ارزشیابی بار شناختی و روش‌های نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی

■ صادق صیادی* ■ جواد مصرآبادی** ■ ابوالفضل فرید*** ■ رامین حبیبی‌کلیبر****

چکیده:

استفاده از آزمون‌های مبتنی بر نقشهٔ مفهومی به‌عنوان ابزارهای ارزشیابی، یکی از روش‌های نوین ارزشیابی است که می‌تواند جایگزین روش‌های سنتی ارزشیابی شود. لذا هدف پژوهش حاضر، ارزشیابی بار شناختی و روش‌های نمره‌دهی این آزمون‌ها بود. آزمودنی‌های پژوهش حاضر شامل دانش‌آموزان چهار کلاس درس زیست‌شناسی از چهار دبیرستان متفاوت بودند (در مجموع ۱۰۹ نفر) که به روش نمونه‌گیری خوشه‌ای چندمرحله‌ای انتخاب شدند. ابزارهای پژوهش شامل یک آزمون چهارگزینه‌ای (محقق‌ساخته)، چهار آزمون مبتنی بر نقشهٔ مفهومی (محقق‌ساخته) و مقیاس بار شناختی «NASA-TLXL» (هارت و استیولند، ۱۹۸۸) بود. طرح پژوهشی مورد استفاده نوعی طرح ارزشیابی بود. روش‌های آماری شامل ضریب همبستگی پیرسون، آزمون t مستقل و تحلیل واریانس یک‌راهه (آزمون ولچ و آزمون تعقیبی گیمز-هاول) بود. بر اساس نتایج، آزمون‌های ترسیم نقشهٔ مفهومی نسبت به آزمون‌های تکمیل نقشهٔ مفهومی، به‌طور معناداری بار شناختی بیشتری روی آزمودنی‌ها تحمیل می‌کردند. همچنین، سه روش نمره‌دهی رابطه‌ای، ساختاری و شباهت به نقشهٔ متخصص، شاخص‌های روایی ملاکی، روایی تمییزی و پایایی قابل قبول و معناداری داشتند.

در مجموع، روش‌های نمره‌دهی رابطه‌ای، ساختاری و شباهت به نقشهٔ متخصص می‌توانند برای نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی به‌خوبی مورد استفاده قرار گیرند. همچنین، استفاده از آزمون‌های تکمیل نقشهٔ مفهومی می‌تواند راهکار مناسبی برای غلبه بر اضافه‌بار شناختی در جریان ارزشیابی‌های مبتنی بر نقشهٔ مفهومی باشد. بنابراین، نتایج پژوهش حاضر می‌تواند به طراحی آزمون‌های مبتنی بر قواعد شناختی کمک کند و پلی بین فرایند ارزشیابی آموزشی و علوم شناختی برقرار کند.

کلید واژه‌ها: **بار شناختی، روش‌های نمره‌دهی، آزمون‌های مبتنی بر نقشهٔ مفهومی**

■ تاریخ دریافت مقاله: ۹۸/۱۲/۹ ■ تاریخ شروع بررسی: ۹۹/۲/۲۰ ■ تاریخ پذیرش مقاله: ۹۹/۶/۲۳

* دانشجوی دکترای روان‌شناسی تربیتی، دانشگاه شهید مدنی آذربایجان. تبریز..... Sadeghsayadi90@gmail.com

** استاد رشتهٔ روان‌شناسی تربیتی، هیئت علمی گروه علوم‌تربیتی، دانشگاه شهید مدنی آذربایجان،تبریز، **(نویسندهٔ مسئول)** mesrabadi@azaruniv.ac.ir

*** استادیار رشتهٔ روان‌شناسی تربیتی، هیئت علمی گروه علوم تربیتی، دانشگاه شهید مدنی آذربایجان، تبریز..... abolfazlfarid@gmail.com

**** دانشیار رشتهٔ روان‌شناسی تربیتی، هیئت علمی گروه علوم تربیتی، دانشگاه شهید مدنی آذربایجان، تبریز..... habibikaleibar@gmail.com

■ *این مقاله از پایان‌نامهٔ دکترا با عنوان «بررسی شاخص‌های روایی، پایایی، بارشناختی و نظام‌های نمره‌گذاری آزمون‌های مبتنی بر نقشهٔ مفهومی» استخراج‌شده است که در اسفند ماه سال ۱۳۹۷ با راهنمایی دکتر جواد مصرآبادی در دانشگاه شهید مدنی آذربایجان به اتمام رسید.*

■ مقدمه

امروزه به‌طور معمول از دوره‌های پایین تحصیلی تا دوره‌های بالا برای ارزشیابی پیامدهای یادگیری از آزمون‌های سنتی (مانند آزمون‌های چندگزینه‌ای، صحیح و غلط، و پرکردن جای خالی) استفاده می‌شود. بیشتر این روش‌های ارزشیابی روی یادآوری ایده‌ها و مفاهیم پراکنده و مجزا تمرکز می‌کنند (شوندیمن[۱]، ۲۰۱۵). اگرچه آزمون‌های سنتی جنبه‌های مثبت دارند، اما معمولاً کارایی لازم را در زمینهٔ ارزشیابی ندارند. برای مثال، آزمون‌های چندگزینه‌ای[۲] معمولاً روابط پیچیدهٔ بین مفاهیم را نشان نمی‌دهند. علاوه بر این، این آزمون‌ها پاسخ‌های شاگردان را بسیار محدود می‌کنند و این موضوع، پی‌بردن به دلایل زیربنایی پاسخ‌ها را مشکل می‌کند. در نهایت، آزمون‌های چندگزینه‌ای می‌توانند مفاهیم مهمی را که طراحان آزمون قصد ارزشیابی آن‌ها را دارند، ارزشیابی کنند، اما برای ردیابی کج‌فهمی‌هایی[۳] که هنوز کشف نشده‌اند، ابزارهای مفیدی نیستند. بنابراین، این آزمون‌ها در تشخیص مؤثر اطلاعات ارزشمند دربارهٔ جنبه‌های اساسی ساختار دانش فراگیرندگان شکست می‌خورند (لئو، لین و تسای[۴]، ۲۰۰۹).

در همین راستا، پژوهشگران آموزشی، به‌منظور کشف شکل‌های مؤثر ارزشیابی، به دلیل نارضایتی از روش‌های سنتی ارزشیابی، تلاش‌هایی انجام داده‌اند. یکی از پیشگامان تغییر این وضعیت، نظریهٔ سازنده‌گرایی[۵] است (آرنسون[۶]، ۲۰۰۵). این نظریه پیشنهاد می‌کند که فهم اطلاعات مستلزم این است که اطلاعات تازه تا حدی به دانش موجود پیوند داده شوند. میزان ارتباط بین اطلاعات قبلاً آموخته‌شده و اطلاعات تازه، میزان درک و فهم را منعکس می‌کند (کاپوزا[۷]، ۲۰۲۰).

یکی از روش‌های جذاب ارزشیابی و هم‌خوان با نظریهٔ سازنده‌گرایی، روش «ارزشیابی مبتنی بر نقشهٔ مفهومی[۸]» است که می‌تواند جنبه‌هایی از یادگیری را اندازه بگیردکه آزمون‌های سنتی نمی‌توانند (استودارت، آبرامس، گاسپر و کانادای[۹]، ۲۰۰۰). در واقع، توانایی توصیف ساخت‌شناختی یادگیرنده از طریق ترسیم نقشهٔ مفهومی، به استفاده از نقشه‌های مفهومی به‌عنوان ابزارهای ارزشیابی منجر شده است (آرنسون، ۲۰۰۵). علاوه بر این، ادغام مفاهیم حیطه‌های گوناگون، یکی از ویژگی‌های اصلی نقشه‌های مفهومی است که نشان می‌دهد این شکل از ارزشیابی می‌تواند در زمینهٔ درک کلی از موضوعات بسیار کمک‌کننده باشد، چرا که در این روش از آزمودنی‌ها خواسته می‌شود رابطهٔ بین مفاهیم و بافت (زمینه) گسترده‌ترشان را درک کنند. بنابراین، آزمودنی‌ها به درک معنادارتر موضوعات تشویق می‌شوند (هو، هریس، کومار و ولان[۱۰]، ۲۰۱۸). اما با وجود مزایای یادشده، آزمون‌های مبتنی بر نقشهٔ مفهومی نیز باید همچون سایر روش‌های ارزشیابی، روش یا روش‌های معینی برای نمره‌دهی داشته‌باشند.

روئیز-پریمو، شولتز و شیولسون[۱۱] (۱۹۹۶) اظهار داشتند، روش نمره‌دهی[۱۲] یک روش منظم و نظام‌دار است که با نقشه‌های مفهومی آن، شاگردان می‌توانند به‌طور دقیق و منسجم ارزشیابی شوند (اسلسون[۱۳]، ۲۰۱۰). روش نمره‌دهی نقشه‌های مفهومی معمولاً ترکیبی از چند معیار را در برمی‌گیرد، چرا که معیارهای متفاوت، جنبه‌های متفاوت را از دانشی که با نقشهٔ مفهومی نشان داده می‌شود، اندازه

می‌گیرند (مثلاً تعداد گزاره‌ها حجم دانش را و تعداد سطوح سلسله‌مراتب عمق دانش را نشان می‌دهد.) (استراوتمین[۱۴]، ۲۰۱۲).

بعضی نویسندگان نقشه‌های مفهومی را با توجه به چند بعد نمره‌دهی می‌کنند؛ مانند جامع بودن، تعداد اجزا و جزئیات و پیچیدگی ارتباطات (ودووزوف و راود[۱۵]، ۲۰۱۵). در همین باره، در ادبیات ارزشیابی‌های مبتنی بر نقشهٔ مفهومی، از روش‌های نمره‌دهی متعددی استفاده شده‌است که از مهم‌ترین آن‌ها می‌توان به روش نمره‌دهی ساختاری[۱۶]، روش نمره‌دهی رابطه‌ای[۱۷] و روش شباهت به نقشهٔ متخصص[۱۸] اشاره کرد.

در روش نمره‌دهی ساختاری که نواک و گوییـن[۱۹] (۱۹۸۴) ابداع کرده‌اند، دانشی که به‌صورت سلسله‌مراتبی سازمان‌دهی شده است، مورد توجه قراردارد. در این روش، به روابطی که ارتباطات سلسله‌مراتبی را نشان می‌دهند، نمرات بالاتری داده می‌شود. بالاترین امتیاز در این روش نمره‌دهی به روابطی داده می‌شود که مفاهیم بین شاخه‌های گوناگون یک ساختار سلسله‌مراتبی را به هم پیوند می‌دهند.

در روش نمره‌دهی رابطه‌ای که مک‌کلور و بل[۲۰] (۱۹۹۰) معرفی کردند، به هر یک از روابط بین مفاهیم (گزاره) به‌صورت جداگانه نمره داده‌می‌شود و بیشترین امتیاز به روابطی تعلق می‌گیرد که به‌درستی پیوند خورده‌اند و روابط بنیادی یک حیطهٔ دانش را بیان‌کرده‌اند.

در روش نمره‌دهی شبیه به نقشهٔ متخصص نیز که مک‌کلور، سوناک و سوئن[۲۱] (۱۹۹۹) ابداع کرده‌اند، نقشه‌های فراگیرندگان در مقایسه با نقشهٔ مفهومی معیار (نقشهٔ معلم یا متخصص) ارزشیابی می‌شود (ماکوپادیای، سار، گوسل و آچاریا[۲۲]، ۲۰۱۸). با گسترش استفاده از نقشه‌های مفهومی به‌عنوان ابزارهای اندازه‌گیری دانش فراگیرندگان، نمره‌دهی آن‌ها به مبحثی چالش‌انگیز تبدیل شده‌است.

بخشی از مشکل موجود در زمینهٔ نمره‌دهی، ناشی از این تنوع روش‌هاست؛ به‌طوری که شیولسون و همکاران (۲۰۰۵) بیش از ۱۲۸ روش ممکن را برای ایجاد و نمره‌دهی نقشه‌های مفهومی تشخیص داده‌اند (مصرآبادی، ۱۳۹۰). با توجه به این تنوع بالا، انتخاب مناسب‌ترین روش نمره‌دهی می‌تواند بخشی از مشکلات مربوط به استفاده از نقشه‌های مفهومی برای اهداف ارزشیابی را رفع کند. از طرف دیگر، روش نمره‌دهی انتخاب‌شده نیز باید با توجه به ملاک‌هایی همچون داشتن روایی و پایایی[۲۳] مناسب انتخاب شود. روایی در ارزشیابی‌های مبتنی بر نقشهٔ مفهومی، کیفیت استنباط‌هایی را منعکس می‌کند که می‌توان در ارتباط با دانش یادگیرنده از فعالیت ترسیم نقشهٔ مفهومی به دست آورد. پایایی نیز در ارزشیابی‌های مبتنی بر نقشهٔ مفهومی به درجه‌ای از توافق و همسانی (ثبات) یا نوسان و تغییرپذیری (بی‌ثباتی) در میان نمرات اشاره می‌کند که به عواملی مانند مهارت یادگیرنده در ترسیم نقشهٔ مفهومی و روشی که توسط آن نقشه‌ها نمره‌دهی می‌شوند، مربوط است (کلاین، بروستر و فل[۲۴]، ۲۰۱۰).

