

## طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی Designing, Implementing, and Validating the Concept Map-Based Tests

تاریخ دریافت مقاله: ۲۵/۰۸/۹۲؛ تاریخ پذیرش مقاله: ۲۰/۰۷/۹۳

J. Mesrabadi (Ph.D)

**Abstract:** Evaluation of learners' cognitive structure is a growing and important part of academic evaluation research. One of the most commonly used methods to measure cognitive structure is concept map-based tests. The purpose of this research is to design, implement, and validate the concept map-based tests in Persian textbooks. First, four types of concept map-based tests (free drawing, the leaded drawing, the free complete, the leaded complete) were presented. The tests were implemented among a sample of 221 students of secondary school students that participated in eight experimental classes in Tabriz, Iran. In the next step, the tests were validated these based on indexes of content validity, criterion validity, and differential validity and inter raters' reliability. The qualitative findings indicated that the Persian textbooks include the potential for concept map-based tests. The findings also showed that these tests have a high content validity. The results of these tests have different values of correlation (average value of .50) with multiple-choice tests (criterion validity). Furthermore, the tests detected a difference between high achievers students and low achievers ones (differential validity). According to the obtained results, the coefficients of scoring agreement in concept maps tests were high (inter raters' reliability).

**Keywords:** test, concept maps, educational assessment, concept map-based test, validity, reliability

جواد مصرآبادی<sup>۱</sup>

**چکیده:** امروزه ارزشیابی از ساختار شناختی فراگیران بخش مهم و رو به گسترشی از پژوهش‌های سنجش تحصیلی را شامل می‌شود. یکی از پرکاربردترین شیوه‌های سنجش ساختار شناختی، آزمون‌های مبتنی بر نقشه مفهومی هستند. هدف از تدوین این مقاله طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی در متن‌های درسی فارسی است. در این پژوهش ابتدا چگونگی طرح چهار نوع آزمون مبتنی بر نقشه مفهومی (ترسیم آزاد، ترسیم هدایت شده، تکمیل آزاد و تکمیل هدایت شده) تشریح می‌شود. سپس به چگونگی اجرای آزمون‌ها پرداخته می‌شود که در این پژوهش این آزمون‌ها در بین یک نمونه ۲۲۱ نفری از دانش‌آموزان سال دوم متوسطه رشته علوم تجربی شهر تبریز که در هشت کلاس حضور داشتند، اجرا شد. در گام آخر در مورد نتایج اعتباربخشی این آزمون‌ها بر اساس شاخص‌های روایی محتوایی، روایی ملاکی، روایی افتراقی و پایایی نمره‌گذاران ارائه می‌شود. یافته‌های کیفی پژوهش نشان دادند که طراحی و اجرای آزمون‌های مبتنی بر نقشه مفهومی در متن‌های درسی فارسی امکان‌پذیر می‌باشد. یافته‌های کمی نیز نشان دادند که این آزمون‌ها از روایی محتوایی بالایی برخوردار هستند. همچنین نمرات حاصل از این نوع آزمون‌ها دارای مقادیر متفاوت همبستگی (با مقدار متوسط ۰/۵۰) با یک آزمون چهارگزینه‌ای بودند (روایی ملاکی). همچنین همه آزمون‌های نقشه مفهومی به‌خوبی از عهده تشخیص تفاوت بین نمرات دو گروه تحصیلی بالا و پائین برآمدند (روایی افتراقی). دیگر نتایج نشان دادند که ضریب توافق نمره‌گذاران آزمون‌های نقشه مفهومی در حد بالایی است (پایایی نمره‌گذاران).

**کلیدواژه‌ها:** آزمون، نقشه مفهومی، ارزشیابی تحصیلی، آزمون‌های مبتنی بر نقشه مفهومی، روایی، پایایی

ارزشیابی تحصیلی یک بخش ضروری برای هر موقعیت آموزشی برنامه‌ریزی شده است که بدون آن فرایند یاددهی-یادگیری کامل نمی‌شود. نوواک<sup>۱</sup> (۲۰۰۵) نقشه مفهومی<sup>۲</sup> را به‌عنوان ابزاری برای ارزشیابی تحصیلی معرفی کرده است. یک نقشه مفهومی متشکل از هسته‌ها و پیوندها است. در هسته‌ها یک مفهوم، عبارت و یا سؤال قرار می‌گیرد که از طریق رابطه‌ها به سایر هسته‌ها وصل می‌شود. جاکوبس - لاوسون و هرشی<sup>۳</sup> (۲۰۰۲) معتقدند که نقشه مفهومی نه تنها می‌تواند به‌عنوان شیوه‌ی ارزشیابی جایگزین و مکمل به کار گرفته شود، بلکه همچنین برتری‌های نقشه مفهومی به‌عنوان یک ابزار ارزشیابی نیز مشخص شده است. از محاسن ارزشیابی به کمک نقشه مفهومی می‌توان به امکان ارزشیابی توانایی بازسازی اطلاعات، ارزشیابی ساختار شناختی، ارزشیابی یادگیری‌های سطح بالا و ارزشیابی برای اهداف مختلف (آغازین، تکوینی و تراکمی) نام برد (روئیر-پریمیو و شیولسون<sup>۴</sup>، ۱۹۹۶).

اما برخلاف روش‌های مرسوم ارزشیابی تحصیلی که در قالب‌های محدود و مشخصی اجرا و نمره‌گذاری می‌شوند؛ شیوه‌های ارزشیابی نقشه مفهومی از تنوع بسیار زیادی هم از لحاظ شکل آزمون و هم از لحاظ روش نمره دهی برخوردار هستند. با توجه به گستردگی و جدید بودن (البته در داخل کشور) حوزه ارزشیابی مبتنی بر نقشه مفهومی لازم است که عملی‌ترین و در عین حال مؤثرترین شیوه ارزشیابی مبتنی بر نقشه مفهومی و نیز مناسب‌ترین شیوه نمره‌گذاری این آزمون‌ها بررسی شود. هرچند که چنین بررسی در این حوزه در متن‌های درسی غیرفارسی قبلاً انجام شده است (ویست، پامروی و ساندول<sup>۵</sup>، ۲۰۰۲؛ شیولسون و همکاران، ۲۰۰۵؛ ویست و همکاران، ۲۰۰۰)، اما بایستی در نظر داشت که همچون استفاده از نقشه مفهومی به‌عنوان یک راهبرد یادگیری که نوع زبان مورد استفاده می‌تواند بر اثربخش بودن آن مؤثر باشد (کلیس<sup>۶</sup>، ۲۰۰۳)، ضروری به نظر می‌رسد که مناسب‌ترین شیوه ارزشیابی مبتنی بر نقشه مفهومی در متن‌های درسی فارسی نیز مورد بررسی قرار گیرد. اهمیت عملی پژوهش حاضر در ارائه

- 
1. Novak
  2. Concept map
  3. Jacobs-Lawson & Hershey
  4. Ruiz-Primo & Shavelson
  5. West, Pomeroy & Sandoval
  6. Kilic

توضیحات کافی به همراه نمونه‌ای از آزمون‌های مبتنی بر نقشه مفهومی است که به مدرسان و برنامه ریزان در تهیه این‌گونه ابزارها می‌تواند کمک کند. از لحاظ نظری نیز پژوهش حاضر به گسترش دانش نظری در مورد شاخص‌های روایی و پایایی آزمون‌های مبتنی بر نقشه مفهومی در متن‌های درسی فارسی منجر می‌شود.

