

معرفی و اعتباربخشی آزمون‌های مبتنی بر نقشه مفهومی در ارزشیابی عملکرد تحصیلی دانش‌آموزان*

دکتر جواد مصرآبادی^۱

چکیده

استفاده از نقشه‌های مفهومی به عنوان ابزارهای ارزشیابی تحصیلی چندین دهه است که در کانون توجه محققان و مربیان قرار دارد. هدف از تدوین این مقاله معرفی و اعتباربخشی شیوه‌های استفاده از نقشه مفهومی به عنوان ابزارهای ارزشیابی تحصیلی است. به طور مشخص در اینجا شیوه‌های ترسیم و تکمیل (آزاد و هدایت شده) نقشه مفهومی معرفی می‌شود و این نقشه‌ها از لحاظ شاخص‌های روایی محتوایی، روایی ملاکی، روایی افتراقی و پایایی نمره‌گذاران مورد بررسی قرار می‌گیرد. نمونه آماری پژوهش ۲۲۱ نفر از دانش‌آموزان سال دوم متوسطه رشته علوم تجربی شهر تبریز در سال تحصیلی ۸۸-۱۳۸۷ بودند که در هشت کلاس حضور داشتند. یافته‌های پژوهش حاضر به طور کلی نشان داد که آزمون‌های مبتنی بر نقشه مفهومی از روایی محتوایی بالایی برخوردارند. همچنین این آزمون‌ها دارای مقادیر متفاوت همبستگی (با مقدار متوسط ۰/۵۰) با یک آزمون چهارگزینه‌ای بودند. همه آزمون‌های نقشه مفهومی به خوبی از عهده تشخیص تفاوت بین نمرات دو گروه تحصیلی بالا و پائین برآمدند. همچنین نتایج نشان داد که ضریب توافق نمره‌گذاران آزمون‌های نقشه مفهومی در حد بالایی است. با مقایسه چهار شیوه آزمون مشخص شد که روش تکمیل هدایت شده در روایی محتوایی و پایایی نمره‌گذاران نسبت به دیگر روش‌ها بهتر است و شیوه ترسیم آزاد در روایی ملاکی و روایی افتراقی، نسبت به دیگر شیوه‌های آزمون نقشه مفهومی، دارای شاخص‌های بالاتری است.

کلید واژه‌ها: ارزشیابی تحصیلی، عملکرد تحصیلی، آزمون، نقشه مفهومی، روایی، پایایی

* تاریخ دریافت مقاله: ۱۳۸۹/۶/۲۱ تاریخ شروع بررسی: ۱۳۸۹/۸/۱۵ تاریخ پذیرش مقاله: ۱۳۹۰/۵/۲۱
۱. استادیار دانشگاه تربیت معلم آذربایجان (این پژوهش بر اساس یافته‌های طرح پژوهشی مصوب سازمان پژوهش و برنامه ریزی آموزشی تدوین شده است).

مقدمه

پژوهش‌ها نشان می‌دهد که فراگیران فاقد فهم منسجم و بلندمدت از محتوای علم هستند و توانایی بسیار کمی برای کاربرد دانسته‌های خود دارند (هزل و پراسر^۱، ۱۹۹۴) و درک آن‌ها از مفاهیم علمی ضعیف است. (لین و هو^۲، ۲۰۰۳). همچنین پژوهش‌ها نشان می‌دهد که دانش‌آموزان و دانشجویان به مقدار کمی از شیوه‌های یادگیری معنادار استفاده می‌کنند (نواک^۳، ۱۹۹۰). بی‌شک یکی از علل مهم این کاستی‌ها کاربست شیوه‌های نامناسب در ارزشیابی از آموخته‌های فراگیران است. به جهت ارتباط نزدیکی که بین روش‌های یادگیری و شیوه‌های ارزشیابی وجود دارد، در صورتی که یادگیری‌های دانش‌آموزان از همان ابتدا با شیوه‌های غلط و نامناسبی ارزشیابی شوند، آنها نیز روش‌های نادرست یادگیری را در پیش خواهند گرفت. به‌عنوان مثال اگر در ارزشیابی تنها بر سؤالات چندگزینه‌ای تأکید شود، طبعاً فراگیران نیز، به جای درک ساختار کلی محتوا، بیشتر به یادگیری‌های جزئی خواهند پرداخت (سیف، ۱۳۷۸).

روان‌شناسان تربیتی و صاحب‌نظران حیطهٔ سنجش تحصیلی با آگاهی از این امر دست به تغییراتی در اهداف و شیوه‌های ارزشیابی تحصیلی زده‌اند. اکنون بیشترین تأکید بر سنجش هدف‌های فرایندی و شناختی یادگیری است تا هدف‌های فرآورده‌ای و رفتاری. همچنین شیوه‌های سنجش عملکردی، توصیفی و تکوینی نیز جایگزین روش‌های قدیمی شده‌اند. بخش اعظمی از این تحولات ناشی از افزایش کارکردهای روزافزون نظریه‌های شناختی در زمینه‌های برنامه‌ریزی درسی، آموزش، یادگیری و ارزشیابی است.

