

## INSBAT Batterie d'analyse de la structure de l'intelligence

L.F. Hornke, M. Arendasy, M. Sommer, J. Häusler,  
M. Wagner-Menghin, G. Gittler, B. Bognar, M. Wenzl

© SCHUHFRIED GmbH

L'INSBAT est une batterie de tests d'intelligence modulaire, construite sur des bases théoriques et destinée à mesurer de manière économique et juste les aptitudes professionnelles.

**Utilisation :**

La mesure du niveau d'intelligence et de la structure de l'intelligence, utilisable à partir de 14 ans.

**Principaux domaines d'application :** Psychologie des ressources humaines, psychologie de l'aviation, psychologie pédagogique

**Contexte théorique :**

L'INSBAT a été construit de façon modulaire, dans l'objectif d'une évaluation psychologique orientée sur la prise de décision. Cela signifie qu'il suffit de présenter uniquement les sous-tests qui fournissent le maximum d'informations utiles pour répondre aux questions.

L'INSBAT est basé sur le modèle hiérarchique de l'intelligence de Cattell-Horn-Carroll (Carroll, 1993 ; Horn, 1989 ; Horn & Noll, 1997). Ce modèle postule que les intercorrélations entre les sous-tests pour la mesure des facteurs primaires peuvent être expliquées par neuf facteurs secondaires au contenu plus large. Les relations entre les facteurs secondaires sont expliquées par un facteur global d'intelligence qui forme le sommet du modèle hiérarchique d'intelligence. La validité de cette structure de facteurs a pu être répliquée dans un grand nombre d'études dans différents pays (p.ex. Brickley, Keith & Wolfe, 1995; Carroll, 1989 ; Gustafson, 1984 ; Horn & Stankov, 1982; Undheim & Gustafsson, 1987).

Dans la construction de l'INSBAT, on a choisi les facteurs secondaires suivants, intéressants dans les domaines d'application pratiques tels que la psychologie du travail, des entreprises et des organisations ou la psychologie pédagogique :

**Intelligence fluide :** C'est la capacité à reconnaître des relations entre les stimuli, de comprendre des implications et de tirer des conclusions logiques (sous-tests : raisonnement numérique inductif, raisonnement figural inductif, raisonnement verbal inductif).

**Intelligence cristallisée :** C'est l'ampleur et la profondeur des connaissances culturelles acquises, notamment la fluidité écrite ainsi que la compréhension parlée et écrite (sous-tests : vocabulaire, fluidité verbale).

**Mémoire à court terme :** C'est la capacité de retenir à court terme des informations visuelles et verbales et de pouvoir les reproduire correctement (sous-tests : mémoire visuelle à court terme, mémoire verbale à court terme).

**Mémoire à long terme :** C'est la capacité de retenir des informations à long terme, de les intégrer dans sa propre base de connaissances et de les rappeler correctement (sous-test : mémoire à long terme).

**Traitement visuel :** C'est la capacité de pouvoir se représenter l'image d'objets suite à une rotation mentale ou une transformation (sous-test : représentation spatiale).

**Vitesse de traitement :** Elle est définie par le nombre d'opérations cognitives de bases qui peuvent être effectuées dans un certain temps (sous-test : temps de visualisation).

**Certitude de la décision :** Elle est définie par le nombre d'opérations cognitives de base nécessaires à une prise de décision à la fois rapide et correcte pour des tâches non triviales (sous-test : certitude de la décision).

**Raisonnement quantitatif :** C'est la compréhension et la capacité d'application de concepts et d'aptitudes mathématiques (sous-tests : capacité d'estimation arithmétique, compétence arithmétique, flexibilité arithmétique, raisonnement algébrique).

Au total, on dispose donc de 15 sous-tests différents. Les tâches des différents sous-tests ont été construites à l'aide de différentes approches de la génération automatique d'items (AIG : Arendasy & Sommer, in press ; Irvine & Kyllonen, 2002), en tenant compte des résultats actuels de la recherche en science cognitive et en psychométrie appliquée. La construction des items a été réalisée soit par des personnes humaines, soit de manière automatique à l'aide de générateurs d'items. Par ailleurs, compte tenu des caractéristiques psychométriques des items, on a veillé à ce que les items des différents sous-tests (1) soient évolutifs sur le modèle de Rasch (Rasch, 1980) et que (2) le modèle théorique qui a servi de base aux items explique au moins 50 pour cent de la variance des paramètres de difficulté des items. Pour l'utilisateur, il en résulte l'avantage d'une justesse de calcul et d'une clarté d'interprétation des différents résultats obtenus lors des sous-tests.

**Réalisation :**

L'INSBAT a été développé sous forme d'une batterie de tests d'intelligence modulaire. Cela signifie qu'il suffit de présenter uniquement les sous-tests pertinents pour répondre aux questions diagnostiques. Il existe pour ce faire la version variable I (S1) et la version variable II (S2). Tandis que la version S1 permet uniquement de sélectionner les sous-tests, la version S2 offre la possibilité d'ajuster parfaitement l'ordre des sous-tests et les conditions finales des tests progressifs.

Les deux versions de tests sont adaptées avant tout à la prévision du succès

professionnel et scolaire dans des domaines ayant des profils d'exigence clairement définis. Pour effectuer un contrôle différencié du niveau d'intelligence et de la structure de l'intelligence, chaque facteur secondaire devrait être mesuré par un sous-test de marquage (raisonnement inductif figural, vocabulaire, mémoire visuelle à court terme, mémoire à long terme, représentation spatiale et temps de visualisation). En complément, on peut utiliser un second sous-test par facteur secondaire afin de garantir la largeur du contenu du facteur secondaire. Cette approche est déjà réalisée dans la version globale (S3). Selon Schmidt et Hunter (1998), une telle méthode est surtout adaptée au pronostic de la performance professionnelle dans des métiers très variés comprenant des tâches hétérogènes. Les ressources temporaires ne s'avèrent pas suffisantes pour effectuer une mesure différenciée de la performance d'une personne, il est possible d'avoir recours à la version courte (S4).

