابی پایانی، در پایان یک واحد آموزشی یا در یک نقطه خاص از زمان تحصیل از دانش‌آموز به عمل می‌آید و دانش و مهارت یادگیرنده را در برابر استانداردها و معیارها مورد مقایسه قرار داده، میزان تسلط یادگیری را مورد سنجش قرار می‌دهد و به جای تمرکز بر یک فرد خاص به چگونگی پاسخ یک جمعیت توجه می‌کند (استیت، دنریچ و کی‌ورت؛<sup>۴</sup> ۲۰۱۸). از آن جایی که نمره‌های عددی ارزشیابی پایانی قادر به سنجش صحیح پیشرفت تحصیلی دانش‌آموزان نیست، نمی‌توان پایش، کنترل و سنجش پیشرفت یادگیری مفهومی دانش‌آموزان را به ارزشیابی پایانی سپرد؛ بنابراین، طراحی یک ارزشیابی مستمر خوب می‌تواند به معلمان کمک کند تا شکاف ایجاد شده را پر کنند (تریدین، بلعود، بن‌مختار، جورجا و ردید؛<sup>۵</sup> ۲۰۱۵). در واقع، ارزشیابی مستمر که با عناوین دیگری از قبیل ارزشیابی تکوینی<sup>۶</sup>، ارزشیابی مرحله‌ای، ارزشیابی بازخورد، ارزشیابی فرایند و نظایر آن مشابهت دارد، بخش جدایی‌ناپذیری از فرایند یاددهی-یادگیری است که به بهبود و رفع مشکلات فرایند مذکور منجر می‌شود. این نوع ارزشیابی، عبارت از فعالیت‌های معلم و فراگیران در فرایند ارزیابی است تا اطلاعاتی را به‌عنوان بازخورد فراهم آورده و در صورت لزوم فعالیت‌های آموزش و یادگیری را تغییر دهند (بلک و ویلیام،<sup>۷</sup> ۲۰۰۹). در حقیقت، ارزشیابی مستمر سنجشی است که در

1. Final

2. Continuous

3. American Educational Research Association

4. States, Detrich &amp; Keyworth

5. Tridane, Belaaouad, Benmokhtar, Gourja, &amp; Radid

6. formative assessment

7. Black &amp; william

حین فرایند یادگیری به منظور بهبود آموزش و یادگیری صورت می‌گیرد (سیزیک، ۱۷-۲۰۱۰، ۳) و نوعی بازخورد کیفی (فراتر از نمره) برای دانش‌آموزان و معلمان است که بر جزئیات محتوا و عملکرد تأکید دارد (هوآتا، ۲۰۱۰).

مطالعات فراتحلیل حوزه ارزشیابی مستمر نشان می‌دهند که ضریب تأثیر<sup>۳</sup> این ارزشیابی در بهبود یادگیری بسیار زیاد است (هاتی، ۲۰۰۹). البته وب و همکارانش<sup>۴</sup> (۲۰۱۸) معتقدند که زمانی این ضریب تأثیر اتفاق می‌افتد که ارزشیابی مستمر گام‌هایی متوالی در جهت رسیدن به ارزشیابی پایانی باشد و هر دو ارزشیابی پیرو مسیر مشترکی در هدف‌ها و برنامه‌ریزی باشند. درواقع، ترکیبی از هر دو نوع آزمون ضروری است تا اینکه معلمان بتوانند با استفاده از آزمون مستمر روند یادگیری دانش‌آموزان را رصد کنند و همچنین با استفاده از آزمون پایانی از این اطمینان حاصل کنند که آیا آزمون مستمر به درستی انجام شده است یا خیر (گلزر، ۲۰۱۴). گاهی نمره‌های ارزشیابی مستمر باعث بهبود عملکرد پایانی نمی‌شود و این موضوع باعث شده که برخی از پژوهشگران چون اسوافیلد<sup>۵</sup> (۲۰۱۱) اظهار کنند که باید بین دو واژه ارزشیابی تکوینی و ارزشیابی برای یادگیری<sup>۶</sup> تمایز قائل شد. در حالی که پژوهشگران دیگری مانند بنت<sup>۷</sup> (۲۰۱۱)، بلک و ویلیام (۲۰۰۹) و برد، هوفنیک، نیوتن، استوبارت و استین-اوتیم<sup>۸</sup> (۲۰۱۴) این دو واژه را هم معنا می‌دانند. این بحث توسط بلک و ویلیام (۲۰۰۹) بدین صورت جمع‌بندی شده که به جای تعریف واژه‌های متفاوت، ما باید بدانیم که چه راهبردهایی در ارزشیابی تکوینی به پیشرفت تحصیلی منجر می‌شود. بنابراین در مطالعات بر راهبردهای اثربخش در ارزشیابی تکوینی تمرکز شد. برای مثال؛ هاتی و تیمپرلی<sup>۹</sup> (۲۰۰۷) بر تأثیر ارائه

1. Cizek

2. Huhta

3. Effect size

4. Hattie

5. Webb et al

6. Glazer

7. Swaffield

8. assessment for learning

9. Bennett

10. Baird, Hopfenbeck, Newton, Stobart, & Steen-Utheim

11. Hattie & Timperley

بازخورد از طریق ارزشیابی مستمر و دیگناس و باتنر<sup>۲</sup> (۲۰۰۸) بر یادگیری خودتنظیمی و برنامه‌ریزی گام به گام تا رسیدن به هدف به‌عنوان عامل اثربخشی ارزشیابی مستمر تأکید کردند. اندرسون و پالم<sup>۳</sup> (۲۰۱۸) معتقدند تا وقتی که همه استراتژی‌ها در هم ادغام شود و در کلاس درس به کار بروند، می‌توان انتظار داشت که ارزشیابی مستمر به ارتقای عملکرد تحصیلی پایانی منجر شود.

در ایران از سال تحصیلی ۱۳۸۱-۱۳۸۰ به منظور افزایش پایداری و کاربست آموخته‌ها در حل مسائل زندگی، ارزشیابی مستمر در همه مقاطع تحصیلی، کنار ارائه نمره‌های پایانی هماهنگ مورد توجه قرار گرفت. اگرچه با بررسی آیین‌نامه‌های امتحانات می‌توان فهمید استفاده از ارزشیابی مستمر از سال ۱۳۳۱ در آیین‌نامه امتحانات تصویب شده بود اما به‌ندرت از آن استفاده می‌شد. همان‌گونه که در حال حاضر در آیین‌نامه ارزشیابی از هر سه نوع آزمون ورودی، تکوینی، پایانی و انواع ارزشیابی‌های کتبی، شفاهی و عملی صحبت شده است اما در عمل از نوع تکوینی بر اساس میزان مشارکت دانش‌آموزان و پایانی برای ارتقای دانش‌آموزان استفاده می‌شود که آن هم به صورت کتبی است. این عدم تطابق آیین‌نامه و عمل، نظام ارزشیابی ما را هم از هدف‌های ارزشیابی خود و هم از ارزشیابی سایر کشورها (آلمان و انگلستان) دور می‌کند (کاظمی، ۱۳۹۰).

بنا بر اعلام مرکز سنجش وزارت آموزش و پرورش در سال ۱۳۹۴ میانگین نمره‌های دانش‌آموزان در آزمون‌های نهایی (کلاس یازدهم) در سه سال اخیر در حدود دوازده بود. به گزارش این مرکز، علت قبولی دانش‌آموزان علی‌رغم نمره‌های پایین در امتحانات نهایی این است که نمره مستمر طول سال با نمره امتحانات نهایی جمع می‌شود و نمره قبولی دانش‌آموزان را می‌سازد. در تأیید این آمار، میانگین نمره‌های دانش‌آموزان شهرستان بستک در سال‌های ۱۳۹۳ تا ۱۳۹۶ در حدود ۱۲/۲۱ است که این میانگین نمره‌ها نگران‌کننده به نظر می‌رسد (اداره کل آموزش و پرورش استان هرمزگان، ۱۳۹۵). بر اساس اطلاعات گردآوری‌شده، در رشته تجربی میانگین نمره‌های مستمر ۱۶/۵۸ و نمره‌های پایانی کتبی ۱۱/۷۷، در رشته علوم انسانی میانگین نمره‌های مستمر ۱۵/۱۵ و

1. Feedback

2. Dignath & Büttner

3. Andesson & Palm

نمره‌های پایانی کتبی ۱۰/۵۰ و در رشته ریاضی میانگین نمره‌های مستمر ۱۶/۵۲ و نمره‌های پایانی کتبی ۱۳/۰۵ است (ادارات آموزش و پرورش بستک و جناح، ۱۳۹۶). اگر نتایج آزمون‌های نهایی و استاندارد نشان‌دهنده عملکرد دانش‌آموزان باشد این عملکرد ضعیف آموزشی به‌مثابه زنگ خطری است که عوامل زیادی در به وجود آمدن آن سهیم‌اند. در نگاه اول، هم‌سطح نبودن آزمون‌های معلم‌ساخته با آزمون نهایی می‌تواند نشانگر نارسایی‌های نظام ارزشیابی در مدرسه باشد. بدین معنی که بخشی از مشکلات ساختاری نظام آموزشی ناشی از شکاف ایجادشده در ناهماهنگی فهمی و عملکردی بین طراحان محتوای آموزشی و دبیران است که ناکارآمدی نظام آموزشی را برجسته می‌کند اما به‌طورکلی، حل مشکلات نیازمند فهم عمیقی از دلایل ایجادکننده هست. به همین منظور، در پژوهش حاضر تلاش شد ابعاد مسئله با روش‌های ترکیبی کنکاش شود؛ بدین صورت که از تکنیک‌های کمی و کیفی در مراحل از پژوهش به شیوه متوالی<sup>۱</sup> یا هم‌زمان<sup>۲</sup> استفاده شد (کرسول و پلانو کلارک،<sup>۳</sup> ۲۰۰۷ و جانسون و کریستینسن؛<sup>۴</sup> ۲۰۰۸). بر اساس این روش‌شناسی، تبیین و تفهم هم‌زمان پیچیدگی‌ها، ابعاد، ساختار و اشکال پدیده‌ها و واقعیات اجتماعی، استفاده از طرح‌ها، روش‌ها، نظریه‌ها و منابع داده‌های گوناگون یک ضرورت است، زیرا همان‌طور که برور و هانتز<sup>۵</sup> (۲۰۰۶) عنوان کرده‌اند: «هیچ روشی به تنهایی کامل نیست». در این خصوص، تلاش شد در گام اول علاوه بر بررسی نمره‌ها به روش کمی، کرانه‌های بالا و پایین نمره‌ها بررسی شود و در ادامه با استفاده از رویکرد مبتنی بر روش‌شناسی تفسیرگرایی، ضمن تحلیل تجربه زیسته معلمان و ادراک دیدگاه‌های آنها به‌عنوان مهم‌ترین عنصر تأثیرگذار در هدایت فرایند یاددهی- ارزشیابی و معمار اصلی نظام آموزشی بازنمایی شود. همچنین از آنجا که اغلب معلمان خود آزمون‌های مستمر را طراحی می‌کنند، بررسی دلایل این تفاوت نمره‌ها از دیدگاه این گروه به درک عمیق‌تری منجر خواهد شد. علاوه بر این، دانش‌آموزان نیز به دلیل افرادی که مورد ارزشیابی قرار می‌گیرند، می‌توانند دلایل خود برای کسب دو نمره متفاوت را به خوبی بیان کنند. کنکاش چرایی این رویداد به

