

## ساخت و هنجاریابی آزمون پیشرفت تحصیلی (MAT) دانش‌آموزان سال اول دبیرستان Construct and standardize an academic achievement test for Year 1 Mathematics at high schools

دریافت مقاله: ۱۳۸۹/۱۲/۴؛ دریافت نسخه نهایی: ۱۳۹۰/۲/۱۶؛ پذیرش مقاله: ۱۳۹۰/۹/۱۷

H. Pasha Sharifi (Ph.D.)

H. Noori., (M.A)

Email: hamed noori 84 @yahoo.com

دکتر حسن پاشا شریفی<sup>۱</sup>، حامد نوری<sup>۲</sup>

**Abstract:** The purpose of this study was to construct and standardize an academic achievement test for Year 1 Mathematics at high schools. A sample of 1000 students (500 girls, 500 boys) was randomly selected and subjects were matched by sex, region, and educational setting. Research instrument was a 400-item Moths scale with four subscales developed by the researchers. Reliability of subscales ranged from .90 to .98. Content validity of the scale was satisfactory. The Kelly method was used to modify the instrument by removing inappropriate items. Results indicated that girls had better academic performance than boys. Items were in congruent with the Bloom taxonomy of cognitive domain. Then, the scale was standardized using percentile ranks. In conclusion, the scale may be used for both summative and formative assessment in high schools to evaluate student's performance on Mathematics.

**KeyWords:** academic achievement test, Mathematics, standardizational.

**چکیده:** هدف پژوهش حاضر بررسی توصیفی است و منظور ساخت و هنجاریابی آزمون پیشرفت تحصیلی درس ریاضی دانش‌آموزان سال اول دبیرستان انجام شده است. گروه نمونه به تعداد هزار نفر (۵۰۰ دختر و ۵۰۰ پسر) دانش‌آموزان سال اول دبیرستان شهرستان سمنجان است. گروه نمونه به شیوه تصادفی انتخاب شده و از نظر متغیرهای جنسیت و منطقه و محل تحصیل هم‌متا گردیدند. ابزار مورد استفاده آزمون محقق ساخته چهار خرده آزمون تحت عنوان آزمون پیشرفت تحصیلی درس ریاضی هر کدام ۱۰۰ سؤال جمعاً با ۴۰۰ سؤال و با محتوای چهار خرده آزمون A, B, C, D با اعتبار ۰/۹۶، ۰/۹۸ و ۰/۹۰ و ۰/۹۲ می‌باشد. آزمون از روایی محتوایی با استفاده از نظر متخصصین و اعتبار میانگین خرده آزمون‌ها با استفاده از فرمول ۲۰ کوردر- ریچاردسون حدود ۰/۹۲ می‌باشد. سؤال‌های آزمون‌های مزبور با حدود ۲۵۰ الی ۳۰۰ نفر پیش از اجرا برای هر خرده آزمون اجرا شده و به روش کلی KELLY گروه‌های قوی و ضعیف به صورت ۲۷٪ گروه بالا و ۲۷٪ گروه پایین ضریب دشواری و ضریب تمیز سؤال‌ها تعیین گردیده شد و سؤال‌های نامناسب حذف و سؤال‌های مناسب جایگزین گردید. یافته‌ها نشان داد عملکرد تحصیلی دانش‌آموزان دختر بهتر از پسران است و همچنین سئوال‌های آزمون‌ها با رعایت هدف‌های رفتاری و طبقه بندی شناختی بلوم سازگار بوده و ۲۲٪ درصد از سئوالها در سطح آسان با ضریب دشواری  $P=0/70$  و ۵۰ درصد سئوالها در سطح متوسط با ضریب دشواری  $P=0/50$  و ۱۸٪ درصد سئوالها در سطح دشوار با ضریب دشواری  $P=0/30$  تهیه شده و ۱۰٪ درصد بقیه در سطح آسان - متوسط تهیه شدند. هنجاریابی (نرم) بر طبق رتبه درصدی با استفاده از نمرات Z و CEEB و AGCT تهیه گردید. این آزمون به عنوان آزمون پیشرفت تحصیلی برای دانش‌آموزان از لحاظ اصول روان‌سنجی و سنجش اندازه‌گیری برای ارزشیابی دانش‌آموزان مناسب می‌باشد و می‌تواند مورد استفاده دبیران متوسطه و دانش‌آموزان بعنوان ارزشیابی تشخیصی تکوینی و تراکمی مورد استفاده قرار گیرد.

**کلید واژه‌ها:** آزمون پیشرفت تحصیلی، درس ریاضی، هنجاریابی (نرم)،

۱. دانشیار دانشگاه آزاد اسلامی، واحد رودهن

۲. کارشناسی ارشد سنجش و اندازه‌گیری (روانسنجی)

## مقدمه

یکی از مقاصد اصلی و عمده هر علم، اندازه‌گیری پدیده‌های مورد نظر و دادن جنبه کمی به آن‌هاست. یعنی تبدیل کیفیت به کمیت است. وقتی پدیده‌ای را با اعداد و ارقام نشان می‌دهیم به راحتی می‌توانیم آن را طبقه‌بندی و قوانین محکمی درباره آن وضع کنیم. در علم روان‌شناسی، مجموعه شیوه‌هایی که به ما کمک می‌کند تا ویژگی‌های روانی انسان را از حالت کیفی به حالت کمی درآوریم اصطلاحاً «روش روان‌سنجی» نامیده می‌شود.

روان‌سنجی، در معنای وسیع کلمه، یعنی استفاده از روش آزمایشی، اما در معنای محدود کلمه یعنی بهره‌گیری از آزمون‌های روانی برای اندازه‌گیری توانایی‌های ذهنی (گنجی، ۱۳۶۷: ۱۹). با توجه به مطلب فوق می‌توان گفت هر علمی به‌طور کلی از یک سلسله مفاهیم یا سازه‌های نظری و رابطه بین این مفاهیم نظری تشکیل می‌شود.

