



مجلة إضاءات علمية

مجلة علمية أكاديمية تصدر عن جامعة الجزائر 2



تصديق اختبارات التقويم وفق المقاربة بالكفايات

Certification of competency assessment assessment tests

الدكتور : أحمد بلاطي

جامعة الجزائر (الجزائر)

ahmedbellati17@gmail.com

المرسل: أحمد بلاطي

النشر: 2022/08/30

القبول: 2022/08/20

الارسال: 2022/04/18

الملخص:

تؤدي المقاربة بالكفايات إلى تقويم مكتسبات التلاميذ عبر وضعيات مركبة تستدعي إنتاجا مركبا ينجزه التلميذ. ومن ثم فإن التقنيات الكلاسيكية المعتمدة في التصديق على اختبارات التقويم، تبدو قاصرة عن تحقيق المؤمل، مما يستدعي تبني مقاربات أخرى للتصديق القبلي والبعدي في إطار احترام مقتضيات الصحة والنجاعة فضلا عن الملاءمة، في المقام الأول من الاعتبار. الكلمات-المفاتيح: تقويم المكتسبات، تقويم الكفايات، الملاءمة، الصحة، النجاعة، المقاربة بالكفايات.

Résumé:

L'approche par les compétences conduit à évaluer les acquis des élèves à travers des situations complexes nécessitant une production complexe de la part de l'élève. Les techniques classiques de validation des épreuves d'évaluation ne peuvent dès lors pas être utilisées telles quelles. D'autres approches de validation, qu'elles soient a priori ou a posteriori, devraient être adoptées, dans le respect des exigences de validité et de fiabilité, mais aussi et avant tout de pertinence.

MOTS CLES : *Evaluation des acquis, évaluation des compétences, pertinence, validité, fiabilité, approche par les compétences*

مقدمة

لقد صارت الممارسات البيداغوجية على الصعيد الدولي تستند أكثر على المقاربة بالكفايات المباشرة للمقاربة بالمحتويات أو الأهداف :

-المقاربة بالمحتويات تنظر إلى التعليم باعتباره لائحة من المواد والمحتويات الدراسية الملقنة أو التي يتعين نقلها إلى المتعلم؛

-المقاربة بالأهداف تعتبر السلوكات الملاحظة والمبنيّة والمنفصلة الواجب تنميتها لدى المتعلم، مدخلا ضروريا لها؛

-المقاربة بالكفايات فتبحث في تطوير قدرات المتعلمين على تعبئة مجموعة مندمجة من الموارد⁽¹⁾، من أجل حلّ وضعية مشكلة، تنتمي إلى فئة معينة من الوضعيات (دوكيتيل 2000، 2001 أ وب؛ دولز وأولاني 2002؛ فوريز 1999؛ جبرار 2005؛ جونايرت 2002؛ لاسني 200؛ لوبوطيرف 1994؛ لوجوندر 2001؛ بيرينو 1997؛ بوايري برولكس 1999؛ راي 1996؛ راي وكاريت ودوفرانس وكان 2003؛ روكيرس 2000، 2003؛ سكالون 2004؛ طارديف وآخرون 1995؛ تيلمان 2000).

تستدعي هذه الممارسات المختلفة ممارسات تقويم مختلفة (De Ketele & Dufays, 2003) :

- من منظور المحتويات، يبني التقويم على أخذ عينة تمثيلية من المحتويات المرجعية الذي تم تدريسها؛
- من منظور الأهداف، يبني التقويم على أخذ عينة تمثيلية من الأهداف الخاصة والإجرائية وتوليد عينة من الأسئلة تترجم هذه العينة من الأهداف على أفضل وجه.

- من منظور الكفايات، يبني التقويم من اقتراح وضعية أو وضعيات مركبة، تنتمي إلى عائلة الوضعيات التي تحددها الكفاية، وستتطلب، من جانب التلميذ، إنتاجاً مركباً بحد ذاته لحلّ الوضعية.

بين الملاءمة والصحة والنجاعة

ستكون الاختبارات المنجزة لغايات التقويم، تبعا لكل واحدة من هذه المقاربات، مختلفة لا محالة، ومع ذلك، فإن الأسئلة المطروحة في هذه الاختبارات هي تلك الأسئلة الكلاسيكية المتعلقة بكل ما نستقيه من معلومات: فهل هي ملائمة وصادقة وناجعة؟

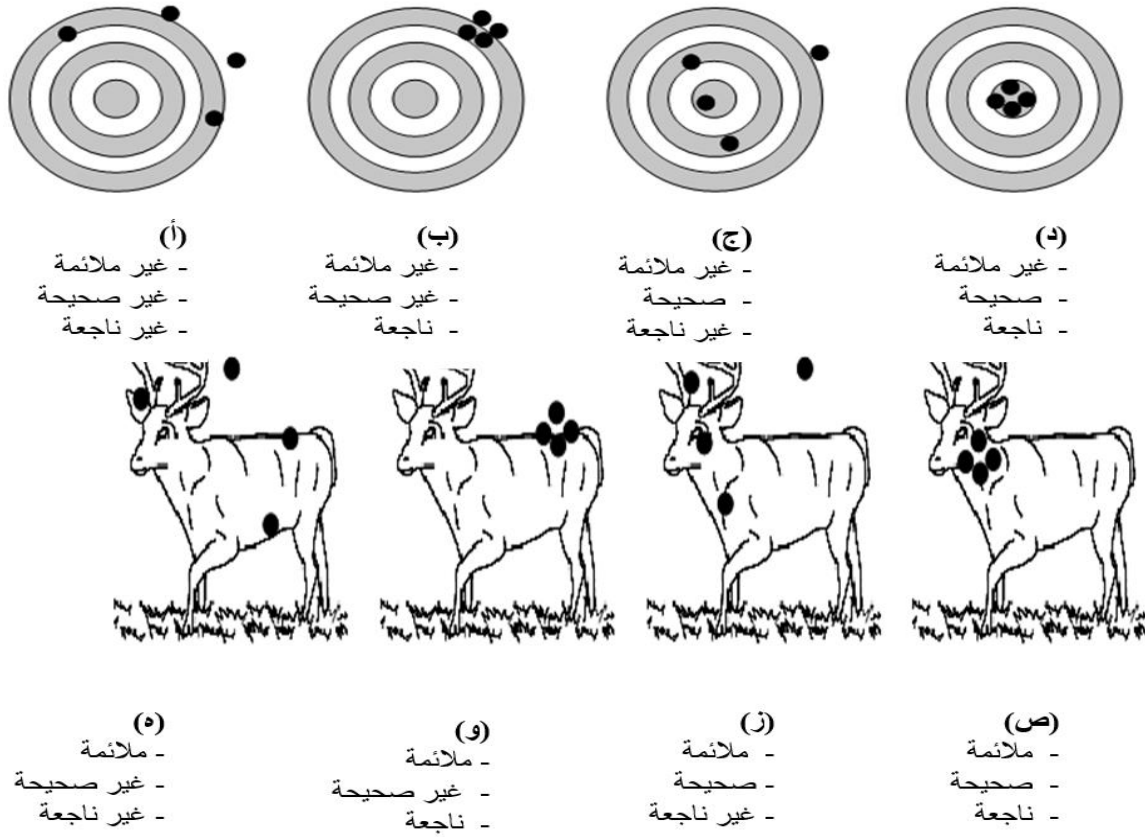
- الملاءمة هي خاصية تناسب الاختبار بشكل أو بآخر، بحسب انخراطها في منحي الأهداف المرجوة (De Ketele, Chastrette, Cros, Mettelin & Thomas, 1989) أي درجة "توافقها" مع العناصر الأخرى للمنظومة

التي تنتهي إليها (Raynal & Rieunier, 2003)

- الصحة هي درجة المواءمة بين ما نعلن فعله (تقويم بعد من الأبعاد) وما نقوم بفعله حقا، أو بين ما تقيسه الأداة، وما تسعى إلى قياس. (Laveault & Gregoire, 2002)

- النجاعة هي درجة الثقة التي نغذقها على النتائج الملاحظة فهل ستبقى هذه النتائج هي نفسها إذا استقينا المعلومة في لحظة مغايرة، متوسلين بأداة أخرى، أو إذا قام بالعمل شخص آخر، ألخ؟ إنها تنبئنا بدرجة العلاقة القائمة بين النقطة المحصل عليها والنقطة الصحيحة (Cardinet & Tournier, 1985; Laveault & Gregoire, 2002)

(Gregoire, 2002) &، ذلك أنه يجب ألا يغرب عن بالنا أن النقطة الصحيحة ليست إلا ضرباً من الوهم باعتبارها نقطة التقاطع المؤمل بمنأى عن المقومين وعن ملابسات التقويم. هذه الأبعاد المستقلة نظرياً تشكل مجتمعة الشروط الضرورية للحصول على اختبار تقويمي جدير بالصفة، ولكنها إذا نظر إلى كل شرط بمعزل عن الآخر فإنها لن تكون أبداً كافية؛ فالأداة الناجعة يحتمل ألا تكون صحيحة، وملائمة، وقد تكون وسيلة ملائمة ولكنها غير صحيحة ولا ناجعة، وهكذا دواليك. وفي مثال "الرمي بالقوس" الذي اقترحه لافو و كريكوار (1997-2002)، لننتخِلُ أننا نريد تقويم كفاية صياد طرائد يستعمل سلاح القوس، يتعين عليه كي ينجح في مسعاه أن يكون قادراً على تصويب رمي مركز في المركز على هدف متحرك، وفي هذا الصدد يكون بإزاء اختبارين في الوقت نفسه: من جهة، تسديد رمية من خمسة سهام نحو هدف تقليدي ثابت، ومن جهة أخرى، إنجاز رمية من خمسة سهام مصوبة نحو هدف متحرك مشابه للطريدة. وعندئذ تكون الحصيلة هذه السلسلة من الوضعيات:



في الوضعية (أ): الرمية غير ناجعة لأنها مشتتة على سطح الهدف الثابت. وهي غير صحيحة نظراً لغياب الدقة في التصويب مما لا يبنى قطعاً ببلوغ الهدف المنشود. ثم إن الرمية أخيراً غير ملائمة ما دامت لا تحقق المبتغى، ولا تقيس في جميع الأحوال ما تصرح بقياسه.

