

The background of the slide features a close-up of a pencil with a yellow eraser and a sharpened lead tip, resting on a white test paper. The paper is filled with multiple-choice bubbles, each containing a letter (A, B, C, D) inside a circle. Some bubbles are filled with a dark grey or black ink. The overall scene is brightly lit, with soft shadows from the pencil.

PSYCHOMÉTRIE ET TESTS

PSYCHOMÉTRIQUES

2^{ÈME} ANNÉE LICENCE

ORTHOPHONIE

- **Université** : ABDERRAHMANE MIRA BEJAIA
- **Chargée de cours** : Mme MEKHOUKH H.
- **Année universitaire** : 2020 / 2021

- **Les objectifs du cours**

- la présentation des notions nécessaires (prérequis) pour comprendre et interpréter des données en psychométrie.
- Saisir et comprendre les concepts liés à la mesure et à l'évaluation en psychologie.

PROGRAMME

- 1- Définitions de psychométrie
- 3- Niveaux de mesure
- 2- Théories de mesure
- 4- Définitions de tests
- 5- Les qualités psychométriques de tests
- 6- Quelques exemples de tests psychométriques

- **Définition de la psychométrie**

- La psychométrie concerne l'ensemble des principes à la base de la mesure en psychologie (Jean-Luc Roulin, 2018)
- La psychométrie est la science qui étudie l'ensemble des techniques de mesures pratiquées en psychologie, ainsi que les techniques de validation et d'élaboration de ces mesures (J.-L. Bernaud, 2007)
- La psychométrie est la science de la mesure en psychologie. Elle se repose sur l'usage de procédures rigoureuses s'appuyant notamment sur l'usage de techniques statistiques variées.
- La psychométrie, on pourrait la définir comme «l'ensemble des théories et des méthodes régissant la construction des tests psychologiques ainsi que l'évaluation de leurs qualités métrologiques et de leurs utilités» (Jean Bégin, 2020)

- **Qu'est ce que mesurer**

- La mesure est partout. De façon générale, on définit la mesure comme étant «un ensemble de méthodes utilisées pour donner une description (quantitative) d'une caractéristique spécifiée d'un phénomène». (J. Bégin, 2020)
- La mesure est une notion complexe qui concerne des caractéristiques d'un objet (le poids, la dimension, la surface, la quantité, la capacité, ...) ou d'une personne (la taille, la tension, l'intelligence, la mémoire la température corporelle, etc.).
- De façon générale, mesurer c'est attribuer des nombres aux objets, selon des règles déterminées.
- Ces règles vont toujours avoir pour objet d'établir une correspondance entre certaines propriétés des nombres et certaines propriétés des objets (Roulin, 2018)

- **Les échelles de mesure**

- 1/ Échelle nominale**

- **Une échelle nominale répartie les observations dans un certain nombre de classes disjointes (distinctes), telles que chaque observation entre dans une seule classe. L'ensemble des classes utilisées constitue l'échelle nominale.**
- **En d'autres termes, on effectue une partition de l'ensemble des observations (application d'une relation d'équivalence) et tous les objets ou les sujets d'une même classe sont considérés comme équivalents.**
- **Attention : ce n'est pas un critère statistique qui définit la partition, c'est un critère empirique. Ce critère détermine la signification à attribuer à la mesure.**

- **2/ Échelle ordinale**

- Les échelles ordinales possèdent les propriétés des échelles nominales (effectuer une partition des observations), mais les objets d'une catégorie ne sont pas seulement différents de ceux d'une autre catégorie, il existe entre les catégories de l'échelle une relation d'ordre stricte ou non.
- Pour construire une échelle ordinale, il faut donc :
- Effectuer une partition de l'ensemble des observations (relation d'équivalence).
- Définir une relation d'ordre stricte ou non.
- **Exemple d'échelle ordinale** : Les échelles d'appréciation par lesquelles on demande aux sujets d'exprimer des jugements sur un « objet » (comme Très bon, Bon, Moyen, Mauvais, Très mauvais), sont des échelles ordinales.

- **3/ Échelle d'intervalle**

- Dans l'échelle d'intervalle, la mesure implique, en plus des propriétés des échelles ordinales (partition des observations et relation d'ordre stricte ou non), la notion de distance. L'unité de distance donne la signification à la mesure (par exemple : le temps en millisecondes).
- Cette unité de distance est stable tout au long de l'échelle, ce qui signifie que l'on peut comparer la différence observée entre deux mesures à la différence observée sur deux autres mesures.
- Les opérations arithmétiques peuvent s'appliquer sur les nombres représentant les classes.
- Dans les échelles d'intervalles le point zéro est arbitraire.
- **Exemple** : un exemple typique est la température mesurée en degrés Celsius.
- Nous pouvons dire qu'une température de 60 degrés est plus élevée qu'une température de 50 degrés, et qu'une augmentation de 30 à 60 degrés est deux fois plus importante qu'une augmentation de 30 à 45 degrés. Le point zéro est par contre arbitraire et on ne peut pas dire que 60° Celsius est deux fois plus chaud que 30° Celsius.