به‌طور خلاصه در این پژوهش محقق در تلاش بود به شش سؤال پژوهشی پاسخ دهد. دو سؤال اول در یک قالب کیفی مطرح و مورد آزمون قرار گرفتند که عبارت بودند از: (۱) آیا برای متن‌های درسی فارسی می‌توان آزمون‌های مبتنی بر نقشه مفهومی طراحی کرد؟، (۲) آیا آزمون‌های نقشه مفهومی در کلاس‌های درسی داخل کشور قابل کاربست هستند؟ سؤال اول به طراحی آزمون‌های مبتنی بر نقشه مفهومی مربوط می‌شود. لزوم انجام این کار همان‌گونه که ذکر شد به‌تازگی این شیوه‌های ارزشیابی در حوزه ارزشیابی تحصیلی داخل کشور مربوط می‌شود. سؤال دوم نیز به امکان اجرای این آزمون‌ها در کلاس‌های درس داخل کشور مربوط می‌شود؛ لزوم بررسی این موضوع نیز از این رو است که این آزمون‌ها قالب‌هایی کاملاً متمایز با آزمون‌های مرسوم دارند که لازم است پیش از اجرای آزمون‌ها به دانش‌آموزان نحوه پاسخ‌دهی به آن‌ها را آموزش داد. اما چهار سؤال پژوهشی دیگر به اعتباربخشی به نمرات حاصل از این ابزارها مربوط می‌شد که در این سؤالات روایی محتوایی<sup>۱</sup>، روایی ملاکی<sup>۲</sup>، روایی افتراقی<sup>۳</sup> و پایایی نمره‌گذاران<sup>۴</sup> آزمون‌های مبتنی بر نقشه مفهومی مورد بررسی قرار گرفت. در ادامه به توضیحات تکمیلی در مورد طراحی، اجرا و اعتبار بخشی به آزمون‌های مبتنی بر نقشه مفهومی پرداخته می‌شود.

### طراحی آزمون‌های مبتنی بر نقشه مفهومی

در این پژوهش برای طراحی آزمون‌های مبتنی بر نقشه مفهومی از تقسیم‌بندی روئیر-پریمبو و شیولسون (۱۹۹۶) برای انواع آزمون‌های مبتنی بر نقشه مفهومی استفاده شد. دلیل انتخاب تقسیم‌بندی این محققین آثار متعدد ایشان در حوزه ارزشیابی نقشه مفهومی است. با مروری به

1. Content validity
2. Criterion validity
3. Discriminative validity
4. Interrator reliability

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی

مقالات حوزه ارزشیابی مبتنی بر نقشه مفهومی متوجه تعدد نظرات و کارهای ماریا آراسلی روئیز-پریمو و ریچارد شیولسون از اساتید دانشگاه استنفورد می‌شویم. روئیز-پریمو و شیولسون (۱۹۹۶) ارزشیابی مبتنی بر نقشه‌های مفهومی را متشکل از سه بخش می‌دانند: (۱) تکلیفی که از دانش‌آموز می‌خواهد که شواهدی از ساختار دانش خود در یک زمینه را مشخص کند، (۲) یک شکل پاسخ‌دهی دانش‌آموز و (۳) یک نظام نمره‌گذاری که به وسیله آن می‌توان نقشه مفهومی دانش‌آموز را به‌طور دقیق و پایایی ارزیابی کرد. بدون این سه مؤلفه نقشه مفهومی را نمی‌توان به‌عنوان یک ابزار ارزشیابی مورد توجه قرار داد. این ویژگی‌ها موجب تنوع شیوه‌های ارزشیابی مبتنی بر نقشه مفهومی می‌شود. از این رو بر اساس این سه مؤلفه ذکر شده می‌توان شیوه‌های مختلفی برای ارزشیابی مبتنی بر نقشه مفهومی ارائه کرد. درک شیوه‌های مختلف ارزشیابی نقشه مفهومی زمانی کامل می‌شود که به روابط بین سه مؤلفه ارزشیابی بیشتر توجه کنیم. در جدول ۱ چارچوبی پیشنهاد شده است که جنبه‌های مختلف ماهیت تکلیف، شیوه پاسخ‌دهی و نظام نمره‌گذاری را مورد توجه قرار می‌دهد. تصمیماتی که در مورد تکلیف نقشه مفهومی و نظام نمره‌گذاری گرفته می‌شود، مستقیماً به مؤلفه‌های نقشه مفهومی - اینکه کدام مؤلفه‌ها و با چه ویژگی‌هایی فراهم شده باشند- وابسته است.

جدول ۱: مؤلفه‌های ارزشیابی‌های مبتنی بر نقشه مفهومی (اقتباس از روئیز-پریمو و شیولسون، ۱۹۹۶)

مؤلفه‌های ارزشیابی	متغیرها	مثال‌ها
تکلیف	خواسته‌های تکلیف	از دانش‌آموزان می‌توان خواست که: یک نقشه را تکمیل کنند. یک نقشه بسازند.
		روابط بین یک جفت از مفاهیم را برقرار کنند.
		یک انشاء بنویسند.
		به یک مصاحبه پاسخ دهند.

مؤلفه‌های ارزشیابی	متغیرها	مثال‌ها
	محدودیت‌های تکلیف	دانش‌آموزان ممکن است که: مفاهیمی تکلیف در اختیار داشته باشند. مفاهیم رابطه‌ها را در اختیار داشته باشند. اجازه به استفاده بیشتر از یک رابطه بین مفاهیم داشته باشند. اجازه جابجایی مفاهیم را داشته باشند. نقشه را به‌طور گروهی تهیه کند.
پاسخ	نحوه پاسخ	پاسخ دانش‌آموزان به کدام نحوه است: قلم - کاغذی شفاهی کامپیوتری
	ترسیم‌کننده	چه کسی نقشه را می‌کشد: فراگیر معلم یا محقق
نظام نمره دهی	مؤلفه‌های نمره دهی	سه مؤلفه زیر یا تعدادی از آن‌ها می‌توانند نمره‌گذاری شوند: گزاره‌ها سطوح سلسله‌مراتبی مثال‌ها
	استفاده از یک ملاک	نقشه ملاک می‌تواند متعلق باشد به: یک یا چند متخصص از حوزه مورد آزمون یک یا چند معلم یک یا چند دانش‌آموز برتر

تکلیف خواسته شده: در مورد تکلیفی که دانش‌آموز در ارزشیابی نقشه مفهومی بایستی انجام دهد غالباً یک ارزشیابی مبتنی بر نقشه مفهومی در گستره‌ای از هدایت‌شدگی یعنی از هدایت بالا تا هدایت پائین بر اساس میزان اطلاعاتی که در اختیار دانش‌آموز قرار داده می‌شود، مشخص می‌شود. تکلیف‌های نقشه مفهومی که از هدایت بالایی برخوردار هستند، مفاهیم، خطوط ارتباطی، عبارت‌های ربط دهنده و ساختار نقشه را در اختیار دانش‌آموزان قرار می‌دهند. در مقابل

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی

در یک تکلیف نقشه مفهومی با هدایت پائین فراگیران برای تصمیم‌گیری در این زمینه‌ها که چه تعداد از مفاهیم را در نقشه‌شان بگنجانند، کدام مفاهیم را به هم ربط دهند، و چه کلماتی را برای تشریح روابط به کار گیرند آزاد هستند. با توجه به درجه هدایت‌شدگی روئیر-پریمیو و شیولسون (۱۹۹۶) پنج روش ترسیم نقشه مفهومی در گستره‌ای از تکلیف ساخت یک نقشه کامل تا تکمیل یک نقشه معرفی کرده‌اند.

