Assessment is
Assessment is one component of teaching and activities
"التقويم هو احدى عناصر التدريس واللنشطه"

By doing assessment. teachers can hopefully gain information about every aspects of their students especially their achievement
" لما المدرس يسوي تقويم لطلابه راح يقدر يحصل على معلومات عن كل جوانب طلابهم خصوصاً انجاراتهم"

An aspect that plays crucial role in assessment is tests.
" والجانب اللي يلعب دور مصيري في التقويم هو الاختبارات"

A good test is constucted by considering the principles of language assessment.
" الاختبار الممتاز يعتمد على المبادئ لتقويم اللغه"

They are 1-Practicality . 2-Validity . 3-Reliability . 4-Authenticity 5- Washback/Backwash
"المبادئ هي
١-التطبيق العملي . ٢-الصلاحيه . ٣-الصلابه. ٤-موثوق.

Practicality :Practicality can be simply defined as the relationship between available resources for the test, ie human resources, material resources, time, etc and resources which will be required in the design, development, and use of the test .
Brown defines practicality is in terms of 1)Cost 2) Time 3) Administration 4) Scoring /Evaluation
التدرب العملي يعرف بـ العلاقه بين الموارد المتاحه للاختبار (اللي هي الموارد البشريه والموارد الماديه والوقت) والوارد اللي راح تكون مطلوبه في مجال تصميم وتطوير واستخدام الاختبار.
شخص اسمه براون عرف التدريب العملي على حسب عدة اشياء اللي هي التكلفه-- والوقت-- والاداره--والنقاط او التقييم-

Cost :The test should not be too expensive to conduct The cost for the test has to stay within the budget Avoid conducting a test that requires excessive budget.
التكلفه : لاينبغي ان يكون الاختبار مكلف للغايه . التكلفه لازم تبقى في حدود الميزانيه . تنجب اجراء الاختبار اللي يتطلب ميزانيه ضخمه .

Time :The test should stay within appropriate time constraints . The test should not be too long or too short.
الوقت : الاختبار لازم يكون ضمن وقت محدد . ومايكون قصير مره ولا طويل مره .

Administration : The test should not be too complicated or complex to conduct. The test should be quite simple to administer
الادارة : الاختبار لازم مايكون مقعد او معقد للاجراء .الاختبار لازم يكون بسيط .

Scoring /Evaluation :The scoring/evaluation process should fits into the time allocation .A test should be accompanied with scoring rubrics, key answers, and so on to make it easy to score/ evaluate.
التقييم واحتساب النقاط : هذي العمليه لازم تناسب الوقت المخصص . ينبغي ان يقترن الاختبار مع عناوين النقاط و مفاتيح الاجابه وغيرها لتجعلها سهله لتقييم واحتساب النقاط .

Validity : Validity of a test is the extent, to which it exactly measures what it is supposed to measure . A test must aim to provide a true measure of the particular skill which it is intended to measure not to the extent that it measures external knowledge and other skills at the same time .
صحة الاختبار هو المدى الذي يقيس بالضبط مايفترض قياسه . ينبغي ان يهدف الاختبار ليأمن المقياس الحقيقي لمهاره معينه ،والذي يهدف لقياس، ليس الى الحد الذي يقيس المعرفه الخارجيه والمهارات الاخرى في نفس الوقت .

Brown proposed five ways to establish validity. They are: 1. Content Validity 2. Criterion Validity 3. Construct Validity 4. Consequential Validity 5. Face Validity
اقترح براون خمس طرق لتاسيس الصلاحيه او صحة الاختبار :١-محتوى الصلاحيه ٢-معيار الصلاحيه ٣-بناء الصلاحيه ٤-منطقية الصلاحيه ٥-مواجهة الصلاحية.

Content Validity :The correlation between the contents of the test and the language skills, structures, etc. with which it is meant to be measured has to be crystal clear
The test items should really represent the course objective

صلاحية المحتوى :الرابط بين محتوى الاختبار ومهارات اللغويه ،والبناء وغير ذلك مع اللي المفروض تكون للقياس وتكون واضحه وضوح الشمس. وعناصر الاختبار لازم تمثل هدف الكورس .

