

# Lignes directrices pour la classification et l'interprétation des résultats aux tests neuropsychologiques

## 1 Remarques préalables

Lors d'une évaluation neuropsychologique, le jugement clinique repose en général sur plusieurs aspects. Parallèlement aux résultats des tests diagnostiques, il faut évidemment aussi tenir compte des constatations médicales et psychopathologiques, de l'anamnèse et parfois également de l'hétéro-anamnèse, des symptômes subjectifs, de l'observation du comportement ou encore de la validité des symptômes afin de définir si un trouble neurocognitif est présent ou non. Les approches cliniques qualitatives continuent à être justifiées en neuropsychologie (p.ex. pour l'anosognosie ou encore l'apraxie) et ne sont nullement en contradiction avec l'approche quantitative des tests psychométriques. Ces lignes directrices constituent une recommandation dans le sens d'un standard actualisé afin de parvenir à une plus grande cohérence interindividuelle dans la classification et l'interprétation des résultats aux tests. Ces lignes directrices sont issues d'un consensus pragmatique et largement soutenu, dans le but d'améliorer la qualité dans ce domaine. Elles complètent les lignes directrices pour la rédaction des rapports ([www.neuropsych.ch](http://www.neuropsych.ch)) dont l'objectif principal est d'améliorer la qualité des comptes-rendus neuropsychologiques. Dans un premier temps, ce document discute du concept de norme qui sous-tend le bilan neuropsychologique ainsi que de l'utilisation de paramètres standardisés.

## 2 Le concept de norme

Les résultats aux tests neuropsychologiques sont obtenus sous forme de scores bruts. Toutefois, les valeurs brutes en elles-mêmes n'ont pas de signification et doivent être intégrées dans un système de référence (valeurs standardisées) permettant de classer les performances au test pour un individu (p.ex. en dessous ou en dessus de la moyenne). Pour ce faire, les scores bruts sont transformés en rangs centiles et en paramètres normalisés (comme le score z ou le score T) dans le but de permettre d'une part la comparaison avec la population de référence, et d'autre part la comparaison avec d'autres tests. Les échelles standard les plus courantes sont les normes de variabilité, les normes d'équivalence et les rangs centiles (Fisseni, 2004). Tandis que des normes de référence sociales (normes de variabilité) sont pertinentes pour l'évaluation neuropsychologique chez l'adulte, il arrive que chez l'enfant et l'adolescent il faille également utiliser des normes d'équivalence pour déterminer l'âge du développement par rapport à l'âge réel. Chez les patients adultes, les normes sociales atteignent parfois leur limite, par exemple lorsqu'un patient âgé doit être évalué quant à son aptitude à la conduite automobile. Dans ce cas, la comparaison avec le même groupe d'âge ne fonctionne pas, car les données normatives de la population de référence (p.ex. le groupe de référence des 90 ans) ne doivent pas être décisives pour la question de la conduite sécuritaire d'une personne âgée sur la route. Il est alors nécessaire d'appliquer une approche basée sur des exigences minimales à remplir (maîtrise). Dans le cadre de sa pratique clinique, le neuropsychologue doit être conscient de cet aspect et choisir le système de normes le mieux adapté pour répondre à la question posée.

## 3 Epreuves et normalisation

La sélection des instruments (tests) appropriés pour répondre d'un point de vue neuropsychologique aux diverses questions posées ne relève pas de ces lignes

directrices. De manière générale, il est toutefois recommandé d'utiliser, autant que possible, des tests standardisés et bien normés qui répondent aux exigences méthodologiques en termes de qualité du test. Par ailleurs, la procédure doit être adaptée à la gravité et à la nature du trouble et être basée sur une hypothèse. Si elles sont disponibles, il faut préférer des méthodes dites « *co-normées* » (i.e. normalisation sur un seul et même échantillon standard) qui permettent une identification plus claire et statistiquement fiable des déficits (différentiels) respectivement des performances aux tests divergentes (voir p.ex. la distinction « *classical dissociation* », « *strong dissociation* » ou « *trend dissociation* » selon Shallice, 1988). Par contre, les écarts de performances ne peuvent pas être objectivement interprétés s'ils ont été mesurés avec des techniques normées sur des échantillons standards de taille ou de qualité différentes ou ayant été réalisés dans des milieux culturels différents.

