

رده‌بندی شناختی سؤالات بر اساس الگوی توان-پیچیدگی مطالعه موردی: سؤالات آزمون آزمایشی حساب دیفرانسیل

سلیمان ذوالفقارنسب*

رضا محمدی**

جواد حاتمی***

چکیده

تهیه سؤالات دشوار با الگوریتم‌های محاسباتی پیچیده، منجر به روندهای نابجا مثل راهبردهای آموزش تست‌زنی و آمادگی برای آزمون در بین آزمون‌شوندگان شده است. داشتن استانداردهای آموزشی دقیق و کاربرد الگوهای رده‌بندی کمک می‌کند تا ارزشیابی بهتری از برنامه‌های آموزشی، محتوای درسی و همین‌طور پیشرفت تحصیلی دانش‌آموزان به دست آید. در این نوشته به بررسی نمونه‌ای کوچک از مجموعه سؤالات چهارگزینه‌ای حساب دیفرانسیل در چهارچوب هدف‌های رفتاری پرداخته شد که مبتنی بر ترکیب رویکرد بلوم و وب بود؛ هدف معرفی رویکرد توان-پیچیدگی در رده‌بندی سؤالات آزمون‌ها بوده است. ۱۲ سؤال از یک آزمون ۳۰ سؤالی که روی یک نمونه ۳۴۰۹ نفری از دانش‌آموزان اجرا شد که بر اساس رویکرد توان-پیچیدگی رده‌بندی گردید. نتیجه مربوط به ۳ سؤال آن در این مقاله آمده است. نتایج نشان داد که سؤالات این آزمون همگی دانش‌روندی را اندازه می‌گیرند و عمدتاً در سطح ۳ پیچیدگی قرار داشته‌اند. همچنین ۹۲٪ سؤالات انتزاعی بوده و برای پاسخ‌گویی صحیح به هر سؤال هم‌زمان باید چندین الگوریتم محاسباتی مختلف فراخوانی شوند. سؤالاتی که تنها از یک الگوریتم برای حل آن‌ها استفاده شده و سطوح پیچیدگی کمتری داشته‌اند از پارامترهای آماری بهتری مثل مقیاس‌پذیری برخوردار بوده‌اند و تابع آگاهی آن‌ها اطلاعات بیشتری از توانایی آزمون‌شوندگان فراهم می‌کردند.

واژه‌های کلیدی: اهداف آموزشی، ارزشیابی، پیچیدگی شناختی، رده‌بندی بلوم، ژرفای دانش وب، حل

مسئله

این مقاله برگرفته از طرح پژوهشی است که در سازمان سنجش آموزش کشور به پایان رسیده است.

* کارشناس پژوهشی مرکز تحقیقات سازمان سنجش آموزش کشور. (نویسنده مسئول)

salarnik2001@yahoo.com

** استادیار مرکز تحقیقات سازمان سنجش آموزش کشور

*** دانشیار دانشکده علوم انسانی، دانشگاه تربیت مدرس

مقدمه

امروزه آزمون‌های مختلفی برای دوره‌های آموزشی تهیه و حجم زیادی از هزینه و زمان برای تهیه سؤالات صرف می‌شود. هدف نهایی این آزمون‌ها عمدتاً ارزشیابی پیشرفت تحصیلی دانش‌آموزان و دانشجویان است، بدون این‌که داوری خاصی درباره اثربخشی سؤالات، آزمون‌ها، مواد آموزشی و کیفیت آموزش صورت بگیرد و یا معلمان و یا دست‌اندرکاران تهیه و توسعه آزمون‌ها برای تجدیدنظر در روش کار خود از این نتایج ارزشیابی استفاده کنند. در ارزشیابی آموزشی عملکرد یادگیرندگان با توجه به هدف‌های آموزشی دقیق و جزئی سنجش می‌شود و همین هدف‌ها هستند که ملاک قضاوت توفیق معلمان و یادگیرندگان قرار می‌گیرند (کراتول^۱، ۲۰۰۲). بخش کوچکی از نظام ارزشیابی آموزشی در ایران، ارزشیابی از پیشرفت تحصیلی دانشجویان و دانش‌آموزان در قالب هدف‌های رفتاری است و دست‌کم سؤالات آزمون‌ها نقش کلیدی در این ارزشیابی دارد. در سال‌های اخیر عمده تلاش‌های سازمان سنجش اجرای یک آزمون سراسری عادلانه در سطح کشور بوده است؛ اما مسائل امنیتی و خطرپذیری بالای چنین آزمونی باعث شده که عمدتاً نقش مسائل فنی در ساخت سؤالات استاندارد کم‌رنگ شود. اغلب معلمان نیز آزمون‌هایی را تهیه می‌کنند که معمولاً انعکاس‌دهنده محتوای خاصی از مطالب آموزش داده شده کتاب است و افراد برای ادامه تحصیل باید آزمون‌هایی بگذرانند که سؤالات آن‌ها بسیار دشوار و بدون پیچیدگی شناختی متناسب با محتوای آموزش داده شده و یا حتی سطح تحصیلی آن‌ها است. سؤالاتی که در تهیه آن‌ها دستورالعمل‌های استاندارد در نظر گرفته شده به‌خوبی کمک می‌کنند تا یک جدول رده‌بندی از ابعاد آن‌ها تهیه شود و بدین‌وسیله بتوان اهداف آموزشی پیش‌بینی شده را ارزیابی کرد. اگر برای تهیه آزمون‌هایی که سطوح بالای تفکر را اندازه بگیرند مرجعی وجود داشته باشد، این کار کمک می‌کند تا معلمان در کلاس‌های خود (به‌جای تست‌زنی) سطوح بالای تفکر را آموزش دهند و آزمون‌سازان در تهیه سؤالات تستی ملاک معین و استانداردهای از پیش تعیین‌شده‌ای برای ترکیب سرفصل‌های مختلف در تهیه یک سؤال به‌ویژه در درس ریاضی خواهند داشت. بر این اساس تهیه

چهارچوب شناختی برای آزمون‌هایی که سطوح بالای تفکر را اندازه بگیرند موجب تسهیل و توسعه تفکر سطح بالا در دانش‌آموزان خواهد شد (راثمن^۱، ۲۰۰۲).

تابه‌حال طبقه‌بندی‌های مختلفی برای هدف‌های رفتاری صورت گرفته که نخستین آن‌ها به‌وسیله گروهی از متخصصان آموزش و ارزشیابی پیشرفت تحصیلی تهیه شده و به نام بنجامین اس. بلوم^۲ (۱۹۵۶) که سرپرستی این گروه را به عهده داشته معروف گشته است. عمده طبقه‌بندی‌های دیگری که بعدها به وجود آمدند با الگوبرداری از طبقه‌بندی بلوم بوده است. در این طبقه‌بندی، هدف‌های آموزشی ابتدا به سه طبقه کلی با نام‌های حوزه شناختی، حوزه عاطفی و حوزه روانی حرکتی تقسیم شده‌اند و هر حوزه نیز دارای طبقه‌بندی‌های مخصوص به خود است (سیف، ۱۳۸۵).

هدف‌های حوزه شناختی که در ادبیات پژوهش توان شناختی^۳ نامیده می‌شود به جریان‌هایی که با فعالیت‌های ذهنی و فکری آدمی سروکار دارند مربوط می‌شوند. از این لحاظ این حوزه مهم‌ترین حوزه یادگیری به شمار می‌آید، زیرا اکثریت فعالیت‌های تحصیلی آموزشگاه‌ها و غالب موضوع‌های درسی و هدف‌های آموزشی به این بخش مربوط می‌شوند. آزمون‌های موردنیاز برای سنجش هدف‌های حوزه شناختی، آزمون‌های توانایی شناختی هستند که اغلب به‌صورت کتبی تهیه و اجرا می‌شوند. بلوم (۱۹۵۶) در شکل اولیه نظریه خود هدف‌های حوزه شناختی را به عبارت بهتر توان شناختی را در شش طبقه کلی شامل: دانش، فهمیدن، به کار بستن، تحلیل، ترکیب و ارزیابی^۴ جای داد و آن‌ها را به‌عنوان هدف‌های حوزه شناختی در یک دوره تحصیلی معرفی کرد.

در طبقه‌بندی تجدیدنظر شده طبقات کاربرد، تحلیل و ارزیابی با همان محتوای اصلی حفظ شدند. اما به‌صورت وجه فعلی آن‌ها یعنی به کار بستن، تحلیل کردن و ارزیابی کردن بیان شدند و ترکیب کردن جای خودش را به ارزیابی داد و به آفریدن تغییر نام داده است؛ بنابراین در رده‌بندی تجدیدنظر شده طبقات به‌صورت به یادآوردن، فهمیدن، به کار بستن، تحلیل کردن، ارزیابی کردن و آفریدن^۵ آمده است.

-
1. Rothman
 2. Bloom
 3. cognitive ability
 4. knowledge, comprehension, application, analysis, synthesis & evaluation
 5. Remembering, understanding, applying, analyzing, evaluation and creating

جدول ۱. ترتیب رده‌بندی اولیه بلوم و نسخه تجدیدنظر شده آن

ارزیابی	ترکیب	تحلیل	کاربرد	درک و فهم	دانش	رده‌بندی بلوم (۱۹۵۶)
آفریدن	ارزیابی کردن	تحلیل کردن	به کار بستن	فهمیدن	به یاد آوردن	رده‌بندی تجدیدنظر شده (آندرسون ^۱ ، کراتول، ایراسیان، کریک شانک، مایر و همکاران، ۲۰۰۱).