Criterion Validity :This kind of validity emphasizes on the relationship between the test score and the outcome
-The test score should really represent the criterion that is intended to measure in the test.
Criterion validity can be established through two ways.
Concurrent Validity A test is said to have concurrent validity f its result is supported by other concurrent performance beyond the assessment itself (B own,2004 24) For example, the vaidity ofa high score on the finalexamination of a foreign language course will be verified by the actual proficiency in the language
2. Predictive Validity :The predictive validity tends to assess and predict a students possible future success . For example, TOEFL or IELTS tests are intended to know how well somebody will perform the capacity of his/ her English in the future
هذا النوع يركز على العلاقه بين درجة الاختبار والنتائج .
درجة الاختبار ينبغى ان تمثل المعاير الذي يهدف الى قياسه في الاختبار-
يمكن تحديد معيار الصلاحيه من خلال طريقتين- :
١-الصحة المتزامنه :نقول ان الاختبار له صحة متزامنه اذا كانت نتائجه تدعم بتزامن اداء الاخرين ماوراء التقييم نفسه . على سبيل المثال / صحة الدرجه العاليه في الاختبار النهائي في مادة لغه اجنبيه راح يتحقق منها بواسطة الكفاءه الحقيقه في اللغه .
٢-الصحة التنبؤيه :
تميل الى تقييم والتنبأ للطلاب بنجاح في المستقبل على سبيل المثال اختبار التوفل والايلتس تميل الى معرفة كيف ان الشخص سيقدم كفاءته في الانجليزي

Construct Validity :Construct validity refers to concepts or theories which are underlying the usage of certain ability including language ability
- Construct validity shows that the result of the test really represents the same construct with the ability of the students which is being measured .
بناء الصلاحيه : يشير الى المفهوم او النظريات التي يقوم عليها استخدام قدرة معينه بما في ذلك قدرة اللغه .
بناء الصلاحيه يظهر ان نتيجة الاختبار في الحقيقه تظهر قدرة الطالب الذي يجري قياسه- .

Consequential Validity :Consequential validity to refer to the social consequences of using a particular test for a particular purpose.
-The use of a test is said to have consequential validity to the extent that society benefits from that use of the test.
تبعية الصلاحية:للاشاره الي الاثار الاجتماعيه المترتبه على استخدام اختبار معين لغرض معين .
استخدام الاختبار له صلاحية التبعيه الى الحد الذي يستفيد المجتمع من هذا الاستخدام للاختبار - .

Face Validity:

- A test is said to have face validity if it looks to other testers, teachers, moderators, and students as if it measures what it is supposed to measure .
-In speaking test for instance face validity can be shown by speaking activities as the main activities in the test. The test should focus on students activities in speaking, not anything else.
-The test can be judged to have face validity by simply look at the items of the test.
-Note that face validity can affect students in doing the test -

-To overcome this, the test constructor has to consider these :
a. Students will be more confident if they face a well- constructed, expected format with familiar tasks.
b. Students will be less anxious if the test is clearly doable within the allotted time limit.
c Students will be optimistic if the items are clear and uncomplicated (simple)
d. Students will find it easy to do the test if the directions are very clear
e. Students will be less worried if the tasks are related to their course work (content validity).
f Students will be at ease if the difficulty level presents a reasonable challenge.
-في الاختبار لديك صلاحية المواجهه اذا كان يبدو للمختبرين والمعلمين والمشرفين والطلاب كما لو انه يقيس مايفترض قياسه .
-في اختبار التحدث على سبيل المثال صلاحية المواجهه يمكن ان تظهر من خلال انشطة التحدث مثل اللنشطه الرئيسيه في الاختبار. يبنبغى ان تركز الاختبار على الانشطه الطلابيه في الحديث وليس شي اخر.
-يستطيع الاختبار ان يحكم اذا كان لديه صلاحية المواجهه بواسطة النظر لعناصر الاختبار ببساطه.
-ملاحظة صلاحية المواجهه ممكن ان توثر على الطلاب اثناء الاختبار.
-منشئ الاختبار لازم ياخذ بالاعتبار هذي النقاط :
أ-الطلاب راح يكونون واثقين اكثر اذا واجهو اسئله منشاها ،وتنسيقها متوقع مع مهام مالوفه .
ب-الطلاب بيكونون اقل قلق اذا كان الاختبار واضح وقابل للتنفيذ خلال الوقت المخصص .
ج-راح يكون الطلاب اكثر تفاؤل اذا كانت العناصر واضحه واقل تعقيد.
د-راح يشوفون الطلاب الاختبار سهل انهم يحلونه اذا كانت التعليمات واضحه.
هـ-راح يكونون الطلاب اقل قلق اذا كان الاختبار مرتبط بالماده اللي يدرسونها .
و_الطلاب راح يكونون هادئين اذا كانت مستويات الصعوبه تقدم تحديات معقوله.

