



Cours Magistral 3^{eme} LFEP

Comment choisir ou créer des tests d'évaluation de capacité et de condition physique

Mme Ben ezzeddine Lamia
Gzorges cazorla/ Luc Léger

Ethique et sécurité

- **Les règles d'éthiques**
- **La sécurité**
- **Conclusion**
- **Pistes de réflexion à propos des règles d'éthique et de sécurité**

Les règles d'éthiques

Le choix des tests se fait en respectant le droit de l'être humain. Ainsi On doit tenir compte de certaines règles d'éthique et de sécurité.

Le droit à l'information et le libre consentement

L'individu doit être informé de la procédure à suivre lors de l'expérimentation des épreuves à subir (type d'évaluation, risques, objectifs, précaution...). L'information doit être accessible et compréhensible par l'individu. Un risque par exemple c'est de négliger de donner toutes les informations ou de faire passer des épreuves supplémentaires (de manière plus ou moins inconsciente).

De plus, il est nécessaire d'obtenir un consentement libre et éclairé de l'individu pour qu'il puisse passer les épreuves, c'est à dire l'individu doit être apte à donner son consentement sans pressions psychiques et en fonction de l'information sur l'épreuve, même si l'épreuve est **passive** (plis cutanés, questionnaire...).

Règles d'éthique et évaluation de la condition physique

1. Retenir les objectifs d'évaluation en fonction du bien-être du sujet.
2. Choisir les épreuves et mesures en fonction des objectifs retenus.
3. S'assurer que l'évaluation n'entraînera pas de risques démesurés pour la santé, le bien être et la dignité du sujet.
4. Informer les principaux intéressés sur les objectifs et la nature des épreuves et mesures proposées.
5. Utiliser un formulaire approprié de consentement libre en connaissance de cause.
6. Communiquer et commenter les résultats au sujet évalué.
7. Adopter une attitude responsable qui dépasse les simples exigences légales.

Le respect de la dignité, de l'honneur et de l'intimité du sujet

L'intimité du sujet peut poser problème avec certaines manipulations (ex: plis cutanés) ou pour des évaluations en groupes (cas de personnes obèses); les réactions individuelles sont aussi à prendre en compte pour éviter les contrariétés de chacun.

Les résultats doivent être confidentiels : c'est le secret professionnel.
Ils doivent être transmis à l'évalué et commentés.
Toute publication de résultat sous entend une autorisation préalable.

On peut donc constater la nécessité pour l'évaluateur de faire preuve de tact pour préserver l'individu de toutes atteintes. L'individu doit donc être informé, consentir librement à passer les tests tandis que de son côté, l'évaluateur lui doit tout le respect.

La sécurité

La sécurité se fait par le choix d'épreuves à risques physiques limités. On évite les épreuves contraires à la santé, celles qui sont dangereuses. Il existe dans ces épreuves des risques liés à la santé de l'individu (athlètes, élèves, adultes, personnes avec problèmes cardiovasculaires ou autres handicaps). Il ne **faut pas omettre** ces risques **ni les exagérer**. Cela vaut autant pour les risques **psychiques** que **physiques**.

Note 1. Le **risque nul** n'existe pas. Le risque est le piment de la vie. Le juste milieu est un risque **socialement, médicalement et légalement acceptable** dans une **culture donnée**, une notion en constante **évolution** nettement tributaire de la **mondialisation**.

La sécurité

Note 2. Les risques physiques, **cardiovasculaires surtout**, sont souvent et nettement **exagérés**, même par des spécialistes, parce que basés sur une **opinion biaisée plutôt que sur des faits documentés**. On en arrive donc à **condamner tout effort maximal** que ce soit dans un contexte d'entraînement ou d'évaluation. On en arrive donc à encourager **paradoxalement** la sédentarité qui comporte **plus de risques à long terme**.

Ces épreuves d'effort maximaux ne sont pas plus risqués que les autres activités : les risques sont minimes et connus.

La sécurité

Note 3. Les risques **psychiques** associés aux efforts maximaux (**sensations désagréables, confirmation d'une aptitude médiocre...**), particulièrement chez des individus **sédentaires**, sont souvent **plus importants que les risques physiques. Sans préparation psychologique** ou **progression adéquate**, on risque **d'éloigner les gens plutôt que de les inciter à adopter des habitudes de vie saines.**

Conclusion

Ainsi il faut respecter l'individu en l'informant et en lui laissant le libre choix de faire ou non l'épreuve d'effort ou même les mesures anthropométriques ou questionnaires.

