دليل الطالب

حول نظم التقويم والامتحانات

عزيزى الطالب/ عزيزتى الطالبة

بين أيديكم دليل تقويم الطلاب ونظم الامتحانات، والذى يتناول الجوانب المرتبطة بالتقويم والنظم المختلفة للامتحانات، ويتكون هذا الدليل من عدة جوانب أهمها :

أ. تعريف التقويم وأنواعه. ب. الاختبارات التحصيلية.
ج. التقويم الحقيقى. د. بنوك الأسئلة.

ويشكل هذا الدليل جزءاً يتكامل مع المحاضرات التى ستقدم حول الجوانب المختلفة للتقويم، ولذا ننصح بقراءة هذه الموضوعات بتعمق وطرح تساؤلاتكم حول كافة الموضوعات المرتبطه بهذه الجوانب ونأمل أن تتحقق الاستفادة المرجوة من هذا العمل.

والله نسأل التوفيق

نظـم التقويم

تتحد نظم التقويم في الأنواع التالية:

التقويم المبدئى:

ويتم قبل تقديم البرنامج التعليمى لكى يتم تحديد ما يتوافر لدى المتعلم من خصائص ترتبط بموضوع التعلم، بهدف الكشف عن حاجة المتعلمين إلى تعلم بعض المهارات المدخلية أو المتطبات اللازمة للبدء فى دراسة موضوع معين.

التقويم البنائى (التكوينى):

ويعنى استخدام التقويم المنظم خلال مسار عملية التدريس، بهدف الكشف عن جوانب القوة والضعف فى البرنامج التعليمى أثناء إعداده وتنفيذه فى مرحلة التجريب بغرض مراجعة مكونات البرنامج وتعديله وتطويرها، وكذلك تشخيص مواطن القوة والضعف فى أداء المتعلمين وتحديد الصعوبات التى يواجهها كل منهم أثناء التعليم واتخاذ ما يلزم من أساليب العلاج، ومساعدة المتعلم فى التعرف على قدراته وإمكاناته وإقتراح سبل ووسائل تحسينها إلى أقصى حد ممكن.

التقويم التجميعى (النهائى):

ويعنى الحكم على مدى إحراز نواتج التعلم بهدف اتخاذ قرارات مثل نقل المتعلم إلى مستوى جديد أو تخرجه.

ولتحقيق الخطوة الأولى لكل من التعليم والتقويم والمتمثلة فى تحديد أهداف التعليم تحديدا واضحا، يجب تحديث الأهداف القومية للتعليم لتلافى أخطاء وعيوب واقعنا التعليمى عن طريق ترجمة هذه الأهداف إلى معايير تربوية.

وغالباً ما يستخدم في التقويم الاختبارات التحصيلية بأنواعها المختلفة ومنها:

أ . الاختبارات الشفهية:

تعتبر الاختبارات الشفهية هى الأداة الوحيدة التى تصلح لقياس مدى تحقق بعض الأهداف التربوية، فاذا أردنا أن نختبر مدى تقدم التلميذ فى نطق بعض اللغات الأجنبية أو إجادته للقراءة مثلا أو مدى قدرته على التعبير عن نفسه فإننا لابد أن نلجأ إلى الاختبارات الشفهية.

والاختبار الشفهى هو مزيج من اختبار المقال والاختبار العملى وله فائدته فى دراسة العمليات المعرفية التى يستخدمها التلميذ فى الإجابة على أسئلة معينة.

والاختبار الشفهى عبارة عن مجموعة من الأسئلة القصيرة إلى حد ما والتى يلقيها المعلم على التلميذ بصورة شفهية، وعلى التلميذ أن يجيب على هذه الأسئلة أيضا بصورة شفهية وفى الحال.

ب . الاختبارات التحريرية:

وهى الاختبارات التى يستخدم فيها التلاميذ الورقة والقلم للاجابة عليها، ومن هذه الاختبارات ما يلى:

1 . اختبارات المقال.

2 . الاختبارات الموضوعية.

وسوف نتناول فيما يلى كل منهما بشئ من التفصيل:

أولا: اختبارات المقال:

وهى عبارة عن مجموعة من الأسئلة قليلة العدد، يتطلب الإجابة عليها أن يكتب التلميذ مقالا حول موضوع السؤال أو الظاهرة، حيث يناقش ويعلل ويفسر الظاهرة فى صورة مقال.

ويستخدم اختبار المقال بطريقة شائعة فى حجرات الدراسة وفى الامتحانات العامة مثل امتحان الثانوية العامة، ويتكون المقال من عدد من الأسئلة التى تتطلب إجابة مكتوبة حرة خاصة بموقف يمثل مشكلة، ونذكر فيما يلى عددا من الأسئلة "كاختبار مقال" مأخوذة من امتحانات عامة:

- لا يحدث التعلم دون شروط معينة. ناقش؟
- ما الفرق بين نظرية ثورنديك ونظرية الجشطلت؟
- كيف يستفيد المدرس من دراسته للدافعية فى حجرة الدراسة؟
- ماذا تعنى بالتقويم ؟ وما الفرق بين التقويم والقياس؟

والمطلوب من التلميذ أن يكتب مقالا كإجابة عن كل سؤال من هذه الأسئلة، ثم بعد ذلك يصحح الممتحن الإجابات، وغالبا ما تعطى الدرجات بإحدى الطريقتين الآتيتين:

- أن يعد الممتحن نموذجا مفصلا للإجابة لإعطاء الدرجات طبقا لمزايا معينة للإجابة، وتسمى هذه الطريقة "التصحيح التحليلى".
- والطريقة الثانية تتلخص فى أن الممتحن يعطى ما يعتقد أنه الدرجة المناسبة بعد قراءة الإجابة بدون أى تحليل مفصل، وهذه الطريقة تسمى "التصحيح الانطباعى".

ثانيا: الاختبارات الموضوعية:

تتميز اختبارات التحصيل الموضوعية بأنها تشتمل على عدد كبير من الأسئلة التى تغطى معظم أجزاء المحتوى الدراسي على العكس من اختبارات المقال، هذا وتختلف مواصفات الاختبار الموضوعي باختلاف الهدف من الاختبار وكذلك العينة التى تطبق عليها.

ونظرا لأن الاستجابة فى اختبار المقال تترك لذاتية المتعلم ومدى فهمه للسؤال فإنها تتسم بالذاتية أو تنقصها الموضوعية، ولذلك فان الاختبارات الموضوعية تحتوى على المثير والاستجابة معا والمثير يتمثل فى مقدمة السؤال أما الاستجابة تكون موضوعية بطريقة تقتضى تميزها ما بين عدد من الاستجابات الخاطئة أو توضح بطريقة تقتضى الحكم على مدى صحتها أو خطئها أو توضح بطريقة تقتضى من المتعلم أن يضيف إليها من عنده إضافة تجعلها استجابة كاملة وصحيحة.

ومن هنا فإن الاختبارات الموضوعية تحتوى على أكثر من نوع، وأهمها:

أ . اختبار الاختيار من متعدد.

ب . اختبار الصواب والخطأ.

جـ . اختبار التكملة.

د . اختبار المزاوجة.

وسوف نشرح الآن طريقة بناء كل من هذه الأنواع:

أ . أسئلة الاختيار من متعدد:

يتكون السؤال فى هذا النوع من الاختبارات الموضوعية من جزئين رئيسيين:

1 . الإثارة وهى مقدمة أو مفتاح للسؤال.

2 . الاستجابات أو الاختيارات ويشمل السؤال عادة ما بين أربع أو خمس استجابات يختار التلميذ من بينها الاستجابة الصحيحة ويتم تدوينها بمجرد وضع علامة مناسبة فى المربع المخصص لذلك، وتكون الإجابة الصحيحة فى اختبار الاختيار من متعدد إما:

- باختيار الاستجابة الصحيحة من بين الاستجابات الأخرى وهذا يعرف باسم البحث عن الصواب.
- اختيار الاستجابة الخاطئة من بين الاستجابات الأخرى الصحيحة وفى هذه الحالة نلاحظ وجود أداة استثناء فى عبارة المثير مثل ما عدا، وإلا ويعرف هذا النوع باسم البحث عن الخطأ.
- اختيار الاستجابة الأكثر أهمية أو قوة وهذا يقتضى من التلميذ المفاضلة بين مجموعة من العوامل والاعتبارات ويقوم بترتيبها فى ذهنه وفقا لمعيار موضوعى ثم الاتفاق عليه أثناء التدريس ويعرف هذا النوع بالبحث عن الأهم.

ب . أسئلة الصواب والخطأ:

فى هذا النوع من الاختبارات تعرض على التلميذ عبارة أو جملة متضمنة لحقائق تاريخية أو رقمية أو فكرية ويطلب منه أن يذكر فيما إذا كانت العبارة صوابا أو خطأ أى أن هذا النوع من الاختبارات يمكن أن تسمية اختبارا ذا اختيارين.

جـ . أسئلة التكملة:

يطلب من التلاميذ فى هذا النوع من الاختبارات إكمال بعض العبارات الناقصة أو إضافة بعض الأرقام أو تكملة بعض الرسوم.

د . أسئلة المزاوجة:

يعتبر اختبار المزاوجة من أهم الاختبارات الموضوعية وذلك لارتفاع درجة الموضوعية فيه بسبب انخفاض عامل التخمين مما يزيد من معامل الثبات لهذا الاختبار.

ويكون اختبار المزاوجة من عمودين أو قائمتين إحداهما تمثل المثيرات والأخرى تمثل الاستجابات وتزيد عدد الاستجابات على عدد المثيرات وعلى التلميذ أن يزاوج . أو يوافق بين العمودين ويمكن اعتباره نوعا مطورا من اختبار الاختيار من متعدد.

خطوات إعداد الاختبارات التحصيلية الموضوعية:

أولا: تحديد الغرض من الاختبار:

وذلك من حيث ميادين استخدامه، المستوى التحصيلي للتلاميذ المراد تطبيقه عليهم، زمن الاختبار.

ثانيا: تحديد الهدف من الاختبار:

يفيد تحديد الأهداف التربوية فى عبارات محددة واضحة فى تهيئة المعلم وتوجيهه إلى الأهداف التى يلتزم بقياسها فى الاختبار وتصاغ الأهداف المراد قياسها فى صورة أنماط سلوكية يمكن تقويمها فى سلوك التلاميذ. ومن أهم هذه الأهداف التى وضعها بلوم ورفاقه فى المجال المعرفي، وهى تتكون من عدة فئات متدرجة من حيث بساطتها من البسيط إلى المعقد ومن حيث تجريدها من المادى إلى المجرد، وتتضمن الأهداف المعرفية الفئات التالية:

1 . المعرفة:

استحضار المادة المناسبة للذهن سواء باستدعائها أو التعرف عليها من خلال إطار من الحقائق والمبادئ والطرق وتنقسم إلى الفئات الفرعية التالية:

1- استدعاء المعلومات والتعرف عليها.
2- التعرف على الرموز ذات الأصل المادى مثل المصطلحات والحقائق التفصيلية.
3- التعرف على طرق التعامل مع الرمز.
4- معرفة التصنيفات والفئات.
5- معرفة المحكات.
6- معرفة الطرق المنهجية الملائمة لدراسة مشكلات معينة.
7- معرفة العموميات والتجريدات فى مجال ما والتى تستخدم لتنظيم الأفكار.
8- معرفة المبادئ.
9- معرفة النظريات كمجموعة منظمة مترابطة من المبادئ والتعميمات والعلاقات المتبادلة.