اما تا وقتی که مفاهیم نظری قابل اندازه‌گیری یا کمی شدن نباشد کشف قوانین و روابط علی به منظور پیش‌بینی پدیده‌ها امکان‌پذیر نیست. در علوم فیزیکی، اندازه‌گیری هم سن و سال تمدن بشری است و در حال حاضر دقت وسایل اندازه‌گیری در این علوم حقیقتاً شگفت‌آور است. اما علوم مربوط به رفتار انسان دیرتر از سایر علوم از فلسفه جدا شدند و فقط در هشتاد سال گذشته است که روش‌های تحقیق در علوم فیزیکی به بررسی مفاهیم نظری علوم رفتاری تعمیم داده شده‌اند. وبر، فخر، گالتون، کتل، بینه و برت نخستین کسانی بودند که امکان اندازه‌گیری خصوصیات روانی را مطرح کردند و وسایلی برای اندازه‌گیری این خصوصیات تهیه نمودند (مگنوسون، ۱۹۶۷، ترجمه براهنی، ۱۳۷۰: ۱). آن‌ها در زمان خود نیاز اساسی انسان را نسبت به وسایل روش‌هایی برای اندازه‌گیری، تا حدودی بر طرف نمودند. امروز بر اثر پیشرفت روزافزون علوم و تکنولوژی و یافته‌های نوین روان‌شناسی و تعلیم و تربیت و با افزایش جمعیت و پیچیده شدن جوامع و با اطلاع‌یافتن از فواید ابزارهای اندازه‌گیری، انسان کنونی نیز بیش از قبل، نسبت به چنین وسایلی احساس نیاز می‌کند که این نیاز در آموزش و پرورش بیشتر از جاهای دیگر احساس می‌شود زیرا در امر تعلیم و تربیت برای ایجاد تغییر در رفتار فراگیر آگاهی از اندوخته‌های قبلی او لازم است. این آگاهی از طریق اندازه‌گیری و ارزشیابی درست حاصل می‌شود تا بر اساس این آگاهی و شناخت معلم بتواند او را در انتخاب واحدهای درسی، رشته‌های تحصیلی و یا برنامه‌ریزی برای شغل آینده هدایت کند. بنابراین، یکی از ابزارهایی که معلمین را در امر شناخت دانش‌آموزان یاری می‌کند، آزمون پیشرفت تحصیلی است که در بعضی از مقاطع تحصیلی مخصوصاً در نظام جدید دبیرستانی چندان فعالیتی جهت تهیه و ساخت آزمون‌های پیشرفت تحصیلی نشده است و به نظر می‌رسد آزمون‌هایی که برای ارزشیابی به کار گرفته می‌شوند از نظر علمی مطابق استانداردهای فنی تهیه نمی‌گردند. در نتیجه ارزشیابی از

فعالیت‌های دانش‌آموزان دقیق نمی‌باشد و ضرورت دارد آزمون‌های مناسبی برای دروس مختلف به روش‌های علمی تدوین و در اختیار دبیران قرار گیرد تا بتوانند ارزشیابی دقیقی از آموخته‌های درسی دانش‌آموزان داشته باشند.

## روش

روش پژوهش توصیفی است. در این پژوهش هزار نفر از دانش‌آموزان سال اول دبیرستان دختر و پسر شهرستان سمنان (۵۰۰ دختر و ۵۰۰ پسر) به روش نمونه‌گیری تصادفی طبقه‌ای مورد بررسی قرار گرفتند. ابزار پژوهش آزمون محقق ساخته تحت عنوان آزمون پیشرفت تحصیلی درس ریاضی سال اول دبیرستان (MAT)<sup>۱</sup> است که شامل ۴ خرده آزمون A, B, C, D با اعتبار ۰/۹۶، ۰/۹۸، ۰/۹۰ و ۰/۹۲ می‌باشد. برای هر آزمون صد سؤال مطابق با مشخصات امتحان درس ریاضی و هدف‌های رفتاری و طبقه‌بندی بلوم در حیطه شناختی تهیه گردید و روایی محتوایی و روائی سازه آزمون به‌وسیله متخصصین مورد بررسی و تأیید قرار گرفت. پس از اجرای فرم تجربی ۳۰۰ نفری، سؤال‌های با ضریب دشواری و ضریب تمیز نامناسب حذف و سؤالات مناسب جایگزین گردیدند. روش‌های آماری توصیفی مورد استفاده برای میانگین رتبه درصدی نما و میانه، جداول و نمودارها می‌باشد و از آمار استنباطی خی‌دو و از آزمون T برای تعیین عملکرد دانش‌آموزان دختر و پسر در درس ریاضی مورد استفاده قرار گرفت.

## یافته‌های پژوهش

### الف- یافته‌های مربوط به بررسی ویژگی‌های روان‌سنجی سؤال‌ها

بررسی سطح دشواری سؤال‌های آزمون پیشرفت تحصیلی ریاضی سال اول دبیرستان بیان‌گر این است که ۲۲ درصد سؤال‌ها در سطح آسان و ۵۰ درصد در سطح متوسط و ۱۸ درصد در سطح دشوار یا مشکل و ۱۰ درصد آسان - متوسط و تمامی سؤال‌های آزمون از ضریب تشخیص معنادار برخوردار بوده‌اند. این یافته‌ها با آنچه سیف (۱۳۸۲) ذکر نموده یعنی ۲۵ درصد سؤال‌ها باید در سطح آسان و ۵۰ درصد متوسط و ۲۵ درصد در سطح دشوار طرح شوند، هماهنگی دارد.

### ب- اعتبار آزمون‌های ریاضی

- اعتبار آزمون‌های A, B, C, D به ترتیب ۰/۹۶، ۰/۹۸، ۰/۹۰ و ۰/۹۲ می‌باشد که با اعتبار آزمون پیشرفت ریاضی آقای نصیب‌اله رفیعی (۱۳۷۴) هماهنگی دارد و از آن بیشتر می‌باشد. با توجه به مقدار قابل قبول اعتبار (بالای ۹۰ درصد)، این اعتبار با دیدگاه شریفی (۱۳۸۱) در خصوص

<sup>۱</sup>. Mathematic Achievement Test

اصول روان‌سنجی مطابقت دارد. اعتبار با استفاده از فرمول ۲۰ کودر - ریچاردسون<sup>۱</sup> به‌دست آمده است.