Il faut aussi assurer sa sécurité avec des épreuves adaptées en connaissant les risques qui sont susceptibles d'arriver en cours d'épreuve et en maîtrisant les précautions et mesures d'urgence appropriées.

L'évaluation peut être considérée comme une habitude de vie saine, par exemple, pour connaître sa condition physique et mérite d'être encouragée.

Pistes de réflexion à propos des règles éthique et de sécurité en Tunisie

Ces règles sont très récentes (1995) et oblige pour toute évaluation sur l'être, vivant à passer par un comité d'éthique. Ce comité se réunit une fois tous les 1 à 3 mois.

-Point positif: ce comité évite les dérives et les «expérimentations sauvages ».

-Point négatif: c'est un frein pour l'expérimentation sur l'être humain et nécessite un délai supplémentaire pour toute expérience comprenant l'individu.

Au niveau de l'EPS, cela pose le problème plus large de la présence d'un médecin (normalement obligatoire lors d'épreuve d'effort maximale ou d'exercice physique maximal). Il réside donc une ambiguïté en ce qui concerne ces règles éthiques au niveau de l'EPS:

Doit on pratiquer l'EPS avec un médecin à nos côtés ?

L'élève a t il vraiment le choix de faire ou non son épreuve d'effort ?

**CONDITIONS DU CHOIX
ET/OU DE LA CONSTRUCTION
D'UNE BATTERIE DE TESTS**

NON REDONDANCE
des tests au sein
d'une batterie

PERTINENCE
Dépend des capacités
à évaluer (à identifier
préalablement)

VALIDITE
avoir la preuve
que le test mesure
bien ce qu'il est
censé mesuré

FIDELITE (reproductibilité)
les résultats du même test
passé plusieurs fois consé-
cutivement avec les mêmes
sportifs doivent rester stables

FONCTIONNALITE
Un test est dit fonctionnel
lorsque ses résultats peuvent
être utilisés dans les processus
d'apprentissage et d'entraînement

ACCESSIBILITE
Facile à mettre en œuvre
et utilisation des outils
disponibles

Caractéristiques métrologiques des épreuves

Validité

Difficulté

Fidélité

Discrimination

Précision

Normes

Validité

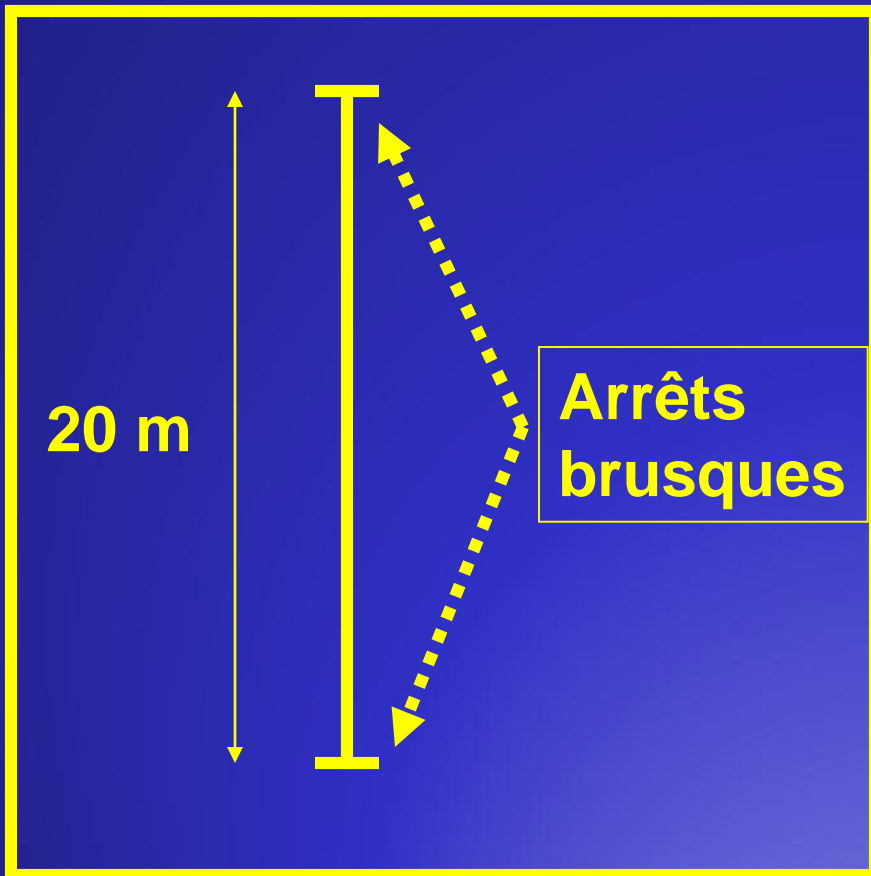
- Concept
- Degrés de validité
- Indices de validité
- Corrélation simple de Pearson
- Corrélation multiple
- Corrélation partielle
- Corrélation biaisée
- Validité interne et validité externe