1. Concurrent

2. Sequential

3. Creswell & Plano Clark

4. Johnson & Christensen

5. Brewer & Hunter

درک بهتر کیفیت نظام ارزشیابی و در چشم‌اندازی وسیع‌تر به پایش و فهم بهتر فرایندهای یاددهی و یادگیری منجر می‌شود. نتایج این پژوهش، ابزاری فراهم می‌کند تا تصمیم‌سازان و سیاست‌گذاران آموزشی در درک بهتر واقعیت‌های نظام ارزشیابی و کیفیت آموزشی توانمند شوند. بنابراین پرسش‌های این پژوهش به شکل زیر مطرح شده است:

- ۱- آیا بین نمره‌های مستمر و نمره‌های نهایی تفاوت وجود دارد؟
- ۲- تفاوت بین نمره‌های مستمر و نمره‌های نهایی از دید دبیران ناشی از چیست؟
- ۳- تفاوت بین نمره‌های مستمر و نمره‌های نهایی از دید دانش‌آموزان ناشی از چیست؟

### روش‌شناسی پژوهش

در این مطالعه، از روش پژوهش آمیخته متوالی تبیینی<sup>۱</sup> استفاده شده است. یکی از ویژگی‌های این روش، توالی استفاده از روش‌های پژوهش کمی و کیفی است. در این روش در مرحله اول، داده‌های کمی و سپس در مرحله دوم، داده‌های کیفی گردآوری و تحلیل می‌شوند. در نهایت، هر دو تحلیل کمی و کیفی یکجا مورد تفسیر قرار می‌گیرند (تشکری و تدلی،<sup>۲</sup> ۲۰۰۳ و محمدپور، صادقی و رضایی، ۱۳۸۹). در پژوهش حاضر، در مرحله اول (بخش کمی)، با مراجعه به اداره آموزش و پرورش، نمره‌های همه دانش‌آموزان به تعداد ۵۳۰ نفر از بخش آمار نمره‌های مستمر و پایانی رشته‌های انسانی، تجربی و ریاضی دانش‌آموزان سوم متوسطه دوم شهرستان بستک، گردآوری شد. در ادامه، تفاوت معنی‌دار میانگین نمره‌های مستمر و نهایی با استفاده از آزمون تی وابسته مشخص شد (جدول ۴) که مبنای اجرای مرحله بعد پژوهش شد. لازم به ذکر است که در بخش کمی، روش نمونه‌گیری سرشماری بود.

در بخش کیفی، برای تعیین حجم نمونه دانش‌آموزان و معلمان از روش نمونه‌گیری هدفمند بر مبنای اشباع نظری استفاده شد، بدین صورت که تعداد ۳۵ نفر (۱۷ پسر و ۱۸ دختر) از دانش‌آموزان سه رشته انسانی، تجربی و ریاضی که نمره آنها در دو طرف کرانه یعنی در مستمر گروه بالا و در پایانی جزء گروه پایین بودند، به صورت هدفمند

1. Sequential-Explanatory Mixed Methods Design

2. Tashakkori & Teddlie

انتخاب شدند. با ۲۰ نفر (۱۲ مرد و ۸ زن) از معلمان مطلع از آسیب‌های مدرسه‌های بستک که مایل به همکاری بودند، مصاحبه ترتیب داده شد. ابزار پژوهش، مصاحبه نیمه ساختار یافته بود که برای دستیابی به اعتبار پرسش‌های مصاحبه از نظرات استادان صاحب‌نظر استفاده شد. برای حصول اطمینان از دقت کار نیز از روش مطالعه مکرر، مقایسه مستمر داده‌ها، خلاصه‌سازی و دسته‌بندی اطلاعات بدون اینکه داده‌ها آسیبی ببینند، استفاده شد. برای تجزیه و تحلیل داده‌ها در این بخش از پدیدارشناسی استفاده شد و تحلیل یافته‌ها به روش دیکلمن<sup>۱</sup> (۲۰۰۱) صورت گرفت. در این روش، برای رسیدن به هدف نهایی هفت مرحله طی می‌شود؛ در مرحله نخست، در پایان هر مصاحبه و ثبت یادداشت‌برداری‌های میدانی، ابتدا گفته‌های ضبط شده شرکت‌کنندگان چندین بار شنیده می‌شود، به صورت متن نوشته و برای فهم احساس و درک کلی از تجارب شرکت‌کنندگان چند بار مطالعه می‌شود. در مرحله دوم، خلاصه تفسیری از هر مصاحبه تدوین می‌شود. در سومین مرحله مصاحبه یا متن نوشته شده منتخب تحلیل می‌شود. در مرحله بعد در صورت بروز تناقض‌های تفسیری برای حل مسئله دوباره به متن رجوع می‌شود. در مرحله پنجم، مقوله‌ها<sup>۲</sup> یا مضمون‌هایی که نمایانگر معنی مشترک و تمرینات به اشتراک گذاشته شده است، تعیین شده و پس از آن روابط بین مضمون‌ها مشخص می‌شود. در پایان، پیش‌نویسی از مضمون‌ها به همراه مثال از متن برای به هم پیوند یافتن پاسخ‌ها و پیشنهادها در پیش‌نویس پایانی ارائه می‌شود (جدول‌های ۵ و ۶). در این مطالعه، برای دستیابی به اعتبار پژوهش جنبه‌های مختلفی در نظر گرفته شد. استاد راهنمای پژوهش به عنوان بازرس خارجی<sup>۳</sup> در تمامی مراحل پژوهش، نظارت و راهنمایی داشت. در روش کیفی، از روش‌های دیگر اعتبارسنجی، اعتبار از طریق اعضا است. در این پژوهش نیز نتایج پژوهش به مصاحبه‌شوندگان برگردانده شد و سپس با نظر آنها مطابقت داده شد.

در ادامه، خلاصه مشخصات دانش‌آموزان (جدول ۱) و خلاصه اطلاعات معلمان (جدول ۲) ارائه شده است.

1. Diekelmann

2. Theme

3. External Audits

جدول (۱) مشخصات دانش‌آموزان شرکت‌کننده در پژوهش

| پایه   | تعداد | جنسیت | رشته   |
|--------|-------|-------|--------|
| یازدهم | ۸     | دختر  | تجربی  |
| یازدهم | ۶     | پسر   | تجربی  |
| یازدهم | ۵     | دختر  | ریاضی  |
| یازدهم | ۴     | پسر   | ریاضی  |
| یازدهم | ۶     | دختر  | انسانی |
| یازدهم | ۷     | پسر   | انسانی |

جدول (۲) مشخصات معلمان شرکت‌کننده در پژوهش

| تعداد | جنسیت | درس تدریسی   | تحصیلات    | درس مورد تدریس        | سابقه تدریس |
|-------|-------|--------------|------------|-----------------------|-------------|
| ۱     | مرد   | فیزیک        | فوق لیسانس | فیزیک                 | ۲۴          |
| ۱     | زن    | فیزیک        | لیسانس     | فیزیک                 | ۹           |
| ۱     | مرد   | شیمی         | لیسانس     | شیمی                  | ۱۹          |
| ۱     | زن    | شیمی         | لیسانس     | شیمی                  | ۱۴          |
| ۱     | مرد   | ریاضی        | فوق لیسانس | ریاضی-انتگرال         | ۱۱          |
| ۱     | مرد   | ریاضی        | لیسانس     | ریاضی                 | ۶           |
| ۱     | زن    | ریاضی        | فوق لیسانس | ریاضی                 | ۱۷          |
| ۱     | زن    | ریاضی        | لیسانس     | ریاضی                 | ۲۹          |
| ۱     | مرد   | ادبیات فارسی | لیسانس     | ادبیات-فلسفه          | ۱۵          |
| ۱     | زن    | ادبیات فارسی | لیسانس     | ادبیات                | ۱۳          |
| ۱     | مرد   | ادبیات فارسی | فوق لیسانس | ادبیات                | ۲۷          |
| ۱     | مرد   | زیست‌شناسی   | لیسانس     | زیست                  | ۱۴          |
| ۱     | زن    | تاریخ        | فوق لیسانس | تاریخ-مطالعات اجتماعی | ۱۲          |
| ۱     | مرد   | الهیات       | لیسانس     | دینی-عربی-فلسفه       | ۲۴          |
| ۱     | زن    | الهیات       | لیسانس     | دینی-عربی             | ۱۷          |
| ۱     | مرد   | کامپیوتر     | لیسانس     | ریاضی-کامپیوتر        | ۱۳          |
| ۱     | مرد   | علوم اجتماعی | لیسانس     | مطالعات-تاریخ         | ۱۴          |
| ۱     | مرد   | زبان انگلیسی | فوق لیسانس | زبان انگلیسی          | ۲۷          |
| ۱     | مرد   | جغرافیا      | فوق لیسانس | جغرافی-زمین‌شناسی     | ۲۷          |
| ۱     | زن    | روان‌شناسی   | فوق لیسانس | روان‌شناسی            | ۲۸          |



یافته‌ها

۱- نتیجه مطالعه کمی

پرسش اول: آیا بین نمره‌های مستمر و پایانی دانش‌آموزان تفاوت وجود دارد؟ در ابتدا برای ارائه درک روشنی از مسئله، میانگین نمره‌های مستمر و پایانی در جدول (۳) ارائه شده است.

جدول (۳) میانگین نمره‌های مستمر و پایانی در سه رشته نظری

| رشته   | میانگین نمره مستمر | میانگین نمره کتبی ورقه | تفاوت نمره‌های مستمر و پایانی |
|--------|--------------------|------------------------|-------------------------------|
| تجربی  | ۱۶/۵۸              | ۱۱/۷۷                  | ۴/۸۱                          |
| انسانی | ۱۵/۱۵              | ۱۰/۵۰                  | ۴/۶۵                          |
| ریاضی  | ۱۶/۵۲              | ۱۳/۰۵                  | ۳/۴۷                          |
| جمع    | ۱۶/۰۸              | ۱۱/۷۸                  | ۴/۳۰                          |

همان‌گونه که در جدول (۳) دیده می‌شود در رشته تجربی، میانگین نمره‌های مستمر ۱۶/۵۸ و لی نمره‌های پایانی کتبی ۱۱/۷۷ است که نشان از تفاوتی به مقدار ۴/۸۱ دارد. در رشته علوم انسانی، میانگین نمره‌های مستمر ۱۵/۱۵ و لی نمره‌های پایانی کتبی ۱۰/۵۰ است که بازهم کاهش نمره‌های پایانی در حدود ۴/۶۵ است. در رشته ریاضی، میانگین نمره‌های مستمر ۱۶/۵۲ و لی نمره‌های پایانی کتبی ۱۳/۰۵ است که نشان از افت عملکرد در حدود ۳/۴۷ دارد. تفاوت بین نمره‌ها به گونه‌ای است که نمره‌های مستمر نمی‌توانند بازگوکننده ارزیابی‌های واقعی به نسبت نمره‌های پایانی باشند. برای بررسی معنی‌داری تفاوت، آزمون تی وابسته انجام گرفت که خلاصه نتایج در جدول (۴) گزارش شده است.