در رده‌بندی اولیه بلوم (۱۹۵۶) تنها یک بعد فرایند شناختی ارائه می‌داد (ردیف بالای جدول ۲) و به‌طور آشکارا به تعیین سطوح پیچیدگی به‌گونه‌ای که در رویکرد وب وجود دارد پرداخته نشده بود. اما سطوح متفاوتی از پیچیدگی شناختی به‌طور ضمنی^۲ در افعالی که برای طبقه‌بندی به کار می‌رود در این رده‌بندی وجود دارد. به عبارتی این رده‌بندی شش سطح پیچیدگی در حوزه شناختی را شناسایی کرد که از سطح دانش با محتوای فراخوانی اطلاعات^۳ یا بازشناسی حقایق^۴ در پایین‌ترین سطح شروع می‌شد و به‌طور افزایشی به سطوح بالای پیچیدگی و انتزاع در رده‌ی ارزیابی^۵ ادامه می‌یافت (هس^۶، کارلوک، جونز و واکاپ، ۲۰۰۹).

در شکل تجدیدنظر شده این طبقه‌بندی که در سال ۲۰۰۱ انتشار یافت شیوه بیانی این شش طبقه از وجه اسمی (به‌عنوان مثال کاربرد^۷) به‌صورت وجه فعلی (به کار بردن^۸) بیان شدند و بعد دانش به‌جای سه طبقه عمده به چهار طبقه افزایش یافت. سه تا از آن دربرگیرنده زیرطبقات اصلی دانش بر اساس همان چهارچوب اصلی بلوم است و زیر طبقه جدیدی به نام دانش فراشناختی به آن اضافه شد که در تهیه طرح اولیه مشخص نشده بود (به جدول ۲ نگاه کنید). دانش فراشناختی اشاره دارد به شناخت فرد از فرایند فکر کردن خودش و آگاهی از این‌که چگونه این فرایند را به‌صورت مؤثری

1. Anderson
2. implicit
3. recall of information
4. or recognition of facts
5. evaluation
6. Hess
7. application
8. applying

دست‌کاری و با مهارت اداره کند (آندرسون، کراتول، ایراسیان، کریک شانک، مایر و همکاران ۲۰۰۱). به عبارتی دانش فراشناختی درباره شناخت به‌طورکلی است همچنین راجع به آگاهی و فهم فرد راجع به شناخت خودش است. پژوهشگران به اهمیت این مسئله پی بردند که دانش‌آموزانی که از فعالیت‌های فراشناختی خودآگاهی دارند و سپس از این دانش برای تعدیل راه‌هایی که فکر یا عمل می‌کردند، به‌طور مناسبی استفاده می‌کنند، افزایش معنی‌داری در عملکرد آن‌ها به وجود می‌آید (مارزانو^۱، ۲۰۰۶).

جدول ۲. ابعاد دانش و فرآیند شناختی

بعد فرایند شناختی						بعد دانش
۶	۵	۴	۳	۲	۱	
آفریدن	ارزشیابی کردن	تحلیل کردن	به کار بستن	فهمیدن	به یادآوردن	

الف-دانش امور واقعی

ب-دانش مفهومی

پ-دانش روندی

ت-دانش فراشناختی

(سیف ۱۳۸۵، اقتباس از پین، ۲۰۰۳؛ آندرسون، کراتول و همکاران، ۲۰۰۱)

بسیاری از آموزگاران به‌اشتباه معتقدند که دانش تنها به حقایق محدودی اشاره دارد و این نوع دانش را می‌توان از طریق حفظ کردن^۲ یاد گرفت. اما بلوم دانش امور واقعی را نیز به‌عنوان گونه‌ای از دانش در این طبقه گنجانده و چند زیرمجموعه به این طبقه اضافه کرد (اوسترهوف^۳، ۲۰۰۸). سپس روان‌شناسان شناختی معاصر اولین طبقه یعنی «دانش» در طبقه‌بندی بلوم را دانش بیانی^۴ نام نهادند و «فهمیدن، به کار بستن، تحلیل، ارزشیابی و آفریدن» را در طبقه دوم یعنی دانش روندی^۵ معرفی کردند. خود بلوم این پنج بخش اخیر را تحت عنوان «توانایی‌ها و مهارت‌های عقلانی^۶» نام نهاد (کانفری^۷،

1. Marzano
2. memorization
3. Oosterhof
4. declarative knowledge
5. procedural knowledge
6. intellectual abilities & skills
7. Confrey

۱۹۹۰؛ اوسترهوف، ۲۰۰۸). تفاوت‌های بین زیرطبقات توانایی‌ها و مهارت‌های عقلانی بلوم و آنچه در روان‌شناسی شناختی به‌عنوان دانش بیانی و دانش روندی معرفی شده بود، بعضی مشکلات را در استفاده از این طبقه‌بندی برای رده‌بندی سؤالات آزمون‌ها هنگام سنجش و اندازه‌گیری در مقیاس بزرگ به وجود آورد.

نرمان آل وب^۱ (۲۰۰۲) الگویی ارائه داد که به‌وسیله آن سطوح پیچیدگی شناختی سؤالات آزمون‌ها را می‌توان در چهار سطح پیچیدگی طبقه‌بندی کرد. الگوی ژرفای دانش^۲ وب، پیچیدگی فرایند ذهنی را توصیف می‌کند که هنگام پاسخ‌گویی به یک سؤال، انجام یک تکلیف و یا تولید یک محصول روی می‌دهد. این رویکرد با رده‌بندی بلوم تفاوت دارد که در آن به سطوح توانایی^۳ اشاره می‌کند. در عوض ژرفای دانش وب به محتوای استانداردهای یادگیری^۴ و آنچه برای تکمیل موفقیت‌آمیز یک تکلیف یا سؤال لازم است تمرکز می‌کند (سرویس خدمات تهیه آزمون‌ها، ۲۰۱۲). منطق طبقه‌بندی سؤالات بر اساس سطوح پیچیدگی وب نه بر توانایی دانش‌آموز و نه بر تعداد گزینه‌های انحرافی سؤال - که می‌تواند دشواری سؤال را تحت تأثیر قرار دهد - بلکه بر انتظاری که از سؤال می‌رود متمرکز است؛ هر سؤال یا تکلیف یک الزام فکری ایجاد می‌کند، یعنی چیزی در سؤال وجود دارد که دانش‌آموز را وادار می‌کند که به یاد آورد، بفهمد، تحلیل کند و انجام دهد. این الزام‌های فکری در هر تکلیف یا در تهیه سؤال باید با در نظر گرفتن این امر که دانش‌آموز با مفاهیم اولیه تکلیف آشنا است ایجاد شود. همچنین سؤالاتی که برای آزمون انتخاب می‌شوند باید بر اساس تناسبشان با سطح کلاسی باشد. اما پیچیدگی سؤالات، از برنامه درسی^۶ ویژه‌ای که دانش‌آموز آن

1. Norman L. Webb

2. depth of knowledge model (DOK)

3. ability levels

۴. استانداردهای یادگیری توصیفات مختصر و نوشته شده‌ای هستند از آنچه که انتظار می‌رود دانش‌آموزان در یک مرحله معین از تحصیل بدانند و بتوانند انجام دهند. استانداردهای یادگیری اهداف آموزشی را توصیف می‌کنند - به عنوان مثال آنچه که دانش‌آموزان می‌بایست در انتهای یک دوره، سطح تحصیلی یا در طول یک پایه تحصیلی یاد بگیرند - اما استانداردهای یادگیری هیچ تمرین درسی خاصی، روش برنامه درسی یا شیوه سنجش خاصی را توصیف نمی‌کنند.

5. general educational development (GED)

6. curriculum

را تجربه کرده مستقل باقی می‌ماند (دپارتمان آزمون سنجش جامع فلوریدا^۱، ۲۰۰۸ و ۲۰۱۲). به‌هرحال این الگو مبتنی است بر پیچیدگی شناختی نه دشواری سؤال. یعنی می‌توان از این الگوی برای شناسایی میزان پیچیدگی فرایند ذهنی که باید هنگام پاسخ به یک سؤال، یا انجام یک تکلیف یا تولید یک محصول رخ دهد استفاده کرد.

به‌هرحال اگرچه این دو رویکرد در رابطه با مشخص کردن پیچیدگی تفکر ماهیت مشابهی دارند اما هم رده‌بندی بلوم و هم ژرفای دانش وب در اهداف و کاربردهایشان با یکدیگر تفاوت دارند. در عمل آموزگاران که از سطوح رده‌بندی بلوم استفاده می‌کنند عملی که در فعل جمله مربوط به تکلیف یا سؤال هست را مبنای رده‌بندی آن تکالیف به سطوح پیچیدگی قرار می‌دهند. برعکس در رویکرد وب فعل جمله باید در متن تکلیف در نظر گرفته شود و آنچه به دنبال فعل جمله قرار می‌گیرد اهمیت بیشتری دارد تا خود فعل جمله. به‌عنوان مثال فعل توصیف کردن در رده‌بندی بلوم احتمالاً در سطح اول رده‌بندی یعنی دانش قرار می‌گیرد در حالی که در رده‌بندی وب فعل توصیف کردن ممکن است آن تکلیف را در بیش از یک رده قرار دهد یا آن را به سطوح بالاتر رده‌بندی جای دهد به‌عنوان مثال سؤال را تبدیل به یک انشا کند یا فرد را ملزم به تحلیل کند (هس و همکاران، ۲۰۰۹).