### 3.1 Prise en compte des effets de l'âge, du sexe et du niveau d'éducation

Lors de l'examen neuropsychologique, la performance aux tests dépend souvent de variables démographiques telles que l'âge, le sexe ou le niveau socio-culturel. Si plusieurs tests existent, il arrive parfois que ces tests diffèrent quant à leur difficulté. Si cela est possible et pertinent par rapport à la question de départ, il faut choisir les normes du groupe de référence le mieux adapté ; si on utilise des normes globales, il est conseillé de faire un ajustement régressif des valeurs brutes. Pour des questions spécifiques portant sur l'intelligence à l'âge adulte ou l'aptitude à la conduite automobile, il est également important que la comparaison à la population de référence ne prenne pas en compte l'âge, le sexe ou le niveau socio-culturel.

### 3.2 Prise en considération des degrés de différenciation des résultats de tests

Les différents résultats d'un seul test ou d'une batterie de tests n'ont pas tous le même degré de différenciation. Les intervalles des résultats effectivement obtenus chez une population de référence peuvent être plus ou moins larges. L'intervalle est particulièrement restreint pour des résultats qui tiennent compte de certains types d'erreurs ou pour des résultats qui sont facilement obtenus par des sujets sains. Pour ce type de tests, seules des valeurs seuils sont généralement disponibles. Des scores inférieurs ou supérieurs à de telles valeurs seuils indiquent alors des performances déficitaires. Pour d'autres tests avec un échantillon de normalisation plus large, il existe des normes plus fines. Pour des tests plus élaborés, l'examineur bénéficie souvent d'informations sur le degré de différenciation des différents résultats par le biais des normes recommandées (valeurs z, valeurs T, valeurs C, valeur seuil).

## 4 Classification diagnostique des résultats

Après une sélection des tests en fonction des questions posées et leur administration, les scores doivent être analysés et interprétés en fonction des normes et standards les plus récents et des connaissances professionnelles. Les étapes de description, de comparaison des résultats et d'interprétation devraient être réalisées séparément (voir *Richtlinien für den diagnostischen Prozess*, Westhoff et al., 2003). L'approche qualitative mentionnée plus haut constituerait donc plutôt une étape descriptive mais qui doit être prise en compte lors de l'interprétation des résultats au même titre que l'approche quantitative. Avant l'interprétation des résultats aux tests ou des éléments cliniques (comportement), il existe une étape de classification des résultats qui est discutée dans la section qui suit.

#### 4.1 Classification des résultats aux tests neuropsychologiques

Les classifications peuvent se faire d'une part sur la base de différents tableaux cliniques et d'autre part à partir de différences de performances (voir p.ex. Goldstein & Shelly, 1987). Ces lignes directrices traitent de la question de la classification des résultats aux tests neuropsychologiques.