**شکل پاسخ‌دهی:** این جنبه از آزمون‌های نقشه مفهومی به‌طور مستقیم به محدودیت‌های ناشی از تکلیف بستگی دارد. برای مثال اگر ساختار نقشه ارائه شده باشد، صرفاً آزمودنی کلمات یا عبارات را در جاهای خالی می‌گذارد (تکمیل نقشه). اگر مفاهیم مهم داده شده باشد؛ لازم است که فراگیر بر اساس آن‌ها به طراحی نقشه مفهومی اقدام کند (ترسیم نقشه). لازم به یادآوری مجدد است که این شیوه در این پژوهش با توجه به تقسیم‌بندی روئیر-پریمیو و شیولسون (۱۹۹۶) از دو شیوه اصلی ترسیم نقشه مفهومی و تکمیل نقشه مفهومی استفاده شد. در ادامه با ارائه یا عدم ارائه مفاهیم کلیدی محتوای مورد سنجش در هر کدام از این دو شیوه چهار نوع آزمون مبتنی بر نقشه مفهومی طراحی شد. این شیوه‌ها عبارت بودند از: ۱) آزمون تکمیل نقشه مفهومی بر اساس مفاهیم داده شده، ۲) آزمون تکمیل نقشه مفهومی آزاد، ۳) آزمون ترسیم آزاد نقشه مفهومی و ۴) آزمون ترسیم نقشه مفهومی بر اساس مفاهیم داده شده که توضیح مختصری از این آزمون‌ها در ادامه ارائه می‌شود.

**آزمون تکمیل نقشه مفهومی ناقص بر اساس مفاهیم داده شده:** در این شیوه یک نقشه مفهومی ناقص از قبل ترسیم شده توسط محقق در اختیار آزمودنی‌ها قرار می‌گیرد. در این پژوهش این نقشه از محتوای کتاب زیست‌شناسی علوم تجربی از مبحث قلب تهیه شده بود. آزمودنی‌ها می‌توانستند هر کلمه یا عبارتی را برای تکمیل بخش‌های ناقص نقشه همچون مفاهیم داخل هسته‌ها، مطالب روی خطوط ارتباطی یا ارتباط‌های ناقص به کار گیرند آزمودنی‌ها برای تکمیل نقشه مفهومی ناقص بایستی از مفاهیمی استفاده می‌کردند که توسط محقق در اختیارشان قرار داده شده بود (شکل شماره ۱).

**آزمون تکمیل آزاد نقشه مفهومی ناقص:** این شیوه کاملاً شبیه به آزمون قبلی است با این تفاوت که مفاهیمی که بایستی آزمودنی از آن‌ها برای تکمیل نقشه استفاده نماید؛ در اختیار وی

قرار داده نمی‌شود. آزمودنی در این آزمون‌ها هم بایستی مفاهیم اصلی محتوا را به یاد بیاورد و هم این که آن‌ها را در جاهای مناسب قرار دهد (شکل شماره ۲).

**آزمون ترسیم آزاد نقشه مفهومی:** در این شیوه آزمودنی‌ها به‌طور آزادانه برای یک مطلب نقشه مفهومی ترسیم می‌کردند. برای این منظور یک برگه A4 با مداد و پاک‌کن در اختیار آن‌ها قرار داده شد و از آن‌ها خواسته شد برای موضوعی که در وسط صفحه نوشته شده است یک نقشه مفهومی ترسیم کنند. موضوع انتخاب شده برای این آزمون مبحث قلب بود که در داخل کادری در وسط صفحه نوشته شده بود. ابتدا دانش‌آموزان بایستی مفاهیم اصلی حوزه مورد آزمون را به یاد بیاورند و در گام بعدی با گسترش موضوع به حاشیه‌های صفحه به ترسیم نقشه مفهومی قلب می‌پرداختند (شکل شماره ۳).

**آزمون ترسیم نقشه مفهومی بر اساس مفاهیم داده شده:** این شیوه مانند شیوه بالا بود با این اختلاف که در این شیوه فهرستی از مفاهیم و رابطه‌های مهم مربوط به مبحث قلب به آزمودنی‌ها داده شد و از آن‌ها خواسته شد که در ترسیم نقشه مفهومی قلب از این مفاهیم و روابط استفاده نمایند (شکل شماره ۴).

**نمره‌گذاری:** پس از اجرای آزمون‌های مبتنی بر نقشه مفهومی لازم است که آن‌ها نمره‌گذاری شوند. تصمیمات مرتبط با نظام نمره دهی به‌طور مستقیم تحت تأثیر دو مؤلفه نوع تکلیف و شکل پاسخ‌دهی است (مصرا‌آبادی، ۱۳۸۹). در سال‌های اخیر روش‌های مختلفی برای نمره‌گذاری آزمون‌های مبتنی بر نقشه مفهومی ارائه شده است. با توجه به مبانی تجربی حوزه ارزشیابی نقشه مفهومی می‌توان کلیه شیوه‌های نمره دهی را در دو قالب کمی (نواک و کوئین، ۱۹۸۴؛ مارکهام، میتنز و جونز، ۱۹۹۴؛ ویست و همکاران، ۲۰۰۲؛ مک کلور، سوناک و سئون، ۱۹۹۹) و کیفی (هرل و باکر، ۱۹۹۶؛ چنگ، سانگ و چن، ۲۰۰۱) گنجانند.

در روش‌های کمی بر تخصیص یک نمره به یک نقشه مفهومی تأکید می‌شود و این کار معمولاً از طریق نمره دهی به بخش‌های مختلف نقشه صورت می‌گیرد. شیولسون و همکاران (۲۰۰۵) نیز در یک جمع‌بندی، حداقل چهار ملاک برای نمره‌گذاری نقشه‌های مفهومی

---

1. Markham, Mintzes & Jones

2. McClure, Sonak & Suen

3. Herl & Baker

4. Chang, Sung & Chen

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی

ذکر کرده‌اند: (۱) شاخص‌های کمی ویژگی‌هایی مانند تعداد گزاره‌ها، (۲) شاخص‌های ساختاری مانند تعداد سطوح سلسله مراتبی که می‌توان در روابط ایجاد شده کشف کرد، (۳) شاخص‌های صحت، که روایی یک گزاره را نشان می‌دهد، (۴) شباهت به یک نقشه مفهومی متخصص. سه روش اول پیشنهادی شیولسون و همکاران (۲۰۰۵) جزء روش‌های کمی و روش آخر از شیوه‌های نمره‌گذاری کیفی است. کپینز و هی<sup>۱</sup> (۲۰۰۸) روش‌های نمره دهی کمی را با عنوان نمره دهی وزنی به بخش‌های مختلف نقشه<sup>۲</sup> نامیده‌اند. در روش‌های نمره دهی وزنی نمرات نسبی را به مفاهیم و یا رابطه‌های بین مفاهیم اختصاص می‌دهند. نمره داده شده به یک نقشه مفهومی برابر است با مجموع نمره‌های نسبی به هر بخش از این نقشه. میزان نمره اختصاصی به هر بخش بستگی به اعتبار آن بخش و یا نوع ساختاری دارد که به کل نقشه می‌بخشد. همچنین ویست و همکاران (۲۰۰۲) دو روش معمول نمره‌گذاری کمی نقشه‌های مفهومی را روش ساختاری<sup>۳</sup> و روش رابطه‌ای<sup>۴</sup> معرفی کرده‌اند. در روش ساختاری براساس سازمان، ساختار و ارتباطات موجود بین مفاهیم نمرات داده می‌شود. در مقابل، در روش ارتباطی بر کیفیت یا اهمیت تک‌تک هر کدام از رابطه‌های مفهومی تأکید می‌شود، بدون این‌که به ساختار کلی نقشه مفهومی توجه شود.

