الفرق بين الملاحظات البنائية والتقويم التجميعي وتركز على تحديد نقاط الضعف في تعلم الطالب لمساعدته على تحسينها، في حين أن التقويم التجميعي يهدف إلى تقديم تقييم شامل للأداء الكلي بعد فترة معينة. الأول موجه لتحسين التعلم الفردي والتحديد الدقيق للنقاط التي تحتاج للتحسين، بينما الثاني يهدف إلى تقديم تقييم عام وشامل لأداء الطالب. مثال توضيحي: مادة (الرياضيات) الحساب الغرض من التقويم التجميعي هو تحديد قدرة وسرعة الطالب على ترجمة المسائل اللفظية إلى حلول كمية عند إجراء القسمة. وأيضًا تحديد مستوى الطلاب وتقديم تقرير لأولياء الأمور والهيئة الإدارية. أو عدم كفاية المفردات اللغوية. يمكن للمعلم الدمج بين التقويم التجميعي والتقويم البنائي في اختبار واحد، لذا فالتوقيت يلعب دورًا أيضًا حيث تجرى الاختبارات البنائية بشكل متكرر خلال فترات أقصر مقارنة بالتجميعية. فالتقويم البنائي يُستخدم لتحسين تعلم التلميذ بعد تعليم مهارة جديدة، فالاختبارات القصيرة التي تُعطى في نهاية الوحدات التعليمية القصيرة تميل لأن تكون بنائية بطبيعتها. ومن الفروق بين التقويم التجميعي والبنائي هو مستوى التعميم، حيث يقوم مبدأ تايلور بالتركيز على التوصيفات السلوكية المرغوبة (كما ورد ذكره في الفصل الثاني)، بينما جانيه ركز على التفاصيل الأولية للمقدرات. فالتقويم البنائي يلاحظ المتطلبات الأولية للعمل، خطوات عامة في بناء الاختبارات بالرغم من أن تفاصيل تكوين هذه الجداول تم شرحها في مواضع أخرى، القاعدة الرئيسية هي النظر إلى مجال المحتوى من منظور الموضوعات الكبيرة للمادة، مثال ذلك أنه في مادة العلوم يمكن استخدام الأبعاد المتجانسة في منهج العلوم البيولوجية، أو تقسيم المادة إلى أنظمة فسيولوجية ووظائف بيولوجية وعلم الحيوان. 2- بناء مفردات ومصطلحات للاختبار التي تتناسب مع الجدول الذي تم بناءه. 3- اختيار عينات مناسبة للاختبارات. سواء من المحتوى أو السلوكيات المعقدة. يتعين على المعلم تحديد الخلايا المهمة لعينته بناءً على أهدافه، سواء كانت عينة عشوائية لكل المفردات أو عينات تركز على أهداف محددة. هذا يساعد في وضع الطلاب في مستويات مناسبة أو منحهم شهادات بناءً على السلوكيات النهائية المطلوبة. 4-تدور حول تجميع المفردات في الاختبارات بناءً على خطة منظمة. يُفضل ترتيبها من السهل إلى الصعب. من المفيد أيضًا تجميع أنواع المفردات المختلفة مثل الاختيار من متعدد والمزاوجة وصواب أو خطأ حسب النوع لتسهيل إعطاء التعليمات للمعلم والتلميذ. 5- ضع خطة لإعطاء الدرجات بحيث تقدم معلومات مفيدة لأغراض الاختبار. يمكن جمع الدرجات الفردية للحصول على تقييم شامل لمهارات أو قدرات محددة. من أجل التغذية الراجعة، النظام يجب أن يسمح بمقارنة مخرجات مجموعات أو معالجات مختلفة ويعكس الفروق في القدرات والمعارف. التعليمات يجب أن تحدد الوقت المخصص للاختبار وأجزاءه، استخدامات الاختبار وبناءه: الاختبار التجميعي يمكن أن يخدم وظائف متنوعة، وأنه عند بناء