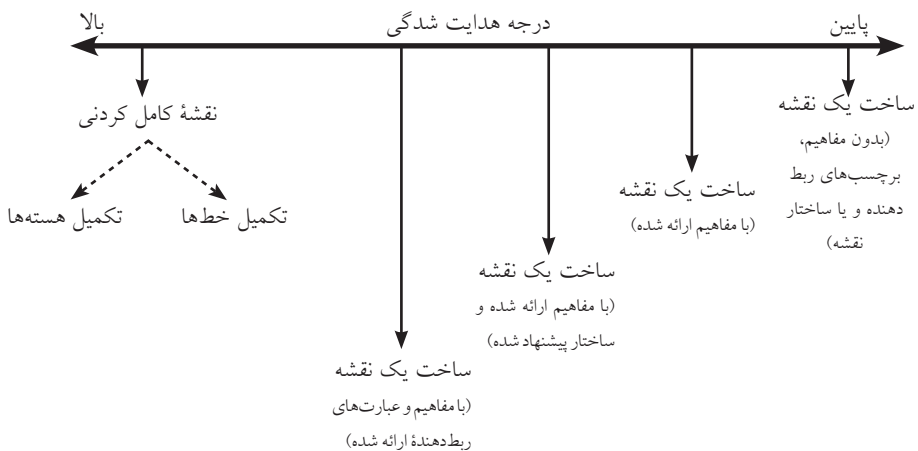
امروزه تأکید بر ساختار شناختی یعنی همان شیوه‌ای که دانش‌آموزان به وسیله آن به یادگیری‌هایشان ساختار می‌هند - بخش مهمی از پژوهش‌های تربیتی را در حوزه سنجش تحصیلی شامل می‌شود. اکنون دیگر تأکیدی بر انبارش اطلاعات نمی‌شود و روان‌شناسان شناختی اساس دانش فرد را ساختار آن می‌دانند نه مقداریا کمیت آن. از این رو «سازمان دانش فرد» در حال حاضر بیشتر مورد تأکید مریبان و نظریه‌پردازان تعلیم و تربیت است. همان‌گونه که چیرگی یک فرد در یک حوزه دانش افزایش پیدا می‌کند اجزای دانش آن حوزه برای او به طور فزاینده‌ای به هم مرتبط می‌شوند. همچنین مشخص شده است که شیوه‌های یادگیری دانش‌آموزان تحت تأثیر ساختار شناختی آنها قرار دارد (بیگز^۴، ۱۹۸۹؛ ادمسون^۵ و نواک، ۱۹۹۳) و ساختار شناختی نشانگر تفاوت‌های کیفی در فهم است (بزی^۶، ۱۹۹۶؛ نواک و موسوندا^۷، ۱۹۹۱، پراسر^۸، ۱۹۸۷). همان‌گونه که پیرسل، اسکیر و مینتز^۹ (۱۹۹۷) خاطر نشان می‌سازند «دانش‌آموزانی که گزارش می‌دهند و از راهبردهای پردازش اطلاعات فعال و عمیق استفاده می‌کنند، از ساختار وسیع و روشنی در مورد چارچوب دانش برخوردارند (ص ۲۱۳)».

در سال‌های اخیر شیوه‌های متنوعی برای بازنمایی‌های ترسیمی مفاهیم ارائه شده است. به‌عنوان نمونه می‌توان مفاهیم را در قالب نگاره‌های راندهوس^۱، نگاره‌های وی^{۱۱} و نگاره‌های هسته-پیوند^{۱۲} ارائه کرد. در بین این نگاره‌ها، نگاره‌های هسته - پیوند، که نقشه مفهومی نیز یکی از آنها است، بیشترین کاربرد را در نمایش تصویری مفاهیم دارند. نگاره‌های هسته - پیوند نیز دارای انواع مختلفی مانند نقشه مفهومی، نقشه دانش^{۱۳}، شبکه معنایی^{۱۴}، نقشه ذهنی^{۱۵}، درخت مفهومی^{۱۶} و شبکه مفهومی^{۱۷} هستند. از بین این شیوه‌ها نقشه‌های مفهومی^{۱۸} به عنوان یک روش مستقیم‌تر برای به تصویر کشیدن روابط درونی بین مفاهیم یک حوزه پیشنهاد شده است. نقشه مفهومی توسط جوزف نواک در دهه ۱۹۸۰ در دانشگاه گرنل برای بازنمایی دانش به شکل ترسیمی ارائه شد. 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. نواک (۱۹۹۰) عقیده دارد که نقشه‌های مفهومی می‌توانند در بهبود حوزه تعلیم و تربیت در چهار سطح یادگیری، آموزش، برنامه‌ریزی درسی و ارزشیابی (حوزه مورد تمرکز پژوهش حاضر) مؤثر واقع شوند. مروری بر ادبیات پژوهشی خارج از ایران نشانگر اثربخشی نقشه مفهومی در این چهار حوزه است (هال و اودانل^{۱۹}، ۱۹۹۶؛ پاتله و روئیت^{۲۰}، ۲۰۰۳؛ هارتون، مک کونی، گالو، ودز و هامیلتون^{۲۱}، ۱۹۹۳؛ آل، هویکه و فیشر^{۲۲}، ۲۰۰۳؛ شرن، تورچیم و لاکمپ^{۲۳}، ۱۹۹۵؛ ویست، پارک، پومروی و ساندول^{۲۴}، ۲۰۰۲؛ شیولسون، روئیز-پریمو و ویلی^{۲۵}، ۲۰۰۵). پژوهش‌های داخل کشور نیز نشانگر اثربخشی نقشه‌های مفهومی به‌عنوان راهبردهای یادگیری - یاددهی هستند (مصرآبادی، فتحی آذر و استوار، ۱۳۸۴؛ مصرآبادی، علوی و استوار، ۱۳۸۵؛ مصرآبادی، حسینی نسب، فتحی آذر و مقدم، ۱۳۸۷؛ رحمانی، ۱۳۸۴). با این وجود پژوهش مدونی در داخل کشور در زمینه ارزشیابی مبتنی بر نقشه مفهومی و اعتباربخشی به آن انجام نشده است. اهمیت پژوهش در این حوزه از این رو است که همچون استفاده از نقشه مفهومی به‌عنوان یک راهبرد یادگیری که نوع زبان مورد استفاده می‌تواند بر اثربخشی آن مؤثر باشد (کلیس^{۲۶}، ۲۰۰۳)؛ ضروری به نظر می‌رسد که امکان استفاده از ارزشیابی مبتنی بر نقشه مفهومی در متن‌های درسی فارسی مورد بررسی قرار گیرد. در پژوهش حاضر پس از معرفی انواع آزمون‌های مبتنی بر نقشه مفهومی، به بررسی شاخص‌های روایی^{۲۷} و پایایی^{۲۸} این شیوه‌ها پرداخته می‌شود.

