

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تقدیم به :

پدرم؛

که امروز خود را مرهون گوشه‌ای از زحمات بی‌ریای او

می‌دانم

مادرم؛

که با دعای خیر او آینده را به انتظار نشسته‌ام.

تقدیر و تشکر:

اینک که به لطف الهی پژوهش حاضر پس از چندین ماه تلاش بی وقفه به اتمام رسیده است، بر خود لازم می دانم از اساتاد ارجمندم جناب آقای دکتر حسین پای کبابکال بربرداری، ایجناب راد

انجام کلیه مراحل این پژوهش را بابتجالی نمودند، تشکر و قدردانی نمایم:

همچنین از اساتاد مشاور محترم جناب آقای حسین الهام صحر که در همه مراحل لطف و محبت خویش را سزودل داشتند، اساتاد ارجمندم جناب آقای حسین الهام صحر که در همه مراحل لطف و محبت خویش را سزودل داشتند، اساتاد ارجمندم جناب آقای حسین الهام صحر که در همه مراحل لطف و محبت خویش را سزودل داشتند، تشکر و قدردانی نمایم:

فصل اول: موضوع تحقیق

۱	مقدمه
۵	بیان مساله
۸	ضرورت و اهمیت تحقیق
۱۰	اهداف تحقیق
۱۰	سوالات تحقیق
۱۲	تعریف مفهومی و عملیاتی متغیرهای تحقیق

فصل دوم: مبانی نظری و پیشینه تحقیق

۱۴	مقدمه
۱۴	تاریخچه و تحول مفهوم ارزشیابی
۱۶	انواع دسته‌بندی از ارزشیابی آموزشی
۱۶	دسته‌بندی ارزشیابیهای آموزشی با توجه به ملاک مورد استفاده
۱۶	ارزشیابی وابسته به ملاک
۱۸	ارزشیابی وابسته به هنجار
۱۸	مقایسه ارزشیابی وابسته به ملاک با ارزشیابی وابسته به هنجار
۱۹	دسته‌بندی ارزشیابیهای آموزشی به توجه به زمان و هدف استفاده از ارزشیابی
۱۹	ارزشیابی آغازین یا سنجش آغازین
۲۰	ارزشیابی تکوینی
۲۱	ارزشیابی تشخیصی
۲۱	ارزشیابی تراکمی
۲۲	هدفهای آموزشی و ارزشیابی پیشرفت تحصیلی
۲۳	هدفهای آموزشی
۲۴	حوزه شناختی
۲۷	حوزه عاطفی
۲۹	حوزه روانی - حرکتی

۳۰	آزمون‌ها و انواع دسته‌بندی از آزمون‌ها
۳۱	دسته‌بندی آزمون‌ها براساس چگونگی تهیه آنها
۳۱	آزمون‌های میزان شده
۳۱	آزمون‌های معلم ساخته
۳۲	دسته بندی آزمون‌ها با توجه به ویژگی‌های مورد اندازه گیری
۳۲	آزمون‌های توانایی
۳۳	آزمون‌های پیشرفت تحصیلی و انواع آنها
۳۴	آزمون‌های صحیح - غلط
۳۵	آزمون‌های جور کردنی
۳۶	آزمون‌های چند گزینه‌ای
۳۸	آزمون‌های ذهنی و انواع آن
۳۸	آزمون‌های تشریحی
۴۰	آزمون‌های کوتاه پاسخ
۴۱	جدول مشخصات و نقش آن در آزمون
۴۲	مدلهای اندازه گیری روانی و تربیتی
۴۲	نظریه کلاسیک آزمون
۴۶	نظریه سوال - پاسخ
۵۰	نظریه تعمیم پذیری
۵۴	ضریب دشواری در نظریه کلاسیک آزمون
۵۷	ضریب تمیز در نظریه کلاسیک آزمون
۶۰	روایی
۶۳	پایایی
۷۴	تحقیقات انجام گرفته در داخل کشور
۷۸	تحقیقات انجام گرفته در خارج از کشور

فصل سوم: روش تحقیق

۸۱	مقدمه
۸۱	روش تحقیق

جامعه، نمونه، و روش نمونه گیری	۸۱
داده‌های تحقیق	۸۲
روش تجزیه و تحلیل داده‌ها	۸۳

فصل چهارم: یافته‌های تحقیق

مقدمه	۸۶
یافته‌های توصیفی	۸۶
یافته‌های سوال ۱	۸۷
یافته‌های سوال ۲	۸۸
یافته‌های سوال ۳	۹۳
یافته‌های سوال ۴	۹۳
یافته‌های سوال ۵	۹۸
یافته‌های سوال ۶	۹۹
یافته‌های سوال ۷	۱۰۰
یافته‌های سوال ۸	۱۰۰
یافته‌های سوال ۹	۱۰۱
یافته‌های سوال ۱۰	۱۰۲