همچون روش‌های نمره‌دهی، ساختار نقشه‌های مفهومی (مثلاً نقشه‌های سلسله‌مراتبی، شبکه‌ای و

زنجیری[۲۵]) و میزان هدایت نقشه‌ها (مثلاً تکمیل‌کردنی یا ترسیم‌کردنی) نیز می‌تواند تفاوت‌هایی را در زمینهٔ استفاده از ارزشیابی‌های مبتنی بر نقشهٔ مفهومی رقم بزند. برای مثال آمادیو، تریکوت و مارین[۲۶] (۲۰۱۰) نشان دادند، نقشه‌های با ساختار سلسله‌مراتبی، به خوانندگان متن کمک می‌کنند در حین خواندن، انسجام و یکپارچگی فهمشان را حفظ کنند و احساس گم‌گشتگی مفهومی‌شان کمتر شود. در حالی که نقشه‌های مفهومی با ساختار شبکه‌ای، برعکس، الزامات و خواست‌های بیشتری را تحمیل و درک و فهم مطالب را مختل می‌کنند.

بنابراین، این تفاوت‌ها می‌توانند استفاده از نقشه‌های مفهومی به‌عنوان ابزارهای ارزشیابی را با چالش‌هایی مواجه کنند. در همین راستا، رویز-پریمو و شیولسون (۱۹۹۶) اصرار داشتند که نظریهٔ شناختی منسجم و کارآمدی نیاز است تا پراکندگی در تکالیف نقشهٔ مفهومی برای اهداف ارزشیابی را محدود کند. شاید این موضوع می‌تواند با به‌کاربردن پیشرفت‌های اخیر در نظریهٔ بار شناختی[۲۷] اصلاح شود و چارچوبی برای بهبود ارزشیابی نقشه‌های مفهومی ساخته شود (ثاین[۲۸]، ۲۰۱۲). در واقع، یکی از نظریه‌های شناختی که به‌منظور توضیح اثربخشی نقشه‌های مفهومی وجود دارد، نظریهٔ بار شناختی است (بری[۲۹]، ۲۰۱۱).

نظریهٔ بار شناختی سویلر و همکارانش[۳۰] در اواخر دههٔ ۱۹۸۰ روی محققان و طراحان آموزشی تأثیر زیادی داشته است. اصولاً نظریهٔ بار شناختی ادعا می‌کند، زمانی که ظرفیت حافظهٔ کاری[۳۱] در یک تکلیف یادگیری از حد معینی تجاوز می‌کند، جریان یادگیری مختل می‌شود (دی‌یانگ[۳۲]، ۲۰۱۰). نکتهٔ اصلی نظریهٔ بار شناختی این است که محدودیت‌های ظرفیت شناختی یادگیرنده باید مورد توجه قرار گیرد تا از ایجاد بار شناختی بیرونی در طراحی مواد یادگیری اجتناب شود (اسکالموسکی و ری[۳۳]، ۲۰۲۰).

فرض اصلی نظریهٔ بار شناختی این است که حافظهٔ کاری انسان ظرفیت محدودی دارد، در حالی که ظرفیت حافظهٔ بلندمدت نامحدود است. هدف یادگیری در نظریهٔ بار شناختی، ساختن ساختارهای دانش در حیطه‌ای مشخص (طرح‌واره‌های شناختی) است، اما ساخت طرح‌واره در حافظهٔ کاری صورت می‌گیرد که ظرفیت محدودی دارد. مبنی بر نظریهٔ بار شناختی، این ظرفیت باید بین منابع متفاوت بار تحمیل‌شود و روی حافظهٔ کاری در طول یادگیری یا پرداختن به تکالیف گوناگون تقسیم شود (کلپش و سیوفرت[۳۴]، ۲۰۲۰).

در واقع، نظریهٔ بار شناختی به‌طور گسترده استفاده شده‌است تا بار نظام شناختی انسان را هنگام انجام‌دادن یک تکلیف مشخص توصیف کند. در دهه‌های اخیر، نظریهٔ بار شناختی به‌عنوان چارچوبی برای محققان، به‌منظور بهبود پیامدهای آموزشی، پذیرفته شده است. با توجه به اینکه منابع شناختی سیستم شناختی انسان، مانند دقت یا حافظهٔ کاری محدودند، یک تکلیف یا وظیفهٔ پیچیده نسبت به تکلیف ساده، بار سنگین‌تری را به سیستم شناختی تحمیل می‌کند (لین و لین[۳۵]، ۲۰۱۴).

محققان پیشنهاد کرده‌اند، استفاده از سازمان‌دهنده‌های تصویری مانند نقشهٔ مفهومی، می‌تواند در

مدیریت بار شناختی کمک‌کننده باشد، چرا که آن‌ها مفاهیم را در یک طرح منسجم سازمان‌دهی می‌کنند و فضایی را برای حافظه‌کاری به‌منظور فراگرفتن اطلاعات جدید فراهم می‌کنند (وکایری[۳۶]، ۲۰۰۲). نقشه‌های مفهومی به یادگیرندگان اجازه می‌دهند یک مفهوم را به‌عنوان یک کل در نظر بگیرند، بدون اینکه استفادهٔ زائدی از حافظهٔ کاری، به‌منظور فهم این مفاهیم، داشته باشند.

نقشه‌های مفهومی، همانند سایر سازمان‌دهنده‌های تصویری، در کاهش بعضی از بارهای شناختی در روند یادآوری مفاهیم علمی مؤثرند. با اینکه ویژگی‌های معمول نقشه‌های مفهومی (آرایش، ساده‌سازی مفاهیم و انطباق‌پذیری) در این راه به آن‌ها کمک می‌کند، مسائل چالش‌انگیزی نیز در استفاده از آن‌ها وجود دارد. روئیز-پریمو و شیولسون (۱۹۹۶) عنوان کردند که نقشه‌های مفهومی ممکن است بار شناختی بیرونی[۳۷] (که مانع یادگیری است) را بیشتر از بار شناختی وابسته[۳۸] (که یادگیری را تسهیل می‌کند) افزایش دهند. با توجه به اینکه نقشه‌های مفهومی به طرز فکر ویژهٔ هر شخص بستگی دارند، بنابراین، زمانی که یادگیرنده نقشه‌ای را خلق نمی‌کند (مانند تکالیف تکمیل‌کردن نقشه‌های مفهومی) و دیاگرام (شکل ترسیمی)، ممکن است بارهای غیرضروری را هنگام تلاش برای فهمیدن معنایی که مدنظر خالق نقشهٔ مفهومی (آموزش‌دهنده و متخصص) است، افزایش دهد. از طرف دیگر، نقشه‌های مفهومی با هدایت بالا (مانند نقشه‌های مفهومی کامل‌کردنی) از این جهت که ساختار ساده‌تری دارند، ممکن است بار شناختی کمتری تولید کنند (ثاین، ۲۰۱۲).

بنا بر آنچه گفته شد، استفاده از تکنیک‌های گوناگون نقشهٔ مفهومی برای اهداف ارزشیابی ساختار دانش یادگیرندگان، با چالش‌هایی روبه‌روست. پژوهش حاضر در پاسخ‌گویی به این چالش‌ها در تلاش است آزمون‌های مختلف مبتنی بر نقشهٔ مفهومی را از نظر میزان ایجاد بار شناختی و روش‌های نمره‌دهی، با توجه به دو بعد ساختار (نقشه‌های مفهومی سلسله‌مراتبی در مقابل شبکه‌ای) و میزان هدایت (تکمیل نقشه‌های مفهومی در مقابل ترسیم نقشه‌ها) مورد ارزشیابی و کنکاش قرار دهد. بنابراین، در پژوهش حاضر دو سؤال اساسی وجود دارد:

۱. آیا روش‌های نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی از روایی و پایایی مناسبی برخوردارند؟

۲. آیا بین بار شناختی آزمون‌های گوناگون مبتنی بر نقشهٔ مفهومی تفاوتی وجود دارد؟

■ روش پژوهش

جامعه، نمونهٔ آماری، روش نمونه‌گیری

جامعهٔ آماری[۳۹] پژوهش حاضر را تمامی دانش‌آموزان سال اول رشتهٔ علوم تجربی دورهٔ متوسطهٔ دوم شهرستان اسلام‌آباد غرب در سال تحصیلی ۱۳۹۷-۱۳۹۶ تشکیل دادند. دانش‌آموزان چهار کلاس درس زیست‌شناسی (در مجموع ۱۰۹ نفر) به روش نمونه‌گیری خوشه‌ای چندمرحله‌ای[۴۰]، به‌عنوان آزمودنی‌های پژوهش انتخاب شدند. در روش نمونه‌گیری خوشه‌ای، به علت زیادبودن حجم جامعه

و غیرممکن بودن تهیهٔ فهرست کامل تمام افراد جامعه، از خوشه‌ها به‌عنوان واحد نمونه‌گیری استفاده می‌شود. به این منظور، از بین تمام دبیرستان‌های شهرستان اسلام‌آباد غرب، چهار دبیرستان به‌صورت تصادفی انتخاب شدند. سپس از هر یک از این دبیرستان‌ها یک کلاس اول رشتهٔ علوم تجربی انتخاب شد. هر کدام از این کلاس‌ها به‌صورت تصادفی در یکی از گروه‌های چهارگانهٔ ترسیم نقشهٔ مفهومی سلسله‌مراتبی (۲۹ نفر)، ترسیم نقشهٔ مفهومی شبکه‌ای (۲۵ نفر)، تکمیل نقشهٔ مفهومی سلسله‌مراتبی (۲۷ نفر) و تکمیل نقشهٔ مفهومی شبکه‌ای (۲۸ نفر) قرار گرفتند. این کلاس‌ها از قبل تشکیل شده بودند و چهار معلم متفاوت آن‌ها را آموزش دادند و ارزشیابی کردند.

■ ابزارهای پژوهش

الف) آزمون چهارگزینه‌ای

در این آزمون، اهداف آموزشی محتوای درسی به شکل ۳۳ سؤال چهارگزینه‌ای و در دو سطح دانش و درک (طبقه‌بندی بلوم از حیطهٔ شناختی) سنجیده شدند. این آزمون به‌عنوان مقیاس ملاک مورد استفاده قرار گرفت. برای محاسبهٔ روایی ملاکی ضریب همبستگی بین نمرات این آزمون با نمرات یک آزمون معلم ساخته محاسبه شد که برای آزمون‌های فرعی دانش و درک مقادیر، ضریب روایی ملاکی به ترتیب برابر با ۰/۶۵ و ۰/۷۶ بودند. برای اطمینان از پایایی آزمون‌های فرعی دانش و درک نیز از روش کودر- ریچاردسون ۲۱ استفاده شد. این ضریب برای آزمون‌های فرعی دانش و درک به ترتیب برابر ۰/۷۰ و ۰/۷۵ بود.

ب) آزمون‌های مبتنی بر نقشهٔ مفهومی

در پژوهش حاضر، از چهار آزمون مبتنی بر نقشهٔ مفهومی استفاده شد که شامل آزمون ترسیم نقشهٔ مفهومی سلسله‌مراتبی (ترسیم نقشه‌های مفهومی سلسله‌مراتبی توسط خود دانش‌آموزان با استفاده از مفاهیم داده‌شده)، آزمون ترسیم نقشهٔ مفهومی شبکه‌ای (ساختن نقشه‌های مفهومی شبکه‌ای توسط خود دانش‌آموزان با استفاده از مفاهیم داده‌شده)، آزمون تکمیل نقشهٔ مفهومی سلسله‌مراتبی (کامل‌کردن نقشه‌های مفهومی سلسله‌مراتبی) و آزمون تکمیل نقشهٔ مفهومی شبکه‌ای (کامل‌کردن نقشه‌های مفهومی شبکه‌ای) بودند. محتوای آزمون‌ها از درس زیست‌شناسی سال اول دورهٔ متوسطهٔ دوم رشتهٔ علوم تجربی (مبحث گوارش غذا) استفاده‌شده بود. (نمونه‌هایی از تکمیل و ترسیم هر کدام از چهار آزمون مبتنی بر نقشهٔ مفهومی در بخش پیوست‌ها ارائه شده است).

ج) مقیاس بارشناختی ناسا- تی‌آل‌ایکس[41]

مقیاس ناسا (هارت و استیولند[42]، ۱۹۸۸) یک ابزار ارزیابی چندبعدی در قالب پرسش‌نامه است که هدف آن ارزیابی و سنجش حجم و فشار ادراکی به‌منظور ارزیابی کار (یا فعالیت خاص) یک سیستم،

کارایی و اثربخشی یک تیم، یا سایر جنبه‌های دیگر عملکرد در نظر گرفته شده است. این ابزار را گروه عملکرد انسانی[۴۳] در مرکز تحقیقات ایمز ناسا[۴۴] در مدت سه سال و انجام بیش از ۴۰ شبیه‌سازی آزمایشگاهی توسعه دادند.