Concept

Un test est valide lorsqu'il mesure bien ce qu'il est censé mesurer. Par exemple, une sprint de 100 m mesure-t-il bel et bien l'aptitude aérobie ?

Si l'on se sert d'un test qui mesure plusieurs facteurs, il faut que le facteur que l'on veut mesurer joue un rôle plus important que les autres.

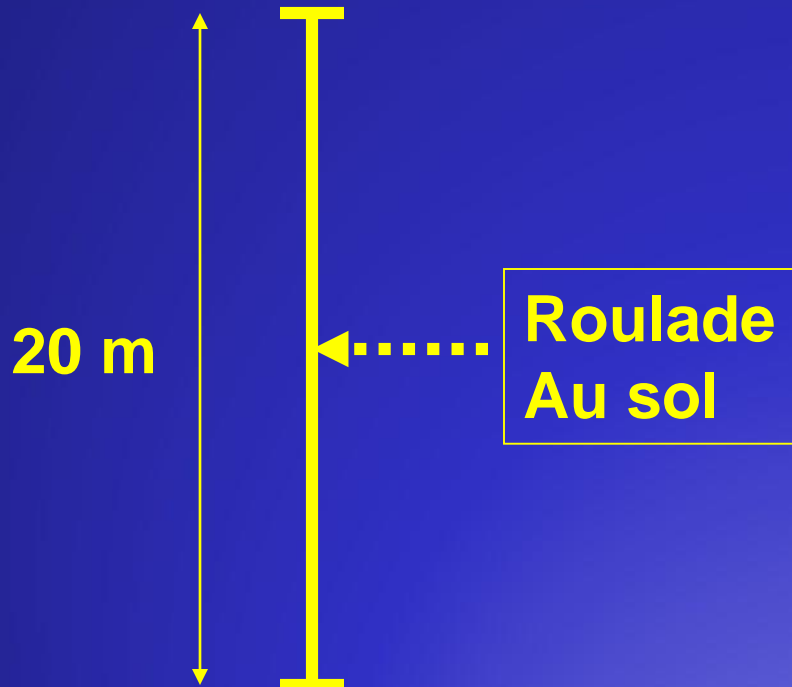
EXEMPLE 1: Navette 20 m en PATINS



- **Mesure-t-on le VO₂max ?**
ou
l'habileté à faire des arrêts brusques ?

Difficulté de départager ces deux facteurs chez les débutants !