فصل پنجم: بحث و نتیجه گیری

مقدمه	۱۰۳
بحث پیرامون ضریب دشواری و تمیز سوالات امتحان درس ریاضیات	۱۰۳
بحث پیرامون ضریب دشواری و تمیز سوالات امتحان درس علوم	۱۰۵
بحث پیرامون ضریب پایایی نمرات	۱۰۶
بحث پیرامون روایی محتوایی سوالات	۱۰۸
بحث پیرامون سطوح حیطة شناختی سوالات	۱۰۸
محدودیت‌های تحقیق	۱۱۰
پیشنهادات تحقیق	۱۱۱
منابع	۱۱۳

فهرست جداول

- جدول ۱-۴- یافته‌های توصیفی امتحان درس ریاضیات ۸۶
- جدول ۲-۴- یافته‌های توصیفی امتحان درس علوم ۸۷
- جدول ۳-۴- یافته‌های مربوط به ضرایب دشواری و تمیز سوالات امتحان درس ریاضیات ۹۰
- جدول ۴-۴- یافته‌های مربوط به ضرایب دشواری و تمیز سوالات امتحان درس ریاضیات ۹۱
- جدول ۵-۴- یافته‌های مربوط به ضرایب دشواری و تمیز سوالات امتحان درس ریاضیات ۹۲
- جدول ۶-۴- یافته‌های مربوط به ضرایب دشواری و تمیز سوالات امتحان درس علوم ۹۵
- جدول ۷-۴- یافته‌های مربوط به ضرایب دشواری و تمیز سوالات امتحان درس علوم ۹۶
- جدول ۸-۴- یافته‌های مربوط به ضرایب دشواری و تمیز سوالات امتحان درس علوم ۹۷
- جدول ۹-۴- یافته‌های مربوط به ضریب پایایی نمرات امتحان درس ریاضیات ۹۸
- جدول ۱۰-۴- یافته‌های مربوط به ضریب پایایی نمرات امتحان درس علوم ۹۹
- جدول ۱۱-۴- یافته‌های مربوط به سطوح حیطة شناختی سوالات امتحان درس ریاضیات ۱۰۱
- جدول ۱۲-۴- یافته‌های مربوط به سطوح حیطة شناختی سوالات امتحان درس علوم ۱۰۲

نام خانوادگی: میر آقایی	نام: علی عباس
عنوان: بررسی و تحلیل شاخص‌های روان‌سنجی و سطوح حیطة شناختی سوالات امتحانات نهایی دروس ریاضیات و علوم پایه سوم راهنمایی شهرستان خرم‌آباد در خرداد ماه ۱۳۸۶.	
استاد راهنما: دکتر حسین سپاسی	استاد مشاور: آقای حسین الهام پور
درجه تحصیلی: کارشناسی ارشد	رشته: تحقیقات آموزشی
محل تحصیل: دانشگاه شهید چمران اهواز	دانشکده: روانشناسی و علوم تربیتی
تاریخ فارغ التحصیلی: ۱۳۸۸/۳/۳۰	تعداد صفحه: ۱۱۹
واژه‌های کلیدی: ضریب دشواری، ضریب تمیز، پایایی، روایی محتوایی، حیطة شناختی	
<p>چکیده:</p> <p>هدف اصلی این تحقیق بررسی و تحلیل شاخص‌های روان‌سنجی (ضریب دشواری، ضریب تمیز، ضریب پایایی، روایی محتوایی) و سطوح حیطة شناختی سوالات امتحانات نهایی دروس ریاضیات و علوم پایه سوم راهنمایی شهرستان خرم‌آباد در خردادماه ۱۳۸۶ بود. روش تحقیق توصیفی و از نوع ارزشیابی بود. از هر کدام از جامعه‌های آماری برگه‌های نمره‌گذاری شده دروس ریاضیات و علوم ۱۰۰۰ برگه به روش تصادفی ساده انتخاب شدند. نتایج محاسبات نشان داد که ضریب دشواری سوالات امتحانی در مقایسه با ملاک پیشنهادی (۰/۳۰ تا ۰/۷۰) مناسب نبودند ولی ضریب تمیز سوالات امتحانی به استثنای تعداد کمی از سوالات در مقایسه با ملاک پیشنهادی (۰/۲۰ یا بیشتر) از ضریب تمیز مناسب برخوردار بودند. محاسبه ضریب پایایی نشان داد که نمرات امتحان دروس ریاضیات و علوم هر کدام دارای ضریب پایایی ۹۰ صدم بودند و به طور معنی داری (در سطح ۰/۰۱) از ملاکی که معمولاً برای آزمونهای معلم ساخته (بیش از ۸۰ صدم) پیشنهاد می‌شود، بیشتر بود. بنابراین، نمرات امتحانات دروس از دقت و ثبات مناسبی برخوردار بودند. برای بررسی روایی محتوایی و سطوح حیطة شناختی سوالات ۲۵ دبیر، با سابقه تدریس بالای ۵ سال در دروس مربوطه، به روش تصادفی ساده انتخاب شدند. نتایج نشان داد که سوالات امتحانات از روایی محتوایی برخوردار بوده‌اند. برای تعیین سطوح حیطة شناختی سوالات از روش توافقی دلفی دو مرحله‌ای استفاده شد. نتایج نشان داد که سوالات امتحان درس ریاضیات به اندازه‌گیری سطوح بالای حیطة شناختی (ترکیب و ارزشیابی) اختصاص نداشتند. بیشتر سوالات امتحان درس ریاضیات مربوط به سطح کاربرد (۸۶ درصد)، فهم (۶ درصد)، دانش و تحلیل (هر کدام با ۴ درصد) مربوط بودند. سوالات امتحان درس علوم به اندازه‌گیری سطوح تحلیل، ترکیب، و ارزشیابی اختصاص نداشتند و بیشتر سوالات به سطح دانش (۵۳ درصد)، سپس سطح فهم (۳۲ درصد)، و سپس سطح کاربرد (۱۵ درصد) مربوط بودند.</p>	