آیتم‌های مقیاس ناسا

ناسا- تی آل ایکس در اصل شامل دو بخش است. کل فشار کاری یک فعالیت به شش زیر مقیاس تقسیم می‌شود که در یک صفحه نشان داده شده است و به‌عنوان بخش نخست پرسش‌نامه به کار می‌رود: الزامات ذهنی[۴۵]؛ الزامات فیزیکی[۴۶]؛ الزامات زمانی[۴۷]؛ عملکرد[۴۸] ؛ تلاش[۴۹] ؛ سرخوردگی (ناکامی)[۵۰].

برای هر یک از این خرده‌مقیاس‌های شش‌گانه توضیحاتی داده می‌شود که کاربر باید این موضوع را قبل از امتیازدهی بخواند. در این پرسش‌نامه، هر زمینه از فعالیت در بازه ۱۰۰ امتیاز با گام‌های پنج امتیازی تقسیم‌بندی شده است. سپس این رتبه‌ها در مقیاس گذاشته شده‌اند. این توضیحات به شرح زیرند:

- **فشار ذهنی:** چه میزان فعالیت ذهنی و ادراکی مورد نیاز است؟ این فعالیت ساده بود یا پیچیده؟
- **فشار فیزیکی:** چه میزان فعالیت فیزیکی لازم است؟ همراه با استراحت بود یا به فعالیت شدید جسمی نیاز داشت؟
- **فشار زمانی:** با توجه به‌سرعت رویداد هر فعالیت، چه میزان فشار زمانی را احساس کردید؟ سرعت رویدادها و جلو رفتن فعالیت کند بود یا تند؟
- **عملکرد کلی:** تا چه میزان در انجام این کار موفق بودید؟ تا چه حد از عملکرد خود راضی بودید؟
- **سطح ناکامی و سرخوردگی:** چه میزان حس رنجش، خشم و استرس در مقابل حس خرسندی، آرامش و رضایت‌مندی در طول کار احساس کردید؟
- **تلاش:** از نظر ذهنی یا فیزیکی تا چه میزان مجبور شدید سخت تلاش کنید تا به سطح عملکرد مورد نظرتان برسید؟

تحلیل و بخش دوم ناسا - تی آل ایکس

در بخش دوم ناسا، این خرده‌مقیاس‌ها به افراد اجازه می‌دهند با مقایسهٔ دوبه‌دوی آن‌ها و بر اساس اهمیت درک خود، وزن فردی و شخصی ایجاد کنند. در این حالت فرد از بین هر دو انتخاب دوتایی، آن موردی را که بیشتر به آن فعالیت ارتباط دارد گزینش می‌کند و هر بار انتخاب، با یک نمرهٔ وزنی برای آن مورد برابر است. حاصل‌ضرب این نمرهٔ وزنی در نمرهٔ مقیاس هر بعد تقسیم بر ۱۵، برابر است با نمرهٔ کل از صفر تا صد.

روبیـو، دیاز، مارتیـن و پونته[۵۱] (۲۰۰۴) در پژوهش خود، روایی همگـرا، روایی هم‌زمان و روایی تشـخیصی[۵۲] مقیاس ناسـا را مورد تائید قرار دادند. به‌منظور برآورد روایی همگرا، ضریب همبستگی پیرسون بین مقیاس ناسا و مقیاس تکنیک ارزیابی بار کاری ذهنی[۵۳] ۰/۹۸۱ و ضریب همبستگی پیرسون بین مقیاس ناسا و مقیاس نیمرخ بار کاری[۵۴] ۰/۹۸۶ به دست آمد که نشان از روایی همگرای بسیار بالای این مقیاس داشت. به‌منظور برآورد روایی هم‌زمان، ضریب همبستگی پیرسون بین امتیازات مقیاس ناسا و دو مقیاس عملکردی (تکلیف ردیابی و تکلیف جست‌وجوی حافظه[۵۵] محاسبه شد. ضریب همبستگی پیرسون بین امتیازات مقیاس ناسا با عملکرد در تکلیف ردیابی ۰/۶۵ و با تکلیف جست‌وجوی حافظه ۰/۷۵ بود. به‌منظور برآورد روایی تشخیصی نیز از ضریب تابع تشخیص بین دو تکلیف واحد و دوگانه[۵۶] اسـتفاده شد. تابع تشخیص محاسبه شـده با درصد واریانس ۹۸ درصد و Rc =۰/۷۲ به‌خوبی بین دو تکلیف واحد و دوگانه تمییز قائل شد که نشان از روایی افتراقی مناسب مقیاس ناسا داشت.

در ایران نیز محمدی، مظلومی، نسـل سـراجی و زراعتی (۱۳۹۲) بـرای تأیید اعتبار صوری، این پرسـش‌نامه را به زبان فارسـی ترجمه و همراه دو نفر از متخصصان رشـتهٔ ارگونومی بررسی کردند و مجدداً به زبان انگلیسـی برگرداندند. سپس نسخهٔ انگلیسی ترجمه‌شده نیز بازبینی شد و پس از اعمال نظرات اصلاحی، اعتبار صوری پرسـش‌نامه مورد تأیید قرار گرفت. پایایی پرسش‌نامه نیز با استفاده از روش آلفای کرونباخ ۰/۸۴۷ گزارش شد.

■ طرح پژوهش

طرح پژوهشی مورد استفاده در این تحقیق یک طرح ارزشیابی (اجراشده در سال ۱۳۹۷) بود. طبق نظر روسی (۱۹۸۲؛ به نقل از سیف، ۱۳۸۷) این دسته از پژوهش‌ها نوعی ارزشیابی هستند که بر شواهد تجربی جمع‌آوری‌شـده به راه‌های قابل تکرار مبتنی هسـتند و در آن‌ها مسائل مربوط به روایی درونی، روایی بیرونی و روایی سـازه مورد توجه قرار می‌گیرد. در پژوهش حاضر نیز شـاخص‌های مربوط به روایی و پایایی آزمون‌های مبتنی بر نقشهٔ مفهومی بررسی شدند.

جدول ۱. طرح پژوهش

گروه	اعمال متغیر مستقل	پس‌آزمون
۱	آموزش تکمیل نقشهٔ سلسله‌مراتبی	اجرای آزمون‌های چهارگزینه‌ای، بارشناختی و تکمیل نقشه
۲	آموزش تکمیل نقشهٔ شبکه‌ای	اجرای آزمون‌های چهارگزینه‌ای، بارشناختی و تکمیل نقشه
۳	آموزش ترسیم نقشهٔ سلسله‌مراتبی	اجرای آزمون‌های چهارگزینه‌ای، بارشناختی و ترسیم نقشه
۴	آموزش ترسیم نقشهٔ شبکه‌ای	اجرای آزمون‌های چهارگزینه‌ای، بارشناختی و ترسیم نقشه

روش اجرای پژوهش

جزئیات اجرای پژوهش در قالب شش مرحله به شرح زیر بود:

۱. آماده‌سازی

در ایـن مرحلـه، مقدمات اجرای پژوهش فراهم شـد. فعالیت‌های این مرحلـه عبارت بودند از: انتخاب مدرسه‌ها و کلاس‌های نمونه، هماهنگی با مدرسه‌ها، آماده‌سازی ابزار و وسایل لازم، و آموزش مدرسان برای آموزش نقشه‌کشی مفهومی.

۲. آموزش

قبل از اجرای آزمون‌ها، دو جلسهٔ ۶۰ دقیقه‌ای آموزشی برای یاددادن نحوهٔ تکمیل یا ترسیم نقشهٔ مفهومی بر روی چهار گروه اجرا شد. در طول این دو جلسه آموزش، فراگیرندگان چهار گروه آزمایشی با شیوه‌های متفاوت، تحت آموزش مطالب آموزشی یکسانی (مبحث گوارش غذا) قرار گرفتند. جزئیات اجرا در این چهار گروه به شرح زیر بود:

رویهٔ آموزش دو گروه ترسیم نقشهٔ مفهومی (سلسله‌مراتبی و شبکه‌ای):

آموزش نقشـهٔ مفهومی در این دو گروه، سـه مرحلهٔ اصلی را در بر گرفت که هر کدام شـامل چند زیرمرحله بود:

- **مرحلهٔ اول: تعریف و توصیف نقشهٔ مفهومی و فواید نقشهٔ مفهومی.**
- **مرحلهٔ دوم: طراحی نقشهٔ مفهومی**

گام اول: در حین (یا بعد از) خواندن یک متن، روی آن قسمت از متن که به آن علاقه‌مند هستید تمرکز کنید و از خودتان بپرسید نقشهٔ مفهومی من دربارهٔ چیست؟

گام دوم: مفاهیم مناسب و مربوطهٔ متن را شناسایی کنید.

گام سـوم: مفاهیم را فهرسـت کنید و دربارهٔ مفاهیم اصلی و کلی و مفاهیمی که زیرمجموعه و فرعی هستند، فکر کنید. مفاهیم کلی را در بالا (در نقشه‌های مفهومی سلسله‌مراتبی) یا در مرکز (در نقشه‌های مفهومی شبکه‌ای) نقشه قرار دهید و مفاهیم با جامعیت و اهمیت کمتر را در پایین نقشه (در نقشه‌های مفهومی سلسله‌مراتبی) یا در حاشیه‌های نقشه (در نقشه‌های مفهومی شبکه‌ای) قرار دهید.

گام چهارم: روابط میان مفاهیم را برقرار کنید. با مفاهیم کلی شروع کنید.

گام پنجم: عبارت‌های ربط‌دهندهٔ مناسب را انتخاب و بررسی کنید آیا روابط بین مفاهیم معنادارند.

گام ششم: با استفاده از فلش‌ها (پیکان‌ها) جهت روابط را تعیین کنید تا روابط به گزاره (جملات معناداری با مفاهیم، عبارت‌های ربط‌دهنده و جهت) تبدیل شوند.

گام هفتم: تمام روابط را با مفاهیمی که شناسایی کردید، به‌صورت نقشه در آورید.

گام هشتم: ساختار نقشه را دوباره بازبینی و اصلاح کنید.
گام نهم: نمونه‌هایی از نقشه‌های مفهومی سلسله‌مراتبی یا شبکه‌ای را مشاهده کنید.

- **مرحلهٔ سوم: تمرین طراحی نقشهٔ مفهومی**

در این مرحله، شاگردان پس از آشنایی با نقشهٔ مفهومی و نحوهٔ طراحی آن، نقشه‌هایی مفهومی دربارهٔ موضوعات درس زیست‌شناسی (سال اول متوسطهٔ دوم) ترسیم کردند.
دو نکتهٔ مورد توجه در این مرحله عبارت بودند از:

۱. شروع از وسط نقشه و حرکت تدریجی به کناره‌ها (در نقشه‌های مفهومی شبکه‌ای) و یا شروع از بالای نقشه و حرکت تدریجی به پایین (در نقشه‌های مفهومی سلسله‌مراتبی)؛
۲. حرکت مطالب در جهت عقربه‌های ساعت در نقشه‌های مفهومی شبکه‌ای و حرکت مطالب از سمت راست نقشه به سمت چپ نقشه در نقشه‌های مفهومی سلسله‌مراتبی.

رویهٔ آموزش دو گروه تکمیل نقشهٔ مفهومی (سلسله‌مراتبی و شبکه‌ای):
در این دو گروه نیز آزمودنی‌ها دو مرحلهٔ اصلی آموزشِ دو گروه قبلی را دریافت کردند، اما در مرحلهٔ سوم، به‌جای تمرین ترسیم نقشه، نقشه‌های مفهومی ناقص از قبل تهیه‌شده را دریافت و کامل‌کردن آن‌ها را تمرین کردند.

۳. ارزشیابی

در مرحلهٔ سوم، آزمودنی‌های تمام گروه‌ها به یک آزمون چهارگزینه‌ای، آزمون مبتنی بر نقشهٔ مفهومی و آزمون بار شناختی پاسخ دادند. مدت‌زمان اجرای این مرحله ۱۰۰ دقیقه بود. در این مرحله، ابتدا هر گروه ظرف مدت ۱۵ دقیقه آزمون چهارگزینه‌ای مربوط به مبحث گوارش غذا را پاسخ دادند. بعد از چند دقیقه استراحت، هر کدام از گروه‌ها آزمون نقشهٔ مفهومی مربوط به خود را انجام دادند. به این صورت که آزمودنی‌های گروه اول باید نقشهٔ مفهومی سلسله‌مراتبی را ترسیم می‌کردند. آزمودنی‌های گروه دوم نیز باید نقشهٔ مفهومی شبکه‌ای را ترسیم می‌کردند. از آزمودنی‌های گروه سوم خواسته شد نقشهٔ مفهومی سلسله‌مراتبی ناقص را کامل کنند. در نهایت از گروه چهارم خواسته شد نقشهٔ مفهومی شبکه‌ای ناقص را کامل کنند. مدت‌زمان لازم برای انجام آزمون‌های مبتنی بر نقشهٔ مفهومی ۶۰ دقیقه بود. در نهایت، پس از اجرای آزمون‌های مبتنی بر نقشهٔ مفهومی و چند دقیقه استراحت، از آزمودنی‌ها خواسته شد در مدت ۱۵ دقیقه به آزمون بار شناختی پاسخ دهند.