- **4/ Échelle de rapport**

- Une échelle de rapport est une échelle d'intervalle dans laquelle le point zéro n'est pas arbitraire (comme le temps de réponse ou une mesure de vitesse).
- Les propriétés de ce type d'échelle sont rarement utilisées en psychologie.
- Les échelles de rapport représentent des rapports, car elles ont une origine absolue (correspondant à l'absence de l'attribut mesuré).
- Exemple, la distance a pour origine 0 (absence de distance) et 40 mètres est deux fois plus loin que 20 mètres. Ce n'est pas le cas d'une échelle d'intervalle comme la température exprimée en Celsius où le 0° est arbitraire.
- Une température de 40° n'est pas deux fois plus chaude que 20°. Pour connaître le rapport entre ces deux températures, il aurait fallu prendre une mesure absolue de la température en Kelvin (qui est une échelle de rapport) et on peut alors comparer les deux mesures en Kelvin et en faire le rapport

- **Les théories de mesure**
- **1/ la théorie classique (dite model du score vrai)**
- L'origine de cette théorie remonte essentiellement aux travaux du psychologue anglais Charles Spearman (1907-1913).
- Dans cette approche, l'intérêt est mis sur le score global de l'individu suite à un test (de Ayala, 2009 ; p.5).
- Le modèle de mesure dit classique permet d'évaluer jusqu'à quel point un score obtenu à un test reflète bien la compétence ou l'aptitude de l'individu en question. (K. Belhadj, 2012)
- L'équation de base du modèle classique est la suivante :

$$X = V + E$$

- **X** est le score observé d'un individu,
- **V** est le score vrai de cet individu
- **E** est l'erreur de mesure.
- Il est bon de souligner que :
 - - le score observé (**X**) est variable d'une répétition à l'autre du test
 - - Le score vrai (**V**) est intuitivement lié à un individu particulier et à un test particulier : ainsi le score vrai changera non seulement d'un individu à un autre mais aussi d'un test à l'autre.
 - - l'erreur de mesure (**E**) est une entité non observable, inconnue, variable d'une répétition à l'autre du test.

- L'équation de base du modèle classique, signifie que:
- Le score observé à un test est constitué de deux composantes additives : V et E.
- Le score observé à un test est obtenu lors d'une administration particulière de ce test.
- Chaque individu qui a passé ce test, à ce moment particulier, a donc un score observé.
- Ce score observé varie d'une répétition à l'autre du même test.
- Typiquement, le score observé peut être une fonction de la somme des items réussis d'un test (...) (Bertrand R., BLAIS J. (2004), p.39).

- **Les limites de la théorie du score vrai**

- Difficultés à comparer les caractéristiques du candidat et celles du test. Il est donc difficile de comparer des candidats évalués par des tests différents, de même qu'il est difficile de comparer les tests lorsqu'ils sont administrés à des populations différentes (exemple de deux candidats qui ont le même score à deux tests mesurant la même habileté, mais dont les difficultés moyennes ne sont pas égales).
- Il est difficile en pratique d'avoir des tests parallèles
- L'erreur standard de mesure est supposée identique pour tous les candidats.
- La théorie classique est orientée sur le test, et non sur l'item. Donc, elle ne fournit pas d'informations sur la manière dont les candidats répondent à un item donné.

- **2/ La théorie de la réponse à l'item**

- Cette théorie développée dans les années 60 est venue pour pallier aux limites de la théorie classique. Ces caractéristiques sont:
- Elle rend possible une mesure plus objective de traits psychologiques.
- Elle permet de décrire simultanément la compétence de la personne et la difficulté de l'item.
- Elle s'efforce de produire une estimation des propriétés de l'item qui soit indépendante d'un groupe particulier d'individus. En d'autres termes, elle cherche à élaborer des instruments de mesure dont les caractéristiques ne soient pas excessivement influencées par tel ou tel autre groupe de référence.

-
- **L'item** est la petite unité du test. Il peut être une question, un exercice, ou toute autre tâche à réaliser

- Elle permet d'obtenir des erreurs d'estimation séparées pour chaque item et pour chaque personne. il sera possible alors de quantifier les sources d'erreurs et d'identifier les plus importantes afin de les corriger. Alors que dans la théorie du score vrai l'erreur de mesure est estimée globalement dans le test elle n'est pas différentiable mais estimée globalement sur les items du test.
- Enfin, elle permet d'ajuster la difficulté des items indépendamment des répondants cibles.
- L'inconvénient de cette théorie est sa relative complexité de mise en œuvre

- **Définition de test**
- **BINET (1895):** “Les tests sont des expériences arrêtées d’avance, qui permettent d’obtenir des résultats relatifs aux différences individuelles pour plusieurs processus psychiques.”
- **Association Internationale Des Psychotechniques (1933)**
- C’est la première définition officielle des tests:
- “Un test est une épreuve définie, impliquant une tâche à remplir, identique pour tous les sujets examinés, avec une technique précise pour l’application du succès ou de l’échec, ou pour la notation numérique de la réussite.”
- **Pichot (1949):** Un test psychologique est une mesure contrôlée du comportement, c'est «une situation expérimentale, standardisée servant de stimulus à un comportement. Ce comportement est évalué par une comparaison statistique avec celui d'autres individus placés dans la même situation, permettant ainsi de classer le sujet examiné, soit quantitativement, soit typologiquement»

- Jean Bégin (2020) explique cette définition de Pichot (1949) comme suit:
- situation expérimentale
- enregistrement (mesure) du comportement «déclenché»
- comparaison du comportement avec celui d'autres individus (ou situations)
- le classement du sujet (but final du test)

- REUCHLIN (1991):
- « Epreuve, utilisée notamment en psychologie différentielle, qui permet de décrire le comportement d'un sujet dans une situation définie avec précision (« consigne » du test) par référence au comportement d'un groupe défini de sujets placés dans la même situation ».