روش نمره دهی ساختاری که به وسیله نواک و کوئین (۱۹۸۴) ابداع شده است، به دانش ساختاری سلسله مراتبی اهمیت می‌دهد. نواک و کوئین (۱۹۸۴) برای نمره‌گذاری نقشه‌های مفهومی چهار ملاک سطوح سلسله مراتبی، اعتبار گزاره‌ها، همبرها و استفاده از مثال‌ها پیشنهاد داده‌اند. در این روش یک نمره نسبتاً پایینی به هر رابطه و مثال معتبر از یک مفهوم داده می‌شود. در عوض نمره بالایی به رابطه‌هایی داده می‌شود که بیانگر یک ارتباط سلسله مراتبی هستند (همچون روابط "هست نوعی از" یا "شامل"). بالاترین نمره به رابطه‌هایی داده می‌شود که مفاهیمی قرار گرفته شده در شاخه‌های مختلف یک ساختار سلسله مراتبی را به هم ربط می‌دهد. در روش نمره‌گذاری ساختاری نواک و کوئین ۱ نمره به هر رابطه معتبر داده می‌شود، ۴ نمره به سلسله مراتب داده می‌شود، ۱۰ نمره به هر همبر داده می‌شود و ۱ نمره به هر مثال داده می‌شود.

- 
1. Keppens & Hay
  2. Weighted component scoring methods
  3. Structural
  4. Relational



مارکهام و همکاران (۱۹۹۴) نیز نقشه‌های مفهومی را با یک ویرایش تغییر یافته‌ای از روش ساختاری نوک و کوئین (۱۹۸۴) نمره‌گذاری کردند. آن‌ها شش جنبه از نقشه‌های مفهومی دانش‌آموزان را نمره‌گذاری کردند: (۱) تعداد مفاهیم، به‌عنوان نشانی از دانش فرد از حوزه موردنظر؛ (۲) روابط مفاهیم؛ که فراهم کننده شواهد اضافی از دانش فرد در حوزه مورد نر بود؛ (۳) شاخه‌ها؛ که آن‌ها به‌عنوان شاهدهی از تمییز پیش‌رونده محسوب می‌شدند. (۴) سلسله‌مراتب، که شواهدی برای دانش مشمول سازی<sup>۱</sup> فراهم می‌آوردند. (۵) همبرها، که نشان دهنده شواهدی از همبسته بودن دانش بودند. (۶) مثال‌ها، که نشان دهنده مشخصات حوزه دانش بودند.

در شیوه‌های کیفی بیشتر اهداف یاددهی - یادگیری در نظر هستند تا اهداف ارزشیابی. شیوه‌های کیفی نمره دهی غالباً مشتمل بر مقایسه نقشه فراگیر با یک نقشه کامل شده متخصص به‌صورت کلی است. نوک و کوئین (۱۹۸۴) علاوه بر روش‌های وزنی پیشنهاد داده‌اند که معلمان یک نقشه ملاک از موضوع تهیه کنند تا بتوان نقشه‌های دانش‌آموزان را با آن مقایسه کرد؛ درجه شباهت بین این نقشه‌ها می‌تواند به‌صورت یک نمره درصدی داده شود. در این زمینه چنگ، سانگ و چن<sup>۲</sup> (۲۰۰۱) از شاخصی به‌عنوان شاخص شباهت<sup>۳</sup> استفاده کرده‌اند. شاخص شباهت نشان دهنده میزان شباهت بین نقشه مفهومی فراگیران به نقشه مفهومی تهیه شده توسط یک فرد متخصص است. شاخص شباهت، شاخصی از درستی و کامل بودن نقشه مفهومی فراگیر است. مک کلور، سوناک و سوان (۱۹۹۹) نیز در پژوهش خود روش ارزشیابی کلی<sup>۴</sup> را پیشنهاد کرده‌اند. در روش نمره دهی کلی نمره‌گذاران آموزش می‌بینند که به یک نقشه مفهومی نمره‌ای از یک دامنه به‌عنوان مثال صفر تا ۱۰ تخصیص دهند که این نمره نشانگر فهم کلی طراح نقشه از یک بعد است.

### اجرای آزمون‌های نقشه مفهومی

هدف از این مرحله امکان سنجی اجرای آزمون‌های مبتنی بر نقشه مفهومی در کلاس‌های درس بود. همانگونه که پیشتر هم به‌نوعی اشاره شد؛ لزوم انجام این کار از این جهت بود که

---

1. Subsumption

2. Chang, Sung & Chen

3. Similarity Index

4. Holistic Method

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی

آزمون‌های مبتنی بر نقشه مفهومی شیوه‌هایی کاملاً متفاوت ارزشیابی هم برای فراگیران و هم برای معلمان آن است. محقق در پژوهش‌های پیشین امکان‌سنجی کاربست نقشه‌های مفهومی را به‌عنوان راهبردهای یاددهی و یادگیری (مصرآبادی، فتحی آذر و استوار، ۱۳۸۴؛ مصرآبادی، علوی و استوار، ۱۳۸۵؛ مصرآبادی، حسینی نسب، فتحی آذر و مقدم واحد، ۱۳۸۶؛ مصرآبادی و استوار، ۱۳۸۸؛ مصرآبادی، الف ۱۳۹۰) مورد بررسی قرار داده بود. تجارب پیشین نشان می‌دهد که در هر دو نوع شرایط در ابتدای کاربندی مقاومت‌هایی از طرف فراگیران و معلمان در مورد این شیوه وجود دارد. برای بررسی امکان کاربست نقشه‌های مفهومی لازم بود آزمون‌های مبتنی بر نقشه مفهومی بر روی دانش‌آموزان مدارس عادی و در ارتباط با متن‌های درسی اجرا شود. ترتیبات اجرایی این بخش در ادامه ارائه شده است.

**آزمودنی‌ها:** آزمون‌های مبتنی بر نقشه مفهومی بر روی یک نمونه ۲۲۱ نفری از دانش‌آموزان کلاس دوم تجربی شهر تبریز در سال تحصیلی ۸۸-۱۳۸۷ اجرا شد که به روش نمونه‌گیری خوشه‌ای چندمرحله‌ای<sup>۱</sup> انتخاب شده بودند. این دانش‌آموزان در هشت کلاس از قبل تشکیل شده حضور داشتند که این کلاس‌ها به صورت دوتایی به‌عنوان چهار گروه ترسیم آزاد نقشه مفهومی، ترسیم هدایت شده، تکمیل آزاد و تکمیل هدایت شده در نظر گرفته شدند.