### ج- روایی محتوایی

- آزمون‌های پیشرفت تحصیلی درس ریاضی سال اول دبیرستان دارای روایی محتوایی می‌باشد زیرا وقتی آزمون‌های اعتباری باشد دارای روایی نیز خواهد بود. سؤال‌های آزمون پیشرفت تحصیلی درس ریاضی سال اول دبیرستان براساس اهداف آموزشی و محتوای درسی کتاب ریاضی تهیه شده است.

### د- مقایسه عملکرد دانش‌آموزان دختر و پسر

- براساس یافته‌های این پژوهش عملکرد پیشرفت تحصیلی دانش‌آموزان دختر شهرستان سنندج در درس ریاضی بهتر از عملکرد پیشرفت تحصیلی دانش‌آموزان پسر می‌باشد. نتایج این پژوهش با تحقیق کشاورز (۱۳۷۶) هماهنگی دارد. نتیجه این پژوهش نیز با نتایج و میانگین امتحانات خرداد ۸۱-۸۲ شهرستان سنندج از نظر عملکرد هماهنگی دارد یعنی میانگین نمرات درس ریاضی سال اول دبیرستان شهرستان سنندج (دخترها در حدود ۱۰ و پسرها در حدود ۸) می‌باشد. نتایج این پژوهش با نتایج پژوهش‌های انجام شده توسط روان‌شناسان مغایرت دارد.

جدول توزیع فراوانی نمره‌های آزمودنی‌های دختر در درس ریاضی سال اول دبیرستان

حدود طبقات	X عدد میانی هر طبقه	F فراوانی مطلق	P فراوانی نسبی	P% درصد فراوانی نسبی	Cf فراوانی تراکمی	Ref فراوانی تراکمی نسبی	Cf% فراوانی تراکمی درصدی
۹۰-۱۰۰	۹۵	۳۰	۰/۰۵۶	۸	۱۲۵	۱	۱۰۰
۸۰-۹۰	۸۵	۳۵	۰/۰۸	۱۱/۲	۱۱۸	۰/۹۴	۹۴
۷۰-۸۰	۷۵	۴۰	۰/۱۱۲	۱۴/۴	۱۰۸	۰/۸۶	۸۶
۶۰-۷۰	۶۵	۵۵	۰/۱۴۴	۱۴/۴	۹۴	۰/۷۵	۷۵
۵۰-۶۰	۵۵	۱۳۰	۰/۲۳۲	۲۳/۲	۷۶	۰/۶۱	۶۱
۴۰-۵۰	۴۵	۵۵	۰/۱۴۴	۱۱/۲	۵۸	۰/۴۶	۴۶
۳۰-۴۰	۳۵	۵۰	۰/۱۱۲	۵/۶	۲۹	۰/۲۳	۲۳
۲۰-۳۰	۲۵	۴۰	۰/۰۵۶	۶/۴	۱۵	۰/۱۲	۱۲
۱۰-۲۰	۱۵	۵۵	۰/۰۶۴	۵/۶	۸	۰/۰۶	۶
۰-۱۰	۵	۳۰	.	.	.	.	.

$$n=500 \quad \sum P=1 \quad 100$$

<sup>1</sup>. Kuder – Richardson

شاخص‌های آماری جدول

کجی	درصد تغییرپذیری	نما	میانه	واریانس	انحراف معیار	میانگین	تعداد
۰/۲۹	۳۳	۵۵	۶۰	۳۹/۷۲	۶/۳۰	۶۵	۵۰۰

مشاهده شاخص‌های آماری محاسبه شده برای آزمون، بیانگر کجی مثبت توزیع نمرات است زیرا میانگین از میانه و نما بیشتر است (میانگین ۶۵، میانه ۶۰ و نما ۵۵). تجمع نمره‌ها بیشتر در قسمت چپ توزیع است. با توجه به میزان نامحسوس کجی (۰/۲۹) می‌توان گفت که توزیع مذکور چندان از توزیع بهنجار (نرمال) فاصله ندارد.

جدول توزیع فراوانی نمره‌های آزمودنی‌های پسر در درس ریاضی سال اول دبیرستان

حدود طبقات	X عدد میانی هر طبقه	F فراوانی مطلق	P فراوانی نسبی	P% درصد فراوانی نسبی	Cf فراوانی تراکمی	Ref فراوانی تراکمی نسبی	Cf% فراوانی تراکمی درصدی
۹۰-۱۰۰	۹۵	۲۰	۰/۰۰۸	۷	۱۲۱	۱	۱۰۰
۸۰-۹۰	۸۵	۳۵	۰/۰۸۸	۱۰	۱۲۲	۰/۹۹	۹۹
۷۰-۸۰	۷۵	۴۰	۰/۱۲۸	۸	۱۱۶	۰/۹۰	۹۰
۶۰-۷۰	۶۵	۶۰	۰/۱۰۴	۱۲	۹۸	۰/۷۸	۷۸
۵۰-۶۰	۵۵	۶۰	۰/۰۱۴	۹	۸۵	۰/۹۷	۶۷
۴۰-۵۰	۴۵	۱۱۵	۰/۱۲۸	۹/۴	۷۱	۰/۵۷	۵۷
۳۰-۴۰	۳۵	۵۰	۰/۰۸۸	۱۱	۵۵	۰/۴۴	۴۴
۲۰-۳۰	۲۵	۴۵	۰/۰۸۸	۸/۸	۴۴	۰/۳۵	۳۵
۱۰-۲۰	۱۵	۴۰	۰/۱۷۶	۱۶/۴	۳۳	۰/۲۶	۲۶
۰-۱۰	۵	۳۵	۰/۰۸۸	۸/۸	۱۱	۰/۰۹	۹