في الوضعية (ب): الرمية مركزة إلا أنها تخطئ الهدف. وهي ناجعة (مادام القياس يثبت أن الرامي يجيد الرمي المركّز) لكنها غير صحيحة، لأن القياس لا يسمح لنا بالتحقق مما إذا كان الرامي قادراً فعلاً على إصابة الهدف.

في الوضعية (ج): تحقق الهدف بما أن الرامي أصاب المحور في الصميم، والقياس صحيح، بيد أنه يلبث مع ذلك غير ناجع مادام لا يسمح لنا بتوقع تكرار النتيجة.

في الوضعية (د): الرمي مكرر بصورة مركزة في صميم الهدف إلا أن هدف التسديد غير دقيق (بالنظر إلى نوع الكفاية المستهدفة). إذن فالقياس ناجع وصحيح بيد أنه غير ملائم.

تمثل الوضعيات المتراوحة بين (هـ) و (و) وضعيات ملائمة بالنسبة إلى الكفاية المستهدفة. ولئن كانت الوضعيات (هـ) و (و) (ز) تعاني من مشاكل متعلقة بالصحة و/ أو النجاعة فإن الوضعية (ص) تقدم على العكس من ذلك قياسا يستوفي في الآن معا شروط الملاءمة والصحة والنجاعة، ومن ثم فهي الوضعية المثلى التي تشهد بأن الرامي كفء.

تسترعي الوضعية (د) الاهتمام من حيث كونها تقدم صورة عن الكيفية التي تجري بها الاختبارات الكلاسيكية نظير تلك التي تستعمل في الاختبارات الدولية من نوع تيمس TIMMS أو ييزا PISA، على الأقل من حيث انخراطها في نظام تربوي مبني على المقاربة بالكفايات. هذه الاختبارات تقترح مجموعة من المواضيع تستند على بنية متينة ومنبثقة أساسا من المقاربة بالمحتوى و/ أو الأهداف، إنها صحيحة وناجعة ولكنها ليست ملائمة بالنسبة للمقاربة بالكفايات وبالتالي بالنسبة لغايات النظم التربوية التي بنيت على أساس هذا التوجه. فإذا كانت تسمح بتقويم الموارد الضرورية (معارف إعادة الإنتاج ومهارات الأداء) فإنها لا تسمح (وإن بنحو ضئيل) بتقويم ملكة تعبئة هذه الموارد الملائمة لحل المشكلات أو إنجاز المهام المركبة.

لذا يجب الاقتدار على إعداد اختبارات تطابق الوضعية (ص)، ولكن من المحتمل أن عددا من الاختبارات الموضوعية حاليا أقرب ما تكون من الوضعية (د) أو (و)، إذ يتعين عليها أن تكون ملائمة لأنها تعرض وضعية مركبة رغم صعوبة إثبات صحتها لأن الطرق المتبعة لإثبات الصحة أو المصادقية إما غير واضحة، أو ببساطة غير معروفة. ولعله من المتعذر على أية حال استعمال الطرق الحرفية المعتمدة في النظرية الكلاسيكية للقياسات. ولئن كنا واعين بأن هذا الوضع وبعض المقترحات المقدمة فيما سيتلو من هذا المقال قد يثيران بعض الجدل، فالأمل معقود على أن تشكل الأسئلة المطروحة محور أعمال الخبراء المختصين في التقويم.

حدود النظرية الكلاسيكية للدرجات

كما أوضح ذلك "كارديني" منذ 1973، فإن هناك عنصرا هاما في هذا الصدد يجب وضعه في الحسبان، يتمثل في وجود حدود منهجية هامة أثناء التصديق على اختبارات التقويم البيداغوجية بغض النظر عن المقاربة المرجعية. تكمن هذه الصعوبة في كوننا نتوسل بتقنيات موظفة أصلا في القياسات النفسية *psychométrie* مع أن مجال اشتغالنا هو القياس التربوي *édumétrie*، وهي لا تتوفر على أدوات متطورة تنسجم وخصوصية حقل التربية.

فالاختلاف يكمن أساسا في التوزيع المنتظر للنتائج مجسدا في منحني مرجعي. فنحن عندما نستعمل رانز القياس النفسي، نتوقع أن تكون المجموعة المستهدفة موزعة وفقا لقانون طبيعي أو لمنحنى كوص Gauss، حيث هناك قلة من النقط المتدنية، وقلة من النقط العالية بالنسبة لغالبية النقط المتوسطة، وحيث متوسط النقطة يتطابق مع المنوال والوسيط. فالاعتبار الذي يحكم هذا المنطق السيكولوجي يتمثل في كونه يتوق إلى وصف مجموعة معينة، مع تحديد موقع فرد ما في التوزيع الإجمالي الذي يضاهي إلى حد بعيد

منحنى كوص. توزيع من هذا القبيل يظهر بخاصة إذا كانت الظاهرة المدروسة وليدة جملة من العوامل المتداخلة، ويختفي إذا كانت الظاهرة المدروسة خلاصة عامل نسقي قوي. وحيث إن التدخل البيداغوجي هنا يكون أقوى من العوامل العشوائية الفردية أو الظرفية المتعددة، وجب بالضرورة أن يكون المنطق التربوي أشد اختلافا نظرا لأنه لا يستهدف وصف المجموعة بقدر ما يسعى إلى التأثير فيها. فالهدف من التربية المدرسية أن يتعلم التلاميذ أو بعبارة أدق أن يتعلموا قاطبة، ومن ثم فالتوزيع المنتظر في نهاية المسار التعليمي - التعليمي لا ينبغي منطقيا أن يكون "طبيعيا"، بل مطابقا لما يمكن تسميته منحنى على شكل I يدل على تحصيل السواد الأعظم من التلاميذ للأهداف الأساسية المسطرة خلافا لأقلية منهم. وبمعنى ما فـ"المربي" هو ذلك الذي عليه بفضل مجهوده التربوي أن يكذب "عالم النفس". ومما يؤسف له، أن هذا المنطق لا يتم استيعابه ولا يأخذه المربون أنفسهم بعين الاعتبار، وهكذا فقد أثبتت عدة دراسات القانون المسمى "بوستميس Posthumus" الذي تمت صياغته بدءا من سنة 1947، والذي بموجبه يميل المدرسون بصورة لا شعورية في معظم الأوقات، إلى إعادة إنتاج توزيع على شكل جرس، في أقسامهم، مما يعكس منطقا انتقائيا (كراهي 1996، ودولاندشير 1980). وليس يعزب عن الذهن أن هذا المنطق ليس محبداً سواء في التوجهات التربوية الراهنة أو في كثير من النصوص الرسمية. والواقع أن منطق "النجاح" هو المرغوب فيه بوضوح في إطار بيداغوجية التحكم التي تسعى إلى أن تتحكم نسبة كبيرة من التلاميذ (80% إلى 90%) في هدف معين، قبل المرور إلى الهدف التعليمي الموالي، وهو الأمر ذاته في إطار المقاربة بالكفايات الأساس (دوكيتيل أ، ب 1993 وروجيرز 2000) التي تحدد عددا حصريا من الكفايات الأساسية المرتبطة بالحياة اليومية يتعين أن يتحكم فيها كل التلاميذ تحكما تاما.

تستند معظم التقنيات الكلاسيكية المعتمدة في تصديق اختبارات التقويم المدرسي على المقاربة القياسية النفسية أي على التوزيع الطبيعي في شكل منحنى كوص. وعلى سبيل المثال فالمؤشر الأكثر استعمالا هو مؤشر التمييز (فيندلي 1956) الذي يسعى إلى التحقق مما إذا كانت بنود الاختبار تستهدف التمييز بين التلاميذ، أي إبراز الفروق بين التلاميذ الضعاف والتلاميذ المتفوقين. ويميز هذا المؤشرين مؤشر صعوبة البنود بالنسبة للمجموعة "المتفوقة"، ومؤشر الصعوبة بالنسبة للمجموعة "الضعيفة": وكلما كان البون شاسعا كلما أسعف البند على التمييز بين التلاميذ. وعليه فالمجموعة المتفوقة هي المكونة من عناصر أحرزت على درجات إجمالية تضعها في خانة الـ 27% أعلى، فيما تندرج المجموعة الضعيفة في خانة الـ 27% أدنى (لافووكريكوار 2002). وهنا نرى بوضوح أن هذه الطريقة لا تتبنى وحدها توزيعا طبيعيا، بل إن الاختبار نفسه وضع فيها خصيصا ليفضي إلى التوزيع الطبيعي للنتائج أملا في تمييز بين التلاميذ يستوفي أعلى قدر من الدقة.

وفي نظر دوكيتيل وديفاي (2003)، فصحة اختبارات التقويم المبنية على المحتويات والأهداف تتسم إذن بـ:

- تمثيلية عينة الأسئلة للمحيط المرجعي؛

- وضعية المعدل (حيث يكون غير بعيد كثيرا عن مركز سلم التنقيط)؛

- تشتت النتائج (متغير أكبر من أجل التمييز بدقة متناهية بين التلاميذ)؛

- شكل التوزيع (منحنى كوص)؛

- البعد الأحادي للسمة المقيسة (كقياس استعداد مدرسي معين، على سبيل التمثيل).