La notion de norme statistique qui sous-tend la norme de variabilité caractérise une certaine variation autour de la moyenne pour une distribution gaussienne qui peut, par définition, être décrite comme « normale ». Il n'existe toutefois pas de consensus entre neuropsychologues sur les valeurs qui limitent l'étendue de cette variation autour de la moyenne, que ce soit en Suisse ou au niveau international. Un sondage non représentatif mené par les auteurs de ces lignes directrices a montré que la limite standard de la norme correspond en général au centile 5 dans l'ouest de la Suisse, en France ou encore en Italie, comme proposé par exemple par Benton et al. (Benton, Van Allen, Hamsher, & Levin, 1978; Benton, Varney, & Hamsher, 1977). Dans les pays germano- et anglophones, la limite de la norme semble généralement être considérée au centile 16, soit un écart-type en dessous de la moyenne de la population générale. Cette limite des normes est probablement due à la large diffusion de la batterie Halstead-Reitan et à ses normes (Heaton, Grant, & Matthews, 1991; Heaton, Miller, Taylor, & Grant, 2004). De plus, les centiles 25 ou 10 sont parfois également utilisés comme limites (cut-off) par d'autres cliniciens. Dans les épreuves de Wechsler, les résultats aux tests sont traditionnellement considérés comme « limites » en dessous du centile 10 et comme « extrêmement faibles » en dessous du centile 2 (Wechsler, 1997a, 1997b). L'absence de consensus à l'égard des valeurs limites des normes nécessite que le lecteur d'un rapport neuropsychologique soit familiarisé à cette problématique de classification des résultats et qu'il connaisse les différences régionales, voire parfois les conventions internes à certaines institutions. Cette situation nuit non seulement à la transparence mais aussi à la fonction première d'un rapport, c'est-à-dire la communication et la documentation.

De nombreuses propositions existent pour classer les résultats dans diverses catégories telles que « inférieur à la moyenne », « dans la moyenne » ou « supérieur à la moyenne » (Brooks et al., 2011). Ces propositions peuvent varier d'un faible niveau de différenciation (trois catégories) à un haut niveau de différenciation (neuf catégories). Dans le cadre de ces lignes directrices, l'ASNP propose une solution praticable en sept catégories (voir le tableau 1). Pour la limite inférieure de la norme, le centile 16 est appliqué et les résultats des tests entre le PR (*percentile rank*) >5 jusqu'au PR < 16 sont classés comme « à la limite inférieure de la norme ». Les résultats entre le PR ≥2 et le PR ≤5 sont classés comme « clairement inférieurs à la norme ». Une performance inférieure au centile 2 est classée comme « très inférieure à la norme ». Un score est « dans la norme » lorsqu'il se situe dans un intervalle d'un écart-type autour de la moyenne de la population (PR ≥16 à PR ≤84). Des scores allant du PR >84 au PR <95 sont classés comme « à la limite supérieure de la norme. Ils seront classés comme « clairement supérieurs à la norme » entre le PR ≥95 et le PR ≤98 et comme « très supérieurs à la norme » au-dessus du PR 98.

Tableau 1: Classification et interprétation clinique des résultats aux tests neuropsychologiques

Percentiles	Valeurs T	Valeurs z	QI	Description des valeurs statistiques (classification)	Interprétation clinique
>98 - 100	>70 - 80	>2.0 - 3.0	>130 - 145	très supérieur à la norme	
≥95 - ≤98	≥66.4 - ≤70	≥1.6 - ≤2.0	≥125 - ≤130	clairement supérieur à la norme	
>84 - <95	>60 - <66.4	>1.0 - <1.6	>115 - <125	à la limite supérieure de la norme	
≥16 - ≤84	≥40 - ≤60	≥-1.0 - ≤1.0	≥85 - ≤115	dans la norme	sans particularité
>5 - <16	>33.6 - <40	>-1.6 - <-1.0	>75 - <85	à la limite inférieure de la norme	possiblement déficitaire
≥2 - ≤5	≥30 - ≤33.6	≥-2.0 - ≤-1.6	≥70 - ≤75	clairement inférieur à la norme	déficitaire
0 - <2	20 - <30	-3.0 - <-2.0	55 - <70	très inférieur à la norme	sévèrement déficitaire

Il est ici important de préciser que les termes « *inférieur à la norme* », « *clairement inférieur à la norme* » et « *très inférieur à la norme* » ne reflètent pas forcément un résultat pathologique ou caractéristique d'une dysfonction cérébrale.