**ابزارهای پژوهش:** در این پژوهش برای اندازه‌گیری میزان پیشرفت تحصیلی دانش‌آموزان از چهار نوع آزمون مبتنی بر نقشه مفهومی (ترسیم آزاد نقشه مفهومی، ترسیم هدایت شده، تکمیل آزاد و تکمیل هدایت شده) و یک آزمون چهارگزینه‌ای به‌عنوان ملاک استفاده شد. محتوای آزمون‌های استفاده شده در پژوهش از متن‌های درسی زیست‌شناسی سال دوم رشته علوم تجربی انتخاب شدند.

**طرح پژوهش و روش اجرا:** برای اعتباربخشی به آزمون‌های مبتنی بر نقشه مفهومی از پژوهش ارزشیابی<sup>۲</sup> استفاده شد. طبق نظر روسی<sup>۳</sup> (به نقل از سیف، ۱۳۸۲) این دسته از پژوهش‌ها نوعی ارزشیابی هستند که در آن مسائل مربوط به روایی درونی، روایی بیرونی، روایی سازه منظور می‌شود. این پژوهش همچنین به شیوه درون‌گروهی انجام پذیرفت. درونی گروهی بودن طرح از این جهت است که افراد داخل هر گروه به دو شیوه اندازه‌گیری شدند (آزمون چهارگزینه‌ای و

- 
1. Multistage cluster sampling
  2. Evaluation research
  3. Rossi

یکی از شیوه‌های ارزشیابی مبتنی بر نقشه مفهومی) و این اندازه‌ها باهم مورد مقایسه قرار گرفتند. در این پژوهش چهار گروه آزمایشی (هر گروه شامل دو کلاس) حضور داشتند. جزئیات اجرایی پژوهش در گروه‌های آزمایشی در قالب چهار مرحله به شرح زیر بود:

(۱) مرحله آموزش نقشه مفهومی: اصلی‌ترین فعالیت این مرحله برای گروه‌های آزمایشی آموزش و تمرین تهیه نقشه مفهومی برای مطالب درسی بود. بر طبق نظر نواک و کوئین (۱۹۸۴) همچون بسیاری از مهارت‌های شناختی، توانایی ترسیم نقشه مفهومی با تمرین بهبود می‌یابد. در این مرحله معلم پس از آموزش کلاسی محتوای درس از دانش‌آموزان می‌خواست که به‌عنوان تکلیف خانه نقشه مفهومی محتوای موردنظر را تهیه کنند. یک دوره ۲ ساعته آموزشی برای ترسیم و تکمیل نقشه‌های مفهومی قبل از اجرای پژوهش اجرا شد. روند عمومی ترسیم نقشه مفهومی به اقتباس از چالاروت و دی باکر<sup>۱</sup> (۲۰۰۳) طی گام‌های زیر آموزش داده شد.

≠ مفاهیم اصلی که باید در نقشه قرار گیرند را انتخاب می‌کنند.

≠ فهرستی از مفاهیم را از کلی‌ترین تا جزئی‌ترین رتبه‌بندی می‌کنند.

≠ مفاهیم را بر طبق دو ملاک خوشه‌بندی می‌کنند: مفاهیمی که از لحاظ کلیت مشابه هستند (ارتباط افقی) و مفاهیمی را که دارای ارتباط نزدیکی باهم هستند (ارتباط سلسله مراتبی).

≠ مفاهیم را برای نشان دادن روابط فی مابینشان به ترتیبی منظم می‌کنند. در این مرحله لازم بود چندین بار به مرتب کردن، تفکر و خوشه‌بندی پرداخت.

≠ مفاهیم را با خطوطی به هم وصل می‌کنند و به هر خط عبارات ربط دهنده اضافه می‌شود.

(۲) مرحله اجرای آزمون‌ها: پس از یک ماه آموزش محتوای موردنظر و تمرین ترسیم نقشه از محتوای آموزش داده شده ارزشیابی به عمل آمد. در این مرحله آموخته‌های فراگیران از محتوای درسی آموزش داده شده مورد سنجش قرار گرفت. برای اندازه‌گیری میزان پیشرفت تحصیلی دانش‌آموزان دو گروه اول از آزمون ترسیم آزاد، دو گروه دوم از آزمون ترسیم هدایت شده، دو گروه سوم از آزمون تکمیل آزاد و برای ارزشیابی دو گروه آخر از آزمون تکمیل هدایت شده

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی

استفاده شد. همچنین همه اعضای این گروه‌ها به یک آزمون چهارگزینه‌ای از محتوای موردنظر پاسخ دادند.

(۳) نمره‌گذاری آزمون‌ها: در این مرحله آزمون‌های اجرا شده به نحو زیر نمره‌گذاری شدند. روش‌های دقیق‌تر نمره‌گذاری آزمون‌های مبتنی بر نقشه مفهومی در پژوهش‌های مصرآبادی (ب، ۱۳۹۰) و مصرآبادی (۱۳۸۹) ارائه شده است. در این مقاله به شرح مختصر این روش‌های نمره دهی می‌پردازیم:

**آزمون چهارگزینه‌ای:** برای نمره دهی این آزمون‌ها به شمارش پاسخ‌های درست از کل ۳۰ سؤال بر اساس یک کلید پاسخ اقدام و سپس نسبت نمرات نهایی از ۲۰ محاسبه و نمره نهایی هر برگه مشخص شد.

**آزمون تکمیل آزاد:** بر اساس یک نمونه نقشه مفهومی کلید با توجه به میزان تکمیل بخش‌های مختلف نقشه مفهومی ناقص به هر جای خالی صحیح پر شده یک نمره ارائه شد و سپس نسبت نمرات نهایی از ۲۰ حساب و نمره نهایی هر برگه مشخص شد.

**آزمون تکمیل هدایت شده:** این آزمون همچون شیوه قبل نمره‌گذاری شد؛ اما در این روش برای عدم استفاده از هر مفهوم داده شده ۰/۵ نمره نیز کسر شد. در نهایت با جمع و تفریق این نمرات، نسبت نمره نهایی از ۲۰ محاسبه و نمره نهایی هر برگه مشخص شد.

**آزمون ترسیم آزاد:** برای نمره‌گذاری این آزمون‌ها از سه شیوه نمره دهی رابطه‌ای، ساختاری و شباهت به نقشه معلم استفاده شد. در شیوه رابطه‌ای به هر رابطه بین مفاهیم به‌طور جداگانه (بدون در نظر گرفتن دیگر رابطه‌ها) نمره‌ای اختصاص داده شد. در روش ساختاری به دانش ساختاری و سلسله مراتبی اهمیت داده می‌شود. در شیوه شباهت به نقشه معلم، با مقایسه نقشه مفهومی دانش‌آموز با یک نقشه مفهومی کامل تهیه شده توسط معلم، برای هر گزاره صحیح در نقشه دانش‌آموز یک نمره اختصاص داده شد. همچنین برای هر گزاره غیر صحیح نیز ۰/۵ نمره از نمره کل کسر شد.

**آزمون ترسیم هدایت شده:** در این شیوه همچون روش قبل عمل شد با این تفاوت که برای عدم استفاده از هر مفهوم داده شده ۰/۵ نمره نیز کسر شد.