جدول (۴) نتایج آزمون  $t$  وابسته و ضریب همبستگی مستمر و نهایی

| رشته تحصیلی | نوع ارزیابی | میانگین | انحراف معیار | t      | df  | سطح معنی‌داری | ضریب همبستگی     |
|-------------|-------------|---------|--------------|--------|-----|---------------|------------------|
| تجربی       | مستمر       | ۱۶/۵۸   | ۱/۲          | ۱۱۲/۴  | ۳۴۹ | ۰/۰۰۰۱        | ۰/۹۵*            |
|             | پایانی      | ۱۱/۷۷   | ۲/۰۲         |        |     |               | * $P \leq ۰/۰۰۱$ |
| ریاضی       | مستمر       | ۱۶/۵۲   | ۲/۳          | ۱۰۹/۶  | ۳۴۹ | ۰/۰۰۰۱        | ۰/۹۸*            |
|             | پایانی      | ۱۳/۰۵   | ۴/۳          |        |     |               | * $P \leq ۰/۰۰۱$ |
| انسانی      | مستمر       | ۱۵/۱۵   | ۱/۶          | ۱۵۶/۶۴ | ۳۴۹ | ۰/۰۰۰۱        | ۰/۸۹*            |
|             | پایانی      | ۱۰/۵    | ۲/۴          |        |     |               | * $P \leq ۰/۰۱$  |

نتایج آزمون  $t$  نشان داد که بین نمره‌های مستمر و پایانی در رشته تحصیلی ریاضی ( $t=۱۰۹/۶$ ،  $P=۰/۰۰۱$ )، در رشته تجربی ( $t=۱۱۲/۴$ ،  $P=۰/۰۰۱$ ) و در رشته انسانی ( $t=۱۵۶/۶$ ،  $P=۰/۰۰۱$ ) تفاوت معنی‌داری وجود دارد. ضریب همبستگی نمره‌های مستمر و نهایی در رشته تجربی، ریاضی و انسانی به ترتیب  $۰/۹۵$ ،  $۰/۹۸$  و  $۰/۸۹$  مثبت و معنی‌دار است.

#### ۱- تحلیل کیفی

پرسش دوم: تفاوت بین نمره‌های مستمر و نهایی از دید دبیران در چیست؟ پس از ارزیابی مصاحبه‌ها و تبدیل محتوای فایل‌های صوتی مصاحبه‌ها به محتوای نوشتاری، مضامین استخراج شد و با مقایسه مستمر بین داده‌ها مقوله‌هایی که با یکدیگر هم‌پوشانی داشتند در یک عنوان قرار گرفتند. در نهایت، از تحلیل یافته‌های به دست آمده در این پژوهش، ۷ مضمون اصلی و ۱۳ مضمون فرعی درباره علل تفاوت بین نمره‌های مستمر و پایانی از دید دبیران پدیدار شده است که می‌توانند تجربه و ادراکات دبیران از پدیده تفاوت بین آزمون مستمر و پایانی را به تصویر بکشند. به مضامین اصلی و فرعی این تفاوت‌ها در جدول (۵) اشاره شده است.

جدول (۵) مضمون‌های اصلی و فرعی از نظر دبیران

| مضامین اصلی حاصل از مصاحبه با دبیران                              | مضامین فرعی  |
|---|--|
| تفاوت در حجم مطالب مورد آزمون                                     | - نسبت کم حجم مطالب درسی در آزمون مستمر<br>- نسبت زیاد حجم مطالب درسی در آزمون پایانی                                  |
| تفاوت در میزان دشواری آزمون                                       | - اصول نمره‌گذاری متفاوت<br>- تفاوت در نوع آزمون (استاندارد نبودن آزمون مستمر)   |
| تفاوت در شاخص‌های روان‌سنجی آزمون                                 | - روایی - پایایی   |
| دستکاری نمره‌های مستمر توسط دبیران                                | - دخالت دیگر عوامل در ثبت نمره مستمر<br>- تأثیر ارزش ضرایب<br>- مصلحت‌گرایی آموزشی                                     |
| عدم شناخت فضای روانی آزمون نهایی (تفاوت جو روانی - اجتماعی آزمون) | - سنجش رشد عاطفی، روانی، هیجانی و اجتماعی<br>- آسودگی زیست آزمون   |
| عدم تناسب کیفیت یاددهی - یادگیری                                  | - نا برخورداری آموزشی<br>- معنی‌داری و پایداری در یادگیری  |
| انگیزه و نگرش دانش‌آموزان   | انگیزش بیشتر در امتحانات مستمر به دلیل پوشش دادن انواع روش‌های سنجش چون عملی، شفاهی و ...<br>نگرش منفی به امتحان نهایی |

### بررسی مضامین محوری حاصل از مصاحبه دبیران

#### ۱- تفاوت در حجم مطالب مورد آزمون

با توجه به موارد مطرح شده در ارتباط با مضمون اصلی (۱)، دو مضمون فرعی شامل حجم مطالب درسی کم در آزمون مستمر و حجم مطالب درسی زیاد در آزمون پایانی به دست آمد که به تحلیل آنها اشاره می‌شود.

#### ۱-۱- نسبت کم حجم مطالب درسی در آزمون مستمر: از آنجایی که آزمون‌های

مستمر در طول سال تحصیلی و معمولاً پس از اتمام هر فصل یا مبحث درسی از دانش‌آموزان به عمل می‌آید، حجم کمی از محتوا را شامل می‌شود و دانش‌آموزان زمان کافی را برای مطالعه در اختیار دارند و در نهایت نمره‌های بالاتری را کسب می‌کنند. در ادامه اظهارات چند تن از دبیران ذکر می‌شود.

پاسخ‌دهنده شماره ۹: «حجم کم مطالب درسی در آزمون مستمر و همچنین زمان مناسب و کافی برای مطالعه باعث می‌شود دانش‌آموز بر موضوع خاصی تمرکز داشته

باشد و بتواند نمرات بالایی کسب کند؛ در حالی که در آزمون پایانی چنین نیست». پاسخ‌دهنده شماره ۱۸: «در آزمون مستمر دروس تقسیم‌بندی شده و فصل به فصل آزمون برگزار می‌شود، بالطبع آن سوالات کمتری را شامل می‌شود و این خود باعث می‌شود دانش‌آموزان نمرات بالاتری بگیرند».

**۲-۱- نسبت زیاد حجم مطالب درسی در آزمون پایانی:** آزمون پایانی یا نهایی در پایان سال تحصیلی به صورت هماهنگ کشوری از کل کتاب درسی تهیه می‌شود و بر خلاف آزمون مستمر، حجم زیادی از محتوا را شامل می‌شود و کسب نمره بالا مستلزم برقراری ارتباط بین همه موضوعات و فصل‌های کتاب درسی است، به‌گونه‌ای که دانش‌آموز به سطح یادگیری در حد تسلط رسیده باشد. یکی از دلایلی که مانع کسب نمره‌های بالا توسط فراگیران در آزمون نهایی می‌شود، حجم زیاد محتوا و ناتوانی دانش‌آموزان در برقراری ارتباط بین موضوعات و قسمت‌های مختلف کتاب درسی است. به دو نمونه از اظهارات دبیران در این خصوص اشاره می‌شود.

پاسخگوی شماره ۹: «وقتی که حجم مطالب در آزمون پایانی زیاد شود به طبع آن فشار روانی بیشتر بر دانش‌آموز وارد می‌شود؛ همین اضطراب و فشار باعث می‌شود که نمرات پایانی ریزش داشته باشد». پاسخ‌دهنده شماره ۷: «حجم مطالب زیاد در یک بازه زمانی اندک در امتحانات نهایی باعث می‌شود که اکثر دانش‌آموزان مطالب را قاطی کنند و فقط آن دسته از دانش‌آموزانی که به صورت مفهومی درس را یاد گرفته‌اند، موفق هستند».

**۲- تفاوت در میزان دشواری آزمون:** با توجه به موارد مطرح شده در ارتباط با مضمون اصلی (۲)، دو مضمون فرعی شامل اصول نمره‌گذاری متفاوت و تفاوت در نوع آزمون به دست آمد که در ذیل به تحلیل آنها اشاره شده است.

**۲-۱- اصول نمره‌گذاری متفاوت:** اساساً اصول نمره‌گذاری در آزمون‌های مستمر و پایانی متفاوت است. در نمره‌گذاری مستمر مجموعه‌ای از عوامل و فاکتورها مانند حضور و غیاب، فعالیت‌های فردی، گروهی، انجام کارهای پژوهشی، اخلاق و انضباط نقش دارند؛ در حالی که نمره‌های پایانی همانی است که به صورت کتبی از ستاد تصحیح اوراق به مدارس ابلاغ می‌شود و هیچ عامل دیگری در آن دخل و تصرف ندارد. برای روشن‌تر شدن موضوع به دو مورد از اظهارات دبیران در این رابطه اشاره می‌شود.

پاسخ‌دهنده شماره ۶: «نمرات پایانی و مستمر متفاوت است چون معیار آزمون نهایی صددرصد کمی ولی مستمر تا حدودی کیفی است و می‌توان گفت ترکیبی از

کمی و کیفی». پاسخگوی شماره ۱: «به علت تفاوت در تعریف مستمر و پایانی تفاوت نمرات این دو آزمون تا حدودی منطقی است ولی نباید از حد معینی بگذرد. نمرات مستمر براساس پرسش‌های شفاهی، آزمونک‌ها، تحقیق، فعالیت‌های کلاسی، حضور فعال در سر کلاس و ... داده می‌شود اما نمره پایانی فقط یک نمره کتبی روی ورقه امتحان نهایی است».

**۲-۲- تفاوت در نوع آزمون (استاندارد نبودن پرسش‌های آزمون مستمر):** سطح، شکل و نوع پرسش‌ها از دیگر نکاتی است که دبیران به آن اشاره می‌کردند و آن را دلیلی بر اختلاف نمره‌های پایانی و مستمر می‌دانستند.

پاسخ‌دهنده شماره ۷: «نوع سؤالات پایانی با مستمر متفاوت است، سؤالات پایانی به نوعی هنجاریابی شده به کرات چندین نوع سؤال روی دانش‌آموزان پیاده شده و میانگینی از این نوع سؤالات فراهم شده است و هر سه سطح دشواری پایین، متوسط و عالی را در بر می‌گیرد». پاسخ‌دهنده شماره ۵: «آزمون نهایی همه انواع سؤالات را در بر می‌گیرد از تکمیل کردنی گرفته تا صحیح و غلط، چندگزینه‌ای، کوتاه پاسخ و بازپاسخ همه را شامل می‌شود ولی در آزمون‌های مستمر این موارد رعایت نمی‌شود». پاسخگوی شماره ۱۲: «طراحی سؤالات استاندارد، تصحیح اوراق و دادن بازخورد به دانش‌آموزان وقت‌گیر و زمان‌بر است. در آزمون مستمر ما به چند سؤال مختصر بسنده می‌کنیم».

**۳- تفاوت در شاخص‌های روان‌سنجی آزمون:** متخصصان اندازه‌گیری و ارزشیابی برای آزمون‌های مختلف، ویژگی‌های زیادی در نظر گرفته‌اند. مهم‌ترین این ویژگی‌ها روایی و پایایی هستند.