حدود الاختبارات في منظور المقاربة بالكفايات

أولى هذه السمات تسمح لنا بالمضي قدما في إشكالية تصديق اختبارات التقويم المرتكزة على المقاربة بالكفايات. فعلى غرار ما يحدث حين نقوم بعملية استطلاع للرأي، حيث نتوخى الحصول على عينة ممثلة للمجموعة، عن طريق القيام بعملية سحب عشوائي، فكذلك تتحقق الخاصية التمثيلية لعينة بنود اختبار التقويم، على نحو أفضل عندما تتم بطريقة عشوائية. وينطبق ذلك إلى حد بعيد على مثال الامتحان النظري في رخصة السياقة باستعمال الحاسوب. فهذا الأخير يقترح بصفة عشوائية سلسلة من الأسئلة المنتقاة من محيط مجموع الأسئلة الممكنة المتعلقة بقانون السير (وتكون منصبة على "المحتويات" وحدها). فلنفترض أن العينة العشوائية منصّدة بحيث تشتمل مثلا على عدد من الأسئلة التي تتناول المخالفات الجسيمة. يتلقى جميع المرشحين استبيانات مختلفة مما يعني أن هؤلاء المترشحين لن يتوفروا على حظوظ متساوية في الامتحان. والحال أن هذا البعد العشوائي لبناء الاستبيان هو الذي يضمن صفاتها التمثيلية ويعزز صحتها على اعتبار أنه كلما كان عدد الأسئلة كافيا كلما أمكن تقليص نسبة الخطأ في القياس. وعندما يتم الالتزام بالضوابط المتعلقة بالصفة التمثيلية للعينة، فنادرا ما يكون هذا الشرط كافيا. الواقع أن خطأ القياس

$$\frac{1}{\sqrt{n}}$$

داخل العينة يمكن تقديره باعتماد صيغة $\frac{1}{\sqrt{n}}$ ، بحيث تمثل n عدد العناصر المتضمنة في العينة. وفي خطأ القياس المحدد في نسبة 5%، ينبغي بالضرورة أن يتضمن اختبار التقويم 400 بند، وهذا متعذر من الناحية العملية، في معظم الحالات. ومثال امتحان رخصة السياقة يجسد الوضعية D من خطاطتنا، حيث تتركز الجهود على النجاعة والمصدقية في تقويم الموارد، على حساب معيار الملاءمة. ولهذا السبب ينظم الامتحان التطبيقي في الوضعية. ونحن نعلم يقينا أن النجاح الباهر في الامتحان الكتابي قد يصاحبه غياب فادح للكفاية في الامتحان التطبيقي.

إن الاختبارات المنجزة وفق المقاربة بالكفايات قد تطرح مع ذلك صعوبات على مستويي المصدقية والنجاعة، بالنظر إلى حجم العينة المنتقاة من المحيط المرجعي. فمن حيث الجوهر، ستركز هذه الاختبارات على تقديم وضعية أو حتى وضعيتين مركبتين للتلميذ تتطلبان منه إنتاجا مركبا أيضا، ويتطلب الحل حيزا زمنيا معيناً². وهكذا فالمهام المطلوبة من التلاميذ تختلف بحسب الوضعيات التالية:

- أحيانا قد يقتصر الإنتاج على مهام وحيدة مركبة مثل كتابة نص يستجيب لوضعية تواصلية محددة، أو اجتياز ممر طرقي معين يقتضي تعبئة مجموعة من الموارد الضرورية، في الاختبار التطبيقي لرخصة السياقة.
- وأحيانا أخرى، يمكن تجزئة الإنتاج إلى عدة مهام، تتضمن بدورها عدة مراحل دونما حصر للعدد بالضرورة، كما هو الشأن بالنسبة إلى حل الوضعيات - المشكلة المركبة في مادة الرياضيات حيث العديد من سبل الحل ممكنة.

غير أنه في جميع الحالات لا يمكن إنشاء العديد من "البنود"، لأنها تعني تفكيك المهام المركبة إلى وحدات فرعية، مما يتعذر معه تقويم الكفاية التي عليها أصلا مدار تدبير هذا التركيب. ففي المقاربة بالكفايات لا يتعلق الأمر في اختبارات التقويم، بالتحقق من مصداقية الأداة، كما نفعل عادة في الاختبارات

الكلاسيكية (المبنية على المحتويات والأهداف) حين تكون "البنود" المكونة للعيينة المنتقاة من المحيط المرجعي عمدة وحدة التحليل. وفي اختبار بالكفايات، لا تكون هذه "البنود" واردة أصلا، إذ ليس ثمة سوى إنتاج مركب لا يشكل قطعا عينة إحصائية تمثيلية لمجمل الإنتاجات المركبة الممكنة ذات الصلة بالكفاية. إن تقنيات تصديق اختبارات التقويم بحسب المقاربة بالكفايات تظل مطروحة للابتكار بواسطة مسلمات أخرى غير المسلمات المعتمدة في النظرية الكلاسيكية للدرجات.

تصديقات الاختبارات وفق المقاربة بالكفايات

هناك أنواع مختلفة من التصديقات يمكن النظر إليها على نحو تكاملي:

- تصديق قبلي بالعودة إلى الحكام؛

- تصديق تجريبي داخلي؛

- تصديق تجريبي خارجي.

التصديق القبلي بالعودة إلى الحكام

التصديق القبلي بالعودة إلى الحكام تقنية استعملت أيضا في الاختبارات الكلاسيكية. وهي هنا تركز على مراجعة الحكام قصد التحقق مما إذا كانت البنود تشكل فعلا جزءا من المحيط المرجعي. وهو الشأن نفسه بالنسبة للمقاربة بالكفايات، مع فارق يتمثل في أن البنود والمحيط المرجعي متباينان. فالبنود عبارة عن وضعيات - مشكلة واقعية تقترح على التلاميذ، أما المحيط المرجعي فليس إلا تلك العائلة من الوضعيات التي تلائم الكفاية. وعلى الحكام أن يتحققوا من أن تلك الوضعية أو الوضعيات تنتمي فعلا إلى عائلة الوضعيات. إنه سؤال تكافؤ الوضعيات البالغ الصعوبة بله المستعصي على الحل. والواقع أن هذه الوضعيات المركبة من حيث المبدأ، يصعب التأكيد على تطابقها وانتمائها بالتبع إلى عائلة بعينها من الوضعيات. والحالة الوحيدة الضرورية لإنجاز عمل من هذا القبيل هو التوفر على وصف دقيق بما يكفي لعائلة الوضعيات تلك، أي على تعريف عناصر الإعداد التي تخصصها.

فإعداد Paramétrage عائلة الوضعيات تحدد الخاصيات التي يجب أن تستجيب لها كل الوضعيات التي تنتمي إلى نفس العائلة. ويمكنها أن تتناول عددا من الجوانب:

- كالمجال المرجعي في علاقته بالموارد القابلة للتعبئة: مثلا، عندما تستلزم جميع الوضعيات حلا مكونا من عملية طرح تليها عملية جمع، وتنصب كلتاها على أعداد طبيعية محصورة بين 0 و 1000؛
- ونوع الوضعيات، مثل وضعيات التواصل الشفوي التي تستلزم حضور شخصين حضورا فيزيقيا (وهي وضعية تستبعد التواصل الهاتفي).
- ونوع وعدد الحوامل، مثل وجود خطاطة على الأقل و / أو نص مؤلف من مائة كلمة؛
- ونوع المهمة المنتظرة، كإنتاج نص استدلالي من عشر جمل على الأقل، يكون مشفوعا بخطاطة توضيحية؛
- وشروط الحل، لا سيما المساعدات الموضوعية رهن إشارة التلميذ (معجم أو استبيان ...).

- ونوع المؤشرات الموظفة في تقويم الإنتاج، فعلى سبيل المثال لحل مسائل رياضية: (1) ملائمة الموارد المعبأة لمواجهة المشكلة المطروحة، (2) سلامة الموارد المعبأة (سواء أكانت ملائمة أم غير ملائمة)، (3) فضلا عن المعنى المقدم للجواب النهائي في إطار المسألة المطروحة.
- الخ..

إن إعداد عائلة الوضعيات لا يجب أن يلتبس بإكساء *habillage* الوضعيات التي تتعلق بكل وضعية على حدة، من حيث كونها تكشف كل مرة عن مظاهر خاصة تجعل منها وضعية قائمة الذات بين وضعيات أخرى. إن الإكساء يرتبط بالسياق الخاص للوضعية، وبما تتضمنه من معطيات ملموسة، وبطبيعة أسناد معينة أيضا. فالإعداد إذن يختلف من وضعية إلى أخرى، على النقيض من الإكساء الذي يظل مشتركا بين كل الوضعيات المشتقة من نفس العائلة التي في غيابها ينعدم التكافؤ بين تلك الوضعيات. إذن، يركز التصديق بواسطة الحكام على التحقق من مدى حضور عناصر الإعداد - والعناصر وحدها - الخاصة بعائلة الوضعيات في الوضعية الملموسة. فتكافؤ الوضعيات ينصب على الإعداد لا على الإكساء. ولعل الصعوبة القصوى إنما تكمن في إيجاد تعريف دقيق وجامع حول عائلة الوضعيات. وفي غيابها يتعذر علينا ضمان التكافؤ المنشود. ومن ثم، يمكننا القول بأنه قبل مطابقة الوضعيات مع رأي الحكام، ينبغي أن تركز عملية تصديق الاختبارات الموضوعة وفقا للمقاربة بالكفايات، على التحقق من توفرنا على عناصر إعداد دقيقة وشاملة بما فيه الكفاية، وضرورة صياغتها إذا لم تكن متوافرة.

التصديقات التجريبية البعدية

يمكننا القيام بتصديق آخر بعدي، على أساس النتائج المحصل عليها. فقديما كنا نميز بين:

- تصديق تجريبي داخلي، إذا كان يهدف إلى التحقق من الوحدة المفهومية للاختبار في حد ذاته.
- وتصديق تجريبي خارجي، إذا كان يروم إبراز التقارب القائم بين الاختبار ونظائره من الاختبارات المصدقة، أو المطروحة نتائجها للمقارنة مع اختبارات أخرى ذات صلة عكسية - على حد علمنا - أو غير عكسية بالبعد المقوم.