### یافته‌های اعتبار بخشی به آزمون‌های نقشه مفهومی

پس از طراحی و اجرای آزمایشی لازم است که بر اساس شیوه‌های رسمی ارزشیابی ابزارهای اندازه‌گیری به بررسی روایی و پایایی آن‌ها پرداخته شود. برای اعتباربخشی به روش‌های ارزشیابی مبتنی بر نقشه مفهومی از شاخص‌های روایی محتوایی، روایی ملاکی، روایی افتراقی و پایایی نمره‌گذاران استفاده شد. برای اطمینان از روایی محتوایی از قضاوت ۵ نفر داور استفاده شد. میزان توافق حاصل از نظرات داوران در مورد این‌که تا چه حدی آزمون مورد نظر توانسته محتوا و اهداف آموزشی را پوشش دهد، شاخصی از روایی محتوا بود. برای آگاهی از روایی ملاکی نیز از ضریب همبستگی بین نمرات حاصل از این آزمون‌ها با یک آزمون چندگزینه‌ای استفاده شد. همچنین برای آگاهی از روایی افتراقی از تمایز گروه‌های بالا و پائین پیشرفت تحصیلی توسط این دسته از آزمون‌ها استفاده شد. برای آگاهی از پایایی نمره‌گذاران نیز از ضریب توافق بین نمرات ارائه شده توسط سه نمره‌گذار بهره گرفته شد. در جدول نتایج اعتباربخشی به آزمون‌های نقشه مفهومی ارائه شده است.

جدول ۲: نتایج اعتباربخشی به آزمون‌های نقشه مفهومی به تفکیک چهار گروه

تکلیف	روایی محتوایی	میزان هدایت	روایی محتوایی	روایی افتراقی	پایایی نمره‌گذاران
تکمیل	٪۸۴	آزاد	۰/۳۴	۲/۷۴	۰/۹۱
		هدایت شده	۰/۴۲	۳/۲۴	۰/۹۶
ترسیم	٪۷۵	آزاد	۰/۷۵	۵/۳۳	۰/۸۲
		هدایت شده	۰/۴۰	۲/۹۹	۰/۸۳

**روایی محتوایی:** در ستون دوم جدول ۲ نتایج روایی ملاکی گزارش شده است. برای اطمینان از روایی محتوایی آزمون‌های مبتنی بر نقشه مفهومی از پنج نفر از معلمان درس زیست شناسی خواسته شد که در قالب یک مقیاس صفر تا ۱۰۰ میزان پوشش محتوا و هدف‌های آموزشی در چهار نوع آزمون را رتبه‌بندی کنند. پس از مشخص شدن رتبه‌های داوران میانگین رتبه‌های داوران محاسبه و در صد ضرب شد تا درصد پوشش محتوا و اهداف آموزشی مشخص شود. برای آزمون‌های تکمیل درصد پوشش محتوا ٪۸۴ و برای آزمون‌های ترسیم این برابر با ٪۷۵ گزارش شده است. همچنین روایی محتوایی یک آزمون چهارگزینه‌ای (آزمون ملاک) برابر با ٪۸۷

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی

گزارش شد. مقایسه این رتبه‌ها با آزمون کروسکال - والیس با مقدار مجذور کای غیرمعنادار برابر با ۵/۲۹ نشان‌دهنده غیرمعنادار بودن تفاوت رتبه‌های روایی محتوایی داوران برای این سه آزمون بود.

**روایی ملاکی:** برای اطمینان از روایی ملاکی آزمون‌های تکمیل و ترسیم آزاد و هدایت شده نقشه مفهومی همبستگی نمرات حاصل از آن‌ها با نمرات یک آزمون چهارگزینه‌ای استفاده شد. در ستون چهارم جدول ۲ ضرایب همبستگی نمرات چهار آزمون مبتنی بر نقشه مفهومی با آزمون چهارگزینه‌ای نشان داده شده است. بیشترین ضریب همبستگی به آزمون ترسیم آزاد (۰/۷۵) و کمترین ضریب به آزمون تکمیل آزاد (۰/۳۴) تعلق دارد. همه این ضرایب معنادار هستند؛ هرچند که به جز آزمون ترسیم آزاد در سه شیوه دیگر ضرایب همبستگی حاصل نشان‌دهنده یک اندازه اثر متوسط به پائین هستند.

**روایی افتراقی:** برای آگاهی از روایی افتراقی چهار آزمون مبتنی بر نقشه مفهومی تفاوت معنادار میانگین نمرات دانش‌آموزان گروه پیشرفت تحصیلی بالا و پائین در نمرات این آزمون‌ها مورد بررسی قرار گرفت. در ستون پنجم جدول ۲ نمرات  $t$  برای اختلاف میانگین‌های این دو گروه از دانش‌آموزان ارائه شده است. تمام این مقادیر معنادار هستند که نشان‌دهنده تمایز دو گروه دانش‌آموزان ضعیف و قوی توسط این آزمون‌ها است.

**پایایی مصححان:** برای آگاهی از پایایی مصححان در شیوه‌های مختلف ارزشیابی مبتنی بر نقشه مفهومی از ضرایب همبستگی نمرات ارائه شده توسط ۳ داور استفاده شد. پس از به دست آوردن این ضرایب همبستگی میانگین این ضرایب به دست آمد که در ستون ششم جدول ۲ ارائه شده است. همانگونه که مشخص است بین نمره‌گذاران در نمره‌گذاری همه آزمون‌های نقشه مفهومی توافق نسبتاً کاملی وجود دارد. برای تمام آزمون‌های ضرایب همبستگی چهار آزمون نقشه مفهومی در سطح بسیار بالا و معنادار است. هرچند کمترین توافق در نمره‌گذاری داوران در آزمون‌های ترسیم مشاهده شد و در آزمون‌های تکمیل به‌ویژه در آزمون تکمیل هدایت شده نقشه مفهومی توافق بالایی بین نمره‌گذاران مشاهده شد.

## بحث و نتیجه‌گیری

این پژوهش با هدف طراحی، امکان‌سنجی اجرا و اعتباربخشی به آزمون نقشه‌های مفهومی اجرا شد. یافته‌های کیفی پژوهش نشان دادند که برای متن‌های درسی فارسی می‌توان به طراحی چهار نوع آزمون مبتنی بر نقشه مفهومی پرداخت. همچنین این ابزارها پس از ارائه راهنمایی‌ها و آموزش‌های کافی در کلاس‌های آزمایشی اجرا شدند. در زمینه اعتباربخشی به این آزمون‌ها نیز این چهار نوع آزمون از لحاظ چهار شاخص روایی محتوایی، روایی ملاکی، روایی افتراقی و پایایی نمره‌گذاران مورد مقایسه قرار گرفته‌اند. نتایج نشان داد که این روش‌ها دارای مقادیر متفاوتی از این شاخص‌ها هستند. با مقایسه این چهار شیوه متوجه می‌شویم که روش تکمیل هدایت شده در دو شاخص روایی محتوایی و پایایی نمره‌گذاران نسبت به دیگر روش‌ها بهتر است. همچنین شیوه ترسیم آزاد در دو شاخص روایی ملاکی و روایی افتراقی نسبت به دیگر شیوه‌های آزمون نقشه مفهومی دارای شاخص‌های بهتری است.