- On retient d'après toutes ces définitions, qu'un test psychologique doit répondre aux **qualités métrologiques** des tests attestant sa fiabilité, à savoir: sensibilité, validité, et fidélité.

- **Les qualités métrologiques des tests**

- **1) La sensibilité:** L'objectif principal d'un test est bien d'observer des différences interindividuelles. Cette capacité de différenciation des sujets correspond à la notion de sensibilité.
- Donc, un test sensible vise à mettre en évidence des différences individuelles au niveau des comportements. Il doit présenter une certaine finesse discriminative.
- dans le cadre de la mesure d'une dimension (intelligence, aptitude, attitude, motivation, ...), le test doit différencier au mieux l'ensemble des individus.
- dans le cadre d'un dépistage (exemple: présence ou absence d'un trouble), le test diagnostic doit détecter au mieux les personnes ayant le trouble de celles qui n'en souffrent pas (J.-L. Roulin, 2018) . C'est la capacité d'un outil de mesure à détecter un changement minimal dans la condition d'un patient qui soit cliniquement significatif (M. Hotton, F. Bergeron, 2014).

- **2) La validité:** Capacité d'un outil de mesurer ce qu'il est censé mesurer ((M. Hotton, F. Bergeron, 2014). « le test mesure-t-il ce qu'il est censé mesurer ? »)
- les comportements mesurés doivent refléter certains processus mentaux, certaines caractéristiques et traits psychologiques non observables directement mais bien définis et bien cernés par l'instrument de mesure (le test)
- le concept de validité renvoie à la relation qui existe entre les éléments théoriques (modèles, définitions, concepts, hypothèses, etc.) et la réalité empirique supposée les représenter.
- La validité réfère donc à l'ensemble des éléments (preuves) qui doit conduire à nous assurer que l'interprétation des scores par les utilisateurs sera correcte. (J.-L. Roulin, 2018)

- **3) La fidélité:** Les mesures effectuées à l'aide d'un test doivent présenter une certaine stabilité.
- « On considère qu'un instrument de mesure est fidèle si le résultat qu'il produit est reproductible (comparable). » (Dickes et al., 1994, p. 165).
- un test fidèle est un test avec une erreur de mesure faible.
- La fidélité est donc un indicateur de la précision et de la constance des scores.
- Plus un instrument est fidèle, plus le score observé sera proche du score vrai (celui qui serait obtenu si le test était parfait !) (J.-L. Roulin, 2018)
- Si un même groupe de sujets passe le même test à quelques jours d'intervalle on doit observer les mêmes résultats ou, au moins, des résultats comparables.

Quelques exemples de tests psychométriques

- **Les tests de Raven**

- Les épreuves Progressive Matrices de Raven sont créées par John Carlyle Raven. Elles sont conçues pour mesurer l'intelligence générale (facteur **G**) et l'intelligence fluide (capacité à résoudre des problèmes nouveaux indépendamment des connaissances déjà acquises).
- Il existe trois versions des tests Progressives Matrices capables de couvrir une large gamme de niveaux différents et adaptées à un large public, allant des enfants aux adultes de haut niveau.
- La première version des Progressives Matrices est la version standard datant de 1938. C'est la version Progressives Matrices Standard ou SPM (encore appelée PM38).
- En 1943, une nouvelle version plus difficile pour les adultes (les *Advanced Progressives Matrices* ou A.P.M)
- En 1947 une version en couleur, assez attrayante pour les enfants (les *Progressives Matrices Couleur*)
- On se limite à la présentation de la 1^{ère} version standard

- **Présentation de la version SPM de Raven**

- La première édition date de 1938. La 2^{ème} de 1956. La dernière de 1998
- Elle destinée à des sujets de niveaux d'études qui correspondent au collège et au lycée.

- **Le principe du test**

- il s'agit de repérer les lois de progression entre différents éléments d'une même série et de les appliquer ensuite afin d'identifier l'élément qui vient compléter la série. (P. Chartier, E. Loarer, 2008)