التصديق التجريبي الداخلي

يتم التصديق التجريبي الداخلي للرائز الكلاسيكي بالجوء أساسا إلى تقنيتين: التحليل العاملي وحساب معاملات التجانس (في الأغلب الأعم بتوظيف معامل "ألفا" لكرونباخ). نسعى من خلال التحليل العاملي إلى التحقق مما إذا كانت جميع بنود الرائز - في الحالة التي يفترض فيها من الاختبار أن يقيس استعدادا مدرسيا معينة - تمنح إشباعا قوية لأول مكون رئيس أو عامل مشتق³، فتساهم بالتالي في الحصول على متغير يكون على درجة عليا من الأهمية. على حين، إذا كان الاختبار مزيجا من العناصر (وعلى سبيل المثال فمسألة التحكم في حقل الرياضيات يفترض فيها أن تتضمن بعدين أحدهما حسابي والآخر هندسي) فعلى التحليل العاملي أن نؤكد أن يعمل جاهدا على التحقق مما إذا كان بإمكاننا، وراء أول مكون رئيس قوي يمثل استعدادا رياضيا عاما، أن نجد بالتعاقب عاملين، أحدهما يمثل البعد الحسابي والآخر البعد الهندسي. والحق أن الاستناد على معامل "ألفا" لكرونباخ يستجيب للانشغالات ذاتها؛ إذ في الحالة الأولى من المثال (عندما تكون جميع البنود حول نفس البعد) نسعى إلى التحقق من أن المعامل المحصل عليه مرتفع بما فيه الكفاية. وإذا

لم نفلح في ذلك عمدنا إلى تعيين البند أو البنود التي لا تقيس الاستعداد المطروح للتقويم. وفي الحالة الثانية من المثال (أي حين يهدف الاختبار إلى قياس عدة أبعاد) فإننا نحسب المعاملات على قدر الأبعاد (وبالتالي زمرة من البنود).

ولتوضيح هذا النوع من الاستدلال بطريقة ميسرة ولغايات بحثية صرف، نلوذ بحساب الترابطات، رغم أن هذه التقنية ليست مناسبة تماما لتحليل البنود. فلنفترض إذن أن الاختبار يقوم بعددين: أن يكون ترابط جميع الأسئلة المتعلقة بنفس البعد أعلى من الترابط القائم بين الأسئلة المنتمية إلى بعد واحد وتلك المتعلقة ببعد آخر.

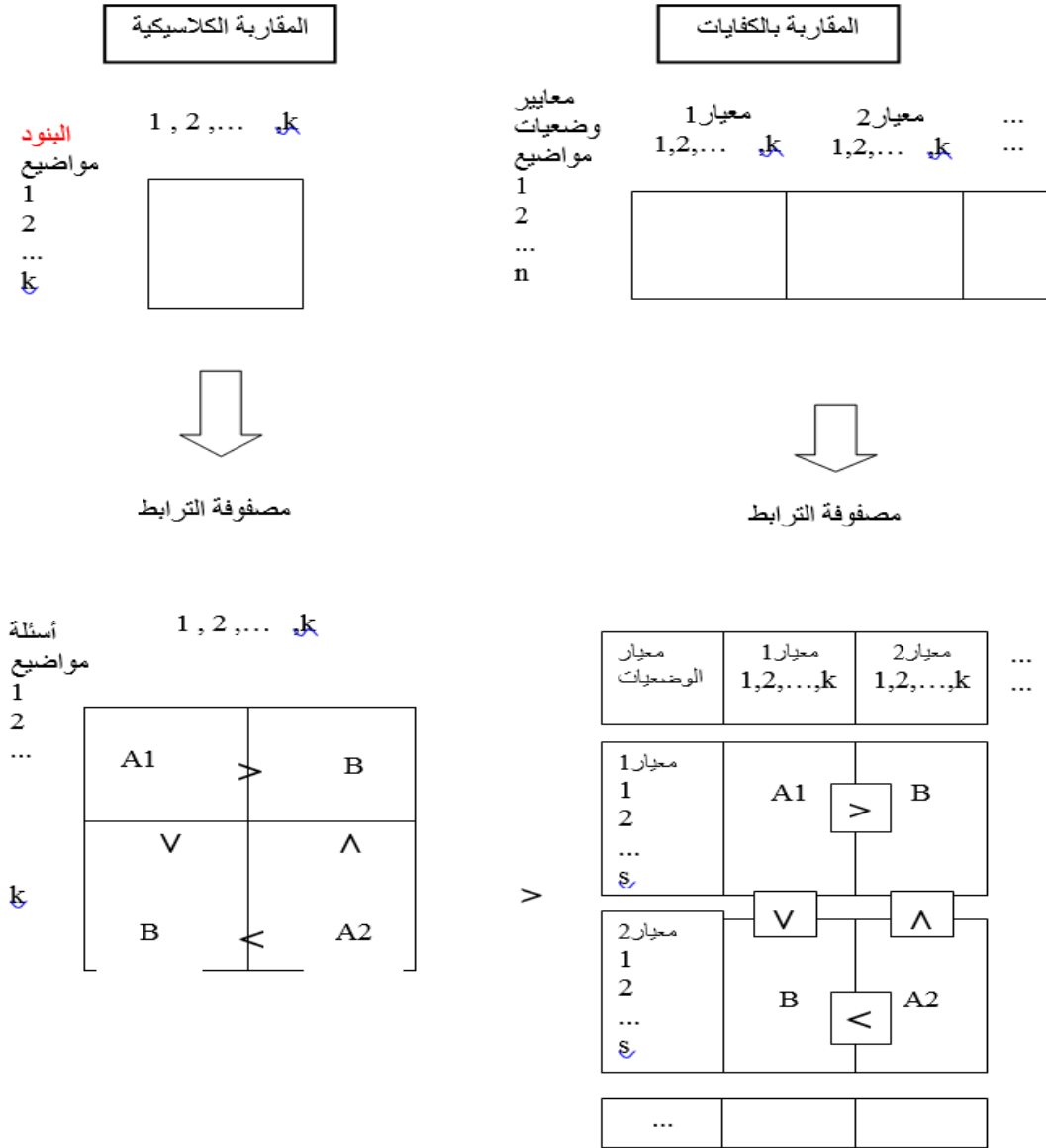
البعد 2	البعد 1	
$<$	r_{ij}	البعد 1
	V	
r_{ij}	R_{ij}	البعد 2

تكمن صعوبة التصديق التجريبي الداخلي للاختبار المهيأ وفق المقاربة بالكفايات والمشتمل على وضعية مركبة واحدة أو أكثر، في عدم وجود بعد وحيد، ولا مجموعة من البنود الكفيلة بقياسه، مادام الأمر متعلقا بإنتاج مركب؛ مما يعني أنه لا توجد استراتيجية تجيب على بند (مهمة وحيدة مركبة) وتسمح ببناء تحليلات للترابط.

فالتصديق إنما يجب أن يبنى على معايير التقويم التي تكون ثابتة في الوضعيات التي تنتمي إلى نفس العائلة، بخلاف بعض المؤشرات التي يتعلق معظمها بكل وضعية على حدة. فالمعايير تعكس زوايا النظر المختلفة التي يرى من خلالها المقوم منتوج التلميذ أو مختلف أنواع الجودة المنتظرة منه. ففي المقاربة الكلاسيكية للروائز، لا يوجد في معظم الأحيان سوى معيار واحد هو صحة الإجابة: وفي حالة الوضعية المركبة، يستعاض عن الإجابة الوحيدة المتعذرة بالإجابة "الجيدة". ولتقويم هذه الأخيرة لا مندوحة من اللجوء إلى معايير من قبيل: هل الإنتاج ملائم للوضعية؟ أهو تام؟ وهل الأدوات الموظفة (من خوارزميات، وصيغ وقواعد، إلخ) وافية بالغرض؟ وهل استعملت استعمالا صحيحا؟ وهل تنم بنية الإنتاج عن وجود انسجام؟ وهل الإنتاج قدم بصورة أنيقة؟ وهل جرا. ويمكن لكل معيار من هذه المعايير أن تتم أجرأته في وضعية ما من خلال بعض المؤشرات التي تكون عناصر قابلة للملاحظة المباشرة، وفي معظم الأوقات خاصة بتلك الوضعية. وعلى سبيل المثال يمكن التحقق من معيار "مناسبة الإنتاج للوضعية التواصلية" في الحالة التي يطلب فيها من التلميذ تقديم معلومات تتعلق بتحديد وجهة معينة، فأشار التلميذ فعلا إلى الطريق الذي يجب اتباعه سواء أكان صحيحا أم لا، معبرا عنه بأسلوب يحترم القواعد النحوية أو يلحن فيها⁴، إلخ.

ففي كل كفاية مقومة ينبغي تحديد المعايير المرشحة للاستعمال في تقويم إنتاجات التلاميذ (وهذا يندرج في نطاق عملية الإعداد). إن تصديق اختبارات مداره على مقارنة الوضعيات المختلفة المتعلقة بنفس الكفاية:

فالتلاميذ الذين حصلوا على درجات رفيعة في أحد المعايير المرتبطة بوضعية معينة، ملزمون بالحصول على درجات راقية، طبقاً لنفس المعيار، في وضعية أخرى تنتمي إلى نفس العائلة، والعكس صحيح. وإذن، فالنوع الأول من الاستراتيجيات موكول إليه التحقق من احترام هذا المبدأ في كل معيار من تلك المعايير. ويجب أن يخضع نفس التلاميذ، بطريقة مثالية، لتشكيلة من الوضعيات المتكافئة، أي تم إعدادها بالطريقة نفسها. لكننا نعلم أن الاختبارات المبنية على المقاربة بالكفايات تتطلب متسعاً من الوقت، ناهيك عن كوننا لا نستطيع، في كثير من الحالات، أن نفرض على نفس الأشخاص أكثر من اختبار واحد دون أن تتخلله إكراهات (كالإرهاق، وانعدام الحافزية، وتأثير التمرين، ولحظة التمرير، الخ) يصعب فك الارتباط بينها. على أن هناك حالات أخرى نستطيع فيها تقديم اختبارين (بل أكثر) لنفس التلاميذ. وهنا يمكننا الاستعانة بالنماذج الترابطية، علماً بأنه لا يحق لنا أن نعقد ترابطاً بين البنود مثنى مثنى، بل يكون الترابط بين الوضعيات مثنى مثنى، فيما يخص التحكم في المعيار المحدد. في الخطاطة أدناه نستطيع أن نعقد مقارنة بين عملية التصديق في المقاربة الكلاسيكية والمقاربة بالكفايات:



وكما نوهنا بذلك آنفا، فنحن نتوقع في المقاربة الكلاسيكية الحصول على ارتباطات داخل البنود أكثر قوة في المصفوفتين الفرعيتين A1 و A2 (بنود تنتمي إلى البعد نفسه) منه في المصفوفة B (بنود تنتمي إلى أبعاد مختلفة). أما في المقاربة بالكفايات فجميع الترابطات القائمة داخل الوضيعات بخصوص نفس المعيار (A1 et A2) يلزمها بالضرورة أن تكون قريبة من 1، وأكبر من تلك التي تنتمي لمعايير مختلفة B.