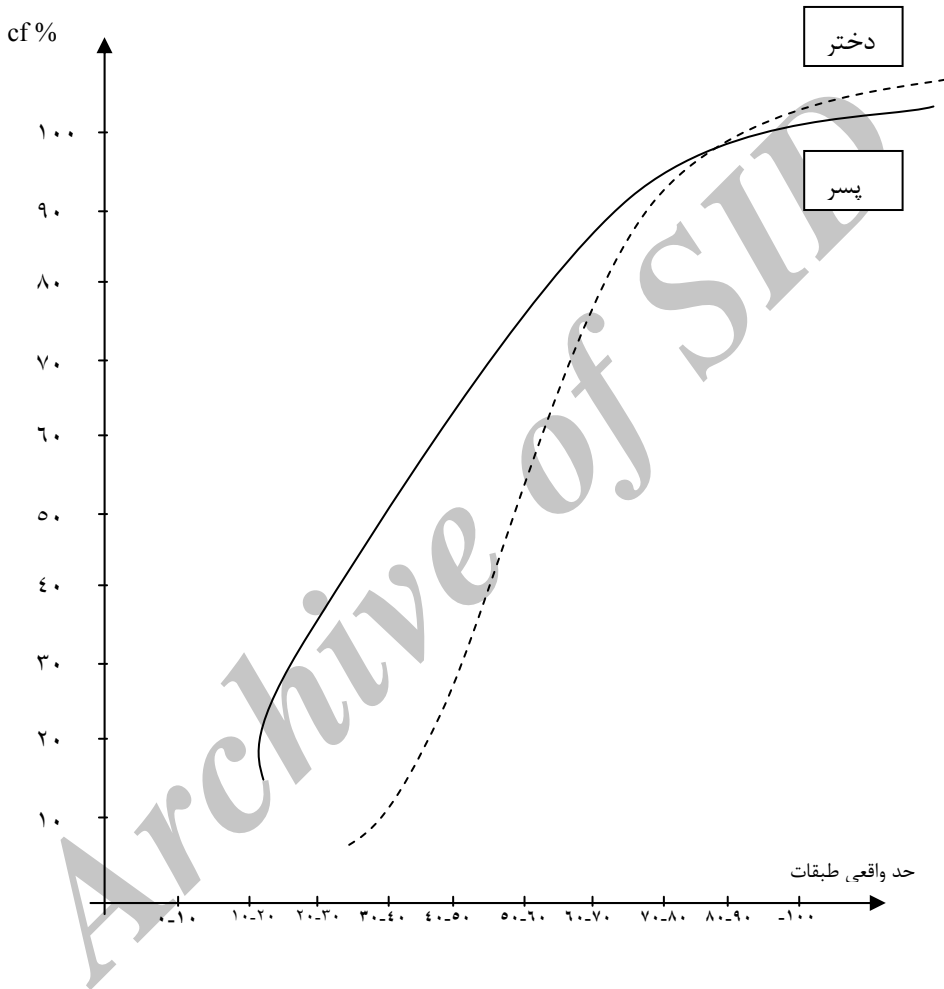
$$n=500 \quad \sum P = 100$$

شاخص‌های آماری جدول

کجی	درصد تغییرپذیری	نما	میانه	واریانس	انحراف معیار	میانگین	تعداد
۰/۱۴	۵۰/۰۳	۴۰	۴۲	۶۰/۳۸	۷/۷۷	۴۵	۵۰۰

چنان‌که مشاهده می‌شود میانگین نمره‌های آزمون ۴۵، از میانه ۴۲ و نما ۴۰ بیشتر است. بنابراین توزیع دارای کجی مثبت یا دارای کجی به راست است. میزان این کجی ۰/۱۴ بوده و

نشان می‌دهد که تجمع نمره‌ها اندکی در سمت چپ توزیع بیشتر است. با این حال می‌توان گفت توزیع یاد شده به توزیع نرمال شبیه است.



**تفسیر:** همان‌طور که تصویر نشان می‌دهد نمودار عملکرد دخترها در درس ریاضی در سمت راست نمودار و نمودار عملکرد پسرها در سمت چپ نمودار در این درس قرار گرفته است. در نتیجه عملکرد دخترها در درس ریاضی بهتر از پسرها بوده است. از مقایسه دو جدول توزیع فراوانی نمره‌های آزمودنی‌های پسر و دختر سال اول دبیرستان در درس ریاضی نتایج زیر حاصل شده است:

- ۱- میانگین عملکرد دخترها در درس ریاضی شهرستان سنندج ۵۶ با انحراف معیار ۶/۳۰ و میانگین عملکرد پسرها در درس ریاضی شهرستان سنندج ۴۵ با انحراف معیار ۷/۷۷ است.  $t$  به‌دست آمده برابر ۴۴/۴۵ است که در سطح ۰/۰۰۱ معنادار است؛ یعنی عملکرد دخترها در درس ریاضی بهتر از عملکرد پسرها در این درس بوده است. این نتایج نزدیک به نتایج عملکرد بهتر دانش‌آموزان دختر شهرستان سنندج در امتحان درس ریاضی می‌باشد زیرا میانگین نمرات دخترها ۱۰ و پسرها ۸ در خردادماه ۱۳۸۲ می‌باشد.
- ۲- نما در جدول توزیع فراوانی نمره‌های آزمودنی‌های دختر در درس ریاضی ۵۵ (عدد ۵۵ بیشترین تکرار را دارد) است و نمای جدول توزیع فراوانی نمره‌های آزمودنی‌های پسر در این درس ۴۵ است. از مقایسه این دو نما برتری نمره‌های دانش‌آموزان دختر در درس ریاضی نسبت به نمره‌های آزمودنی‌های پسر کاملاً مشخص است.
- ۳- توزیع نمره‌های دخترها دارای کجی مثبت است و این بدان معناست که بیشتر دخترها نمره بالاتر از میانگین (۵۵) گرفته‌اند. در صورتی که نمودار نمره‌های پسر دارای کجی منفی است یعنی بیشتر دانش‌آموزان پسر در درس ریاضی نمره کمتر از میانگین (۴۵) گرفته‌اند. اما از آنجایی که اندازه هردو کجی ناچیز است، مقدار کجی نمودار توزیع فراوانی آزمودنی‌های دختر ۲۹ و این مقدار برای نمودار توزیع فراوانی آزمودنی‌های پسر ۱۴/۰+ است. نمودار توزیع نمره‌های دانش‌آموزان دختر و پسر در درس ریاضی به توزیع بهنجار شبیه است.