**۳-۱- روایی آزمون:** روایی مربوط است به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است؛ بنابراین، آزمونی دارای روایی است که برای اندازه‌گیری آنچه مورد نظر است کافی و مناسب باشد؛ به عبارت دیگر، روایی آزمون یعنی آیا آزمون آنچه را که قرار است بسنجد، می‌سنجد. از نظر روایی محتوایی آزمون‌های مستمر با توجه به اینکه بخش‌های کمتری از محتوا را در بردارد می‌تواند معرف جامعه پرسش‌های بیشتری باشد و آزمون پایانی با توجه به محدودیت پرسش‌ها و حجم زیاد کتاب از پوشش کلیت جامعه پرسش‌های ممکن کمتری برخوردار است. با توجه به روایی ملاکی میزان ارتباط بین نمره‌های حاصل از یک آزمون با نمره‌های به دست آمده از آزمون دیگر در آزمون‌های مستمر و پایانی از شدت پراکندگی کم و بیش شبیه هم برخوردار هستند.

پاسخ‌دهنده شماره ۱۴: «آزمون پایانی بیشتر بخش مربوط به دانش را می‌سنجد اما آزمون مستمر علاوه بر دانش، مهارت، نگرش و سایر فعالیت‌های دانش‌آموز را می‌سنجد».

**۲-۳- پایایی آزمون:** پایایی به دقت، اعتمادپذیری، ثبات یا تکرارپذیری نتایج آزمون اشاره دارد. اگر ارزشیابی مستمر به شیوه‌های مختلف (شفاهی، عملی و آزمونک‌های کتبی) در طول سال تحصیلی و از مطالب تدریس شده گرفته شود و نمره‌های پایانی هر دانش‌آموز بر اساس نمره‌های طول سال داده شود این نمره‌ها از اعتبار و ثبات بهتری برخوردار می‌شوند.

پاسخ‌دهنده شماره ۱۰: «از آنجایی که در ارزشیابی مستمر دست دبیر در گرفتن نوع آزمون باز است و بیشتر از نوع کیفی است بدون شک از اعتبار کمتری نسبت به آزمون پایانی برخوردار است».

**۴- دستکاری نمره‌های مستمر توسط دبیران:** در رابطه با نمره‌های مستمر در ارزشیابی‌های مستمر، مهم‌ترین نکته مستندسازی است. بدین‌گونه که دبیران در رابطه با نمره مستمر اسناد و مدارکی (چک‌لیست، گزارش فعالیت‌ها، کارپوشه، امتحانات کلاسی و ...) دال بر فعالیت‌های دانش‌آموز در طول سال تحصیلی ارائه دهند، در حالی که چنین نیست و نمره‌های مستمر با آزادی عمل و اعمال نظرات شخصی معلمان و سلیقه‌ای عمل کردن آنان همراه است. از این مضمون اصلی، سه مضمون فرعی شامل دخالت دیگر عوامل در ثبت نمره‌های مستمر، تأثیر ارزش ضرایب و مصلحت‌گرایی آموزشی از دیدگاه دبیران استخراج شده است.

**۴-۱- دخالت دیگر عوامل در ثبت نمره مستمر:** در حال حاضر موضوعی که به یک معضل تبدیل شده و نظام ارزشیابی ما را به چالش کشیده است مسئله‌ای به نام «درصد قبولی» است. تلاش برای بالا بردن درصد قبولی با دادن نمره‌های مستمر بالا و بدون دلیل به این امید که دانش‌آموز در آینده جبران کند به رویه‌ای نامناسب در نظام آموزشی ایران تبدیل شده است که اثرات نامطلوب و نگران‌کننده آن در سال‌های نه‌چندان دور آینده خود را نشان خواهد داد. در این بخش، بیشتر دبیران تأکید بر درصد قبولی از طرف عالی‌ترین مقام گرفته تا مدیر مدرسه را مهم‌ترین چالش‌ها و معضلات نظام ارزشیابی مستمر می‌دانند که آثار مخربی را به دنبال دارد. در ادامه به تجربیات چند تن از دبیران اشاره می‌شود.

پاسخگوی شماره ۱۱: «از اول سال تحصیلی فرم قرارداد درصد قبولی و میانگین بین مدیر مدرسه و دبیران مربوطه بسته می‌شود و از دبیران تضمین افزایش درصد و میانگین می‌خواهند بدون توجه به اینکه چه دانش‌آموزانی و با چه سطحی سر کلاس حضور دارند و این به نظر غیرمنطقی می‌رسد. اصلاً تقدیرنامه‌ها بر مبنای افزایش درصد قبولی و میانگین صادر می‌شود و تلاش معلم در طول سال تحصیلی نادیده گرفته می‌شود». پاسخگوی شماره ۷: «ارزیابی عملکرد دبیران یک بعدی شده و صرفاً با نمره دبیر را ارزیابی می‌کنند. توجه صرف به بعد شناختی و ارزیابی دبیر با درصد قبولی فشار روانی زیادی بر دبیر وارد می‌کند و دبیر برای فرار از زیر این فشار دست نمره را بالا می‌گیرد تا اینکه مؤاخذه نشود». پاسخگوی ۱۲: «دانش‌آموزسالاری به حد افراط رسیده، دانش‌آموز به معلم می‌گوید اگر نمره مستمر من کم داده شود من دیگه درس نمی‌خوانم و این برای شمای معلم ضرر دارد. از طرف دیگر رقابت بین مدارس باعث شده است مدیران بجای اجرای برنامه‌های آموزشی مناسب، به دبیران تأکید کنند که نمرات مستمر را بالاتر از نمره واقعی بدهند». پاسخ‌دهنده شماره ۲۰: «تأکید زیاد بر میانگین و درصد قبولی باعث شده است متأسفانه معلمان از درج نمرات واقعی در لیست مستمر خودداری نمایند و نمره‌ای به مراتب بالاتر را جایگزین نمره واقعی نمایند، در واقع در آزمون مستمر امکان دستکاری نمرات نسبت به پایانی به مراتب بیشتر است. اگر مستمر واقعی داده شود، چه بسا شکاف عمیق بین واقعیت و ظواهر موجود آشکار خواهد شد».

۴-۲- تأثیر ارزش ضرایب: یکی از مهم‌ترین موانع عدم توجه به ارزشیابی مستمر در نظام آموزشی، اهمیت دادن به ارزشیابی پایانی است. برای مثال، در آیین‌نامه آموزشی سه ساله متوسطه در ماده ۵۵ و ۵۶ ضریب نمره‌های ارزشیابی مستمر در تمام نوبت‌ها ۱ در نظر گرفته شده است در حالی که ضریب نمره‌های پایانی در نوبت اول و دوم و تابستان به ترتیب ۲ و ۶ و ۴ در نظر گرفته شده است. این بی‌توجهی مسئولان به نمره‌های مستمر و از طرفی دیگر توجه بیش از حد به ارزشیابی پایانی باعث اضطراب و مشکلات امتحانات پایانی خواهد شد و پدیده منفی دیگری به نام شب امتحان درس خواندن را به دنبال خواهد داشت. نظام ارزشیابی حاکم به نوعی ارزش و اهمیت خاصی به صورت از پیش تعیین شده برای آزمون پایانی متصور شده است که این خود می‌تواند از اهمیت نمره‌های مستمر بکاهد و دبیران با بالا دادن نمره‌های مستمر به گونه‌ای در صدد جبران آن هستند. در ادامه اظهارات تعدادی از دبیران ذکر می‌شود.

پاسخگوی شماره ۱۵: «خود تصمیم‌سازان نظام آموزشی به ارزشیابی پایانی بیشتر اهمیت می‌دهند به‌گونه‌ای که ضریب نمرات پایانی ۶ ولی ضریب نمرات مستمر ۱ است که این خود‌گویای توجه بیشتر به ارزشیابی پایانی است». پاسخگوی شماره ۱۴: «اصلاً دستورالعمل به گونه‌ای است که دبیران را تحریک به بالا دادن نمره مستمر می‌کند تا تأثیر کم آن را جبران کند».

۴-۳- **مصلحت‌گرایی آموزشی:** مصلحت‌گرایی، رفتاری است که متناسب با مقتضیات زمان و بر اساس فایده‌طلبی و سودخواهی انجام می‌گیرد. در این پژوهش، مصلحت‌گرایی آموزشی به صورت رقابت‌های بین مدرسه‌ای و با توجه به ارزش‌گذاری کمی به نتایج عملکرد تحصیلی مدرس‌ها از نظر نهادهای بالادستی و جامعه انجام می‌گیرد. بنابراین، نمره‌های مستمر به جای اینکه ابزاری برای سنجش آموخته‌های واقعی یادگیرندگان باشد به ابزاری برای پاس کردن درس‌ها تبدیل می‌شود. در ادامه به اظهارات چند تن از دبیران برای تأیید این مطالب اشاره می‌شود.

پاسخ‌دهنده شماره ۶: «بالا دادن نمره مستمر باعث قبول شدن دانش‌آموز در امتحانات نهایی می‌شود و همین قبولی باعث بالا رفتن پرستیژ معلم، مدیر مدرسه و اداره مربوطه می‌شود». پاسخ‌دهنده شماره ۳: «به دلیل رقابت بین مدارس و مناطق آموزشی مدیران تأکید بر بالا دادن نمرات مستمر دارند به این امید که نمرات مستمر بالا باعث قبولی دانش‌آموزان در آزمون پایانی شود».

۵- **عدم شناخت فضای روانی آزمون نهایی (تفاوت جو روانی - اجتماعی آزمون):** با توجه به گستردگی ابعاد عواطف و اثرگذاری آن بر موفقیت زندگی، نقش آن در پدیدارهای آموزشی غیرقابل انکار است. پیشرفت تحصیلی به‌عنوان شاهد اصلی بر سلامت روانی دانش‌آموزان، قابل پیگیری و تأمل است. تجارب موفقیت‌آمیز در محیط یادگیری علاوه بر انگیزش درونی، موجب ارزش‌گذاری در بین گروه همسالان و در بسته بزرگ‌تر اجتماع می‌شود. اما در مقابل آن، باور به شکست در آزمون موجب افسردگی، ناراحتی، ناسازگاری و بی‌قراری می‌شود. در تحلیل این مضمون اصلی، دو مضمون فرعی یکی با عنوان «اثر رشد عاطفی، روانی، هیجانی و اجتماعی» و دیگری «آسودگی زیست‌آزمون» مورد بررسی قرار گرفته است.

۵-۱- **سنجش رشد عاطفی، روانی، هیجانی و اجتماعی:** ارزشیابی پایانی رشد دانش‌آموزان را از نظر عاطفی، روانی، هیجانی و اجتماعی نمی‌سنجد. ارزشیابی پایانی تعامل دانش‌آموزان با همدیگر، حس نوع‌دوستی، صمیمیت، هنر دانش‌آموز، فداکاری و



حس باهم زیستن را نمی‌سنجد اما در طول سال تحصیلی به‌طور مستمر معلم می‌تواند پارامترهای فوق را بسنجد. در آزمون پایانی به علت ناآشنایی مصحح با آزمون‌دهنده و نداشتن شناخت پیشین با آن در نمره‌گذاری عوامل عاطفی و هیجانی برآمده از تعامل دبیر و دانش‌آموز چندان اثرگذار نیست اما بر عکس نقش عوامل عاطفی، روانی، هیجانی و اجتماعی در آزمون‌های مستمر با توجه به زیست پیوسته درون‌کلاسی و برون‌کلاسی بین شاگرد و دبیر از اثر زیادی برخوردار است. شواهد گویای این رویداد به صورت کم یا زیاد کردن نمره‌های مستمر با توجه به نوع خوانش رفتار دانش‌آموز در ذهن معلم جابجا می‌شود. تعدادی از دبیران به تأثیر رشد عاطفی، روانی و اجتماعی آزمون اشاره کردند که به نمونه‌هایی از آنها اشاره می‌شود.