وفيما يلى عدد من الأمثلة التى تساعد على توضيح تلك الأهداف:

1- التعرف على بعض خصائص ثقافة ما.
2- معرفة الطرق العلمية لفحص ودراسة المفاهيم المرتبطة بالصحة.
3- معرفة المبادئ التى تضمنتها نظرية ثورنديك للتعلم.

ويجب ملاحظة أنه لوضع هذه الأهداف فى صورة اجرائية محددة يجب أن تستخدم فى الأفعال المصدرية مثل "أن يتعرف على" "أن يتذكر" "أن يعرف" "أن يكتسب".

2. الفهم:

تتضمن هذه الأهداف أن يستطيع التلميذ شرح وتلخيص معلومات معينة ومحاولة تطبيق ما تعلمه على مجالات أخرى وتنقسم هذه الفئة إلى الفئات الفرعية التالية:

1- فهم المادة المراد تعلمها بدون الحاجة إلى ربطها بغيرها.
2- الترجمة من رموز معينة إلى رموز أخرى.
3- التفسير فى صورة تلخيص أو توضيح.

ومن الأمثلة التى تساعد على توضيح تلك الأهداف:

- المهارة فى ترجمة المفاهيم الرياضية اللفظية إلى صورة رمزية عددية والعكس.
- تلخيص فقرة من فقرات موضوع معين أو تلخيص موضوع بأسره، ومن الأفعال المصدرية التى يمكن استخدامها لوضع تلك الأهداف "أن يترجم" "أن يعبر بعبارة من عنده" "أن يشرح" "أن يعيد صياغة " "أن يميز" "أن يوضح" "أن يقدر" "أن يستنتج" "أن يستكمل" "أن يستوفى" "أن يستخرج".

3. التطبيق:

القدرة على اختيار وتطبيق تجريدات معينة (كالأفكار . والقواعد . والإجراءات) على المواقف الجديدة. وتتضمن هذه الفئة استخدام المجردات فى مواقف مادية ومن أمثلة ذلك القدرة على التنبؤ بالنتائج المحتملة لتأثير

عامل بيولوجى على حالة الكائن العضوية. أو القدرة على تعميم قاعدة أن المعادن تتمدد بالحرارة على معادن لم تدرس بالمعمل.

ومن الأفعال المصدرية المناسبة "أن يطابق" "أن يربط" "أن يختار" "أن يطور" "أن يستخدم" "أن يصنف" "أن يعمم".

4. التحليل:

القدرة على تحليل وتجزئه المفهوم أو التنظيم إلى مقوماته وعناصره لكى تتضح العلاقات القائمة بين تلك المقدمات، وتظهر أسس تنظيمها، وكيفية تأثير تلك العناصر على عوامل وأشياء أخرى، ويتضمن التحليل:

1- تحليل العناصر للتعرف على المقومات الأساسية التى يقم عليها.
2- تحليل العلاقات القائمة بين أجزاء المحتوى لتبين العوامل المنظمة.
3- تحليل المبادئ التنظيمية التى تساعد على وضع المحتوى فى صورة بناء متكامل ومن أمثلة ذلك:
   - دراسة الأدب والأعمال الفنية للكشف عن خصائصها.
   - التمييز بين المبادئ والفروض.
   - ادراك العلاقة بين عناصر المسألة الحسابية.

ومن الأفعال المصدرية المناسبة "أن يميز" "أن يصنف" "أن يحلل" "أن يقارن" "أن يستنتج".

5. التركيب:

إعادة تنظيم العناصر والمقومات فى صورة كلية جديدة وتتضمن:

1- إنتاج بناء جديد يمكن نقله للآخرين.

2- وضع خطة ما موضع التنفيذ.

3- اشتقاق علاقات مجردة.

ومن أمثلة ذلك:

1- كتابة قصة أو قصيدة شعرية.

2- وضع خطة عمل لتنفيذ نموذج من الورق أو الصلصال.

3- الوصول إلى طريقة جديدة لحل مسألة رياضية.

ومن الأفعال المصدرية المناسبة : "أن يكتب" "أن يقص" "أن يكون" "أن يعدل" "أن يقترح" "أن يصمم" "أن يربط بين" "أن ينظم" "أن يؤلف بين عناصر".

6. التقويم:

استخدام محكات موضوعة لإصدار أحكام كيفية تبين مدى ملائمة مادة معينة أو طريقة ما للغرض الذى وضعت من أجله وتتضمن:

1- الحكم باستخدام أدلة داخلية كالاتساق المنطقى.

2- الحكم باستخدام أدلة خارجية كالاتساق مع حقائق موضوعة.

ومن أمثلة ذلك:

1- الحكم على الخطأ المنطقى فى قياس من الأقيسة المنطقية.

2- الحكم على الخطأ فى الاجابة على سؤال ما بالقياس إلى ما هو متعارف عليه علميا.

ومن الأفعال المصدرية المناسبة "أن يحكم على" "أن يثبت صدق" "أن يقيم" "أن يفحص" "أن يقارن" "أن يقدر".

يلاحظ على التصنيف السابق أنه تصنيف متدرج هرمى فتحقيق الأهداف العليا متوقف على تحقيق الأهداف الدنيا. وبالتالي فان المدرس أو الأخصائي التربوي يستطيع أن يحدد الصعوبة التربوية التى يعانى منها التلميذ وأن يحدد الأساس الأول الذى قامت عليه تلك الصعوبة فى ضوء الطبيعة الهرمية لتلك الأهداف وبالتالى يستطيع أن يحدد طريقة مواجهة تلك الصعوبة وكيفية التغلب عليها، كذلك فإن التصنيف السابق يساعد المدرس والأخصائى على تحليل الامتحانات والاختبارات الموضوعة لمعرفة درجة تأكيدها على أهداف معينة دون أهداف أخرى، وبالتالى الاسهام فى تشخيص صعوبات التعلم بالنسبة للتلاميذ وفى بناء الاختبارات التى تقيس المعايير المختلفة.

ثالثا: اعداد تخطيط عام لمحتويات الاختبار:

وفيها يوضع تخطيط عام بمحتويات الاختبار وبحيث يشمل هذا التخطيط مادة الاختبار وأسئلته ونوع القدرات التى تتطلبها الاجابة عنها، ويجرى ذلك من خلال الخطوات التالية:

أ . تحليل محتوى مادة الاختبار:

وذلك بإعداد قائمة بالموضوعات التى تتضمنها المادة الدراسية بأكبر قدر من التفصيل، ثم تحديد الوزن النسبي لكل موضوع دراسي وذلك على أساس الوقت المستغرق فى تدريس كل موضوع، أيضا تحديد الأوزان النسبية للأهداف التعليمية.

ب . إعداد جدول المواصفات:

بعد الانتهاء من تحديد الوزن النسبي لكل موضوع دراسي وكذلك الأهداف التعليمية (السلوكية)، يتم تحديد مواصفات الاختبار وذلك فى جدول ثنائي الاتجاه، وفيه تدون المحتويات رأسيا والأهداف أفقيا وفيما يلى مثالا عمليا لجدول مواصفات وحدة ما فى مادة العلوم للصف التاسع من التعليم الأساسي، وفى ثلاث مستويات فقط من مستويات بلوم):

جدول مواصفات لاختبار تحصيلي فى وحدة الكهربية

فى مادة العلوم للصف التاسع من التعليم الأساسي

موضوعات ***المحتوى***	**الأهداف السلوكية**							**المجموع**
	المعرفة			**الفهم**			**التطبيق**	
	معرفة حقائق	**معرفة مصطلحات**	**معرفة طرق التعامل مع المصطلحات**	**ترجمة**	**تفسير**	**تقدير استقرائى**		
أهمية الكهرباء فى حياتنا							2	2
الكهربية الساكنة	1	2		1	1	1		6
بعض مصادر التيار الكهربى					3			3
الاستدلال على مرور تيار كهربى				2			1	3
الجهد الكهربى		1	2	1	1			5
شدة التيار		1	2					3
البطارية الكهربية			2		1			3
المقاومة الكهربية		1	2		2			5
المجموع	1	5	8	4	8	1	3	30

رابعا: صياغة أسئلة الاختبار:

بعد إعداد جدول المواصفات، تصاغ أسئلة الاختبار ويراعى فى صياغتها الدقة ويتحدد عدد الأسئلة التى يتكون منها الاختبار بناء على الأوزان النسبية للموضوعات التى يشملها الاختبار، وكذلك الزمن المتاح للاختبار ومن الجدير بالذكر أنه كلما ازداد الزمن وعدد الأسئلة كلما كان الاختبار أكثر ثباتا.

ويجب على واضع الاختبار أن يتخير بدقة نوع الأسئلة التى يتكون منها الاختبار بحيث تكون مناسبة تماما لطبيعة المادة الدراسية وكذلك المرحلة الدراسية للتلاميذ.

خامساً: ترتيب الأسئلة

يجب على مصمم الاختبار أن يقوم بترتيب الأسئلة تبعا لمستوى صعوبتها ولكن هذا الترتيب لا يكون نهائيا إلا بعد تطبيق الاختبار وتجريبه وتحليل نتائجه.

سادسا: صياغة التعليمات:

وفى هذه الخطوة توضع بعض التعليمات التى تحدد المطلوب من التلاميذ بصورة واضحة، والخطوات الواجب إتباعها فى الإجابة على أسئلة الاختبار وطريقة الإجابة، وكذلك أى معلومات أخرى يرى مصمم الاختبار ضرورتها بحيث لا يحتاج أى تلميذ لأية إيضاحات أخرى بعد بدء الإجابة كذلك يوضح زمن الإجابة على الاختبار كما يجب أن يضع مصمم الاختبار مثالا توضيحيا للأسئلة. وفى التعليمات ليسهل فهم التلاميذ لطريقة الإجابة، كما يجب تخصيص مكان أيضا يسجل فيه

التلميذ البيانات الخاصة به كالاسم والمدرسة والفرقة الدراسية والجنس وتاريخ اجراء الاختبار وغيرها من البيانات.

تفسير درجات الاختبار:

تطورت الأنظمة المرجعية التى يعتمد عليها تفسير الدرجات المستمدة من أدوات القياس النفسي والتربوي، وكذلك أساليب بناء الاختبارات والمقاييس التربوية والنفسية، وما يترتب على ذلك من توجهات جيدة فى استخدامات هذه الاختبارات والمقاييس وتطبيقاتها المتنوعة.

ويقصد بالنظام المرجعي الأساس الذى يبنى عليه هذا التفسير وتقارن فى ضوئه الدرجات بحيث يمكننا الافادة من المعلومات المستمدة من هذه الأدوات فى الأغراض المختلفة للتقويم، وتشير أدبيات القياس وتطبيقاته التربوية والنفسية إلى أربعة أنظمة مرجعية رئيسية مؤثرة فى حركة تطوير أساليب التقويم وأدواته فى وقتنا الحاضر سوف نعرضها فيما يلي:

1- النظام الذى ينسب أداء الفرد إلى معيار مستمد من الجماعة Norm-Referenced (القياس جماعى المرجع).
2- النظام الذى ينسب أداء الفرد إلى نفسه Self-Referenced (القياس ذاتى المرجع).
3- النظام الذى ينسب أداء الفرد إلى محك Criterion-Referenced (القياس محكى المرجع).
4- نظام القياس الموضوع Objective Measurement .