### جدول طبقه‌بندی نمره‌های خام آزمودنی‌ها

در درس ریاضی سال اول دبیرستان شهرستان سنندج

طبقه‌بندی نمره‌های خام	درجه‌ها
$M$ = کمترین نمره - بالاترین نمره	عالی
فاصله طبقاتی = $M \div 7$	بسیار خوب
حدود نمره‌های خام	خوب
۸۵ - ۱۰۰	متوسط
۷۰ - ۸۵	ضعیف
۵۰ - ۶۵	بسیار ضعیف
۳۵ - ۵۵	عقب‌مانده شدید تحصیلی
۱۵ - ۳۰	
۱۰ - ۱۵	
۰ - ۱۰	

## پروفایل (نیم‌رخ) آزمون پیشرفت تحصیلی درس ریاضی سال اول دبیرستان

الف - استفاده از نرم صدکی (رتبه درصدی)<sup>۱</sup> :

نرم‌های درصدی براساس نمره‌های خام آزمودنی‌های گروه نرم که با فاصله طبقه‌ای معین طبقه‌بندی می‌شوند، به دست می‌آید. هر نمره خام را نمره درصدی و درصد افراد گروه نرم را که زیر نمره معینی قرار می‌گیرد رتبه درصدی می‌نامند. برای روشن شدن و دست یافتن به نمره‌های درصدی اقدامات زیر صورت گرفته است.

۱- فاصله طبقات و نمراتی که بین حد پایین و حد بالای هر طبقه قرار دارند.

۲- نمره‌های میانی هر طبقه که از تقسیم مجموع حد بالا و پایین هر طبقه بر عدد ۲ به دست می‌آید.

۳- فراوانی هر طبقه، تعداد افرادی که در آن طبقه قرار می‌گیرند.

۴- فراوانی تراکمی زیر عدد میانی (چند نفر در زیر این طبقه قرار دارند).

۵- رتبه درصدی هر طبقه که از حاصل ضرب فراوانی تراکمی زیر عدد میانی آن طبقه

$$P.R = \frac{100}{N}$$

به دست می‌آید.

ب) Q: سؤال‌های هر خرده آزمون (فصلهای ۱ تا ۴) کتاب ریاضی سال اول دبیرستان با توجه

به تعداد کل سؤالات دفترچه‌های A, B, C, D از هر ۱۰۰ سؤال ۲۵ سؤال به هر فصل قراردا رد.

یعنی فصل ۱، ۲۵ سؤال، فصل ۲، ۲۵ سؤال، فصل ۳، ۲۵ سؤال، فصل ۴، ۲۵ سؤال که روی

هم کل MAT را شامل می‌شوند.

پ) L: سطح آزمودنی با توجه به جدول طبقه‌بندی نمره‌های خام آزمودنی‌ها در درس ریاضی

سال اول دبیرستان شهرستان سنندج تعیین سطح گردیده است. به عنوان مثال اگر دانش‌آموزی

در آزمون پیشرفت تحصیلی ریاضی MAT در پروفایل (نیم‌رخ) نمره‌اش بین ۲۹-۱۵ باشد ضعیف

و اگر بین ۱۰۰-۸۰ باشد عالی محسوب می‌شود.

ج) امتیاز مشاوره‌ای: این پروفایل (نیم‌رخ) می‌تواند به عنوان امتیاز مشاوره‌ای طبق نظر معلم و

عملکرد دانش در آزمون MAT امتیازات ۱۰، ۸، ۶، ۴ را دریافت نماید و معادل سطح L

محسوب می‌شود. با توجه به پروفایل آزمون پیشرفت تحصیلی می‌توان وضعیت یک دانش‌آموز را

از نظر توانایی انجام هدف‌های آموزشی فصول مختلف کتاب ریاضی سال اول دبیرستان مشخص

و نقاط ضعف و قوت دانش‌آموز را با توجه به نیم‌رخ او در پروفایل (نیم‌رخ) نسبت به همکلاسی‌ها

<sup>۱</sup> . Percentile Rank

<sup>۲</sup> . Questions

<sup>۳</sup> .Level



و سطح شهرستان محاسبه و مشخص نمود. این پروفایل و آزمون MAT برای اتمام هر قسمت کتاب نیز معتبر است زیرا خرده آزمون‌های MAT با کل آزمون دارای روایی سازه می‌باشد. همچنین از روی ستون امتیاز مشاوره‌ای که ملاک عملکرد دانش‌آموز در MAT می‌باشد علاوه بر شرایط احراز رشته ریاضی فیزیک و تجربی را می‌توان سنجش نمود. با برگزاری این آزمون در آبان‌ماه و آذرماه دانش‌آموزانی را که نیاز به کلاس جبرانی ریاضی دارند مشخص نمود.