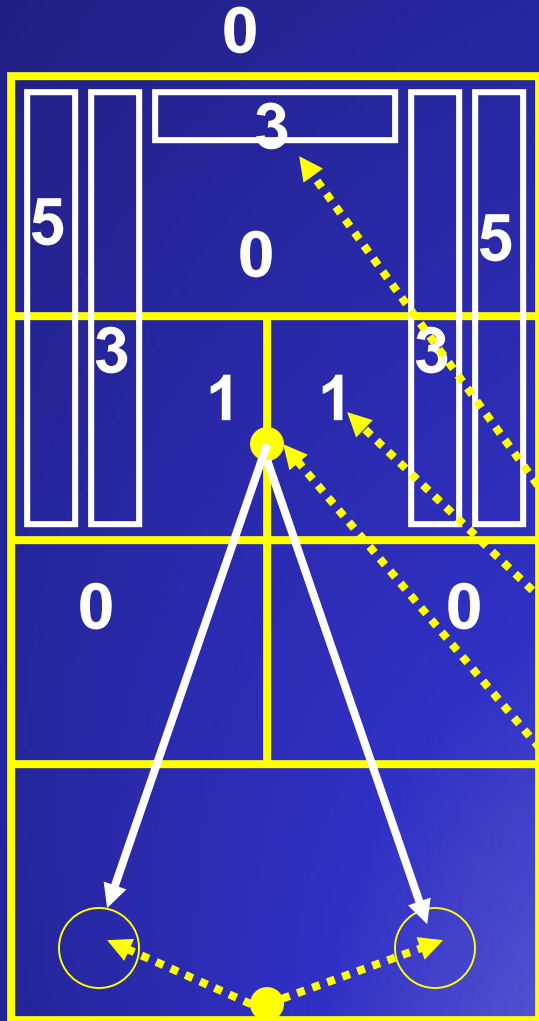
EXEMPLE 2: Navette 20 m au JUDO



- **Mesure-t-on le VO₂max ?**
ou
l'habileté à faire des roulades ?

Difficulté de départager ces deux facteurs !

EXEMPLE 3: Test Navette Aérobie Intermittent au TENNIS



- Cette fois le fait de frapper la balle n'affecte pas le nombre de paliers atteints
- D'autre part, on mesure aussi la précision en fatigue croissante

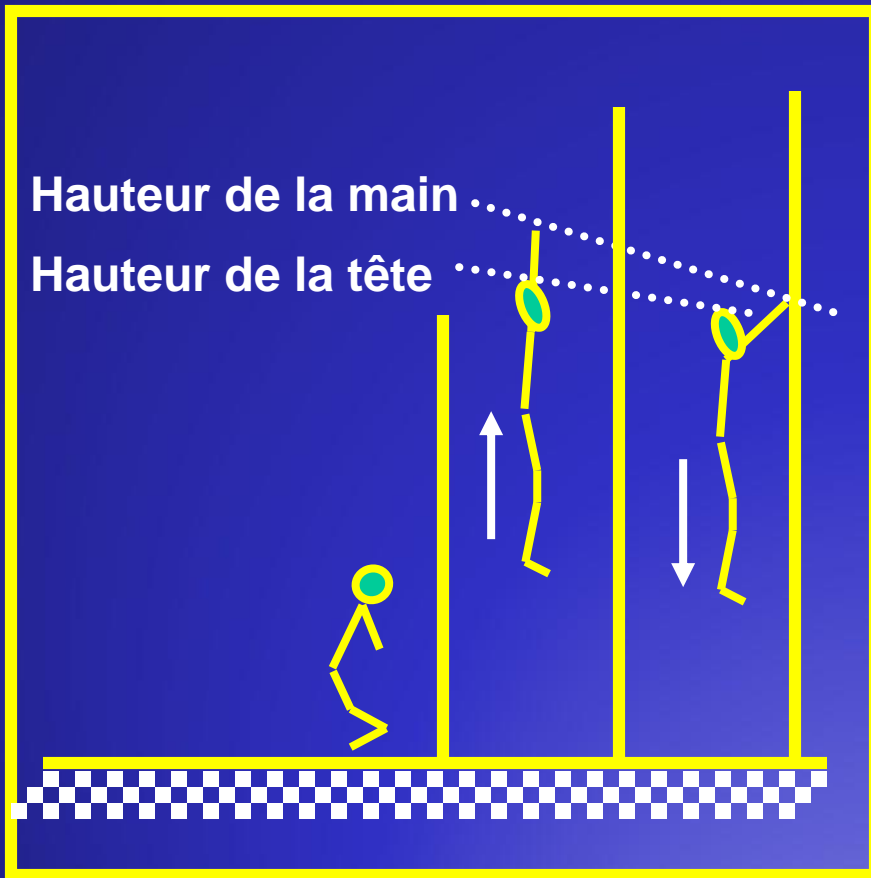
Pointage pour
Retour de balles

Placement cadencé
de balles par l'entraîneur

Déplacement G-D aléatoire

Adapté de Précigout & Fargeas,
Limoges, 2000 (Léger,2001)