- **En quoi consiste le test?**

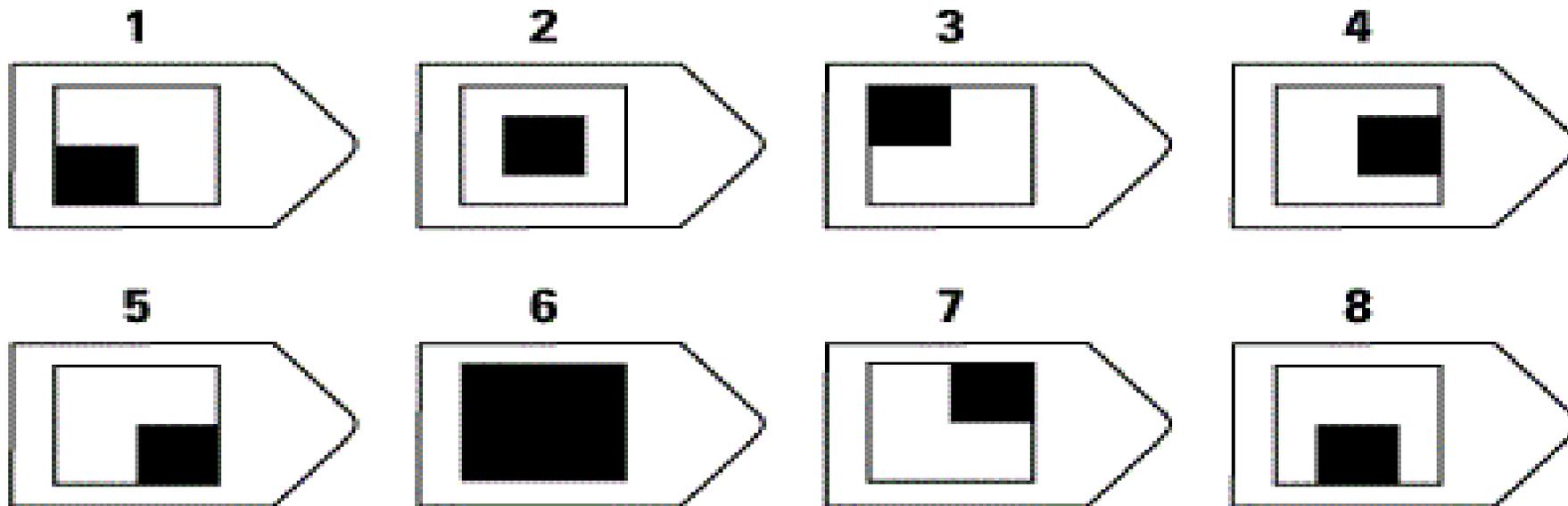
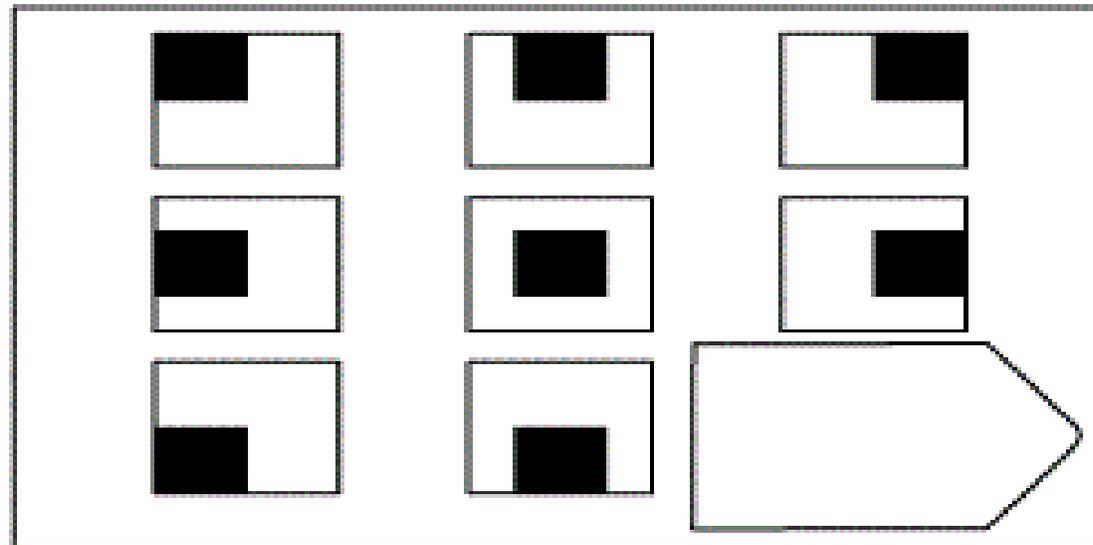
- La tâche consiste pour le sujet à sélectionner, parmi plusieurs possibilités offertes, l'élément qui vient le mieux compléter une série proposée ;
- Les matrices (doubles séries en lignes et en colonne) comportent quatre éléments (2 lignes et 2 colonnes) ou neuf éléments (3 lignes et 3 colonnes), l'ensemble correspondant au problème à résoudre ;
- La tâche du sujet consiste à découvrir les règles d'organisation (de transformation) de ces différents éléments, puis à appliquer ces règles afin d'identifier la réponse pertinente ;
- La réponse est donnée en choisissant un élément dans un ensemble comportant 6 à 8 réponses possibles.