۴. آموزش نمره‌گذاران

در پژوهش حاضر، برای نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی، از هشت نمره‌گذار مستقل استفاده شد. این هشت نمره‌گذار از معلمان زیست‌شناسی استان کرمانشاه بودند که در یک جلسهٔ آموزش ۹۰ دقیقه‌ای با روش‌های متعدد نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی آشنا شدند.

۵. نمره‌دهی آزمون‌ها

- **آزمون چهارگزینه‌ای:** برای نمره‌دهی این آزمون، به شمارش پاسخ‌های درست از کل ۳۳ سؤال بر اساس یک کلید پاسخ سؤالات اقدام و سپس نسبت نمرات نهایی از ۲۰ محاسبه و نمرهٔ نهایی هر آزمودنی مشخص شد.
- **دو آزمون ترسیم نقشهٔ مفهومی سلسله‌مراتبی و شبکه‌ای:** برای نمره‌دهی این آزمون‌ها از سه روش نمره‌دهی رابطه‌ای، ساختاری و شباهت به نقشهٔ متخصص استفاده شد.

روش رابطه‌ای: در این روش به هر رابطهٔ بین مفاهیم به‌طور جداگانه (بدون در نظر گرفتن دیگر رابطه‌ها) نمره‌ای اختصاص داده می‌شود. ترتیب دقیق‌تر این روش به شرح جدول ۲ است:

جدول ۲. روش نمره‌دهی رابطه‌ای نقشهٔ مفهومی (اقتباس از وست، پارک، پومروی و ساندوال[۵۷]، ۲۰۰۲)

بخش‌های نقشهٔ مفهومی	نمره
رابطهٔ معتبر بین مفاهیم که برچسب نادرست دارد	هر کدام ۱ نمره
رابطهٔ معتبر بین مفاهیم که برچسب درست دارد، اما رابطهٔ سلسله‌مراتبی، علی یا توالی بین مفاهیم را نشان نمی‌دهد	هر کدام ۲ نمره
رابطهٔ معتبر بین مفاهیم که برچسب درست دارد و رابطهٔ سلسله‌مراتبی، علی یا توالی بین مفاهیم را نشان می‌دهد.	هر کدام ۳ نمره
رابطهٔ بین مفاهیمی که هیچ رابطه‌ای با هم ندارند.	هر کدام ۰ نمره

روش ساختاری: در این روش، به دانش ساختاری سلسله‌مراتبی اهمیت داده می‌شود. ترتیب دقیق‌تر این روش به شرح جدول ۳ است:

جدول ۳. روش نمره‌دهی ساختاری نقشهٔ مفهومی (اقتباس از وست و همکاران، ۲۰۰۲)

بخش‌های نقشهٔ مفهومی	نمره
رابطهٔ سلسله‌مراتبی معتبر بین مفاهیم	هر کدام ۵ نمره
رابطهٔ سلسله‌مراتبی معتبر با عبارت ربط دهندهٔ نامعتبر	هر کدام ۳ نمره
گزاره‌های معتبر	هر کدام ۲ نمره
گزاره‌های معتبر با عبارت ربط‌دهندهٔ نامعتبر	هر کدام ۱ نمره
مثال‌های مفاهیم	هر کدام ۱ نمره
مفهوم یا رابطهٔ غیرمعتبر	هر کدام ۰ نمره

روش شباهت به نقشهٔ متخصص: در این رویکرد، نقشه‌های دانش‌آموزان در مقایسه با نقشهٔ مفهومی معیار که متخصصان ساخته بودند، مقایسه و ارزشیابی شد. در پژوهش حاضر، گزاره‌های نقشه‌های مفهومی شـاگردان با گزاره‌های نقشهٔ مفهومی متخصصان مورد مقایسه قرار گرفت. به این صورت که هم‌خوانی هر گزاره در نقشهٔ مفهومی آزمودنی با گزارهٔ موجود در نقشهٔ مفهومی معیار، یک نمره برای آزمودنی به همراه داشت.

۶. تجزیه‌وتحلیل داده‌ها

در مرحلهٔ نهایی پژوهش، یافته‌های به دست‌آمده مورد تجزیه‌وتحلیل آماری قرار گرفتند.

روش‌های آماری: برای آگاهی از روایی ملاکی، از ضریب همبسـتگی پیرسون بین نمرات حاصل از نقشه‌های مفهومی ترسیم‌شده (در دو گروه ترسیم نقشهٔ مفهومی سلسله‌مراتبی و شبکه‌ای) با آزمون چهارگزینه‌ای پیشرفت تحصیلی استفاده شد. به‌منظور آگاهی از روایی تمییزی، از آزمون t دونمونه‌ای مستقل برای بررسی تفاوت گروه‌های دارای پیشرفت تحصیلی بالا و پایین استفاده شد. برای آگاهی از پایایی روش‌های نمره‌دهی، از پایایی بین نمره‌گذاران[۵۸] (با استفاده از روش همبستگی پیرسون) استفاده شـد. به‌منظور تعیین تفاوت بار شـناختی آزمون‌های متعدد مبتنی بر نقشهٔ مفهومی نیز از آزمون تحلیل واریانس یک‌راهه[۵۹] (با استفاده از آزمون ولچ[۶۰] و آزمون تعقیبی گیمز- هاول[۶۱]) استفاده شد.

■ یافته‌ها

در این پژوهش، برای آگاهی از میزان مطلوب‌بودن روایی ملاکی روش‌های گوناگون نمره‌دهی، از نمرات دانش‌آموزان در یک آزمون چهارگزینه‌ای به‌عنوان آزمون ملاک استفاده شد. برای به‌دست‌آوردن نمرات آزمون‌های ترسـیم نقشهٔ مفهومی نمرات هشـت نمره‌گذار با هم جمع و بر تعداد نمره‌گذاران (هشت) تقسیم شد تا نمرهٔ میانگین به‌دست آید. سپس ضریب همبستگی بین این نمرهٔ میانگین با نمرهٔ آزمون چهارگزینه‌ای بر اساس هر کدام از روش‌های نمره‌دهی محاسبه شد.

جدول ۴. ضریب همبستگی نمرات آزمون چهارگزینه‌ای با روش‌های نمره‌دهی

		روش نمره‌دهی شباهت	روش نمره‌دهی ساختاری	روش نمره‌دهی رابطه‌ای
آزمون چهارگزینه‌ای	همبستگی	۰/۵۳	۰/۵۳	۰/۵۵
	معناداری	۰/۰۰۰۱	۰/۰۰۰۱	۰/۰۰۰۱
	n	۵۴	۵۴	۵۴

همان‌گونه که در جدول ۱ مشـــاهده می‌شود، هر سه روش نمره‌دهی، ضرایب همبستگی متوسط و معناداری با آزمون چهارگزینه‌ای پیشرفت تحصیلی دارند.

برای اطمینان از روایی افتراقی (تمییزی) روش‌های نمره‌دهی در پژوهش حاضر، توان این روش‌ها در تمایز گروه‌های بالا و پایین پیشـــرفت تحصیلی بررسی شـــد. برای تعیین گروه‌های دانش‌آموزان دارای پیشـــرفت تحصیلی بـــالا و پایین، ابتدا میانگین نمرات دانش‌آمـــوزان در آزمون چهارگزینه‌ای محاســـبه شد و ســـپس نمرات بالاتر از میانگین به‌عنوان گروه دارای پیشرفت تحصیلی بالا و نمرات پایین‌تر از میانگین به‌عنوان گروه دارای پیشـــرفت تحصیلی پایین در نظر گرفته شد. از یک آزمون t دونمونه‌ای مســـتقل نیز برای بررســـی تفاوت میانگین دو گروه بالا و پایین در هر کدام از روش‌های نمره‌دهی استفاده شد.

جدول ۵. نتایج سه آزمون مستقل برای بررسی تفاوت میانگین گروه‌های بالا و پایین بر اساس روش‌های نمره‌دهی

روش نمره‌دهی	آزمون لون		t	درجه آزادی	معناداری	تفاوت میانگین	فاصلهٔ اطمینان ۹۵% تفاوت	
	F	معناداری					پایین‌تر	بالاتر
شباهت	۵/۲۴	۰/۰۲۶	۳/۱۶	۴۲/۴۱	۰/۰۰۲	۳/۴	۱/۲۹	۵/۵۱
ساختاری	۴/۶۳	۰/۰۳۶	۳/۷۳	۴۱/۹۶	۰/۰۰۱	۳/۶۴	۱/۵۶	۵/۷۲
رابطه‌ای	۵/۳۸	۰/۰۲۴	۳/۴۹	۴۰/۸۳	۰/۰۰۱	۳/۷۳	۱/۵۷	۵/۸۹

همان‌گونه که در جدول ۵ مشـــاهده می‌شـــود، نتایج هر سه آزمون t مستقل معنادار است و هر سه روش نمره‌دهی به‌خوبی توانسته‌اند بین دو گروه دارای پیشرفت تحصیلی بالا و پایین تمییز قائل شوند. این نتیجه نشان‌دهندهٔ روایی تمییزی مناسب سه روش نمره‌دهی رابطه‌ای، ساختاری و شباهت به نقشهٔ متخصص است.

برای بررسی پایایی آزمون‌های ترسیم نقشهٔ مفهومی نیز از روش پایایی بین نمره‌گذاران استفاده شد. به این منظور از ضرایب همبستگی پیرسون بین نمرات هشت نمره‌گذار استفاده شد.

جدول ۶. ضرایب همبستگی نمرات نمره‌گذاران در هر کدام از روش‌های نمره‌دهی

روش نمره‌دهی	ضریب همبستگی پیرسون	تعداد	معناداری
شباهت	۰/۹۵	۵۴	۰/۰۰۰۱
ساختاری	۰/۹۹	۵۴	۰/۰۰۰۱
رابطه‌ای	۰/۹۹	۵۴	۰/۰۰۰۱

همان‌گونه که در جدول ۶ مشـــخص است، ضرایب همبستگی بین نمرات نمره‌گذاران در هر سه روش نمره‌دهی رابطه‌ای، ساختاری و شباهت به نقشهٔ متخصص، بالا و معنادار هستند. این نشان می‌دهد هر سه روش نمره‌دهی از پایایی مناسبی بین نمره‌گذاران برخوردارند.

در پژوهش حاضر، به‌منظور تعیین تفاوت بار شـــناختی آزمون‌های متعدد مبتنی بر نقشهٔ مفهوم، از تحلیل واریانس یکطرفه استفاده شد. ابتدا نتایج آزمون همگنی واریانس‌ها[۶۲] ارائه شده است.

جدول ۷. آزمون همگنی واریانس‌ها

آمارهٔ لون	درجهٔ آزادی ۱	درجهٔ آزادی ۲	معناداری
۹/۸۰	۳	۱۰۵	۰/۰۰۰۱

همان‌گونه که جدول ۷ نشـــان می‌دهد، آزمون همگنی واریانس‌ها معنادار اســـت. معناداربودن این آزمون به این معناست که فرض همگنی واریانس‌ها رد شده‌است و بنابراین واریانس‌ها برابر نیستند. در شرایطی که واریانس‌ها همگن نباشند، به‌جای آزمون F باید از آزمون ولچ و به‌جای آزمون‌های تعقیبی رایج باید از آزمون‌های تعقیبی مانند گیمز- هاول استفاده کرد.

جدول ۸. آزمون ولچ

آماره	درجهٔ آزادی ۱	درجهٔ آزادی ۲	معناداری
۲۱/۵۹	۳	۵۳/۹۱۲	۰/۰۰۰۱

همان‌گونه که در جدول ۸ مشـــاهده می‌شود، آزمون ولچ به‌منظور بررسی تفاوت میانگین‌ها معنادار است. بنابراین از لحاظ بار شناختی، بین گروه‌ها تفاوت معناداری وجود دارد.

پس از معناداری تفاوت میان میانگین‌های گروه‌ها، از آزمون تعقیبی گیمز-هاول به‌منظور بررســـی اینکه تفاوت بین کدام گروه‌ها معنادار است، استفاده شد.