انواع آزمون‌های مبتنی بر نقشه مفهومی

طبق نظر مک کلور، سوناک و سوئن^{۲۹} (۱۹۹۹) آزمون نقشه مفهومی رویه‌ای را فراهم می‌سازد که نتیجه آن ساختن یک نقشه مفهومی است که بازنمایی از سازمان دانش فراگیر است. مرور

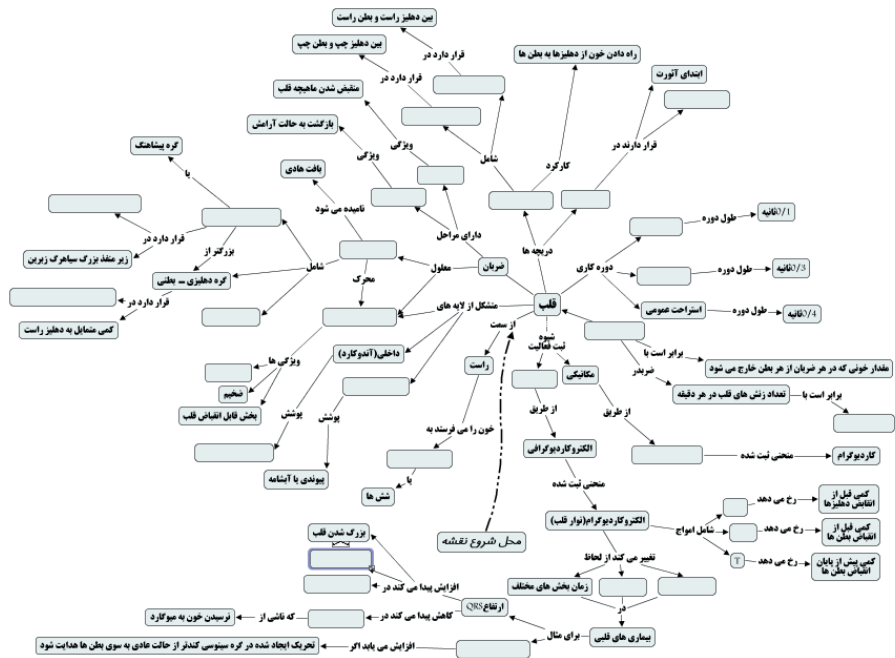
ادبیات پژوهشی ارزشیابی مبتنی بر نقشه مفهومی نشان می‌دهد که آزمون‌های نقشه مفهومی را می‌توان به دو شیوه اصلی، یکی ترسیم یک نقشه مفهومی و دیگری تکمیل یک نقشه مفهومی ناقص اجرا کرد. در روش ترسیم یک نقشه مفهومی فراگیر بایستی به طور کامل مفاهیم و روابط بین آنها را در یک نقشه نشان دهد؛ اما در روش تکمیل فراگیر تنها به کامل کردن بخش‌های ناقص نقشه ناقص می‌پردازد (همچون تکمیل مفاهیم داخل هسته‌ها یا عبارات روی خطوط ربط دهنده). در مورد تکلیفی که دانش‌آموز در ارزشیابی نقشه مفهومی بایستی انجام دهد (ترسیم یک نقشه کامل یا تکمیل یک نقشه ناقص) غالباً یک آزمون نقشه مفهومی در گستره‌ای از هدایت‌شدگی بر اساس میزان اطلاعاتی که در اختیار دانش‌آموز قرار داده می‌شود؛ قرار دارد. آزمون‌های نقشه مفهومی که از هدایت بالایی برخوردارند مفاهیم، خطوط ارتباطی، عبارت‌های ربط دهنده و ساختار نقشه را در اختیار دانش‌آموزان قرار می‌دهند. در مقابل، در یک آزمون 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 نوع تکلیف و درجه هدایت‌شدگی روئیز - پریمو و شاولسون^{۳۰} (۱۹۹۶) پنج روش ترسیم نقشه مفهومی در گستره‌ای از تکلیف ساخت یک نقشه کامل تا تکمیل یک نقشه معرفی کرده‌اند (شکل ۱). البته مشخص است که شیوه‌های مختلف ارزشیابی نقشه مفهومی اطلاعات یکسانی در مورد فهم منسجم و مرتبط فراگیران فراهم نمی‌کنند.