**پروفایل آزمون پیشرفت تحصیلی ریاضی سال اول دبستان**

دانش‌آموز: **A**  
 شماره پایه: **۷۵-۵۵-۷۰-۲۵-۵۵/۷۵**  
 نام دبیر: **X**  
 نام آزمونگر: **B**  
 تاریخ: **۸۷/۱۱/۲۰**

سطح	Percentile J-		سطح				Percentile J+				سطح	
	PR	Q	Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>	Y <sub>4</sub>	PR	Q				
عالی	100	25	-	-	-	-	-	-	-	100	25	<b>تفسیر MAT</b> دانش‌آموز ۸ طبق پروفایل با میانگین رتبه درصدی چهار خرده آزمون فصل (۴، ۲، ۴، ۴) P.R-59 می‌باشد یعنی از ۹۶ درصد دانش‌آموزان عملکرد تحصیلی بهتری داشته‌اند. جالب این بیشترین نطفه اوج پروفایل وی در فصل اول مجموعه و کمترین آن در معادله بوده است. در کل دانش‌آموز عملکرد خوبی می‌باشد و امتیاز ۶ مشاوره‌ای بر اساس ورودیه رشته ریاضی و تجربی را جهت آورده است. دانش‌آموز ۸ به ۷ سوال پاسخ داده است از ۱۰۰ سوال و بیشترین سوال‌های پاسخ داده وی فصل اول مجموعه (۵) با ۲۰ سوال و کمترین سوال‌های پاسخ داده شده معادله با ۱۲ سوال. جالب این دانش‌آموز چنانچه فصل ۶ را بیشتر تمرین نماید در ضمن نسبت به فصل اول پیشرفت تحصیلی بهتری در فصل ۴ کسری خواهد داشت. Arithmetic Achievement test profile
	95	24	10	-	-	-	-	-	-	95	24	
	90	23	-	-	-	-	-	-	-	90	23	
	85	22	-	-	-	-	-	-	-	85	22	
	80	21	-	-	-	-	-	-	-	80	21	
خیلی خوب	75	20	8	-	-	-	-	-	-	75	20	
	70	19	-	-	-	-	-	-	-	70	19	
	65	18	-	-	-	-	-	-	-	65	18	
خوب	60	17	6	-	-	-	-	-	-	60	17	
	55	16	-	-	-	-	-	-	-	55	16	
	50	15	-	-	-	-	-	-	-	50	15	
متوسط	45	14	4	-	-	-	-	-	-	45	14	
	40	13	-	-	-	-	-	-	-	40	13	
	35	12	-	-	-	-	-	-	-	35	12	
	30	11	-	-	-	-	-	-	-	30	11	
	25	10	-	-	-	-	-	-	-	25	10	
ضعیف	20	9	2	-	-	-	-	-	-	20	9	
	15	8	-	-	-	-	-	-	-	15	8	
بسیار ضعیف	10	7	-	-	-	-	-	-	-	10	7	
	5	6	-	-	-	-	-	-	-	5	6	
خیلی ضعیف نیازمند تحصیلی	4	5	2	-	-	-	-	-	-	4	5	

### بررسی اعتبار آزمون پیشرفت ریاضی (MAT)

برای محاسبه اعتبار آزمون‌های مورد نظر از فرمول ۲۰ کودر - ریچاردسون استفاده شد. ابتدا اعتبار آزمون A از فرمول ۲۰ کودر - ریچاردسون بررسی شد. پولمن تیر<sup>۱</sup> فرضیه‌هایی را تهیه و کاربرد آزمون‌های استاندارد شده در آن‌ها را به شرح زیر خلاصه کرده است:

$$r_{tt} = \frac{K}{K-1} \left[ 1 - \frac{\sum pq}{s^2} \right]$$

$$A \rightarrow Test \quad r_{tt} = \frac{100}{100-1} \left[ 1 - \frac{23/42}{(23/42)^2} \right] = 0/96$$

$$B \rightarrow Test \quad r_{tt} = \frac{100}{100-1} \left[ 1 - \frac{40/39}{(38/12)^2} \right] = \frac{100}{99} (0/9722) = 0/98$$

$$C \rightarrow Test \quad r_{tt} = \frac{100}{100-1} \left[ 1 - \frac{31/15}{(30/16)^2} \right] = \frac{100}{99} (0/966) = 0/90$$

$$D \rightarrow Test \quad r_{tt} = \frac{100}{100-1} \left[ 1 - \frac{23/28}{(24/2)^2} \right] = \frac{100}{99} 0/96 \cong 0/92$$

نمونه‌ای از پیشرفت افراد نشان‌دهنده پیشرفت کلی آن‌ها در سطح یا سطوحی از یک موضوع یا زمینه‌ای است که محتوای آزمون را تشکیل می‌دهد.

هر آزمون در حقیقت یک وسیله اندازه‌گیری برای تعیین مقدار صفت بخصوصی در افراد است. اما قبل از استفاده عملی از یک آزمون باید از راه تحقیقات تجربی نشان داد که نمره‌های حاصل از این آزمون، تا چه اندازه شاخص دقیق صفت مورد اندازه‌گیری هستند. بنابراین «داده‌های حاصل از اندازه‌گیری باید از دو لحاظ قابل انطباق باشند: اولاً معنای مشخصی داشته باشند و ثانیاً باز یافتنی باشند» (مگنسون، ۱۹۶۷، ترجمه براهنی، ۱۳۷: ۷۸). هرگاه آزمون از لحاظ اول کارآمد باشد می‌گوییم دارای اعتبار است و هرگاه مقصود دوم را برآورده سازد می‌گوییم دارای پایایی است.

<sup>1</sup> - Polmantir.

این دو از خصایصی هستند که برای مفید و مؤثر شدن روش‌های جمع‌آوری داده‌ها شرط اساسی به شمار می‌روند.

### مقایسه عملکرد دانش‌آموزان دختر و پسر در درس ریاضی سال اول دبیرستان

برای مقایسه پیشرفت تحصیلی دانش‌آموزان دختر و پسر در درس ریاضی سال اول دبیرستان شهرستان سنندج، از آزمون  $t$  برای گروه‌های مستقل با حجم زیاد استفاده شده است.