وفى الجزء التالي سوف نتناول كل من الأنظمة المرجعية الأربع.

أولا: القيـــاس جمـــاعي المرجـــع Norm-Referenced Measurement:

يقوم القياس جماعي المرجع على تقدير الفروق بين الأفراد والتمييز بينهم فى الأداء، حيث تكتسب درجة الفرد على الاختبار معناها عندما تسند إلى معيار مستوى الجماعة التى ينتمي إليها الفرد، ويحدد مستوى الجماعة بمتوسط درجاتها على الاختبار، ويتغير هذا المعيار بتغير الجماعة، حيث تعتمد فلسفة القياس على التمييز بين الأفراد.

ثانيا: القياس ذاتى المرجع:

هذا النظام لا يهتم بموازنة أداء الفرد بأداء أقرانه وإنما يركز الاهتمام على موازنة المعلومات المتعلقة بفرد ما بمعلومات أخرى تتعلق بالفرد نفسه.

وتستمد هذه المعلومات عادة من عدة مقاييس أو من بطارية من الاختبارات تطبق على الفرد نفسه فى وقت معين أو فى أوقات متباعدة، وذلك لاتخاذ قرارات تتعلق بالفروق داخل الفرد وليس بالفروق بين الأفراد كما هو الحال فى النظام جماعى المرجع. ويفيد هذا النظام فى قياس التغير الذى يحدث فى بعض خصائص الفرد مثل تغير الاتجاهات والاكتساب والتحسن الذى يطرأ على سلوكه.

ثالثـاً: القيـاس محكـى المرجـع Criterion-Referenced Measurement :

يهدف القياس محكى المرجع إلى تقويم أداء الفرد وتحديد مدى تقدمه بالنسبة لمجموعة من المهارات أو الأهداف ومعرفة جودة الأداء فى ضوء محكات محددة مسبقاً،والاختبار محكى المرجع هو الاختبار الذى

صمم عن قصد لإعطاء قياسات تفسر مباشرة بالنسبة إلى معايير أداء محددة. بمعنى أنه يهدف إلى تقدير الفرد بصرف النظر عن مستوى أداء أقرانه، حيث تفسر درجة الفرد على الاختبار فى ضوء محك، يحدد حسب الأهداف الموضوعة للقياس.

أهمية التقويم الحقيقى:

يستخدم التربويون التقويم الحقيقى فى تحديد ممارسة المتعلم الواقعية للمشاركة فى تقويم تحصيله بنفسه. فالتقويم الحقيقى يمثل تقويماً واقعياً للأداء بصورة تعكس المستوى الفعلي لما تم تعلمه، ويعرفه البعض بأنه تقويم أداء المتعلم من خلال مواقف الحياة الواقعية، وطرق التقويم الواقعى ليست محددة البنية كما هو الحال فى اختبارات الاختيار من متعدد أو الاختبارات المقننة ولكنها تحتوى على مشروعات تتطلب من الطلاب استعراض مهارات حل المشكلات لديهم فضلاً عن مهارات تحليل المعلومات وتركيبها فى سياق جديد، ولتحقيق تقويم أفضل يؤدى إلى مستوى ممتاز من الأداء فإن الطلبة تحتاج ما يمكن أن نطلق عليه التقويم لتحسين التعلم، ولهذا النوع من التقويم خاصيتان على الأقل هما:

- أنه يقوم على مهمات أصيلة أى المهمات التى تُعلم الطلبة الأعمال التى تواجه الكبار فى مجال عملهم.
- أنها توفر للطلبة والمعلمين التغذية الراجعة والفرص التى باستطاعتهم استخدامها لمراجعة أدائهم بالنسبة لهذه الأعمال أو أعمال مشابهة لها.

ويكـون التقـويم حقيقيـا عنـدما تكـون الاختبـارات قائمـة علـى نـوع الأعمال التى يقوم بها الناس، بدلا من استثارة إجابات على أسئلة سهلة. والتقويم الحقيقى هو التقويم الفعلى للأداء لأننا بذلك نعلم إذا ما كان الطلبة قادرين على استخدام ما تعلموه فى مواقف الحياة التى تقترب كثيراً من مواقف الحياة الفعلية، وإذا ما كانوا قادرين على التجديد والابتكار فى المواقف الجديدة.

وأهم خصائص المهمات الحقيقية ما يلى:

واقعيـة: فالموقف أو المواقف المستخدمة تطابق الطرق التى تختبر بها معرفة الفرد وقدراته فى مواقف الحياة الفعلية.

تتطلب الحكم والتجديد: إذ يجب على الطالب أن يطبق المعرفة والمهارة بحكمة وفاعلية لحل المشكلات غير المنظمة، مثال ذلك عندما يطلب منه تصميم خطة عمل، تتطلب أكثر من مجرد إتباع قواعد أو إجراءات روتينية، أو استدعاء معلومات محددة.

يتطلب من الطالب العمل فى الموضوع بدلا من تسميع أو استرجاع ما تعلمه، بل يجب عليه أن يكتشف ويعمل ضمن المقرر الذى يدرسه سواء كان فى العلوم أو الاجتماعيات أو غيرها من المواد.

تحاكى المضمون الذى تختبر فيه أعمال الكبار سواء كان ذلك فى مكان العمل أو الحياة المدنية أو الحياة الشخصية. ويحتوى هذا المضمون على مواقف محددة من حيث القيود والأغراض والجمهور. والاختبارات المدرسية عادة ما تكون من غير مضمون عملى، إذا يحتاج الطلبة أن يختبروا بأنفسهم كيف يكون عليه الوضع عندما يمارسون العمل فى مواقف تحاكى مواقف الحياة العملية. ومثل هذه الأعمال الحقيقية تلغى

مضار الامتحانات المدرسية من حيث السرية وغياب المصادر والتغذية الراجعة الذى تتميز بها الاختبارات المدرسية المكثفة.

تقوم قدرة الطالب على استخدام معينة من المعلومات والمهارات بفاعلية ومهارة للتعامل مع مهمة معقدة . فمعظم مفردات الاختبارات التقليدية عبارة عن عناصر منعزلة من الأداء، وتبعد عن الطبيعة التكاملية للمهارات التى يتطلبها العمل فى مواقف الحياة الفعلية.

تسمح بفرص للتدريب والممارسة واستشارة المصادر، والحصول على التغذية الراجعة لما يمارسه من أعمال.

وليس معنى ذلك إلغاء دور الاختبارات التقليدية، فهذه لها مكانها من حيث السرية والبعد عن مصادر المعلومات. ولكن المقصود من ذلك أن توجد هذه الاختبارات بجانب التقويم الذى يعتمد على أداء الطالب، وقدرته على تطبيق مهارته ومعلوماته فى مواقف شبيهة بمواقف الحياة العملية.

أهداف التقويم الحقيقى:

تتعدد أهداف التقويم الحقيقي وفقاً لما أورده المنظرون البنائيون (Pett, 1990)، ويمكن إيجازها فيما يلي:-

تنمية المتعلم على الاستجابة لمهام التعلم أو مشكلات الواقع الحياتية، وليس مجرد الاختيار من بين عدة بدائل أو اختبارات تم تحديدها مسبقاً.

- اختبار مهارات التفكير العليا والأساسية.
- تقويم المشاريع الجماعية بصورة مباشرة وحقيقية.

- استخدام عينات من عمل الطلاب، والتى يتم تجميعها خلال فترات زمنية ممتدة.
- الاعتماد على معيار واضح، وهذا يجعل رؤية التلاميذ أكثر وضوحاً.
- السماح بتعددية الأحكام الإنسانية وتنوع احتمالتها.
- إتاحة الفرصة للمتعلمين لأن يقيموا عملهم الخاص بأنفسهم.
- التركيز على أبعاد متعددة للقياس Multidimentionality Scaling بدلاً من بعد واحد كما هو الحال فى الاختبارات التقليدية.
- الدخول فى جوهر التعلم، يهدف مساعدة الطلاب فى التعلم.

وعليه فهو تقويم بنائى يمارس الطلاب من خلاله العمليات العقلية.

ويقتضى التقييم الحقيقي أن تتسم المشكلات أو المهام بالواقعية وباتصالها بشئون الحياة العقلية للطلاب. كما يتطلب شكلاً من أشكال التعاون Cooperation بين المتعلمين، وربما يمكن أن يفيد هنا حدوث قدر من التفاوض الاجتماعي بينهم، فضلاً عن أنه يتجنب المقارنات بين الطلاب والتى تعتمد بدورها على معايير أداء الجماعة، كما يؤكد مسؤلية المتعلم عن عملية التقييم من خلال ما يسمى بالتقويم الذاتي - Self evaluation .

ملامح التقويم الحقيقى :

تتعدد أنماط أو إجراءات التقويم الحقيقى التى يستخدمها المعلمون وتستخدمها المدارس التى تتبنى الفكر البنائي (1991,1994) Paulson & Paulson لتشمل:

تقييم الأداء Performance Assessment: وتختص بقياس قدرات المتعلمين فى إنجاز مهام بتطبيقهم للمعرفة التى بحوزتهم وبعرضهم لإمكانات استخدامها فى حل المشكلات التى واجهتهم بصورة ذات مغزى.

اختبارات الكتابة Writing Tests: وتختص بقياس كل من مهارات فنون اللغة وقياس المحتوى المحتوى المعرفى لمجالات عدة، وذلك حين يطلب من المتعلم كتابة موضوع محدد أو أن يستخدم أنماط مختلفة من مهارات الكتابة "ككتابة تقرير أو مقال"

سجلات الأداء Portfolios: عبارة عن سجلات للتعلم والتقويم يتجمع فيها عينات ممثلة من أعمال المتعلمين (التي توضح، تحصيلهم، وتقدمهم، وجهدهم) وتشمل كل من مخرجات التعلم، إلى جانب عملياته، وقد تركز على مجال دراسي معين أو أكثر من مجال، ويشير نمو سجل الأداء إلى مدى مشاركة المتعلم فى إنتقاة للمواد المختارة، ويوضع فى الحسبان التقويم الذاتى للمتعلم باعتباره أهم مكون فى نمو سجل الأداء. بما يسمح بإطلاع الأباء على الأنشطة التى يمارسها أبناءهم خلال عملية التعلم ومعدل نموهم.

معارض بلوغ المنتهى Culminating Exhibitions: والتي تتطلب من المتعلمين عرض ما تعلموه من محتوى ومهارات أساسية في الفصل أو أمام مجموعة من الرفقاء، وغرضها أن يظهر المتعلم تمكنه من التحديات التي واجهته، والتي تطلبت تحليله البيني للمجالات الدراسية المختلفة، أو تحليله لمجال دراسي محدد.

خرائط المفاهيم Concept Mapping: وهى رسوم تخطيطية ثنائية البعد أو متعددة الأبعاد تعكس مفاهيم بنية محتوى النص، يتم تنظيمها

بطريقة متسلسلة تتخذ شكلاً هرمياً، إذ يوضع المفهوم الرئيسي فى قمة الخريطة وتندرج تحته المفاهيم الأقل عمومية فى المستويات الأدنى، مع وجود روابط توضح العلاقات بين المفاهيم الرئيسية والفرعية.