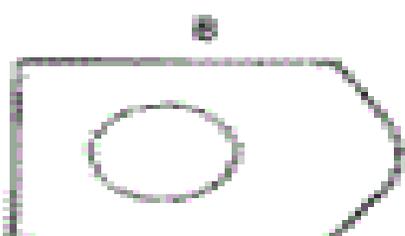
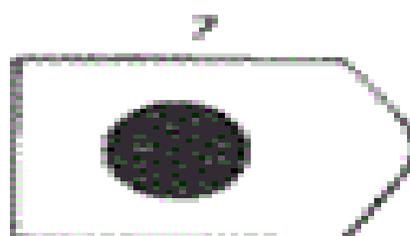
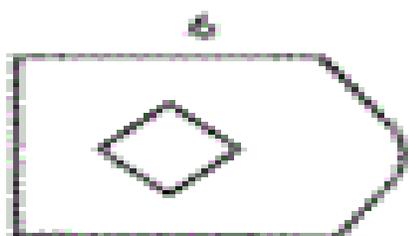
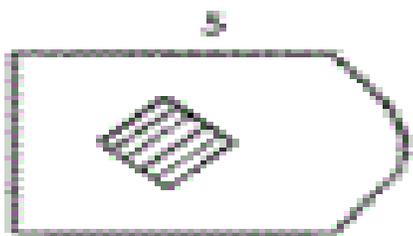
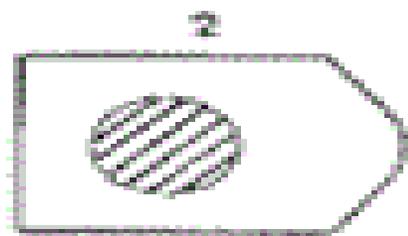
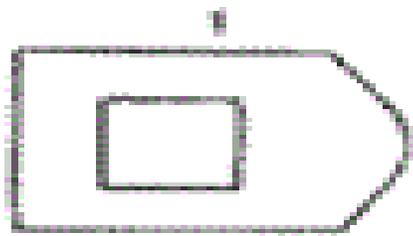
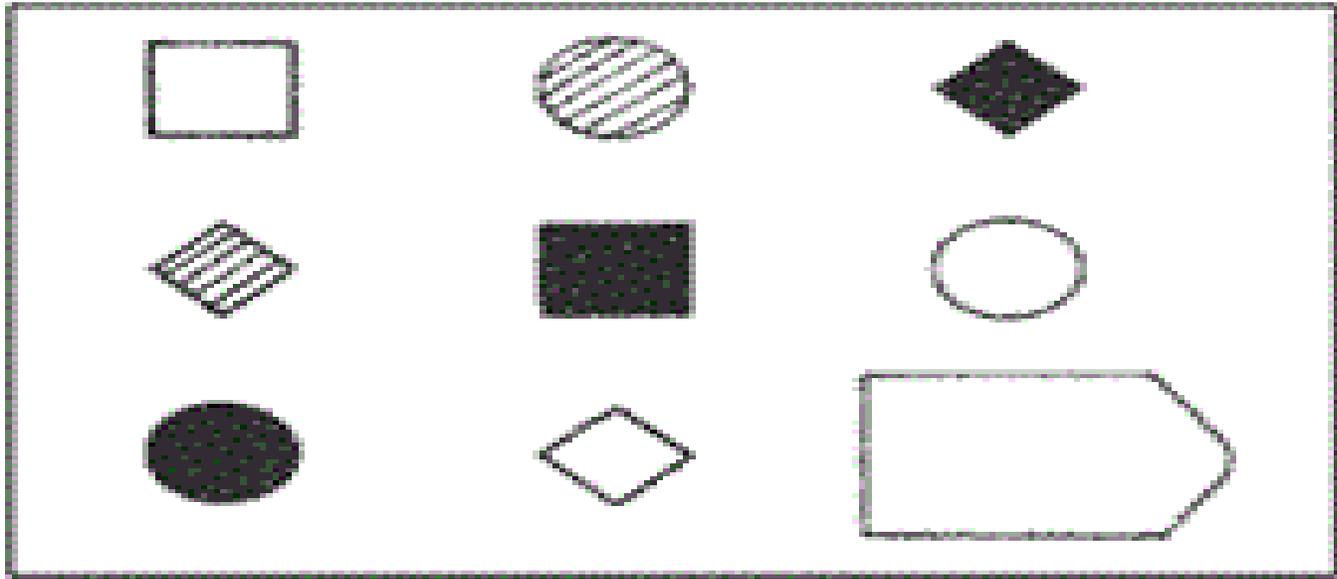
- **Les items de la version SPM**