جدول ۹. آزمون تعقیبی گیمز- هاول به‌منظور بررسی تفاوت بین میانگین‌های گروه‌ها

گروه	تفاوت میانگین	انحراف معیار	معناداری
تکمیل شبکه‌ای تکمیل سلسله‌مراتبی	۱۱/۲۱	۵/۲۳	۰/۱۵۸
ساخت شبکه‌ای	- ۳۵/۹	۸/۱۶	۰/۰۰۰۱
ساخت سلسله‌مراتبی	- ۱۷/۷۶	۶/۴۵	۰/۰۳۹
تکمیل سلسله‌مراتبی ساخت شبکه‌ای	- ۴۷/۱۲	۷/۱۱	۰/۰۰۰۱
ساخت سلسله‌مراتبی	- ۲۸/۹۸	۵/۰۵	۰/۰۰۰۱
ساخت شبکه‌ای ساخت سلسله‌مراتبی	۱۸/۱۳	۸/۰۵	۰/۱۲۶

بر اساس جدول ۹ بین دو آزمون تکمیل نقشهٔ مفهومی شبکه‌ای و تکمیل نقشهٔ مفهومی سلسله‌مراتبی تفاوت معناداری در زمینهٔ بار شناختی وجود ندارد، اما بین آزمون تکمیل نقشهٔ مفهومی شبکه‌ای با دو آزمون ساخت نقشهٔ مفهومی شبکه‌ای و ساخت نقشهٔ مفهومی سلسله‌مراتبی تفاوت معناداری از لحاظ بار شناختی وجود دارد. یعنی بار شناختی آزمون تکمیل نقشهٔ مفهومی شبکه‌ای از بار شناختی دو آزمون ساخت نقشهٔ مفهومی شبکه‌ای و ساخت نقشهٔ مفهومی سلسله‌مراتبی به‌طور معناداری کمتر است. همچنین، بر اساس جدول ۹، بین آزمون تکمیل نقشهٔ مفهومی سلسله‌مراتبی و دو آزمون ساخت نقشهٔ مفهومی شبکه‌ای و ساخت نقشهٔ مفهومی سلسله‌مراتبی تفاوت معناداری از لحاظ بار شناختی وجود دارد. یعنی بار شناختی آزمون تکمیل نقشهٔ مفهومی شبکه‌ای از بار شناختی دو آزمون ساخت نقشهٔ مفهومی شبکه‌ای و ساخت نقشهٔ مفهومی سلسله‌مراتبی به‌طور معناداری کمتر است.

از دیگر نتایج آزمون تعقیبی این است که بین دو آزمون ساخت نقشهٔ مفهومی شبکه‌ای و ساخت نقشهٔ مفهومی سلسله‌مراتبی، از لحاظ بار شناختی تفاوت معناداری وجود ندارد؛ هر چند که میانگین بار شناختی گروه ساخت نقشهٔ مفهومی سلسله‌مراتبی پایین‌تر بود.

■ بحث و نتیجه‌گیری ■

هدف اصلی پژوهش حاضر بررسی بار شناختی و روش‌های نمره‌دهی آزمون‌های گوناگون مبتنی بر نقشهٔ مفهومی بود. نتایج نشان داد بار شناختی آزمون‌های گوناگون مبتنی بر نقشهٔ مفهومی با یکدیگر متفاوت است و آزمون‌های ترسیم نقشهٔ مفهومی نسبت به آزمون‌های تکمیل نقشهٔ مفهومی به‌طور معناداری بار شناختی بیشتری بر آزمودنی‌ها تحمیل می‌کنند. در زمینهٔ روش‌های نمره‌دهی نیز نتایج پژوهش نشان داد سه روش نمره‌دهی رابطه‌ای، ساختاری و شباهت به نقشهٔ متخصص، از روایی ملاکی، تمییزی و پایایی مناسبی برخوردارند.

در رابطه با بار شناختی بالاترِ نقشه‌های مفهومی ترسیم‌شده توسط یادگیرنده نسبت به نقشه‌های مفهومی تکمیل شدنی، محققان اظهار داشته‌اند که تکالیف نقشه‌کشی مفهومی با هدایت کم (مانند فرایند ترسیم نقشهٔ مفهومی) دانش محتوایی بیشتری را از شاگردان می‌طلبد. بنابراین، این تکالیف ممکن است برای شاگردان بسیار چالش‌آمیز و دشوار باشند (استراوتمین، ۲۰۱۲). ساگلام[۶۳] (۲۰۰۶) نیز در پژوهش خود در این زمینه به این نتیجه دست یافت که آزمودنی‌ها کامل‌کردن نقشه‌های مفهومی را آسان‌تر از ترسیم نقشه‌های مفهومی می‌دانند. از طرف دیگر، چانگ، سانگ، و چن[۶۴] (۲۰۰۱) نیز در پژوهش خود، اثربخشی نقشه‌های مفهومی کامل‌کردنی و نقشه‌های مفهومی رسم‌کردنی را مقایسه کردند. نتایج نشان داد، ۹۴ درصد از آزمودنی‌ها معتقدند که تکلیف ترسیم نقشهٔ مفهومی کاری بود که تلاش بیشتری را می‌طلبید و نیازمند توجه به جزئیات پیچیده‌ای بود. بنابراین، نقشه‌های مفهومی کامل‌کردنی بار کاری کمتری را تحمیل می‌کردند.

پاس[۶۵] (۱۹۹۲؛ نقل از چانگ و همکاران، ۲۰۰۱) نیز پیشنهاد کرده است که باید از تحمیل بار اضافی بالقوه بر فراگیرندگان اجتناب کرد. وی اظهار داشت، راهبردهای تکمیل (مانند روش کامل‌کردن نقشه‌های مفهومی) روش مناسبی برای ارائهٔ مواد و مطالب یادگیری هستند و از بار شناختی می‌کاهند، چرا که انرژی ذخیره‌شده در این تکالیف، برای کارکردن روی جاهای خالی و مهارت در زمینهٔ کل محتوا مصرف می‌شود.

راهبرد تکمیل را در اصل ون‌مرینبوئر[۶۶] (۱۹۹۰) پیشنهاد داده است که یک راه‌حل بالقوه برای چیره‌شدن بر اضافه‌بار شناختی هنگام حل‌کردن مسائل است (لیم[۶۷]، ۲۰۰۸). بعضی از محققان (مانند شائو، و میترن[۶۸]، ۱۹۹۷) اظهار داشته‌اند، اگر از یادگیرندگان بخواهیم یک نقشهٔ مفهومی را ترسیم کنند (به‌منظور

اینکه دانششان را به‌صورت معناداری به نمایش بگذارند)، الزامات شناختی بسیار بالایی را بر آن‌ها تحمیل می‌کنیم. به همین دلیل، آن‌ها تکنیک کامل‌کردن نقشه‌های مفهومی را پیشنهاد داده‌اند (روئیز – پریمو، ۲۰۰۰).

بعضی از محققان اظهار داشته‌اند، فرایند ترسیم نقشهٔ مفهومی می‌تواند ملال‌آور و ناامیدکننده باشد، چرا که ایجاد یک نقشهٔ مناسب، غالباً به بازبینی‌های زیادی نیاز دارد. زمانی که از یادگیرندگان می‌خواهیم کل یک نقشه را ترسیم کنند، این الزامات شناختی در سطح بالایی (هم در سطح کلامی و هم در سطح بصری–فضایی) به شاگردان تحمیل می‌شود و حتی زمانی که تکلیف ارزشیابی از یادگیرندگان می‌خواهد با استفاده از مفاهیم داده‌شده نقشه را ترسیم و روابط بین مفاهیم را برقرار کنند (مانند تکلیف ارزشیابی ترسیم نقشهٔ مفهومی در پژوهش حاضر)، سطح الزامات شناختی همچنان بالاست. این محققان اظهار داشتند، این سطح بالای الزامات شناختی ممکن است به‌واسطهٔ موفقیت کمتر یادگیرندگان و همچنین به‌واسطهٔ تفاوت زبانی یادگیرندگانی که زبان اول آن‌ها زبان مورد استفاده در تکلیف ارزشیابی نیست (مانند شرایط پژوهش حاضر) مشکل‌آفرین و پیچیده‌تر نیز بشود (شائو، و میترن، ۱۹۹۷).

علاوه بر این، محققان عنوان کرده‌اند که آزمودنی‌ها در زمان آزمون مجبورند در ذهن خود، بین مؤلفه‌های آموخته‌شده، تعامل بیشتری برقرار کنند. بنابراین، درگیری ذهنی بیشتری با سؤالات خواهند داشت و این خود می‌تواند بر بار شناختی آنان بیفزاید. از طرف دیگر، در زمان آزمون، درگیری ذهنی ناشی از پیامدهای آزمون و اضطراب ناشی از آن سبب می‌شود آن‌ها تمام تلاش ذهنی خود را برای پاسخ به آزمون به‌کار گیرند و این موضوع می‌تواند بر بار شناختی آن‌ها هنگام آزمون بیفزاید (احدی و سلیمانی، ۱۳۹۳). بنابراین، فرایندهای ایجاد تعامل بین مؤلفه‌ها و اضطراب و درگیری ذهنی، در آزمون‌های ترسیم نقشهٔ مفهومی، نسبت به آزمون‌های تکمیل نقشهٔ مفهومی، به‌مراتب بیشتر خواهد بود، چرا که پاسخ‌گویی به آزمون‌های ترسیم نقشهٔ مفهومی نیازمند دانش محتوایی بیشتر، توجه به جزئیات ویژه و استفاده از مهارت ترکیب (برای خلق نقشهٔ مفهومی) است که می‌تواند به ایجاد اضافه‌بار شناختی بر آزمودنی‌ها منجر شود.

از دیگر دلایلی که بتوان در این رابطه عنوان کرد، این است که در نظام آموزشی کشور، به‌طور معمول در بیشتر دوره‌های تحصیلی در کلاس‌های درس برای ارزشیابی آموخته‌های فراگیرندگان از آزمون‌های چندگزینه‌ای

استفاده می‌شود و آزمودنی‌ها به‌نوعی به این‌گونه ارزشیابی‌ها عادت کرده‌اند؛ ارزشیابی‌هایی که فرایند پاسخ‌گویی به آن‌ها (همچون آزمون‌های تکمیل نقشهٔ مفهومی) معمولاً نیازمند صرف تلاش ذهنی و فیزیکی و زمان کمتری است. لذا فراگیرندگان هنگام رویارویی با آزمون‌هایی که ساختار متفاوتی دارند (همچون آزمون‌های ترسیم نقشهٔ مفهومی) احساس سردرگمی و خستگی می‌کنند و نظام شناختی آن‌ها بار شناختی بیشتری را متحمل می‌شود.

در رابطه با روش‌های نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی نیز باید اظهار داشت، هر کدام از روش‌های نمره‌دهی نقشه‌های مفهومی جنبه‌های ویژه‌ای از ساختار دانش و درک شاگردان را اندازه‌گیری می‌کنند. برای مثال، روش نمره‌دهی رابطه‌ای که روی کیفیت گزاره‌ها تأکید می‌کند، به ارزشیابی میزان مطابقت درک شاگردان از مفاهیم، با مفاهیم ارائه‌شده در جریان آموزش می‌پردازد. در مقابل، روش نمره‌دهی ساختاری که روی ساختار سلسله‌مراتبی نقشه‌های مفهومی تأکید می‌کند، میزان پیچیدگی ساختارهای ذهنی شاگردان را مورد ارزشیابی قرار می‌دهد.

با وجود اینکه روش ساختاری ممکن است میزان جذب دانش جدید به داخل دانش موجود شاگردان را به‌گونه‌ای مؤثر مورد آزمون قرار دهد، اما ممکن است در تحلیل اینکه شاگردان تا چه میزان دانش مفهومی را به کمک آموزش درک کرده و دانششان را بهبود بخشیده‌اند، نتایج رضایت‌بخشی فراهم نیاورد. بنابراین، استفاده از چند روش نمره‌دهی به‌طور هم‌زمان، ارزیاب را قادر می‌سازد جنبه‌های متفاوت دانش مفهومی شاگردان را ارزشیابی کند. برای ارزشیابی رشد مفهومی شاگردان چند روش نمره‌دهی می‌توانند در ارزشیابی نقشه‌های مفهومی به‌کار گرفته شوند (اودا[۶۹]، ۲۰۱۱).

انتقادهای دیگری نیز از روش‌های نمره‌دهی ساختاری و شباهت صورت گرفته‌اند. برای مثال، کینچین، هی و آدامز[۷۰] (۲۰۰۰) نگرانی‌هایی را در رابطه با نمرات نقشه‌هایی که با جمع‌کردن عناصری مانند رابطه‌های معتبر، همبرها و ساختار سلسله‌مراتبی به دست می‌آیند، اظهار داشته و ادعا کرده‌اند، نمراتی که به‌صورت کلی به‌دست می‌آیند، مبهم و نامشخص هستند (رأی و روبا[۷۱]، ۲۰۰۲). پلومر[۷۲] (۲۰۰۸) نیز عنوان داشت، پیچیدگی شناختی در روش نمره‌دهی ساختاری معمولاً بالاست و نمره‌گذاران باید تعیین کنند چه چیزی ساختارهای سطح بالاتر را در یک نقشهٔ مفهومی تشکیل می‌دهد (مثلاً چه چیزی یک همبر

یـا سلسـله‌مراتب را تشـکیل می‌دهد). برای مثال سلسـله‌مراتب ممکن اسـت ساختاری پیچیده داشته باشد و شناسایی و ارزشیابی این ساختارها ممکن است محدودیت‌هایی را روی حافظهٔ کاری نمره‌گذاران تحمیل کند.