وبما أنه يستحيل إخضاع نفس الأشخاص لوضيعات متكافئة تماما، فإنه من الممكن اللجوء إلى استراتيجيات غير مباشرة انطلاقا من الوضيعات التي تتم مواجهتها عند الممارسة. وفي الواقع، فالتعلم

بالمقاربة بالكفايات يستلزم تقديم وضعيات متدرجة في التركيب للتلاميذ؛ حيث بوسع التلميذ أن يوظف في الوضعية $x+1$ مصادر سبق له توظيفها في الوضعية x . وبالنظر إلى معايير التقويم، تُطرح حالتان بحسب انتماء أو عدم انتماء المعايير المشتركة بين الوضعيتين x و $x+1$ ، إلى نفس مستوى المتطلبات (والمثال الذي يمكن تقديمه بخصوص هذه الحالة الأخيرة، يجسده معيار "التصحيح الإملائي"، حيث يمكننا أن نأخذ بعين الاعتبار قواعد أكثر تعقيدا).

فلنتصور وضعيتين تقويميتين Sx و $Sx+1$ ، حيث $(C1, C2) \subseteq Sx$ عبارة عن معيار من نفس مستوى المتطلبات التي هي $(C'1, C'2) \subseteq Sx+1$ الخاصة ب $Sx+1$ ، وحيث $C3$ و $C4$ معياران خاصان بالوضعية $Sx+1$ ، أو يمثلان مستوى المتطلبات أعلى من $C1$ و $C2$. فالخطاطة أدناه تمثل على التوالي، المعطيات المطلوب جمعها، والمؤشرات المطلوب حسابها، والنتائج المنتظرة:

النتائج المنتظرة:
معدل $C1 \leq C'1$ معدل
$SDC'1 \leq SDC1$
معدل $C2 \leq C'2$ معدل
$SDC'2 \leq SDC2$
معدل $C3$ أو $C4 \geq$ معدل $C'1$ أو $C'2$
أو $SDC'1 \leq C4$ أو $C'2$ معدل $SDC3$
$r(C2, C'2)$ و $r(C1, C'1)$
جد مرتفعة و أكبر من مصفوفات التلازم الأخرى.

الوضعيات	1+Sx	1S
المعايير	1'C 2'C 3C 4C	1C 2 C.
المواضيع		
1		
2		
.....		
n		
معدل		
SD	مصفوفة التلازم	

+

المعايير	C1	C2	C'1	C'2	C3	C4
C1						
C2						
C'1						
C'2						
C3						
C4						

النتائج المنتظرة:

معدل $C'1 \leq$ معدل $C1$
 $SDC'1 \leq SDC1$

معدل $C'2 \leq$ معدل $C2$
 $SDC'2 \leq SDC2$

معدل $C'3$ أو $C4 \geq$ معدل $C'1$ أو $C'2$
 $SDC'1 \leq C4$ أو $SDC'2 \leq C3$

$r(C2, C'2)$ و $r(C1, C'1)$
 جد مرتفعة و أكبر من مصفوفات التلازم الأخرى.

كما هو مبين في الخطاطة، فإننا نتوقع التمكن من المعايير المشتركة في $Sx+1$ مساو على الأقل لـ Sx نظرا لتدعيمهما عبر الزمن. وبالمقابل، لا ينبغي أن تكون الانحرافات المعيارية عالية في $Sx+1$ ، لأننا نفترض أن الامتلاك appropriation الجيد في المعايير عبر الزمن يؤدي إلى تجانس النتائج. والمثال التالي يشرح هذه الظاهرة: فإذا قمنا، خلال عام من الفرق، بتقديم مشكلات تتطلب اللجوء إلى طريقة الجمع، فسوف نحصل على أفضل إنجاز في $Sx+1$ (حيث يكون معدل التلاميذ جد مرتفع) وحتى في الجميع (بمعنى أن معدل SD سيكون ضعيفا جدا). مما يفسر أن الترابطات الموحدة المعيار، بين الزمنين، تكون بالضرورة أقوى مما هي عليه في حالة اختلاف المعايير.

التصديق التجريبي الخارجي

إن الاستراتيجيات التي أتينا على بسطها هي محاولات للتصديق التجريبي الداخلي. بيد أنه قد يكون من الأفيد اللجوء إلى استراتيجيات مكلمة للتصديق التجريبي الخارجي، أي الاستعانة بمعايير خارجية. وثمة العديد من الإمكانيات المتاحة في هذا الصدد يمكن تصنيفها بحسب ما تقتضيه العدة من دراسات مقارنة أو ترابطية. فالدراسات المقارنة تركز على مقارنة مجموعة تجريبية (المقاربة بالكفايات) بمجموعة شاهدة مكافئة. هنا تلوح حالتان تخصان تشكيل المجموعة الشاهدة بحسب معيار تعميم أو عدم تعميم المقاربة بالكفايات. ففي حالة عدم تعميم المقاربة بالكفايات، يتعين إنشاء عينة قياسا إلى العينة التجريبية، وذلك بالعودة إلى المدارس التي لا تطبق المقاربة بالكفايات، ولم يستفد مدرسوها من تكوين فيها (ولنحاذر هنا من ظواهر العدوى). أما في حالة تعميمها فينبغي أن تتكون المجموعة الشاهدة من تلاميذ اكتسبوا الموارد الضرورية لحل الوضعية؛ ولم يتعلموا بعد تعبئتها في هذا النمط من الوضعيات⁽⁵⁾. فحتى لو تكافأت المجموعتان التجريبية والشاهدة عند تقديم نفس الوضعية لهما، فإن إتقان أحد المعايير المقومة للكفاية المعنية سيكون لا محالة أعلى في المجموعة التجريبية. وسيكون مفيدا بناء اختبار من جزئين: أحدهما يختص بتقويم الموارد التي لا يستغني عنها الجزء الثاني الذي عليه أن يقوم الكفاية. وعندما يعجز الجزء الثاني عن التحكم الجيد في كل معيار، فإن المجموعة التجريبية لا يجب أن تكون، من حيث المبدأ، أقل إتقانا للجزء الأول من الاختبار (آدن وروجرز، 2003؛ أو/ديدي، والحاج عمار، وجيرار وروجرز 2005؛ راجونسون، وراميلجأونا، وراندريانيرينا، ورازافيندرالانبو، ورازافيندرانوفونا، وراوروفولولونا وجيرار 2005). وفيما تشغل الدراسات المقارنة على زممر مختلفة من التلاميذ، تعتمد الدراسات الترابطية على قياسات مكررة، أي أن نفس الزمرة من التلاميذ يتم تقويمها في اختبار الكفاية على أساس متغير معياري واحد أو أكثر. ويفترض في هذا الأخير أن تكون صلته بالكفاية المطروحة للتصديق مبرهنا عليها بصورة كافية ومعلومة. وهكذا ومادام إحراز صفة الكفاية يستلزم القدرة على حل المشكلات، أو إنجاز مهام مركبة، فإن التلميذ الأكثر (أو الأقل) إجادة للكفاية X، تكون حظوظه أوفر لتحسن إجادته للكفاية أو المتغير الملائم Y. ويمكن هنا إعمال العديد من المتغيرات المعيارية التي يكون بعضها أقرب، والبعض الآخر أبعد من الكفاية التي يتعين التصديق عليها، تعقبها عملية التحقق من مدى مطابقة النتائج الملحوظة (إتقان كل معيار من معايير تقويم الكفاية) للفرضيات الموضوعة.

الإشهاد بدليل النقطة المرقمة؟

كل الاستراتيجيات المقترحة لحد الآن هامة بالنظر إلى كونها تشكل حزمة من مؤشرات التصديق (وفي قاموس التحقيق البوليسي، فهي عبارة عن تخمينات)، بيد أنها لا تشكل أدلة (أي أنها لا تنتصب دليلا قاطعا على الجرم أو البراءة). وفي الواقع وكما أثبتنا أن التحكم في الموارد (في التصديق بناء على مقارنة كلاسيكية) ليس كافيا، فكذاك يعتبر إتقان كل معيار من معايير التقويم في اختبارينبي على المقاربة بالكفايات (تصديق بالاعتماد على الاستراتيجيات الواردة أعلاه) غير كاف هو الآخر. وطبعا ليس مؤكدا أن إتقان كل معيار من المعايير ينهض دليلا على إتقان الكفاية المستهدفة، ليس فقط لأنها تنتمي إلى نظام مركب وتستدعي من ثم أبعادا لا نعرف جيدا درجة التعالقات التي تنشأ بينها، ولكن أيضا لأن تقويمها يظل رهين المعايير المستعملة فيها⁽⁶⁾ (عددها، وأجراتها، ونظامها).

تبرز أهمية هذه المسألة عندما نجتهد في محاولة ترجمة نسبة إتقان الكفاية التي يجري تقويمها على أساس وضعية محددة بدقة، عبر إسناد نقطة وحيدة لهذه النسبة، كما تنص على ذلك معظم النصوص الوزارية أو يتوخاه العديد من الخبراء، لغايات المقارنة. حين يتعلق الأمر بترجمة نسبة التمكن إلى مجموعة من البنود البسيطة، عبر سلم رياضي أحادي الأبعاد (على نحو تقريبي، يقارن درجة التمكن في بند ما، بدرجة التحكم في بند آخر؛ كما جرت العادة في الاختيار من متعدد QCM، المبنية جيدا)، فإننا نتوفر هنا على نماذج رفيعة في القياس النفسي (انظر نظرية الإجابة عن الأسئلة⁽⁷⁾). هنا تكون النتيجة المرقمة والشيء المقوم متماثلين. على أن قاعدة التماثل هذه لا تغدو سائغة عندما نقوم واقعا متعدد الأبعاد، كما هو الشأن في المقاربة بالكفايات.