رده‌بندی بلوم مهارت‌های شناختی لازم برای مغز در انجام یک تکلیف را طبقه‌بندی می‌کند و «نوع فرایند تفکر» لازم برای پاسخ گفتن به یک سؤال را توصیف می‌کند. همچنین دربرگیرنده مفروضه‌هایی در رابطه با «رویکرد دانش‌آموز به سؤال^۲» بوده است که درک و مشخص کردن آن مشکل و ذهنی است. اما ژرفای دانش وب بیشتر ارتباط دارد به عمق دریافت فرد از محتوای آموزش داده شده که در مهارت‌های موردنیاز برای تکمیل یک تکلیف از آغاز تا پایان (مثل مهارت طراحی کردن، تحقیق و نتیجه‌گیری) خود را نشان می‌دهد. هم فرایند تفکر بلوم و هم ژرفای دانش وب در طراحی برنامه درسی، ارائه درس و تهیه برنامه سنجشی و اجرای آن کاربردهای مستقیم دارند (هس، ۲۰۰۵؛ هس و همکاران، ۲۰۰۹).

-
1. florida comprehensive assessment test (FCAT)
 2. the student's approach to the item

وب معتقد بود که در آزمون‌هایی که در سطح ملی اجرا می‌شوند (مثل آزمون‌های ورودی دانشگاه‌ها) به علت محدودیت زمانی برای چنین آزمون‌هایی سه سطح الگوی وی برای رده‌بندی کافی است و سطح چهارم را می‌توان در آزمون‌های معلم ساخته و یا کلاسی پیاده کرد. این رویکرد تحت عنوان الگوی ژرفای دانش با کمی تعدیل مورد استفاده مراکز آزمون‌سازی مختلفی قرار گرفته است. وب چهار سطح از پیچیدگی را برای دروس مختلف مشخص کرد (وب، ۲۰۰۲، ۲۰۰۵ و ۲۰۰۶). به‌عنوان مثال سطوح پیچیدگی برای دروس ریاضی عبارت‌اند از:

سطح ۱: بازخوانی و بازتولید^۱ (پیچیدگی کم). این طبقه به‌طور عمده بر یادآوری یا بازشناسی مفاهیم و اصول یاد گرفته شده قبلی متکی است. سؤالات عموماً آنچه دانش‌آموز باید انجام دهد را مشخص می‌کنند که اغلب انجام بعضی از روش‌هایی است که می‌توان آن را به‌صورت مکانیکی انجام داد. عمده‌ترین ویژگی این سطح آن است که به عهده دانش‌آموز واگذار نمی‌شود که یک راه‌حل یا روشی ابتکاری با پیچیدگی کم پیدا کند. به‌عنوان مثال می‌توان از بازخوانی حقایق و اصطلاحات و استفاده از کلمات برای بیان مفاهیم و روابط علمی و انجام فرایندهای عادی مثل اندازه‌گیری طول نام برد (وب و هرمان، ۲۰۰۶؛ اسفورزا و داریو^۲، ۲۰۱۴). مهارت‌های لازم برای پاسخگویی به سؤالاتی با پیچیدگی کم عبارت‌اند از:

- حل سؤالاتی با یک مرحله
- محاسبه جمع، منها و ضرب و تقسیم
- ارزیابی یک متغیر با توجه به مقادیر ویژه‌ای برای آن متغیر
- تشخیص یا ساختن یک نتیجه معادل
- یادآوری یا تشخیص یک حقیقت، عبارت یا ویژگی
- بازیابی اطلاعات از یک نمودار، جدول یا تصویر
- تشخیص واحدها یا ابزار مناسب برای اندازه‌گیری‌های متداول
- انجام یک تبدیل یک‌رقمی

1. recall & reproduction
2. Sforza & Dario

سطح ۲: مهارت‌ها و مفاهیم^۱ (پیچیدگی متوسط). سؤالاتی که در طبقه پیچیدگی متوسط قرار می‌گیرند نیازمند تفکر انعطاف‌پذیر بیشتری هستند و نسبت به سؤالاتی که دارای پیچیدگی کم هستند پاسخ‌گویان امکان انتخاب بیشتری دارند. پاسخ‌گویی به این‌گونه سؤالات فراتر از پاسخ‌های مبتنی بر مهارت هستند، پاسخ مشخص و ویژه نیست و به‌طور معمول پاسخ‌گویی به آن‌ها از یک گام فراتر می‌رود و دانش‌آموز باید با به‌کارگیری روش‌های غیررسمی استدلالی و راهبردهای حل مسئله و یکپارچه‌سازی مهارت و دانش در زمینه‌های مختلف تصمیم بگیرد چه کاری انجام دهد. به‌عنوان مثال می‌توان از توضیح روابط بین حقایق و متغیرها، توصیف و توضیح مثال‌های مربوط به یک مفهوم علمی و سازمان‌دهی و تفسیر داده‌ها و اطلاعات نام برد (وب و هرمان، ۲۰۰۶؛ اسفورزا و داریو، ۲۰۱۴). مهارت‌های موردنیاز برای پاسخ‌گویی به سؤالاتی که پیچیدگی متوسط دارند عبارت‌اند از:

- حل تکلیف یا مسئله نیازمند چندین عملیات است
- حل تکلیف یا مسئله دربرگیرنده تصویرسازی فضایی و یا استدلالی است
- انتخاب و یا به‌کارگیری ایده‌های مختلف در رابطه با موقعیت و هدف
- بازیابی اطلاعات از نمودار، جدول و یا اشکال و به‌کارگیری آن برای حل یک مسئله یا تکلیف
- تعیین یک برآورد منطقی
- بسط یک الگوی جبری یا هندسی
- فراهم ساختن یک توجیه برای گام‌هایی که در یک فرایند حل مسئله یا تکلیف برداشته می‌شود
- مقایسه اشکال یا اظهارات و گفته‌ها
- بازنمایی یک موقعیت از لحاظ ریاضی به بیش از یک روش
- فرمول‌بندی یک مسئله عادی با توجه به داده‌ها و شرایط

سطح ۳: تفکر راهبردی^۲ (پیچیدگی بالا). سؤالاتی که پیچیدگی بالایی دارند دانش‌آموز را ملزم به تفکر بسیار بیشتری می‌کنند. دانش‌آموزان باید در زمینه‌هایی مثل استدلال، برنامه‌ریزی، تحلیل، قضاوت و تفکر خلاق تفکر انتزاعی خیلی بیشتری داشته باشند.

1. skills & concepts
2. strategic thinking

سؤالاتی که پیچیدگی بالایی دارند دانش‌آموز را ملزم می‌کنند که به روش‌های انتزاعی‌تر و پیچیده‌تری فکر کنند. به‌عنوان مثال می‌توان از مواردی همچون شناسایی یک سؤال و مسئله علمی و طراحی یک تحقیق برای یک مسئله علمی و حل یک مسئله غیرعادی و فرمول‌بندی نتایج برای داده‌های تجربی نام برد (وب و هرمان، ۲۰۰۶؛ اسفورزا و داریو، ۲۰۱۴). مهارت‌هایی که لازم است تا به سؤالاتی با پیچیدگی بالا به‌درستی پاسخ داد عبارت‌اند از:

- انجام روشی که دارای مراحل چندگانه است و مسیر تصمیم‌گیری چندمرحله‌ای دارد
 - حل مسئله غیرتکراری
 - حل مسئله به بیش از یک روش
 - توصیف چگونگی استفاده از ایده‌ها و افکار متفاوت برای اهداف مختلف
 - بسط یک الگوی جبری یا هندسی
 - توضیح و توجیه یک راه‌حل برای یک مسئله
 - توصیف، مقایسه و برابرسازی روش‌های حل (مسئله یا تکلیف)
 - فراهم‌سازی یک توجیه از لحاظ ریاضی
 - تحلیل شباهت‌ها و تفاوت‌ها بین روش‌ها و مفاهیم
 - فرمول‌بندی یک مسئله تازه و ارائه یک راه‌حل
 - فرمول‌بندی یک مدل ریاضی برای یک موقعیت پیچیده
 - تجزیه و تحلیل یا ارائه یک بحث استنتاجی^۱
- سطح ۴: تفکر گسترش‌یافته^۲. تفکر گسترش‌یافته مستلزم استدلال، طرح‌ریزی و توسعه پیچیده در یک دوره طولانی‌مدت است. اگر کار موردنظر مستلزم یک عملیات تکراری است و در آن تفکر ادراکی و انتزاعی و سطوح بالا به کار گرفته نمی‌شود دوره طولانی‌مدت به‌تنهایی نمی‌تواند عامل تمایزگذار باشد. در سطح ۴ الزام شناختی^۳ تکلیف باید بسیار بالا و کار باید بسیار پیچیده باشد. دانش‌آموز باید ایده‌های مختلف در حوزه‌های محتوایی و موضوعی متفاوت را با همدیگر مرتبط کند و باید در بین رویکردهای مختلف روشی را انتخاب کند که موجب حل موقعیت

1. deductive argument
2. Extended thinking
3. cognitive demand

پیش‌آمده شود. وب (۲۰۰۲) خاطر نشان کرد که به انجام رساندن یک تکلیف در سطح ۴ احتمالاً مستلزم یک دوره طولانی و فرصت‌های چندگانه برای مشاهده فرد است. همچنین تفکر گسترش‌یافته می‌تواند دربرگیرنده استدلال پیچیده، شرح و بیان ایده‌هایی که درون حوزه‌های محتوایی مختلف یا میان‌رشته‌ای است و طراحی یک راه‌حل برای یک مسئله پیچیده در موقعیتی که چندین راهبرد بالقوه می‌تواند وجود داشته باشد نیز هست. مثال وب از تکلیف سطح ۴ دربرگیرنده طراحی و اجرای یک بررسی و تحقیق علمی است که دارای بخش‌هایی مثل مشخص کردن مسئله، طراحی و اجرای یک آزمایش، تحلیل داده‌ها و تهیه نتیجه‌گیری‌ها است. در رده‌بندی وب همچون رده‌بندی بلوم حرکت از تفکر عینی و ساده به سطوح بالای تفکر انتزاعی و پیچیده است (وب و هرمان، ۲۰۰۶: اسفورزا و داریو، ۲۰۱۴).