وتصحيح الاختبارات بشكل صحيح، يجب تحديد الأهداف المقصودة قبل بناء الاختبار لضمان استفادة أكبر من النتائج. لا يمكن تطبيق مخرجات الاختبار بنفس الطريقة في جميع الاستخدامات والمجالات. ـ ومقارنة مخرجات مجموعات مختلفة. وذلك من المدرسة الابتدائية إلى مرحلة ما بعد التخرج. الهدف هو تصنيف التلاميذ حسب مستوى تعلمهم مقارنة ببقية التلاميذ. لتحقيق انتشار فعّال، صعوبة المفردة تتعلق بنسبة الطلاب الذين يجيبون عليها بشكل صحيح، لتحقيق توزيع عادل، ولتعيين المستويات بشكل فعال، يجب تجريب كل اختبار على تلاميذ مماثلين قبل استخدامه. إذا كان ذلك غير عملي، يمكن أيضاً تجميع مفردات الاختبار ببطء لضمان توازنها في الصعوبة والتمييز. إن التربية الموجهة نحو التمكن تتطلب تغيير في الوقت ومنهجية التعليم لتحقيق هذا الهدف. في هذه الحالة، قد يشير ذلك إلى خلل في منهجية التعليم. مع تحسين التعليم، هذا هو الرأي الذي يتبناه التقويم البنائي. يجب الانتباه إلى بناء خطوات معينة مثل الخطوات 3، و5، كما هو موضح في جدول مواصفات العلوم البيولوجية للمرحلة الثانوية. من المهم تحسين المفردات وترتيبها من السهل إلى الصعب. ومع ذلك، لتوضيح النقطة بشكل أفضل، تم عرض مثال على معلم قرر أن يركز الامتحان التجميعي في نهاية المقرر على استخدام المهارات في فهم مشكلات العلوم وعرض العلاقات بين تراكيب المعرفة، مع التركيز على محاور رئيسية مثل التطور والتكامل الحيوي والبيئي. ولمنح الشهادات. الدرجات المتعددة تعطي معنى أكبر مقارنة بالدرجة الواحدة. والمعيار القائم على الآراء. بينما تعتمد المعايير المقدرة على توقعات الخبراء. المستويات القائمة على تحليل المهمة أو أفكار التمكن تركز على الأداءات المطلوبة لنجاح الطالب في مهام معينة. إجازة المهارات والقدرات أو المهارات الرياضية للإحصاء التربوي، أو مهارات معمل العلوم. في هذه المواقف، ويمكن أن تكون الأحكام الذاتية للخبراء بديلاً حتى تتوفر الأدلة التجريبية. أهمية استخدام الاختبارات التجميعية لتنبؤ نجاح الطلاب في المقررات اللاحقة. يؤكد على أن تصميم هذه الاختبارات يتطلب استشارة متخصصين تقنيين، لكن يجب على المدرس تحديد أهداف الامتحانات. يوضح النص أن الاختبارات التجميعية في نهاية المقرر يمكن أن تتنبأ بالنجاح في المقررات ذات الصلة، بشرط وجود دليل على العلاقة بين الأداء في الاختبار والمقرر اللاحق. لكن هناك

Summarized by © lakhasly.com

عوامل عديدة غير معروفة تساهم في هذه العلاقة. نقطة البداية في تدريس مقرر لاحق أهمية الاختبارات التجميعية التي تُعطى في نهاية العام الدراسي في تحديد مستوى الطلاب لبداية تدريس المقرر التالي، التغذية المرتجعة للتلاميذ يجب أن تكون الاختبارات مبنية بشكل جيد لتوجيه التلميذ نحو تحسين نواحي النقص. من المهم أن تكون مفردات الاختبار دقيقة وتعكس جوانب متنوعة من المحتوى والسلوك. التغذية المرتجعة تكون أكثر فعالية عندما يتم استخدامها لتحسين عملية التعلم والتدريس بشكل مستمر.