شکل ۱: روش‌های مختلف ارزشیابی نقشه مفهومی (اقتباس از روئیز - پریمو و شاولسون، ۱۹۹۶)

در این پژوهش با ترکیب نوع تکلیف آزمودنی (ترسیم یا تکمیل یک نقشه) و میزان هدایت‌شدگی این تکالیف (هدایت شده یا آزاد) به چهار شیوه اجرایی آزمون‌های مبتنی بر نقشه مفهومی به دست آمد. این شیوه‌ها به همراه توصیف مختصری در ادامه آمده است:

۱. **آزمون تکمیل آزاد نقشه مفهومی ناقص:** در این شیوه یک نقشه مفهومی ناقص از قبل ترسیم شده در اختیار دانش‌آموزان قرار داده می‌شود و آنها آزادند که هر کلمه یا عبارتی را برای تکمیل بخش‌های ناقص به کار گیرند. یک نمونه از این آزمون‌ها را در شکل ۲ می‌بینید.



شکل ۲: نمونه‌ای از آزمون تکمیل آزاد نقشه مفهومی ناقص (اقتباس از مصرآبادی، ۱۳۸۶)

۲. **آزمون تکمیل نقشه مفهومی ناقص بر اساس مفاهیم داده شده:** این آزمون همانند آزمون قبلی است با این تفاوت که فراگیران برای تکمیل نقشه مفهومی ناقص بایستی از مفاهیمی استفاده کنند که معلم در اختیارشان قرار می‌دهد. به عبارت دیگر مفاهیم لازم جهت تکمیل نقشه مفهومی در اختیار فراگیران قرار می‌گیرد و آنها بایستی با استفاده از این مفاهیم نقشه را کامل کنند.

۳. **آزمون ترسیم آزاد نقشه مفهومی:** در این شیوه فراگیران به‌طور آزادانه برای یک مطلب

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 آزمون‌ها براساس متن‌های درسی زیست‌شناسی سال دوم رشته علوم تجربی تهیه شد.

طرح پژوهش و روش اجرا

این تحقیق از نوع پژوهش ارزشیابی^{۳۳} بود. طبق نظر روسی^{۳۳} (به نقل از سیف، ۱۳۸۲) این دسته از پژوهش‌ها نوعی ارزشیابی هستند که در آن مسائل مربوط به روایی درونی، روایی بیرونی و روایی سازه منظور می‌شود. این پژوهش همچنین به شیوه درون گروهی انجام پذیرفت. درونی گروهی بودن طرح از این جهت است که افراد داخل هر گروه به دو شیوه اندازه‌گیری شدند (آزمون چهارگزینه‌ای و 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. مراحل اجرایی پژوهش در گروه‌های آزمایشی شامل چهار مرحله به شرح زیر بود:

۱. **مرحله آماده‌سازی و تمرین:** در این مرحله مقدمات کاربندی آزمایشی فراهم شد. 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. برای ترسیم نقشه‌های مفهومی از متن‌های آموزش داده شده، آزمودنی‌ها یا همان فراگیران، پس از مطالعه متن باید مراحل زیر را طی می‌کردند (اقتباس از چالاروت و دی باکر^{۳۴} ۲۰۰۳):

- انتخاب مفاهیم اصلی که باید در نقشه قرار گیرند؛
- رتبه‌بندی فهرستی از مفاهیم، از کلی تا جزئی؛
- خوشه‌بندی مفاهیم بر طبق دو ملاک؛ یکی براساس مفاهیمی که از لحاظ کلیت مشابه‌اند (ارتباط افقی) و دیگر براساس مفاهیمی که دارای ارتباط نزدیکی باهم هستند (ارتباط سلسله مراتبی)؛
- منظم کردن مفاهیم برای نشان دادن روابط فی‌مابین شان به ترتیب، در این مرحله لازم است چندین بار به مرتب کردن، تفکر و خوشه‌بندی پردازند.

- وصل کردن مفاهیم با خطوطی به یکدیگر؛ که 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 دو گروه آخر از آزمون تکمیل هدایت شده 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:
- آزمون چهارگزینه‌ای:** برای نمره‌دهی این آزمون‌ها به شمارش پاسخ‌های درست از کل ۳۰ سؤال براساس یک کلید پاسخ اقدام و سپس نسبت نمرات نهایی از ۲۰ محاسبه و نمره نهایی هر برگه مشخص شد.
- آزمون تکمیل آزاد:** براساس یک نمونه نقشه مفهومی کلید، با توجه به میزان تکمیل بخش‌های مختلف نقشه مفهومی ناقص، به هر جای خالی صحیح پر شده یک نمره داده شد، سپس نسبت نمرات نهایی از ۲۰ حساب و نمره نهایی هر برگه مشخص گردید.
- آزمون تکمیل هدایت شده: این آزمون همچون شیوه قبل نمره‌گذاری شد؛ اما برای عدم استفاده از هر مفهوم داده شده ۰/۵ نمره نیز کسر شد. درنهایت با جمع و تفریق این نمرات، نسبت نمره نهایی از ۲۰ محاسبه و نمره نهایی هر برگه مشخص گردید.
- آزمون ترسیم آزاد:** برای نمره‌گذاری این آزمون‌ها از سه شیوه نمره‌دهی رابطه‌ای^{۳۵}، ساختاری^{۳۶} و شباهت^{۳۷} به نقشه معلم استفاده شد. در شیوه رابطه‌ای به هر رابطه بین مفاهیم به طور جداگانه (بدون در نظر گرفتن دیگر رابطه‌ها) نمره‌ای اختصاص داده شد؛ در روش ساختاری به دانش‌ساختاری و سلسله مراتبی اهمیت داده می‌شد، و در شیوه مشابهت 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 ۰/۵ نمره از نمره کل کسر گردید. برای آگاهی از ترتیبات دقیق‌تر نمره‌گذاری آزمون‌های ترسیم به مصرآبادی (۱۳۸۸) و ویست و همکاران (۲۰۰۲) مراجعه شود.
- 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 کسر شد.