هل يترتب على ذلك وفي جميع الحالات، ضرورة العزوف عن ترجمة تقويم الكفاية، أو بتعبير أدق، تقويم إنجاز مركب يؤشر على الكفاية، إلى نقطة وحيدة؟ ألا يمكننا أن نجد في التقويم الإشهادي استراتيجية إجمالية⁽⁸⁾ تسمح بالتأسيس لقرار النجاح أو الفشل؟

إن الرهان هنا كبير ما دام الأمر يتعلق بالتقويم الإشهادي وأن هذا النوع من التقويم يفضي إلى اتخاذ قرارين: أولهما، في غاية الأهمية؛ ويكون ثنائيا حيث يركز على الإشهاد - اجتماعيا - على النجاح أو عدمه؛ أما الثاني فهو ثانوي⁽⁹⁾ في تصورنا، ويرتكز على تصنيف التلاميذ بحسب درجة الإنجاز. فالقرار الثاني لا يهمننا كثيرا لعدة أسباب منها: استحالة التحقق من مدى صحة التصنيف المبني على سلم رياضي أحادي الأبعاد، لإنجاز متعدد الأبعاد. فعملية من هذا القبيل تصلح لنموذج انتقائي لا يلائم البتة بيداغوجيا النجاح، الخ. على حين ينبغي أن يستأثر القرار الأول بأقصى درجات اهتمامنا، ومن ثم ضرورة التساؤل حول دلالة قرار النجاح مقابل الفشل: إنه سؤال الملاءمة.

التقويم الإشهادي لقرار النجاح أو الرسوب

إن قرار إنجاز التلميذ سيكون ملائما إذا شهدنا أن التلميذ يتحكم، في عتبة كافية، في الكفايات الدنيا اللازمة، حتى يتأتى له أن يستأنف بحصيلة مثمرة، تعلّمت الموسم الموالي، أو إذا تعلق الأمر بالتأهيل، أن ينجز المهام المهنية التي تم تكوينه من أجلها (مع احتمال إضافة شق ميداني إلى التكوين). وسيكون قرار النجاح صحيحا إذا برهنا على أن التلميذ يتحكم فعلا فيما يفترض فيه أن يتحكم فيه. ثم يكون القرار ناجعا عندما لا يكون مرهونا بالمقوم أو بظروف التقويم.

وإذن، فتصديق التقويم الإشهادي يمر أولا بتقويم قبلي يركز - مسنودا بخبراء - على التحقق مما: (1) - إذا كان التقويم الإشهادي ينصب فعلا على الكفاية أو الكفايات الدنيا⁽¹⁰⁾؛ (2) - إذا كنا قد ميزنا أصلا بين معايير التقويم الحد الأدنى، ومعيّار أو معايير الإتقان⁽¹¹⁾؛ (3) - إذا كنا قد حددنا، على النحو الوافي بالغرض، عتبات التحكم الدنيا. ويتعين، في الإطار المدرسي، إضافة شرط رابع: (4) - لا يمكن اعتماد التقويم الإشهادي لكفاية معينة إلا في نهاية مرحلة التعلم. والجدير بالذكر أنه عادة ما يقع خرق هذا الشرط الأخير، إذ يعتمد المدرسون إلى استعمال تقويمات يفترض فيها أن تكون أصلا تكوينية لاتخاذ قرارات إشهادية. والتباس الوظائف هذا، يفضي إلى تفضيل تلاميذ هم في غنى عن تدخل المدرس، على حساب الآخرين.

فهذا التصديق/القبلي للملاءمة يمكن أن يكون موضوع تصديق تجريبي خارجي. وفي هذه الحالة، فالمتغير المعياري يدل على أن التلميذ المعني بالأمر قد تعترضه مشاكل أو قد يكون في حل منها، عندما يباشر التعلم الجديدة أو الحياة المهنية. ويمكن صياغة معطيات هذا التصديق في الجدول التالي:

المتغير المعياري			
يسر	بصعوبة		
B	A	النجاح	الشهادة
D	C	الرسوب	

النتائج المنتظرة:

a: يرتقب أن يكون صفريا أو قريبا من الصفر حسب العدد الإجمالي

b: قريب من $a+b$

c: قريب من c و d

d: قريب من الصفر

إن وضع عدة تصديق تجريبي خارجي من هذا النوع، تعتوره صعوبة كبرى، ما دام التلميذ الراسب لن يجتاز، من حيث المبدأ، إلى القسم الموالي في السنة اللاحقة. غير أنه يمكن تفادي هذه الصعوبة إذا اختير التلاميذ الذين استفادوا من قرار اعتباطي في النجاح (لأن التقويم الختامي لم يعتمد على عُدّة ملائمة¹²)، وهذا يفترض أن يتم تحديد هؤلاء، على أساس الاختبار المعياري الذي يراد تصديقه.

إن القرار المتخذ بخصوص الوضعية x، سيكون صحيحا إذا كان القرار المتخذ بصدد الوضعية y أو الوضعية z المفترض انتماؤهما إلى نفس عائلة الوضعيات، وتقويمهما في نفس ظروف الإجراء، قرارا مماثلا. ويوضح الجدول التالي ما ننتظره حين يتعلق الأمر بالصحة:

		الوضعية Y	
		رسوب	نجاح
الوضعية X	نجاح	0-	
	رسوب		0-

يكون القرار المتخذ بخصوص الوضعية x ناجعا، إذا اتخذ مقومون مختلفون نفس القرار. ويمكن تجميع المعلومات ومعالجتها كالآتي:

		المقوم ب	
		النجاح	الرسوب
المقوم أ	النجاح	A	B
	الرسوب	C	D

$$\text{PHI de Guilford} = \frac{(b + c) - (a - d)}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$

$$p_e = \frac{p_o - p_c}{1 - p_c}$$

$$p_e = \frac{b + c}{b + c + a + d}$$

$$p_e = \text{proportion d'accords attendus par le hasard}$$

$$= \frac{(a + b)(b + d)(a + c)(c + d)}{N^2}$$

معامل PHI (في) كيلفورد = نسبة الاتفاقات المنتظرة بمحض المصادفة p_e

ويمكن التحقق من النجاعة عن طريق معامل كيلفورد (في) PHI الذي يجب أن يؤول إلى 1 (الاتفاق الكامل بين مجموع الاختبارات المقومة على أساس نفس الوضعية x). كما يمكن التحقق منها أيضا عن طريق معامل الاتفاق عند كوهن Cohen، وهو المتراوح بين 1- (أي عدم الاتفاق التام؛ حيث يتخذ المقومون قرارات مختلفة تعميم سائر التلاميذ) و+1 (أي حالة الاتفاق التام؛ حيث القرارات موحدة بين كل التلاميذ) مرورا ب 0 (حيث التوزيع عشوائي للقرارات المتفقة وغير المتفقة). وعلى العموم، ينبغي هنا أن يؤول معامل كوهن K إلى القيمة +1.

التقويم الإشهادي لمنح النقطة

في الفقرات السابقة، قدمنا التقويم الإشهادي في صورة قرار ثنائي. ومع ذلك، هل في وسعنا أن نمنح نقطة تحترم، في الآن معا، هذا النوع من القرار، وتسمح بإحداث ترتيب في صفوف التلاميذ (علما بأن هذا الإجراء الأخير يطرح مجموعة من المعضلات)؟

لقد حاولنا أن نقوم بذلك في حرص على وضع قواعد تحترم صحة القرار (نجاح - رسوب) في المقام الأول من الاعتبار. وهذه القواعد التي عرضنا آنفا لأربع منها هي كالتالي:

- أن يتمحور التقويم الإشهادي حول الكفاية أو الكفايات الدنيا؛
- أن تكون معايير التقويم الدنيا محددة سلفا، ومميزة بوضوح عن معيار أو معايير الإتقان؛
- أن تحدد عتبات التحكم الدنيا على النحو الوافي بالغرض؛
- ألا يوضع اختبار الإشهادي لكفاية معينة إلا بعد إكمال التعلم؛

- أن تخصص نسبة 75% على الأقل من سلم التنقيط للمعايير الدنيا، و25% أو أكثر للمعايير الإتقان (قاعدة 3/4)؛
- أن تحدد عتبة التحكم الدنيا الخاصة بالمعيار الأدنى في 2/3، وهي القاعدة التي تذهب إلى أن ارتكاب خطأ واحد على ثلاثة (كارتكاب خطأ على ثلاثة في عملية حسابية) لا يدل بالضرورة، في غالب الأحيان، على عدم التحكم، بقدر ما يعزى إلى السهو أو عوامل أخرى. وكلما اعتمدنا قاعدة 2/3 أقلحنا في تقليص حالات الرسوب التعسفية¹³.
- أن تحدد عتبة النجاح في نسبة 50% من سلم التنقيط؛ ومعناه أن 2/3 من 75% تساوي 50% من سلم التنقيط، وهذه العملية سنقلص حالات الرسوب والنجاح الاعتبارية. والجدول أدناه يوضح هذه الشروط:

مستويات التحكم		المعايير الدنيا			معايير الإتقان	
		C1	C2	C3	C4	C5
غياب تام		0	0	0	0	0
تحكم جزئي		2	2	2	1	0
تحكم أدنى		4	4	4	2	1
تحكم أعلى		5	5	5	3	2
المجموع		15/20			5/20	

ملحوظة: التحكم الأدنى في المعايير الدنيا، المكتوب بحروف بارزة يؤدي إلى النجاح (هنا بنقطة 12/20).