Pour chaque sous-test, on dispose d'une consigne standard avec des exemples d'exercices basés sur les principes de cours programmés et du « mastery learning ». La saisie des réponses se fait, en fonction du sous-test, sur le principe du choix multiple ou sur la base de réponses libres automatisées. Les tâches des différents sous-tests sont présentées en partie sous forme de test de puissance et en partie avec une limite de temps par item. Dans 11 sous-tests, les tâches sont présentées sous forme de test adaptatif (CAT) dont l'accès est optimisé à l'aide de données sociodémographiques, ce qui permet de maximiser l'information recueillie pour les différentes tâches sans être trop ou pas assez exigeant envers le sujet.

## Versions du test :

Au total, quatre versions du test sont disponibles :

Version S1 (version variable I), version S2 (version variable II), version S3 (version globale) et version S4 (version courte)

## Evaluation :

Pour chacun des sous-tests sélectionnés, le paramètre d'aptitude selon le modèle de Rasch est édité. Alors que le nombre d'items traités correctement mesure seulement la performance de l'individu, le paramètre d'aptitude permet d'estimer la dimension d'aptitude sous-jacente, latente. C'est là un pas d'inférence essentiel et nécessaire à l'évaluation diagnostique. Ce pas d'inférence est lié à l'ajustement au modèle de Rasch qui a pu être démontré pour les sous-tests de l'INSBAT (cf. à ce sujet van der Linden & Hambleton, 1997).

Si plus d'un sous-test a été sélectionné pour un facteur secondaire, un paramètre d'aptitude est calculé pour le facteur secondaire correspondant. Le paramètre d'aptitude pour le facteur global Intelligence générale (G) est calculé si au moins un sous-test a été sélectionné pour quatre des six facteurs secondaires intelligence fluide (Gf), intelligence cristallisée (Gc), traitement visuel (Gv), raisonnement quantitatif, mémoire à court terme (Gstm) et mémoire à long terme (Gltm). En parallèle à l'édition du paramètre d'aptitude et des scores de facteurs, une comparaison de normes est effectuée (percentiles et QI ; intervalle de confiance). À la fin du test, les résultats sont présentés sous forme de tableau ainsi que sous forme d'un profil. Le protocole des différents sous-tests permet de montrer les temps de latence par item, la réponse choisie et son évaluation ainsi que des statistiques du processus adaptatif.

Par ailleurs, l'INSBAT offre également la possibilité de transférer automatiquement les résultats du test dans un modèle de rapport.

## Fiabilité :

En raison de la présentation adaptative des items des différents sous-tests ainsi que de la validité nécessaire d'un modèle de test de probabilité, il est possible d'atteindre toute précision souhaitée. Pour des raisons d'économie, la fiabilité des différents groupes d'items se situe entre  $r=.70$  et  $r=.95$ . Au bout d'environ 15 mois, la stabilité des sous-tests varie entre  $r=0.63$  et  $r=0.87$ .

## Validité :

La représentation de construit (Embretson, 1983) des différents sous-tests de l'INSBAT a été démontrée par des études dans lesquelles les difficultés des items ont été prédites à l'aide de caractéristiques de tâches qui peuvent être déduites des modèles théoriques pour la résolution de ce type d'items. Les multiples corrélations entre les paramètres de difficultés des items du modèle de Rasch (Rasch, 1980) et les caractéristiques des items ainsi obtenues varient pour les différents sous-tests entre  $R=0.72$  et  $R=0.97$ . Cela signifie que les représentations théoriques ayant servi de base à la construction des items dans les différents sous-tests expliquent 50 à 94 pour cent des différences du niveau de difficulté des items.

Par ailleurs, on dispose aujourd'hui de nombreuses études sur l'approche nomothétique (Embretson, 1983) des différents sous-tests. La validité de construit pour la batterie de tests dans son ensemble ainsi que pour la version globale et la version courte a pu être démontrée dans une étude de Sommer et Arendasy (2005; Sommer, Arendasy & Häusler, 2005). Les auteurs ont pu confirmer l'attribution, sur des bases théoriques, des différents sous-tests aux facteurs secondaires du modèle de Cattell-Horn-Carroll, grâce à une analyse factorielle confirmatoire. Ces résultats sont complétés par les études de Arendasy et Sommer (2007), ainsi que de Arendasy, Hergovich et Sommer (2008) dans lesquelles les résultats précédemment explicités ont pu être répliqués par association de sous-tests alternatifs à l'aide d'un échantillon indépendant.

Il existe des références concernant la validité des critères des différents sous-tests de l'INSBAT dans le domaine de la psychologie de l'aviation (sélection des futurs pilotes) et dans le conseil pédagogique (pronostic de réussite en haute école spécialisée).

## Normes :

Il existe des normes pour 904 adultes âgés de 16 à 73 ans ainsi que pour 1595 jeunes âgés de 12 à 15 ans. Ces deux normes sont également disponibles séparément en fonction de l'âge, de la formation et du sexe.

## Durée de passation :

La durée de passation de la version variable I (S1) et de la version variable II (S2) dépend du choix des sous-tests et de la précision sélectionnée. La durée des différents sous-tests varie entre 3 et 40 minutes. Pour les deux versions de test S3 et S4 pour lesquelles l'ordre des sous-tests est fixe, la durée de passation est respectivement de 2 heures 52 minutes et de 1 heure et 27 minutes.