شاخص‌های آماری حاصل از توزیع نمرات دانش‌آموزان دختر و پسر

شاخص‌های آماری	n	M	S <sup>2</sup>
دانش‌آموزان پسر	۵۰۰	۴۵	۶۰/۳۸
دانش‌آموزان دختر	۵۰۰	۶۵	۳۹/۷۲

$$t = \frac{|M_1 - M_2|}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} \Rightarrow \frac{|20|}{\sqrt{\frac{(39/69)^2}{500} + \frac{(60/38)^2}{500}}} = 2/21$$

چون  $t$  محاسبه شده از  $t$  بحرانی در سطح  $0/01$  ( $2/21$ ) با درجه آزادی  $119$  بزرگ‌تر است در نتیجه فرض صفر در سطح  $0/01$  رد می‌شود و فرض خلاف تأیید می‌گردد. یعنی بین پیشرفت تحصیلی دانش‌آموزان دختر و پسر در درس ریاضی سال اول دبیرستان تفاوت معناداری وجود دارد. چون میانگین عملکرد دخترها در درس ریاضی بیشتر از میانگین عملکرد پسرها در این درس است، نتیجه می‌شود که پیشرفت تحصیلی دخترها بهتر از پسرها می‌باشد. نتیجه این که پیشرفت تحصیلی دخترها بهتر از پسرها می‌باشد. در استعداد تجسم فضایی تشخیص جهت و هدف‌گیری پسرها بهتر از دخترها است. در استعداد ریاضی پسرها به مراتب بهتر از دخترها هستند (گنجی، ۱۳۷۱).

### بحث

تعریف اندازه‌گیری به طوری که مورد قبول همه باشد، کار آسانی نیست. از تعریف بسیار ساده‌ای که در لغت‌نامه دهخدا آمده است شروع می‌کنیم. «اندازه‌گیری یعنی عمل اندازه گرفتن، یعنی طول و عرض و عمق را تعیین کردن، پیمانانه کردن، کشیدن، قیاس کردن مقدار و حد چیزی را سنجیدن» (گنجی، ۱۳۷۳: ۲۵) و یا چنان که در کتاب روش‌های آماری در روانشناسی و علوم

تربیتی (دلور، ۱۳۷۰: ۲۰) آمده است، اندازه‌گیری فرایندی است که از طریق آن مشاهده‌ها به عدد تبدیل می‌شود. به‌طور کلی اندازه‌گیری عبارت است از نسبت عددی دادن به یک صفت یا حادثه بر اساس یک قانون معین (هومن، ۱۳۶۸). اندازه‌گیری در تعلیم و تربیت را می‌توان به‌عنوان وسیله‌ای برای گردآوری یک رشته اطلاعات منظم درباره ویژگی‌های رفتار آدمی و تجزیه و تحلیل این اطلاعات به‌عنوان پایه و اساس یک سلسله تصمیم‌گیری‌های خاص تعریف کرد (همان: ۴).

در کتاب مبانی نظری آزمون‌های روانی تألیف مگنوسون (۱۹۶۷، ترجمه براهنی، ۱۳۷۰) آمده است: مقصود از اندازه‌گیری این است که مقدار صفت معینی را در اشیاء مختلف با دستگاه اعداد نشان دهیم. به‌عبارتی دیگر جمع‌آوری اطلاعات منظم درباره پدیده‌های مورد مطالعه و نمایش کمی این اطلاعات را می‌توان اندازه‌گیری نامید. اندازه‌گیری زیربنای اکثر فعالیت‌های ارزشیابی را تشکیل می‌دهد (سیف، ۱۳۷۳: ۲۱-۱۸). بنا به تعریفی که در کتاب معروف نظریه روان‌سنجی (نانالی، ۱۹۷۸: ۳) آمده است، «اندازه‌گیری از قواعدی تشکیل می‌یابد که برای نسبت دادن اعداد به اشیاء [یا افراد] به کار می‌رود.

- براساس یافته‌های این پژوهش عملکرد پیشرفت تحصیلی دانش‌آموزان دختر شهرستان سنندج در درس ریاضی بهتر از عملکرد پیشرفت تحصیلی دانش‌آموزان پسر می‌باشد. نتایج این پژوهش با تحقیق کشاورز (۱۳۷۶) هماهنگی دارد.

نتیجه این پژوهش نیز با نتایج و میانگین امتحانات خرداد ۸۱-۸۲ شهرستان سنندج از نظر عملکرد هماهنگی دارد. یعنی میانگین نمرات درس ریاضی سال اول دبیرستان شهرستان سنندج (دخترها در حدود ۱۰ و پسرها در حدود ۸) می‌باشد.

نتایج این پژوهش با نتایج پژوهش‌های انجام‌شده توسط روان‌شناسان مغایرت دارد. هر آزمون در حقیقت یک وسیله اندازه‌گیری برای تعیین مقدار صفت به‌خصوصی در افراد است. اما قبل از استفاده عملی از یک آزمون باید از راه تحقیقات تجربی نشان داد که نمره‌های حاصل از این آزمون، تا چه اندازه شاخص دقیق صفت مورد اندازه‌گیری هستند. بنابراین «داده‌های حاصل از اندازه‌گیری باید از دو لحاظ قابل انطباق باشند:

اولاً معنای مشخصی داشته باشند و ثانیاً باز یافتنی باشند» (مگنوسون، ۱۹۶۷، ترجمه براهنی، ۱۳۷۰: ۷۸). هرگاه آزمون از لحاظ اول کارآمد باشد می‌گوییم دارای اعتبار است و هرگاه مقصود دوم را برآورده سازد، می‌گوییم دارای پایایی است. این دو از خصایصی هستند که برای مفید و مؤثر واقع شدن روش‌های جمع‌آوری داده‌ها شرط اساسی به‌شمار می‌روند. فرض دیگر آن است که نرم یا هنجارهای سنجش نتایج آزمون‌ها بر نمونه‌هایی از جامعه آماری مبتنی است که نمونه‌هایی از طبقات مختلف افراد را شامل می‌شود. در این زمینه باید دلایل و فرضیه مذکور مورد بررسی

قرارگیرد و هم‌چنین نسبت به تهیه نرم‌های محلی که با ویژگی‌های دانش‌آموزان مدارس و افراد جامعه محلی انطباق بیشتری دارند، اقدام شود. برای مصرف‌کنندگان آزمون‌های روانی این سؤال مطرح می‌شود که ضریب اعتبار یک آزمون چقدر باید باشد تا با اطمینان کافی از آن استفاده شود. پاسخ این سؤال به هدف اندازه‌گیری بستگی دارد. اگر آزمون برای هدف‌های پژوهشی بکار می‌رود ضریب اعتبار بین ۰/۷۰ تا ۰/۸۰ کافی به نظر می‌رسد. در بیشتر پژوهش‌ها، هدف پژوهشگر این است که میزان همبستگی دو متغیر را برابر کند. در این گونه موارد استفاده از آزمون‌های خیلی معتبر که با صرف هزینه نیروی انسانی زیاد ساخته می‌شود ممکن است چندان ضروری نباشد (نانالی، ۱۹۷۸).