ونظراً لشيوع ممارسة بعض أنماط التقويم الحقيقي السابقة مثل اختبارات الكتابة وتقديرات الأداء فى رصد مخرجات التعلم وبخاصة المعرفية منها لبعض المواد الدراسية مثل اللغات والرياضيات ضمن ممارسات التقويم التي تتم بمدارسنا فى مصر فى حين لا تمارس باقي أشكال أو صور التقويم الحقيقي الأخرى فى الواقع التربوي، مما يستدعينا التركيز عليها فى هذا البحث وخاصة تقييم الأداء وسجلات الأداء (*) Portfolios، وخرائط المفاهيم وما يتصل بهما من مشكلات الثبات والصدق وفيما يلى تفصيلاً لكل منهما باعتبارهما أدوات بديلة فى إطار التقويم الحقيقي من منظور الفكر البنائي.

تقييم الأداء:

تدعو عمليات تقييم الأداء فى العلوم الطلبة إلى بحث تجريبي لاختبار فرض أو حل مشكلة فالطلبة يعدون ويجرون البحث ثم يكتبون تقريراً ويفسرون النتائج. وتقدم عمليات تقييم الأداء دليلاً يرتبط بالمعرفة

(*) نظراً لحداثه مصطلح Portfolios فى مجال تقويم التعلم البنائى فإن هذا المصطلح ترجم تحت مسميات عدة منها: ملفات التعلم وملفات التقويم، الصحائف الوثائقية، ويفضل الباحثان ترجمته بمعنى " سجلات الأداء " لشمول المعنى لأنواع عديدة من الأداء، وعلى اعتبار أن السجل لفظة مشبعة بالمعانى التى وردت فى الترجمات الأخرى ومنها ملف أو وثيقة أو صحيفة وبالتالى اعتبارها جميعا تنتمى إلى مفهوم فوقى هو المجلد أو الدوسيه أو حافظة الأوراق Folder بما يجعله كمصطلح قادر على استيعاب العديد من المفاهيم الأخرى التى تنطوى تحته فى ضوء علاقة الكل بالجزء وبالتالى فإن مصطلح " سجلات الأداء " يحمل جميع المعانى النفسية والعمليات المتضمنة التى عبر عنها المفهوم كما ورد فى أدبيات التقويم من منظور الفكر البنائى.

الإجرائية والإستراتيجية (باكستر، إلدر، جلاسر سنة 1996م) ويرتبط أحياناً بالمعرفة الافتراضية.

ويتكون تقييم الأداء فى العلوم من (أ) نشاط يفترض مشكلة لها معنى. ويتطلب حلها استخدام مواد محسوسة تتفاعل مع ما يقوم الطلبة بأدائه من أفعال، (ب) صيغة الاستجابة والتى تركز على تقرير الطلبة عن البحث (مثل: تسجيل الاجراءات، رسم مخطط بيانى، رسم جدول كتابة نتيجة)، (جـ) نظام تقدير الدرجات والذى يشمل مهنيين يصدرون حكما على كل من معقوليه القدرة على الدفاع العلمى عن الإجراءات المستخدمة لتنفيذ النشاط، ودقة النتائج (رويز – بريمو – شافيلسون سنة 1996 ب).

ولطرح مثال لهذا التعريف، نستخدم تقييم أداء الـ "جراثيم" (شافيلسون، باكستر، باين سنة 1991م) يستخدم الطلبة، فى هذا التقييم، أدوات معملية فى النشاط (مثل: جراثيم ، طبق، ورق نشاف، ورق أسمر، لمبة، زجاجة رش) لإعداد وإجراء البحث لمعرفة ما هيه البيئة (مثل: رطب أو جاف) التي تساعد على انتشار الجراثيم. ويستجيب الطلبة بالوصول إلى نتيجة عن البيئة وبوصف الإجراءات التي يستخدمها لإجراء البحث. ويسجل أداء الطلبة بوضع الصدق العلمي للإجراء المستخدم فى الاعتبار (مثل: مقارنة البيئات البديلة فى نفس البحث) والنتيجة التى تم التوصل إليها عن طريق البحث.

ولقد ظهر اتجاه قوى فى العقد الأخير من القرن العشرين بين عدد كبير من التربويين من صانعي السياسة التربوية بالولايات المتحدة الأمريكية يؤيد استخدام الاختبارات العملية (Tanner,2001)

والاختبارات العملية (أو ما يطلق عليه أحيانا اختبارات الأداء) طريقة فى تقويم حالة الطالب بناء على الطريقة التى يكمل بها عملا محددا. ومن الناحية النظرية بالطبع، فإن الطالب عندما يختار فى سؤال ثنائي الاستجابة بين إجابة صحيحة أو إجابة خاطئة فإنه يقوم بمهمة ما، وإن كان من الواضح أنها مهمة متواضعة. إلا أن مؤيدي الاختبارات العملية كان فى ذهنهم مخططات تختلف فى معناها عن مجرد الاختيار الثنائي للاستجابة فى أسئلة الصواب والخطأ، أو حتى أسئلة الاختيار من متعدد. والواقع أن ما أوجد هذا الاتجاه فى الولايات المتحدة الأمريكية هو عدم رضا التربويين عن اختبارات الورقة والقلم التقليدية التي أدت إلى الهرولة نحو الاختبارات العملية.

ما هو الاختبار العملي:

رغم أن جميع الاختبارات التربوية تتطلب من الطالب أن يقوم بعمل شيء ما، إلا أن معظم التربويين عندما يتكلمون عن الاختبارات العملية فإنهم إنما يقصدون تلك الاختبارات التى يطلب فيها من الطالب أن ينتج استجابة أصلية. وفى الأغلب الأعم يقوم الممتحن (المدرس مثلاً) بملاحظة الطالب أثناء أدائه حتى يحكم على ذلك الأداء. وقد لاحظ كل من فيتز باترك وموريسون (Fitzpartick & Morrison,1971) منذ حوالي ثلاثة عقود من الزمان أنه لا يمكن التمييز تمييزاً مطلقا بين اختبارات الأداء وغيرها من أنواع الاختبارات. وذكراً أن الفرق الرئيسي بين اختبارات الأداء وغيرها من الاختبارات التقليدية هو فى الدرجة التى يقترب فيها الامتحان من محاكاة الموقف المحك أى فى اقتراب الامتحان من

المجال السلوكي الذى نرغب فى قياسه، وعلى هذا يمكن تصنيف اختبارات الأداء فى أربعة أنواع:

- اختبارات الورقة والقلم.
- اختبارات التعرف.
- اختبارات المحاكاة.
- اختبار عينات العمل.

وبالرغم من أن هذه الأنواع تتداخل مع بعضها البعض إلى حد ما فإنها توفر لنا وسائل مفيدة لوصف الطرق المختلفة التى يمكن أن تستخدم فى قياس المهارات العملية التطبيقية. وفى بعض الحالات قد يكفى استخدام نوع واحد من هذه المهارات العملية التطبيقية. وفى حالات أخرى قد نحتاج إلى استخدام أكثر من نوع من هذه الاختبارات لقياس المهارة المطلوبة. وفى معظم الحالات قد يحتاج الأمر إلى استخدام نوعين أو أكثر فى مراحل مختلفة فى عملية التدريس.

وقد يرغب مدرس مثلا أن يعلم الطلبة كيفية حل المشكلات بشكل جماعي تعاوني، وأراد أن يعرف هل اكتسب الطلبة هذه المهارة الجمعية أم لا، وفى هذه الحالة نجد أن أساس تقدير درجة الطالب يرتكز على مدى إتقانه لتلك المهارة. وقد يرتبط بذلك قرار تربوي يتعلق بمدى حاجة الطالب لتعليم إضافي، أو أن الوقت قد حان للانتقال إلى درس أخر. أى أن اهتمام المدرس الأول هو مدى إتقان الطلبة للعمل مع الطلبة الآخرين للوصول بشكل تعاوني إلى حلول للمشكلات. ويوجد عدة إجراءات للتقويم يمكن تسليطها على مهارات الطالب الجمعية فى كل المشكلات. إلا أنه يلاحظ أن الإجراءين الأولين وهما اختيار استجابة من بين استجابات

أخرى (رقمي 1و2) لا تطلب من الطلبة إنتاج أى شيء. أما بالنسبة للإجراءات الثلاثة الأخرى (أرقام 3و4و5) والتي تطلب من الطلبة بناء الاستجابة فإنها تختلف فيما بينها من حيث درجة تطابقها مع نوع المهمة التي تدعو إليها الأهداف التي وضعها المدرس. وهنا نجد أن الإجراء التقويمي الخامس هو أكثرها تطابقا مع الهدف المطلوب. ومع ذلك فإن الإجراء الرابع أقرب إلى الاختبار العملي من الإجراء التقويمي الأول.

أهداف التدريس

يمكن للطلبة حل المشكلات تعاونيا

5- يعمل الطلبة فى مجموعات صغيرة لحل مشكلات جديدة ويلاحظ المدرس جهودهم ويحكم عليها.

4- يعطى للطلبة مشكلة جديدة ويطلب منهم كتابة مقال عن كيفية قيام الطلبة بحلها بشكل تعاونى.

3-يسأل الطلبة سلسلة من الأسئلة حول طرق حل المشكلات تعاونيا، ثم يطلب منهم إعطاء إجابات قصيرة للأسئلة.

2- يجيب الطلبة على سلسلة من أسئلة الاختيار من متعدد حول الخطوات التى تتبع عند حل المشكلات فى مجموعة.

1- يستجيب الطالب إلى أسئلة صواب وخطأ حول أفضل إجراء يتبع فى حل المشكلة الجمعى.

مجموعة من اختبارات التقويم التى تختلف فى الدرجة التى يقترب بها عمل الطالب من السلوك المحكي المرغوب.

ومن الواضح إذا اختلاف المربين فيما بينهم على تفسير المقصود بالاختبار العملي، حيث يشير مفهوم "الاختبار العملي" إلى إجراءات تقويم

مختلفة. فبعض المدرسين مثلا يعتبرون الاختبارات ذات الإجابات القصيرة واختبارات المقال نوعا من الاختبارات العملية. وبمعنى آخر فإن الاختبارات العملية بالنسبة لهم تعنى أي اختبار يستدعى بناء الاستجابة من جانب الطالب. فى حين نجد مدرسين آخرين أكثر تشددا فى وصف المقصود من الاختبار العملي (Gronlund, 1998)إذ يرى بعض مؤيدي الاختبار العملي ضرورة أن الاختبار يجب أن يتضمن على الأقل المظاهر الثلاثة التالية:

محكات تقويم متعددة: يجب الحكم على أداء الطالب باستخدام أكثر من محك تقويمى. مثال ذلك الحكم على قدرة الطالب على التحدث باللغة الإنجليزية على أساس سلامة النطق والجمل والمفردات.

معايير جودة محددة مسبقا: يجب تحديد المستوى المطلوب لكل محك من محكات أداء الطالب قبل الحكم على جودة أداء الطالب.

الحكم التقويمي: على العكس من طريقة تقدير اختبارات الورقة والقلم الموضوعية والذى يمكن تنفيذه باستخدام الحاسب الآلي، فإن الحكم على أداء الطالب لا يتم إلا عن طريق الملاحظة والحكم الإنساني لتحديد الدرجة التى يمكن بها قبول ما قام به الطالب.