پاسخگوی شماره ۲۰: «آزمون پایانی نمی‌تواند برای رفتار روانی، اجتماعی و عاطفی دانش‌آموز نمره‌ای در نظر بگیرد در حالی که در آزمون مستمر چنین شاخص‌هایی دیده شده و می‌توان برای آن نمره‌ای لحاظ کرد».

۵-۲- آسودگی زیست آزمون: یکی از عوامل شکست در آزمون‌ها آشنایی نداشتن دانش‌آموزان با شرایط و نحوه برگزاری آزمون‌ها است. آزمون‌های مستمر در یک فضای آشنا با هدایت یک‌راست دبیر انجام می‌پذیرد. در آزمون‌های مستمر، امکان همیاری گروهی در آموزش میسر است و راهنمایی‌های دبیر از احساس ترس یادگیرنده از شکست می‌کاهد. اما آزمون‌های پایانی در یک فضای ناشناخته و با عوامل اجرایی ناشناس برگزار می‌شود. همین تفاوت در زمینه و فرایند باعث ایجاد اضطراب و ناامنی روانی دانش‌آموز می‌شود و بر کیفیت عملکرد او تأثیرگذار است. در ادامه به اظهارات چند تن از دبیران در این خصوص اشاره می‌شود.

پاسخگوی ۲۱: «دانش‌آموز باید در محیطی که درس را فراگرفته در همان محیط آزمون پس دهد، چون یادآوری در آن محیط راحت‌تر است. عوامل درگیر در آزمون مستمر فقط دبیر مربوطه و خود دانش‌آموزان هستند ولی در آزمون‌هایی عوامل مختلفی درگیر هستند که بعضاً دانش‌آموزان با آنها هیچ‌گونه شناخت و آشنایی ندارند که این باعث ایجاد فشار روانی و اضطراب بر دانش‌آموزان و در نهایت افت نمرات پایانی می‌شود». پاسخگوی شماره ۲: «در آزمون مستمر، دبیر مربوطه حضور دارد و این حضور دبیر قوت قلبی است برای دانش‌آموز. اگر مطلبی یا موضوعی نامفهوم باشد دبیر آن را توضیح می‌دهد، اما سر جلسه آزمون نهایی معمولاً دبیران نامرتبط یا از مقاطع دیگر هستند و این امر باعث اضطراب دانش‌آموزان می‌شود». پاسخگوی شماره ۱۵: «خود

واژه امتحان نهایی برای دانش‌آموزان استرس‌زا است و همین پشت سر هم قرار گرفتن امتحانات نهایی به تعداد زیاد باعث ایجاد خستگی و بعضاً ناامیدی و نوعی فشار روانی بر دانش‌آموزان می‌شود و در نتیجه نمرات پایانی با کاهش مواجه می‌شود».

**۶- عدم تناسب کیفیت یاددهی - یادگیری:** کیفیت یک اتفاق نیست بلکه همواره نتیجه تلاشی آگاهانه است. بنابراین باید نظام تعلیم و تربیت به گونه‌ای مدیریت شود که هدف آن ایجاد استاندارد برای ارتقای مدام کیفیت، روش‌ها و استفاده بهینه از امکانات محدود باشد. ایده اصلی مدیریت کیفیت جامع کلاس درس، به کارگیری مجموعه اصولی است که معلم کلاس برای بهبود یادگیری مستمر به کار می‌برد تا کیفیت یاددهی و یادگیری بهتر شود. به عبارت دیگر، مدیریت کیفیت جامع فرایندی است که معلم را درگیر تلاشی گسترده برای بهبود کیفیت یاددهی یادگیری می‌کند تا نیازهای دانش‌آموزان به بهترین شکل ممکن رفع شود. از دیگر ادراکات و تجربیات دبیران دو مضمون فرعی شامل نابرخورداری آموزشی و معنی‌داری و پایداری در یادگیری در ارتباط با کیفیت یاددهی-یادگیری استخراج شده است که در ادامه، تحلیل آنها ارائه می‌شود.

**۶-۱- نابرخورداری آموزشی:** اجرای هر برنامه آموزشی نیازمند امکانات و تجهیزات خاص خود است. اجرای ارزشیابی مستمر، نیازمند امکاناتی از قبیل آزمایشگاه، لابراتوار زبان، وسایل تکثیر، فضای آموزشی مناسب برای کارهای عملی و گروهی است که برنامه‌ریزان آموزشی باید به‌طور جدی در تدارک و تهیه آن تلاش کنند. تعدادی از دبیران نیاز به امکانات و فضاهای آموزشی را در راستای اجرای صحیح ارزشیابی مستمر مطرح کردند که در ادامه به نمونه‌هایی از آنها اشاره می‌شود.

پاسخگوی شماره ۱۷: «به دلیل کمبود امکانات آزمایشگاهی درس زیست شناسی آزمون جامعی از این قسمت به عمل نمی‌آوریم و با توجه به شناخت کلی نمره‌ای برای دانش‌آموز در نظر می‌گیریم». پاسخگوی شماره ۴: «کمتر توسعه‌یافتگی، کمبود وسایل و امکانات آزمایشگاهی و آموزشگاهی و همچنین ورود دانش‌آموزان با پایه ضعیف باعث می‌شود زمان و انرژی زیادی از دبیر گرفته شود و دبیر نتواند آن‌گونه که باید و شاید به اجرای ارزشیابی مستمر بپردازد و نمره‌ای بالاتر از آنچه نمره واقعی است برای دانش‌آموزان در نظر می‌گیرد».

**۶-۲- معنی‌داری و پایداری در یادگیری:** معنی در یادگیری، جایگاه مهمی دارد. یعنی وقتی مفهومی قابل ارتباط دادن با مفاهیمی باشد که از قبل در ساخت شناختی فرد وجود دارد، آن مفهوم معنی‌دار است. یعنی یادگیری معنی‌دار، مستلزم آن است که

یادگیرنده از قبل مفاهیمی را آموخته باشد که مفهوم جدید قابل ربط دادن به آن است. حال اگر یادگیرنده بتواند مطالب جدید را به مطالبی که قبلاً آموخته است ارتباط دهد، یادگیری او معنی‌دار می‌شود. به عکس، اگر مطالب جدید فقط بر اثر تمرین و تکرار حفظ شود و بدون ارتباط با آموخته‌های قبلی به دست آید و پراکندگی و نبود انسجام در آن مشهود شود، آن یادگیری طوطی‌وار و بی‌معنی خواهد بود. در ادامه به اظهارات پاسخ‌دهنده‌ای در این خصوص اشاره می‌شود.

پاسخ دهنده شماره ۷: «دلیل افت نمرات پایانی نسبت به مستمر عمقی مطالعه نکردن دانش‌آموزان و اصرار بر ادامه روش حفظی طوطی‌وار گونه است. دانش‌آموزان زحمت مطالعه کل یک فصل را حتی به خود نمی‌دهند تا موضوع و مفهوم درس را بهتر درک کنند و مدام خواهان برجسته کردن سؤالات امتحانی توسط دبیر هستند، حال آنکه آزمون پایانی مستلزم فهم عمیق دروس است و اگر سؤالات پایانی اندکی متفاوت از سؤال معلم ساخته باشد دانش‌آموزان از پاسخ دادن به آن درمانده می‌شوند».

۷- انگیزه و نگرش: اگر انگیزش را فرایندهای جهت‌دهنده به رفتار بدانیم، تابعی از نیازها یا شرایط درونی فرد و رویدادهای بیرونی مشوق رفتار فرد است. این رویدادها ممکن است نشان‌دهنده پیامدهای خوشایند یا آزاردهنده باشد. آزمون‌های مستمر با توجه به درگیری بیشتر مشوق‌های بیرونی و درونی از انگیزش عملکردی سطح بالاتری برخوردار است اما آزمون‌های پایانی با توجه به گستردگی مطالب درسی، مدیریت خودخوانی دانش‌آموز و فشردگی زمان آزمون‌ها باعث خستگی شده و این خستگی یکی از عوامل کاهش انگیزش یادگیرنده است. در ادامه به اظهارات دبیران اشاره می‌شود.

پاسخ‌دهنده شماره ۸: «خستگی و فشار امتحانات پایانی باعث کاهش انگیزه دانش‌آموزان و در نتیجه افت نمرات امتحان نهایی را به دنبال دارد». پاسخ‌دهنده شماره ۱۲: «به ما معلمان گفته می‌شود نمرات مستمر را بالا بدهید که در آزمون پایانی دانش‌آموز خودش نمره قبولی را می‌گیرد اما در پایان وقتی کارنامه‌ها را نگاه می‌کنیم دیگر نه انگیزه‌ای برای من دبیر و نه دانش‌آموز باقی می‌ماند. دانش‌آموز وقتی می‌بیند بهش ارفاق می‌شود این تصور در ذهنش ایجاد می‌شود که چه درس بخواند و چه نخواند قبول خواهد شد و این نگرش در واقع آفت نظام آموزشی ما است».

پرسش سوم: تفاوت بین نمره‌های مستمر و نهایی از دید دانش‌آموزان در چیست؟

پرسش سوم پژوهش، به تعیین علل تفاوت بین آزمون‌های معلم‌ساخته و آزمون‌های نهایی از دید دانش‌آموزان اشاره دارد. با مطالعه مکرر متون نوشته شده و متن شنیداری برآمده از مصاحبه دانش‌آموزان به شیوه نیمه ساختاریافته در طی این پژوهش، ۴ مضمون اصلی و چندین مضمون فرعی استخراج شد. زمینه کلی، مضامین اصلی و فرعی در جدول (۶) ارائه شده است.

جدول (۶) مضامین اصلی و فرعی از نظر دانش‌آموزان

| مضامین اصلی  | مضامین فرعی   |
|--|---|
| یکسان نبودن شاخص‌های روایی و پایایی در آزمون مستمر و نهایی | <ul style="list-style-type: none"> <li>• نمونه‌گیری متفاوت از برنامه درسی و محتوای درس</li> <li>• تفاوت در هدف‌های حوزه شناختی</li> <li>• تفاوت در شاخص‌های دشواری آزمون</li> <li>• شرایط متفاوت اجرای آزمون مانند: زمان، فضا، تعدد مراقبان و امکان تقلب</li> </ul> |
| نارضایتی از کیفیت تدریس معلمان                             | <ul style="list-style-type: none"> <li>• نارضایتی از کیفیت تدریس معلمان</li> <li>• تأکید بر حفظ مطالب</li> <li>• ناکافی بودن تمرینات ارائه شده</li> <li>• استفاده از روش سنتی و قدیمی</li> </ul>  |
| عوامل فردی   | <ul style="list-style-type: none"> <li>• خستگی</li> <li>• نبود فضای مطالعه مناسب</li> <li>• اهمال‌کاری تحصیلی</li> </ul>  |
| تفاوت در حجم مطالب مورد آزمون                              | <ul style="list-style-type: none"> <li>• نسبت کم حجم مطالب درسی در آزمون مستمر</li> <li>• نسبت زیاد حجم مطالب درسی در آزمون پایانی</li> </ul>   |

### بررسی مضامین محوری حاصل از مصاحبه با دانش‌آموزان

مضمون اصلی ۱: یکسان نبودن شاخص‌های روایی و پایایی در آزمون مستمر و نهایی: یکی از پرسش‌ها این بود که دلیل اختلاف نمره‌های مستمر و پایانی خود را در چه می‌بینند. درون‌مایه‌های استخراج شده عبارت‌اند از: نمونه‌گیری متفاوت از برنامه

درسی و محتوای درس، تفاوت شاخص دشواری آزمون، تفاوت در هدف‌های حوزه شناختی.