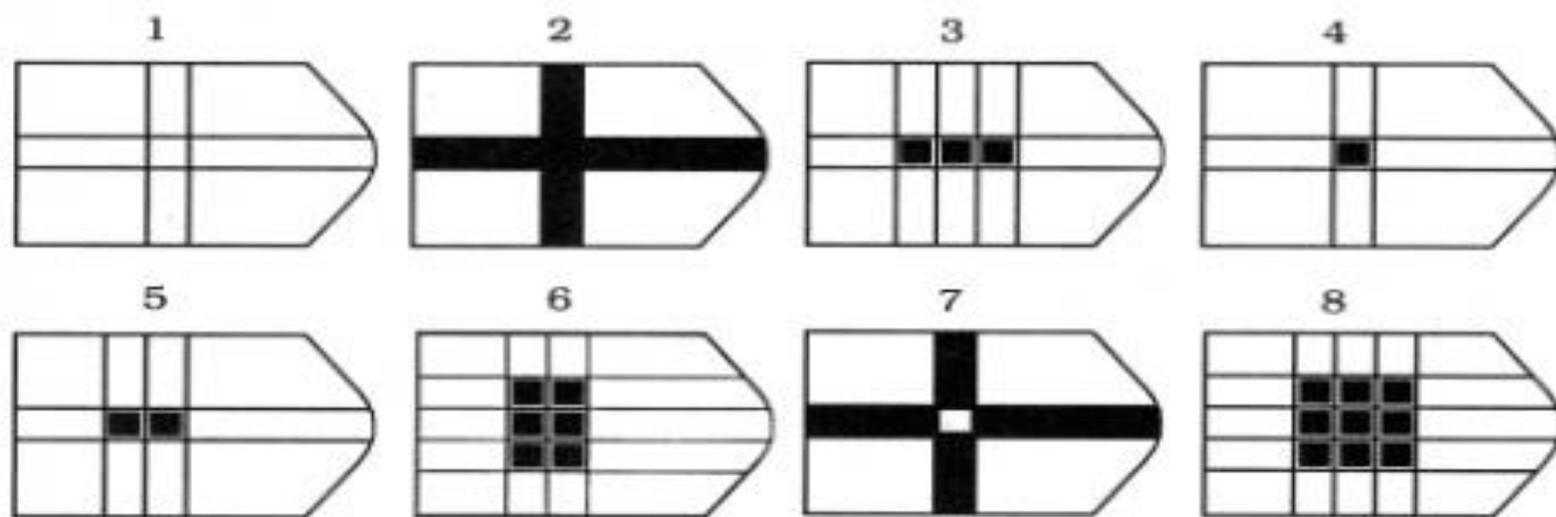
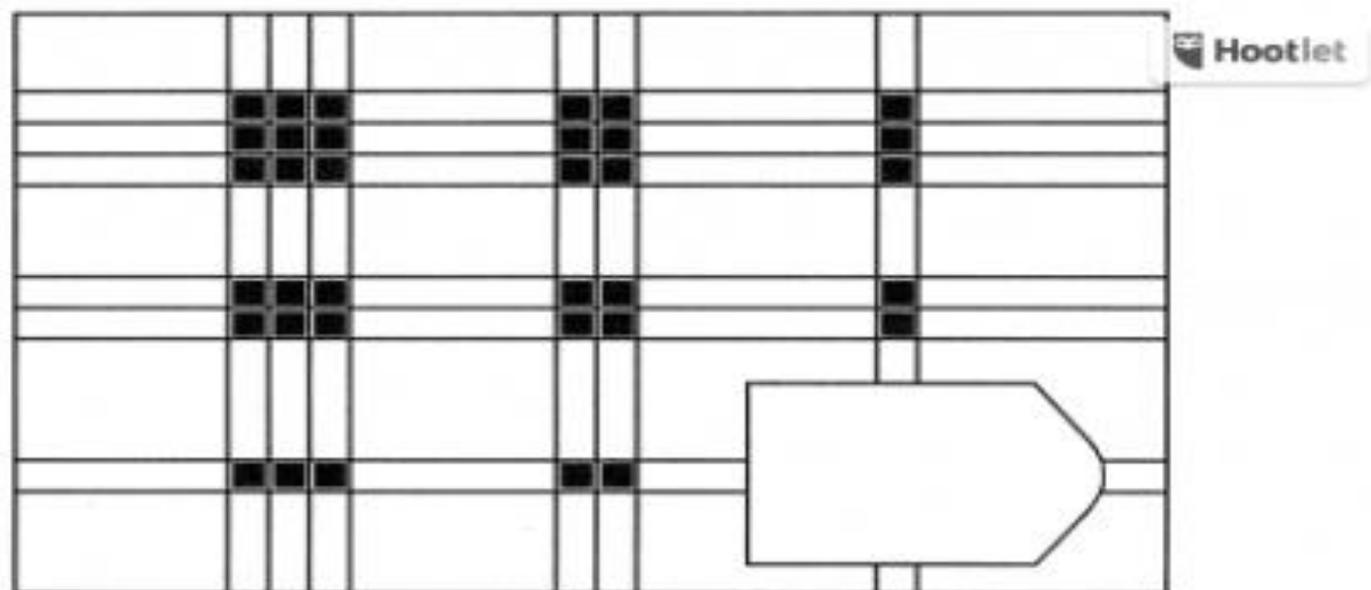
- Cette version est composée de 60 items, organisés en cinq séries de 12 items.
- Rien n'indique cependant au sujet cette organisation, car les items se suivent de façon continue.
- Quelle que soit la série d'items, la tâche consiste toujours à sélectionner, parmi plusieurs possibilités (6 ou 8 possibilités selon les séries), la figure qui vient compléter l'ensemble.
- L'épreuve est organisée selon un ordre de difficulté croissant (du plus facile au plus difficile), aussi bien au sein d'une même série qu'entre les séries.

- **La cotation de la version SPM**
- la cotation est très facile. On 1 point pour chaque bonne réponse. Le score peut donc varier de 0 à 60 points.
- la cotation se fait sur la feuille de cotation, conçue pour cet effet, et ne prend que quelques minutes.