روش‌های تحلیل داده‌ها

برای اعتباربخشی به روش‌های ارزشیابی مبتنی بر نقشه مفهومی از روایی محتوایی^{۳۸}، روایی ملاکی^{۳۹}، روایی افتراقی^{۴۰} و پایایی نمره‌گذاران^{۴۱} استفاده شد. برای اطمینان از روایی محتوایی ۵ نفر داور به قضاوت خوانده شدند. میزان توافق حاصل از نظرات داوران در مورد این که آزمون مورد نظر تا چه حدی توانسته است 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.

یافته‌ها

روایی محتوایی: برای اطمینان از روایی محتوایی، از قضاوت پنج نفر از دبیران درس زیست‌شناسی به‌عنوان داور استفاده شد. داوران بایستی مشخص می‌کردند که آزمون‌های استفاده شده در تحقیق تا چه حدی توانسته‌اند محتوا و اهداف آموزشی مبحث تدریس شده را پوشش دهند. برای این منظور از آنان خواسته شد در قالب مقیاس صفر تا ۱۰۰ میزان پوشش محتوا و هدف‌های آموزشی در آزمون‌ها را رتبه‌بندی کنند. پس از رتبه‌بندی، داوران میانگین رتبه‌های داوران محاسبه و در صد ضرب شد تا درصد پوشش محتوا و اهداف آموزشی مشخص شود. همان‌گونه که در جدول ۱ مشخص است، داوران بر این باورند که آزمون‌های چهارگزینه‌ای بیش از دیگر آزمون‌های تکمیل و ترسیم نقشه مفهومی، محتوا و اهداف آموزشی را پوشش می‌دهند.

جدول ۱: میزان پوشش محتوا و هدف‌های آموزشی در آزمون‌های مختلف

نوع آزمون	داور ۱	داور ۲	داور ۳	داور ۴	داور ۵	درصد رتبه داوران
آزمون‌های چهارگزینه‌ای	۹۰	۸۵	۹۰	۹۵	۷۵	٪۸۷
آزمون‌های تکمیل	۸۰	۸۰	۹۵	۹۰	۷۵	٪۸۴
آزمون‌های ترسیم	۸۰	۸۰	۷۰	۷۵	۷۰	٪۷۵

جدول ۲ نتایج آزمون کروسکال - وایس برای بررسی تفاوت رتبه‌های داوران در مورد روایی محتوایی سه آزمون مورد استفاده در تحقیق را نشان می‌دهد. همان‌گونه که مشخص است

آزمون‌های چهارگزینه‌ای از بالاترین میانگین رتبه و آزمون‌های ترسیم از پائین ترین میانگین رتبه برخوردارند، با این وجود مقدار مجذور کای (χ^2) محاسبه شده برابر با ۵/۲۹ است که از لحاظ آماری معنادار نیست.

جدول ۲: آزمون کروسکال - والیس برای تفاوت رتبه‌های داوران از روایی محتوایی آزمون‌ها

نوع آزمون	تعداد	میانگین رتبه	مجذور کای	درجه آزادی	سطح معناداری
آزمون چهارگزینه‌ای	۵	۱۰/۵	۵/۲۹	۲	۰/۰۷
آزمون‌های تکمیلی	۵	۹/۱			
آزمون‌های ترسیم	۵	۴/۴			

روایی ملاکی: در جدول ۳ ضرایب همبستگی نمرات چهار آزمون مبتنی بر نقشه مفهومی با آزمون چهارگزینه‌ای نشان داده شده است. بالاترین ضریب همبستگی به آزمون ترسیم آزاد و کمترین ضریب به آزمون تکمیل آزاد تعلق دارد.

جدول ۳: ضرایب همبستگی نمرات چهار آزمون تکمیل و ترسیم آزاد و هدایت شده نقشه مفهومی با آزمون چهارگزینه‌ای

نقشه مفهومی								نوع آزمون
ترسیم				تکمیل				
هدایت شده		آزاد		هدایت شده		آزاد		
r	n	r	n	r	n	r	n	چهارگزینه‌ای
*۰/۴۰	۵۶	*۰/۷۵	۴۹	*۰/۴۲	۴۲	*۰/۳۴	۶۷	

* ضریب همبستگی در سطح ۰/۰۱ معنادار است.