همچنیـن، به این خاطر که محققان فرض گرفته‌اند سـاختار دانش به‌صورت سلسله‌مراتبی است (آزوبل[۷۳]، ۱۹۶۰)، روش نمره‌دهی ساختاری روی استفاده از سـطوح سلسله‌مراتب و همبرها تأکید بیشتری دارد. بنابراین، در این روش، نمرهٔ بیشتری به این مؤلفه‌ها داده می‌شود. به‌عنوان یک نتیجه، این روش نمره‌دهی فقط برای استفاده در نقشه‌های مفهومی سلسله‌مراتبی مناسب است (میسا[۷۴]، ۲۰۱۰).

عده‌ای از محققان هم از روش نمره‌دهی شـباهت به نقشـهٔ متخصص انتقاد کرده‌انـد. آن‌ها اظهار داشـته‌اند، مقدار (نرخ) شـباهت فقط به شـرح‌دادن این موضـوع کمک می‌کند که شـاگردان یک موضوع مشـخص را می‌دانند و به‌طور واضـح وضعیـت یادگیری شـاگردان را اندازه‌گیـری نمی‌کنـد. بنابراین، چنین ارزشـیابی‌هایی تفاوت وضعیت یادگیری در شاگردان متفاوت را به‌طور کامل به تصویر نمی‌کشـند. برای مثال، دو شـاگرد با مقدار (نرخ) شباهت یکسان در یک ارزشـیابی، ممکن است درک و فهم متفاوتی از موضوع داشته باشند. از آنجا که هنگامی که نرخ شـباهت یکسـان اسـت هیچ راهی برای تصمیم‌گیری در مورد تفاوت نتایج یادگیری وجود ندارد، بازخورد فراهم‌شـده توسـط نرخ شـباهت، کمک مناسبی به این شاگردان نیست (چانگ، سانگ، چانگ و لین[۷۵]، ۲۰۰۵).

برخی از محققان نیز (ییـن، وانیدس[۷۶]، روئیز-پریمو و شیولسـون، ۲۰۰۵؛ و روئیز-پریمو و همکاران، ۲۰۰۹؛ نقل از شـوندیمان، ۲۰۱۱) اظهار داشـته‌اند، با توجه به اینکه نقشه‌های مفهومی می‌توانند تعداد زیادی رابطهٔ به‌نسبت سطحی و کم‌اهمیت نیز داشـته باشند، لزومی به نمره‌دهی به تمام رابطه‌های موجود نیست و می‌تـوان تعدادی از رابطه‌هـای مهم را نمره‌دهی کرد. آن‌ها پیشـنهاد می‌کنند، نمره‌دهی رابطه‌های مهم نسـبت به اندازه‌گیری تغییرات در درک و فهم شاگردان حساس‌تر است، چرا که در این روش فقط روی مفاهیم و ایده‌های کلیدی نقشهٔ مفهومی تمرکز می‌شـود. این روش نمره‌دهی از لحـاظ صرفه‌جویی در زمان نیز روش مؤثری محسوب می‌شود و آزادی عمل بیشتری برای شاگردانی که نقشه‌ها را می‌سـازند فراهم می‌کند. بنابراین، می‌تواند در ارزشـیابی‌های مبتنی بر نقشـهٔ مفهومی در مقیاس وسیع نیز کمک‌کننده باشد.

نکتهٔ قابل ذکر در زمینهٔ روش‌های نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی

این است که روش نمره‌دهی رابطه‌ای از چندین لحاظ نسبت به دیگر روش‌ها برتری دارد: اول اینکه این روش نمره‌دهی از نظر زمانی مقرون به‌صرفه است، چرا که پیچیدگی کمتری دارد و فرایند نمره‌دهی نقشه‌های مفهومی با سرعت بیشتری صورت می‌گیرد. دوم اینکه با توجه به نظر نمره‌گذاران نقشه‌های مفهومی در پژوهش حاضر، این روش نمره‌دهی فشار کاری کمتری بر نمره‌گذاران وارد می‌کند و راحت‌تر و جذاب‌تر است. سوم اینکه این روش نمره‌دهی انعطاف‌پذیری بالایی دارد و می‌تواند در انواع تکالیف ترسیم نقشهٔ مفهومی مورد استفاده قرار گیرد. برای مثال، این روش می‌تواند در نقشه‌هایی مفهومی که در آن‌ها سلسله‌مراتب مفاهیم وجود ندارد (مانند نقشه‌های مفهومی سیستمی یا زنجیری) نیز به کار رود، چرا که این روش نمره‌دهی فقط با گزاره‌ها (کوچک‌ترین واحد ساخت دانش) سر و کار دارد. چهارم اینکه این روش نمره‌دهی از انتقاد ابهام و کلیبودن مصون است.

در مجموع، بر اساس نتایج پژوهش حاضر باید اظهار داشت، سه روش نمره‌دهی رابطه‌ای، ساختاری و شباهت به نقشهٔ متخصص، از لحاظ شاخص‌های روایی و پایایی شرایط مناسبی دارند و می‌توانند به‌خوبی برای نمره‌دهی آزمون‌های مبتنی بر نقشهٔ مفهومی به کار روند. همچنین، هنگام استفاده از نقشه‌های مفهومی به‌عنوان ابزارهای ارزشیابی، باید بار شناختی نقشه‌های مفهومی را نیز مد نظر داشت تا ارزیابی مؤثرتری از سطح دانش و درک یادگیرندگان به‌عمل‌آید و نتایج دقیق‌تر باشند، چرا که اضافه‌بار شناختی، همان‌گونه که قبلاً ذکر شد، مختل‌کنندهٔ جریان یادگیری و ارزشیابی است.

همان‌گونه که نتایج ما نشان داد، استفاده از آزمون‌های تکمیل نقشهٔ مفهومی می‌تواند راهکار مناسبی برای غلبه بر اضافه‌بار شناختی در جریان ارزشیابی باشد، چرا که این آزمون‌ها نسبت به آزمون‌های ترسیم نقشهٔ مفهومی بار شناختی کمتری را بر نظام شناختی آزمودنی‌ها تحمیل می‌کنند. لذا نتایج به‌دست‌آمده از پژوهش حاضر می‌تواند در طراحی آزمون‌های مبتنی بر اصول و قواعد شناختی کمک مناسبی باشد و بین آموزش و ارزشیابی با علوم شناختی پلی برقرار کند.

در واقع، عاریه‌گرفتن نظریات روان‌شناسی شناختی (مانند نظریهٔ بار شناختی) و واردکردن آن‌ها به داخل فرایند اندازه‌گیری آموزشی، امری ممکن و مطلوب است. گام‌های اولیه برای پل زدن شکاف بین علوم شناختی و ارزیابی آموزشی با شکل‌دهی چارچوب نظری به‌منظور ارزیابی تشخیصی شناختی[۷۷] شروع شد که

در آن ارزیابی آموزشی باید زیر «میکروسکوپ شناختی[۷۸]» قرار گیرد (لایتون و گایرل[۷۹]، ۲۰۰۷). این یعنی مسئولیت پژوهشگران آموزشی این است که روش‌ها، تکنیک‌ها و ابزارهای ارزشیابی آموزشی را با نظریه‌های شناختی وفق دهند و مهم‌تر از آن، برای دستیابی به اهداف ارزشیابی آموزشی، نظریه‌های شناختی را به‌صورت فعالانه خلق، اصلاح و آزمایش کنند.

روی هم رفته، نقشه‌های مفهومی روش‌های نویدبخشی برای ارزشیابی سطح دانش و تخصص یادگیرندگانی از تمام گروه‌های فرهنگی و تمام طبقات اقتصادی هستند، چرا که نسبت به سایر مقیاس‌ها درک عمیق‌تری از محتوا فراهم می‌آورند. زیرا در حین فرایند ارزشیابی، فراگیرندگان باید روش‌های سازمان‌دهی، تحلیل، فکرکردن و ارتباط‌دهی مفاهیم را نیز شرح دهند. در این نوع ارزشیابی، علاوه بر توانایی‌های تحلیلی، توانایی‌های عملی و خلاقیت نیز سنجیده می‌شوند. همچنین، بافت ارزشیابی‌ها نیز هنگام ساخت و نمره‌دهی نقشه‌های مفهومی با استفاده از بیان استانداردها و طرح‌های انتخاب مفاهیم مورد توجه قرار می‌گیرد (میکر و زیمرمان[۸۰]، ۲۰۲۰). بنابراین، تمام موارد فوق می‌تواند سرمایه‌گذاری‌های کلان در زمینهٔ استفاده از نقشه‌های مفهومی به‌عنوان ابزارهای ارزشیابی نوین را توجیه کند.

منابع

- احدی، فاطمه و سلیمانی، محسن. (۱۳۹۳). مقایسهٔ تأثیر دو روش تدریس به شیوهٔ ارائهٔ مثال آموزشی حل شدهٔ کامل و ارائهٔ مثال به شیوهٔ حل شدهٔ ناقص بر بار شناختی دانشجویان در درس زبان تخصصی پزشکی. *آموزش در علوم پزشکی*، ۱۴(۴)، ۲۹۱ – ۳۰۲.
- سیف، علی‌اکبر. (۱۳۸۷). *اندازه‌گیری، سنجش و ارزشیابی آموزشی* (چاپ بیست و هفتم). تهران: انتشارات دوران.
- محمدی، محسن.، مظلومی، عادل.، نسل‌سراجی، جبرائیل و زراعتی، حجت. (۱۳۹۲). طراحی پرسش‌نامهٔ ارزشیابی بار کاری ذهنی و تعیین روایی و پایایی آن در پرستاران بخش‌های ویژه یکی از بیمارستان‌های دانشگاه علوم پزشکی تهران. *دانشکده بهداشت و انستیتو تحقیقات بهداشتی*، ۱۱(۲)، ۸۷-۹۶.
- مصرآبادی، جواد. (۱۳۹۰). بررسی روایی و پایایی شیوه‌های نمره‌گذاری آزمون‌های مبتنی بر نقشهٔ مفهومی. *اندازه‌گیری تربیتی*، ۲(۵)، ۱-۲۴.
- Amadieu, F., Tricot, A., & Mariné, C. (2010). Interaction between prior knowledge and concept-map structure on hypertext comprehension, coherence of reading orders and disorientation. *Interacting with computers, 22*(2), 88-97. doi:10.1016/j.intcom.2009.07.001.
- Arneson, B. T. (2005). *On the role of concept mapping assessment in today□s constructivist classroom* (Doctoral dissertation, The University of Texas, Austin). Retrieved from https://repositories.lib.utexas.edu/handle/2152/23435
- Ausubel, D. P. (1960). The use of advance organizersin learning and retention of meaningful information. *Educational Psychology, 51*(5), 267-272.
- Berry, J. L. (2011). *The effects of concept mapping and questioning on students organization and retention of science knowledge while using interactive informational read-alouds* (Doctoral Dissertation, Texas A&M University). Retrieved from https://oaktrust.library.tamu.edu/handle/1969.1/ETD-TAMU-2011-08-10164.
- Chang, K-E., Sung, Y-T., Chang, R.-B., & Lin, S-C. (2005). A New Assessment for Computer-based Concept Mapping. *Educational Technology & Society, 8*(3), 138-148.
- Chang, K. E., Sung, Y. T., & Chen, S. F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Computer Assisted Learning, 17*, 21-33.
- Cline, B. E., Brewster, C. C., & Fell, R. D. (2010). A rule-based system for automatically evaluating student concept maps. *Expert systems with applications, 37*(3), 2282-2291.
- De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional science, 38*(2), 105-134.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. Hancock, and N. Meshkati (Eds), *Human Mental Workload* (Vol. 52, pp. 139-183). North Holland, Amsterdam: Elsevier Science Publishers B.V.
- Ho, V. W., Harris, P. G., Kumar, R. K., & Velan, G. M. (2018). Knowledge maps: a tool for online assessment with automated feedback. *Medical education online, 23*(1), 1457394.
- Kapuza, A. (2020). How Concept Maps with and without a List of Concepts Differ: The Case of Statistics. *Education Sciences, 10*(4), 91. https://doi.org/10.3390/educsci10040091
- Kinchin, I. M., Hay, D. B., & Adams, A. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational research, 42*(1), 43-57.
- Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science, 48*(1), 45-77.
- Leighton, J. P. & Gierl, M. J. (2007). *Cognitive diagnostic assessment for education: Theory and practices*. New York, NY: Cambridge University Press.
- Lim, Y. K. (2008). *The effect of concept mapping with different levels of generativity and learners self-regulated learning skills on knowledge acquisition and representation* (Doctoral dissertation, Pennsylvania State University). Retrieved from http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:3431422
- Lin, S. S. (2014). Cognitive load for configuration comprehension in computer-supported geometry problem solving: An eye movement perspective. *International Journal of Science and Mathematics Education, 12*(3), 605-627.
- Liu, T. C., Lin, Y. C., & Tsai, C. C. (2009). Identifying senior high school students, misconceptions about statistical correlation, and their possible causes: an exploratory study using concept mapping with interviews. *International*

Journal of Science and Mathematics Education. 7(4), 791-820.