EXEMPLE 4: Saut vertical en touchant le mur de côté le plus haut possible

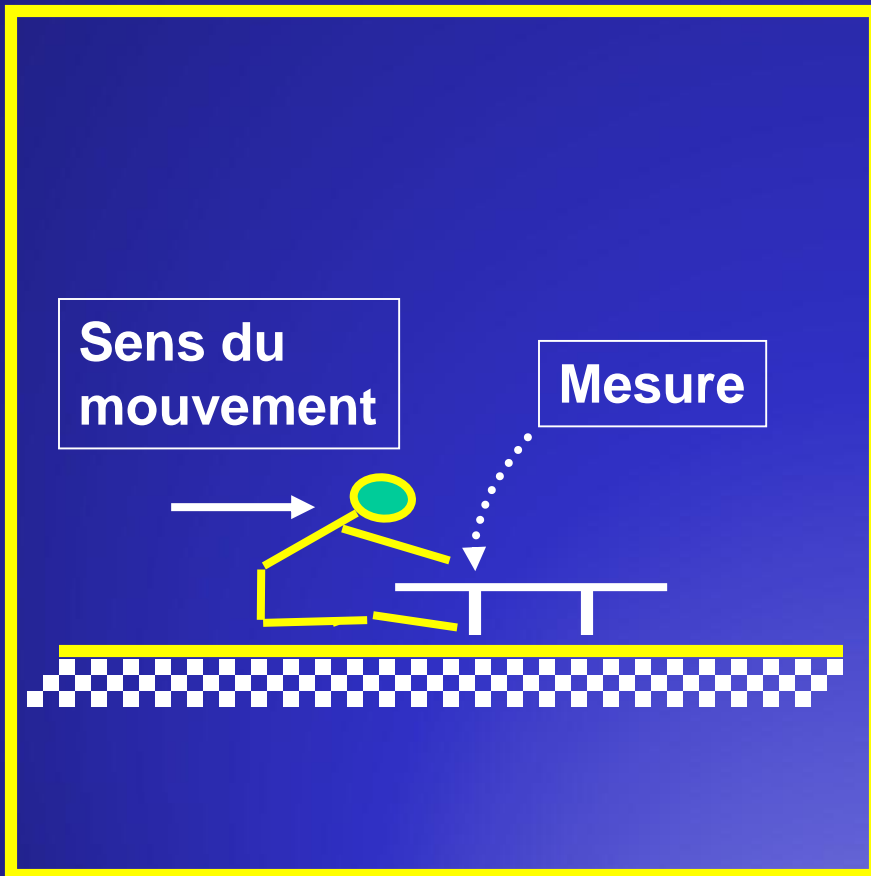


- **Mesure-t-on la puissance ?**
ou
la coordination ?
 - Toucher du mur en descente
 - Trop loin du mur

Difficulté de départager ces deux facteurs !

Plusieurs essais et/ou apprentissage

EXEMPLE 5: FLEXIBILITÉ – Flexion du tronc sur les jambes



- **Mesure-t-on la flexibilité ?
ou
la taille relative
des segments ?**

**Difficulté de départager
ces deux facteurs !**

**SI ON VEUT MESURER
PLUSIEURS QUALITÉS LORS
D'UN MÊME TEST, IL FAUT
S'ASSURER D'UN MINIMUM
D'INTERACTIONS ENTRE CES
QUALITÉS DANS LA TÂCHE
EXÉCUTÉE**

**C'EST POUR CELA QU'EN GÉNÉRAL,
LES TESTS SONT FAITS DE TÂCHES
SIMPLES FACILES À STANDARDISER
ET FACILES À MESURER**

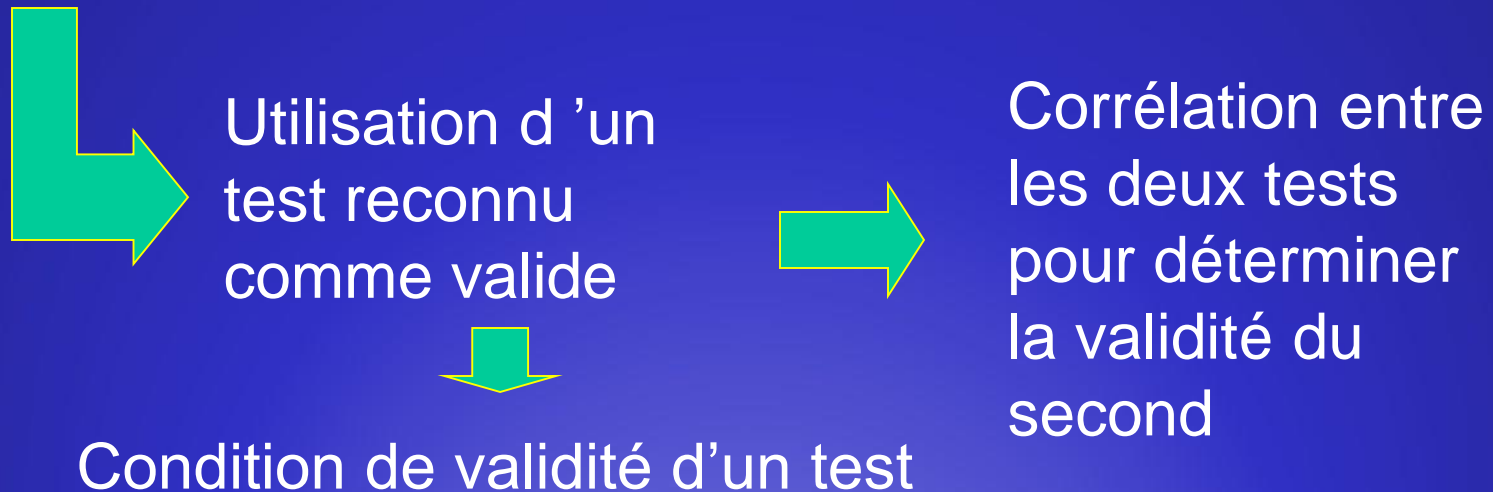
- Fiabilité des appareils de mesures,**
- Objectivité des évaluateurs,**
- Objectivité de mesure unidimensionnelle**