الهدف هو تقويم طرق مختلفة أو مواد أو أنواع مختلفة من التلاميذ. لتحقيق تفسيرات مفيدة، إذا كان الغرض هو مقارنة المخرجات، يمكن تقسيم الفصل إلى عينات متعددة وإجراء اختبارات منفصلة لكل منها، والتي تغطي خلايا السلوك والمحتوى كما هو مبين في الجدولين 4–1 و4–2. في الجدول 4–2، التسع مفردات الأولى تختبر معرفة المصطلحات المتعلقة بالتطور، بينما المفردات من 10 إلى 14 تختبر استرجاع الاتجاهات والتتابع في المحتوى المرتبط بالتباعد في الأنواع. باقي المفردات تختبر ثلاث خلايا أخرى. وذلك يعتمد على تركيز صانع الاختبار على سلوك أو محتوى معين. مما يوفر نوعًا من المراجعة ويعرف بالمفردات المرتكزة. هذا التداخل يسمح بمقارنة أداء المجموعات. فهذا يثير التساؤل عن مدى تمثيل هؤلاء المفحوصين للمجموعة بأكملها. بتقسيم الفصل إلى أربع فئات جزئية تحتوي كل منها على 11 مفردة، مما يغطي مجموعة واسعة من الأهداف. يمكن زيادة عدد المفردات إلى 40–60 مفردة حسب الحاجة، هذا يساعد في تحسين دقة التقييم وتغطية المحتوى بشكل أفضل. أن من مميزات أخذ العينات لغرض المقارنات هو الوضوح في التقييم، بدلاً من ذلك، مع مراعاة عدد المفردات وأوقات الاختبار المتعددة لتحقيق تغطية شاملة ودقيقة. خطوات تنقيح الاختبارات: 1. إعطاء الطلاب الاختبار بشكل فردي وتسجيل سلوكهم: يعد أداة فعّالة لتقييم تعلم الطلاب ويسمح بجمع بيانات مفصلة عن استجاباتهم، ويقدم وضوحًا حول فهمهم للتعليمات والأسئلة. يتم تطبيق هذا الأسلوب بشكل ميداني في مستويات صفية متعددة لتحديد نقاط القوة والضعف في تعلم الطلاب، 2. عرض مفردات الاختبار وجدول المواصفات على محكم كفء وعلى متخصص أخر في المجال والطلب منه وضع الجزئية من الجدول التي يربطها بكل مفردة، إن كلا من صانع الاختبار والمحكم يجدوا ان نظام الكروت من العوامل المساعدة وتساعد أيضا في وضع بنك مفردات لبناء اختبارات جديدة. ويعتبر المتوسط ضرورياً للإجراءات الموصوفة التي تعنى بالانحراف المعياري والثبات وتعتبر أيضا مقاييس النزعة المركزية ذات فائدة في مقارنة أداء مجموعتين مختلفتين في نفس الاختبار. الوسيط: هو القيمة المتوسطة للدرجات في اختبار معين بعد ترتيبها تصاعدياً (من الأصغر إلى الأكبر). 4. احسب الانحراف المعياري لدرجات الاختبار: ويعد هذا المقياس ذو فائدة مع المتوسط في إجراءات الحصول على قيمة للثبات، والانحراف المعياري هو مقياس يُظهر مدى تشتت أو انتشار درجات الطلاب حول المتوسط الحسابي. فهذا يعني أن معظم درجات الطلاب قريبة من المتوسط. وإذا كان مرتفعاً، فهذا يعني أن هناك اختلافاً كبيراً بين درجات الطلاب. 5. احسب مقياس الصعوبة لكل مفردة: هو مقياس يُعبّر عن نسبة الطلاب الذين أجابوا على البند بشكل صحيح. يتم تقسيم عدد الطلاب الذين أجابوا بشكل صحيح على العدد الكلي للطلاب الذين حاولوا الإجابة على البند. ويعتبر معرفة مستوى صعوبة كل بند مفيدًا في تحسين جودة الاختبار. البنود التي تكون سهلة أو صعبة للغاية قد لا تكون مفيدة لأنها لا تساعد في تحفيز الطلاب بشكل كافٍ، مع الحفاظ على الهدف التعليمي للبند. 