در زمینه روایی محتوایی یافته‌های این تحقیق نشان داد که آزمون‌های نقشه مفهومی به اندازه آزمون‌های چهارگزینه‌ای می‌توانند از روایی محتوایی قابل قبولی برخوردار باشند. این نتیجه خیلی دور از انتظار نیست، چرا که در نیمی از آزمون‌های نقشه مفهومی (روش‌های ترسیم و تکمیل هدایت شده) مفاهیم مهم حوزه موردنظر از قبل ارائه شده بود تا فراگیر آن‌ها را در نقشه ترسیمی خود بگنجانند. همین امر خود موجب به کارگیری این مفاهیم که محتوای مهم حوزه مورد سنجش را تشکیل می‌دادند، شد. در زمینه نصف دیگر آزمون‌ها که به صورت ترسیم و تکمیل آزاد بودند، انتظار این بود که فراگیران مفاهیم اصلی را در نقشه‌هایشان بیاورند؛ چرا که اساساً نقشه مفهومی ماهیتش به‌گونه‌ای است که از مفاهیم مهم حوزه مورد آزمون تشکیل می‌یابد.

همبستگی‌های متوسط معناداری که بین نمرات حاصل از انواع شیوه‌های ارزشیابی مبتنی بر نقشه مفهومی با یک آزمون چهارگزینه‌ای به دست آمد، نشان از روایی ملاکی قابل قبول این آزمون‌ها دارد. این یافته‌ها با نتایج دیگر پژوهش‌های این حوزه همخوانی دارد (لیو و هینچی<sup>۱</sup>، ۱۹۹۳؛ رایس، رایان و سامسون<sup>۲</sup>، ۱۹۹۸). از لحاظ شدت همبستگی نیز مقادیر همبستگی به

1. Liu & Hinchey
2. Rice, Ryan & Samson

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی

دست آمده در این پژوهش که به طور متوسط در حدود ۰/۵۰ بود با مقادیر همبستگی پژوهش فاروخ و کراسه<sup>۱</sup> (۱۹۹۶) بسیار نزدیک است. البته وجود همبستگی متوسط بین آزمون‌های نقشه مفهومی و آزمون چهارگزینه‌ای را شاید به راحتی نتوان به حساب پائین بودن نسبی روایی ملاکی گذاشت؛ چرا که اساساً این دو نوع آزمون به شیوه‌ای متفاوت سطح دانش فراگیران را اندازه می‌گیرند. یافته‌های بخش دیگری از تحقیق نشان داد که در همه آزمون‌های نقشه مفهومی دانش‌آموزان گروه بالا عملکرد بهتری نسبت به دانش‌آموزان گروه پائین داشتند. این یافته نشان دهنده روایی افتراقی آزمون‌های نقشه مفهومی است. در مورد روایی افتراقی آزمون‌های مبتنی بر نقشه مفهومی نیز نتایج پژوهش ویست و همکاران (۲۰۰۲) نشان می‌دهد که نقشه‌های مفهومی در این زمینه ابزارهای قابل اعتمادی هستند. یافته‌های آن‌ها نشان می‌دهد که آزمون‌های مبتنی بر نقشه مفهومی قادر هستند که تفاوت‌های مورد انتظار بین آزمودنی‌ها را به درستی بعد از مداخله یک آموزش تشخیص دهند.

در زمینه پایایی نمره گذاران نیز یافته‌های پژوهش نشان داد که بین نمره گذاران از لحاظ میزان نمره‌ای که بایستی به هر نقشه مفهومی تعلق گیرد توافق بالایی وجود دارد. این یافته با نتایج پژوهش شیولسون و همکاران (۲۰۰۵) همخوان است. یافته‌های آن پژوهش نشان از همبستگی بالای (حدود ۰/۹۰) بین نمرات نمره گذاران داشت. نکته‌ای که بایستی مورد توجه قرار گیرد این است که پایایی نمره گذاران به شدت تحت تأثیر شیوه نمره‌گذاری است. یافته‌های پژوهش مک کلور و همکاران (۱۹۹۹) نشان می‌دهد در شرایطی که از شش روش نمره دهی برای نمره‌گذاری نقشه‌های مفهومی استفاده شود مقدار پایایی این شش روش نمره‌گذاری می‌تواند دارای دامنه‌ای از ۰/۲۳ تا ۰/۷۶ قرار داشته باشند.

مقایسه شاخص‌های روایی و پایایی شیوه‌های چهارگانه ارزشیابی مبتنی بر نقشه مفهومی نشان می‌دهد که روش تکمیل هدایت شده در روایی محتوایی و پایایی نمره گذاران و شیوه ترسیم آزاد در روایی ملاکی و روایی افتراقی نسبت به دیگر شیوه‌های آزمون نقشه مفهومی دارای شاخص‌های بالاتری هستند. این یافته در راستای مبانی نظری و تجربی موجود در زمینه ارزشیابی مبتنی بر نقشه مفهومی است. بر اساس نظرات روئیر-پرمیو و شیولسون (۱۹۹۶) شیوه‌های مختلف آزمون نقشه مفهومی جنبه‌های مختلفی از دانش فراگیران را اندازه می‌گیرند.



به جهت این جنبه‌های متفاوت دانش فراگیران و نیز نوع تکلیفی که آزمون به فراگیر ارائه می‌دهد شاخص‌های روایی و پایایی می‌توانند متفاوت باشند. بنابراین می‌توان گفت که الزاماً بالا بودن شاخص‌های روایی و پایایی در تعدادی از این شیوه‌ها ناشی از ویژگی‌های خاص آن‌ها است که البته می‌تواند در حوزه‌ای دیگر محدودیت محسوب شود. به‌عنوان مثال، بالا بودن پایایی نمره گذاران در روش‌های تکمیل به جهت ساختار از قبل تعریف شده و پاسخ‌های مشخص این روش است که در روش‌های ترسیم وجود ندارد.

در یک جمع‌بندی از یافته‌های پژوهش حاضر می‌توان گفت که آزمون‌های نقشه مفهومی از پایایی و روایی نسبتاً بالایی برخوردار هستند. هرچند بایستی دقت داشت که در این پژوهش ملاک مورد مقایسه در بیشتر شرایط نتایج یک آزمون چهارگزینه‌ای بود؛ که خود این نوع از آزمون‌ها نیز مورد انتقاد هستند. بنابراین الزامی نیست این‌گونه فرض شود که همخوانی بالایی نمرات حاصل از آزمون‌های نقشه مفهومی با چنین ملاکی به معنای قابل اعتماد بودن آزمون‌های مبتنی بر نقشه مفهومی است و یا برعکس.

حال سؤالی که اینجا مطرح می‌شود این است که آیا مزیت‌های استفاده از نقشه‌های مفهومی ارزش آن را دارد که محدودیت‌ها و مشکلات ذاتی کاربری شیوه‌های متنوع این روش‌ها و نیز نمره دهی به این آزمون‌ها را قبول کنیم؟ آشکارا پاسخ به این سؤال به نقش نقشه‌های مفهومی در یک بستر آموزشی مربوط می‌شود. برای بسیاری از افراد مزایای استفاده از نقشه‌های مفهومی بیشتر بهبود فرایند یادگیری و فراهم ساختن بازخورد است تا به دست آوردن اطلاعات کمی که قابل تبدیل به نمره باشند. اگر تمایز دو اصطلاح سنجش برای یادگیری و سنجش از یادگیری (سیف، ۱۳۸۷) را در نظر داشته باشیم استفاده از شیوه‌های ارزشیابی مبتنی بر نقشه مفهومی را تا حد زیادی می‌توانیم در طیف سنجش برای یادگیری بدانیم، چرا که آزمون‌های نقشه مفهومی در مقایسه با آزمون‌های چهارگزینه‌ای قابلیت بیشتری در کشف نواقص یادگیری دانش‌آموزان و بالطبع هدایت دانش‌آموزان به سوی شیوه‌های صحیح یادگیری دارند. این بحث زمانی بهتر قابل درک است که بدانیم ارزشیابی از یادگیری دانش‌آموزان خود بخشی جدایی ناپذیر از فرایندهای بزرگ‌تری به نام یادگیری و تدریس است. بنابراین بر اساس یافته‌های این پژوهش و نکات طرح شده در بالا می‌توان پیشنهاد کرد که از نقشه‌های مفهومی به‌عنوان ابزارهای ارزشیابی به ویژه جهت اهداف تکوینی استفاده شود. همچنین می‌توان تعدادی از آزمون‌های نقشه مفهومی

طراحی، اجرای آزمایشی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی به‌ویژه آزمون‌های تکمیل در کتب درسی جهت آشنایی دانش‌آموزان با این شیوه ارزشیابی ارائه نمود؛ تا به مرور فراگیران و معلمان با این شیوه ارزشیابی آشنا شوند.