فصل اول

موضوع پژوهش

مقدمه

دنیای حاضر، دنیای تصمیم‌گیری است. پاره‌ای از تصمیمات در سطح ملی و جهانی گرفته می‌شود که با سیاست، بودجه، قراردادهای، انرژی و مانند آن سروکار دارد. برخی از تصمیم‌ها به اشیاء ارتباط پیدا می‌کند. برای مثال آیا لازم است یک اتومبیل جدید خریداری شود؟ اما، بسیاری از تصمیم‌ها به مردم مربوط می‌شود؛ یا به صورت انفرادی و یا به صورت گروهی. اینها تصمیماتی هستند که به نحوی با اندازه‌گیری و سنجش‌های روانی و تربیتی ارتباط پیدا می‌کنند (هومن، ۱۳۷۱). هدف نهایی اندازه‌گیری‌های روانی و تربیتی شناخت فرد است، به نحوی که بتوان او را به طور کامل و جامع توصیف، و رفتار او را در موقعیت‌های احتمالی آینده پیش‌بینی کرد. برای شناسایی کامل فرد نه تنها باید ویژگی‌ها و مشخصات ظاهری او را توصیف کرد، بلکه لازم است درباره آنچه وی می‌تواند انجام دهد یا در آینده خواهد داد، نیز بتوان پیش‌بینی به عمل آورد. به منظور شناسایی کامل و توصیف قابل درک رفتار فرد لازم است فرآیند گردآوری اطلاعات به یک روش خاص منحصر نشود. هر اندازه اطلاعاتی که درباره فرد گردآوری می‌شود بیشتر و میزان دقت آن زیادتر باشد، احتمال اخذ تصمیم درست در مورد فرد بیشتر خواهد بود (نوروزی، ۱۳۷۴).

روانشناسان و متخصصان تعلیم و تربیت به طور کلی از دو جهت به اندازه‌گیری و سنجش علاقه‌مند بوده‌اند. نخست آنکه بدانند فرد معینی چه می‌کند. در این مورد نتایج اندازه‌گیری مربوط به آن قسمت از رفتار فرد می‌شود که شخصیت نام دارد و موضوع اندازه‌گیری در واقع عمل فرد است. دوم آنکه بدانند فرد چه می‌تواند انجام دهد. در این نوع اندازه‌گیری، مقصود بدست آوردن شواهد و قراین درباره اعمالی است که اگر فرد تلاش کند می‌تواند از عهده انجام آن برآید. نتایج این نوع اندازه‌گیری که در واقع مربوط به توانایی فرد می‌شود ممکن است به دو قسمت اصلی تقسیم شود:

۱- اندازه‌های مربوط به استعداد^۱ که معرف قابلیت یادگیری فرد است؛

۲- اندازه‌های مربوط به معلومات یا پیشرفت تحصیلی^۲ که معرف اموری است که فرد قبلاً یاد گرفته و اکنون می‌تواند از عهده انجام آن برآید.

منظور از اندازه‌گیری استعداد آن است که مشخص کنیم فرد معینی تحت شرایط مساعد، در آینده چه کاری را بهتر می‌تواند انجام دهد؛ در حالی که مقصود از اندازه‌گیری معلومات یا پیشرفت، تعیین چیزی است که فرد در گذشته فرا گرفته است (نوروزی، ۱۳۷۴).

نقش اساسی تعلیم و تربیت، در همه زمینه‌های که شناسایی فرد مورد توجه است، ایجاد تغییرات مطلوب در رفتار فرد است. معلم باید مشکل دانش‌آموزان را تشخیص دهد و به وسیله تغییر در رفتار راهی برای مرتفع نمودن مشکل او توصیه کند. امتحان یکی از وسایل و تسهیلاتی است که به کمک آن معلم می‌تواند، نتایج تغییراتی را که در فرد ایجاد شده است، ارزشیابی کند و منظور از آن در واقع این است که معلوم شود آیا تغییرات معینی در فرد بوجود آمده است یا نه؟ آیا این تغییرات، مطلوب و رضایت‌بخش است؟ و میزان این تغییرات نزد افراد مختلف تا چه اندازه متفاوت است؟ (باقری، ۱۳۸۵).