بنابراین طبقه‌بندی‌هایی تحت عنوان پیچیدگی کم، پیچیدگی متوسط و پیچیدگی بالا، یک توصیف از الزامات شناختی است که یک سؤال برای دانش‌آموز ایجاد می‌کند. برای مثال پیچیدگی کم سؤال ممکن است دانش‌آموز را ملزم کند به این که سؤال را در یک گام حل کند. پیچیدگی متوسط ممکن است مستلزم گام‌های چندگانه‌ای باشد. سؤالاتی با پیچیدگی بالا دانش‌آموزان را ملزم به تحلیل و ترکیب اطلاعات می‌کنند. تمایزاتی که در پیچیدگی سؤال ایجاد می‌شود این اطمینان را حاصل می‌کند که سؤالات می‌توانند ژرفای دانش یادگیرندگان را در هر برنامه آزمون‌گیری موردسنجش و ارزشیابی قرار دهد (دپارتمان آزمون سنجش جامع فلوریدا، ۲۰۰۸ و ۲۰۱۲).

رویکرد وب بیشتر می‌تواند فرایند شناختی درگیر در حل سؤال، یعنی ژرفای دانش را معین کند. این الگو نوعی رده‌بندی است که هم با «عمق درک مطلب»^۱ از سوی فرد و هم با «هدف» یک فعالیت یادگیری ارتباط نزدیکی دارد که در مهارت‌های موردنیاز برای کامل کردن یک تکلیف از آغاز تا پایان (مثل برنامه‌ریزی، پژوهش و نتیجه‌گیری) آشکار می‌شود (هس، ۲۰۰۹). بعلاوه مبتنی بر «الزام یا نیاز شناختی در ذات سؤالات آزمون^۲ یا یک تکلیف» است [یعنی پیچیدگی] و لازم است که به «انتظارات برخاسته از سؤال یا تکلیف تمرکز کرد و نه به توانایی دانش‌آموز». بعلاوه سطوح ژرفای دانش وب

-
1. depth of content understanding
 2. cognitive demand inherent in the test items

باید انعکاس‌دهنده پیچیدگی فرایند شناختی مورد نیاز برای حل تکلیفی باشد که به منظور یک هدف طرح‌ریزی شده است نه دشواری تکلیف که معمولاً از طریق ضریب دشواری سؤال به دست می‌آید (راهنمای ژرفای دانش وب، ۲۰۰۵).

در توضیح استفاده از الگوی ژرفای دانش وب هنگامی که به یادگیری در مدرسه ارتباط دارد، باید فرایند دومرحله‌ای به کار برد. نخست باید دانش و مهارت‌ها را به سه طبقه عمده دانش بیانی، دانش روندی و حل مسئله^۱ تفکیک کرد و سپس سطوح پیچیدگی آن را آشکار ساخت؛ دانش بیانی دربرگیرنده اطلاعاتی است که می‌توان آن را به صورت زبانی بیان کرد و یا نوشت. دانش روندی به عنوان «چگونه انجام دادن چیزی» شناخته می‌شود که کاملاً با «توضیح و یا توصیف» آنچه شناخته شده متفاوت است. به عنوان مثال یک دانشجوی فیزیک اگر بتواند قانون اهم^۲ را توضیح دهد (ارتباط بین مقاومت الکتریکی، جریان و ولتاژ) مثالی است از دانش بیانی. درحالی که اگر قادر باشند قانون اهم را به کار گیرد بر دانش روندی ناظر است (برای مثال این که آیا مقاومت یک سیم الکتریسیته باعث می‌شود جریان و ولتاژ خاصی را در خودش عبور دهد؟). ماهیت دانش روندی پیچیده است. اوسترهوف (۲۰۰۸) دانش روندی را تحت عناوینی مثل ایجاد تمایزات، فهم و درک مفاهیم و به کارگیری قوانینی که بر روابط حاکم هستند سازمان‌دهی کرد. روان‌شناسان شناختی به این نتیجه رسیده‌اند که دانش روندی اغلب دربرگیرنده مهارت‌های حرکتی و راهبردهای شناختی هستند. پایه حل مسئله روی دانش بیانی و دانش روندی بنا می‌شود. حل مسئله در زمینه‌های خاص ممکن است دربرگیرنده راهبردهای متفاوت در آن حوزه خاص^۳ باشد و هنگامی که حل مسئله در حوزه‌های محتوایی متفاوتی باشد مثل نوشتن انشا و یا علوم، راهبردهای متفاوتی به کار گرفته می‌شود. گاگنی و یکویچ^۴ (۱۹۹۳) نشان دادند که حل مسئله زمانی وجود دارد که فرد یک هدف دارد اما هنوز نتوانسته یک ابزار برای دستیابی به آن مشخص کند. برای مثال حل مسئله می‌تواند دربرگیرنده فرایند نوشتن یک مقاله توصیفی، تهیه یک سخنرانی اقناع‌کننده و یا اجرای یک تحقیق و بررسی علمی باشد.

۱. در این تحقیق حل مساله به معنای حل تکالیف ریاضی مرسوم نیست.

2. Ohm's law
3. domain – specific strategies
4. Gagné & Yekovich

در فرایند سنجش، تهیه سؤالاتی که بتوانند مهارت‌های موردنظر در یک دوره آموزشی را اندازه بگیرند برای روایی^۱ تفسیر نمرات و تنظیم یا رده‌بندی سؤالات آزمون، تعیین این‌که آزمون موردنظر چه مهارت شناختی را اندازه می‌گیرد و نیز برای تهیه محتوای کتاب‌های آموزشی مهم است. هدف این پژوهش بررسی نمونه کوچکی از سؤالات آزمون حساب دیفرانسیل، به‌منظور مشخص کردن گستره ابعاد تفکر بر اساس رویکردی است که هسته اصلی از ترکیب رویکرد بلوم (یعنی توان) و رویکرد وب (یعنی پیچیدگی) است. از این‌رو، پرسش این پژوهش عبارت است از: نمونه موردی سؤالات حساب دیفرانسیل آزمون سراسری در چه سطحی از فرایندهای شناختی مبتنی بر رویکرد بلوم-وب طبقه‌بندی می‌شوند؟

روش

طرح این پژوهش به‌صورت موردی است. یک آزمون حساب دیفرانسیل ۳۰ سؤالی^۴ گزینه‌ای انتخاب شده که عمدتاً سؤالات آزمون‌های کنکور سال‌های قبل و یا شبیه به آن‌ها بوده است. این آزمون در سال ۱۳۹۲ روی یک گروه نمونه ۳۴۰۹ نفری دختر و پسر به‌وسیله مرکز خدمات آموزشی کارکنان سازمان سنجش به‌طور آزمایشی اجرا شده است. کارشناسان موضوعی که دانشجویان کارشناسی دانشگاه آمار و انفورماتیک سازمان سنجش با گرایش آمار و سنجش آموزشی بوده‌اند و یک دوره آموزش در زمینه اصول و مبانی سنجش و ارزشیابی آموزشی دیده بودند با راهنمایی گام‌به‌گام بر اساس الگوی توان- پیچیدگی^۲ با جزئیات زیادی که در پیوست (۱) آمده است هر سؤال رده‌بندی کردند. لازم به ذکر است که نمونه سؤالات دروس متفاوتی از دوره دبیرستان -با توجه به علاقه دانشجویان- به همین شیوه رده‌بندی شده است. این سؤالات در دو مرحله رده‌بندی شده‌اند؛ ابتدا نوع دانش درگیر در حل تکلیف باید مشخص می‌شد و سپس سطوح پیچیدگی آن‌ها. به‌رحال ماهیت عمده تکالیف درس ریاضی مبتنی بر دانش روندی است. نتایج به‌دست‌آمده از رده‌بندی هر کارشناس در مراحل مختلف موردبررسی قرار گرفته و ناهمسانی رده‌بندی سؤالات هر فرد با دیگر کارشناسان در

1. validity

2. capabilities-complexity model

کلاس مورد بررسی مجدد قرار گرفته تا یک هماهنگی و توافق کامل برای رده‌بندی هر سؤال به دست آید. خلاصه‌ای از رده‌بندی ۳ سؤال در بخش نتایج آمده است (جزئیات در پیوست ۱). برای بررسی بیشتر ویژگی‌های روان‌سنجی این سؤالات پارامترهای آن‌ها با نرم‌افزار Bilog-MG (زیموروفسکی، موراکی، میس لوی و بوک، ۱۹۹۰) به دست آمده است. با کنار هم قرار دادن اطلاعات مربوط به رده‌بندی و نتایج آماری، یافته‌هایی به دست آمده که در ادامه بیان شده است.

یافته‌ها

جدول ۳ برای طبقه‌بندی سؤالات مربوط به دانش روندی به سه سطح پیچیدگی به کار رفته است. همان‌طور که پیش‌تر بیان شد، برای استفاده از الگوی توان-پیچیدگی ابتدا نوع توان (دانش مفهومی، دانش روندی و حل مسئله) درگیر در فرایند حل تکلیف مشخص می‌شود. سپس سطوح پیچیدگی هر تکلیف با استفاده از شاخص‌ها و نشانه‌هایی که در جدول ۳ فهرست شده تعیین می‌شود.