Reliability :
-Reliability refers to the consistency of the scores obtained .
-It means that if the test is administered to the same students on different occasions (with no language practice work taking place between these occasions) then it produces (almost) the same results.
- Reliability actually does not really deal with the test itself It deals with the results of the test. The test results should be consistent.
تشير الى تطابق الدرجات التي تم الحصول عليها.
-معنى هذا اذا كان الاختبار يقدم لنفس الطلاب في مناسبات مختلفه (مع عدم وجود عمل ممارسة اللغه تكون بين هذي المناسبات) ثم راح تعطينا نفس النتايج تقريباً.
-الموثقيه في الحقيقه لاتعمل في الحقيقه مع الاختبار نفسه ،تعمل مع نتايج الاختبار. نتايج الاختبار لازم تكون ثابته.

Reliability falls into 4 kinds.
They are:
1)Student-Related Reliability
2) Rater Reliability
3) Test Administration Reliability
4) Test Reliability

Student-Related Reliability This kind of reliability refers to temporary illness, fatigue, a bad day, anxiety and other physical or psychological factors of the students. Thus, the score obtained of the student may be not his/her actual score.
هذا النوع من الثقه يشير الى مرض موقت او التعب او يوم سي او القلق او اعراض جسديه او نفسيه اخرى من الطالب .
وبالتالي ،النتيجه التي حصل عليها قد لاتكون النتيجه الحقيقيه .

Rater Reliability :
Rater reliability deals with the scoring process. Factors that can affect the reliability might be human error subjectively, and bias in scoring process.
This kind of reliability fall into two categories. They are:
1. Inter-rater reliability It occurs when two or more scorers yield inconsistent scores of the same test, possibility for lack of attention to scoring criteria, inexperience, inattention, or even biases.
2. Intra-rater reliabiliy It is a common occurence for classroom teacher because of unclear scoring criteria, fatigue, and bias toward particular good or bad students or simple carelessness.

التصنيفات الموثقيه تتعامل مع عملية احتساب التقاط.العناصر اللي ممكن تاثر في الموثقيه قد تكون خطا بشري ذاتي ،والتحيز في عملية احتساب النقاط. ولها نوعان : ١ -الموثقيه بين المقيم :
يحدث عندما اثنين او اكثر من محتسبين النقاط يخضع مجموعات غير متناسقه من نفس الاختبار، امكانية لعدم الاهتمام بمعايير احتساب النقاط ،عدم الخبره الغفله،التحيز
٢الموثقيه داخل التصنيفات :
ظهورها شائع لمدرسين الفصوب بسبب عدم وضوح معايير احتساب النقاط التعب والتحيز اتجاه طلاب ممتازين او كسلانين او ببساطه عدم المبالاه.

Test Administration Reliability:
- Test administration reliability concerns with the condition and situation in which the test is administered
-To increase the degree of this kind of reliability test, teachers as the administrators should consider all the things related to the test administration.
-For instance, if we want to conduct a listening test, we should provide a room which is very comfortable to listening environment. The noise from outside the room cannot enter the room. The audio system should clear to all students Even, we have to consider the lighting, the condition of the desks and chairs as well.
-يهتم بالحاله والموقف التي يخصع فيها الاختبار .
-لرفع درجات هذا النوع من اختبار الموثقيه، المعلميين كمشرفين لازم ياخذون بعين الاعتبار كل الاشياء المرتبطه بادارة الاختبار.
-على سبيل المثال ،اذا اردنا ان ننتج اختبار السمع ،لازم نامن غرفه اللي تكون مريحه جداً لمحيط السمع .الازعاج خارج الغرفه لايمكن ان يدخل هذي الغرفه . نظام السمع لازم يكون واضح لكل الطلاب لازم ناخذ بعين الاعتبار الاضواء وحالة الطاولات والكراسي .