معلمی که به اتکای دانش، معلومات و مهارت‌های پیشین شاگردان عهده‌دار تدریس و ارائه درس بخصوصی در کلاس درس می‌شود، لازم است بداند که آیا شاگردان واقعاً مواردی را که قبلاً به آنها آموزش داده شده را یاد گرفته‌اند یا نه. برای نیل به این هدف معلم می‌تواند شاگردان خود را قبل از شروع آموزش جدید، امتحان و نتایج آن را بررسی کند. چنانچه نتایج حاصل، دال بر آن باشد که شاگرد هنوز مهارت‌های اساسی مورد نیاز آموزش جدید را فرا نگرفته است، ممکن است انجام اقدامات اصلاحی و جبرانی را پیشنهاد کند. گاه نیز منطقی به نظر می‌رسد که معلم معلوم کند که آیا شاگردان از لحاظ دانش، معلومات، و مهارت‌ها به گونه رضایت‌بخش به آن سطح که در برنامه و طرح تدریس او گنجانیده شده است رسیده‌اند یا نه. چنانچه موفقیت شاگردان در سطح مواد تدریس

1. aptitude
2. achievement

شده رضایت بخش باشد می‌تواند آن مواد را از طرح تدریس خود حذف و به تدریس سایر مواد پردازد. در هر دو مورد، یک امتحان قبلی می‌تواند برای طرح تدریس معلم کمک مؤثر و مفیدی باشد (هومن، ۱۳۷۱).

از یک سو، یک معلم خوب مایل است بداند تا چه حد در رسیدن به اهداف آموزشی موفقیت حاصل کرده است. امتحانی که با دقت تهیه شده و همه هدف‌های آموزشی را می‌پوشاند، می‌تواند اطلاعات وسیع و سودمندی را در اختیار معلم قرار دهد. بعد از اجرای آزمون، معلم می‌تواند سؤال‌های امتحان را تجزیه و تحلیل و مشخص کند که اکثریت شاگردان به کدام سؤال پاسخ درست داده‌اند و به کدام سؤال پاسخ درست نداده‌اند. فرض کنید، معلمی هم به اطلاعات و معلومات شاگردان در زمینه درس بخصوصی و هم به افزایش توانایی آنها در کاربرد اصول آن درس در زندگی روزمره علاقه‌مند است. چنانچه معلم مزبور امتحانی را که محتوای آن منطبق بر هر دو فرآیند مذکور است به شاگردان بدهد و نتایج آن را تجزیه و تحلیل کند، در این صورت با شمارش تعداد شاگردانی که به هر یک از انواع سؤال‌ها پاسخ درست داده‌اند، می‌تواند مواد و روش‌هایی را که در امتحان بکار برده است بررسی و معلوم کند که اصولاً هر سؤال تا چه اندازه در سنجش هدف مورد نظر موفق بوده است (هومن، ۱۳۷۱).

از سوی دیگر، معلم در ارزشیابی تفاوت‌های فردی بین شاگردان خود از لحاظ نیل به هدف‌های گوناگون تربیتی مسئولیت مستقیم دارد. اگر این مطلب را بپذیریم که مدرسه باید مناسب‌ترین محیط یادگیری را برای کودک فراهم سازد، بنابراین معلم باید تفاوت‌های بین شاگردان خود را به صورت نمره یا برگ معرفی‌نامه، به معلم بعدی گزارش کند تا از این طریق بتواند در بهبود شرایط یادگیری کودک بکوشد. هرگونه تصمیم معلم در مورد دانش‌آموز به گونه‌ای قابل ملاحظه‌ای بستگی به داوری معلمان قبلی دانش‌آموز دارد، که در پرونده تحصیلات او ثبت و منعکس شده است، و بسیاری از اطلاعات که داوری معلمان مبتنی بر آنهاست نیز تا اندازه‌ای زیادی از طریق اجرای امتحانات مختلف بدست آمده است (سعیدی آرانی، ۱۳۷۷).

به گونه‌ای کلی نتایج آزمون‌ها تا حدودی مشخص می‌کند که شاگردان باید چه موقع مطالعه کنند، چه مطالبی را فرا گیرند و چگونه یاد بگیرند، امتحانی که خوب تهیه شده باشد به شاگردان این فرصت را می‌دهد که خود را بیازمایند و معلومات خود را ارزشیابی کنند، و بنابراین به عنوان بازخوردی سازنده و یک وسیله ترغیب، تأثیر مهمی در بهبود و بالابردن میزان کارکرد و کارآئی آنها دارد. این امتحانات اشتباهات شاگردان را تصحیح و فعالیت‌های آنان را در جهت نیل به هدف‌های مناسب هدایت می‌کنند. اما چنانچه امتحان طبق اصول صحیح تهیه نشده باشد یا از آن به عنوان یک وسیله تنبیه استفاده شود، به گونه موثر موجب دلسردی یا گمراهی شاگردان شده یادگیری آنها را به مخاطره می‌اندازد. در واقع، امتحان خود محرکی است برای یادگیری، و هر چه آموزش و ارزشیابی نزدیک‌تر باشد رغبت به یادگیری نیز بیشتر خواهد بود (یاقوتیان، ۱۳۷۴).