### کتاب‌نامه

- سیف، علی اکبر. (۱۳۸۲). اندازه‌گیری، سنجش، و ارزشیابی آموزشی. تهران: نشر دوران.
- سیف، علی اکبر. (۱۳۸۷). سنجش فرایند و فراورده تحصیلی: روش‌های قدیم و جدید. تهران: نشر دوران.
- مصرآبادی، جواد، فتحی آذر، اسکندر، و استوار، نگار. (۱۳۸۴). اثربخشی ارائه، ساخت فردی و ساخت گروهی نقشه مفهومی به‌عنوان یک راهبرد آموزشی. فصلنامه علمی - پژوهشی نوآوری‌های آموزشی، ۱۳، ۳۱-۱۱.
- مصرآبادی، جواد، علوی، الهه، و استوار، نگار. (۱۳۸۵). مقایسه اثربخشی استفاده از نقشه مفهومی به‌عنوان یک راهبرد یاددهی - یادگیری در پیشرفت درسی دروس مختلف. چکیده مقالات و طرح‌های نوآورانه در همایش ملی نوآوری‌های آموزشی. تهران: دی‌ماه ۱۳۸۵، مؤسسه پژوهش برنامه‌ریزی درسی و نوآوری‌های آموزشی.
- مصرآبادی، جواد، حسینی نسب، داوود، فتحی آذر، اسکندر. و مقدم واحد، محمد. (۱۳۸۷). اثربخشی راهبرد یاددهی - یادگیری نقشه مفهومی بر بازده‌های شناختی - عاطفی در یادگیری درس زیست‌شناسی. مجله روانشناسی دانشگاه تبریز. شماره ۸، ۱۲۸-۱۰۵.
- مصرآبادی، جواد، حسینی نسب، داوود، فتحی آذر، اسکندر. و مقدم واحد، محمد. (۱۳۸۶). تأثیر ساخت و ارائه نقشه مفهومی و سبک یادگیری بر یادداری، درک و حل مسئله در زیست‌شناسی. مطالعات تربیتی و روان‌شناسی، شماره ۱۰(۳)، ۱۶۲-۱۴۱.
- مصرآبادی، جواد. (۱۳۸۹). امکان سنجی به‌کارگیری روش‌های ارزشیابی مبتنی بر نقشه مفهومی در زیست‌شناسی. طرح پژوهشی چاپ نشده: مؤسسه پژوهشی برنامه‌ریزی درسی.
- مصرآبادی، جواد، فتحی آذر، اسکندر، و استوار، نگار. (۱۳۸۴). اثربخشی ارائه، ساخت فردی و ساخت گروهی نقشه مفهومی به‌عنوان یک راهبرد آموزشی. فصلنامه علمی - پژوهشی نوآوری‌های آموزشی، شماره ۱۳، ۳۱-۱۱.

مصرآبادی، جواد، استوار، نگار. (۱۳۸۸). تأثیر اثربخشی نقشه مفهومی بر پیشرفت تحصیلی دانش‌آموزان در دروس زیست شناسی، روان شناسی و فیزیک. اندیشه‌های نوین تربیتی، ۵(۱)، ۹۳-۱۱۴.

مصرآبادی، جواد. (الف، ۱۳۹۰). پیامدهای شناختی و عاطفی استفاده از اسلاید در کلاس درس. فناوری آموزش، ۵(۴)، ۲۹۹-۳۰۸.

مصرآبادی، جواد. (ب، ۱۳۹۰). بررسی روایی و پایایی شیوه‌های نمره‌گذاری آزمون‌های مبتنی بر نقشه مفهومی، زیر چاپ در فصلنامه اندیشه‌های نوین تربیتی.

Chang, K.E., Sung, Y.T., & Chen, S.F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning*, 17, 21-33.

Chularut, P., & DeBacker, T.K. (2003). The influence of concept mapping on achievement, self-regulation, and self-efficacy in students of English as a second language. *Contemporary Educational Psychology*, 29, 248-263.

Farrokh, K. & Krause, G. (1996). The relationship of concept-mapping and course grade in cell biology. Meaningful Learning Forum. *On the World Wide Web: <http://www2.ucsc.edu/mlrg>*.

Herl, H., & Baker, E.D. (1996). Construct validation of approach to modeling cognitive structure of US history knowledge. *Journal of Education Research*, 89, 213-230.

Herl, H., & Baker, E.D. (1996). Construct validation of approach to modeling cognitive structure of US history knowledge. *Journal of Education Research*, 89, 213-230.

Jacobs-Lawson, J.M., & Hershey, D.A. (2002). Concept Maps As an Assessment Tool in Psychology Courses. *Teaching of Psychology*, 29(1), 25-29.

Keppens, J., & Hay, D. (2008). Concept map assessment for teaching computer programming. *Computer Science Education*, 18(1), 31-42.

Kilic, G.B. (2003). Concept maps and language: a Turkish experience. *International Journal of Science Education*, 25, 1299-1311.

Liu, X., Hinchey, M. (1993). Validity and reliability of concept map as an alternative science assessment. *Published in the proceedings of international Seminar on Misconceptions and Educational Strategies in Science and Mathematics*. Department of Education, Ithaca, NY: Cornell University.

Markham, K.M., Mintzes, J.J. & Jones, M.G. (1994). The concept map as a research and evaluation tool: further evidence of validity. *Journal of Research in Science Teaching*, 31, 91-101.

- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 4, 475–492.
- Novak, J. D., & Gowin, D. B. (1984). *Concept mapping for meaningful learning*. Cambridge University Press, Cambridge.
- Rice, D.C., Ryan, J.M. & Samson, S.M. (1998). Using concept mapping to assess student learning in the science classroom: Must different methods compete? *Journal of Research in Science Teaching*, 35(10), 1103-1127.
- 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*, 33, 569–600.
- Shavelson, R.J., Ruiz-Primo, M., & Wiley, E. (2005). Windows into the mind. *Higher Education*, 49, 413–430.
- West, D., Park, J., Pomeroy, J., & Sandoval, J. (2002). Concept mapping assessment in medical education: A comparison of two scoring systems. *Medical Education*, 36, 820–826.
- West, D.C, Pomeroy, J.R, Park, J.K, Gerstenberger, E.A, & Sandoval, J.H. (2000). Critical thinking in graduate medical education: a role for concept mapping assessment? *Journal of American Medicine Association*, 284, 1105–10.