**مضمون فرعی ۱: نمونه‌گیری متفاوت از برنامه درسی و محتوای درس:** برای طراحی پرسش‌ها، ابتدا برنامه درسی مورد نظر به دقت واری می‌شود و از محتوای درس‌ها نمونه‌گیری به عمل می‌آید. در این پژوهش، چگونگی توزیع پرسش‌ها از بخش‌های مختلف، یکی از دلایل تفاوت نمره‌های مستمر و نهایی بیان شد. تعدادی از مشارکت‌کنندگان در پژوهش حاضر عنوان داشتند که نکات مهم از دید دبیران آنها با سازندگان نمره‌های نهایی تفاوت داشته است. برای مثال، یکی از دانش‌آموزان اظهار داشت: «در امتحانات نهایی از جاهایی سوال آمده بود که حتی معلم ما آن را درس نداده بود» یا دانش‌آموز دیگری گفت: «جاهای مهم کتاب از دید معلم ما با سؤالات امتحانی فرق داشت».

**مضمون فرعی ۲: تفاوت در هدف‌های حوزه شناختی:** هدف‌های شناختی در شش طبقه کلی قرار می‌گیرند که این شش طبقه به‌طور سلسله‌مراتبی از عینی و ساده به انتزاعی و پیچیده مرتب شده‌اند. تفاوت در انتخاب هدف‌ها برای سنجش حوزه شناختی از دیگر دلایل تفاوت عملکرد دانش‌آموزان در امتحانات مستمر و نهایی بود. یکی از دانش‌آموزان گفت: «سطح سؤالات نهایی کجا و سطح سؤالات معلم کجا؟ سؤالات امتحان نهایی خیلی پیچیده بود». دیگر دانش‌آموزی نیز عنوان داشت: «سؤالات نهایی خیلی جامع بود و باید بین فصل‌ها ارتباط برقرار می‌کردی ولی سؤالات معلم همان سؤالات آخر فصل بود و نیاز به درک مطلب نداشت».

**مضمون فرعی ۳: تفاوت در شاخص‌های دشواری آزمون:** سطح دشواری آزمون از دیگر نکاتی بود که دانش‌آموزان به آن اشاره کردند. پرسش‌های مستمر به دلیل اینکه معلم موقع درس پرسیدن مرتب از آنها استفاده کرده است، برای دانش‌آموز قابل پیش‌بینی است و با آنها آشناست. به اظهارات برخی از مشارکت‌کنندگان اشاره می‌شود. - «من تقریباً با سؤالات مستمر آشنا بودم و اصلاً نمونه سؤالات دبیر قابل تشخیص بود». - «سؤالات دبیر همان‌ها بود که می‌پرسید و یا می‌گفت اینجا مهمه ولی سؤالات نهایی معلوم نیست از کجا میاد».

**مضمون فرعی ۴: شرایط متفاوت اجرای آزمون (زمان، فضا، تعدد مراقبان و امکان تقلب):** شرایط اجرای آزمون در امتحانات نهایی و مستمر، کاملاً متفاوت است و این عامل، عملکرد تعدادی از فراگیران را تحت تأثیر قرار داده بود.

یکی از افراد در شرح احساس خود در خصوص محدود بودن زمان گفت: «زمان امتحان نهایی کم بود. من عادت دارم جلوی سوال جواب بدهم اما در نهایی باید روی پاسخنامه می‌نوشتیم و من باید جواب‌ها را انتقال می‌دادم به پاسخنامه و وقت کم می‌آوردم». دیگری بیان داشت: «در امتحان مستمر با خواهش و تمنا از معلم می‌شود وقت گرفت ولی نهایی اینجوری نیست». دانش‌آموز دیگری ابراز داشت: «در امتحان نهایی تعداد زیاد مراقبان سر جلسه باعث کاهش امکان تقلب می‌شود در صورتی که در امتحانات مستمر تقلب راحت‌تر است». فراگیر دیگری گفت: «در امتحان نهایی ما هیچ شناختی نسبت به عوامل اجرایی نداریم در حالی که در آزمون مستمر دبیر خودمان بالا سرمون هست و اگر سؤالی نامفهوم باشد، توضیح می‌دهد».

**مضمون اصلی ۲:** نارضایتی از کیفیت تدریس معلمان: دانش‌آموزان در پاسخ به علت تفاوت نمره‌های خود در مستمر و نهایی از روش تدریس معلمان نیز گلایه داشتند و تأکید کردن معلمان بر حفظ مطالب، استفاده از روش‌های قدیمی و ناکافی بودن تمرینات کلاسی را علت کاهش نمره‌های خود قلمداد می‌کردند.

**مضمون فرعی ۱:** تأکید بر حفظ مطالب: دانش‌آموزان اظهار می‌کردند که در کلاس‌ها معلمان از آنها ارتباط مطالب را نمی‌خواهند و تأکید بر حفظ دارند اما در امتحان نهایی تأکید بر یادگیری عمیق مطالب است. یکی از فراگیران، نارضایتی خود را اینگونه اظهار داشت: «ما در کلاس حفظ می‌کردیم و نهایت تمرینات کتاب کار را حل می‌کردیم. کسی از ما نخواستہ بود که عمیق درس را بفهمیم ولی امتحان نهایی پیچیده بود». دانش‌آموز دختری گفت: «دانش‌آموز مطالب را یادنگرفته معلم از کلاس خارج می‌شود و اگر هم سؤالی بپرسیم سرسری جواب می‌دهند».

**مضمون فرعی ۲:** استفاده از روش سنتی و قدیمی: مشارکت‌کنندگان با انتقاد از روش‌های تدریس معلمان خود، آنها را دلیلی بر کم بودن نمره‌های نهایی خود می‌دانستند. مثلاً خانمی اظهار کرد: «معلم ما هنوز از گچ و تخته استفاده می‌کند و ما در کلاس هیچ فعالیتی نداریم. آدم غیرفعال چطوری مطالب را به هم ربط بدهد؟».

**مضمون فرعی ۳:** ناکافی بودن تمرینات ارائه شده: کافی نبودن تمریناتی که معلم به دانش‌آموزان می‌دهد نیز علت افت نمره‌ها از دید مصاحبه‌شوندگان بود. دانش‌آموزی گفت: «تمرین‌های کتاب اصلاً کافی نبود برای امتحان نهایی و کاش معلم تمرین‌های دشوارتری با ما حل می‌کرد».

**مضمون اصلی ۳: عوامل فردی:** دانش‌آموزان همچنین تعدادی از عوامل شخصی را در افت نمره‌های خود دخیل می‌دانستند. اهمال‌کاری تحصیلی، نبود فضای مطالعه مناسب و خستگی تعدادی از مواردی بود که دانش‌آموزان به آنها اشاره داشتند.

**مضمون فرعی ۱: اهمال‌کاری تحصیلی:** کم‌کاری خود فرد نیز در بیان دانش‌آموزان از درون‌مایه‌های پرتکرار بود. شرکت‌کنندگان تجارب خود را این‌گونه بیان کردند: «خودم هم کم‌کاری کردم. می‌توانستم نمونه سؤال بیشتری کار کنم» یا «اگر در روزهای پایانی سال تحصیلی از زمان استفاده بهتری می‌کردم، نمراتم بهتر از این می‌شد».

**مضمون فرعی ۲: نبود فضای مطالعه مناسب:** دانش‌آموزان کمبود فضا و مکان مناسب مطالعه را یکی دیگر از عوامل کاهش نمره‌های پایانی خود می‌دانستند. دانش‌آموز دخترتری ابراز داشت: «چون آزمون پایانی به نسبت آزمون مستمر به مطالعه بیشتری در حدود ۷ الی ۸ ساعت در روز نیاز دارد متأسفانه چنین فضای ساکت و آرامی در منزل ما یافت نمی‌شود». نظر شرکت‌کننده دیگری این بود: «در شهر ما بر خلاف بسیاری از شهرها هیچ سالن مطالعه‌ای وجود ندارد و فرهنگ دید و بازدید از فامیل و همسایه‌ها باعث شلوغی منازل ما شده است».

**مضمون فرعی ۳: خستگی:** امتحانات نهایی پشت سرهم و فشرده هستند و دانش‌آموزان اظهار می‌کردند که خسته شده‌اند و این امر به کاهش نمره‌های آنها منجر شده است. مثلاً این عبارت می‌تواند حاکی از این مطلب باشد: «آن‌قدر پشت سرهم امتحان دادیم که من واقعاً خسته شده بودم و این باعث کاهش نمراتم شد».

**مضمون اصلی ۴: تفاوت در حجم مطالب مورد آزمون:** حجم مطالب درسی در آزمون‌های مستمر و پایانی یکی از پاسخ‌های پرتکرار شرکت‌کنندگان بود. دانش‌آموزان، حجم کم مطالب درسی در آزمون مستمر و حجم زیاد مطالب درسی در آزمون پایانی را دلیل دیگری بر اختلاف نمره‌های مستمر و پایانی خود می‌دانستند.

**مضمون فرعی ۱: حجم کم مطالب درسی در آزمون مستمر:** آزمون‌های مستمر در طول سال تحصیلی و پس از اتمام هر فصل یا مبحث درسی برگزار می‌شود، حجم کمی از محتوا را شامل می‌شود و دانش‌آموزان زمان کافی را برای مطالعه در اختیار دارند و در نهایت نمره‌های بالاتری را کسب می‌کنند. دانش‌آموزان نظرات خود را این‌گونه اظهار کردند: «آزمون مستمر خیلی به ما فشار نمی‌آورد چون حجم مطالب کم و زمان کافی برای مطالعه در اختیار داریم». شرکت‌کننده دیگری گفت: «آزمون مستمر چند

صفحه‌ای بیشتر نیست و چون فاصله زمانی بین تدریس و امتحان نزدیک است می‌توانیم نمرات بهتری کسب نماییم».

**مضمون فرعی ۲: نسبت زیاد حجم مطالب درسی در آزمون پایانی: امتحان نهایی**  
در پایان سال تحصیلی برگزار می‌شود و شامل تمام کتاب درسی است. همین حجم زیاد محتوا و تعدد کتاب‌های درسی در امتحانات نهایی باعث ایجاد خستگی و فشار بر دانش‌آموزان می‌شود و در نهایت، نمره‌های پایانی نسبت به مستمر کاهش می‌یابد. دانش‌آموزی تجربه خود را در این خصوص چنین بیان کرد: «من عادت نداشتم این همه مطالب را یکجا و در این زمان اندک بخوانم و تا حدودی دچار سردرگمی و استرس شدم و این باعث کاهش نمراتم نسبت به مستمر شد».