Exemple d'items de la version SPM







- **Test de la figure de REY**

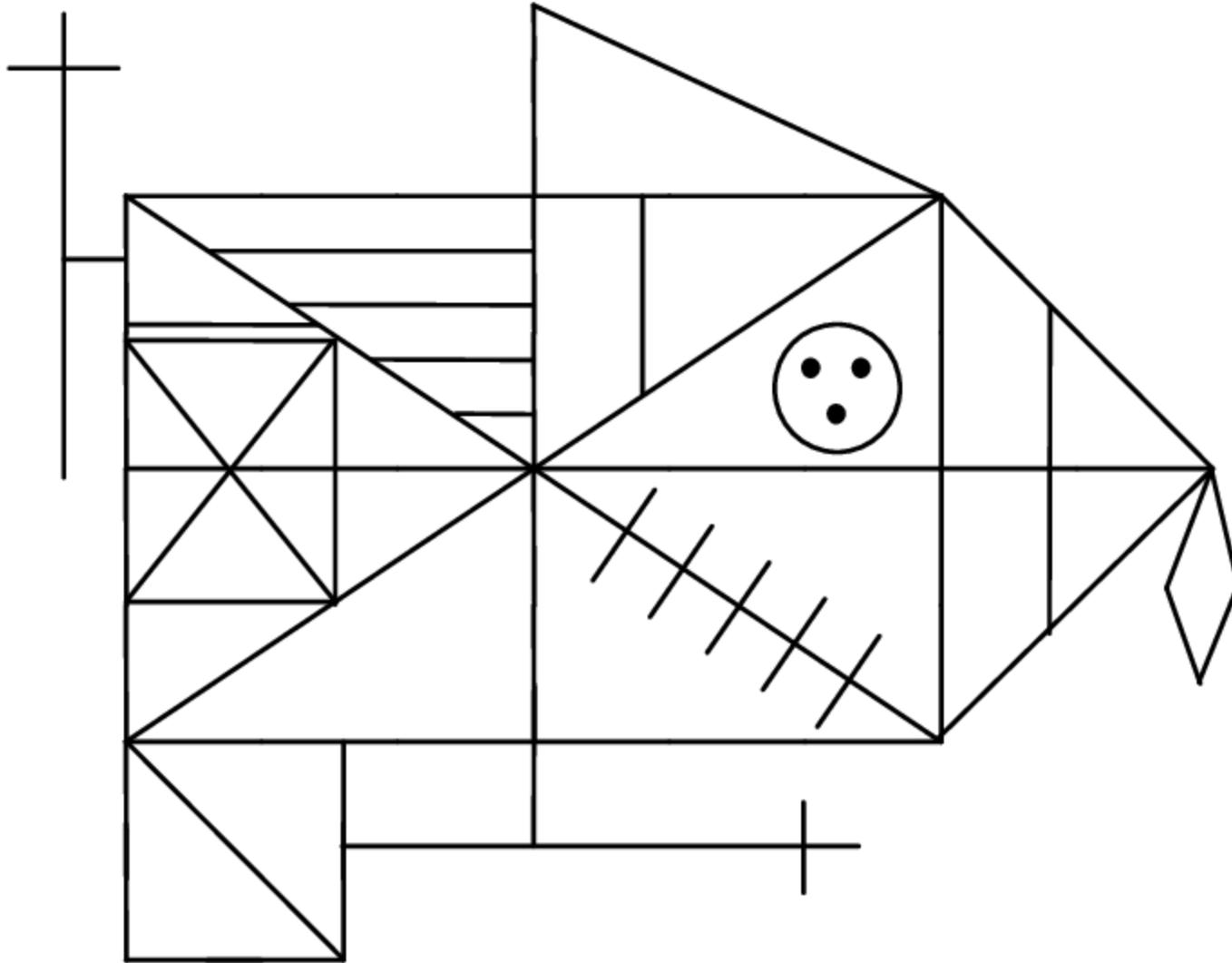
- Le test de la figure complexe de Rey, a été réalisé en 1939 par André Rey. Il est connu sous le nom de "Figure de Rey"
- Il est composé de deux versions: la version de la "Figure Complexe de Rey", destinée à la tranche d'âge 8 ans et plus, ainsi que la version « figure simplifiée de Rey », destinée à la tranche d'âge 3- 7 ans.