جدول ۳. تعیین سطح پیچیدگی تکالیفی که مبتنی بر دانش روندی هستند در دروس ریاضی

(برگرفته از اوسترهوف و همکاران، ۲۰۰۸)

سطح پیچیدگی	تا چه حد دستورالعملی	تا چه حد توضیحات یا متغیرها	با چه دقتی روش انجام کار از طریق متغیرها هدایت شده است	چقدر شباهت با زمینه یا متنی که فرد در آن روش یا فرایند حل مسئله را یاد گرفته وجود دارد
۱	تکلیف یا عمل خاصی که انجام می‌گیرد به‌طور مستقیم بیان شده	محسوس	تنها یک راه خاص برای روش یا فرایند حل سؤال وجود دارد	احتمالاً متن شبیه به آن چیزی است که در آموزش به کار رفته
۲	تکلیف یا عمل خاصی که انجام می‌گیرد نیازمند مقداری استنباط است	ممکن است انتزاعی باشد	مقداری تغییر و یا چندگانگی در انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن شبیه یا معادل با آن چیزی است که در آموزش به کار رفته

گام‌های چندگانه‌ای به‌طور متوالی یا هم‌زمان برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌شود باید از دستورالعمل استنباط شود	تغییر و یا چندگانگی آشکاری در روش‌های انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن به‌طور کامل متفاوت با آن چیزی است که در آموزش به کار رفته
--	--	---	--

در ادامه نتایج مربوط به سه سؤال ۱، ۲۲ و ۲۴ از آزمون ۳۰ سؤالی آمده است. ابتدا دانش زیربنایی مربوط به پاسخ‌گویی صحیح به هر سؤال و الگوریتم محاسباتی آن بیان شده و بر همین اساس یک جدول سطوح پیچیدگی برای رده‌بندی هر سؤال تهیه شده و سپس نمودار تابع آگاهی در سمت راست و تابع ویژگی آن‌ها در سمت چپ ارائه شده است. از لحاظ پارامترهای آماری سؤال ۱ این آزمون به‌عنوان یک سؤال متوسط رو به ضعیف، سؤال ۲۲ به‌عنوان یک سؤال بد کارکرد و سؤال ۲۴ به‌عنوان نمونه‌ای از یک سؤال خوب با پارامترهای بهینه مشخص شده‌اند.

۱- اگر $a_n = \sqrt{n^2 + n}$ و $b_n = \frac{n^2 + 1}{n}$ ، هریک از دو دنباله $\{a_n - b_n\}$ و $\left\{\frac{a_n}{b_n}\right\}$ به

ترتیب چگونه‌اند؟

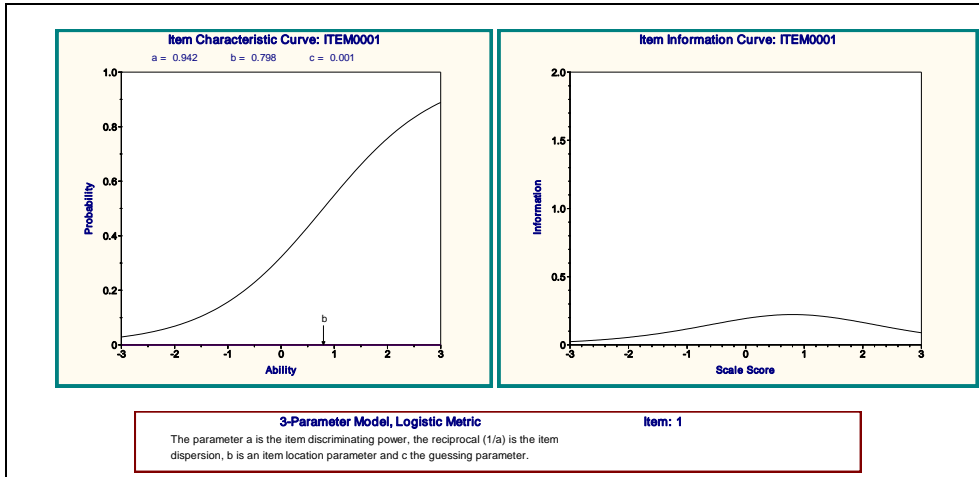
- (۱) همگرا- همگرا
 (۲) همگرا- واگرا
 (۳) واگرا- همگرا
 (۴) واگرا- واگرا

این سؤال نیازمند دانش مربوط به توابع، دنباله‌ها، همگرا یا واگرا بودن دنباله‌ها، توانایی انجام عملیات تقسیم و تفریق بر روی توابع است (برای مشاهده جزئیات بیشتر به پیوست ۱ نگاه کنید)

جدول ۴. سطوح پیچیدگی سؤال ۱

سطح پیچیدگی	تا چه حد	تا چه حد	با چه دقتی روش	چقدر شباهت با
چندگام یا عملیات طی می‌شود	دستورالعملی که در سؤال ارائه شده برای انجام عملیات حل سؤال مستقیم است	توضیحات یا متغیرها انتزاعی هستند	انجام کار از طریق متغیرها هدایت شده است	زمینه یا متنی که فرد در آن روش یا فرایند حل مسئله را یاد گرفته وجود دارد
۱- دربرگیرنده تنها یک گام یا عملیات است	تکلیف یا عمل خاصی که انجام می‌گیرد به‌طور مستقیم بیان شده	محسوس	تنها یک راه خاص برای روش یا فرایند حل سؤال وجود دارد	احتمالاً متن شبیه به آن چیزی است که در آموزش به کار رفته
۲- گام‌های چندگانه‌ای به‌طور متوالی برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌گیرد نیازمند مقداری استنباط است	ممکن است انتزاعی باشد	مقداری تغییر و یا چندگانگی در انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن شبیه یا معادل با آن چیزی است که در آموزش به کار رفته
۳- گام‌های چندگانه‌ای به‌طور متوالی یا هم‌زمان برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌شود باید از دستورالعمل استنباط شود	انتزاعی	تغییر و یا چندگانگی آشکاری در روش‌های انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن به‌طور کامل متفاوت با آن چیزی است که در آموزش به کار رفته

سمت راست نمودار ۱ بیانگر تابع آگاهی سؤال ۱ است و نشان می‌دهد سؤالاتی با چنین ویژگی‌هایی قدرت کمی (کمتر از ۰/۲۵) در به دست دادن آگاهی از توانمندی دانش‌آموزان و به عبارتی صفت مکنون زیربنایی ریاضی دارند. نمودار سمت چپ تابع ویژگی سؤال است. به علت شیب $a=0/94$ کم این سؤال طبیعتاً نمی‌تواند معیار خوبی برای تفکیک افراد قوی و متوسط و ضعیف باشند.



نمودار ۱. تابع ویژگی سؤال (چپ) و تابع آگاهی سؤال (راست)

۲۲- در تابع با ضابطه‌ی $f(x) = \begin{cases} x^3 + 1 & , x \geq 1 \\ 2x^2 - 1 & , x < 1 \end{cases}$ حاصل $\lim_{h \rightarrow 0^+} \frac{f(1-h) - f(1)}{h}$ کدام

است؟

(۱) -۴ (۲) ۴

(۳) -۳ (۴) وجود ندارد

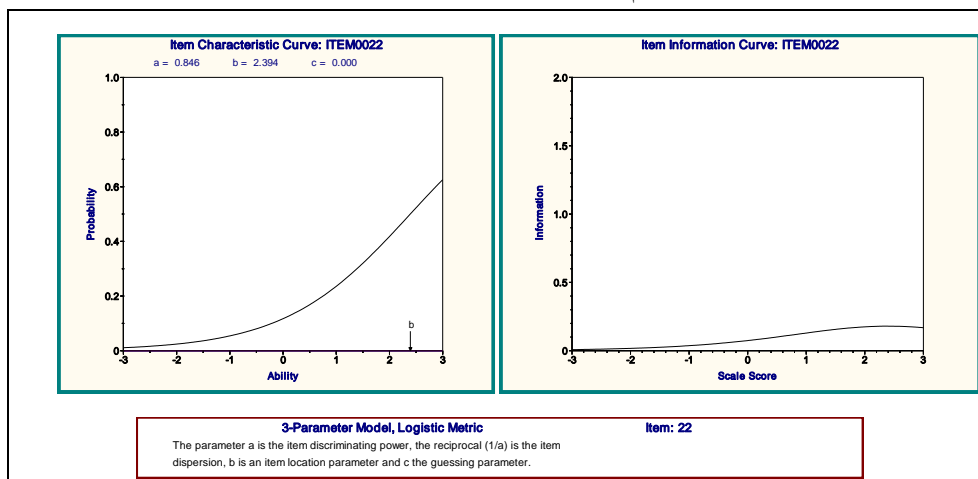
این سؤال نیازمند دانش در زمینه تابع، مشتق و حد است.