### بحث و نتیجه‌گیری

با توجه به تحلیل برآمده از گفتگو با دبیران و دانش‌آموزان در خصوص عنوان این پژوهش، تفاوت در حجم دو آزمون، تفاوت در میزان دشواری دو آزمون، تفاوت در شاخص‌های روان‌سنجی دو آزمون، تفاوت در جو روانی اجتماعی دو آزمون، عدم تناسب کیفیت یاددهی - یادگیری، انگیزه و نگرش (عوامل مربوط به دانش‌آموزان) و دستکاری نمره‌ها توسط دبیران از جمله عوامل مؤثر بر تفاوت مشهود نمره‌های مستمر و پایانی بود.

یکسان نبودن شاخص‌های روایی و پایایی در آزمون مستمر و نهایی (نمونه‌گیری متفاوت از برنامه درسی و محتوای درس) دلیل و مضمون اولیه از دید دانش‌آموزان بود. در تبیین این یافته می‌توان به پژوهش حاجبی (۱۳۹۶) اشاره کرد. وی کوشید که از ارزشیابی مستمر در ارتقای نمره‌های تحصیلی در درس زبان استفاده کند. در مرحله اول پژوهش مذکور ۱۲ سری پرسش با توجه به جدول مشخصات طراحی پرسش و هدف‌های شناختی، توسط دو دبیر تهیه شد. این آزمون‌ها برای بررسی روایی محتوایی و صوری به افراد مجرب داده شد و سپس پایایی به روش تصنیف محاسبه شد. استفاده از این آزمون‌ها به پیشرفت تحصیلی، یعنی افزایش ۵ نمره در طی شش ماه استفاده از آزمون‌ها منجر شد. بنابراین، گام اول در اثربخشی امتحان مستمر، صرف وقت و دقت در طراحی پرسش‌هاست. در تأیید این مطلب، دوینگ (۲۰۰۵) هم معتقد بود که اگر

<sup>1</sup> . Downing



اصول صحیح طراحی پرسش‌ها در ارزشیابی تکوینی رعایت نشده باشد، اعتباری در پیشرفت تحصیلی پایانی نخواهد داشت. بنابراین یکی از دلایل تفاوت نمره‌ها در پژوهش حاضر، عدم بودجه‌بندی صحیح طبق جدول مشخصات طراحی آزمون، رعایت نکردن شاخص‌های روایی و پایایی پرسش‌ها توسط دبیران است. ذکر این نکته هم مهم است که اشکالات طراحی پرسش‌ها منحصر به این منطقه نیست و ۳۰ تا ۵۰ درصد پرسش‌های طراحی شده در نظام‌های آموزشی دارای اشکال است (کمیلی و رضایی، ۱۳۸۱؛ کاظمی، احسان‌پور و حسن‌زاده، ۱۳۸۸؛ شکوری‌نیا، مظفری و خسروی بروجنی، ۱۳۸۸). یکی از دلایل و مضامین مهم که دانش‌آموزان ذکر کردند «تأکید معلمان بر حفظ طوطی‌وار مطالب در طول سال تحصیلی» بود. در واقع، در ارزشیابی مستمر، معلمان شرکت‌کننده در این پژوهش، به سطح دانش از حوزه شناختی بلوم اکتفا کرده بودند. این موضوع باعث افت عملکرد در امتحان نهایی می‌شود چراکه در آزمون نهایی، هدف‌های چندگانه بلوم لحاظ شده است. در این خصوص، سیف (۱۳۹۰) اظهار داشته که وقتی دانش‌آموزان یک کلاس، از پیش می‌دانند که در ارزشیابی از یادگیری آنها از یک آزمون چندگزینه‌ای استفاده خواهد شد، هنگام مطالعه و یادگیری مطالب درسی، در جهت پیدا کردن و به یاد سپردن نکات جزئی و پراکنده درسی خواهند کوشید و به مسائل کلی و ساختاری آن توجه زیادی نخواهند کرد. در مقابل، یادگیرندگانی که خود را برای شرکت در یک امتحان تشریحی آماده می‌کنند به اطلاعات پراکنده درسی چندان اهمیتی نمی‌دهند و بیشتر سعی می‌کنند تا مسائل کلی و استنباطی را فرابگیرند. در واقع، اثرگذاری آنچه معلم در کلاس درس انجام می‌دهد (چه طرح پرسش شفاهی و چه نوع پرسش‌های کتبی)، استراتژی کلیدی و مهمی در بهبود کیفیت یادگیری برای همه فراگیران است. عدم کیفیت یاددهی می‌تواند در انگیزه، نگرش و عوامل فردی تأثیر گذاشته و شکاف بین نمره‌های مستمر و نهایی را ایجاد کند (اومیرا، وایتینگ و استیل مالی، ۲۰۱۵). البته این مشکل امروز نیست و در سال ۱۳۷۶ نیز کیامنش و نوری با بررسی ارزشیابی‌های بین‌المللی طرح تیمز و طرح A-B-C نشان دادند که محصلان نظام آموزش و پرورش ایران به دلیل محروم بودن از مهارت‌های اساسی تفکر، از نظر پایداری و کاربست معلومات و حل مسائل جدید بسیار ناتوانند؛ حال آنکه از جهت

1. O'Meara, Whiting & Steele-Maley

معلومات و محفوظات، اندوخته‌های قابل ملاحظه‌ای دارند (کیامنش و نوری، ۱۳۷۶، ۷۰).

تا اینجا می‌توان نتیجه گرفت، ارزشیابی باید بخشی از فرایند یاددهی - یادگیری محسوب شود. ارزشیابی صحیح است که اگر در مسیر فرایند آموزشی انجام گیرد، تعیین کند که چه چیزی، به چه کسی، در چه زمانی، به چه منظوری و چگونه باید آموزش داده شود تا رفتارهای مطلوب و منطبق با هدف‌ها فرا گرفته شوند. ارزشیابی مستمر معلمان پژوهش حاضر، به‌عنوان علتی در مسیر موفقیت آزمون نهایی نیست.

همچنین دانش‌آموزان از کمبودهای موجود در کیفیت آموزشی، کیفیت تدریس معلمان، خلاقیت معلمان و انگیزه‌بخشی به‌عنوان دلایل کاهش نمره‌ها یاد کردند. در تأیید نظرات دانش‌آموزان، پژوهش‌ها نشان داده است که رشد مهارت دبیران و روحیه تعهد آنها و همچنین افزایش دانش و توسعه خلاقیت آنها با رشد دانش‌آموزان و عملکرد تحصیلی آنها همبستگی مثبت و قوی دارد (پارسا، ناخدا و مرعشی، ۱۳۸۹ و بلک و ویلیام، ۲۰۰۹). از آنجا که استان هرمزگان از مناطق محروم و بستک نیز محروم‌تر است، بنابراین معلمان تازه‌کار در این مناطق مشغول خدمت می‌شوند و دور از انتظار نیست که کیفیت تدریس آنها مناسب نباشد. در کنار نارضایتی دانش‌آموزان از کیفیت خدمات، معلمان این پژوهش هم در همین خصوص، «نابرخورداری آموزشی» را از دلایل برجسته افت نمره نهایی دانستند. معلمان عنوان داشتند که به دلیل این نابرخورداری، به ناچار با دادن نمره مستمر بالا سعی در جبران آن دارند. در تأیید نظر معلمان و دانش‌آموزان می‌توان به پژوهش عربشاهی کریزی (۱۳۹۵) استناد کرد. وی به تعیین درجه توسعه‌یافتگی شهرستان‌های استان هرمزگان براساس شاخص‌های آموزشی اقدام کرد. سطح توسعه‌یافتگی براساس شاخص‌های آموزشی در شهرستان‌های استان هرمزگان به ترتیب: بندرعباس، میناب، رودان، بندرلنگه، حاجی‌آباد، قشم، بستک، جاسک، پارسیان و ابوموسی است. همان‌طور که مشخص است، بستک درجات آخر را به لحاظ توسعه‌یافتگی دارد و این نتیجه‌گیری منطقی به نظر می‌رسد.

برخی دانش‌آموزان، نداشتن امکان تقلب در امتحان نهایی را دلیل افت نمره‌ها عنوان کردند. بدان معنا که نمره‌های حاصل در ارزشیابی تکوینی، حاصل تقلب بوده است و نه یادگیری. این نتیجه‌گیری با یافته‌های فراست، مه‌رام و امین‌خندقی (۱۳۹۶) همسویی دارد. مطابق با مطالعات این پژوهشگران، ۹۶/۱ درصد از دانش‌آموزان مرتکب تقلب می‌شوند و تنها ۳/۹ درصد تا کنون تقلب نکرده‌اند. ۳۵/۹ درصد از افراد میزان ارتکاب

به تقلب را در صورت وجود شرایط، به میزان زیاد و همیشگی اعلام داشتند. همچنین دانش‌آموزان نگرش مثبتی به تقلب داشته و آن را یک مهارت قلمداد می‌کردند. بنابراین نمره‌های ارزشیابی مستمر در این پژوهش، نمره‌ای واقعی نیست و در شرایط آزمون نهایی که امکان تقلب کمتر می‌شود، نمره‌ها کاهش می‌یابد. علاوه بر تقلب، با توجه به مضمون انگیزه و نگرش از منظر دبیران که باید نمره‌های مستمر را دستکاری کنند و این باعث افت انگیزه و عملکرد دانش‌آموز می‌شود و به هر جهت می‌دانند که نمره قبولی را کسب می‌کنند. بنابراین دستکاری نمره‌ها توسط معلمان و تقلب خود فراگیران، افزایش نمره‌های مستمر را به صورت کاذب نشان داده است و دلیل افت در نمره‌های نهایی از بین رفتن این دو امکان است.

در رابطه با دستکاری نمره‌ها، در حال حاضر مدرسه‌ها در رقابتی برای افزایش درصد قبولی هستند. مسئله رتبه‌بندی مدرسه‌ها با ایجاد حس رقابت، باعث شده که مدیران برای جلوگیری از امتیاز نامناسب در ارزشیابی‌ها، معلمان را برای دادن نمره غیرواقعی تحت فشار قرار دهند و گاهی با تهدید معلمان به کم کردن نمره ارزشیابی سالانه معلم به خواسته خود برسند. بعضی از معلمان نیز برای موفق نشان دادن خود و شاید گرفتن ابلاغ بی‌دردسر سال آینده، بدون سفارش و اعمال فشار و به‌طور خودکار با دادن نمره‌های مستمر بالا و راهکارهای دیگر درصد قبولی خوبی ارائه می‌دهند. هدف از این افزایش نمره بالا بردن امتیاز مدرسه در سطح منطقه است و هدف مدیران منطقه نیز بالا بردن امتیاز منطقه در بین مناطق و به همین ترتیب در سطح استان و کشور. مسئله نمره‌های غیرواقعی، در امتحانات نهایی خود را آشکار می‌سازد؛ وقتی که تمام نمره‌های مستمر بالا هستند؛ اما دانش‌آموزان در برگه امتحان نهایی، نمره‌های ضعیفی می‌گیرند. با این حال، تاکنون به صورت جدی کسی پیگیر چرایی و شناسایی ابعاد این موضوع نشده و فقط نگرانی از سوی مسئولان ابراز می‌شود. فقط در سال ۱۳۹۶ آموزش و پرورش استان لرستان به برگزاری امتحانات شبیه‌سازی نهایی، آزمون‌های پیشرفت تحصیلی، جلسات تجزیه و تحلیل شهرستان‌ها و مناطق، تشکیل کلاس‌های کارگاهی مدیران دبیرستان‌های متوسطه دوم در سطح استانی و توجه به پرسش‌های تدریجی اقدام کرد و توانست برای نخستین بار میانگین معدل امتحان نهایی را بالاتر از ۱۱ برساند که لزوم پرداختن به برنامه‌های مدون را نشان می‌دهد (خبرگزاری مهر، ۱۳۹۶).