نرى أنه من الواضح إذا طبقنا المتطلبات الثلاثة السابقة على اختيارات التقويم الخمسة، فإن الاختيار الخامس يحقق متطلبات الاختبار العملي، ومن المحتمل كذلك أن الاختيار الرابع يمكن أن يكون اختبارا عمليا كذلك. أما الاختيارات الثلاثة الأخرى فإنها لن تحقق المطلوب من

الاختبار العملي إذا كان تعريفنا لاختبار الأداء يتطلب تضمين محكات التقويم المتعدد، ومعايير الجودة، والحكم التقويمي.

وكثير من مؤيدي التقويم باستخدام الاختبارات العملية يفضلون أن تكون المهمات المقدمة للطالب ممثلة لمواقف واقعية من الحياة وليس مشكلات من المواقف مدرسية. فى حين أن بعض المؤيدين الآخرين للاختبارات العملية قد يسعدهم أن يتضمن الاختبار العملي مشكلات تتطلب أن يعطى الطالب الاستجابة، لا أن يختارها. وهناك فريق آخر مؤيد للاختبارات العملية يرى أن تكون الاختبارات ذات مستوى مرتفع، أى أنها تحقق المستويات العليا من تصنيف بلوم (Linn & Gronlund,2000). وباختصار فإن مؤيدى الاختبارات العملية (أى اختبارات الأداء) يختلفون فيما بينهم حول طريقة قياس الأداء الفعلي للطلبة.

وأحيانا ما نقابل بعض المربين الذين قد يستخدمون عبارات أخرى لوصف تقويم الأداء، مثال ذلك استخدام عبارة "التقويم الأصيل أو التقويم الحقيقي (لأن مهمات التقويم تتفق بشكل أدق مع مواقف الحياة الطبيعية، غير المدرسية) أو "التقويم البديل" (لأن هذا النوع من التقويم يعتبر بديلا للتقويم التقليدي الذي يستخدم اختبارات الورقة والقلم)، وهناك من يستخدم "تقويم الحقيبة الوثائقية" الذى يعتبر نوعا خاصا من تقويم الأداء، ولكن لا يجب اعتباره وصفا مرادفا لأسلوب تقويم الأداء المستخدم فى القياس التربوي.

أنواع تقييمات الأداء:

توجد أنشطة للأداء وبحوث كثيرة، وفى محاولة للتقريب بينهم بتصنيفهم طبقا للتوافق المنتظم فى خصائصهم (شكل 3، أنظر، رويز – بريمو، شافيلسون سنة 1996 شافيلسون، سولانو، فلوريز، رويز – بريمو ، تحت الطبع) ورغم أن تصنيفاتنا (مثل أى نظام تصنيف آخر) تفرط فى تبسيط تعقد وتفرد كل تقييم، فهى تركز على الجماعيات التى أثبتت أنها مفيدة فى تطوير تقييمات أخرى داخل نفس التصنيف، وتسلم قائمة التقييمات الأخرى، والمتضمنة فى التصنيف الذى أعددنا، بجهلنا وباحتمال اكتشاف أنواع أخرى من الأنشطة.

شكل (1)

أنواع الأنشطة وأنظمة تقدير الدرجات المستخدمة فى تقييمات الأداء

أنواع الأنشطة						أنواع أنظمة تقدير الدرجات	
أشياء أخرى	ملاحظات	تصنيف	تحديد العنصر	بحث مقارن			
				– مناشف ورقية. – جراثيم. – مستويات مائلة. – تشبع.	مبنى على الإجراءات	تحليلى	
			– ألغاز كهربائية. – مساحيق ملغزة.		مبنى على الدليل		
		– صخور وخرائط. – يغوص، ويطفو.			مبنى على الأبعاد		
	– علم فلك النهار				مبنى على دقة البيانات		
؟					أشياء أخرى		
	أوراق (تقيم أ 28)				عناوين	كلى	
؟					أشياء أخرى		

ويعرض لنا جدول (1) أمثلة لكل نوع من التقييم ولتركيز صيغة استجابته الملائمة، ولنظام تقدير الدرجات. ونحدد هنا باختصار كل تصنيف ونصف كل مثال:

(جدول (1): أمثلة لأنواع التقييمات المختلفة)

نظام تقدير الدرجات	صيغة الاستجابة	النشاط	نوع التقييم
* يبنى على أساس الإجراءات. * يركز على قدرة الدفاع العلمى عن الإجراءات المستخدمة ودقة النتائج.	تطلب من الطلبة الكتابة تفصيليا عن كيفية إجراء البحث ونتائجه.	* قدم للطلبة 3 مساحيق. * يحدد الطلاب أى المساحيق لديها استعدادا للتشبع بالماء وأيها التى ليس لديها استعدادا لذلك.	بحث مقارن: محاليل مشبعة
* يبنى على تقديم الأدلة. * يركز على الدليل الذى يؤكد أو لا يؤكد وجود مسحوق خاص ويركز كذلك على دقة النتائج.	تطلب من الطلبة كتابة تقرير عن الاختبارات التى استخدموها للتأكد وعدم التأكد من وجود المادة، وكذلك كتابة ملاحظاتهم.	* قدم للطلبة أكياس بها خليط من المساحيق. ويحدد الطلاب أى المساحيق فى كل كيس.	تحديد العنصر: مساحيق **ملغزة**.
* يبنى على أساس الأبعاد * يركز على ملاءمة الخصائص المختارة لتصميم المخطط، ويركز كذلك على دقة استخدام المخطط التصنيفى.	يطلب من الطلاب أن يشرحوا مخطط التصنيف الذى صمموه لتفسير سبب اختيارهم للخواص المستخدمة فى مخططهم التصنيفى.	* قدم للطلبة بعض الصخور. * يصمم الطلبة مخططا تصنيفياً بإختيار الخصائص المناسبة التى تساعد فى تصنيف هذه القطع من الصخور وقطع أخرى.	تصنيف: الصخور والخرائط
* يبنى على دقة	يطلب من الطلبة أن	* قدم للطلبة نموذج	الملاحظة:

البيانات. * يركز على ملاءمة النموذج المستخدم لجميع البيانات، ودقة البيانات التى حصلوا عليها.	يقدموا نتائج ملاحظاتهم، ويفسروا طريقة جمع المعلومات.	للكرة الأرضية * يبين للطلبة مسار الشمس من وقت الشروق لوقت الغروب ويستخدمون الاتجاه والطول وزاوية الظل لحل مشاكل المكان.	علم فلك النهار

يطلب من الطلبة فى البحوث المقارنة أن يقارنوا بين أداتين أو أكثر فى خاصية ما أثناء التحكم فى المتغيرات الأخرى. ويوجد بحث "المحاليل المشبعة " داخل هذه القائمة (شكل 3) ويطلب النشاط من الطلبة فى هذا البحث أن يقارنوا قابلية المساحيق الثلاثة على التحلل فى الماء (جدول 2)، وتدعو صيغة الاستجابة للمحاليل المشبعة الطلبة إلى صياغة ليس نتيجة التشبع فقط بل وصف كيفية إجرائهم للبحث كذلك (بمعنى الإجراء المستخدم)، ويبنى نظام تقدير الدرجات في البحث المقارن على الإجراءات ولكي نقوم بعملية تقدير درجات أداء الطلبة، فالمعلومات عن كل من جودة الإجراءات المستخدمة ودقة حل المشكلة ضرورية، وعلى سبيل المثال، فالدرجة التي يحصل عليها الطالب في تقييم "المحاليل المشبعة"، تبنى على قدرة الدفاع العلمي عن الإجراء المستخدم (**مثل:** هل استخدم الطالب نفس كمية المياه مع كل من المساحيق الثلاثة؟ هل يقيس الطالب بوعي كمية المسحوق المستخدم فى كل محلول؟ وتبنى كذلك على دقة نتائج الطالب (بمعنى أى المساحيق مطلوبة أكثر وأى منها مطلوبة

بشكل أقل حتى تتشبع بالماء؟)، وإذا أهمل الطالب عند قياس الماء أو كمية المسحوق فلابد أن تتصدع عملية مقارنة قابلية المسحوق للتحلل.

ويطلب بحث تحديد العنصر من الطلبة أن يحددوا العناصر التي تشكل المجموع فعلى سبيل المثال ففي بحث "المساحيق الملغزة" تكون مهمة الطالب تحديد الأغراض المنزلية **(مثل:** بيكربونات الصودا، نشا الذرة، ملح**)** التى توجد فى حقيبة "ألغاز"، وتطلب صيغة الاستجابة من الطلبة أن يقدموا معلومات عن الاختبارات **(مثل:** الأيودين، الخل**)** التى استخدموها لتأكيد وعدم تأكيد وجود المادة **(مثل:** بيكربونات الصودا**)** وملاحظاتهم عما يحدث عند استخدام كل اختبار ويبنى نظام تقدير الدرجات على الأدلة – فهو يركز على الدليل الذى يستخدمه الطالب لتأكيد وجود عنصر واحد و / أو لعدم تأكيد وجود عنصر آخر، كما يركز على دقة استجابته لمحتوى حقيبة الألغاز.

ويطلب بحث التصنيف من الطلبة أن يبتكروا مخططا تصنيفيا باستخدام خصائص مجموعة أدوات يمكن أن تصلح للغرض العملي أو المفاهيمى، وعلى سبيل المثال، فيطلب النشاط فى بحث "الصخور والخرائط" من الطلبة أن يضعوا فى اعتبارهم الخصائص المختلفة للصخور **(مثلا:** الصلابة، الطبقات، السلسلة المتصلة ببعضها**)** لابتكار مخطط تصنيفي ولاستخدامه لتصنيف عينة من الصخور وفى صيغة الاستجابة النموذجية يطلب من الطلبة كتابة تقرير عن مخطط التصنيف، وتقديم تبرير للأبعاد المختارة فى هذا المخطط، واستخدام المخطط المبتكر لتصنيف عينة الأدوات المتوفرة لإجراء التجربة عليها. ويبنى نظام تقدير الدرجات على أساس الأبعاد، ويركز على ما إذا كان الطالب استخدم

الخصائص الوثيقة الصلة بالغرض بشكل فردى أو بشكل جماعي لتصنيف الأدوات. فعلى سبيل المثال، إن نظام تقدير الدرجات للـ "الصخور والخرائط" يركز على ارتباط خصائص الصخور التى اختارها الطالب لتطوير مخطط التصنيف، وكيفية استخدامه للمخطط بدقة لتصنيف عينة الصخور المتوفرة لإجراء التجربة عليها.

ويطلب بحث الملاحظة من الطلاب أن يلاحظوا ثم يسجلوا بشكل منتظم خاصية الأداه عبر فترة زمنية، فعلى سبيل المثال، فالنشاط في بحث "علم فلك النهار" تطلب من الطلبة أن يستخدموا معرفتهم بعلاقات الأرض والشمس بالنسبة للظلال النموذجية في أوقات مختلفة من النهار فى نصف الكرة الأرضية الشمالي ونصف الكرة الأرضية الجنوبي لحل مشكلات الموقع، وفى صيغة الاستجابة، يقدم الطلبة معلومات عن البيانات التي جمعوها أثناء ملاحظتهم، وكذلك الطرق التي استخدموها لجمع البيانات، ويبنى نظام تقدير الدرجات على أساس دقة البيانات فيركز على تقدير الدرجات على ما إذا توصل الطلاب إلى نتائج دقيقة على أساس كل من ملاحظة الظواهر الطبيعية، وتطوير نماذج لتفسير ملاحظاتهم ويركز نظام تقدير الدرجات فى بحث " علم فلك النهار " على دقة الملاحظات المتجمعة، والإجراء المستخدم لعمل نموذج لظلال الشمس باستخدام ضوء ساطع ونموذج كرة أرضية.