متأسفانه با اینکه شاهد اجرای آزمون‌های بسیاری در نظام‌های آموزشی هستیم ولی تاکنون از آزمون‌های مناسب، کم و از آزمون‌های بسیار مناسب خیلی کمتر استفاده کرده‌ایم. بسیاری از سازندگان آزمون بدون اینکه از آموزش کافی در سنجش و اندازه‌گیری و شیوه ساخت آزمون‌ها برخوردار باشند به ساخت آزمون مبادرت می‌کنند، به ویژه این امر در مورد معلمان که اطلاعات آنها درباره ساخت و ویژگی‌های امتحان به زمان دانشجویی خلاصه می‌شود شایع‌تر است. علاوه بر این، وقتی از دانش‌آموزان درباره سؤالات امتحانی معلمان‌شان سؤال می‌شود پاسخ می‌دهند که سؤالات نامفهوم بودند، سؤالات بسیار دشوار بودند، سؤالات به پرسش مفاهیم ساده و پیش پا افتاده که فقط به حفظ طوطی وار مطالب نیاز داشتند اختصاص داشت (مهرنز و لهمان، ۱۹۸۴).

توضیح بیشتر اینکه، مطالعات متعدد نشان داده است که بیشتر سؤالات امتحانات معلم ساخته تنها نیاز به یادآوری اطلاعات بسیار جزئی و خاص دارند، در حالی که قسمت بیشتر هدف‌هایی که به وسیله همان معلمان تنظیم شده، مهارت‌های اساسی نظیر کار بستن، تجزیه و تحلیل و ترکیب را در برداشته است. به بیان دیگر سؤال‌های که معلمان در امتحانات به کار می‌برند به هدف‌های اصلی تدریس آنان هماهنگ نیست. قراین و شواهد نشان می‌دهند که آموخته‌های شاگردان بیش از آنکه

مبتنی بر مهارت‌های پیچیده و اساسی معلمان باشد، شامل محفوظات و یادآوری مطالب جزئی و کم اهمیت است. به طور خلاصه معلم بدون آنکه خودش آگاه باشد از روی سؤال‌های که برای امتحان تهیه کرده است می‌توان شناخت و از روی همین امتحانات می‌توان تشخیص داد که معلم واقعاً چه هدف‌هایی را دنبال کرده، برای چه مطالبی اهمیت قائل بوده و یادگیری شاگردانش نیز بیشتر تحت تأثیر کدام قسمت از مطالب درسی بوده است (هومن، ۱۳۷۱).

در یک مطالعه تطبیقی که توسط سازمان بین‌المللی سنجش و ارزشیابی (EIA) بین ۵۰ کشور در مقاطع ابتدایی و راهنمایی از جمله دانش‌آموزان ایرانی در دروسی مثل ریاضیات و علوم انجام گردید، دانش‌آموزان ایرانی توانستند به کمتر از ۵۰ درصد از سوالات این آزمون‌ها پاسخ درست بدهند (کیامنش، ۱۳۷۷).

در یک مدل آموزش، اندازه‌گیری و ارزشیابی از آموخته‌های دانش‌آموزان در قالب امتحانات نقش بسیار مهمی بر انگیزه و عملکرد دانش‌آموزان و معلمان دارد. پرداختن به عوامل عدم موفقیت آزمون‌ها در دسترسی به اهداف قصد شده به مطالعات گوناگونی نیاز دارد که از حوصله یک تحقیق خارج است. این تحقیق از بین این عوامل به بررسی و تحلیل شاخص‌های روان‌سنجی (دشواری، تمیز، پایایی و روایی محتوایی) و سطوح حیطة شناختی سؤالات امتحانات نهایی دروس ریاضیات و علوم پایه سوم راهنمایی شهرستان خرم‌آباد در خردادماه ۱۳۸۶ پرداخت. چنین به نظر می‌رسد که قوت و ضعف بسیاری از سؤالات امتحانات را می‌توان با محاسبه این شاخص‌های روان‌سنجی شناسایی کرد. علاوه بر این، با محاسبه شاخص‌های روان‌سنجی و سطوح حیطة شناختی سؤالات، می‌توان برای تشخیص ویژگی‌های آنها و تشخیص نواقص موجود در سؤالات امتحانات به منظور ایجاد اصلاحات لازم در آزمون‌ها و اجرای مجدد آنها استفاده کرد.

بیان مسأله

امروزه بر اثر پیشرفت روزافزون علم و تکنولوژی و یافته‌های روانشناسی و علوم تربیتی، مفهوم آموزش و پرورش دگرگون شده است و دامنه گسترش آن نیز افزایش یافته است. نقش آموزش و