#### 4.2 Interprétation des résultats aux tests

Lors de l'interprétation des résultats, il faut considérer une multitude d'aspects qui ne peuvent être entièrement traités dans ces lignes directrices. Seules certaines caractéristiques méthodologiques telles que la sensibilité et la spécificité du bilan neuropsychologique seront abordées ici. Pour identifier un patient avec un dysfonctionnement cérébral acquis, il est nécessaire d'une part de définir des limites de la normalité relativement serrées, et d'autre part de s'assurer que les personnes normales (c'est-à-dire qui ont des résultats normaux aux tests) soient effectivement classées comme telles. Pour satisfaire aux exigences d'une sensibilité et d'une spécificité adéquates, il s'agit de respecter quelques principes de bases de la psychométrie, comme évoqué par exemple par Binder, Iverson & Brooks (2009) dans leur article « *To Err is Human: 'Abnormal' Neuropsychological Scores and Variability are Common in Healthy Adults* ». Ces auteurs soulignent que plus le nombre de tests utilisés augmente, plus la probabilité d'obtenir des résultats faibles (i.e. inférieurs à la moyenne) est élevée. Étant donné que l'utilisation d'une multitude de tests dans le contexte d'un bilan neuropsychologique est plutôt la règle que l'exception, et que la probabilité susmentionnée est d'autant plus élevée, il faut alors donner priorité aux batteries de tests dont les détails des données normatives sont disponibles afin de savoir combien de résultats inférieurs à la moyenne sont également observés chez des personnes saines de l'échantillon normatif. Une alternative pour estimer cette probabilité est l'utilisation de modèles statistiques telles que la distribution probabilistique binomiale ou encore la simulation Monte-Carlo.

Lors de l'interprétation des résultats aux tests, il faut généralement identifier toute inconsistance. Dans ce cadre, des procédures pour l'analyse de profils individuels sont très utiles. En fonction de la question posée, l'utilisation de procédures évaluant la

validité des symptômes pourra aussi s'avérer nécessaire. Par ailleurs, il est très important de s'assurer que les résultats obtenus sont compatibles avec une étiologie cérébrale connue ou supposée ou avec les données de la neuroimagerie. Il faut garder à l'esprit qu'une lésion neuroradiologique, par exemple une lésion frontale droite, n'entraînera pas forcément de dysfonction neurocognitive, et qu'à l'inverse l'absence de lésion neuroradiologique n'empêche pas forcément la présence d'une dysfonction neurocognitive (p.ex. amnésie après hypoxie cérébrale). En cas d'informations anamnétiques manquantes ou incomplètes, il serait clairement inadmissible de conclure à une dysfonction cérébrale ou de spéculer là-dessus sur la seule base des performances cognitives diminuées.

## 5 Interprétation de l'examen

L'interprétation de l'ensemble des résultats à l'examen neuropsychologique implique que les informations obtenues soient pondérées et intégrées en fonction de leur importance relative. Cette interprétation des données ne doit pas être laissée au lecteur. Les conclusions sont à présenter de manière claire et les conclusions provisoires à signaler comme telles. Il est recommandé de séparer les différentes fonctions cognitives en sous-catégories (attention, mémoire, fonctions exécutives, perception visuelle, perception visuo-constructive, etc.) et de les évaluer séparément.

L'interprétation se termine en général avec une estimation de l'ampleur du déficit neurocognitif. Pour la classification de la sévérité du trouble référence nous vous renvoyons aux lignes directrices de l'ASNP (Frei, Balzer, Gysi, Leros, Plohmann & Steiger, 2016). Depuis l'introduction du DSM-V (*American Psychiatric Association*, 2013), une alternative existe avec le diagnostic de « *trouble neurocognitif mineur ou majeur* ». Pour les troubles neurocognitifs majeurs, il existe en outre la possibilité d'une distinction en trois niveaux selon la gravité : léger, modéré ou sévère. Les troubles neurocognitifs majeurs impliquent des performances aux tests inférieures à deux écart-types par rapport à la moyenne de la population générale, alors que pour un trouble mineur ces performances se situent entre un et deux écart-types. Les guidelines ASNP proposées dans le tableau 1 rejoignent cette appréciation. Le DSM-V souligne à plusieurs reprises que l'interprétation doit se faire dans le contexte du niveau de performance pré-morbide. Une diminution de la performance par rapport à avant l'atteinte doit donc être constatée afin de pouvoir poser un tel diagnostic. Il va de soi que ce postulat n'est pas pertinent pour les interprétations en neuropsychologie développementale chez l'enfant et l'adolescent ou lorsque plusieurs événements (situations cliniques) se succèdent. Dans ces situations, les déficits sont plutôt évalués en fonction de l'âge de développement (normes équivalentes) ou respectivement par rapport aux normes sociales. L'interprétation des résultats de patients avec des niveaux d'éducation/d'intelligence très élevés ou très faibles y représente une difficulté particulière, de même que pour des patients issus d'un autre milieu culturel ou linguistique.