طرق تقييم الأداء:

لقد تم قياس أداءات الطلبة بطرق مختلفة: الملاحظة المباشرة – يسجل الملاحظ أداء الطالب خطوة خطوة مع البحث، دفتر ملحوظات الطالب – يسجل الطالب إجراءاته ونتائجه فى دفتر، محاكاة الكمبيوتر –

يجرى الطالب بحثا على شاشة الكمبيوتر، الورقة والقلم (تحريريا) – يحل الطلبة مشاكل فى تخطيط وتصميم و/ أو تفسير بحثا تجريبيا افتراضيا (بمعنى اختبارات الإجابة القصيرة والاختيار من متعدد، باكستر وشافيلسون سنة 1994، شافيلسون وآخرون سنة 1991، شافيلسون وباكستر سنة 1992)، وتعتبر الملاحظة المباشرة طريقة القياس المثالية أو وضع علامات للأداء والملاحظ جالس بمكانه، وتعتبر الطرق الأخرى بدائل، ولأن الملاحظة المباشرة مستهلكة للوقت وللمصادر البشرية فالطرق البديلة تستخدم فى الفصول بشكل عام وفى عمليات تقييم شاملة الجودة الفنية ارتبط الاهتمام الكبير الأولى عن عمليات تقييم الأداء بمتغيرية أخذ عينة من المقدرين وكان يفترض أن السلوك المعقد صعب جداً عند الحكم عليه إما فى الوقت الأصلى أو من سجل مكتوب، والبحوث واضحة حول هذه القضية: فيمكن تدريب المقدرين لتقويم الأداء المعقد بثبات (شافيلسون وآخرون سنة 1993) ومع هذا لا يمكن تدريب جميع الأفراد على تقدير درجات الأداء بشكل متوافق، ويجب مراجعة وإعادة تقويم المقدرين باستمرار (ويجدور، جرين سنة 1991).

إن النتائج متوافقة بشكل ملحوظ حول متغيرية أخذ عينة من النشاط من خلال أبعاد متنوعة مثل التحصيل فى الكتابة والرياضيات والعلوم (باكستر وآخرون سنة 1993، دونبار، كوريتز، هوفر سنة 1991، شافيلسون وآخرون سنة 1993)، ومثل أداء الأفراد العسكريين (ويجدور، جرين سنة 1991). إن متغيرية اختيار عينة الأنشطة عظيمة، فعدد كبير من الأنشطة مطلوب للوصول إلى قياس لأداء الطلبة يمكن تعميمه.

إن دراسة واحدة أو ربما الدراسة الوحيدة لمتغيرية اختيار عينة التوقيت مع عمليات تقييم الأداء فى العلوم تشير إلى أن هذا المصدر للمتغيرين ربما يكون كبيراً أيضاً (رويز، بريمو، باكستر، شافيلسون سنة 1993، شافيلسون وآخرون سنة 1993) وفى الواقع، فإن متغيرية أختيار عينة الموعد تندمج مع متغيرية اختيار عينة النشاط لأن عمليات التقييم تجرى فقط على موضوع واحد فى وقت معين (كرونباك ، لين ، برينان، هيرتيل سنة 1997). ولقد وجدنا أن كلا من متغيرية اختيار عينة النشاط والموعد توحدتا لتكون المصدر الرئيسى لخطأ القياس عند تقييم الأداء (شافيلسون، رويز – بريمو، تحت الطبع).

أخيرا، فإن متغيرية إختيار عينة الطريقة هامة جداً لكى تقترح أن الطرق المختلفة ربما تركز على الأوجه المختلفة للتحصيل فى العلوم (باكستر وشافيلسون سنة 1994)، وتعتمد درجة الطلبة على النشاط الخاص المؤدى كعينة والطريقة الخاصة المستخدمة لتقييم الأداء (أنظر باكستر وشافيلسون سنة 1994)، وتقترح البحوث أن طريقة الورقة والقلم أقل قدرة على التبادل مع الملاحظة المباشرة (30 (r <.من الطرق الأخرى. فالملاحظة المباشرة، ودفاتر الملاحظات، والكمبيوتر – جميع الطرق التى تتفاعل مع الأعمال التى يؤديها الطلبة عند إجراء البحث – تبدو أنها أكثر قدرة على التبادل (50) (r <.شافيلسون، رويز – بريمو، تحت الطبع). والدرس المهم المستفاد من هذا البحث هو أن درجات تقييم الأداء تتقلب مع الطريقة المستخدمة.

نقاط الخلاف المثارة حول استخدام تقييمات الأداء. تواجه تقييمات الأداء تحديات فى الوقت الحالى، تحدد الإجرائية فى برامج الاختبارات

الواسعة الانتشار. فعلى سبيل المثال تقييمات الأداء مستهلكة للوقت ومكلفة عند تطويرها وتنفيذها (ستيتشر ، كلين سنة 1997) والتكنولوجيا المتاحة محدودة، ولا تقترب من كفاءة اختبارات الاختيار من متعدد. وبرغم إجراء بعض المحاولات لتطوير نوع من التكنولوجيا (سولانو – فلوريز، جوفانو فيك، شافيلسون سنة 1997، سولانو – فلوريز وشافيلسون سنة 1997، فلم نكن قريبين من تقييم أداء عالي الجودة وإنتقائى مشابه للتقييم المرتبط باختبارات الاختيار من متعدد.

بنوك الأسئلة:

يعد مفهوم بنوك الأسئلة من المفاهيم التربوية المستحدثة والتي شهدت تطورات متلاحقة منذ منتصف الستينات حتى وقتنا الحاضر نتيجة للدراسات والبحوث في مجال القياس والتقويم التربوي.

ويعد المركز التأسيسي القومي للبحوث التربوية في إنجلترا وويلز (NFER) أول من اهتم بإلقاء الضوء علي هذا المفهوم، عندما قرر مجلس التعليم البريطاني عام 1966م تمويل بعض البحوث لدراسة إمكانية بناء أنظمة بنوك الأسئلة في المملكة المتحدة، وتعد هذه البحوث رائدة في هذا المجال حيث انبثق عنها تقرير قدمه Wood and Skurnlk عام 1969م أوضح فيه المفاهيم المختلفة لبنوك الأسئلة، وأساليب بناء أنظمتها. كما انبثقت عنه دورية الامتحانات التي أصدرها مجلس التعليم حول بنوك الأسئلة واستخداماتها في امتحانات المدارس الإنجليزية. **(صلاح علام، 1989: 1–3)**

وقد ميز هذا التقرير بيم مفهوم بنوك المفردات Item Banks ومفهوم بنوك الأسئلة Question Banks حيث يقتصر المفهوم الأول

علـــي أنظمـــة البنـــوك التـــي تشـــتمل علـــي مفـــردات اختيـــار مـــن متعـدد Multiple-Choice item لـذلك تسـمى هـذه البنـوك ببنـوك المفردات الموضوعية Objective Item Banks أما المفهوم الآخر فهو أكثر عمومية واتساعا، ويطلق علي البنوك التي تشتمل علي جميع أنواع الأسئلة سواء التي تتطلب إجابة مقيدة (الأسئلة الموضوعية) أو إجابة حرة (أسئلة المقال بأنواعها المختلفة).

كما أشار التقرير إلي إمكانية اشتمال بنوك الأسئلة علي أنواع أخري من المهام تختلف عن الأسئلة المتعارف عليها ، غذ يمكن أن يشتمل بنك الأسئلة علي تجارب عملية، وموضوعات بحثية، وأسئلة شفوية، وقطع إملائية وغير ذلك، وهنا يشترط أن تدعم هذه المهارات بالإجراءات التي تسمح بتقويمها تقويما كميا.

وعلي الرغم من تباين المفاهيم المتعلقة ببنوك الأسئلة من دولة إلي أخرى غلا أنه يجمعها عناصر مشتركة أهمها:

- توافر رصيد كبير من المفردات أو الأسئلة أو الفقرات الاختبارية لمحتوى أو محتويات دراسية يتم تصنيفها وتنظيمها وفقا لمحكات معينة.
- تحديد الخصائص الإحصائية (القياسية) لمفردات البنك اعتمادا علي البيانات التي تجمع في مرحلة التجريب الميداني لهذه المفردات.
- تشكيل اختبارات ذات خصائص معينة من عينات من هذه المفردات واستخدامها كأدوات لقياس التحصيل الدراسي للطلاب. **(صلاح علام، 1989:4)**.

أهداف واستخدامات بنوك الأسئلة:

من الممكن أن نجمل أهم أهداف واستخدامات بنوك الأسئلة فيما يأتي:

1- تحقيق الموضوعية في تقويم التحصيل الدراسي للطلاب، ويعتبر هذا هو الهدف الأساسي لبنوك الأسئلة، ويعتمد تحقيق هذا الهدف علي مجموعة من العوامل التي يمكن أن تؤثر علي بناء وتكوين بنك الأسئلة، مثل فلسفة القياس التي يعتمد عليها، ومدى الدقة في صياغة البنود (المفردات أو الأسئلة) وتحليلها وتحديد مواصفاتها الإحصائية والتربوية، ومدى تغطيتها لمجال القياس من حيث المحتوى والمستوى، وكيفية تخزينها واستدعائها، والحفاظ علي سريتها وإلي غير ذلك من العوامل.

2- إتاحة الفرصة لقياس العمليات العقلية العليا التي طالما أغفلت في تقدير تحصيل الطلاب وذلك نظرا لاتساع المدى الذي يمكن أن يغطيه البنك من حيث محتوى المقرر أو من حيث المستويات المعرفية المختلفة.

3- توفير وقت وجهد المعلم في بناء الاختبارات التحصيلية وذلك بحصوله علي اختبارات جديدة من بنوك الأسئلة طبقا لمواصفات وأهداف المقرر الدراسي ويتيح له هذا أن يركز جهوده للجوانب الأخرى من العملية التعليمية.

4- تخفيض عوامل القلق والرهبة والتوتر المصاحبة لموقف الامتحان وذلك باستخدام عينات من بنود البنك للتقويم المستمر

أثناء العام الدراسي حتي يألف الطلاب هذا النوع من الأسئلة الموضوعية ولا يؤثر هذا علي سرية الامتحانات.

5- مرونة القياس حيث يسهل تشكيل أي اختبار في أي وقت بسرعة وسهولة.

6- سرعة الحصول علي صور متكافئة من الاختبارات تبعا لمواصفات البنود المكونة للبنك.

7- إمكانية المقارنة الموضوعية بين مستوى تحصيل الطلاب مهما اختلفت الأعوام، طالما أنم البنود قد تم تدريجها ومعايرتها ووضعها في بنك واحد وذلك تبعا لما يسمى بعملية التعادل Equating . ويعتبر تحقيق هذا الهدف منعطفا أساسيا في تاريخ القياس السلوكي ونقله هامة في اتجاه تحقيق الموضوعية.