روایی افتراقی: در جدول ۴ میانگین و انحراف معیار دانش‌آموزان دارای پیشرفت تحصیلی بالا و پائین در انواع آزمون‌های نقشه مفهومی ارائه شده است. همان‌گونه که مشخص است در تمام این آزمون‌ها دانش‌آموزان دارای پیشرفت تحصیلی بالا از میانگین نمرات بالاتری برخوردارند. تفاوت‌های بین دو گروه در هر چهار نوع آزمون معنادار است. با توجه به مقدار نمره t در آزمون ترسیم آزاد در مقایسه با دیگر آزمون‌ها می‌توان استنباط کرد که این آزمون از توان بیشتری در متمایز کردن دانش‌آموزان قوی از ضعیف برخوردار است.

جدول ۴: آزمون t مستقل برای تفاوت میانگین نمرات گروه پیشرفت تحصیلی بالا و پایین در نمرات آزمون‌های نقشه مفهومی

نوع آزمون	سطح پیشرفت تحصیلی	میانگین	انحراف معیار	تعداد	t	df	سطح معناداری
ترسیم آزاد نقشه مفهومی	گروه بالا	۱۴/۰۴	۴/۳۹	۲۶	۵/۳۳	*۴۳/۱۷	۰/۰۰۱
	گروه پایین	۸/۴۸	۲/۸۳	۲۳			
ترسیم هدایت شده نقشه مفهومی	گروه بالا	۱۱/۹۷	۲/۸۶	۴۰	۲/۹۹	۵۴	۰/۰۰۴
	گروه پایین	۸/۵	۳/۵۸	۱۶			
ترسیم آزاد نقشه مفهومی	گروه بالا	۲۰/۸۴	۹/۵۵	۱۹	۲/۷۴	۶۵	۰/۰۰۸
	گروه پایین	۱۴/۵۸	۷/۴۸	۴۸			
تکمیل هدایت شده نقشه مفهومی	گروه بالا	۲۵/۴۶	۵/۷۱	۱۳	۳/۲۴	*۳۵/۵۵	۰/۰۰۳
	گروه پایین	۱۸/۰۳	۹/۲۱	۳۱			

* درجه آزادی با توجه به رعایت نشدن مفروضه همگنی واریانس‌های دو گروه مورد مقایسه تعدیل شده است

پایایی نمره‌گذاران: برای آگاهی از پایایی مصححان در شیوه‌های مختلف ارزشیابی مبتنی بر نقشه مفهومی از ضرایب نمرات ارائه شده توسط ۳ داور استفاده شد. پس از به دست آوردن این ضرایب‌های همبستگی، میانگین آنها به دست آمد که به صورت جدول ۵، با عنوان «ضرایب توافق نمره‌گذاران» ارائه شده است. همان‌گونه که مشخص است در نمره‌گذاری همه آزمون‌های نقشه مفهومی توافق نسبتاً کاملی بین نمره‌گذاران وجود دارد. برای تمام آزمون‌ها، 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 به ویژه در آزمون تکمیل هدایت شده نقشه مفهومی توافق بالایی بین نمره‌گذاران مشاهده شد.

جدول ۵: ضرایب همبستگی بین نمرات نمره‌گذاران روش‌های مختلف ارزشیابی مبتنی بر نقشه مفهومی

نوع آزمون	ضریب همبستگی	تعداد	سطح معناداری
آزمون تکمیل آزاد	۰/۹۱	۶۷	۰/۰۱
آزمون تکمیل هدایت شده	۰/۹۶	۴۴	۰/۰۱
آزمون ترسیم آزاد	۰/۸۲	۴۹	۰/۰۱
آزمون ترسیم هدایت شده	۰/۸۳	۵۶	۰/۰۱

مقایسه شاخص‌های اعتباربخشی انواع آزمون‌های نقشه مفهومی: در جدول ۶ همه آزمون‌های نقشه‌های مفهومی، از لحاظ چهار شاخص روایی محتوایی، روایی ملاکی، روایی افتراقی و پایایی نمره‌گذاران، با هم مقایسه شده‌اند. 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.

جدول ۶: مقادیر روایی ملاکی، روایی افتراقی و پایایی نمره‌گذاران چهار شیوه آزمون نقشه مفهومی

شیوه نمره‌گذاری	روایی محتوایی	روایی ملاکی	روایی افتراقی	پایایی نمره‌گذاران
آزمون تکمیل آزاد	۰/۸۱	۰/۳۴	۲/۷۴	۰/۹۱
آزمون تکمیل هدایت شده	۰/۸۷	۰/۴۲	۳/۲۴	۰/۹۶
آزمون ترسیم آزاد	۰/۷۰	۰/۷۵	۵/۳۳	۰/۸۲
آزمون ترسیم هدایت شده	۰/۸۰	۰/۴۰	۲/۹۹	۰/۸۳

بحث و نتیجه‌گیری

یافته‌های تحقیق نشان داد که آزمون‌های نقشه مفهومی به همان 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 مورد نظر 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 موردآزمون 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؛ زیرا اساساً ماهیت نقشه مفهومی به گونه‌ای است که از مفاهیم مهم حوزه مورد آزمون تشکیل می‌یابد. در زمینه آزمون تکمیل آزاد نیز نقشه مفهومی از قبل توسط مدرس ساخته شده بود و از این رو مفاهیم مهم در آن لحاظ شده بود.