Indice de validité d'un test

Rapport entre ce que le test mesure réellement et la qualité physique que l'on veut mesurer

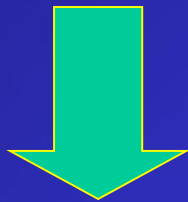
Corrélation simple de Pearson

On peut objectiver la démarche en quantifiant statistiquement la relation qui existe entre la mesure du test et la qualité mesurée.

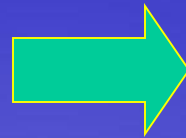


Corrélation multiple

Il arrive que l'on veuille prendre plusieurs tests ou variables pour prédire telle ou telle qualité.



Utilisation de la statistique pour traiter les mesures combinées comme un tout et les mettre en corrélation avec le test retenu comme critère.



Interprétation comme une corrélation simple

Corrélation Biaisée

Deux tests peuvent avoir une corrélation élevée sans que l'on puisse prédire l'un par l'autre.

Exemple 1: VO_2 max exprimé en L/min est fortement corrélé à la taille et semble donc fortement dépendant de celle-ci. En fait, cette mesure dépend surtout de la masse musculaire totale. Mais cette masse musculaire totale est aussi dépendante de la taille surtout en période de croissance. Mais si on recrute des sujets, de même poids, VO_2 max n'est plus lié à la taille alors que pour des sujets de même taille, VO_2 max reste fortement corrélé au poids. C'est ce que la corrélation partielle permet de départager mathématiquement.

Corrélation Biaisée

Deux tests peuvent avoir une corrélation élevée sans que l'on puisse prédire l'un par l'autre.

Exemple 2: VO_2 requis pour courir à une vitesse donnée, exprimé en $ml\ kg^{-1}\ min^{-1}$ est fortement corrélé à l'âge et semble donc fortement dépendant de celui-ci. En fait, cette mesure dépend surtout de la masse corporelle totale. Mais cette masse musculaire totale est aussi dépendante de l'âge surtout en période de croissance. Mais si on recrute des sujets, de même poids, VO_2 requis n'est plus lié à l'âge alors que pour des sujets de même âge, VO_2 requis reste fortement corrélé au poids. C'est ce que la corrélation partielle permet de départager mathématiquement.

Corrélation partielle

Exprime la contribution partielle d'un facteur ou d'une mesure en éliminant mathématiquement l'influence des autres pour prédire la mesure critère.

Validité interne et validité externe

Validité interne : voir si les individus ont les mêmes performances entre deux tests équivalents. Dépend de l'erreur aléatoire.

Validité externe : vérifier si les mêmes résultats seraient obtenus si l'étude était reprise sur un échantillon différent ou sur le même échantillon. Dépend de l'erreur systématique.

Validité interne et validité externe

Validité interne : Un lien fort est observé entre une variable ou test prédicteur et la variable prédite.

Validité externe : Le lien fort observé antérieurement dans l'étude originale est confirmé par plusieurs autres études, dans des conditions similaires et à plus forte raison, dans des conditions différentes (populations, environnement...). On parle alors de **robustesse**, d'approche « **généralisable** ».

Cela concerne autant les **erreurs aléatoires** (corrélation, erreur type de l'estimé) que les **erreurs systématiques** (les équations développées donnent des valeurs prédites semblables d'une étude à l'autre).