- **Objectif du test**

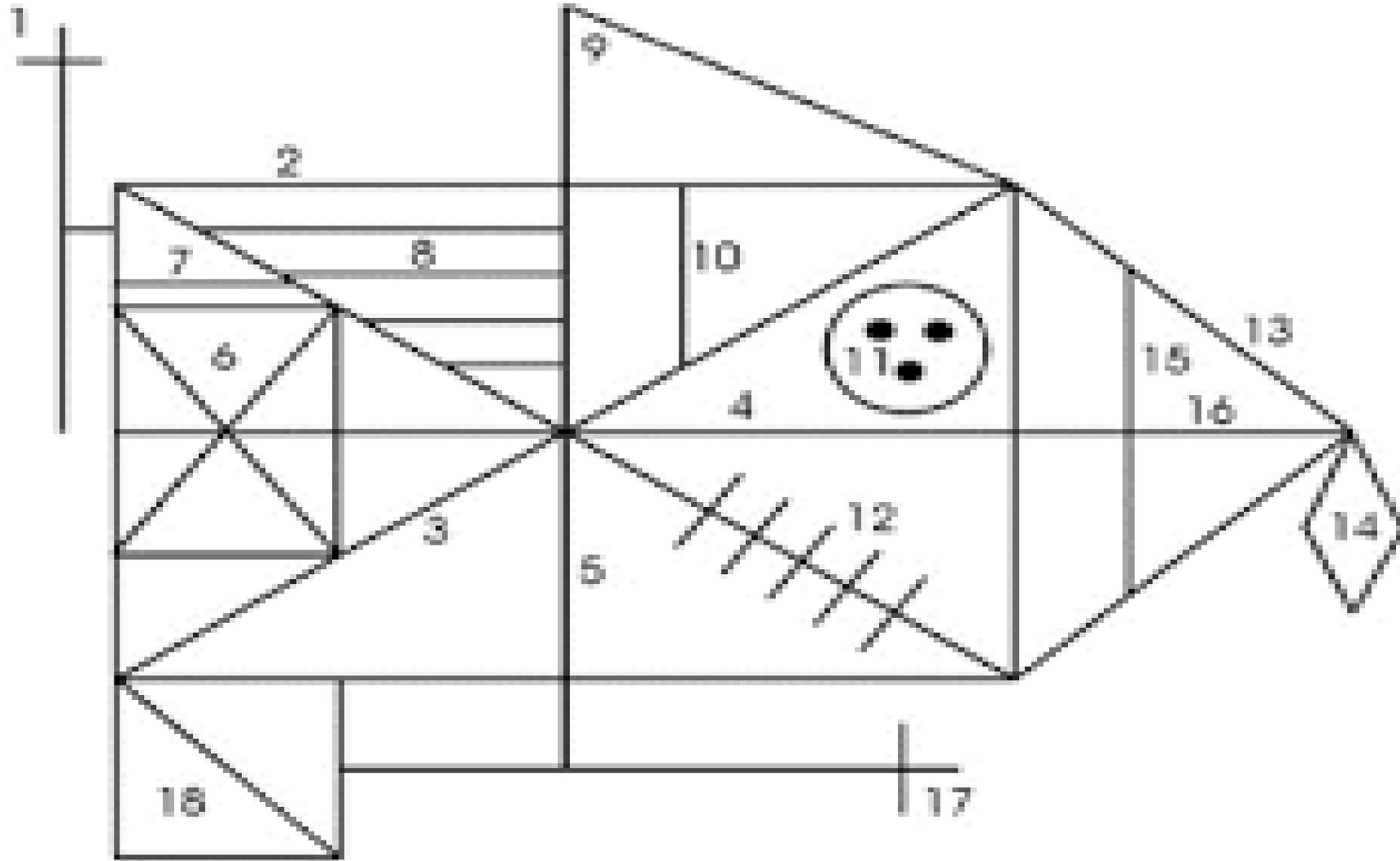
- D'abord destiné à l'évaluation de la mémoire visuelle et des capacités visuo-spatiales dans le cas de traumatisme crânien (Rey, 1941), ce test a ensuite été utilisé par Osterrieth (1945) auprès des sujets suspectés de déficit mnésique .
- Aujourd'hui, le test de la figure de Rey reste un des tests "papier-crayon" les plus utilisés pour l'évaluation cognitive (Camara et coll, 2000).
- Plusieurs professionnels (psychologues, psychomotriciens, orthophonistes) peuvent l'utiliser, à des fins bien différentes. (THINE-COUSSIÉ A., 2012)

- **Description du test de la figure complexe de REY**
- Rey propose une figure n'ayant aucune signification évidente, de réalisation graphique facile mais dont la structure globale est suffisamment compliquée pour mettre en jeu une activité perceptive et organisatrice.
- Elle ne demande pas de capacités graphiques particulières, la difficulté se trouvant essentiellement dans l'agencement des différents éléments entre eux.
- La figure est composée de 18 éléments, et est organisée en trois parties : une forme globale (le grand rectangle), des éléments externes (carrés, croix, triangles), et des éléments internes à la forme globale (lignes, ronds...). (Noël, 2007)

- La figure complexe de REY



- Les 18 détails de la figure complexe de REY



- **La passation du test**
- La passation du test de la figure de REY se déroule en 2 étapes: étape de la copie, et étape de la reproduction en mémoire.
- **1- étape de la copie**
- Le modèle est présenté horizontalement au sujet, où le losange (détail n° 14) doit être orienté en bas à droite
- Une feuille de format A4 est placée horizontalement devant le sujet.
- La consigne pour l'adulte est "Je vous demande de copier ce dessin".
- L'examineur doit mettre à la disposition de son sujet un crayon, ou plusieurs crayons de couleur pour se faciliter la tâche à retracer le processus de l'agencement des différents éléments de la figure utilisé par le sujet.
- L'utilisation discrète du chronomètre est recommandée pour calculer le temps de production

- 2- étape de reproduction de mémoire
- Une fois que l'étape de production est achevée, l'examineur accorde au sujet un temps de 3 minutes avant d'entamer l'étape de la reproduction de mémoire
- L'examineur demande alors au sujet : "Réalisez maintenant le même dessin sans le modèle".
- Comme à la 1^{ère} étape, l'examineur mis à la disposition du sujet: une feuille blanche au format A4, un crayon ou des crayons de couleur. Comme il doit aussi utiliser discrètement le chronomètre.

- **La cotation de la figure complexe de REY**
- La cotation se fait en 2 étapes:
- **La cotation du type de la production**: permet de comprendre la méthode utilisée par le sujet pour copier, puis reproduire la figure.
- Pour connaître le type d'organisation utilisé, il faut lors de la passation de l'épreuve, avoir recueilli avec attention la succession des traits.
- L'objectif de cette cotation est de permettre de juger des capacités perceptives et visuo-attentionnelles du sujet.
- **La cotation numérique**: elle permet d'établir un score numérique. Elle repose sur la qualité de copie et de reproduction du modèle.
- La figure est divisée en 18 éléments (voir les détails de la figure). Ils sont cotés un par un, de la manière
- suivante :

$$[\text{Présence (p)} + \text{Exactitude (E)}] \times \text{Place (Pl)} = \text{Note (N)}$$

- La feuille de dépouillement