جدول ۵. سطوح پیچیدگی سؤال ۲۲

سطوح پیچیدگی	چندگام یا عملیات طی می‌شود	تا چه حد دستورالعملی که در سؤال ارائه شده برای انجام عملیات حل سؤال مستقیم است	تا چه حد توضیحات یا متغیرها انتزاعی هستند	با چه دقتی روش انجام کار از طریق متغیرها هدایت شده است	چقدر شباهت با زمینه یا متنی که فرد در آن روش یا فرایند حل مسئله را یاد گرفته وجود دارد
	دربرگیرنده تنها یک گام یا عملیات است	تکلیف یا عمل خاصی که انجام می‌گیرد به‌طور مستقیم بیان شده	محسوس	تنها یک راه خاص برای روش یا فرایند حل سؤال وجود دارد	احتمالاً متن شبیه به آن چیزی است که در آموزش به کار رفته

گام‌های چندگانه‌ای به‌طور متوالی برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌گیرد نیازمند مقداری استنباط است	مقداری تغییر و یا چندگانگی در انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن شبیه یا معادل با آن چیزی است که در آموزش به کار رفته
گام‌های چندگانه‌ای به‌طور متوالی یا هم‌زمان برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌شود باید از دستورالعمل استنباط شود	تغییر و یا چندگانگی آشکاری در روش‌های انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن به‌طور کامل متفاوت با آن چیزی است که در آموزش به کار رفته

تابع آگاهی سؤال ۲۲ نیز بیانگر ضعیف بودن این سؤال در به دست دادن آگاهی از توانمندی دانش‌آموزان در درس مربوطه است (کمتر از ۰/۲۵). بعلاوه شیب $a=0/85$ تخت این سؤال نشانگر ضعف آن در تفکیک افراد است به‌رغم این‌که این سؤال بسیار دشوار بوده $b=2/39$ و درصد کمی از پاسخ‌دهندگان به آن جواب درست داده‌اند. تهیه چنین سؤالاتی با ترکیب الگوریتم‌های محاسباتی مختلف کارایی کمی دارند.



نمودار ۲. تابع ویژگی سؤال (چپ) و تابع آگاهی سؤال (راست)

۲۴- در نقطه‌ای با کدام طول، خط مماس بر منحنی به معادله $y = x\sqrt{x}$ ، بر خط $2x + 3y = 1$ عمود است؟

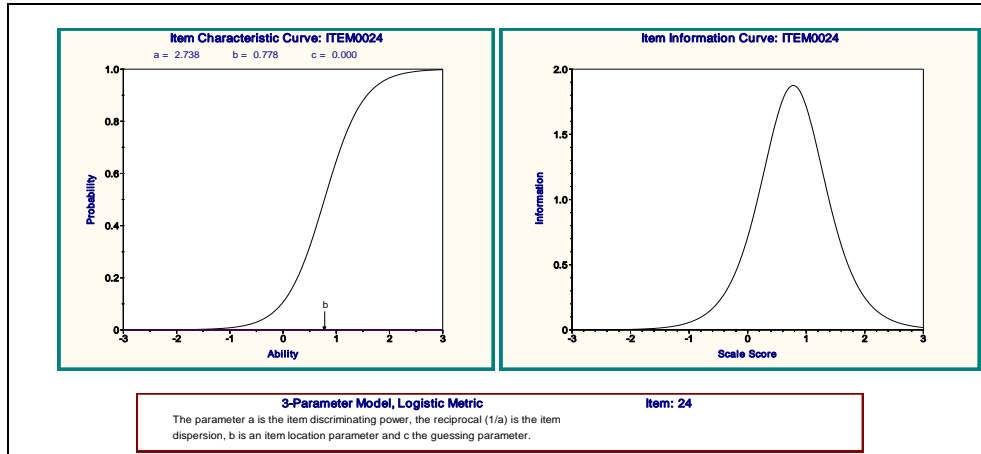
- ۱ (۱) ۲ (۲) ۳ (۳) ۴ (۴)

این سؤال تنها به دانش در زمینه مشتق نیاز دارد

جدول ۶. سطوح پیچیدگی سؤال ۲۴

سطح	تأثیر	تأثیر	تأثیر	تأثیر
۱	چندگام یا عملیات طی می‌شود	دستورالعملی که در سؤال ارائه شده برای انجام عملیات حل سؤال مستقیم است	تا چه حد توضیحات یا متغیرها انتزاعی هستند	با چه دقتی روش انجام کار از طریق متغیرها هدایت شده است
۲	دربرگیرنده تنها یک گام یا عملیات است	تکلیف یا عمل خاصی که انجام می‌گیرد به‌طور مستقیم بیان شده	محسوس	تنها یک راه خاص برای روش یا فرایند حل سؤال وجود دارد
۳	گام‌های چندگانه‌ای به‌طور متوالی برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌گیرد نیازمند مقداری استنباط است	ممکن است انتزاعی باشد	مقداری تغییر و یا چندگانگی در انجام فرایند و یا روش حل سؤال وجود دارد
۴	گام‌های چندگانه‌ای به‌طور متوالی یا هم‌زمان برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌شود باید از دستورالعمل استنباط شود	انتزاعی	تغییر و یا چندگانگی آشکاری در روش‌های انجام فرایند و یا روش حل سؤال وجود دارد

سؤال ۲۴ به‌عنوان بهترین سؤال مجموعه ۳۰ سؤالی شناخته می‌شود و تنها به یک الگوریتم محاسباتی یعنی دانش در زمینه مشتق نیاز دارد. تابع آگاهی آن نزدیک به ۲ و ضریب دشواری آن $b=0/78$ بوده و شیب آن در جدا کردن افراد از یکدیگر $a=2/74$ است که برای یک سؤال چهارگزینه‌ای ایده آل است.



نمودار ۳. تابع ویژگی سؤال (چپ) و تابع آگاهی سؤال (راست)

در رابطه با پاسخ‌گویی به سؤال این پژوهش، تحلیل ۱۲ سؤال نشان می‌دهد که عملاً همه سؤالات این آزمون آزمایشی حساب دیفرانسیل مبتنی بر دانش روندی هستند. درصد مربوط به سطوح پیچیدگی آن‌ها را می‌توان در جدول ۷ دوبعدی دید.

جدول ۷. درصد سطوح پیچیدگی ۱۲ سؤال مبتنی بر دانش روندی

سطح پیچیدگی	چندگام یا	تا چه حد دستورالعملی که	تا چه حد	با چه دقتی روش	زمینه یا متنی که	چقدر شباهت با
	عملیات طی می‌شود	در سؤال ارائه شده برای انجام عملیات حل سؤال مستقیم است	توضیحات یا متغیرها انتزاعی هستند	انجام کار از طریق متغیرها هدایت شده است	فرد در آن روش یا فرایند حل مسئله را یاد گرفته وجود دارد	
-	دربرگیرنده تنها یک عملیات است	تکلیف یا عمل خاصی که انجام می‌گیرد به‌طور مستقیم بیان شده (۱۷٪)	محسوس	تنها یک راه خاص برای روش یا فرایند حل سؤال وجود دارد	احتمالاً متن شبیه به آن چیزی است که در آموزش به کار رفته (۳۳٪)	

گام‌های چندگانه‌ای به‌طور متوالی برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌گیرد نیازمند مقداری استنباط است	ممکن است انتزاعی باشد (۰/۰۸٪)	مقداری تغییر و یا چندگانگی در انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن شبیه یا معادل با آن چیزی است که در آموزش به کار رفته (٪۶۷)
۲ (٪۴۲)	(٪۵۸)		(٪۵۸)	
گام‌های چندگانه‌ای به‌طور متوالی یا هم‌زمان برداشته می‌شود	تکلیف یا عمل خاصی که انجام می‌شود باید از دستورالعمل استنباط شود	انتزاعی (٪۹۲)	تغییر و یا چندگانگی آشکاری در روش‌های انجام فرایند و یا روش حل سؤال وجود دارد	احتمالاً متن به‌طور کامل متفاوت با آن چیزی است که در آموزش به کار رفته
۳ (٪۵۸)	(٪۲۵)			

بحث و نتیجه‌گیری

طبقه‌بندی بنجامین اس. بلوم (۱۹۵۶) و الگوهای برخاسته از آن مثل ژرفای دانش وب (۲۰۰۲) نقش زیادی در استانداردسازی اهداف آموزشی به‌ویژه طبقه‌بندی سؤالات آزمون‌ها داشته است. در این راستا بلوم رویکرد رده‌بندی ارائه داد که به‌وسیله آن می‌توان تحلیلی از اهداف درسی و یا دوره آموزشی به دست آورد که نشان می‌دهد تا چه اندازه ابعاد چهارگانه دانش و فرایندهای شناختی شش‌گانه در یک دوره آموزشی وجود داشته است (هس، ۲۰۰۵). رویکرد بلوم در ابتدا برای مشخص کردن سطح دشواری سؤالات آزمون‌ها به کار می‌رفت^۱ (کینگ و رازوک، ۲۰۱۴). در ادامه نورمان آل وب (۲۰۰۲، ۲۰۰۵) رویکرد چهار سطحی ژرفای دانش را ارائه داد که می‌توان به‌وسیله آن الزامات شناختی که در «استانداردهای یک واحد درسی» عنوان شده را مورد ارزیابی قرار داد. این رویکرد به‌عنوان یک روش هم‌ترازی^۲ برای بررسی هماهنگی بین الزامات شناختی استانداردهای آموزشی و الزامات شناختی برنامه‌های سنجشی^۳ توسعه یافته است^۴ (کینگ و رازوک، ۲۰۱۴). تفاوتی که رویکرد بلوم با رویکرد وب ایجاد می‌کند این است که در رویکرد بلوم فعل سؤال به‌عنوان مثال تحلیل کنید، توضیح دهید و یا توصیف کنید نقش اساسی در دشواری سؤال ایجاد می‌کند. درحالی‌که در رویکرد

1. identifying FCAT item difficulty
2. alignment method
3. cognitive demands of assessments
4. identifying FCAT item difficulty

وب تفکر سطح بالا مدنظر است و اشاره دارد به این که چند گام لازم است تا به یک سؤال پاسخ درست داد. در الگوی وب ژرفای دانش یادگیرنده بر اساس فعل سؤال - که در رده‌بندی بلوم نقش اساسی دارد - مشخص نمی‌شود بلکه متنی که در آن فعل به کار رفته و عمق تفکری که در پاسخ‌گویی به سؤال لازم است از اهمیت برخوردار است. به‌رحال «دشواری» اشاره دارد به اینکه چند دانش‌آموز از کل دانش‌آموزان به یک سؤال به‌درستی جواب داده‌اند. درحالی‌که «تفکر سطح بالا» اشاره دارد به الزام‌های شناختی لازم برای پاسخ‌گویی درست به یک سؤال و این که چند گام لازم است تا به سؤال پاسخ داد (هس، کارلوک، جونز و واک آپ، ۲۰۰۹؛ ماتیوس، ۲۰۱۰).