6. حدد قوة التمييز لكل مفردة: معامل التمييز في أبسط مستوياته يتضمن الحصول على معامل الصعوبة لكل من المجموعتين من التلاميذ: اللذين أدوا بطريقة جيدة في الاختبار وبطريقة رديئة، وإذا لتميز المفردة بطريقة فعلية فيمكن حذفها، أما ثبات الأداء عبر الزمن، صيغة "كوردر ريتشاردسون" هي إحدى الطرق المستخدمة لحساب معامل الثبات (الصدق الداخلي) للاختبارات التي تتكون من أسئلة ثنائية الإجابة (صح أو خطأ، تُستخدم هذه الصيغة لتقدير مدى اتساق الأسئلة في قياس نفس الخاصية أو المفهوم. الخصائص الفنية للاختبارات: وفيما يلي جدول يوضح أنواع الصدق والفروقات بينها. نوع الصدق الوصف أمثلة تفصيلية الشروط اللازمة لتحقيق هذا النوع من الصدق الأدوات صدق المحتوى اختبار رياضيات يحتوي على أسئلة تتعلق بمفاهيم دروس الرياضيات الحالية يعتبر الاختبار صالحاً إذا وافق المحكمون على أن معظم الأسئلة تتعلق بمحتوى المقرر الدراسي بنسبة اتفاق عالية 75% أو أكثر. محكمون ذو كفاءة جدول المواصفات يهتم بالتأكد من أن مكونات الاختبار تقيس سمة أو مفهوم معين بشكل مترابط، مقارنة أداء التلاميذ في أنواع مختلفة من الأسئلة التي تقيس نفس القدرة، لتقييم قدرتهم على تطبيق المعرفة في مواقف جديدة. تحليل العناصر الإحصائي الصدق التنبؤي يشير إلى قدرة الاختبار على التنبؤ بأداء الطالب في اختبارات لاحقة أو مواقف مشابهة. أن تكون العناصر مرتبطة ارتباط وثيق بالأهداف التعليمية الفعلية للموضوع الدراسي. معامل الارتباط الصدق المتلازم يقيس توافق نتائج اختبار مع نتائج اختبار آخر يقيس نفس السلوك أو القدرة، اختبار جديد للذكاء يتم مقارنته مع اختبار "بينيه" التقليدي يمكن أيضاً

Summarized by © lakhasly.com

استخدامه لمقارنة قدرة الطلاب على الكتابة باستخدام مهارات تصحيح النصوص مقارنة بمهارات الكتابة المباشرة للتأكد من إذا أظهر التلاميذ ترتيباً مماثلاً في كلا الاختبارين، فإن الصدق المتلازم يتحقق معامل الارتباط هناك معيارين عامان يمكن استخدامهما لاختيار العينة المناسبة من الأسئلة التي تتيح تقييم القدرات المعطاة للطلاب، وهما: هذه الجداول تساعد في تنظيم أنواع الأسئلة وموضوعاتها، السؤال الجوهري في هذا المعيار: كيف نحدد الأوزان النسبية لمختلف مجالات المادة الدراسية؟ التقدير الشخصي: يعتمد بعض المدرسين على حكمهم الشخصي لتحديد أهمية كل مجال، بناءً على ما يعتقدون أنه أكثر فائدة أو أهمية للطلاب، أو ما تركز عليه المراجع بدلالة عدد الصفحات المكرسة للموضوع أو ماذا قالت سلطة أخرى. والاستناد إلى حكم سلطة محلية معينة مثل: رئيس القسم أو مشرف أو لجنة تنظيم المنهج. الترابطات المنطقية: يركز البعض على الروابط بين مختلف مجالات المادة، فالمجالات التي ترتبط بشكل وثيق بمواضيع أخرى تعتبر أكثر أهمية، وهم أيضاً يهاجمون مشكلة إعطاء القيمة لمجالات المحتوى. التحليل العلمي: يلجأ بعض المدرسين إلى تحليل الأبحاث والدراسات العلمية لتحديد المفاهيم الأكثر شيوعاً والأكثر أهمية في المجال. 