پرورش در دنیای متحول امروز آن است که استعداد و توانایی دانش‌آموزان را تا جایی که امکان دارد پرورش دهد و آنان را برای شرکت در فعالیت‌های سازنده جوامع پویایی انسانی آماده سازد، به نحوی که اسباب سعادت واقعی افراد و جامعه هر دو تأمین شود. در فرایند آموزش و پرورش عناصر و عوامل متعددی وجود دارند، که هر کدام به نوبه خود در دستیابی به هدف‌های آن نقش مهمی دارند. یکی از این عوامل ارزشیابی است. قضاوت درباره کارکردهای نظام آموزشی، کارایی یا عدم کارایی، به این عامل مربوط است. برای عمل ارزشیابی از ابزارهای گوناگونی استفاده می‌شود، یکی از ابزارهای که به طور گسترده مورد استفاده قرار می‌گیرد، آزمون است (باقری، ۱۳۸۵). مانند سایر پدیده‌های آموزشی اهمیت آزمون‌ها با گذشت زمان دوران فراز و نشیبی را پشت سر گذاشته است. معلمان به طور معمول برای ارزشیابی نتایج تدریس و فراهم آوردن انگیزه یادگیری برای دانش‌آموزان در کلاس‌های درس به انجام آزمون‌های می‌پردازند. در بسیاری از مدارس در ایالات متحده امریکا آزمون‌های معلمان و نواحی به وسیله آزمون‌های دیگری که به وسیله‌ای کارشناسان ایالتی و یا ناحیه-ای پیگیری می‌شود، تکمیل می‌گردد (ایبل، ۱۹۹۱، ترجمه سپاسی، ۱۳۸۱).

به طور کلی، معلمان آزمون‌ها را برای چهار نوع ارزشیابی به کار می‌گیرند. این چهار نوع ارزشیابی عبارتند از آغازین، تکوینی، تشخیصی و تراکمی. ارزشیابی آغازین^۱ در ابتدای آموزش اجرا می‌شود و دارای دو هدف است، این اهداف عبارتند از این که آیا دانش‌آموزان برای یادگیری مطالب جدید دانش و مهارت‌های پیش‌نیاز را کسب کرده‌اند و اینکه آیا دانش‌آموزان مطالب جدید را قبلاً یاد گرفته‌اند. ارزشیابی تکوینی^۲ در ضمن آموزش اجرا می‌شود و هدف از کاربرد ارزشیابی تکوینی در ارتباط با پیشرفت تحصیلی دانش‌آموزان، آگاهی یافتن از میزان و نحوه یادگیری آنان برای تعیین نقاط قوت و ضعف یادگیری و نیز تشخیص مشکلات روش آموزشی معلم در رابطه با هدف‌های آموزشی است. به طور خلاصه، هدف از آن رفع مشکلات یادگیری دانش‌آموزان و برطرف کردن نواقص روش آموزش است. ارزشیابی تشخیصی^۳ نیز در جریان آموزش اجرا می‌شود و هدف از آن،

1. pre assessment evaluation
2. Formative evaluation
3 diagnostic evaluation

همانند ارزشیابی تکوینی، تشخیص مشکلات یادگیری دانش‌آموزان است، ولی وقتی مورد استفاده قرار می‌گیرد که معلم با مشکلات مبرم و مکرری در یک یا چند دانش‌آموز روبه‌رو شود. ارزشیابی تراکمی^۱ در پایان دوره اجرا می‌شود و هدف از آن تعیین این است که آیا دانش‌آموزان صلاحیت‌های لازم را برای کسب مدرک تحصیلی و ورود به پایه تحصیلی بالاتر کسب کرده‌اند یا خیر (سیف، ۱۳۸۳).

متأسفانه با این‌که شاهد اجرای آزمون‌های بسیاری در نظام‌های آموزشی هستیم ولی تاکنون از آزمون‌های مناسب، کم و از آزمون‌های بسیار مناسب خیلی کمتر استفاده کرده‌ایم. بسیاری از سازندگان آزمون بدون اینکه از آموزش کافی در سنجش و اندازه‌گیری و شیوه ساخت آزمون‌ها برخوردار باشند به ساخت آزمون مبادرت می‌کنند، به ویژه این امر در مورد معلمان که اطلاعات آنها درباره ساخت و ویژگی‌های آزمون به زمان دانشجویی خلاصه می‌شود شایع‌تر است. علاوه بر این، وقتی از دانش‌آموزان درباره سؤالات امتحانی معلمان‌شان سؤال می‌شود پاسخ می‌دهند که سؤالات نامفهوم بودند، سؤالات بسیار دشوار بودند، سؤالات به پرسش مفاهیم ساده و پیش‌پا افتاده که فقط به حفظ طوطی وار مطالب نیاز داشتند اختصاص داشت (مهرنزو لهما، ۱۹۸۴).

اگر سؤالات آزمون‌ها نامفهوم و یا بسیار دشوار یا بسیار آسان باشند، نمرات حاصل از آنها منعکس‌کننده توانایی واقعی دانش‌آموزان در موضوع مورد مطالعه نخواهند بود و سؤالات آزمون قادر نیستند بین دانش‌آموزان قوی و ضعیف تفاوت بگذارند و چنانچه دانش‌آموزان تلاش زیادی برای یادگیری بکنند ولی به نتایج مورد انتظارشان دست نیابند، پس از یک دوره زمانی انگیزه آنان برای یادگیری پایین خواهد آمد و اگر این عدم موفقیت تکرار شود دانش‌آموزان به نوعی درماندگی مبتلا خواهند شد و امکان دارد از تلاش برای جبران نقاط ضعف خود باز بمانند.