## Sources

- Benton, A. L., Van Allen, M. W., Hamsher, K. S., & Levin, H. S. (1978). Test of facial recognition, form SL. Iowa City, IA: Department of Neurology, University of Iowa Hospitals and Clinics.
- Benton, A. L., Varney, N. R., & Hamsher, K. S. (1977). Judgment of line orientation manual. Iowa City, IA: Department of Neurology, University of Iowa Hospitals and Clinics.
- Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human: "Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, 24, 31–46.
- Brooks, L. B., Sherman, E. M. S., Iverson, G. L., Slick, D. J. & Strauss E. (2011). Psychometric Foundations for the Interpretation of Neuropsychological Test Results. In: *The little Black Book of Neuropsychology*, Schoenberg, M. R. & Scott, J. G. (Eds). Springer, New York.
- DSM-5 (2013). Diagnostic and Statistical Manual of Mental Disorders. Fifth Edition. American Psychiatric Association, Washington, DC.
- Fisseni, H.-J. (2004). Lehrbuch der psychologischen Diagnostik. Mit Hinweisen zur Intervention. Ausgabe 3. Hogrefe Verlag GmbH, Göttingen.
- Frei, A., Balzer, Ch., Gysi, F., Leros, J., Plohmann, A. & Steiger, G. (2016). Kriterien zur Bestimmung des Schweregrades einer neuropsychologischen Störung sowie Zuordnungen zur Funktions- und Arbeitsfähigkeit. *Zeitschrift für Neuropsychologie*, 27[2], 107-119.
- Goldstein, G., & Shelly, C. (1987). The classification of neuropsychological deficits. *Journal of Psychopathological and Behavioral Assessment*, 9, 183-202.
- Heaton, R. K., Grant, I., & Matthews, C. G. (1991). Comprehensive norms for an extended Halstead–Reitan Battery: Demographic corrections, research findings, and clinical applications. Odessa, FL: Psychological Assessment Resources.
- Heaton, R. K., Miller, S. W., Taylor, M. J., & Grant, I. (2004). Revised comprehensive norms for an expanded Halstead–Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults professional manual. Lutz, FL: Psychological Assessment Resources.
- Shallice, T. (1988). From neuropsychology to mental structure. New York, NY: Cambridge University Press
- Integritätsschaden bei psychischen Folgen von Hirnverletzungen. *Medizinische Mitteilungen*, 2870/8.d.
- Wechsler, D. (1997a). Wechsler Adult Intelligence Scale: Third Edition. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997b). Wechsler Memory Scale: Third Edition. San Antonio, TX: The Psychological Corporation.
- Westhoff, K., Hornke, L.F., Westmeyer, H. (2003). Richtlinien für den diagnostischen Prozess - Zur Diskussion gestellt.[Deutsche Adaptation von: Fernandez-Ballesteros, R, De Bruyn, E.E.J., Godoy, A., Hornke, L.F., Ter Laak, J., Vizcarro, C., Westhoff, K., Westmeyer, H., & Zaccagnini, J.L. (2001). Guidelines for the Assessment Process (GAP): A Proposal for Discussion. *European Journal of Psychological Assessment*, 17, 187-200.] *Report Psychologie*, 28 (9), 504-517.