در جمع‌بندی نهایی به دو پژوهش همسو با پژوهش حاضر اشاره می‌شود. پاهنگ، مهدیون و یاریقلی (۱۳۹۶) ویژگی‌های معلم، امکانات و تجهیزات، ویژگی‌های دانش‌آموز، هدف‌های آموزشی و محتوای تدریس، منابع مالی و انسانی و روش تدریس

را از عوامل مؤثر بر کیفیت مدرسه می‌دانند. کاکوجویاری، سلسبیلی و پرتو (۱۳۸۲) نیز مهم‌ترین مسائل آموزش و پرورش متوسطه را ارائه ندادن آموزش‌های ضمن خدمت با کیفیت مطلوب، بی‌توجهی به سبک یادگیری و ویژگی‌های شخصیتی شاگردان، غیراستاندارد بودن فضای آموزشی، استفاده نکردن پیوسته و پویا از وسایل کمک‌آموزشی، ضعف توان علمی و تخصصی معلمان و نبود آموزش‌های تخصصی و مستمر برشمردند؛ پس با وجود این مسائل، این نمره‌ها دور از ذهن نیست.

نکته برجسته در نتایج، تسلط نداشتن معلمان به اصول طرح پرسش استاندارد، تکیه بر روش‌های سنتی، آگاهی نداشتن معلمان از نحوه طرح پرسش‌های نهایی و ندانستن هدف‌های شناختی مورد نظر امتحانات کشوری بود. بنابراین، پیشنهاد می‌شود که مدیران بالادستی برای آموزش اصول طراحی پرسش بر اساس هدف‌های شناختی بلوم، روایی صوری، محتوایی و ضریب دشواری، برای معلمان برنامه آموزشی مدونی را طراحی و تأثیر آن را بر نمره‌ها بررسی کنند. همچنین به‌عنوان پیشنهاد کاربردی، با توجه به امکان تقلب گسترده در ارزشیابی مستمر پیشنهاد می‌شود، علل و راه‌های تقلب کردن مورد بررسی قرار گیرد تا فراگیران کوشش خود را افزایش دهند و نظام آموزشی بتواند هدف‌های خود را به‌درستی بررسی و اجرا کند. با توجه به انتقاد بر روش‌های آموزش معلمان از طریق تأکید بر روش‌های سنجش سطحی، حفظی و مانند آن، پیشنهاد می‌شود که به برخی از شیوه‌های جدید ارزشیابی توجه شود و ضمن آن بر آموخته‌های واقعی آزمودنی‌ها تأکید شود. حضور معلمان در دوره‌های توسعه حرفه‌ای که شامل فعالیت‌ها و فرایندهای برنامه‌ریزی شده برای افزایش دانش، مهارت نگرش حرفه‌ای معلمان است، می‌تواند موجبات بهبودی یادگیری دانش‌آموزان را فراهم سازد.

منابع

- پارسا، عبدالله؛ ناخدا، طوبی؛ و مرعشی، سیدمنصور (۱۳۸۹). بررسی نقش جشنواره فعال تدریس بر بهبود کیفیت آموزش و پیشرفت تحصیلی دانش‌آموزان (مطالعه نگرش دبیران مقطع متوسطه استان خوزستان). *مجله علوم تربیتی دانشگاه شهید چمران اهواز*، ۶(۱)، ۷۹-۱۰۲.
- پاهنگ، نظام‌الدین؛ مهدیون، روح‌اله؛ و یاریقلی، بهبود (۱۳۹۶). بررسی کیفیت مدارس و شناسایی عوامل مؤثر بر آن: پژوهشی ترکیبی. *دوفصلنامه مدیریت مدرسه*، ۵(۱)، ۱۷۳-۱۹۳.
- حاجبی، مصطفی (۱۳۹۶). ارتباط بین ارزشیابی مستمر و پیشرفت تحصیلی دانش‌آموزان در درس زبان انگلیسی. *فصلنامه تحقیقات جدید در علوم انسانی*، ۳، ۶۵-۷۲.
- خبرگزاری مهر (۱۳۹۶). *ارتقای میانگین نمره‌های نهایی دانش‌آموزان لرستانی*. [۱۳ تیر ۱۳۹۶ - ۱۳:۵۵: کد خبر: ۴۰۲۱۴۰۰].
- دولتی، علی‌اکبر؛ جمشیدی، لاله؛ و امین بیدختی، علی‌اکبر (۱۳۹۴). بهبود فرایند یاددهی - یادگیری مدارس هوشمند از منظر ارزشیابی. *دوفصلنامه مطالعات آموزش و یادگیری*، ۷(۲)، ۱-۲.
- سیف، علی‌اکبر (۱۳۹۰). *روان‌شناسی پرورشی نوین*. تهران: دوران.
- شکوری‌نیا، عبدالحسین؛ مظفری، علیرضا؛ و خسروی بروجنی، آذر (۱۳۸۸). بررسی سؤالات چندگزینه‌ای آزمون دستیاری در دانشگاه علوم پزشکی جندی‌شاپور اهواز. *مجله علمی پزشکی*، ۸(۴)، ۴۹۱-۵۰۳.
- عرب‌شاهی کریزی، احمد (۱۳۹۵). تعیین درجه توسعه‌یافتگی شهرستان‌های استان هرمزگان براساس شاخص‌های آموزشی. *پژوهشنامه فرهنگی هرمزگان*، ۱۳(۹)، ۱۱۳-۱۳۳.
- فراست، مریم؛ مهram، بهروز؛ و امین‌خندقی، مقصود (۱۳۹۶). نگرش و میزان ارتکاب دانش‌آموزان به تقلب و پیش‌بینی‌پذیری نگرش دانش‌آموزان به تقلب بر اساس برنامه درسی اجرا شده. *فصلنامه مطالعات اندازه‌گیری و ارزشیابی آموزشی*، ۷(۱۷)، ۱۳۳-۱۵۵.
- کاظمی، اشرف؛ احسان‌پور، سهیلا؛ و حسن‌زاده، اکبر (۱۳۸۸). وضعیت ارزشیابی تحصیلی دروس تخصصی نظری دوره کارشناسی پیوسته مامایی. *مجله ایرانی آموزش در علوم پزشکی*، ۹(۴)، ۳۴۶-۳۵۵.

- کاظمی، سارا (۱۳۹۰). مطالعه تطبیقی نظام ارزشیابی پیشرفت تحصیلی و مدارج تحصیلی در مقطع ابتدایی، راهنمایی و متوسطه کشورهای ایران، آلمان و انگلستان. پایان‌نامه کارشناسی ارشد، دانشگاه علامه طباطبایی.
- کاجوبیاری، علی‌اصغر؛ سلسبیلی، نادر؛ و پرتو، ناصر (۱۳۸۲). سند و منشور اصلاح نظام آموزش و پرورش ایران، پیش‌نویس دوم. تهران: پژوهشکده تعلیم و تربیت. کمیلی، غلامرضا؛ و رضایی، قربانعلی (۱۳۸۱). نحوه ارزشیابی دانشجویان توسط اعضای هیئت علمی دانشگاه علوم پزشکی زاهدان. *مجله ایرانی آموزش در پزشکی*، ۴(۹)، ۵۳-۴۹.
- کیامنش، علیرضا؛ و نوری، رحمان (۱۳۷۶). یافته‌های سومین مطالعات بین‌المللی تیمز. تهران: پژوهشکده تعلیم و تربیت.
- محمدپور، احمد؛ صادقی، رسول؛ و رضایی، مهدی (۱۳۸۹). روش‌های تحقیق ترکیبی به عنوان سومین جنبش روش‌شناختی: مبانی نظری و اصول عملی. *جامعه‌شناسی کاربردی*، ۳۸(۲)، ۷۷-۱۰۰.

- American Educational Research Association, American Psychological Association, & the National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Andesson, C. & Palm, T. (2018). Reasons for teachers' successful development of a formative assessment practice through professional development a motivation perspective. *Assessment in Education: Principles, Policy & Practice*, 25(6), 576-597.
- Baird, J.; Hopfenbeck, T.; Newton, P.; Stobart, G. & Steen-Utheim, A. (2014). *State of the field review: Assessment and learning*. Oslo: Report for the Norwegian Knowledge Centre for Education, case number 13/4697. Retrieved from <http://forskingsradet.no> [Google Scholar].
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18 (1), 5-25.
- Brewer, J. & Hunter, A. (2006). *Foundations of Multi method Research: Synthesizing Styles*. Second Edition, Thousand Oaks, CA: Sage Publications.

- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation & Accountability*, 21(1), 5–31.
- Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of formative assessment*. New York, NY: Routledge.
- Creswell, J., W. & Plano Clark, V. L. (2007). *Designing and Conducting Mixed Methods Research*. London: Sage Publication Inc.
- Dignath, C. & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition & Learning*, 3(3), 231–264.
- Diekelmann, N. L. (2001). Narrative pedagogy: heidegerian hermeneutical analysis of lived experiences of student, teachers, and clinicians. *Advances in Nursing Science*, 23(3), 53-71.
- Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education*, 10(2), 133-143.
- Glazer, N. (2014). Formative Plus Summative Assessment in Large Undergraduate Courses: Why Both? *International Journal of Teaching & Learning in Higher Education*, 26(2), 276-286.
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon: Routledge.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Huhta, A. (2010). Diagnostic and Formative Assessment. In Spolsky, Bernard; Hult, Francis M. *The Handbook of Educational Linguistics*. Oxford, UK: Blackwell. pp. 469–482.
- Johnson, B. & Christensen, L. (2008). *Educational Research: Quantitative, Qualitative and Mixed Approaches*. Needham Heights: Allyn and Bacon.
- O'Meara, J.; Whiting, S. & Steele-Maley, T. (2015). The contribution of teacher effectiveness maps and the TACTICS framework to teacher leader professional learning, leader professional learning. *Journal of Education for Teaching*, 41(5), 529–540.
- States, J.; Detrich, R. & Keyworth, R. (2018). *Overview of Summative Assessment*. Oakland, CA: The Wing Institute. <https://www.winginstitute.org/assessment-summative>.

- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 433–449.
- Tridane, M., Belaouad, S., Benmokhtar, S., Gourja, B., & Radid, M. (2015). The impact of formative assessment on the learning process and the Unreliability of the mark for the summative evaluation. *7th World Conference on Educational Sciences*, (WCES-2015), 186(4), 578–595.
- Tashakkori, A. & Teddlie, C. (2003). *Handbook of Mixed Methods in Social and Behavioral Research*. First Edition, Sage Publication Inc.
- Webb, M. E. & et al (2018). Challenges for IT-Enabled Formative Assessment of Complex 21st Century Skills. *Technology, Knowledge & Learning*, 23(3), 441-456.