چهارچوبی که با ترکیب این دو روش به وجود آمده تحت عنوان الگوی توان پیچیدگی بدون توجه به این که در کل دانش‌آموزان چند نفر به یک سؤال پاسخ داده، سؤالات آزمون را بر اساس ظرفیت یا توان شناختی از جمله دانش بیانی، دانش روندی و حل مسئله و سه سطح پیچیدگی شناختی به همراه جزئیاتی که در جداول آمده طبقه‌بندی می‌کند. برای استفاده از مدل توان-پیچیدگی ابتدا نوع توان یا ظرفیت شناختی مشخص می‌شود و سپس معیارهای مربوط به رده‌بندی سطوح پیچیدگی به کار می‌رود. مزایای این الگو آن است که برای طبقه‌بندی مهارت‌های فکری فراخوانی شده در سؤالات به هر دو مدل شناختی درگیر در فرایند حل تکلیف، یعنی توانایی یا ظرفیت شناختی^۱ درگیر در حل مسئله که به‌وسیله بلوم (۱۹۵۲) تحت عنوان کلی دانش و در کارهای گاگنی، یوکوویچ (۱۹۹۳) تحت عناوین دانش بیانی، دانش روندی و حل مسئله بیان شده‌اند و هم پیچیدگی مهارت مورداندازه‌گیری که در طبقه‌بندی وب تحت عنوان ژرفای دانش به سه سطح پیچیدگی کم، پیچیدگی متوسط و پیچیدگی بالا تفکیک شده می‌پردازد. این مدل‌ها مبنای استانداردهای آموزشی ایالتی^۲ و همچنین برنامه‌های آزمون‌گیری ایالتی در آمریکا است و در بسیاری از آزمون‌های جامع فارغ‌التحصیلی دبیرستان مثل آزمون سنجش جامع فلوریدا و توسط افرادی همچون کراتول (۲۰۰۲) وب و همکاران (۲۰۰۲، ۲۰۰۵ و ۲۰۰۶) اوسترهوف و همکاران (۲۰۰۸)، با اندکی تغییرات مورد استفاده قرار گرفته‌اند.

1. cognitive ability
2. sunshine state standards (SSS)

بررسی هم‌زمان توانایی‌ها و سطوح پیچیدگی کمک می‌کند تا آموزگاران متوجه شوند که چرا دانش‌آموزان در بعضی از محتوای آموزشی مشکلات یادگیری دارند و هنگامی که حتی آمادگی کافی برای آزمون داشته‌اند بر روی بعضی از سؤالات عملکرد ضعیفی از خود نشان می‌دهند. یکی از عمده‌ترین دلایل عملکرد ضعیف در چنین آزمون‌هایی که عمدتاً سؤالات آن مبتنی بر دانش روندی و بسیار انتزاعی هستند، گام‌های هم‌زمان و متوالی همراه با ترکیب اطلاعات حاصل از چند روش یا الگوریتم ریاضی است. به‌عنوان مثال سؤال ۲۲ نیازمند دانش در زمینه حد، تابع و مشتق است و استفاده هم‌زمان از این سه موضوع برای دستیابی به پاسخ صحیح سؤال است. سؤالی که مطرح می‌شود این است که آیا لزومی دارد برای تهیه یک سؤال از ترکیب چنین سخت و غیرمنعطف از موضوعات ریاضی استفاده شود. آیا در عالم واقعی می‌توان از چنین سؤالی برای پیدا کردن یک راه‌حل منطقی در موضوعات علوم و تکنولوژی متناسب با قرن ۲۱ استفاده کرد؟ و اگر احتمالاً جواب چنین سؤالی بله باشد آیا برای یک آزمون ورودی رشته‌های دانشگاهی مختلف این نوع سؤالات برای پذیرش در هر رشته دانشگاهی روایی دارد و یا اثر بخش هستند؟ طبیعتاً خیر! تحلیل بیشتر چنین سؤالاتی نشان می‌دهد که عمدتاً آزمون‌شوندگان برای پاسخ‌گویی بیشتر به حدس و گمان متکی می‌شوند تا به کارگیری یک روش خوب آموزش داده شده. همچنین تحلیل‌های آماری سؤالات نشان داده‌اند که هرچه ویژگی‌های این سؤالات به ردیف اول جدول دو بعدی رده‌بندی بالا می‌رود ضریب مقیاس پذیری H^1 آن‌ها که نشانگر قدرت تمیز یک سؤال در مجموعه سؤالات است بهتر می‌شود و الگوی خطای گاتمن آن‌ها (که در آن افراد ضعیف به سؤال سخت درست پاسخ می‌دهند به علت حدس شانسی و به سؤال آسان پاسخ نادرست) کمتر می‌گردد، دارای شیب a بهتری برای جدا کردن افراد روی پیوستار توانایی می‌شوند و دانش‌آموزان را برای پیدا کردن یک پاسخ درست بیشتر به تفکر و انجام عملیات منجر به پاسخ درست تشویق می‌کنند تا حدس.

c.

از طرف دیگر، تهیه چنین سؤالاتی تنها بر اساس دانش روندی، ممکن است ناشی از ماهیت تکالیف ریاضی نیز باشد که تهیه‌کنندگان آن به ناچار سؤالاتی در رده دانش

روندی تهیه می‌کنند که عمدتاً مبتنی است بر «به کار بستن یک اصل یا قاعده ریاضی در یک موقعیت جدید اما مشابه آنچه قبلاً آموخته شده». از نظر بلوم (۱۹۷۱) چنین سؤالاتی که تنها تغییرات بسیار کوچکی در آن صورت گرفته ولی فرایند همان است که در کلاس تمرین شده است ربطی به سطوح بالای تفکر ندارند. سؤالات آزمونی که در کلاس تمرین شده را می‌توان تحت عنوان فهمیدن یا احتمالاً دانش مطرح کرد. تفاوت‌های عمده‌ای که بین تفکر سطح بالا و تفکر سطح پایین در ادبیات پژوهش وجود دارد بیشتر برمی‌گردد به آشنایی دانش‌آموز به تکلیفی که باید انجام دهد و این که آیا دانش‌آموز هم‌اکنون یک الگوریتم یا راهبرد حل مسئله‌ای می‌داند که وقتی آن را به‌طور مناسب به کار بُرد منجر به یک راه‌حل درست شود؟ معمولاً سؤالاتی که برای آزمون‌های ورودی و یا پذیرش تهیه می‌شوند برای جبران کاستی پیچیدگی شناختی از ترکیب چندین الگوریتم ریاضی استفاده می‌کنند و سؤال را دشوار و چالش برانگیز می‌کنند به جای پیچیده کردن آن از لحاظ شناختی و انعطاف بیشتر در راه‌های پاسخ‌گویی (بلوم، هستینگ و ماداوس^۱، ۱۹۷۱).

پیشنهاد می‌شود سازمان‌ها و یا مؤسساتی که سؤالات آزمون‌های ویژه را تهیه می‌کنند چهارچوب‌های استاندارد یادگیری داشته باشند که در آن سطوح پیچیدگی و دشواری سؤالات به‌تناسب برای تهیه آزمون رعایت شده باشد. همچنین تهیه‌کنندگان سؤالات باید آموزش ببینند تا بتوانند انواع مختلفی از سؤالات آسان، متوسط و دشوار را برای گنجینه سؤالات آزمون موردنظر تهیه کنند و هم‌زمان سطوح مهارت‌های فکری لازم برای پاسخ‌گویی به هر سؤال مشخص شود. در نهایت سؤالات آن‌ها توسط متخصصان محتوایی و کارشناسان موضوعی مورد ارزیابی قرار گیرد و در رابطه با کیفیت سؤالات بر اساس استانداردهای یادگیری مشخص شده به تهیه‌کنندگان سؤالات بازخورد داده شود (بازمور، کرامر و گالاگر، انگلهارت و براون^۲، ۲۰۰۸).