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 هیئچی^{۴۲}، ۱۹۹۳؛ رایس، رایان و سامسون^{۴۳}، ۱۹۹۸). از لحاظ شدت همبستگی نیز مقادیر همبستگی به دست آمده در این پژوهش که به طور متوسط در حدود ۰/۵۰ بود با مقادیر همبستگی پژوهش‌های فرخ و کراپس^{۴۴} (۱۹۹۶) و واند رسون و هانگ^{۴۵} (۱۹۸۹) بسیار نزدیک است. البته وجود همبستگی متوسط بین آزمون‌های نقشه مفهومی و آزمون چهارگزینه‌ای را شاید به راحتی نتوان به حساب پایین بودن نسبی روایی ملاکی گذاشت؛ چرا که اساساً این دو نوع آزمون به شیوه‌ای متفاوت سطح دانش فراگیران را اندازه می‌گیرند. حتی شاید بتوان گفت نوع دانش مورد اندازه‌گیری نیز در این دو نوع شیوه متفاوت است. آزمون‌های مبتنی بر نقشه مفهومی هم اطلاعاتی در مورد این که دانش‌آموزان از محتوای موردنظر چه می‌دانند به دست می‌دهند. این کاری است که آزمون‌های چهارگزینه‌ای به خوبی از عهده آن بر می‌آیند. علاوه بر این، آزمون‌های نقشه مفهومی اطلاعاتی در مورد این که فراگیران چگونه به اطلاعات و دانش خود ساختار می‌بخشند، فراهم می‌کنند. بنابراین می‌توان گفت که نقشه‌های مفهومی جنبه‌هایی از یادگیری را اندازه می‌گیرند که آزمون‌های مرسوم آنها را به خوبی اندازه نمی‌گیرند.

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 وست^{۴۶} 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 می‌توانند تفاوت‌های مورد انتظار بین آزمودنی‌ها را به درستی بعد از مداخله یک آموزش تشخیص دهند.

در زمینه پایایی نمره‌گذاران نیز یافته‌های پژوهش نشان داد که بین نمره‌گذاران از لحاظ میزان نمره‌ای که بایستی به هر نقشه مفهومی تعلق گیرد توافق بالایی وجود دارد. این یافته با نتایج پژوهش شاولسون و همکاران (۲۰۰۵) همخوان است. یافته‌های آن پژوهش نشان از همبستگی بالایی (حدود ۰/۹۰) بین نمرات نمره‌گذاران داشت. نکته‌ای که بایستی مورد توجه قرار گیرد این است که پایایی نمره‌گذاران به شدت تحت تأثیر شیوه نمره‌گذاری است. یافته‌های پژوهش مک کلور^{۴۷} و همکاران (۱۹۹۹) نشان می‌دهد در شرایطی که از شش روش نمره‌دهی برای نمره‌گذاری نقشه‌های مفهومی استفاده شود مقدار پایایی این شش روش نمره‌گذاری می‌تواند دارای دامنه‌ای از ۰/۲۳ تا ۰/۷۶ باشد.

مقایسه شاخص‌های روایی و پایایی شیوه‌های چهارگانه ارزشیابی مبتنی بر نقشه مفهومی نشان می‌دهد که روش تکمیل هدایت شده در روایی محتوایی و پایایی نمره‌گذاران 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. براساس نظرات روئیز- پرمیو و شاولسون (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 آنهاست که البته می‌تواند در حوزه‌ای دیگر محدودیت محسوب شود. به‌عنوان مثال، بالا بودن پایایی نمره‌گذاران در روش‌های تکمیل، به دلیل 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 برخوردارند؛ 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 این گونه 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 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 کاربردی شیوه‌های متنوع این روش‌ها

و نیز نمره‌دهی 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 جدایی‌ناپذیر از فرایندهای بزرگ‌تری به نام یادگیری و تدریس است.

منابع

- رحمانی، آزاد. (۱۳۸۴). *تاثیر آموزش بر مبنای نقشه مفهومی در یادگیری درس فرایند پرستاری دانشجویان ترم دوم پرستاری*. پایان‌نامه کارشناسی ارشد چاپ نشده. دانشگاه علوم پزشکی تبریز.
- سیف، علی اکبر. (۱۳۸۲). *اندازه‌گیری، سنجش، و ارزشیابی آموزشی*. تهران: نشر دوران.
- سیف، علی اکبر. (۱۳۸۷). *سنجش فرایند و فرآورده تحصیلی: روش‌های قدیم و جدید*. تهران: نشر دوران.
- مصراآبادی، جواد، فتحی آذر، اسکندر، و استوار، نگار. (۱۳۸۴). *اثربخشی ارائه، ساخت فردی و ساخت گروهی نقشه مفهومی به عنوان یک راهبرد آموزشی*. فصلنامه علمی - پژوهشی *نوآوری‌های آموزشی*، ۱۳، ۳۱-۱۱.
- مصراآبادی، جواد، علوی، الهه، و استوار، نگار. (۱۳۸۵). *مقایسه اثربخشی استفاده از نقشه مفهومی به عنوان یک راهبرد یاددهی - یادگیری در پیشرفت درسی دروس مختلف*. چکیده مقالات و طرح‌های نوآورانه در همایش ملی نوآوری‌های آموزشی. تهران: مؤسسه پژوهش برنامه‌ریزی درسی و نوآوری‌های آموزشی.
- مصراآبادی، جواد. (۱۳۸۶). *اثرات نقشه مفهومی (ساخت و ارائه) و ویژگی‌های ورودی فراگیران بر بازده‌های شناختی-عاطفی یادگیری زیست‌شناسی*. رساله دکتری چاپ نشده. دانشگاه تبریز.
- مصراآبادی، 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.
- مصراآبادی، جواد. (۱۳۸۸). *امکان سنجی به کارگیری روش‌های ارزشیابی مبتنی بر نقشه مفهومی در زیست‌شناسی دوره متوسطه*. سازمان پژوهش و برنامه‌ریزی درسی.