- Maker, C. J., & Zimmerman, R. H. (2020). Concept maps as assessments of expertise: Understanding of the complexity and interrelationships of concepts in science. *Journal of Advanced Academics, 31*(3), 254-297.
- McClure, J. R., & Bell, P. E. (1990). *Effects of an Environmental Education-Related STS Approach Instruction on Cognitive Structures of Preservice Science Teachers*. (ERIC Number: ED341582).
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practically. *Research in Science Teaching, 36*(4), 475–492.
- Mesa, J. C. (2010). *The development and use of a concept mapping assessment tool with young children on family visits to a live butterfly exhibit* (Doctoral dissertation, University of Florida). Retrieved from https://ufdc.ufl.edu/UFE0042067/00001/pdf
- Mukhopadhyay, A., Sur, S., Ghosal, A., & Acharya, A. (2018). A New Score Generation Algorithm for Student Concept Map Evaluation. *Educational Technology Systems, 47*(2), 267-286.
- Novak, J. D., & Gowin, B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- Oda, K. (2011). *College students GIS spatial concept knowledge assessed by concept maps* (Doctoral dissertation. Texas A & M University). Retrieved from http://oaktrust.library.tamu.edu/handle/1969.1/ETD-TAMU-2011-05-9213
- Plummer, K. J. (2008). *Analysis of the psychometric properties of two different concept-map assessment tasks* (Doctoral dissertation, Brigham Young University). Retrieved from https://scholarsarchive.byu.edu/etd/1352
- Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An international review, 53*(1), 61-86.
- Ruiz-Primo, M. A. (2000). On the use of concept maps as an assessment tool in science: What we have learned so far. REDIE. *Revista Electrónica de Investigación Educativa, 2*(1), 29-53.
- Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 33*(6), 569-600.
- Rye, J. A., & Rubba, P. A. (2002). Scoring concept maps: An expert map-based scheme weighted for relationships. *School Science and Mathematics, 102*(1), 33-44.
- Saglam, Y. (2006). *A phenomenographic case study: concpt maps from the perspectives of midlle school students* (Doctoral dissertation, Purdue University, West Lafayette, Indiana). Retrieved from https://docs.lib.purdue.edu/dissertations/AAI3232233/
- Schau, C., & Mattern, N. (1997). Use of map techniques in teaching applied statistics courses. *The American Statistician, 51*(2), 171-175.
- Schwendimann, B. A. (2011). *Mapping biological ideas: Concept maps as knowledge integration tools for evolution education* (Ductoral Dissertation, University of California, Berkeley).Retrieved from https://infoscience.epfl.ch/record/211031/files/Dissertation%20Beat%20Schwendimann.pdf
- Schwendimann, B. A. (2015). Concept maps as versatile tools to integrate complex ideas: From kindergarten to higher and professional education. *Knowledge Management & E-Learning: An International Journal, 7*(1), 73-99.
- Skulmowski, A., & Rey, G. D. (2020). Subjective cognitive load surveys lead to divergent results for interactive learning media. *Human Behavior and Emerging Technologies, 2*(2), 149-157.
- Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. W. (2005). Windows into the mind. *Higher Education, 49*(4), 413-430.
- Snelson, L. J. (2010). *Estimating the Reliability of Concept Map Ratings Using a Scoring Rubric Based on Three Attributes of Propositions* (Doctoral dissertation, Brigham Young University). Retrieved from https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=3283&context=etd
- Stoddart, T., Abrams, R., Gasper, E., & Canaday, D. (2000). Concept maps as assessment in science inquiry learning: A report of methodology. *International Journal of Science Education, 22*(12), 1221-1246.
- Strautmane, M. (2012). concept map-based knowledge assessment tasks and their scoring criteria: an overview. In A. J. Cañas, J. D. Novak and J. Vanhear (Eds), *Concept Maps: Theory, Methodology, Technology Proceedings of the Fifth International Conference on Concept Mapping* (pp. 80-88). Valletta, Malta: University of Malta.

■ Thain, D. (2012). A study of semi-hierarchical organization in the construction of concept maps using the framework of cognitive load theory (Doctoral Dissertation, Queen's University, Kingston, Ontario, Canada). Retrieved from https://qspace.library.queensu.ca/handle/1974/7239.

■ Van Merriënboer, J. J. (1990). Strategies for programming instruction in high school: Program completion vs. program generation. *Journal of educational computing research, 6*(3), 265-285.

■ Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*(3), 261-312.

■ Vodovozov, V & Raud, Z. (2015). Concept Maps for Teaching, Learning, and Assessment in Electronics. *Education Research International*, Article ID 849678, 9 pages. http://dx.doi.org/10.1155/2015/849678.

■ West, D. C., Park, J. K., Pomeroy, J. R., & Sandoval, J. (2002). Concept mapping assessment in medical education: a comparison of two scoring systems. *Medical education, 36*(9), 820-826.

■ Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of Two Concept-Mapping Techniques: Implications for Scoring, Interpretation, and Use. *Research In Science Teaching., 42*(2), 166–184.

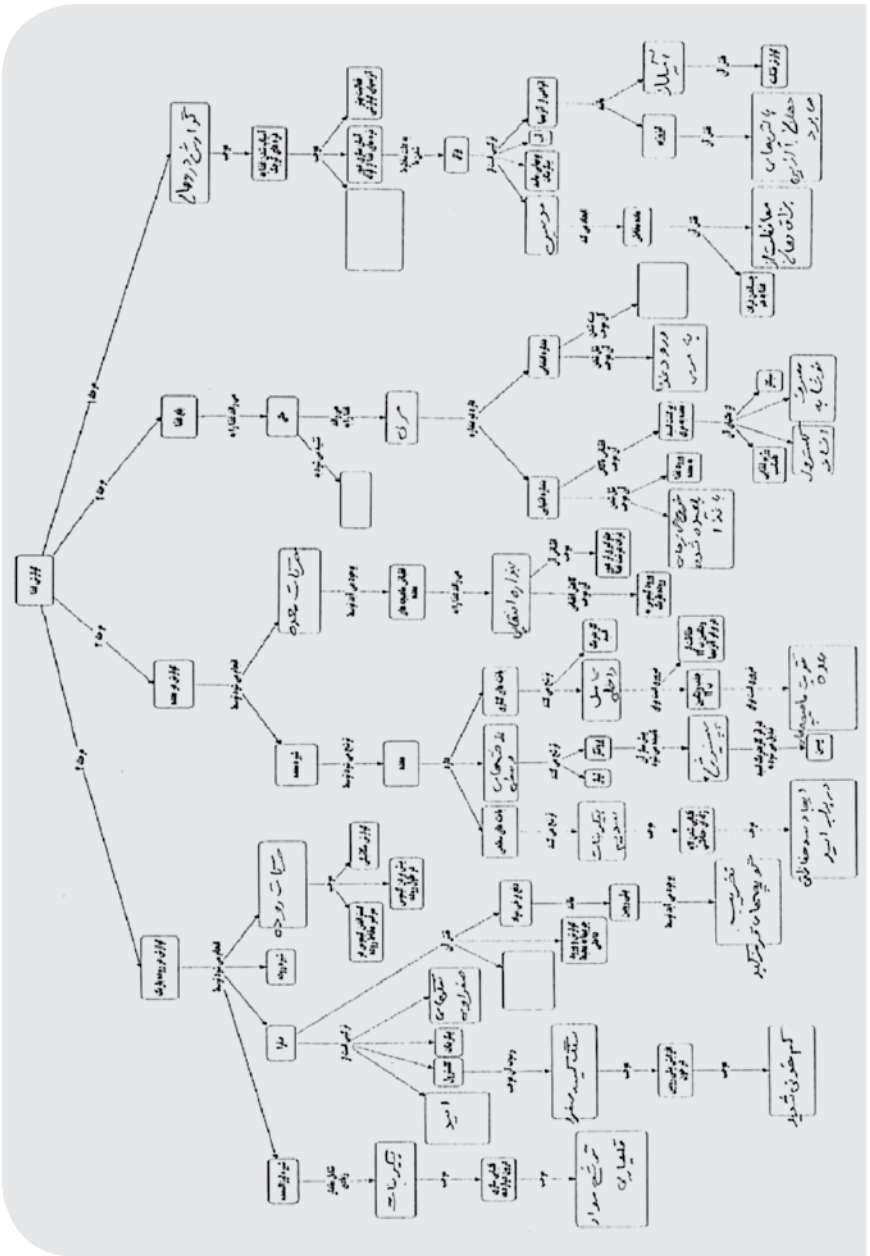
پی‌نوشت‌ها

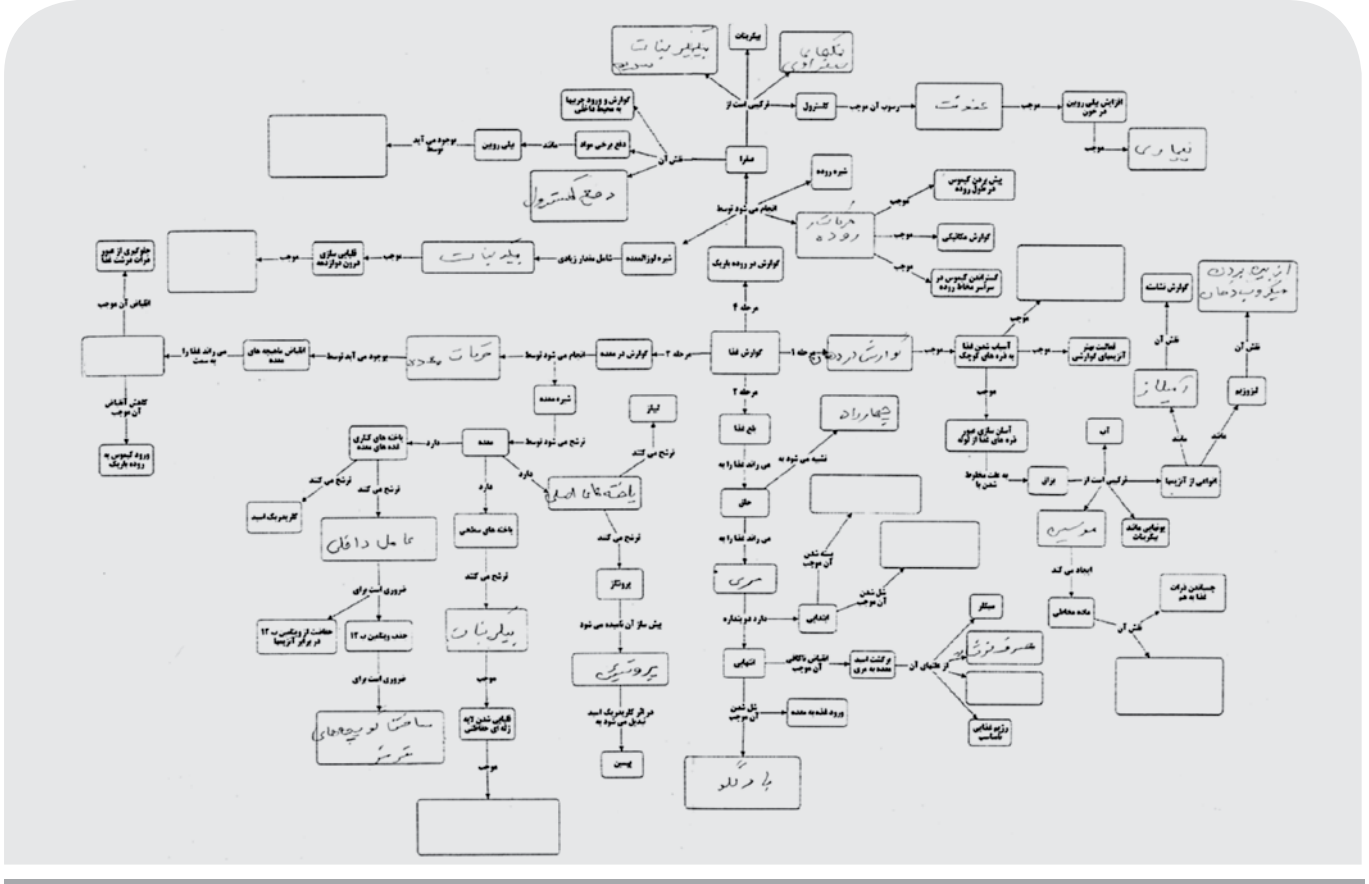
1. Schwendimann
2. Multiple-choice tests
3. Misconceptions
4. Liu, Lin, & Tsai
5. Constructivism
6. Arneson
7. Kapuza
8. Concept map-based evaluations
9. Stoddart, Abrams, Gasper & Canaday
10. Ho, Harris, Kumar & Velan
11. Ruiz-Primo, Schultz & Shavelson
12. Scoring method
13. Snelson
14. Strautmane
15. Vodovozov & Raud
16. Structural
17. Relational
18. Similariry to expert map
19. Novak & Gowin
20. McClure & Bell
21. Sonak & Suen
22. Mukhopadhyay, Sur, Ghosal & Acharya
23. Validity & Reliability
24. Cline,,Brewster & Fell
25. Hierarchy, network & chain maps
26. Amadieu, Tricot & Mariné
27. Cognitive Load Theory (CLT)
28. Thain
29. Berry
30. Sweller et al.
31. Working memory
32. De Jong
33. Skulmowski & Rey
34. Klepsch & Seufert
35. Lin & Lin
36. Vekiri
37. Extraneous cognitive load
38. Germane cognitive load
39. Statistical population
40. Multi-stage cluster sampling
41. NASA-Task Load Index (TLX)
42. Hart & Staveland
43. Human Performance Group
44. NASA-Ames Research Center
45. Mental Demands
46. Physical Demands
47. Temporal Demands
48. Performance
49. Effort
50. Frustration
51. Rubio, Díaz, Martín & Puente
52. Convergent validity, Concurrent validity & Discriminant validity
53. Subjective Workload Assessment
54. Workload Profile
55. Tracking Task & Memory Searching Task
56. Single and Dual tasks
57. West, Park, Pomeroy & Sandoval
58. Inter-rater reliability
59. One way ANOVA
60. Welch
61. Games-howell
62. Test of homogeneity of variances
63. Saglam
64. Chang, Sung, & Chen
65. Paas
66. Van Merrienboer
67. Lim
68. Schau & Nancy Mattern
69. Oda
70. Kinchin, Hay & Adams
71. Rye & Rubba
72. Plummer
73. Ausubel
74. Mesa
75. Chang, Sung, Chang, & Lin
76. Yin, Vanides
77. Cognitive Diagnostic Assessment (CDA(
78. Cognitive microscope
79. Leighton & Gierl
80. Maker & Zimmerman