علاوه بر این‌که می‌توان از نتایج آزمون‌ها برای تعیین پیشرفت تحصیلی دانش‌آموزان و فراهم کردن انگیزه برای یادگیری آنها استفاده کرد، برای ارزشیابی عملکرد معلمان در کلاس‌های درس نیز کاربرد دارد. بنابراین، ضروری است که شیوه و کیفیت وسایل سنجش مورد ارزشیابی قرار گیرد. از

1. summative evaluation

طریق این ارزشیابی می‌توان به نقاط قوت و ضعف ساخت سؤالات و خطاهای ناشی از قضاوت طراح سؤال پی‌برد. متخصصان آزمون‌سازی برای ارزشیابی سؤالات آزمون شاخص‌هایی تحت عنوان شاخص‌های روان‌سنجی پیشنهاد کرده‌اند که از طریق آنها سازندگان آزمون می‌توانند با شناسایی کردن نقایص و برطرف کردن آنها مهارت‌های خود را در نوشتن سؤالات امتحانی افزایش دهند و در کیفیت سؤالات آزمون‌ی که در آینده مورد استفاده قرار خواهند داد، تجدیدنظر کنند. با توجه با مطالب فوق می‌توان گفت که همه سؤالاتی که توسط معلمان برای اندازه‌گیری پیشرفت دانش‌آموزان ساخته می‌شود، بویژه سؤالات امتحان نهایی که پیشرفت تحصیلی تعداد زیادی از دانش‌آموزان را اندازه‌گیری می‌کنند و نتایج آنها دارای پیامدهای بسیاری است، لازم است که شاخص‌های روان‌سنجی (دشواری، تمیز، پایایی و روایی محتوایی) و اهداف آموزشی که سؤالات این امتحانات می‌سنجند، مورد بررسی قرار گیرند. با توجه به این که در امتحانات نهایی خرداد ماه ۱۳۸۶ پایه سوم راهنمایی در شهرستان خرم‌آباد از میان دروس نهایی پایه سوم راهنمایی دروس ریاضیات و علوم، به ترتیب با ۸۵ و ۸۷ درصد قبولی، دارای کمترین درصد قبولی بودند. تحقیق حاضر به بررسی و تحلیل شاخص‌های روان‌سنجی و سطوح حیطة‌شناختی سؤالات این امتحانات پرداخت.

ضرورت و اهمیت تحقیق

آزمون‌های کلاسی معلم ساخته به طور وسیعی در سیستم آموزش و پرورش برای بهبود بخشیدن به تمرین‌های کلاسی، گزارش دادن پیشرفت دانش‌آموزان و آزمودن شمار زیادی از دانش‌آموزان استفاده می‌شوند. طراحی این آزمون‌ها به طور وسیعی به تجارب کاری معلمان در ساخت سؤالات آزمون‌ها و ادراک آنها از دشواری سؤالات وابسته است (گادو، ۲۰۰۴). از انواع مختلف امتحانات می‌توان به عنوان یک وسیله مفید و ارزنده به منظور گزینش دانش‌آموزانی که احتمال موفقیت تحصیلی آنان در آینده بیشتر است یا برای راهنمایی شخصی، تحصیلی و شغلی آنان استفاده کرد (قندیلی، ۱۳۷۲).

برای انجام این تحقیق می‌توان دلایل زیر را ذکر کرد: اول، امتحان جزئی از فرآیند آموزش و قسمت جدانشدنی از آن است و در واقع یکی از وسایل برانگیزاننده و جالب فرآیند یادگیری بشمار می‌آید. نقش عمده آن مشخص کردن کارآیی شاگرد هم برای خود او و هم برای معلم است. دوم،

نتایج امتحان می‌تواند مسیر آموزش را جهت بهبود و یا تغییر، برای معلم روشن سازد و علاوه بر این مواردی را که احتیاج به آموزش بیشتر دارد و نیز میزان آمادگی شاگرد را برای آموزش بعدی مشخص سازد (سعیدی آرانی، ۱۳۷۷). سوم، مدل بازخورد نشان می‌دهد که باید کیفیت وسایل سنجش مورد ارزشیابی قرار گیرد. چنانچه عملکرد دانش‌آموزان پایین‌تر از حد انتظار ارزیابی شود، شاید به جای این‌که به یادگیری غیر کافی از طرف یادگیرندگان مربوط گردد به دلیل ضعف روش‌های ارزشیابی باشد. پس برای اینکه نمره‌ها برای هدفی که در اصل مورد نظر بوده ارزشمند باشند، آزمون‌ها باید مورد ارزشیابی قرار گیرند. چهارم، اطلاع و آگاهی معلمان از عملکرد ضعیف برخی سؤالات امتحان توانایی بیشتری را برای کنترل بحث‌های کلاسی در اختیار معلمان قرار می‌دهد. پنجم، آزمون و نمره سؤالات می‌توانند ضعف‌های ساخت سؤالات و خطاهای ناشی از قضاوت نویسنده سؤال را روشن سازند. معلمان می‌توانند با برطرف کردن این نقائص و قابل استفاده مجدد نمره‌ها، مهارت‌های نوشتن سؤال خود را پیشرفت دهند و در کیفیت آزمون‌ها که در آینده مورد استفاده قرار خواهد گرفت تجدید نظر کنند (ایبل، ۱۹۹۱، ترجمه سپاسی، ۱۳۸۱). ششم، با تعیین سطح حیطه شناختی که هر کدام از سؤالات امتحان اندازه‌گیری می‌کنند، می‌توان در مورد اهدافی که دانش‌آموزان به آن دست یافته‌اند، اطلاع پیدا کرد.