- All, C.A., Huycke, L.I., & Fisher, M.J. (2003). Instructional tools for nursing education: Concept maps, *Nursing Education Perspectives*, 24(6), 311-317.
- Anderson, T.H., & Huang, S-C.C. (1989). On using concept maps to assess the comprehension effects of reading expository text. (ERIC Document Reproduction Service No. ED 310 368).
- Bezzi, A. (1996). Use of repertory grids in facilitating knowledge construction and reconstruction in geology. *Journal of Research in Science Teaching*, 33(2), 179-204.
- Biggs, J. B. (1989). Institutional learning and the integration of knowledge. In J. I. Balla, M.Gibson, & A. M. Chang (Eds.), *Learning in medical school* (pp. 21-38). Hong Kong: Hong Kong University Press.
- 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.
- Edmondson, K. M., & Novak, J. D. (1993). The interplay between epistemological views, learning strategies, and attitudes of college students. *Journal of Research in Science Teaching*, 32(6), 547-559.
- Farrokh, K. & Krause, G. (1996). *The relationship of concept-mapping and course grade in cell biology*. Retrieved from <http://www2.ucsc.edu/mlrg>.
- Hall, R.H. & O'Donnell, A.M. (1996). Cognitive and affective outcomes of learning from knowledge maps. *Contemporary Psychologist*, 21, 94-101.
- Hazel, E., & Prosser, M. (1994). First year university student's understanding of photosynthesis, their study strategies and learning context. *American Biology Teacher*, 56, 274-279.
- Horton, P.B., McConny, A.A., Gallo, M., Woods, A.L., & Hamelton, O. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77(1), 95-111.
- Kilic, G.B. (2003). Concept maps and language: a Turkish experience. *International Journal of Science Education*, 25, 1299-1311.
- Lin, C. Y., & Hu, R. (2003). Students' understanding of energy flow and matter cycling in the context of the food chain, photosynthesis and respiration. *International Journal of Science Education*, 25(12), 1529-1544.
- 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*. Ithaca, NY: Cornell University.
- 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., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, 28(1), 117-153.
- Novak, J. D., & Gowin, D. B. (1984). *Concept mapping for meaningful learning*. Cambridge: Cambridge University Press.
- Novak, J.D. (1990). Concept maps: A useful tool for science education. *Journal of Research in Science Teaching*, 27, 937–949.
- Novak, J.D. (1993). How do we learn our lesson? Taking students through the process. *The Science Teacher*, 60, 50–55.
- Pearsall, N. R., Skipper, J. E. J., & Mintzes, J. J. (1997). Knowledge restructuring in the life sciences: A longitudinal study of conceptual change in biology. *Science Education*, 81(2), 193-215.
- Potelle, H., & Rouet, J.F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58, 327-345.
- Prosser, M. (1987). The effects of cognitive structure and learning strategy on student achievement. In J. T. E. Richardson, M. W. Eysenck, & D. W. piper (Eds.), *Student learning* (pp. 29-38). Milton Keynes, UK: Open University Press.
- 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.
- Shern, D., Trochim, W.M., & LaComb, C.A. (1995). The use of concept mapping for assessing fidelity of model transfer: An example from psychiatric rehabilitation. *Evaluation and Program Planning*, 18, 143–153.
- 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.

زیر نویس

1. Hazel & Prosser
2. Lin & Hu
3. Novak
4. Biggs
5. Edmondson
6. Bezzi
7. Musonda,
8. Prosser
9. Pearsall, Skipper & Mintzes
10. Roundhouse diagrams
11. Vee diagrams
12. Node-link diagrams
13. Knowledge map
14. Semantic networks
15. Mental map
16. Concept tree
17. Concept network
18. Concept maps
19. Hall & O'Donnell
20. Potelle & Rouet
21. Horton, McConny, Gallo, Woods & Hamelton
22. All, Huycke & Fisher
23. Shern, Trochim & LaComb
24. West, Park, Pomeroy, & Sandoval
25. Shavelson, Ruiz-Primo & Wiley
26. Kilic
27. Validity
28. Reliability
29. McClure, Sonak & Suen
30. Ruiz-Primo & Shavelson
31. Multistage cluster sampling
32. Evaluation research
33. Rossi
34. Chularut & DeBacker
35. Relational
36. Structural
37. Similarity
38. Content validity
39. Criterion validity
40. Discriminative validity
41. Interrator reliability
42. Liu & Hinchey
43. Rice, Ryan & Samson
44. Farrokh & Krause
45. Anderson & Huang
46. west
47. Mc clure