يرتكز التصديق القبلي على التحقق من استيفاء تلك الشروط. أما التصديق التجريبي فينبني على بلورة الإجراءات النوعية التالية:

- التحقق من حصولنا فعلا على توزيع على شكل z ⁽¹⁴⁾ وليس على شكل توزيع كوص؛ وإلا فعلينا التشكيك في عدم احترام بعض الشروط (خاصة الشروط الأربعة الأولى) وهذا يندرج تحت الإخلال بالملاءمة؛
- مقارنة توزيعات نتائج نفس التلاميذ (أو عينتين من التلاميذ، يتم إعدادهما عند الاقتضاء) تكون قد قوّمت في وضعية x ووضعية y ، تنتمي إلى نفس عائلة الوضعيات (انظر النماذج الإحصائية المسماة بنماذج الضبط)؛ مع حساب الترابط بين التوزيعين (حيث يجب أن تؤول قيمته إلى $1+$).
- دراسة الترابط القائم بين توزيع النتائج في وضعية x من جهة، والنتائج المحصل عليها باستعمال متغير معياري واحد أو أكثر، من جهة أخرى (التصديق التجريبي الخارجي)
- حساب معاملات النجاعة (الترابطات) بين النقط التي منحها مقومون مثنى مثنى، أو معامل التوافق الخاص بـ "Kendall".

الوضعية Y				الوضعية X	
(d)	(c)	(b)	(a)		
					تلاميذ لا يعانون من أي مشكل
					تلاميذ ناجحون لكنهم محتاجون إلى دعم
					تلاميذ يعانون من صعوبات
				تلاميذ يعانون من صعوبات أكبر	

- 56 -

خلاصات

يعتبر تصديق الاختبارات المبنية في انسجام مع المقاربة بالكفايات مشكلا عصبيا على الحل بكيفية مرضية، كما أن الاقتراحات المعروضة في هذا المقال لا تعدو أن تكون مجرد مقاربات للمشكلة المطروحة. وإذا شئنا استرجاع مثال التحقيق البوليسي، فنحن لا نتوفر على أي دليل مادي مادام الظنن لم يقبض عليه في حالة تلبس ولا اعترف بالجرم المنسوب إليه. فليس تحت أيدينا إلا حزمة من التخمينات غير الموثوق بها. وعلى هذا الأساس نرجو أن تفهم اقتراحاتنا هنا.

هل لنا أن نتوق يوما إلى الحصول على طريقة قياس تربوية مرضية لتقويم الكفايات المركبة؟ لا شيء يبعث على الشك إطلاقا في ذلك، حسب معارفنا الراهنة. لكن، هل ينبغي لذلك الاقتصار فقط على تقويم الموارد المكتسبة (معرفة إعادة الإنتاج والمهارات الأوليتين) ما دامت هي الوحيدة القابلة للتقويم بطريقة صحيحة وناجعة؟ إن الأمر ليدعو حقا إلى ركوب المحاولة، ما دامت الترسنة التقنية جاهزة بالشكل المطلوب. ونحن نعلم هالة السحر، والشعور بالصرامة والثقة، والانطباع العلمي الذي يمنحنا إياه توظيف التقنيات المتطورة.

لقد دافعنا في هذا المقال بقوة على مبدأ أولوية الملاءمة على الصحة والنجاعة. فهاتان الأخيرتان لئن كانتا في نظرنا - قطعاً - هامتين، فنحن لا نستطيع الاكتفاء بهما بصورة حصرية. فتكوين صياد كفاء، يقتضي بالضرورة جعله قادرا على تسديد رمية مركزة في مركز الهدف المتحرك، وليس فقط صوب مركز الهدف الثابت.

وفي انتظار إيجاد طريقة قياس تربوية أكثر مواءمة لتقويم حل المشكلات المركبة، أو إنجاز المهام المركبة (وهذا أسُّ مفهوم الكفاية)، فالأجدي - على غرار صورة المحقق الذي تنقصه الأدلة الراجعة - تبني سلسلة من الاستراتيجيات التي تشكل حزمة من القرائن الكافية حول قيمة اختباراتنا التقويمية (الملائمة، والصحيحة، والناجعة في ذات الآن).

المراجع

- Aden, H.M., & Roegiers, X. (2003). A quels élèves profite l'approche par les compétences de base ? Étude de cas à Djibouti. Document inédit, accessible à l'adresse suivante : [http://www.bief.be].
- Bloom, B.S., Hastings, J.Th.H., & Madaus, C.F. (éds) (1971). Handbook on formative and summative evaluation of student learning. New York.
- Cardinet, J. (1973). L'adaptation des tests aux finalités de l'évaluation. Les sciences de l'éducation – Pour l'ère nouvelle, 2-3(1973), 148-182.
- Cardinet, J., & Tourneur, Y. (1985). Assurer la mesure. Berne – Francfort-s.Main – New York : Peter Lang.
- Carver, R.P. (1974). Two dimensions of tests: psychometric and edumetric. American Psychologist, 29, 512-518.
- Crahay, M. (1996). Peut-on lutter contre l'échec scolaire ? Bruxelles : De Boeck Université.
- De Ketele, J.-M. (1993a). L'évaluation conjugée en paradigmes. Revue française de pédagogie, 103, 59-80.
- De Ketele, J.-M. (1993b). Objectifs terminaux d'intégration et transfert des connaissances. In R. Hivon (ed.), L'évaluation des apprentissages. Réflexions, nouvelles tendances et formation, (chap. 1, pp. 15-26). Sherbrook (Canada): CRP.
- De Ketele, J.-M. (2000). En guise de synthèse: Convergences autour des compétences. In C. Bosman, F.-M. Gerard & X. Roegiers (éds), Quel avenir pour les compétences ? (pp. 187-191). Bruxelles : De Boeck Université.
- De Ketele, J.-M. (2001a). Place de la notion de compétence dans l'évaluation des apprentissages. In G. Figari & M. Achouche (éds), L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels (pp. 39-43). Bruxelles: De Boeck Université
- De Ketele, J.-M. (2001b). Enseigner des compétences : repères. In J.-L. Jadoulle & M.Bouhon (éds), Développer des compétences en histoire (pp. 13-22). Louvain-la-Neuve-Bruelles: Université catholique de Louvain et Ministère de l'éducation, de la recherche et de la formation.
- De Ketele, J.-M., Chastrette, M., Cros, D., Mettelin, P., & Thomas, J. (1989). Guide du formateur (2e ed.) Bruxelles : De Boeck Université.

- De Ketele, J.-M., & Dufays, J.-L. (2003). Vers de nouveaux modes d'évaluation des compétences. In L. Colles, J.-L. Dufays & C. Maeder, Enseigner le français, l'espagnol et l'italien. Les langues romanes à l'heure des compétences (pp. 171-182). Bruxelles-Paris : Editions De Boeck Duclos.
- De Ketele, J.-M., & Roegiers, X. (1991, 2è éd. 1993). Méthodologie du recueil d'informations. Bruxelles : De Boeck Université.
- De Landsheere, G. (1980). Examens et évaluation continue. Précis de docimologie. Bruxelles-Paris : Labor-Nathan.
- Dolz, J., & Ollagnier, E. (éds) (2002). L'énigme de la compétence en éducation. Bruxelles : De Boeck Université.
- Findley, W. G. (1956). A rational for evaluation of item discrimination statistics. Educational and Psychological Measurement, 16, 175-180.
- Fourez, G. (1999). Compétences, contenus, capacités et autres casse-têtes. Forum des pédagogies, mai, 26-31.
- Gerard, F.-M. (2000). Savoir, oui... mais encore ! Forum - pédagogies, mai, 29-35.
- Gerard, F.-M. (2001). L'évaluation de la qualité des systèmes de formation. Mesure et évaluation en éducation, 24(2-3), 53-77.
- Gerard, F.-M. (2002). L'indispensable subjectivité de l'évaluation, Antipodes, 156, avril, 26-34.
- Gerard, F.-M. (2005), Evaluer des compétences, ou ne pas se tromper de cible. Liaisons, 40, février, Beyrouth, Liban, 7-9.
- Gerard, F.-M., & Braibant, J.-M. (2003). Activités de structuration et activités fonctionnelles, même combat ? Le cas de l'apprentissage de la compétence en lecture à l'école primaire, Français 2000, 190-191, avril 2004, 24-38.
- Jonnaert, Ph. (2002). Compétences et socioconstructivisme - Un cadre théorique. Bruxelles: De Boeck.
- Lasnier, F. (2000). Réussir la formation par compétences. Montréal : Guérin.
- Laveault, D., & Gregoire, J. (1997, 2e éd. 2002). Introduction aux théories des tests en psychologie et en sciences de l'éducation. Bruxelles : De Boeck Université.
- Le Boterf, G. (1994). De la compétence. Essai sur un attracteur étrange. Paris: Editions de l'organisation.

- Legendre, M.-F. (2001). Sens et portée de la notion de compétence dans le nouveau programme de formation. *Revue de l'AQEFLS*, 23(1), 12-30.
- O/ Didiye, D., El Hadj Amar, B., Gerard, F.-M., & Roegiers, X. (2005). étude relative à l'impact de l'introduction de l'APC sur les résultats des élèves mauritaniens. Document inédit, accessible à l'adresse suivante: [<http://www.bief.be>].
- Perrenoud, Ph. (1997). Construire des compétences dès l'école. Paris: ESF.
- Poirier Proulx, L. (1999). La résolution de problèmes en enseignement. Cadre référentiel et outils de formation. Bruxelles : De Boeck Université.
- Rajonhson, L., Ramilijaona, F., Randrianirina, P., Razafindralambo, M.H., Razafindranovona, O., Ranorovololona, E., & Gerard, F.-M. (2005). Premiers résultats de l'APC : invitation à continuer... Document inédit, accessible à l'adresse suivante : [<http://www.bief.be>].
- Rand McNally. Tardif, J., et al. (1995). Le développement des compétences, cadre conceptuel pour l'enseignement. In J.-P. Goulet, Enseigner au collégial (pp. 157-168).
- Raynal, F., & Rieunier, A. (1997, 4^èéd. 2003). *Pédagogie : dictionnaire des concepts clés*. Paris : ESF éditeur.
- Rey, B. (1996). Les compétences transversales en question. Paris : ESF.
- Rey, B., Carette, V., Defrance, A., & Kahn, S. (2003). Les compétences à l'école – Apprentissage et évaluation. Bruxelles : De Boeck.
- Roegiers, X. (2000). Une pédagogie de l'intégration. Bruxelles : De Boeck.
- Roegiers, X. (2003). Des situations pour intégrer les acquis. Bruxelles : De Boeck.
- Scallon, G. (2004). L'évaluation des apprentissages dans une approche par les compétences. Bruxelles : De Boeck Université.
- Scriven, M. (1967). The methodology of evaluation, AERA Monograph Series on Curriculum Evaluation, 1. Chicago- Montréal : Association québécoise de pédagogie collégiale.
- Tilman, F. (2000). Qu'est-ce qu'une compétence ? *Exposant neuf*, 2 (2000), 28-31.