8- المساعدة في تطوير العملية التعليمية وذلك لما تحظي به بنوك الأسئلة من مميزات يمكن إجمالها فيما يأتي:

- إن السرعة والسهولة التي يمكن بها باستخدام الحاسب الآلي تكوين اختبارات موضوعية جيدة مسحوبة من بنك الأسئلة، تتيح فرصة التقويم المستمر أثناء العام الدراسي بما يساعد علي تشخيص الصعوبات ومحاولة علاجها، وقد يؤثر هذا علي تطوير المناهج وطرق التدريس وربما الأهداف.
- إن تدرج صعوبة البنود ومعايرتها لقياس قدرة الطلاب علي التحصيل الدراسي ، تتيح الفرصة لاستخدام صورة مختصرة من الاختبارات تضم عددا قليلا من البنود تناسب

مستوى الطلاب وتحقق أهداف القياس مما يساعد على سرعة عملية التقويم.

- إن سهولة تكوين الاختبارات المسحوبة من بنك الأسئلة وسهولة طباعتها وتصحيحها باستخدام الحاسب الآلي بالإضافة إلى إمكانية المقارنة بين النتائج مهما اختلفت الاختبارات المستخدمة أو مستويات الأفراد ، يساعد على استخدامها للأعداد الكبيرة والأماكن المترامية الأطراف خاصة انه من الممكن توصيل وحدات الكمبيوتر في تلك المناطق بالحاسب الآلي المركزي ، فبعد اختيار الاختبارات يمكن طباعتها في عدة أماكن في وقت واحد على وحدات الكمبيوتر المختلفة ، وكما سبق أن ذكرنا فانه حتى لو اختلفت بنود الاختبارات المسحوبة فان ذلك لا يؤثر على تقدير مستوى تحصيل الطالب ولا على مقارنته بزملائه في نفس العام أو الأعوام السابقة.

(أمينة محمد كاظم ، 1995:248،249)

مميزات بنوك الأسئلة :

ترتبط مميزات بنوك الأسئلة بأهدافها واستخداماتها السابق توضيحها ، ومن الصعب الفصل بينها ذلك لأن مميزاتها متعددة وهى نفسها استخدامات للبنوك .

ومن أهم المميزات ما يلي:

1- دقة قياس تحصيل التلاميذ باستخدام أسئلة جيدة سبق تدريجها ومعايرتها.

2- استخدام عدد قليل من الأسئلة التى تم تدريجها ومعايرتها لقياس قدرات التلاميذ فى التحصيل الدراسى ، وذلك عن طريق استخدام الأسئلة ذات التدريج القريب من مستويات التلاميذ.

3- وجود مستوى محدد لكل سؤال من الأسئلة التى تم تدريجها ومعايرتها وبذلك فان وحدات القياس المستخدمة للأسئلة هي وحدات ثابتة بدرجة كبيرة.

4- سهولة الحصول على الاختبارات المطلوبة فى وقت قصير بالاستعانة بالحاسب الالى واستخدام برامج مصممة لاستدعاء الأسئلة طبقا لمواصفات معينة.

5- الحصول على أسئلة جيدة واختبارات مطابقة للمواصفات دون معرفة أي فرد بذلك حيث يقوم الحاسب الالى بالاختيار من البنك، كما يمكن تصحيح الاختبارات وتحليل نتائجها.

6- سهولة تنفيذ الامتحانات فى المناطق المترامية الأطراف عن طريق توصيل وحدات الكمبيوتر فى تلك المناطق بالحاسب الالى المركزي ثم توضع شرائط الأسئلة ويتم اختيار الاختبار المرغوب طبقا للمواصفات حيث يمكن طباعته فى عدة أماكن فى وقت واحد على وحدات الكمبيوتر المختلفة، حتى إذا اختلفت أسئلة الاختبارات المسحوبة من البنك فى منطقة عن الأخرى فان ذلك لا يؤثر على تقدير مستوى الطالب ومقارنته بزملائه فى الأعوام المختلفة.

7- تستخدم الاختبارات من بنوك الأسئلة لقياس التحصيل الدراسي للأعداد الكبيرة بسبب سهولة عمل الاختبارات وطباعتها

وتصحيحها التى لا تستغرق وقتا طويلا. **(أمينة كاظم وآخرون، 17،16:1989)**

بنوك الأسئلة من وجهتي النظر جماعية المرجع ومحكية المرجع:

في هذه الأحوال تعتبر بنوك الأسئلة مخزنا كبيرا لمئات البنود التى تتخذ ما اصطلح عليه بالشكل الموضوعي للأسئلة، من حيث صياغتها وتقنينها واستخدامها وتصحيحها وتحليلها وحساب معاييرها، سواء استخدمت فى ذلك وجهة النظر جماعية المرجع أو وجهة النظر محكية المرجع.

وفى كلتا الحالتين ينبغي أن تغطى البنود جميع جوانب المقرر الدراسي، شاملة لجميع أجزائه وأهدافه التقويمية، كما ينبغي أن تغطى بنود البنك جميع الأهداف المعرفية ولا تركز فقط على مستوى استدعاء المعلومات بل ترقى إلى قياس المستويات الأعلى من المعرفة.

أما من حيث صياغة الأسئلة فلا يصح أن ينفرد بصياغتها فرد واحد، بل ينبغي أن يشترك فى صياغتها مجموعة متمرسة من المتخصصين ثم تقوم لجنة ذات خبرة تربوية كبيرة خاصة فى مجال التقويم، بمراجعة البنود وتبويبها وتصنيفها وإعدادها للتجريب على عينات من الطلاب ثم تحليل نتائج الاستجابات وحساب معاملات الصعوبة والتمييز وغير ذلك، ويحفظ الجيد منها فى شرائط أو اسطوانات الحاسب الآلي، أما البنود التى بها عيوب فيعاد فحصها وتعديلها وتجريبها ثانية وهكذا.

بعد ذلك من الممكن أن تقوم اللجنة المكلفة بوضع الامتحان بتحديد المواصفات العامة لبنود الاختبار، ويقوم الحاسب الآلي على برنامج معد

خصيصا لذلك باختيار البنود المطلوبة طبقا لمواصفات المحددة وطباعتها. دون أن تعلم لجنة وضع الامتحان أي البنود سوف تتضمنها ورقة الامتحان.

ويتيح هذا النظام التقنين الكامل للاختبار كما يتيح أيضا تكوين صور مختلفة من الاختبار قد تتباين فى عدد بنودها أو مستوياتها أو مواصفاتها الإحصائية ولكن من الممكن التوصل إلى نوع من التعادل بينها من حيث درجات النهاية الصغرى للنجاح (الدرجة الفاصلة للنجاح) أو بين الدرجات التى تعبر عن التقديرات الجامعية المختلفة (مثال على ذلك كلية طب جامعة عين شمس)، وبالإضافة لذلك فان الوقت المستغرق فى أداء الامتحان وتصحيحه وتحليله النتيجة مفصلة لا يستغرق سوى دقائق لكل مادة دراسية، وسوى ساعات قليلة لجميع المواد الدراسية.

وعندما تستخدم وجهة النظر جماعية المرجع فى بناء بنوك الأسئلة يكون الهدف عندئذ، قياس الفروق بين الأفراد فى التحصيل والتمييز بينهم، هنا يكون اختيار بنود البنك بناء على مستويات الصعوبة ومعاملات التمييز علاوة على الصدق والثبات بحيث تحقق هذا الهدف، وعندما تستخدم وجهة النظر محكية المرجع فى بناء بنوك الأسئلة يكون الهدف عندئذ هو قياس مدى الاكتساب أو النمو أو التحصيل فى تحقيق الأهداف التقويمية بصرف النظر عن مستوى الأفراد، لذا يكون اختيار بنود البنك من حيث مستوى صعوباتها وصدقها وثباتها بالإضافة إلى اختيار المستويات المناسبة لمحكات القياس بحيث تحقق الهدف من وجهة النظر محكية المرجع. **(أمينة كاظم، 1995: 253–255)**

بنك الأسئلة باستخدام نموذج(راش):

تعتبر بنوك الأسئلة من أهم التطبيقات العملية لنموذج (راش) حيث يتيح استخدام النموذج تكوين بنك للأسئلة، بحيث يضم البنك عدة اختبارات تتدرج جميعها في تدريج واحد مشترك وصفر واحد مشترك، وبحيث تعرف هذه البنود مدى واسعا من مستويات المتغير موضوع القياس.

وعندما يستخدم نموذج راش في بناء بنوك الأسئلة، يمكن عندئذ التغلب على مشكلتين من أهم مشكلات التقويم الأكاديمي وهما:

- موضوعية تقدير المستوى الأكاديمي للطالب بالرغم من اختلاف الاختبارات المستخدمة وتباين مستوياتها.
- موضوعية تقدير المستوى الأكاديمي للطالب بالرغم من اختلاف مستوى الشعبة أو المجموعة التى ينتمي إليها.
- ويبدأ بنك الأسئلة بدمج اختبارات في تدريج واحد ، وينتهي بشبكة من الاختبارات التى تغطى المدى الواسع من متغير القياس، وتقوم فكرة تكوين بنك الأسئلة على ما يتمتع به النموذج من خاصية استقلال القياس عن كل من تأثيرات عينة الأفراد ومجموعة البنود المستخدمة، وستتناول الإشارة هنا النقاط الآتية:
    - دمج بنود اختبارين فى تدريج واحد.
    - تكوين بنك الأسئلة.
    - سحب الاختبارات الفرعية من بنك الأسئلة.
    - حبك الاختبار.

دمج بنود اختبارين في تدريج واحد:

وتهدف هذه العملية إلى تحويل التدريج المستقل لكل من الاختبارين إلى تدريج واحد مشترك، ويتطلب هذا التحويل موازنة أو معادلة لتدرج البنود المكونة لكل من الاختبارين. وتتم هذه العملية بأسلوبين يقوم أولهما (وهو الأكثر استخداما) على استخدام بعض البنود المشتركة كرابطة بين الاختبارين ويقوم الأسلوب الثاني (وهو الأقل شيوعا) على استخدام بعض الأفراد المشتركين في أداء كل من الاختبارين وقد فصل هذا فى (**أمينة كاظم، 1988: 108–119**)

تكوين بنك الأسئلة:

تتضمن اغلب بنوك الأسئلة مئات من الأسئلة أو البنود المدرجة على ألاف من الأفراد. ولما كان من المستحيل على كل فرد من الأفراد أن يؤدى كل بند من هذه البنود فان بناء البنك يقوم على دمج المجموعات المختلفة من البنود (أى الاختبارات المختلفة التي تعرف متغيرا واحدا) وذلك في تدريج واحد مشترك، ويوفر نموذج (راش) طريقة خاصة واضحة لبناء بنك الأسئلة ويبدأ الأسلوب الأكثر شيوعا بتدريج عدد كبير من البنود على متغير واحد، وذلك باستخدام مجموعة من البنود المشتركة بين اختبارين مختلفين تعمل كرابطة، تضمنها في تدريج مشترك واحد له صفر مشترك واحد تشكل بنكا للبنود أو الأسئلة، ويتطور الرباط بين الاختبارات المختلفة حتى يصل إلى شبكة من الارتباطات العديدة، كل رباط منها يربط بين اختبارين وتشكل هذه الشبكة من الارتباطات نسيجا من الاختبارات المدرجة جميعها على تدريج واحد، ويغطى بنك

الأسئلة في هذه الحالة مدى واسعا من المتغير، كما يتضمن أيضا صورا متكافئة من الاختبارات التي تغطى المستويات المختلفة من المتغير موضوع القياس.