با توجه به مطالب ذکر شده، نتایج آزمون‌ها دارای اثرات بسیار مهمی هستند. برای این‌که این نتایج از صحت و اعتبار کافی برخوردار باشند سازندگان آزمون‌ها باید از دانش و تجربه کافی برای ساختن آزمون‌ها برخوردار باشند، با این وجود بسیاری از متخصصان سنجش و اندازه‌گیری معتقدند برخی معلمان فاقد این دانش و مهارت می‌باشند (ایبل، ۱۹۹۱، ترجمه سپاسی، ۱۳۸۱). چنانچه سازندگان آزمون‌ها از دانش و مهارت درباره شیوه‌هایی ساخت آزمون و محاسبه شاخص‌های روان-سنجی برخوردار نباشند، بدون شک تصمیم‌گیرها براساس نتایج این آزمون‌ها از صحت و اعتبار کافی برخوردار نخواهد بود. با اهمیتی که آزمون‌ها در آینده دانش‌آموزان و معلمان و جامعه دارند، ضرورت دارد که شاخص‌های روان‌سنجی و سطوح حیطه شناختی سؤالات امتحانات مورد بررسی قرار گیرند تا هم به سازندگان آزمون کمک کند تا به نقاط قوت و ضعف ساخت آزمون‌ها آشنایی

پیدا کنند و هم به آنها در این که چه نکاتی را باید برای ساخت آزمون‌های آینده باید مورد توجه قرار دهند، کمک کند.

اهداف تحقیق

هدف کلی. هدف اصلی این تحقیق بررسی و تحلیل شاخص‌های روان سنجی (ضریب دشواری، ضریب تمیز، پایایی و روایی محتوایی) و سطوح حیطة شناختی سؤالات امتحانات نهایی دروس علوم و ریاضیات پایه سوم راهنمایی شهرستان خرم‌آباد در خردادماه ۱۳۸۶ بود.

اهداف فرعی

- (۱) تعیین ضریب دشواری و ضریب تمیز سؤالات.
- (۲) تعیین پایایی نمرات.
- (۳) بررسی اعتبار محتوایی سؤالات.
- (۴) بررسی میزان اندازه‌گیری سطوح مختلف حیطة شناختی توسط سؤالات.

سؤالات تحقیق

سؤالات اصلی

- (۱) ضریب دشواری و ضریب تمیز سؤالات امتحان نهایی درس ریاضیات، پایه سوم راهنمایی به چه میزان بوده است؟
- (۲) ضریب دشواری و ضریب تمیز سؤالات امتحان نهایی درس علوم، پایه سوم راهنمایی به چه میزان بوده است؟
- (۳) ضریب پایایی نمرات امتحانات نهایی، دروس ریاضیات و علوم پایه سوم راهنمایی تا چه اندازه بوده است؟

۴) آیا سؤالات امتحانات نهایی، دروس ریاضیات و علوم پایه سوم راهنمایی، دارای روایی محتوایی بوده‌اند؟

۵) سؤالات امتحانات نهایی دروس ریاضیات و علوم پایه سوم راهنمایی، کدام یک از سطوح حیطة شناختی را اندازه‌گیری کرده‌اند؟

سؤالات فرعی تحقیق عبارتند از:

۱) ضریب دشواری سؤالات امتحان نهایی درس ریاضیات به چه میزان بوده‌است؟

۲) ضریب تمیز سؤالات امتحان نهایی درس ریاضیات به چه میزان بوده‌است؟

۳) ضریب دشواری سؤالات امتحان نهایی درس علوم به چه میزان بوده‌است؟

۴) ضریب تمیز سؤالات امتحان نهایی درس علوم به چه میزان بوده‌است؟

۵) ضریب پایایی نمرات امتحان نهایی درس ریاضیات تا چه اندازه بوده‌است؟

۶) ضریب پایایی نمرات امتحان نهایی درس علوم تا چه اندازه بوده‌است؟

۷) آیا سؤالات امتحان نهایی درس ریاضیات دارای روایی محتوایی بوده‌اند؟

۸) آیا سؤالات امتحان نهایی درس علوم دارای روایی محتوایی بوده‌اند؟

۹) سؤالات امتحان نهایی درس ریاضیات کدام یک از سطوح حیطة شناختی را اندازه‌گیری کرده-

اند؟

۱۰) سؤالات امتحان نهایی درس علوم کدام یک از سطوح حیطة شناختی را اندازه‌گیری کرده‌اند؟