Test Reliability: Tests reliability refers to the test itself Whether the test fits into the time constraints.
-It means that the test should not be too long or short
-The items of the test should be crystal clear that it will not end with ambiguity.

اختبار الموثقيه:يشير الى الاختبار نفسه سواء كان يناسب قيود الوقت .
-يعني ان الاختبار لايكون طويل مره او قصير مره
-وعناصر الاختبار لازم تكون واضح وضوح الشمس والتي لاتنتهي بالغموض.
Authenticity:
- Authenticity deals with the "real world"

-Authenticity is the degree of correspondence of the characteristics of a given language test task to the features of a target language task
-Teachers should construct a test with the test items are likely to be used or applied in the real contexts of daily life.
-Brown (2004:28) also proposes considerations that might be helpful to present authenticity in a test They are:
1. The language in the test is natural as possible.
2. Items are contextualized rather than isolated.
3, Topics are meaningful (relevant interesting) to the learners
4. Some thematic organization to items is provided, such as through a story or episode.
5. Tasks represent or closely approximate real-world tasks.

الاصاله (من الاصلي)

تتعامل مع العالم الحقيقي

- هي درجة من توافق خصائص اختبار اللغه المعطى الى خصائص اللغه الهدف .

- ينبغي على المعلمين انشاء اختبار مع عناصر الاختبار من المرجح استخدامها في الحياة اليوميه.

- يقترح براون الاعتبارات التي ممكن ان تكون مفيده لصحة الاختبار وهي :

١- ان تكون اللغه في الاختبار طبيعه قدر الاستطاعه

٢-ان تكون العناصر ضمن نصوص بدلاً من ان تكون معزوله

٣-المواضيع تكون ذا معنى للطلاب

٤-تقدم بعض المواضيع المنظمه للعناصر من خلا القصص او الحلقات

٥-تمثل المهام تقريباً مهام العالم الحقيقي

washback/Backwash :
-The term wasback is commonly used in applied linguistics it is rarely found in dictionaries
-However, the word backwash can be found in certain dictionaries and it is defined as an effect that is not the direct result of something
-In dealing with principles of language assessment, these two words somehow can be interchangeable
- Washback (Brown, 2004) or Backwash (Heaton, 1990) refers to the influence of testing on teaching and learning
-The influence itself can be positive or negative (Cheng et al. (Eds) 2008: 7-10

مصطلح الووش باي يستخدم في اللغويات التطبيقه ونادراً مايوجد في القاموس .

-نحصل معنى كلمة الووش باك في قاموس معين ويعرف بـ التاثير الذي لايكون بنتائج مباشره لشيء

-بالتعامل مع مبادئ تقييم اللغه هذي الكلمتين نقدر نستخدمها بشكل تبادلي يعني نقدر نقول ووش باك او باك ووش

-يشير الى تاثير الاختبارات على التعليم والتعلم

والتاثير قد يكون ايجابي او سلبي

Positive washback
Positive washback has beneficial influence on teaching and learning.
- It means teachers and students have a positive attitude toward the examination or test and work willingly and collaboratively towards its objective
-A good test should have a good effect.
-For example, UN (National Examination) will require students to pay attention to the lessons more attentively prepare everything dealing with UN more thoroughly learn the lessons by heart, and so on. UN will also require teachers to teach the lessons harder than before, give their students extra lessons, and give tips and tricks to study effectively and efficiently the extent that these activities increase such activity and motivation, the UN can be said it has positive backwash

تاثير الووش باك الايجيابي على التعليم والتعلم . يعني ان المعلم والطالب لهم تاثير ايجابي اتجاه عمليه الاختبار ،والعمل بعزيمه وتعاون باتجاه هدفها .