(أمينة كاظم، 119:1988)

سحب الاختبارات الفرعية من بنك الأسئلة:

كما سبق فان بنك الأسئلة يتضمن عددا هائلا من المفردات التى تشترك جميعها فى تدريج واحد وصفر واحد، والتي يتسع مدى صعوبتها ويزيد عما يستطيع أداؤه فرد واحد من الأفراد، ويتحكم فى بناء البنك عدد البنود التى نود تدريجها فى البنك، والحد الأعلى لعدد البنود المكونة لكل اختبار من الاختبارات الفرعية، ومدى الصعوبة التى نود أن يغطيه هذا البنك.

ويتميز بنك الأسئلة هذا بما يتميز به نموذج (راش) من استقلالية القياس.

ويتحرر بذلك تقدير الأفراد من تأثير البنود المستخدمة، وهذا يعنى تكافؤ تقدير الأفراد (باعتبار الخطأ المعياري) مهما اختلفت البنود المستخدمة، المسحوبة من البنك بشرط مناسبتها لمستوى الأفراد.

ويتيح المدى الواسع من القدرة الذي يغطيه بنك الأسئلة فرصه اختيار المجموعات المختلفة من البنود التى تشكل الاختبارات المناسبة لمستويات الأفراد المتباينة.

كما يتيح ما يتضمنه البنك من ذلك المدى العريض من البنود المتناظرة والمتكافئة الصعوبة , الفرصة لاختيار مجموعات البنود التى

تشكل الصور المختلفة من الاختبارات التى تناسب مجموعات الأفراد المتشابهة فى مستوى القدرة.

وبذا يشكل بنك الأسئلة مصدرا مفيدا لمجموعة من الاختبارات التى تقيس مدى واسعا من مستويات متغيره وسواء أكانت هذه الاختبارات طويلة أو قصيرة سهله أم صعبه واسعة من حيث مدى الصعوبة أو ضيقه فإنها تتعادل فى تقديرها لمستوى قدره الأفراد وبهذا يمكن اعتبار مشكله بناء الصور المختلفة من الاختبارات الموضوعية فى طريقها للحل وذلك بما يتيحة بنك الأسئلة من مرونة في اختيار الاختبارات المختلفة التى تتعادل تقديراتها للأفراد بصوره مباشره.

وهكذا يمكن أن نقارن بين مستويات القدرة للأفراد أو المجموعات المختلفة كما يمكن أيضا قياس التغير الذى يحدث فى مستوى الفرد أو مستوى الأفراد وتكون هذه المقارنة أو قياس هذا التغير باستخدام أي من مجموعات البنود (الاختبارات الفرعية) طالما أنها مسحوبة من بنك واحد للأسئلة وطالما أنها مناسبة للأفراد الذين يؤدون الاختبار وكلما كان الاختبار مناسبة للفرد كلما كان تقدير القدرة أقرب للدقة.

وعلى هذا فعندما يسحب اختيار فرعى ما من بنك الأسئلة ينبغي أن يكون مناسبا للفرد أو مجموعه الأفراد الذين يقومون بأدائه ويكون الاختيار مناسبا عندما تقترب المميزات الإحصائية للاختبار من المميزات التى يمكن أن يتصف بها الفرد أو الأفراد الذين تهدف إلى تقديرهم وعلى هذا الأساس يمكن اختيار البنود التى يكون الاختيار المناسب قدراتهم من بين البنود التى يضمها بنك الأسئلة.

ويكون تقدير المميزات الخاصة بالفرد أو الأفراد بصوره تقريبية إما عن طريق خبره الباحث وتوقعه وإما عن طريق تجربته لبعض الصور المختصرة من المقياس أما المميزات الإحصائية التي ينبغي توافرها فى لاختيار المناسب فتتضمن مستوى صعوبته ومدى هذه الصعوبة وعدد بنوده وقد لخص wright&stone1979 الخطوات التى يمكن بها تصميم الاختبار المناسب للأفراد الذين نهدف إلى تقديرهم وسحب بنوده من بنك الأسئلة. **(أمينة كاظم . 1988 :119–123)**

حبك الاختبار:

تبدو الحاجة واضحة إلى ضرورة تعيين أو تخصيص الاختبار الذى يناسب الفرد حتى لا يجابه الباحث أو المعلم ذلك الموقف الذى يكون فيه البنود غير مناسبة للفرد الذى يجيب عنها لذا كان من المهم اختيار أو سحب بنود من بنك الأسئلة تقترب صعوبتها من قدره الفرد أي تكون هذه البنود محبوكة على مستوى الفرد استطعنا التوصل إلى الدقة فى القياس وتتم عمليه حبك الاختبار للفرد أو لمجموعه الأفراد بعدة طرق (أمينة كاظم و آخرون 1989:30)

وقد تعرضت دراسات كل من: (أمينة كاظم . 1988, 1996 ب), (صلاح علام, 1989: 451 –496), (صلاح علام, 1995), (أمينة كاظم وآخرون, 1989)، (سليمان الخضرى, أمينه كاظم, 1997،1998،2000) إلى أهمية بنوك الأسئلة المكونة من مجموعات من الأسئلة تم تدريجها بواسطة نموذج "راش" بحيث تشكل مقياسا واحدا له صفر واحد مشترك وتدريج واحد مشترك ولهذا التدرج وحده قياس واحده ثابتة على جميع مستويات المتغير كما أوضحت تلك الدراسات كيف

يؤدى استخدام نموذج " راش " فى تدريج البنود أو الاسئله إلى أن تتميز هذه البنوك بأحاديه القياس بمعنى تدرج بنود أسئلة البنك تبعا لصعوبتها بحيث تعرف فيما بينهما متغيرا واحدا كما تتميز بنوك الاسئله البنك باستقلالية القياس بمعنى استقلال قدره الأفراد وصعوبة البنود عن جميع البنود كما يستقل تقدير صعوبة البنود عن مجموعه الأفراد والذين يجرون الاختبار بمعنى توفر موضوعيه القياس الأمر الذى يؤدى إلى الثقة العلمية فى نتائج القياس المبنية على اختبارات تم اشتقاقها من بنوك الاسئله من بنوك الاسئله التى تم تدريج بنودها باستخدام نموذج "راش".

كما أسفرت نتائج دراسة (أمينه كاظم,1996 جـ) عن فاعليه نموذج "راش" فى تحويل التدريج المستقل لاختبارين إلى تدرج واحد مشترك من خلال عمليه معادله لتدريج البنود المكونة لاختبارين (سهل– صعب) فى ضوء بيانات لعينتين مختلفتين من حيث المستوى (مرتفعه القدرة– منخفضة القدرة) أى أن نموذج راش بما يوفره من أحاديه واستقلالية القياس يسهم فى دقه التعادل الراسي vertical equating لتدريج اختبارين.

وهدفت دراسة سون وتشنج (Davison and change 1992) إلى بحث مدى ملائمة النموذج أحادى البعد "راش" والنماذج الثنائية ومتعددة إلا بعد ا لنظرية الاستجابة المفردة فى تدرج مفردات اختبارات القراءة والرياضيات من بطارية woodcock-Johnson النفس – تربوية على عينه من 1000 فرد وقد أسفرت الدراسة عن ملائمة النموذج الثنائي متعدد الأبعاد فى تدريج مفردات الاختبارات التى تتضمن قدرات أحاديه وثنائيه مع خطا معياري محدود كما لا ترتبط درجات النموذج أحادى

البعد خطيا مع تقديرات القدرة متعددة الأبعاد للاختبارات الفردية التي تتضمن أكثر من قدره كما تعطى خطا معياريا كبيرا وهنا تظهر محدودية نموذج "راش" عندما تتضمن الاختبارات أكثر من بعد يقيس أكثر من قدرة بينما تظهر فاعليه النماذج المتعددة رغم محدوديه انتشارها لصعوبة التعامل معها كونها على تعقيدات رياضيه تحتاج الى تدريب.

ودراسة (bay-luz,1998) هدفت إلى بحث الفروق فى أداء الطلاب على مفردات اختبار اختيار من متعدد (انتقاء الاستجابة) والاستجابة المبنية (إنتاج الاستجابة) على المفردات المرتبطة بالتقييم القومي لمستويات التحصيل المتقدم فى التربية (إنتاج الاستجابة) حيث يتم تقدير أداء الطالب فى إطار نظريه الاستجابة المفردة IRT وقد حددت سبع تجمعات للمفردات وحددت نقاط قطع cut point لكل منها لتحديد مستويات التحصيل وذلك بأخذ متوسط تقديرات المشاركين لتحديد النسب المئوية المتضمنة والتى تستخدم لتحديد المنحنيات المميزة للاختبارات حيث أسفرت النتائج عن ان أداء الطلاب على مفردات الاختيار من متعدد يكون أفضل في ظل نظريه IRT والمرتبطة بنقاط قطع محدده من أدائهم على اختبارات إنتاج الاستجابة كما وجد ان توقعات الأداء لتحليل البيانات للمفردات للاختبار (MC) تكون اقل ارتباطا بالأداء الواقعي أو الحقيقي بينما توقعات الأداء للمفردات المرتبطة بإنتاج الاستجابة تكون أكثر مناسبة للأداء الحقيقي.

ويحدد (صلاح علام 2000: 683–864) عددا من المشكلات المرتبطة بالنظرية الحديثة فى القياس وهى:

- المشكلات المتعلقة بتقدير بارامترات النماذج وتقييم حسن مطابقة البيانات لهذه النماذج وإمكانية استخدامها عندما تكون المتغيرات التي تقيسها الاختبارات متعددة الأبعاد وغير ذلك من القضايا السيكومترية والتطبيقية.
- تحتاج النظرية إلى قدر كبير من الرياضيات عالية الإحصائية متعددة المتغيرات.
- البحوث والدراسات حول نظريه القياس وجهت إلى منظري القياس وليس الذين سيقومون بتطبيقه فى بناء الاختبارات وتحليل وتفسير نتائجها.
- عدم توفر برامج للحاسب الآلي فعالة ومناسبة لتقدير بارامترات نماذج السمات الكامنة التي تتضمن النظرية الحديثة IRT.

وقد توصلت دراسة (أمينة كاظم 1996 أ) إلى عدد من المشكلات التى ترتبط بنتائج الدراسات والبحوث السابقة حول تحذير الباحثين عند استخدامهم لنموذج "راش" واعتمادهم عليه كوسيلة موضوعيه لقياس حيث يمكن التحذير فى اختيار عينه التدريج عند بناء اختبار ما وكذلك فى اختبار مجموعه البنود التى يقدر بها قدرات الأفراد صعوبة تطويع النموذج للتطبيق وبنا الاختبارات وتفسيرها بالنسبة للمدرس العادي الذى اعتاد على بناء الاختبارات الموضوعية الشائعة

وتعرضت دراسات كل من (صلاح علام1989) , (أمينة كاظم واخرون1989) ,(شكري سيد احمد 2001) لبعض المشكلات المرتبطة بنظرية الاستجابة المفردة IRT والنماذج المنبثقة عنها وهى :

المقاومة ورفض التجديد والإبقاء على ما هو شائع وسبق قبوله.

الكلفة المالية العالية (كلفه اقتناء الحاسبات الآلية – كلفه تصميم أو شراء البرمجيات الحديثة – كلفه تدريب الخبراء وغير متخصصين على إعداد وبناء بنوك الأسئلة وتطبيقها والتحقق من شروطها)

الحاجة إلى خبراء ومتخصصين على مستوى عال من الخبرة فى مجالي نظريه IRT.