به‌هرحال رده‌بندی وسیله‌ای است برای طبقه‌بندی اهداف، موضوعات و اخیراً استانداردهای آموزشی. این کار یک ساختار سازمان‌یافته ایجاد می‌کند که معنای ادراکی عامی را از هر هدف طبقه‌بندی‌شده در یکی از طبقات به دست می‌دهد و از این طریق

-
1. Hastings & Madaus
 2. Bazemore, Kramer, Gallagher, Englehart & Brown

باعث افزایش روابط بین دست‌اندرکاران آموزشی می‌شود. استفاده از این جداول برای طبقه‌بندی اهداف، فعالیت‌ها و اندازه‌گیری‌ها یک بازنمایی روشن، مختصر و دیداری از یک دوره و یا واحد درسی خاص ارائه می‌دهد. وقتی این جداول کامل شدند، چیزهای واردشده در این جداول را می‌توان برای بررسی اهمیت نسبی چیدمان برنامه درسی و فرصت‌های آموزشی فراموش‌شده مورد استفاده قرار داد. بر اساس این بررسی معلمان می‌توانند تصمیم بگیرند که کجا و چگونه طراحی برنامه درسی، اجرای آموزش و شیوه بیان مطالب را بهبود بخشند (کراتول، ۲۰۰۲). بعلاوه کمک می‌کند تا با توجه به نوع یادگیری و تفکر دانش‌آموزان و به علت تفکیک انواع موضوعات آموزشی در جدول رده‌بندی، گونه‌های متعددی از روش‌های ارزشیابی و سنجش هم از عملکرد دانش‌آموزان و هم از اهداف و موضوعات آموزشی صورت بگیرد (ایراسیان و میراندا، ۲۰۰۲).

به‌عنوان پیشنهادی دیگر، آموزگاران که مایل‌اند آموزش‌های خود را ارتقا ببخشند می‌توانند از جدول مشخصات برای بازبینی برنامه‌های آموزشی خود استفاده کنند تا اطمینان یابند که چیدمان اهداف، فعالیت‌ها و سنجش و ارزیابی به‌خوبی و به‌تناسب ایجاد شده است. به‌ویژه اهداف یادگیری آموزش خود را به فرایندهای شناختی پیچیده‌تر مثل تحلیل، ارزشیابی و آفرینش برسانند تا صرفاً حفظ مطالب و موضوعات درسی (راث، ۲۰۰۲). راه دیگری که آن‌ها می‌توانند اهداف یادگیری دانش‌آموزان را ارتقا بخشند حرکت به فراسوی سه گونه دانش (امور حقیقی، مفهومی و روندی) است و اهدافی را مدنظر قرار دهند که بر دانش فراشناختی تأکید بیشتری دارند. این امر به دانش‌آموزان کمک می‌کند تا درک کنند چگونه آموزش و یادگیری دانشگاهی (و یا مدرسه‌ای) با آن‌ها ارتباط می‌یابد و برعکس ارتباط آن‌ها با این یادگیری‌های دانشگاهی چیست. از طریق دانش فراشناختی است که آن‌ها دانش استراتژیک به دست می‌آورند و می‌توانند آن‌ها را در یادگیری علوم، ریاضی و نظایر آن به کار برند (راث، ۲۰۰۲).

یکی از محدودیت‌های این پژوهش آن بود که فرایند دومرحله‌ای رده‌بندی بر مبنای الگوی توان-پیچیدگی و یا الگوهای رده‌بندی اهداف آموزشی گاه منجر به عدم

1. Airasian & Miranda

2. Rath

توافق در بین کارشناسان موضوعی و یا آموزگاران و یا تهیه‌کنندگان سؤالات برای رده‌بندی یک سؤال در یک طبقه می‌شود. با آموزش‌های بیشتر کارشناسان موضوعی و کسب تجربه آن‌ها می‌تواند روی گونه‌ای از توانایی که سؤالات آزمون اندازه می‌گیرند، دشواری و همین‌طور سطوح پیچیدگی سؤالات سریع‌تر به توافق برسند. اگر به نمودار و جدول سطوح پیچیدگی سؤال ۲۴ نگاه کنید می‌توان دید که هرچقدر سؤالات از ترکیب ساده‌تری از الگوریتم‌های محاسباتی استفاده کنند بهتر می‌توانند توانایی زیربنایی آزمون‌شوندگان را اندازه بگیرند. از طرف دیگر تهیه‌کنندگان آزمون و یا کارشناسان موضوعی راحت‌تر می‌توانند سؤالات را رده‌بندی کنند. وب، هرمان^۱ و وب (۲۰۰۶) بیان می‌کنند که عدم توافق بین کسانی که سؤالات را رده‌بندی می‌کنند می‌تواند دل‌نگرانی مهمی در بررسی‌های مربوط به تنظیم آزمون باشد. اگر ناهماهنگی مسئله‌ساز شود این امر به هماهنگی و هم‌ترازی بین استانداردهای آموزشی و سؤالات آزمون که برای ارزیابی پیشرفت دانش‌آموزان در دستیابی به آن استانداردها طراحی شده است آسیب می‌رساند (وب و هرمان، ۲۰۰۶).

منابع

- سیف، علی‌اکبر. (۱۳۸۵). روش‌های اندازه‌گیری و ارزشیابی آموزشی. تهران: دوران.
- Airasian, P. W., & Miranda, H. (2002). The role of assessment in the revised taxonomy. *Theory into practice*, 41(4), 249-254.
- Anderson, L. W. (Ed.), Krathwohl, D. R. (Ed.), Airasian, P.W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M.C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational objectives* (complete edition). New York: longman
- Bazemore, M., Kramer, M. B., Gallagher, M. P., Englehart, T. P. & Brown, R. A. (2008). The North Carolina Mathematics Tests. Associate State Superintendent Office of Innovation and School Transformation. 6301 Mail Service Center. Raleigh, NC 27699-6301
- Bloom, B. S., Hastings, J. T. & Madaus, G. F. (1971). *Handbook on the Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill.
- Bloom, B. S. (1956). *Taxonomy of educational objectives*, handbook 1: Cognitive domain. New York: McKay.
- Confrey, J. (1990). *A review of the research on student conceptions in mathematics, science, and programming*. In C. B. Cazden (Ed.), *Review of research in education: Vol. 16* (pp. 3-56). Washington, DC: American Educational Research Association.

- Florida Department of Education. (2009). Draft fcatt Mathematics Test Item Specifications Grades 3–5. The Administrator Office of Assessment Florida Department of Education Tallahassee, Florida 32399-0400
- Florida Department of Education. (FCAT). (2008). Cognitive Complexity Classification of Fcatt Test Items. Florida Comprehensive Assessment Test (FCAT).
- Florida Department of Education. (2012). Cognitive Complexity Classification of the 2012-13 Statewide Assessment Test Items. Office of Assessment available at: <http://www.fldoe.org/core/fileparse.php/3?url=cognitivecomplexity.pdf>
- Gagné, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). *The cognitive psychology of school learning*. New York: Harper Collins.
- GED Testing Service. (2012). Webb's Depth of Knowledge: Transiting to the 2014 GED Test. American Council on Education (ACE)
- Hess, K. (2005). Introduction to Depth of Knowledge (DOK) - Based on Norman Webb's Model. Kentucky Department of Education. Center for Assessment/NCIEA, 2005.
- Hess, K., Jones, B. S., Carlock, D., & Walkup, J. R. (2009). Cognitive rigor: Blending the strengths of Bloom's Taxonomy and Webb's Depth of Knowledge to enhance classroom-level processes. ERIC: ED517804 NYC Department of Education. Introduction to Depth of Knowledge [Video].
- Kentucky Department of Education (2007). *Support materials for core content for assessment version 4.1: Mathematics*. Retrieved from: <http://education.ky.gov/curriculum/docs/pages/content-specific-core-content-for-assessment-dok-support-materials.aspx>
- King I., & Razzouk R. (2014). Content Complexity-Florida Standards: Definitions. Florida State University
- Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. Theory into Practice 41 no4 2002. <http://coe.ohio-state.edu/>.
- Linn, R. (1989). Has Item Response Theory Increase the Validity of Achievement test scores? University of Colorado, UCLA Center for Research on Evaluation, Standards, and Student Testing. CSE Technical Report 302
- Marzano, R. J. & Kendall, J. S. (2006). *The New Taxonomy of Educational Objectives*. Publisher: Corwin; 2nd edition (December 20, 2006). ISBN-10: 1412936292
- Matthews, B. (2010). Developing Higher Order Thinking Questions Based on Webb's DOK and FCAT Content Complexity Presented by Brenda Matthews, District Literacy Coach November 22 & 23, 2010
- Oosterhof, A., Rohani, F., Sanfilippo, C., Stillwell, P. & Hawkins, K. (2008). The Capabilities-Complexity Model. Center for Advancement of Learning and Assessment. Florida State University, Tallahassee, FL. www.cala.fsu.edu
- Raths, J. (2002). Improving instruction. *Theory into Practice*, 41(4), 233-237. College of education, The Ohio State University
- Rothman, R. (2002). Statewide Testing Programs, In: *Encyclopedia of education* (Vol. 8) edited by James W Guthrie, New York: Macmillan Reference USA.
- Sforza, D. (2014). The Inclusion of Cognitive Complexity: A Content Analysis of New Jersey's Current and Past Intended Curriculum. Seton Hall University Dissertations and Theses (ETDs). Paper 2020.
- Thompson, T. (2011). An Analysis of Higher-Order Thinking on Algebra I End-of Course Tests. *International Journal for Mathematics Teaching & Learning*.
- Webb, N. L. (2005). Web Alignment Tool. 24 July 2005. Wisconsin Center of Educational Research. University of Wisconsin-Madison. <<http://www.wcer.wisc.edu/WAT/index.aspx>>.

- Webb, N. L. (2002). Depth-of-knowledge levels for four content areas. *Language Arts*. March 28.
- Webb, N. M., Herman, J. L., & Webb, N. L. (2007). Alignment of mathematics state-level standards and assessments: The role of reviewer agreement. *Educational Measurement: Issues and Practice*, 26(2), 17-29.
- Zimowski, M., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). BILOG-MG 3: Item analysis and test scoring with binary logistic models. *Chicago, IL: Scientific Software*.