2. التساوي بين المهام المتضمنة: إذا علم المدرس بأن بعض المصطلحات والرموز المستخدمة في مادته تستخدم في مجال آخر، فإنه يتفادى عمل عينات من كليهما إذا كان محدداً بعدد من المفردات التي يمكن استخدامها. ● لا تعني بالضرورة الأعداد المتساوية للمفردات أنّ كل جزء يسهم بنفس الوزن للدرجة الكلية. ● إذا كان هناك تفاوت كبير في أداء الطلاب في جزء معين من الاختبار، فإن هذا التفاوت سينعكس بشكل واضح في النتيجة النهائية. ● لا يمكن الجزم بشكل كامل بمدى ملائمة الأوزان النسبية المحددة قبل الاختبار إلا بعد رؤية نتائج الطلاب الفعلية. ● القيمة المعطاة لمركبات أي اختبار تعكس أوزان هذه المركبات في الدرجة الكلية في حال وجود شرطين هما: تباين متساو في إجابات المفحوصين، وهذا نادر الحدوث في أي اختبار. الخاصية الفنية الثانية للاختبارات الثبات: إذا تم تطبيق نفس الاختبار على مجموعة من الطلاب مرتين (مرة في الصباح ومرة في المساء)، وإذا لم تكن نقول بأن النتائج لم تكن متسقة. في المثال الأول: نفس الاختبار يُعطى مرتين خلال فترة زمنية قصيرة جداً يعكس اتساقاً في إجراءات عملية الامتحان أو الأداة، مما يعكس على استجاباتهم، وطريقة منح الدرجات كجزء من إجراءات الامتحان قد تسبب في تباين النتائج لنفس الاستجابات عند التقييم. مما يجعلها أقل عرضة للتباين، ولتحسين الإتساق في منح الدرجات، ينصح باستخدام مفتاح تصحيح موحد، في المثال الثالث: استقرار السمة أو القدرة في التلاميذ: يعكس هذا المثال الفارق الزمني الكبير نسبياً قبل إجراء الاختبار الثاني بنفس العينة من المفردات أو مفردات مختلفة؛ فإذا كنا نتعامل مع اهتمامات التلاميذ لأغراض النجاح المهني، فإننا نتوقع الاستقرار عبر الزمن، وإذا كنا نتعامل مع القدرة الكتابية خلال فترة التعليم، ويجب أن يؤخذ الثبات في الاعتبار عند استخدام الاختبارات في منح الشهادة أو في التنبؤ بالنجاح. العلاقة بين الصدق والثبات: ● الصدق يعتمد جزئياً على الثبات، بحيث يستطيع الفرد بناء اختبار غير صادق، ● عند وجود اتفاق عام بين محكمين مناسبين على أن فئة معينة من المفردات تقيس القدرة على تفسير البيانات، وعليه يبنى صدق المحتوى، ● إن الثبات يحد من الصدق؛ بعض الاقتراحات لتحسين ثبات اختبار ما: 2. إن زيادة الاتفاق بين المصححين أو مصحح واحد في مناسبات مختلفة سوف يحسن من الثبات، الموضوعية في تصحيح الاختبارات ليست مقتصرة فقط على التقويم التجميعي، بل تمتد لتشمل أيضاً الاختبارات التشخيصية والتقويم البنائي، الموضوعية هنا لا تتعلق فقط بجودة التصحيح، بل تؤثر كذلك على ثبات وصدق النتائج. وعلى الرغم من إعداد الاختبارات بناءً على جدول المواصفات لتحقيق أهداف محددة، فإن عدم تطبيق الموضوعية في التصحيح قد يؤثر سلباً على صدق المحتوى، إذ قد يؤدي التحيز في التقييم إلى إعطاء درجات غير متساوية لنفس السلوك للطلاب المختلفين، مما يضعف كل من الثبات والصدق في نتائج الاختبار. ليتحقق من دقة العبارات حتى عند وجود ترابط قوي بين المفردات والمواصفات. في حال عدم الاتفاق على مفردة معينة، قد يكون هناك حاجة إلى تعديل الصياغة لضمان دقة المحتوى المستهدف، ومن المؤكد أن مفتاح الإجابة للاختبارات الاختيارية ضروري، بحيث لا يكون هناك اختلاف في الدرجات إلا إذا كان نتيجة خطأ في تطبيق المفتاح، مثل التكميل أو كتابة مقالة،

Summarized by © lakhasly.com