معجم:

A posteriori	بعدي
a priori	قبلي
Abusivement	تعسفي
Adequat	وافية بالغرض
Adéquation	مواءمة
Analyse factorielle	تحليل عاملي
Appoprie	تناسب
Appropriation	تملك
Coefficients d'homogénéité	معامل التجانس
Cohérence	منسجم
Compétences minimales	الكفايات الدنيا
Convergence	تقارب
Corrélations	الترابطات
Correspond	تلائم
Courbe de Gauss	منحنى كوص
Courbe de référence	منحنى مرجعي
Courbe en J	منحنى في شكل J
Critères	معايير
Critères de perfectionnement	معايير الإتقان
Degré de compatibilité	درجة التوافق
Démotivation	انعدام الحافزية
Déviation standards	الانحرافات المعيارية
Dichotomique	ثنائي
Dispositif	العدة
Échantillon aléatoire	عينة عشوائية
Échantillon représentatif	عينة تمثيلية
Edumétrie	قياس تربوي
Effet d'entraînement	تأثير التمرين
effet systématique	عامل نسقي
Empirique	تجريبي

Équivalente	مكافئ وتكافؤ
Être moins performant	أقل إتقاناً
Études comparatives	الدراسات المقارنة
Études corrélationnelles	الدراسات الترابطية
Faisceau	حزمة
Famille de situations	عائلة الوضعيات – فئة الوضعيات
Fiabilité	النجاعة
Finalités	غايات
Groupe contrôle équivalent	مجموعة شاهدة مكافئة
Groupe expérimental	مجموعة تجريبية
Groupé	مركزة
Groupe	مجموعة زمرة -
Habillage	إكساء وتكسية وتهيئة
Inadéquat	غير ملائم
Incompétence	غير كفء
Indépendants	مستقلة
Indicateurs	مؤشرات
Indice	مؤشر
Infaisable	متعذر
Interrelations	تعالقات
Inter-situations	داخل الوضعيات
Isomorphes	تماثل
Items	بنود
Itinéraire	مسار – وجهة معينة
Loi de Posthumus	قانون بوستموس
Maitrise	تحكم – تمكن – إتقان
Matrice	مصفوفة
Médiane	الوسيط
Même niveau d'exigence	نفس مستوى المتطلبات
Mobiliser	تعبئة
Mode	منوال
Modelés corrélationnels	النماذج الترابطية

Modelés psychométriques	نماذج القياس النفسي
Modelés statistiques d'ajustement	النماذج الإحصائية للضبط
Moment de la passation	لحظة – فترة التمرير
Multidimensionnelle	متعدد الأبعاد
Non-maitrise	عدم التمكن – عدم التحكم
Opérationnalisé	أجراء
Outil fiable	أداة ناجعة
Paramétrage	إعداد
Paramètres	عناصر الإعداد
Pertinence	ملاءمة
Pratiques pédagogiques	ممارسات بيداغوجية
Psychométrie	قياسات نفسية
Recours	عودة
Réponse items théorie	نظرية الإجابة على الأسئلة
Ressources	الموارد
Savoir-faire	المهارات – معارف التصرف
Savoir-reproduire	معارف إعادة الإنتاج
Scores	الدرجات
Seuils de maitrise	عتبة التحكم
Situationnels	ظرفي
Situation-problème	وضعية مشكلة
Situations-problème équivalentes	وضيعات مشكلة متكافئة
Situations complexes	وضيعات مركبة
Stratifié	منضدة
Supports	أسناد
Teste	روائز
Unidimensionnalité	أحادي البعد
Univers de référence	الكون المرجعي
Validité	المصدقية
Variables critères	المتغيرات المعيارية
	Vraisemblable محتمل

¹ - تتفرع هذه الموارد في الآن نفسه إلى داخلية خاصة بالشخص، وخارجية (جيرار وبريبان، 2004): فالداخلية خاصة بالفرد، توجه عمله لحل المشكلة. يتعلق الأمر أساساً بمعرفة إعادة الإنتاج، ومعرفة الفعل، ومعرفة الكينونة التي توظف جميعاً في المجالات المعرفية، والنفسية الحسية - الحركية، أو السوسيو وجدانية (جيرار 2000، 2001). أما الموارد الخارجية فتتحلل على كل ما يكون قابلاً للتعنُّب خارج الشخص. وهي تتعلق من ثم بالموارد المادية (وسيلة، نص مثلاً)، والموارد الاجتماعية (اجتماع، شبكة علاقات مثلاً)، والموارد الإجرائية (خوارزم، قانون على سبيل المثال).

² - مما يفيد صعوبة اقتراح أكثر من وضعية واحدة أو وضعيتين.

³ - مصطلح التحليل العاملي مصطلح نوعي يكثر استعماله لكنه يحجب الاختلاف، الذي قليلاً ما يكون معروفاً، بين دلالة المكون الرئيس، ودلالة العامل. وفي حالتنا هاته، فالمكون الرئيس الأول هو الأكثر وفاء بقياس الاستعداد المدرسي الشامل الذي يقيسه الاختبار.

⁴ - يوضح ذلك قاعدة هامة تتمثل في استقلال معايير التقويم.

⁵ - يتمثل المبدأ المنهجي للمقاربة بالكفايات الأساسية في أن يتعلم التلميذ - بطريقة كلاسيكية نسبياً - مختلف الموارد المرتبطة بالكفاية، خلال خمسة أسابيع مثلاً. وبعدها يخصص أسبوع لإدماج الموارد حيث يتعلم التلاميذ، في وضعيات مركبة، تعيين الموارد الضرورية لحل الوضعية؛ وتعبئتها بطريقة مدمجة لحل الوضعية المطروحة (جيرار وبريبان 2004؛ راي وكاريت ودوفرانس وكان 2003؛ روكيرس 2000).

⁶ - يرتبط هذا بذاتية المقوم المحتومة والتي لا مفر منها (جيرار 2002).

⁷ - تشير نظرية الإجابة على الأسئلة *Response items theorie* التي يرمز إليها بـ (IRT)، كما تعرف باسم نظرية الإجابة الكامنة، إلى مجموعة من النماذج الرياضية تشرح العلاقة بين خاصية أو سمات كامنة غير قابلة للرصد (مثل الإجهاد والمعرفة والمواقف) ومظاهرها من خلال القياسات (أي النتائج المرصودة أو الاستجابات أو الأداء). (المرجمون)

⁸ - لأسباب تاريخية (وخاصة أعمال سكريفن وبلوم) فإن خلطاً ما فتي يسود حتى في صفوف خبراء معاصرين كثر، بين وظائف التقويم وإجراءات التقويم. فالتقويم قد يخدم وظائف التوجيه (قرارات بشأن عمل جديد يتعين إنجازه)، والمعالجة (قرارات تخص تطوير عمل جارٍ كما هو الشأن بالنسبة للتقويم المسمى تكوينياً)، والإشهاد (قرارات تتعلق بأهداف التعرف الاجتماعي والإداري). وواحدة من هذه الوظائف يمكن إنجازها بأحد الإجراءات التالية: إجراء توصيفي (فأنا أصف وقائع وإنجازات، ألخ)، أو إجمالي (أجمل بأسلوب رياضي وقائع وإنجازات، ألخ)، أو تأويلية (أعطي معنى بصورة حدسية لحزمة من العلامات). والخلط مرده إلى المباشرة التعسفية بين التقويم الإشهادي والتقويم الإجمالي؛ كما لو كان الأول متعذر الحصول إلا بواسطة التقويم الإجمالي.

⁹ - هي ثانوية في نظرنا، غير أنها ليست كذلك في نظر الكثير من الآباء.

¹⁰ - الرأي عندنا أن التقويم الإشهادي سيكون أفضل إذا أقرنا بوجود كفاية كبرى تلزم التلميذ بتعبئة مجموع الكفايات الدنيا المطلوبة: نسمي هذا الإجراء هدفاً نهائياً للإدماج (دوكيتيل 1993 أ-ب).

¹¹ - يكون المعيار أدنى إذا كان متحكماً فيه بالضرورة حتى يشهد بتحقيق الكفاية بنجاح؛ خلافاً لمعيار الإتقان الذي لا يفترض التحكم فيه بالضرورة. وإذا تحكّم فيه التلميذ فذلك أحسن لكنه ليس لازماً. وعلى سبيل المثال، ففي حالة تقويم كفاية حلّ المشكلات الرياضية، يمكن أن يكون معيار "الاستعمال السليم لوسائل المادة" معياراً أدنى، بينما تكون "جودة العرض" معياراً للإتقان. يتعلق الأمر هنا في الواقع، بقرار المقومين السابق على التقويم والذي يشكل قسماً من إعداد عائلة الوضعيات.

¹² - يكثر هذا الوضع في الدول التي يكون التباين فيها بين المدارس قوياً.

¹³ - نتحدث عن فشل تعسفي حين يرسب التلميذ ليس بسبب عدم تحكمه في الكفايات الدنيا المطلوبة، ولكن بسبب منظومة التقويم نفسها. إنه يرسب حيث لا يجب، لأنه كان بمقدوره مواصلة التعلم في السنة الموالية. أما عندما ينجح التلميذ حيث لا يجب، فتحدث حينئذ عن النجاح التعسفي لأن عدم تحكمه في الكفايات الأساسية المطلوبة لا يؤهله لمباشرة التعلم الجوهري في السنة الموالية.

¹⁴ - يستخدم تعبير منحني على شكل J في الميدان الاقتصادي لوصف ظاهرة التدهور الأولي يليها الانتعاش. (المرجمون)