-الاختبار الجيد لازم يكون له تاثير جيد .

-تتطلب من الطلاب انهم ينتبهون للدرس ،ويجهزون كل شي يتعامل مع (الاختبارات المحليه)UM على سبيل المثال ،الامم المتحده الامم المتحده من خلال تعلم الدورس من قلب وغير ذلك .والامم المتحده راح تطلب مدرسين يدرسون الدروس اصعب من قبل ،يعطون الطلاب حصص اضافيه،ويعطون نصائح وخدع للدراسه بفاعليه وكفاءه الى الحد الذي هذه النشاطات تزيد نشاطهم ودوافعهم ،الامم المتحده نقدر نقول انها لها تاثير ايجابي للووش باك

Negative washback :

Negative washback does not give any beneficial influence on teaching and learning (Cheng and Curtis, 2008 9).

Tests which have negative washback is considered to have negative influence on teaching and learning

-For example, if the UN affects teachers to give more the UN lessons then they tend to ignore other lessons that do not contribute directly to pass the exam, the UN can be considered having negative washback Moreover, if the UN threatens the students in facing the exam, the students will feel a lot of anxiety about the exam. They have to pass the UN otherwise they will fail. The consequence surely will affect students in facing the UN They will fell that they are under pressure. Thus, it will affect the students performance. Then, it can be inferred that if the UN has this kind of effect, it has negative washback

سلبية الووش باك : لاتعطي اي تاثير مفيد للمعلم والمتعلم

-الاختبارات اللي لها ووش باك سيئ تعتبر لديها تاثير سلبي للتعليم والتعلم

-على سبيل المثال اذا الامم المتحده تاثر على المدرسين ليعطون دروس كثيره ،فانك بالتالي راح تتجاهل بعض الدروس التي لاتاثر مباشر في اجتياز الاختبار، الامم المتحده تعتبر ان فيه ووش باك سلبي اذا كان الامم المتحده تهدد الطالب لمواجهة الاختبار الطلاب راح يحسون بقلق شديد باتجاه الاختبار . يجب عليهم ان يجتازونه وبعض راح يفشل .العوقب بتاكيد راح تاثر على الطلاب في مواجهة الامم المتحده راح يحسون انهم تحت الضغط . وبالتالي راح تاثر على اداءهم .وبالتالي راح نعرف ان الووش باك سلبي .

The quality of washback might be independent of the quality of the test

-An ordinary test will have positive washback if it has beneficial effect on students as well the as the teachers,

-A standardized test, for instance UN will have negative washback if it does not give beneficial effect on students or teachers

-It seems that there is no way to generalize about washback at the present time Teaching and learning will be impacted in many different ways depending upon the variable at play at specific contexts. What these variables are, how they are to be weighted, and whether we can discover atterns of interaction that may hold steady across contexts is a matter for ongoing research .

- In the meantime, teachers as the test constructor need to consider the probability of the washback of tests which will be constructed and what the future impact on teaching and learning later on.

جودة الووش باك تعتمد على جودة الاختبار.

-الاختبار العادي راح يكون له تاثير ايجابي ،اذا كان له تاثير ايجابي على الطلاب زي ماهو على المدرسين .

-الاختبار الموحد ،على سبيل المثال الامم المتخده راح يعطينا تاثير سلبي اذا كان مياثر بالايجابيه على المعلمين والطلاب

-يبدو انه لايوجد اي طريقة لتعميم الووش باك في الرقت الحاضر .التعليم والتعلم راح يتاثر بطرق مختلفه اعتماد على المتغيرات التي تلعب دور معين في سياق محدد . ماهي هذي المتغيرات ،وكيف لها ان تكون موزونه واذا كنا نستطيع ان نكتشف صيغ من التفاعل التي قد تبقى ثابته من خلال مختلف السياقات .انها مساله للبحث المستمر .

- في هذه الاثناء الاساتذه كمنشئين للاختبارات يحتاجون للاخذ بعين الاعتبار الاحتمال الوش باك للاختبارات التي سوف تنشأ و ماهي السمات التي ستاثر في التعليم والتعلم لاحقاً.