اما در مطالعات بالینی، بالابودن ضرایب اعتبار آزمون‌ها اهمیت فراوانی دارد. هنگامی که آزمونی برای تصمیم‌گیری مهم درباره آینده افراد به کار بسته می‌شود، باید خطای اندازه‌گیری به حداقل ممکن کاهش داده شود. در این صورت آزمونی با ضریب اعتبار ۰/۹۰ ممکن است آزمون مناسبی نباشد. اغلب گفته شده است هنگامی که تصمیم‌گیری مهم درباره آینده افراد مطرح است، ضریب اعتبار آزمون باید بیش از ۰/۹۵ باشد. در جدول زیر ضرایب اعتبار به‌دست آمده در مورد انواع آزمون‌های مختلف نشان داده شده است.

نوع آزمون	ضرایب اعتبار		
	کم	متوسط	زیاد
مجموعه آزمون‌های پیشرفت تحصیلی	۰/۶۶	۰/۹۲	۰/۹۸
آزمون‌های توانایی تحصیلی	۰/۵۶	۰/۹۰	۰/۹۷
مجموعه آزمون‌های استعداد	۰/۲۶	۰/۸۸	۰/۹۶
آزمون‌های عینی شخصیت	۰/۴۶	۰/۸۵	۰/۹۷
پرسش‌نامه‌های علائق	۰/۴۲	۰/۸۴	۰/۹۳
مقیاس‌های نگرش‌سنج	۰/۴۷	۰/۷۹	۰/۹۸

پولمن تیر فرضیه‌هایی را تهیه و کاربرد آزمون‌های استاندارد در آن‌ها را نشان داده است. پیشرفت افراد نشان‌دهنده پیشرفت کلی آن‌ها در سطح یا سطوحی از یک موضوع یا زمینه‌ای است که محتوای آزمون را تشکیل می‌دهد (شریفی، ۱۳۸۱).

## منابع

- اردبیلی، یوسف (۱۳۴۶). *اصول علمی تهیه، اجرا و استاندارد کردن تست*. تهران: مؤسسه تحقیقات تربیتی دانشسرای عالی تهران.
- بابلیان، اسماعیل؛ جلیلی، میرزا؛ شهریاری، رضا؛ و مدقالچی، علیرضا (۱۳۸۲). *ریاضیات سال اول نظام جدید*. تهران: دفتر برنامه‌ریزی کتب درسی.
- آناستازی (۱۳۷۹). *روان‌آزمایی*. ترجمه محمد تقی براهنی. تهران: انتشارات دانشگاه تهران، چاپ چهارم.
- بست، جان (۱۳۶۷). *روش‌های تحقیق در علوم تربیتی*. ترجمه حسن پاشا شریفی، و نرگس طالقانی، تهران: انتشارات وزارت فرهنگ و ارشاد اسلامی.
- بهرامی، هادی (۱۳۷۲). *کاربرد و تشخیص آزمون‌های فرافکنی شخصیت*. تهران: انتشارات دانا.
- بهرشاد، محمد (۱۳۵۹). *روان‌سنجی. تحول در امتحانات*، شماره ۱۷، مردادماه.
- پورکاظمی (۱۳۶۸). *ریاضیات عمومی*. تهران: نشر نی، جلد اول.
- گنجی، حمزه (۱۳۷۱). *تفاوت‌های فردی کودکان*. تهران: انتشارات دانشگاه پیام نور.
- دلاور، علی (۱۳۷۰). *روش‌های آماری در روان‌شناسی و علوم تربیتی*. تهران: انتشارات دانشگاه پیام نور.
- رمضانی، خسرو و غفاری، ابوالفضل (۱۳۷۳). *هنجاریابی مقیاس عزت‌نفس*. پایان‌نامه کارشناسی ارشد روانشناسی، دانشگاه علامه طباطبائی.
- سیف، علی اکبر (۱۳۷۱). *اندازه‌گیری و ارزشیابی پیشرفت تحصیلی*. تهران: انتشارات آگاه.
- شریفی، حسن پاشا (۱۳۸۱). *اصول روان‌سنجی و روان‌آزمایی*. تهران: انتشارات رشد.
- هومن، حیدرعلی (۱۳۷۸). *اندازه‌گیری روانی و تربیتی و فن تهیه تست*. تهران: انتشارات دانشگاه تهران.
- هومن، حیدرعلی (۱۳۷۸). *روان‌سنجی کاربردی*. تهران: انتشارات دانشگاه تهران.
- هومن، حیدرعلی (۱۳۷۲). *استنباط آماری*. تهران: نشر دیبا.
- مگنوسون (۱۳۷۰). *مبانی نظری آزمون‌های روانی*. ترجمه محمد تقی براهنی، تهران: انتشارات دانشگاه تهران.
- نادری، عزت‌ا... و سیف‌نراقی، مریم (۱۳۷۰). *سنجش و اندازه‌گیری و بنیادهای تحلیلی ابزارهای آن در علوم تربیتی و روانشناسی*. تهران: انتشارات معین.

Alken, L.R. (1964). *Psychological testing and assessment*.  
 Rath, L. *Pre- Algebra: A problem – solving approach*. Charlese Merrill Publications.  
 Douglas, B. & Richard, K. *Mathematics 1 and 2*. Charles Merrill Publications.  
 Thorndike, R.L. & Hagen, E. (1977). *Measurement and evaluation in psychology and education*. New York: John Wiley.  
 Educational measurement. Thorndike.