Fidélité

Définition du concept

Indice statistique de fidélité

Fidélité interne et externe

Exactitude et précision

Définition du concept

Fidélité mais aussi stabilité-objectivité

- Un test est fidèle lorsque, exécuté à deux occasions, il donne les **mêmes résultats**.
- Lorsqu'il s'agit de **constance** dans le temps on parle aussi de stabilité **dans la mesure**.
- Lorsqu'il s'agit de **constance** entre deux évaluateurs ou appareils de mesure on parle **d'objectivité du test**.

Indice statistique de fidélité

- Comme pour évaluer la validité d'un test un des indice de fidélité est **le coefficient de corrélation (simple de Pearson)**.
- Mais les résultats sont toujours **plus élevés** que ceux de la validité

CEPENDANT

- Les deux paires de donnés qui sont mises en corrélation pour établir la fidélité et la validité d'un test sont **différentes**.
- Le coefficient de corrélation de fidélité est appréhendé soit :
 - entre un test et un retest
 - à partir d'une même mesure obtenue à deux occasions différentes

ATTENTION

- Un simple coefficient de corrélation ne peut déterminer si un test est fidèle. On peut obtenir une corrélation parfaite, avec des résultats non identiques
- C'est **pourquoi**, il faut également faire un test de **différence entre les moyennes des deux essais**.

Fidélité interne et externe

FIDELITE EXTERNE

La source des variations (ou erreur extérieure au sujet). Elle dépend plutôt du test, de l'appareil de mesure ou de l'évaluateur.

—————> Moyenne des essais.

FIDELITE INTERNE

La source des variation provient de l'individu lui même. Variation physiologique, psychologique etc

—————> Meilleur score

Les différentes erreurs

Les erreurs pouvant compromettre la validité et la fidélité d'un test sont de deux types :

- **L'erreur aléatoire** : rarement identifiable lors d'un test unique elle est non définie. La solution pour diminuer cette erreur serait de faire plusieurs essais pour un même test.
- **L'erreur systématique** : elle est facilement corrigible car elle est identifiable. En fait elle concerne essentiellement les appareils de mesures.