■ پیوست‌ها ■

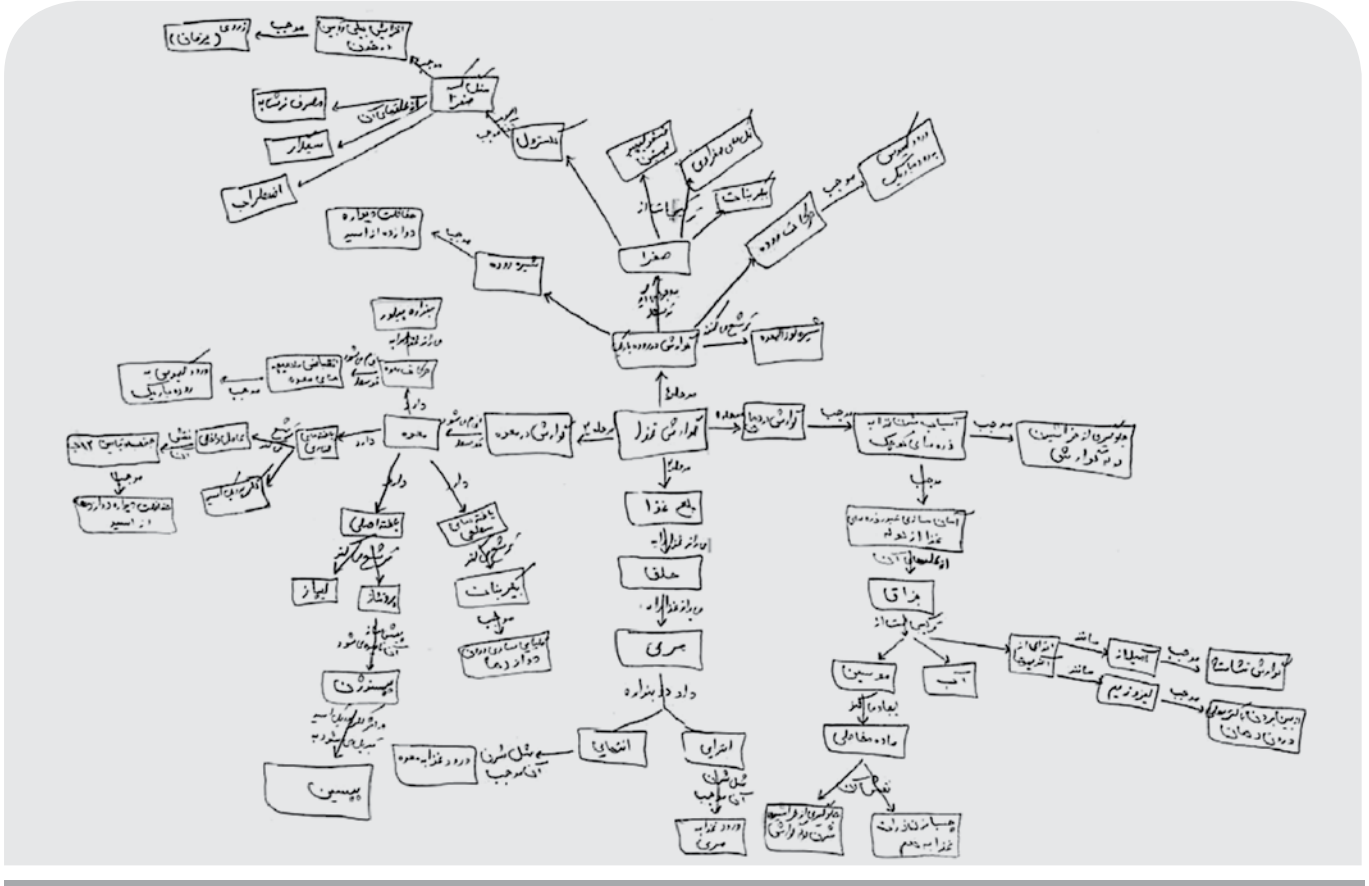
نمونه‌هایی از تکمیل و ترسیم آزمون‌های مبتنی بر نقشهٔ مفهومی توسط آزمودنی‌ها:

آزمون تکمیل نقشهٔ مفهومی سلسله‌مراتبی (کامل‌شده توسط یک دانش‌آموز)

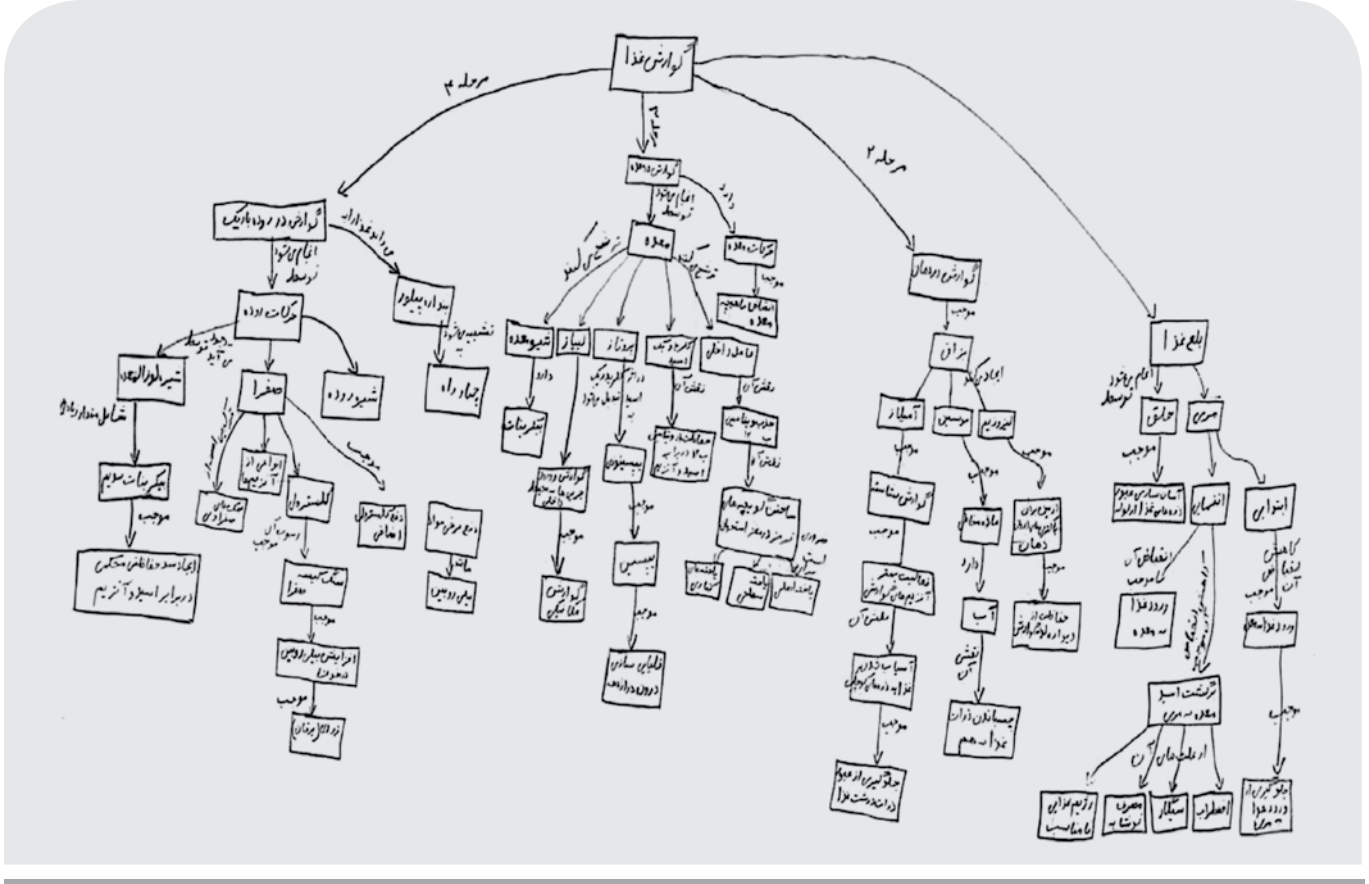




آزمون تکمیل نقشهٔ مفهومی سلسله‌مراتبی (کامل‌شده توسط یک دانش‌آموز)



آزمون تکمیل نقشهٔ مفهومی سلسله‌مراتبی (کامل‌شده توسط یک دانش‌آموز)



آزمون تکمیل نقشهٔ مفهومی سلسله‌مراتبی (کامل‌شده توسط یک دانش‌آموز)

Evaluation of cognitive loadand scoring procedures of concept map-based tests

■ Sādegh Sayyādi, PhD Candidate in Educational Psychology at Āzarbāijān Shahid Madani University[1]
■ Javād Mesrābādi (PhD), Āzarbāijān Shahid Madani University [2]
■ Abolfazl Farid (PhD), Āzarbāijān Shahid Madani University[3]
■ Rāmin Habibi Kaleibar (PhD), Āzarbāijān Shahid Madani University[4]

Abstract

Using concept map-based tests as assessment tools is one of the new assessment methods that can be an alternative for traditional assessment methods. So, the goal of the present study was evaluating cognitive load and scoring procedures of these tests. The research sample consisted of students of four biology classrooms from four different high schools (totally 109 students) who were selected via a multi-stage cluster sampling method. The research tools were a multiple-choice test (researcher-made), four concept map-based tests (researcher-made) and the NASA-TLX scale (Hart & Staveland, 1988). The study was done with some type of evaluation design. The statistical methods included Pearson's correlation coefficient, independent samples t-test and one-way ANOVA (Welch test and Games-Howell post hoc). Findings showed that construct-a-map tests imposed significantly more cognitive load on the subjects than fill-in-a-map tests. Also, relational, structural and similarity to expert map scoring procedures had acceptable and significant criterion validity, discriminant validity and reliability. All in all, relational, structural and similarity scoring procedures can be used appropriately for scoring the concept map-based tests. Also, using the fill-in-a-map tests can be an appropriate strategy to overcome the cognitive overload during the concept map-based tests. Thus, the results of the present research can be useful to design tests based on the cognitive rules and to bridge the gap between academic evaluation and cognitive sciences.

Keywords **Cognitive Load, Scoring Procedures, Concept Map- Based Tests**

1. Sadeghsayadi90@gmail.com **2.** Email: mesrabadi@azaruniv.ac.ir **(Corresponding Author)** **3.** abolfazlfarid@gmail.com
4. habibikaleibar@gmail.com