Les normes

Utilité

Moyenne et écart types

Normes centiles

Distribution normale

Grandeur de l'échantillon

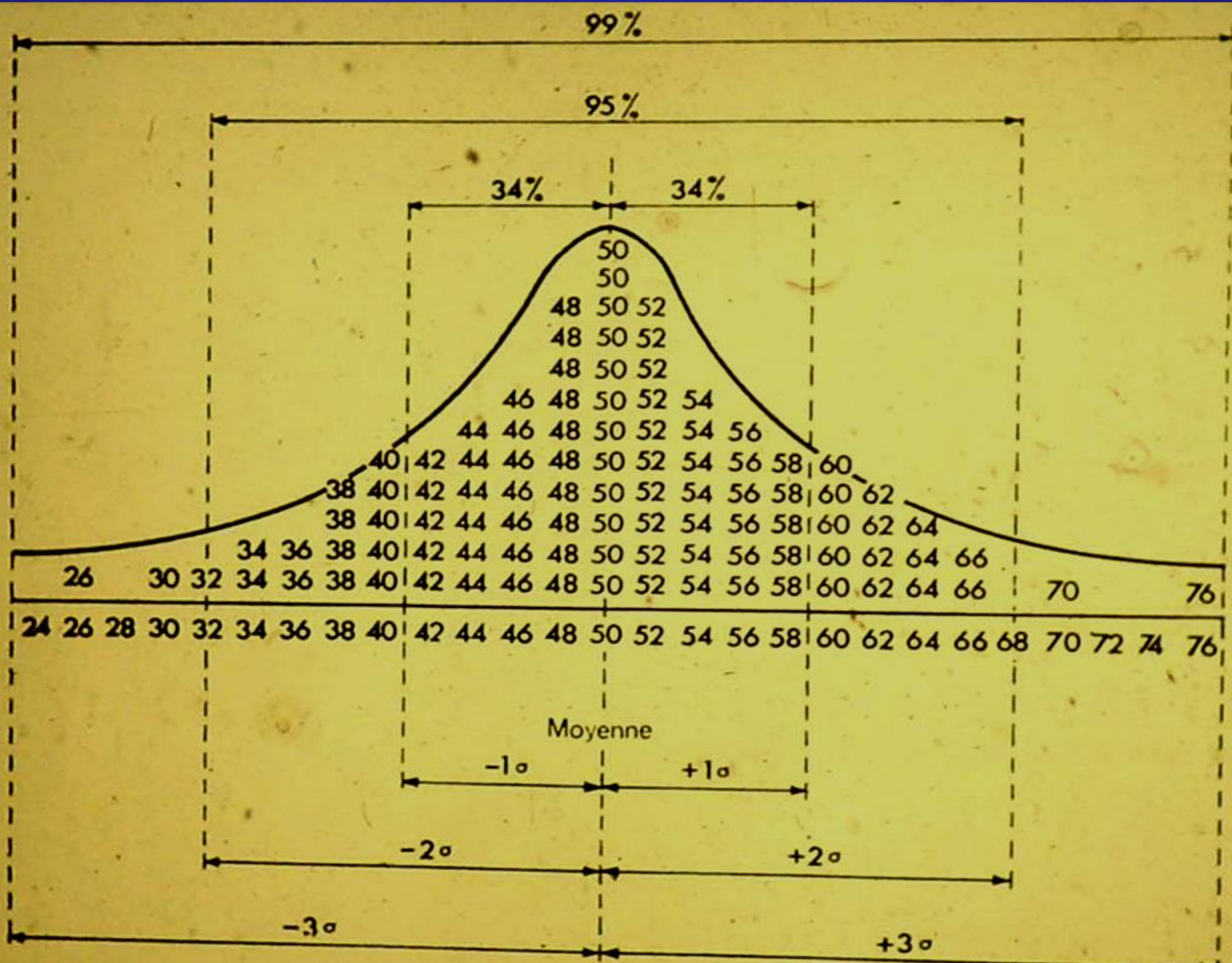
Utilité

Les normes sont utilisées pour situer les individus par rapport aux autres. Ces « autres » peuvent représenter la population générale ou une sous population (exemple : élite de différentes discipline sportive, adulte non compétiteur 30-40 ans, femmes...).

Moyennes et écart-types

La moyenne du groupe est une norme explicite :

- Le sujet est-il près ou loin de la moyenne ?
Pour cela il est important de définir l'écart-type pour savoir comment les individus se répartissent autour de la moyenne.
- Exemple : si la moyenne du $\dot{V}O_2$ max des hommes de 30 à 39 ans est de $35 \text{ ml d 'O}_2 \cdot \text{min}^{-1} \cdot \text{kg}^{-1}$ et l'écart type $5 \text{ ml O}_2 \cdot \text{min}^{-1} \cdot \text{kg}^{-1}$ cela veut dire que 68% de ces individus se situent entre 30 et 40 $\text{ml d 'O}_2 \cdot \text{min}^{-1} \cdot \text{kg}^{-1}$.
- N. B. : 99% des gens se situe + ou - 3 écarts types de la moyenne.



Normes centiles

Une autre façon de décrire la moyenne d'une population et sa distribution autour de la moyenne, est de rapporter les différents centiles et les scores correspondants. La moyenne correspond généralement au 50^{ème} centile, c'est à dire qu'il y a 50% des individus au dessus et 50 % au dessous de la moyenne ; aussi, si la catégorie ou classe faible est situé entre le 20^{ième} et le 40^{ième}, elle représentera alors une classe de 20% des individus (20% au dessous et 60 au dessus de cette classe).

On situe donc quantitativement un individu par rapport au groupe.

Normes centiles

Cependant, il faut se garder de porter un jugement de valeur (exemple classe forte, classe faible) mais plutôt s'intéresser au centile d'une catégorie. Ceci, parce que les normes centiles ne sont pas toujours identiques. En effet, certains vont vouloir rendre les classes extrêmes moins accessibles par préférence, d'autres voudront avoir des catégories similaires par exemple 5 classes de 20% chacune.

Distribution normale

Les trois concepts résumés précédemment ne prennent de sens que si l'on distribue « normalement » la population autour de la moyenne, pour ce faire, on part de 2 postulats de base suivants :

- autant de sujets sont situés de part et d'autre de la moyenne
- la distribution est unimodale (une seule population distincte).

Ainsi, si la distribution est normale on a une courbe en forme de cloche, telle que 68% des sujets se situent autour de la moyenne dans l'écart-type correspondant.

Le graphique ci-dessous permet une comparaison aisée des quatre types de normes.

